



GÉNTÉCHNOLÓGIA ÉS FEHÉRJEMÉRNÖKSÉG

szerkesztő:
NYITRAY LÁSZLÓ

Géntechnológia és fehérjemérnökség

Szerkesztette: Nyitray László

Alexa Anita (12. és 13. fejezet)
Fodor Krisztián (3. és 9. fejezet)
Garai Ágnes (4. és 5. fejezet)
Glatz Gábor (6. és 7. fejezet)
Radnai László (1. és 2. fejezet)
Rapali Péter (11. fejezet)
Szakács Dávid (8., 16. és 17. fejezet)
Várkuti Boglárka (15. fejezet)
Zeke András (10. és 14. fejezet)

Géntechnológia és fehérjemérnökség: Szerkesztette: Nyitray László

írta Alexa Anita (12. és 13. fejezet), Fodor Krisztián (3. és 9. fejezet), Garai Ágnes (4. és 5. fejezet), Glatz Gábor (6. és 7. fejezet), Radnai László (1. és 2. fejezet), Rapali Péter (11. fejezet), Szakács Dávid (8., 16. és 17. fejezet), Várkuti Boglárka (15. fejezet), és Zeke András (10. és 14. fejezet)

Lektorálta:

Várallyai Éva

Szerzői jog © 2013 Eötvös Loránd Tudományegyetem

E könyv kutatási és oktatási célokra szabadon használható. Bármilyen formában való sokszorosítása a jogtulajdonos írásos engedélyéhez kötött.

Készült a TÁMOP-4.1.2.A/1-11/1-2011-0073 számú, „E-learning természettudományos tartalomfejlesztés az ELTE TTK-n” című projekt keretében. Konzorciumvezető: Eötvös Loránd Tudományegyetem, konzorciumi tagok: ELTE TTK Hallgatói Alapítvány, ITStudy Hungary Számítástechnikai Oktató- és Kutatóközpont Kft.

Nemzeti Fejlesztési Ügynökség
www.ujjszechenyiterv.gov.hu
06 40 638 638



A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.



Köszönetnyilvánítás

Az elektronikus jegyzet a TÁMOP-4.1.2.A/1-11/-1-2011-0073 (E-learning természettudományos tartalomfejlesztés az ELTE TTK-n) és a TÁMOP 4.2.4.A/1-11-1-2012-0001 (Nemzeti Kiválóság Program - Hazai hallgatói, illetve kutatói személyi támogatást biztosító rendszer kidolgozása és működtetése országos program; támogatott: Várkúti Boglárka) pályázatok támogatásával készült.

Tartalom

Előszó	vii
1. A géntechnológia történetének mérföldkövei	1
1.1. Az örökítőanyag felfedezésétől a genetikai kód megfejtéséig	1
1.1.1. A DNS felfedezése és a kólibaktérium leírása	1
1.1.2. A DNS örökítőanyag szerepének tisztázása	2
1.1.3. A DNS kettőshélix felfedezése, a genetikai kód megfejtése	4
1.2. A DNS „manipulálására” alkalmas enzimek felfedezése	5
1.3. A rekombináns DNS technika megjelenése	6
1.4. Hibridizáción alapuló technikák megjelenése	7
1.5. A szekvenálástól a genom korszakig és tovább	7
1.5.1. A polimeráz láncreakció felfedezése	8
1.5.2. Géntechnológiai eljárások magasabbrendű szervezetekben	9
1.6. További olvasnivaló a fejezethez	10
2. A géntechnológia alapja: molekuláris klónozás	11
2.1. <i>In vitro</i> rekombináció	11
2.2. Rekombináns DNS létrehozása	12
2.3. Molekuláris klónok kialakítása	13
2.4. További olvasnivaló a fejezethez	14
3. DNS módosító enzimek és felhasználásuk	15
3.1. DNS-polimerázok	15
3.1.1. A DNS-polimeráz I és a Klenow-fragmentum	17
3.1.2. T4 DNS-polimeráz	18
3.1.3. T7 DNS-polimeráz	18
3.1.4. Hőstabil DNS-polimerázok	19
3.1.5. Reverz transzkriptáz	19
3.2. Nukleázok	20
3.2.1. Restriktív endonukleázok	20
3.2.2. Egyéb nukleázok	22
3.3. DNS-ligázok	23
3.4. Végmódosító enzimek	24
3.4.1. Terminális transzferáz	24
3.4.2. Alkalikus foszfatáz	25
3.4.3. T4 polinukleotid-kináz	25
3.5. További olvasnivaló a fejezethez	26
4. Alapmódszerek és rekombináns DNS konstrukciók tervezése	27
4.1. Alapmódszerek	27
4.1.1. A DNS tisztítása és analízise	27
4.1.2. Gélelektroforézis	30
4.1.3. Hibridizációs technikák és nukleinsav próbák	34
4.2. Rekombináns konstrukciók tervezése	40
4.2.1. Az inszert előállítása	41
4.2.2. A vektor előkészítése	42
4.2.3. Ligálás	42
4.2.4. TA/Topo-klónozás	44
4.2.5. Ligálás-független klónozás	45
4.2.6. Restriktív hasítás nélküli ligálás rekombinációval (attB, attP)	46
4.3. Génbeviteli eljárások	46
4.3.1. Transzformálás	46
4.3.2. Transzfektálás	47
4.3.3. Elektroporáció	47
4.3.4. Infekció	48
4.3.5. Gépüska	48
4.4. További olvasnivaló a fejezethez	48
5. DNS szekvenálás	50
5.1. Kémiai szekvenálás: radioaktív végjelölt templát	50

5.2. Enzimatiszus szekvenálás: láncterminációs módszer	50
5.2.1. Manuális láncterminációs szekvenálás	51
5.2.2. Automata fluoreszcens szekvenálás	53
5.3. Új-generációs szekvenálás	55
5.3.1. Piroszekvenálás	55
5.3.2. Illumina/Solexa szekvenálás	56
5.3.3. SOLiD (<i>Sequencing by Oligonucleotide Ligation and Detection</i>)	58
5.3.4. TSMS (<i>True Single Molecule Sequencing</i> : Valódi egymolekulás szekvenálás)	58
5.3.5. SMRT (<i>Single Molecule Real-time</i> : Egymolekulás valósídejű) szekvenálás	59
5.4. További olvasnivaló a fejezethez	59
6. Polimeráz láncreakció (PCR)	60
6.1. A PCR lépései	60
6.1.1. Denaturálás	60
6.1.2. Anellálás	61
6.1.3. Polimerizáció/lánchosszabbítás (extenzió)	61
6.2. A PCR reakció komponensei	62
6.2.1. PCR templát	62
6.2.2. PCR enzimek	63
6.2.3. PCR primerek	63
6.2.4. Egyéb PCR komponensek	64
6.3. A PCR termék (amplikon) detektálása és kvantitatív PCR	64
6.3.1. Agaróz és poliakrilamid gélelektroforézis	64
6.3.2. Valós idejű (<i>real-time</i>) vagy kvantitatív (q-) PCR	65
6.4. A PCR az alap kutatásban	66
6.4.1. A PCR termék (amplikon) klónozása és klónozás PCR-rel	66
6.4.2. Szekvenálás PCR készülékben	67
6.4.3. Mutagenézis	67
6.4.4. Reverz transzkripció (RT-) PCR	67
6.4.5. További PCR módszerek (random, fúziós, multiplex, fészkes és inverz)	68
6.5. PCR az alkalmazott kutatásban	70
6.5.1. Orvosi diagnosztika	70
6.5.2. Igazságügyi orvostan (DNS „ujjlenyomat”)	71
6.5.3. „Ős”-PCR	71
6.6. További olvasnivaló a fejezethez	72
7. Vektor-gazda rendszerek	73
7.1. A vektorok általános jellemzői	75
7.1.1. Klónozó és expressziós vektorok	75
7.1.2. Antibiotikum rezisztencia gének	76
7.2. Plazmid vektorok	77
7.2.1. A plazmidok általános jellemzői	77
7.2.2. A pBR322 vektor	78
7.2.3. A pUC vektorcsalád	78
7.2.4. Fágmidok	80
7.3. Bakteriofág vektorok	81
7.3.1. λ-fág vektorok	81
7.3.2. M13-alapú vektorok	84
7.4. Kozmidok	84
7.5. Mesterséges kromoszómák (PAC, BAC, YAC, HAC)	85
7.6. Élesztő és egyéb eukarióta vektorok és gazdasejtek	86
7.7. További olvasnivaló a fejezethez	87
8. Rekombináns DNS könyvtárak	88
8.1. A könyvtárak típusai	88
8.2. Genomiális DNS könyvtárak	88
8.2.1. A genomiális könyvtárak létrehozásáról általában	88
8.2.2. Könyvtár vektorok	90
8.2.3. λ-fág könyvtárak	90
8.2.4. Kozmid könyvtárak	95
8.2.5. Mesterséges kromoszóma könyvtárak (BAC és YAC)	96

8.2.6. Genomiális könyvtárak felhasználási területei	98
8.3. cDNS könyvtárak	99
8.3.1. cDNS könyvtárak létrehozása	99
8.3.2. cDNS könyvtárak szűrése és fenntartása	100
8.3.3. cDNS könyvtárak felhasználási területei	100
8.4. További olvasnivaló a fejezethez	100
9. Genom projektek	101
9.1. A genomszekvenálás alapjai	101
9.1.1. Térképezés	101
9.1.2. Genomszekvenálás térkép-alapú módszerrel	103
9.1.3. A <i>shotgun</i> módszer	104
9.2. Egy példa: a <i>Haemophilus influenzae</i> baktérium szekvenciájának meghatározása	104
9.2.1. A lefedettség (<i>coverage</i>) jelentősége	107
9.2.2. <i>Paired-end</i> szekvenálás	107
9.3. A Humán Genom Projekt (HGP)	108
9.3.1. Nehézségek a humán genom vizsgálatában	108
9.3.2. Elindul a HGP	109
9.3.3. Megjelenik egy versenytárs	109
9.3.4. A humán genom vázlatos (<i>draft</i>) szekvenciája	110
9.3.5. 2003-ra elkészül a szekvencia	110
9.3.6. A szekvenálás technikájának fejlődése	110
9.3.7. Etikai kérdések avagy kinek a DNS-ét használták a szekvenáláshoz?	110
9.4. További genom projektek	111
9.5. További olvasnivaló a fejezethez	111
10. Bevezetés a bioinformatikába	112
10.1. Bioinformatikai adatbázisok	112
10.1.1. Bibliográfiai gyűjtemények	112
10.1.2. DNS szekvencia-adatbázisok: GenBank és Ensembl	112
10.1.3. Fehérje-szekvencia adatbázisok: UniProt	113
10.1.4. Az "omikák" világa. Fajonkénti adatbankok.	114
10.1.5. Szerkezeti és funkcionális adatbázisok	114
10.2. Szekvencia-illesztések és egyéb fontos algoritmusok	116
10.2.1. Páronkénti szekvencia-illesztések (<i>pairwise alignments</i>)	116
10.2.2. A szekvencia-illesztés paraméterei	117
10.2.3. A BLAST algoritmus	118
10.2.4. Szekvenálás ellenőrzése illesztésekkel. Plazmid szerkesztők	118
10.2.5. Többszörös illesztések (<i>multiple alignments</i>). A Clustal programsomag	119
10.2.6. Evolúciósan konzervált elemek azonosítása	119
10.2.7. Automatizált módszerek a szekvenciák elemzésére	120
10.3. Nukleinsavak tulajdonságainak jóslása	123
10.3.1. Leolvasási keretek és gének jóslása	123
10.3.2. Gén meghatározás cDNS alapján. Eukarióta gének analízise	124
10.3.3. DNS és RNS szerkezetek jóslása. Oligonukleotid tervezés	124
10.3.4. Nukleinsavak "olvasáspontja"	125
10.4. Fehérjék tulajdonságainak jóslása	126
10.4.1. A fehérjék egyszerű fizikai tulajdonságainak becslése. Reverz transzláció	126
10.4.2. Lineáris motívumok keresése fehérjékben (és nukleinsavakban)	127
10.4.3. Fehérjeszerkezetek jóslása. Rendezett és rendezetlen régiók	127
10.5. <i>In silico</i> szerkezeti analízisek	129
10.5.1. Makromolekulák grafikai megjelenítése	129
10.5.2. Homológia-modellezés és <i>in silico</i> dokkolás	131
10.6. A predikciós módszerek pontossága	131
10.6.1. Egy predikció "jóságát" leíró paraméterek. Optimális döntések.	131
10.6.2. Adatbázisok minősége: elsődleges és másodlagos hibák	132
10.6.3. Ellenőrzött adatbázisok. Ellentmondó kísérleti eredmények kezelése	133
10.6.4. Mikor használjunk jóslásokat?	133
10.7. További olvasnivaló a fejezethez	133
11. <i>In vitro</i> mutagenézis	134

11.1. Nem-irányított <i>in vitro</i> mutagenesis módszerek	134
11.1.1. Mutátor baktérium sejtvonalak	134
11.1.2. Hibát ejtő (<i>error prone</i>) PCR mutagenesis	135
11.2. Helyspecifikus <i>in vitro</i> mutagenesis módszerek	136
11.2.1. DNS kazetta mutagenesis	136
11.2.2. Szelektív templát lebontáson alapú mutagenesis módszerek	137
11.2.3. PCR alapú irányított mutagenesis módszerek	141
11.3. További olvasnivaló a fejezethez	143
12. <i>In vitro</i> és prokarióta expressziós rendszerek	144
12.1. <i>In vitro</i> expressziós rendszerek	144
12.1.1. Az <i>in vitro</i> transzlációs rendszer komponensei	144
12.1.2. Előnyök és hátrányok	146
12.1.3. Alkalmazások	147
12.2. Prokarióta expressziós rendszerek	147
12.2.1. Bakteriális promóterek és a riboszóma kötőhely	148
12.2.2. A prokarióta fehérje expresszió optimalizálása	151
12.2.3. Fúziós rekombináns konstrukciók	154
12.2.4. Rekombináns fehérjék termelése és izolálása	158
12.2.5. A prokarióta expresszió előnyei, hátrányai	161
12.3. További olvasnivaló a fejezethez	161
13. Eukarióta expressziós rendszerek	162
13.1. Élesztő expressziós rendszer	162
13.1.1. Az expressziós kazetta	163
13.1.2. Szelekciós markerek	164
13.1.3. Szekréció	164
13.1.4. Poszttranszlációs módosítások	164
13.1.5. Az expresszió lépései	164
13.1.6. Élesztő expressziós vektorok és rendszerek	165
13.1.7. Az élesztő expressziós rendszer előnyei és hátrányai	166
13.2. Bakulovírus expressziós rendszer	167
13.2.1. „Bac-to-Bac” rendszer	167
13.2.2. Közvetlen bakulovírus (Direkt Bac) rendszer	168
13.2.3. A bakulovírus rendszer előnyei és hátrányai	168
13.3. Emlős expressziós rendszer	169
13.3.1. Emlős expressziós vektorok	169
13.3.2. Sejtvonalak	175
13.3.3. Génbevitel emlős sejtbe	176
13.3.4. Indukálható emlős expressziós rendszerek	176
13.4. Egyéb eukarióta expressziós rendszer	178
13.4.1. <i>Dictyostelium discoideum</i> expressziós rendszer	178
13.4.2. <i>Drosophila melanogaster</i> expressziós rendszer	179
13.5. Eukarióta sejttenyésztés	179
13.6. További olvasnivaló a fejezethez	180
14. Célzott génmódosítás (gén- <i>targeting</i>) módszerek	181
14.1. Bevezetés	181
14.2. DNS hibajavítás és homológ rekombináció	181
14.3. Gén- <i>targeting</i> konstrukciók előállítása	184
14.3.1. A <i>recombineering</i> alapjai. A Red rekombinációs rendszer	184
14.3.2. Gateway klónozás	185
14.3.3. <i>Targeting</i> vektorok tervezése	186
14.3.4. A gén- <i>targeting</i> konstrukciók elkészítése	187
14.4. A <i>targeting</i> -konstrukció bejuttatása a célszervezetbe	188
14.4.1. Sejtek és sejtvonalak transzfektálása. Emlős embriók manipulációja	188
14.4.2. Stabil transzfekció. Transzgének célzás nélküli bevitel	190
14.4.3. Növényi és gomba sejtek transzformálása. Homológ rekombináció növényekben	190
14.4.4. Növények genetikai módosítása Ti plazmid segítségével	190
14.5. <i>Knock-out</i> és <i>knock-in</i> technikák	192
14.5.1. <i>Homing</i> endonukleázok. Cre/Lox és Flp/Frt rendszerek	192

14.5.2. Célszekvenciára tervezett hasító enzimek. Cink-ujjas és TALE-nukleázok	194
14.5.3. Génkiütési technikák (<i>knock-out</i>) és új gének célzott beépítése (<i>knock-in</i>)	195
14.5.4. Feltételes génkiütés, génjavítás	197
14.5.5. Genetikai elemek beépítése transzpozonok segítségével	198
14.6. Géncsendesítés (<i>knock-down</i>) RNS interferenciával	201
14.7. Stabil episzomális rendszerek	203
14.8. További olvasnivaló a fejezethez	203
15. Transzgenikus élőlények. Génterápia	204
15.1. A transzgenikus élőlények típusai és felhasználásuk	204
15.1.1. Transzgenikus mikrobák	204
15.1.2. Transzgenikus növények	206
15.1.3. Transzgenikus állatok az alap kutatásban és a gyógyászatban	210
15.2. Génterápia	217
15.3. A transzgenikus élőlények felhasználásának tudományetikai és környezetbiztonsági kérdései	218
15.4. További olvasnivaló a fejezethez	219
16. Irányított evolúciós technikák	220
16.1. Az irányított evolúciós megközelítésről általában	220
16.2. Könyvtárkészítési technikák	222
16.2.1. PCR alapú könyvtárkészítési eljárások	222
16.2.2. Helyspecifikus mutagenézisen alapuló eljárások	223
16.3. Szelekciós eljárások	226
16.3.1. <i>In vitro</i> szelekciós módszerek	226
16.3.2. A fág-bemutató	227
16.3.3. Rekombináns fehérje könyvtár bemutatása és szelekciója élő sejtek felszínén	229
16.4. További olvasnivaló a fejezethez	230
17. Kölcsönhatások kimutatása géntechnológiai módszerekkel	231
17.1. Az élesztő két-hibrid rendszer bemutatása	231
17.2. Az élesztő két-hibrid rendszer gyengeségei	232
17.3. Hasonló alapokon működő technikák	233
17.3.1. Fordított két-hibrid rendszerek	233
17.3.2. Egy-hibrid rendszerek	233
17.3.3. Három-hibrid rendszerek	234
17.4. További olvasnivaló a fejezethez	234
18. Függelék	235

Előszó

A **géntechnológia és fehérjemérnökség** című elektronikus jegyzet elsősorban a mesterképzésben molekuláris szintű biológiát tanuló egyetemi hallgatók számára készült, de bármely, a rekombináns DNS technológia iránt érdeklődő hallgató, vagy már végzett kutató számára is hasznos információkat tartalmaz.

Mivel a jegyzet fő célközönsége mesterszintű képzésben tanul, feltételezzük, hogy az olvasó korábban már sikeresen elvégzett biokémia, genetika és mikrobiológia kurzusokat. Ezen tudás nélkülözhetetlen a jegyzetben leírt módszerek elméleti hátterének megértéséhez.

A jegyzet szerkezetét és tartalmát az ELTE-n idestova több mint húsz éve tanított géntechnológia tárgy oktatása során szerzett tapasztalatok alapján állítottuk össze. Ezúton fejezem ki köszönetemet ma már nyugdíjas kollégámnak, **Szilágyi Lászlónak**, aki sok évig oktatta a géntechnológia tárgyat, aki az elsők között alkalmazta a Biokémiai Tanszéken a géntechnológiai módszereket a kutatómunkájában, s akivel közösen alakítottuk ki a tárgy tematikáját. Mivel a jegyzetírásban közreműködő szerzőgárda minden tagja az ELTE Biokémia Tanszék fiatal munkatársa, ahol hagyományosan fehérjetudománnyal foglalkozunk, ezért a könyv hangsúlyozottan „biokémikusi” szemléletre épül. Ebből a tényből adódik, hogy a fehérjetudományban általunk is használt géntechnológiai módszereket nagyobb részletességgel tárgyaljuk, viszont nem megyünk bele mélyen például a géntechnológiában nélkülözhetetlen mikrobiális és eukarióta sejtenyésztési módszerek részletezésébe. Szintén kevesebb teret szentelünk a molekuláris genetikai, funkcionális genomikai, molekuláris sejtbiofizikai, molekuláris mikrobiológiai, virológiai, növénybiológiai és más molekuláris szintű biológiai kutatásokban alkalmazható „speciális” módszerek bemutatásának.

A jegyzet címe is a biokémikus szemléletre utal, hiszen a fehérjemérnökség a fehérjetudomány része. Ennek, a neve alapján alkalmazott tudományágnak a célja valamilyen szempontból jobb, hasznosabb fehérjék előállítására. De nem csak alkalmazott tudományról van szó, hiszen a fehérjék szerkezet-funkció összefüggéseinek vizsgálata ma már legnagyobb részt rekombináns fehérjékkel történik, és sokszor a fehérjeláncokat kódoló DNS szekvenciájának megváltoztatásán keresztül (*in vitro* helyspecifikus mutagenézissel) igyekszünk megérteni a fehérje biológiai szerepét – s módszertani megközelítésben ekkor is fehérjemérnökséget művelünk. A jegyzetnek nem célja egy szisztematikus „fehérjemérnökség” tárgy leírása, ezért nem foglalkozunk részletesen azokkal a konkrét „tervezett fehérjékkel”, amelyeket a géntechnológia ezen ága máig előállított. A hangsúly a fehérjemérnökséghez szükséges géntechnológiai eljárások leírásán lesz. Igen alaposan tárgyaljuk a rekombináns fehérjék előállításának lehetőségeit (expressziós rendszerek) és a szerkezet megváltoztatásának lehetőségeit (mutagenézis módszerek). Külön fejezetet szentelünk az irányított evolúció - izgalmas lehetőségeket megteremtő - módszereinek bemutatására. Bizonyos értelemben a fehérjemérnökség termékei azok az egyre nagyobb számban felhasználásra kerülő fehérjegyógyszerek is, amelyek előállításával, ha nem is részletekbe menően, de szintén foglalkozunk.

A géntechnológia a biotechnológiai iparág része. Nem tárgya a jegyzetnek a géntechnológia ipari alkalmazásainak mélyebb tárgyalása, mivel hangsúlyozottan a biológus kutatóképzésben résztvevő hallgatókat céloztuk meg. Ennek ellenére úgy véljük, hogy a vegyész, biomérnök, bionikai, orvosi és biotechnológia szakemberképzés területén is haszonnal forgatható jegyzetet kap kézhez az olvasó.

Elektronikus jegyzetről lévén szó, igyekeztünk az ebből a tényből adódó lehetőségeket kihasználni. A jegyzet nagyszámú kereszthivatkozás, internetes hivatkozást tartalmaz. Az utóbbiak segítségével az olvasó a jegyzet ismeretanyagát messze meghaladó, a tanulás hatékonyságát növelő további témérdek információhoz juthat hozzá. A módszerek gyakorlatban történő használhatóságát segíti a jegyzetbe integrált néhány videofilm, de ezt a célt szolgálja egyes módszerek praktikus szempontokat is megemlítő leírása. Jegyzetünk elméleti anyag tehát nem tartalmaz részletes gyakorlati „recepteket”, bár több helyütt hiperhivatkozások formájában ilyen weboldalakat is elérhet az olvasó.

Szerkezetét tekintve a jegyzet egy, a **géntechnológia történetét** a szűkebb értelemben vett molekuláris biológia (a nukleinsavak kutatása) történetével együtt felvázoló fejezettel kezdődik. A géntechnológia alapját jelentő **molekuláris klónozás** bemutatása után sorra vesszük a terület műveléséhez szükséges eszközöket és alapmódszereket. Bemutatjuk a **DNS „manipuláció” enzimeit**, a DNS molekulák szeparálásához szükséges **elektroforézis technikákat**, a nélkülözhetetlen **hibridizációs és szekvenáló módszereket**, a klónozáshoz használható **vektorok** típusait. Sorra vesszük a **rekombináns DNS tervezésének** egyes lépéseit, a klónozási stratégiákat, a DNS konstrukciók sejtbe juttatási módszereit (**génbevitel**), az *in vitro* DNS felszaporítást lehetővé tevő **polimeráz**

láncreakció alapjait és felhasználási területeit. Tárgyaljuk a különböző **génkönyvtárakat**, bemutatjuk a **genom projektek** módszertani hátterét és néhány konkrét genom programot is (kiemelten a Humán Genom Programot). Igen alapos és részletes fejezet foglalkozik a géntechnológia műveléséhez szintén nélkülözhetetlen **bioinformatika** releváns területeivel és alkalmazásaival. A **fehérjemérnökség tárgykörébe** elsősorban az *in vitro* mutagenézist és az összes lehetséges **expressziós rendszert** (*in vitro*, **prokarióta**, **eukarióta** **gazdasejtek**, **transzgenikus élőlények**) ismertető fejezetek tartoznak. Külön fejezetet szentelünk az igen fontos ún. **gén targeting (célzott génbeviteli) módszereknek**, amelyek segítségével organizmus szinten módosíthatunk célzottan egy-egy gént (*knock-out*, *knock-in* technika), illetve a nem túl régen felfedezett RNS csendesítés jelenségét (*knock-down*) kihasználva tehetünk tönkre egy-egy mRNS molekulát. Szintén külön fejezet foglalkozik a **transzgenikus élőlények** előállításának hátterével, s néhány GMO példával a mikroorganizmusok, a növények és az állatok köréből - de nem próbálunk átfogó leírást adni erről a területről. Röviden bemutatjuk a ma még inkább jövőbeni alkalmazást jelentő **génterápiás** eljárások alapjait. Végül sort kerítünk a már említett **irányított evolúciós módszerek** bemutatására és ismertetünk néhány specifikus, a fehérjetudományban alkalmazott géntechnológiai ún. **interakciós módszert** is. Az egyes fejezetek végén a témához kapcsolódó néhány szakirodalmi forrást, további olvasnivalót sorolunk fel.

Bízunk benne, hogy az e-jegyzetünk eléri a megfelelő célközönséget, s mind egyetemi hallgatók, mind a géntechnológia iránt érdeklő szakemberek hasznosnak találják majd az elektronikus oldalak lapozgatását. Az egyik fő célunk, hogy a hallgatóink a **Géntechnológia kurzus** vizsgakövetelményét az e-jegyzet használatával könnyebben és eredményesebben teljesítsék. A másik fő célunk, hogy az olvasó ne csak elméleti háttéranyagként használja az e-jegyzetet, hanem a kísérleti munkájának megtervezéséhez és kivitelezéséhez is hasznosítható információkat találjon benne, legyen szó laboratóriumi gyakorlatokról, tudományos diákköri és szakdolgozati munkáról vagy géntechnológiai módszerek használatát megkívánó alap- és alkalmazott kutatási projektekről.

Budapest, 2013. május

Nyitrai László, szerkesztő

1. fejezet - A géntechnológia történetének mérföldkövei

A géntechnológia születését megelőzte a DNS szerepének és szerkezetének megismerése, az öröklődés molekuláris mechanizmusainak feltárása. Ezért az alábbi történeti áttekintést a DNS felfedezésénél kezdjük, s a nukleinsav biokémia mérföldköveinek (és a kulcs kísérletek) felvázolásán át jutunk el azoknak a módszereknek a leírásáig, amelyek lehetővé tették az első rekombináns DNS létrehozását. Áttekintjük a technológia fejlődését jelző fő módszertani újításokat, azt a fejlődést, ami az egyes gének vizsgálatától elvezetett a genom projektekig és a nagy áteresztőképességű módszerek megjelenéséig.

1.1. Az örökítőanyag felfedezésétől a genetikai kód megfejtéséig

1.1.1. A DNS felfedezése és a kólibaktérium leírása

A géntechnológia kialakulásához, a rekombináns DNS technikák megjelenéséhez vezető úton az első fontos lépést **Friedrich Miescher** (1844-1895) tette meg **1869**-ben a **DNS felfedezésével**. Miescher a leukociták kémiai összetételét vizsgálta. A sejtmagokból egy savban oldhatatlan, viszont lúgban oldódó, magas foszfor tartalommal jellemezhető anyagot sikerült elkülönítenie, ami a fehérjékkel ellentétben hőkezelés hatására nem csapódott ki az oldatból. Később ugyanezt az anyagot sikerült izolálnia úgy is, hogy pepszines kezeléssel lebontotta a mintában található fehérjét. A fehérjétől eltérő összetétel és viselkedés, valamint a proteáz kezeléssel szembeni rezisztencia arra mutattak, hogy Miescher egy új, korábban ismeretlen anyagot talált. Ezt a sejtmagokra jellemző anyagot „nukleinnek” nevezte el.



1.1. ábra: Friedrich Miescher

Theodor Escherich (1857-1911), német gyermekorvos és mikrobiológus a bélrendszer baktériumflóráját vizsgálva számos addig ismeretlen baktériumot azonosított, közöttük egy Gram-negatív, fakultatív anaerob baktériumot, melynek a „*Bacterium coli commune*” nevet adta. Eredményeit **1886**-ban publikálta. A halála után ***Escherichia coli*** nevet kapott baktérium napjainkra a géntechnológiai kutatások és a géntechnológiai ipar egyik legfontosabb eszközévé vált, melyben könnyű tenyésztetősége, gyors osztódása (ideális körülmények között 20 percenként osztódik), egyszerű genetikai manipulálhatósága játszottak szerepet.

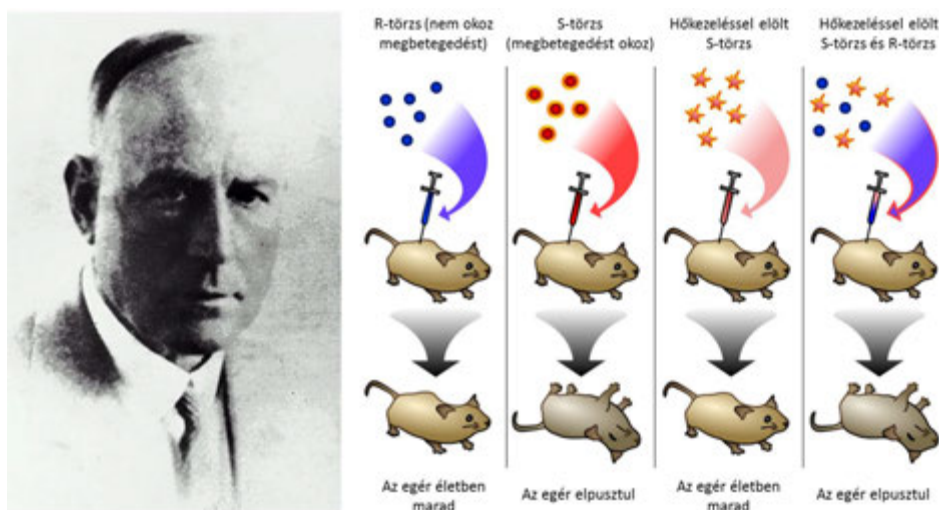


1.2. ábra: Theodor Escherich arcképe és az *Escherichia coli* elektronmikroszkópos képe (mérete 2 x 0,5 mm)

1.1.2. A DNS örökítőanyag szerepének tisztázása

Frederick Griffith (1879-1941) brit bakteriológus, a tüdőgyulladást (pneumonia) okozó *Streptococcus pneumoniae* baktériumot vizsgálva **1928**-ban felfedezte a transzformáció jelenségét. A baktérium két törzsét használta fel kísérletei során, melyeket a telepek morfológiája alapján tudott elkülöníteni: a „sima” törzs (S: *smooth*) és a „durva” törzs (R: *rough*). Az S-törzs képes tüdőgyulladást okozni, míg az R-törzs nem. Egerekbe injektálva a baktériumokat az S-törzs esetén a kialakuló tüdőgyulladás az egér pusztulását okozza, míg az R-törzs esetén az egér túlél. Amikor Griffith hőkezeléssel elpusztította az S-törzsbe tartozó baktériumokat, majd egerekbe injektálta az így nyert anyagot, az egerek nem betegedtek meg. Ha azonban az elpusztított S-törzshöz élő R-törzsbe tartozó baktériumokat kevert, az egér megbetegedett és elpusztult, annak ellenére, hogy önmagában az R-törzs és az élőlt S-törzs is ártalmatlannak bizonyult a korábbi kísérletekben. Az elpusztult egerek szervezetében mindkét törzs jelenlétét ki tudta mutatni. Az elpusztított baktériumokból kikerülő anyagok tehát „átalakították” valamilyen módon az ártalmatlan R-törzsbe tartozó baktériumokat, vagyis a fertőzés képessége átadódott. Griffith a jelenséget „transzformációnak” nevezte el.

(Ma már tudjuk, hogy az S-törzsbe tartozó baktériumok rendelkeznek egy poliszacharid burokkal, mely védelmet biztosít a gazdaszervezet immunrendszere ellen. Az R-törzs esetén ez hiányzik. A fertőzés képessége, vagyis a poliszacharid burok létrehozásához szükséges genetikai információ, a hőkezelés során épségben megmaradt, majd az R-törzs baktériumai által a külvilágból felvett DNS molekulák közvetítésével jutott át a korábban fertőzni nem képes sejtekbe.)



1.3. ábra: Frederick Griffith és Griffith transzformációs kísérlete (GNU Free Documentation License, szerző: Madeleine Price Ball)

A transzformációért felelős, vagyis a genetikai információt hordozó anyag kémiai természetének azonosítása **Oswald Avery** (1877-1955), **Colin MacLeod** (1909-1972) és **Maclyn McCarty** (1911-2005) munkájának köszönhető. Az **1944**-ben publikált kísérletben Griffith kutatásaihoz hasonlóan ők is a *Pneumococcus* baktérium R- és S-törzszeit

használták. Az S-törzsből sikerült izolálniuk a transzformációért felelős anyagot. Különböző kémiai módszerekkel vizsgálva a preparátumot kimutatták, hogy az nagy valószínűséggel DNS-ből áll, és nem tartalmaz fehérjét. (Akkoriban a fehérjét tartották a genetikai információ hordozójának.) Ezt követően különböző enzim preparátumokkal kezelték az anyagot (pl. proteáz, RN-áz, DN-áz). Transzformációt csak akkor nem tapasztaltak, ha a preparátumnak DN-áz aktivitása volt. Egyéb enzimek viszont nem gátolták transzformációt.



1.4. ábra: Oswald Avery

Avery és munkatársai eredményeit akkoriban a tudományos közvélemény egy része kritikusan fogadta, mivel logikusabbnak tűnt a 20 féle aminosavból álló fehérjékről feltételezni, hogy örökítőanyagként funkcionálnak. A perdöntő kísérletet **1952-ben Alfred Hershey** (1908-1997) és **Martha Chase** (1927-2003) hajtották végre, akik T2 bakteriofággal kísérleteztek. A fágreszcsekről ismert volt, hogy fehérje burokkal rendelkeznek, melynek belsejében DNS található. Kihasználva, hogy a fehérjealkotó aminosavak a DNS-sel ellentétben tartalmaznak kén, viszont nem tartalmaznak foszfort, Hershey és Chase szelektíven jelölhette radioaktív kén (^{35}S) és foszfor (^{32}P) izotópok felhasználásával a fehérjeburkot és a DNS-t. Ehhez olyan baktériumokon szaporították a fágokat, melyek vagy ^{35}S , vagy ^{32}P tartalmú táptalajon növekedtek. Az ilyen módon jelölt fágreszcsekkel azután olyan baktériumokat fertőztek, melyek jelöletlen táptalajon növekedtek. A ^{35}S izotóp nagyrészt a baktériumsejteken kívül maradt, míg a ^{32}P a sejtek belsejébe került, vagyis fertőzéskor a fág DNS bejutott a sejtekbe, míg a fehérjeburok elemei nem. Ráadásul a fertőzés után újonnan képződő fágreszcsek a ^{32}P -t „megörökölték” a szülői generációtól, míg a ^{35}S -t nem.

Eredményeik elismeréseképpen Hershey (Max Delbrück-kel és Salvador Luria-val megosztva, akik szintén bakteriofágokkal végzett kísérleteikben a replikáció és a mutáció problémakörét vizsgálták) 1969-ben fiziológiai és orvostudományi Nobel-díjat kapott.

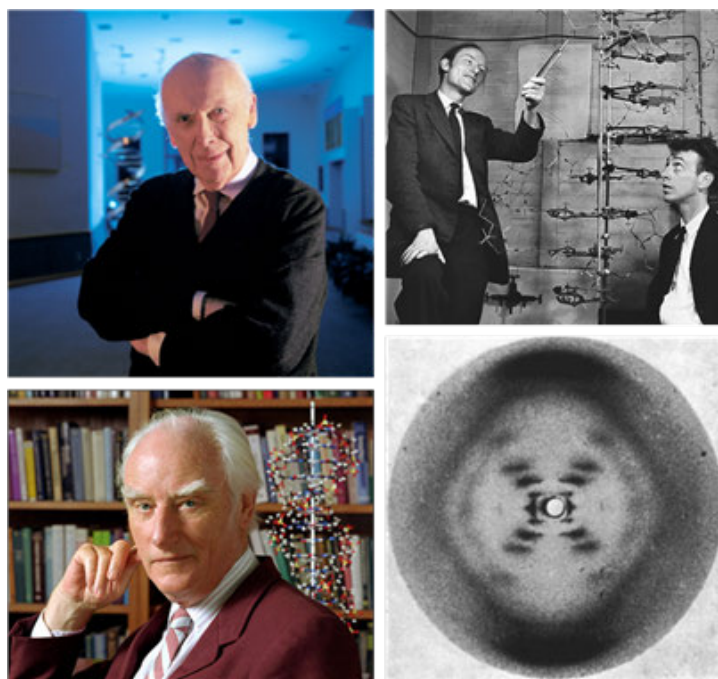


1.5. ábra: Alfred Hershey és Martha Chase

1.1.3. A DNS kettőshélix felfedezése, a genetikai kód megfejtése

A géntechnológia szempontjából kulcsfontosságú esemény volt a DNS szerkezetének felfedezése, amely **James D. Watson** (1928-) és **Francis H. C. Crick** (1916-2004) nevéhez fűződik. Watson és Crick a **DNS kettőshélix modelljét** Maurice Wilkins (1916-2004) és Rosalind Franklin (1920-1958) röntgendiffrakciós adatai alapján állította fel **1953**-ban. A modell helyesen feltételezte, hogy a DNS két, egymással ellentétes irányba futó szálát tartalmaz. Az **Erwin Chargaff** (1905-2002) által megfigyelt **szabályok** alapján (az adenin és a timin, továbbá a citozin és a guanin aránya megegyezik a DNS-ben) helyesen feltételezték azt is, hogy a szálak közötti kapcsolatot a bázisok között specifikus módon kialakuló hidrogén-kötések jelentik. Watson és Crick modellje rávilágított a DNS megkettőződésének mechanizmusára is: mivel bármelyik szál bázissorrendje egyértelműen meghatározza a másik szál bázissorrendjét, a két szál elválasztásával nem vesz el információ. Az elválasztott szálak mintaként (templátként) szolgálhatnak új, komplementer szálak szintéziséhez. Így két új kettőshélixhez jutunk, melyeknek csak az egyik szála újonnan szintetizált. Ezt a mechanizmust szemikonzervatív replikációnak nevezzük.

Watson, Crick és Wilkins 1962-ben fiziológiai és orvostudományi Nobel-díjat kapott a DNS szerkezetének felfedezéséért. Rosalind Franklin, akinek a hozzájárulása az eredményekhez szintén kulcsfontosságú volt, sajnos nem részesülhetett ebben az elismerésben, ugyanis négy évvel korábban, 1958-ban elhunyt.



1.6. ábra: A. James Watson (balra fent), Francis Crick (balra lent); Watson, Crick és a DNS kettőshélix modellje (jobbra fent); A Rosalind Franklin által készített DNS szál-diffrakciós felvétel, mely az egyik kulcsfontosságú bizonyíték volt a modell helyességét illetően (jobbra lent).

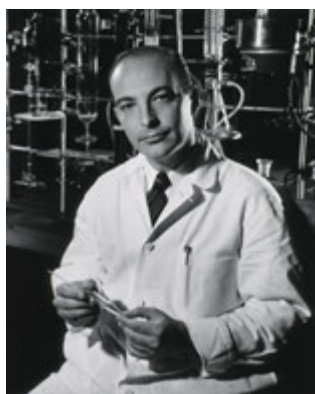
A DNS szerkezetének megismerésével azonban még mindig nyitva maradt egy nagyon fontos kérdés: milyen módon hasznosul a DNS molekulában tárolt genetikai információ a sejtekben, vagyis hogyan képes a DNS a fehérjeláncok szintézisét kódolni? **Marshall Nirenberg** (1927-2010) volt az első, aki sikereket ért el a genetikai kód megfejtése terén. Kísérleteihez *E. coli* sejtek citoplazmáját használta fel, melyben jelen volt a fehérjeszintézishez szükséges minden komponens. A mintához szintetikus RNS-t (poli-U) és a 20 fehérjealkotó aminosav közül egyet-egyet radioaktívan jelölve adott. Megfigyelte, hogy poli-U templát esetén csak a fenilalanin épül be a fehérjeláncba. További szintetikus RNS molekulákkal is elvégezte a kísérletet; ekkor más-más aminosavak épültek be. **Har Gobind Khorana** (1922-2011) rövid oligonukleotidok szintézisét kombinálta enzimatikus módszerekkel, így képes volt előállítani olyan RNS polimereket, melyekben kettő, három, illetve négy nukleotid egység ismétlődött. Ezeket – Nirenberg kísérleteihez hasonlóan – sejtmentes expressziós rendszerben tesztelte. Két nukleotid szabályos ismétlődése esetén a polipeptidláncba két féle aminosav épült be, szabályosan váltakozva. Három nukleotid esetén egyetlen aminosavból álló szabályos polimer jelent meg. Négy bázis szabályos ismétlődése négy aminosav

szabályosan ismétlődő polimerét eredményezte. Mindez igazolta, hogy a genetikai kódban egy aminosavat három bázis definiál, továbbá, hogy a bázishármasok nem fednek át.

Nirenberg, Khorana és Holley (az utóbbi a tRNS-t fedezte fel) 1968-ban a genetikai kód megfejtéséért megkapta a Nobel-díjat.

1.2. A DNS „manipulálására” alkalmas enzimek felfedezése

Az örökítőanyag kémiai természetének, szerkezetének megismerése, valamint a genetikai kód megfejtése mellett a géntechnológia kialakulásához szükséges volt a DNS „manipulálására” alkalmas enzimek felfedezése is. Az egyik legfontosabb enzim a DNS szintéziséért felelős **DNS-polimeráz**, amit **Arthur Kornberg** (1918-2007) fedezett fel **1957**-ben. Kornberg munkásságát 1959-ben Nobel-díjjal jutalmazták (Severo Ochoa-val megosztva, aki először bizonyította, hogy kémcsőben RNS-t is lehet szintetizálni)



1.7. ábra: Arthur Kornberg

Werner Arber (1929-) a **bakteriális restrikciós-modifikációs rendszer** vizsgálatával **1962**-ben olyan enzimeket fedezett fel, melyek nélkül a rekombináns DNS technológia nem jöhetett volna létre. Arber a bakteriofágok gazdasejt specifikálásának okait vizsgálta. Megfigyelte, hogy az *E. coli* K12 törzsben felszaporított bakteriofágok alacsony határfokkal fertőznek egy másik törzset, az *E. coli* B-t, és fordítva: a B törzsben-ben szaporított fágok rossz hatékonysággal fertőznek meg a K12-es törzset. Ha viszont megtörténik a keresztfertőzés az újonnan termelt fágok már jól szaporodnak az új gazdasejtben és alacsony határfokkal fertőzik az eredeti gazdasejtet, amiben korábban jól szaporodtak. Arber felfedezte, hogy a jelenség hátterében a DNS metilációja (modifikáció), és endonukleázok általi, metilációtól függő hasítása áll.

A restrikciós-modifikációs rendszernek tehát két eleme van: egy **restrikciós endonukleáz** és egy annak megfelelő, **DNS-metiláz** enzim. A restrikciós endonukleáz specifikusan felismer egy 4-8 bázisnyi DNS szekvenciát és hasítja a DNS mindkét szálát. A módosító metiláz ugyanebben a szekvenciában metilál egy adenint (az aminocsoporton), vagy egy citozint (az 5. szénen vagy az aminocsoporton). A restrikciós endonukleáz nem tudja elhasítani az így módosított szekvenciát. A két enzim együttesen van jelen a sejtben, és jelenlétük törzs-specifikus. A baktérium DNS-e a metiláznak köszönhetően állandóan metilált, tehát védve van a restrikciós endonukleáz működésével szemben. A külvilágból, pl. egy fágból származó DNS – feltéve hogy nem ugyanabból a baktérium törzsből érkezik – nincs metilálva ugyanezeknél a specifikus szekvenciáknál, ezért a restrikciós endonukleázok nagy valószínűséggel elhasítják és így lebomlik.

Az **első restrikciós endonukleázt** **Hamilton Smith** (1931-) és **Daniel Nathans** (1928-1999) izolálta az 1960-as évek végén. Nathans **1972**-ben elsőként használta a restrikciós enzimeket genetikai térképezésre. (Az SV40 vírus restrikciós térképezését végezte.) Napjainkra már több ezer restrikciós enzimet ismerünk, melyek több száz különböző szekvencia specifikus elhasítását teszik lehetővé. Arber, Smith és Nathans a restrikciós endonukleázokkal végzett kutatásokért 1978-ban megkapta a Nobel-díjat.

Martin Gellert (1929-) fedezte fel a **DNS-ligáz** nevű enzimet **1967**-ben, amely képes a DNS molekulák végeit foszfodiészter kötésekkel kovalensen összekapcsolni, ezzel megteremtve a lehetőségét a különböző forrásokból származó DNS molekulák egyesítésének, így a rekombináns DNS molekulák előállításának.



1.8. ábra: Werner Arber, Hamilton Smith és Daniel Nathans (HS arcképe: Creative Commons License, szerző: Jane Gitschie; DN arcképének forrása: National Library of Medicine)

1.3. A rekombináns DNS technika megjelenése

Az első, valódi rekombináns DNS molekulák létrehozása az **1972**-ben történt meg. **Paul Berg** (1926-) elsőként dolgozott ki egy enzimekre épülő módszert, két, különböző eredetű DNS molekula (SV40 vírus és λ -fág) egyesítésére. A cirkuláris DNS molekulákat restrikciós endonukleázal hasította. A lineáris DNS-t terminális dezoxiribonukleotid-transzferáz enzimmal kezelte, mely enzim a kétszálú DNS egyik szálát képes meghosszabbítani a szubsztrátként adott dNTP-kel. Csak egyféle monomert adva az enzim egyszálú, homopolimer „farkat” készít a DNS-en (pl. dATP esetén poli-A). A kétféle forrásból származó DNS molekulán így egymással komplementer (poli-A és poli-T) túlnyúló végeket tudott létrehozni, melyek Watson-Crick féle bázispárosodással felismerték egymást. A DNS láncok egyesítéséhez ezután DNS-polimerázt és DNS-ligázt használt. Munkásságát 1980-ban Nobel-díjjal jutalmazták.

Stanley Cohen (1935-) **1973-ban** *E. coli* sejtek transzformációjára dolgozott ki hatékony eljárást, mellyel plazmidokat (extrakromoszomális, cirkuláris, kétszálú DNS molekulák, melyek önállóan replikálódnak) tudott bejuttatni a sejtekbe. A baktériumok a plazmidon található antibiotikum rezisztencia gén hatására ellenállóvá váltak az antibiotikummal szemben; annak jelenlétében is tudtak növekedni. Cohen a következő évben **Herbert Boyer**-rel (1936-) együttműködve létrehozott egy **rekombináns plazmidot** két másik plazmid restrikciós endonukleáz kezelés hatására létrejött darabjaiból. A fragmentumokat DNS-ligáz segítségével illesztették össze. Az új plazmiddal sikerült *E. coli* sejteket transzformálni, melyek mind a két kiindulási molekulától származó gént (antibiotikum rezisztencia) hordozták. Ezzel amellet, hogy rekombináns DNS-t hoztak létre, demonstrálták azt is, hogy **a plazmidok vektorként (hordozó DNS) használhatók**, vagyis más forrásból származó DNS darabok beléjük illeszthetők, majd stabilan fenntarthatók *E. coli* sejtekben, ahol a plazmid jelenlétére a rajta található antibiotikum rezisztencia gének segítségével lehet szelektálni. Herbert Boyer és Robert Swanson **1976-ban** megalapították a **Genentech** nevű céget, mely a **géntechnológiai ipar** kezdetét jelentette. A Genentech a rekombináns DNS technika gyakorlati alkalmazását valósította meg. A legelső sikeres fejlesztés a humán inzulin hormon *E. coli* baktériumban történő termeltetése volt 1978-ban („Humulin” néven került forgalomba 1982-ben.)



1.9. ábra: Paul Berg, Herbert Boyer és Stanley Cohen

A rekombináns DNS technika megjelenése azonban a tudományos közvélemény jelentős részében aggodalmakat is keltett. Felismerték a különböző eredetű DNS molekulák összekapcsolásával létrehozható új vírusok illetve baktériumok veszélyeit. Az ilyen mesterségesen összeállított genommal rendelkező élőlények igen komoly humán egészségügyi kockázatot jelenthetnek. Emellett a laboratóriumból kikerülve előre meg nem jósolható hatásuk lehet más élőlényekre is. Paul Berg az első rekombináns DNS molekula létrehozójaként 1974-ben javasolta, hogy a veszélyekre való tekintettel egészen addig ne folytassanak kísérleteket rekombináns DNS-sel, amíg a megfelelő szabályozás és a biztonsági előírások meg nem születnek. A kérdés megvitatására **1975** februárjában megszervezte a híressé vált **Asilomar-konferenciát**. A kísérletek felfüggesztése (moratórium) végül 16 hónapig, a megfelelő biztonsági előírások és irányelvek kidolgozásáig tartott.

1.4. Hibridizáción alapuló technikák megjelenése

Julius Marmur (1926-1996) és **Paul Doty** (1920-2011) a DNS hő hatására bekövetkező denaturációját vizsgálták az **1960-as évek elején**, mely munkával igen komoly befolyást gyakoroltak a rekombináns DNS technológia fejlődésére. Hőkezelés hatására a DNS két szála elválik egymástól, és a minta fényelnyelése megnő (260 nm-es hullámhossz környékén a bázisok elnyelik az UV sugárzást). A jelenség **hiperkróm effektus** néven ismert. A hőmérséklet csökkentésével a DNS minta renaturálódik, vagyis egymással komplementer szekvenciák esetén a hidrogénkötések újra kialakulnak a megfelelő bázispárok között. A DNS denaturáció, renaturáció (különböző eredetű, de komplementer szekvenciával rendelkező szálak esetén **hibridizáció**) jelenségét számtalan, a mai napig használatos módszer kihasználja, így például az **Edwin Southern** (1938-) által **1975**-ben kidolgozott – és róla elnevezett – Southern-blot (lenyomat) módszer. A Southern-blot egy adott szekvenciával rendelkező DNS molekula specifikus kimutatására szolgáló eljárás. (A hasonló elven működő kimutatási eljárást RNS minta esetén Northern-blotnak nevezzük.)

Michael Smith (1932-2000) a **helyspecifikus mutagenézis** módszerének kidolgozásáért (**1978**), míg **Kary Mullis** (1944-) a **polimeráz láncreakció** (PCR: *Polymerase Chain Reaction*) módszerének megalkotásáért (**1983**) kapott Nobel-díjat 1993-ban. Mindkét módszer kihasználja a DNS láncok hibridizációs képességét. Smith a Φ X174 bakteriofág cirkuláris, egyszálú DNS-ének rövid szakaszaival komplementer szintetikus oligonukleotidokat szintetizált, melyeket úgy tervezett meg, hogy mind az 5'-, mind a 3'-végeik hibridizáltak a fág DNS egy-egy szakaszával, ám a középső szakasz mutációt tartalmazott. DNS-polimeráz segítségével ezután a hibridizált oligonukleotid 3'-végétől egészen annak 5'-végéig szintetizálta a komplementer DNS láncot, majd DNS-ligázzal egyesítette az új szál végeit. Az így keletkezett cirkuláris DNS-nek tehát csak az új szála hordozta a mutációt. A DNS-t baktériumokba juttatva a keletkező fágok egy része az eredeti, míg más része a mutáns szekvenciát hordozta. Ezzel a módszerrel tehát célzottan, bármely két ismert szekvenciarészlet között mutáció hozható létre. A **polimeráz láncreakció** esetén egy kétszálú DNS molekula („templát”) tetszőleges szakaszáról készíthetünk specifikus módon nagyszámú másolatot. A szakaszt **két**, a templát DNS egy-egy szálához hibridizáló **szintetikus oligonukleotid** jelöli ki. A módszer felfedezését az **1.5.1. fejezetben** részletezzük.

A DNS hibridizáció jelenségét használja ki a nagy áteresztőképességű kísérletek elvégzésére alkalmas „**DNS-chip**” (**DNS microarray**) technológia is, amelyeknek egyik első változatát **Stephen Fodor** **1991**-ben fejlesztette ki. A DNS-chip lényege, hogy egy szilárd hordozó felszínén *in situ* kovalens kötéssel kapcsoltan oligonukleotid próbákat szintetizálnak fotolitográfiai eljárással mikroszkopikus méretű foltokban, előre meghatározott elrendezésben. Adott területen csak egyféle oligonukleotid található, melynek ismert a szekvenciája. Több tízezer ilyen folt is lehet egyetlen chip felszínén. Az így létrehozott szintetikus próbákat felhasználhatjuk például arra, hogy sok gén expressziós szintjét egyidejűleg, párhuzamos módon meghatározzuk.

1.5. A szekvenálástól a genom korszakig és tovább

A DNS bázissorrendjének meghatározására elsőként **Frederick Sanger** (1918-) dolgozott ki általánosan használható eljárást **1975**-ben. A DNS-t felépítő nukleotidokat szelektív módon adta különböző reakcióelegyekhez, melyekben

a szekvenálandó DNS-sel komplementer szál szintézisét DNS-polimeráz enzim végezte. A szintézis leállt ott, ahová az enzimnek a hiányzó nukleotidot be kellett volna építenie (ez volt az ún. „plusz-mínusz” módszer). **Walter Gilbert** (1932-) és **Allan Maxam** (1942-) **1977-ben** kidolgozott egy alternatív eljárást, amelyben a DNS szálakat különböző reagensekkel a négyféle nukleotid után specifikusan elhasította (**kémiai szekvenálás**). A létrejövő termékek elválasztása poliakrilamid gélen, a fragmentumok kimutatása pedig radioaktivitás alapján történt.

Ugyanebben az évben **Sanger** is kidolgozott egy új eljárást, melyben ismét a DNS-polimeráz enzimet használta fel a DNS szekvenálására. A DNS szintézis *in vitro* zajlik, ha van a reakcióelegyben DNS-polimeráz, templát DNS, a templáthoz hibridizált oligonukleotid („primer”; a 3'-OH vég szükséges a DNS szintézishez!), és dNTP (a négyféle dezoxinukleotid: dATP, dCTP, dGTP, dTTP). A DNS-polimeráz enzim képes dNTP analógokat is szubsztrátként elfogadni. Így például a 2',3'-didezoxinukleoizid-trifoszfát (ddNTP) analógok is beépülnek, de mivel ezeknek nincs 3'-OH csoportja, a további szintézis leáll. A módszer ezért a **láncterminációs szekvenálás** elnevezést kapta (hívjuk még **enzimatikus-, didezoxi- vagy Sanger-módszernek** is). Sanger négy különálló reakcióelegyet állított össze, ugyanazzal a templát DNS-sel. Mindegyik reakcióelegy tartalmazta a négyféle dezoxi-nukleotidot, viszont a didezoxi-nukleotidok közül csak egyet-egyet. Az újonnan szintetizálódott szálakat gélelektroforézissel szeparálta és az ún. szekvenáló létrát autoradiogramon tudta leolvasni. Sanger az általa kidolgozott módszerrel elsőként határozta meg egy természetben előforduló DNS molekulának, a **ΦX174 bakteriofág teljes genomjának** (5375 bp) szekvenciáját. Ezzel elkezdődött a **genomszekvenálások korszaka**. Sanger és Gilbert 1980-ban a DNS szekvenálás módszereinek kidolgozásáért kémiai Nobel-díjat kapott. Sanger esetén ez már a második Nobel-díj volt; az első a fehérjeszekvenálásért kapta 1958-ban.



1.10. ábra: Frederick Sanger

Kezdetben csak kisebb **genomok szekvenálása** történt meg: **1983-ban** a **λ-fágé** (48 kbp), majd **1995-ben** a ***Haemophilus influenzae*** (1,9 Mbp) baktériumé, ami az első teljes prokarióta genom megismerését jelentette. Az ***E. coli*** genom szekvenciáját **1997-re** ismertük meg (4,6 Mbp, ~4.300 gén). Az eukarióták köréből az élesztő (***Saccharomyces cerevisiae***) volt az első, melynek a teljes genom szekvenciája **1996-ra** készült el (12 Mbp, 5885 gén). A ***Caenorhabditis elegans*** genomjának szekvenálása **1998-ban** fejeződött be (97 Mbp, 19100 gén). A **humán genom** (3×10^9 bp, ~23 ezer gén) szekvenálása (Humán Genom Program) **2004-ben** fejeződött be (mindössze 14 év munkával), bár az első „vázlatos” szekvenciát már 2000-ben publikálták. Sanger módszerének továbbfejlesztésével (automata fluoreszcens szekvenálás), majd a közelmúltban ún. **új-generációs szekvenálási módszerek** megjelenésével, mint pl. a **piroszekvenálás (2004)** tovább gyorsult a genom szekvenálások üteme (**ld. 5.3. fejezet**). A robbanásszerű fejlődés eredményeképpen több mint 4000 teljes genom szekvenálása készült el 2013-ra, melyek között majdnem 200 eukarióta genomot is találunk.

1.5.1. A polimeráz láncreakció felfedezése

Az ***in vitro* DNS sokszorosítás** lehetőségét a polimeráz láncreakció felfedezése teremtette meg **1983-ban**. A felfedezés elméleti háttere már régen ismert volt, hiszen Arthur Kornberg 1957-ben felfedezte a DNS biológiai szintézisének mechanizmusát. Munkája során ***E. coli*** baktériumból izolált DNS-polimeráz I enzimmel dolgozott, amely segítségével új DNS szálakat tudott szintetizálni. A tanulmány megmutatta, hogy a folyamathoz kell egy templátként szolgáló DNS lánc, továbbá, hogy a DNS-polimeráz I enzim működéséhez szükség van Mg^{2+} ionokra, valamint mind a négy fajta DNS építőköre, aktivált monomerekre (dNTP-k). Már ma tudjuk, hogy a DNS-polimeráz enzim önmagában nem tudja elkezdni az új lánc szintézisét, ehhez egy oligonukleotid primerre van szükség (amely

a sejtekben egy rövid RNS darab lesz, amit a primáz nevű RNS-polimeráz állít elő), s ezt a primert képes a DNS-polimeráz 5'→3' irányban meghosszabbítani.

Ezen háttér információk ismeretében a PCR eljárást **Kary Mullis** dolgozta ki. Mullis érdeme abban rejlik, hogy ciklikussá tette az *in vitro* DNS polimerizációt és egyszerre két oligomer primert használt, amelyekről egymás felé történik az új DNS szálak szintézise. Meg kell említeni azt is, hogy az ötlet már Kjell Kleppe fejében is megfordult, azonban azt csak elméletben fogalmazta meg, így minden szabadalom Mullis, illetve az őt alkalmazó Cetus Corporation biotechnológiai cégnél kötött ki. Kezdetben Mullisnak mindössze egy 25 bázispár hosszú szakaszt sikerült szintetizálnia. Először az *E.coli*-ből származó DNS-polimeráz Klenow-frAGMENTUMÁVAL dolgozott, később azonban már a hőtűrő *Thermus aquaticus* baktérium DNS-polimerázát (Taq-polimeráz) használták. Ennek következményeként már nem kellett a magas hőmérséklet miatt minden ciklus után új adag polimeráz enzimet adni a reakcióhoz, és így 1988-ban már 2 ezer bázispár hosszú DNS szálakat tudtak szintetizálni.

A módszer a 90-es évek végére már széles körben elterjedt és számos tudományterületen tett lehetővé áttörő sikereket (pl.: archeológia, igazságügyi orvostan; ld. 6.4. fejezet).

1.5.2. Géntechnológiai eljárások magasabbrendű szervezetekben

Az **első transzgenikus élőlényt** Rudolf Jaenisch (1942-) és Beatrice Mintz (1921-) hozta létre 1974-ben. Jaenisch korai egér embriókba (blasztocisztákba) injektált SV40 vírus DNS-t. (Az SV40 egy majomvírus, egérben nem képes szaporodni.) Az embriókat nőstény egerek méhében ültette, ahol azok egészséges utódokká fejlődtek.

Ralph **Brinster** és Richard **Palmiter** 1982-ben úttörő munkát végeztek a transzgenikus élőlények előállításának területén. Megtermékenyített egér petesejtekbe injektáltak plazmid DNS-t, mely idegen gént hordozott egy indukálható promóter (az egér metallothionein gén promótere, amely nehézfémek jelenlétében aktiválódik) szabályozása alatt. Az ilyen módon kezelt petesejteket nőstény egerek méhében ültették. Az így létrehozott egerek szöveteiben az idegen DNS azonos mennyiségben volt jelen (az egér nem volt mozaikos). A bevitt DNS konstrukció integrálódott a genomba, sőt az idegen génről fehérjeszintézis történt. A módszer tehát lehetőséget teremtett arra, hogy bármilyen gént bejuttassanak az állatok genomjába úgy, hogy az minden szövetben és sejtben jelen legyen. Az így módosított organizmusban a gént szabályozó régiók megfelelő megválasztásával lehetséges indukált, vagy konstitutív fehérje expressziót létrehozni, akár szövet- vagy sejtspecifikus módon.

A transzgén technológiával párhuzamosan, szintén az **1980 években** jelent meg az úgynevezett „*knock-out*”, azaz génkiütés technológia. A módszer kifejlesztésében kulcsszerepet játszó Mario **Capecchi** (1937-), Oliver **Smithies** (1925-) és Martin **Evans** (1941-) 2007-ben fiziológiai és orvostudományi Nobel-díjat kapott. Evans elsőként izolált embrionális őssejteket korai egér embriókból és azokat blasztocisztákba injektálva sikeresen hozott létre **kiméra állatokat** (olyan állatok, melyek bizonyos sejtjei, szövetei az őssejtekből fejlődnek, így genetikai értelemben mozaikosak). Capecchi és Smithies egymástól függetlenül bizonyították, hogy az emlős sejtekben lejátszódó **homológ rekombináció** felhasználható arra, hogy adott géneket **célzottan módosítsanak** a genomban. Később az Evans által kifejlesztett módszert használva, embrionális őssejtek **genomját** sikerült célzottan módosítaniuk. Capecchi később kidolgozott egy szelekción alapuló módszert, melynek segítségével tesztelhetővé vált a homológ rekombináció sikeressége bármely megcélzott DNS szakasz esetén. A módosítás nem feltétlenül jelenti a megcélzott gén inaktiválását (*knock-out*); lehetséges a megcélzott helyre egy teljes, funkcióképes új gén beépítése is (*knock-in*). A technológia forradalmasította a biológia számos területét, hiszen lehetőség nyílt bármely gén funkcióját közvetlenül, a teljes élőlényben vizsgálni (ld. 14. fejezet). Talán a legjobb példa erre, hogy ma már számos genetikai eredetű emberi betegség kutatásához elengedhetetlen modellrendszer a megfelelő **KO (knock-out) egér**.

Egy másik fontos módszer, mellyel gének működésének célzott módosítására van lehetőség, az úgynevezett **RNS interferencia**. A más néven gélcsendesítésnek nevezett módszer felfedezéséért 2006-ban Andrew **Fire** (1959-) és Craig **Mello** (1960-) kapott Nobel-díjat. A két kutató modellorganizmusként a *Caenorhabditis elegans* nevű fonálférget használta. A kutatók kimutatták, hogy az állat testébe injektált kétszálú RNS az adott gén funkcióvesztéses mutációjához hasonló fenotípust eredményezett. (Az egyszálú minták esetén alig volt kimutatható hatás.) A kétszálú RNS tehát „interferált” a gén működésével. Fire és Mello más géneket „megcélözva” is megismételte a kísérletet, így kimutatták, hogy a jelenség általános, továbbá specifikus a megcélzott génre. Mivel az RNS interferenciához igen kis mennyiségű kétszálú RNS is elég volt, feltételezték, hogy valamilyen katalitikus komponens is szerepet játszik a jelenség mögötti molekuláris mechanizmusban. Ma már bizonyított, hogy az RNS interferencia a legtöbb

eukariótában működő jelenség, mely mögött a sejtekben szintetizálódó mRNS molekulákat lebontó enzimátikus mechanizmus áll. Ez az enzimrendszer számos fontos funkcióval rendelkezik, így például védelmet nyújt a vírusokkal, vagy a transzpozonokkal (mobil genetikai elemek) szemben, de elengedhetetlen számos gén működésének szabályozásához is. Napjainkra az RNS interferencia a génműködés tanulmányozásának fontos eszközévé vált, sőt várható, hogy a jövőben a humán gyógyászatban is szerepet kap (**ld. 15. fejezet**).

1.6. További olvasnivaló a fejezethez

Avery, OT, MacLeod, CM, and McCarty, M (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types. *Journal of Experimental Medicine* 79: 137-158.

Watson JD, Crick FH (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171 (4356): 737–738.)

Hershey A, Chase M (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* 36 (1): 39-56.

Nirenberg MW, Matthaei JH. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc Natl Acad Sci USA*. 47:1588-602

Cohen, SN, Chang, ACY, Boyer, HW & Helling, RB (1973) Construction of biologically functional bacterial plasmids in vitro. *Proc. Nat. Acad. Sci. USA* 70, 3240–3244

Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74: 5463–7

Palmiter RD, Brinster RL, Hammer RE, Trumbauer ME, Rosenfeld MG, Birnberg NC, Evans RM (1982) Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature*. 300(5893):611-5.

Saiki, R.; Gelfand, D.; Stoffel, S.; Scharf, S.; Higuchi, R.; Horn, G.; Mullis, K.; Erlich, H. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239 (4839): 487–491

Fire A, Xu SQ, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806-811.

2. fejezet - A géntechnológia alapja: molekuláris klónozás

A **rekombináns DNS technológia** („génésebészet”, *genetic engineering*) eszköztára lehetővé teszi az élőlények egyes tulajdonságait meghatározó gének azonosítását, jellemzését és szabadon történő megváltoztatását. Megismerhetjük akár egy adott élőlény teljes genetikai állományának (genom) információtartalmát is, hiszen napjainkra a **DNS szekvenálása** rutinszerű feladattá vált. Feltárhatjuk a gének működését szabályozó régiókat. Célzottan, **irányított mutációkat** hozhatunk létre, sőt, akár teljes géneket is eltávolíthatunk az adott élőlényből, illetve módosított, mutáns géneket vihetünk be kísérleti élőlényekbe (**reverz**, azaz fordított irányú **genetika**). Ilyen módon megérthetjük az egyes gének, illetve az általuk kódolt fehérjék működését. A genetikai kód univerzális mivoltát kihasználva lehetséges géneket „átültetni” más organizmusokba, ahol megtörténhet a bevitt szekvencia transzkripciója és translációja. Az ilyen organizmusokat felhasználhatjuk arra, hogy a kódolt fehérjét nagy mennyiségben és tisztaságban állítsuk elő (**heterológ fehérje expresszió**). Ez a lehetőség az alaputatások mellett (a termelő fehérje működésének és szerkezetének vizsgálata) a biotechnológiai ipar illetve a gyógyszeripar szempontjából is kulcsfontosságú. Az expresszált fehérjét kódoló szekvenciát módosítva számunkra hasznos változatokat hozhatunk létre az adott fehérjéből. Akár az eredetitől eltérő funkciót, vagy aktivitást is „hozzárendelhetünk” ilyen módon egy fehérjéhez (**fehérjemérnökség** vagy **protein engineering**). Arra is lehetőségünk nyílik, hogy saját sejtjeinkbe juttassunk be új genetikai információt bizonyos betegségek gyógyítása érdekében (**génterápia**).

Ahhoz, hogy a géntechnológiát ténylegesen alkalmazni lehessen, rekombináns DNS molekulákat kell előállítanunk. A **rekombináns DNS** egy **hordozó DNS molekulából (vektor)** és a vizsgálni kívánt **DNS-ből (inszert)** álló „hibrid” (**kiméra**) molekula, amelyet a kísérletekhez szükséges mennyiségben molekuláris klónozással állítunk elő.

rekombináns DNS = vektor DNS + inszert DNS

A molekuláris klónozás nem más, mint egy specifikus inszertet tartalmazó rekombináns DNS felszaporítása megfelelő gazdasejtben (*E. coli*). Ezeknek a sejteknek a klónjaiból (bakteriális kolóniák) izolálhatjuk a molekuláris klón rekombináns DNS-t. Létezik *in vitro* DNS szaporítási módszer is, a polimeráz láncreakció (PCR), azonban ezzel a módszerrel szigorúan véve nem molekuláris klónokat állítunk elő (ld. 6. fejezet).

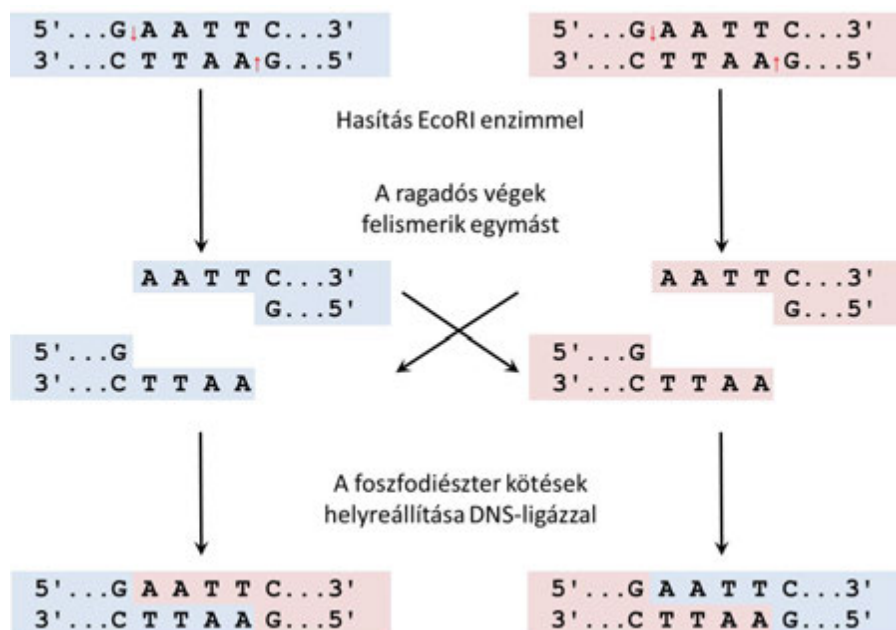
2.1. *In vitro* rekombináció

A történeti összefoglalóból már kiderült, hogy a molekuláris klónozáshoz egyrészt fel kellett fedezni azokat az enzimeket, amelyekkel a DNS molekulát *in vitro* „szabni-varrni” lehet, másrészt találni kellett olyan DNS molekulákat, amelyeket vektorrá lehet alakítani (a vektorokkal szemben támasztott kritériumokról és a típusaikról részletesen a 7.1. fejezetben lesz szó). Végül az *in vitro* előállított rekombináns DNS-t egy gazdasejtbe kell juttatnunk (ld. 4.3. fejezet) és valamilyen úton-módon ki kell szelektálni azokat a sejteket, amelyekbe az általunk megtervezett rekombináns DNS került be.

A DNS „manipuláció”, a molekuláris klónozás *in vitro* szakaszának legfőbb eszközei a restrikciós endonukleázok és a DNS-ligáz enzim (a géntechnológiában használatos DNS-módosító enzimekkel részletesen a 3. fejezet foglalkozik).

A **restrikciós endonukleázok** (röviden **restrikciós enzimek**) foszfodiészter kötést hidrolizáló enzimek. A baktériumokat az idegen DNS-ekkel szemben „védelmező” molekuláris rendszer (ld. 1.2. fejezet), a restrikció-modifikáció DNS bontó elemei, amelyek nagy számban és sokféleségben állnak a géntechnológus rendelkezésére. Az enzimek többsége középpontosan szimmetrikus, ún. „**palindrom**” **szekvenciákat ismer fel**. Amennyiben a DNS két szálának elhasítása egymással szemközi foszfodiészter-kötéseknél történik, úgynevezett **tompá végek** (*blunt end*) keletkeznek (pl. *DraI*, ld. 2.2. ábra). Ha azonban egymáshoz képest elcsúsztatva, néhány bázissal távolabb a felismerőhely szimmetriatengelyétől, úgy **ragadós végek** (*sticky end*) jönnek létre (pl. *EcoRI*, ld. 2.1. ábra). A **DNS-ligáz** enzim foszfodiészter kötés létrehozásával képes két DNS molekulát, vagy egy molekula két végét **kovalensen összekapcsolni**. Ha két különböző eredetű DNS molekulát *EcoRI* enzimmel hasítunk, majd

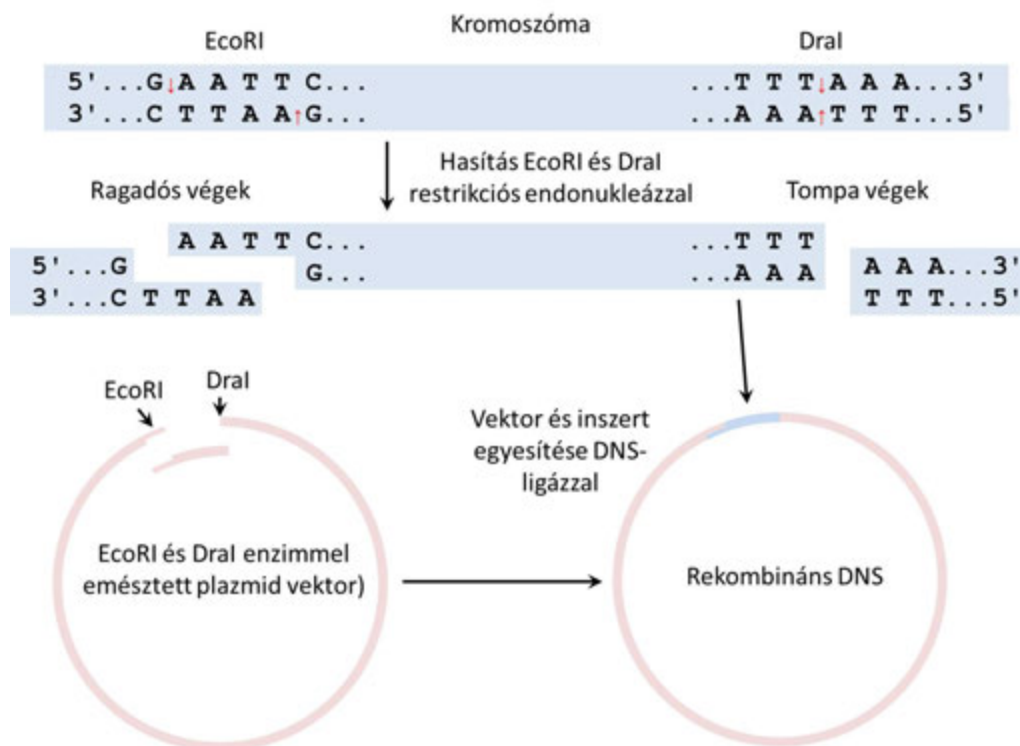
összekeverünk, akkor a **DNS-ligáz *in vitro* rekombináció**val létrehozhat hibrid, azaz rekombináns DNS molekulákat, mint azt a **2.1. ábra** szemlélteti. A restrikciós endonukleázokkal részletesen a **3.2.1. fejezet** foglalkozik.



2.1. ábra: ***In vitro* rekombináció.** Két különböző eredetű DNS hasítása EcoRI restrikciós endonukleázzal. A létrejövő DNS fragmentumok egymást felismerő ragadós végekkel rendelkeznek. Összekapcsolásuk DNS-ligáz segítségével történik. Amennyiben a két különböző eredetű DNS-vég kapcsolódik össze, *in vitro* rekombináció történik. Megjegyzendő, hogy egy DNS molekula két végén található ragadós végeket is összekapcsolhatja a DNS-ligáz. Ebben az esetben nem jön létre rekombináció, hanem pl. ha egy linearizált vektor DNS-ről van szó, az cirkularizálódik. Az ún. vektor „önzáródás” kiküszöbölhető, ha a hasított DNS 5'-végéről a foszfát-csoportokat foszfátáz enzimmal eltávolítjuk (**ld. 3.4.2. fejezet**)

2.2. Rekombináns DNS létrehozása

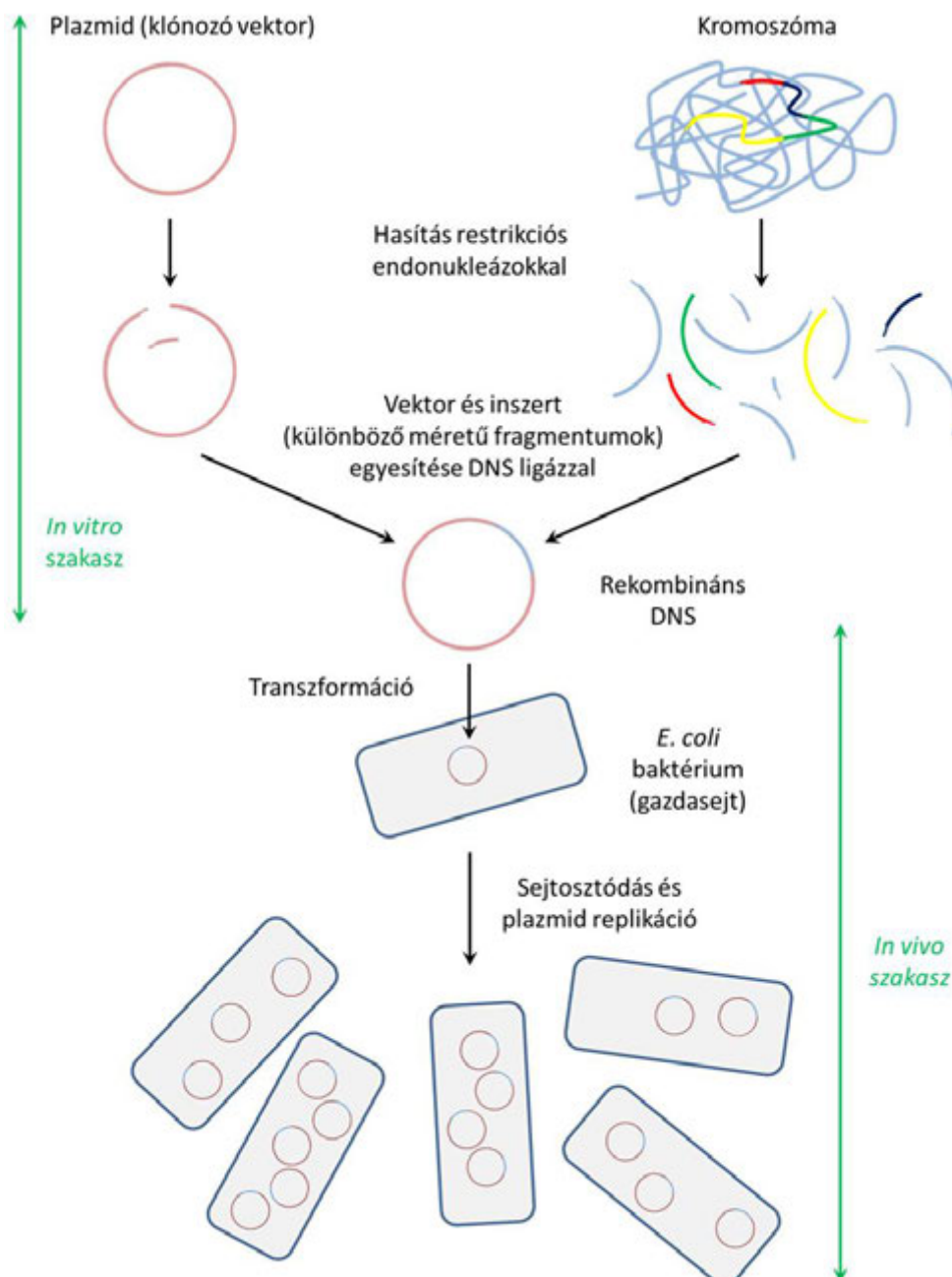
A molekuláris klónozás *in vitro* szakaszában alakítjuk ki a rekombináns DNS-t. A hordozó DNS leggyakrabban egy cirkuláris plazmid vektor (a plazmidok extrakromoszomális elemek, a baktériumsejtekben önállóan replikálódó ún. replikonok, **ld. 7.2. fejezet**), míg a vizsgálandó DNS lehet egy eukarióta kromoszóma. Példaként egy olyan klónozási stratégiát mutatunk be, ahol az inszertet a kromoszómából két különböző restrikciós enzimmal hasítjuk ki. Ha a vektor DNS-t ugyanazon két restrikciós endonukleázzal hasítjuk és a tisztított hasítási termékeket összekeverjük, a DNS-ligáz a kémcsőben lezajló reakció során gyakorlatilag csak ezt a kétféle eredetű DNS-t képes egymáshoz kapcsolni, ezáltal **rekombináns DNS** jön létre, mint azt a **2.2. ábra** mutatja. Ez a **direkcionális vagy irányított klónozás**, amit részletesen a **4.2. fejezet**ben ismertetünk.



2.2. ábra: **Rekombináns DNS létrehozása irányított klónozással.** A plazmid vektort és a kromoszomális DNS-t ugyanazzal a kétféle restriktív enzimmel hasítjuk. Az egyik végén "ragadós" (EcoRI), a másik végén "tompá" végű (DraI) DNS molekulákat a DNS-ligáz kovalens összekapcsolja, ezzel létrejön a rekombináns DNS.

2.3. Molekuláris klónok kialakítása

A DNS klónozás *in vivo* szakasza a rekombináns DNS gazdasejtbe juttatása és felszaporítása. A DNS sejtekbe juttatására többek között a **bakteriális transzformáció** jelenségét lehet kihasználni (további génbeviteli eljárások leírása a **4.3. fejezetben** olvasható). Egy baktérium sejt általában csak egyetlen rekombináns DNS-t vesz fel. Mivel a transzformáció hatásfoka alacsony, szükségünk van valamilyen **szelekciós** eljárásra. A vektor tartalmaz egy antibiotikum elleni védelmet biztosító gént (**antibiotikum rezisztencia gén**), amely megfelelően választott körülmények között szelektív túlélést tesz lehetővé a rekombináns DNS-t hordozó sejtek számára. A transzformálás után a sejteket olyan, antibiotikumot tartalmazó, szelektáló táptalajra helyezzük, ahol csak a rekombináns plazmidot (és így az adott antibiotikumra rezisztenciát biztosító gént) tartalmazó sejtek képesek szaporodni. Minden sejtből egy **kolónia** keletkezik (amely genetikailag azonos sejtekből, **klónokból** áll). Sok ezer ilyen, a kromoszóma különböző darabját tartalmazó kolónia együttesen egy genomális könyvtárat reprezentál (**ld. 8.2. fejezet**). Egy-egy kolónián belül azonos rekombináns DNS-t tartalmazó sejtek vannak, melyeket tovább lehet szaporítani, és a klónozott DNS-t izolálni lehet. A folyamat sémáját a **2.3. ábra** mutatja be.



2.3. ábra: **A molekuláris klónozás elvi vázlata.** A rekombináns DNS-t transzformációval juttatjuk az *E. coli* gazdasejtbe, ahol az a baktériumsejt kromoszómájától (amit nem ábrázoltunk) függetlenül replikálódik és több kópiában lesz jelen az utódsejtekben, ezzel is hozzájárulva a DNS klón felszaporításához (amplifikáció). A különböző inszerteket tartalmazó baktériumok egyedi klónok. A kromoszóma egészét darabokban tartalmazó *E. coli* klónok összessége a genomális könyvtár (ld. 8.2. fejezet). A rekombináns DNS a sejtekből tiszta formában egyszerűen kinyerhető (ld. 4.1.1. fejezet).

2.4. További olvasnivaló a fejezethez

Watson, JD et al. (2007) Recombinant DNA: Genes and Genomes. A Short Course. W. H. Freeman. ISBN-10: 0716728664

Brown TA (2010): Gene Cloning and DNA Analysis: An Introduction. Wiley-Blackwell, 6th Edition

3. fejezet - DNS módosító enzimek és felhasználásuk

A manapság rutinszerűen alkalmazott biotechnológiai és molekuláris biológiai eljárások jelentős része a DNS módosítását célozza. Ezeknek a módszereknek a kialakítása az 1970-es években kezdődött, amikor a kutatók elkezdték a DNS manipulálására szolgáló enzimeket tisztított formában előállítani és vizsgálni. Később ezeket az enzimeket, amelyek a természetben a **rekombináció, replikáció és hibajavítás folyamataiban vesznek részt**, sikeresen használták a DNS molekuláról történő másolatok készítésére, a DNS szétvágására vagy éppen összeragasztására. A módszerek fejlesztése lehetővé tette, hogy megszülessen a rekombináns DNS technológia, vagyis különböző DNS szakaszok „újrakombinálására”, összeépítésére nyílt lehetőség plazmidok és más eredetű DNS fragmentumok (pl. kromoszómák darabjainak) felhasználásával. Mindez megteremtette a DNS klónozásának lehetőségét is (**ld. 2. és 4.2. fejezet**). Ehhez elegendő egy gént vagy más DNS fragmentumot plazmidba vagy egyéb hordozóba (vektor) építeni, majd megfelelő gazdaszervezetbe juttatni, és ott a rekombináns DNS replikálódni képes.

A DNS módosítására szolgáló enzimeknek négy nagy csoportját különböztethetjük meg.

- **DNS polimerázok:**

Funkciójuk egy, a templát DNS-sel vagy RNS-sel komplementer új polinukleotid szál szintézise.

- **Nukleázok:**

Képesek a nukleotidokat összekötő foszfodiészter kötések elhasítani (hidrolizálni), és így a DNS-t (le)bontani.

- **Ligázok:**

Foszfodiészter kötés létrehozásával képesek két DNS molekulát, vagy egy molekula két végét összekapcsolni.

- **Láncvég módosító enzimek:**

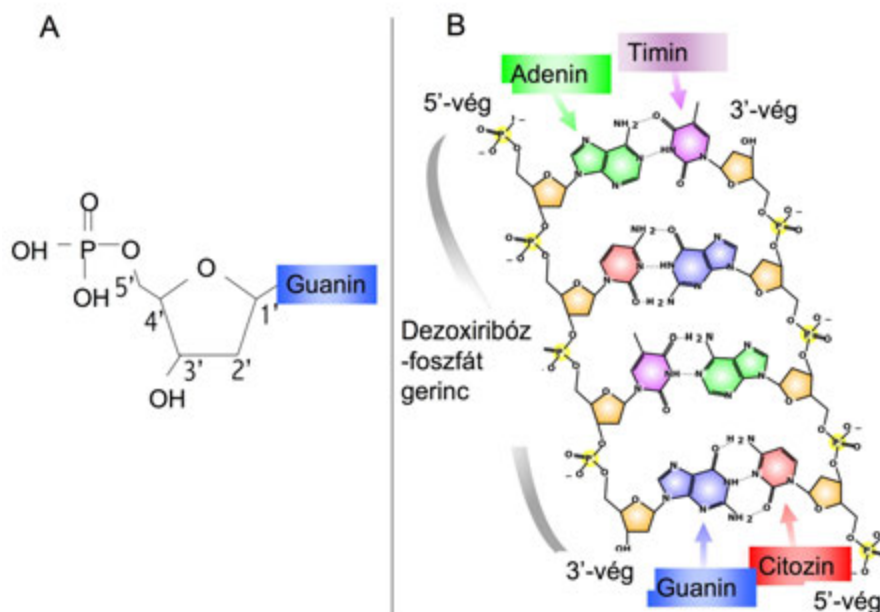
A DNS lánc végeinek módosításával egyes klónozási technikákban van szerepük.

3.1. DNS-polimerázok

Számtalan molekuláris biológiai módszerhez elengedhetetlen a DNS-polimerázok alkalmazása, gondoljunk akár csak a PCR-re (*Polymerase Chain Reaction*; **ld. 6. fejezet**) vagy a DNS szekvenálásra (**ld. 5. fejezet**). Általánosságban **DNS-polimeráznak** nevezzük azokat az enzimeket, amelyek képesek **DNS szintézisére dezoxiribonukleozid-trifoszfát (dNTP) építőkövek** (aktivált monomerek) **felhasználásával**. Ezen belül megkülönböztethetjük a **templát-függő DNS-polimerázokat**, amelyek egy „minta”, templát DNS vagy RNS alapján, annak másolatát készítik el, valamint léteznek **templát-független DNS-polimerázok** is (ilyen enzim pl. a terminális dezoxinukleotid-transzferáz).

A DNS-polimerázok multifunkciós enzimek, amelyek a szintézis mellett képesek a DNS láncvégeinek hidrolízisére is. Míg a **szintézis** szigorúan **egy irányban, az 5'-végtől a 3'-vég irányában zajlik**, a lebontás a lánc mindkét vége felől megvalósulhat, az adott polimeráz aktivitásától függően. Az egyszerűbb tárgyalhatóság érdekében érdemes megismerkedni a DNS molekula kapcsán a direkcionalitás fogalmával.

A molekuláris biológiában **direkcionalitásnak**, vagy egyszerűen **iránynak** nevezzük az egyszálú nukleinsavak hosszanti kémiai orientációját (a DNS, RNS, a fehérjeláncok **direkcionális polimerek**). A DNS szál végeinek elnevezése a nukleotidok cukorgyűrűjének konvencionális számozásából ered (**ld. 3.1. ábra**), ily módon egyértelműen meghatározható a szál **5'**- (ejtsd: öt-vessző) és **3'**- (három-vessző) vége. Amennyiben egy hosszú DNS szakasz régióit szeretnénk megkülönböztetni, beszélhetünk „**upstream**” (felfelé irányuló) szakaszokról, amelyek az 5'-véghöz esnek közelebb, illetve „**downstream**” (lefelé irányuló) szakaszokról a 3'-véghöz közelebb. A templát-függő DNS-polimerázok azon tulajdonsága, hogy új foszfodiészter kötés kialakítására csak a 3'-hidroxil csoporttal képesek, meghatározza a szintézis 5'→3' irányát.

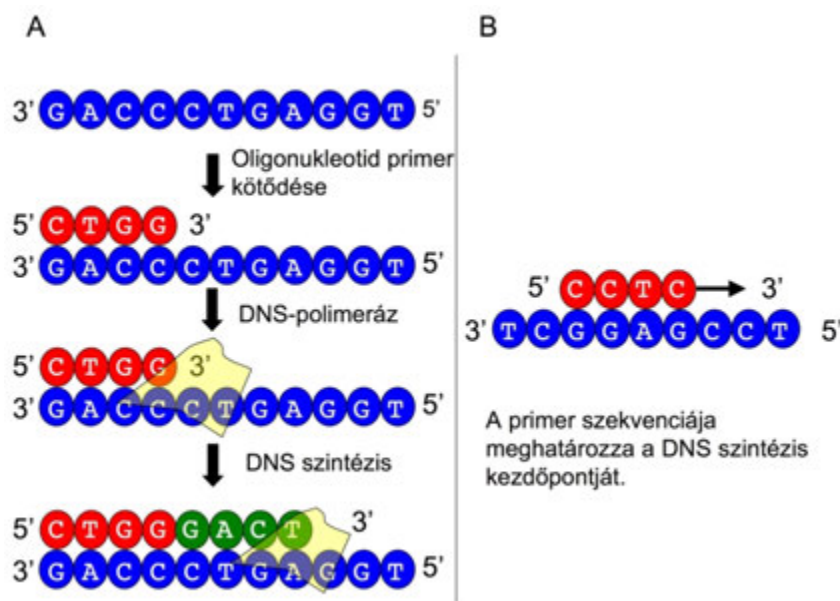


3.1. ábra: **A DNS molekula irányultsága, 5'- és 3'-vége.** A) A guanin nukleotid a dezoxiribóz gyűrű számozásával, aminek alapján meghatározható a DNS lánc 5'- és 3'-vége. B) A DNS molekula kémiai szerkezete, külön jelölve a DNS felépítésében résztvevő dezoxiribonukleotidokat (részlet).

A lánc bontása kapcsán kétféle aktivitást különböztethetünk meg, amelyekkel külön-külön vagy együtt szintén mindegyik DNS-polimeráz rendelkezik.

- **3'→5'-exonukleáz aktivitás:** Az irány megjelölés arra utal, hogy az enzim a DNS szálon visszafelé haladva képes a már szintetizált nukleotidot eltávolítani. A polimeráz ezen tulajdonságát „*proofreading*” vagy hibajavító aktivitásnak is nevezzük, mert ez biztosítja, hogy egy esetlegesen rosszul, nem a bázispárosodás szabályainak megfelelő nukleotid beépülés javítása megtörténjen.
- **5'→3'-exonukleáz aktivitás:** Kevesebb polimeráz rendelkezik ezzel az aktivitással, aminek során az enzim képes a láncon előre haladva az esetleg már ott található, a templáttal komplementer DNS szálat elbontani (ezt a „hasadás-elmozdulás” kettős aktivitást nevezi a szakirodalom „*nick* transláció”-nak, ld. 3.3. ábra). A természetben a DNS replikációja során van szükség erre a funkcióra.

A templát függő DNS-polimerázok működésének fontos jellemzője, hogy *de novo* szintézisre nem képesek. Új **DNS szál szintéziséhez szükséges egy rövid kétszálú szakasz**, és így az új szál szintézise egy már meglévő 3'-végről indulhat. A természetben az új szál szintézise mindig egy rövid RNS szakasszal, a **primerrel** indul. A biotechnológiai gyakorlatban **primerként** általában egy rövid **szintetikus** oligo-dezoxiribonukleotid (röviden **oligonukleotid**) molekulát használnak, ami a templát DNS-hez kapcsolódva biztosítja a szükséges kétszálú szakaszt (ld. 3.2. ábra). Ennek nagyon fontos gyakorlati jelentősége, hogy így a DNS polimerizáció kezdőpontja beállítható, vagyis bármilyen, hosszabb DNS szál szintetizálható a megfelelő helyre tervezett primer segítségével. (Mindez természetesen nem lenne lehetséges, ha a DNS polimerizáció véletlenszerű helyről indulna.)



3.2. ábra: **A primerek jelentősége.** A) A templát-függő DNS-polimerázok működéséhez kétszálú DNS-re van szükség, ami a templáttal komplementer rövid oligonukleotid, primer hozzáadásával biztosítható. B) A primer a bázissorrendjének megfelelő komplementer szakaszhoz hibridizálódik, ezzel meghatározva a DNS szintézis kezdőpontját.

3.1.1. A DNS-polimeráz I és a Klenow-fragmentum

A legkorábban felfedezett DNS-polimeráz, amelyet eredetileg *Escherichia coli* baktériumból izolált Arthur Kornberg, de minden prokariótában megtalálható. Fontos szerepe van a bakteriális genom replikációjában, eltávolítja az RNS primert a követő szárról és az Okazaki-fragmentumok összekötését végzi (ld. A biokémia és molekuláris biológia alapjai elméleti e-jegyzet).

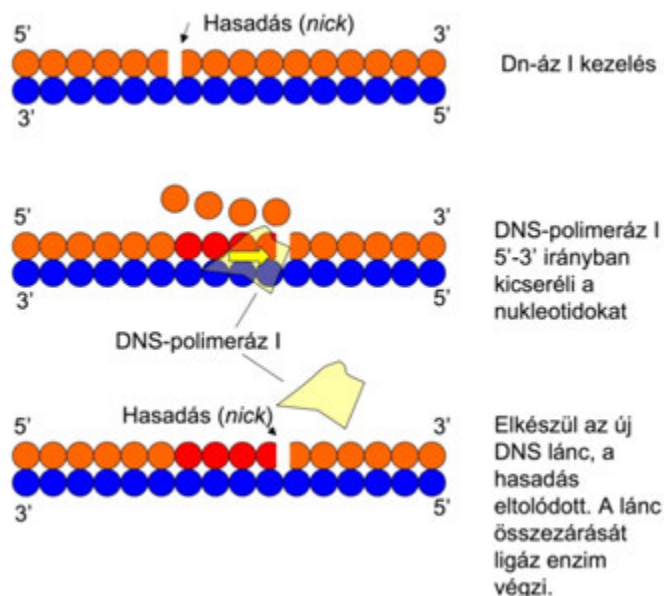
Négyféle enzimátikus aktivitással rendelkezik:

- 5'→3' DNS-függő DNS-polimeráz aktivitás
- 3'→5' exonukleáz aktivitás (*proofreading* aktivitás)
- 5'→3' exonukleáz aktivitás (RNS primer eltávolítása és kivágásos hibajavítás)
- 5'→3' RNS-függő DNS-polimeráz aktivitás (valószínűleg nincs biológiai jelentősége)

A molekuláris biológiai gyakorlatban a natív enzim hasadás-elmozdulás, angolul „*nick-translation*” képességét használják ki (ld. 3.3. ábra). A természetben az RNS primerek eltávolítására és DNS-re cserélésre szolgál a folyamat. A laboratóriumban a kétszálú DNS-en mesterségesen, DN-áz kezeléssel „nick”-eket, bevágásokat ejtenek, majd a DNS-polimeráz I 5'→3' exonukleáz és 5'→3' polimeráz aktivitását kihasználva, a hasadástól indulva a DNS egy szakaszát kicserélik. Ennek során lehetőség van például radioaktívan vagy fluoreszcensen jelölt nukleotid analógok DNS láncba építésére.

Míg a fenti példában az enzim 5'→3' exonukleáz aktivitását lehet kihasználni, azoknál az alkalmazásoknál, ahol kizárólag a polimeráz aktivitásra lenne szükség (mint például a szekvenálás), nem szerencsés, hogy az enzim képes az újonnan szintetizált DNS láncot lebontani. Az enzim további vizsgálatok kiderült, hogy proteolitikus hasítása két fragmentumot (76 kDa és 36 kDa) eredményez, amelyek közül az egyik, a felfedezőjéről **Klenow-fragmentum**nak nevezett darab rendelkezik mind a polimeráz, mind a *proofreading* aktivitással, de hiányzik belőle az 5'→3' exonukleáz aktivitás. A Klenow-fragmentumot manapság rekombináns módon, *E. coli* sejtekben termelik. A Klenow-fragmentum felhasználható a Sanger-féle didezoxi szekvenálásban (ld. 5. fejezet), valamint restrikciós

endonukleázok ragadós végének feltöltésére, illetve 3'-túlnyúló végek lebontására (tompá vagy „*blunt-end*” vég kialakítására).



3.3. ábra: **Nick transláció (hasadás-elmozdulás)**. A DNS-polimeráz I enzim az 5'→3' irányú polimeráz aktivitás mellett 5'→3' irányú exonukleáz aktivitással is rendelkezik, és ezek kombinációja teszi lehetővé a hasadás-elmozdulást. A laboratóriumi gyakorlatban jelölt nukleotidok beépítésére használható a jelenség. Elsőként a dezoxiribonukleáz I (DN-áz I) enzimmel kis bevágásokat (*nick*) ejtenek a DNS szálain. Ezután a DNS-polimeráz I a bevágástól indulva a sorban következő nukleotidokat kivágja, és a reakcióelegyhez adott jelölt nukleotidokat építi be. A reakció eredményeként kapott DNS szálon a *nick* elmozdul előre felé addig a pontig, ahol a DNS-polimeráz I leválik a DNS láncról. A lánc folytonossága ligálással állítható helyre.

3.1.2. T4 DNS-polimeráz

Kétféle enzimatis aktivitással rendelkezik:

- 5'→3' DNS-függő DNS-polimeráz aktivitás
- 3'→5' exonukleáz aktivitás (*proofreading* aktivitás)

A T4 jelű *E. coli* bakteriofág genomjából származik. Több, mint 200-szor erősebb 3'→5' exonukleáz aktivitással rendelkezik, mint a DNS-polimeráz I vagy a Klenow-fagmentum. Tompa vég kialakítására alkalmas, valamint a ligáz-független klónozásban is használják a vektor megfelelő ragadós végeinek kialakítására (ekkor dTTP hozzáadásával biztosítják, hogy a kívánt nukleotidig „vágja vissza” az enzim a DNS megfelelő szálát) (ld. 4.2.5. fejezet).

3.1.3. T7 DNS-polimeráz

A T7 DNS-polimeráz *in vivo* a T7 bakteriofág DNS-ének gyors replikációjáért felelős. Az enzim két fehérje, a T7 bakteriofág „5-ös gén” fehérjetermékének (80 kDa) és az *E. coli* tioredoxin fehérjéjének (12 kDa) komplexe.

Kétféle enzimatis aktivitással rendelkezik:

- 5'→3' DNS-függő DNS-polimeráz aktivitás
- 3'→5' exonukleáz aktivitás (*proofreading* aktivitás)

A 3'→5' exonukleáz aktivitása jelentős, körülbelül 1000-szerese a Klenow-fagmentuménak, ami egyben jelzi, hogy nagyon **nagy hűségű (high fidelity) enzim**, vagyis a DNS szintézis nagy pontossággal zajlik. Kiemelkedően

magas az enzim processzivitása is, ami annyit jelent, hogy hosszú ideig a DNS szálon marad, és így képes akár több ezer nukleotid szintézisére. Szekvenálásra ez az enzim, bár magas processzivitása fontos előny, eredeti formájában nem alkalmas, mert a szekvenálás során elkerülendő bármifajta exonukleáz aktivitás, nehogy az enzim kivágja a láncterminációs nukleotidot. Ezért erre a célra az enzim egy módosított, **Sequenase** néven forgalmazott formáját használják, amelyből teljesen hiányzik az exonukleáz aktivitás. A T7 DNS-polimerázt ezen kívül alkalmazzák helyspecifikus mutagenézis módszereknél is (ld. 11. fejezet).

3.1.4. Hőstabil DNS-polimerázok

Bár már viszonylag korán, a 1960-as és 1970-es években izoláltak hőstabil DNS-polimerázokat, igazi jelentőségük a polimeráz láncreakció (PCR) elterjedésével (ld. 6.2.2. fejezet) mutatkozott meg, amihez magas hőmérsékleten is aktív enzimre van szükség.

3.1.4.1. Taq-polimeráz

Eredetileg a Yellowstone Nemzeti Parkban található hőforrásokban élő *Thermus aquaticus* nevű termofil baktériumból izolálta Thomas D. Brock 1965-ben (a baktérium nevéből származik a polimeráz neve, ugyanazt a logikát követve, mint a restrikciós endonukleázok elnevezése; ld. 3.2.1.). Kiválóan használható PCR-hez, mert az enzim féléletideje még 95°C-on is 40 perc. Mivel nem rendelkezik 3'→5' exonukleáz aktivitással, **nem** tartozik a **nagy hűségű polimerázok** közé. A hibagyakorisága ~10⁻⁴, de a pontos érték erősen függ az adott kísérleti elrendezéstől. 1 kb méretű célszekvenciát ~30 sec alatt másol le. Más, 3'→5' exonukleáz aktivitás nélküli enzimekhez hasonlóan templát-független **dezoxinukleotid-transzferáz aktivitással rendelkezik**, aminek következtében (preferáltan) adenin nukleotidokat épít be a DNS 3'-végére. Ezt a tulajdonságát használják ki a TA klónozásnál (ld. 4.2.4. fejezet). A gyakorlatban manapság leginkább rutin PCR-hez, főként kolónia PCR-hez használják, amikor a cél a baktérium kolóniák közül azokat kiválasztani, amelyek egy adott klónt tartalmaznak. Ilyen esetben nem probléma egy-egy mutáció a PCR termékben.

3.1.4.2. Pfu-polimeráz

A hipertermofil *Pyrococcus furiosus* ősbaktériumból származó DNS-polimeráz. A Taq-polimeráznál jelentősen **nagyobb hőstabilitással** rendelkezik. Ezen kívül legfontosabb tulajdonsága, hogy **rendelkezik proofreading aktivitással**, ezért minden olyan alkalmazásra használható, ahol fontos a pontos másolás (átlagosan kb. 350-400 ezer bázispáronként épít be hibás nukleotidot). Valamivel lassabb, mint a Taq-polimeráz, 1-2 perc alatt amplifikál 1 kbp méretű DNS-t 72°C-on. Nagyon sok laboratóriumban a Pfu-t használják elsődleges DNS-polimerázként, mind PCR-hez, mind például helyspecifikus mutagenézishez.

3.1.4.3. Phusion-polimeráz

A Phusion-polimeráz egy **proofreading aktivitású** *Pyrococcus*-szerű polimeráz és egy DNS-kötő domén „fúziós” (innen a név) terméke. Egyesíti magában a Taq- és a Pfu-polimerázok előnyeit. Rendkívül **nagy másolási hűségű (fidelity)** enzim, több, mint 2 millió bázispáronként épít be hibás nukleotidot. Ezen kívül kifejezetten gyors is, ami jelentősen csökkenti a PCR reakció idejét, valamint a sikeres PCR-hez gyakorlatilag nem szükséges optimalizáció sem. Alkalmas mind rutin, mind komplikált, például hosszú vagy nehéz templátokkal való munkára is. Egyetlen hátránya, hogy az említett polimerázok közül ez a legdrágább.

Néhány további, a PCR-hez használt polimeráz enzimet az 6.2.2. fejezetben ismertetünk.

3.1.5. Reverz transzkriptáz

RNS-függő DNS-polimerázok, amelyek a retrovírusok replikációs ciklusában játszanak szerepet. A retrovírusok genomja RNS-ből áll, amiről DNS másolat készül a gazdaszervezet fertőzése után. A laboratóriumi gyakorlatban reverz transzkriptázok használhatók mRNS molekulákról DNS másolat készítésére. Az így szintetizált, a templát **mRNS-el komplementer szálát** nevezik **cDNS-nek**. A cDNS templátként szolgálhat további PCR reakcióhoz. Az mRNS-ből kiinduló, és végül amplifikált DNS terméket eredményező reakciót RT-PCR-nek, reverz transzkriptációs PCR-nek hívják (ld. 6.4.4. fejezet). (Az RT-PCR nem összekeverendő a valós idejű, *real-time* PCR-rel, aminek során a fluoreszcensen jelölt DNS amplifikációjának kvantitatív meghatározása történik.) A reverz transzkriptáz fontos felhasználási területe a cDNS könyvtárak készítése is (ld. 8.3. fejezet).

3.2. Nukleázok

Számos nukleázt használnak rutinszerűen a mindennapos molekuláris biológiai laboratóriumi gyakorlatban. A legtöbb nukleázt két csoportra lehet osztani:

- **Exonukleázok:** olyan enzimek, amelyek a DNS vagy RNS két végéről képesek nukleotidok eltávolítására.
- **Endonukleázok:** a nukleinsavak közbülső részein található foszfodiészter kötéseket tudják elhasítani.

3.2.1. Restriktációs endonukleázok

A restriktációs endonukleázok baktériumokban, archeákban, és egyes vírusokban megtalálható enzimek, és a **restriktációs-modifikációs rendszer komponensei (ld. 1.2. fejezet).**

A restriktációs endonukleázok (vagy röviden **restriktációs enzimek**) specifikus helyeket ismernek fel a DNS molekulán, majd a DNS mindkét szálát átvágva képesek azt elhasítani (foszfodiészter kötéseket hidrolizálnak). Elnevezésüknek külön **nomenklatúrája** van. A HindIII enzim nevének magyarázata, hogy a *Haemophilus influenzae* baktérium **d** szerotípusából harmadikként izolálták. Hasonlóképpen képezhető az EcoRI enzim neve is: az *Escherichia coli* **RY13** törzséből elsőként izolálták.

Egy random bázissorrendű DNS-t tekintve a statisztikai hasítási gyakoriság attól függ, hogy milyen hosszú az enzim felismerési szekvenciája. Ha az emésztőenzim felismerő helye 4 bázispárt tartalmaz, akkor, mivel minden egyes bázispár helyén négyféle variáció (A, T, G, C) lehetséges, $4^4 = 256$ bázispáronként számíthatunk hasítóhelyre. Ha a felismerő hely 6 nukleotid hosszúságú, a hasítás valószínűsége $1/4096$ -ra módosul ($4^6 = 4096$). A valóságban a DNS szekvenciája nem véletlenszerű, ezért a hasítóhelyek eloszlása sem teljesen követi a fentieket. Vannak restriktációs endonukleázok, amelyek felismerőhelye megegyezik, de hasítóhelyük különböző, ezeket **neoskizomereknek** nevezik. Azokat az enzimeket, amelyeknél mind a felismerő, mind a hasítóhely egyforma, **izoskizomereknek** hívják.

A restriktációs enzimek leggyakoribb felhasználási területe a molekuláris klónozás (**ld. 2. és 4.2. fejezetek**), valamint a DNS fizikai térképezése (**ld. 9.1.1. fejezet**) és az ahhoz kapcsolódó területek (pl. diagnosztika, bűnügyi orvosszakértői vizsgálatok). A restriktációs endonukleázokat hagyományosan négy csoportba osztják a felismert szekvencia, a hasítóhely, az alegység kompozíció és a kofaktor szükséglet szerint.

3.2.1.1. I-es típusú restriktációs endonukleázok

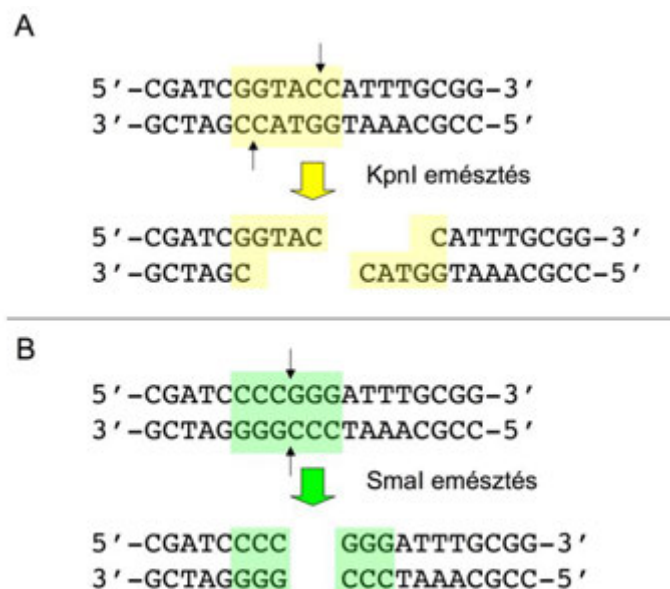
Multifunkciós fehérjék restriktációs és metil-transzferáz aktivitással rendelkeznek. Komplex, három alegységből felépülő szerkezetük van, Mg^{2+} szükséges a működésükhöz, valamint kofaktoruk még az ATP és az S-adenozil-L-metionin (AdoMet) is. Hasítóhelyük a felismerési helytől távol (több ezer bázispámyira) helyezkedik el. Korábban úgy gondolták, hogy az I-es típusú enzimek ritkák, azonban a teljes genomok vizsgálatából kiderült, hogy valójában nagy számban találhatók meg. Bár a szervezet biokémiai folyamataiban fontos szerepük van, gyakorlati hasznosíthatóságuk kicsi, mert nem szigorúan meghatározott a hasítás pozíciója a felismerő helyhez képest.

3.2.1.2. II-es típusú restriktációs endonukleázok

Változatos szerkezetű enzimek, csak restriktációs funkcióval rendelkeznek (a metiltranszferáz komponenst külön gén kódolja). A működésükhöz Mg^{2+} -ot igényelnek. **Hasítóhelyük a felismerő helyben vagy annak közvetlen közelében található.** Ebbe a családba tartozik a restriktációs endonukleázok túlnyomó többsége, eddig körülbelül 4000 ilyen típusú enzimet írtak le. A **felismerőhely** általában **4-8 bázispár** hosszúságú, és jellemzően **palindrom szekvencia**. A palindrom kifejezés azt jelenti, hogy a DNS bázissorrendje előre és hátrafelé olvasva (a két, komplementer láncra vonatkoztatva) ugyanaz. Például a KpnI nevű enzim felismerő helye az 5'-GGTACC-3' szekvencia. Ennek a komplementer párja 3'→5' irányban CCATGG, ami visszafelé olvasva éppen GGTACC.

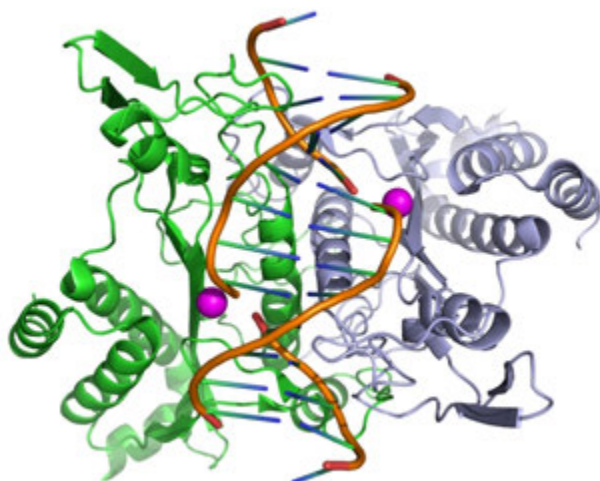
A **restriktációs enzimek hasítóhelye eredményezhet ragadós véget vagy tompa véget.** Előbbinél a hasítás eredményeként az egyik szálon rövid túlnyúló vég jön létre, például az említett KpnI az ötödik nukleotid után hasít (GGTAC/C), és így ragadós véget eredményez, 3'-túlnyúló véggel (**ld. 3.4. ábra**). Ezzel szemben a SmaI enzim, amelynek felismerő helye a CCCGGG szekvencia, a harmadik nukleotid után hasítva (CCC/GGG) tompa véget

hoz létre. A restrikciós enzimek egyik legismertebbje, az EcoRI a GAATTC szekvenciát ismeri fel és az első nukleotid után hasítja (G/AATTC), ezáltal 5'-túlnyúló vég jön létre (ez a gyakoribb túlnyúló vég típus).



3.4. ábra: **Ragadós és tompa végek.** A restrikciós enzimek hasításának terméke lehet ragadós végű (*sticky end*) vagy tompa végű (*blunt end*) DNS. A) A KpnI enzim ragadós végeket hoz létre, mert a palindrom hasítóhelyet aszimmetrikusan vágja szét. A kapott ragadós végek könnyen ligálhatók. B) A SmaI enzim a hasítóhely közepén vág, ezért nem keletkezik túlnyúló vég.

A II-es típusú enzimek meghatározott helyen hasítják a DNS-t, így **az emésztés reprodukálható és a keletkezett termékek szekvenciája pontosan meghatározott lesz.** Ma már több ezer ilyen enzimet ismerünk (>3000 enzim, >250 felismerőhellyel 750 baktérium törzsből) és sok száz kapható kereskedelmi forgalomban. A II-es típusú endonukleázok legtöbbje homodimer szerkezetű (**ld. 3.5. ábra**), amely szerkezeti tulajdonság egyrészt összecseng a felismerőhely szimmetrikus szerkezetével, másrészt a két azonos DNS-kötőhely (bivalencia) erősebb DNS kötődést tesz lehetővé.



3.5. ábra: Az **EcoRI** restrikciós enzim térszerkezete. A II-es típusú restrikciós endonukleázok egyik tipikus képviselője az EcoRI enzim. Homodimer szerkezetű, egyenként 31 kDa méretű alegységekből áll. Mindkét alegység tartalmaz egy hurok régiót, amelyek körbeölelik a DNS-t (narancssárga). Az ábrán látható, hogy a DNS aszimmetrikus elhasadása révén ragadós végek jönnek létre. Az enzim kristályosítása mangán-ionok jelenlétében történt (lila gömbök), amelyek a natív enzimben megtalálható magnézium-ionok helyét foglalják el.

Később több alkategóriáját állították fel ennek az enzimsaládnak, ahová azok az enzimek tartoznak, amelyek jellegzetes tulajdonságaikban különböznek a tipikus II-es típusú endonukleázoktól. A II.M típusú enzimek például képesek metilált DNS felismerésre és elhasítására is. Ilyen a DpnI nevű restrikciós endonukleáz, amelyet helyspecifikus mutagenézisnél használnak a templát DNS elbontására (**ld. 11.2.2. fejezet**). A II.M típusú enzimek meghatározott helyen hasítják a DNS-t, a DpnI hasítólye például GA/TC. A IV-es típusú restrikciós endonukleázoknak is metilált DNS a szubsztrátja, azonban a II.M típustól eltérően nem specifikus helyen hasítanak.

3.2.1.3. III-as típusú restrikciós endonukleázok

Nagyméretű, két alegységből felépülő fehérjék (hidrolizáló alegység és metiláló alegység). Szintén Mg^{2+} -függő enzimek, kofaktoruk az ATP, az S-adenozil-L-metionin segíti a metil transzfer reakciót, de nem feltétlen szükséges a működésükhöz. Hasítóhelyük a felismerési hely közelében (általában 25-27 bázispárnira) található, nem palindrom. Akárcsak az I-es típusú enzimek, itt is két szekvencia részből áll a felismerő hely, és az enzim képes a DNS szál transzlokációjára. A hasítás az egyik felismerő hely közelében történik.

3.2.1.4. IV-es típusú restrikciós endonukleázok:

Módosított, általában metilált DNS-t ismernek fel és hasítanak, specifitásuk alacsony. Nem a restrikciós-modifikációs rendszer részei.

3.2.2. Egyéb nukleázok

3.2.2.1. S1-nukleáz

Az *Aspergillus oryzae* gombából származó S1-nukleáz olyan endonukleáz, amelyik egyszálú DNS-t vagy RNS-t képes lebontani. Az enzim ötször aktívabb DNS-en, mint RNS-en. Kétszálú polinukleotid is lehet az S1-nukleáz szubsztrátja, ha tartalmaz egyszálú szakaszokat, ilyen lehet egy hurok régió vagy egy *gap*. Kofaktora a Zn^{2+} , pH optimuma a savas tartományban, 4.0-4.3 között van. Felhasználható tompa vég kialakítására, illetve egyszálú DNS és RNS szelektív lebontására.

3.2.2.2. Dezoxiribonukleáz (DN-áz) I

A dezoxiribonukleáz I (DN-áz I), ellentétben a fentebb tárgyalt S1-nukleázzal, **egy- és kétszálú DNS lebontására** egyaránt képes. A 31 kDa méretű enzim négy izoenzim keveréke, és maximális aktivitást kalcium-, magnézium- vagy mangán-ion jelenlétében mutat. Specifitása és hatásmechanizmusa a jelen levő kofaktortól függ. Míg például Mg^{2+} jelenlétében a két szálát random helyeken hasítja, Mn^{2+} jelenlétében nagyjából egy helyen történik a hasítás, tompa vagy egy-két nukleotidnyi túlnyúló véget kialakítva. A molekuláris biológiai gyakorlatban sokrétűen felhasználható, például didezoxi- (Sanger-féle) szekvenáláskor véletlenszerű DNS fragmentumok kialakítására, vagy jelölt DNS létrehozására hasadás-elmozdulás (*nick translation*) technikával (**ld. 3.3. ábra**). Nagyon gyakran alkalmazzák fehérje vagy RNS preparátumok DNS-től való megtisztítására is.

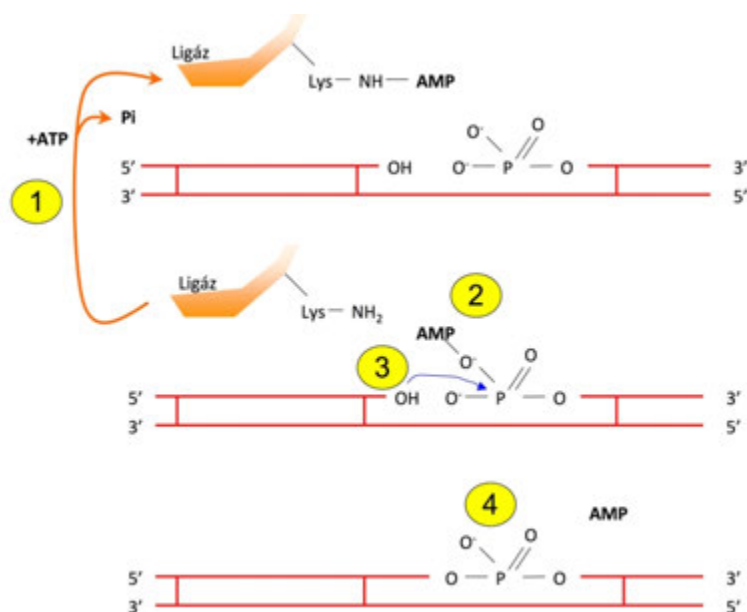
3.3. DNS-ligázok

A restriktációs endonuklázal kezelte DNS fragmentumok újra összeköthetők kovalensen **DNS-ligáz** hozzáadásával. A ligáz típusától függően a reakció kofaktora ATP (eukarióták, vírusok és bakteriofágok) vagy NAD^+ (baktériumok).

Az emlősökben a DNS-ligázok négy típusa fordul elő, szerepük van többek között az Okazaki-fragmentumok összekapcsolásában (DNS-ligáz I), illetve különböző hibajavító mechanizmusokban, mint például a nukleotid-kivágással járó DNS hibajavítás (DNS-ligáz III), vagy a kétszálú DNS törése és javítása (DNS-ligáz IV).

A laboratóriumi gyakorlatban leggyakrabban a **T4 bakteriofággal** fertőzött *E. coli*-ből izolált **T4-ligázt** használják. Hatásmechanizmusának lényege, hogy képes a DNS egy szálán található "lyukakat" foszfodiészter **kötés** létrehozásával betömni (**ld. 3.6. ábra**). Kicsit bonyolultabb eset, ha egy restriktációs hasítás eredményeként kapott két DNS véget szeretnénk újra összekapcsolni, vagy egy plazmidba szeretnénk új inszertet ligálni. Ekkor az enzimnek két foszfodiészter kötés szintézisét kell elvégeznie. Aktivitásához ATP-re van szüksége (ellentétben az *E. coli* DNS-ligázzal, amely érdekes módon energiaforrásként NAD-ot használ). Ezen kívül az is feltétele a sikeres reakciónak, hogy **a ligálandó két DNS vég egymáshoz megfelelően közel kerüljön**, mert az enzim csak a szintézist végzi, a DNS két végét nem tartja egymás közelében. Ragados végek esetén egyszerűbb a helyzet, mert a véletlenszerűen egymáshoz közel kerülő végek gyengén (a bázispárok hidrogénkötéseinek keresztül) összekapcsolódhatnak, és ha ez a kölcsönhatás elegendő ideig fennáll, a ligáz hozzá tud kapcsolódni a DNS-hez, és létrehozza a hiányzó foszfodiészter kötések (kondenzációs reakcióval). Tompa végek esetén nincs lehetőség az említett átmeneti kölcsönhatás kialakulására, ezért a ligálási reakció hatékonysága ebben az esetben alacsony (az enzimkoncentráció növelésével javítható). Megjegyzendő, hogy ellentétben a T4 DNS-ligázzal, az *E. coli* DNS-ligáz a tompa végek összekapcsolására nem képes.

A laboratóriumban a ligálási reakció egyik legfontosabb faktora a hőmérséklet. Az enzim hőmérsékleti optimuma 25°C , ezen a hőmérsékleten azonban csökken a DNS végek átmeneti kölcsönhatásának esélye. Ezért a ligálást tipikusan alacsonyabb hőmérsékleten, 16°C -on végzik, de a túlnyúló DNS végektől függően akár arra is szükség lehet, hogy $4-8^{\circ}\text{C}$ -on történjen a reakció. Ilyenkor a szokásos néhány perc helyett több óra szükséges a hatékony ligáláshoz.



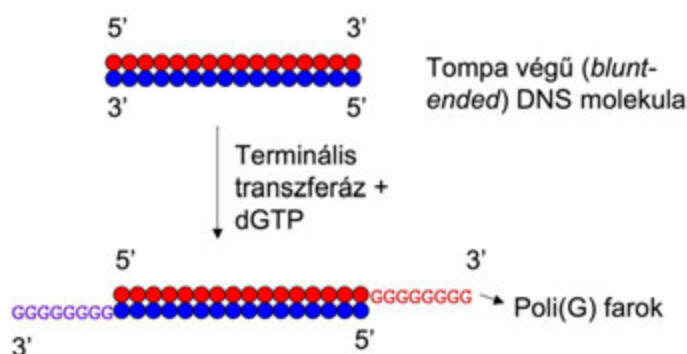
3.6. ábra: A DNS-ligáz mechanizmusa. A DNS ligázok által katalizált reakcióknak három alapvető lépése van.

1) Elsőként kovalens enzim-AMP intermedier jön létre pirofoszfát keletkezése mellett (feltéve, hogy ATP a kofaktor), az enzim egy lizin aminosaván keresztül. 2) Következő lépésként az adenil-csoport (AMP) a DNS szabad 5'-foszfát csoportjának adódik át, és 3) aktiválja azt. Végül a 4.) lépésben a 3'-hidroxil csoport nukleofil támadóként reagál az aktivált 5'-foszfát csoporttal, így összekötve a foszfátgerincet AMP felszabadulása mellett.

3.4. Végmódosító enzimek

3.4.1. Terminális transzferáz

A biotechnológiai gyakorlatban borjú csecsemőmirigyéből izolált terminális transzferáz enzim **homopolimer ragadós végek kialakítására** használható ("farkazás"). Az enzim gyakorlatilag egy **templát-független DNS-polimeráz**. Ahogy fent említettük, a tompa végű ligálás hatásfoka igen alacsony, amin segíteni lehet ragadós vég mesterséges kialakításával. A tompa véget eredményező hasítás után a terminális transzferáz aktivitását kihasználva például poli(G) "farkat" lehet a DNS-re szintetizálni dGTP hozzáadásával, ami egy hasonló reakcióban előkészített, de dCTP hozzáadásával kialakított, immár ragadós végű DNS darabhoz ligálható (ld. 3.7. ábra).

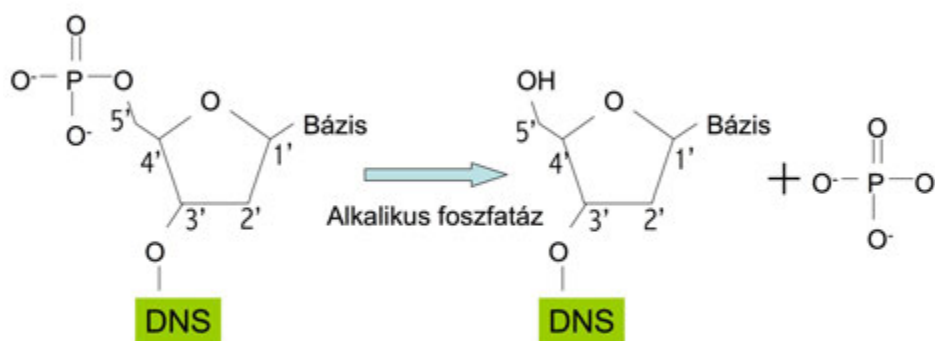


3.7. ábra: A terminális transzferáz enzim aktivitása. A terminális transzferáz aktivitása használható tompa végű DNS molekulák módosítására, mivel templát nélkül is képes polinukleotid lánc szintézisére. Az ábrán bemutatott példán dGTP hozzáadásával az enzim poli(G) láncot szintetizál a DNS 3'-végére, amely könnyen ligálható a komplementer szálon poli(C) farkat tartalmazó molekulával.

3.4.2. Alkalikus foszfatáz

A végmódosító enzimekhez sorolható a hidrolázok közé tartozó alkalikus foszfatáz is. Általánosságban az alkalikus foszfatáz funkciója **foszfátcsoport eltávolítása** különböző makromolekulákról (nukleinsavak, fehérjék, alkaloidok). Az alkalikus foszfatáz által katalizált folyamatot defoszforilálásnak nevezik (**ld. 3.8. ábra**).

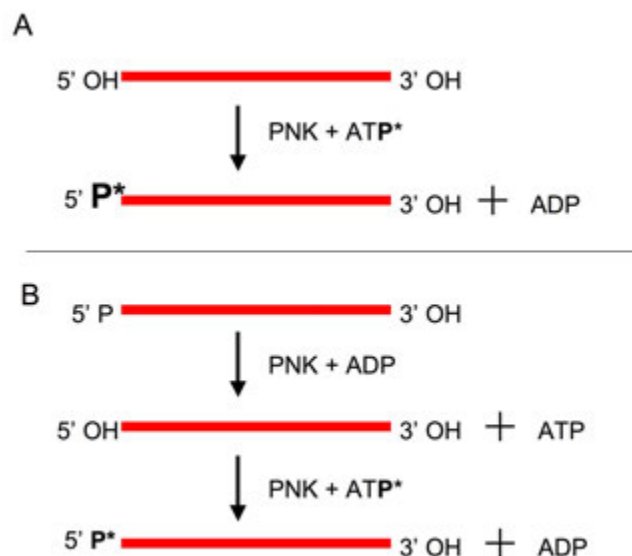
A molekuláris biológiai gyakorlatban leggyakrabban borjúbélből izolált alkalikus foszfatázt használnak (CIAP: *Calf Intestinal Alkaline Phosphatase*) a DNS 5'-végén levő foszfát-csoport eltávolítására. A plazmid vektorba való klónozás során az egyik lehetséges hibaforrás, hogy a felnyitott plazmid üresen ligálódik össze ("önzáródás"). A foszfátcsoport eltávolítása ezt akadályozza meg, növelve ezzel az inszertet tartalmazó konstrukciók számát és **csökkentve az úgynevezett "üres vektor" háttérét**.



3.8. ábra: **Az alkalikus foszfatáz által katalizált reakció.** Klónozáskor az üres plazmid önzáródásának kivédésére alkalmazható az alkalikus foszfatázzal való kezelés, amely monofoszfát észterek (mint pl. az 5'-foszforilált DNS) defoszforilálását katalizálja.

3.4.3. T4 polinukleotid-kináz

A T4 polinukleotid-kináz (PNK) az ATP **g-foszfát csoportjának a DNS szabad 5'-hidroxil végére történő átvitelét katalizálja**. (A DNS-ről hiányozhat az 5' foszfát-csoport előzetes defoszforiláció miatt, vagy ha például az oligonukleotid primereket kémiai szintézissel állították elő.) A PNK ezen kívül ADP jelenlétében képes a foszfát csoport cseréjére is (**ld. 3.9. ábra**). A reakció felhasználható oligonukleotidok, DNS vagy RNS 5'-végének jelölésére (g-³²P) ATP használatával.



3.9. ábra: **A T4 polinukleotid-kináz enzim működése.** A) A T4 polinukleotid kináz (PNK) az ún. “forward” reakció során a polinukleotid szál 5'-végére viszi át az ATP g-foszfát csoportját. Radioaktív ATP-t alkalmazva a reakcióval foszfát-jelölt (P*) nukleotidot lehet előállítani. B) A kicserélődési reakció akkor megy végbe, ha feleslegben ADP-t adunk a reakcióhoz. Ekkor az enzim első lépésben a nukleinsav 5'-foszfátját viszi át az ADP-re, majd a jelölt ATP g-foszfátját a nukleinsavra.

3.5. További olvasnivaló a fejezethez

Tindall KR, Kunkel TA (1988) Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry* 27, 6008-6013.

Pingoud A, Fuxreiter M, Pingoud V and Wende W (2005) Type II restriction endonucleases: structure and mechanism. *Cell. Mol. Life Sci.* 62, 685-707.

Martin IV, MacNeil SA (2002) ATP dependent DNA ligases. *Genome biology* 3, reviews 3005-reviews3005.7

Oakeley EJ (1999) DNA methylation analysis: a review of current methodologies. 1999. *Pharm. Ther.* 84, 389-400.

4. fejezet - Alapmódszerek és rekombináns DNS konstrukciók tervezése

A géntechnológiai módszerek fejlődése lehetővé tette, hogy bármilyen organizmusból izolálhassunk bármilyen tetszőleges gént, az izolált gént klónozással felszaporítsuk, és szekvenálással „kiolvassuk” a genetikai információt. Ebben a fejezetben a géntechnológia alapmódszereit ismertetjük, kezdve a DNS tisztítástól és a szeparációs technikáktól (gélelektroforézis, ld. 4.1.2.), a hibridizációs technikákon (ld. 4.1.3.) át a DNS konstrukciók tervezéséig (ld. 4.2.) és a DNS gazdasejtbe juttatásának módszereit (génbeviteli eljárások, ld. 4.3.).

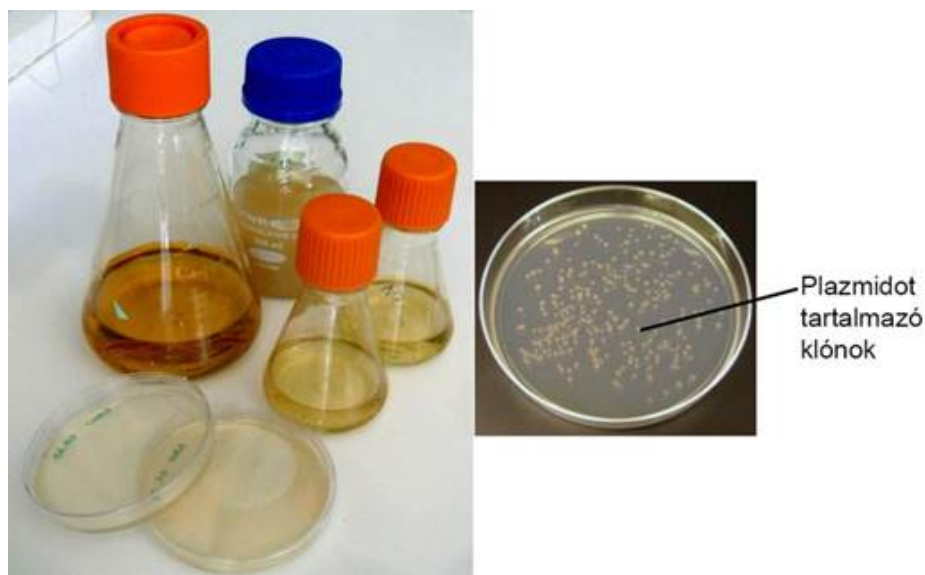
4.1. Alapmódszerek

4.1.1. A DNS tisztítása és analízise

A rekombináns DNS-ek előállítása a vektor és a majdani inszert DNS tisztításával kezdődik. Ezt követően meg kell győződnünk a DNS pontos méretéről, mennyiségéről és tisztaságáról. A DNS izolálása az RNS-hez képest egyszerűbb, mert az RNS esetében nehéz megszabadulni a szinte mindenhol jelen lévő RNáz enzimektől. Ez az alfejezet a DNS előállításához és izolálásához szükséges oldatokat, reagenseket, a tisztításhoz használt módszereket (pl. „miniprep”) és a DNS koncentrációmérés lehetőségeit mutatja be.

4.1.1.1. Oldatok, reagensek

A vektor DNS-t baktérium tenyészetekben állítjuk elő, s ugyanezeket a tenyésztési körülményeket használjuk a rekombináns DNS felszaporításához is a molekuláris klónozás során. A baktériumok tenyésztéséhez, növekedéséhez táptalajra van szükségünk. Leggyakrabban steril **LB** (Luria-Bertrani; összetétele: tripton, élesztő kivonat, NaCl) vagy **2YT** (több élesztő kivonatot és triptonot tartalmaz) **médiум**mal dolgozunk. A mikrobiológiában használatos táptalajok halmazállapotuk szerint lehetnek folyékonyak és szilárdak. Szilárd táptalajt úgy kaphatunk, ha a sterilizést megelőzően valamilyen szilárdító anyagot keverünk a tápleveshez, leggyakrabban agart. Ilyenkor magas hőmérsékleten az agar feloldódik, majd a kihűlés előtt Petri-csészékbe öntjük, és hagyjuk megdermedni (**agar lemez**). A **4.1. ábrán** folyadék és szilárd halmazállapotú táptalajokat mutatunk be. Nagy mennyiségű baktériumsejt felneveléséhez folyékony táptalajra van szükség, de abból az egyedi klónokat nem tudjuk szétválasztani. Azért kell a szilárd táptalaj, hogy a sejtszuszpenziót hígán elosztatva, az egyedi klónok egymástól szétválasztható egyedi telepeket képezzenek.



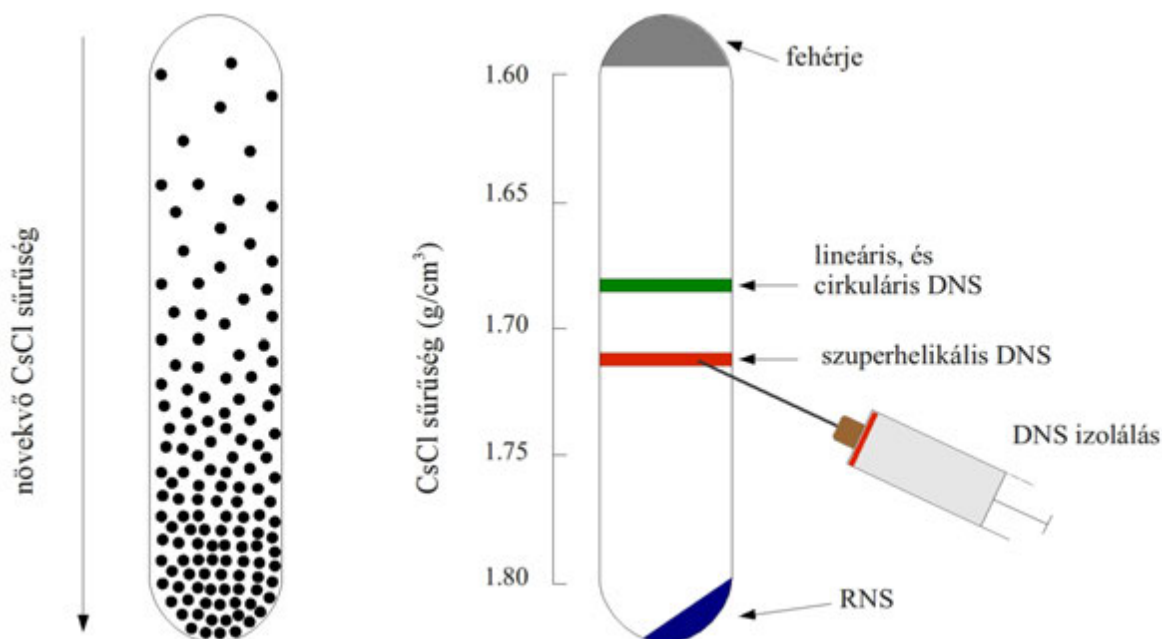
4.1. ábra: **Folyadék és szilárd halmazállapotú táptalajok.** A jobb oldali lemezen az egyedi baktériumtelepek (klónok) is látszanak.

Nagy mennyiségű plazmid DNS előállításához először a felszaporítani kívánt plazmidot (**ld. 7.2. fejezet**) kompetens *E.coli* baktériumokba kell transzformálni (**ld. 4.3.1. .**). Ahhoz, hogy kizárólag a plazmidot felvett baktériumsejtek növekedjenek, a táptalajhoz valamilyen **antibiotikumot** is kell adni, hiszen a plazmidot felvett sejtek rendelkeznek legalább egy antibiotikum rezisztenciát kódoló génnel (**ld. 7.1.2. fejezet**), és így tovább tudnak osztódni, míg az „üresen” maradt sejtek elpusztulnak (**szelekció**).

4.1.1.2. Plazmid DNS preparálás

A felszaporítandó plazmidot tartalmazó baktériumsejtekkel folyékony táptalajt inokulálunk, és egy éjszakán át, 37°C-on növesztjük, majd a sejteket centrifugálással összegyűjtjük. A baktériumsejtekből történő DNS izolálás klasszikus módszere az ún. **alkalikus lízis**, amely során SDS-t tartalmazó NaOH oldattal tesszük tönkre a sejtmembránt és denaturáljuk a makromolekulákat. (Birboim and Doly dolgozták ki ezt a módszert 1979-ban). A sejtek feltárását elérhetjük enzimekkel (lizozim, proteináz K) vagy mechanikai hatással is (pl. szonikálás ultrahanggal, nagy nyomás az ún. French-press módszernél, többszörös fagyasztás-felolvasztás). Az oldatba került DNS molekulákat el kell választani a többi sejttermeléktől és fehérjéktől. Az óriási méretű kromoszomális DNS-t (3,6 Mbp) könnyen ki lehet csapni (neutrális pH-n), az RNS-eket és a különböző fehérjéket enzimekkel lehet degradálni. A fenol-kloroform-izoamilalkoholos extrakció esetén a fenol és a kloroform, mint szerves oldószerek denaturálják a fehérjéket, a plazmid méretű (2-10 kbp) nukleinsavak viszont a vizes fázisba kerülnek. Az izoamilalkohol csökkenti a bifázisos oldat habosodását. Végül a DNS-től etanollal vagy izopropanollal vonjuk el a vizet és ezzel csapjuk ki, majd centrifugálással kiülepítjük az extrakromoszomális plazmid DNS-t egy műanyag Eppendorf-cső falára. Az alkohol eltávolítása után vízben vagy Tris-EDTA (TE) pufferben visszaoldjuk a tiszta DNS-t.

A nyers lizátum tisztításának egy másik hagyományos módja a **CsCl-os sűrűséggradiens centrifugálás**, amit manapság már ritkábban használnak, mivel fél napos nagy fordulatszámú (350.000 g) ülepítés szükséges hozzá. Viszont nagy tisztaságú DNS-t lehet vele kinyerni, és épségben maradnak a nagyobb méretű (>10 kbp) DNS molekulák is, amelyek mechanikailag érzékenyek a pipettázásnál is fellépő nyíróerőkre. A **4.2. ábra** a CsCl sűrűséggradiens centrifugálás sémáját mutatja be.



4.2. ábra: **A sűrűséggradiens centrifugálás sémája.** Erős centrifugális erő hatására a nehéz Cs^+ ionok a centrifugacső aljába kényszerülnek, ezzel egy vékony sűrűséggradienst képezve. A DNS molekulák ebben a gradiensben addig a pontig fognak vándorolni, amíg el nem érik a gradienssel ekvivalens sűrűséget (izopiknikus pont; 1,6-1,8 mg/ml CsCl koncentrációnál). Centrifugálás hatására a CsCl oldatban lévő makromolekulák különböző helyzetű sávokat alkotnak a gradiensben. Az alig különböző méretű és alakú molekulák is képesek szeparálódni (például különböző GC gazdagságú vagy térbeli szerkezetű DNS molekulák). A DNS-t etídium-bromiddal kezeljük, amely interkalálódik a DNS bázispárjai közé. A DNS sávok elhelyezkedését UV fényben lehet detektálni a fluoreszcens festéknek köszönhetően. Az izolált DNS-ről végül n-butanollal az etídium-bromid, dializálással a CsCl választható el.

Az eddig ismertetett módszerek hátránya, hogy nagy elővigyázatosságot igényel a mérgező reagensek (pl. fenol, etídium-bromid) és a veszélyes hulladék miatt, drága eszközök (pl. ultracentrifuga) szükségesek hozzá és rendkívül időigényesek. A kis mennyiségű bakteriális DNS („miniprep”) alkalikus lízisen alapuló kivonására manapság számos, kereskedelmi forgalomban kapható kit áll rendelkezésre, melyekkel nagy tisztaságban és mennyiségben lehet kinyerni a plazmid DNS-t. Ezek a kitek alapvetően a részletesen ismertetett alkalikus lízis módszerén alapulnak, de a fenolos extrakciós lépés helyett egy **szilikát tartalmú oszlopon (vagy membránon) történő tisztítást** tartalmaznak. A minipreppel általában 10-100 μg DNS-t lehet kinyerni. Vannak olyan kitek is, amelyekkel jóval nagyobb mennyiségű DNS-t is lehet tisztítani. A „midiprep” módszernél kb. 25-50 ml sejtkultúrából indulunk ki, és 100-350 μg DNS-t kaphatunk. Ennél több DNS „maxi-” „mega-” és „gigaprep”-pel (100 ml-5l sejtkultúra, 500 μg -10 mg DNS) állítható elő. „Miniprep” protokoll-leírásokat többek között itt találunk az interneten.

4.1.1.3. A DNS koncentráció meghatározása

A plazmid DNS koncentrációját leggyakrabban **fotometriával** mérjük. A nukleinsavak elnyelési maximuma 260 nm-nél van az aromás gyűrűk miatt. Az elnyelés alapján megadható a DNS koncentrációja. Ha az abszorbancia $A_{260} = 1$, ez az érték megfelel 50 $\mu\text{g}/\text{ml}$ duplaszálú DNS-nek (37 $\mu\text{g}/\text{ml}$ egyszálú DNS-nek vagy 40 $\mu\text{g}/\text{ml}$ RNS-nek). A fehérjék elnyelési maximuma 280 nm-nél van, tehát különbözik a DNS-étől, ezért könnyedén megállapítható a minta fehérje szennyezettsége is, melyet az A_{260} / A_{280} hányados határoz meg pontosan. Ha ez az érték kevesebb, mint 1,8, akkor a minta fehérjével szennyezett. Ha viszont 2 felett van, akkor RNS-t is tartalmazhat. A mérést elvégezhetjük hagyományos UV fotométerben, de ehhez sok anyagra van szükségünk. Ma már legtöbbször ún. **nanodrop technikát** alkalmazunk (ld. 4.3. ábra), melynél egy csepp (~2 μl) DNS koncentrációját mérjük egy speciális spektrofotométerben.



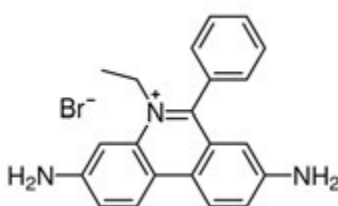
4.3. ábra: DNS koncentráció mérése nanodrop technikával

4.1.2. Gélelektroforézis

Mindenfajta gélelektroforézis azon az elven alapul, hogy a töltéssel rendelkező molekulák az elektromos térben az ellentétes töltésű elektróda felé vándorolnak. A vándorlás sebességét a töltésen kívül a molekula alakja és mérete határozza meg. Ez a három tulajdonság teszi lehetővé, hogy különböző minőségű és pórusméretű gélekben az eltérő tulajdonságú biomolekulákat elválasszuk egymástól.

4.1.2.1. Agaróz gélelektroforézis

Az agaróz a vörösmoszatok agar-agarjának egyik poliszacharid összetevője, mely a D-galaktóz és 3,6-anhidro-L-galaktopiranoz lineáris polimere. Rendkívül jól alkalmazható nukleinsavak szeparálására. A biokémiában az agaróz gélelektroforézis DNS molekulák, elsősorban méret és alak szerinti elválasztására szolgál. Agaróz gélben **néhány 10 bp-tól több 10 kbp-ig terjedő méretű** (hosszúságú) **DNS molekulákat lehet szétválasztani**. A DNS molekula negatív töltése cukorfoszfát gerince miatt egyenesen arányos a hosszúságával. Ezért egy vízszintesen elhelyezett agaróz gélben (**horizontális gélelektroforézis**) a pozitív pólus felé vándorló DNS molekulák mobilitása arányos lesz a hosszukkal. Az agaróz 0,5-2%-os oldata forralással készül. Dermedéskor az agaróz térhálós szerkezetű lesz. A forralást követően a gélhez **interkaláló fluoreszcens festéket** is adunk, amely a DNS-t UV fényben láthatóvá teszi. Régebben a már korábban is említett etidium-bromidot (**ld. 4.4. ábra**) alkalmazták erre a célra, azonban ez erősen mutagén hatású, ezért manapság más, biztonságosabb festékeket használunk helyette (pl. SYBR Safe, Gel Green).

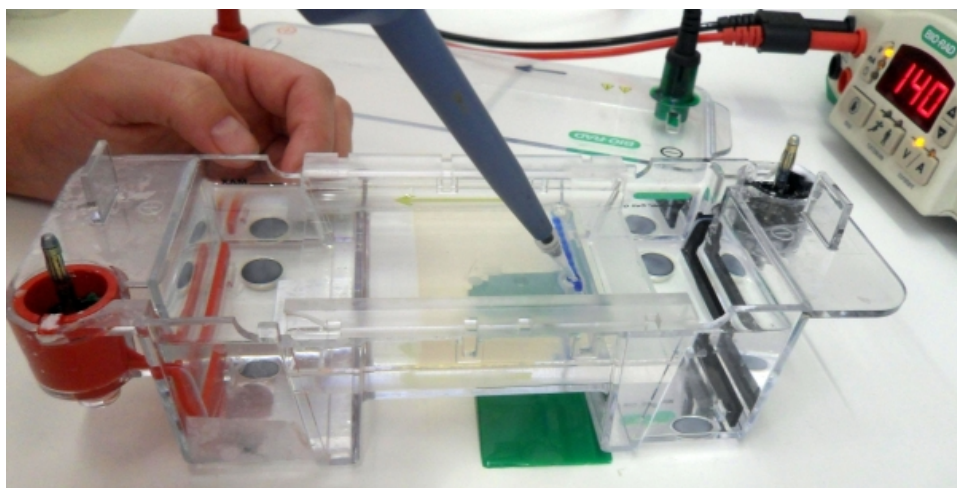


4.4. ábra: Az etidium-bromid fluoreszcens DNS-festék szerkezeti képlete

A gél öntésekor, az öntő formára merőlegesen, felülről egy ún. "fésűt" is elhelyezünk, ami a mintafelvitelhez szükséges zsebeket fogja kialakítani. A fésű méretét a minta mennyiségétől függően választjuk ki. A dermedt gélből óvatosan kivesszük a fésűt, és a zsebekbe pipetázzuk a mintánkat, majd egy elektródákkal ellátott kádba helyezzük a géllel. Az elektroforézishez szükséges egyenáramot a **tápegység** biztosítja. Minden egyes mintához "DNS-kezelő" oldatot is keverünk, mely glicerint és általában kétféle festéket is tartalmaz. A glicerint arra szolgál, hogy nagyobb sűrűsége révén a mintákat a zseb alján tartsa, ne diffundáljanak ki a pufferbe. A festékek segítségével pedig megállapítható, hogy hol tart az elektroforézis. A brómfenolkék láthatóvá teszi felvitelkor a mintát, és kiváló mobilitási képességei következtében (1%-os gélben egy 300 bp-os DNS fragmentummal fut együtt) sötétkék színe a futási frontot jelzi. A xilén-cianol nagyon lassan halad a gélben (kb. 4 kbp DNS-sel fut együtt).

Az elektroforézishez használt puffer (TBE: Trisz-borát-EDTA vagy TAE: Trisz-acetát-EDTA) ionerőssége fontos tényező a DNS elektroforetikus mobilitása szempontjából. Túl alacsony ionerőn a DNS csak lassan képes haladni, míg ha túl magas az ionerő, akkor a vezetőképesség is növekszik, túl sok hő képződik, amitől a gél meg is olvad. Alacsony feszültségen a lineáris DNS vándorlási sebessége arányos a feszültséggel. A megfuttatott gél végül UV fénnel világítjuk meg, s a gélhez adott fluoreszcens festéknek köszönhetően világítani fognak a DNS-ek sávjai a gélben. A láthatóvá vált DNS sávokat ki is vághatjuk a gélből további tisztításra, felhasználásra.

Az agaróz gélelektroforézis kivitelezését a **4.5. ábrán** és a 4.1. videon szemléltetjük.

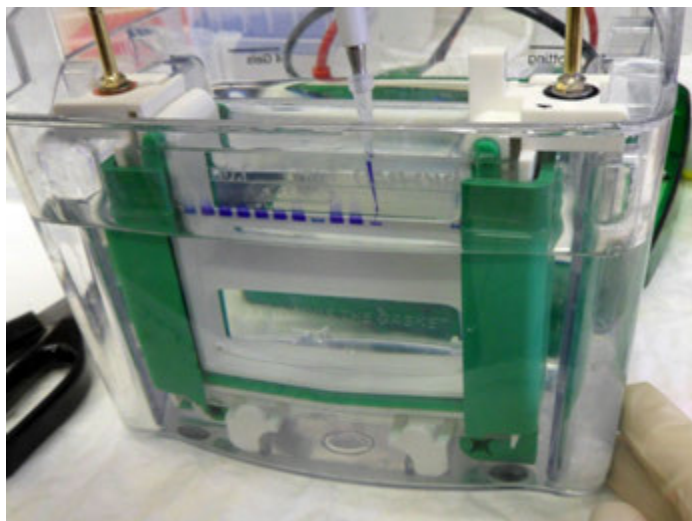


4.5. ábra: Agaróz gélelektroforézis

4.1. video: Az agaróz gélelektroforézis kivitelezése

4.1.2.2. Poliakrilamid gélelektroforézis (PAGE)

A nukleinsavakat nemcsak agaróz, hanem poliakrilamid gélelektroforézissel (PAGE) is el tudjuk választani egymástól. A PAGE-t a horizontális agaróz elektroforézistől eltérően vertikális (függőleges) elrendezésű készülékben végezzük (**ld. 4.6. ábra**). A poliakrilamid gél egyik legjelentősebb eltérése az agaróz géltől, hogy sokkal kisebb a pórusmérete. **Felbontása 5-2000 nukleotidig terjed**, és képes elválasztani az akár 1 nukleotid hosszban különböző oligo- vagy polinukleotid szálakat is egymástól. A Sanger-féle klasszikus DNS szekvenálás során PAGE-vel választják el a különböző hosszúságú lineáris, újonnan szintetizált egyszálú DNS-eket egymástól. Megjegyzendő, hogy az automata DNS szekvenátorokban kapilláris elektroforézist használnak (**ld. 5.2.2. fejezet**), amihez más szintetikus polimereket alkalmaznak. Az akrilamid vizes közegben, megfelelő iniciátorok (pl. peroxidiszulfát) és katalizátorok (tetrametil-etilén-diamin, TEMED) jelenlétében gyökös polimerizációra képes, és nagy molsúlyú lineáris polimer, poliakrilamid keletkezik. A térhálós szerkezet a keresztkötő ágens (N,N-metilén-bisz-akrilamid) jelenlétében alakul ki, aminek segítségével a hosszú poliakrilamid láncok között hidak képződnek. Megjegyzendő, hogy a fehérjék elválasztására is PAGE-t használunk, mivel még a nagyobb méretű fehérjék (>100 kDa) is jóval kisebbek, mint egy kisméretű plazmid (1 kbp ~ 600 kDa).



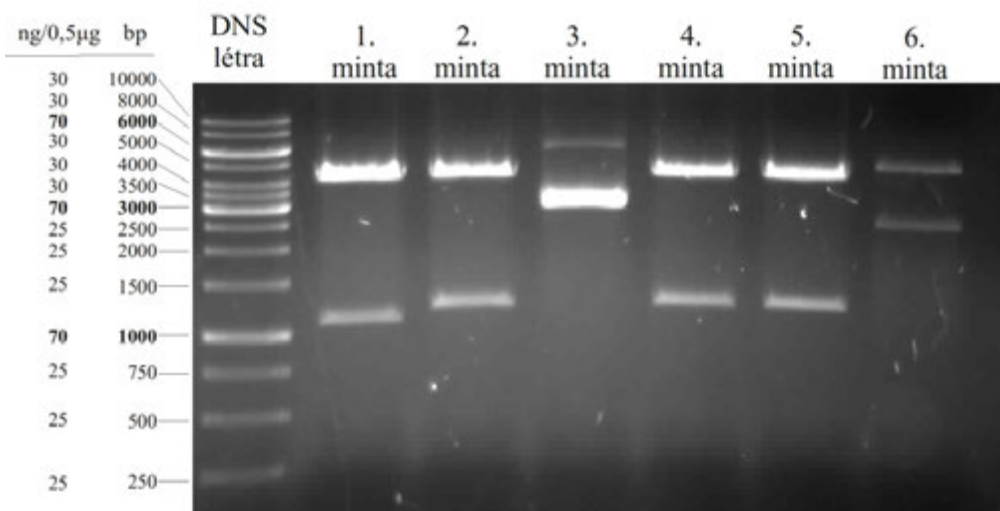
4.6. ábra: Poliakrilamid gélelektroforézis készülék „kamráinak” felöltése mintákkal.

Mivel a különböző hosszúságú DNS molekulák relatív töltése azonos (a foszfát-csoportok negatív töltése miatt), ezért a lineáris molekulák mindig méret szerint válnak szét. Ha valamilyen denaturáló ágenszt is keverünk a gélhez, például ureát, akkor a lineáris DNS szálak alakja is egyforma lesz, így ténylegesen csak a méret alapján szeparálódnak. Fehérjék esetében is hasonló az elválasztás elve, de fehérjék szeparálására a PAGE számos más változatát is használjuk (2D PAGE, SDS-PAGE, izoelektromos fókuszálás).

4.1.2.3. A DNS-ek méretének meghatározása

A DNS méretét ismernünk kell, ha azt analizálni vagy preparálni szeretnénk. A lineáris DNS molekula méretének logaritmusa és az elektroforetikus mobilitása között fordított arányosság van (a mobilitás és a méret logaritmusa között lineáris az összefüggés). Így egy ismeretlen molekula mérete meghatározható a futási távolság ismeretében, ha mellé olyan DNS molekulákat is felviszünk a gélre, amelyeknek a mérete pontosan ismert. Ez a **DNS „létra”** vagy molekulasúly marker, ami különböző méretű lineáris DNS molekulák keveréke. Az agaróz gél alapvetően nagyobb méretű nukleinsavak elválasztására használják, mivel pórusmérete jóval nagyobb a poliakrilamid gélhez képest. A gél agaróz vagy akrilamid koncentrációjától függően, különböző méretű polinukleotidokat lehet elválasztani. Minél hígabb, annál nagyobb, minél töményebb, annál kisebb DNS szálakat lehet szeparálni. Minden vizsgálandó mintához tartozik egy optimális pórusméret, amelyben a molekulák bizonyos ellenállással vándorolnak, de nem akadnak fel, hanem képesek haladni benne.

A hosszabb molekulák a gélben „nehezebben” haladnak, így ezek lemaradnak a rövidebb, kisebb molekulákhoz képest. A polinukleotidok gélben való **mobilitásánagyban befolyásolja az alakjuk is**. A cirkuláris/relaxált plazmid formák széttekeredtségük miatt lassabban haladnak, mint az erősen kondenzált, szuperhelikális DNS-ek. A linearizált plazmidok pedig az előbb említett két forma között mozognak (a molekulák alak-szerinti mobilitását a gélben alapvetően a hidrodinamikai térfogatuk dönti el. DNS minták agaróz gélelektroforetikus képét a **4.7. ábra** mutatja be.



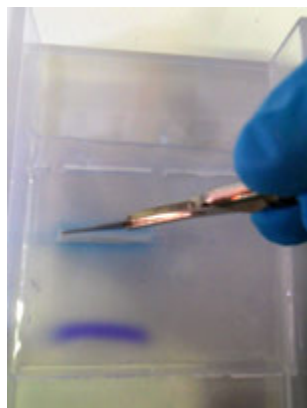
4.7. ábra: Agaróz gél DNS létrával és különböző DNS mintákkal. A mintákat valamilyen restrikciós endonukleázzal (1- vagy 2-féle) kezeltük. A 3. minta kivételével a DNS minták kettévágódtak, ezért két különböző méretű sávban futnak. A 3-as plazmid DNS minta cirkuláris maradt. A kondenzáltabb szuperhelikális formák gyorsabban, míg a relaxált, kitekeredett molekulák lassabban vándorolnak a gélben. A DNS létrával a molekula méreten kívül a DNS mennyiségét is meg lehet becsülni.

4.1.2.4. DNS-ek izolálása agaróz gélből

A megfelelő DNS-t (pl. vektor vagy inszert DNS), melyet pl. restrikciós emésztés vagy PCR reakció eredményeként kaptunk, az agaróz gélből izolálnunk kell ahhoz, hogy megfelelően elválasszuk a DNS mellett maradt enzimektől (endonukleázok, DNS-polimeráz) vagy a hasítatlan vektortól.

Az agaróz gélelektroforézist követően a gélben az esetben UV fény helyett kék fényben világítjuk meg, mert az UV roncsolná az izolálni kívánt DNS-t. A gélhez adott DNS festék (etídium-bromid helyett pl. SYBR Safe) ugyanis kék fényben is fluoreszkál, ami lehetővé teszi a megbízható, károsodásmentes DNS izolálást. A gélből a megfelelő csíkot egy szikével vágjuk ki (ld. 4.8. ábra), amiből a DNS-t többféle módon is kinyerhetjük.

Az izolálás történhet **elektroelúcióval**, amikor a kivágott géldarabot egy olyan dializáló membránba helyezzük, ami impermeábilis egy bizonyos DNS molekula méretre, és permeábilis a folyadékokra. TE pufferbe áztatjuk a géldarabot tartalmazó dialízis zacskót, és elektromos teret hozunk létre a dialízis cső körül. Ugyanúgy, ahogy a gélelektroforézisnél is, a DNS elkezd vándorolni a pozitív elektróda felé. Ebben az esetben kivándorol a gélből, de nem juthat át a membránon. A dialízis zacskóból ezután csak ki kell pipettázni a már tiszta DNS-t tartalmazó folyadékot.

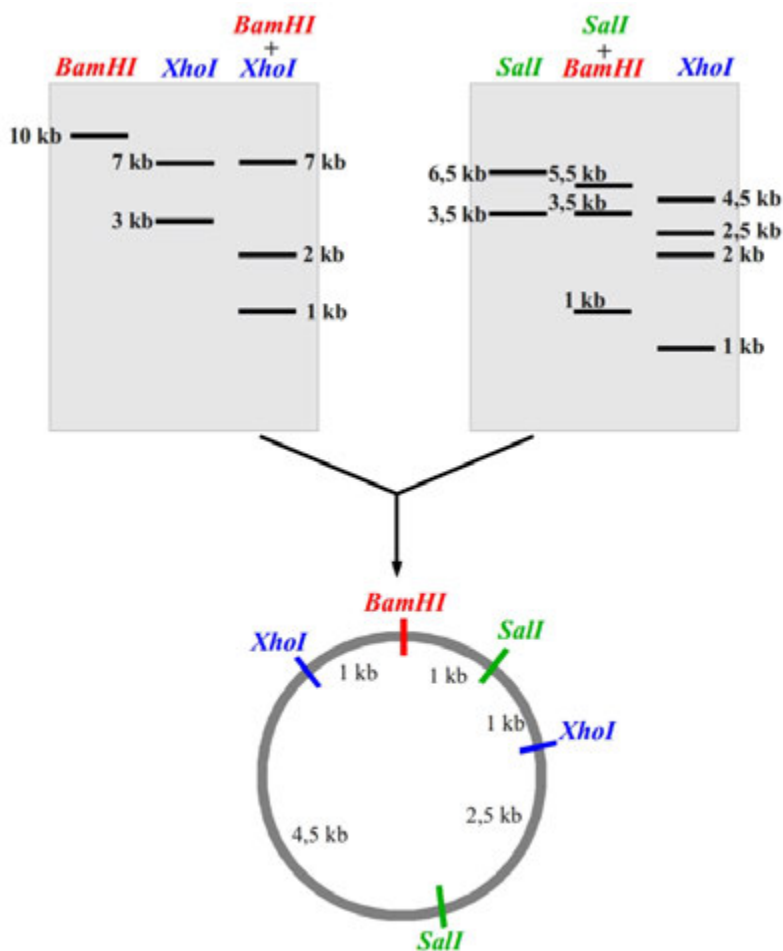


4.8. ábra: DNS agaróz gélből izolálása steril szikével

A DNS-t tovább lehet még tisztítani fenolos extrakcióval és etanolos kicsapással. Egy másik hagyományos módszer a géldarab -80°C -on történő **fagyasztása**. Ezen a hőmérsékleten az agaróz szerkezete már roncsolódik, így egy szűrőn keresztül, centrifugálással kinyerhető a DNS. Egy további kivonási eljárás az **alacsony hőmérsékleten megolvadó** (*low melting*) **agaróz** használata, mivel az oldatfázisba került DNS-t a miniprep módszernél ismertetett szilika membránhoz történő adszorpcióval könnyen tisztíthatjuk. A DNS (és RNS) gélből történő extrakciójára sokféle kit kapható a kereskedelemben.

4.1.2.5. Restriktációs térképezés

A különböző restriktációs endonukleázok eltérő felismerő szekvenciákkal rendelkeznek, így ugyanazon DNS molekulát különböző enzimekkel emésztve különböző restriktációs fragmentumok keletkeznek. A restriktációs enzimeket önmagukban vagy kombinálva alkalmazzuk, és a kapott DNS fragmentumokat gélelektroforézissel választjuk el. A fragmentumok mérete alapján össze tudjuk állítani az eredeti DNS molekula ún. **restriktációs térképét** (ld. 4.9. ábra).



4.9. ábra: Rekombináns DNS konstrukció térképezése restriktációs endonukleázokkal

Ez egy „kirakós játék”, ami azon alapszik, hogy a hasítóhelyek egymáshoz viszonyított helyzete dönti el, hogy melyik emésztésnél milyen méretű fragmentumok keletkeznek. Akár egy teljes genomot is fel lehet így térképezni (ld. 9.1.1. fejezet). Az ilyen mintázat minden embernél más és más – ezt DNS ujjlenyomatnak nevezik és több célra is fel tudják használni (ld. 4.1.3.2. fejezet).

4.1.3. Hibridizációs technikák és nukleinsav próbák

A molekuláris biológiában a hibridizációs technikákkal valamilyen ismeretlen biológiai mintából akarunk kimutatni egy adott DNS, RNS vagy fehérje szakaszt. Meg akarunk bizonyosodni annak jelenlétéről, mennyiségéről,

minőségéről. A hibridizációs technikáknak köszönhetően nincs szükség feltétlenül szekvenálásra, ha a genomban egy mutációt vagy strukturális változást akarunk azonosítani. A hibridizációs technikák molekuláris hátterét az a tény adja, hogy **a DNS két komplementer szála reverzibilis módon denaturálható**, s a renaturáció nem csak a két eredeti szál között megy végbe, hanem egy oligonukleotid próba jelenlétében az is anellál (hozzátapad) a vele komplementer szálhoz.

Ezekhez az eljárásokhoz tehát mindig szükséges valamilyen ismert szekvenciájú **DNS vagy RNSpróbára**. Próbának nevezzük a radioaktívan vagy más módon (leginkább fluoreszcenciával) jelölt (az esetek többségében oligonukleotid méretű) oligo- vagy polinukleotidot, ami **komplementer a vizsgálni kívánt nukleinsav egyik láncának szekvenciájával**. Az egyszálú próba csak egyszálú cél nukleinsav lánchoz képes kapcsolódni. A hibridizáció „szigorúsága” (*stringency*) attól függ, hogy milyen só koncentrációt és hőmérsékletet alkalmazunk. Magas anellációs hőmérsékleten és alacsony ionerő ún. szigorú (*stringent*) körülményeket jelent, ilyen esetben csak a próbával tökéletesen egyező nukleotid szálak fognak hibridizálni. Azonban magas ionerőn és alacsonyabb hőmérsékleten a csak részben hasonló szekvenciákhoz is hozzátapadhat a próba.

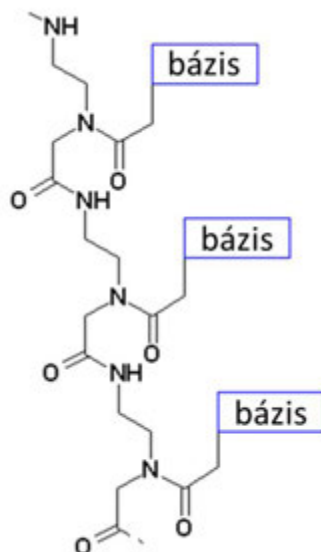
A fejezet első részében a próbakészítés és jelölés lehetőségeit vázoljuk fel, a további alfejezetekben a **Southern- és Northern-lenyomat (blot) módszerről**, az **in situ hibridizációról** és a **DNS chipokról** lesz szó. Néhány további hibridizációs módszert (pl. kolónia- vagy plakkk-hibridizáció) az azt alkalmazó módszert leíró fejezetben tárgyalunk (**ld. 8. fejezet**). (A hibridizációról még **ld. 18.1. animáció**.)

4.1.3.1. Nukleinsav próbák előállítása és jelölése

Korábban szinte kizárólag **izotópjelölést** használtak. A jelölt molekula a próbakészítés módszerétől függően (**ld. alább**) lehet g-ATP, a-ATP, a-CTP. Az izotópot tekintve használhatunk ^{32}P -t (felezési idő = 14 nap), ^{33}P -t ($t_{1/2}$ = 25 nap), ^{35}S -t ($t_{1/2}$ = 87 nap). A hibridizálás eredményét **radioaktív próba** esetén **autoradiográfiával** tesszük láthatóvá. Később megjelentek egyéb jelölési módok is, amelyek között a radioaktivitással azonos érzékenyséű **fluoreszcens próbák** terjedtek el leginkább. A fluoreszcens próbákról részletesebben a szekvenálás (**ld. 5. fejezet**) és a PCR fejezetben (**ld. 6.3.2. fejezet**) lesz szó. A nem-radiotív és nem-fluoreszcens próbák között megemlítjük a **digoxigenint**, egy erősen immunogén (ami azt jelenti, hogy igen nagy affinitással hozzákötődő antitestek állnak rendelkezésünkre), növényi eredetű szteroid, ami nukleotidokra konjugálható, s beépíthető nukleinsav próbákba (az előhívás az antitest jelölésével vagy a hozzá konjugált enzim által termékként előállított festékkel vagy kemiluminszcenciával történhet).

Az **oligonukleotid próbákat szilárdfázisú szintézissel** (szintetizátorokban), míg a hosszabb **DNS vagy RNS próbákat enzimatis úton** állítunk elő. Az egyik lehetőség a DNS molekula végjelölése. Az **5'-végjelölés**hez két enzimre van szükség. Alkalikus foszfátázzal eltávolítjuk a foszfát-csoportot, majd polinukleotid-kináz segítségével ATP-ből izotóp jelölt foszfát-csoport (g-foszfát jelölt!) építhető be a DNS 5'-végére. A **3' végjelölés**hez a DNS-polimerázok 3'→5' exonukleáz aktivitását használjuk ki. A „visszarágott” szálhoz a polimeráz aktivitás révén dNTP-ből izotóp vagy fluoreszcens jel épül be a 3'-vég közelébe. Ebben az esetben általában jelölt dCTP-t használnak (ennél a módszerrel az α -foszfátot kell jelölni!). Készíthetünk enzimatisan **random DNS próbákat** úgy is, hogy DNS-polimeráz segítségével dCTP-ből épül be a jel az újonnan szintetizált láncba. Ugyanezen az elven RNS-polimerázzal DNS-templárról jelölt **RNS-próbák** is előállíthatók **in vitro transzkripcióval**.

Megjegyzendő, hogy léteznek szintetikus ún. **peptid-nukleinsav (PNA) próbák** is, amelyekben a cukorfoszfát gerinc és a foszfodiészter kötések helyett poli N-(2-aminoetil)-glicin váz található és a bázisok ehhez metilén-karbonil csoporton keresztül kapcsolódnak (a bázisok távolsága megegyezik a foszfátgerinchez kapcsolódó bázisok távolságával, tehát a PNA is tud bázispárosodni egy komplementer láncsal). Egy PNA részletét a **4.10. ábrán** mutatjuk be



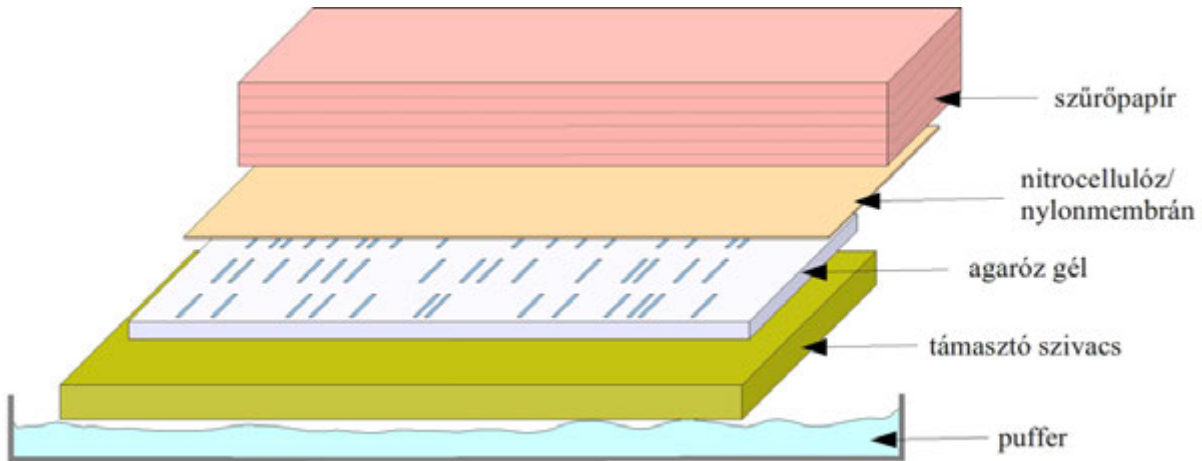
4.10. ábra: Egy peptid-nukleinsav részlete

4.1.3.2. Southern-blot (lenyomat) technika és az RFLP módszer

A Southern-blot (magyarul lenyomat) technika **speciális DNS szakaszok kimutatására** szolgál egy komplex DNS mintában. Például egy adott gén jelenlétét tudjuk vizsgálni különböző fajok teljes genomjában. Ezzel a technikával a gén méret vagy szerkezetbeli megváltozása is kimutatható. A „lenyomatra” azért van szükségünk, mivel a gélből a DNS diffúzióval az oldatba kerülne a hibridizálás során, másrészt a gél törékeny és nehezen kezelhető.

A Southern-blot kivitelezésekor először a nagyobb méretű DNS molekulát, akár egy teljes genomot, restrikciós endonukleázokkal kisebb darabokra hasítunk, majd agaróz gélben futtatjuk a mintát. A hibridizáció miatt egyszálú DNS molekulákra van szükségünk, ezért NaOH segítségével denaturáljuk a kettős szálát, majd K-acetáttal semlegesítjük. A gél elektrolit oldatba helyezzük (ami lehet egy szivacs, vagy a blottoló folyadékkal átitatott szűrőpapír), és a tetejére simítjuk a **nitrocellulóz- vagy nylonmembránt**. A membrán fölé további nedvzívó anyagot (pl. szűrőpapír) rétegezzük. A kapillárishatás miatt a DNS-molekulák a gélből a membránba vándorolnak. A membrán csak az egyszálú DNS-eket köti meg ugyanabban a pozícióban, ahogyan azok eredetileg a gélben elhelyezkedtek (ld. 4.11. ábra). A nitrocellulóz membrán esetében száraz hőkezeléssel (60-100°C-os vákuumban), a nylonmembrán esetében pedig UV-kezeléssel kovalensen hozzákötjük a DNS-t a membránhoz. Ezt követi a hibridizációs próbával történő inkubálás. Majd a membránról lemossuk a nem kötődött nukleinsav próbákat és röntgenfilmen autoradiográfiával, fluorométerrel vagy fotométerrel detektáljuk a jelet (a próba jelölésétől függően).

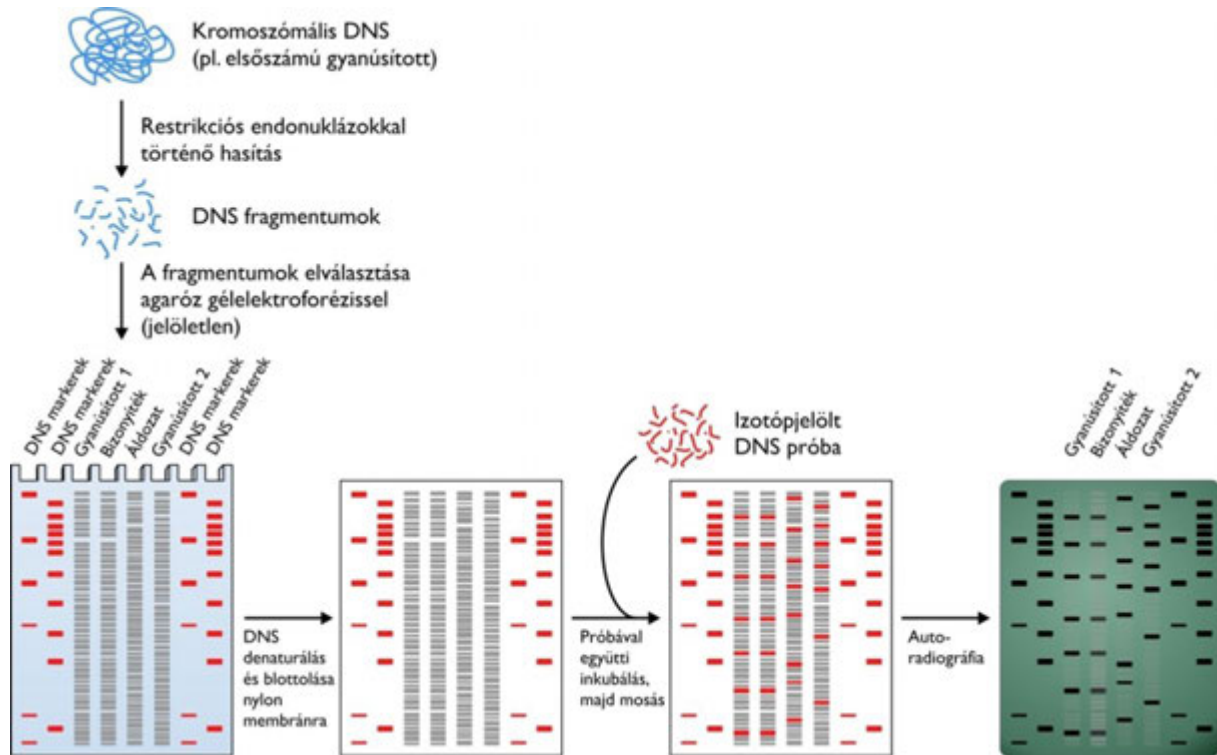
Radioaktív próba esetén előhíváskor a membránra egyszerűen egy röntgenfilmet helyezünk, és ahol a próba hozzákötődött a mintához, elfeketedik a film. Ezek a sötét sávok tökéletesen egybeesnek a hibridizált, azaz azonosítani kívánt DNS sávjával. Ma már elterjedtek az olyan eljárások is (az ún. **foszfo-image screen**), ahol a radioaktív jelet közvetlenül digitális jellel lehet alakítani. A radioaktivitás használata a sugárzás veszélyessége miatt ma már kevésbé használatos. A detektálás történhet még **kolorimetriával**, amikor a próbán valamilyen festékanyag található, amely enzim (pl. alkalikus foszfatáz, peroxidáz) jelenlétében oldhatatlan színes csapadékká alakul, és megfesti a membránt. Ez szabad szemmel is jól látható. Egy másik nem izotópos detektálási mód a **fluoreszcens kimutatás**, amikor biotinált próbát alkalmazunk, és a membránt fluoreszcens reagenssel jelölt avidinnel vagy sztreptavidinnel kezeljük, majd fluorimetriával detektáljuk. Manapság jóval elterjedtebb és közkedveltebb a **kemilumineszcencián alapuló előhívás** (ECL: *Enhanced Chemiluminescence*: megnövelt kemilumineszcencia). Ennél a módszernél a próbára kovalensen valamilyen enzimet, általában torna-peroxidázt konjugálnak. Ezután a membránt olyan anyaggal kezeljük, ami a torna-peroxidáz szubsztrátját, leggyakrabban luminolt tartalmaz. Az oxidáció során, hidrogén-peroxid jelenlétében, a luminol gerjesztett állapotba kerül, és egy fotont bocsát ki, amit megfelelő detektorral vagy röntgenfilmen azonosíthatunk. Fenol vázzal rendelkező molekulák jelenlétében a fénykibocsátás akár 1000-szeresére is növelhető. Ettől lesz „megnövelt” kemilumineszcencia az eljárás.



4.11. ábra: **Southern-blot (lenyomat) technika.** A gélben lévő DNS-t nitrocellulóz vagy nylon membránra transzferáljuk, így az elválasztott DNS mintázatának replikáját kapjuk. A transzfer általában egy szivacson keresztül történik, amit egy puffer tartályba helyezünk. A gél és a membránt a szivacsra tesszük, fölé pedig vastag rétegben szűrőpapírt helyezünk, hogy felszívja a puffert a szivacson, gélen és membránon keresztül. Ezáltal a DNS a gélből a membránba vándorol, és erősen immobilizálódik.

Southern-blot technika segítségével analizáljuk az **RFLP** (*Restriction Fragment Length Polymorphism*: **restrikciós fragmentum hossz polimorfizmus**) módszer alkalmazása során kapott DNS fragmentumokat is. Az RFLP módszer lényege, hogy homológ DNS molekulák variációira (SNP: *single nucleotide polymorphism*, kiejtve „sznip”) világíthat rá. Az eljárás során teljes genomból indulunk ki, melyet restrikciós endonukleázokkal kezelünk. Az különböző endonukleázok (**ld. 3.2.1.2. fejezet**) eltérő DNS szakaszokat ismernek fel. Az endonukleázos kezelés után kapott DNS fragmentumok száma az enzim által felismert szekvenciák számától, mérete pedig ezek távolságától függ. A restrikciós fragmentumokat ezt követően gélelektroforézissel szeparáljuk, és **Southern-blot hibridizációval detektáljuk**. RFLP abban az esetben jelenik meg, ha a vizsgált DNS szálon található inszerció, delécio vagy SNP egy meglévő restrikciós helyet tönkretesz vagy egy újat hoz létre. Ilyenkor gélelektroforézissel eltérő hosszúságú DNS fragmentumokat kapunk. Az RFLP analízis volt az első **DNS „ujjlenyomatot”** feltérképező, széles körűen elterjedt technika (**DNA profiling, DNS tipizálás**), amelyet bevezettek az igazságügyi orvostani és a bírósági gyakorlatba is. Az RFLP technika nagyon fontos eszköz volt a genomok fizikai feltérképezésében, különböző mutációk kimutatásában. Ezzel a módszerrel különböző betegségek, vagy az azokra való hajlam is kimutatható.

A **4.12. ábrán** egy bűntény elkövetőjének Southern-blot módszerrel előállított RFLP mintázat alapján történő azonosítását mutatjuk be. Ezt az alkalmazást Alec Jeffrey vezette be a gyakorlatba az 1980-as évek közepén (a szexuális-indíttatású gyilkosság elkövetőjét a bíróság a DNS ujjlenyomat alapján 1992-ben életfogytiglani börtönbüntetésre ítélte – ez volt a bírósági gyakorlatban az első eset, ahol géntechnológiai bizonyíték alapján született meg az ítélet).



4.12. ábra: **RFLP, mint DNS ujjlenyomat előállítás.** Az eljárás során az egyénekből vett DNS mintát restriktions endonukleázokkal emésztik, gélelektroforézissel elválasztják, majd Southern-blot hibridizálással állítják elő a DNS ujjlenyomatot, amellyel akár egy bűntény elkövetőjét is azonosítani lehet (ld. a bemutatott ábrát).

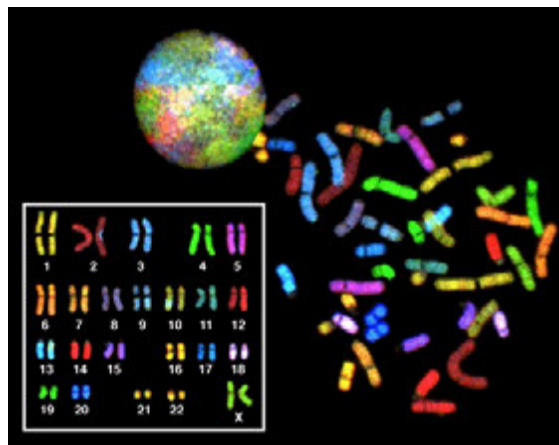
4.1.3.3. Northern-blot

A Southern-blottal rokon technika a Northern-blot, csak ebben az esetben DNS helyett **RNS molekulákat analizálunk**. Ez az eljárás jóval nagyobb elővigyázatosságot igényel, mivel az RNS rendkívül bomlékony. Ezért minden oldatot és eszközt RN-áz mentesíteni kell. Segítségével megállapítható egy adott sejt pillanatnyi génexpressziós állapota és annak változása például differenciálódás során vagy patológias esetekben. Az izolált mRNS mintát agaróz gélben választjuk el, de a gél formaldehidet vagy glioxált is tartalmaz, amelyek denaturálják az RNS-t, megszüntetik az esetleges másodlagos szerkezeti elemeit. Kisebb méretű RNS molekulák esetén (pl. siRNS, miRNS) ureát tartalmazó poliakrilamid gél alkalmaznak a tökéletesebb elválasztás érdekében. Az így elválasztott RNS-t tartalmazó gél nylon membránra blotolják a Southern-blothoz hasonló módon. A blotolás után az RNS membránhoz rögzítését UV keresztkötéssel végzik. A hibridizáció körülményeit a kimutatandó RNS, a használt próba (végjelölt oligonukleotid, random jelölt DNS vagy RNS próba) és a használt hibridizációs puffer együttesen határozzák meg. A kapott jelet a Southern blottnál leírtakhoz hasonlóan hívhatjuk elő. A használt radioaktív próba megfelelő körülmények között (alacsony sókoncentráció, magas hőmérséklet) a membránról lemosható. Az így „lefőzött” membrán RNS-eket detektáló új radioaktív próbával újrahibridizálható.

4.1.3.4. *In situ* hibridizáció, kariotipizálás

Az *in situ* hibridizációs eljárás során, sejten, szöveten belül tudunk komplementer nukleotid szakaszokat hibridizáltatni jelölt próbákkal. Az *in situ* DNS hibridizálás általában a kromoszóma struktúrák megfigyelésére használatos módszer, míg az *in situ* RNS hibridizálással az adott szövetben expresszált mRNS-eket, illetve miRNS-eket tudjuk kimutatni mikroszkopikus metszeteken. Az egyik legelterjedtebb változata ennek a technikának a **fluoreszcens *in situ* hibridizáció (FISH)**, amellyel a **genom egy-egy kisebb szakaszát lehet megjeleníteni**. A FISH technikák fejlődésével, nagy pontossággal és biztonsággal lehet citogenetikai diagnózisokat felállítani, tumoros sejteket azonosítani. A különböző méretű és típusú DNS szakaszok kimutatásához többféle próba áll rendelkezésre. A próbákat egy hapténnel (pl. biotin, digoxigenin, ösztadiol) jelölik, amelyhez különböző fluorokrómok (pl. FITC, rodamin) kapcsolódnak. Teljes hosszúságú kromoszómák kimutatásához az ún. *painting* próbák alkalmasak. Ha a próbákat eltérő színű fluorokrómmal jelölik, akkor a különböző kromoszómák más-más

színnel fognak világítani. Az ún. **multiplex-FISH** (M-FISH) technikával egyidejűleg jelölhető az összes kromoszóma (**ld. 4.13. ábra**), és ezáltal sokféle DNS átrendeződés közvetlenül kimutathatóvá válik.



4.13. ábra: **Multiplex-FISH**. A humán kromoszómákat különböző fluoreszcens festékekkel jelölték (forrás: JackeyCheng.com)

A kromoszómák telomer és centromer régióiban egy adott kromoszómára jellemző ismétlődő szakaszok találhatók, így ezekre a szekvenciákra találtak ki a repetitív próbákat. Ezek a próbák főként a kromoszómák számbeli eltéréseinek kimutatására alkalmasak. A génspecifikus és töréspont-specifikus próbákkal speciális, rövidebb szakaszok jelenlétét, hiányát vagy struktúráját lehet feltérképezni. A FISH-sel előállított jel mikroszkópban is megfigyelhető. Azonban a detektálás legérzékenyebb módja ún. CCD kamerákkal történik. Ez a kamera -20°C-on olyan fluoreszcens jeleket is érzékel, amelyre az emberi szem már nem képes. A FISH felbontási képessége 50 kbp-tól 2 Mbp-ig terjed, a DNS kondenzáltságától függően. (A FISH módszerről még **ld. 18.2. animáció.**)

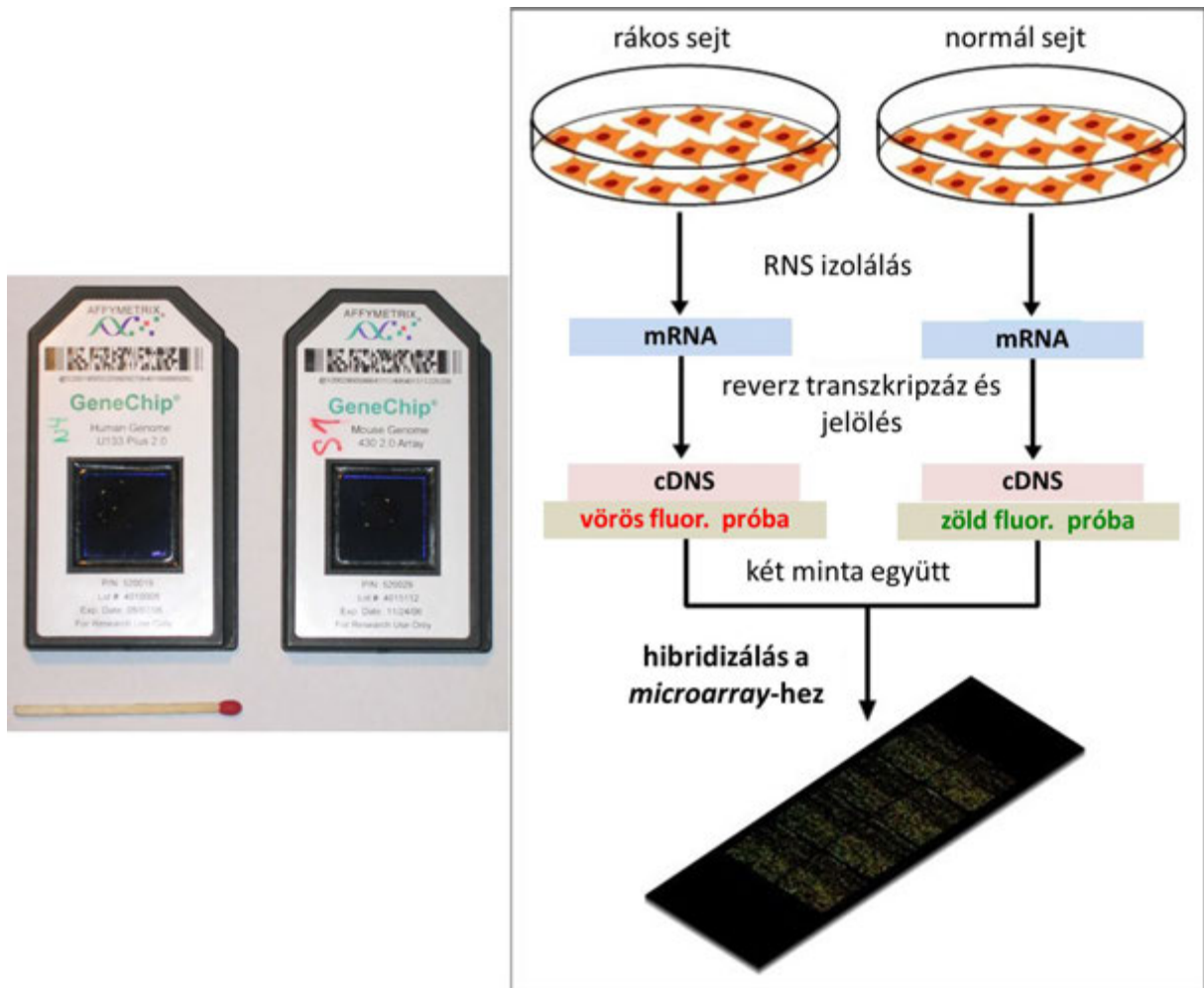
4.1.3.5. DNS-chip (*microarray*) technika

A hibridizációs technikák legújabb változata a **DNS-chip** vagy *microarray*. Egy **kis méretű (1-2 cm²) szilárd hordozó** (pl. szilikon, üveg) felületére szabályos elrendezésben **több 10000**, eltérő szekvenciájú **DNS próbát** rögzítenek kovalensen. Ez történhet pl. epoxi- vagy amino szilánnal, de gyakoribb a **fotolitográfias** eljárással történő *in situ* szintézis (oligonukleotid próbák esetén). A chipen 1 pont kb. 1 pmól mennyiségű DNS-t tartalmaz. A próbák 20-5000 nukleotid hosszúságúak, génekre vagy cDNS-ekre specifikus oligonukleotidok vagy *in vitro* szintetizált DNS-fragmentumok. Az eljárás lényege, hogy mikroszkóp segítségével detektáljuk azokat a próbákat a chipen, amelyekkel komplementer DNS vagy RNS van jelen a mintában. A klasszikus, az előző alfejezetekben ismertetett hibridizációs módszerekhez képest a chip technológiánál megfordult a próba és a minta viszonya: itt **a próbát**, míg az előző módszereknél a mintát **immobilizálják** (blottolással).

A chipkekhez használatos cDNS minták általában úgy készülnek, hogy a vizsgálni kívánt szövetet vagy sejtkultúrát feltárják, és kivonják az mRNS-t, amiből reverz transzkriptázzal cDNS-t szintetizálnak. Ez a mennyiségű cDNS azonban nagyon kevés lenne a további munkákhoz. Ezért PCR reakcióval, fluoreszcens módon jelölt primerek segítségével (pl. Cy5: vörös; Cy3: zöld) sokszorozítják a target DNS szálakat, majd ezt a cDNS-t tartalmazó oldatot inkubálják a chipen. A nem kötődött DNS darabokat lemossák, és a chipet általában egy speciális fluoreszcens előhívó készülékkel (*chip reader*) vizsgálják. A chip pontjai által kibocsátott fény színét és intenzitását detektálják. A kapott képet számítógépes programok segítségével elemezik; összehasonlítják a különböző minták hibridizációs értékeit, így egy adott kezelésre történő expressziós változásokhoz olyan számok rendelhetőek, melyek a relatív expressziót fejezik ki. A kiértékelés részleteire ebben a jegyzetben nem térünk ki.

A DNS *microarray* módszert széleskörűen alkalmazzák a funkcionális genomikában; génexpressziós változásokat, valamint az SNP-ktől kezdve a teljes genomot átfogó genetikai különbségeket lehet a módszerrel analizálni. További érdekességeket a módszer alkalmazásáról itt olvashatnak. (A DNS *microarray*-ról még **ld. 18.3. animáció.**)

Egy gén-chipet és egy expressziós összehasonlító vizsgálat sémáját a **4.14. ábra** mutatja be.



4.14. ábra: **Gén-chip** (az Affymetrix biotechnológiai cég terméke). A bemutatott két chip-en egyidejűleg több mint 40.000 humán (bal) és egér (jobb) transzkriptumot lehet analizálni. Az mRNS-eket a chip-en *in situ* fotolitográfiával szintetizált specifikus komplementer DNS próbák reprezentálják (bal) (forrás: Wikipedia, GNU Free Documentation License). Összehasonlító génexpressziós vizsgálat (rákos és egészséges sejt) hibridizációs chipen (jobb)

4.2. Rekombináns konstrukciók tervezése

A **rekombináns DNS** molekula egy olyan mesterséges konstrukció, amely két (vagy több) DNS molekulából állítható össze. Az egyik maga a vizsgálni kívánt DNS szekvencia, a másik a hordozó DNS, amely lehetővé teszi a replikációt (és egyes esetekben a génexpressziót is). Tehát ha egy idegen DNS darabot (**inszert**) egy replikálódni képes DNS szekvenciával (**vektor**) kapcsolunk össze kovalensen, és bejuttatjuk egy, a vektornak megfelelő gazdasejtbe, akkor az idegen szekvencia is le fog másolódni. A rekombináns DNS-ek nem csak a vektor replikonjának megfelelő gazdasejtbe, de akár teljes organizmusokba is bejuttathatók. Az utóbbi esetben genetikailag módosított élőlényt (GMO) hozunk létre (**ld. 15. fejezet**).

Ebben az alfejezetben a molekuláris klónozáshoz szükséges rekombináns DNS elkészítésének fő lépéseiről, a konstrukciók tervezéséről lesz szó. Honnan származhat, és hogyan állítjuk elő a vizsgálandó DNS fragmentumot, az inszertet és a hordozó DNS-t, a vektort. Részletezzük az inszert és a vektor endonukleázokkal történő hasítását; az inszert és a vektor ligálását, a rekombináns DNS-t tartalmazó gazdasejtek szelektálási lehetőségeit és végül a rekombináns DNS szekvenciájának ellenőrzési módszereit.

4.2.1. Az inszert előállítása

A rekombináns DNS tervezésének első lépése, hogy kiválasztjuk a vizsgálandó DNS szakaszt. Amennyiben a vizsgálandó DNS szekvenciája nem ismert, mert például egy korábban nem kutatott élőlény specifikus funkciójáért felelős génjét keressük, valószínűleg genomiális vagy cDNS könyvtárat kell létrehozunk (**ld. 8. fejezet**). Bár manapság technikailag jelentősen kisebb kihívás egy ilyen munka, mint egy évtizeddel ezelőtt, komoly idő és pénz ráfordítással kell számolnunk. Jelentősen egyszerűbb a helyzet, ha olyan élőlény DNS-éről van szó, amely genomját már szekvenálták (**ld. 9. fejezet**), tehát az adott DNS szakaszok szekvenciája ismert. Ekkor a kiindulási szekvencia információt DNS adatbázisokban találjuk meg (**ld. 10.1.2. fejezet**). Három lehetőség áll előttünk. Az első, hogy mi előállíthatjuk az inszertet, amennyiben rendelkezésünkre áll genomiális DNS vagy RNS minta a megfelelő szervezetből vagy szövetből. A második lehetőség, hogy **szintetizáltatjuk** az inszertet, ami mára, bár még mindig elég drága szolgáltatás, az előnyei miatt (minimális munka és időráfordítás) egyre elterjedtebb. A harmadik eset a legolcsóbb változat, amikor egy klón adatbázisból szerezzük be a bennünket érdeklő inszertet tartalmazó kész rekombináns DNS-t (ilyen „klónmegosztó” szolgáltatás például az **Addgene**, amelybe bárki beteheti az általa elkészített konstrukciót, s azt bárki megkaphatja, csak a szállítási költséget kell kifizetnie). Ebben az esetben a rekombináns DNS-ből restrikciós enzimekkel közvetlenül kivághatjuk, vagy PCR segítségével (**ld. 6.4.1. fejezet**) „szubklónozzhatjuk” az inszertet.

Az inszert előállítása attól is függ, hogy egy **gént**, **gén fragmentumot** vagy **cDNS-t** akarunk klónozni. A **cDNS** (*complementary*, azaz **komplementer DNS**) az mRNS-ről reverz transzkriptáz enzimmel készített másolat. Ez utóbbira akkor van szükségünk, ha vagy cDNS könyvtárat akarunk előállítani (**ld. 8.3. fejezet**) vagy expressziós konstrukciót akarunk előállítani, azaz a rekombináns fehérje a vizsgálatunk tárgya (**ld. 12. és 13. fejezet**). Fontos emlékeztetnünk rá, hogy az eukariótában a gének általában intronokat is tartalmaznak, azaz a gén által kódolt mRNS szekvencia nem egyezik meg a gén DNS szekvenciájával, ami tényre figyelembe kell vennünk a DNS konstrukciók tervezésekor.

4.2.1.1. Genomiális DNS inszert

Ha a génnel akarunk dolgozni, akkor vagy az adott faj genomiális könyvtárára van szükségünk vagy genomiális DNS-t kell izolálnunk. A genomiális könyvtárból kikereshetjük a bennünket érdeklő klónt valamilyen hibridizációs eljárással (kolónia- vagy plak- hibridizáció; **ld. 8.2. fejezet**), de ma már sokkal egyszerűbb két specifikus primer segítségével a specifikus DNS-t **génkönyvtárból** vagy az általunk izolált **genomiális DNS-ből** származó templátról **PCR-rel amplifikálni** (**ld. 6. fejezet**). A PCR primerek 5' végébe beépíthetünk restrikciós endonukleáz felismerőhelyeket, így a klónozás a továbbiakban egy lépésben megvalósítható lesz (**ld. 6.4. ábra a 6.4.1. fejezetben**).

Ahogy korábban már bemutattuk, a legcélravezetőbb klónozási stratégia a **direkcionális klónozás** (**ld. 2.2. fejezet**). A PCR-t követően agaróz gélben futtatjuk a mintát, és izoláljuk a DNS terméket, ha az a DNS létrához viszonyítva a megfelelő méretű. Az izolálást követően olyan restrikciós endonukleázokkal hasítjuk az inszertet, amelyek felismerő helyét a primerekkel beépítettük az amplifikált DNS-be. Az inszertünk a hasítást követően ragadós végű lesz, azaz egy komplementer szekvenciájú, ragadós végű DNS darabbal össze tud tapadni (lehetséges tompa véget eredményező restrikciós enzimet is használni, de a tompa-végű DNS-ek ligálása, mint az korábban már említettük, nehezebb feladat). Végül újból megtisztítjuk a DNS- egy szilika oszlopon, hogy elválasszuk az enzimektől és apró nukleotid törmeléktől, és ezzel elkészítettük a ligálásra kész inszertet.

4.2.1.2. mRNS izolálás és cDNS inszert

Amennyiben a gén által kódolt fehérje a vizsgálataink tárgya, akkor az esetek többségében csak **cDNS szintézis** vagy esetleg szintetikus DNS jöhet szóba. Kivétel, ha be tudunk szerezni (vagy mi készítünk) egy cDNS könyvtárat, amiből vagy expressziós szkríneléssel keressük meg a szekvenciánkat tartalmazó klónt (**ld. 8.3. fejezet**) vagy PCR-rel állítjuk elő az inszertet.

Ha cDNS-t akarunk előállítani, tudnunk kell, hogy milyen sejtekben, szövetekben expresszálódik nagyobb mennyiségben az adott mRNS. Kiválasztjuk a megfelelő szövetet vagy sejt kultúrát, amiből először **totál RNS-t izolálunk**. Az RNS izolálásra számtalan módszer és kit áll rendelkezésre, amelyek közül itt csak egy klasszikus módszert említünk. Ennél a módszernél az RNS izolálás során egy lépésben történik a sejt feltárás és a szerves-vizes fázis alapú extrakció guanidinium izotiocianát-fenol-kloroform oldattal. A vizes fázis tartalmazza a totál RNS-t, amit DN-ázzal kell kezelni a kromoszomális DNS szennyezés eltávolítása érdekében. Ha cDNS könyvtárkészítés

a célunk, a totál RNS-ből **mRNS**-t kell izolálni, ami nehéz feladatnak tűnik, mivel az mRNS a sejt teljes RNS készletének csak néhány százaléka. Azonban, legalábbis az eukarióta mRNS-eknek van egy tulajdonsága, a **poli-A farok**, amelyet kihasználhatunk egy egylépéses **affinitás kromatográfiás** tisztítási eljáráshoz. Oligodezoxitimidint (oligo-dT) kötünk kovalensen egy szénhidrát mátrixhoz (pl. **oligo-dT cellulóz**), s a teljes RNS készletből csak az mRNS-ek kötődnek az oszlophoz. A cDNS-t **reverz transzkriptáz** enzimmel (**ld. 3.1.5. fejezet**) állítjuk elő, oligo-dT primert alkalmazva.

Amennyiben egy expressziós konstrukció elkészítése a cél, közvetlenül a totál RNS-ről is írhatunk át cDNS-t amelyet PCR-rel sokszorozíthatunk (**RT-PCR, ld. 6.4.4. fejezet**). A PCR-hez ebben az esetben is olyan primereket célszerű tervezni, amelyek a tapadó régió kívül tartalmaznak egy restrikciós endonukleáz felismerőhelyet az 5' túlnyúló végükön.

4.2.2. A vektor előkészítése

Nagyon sokféle klónozó és expressziós vektor áll rendelkezésünkre. Fontos tudnunk, hogy „csak” klónozni akarunk *E. coli* sejtekben vagy más gazdasejttel is dolgozni akarunk, mivel a vektorok replikonjai specifikusak a gazdasejtekre. Amennyiben a rekombináns DNS konstrukciónkat több gazdába is be akarjuk juttatni, akkor bináris (*shuttle*: „ingázó”) vektort kell használnunk, ami két gazdában is képes replikálódni. Figyelembe kell venni a rendelkezésre álló klónozóhelyeket, a szelekciós markereket, esetleg további szabályozó elemek jelenlétét vagy hiányát az alkalmazni kívánt vektorban. A céljainknak legmegfelelőbb vektort kiválasztjuk (**ld. 7. fejezet**), a vektor DNS-t felszaporítjuk és előkészítjük az inszert befogadására.

Az üres, vagyis inszertet még nem tartalmazó vektorból nagyobb mennyiségre van szükségünk. Ehhez *E. coli* baktériumokba transzformálunk egy kis mennyiségű DNS-t, majd miniprepet készítünk (**ld. 4.1.1.2.).** A vektor DNS-t ezután az inszert készítésekor használt restrikciós endonukleázokkal hasítjuk (pl. itt is BamHI és XhoI), így az is ragadós végű lesz, komplementer az inszert DNS 5'- és 3'-végével. Ezzel a vektor is készen áll a már említett **direkcionális klónozásra**. Ha csak **egyfajta restrikciós endonukleázzal** hasítunk, fennáll a veszélye a vektor önzáródásának vagy konkatamerek képződésének. Ezért **a hasított vektort alkalikus foszfáttal (CIAP) kezeljük**, ami eltávolítja a linearizált vektor DNS 5'-végéről a foszfát csoportot (ezzel **megakadályozhatjuk a vektor-önzáródást**). A vektort ezt követően gélből izoláljuk (**ld. 4.1.2.4. fejezet**).

4.2.3. Ligálás

A ligálás az a folyamat, melynek során két kompatibilis, ragadós vagy tompa végű DNS molekulát összekapcsolunk. Ezzel a lépéssel építjük be a vizsgálni kívánt DNS szekvenciát, az inszertet a vektorba. A ligálás körülményeit gondosan kell megválasztani, hogy sikeres legyen. Először meg kell bizonyosodnunk arról, milyen koncentrációjú az inszert és vektor DNS oldatunk.

Elvileg megbecsülhető, hogy milyen valószínűséggel jönnek létre rekombináns molekulák, ha egyféle enzimmel nyitottuk fel a cirkuláris vektort és ugyanazok a ragadós végek találhatók az inszertünk két végén. Az ún. effektív végkoncentráció kiszámolható:

$$j = 51,1 \times [\text{Ms}]^{1/2}$$

(Ms: molekulatömeg; tehát pl. egy linearizált pBR322 fág *j* értéke 32 µg/ml). Ha a vektor koncentrációja jóval kisebb, mint az effektív koncentráció ($c \ll j$), akkor a cirkularizáció kerül túlsúlyba (azaz üres vektorok keletkeznek), ellenben ha magasabb ($c \gg j$), akkor elsősorban konkatamerek jönnek létre. A teoretikusan legtöbb rekombináns DNS kb. 1:2,5 vektor:inszert **koncentrációarány**nál keletkezik, tehát ebben az esetben kb. 2-3-szoros inszert felesleg jelenlétében érdemes ligálni. Ha a vektor önzáródást kivédtük a **4.2.2. fejezetben** ismertetett módon, az inszert felesleget 3-4-szerese emelhetjük.

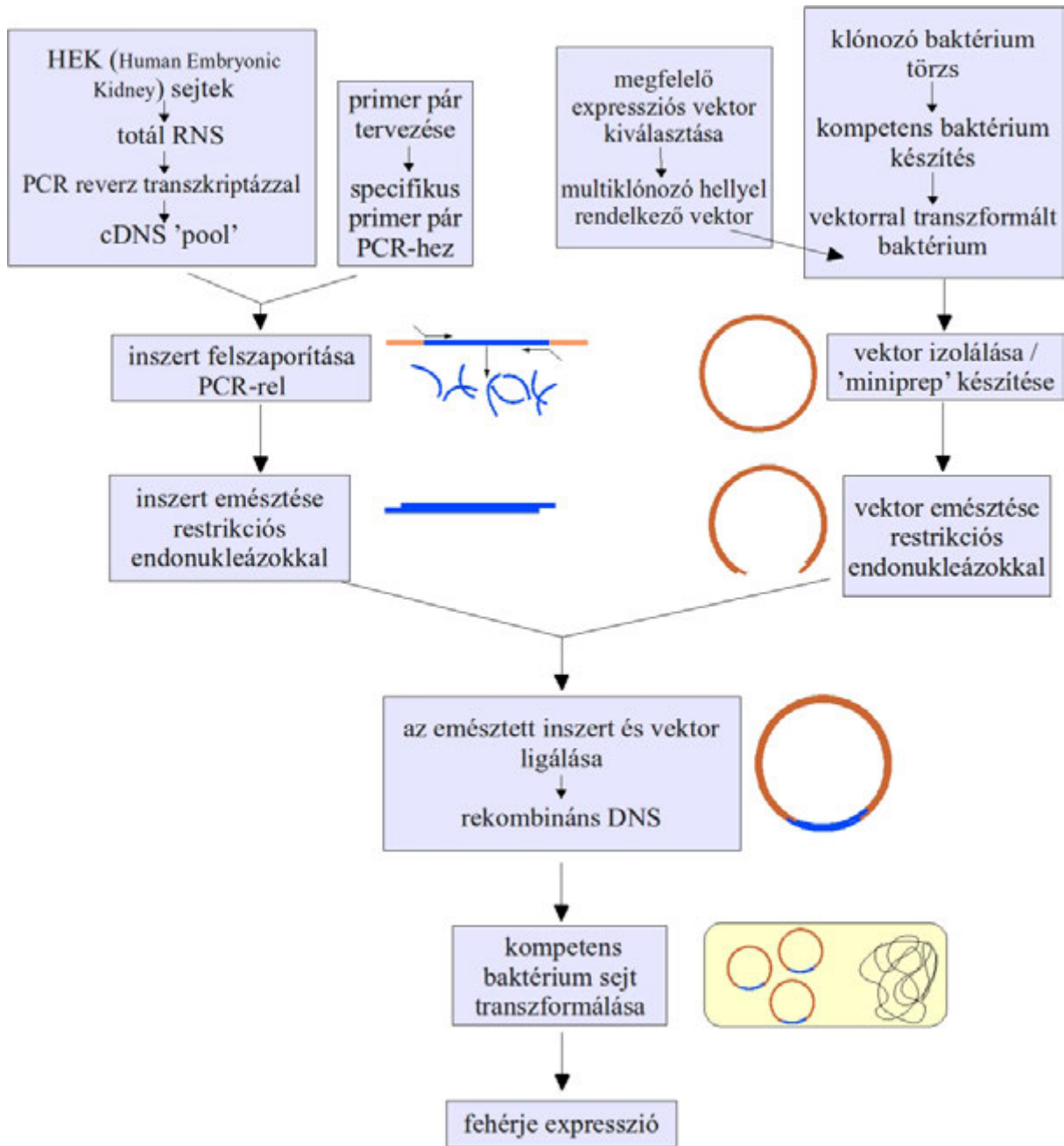
A reakció elegybe a két DNS-en kívül a DNS-ligáz enzimet és puffert kell adni. Enzimként leggyakrabban a T4 bakteriofágból származó DNS-ligázt használjuk, amely a foszfodiészter kötés létrehozásához szükséges energiát ATP-ből nyeri (az ATP-t általában az enzimmel együtt vásárolt puffer tartalmazza). Előnye ennek az enzimnek az *E. coli* saját DNS-ligázával szemben, hogy tompa végű DNS-eket is össze tud ligálni (**ld. 3.3. fejezet**). Az enzimreakció lezajlása után *E. coli* sejtekbe kell juttatnunk a cirkularizálódott rekombináns DNS-t, **transzformálással**. Ezt követően agar táptalajt tartalmazó Petri-csészékre szélesztjük a transzformált baktérium kultúrát (kb. 200 µl).

Mivel a vektorok mindig tartalmaznak valamilyen **antibiotikum rezisztencia** gént, ezért az agarba antibiotikumot keverve szelektálhatóak azok a baktérium sejtek, amelyek felvették a plazmidot. Mivel az üresen összezáródott, inszertet nem tartalmazó, de az antibiotikum rezisztencia gént hordozó plazmidot tartalmazó baktérium-telepek is megjelenhetnek a lemezen, a rekombináns DNS-ek kiválasztásához egy másodlagos szelekcióra is szükségünk lesz, amire példa a később ismertetendő ún. „**kék-fehér**” szelekció (**ld. 7.2.3.1. fejezet**).

A rekombináns klónokból néhányat egy steril pipettahegy vagy fogpiszkáló segítségével 3-5 ml-es térfogatú tápvesbe oltunk. Ezekből miniprepet készítünk. A **tisztított rekombináns DNS-t** nagyon fontos újból **ellenőriznünk**, hogy valóban tartalmazza-e a megfelelő inszertet. Ezért azzal a két endonukleázzal kezeljük, amelyekkel a klónozás elején a vektor felnyitottuk. Ha agaróz gélben futtatjuk az emésztést követően a DNS mintát, akkor sikeres konstrukció esetén két DNS sávot kell látnunk. Az egyik a linearizált üres vektor, a másik pedig az inszert DNS a megfelelő méretnél. A gélen mintát vihetünk fel negatív kontrollként az üres vektor két adott (itt BamHI és XhoI) restriktációs enzimmel emésztett elegyből, pozitív kontrollként pedig az izolált inszertből is.

A **4.15. ábrán** példaként egy expressziós vektor konstrukció elkészítésnek sémáját, fő lépéseit mutatjuk be:

- a fentieknek megfelelően előállítjuk az inszertet (ebben az esetben cDNS-t készítünk reverz transzkripció, majd PCR segítségével; további részletek a **6.4.4. és a 8.3.1. fejezetekben**)
- előkészítjük a vektort (ebben az esetben egy expressziós vektort használunk)
- *in vitro* ligáljuk a vektort és inszertet
- a rekombináns DNS-t gazdasejtbe juttatjuk transzformálással (**ld. 4.3.1. fejezet**)
- a rekombináns DNS-t felvett sejteket felszaporítjuk és a rekombináns fehérjét expresszáljuk (**ld. 12.2. fejezet**).

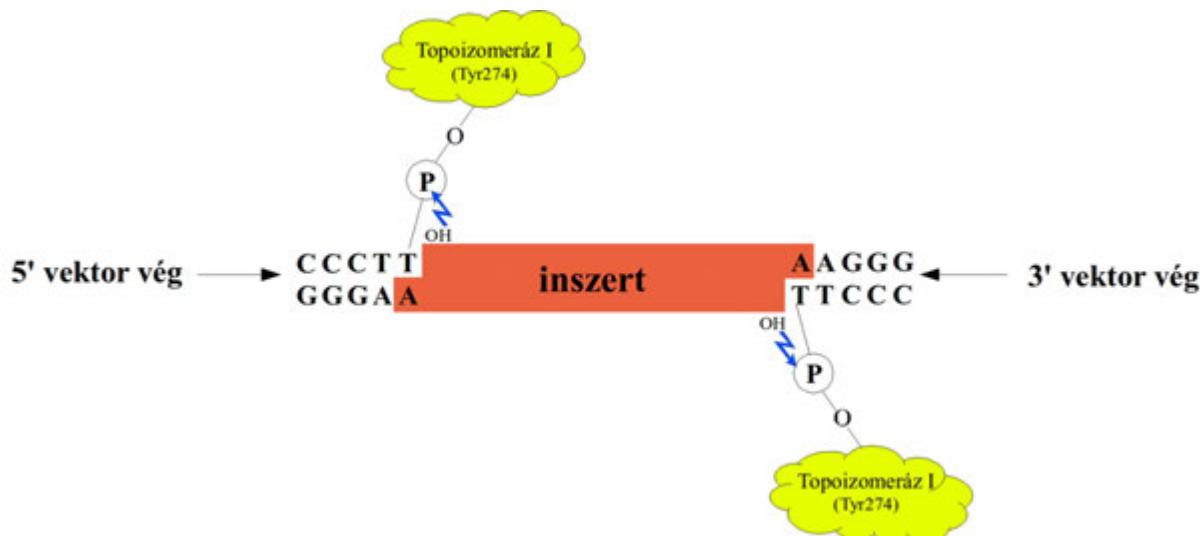


4.15. ábra: **Egy expressziós konstrukció elkészítésének sémája.** A példa egy expressziós vektorba ligált cDNS inszert előállításának lépéseit mutatja: inszert elkészítése, vektor előkészítése, ligálás, transzformálás, sejtnövesztés és fehérje expresszió

4.2.4. TA/Topo-klónozás

A TA- vagy más néven Topo-klónozás egy olyan klónozási technológia, melynek során **sem DNS-ligázra, sem restrikciós endonukleázokra nincs szükség.** A TA-klónozás alapja, hogy az inszertet olyan DNS-polimerázzal amplifikáljuk, aminek hiányzik az 5'→3' hibajavító nukleáz aktivitása, viszont van templat független polimeráz aktivitása (terminális-transzferáz aktivitása) és meg tudja hosszabbítani egy nukleotiddal a DNS szálak 3'-végét. Ilyen enzim a Taq-polimeráz, amely preferenciálisan egy adenin nukleotidot ad hozzá a láncvégekhez. Szükségünk van még egy speciális vektorra is. A TA-klónozó vektorok linearizált, 3'-túlnyúló timin ragadós végeket tartalmaznak, amelyek bázispárosodnak az inszerten lévő 3'-túlnyúló adeninnal. A linearizált vektor 3'-végéhez a vakcinia vírus DNS-topoizomeráz I enzimet kovalensen hozzákapsolták, amely elősegíteni a két DNS szál

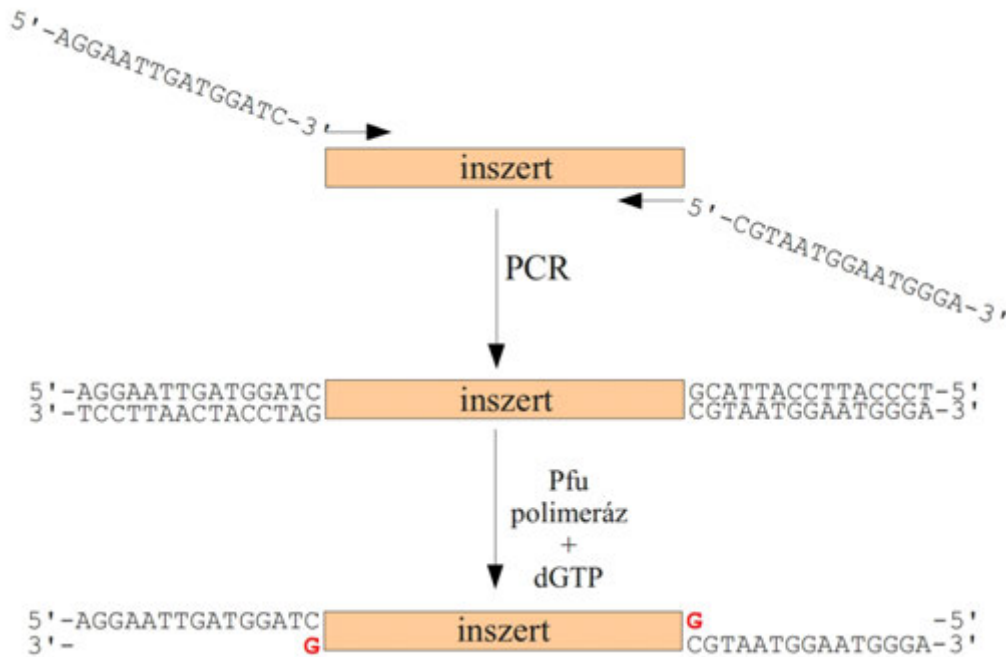
ligálódását (ld. 4.16. ábra). A konstrukciókat a transzformálást követően itt is a kék-fehér szelekció elvén lehet szelektálni (ld. 7.2.3.1. fejezet).



4.16. ábra: **Az ún. TA-klónozás sémája.** A vakcinia vírus DNS-topoizomeráz enzime egyaránt funkcionál restrikciós enzimként és ligázként. Egy speciális szekvenciát ismer fel a dupla szálú DNS-en: 5'-CCCTT-3'. Az enzim a replikáció során ennél a szekvenciánál hasítja a DNS foszfodiészter kötését az egyik szálon egy túlnyúló timint hagyva, majd a szuperhelikális DNS feszültsége lecsökken, és az enzim újra ligálja a láncot a 3' timin foszfát csoportjánál. Az elvágott szál 3'-végének foszfátja és a topoizomeráz tirozinje között kovalens kötés alakul ki. Ha a vektort és az inszertet összekeverjük, akkor a PCR termék 5'-végi OH- csoportja támadja a foszfortirozil kötést, elősegítve a ligáz reakciót és a topoizomeráz leválását. A kereskedelembe kapható vektorokat linearizálva lehet megvásárolni, topoizomerázzal kötve. Így a vektort 5 perc alatt ligálja a megfelelő kompatibilis végű PCR termékkel.

4.2.5. Ligálás-független klónozás

Ennél a módszernél nincs szükség DNS-ligázra és restrikciós enzimre. A módszer egyes DNS-polimerázok (pl. T4, Pfu) 3'→5' exonukleáz aktivitására épül, amelynek köszönhetően hosszú, ragadós végek keletkeznek az inszerten és a linearizált vektoron egyaránt. Az inszertet PCR reakcióval amplifikáljuk. A primereket úgy kell megtervezni, hogy legalább 15 nukleotid hosszúságú túlnyúló véggel rendelkezzenek, melyek majd kompatibilisek lesznek a vektor túlnyúló végeivel. Az inszertet tartalmazó vektor ligálás nélkül is transzformálható kompetens sejtekbe, ahol az endogén DNS-ligáz hozza létre a foszfodiészter kötéseket. A primer túlnyúló végeiből hiányzik valamelyik nukleotid, pl. a dCMP. Ebben az esetben egy 3'→5' exonukleáz aktivitású polimeráz dGTP jelenlétében a 3'-végtől kezdve eltávolítja a nukleotidokat, amíg nem találkozik az első citozinnal a templáton, ahova beépítheti a guanin nukleotidot (ld. 4.17. ábra).



4.17. ábra: A ligálás-független klónozáshoz előkészített inszert

A vektort is hasonló módon kell előkészíteni, de előbb hasítani kell egy restrikciós endonukleázzal (pl. tompa véget kialakító enzimmel), amelynek környezetében ugyanolyan távolságra helyezkedik el az első guanin nukleotid, mint az inszert esetében a citozin nukleotid. A vektort is kezeljük a 3'→5' exonukleáz aktivitású polimerázzal dCTP jelenlétében, így komplementer ragadós véget fogunk kapni.

4.2.6. Restrikciós hasítás nélküli ligálás rekombinációval (attB, attP)

Ez a módszer a helyspecifikus rekombináció elvén alapszik, ahol a DNS szakasz kivágódását vagy beépülését rövid, specifikus bázissorrendű szakaszok teszik lehetővé. Ilyen például a λ -fág attB és baktérium attP (*attachment site*) régiói. Az inszertbe a primerrel visszük be ezeket a szakaszokat. Ezt a reakciót a Invitrogene cég BP-klonáz enzime katalizálja (a klónozó rendszert **Gateway**-nek nevezték el). A Cre-Lox rendszerrel könnyedén ki lehet alakítani deléciót, inserciót, inverziót és transzlokációt sejten belül attól függően, hogy a Lox helyek hogyan helyezkednek el. A P1 bakteriofág Cre-rekombinááz enzime elhasítja a DNS-t, ha LoxP (34 nt) felismerési szekvenciát talál, és rekombinációt katalizál. Ez a módszer főleg transzgenikus élőlények létrehozásához használatos (ld. 14.3.2. fejezet). A Gateway klónozásról további részleteket a 14.3.2. fejezetben olvashatnak.

4.3. Génbeviteli eljárások

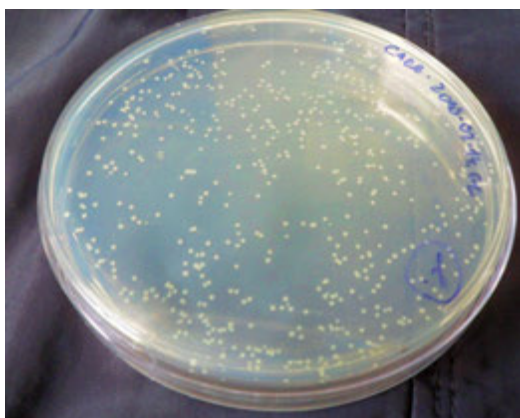
A plazmid DNS baktérium, élesztő vagy egyéb eukarióta (élesztő, növény, állat) sejtbe való bejuttatására számos lehetőség áll rendelkezésünkre. Ebben a fejezetben a génbeviteli módszerek közül a baktérium sejtek **transzformálását**, az eukarióta sejtek **transzfektálását**, az **infekciót**, az **elektroporációt** és a **génpuska** használatát mutatjuk be.

4.3.1. Transzformálás

A baktérium sejtek képesek a környezetből DNS molekulák (pl. plazmid DNS) felvételére. A jelenséget transzformációnak nevezzük, s ez a horizontális géntranszfer egyik formája (ld. 1.1.2. fejezet). A géntechnológiában a transzformáció a rekombináns DNS baktérium sejtbe juttatásának egyik legáltalánosabban használt módja. A sikeres génbevitelhez először ún. **kompetens sejteket** kell létrehoznunk (az a sejt kompetens, amelyik képes a

környezetéből DNS-t felvenni). Mivel az *E. coli* sejtek viszonylag könnyen kompetenssé tehetők, ez a tény is hozzájárult, hogy a géntechnológiában a legfontosabb gazdasejt ez a baktérium faj lett.

Kompetens sejteket legegyszerűbben úgy hozunk létre, hogy a sejteket CaCl_2 -ot tartalmazó oldatban inkubáljuk (táptalaj összetevők nélkül). Ez elősegíti a plazmid DNS kötődését a baktériumfal lipopoliszacharidjához. A pozitív töltésű Ca^{2+} -ok vonzzák a negatív töltésű DNS-t. Hő sok hatására pedig a sejtmembrán fluidabb lesz, és a DNS be tud jutni a citoplazmába (a pontos mechanizmust nem ismerjük). Ilyenkor a sejteket 4°C -ról 42°C -ra melegítjük 30-60 sec-ra (**hő sok**). A sejteket ezután tápanyagdús közegben hagyjuk regenerálódni, majd szelektív táptalajra szélesztjük őket. A kompetens sejt transzformációs hatékonyságát számadattal úgy fejezzük ki, hogy kis mennyiségű (néhány ng) plazmidot transzformálunk a gazdasejtbe, megszámoljuk a kolóniák számát az agar lemezen, s azt 1 μg plazmid DNS-re vonatkoztatjuk. 10^7 - 10^8 telep/ μg plazmid DNS esetén **magas kompetenciáról** beszélhetünk. Tapasztalati alapon kidolgoztak olyan protokollokat is (különböző fémion keverékkel kezelve az *E. coli* sejteket), amelyekkel még ennél is erősebb kompetencia érhető el (max. 10^{10} telep/ μg plazmid DNS). A **4.18. ábrán** baktérium telepek látszanak agar lemezen.



4.18. ábra: *E. coli* telepek agar lemezen

4.3.2. Transzfektálás

A transzfektáció alapja ugyanaz, mint a transzformálásé, csak itt a célsejtek emlős sejtekből származnak. Az elsőként alkalmazott és legelterjedtebb módszer, a **kémiai** úton történő **transzfektálás**, amikor valamilyen kationos hordozóhoz kötődik a negatív töltésű DNS, és ezt a komplexet juttatjuk a sejtbe. A kémiai módszerek hátránya, hogy sok faktortól függ a transzfektációs hatékonyság, mint például a DNS nukleinsav összetétele, az oldat pH-ja és a sejtmembrán aktuális állapota. Leggyakrabban olyan **kationos lipidet** használunk (**pl. lipofektamin**), amely liposzómákat képez, magába foglalva a DNS-t. A liposzóma fuzionál a sejtmembránnal és ez által juttatja be a DNS-t a sejt citoplazmájába.

Igen olcsó és egyszerű módja a kémiai transzfektálásnak, amikor **kalcium-foszfátot** alkalmazunk. Foszfát ionokat tartalmazó pufferhez (pl. HEPES) kalcium-klorid oldatot adunk a transzfektálni kívánt DNS-sel. Ez által egy precipitátum képződik a pozitív töltésű kalciumból és a negatív töltésű foszfátból, ami megköti a felületén a DNS-t. Ezt a szuszpenziót adjuk végül a sejtekhez, amelyek felveszik a precipitátumot a DNS-sel együtt. Egy másik módszer a kationos polimerekkel (pl. DEAE-dextrán, polietilénimin) történő transzfektálás. A negatív töltésű DNS a polikationhoz kötődik, majd a sejt endocitózissal veszi fel a komplexet.

A jóval hatékonyabb, és újabban használatos fizikai transzfektálás módszerei közé tartozik még az optikai transzfektálás és az elektroporáció. Az optikai transzfektálás esetében lézer segítségével kisméretű ($\sim 1 \mu\text{m}$) lyukakat képezünk a sejt plazmamembránján. A médium és a citoplazma közötti ozmotikus különbség miatt a DNS beáramlik a sejtbe.

4.3.3. Elektroporáció

Az elektroporáció az egyik leghatékonyabb módszer a cirkuláris DNS sejtbe juttatáshoz.

Ez a technika a sejt lokális transzmembrán- feszültségétől függ. Az elektroporációt egy **elektroporátor** nevű berendezéssel végezzük, amely egy elektromágneses teret hoz létre a sejteket tartalmazó oldatban.

A sejtszuspenziót, a bejuttatni kívánt DNS-sel egy olyan küvetába pipetázzuk, aminek a két oldalán alumínium elektródák találhatók, és nagyfeszültséget hozunk létre (10,000-100,000 V/cm néhány μ sec-msec-ig, a sejt típusától és méretétől függően). Nagyon fontos a megfelelő sókoncentráció. Ha túl magas, akkor elektromos kisülés jöhet létre, ami nagyarányú sejtpusztulást okoz. A nagy térerő ideiglenesen lyukakat képez a foszfolipid kettős rétegen és a sejtfalon. A sejten kívüli nagy plazmid koncentráció miatt hirtelen megindul a lyukakon keresztül a DNS beáramlása a sejt belsejébe. Majd a membrán töltöttsége megszűnik, a foszfolipid kettős réteg összeáll ismét, így gyorsan megszűnnek a pórusok. Ezt követően az életben maradt sejtek nagy arányban tartalmazzák a plazmidot, de azonnal médiumba kell őket helyezni, hogy regenerálódhassanak. Az elektroporáció nagy előnye, hogy szinte minden sejtípusra alkalmazható, és nagy hatékonysággal lehet vele bejuttatni a DNS-t a sejtbe. Hátránya viszont, hogy ha az elektromos pulzusok hossza vagy intenzitása túl nagy, akkor nagy lyukak keletkeznek a sejteken, és elpusztulnak. A **4.19. ábrán** mutatunk be egy elektroporátort és a sejteket tartalmazó küvetát.



4.19. ábra: Elektroporáló készülék és küvetta

4.3.4. Infekció

Az infekció rekombináns fág illetve rekombináns eukarióta, vírus által közvetített géntranszfer baktérium vagy eukarióta (elsősorban rovar és emlős) sejtekbe. Egy donor baktérium populációban szaporítjuk fel a fágokat és a recipiens baktériumokat fertőzzük a rekombináns fág populációval. Az eukarióták esetében is csak a vírus specifikus gazdasejtjébe lehet rekombináns DNS-t bejuttatni. A módszer nagy előnye, hogy a transzformációval ellentétben, ami igen kis hatékonyságú génbeviteli módszer, **az infekció gyakorlatilag 100% hatékonyságú** (növényi gazdasejtek kivételével).

4.3.5. Génpuska

Ez a módszer egy közvetlen géntranszfer, melyet eredetileg növényi sejtek transzformálására fejlesztettek ki. Egy fémről készült (pl. arany) elemi hordozót bevonnak a bejuttatni kívánt DNS-sel. A génpuska készülék segítségével szinte akármilyen sejtbe be lehet lőni a géneket. A puska hajtóereje nagynyomású hélium gáztartály, ezzel gyorsítjuk fel a lövedéket. A makrolövedékek egy fékező lemezen fennakadnak, míg a mikrolövedékek meghatározott sebességgel szétszóródva becsapódnak a célszövetbe. Az eljárás sikerességét a DNS-sel bejuttatott markerek megjelenése jelzi. Egy génpuska vázlatos rajzát a **15.1.2.1. fejezetben**, a **15.3. ábrán** mutatjuk be.

4.4. További olvasnivaló a fejezethez

Brown TA (2010): Gene Cloning and DNA Analysis: An Introduction. Wiley-Blackwell, 6th Edition. ISBN-10: 1405181737.

Sambrook, J and Russel, DW (2001) Molecular Cloning. A Laboratory Manual. CSHL Press, volume 1-3, Third edition. ISBN: 0879695773.

5. fejezet - DNS szekvenálás

A DNS molekula felfedezését követően intenzíven foglalkoztatta a tudományt, hogy a különböző fajok vagy fajok belüli egyedek DNS-e milyen mértékben mutat egyezést vagy eltérést. Mi az, ami befolyásolhatja különböző tulajdonságok meglétét vagy hiányát, illetve bizonyos betegségek kialakulását. Ezért egyre fontosabbá vált olyan módszerek kifejlesztése, melyek segítségével meg lehet határozni a DNS szekvenciáját. **Definíció szerint a DNS szekvenálás az a folyamat, melynek során meghatározzuk egy DNS molekula nukleotid sorrendjét.** Napjainkban a DNS szekvenálás a molekuláris biológia egyik legmeghatározóbb eszközévé vált. A modern szekvenáló technikákkal már számos élőlény teljes genomját, köztük az emberét is sikeresen meghatározták (**ld. 9. fejezet**).

Az 1970-es évektől kezdve két módszert fejlesztettek ki, amelyek közül a **Sanger-féle láncterminációs (enzimatis) módszer** szinte egyeduralkodóvá vált. A fluoreszcens detektálás megjelenésével és automatizálásával könnyebbé és gyorsabbá vált a DNS szekvenálás. Végül az ezredfordulón jelentek meg a teljesen más mechanizmuson alapuló, még nagyobb hatékonyságú és olcsóbb ún. **új-generációs szekvenáló módszerek** (*next-generation sequencing*), amelyek az információszerzés mennyiségét tekintve újabb forradalmi változást indítottak el a molekuláris biológiába. Jelenleg ezeket a módszereket és az automata láncterminációs szekvenálást párhuzamosan használja a géntechnológia. A fejezetben ezeket a módszereket tekintjük át.

5.1. Kémiai szekvenálás: radioaktív végjelölt templát

Az eljárást kidolgozóiról Maxam-Gilbert-féle szekvenálásnak is nevezzük. 1977-től ez volt az első széles körben elterjedt szekvenáló eljárás. Manapság azonban már nem igazán használják, mert kiszorult a Sanger-féle módszer kidolgozását követően.

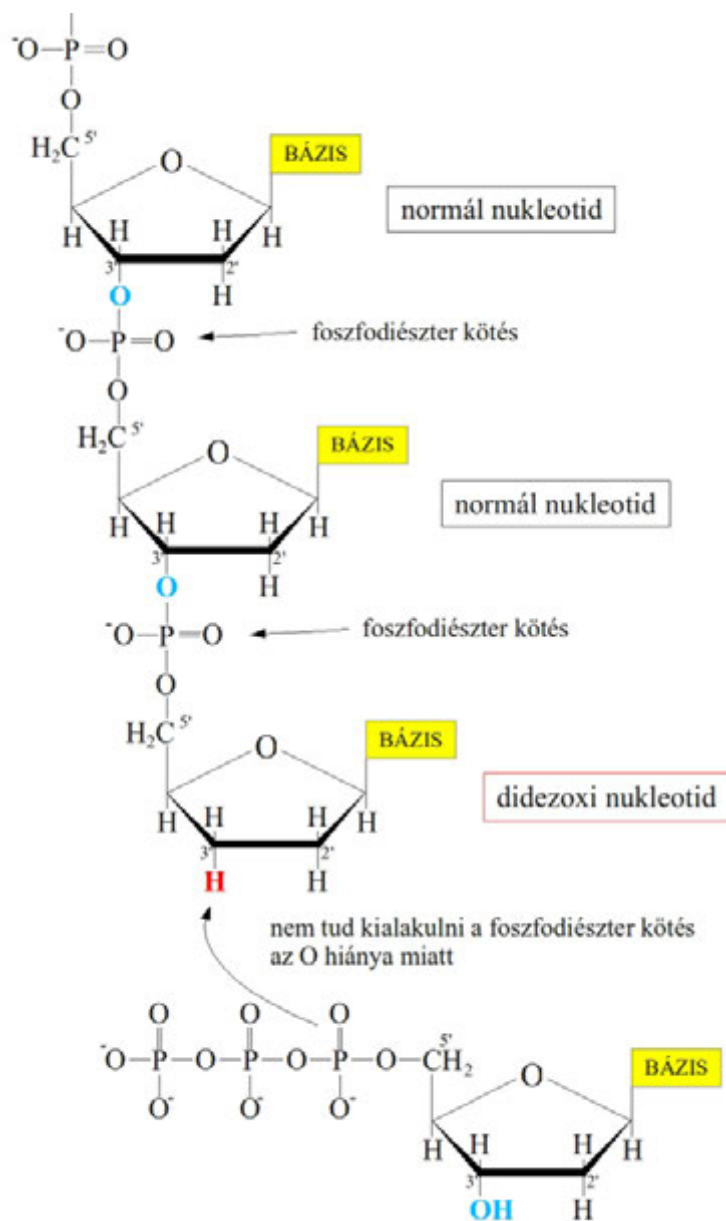
A módszer a DNS kémiai módosításán és az azt követő specifikus bázis mentén történő hasításon alapult. A DNS szálat 5'-végét radioaktív ^{32}P -tal jelölték (polinukleotid-kináz enzimmel, **ld. 3.4.3. fejezet**), majd a két komplementer szálat gélelektroforézissel izolálták. Ezután négy részre osztották a mintát és **négyféle kémiai kezelés** hatására, meghatározott nukleotidok vagy nukleotid párok mentén (A+G, G, C+T, C) törések keletkeztek a DNS-ben. A négyféle reakció: a purinok (A+G) hangyasavval depurinálhatóak; a guanin dimetil-szulfáttal metilálódik; a pirimidinek (C+T) hidrazinnal metilálódnak; a hidrazinhoz adott NaCl gátolja a timin metilálódását, így csak a citozinok módosulnak. A reakció lezajlása után mindegyik mintához forró piperidint adnak, aminek a hatására a cukor-foszfát gerinc elhasad. A reagensek koncentrációját úgy állították be, hogy átlagosan egy DNS molekula csak egyszer módosult. Végül denaturáló poliakrilamid gélelektroforézis segítségével elválasztották a különböző méretű, akár egy bázis különbségű DNS-fragmentumokat, melyek azonosítása autoradiográfiával történt. A gélben lévő négy reakciósor mintáinak nukleotid sorrendjére a sávok helyzetéből lehetett következtetni.

5.2. Enzimatis szekvenálás: láncterminációs módszer

Történetileg az első DNS szekvenálási módszert Sanger már 1975-ben kidolgozta (ez volt az ún. „plusz-mínusz” módszer, amelyet itt nem ismertetünk), de az igazi áttörést az 1977-ben publikált **láncterminációs** (más néven **didezoxi-) szekvenálás** jelentette. Egyszerűségének és megbízhatóságának köszönhetően hamar elterjedt, és nagyon népszerűvé vált. Nagy előnye volt a kémiai módszerrel összevetve, hogy maga az eljárás kevesebb mérgező vegyszert és radioaktív vegyületet igényelt.

A DNS-t ennél a módszernél klónozni kell, az eredeti módszer szerint egyszálú terméket eredményező vektorba. Erre a célra az M13-alapú bakteriofág vektorok feleltek meg (**ld. 7.3.2. fejezet**), ma azonban már használhatunk bármilyen kétszálú rekombináns DNS-t is a Sanger-féle szekvenálás céljára, sőt PCR termékeket is közvetlenül szekvenálhatunk. A templátként szolgáló szálról **DNS-polimeráz** segítségével egy **szintetikus primertől** kezdődően jelölt új szálat szintetizálunk, de a **négyféle osztott reakcióhoz** olyan **didezoxinukleotid analógokat** használunk, amelyeket ugyan az enzim elfogad szubsztrátként, de láncevégződést (láncterminációt) okoznak. A didezoxinukleotidokról hiányzik a 3'-hidroxil csoport, ezért nem képesek foszfodiészter kötést létrehozni a templátnak

megfelelő következő aktivált nukleotiddal, azaz a DNS-polimeráz működése egy ilyen nukleotid analóg beépülése után leáll. A didezoxinukleotidokat a „normál” dNTP-hez képest olyan arányban kell hozzáadni a szekvenáló reakciókhoz, hogy a szintézis egy adott ddNTP jelenlétében a templáton található komplementer nukleotidnak megfelelő összes pozícióban leálljon. Ily módon annyiféle hosszúságú új lánc keletkezik, ahány pozícióban a ddNTP-vel komplementer bázis a templát szálán előfordul. A didezoxiribonukleozid-trifoszfátok szerkezetét az **5.1. ábrán**, a random lánctermináció sémáját pedig az **5.2. ábrán** mutatjuk be.

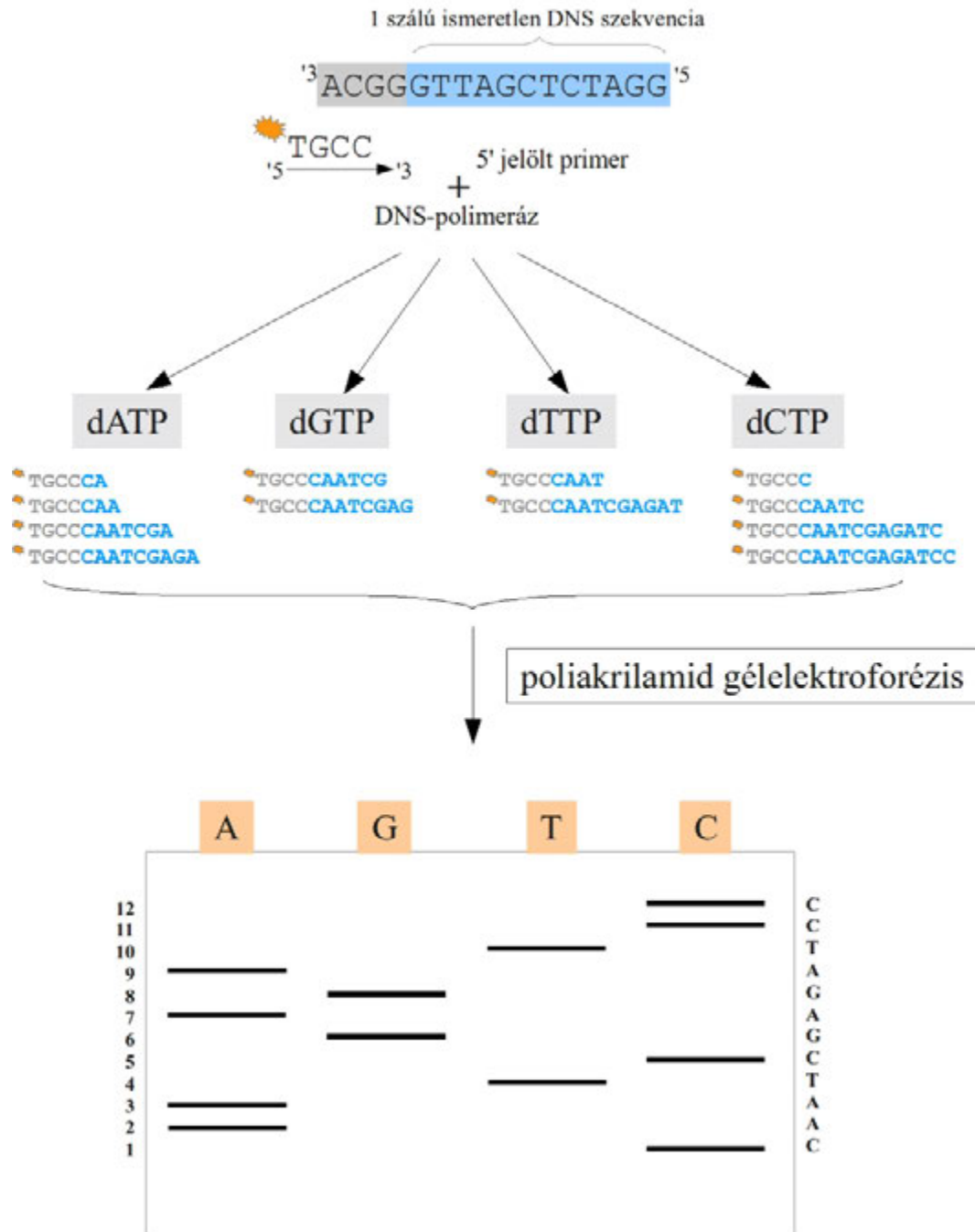


5.1. ábra: A **didezoxinukleotid hatására bekövetkező DNS-polimeráz lánctermináció**. A ddNTP-hez, a 3'-hidroxil csoport hiánya miatt, nem képes további aktivált nukleotid kapcsolódni a szintézis során, így a DNS polimerizáció megszakad.

5.2.1. Manuális láncterminációs szekvenálás

A manuális szekvenálás eredményének láthatóvá tétele **autoradiográfiával** történt. A radioaktív jelölés lehetett a primeren, a dNTP-ken (leggyakrabban a dCTP-t jelölték) vagy a ddNTP-ken egyaránt. A módszer hőskorában kizárólag a viszonylag rövid felezési idejű (13,4 nap), béta-sugárzó ^{32}P izotópot használták. Hátránya volt, hogy a jelölt nukleotid minta hamar lefeleződött és csak szigorú biztonsági intézkedéseket betartva lehetett a szekvenálást végezni. Később ezt az izotópot felváltotta az alfa-sugárzó, ezáltal kevésbé veszélyes ^{35}S izotóp (a foszfát-csoport

egyik oxigén atomját lehet kicserélni kénre), ami ráadásul sokkal hosszabb felezési idejű (87,5 nap) is. Ennek az izotópnak a hátránya viszont az volt, hogy a gyengébb sugárzás miatt az autoradiogramok előhívása hosszabb időt igényelt. A Sanger-féle láncterminációs szekvenálás sémáját az **5.2. ábrán** láthatjuk.



5.2. ábra: A láncterminációs enzimatis szekvenálás sémája. Az ábra alján a szekvenáló létráról olvasható le az újonnan szintetizált DNS lánc nukleotid sorrendje (a primertől kezdve, ami nem látszik) alulról felfelé

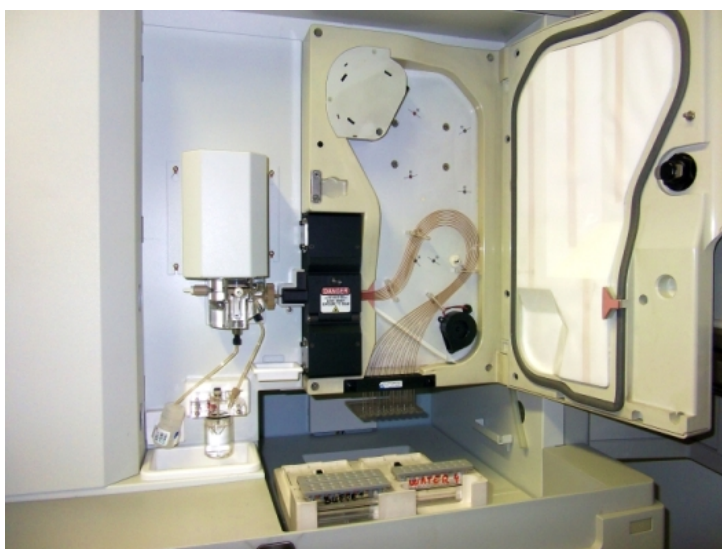
Az eljárás azzal kezdődik, hogy a DNS mintát (templát) négyfelé osztják. Az egyes mintákhoz a négyféle dNTP-t és sorra az egyik ddNTP-t adják. Például az első csőbe a négyféle dNTP mellett ddATP is tesznek, így az új szál szintézise a templáton levő timin nukleotidok után random leállhat. A négy reakcióelegy tehát abban különbözik, hogy különböző didezoxi-nukleotidot tartalmaz, de elvileg nem lehet bennük azonos hosszúságú, újonnan szintetizálódott szál. A reakció eredményeként kapott, különböző hosszúságú DNS-fragmentumok végül egy nagyméretű, lap alakú („slab”) **denaturáló poliakrilamid-urea elektroforézis** gélen választhatók el egymástól. Azért kell a gélhez urea denaturálószer adni, hogy a jelölt új szálon esetleg kialakuló bázispárosodó régiók vagy

a jelen levő templát szállal kialakuló duplex DNS (térszerkezete lévén) ne befolyásolja a DNS szálak mobilitását az elektroforézis során.

A reakcióelegyeket külön futtatják a négy ddNTP stop nukleotidnak megfelelően. Egy adott reakció csak egyféle ddNTP-t tartalmaz, például a ddCTP-t tartalmazó elegy mindig a citozinok, a ddGTP-t tartalmazó pedig a guaninok elhelyezkedését mutatja meg, így a négy mintából összeáll a DNS szekvenciája (lényegében mindegy, hogy az újonnan szintetizált vagy a komplementer templát szál szekvenciáját „olvassuk le”). Az elektroforézist követően a sávokat autoradiográfiával teszik láthatóvá. A nukleotidok sorrendjének leolvasása az egyes mintáknak megfelelő oszlopoknál alulról felfelé történik (ezt nevezzük **szekvenáló létrának**). Ha a legrövidebb (a gélen a legnagyobb mobilitású) csík a didezoxi-citidines reakcióelegynél található, akkor a DNS szekvenciája C-vel folytatódik a primer után (bár a gyakorlatban a primer utáni első néhány nukleotid általában nem detektálható). Ha a következő csík a didezoxi-guaninos elegynél látható, akkor a szekvencia guaninnal folytatódik. A leolvasás így megy tovább, amíg még a csíkok megbízhatóan elválnak egymástól (**500-1000 nukleotid hosszúságig**).

5.2.2. Automata fluoreszcens szekvenálás

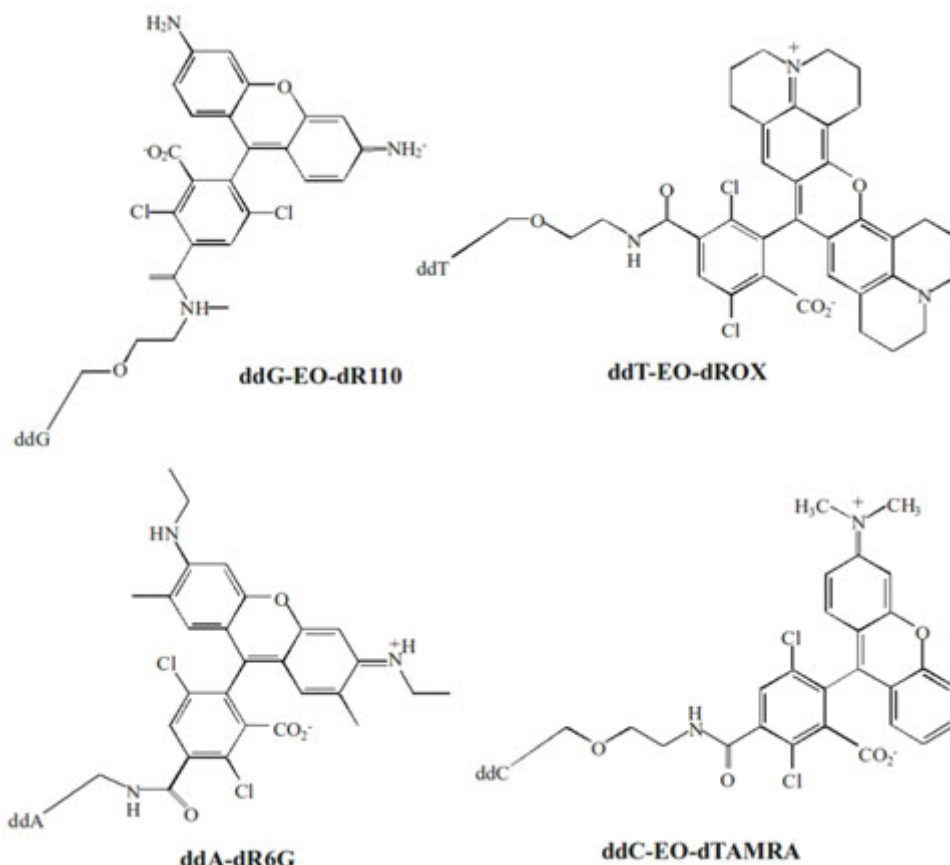
Az automata fluoreszcens szekvenálás kidolgozása és az automata szekvenátorok megjelenése (a technika kidolgozása Leroy Hood nevéhez fűződik) tette lehetővé a genom projektek megszületését, mivel nagymértékben felgyorsította a Sanger-féle szekvenáló reakciók futtatását (**ld. 5.3. ábra**).



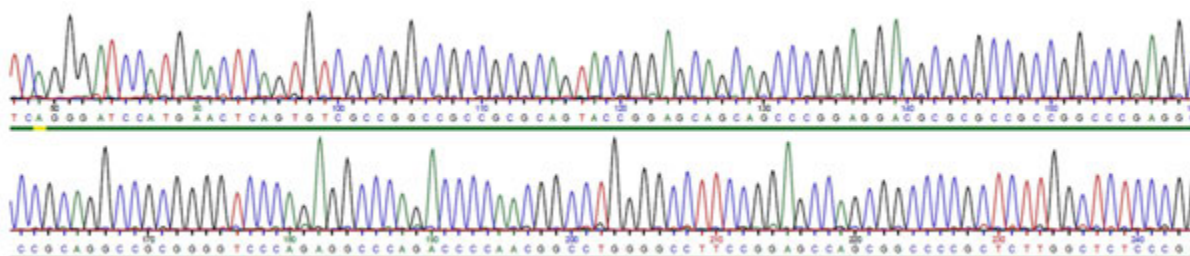
5.3. ábra: Egy automata fluoreszcens DNS szekvenátor (a Biomi Kft.-nek köszönjük a fotót)

További újítás volt, hogy a reakció „fokozása” érdekében egy PCR készülékben zajlik a láncszintézis, ezáltal kevesebb templáttal lehet indítani a reakciót, illetve több jelölt új DNS fragmentum keletkezik, így a detektálás is könnyebbé válik. Az automata fluoreszcens szekvenálás során radioaktív jelölések helyett kizárólag fluoreszcens festéket alkalmaznak. A **fluorofórt** vagy a ddNTP-kre vagy a szekvenáló primerekre lehet bevenni. Ma leggyakrabban a BigDye nevű, négyféle fluoreszcens tulajdonsággal (ez négyféle hullámhosszon megjelenő emittált fényt jelent) rendelkező ddNTP analógokat használják (**ld. 5.4. ábra**).

A módszer egyszerűségét az adja, hogy négy reakcióelegy helyett elegendő eggyel dolgozni, köszönhetően a négyféle eltérő fluoreszcens festékkel jelölt didezoxi-nukleotidoknak. A jelölt DNS fragmentumokat **kapilláris gélelektroforézissel** választják szét. Az eltérő méretű fragmentumok kimutatása fluoreszcencia-spektroszkópiával történik. A gél UV-lézerrel világítják meg (bele van építve a detektorba), és amikor áthalad egy jelölt oligonukleotid, akkor az emittált fény amplitúdóját négyféle hullámhosszon egy detektor érzékeli. Az adatokat végül számítógéppel lehet analizálni, amely azonnal megmutatja a nukleotidok sorrendjét (**kromatogram** vagy szekvenogram). Négy szín reprezentálja a különböző nukleotidokat: piros = T, zöld = A, sárga = G, kék = C. Egy automata szekvenálás eredményét (szekvenogram) az **5.5. ábra** mutatja be.



5.4. ábra: A BigDye fluoreszcens ddNTP analógok szerkezete



5.5. ábra: Automata fluoreszcens láncterminációs szekvenálás eredménye (kromatogram vagy szekvenogram)

Szekvenátortól függően egyszerre akár 96 mintát is tudnak szekvenálni kapilláris géllel, és egy futással körülbelül **900-1000 nukleotidig** lehet leolvasni a szekvenciát. Enyhe hátránya a módszernek, hogy a lánctermináló nukleotidok beépülése a DNS fragmentumokba eltérő mértékű lehet, így egyenetlen alakú és magasságú csúcsokat lehet kapni az elektronikusan kiértékelt kromatogramon. Ezeket a problémákat úgy próbálják meg kiküszöbölni, hogy módosított DNS-polimeráz enzim rendszert és festéket alkalmaznak, amely minimalizálja az egyenlőtlen beépülés okozta hibákat. További elvárás minden DNS szekvenáló projekt esetében, függetlenül a módszertől, hogy **mindkét DNS szálát szekvenálni kell**, mivel így, ha a két független információ a komplementaritásnak megfelelő lesz, biztosak lehetünk benne, hogy a templát valós szekvenciáját olvastuk le, vagyis fejtettük meg.

Manapság több biotechnológiai cégtől is vásárolhatunk láncterminációs szekvenáló kitet, amely tartalmazza a szekvenáláshoz szükséges reagenseket, így a reakciót magunk is el tudjuk végezni, csak a futtatáshoz veszünk igénybe szolgáltatást. Alternatív megoldás, hogy csak a szekvenálandó templátot (és esetleg oligonukleotid primert) küldjük el a szolgáltatónak, akitől egy szekvencia kromatogramot kapunk vissza. (A láncterminációs szekvenálásról még ld. 18.4. animáció.)

5.3. Új-generációs szekvenálás

Az új generációs szekvenálások közé tartoznak azok a módszerek, amelyek párhuzamosan sok mintát képesek szekvenálni, azaz nagy áteresztőképességű (HTP: *high-throughput*) módszerekről van szó (angolul más néven *deep sequencing*-nek, „mély” szekvenálásnak és *massively paralel sequencing*-nek is hívják az új módszereket). Az új-generációs módszerek mind a láncszintézis, mind a detektálás terén lényegileg térnek el a hagyományos Sanger-féle szekvenálástól. Használatukhoz fejlett robottechnika és igen nagy számítógép-kapacitás szükséges. Hátrányuk, hogy valamivel több hibát ejtenek és egy reakció során viszonylag rövid (néhány 100 bázis) DNS-darabot szekvenálnak, igaz, hogy párhuzamosan akár 1 milliót is. Az új-generációs módszerek bevezetésével robbanásszerűen megnövekedett a DNS-alapú vizsgálatok hatékonysága és csökkent a fajlagos költsége. Egy új-generációs szekvenátorral pár hét vagy akár pár óra alatt szekvenálni lehet akár az emberi genomot is. Elsősorban a funkcionális és a környezeti genomika (metagenomika), valamint a transzkriptomok vizsgálata területén alkalmazzák őket. Fontos megjegyezni, hogy ezek a módszerek nem csak **DNS szekvenálására**, hanem **mRNS (RNAseq)**, **kis RNS-ek (small RNA-Seq)**, és az ún. **CHIPseq szekvenálásra** (ennél a módszernél az ún. kromatin immunoprecipitációval előállított génexpresszió szabályozásban résztvevő DNS régiókat szekvenálják új-generációs módszerekkel) is alkalmasak, melyből közvetlenül következtethetünk egy sejt, vagy a vizsgálandó anyag expressziós állapotára, és DNS szabályozó régióira.

5.3.1. Pirosekválás

A **pirosekválást** 1996-ban dolgozták ki, és néhány éve már teljesen automatizálttá vált. A módszer alapelve teljes mértékben eltér a Sanger-féle szekvenálástól. A pirosekválás a „szekvenálás szintézissel” elvén alapul, vagyis egy egyszálú DNS templátról enzimatikusan komplementer szálakat szintetizálnak. **Valós időben a DNS-polimeráz aktivitását detektálják egy kemilumineszcens enzim segítségével.** Amikor egy nukleotid beépül a DNS szálba, **pirofoszfát (PPi)** keletkezik, és ennek a **mennyiségét mérjük egy kapcsolt reakcióval**, ami végső soron a luciferáz enzim felvillanással jelez. Egyszerre csak egyféle nukleotidot adnak a rendszerhez, így biztos, hogy csak egy nukleotid fog beépülni növekvő szálba. Ha több nukleotid épül be, akkor arra a fényintenzitás növekedéséből lehet következtetni. Kemilumineszcencia csak a megfelelő, komplementer nukleotid esetében következik be, hiszen csak akkor szabadulhat fel a pirofoszfát. A nem komplementer, felhasználatlan nukleotid, még a következő bázis beépülése előtt degradálódik. Alapvetően azt kell detektálni, hogy éppen milyen nukleozid-trifoszfátot adtak a rendszerhez, és az beépült-e az egyszálú DNS-láncba vagy sem.

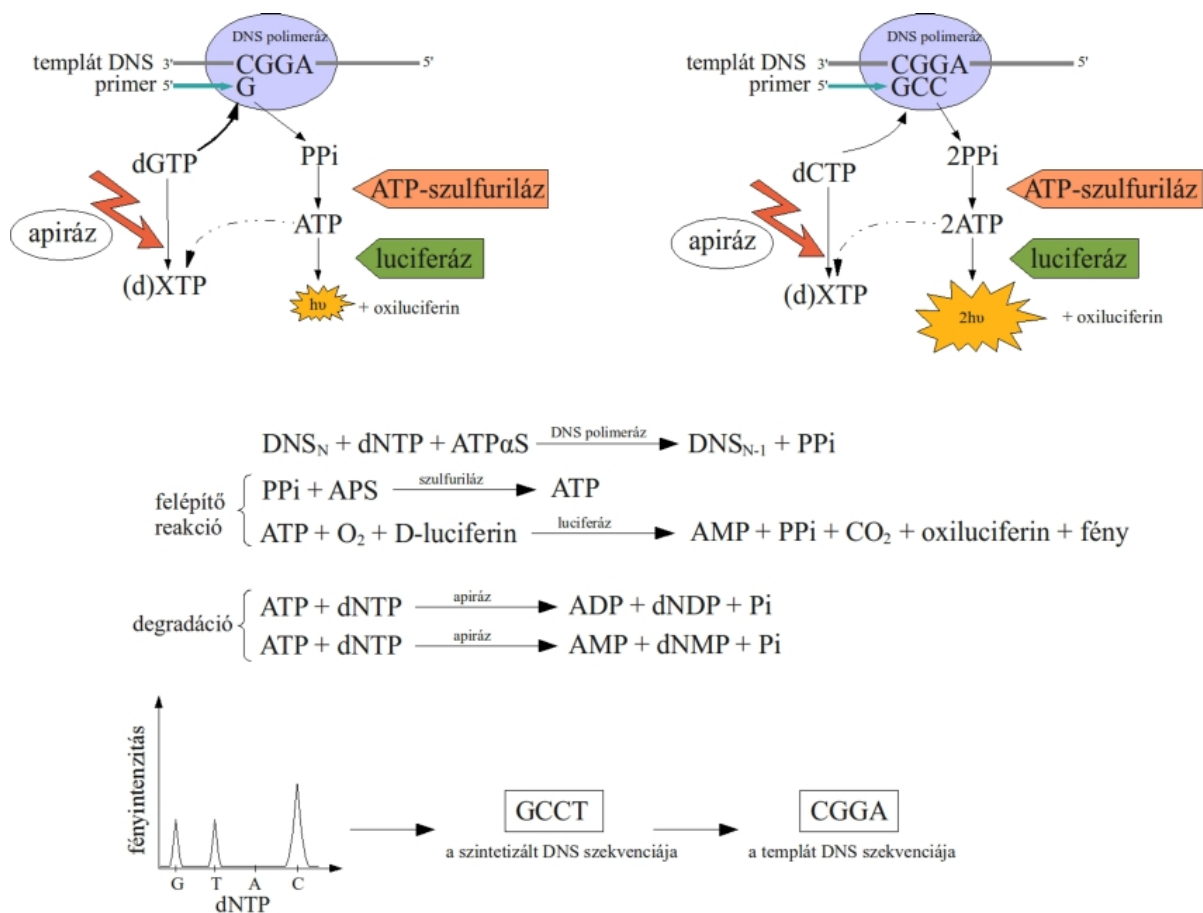
Az eljárás egy ún. **emulziós PCR** segítségével párhuzamosítható, aminek a segítségével már nincs szükség arra, hogy a szekvenálandó DNS darabokat klónozzák. A pirosekválásnak két típusa van, a szilárdfázisú és a folyadék fázisú eljárás. Az első esetben a templátokat szilárdfázisú polisztrén gyöngyökre rögzítik, míg a második esetben enzimatikusan, apiráz és exonukleáz kezeléssel készítik elő. Rögzített DNS templát esetében a vizsgálandó DNS szálakat először feldarabolják, PCR-rel felszaporítják miközben a primerrel biotinálják. Majd a szálakat aprócska méretű, streptavidint tartalmazó műgyanta (polisztrén) gyöngyökhöz kötik. Egy gyöngyökhöz mindig csak 1-1 DNS darabot kapcsolnak. Ezt követően a gyöngyöket egy olaj/víz emulzióhoz adják. A vízcseppek - melyek a PCR-hez szükséges anyagokat tartalmazzák - mérete pontosan akkora, hogy mindegyikbe 1 db gyöngy férjen el. A gyöngyöket egy olyan szálóptikából készített lemezre viszik fel, ami néhány cm² ugyan, de kb. másfél millió nanométeredést tartalmaz. Egy ilyen mélyedés kb. 15 pikoliteres térfogatú, és csak egy gyöngy fér bele. Ezekbe centrifugálással juttatják be a DNS-t, majd hozzáadják a reakcióhoz a többi szükséges összetevőt. A vízcseppekben egymástól elválasztva, külön-külön zajlanak le a reakciók. Minden gyöngyön 100-500 bázis hosszúságú DNS szakasz szaporítható fel. A másfél millió miniatűr csőben egyszerre zajlanak le a reakciók, és ez által nagyon gyorsan, nagyszámú nukleotidot lehet szekvenálni.

Az egyszálú DNS templátot először egy primerrel hibridizáltatják, majd **DNS-polimeráz**, **ATP-szulfuriláz**, **luciferáz**, **apiráz**, **adenozin-5'-foszfoszulfát (APS, mint szubsztrát)** és **luciferin** (szintén szubsztrát) jelenlétében inkubálják. A reakció első lépése, hogy valamilyen dNTP-t adnak a rendszerhez. Ha az komplementer a templáttal, akkor beépül a szálba, leválik róla a pirofoszfát, melyet a szulfuriláz megköt és az APS-ből ATP-t állít elő. Az ATP-ből luciferin jelenlétében, a luciferáz oxiluciferint hoz létre, amely egy fényvillanással járó folyamat (kemilumineszcencia). Fontos megjegyezni, hogy az ATP-nek egy módosított változatát, (ATP α S) alkalmazzák, melyet a DNS-polimeráz igen, de az ATP-szulfuriláz nem tud hasznosítani. Ennek köszönhetően a keletkező ATP biztosan a pirofoszfátból származik, és nem a mesterségesen hozzáadott ATP-t használja fel a szulfuriláz. Minden egyes felvillanást a szálóptika egy kamerához vezet, ami detektálja és rögzíti a képet. Ismétlődő bázisok esetében (pl.: TTTT), öt darab dATP fog beépülni a szintéziskor, és öt luciferint fog felhasználni a luciferáz. A kamera

képes arra, hogy ezt pontosan érzékelje, hiszen a keletkezett fénymennyiség arányos a beépülő bázisok számával, tehát ötször erősebb lesz.

A be nem épült dNTP-eket és felesleges ATP-t, rögzített DNS-templát esetében lemossák, szabad templát esetében pedig apiráz enzimmal lebontják, még a következő nukleotid hozzáadása előtt, és kezdődhet újra a ciklus. Egy reakció kb. 10 mp alatt megy végbe, így egy lapocskán kb. 20 millió nukleotidot lehet meghatározni. A módszerhez elengedhetetlen volt a megfelelő informatikai szoftver(ek) kifejlesztése, ami pontosan értelmezni tudja a kapott jeleket, melyeket nukleotid sorrendre fordít, majd a rövid szekvenciák átfedései által, összerakja a teljes DNS molekula pontos szekvenciáját. A piroszekvenálás sémáját az **5.6. ábra** szemlélteti.

Amellett, hogy a piroszekvenálás egy rendkívül költséges módszer, további hátránya, hogy mivel nagyon kis térfogatban zajlanak a reakciók, teljesen egyedi DNS templátokkal, **egy reakcióban 300-500 nukleotidnál** hosszabb templát DNS-t nem lehet alkalmazni, ami jóval rövidebb a Sanger-féle láncterminációs módszernél szintetizált új szálhoz képest (800-1200 nt). Azonban a legújabb technikákkal 5-10 óra leforgása alatt meghatározható akár 400 Mbp hosszúságú szekvencia is egyetlen készülékkel (egy ilyen futtatás kb. 5-7 ezer dollárba kerül). Baktériumok esetében 4-5 óra, eukarióták esetében pár hét alatt meg lehet szekvenálni egy teljes genomot.

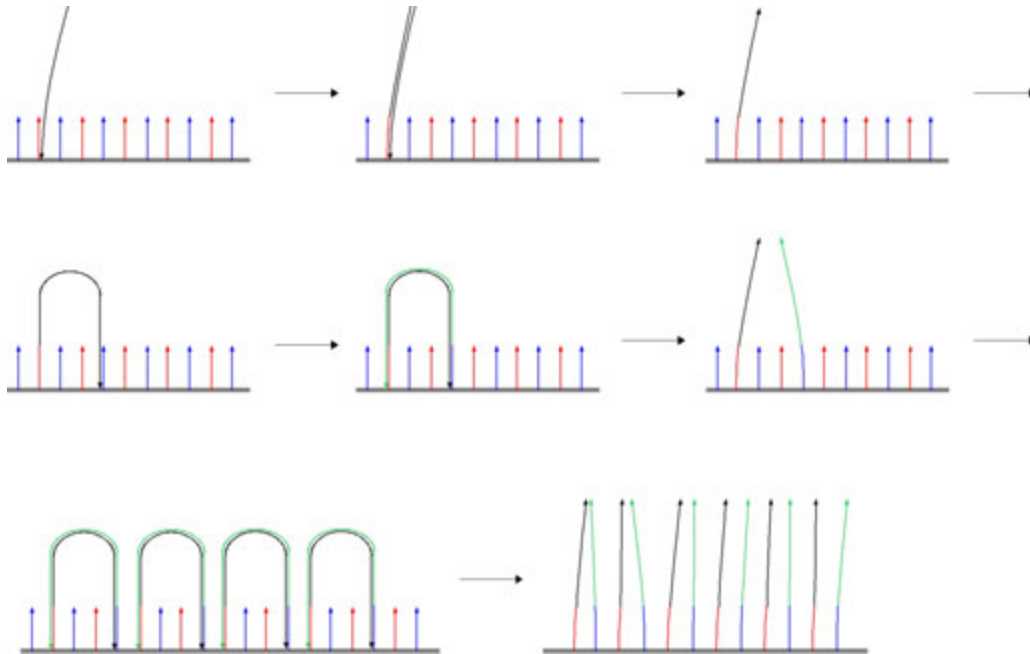


5.6. ábra: A DNS piroszekvenálás sémája

5.3.2. Illumina/Solexa szekvenálás

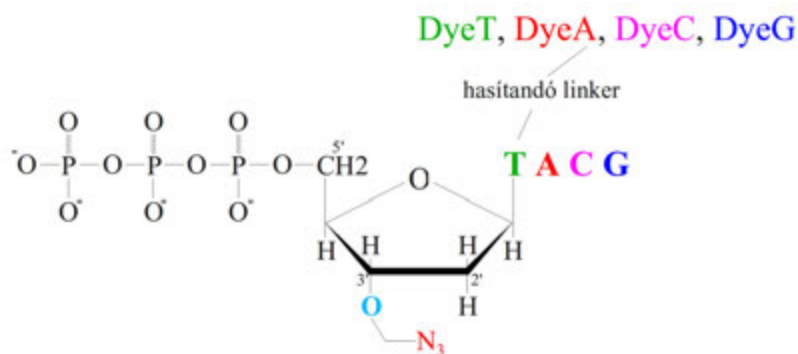
Ezt a módszert a Solexa cég fejlesztette ki, melyet az Illumina cég később felvásárolt. Jellemzősége, hogy ez volt az első rövid leolvasásokat végző technika. A szekvenálni kívánt DNS-t (akár teljes genomot) apró, **100 bp-os darabokra** fragmentálják. A dupla szálú DNS két végét kijavítják (ragadós végek eltüntetése), és egy adeninnel töltik meg az 3' végeket. Ehhez egy timin túlnyúló véggel rendelkező adapter DNS-t ligálnak. A ligált DNS-darabokat ezt követően szelektálják, és egyszálúsítják NaOH-dal. Majd a mintát egy szekvenáló lemezre viszik fel, ahova az adapterrel komplementer oligonukleotidokat (primerek) horgonyoznak ki. Ezekhez hibridizáltatják

a DNS fragmentumokat. Ezután az ún. **híd-amplifikációs** (*bridge amplification*) módszerrel (ld. 5.7. ábra) klasztereket képeznek az egyes szálak között és felszaporítják a meghatározni kívánt DNS-szakaszt.



5.7. ábra: A **Bridge-amplifikációs módszer**. A felszínen történő amplifikáció állandó hőmérsékleten (60°C) zajlik megfelelő puffer és denaturálószer jelenlétében. Minden reagens egy ciklus erejéig van jelen, aztán lemosódik. A PCR reakció végén kapott dupla szálú DNS templát szálát (amit eredetileg hibridizáltak a horgonyra) formamiddal leválasztják és lemosják. Az egyszálú DNS elhajlik, és a szabad végével egy másik lehorgonyzott primerrel hibridizál. Ezzel is végbemegy a PCR, és a végső denaturáció. Ezt követően az átírt szálak egymással vagy másik primerrel hibridizálhatnak és duplikálódnak, majd további amplifikáció után készen állnak a szekvenálásra.

Maga a szekvenálás itt is szintézissel történik. A templáthoz egy primert ligálnak, majd a reakcióhoz egyszerre adják hozzá a 4 nukleotidot (**reverzibilis terminátorok**, ld. 5.8. ábra), melyeket különböző fluoreszcens jelöléssel látnak el.

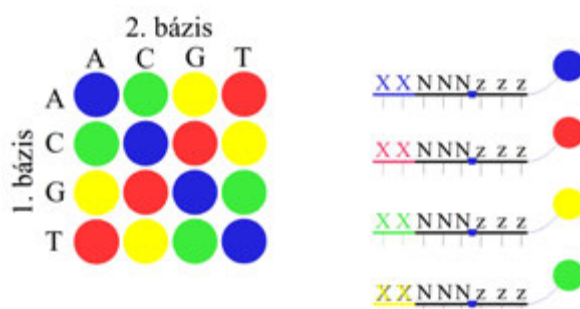


5.8. ábra: Egy ún. reverzibilis nukleotid analóg szerkezet

A beépült nukleotid fluoreszcens jelét CCD kamerával detektálják. A nukleotid beépülése után kémiai úton levágnak róla a fluorofórt és a 3' blokkolót, majd lemosják a be nem épült nukleotidokkal együtt. Ezzel a módszerrel a humán genomot 7-10 nap alatt lehet szekvenálni, ráadásul a költségek is alacsonyabbak.

5.3.3. SOLiD (*Sequencing by Oligonucleotide Ligation and Detection*)

A szekvenálás oligonukleotid ligálással és detektálással (SOLiD: *Sequencing by Oligonucleotide Ligation and Detection*) technika 2006-ban született meg. Előnye, hogy nagyságrendekkel alacsonyabb költségekkel dolgozik, mint a korábbi módszerek, viszont hátránya, hogy rövid leolvasott szekvenciákat eredményez (ezért például genom projektekhez nem igazán alkalmas). Az eljárás ebben az esetben is azzal kezdődik, hogy feldarabolják a genomot (**35-50 bázis hosszú fragmentumokra**), és egy fragmentum **könyvtárát** hoznak létre. Először a fragmentumok két végére különböző adaptereket (pl. A1 és A2) ligálnak. Második lépésként a szálakat mágneses gyöngyökhöz kapcsolják úgy, hogy az 1 µm-es gyöngyökhöz kapcsolt P1 és P2 primerek hibridizálnak a komplementer A1 és A2 templát végekkel. Minden gyöngyökhöz csak egyféle DNS darab kapcsolódik. Ezt követően **emulziós PCR**-rel felszaporítják a szálakat a gyöngyök felszínén, denaturálják a templátot és a gyöngyöket kovalensen szekvenáló lemezre kötik. A **szekvencia meghatározás itt ligálással** történik olyan oligonukleotidok segítségével, melyek oktamerek, vagyis 8 nukleotidból állnak, és négyféle fluorofórral jelölik őket (**ld. 5.9. ábra**).



5.9. ábra: A SOLiD szekvenáláshoz használt primerek. A próbák 8 nukleotidból épülnek fel. Az 1. és 2. nukleotid specifikusak (X; komplementerek a szekvenálandó templáttal), a 3., 4., 5. nukleotidok degeneráltak (N; képesek bármilyen nukleotiddal párosodni a templáton) és végül a 6., 7., 8. nukleotidok (z) univerzálisak. A próbák rendelkeznek egy 3'-hidroxil csoporttal, az 5'-végen egy fluoreszcens jellel, valamint az 5. és 6. nukleotid közötti hasítóhellyel, ahonnan a fluoreszcens jelölés, a próba beépülését követően lehasad. A próbáknak, a két darab specifikus nukleotidot figyelembe véve tehát 16-féle (4x4) kombinációja lehetséges.)

Egy templát szál ligálásához 7 ciklus szükséges, ilyenkor az 5'-végtől számított 1. és 2., a 6. és 7., valamint hasonlóan a további pozíciók párosodnak. Ahhoz, hogy az összes bázis elhelyezkedését meg lehessen ismerni, ötször kell megismételni a hétciklusos reakciót. Minden ismétlésnél új primert adnak a reakcióhoz, ami az 3'-vég felé egy-egy bázissal elcsúszva (n-1, n-2, stb.) anellál a templáthoz, így végezetül minden bázist kétszer olvas le a rendszer.

5.3.4. TSMS (*True Single Molecule Sequencing: Valódi egymolekulás szekvenálás*)

2007-ben fejlesztették ki ezt a technikát (Heliscope készülék), amely **amplifikáció nélkül** is képes szekvenálni akár egyetlen molekula DNS-t is. Hasonlóan az eddig bemutatott eljárásokhoz, a szekvenálni kívánt DNS-t először 100-200 bp-os darabokra hasítják. Majd 94°C-on denaturálják a kettőshéliceket. Az egyszálú fragmentumokhoz adapter DNS-molekulákkal fluoreszcensen jelölt poli-A farkakat kötnek. Erre azért van szükség, mert a DNS-darabokat egy kamra felületéhez kötik, amihez hatalmas sűrűségben (100 millió/cm²) oligo-T-ket kapcsoltak, és ehhez hibridizáltatva lehet a fragmentumokat a felszínhez rögzíteni. A szekvenáláshoz előkészített kamrát egy készülékbe helyezik, ami lézerrel világítja meg a felületet. Azok a pontok fognak világítani, ahová sikeresen hibridizáltak a DNS-darabok. Ezzel tulajdonképpen a detektor feltérképezi, hogy hol vannak a szekvenálásra vagyis szintetizálásra alkalmas pontok. Ezt követően lehasítják, és lemossák a fluoreszcens molekulákat a felszínről, majd megkezdik a szintézist. Ehhez DNS-polimeráz és fluoreszcensen jelölt dNTP-k szükségesek, primerként az oligo-T szolgál. Ahogyan a piroszekvenálásnál, ennél a technikánál is egyszerre csak egyféle nukleotidot adnak a rendszerhez (pl. dGTP), ami a poli-A fark utáni első citozinnal párosodni fog. Gerjesztés hatására azok a szálak fognak világítani, ahova az adott nukleotid beépült. Majd lemossák a felszínt, és jöhet egy másik jelölt nukleotid.

A négyféle nukleotidot mindig egymást követve, meghatározott sorrendben adják a felszínhez. Minden lépés végén csak azok a szálak világítanak, ahová beépült az éppen aktuális dNTP. A készülék minden egyes lépést követően egy fényképet készít a lemezről. Ezzel a módszerrel naponta 1 milliárd bázist is lehet szekvenálni. Hátránya, hogy 0,5%-os hibával dolgozik.

5.3.5. SMRT (*Single Molecule Real-time*: Egymolekulás valósídejű) szekvenálás

2011-ben fejlesztette ki egy amerikai cég a SMRT készüléket. Ezzel a módszerrel a humán genomot akár 15 perc alatt meg lehet szekvenálni kb. 100 dollárért. Ilyen készülékből jelenleg kb. 10 darab üzemel a világon. 1,5-2,9 kbp hosszúságú szakaszokkal dolgoznak, ami által pontosabb szekvencia meghatározás érhető el. A reakciók apró, 'egymolekulás' kamrákban zajlanak, a DNS szintézis pedig fluoreszcensen jelölt nukleotidokkal történik. Ez a jelölés abban különbözik az eddig leírtaktól, hogy a festéket nem a bázisra, hanem a foszfátsoporra rakják, amely a nukleotid beépüléskor lehasad a pirofoszfáttal együtt. A leváló fluoreszcens pirofoszfátot egy nanofotonikus vizualizáló kamra detektálja. Egy-egy kamrában 20 zeptoliter (20^{-21}) mintát lehet érzékelni. Minden kamra alá egyetlen DNS-polimeráz van rögzítve, ami a szintézist végzi. A módszer igazi erejét az adja, hogy egy lemezen több ezer ilyen kamrácska található, ezért hosszú DNS-szálakkal, nagyon nagyszámú minta szekvenálható meg párhuzamosan. Mint minden szintézisen alapuló szekvenáló technikánál, itt is a homopolimerek pontos meghatározásakor akadnak nehézségek, ugyanis ebben az esetben csak a fluoreszcencia intenzitásbeli különbségeire lehet hagyatkozni.

5.4. További olvasnivaló a fejezethez

Schuster SC (2008) Next-generation sequencing transforms today's biology. *Nature Methods* 5:16-18.

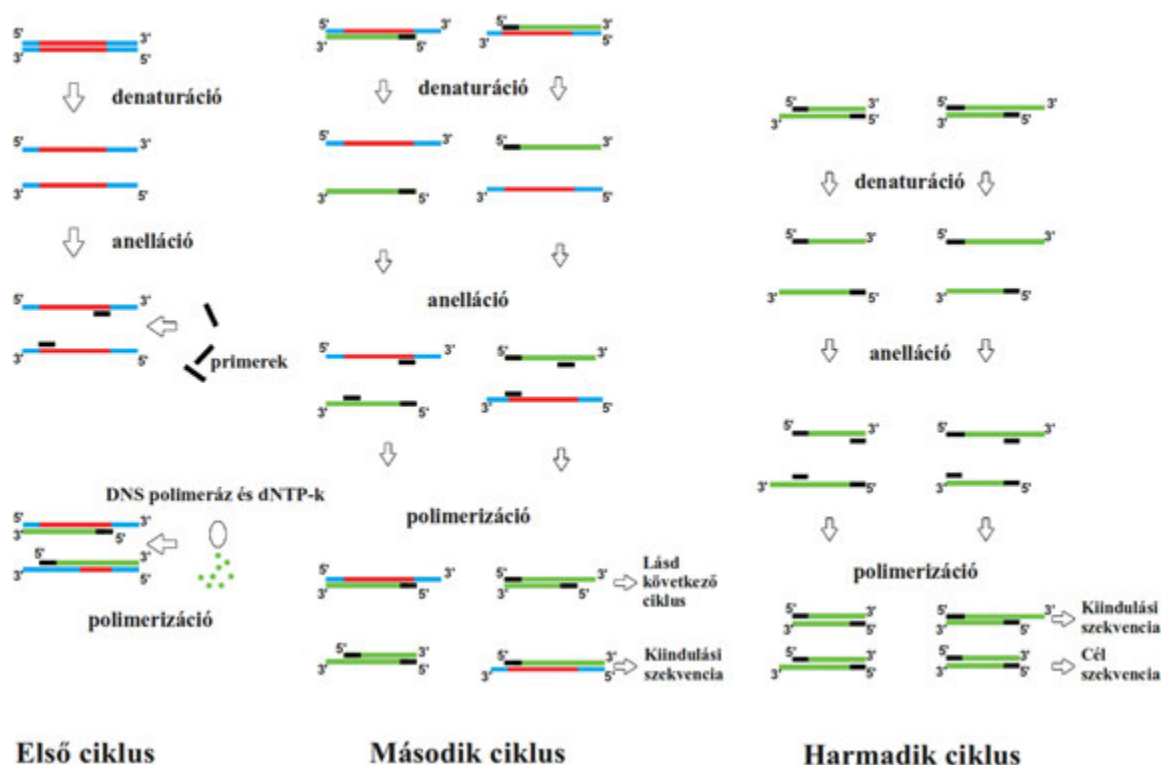
Harbers M, & Kahl G (2012) Tag-based Next Generation Sequencing. Wiley Online Library

6. fejezet - Polimeráz láncreakció (PCR)

6.1. A PCR lépései

A polimeráz láncreakció 1983-as felfedezése óta (**ld. 1.5.1. fejezet**) mára a modern biológiai kutatások és számos alkalmazott tudományterület nélkülözhetetlen módszerévé vált. A PCR jelentőségét az adja, hogy DNS-polimeráz enzimet használva laboratóriumban, tehát *in vitro*, molekuláris klónozás nélkül is képesek vagyunk specifikus DNS molekulákat sokszorozítani. A PCR eljárás során egy ciklusban három fázist különböztetünk meg (**ld. 6.1. ábra**).

1. **Denaturálás:** a kettős szálú DNS szétválasztását biztosítja.
2. **Anellálás:** szintetikus, rövid DNS oligonukleotidok tapadása (bázispárosodással) a templát szálhoz, amelyek a DNS-polimeráz számára primerként szolgálnak.
3. **Polimerizáció (extenzió):** a cél DNS szintetizálása aktivált monomerek (dNTP-k) beépítésével.



6.1. ábra: A PCR első három ciklusának sematikus ábrázolása. Színkód: piros, a többszörözni kívánt célszekvencia; kék, a teljes DNS szekvencia; fekete, primerek; zöld, újonnan szintetizált DNS. A második ciklusban a 3 fázist megismételve új végtermék keletkezik, amelyeknek egyik vége már egybeesik a kívánt minta végével. Végül láthatjuk, hogy a 3. ciklusban végbemenő folyamatok a célszekvencia (az amplikon) megjelenéséhez vezetnek.

6.1.1. Denaturálás

A DNS replikáció első fázisában a kettőshélixnek szét kell válnia, hogy az egyes szálak templátként tudjanak szolgálni a polimerizáció során. Ez a folyamat a sejten belül enzimatikusan megy végbe, a kettőshélix széttékerésében motorfehérjék, DNS-helikázok vesznek részt. A helikázok ATP felhasználásával a kémiai energiát mechanikai

munkává alakítják át, azaz a DNS szálon végighaladva a bázisok közötti H-kötéseket szétbontják. A PCR során a DNS két szálának szétválasztása nem enzimatikusan, hanem termodinamikai úton történik - magas hőmérséklet hatására a bázisokat összetartó hidrogénkötések felszakadnak. A másodlagos kötések széteséséhez szükséges hőmérséklet függ az adott DNS hosszától - minél több bázispárt tartalmaz a polimer, annál magasabb hőfok szükséges a teljes felbomlásához. A PCR első fázisában, **a denaturáció során 92-98°C közötti hőmérsékletet** használunk. Ez a hőmérséklet tartomány elegendő ahhoz, hogy a DNS templát hosszától (és komplexitásától) függetlenül teljes mértékben szétválasszuk a két szálát. A PCR „hőskorában” ez a magas hőmérsékletű fázis okozta azt a problémát, hogy minden egyes ciklust követően új DNS-polimeráz enzimet kellett a reakcióhoz adni. Ezt a problémát sikerült a Taq-polimeráz használatával megoldani. A *Thermus aquaticus* baktériumot a Yellowstone Nemzeti Park egyik hőforrásában fedezték fel. A hőforrás vize 50-80°C között ingadozik (a baktériumnak optimális hőmérséklet pedig 72°C-on van), ennek köszönhetően a benne található fehérjék ezekhez a szélsőséges körülményekhez adaptálódtak, így közel 100°C-on sem veszítik el natív szerkezetüket (pontosabban ezen a hőmérsékleten lassan már csökken az aktivitásuk, amit a PCR ciklusok előrehaladtával folyamatosan kompenzálni lehet). A manapság használatos PCR enzimek között több hőtűrő (termofil) és extrém hőtűrő (extremofil) baktériumból származó természetes vagy rekombináns, esetleg további előnyös tulajdonsággal „feljavított” DNS-polimeráz található (**ld. 3.1.4. és 6.2.2. fejezetek**).

6.1.2. Anellálás

Az PCR második fázisában történik a rövid **DNS primerek** templáthoz tapadása, anellálása. Mivel a denaturációhoz használt magas hőmérséklet megakadályozná a DNS primerek megfelelő helyre való tapadását, ebben a fázisban a hőmérsékletet le kell csökkenteni. A megfelelő tapadáshoz szükséges hőmérséklet függ az adott primerek hosszától és szekvenciájától, tehát a reakció során ennek függvényében állítjuk be az **anellációs hőmérsékletet**. Ha nem megfelelően állítjuk be, akkor attól függően, hogy alacsonyabb vagy magasabb hőmérsékletet használunk, több problémába ütközhetünk. Először is, ha az optimálisnál alacsonyabb hőmérsékletet állítunk be, akkor a primerek tapadása aspecifikussá válhat, és nem a megfelelő szekvenciájú részhez vagy több helyhez kötődve indul el a DNS polimerizáció, így nem a megfelelő, vagy többféle terméket kapunk (ez a *mispriming*-nak nevezett jelenség). Ha magasabb hőmérsékletet állítunk be az anellációhoz, akkor a DNS primerünk nem lesz képes kötődni a templát szálhoz. A gyakorlatban tapasztalt **ideális hőmérséklet tartomány a DNS primerek kötődésére 55-65°C között van** (az anellálási hőmérsékletet megbecsülhetjük az adott oligonukleotid olvadási hőmérsékletének kiszámításával, **ld. 6.2.3. és 10.3.4. fejezet**).

6.1.3. Polimerizáció/lánchosszabbítás (extenzió)

A DNS lánc szintézise, a polimerizáció vagy lánchosszabbítás (extenzió) a DNS-polimeráz elsődleges feladata. Az élővilágban számtalan különböző tulajdonságú DNS-polimeráz enzim található (**ld. 3.1. fejezet**). A közös jellemzőjük, hogy a templát szálra szintetizált primerekből kiindulva, azok 3' OH-csoportjára képesek a következő dNTP-t a templátnak megfelelően beépíteni. Az egyes polimeráz enzimek rendelkezhetnek 3'→5' és 5'→3' exonukleáz aktivitással, amelyek közül az első tulajdonság a hibajavításért, míg a második elsősorban az RNS primerek elbontásáért felelős (**ld. 3.1. fejezet**). A különböző polimerázok templátról történő másolási hűsége is nagyban különbözhet egymástól, ezért fontos hogy a PCR során olyan enzimmel dolgozzunk, aminek megfelelő a hatékonysága.

A PCR harmadik fázisában hőtűrő baktériumokból izolált, illetve manapság már rekombináns módszerekkel előállított polimerázok végzik a DNS szintézisét. Ezeknek az enzimeknek a megfelelő működéséhez optimális hőmérséklet szükséges, ami 70-80°C között van. A módszer mindhárom fázisában fontos a megfelelő időintervallum beállítása (elég idő kell mind a teljes kettős szálú DNS szétnyílásához, mind a primerek megfelelő helyre való letapadáshoz), de főként a polimerizációs fázisban kell odafigyelni, mivel minél hosszabb DNS szakaszt akarunk felerősíteni, annál több időre van szüksége az enzimnek. Általánosságban számolhatunk úgy, hogy **1 perc alatt a DNS-polimeráz ezer nukleotidot épít be**, de ma már kaphatóak ennél gyorsabb enzimek is.

A DNS termék mennyisége nagyban függ a ciklusok számától. A láncnövekedés mértéke elméletileg exponenciális, amit a következő egyenlettel lehet leírni:

$$N_t = N_k(1+H)^{n-2}$$

N_k kiindulási célszekvenciák számából n ciklust követően szintetizálódott N_t (teljes) célszekvenciák száma a totális végtermék (amit **amplikonnak** is nevezünk), H hatékonyság mellett. Ideális esetben $H = 1$, azonban a DNS-polimeráz több ciklust követően veszít a hatékonyságából, valamint a reakció során fogyó reagensek következtében a polimerizáció hatékonysága az idővel csökken. Végül figyelembe kell venni azt is, hogy a PCR molekuláris sajátsága miatt a reakció első két ciklusában még nem keletkezik a kívánt méretű DNS szál (ld. 6.1. ábra).

Látható, hogy a PCR folyamán egy ciklus alatt a különböző fázisokhoz eltérő hőmérséklet szükséges, ezért az automatizált PCR készülék egy **programozható termosztát** (ld. 6.2. ábra). (A polimeráz láncreakcióról még ld. 11.1.2., 11.2.3., 16.2.1. fejezetek és a 18.5. animáció.)



6.2. ábra: PCR készülék

6.2. A PCR reakció komponensei

Az *in vitro* DNS sokszorozáshoz a fent leírtak szerint számos komponensre van szükség. Ezeket részletezzük ebben a fejezetben, elsősorban praktikus szempontokat előtérbe helyezve.

6.2.1. PCR templát

A kívánt célszekvencia felerősítéséhez szükséges templát DNS számos forrásból származhat. A molekuláris biológiai alap kutatás során számtalan célból végezhetünk PCR reakciót, például rekombináns fehérjét kódoló gén felerősítésére (expressziós konstrukció készítéséhez), vagy taxonómiai azonosításhoz (pl. „környezeti” PCR). Az első esetben templátként különböző szövetekből vagy szervekből származó RNS-ről készült cDNS-t használhatunk (ld. 6.4.4.). Ha nem áll rendelkezésünkre megfelelő minőségű cDNS (akár cDNS könyvtár, akár saját magunk által totál RNS preparátumról készített cDNS), akkor az egyes DNS konstrukciókhoz használt templátot meg is rendelhetjük valamely biotechnológiai cégtől vagy a non-profit rekombináns DNS „klónmegosztó” szervezettől, az Addgene-től.

A faj alapú azonosításokhoz ma már elengedhetetlen a PCR használata. A mikrobiológia ennek a módszernek a segítségével képes különböző forrásból származó mintákban előforduló baktériumokat nagy pontossággal azonosítani. Az orvosi diagnosztikában például a páciensből származó mintákból felerősített PCR termék segítségével lehet megmondani, hogy a fertőzést milyen kórokozó okozhatta (legyen az akár egysejtű eukarióta, akár baktérium, akár vírus). Az igazságügyi orvostanban a DNS-t minta alapján lehet egy apasági perben a biológiai apát, egy bűnügyben pedig az elkövetőt (amennyiben DNS-t tartalmazó biológiai mintát hagy hátra a tett színhelyén) azonosítani.

A PCR, mint DNS amplifikációs módszer óriási előnye, hogy **a templát DNS-nek nem kell tisztított preparátumnak lenni**, azaz sejt- vagy szövetminta, DNS-t tartalmazó bármilyen, akár „koszos” minta PCR templátként szolgálhat. Ügyelnünk kell azonban egyes mintákban (pl. vér, talaj) a „PCR inhibitorok” jelenlétére. Ezeknek a negatív hatását számos módon lehet csökkenteni (pl. a vér esetében hozzáadott BSA-val vagy Chelex gyantával).

6.2.2. PCR enzimek

A forgalomba került, magas hőmérsékletet elviselő, röviden hőstabil DNS-polimeráz enzimek száma az utóbbi két évtizedben megtöbbszörözött. Ezeket leggyakrabban hőforrásokban élő baktériumokból, mélytengeri, vagy óceánaljzaton élő hipertermofil ösbaktériumokból izolálják.

A hőtűrő baktériumokból izolált enzimeket rendszerint klónozzák, annak érdekében, hogy rekombináns fehérjeként nagyobb mennyiségben és nagyobb tisztasággal tudják előállítani megfelelő baktériumokban, például *E. coli*-ban. Ilyen enzim a Taq-polimeráz tulajdonságaival rendelkező **Amplitaq**→, amelyet szintén a Cetus szabadalmaztatott, és még ma is ideális rövid DNS szakaszok többszörözésére. Ugyanakkor probléma, hogy az eltérő gyártóktól kapható ugyanolyan fajta enzimek is eltérnek egymástól a felerősíthető termék hosszában, mennyiségében és másolási pontosságában. A **Taq-polimeráz** hátránya az alacsony másolási pontosság (**1 hiba/10⁴ bp**), aminek az az oka, hogy az enzim nem rendelkezik 3'→5' hibajavító exonukleáz aktivitással. Széles körben elterjedt a Vent- és a Pfu-polimeráz enzimek, amelyek 3'→5' exonukleáz aktivitású DNS-polimerázok. A **Vent DNS-polimerázt** a *Thermococcus litoralis*, míg a **Pfu DNS-polimerázt** a *Pyrococcus furiosus* hipertermofil, mélytengeri hidrotermális kúrtőkben élő *Archea* baktériumokból klónozták. Az utóbbi a jelenleg ismert legnagyobb másolási hűséggel működő polimeráz enzim (**~10⁻⁶ hiba/bp**). Előnyük még, hogy magas hőmérsékleten jóval alacsonyabb a felezési idejük a Taq enzimhez képest. A 3'→5' exonukleáz aktivitású enzimek kiemelkedő másolási hűségüknek köszönhetően jóval hosszabb láncú DNS szekvenciák többszörözésére is alkalmasak, mint a Taq enzim, ugyanakkor általában lassabb a lánchosszabbító sebességük.

Helyspecifikus mutációval a polimeráz enzimek pontossága, hőstabilitása tovább fokozható. Például a Pfu polimerázt úgy módosították, hogy a polimeráz doménjéhez egy kettős-szálú DNS-kötő domént fuzionáltattak, ezzel megnövelve az enzim pontosságát és hatékonyságát (ez a **Phusion DNS-polimeráz**). (A hőstabil DNS-polimerázokkal a **3.1.4. fejezet** is foglalkozott.)

Bizonyos DNS-polimerázok, megfelelő körülmények mellett, RNS reverz transzkripciójára is képesek; ilyen enzim a *Thermus thermophilus*-ból izolált **Tth DNS-polimeráz**. Ennek az enzimnek a használatával nem szükséges külön reverz transzkriptáz enzimet adni a reakcióhoz, amennyiben RNS mintából indulunk ki és RT-PCR (**ld. 6.4.4. fejezet**) vagy kvantitatív, valós idejű PCR-t (**ld. 6.3.2. fejezet**) végzünk. A végtermék tisztaságának növelése érdekében egyes DNS-polimerázokat úgy módosítottak, hogy csak a denaturációs hőmérsékleten aktiválódjon. Ezt a módszert hívjuk „**Hot Start**” PCR-nek. Történhet az ideiglenes inaktiválás ellenanyaggal is, amely a DNS-polimerázhoz kapcsolódik, majd a denaturációs hőmérsékleten elválik tőle, és visszafordíthatatlanul denaturálódik. A **Hot Start** PCR-rel működő enzimek között igen széles a választék, ilyen DNS-polimeráz például az **AmpliTaQ Gold** néven forgalmazott módosított Taq-polimeráz.

6.2.3. PCR primerek

A PCR során történő DNS sokszorozáshoz **szintetikus (dezoxi)oligonukleotid primereket** használunk. A reakcióhoz egy ún. „*forward*” (5'-) és egy „*reverse*” (3'-) primer szükséges. Az első a felerősíteni kívánt DNS szakasz 5'-végén a komplementer szálhoz fog anellálni, míg az utóbbi az átírandó DNS szakasz 3'-végével komplementer. A mindennapi géntechnológiai alkalmazás során általában **18-25 nukleotid hosszúságú primereket** érdemes tervezni, ennél hosszabb primerek már nem növelik jelentősen a reakció fajlagosságát. A primerek szekvenciájának a 3'-végtől legalább 10 nukleotid távolságra tökéletesen komplementernek kell lennie a templát szálal, a primer 5'-végén viszont nem feltétel a komplementaritás (ezt a lehetőséget lehet kihasználni például a PCR termék vektorba történő beépítése során, **ld. 6.4.1.**)

A primerek tervezése során több szempontot figyelembe kell venni. Először is a két tervezett primer (*forward*, *reverse*) ne legyen egymás komplementere, még részlegesen sem. Ekkor ugyanis ún. **primer dimerek** jönnek létre, mivel a két primer kapcsolódását követően a DNS-polimeráz az egyiket templátként használva rövid szakaszokat szintetizál. A primer dimerek kialakulása nagyban csökkenti a helyes PCR termék keletkezését. Ügyelni

kell arra is, hogy a tervezett primeren belül ne alakuljon ki másodlagos szerkezet, például hajtűkanyar, mert az erősen gátolja a templát DNS-re való kötődését. A reakció során az alkalmazott primer koncentráció is döntő fontosságú, mivel túl alacsony koncentráció esetén csökken a letapadás esélye az anellálás során, míg túl magas primer mennyiséget alkalmazva nem kívánt, aspecifikus termékeket is kaphatunk. Az anellálás hőmérsékletének meghatározásához kiemelkedően fontos a primerek szekvenciája. A templáthoz kötődés jellemző tulajdonság a T_m , más néven **olvadási hőmérséklet**. Ez a paraméter jelzi azt a hőmérsékletet, ahol a kettős szálú DNS-ek fele szétválik. Ennek meghatározásához figyelembe kell venni a primerek hosszát és nukleotid tartalmát. Körülbelül 20-40 nukleotid hosszúságig a következő egyszerű formulával (ún. Wallace-szabály, ld. 10.3.4. fejezet) számolhatjuk ki az olvadási hőmérsékletet:

$$T_m(^{\circ}\text{C}) = 4(\text{G} + \text{C}) + 2(\text{A} + \text{T})$$

A képlet szerint a guanin-citozin (G-C) bázispárok 4°C -kal, míg az adenin-timin (A-T) bázispárok 2°C -kal járulnak hozzá az olvadási hőmérséklethez. A primerek olvadáspont számításáról további részletek találhatók a 10.3.4 fejezetben. A PCR reakció alatt az anellációhoz szükséges hőmérsékletet (T_a) általában a T_m -nél 5°C -kal alacsonyabban határozzuk meg. A tervezés során ügyelni kell arra, hogy a két primer anellációs hőmérséklete ne térjen el nagyon egymástól, így a PCR során mindkét primer kapcsolódási esélye közel azonos lesz.

6.2.4. Egyéb PCR komponensek

A megfelelő PCR-hez a fent leírtakon kívül számos más reagens is szükséges. Az egyes DNS-polimerázok helyes működéséhez az adott enzimhez **optimális puffert** kell használni, amit általában a megvásárolt enzimhez adnak. A DNS-polimerázokhoz kapható különböző pufferoldatokban általában Tris-HCl és valamilyen só, például NaCl, KCl található. Az oldatok pH-ja általában 8,5 körül van, ami majd a reakció során (a magasabb hőmérsékleten) $\sim 7,2$ -re csökken. A kereskedelmi forgalomban lévő pufferoldatok általában tízszer töményebbek a végső koncentrációkhoz képest, így hígítani kell őket a reakció összemérésekor. A pufferoldatok ezen kívül tartalmazhatnak egyéb, a reakció hatékonyságát növelő anyagokat. Például redukálószerket (pl.: ditiotritolt), detergenset (pl.: Triton X-100) vagy más ún. erősítő szereket, például: glicerint, zselatint vagy DMSO-t. Ezek az anyagok főleg a DNS-polimerázt stabilizálják, növelik a hatásfokát, de magas koncentrációban gátolhatják is a PCR reakciót.

A reakcióhoz szükséges még mind a négy fajta dezoxiribonukleotid (**dNTP**), egyenlő arányban, ugyanis az eltolt dNTP arány negatívan befolyásolhatja az amplikon keletkezését. A dNTP-k egyfelől a DNS szintéziséhez szükséges aktivált monomerek, másrészt energiaforrásként szolgálnak a DNS-polimeráz motor funkciójához (a polimeráz végighalad a templáton, tehát nem csak kémiai, hanem mechanikai munkát is végez). Kereskedelmi forgalomban kiváló minőségű dNTP keverékekhez juthatunk hozzá, 2-10 mM koncentrációjú törzsoldatban. A PCR során általában 50-200 mM koncentráció között dolgozunk

A DNS-polimerázok helyes működéséhez az enzimnek szüksége van valamilyen kétértékű kationra. A legideálisabb a Mg^{2+} , de a fehérje képes Mn^{2+} mellett is működni, igaz nem a legjobb hatásfokon. Emellett ezek az ionok a dNTP-kel komplexet képezve szolgálnak szubsztrátként az enzim számára. Megjegyzendő, hogy sokszor a kétértékű kation koncentrációja befolyásolja az amplikon mennyiségét és specitását is, ezért nem minden pufferoldatban van benne előre a Mg^{2+} , mi is hozzáadhatjuk a reakcióhoz, amennyiben azt optimalizálnunk kell.

6.3. A PCR termék (amplikon) detektálása és kvantitatív PCR

6.3.1. Agaróz és poliakrilamid gélelektroforézis

A PCR helyes végbemenetelének ellenőrzésére több módszer áll rendelkezésünkre. A legegyszerűbb és a legáltalánosabban használt módszer az **agaróz vagy poliakrilamid gélelektroforézis**, amelyek segítségével méret alapján tudunk DNS fragmentumokat elválasztani. Mindkét módszer esetén a polimerizált szilárd fázis töménységével lehet beállítani az elválasztási tartományát (ld. 4.1.2. fejezet), így a várt PCR termék hosszának megfelelően tudunk detektálni. A DNS szálak láthatóvá tételét valamilyen interkaláló festékkel (pl.: etídium-bromid, SYBR Green) végezzük el, amelyek a kettős szálba épülve UV-fény hatására megjelölik a terméket. A PCR reakciók döntő

többségét (amennyiben az ampikon hossza >100 bp) agaróz elektroforézissel detektáljuk (ennél kisebb PCR termék esetén jön szóba a PAGE)

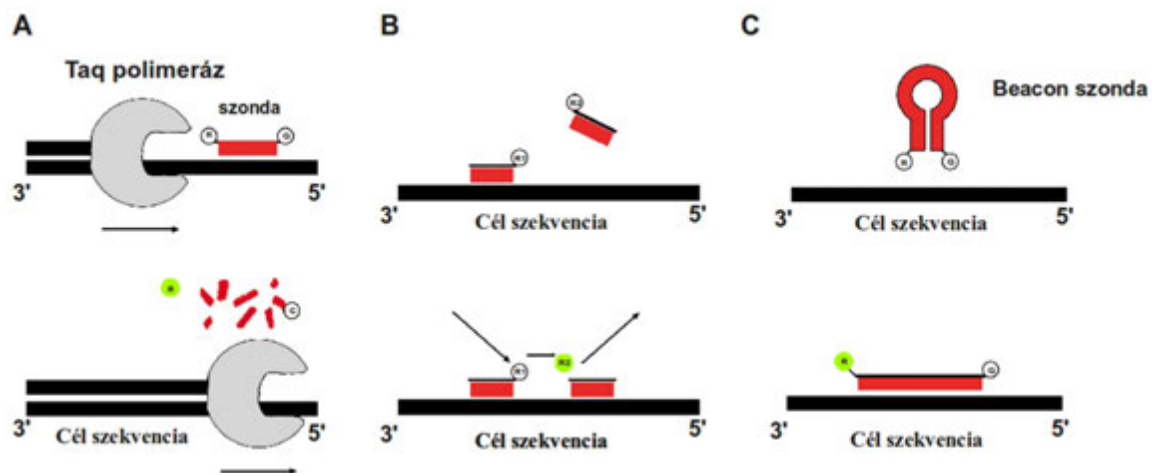
6.3.2. Valós idejű (*real-time*) vagy kvantitatív (q-) PCR

A **valós idejű**, angolul ***real-time* PCR** módszer lényege, hogy a DNS sokszorosítást **kvantitatívan** tudjuk nyomon követni. Ezzel a technikával nagy pontossággal meg tudjuk határozni a DNS templát kiindulási mennyiségét, ami sok szempontból igen hasznos lehet (pl. az orvosi diagnosztika területén, amikor egy vírusfertőzés mértékére vagyunk kíváncsiak). A *real-time* PCR-t többféle módon is meg lehet valósítani.

A valós idejű PCR során a DNS mennyiségének meghatározás azon a közös elven alapszik, hogy a célszekvenciához olyan **primer próbát** (szondát) készítenek (**harmadik oligonukleotid**), amely nem a *forward* vagy *reverse* primer letapadási helyére kötődik, hanem attól különböző helyre, valahol közöttük. Kezdetben olyan szintetikus oligonukleotidot alkalmaztak, amely 5'-végén radioaktív izotóp foszfátot tartalmazott, a 3'-véget viszont úgy módosították, hogy arról ne tudjon a polimeráz szálhosszabbítást indítani. A polimerizáció során, mikor az enzim a szondához elért, annak 5'→3' exonukleáz aktivitása elbontotta azt és később az oldatban megjelenő izotópjelet mérni lehetett. Napjainkban már olyan szondákat alkalmaznak, amelyek az izotóp használatával szemben jóval biztonságosabbak. A **Taqman** néven ismert módszernél a szonda oligonukleotid végein különböző fluoreszcens csoportokat tartalmaznak, egy riporter és kioltó (*quencher*) fluorofórt, amelyek a fluoreszcens rezonancia-energiatranszfer (FRET) jelenséget kihasználva biztosítják a detektálás lehetőségét. A *quencher* molekula a riporter csoporthoz közel kioltja annak emisszióját, ha azonban a DNS-polimeráz elbontja a szondát, a két fluoreszcens molekula az oldatban távol kerül egymástól, és így már képesek vagyunk a riporter molekulán keresztül detektálni a mennyiségi növekedést (a fluoreszcencia jel növekedése és az újonnan szintetizálódó DNS mennyisége egyenes arányos). A gyakorlatban a Taqman módszer során használhatunk két szondát is, amelyek letapadási helye egymáshoz közel kell hogy legyen. Ebben az esetben az egyik a kioltót, míg a másik oligonukleotid próba pedig a riporter fluorofórt tartalmazza és a polimeráz 5'→3' exonukleáz aktivitása biztosítja a mért jel növekedését.

A Taqman módszer egy alternatívája a *Molecular Beacon* módszer, ahol a szonda egy hajlítékanyar másodlagos szerkezetet tartalmaz, és ennek következtében kerül egymáshoz közel a kioltó és riporter fluoreszcens csoport. A DNS polimerizációt követően a szonda képes kötődni az újonnan szintetizálódott szálra, így ez a másodlagos szerkezet kinyílik, a két fluorofór távol kerül egymástól és a riporter emisszióját követni tudjuk.

A valós idejű PCR megvalósításának három lehetséges módját a **6.3. ábra** mutatja be.



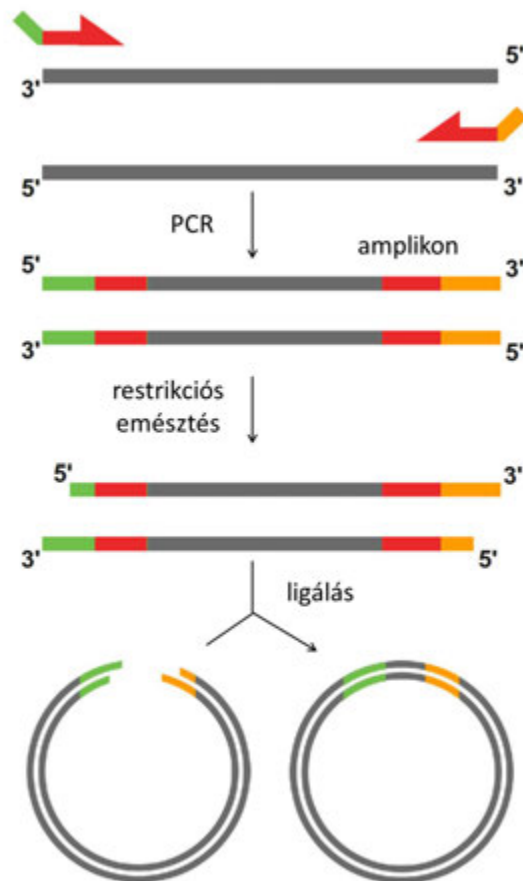
6.3. ábra: **Valós idejű PCR módszerek.** A különböző panelek (A: TaqMan, B és C: *Molecular Beacon*) az eltérő módszereket ábrázolják sematikusan. Az *R*-rel jelölt molekula a fluoreszcens riporter, a *Q* az elnyomó fluoreszcens molekula, *R1* és *R2* eltérő fluoreszcens festékek.

6.4. A PCR az alap kutatásban

6.4.1. A PCR termék (amplikon) klónozása és klónozás PCR-rel

A mindennapi géntechnológiai gyakorlatban a **PCR terméket, az amplikont** sokszor „hagyományos klónozással” rekombináns DNS-sé alakítjuk. Megjegyzendő, hogy az amplikon szigorúan véve nem molekuláris klón, mivel nem feltétlenül egyféle szekvenciát tartalmaz (emlékezzünk, hogy a hibajavító aktivitással nem rendelkező PCR enzimek több hibát vétenek, s a PCR során, lévén *in vitro* módszer, hiányoznak a sejtek nagy replikációs hűségét biztosító további hibajavító mechanizmusok is).

Az amplikont talán legtöbbször a rekombináns fehérjegyártáshoz szükséges konstrukció elkészítéséhez használjuk fel (azaz a **PCR segítségével klónozzunk, ld. 6.4. ábra**). Ahhoz, hogy valamely prokarióta vagy eukarióta expressziós rendszerben fehérjét tudjunk előállítani, az adott fehérjét kódoló DNS szakaszt egy expressziós vektorban kell bejuttatnunk a gazdasejtbe. PCR segítségével az amplifikáció után két egyszerű lépéssel elő tudjuk állítani a megfelelő konstrukciót. Először a kívánt DNS szakasz felerősítése során **a primerekbe** már a tervezés során **restrikciós endonukleáz helyeket építünk be**, amelyek megfelelnek az adott vektor klónozó helyeinek. Ezután hasítjuk a megfelelő enzimekkel az amplikont, s máris jöhet a ligálás a megfelelő módon előkészített expressziós vektorba (ld. 4.2. fejezet).



6.4. ábra: PCR termék klónozásának első lépései a primerekbe beépített restrikciós helyek (zöld és sárga) segítségével

6.4.2. Szekvenálás PCR készülékben

A Sanger által 1977-ben kidolgozott láncterminációs DNS szekvenálás (**ld. 5.2. fejezet**) jóval hatékonyabbá és gyorsabbá vált a PCR módszerrel kombinálva. A Sanger-féle szekvenálás során a denaturációt követően az alap PCR reakcióval szemben itt csak egy primerről indul a DNS szintetizálás. Ennek megfelelően a templát mennyisége minden ciklusban szaporodni fog (de természetesen nem exponenciálisan, mint a két primeres PCR reakciónál).

6.4.3. Mutagenesis

A PCR segítségével a fehérjemérnökség eszköztára is számottevően bővült. Mind az *in vitro*, mind az *in vivo* vizsgálatokhoz képesek vagyunk olyan fehérje mutánsokat tervezni, vagy random módon előállítani, amelyek segítségével jobban megismerhetjük az egyes fehérjék szerkezeti és funkcionális tulajdonságait. Ebben a fejezetben csak röviden említjük azokat a mutagenesis módszereket, amelyeket PCR segítségével lehet elvégezni. Az *in vitro* mutagenesis módszerekről további részletek a **11.2.3. fejezetben** olvashatnak.

6.4.3.1. Random (*error prone*: hibát ejtő) PCR mutagenesis

A random mutagenesis során véletlenszerűen okozunk mutációt a DNS szálban. Ezt több módon is elérhetjük. Például ha UV-fénnyel világítjuk meg, vagy mutagén szert alkalmazunk a PCR során. A random mutációk úgy is bekövetkeznek a DNS sokszorozás során, ha a PCR reakciót összeállítva nem optimális körülményeket biztosítunk a Taq-polimeráz számára (***error prone* PCR**). Példa erre, ha egyenlőtlen dNTP arányt használunk, vagy Mg^{2+} helyett Mn^{2+} -t teszünk a reakciópufferbe. A módszert részletesebben a **11.1.2. fejezetben** mutatjuk be.

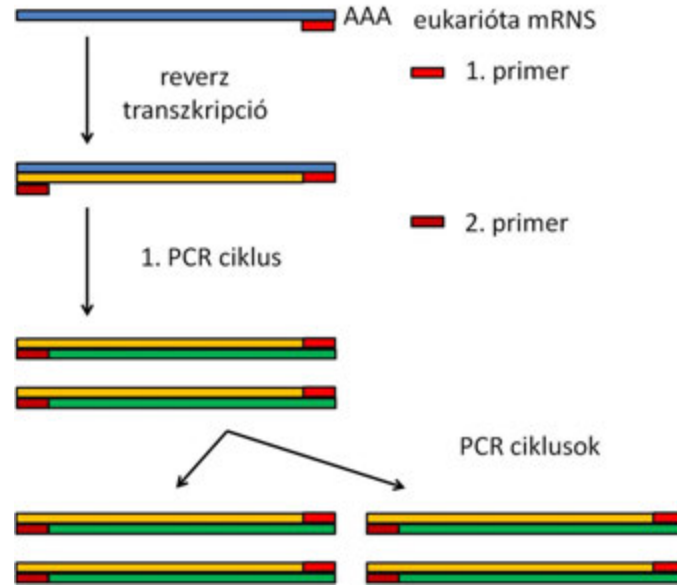
6.4.3.2. Helyspecifikus mutagenesis PCR-rel

A helyspecifikus vagy más néven irányított mutagenesis során a DNS szálon belül jól meghatározott helyre tudunk tervezni mutációt. Az ún. **megaprimer PCR** és az **overlapping PCR** módszerek részleteit a **11.2.3. fejezetben** ismertetjük.

6.4.4. Reverz transzkripció (RT-) PCR

A reverz transzkripció (RT-) PCR módszer alapja, hogy a reakcióban lévő reverz transzkriptáz enzim képes az RNS templátról DNS-t szintetizálni. Ebből adódóan ezzel a módszerrel tudunk mRNS molekulákat sokszorozni. Az orvosi diagnosztikában az RT-PCR és a valós-idejű PCR kombinálása ideális például az RNS alapú vírusfertőzések vagy rákos sejtmakerek azonosítására. Az RT-PCR alkalmazásával lehetőség van arra is, hogy kvantitatív módon határozzunk meg transzkripció szinteket, valamint a módszer segítségével képesek vagyunk sejt- vagy szövetspecifikus cDNS könyvtárakat létrehozni. Amennyiben egy adott gén expressziós szintjét akarjuk meghatározni, az mRNS-re specifikus 3'- és a komplementer szála specifikus 5'-oligonukleotidot kell terveznünk. Ha cDNS könyvtár készítése a cél, akkor az érett, tehát intronokat nem tartalmazó RNS-ek átírását kell biztosítani. Ehhez a reakcióhoz olyan oligo-dT primert adunk, ami az érett mRNS 3' végén található poli-A farkával komplementer és így biztosítja annak egy szálú DNS-sé való átírását.

Az RT-PCR során az **első ciklusban** a **reverz transzkriptáz** enzim csak az RNS-szála specifikus primert tudja meghosszabbítani, de a második ciklus során már a komplementer DNS lánc is templáttá válik, s mindkét primerrel beindulhat a „láncreakció” (**ld. 6.5. ábra**). A probléma csak az, hogy ennek az enzimnek nincs könnyen hozzáférhető, termofil élőlényből származó változata, ezért a reakcióhoz egy **hőstabil DNS-polimerázra** is szükségünk van, amely a **második ciklustól** kezdheti meg a működését. A RT-PCR során figyelembe kell venni azt a molekuláris sajátosságot, hogy a reverz transzkriptáz enzim dNTP kötése gyengébb a DNS-polimerázoknál, így magasabb koncentrációban kell a reakcióhoz adni.



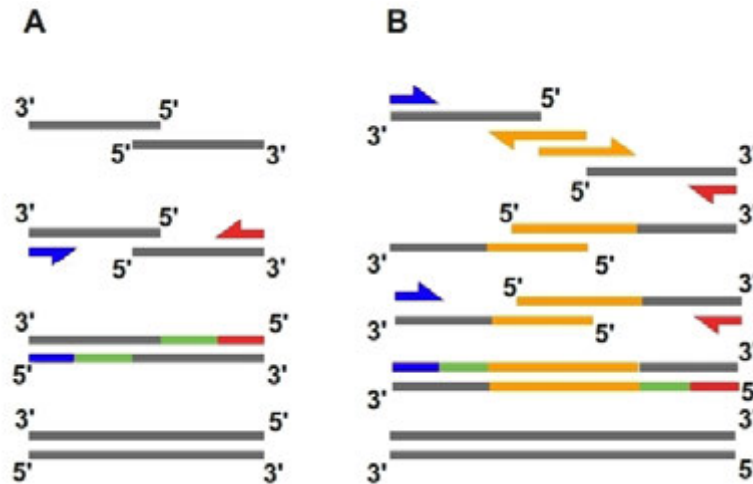
6.5. ábra: A reverz transzkripció (RT-) PCR sémája

6.4.5. További PCR módszerek (random, fúziós, multiplex, fészkes és inverz)

A PCR alkalmazásainak még számos változata ismert, amelyekből itt néhányat ismertetünk.

Az ún. **random PCR** (más néven **RAPD: random amplification of polymorphic DNA**) során olyan DNS fragmentumok amplifikálása is megoldható, amelyekre nem tudunk specifikus primert tervezni. Ilyen esetben különböző, 8-12 nukleotid hosszúságú, tetszőleges (*arbitrary*) szekvenciájú primereket használnak (egy PCR-hez egyféle primert), és ezzel végzik az általában genomiális eredetű DNS sokszorosítását, annak ellenére, hogy a primerek letapadási helyét nem ismerjük. Ennek a módszernek az előnye, hogy képesek vagyunk a pontos szekvencia ismerete nélkül is polimorf DNS-ek összehasonlítását elvégezni.

A **fúziós PCR** (angolul *jumping*, azaz „ugráló” PCR-nek is nevezik) során lehetőségünk nyílik két különálló DNS fragmentumot az amplifikáció során összekapcsolni, vagy részben degradálódott DNS templátról specifikus, hosszabb DNS régiókat amplifikálni. A módszer alkalmazása során, ha a két különálló fragmentum tartalmaz átfedő szekvenciát, akkor a két DNS szál egy-egy végére tervezett, egy *forward* és egy reverz primerrel elérhetjük, hogy a keletkező DNS szál már mindkét darabot egyszerre magában foglalja. Ha a fuzionáltatni kívánt fragmentumok nem tartalmaznak átfedő szekvenciát, akkor első lépésként a szálakra egyenként olyan primereket tervezünk, amelyek közül az egyik szál reverz, míg a másik szál *forward* primere tartalmazza ugyanazt a szekvenciát. Ezt követően olyan ampliconokat kapunk, amelyek már rendelkeznek azonos, átfedő régiókkal és ezek után már a fent leírtaknak megfelelően tudjuk egyesíteni a két DNS darabot. Ennek a módszernek a segítségével az ősmaradványokból származó fragmentált genomiális DNS-t képesek vagyunk (legalábbis egy bizonyos határig) újra egy molekulává alakítani, valamint a fúziós PCR alkalmazásával különböző gének egyesítésével egy lépésben tudunk kiméra fehérjéket kódoló konstrukciókat (**ld. 12.2.3. fejezet**) előállítani (**ld. 6.6. ábra**).



6.6. ábra: **A fúziós PCR sematikus ábrázolása.** Az első esetben (A) a két fuzionáltatni kívánt DNS szál tartalmaz átfedő régiót. A második esetben (B) az első körben PCR-rel létre hozzuk átfedő szakaszokat, amit a reakció során egy-egy primer biztosít (narancs), majd ezt követően egy újabb PCR során már a kívánt terméket kapjuk.

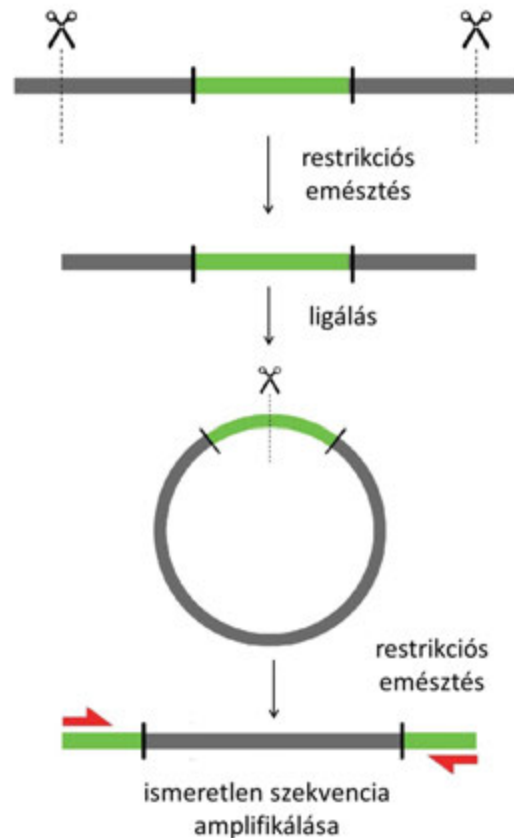
A **multiplex PCR** módszer lényege, hogy a reakcióhoz több különböző primer párt is adunk, így több amplikont kapunk egy amplifikálás során. A több primer pár alkalmazása miatt az egyes amplikonok mennyisége arányosan kevesebb lesz, mintha két specifikus primert alkalmaznánk. A multiplex PCR további technikai nehézsége, hogy el kell kerülni a különböző primerek egymással kialakított kölcsönhatását, mert azok gátolják a reakció megfelelő végbemenetelét. Fontos a detektálás szempontjából, hogy az amplikonokat különböző méretűre tervezzük. A módszert többek között patogének azonosítására, DNS-profil készítésére (ld. 6.5.2. fejezet) és nagy teljesítményű genotipizálásra alkalmazzák.

A **nested (fészkes) PCR** segítségével a nem kívánatos termékek (*mispriming*) megjelenését küszöbölhetjük ki, amely abból eredhet, hogy az egyik alkalmazott primer aspecifikusan (is) köt a célszekvenciához. A módszer során két, egymást követő PCR-t alkalmaznak, mely során a második primer pár az első páron belül anellál az első termékhez, ezzel nagyban növelve a végső termék specificitását. (ld. 6.7. ábra)



6.7. ábra: **A "nested" PCR sematikus ábrázolás.** Az első PCR során a nem kívánatos terméket, valamint a célszekvenciát tartalmazó szakaszt erősítjük fel. A második körben a két termék közül a cél szekvenciát tartalmazó terméket használjuk a továbbiakban templátként.

Az **inverz PCR** olyan esetekben alkalmazható, amikor egy ismert DNS régió környezetében lévő ismeretlen szekvenciát szeretnék feltérképezni (ld. 6.8. ábra).



6.8. ábra: Az inverz PCR sémája

A módszer lényegi lépése, hogy az ismert szekvenciától nagy távolságra valamely restrikciós endonukleázzal hasítjuk a DNS szálát, majd a mintát hígítjuk, és DNS-ligázzal a DNS fragmentumokat cirkulárizáljuk. Ezt követően a cirkuláris DNS-t az ismert szekvencia régió belül hasítjuk egy másik endonukleázzal, így olyan lineáris terméket kapunk, amelynek két végét pontosan ismerjük. Erre a templátra már képesek vagyunk PCR-t alkalmazni az ismert régióknak megfelelő primerek segítségével. A módszer gyakorlati szempontból főképp mobilis DNS elemek (retrovírusok, transzpozonok) genomba való beékelődésének pontos helyének meghatározására alkalmazható.

6.5. PCR az alkalmazott kutatásban

Napjainkra a PCR-rel számos alkalmazott tudományágban érték el átütő sikereket. Az orvosi diagnosztikában, az igazságügyi orvostanban ma már alapmódszerré vált, de érdekes módon a régészet, sőt az őslényt is alkalmazza (amennyiben a kérdéses mintában DNS-t vagy legalábbis nukleinsav töredéket lehet detektálni). Ezen kívül használják a mezőgazdaságban, az élelmiszeriparban is. A következőkben röviden áttekintjük a PCR szerepét az alkalmazott tudományokban.

6.5.1. Orvosi diagnosztika

Az orvosi diagnosztikában egyes betegségek kimutatása a diagnosztikához használt módszer érzékenységétől függ. Minél érzékenyebb az eljárás, annál hamarabb meg lehet kezdeni a szükséges beavatkozást, valamint jobban elkerülhetők a hibás negatív eredmények (igaz, hogy a fals pozitív esetek száma növekedhet). Vírusfertőzéseknel, mint például a HIV, a vírusos hepatitisz, különösen fontos a PCR-rel történő érzékeny diagnosztika használata. Tumorok vizsgálatára a valós-idejű kvantitatív reverz transzkripció PCR a megfelelő eszköz. A rákos sejtek vagy a vírusfertőzések transzkriptumaik alapján kimutathatóak.

Az onkológiai diagnosztikában a génátrendeződések következtében megjelenő fúziós mRNS termékek kimutatására is használnak PCR-t. Akut mieloid leukémiára utal például az AML1–ETO fúziós transzkriptum megjelenése. A gyógyult betegeknél különösen fontos kimutatni a minimális reziduális betegséget (MRD), a kis mennyiségű visszaeséshez vezető maradék rosszindulatú sejtet. PCR-rel az AML1–ETO fúziós transzkriptum mennyisége pontosan megbecsülhető. Az MRD megállapítása alapulhat még klonális tumor markereken, kromoszomális transzlokációs markereken és egyéb, újonnan létrejött fúziós transzkriptumokon. A valós-idejű PCR technikák használatával a felsorolt markereket vizsgálva életbevágó előrejelzések születhetnek. A telomeráz aktivitás a tumorok 85%-ban kimutatható, míg az egészséges szövetekben általában jóval alacsonyabb mértékben vagy egyáltalán nem fordul elő, így több rákos megbetegedésnél is ígéretes diagnosztikai marker lehet a PCR-rel történő kimutatásuk. Hasnyálmirigyrák vizsgálatánál a hasnyálmirigy nedvben mérve is sikerült valós-idejű PCR-rel a telomeráz hTERT alegység mRNS-ének szintjét mérni. A kutatásban a kimutatott hTERT mRNS szintje korrelált a hasnyálmirigyrák megjelenésével. Sok más példát is fel lehetne még sorolni a PCR tumordigasztikai alkalmazásáról.

6.5.2. Igazságügyi orvostan (DNS „ujjlenyomat”)

Az igazságügyi orvostanban (*forensic science*) komoly fejlődést eredményezett a DNS profilok megjelenése. A humán genomban több tízezer **rövid, tandem ismétlődő (STR: Short Tandem Repeat)** **szekvencia** található, amelyek személyenként eltérőek. Napjainkban a **DNS ujjlenyomat** készítésére főképp az STR szekvenciákat vizsgálják PCR segítségével. Az eljárás során multiplex, azaz egyszerre több, lókuszt-specifikus primer párt használva amplifikálják az STR szekvenciákat, s az így kapott, (általában kapilláris) elektroforézissel kimutatható, eltérő hosszúságú fragmentumok képezik az azonosítás tárgyát. Az USA-ban és Európában bizonyos eltérések vannak a vizsgált STR lókusztok használatában. Az azonos DNS ujjlenyomatok elkerülésének érdekében a '90-es években használt STR lókusztok számához képest mára jóval többet, akár 16 STR lókuszt is amplifikálnak egyszerre. A DNS ujjlenyomat használatával apasági tesztet is végezhetünk, amennyiben a gyerek apjának személye kérdéses. Rokonsági viszonyok feltárása is lehetséges PCR analízis (DNS profil) segítségével. Hogy csak egy példát említsünk: néhány évvel ezelőtt ezzel a módszerrel azonosították a Jekatyerinburgban 1918-ben lemészárolt Romanov orosz cári család földi maradványait.

6.5.3. „Ős”-PCR

Az ősidők kutatóinak is nagy segítséget nyújt a PCR módszer. Egyrészt a paleontológiai vizsgálatoknál az ősmaradványokból kinyert DNS felsokszorozásával és további módszerek alkalmazásával pontosabb információkat nyerhetünk az adott élőlény taxonómiai azonosításához, és ezzel evolúciós léptékben is tisztább képet kaphatunk az adott őstörténeti kor bioszférájáról. Ép DNS-t nyilvánvalóan nem lehet egy ősmaradványból remélni, de pl. a fúziós PCR segítségével kisebb darab DNS-eket is lehet amplifikálni, a bennük tárolt információt szekvenálással kinyerni, majd evolúciós következtetéseket levonni. A PCR legnagyobb hírverést kapott őslénytani alkalmazásának a dinoszauruszok korából származó DNS minta (borostyánkőbe zárt rovarból) PCR sokszorozása és szekvenálása tűnt, de később kiderült, hogy a minta későbbi korok DNS-ével volt szennyezve. Viszont valós eredményeket kaptak több ezer éves mintákból. Az eddigi legérdekesebb eredmény a legalább 50 ezer éve kipusztult **neandervölgyi ősember** gyakorlatilag teljes **genomjának** feltárása volt (**ld. 9.4. fejezet**), annak ellenére, hogy töredezett DNS-t lehetett csak a csontmaradványokból kivonni.

Az archeológiai kutatások esetén a PCR-t több szempontból is használhatjuk. Az egyik fő irány az emberi populációk, népcsoportok csontmaradványaiból származó DNS analízise, másrészt a régészeti lelőhelyeken talált további biológiai eredetű minták (háziállatok, termesztett növények, ételmaradványok, szerves eredetű tárgyak) PCR sokszorozása, szekvenálása és azonosítása. Másrészt azt a biológia sajátosságát használhatjuk ki, hogy a megtermékenyítés után a fejlődő embrió mindig az anyai mitokondriumot örökli. A mitokondriumban található DNS alapján pedig képesek vagyunk az egyes népcsoportok földrajzi elhelyezkedését és ennek változásait nyomon követni, valamint a népcsoportok között található rokonsági kapcsolatokat is felderíthetjük ezzel a módszerrel. Például archeo-PCR vizsgálatsorozattal igyekeznek genetikusok és archeológusok együttműködésében feltárni a honfoglalás kori magyarság genetikai eredetét. A honfoglalás időszakából származó csontmaradványokból vonnak ki DNS-t és ebben az esetben az apai öröklődésű Y kromoszómára jellemző marker régiókat amplifikálnak (a töredezettség miatt a mintákat először random PCR-nek vetik alá) multiplex PCR-rel.

6.6. További olvasnivaló a fejezethez

Erlich HA, Arnheim N. (1992) Genetic analysis using the polymerase chain reaction. *Annu Rev Genet.* 26:479-506

Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. (2001) Ancient DNA. *Nat Rev Genet.* 2(5):353-9.

Valasek MA, Repa JJ. (2005) The power of real-time PCR. *Adv Physiol Educ.* 29(3):151-9.

Saunders, NA and Lee, MA, editors (2013) *Real-Time PCR: Advanced Technologies and Applications*. Caister Academic Press. **ISBN:** 978-1-908230-22-5

7. fejezet - Vektor-gazda rendszerek

Korábban bemutattuk, hogy rekombináns DNS előállításához szükségünk van egy hordozó DNS vázra, a vektorra, amibe *in vitro* a molekuláris klónozás tárgyát, az inszert DNS-t beépíthetjük (**2.2. és 4.2. fejezetek**). A klónozás *in vivo* részéhez a rekombináns DNS-t megfelelő gazdaszervezetbe kell juttatnunk (génbevitel, **ld. 4.3. fejezet**). Ebben a fejezetben a vektorok típusait ismertetjük. A **vektor** legfőbb tulajdonsága, hogy képes önálló szaporodásra valamilyen gazdaszervezetben, azaz egy **replikon**. A replikációhoz szükség van egy replikációs origóra és az azt körülvevő szabályozó szekvenciákra. A replikon milyensége dönti el, hogy az adott vektor milyen gazdasejtben képes szaporodni. Vektorokat eredetük szerint **plazmidokból**, **bakteriofágokból**, **vírusokból**, **transzpozonokból** és **kromoszómákból** (bakteriális, élesztő, sőt humán kromoszómák) lehet előállítani, géntechnológiai úton. A plazmidok és a fágok hasznos tulajdonságainak ötvözéséből születtek meg a kevert típusú **kozmid** és **fágmid** vektorok.

A gazdasejteket tekintve a mindennapi géntechnológiai gyakorlatban a Gram-negatív *Escherichia coli* baktérium a legelterjedtebb gazdaszervezet. Klónozni gyakorlatilag csak *E. coli* sejtekben szoktunk. Egyes alkalmazásokhoz (például fehérjék túltermeltetésére) kutatási és ipari célra más baktérium fajokat is használnak, mint pl. a Gram-pozitív *Bacillus subtilis*, de ebben a jegyzetben csak *E. coli* gazdatörzsekkel foglalkozunk. A bakteriális rendszer előnye, hogy sokféle vektor típushoz használható, egyszerű és relatíve olcsó a laboratóriumi fenntartása, gyorsan szaporodik (optimális körülmények között 20 perc alatt osztódik), és mivel az emberi vastagbél bioflórájának tagja, ezért nem allergén. Ugyan léteznek patogén *E. coli* törzsek is (pl. az O104:H4), a géntechnológiai gyakorlatban elsősorban használt **K-12-es törzs** és több száz genetikai variánsa igen biztonságosnak tekinthetők, az emberi bélcsatornában nem képesek szaporodni. A géntechnológiában használatos legfontosabb *E. coli* K-12 altörzseket, genotípusukat és jelentőségüket, valamint a fontosabb genetikai markereket a **7.1. táblázatban** és **7.2. táblázatban** foglaljuk össze (nem mindegyik fenotípussal foglalkozunk a jegyzetben, ezekről mikrobiológiai kézikönyvekben, illetve ezen a weblapon található további információ). A K-12 törzs mellett a **B törzs** néhány variánsát is használjuk a géntechnológiában, pl. a fehérjeexpresszióhoz a **BL21-et** és származékait. Ezekről részletesebben a **ld. 12.2.2.2. fejezet** olvashatnak.

Az eukarióta gazdasejtek közül a laboratóriumi tenyésztést tekintve az élesztő áll legközelebb az *E. coli* sejtekhez. Különböző géntechnológiai célokra használunk még rovarsejteket, növényi és állati sejteket is, amelyekkel a jegyzet későbbi fejezeteiben foglalkozunk. Fontos megjegyezni, hogy az eukarióta gazdasejtek esetében gyakorlatilag minden esetben ún. „shuttle” (ingázó) vektorokat használunk, amelyek kétféle gazdában is szaporodni tudnak (bináris = kettős replikonnal rendelkeznek). Ennek az az oka, hogy mint azt már említettük, a molekuláris klónozás, a rekombináns DNS elkészítése minden esetben *E. coli* gazdában történik, s a már felszaporított konstrukciót juttatjuk be az eukarióta gazdába.

7.1. táblázat: *E. coli* K-12 altörzsek

<i>K-12 altörzs</i>	<i>fenotípus</i>
χ 1776	(diaminopimelinsav igény; UV, 37°C, detergens szenzitivitás)
HB101 ⁺	leuB6 proA2 <u>recA13</u> thi-1 lacY1 supE44 <u>hsdS20</u> Δ (mcrC-mrr)
JM105	[F' <u>lacI^qZΔM15</u> traD36 proAB] Δ (lac-proAB) thi-1 <u>endA</u> <u>hsdR4</u>
DH5 α F'	[F' <u>lacZΔM15</u>] <u>recA1</u> <u>endA1</u> gyrA96 thi-1 <u>hsdR17</u> supE44 relA1
XL1-Blue	[F' <u>lacI^qZΔM15</u> proAB Tn10(Tet ^r)] <u>recA1</u> <u>endA1</u> gyrA96(Nal ^r) thi-1 <u>hsdR17</u> supE44 relA1 lac
XL1-B. MRA	Δ (mcrA) Δ (mcrBC-hsdRMS-mrr) (+XL1-Blue)
SURE	recB recJ sbcC umuC::Tn5(Kan ^r) uvrC (+XL1-Blue MRA)
TOP10	F' <u>mcrA</u> Δ (mcrBC-hsdRMS-mrr) <u>lacZΔM15</u> <u>recA1</u> <u>endA1</u>

7.2. táblázat: Néhány fontos *E. coli* genetikai marker

<i>gén</i>	<i>fenotípus (kódolt fehérje)</i>	<i>jelentőség</i>
endA	DNS-endonukleáz I mutáció	plazmid izolálás
lacI ^q	lac-represszor túltermelés	lac szuperrepresszió
lacZΔM15	β-galaktózidáz N-term. delécio (11.-41. as)	kék-fehér szelekció
recA	homológ rekombináció és repair defekt (ss DNS-kötő ATPáz, "rekombináz")	plazmid stabilitás
recBCD	homológ rekombináció defekt (exoV)	plazmid stabilitás
supF	amber szupresszió (UAG→Tyr tRNS)	szelekció
thi1	B1 vitamin igény (tiamin enzim)	szelekció
umuC	SOS válasz mutáció	plazmid stabilitás
uvrC	UV excision repair (UvrABC endonukleáz)	plazmid stabilitás
Tn5	kanamicin reziszt. transzpozon	szelekció
Tn10	tetraciklin reziszt. transzpozon	szelekció
hsdMRS	EcoK restr. modif. komplex (metiláz, endonukleáz, specifitás alegység)	transzformálás
mcrABC	C metiláció specifikus restrikció	transzformálás
mrr	A metiláció specifikus restrikció	transzformálás
F'	F-faktor+kromoszóma darab	filamentózus fágok

7.1. A vektorok általános jellemzői

A vektorok az egyes gazda szervezetek DNS információjának módosításához használt általános géntechnológiai eszközök, a klónozendó DNS (az inszert) hordozómolekulái.

Milyen tulajdonságokkal kell rendelkezni egy DNS molekulának, hogy vektor céljára használhassuk? **Önálló replikációra** kell képesnek lennie. **Kicsi méretűnek** kell lennie, mivel akkor könnyű dolgozni vele a laboratóriumban; az ideális méret 2-5 kbp (1,4-3,5 MDa). **Egyedi restrikciós helyekkel** kell rendelkeznie, s végül kell tartalmazni legalább egy **szelekciós marker gént**.

7.1.1. Klónozó és expressziós vektorok

A vektorokat felhasználásuk alapján két nagy csoportra oszthatjuk. Az első csoportba a **klónozó vektorok** tartoznak, melyeket **specifikus DNS információ tárolására, sokszorosítására használhatunk** fel. Ezek a vektorok a gazdaszervezetre jellemző **replikonnal** kell, hogy rendelkezzenek. A „jó vektorok” alapkövetelménye, hogy az organizmus DNS állományának replikációjától függetlenül képesek sokszorozódni, ebből kifolyólag **több kópiában**

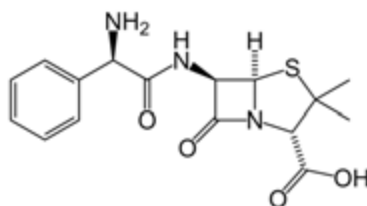
is előfordulhatnak. Minden vektornak, amelybe mesterségen kívánunk idegen DNS információt építeni, rendelkeznie kell klónozó hellyel. A legtöbbször ezek ún. **multiklónozó (poliklónozó) helyek** (MCS: *Multiple Cloning Site*), amelyek a vektorban csak egyszer előforduló, egymás mellett elhelyezkedő különböző restriktációs endonukleáz felismerő helyek. **Szelekciós marker** génként a legtöbb vektorban **antibiotikum rezisztencia gént** használnak.

A második nagy csoportba az **expressziós vektorok** tartoznak. Ezek a DNS konstrukciók olyan szabályzó elemeket tartalmaznak, amelyek lehetővé teszik, hogy a bennük tárolt DNS információ (az inszert) képes legyen fehérjeként kifejeződni, expresszálni. Az expressziós vektoroknak tartalmazniuk kell a **klónozó hely előtt egy promóter régiót**, ami lehetővé teszi a transzkripciót és ami leginkább meghatározza az expresszálandó fehérje mennyiségét. A promóter működésének szabályozhatósága kulcsfontosságú a gyakorlatban - ennek segítségével elkerülhetjük az „idő előtti” fehérje expressziót, illetve azt kívánalmainknak megfelelően indíthatjuk el. Az egyik legismertebb és leginkább alkalmazott **szabályzó mechanizmus az *E. coli* lac operon** elemeire épül, amelynek lényege röviden, hogy a génexpresszió csak allolaktóz jelenlétében indul be. A gyakorlatban az allolaktózt metabolikusan nem bontható analóg molekulával, izopropil β -D-tiogalaktóziddal (IPTG-vel) helyettesítjük.

7.1.2. Antibiotikum rezisztencia gének

A plazmidok felfedezésekor hamar fény derült arra, hogy ezek az extrakromoszomális elemek gyakran tartalmaznak antibiotikum-rezisztencia géneket, amelyek horizontális géntranszferrel nem csak a fajon belül, de különböző fajok között is vándorolhatnak. A gyakorlatban az antibiotikum-rezisztencia géneket szelekciós marker génekként használjuk. Ha a vektor rendelkezik valamely antibiotikum-rezisztencia génnel, az adott antibiotikumot tartalmazó táptalajon csak azok az egyedek lesznek életképesek, amelyekbe bekerült a vektor DNS (azonban az antibiotikum az üres és az inszertet tartalmazó vektor között nem szelektál, arra például a „kék-fehér” szelekció alkalmas (**ld. 7.2.3.1**).

Az egyik legelterjedtebb antibiotikum az **ampicillin**, ami az elsőként felfedezett penicillin félszintetikus változata (**ld. 7.1. ábra**). Az ampicillin, elődjéhez hasonlóan, béta-laktám antibiotikum, amely kötődik és ún. öngyilkos szubsztrátként gátolja a sejtfall peptidoglikán szintézisben szerepet játszó transzpeptidáz enzimet, és ezáltal gátolja a baktérium növekedését. Szerkezeti sajátága, hogy egy aminocsoport kapcsolódik a penicillin alapstruktúrájához (aminopenicillin), így néhány Gram-negatív baktérium fehérjeburokkal védett membránján is át tud hatolni, ezzel szélesebb baktériumölő spektrummal rendelkezik. Egy másik penicillin változat a **karbenicillin** (karboxipenicillin), ami a fent leírt ampicillinhez képest a Gram-negatív baktériumok felé szélesebb hatásspektrummal bír. Annak ellenére, hogy a karboxipenicillinek a béta-laktamáz enzimekkel szemben kevésbé ellenállóak, a könnyen degradálódó ampicillinhez képest jóval stabilabbak. Mindkét gátlószer elleni **rezisztenciát** a már fent említett **béta-laktamáz enzim** biztosítja, amely képes elhasítani a béta-laktám-gyűrűt, ezáltal hatástalanítja a molekulát.



7.1. ábra: Az ampicillin szerkezete

A kloramfenikol antibiotikumot 1949-ben állították elő először. Alkalmazása igen elterjedt, mert gyártása olcsó és nagyon hatékony baktérium-növekedést gátló szer, hatásspektruma széles. Bakteriosztatikus képességének molekuláris háttere, hogy a transláció során a riboszóma komplex 50S alegységéhez köt, így gátolja a fehérjeszintézist. A laboratóriumi gyakorlatban használt **kloramfenikol rezisztencia gén** egy **kloramfenikol-acetiltranszferáz** enzimet kódol, amely úgy képes inaktiválni az antibiotikumot, hogy arra két darab acetyl-csoportot köt, így a szer már nem képes kifejteni a hatását.

Egy másik fehérjeszintézis gátló antibiotikum a **tetraciklin**, amelyet a *Streptomyces* baktériumok termelnek és egy széles spektrumú poliketid típusú természetes antibiotikum. A kloramfenikollal ellentétben az aminoacyl-tRNS mRNS-riboszóma komplexhez való kötődését gátolja a 30S alegységhez kötődve. A **tetraciklin rezisztenciát** egy 40 kDa-os **membránfehérje biztosítja**, amely aktív módon **kipumpálja az antibiotikumot** a sejtből.

A **kanamicin** nevű antibiotikumot a *Streptomyces kanamyceticus* baktérium termeli, és a tetraciklinhez hasonlóan a prokarióta riboszóma 30S alegységéhez köt, és ott translációs hibát indukál, amellyel gátolja a fehérjék átíródását. A kanamicinnel közeli rokon molekula a **neomicin**. Mindkét antibiotikum aminoglikozid típusú szer, továbbá mindkettő hatékonyan gátolja főképp a Gram-negatív, kisebb részben a Gram-pozitív baktériumok növekedését. **Inaktiválásukat egy acetil-transzferáz enzim végzi**, amely egy vektorban kódolva biztosítja a rezisztenciát.

A **sztreptomycin** nevű antibakteriális szert, ahogyan nevéből is kiderül, a *Streptomyces griseus* nevű baktérium bioszintézisének természetes termékeként fedezték fel 1943-ban. Gyógyászati felhasználásban a tuberkulózis kezelésében vált igen népszerűvé. A kanamicinhez és neomicinhez hasonlóan ez az antibiotikum is az aminoglikozidok közé tartozik, és a fehérjeszintézist gátolja. Hatását a 30S riboszóma alegység 16S rRNS-éhez kötve fejt ki. Ennek következtében kodon elcsúszás alakul ki, ami végső soron a baktérium halálához vezet. A **rezisztencia** génje **sztreptomycin-6-foszfotranszferáz enzimet** kódol, ami az antibiotikumot foszforilálva gátolja annak működését.

Az utolsóként bemutatott antibakteriális molekula a **zeocin**, amely egy glikopeptid, és mikroba ölő képessége abban rejlik, hogy a kettősszalú DNS-be interkalálódva, kettősszal-törést okoz. Ezen tulajdonságának köszönhetően széles spektrumú antibiotikum. A sejten kívül a zeocin réz-ionnal komplexet képezve inaktív formában van, míg a sejten belül a redukáló környezetben a réz-ion leválik a glikopeptidről és így aktív vá válik. A **zeocin-rezisztencia** gén egy *Streptoalloteichus hindustanus* nevű organizmusból származó 14 kDa molekulatömegű fehérjét kódol, ami **megköti az antibiotikumot**, és így az nem képes a DNS-t hasítani.

7.2. Plazmid vektorok

7.2.1. A plazmidok általános jellemzői

Az első plazmidot az 1940-es évek elején azonosították egy *Shigella* baktérium fajból. A plazmidok rövid, néhány kbp és 200 kbp közötti méretű, **extrakromoszomális, kettős szálú cirkuláris DNS molekulák**, amelyek az organizmus kromoszómájától függetlenül képesek replikálódni (azaz replikonok).

Replikációjukat a baktérium kromoszómájának replikációjához szükséges enzimek egy része végzi (DNS-polimeráz I, DNS-polimeráz III, stb.). A replikáció kezdőpontját itt is a replikációs origó jelöli ki. Az origó környékén található szekvenciák szabályozzák a kópiaszámot. A legtöbb plazmid a kromoszomális DNS-ek általános bidirekcionális **replikációs mechanizmusától** eltérően **unidirekcionálisan** másolódik (azaz az origótól csak egy replikációs villa indul el). A ma használatos plazmid vektorok replikációja független a sejtosztódástól, ún. **relaxált kontroll** alatt áll. Ilyen például a pMB1 plazmidban található replikon. (Ezt a plazmidot először egy klinikai mintából izolálták 1974-ben). Az eredeti szekvenciájú pMB1 replikont tartalmazó vektorok kópiaszáma sejtenként 15-20. Az *E. coli* plazmidok között több különböző replikon szekvenciát is találtak (ColE1, p15A, pSC101). Két különböző szekvenciájú, különböző típusú replikont tartalmazó plazmid egymás mellett stabilan fenntartható ugyanabban a sejtben. Az **azonos replikációs origót hordozó plazmidok** azonban egymással **inkompatibilisek**, ezekből csak egyféle fordulhat elő egy sejtben (kivéve, ha két különböző antibiotikummal kettős szelekciót alkalmazunk a két azonos replikonnal rendelkező plazmid esetén).

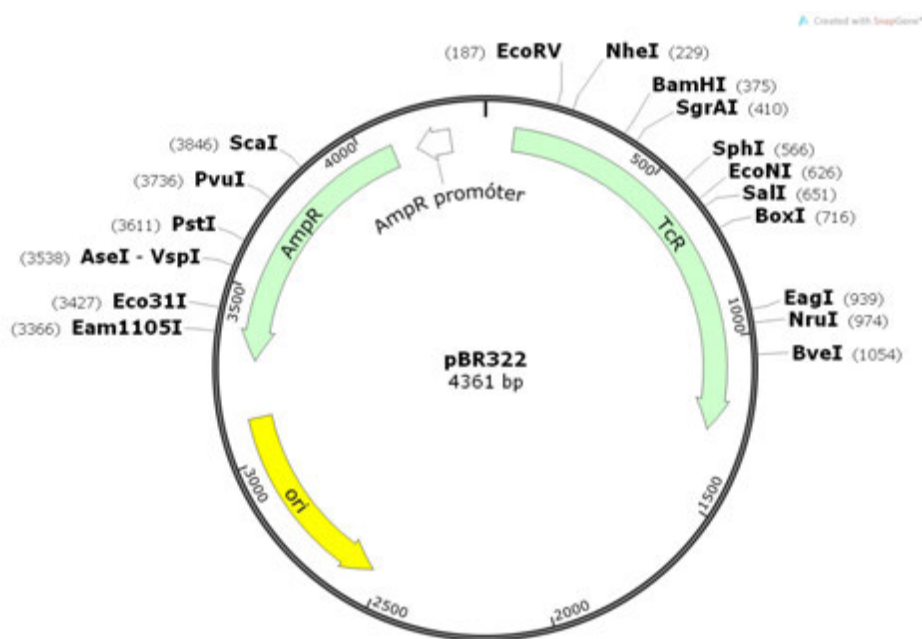
Mint hogy a plazmidok replikációja nem függ fehérjék expressziójától (a szükséges enzimek hosszú életidejűek a sejtben), a gazdasejt fehérjeszintézisének **kloramfenikollal** való gátlásával el lehet azt érni, hogy a sejt nem lesz képes a kromoszomális DNS-et replikálni, de a plazmid szintézise tovább folyik. Ezzel az eljárással, melyet **amplifikálásnak** nevezünk, tovább növelhető egy plazmid kópiaszáma.

A plazmidokat funkcionális szempontból több csoportba oszthatjuk. Az elsőbe az ún. konjugatív plazmidok tartoznak, mint pl. az *E. coli* **F-faktora** (fertilizációs-faktor), amelyek a baktériumok közötti ivaros szaporodást, a konjugációt biztosító episzómák (kromoszómán kívüli genetikai elemek). A konjugáció során a két baktérium az F-faktor-kódolta pilin oligomer fehérjéből álló piluson keresztül kapcsolódik össze. A plazmidok nagy másik csoportjába tartoznak a már említett **rezisztencia géneket hordozó plazmidok**, amelyek szintén képesek horizontális géntranszferrel más egyedekbe átjutni és biztosítani az adott antibiotikummal szemben ellenállóságot. A harmadik plazmid csoport az *E. coli* baktériumra jellemző, általuk termelt és más *E. coli* vonalakra (és más baktériumokra) ható toxinokat kódol.

Ebben a fejezetben néhány „klasszikus” plazmid-alapú vektort és vektorcsaládot fogunk bemutatni, további plazmid vektorokról a későbbi fejezetekben lesz szó.

7.2.2. A pBR322 vektor

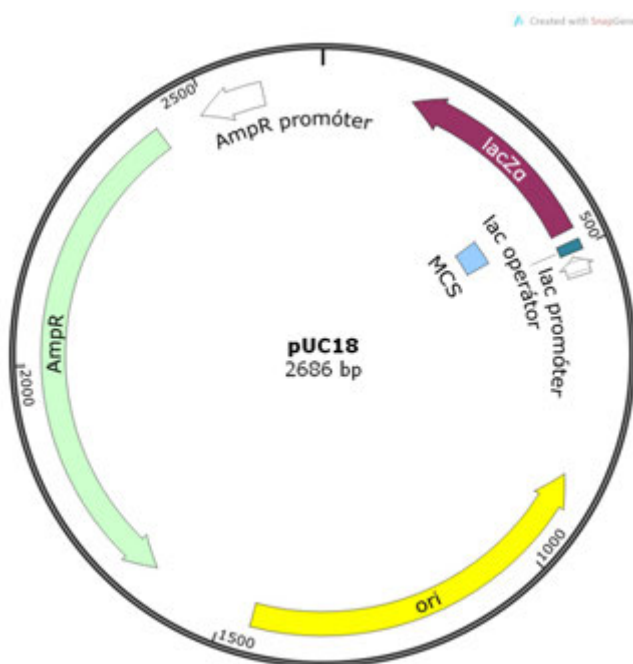
Az első széles körben ismertté vált *E. coli* kompatibilis klónozó vektort 1977-ben állították elő a Kaliforniai Egyetemen, amelyet pBR322 névre kereszteltek el (BR a feltaláló Bolivar Rodriguez monogramja). Ez a vektor 4361 bázispárból áll, megtalálható benne egy ampicillin és egy tetraciklin rezisztencia gén, valamint egy természetes *E. coli* plazmidból, a pMB1-ből származó replikációs origó. A plazmidon belül 40 egyedi restrikciós hasítóhely található, amelyekbe inszertet lehet beépíteni. 11 klónozó hely a tetraciklin rezisztenciát, míg 6 darab az ampicillin rezisztenciát kódoló régióban található, így az inszert beépülésére az adott rezisztencia gén funkcióvesztése mellett lehet szelektálni. (Ebben a vektorban még nem volt multiklónozó hely) (ld. 7.2. ábra).



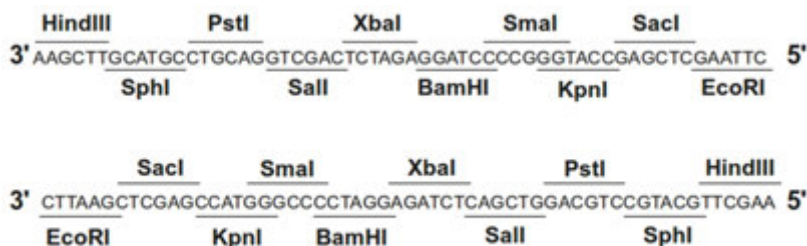
7.2. ábra: A pBR322 vektor térképe. A replikációs origón kívül (ori) látható a két antibiotikum rezisztencia gén (AmpR és TcR), valamint a bennük található 6, illetve 11 restrikciós endonukleáz felismerő hely.

7.2.3. A pUC vektorcsalád

A pUC klónozó vektorcsalád is a Kaliforniai Egyetemen látott napvilágot (az UC rövidítés a *University of California* névből ered). Sikerét egy igen ötletes tulajdonságának köszönheti, ami jócskán leegyszerűsítette a klónozás sikerét bizonyító másodlagos szelekciót (az elsődleges szelekció az antibiotikum rezisztencia, viszont annak eldöntésére, hogy üres vektor vagy rekombináns DNS került a gazdasejtbe, szükségünk lehet egy másodlagos szelekcióra is).



7.3. ábra: A pUC18 vektor térképe. A replikációs origó (ori) mellett ampicillin rezisztencia gén (AmpR), a kék-fehér szelekciót biztosító β-galaktozidáz (LacZα) α-fagmentumát kódoló gén és az utóbbin belül egy multiklónozó helyet (MCS) tartalmaz.



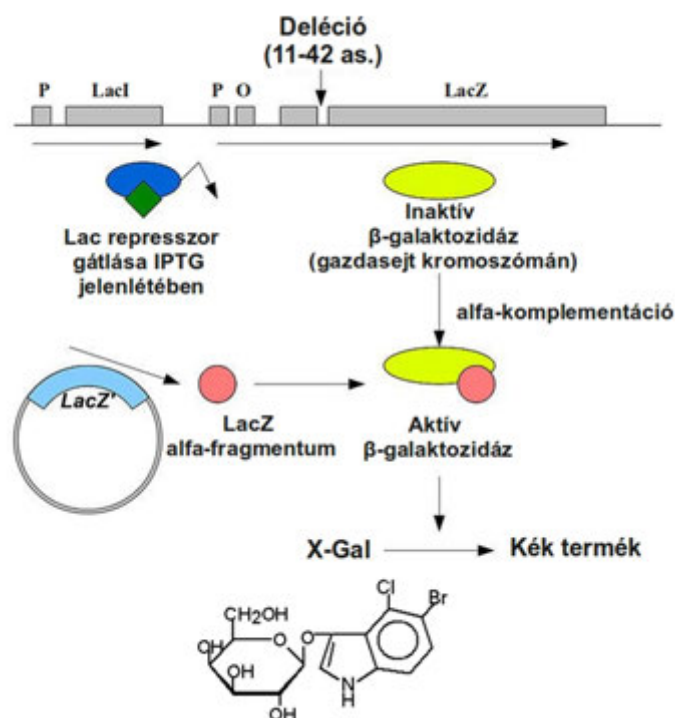
7.4. ábra: A pUC vektor család két tagjának (pUC18 és pUC19) multiklónozó helyei, amelyek egymáshoz képest reverz irányban helyezkednek el

A pUC vektor egy 2686 bp méretű plazmid, ami a pMB1 replikációs origón kívül tartalmaz még egy ampicillin rezisztencia gént, valamint a β-galaktozidáz (lacZ) enzimét kódoló gén egy darabját (lacZ'). Ennek a génnek köszönhető a másodlagos, ún. kék-fehér szelekció (ld. lentebb). További előnye, hogy a pMB1 replikon szekvenciájának módosításával a kópiaszám 500-700-ra növekedett. A vektorcsaládon belül a pUC18 és a pUC19 abban különböznek, hogy a multiklónozó helyük (MCS) egymáshoz képest reverz módon helyezkedik el (ld. 7.3. ábra és 7.4. ábra).

7.2.3.1. A „kék-fehér” szelekció

A kék-fehér szelekció segítséget nyújthat annak eldöntésében, hogy a létrehozni kívánt rekombináns plazmid tartalmazza-e az inszertet. Azok a baktérium telepek, amelyekben **sikeresen klónozott, inszertet tartalmazó plazmid van, fehérek lesznek, az üres vektort tartalmazó telepek viszont kék**. A látványos teszt molekuláris alapja a β-galaktozidáz enzim működésében rejlik és az ún. a-komplementáción alapszik. A β-galaktozidáz négy azonos alegységből álló fehérje, és csak ebben a formában aktív. Az egyes alegységek két peptidre bonthatók szét (egy rövidebb N-terminális LacZ a-fagmentumra és egy nagyméretű LacZ w-fagmentumra), aminek következtében felbomlik a homotetramer szerkezet, és az enzim inaktív lesz. Ha a két peptidet újra összekeverjük (vagyis az w-fagmentumot kiegészítjük az a-fagmentummal, amit genetikai kifejezéssel a-komplementációnak hívunk), spontán kialakul az aktív enzimszerkezet.

A szelekció során az enzimnek ezt a spontán összeszerelődési képességét használjuk ki oly módon, hogy az a-fragmentumot a plazmid kódolja, míg az w-fragmentumot a gazdasejt genomja. Mivel a cél az inszert beépülésének a vizsgálata, a vektorban a **multiklónozó hely (MCS)** az a-fragmentumot kódoló szakaszba van beépítve. **Vagyis sikeres klónozás esetén az a-fragmentumot kódoló szakasz elromlik, és az a-fragmentum nem tud kifejeződni.** Ebben az esetben tehát az IPTG-vel való indukció eredményeként az inszert által kódolt fehérje fog expresszálódni, és a telepek fehérek lesznek. Az **üres vektort** tartalmazó baktériumok IPTG indukció hatására az enzim mindkét fragmentumát szintetizálják, ami azt eredményezi, hogy egy kromogén szubsztrát, az 5-bróm-4-klór-3-indolil- β -D-galaktozid (**X-gal**) jelenlétében **kék színű baktérium telepek** keletkeznek. (ld. 7.5. ábra).

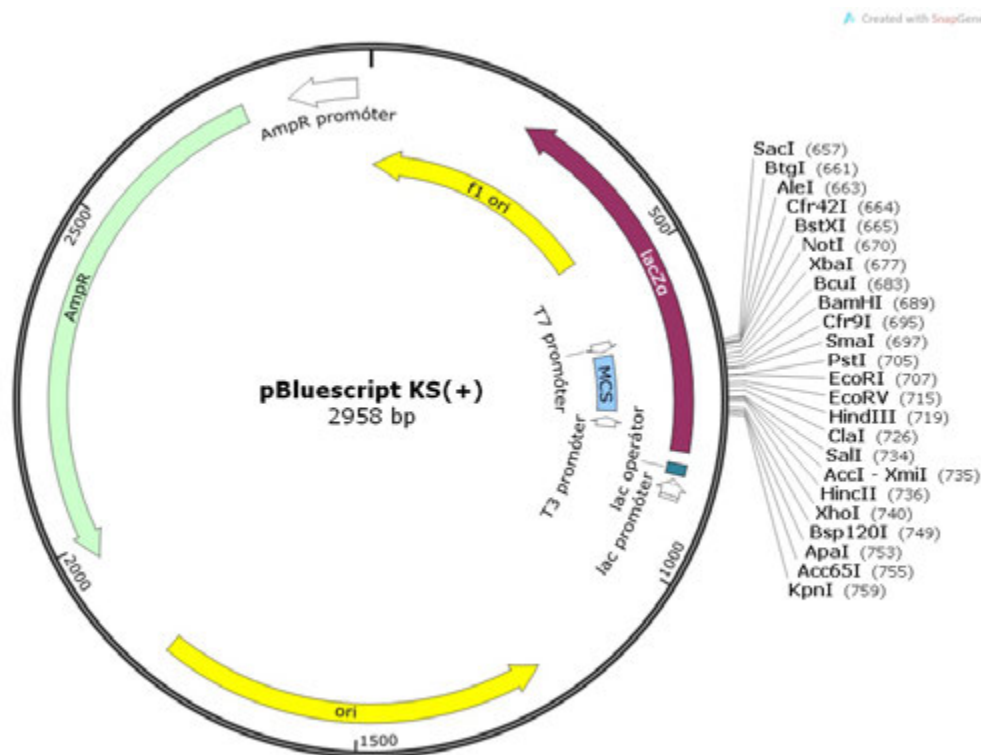


7.5. ábra: A „kék-fehér” szelekció molekuláris háttere. A csonkolt β -galaktozidáz enzim (sárga) IPTG jelenlétében a gazdasejtben termelődik, de inaktív marad, míg a vektorban kódolt enzim a-fragmentum (piros) ki nem egészíti, és aktívvá nem teszi. Ez az a-komplementáció jelensége. Az aktív enzim az X-gal szubsztrátot bontja, és kék csapadék keletkezik, ami kék színűvé teszi a telepet is. Ha az inszerció sikeres, akkor az enzim fragmentum nem termelődik, így a kolónia fehér színű marad.

7.2.4. Fágmidok

A **fágmidok**, mint képzett szó fonalas (bakterio)fág és plazmid kombinációjára utalt. Ebből adódóan jellemzői az előbbi kettő tulajdonságainak keveréke. A fágmidok a gyakorlatban klónozó vektorként használhatók. Egyrészt tartalmaznak egy **plazmid replikációs origót**, másrésztől megtalálható bennük az f1 fágól származó **f1 origó** is. Ez utóbbi a fonalas (fibrózus) fágokra jellemző és biztosítja azt, hogy a fágmidról **egyszálú DNS** is átíródjon, majd az virionba pakolódva, a későbbi gazda organizmusba átvihető legyen. A fág részecskébe csomagoláshoz szükség van ún. **segítő (helper) plazmidokra** is, amelyek az egyszálú DNS replikációját biztosítják és a virionba való pakoláshoz szükséges géneket kódolják. A fágmidok, a fibrózus fágokhoz hasonlóan nem lítikusak, tehát kiszabadulásukkor nem esik szét a gazdasejt, hanem a sejtet folyamatosan új vírusrészecskék kiválasztására készíti. A fonalas fágok, s ebből adódóan a fágmidok fertőző képességéhez az is szükséges, hogy a gazdasejt rendelkezzen pilussal. A fágmidok tehát kizárólag olyan gazdasejttel alkalmazhatóak, amelyben megtalálható vagy F-faktor, vagy annak valamilyen változata. A fágmidokat korábban szekvenálásra, valamint irányított mutagenézishez alkalmazott konstrukciókhoz használták, mivel egyszálú templátot lehet róla egyszerűen előállítani (maga a rekombináns fágmid a fág részecskében egyszálú).

Az egyik legismertebb klónozó fágmid a **pBlueScript (pBS) vektor**, amely a pUC vektorcsaládra nagyon hasonlít (ld. 7.6. ábra). Egyrészt tartalmaz egy **plazmid** és egy **f1 replikációs origót**. Ezen kívül ampicillin rezisztencia gént kódol, valamint 21 restrikciós endonukleáz kompatibilis MCS-t, amely a pUC vektor családra jellemző β -galaktozidáz α -peptidet kódoló régióján belül található. Utóbbinak köszönhetően az inszert beépülése nyomon követhető a **kék-fehér szelekció** segítségével. Megemlítendő még, hogy a poliklónozó hely két végén egy-egy **T7** illetve **T3 fág-specifikus promóter** található, amelyekről az adott promóterre specifikus RNS-polimerázzal *in vitro* transzkripciót lehet végezni az inszert mindkét száláról.



7.6. ábra: A pBlueScript fágmid vektor térképe. A vektor tartalmazza, mind a plazmid (ori), mint a fonalas fágokra jellemző f1 replikációs origót. Ezen kívül található benne egy ampicillin (AmpR) rezisztencia gén, valamint a "kék-fehér" szelekciót biztosító β -galaktozidáz (LacZ α) fragmentum kódoló régiója, és ez utóbbin belül egy MCS. A klónozott gén kifejeződését T7, illetve T3 promóterrel is el lehet indítani

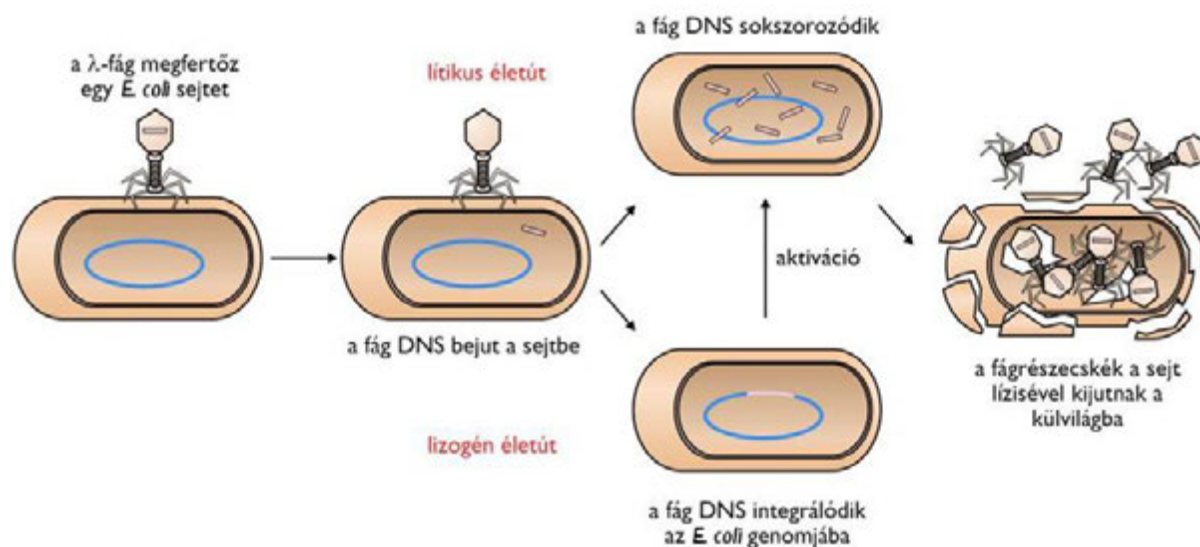
7.3. Bakteriofág vektorok

A bakteriofág-alapú vektorok előnye a plazmidokkal szemben, hogy fertőző tulajdonságuk révén sokkal hatékonyabban képesek az idegen DNS-t a gazdasejtbe bejuttatni (a génbeviteli módszereket részletesen a 4.3. fejezetben tárgyaljuk). A tipikus bakteriofág egy külső, fehérjéből álló kapszidból (a virion burka), és a benne lévő örökítőanyagból áll. A genomjukat alkothatja RNS és DNS is, egy- vagy kétszálú formában is. A géntechnológiában vektor előállítása céljából elsősorban az ikozaéderez **λ -fágot** és a filamentózus **M13 fágot** vonták be.

7.3.1. λ -fág vektorok

A géntechnológia alkalmazásban az egyik legelterjedtebb bakteriofág az *E. coli* specifikus λ -fág (más néven enterobakteriofág- λ). Szerkezetében tipikus fág struktúrát mutat, rendelkezik fej, más néven egy kapszid régióval, amelyben lineáris **kettős szálú DNS örökítőanyag** található, valamint farok régióval, amin keresztül a DNS bejuthat a gazdasejtbe. A vírus genomja ~48500 bp-ból áll és közel 60 gént kódol, amelyek biztosítják a fág szaporodásához és becsomagolásához szükséges enzimeket, azok szabályzó fehérjéit, valamint burokfehérjéket. A vírus lineáris genomja képes cirkularizálódni a gazdasejten belül, ezt a DNS szálak végén található két **ragadó cos régió** biztosítja. Az λ -fág kétféle életciklust mutat. A **lítikus életciklus** során a vírus örökítőanyaga miután

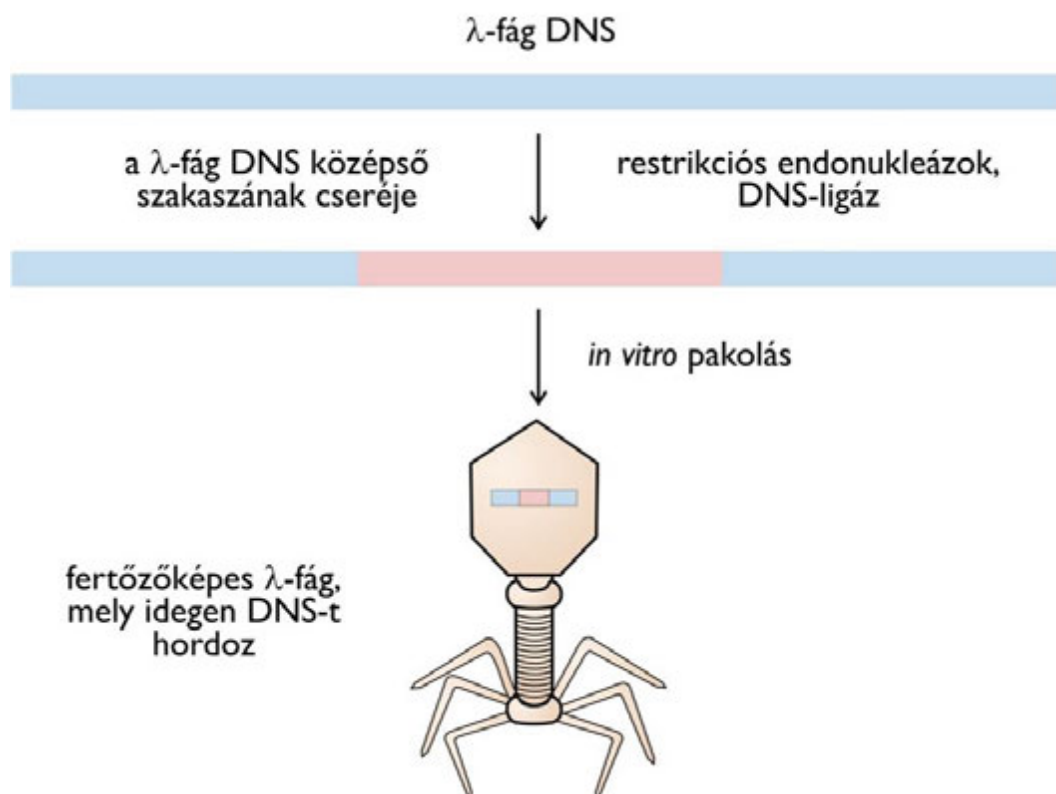
bejutott a gazdasejtbe, ott többször replikálódik, majd a fej és farok struktúrát képző fehérjék expresszióját követően új vírusrészecskék képződnek, amik a gazdasejt elpusztulása után további sejteket fertőznek. Ezzel ellentétben a **lizogén életciklus**ba akkor lép a fág, amikor az örökítőanyaga beépül az *E. coli* kromoszómájába, azzal együtt replikálódik. A sejt osztódásával tovább öröklődő, látens fágot **profágnak** nevezzük (ld. 7.7. ábra).



7.7. ábra: A λ -fág lítikus és lizogén életciklusa

A λ -fág-alapú vektorokon belül megkülönböztetünk **inszerciós** és szubsztitúciós (helyettesítő) **vektorokat**. Előbbi esetén a lineáris DNS-ben található egy restrikciós endonukleáz hely, ahova maximum **7 kbp méretű DNS inszert**et építhetünk be. Ezt a típust főként **cDNS könyvtárak előállításához** használják (ld. 8.3. fejezet). A **helyettesítő vektorok** esetén a lineáris DNS-en belül található egy "kitöltő" szekvencia (amely régió eredetileg a vírus lítikus életciklusához nem esszenciális géneket tartalmaz), amelyet elsőként ki kell vágni a vektorból és annak helyére tudjuk beépíteni az idegen génszakaszt). Szubsztitúciós vektorok alkalmazásánál a maximális beépíthető **inszert méret 20 kbp**, és főleg **genomiális könyvtárakat** állítanak elő belőlük (ld. 8.2. fejezet). Egy helyettesítő λ -fág vektor konstrukciót ábrázol a 7.8. ábra.

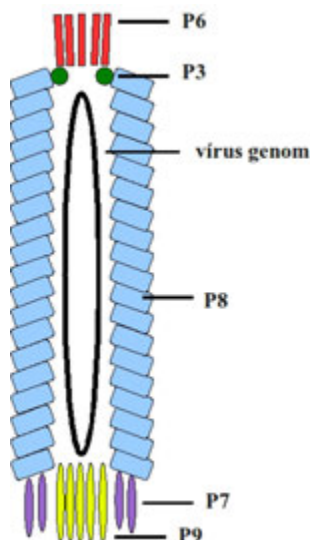
Az λ -fág alapú vektorok alkalmazásakor ügyelni kell a beépíteni kívánt inszert méretére, ugyanis ha az inszert túl kicsi, akkor nem alakul ki a fág fejében a fertőzéshez nélkülözhetetlen nyomás, míg ha túl nagy, akkor nem fér el a DNS a virion feji részében. A rekombináns konstrukciók elkészítésénél nehézséget jelent, hogy a vektor „bal” és „jobb” karját kell az inszerttel ligálni, azaz hármass ligálás történik. A λ -fág vektorok használatának viszont nagy előnye, hogy a **rekombináns DNS *in vitro* „becsomagolható”** a vírusrészecskékbe (ehhez olyan „csomagoló kit”-eket használunk, ami a virionok kialakulásához szükséges összes komponens tartalmazza), s ezáltal infekcióval a megfelelő gazdasejtbe juttatás határfoka kiváló.



7.8. ábra: **Rekombináns DNS helyettesítő λ-fág vektorban.** A λ-fág genomjának középső szakasza (piros) a fág működéséhez nem esszenciális, ezért az inszerttel helyettesíthető. Mivel a DNS fágba pakolásához a DNS méretének egy szűk tartományba kell esnie, így az idegen DNS nélkül a visszazáródó forma túl kicsi ahhoz, hogy a specifikus fehérjék fágba építsék. Az *in vitro* pakolórendszer a fág genom két végén ismer fel egy rövid DNS szekvenciát.

7.3.2. M13-alapú vektorok

Az **M13** fonalas bakteriofág egy 6,4 kbp-ból álló, **egyszálú DNS örökítőanyaggal** rendelkező, a λ -fághoz hasonlóan *E. coli*-t fertőző vírus. A viriont kb. 2700 kópiaszámú kapszidburok fehérje (P8) alkotja, melynek egyik végét a P9 (5 kópiában) és P7, míg másik végét a P6 (5 db) és P3 fehérjék zárják le (**ld. 7.9. ábra**).



7.9. ábra: Az **M13 fág sematikus ábrázolása**. Az egyszálú cirkuláris DNS-t körülvevő kapszidot a P8 (kék) fehérje alkotja. A viriont az egyik oldalról a P6 (piros) és P3 (zöld), a másiktól a P9 (sárga), és a P7 (lila) fehérjék zárják le.

A P3 fehérje szerepe igen fontos a gazdasejt pilusán át történő fertőzésében. Az M13 fág **nem lítikus** a gazdasejtre nézve. A fág egyszálú DNS-e a replikáció során **kétszálú replikatív** formává alakul, amely a piluson át történő transzfer során újra egyszálúvá válik és így fertőz tovább. A laboratóriumi gyakorlatban a kétszálú **replikatív formát** használják **klónozó vektor**ként. A vírus restrikciós endonukleáz hasító helyeket tartalmaz (pl.: az **M13mp18** fág-vektor a pUC18 plazmid vektor tulajdonságait örökölte), ahova könnyen be lehet építeni az inszertet. A **vírusrészecske mérete nem limitált**, mivel a P8 fehérjéből rendkívül sok darab is képes kapszidburkot képezni, ezért a beépíteni kívánt DNS mérete igen nagy lehet. Legfontosabb felhasználása az M13 fág-alapú vektoroknak, hogy egyszálú templátot szolgáltat a Sanger-féle DNS-szekvenáláshoz és az irányított mutagenézis egyes módszereihez. Ezek mellett az M13 fág kiválóan alkalmazható irányított evolúciós módszerekhez, mint például a fág-bemutató (**ld. 16. fejezet**).

7.4. Kozmidok

A kozmid vektorok ötvözik a plazmidok és a λ -fág vektorok előnyeit. A plazmid vektorok megfelelő gazdasejtekben való fenntartásához nincs szükség másra, csak replikációs origóra, antibiotikum-rezisztencia génre és egy poliklónozó helyre. Már említettük, hogy a λ -fág nagy előnye, hogy rendkívül hatékonyan képes a fág részecskébe pakolt DNS sejtekbe juttatására. Ezen kívül a fág részecskében hosszú inszerteket tartalmazó klónok is biztonságosan fenntarthatóak annak veszélye nélkül, hogy az inszert egy része rekombinációs folyamatok következtében elveszik. A kozmid vektorokat úgy alakították ki, hogy a sejtekben plazmidként fenntarthatóak legyenek, ugyanakkor lehetőség legyen a vektorok fertőzőképes λ -fág részecskébe történő becsomagolására, és ez által a rossz hatékonyságú transzformáció fág fertőzéssel történő kiváltására. Ezekhez a funkciókhoz nincs szükség 5-6 kbp DNS-nél többre. Ennek köszönhetően a kozmid vektorokba akár **45 kbp-os inszert** is beépíthető úgy, hogy a keletkező vektor még fág részecskébe pakolható méretű maradjon. A **kozmid vektorok tulajdonképpen I-fág cos szekvenciát tartalmazó plazmidok**. Fontos kiemelni, hogy a kozmid vektorok a cos szekvencián kívül nem tartalmaznak a λ -fág genomjából származó szekvenciát.

A kozmidok rendelkezhetnek akár többféle replikációs origóval is. Ezek megléte határozza meg, hogy az adott kozmid vektor milyen sejtekben és milyen módon szaporítható. A kozmid vektor rendelkezhet például az emlős

sejtekben működő SV40 origóval és ColE1 origóval, ami a kétszálú replikációhoz szükséges prokarióta rendszerben. Bizonyos kozmid vektorok az F-plazmid replikációs origóját tartalmazzák. Ez az origó a plazmid kópiaszámát 1-2 példányban határozza meg. Ezeket a vektorokat kozmidoknak is nevezik. Egyes vektorok tartalmazzák az f1 origót is, ami a vektor egyik (adott) szálának szaporítását teszi lehetővé baktérium sejtekben fonalas fágok (pl. M13 fág) fehérjéinek jelenlétében. A megfelelő pakoló szignállal ellátva ezek a vektorok fonalas fág virionba is becsomagolhatóak.

7.5. Mesterséges kromoszómák (PAC, BAC, YAC, HAC)

A géntechnológia fejlődésével megjelent az igény egy olyan vektor típusra, amelybe genom-szintű analízisekhez szükséges, méretét tekintve igen nagy inszerteket (>100 kbp) lehet beépíteni és klónozni. Ezt a célt elérendő tervezték meg az ún. mesterséges kromoszómákat.

Az első ilyen mesterséges kromoszómához a P1 *E. coli* fág genomja adta az alapot. A P1 fág a λ -fággal ellentétben lizogén életciklusa alatt a saját DNS-ét nem integrálja a gazdaszervezet genomjába, hanem extrakromoszómálisan, episzómaként szaporodik a gazdasejtrel együtt. A vírus genomja 93 kbp méretű kettős szálú lineáris DNS, az *E. coli*-ba bejutva azonban cirkularizálódik. A vírus genomját felhasználva kb. 100-300 kbp mérettartományban lehet idegen DNS-t beépíteni a **P1 fág-alapú mesterséges kromoszómába** (PAC: *P1-derived Artificial Chromosome*). A gyakorlatban a PAC konstrukcióba restriktációs endonukleáz hasítóhelyek segítségével beépítik az inszerteket, majd elektroporációval a sejtbe juttatják.

A **bakteriális mesterséges kromoszóma (BAC)** család az *E. coli* baktérium F-faktor episzómájának módosított változata. A BAC-t szintén nagyméretű **inszertek** befogadására fejlesztették ki, s még a PAC-nál is nagyobb, akár **350 kbp** méretű DNS darabokat is bele lehet építeni. A BAC és PAC közös vonása, hogy mindkét vektor tartalmaz a replikációhoz szükséges bakteriális géneket, valamint szelekciós markereket (pl.: antibiotikum-rezisztencia géneket), amelyek lehetővé teszik a mesterséges kromoszómát tartalmazó sejtek kiválasztását. A replikációjuk a bakteriális kromoszómához kötött, azaz egy sejtben általában csak egy kópiában fordulnak elő („szigorú”, azaz **stringens replikáció**). Ez a tulajdonság nagyméretű rekombináns DNS-ek esetében hasznosabb, mint a nagy kópiaszám, ugyanis így kisebb valószínűséggel történik homológ rekombináció, a konstrukció stabilabb marad a felszaporítás során.

A harmadik mesterséges kromoszóma típus eukarióta eredetű, ugyanis a sörélesztő gomba, a *Saccharomyces cerevisiae* genom tulajdonságait örökölte meg. **Élesztő mesterséges kromoszómákat** (YAC: *Yeast Artificial Chromosome*) használtak eleinte a Humán Genom Program során is, mivel az előbb említett PAC és BAC vektorokkal szemben akár **Mbp** nagyságú DNS **inszertek**et is képes tárolni, így az emberi genom teljes lefedéséhez elegendő kb. 10000 élesztőklón. (Később azonban a YAC klónok instabilitása miatt teljesen elhagyták ezek használatát.) Az előbbi kettő mesterséges kromoszómával szemben a YAC nem cirkuláris, hanem **lineáris** felépítésű. Rendelkezik az eukariótákra jellemző **centromer és telomer régiókkal** is, amelyek megakadályozzák a vektor letörését és más DNS darabbal való kapcsolódását. A YAC tartalmaz minden olyan szakaszt, amely a kromoszóma replikációjához és az utódsejtbe kerüléshez (partíció) szükséges. Ezekben a szakaszokon kívül a YAC vektor rendelkezik még egy, a gazdasejtben működőképes replikációs origóval is. A YAC vektor hátránya a másik két mesterséges kromoszómával szemben, hogy mivel eukarióta gazdasejtben tartják fenn, így a rekombinációs folyamatoknak jobban ki van téve, ezen kívül a konstrukció létrehozásakor gyakrabban alakulnak ki konkatamer inszertek.

Végül meg kell emlékeznünk a **humán mesterséges kromoszómákról** (HAC: *Human Artificial Chromosome*) is, mint a legkomplexebb vektor típusról, amely ígéretes eszköz a génterápia számára is, mint „génszállító” molekula. Megjegyzendő, hogy az első emlős-alapú mesterséges kromoszómát a Szegedi Biológiai Központban Hadlaczy Gyula (1948-2013) állította elő, akinek több évtizedes munkája lehetővé tette, hogy ezt a géntechnológiai eszközt ma már emlős sejtvonalakban gyógyszer-molekulák termelésére, növényekben a bioüzemanyag-előállítására is fel lehet használni (**ld. 15. fejezet**). Mitotikusan stabil HAC-ot létre lehet hozni meglévő kromoszómákból és *de novo* is (telomer és centromer szekvenciák kellenek hozzá, az utóbbiakon funkcionális kinetokórral). Méretük 1-10 Mbp között lehet, és a HAC-ba beépíthető inszertnek potenciálisan nincs mérethatára.

7.6. Élesztő és egyéb eukarióta vektorok és gazdasejtek

Az eukarióta rendszerek bevonása a géntechnológiába hamar megkezdődött, hiszen világos volt, hogy amennyiben a rekombináns DNS-t be lehet juttatni egy eukarióta sejtbe, a baktériumoknál megismert elvek alapján bármilyen sejtben fenn lehet tartani és akár ki is lehet fejteni idegen géneket. Az első nem prokarióta vektor típus a *Saccharomyces cerevisiae* sörélesztőben talált, ún. **2-mikronos** gyűrű nevű **plazmid**, amely egy autonóm módon replikálódó DNS molekula. Van **replikációs origója** (**ARS: Autonomously Replicating Sequence**), és ún. REP génei, amelyek kis extrakromoszomális kópiaszámot biztosítanak úgy, hogy csak akkor választják szét a plazmid replikációt az élesztő genomjának megduplázásától, amikor a plazmid kópiaszáma lecsökkent a sejtben. Ekkor beindul a plazmid független replikációja, amely addig tart, amíg a **kópiaszám** el nem éri a sejtenkénti **30-50** darabot. A 2-mikronos gyűrűt **YE_p**-nek is nevezik, azaz **élesztő episzomális plazmid**nek (*yeast episomal plasmid*), amelyeken belül **két típust** különböztetünk meg. Az első az élesztő **replikálódó plazmid**, **YRp** (*yeast replicating plasmid*), amelyek tartalmazzák a 2-mikronos gyűrűre jellemző ún. **ARS** szekvenciát, ami biztosítja az autonóm replikációt. Ennek a típusnak a hátránya, hogy a sejt osztódásakor nem biztosított a nagy kópiaszám, és a vektor utódsejtbe történő átkerülése. A másik csoport az élesztő **integrálódó plazmid**, **YIp** (*yeast integrating plasmid*), amely a YRp-vel ellentétben a hordozott információt beépíti az élesztő genomjába és azzal együtt replikálódik, valamint osztódáskor az utódsejtbe is átkerül. Az élesztő vektorokkal az expressziós rendszerek kapcsán a **13.1.6. fejezetben** foglalkozunk részletesebben.

A plazmid típusú vektorokon kívül az eukarióta gazdaszervezeteket is transzformálhatjuk **vírus alapú vektorokkal**. A legtöbb vírus alapú rendszer esetén szükséges egy köztes prokarióta átmeneti rendszer használata is, mivel a klónozás mindig *E. coli*-ban történik. Ezeket a kettős replikon típusú vektorokat **shuttle bináris vektornak** hívják.

Az eukarióta vektor-gazda rendszerek között említésre érdemes a rovarsejtek használata bakulovírus, azaz rovar-specifikus vírus-alapú vektorok alkalmazásával. Mivel a fő felhasználási területük a heterológ génexpresszió, részletesebben ott lesz róluk szó (**ld. 13.2. fejezet**).

Növényi sejtek transzformálására (transzgenikus növények előállítására) egy növényi parazita, az *Agrobacterium tumefaciens* tumor indukáló (**Ti**) **plazmidj**át alakítják bináris vektorrá. A Ti plazmid vektorként való alkalmazását részletesen a **14.4.4. fejezetben** tárgyaljuk.

Az **emlős gazdaszervezethez** használt vektorok az esetek többségében vírus eredetűek. A leggyakrabban használt vektorok SV40-, adeno-, retro-, vagy vakcina vírusokon alapulnak. Az **SV40** a legkorábban leírt és legáltalánosabban használt emlős vírus vektor. Mérete kicsi, a gömb alakú kapszidfejen egy 5,2 kbp méretű cirkuláris DNS található. A kis méret előnyös, ha *in vitro* beavatkozást végzünk, de hátrányos olyan szempontból, hogy a genomon nincsenek felesleges részek, melyeket helyettesíteni lehetne a klónozendó DNS-sel. Ennek megoldása az volt, hogy a vírus késői géneit a klónozendó DNS-re cserélték. A késői szakasz hiánya a vírusok termelésének hiányához vezet. Ennek elkerüléséhez a vírustermelést egy **segédvektor** (**helper vírus**) alkalmazásával oldották meg, amely tartalmazza a rekombináns vírusból hiányzó géneket. Ezt követően a vírus lítikus életciklusából adódó sejtpusztító tulajdonságát kellett kiküszöbölni, amelyet úgy oldottak meg, hogy kifejlesztették a nem lítikus SV40 vírustörzset és mellé egy speciális gazdasejt vonalat, amelyben fenntartható a vektor.

Emlős vektorokként **adenovírusokat** is használhatunk. Az adenovírusok az emlős szervezetekben normális körülmények között enyhe lefolyású megbetegedéseket okoznak. A laboratóriumi gyakorlatban az adenovírusokat általában expressziós vektorként használják fel. Alkalmazásuk során a klónozendó fehérje génjét beépítik a vírusgenomba, majd az *in vitro* becsomagolt vírussal fertőzik a gazda organizmust. Az adenovírusok más vírusokhoz hasonlóan géneiket képesek úgy bekapcsolni, hogy nagy kópiaszámban replikálódjanak, ebből adódóan a gének által kódolt fehérje is nagy mennyiségben fog termelődni. A **vakcinia vírus** a bárányhimlő vírus rokona, azonban ártalmatlan és biztonságos, amelyet a gyakorlatban több területen is alkalmaznak. Számos sejtvonalat képes fertőzni és nagy előnye, hogy nagyméretű idegen DNS fragmentumokat lehet beleépíteni. A **retrovírusok** RNS genommal rendelkeznek. A különböző RNS vírusok felépítése igen hasonló, az RNS genom egy középső magban helyezkedik el, ahol a reverz transzkriptáz enzim és különböző szerkezeti fehérjék is megtalálhatók. Ezt a magot lipidburok veszi körül. A retrovírusok vektorként való alkalmazása igen sokrétű, mivel szinte minden állatfajnál alkalmazhatók célgének sejtmagba juttatására, ahol az a gazdasejt genomjába integrálódva kifejeződik. Az emlős gazdasejtekben használható vírus-alapú vektorokról további részleteket a **13.3. fejezetben** olvashatnak.

7.7. További olvasnivaló a fejezethez

Brown TA (2010): Gene Cloning and DNA Analysis: An Introduction. Wiley-Blackwell, 6th Edition. ISBN-10: 1405181737.

Sambrook, J and Russel, DW (2001) Molecular Cloning. A Laboratory Manual. CSHL Press, volume 1-3, Third edition. ISBN: 0879695773.

Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc Natl Acad Sci USA 89: 8794–8797.

Kim JH, Kononenko A, et al (2011) Human artificial chromosome (HAC) vector with a conditional centromere for correction of genetic deficiencies in human cells. Proc Natl Acad Sci U S A. 108:20048-53

8. fejezet - Rekombináns DNS könyvtárak

A gének és az azokról termelődő fehérjék vizsgálata útjában az 1970-es évekig komoly akadályt jelentett az, hogy a kutatók nem tudták a géneket megfelelő tisztaságban és mennyiségben izolálni. A rekombináns DNS technológiák fejlődése az 1970-es évek végére lehetővé tette az egyes gének különböző vektorokba történő klónozását (**ld. 2. és 4. fejezet**). A vizsgálni kívánt gént tartalmazó vektort a megfelelő gazdasejtbe juttatva a gén illetve az arról termelődő fehérje nagy mennyiségben, tisztán előállítható. Ez lehetővé tette akár egész genomok DNS szekvenciájának meghatározását, a gének működésének, szabályozásának megismerését és a génekről termelődő fehérjék biokémiai módszerekkel történő vizsgálatát.

A rekombináns DNS technológiák nemcsak egyedi gének és fehérjék vizsgálatára alkalmas rendszerek létrehozására használhatóak. Ugyanezek a technikák alkalmasak rekombináns DNS könyvtárak létrehozására is. Ezek a könyvtárak hatalmas mennyiségű DNS-ben kódolt információt tartalmazhatnak. A **könyvtár létrehozása során először a klónozni kívánt DNS megfelelő méretű fragmentumait állítják elő, majd ezeket a fragmentumokat klónozzák a választott vektorba**. Az így létrejövő könyvtár tehát a DNS különböző részleteit tartalmazó klónok összessége. Az egyes klónok, amelyek a klónozott DNS más-más részletét tartalmazzák, megfigyelhetők egy hatalmas könyvtárban elhelyezett különböző könyveknek, innen ered a rekombináns DNS könyvtár elnevezés. A rekombináns DNS könyvtárból a vizsgálni kívánt gént tartalmazó klónt vagy klónokat különböző, később részletezett technikák segítségével ki lehet választani, és a többi klóntól el lehet különíteni. Ez az eljárás a **könyvtárak szűrése**, angol eredetű kifejezéssel **szkrínelés**. Ezáltal a vizsgálatokat tiszta klónokon lehet elvégezni. Az egész DNS könyvtár vagy az egyes kiválasztott klónok a megfelelő gazda rendszerben szaporíthatóak, azaz a DNS könyvtár bizonyos értelemben állandó forrása a specifikus klónoknak is.

8.1. A könyvtárak típusai

Ahogy a hagyományos könyvtáraknak, úgy a DNS könyvtáraknak is különböző típusai vannak. A könyvtárak csoportosításának legfontosabb szempontjai a könyvtár készítése során felhasznált DNS forrása, és a felhasznált vektor. Az úgynevezett **genomiális könyvtárak** klónjainak összessége egy adott élőlény teljes genomját, vagy annak egy nagyobb részét, például egy kromoszómát tartalmaz. Az úgynevezett **cDNS (complementary DNS) könyvtárak** egy adott sejttypusban egy adott pillanatban kifejeződő fehérjéket kódoló érett mRNS-ek szekvenciáit tartalmazzák. Egy élőlény különböző típusú sejtjeiben különböző fehérjék fejeződnek ki, így az ezekből készített cDNS könyvtárak is különbözőek lesznek. Egy adott DNS könyvtár tulajdonságait és felhasználhatóságát meghatározó másik fontos tényező az, hogy az adott könyvtár milyen vektor használatával készült. Az egyes vektorok más-más méretű inszertek befogadására alkalmasak. Ez alapvetően meghatározza azt, hogy egy könyvtár milyen jellegű kísérletek során használható. Ebben a fejezetben a különböző típusú könyvtárak létrehozásának lépései, az alkalmazott vektorok, valamint a könyvtárak felhasználásának módjai kerülnek terítékre.

8.2. Genomiális DNS könyvtárak

8.2.1. A genomiális könyvtárak létrehozásáról általában

Genomiális könyvtár létrehozása során az előzőleg izolált **genomiális DNS-t megfelelő méretű fragmentumokra kell darabolni**. A DNS fragmentálása kivitelezhető **restrikciós endonukleázokkal, vagy fizikai módszerekkel**, például ultrahang alkalmazásával (**szonikálás**). A felhasználható fragmentumok méretét a könyvtár készítés során alkalmazott vektor határozza meg, ezért előnyös, ha a folyamat során keletkező átlagos fragmentum méretet a kutató meg tudja határozni.

Meg lehet-e becsülni, hogy adott inszert méretnél és ismert genom méretnél hány rekombináns klónból kell állnia a könyvtárunknak? Igen, sőt azt is megadhatjuk előre, hogy milyen valószínűséggel fogja tartalmazni a könyvtár a teljes genom szekvenciát. Az ideális cél 100%, ám ez nem érhető el, csak megközelíthető. Legyen a kívánt valószínűség 99%, ez gyakorlatilag teljes lefedettséget fog eredményezni. Válasszuk a humán genomot (3×10^9 bp), valamint egy olyan vektort, amelybe 10^5 bp méretű inszerteket lehet beépíteni. Az alábbi képlettel számolunk:

$$N = \ln(1-p)/\ln[1-(I/G)]$$

ahol N a rekombináns klónok száma, p a teljes lefedettség valószínűsége, I az inszert méret és G a genom mérete. A fenti példa esetén:

$$N = \ln(1-0,99)/\ln[1-(2 \times 10^5/3 \times 10^9)] = 70.000$$

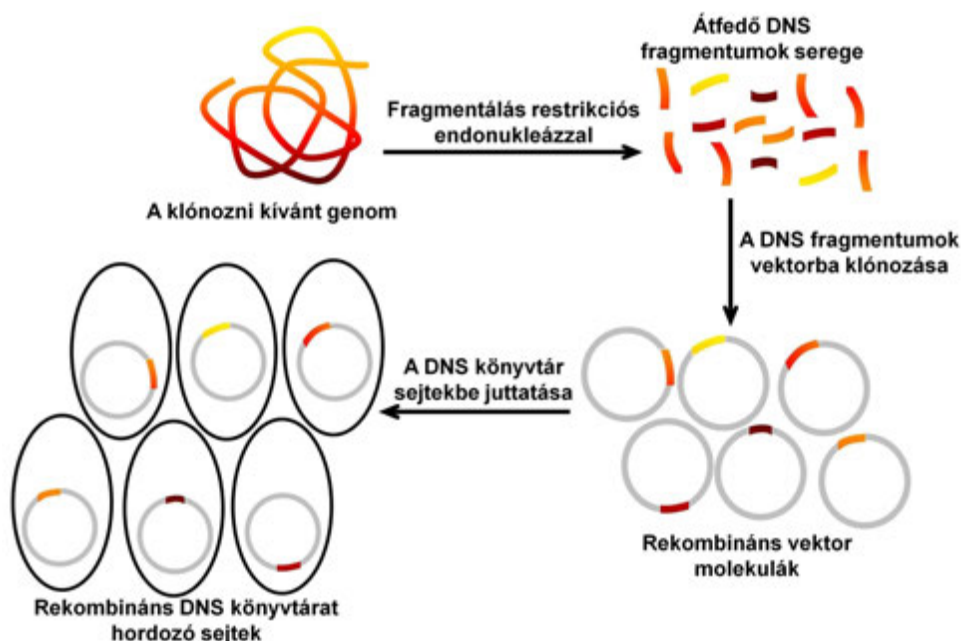
azaz 70 ezer egyedi klónból kell állnia a könyvtárnak, hogy nagy valószínűséggel lefedje a 23 kromoszómánkon tárolt teljes genetikai információt.

A **restrikciós endonukleázokkal** történő fragmentálás esetén a megfelelő enzim kiválasztásával elérhető, hogy a legtöbb fragmentum a kívánt mérettartományban keletkezzen. Az enzim kiválasztásához tudni kell, hogy az egyes enzimek hasítóhelyei eltérő gyakorisággal fordulnak elő a fragmentálni kívánt célpont DNS-en.

Amennyiben a bázisok véletlenszerű eloszlásban fordulnának elő a DNS szekvenciában, úgy egy tetszőleges 4 nukleotidból álló restrikciós hasítóhely 256 bázisonként ismétlődne, egy 6 karakteres megközelítőleg 4 kilobázisonként, míg egy 8 karakteres nagyjából 66 kilobázisonként fordulna elő. A valóságban a helyzet az, hogy a bázisok eloszlása nem véletlenszerű. Például egy GC gazdag genomban egy AT gazdag szekvencia motívumot felismerő restrikciós enzim hasítóhelyei véletlen eloszlás esetén a vártnál távolabb lehetnek egymástól. Az emberi genomban például a 8 karakteres szekvenciát felismerő enzimek közül az SmaI hasítóhelye a véletlenszerű eloszlás esetén vártnál gyakrabban fordul elő, míg az FseI, AscI és NotI enzimek hasítóhelye jelentősen ritkább. A helyzetet tovább bonyolítja, hogy a restrikciós endonukleázok nem egyforma valószínűséggel hasítják a DNS-t az összes hasítóhelynél. Ezt például a szubsztrát DNS komplex szerkezete határozza meg. Ezért az alkalmazott enzim a könnyen hozzáférhető hasítóhelyeket hamar, míg a nehezen hozzáférhető hasítóhelyeket csak hosszabb inkubálás alatt képes elhasítani.

A fent leírtak illusztrálják, hogy számos szempontot figyelembe kell venni a kívánt eredmény eléréséhez. Ilyen szempontok a megfelelő restrikciós enzim kiválasztása, valamint a DNS fragmentálás körülményeinek, nevezetesen a reakció időtartamának vagy az alkalmazott enzim koncentrációjának megválasztása. Mindazonáltal a restrikciós endonukleázok felhasználásával lehetőség van a kívánt méretű és könnyen vektorba építhető DNS fragmentumok előállítására. A genomiális DNS könyvtárak előállítása során általában arra törekednek, hogy a hasítóhelyeknek csak egy részén következzen be hasítás (**részleges fragmentálás**). Így egymással átfedő DNS szakaszok jönnek létre, azaz a keletkező könyvtár tagjai egymással átfedő inszerteket hordoznak. Az egyes klónok által hordozott inszertek közti átfedés a genomiális könyvtárak fontos tulajdonsága.

A megfelelő méretű fragmentumokat ezután a többitől elválasztják. Ez általában agaróz gélelektroforézissel történik. A DNS-t a gélből izolálják, majd a választott vektorba klónozzák. Ezzel létrejön a rekombináns **DNS könyvtár**, azaz **a klónozott DNS különböző szakaszait tartalmazó vektor klónok (illetve a rekombináns DNS-ek) összessége**. A könyvtárat ezután a vektornak megfelelő **gazdaszervezetbe**, általában *E. coli*, vagy ritkábban *S. cerevisiae* sejtekbe juttatják (**ld. 4.3. fejezet**) a könyvtárat alkotó vektorok szaporítása céljából (az utóbbira példa az élesztő két-hibrid rendszer használata, **ld. 17.1. fejezet**). A genomiális könyvtárak létrehozásának általános sémáját a **8.1. ábrán** mutatjuk be.



8.1. ábra: Genomiális könyvtárak létrehozásának általános sémája

8.2.2. Könyvtár vektorok

A könyvtár legfontosabb paramétereit a választott vektor határozza meg. A befogadható inszert méreten kívül a vektortól függ többek között az is, hogy milyen gazda rendszerben tartható fenn és szaporítható a könyvtár. Ezért mindenképpen szót kell ejteni az egyes vektorcsaládok könyvtárkészítés szempontjából releváns tulajdonságairól.

A korábban részletesen bemutatott plazmid vektorok (**ld. 7.2. fejezet**) elvileg alkalmasak bármilyen típusú könyvtár befogadására, azonban van néhány olyan alkalmazhatósági korlátjuk, ami miatt mégsem használjuk őket genomiális könyvtárak készítésére. Az egyik ilyen probléma az, hogy a plazmidokkal való transzformálás hatékonysága nagyon alacsony, és a plazmid méretének növekedésével ez a hatékonyság tovább romlik. Ezen kívül a nagy inszertet tartalmazó plazmidok nem tarthatók fenn stabilan, ugyanis több példányban vannak jelen a sejten belül, és egymás közti rekombinációs folyamatok révén hajlamosak elveszteni az inszert egy részét, illetve átrendezni azt. A plazmidok egyébként alkalmasak nagyobb, akár 10-15 kbp méretű inszertek befogadására, azonban a transzformáció rossz hatékonysága és a nagy inszertek instabilitása miatt kielégítően működő, azaz változatlan információtartalommal fenntartható plazmid könyvtárat csak sokkal kisebb, 3-5 kbp méretű inszertek alkalmazásával lehet létrehozni. A plazmidoknak ezek a hátrányos tulajdonságai ösztönözték a kutatókat a hatékonyabban transzformáló, és nagyobb inszertek befogadására képes vektorok fejlesztésére (**ld. 7. fejezet**).

8.2.3. λ -fág könyvtárak

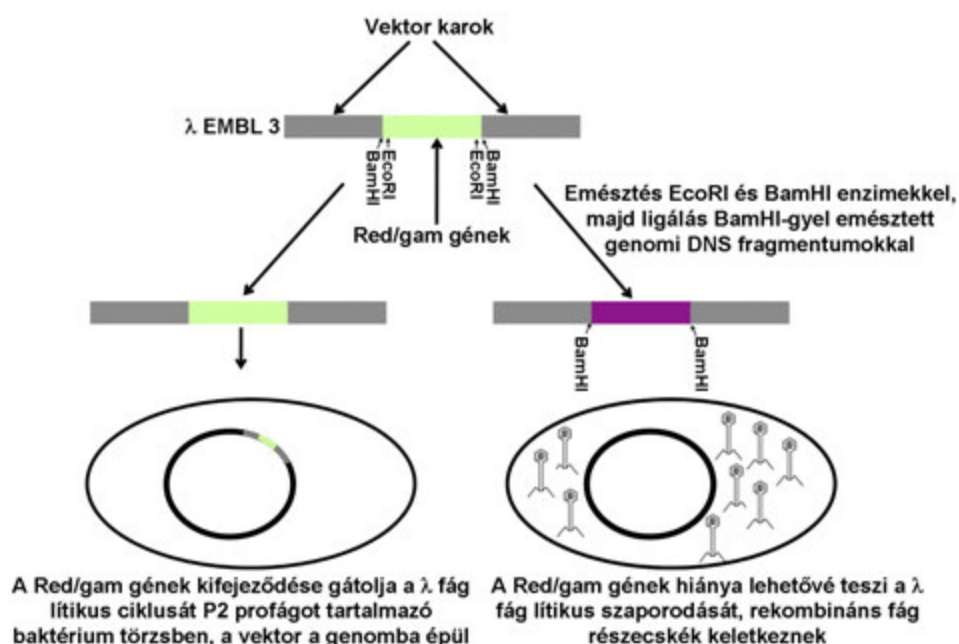
A λ -fágot az 1950-es évek óta intenzíven tanulmányozzák. Életciklusa, génjeinek szabályozása részleteiben is jól ismert. Ez tette lehetővé, hogy módosításával számos, a mikrobiális genetikai, molekuláris genetikai és molekuláris biológiai kutatásokban használható eszközt fejlesszenek ki (**ld. 7.3.1. fejezet**). A λ -fág helyspecifikus (int) és homológ (Red) rekombináns rendszereit elterjedten alkalmazzák rekombináns DNS konstrukciók és genetikailag módosított élőlények előállítására (**ld. 14. és 15. fejezetek**). Ezen kívül a λ -fág genom egyszerű módosításokkal klónozó vektorrá alakítható át. A λ -fág alapú klónozó rendszerek nagy előnye a plazmidokkal szemben az, hogy segítségével a rekombináns DNS konstrukciót a λ -fág fertőzés folyamatának hatékonyságát kihasználva lehet a gazdasejtekbe juttatni. A fág fertőzés hatékonysága a plazmiddel való transzformálásnál kb. három nagyságrenddel nagyobb.

8.2.3.1. λ -fág könyvtárak létrehozása

A λ -fág alapú **helyettesítő** vagy szubsztitúciós klónozó **vektorok 25 kbp méretű inszertek** befogadására képesek annak függvényében, hogy a fág genomjának mekkora, a lítikus ciklusban való szaporodáshoz nem nélkülözhetetlen részét távolították el az adott rendszerben. Genomiális könyvtárak létrehozására alkalmas vektorok pl. a IEMBL3/4, a IDASH II és a IFIX II.

A λ -fág könyvtár létrehozása során a genomiális DNS-t ragadós véget eredményező restrikciós endonukleázokkal hasítják, majd a megfelelő méretű DNS fragmentumokat izolálják. A λ -fág vektort olyan restrikciós enzimmel kezelik, ami a genomiális DNS fragmentumain kialakított ragadós végekkel komplementer végeket eredményez. A karokat ezután megtisztítják a köztes, úgynevezett kitöltő (*stuffer*) szekvenciától. A kitöltő szekvenciára azért van szükség, hogy az „üres”, eredeti állapotú vektor mérete megfelelő legyen ahhoz, hogy az azt tartalmazó fágészecskék fertőzőképesek legyenek.

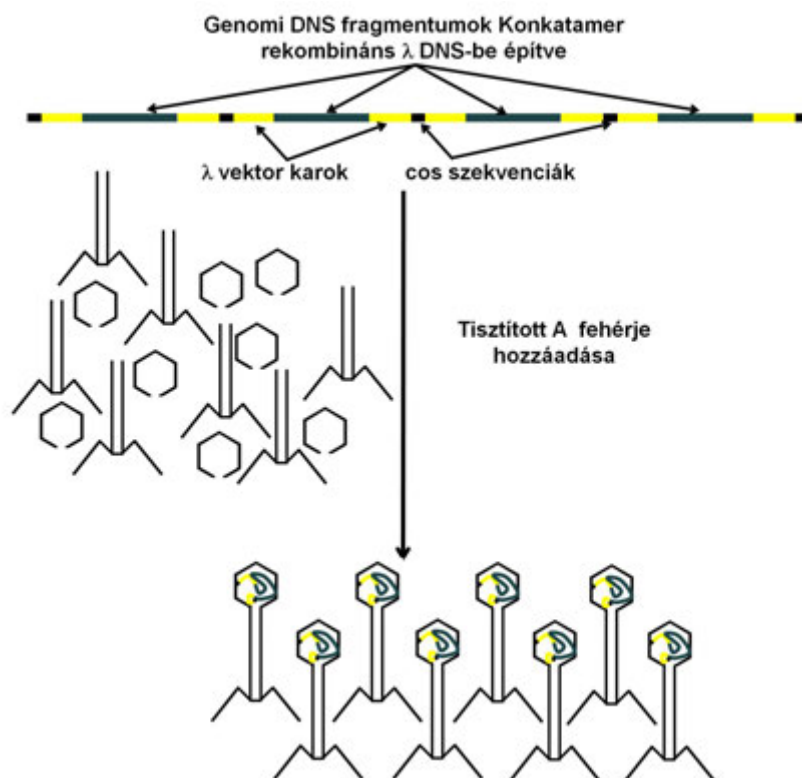
A nagyobb inszertek befogadására képes úgynevezett helyettesítő λ -fág vektorok olyan tulajdonságokkal vannak felruházva, amelyek lehetővé teszik a kitöltő szekvenciát tartalmazó, eredeti állapotú klónok elleni szelekciót. Az egyik lehetőség, hogy a kitöltő szekvenciáról annak vektoron belüli orinotációjától függetlenül termelődnek azok a fehérjék, amelyek jelenléte gátolja a fágok lítikus ciklusát úgynevezett P2 profágot tartalmazó baktérium törzsekben. Ezeket a vektorokat P2 profágot tartalmazó törzsben használva az *E. coli* pázsiton nem képeznek plakkt azok a klónok, amelyek véletlenül a kitöltő szekvenciát tartalmazzák. Ilyen vektort eredeti állapotában előállítani nyilvánvalóan csak P2 profágot nem tartalmazó törzsben lehetséges. Egy másik lehetőség a kitöltő szekvencia újra beépülése elleni szelekciónak az, ha a kitöltő szekvencia két oldalán egymással megegyező poliklónozó helyek vannak egymáshoz képest fordított orientációban. Egy ilyen vektort két, a poliklónozó helyen belül hasító enzimmel kezelve a keletkező karok az egyik, míg a kitöltő szekvencia a másik enzim által generált, egymással nem kompatibilis tapadós végekkel fognak rendelkezni. Az így kezelt vektor karok közé a kitöltő szekvencia nem tud újra beépülni. A λ EMBL3 vektor lehetőséget ad a kitöltő szekvencia újra beépülése elleni szelekció mindkét módjára (ld. 8.2. ábra).



8.2. ábra: Rekombináns λ -fág konstrukcióra történő szelekció

A genomiális DNS-t restrikciós enzimekkel darabolják, majd a fragmentumok DNS számainak 5'-végéről a foszfátsoprotot eltávolítják, hogy a ligálás során eredetileg nem szomszédos szakaszok ne kapcsolódhassanak össze. A vektor karokat és a genomiális DNS-t a kezelést követően összekeverik, és DNS-ligáz segítségével kovalensen összekapcsolják. A reakció során hosszú **konkatamer** DNS molekulák keletkeznek, amelyek a fág lítikus ciklusban való szaporodáshoz szükséges összes génből, a virionba pakolódáshoz szükséges cos szekvenciákból, valamint a vizsgált genom egy-egy részletéből felépülő ismétlődő egységekből állnak.

A DNS könyvtárat ezután *in vitro* fág burokba csomagolják (***in vitro* pakolás; ld. 8.3. ábra**). Ez úgy történik, hogy a könyvtárat összekeverik *E. coli* sejtek kivonatával, ami tartalmazza a preformált fág burok fej és farok egységeket, valamint a DNS fejekbe pakolását elvégző fehérjét (***in vitro* pakoló extraktum**). Ezt a kivonatot olyan fággal (IA⁻) fertőzött sejtekből lehet készíteni, amelyek genomjában mutációval elrontották a DNS virionba csomagolásához nélkülözhetetlen „A” fehérje génje, azaz a burok létrejön a fertőzött sejtben, de a fág DNS nem tud pakolódni. Ezt a sejtkivonatot kiegészítik tisztított A fehérjével, majd a rekombináns DNS-sel. A DNS a cos szekvenciáknál elhasad és a létrejövő molekulák burokba csomagolódnak. Ezzel létrejön a genomális λ -fág könyvtár.



8.3. ábra: Rekombináns λ -fág *in vitro* pakolása

A rekombináns DNS-t tartalmazó fágok könyvtárából hígítási sort készítenek, majd ennek egyes tagjaival *E. coli* sejteket fertőznek és a sejteket 37°C-on inkubálják. A fertőzést követően 10-15 perccel a sejteket 45-48°C hőmérsékletű, 0,6-0,8% agart tartalmazó tápoldatban, ún. **top agarban** felveszik, majd előmelegített agarlemez felszínére terítik szét. Így a sejtek pázsitot alkotnak, amin **tarfoltnak** (más néven **plakk**) nevezett világosabb vagy áttetsző foltok figyelhetők meg ott, ahová a szétterítés során faggal fertőzött sejt került. A fertőzött sejtől ugyanis a sejt pusztulásakor a környezetébe jutnak az új fágok, amik a szomszédos sejteket fertőzik és elpusztítják azokat, miközben újabb fág részecskék termelődnek. Ezért a fertőzött sejtek közelében nem tudnak osztódni a sejtek, ez figyelhető meg plakkok formájában. A plakkokból izolálható fágok egy rekombináns DNS klónt képviselnek. A fertőzéshez használt fág hígítás és a fertőzés után keletkezett plakkok száma alapján meghatározható a **könyvtár fág koncentrációja**, azaz **titere** is, aminek ismeretében kiszámítható, hogy a könyvtárat milyen mértékben kell hígítani, és hány agar lemezt kell készíteni, hogy az összes klón megjelenjen a lemezekon plakkok formájában. Így előállítható a **könyvtár az *E. coli* pázsiton plakkok formájában**.

8.2.3.2. λ -fág könyvtárak szaporítása, a könyvtár fenntartása

A DNS könyvtár összes tagjának nagy mennyiségben való előállításához a könyvtárral *E. coli* sejteket kell fertőzni. A fertőzött sejtek az őket fertőző klónt 100-200 példányban állítják elő, és a frissen termelt fágok újabb sejteket képesek fertőzni. Így a könyvtár tagok könnyen szaporíthatóak. Az viszont korántsem biztos, hogy minden klón ugyanolyan mértékben szaporodik a könyvtár növesztése során. A könyvtárban minden klón egyedi, a baktérium sejteknek különböző nehézségű feladatot jelent ezek szaporítása. A felnövesztett könyvtárban általában vannak az átlagosnál jobban és rosszabbul termelődő, azaz **alul- és felülreprezentált klónok** egyaránt. Egyes klónok akár ki is halhatnak, ha a könyvtár többször is felnövesztésre kerül, ami értékes információ elvesztését jelenti.

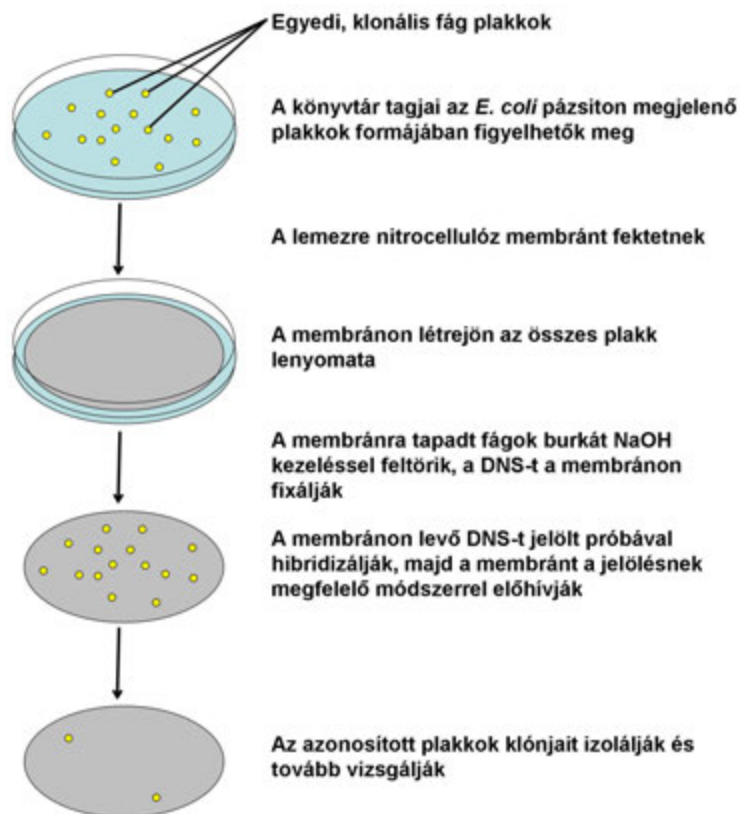
A λ -fág könyvtárak esetében az a legjobb megoldás, hogy a könyvtárat csak egyszer növesztik fel nagy mennyiségben, majd a fágokat izolálják és több adagban lefagyasztják. Így sok kísérletre elegendő mennyiségben és azonos minőségben áll rendelkezésre a könyvtár hosszú időn át. Az együttműködő kutatócsoportok akár meg is tudják osztani egymással a könyvtáraikat, ehhez elég egy lefagyasztott adag fágot elküldeni. Az összes könyvtárra igaz, hogy megfelelő, glicerinnel kiegészített oldatban lefagyasztva hosszú ideig tárolhatóak.

8.2.3.3. A λ -fág könyvtárak szűrése (szkrínelése) plakk hibridizációval

A korai könyvtárakat, és a modern könyvtárak nagy részét is azért állítják, állították elő, hogy adott géneket és azok szabályozó régióit izolálják további vizsgálatok céljából. A keresett gén szerencsés esetben akár több klón inszertjében is jelen lehet teljes hosszúságában, azonban gyakran előfordul, hogy a keresett információt több, egymással átfedő inszertet tartalmazó klón együttese tartalmazza. Ez az egyik oka annak, hogy általában arra törekednek a könyvtárkészítés során, hogy az inszertek egymással átfedjenek.

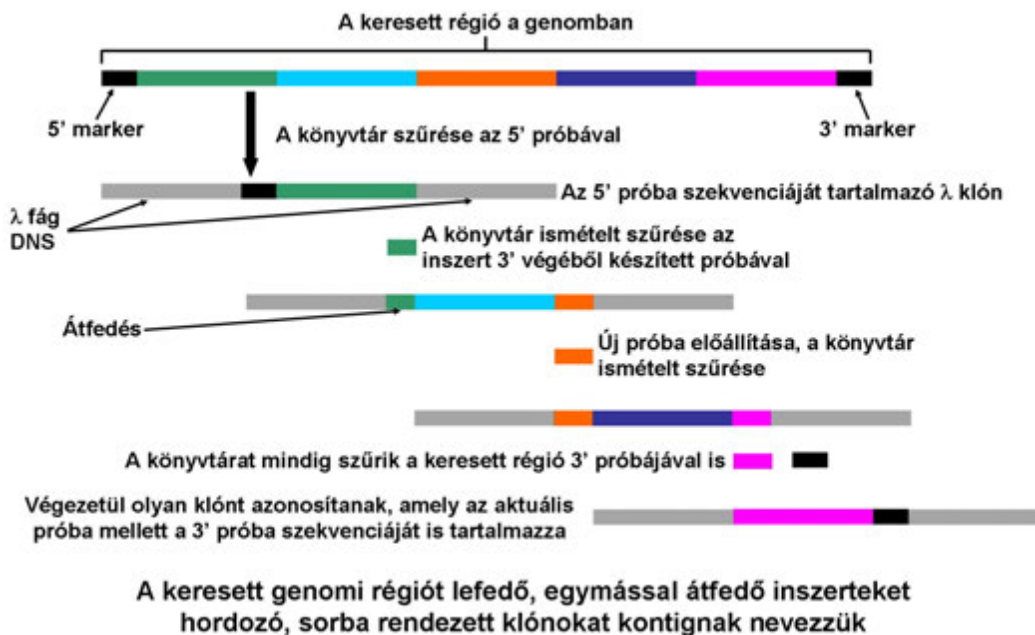
A könyvtárat **szűrni (szkrínelni)** kell, hogy a rengeteg klón közül azonosítsuk azokat, amelyek a keresett információ egészét vagy részleteit tartalmazzák. A szkrínelés gyakran alkalmazott módja specifikus egyszálú **DNS vagy RNS próbák** használatán alapul és **plakk (tarfolt) hibridizációnak** hívjuk (ld. 8.4. ábra). A szűrés első lépéseként a fág könyvtárat *E. coli* sejtekbe kell juttatni, és olyan sejtpázsitot tartalmazó lemezeket kell készíteni, amelyeken 150-200-fág plakk keletkezik. Ezt követően a lemezekre **nitrocellulóz membránt** fektetnek. Minden plakkból kerülnek fágok a membránra, így azon létrejön az eredeti lemez lenyomata. Ezután a membránokat nátrium-hidroxiddal kezelik, hogy a fágok fehérje burkát és a DNS-t denaturálják. A DNS-t ezt követően hevítéssel fixálják a membránokon. Az előkészített membránokat radioaktívan vagy fluoreszcensen jelölt DNS vagy RNS próba oldatába áztatják. A próba szekvenciája komplementer a keresett szekvencia egy részletével, tehát a megfelelő klónok DNS-éhez hibridizál. Radioaktív próba esetén a membránt röntgenfilm segítségével hívják elő, míg fluoreszcens próba esetén a megfelelő hullámhosszon történt gerjesztést követő emissziót detektálják az egyes pontokon. A röntgenfilmen megjelenő, vagy fluoreszcens jelet adó pontok pozíciója a membránon megegyezik az eredeti plakk pozíciójával az agar lemezen, azaz a keresett szekvenciát, vagy annak egy részét tartalmazó klón vagy klónok ezzel a módszerrel azonosíthatók. A módszerhez szükség van a keresett gén szekvenciájának vagy a keresett génhez a genomban közel elhelyezkedő marker ismeretére.

A kiválasztott fágokkal baktérium sejteket fertőznek és friss fágokat termeltetnek, hogy az általuk hordozott genomiális DNS szakasz megfelelő mennyiségben és minőségben álljon rendelkezésre a további vizsgálatokhoz.



8.4. ábra: Plakk hibridizáció

Ha nincs jelen olyan klón a kiválasztottak között, amelyek a teljes keresett szekvenciát tartalmazza, akkor a „**kromoszóma séta**” (*chromosome walking*) módszerével azonosítanak és sorba rendeznek több olyan átfedő inszertet tartalmazó klónt, amelyek együttesen tartalmazzák az összes keresett információt (ld. 8.5. ábra).



8.5. ábra: Kromoszóma séta

A kromoszóma séta során szükség van két ismert markerre, amelyek a keresett szekvencia két végén találhatóak. Az egyik próbával szkrinelik a könyvtárat, és olyan klónokat azonosítanak, amelyek a keresett információnak ezt a részét tartalmazzák. Ezután restrikciós térképezéssel feltérképezik az adott fragmentumot, és az első próbától legtávolabbi végével komplementer próbát készítenek, amivel újra szűrnek a könyvtárat. Így újabb klónokat találnak, amelyek az új próba szekvenciáját tartalmazzák. A folyamat során kiválasztott, egymással átfedő és sorba rendezett klónok összességét kontignak nevezzük.

8.2.4. Kozmid könyvtárak

A λ -fág vektorok és a felhasználásukkal készített könyvtárak lehetővé tették, hogy 20-25 kbp átlagos inszert méretű könyvtárakat használjanak a genetikai és genomikai kutatások során. Azonban számos gén hossza ennél jelentősen nagyobb lehet, ezért továbbra is fennállt az igény nagyobb inszertet befogadni képes vektorok fejlesztésére, és könyvtárak készítésére való felhasználására. Erre a célra először a kozmid vektorokat hozták létre (ld. 7.4. fejezet).

8.2.4.1. Kozmid könyvtárak létrehozása

Az üres, eredeti állapotú kozmid vektor *E. coli* sejtekben plazmidként szaporítható, és a plazmidokkal azonos módszerekkel izolálható. A könyvtárkészítési folyamat során a vektort restrikciós endonukleázokkal felnyitják, majd a keletkező 5'-végekről a foszfát-csoportot eltávolítják. A vektort ezután egy másik enzimmel is megemésztik. A reakció során vektor karok jönnek létre, amelyek végeiken cos szekvenciákat tartalmaznak. A vektor molekulákat méret szerint elválasztják a nem megfelelően hasadt DNS molekuláktól. Ezután az emésztett vektort összekeverik a klónozásra előkészített, azaz a megfelelő enzimmel kezelt és defoszforilált genomiális DNS-sel, és ligálják a komplementer DNS végeket. A ligálás során a λ -fág vektorokhoz hasonlóan konkatamer DNS molekulák jönnek létre. Ezekhez a hosszú DNS molekulákhoz preformált λ -fág fej és fark egységeket és a becsomagoláshoz szükséges fág fehérjéket tartalmazó sejt kivonatot adnak. Ekkor a hosszú DNS molekulák a **cos helyeknél elhasadnak és egyesével fág burokkba pakolódnak**. A megfelelő méretű inszertet tartalmazó vektorok becsomagolódása fertőzőképes fág részecskét eredményez.

A rekombináns kozmid vektorokat tartalmazó fágokat, azaz a kozmid könyvtárat izolálják, majd azokkal baktérium sejteket fertőznek. A sejtben a vektor kör alakú molekulává záródik, és plazmidként replikálódik. Sajnos a kozmid vektorok esetében is fennáll annak a veszélye, hogy a nagy inszertek egy része elveszik, amikor a könyvtár tagjai sejtekben vannak fenntartva. A rendszernek megvan viszont az az előnye, hogy ez a veszély a fág burokkba csomagolt könyvtár esetében nem áll fenn.

8.2.4.2. Kozmid könyvtárak szaporítása, a könyvtár fenntartása

A kozmid könyvtár klónjainak sokszorosításának első lépéseként a klónokat tartalmazó vírusokkal *E. coli* sejteket fertőznek, majd rövid növesztés után a sejteket a vektornak megfelelő antibiotikumot is tartalmazó agar lemezekre szélesztik. A lemezeket 37°C-on inkubálják. Az antibiotikum szelekciónak köszönhetően csak a kozmid vektor klónnal transzformált sejtek tudnak osztódni a lemezeken. Másnap a lemezekről a sejteket lemossák a megfelelő tápoldattal és a kozmid vektorokat hordozó sejteket tartalmazó táplevest egy csőben összegyűjtik és alaposan összekeverik. Ezt követően glicerinnel kiegészítik a mintát, szétosztják több adagra, lefagyasztyják és -80°C-on tárolják.

A kozmid könyvtárakat baktérium sejtekben, agar lemezek felszínén fenntartani, de a fent leírt módon, „ömlesztve” amplifikálni nem könnyű feladat. Egyrészt a nagy inszerteket hordozó klónok instabilak, másrészt a könyvtárban maradt nem rekombináns, eredeti állapotú kozmiddal, vagy kisebb inszertet tartalmazó vektorral transzformált sejtek kisebb terhet viselnek, ezáltal gyorsabban osztódnak a nagy rekombináns vektort hordozó társaiknál. Vizsgálatok kimutatták, hogy egy 2-3% eredeti vektort tartalmazó kozmid könyvtárban egy amplifikációt követően a sejteknek akár 40%-a eredeti vektort hordozhat. Ezt kiküszöbölni rendkívül sok munkát igénylő módon lehet. Egy lehetséges módszer az összes klón megőrzésére az, hogy az elkészített könyvtárat sejtekbe juttatják, és a sejteket szélesztik agar lemezekre úgy, hogy a lemezekre egyedi, klonális telepek nőjenek. A telepeket ezután egyesével leoltják 96 lyukú sejtenyésző lemezek táplevestel töltött mintahelyeire. Egy nagy könyvtár esetében több tízezer klónt kell kézzel leoltani, de így biztosítható, hogy minden klón fennmaradjon. A lemezek összes mintahelyéről időnként a sejtek egy része új lemezekre átvihető, és így a könyvtár bizonyos ideig fenntartható. Ez utóbbi művelet robotok segítségével automatizálható. Összességében elmondható, hogy a kozmid könyvtárak

esetében megfontolandó lehet a könyvtár ismételt elkészítése bizonyos időközönként, vagy akár minden újabb szkrínélést megelőzően.

8.2.4.3. A kozmid könyvtárak szűrése kolónia hibridizációval

A kozmid könyvtárak szkrínélése során az agar lemezek felszínén növekedő baktérium telepek egy részét nitrocellulóz membránra viszik át, majd a membránt néhány példányban sokszorosítják. A telepek lenyomatát tartalmazó egyik membránt NaOH oldattal kezelik, hogy a sejtek feltáródjanak, majd a DNS-t kémiai módszerrel, vagy hevítéssel fixálják a membránra. A fixált membránokhoz radioaktívan vagy fluoreszcensen jelölt DNS próbákat hibridizáltatnak, így azonosítják a keresett információt tartalmazó klónokat. Az eljárást **kolónia hibridizációnak** nevezzük. A folyamat tulajdonképpen a legtöbb lépésében nagyon hasonlít a λ -fág könyvtárak plakkk hibridizációjához.

A szkrínélés során pozitívnak bizonyult klónok felszaporítása általában nem magától értetődő feladat a kiindulási lemezeken tapasztalt nagy kolónia sűrűség miatt. Ilyen esetekben az eredeti membránra az autoradiogram vagy az előhívott membránról készült fluoreszcens fotó alapján azonosítják azt a régiót, ahol a pozitív klón megtalálható. Ezután a kiválasztott régióban található sejtekkel táptalajt oltanak be és a sejteket szaporítják. A sejteket ezután hígítják és agar lemezen szélesztik úgy, hogy kis klónsűrűségű lemezek keletkezzenek, azaz a másnapra megjelenő telepeket alkotó sejtek egyetlen kozmid klónt hordozzanak. A kis klónsűrűségű lemezeket ezután ismét szkrínelik. A második szűrés során azonosított pozitív klónok már izolálhatóak tiszta formában.

A fent leírt két ciklusban történő szkrínélésre azért van szükség, mert ha a könyvtár összes klónját akarnák széleszteni úgy, hogy klonális telepek nőjenek fel a szűrendő lemezekre, akkor a lemezek száma csillagászati lenne. Összességében sokkal kevesebb munkát jelent az összes klónt szűrni nagy klónsűrűségű lemezekre, majd a pozitív klónokat a fent leírt módon megkeresni a szűrés során kiválasztott régiókban jelen levő klónok közül.

A λ -fág és a kozmid könyvtárak összehasonlítása során a legfontosabb szempontok szerint arra a megállapításra juthatunk, hogy a kozmid könyvtárak a nagyobb befogadható inszert méretének köszönhetően feleannyi klónnal lefedhetnek egy adott genetikai információt.

8.2.5. Mesterséges kromoszóma könyvtárak (BAC és YAC)

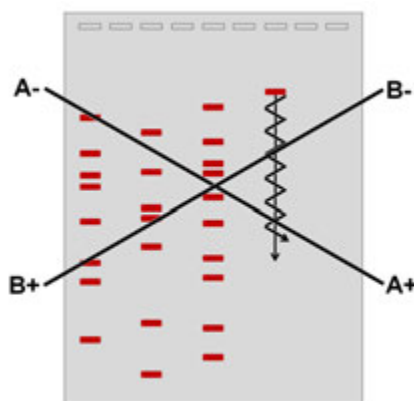
Az eddig bemutatott rendszerek által befogadható inszert méret még mindig nem elegendően nagy ahhoz, hogy belátható és könnyen kezelhető mennyiségű klón létrehozásával a vizsgált genomot, vagy annak kiválasztott részét reprezentáló könyvtárakat lehessen készíteni, valamint hogy a leghosszabb géneket, együtt szabályozott géncsoportokat egészében hordozó klónokat lehessen előállítani. Ezt az igényt a mesterséges kromoszóma vektorok kifejlesztése és genomális könyvtárak készítésére történő alkalmazása elégítette ki. A **mesterséges kromoszóma vektorok** sokkal nagyobb, **néhány száz kbp (BAC), vagy akár két-három Mbp (YAC) méretű inszertek befogadására alkalmasak**. Ez többszöröse a korábban bemutatott rendszerek által befogadható inszert méretének

8.2.5.1. Mesterséges bakteriális kromoszóma (BAC) könyvtárak létrehozása

A BAC vektorok az *E. coli* F-plazmidjának replikációs origóját tartalmazzák (**ld. 7.5. fejezet**). Ennek köszönhető, hogy a BAC vektor példányszáma szigorúan szabályozott (stringens), és egy sejtben az osztódást megelőző replikációt leszámítva csak egyetlen példányban fordul elő. A BAC könyvtárakat elterjedten alkalmazzák a genom projektek során a genom nagyobb részleteinek amplifikálása, fenntartása, a genom fizikai térképezése és szekvenálása során (**ld. 9.1.2. fejezet**).

A BAC könyvtárak létrehozása során alapvetően a korábban már ismertett technikákat használják. A nagy inszert méret miatt ugyanakkor bizonyos dolgokra nagyobb figyelmet kell fordítani. A genomális DNS izolálása során kerülni kell, hogy a DNS-re ható nyíróerők következtében, vagy bármi más okból a DNS károsodjon. A DNS szálon bekövetkező nem célzott törések rontják a klónozás hatékonyságát és csökkenthetik az elérhető maximális inszert méretet is. A restriktions endonukleázzal kezelt DNS fragmentumok méret szerinti elválasztása is más módszerrel történik a száz kbp-os mérettartományban. A kívánt **100-300 kbp tartományban** a DNS molekulákat

nem lehet méret szerint elválasztani hagyományos gélelektroforézis technikák segítségével. A megoldás a **pulzáló-terű gélelektroforézis** (PFGE: *pulsed field gel electrophoresis*) technika alkalmazása (ld. 8.6. ábra).



8.6. ábra: Pulzáló-terű gélelektroforézis

A PFGE során az alkalmazott feszültség iránya adott időközönként változik. Az egyik legegyszerűbb elrendezésben a gél hossz tengelyére merőleges irányon kívül a tengelytől mindkét irányban 60°-os szöget bezáró feszültséget alkalmaznak még. A módszer alapja, hogy a feszültség irányának megváltozásakor a kisebb molekulák gyorsabban reagálnak, és kezdenek el a megváltozott iránynak megfelelően mozogni, ezáltal gyorsabban haladnak. Mivel a különböző irányokban egyforma ideig alkalmazzák a feszültséget, ezért a haladás iránya összességében a gél tengelyével párhuzamos marad.

A megfelelő mérettartományba eső DNS molekulák izolálására is több lehetőség van. Az egyik, a DNS molekulákat tartalmazó agaróz géldarab kezelése a polimerizált agarózt lebontó agaráz enzimmel. A DNS-re ható nyíróerők ezzel a módszerrel kizárhatóak. Egy másik módszer az úgynevezett elektroelúció. A DNS-t az izolálást követően koncentrálni is kell. Ez általában dialízissel történik a nyíró erők fellépésének elkerülése miatt.

A BAC vektorokat a klónozást megelőzően restriktációs endonukleázzal felnyitják, majd a DNS végekről a foszfát csoportot alkalikus foszfatáz enzimmel eltávolítják, hogy megelőzzék a vektor önzáródását (ld. 3.4.2. fejezet). A kezelt vektort és a genomiális DNS fragmentumokat ezután összekeverik és ligálják. A könyvtár sejtekbe juttatása vagy kémiai transzformálással, vagy a létrejövő vektorok nagy mérete miatt elektroporálással történik.

8.2.5.2. BAC könyvtárak szűrése és fenntartása

A BAC könyvtárakat a kozmid könyvtárakhoz hasonlóan lehet szkrínálni. A különböző eljárások során jelölt DNS próbákat hibridizáltatnak a klónok membránon fixált DNS-éhez, majd a kiválasztott klónokat izolálják, és tiszta formában szaporítják. A BAC könyvtárak fenntartására szintén alkalmazhatóak a kozmid könyvtárak esetében ismertetett módszerek. Az eredeti agar lemezek lenyomatait tartalmazó membránok, vagy a klonálisan, sejtenyészítő lemezekon növesztett klónok formájában a könyvtár hosszabb időn keresztül, megfelelő minőségben fenntartható.

8.2.5.3. Élesztő mesterséges kromoszóma (YAC) könyvtárak és létrehozásuk

A YAC vektorok a BAC vektorokhoz hasonlóan eredeti állapotukban *E. coli* sejtekben fenntartható, plazmid szerű cirkuláris DNS molekulák (ld. 7.5. fejezet). Tartalmaznak bakteriális és élesztő replikációs origót, szelekciós markereket a YAC-ot hordozó baktérium illetve élesztő sejtek szelekciójához, valamint multiklónozó helyet is. A **YAC vektoroknak** rendelkezniük kell a **telomer** és **centromer** régiókból származó szekvenciákkal is, ezek ugyanis nélkülözhetetlenek a genomiális DNS szakaszt tartalmazó YAC vektor élesztősejtekben való stabil fenntartásához és az osztódáskor történő pontos szétválásához. A YAC vektorok által befogadható **2-3 Mbp méretű inszertek** ideális eszközké tehetnék őket a genom projektek, genomok térképezése és szekvenálása során, azonban számos olyan hátrányos tulajdonsággal rendelkeznek, ami megnehezíti a YAC könyvtárakkal való munkát.

A YAC könyvtárakban tárolt információ jóval kevésbé stabil, mint a BAC könyvtárakban tárolt. A YAC könyvtárakban gyakran előfordul, hogy az egyes klónok által hordozott inszert nem a genom egyetlen, folytonos

régiójának felel meg, azaz kiméra. Ennek oka lehet az, hogy egyetlen YAC vektorba több inszert is beépül a ligálás során, vagy az, hogy a könyvtárral való transzformálás során egyetlen élesztősejtbe több YAC vektor is bejut, és a vektorok között rekombinációs folyamatok játszódhatnak le. Ezen kívül az inszert egy részének delécioja vagy inverziója is gyakran megfigyelhető. Egyes vizsgálatok szerint a YAC könyvtárakban a klónok akár felére igaz lehet, hogy a hordozott inszert nem felel meg egy folytonos genomiális régiónak. Ez a jelenség kísérletes műtermékeket eredményez, és téves következtetéseket okozhat. A fent felsorolt hátrányok ellenére bizonyos esetekben a YAC vektorba történő klónozás az egyetlen megoldás. Az élesztőben például szélsőségesen magas A-T tartalmú DNS szakaszok stabilabbak, mint bakteriális rendszerekben.

8.2.5.4. YAC könyvtárak fenntartása és szűrése

A YAC könyvtárakat 96 vagy 384 mintahelyes sejtenyésző lemezeken lehet fenntartani. Ehhez a könyvtárral transzformált és táplemez felszínén növesztett élesztő klónokat egyesével le kell oltani és a sejtenyésző lemez mintahelyeire kiosztott táptalajba kell oltani. Adott időközönként a sejtenyésző lemezen nőtt sejtek egy részét új sejtenyésző lemezekre kimért friss táplevesbe viszik át. A könyvtár ilyen formában 4°C-on néhány hétig tartható el, vagy a mintahelyeket glicerinnel kiegészítve a lemezek le is fagyaszthatóak, így a könyvtár hosszabb ideig is eltartható.

A YAC könyvtárak szkrinelésének leghatékonyabb módja szintén DNS próbák hibridizálásán alapul. A hibridizálás során használt membránt agar lemezek felszínére fektetik, majd a sejtenyésző lemezek fenntartott könyvtár klónokat automatizáltan, robotok segítségével felviszik a membránra meghatározott elrendezésben. A klónokat az agar lemezen növesztik, majd a membránt eltávolítják. A membránon az összes klón jelen van szabályos elrendezésben. A membránon levő sejteket ezután a korábban leírtakhoz hasonlóan feltárják, majd a DNS-t denaturálják és fixálják a membránon. A próba hibridizálása és előhívása után azonosíthatóak a membránon és a sejtenyésző lemezen azok a klónok, amelyek a próba szekvenciáját hordozzák.

8.2.5.5. A BAC és YAC rendszerek összehasonlítása

A mesterséges kromoszóma rendszerek összehasonlítása alapján megállapítható, hogy a könyvtárkészítés során kritikus transzformálási hatékonyság a BAC könyvtárak esetében jelentősen, mintegy két nagyságrenddel jobb, mint a YAC könyvtáraknál. A cirkuláris BAC vektorok kompakt, szuperhelikális szerkezetüknek köszönhetően egyszerűen izolálhatóak, míg a lineáris és nagyobb méretű YAC vektorok gyakran károsodnak a tisztítási folyamat során. A YAC vektorok előnye viszont, hogy még a BAC vektorokhoz képest is hatalmas a befogadható inszertek mérete, valamint bizonyos esetekben a DNS YAC vektorba építve és élesztő sejtekben fenntartva stabilabb, mint bakteriális rendszerben.

8.2.6. Genomiális könyvtárak felhasználási területei

A genomiális könyvtárakat széles körben használják genetikai és genomikai kutatásokban, kiemelten a **genom projektekben**. A felhasználás célja lehet egy gén, gén klaszter vagy akár a teljes genom fizikai térképének elkészítése, új markerek azonosítása és elhelyezése a térképen, vagy akár a teljes genom szekvenciájának meghatározása. Bármilyen alkalmazás során az első feladat, hogy az éppen keresett információt tartalmazó klónt vagy klónokat a korábban bemutatott szűrési technikák segítségével megkeressék és izolálják. A szkrinelés során kiválasztott klónok által hordozott inszertek vizsgálatának első lépése gyakran a **restrikciós térképezés**. Ez a módszer hasznos információkat szolgáltat az inszertől a genom szekvenálása során, és akkor is, ha a cél a genom egy adott régiójának azonosítása, amplifikálása és további vizsgálata. A restrikciós térképezés folyamán információ gyűjthető arról, hogy az adott inszertben bizonyos restrikciós endonukleázoknak hány hasítóhelye van, ezek milyen sorrendben, és egymástól milyen távolságban helyezkednek el. Ez az információ azért hasznos, mert a hasítóhelyek helyzetének ismerete egy korábban nem térképezett DNS szakasz **fizikai térképének** első ismert pontjait adhatja. Ezek a pontok hatalmas segítséget nyújthatnak az inszert több részletben történő szekvenálását követően az egymással átfedő szekvenciák sorba rendezésében, ha a létrehozott restrikciós térkép felbontása megfelelő. (ld. 9.1.2. fejezet)

8.3. cDNS könyvtárak

A cDNS könyvtárak készítésének célja alapvetően más, mint a genomális könyvtárak esetében. A **cDNS könyvtárak** a könyvtárkészítés során használt **sejttípusban kifejeződő** fehérjék érett **mRNS szekvenciáiról** készült komplementer szálat tartalmaznak. Ilyen könyvtárakat leginkább adott fehérjék kódoló szekvenciájának előállítására hoznak létre. A cDNS könyvtárból klónozott eukarióta gén intronokat nem tartalmaz. Ennek köszönhetően lehetséges a fehérje bakteriális expressziós rendszerben történő előállítása. A cDNS könyvtárak nem alkalmasak az egyes gének kifejeződésének, az expresszió szabályozásának vizsgálatára, mert a génekhez tartozó szabályozó elemek nincsenek jelen. A cDNS könyvtár elkészítése előtt körültekintően, előzetes ismeretek alapján kell megválasztani azt a sejttípust, amelyben a vizsgálni kívánt fehérjék nagy mennyiségben jelen vannak. A sejttípus megfelelő kiválasztása növeli annak az esélyét, hogy a keresett gének jelen lesznek a könyvtárban.

8.3.1. cDNS könyvtárak létrehozása

A cDNS könyvtárak létrehozásának első fontos lépése a megfelelő sejttípus vagy szövet, illetve az alkalmazott vektor kiválasztása. A cDNS könyvtárak létrehozása során gyakran használt vektorok közé tartoznak az inszerciós és szubsztitúciós λ -fág vektorok, valamint a plazmidok és kozmidok. A cDNS könyvtárak átlagos inszert mérete nem indokolja a mesterséges kromoszóma vektorok használatát.

A kiválasztott sejtekből izolálják az RNS-t, majd az érett mRNS-ek 3' poli-A farkához hibridizáló oligo-dT primer és **reverz transzkriptáz enzim felhasználásával** elkészítik az mRNS-ekkel **komplementer DNS (cDNS)** szálat. A cDNS molekulák vektorba klónozásához szükség van a létrejövő RNS/DNS hibrid molekulák kétszálú DNS megfelelőinek előállítására. Erre is több módszer áll rendelkezésre. Az egyik leggyakrabban alkalmazott eljárás során a reverz transzkripció termékéhez RN-áz H és DNS-polimeráz enzimet adnak. Az RN-áz H *nick*-eket (bevágásokat) hoz létre az RNS szálon. A DNS-polimeráz a *nick*-eknél megjelenő szabad 3'-OH csoportoktól indulva emészt az RNS szálat és létrehozza a DNS másik szálát.

A létrejövő kétszálú DNS-t a vektorba ligálást megelőzően még további kezelésnek kell alávetni. A molekulák végein az egyik szál esetleges túlnyúlását egyszálú-specifikus DNS-exonukleáz enzimmel emésztik, vagy DNS-polimeráz enzimmel feltöltik. A DNS vektorba ligálásához a terméket restrikciós endonukleázzal kell hasítani. Ehhez el kell látni a DNS molekulák végeit a megfelelő enzim hasítóhelyével, valamint biztosítani kell, hogy az enzim csak a végeken hasítson. Ennek érdekében először metil-transzferáz enzimmel kezelik a DNS-t, ami az alkalmazandó restrikciós endonukleáz hasítóhelyeit metilálja, ezáltal megóvjva azokat a későbbi hasítástól. Ezt követően a DNS molekulák végeire tompa vég ligálással linkereket illesztenek, amelyek tartalmazzák a restrikciós endonukleáz metilálatlan hasítóhelyet. A választott vektort és a cDNS-t restrikciós enzimmel kezelik, majd az inszerteket a vektorba ligálják. A vektor önmagával való záródásának megakadályozására alkalmazható a vektor foszfátáz enzimmel való kezelése. A cDNS könyvtárkészítés sémáját a **8.7. ábrán** mutatjuk be.



8.7. ábra: cDNS könyvtár készítése

8.3.2. cDNS könyvtárak szűrése és fenntartása

A cDNS könyvtárak fenntartásának módját az alkalmazott vektor típusa határozza meg. A különböző vektorok felhasználásával készült könyvtárak fenntartásával kapcsolatos tudnivalók a genomiális könyvtárakkal foglalkozó fejezetben találhatók.

A cDNS könyvtárak **szkrínelésére** is alkalmasak a korábban leírt, **jelölt DNS vagy RNS próbák hibridizálásán** alapuló eljárások. Bizonyos vektorok lehetőséget adnak a könyvtár szűrésének más módjaira is. Az egyik lehetőség az, hogy a vektorba ligált **inszerctegy promóter mögött** helyezkedik el, transzkripcióval **átíródik** mRNS molekulákba és végül fehérje formában megjelenik a sejtben. Ilyen esetekben lehetőség van **specifikus, jelölt ellenanyaggal** azonosítani azokat a könyvtár kolóniákat, amelyekben termelődik a klónozni kívánt fehérje, vagy annak az ellenanyag által felismert részlete.

A cDNS könyvtárban előforduló szekvenciák nagy része nem tartalmazza a fehérjét kódoló mRNS teljes szekvenciáját, csak annak egy részét. A teljes hosszúságú cDNS előállításához vagy megkereséséhez a könyvtárat újra kell szkrínelni. Ennek során általában a korábban már azonosított, de nem teljes hosszúságú klón szekvenciájából készített próbát használják. Számos lehetőség van annak megállapítására, hogy az azonosított klón vagy klónok szekvenciája tartalmazza-e a teljes mRNS szekvenciát. Talán a legegyszerűbb az 5' és 3' nem-átíródó régiók meglétének vizsgálata. Ha valamelyik hiányzik, akkor biztos, hogy a megtalált információ még nem tartalmazza a teljes mRNS szekvenciát. Ugyanakkor fontos számolni azzal is, hogy egy adott génről alternatív *splicing* (kivágódás) folyamatoknak köszönhetően többféle mRNS is keletkezhet. Az azonosított és nem-átíródó régiókkal is rendelkező szekvenciákról meg lehet állapítani, hogy a keresett fehérje génjének felelnek-e meg, vagy esetleg alternatív *splicing* termékek. Ha ismert a fehérje molekulatömege, akkor ellenőrizhető, hogy az azonosított szekvenciáról termelődő fehérje tömege megegyezne-e ezzel.

8.3.3. cDNS könyvtárak felhasználási területei

A cDNS könyvtárak nagymértékben megkönnyítik a vizsgálni kívánt gének fehérje kódoló szekvenciájának izolálását. Az izolált cDNS szekvenciák később változatos kutatásokban, sokféle céllal felhasználhatók. A cDNS könyvtárakat széles körben használják arra, hogy a vizsgálni kívánt fehérjék génjeit olyan formában izolálják, ami alkalmas a fehérje expressziós rendszerben való előállítására.

Az izolált cDNS szekvenciákból jelölt próbák készíthetők. A próbák segítségével vizsgálható az adott gén expressziójának szintje különböző szövetekben, vagy az expresszió megváltozása különböző kezelések, beavatkozások hatására. Ilyen próbákat széles körben alkalmaznak genomiális DNS könyvtárak szkrínelése során is. Így lehetőség nyílik olyan klónok azonosítására, amelyek a vizsgált fehérje génjét szabályozó régiókkal és intronokkal együtt tartalmazzák. Így a gén elhelyezhető az élőlény genomjának fizikai térképén, valamint lehetőség adódik bizonyos fenotípusok, például öröklődő betegségeket markereivel való kapcsolat vizsgálata. A génszabályozó régiókkal együtt történő izolálása lehetővé teszi a génkifejeződés szabályozásának vizsgálatát is.

8.4. További olvasnivaló a fejezethez

Deaven L. L. (1992) Recombinant clones for mapping and sequencing. Los Alamos Science 20:218-249.

She K. (2003) So you want to Work with Giants: The BAC Vector. BioTeach Journal 1:69-74.

Seidman J. G. (editor) (2003) Construction of Recombinant DNA Libraries. Current Protocols in Molecular Biology, Ch 5.

9. fejezet - Genom projektek

A **genom** egy adott organizmusban megtalálható örökletes információ összessége ami általában a DNS-ben, de pl. egyes vírusok esetén az RNS-ben kódolt információ. A **genom projektek** olyan tudományos vállalkozások, amelyek célja, hogy egy adott organizmus **teljes genomját** meghatározzák, vagyis az organizmus teljes DNS-ét szekvenálják. A genom meghatározásának következő része, hogy a **genomhoz biológiai információt rendelnek** hozzá, a genomot **annotálják**. Ennek során meghatározzák, hogy a genom mely szakaszai tartalmaznak fehérjék kódolásáért felelős géneket, vagy egyéb fontos szakaszokat (pl. promóterek, enhanszerek stb). Ebben a fejezetben bemutatjuk a teljes genomok vizsgálatára irányuló kutatások történetét, a legfontosabb technológiai újításokat, és genom projektek néhány eredményét.

9.1. A genomszekvenálás alapjai

A genom szekvenálása, vagyis nagyon hosszú, összefüggő DNS szakaszok szekvenciájának meghatározása jóval komplexebb feladat, mint egy rövid DNS darab bázissortrendjének meghatározása. Ennek oka, hogy a szekvenálás eredményeként kapott néhány száz bázispár hosszúságú DNS szakaszokból össze kell állítani a nagyságrendekkel nagyobb genom (az egy-egy kromoszómának megfelelő DNS molekulák) teljes szekvenciáját. Az összeállításhoz alapvetően két technikát különböztethetünk meg. Az egyik, korábban kifejlesztett és nagyon munkaigényes technika a **térkép-alapú módszer**. A másik az úgynevezett **teljes genom shotgun** („sörétes puska”) **módszer**, amely egyszerűbb, de nagyon komoly számítástechnikai kapacitást igényel. A térkép-alapú módszert szokták „felülről-lefelé” (*top-down*) módszernek is hívni, mivel ebben az esetben a nagyobb egységből (genom vagy kromoszómák térképezése) indulnak ki. A *shotgun* módszert ellenben „alulról-felfelé” (*bottom-up*) módszernek is nevezik, mert a genom meghatározása gyakorlatilag vakon történik, néhány száz bp hosszúságú szekvencia részletek (*read*-ek) összerakásából áll össze a genom. Nagyméretű genomok esetén a két technika kombinációja a leghatékonyabb.

9.1.1. Térképezés

A **térkép-alapú** (vagy más szóval **klón-kontig**) **módszer** esetén a szekvenálás előtt elkészítik a genom fizikai és/vagy genetikai térképét, vagy már meglevő térképet használnak. A genom szekvenálás során a térkép segít a szekvenált DNS darabok genomban történő elhelyezésében. A fizikai térképezés egyik módszere, hogy a genomot, vagy nagyméretű genomok esetén a genom feldarabolásával létrehozott óriási DNS szakaszokat **restrikciós enzimmel** (pl. HindIII) **hasítják**, és a kapott darabokat gélelektroforézissel elválasztják. Az eredményként kapott **hasítási mintázat** („ujjlenyomat”: *DNA fingerprint*) elárulja, hogy mely DNS szakaszok tartalmaznak egymással átfedő részeket (hiszen ezeken a helyeken az ujjlenyomat megegyezik), és ez alapján meg lehet mondani, hogy az adott DNS szakasz hol helyezkedik el a genomban. A **fizikai térképezés** így **abszolút módon meghatározza**, hogy egyes DNS szakaszok egymáshoz képest hogyan helyezkednek el, és hol találhatók (ld. 9.1. ábra).

Fizikai és genetikai térképezés

A) Fizikai térkép restrikciós hasítóhelyek alapján

I. Nagyméretű, térképezett klónok



II. Összeállított kontig



I Hasítóhely

B) Genetikai térkép génmarkerek alapján

GÉNEK

m - mini szárnyak
V - cinóber szemek
W - fehér szemek
Y - sárga test



Rekombinációs frekvenciák

m - v: 3%
m - y: 33.7%
v - w: 29.4%
w - y: 1.3%

Gének pozíciója a térképen

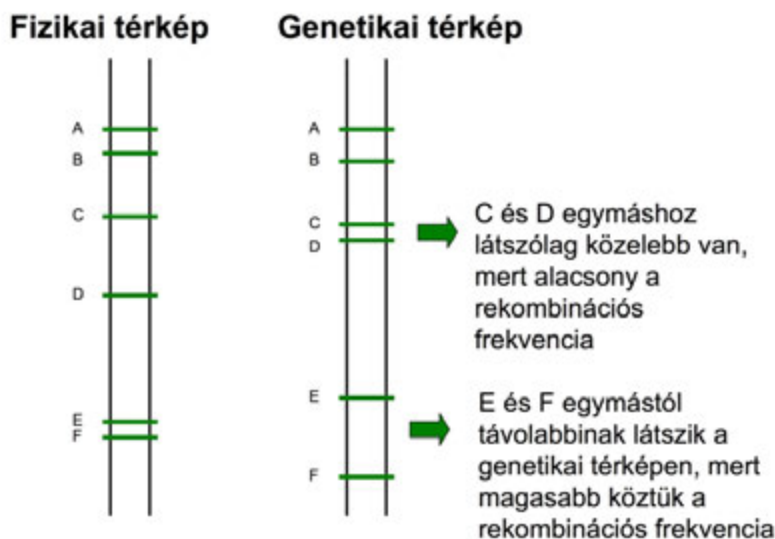


9.1. ábra: A genom szekvencia összeállítását jelentősen megkönnyíti, ha rendelkezésre áll genetikai és/vagy fizikai térkép. A) Fizikai térképezésre használhatók restrikciós endonukleáz hasítóhelyek. A nagyméretű klónok hasítási mintázata alapján összeállítható egy klón kontig, és az egyes szakaszok helye a genomban kijelölhető. B) A gének és DNS markerek rekombinációs gyakorisága alapján megállapítható az egymáshoz viszonyított helyzetük.

A **genetikai térképezés** ezzel szemben az egyes DNS szakaszok egymáshoz viszonyított **relatív helyzetéről** ad információt. Az egymáshoz közel elhelyezkedő gének illetve régiók nagyobb valószínűséggel öröklődnek együtt, mint az egymástól távolabbiak. Vagyis az **egymáshoz közeli DNS szakaszok rekombinációs frekvenciája alacsony, míg a távolabbiaké magasabb**. Több generáció vizsgálatával és megfelelő DNS markerek használatával feltérképezhető különböző gének egymáshoz viszonyított távolsága. Ilyen markerek lehetnek fenotípusosan megjelenő tulajdonságok is (ld. 9.1. ábra). Tipikus DNS **markerek** a restrikciós fragment hosszúság polimorfizmusok (**RFLP**: *Restriction Fragment Length Polymorphism*), a szekvencia hosszúság polimorfizmusok (**SSLP**: *Simple Sequence Length Polymorphism*), vagy az egyedi nukleotid polimorfizmusok (**SNP**: *Single Nucleotide Polymorphism*). Az RFLP-k esetén egy restrikciós endonukleáz hasítási mintázatát vizsgálják a két allélon. Ha ezek eltérnek, vagyis az egyik allélon hiányzik a hasítóhely egy mutáció miatt, az adott helyet markerként lehet használni. Az SSLP-k valamilyen ismétlődő szekvenciárészletet tartalmaznak eltérő számban az allélokon, míg az SNP-k egy bázispár különbséget jeleznek (pl. az egyik allélon C, a másikon T található).

A genetikai és a fizikai térképezés során kapott térkép, a két módszer közötti elvi különbség miatt, a legtöbb esetben jelentősen különbözik (ld. 9.2. ábra).

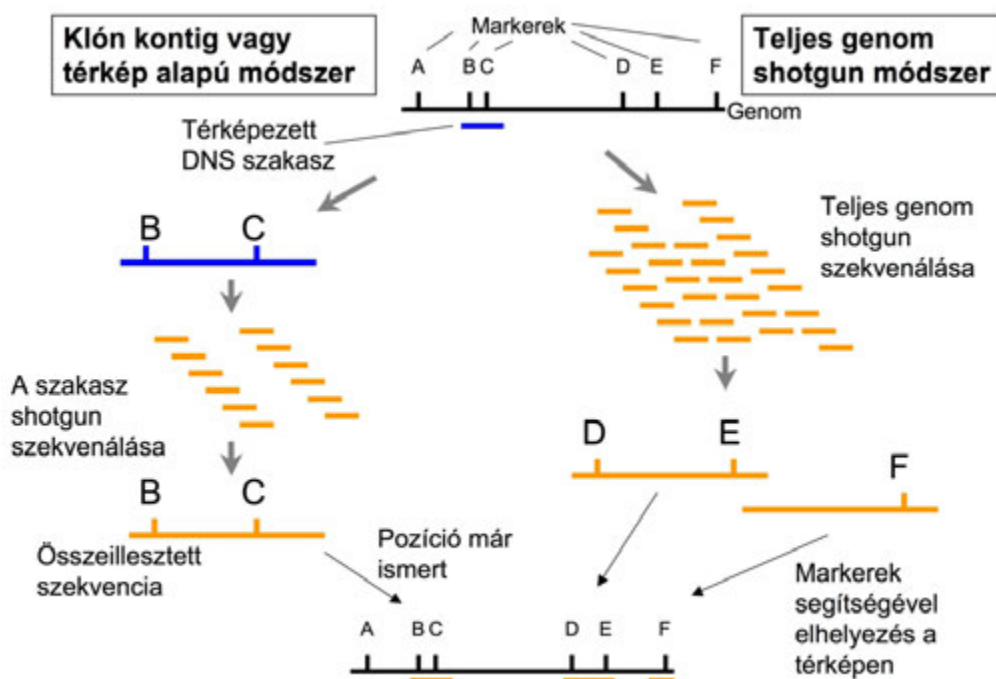
A fizikai és a genetikai térkép különböző lehet



9.2. ábra: A **genetikai és fizikai térkép viszonya**. A genetikai és fizikai térképek nem teljesen egyformák, mert a genetikai térkép csak a markerek egymáshoz viszonyított helyzetéről ad információt.

9.1.2. Genomszekvenálás térkép-alapú módszerrel

A térkép-alapú szekvenáláskor a genomot nagyméretű, 40-200 kbp méretű darabokra hasítják. Az így kapott DNS fragmentumokat ezután olyan vektorokba klónozzák, amelyek képesek ilyen méretű szakaszokat is befogadni. Tipikusan ilyen vektorok a **BAC vektorok**, de kozmidok vagy YAC vektorok is használhatók (ld. 8.2.5. fejezet). Mivel a genom hasítása véletlenszerűen történt, a kapott DNS fragmentumok is véletlenszerűen találhatók meg a BAC-könyvtárban, és a fragmentumok egymással átfedő szakaszokat tartalmaznak. A következő lépés tehát az, hogy megállapítsák, mely DNS darabok fednek át egymással. Ennek során a könyvtárban található klónok „ujjlenyomatát” készítik el (ld. 9.1.1.). Ha ez rendelkezésre áll, már meg lehet határozni, hogy a sok DNS fragmentumból melyek azok, amelyeket felhasználva fel lehet építeni a teljes genomot. Ezt angol kifejezéssel „*minimum tiling path*”-nak hívják (magyarul a „legrövidebb kirakható út”). Miután a fizikai térkép elkészült, a hatalmas, már térképezett inszerteket tartalmazó **BAC klónokat shotgun módszerrel szekvenálják**. Ehhez a klónokat véletlenszerűen tovább darabolják maximum 1500 bp nagyságú szakaszokra, szubklónozzák (pl. M13 fág vagy plazmid vektorba), majd szekvenálják. A kapott szekvenciákat az átfedő régiók alapján egymáshoz illesztve meg lehet határozni az eredeti inszertek szekvenciáját, majd a térkép alapján a teljes genomszekvenciát (ld. 9.3. ábra).



9.3. ábra: **A genomszekvenálás módszerei.** A klón-kontig stratégia esetén a kiinduló nagyméretű inszert helyzete ismert a genomon belül. Ezt *shotgun* módszerrel szekvenálják, majd a genomban elhelyezik. A kontigok genombeli pozícióját a rajtuk található markerek segítségével próbálják meghatározni. A markerek nagyobb száma mindkét esetben jelentősen segítheti a végső szekvencia összeállítását.

Az összeillesztés természetesen nem mindig egyszerű. Ha az egyes DNS szakaszokat a rendelkezésre álló információk alapján nem lehet összeilleszteni, itt a szekvencia egy lyukat, „gap”-et tartalmaz. Azokat a nagyméretű, térképezett DNS szakaszokat, amelyeket az átfedő klónok alapján, „gap-mentesen” össze lehet illeszteni, **klón-kontignak** nevezik (a *contig* elnevezés a *contiguous*, magyarul összefüggő szóból ered). A klón-kontigok némileg különböznek a szekvencia kontigoktól, amikről a következő fejezetben lesz szó. Ezt a módszert használták a sörélesztő (*Saccharomyces cerevisiae*) és a *Caenorhabditis elegans* nevű fonálféreg genomjának meghatározásához, illetve a Humán Genom Program során is ezt a technikát használta az egyik kutatócsoport.

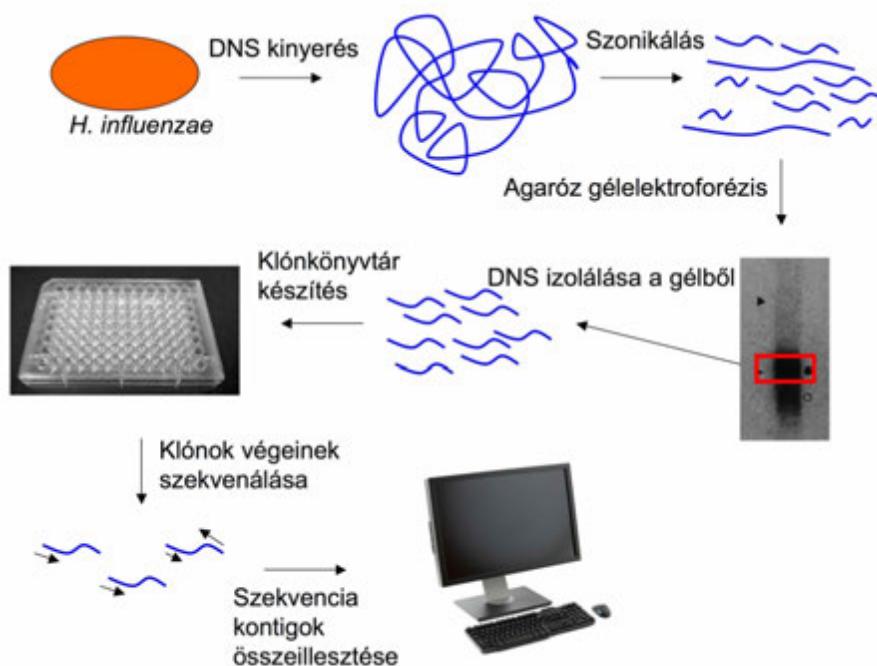
9.1.3. A *shotgun* módszer

A *shotgun* módszer esetén **előzetes, térképezésből származó információ nélkül történik a szekvenálás**, ezért sokkal gyorsabb. A DNS-t itt is feldarabolják, de jóval kisebb szakaszokra. Készítenek egy kb. 2 kbp és egy kb. 10 kbp darabokból álló könyvtárat. Mivel a rövidebb inszertek miatt nincs szükség speciális vektorokra, **plazmid vektorokat használnak**, és a végeredmény egy sok százezer vagy akár millió tagból álló plazmid könyvtár. Az így kapott klónokat szekvenálják, és az átfedő szekvenciariészletek alapján **szekvencia-kontigokat** állítanak össze, majd a teljes genom „mester” szekvenciáját. Fontos megemlíteni, hogy nagyméretű, komplex genomok esetén a végső szekvencia összeállításánál térképek felhasználása nagymértékben segíti a munkát (ld. 9.3. ábra).

9.2. Egy példa: a *Haemophilus influenzae* baktérium szekvenciájának meghatározása

Eleinte a kutatók kételkedtek abban, hogy egyáltalán lehetséges a nagyszámú egyedi szekvenciából egy teljes genomot sikeresen összeállítani. 1995-ben azonban Craig Venter és munkatársai publikálták a *Haemophilus influenzae* baktérium 1830 kb hosszúságú genomját, amelyet **shotgun módszerrel** határoztak meg (ld. 9.4. ábra). A genom szekvenálásának első lépéseként a baktérium DNS-ét szonikálással (magas frekvenciájú hanghullámok

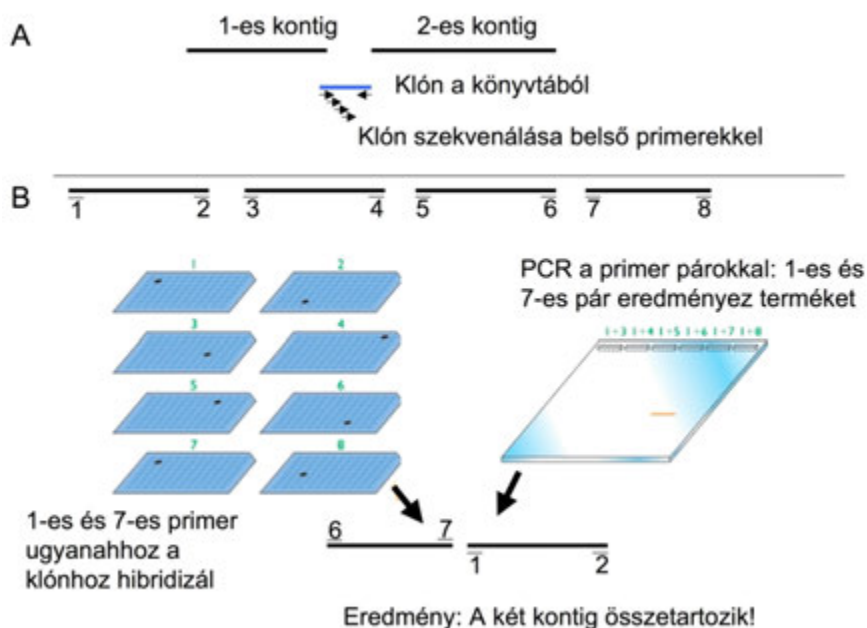
segítségével) kis darabokra törték. Az így kapott DNS szakaszokat elektroforetizálták, és 1,6-2 kb nagyságú darabokat izoláltak, majd klónoztak, létrehozva a **plazmid könyvtár**at. Ezután összesen 19.687 klónt véletlenszerűen kiválasztva 28.643 szekvenálást végeztek, melynek során néhány száz bázispár hosszan tudták leolvasni az egyes klónok szekvenciáját (vagyis nem az egész klón bázissorrendjét határozták meg, csak a klónok „végét”). Többnyire egyetlen irányból történt a szekvenálás, azonban a szekvenálások körülbelül egyharmadát a klónok másik végéről is elvégezték (lásd 19.687 klón és kb. 30%-kal több szekvenálási reakció). Ez a megoldás komoly segítséget jelentett a teljes szekvencia összeállításában. Az összes szekvenálási reakció 16%-át sikertelennek ítélték, mert a kapott szekvencia kevesebb, mint 400 bázispárt tartalmazott. A maradék 24.304 szekvenálás végül 12.631.485 leolvasott bázispárt jelent, ami mintegy **hatszorosa a *H. influenzae* genomjában található összes bázispárnak**. Ekkora ráhagyás (~6x) szükséges is ahhoz, hogy teljes lefedettséget lehessen elérni a genom meghatározásakor. Ezek után következett a kapott szekvenciák összeillesztése, vagyis a rövid szekvenált DNS szakaszok átfedő részeit felhasználva a szekvencia-kontigok összeállítása, ami a számítástechnika akkori fejlettségét figyelembe véve, komoly feladat volt. A konkrét kísérletben 30 órán keresztül dolgozott a számítógép, és végül 140 szekvencia-kontigot sikerült meghatározni.



9.4. ábra: A **shotgun módszer** a *H. influenzae* példáján bemutatva. A baktériumból izolált genomális DNS-t szonikálással darabolják, majd agaróz gélelektroforézissel elválasztják, hogy 1,6-2 kb nagyságú darabokat tudjanak izolálni. Ezekből készül a klón könyvtár, amelyből nagyszámú, véletlenszerűen kiválasztott klónt szekvenálnak.

A szekvencia kontigok összeillesztésével készül el a genom teljes szekvenciája.

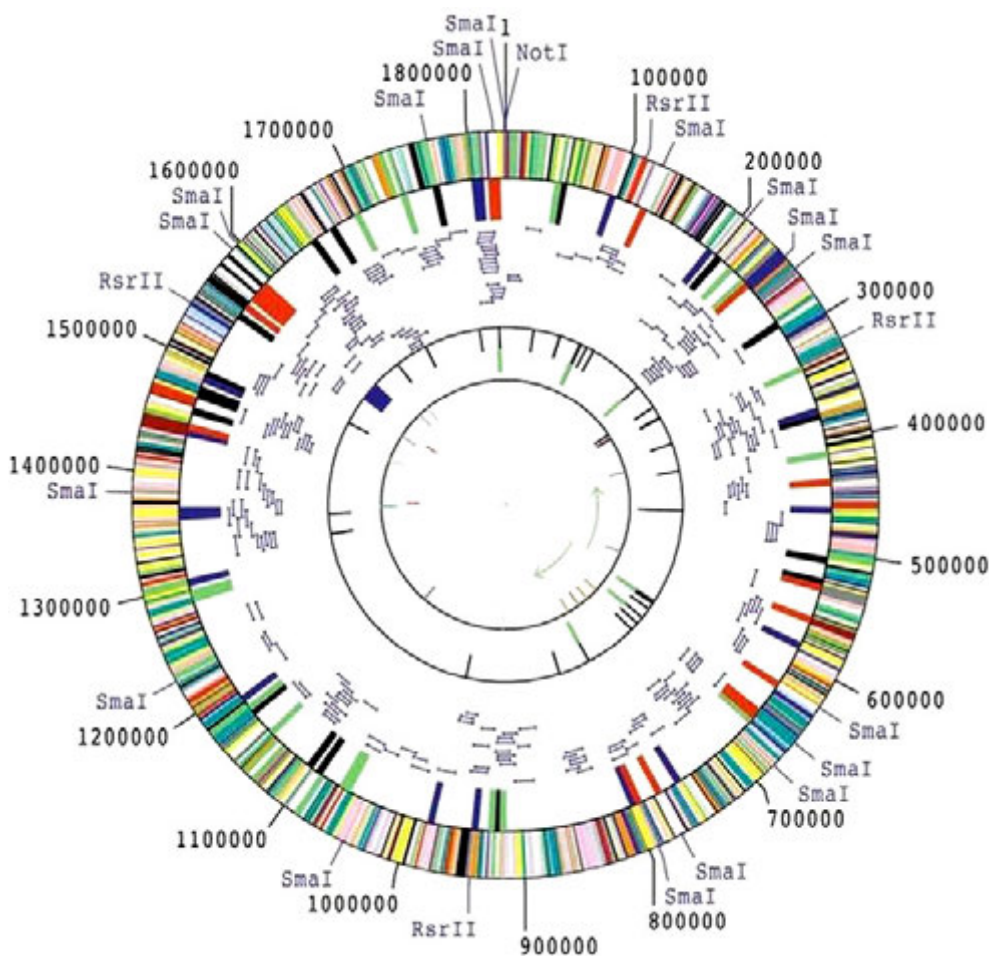
Természetesen ezek után a feladat a **kontigok közötti lyukak betömése**. A hiányzó DNS szakaszok egy részét viszonylag egyszerűen meg lehetett határozni. Ehhez újra végignézték (szkrínelték) a könyvtár összes klónját, és olyan klónokat kerestek, amelyekben egyszerre található meg két különböző kontig végszekvenciája. Ekkor ugyanis elegendő volt végigszekvenálni az egész klónt, és meg lehetett határozni a hiányzó részletet a két kontig között. Ezzel a módszerrel a 140 lyukból 99-et sikerült is betömni. Még mindig maradt 42 lyuk, amelyet meg kellett szüntetni. Mivel valószínűleg a lyukakat tartalmazó DNS szakaszok nem voltak jelen a könyvtárban (mert instabil klónokat eredményeztek), **új könyvtárat hoztak létre**, λ -fág vektort használva. A 42 *gap* szélein található szekvenciákat felhasználva 84 oligonukleotidot szintetizáltak, amelyekkel végigpróbálták az új könyvtárat. Ha egy klónhoz két oligonukleotid hibridizált, akkor azt végigszekvenálták, mivel az adott klónnak tartalmaznia kellett két kontig közötti szekvenciát. Ilyen módon további 23 lyukat sikerült betömni (ld. 9.5. ábra).



9.5. ábra: **A genom szekvenálás befejező lépése a szekvencia javítása és a lyukak eltüntetése.** A) A szekvencia lyukak javítására a könyvtárban megtalálható klónok közül azokat lehet használni, amelyek két kontig végszekvenciáját tartalmazzák. B) Fizikai lyukak esetén a még nem összeállított kontigok végszekvenciájáról szintetizálnak primereket, és ezeket egy teljesen új könyvtár klónjaihoz hibridizáltatják. Az ugyanazon klónhoz hibridizáló primerek jelzik, hogy az adott klón tartalmazza a hiányzó szekvenciárészletet (bal oldali panel). A másik lehetőség, hogy az oligonukleotidokat párba állítva sok PCR-t végeznek, és a keletkező terméket szekvenálják.

A hiányzó szakaszok meghatározásához ugyanezt a 84 oligonukleotidot használták. Egyrészt véletlenszerűen összepárosították az oligonukleotidokat, majd PCR elvégzésével megnézték, hogy kapnak-e terméket a *H. influenzae* genomját templátként használva. Ha igen, akkor a PCR termék szekvenálásával meg lehetett határozni egy-egy hiányzó DNS szakaszt. Mivel a lehetséges oligonukleotid párok száma túl magas lett volna ahhoz, hogy mindegyiket mindegyikkel kipróbálják, egy másik módszert is alkalmaztak. A primereket olyan *H. influenzae* DNS szakaszokhoz hibridizáltatták, amelyeket a genom restrikciós endonukleázzal történő emésztésével állítottak elő. Azon oligonukleotidok, amelyek azonos restrikciós fragmentumhoz kötődnek, nagy valószínűséggel egymáshoz közeli DNS szakaszokhoz tartoznak. Az így kapott primerpárokkal PCR-t végeztek, majd a terméket szekvenálták. Végül az összes hiányzó DNS szakaszt sikerült meghatározni, és ezzel megszületett a *H. influenzae* teljes genomja (**ld. 9.6. ábra**).

A fenti példával sikerült bebizonyítani, hogy a *shotgun* stratégia működőképes és nincsen szükség előzetes információra a genomról ahhoz, hogy a szekvenciáját meg lehessen határozni. Az azóta eltelt évek során ezres nagyságrendű prokarióta szervezet genomját szekvenálták, köztük olyan fontos patogénét, mint pl. a TBC-t okozó *Mycobacterium tuberculosis*, vagy a kolera, a tüdőgyulladás, az anthrax (lépfene) kórokozóját. A kapott genomszekvenciák alapján meghatározhatók a kórokozók génjei által kódolt fehérjék, és új terápiás megoldások kidolgozására nyílik lehetőség.



9.6. ábra: A *H. influenzae* genomtérképe.

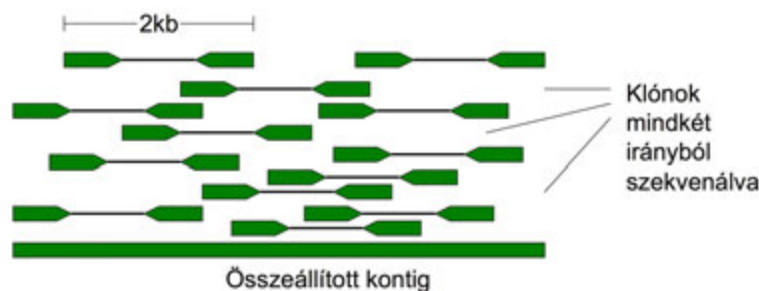
9.2.1. A lefedettség (coverage) jelentősége

A *H. influenzae* által bemutatott példán már említettük, hogy az összes megszekvenált bázispár körülbelül hatszorosa volt a baktérium teljes genetikai állományának. Általánosságban elmondható, hogy a szubklón könyvtárak létrehozásának fontos szempontja, hogy a klónok száma **optimális lefedettséget (coverage)** vagy más szóval többletet (*redundancy*) biztosítson. A két kifejezés arra utal, hogy a szekvenálás során az eredeti genom méreténél jóval nagyobb számú bázispár meghatározására van szükség ahhoz, hogy a teljes genom reprezentálva legyen. A lefedettség jellemzésére egy számot szoktak megadni. Az 1x lefedettség nem jelenti azt, hogy a teljes genom minden régiója szekvenálásra került. Mivel a klón könyvtárakból véletlenszerűen kerülnek kiválasztásra a szekvenálandó klónok, előfordul, hogy ugyanazt a szakaszt többször is megszekvenálják, míg más régiók teljesen kimaradhatnak. A kutatók **jellemzően 6x-10x lefedettséget** alkalmaznak, ezzel biztosítható, hogy a genom 99,8-99,99%-a legyen szekvenálva. További előnye a nagy feleslegnek, hogy így megnövekszik az átfedő szekvenciák száma, ami a kontigok összeállításában jelent komoly segítséget.

9.2.2. Paired-end szekvenálás

Ugyancsak a *H. influenzae* példáján láttuk, hogy a szekvenálási reakciók száma nagyobb volt, mint az összes kiválasztott klón száma, mert a szekvenálások körülbelül egyharmadát a másik irányból is elvégezték. A szekvencia *read*-ek egymáshoz rendelése, egymással való összekapcsolása komoly feladatot jelent, és rengeteg hibalehetőséget is rejt magában, elég csak például az ismétlődő szekvenciariészletek okozta problémákra gondolni. A könyvtárakat úgy hozzák létre, hogy az ezekben található DNS szakaszok körülbelül egyforma hosszúságúak legyenek. **A paired-end szekvenálás során a klónozó vektor mindkét vége felől készül egy szekvencia.** Ennek hatalmas előnye, hogy mivel az adott DNS szakasz hossza ismert, azt is tudni lehet, hogy a két *read* milyen távolságra van egymástól. A teljes szekvencia összeállítása során figyelembe tudják venni ezt a távolság információt, így ha egy *paired-end*

szekvenálási reakció két termékét a számítógép egymástól százezer bázispárra helyezi, akkor egyértelműen kiderül, hogy hiba történt az egymáshoz rendelés közben (**ld. 9.7. ábra**).



9.7. ábra: **Paired-end szekvenálás**. A 2-3 kb hosszúságú klónok végeiről néhány száz bp hosszúságú *read*-ek készülnek. Ha mindkét oldalról szekvenálják a klónokat, akkor az ellentétes oldali szekvenciák közötti távolság körülbelül megegyezik, feltéve, hogy nagyjából egyforma hosszúságú inszerteket szekvenálnak. Ez az információ felhasználható a szekvenciárészletek összeillesztésekor.

9.3. A Humán Genom Projekt (HGP)

A molekuláris biológia rendkívül gyors fejlődésének egyik legjobb példája a Humán Genom Projekt, amely nem kisebb célt tűzött ki maga elé, mint meghatározni a teljes emberi genomot. 1953 áprilisában jelent meg a DNS szerkezetét leíró publikáció James Watson és Francis Crick szerzőségével a *Nature*-ben, és a hetvenes évek közepén fedezték fel a DNS szekvenálás technikáját. Alig harminc évvel később pedig meghatározták a teljes emberi genomot. Érdekesség, hogy Watson 1990-ben a Humán Genom Projekt vezetője lett és 1992 áprilisáig dolgozott ebben a pozícióban. (Azért mondott le ilyen gyorsan, mert összekülönbözött a főnökével a gének szabadalmaztatásának ügyében. Watson álláspontja az volt, hogy a projekt összes eredményét szabadon elérhetővé kell tenni.) A gyors technikai fejlődés azt is lehetővé tette, hogy a projekt az eredetileg tervezettnél két évvel korábban befejeződhessen. A mikrobiális genomok átlagosan 3 millió bp-ból állnak, míg az emberi genom ennél három nagyságrenddel több, mintegy 3 milliárd bp-t tartalmaz. A Humán Genom Projekt tizenhárom éven keresztül zajlott, és 2003-ban zárták le. A projekt célját az alábbiakban határozták meg:

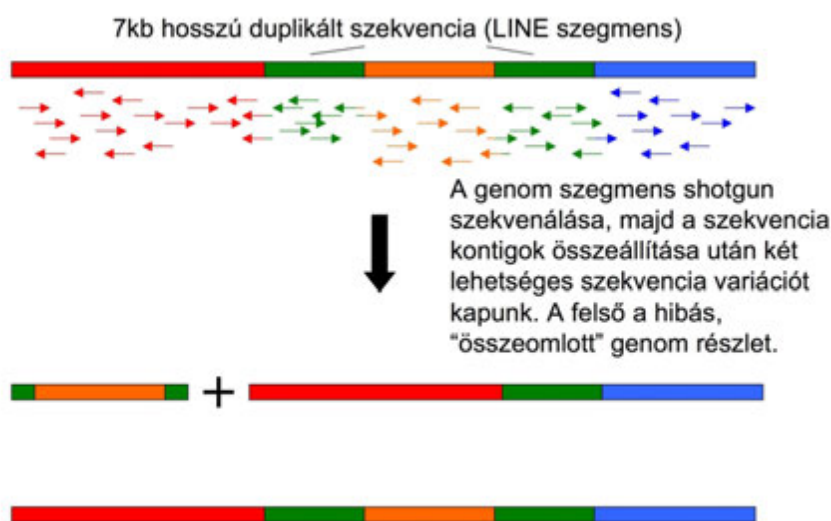
1. A humán DNS **összes génjének** (a korábbi becslés ~100.000 volt, a mai ismert szám <23.000) **azonosítása**.
2. A humán DNS nagyjából **3 milliárd bázispárjának meghatározása**.
3. A kinyert információmennyiség **adatbázisban történő tárolása**.
4. Az **adattfeldolgozási eljárások korszerűsítése**.
5. A projekt során létrehozott illetve korszerűsített **technológiák átadása a magánszektor**nak.
6. A felmerülő **etikai, jogi és társadalmi kérdések tisztázása**.

9.3.1. Nehézségek a humán genom vizsgálatában

Már a projekt legelejétől világos volt, hogy a biológiai tudományok eddigi legnagyobb volumenű és rengeteg technikai nehézséggel járó munkájába vágnak bele. Az első nehézség mindjárt a vizsgálni kívánt **genom mérete** volt, amely több nagyságrenddel meghaladta a korábban vizsgált genomok méretét. Másik komoly nehézség volt az **ismétlődő szekvenciák** jelenléte.

Egyrészt a DNS-polimeráz nem olvassa le a teljes szekvenciát, megakad vagy átugrik szakaszokat, ha túl sok A és T vagy G és C követi egymást akár több száz nukleotidon keresztül. Még komolyabb problémát jelentenek a hosszabb (a szokásos 500 bp-os leolvasott szekvencia szakaszoknál hosszabb) ismétlődő régiók, hiszen ezek egymástól megkülönböztethetetlenek. Duplikációk esetén az is előfordulhat, hogy az összeillesztés csak egyszer

veszi figyelembe az adott szekvenciárészletet, és így a végső szekvencia „összeomlik“, rövidebbnek adódik, mint a valóságban (**ld. 9.8. ábra**). Mindezen problémák miatt a kutatók fontosnak tartották, hogy jó minőségű genetikai és fizikai térképek készüljenek a humán genomról.



9.8. ábra: **A kontigok összeillesztésének problémái.** Ha a genom hosszú ismétlődő szekvenciákat tartalmaz, mint például egy 7 kb hosszúságú LINE szegmens (*long interspersed element*), az ismétlődő részek egymástól nehezen megkülönböztethetők a szekvencia kontigok összeillesztése során. A rossz egymáshoz rendelés eredménye a szekvencia összeállítás „összeomlása“ lehet, amikor az egymáshoz tartozó, de két LINE régiót tartalmazó szakasz helyett két különböző kontigot állít össze a szoftver, és a duplikáció eltűnik.

9.3.2. Elindul a HGP

A teljes humán genom vizsgálatának ötlete a 80-as években merült fel komolyabban, mikor lehetővé vált a DNS automata szekvenálása. 1987-ben az Egyesült Államok Energia Hivatala (*U.S. Department of Energy*) részvételével indult el egy korai genom projekt, amelynek célja a radioaktív sugárzás hatásának vizsgálata volt, különös tekintettel arra, hogyan lehet megvédeni a genomot a mutagén hatásoktól. További megbeszélések következtek, melyben már az NIH (*National Institute of Health*) is részt vett, és a két kutatóközpont közös erőfeszítésének eredményeként **1990 októberében hivatalosan is elindult a HGP**. Szinte azonnal nemzetközi együttműködés alakult ki, amiben Anglia, Németország, Franciaország, Kína és Japán is részt vett, kialakítva a **Humán Genom Konzorciumot**. Bár a projekt indulásakor még sok kutató komoly ellenérzéssel szemlélte a munkát, megkérdőjelezve az egész projekt hasznosságát, már a 90-es évek közepére jelentős mennyiségű adatot sikerült gyűjteni. A humán genom egyes részeinek **fizikai térképe** mellett **humán és egér genetikai térképek** születtek, **nagyszámú cDNS szekvencia**, és **teljes bakteriális genom szekvenciák** váltak elérhetővé. Kialakult a genomika tudománya. Érdekesség, hogy James Watson már a projekt legelején felhívta a figyelmet arra, hogy a megszülető eredményeknek jelentős bioetikai, szociális hatási lehetnek. Ezért a projekt összköltségvetésének először három, majd később öt százalékát az ezzel kapcsolatos kutatásokra fordították.

9.3.3. Megjelenik egy versenytárs

1998-ban Craig Venter (aki a *H. influenzae* genom projektet is vezette) bejelentette, hogy **Celera Genomics** néven létrehoz egy vállalatot, ami 2-3 éven belül végez a teljes humán genom szekvenálásával. Tervük szerint **shotgun módszer** alkalmaznak, és így el tudják kerülni az idő és pénzigényes térképezést. Nem csak gyorsabban, de olcsóbban is dolgoznak: a 3 milliárd dolláros államilag finanszírozott projekttel szemben 300 millió dolláros költségvetést tartott reálisnak. Összességében a kialakult verseny jót tett a teljes HGP-nek. Mindkét kutatói csapat nagy erőfeszítéseket tett, hogy minél hamarabb elkészüljön az első „draft“, vázlatos szekvencia. Ebben a legnagyobb nehézséget a fentebb már említett szekvencia összeillesztés okozta. Mindkét csoportnak sikerült olyan algoritmusokat kidolgozni, amik a korábbinál jóval nagyobb hatékonysággal tudták az egybefüggő szekvenciárészleteket

összeállítani. A Celera kutatói hamarosan publikálták a *Drosophila melanogaster* vázlatos szekvenciáját, megmutatva, hogy a *shotgun* módszernek a bakteriális genomoknál jóval nagyobb méretű genomok esetén is van létjogosultsága.

9.3.4. A humán genom vázlatos (*draft*) szekvenciája

A HGP egyik legfontosabb mérföldköve egy 2000 júniusában tartott sajtókonferencia volt, ahol a Humán Genom Konzorcium és a Celera együtt mutatta be a **humán genom első vázlatos szekvenciáját**. A humán genom projekt ekkor még messze volt attól, hogy késznek lehessen tekinteni. Az eukromatinnak, vagyis a kromatin génekben gazdag részének körülbelül 90%-a volt megtalálható a bejelentett szekvenciában, de sok hibával és nagyon sok *gap*-el. A két szekvenciát 2001 februárjában publikálták. A Humán Genom Konzorcium a Nature-ben, a Celera Genomics pedig a Science-ben, mivel utóbbi nem volt hajlandó a teljes szekvenciát teljesen elérhetővé tenni, amibe pedig a Nature szerkesztői nem egyeztek bele. Bár mindkét szekvencia erősen hiányos volt, mégis rengeteg új információt tartalmazott. Sok újdonság derült ki az ismétlődő szekvenciák és a gének genomon belüli elhelyezkedéséről, illetve fény derült arra is, hogy a korábban előrejelzett 100 000 helyett csak alig 23000 gén található meg a humán genomban, körülbelül annyi, mint az *ecetmuslica* genomjában.

9.3.5. 2003-ra elkészül a szekvencia

A *draft* szekvenciák elkészülte után a feladat a **hibák kijavításának** és a **lyukak betömésének** szisztematikus munkája volt. A hibajavításban nagy segítséget jelentettek a különböző megoldásokkal készült térképek. Egyedi DNS markereket kerestek, amiket a térképekből származó információkkal összepárosítva meg lehetett állapítani, hogy egyes kontigok mely más kontigokkal kapcsolódnak össze. A *H. influenzae* leírásánál említett módszereket is alkalmazták a lyukak betömésére. Voltak olyan *gap*-ek is, amelyek éppen egy *paired-end* módszerrel szekvenált 3kb-os klón közepére estek. Ebben az esetben **primer walking** módszert alkalmaztak, vagyis a klón két végén található, nagyjából 500bp-os már megismert szekvenciára terveztek egyedi oligonukleotidokat. Megszekvenálták még 500 bp-t, majd újabb oligonukleotidokat terveztek, és így „végigsétálva” az adott klónon annak szekvenciája ismertté vált. Az egyes javítások után újabb és újabb szakaszok váltak teljessé, amit megint újabb javítások követhettek a kapott új szekvencia információk alapján. Ezzel az **iteratív**, az előző eredményekre épülő **módszerrel** végül sikerült megalkotni egy majdnem teljes, **jó minőségű genom szekvenciát**.

A projekt 2003-as befejezésére az **eukromatin 99,3%-áról** állt rendelkezésre szekvencia információ, és a 2001-es, tehát a *draft* szekvencia bejelentésekor még meglevő 145 514 lyukat mindössze 281-re sikerült leszorítani. A kész szekvencia bejelentése 2003. április 25-én történt meg, éppen 50 évvel a híres Watson és Crick publikáció megjelenése után.

9.3.6. A szekvenálás technikájának fejlődése

A Humán Genom Projekt sikeréhez nagyban hozzájárult, hogy az évek során a DNS szekvenálás technikája is jelentős fejlődésen ment át. A **fluoreszcens technika** és az **automata szekvenátorok** megjelenése komoly előrelépésnek számított a korábbi manuális technikához képest, hiszen nem kellett radoaktívan jelölt nukleotidokat és az előhíváshoz röntgenfilmet használni. Mindez azonban nem lett volna elegendő egy ilyen nagyszabású projekthez, mert a teljes humán genom szekvenálásához körülbelül 1 millió gélelektroforézist kellett volna végezni, mindegyik gélen 48 szekvenálási reakcióval. Az áttörést a **kapilláris elektroforézis** alapuló technika hozta meg, amely jóval hatékonyabbnak bizonyult, mert így nem volt szükség akrilamid gélekre a szekvenáláshoz. A 2001-ben bemutatott második generációs szekvenáló készülékek tovább növelték a hatékonyságot, napjainkra pedig a „harmadik” vagy „következő” generációs készülékek és az új-generációs szekvenálási eljárások (**ld. 5.3. fejezet**) bizonyítják, hogy egészen olcsó és gyors is lehet a teljes genom szekvenálása.

9.3.7. Etikai kérdések avagy kinek a DNS-ét használták a szekvenáláshoz?

Ahogy korábban említettük, már a projekt legelején felmerült, hogy a HGP komoly etikai, jogi és szociális vonatkozásokkal is rendelkezik. Az egyik legelső fontos etikai kérdés az volt, hogy **kinek a DNS-ét** használják a szekvenáláshoz, illetve hogyan kezeljék a szekvenálás során keletkező bizalmas adatokat. A Humán Genom Konzorcium munkatársai különösen nagy hangsúlyt fektettek arra, hogy **teljesen anonim** legyen a szekvencia, és

még utólag se lehessen visszakövetni, hogy kinek a DNS-ét használták a munkához. Ehhez például a már eredetileg is anoniman kezelt DNS minták közül 10-ből egyet választottak ki, amelyekről eltávolítottak mindenfajta jelzést. Mára annyit lehet tudni, hogy a HGP során meghatározott genomszekvencia **70%-a egy donortól, míg a maradék 30% több különböző donortól** származik. A Celera Genomics is anonim módon tervezte kezelni a mintákat. Huszonegy, különböző etnikai csoporthoz tartozó donortól származó mintából véletlenszerűen választották ki a ténylegesen felhasznált DNS-t. A Celera akkori vezetője, Craig Venter a vállalatától történt távozása után bevallotta, hogy variált a mintákkal, és az ő DNS-e volt az egyik, amit a munkához használtak. Bejelentette, hogy a szekvencia alapján szívbetegségre és Alzheimer-kórra van hajlama. A tudományos közvélemény rendkívül etikátlannak tartotta az eljárást, mert ugyan Venter-nek joga van ahhoz, hogy a világ elé tárja a saját DNS-szekvenciáját, ezzel a rokonairól is bizalmas információkat tesz közzé, amihez azok nem járultak hozzá. Bár Venter szerint az ő példája megmutatja, milyen fantasztikus eszköz a saját genetikai információknk ismerete (aminek hatására ő egyébként azóta több testmozgást végez és diétázik), egyben arra is rávilágít, milyen **komoly etikai, és az egyének életére közvetlenül kiható problémákat vet fel a genom szekvenálás**. Különös tekintettel arra, hogy a szekvenálás technológiájának fejlődése hamarosan elhozza az „**1000 dolláros genom**” időszakát, amikor gyorsan és olcsón lehet meghatározni az egyes emberek teljes genetikai állományát. (Valójában azt is mondhatjuk, hogy ez a korszak el is érkezett. Van már olyan amerikai cég, amely pontosan 1000 dollár ellenértékéért bárki teljes „exom”-ját, azaz a teljes fehérjekódoló DNS szekvenciát elkészíti; a teljes eukromatint tartalmazó genom pedig 1500 dollárba kerül).

9.4. További genom projektek

A szekvenálási technikák fejlődésével, ami a genom meghatározás idejét és költségét egyaránt jelentősen csökkentette, további genom projektek megvalósítására nyílt lehetőség. Ezek közül érdemes megemlíteni a **Csimpánz genom projektet**, aminek célja, hogy össze lehessen hasonlítani a főemlősök és az ember genomját, és ezáltal többet tudjunk meg arról, mi különbözteti meg az embert más fajoktól. Különösen érdekes vállalkozás volt a **Neandervölgyi ember genom projekt**, aminek célja fossziliákból meghatározni a neandervölgyi ember genomját. A *draft* szekvenciát 2010-ben közzétették a Science-ben, és ez alapján a genom 99,7%-a megegyezik a *Homo sapiens*-ével (míg a csimpánz esetén ez 98,8% volt). A neandervölgyi projektben a legnagyobb nehézséget a minták kezelése jelentette, hiszen a fossziliák egyrészt nagyszámú mikrobával szennyezettek, másrészt a fossziliákkal kapcsolatba kerülő kutatók DNS-e is beszennyezhetette őket. A kutatók szerint azonban ezek a problémák jól kezelhetők, és pl. az emberi szennyező DNS mértéke nem több, mint 1%. A genom projektek jövője az egyes fajokon belül minél több egyed genetikai állományának meghatározása, amellyel közelebb juthatunk az egyedek közötti genetikai diverzitás jobb megértéséhez. (Az „egyéni” humán genomok között az elsők között szekvenálták James Watson-ét).

9.5. További olvasnivaló a fejezethez

Fleischmann RD, Adams MD, White O, Kirkness EF, et al. (1995) Whole-genom random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269: 496-512.

Bentley DR (1996) Genomic sequence information should be released immediately and freely in the public domain. Science 274: 533-534.

International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409 (6822): 860-921.

Venter, JC et al. (2001) The sequence of the human genome. Science 291 (5507): 1304-1351

Wolfsberg TG, Wetterstrand KA, Guyer MS, Collins FS and Baxevanis AD (2003) A user's guide to the human genome. Nat. Genet. 35: 4.

Green RE, Krause J, Briggs AW et al (2010) A draft sequence of the Neandertal genome. Science 328 (5979): 710-22.

10. fejezet - Bevezetés a bioinformatikába

A biológia, különösen a molekuláris biológia 20. század végi és 21. századi fejlődése elképzelhetetlen lett volna megfelelő informatikai háttér nélkül. Egyedül a számítógépek képesek ugyanis áttekinthetően tárolni azt a hatalmas mennyiségű adatot, ami a biológiai polimerek szerkezetének és működésének megértéséhez szükséges. Ugyanezek a számítógépek teszik lehetővé a genomok automatikus összeállítását, gének azonosítását, a nukleinsavak és fehérjék tulajdonságainak jóslását - évről-évre egyre nagyobb pontossággal. A számítógépeket összekapcsoló világháló segítségével a publikációk, kísérletes eredmények, a genomok, gének, nukleinsavak és fehérjék adatai ma már bárhol elérhetőek. Ezen felül számos, szabadon hozzáférhető program áll rendelkezésre az interneten, amelyekkel nagyon könnyű molekuláris biológiai elemzéseket végezni. A közvetett információkon alapuló jóslat természetesen nem helyettesíti a pontos mérési eredményeket. Ennek ellenére, a molekuláris szintű kísérletek tervezéséhez és értékeléséhez a megfelelő teljesítményű és tároló kapacitású számítógépek ma már nélkülözhetetlenek.

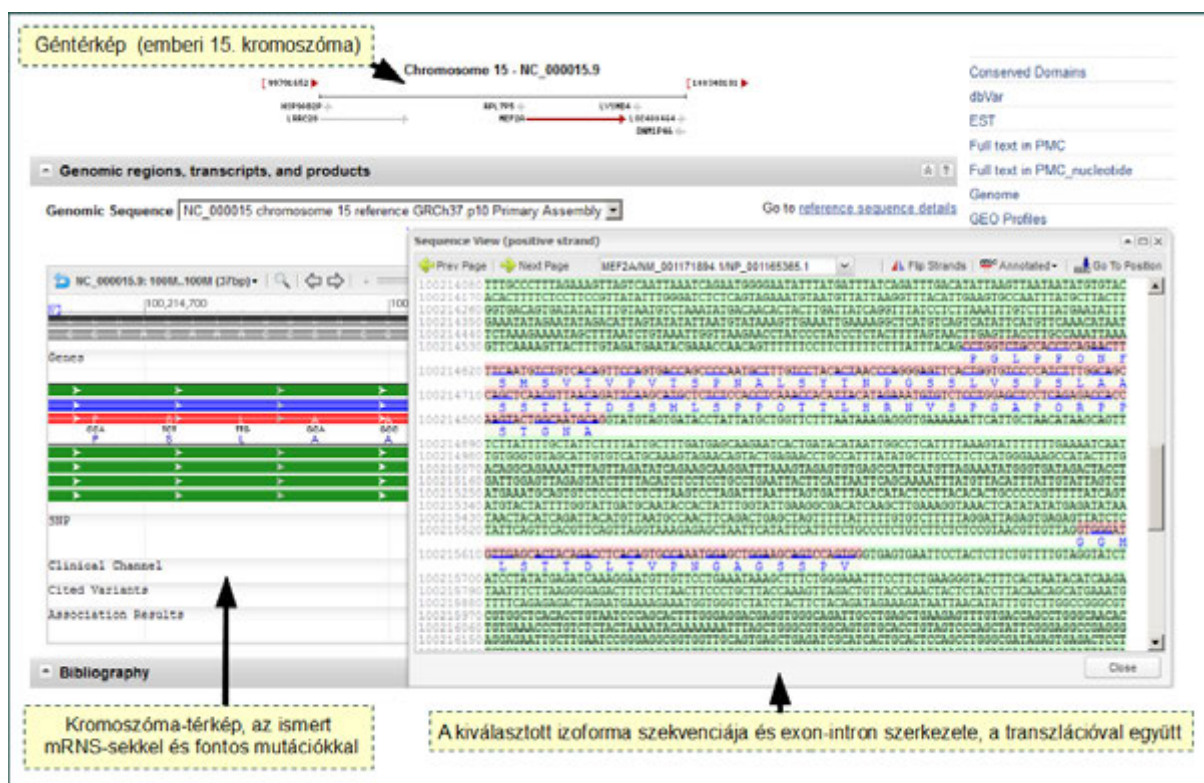
10.1. Bioinformatikai adatbázisok

10.1.1. Bibliográfiai gyűjtemények

Az olvasóban első kérdésként bizonyára felmerül, hogy mi lehet az internetes oldalak haszna a biológiai kutatás számára. A válasz igen sokrétű. A legnagyobb előny a kísérletes adatokhoz való szabad hozzáférés - akár teljes cikkek alakjában, akár elemi adathalmazok formájában. Mielőtt hozzáférnénk bármilyen kísérletsorozathoz, azt általában jelentős tervezői munka kell, hogy megelőzze. Ebben az első és legfontosabb forrásunk az irodalomban már publikált eredmények lesznek. A világ legnagyobb ilyen irodalmi adatbázisa az amerikai NCBI (*National Center for Biotechnology Information*) **PubMed** nevű gyűjteménye. A keresőprogramja nem túl látványos, de nagyon intelligens: kereshetünk kulcsszavakra, szerzők, vagy folyóiratok szerint is. Az egymásra hivatkozó, rokon cikkeket össze is fűzi a keresőprogram (a "*related*" oszlopban tekinthetők meg). Mindemellett egyéb keresők, köztük a Google Scholar keresőfelület hasznosságát sem szabad lebecsülni, ha irodalmat kutatunk. A már leírt eredmények ismerete már csak azért is fontos, hogy ne ismételjük meg feleslegesen a már jól ismert tényeket bizonyító kísérleteket, illetve hogy ne állítsunk fel elvben is megvalósíthatatlan kísérleti terveket.

10.1.2. DNS szekvencia-adatbázisok: GenBank és Ensembl

Egy molekuláris biológiai laboratórium mindennapi munkája számára az egyik legfontosabb forrásként a gének és fehérjék pontos "képletei", azaz szekvenciái szolgálnak. Az ezeket tároló és interneten hozzáférhető gyűjteményeket nevezik **elsődleges szekvencia-adatbázisoknak**. Az ismert fajokból nyert elsődleges genetikai információ három nagy nemzetközi oldalon férhető hozzá: Ilyen az NIH intézménye, az NCBI által fenntartott **GenBank** adatbázis (USA), az EMBL (*European Molecular Biology Laboratory*) gyűjteménye, az **ENA** (*European Nucleotide Archive*) illetve a DDBJ (*DNA Data Bank of Japan*) adatbankja. A gyakorlatban leginkább a GenBank adatbázist szoktuk használni. Itt a **genomikus adatok**on túl **cDNS információkat** is tárolnak. A cDNS szekvenciákat könnyen lefordíthatjuk aminosav szekvenciákra, de a klónozási feladatokhoz is elengedhetetlen az ismeretük. A primerek tervezéséhez is ismerni kell a cDNS pontos szekvenciáját, akár a nem kódoló szakaszokat is beleértve. A GenBank adatbázisa lehetővé teszi a gének látványos megjelenítését is: ez a **Map GeneViewer**, amely a Google térképek szoftveres felületén alapszik (**ld. 10.1. ábra**). A kijelölt génszakaszt felnagyítva (*zoom*) láthatóvá válik az egyedi nukleotid sorrend. A program külön jelöli az mRNS-re átíródó szakaszokat (az összes ismert variáns), és nemcsak az exonokat, de azok pontos aminosav-fordítását is megadja. Sőt, még a fiziológiás variánsokat (polimorfizmusokat) és a patológiás mutációkat is jelöli. Ez azonban nem az egyetlen genom-vizualizációs program.



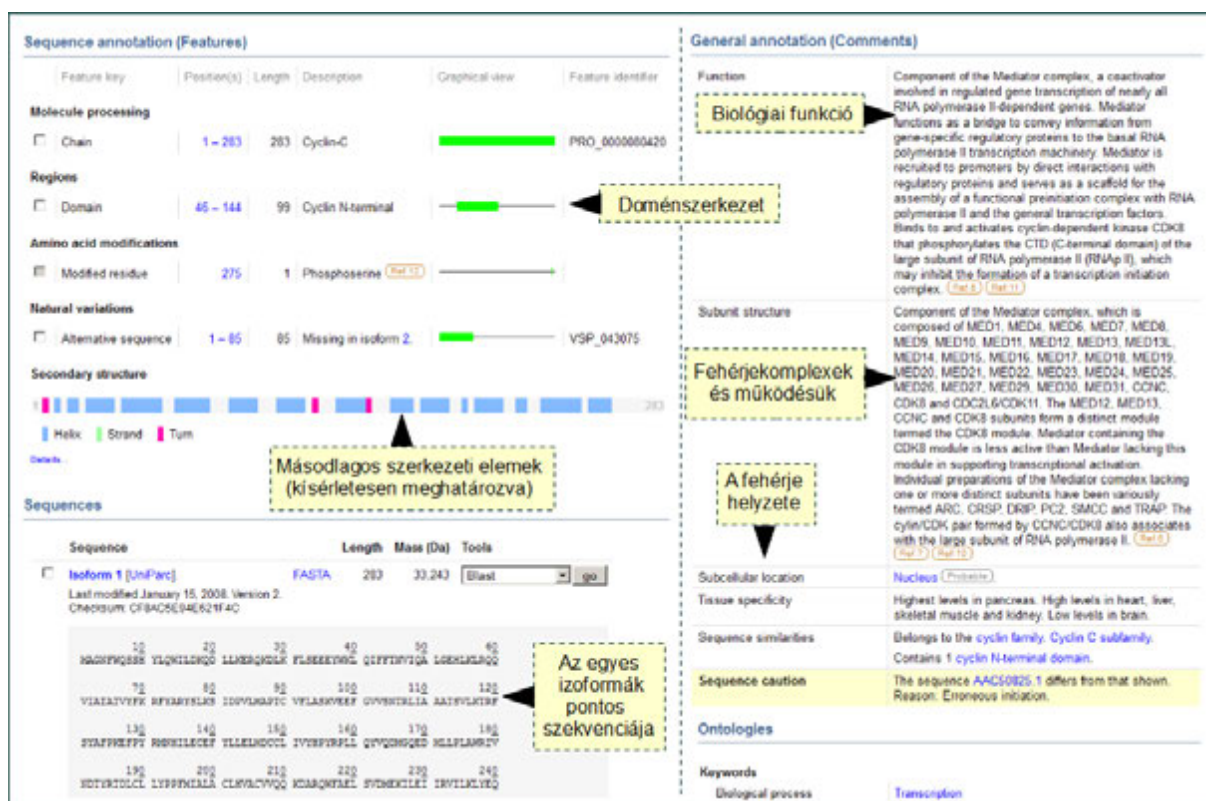
10.1. ábra: Az NCBI grafikus „nézegető” felülete, a GeneViewer. A humán MEF2A gén részlete

Az európai **Ensembl** genom-adatbázis (amely főleg gerincesek és néhány modell gerinctelen élőlény teljes genom szekvenciáit tartalmazza) internetes felülete szintén könnyen kezelhető. Az utóbbi előnye, hogy tartalmazza például a génekben az eddig leírt potenciális mRNS szekvenciákat (azokat is, amelyekről nem történik fehérjeszintézis), az ismert epigenetikai módosításokat, továbbá rengeteg bioinformatikai jóslással (predikció) kapott információt is.

10.1.3. Fehérje-szekvencia adatbázisok: UniProt

A nukleotidokkal ellentétben, a jelenleg létező legjobb fehérje adatbázis nem az USA-ban található, hanem Európában. Ez nem más, mint a svájci **UniProt** (régebben **SwissProt**) gyűjteménye. Ezen adatbank nemcsak a sokféle előrejelző programnak köszönheti kiválóságát, hanem az adatok manuális feldolgozásának is. Az adatbázisban módunk van a fehérjék neve (akár szinonimák) vagy azonosítója alapján kutatni. A UniProt adott fehérjei szekvencia hasonlóság alapján, a BLAST programmal is megkereshetők (ld. 10.2.3. fejezet). A bennünket érdeklő találatokat a kijelölésük után kötegelve ("batch" formájában) is letölthetjük, további analízis céljából. A szekvenciák **FASTA formátumban** fognak megjelenni: a kötelező ">" jel után a fejléc-sor tartalmazza a szekvencia azonosítóját, a fehérje nevét, és az organizmus nevének betűkódját. A második sortól következik csak a tényleges aminosav sorrend, egybetűs kódban.

A UniProt azonban nem csak a fehérjék szekvenciáját adja meg: az egy génhez rendelhető különféle (alternatív *splicing* és alternatív iniciáció következtében keletkező) izoformák is közös oldalon találhatóak (a különbségekre az oldal külön felhívja a figyelmünket). Ezenkívül információt kapunk a kísérletesen azonosított kötőpartnerekről, az egyes szerkezeti elemekről, és másodlagos (poszttranszlációs) módosításokról is. A weboldalak rengeteg hasznos, direkt hivatkozást is tartalmaznak, más adatbázisokra, például a gén, mRNS és a térszerkezet kapcsán (ld. 10.2. ábra). Jóllehet a GenBank nukleotid szekvenciáihoz is megadják az aminosav szekvencia fordítást (*translation*), általában jobb, ha a szekvenciának egy dedikált fehérje adatbázisban is utánanézzünk. A UniProt **manuálisan ellenőrzött** fehérje-szekvenciái már jórészt mentesek a predikciós vagy szekvenálási hibák során adódó kellemetlen következményektől (pl. nem létező izoforma, esetleg leolvasási keret eltolódás, angolul *frameshift* vagy korai stop kodon). Ha egy fehérjével problémák vannak, például csak egy gén alapján jósolt (*predicted*) szekvencia, vagy még nem is ismert a teljes hossza (*incomplete*), erre a UniProt külön figyelmeztetést ad. Így nagyobb biztonsággal tervezhetjük az adott szekvenciával a további géntechnológiai munkát (pl. klónozás expressziós vektorba).



10.2. ábra: Az UniProt adatbázis egy oldala. A human ciklin C (CCNC) fehérje aminosav szekvenciája és az annotáció egyes elemei.

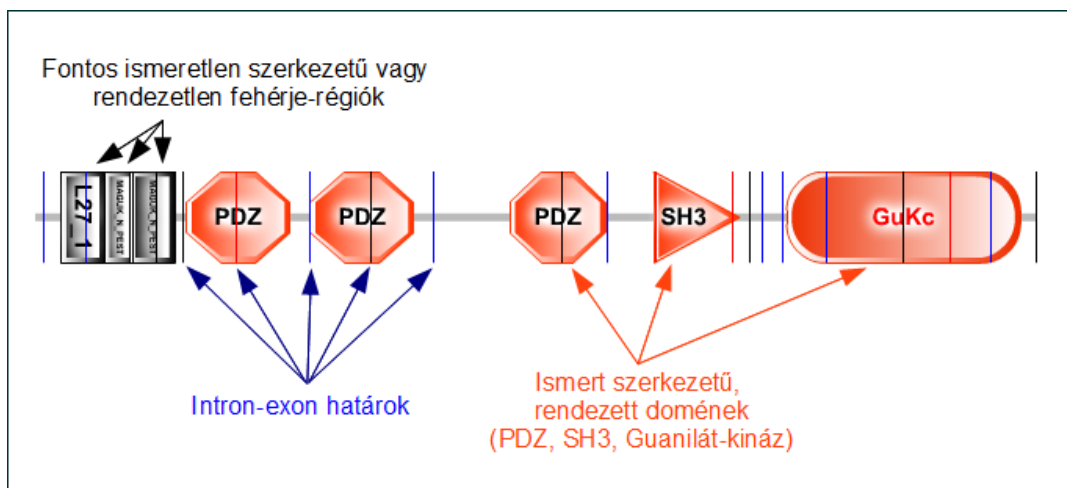
10.1.4. Az "omikák" világa. Fajonkénti adatbankok.

Az eddig említett néhány adatbázis csupán példa a molekuláris biológiai kutatási területek egyedi adatgyűjteményeire. Az ilyen területek neveit divatos szóval "-omika" végződéssel szokás ellátni (s a mára a biológiai kutatások egyik fő irányává vált rendszerbiológia részét képezi), magát az adathalmazt pedig (etimológiailag nem teljesen helyes módon) "-om" végződéssel. A **genomika** a genomok vizsgálatának tudománya, a **transzkriptomika** az átírt nukleotidoké (mRNS), a **proteomika** a fehérjéké. Ezenkívül beszélhetünk többek között **epigenomikáról** (DNS-módosítások), **metabolomikáról** (anyagcsere köztestermékek), **lipidomikáról**. Ezen rész tudományok számára külön adatbázisok is rendelkezésre állnak. Például a **kinom** (ami az összes protein-kináz enzimet jelöli) részletes leírására ott vannak a **KinBase** vagy a **PhosphoSitePlus** adatbázisok. Hasonlóan, a **glikom** az összes biológiailag érdekes szénhidrátot jelöli (amivel a glikomika tudományterülete foglalkozik), külön adatbázisokkal (ilyen például a cukorlánc szerkezete alapján kereshető **GlycomeDB**). A lipideket kutató szakemberek szintén saját adatbankokat használhatnak (pl. **Lipidmaps**), amire elsősorban a poliketidek nagy variabilitása miatt van szükség. Az anyagcsere-utak nyomon követésére pedig ott a jól ismert **KEGG** (*Kyoto Encyclopedia of Genes and Genomes*), amely enzimenként és lépésenként tekinti át az egész intermediér anyagcserét, interaktív térképekkel. Szak-adatbázisokat azonban nem csak az egész élővilágot áttekintő, integratív jelleggel hoztak létre. Léteznek egy-egy modell-organizmusra fókuszáló genetikai, proteomikai, és egyéb kísérletes adatokat összesítő adatbankok is. Ilyen például a *Caenorhabditis elegans*-ra fókuszáló **WormBase**, a *Drosophila melanogaster* (ecetmuslica) molekuláris biológiáját taglaló **FlyBase**, vagy a *Saccharomyces cerevisiae* (sörélesztő) genetikájával foglalkozó **SGD** adatbázis. Ezek az oldalak összehasonlításokkal is szolgálnak: az **Interactive Fly** például a rovaroktól emberig konzervált fehérje-rendszerek működésének az egyik létező legjobb leírását adja.

10.1.5. Szerkezeti és funkcionális adatbázisok

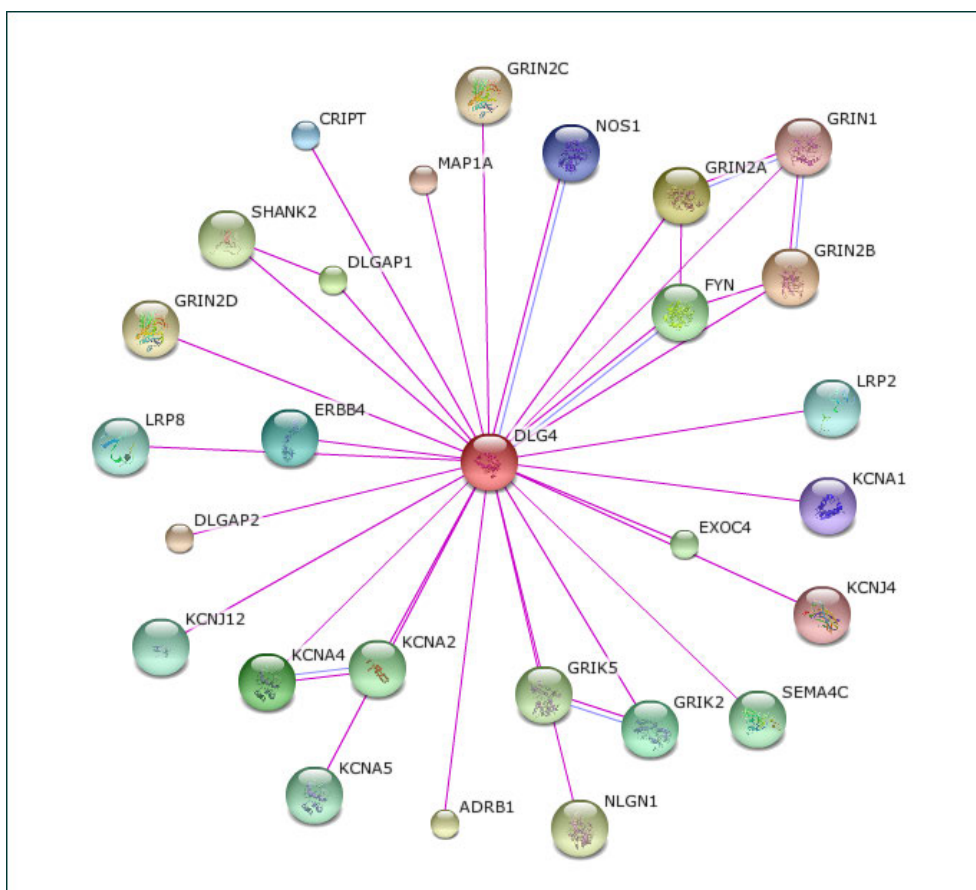
Az előző alfejezetben láttuk, hogy adatbázisokat nem csak az elsődleges DNS vagy fehérjeszekvenciákból lehet létrehozni. Rengeteg génről (illetve annak fehérjetermékéről) állnak rendelkezésre a szekvencián messze túlmutató szerkezeti és működési adatok. Sok fehérje harmadlagos szerkezetét is már meghatározták. Ezeket (de a nukleinsavakét is) a **Protein Data Bank (PDB)** adatbázis összesíti. A PDB valójában három nagy szerkezeti

genomikai kezdeményezés (USA, EU, Japán) kooperációjával jött létre; célja, hogy a makromolekulák térszerkezetei minden kutató számára szabadon elérhetőek legyenek. Ennél persze sokkal feldolgozottabb (másodlagos) adatbázisok is léteznek. A fehérjék szerkezettel bíró doménjeit a legtöbb esetben elegendő egy egyszerű doboz-moddal ábrázolni. Ilyeneket találunk a **SMART** (*Simple Modular Architecture Research Tool*) másodlagos adatbázisban, amely - nevének megfelelően - nagyon intelligens keresővel rendelkezik: a megkeresett fehérjében (amit akár szekvenciájával is megadhatunk) automatikusan bejelöli az ismert doméneket, az intron-exon határokat, sőt még listázni is képes a nagy vonalakban hasonló domén felépítésű, rokon fehérjéket (**ld. 10.3. ábra**).



10.3. ábra: A SMART adatbázis részlete. A human PSD-95 idegi szinapszis-szervező fehérje domén felépítése.

Ha csak a fehérje-fehérje kölcsönhatásokra vagyunk kíváncsiak, erre is számos, kiváló weboldal áll rendelkezésünkre. Ilyenek az **IntAct** vagy a **BioGRID**; de az egyik leglátványosabb (mind adatok, mind grafika tekintetében) a **STRING** adatbázis (**ld. 10.4. ábra**). Ez a kölcsönhatásokat, a SMART elemzéseit, sőt, a háromdimenziós szerkezeteket is egyetlen hatalmas gráf-modellbe integrálja. Minden esetben mód van szűrők beállítására: ha például csak a kísérletekkel megerősített kölcsönhatások után érdekelődünk, a nagy áteresztőképességű (*high-throughput*) mérések és az adatbányászattal (*data mining*) nyert automatikus következtetésekre vagy jóslatokra nem, erre is megvan a mód. Az interakciós gráfot pedig tetszés szerint végigjárhatjuk, így az azonos funkcióhoz tartozó fehérjék biokémiai hálózata is megfigyelhető lesz. A fehérjék működésének legrészletesebb osztályozását egy másik adatbázis, a **GO** (*Gene Ontology*) ill. annak keresője (**AMIGO**) tartalmazza. A GO minden egyes fehérjéhez egy hierarchikusan felépített működési leírást ad. Például az emlős izomzat fejlődését szabályozó MEF2A transzkripciós faktorra egy alapszintű leírás a "*transcription, DNA dependent*" lehet, egy ennél jóval specifikusabb pedig a "*heart development*". Ha a sejten belüli szabályozási mechanizmusok finom részletei érdekelnek bennünket, felkereshetjük a **Reactome** vagy a **Signalink** oldalát is. Itt minden egyes biokémiai útvonalra külön gráf-modellt találunk. A színes folyamatábrák nagyban segíthetik a komplex biokémiai folyamatok megértését, de a kísérletek tervezését is.

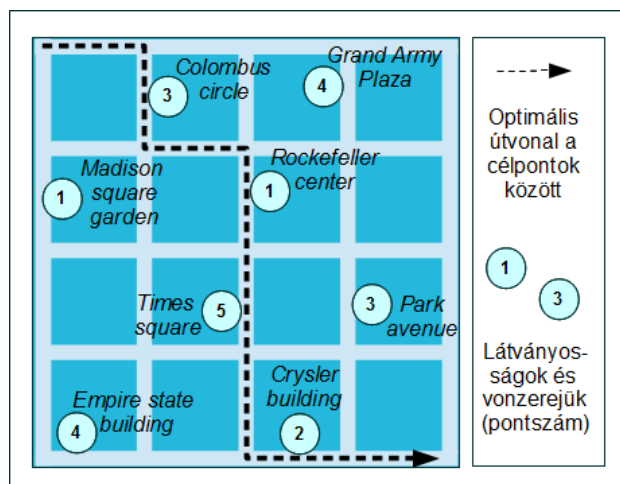


10.4. ábra: A STRING adatbázis részlete. A human PSD-95 fehérje fontosabb ismert kötőpartnerei.

10.2. Szekvencia-illesztések és egyéb fontos algoritmusok

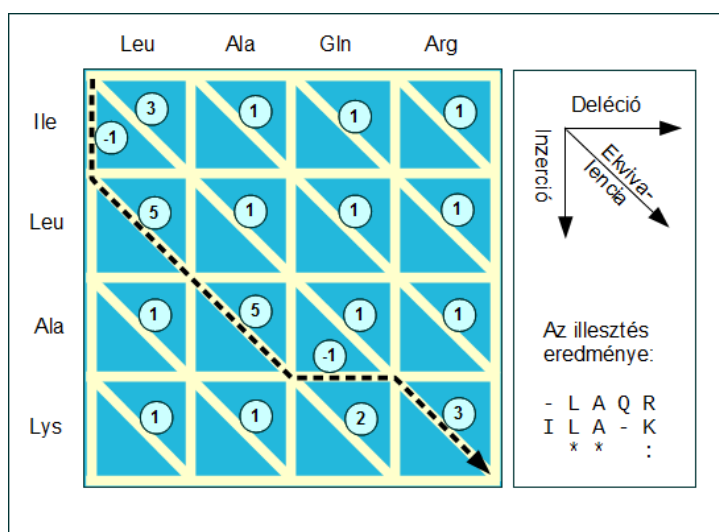
10.2.1. Páronkénti szekvencia-illesztések (*pairwise alignments*)

A páronkénti szekvencia-illesztés a molekuláris biológia egyik leggyakrabban használt algoritmus. A feladat az, hogy két, hasonló kódsort a lehető legpontosabban megfeleltessünk egymásnak - ha szükséges, akár részek (deléciók) és beillesztések (inszerciók) segítségével. A kódsor egyedi betűi lehetnek nukleotidok (nukleinsavak esetén) vagy aminosavak (fehérjék esetén): az algoritmus mindkét esetben nagyon hasonló lesz. A két kódsor megfeleltetésének problémája visszavezethető egy híres feladatra, a **“manhattani turista problémája”**.



10.5. ábra: A „manhattani turista” problémája

A kérdés: hogyan tudunk eljutni egy teljesen négyzethálós térképen az utcák (élek) mentén az utcasarkokon (csúcsokon) keresztül a bal felső sarokból a jobb alsóba, úgy hogy közben a lehető legtöbb nevezetességet érintjük? (ld. 10.5. ábra). A problémát nagyon könnyű számítógéppel megoldani, amely pontozással megtalálja az optimális utat. Ugyanígy lehet DNS vagy fehérjeszekvenciákat pontos párokba állítani, itt a “bejárt út” az elfogadott vagy elvetett párokat mutatja meg (ld. 10.6. ábra). Erre a feladatra számos program áll rendelkezésre: pl. a **CLUSTAL** programcsomag (ld. 10.2.5. fejezet).



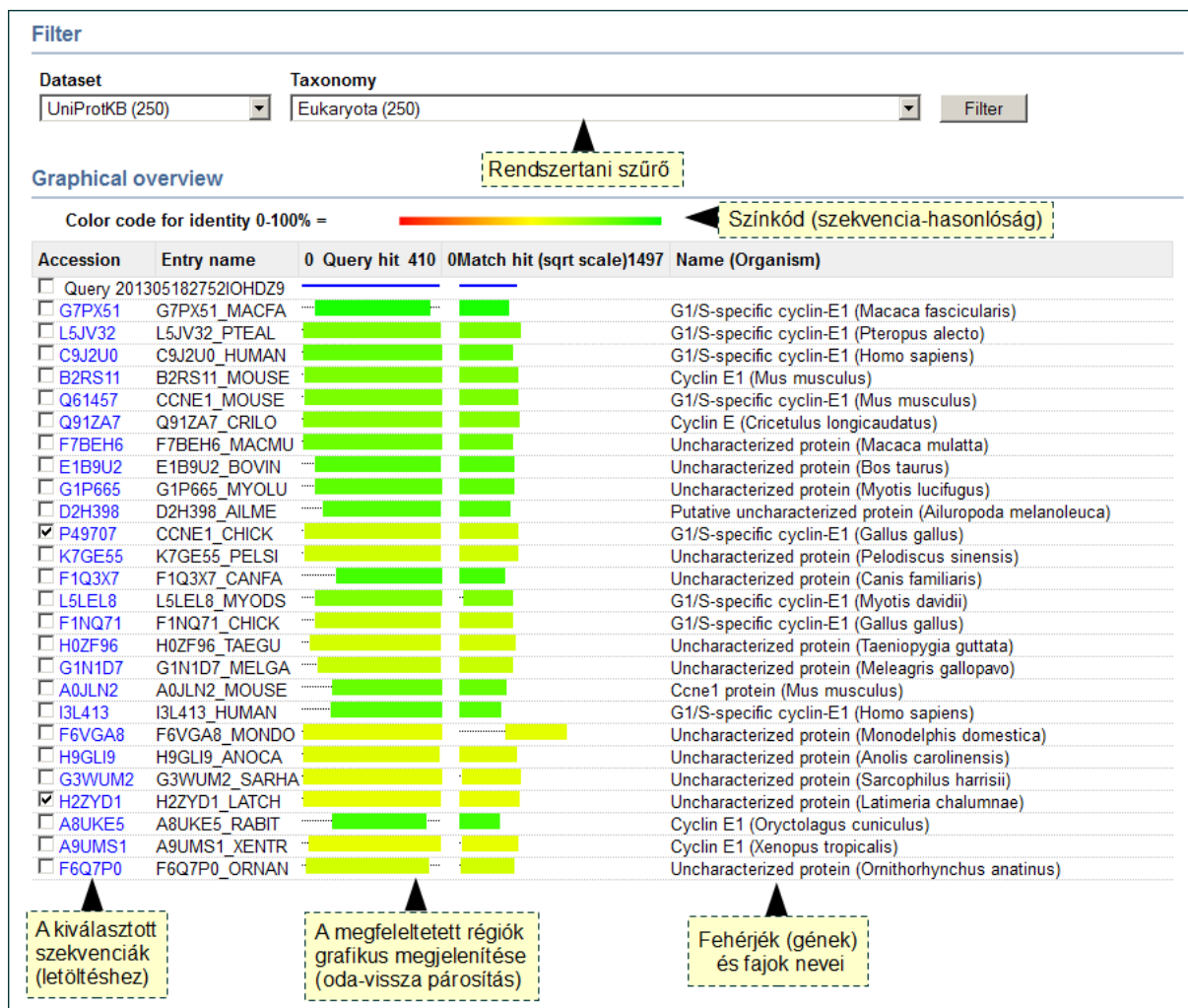
10.6. ábra: A páronkénti szekvencia-illesztések problémája

10.2.2. A szekvencia-illesztés paraméterei

Az illesztés jósága a pontozási rendszertől függ. Egyrészt súlyozni kell az egyes betűk hasonlóságát: egy leucin aminosav például jobban hasonlít egy izoleucinra, mint egy argininre. Ezt különféle hasonlósági mátrixok segítségével lehet megtenni (a részletekre itt most nem térünk ki). A gyakorlatban a fehérjék számára a **BLOSUM**, **GONNET** vagy **PAM mátrixok**at használjuk, a nukleinsavaknak egyszerűbb mátrixok is megfelelnek (ID). A hosszúságbeli eltéréseket, inszerciókat/deléciókat szintén pontozni kell, méghozzá negatív (büntető) pontokkal. A jobb illesztőprogramok általában megengedik ezen paraméterek változtatását, erre pedig a következők miatt lehet szükség: alacsonyabb fokú hasonlóság vagy nagyon eltérő hosszúságú párok esetén a módszer sokszor nem találja meg az összetartozó szakaszokat. Ilyen problémák gyakoriak például rendezetlen fehérjeszekvenciák evolúciós vizsgálata esetén (ahol nagyon alacsony a konzerváltság). Tehát ha felmerül a gyanú, hogy a módszer tévedett, ne sajnáljunk "játszani" a paraméterekkel, egy jobb minőségű illesztés érdekében!

10.2.3. A BLAST algoritmus

A páronkénti illesztést nem csak egyenként lehet végrehajtani. A legtöbb adatbázis lehetőséget kínál arra, hogy az összes ott tárolt szekvencia ellenében automatikusan lefuttassuk a páronkénti illesztést, és a találatokat jóság alapján rangsorolva visszakapjuk: ez a **BLAST** algoritmus (*Basic Local Alignment Search Tool*). A BLAST-nak számos változata létezik: fehérjékre és nukleinsavakra egyaránt használható. A program a megadott szekvenciában előbb (**heurisztikus** módon) kellően egyedi, kicsi darabokat azonosít, ezzel keres, majd az elsődleges találatokon végzi el a teljes illesztéseket. Az NCBI és a UniProt oldalai automatikusan felkínálják a BLAST programot minden egyes oldalon, az éppen aktuálisan kijelölt szekvenciára. A visszaadott találatokon általában jelölik a megfeleltetett szakaszt, az azonosság/hasonlóság mértékével együtt (szinkóddal is). A véletlen egyezés (vagyis a tévedés) valószínűségének értékét is megadják (ez a p-érték), a rangsor ennek alapján készül (**ld. 10.7. ábra**).



10.7. ábra: Protein BLAST keresés a UniProt szerveren (humán ciklin E1 fehérje)

10.2.4. Szekvenálás ellenőrzése illesztésekkel. Plazmid szerkesztők

A gyakorlatban a páronkénti illesztéseket leggyakrabban a molekuláris klónok ellenőrzésére használjuk. A szekvenálással meghatározott bázissorrendet ugyanis össze kell hasonlítani az eredetivel. Kényelmes megoldás lehet egy-egy BLAST futtatás akár DNS szinten, akár a lefordítás után fehérje szinten. Ennek azonban vannak bizonyos hátulütői: a BLAST elsősorban kisebb darabonkénti egyezéseket keres, így a mutációk, inszerciók és deléciók esetleg elkerülhetik a figyelmünket. Továbbá a BLAST nem tud mit kezdeni a szintetikus, nem természetes génekkel (pl. emberi vagy egér eredetű gendarab egy bakteriális plazmidban), így az eredmény néha meglehetősen

10.2.5. Többszörös illesztések (*multiple alignments*). A Clustal programcsomag

Az apoptózt szabályozó PEA15 fehérjék többszörös illesztése (a *ClustalW2* szerveren)

```

sp |Q15121|PEA15_HUMAN      MAEYGLT LQDL TNNIT LED LEQ LKSACE DIPSEKSEE ITT GSAWFSFLE 50
sp |Q62048|PEA15_MOUSE     MAEYGLT LQDL TNNIT LED LEQ LKSACE DIPSEKSEE ITT GSAWFSFLE 50
tr |Q3KPT6|Q3KPT6_XENLA    MEEDSL QLEL TEN ITT NED LD L LKSACE DIPSEKSEE QIAS QCEWFSFLE 50
tr |H3A692|H3A692_LATCH    MAEYNAL LDL MDN ITT NED LEQ LKSACE DIPSDKSEE ITT CSKEWNFLE 50
tr |E7FEH9|E7FEH9_DANRE    MSEYSSL SDL SEN ITT NED LEQ LKSACE DIP EEQSN SI AC SRDWF SYLE 50
tr |H2SKB5|H2SKB5_TAKRU    MAEYTSLS DL SEN ITT NED LEQ LKSACE DIPEDQSN ITT TSKEWFSYLE 50
                                * * * * * : : * * * * * : : * * * * * : : * * * * * : : * * * * *

```

```

sp |Q15121|PEA15_HUMAN      SHNKLKDKNLSYIEHIFEI SRRPDL T MWIVYRT RV L K I SEED ELD T K L T 100
sp |Q62048|PEA15_MOUSE     SHNKLKDKNLSYIEHIFEI SRRPDL T MWIVYRT RV L K I SEED ELD T K L T 100
tr |Q3KPT6|Q3KPT6_XENLA    KHEKLKDKNLSYIEHIFEI SRRPDL T MWIVYRT RV L K I SEED ELD T K L T 100
tr |H3A692|H3A692_LATCH    KHEKLKDKNLSYIEHIFEI SRRPDL T MWIVYRT RV L K I SEED ELD T K L T 100
tr |E7FEH9|E7FEH9_DANRE    KNDKLAQNLSYIEHIFEI SRRPDL T MWIVYRT RV L K I SEED ELD T K L T 100
tr |H2SKB5|H2SKB5_TAKRU    KNDKLAQNLSYIEHIFEI SRRPDL T MWIVYRT RV L K I SEED ELD T K L T 100
                                : : * * * * * : : * * * * * : : * * * * * : : * * * * *

```

```

sp |Q15121|PEA15_HUMAN      RIPS AKYKDI IRQPS EEE I I K L A P P K K A 130
sp |Q62048|PEA15_MOUSE     RIPS AKYKDI IRQPS EEE I I K L A P P K K A 130
tr |Q3KPT6|Q3KPT6_XENLA    RIPS AKYKDI IRQPS EEE I I K L A P P K K A 130
tr |H3A692|H3A692_LATCH    RIPS AKYKDI IRQPS EEE I I K L A P P K K A 130
tr |E7FEH9|E7FEH9_DANRE    RIPS AKYKDI IRQPS EEE I I K L A P P K K A 130
tr |H2SKB5|H2SKB5_TAKRU    RIPS AKYKDI IRQPS EEE I I K L A P P K K A 130
                                * * * * * : : * * * * * : : * * * * * : : * * * * *

```

Az eredményül kapott hasonlósági fa (kladogram)

A páronkénti illesztések sorrendje

1. 3. 4. 5.

10.2.6. Evolúciósan konzervált elemek azonosítása

119

fonálféreg és ecetmuslica), ahol a véletlen mutációk már szinte minden olyan pozíciót kicseréltek, amely nem volt abszolút nélkülözhetetlen, akkor a megfordítás is igazzá válik. Ilyen esetekben a többszörös szekvencia-illesztések rendkívül értékes adatokkal szolgálnak. Ezekhez az elemzésekhez nincs szükség semmiféle szerkezeti ismeretre; így egy hasznos kiegészítő technika lehet a kísérletes eredmények értelmezésére vagy akár ellenőrzésére. Például a DNS-ben talált **hosszú konzervált szakaszok** legtöbbször az **exonoknak** felelnek meg, míg a köztes, alig-alig konzervált szakaszok lesznek az intronok. Hasonlóan, a fehérjékben látható, **hosszan konzervált blokkok** általában a szerkezettel rendelkező **doméneket** mutatják meg, míg a megőrzött aminosavakat nem mutató szakaszok legtöbbször rendezetlenek. Az egyébként teljesen változatos fehérjeszekvencia közepén itt-ott található, **rövid konzervált szakaszok** a molekulák összekapcsolódását segítő felismerő elemeknek (**lineáris motívumoknak**) felelhetnek meg. Ugyanígy, a gének promóterében található konzervált "szigetek" igen gyakran a transzkripció faktor kötőhelyeket jelzik. A módszer használatához mindössze arra van szükség, hogy megtaláljuk azt az optimális evolúciós távolságot, ahol a kérdéses gén vagy fehérje működése még közel azonos, viszont a szekvencia már eléggé különböző. Feltéve, hogy a szekvenciák megbízhatóak, az adatbázisból BLAST segítségével nyert találatok óvatos illesztése ilyenkor nagyon sok információt nyújthat. (ld. 10.9. ábra)



10.9. ábra: Evolúciós analízis többszörös szekvencia-illesztés alapján (szerkezeti és szekvenciális predikciók)

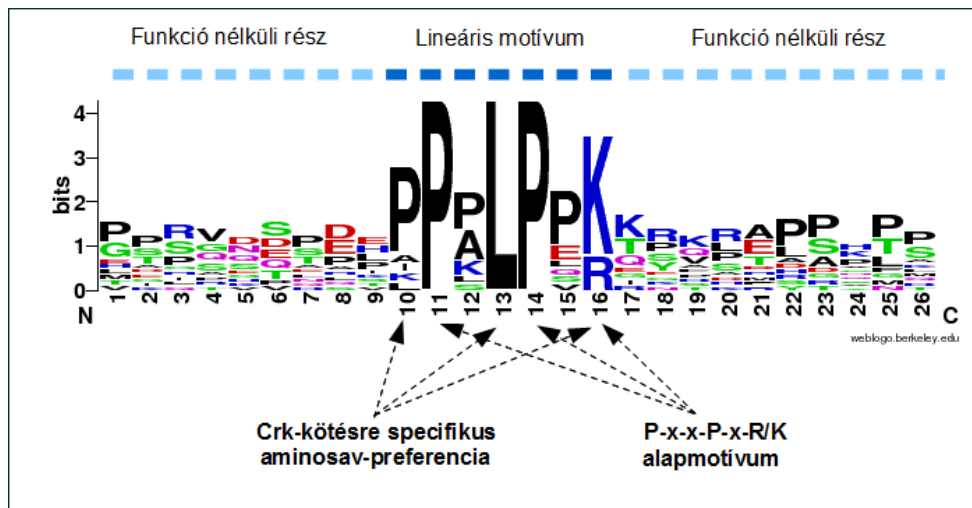
10.2.7. Automatizált módszerek a szekvenciák elemzésére

A bioinformatika egyik legnagyobb kihívása, hogy ismeretlen makromolekulák tulajdonságait kellő biztonsággal meg lehessen jósolni. Ezekre az előrejelzésekre (predikciókra) számos módszer áll rendelkezésre, itt csak a leggyakrabban használt néhány algoritmus alapjait fogjuk áttekinteni.

10.2.7.1. Pontozómátrixok (PSSM: *Position Specific Scoring Matrices*). Logók

Abban az esetben, ha csak egy-egy rövid, jól definiált szakasz felelős a DNS, RNS vagy a fehérje adott biológiai funkciójáért, viszonylag egyszerű feladat jóslatokat adni. Ezeket az elemeket szokták **lineáris motívumoknak** nevezni. Például az állatvilágban megtalálható úgynevezett SH3 domén egy konzervált szerkezeti elem, és olyan partnerfehérjékhez kötődik, ahol (a fehérje rendezetlen részén) két prolin egymástól három aminosav távolságra van, őket pedig megfelelő pozícióban egy arginin követi. Ilyesfajta szekvenciák azonban óriási számban vannak a fehérjékben: nagy részük nem tud SH3 doménekhez kötődni. Hogyan tudjuk mégis jó eséllyel megmondani, hogy melyek a valódi motívumok? Minden ilyen interakció a polipeptidlánc speciális szerkezetével függ össze, amely „összefér” a partner domén felszínével. A szerkezetnek azonban pontosan tükröződnie kell az aminosav-összetételben és sorrendben. Ha az ismert példák szekvenciáit megfelelő módon illesztjük, látni fogjuk, hogy nemcsak a Pro vagy Arg aminosavak a rögzítettek, hanem a köztes pozíciókban is határozott preferenciák láthatóak. Ezeket pozíció szerinti (20 tagú) vektorokkal lehet kifejezni, az összes vizsgált pozíció vektorai pedig egy pozíció specifikus mátrixot fognak kiadni, amelyet **PSSM**-nek rövidítünk (*Position Specific Scoring Matrix*). A szekvencia-preferenciákat leglátványosabban ún. **szekvencia logó** formájában lehet ábrázolni: Itt minden egyes betű (szimbólum) magassága arányos a mátrix neki megfelelő súlyával (ezt a gyakoriság értéket az y tengelyen az adott szimbólum

információ tartalmával, bit-ekben fejezik ki). (ld. 10.10. ábra) Magát a mátrixot pedig fel lehet használni a többi, potenciális motívum jóságának vizsgálatára is: azok a találatok lesznek a legesélyesebbek arra, hogy valóban működőképesek legyenek, amelyek szekvencia-eloszlása leginkább hasonlít a már ismert példákéra. Sajnos a módszer alkalmazása feltételezi azt, hogy nagyszámú, kísérletesen azonosított motívumot ismerünk, amelyek mind egyforma módon kötődnek a partner molekulák felszínéhez: a gyakorlatban ezért gyakran pontatlan eredményeket kapunk.

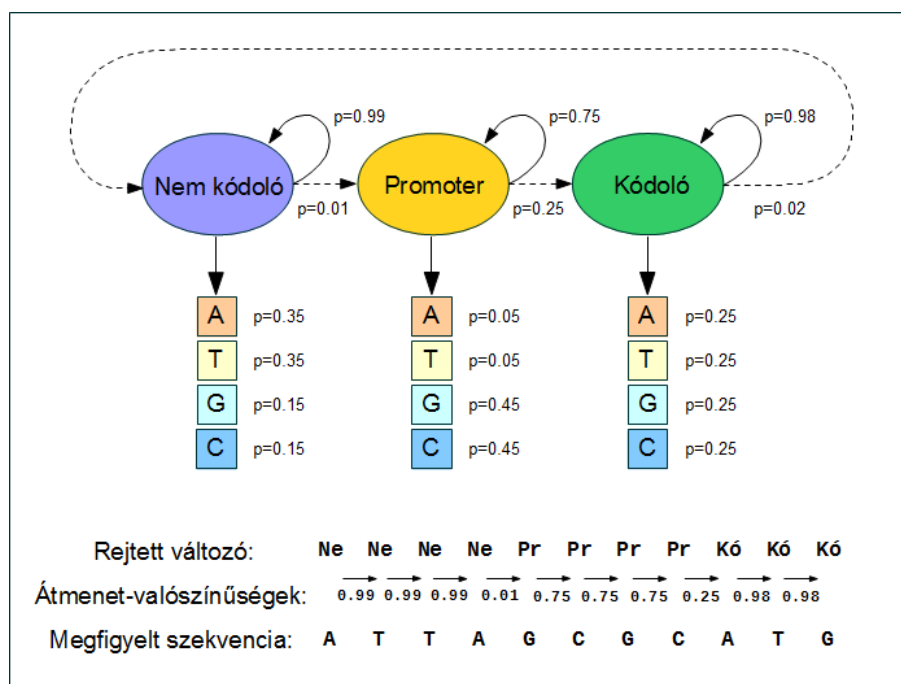


10.10. ábra: A humán Crk fehérjéhez kötődő peptidek szekvencia logója

10.2.7.2. Rejtett Markov-modellek (HMM: *Hidden Markov Models*)

A legtöbb biológiai funkció nemcsak egy-egy rövid motívum jelenlététől vagy hiányától függ, hanem nagyobb szerkezeti blokkok jelenlétét feltételezi. Ezek a blokkok azonban minden esetben lebonthatóak elemi motívumokra: a motívumok sorrendje pedig általában rögzített, a hossza viszont változatos lehet. Például egy eukarióta gén esetében a promóter régió után exonok következnek, közöttük intronok, azok után pedig egy poliadenilációs hely, mindegyik jellemző szekvencia-összetétellel. Ezen ismereteket fel lehet használni például a gének automatizált jóslására. A **rejtett Markov-modellek (Hidden Markov Models)** a bioinformatika "nagyágyú", amelyeket pontosan ilyen problémák megoldására dolgoztak ki (ld. 10.11. ábra).

Maga a modell nem más, mint egy gráf, amely állapotokból (csúcsokból) és állapot-átmenetektől (élekből) áll. Az egyes állapotok közötti átmenet mindig megadott valószínűséggel lehetséges. Az egyes állapotátmenetek pedig függetlenek a megelőzőektől (a fenti példában: az exonban haladva minden nukleotid után egyforma valószínűséggel folytatódhat az exon vagy következhet egy intron, és ez nem függ attól, hogy mennyi nukleotid volt már előtte). A statisztikában az ilyen rendszereket hívják **Markov-láncnak**.



10.11. ábra: Példa a rejtett Markov-modellek alkalmazására

Ez a lánc pedig azért rejtett, mert nem figyelhetjük meg közvetlenül a belső állapotokat (pl. intron/exon), csak egy attól függő külső változót (például a nukleinsav-összetételt). Vegyük a következő, egyszerű modellt: egy képzeletbeli organizmus génjeiben a nem kódoló régiók A/T gazdagok, a promóter régió nagyon G/C gazdag, míg a kódoló régió nukleotid eloszlása egyenletes. Az erre az ismeretekre felépített rejtett Markov-modell alapján pontosan ki lehet számolni hogy a genomban legvalószínűbben hol vannak a fehérje-kódoló gének. A trükk az, hogy minden egyes állapot átmenetkor összeszorozódnak a valószínűségek, és a végén vissza lehet számolni (számítógéppel) az egyes állapot-utakhoz tartozó egyedi valószínűségeket. Az efféle modelleket nagyon széles körben használják: de ezek felépítéséhez komoly mennyiségű kísérletes adatra van szükség.

10.2.7.3. Neurális háló módszerek (*Neural net*)

A biológiai gyakorlat tele van olyan problémákkal, ahol viszonylag kevés és nehezen rendszerezhető kísérleti adat áll rendelkezésre, és ezek alapján kell komplex jóslatokat adni egy másik rendszer viselkedésére. Az ilyen esetekre nehéz egzakt statisztikai modelleket alkotni. Viszont e nélkül is lehetséges jóslatokat adni, méghozzá olyan módszerekkel, mint a neurális háló algoritmusok. Maga a neurális háló nem más, mint az agykéregről "lemásolt" számítógépes modell. Az egyes **csomópontok** ("neuronok") egymás feletti **rétegekben** találhatók: a legfelső réteg kapja a direkt bemenetet (a nyers információt hordozó változók formájában), az ez alatti rétegek pedig mindig a megelőző réteg csomópontjaiban levő (számszerű) változók különféle **logiai kombinációt** kapják (pl. "és" operátor, "vagy" operátor, stb.). A végén, a legalsó réteg kimenete fogja megadni a kívánt jóslatot. A rendszer működéséhez előbb **"tanításra"** van szükség: ismert tulajdonságú példák feldolgozásával a neurális hálóban előbb rögzítjük a rétegek közötti logikai kombinációkat. Csak ez után következhet az ismeretlen minták elemzése. A neurális háló működésének hatékonysága a tanító példáktól függ: elegendő számú példával megfelelő minőségű modellt és viszonylag megbízható jóslatokat kaphatunk. Megjegyzendő azonban, hogy a neurális háló nem csodaszer: abban az esetben, ha egzakt modell felállítására megvan a mód, az gyakran jobb eredményeket nyújt, mint a neurális háló módszer.

10.2.7.4. A Bayes-statisztika és a döntésemélet alapjai (*Bayesian inference*)

A legutoljára említett, fontos matematikai módszer a **Bayes-statisztika**. Ezt nemcsak a bioinformatikai jóslásokra, hanem a kísérleti eredmények kiértékelésére is fel lehet használni (vagyis az elméleti modellek ellenőrzésére). Az egész Bayes-statisztika egyetlen fontos fogalommal kapcsolatos: ez a feltételes valószínűség. Amikor a kísérleti rendszerünkben egynél több véletlen változót figyelünk meg egyidejűleg, akkor nemcsak az egybeesések közös

valószínűségét lehet kiszámítani ($P(A=x, B=y)$): *Mi a valószínűsége hogy A és B egyszerre x illetve y értékű?*, hanem a **feltételes valószínűségeket** is, mint például $P(A=x|B=y)$ (*Mi a valószínűsége hogy A=x, ha azt találjuk hogy B=y?*). A feltételességet a képletbe tett függőleges vonal jelöli, és mindig az első változóról van szó, a másodiktól függően. A kétféle, "fordított" $P(A=x|B=y)$ és $P(B=y|A=x)$ feltételes valószínűségek általában különbözőek: itt a két ellenkező irányú **logikai következtetés**ről van szó. Vagyis: ha A-ból következik B, ez még nyilvánvalóan nem jelenti azt, hogy B-ből is következik A. Tegyük fel, hogy a kísérletünkben az A változó **rejtett**, nem mérhető (de ezt akarjuk tudni), viszont van egy tőle $P(B|A)$ -szerint függő B változó, amit meg tudunk **mérni**. Az általunk tanulmányozott B változó méréseiből szeretnénk az A-ra következtetni. Ha csak a B változót tudjuk megfigyelni, akkor a kísérleti rendszerből adódó $P(B|A)$ feltételes valószínűség önmagában még nem mondja meg hogy mi lehet az A változó legvalószínűbb állapota. Ehhez a fordítottját, $P(A|B)$ -t kellene tudnunk! Szerencsére van egy képlet, amely segítségével a kétféle feltételes valószínűség: $P(A|B)$ és $P(B|A)$ átszámolhatóak egymásba. Bayes tétele szerint:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Az egész döntésmélet lényegében ezen az egyetlenegy képletben alapszik. Mindig az A változó azon értékére kell szavaznunk, amelyre a **$P(A|B)$ valószínűség** (a mért B-vel) **maximális**. Ekkor lesz a döntésünk **hibája a legkisebb**. A fordított, $P(B|A)$ értékek a kísérleti rendszer alapján határozhatóak meg, a $P(A)$ valószínűségek pedig előzetes mérésekből, vagy akár az irodalomból becsülhetőek. Így **dönthetünk** két vagy több rivális hipotézis, vagy akár tudományos elmélet között is.

10.3. Nukleinsavak tulajdonságainak jóslása

10.3.1. Leolvasási keretek és gének jóslása

A modern világban, ahol az eddig ismeretlen organizmusok teljes genomjának szekvenálása ma már rutinszerű eljárásnak számít, égető szükség van olyan módszerekre, amelyekkel a nyers szekvenciákban teljesen automatizált módon meg lehet keresni a fehérjekódoló géneket. Az első és legfontosabb feladat minden ilyen esetben a helyes **nyitott leolvasási keret (ORF: open reading frame)** megtalálása. Mindez egy nagyon egyszerű elven alapszik: egy START kodon jelenlétén és a STOP kodonok hiányán. Mivel a nem-kódoló régiók nincsenek semmiféle nyomásnak kitéve a kodon-összetételüket tekintve, a STOP-ok véletlenszerűen gyakran fordulnak bennük elő. Ezt egyszerű módon mi is kiszámíthatjuk. Tegyük fel, hogy a vizsgált genom nukleotid-eloszlása (jó közelítéssel) egyenletes. Ebben az esetben minden három nukleotidonként $3/64$ valószínűséggel találunk stopot, tehát $p=0.047$ valószínűséggel ér véget a leolvasás és $1-p=0.953$ valószínűséggel folytatódik. Legalább két kodont $(1-p)^2$ valószínűséggel olvashatunk le, legalább hármat $(1-p)^3$ -nal, stb. A látszólagos ORF-ek hosszúsága így exponenciális eloszlást fog követni: $P(x) \sim (1-p)^x$, vagyis pontosabban:

$$P(x) = -\ln(1-p) \cdot \exp(x \cdot \ln(1-p)).$$

Ennek várható értéke:

$$-1/\ln(1-p) = 20,8 \text{ kodon}$$

(azaz 62,5 nukleotid). Ez lesz az átlagosan leolvasható szekvencia hossz, amennyiben véletlenszerű szekvenciát próbálunk lefordítani. Ugyanez a helyzet, ha egy kódoló régiót nem a megfelelő keretben próbálunk leolvasni. Ha a gyakorlatban azt találjuk, hogy több száz, esetleg több ezer nukleotid hosszan értelmes leolvasást kapunk, annak a valószínűsége, hogy ez véletlen lesz, igen kicsiny (a fenti eloszlás birtokában ezt pontosan ki is számolhatjuk). A módszer így nagyon alkalmas arra, hogy akár kézzel meghatározzuk a DNS szekvenciánk helyes leolvasási keretét, például a klónunk szekvenálása után. Ha nincs deléciónk vagy inszerciónk, általában az a keret lesz a helyes, amelyben a leghosszabban értelmes szekvenciát kapjuk. Így működnek az ORF-kereső prediktorok (pl: **ORF Finder**), és ezt használják ki az automatikus gén-annotáló szoftverek is (a prokarióta génekben nagyon ritkák az intronok). Ráadásul a legtöbb bakteriális gén igen jellemző promóter szekvenciákkal kezdődik (-35 régió és a -10 "Pribnow box" motívumok, Shine-Dalgarno szekvencia a transzlációs start hely előtt). A rejtett Markov-modell alapú modern génjósoló programok, mint pl. a GLIMMER így képesek a bakteriális genomok csaknem teljes génállományát előre jelezni.

10.3.2. Gén meghatározás cDNS alapján. Eukarióta gének analízise

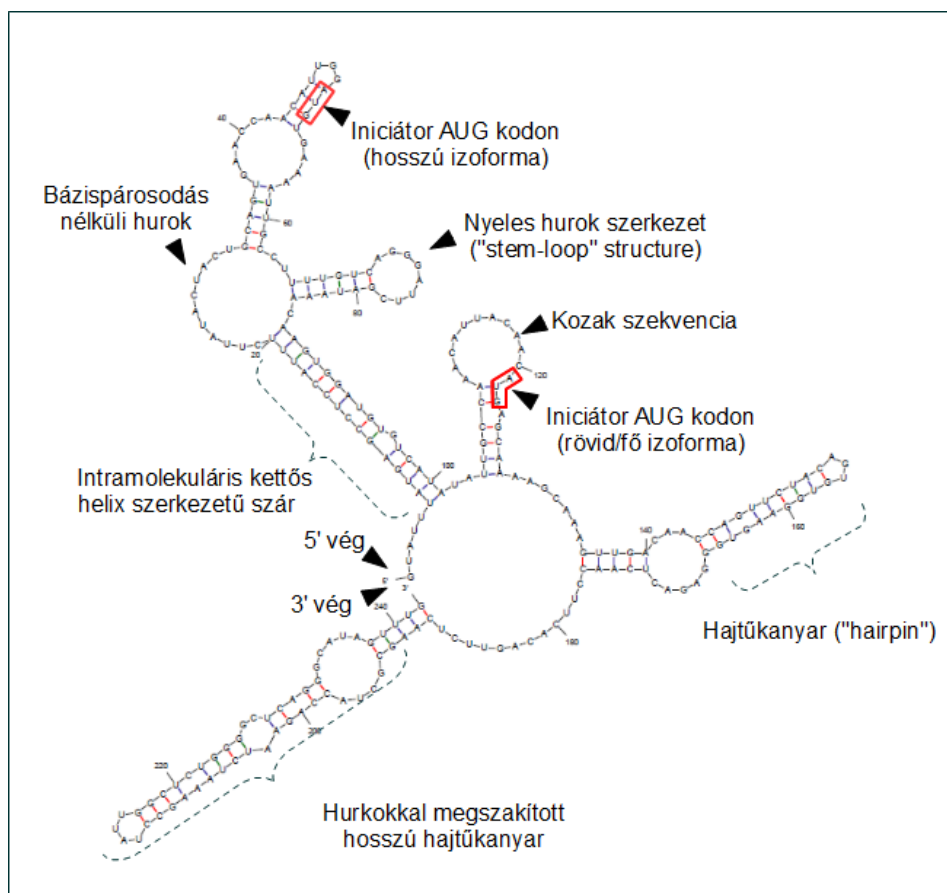
A fehérjét kódoló géneket természetesen kísérletes úton is meghatározhatjuk. A reverz transzkripcióval cDNS-re átírt, érett mRNS-ek szekvenciáit összevetve a genomikus DNS-sel, pontosan meghatározható az átíródott régió kiterjedése. Amennyiben a cDNS könyvtár nem tartalmazza a teljes hosszúságú mRNS-t, csak rövid darabokat, úgynevezett **EST**-ket (*Expressed Sequence Tag*), a helyzet valamivel nehezebb. Ilyenkor előbb számítógépes analízissel össze kell rakni az összetartozó EST-kből a teljes mRNS várható szekvenciáját (ahogy a genomi szekvenciát is a rövidebb kontigokból lehet összeállítani). A teljes cDNS ismeretére azért van szükség, mert így könnyen megadhatóak az intron-exon határok is. Noha az intronok kivágódási helyei (*splice sites*) jellemző szekvenciákkal rendelkeznek, az intronok és exonok - csak a genomi DNS-sen alapuló - jóslása ma is nagyon nehéz feladat. Ennek az egyik fő oka, hogy a szerkezetileg "unalmas" DNS-sel ellentétben az egyszálú RNS-eknek nagyon komplex térbeli szerkezete lehet, ráadásul rengeteg, ma még kevésbé ismert RNS-kötő fehérje befolyásolja az intronok felismerését. További nehézség, hogy az eukarióta promóterekben nincsenek kötelező elemek (a Pribnow-box homológja, a TATA-box az egyik leggyakoribb elem, de ez is csak a gének 10%-ában található meg). Ezen okok miatt az eukarióta gének automatikus analízise sokkal nagyobb kihívás, mint a prokariótáké. A kivágási (*splice*) helyek predikciói ma is meglehetősen pontatlanok: az adatbázisokban szereplő, genomikus DNS-ből jósolt fehérjékben bizony gyakoriak a hibásan megadott exonok. A translációs start hely jóslása is bizonytalan lehet. A legtöbb eukarióta mRNS-en a metilguanozin-sapka utáni első AUG kodont tekintik a translációs start helynek, de ez a sejtben nem mindig van így. Az eukarióta gének translációs kezdőpontját az első kodont körbevevő szekvencia környezet (az úgynevezett **Kozak-szekvencia**) határozza meg. Nagyon erős kontextus esetén előfordulhat, hogy nem AUG, hanem más (pl. CUG) kodonnal indul el a fehérjeszintézis. Gyenge kontextus esetén pedig a riboszóma "átugorhatja" az első AUG kodont. Sajnos ezeket a különlegességeket a mai prediktorok egyáltalán nem veszik figyelembe, ahogy az **IRES**-eket (*Internal Ribosome Entry Site*) sem. A szokatlan helyen kezdődő fehérjeszintézis kísérleti bizonyítása sem egyszerű (N-terminális fehérje szekvenálásra van szükség). Eukarióta gének elemzésére ezért újabban egyre nagyobb mértékben használnak evolúciós módszereket is: így az átírt szekvencia, az exonok pozíciója és a translációs start helyek jóval nagyobb megbízhatósággal jósolhatóak.

10.3.3. DNS és RNS szerkezetek jóslása. Oligonukleotid tervezés

A dezoxiribonukleinsav volt az első a biológiai makromolekulák közül, amelynek a szerkezetét sikerült exakt módon, röntgenkristallográfiával meghatározni. Ebben az is szerepet játszott, hogy a DNS szerkezete meglehetősen szigorúan rögzített: a dimer két szála közötti kapcsolódást az egymással szemben lévő bázisok párosodása és az egymás alatti bázisok közötti másodlagos kötőerők (delokalizált elektronpálya kölcsönhatások: π - π *stacking*) adják. Ez a szerkezet annyira stabil, hogy csak nagy ritkán - például az eukarióta telomereken - képesek az egymással szembeni monomerek "kihurkolódni". Mivel a sejtben megtalálható RNS-ek oroszlánrészének nincs komplementer párja, az ő esetükben a stabilitás alapja teljesen más lesz. A fő kölcsönhatásokat itt is a bázispárosodások és a bázisok átlapolódása fogja adni, de ezek az egyszálú molekulán belül, önmagukba hurkolódva fognak kialakulni. Az eredmény minden esetben egy komplex térszerkezet. A rövid, egymással kettőshélixben párosodó, hajtúkanyar szakaszok (*stems*) közötti részek kihurkolódhatnak (*loops*), néha pedig - ahol többféle szakasz párosodik egymás után - komplex csomók (*knots*) is létrejöhetnek. A belső szerkezetnek komoly szerepe van az RNS-ek biológiai funkciójában: elég csak a riboszómákra vagy a tRNS-ekre gondolni. A korábban említett Kozak szekvencia is egy "nyeles hurkot" (*stem-loop structure*) képez az iniciátor kodon körül: a szerepe nem közvetlenül a riboszóma kötése.

Mivel a nukleinsavak másodlagos (és harmadlagos) szerkezetének kialakulását lényegében csak a bázis-komplementaritás vezérli, kisebb molekulák esetén az alapvázat sokszor könnyű *in silico* megjósolni. Ilyenkor mód van az egyes lehetséges alternatív konformációk energiáinak becslésére, és a legstabilabbak kiválasztására. Számos szoftver áll rendelkezésre a nukleotidszálak szerkezetének kiszámítására (**RNA123**, **KineFold**). Talán az egyik legkényelmesebb módszer az **mFold** szerver használata (**ld. 10.12. ábra**). A természetes RNS-hez hasonlóan, a szintetikus, rövid DNS-oligonukleotidoknak szintén lehet belső szerkezete (**Primer Premier**, **DNASTAR**). Ha egy primernek szánt DNS oligonukleotid önmagába hurkolódik, az ronthatja vagy esetleg lehetetlenné is teheti a PCR reakció kivitelezését. Ezért újabban több erre szakosodott cég is javasolja, hogy *in silico* módszerekkel ellenőrizzük az oligonukleotidunk lehetséges belső szerkezeteit, az ahhoz tartozó elméleti energiákat és

"olvasáspontokat", mielőtt megrendelnénk a PCR primereinket. Jó esetben ezeknek a szerkezeteknek illik megszűnnie a PCR-hez használt anellálási hőmérsékleten



10.12. ábra: A humán JNK mRNA jóslott szerkezete az mFold programmal

10.3.4. Nukleinsavak "olvasáspontja"

A nukleotidok szerkeztőlásában szerepet játszó energiatényezőket egyszerűbb, és sokkal gyakorlatiasabb célokra is fel lehet használni. Ki lehet számolni például a nukleotidok és nukleotid-párok denaturációs hőmérsékletét: ezt a gyakorlatban legtöbbször csak "olvasáspontnak" (T_m : *melting temperature*) nevezik. Ez természetesen érzékenyen függ a DNS (vagy RNS) monomer-összetételétől: a C-G párok három, míg az A-T párok csak két hidrogénkötésre képesek, ami alacsonyabb kötési energiákat fog okozni az A/T gazdag molekulapárookban. A legrégebben ismert közelítés az oligonukleotid-párok olvasáspontjára az ún. **Wallace-szabály**:

$$T_m (^{\circ}\text{C}) = 2 \cdot N(\text{A}+\text{T}) + 4 \cdot N(\text{G}+\text{C}).$$

Itt az olvasáspontot közvetlenül az egyes párok típusából becsülhetjük ($N(\text{G}+\text{C}) = \text{G/C párok száma}$, $\text{GC}\% = \text{G/C párok aránya}$). Ez a szabály csak 20-40 nukleotid hosszúságú tartományban igaz, de ott is elég gyenge közelítésnek számít. Ennél sokkal jobb becslést ad - különösen az 50 nukleotidnál hosszabb párokra - az alábbi, szintén közismert képlet:

$$T_m (^{\circ}\text{C}) = 81.5 + 16.6 \cdot \log([\text{Na}^+]) + 41 \cdot \text{GC}\% - 600/L.$$

Itt figyelembe vettük nemcsak az oligonukleotid hosszát (L), de az oldat ionerősségét (pl. $[\text{Na}^+]$) is. Az eddig említett képletek azonban mind csak empirikus közelítések és meglehetősen korlátozott pontossággal igazak csak. Belső szerkezetek jóslására egyik sem alkalmas. A mai korszerű T_m becslések a kötési energia közvetlen számításán alapulnak. Tudjuk, hogy egy-egy komplementer nukleotid pár önmagában még nem képes stabil kötődésre (mert nem tud pi-pi kölcsönhatásra lépni a szomszédjával). Így nem az egyedi nukleotid párokat, hanem a párok párjait

vesszük figyelembe. A közelítés neve ezért "a legközelebbi szomszéd módszere" (*Nearest-neighbour method*). A kötési energiákat már minden szomszéd-párra megmérték, így könnyen rendelkezésre állnak (az összegük is). A számítás pedig a következő, ismert termodinamikai összefüggésen alapszik:

$$\Delta G = - R \cdot T \cdot \ln([A] \cdot [B] / [AB]).$$

Mivel T_m az a hőmérséklet, ahol a párok pontosan fele disszociált, ezért ha csak a dimer DNS-ből indultunk ki, akkor

$$[A] = [B] = [AB]_{\text{kiindulási}} / 2$$

Így a

$$T_m = - \Delta G / (R \cdot \ln([AB]_{\text{kiindulási}} / 2)).$$

A megfelelő, egyedi ΔS és ΔH értékek ismeretében ez számítógéppel könnyen kiszámítható. A só koncentrációval valamint a terminális nukleotidok gyengébb kötésével (A) korrigálva jó egyezést kapunk a kísérletes adatokkal. Maga a teljes képlet a következő :

$$T_m (^{\circ}\text{C}) = \Delta H \cdot 1000 / (A + \Delta S + R \cdot \ln([DNS]_{\text{dimer}} / 2)) - 273.15 + 16.6 \cdot \log[Na^+].$$

Ezt a képletet használják az internetes "oligó kalkulátorok", és a nukleotid szintézisre szakosodott cégek oldalai is. Sajnos 50-nél hosszabb nukleotidok esetében a szomszédossági módszer hatékonysága csökken: így ez a képlet nem jósolja meg egész plazmidok vagy kromoszómák denaturációs hőmérsékletét

10.4. Fehérjék tulajdonságainak jóslása

10.4.1. A fehérjék egyszerű fizikai tulajdonságainak becslése. Reverz transzláció

Hasonlóan a nukleinsavakhoz, a fehérjék számos elemi tulajdonságát is könnyen kiszámolhatjuk a szekvencia ismeretében. A molekulatömeg számítása akár papíron is elvégezhető, ugyanígy a moláris abszorbancia (extinkciós koefficiens) kiszámítása. Utóbbihoz csak az egyes, az UV tartományban elnyelő aminosavak elemi extinkciós együtthatóira van szükség. Mindezeket azonban könnyebb számítógépes programmal elvégezni. Vannak azonban olyan paraméterek, amelyeket "kézzel" kiszámolni meglehetősen fáradságos lenne: ilyen például az **izoelektromos pont** számítása. Egyedi, semleges aminosavak esetén, ahol az egyes savi (pK_a) és bázis (pK_b) állandók messze esnek egymástól, az állandók számtani közepével a **pI** könnyen megbecsülhető. Ez azonban nem így van a nagyobb peptideknél, vagy fehérjéknél. Mivel az egzakt számítás - még elhanyagolásokkal együtt is - hosszadalmas, számítógépes programok használata szükséges. Ezek az algoritmusok is általában közelítésekkel találják meg a pI értékét: nevezetesen, felveszik a becsült titrálási görbét, és megkeresik annak metszéspontját az x-tengellyel. Ehhez a számításhoz persze bizonyos elhanyagolásokat kell tenni. Ilyenkor is általában azt szoktuk feltételezni, hogy a pK_a és pK_b értékeket nem befolyásolja a molekuláris környezet, valamint azt hogy minden aminosav érintkezik az oldószerrel. Ezért a becsült pI is inkább csak a peptidekre érvényes, nagyobb fehérjékre nem, legalábbis natív állapotban nem. Bizonyos programok megjelenítik a teljes titrálási görbét: ez sokkal jobban alkalmas a tisztítási folyamat (pl. ioncsere) tervezéséhez, mint a pI egymagában, hiszen megadja az elméletileg várt, átlagos töltésszámot az illető pH tartományban. Ezek és sok más fizikai paraméter kiszámítására alkalmas internetes programcsomagot tartalmaz az **ExPaSy** oldala. Nukleinsavak és fehérjék tulajdonságait egyaránt jósolhatjuk a *Sequence Manipulation Suite* (**SMS2**) javascript alapú programjaival. Itt még komplexebb feladatokat is elvégezhetünk. Példaként említjük a reverz transzlációt: meg kell keresni egy ismert fehérje-szekvenciához tartozó legvalószínűbb DNS szekvenciát. Mivel a DNS-kódban gyakran többféle kodon kódol egy-egy aminosavat, ez a megfeleltetés nem egyértelmű. Ha azonban az a feladatunk, hogy tervezzünk egy olyan mesterséges gént, ami adott (esetleg szintén mesterséges) fehérjeszekvenciát kódol, akkor statisztikai megközelítést érdemes használni. Az egyes élőlények kodon-eloszlása ugyanis nem egyenletes: mindig vannak ritka kodonok, amelyhez tartozó tRNS is csak kis mennyiségben van jelen az illető organizmus sejtjeiben. Így tehát, a megfelelő mértékű fehérjeszintézishez érdemes a leggyakrabban használt kodont választani. Ez a megfelelő program és kodon-használati adatbázis segítségével könnyen kivitelezhető.

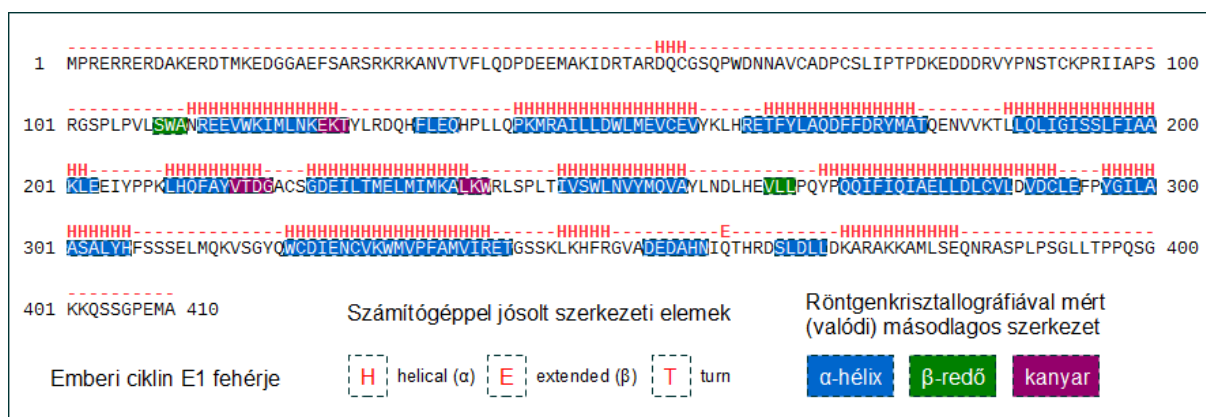
10.4.2. Lineáris motívumok keresése fehérjékben (és nukleinsavakban)

Ahogy azt korábban említettük, a fehérjék tulajdonságainak jó részét a bennük található aminosavak adott kombinációi (szekvencia-motívumok) határozzák meg. Amennyiben ezek egymás mellett, rögzített távolságra találhatóak, **lineáris motívumokról** beszélünk. Ezek az esetek többségében a fehérjék rendezetlen részein találhatóak és más molekulákhoz való kapcsolódásra szolgálnak. Az ilyen motívumok azonban alapvetően befolyásolják a polipeptidlánc sorsát és lehetséges módosításait. Például ha egy eukarióta fehérje N-terminális vége megfelelően elrendezett Arg/Lys és hidrofób aminosavakkal kezdődik (szignál peptid), azt a transzlokónok felismerik, így az extracelluláris vagy transzmembrán helyzetű lesz. A szignál peptid előrejelzésére ma már nagyon jó algoritmusaink vannak (pl. a neurális háló alapú **SignalP**). Ha a fehérje ezenkívül több rövid (~15-30 aminosav), tisztán hidrofób szakaszt is tartalmaz, akkor valószínűsíthetően többszörösen transzmembrán szerkezetű lesz a sejtekben. Meglepő módon, a membránfehérjék **topológiáját** (intracelluláris / extracelluláris szakaszokat) is relatíve magas megbízhatósággal tudjuk jóslni, olyan programokkal, mint a **Phobius** (amely rejtett Markov-modell alapú). Van mód a sejten belüli lokalizáció megjósolására ismert példák alapján, neurális háló algoritmusával is: ilyen a **WolfPSort** program. Az exportra szánt fehérjén glikozilációs helyek lehetnek, amelyek szintén előre jelezhetők (**NetNGlyc**). Az endoplazmás retikulum fehérjéire pedig a „visszatartó” motívum (idealizálva: KDEL, Lys-Asp-Glu-Leu) jellemző, ez például pontozómátrixokkal található meg könnyűszerrel. Egyelőre azonban nincs egyszerű módszer a Golgi-apparátus fehérjéinek megtalálására. A mitokondriális, kloroplaszt és peroxisomális fehérjék viszont mind-mind jellemző N-terminális vagy C-terminális tranzit- vagy import-szignálokkal rendelkeznek, amelyek segítségével *in silico* is nagy valószínűséggel azonosíthatók (pl. **TargetP**-vel).

A citoplazmában maradó fehérjéken is különféle lineáris motívumok irányítják a legtöbb poszt-transzlációs módosítást. A foszforiláció vagy ubiquitiniláció általában szigorúan szabályozottan történik, de az enzimet felismerő (kináz-kötő vagy ubiquitin-ligáz kötő) motívumok gyakran távol esnek a ténylegesen módosított helytől. Más motívumok, például a proteázok felismerő helyei, leggyakrabban az aktuális módosítás helyén találhatóak. Ezeket a rendkívül sokrétű módosításokat lehet akár egyszerre is jóslni. Az **ELM** (*Eukaryotic Linear Motifs*) szervere erre a feladatra specializálódott. A megadott fehérjeszekvenciában megkeresi az ismert motívumokkal összeférő szakaszokat, sőt még pontozást is végez, mégpedig szerkezeti (rendezetlenségi) és evolúciós konzerváltsági alapon. Az ismert kötőelemeket és módosítási helyeket a program külön kiemeli. Az eddig azonosított több száz különféle lineáris motívum többsége nem a módosítások irányítására szolgál, hanem egyszerűen csak adott fehérje komplexek összeszerelésére. Ilyen motívumokat keresni a DNS-en is lehet, sőt érdemes is: A transzkripció faktorok kötőhelyei például DNS-alapú lineáris motívumok, amelyeket pontozómátrix módszerrel lehet azonosítani a genom különféle helyein. Ugyanígy jóslhatóak a másodlagos (például DNS metiláció) módosítási helyek is (pl. az úgynevezett CpG szigetek)

10.4.3. Fehérjeszerkezetek jóslása. Rendezett és rendezetlen régiók

A 21. század egyik legnagyobb informatikai kihívása a fehérjék szerkezetének szekvencia alapú jóslása. Ellentétben a nukleinsavakkal, a fehérjék szerkezetét a legváltozatosabb oldallánc-oldallánc kölcsönhatások tartják össze, így a lánc helyi összetétele nem jóslja meg még a másodlagos szerkezetet sem: a háromdimenziós kényszerfeltételek ismeretére van szükség. Amennyiben mindenféle előzetes ismeret nélkül szeretnénk **térszerkezetet jóslni** (*ab initio folding*), egy nagyon erősen számítógép igényes feladat elé nézünk. Ilyen szimulációkat például a **ROSETTA** programmal lehet végeztetni, de a futásidők ma még általában elégtelenek a nagyobb, stabil szerkezetek meghatározására.

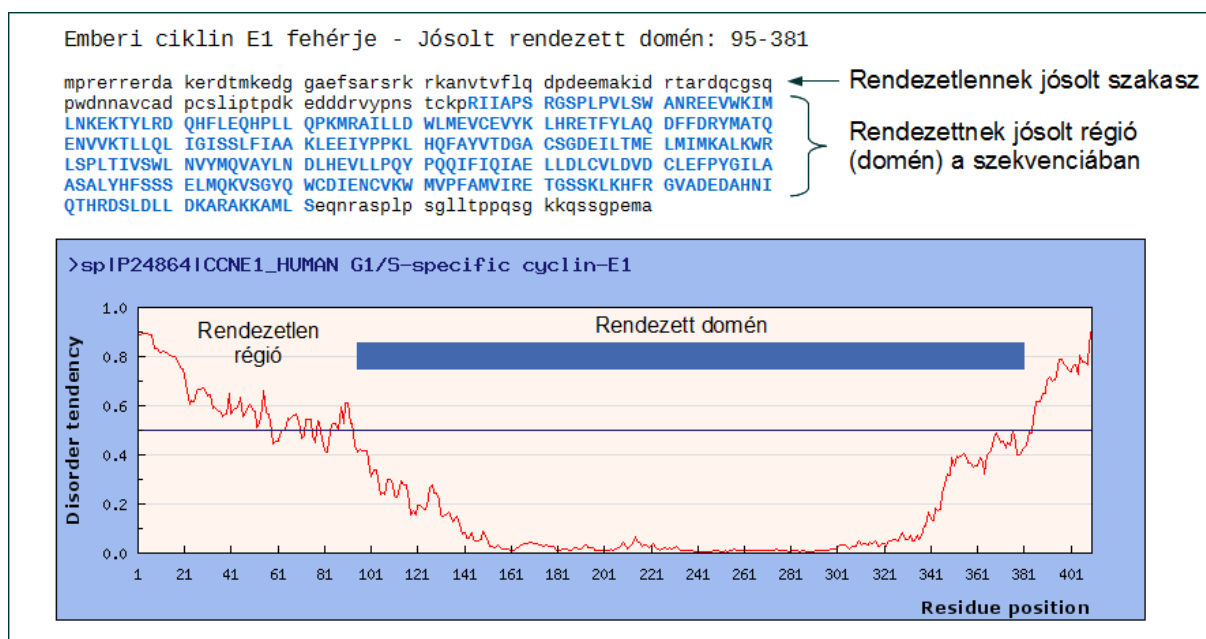


10.13. ábra: A ciklin E1 fehérje Jpred programmal jóslt másodlagos szerkezete és a valós másodlagos szerkezet összehasonlítása

Valamivel könnyebb dolgunk van, ha csak azt tűzzük ki célul, hogy szeretnénk a másodlagos szerkezeti elemeket megjósolni. Az olyan prediktorok, mint a **JPred** könnyen megtalálják a másodlagos szerkezeti elemeket, de a valódi típusukat (α-hélix, β-redő, kanyar) sokszor elvétik (**ld. 10.13. ábra**).

Például a β-redőket gyakran hélixnek jósolják, vagy fordítva. Általában elmondható, hogy a stabil szerkezet jelenlétének ténye sokkal könnyebben jelezhető előre, mint maga a konkrét szerkezet. Ha csak az a célunk, hogy elkülönítsük a rendezett doméneket a szerkezet nélküli, flexibilis régióktól, már ma is számos megbízható program áll rendelkezésre. Ezeket elsősorban a rendezetlenség - aminosav-összetétel alapú - jelzésére dolgozták ki (pl. **DisoPred**, **PondR**), de a módszer megfordításával képesek a domén jellegű, rendezett régiókat is megadni (pl. **IUPred**). Az eljárás előnye, hogy a teljesen ismeretlen doméneket is előre lehet vele jelezni (de a szerkezetüket természetesen nem) (**ld. 10.14. ábra**).

Az ismert szerkezetű domének azonosítása leghatékonyabban a homológia-alapú keresőkkel végezhető el. Az egyik legszélesebb körben használt ilyen kereső szerver a **PFAM** (*Protein Families*). A PFAM az összes fontosabb ismert domén egyedi motívum-lenyomatát tartalmazza, amelyeket (rejtett Markov-modellek segítségével) végigfuttat az általunk megadott szekvencián. A statisztikailag erősen szignifikáns találatokat visszaadva, megkaphatjuk a fehérjénk homológia alapú szerkezeti modelljét. Ilyen kereséseket más szervereken is lehet végeztetni. A **ProSite** (ExPaSy) programcsomag például nem csak a lineáris motívumok keresésére alkalmas: a keresője számára megadhatunk akár teljesen egyedi domén definíciókat is, ha újszerű doméneket szeretnénk azonosítani. Persze az a tény, hogy két domén belseje hasonlít egymásra, még nem adja meg automatikusan a választ, ha a rendezettség pontos külső határait szeretnénk tudni: ehhez konzerváltsági analízisekre is szükség lehet. A domének illetve azok határainak ismerete nagyon fontos a géntechnológiai alkalmazások számára: A fehérje fragmentumok klónozását mindig úgy kell tervezni, hogy ne vágjunk bele a rendezett domének belsejébe. Ilyenkor ugyanis (jó eséllyel) egy nagyon instabil, nehezen oldható és mindenhez aspecifikusan, véletlenszerűen kötődő fehérjét kapnánk. E miatt korábban sajnos nagyon sok fehérje-fehérje kölcsönhatás-térképezési kísérlet bukott meg.



10.14. ábra: Rendezett domének és rendezetlen régiók jósálása az IUPred programmal

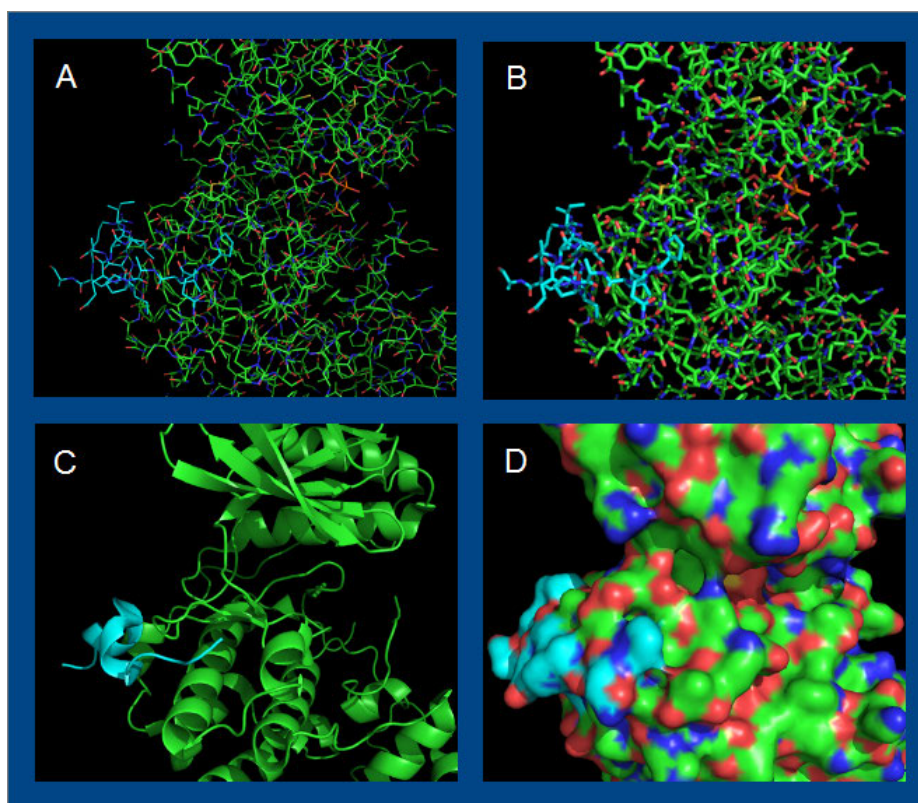
10.5. *In silico* szerkezeti analízisek

10.5.1. Makromolekulák grafikai megjelenítése

Bár a makromolekulák térszerkezetének kísérletes meghatározása atomi felbontásban nehéz feladat, ennek ellenére egyre növekszik az ismert fehérje, fehérje-fehérje és fehérje-nukleinsav komplex szerkezetek száma. A legfontosabb szerkezeti adatbázis a **PDB**. Az itt elhelyezett szerkezeteket általában kísérletes úton, röntgenkristallográfiával (*X-ray diffraction*), vagy magi mágneses rezonancia-spektroszkópiával (NMR) határozták meg. Az adatbázisból letölthető file *.pdb kiterjesztéssel rendelkeznek: ez tartalmazza az atomok elemi koordinátáit. Ezen felül számos egyéb fontos adatot szoktak mellékelni (annotáció): ilyenek például a szerkezet „jóságát” jelző statisztikai adatok, a szerkezeti szimmetriák, és a méréshez felhasznált fehérje fragmentumok pontos szekvenciája. Az atomi koordinátákat erre a célra alkotott molekulás rajzoló programokkal jeleníthetjük meg „élvezhető” formában. Rengeteg ilyen vizualizációs program létezik: **SwissPDBviewer**, **Chimera**, **RasMol**, **JMol**, stb. A legtöbbjük ingyenesen letölthető és használható.

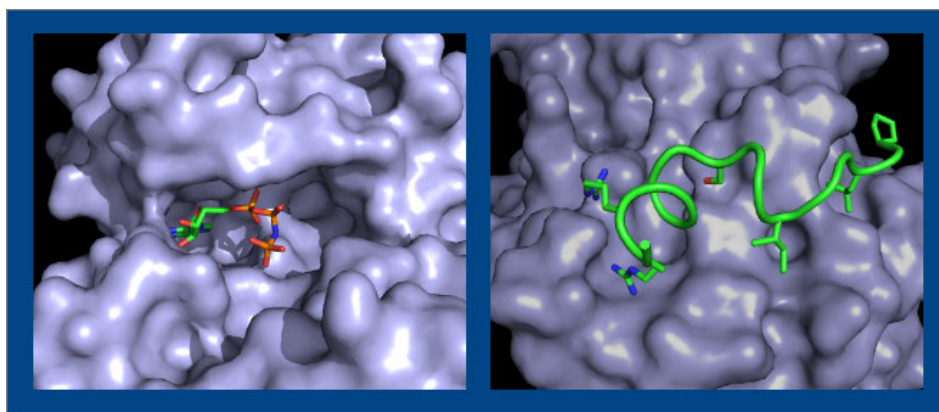
A gyakorlatban talán a legelterjedtebbnek számít a **PyMol** program használata (ennek használatához előbb a Python programnyelvet is telepíteni kell). A megjelenítésnek számos fajtája lehetséges (ld. 10.15. ábra).

A legegyszerűbb módon csak az atomokat összekötő kémiai kötések ábrázoljuk, színes vonalak formájában (*lines*). A színeket a vonalak egyes atomok alapján kapják, de mód van az átszínezésre az egyedi láncok alapján is: a komplexek így könnyebben megfigyelhetők. Pusztán az a tény, hogy egy szerkezetben egynél több fehérjelánc látszik, önmagában még nem igazolja a komplexek létét: a molekulák a kristályosodás során is összeállhatnak (mesterséges dimerekké, oligomerekké). A számunkra kevésbé érdekes láncokat, aminosavakat vagy atomokat a kijelölésük után egy kattintással átlátszóvá téve akár el is tüntethetjük.



10.15. ábra: **Fehérjeszerkezetek megjelenítése PyMol programmal.** Drótmodell (A), pálcika modell (B), *cartoon* („rajzos”) modell (C) és a molekulafelszín ábrázolása (D)

A vonalas modelleknél persze vannak sokkal szebbek is: ilyen például az "iskolai kémiai építőkészlet" golyó és pálcika (*stick*) modellje: ez főleg kis molekulák esetén nagyon látványos (**ld. 10.16. ábra**). A vonalas vagy pálcika-modellek nem igényelnek nagy számítási kapacitást, de az oldalláncok már a néhány tucat monomerből álló makromolekulák esetén is áttekinthetetlen "dzsungelt" eredményeznek. Nagy molekulák esetén ezért gyakrabban használjuk a fő láncot szalaggal ábrázoló (*ribbon*) vagy a rajzos (*cartoon*) módot.



10.16. ábra: Kismolekulás ligandumok (A) és fehérje-fehérje kölcsönhatások (B) látványos megjelenítése különféle modellek kombinációjával, PyMol programmal

Rajzos modellben a nukleinsavak cukor-foszfát gerince vastag csőnek, míg a nukleinsav bázisok laposcskáknak fognak látszani. A fehérjék béta-redős régióit lapos nyilakkal, az alfa-héliceket "hajcsavarokkal", míg a többi fehérjeszakaszt vékony csővel jelölik a programok. A publikációkban is leggyakrabban ezzel a megjelenítési móddal találkozhatunk. Fontos megjegyezni, hogy az eddig említett modellek egyike sem veszi figyelembe az atomok valódi térkitöltését. De ennek ábrázolására is van mód: akár hálóval (*mesh*), akár teljes felszínnel (*surface*). Fehérjék és kisebb molekulák kölcsönhatása különösen látványossá tehető, ha a kis molekulát pálcikákkal, a fehérjét

pedig a felszint mutató módban ábrázoljuk. A felszín ábrázolásakor színezhethetünk is: nem csak az egyes atomok típusai, hanem a felszín karaktere szerint is: például ábrázolhatjuk a hidrofobicitást vagy a töltéssűrűséget. Ha pedig két rokon molekulát szeretnénk összehasonlítani, arra is van mód: a legtöbb ilyen megjelenítő program tartalmaz "térbeli illesztést" lehetővé tevő algoritmusokat. A PyMol "*align*" és "*super*" parancsa az elsőként megadott molekulát addig tolja és forgatja a térben, amíg az a másiktól minimális távolságra és orientációba nem kerül (az atomi koordináták "euklideszi" különbségeit tekintve).

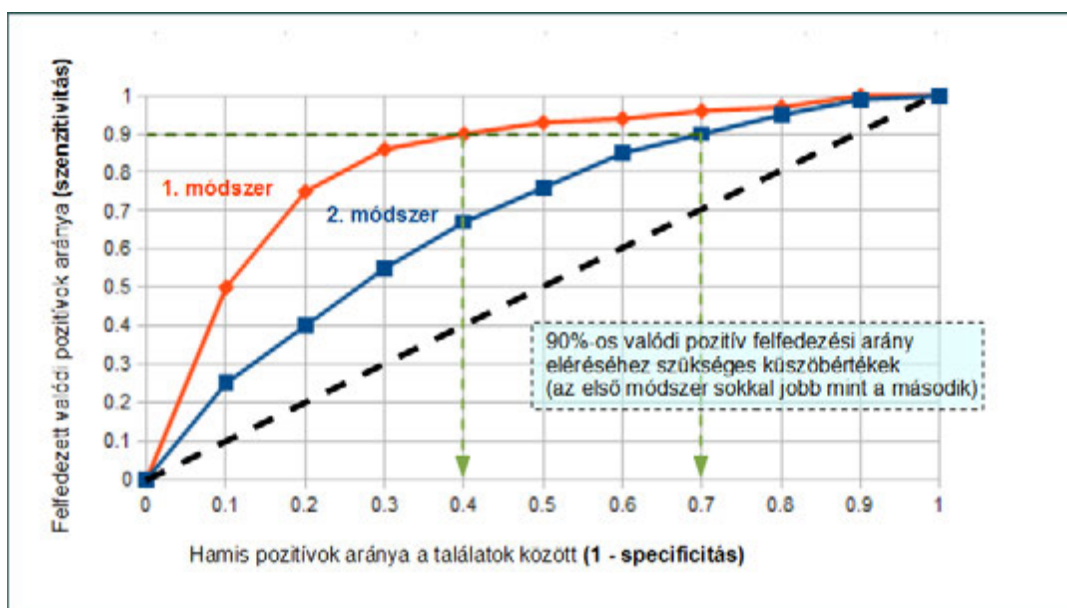
10.5.2. Homológia-modellezés és *in silico* dokkolás

A makromolekulák térbeli szerkezetének elemzése nem csak "kézzel" lehetséges. Az egyes atomi koordináták ismerete lehetőséget teremt arra, hogy különböző szerkezeteket automatizált módon összehasonlítsunk, sőt, akár ismeretlen szerkezeteket is előre jósoljunk, méghozzá a **homológia-modellezés** módszerével. Erre elsősorban akkor van mód, ha az általunk vizsgált fehérje domén szerkezete ugyan nem ismert, de egy vagy több közeli rokonáé igen. A homológia-modell felállítására jól használható például a **SwissModel** szerver: Bemeneti információnak egy minél pontosabb páronkénti illesztést kell megadnunk, és az ismert partner szerkezetét mutató fájlt (vagy pdb kódot). Az eljárás során a program nagyrészt változatlanul hagyja a fehérje fő láncának lefutását, de az oldalláncokat az új domén szekvenciája alapján fogja újjáépíteni. Természetesen kezelnie kell az inszerciók/deléciók esetét is, méghozzá a polipeptidlánc lefutásának észszerű korrekciójával. De a szerkezetek ismerete nemcsak homológia-modellezésre használható, hanem ismeretlen partnerek kötődésének modellezésére is. Ennek mindenekelőtt a gyógyszermolekulák tervezésében van óriási jelentősége. A legtöbb, kis szerves molekula esetén a belső konformációs átmenetek szerepe elhanyagolható, így a kis molekula teljesen rigid darabként kezelhető. A módszer neve ezért "merev-test dokkolás" (*Rigid body docking*). A számítógép a gyógyszermolekulát a "végtelenből" közelítve a célzott felszínhez, megkeresi a lehetséges helyi energiaminimumokat, beillesztve a kis molekulát a fehérje megfelelő kötőzsebébe (**AutoDock**). Ilyen szimulációkat peptidok esetére is végezhetünk: itt azonban a lehetséges konformációk nagy száma nehezíteni fogja a számítást. Ha rendelkezünk egy megbízható (homológia-) modellel a várható végállapotról, a kötési energia akár a fő lánc konformációjának megváltoztatása nélkül is megbecsülhető: így működik a **FoldX** nevű szimulátor program. Ez főleg olyankor használható, ha egy szerkezetileg jól ismert fehérje-fehérje kölcsönhatásról van szó. Természetesen arra is van mód, hogy a peptidlánc konformációja véletlenszerűen változtatható legyen. Ilyet kínál fel a (Rosetta alapú) **FlexPepDock** szerver. Noha helyesebb modellnek tűnik, ezen megközelítésnek van egy komoly hátránya: a peptidlánc modellje hajlamos beragadni a lokális energia-minimumokba, így a módszer általában alábecsüli a valóságban mérhető kötési energiát.

10.6. A predikciós módszerek pontossága

10.6.1. Egy predikció "jóságát" leíró paraméterek. Optimális döntések.

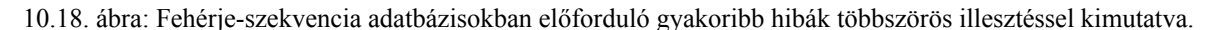
Az előrejelző módszerek részletes tárgyalása után nem kerülhetjük el, hogy magáról az előrejelzés minőségéről ne beszéljünk, ezt kell ugyanis mérlegelni akkor, amikor arról döntünk, hogy elfogadjuk-e egy módszer jóslatát vagy sem. A predikciók jóságának tárgyalásához előbb néhány statisztikai fogalmat kell bevezetnünk. Egy két-kimenetelű (+/-) mérés, elemzés vagy döntés esetén szenzitivitásnak nevezzük a sikeresen felismert elemek arányát (az összes pozitív elem közül), specificitásnak pedig a sikeresen elvetett elemek arányát (az összes negatív elem közül). A legtöbb eddig leírt algoritmus rendelkezik egy vagy több változtatható paraméterrel, ami befolyásolja a módszer jóságát. Például az illesztések magas p-érték mellett nagyon magas szenzitivitással (megtalálják az összes egyezést), de alacsony specificitással (rengeteg véletlen egyezést is találnak) fognak rendelkezni. Alacsony p-érték (mint határérték) esetén megfordul a helyzet: az összes egyezés jó lesz (magas specificitás), de a valódi egyezések csak kis hányadát találjuk meg (alacsony szenzitivitás). A döntésre használt határértéket szoktuk levágási értéknek (*cutoff*) nevezni. Egy-egy levágási érték jóságát pedig - a mérnöki tudományokból kölcsönvett módszerrel - az úgynevezett ROC görbével fejezhetjük ki. Az **ROC** (*Receiver Operating Characteristics*) **grafikon** pontosan mutatja, hogy egy adott paraméter-érték, mint levágási érték esetén a valódi pozitívaknak mely részét találtuk meg, illetve hogy az összes lehetséges elem hányad részét fogadtuk el (**ld. 10.17. ábra**). Az optimális döntéseket a valódi pozitívak arányának becslésével (Bayes-analízis alkalmazásával) ilyen görbék alapján már viszonylag könnyű megadni (főleg a tévedés "költségének" ismeretében). E szerint minden módszer rendelkezik egy rá jellemző, fix hibavalószínűséggel, még az optimális döntés esetén is: ezek léte a módszer lényegéből fakad.



10.17. ábra: ROC görbék használata bioinformatikai jósló módszerek jóságának elemzésére

10.6.2. Adatbázisok minősége: elsődleges és másodlagos hibák

A jóslataink helyességét nemcsak az eredendő hiba befolyásolja. Az adatbázisokban tárolt szekvenciák minősége is lehet hibaforrás. Könnyű belátni, hogy ilyen mennyiségű biológia adat mérésénél és elemzésénél hiba nélkül dolgozni szinte lehetetlen. Az adatbázisokban előfordulhatnak elsődleges hibák: ezek a mérési hibákból fakadnak, például egy-egy kétes minőségű szekvenálásból. Érdekes tény például (és könnyű megfigyelni) hogy a legkorábban szekvenált modell- organizmusok (pl. *Mus musculus*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*) genomi és cDNS szekvenciája sokkal több hibával rendelkezik, mint az újabban szekvenáltaké. Így ha azt találjuk, hogy az egér egy adott génjének adott aminosava eltér az összes többi emlősben megfigyelhető, mindig konzervált aminosavtól, az eltérés eredete vélhetően szekvenálási hibában keresendő: nem arról van szó ugyanis, hogy az egér "feltalált" volna számos, amúgy teljesen szokatlan biológiai megoldást. Mivel a legtöbb élőlény cDNS-szekvenciái nem, vagy csak darabokban hozzáférhetőek, nagyon sok, az adatbázisokban található fehérjeszekvencia valójában csak jóslat (ez a tény fog remélhetőleg megváltozni az RNAseq új-generációs szekvenálás elterjedésével; **ld. 5.3. fejezet**). Az ilyen, prediktált szekvenciákban pedig viszonylag gyakoriak a másodlagos hibák: tévesen jóslt transzlációs kezdőhelyek, eltévesztett exonok (valójában intron-szekvenciák, exonként átírva) és hiányzó szakaszok. Ezek az annotáló-program tévedéséből fakadnak. Az ilyen hibákat leginkább a biológiai korrelációk segítségével, többszörös szekvencia-illesztésekkel szűrhetjük ki. Ha például a ló egy adott, genomi szekvenciából jóslt fehérjéjének közepén található egy szokatlan, hosszú darab, ami az összes emlőstől különbözik (de máshol mindig konzervált), akkor ebben az esetben egy hibásan jóslt exonról lehet szó (**ld. 10.18. ábra**).



Az elsődleges adatbázisok minősége javítható, ha nem csak automatizált módszerekkel előállított adatokat tartalmaznak. Az ilyen, szakemberek által ellenőrzött rendszerek (*manually curated databases*) közé tartozik például a UniProt fehérje adatbázis (**Id. 10.1.3.**). Itt a gépi analízist kísérleti adatokkal váltják fel, amikor azok a szakirodalomban elérhetővé válnak. Ezek a kézi annotációk jelentős emberi munkaerőt kötnek le, cserében viszont a legmegbízhatóbb adathalmazt nyújtják. Persze a szakirodalomban leírt kísérletek sem mindig mentesek a hibáktól: mérési hibák éppúgy előfordulhatnak, mint hibás értelmezések. Ebből fakadóan találkozhatunk egymásnak tökéletesen ellentmondó publikációkkal is: hol az egyik, hol a másik, esetleg mindkét kutatócsoport tévedett.. Mindezek rontják a "kézi" annotációk pontosságát is. Ezért a szakemberek által ellenőrzött adatokat sem szabad feltétlenül hibátlanként kezelni.

Mindezek után bizonyára felmerül az olvasóban a kérdés: érdemes-e egyáltalán előrejelző programokat használnunk? És ha igen, akkor mikor? Általános arany szabálynak annyit mondhatunk, hogy a kísérletes adatok általában pontosabbak, mint a jóslatok. A prediktáló módszereket is mindig a rendelkezésre álló információ alapján válasszuk meg. Tehát amennyiben mód van pl. egy fehérje doménjének homológia alapú elemzésére, ne kezdjünk *ab initio* folding szimulációkba. Fontos megjegyezni, hogy a kutatás maga kétirányú: a már kísérletesen igazolt eredményeket is érdemes újra számítógépes analíziseknek alávetni, és új modelleket építeni rájuk.

Edwards, D, Stajich, J, & Hansen D (2009) Bioinformatics: Tools and Applications. Springer Science. ISBN: 978-0-387-92737-4.

De Fonzo, V, Aluffi-Pentini, F, Parisi, V (2007) Hidden Markov Models in Bioinformatics. *Current Bioinformatics*. 2: 49-61.

Dickter Wickmann (1999) Bayes-statisztika. ELTE Eötvös Kiadó. ISBN: 978-963-463-311-2

133

11. fejezet - *In vitro* mutagenézis

Az 1980-as években kialakuló és napjainkig töretlenül fejlődő géntechnológiai módszerekkel a rekombináns fehérjéket kódoló DNS-ben – így a humán fehérjéket kódoló DNS-ben is – mára egyszerűen, olcsón és gyorsan hozhatók létre mutációk, akár irányítottan is. A DNS-en történő változtatás lehet delécio, amikor tetszőleges számú nukleotid eltávolítása történik; inszercio, amikor tetszőleges számú nukleotid beillesztése történik; vagy szubsztitúció, ami egy nukleotid másik nukleotidra való kicserélését jelenti. Az *in vitro* mutagenézis kidolgozásával beindult a **reverz genetika** korszaka.

Miért lehet szükség a fehérjéket kódoló DNS vagy a vizsgálandó **gének szabályzó elemeinek célzott vagy véletlenszerű megváltoztatására**? Egyrészt azért, mert az egyes biológiai folyamatok vizsgálata és megértése gyakran a folyamatok részlépéseinek defektusain keresztül történik. A klasszikus genetikai vizsgálatokban az egyes gének vagy szabályzó elemeinek funkciójára, az adott folyamatok működésére csak azok természetes mutációinak és változásainak fenotípus vizsgálatával volt lehetőség. Az örökítőanyagban bekövetkező változások a természetben azonban nem célzottak és nem érintik az összes régiót, ezért számos esetben a fehérjék, fehérjerészletek vagy génjük szabályzó elemeinek funkciója nem vizsgálható. A kutatók érdeklődésének központjában gyakran az ember áll. A géntechnológiai módszerek kifejlesztése előtt a humán fehérjék funkciójának vizsgálata is csak a természetben előforduló mutációk és rendellenességek vizsgálatával volt lehetséges. Az **irányított mutagenézis módszerekkel rekombináns fehérjékben** – akár humán fehérjékben is – gyorsan és hatékonyan tudunk véletlenszerű vagy akár **célzott változásokat** létrehozni. Az egyes aminosavak más kémiai karakterrel rendelkező aminosavakra cserélésével vagy akár nagyobb fehérjerészek mutációjával a fehérjék tulajdonságait megváltoztathatjuk. A megváltozott tulajdonságból következtetni tudunk az adott aminosavnak vagy a fehérjerésznek a fehérje enzimaktivitásában, szerkezetében vagy funkciójában betöltött szerepére.

Az *in vitro* mutagenézis másik felhasználási területe a **racionális fehérjemérnökség**, vagy az irányított fehérje evolúció, melyek során irányított vagy nem-irányított mutációs módszerek segítségével fehérje variáns sereget hozhatunk létre (**ld. 16. fejezet**). Ebből a variáns seregből általunk meghatározott szelekciós körülmények között végeredményben megváltoztatott, valamilyen szempontból számunkra előnyösebb, vagy akár újfajta tulajdonsággal rendelkező fehérjék, enzimek, fehérje inhibitorok szelektálhatók (pl.: proteáz inhibitorok; stabilabb és gyorsabb enzimek).

A rekombináns fehérjék *in vitro* mutációs módszerekkel történő vizsgálata nagymértékben hozzájárult a modern biológia mai sikerességéhez. A DNS *in vitro* mutagenézisére mára számos módszert kifejlesztettek. Ebben a fejezetben a teljeség igénye nélkül ismertetésre kerülnek a leggyakrabban használt, a legismertebb és a leghatékonyabb nem-irányított és irányított mutagenézis módszerek. A mutagenézis módszerek kihagyhatatlan lépése és szerves része a mutáns DNS szekvencia ellenőrzése, aminek a legbiztosabb módszere a DNS szekvenálása. A különböző szekvenálási és ellenőrzési módszerekről részletes információ a **5. fejezetben** található.

11.1. Nem-irányított *in vitro* mutagenézis módszerek

Nem-irányított mutagenézis módszerek gyakran használatosak irányított fehérje evolúciós módszerek esetében, mint például az *in vitro* kompartmentáció, a **fág bemutatás** vagy a **baktérium bemutatás** (**ld. 16. fejezet**). Az irányított fehérje evolúciós módszerek során a természetben történő evolúciós történéseket utánozva (diverzifikáció, szelekció és replikáció) egy variáns seregből általunk meghatározott körülmények között szelektálódnak a számunkra megfelelő tulajdonsággal rendelkező könyvtárak. A módszer első lépése a megfelelő variáns sereg létrehozása. A nem-irányított *in vitro* mutagenézis módszerekkel létrehozott variánsok tulajdonságai a mutáció fajtájától függően akár nagymértékben is eltérhetnek a kezdeti templátként (vázként) használt fehérje tulajdonságaitól. Végeredményben szolubilisabb, szerkezetileg stabilabb vagy akár teljesen újfajta tulajdonsággal, újfajta enzimaktivitással rendelkező fehérjék, enzimek is létrehozhatók.

11.1.1. Mutátor baktérium sejtvonalak

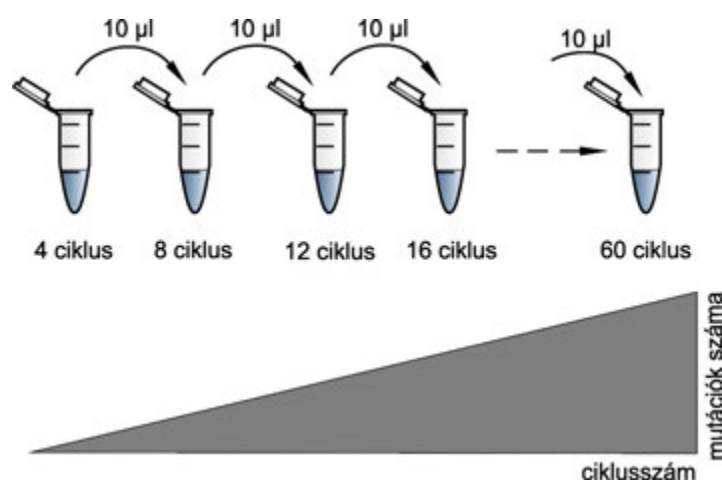
Különféle mutátor baktérium sejtvonalak használatával gyorsan és egyszerűen lehet nem-irányított mutációt létrehozni a vizsgálandó rekombináns fehérje DNS szekvenciájában. A módszer alapja, hogy a rekombináns fehérjét

kódoló DNS-t tartalmazó plazmid sokszorosításához használt baktérium sejtvonalakban nem működnek megfelelően a DNS hibajavító rendszerek. Például plazmid vektorok sokszorosításához általánosan használt *Escherichia coli* baktérium egyik sejtvonalaiban, az **XL1-Red**-ben (*endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac mutD5 mutS mutT Tn10 (Tet^r)*) hiányzik a DNS hibajavító útvonalban szerepet játszó három „kulcs” fehérje; a mutS, a mutD és a mutT. A mutS fehérje az „error prone mismatch” hibák javításban, a mutD fehérje a DNS-polimeráz III 3'→5' exonukleáz aktivitásában, a mutT fehérje az oxidatív folyamatok során történő DNS károsodás javításában vesz részt. A nem megfelelő DNS hibajavítás 5000-szer nagyobb mutációs gyakoriságot eredményez ennél a sejtvonalnál a vad típusú *Escherichia coli* sejtvonalakhoz képest. A módszer előnye, hogy két egyszerű lépéssel, a plazmidnak a mutátor sejtekbe transzformálásával, majd a sejtek felnevesztése után a már mutáns plazmid DNS-ének kinyerésével gyorsan, sokfajta mutáció (deléció, szubsztitúció, inszerció) létrehozása lehetséges. Egyéb géntechnológiai módszerek, mint restriktációs emésztés, PCR vagy ligálás nem szükséges. A módszer hátránya, hogy a mutációs ráta a mutátor sejtvonalak használata mellett is igen alacsony (0,5 mutáció/kbp). Nagyobb mutációs rátát a sejt kultúra fenntartási idejének megnövelésével lehet elérni. A mutátor sejtvonalak a DNS hibajavító rendszer hiánya miatt azonban igen gyorsan halmoznak fel mutációkat saját genomjukban és a plazmidnak nem a vizsgálandó fehérjét kódoló DNS részében is. Megfelelő méretű variáns sereg létrehozása ezért gyakran ismételt transzformálási és plazmid tisztítási lépéseket igényel.

11.1.2. Hibát ejtő (error prone) PCR mutagenézis

Az egyik leggyakrabban használt nem-irányított mutagenézis módszer az úgynevezett hibát ejtő (*error prone*) PCR. Alapja, hogy a vizsgálandó fehérjét kódoló DNS szakasz PCR általi sokszorosítása során a DNS-polimeráz számára nem optimális körülményeket használnak. Nem megfelelő körülmények között a DNS-polimeráz mutációs frekvenciája megnő. A felszaporított DNS-ben a mutációk véletlenszerűen keletkeznek, számuk a ciklusszámmal nő. *Error prone* PCR során általában a *Thermus aquaticus* baktériumból izolált, **alacsony hűségű** („*low fidelity*”) Taq DNS- polimerázt használnak. A Taq-polimeráznak hiányzik a 3'→5' exonukleáz aktivitása, ezért ha a polimeráz a lánchosszabítás során nem a DNS bázispárosodás szabályainak megfelelő bázist épít be, akkor nincs lehetősége a hiba javítására. A Taq-polimeráz optimális reakciókörülmények közötti működése során átlagosan egy szubsztitúciós mutáció történik 10^4 bázispáronként. A Mn^{2+} , a Mg^{2+} , a DNS-polimeráz és a dNTP koncentrációjának megnövelésével stabilizálni lehet a nem komplementer bázispárosodást, csökkenthető a bázispárosodás specificitása, így továbbá növelhető a rosszul anellálódott primer oligonukleotidról történő lánchosszabítás valószínűsége. Összességében a Taq-polimeráz mutációs gyakorisága elérheti ciklusszámonként a 0,066%-ot (30 ciklus során). A mutációk feldúsulását tovább növelhetjük a PCR során alkalmazott ismételt hígítási lépésekkel és a ciklusszám növelésével (60 ciklus; ld. 11.1. ábra). Ezzel a módszerrel összességében a felsokszorozandó DNS szekvenciájának akár a 3,5%-a is megváltozhat.

Az *error prone* PCR során deléciók és inszerciók is előfordulhatnak, jellemzően azonban szubsztitúciók történnek. A legtöbb polimeráz esetében nagyobb valószínűséggel történik A→G vagy T→C tranzíció (purin→purin, piridin→piridin), mint transzverzió (purin→piridin, purin→piridin). A négy dezoxiribonukleozid-trifoszfát arányának eltolásával és módosított Taq-polimeráz enzim (Mutazyme) használatával ez a mutációs aránytalanság (*mutational bias*) javítható. Az *error prone* PCR használatával létrehozott DNS inszert ezután megfelelő restriktációs enzimmel való emésztés után a kívánt plazmidba ligálható. A módszer hátránya, hogy a létrehozható variáns sereg méretét korlátozza a szubklónozási lépés során alkalmazott ligálás hatékonysága.



11.1. ábra: A hibát ejtő (*error prone*) PCR során alkalmazott hígítási módszer. A hibát ejtő (*error prone*)

PCR mutagenézis során nem optimális körülmények között működő DNS-polimeráznak megnő a mutációs frekvenciája (ciklusszámonként a DNS 0,066%-a változhat meg). A ciklusszám növelésével növelhető a DNS-ben keletkező véletlenszerű mutációk száma. A hígítási módszer minden 4. ciklusa után a minta körülbelül 10%-át átoljták egy következő reakcióelegybe. Az átoljtással frissítik a reakciót, így növelhető a ciklusszám egészen 60-ig. Összességében 60 ciklus után a DNS szekvenciájának akár 3,5%-a is megváltozhat.

11.2. Helyspecifikus *in vitro* mutagenézis módszerek

Clyde Hutchison és Marshall Edgell már 1971-ben felvetette a kis DNS fragmentekre épülő irányított mutagenézis ötletét (DNS kazetta mutagenézis, ld. 11.2. ábra). Hutchison és Michael Smith 1978-ban kifejlesztett egy általánosan használható, szintetikus oligonukleotidokra épülő módszert, mellyel mind inszerció, delécio és szubsztitúció is létrehozható. Az eredeti protokoll szerint a mutálandó DNS fragmentumot először FX174 fág DNS-be klónozták. A mutációt tartalmazó szintetikus oligonukleotidot ezután az egyszálú FX174 fág DNS-re anellálták, majd DNS polimeráz és ligáz használatával szintetizálták és cirkularizálták a mutációt tartalmazó új szálát. Az azóta kifejlesztett helyspecifikus *in vitro* mutagenézis módszerek mindegyike szintetikus oligonukleotidot használ a rekombináns fehérjék tetszőleges módosítására.

A szintetikus oligonukleotidok tervezésénél ajánlott, hogy az oligonukleotid a mutációtól mind 5'-, mind 3'-irányban is legalább 15 bázispárnyi szakaszon komplementer legyen a templát szállal és a 3' végén lehetőleg G/C bázisokat tartalmazzon. A mutációtól 5'- és 3'-irányban található 15 bp általában elegendő a mutáció specifikus kijelölésére. A mutáció jelentheti egy bázispár megváltoztatását vagy az oligonukleotid „kihurkolódásával” érinthet akár egy nagyobb régiót is. Ezekkel a módszerekkel általunk szabályozott módon bármely pozícióban megváltoztatható a fehérjét kódoló DNS szekvenciája. Az irányított mutagenézis során bármilyen „szabás-varrás” kivitelezhető; inszerció, delécio vagy szubsztitúció. Lehetőség van fehérjék vagy fehérjerészek fúziójára, kimérák létrehozására. is Az alábbiakban a legelterjedtebb és a leghatásosabb módszerek kerülnek bemutatásra.

11.2.1. DNS kazetta mutagenézis

Hutchison és Edgell által leírt DNS kazetta mutagenézis a legegyszerűbb és leghatékonyabb irányított helyspecifikus mutagenézis módszer (ld. 11.2. ábra). A mutálandó DNS fragmentumot egy dupla szálú plazmid tartalmazza. Két restriktions endonukleáz emésztésével hasítjuk ki a mutálandó DNS szakasz környezetét, majd a megfelelő ragadós véggel rendelkező, mutációt tartalmazó szintetikus oligonukleotidokkal elkészített DNS kazettát T4 ligáz segítségével a plazmidba ligáljuk és *E. coli* baktériumba transzformáljuk. A DNS kazetta mutagenézis egy gyors, hatékony és könnyen kivitelezhető módszer. Fontos szempont, hogy a restriktions enzim felismerőhelyek a mutálandó DNS szekvenciához közel helyezkedjenek el, így csak 60-80 bp hosszú DNS kazetta szintézise szükséges. Ilyen méretű oligonukleotid szintézise olcsó és gyors. Ha nagyobb méretű DNS kazetta szükséges (100-1000 bp), a szintézis költségei jelentősen megnövekednek. Ezekben az esetekben az alábbiakban is említett, rövid oligonukleotidokat igénylő módszerek költséghatékonyabbak.



11.2. ábra: **DNS kazetta mutagenézis módszer.** A dupla szálú DNS plazmidot a mutálandó DNS szekvencia közvetlen környezetében restrikciós endonukleázzal emésztjük. A kieső DNS fragmentum helyére ligáljuk a mutációt tartalmazó szintetikus DNS kazettát.

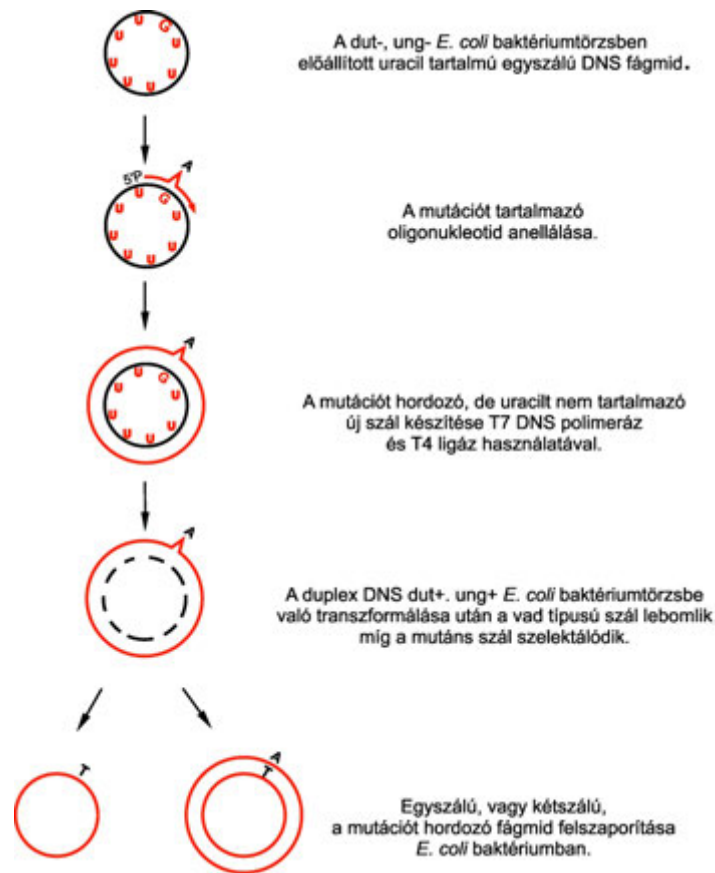
11.2.2. Szelektív templát lebontáson alapú mutagenézis módszerek

A Smith és Hutchinson által kifejlesztett irányított mutagenézis módszer (**ld. 11.2.**) hatékonysága igen kicsi. Elméletben a mutációs ráta maximum 50%-os lehet, de a gyakorlatban a nem megfelelő oligonukleotid letapadás, a polimerizáció és a ligálás hiányos működése miatt a legjobb esetben is csak pár százalékot ér el. A módszer hatékonyságának növelése érdekében az újonnan **létrejövő mutációt tartalmazó plazmid DNS-t szelektálni kell a vad típusú plazmid DNS-től.**

11.2.2.1. Kunkel mutagenézis

A rekombináns fehérjék irányított mutagenézisének egyik legrégebbi, de néhány területen még a mai napig is használatos módszerét Thomas Kunkel 1985-ben közölte (**ld. 11.3. ábra**). A módszer nagy hatékonysággal képes általunk meghatározott pozícióban inszerciók, deléciók és pontmutációk létrehozására. A módszer alapja, hogy a mutálandó DNS inszertet egy **M13 bakteriofág alapú egyszálú DNS fágmid** tartalmazza. Erre, mint templátra, a **mutációt tartalmazó oligonukleotidot** anellálnak, majd T7 DNS-polimeráz használatával szintetizálják a már mutációt tartalmazó új szálát. A mutációt tartalmazó plazmid végül megfelelő baktérium törzsekben szelektálható.

Az első lépés a mutálandó DNS inszert M13-alapú fágmidba klónozása. Az inszertet tartalmazó fágmidból ezután egyszálú DNS-t hoznak létre, ami a mutagenézis során templátként fog szolgálni. A M13 fágmid vektor ötvözi a kétszálú DNS plazmidok és az egyszálú DNS fág tulajdonságait, amiről nevét is kapta (**ld. bővebben a 7.3.2. fejezetben**). *E. coli* sejtekben képes kétszálú DNS plazmidként szaporodni (ez az ún. replikatív forma). M13 *helper* fág hatására azonban egyszálú DNS másolatok jönnek létre, amelyek a helper fág által termelt M13 bakteriofág burokkészítménybe csomagolódva M13 virionokat alakítanak ki. Az M13 bakteriofág egy nem-litikus bakteriofág, a kialakuló virionok a baktérium kultúra felülszójába szekretálódnak, feldúsulnak, ahonnan könnyen tisztíthatók és belőlük egyszálú DNS izolálható.



11.3. ábra: **Kunkel mutagenézis módszer sematikus ábrája.** A mutálandó DNS fragmentumot hordozó fágmidból egyszálú, uracil tartalmú DNS templátot állítunk elő dut-, ung- *E. coli* baktériumtörzsben. A mutációt tartalmazó, szintetikus előállított oligonukleotidot a templát DNS-re anelláljuk, majd T7 DNS polimerázzal „körbeírjuk” azt, végül az újonnan szintetizált szálát T4 ligázzal cirkularizáljuk. A duplex DNS-t dut+, ung+ *E. coli* baktériumtörzsbe transzformáljuk, melyekben a mutáns szál szelektálódik és a vad típusú szál lebomlik. A mutáns fágmid tovább szaporítható, belőle kétszálú vagy egyszálú DNS hozható létre.

Második lépésben az egyszálú templát DNS-hez a mutációt tartalmazó oligonukleotidot anellálnak. Egy mutációs lépésben akár több oligonukleotid is használható, így egymástól távolabb található DNS szakaszok egyszerre is megváltoztathatók. A következő lépésben T7 DNS-polimeráz használatával történik meg az egyszálú templát fágmid „körbeírása”, a kétszálú heteroduplex DNS létrehozása (az egyik szál az eredeti „vad típusú” templát, a másik szál a mutáns). A T7 DNS polimeráz egy processzív polimeráz, azaz képes hosszú, akár másodlagos szerkezettel rendelkező templát átírására is anélkül, hogy a templát szárlól disszociálna. Érdekes az oligonukleotidok 5'- és a 3'-végére lehetőleg G/C bázisokat tervezni, hogy elkerüljük a rossz helyre anellálást (*mispriming*) és ezáltal a nem-megfelelő termékek keletkezését. A 3'-végen található G/C bázispárok biztosítják a T7 DNS- polimeráz pontos működéséhez szükséges kezdő pozíciót. Az oligonukleotid 5'- végén található G/C bázispárok pedig megakadályozzák a T7 DNS polimeráz túlműködését, az oligonukleotid 5' végének a templát DNS szárlól való „letérését”. Az újonnan szintetizált, már a mutációt tartalmazó DNS szálát T4 DNS-ligáz segítségével cirkularizálják. A T4 DNS-ligáz működéséhez szükséges a szintetikus előállított oligonukleotid 5' végének foszforilálása, amit az oligonukleotidnak a templát szárhoz való hozzáadása előtt polinukleotid-kináz enzimmel lehet elvégezni.

A kialakuló kétszálú heteroduplex DNS fágmidot *E. coli* sejtekbe transzformálják és így a továbbiakban plazmidként szaporítható. Smith módszeréhez hasonlóan az ilyen módon létrehozott fágmidoknak a legjobb esetben is csak az 50%-a fogja tartalmazni a mutációt. A Kunkel-módszer során ez az 50%-os mutációs ráta a templát DNS szál szelektív lebontásával csaknem 100%-ig eltolható. Ennek alapja, hogy a templátként szolgáló egyszálú fágmid timin helyett nagyrészt uracilt tartalmaz. Ezt úgy érik el, hogy az egyszálú DNS templátot olyan *E. coli* törzsben szaporítják, ami **dUTP-áz és uracil N-glikozidáz deficiens** (*E. coli* CJ236 K12 dut-, ung-, *F pilus*+). A dUTP-áz hiánya miatt a fágmid DNS szintézise során a sejtekben a dUTP koncentrációja megnövekszik és timin helyett gyakran **uracil is beépül a DNS-be**. Az uracil N-glikozidáz felelős a DNS-be már beépült uracil eltávolításáért,

ezért az uracil N-glikoziláz deficiens törzs nem tudja kijavítani a DNS-be már beépült uracilt. A mutációt tartalmazó második szál T7 DNS-polimerázzal történő *in vitro* elkészítéséhez dUTP-t nem használnak fel, ezért az újonnan szintetizált szál uracilt nem tartalmaz. A következő lépésben a vad típusú, uracilt tartalmazó, és a **mutáns, uracilt nem tartalmazó heteroduplex DNS-t vad típusú *E. coli* törzsbe (*dut*⁺, *ung*⁺) transzformálják** (*E. coli* XL1-Blue vagy DH5-alpha). Az uracilt tartalmazó vad típusú DNS a *dut*⁺, *ung*⁺ *E. coli* sejtekben az uracil-N-glikoziláz enzim hatására preferenciálisan lebomlik, míg a mutáns változat jóval nagyobb arányban fennmarad, így növelve meg a mutációs rátát. A sejtekből a mutáns kétszálú fágmid plazmid tisztítható, illetve M13 helper fág használatával fág és egyszálú fágmid is termelhető.

A módszer előnye, hogy ha a mutálandó inszertet a továbbiakban is fágmid vektorban szeretnénk fenntartani, akkor ez nem igényel további klónozási lépéseket, mint például restrikciós endonukleázzal való emésztést vagy ligálást. Hátránya, hogy a fágmid vektoroknak korlátozott az alkalmazhatósága, így például ha a vizsgálandó rekombináns fehérjénket heterológ expresszióval *E. coli* baktériumban szeretnénk túltermeltetni, akkor a mutációt tartalmazó inszertet szubklónoznunk kell megfelelő expressziós vektorba (pl. pET). A heterológ expresszióról a **12. fejezetben** találhatók további részletek.

11.2.2.2. QuickChange mutagenézis

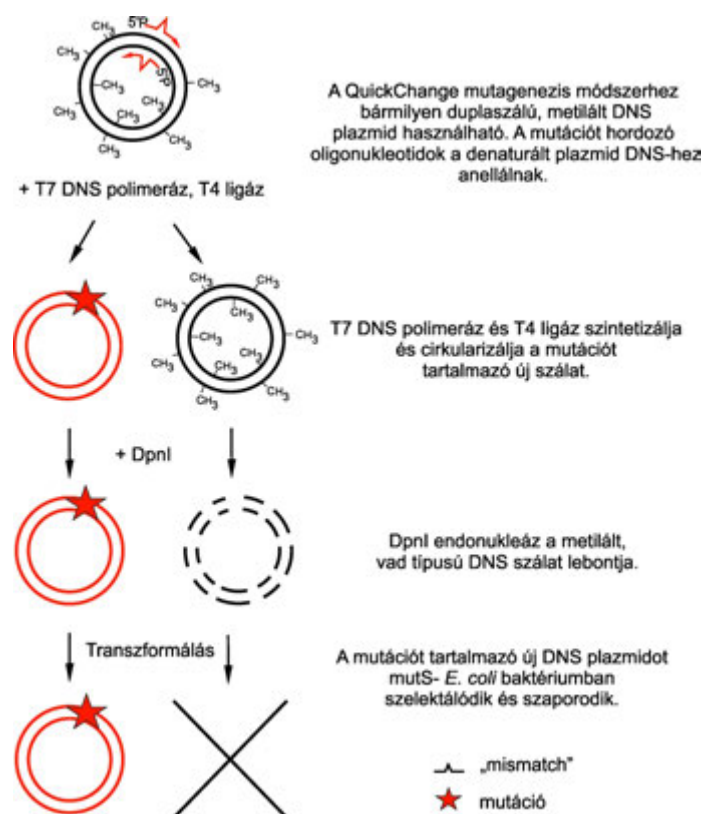
A Stratagene cég által kifejlesztett „QuickChange” mutagenézis az irányított mutagenézis módszerek közül, mint neve is mutatja, talán a leghatékonyabb és leggyorsabb módszer (**ld. 11.4. ábra**). Deléció, inszerció és szubsztitúció is létrehozható vele. Előnye, hogy nem igényel speciális vektorrendszert, mint a Kunkel mutagenézisnél alkalmazott M13 fágmid, valamint összehasonlítva a PCR-alapú „megaprimer”, vagy az „overlapping” mutagenézis módszerekkel (**ld. 11.2.3.1. és 11.2.3.2.**), nincs szükség további klónozó lépésekre (restrikciós emésztés, ligálás) sem.

A módszer során a mutálandó DNS fragmentumot bármilyen kétszálú plazmid tartalmazhatja. A kivitelezésnek **előfeltétele**, hogy a használni kívánt plazmid **vektor guanin bázisai metiláltak legyenek**. E feltétel teljesítése bármely géntechnológiában általánosan használt (*dam*⁺) *E. coli* törzs esetében teljesül, mint például az **XL1-Blue vagy a DH5-alpha törzsek** esetében (**ld. 7. fejezet**). A plazmidot vagy lúgos oldattal (NaOH kezeléssel), vagy hőkezeléssel denaturálják. A mutagenézishez olyan **oligonukleotid primer párt** használnak, melyek ugyanazon DNS-szakasz két szálához hibridizálnak és tartalmazzák a kívánt mutációt. A primerek tehát egymás komplementerei. Érdemes nagyon magas ($T_m > 78^\circ\text{C}$) olvadási hőmérsékletű oligonukleotidot tervezni. Az olvadási hőmérsékletet megbecsülhetjük a következő egyszerű képlet használatával:

$$T_m = 81.5 + 0.41(\text{GC}\%) - 675/N - \text{mismatch}\%$$

ahol az N az oligonukleotid hossza, a GC% és a mismatch% egész számok. Egyszerre több primerpár használatával több, akár egymástól távolabbi DNS szakaszon is létrehozhatunk változást. A következő lépésben egy PCR során DNS-polimeráz (Pfu, KODI, Fusion/Phusion) használatával történik meg a mutációt tartalmazó új szálak szintézise.

A reakció végül vad típusú és mutációt tartalmazó plazmidokat és ezek heteroduplexeit eredményezi. A következő lépésben a **reakcióterméket DpnI restrikciós endonukleázzal kezelik**, ami képes magas sókoncentráció mellett (150-200 mM NaCl) **specifikusan a metilált DNS lebontására**, így a vad típusú templát plazmidok eliminálására (célszekvenciája: 5'-Gm6ATC-3'). A kezelés után a mutációt tartalmazó plazmidok intaktak maradnak. A DpnI enzimmel kezelt mintát, amely elvileg már csak a metilálatlan, mutáns plazmidokat tartalmazza, *E. coli* baktériumba transzformálják, ahol megtörténik a plazmid két végének ligálódása. A DpnI endonukleáz és a DNS-polimeráz minőségétől függően általában maximum 50%-ban sikeres a reakció. A mutáció hatékonysága 80-100%-ra növelhető, ha uracilt tartalmazó templátot használunk és a mutagenézis után a DNS termékünket *dut*⁺, *ung*⁺ *E. coli* baktérium sejtekbe transzformáljuk. Az uracil tartalmú templátot dUTP-áz és uracil N-glikoziláz deficiens *E. coli* baktérium sejtekben állíthatjuk elő (*E. coli* CJ236 K12 *dut*⁻, *ung*⁻, *F pilus*⁺).

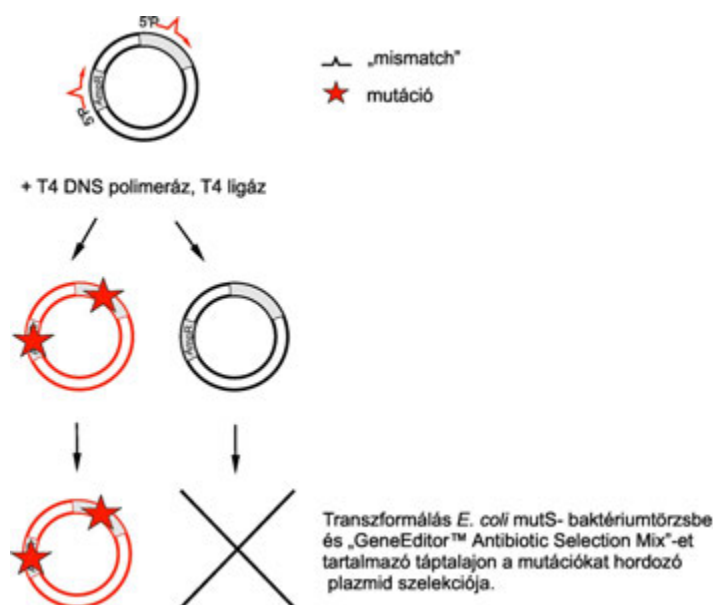


11.4. ábra: **QuickChange mutagenesis sematikus ábrája.** A módszer során bármilyen, dam⁺ *E. coli* sejtben termelt kétszájú, metilált plazmid használható. A plazmidok hő- vagy lúgos denaturációja után mindkét szálra komplementer, a mutációt hordozó oligonukleotidot anellálnak. T7 polimeráz és T4 ligáz szintetizálja és cirkularizálja az újonnan képződő, a mutációt már hordozó szálát. A metilált, vad típusú DNS-t DpnI endonukleázzal magas sókoncentráció mellett specifikusan lebontják. A reakcióterméket mutS- *E. coli* sejtekbe transzformálják, melyekben a mutáns szál szelektálódik és szaporodik. Az ábrán a piros csillag a mutációt jelöli.

A „QuickChange” mutagenesis előnye a Kunkel-módszerrel szemben, hogy nem igényel speciális sejt vonalakat és vektorrendszereket. A PCR alapú módszerekkel szembeni előnye pedig, hogy nincs szükség további klónozási lépésekre. Hátránya, hogy az egész plazmid átírásra kerül, ezért nagyon jó minőségű, nagy hűségű DNS-polimerázt igényel. Nagyméretű plazmidok esetén (> 8 kbp) ezért speciális körülményeket és enzimeket kell használni (részletesebb leírást a gyártó cég honlapján itt olvashatnak).

11.2.2.3. GeneEditor mutagenesis

A Promega cég által kifejlesztett „szelekciós” oligonukleotidra épülő „GeneEditor” módszere a „QuickChange” mutagenesis egyik alternatívájaként szolgál. Minden általánosan használt ampicillin rezisztenciát (TME1 b-laktamáz) hordozó egyszájú és kétszájú plazmidban kódolt DNS fragmentum irányított megváltoztatására alkalmas (ld. 11.5. ábra). A mutáció során egyszerre két oligonukleotidot használnak fel. Az egyik magát a kívánt mutációt tartalmazza. Egyszerre egymástól távolabb található több mutáció is létrehozható több mutációs oligonukleotid együttes használatával. A másik típusú, az úgynevezett „antibakteriális szelekciós” oligonukleotid, a plazmidon kódolt ampicillin rezisztencia gént (TME1 b-laktamáz) módosítja. A módosított rezisztens plazmidot tartalmazó sejtek így nemcsak ampicillint tartalmazó táptalajon, hanem „GeneEditor™ Antibiotic Selection Mix”-et tartalmazó táptalajon is fenntarthatók. Első lépésben a plazmidot denaturálják, majd az oligonukleotidok anellálása után T4 DNS-polimeráz enzimmal és T4 DNS-ligáz enzimmal szintetizálják és cirkularizálják a mutációkat tartalmazó új szálát (mutáció a b-laktamázban és a vizsgálandó DNS inszertben). A ligáláshoz itt is szükséges az oligonukleotidok 5' végének előzetes foszforilálása. A mutagenesis terméket kompetens *E. coli* sejtekbe transzformálják, melyeket ampicillin és „GeneEditor™ Antibiotic Selection Mix”-et tartalmazó táptalajon növesztenek. Ezen a táptalajon csak azok a sejtek nőnek fel, melyek a kívánt, mutáns plazmidot tartalmazzák. A mutagenesis hatásfoka növelhető, ha a reakcióterméket *E. coli* BMH 71-18 mutS baktériumtörzsbe transzformálják. Az így elérhető mutációs ráta gyakran nagyobb, mint 90%.



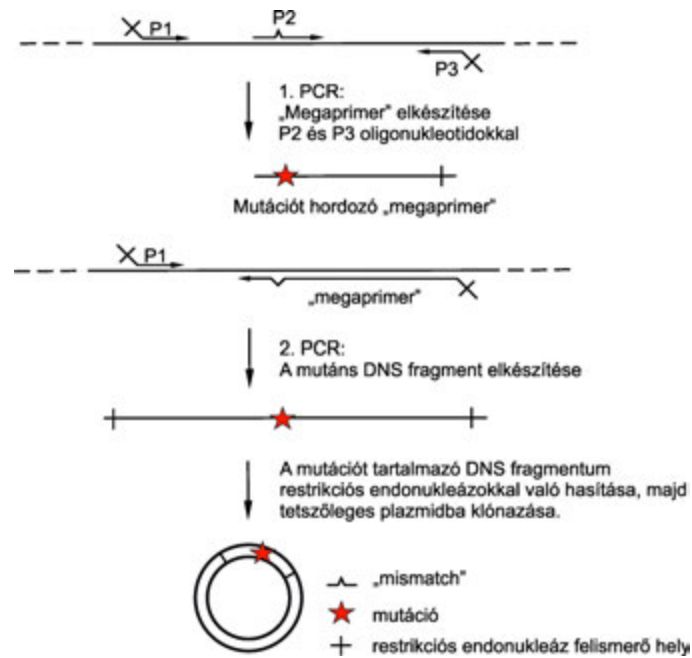
11.5. ábra: A GeneEditor mutagenesis sematikus ábrája. A módszerrel TME1 b-laktamáz alapú ampicillin rezisztenciát hordozó plazmid által kódolt DNS inszert szekvenciáját lehet tetszőlegesen változtatni. A vektor hő- vagy lúgos denaturálása után oligonukleotidok használatával két parallel mutációt történik. Az egyik a vizsgálandó DNS inszerten, a másik a b-laktamázt kódoló génen. A szintézis (T4 DNS-polimeráz) és a cirkularizáció után (T4-ligáz) vad típusú és kétszeresen mutáns szálat kapunk. A terméket *E. coli* mutS- baktériumtörzsbe transzformáljuk, majd „GeneEditor™ Antibiotic Selection Mix”-et tartalmazó táptalajon növesztjük, így csak a megváltoztatott b-laktamáz szekvenciát kódoló plazmidot tartalmazó sejtek szelektálódnak.

11.2.3. PCR alapú irányított mutagenesis módszerek

A **polimeráz láncreakció** egy tetszőleges DNS szekvencia hatékony és egyszerű felszaporítását teszi lehetővé (ld. 6. fejezet). A PCR-re alapozva az utóbbi évtizedekben számos irányított mutagenesis módszert fejlesztettek ki, melyek mindegyike több lépéses PCR-t és további klónozási lépéseket igényel. Mint minden klónozási lépés során, így a PCR alapú irányított mutagenesis módszerek esetében is általánosan ajánlott a „nagy hűségű”, azaz hibajavító aktivitással (*high fidelity, proofreading*) rendelkező DNS-polimerázok használata. Ilyen DNS-polimerázok például a Pfu-, KODI- vagy a Phusion-polimeráz (ld. 3.1.4. fejezet)

11.2.3.1. Megaprimer PCR mutagenesis

A PCR alapú *in vitro* mutagenesis módszerek közül az 1990-es évek elején Sarkar és munkatársai által leírt „megaprimer” PCR módszer talán a leggyorsabb, legegyszerűbb és legsokoldalúbb (ld. 11.6. ábra).

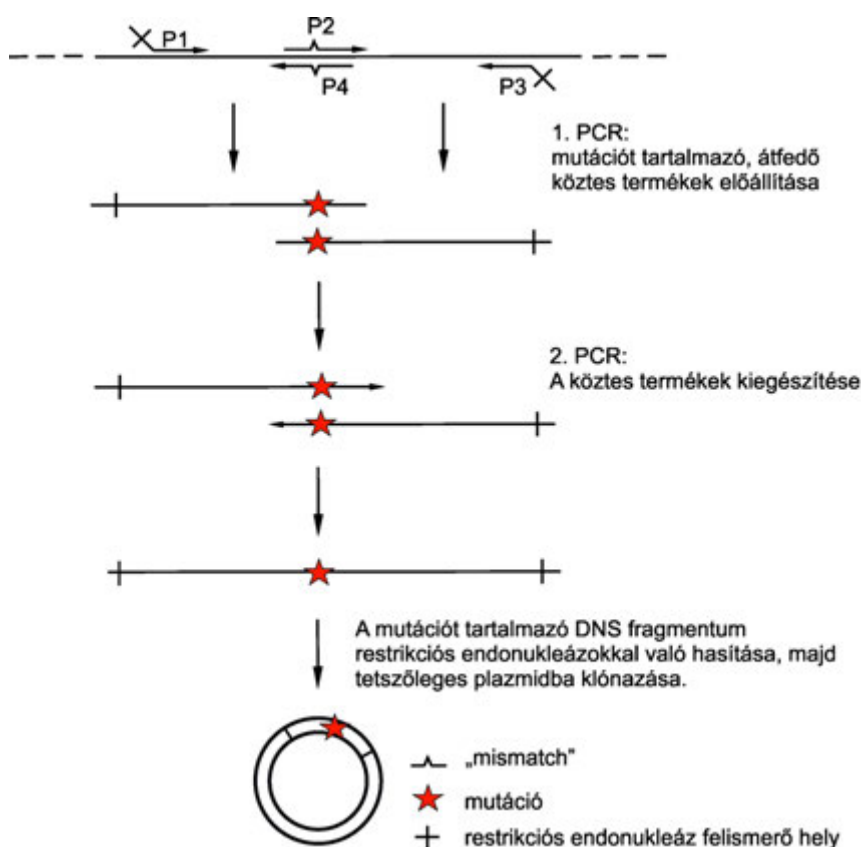


11.6. ábra: „Megaprimer” PCR mutagenézis sematikus ábrája. A módszer során két egymást követő PCR segítségével irányítottan, bárhol létrehozható változás az vizsgált DNS szekvenciájában. Az 1. PCR lépésben a mutációt hordozó P2 és P3 oligonukleotidok használatával létrejövő termék az ún. „megaprimer”. A 2. PCR során a P1 primer és maga a „megaprimer” felhasználásával amplifikálódik a mutációt tartalmazó DNS szakaszunk, melyet a P1 és a P3 oligonukleotidok által kódolt restrikciós enzim felismerőhelyeken való emésztés után szubklónozzhatunk.

A megaprimer módszer során **három primer oligonukleotid** és a vad típusú DNS templát felhasználásával két PCR történik. A három oligonukleotidból **kettő, a határoló vagy „flanking” oligonukleotidok** jelölik ki a felszaporítandó DNS szekvenciát és tartalmazzák a későbbi klónozáshoz szükséges restrikciós enzim felismerőhelyeket. A **harmadik oligonukleotid** tartalmazza magát a **mutációt**, ami lehet inszerció, delécio, szubsztitúció vagy ezek bármilyen keveréke. A mutáció a target génen belül bárhol létrehozható. Az **első PCR** során a mutációt hordozó oligonukleotid és az egyik határoló primer segítségével, valamint a vad típusú templát DNS felhasználásával előállítjuk a **mutációt tartalmazó „megaprimer”**. Az elkészült megaprimer mérete elérheti a több száz, vagy akár ezer bp hosszúságot is. A megaprimer agaróz gélelektroforézissel elválasztjuk az aspecifikus termékektől és a PCR során fel nem használt primerektől, majd gélből izoláljuk. A tisztított **megaprimer**et, mint egy „általános” primert **használjuk fel a másik határoló primerrel a második PCR során** a már mutációt tartalmazó DNS szekvenciánk sokszorosítására. A második PCR során beállított anellálási hőmérsékletnél figyelemmel kell lenni a megaprimer extrém méretére és a másodlagos szerkezet kialakítására való hajlamára. Érdeemes a primer párját úgy tervezni, hogy annak olvadási hőmérséklete 55-60°C környékén legyen. A második PCR során elkészülő mutáns DNS szekvenciának a mennyiségét a reakcióelegybe bemért megaprimer mennyisége fogja meghatározni, ezért ajánlott nagy mennyiségű megaprimer előállítása és a második PCR reakcióba a lehető legnagyobb mennyiségben való hozzáadása. A 2. PCR hozamát nagy mennyiségű templát használatával is növelni lehet. Az elkészült mutáns DNS termék agaróz gélelektroforézissel választható el az esetleg megmaradó megaprimerrel és a templát DNS-t tartalmazó plazmidtól. A tisztított mutáns DNS inszertet a határoló oligonukleotidok által kialakított restrikciós enzim felismerőhelyeken emésztik és a megfelelően előkészített, restrikciós enzimekkel emésztett plazmidba ligálják. A ligátumot kompetens sejtekbe transzformálják, a mutáns DNS inszertet tartalmazó plazmidot szaporítják, majd a szekvenciáját szekvenálással ellenőrzik (a mutációt hordozó inszert plazmidba klónozásának folyamatáról további részletesebb leírás megtalálható a **4. fejezetben**).

11.2.3.2. Overlapping PCR mutagenézis

A PCR alapú irányított mutagenézis módszerek közül az *overlapping* PCR mutagenézis használatával a vizsgálandó fehérje DNS-én belül bárhol létrehozható mutáció. A módszer a megaprimer PCR mutagenézis alternatívájaként használható. Hosszabb időigénye, az egymást követő három PCR és a négy primer miatt – szemben a „megaprimer” PCR során használt 3 primerrel – kevésbé terjedt el (**ld. 11.7. ábra**).



11.7. ábra: Az **Overlapping PCR mutagenesis** sematikus ábrája. A módszer során két egymást követő PCR segítségével irányítottan, bárhol létrehozható változás a vizsgált DNS szekvenciájában. Az 1. PCR lépésben a mutációt kódoló, egymással átfedő köztes termékeket állítunk elő. A 2. PCR lépésben az átfedő DNS szálak kiegészítése történik. Végül egy mutációt tartalmazó, 5' és 3' végén restrikciós enzimfelismerő helyet kódoló inszertet kapunk, amelyet megfelelő plazmidba szubklónozhatunk.

Az első két PCR során két átfedő fragmentum jön létre, melyek tartalmazzák a mutációt és a klónozáshoz szükséges restrikciós enzim felismerőhelyeket. A következő PCR során ezen tisztított termékek anellálásából és lánchosszabbításából kialakuló, mutációt tartalmazó termék szolgál templátként a restrikciós enzim felismerőhelyét tartalmazó, a primer oligonukleotidok számára. A mutációt tartalmazó inszert megfelelő plazmidba klónozása a korábbiakban leírtak alapján történik (ld. 6.4.1.. fejezet).

11.3. További olvasnivaló a fejezethez

Hutchison III, C. A. , Phillips, S., Edgell, M. H., Gillam, S., Jahnke, P. and Smith, M. (1978) Mutagenesis at a specific position in a DNA sequence. *J Biol Chem.* 253(18), 6551-60.

Cadwell, R. C. and Joyce, G. F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl* 2:28–33

Kunkel, T. A , Roberts, J. D., and Zakour, R. A. (1987) Rapid and efficient sitespecific mutagenesis without phenotypic selection. *Methods Enzymol.* 154, 367-382

Sarkar, G. and SomMer, S. S. (1990) The “megaprimer” method of site-directed mutagenesis. *BioTechniques* 8, 404–407.

12. fejezet - *In vitro* és prokarióta expressziós rendszerek

A rekombináns DNS technológia egyik legfontosabb vívmánya, hogy a vizsgálataink tárgyát képező fehérjét nagy mennyiségben, tisztán és viszonylag egyszerűen elő tudjuk állítani. A „tisztán” kifejezés itt nemcsak azt jelenti, hogy a preparátum kb. 90-95%-ban a bennünket érdeklő fehérjénket tartalmazza, hanem azt is, hogy például nem kell 1 mg neuronális fehérje előállításához 25-30 db patkányt megölni, mint a géntechnológiát megelőző korszakban. Az expressziós rendszerek felhasználásával a vizsgálandó fehérjét nemcsak egyszerűbben és nagyobb mennyiségben, hanem a céljainknak megfelelően módosítva - irányított mutációkkal, önálló doménként, csonkolt formában, különböző címkékkel (*tag*-ekkel) ellátva - is tudjuk termeltetni.

Nem mindegy, hogy milyen expressziós rendszert használunk. Ennek kiválasztásában jelentős szempont a szükséges fehérje mennyisége, tisztasága, poszttranszlációs módosítása, aktivitása, a ráfordított idő és pénz, illetve milyen infrastruktúra (és mennyi kutatási forrás) áll mögöttünk. Ebben a fejezetben a **sejtmentes (*in vitro* expressziós rendszer)** és a **prokariótaexpressziós rendszereket** ismertetjük. Megbeszéljük az alapjaikat, milyen vektorokat, gazdasejteket használhatunk, mik az előnyeik és hátrányaik, milyen célra melyik rendszert célszerű használni, valamint milyen alkalmazási lehetőségeik vannak. Az **eukarióta expressziós rendszereket** a következő fejezetben tárgyaljuk (**ld. 13. fejezet**)

Kezdjük az alapfeltételekkel, azaz mi az expressziós rendszerek két pillére? A **gazdasejt**, ami termeli a **rekombináns fehérjét**, illetve az **expressziós konstrukció**, azaz a rekombináns DNS, ami meghatározza a termelendő polipeptidlánc(ok) aminosav sorrendjét. A **rekombináns fehérje kifejezés annyit jelent, hogy rekombináns DNS segítségével expresszáltatunk egy fehérjét**, ami az esetek többségében megegyezik a natív fehérjével (azaz nem két polipeptidlánc „rekombinálásával” állítjuk elő). A fenti sémától eltér a sejtmentes, azaz *in vitro* expressziós rendszer, mivel ebben az esetben nincs szükségünk gazdasejtre, csak a pro- vagy eukarióta sejtől származó riboszómákra, valamint az *in vitro* transzkripcióhoz és transzlációhoz szükséges egyéb komponensekre.

A gazdasejt lehet baktérium (*Escherichia coli*, *Bacillus subtilis*), élesztő (*Saccharomyces cerevisiae*, *Pichia pastoris*) vagy más eukarióta sejtvonala vagy akár transzgenikus állat és növény is. Ezek a gazdasejtek **heterológ expresszióra** képesek, vagyis arra, hogy nem a saját, hanem más fajból származó fehérjét termeljenek. Mindegyik rendszernek vannak előnyei, és hátrányai, és megvannak a saját expressziós vektorai. Ezek az **expressziós vektorok** tartalmazzák azokat a szekvenciákat, amelyek a gazdasejt felismer, és elindítja a kifejezendő gén transzkripcióját, majd transzlációját. A fejezetben támaszkodunk az 7. fejezet ismeretanyagára.

12.1. *In vitro* expressziós rendszerek

Az *in vitro* vagy **sejtmentes expresszió** azon alapszik, hogy *in vitro* körülmények között a sejtől kivont transzlációs rendszer segítségével szintetizáltatjuk a kívánt fehérjét. (A sejtmembránt, illetve eukarióta sejteknél az organellumokat is eltávolítva, a citoplazma és a benne található komponensek alkotják a „sejtmentes” rendszert). A rendszer működéséhez szükség van a genetikai információra, azaz a kívánt fehérjét kódoló mRNS-re, illetve általában egy expressziós vektorban lévő rekombináns DNS-re, valamint a reakció oldatra, mely tartalmazza a szükséges transzkripció és transzláció aktivitást. A legelső sejtmentes transzlációs rendszer **prokarióta kivonat** volt, ezt követték az **eukarióta sejtmentes transzlációs rendszerek**, rovarsejt (Sf9, Sf21-sejt), nyúl retikulocita (RRL), búzacsíra, és ma már humán sejtvonalból származó transzlációs rendszer is elérhető. További információ a kereskedelmi forgalomban elérhető rendszerekről itt olvasható.

12.1.1. Az *in vitro* transzlációs rendszer komponensei

Mi szükséges a sikeres *in vitro* transzlációhoz?

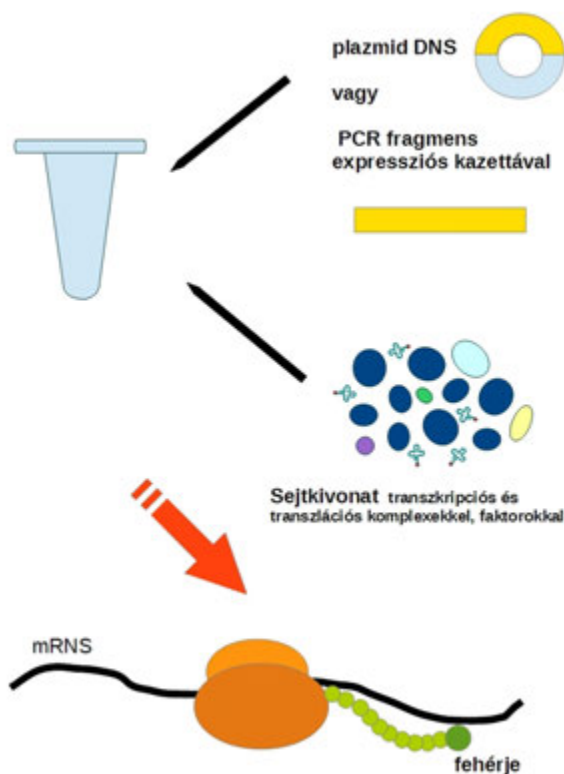
- RNS-polimeráz és NTP-k az mRNS transzkripciójához
- riboszómák a polipeptidlánc transzlációjához
- tRNS-ek és aminosavak

- transzlációs kofaktorok
- ATP, GTP és energiaregeneráló rendszer
- dajkafehérjék és más faktorok a polipeptidlánc feltekeredéséhez
- templát nukleinsav

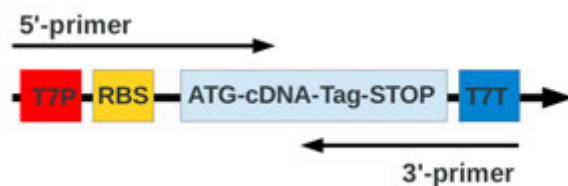
A fehérje-komplexeket és faktorokat a kereskedelmi forgalomban kapható sejtmentes kivonatok tartalmazzák, csak **energiaforrást (ATP, GTP, energiaregeneráló rendszert)** és **aminosavakat** kell hozzáadni, valamint értelemszerűen a **templát nukleinsavat**.

A **templát nukleinsav** mRNS vagy DNS lehet. Amennyiben a templát DNS, azt először át kell írni mRNS-sé, és majd az azt követő transzlációs reakció során lesz belőle fehérje. Ezért ezt a rendszert kapcsolt (*linked* vagy *coupled*) rendszernek hívják (ld. 12.1. ábra).

A kapcsolt rendszer esetén ugyanabban a csőben játszódik le egymás után a transzkripció és transzláció. DNS templát esetén a bemenő DNS lehet lineáris, cirkuláris, vagy PCR termék. A DNS-nek tartalmaznia kell a transzkripció és transzláció elindításáért felelős elemeket, tehát a **promóter** szekvenciát, és prokarióta rendszernél a **riboszóma kötőhelyet (RBS)**, azaz a Shine-Dalgarno szekvenciát. Eukarióta rendszernél ennél bonyolultabb a helyzet, mert a splicing folyamatát (DNS templát esetén), továbbá az mRNS érését és transzlációjának iniciálását is nekünk kell megoldani. Ezért a megfelelő génexpressziós elemeket (pl. Kozak-szekvencia, IRES) is tartalmaznia kell a templátnak, illetve intron mentes DNS-t kell bejuttatni a rendszerbe. Ennek megfelelően speciális prokarióta vagy eukarióta **expressziós vektorokba** kell a target gén cDNS-t bevinni. Ezek az expressziós vektorok **expressziós kazettájukban** már eleve tartalmazzák a génexpresszióhoz szükséges elemeket (ld. 12.2. ábra).



12.1. ábra: A kapcsolt (*linked* vagy *coupled*) *in vitro* transzlációs rendszer sémája



12.2. ábra: **Prokarióta expressziós kazetta sematikus ábrája.** T7P: T7-promóter szekvenciája, RBS: riboszóma kötőhely szekvenciája, ATG: az első metionin aminosav kodonja, Tag: fúziós címke (a példában C-terminális) szekvenciája, T7T: T7 transzkripció terminációs szekvencia

Ilyen vektorok, illetve *in vitro* translációs kit-ek megvásárolhatók, pl. a NEB (New England Biolabs), Invitrogen vagy a ThermoScientific biotechnológiai cégektől. Mivel lineáris DNS-t is használhatunk templatként, ezért PCR reakcióval mi magunk is elkészíthetjük a templatunkat. Ilyenkor olyan primerekkel kell felszaporítani a cDNS-t, amely tartalmazza a megfelelő szabályozó elemeket. A 12.2. ábra alapján látható, hogy az 5' primernek tartalmaznia kell a promótert az RBS-sel együtt, a 3'-primernek pedig a terminációs szekvenciát. Azaz a használandó primerek sematikus felépítése:

- 5'-primer: T7P szekvencia-RBS szekvencia-cDNS-re illeszkedő szekvencia.
- 3'-primer: cDNS-re illeszkedő szekvencia-T7T szekvencia (természetesen ennek a reverz komplementere kell).

A Roche cég fejlesztése, a **Rapid Translation System (RTS 500)**, *E. coli* extraktumot használ, amely akár egy napig is képes folyamatos szintézisre, így nagyobb mennyiségű fehérjét tud előállítani. Ezt úgy tudták megoldani, hogy gondoskodtak az energia hosszú távú regenerálásáról, illetve a kofaktorok, alapanyagok (nukleotidok, aminosavak) folyamatos betáplálásáról és a gátlóanyagok elvezetéséről. Ezt úgy oldották meg, hogy féligáteresztő hártyával elválasztják a reakciótérrel a keletkező gátló anyagokat, ami a konvencionális módszer egyik legnagyobb problémája. A rendszerről további információ itt olvasható.

12.1.2. Előnyök és hátrányok

Előnyök:

- A leggyorsabb expressziós rendszer, néhány órát vesz csak igénybe.
- Egyszerű, nem igényel sejtenyésztést, transzformálást, intenzív fehérjetisztítást.
- Független a gazdasejttől (nincs gazdasejt).
- Toxikus fehérjékre is használható (nincs sejt, ami elpusztulna).
- Proteolízisre nagyon érzékeny fehérjék expressziója is megoldható (bár a sejtmentes kivonathoz is szüksége lehet proteáz enzim gátlószerek hozzáadására).

Hátrányok:

- Kevés fehérje állítható elő, általában egy reakcióval 0,1 µg/ml fehérjét nyerhetünk ki. Gyakran ki sem tisztítjuk a fehérjét, hanem a fehérjét tartalmazó reakcióoldattal dolgozunk tovább.
- Nem megfelelő a poszttranszlációs módosítás (glikoziláció, foszforiláció), bár ez a probléma megfelelő *in vitro* translációs rendszer megválasztásával megoldható: pl. humán fehérjét humán sejtvonalból nyert sejt extraktummal szintetizálthatunk.
- A sejtkivonat endogén mRNS-étől meg kell szabadulni (kezelés *micrococcus* nukleázokkal).

12.1.3. Alkalmazások

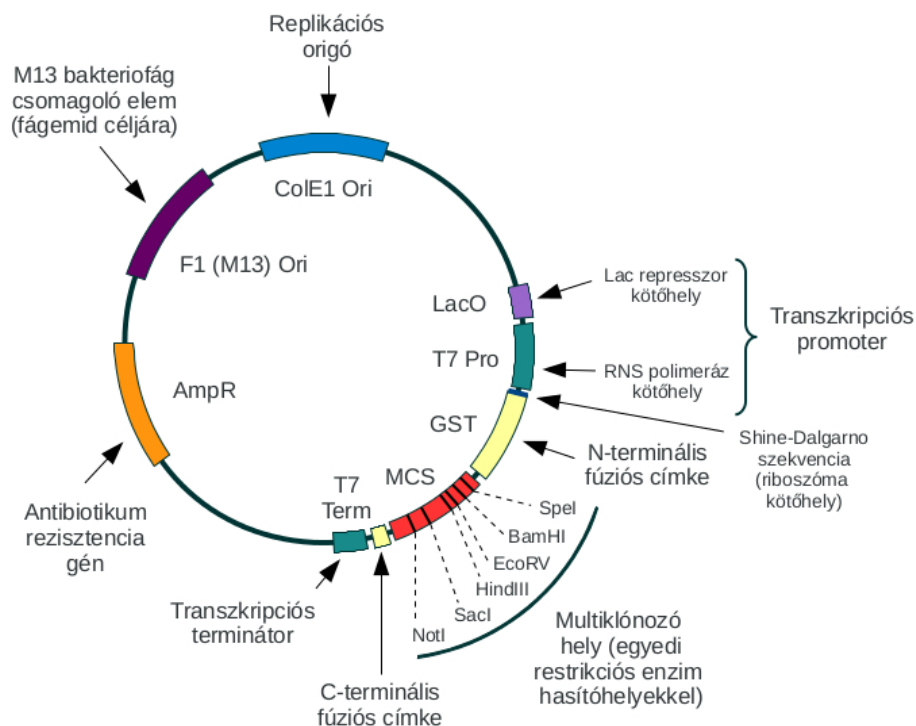
- Gyors és nagyszámú mutáns vagy csonkolt fehérjék előállítására funkcionális analízis céljából (nagy áteresztőképességű, ún. *high throughput*, HTP módszer)
- Megfelelő poszttranszlációs modifikációval rendelkező humán fehérjék előállítására.
- Fehérjék stabil izotópos jelölésére, nem-természetes aminosavak bevitelére, mivel a reakcióhoz szabadon hozzáadhatók a jelölt vagy nem természetes aminosavak.
- Funkcionális virionok (vírusfehérjék és nukleinsavak, amelyek spontán virionná állnak szerveződnek), illetve toxikus polipeptidek előállítására.
- Adott fehérje feltekeredésének (folding), stabilitásának és degradációjának tanulmányozására és optimalizálására.

12.2. Prokarióta expressziós rendszerek

Nézzük meg a **prokarióta expressziós vektorok** általános felépítését, amelynek alapelemei a riboszóma kötőhely kivételével az eukarióta expressziós vektorokban is megtalálhatóak (**ld. 12.3. ábra**).

- **Replikon**, azaz a replikációs origó és szabályozó szekvenciák, amelyek a plazmid baktérium sejtben történő fenntartásához, replikációjához szükségesek.
- Antibiotikum rezisztencia gén a szelekcióhoz.
- **Promóter és szabályozó elemei** a kívánt cDNS transzkripciójához.
- **Riboszóma kötőhely** (Shine-Dalgarno szekvencia) a kívánt cDNS-ről átírt mRNA transzlációjának iniciálásához.
- **MCS**: poli- vagy multiklónozó hely a kívánt cDNS beviteléhez.

Korábbi fejezetekben már ismertettük a vektorok és ezen belül plazmidok általános felépítését (**ld. 7. fejezet**), ezért ebben a fejezetben csak a génexpresszió szempontjából legérdekesebb elemekről, a promóterekről, és a riboszóma kötőhelyről (RBS) lesz szó.



12.3. ábra: Egy tipikus prokarióta expressziós vektor felépítése

12.2.1. Bakteriális promóterek és a riboszóma kötőhely

12.2.1.1. Promóterek

A bakteriális expressziós vektorok általában plazmid-alapúak, amelyek legtöbbször indukálható promótert tartalmaznak. Már korábban megtanultuk, hogy a prokarióta promóternek két fontos eleme van: a **Pribnow-box** vagy **-10 elem** (TAAATA konszenzus szekvencia), ami 10 bp-ra található 5'-irányban az első átírandó bp-hoz képest (amely a +1 pozíciót jelenti). Ezen elem jelenléte esszenciális a transzkripcióhoz. A másik elem egy TTGACA konszenzus szekvencia, amit **-35 szekvenciának** is hívunk, mert a start ponttól 35 bp-ra található *upstream*, azaz 5'-irányban az átírandó szekvenciától. Jelenléte a promóterben nem esszenciális, de jelentős mértékben hozzájárul a gén átírásához. Végül sok prokarióta promóterben a -35 elemtől 5'-irányban további nem esszenciális szekvenciák is találhatóak („UP” elemek). Egy tipikus bakteriális promóter fő elemeit a **12.4. ábra** mutatja be.

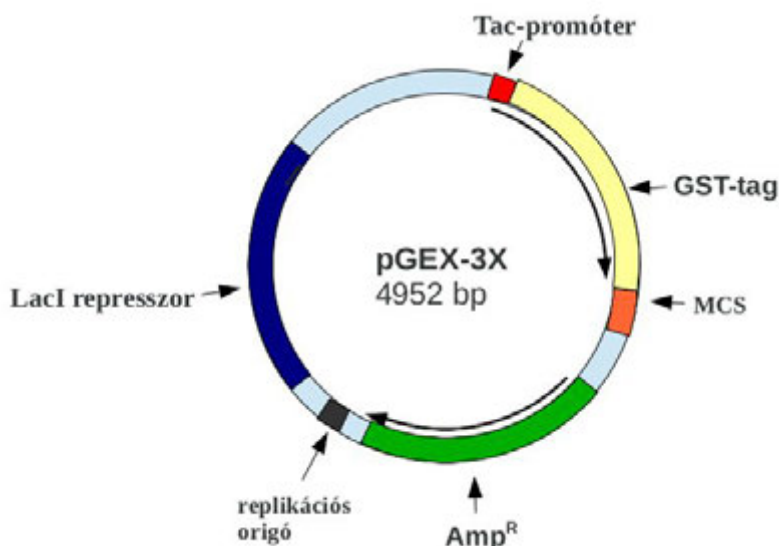


12.4. ábra: Egy bakteriális promóter sémája

A bakteriális promóterek lehetnek állandó kifejeződést biztosító, azaz **konstitutív**, vagy **indukálható** promóterek. Az indukció történhet hőmérséklettel, vagy valamilyen kémiai ágenssel, induktorról. Az indukció leggyakrabban IPTG-vel (izopropil-tiogalaktozid) történik, ami egy nem metabolizálódó laktóz analóg. Így az indukció a lac-operon szabályozási elvén alapul, azaz az expressziós vektor a promóter mellett tartalmazza a lac operátor szekvenciát is (a -10 elem és az átírandó szekvencia között). Az induktor (allolaktóz vagy IPTG) hiányában a lac-represszor fehérje kötődik ehhez a szekvenciához, így nem hozzáférhető az RNS-polimeráz számára a promóter, tehát nincs mRNS átírás. Allolaktóz vagy IPTG jelenlétében a lac-represszor leválik a promóterről, megtörténik a génátírás. Ráadásul, az IPTG mennyiségével még az átírás mértékét is szabályozni tudjuk.

Lac promóter (lavUV5 promóter): Általánosan elterjedt, ilyen van a pUC katabolit, pTZ, pSK, pBluescript, pGEM vektorokban. Bár a promóter viszonylag gyenge, mégis megfelelő mértékű fehérje kifejeződés érhető el, mivel a plazmid nagy kópiaszámú. Akkor működik a promóter maximálisan, ha a baktérium kultúra médiuma nem tartalmaz glükózt, ugyanis ekkor működik a cAMP/aktivátor fehérje (CAP), ami szükséges a promóter működéséhez.

Trp-lac (tac) promóter: hibrid promóter, -35 régió tartalmazza a trp (trp operon) -35 szekvenciáját, majd -10-nél a lavUV5 promóter (*lac-operon*) szekvenciáját és a lac-operátor szekvenciáját. Így a promóter regulálható a lac-represszorral, azaz a génátírás indukálható IPTG-vel, és független a cAMP regulációtól. Erős promóter, nagymértékű expresszió érhető el vele. Ilyen promóter található a pGEX vektorokban, amelyek sematikus felépítését a **12.5. ábra** mutatja be.



12.5. ábra: A pGEX-3X expressziós vektor térképe

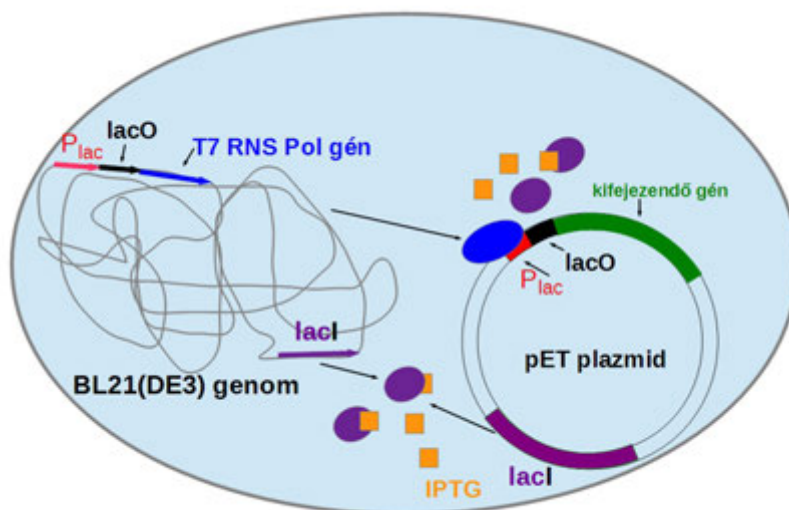
A pGEX vektorok replikonja (pMB1) a pBR322 vektorból származik, ezért ezek alacsony kópiaszámú vektorok. Általában ampicillin rezisztenciájuk van, amit a b-laktamáz enzim, a rezisztencia gén terméke biztosít. A *trp* promóterrel való hibridizáció miatt sokkal erősebb indukálható expressziót biztosít, mint a *lac* promóter, 3-szor erősebb a *trp*, és 10-szer a *lac*-promóternél. Ez egy GST-fúziós fehérje termeltetésére alkalmas vektor (**ld.12.2.3.1. fejezet**)

Trp-lac (trc) promóter: szintén hibrid promóter, a tac promóter egyik változata. A különbség a két promóter között a -35 és -10 elemeket elválasztó távolságban van, a trc-promóternél rövidebb. Ilyen promóter található a pTrc vektorokban.

T7 promóter: A legerősebb promóterek közé tartozik. A T7 bakteriofág specifikus promótere, 5-ször erősebb, mint az *E. coli* baktérium konszenzus promótere, ezért nagymértékű expressziót biztosít. Speciális baktériumtörzs kell hozzá, amely képes felismerni a T7 promótert, azaz termeli a T7 RNS-polimeráz enzimet (az *E. coli* RNS-polimeráz enzim csak a saját génei promótereire kötődik, miként a T7 bakteriofág RNS-polimeráz csak a fág gének promótereire kötődik). Ezért az *E. coli* BL21 baktériumtörzs genomjába fág-fertőzéssel bejuttatták a T7 fág RNS-polimeráz génjét úgy, hogy a lacUV5 promóter szabályozása alatt legyen. Az így kapott baktériumtörzset hívjuk BL21(DE3)-nak. A DE3 profág egy λ -fág származék, amelynek - a kromoszomális integrációért felelős -*int* génjébe juttatták be a T7 RNS-polimeráz és a lacI represszor fehérje géneket. Ebben a törzsben a laktóz analóg IPTG hatására beindul a T7 RNS-polimeráz termelése, ami felismeri a plazmidon lévő T7-promótert, és elkezd átírni a promóter után lévő inszertet.

Ilyen promóter található a **pET vektorokban** (plasmid for Expression of **T**7 RNA polymerase) (eredetileg a Novagen biotechnológiai cég terméke), amelyet Studier és társai fejlesztettek ki a '90-es évek elején. Ezek a vektorok colE1 origót (replikon) tartalmaznak, ami alacsony kópiaszámot biztosít, és általában ampicillin vagy kanamicin rezisztencia gént hordoznak. A promóterük vagy a T7 fág természetes promótere, a **ϕ 10 promóter**, vagy a természetes ϕ 10 promóter a lac operátorral (*LacO*) együtt. Az így kapott **T7lac promóter** már indukálható

IPTG-vel, és nagyobb biztonságot jelent, mert mind a polimeráz, mind a termelendő fehérje génje a lac operátor szabályozása alatt áll, és IPTG-vel indukálható. Tehát indukció nélkül elméletileg nincs génátírás. De a **BL21 pLysS** és **pLysE** baktériumtörzsekben még egy extra biztosíték is van; ezek a törzsek külön plazmidon (pLys) hordozzák a T7 RNS-polimeráz természetes inhibitorának, a lizozimnek a génjét. IPTG nélkül, az esetleg keletkező RNS-polimerázt a lizozim gátolja, így biztos nincs átírás. Ezt akkor érdemes kihasználni, ha toxikus a fehérje a baktérium számára, így csak az indukció hatására kezd a fehérje expresszálni. Nézzük meg a **12.6. ábrán** az igen gyakran használt pET-expressziós rendszer működését. Itt találhatók további elméleti és gyakorlati részletek a legmegfelelőbb vektorok kiválasztásáról.



12.6. ábra: A pET-expressziós rendszer sémája

λpL promóter: Magas hőmérséklettel indukálható, mert hőmérséklet érzékeny represszort tartalmaz. A λ-fág represszor fehérje alacsony hőmérsékleten a promóterhez kötődik, míg magasabb hőmérsékleten leválik, így elkezdődhet a génátírás. A baktérium kultúrát 30°C-on növesztjük, majd az indukció 42°C-on történik. Ez a promóter különösen hasznos toxikus fehérjék kifejeztetésénél. Ilyen vektorok a pHUB, pPLc, pKC30, pAS1, pRM1, pTrxFus sorozatok. De ehhez a promóterhez is speciális *E. coli* törzsre van szükség: az M5219 törzs a λ-represszor gén hőmérséklet érzékeny mutációját hordozza.

araBAD promóter: Az arabinóz operon operátor helyét tartalmazza, így arabinózzal indukálható, dózis-függő módon. A promóter, mivel arabinóz hasznosítású operon része, glükózzal gátolható. Ilyen promóterrel a pBAD sorozat (Invitrogen) vektorai működnek.

lpp promóter: Az expressziós vektor a kifejezendő gén előtt egy külső membrán lipoprotein promóterét tartalmazza. Ez a fehérje nagy mennyiségben és állandóan kifejeződik a baktérium sejtben, így a promóter erős és konstitutív lesz.

phoA promóter: A bakteriális alkalikus foszfatáz gén (phoA) promóterét és N-terminális leader szekvenciáját tartalmazza. Így a fehérjénk olyan N-terminális szekvenciával fog kifejeződni, ami a periplazmatikus térbe küldi. A promótert lehet indukálni foszfát megvonással, vagy akár éheztetéssel. Ilyen promóter található a pTA vagy pBace vektorokban.

rnnB promóter: Magyar vonatkozású promóter; Venetiánér Pál akadémikus és munkatársai (MTA SZBK, Szeged) fejlesztették ki. Azon alapul, hogy a transzkripció mértékéhez nagyban hozzájárulnak a promóterhez képest *upstream* irányban lévő szekvenciák is. Ezért az 5S rRNS gén promóteréhez, ami magában is egy nagyon erős promóter, még egy extra 5'-szekvenciát adtak -40-60 régióban. Ez az *upstream* szekvencia (UP promóter modul) az RNS-polimeráz a-alegység C-terminális részének kötőhelyét tartalmazza, így értelemszerűen megnöveli a génátírás hatékonyságát (kb. 30-szorosára).

12.2.1.2. Riboszóma kötőhely (RBS)

Az RBS vagy Shine-Dalgarno szekvencia (amely az átírt mRNS 5'-UTR, azaz nem transzlálódó régiójában található) is nagyon fontos a gén kifejeződéséhez. Ez az 5-9 nukleotidból álló szekvencia komplementer a 16S rRNS 3' régiójával, így segíti a riboszóma kötést és a transzláció iniciálását. Az RBS és a transzlációt iniciáló ATG (start) kodon távolsága 3-11 nukleotid, s szintén jelentősen befolyásolja az expresszió mértékét (azaz a távolság változtatásával is optimalizálni lehet a fehérjetermelés szintjét – ld. később). Az expresszió hatékonyságát a RBS-től még távolabb lévő 5'-nem kódoló régiók is befolyásolják és a gén 3'-végén található **transzkripcióterminációs szekvenciák** is növelhetik (bár nem esszenciálisak a legtöbb rekombináns fehérje előállításához).

A kereskedelembe kapható expressziós vektorok tartalmazzák a megfelelő szabályozó elemeket. Nekünk csak arra kell ügyelnünk, hogy az expresszálandó génünket a megfelelő helyre klónozzuk be; a promóter és a riboszóma kötőhely után, ne vágjuk ki a riboszóma-kötőhelyet, és a leolvasási keret se tolódjon el. A START kodont tartalmazhatja a vektor szekvencia (pl. az NdeI restrikciós enzim felismerőhelyében van egy START kodon: ATGCAT), de a cDNS-sel is bevihetjük a vektorba, csak az RBS-hez való távolságára kell ügyelni. A **promóter és ariboszóma-kötőhely** létfontosságúak a gén expressziójához, a szabályozó elem (operátor régió) nélkül ugyan működhet az expressziós rendszer, de mégis ajánlott erre is gondolkodni, főleg az expresszió optimalizálásához (ld. alább). A kiválasztott expressziós rendszer vektorai tartalmazzák ezeket az elemeket. Azt kell még meggondolnunk, hogy a kívánt fehérje utolsó aminosava után szeretnénk-e befejezni a fehérje szintézist, vagy fúziós címkét akarunk a fehérje C-terminális részére. Első esetben **STOP** kodont kell tennünk a cDNS végére (természetesen használhatjuk a gén saját STOP szekvenciáját is), a második esetben ki kell ezt hagyni. Ezután csak az MCS (multiklónozó hely) megfelelő helyére kell bevinnünk az expresszálandó gén cDNS-ét. Az így elkészített vektort be kell juttatnunk a megfelelő **gazdasejtbe** (ld. 4.3. fejezet).

12.2.2. A prokarióta fehérje expresszió optimalizálása

Feltételezzük, hogy sikerült kiválasztanunk a megfelelő prokarióta expressziós vektort és a vele kompatibilis baktérium törzset. Az inszertünk (a kívánt fehérje cDNS-e) megfelelő helyre bekerült az expressziós vektorba. Sikerült a plazmid transzformálása az expressziós törzsbe, a transzformált baktériumok jól osztódnak. Az indukciót is jól elvégeztük, de vagy nincs termékünk, vagy nem megfelelő a mérete, vagy nem elegendő a mennyisége. Mi lehet a gond, mi a teendő?

1. Nézzük újra át a vektor térképét, nem klónozó vektort használunk-e, ami nem tartalmaz riboszóma kötőhelyet! Megfelelő leolvasási keretben (*frame*-ben) van-e az inszert. Ez különösen az N-terminális fúziós címkét tartalmazó konstrukcióknál fontos (ld. 12.2.3.). A hárommal nem osztható számú hasítóhelyekkel óvatosan kell bánni (ilyen pl. a NotI), mivel ezek könnyen eltolhatják a leolvasási keretet.
2. Szekvenáljuk a plazmidban lévő inszertet! Nincs-e benne deléció, inszerció, ami *frame* eltolást vagy STOP-kodon beépülést okoz. A plazmid promótere és a transzkripció terminációs szekvencia kiváló primerül szolgál a szekvenáláshoz. Ez a pET-plazmidoknál maga a T7-promóter (*T7forward*), vagy a T7Term (*T7Term-reverse*).
3. Kövessük nyomon a fehérjénk sorsát! Vegyünk mintát minden lépésnél, futassuk meg a mintákat SDS-poliakrilamid gélelektroforézissel (SDS-PAGE) és festéssel, vagy Western-blottal ellenőrizzük a kapott terméket. Legyen mintánk az indukálatlan sejtekből, majd az indukció után is a sejtekből, a sejt feltárása és centrifugálása után a felülúszóból és a pelletből (csapadék, ami a centrifugácsolás alatt összegyűlt), majd az összes tisztítási lépés után is. (A fehérjék tisztítását ld. 12.2.4. fejezet.)
4. **Van fehérjénk**, csak sajnos **nem az oldható frakcióban**, ami a sejt lizátum felülúszója, hanem a csapadékban. Ez azt jelenti, hogy a fehérjénk oldhatatlan formában van a zárvány testben (más néven inklúziós test). Ekkor két lehetőség áll előttünk: az inklúziós testből preparáljuk tovább a rekombináns fehérjét minden hátrányával és előnyével együtt, vagy megpróbáljuk eltolni a fehérje termelést az oldható forma irányába.
5. Nagyon **kevés** fehérjénk van. Miért nem expresszálódik a fehérje?
 - Próbáljuk optimalizálni az indukciós körülményeket; hőmérséklet, idő, indukciós ágens koncentrációja.
 - Korai terminálódás, mert ritka kodonok vannak az inszertben közel egymáshoz (olyanok, amelyekre nincs elég antikodon hordozó tRNS a gazdasejtben) - pl. humán fehérjét szeretnénk expresszáltatni prokariótában,

ahol más bizonyos aminosavak kodon preferenciája. Használjunk olyan baktérium törzset, amely ritka kodonokat is használ (pl. **Rosetta**, a Novagen cégtől, vagy **CodonPlus**, a Stratagene cégtől). Emlékezzünk, fajonként más és más a kódszótár **kodon használati gyakorisága**! Az *E. coli* és a humán kódszótárak kodon gyakoriságát a következő táblázatban hasonlítjuk össze. Sárgával emeltük ki az *E. coli* által használt ritka kodonokat (**ld. 12.1. táblázat**)

12.1. táblázat: Az *E. coli* és a *Homo sapiens* kódszótár összehasonlítása

<i>kodon</i>	<i>E. coli</i>	<i>humán</i>	<i>kodon</i>	<i>E. coli</i>	<i>humán</i>	<i>kodon</i>	<i>E. coli</i>	<i>humán</i>	<i>kodon</i>	<i>E. coli</i>	<i>humán</i>				
UUU	F	0,57	0,46	UCU	S	0,11	0,19	UAU	Y	0,53	0,44	UGU	C	0,42	0,46
UUC	F	0,43	0,54	UCC	S	0,11	0,22	UAC	Y	0,47	0,56	UGC	C	0,58	0,54
UUA	L	0,15	0,08	UCA	S	0,15	0,15	UAA	STOP	0,64	0,30	UGA	STOP	0,36	0,47
UUG	L	0,12	0,13	UCA	S	0,16	0,05	UAG	STOP	0,00	0,24	UGG	W	1	1
CUU	L	0,12	0,13	CCU	P	0,17	0,29	CAU	H	0,55	0,42	CGU	R	0,36	0,08
CUC	L	0,1	0,2	CCC	P	0,13	0,32	CAC	H	0,45	0,58	CGC	R	0,44	0,18
CUA	L	0,05	0,07	CCA	P	0,14	0,18	CAA	Q	0,3	0,27	CGA	R	0,07	0,11
CUG	L	0,46	0,4	CCA	P	0,55	0,11	CAG	Q	0,7	0,73	CGG	R	0,07	0,2
AUU	I	0,58	0,36	ACU	T	0,16	0,25	AAU	N	0,77	0,47	AGU	S	0,14	0,15
AUC	I	0,35	0,47	ACC	T	0,47	0,36	AAC	N	0,53	0,53	AGC	S	0,33	0,24
AUA	I	0,07	0,17	ACA	T	0,13	0,28	AAA	K	0,73	0,43	AGA	R	0,02	0,21
AUG	M	1	1	ACG	T	0,24	0,11	AAG	K	0,27	0,57	AGG	R	0,03	0,21
GUU	V	0,25	0,18	GUU	A	0,11	0,27	GUU	D	0,65	0,46	GGU	G	0,29	0,16
GUC	V	0,18	0,24	GUC	A	0,31	0,40	GUC	D	0,35	0,54	GGC	G	0,46	0,34
GUA	V	0,17	0,12	GUA	A	0,21	0,23	GUA	E	0,70	0,42	GGA	G	0,13	0,25
GUG	V	0,40	0,46	GUG	A	0,38	0,11	GUG	E	0,30	0,58	GGG	G	0,12	0,25

- Rövid az mRNS élettideje, használjunk olyan törzset, amelyben bizonyos RN-ázok hiányoznak. (pl. BL21Star)
- Nem elég stabil a fehérjénk és degradálódik a tisztítás során. Használjunk proteáz gátlószereket, táruk fel a sejteket óvatosabban, és tartsuk minden tisztítási lépésnél a preparátumot jégen, vagy 4°C-on.
- Fejeztessük ki a fehérjénket expressziót és oldhatóságot segítő N-terminális címkével, ilyen pl. a GST illetve MBP címke (**ld. 12.2.3.**).
- Gondoljuk meg, hátha elég a céljainknak a fehérje kisebb része, N- vagy C-terminális csontolt része, fragmentuma. Expresszáltassuk a feltehetően rendezetlen N- vagy C-terminális részeket kódoló régió nélküli fehérjét. A fragmentumot kódoló templát DNS-t szubklónozással, *in vitro* mutagenézissel, PCR módszerrel állíthatjuk elő.
- Toxikus lehet a fehérjénk a gazdasejt számára, ezért használjunk olyan rendszert, amelyben nagyon szigorú szabályozás alatt van a promóter. Ilyenek a pLysS és pLysE plazmidot tartalmazó törzsek.
- Próbáljunk ki más promóterrel rendelkező expressziós vektorokat, más baktérium törzseket.
- Használjunk másik táptalajt a baktérium kultúra növesztéséhez. Általában LB táptalajban fejeztetjük ki a fehérjét, de válthatunk YT, 2YT, Terrific Broth, NzCYM, Nzamin tápra is. Változtathatjuk a táptalaj NaCl

koncentrációját is: 0,2-0,6 M között. Adjunk a növesztéshez glükózt (vigyázat, mert lac-operátort tartalmazó vektoroknál ekkor meg kell növelni az IPTG koncentrációját!), vagy más cukrot

- Fejeztessünk ki bakteriális dajkafehérjéket (chaperonokat) a fehérjénkkal együtt a hatékonyabb folding érdekében. Például a DnaK-DnaJ chaperonokat vagy a GroEL-GroES chaperoninokat hordozó expressziós vektorokkal együtt expresszáltassuk a fehérjénket (**ld. 12.2.2.4.**). Elképzelhető, hogy a fehérjénket csak a kölcsönható partnerével együtt lehet expresszáltatni – ezt is érdemes kipróbálni.
- Váltunk expressziós rendszert, pl. eukarióta expressziós rendszerre térjünk át (**ld. 13. fejezet**).

Ne felejtjük el, hogy a rekombináns fehérjék expressziója, bár vannak általános elvek, mégis a tapasztalaton alapul. Kísérletezni kell, hogy melyik körülmény, forma a legoptimálisabb az adott fehérje és az adott konstrukció előállításához.

12.2.2.1. Az indukciós körülmények optimalizálása

Az indukciós körülmények optimalizálása az **expressziós idő**, a **hőmérséklet** és az **induktor koncentrációjának** a változtatását jelenti, miközben monitorozzuk az expresszáldott fehérje mennyiségét mind a sejtlizátum oldhatatlan (centrifugálás utáni csapadék), mind az oldható frakciójában (felülúszó). Az expresszált fehérje mennyiségét nyomon követhetjük SDS-PAGE-val, vagy Western blotlalt. A Western blot előnye, hogy a specifikus ellenanyag megmondja, melyik „csík” felel meg a gélen a fehérjénknek. Ez akkor különösen nagy segítség, ha nem vagyunk biztosak a fehérje elektroforetikus mobilitásában, vagy kevés a mennyisége. Használhatunk a fehérjénkre, vagy a fúziós címkére specifikus ellenanyagot (anti-His, anti-GST, anti-MBP, anti-Flag, anti-HA). Ez a kísérlet arra a kérdésre is választ ad, hogy termelődik-e (nem történt-e korai termináció) és mennyire stabil (degradálódik-e) a fehérjénk.

Hogyan függ a vizsgált három paramétertől az expresszió mennyisége, sebessége? Minél magasabb a hőmérséklet, és nagyobb az induktor koncentrációja, annál gyorsabban fejeződik ki a fehérje. Ez olyan gyors lehet, hogy a fehérje megfelelő feltekeredésére, harmadlagos szerkezetének felvételére nem lesz idő, így a fehérje belső hidrofób magjai kívülre kerülnek, s ezen keresztül a fehérjék összetapadnak, aggregátumokat képeznek és oldhatatlan formában kicsapódnak. Ezt általában el kell kerülnünk, mert nehéz a fehérjét visszaoldani, illetve renaturálni úgy, hogy visszanyerje a biológiai aktivitását (kivéve, ha eleve zárványtestből szeretnénk preparálni). IPTG-vel történő indukció esetén, ami a legelterjedtebb indukciós módszer, általában 0,01-1 mM között változtatjuk az IPTG koncentrációját. A hőmérséklet hatása is könnyen értelmezhető, hiszen minél magasabb hőmérsékleten történik az indukció, annál gyorsabban megy végbe minden folyamat, így a transzkripció és a transláció is a sejtben. Általában 15-37°C között történik az expresszió. Az optimális hőmérséklet mellett érdemes beállítani az induktor koncentrációját és az expresszió idejét is. Minél hosszabb ideig indukálunk, annál több fehérje keletkezik, ami növeli az esélyét annak, hogy oldhatatlan fehérjét kapunk. Szisztematikusan meg kell nézni a fehérje mennyiségét a paraméterek változtatásával mind az oldható, mind az oldhatatlan frakcióban.

12.2.2.2. Expressziós törzsek (BL21pLysS, BL21pLysE, Rosetta, BL21Star, stb.)

Sokszor segítséget jelent, ha más törzsben fejeztetjük ki a fehérjénket. Ha például arra gyanakszunk, hogy korai termináció miatt leválik a riboszómáról a mRNS, ezért nincs megfelelő méretű termékünk az expresszió során. Ha a fehérjénk cDNS-ének bioinformatikai vizsgálata is azt mutathatja, hogy több ritka kodon van egymás után, amin nem áll módunkban módosítani, akkor érdemes olyan gazdatörzset használnunk, amely tartalmaz ritka kodonú tRNS-eket. Ilyen a BL21 származék **Rosetta** törzs (Novagene). Ha rövid az mRNS élettideje a sejtben, akkor fejeztessük ki a fehérjénket **BL21(DE3)Star** sejtekben. Ennek a törzsnek a fenotípusa: F– *ompT hsdSB*(rB–, mB–) *gal dcm rne131* (DE3). Ebből látszik, hogy alkalmas IPTG indukálta génexpresszióra T7-promóteren keresztül, az *ompT*- és *lon*- proteázok inaktívak (defektesek), ami fontos, hogy ne degradálódjon a rekombináns fehérjénk. A fenti tulajdonságok a **BL21(DE3)** sejtekben is megtalálható, de a Star-nál még a *rne131* jelzés is szerepel, ami azt jelenti, hogy az RNáz-E gén csonkolt, ezért inaktív. Így megnő a sejtben a mRNS stabilitása, élettideje.

Ha toxikus a fehérjénk (pl. aktív protein-kinázt, foszfatázt, proteázt fejeztetünk ki), használjunk **pLysS** törzset, így csak az indukció megkezdésével expresszáldódik a fehérje, addig zavartalanul szaporodhatnak a sejtek. A rekombináns fehérjék előállításához használatos törzsekről kitűnő összefoglaló olvasható itt. Az említett törzsek genotípusáról pedig ezen az oldalon talál az olvasó további részleteket.

12.2.2.3. Promóter optimalizálás

Amennyiben mindent kipróbáltuk, de még így sem expresszálódik megfelelően a fehérje, akkor érdemes más promóterrel is próbálkozni. Ez azt jelenti, hogy más promótert tartalmazó expressziós plazmidba kell az inszertünket szubklónozni. Lehet, hogy a T7lac promóter olyan mértékű expressziót eredményez, ami összes fehérjénket zárványtestbe viszi. Ilyenkor megoldást jelenthet a gyengébb tac vagy lac promóter használata. Ne felejtsük el, a promóterek csak a nekik megfelelő expressziós törzsben működnek. Például a T7 promóter nem működőképes, **ha nincs a baktériumban T7 RNS-polimeráz**, ezért ehhez a promóterhez **módosított BL21 törzset kell használnunk, mint a BL21(DE3) vagy a HMS174(DE3)**. (A DE3 profág hordozza a lac-promóter szabályozása alatt a T7 RNS-polimeráz génjét.)

12.2.2.4. Koexpresszió

A koexpresszió azt jelenti, hogy egyszerre egy sejten belül több polipeptidláncot fejeztetünk ki. Szükségünk lehet erre, ha például egy heterodimer fehérjét kell előállítanunk rekombináns formában. További előny lehet, hogy már a sejten belül létrejön a két polipeptidlánc (vagy fehérje) közötti kölcsönhatás, ami segíthet a láncok feltekeredésében, vagy ha a bennünket érdeklő fehérjével együtt egy módosító enzimet expresszálunk, a poszttranszlációsán módosított rekombináns fehérjét izolálhatjuk a gazdasejtből. Jó példa erre, ha egy fehérjét aktív formában szeretnénk előállítani (pl. egyes MAPK enzimeket), de az aktiváló foszforiláció nélkül ez nem sikerülhet. Megoldást jelenthet, ha a MAPK fehérjét az aktiváló protein-kinázzal (MAPKK) együtt fejeztetjük ki. Így megkapjuk a MAPK fehérjét aktív, foszforilált formában. Koexpresszióra prokariótákban többféle lehetőségünk van:

- Egy promóter, több RBS és inszert cDNS, azaz egy **policisztronos mRNS-t** kapunk, amelyről több fehérje expresszálódik (mint a bakteriális operonoknál). Ilyenkor a mRNS-hez több riboszóma is kötődik az RBS-eknek megfelelően, és egyszerre kezdi a fehérjék szintézisét. Megjegyzendő, hogy a promóterhez közelebb elhelyezkedő ORF-ről több polipeptidlánc fog szintetizálódni.
- **Több azonos promóter** egy plazmidon, amelyről így többféle mRNS íródik át. A módszert például heterodimer fehérje előállítására használhatjuk.
- **Több plazmid**, amely különböző inszertet hordoz. De ennek feltétele, hogy más legyen a plazmidok replikonja, azaz különböző **kompatibilitási csoportba** tartozzanak és különböző antibiotikum rezisztencia gént hordozzanak (ld. 7.2.1. fejezet).
- Egy promóterhez fúziós fehérjeként kapcsoljuk a két különböző fehérjeláncot. Azaz nem használunk **STOP kodont** az első fehérjét kódoló szekvencia után, hanem rögtön következik a másik polipeptidet kódoló szekvencia.

12.2.3. Fúziós rekombináns konstrukciók

Sokszor az is megoldás lehet a sikertelen expresszióra, ha a kívánt fehérjénket fúziós fehérjeként fejeztetjük ki. Ekkor a fehérjénk vagy N-, vagy C-terminális részére, vagy mindkettőre tervezhetünk címkét (*tag-et*). Ezek a címkék segíthetik a

- fehérjénk oldhatóságát, expresszióját,
- olyan szignál szekvenciákat tartalmazhatnak, amelyek megfelelő kompartmentekbe irányítják a fehérjét,
- segítik a fehérje tisztítását affinitás kromatográfiával
- lehetővé teszik a fehérje nyomonkövetését is, pl. enzimaktivitásuk révén vagy antigénként szolgálhatnak Western-blot esetén

A **12.1. videon** egy fúziós rekombináns fehérje tisztítása mutatjuk be affinitás kromatográfiával, FPLC készülék segítségével



12.1. video: Fúziós rekombináns fehérje tisztítása affinitás kromatográfiával (FPLC)

12.2.3.1. GST-címke

A GST, azaz a glutation-S-transzferáz olyan enzim, amely glutation felhasználásával a sejtben lévő kis toxikus molekulákat lebontja. A GST kompakt, globuláris, nagyon stabil fehérje, 25 kDa molekulatömeggel. Nagyon jól kifejeződik és oldatban marad prokarióta sejtekben, ráadásul segíti a hozzá kovalensen kapcsolt fúziós partner kifejeződését is, ha a kifejezendő fehérje N-terminálisára tesszük. Továbbá a GST-címken keresztül a fehérje nagyon jól tisztítható affinitás kromatográfiával. Az előre gyártott glutation-agaróz gyantához (számos cég árulja) ugyanis a GST nagyon nagy affinitással kötődik (szubmikromólos a K_D). A gyantához nagy affinitással kötődő fehérjét szabad glutationt tartalmazó oldattal (10-20 mM) eluálhatjuk. Így elég tiszta fehérje preparátumot kapunk. A GST-gyanta nagy kapacitácú; 1 ml tiszta gyanta kb. 6-8 mg fehérjét tud megkötni, és használat után regenerálható. GST-címkevel expresszálnak a fehérjénket, ha pl. pGEX expressziós vektorok valamelyikével dolgozunk. Részletesebben itt olvashatnak róla.

12.2.3.2. MBP-címke

Az MBP (*maltose-binding protein*), vagyis a **maltóz-kötő fehérje** az *E. coli* malE génjének 42 kDa tömegű terméke. A fehérje a baktérium maltóz transzportjában játszik nagy szerepet, emiatt erős affinitással kötődik a maltózhhoz (a K_D kisebb, mint 1 μ M). Ez a tulajdonsága lehetővé teszi, hogy az agaróz gyöngyhöz immobilizált maltóz a sejtizlátumból megkösse az MBP-fúziós fehérjénket, amelyről szabad maltózzal eluálható a fehérje. Az elúciós oldatunk kb. 10-30 mM maltózt kell tartalmazzon. Mivel a fehérje a baktérium saját fehérjéje, ezért az N-terminálisan hozzákapcsolt fehérje expressziós tulajdonságait (pl. oldhatóság) is jelentősen javíthatja. MBP-fúziós fehérjét a pMal sorozat (New England Biolabs) vektoraival lehet előállítani.

12.2.3.3. Polihisztidin-címke

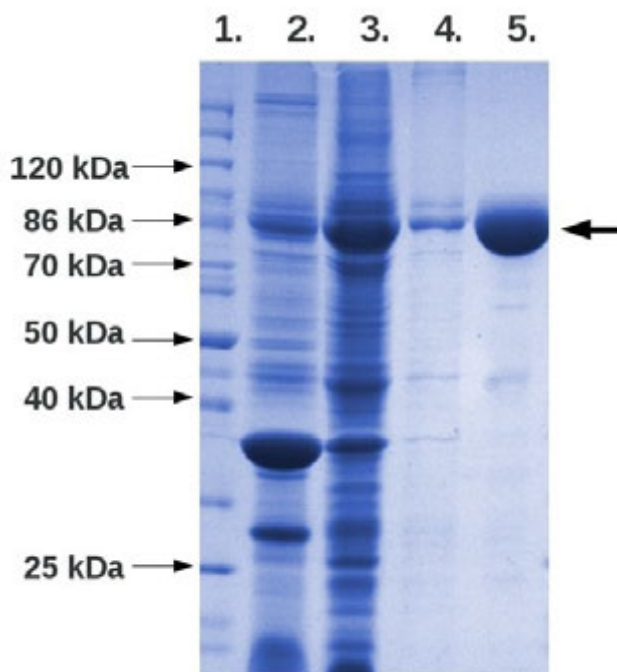
Talán a **leggyakrabban használt fúziós rendszer**. 6-10, de leggyakrabban 6 hisztidin aminosavból álló címke, amit tervezhetünk a fehérje N-, vagy C-terminális végére is. Kisméretű címke, ezért feltehetően a legkevésbé zavarja a fehérje működését. Nagy affinitással kötődik a Ni-gyantához (Ni^{2+} -ionokat kötnek koordinációs kötéssel a gyanta speciális csoportjaihoz). A kötődés azon alapul, hogy a Ni^{2+} -ion koordinációs szférájának a felét foglalják csak el a gyanta funkció csoportjai, így a hisztidin imidazol gyűrűjének nitrogénatomja koordinációs kötést hoz létre a fémionnal is. Ez a kötés elég erős, de fémkelátorokra (EDTA, EGTA) nagyon érzékeny. Ezért az affinitás kromatográfiai lépések során ezeket a komponenseket kerülni kell, mert „leszedik” a fémiont a gyantáról. Szintén kerülni kell a Ni-ionnal csapadékot képező vegyületeket, ilyenek a biokémiai redukálószer, a DTT, vagy a béta-merkaptotetanol, mivel ezek tönkreteszik a gyantát, illetve az oldat pH-ja sem lehet túl lúgos, vagy savas. Viszont ezt az affinitás kromatográfiai lépést lehet erős detergens jelenlétében, vagy denaturáló körülmények között is használni (pl. preparálás inklúziós testből). Nemcsak Ni^{2+} -ion, hanem Co^{2+} -ion is lehet az agaróz gyantához koordinálva, ezért ezt a tisztítási módszert fém-affinitás kromatográfiának (*metal-affinity chromatography*) hívjuk.

Az elúció itt is kompetíción alapul, imidazol oldattal (50-500 mM) szorítjuk le a gyantáról a fehérjénket. Ez a legelterjedtebb módszer, de az elúció megoldható a pH csökkentésével is. Ilyenkor lassan csökkentjük a pH-t,

először a kevésbé specifikus kötések, majd az erősebb kötések szűnnek meg. Az ok, hogy a pH csökkenésével a hisztidin nem-kötő elektronpárral rendelkező N-atomja protonálódik, és nem tud koordinálódni a Ni^{2+} -hez. Ezáltal fokozatos elúciót érhetünk el, így tisztább lehet a fehérje preparátumunk (a Ni^{2+} -gyanta nemcsak polihisztidin-címkés fehérjéket köt, hanem más fehérjéket is, ahol térszomszágban vannak hisztidin aminosavak, persze azokat kisebb affinitással). A **12.7. ábrán** látható egy 85 kDa tömegű fehérje Ni^{2+} -affinitás kromatográfiával történő tisztítása.

A tisztítás különböző lépéseiben mintát vettünk, és megfuttattuk SDS-poliakrilamid gélen, amit fehérjefestékkel (Coomassie Brilliant Blue) hívtunk elő. Látható, hogy a 100 mM imidazollal történő elúció nem elég ahhoz, hogy leszorítsa a gyantáról a fehérjénket, míg 400 mM imidazollal nagy mennyiségű, viszonylag tiszta preparátumot kapunk.

A His-taget gyakran kombinálják más címkével is, pl. N-terminális GST-, vagy MBP-fúzióval. A tisztításnál először Ni^{2+} -affinitás kromatográfiával izoláljuk a sejtizrátumból a fehérjét, majd a másik címke felhasználásával a másik kromatográfiás tisztítást is elvégezzük. Ezáltal jelentősen megnövekszik a fehérje tisztasága és ha a két címke a fehérje két oldalán található, ez lehetőséget teremt arra is, hogy csak a teljes hosszúságú fehérje maradjon a preparátumban.



12.7. ábra: Ni^{2+} -affinitás tisztítás gélelektroforetikus képe. 1. sáv: fehérje marker; 2. sáv: a sejt nem oldható frakciója (centrifugálás utáni csapadék); 3. sáv: a sejt oldható frakciója (felülúszó); 4. sáv: elúció 100 mM imidazollal; 5. sáv: elúció 400 mM imidazollal

12.2.3.4. Egyéb címkék

Nagyon sok expressziós vektor kínál fel a lehetőséget, hogy N-, vagy C-terminális címkével fejeztessük ki a fehérjénket (ld. 12.2. táblázat).

12.2. táblázat: Fúziós címkék és alkalmazásai

<i>Affinitás címke</i>	<i>Méret (aminosavak száma)</i>	<i>Helye</i>	<i>Affinitás-gyanta</i>
poli-hisztidin	6-8	N, C, belső	Ni ²⁺ -
Strep-tag	8	N, C	sztreptavidin
T7-tag	11 vagy 16	N, belső	
Flag-tag	8	N, C	anti-Flag
HA-tag	9	N, C	anti-HA
S-tag	15	N, C	S-protein
C-myc-tag	11	N, C	anti-myc
poli-arginin	25 kD	N	S-Sepharose
Kitin kötő domén	51	N, C	
Kalmodulin kötő domén	26	N, C	kalmodulin
Cellulóz kötő domén	27-129	N, C	cellulóz
<i>Oldhatóságot növelő címkék</i>			<i>Felhasználás</i>
GST	211 (26 kD)	N	tisztítás (GST-gyanta) +oldhatóság
MBP	396 (40 kD)	N, C	Tisztítás (MBP- gyanta)+oldhatóság
SUMO	100	N	oldhatóság
NusA	495 (55 kD)	N	oldhatóság
Ubiquitin	76	N	oldhatóság
Thioredoxin	109	N, C	oldhatóság

A kérdés csak az, mi a célunk. Növelni akarjuk a fehérje oldhatóságát, expresszióját vagy nagyobb mértékű tisztítást szeretnénk elérni? Nyomon szeretnénk követni a fehérjénk kifejeződését, lokalizációját pl. eukarióta sejtekben? Általában elmondhatjuk, hogy egy címkének több funkciója is van, és csoportosításuk egyik alapja éppen a felhasználás lehet. Ez alapján beszélhetünk **affinitás**, **szolubilizációs**, **epitóp**, vagy akár **fluoreszcens** címkéről is. Egy-egy címke több csoportba is besorolható, pl. a GST-címke növeli a fehérje oldhatóságát, lehet affinitás-alapú tisztításra használni, és még ellenanyag is készült rá, amivel nyomon lehet követni. Az alkalmazott címkék ellen általában létezik ellenanyag, így Western-blottal vagy eukarióta sejtek esetén immunfestéssel láthatóvá tehetők. Léteznek olyan eukarióta vektorok is, amelyek fluoreszcens fehérjéket, vagy enzimaktivitással rendelkező fehérjéket kínálnak fúziós címkének. Ilyen címke a **GFP** (*green fluorescent protein*) és változatai, vagy az enzimaktivitással rendelkező β -galaktozidáz, illetve luciferáz. Ezeket azonban felhasználásuk miatt az eukarióta expressziós rendszerrel tárgyaljuk.

Az expressziós címkék használatáról itt találnak további hasznos információkat.

12.2.3.5. A fúziós címkék eltávolítása

A nagyobb méretű címke jelenléte befolyásolhatja a fehérje biológiai funkcióját pl. aktivitását, kötődését más fehérjéhez, vagy a harmadlagos, negyedleges szerkezetének kialakulását. Például a GST dimerizációra hajlamos, így már eleve rákényszerítheti a fúziós partnerét is, hogy dimerizáljon. Ezen okok miatt érdemes tisztítás után a fehérjéről valamilyen proteázzal levágni a fúziós címkét. Számos vektorban már eleve betervezték a proteáz hasító/felismerőhelyét. Természetesen olyan enzimeket kell használni, amelyeknek nagyon speciális hasítóhelye van, és ritkán fordul elő a fehérjékben, így nem hasítja más helyeken is a rekombináns fehérjénket. Néhány esetben a kémiai hasítás is szóba jöhet; pl. a bróm-cián AspPro aminosavak között hasít. A **12.3. táblázat** foglalja össze a leggyakrabban használt proteolitikus enzimeket hasítóhelyeikkel együtt.

A proteázzal való hasítás történhet az elúció után, oldat fázisban. Ilyenkor a fúziós címkétől, a nem hasított fúziós fehérjétől és a proteáztól is meg kell tisztítanunk a fehérjénket valamilyen kromatográfiával. Kisméretű (peptid) címkéket el lehet távolítani dialízissel, vagy gélszűrővel. Másik lehetőség, hogy a gyantára felkötött és mosással tisztított fehérjéről az oszlopon hasítjuk le a fúziós címkét. Ilyenkor a címke az affinitás oszlopon marad, míg a fehérjénk „lecsöpög” az oszlopról. Ekkor csak a proteáztól kell megszabadulnunk. Vigyáznunk kell, hogy a proteáznak és a fehérjénknek is megfelelő puffert alkalmazzunk a hasítás során.

12.3. táblázat: Fúziós címke lehasítására alkalmas proteolitikus enzimek és jellemzőik

Proteáz	Hasítóhely	pozíció	pH	Detergens érzékenység	Só- érzékenység	Enzim/szubsztrát tömegarány
TEV	ENLYFQ/(G/S)	N, C	5,5- 8,5	2 M Urea	Max. 0,1 M	1-3 %
Thrombin	LVPR/G	N, C	6-9	Max. 0,1 M Urea	Max. 0,15 M	1-10%
Xa factor	IEGR/X	N, C	6-9	Max. 0,1 M Urea	Max. 0,15 M	1-10 %
Enterokináz	DDDDK/X	N, C	7-8			0,1 %
SUMO proteáz	SUMO C-term után	N	6-9	Max. 2 M Urea	Max. 0,5 M	0,1 %
Szubtilizin	Propeptid C-term. után	N	7,2			1:1, mól:mól
PreScission	LEVLEQ/GP					

12.2.4. Rekombináns fehérjék termelése és izolálása

Ha optimalizáltuk fehérjénk expresszióját, és tudjuk azt is, hogy a sejt mely részéből kell tovább tisztítanunk, akkor elkezdhetjük a nagyobb mennyiségű fehérje termelését célzó, nagy léptékű expresszió végrehajtását, amely a következő alfejezetekben ismertetett lépésekből áll.

12.2.4.1. Indító (starter) kultúra elkészítése

Az elkészült rekombináns DNS konstrukciót transzformáljuk a megfelelő expressziós törzsbe (pl. BL21(DE3)). A transzformátumot megfelelő antibiotikumot tartalmazó szilárd táptalajra szélesztjük és 37°C-on inkubáljuk. Másnap egy jól körülhatárolt telepet antibiotikumot tartalmazó 50 ml LB-tápra oltunk. Ez lesz a starter kultúra, amit 37°C-on rázatás mellett növesztünk. 12-16 óra múlva, a starter kultúra megfelelően felszaporodik. 10-20 ml-ét 1 l megfelelő antibiotikumot tartalmazó folyékony táptalajba oltjuk, amivel elindítjuk az expressziós kultúránk növesztését. A megfelelő expressziós klón folyadékkultúrában felnövesztett szuszpenziója 10% glicerinnel (amely ez esetben a fagyálló szerepét ölti be) elkeverve lefagyasztható és -80°C-on hosszú ideig tárolható. Ilyen glicerines expressziós klontár létrehozásával kiválthatjuk a gyakran használt klónok expressziós sejtbe való transzformálását, elég a fagyasztott kultúra töredék részét szilárd táptalajra inokulálni.

12.2.4.2. Az expressziós kultúra növesztése és indukciója

A megfelelően inokulált (ld. 12.2.4.1.) baktérium tenyészetet 37°C-on, folyamatos rázatás (140-200 rpm) mellett inkubáljuk. Egy rázóinkubátort a 12.8. ábrán, míg az expressziós kultúrához szükséges eszközök csíramentesítését és a sejt kultúra növesztését a 12.2. videon mutatjuk be. Inkubáció közben nyomon követjük az osztódó baktériumtenyészet sűrűségét. A tenyészet sűrűsége jelzi az osztódások előrehaladását, a baktériumpopuláció szaporodását és egyszerűen, az optikai denzitás 600 nm-en való mérésével, spektrofotométeren mérhető. $OD_{600\text{ nm}} = 0,8$ körüli értéknél a baktériumkultúra készen áll az indukcióra. Ha szükséges, hűtjük le az indukció hőmérsékletére és vegyünk belőle mintát a későbbi SDS-gél futtatáshoz. Ez lesz az indukció előtti minta. Ha a termosztátunk hőmérséklete elérte az indukcióhoz szükséges értéket tegyük bele a megfelelő mennyiségű induktort, ami a leggyakrabban IPTG. Inkubáljuk és rázassuk a megfelelő ideig, majd vegyünk belőle mintát. Ez lesz az indukció utáni minta. Az expressziós kultúrából a következő lépések során tisztítjuk a termelt rekombináns fehérjét.



12.8. ábra: Rázató inkubátor prokarióta és eukarióta szuszpenziós kultúrához

12.2. video: Expressziós kultúrához szükséges eszközök csíramentesítése és a sejt kultúra növesztése

12.2.4.3. A sejtek összegyűjtése (*harvesting*)

Ahhoz, hogy a sejtekből a fehérjét a megfelelő tisztítási lépésekkel izoláljuk, először össze kell gyűjtenünk a sejteket a táptalajból. Ez a lépés centrifugálással történik (5000 g x 10 min, 4°C-on). A tisztítás során mindig jégben, vagy 4°C-on dolgozunk, hogy ne degradálódjon a fehérje. A sejteket meg kell tisztítani a táptalajtól, amit legtöbbször egy pufferrel (pl. PBS) való mosással végzünk. A sejteket ezután újra centrifugálással gyűjtjük össze.

12.2.4.4. A sejtek feltárása

Az összegyűjtött sejteket lizálva a sejtben levő fehérje kinyerhető. A lízist a sejtek lízis pufferben való inkubálásával végezzük. 1g baktérium sejtre általában 3-4 ml lízis puffert számolunk.

A lízis puffer összetétele:

- Megfelelő pH-t biztosító puffer (általában 7-9 pH-tartományban): pl. 20 mM TRIS, foszfát puffer.
- Só, ami ionerősséget biztosít, általában NaCl, 0,15-0,5 M.
- Detergens, ami segíti a sejtek feltárását, illetve a fehérjék oldatban tartását és csökkenti az aspecifikus kölcsönhatásokat, pl. 0,1-0,2% Tween20, CHAPs, TritonX 100.
- Glicerín 5-10%-ban, ami stabilizálja a fehérjéket.

- Redukálószer, a nem megfelelő diszulfidhidak redukálásához: 2-10 mM béta-merkaptotanol (a DTT-re és DTE-re vigyázni kell fém-affinitás kromatográfia esetén).
- Proteáz gátlók, a sejtek degradálódásának kivédésére: pl. 1 mM PMSF, 1-5 mM benzamidin.

A lízist segíti, ha többszörös fagyasztást-felolvasztást alkalmazunk, és/vagy ha a felfuszpendált sejt kultúrához **lizozimot adunk**. A lizozim képes a baktériumok sejtfalát bontani és ezért alkalmas a lízisre. Nagyon hatásos a Gram-pozitív baktériumoknál, de a Gram-negatív baktériumoknál is hasznos, ha detergenssel, vagy ozmotikus sokkal együtt alkalmazzuk. **Szonikálással** (a sejtek ultrahang segítségével történő feltárásával) is segíthetjük a sejtek lízisét, és a benne lévő DNS darabolását. Vigyázzunk azonban, mert szonikálás hatására melegszik az oldat, ami a preparátumban lévő proteázok aktivitásának növekedését, így a fehérjénk degradálódását okozhatja, és emellett a fehérjénk denaturációjához vezethet. Ezért ezt a folyamatot jégben és szünetekkel végezzük (1 perc szonikálás meghatározott frekvenciával, majd egy perc szünet). Ezt a lépést ismételjük 3-4-szer, míg a preparátum hígán folyóssá nem válik. Ezután DN-áz I enzimmel kezeljük, ami a DNS-t kisebb darabokra vágja. A sejtek feltárását végezhetjük erős detergenssel (TritonX 100 1-2%), deoxikólsavval vagy az ún. French-Press készülékkel is (ez utóbbiban a nagy nyomás tárja fel a sejteket). A feltárás után centrifugálással elválasztjuk az oldható és az oldhatatlan frakciót. A **felülúszó** a lizispufferben oldódó fehérjéket, komponenseket, míg az oldhatatlan frakció (**csapadék**) a sejtörmelékét és az inklúziós testet tartalmazza. A további lépések attól függenek, hogy hol található a fehérjénk; a felülúszóban oldható formában, inklúziós testben oldhatatlan formában, vagy a periplazmás térben.

12.2.4.5. Izolálás oldható frakcióból

Ha a fehérjénk az oldható frakcióban van jelen, feltehetően natív formában, akkor ezt a protokollt kell követnünk. A baktérium lizátumban számos fehérje van, köztük proteázok is, amelyek emészthetik a rekombináns fehérjénket. A tisztítás során ezeket el kell választanunk. A sejtek összegyűjtése után a felülúszót visszük tovább, és a címkének megfelelő affinitás kromatográfás tisztítási lépés következik. Amennyiben a fehérjénk tartalmaz polihisztidin címkét, ezt a legcélszerűbb alkalmazni, mert nagy affinitású és denaturáló körülmények között is működik.

12.2.4.6. Preparálás zárványtestből (*inclusion body*)

Ha a fehérjénk nem az oldható fázisban, hanem a csapadékban van jelen, akkor fel kell oldanunk, amire számos lehetőség kínálkozik: 5-8 M guanidin-HCl, 6-8 M urea, SDS-oldat, lúgos pH. A csapadékot először mossuk 1-2% Triton X 100 oldattal, majd centrifugáljuk. A mosott csapadékot szolubilizáló oldatban (5-8 M guanidin-HCl vagy 6-8 M urea) vesszük fel. A csapadékot fel-le pipettázással oldjuk, és 4°C-on több órán át állni hagyjuk. Inkubálás után tisztítás céljából ismét lecentrifugáljuk az oldatot, de most már a felülúszó tartalmazza a feloldott fehérjénket, igaz a denaturálószerben. Választhatunk, hogy elvégezzük denaturáló körülmények között a fém-affinitás kromatográfiát, vagy megpróbáljuk renaturálni a fehérjét, és utána alkalmazzuk a tisztítási lépést. Legcélszerűbb az először alkalmazott fém-affinitás kromatográfia, hisz ez, mint már említettük, denaturáló körülmények között is jól működik, feltétele viszont, hogy a fehérjénk tartalmazzon His-címkét. A gyantára felkötött és mosott fehérjét vagy olyan elúciós pufferrel eluáljuk, ami még tartalmazza a denaturálószer, vagy lecseréljük a mosó-eluáló puffert denaturálószer nélküli pufferre.

Amennyiben az elúciós puffer is tartalmaz denaturálószer, akkor meg kell szabadulni tőle, hogy a fehérjénk renaturálódni tudjon. Erre legjobb módszer a dialízis. Ez történhet lépésenként, azaz fokozatosan csökkentjük a dializáló pufferben a denaturálószer koncentrációját, és közben emeljük a NaCl mennyiségét, hogy ne aggregálódjon a fehérje. A dialízis során a fehérje visszanyeri a szerkezetét, és remélhetőleg a biológiai aktivitását is. Ezt persze függ a fehérjétől, minél több doménből áll egy fehérje, annál reménytelenebb a renaturáció. Szintén reménytelen a renaturáció, ha a fehérjénk feltekeredéséhez az eredeti sejt környezetben dajkafehérjékre van szükség. A dialízis során a ciszteint tartalmazó fehérjénél mindenképpen alkalmazzunk redukálószer: DTE, DTT, béta-merkaptotanol, TCEP (Trisz-karboximetil-foszfát). Ennél a módszernél nagyon fontos, hogy a fehérje elég jól renaturálódjon, és visszanyerje biológiai aktivitását. Előnye viszont, hogy nagy tisztaságú és általában nagy mennyiségű fehérjét lehet előállítani. Az inklúziós test nagyon kevés egyéb fehérjét (és így proteázt) tartalmaz, a rekombináns fehérjénk akár az összefehérje 90%-a is lehet.

12.2.4.7. Izolálás periplazmából

A periplazmás tér a Gram-negatív baktériumok külső és belső membránja közötti tér. Itt jobb lehetőségek kínálkoznak a megfelelő foldingra és a diszulfidhíd képződésre. Ide megfelelő N-terminális szignál szekvenciát tartalmazó

fehérjék jutnak el. Ilyen szekvencia az ompA, pelB, vagy a phoA-fúziós fehérjék. Ilyenkor a tisztításhoz használt címkét a C-terminálisra kell tenni.

A sejt feltárása enyhébb körülmények között, nem szonikálással, hanem ozmotikus sokkal történik. A sejteket 80 ml puffer/1 g sejtsapadék mennyiségű lízis pufferben kell felvenni, amely tartalmaz 20% szacharózt is. Az izoozmotikus oldatban lévő sejteket centrifugáljuk (8000 g x 20 min, 4°C), majd szuszpendáljuk őket nagy mennyiségű 5 mM MgCl₂ oldatban, végül újabb centrifugálás következik. Az így biztosított hipoozmotikus körülmények között a periplazmás fehérjék már a felülúszóba kerülnek.

Másik bevált módszer, hogy az összegyűjtött sejteket rögtön a MgCl₂ oldatban vesszük fel, ami hipoozmotikus sokk a sejteknek, és óvatosan lefagyasztjuk -20°C-on, majd óvatosan felengedjük. Ilyenkor a sejt külső membránja megsérül, de a belső épen marad. Így centrifugálással elválaszthatjuk a sejteket, míg a felülúszóban marad a periplazmatikus tér tartalma.

12.2.5. A prokarióta expresszió előnyei, hátrányai

Egyik legfontosabb előnye, hogy viszonylag **gyors** (ha kész a konstrukció, 3-4 nap alatt termelhető a fehérje), **olcsó**, és **nagy mennyiségű** fehérje állítható el vele. Hátránya, hogy a fehérjéről hiányzik az esetleges **poszttranszlációs módosítás** (foszforiláció, glikoziláció, acetilálás), problémás lehet az S-S hidak képződése, illetve a fehérje feltekeredése. Sokszor a nagyobb, több doménből álló fehérjéket is rosszul lehet kifejezteni ebben a rendszerben. Végül emlékeztetünk a kodon preferenciára, ami ha nem megfelelő, korai termináció jöhet létre (bár ez a veszély megfelelő törzzsel csökkenthető (**ld. 12.2.2. fejezet**)).

12.3. További olvasnivaló a fejezethez

Sambrook, J and Russel, DW (2001) Molecular Cloning. A Laboratory Manual. CSHL Press, volume 1-3, Third edition. ISBN: 0879695773.

Terpe K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 72:211–222

Graslund S.,Sagemark J., Berglund H., Dahlgren L. G., Flores A et al. (2008) The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. Protein Expr Purif. 58(2), 210-21.

Structural Genomics Consortium et al. (2008) Protein production and purification. Nature Methods 5(2):135-146

13. fejezet - Eukarióta expressziós rendszerek

Napjainkban a fehérjemérnökségnek egyre nagyobb jelentősége van nemcsak a kutatásban, hanem a gyógyászatban is. Biotechnológiai cégek alakulnak, hogy ellenanyagokat, hormonokat, növekedési faktorokat és egyéb polipeptid típusú molekulákat állítsanak elő (sokszor közvetlenül gyógyászati célból; ld. pl. 15.1.3. fejezet). A bakteriális expresszió általában ellát bennünket nagy mennyiségű rekombináns fehérjével, de a fehérjénk nem biztos, hogy fel tudja venni a biológiailag aktív formáját. Ennek oka lehet egyrészt, hogy a baktériumban kevés és más dajkafehérjék (chaperonok) működnek, így nem lesz megfelelő a fehérjelánc feltekeredése (foldingja) és a szerkezete. Másrészt a poszttranszlációs módosítások sem érvényesülnek, pl. foszforiláció, glikoziláció, stb. Harmadrészt, a citoplazma redukzív környezete miatt a megfelelő diszulfid-hidak sem tudnak kialakulni. Nagyméretű, több doménből álló fehérjék előállítására is problémás lehet prokarióta rendszerben. Erre az *in vitro* rendszer részben megoldást jelenthet, hisz elérhető emlős, vagy akár humán sejtmentes kivonat is. Ezek hátránya azonban, hogy kevés fehérje állítható elő velük. Ha ilyen problémával találkozunk, meg kell keresni a megfelelő eukarióta expressziós rendszert. Milyen feltételei vannak egy jó expressziós rendszernek?

- Gyorsan és hatékonyan szaporodjon az organizmus
- Nagy mennyiségű fehérjét termeljen
- Megfelelő formában állítsa elő a fehérjét
- Könnyen kezelhető, olcsó és biztonságos legyen

13.1. Élesztő expressziós rendszer

A fejezet bevezetőjében említett céloknak majdnem teljesen megfelel az élesztő, a legegyszerűbb eukarióta gazdasejt. Genetikája, biokémiája nagyon jól ismert, már több mint 30 éve használják rekombináns fehérje termelésre. Nagyon sok élesztő faj ismert (kb. 800), ebből sok alkalmas fehérjetermelésre. A két leggyakrabban használt élesztő a *Pichia pastoris*, illetve a közönséges sörélesztő, a *Saccharomyces cerevisiae*. Ezen a két rendszeren keresztül mutatjuk be az élesztő expressziós rendszerek jellemző tulajdonságait. Az élesztőben használható vektor típusokat már röviden áttekintettük a 7.6. fejezetben.

Minden expressziós rendszer két alapvető egységből áll: a megfelelő gazdasejtből és az expressziós vektorból. Az expressziós vektornak hordoznia kell azokat a szekvenciákat, amely biztosítja a gazdasejtben a replikációt, és a kívánt fehérje kifejeződését. Az élesztő vektorok **shuttle („ingázó”) vektorok**, mivel a baktériumsejtek és az expressziós sejtjeink (élesztő, rovarsejt, emlőssejt) is replikálni tudják. Ez úgy érhető el, hogy mindkét gazdasejt típusra specifikus replikonnal rendelkeznek. Mivel minden fehérjetermelés első lépése az expressziós konstrukció létrehozása és a molekuláris klónozás, a rekombináns DNS konstrukció felszaporítása *E. coli* gazdában a legegyszerűbb, leggyorsabb és legolcsóbb. Ezt a műveletet bármilyen expressziós gazdasejtet is használunk, a bakteriális rendszerben végezzük el.

Alapvetően kétféle típusú élesztő vektor létezik: replikálódó és integrálódó. A **replikálódó (episzomális) vektor** extrakromoszomális formában szaporodik a gazdasejtben. Ilyen vektor használatakor általában csak **átmeneti (tranzienti)** expresszióról beszélünk és a vektornak tartalmaznia kell egy élesztőben működő replikációs start szignált is. A másik esetben a kifejezendő fehérje cDNS-e homológ rekombinációval az élesztő genomjába integrálódik. Ekkor **stabil expresszióról** beszélünk, azaz a rekombináns transzformált törzs megőrzi a transzgént, a törzs fenntartható akár hosszú távon is. A homológ rekombinációt **integrálódó vektorokkal** érjük el, amelyek tartalmazzák a genomi DNS szekvenciával homológ szekvenciákat. A homológ rekombináció hatásfoka elég kicsi, ezért mindenképpen ki kell válogatni a transzformálódott sejteket. Erre a célra **szelekciós markereket** használunk, amit a vektorunk kódol. Az **expressziós kazetta** hordozza a kívánt fehérjét kódoló szekvenciát és annak kifejeződéséhez szükséges elemeket; promóter, transzkripció terminációs szekvencia, riboszóma kötőhely (Kozak szekvencia). Itt található a multiklónozó hely (MCS), ahová a klónozás során bevisszük a fehérjét kódoló szekvenciát.

Összefoglalva, az élesztő expressziós vektorok fő elemei:

1. bakteriális elemek: bakteriális replikációs origó, bakteriális antibiotikum rezisztencia gén
2. élesztő genetikai elemek, amelyek a homológ rekombinációért vagy a vektor replikációjáért felelősek (ARS: *autonomous replication sequences*)
3. szelekciós markerek
4. expressziós kazetta

13.1.1. Az expressziós kazetta

A fehérje expresszió legmeghatározóbb eleme a promóter. Milyen legyen egy jó promóter? Nagymértékű és speciális expressziót biztosítson, amit lehet szabályozni, indukálni. Erre legalkalmasabbak a glikolitikus, vagy más metabolikus utak enzimeit kódoló gének promóterei. A következőkben a leggyakrabban használt promótereket foglaljuk össze:

S. cerevisiae promóterek:

- Nagyon erős, de csak gyengén indukálható promóterek: **PGK** (foszfoglicerát-kináz), **GAPD** (glicerín aldehid-3-foszfát-dehidrogenáz); induktoruk a glükóz
- Erős promóter: GAL1 (galaktokináz), galaktózzal indukálható.
- Közepesen erős promóter: ADH2 (alkohol dehidrogenáz), a glükóz represszálja.
- Közepesen erős promóter: PHO5 (foszfátáz), a foszfát represszálja.

Pichia pastoris promóterek:

- **AOX1 promóter:** a *Pichia pastoris* alkohol-oxidáz génjének promótere.

Az alkohol-oxidáz enzim a levegő oxigénjének felhasználásával a metilalkoholt formaldehiddé oxidálja, miközben hidrogén-peroxid keletkezik. (A folyamat a peroxisómában játszódik le, így nem káros a sejtnak.) Ez a promóter nagyfokú expressziót biztosít, metilalkohol hatására akár 5% lehet az enzim mRNS-ének a mennyisége a teljes mRNS mennyiségre vonatkoztatva, a fehérjeszint pedig az összes fehérje 30%-át is elérheti. Tudnunk kell, hogy a *Pichia* metilotróp élesztő, azaz metanolt tud használni egyedüli szén-és energiaforrásként, ha nincs a médiumban glükóz és glicerín. Tehát a glükóz és a glicerín represszálja, a metilalkohol pedig indukálja a promóter után lévő szekvencia átírását. Ez egy ideális promóter, hiszen indukálható módon nagymértékű expressziót biztosít.

- AOX2: a másik alkohol-oxidáz enzim génjének promótere. Szintén indukálható metilalkohollal, de 10-szer kisebb expressziót biztosít.
- DHAS: a dihidrox-aceton-szintetáz gén promótere. Ez a promóter is a metanol metabolizmusban játszik szerepet, azaz metanollal indukálható, de nincs olyan erős, mint az AOX1.
- FLD1: a formaldehid dehidrogenáz gén promótere szintén a metanol hasznosításban játszik szerepet. De metanol mellett metilaminnal is indukálható, így a metilamin indukció esetén glükóz és glicerín is használható szénforrásként.
- GAP1: a glicerinaldehid-3-foszfát-dehidrogenáz promótere állandó expressziót eredményez, azaz konstitutív promóter. Akkor használható, ha a termelt fehérje nem toxikus a sejtre, illetve nem gátolja a sejt osztódását.
- AOX1-GAP1: tandem promóter, amellyel még nagyobb expresszió érhető el.

További szabályozó elemek:

A promóter mellett ezeknél a vektoroknál is fontos, hogy az expressziós kazetta tartalmazzon **transzkripció terminációs szekvenciát**, valamint a mRNS processzálásához és a poliadenilációhoz szükséges elemeket.

Legtöbbször a PGK, GAPD és CYC1 gének transzkripció terminációs szekvenciáját használják. Fontos még a transláció iniciációjához szükséges szekvenciák jelenléte is. Ezt a szekvenciát eukariótákban **Kozak szekvenciának** nevezzük, ami nélkül nem, vagy nagyon gyengén megy az expresszió. Egyes vektorokban (különösen N-terminál fűzős konstrukciókban) ez már adott, különben nekünk kell beépítenünk.

13.1.2. Szelekciós markerek

Szelekciót két különböző elv alapján végezhetünk. Az egyik esetben az expressziós vektor bioszintetikus utak enzimeit kódolja, pl. **HIS3, HIS4, ADE1, URA3, LEU2, TRP1**. Ebben az esetben olyan törzset kell használnunk, amelyben az adott enzim defektes (mutáns) és ezért nem tud szaporodni olyan tápoldatban, ami nem tartalmazza a szintézisében gátolt anyagot, pl. hisztidint, leucint, triptofánt vagy az uracilt (ezeket a törzsetek auxotrófnak hívjuk). Például HIS3 szelekciós marker gén használata esetén a hisztidin bioszintézisében gátolt törzset hisztidint tartalmazó tápoldatban tartjuk fenn. Transzformáció után azonban hisztidin-mentes tápfolyadékban növesztjük. Ekkor csak azok a sejtek maradnak fenn (szaporodnak), amelyek felvették a vektort, és elő tudják állítani a hisztidin bioszintéziséhez szükséges enzimet.

A másik lehetőség, ha a vektor antibiotikum rezisztencia gént kódol. Azaz a baktérium törzsekhez hasonlóan a rezisztenciagén terméke hatástalanítja a médiumban lévő antibiotikumot (pl. **G418**, amely egy aminoglikozid antibiotikum, s a neomicin-rezisztencia génnel lehet inaktiválni, vagy **zeocin**, amely egy glikopeptid és Zeocin rezisztencia géntermék inaktiválja). Ezek a szelekciós ágensek minden eukarióta sejtben működnek, és a rezisztencia dóziszfüggő. Így az expressziós vektorból több kópiát tartalmazó törzsek is szelektálhatóak növelve a médium antibiotikum tartalmát.

13.1.3. Szekréció

Az élesztő rendszer egyik legnagyobb előnye, hogy a fehérjék szekrécióját könnyű megoldani. Mivel az élesztő nem nagyon szekretál fehérjét, ezért a rekombináns fehérje szinte az egyedüli, ami megjelenik a táptalajban. A szekréciós szignál ráadásul le is hasítja (szignál-peptidáz), így azt sem kell eltávolítanunk. Ilyen szignál-szekvencia lehet a *Saccharomyces cerevisiae* α -mating faktorának prepro-szignálja, vagy a *Pichia pastoris* PHO1 szignálja. Számos esetben a heterológ fehérje saját szignál szekvenciája a legjobb megoldás.

A szekréció nemcsak a tisztítás szempontjából kedvező, hanem a sejt intracelluláris proteázaitól is védelmet jelent. Intracelluláris expresszió esetén proteáz-deficiens törzsek használata jelenthet megoldást. (A két legfontosabb proteáz ebből a szempontból a proteináz-A és -B enzimek). A tápfolyadékban megjelenő proteázok hatását pH optimalizálásával (minden enzim működésének van pH-függése), vagy a fermentációs hőmérséklet csökkentésével lehet minimalizálni (alacsonyabb hőmérsékleten kevésbé működnek az enzimek, illetve kevesebb proteáz szekretálnak.)

13.1.4. Poszttranszlációs módosítások

Az eukarióta **élesztő poszttranszlációs módosításai közelebb állnak az emlősökhöz.** A termelt fehérjék foldingja megfelelő, a diszulfid-hidak képzése is megtörténik a megfelelő helyen. A legnagyobb bajt a glikoziláció jelenti. Emlősökben O-glikoziláció történik, azaz Ser és Thr aminosavak hidroxil-csoportján keresztül, míg élesztőben N-glikoziláció aszparagin oldalláncon keresztül. Ráadásul a cukor egységek mennyisége és fajtája is nagyon különbözik. Emlősökben a cukoregység N-acetil-galaktózamin, galaktóz és szialinsav, míg élesztőben kizárólag mannóz, és az élesztő fehérjék hiperglikoziláltak. Így az élesztőben gyártott fehérjék emlősökben nagyon erős antigén tulajdonsággal rendelkeznek.

13.1.5. Az expresszió lépései

Az eukarióta expresszió lépései alapvetően megegyeznek a prokarióta rendszer megfelelő lépéseivel (**ld. 12.2.4. fejezet**). Csak a különbségeket emeljük ki.

1. Az expressziós kazettában a promóter és transzkripció terminációs szekvencia között van az expresszálandó fehérje kódoló régiója. A terminációs szekvencia ebben a rendszerben esszenciális.

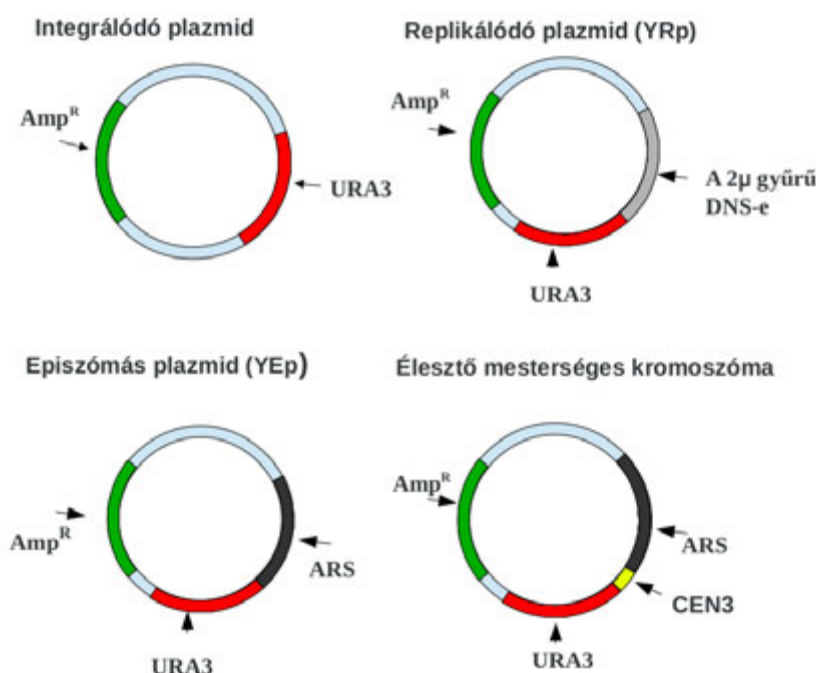
2. Integrálódó vektor esetében a homológ rekombináció hatékonyságának növelésére a vektort linearizáljuk restrikciós hasítással. Fontos, hogy a hasítás a vektor kromoszomális szekvenciájánál legyen, hogy itt történjen meg a homológ rekombináció.
3. Az expressziós vektor bevitele az élesztősejtekbe. Az élesztő cellulózt tartalmazó sejtfala miatt nehéz a transzformálás, ezért gyakran a szferoplasztot (a sejtfalától megfosztott élesztősejtet) transzformálják. Több transzformációs eljárás is létezik, pl. a lítium-kloridos és polietilén-glikolos transzformálás. A leghatékonyabb génbevétel azonban az elektroporálás. A replikálódó vektoroknál a fehérjénket kódoló szekvencia extrakromoszomális marad, míg integrálódó vektoroknál megtörténik a homológ rekombináció az élesztő genomba.
4. A transzformált törzsek kiválasztása a szelekciós markerek segítségével, pl. hisztidin-mentes táptalaj.
5. Ellenőrző PCR annak kimutatására, hogy a genomban van-e a kívánt DNS. Ez a vektorban lévő primer szekvenciák segítségével, vagy egy vektor-specifikus és egy inszert-specifikus primer pár segítségével történik. A keletkezett PCR-termék jelenléte és mérete mondja meg, hogy a megfelelő helyen van-e az inszertünk.
6. Ha készen van a rekombináns törzsünk, elkezdhetjük a sejtek növesztését. Az élesztő sejteket 28-30°C-on növesztjük 220-250 rpm-es rázatás mellett. A viszonylag nagy fordulatszámra való rázatás a megfelelő levegőztetés miatt fontos.
7. Amennyiben a sejtek elérték a megfelelő sejtsűrűséget, indukálhatjuk a tenyészetet. Pl. *Pichia* sejteknél metanollal, közben állandóan mintákat veszünk (0, 6, 12, 24, 36, 48, 72, 96 órában), és SDS-gélelektroforézissel nyomon követjük a fehérjénk termelődését.
8. Ezt követően kell a sejteket feltárunk és a fehérjét izolálnunk, tisztítani. A sejtek feltárása, amennyiben nem szekretált fehérjéről van szó, a sejtfal jelenléte miatt nagyon nehéz. A fel-leolvasztgatás és a savas-üvegyöngyös ismételt vortexelés azonban sokat segíthet. Alternatív megoldásként használhatunk különböző glikoziláz enzimeket is (zimoláz, lizozim stb.). Szekretált fehérjéket a médiumból gyűjtjük össze.

13.1.6. Élesztő expressziós vektorok és rendszerek

Az élesztő sejtekben a baktériumokhoz hasonlóan előfordulnak plazmidok. Legismertebb a sörélesztő (*S. cerevisiae*) „2-mikron” plazmidja, amelyik egy 2 mm átmérőjű cirkuláris DNS, mely autonóm módon replikálódik. Replikonja az origóból és a REP (represszor) szabályozó génből áll. Amíg a plazmid kópiaszáma magas, a kromoszomális DNS-sel együtt replikálódik. Amikor ez lecsökken, beindul a plazmid független replikációja, míg el nem éri a sejtenkénti 30-50 kópiát. Ezt a plazmidot **YE_p**-nek (*yeast episomal plasmid*) is nevezik. Ebből származnak a YE_p sorozat vektorok, amelyek stabil plazmidok, egyenlő mennyiségben válnak szét (szegregálódnak) a leánysejtek között.

Egy másik lehetőség arra, hogy függetlenül replikálódó élesztő vektort kapjunk, hogy egy cirkuláris DNS tartalmazza azt a kromoszomális szekvenciát, ami a replikációért felelős. Ezt az elemet **független replikálódó szekvenciának** hívják (**ARS**: *autonomously replicating sequences*). Nagy kópiaszámot eredményez, de gyakran instabil, könnyen elveszthetik a sejtek a rekombináns plazmidot mitóziskor. Az ilyen típusú élesztő vektorokat **YR_p**-nek (*yeast replicating plasmid*) nevezzük. A YE_p és YR_p sorozat tagjai a **replikálódó plazmid-alapú vektorok**, míg a **YI_p** plazmidok (*yeast integrating plasmid*) a genomba integrálódóak. Ha az ARS-t tartalmazó plazmidba beépítünk a kromoszóma centromerjéből származó CEN szakaszokat, amelyek a kromoszómák mitotikus orsóhoz való kapcsolódásáért felelősek, ekkor stabilizálni tudjuk vele a plazmidot, nem vész el osztódás során. Így kapjuk a **YAC** vektorokat, amelyek a replikációhoz és az utódsejtbe jutáshoz minden elemet tartalmaznak (**ld. 7.5. fejezet**).

Az ismertetett négy élesztő vektor típust (integrálódó, replikálódó, episzomális és élesztő mesterséges kromoszóma) élesztő expressziós vektorként használhatjuk (**ld. 13.1. ábra**)



13.1. ábra: Az élesztő vektorok típusai

A sörélesztő gazdasejtben használható expressziós vektorok pl. az Invitrogen cégtől beszerezhető pYES2 vektorok vagy a Novagen cég pYX vektorsorozata (sokszor a név is utal a vektor fajtájára, pl. pYES2 egy replikálódó „*yeast episomal*” plazmid). A *Pichia* rendszer plazmidjai általában integrálódó plazmidok, pl. pHIL-D2. Ezen a linken számos élesztőben vagy más gombában használható, különböző elven működő vektorról olvashatnak további részleteket. Expressziós kitek is beszerezhetők, amelyek az élesztő rendszerben használhatók. Ezek különböző expressziós vektorokat, némelyek élesztő törzseket, a szelekcióhoz szükséges anyagokat, szekvenáló primereket és más segédanyagokat tartalmaznak. Ráadásul nagyon pontos leírást adnak arról, hogyan kell lépésről-lépésre végrehajtani a rekombináns fehérjék termeltetését. Részletes ismertetőjük itt található.

13.1.7. Az élesztő expressziós rendszer előnyei és hátrányai

Röviden összefoglalva az élesztő expressziós rendszer előnyei:

- olcsó, könnyen kezelhető
- alkalmas fehérje szekrécióra, és a médiumból könnyen kinyerhető a termelt fehérje
- nagy a kitermelés; egy liter kultúrából 5-10 g fehérje is kinyerhető (a *Pichia* rendszerrel)
- általában jó az emlős fehérjék foldingja és a diszulfid-hidak kialakulása is megfelelő

A rendszer hátrányai:

- egyes poszttranszlációs módosulások, elsősorban a glikoziláció nem teljesen felel meg az emlősök hasonló módosulásainak; például hiperglikoziláció lép fel, és más oligoszacharid egységek is beépülhetnek a rekombináns fehérjéláncra (pl. mannóz)
- bizonyos fehérjék heterológ expressziójára nem alkalmas

13.2. Bakulovírus expressziós rendszer

A bakulovírus expressziós rendszerrel a **gazdasejtek** egy **molylepke** fajból (*Spodoptera frugiperda*) származnak, a vektor pedig módosított bakulovírus. A **bakulovírusok rovarokat fertőző** kétszájú cirkuláris **DNS-vírusok**, pl. *Autographa californica* és *Bombyx mori* nukleáris polihedrózis (NPV). Genomjuk kb. 80-100 kbp méretű cirkuláris DNS-ből áll, mely legalább 154 fehérjéjét tartalmaz, amiből 20 virális struktúrféherjét kódol. A vírusrészecskét (virion) egy polihedrin nevű fehérjéből álló mátrix burkolja be, ami biztosítja a vírus túlélését a gazdasejten kívül. A **polihedrin** burokfehérje a vírus fertőzés kései szakaszában termelődik **extrém erős promóterről**. Ezzel az erős promóterrel (pPolh) indítjuk el a célgénünk kifejeződését. A transzkripció iniciációjához kulcsfontosságú a TAAG-szekvencia (innen indul a transzkripció), illetve a pPolh promóter és a TAAG között lévő 50 bp-nyi szakasz szekvenciája is erősen befolyásolja az expressziós szintet. Erről a kései promóterről a fertőzés után 20 órával kezd elindulni a génexpresszió, és a sejtek a fertőzés utáni 72-94 óráig még nem lizálnak. A várható legnagyobb expresszió 30-50 h közé tehető, ilyenkor a sejt legaktívabb promótere a polihedrin promóter, s az általa vezérelt heterológ fehérjék mennyisége elérheti az összfehérje tartalom 1-10%-át.

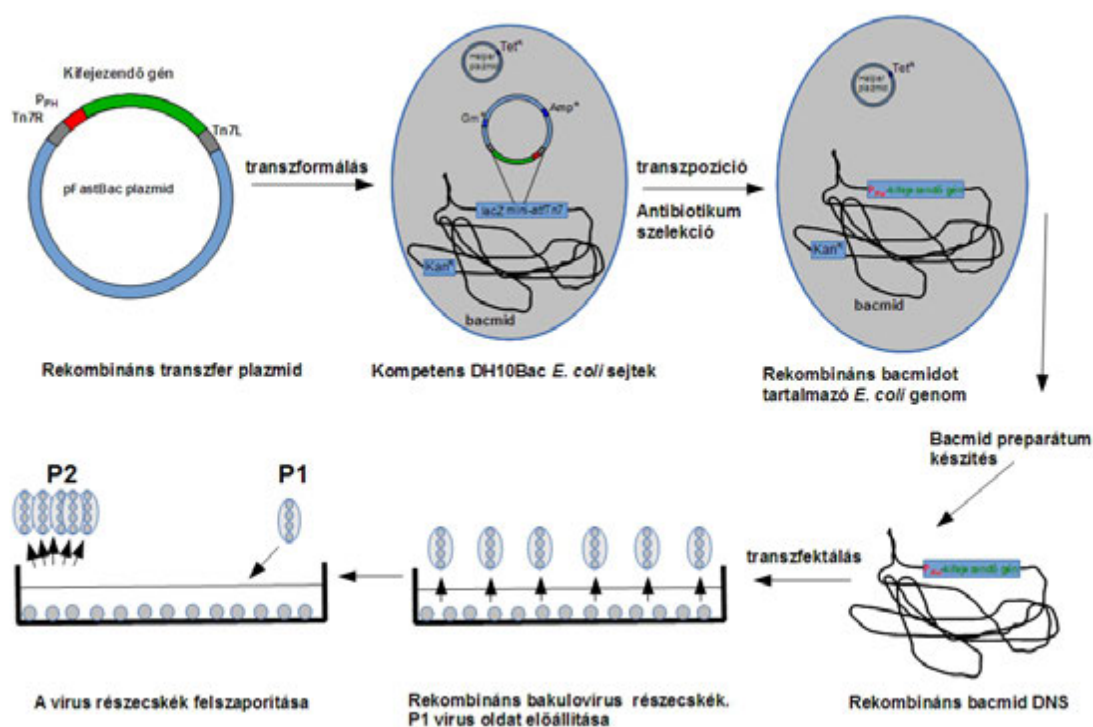
Az expresszió első lépése itt is az expressziós vektor, azaz a rekombináns vírus előállítás. A célgénünket elsőként molekuláris klónozással beviszük egy **transzfer vektorba**, amiből az expressziós kazettánk helyspecifikus transzpozícióval vagy homológ rekombinációval átkerül a vírusgenomba. Majd a célgénünket tartalmazó vírusgenommal megfertőzzük a gazdasejtet, ahol a fertőzés kései szakaszában beindul a génünk kifejeződése. Két módszert fejlesztettek ki eddig a bakulovírus expressziós rendszer hatékony alkalmazására.

13.2.1. „Bac-to-Bac” rendszer

A génünket tartalmazó **transzfer plazmiddal** (pFastBac) módosított *E. coli* törzset, ún. DH10Bac törzset transzformálunk. Ez a törzs tartalmazza a vírus genomját (Bacmid), és egy tetraciklin rezisztenciát hordozó segédplazmidot, ami egy transzpozáz enzimet kódol. Miután a DH10Bac felvette a transzfer plazmidot, a transzpozáz segítségével megtörténik a megfelelő helyek között a helyspecifikus transzpozíció, így a célgénünk a vírusgenomba kerül, elkészül a rekombináns vírus. A vírus szaporítását *E. coli*-ban is végezhetjük, a vektor tehát „shuttle” vektorként szolgál. A rekombináns bacmidot hordozó baktérium telepeket felszaporítás előtt szelektálni kell. Erre a célra antibiotikum és kék- fehér szelekciót alkalmazunk. Transzformálás után a DH10Bac sejteket X-GAL-t, IPTG-t, kanamicint, tetraciklint és gentamicint tartalmazó lemezen szélesztjük. (A transzpozíció után a transzfer vektor gentamicin rezisztencia génje átkerül a vírus genomba, ami eleve már kanamicin rezisztenciát hordozott.) A kék-fehér szelekció azon alapszik, hogy a bacmid hordozza a LacZ azon részét, ami hiányzik a baktérium genomjából (a-komplementáció, ld. 7.2.3.1 . fejezet), de a sikeres transzpozíció hatására ez a kódoló régió tönkremegy, így nem áll össze a sejtben az aktív β -galaktozidáz enzim. A rekombináns bacmidot hordozó baktérium telepek fehérek maradnak. Ezekből a telepekből bacmid-preparátumot készítünk. (Vigyázni kell, mert a bacmid-DNS a plazmidénál legalább 1 nagyságrenddel nagyobb, könnyebben sérül fizikai hatásokra.)

A rekombináns bacmiddal megfertőzzük a gazdasejtet; a gazdasejt lehet SF9, SF21, HiFi5. Először lipofekcióval transzfektáljuk a sejteket. A bacmidot felvett sejtekben idővel kialakulnak a fertőzőképes vírusok, amelyek kijutva a sejtől (bimbózással, vagy kizárással) megfertőzik az összes sejtet. Így a sejtek közel 100%-a tartalmazza és szaporítja a vírusokat. A felszaporított vírusok a transzformálást követő 3-6 napban már a médiumban vannak, ezt hívjuk **P1** vírus oldatnak. A P1-et miután centrifugálással és sterilizációval mentesítettük a sejtektől, hosszabb ideig hűtőben tárolhatjuk. Ez az oldat még kevés vírust tartalmaz, és kicsi a térfogata, ezért ebből egy több vírust tartalmazó, ún. **P2** -t készítünk. A P2 egy P1-nél több vírust tartalmazó sejtmentes médium, melyet nagyobb mennyiségű sejt P1 vírus oldattal való fertőzésével és inkubálásával kapunk. A vírus mennyiségét további, megnövelt sejtmennyiségű tenyésztéssel fertőzéssel növelhetjük, **P3, P4**, stb. A P3 oldat vírus tartalma már megfelelő a fehérje expresszióhoz. Ezzel fertőzzük a sejteket, majd 36-72 óra között „szüreteljük” (*harvesting*) a sejteket és izoláljuk az expresszált fehérjénket. Itt is érdemes először kis térfogatban végezni az expressziót, majd 36-72 óra között mintát venni és megnézni termelődik-e fehérjénk (Western-blot). Az optimalizálás során változtathatjuk a hozzáadott vírus oldat (P3) mennyiségét (0,1-10%), és a sejtek összegyűjtéséig eltelt inkubációs időt.

A Bac-to-Bac expressziós rendszer sémáját a **13.2. ábra** foglalja össze. Az Invitrogen cég Bac-to-Bac kitjei nagyon megkönnyítik a rekombináns vírusok előállítását, erről itt található bővebb információ.



13.2. ábra: A “Bac-to-bac” expressziós rendszer sémája (a P2 víruspopulációgig)

13.2.2. Közvetlen bakulovírus (Direkt Bac) rendszer

Ennél a rendszernél a **linearizált deficiens vírus** genomba a célgénünk a transfer vektorból homológ rekombinációval kerül be. Először elkészítjük a rekombináns transfer vektort, ami tartalmazza a génünk szekvenciáját, majd a linearizált vírus genommal együtt transzfektáljuk a rovarsejteket. A sejten belül létrejön a homológ rekombináció, és csak azok a sejtek tudnak szaporodni, amelyeknél megtörtént a rekombináció, így szelektálni sem kell. A reprodukcióra képes vírusok, mikor kiszabadulnak a sejtéből, tovább fertőzik a sejteket, így nő a vírusok mennyisége, amit még tovább lehet szaporítani, vagy fehérjetermeléshez lehet használni. Bővebb információ itt olvasható.

A transfer vektor szintén szaporítható prokarióta sejtben, lehet benne több vírus is promóter is, így egyszerre több polipeptidláncot is lehetséges expresszálni, pl. **BaculoGold** vektorok, vagy a PAA cég BaculoOne vektorai. Az expresszióhoz a pPolh promóter mellett egy másik kései promótert, a p10 promótert is használják, amely szintén erős promóter. A bakulovírus expressziós kazetták is tartalmazzák a transzkripcióhoz és translációhoz szükséges elemeket: transzkripció terminációs szekvenciák, poli-A szignál, Kozak szekvencia, N-terminális fúziós címke.

13.2.3. A bakulovírus rendszer előnyei és hátrányai

Az *E. coli* sejtben rosszul kifejeződő fehérjékre a következő lépés vagy az élesztő, vagy a bakulovírus rendszer. Itt humán fehérjék is kifejezhetők aktív formában. A kitermelés nem lesz soha olyan jó, mint amit el lehet érni egy jól expresszálódó fehérjére *E. coli* rendszerben, de 1-50 mg fehérje kinyerhető 1 L kultúrából. A **fehérje foldingja jó lesz** (eukarióta dajkafehérjék vannak jelen), a **szelekciós, kompartmentalizációs mechanizmusok működnek**. A **poszttranszlációs módosítások** (foszforiláció, proteolitikus módosítás, glikoziláció, acetilezés, diszulfid-hidak) az emlős rendszerekéhez hasonlóak. Még hetero-dimerizáció is elérhető. Ennek módja ko-infekció, azaz fertőzés egyszerre több vírussal, amelyek különböző célgént tartalmaznak, vagy azt is megtehetjük, hogy a rekombináns vírusba több expressziós kazettát építünk be. A sejtek könnyen fenntarthatók, és gyorsan nőnek szuszpenziós vagy letapadó kultúrában is. Tenyésztésük rázóinkubátort igényel (27-28°C), manipulálásuk pedig steril fülkét és

mikroszkópot. Mikroszkóppal követjük nyomon a sejtek állapotát; a fertőzött sejtek megnőnek és lekerekednek. A táp viszont nagyon drága, még az emlőssejtekhez használt bazális (ld. később) táphoz képest is kb. ötször drágább. Másik hátrány lehet, hogy a fertőzőképes vírus oldat előállítása munka és időigényes; szerencsés esetben is a kész transzfer vektortól a P3 vírus oldatig történő eljutásig két hét szükséges.

13.3. Emlős expressziós rendszer

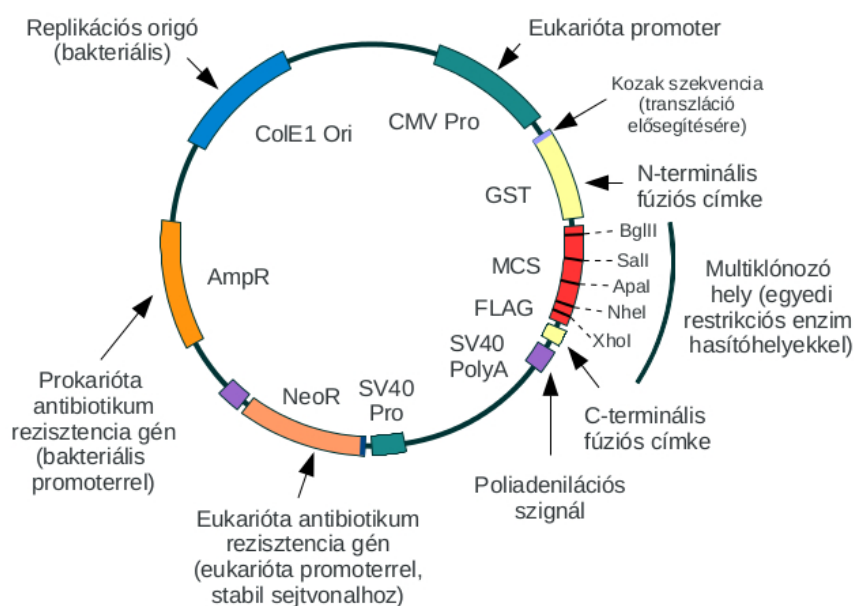
Emlős expressziós rendszer esetén emlős szövetekből származó sejtvonalakban fejeztetjük ki a kívánt génünket. Ennek a rendszernek is a gazdasejt és a vele kompatibilis expressziós vektor az alapja, ami tartalmazza a kifejezendő fehérjénket kódoló szekvenciát. A következőkben összefoglaljuk az emlős expressziós vektor lényeges elemeit, az alkalmazott sejtvonalakat, a génbevitel módját, és az expressziós rendszer előnyeit és alkalmazhatóságát

13.3.1. Emlős expressziós vektorok

Az emlős expressziós vektorok is „shuttle” vektorok; a klónozáshoz és szaporításhoz *E. coli* rendszert használunk, így a vektornak tartalmaznia kell a baktériumban való fenntartáshoz és szaporításhoz szükséges elemeket; *E. coli* replikon, antibiotikum rezisztencia gén (ld. 12.2. fejezet). Ezen kívül a transzkripcióhoz szükséges **promótert**, a mRNS szintézis megfelelő **terminációjához** és a processzázáért (**poszttranszkripciós módosítások**) felelős elemeket, a translációhoz fontos szekvenciákat. Ezeket a cisz-reguláló szekvencia elemeket ismertetjük a következőkben. A 13.3. ábrán egy általános eukarióta expressziós vektort mutatunk be.

13.3.1.1. Emlős promóterek és enhanszerek

Az eukarióta promóterek fontos eleme a TATA-box, mely az első átírandó nukleotidhoz képest (5'-irány) -30 és -25 bp-nál található. További, az 5'-irányban elhelyezkedő elemek pl. a -50-tól -100-ig elhelyezkedő GC-boxok, melyek száma és elhelyezkedése eléggé eltérő lehet génről-génre, hiányuk, vagy mutációjuk viszont bizonyos gének esetén drasztikusan befolyásolhatja a gén kifejeződését. Ezen elemek vizsgálatára fejlesztették ki a **riportergéneket**. A riportergént a vizsgálandó promóter/szabályozó elem után fuzionáltatják, majd valamilyen könnyen detektálható módszerrel nyomon követik a kifejeződését. A kifejeződés mértéke arányos a promóter/szabályozó elem működésével. A riportergének olyan fehérjét kódolnak, amelynek aktivitása mérhető (abszorbancia vagy fluoreszcencia alapon). A riportergének fehérje termékeit könnyebb nyomon követni, mint a génről keletkező mRNS-t, ami viszont közvetlenebbül ad információt a promóter aktivitásáról. Az elsők között alkalmazott riportergén a kloramfenikol-acetiltranszferáz (**CAT**) volt, amely közömbösíti (acetilálással) a sejtek számára toxikus antibiotikumot (azaz egy antibiotikum rezisztencia gén). További riportergének a **β-laktamáz-t**, **β-galaktozidáz-t**, **luciferázt**, valamint a **GFP-t** és változatait kódoló gének.



13.3. ábra: Az eukarióta expressziós vektorok általános felépítése

Általánosan elmondható, hogy az eukarióta gének promótere komplexebb, mint a prokariótáké. Kevesebb bennük a szabályszerűség - sokszor még a TATA-szekvenciák is hiányoznak -, változatosabbak, és rengeteg cisz- és transz-elem befolyásolja a működésüket. Az RNS-polimeráz II sem közvetlenül kapcsolódik a promóterhez, hanem a bazális transzkripció faktorokon keresztül. Vannak „kötelezőbb” szekvenciaelemek, vannak amelyek génről-génre változnak, sőt olyanok is, amelyek több mint 1000 bp-ra is lehetnek a struktúrgéntől, és mégis jelentősen befolyásolják a génátírást. Ezeket az elemeket hívjuk **enhanszereknek** (az *enhancer* szó jelenetése megnövel, megemel). Az enhanszerek lehetnek 5' vagy 3' (*upstream* vagy *downstream*) irányban is a gén promóterétől, sőt a gén egy intronján belül is. A távolságuk is változhat, a promótertől akár több ezer bp-ra, de akár más kromoszómán is lehetnek. Működésük lényege, hogy aktivitásuk kifejtésekor térben közel kerülnek ahhoz a promóterhez, amelynek a működését befolyásolják. Az emlős expressziós vektorok ezért a megfelelő mértékű expresszió érdekében az enhanszerek jelentősége miatt nemcsak a megfelelő promótert, hanem annak enhanszerét is tartalmazzák.

De milyen legyen a célgénünk promótere? A válasz abban rejlik, hogy mit szeretnénk vizsgálni. Amennyiben a gén kifejeződését, működését szeretnénk vizsgálni, mindenképpen a saját promóterét érdemes használni, ráadásul a saját környezetében. Ha viszont az a célunk, hogy nagyobb mennyiségben állítsuk elő a génterméket, a fehérjét, akkor nagy expressziót biztosító, erős promótert használjunk. Erre legalkalmasabbak a vírus promóterek, illetve olyan endogén promóterek, amelyek állandóan aktívak és nagy mennyiségű fehérje expressziót tesznek lehetővé. Nézzük meg az emlős expressziós rendszerekben leggyakrabban használt promótereket:

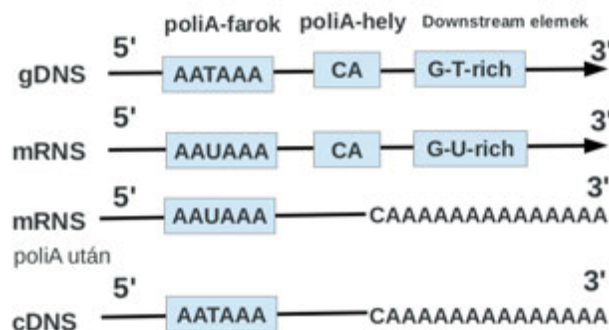
- **SV40 promóter:** SV40 (*Simian vacuolating virus 40*) egy majmokat és az embert is megfertőző DNS vírus. Ezt a polióma családba tartozó vírust *Rhesus* majom vese sejtenyészet vizsgálata közben fedezték fel 1960 körül, és kiderült róla, hogy számos különféle tumor képződését képes elindítani. A vírus a gazdasejt MHCI receptorán keresztül a VP1 glikoproteinjének segítségével jut be a sejtbe. A sejtmagban a sejt II-es típusú RNS-polimeráz enzime kötődik az 5 kb nagyságú vírusgenom SV40 promóteréhez, és elkezdí átírni a vírus korai génjeit. Ezek a gének két fehérjét kódolnak: kis antigént, és a **nagy T-antigént**. A nagy T-antigénnek több szerepe is van: kikerül antigénként a sejt plazmamembránjába; szabályozza a virális gének mRNS szintézisét; és kötődik a vírusgenom replikációs origójához, iniciálva ezzel annak replikációját. Az SV40 promóter TATA-szekvenciája, s a tőle 5' irányban lévő GC-boxok és az enhanszer szabályozza a vírus korai, majd a kései génjeinek az expresszióját. (A kései gének a korai gének komplementer szálain találhatóak ellentétes leolvasásban; ugyanannak a promóternek a szabályozása alatt, s a GC-boxokhoz kötődő fehérjék döntik el a szintézis idejét és irányát.)

- **CMV promóter** (*cytomegalovirus immediate-early promóter*): a citomegalovírusok a herpesz vírusok családjába tartoznak. Promóterük nagyon erős, nagymértékű expressziót biztosít, ezért is használjuk az expressziós vektorainkban. Állandó kifejeződést biztosít, de a tapasztalatok szerint működése eléggé szövet- és sejtfüggő.
- **CAG promóter**: Egy hibrid promóter, amelyben a csirke β -aktin promóterét a CMV korai enhanszerével fuzionáltatták. Ez a promóter a CMV promótertől eltérően nem sejtspecifikus, így bármelyik sejtvonalban használható. Emellett a sejt magas aktin szintje és CMV enhanszer aktiváló hatása miatt erőteljes expressziót eredményez.
- **UbC promóter** (*human Ubiquitin C*): Egy szintén magas szinten expresszálandó, minden sejtípusban kifejeződő promóter. Akkor használják, ha a CMV promóter nem működik.
- **PGK** (egér 3-foszfoglicerát kináz 1 promóter): minden sejtben működik a glikolitikus út, ezért ez a promóter minden sejtípusban kifejeződik, csak a kifejeződés mértékében van különbség. Azokban a sejtekben, amelyekben nagyobb a glikolízis sebessége, aktívabb a promóter is, ezért nagy expressziót eredményez. Transzgenikus egerekben és transzdukciós vektorokban használják (ld. később).
- **EF1A** (humán elongációs faktor 1 α promóter): Szintén akkor használják ezt a promótert, amikor a CMV promóter nem működik (embrionális őssejtek) vagy csökkent az aktivitása.

Az eddig felsorolt promóterek mind konstitutív, erős expressziót biztosítanak, de léteznek ezen kívül szabályozható és indukálható promóterek is.

13.3.1.2. Transzkripció termináció és poliadenilációs szignál

A prokarióta expressziós vektorokban található transzkripció terminációs szekvencia nem feltétlenül szükséges az expresszióhoz, viszont az eukarióta rendszerekben a mRNS megfelelő processzálasához ez a szekvencia esszenciális. Az eukarióta mRNS végét a gén STOP kodonja után következő 3' UTR szekvenciában (3' *untranslated region*) található elemek jelölik ki. A legfontosabb elem a **poliadenilációs szignál**, ami az átiródott RNS-en AAUAAA szekvenciának felel meg (ld. 13.4. ábra)



13.4. ábra: A poli-A farok és kialakulása

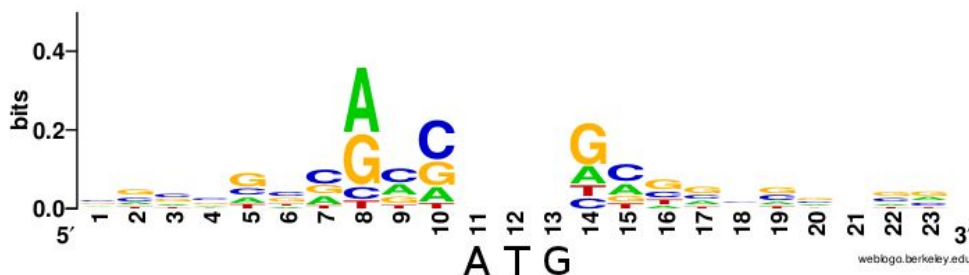
Ez a konszenzus szekvencia jelzés egy RN-áz komplex számára, hogy a szekvencia után 10-30 bázissal hasítson bele az RNS-be, majd a képződött pre-mRNS így kialakult 3'-végére a poli-A-polimeráz (PAP) szintetizál egy kb. 200 bázisnyi **poli-A farkat**. A poli-A szekvenciát nem a vektorunk, és nem az inszertünk kódolja, hanem a sejt enzimeit helyezik fel a megfelelő poliadenilációs szignál felismerése után. A poli-A szignál, de különösen a poli-A hely körüli szekvenciák is változatosak az eukariótákban. Az alkalmazott eukarióta vektorokban a leggyakoribb az **SV40 kései poliadenilációs szignált** (SVLPA) és a **Herpes simplex vírus timidin kináz** (HSV-TK) poli-A szignált használjuk, amelyek be vannak építve a vektorunkba. A poli-A farok és a hozzákötődő fehérjék stabilizálják a mRNS-t, ezek nélkül hamar degradálna. Szepe van az RNS stabilizálásán túl az érett mRNS citoplazmába való kijutásában és a translációban is. A poli-A farok idővel egyre rövidül a sejtben, így egyre csökken a mRNS stabilitása és a transláció mértéke.

13.3.1.3. *Splicing* szignál

Az eukarióta gének az 5' és 3' UTR szekvenciák mellett intronokat is tartalmaznak. Az intronok kivágása és a **pre-mRNS érése** a **sejtmagban** történik. Csak az érett mRNS jut ki a citoplazmába, ahol elkezdődhet a transzláció, a fehérje szintézise. Minden szabályozó elem, ami segíti a mRNS érését hozzájárul a fehérjének expressziójának a növekedéséhez. Fehérjetermelés céljából nem a teljes gént, hanem az intron-mentes cDNS szekvenciát szoktuk klónozni az expressziós vektorba, ami így nem tartalmaz „*splice*”-helyeket. A *splice*-helyek azonban szerepet játszanak a mRNS érésében, meggyorsítják a mRNS képződésének sebességét. Ezért számos expressziós vektor tartalmaz intron szekvenciákat 5' és 3' *splice*-helyekkel, ami megnöveli a fehérje expresszió mértékét

13.3.1.4. Transzláció iniciációs (Kozak-) szekvencia

Az eukarióta gén transzlációjánál is nagyon fontos a mRNS szekvenciája az iniciációs kodon környezetében. Ezek a szekvenciák segítik, hogy jól beilleszkedjen az iniciációs Met kodon a riboszóma kötőhelyére. Marilyn Kozak vizsgálta az eukarióta gének transzlációjának iniciációjáért felelős szekvenciákat, és azt találta, hogy vannak bizonyos helyek, ahol fontos, hogy milyen nukleotid található. Vizsgálatainak eredményeként megállapított egy konszenzus szekvenciát, amely nagy valószínűséggel kedvező az iniciációs kodon (AUG) riboszómához való kötődéséhez. Ezt a szekvenciát hívjuk **Kozak-szekvenciának**. Ezek alapján az iniciációs AUG mellett nagyon fontos, hogy a +4 helyen G, -3 helyen purin bázis (A vagy G) legyen: **(gcc)gccRccAUGG**. A 12.13. ábrán látható az a szekvencia logó, ami feltünteti, hogy az iniciációs kodon környékén az eukarióta gének milyen nukleotidokat tartalmaznak. Látható az ábrán, hogy az iniciációs kodon a DNS templátban szinte mindig ATG (egyes esetben CTG is lehet helyette), +4 helyen leginkább G található, míg a -3 helyen A vagy G nukleotid. Érdekes tehát a nagyobb expresszió érdekében betervezni a feltehetően erős konszenzus Kozak-szekvenciát. Amennyiben az iniciációs szekvenciák környéke nem elég jó, akkor a riboszóma onnan fogja indítani a transzlációt, ahol erősebb Kozak-szekvenciát talál. Ha N-terminális fúziós fehérjét fejeztetünk ki, akkor a vektor már tartalmaz a fúziós-tag kódoló szekvenciája előtt egy erős Kozak-szekvenciát (ld. 13.5. ábra), amiben megbízhatunk. Ha nincs a vektorban, akkor a gyártó ezt közli velünk, és nekünk kell betervezni.



13.5. ábra: A Kozak-szekvencia logója (a konszenzus szekvencia egyes pozíciókban megtalálható nukleotidok gyakorisága)

Az eukarióta gének monocisztronosak, de előfordul, hogy egyszerre két polipeptidláncot szeretnénk kifejeztetni egy promóter irányítása alatt. Ilyenkor policisztronos vektort kell használni, amiben a második fehérjét kódoló szekvencia előtt belső riboszóma kötőhelynek kell állnia. Ezt a szekvenciát hívjuk **IRES-nek**, azaz angolul *internal ribosome entry site*-nak. Ebben az esetben a transzláció az 5'-sapka felismerése nélkül indul a mRNS közepéről.

13.3.1.5. Szelekciós markerek

Az emlős sejtek szelektálásának legelterjedtebb módja az **antibiotikum szelekció**. A szelekciós marker gén (miként a prokarióta sejteknél is) az antibiotikum átalakítását és ezzel közömbösítését végző enzim, amely az expressziós vektorról egy erős emlős promóter irányítása alatt termelődik. A leggyakrabban használt antibiotikum a genetin vagy **G418, higromicin, zeocin és blaszticidin**. A G418, egy aminoglikozid-típusú antibiotikum, amely gátolja mind prokarióta, mind eukarióta sejtekben a fehérjeszintézist. Ebbe az antibiotikum családba tartozik még a gentamicin, neomicin és a kanamicin is. Ezek semlegesítését a neomicin gén által kódolt aminoglikozid-3'-foszfotranszferáz (APH) végzi, amely foszfátcsoportokat visz fel az antibiotikumokra, így azok nem tudnak többé kötődni a riboszómához. A higromicin-B is egy aminoglikozid típusú antibiotikum, amely szintén fehérjeszintézist gátol és használható prokarióta, eukarióta sejtekben egyaránt. Közömbösítéséhez higromicin-foszfotranszferáz

génre (hpt) van szükség. A hatásmechanizmus azonban ugyanaz, az antibiotikum foszforilálódik, így nem tud többé kötődni a riboszómához. Az antibiotikum szelekció során meg kell keresnünk azt a koncentrációt, amely elpusztítja a nem transzfektálódott sejteket, míg a transzfektált sejtek túlélnek és szaporodnak benne. Antibiotikum szelekció felhasználásával kiválogathatjuk azokat a sejteket, amelyek felvették az expressziós vektort és így rezisztenciát szereztek.

13.3.1.6. Az emlős expressziós vektorok típusai

Az emlős expressziós vektorok legtöbb eleme különböző emlős sejteket fertőző vírus genomából származik, sőt egyes típusoknál magát a rekombináns vírust használjuk génhordozónak. Ez alapján csoportosítva nézzük át a legelterjedtebb emlős vektorokat.

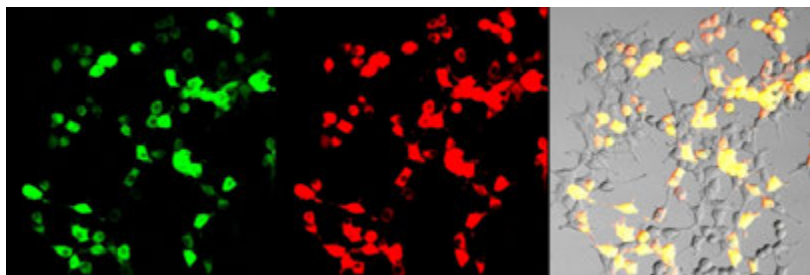
Plazmidok: általában vírus genetikai elemeket használnak bakteriális elemekkel együtt, valamint erős promótereket tartalmaznak, amelyek a legtöbb sejttypusban működnek. Általában nem integrálódnak a gazdasejt genomjába, lehet viszont replikációt biztosító origójuk (SV40, vagy EBV origó).

- pCDNA: CMV promóter
- pEGFP-N1: CMV promóter, eGFP fluoreszcens fúziós címke kerül a fehérjénk C-terminálisára.
- pEGFP-C1: CMV promóter, eGFP kerül a fehérje N-terminálisára (ld. 13.6. ábra). A kívánt génünket CMV promóterrel fejeztetjük ki, SV40 PA terminátor szekvenciával. Mivel ez egy *shuttle* vektor, prokariótában a replikációját a pBR322 replikon teszi lehetővé; a szelekciót pedig a prokarióta AmpR-promóter vezérelte NeoR rezisztencia gén terméke biztosítja (kanamicin antibiotikumot használva). A rezisztencia gén emlős sejtekben való kifejeződését az SV40 promóter és a TK_PA szignál biztosítja (ilyenkor G418 antibiotikumot használunk). Számos (kb.16) különböző restrikciós hellyel lehet a cDNS-ünket bevinni az eGFP-fúziós címke mögé. Az inszertünket EGFP_C_primerrel tudjuk *forward* irányból szekvenálni, de más szekvenálás lehetőségét is kínál a vektor.



13.6. ábra: **pEGFP-C1 expressziós vektor térképe.** A vektor térképet az Addgene Vector Database készletéből töltöttük le. Ez egy szabad vektor megosztó központ, amely lehetőséget teremt a kereskedelmi forgalomban nem lévő vektorok csere-beréjére minimális kezelési költséggel, valamint a saját vektorunk térképét is el tudjuk készíteni.

Léteznek olyan vektorok is, amelyek a GFP különböző variánsait kódolják fúziós címkeként, pl. YFP, CFP; vagy más fluoreszcens fehérjét (mCherry, DsRed). Ezek a fúziós címkek más-más hullámhosszú fényel gerjeszthetők, és máshol van az emissziójuk. Segítségükkel a fehérjénk expresszióját, az eukarióta sejten belüli lokalizációját (kolokalizációját) könnyen nyomon tudjuk követni. Egy ilyen kísérlet eredményét mutatja be a 13.7. ábra.



13.7. ábra: **Fehérje expresszió nyomonkövetése fluoreszcens fúziós címkével.** HEK293T sejteket transzfektáltunk egyszerre két különböző fluoreszcens címkét hordozó plazmiddal (kotranszekció), majd transzekció után 36 órával fluoreszcens mikroszkóppal követtük a plazmidokban kódolt fehérjék expresszióját. Az első képen látható, hogy a sejtek kifejezik CFP-címkét hordozó fehérjét (gerjesztés: 456 nm, emisszió: 486 nm), a második képen pedig a YFP-fehérjét kifejező sejtek világítanak piros színnel (gerjesztés: 510 nm, emisszió: 535 nm). A harmadik ábrán egymásra vetítettük a két képet, így látható, hogy a narancssárga sejtek felvették mindkét plazmidot. Ez a módszer alkalmas *in vivo* fehérje kolokalizáció vagy kölcsönhatás kimutatására.

Rekombináns vírusok: nagy hatékonyságú, **replikáció-inkompetens** virális részecskék. Ezek a rendszerek használhatók fehérjék expressziójára, vagy géncsendesítési kísérletekre is. A génterápiában is a rekombináns vírus rendszereket használják.

- SV40: 5 kbp nagyságú genom, kb. 5 kbp nagyságú inszertet tud befogadni, transzdukáló vektor, de speciális gazdasejt kell hozzá, pl. CV-1 vagy COS sejtek.
- Papillomavírus: 8 kbp nagyságú cirkuláris kétszálú DNS vírus. Stabil extrakromoszomális a replikációja, 20-50 kópia/sejt.
- Retrovírusok: RNS-vírusok, pl. Rouse szarkóma (RSV), egér leukémia (MLV), egér emlőtumor (MMTV) vírusok. Integrálódnak a gazdasejt genomjába, stabil transzekcióra, ezáltal permanens expresszió képesek.
- Adenovírusok: 36 kbp nagyságú genom, extrakromoszomális a replikációjuk, tranziens expresszióhoz, génterápiához használják.
- Lentivírus: RNS-vírus, stabil integráció a genomba, permanens expresszióra alkalmas.

13.3.1.7. Stabil vagy tranziens transzekció

Az expressziós vektorok másik csoportosítása azon alapul, hogy képes-e a vektor, vagy valamely része integrálódni a gazdasejt genomjába, s ezzel stabil klónokat létrehozni. Ezeket a vektorokat hívjuk (ld. élesztőnél) **integrálódó vektoroknak**. Az integrálódó vektorból általában homológ rekombinációval történik meg az expresszálandó fehérjék szekvenciájának beépülése a gazdasejt genomjába. Ennek a folyamatnak a hatásfoka nagyon kicsi, ezért mindenképpen szelektálni kell szelekciós markerek segítségével. A szelekciót hosszabb ideig fenn kell tartani, így közel 100% lesz a transzformálódott sejtek mennyisége. A **stabilan transzfektált** sejtek már egy önálló sejtvonalat alkotnak és fenntarthatók, lefagyaszthatók. Előnyük a homogenitás, a stabilitás (nem vész el a transzgén), illetve a nem túl magas expressziós szint. A nem túl magas expressziós szint sokszor előnyt jelent, mert kevésbé „szól” bele a sejt normál működésébe, és nem pusztul el a sejt. **Tranziens transzekciónál** sok fehérje termelődik, ez jelentősen befolyásolja a sejt működését, ami stresszt jelent a sejtnek és ez sokszor problémát okoz. A tranziens transzekció előnye viszont, hogy nagyon gyors, a transzekció után 12-72 óra elteltével már lehet vizsgálni a sejtet, vagy izolálni a kifejeződött fehérjét.

A transzekció hatásfokát és a fehérje kifejeződését fúziós fehérjékkel lehet nyomon követni. Ilyenkor N- vagy C-terminálisan egy enzimaktivitással (luciferáz, β -laktamáz, β -galaktozidáz) rendelkező fehérjével vagy fluoreszcens fúziós fehérjével együtt fejeztetjük ki a fehérjénket, ezáltal detektálhatjuk az expresszió mértékét, hatásfokát (a sejtek hány százaléka transzfektálódott), valamint a kifejeződő fehérje lokalizációját.

A tranziens transzekció előnye a gyorsaság és a magas expressziós szint, így kiválóan alkalmas fehérje termelésre, illetve nagyszámú különböző fehérje vizsgálatára. 3-4 nap múlva viszont a sejt vagy elpusztul a nagy expressziós szinttől, vagy annyira „kihígul” benne az expressziós vektorunk, hogy jelentősen lecsökken a transzformált sejtek

mennyisége. Ennek oka, hogy a vektorunk nem replikálódik, így mitóziskor kihígul, vagy degradálódik. Ennek kiküszöbölésére használhatunk **replikatív episzomális vektorokat**, mint pl. a pSV-vektorok. Ezek a vektorok virális replikációs origót hordoznak, és ha a sejtben kifejeződik a replikációt szabályozó fehérje is, akkor a vektorunk extrakromoszomálisan fog replikálódni. Ennek következtében kevésbé hígul ki, ráadásul antibiotikum szelekciót alkalmazva még fel is lehet dúsítani a transzfektálódott sejteket. Ez egy köztes megoldás az átmeneti és a stabil transzfekció között, azaz a sejtvonal szelektálható és fenntartható hosszabb ideig, de nem stabil. Az antibiotikum rezisztenciát, amit a szelekciónál kihasználunk, bejuttathatjuk ugyanazon a vektoron, ami a kívánt fehérje cDNS-ét hordozza, de egy másik expressziós vektoron is. Ilyenkor együtt transzfektálunk a két plazmával, azaz **ko-transzfektálunk**, s azzal a feltevessel élünk, hogy ha egy sejt felvette a szelekciós markergént tartalmazó plazmidot, akkor nagy valószínűséggel felvette a másikat is.

13.3.2. Sejtvonalak

A fehérje expresszióra használt sejtvonalakat primer sejtenyészetekből nyerik, többszörös passzálás után. A kiindulási szövet lehet **nem transzformált** felnőtt szövet, azonban az ezekből származó sejtek csak ~30 osztódást bírnak ki. Lehet normál embrionális szövet, amelyből származó sejtek ~50 alkalommal képesek osztódni. Ezeket a sejtenyészeteket tehát nem lehet korlátlan ideig fenntartani. Expressziós kultúráként használnak még daganatos sejtekből kivont **immortalizált**, vagy mesterségesen immortalizált sejtvonalakat is. A mesterségesen immortalizált sejtvonalakat vírus általi fertőzéssel lehet előállítani, ilyen például a COS1 és a COS7 sejtvonal. Ezeket **transzformált sejtvonalaknak** hívjuk, és előnyük, hogy korlátlan ideig képesek osztódni. A nem transzformált sejtvonalak is idővel spontán transzformálódnak, így korlátlan ideig fenntarthatóvá válnak. Ennek oka, hogy nincs immunrendszer, ami „kordában tartaná” őket, és gyorsan mutálódnak, szelektálódnak az *in vitro* körülmények hatására.

A leggyakrabban használt sejtvonalak:

- **CHO:** (*Chinese hamster ovary*) kínai csíkos hörcsög petefészek sejtek; monolayer kultúrában (letapadva) tarthatók, gyorsan nőnek és hatékonyan termelik a rekombináns fehérjét; az emlős sejtek „kólija”. Az ipari termelésre leggyakrabban használt emlős expressziós rendszer.
- **COS:** afrikai zöld majom veséből származó fibroblaszt-szerű sejtvonal, amelyet CV-1 sejtek SV40 vírussal való fertőzésével nyertek. (A CV-1 sejtek afrikai zöld majom vese fibroblaszt sejtek.) Gyakran használják rekombináns fehérjék előállítására. Amennyiben **SV40 vektorral** fejeztetjük ki a fehérjénket, a vektor replikációját a nagy T-antigén szabályozza, így önállóan replikálódik.
- **NIH-3T3:** egér embrionális fibroblasztból származó, letapadó (*monolayer* kultúra) immortalizált sejtvonal.
- **HeLa:** a legrégebben fenntartott humán sejtvonal, méhnyakrák sejtekből származik (**Henrietta Lack** amerikai nő tumorjából 1951-ben vették ki az eredeti sejteket). Rendkívül gyorsan osztódó letapadó sejtvonal, a rákkutatás leggyakrabban használt sejtvonala.
- **L6:** patkány vázizomból izolált sejtvonal
- **HEK293:** embrionális vesesejtekből származó immortalizált sejtvonal. Normál embrionális vesesejteket fertőztek adenovírus-sal, amelynek genomja integrálódott a sejt genomjába. Nagyon könnyen fenntartható, és transzfektálható. Előszeretettel használják sejtbiológiai kutatásokban és biotechnológiában terápiás fehérjék és vírusok előállítására. A **HEK293T** sejtekben kifejeződik az SV40 nagy T antigén, így SV40 replikációs origójú expressziós vektort használva sokáig fenntartható a vektor replikációja és így a fehérje expresszió is.

A megfelelő körülmények között tartott sejtek viszonylag gyorsan osztódnak, így könnyen változhat a geno-és így a fenotípusuk. Ügyelnünk kell arra, hogy ne változzon a genotípus a kísérleteink során (ugyanazzal a sejttel kezdjük és fejezzük be a kísérleteinket), ezért mindig fagyasszuk le a kultúrát, mikor először kézhez kapjuk a sejteket. Így bármikor visszanyúlhatunk az eredeti, ismert genotípusú kiindulási sejthez. Ehhez be kell tartani a fagyasztás és a felolvasztás szabályait, és meg kell találni a sejteknek a legmegfelelőbb fagyasztási médiumot. Általában jellemző, hogy a fagyasztást magas szérum (20-40%), és kb. 10% DMSO mellett végezzük. A fagyasztás lassan történik, míg a felolvasztás gyorsan és utána gyorsan meg kell szabadulni a DMSO-tól, mivel könnyen átjut a plazmamembránon, és toxikus hatású a sejtekre.

13.3.3. Génbevitel emlős sejtbe

Az expressziós vektor bevitele a sejtekbe biokémiai, fizikai módszerekkel és vírusfertőzéssel történhet. Ezekkel a módszerekről a 3. fejezetben foglalkoztunk részletesen. Az emlős sejtekhez használt génbeviteli lehetőségek összefoglalva:

- DEAE-dextrán-mediálta transzfekeció
- transzfekeció Ca^{2+} -foszfátos/DNA csapadékkal
- kationos-lipofekció
- elektroporálás
- transzdukció
- mikroinjektálás

Nagyon fontos a beviendő gén tisztasága. Minél tisztább és töményebb a rekombináns DNS, annál nagyobb a transzfekeció hatékonysága. Általában a szilika oszlopos tisztítás elegendő (kaphatók nagy tisztaságú, mikoplazmamentes DNS-t ígérő kitek is), de még jobb hatásfok érhető el alkoholos-lecsapás után, illetve CsCl-os gradiens centrifugálással.

13.3.4. Indukálható emlős expressziós rendszerek

Minden expressziós rendszerrel nagy előnyt jelent, ha a kívánt gén (target vagy célgén) expresszióját szabályozni, indukálni lehet. Ez különösen akkor jó, ha a termelendő fehérje toxikus, vagy erősen befolyásolja a sejt működését. Az emlős expressziós vektorok indukcióját különösen a transzgenikus állatok esetében használják ki. Ennek egyik módja, ha a génünk promóterébe olyan elemeket építünk, amelyek a megfelelő faktorok megkötése után a gén transzaktivációját idézik elő. Ezek az elemek lehetnek a gazdasajtban már eleve megtalálhatóak, vagy mesterségesen bevittek is. (bakteriális, más eukariótából származó elemek). Az alábbiakban néhány példát sorolunk fel:

· **Glükokortikoid** hormon hatására aktiválódó elem, **GRE** (*glucocorticoid response element*): A **hormon válasz elemek** vagy angolul *Hormon Response Element* (**HRE**) nagyon elterjedtek a humán genomban, több száz található belőlük és nagyon fontos folyamatokat szabályoznak. Működésük a következő: a megfelelő hormont megkötő a sejtben található receptora, ezután a receptor-hormon komplex dimerizál és transzlokálódik a sejtmagba. Itt hozzákötődik a DNS-en található megfelelő HRE konszenzus DNS-szekvenciájához, majd kölcsönhatásba lép a promóterhez kötődő más transzkripció faktorokkal és elindítja a target gén transzkripcióját. Ilyen HRE a GRE is, amelyet különböző szteroid-hormon receptorok képesek felismerni, majd hozzákötődve transzaktivációt elindítani. Indukálható expressziós vektorokban adenovírus „late” promóterben GRE szekvenciákat helyeznek el a megfelelő módon, amely a hormon, vagy származékainak a hatására a promóter után kapcsolt gén expresszióját akár 1000-szeresére is növelheti. Érthető, hogy azok a sejtek mutatnak nagyobb génexpressziót, amelyek több hormon receptort tartalmaznak. Ezért a gazdasajt egy glükokortikoid receptor túltermelő sejt (pl. CHO), amelyet dexametazonnal indukálnak a megfelelő génexpresszió érdekében.

- **Ekdizon-indukció:** Az ekdizon egy ízeltlábú „hormon”, amely nagy szerepet játszik az állat egyedfejlődésében, és ekdizon válasz elemeken („*response factor*”) keresztül képes szabályozni bizonyos gének kifejeződését. Emlőssejtek indukciója esetén be kell vinni ezeket az elemeket és külön a receptort is be kell építeni a promóterbe. Így az indukció két vektorral oldható meg:
 1. vektor: HSP promóter + ekdizon response elem + inszert
 2. vektor: CMV és RSV promóter + ekdizon receptor kódoló régió
- Bakteriális elemmel is lehet emlőssejtekben indukált génexpressziót elérni, méghozzá a jól ismert **lac-operátor** segítségével. Az indukció IPTG-vel történik, de meg kell oldani az operátor és a lac-represszor fehérje szekvenciájának a bevitelét a sejtekbe, ami általában szintén két vektor használatával oldható meg:
 1. vektor: RSV-LTR promóter + lac operátor (3x) + inszert
 2. vektor: CMV promóter + lacI
- **Tetraciklin indukció:** A tetraciklin nagyon hatásos antibiotikum Gram-pozitív, Gram-negatív baktériumok és protozoák ellen is. Legalább 1000 származékát állították elő és ismert a hatásmechanizmusa is. Könnyen bejut a prokarióta sejtekbe, ott hozzákötődik a riboszómához és tönkreteszi a kodon-antikodon kölcsönhatást. A rezisztens baktérium egy multimer antiporter fehérjével (**Tet**) kipumpálja az antibiotikumot a citoplazmából. Számos Tet fehérje létezik, többek között a Tn10 transzpozon által kódolt TetA. A TetA nagy mennyisége esetén, mivel a fehérje működése közben kationokat pumpál ki a sejtől, a membrán depolarizálódik, ami a sejt pusztulásához vezethet. Ezért a sejt erősen szabályozza a TetA expresszióját egy **TetR represszor fehérjével**. Tetraciklin hiányában a TetR erősen kötődik a tetA és tetR gének operátor szekvenciáihoz, meggátolva a fehérjék expresszióját. Tetraciklin hatására azonban a TetR leválik az operátorhelyről és beindul az expresszió. Ez utóbbi indukciós rendszert részletesen is ismertetjük (ld. 13.3.4.1.)

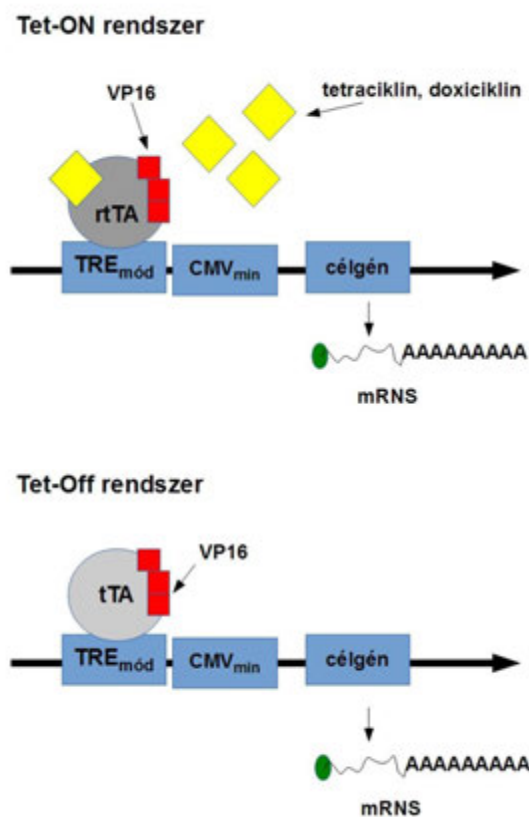
13.3.4.1. A Tet-represszoron alapuló indukciós rendszerek

A Tet operon működését kihasználva fejlesztettek ki több frappáns és széles körben használt indukálható emlős expressziós rendszert. A szabályozott rendszer előnye, hogy az elemei bakteriálisak, tehát nem interferálnak a gazdasejttel. Ezekben a rendszerekben a célgén transzkripcióját a TetR szabályozása alá helyezik úgy, hogy néhány kópiában tetO (operátor) szekvenciák találhatók a promóter és a transzkripció kezdőhelye között. Így tetraciklin (vagy származékai) nélkül nincs expresszió. Megfelelő szövetspecifikus promóter és a tetO elem kombinációjával szövetspecifikus és indukálható génexpresszió érhető el. Ez az ún. **Tet-represszor** rendszer jól működik transzgenikus növényekben, de kevésbé sikeres emlőssejtekben. Ennek oka, hogy emlőssejtekben nem elég nagy a represszor koncentrációja, ami gátolni tudná az erős promótert.

Kreatív kutatók elkészítették az előbbi rendszer a fordítottját, ahol a TetR nem represszor, hanem aktivátor lesz. Először fuzionáltatták egy transzaktivátor fehérjével, a HSV-1 (*Herpes Simplex Virus-1*) VP16-tal, így mikor a represszor köt a promóterhez, akkor a transzaktivátor fehérje génátírást indít el. Amennyiben megjelenik a sejtben az effektor (tetraciklin), akkor a **VP16-TetR transzaktivátor (tTA)** kiméra leválik a promóterről és jelentősen lecsökken a génátírás. Ez a rendszer akkor működik jól, ha stabil transzfekció során a transzaktivátor és a target gén több kópiában van jelen a genomban. A tTA rendszerben tehát akkor indul el a génátírás, ha nincs a sejtben tetraciklin, ezért ezt **Tet-Off rendszernek** is hívják. Az indukálatlan állapot fenntartásához alacsony-közepes koncentrációban kell az effektort (tetraciklin, doxiciklin) adni a sejtekhez.

Később a rendszert megfordították, s ebből alakult ki a mutáns TetR fehérjével, ahol a represszor csak az effektor jelenlétében indítja el az expressziót, annak hiányában nem. Ez a **Tet-On** rendszer sokkal érzékenyebb doxiciklinre és anhidro-tetraciklinre, mint tetraciklinre. A rendszer jól működik ebben a formában is, de még továbbfejlesztették az expressziós szint növelése és az állandóan termelődő tTa toxikus hatása miatt. A vad-típusú tTa kiméra represszor fehérjét is a tetO elem szabályozása alá tették, így az effektor nemcsak a target gén, hanem a tTA gén expresszióját is szabályozza: antibiotikum jelenlétében alacsony a szabad tTa szintje, így nem tud kötődni az operátorhelyre, nincs génátírás. Mikor nincs jelen az antibiotikum, akkor viszont hozzákötődik és elindítja önmaga és a target gén szintézisét. Ráadásul minél több tTA termelődik, az annál jobban indukálja önmaga és a target gén termelődését. Ezt a rendszert **autoregulátor Tet rendszernek** hívjuk.

Számos biotechnológiai cég kínál tetraciklin regulálta emlős expressziós rendszert. Például az Invitrogen **T-Rex** rendszere Tet-On rendszer, tehát tetraciklin és származékaira kapcsol be az expresszió. Bővebben itt lehet olvasni róla. A Tet-On és Tet-Off rendszer sémáját a **13.8. ábrán** mutatjuk be.



13.8. ábra: A Tet-On és Tet-Off rendszer sémája

13.4. Egyéb eukarióta expressziós rendszer

13.4.1. *Dictyostelium discoideum* expressziós rendszer

A *Dictyostelium discoideum* egy sejtés nyálkagomba. A talajban él, többnyire baktériumokkal táplálkozik. Laboratóriumban könnyen tartható akár folyadékkultúrában, akár szilárd táptalajon. Rövid az életciklusa, a genomja (34 Mbp haploid genom, 6 kromoszóma, 12500 fehérje gén) teljesen ismert, sok génje homológ humán génekkel, a gének szabályozása is hasonlóan történik. Génjeinek szekvenciája a **DictyBase** adatbázisban található. Nagyon jó modell organizmus genetikai, sejtbiológiai folyamatok tanulmányozására, de a rákkutatásban, immunfolyamatok megértésében is nagy hasznát veszik.

A rekombináns fehérje expressziójához itt is „shuttle” vektorral visszük be a kívánt fehérjénk cDNS-ét a gazdaszervezetbe. A vektor tartalmazza a baktériumban történő szaporodáshoz szükséges elemeket (origó, rezisztencia gén), és a *Dictyostelium* sejtben történő replikációért felelős szekvenciákat. Ennek az eukarióta élőlénynek is van természetes plazmidja, aminek a replikációs origóját használják ki az expressziós vektorok. Az élesztőhöz hasonlóan vannak **integrálódó** és **replikálódó** expressziós vektorai. A *Dictyostelium* expressziós rendszert leginkább citoskeleton és fehérjéinek tanulmányozására használják. Kiválóan alkalmas rekombináns miozin fragmentumok előállítására. Amennyiben a rekombináns fehérje sejtbiológiai szerepét vizsgálják, akkor GFP-fúziós címkével látják el, fehérje preparálás céljából pedig a szokásos affinitás címkékkel (pl. His, GST, Flag). A pDXA-3H sorozat plazmid vektorokat kimondottan fehérje expresszióra dolgozták ki.

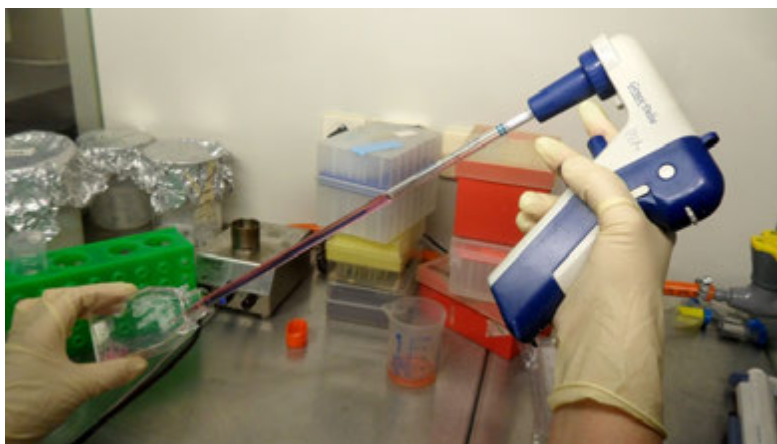
13.4.2. *Drosophila melanogaster* expressziós rendszer

A gyümölcslégy (*Drosophila melanogaster*) **Schneider 2 sejtek** (S2) 20-24 órás lárva primer sejtenyészetből származnak. Könnyen kezelhetők, nem igényelnek széndioxidot, 25-28°C-on letapadó és szuszpenziós kultúrában egyaránt könnyen szaporíthatók. Kiválóan alkalmasak tranziens és permanens heterológ fehérje expresszióra. A sejtek transzfektálása általában kationos-liposzómás (Cellfectin) vagy Ca^{2+} -foszfátos módszerrel történik. Tranziens transzfekciónál 1-4 nap múlva lehet vizsgálni a transzfektált sejteket, vagy izolálni az expresszált fehérjét. Stabil (permanens) transzfekciónál szelektálni kell a megfelelő antibiotikum rezisztens klónokat. Szelekcióra leggyakrabban higromicin és blaszticidin antibiotikumot használnak, melyek eukarióta és prokarióta sejtekre egyaránt toxikusak. A blaszticidin közömbösítésére használt rezisztencia gént (blaszticidin-S-dezamináz, *bsd*) a pCoBlast nevű plazmid, míg a higromicin rezisztencia gént (*hpt*) a pCoHygro plazmid kódolja. Mint a névből is látszik, a rezisztencia gént az expressziós vektorral való ko-transzfekcióval állítják elő. A stabil sejtvonal létrehozása 2-4 hetet vesz igénybe. A stabil klónok eltarthatóak, le lehet őket fagyasztani. Az expresszált fehérje mennyiség viszonylag magas (1-20 mg/L kultúra). Az emlős szervezetekben található glikoziláció és a megfelelő diszulfid-hidak kialakulnak. A szekréció működik mind az inszert saját, mind a vektor BiP-szignáljával. Használható akár virális fehérjék, akár humán fehérjék előállítására, transzkripciós faktorok, vagy nem *Drosophila*-homológ transzkripciós faktorok tanulmányozására

***Drosophila* expressziós vektorok:** a vektoroknak hordozniuk kell az eukarióta vektorok kötelező elemeit (promóter, poliA szignál, Kozak szekvencia, stb.), illetve mivel szintén „shuttle” vektorok, a megfelelő bakteriális elemeket is. Néhány példa: pMT/V5-His (metallotionein promótert tartalmaz, amely fémionokkal indukálható), pCoBlast és pCoHygro vektorok (rezisztencia gént kódolnak egy erős, konstitutív promóter)

13.5. Eukarióta sejtenyésztés

A sejtenyészetek manipulálását **steril körülmények** között kell végezni. Erre alkalmas a lamináris fülke, mivel nagyon kis csíraszámú légtérrel biztosít (ld. 13.9. ábra). A tenyésztéshez használt edényeknek és eszközöknek megbízhatóan sterilnek kell lenniük. Ezért előre csomagolt, általában ionizáló sugárzással sterilizált tenyésztőedényeket, szerológiai pipettahegyeket használunk. Ha erre nincs mód, pl. nagy léptékű szuszpenziós kultúránál, vagy nagyon kis térfogatok kiméréséhez pipettahegyre van szükség, akkor többszöri vízgőz-sterilizálást használunk (20 perc, 121°C, 2 atm nyomás). A laboratóriumban lévő felületeket tartjuk tisztán, fertőtlenítünk gyakran. A sterilfülkét használat előtt és után töröljük át 70%-os alkohollal (etanol, izopropanol), használjunk gumikesztyűt magunk és a sejtek védelmében is.



13.9. ábra: Munka a sterilfülkében

Alapvetően kétfajta tenyésztési mód létezik; **szuszpenziós kultúra** és **letapadó** (*monolayer*) sejtenyészet. Egyes szuszpenziós kultúrákat állandóan rázatni kell, pl. Sf9 és élesztő sejtek. Az élesztő sejtek nevelése egyszerű, hasonló a bakteriális sejtekéhez. A táptalaj összetétele egyszerű, általunk is összeállítható. A rovarsejtek és az emlőssejtek azonban általában drága, komplex táptalajt igényelnek. Jobban járunk, ha megvásároljuk őket a megfelelő gyártótól. Nekünk csak antibiotikumot, táptól függően L-glutamint, és szérumot kell beletenni használat előtt.

Az **alap vagy bazális** táptalajokat kevésbé igényes sejtekhez használjuk, pl:

- **MEM** (*Minimum Essential Medium*);
- **BME** (*Basal Medium; Eagle*),
- **DMEM** (*Dulbecco's modified MEM*, emelt aminosav/ és vitaminszint),

A **komplex médiumok** több komponenst is tartalmaznak, magasabb igényhez igazodnak, ilyen pl. az **RPMI1640**: nagyon sok sejtfeleséghez használható.

A táptalajok pufferolt (bikarbonát-HEPES rendszer, pH 7,2-7,4) körülmények között **aminosavakat** (esszenciálisakat, vagy az összeset), **szervetlen sókat** (Na^+ , K^+ , Mg^{2+} , Ca^{2+}), **nyomelemeket** (Fe, Cu, Zn szerves vegyületek formájában), **vitaminokat**, **D-glükózt** (esetleg fruktóz, galaktóz, mannóz formájában) és a megfelelő pH nyomon követésére fenolvörös sav-bázis indikátort tartalmaznak, amihez még szérumot adhatunk. A **szérum** tartalmazza a sejtek túléléséhez és szaporodásához szükséges növekedési faktorokat és hormonokat. Bizonyos sejtek beérik felnőtt marha/ló vérérszérummal, igényesebbeknek magzati szarvasmarha (borjú) szérumra van szükségük (**FBS: Fetal Bovine Serum**), sőt néhány humán sejt fenntartásához emberi vérsavóra van szükség. A vérszérum tartalmaz még zsírsavakat, lipideket, szállítófehérjéket az ionok és hidrofób molekulák számára (pl. transferrin a vas számára). A szérum (minimum 5-20%-ban) bizonyos sejteknek esszenciális, de jelenléte minden sejtnél kedvező. Végül igen fontos, hogy a médium tartalmazzon **antibiotikumot**. A fertőzések elkerülésére általában penicillint és sztreptomcint (Pen-Strep 0,5-1%-ban) adunk a médiumhoz, de használhatunk gentamicint, és (ritkábban) antimikotikumot is.

13.6. További olvasnivaló a fejezethez

Romanos MA, Scorer CA, Clare JJ (1992) Foreign Gene Expression in Yeast: a Review. *Yeast*. 8:423-488.

Hartley, JL (2006) Cloning technologies for protein expression and purification. *Current Opinion in Biotechnology* 17:359-366

Barnes, LM and Dickson, AJ (2006) Mammalian cell factories for efficient and stable protein expression. *Current Opinion in Biotechnology* 17:381-386

Structural Genomics Consortium et al. (2008) Protein production and purification. *Nature Methods* 5(2):135-146

Knetsch ML, Tsiavaliaris G, Zimmermann S, Rühl U, Manstein DJ, (2002) Expression vectors for studying cytoskeletal proteins in *Dictyostelium discoideum*. *J Muscle Res Cell Motil.* 23 (7-8): 605-11.

Kato T, Kajikawa M, Maenaka K, Park EY (2010) Silkworm expression system as a platform technology in life science. *Appl Microbiol Biotechnol.* 85(3):459-470.

14. fejezet - Célzott génmódosítás (gén-*targeting*) módszerek

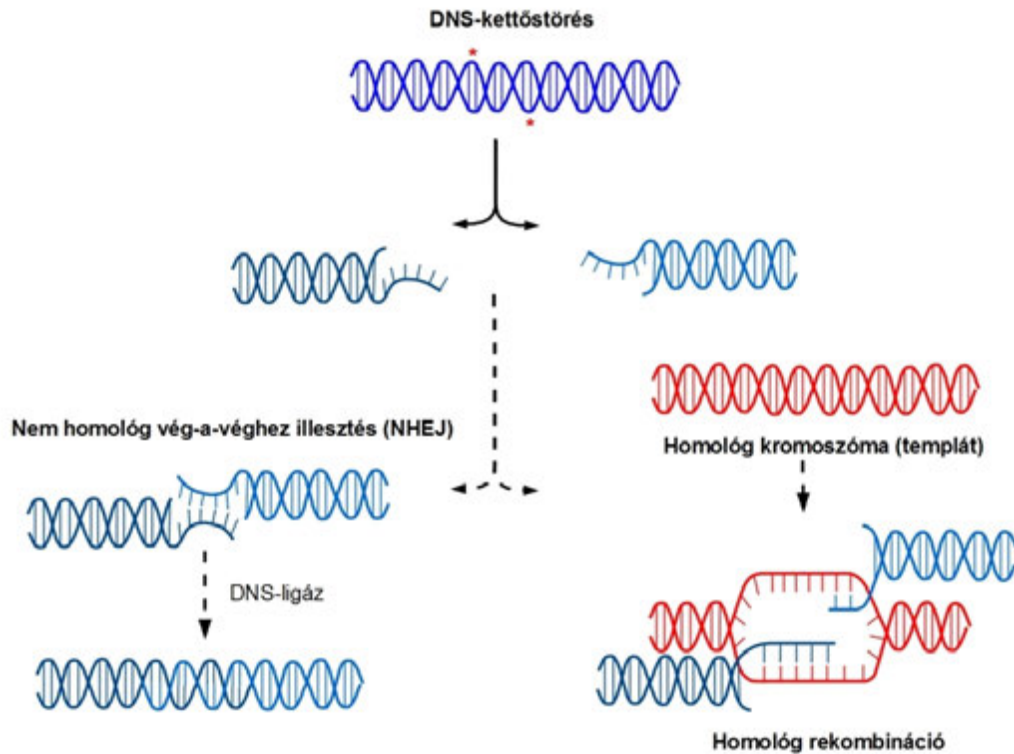
14.1. Bevezetés

Az eddigi fejezetekben megismerkedtünk a géntechnológia sokféle módszerével. Az eddig leírt módszerek azonban az élő rendszer egészét tekintve csak korlátozott alkalmazásokat tesznek lehetővé. Plazmidokkal és egyéb kívülről bevitt, genetikailag **nem stabilelemekkel** nem lehet a legtöbb organizmus saját, belső működését tartósan megváltoztatni. Ehhez olyan technológiákra van szükség, amelyek lehetővé teszik az élőlények **saját génjeinek** DNS-szintű átalakítását. A genomiális DNS módosítása viszont minden tekintetben sokkal nehezebb, mint - a viszonylag kicsi - plazmidok és mesterséges kromoszómák szabása-varrása. A legnagyobb probléma a **kromoszómák mérete**. Egy sok megabázis (millió bázispár) vagy esetleg gigabázis (milliárd bázispár) méretű DNS-szekvenciában nem tudunk kellően rövid, de mégis egyedi szekvenciákat találni. A hagyományos, restrikciós endonukleázokon alapuló technikák itt egyszerűen csődöt mondanak. Tehát a kromoszómák DNS-ének módosításához teljesen más, újszerű módszereket kell használnunk. Ezek nem 5-10 bázispár szekvenciájának azonosságára épülnek, hanem jóval hosszabb darabokra, motívumokra. A másik probléma a genomiális gének és kromoszóma-darabok **fizikai mérete**: egy-egy ekkora makromolekula manipulálása, és bejuttatása a sejtbe önmagában hatalmas kihívást jelent. Ezen problémák még ma sem teljesen megoldottak, de a genom-szintű módosító módszerek évről-évre fejlődnek, s mint látni fogjuk, az alkalmazások egész tárházát nyitják meg a molekuláris biológia számára. A "**gene-targeting**" kifejezést (magyarul **célzott génmódosításnak** fordíthatjuk) itt most tág értelemben használjuk: beleértjük exogén DNS-darabok, vagy egész gének bejuttatását a genomiális DNS-be spontán integrációs mechanizmusok révén is.

14.2. DNS hibajavítás és homológ rekombináció

Hogy megértsük a genomiális DNS-módosító eljárások működését, előbb a **DNS-hibajavítás** mechanizmusait kell áttekintenünk. Az örökítőanyagban sokféleképpen keletkezhet hiba, ennek megfelelően számos mechanizmus létezik a kijavításukra is. Az egyszerű bázis-cseréket viszonylag könnyű kijavítani. Attól függően, hogy a bázis kémiai módosulásáról, hidrolíziséről vagy egy (a replikációs során keletkező) hibás beépülésről van szó, különféle bázis vagy nukleotid kivágó (excíziós) mechanizmusok állnak rendelkezésre. Ha a DNS-dimer egyik felében a cukor-foszfát gerinc "eltörik" (hidrolizál), akkor ezt egy megfelelő DNS-ligáz enzim könnyedén helyre tudja állítani. Más a helyzet, ha néhány nukleotid távolságon belül mindkét lánc eltörik. Ilyenkor a hőmozgás következtében már néhány msec elteltével igen messze kerülnek a szabad végek egymástól. Ahhoz, hogy ilyenkor is helyre lehessen állítani a DNS-t, összetett javító enzim-rendszerekre van szükség. Először is az összetartozó végeket (hosszas kereséssel) meg kell találni, utána pedig lehetőség szerint hiba nélkül egyesíteni kell őket.

A **kettős törések** utáni javításra is többféle *in vivo* mechanizmus áll a sejt rendelkezésére. A géntechnológiai alkalmazásai miatt a két legfontosabb útvonalról részleteiben is kell beszélnünk. Az egyik a nem-homológ vég-a-véghez illesztés (NHEJ: **Non-homologous end joining**). Ilyenkor a következőről van szó: ha a két „törvéget” (eredetileg a csonttörésekre használt kifejezés) meg tudja találni az enzimkomplex és azok tökéletesen illeszkednek, akkor az NHEJ javító rendszer DNS-ligáz aktivitása révén egyesíteni fogja őket. A nem-homológ hibajavítást (ami sok baktériumban eléggé alárendelt, de emlősökben domináns szerepű) a bázispárosodások vezérlik. Amennyiben erre nincs mód (pl. tompa véggel vált szét egymástól a két DNS vég, vagy a nukleotidok kémiaiilag megváltoztak), akkor más mechanizmust kell választani. Ilyenkor elkezdődik a törvégek 5' szálának visszavágása, megfelelő exonukleázok által. Még mindig van lehetőség nem-homológ egyesülésre (MMEJ: **Microhomology-mediated end joining**), de a tökéletes javítás csak a végek egyesítésével ilyenkor már nem lehetséges. Minden remény azonban nem veszett el. Ha a sejtnek sikerül egy megfelelő, ép kettős szálú DNS templátot találnia a javításra, akkor kipótolhatja az elveszett szakaszokat, és hibátlanul helyreállíthatja a sérült molekulát. Ez az alapja a **homológ rekombinációnak** (ld. 14.1. ábra).

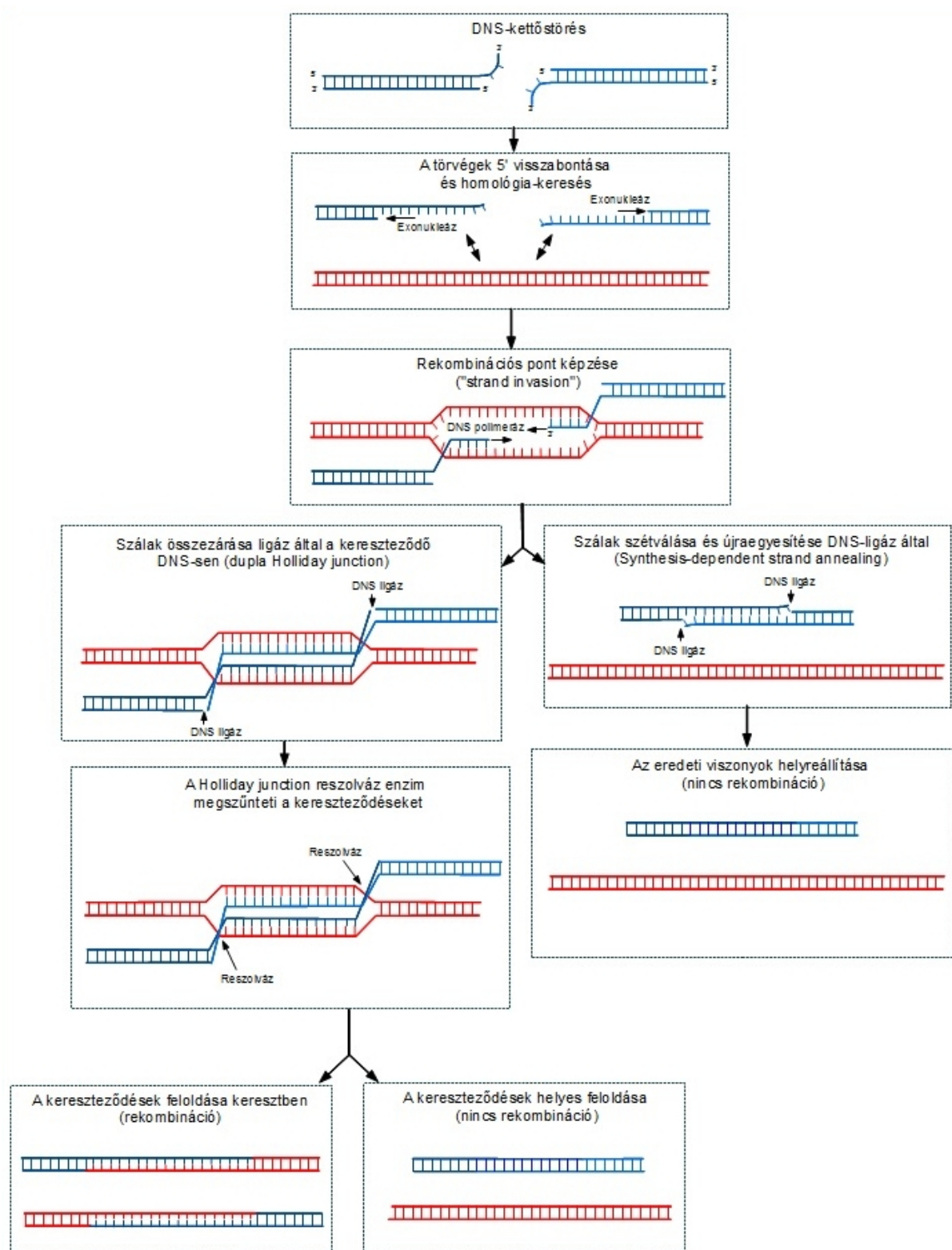


14.1. ábra: DNS kettőtörések javítása: nem homolog „vég-a-véghez” illesztés és homolog rekombináció

Ilyenkor a törvégeket felismerő fehérjekomplex előbb hosszan széttekeri a DNS-t, és lebontja az 5'- végű szálát. Az egyszálú DNS-t erre specializált fehérjék burkolják be, hogy lineáris maradjon. Ezzel egyidejűleg elkezdődik a homologia-keresés. Ha sikerül egy, a törvéghöz hasonló szekvenciájú kettős szálú DNS-t találni, akkor azt a megfelelő fehérjék széttekerik, és a törött 3' szál benyomul az ép párjának a helyére (*strand invasion*). Ez után a DNS-polimeráz segítségével megkezdődik a sérült szál kiegészítése. A rekombinációs pont mögött a négy DNS-szál egy furcsa, "vasúti kereszttváltóra" emlékeztető kereszteződést alkot (**Holliday junction**). Miután minkét szál megtalálta a homolog párját, elegendően hosszú szintézis után az összes nukleotid visszapotlódik. DNS-ligázok ezután bezárják a nukleotidszálak cukor-foszfát gerincét. Marad azonban egy kis probléma: a két kromoszóma két helyen még mindig kereszteződik egymással. Ezen pontok feloldására speciális resolváz enzim áll rendelkezésre (*Holliday junction resolvase*). A resolváz elvágja a négyes DNS-t majd átellenben egyesíti, ezáltal két külön párt hoz létre.

A kereszteződések egyébként a keletkezésük után vándorolhatnak is: ez segíthet akkor, ha csak az egyik 3'-vég találta meg a homologját. A hibrid ilyenkor szétválhat, és így a kiegészített törtvég is megkeresheti a párját (SDSA: *Synthesis dependent strand annealing*). Ebben az egy esetben nem gond az eredeti viszonyok visszaállítása. Minden egyéb esetben viszont - mivel a Holliday junction tökéletesen szimmetrikus képződmény - a resolváz enzim hasítása a két eredetileg különböző kromoszóma hibridjét hozhatja létre (**crossing over**). Ha a két kromoszóma valóban teljesen homolog, akkor ez nem probléma. Más esetekben viszont rekombinációhoz fog vezetni. Ilyen rekombinációs eseményeket használnak a **meiózis** során az eukarióta sejtek (szándékosan eltörve a kromoszómákat) hogy az apai és anyai eredetű génjeikből új kombinációkat hozzanak létre. Ugyanezt a jelenséget használják a homolog rekombinációra alapozott génmódosító módszerek is.

A homolog rekombináció vázlatos mechanizmusait a **14.2. ábra** mutatja be.



14.2. ábra: A homológ rekombinációs javítás útjai és vázlatos mechanizmusaik

A sejtekbe juttatott **linearizált rekombináns DNS-sel** sokkal nagyobb hatékonyságú rekombinációt lehet elérni, mint a cirkuláris eredetivel. A végeket ugyanis - mivel nincs rajtuk a telomerekre jellemző védő szekvencia - az eukarióta sejtek törvégekként fogják érzékelni, és ez aktiválni fogja a kettőtörést javító mechanizmusokat. Kis szerencsével a sejt a végek visszabontása után meg fogja találni a hosszú homológ részek egyikét, és kereszteződést képez a saját kromoszómájával. Általában ez a kritikus lépés: ha az egyik véget sikerült megilleszteni, akkor a

többi homológ szakaszt is nagyon könnyen megtalálja. Bár ilyenkor a "javítás" igazából lehetetlen feladat - hacsak a sejt nem kapcsolja össze a plazmidok végeit - a Holliday-junkció így is bizonyos valószínűséggel feloldásra kerül, így a sejt a saját, jó genomiális szakaszát adott eséllyel eldobhatja egy rossz templát kedvéért. Az eredmény egy genomjában **célzottan módosított** sejt lesz.

14.3. Gén-*targeting* konstrukciók előállítása

A legtöbb ma használatos célzott génmódosító eljárás a homológ rekombináción alapul. Ez azonban nem hasonlít az *in vitro* klónozáshoz. Sosem szabad elfelejteni, hogy itt egy spontán folyamatról van szó, amelynek hatásfoka még ideális esetben is meglehetősen alacsony. A leghatékonyabb ma használatos technikák is csak kb. 1% valószínűséggel eredményeznek "jó" célegyedeket, benne a célmolekulával. Ahhoz, hogy ezt ki tudjuk használni, nem elég a megfelelő DNS-darabot egyszerűen csak bejuttatni a célsejtbe. A valóságban a különböző **pozitív és negatív szelekciós** stratégiák ravasz kombinációjára van szükség. A gyakorlatban nagyon sokféle technikával lehet gén-*targeting* konstrukciókat előállítani. Itt csak néhány főbb módszert fogunk bemutatni, a teljesség igénye nélkül.

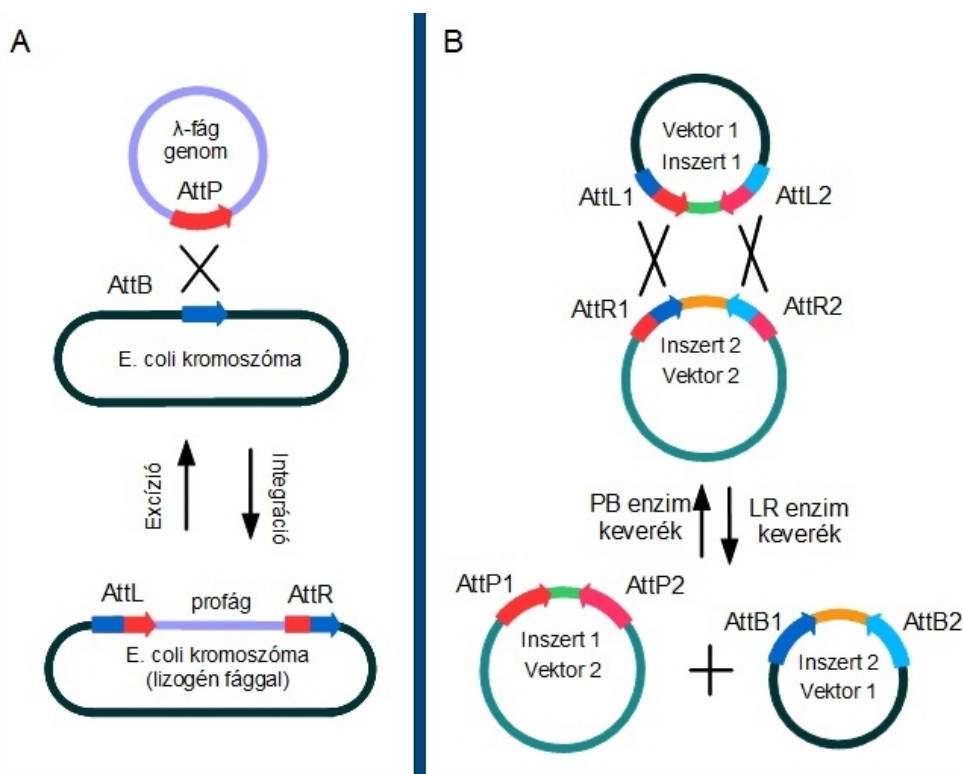
14.3.1. A *recombineering* alapjai. A Red rekombinációs rendszer

A homológ rekombinációt nemcsak eukarióta szervezetekben lehet használni: ilyen folyamatok prokariótákban is léteznek. Ezt a jelenséget fel lehet használni a plazmidok és mesterséges kromoszómák közötti DNS-darabok csereberéléséhez. A homológiát mutató DNS-darabok rekombináció révén például egyesíthetnek két különféle DNS-darabot, vektort (különösen ha csak egy homológ szakasz van). Ha a két DNS-darab több homológ szakasszal is bír, akkor a rekombináció átviheti a közöttük lévő darabot az egyikből a másikba ("bepattintás", angolul "**pop in**"). Ugyanígy, két homológ szakasz vektoron belüli párosodása a köztes darab kivágódásához, esetleg elvesztéséhez vezethet ("kipattintás", "**pop out**"). Pozitív (*pop in* esetén) vagy negatív (*pop out* esetén) szelekciós markerek beépítése a kérdéses darabba felgyorsítja a folyamatot. Ezeket a módszereket szokták "**recombineering**"-nek (*recombinogenic engineering*) nevezni. Noha a baktériumok rekombinációs képessége önmagában viszonylag alacsony, az őket megfertőző vírusok némelyike jelentősen fokozza a homológ rekombináció esélyét: olyannyira, hogy akár a transzformációval bejuttatott rövid, lineáris DNS-darabok is rögtön beépülhetnek a megcélzott plazmidba. Ezt bizonyos törzsek módosításával nagy hatékonyságú ***in vivo* klónozásra** lehet használni.

A legelterjedtebb a **Red rekombinációs rendszer**. Ennek működéséhez a λ -fág három különböző génjére van szükség (*Exo*, *Bet* és *Gam*). Az *Exo* gén egy exonukleáz kódol, ami a bejuttatott kétszálú DNS-seket 5'→3' irányban visszanyesi, ragadós végeket hozva létre. A *Bet* fehérje a szabad, egyszálú DNS-végekhez kötődik, és megvédi őket a degradációtól. Végül, a *Gam* gén terméke blokkolja a baktérium saját RecBCD javító enzimét, amely egyébként szétszedné a bejuttatott lineáris DNS-t. A rekombináció létrejöttéhez a sejteknek aktívan növekednie kell, a beépülés ugyanis a DNS-szintézis közben, közvetlenül a replikációs villában történik. Mivel a virális gének aktiválódása akadályozza a baktérium normális életciklusát, ezért hőmérséklet-szenzitív rendszert használunk. Az *E. coli* tenyészetet 32°C-on tenyésztve a (defektív) profág formájában beépült λ -fág gének teljesen ki lesznek kapcsolva, de egy rövid hősokk révén (42°C) aktiválhatóak. Ezután a baktériumokat azonnal transzformálni kell. A Red rendszert fel lehet használni ***in vivo* mutagenézisre** is: ha egyszálú DNS-t juttatunk a baktériumba (például elektroporációval), akkor az be fog épülni a születő szál helyére. A módszer akkor a leghatékonyabb, ha az oligonukleotidunkat úgy tervezzük meg, hogy a szintézis során elmaradó szállal legyen homológ. A Red rekombinációs rendszer akár plazmiddal is bevihető a célsejtbe (a pSIM plazmidcsalád révén), ezért alkalmas arra, hogy bakteriális mesterséges kromoszómákból (BAC klónokból, illetve az azokat hordozó baktériumokból) genomiális darabokat klónozzunk át egy általunk meghatározott vektorba. Így akár egér vagy emberi kromoszómák sok ezer nukleotidos darabjai is könnyen átvihetők a kívánt plazmidba. A Red rendszer bevitelét követően csupán egy feladatunk van: A megfelelő, a tervezett végekkel homológ kicsi szakaszokat kell beépítenünk a vektorunkba. De a célplazmidot a bejuttatás előtt linearizálni is kell, úgy, hogy a klónozni kívánt rész végeivel homológ rövid darabkák a vektor két szabad végére kerüljenek. Ezért a legkényelmesebb megoldás, ha a teljes vektor PCR termékét transzformáljuk, amit megfelelő (a célgénnel homológ) túlnyúló végekkel láttunk el, a PCR primerek ügyes megválasztásával. A rekombináció ezek után spontán ki fogja szedni a kívánt darabot a BAC klónból, és beépíti a vektorba, bezárva a plazmidot. Ugyanígy, a kívülről bevitt, PCR-rel létrehozott inszertek is egy lépcsőben (mindenféle restriktációs emésztés nélkül) "bepattinthatóak" a kívánt vektorba, ha a tervezett beépülési helyel megfelelő homológia van mindkét végén. Ezen előnyök miatt a Red rendszer a *targeting* vektorok építésénél ma az egyik leggyakrabban használt technika.

14.3.2. Gateway klónozás

A targeting vektorok építésénél gyakori probléma, hogy vektor-alapvázat kell cserélni a klónozás során. Ha itt is el szeretnénk kerülni a restriktációs emésztéseket, erre a leggyorsabb és leghatékonyabb technika az ún. "**Gateway**" klónozás. Ezt az Invitrogen cég által kifejlesztett módszert, mint hasításmentes klónozást röviden már említettük a 4.2.6. fejezetben. Itt is a λ -bakteriofág életciklusát használjuk ki: mikor lizogén fázisba lép, a fág integrációját a gazdakromoszómába egy specifikus enzim (integráz, *Int*) irányítja, amely a fág cirkuláris DNS-sét egy adott helyen (**AttP**) felnyitja és egyesíti a bakteriális genom egy specifikus helyével (**AttB**). Az AttP és AttB helyek összekapcsolódásával hibrid szekvenciák keletkeznek: **AttL** ("bal vég") és **AttR** ("jobb vég"). A reakció maga egyirányú, de megfordítható: méghozzá a fág excíziós enzime (*Xis*) által. Ez az enzim a profág két végén található AttL és AttR szekvenciákat ismeri fel, szétvágja a DNS-t és keresztsben újraegyesíti. Ezzel kivágja a fág-genomot és cirkularizálja azt, helyreállítva az eredeti AttB és AttP helyeket (ld. 14.3. ábra).



14.3. ábra: A λ -fág természetes lizogén életciklusa (A) és a Gateway klónozás sémája

A reakció mechanizmusa hasonlít a restriktációs klónozásra: a hasításkor itt is "ragadós" végek keletkeznek, amelyek a bázis-komplementaritás segítségével fognak újra egyesülni. Az AttB helyek általában rövidek, a beépülés során az AttP hely hosszú, nem szimmetrikus 5'- és 3'-végei megoszlanak az AttL és AttR között. Fontos észrevennünk, hogy egyik felismerő hely sem teljesen azonos a másikkal, mivel nem palindrom szekvenciákról, és csak részben komplementer párokról van szó. Viszont az enzimek érzékenyek a felismerő helyek homológia viszonyaira: ha az összetartozó pár egyikében például az AttP-ben) a homológ szakaszt kicseréljük, akkor a reakció csak akkor fog működni, ha ugyanezt a cserét az AttB-ben is végrehajtjuk. Az egész technológia ezen az észrevételen alapszik: a megfelelő, kompatibilis párokat számozással szokták jelölni. Ha tehát a plazmidunk tartalmaz két darab, megfelelően orientált AttL helyet (AttL1 és AttL2), amelyek egy "picikét" különböztetnek egymástól (a komplementer szakaszuk más), de pontosan megfelelnek egy másik vektor AttR helyeinek (AttR1 és AttR2), akkor az összetartozó helyek páronkénti egyesítésével a köztes darabok átkerülnek az egyik vektorból a másikba. A reakció *in vitro* kivitelezhető: ehhez csak a megfelelő **LRenzim**keverékre van csak szükség (a valóságban nemcsak a *Xis* enzim kell, hanem az *E. coli* saját **IHS-Integration Host Factor**-fehérjéjére, sőt az *Int* integrázra is). A végtermék vektorokban az inszert két végén AttP1, és AttP2, illetve AttB1 és AttB2 helyek keletkeznek. Természetesen, ha a másik fajta, **BPenzim** keveréket használunk a végterméken (ami csak az *IHF* és az *Int* fehérjét tartalmazza, *Xis* nélkül), akkor a reakció iránya megfordul, és visszkapjuk az eredeti konstrukcióinkat. (ld. 14.3. ábra) Nem szabad elfelejteni, hogy az AttL, AttR, AttB és AttP végek mind-mind érzékenyek az irányításra: ha az egyik véget véletlenül fordítva építjük

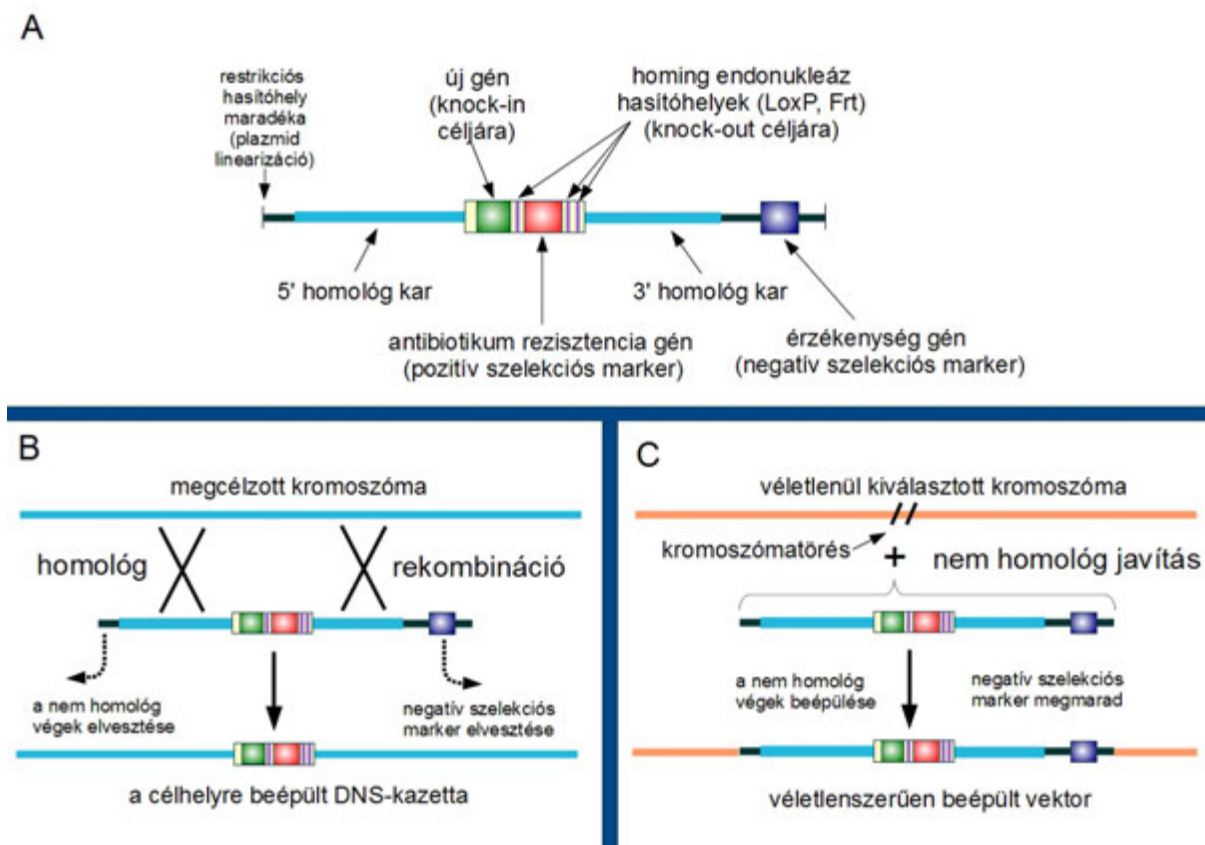
be, akkor a két vektor fog egyesülni egymással, ahelyett hogy az inszertet vinnénk át - ezért nagyon figyeljünk oda a tervezésnél!

A *Gateway* klónozás különösen akkor előnyös, ha az inszertünket később nagyszámú, különféle vektorba szeretnénk átvenni. Ilyenkor az inszertet előbb egy alapvektorba (*entry vector*) építjük be, amely tartalmazza a megfelelő AttL1 és AttL2 végeket. A bevitel történhet magával a BP reakcióval is (**ligálás-független klónozás, ld. 4.2.5. fejezet**): ilyenkor a PCR-termék túlnyúló végei tartalmazzák az AttB1 és AttB2 helyeket, a plazmid pedig az AttP1 és AttP2 célszekvenciákat. Az inszertet utána nagyszámú különböző, AttR1 és AttR2-tartalmú célvektorba lehet átvenni - például a fehérjetermelés optimalizálása céljából. Az AttP, AttL, AttR helyeket tartalmazó "gyári" vektorokat úgy alkották meg, hogy a két rekombinációs hely között egy speciális negatív szelekciós marker legyen: a *ccdB* gén (F-plazmid eredetű DNS-giráz inhibitor) - ha a helyén marad - akadályozni fogja az átlagos *E. coli* sejtek növekedését, így a csak a kicserélődött inszertet tartalmazó klónokat fogjuk visszakapni. A különféle AttR és AttL párok összetettebb szubklónozást is lehetővé tesznek: trükkös felhasználásukkal akár kis inszertek sorozatából is építhetünk egy nagy inszertet. A *Gateway* klónozás a *targeting* vektorok esetében a negatív rezisztencia-gének beépítésénél különösen hasznos: ezeket a végső célvektor tartalmazza, és ide lehet az összes különböző célzó (*targeting*) konstrukció magját berakni, mindenféle **restriktions emésztés nélkül**, LR klónozással.

14.3.3. *Targeting* vektorok tervezése

Az összes szükséges főbb módszer ismeretében most már abban a helyzetben vagyunk, hogy akár meg is tervezhetjük a saját *targeting* vektorunkat. Elsőként a következő feltételt érdemes szem előtt tartani: az eukarióta rendszerek homológ rekombinációs módosításához nagyon hosszú homológ szakaszok kellenek. Akár génkiütést tervezünk, akár gént akarunk célzottan bevinni egy organizmusba, a konstrukciónk két homológ végének külön-külön legalább 1,5-2 kbp méretűnek kell lennie. Bár a két vég nem kell hogy egyforma hosszú legyen, jó ha együttes hosszuk legalább ~6 kbp. Egy másik fontos feladat a szelekciós markerek megválasztása. A tervezett konstrukció belsejébe (a homológ "karok" közé) legalább egy olyan rezisztencia-gén beépítése szükséges, amely a célszerkezetben is aktív: ez lesz a **pozitív szelekciós marker**ünk.

Eukarióta sejtek szelekciójára természetesen nem használhatunk olyan vegyületeket, amelyek csak a baktériumokra hatnak (pl. ampicillin): **neomicin, blaszticidin, puromicin vagy zeocin** antibiotikumokat szoktunk alkalmazni. Ezek többsége mind prokariótákban, mind eukariótákban hatékony: tehát lehetővé teszi, hogy olyan rezisztencia-génekkel szelektáljunk, amelyeket úgy terveztünk meg (dupla promóterrel, Shine-Dalgarno és Kozak szekvenciákkal), hogy mind a klónozásra használt *E. coli*-ban, mind pl. egér sejtekben expresszáldjanak. A pozitív szelekciós marker azonban önmagában nem elegendő. A vektorból eredő, a homológ karokon túlnyúló végekre kell legalább egy **negatív szelekciós marker**, hogy a nagyon gyakori nem-homológ rekombinációs eseményeket (NHEJ) megpróbáljuk kiszűrni. A homológ rekombináció a rezisztenciát hordozó kazetta (kazetták) két végén lévő hosszú homológ karokon fog bekövetkezni, így a negatív szelekciós marker ("érzékenységi gén") ilyenkor (és csak ilyenkor) kiesik (**ld. 14.4. ábra**).



14.4. ábra: Gén targeting vektorok vázlatos felépítése és a kromoszómába való beépülés lehetséges módjai: homológ rekombináció (A) és random beépülés (B)

A gyakrabban használt érzékenységi gének közé tartozik a **diftéria toxin A-alegység** (növényekben is használható), vagy a **citozin-dezamináz** enzim. Emlős rendszerekben leginkább a **herpeszvírustimidilát-kináz (TK)** gént használják, ez ganciklovir jelenlétében a sejtek pusztulásához vezet. Mivel az egyoldali homológ beépülés nagyon ritka folyamat, ezért a legtöbbször elegendő, ha csak az egyik nem homológ karon található egy negatív szelekciós marker.

Végül, de nem utolsósorban a génmódosítást **ellenőrző** tesztekrol (*assay*) kell gondoskodnunk. Ezekkel tudjuk ugyanis bebizonyítani, hogy a targeting vektorunk megfelelően célba talált-e. A legtöbb vektor-építési lépés ellenőrzésére a PCR is megfelel, a célszerkezetben azonban praktikusabb **Southern-blot** eljárást használni. Egy jó stratégia lehet például a következő: a próbáinkat tervezzük úgy a célszerkezet kromoszómájára, hogy éppen kívül essenek a targeting vektorban használt hosszú karokon. Válasszunk olyan restrikciós enzimet, ami belehasít az általunk bevitt kazetába, de az eredeti genomiális szakaszba nem. Az emésztett kromoszómális DNS-hez így a helyes és rossz beépülés esetén máshol fog hibridizálni a próbánk. Mivel a vektor építése is meglehetősen bonyolult, ezért mindenképpen ellenőrizzük a konstrukciónkat a célszerkezetbe juttatás előtt! Csak teljesen szekvenált vektorral ajánlatos dolgozni, és a homológ karok, valamint a beépített kazetták hosszát és orientációját is mindig ellenőrizni kell, megfelelően tervezett Southern-blot próbák, restrikciós térképezés és/vagy PCR segítségével.

14.3.4. A gén-targeting konstrukciók elkészítése

Egy targeting vektor építése nem egyszerű feladat. Ehhez szükségünk van a célszerkezet genomiális DNS-sének megfelelő szakaszára. Egy megfelelő, genomiális DNS-t tartalmazó mesterséges baktérium-kromoszóma (BAC klón) megvásárlása ilyenkor nagy segítség lehet. Ha a hagyományos, **PCR-alapú** eljárást választjuk, akkor keresnünk kell olyan restrikciós helyeket, amelyek nincsenek jelen a megcélzott genomiális szakaszon. Ezután három külön PCR-rel felszaporítjuk a két homológ karnak megfelelő szakaszt, továbbá a közéjük beépítendő (szelekciós és módosító) kazettát. A restrikciós emésztések után (amely helyeket a primerekkel vittük be) egy **hármass ligálást** kell végrehajtanunk (három inszert egy plazmidba)! Természetesen haladhatunk lépésenként is, de akkor nincs

mód sehol a szelekcióra, szűrni kell a klónokat; ezért ez az eljárás meglehetősen fáradságos és lassú. Ha azonban sikerült a ligálás, akkor a plazmid elkészült, mivel tartalmazza a negatív szelekciós markert és a linearizációhoz szükséges extra hasítóhelyet. A hagyományos módszer gyengesége a restrikciós hasítóhelyekben rejlik: nem minden esetben van mód ugyanis ezek alkalmas megválasztására. Továbbá a nem-kódoló genomiális szekvenciában polimorfizmusok lehetnek, amelyek előre nem látható módon hasítóhelyeket hozhatnak létre. Szintén fontos hogy a PCR után minden klónt elejétől-végéig ellenőrizni kell (szekvenálással), hogy ne vigyünk be a rekombinációval nem kívánt mutációkat.

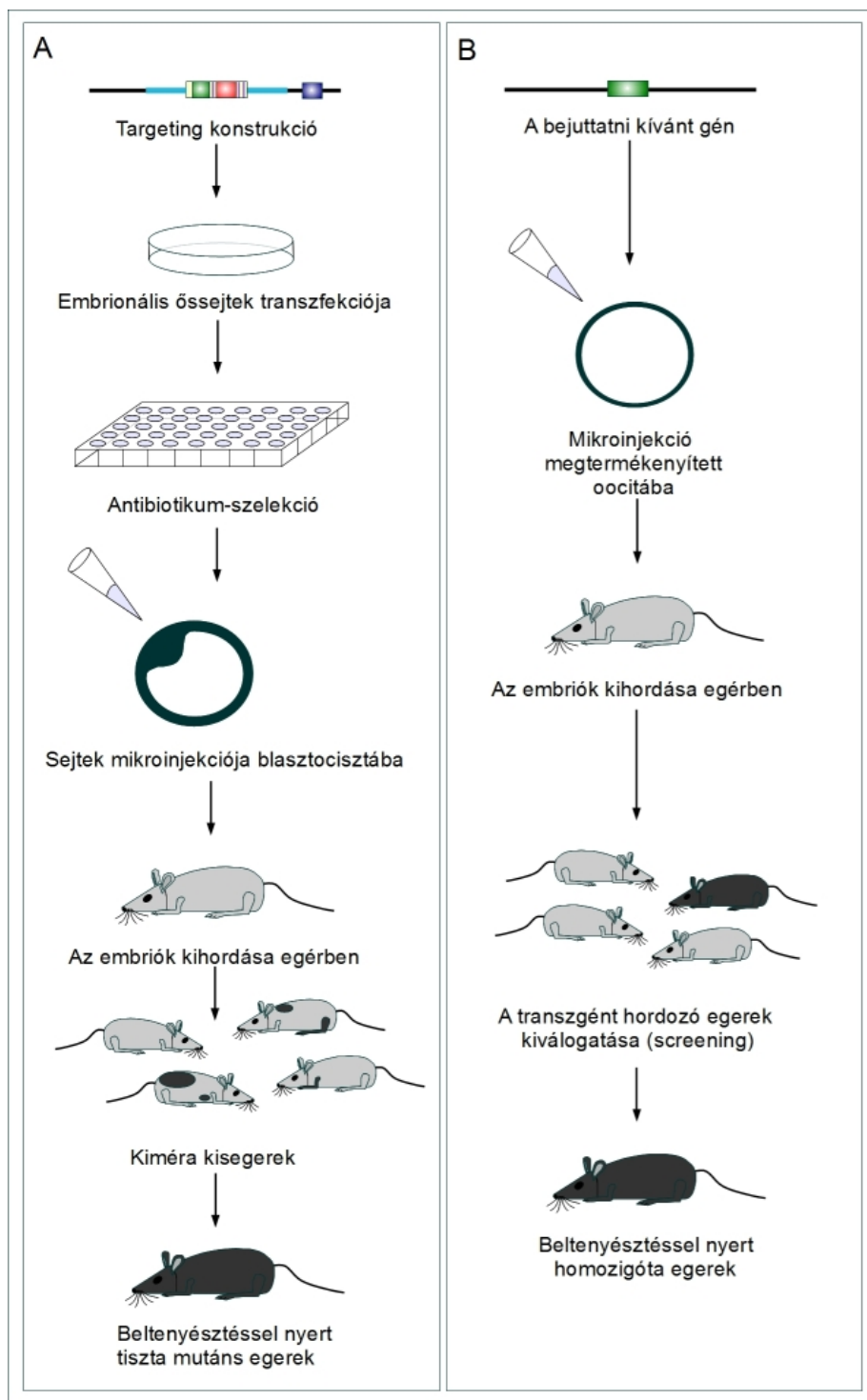
A modernebb, **recombineering-alapú** eljárások ezen hátrányokat jórészt kiküszöbölik. Ilyenkor a Red rekombinációt használjuk a konstrukció megépítéséhez. A BAC-klónt tartalmazó baktériumokat a pSIM plazmiddal transzformálva bevisszük a **Red rendszert**. A módosító kazetták beépítése és a kiszemelt vektorba való átemelés egyaránt ***in vivo*** történik, a hősokkolt baktériumok **PCR termékekkel** való elektroporálása után. A kazetták beépítését egyszerűen antibiotikum-szelekcióval ellenőrizhetjük. Ezután a vektoron is PCR-t hajtunk végre - a tervezett homológ karkok két szélével röviden azonos, kb. 40-50 bázispáros túlnyúló véggel - és ezzel transzformálunk. Így tudjuk kiemelni a BAC klónból a kívánt szakaszt. Ennél az egy lépésnél nem tudunk szelektálni: ezért plazmidot kell preparálnunk és ellenőrzés után újra transzformálnunk. Ezt a lépést ki lehet váltani köztes vektor használatával (amely pl. lehetővé teszi a kék-fehér szelekciót az átemeléskor). Ebben az esetben megcserélhetjük a klónozási sorrendet, és a kiemeléssel kezdhethetünk. Ha az összes kívánt elemünk a helyére került, akkor következhet a **vektor-alapváz cseréje**, így a rendszerbe belekerül a negatív szelekciós marker, és egy vagy több megfelelő, ritka restrikciós hasítóhely is. Ha az első lépésben olyan vektort használtunk, amely *Gateway*-kompatibilis, pl. AttL végeket tartalmaz a klónozó hely két végén, akkor ezt a végső lépést nagyon egyszerű lesz végrehajtani: az inszertünk a végleges vektor AttR végeivel fog rekombinálódni egy *in vitro* LS klónozási reakció után. A célszervezetbe való bejuttatás előtt azonban restrikciós emésztéssel linearizálni kell a vektort.

14.4. A *targeting*-konstrukció bejuttatása a célszervezetbe

14.4.1. Sejtek és sejtvonalak transzfektálása. Emlős embriók manipulációja

Ha a génmódosítást állati sejtenyészetben kívánjuk végrehajtani, akkor a DNS bejuttatása viszonylag egyszerű feladat. Akár **embrionális őssejteket** (ESC: *embryonic stem cell*), akár daganatos **sejtvonalakat** kívánunk megcélózni, a protokoll minden esetben nagyon hasonló. Ugyanúgy kell eljárunk, mint a plazmid-alapú fehérje expressziókor (ld. 4.3.2. fejezet). A sejteket lehetőségünk van **elektroporálni** is, de ennél gyakrabban használjuk a kémiai ágenseken alapuló transzfekciós módszereket. Ilyenkor vagy **kationos lipideket** vagy kationos, elágazó szerves polimereket (**dendrimereket**) keverünk a linearizált plazmidhoz, és a sejtekhez. Ezek komplexet képeznek a DNS-sel, de egyúttal destabilizálják a sejtmembránt is, így a plazmid-reagens komplex be tud jutni a membránon keletkező lyukakon.

A transzfekció után antibiotikum (pl. a neomicin-analóg G418) használatával szelektálnunk kell a transzfektált sejteket. A be nem épült plazmidok nem szaporodnak, így az osztódások során előbb-utóbb "kihígnak" a tenyészetből: csak a kromoszómába beépült rezisztencia-gén fog stabilan (hetek-hónapok múlva is) megmaradni. Mivel a véletlen integráció (NHEJ) lesz a beépülés leggyakoribb módja, a pozitív mellett negatív szelekcióra is szükségünk lesz: ezt állati sejteken leginkább a médiumhoz adott ganciklovirrel szokták végezni (a vektorok timidilát-kináz gént tartalmaznak). Csak az a sejtpopuláció esélyes arra, hogy hordozza a transzgenikus kazettát a genom megfelelő helyén, amely túlélte a dupla szelekciót. Ha **knock-out** vagy **knock-in** egeret szeretnénk létrehozni, a következő lépésben az **embrionális őssejteket** 3,5 napos blasztociszta stádiumú **embriókba** kell injektálnunk (mikroinjektorral). Az embriókat általában *in vitro* fertilizációs eljárással állítjuk elő, de hogy a folyamat sikeres legyen, a beültetésről is gondoskodnunk kell. Steril hímmel való pároztatással a befogadó nőtény egér méhét a megfelelő fázisba hozzuk (álderhesség).



14.5. ábra: **Célzott génmódosítás** (A) és célzás nélküli génmódosítás létrehozása egérben (B). Az utóbbi eljárással transzgenikus egeret hoztunk létre.

A kezelt embriókat a befogadó, ártérhes egér uterusába fecskendezve valódi terhességet hozunk létre. A születő kisegerek genetikailag **kimérák** lesznek, amelyek szerencsés esetben a csírvonalukban is hordozzák az eredeti, módosított sejtek utódait. Az egerek pároztatásával, majd (szükség esetén) beltenyésztésével hozhatóak létre a tisztán transzgenikus állatok (az állatokat természetesen genotipizálni kell, pl. Southern-blottal vagy PCR-rel). Az eljárás sémáját a **14.5. ábra** foglalja össze.

14.4.2. Stabil transzfekció. Transzgének célzás nélküli bevitele

Amennyiben a transzgént célzás nélkül szeretnénk beépíteni a célszervezet genomjába, a fentieknél jóval egyszerűbb dolgunk lesz. Ilyenkor nincs szüksége sem homológ végekre, sem negatív szelekcióra: egyszerűen a cél-kazettával (ill. az azt hordozó plazmiddal) transzfektálunk, amely tartalmazza a bevinni kívánt gént és egy megfelelő rezisztencia-gént. Az antibiotikum-szelekció ki fogja szűrni azokat a klónokat, ahol (főleg NHEJ révén) a genomialis DNS-be beépült a kazettánk. Sejtvonalak esetén ezt hívják **stabil transzfekciónak**: ennek főleg a gyógyszeripari biotechnológiában van nagy jelentősége (ld. a következő fejezetet). A leggyakrabban használt CHO (*Chinese hamster ovary*) sejtekből például hiányzik a **dihidrofolat-reduktáz** gén. Ha ezt a vektorunkkal beviszük, akkor **metotrexát** jelenlétében csak azok a sejtek lesznek képesek tartósan növekedni, amelyek genomjába integrálódott a konstrukciónk. Így elkerülhető a drága antibiotikumok használata is. Mivel a nem-homológ rekombináció emlősökben nagy hatékonyságú folyamat, ez könnyűvé teszi a **célzás nélküli transzgén** egerek létrehozását. Ilyenkor nincs is szükség az előzetes szelekcióra: A vektor DNS-t közvetlenül az *in vitro* megtermékenyített petesejtbe juttatjuk be (mikroinjektorral). A beültetés itt is a megfelelő átherhes egérbe történik, s a megszületett kisegereket alkalmas markerekkel genotipizálva lehet megtalálni a sikeres transzformánsokat.

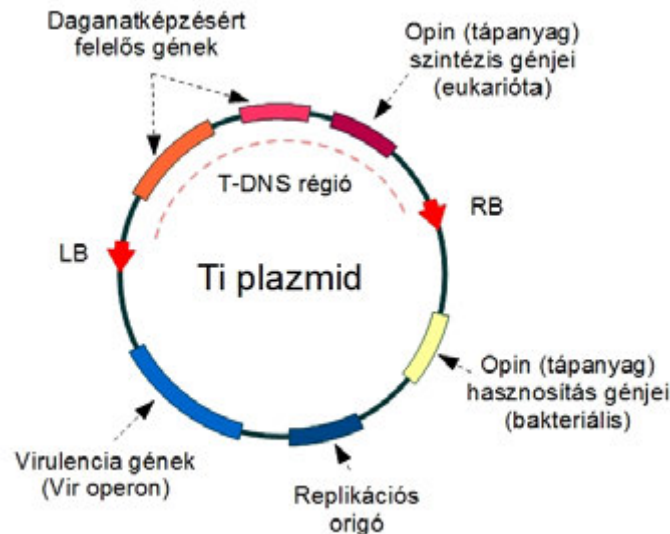
14.4.3. Növényi és gomba sejtek transzformálása. Homológ rekombináció növényekben

Az állati sejtekkel ellentétben, amelyek gyakran "csupaszak", a többi eukarióta élőlény membránját kívülről vastag **sejtfal** burkolja. Ez a DNS bejuttatása szempontjából is komoly akadályt jelent. Így növények és gombák genetikai módosításához a DNS-t nem lehet önmagában csak az eddig említett kémiai módszerekkel bevinni. A sejteket természetesen elektroporálhatjuk, sőt, a (később említésre kerülő) génpuskát is bevethetjük rajtuk, de van ennél egyszerűbb megoldás is: el kell távolítani a sejtfalat. **Gombáknál** ezt hívjuk **szferoplaszt módszernek**. Ilyenkor enzimatisz emésztéssel, **kيتينázzal** szabadulunk meg részben a sejtfaltól; a keletkező sejteket nagy ozmolaritású cukoroldatban (1 M szorbitol) kell reszuszpendálni, és nagyon óvatosan szabad csak kezelni (spontán kilyukadhat a membránjuk). Élesztőgomba sejtek (pl. *Saccharomyces cerevisiae*, *Pischia pastoris*) sikerrel transzformálhatóak a sejtfal leemésztése nélkül is, még hozzá lítium-acetát, nagy molekulatömegű polietilén-glikol (PEG 4000) és hő sokk együttes alkalmazásával (**lítium-módszer**). **Növényektranszformálásakor** szintén előbb a sejtfalat távolítjuk el: **celluláz** és **pektináz** együttes alkalmazásával. A (például levelek szét darabolása után) keletkező, emésztett sejteket itt is megfelelő ozmolaritású oldatban (pl. 0,6 M mannitol vagy 0,25 M CaCl₂) kell tartani. A gombákhoz hasonlóan, a lecsupaszított növényi sejtek (pl. *Arabidopsis* protoplasztok) könnyen transzformálhatók csupán a plazmid DNS és polietilén-glikol hozzáadásával (pl. 25% PEG 6000) is. A növényi sejtek azután megfelelő körülmények mellett visszaépítik a sejtfalat, osztódásnak indulnak (**protoplaszt regeneráció**), majd teljes növényi embrióvá alakulhatnak át (**szomatikus embriogenezis**). A regeneráció közben a sejteket a bevitt kazetta rezisztenciagénjével szelektálni lehet, de egyéb szűrés is megvalósítható (pl. GFP gén bevitelével). Maga a protoplaszt regeneráció alacsony hatékonyságú folyamat (nagyjából 10000, vagy több sejtől lesz egy teljes növény). A homológ rekombinációs megközelítés sikerrel használható pl. mohák (*Physcomitrella patens*) genetikai módosítására, de a magvas növényekben ezek a módszerek sajnos csak alacsony rekombinációs hatékonyságot tesznek lehetővé: ma is van igény a célzást elősegítő új módszerekre.

14.4.4. Növények genetikai módosítása Ti plazmid segítségével

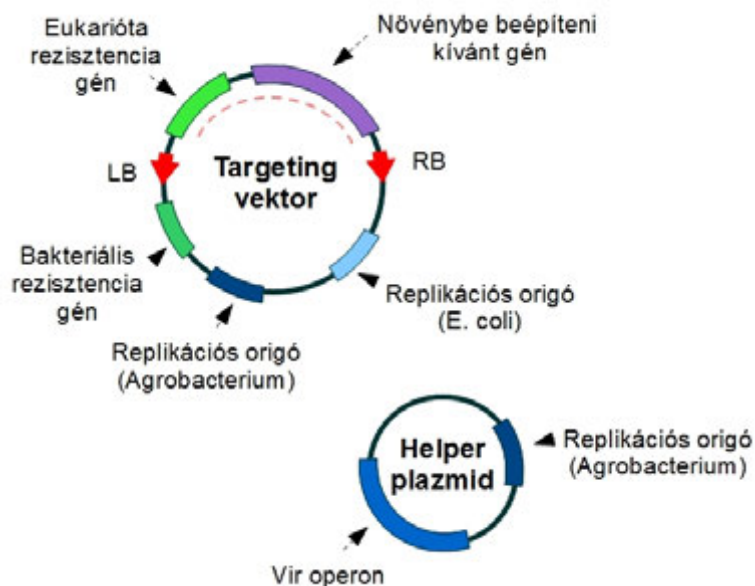
Növényi sejtek genetikai módosítására a protoplaszt-regenerációnál hatékonyabb módszerek is léteznek. A legrégebben ismert, és ma is a leggyakrabban használt eljárás egy különleges növényi parazita, az *Agrobacterium tumefaciens* tumor indukáló (**Ti**) **plazmidjét** használja. Ez a baktérium többféle **kétszikű növény** gyökerén élősködve, az *E. coli*-ból ismert F-plazmidhez hasonló **konjugációs rendszert** használ a gazda megbetegítésére. Ebben az esetben a konjugáció nem két baktérium között következik be, hanem az *Agrobacterium* a gazdanövény sejtjébe viszi át a plazmid egy részét, ahol az a genomba **integrálódik**. A Ti plazmid nagyméretű (~200 kbp), és számos operont tartalmaz. A daganatképzésért illetve a baktériumot ellátó tápanyagok szintéziséért felelős (eukarióta rendszerű) géneket a mesterséges konstrukciókban általában eltávolítjuk (*disarming*), ezek helyére lehet beépíteni a bevinni kívánt géneket. Az átvitelhez illetve a beépüléshez csupán a plazmid két speciális szekvenciájára ("bal

szél", **LB** és "jobb szél", **RB**) valamint a virulenciát biztosító **Vir operon**ra van szükség. A Vir operon valójában legalább 18 különböző gént tartalmaz: a plazmid átvitele ezek kooperációjával valósul meg. A *VirA* gén egy receptor-hisztidin-kinázt kódol, ami érzékeli a növény sérülésekor felszabaduló anyagot (acetosziringon). A *VirG* fehérje foszforilációja révén aktiválódik és bekapcsolja az operon többi génjét. A *VirC1* (felismerő) és *VirD1/2* (endonukleáz) fehérjék a plazmid megfelelő végeihez (LB és RB) kötődve annak hasítását eredményezik. A letekeredő, egyszálú DNS-t a *VirD2* és *VirE2* fehérjék védik meg a degradációtól, egyben a *VirB* gén terméke által létrehozott pórushoz is irányítják, amelyen keresztül az a növényi sejtbe exportálódik. A *VirD2* és *VirE2* fehérjék továbbra is kapcsolódnak a DNS-hez: ezek a növény sejtmagjába irányítják a DNS-darabot, és lehetővé teszik az integrációját (a *VirD2* a gazda DNS-javításába is beleszól). Ez jórészt **véletlen rekombináció** révén megy végbe. A **bal szél és jobb szél közé** (az úgynevezett T-DNS régióba) beépített gének így a gazdanövény kromoszomális DNS-ébe integrálódnak (ld. 14.6. ábra).



14.6. ábra: **Ti plazmid térképe.** A T-DNS régió helyettesíthető az inszerttel.

Az átviteli mechanizmus azonos a nopalín és oktopin típusú Ti plazmidoknál, amelyek egyébként sokban különböznek egymástól, de mindkettőt lehet génbevitelre használni. Bár a Ti plazmid **túl nagy** a közvetlen klónozáshoz, az e célra tenyésztett *Agrobacterium* törzsek plazmidjait manipulálhatjuk többféle módon is: bakteriális homológ rekombinációval (**pop-in**) be lehet vinni a kívánt géneket a "lefejezített" (**disarmed**) Ti plazmid T-DNS régiójába egy kisebb köztes vektorból is, amit *E. coli*-ban szaporítunk. De e helyett ma már inkább dupla vektorrendszert (**binary vector**) használunk: ilyenkor az *Agrobacterium* törzs az eredeti Ti plazmid egy kis darabját tartalmazza, amely szinte csak a Vir operonból áll (helper plazmid). Az RB és LB végeket csak a másik plazmidba építjük be, ami a célgént is tartalmazza, és *E. coli*-ból preparálható (ld. 14.7. ábra) Ezt a plazmidot *Agrobacterium*-ba transzformálva, a baktériumokat antibiotikummal szelektálva, majd a növényi sejtenyészeteket fertőzve nagy hatékonyságú **génbevitel** érhető el. A megtámadott sejteket tartalmazó **kalluszokból** pedig transzformált embriók nyerhetők.

14.7. ábra: Az *Agrobacterium* kettős vektorrendszere

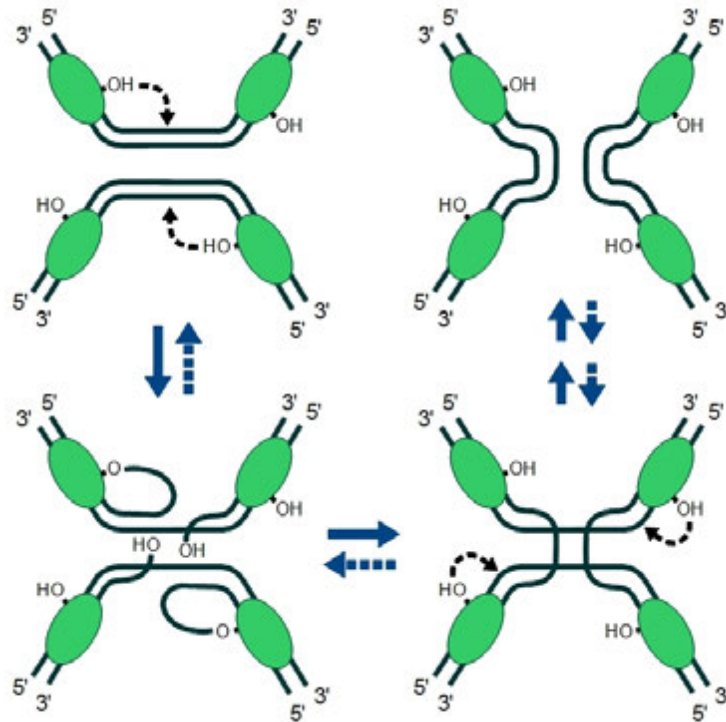
A fejlődő virágok fertőzésével a csirasejtek direkt módosítása is megvalósítható ("virágbemártás", **floral dip** módszer). A legutóbbi időkben sikerült olyan *Agrobacterium* törzseket létrehozni, amelyek **egyszikű növényeket** (pl. gabonákat) is képesek fertőzni. Bár a rendszer önmagában célzás nélkül működik, ez nem jelenti azt, hogy nem lehet **homológ rekombinációra** is használni. Megfelelő pozitív-negatív szelekciós konstrukciókkal elfogadható hatékonyságú (>0,1%) homológ rekombináció és célzott génmódosítás érhető el. Mindezt azonban a jelenlegi növényi géntechnológia alig-alig használja ki: a mezőgazdasági biotechnológia legtöbb mai céljára a célzás nélküli génbevitel is általában megfelel.

14.5. Knock-out és knock-in technikák

A célzott génmódosító eljárások egyik első célja (és ma is az egyik fő felhasználása) a génjeikben elrontott, **null mutáns(KO: knock-out)** élőlények előállítása volt. Ezen mutánsok tanulmányozása sokat segíti a molekuláris biológiai folyamatok megértését. Továbbá a genetikailag módosított emlősállatok számos genetikai eredetű emberi betegség modelljéül szolgálhatnak.

14.5.1. Homing endonukleázok. Cre/Lox és Flp/Frt rendszerek

A modern génkiütési technikák speciális endonukleázok használatán alapulnak. Jóllehet egyszerű funkcióvesztéses mutánt létre lehet hozni csupán homológ rekombinációval is, a rezisztencia-gén jelenléte az eukarióta genomban nem kívánatos jelenségeket eredményezhet (a kromatin módosítása miatt befolyásolja a környező gének expresszióját). Így mindenképp szükség van olyan enzimekre, amelyek *in vivo* is képesek - a bakteriális restrikciós enzimeknél jóval nagyobb célzási pontossággal - specifikus módon szabni-varrni a célszervezet genomiális DNS-sét. Az ilyen tulajdonságokkal rendelkező, természetben is megtalálható enzimeket hívjuk **homing endonukleázoknak**. Ellentétben a bakteriális restrikciós enzimekkel, ezek hosszú (14-44 nukleotidos) DNS-szakaszokat ismernek fel, és nagyon specifikus folyamatokban játszanak szerepet. Jó néhányuk "önző" genetikai elemek: vírusok és transzpozonok részét képezi. Ilyen például a *Chlamydomonas* zöldalga **I-CreI** enzime (nem keverendő össze a Cre-rekombinázzal): ez abból az intronikus szekvenciából íródik át, amely a saját aktivitása következtében tud integrálódni a genomba (transzpozon-szerűen). A homing endonukleázoknak számos családja létezik, a mindennapi gyakorlatban két ilyen fehérjét használunk a célzott génmódosításokhoz: a Cre-rekombinázt (**ld. 14.8. ábra**) és az Flp ("flippáz") enzimet.

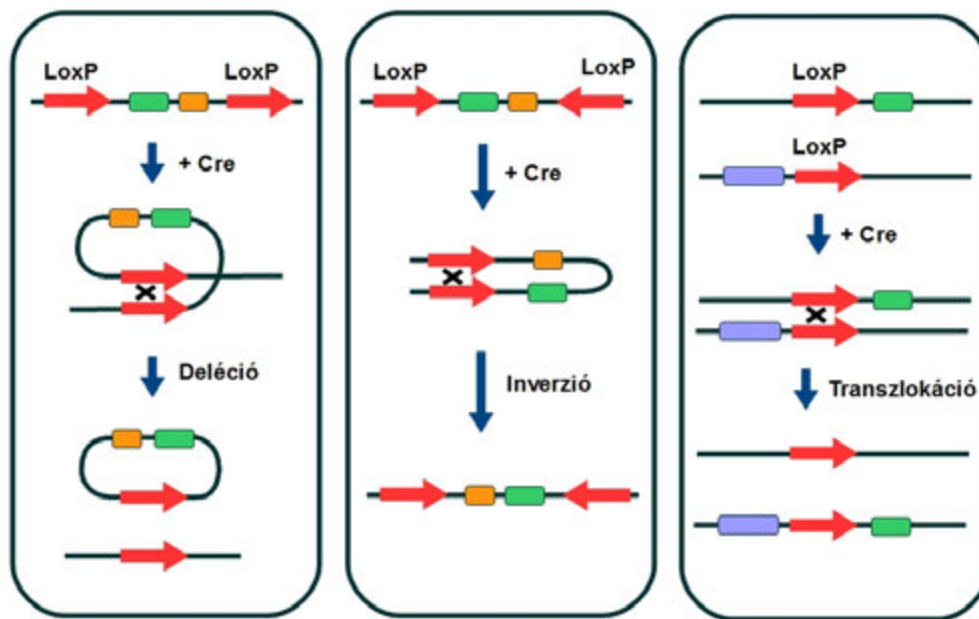


14.8. ábra: A Cre-rekombináz katalitikus ciklusa

A **Cre-rekombináz** eredetileg a P1 bakteriofágból ered, szerkezetileg hasonló a λ -fág integráz enziméhez. A felismerő helyének külön neve van: **LoxP** (*locus of crossover-P1*). Ez a két szélén palindrom, középen viszont aszimmetrikus nukleotid szekvenciát jelent (ATAACTTCGTATA-GCATACAT-TATACGAAGTTAT), ezért a LoxP helyek **nem szimmetrikusak**, irányultságuk van. A hasítás a nem palindrom szakaszon belül fog bekövetkezni. A Cre-enzim azonban nem csak endonukleáz aktivitással rendelkezik, hanem (a tetramerizáció révén, saját DNS-ligáz aktivitása révén) elősegíti a szabadabbá vált végek egyesülését is (**site-specific recombination**). Ez által a két (megfelelően irányított) LoxP hely közötti DNS-darab a Cre-enzim hatására kivágódik és cirkularizálódik. A kivágott darab az osztódások során általában elvész, deléciót hozva létre. A reakció "keresztben" is végbemehet: két kromoszóma karjait például fel lehet cserélni, ha mindkettő tartalmaz egy-egy LoxP helyet (**transzlokáció**). Az egy kromoszómán belüli két fordított LoxP hely rekombinálása pedig **inverzióhoz** vezethet. Így módon a Cre gént hordozó plazmid átmeneti bevitelle egy sejtbe vagy sejtvonalba (transziens transzfekció) **permanens genetikai változásokat** tud okozni a szervezet genomjában, ráadásul célzottan, a LoxP helyek pozíciójától függően. Ezen helyek ügyes beépítésével tehát képesek vagyunk késleltetett *in vivo* rekombinációs folyamatokat vezérelni.

Egyszerű kísérletekhez egy rekombinációs rendszer is elegendő, de az összetettebbekhez (mint a feltételes génkiütés) több, független enzimszisztemre van szükség. A másik gyakran használt rendszer az **Flp** ("flippáz") enzim alapszik. Ez a sörélesztőből izolált enzim felelős a *Saccharomyces* törzsek plazmid szerű törpe DNS-gyűrűinek („2-mikron” plazmid, **ld.13.1.6. fejezet**) fenntartásáért. A felismerő helyeinek neve **Frt** (Flp recognition target), ez a LoxP helyekhez hasonlóan középen nem szimmetrikus, a két szélén viszont majdnem vagy teljesen palindrom (GAAGTTCCTATTC-TCTAGAAA-GTATAGGAAGTTC). Az Flp-enzim is dimerizálódik a cél DNS-en, és középen, a nem szimmetrikus szekvenciánál hasít. A Cre-rekombinázhoz hasonlóan, a középső szakaszra igazából csak a rekombinációhoz van szükség, az endonukleáz aktivitáshoz nem kell. Így ezen szekvencia megváltoztatásával lehetőség van egy-egy rendszerből önmagában is több, keresztben nem kompatibilis felismerő helyet létrehozni (hasonlóan a Gateway vektorok AttR és AttL helyeihez). Ezeket a mesterségesen létrehozott, tervezett rekombinációs helyeket az eredeti neve mögé (mínusz egy betűvel) tett számozással szokták jelölni. Így például egy **Fr3** hely csak egy másik Fr3 hellyel tud rekombinálódni (egy **Fr5** hellyel nem), de ugyanúgy az Flp-enzim irányítja a folyamatot, mint az eredeti Frt helyen. Így egy enzim (átmeneti) expressziójával potenciálisan több, jól irányított rekombinációs eseményt is tudunk vezérelni. A módszer egyetlen korlátozása a felismerő helyekben rejlik: ezeket a helyeket előbb be kell építenünk a genomba (megfelelő targeting vektorral), hogy képesek legyünk a célzott génmódosításra. Ezért az utóbbi időben komoly erőfeszítés történt tervezetten átalakított homing endonukleázok

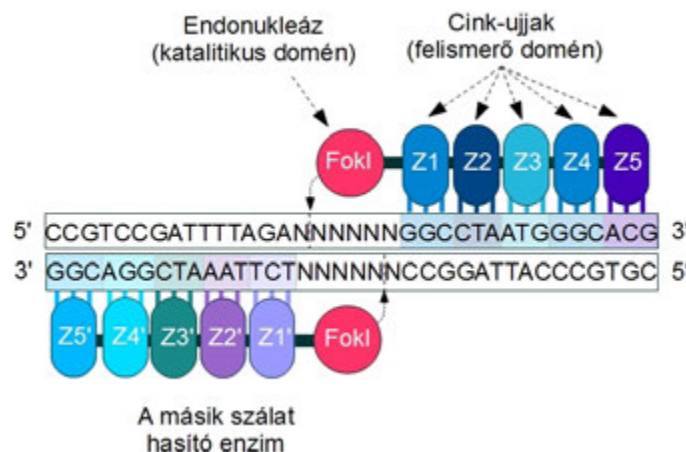
létrehozására (pl. I-CreI mutánsok), amelyekkel a genomális szekvenciák sokkal nagyobb tárháza célozható meg; akár természetes genomális szakaszok is célba vehetők. A *homing* endonukleázokkal kiváltható mutációk sémáját a 14.9. ábra foglalja össze.



14.9. ábra: Kromoszomális mutációk kiváltása *homing* endonukleázokkal

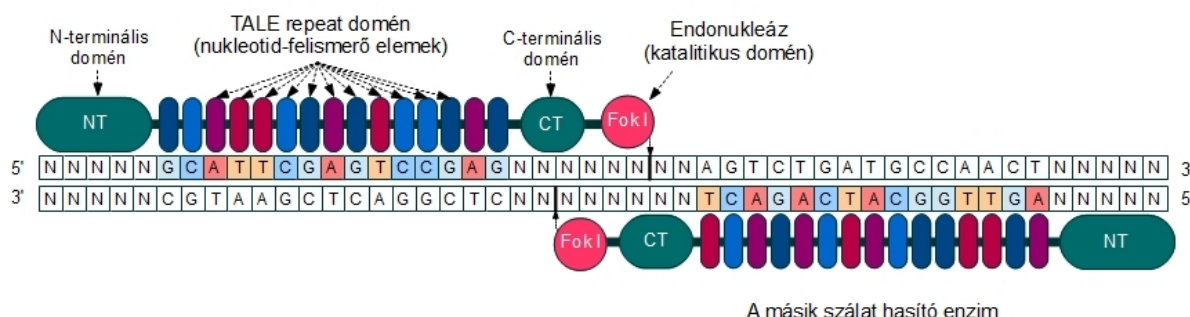
14.5.2. Célszekvenciára tervezett hasító enzimek. Cink-ujjas és TALE-nukleázok

A kromoszomális DNS célzott átszabása iránti nagy igény egészen újszerű módszerek kifejlesztéséhez vezetett. A 2010-es években jelentek meg először olyan technológiák, amelyekkel mód van lényegében **tetszőleges genomális régió** hasítására illetve rekombinálására. Erre a homing endonukleázok általában nem alkalmasak, hiszen a fehérje szubsztrátfelismerő és katalitikus régiója egymástól nem választható el. E helyett olyan mesterséges enzimeket kellett tervezni, amelyek egy DNS-felismerő elem és egy laza specificitású endonukleáz katalitikus domén fúziójával hozhatóak létre (**moduláris felépítés**). Jelenleg két ilyen megközelítés létezik, amelyet a gyakorlatban is használnak: az egyik közülük az úgynevezett **cinkujjas nukleázokon** (ZFN: *Zinc finger nucleases*) alapszik (ld. 14.10. ábra). A természetes cinkujjas transzkripció faktorok a DNS-hez jól meghatározott szekvencia-specificitással kötnek, méghozzá leggyakrabban három cinkujjal. A cinkujjak irányított mutációival ez a specificitás megváltoztatható, így a fehérjék nagyon sokféle célszekvenciához hozzádomíthatóak. Hosszabb "cinkujj-fűzésekkel" még nagyobb specificitás hozható létre: akár a genom egyedi szakasza is célba vehető. Ilyenkor a megfelelő nukleáz (pl. FokI) képzett fúziós fehérje célzottan el tudja vágni a kromoszomális DNS-t a felismerő hely közelében. A ZFN technológia tehát szükségtelemmé teszi a módosító helyek beépítését: a vad típusú genomokban is működik. A mesterséges nukleázok átmeneti bevitele (pl. plazmiddal) önmagában is rekombinációkat okoz: tartósabb expresszió esetén a nagyszámú hasítás-javítás ciklussal járó hibákkal könnyű **célzott pontmutációkat** vagy deléciókat létrehozni. Több, különböző célszekvenciára tervezett nukleázal pedig komplex átrendezéseket lehet végrehajtani a illető organizmus genomjában. A módszer kombinációja a targeting vektorokkal (miszerint célzott kettőstörést okozunk a kiszemelt génben, amihez a bejuttatott DNS tökéletesen passzol) nagyságrendekkel emeli a **homológ rekombinációs** módosítások hatékonyságát is.



14.10. ábra: A cink-ujjas nukleázok felépítése

A módszer hátulütője, hogy itt minden egyes genomiális szakaszra **külön tervezett enzimre** van szükség. Maguk a cinkujjas nukleázok is bizonyos szempontból korlátozottak, egyes DNS-szekvenciákon nem elég szelektív a működésük. Ez abból a tényből fakad, hogy a cink-ujjak a célszekvenciát blokkokban ismerik fel, nem pedig egyedi nukleotidoként. Léteznek viszont olyan természetes fehérjék, amelyek ezen a szinten is tökéletesen modulárisak. Az erre épített technológiát hívják **TALE-nukleázoknak** (TALEN: *transcriptional activator-like effector nucleases*; ld. 14.11. ábra). A fúziós fehérje alapja itt egy, a parazita *Xanthomonas* baktériumokból izolált fehérjecsatlád. Ezeket a fehérjéket a baktérium termeli, és juttatja be a gazdanövénybe azért, hogy ott bizonyos géneket célzottan aktiváljanak. Innen kapták a nevüket (transzkripció faktorokhoz hasonlítanak). A szerkezetük erősen **repetitív**: minden egyes rövid, 34 aminosavas szerkezeti blokk (*repeat*) **egy-egy nukleotid** kötéséért felelős, mégpedig nagy specificitással. Így az általunk megtervezett *repeat*-sorrenddel az egész TALE fehérje a kívánt DNS-szekvencia "tükörképe" szerint építhető fel (általában párban, hogy mind a két DNS-szálat elhasítsuk). Mindezt nem-specifikus endonukleáz doménekhez (*FokI*) kapcsolva egy új fajta mesterséges endonukleáz lesz az eredmény. Az egyetlen nehézség a gén mesterséges előállításával (gén-szintézissel) van: ez ugyanis a repetitív DNS-szekvenciák miatt nem a legegyszerűbb feladat, de általában megoldható.

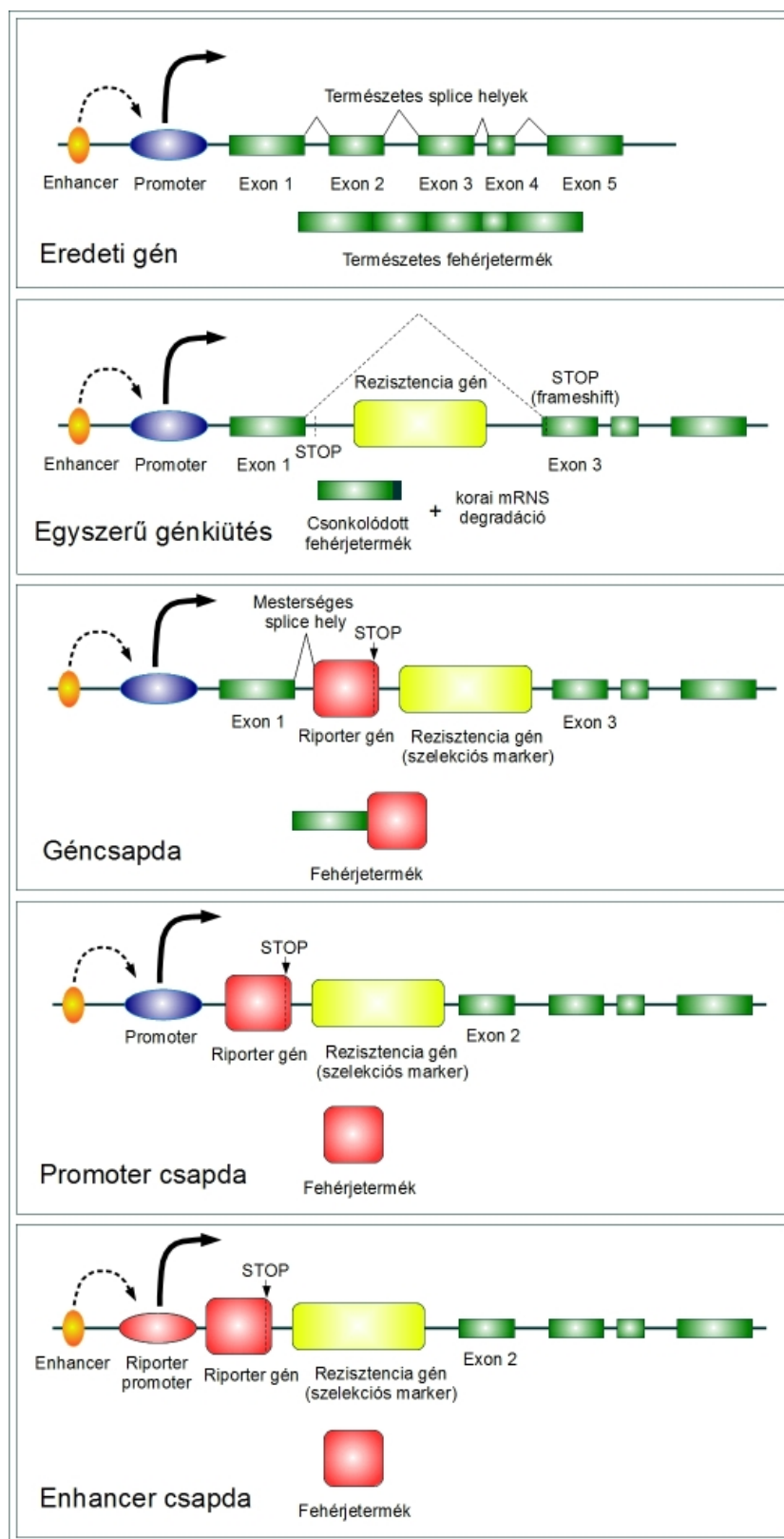


14.11. ábra: TALE-nukleázok és felismerőhelyük

14.5.3. Génkiütési technikák (*knock-out*) és új gének célzott beépítése (*knock-in*)

Ahogy azt láttuk, a **génkiütési** (*knock-out*) technikák ma már általában túlmutatnak a genomiális szekvenciák egyszerű elrontásán. A gén tönkretétele céljából elegendő csupán **egyetlenexont** - amely az összes izoformában jelen van - valami egyébvel, pl. a rezisztencia-gén kazettájával kicserélni. A módszer akkor biztos, ha *frameshift*-et

(leolvasási keret eltolódást) is okozunk az exon kiesésével. Ez az eljárás hatékony, de sokszor önmagában még nem elegendő.



14.12. ábra: A génscapdázás vázlata. Kombinált *knock-out* és *knock-in* módszerek

Felmerülhet az az igény, hogy a célszervezetben csak a kívánt módosítást hajtsuk végre: ne hagyjunk magunk után sok "DNS-szemetet", például a rezisztencia-géneket. Erre a célra (amelyet néha a "*delitto perfetto*" - olaszul a "tökéletes bűntény" - elvének is neveznek) a rezisztencia-gének köré általában egy Frt vagy LoxP párt építünk be. A megfelelő stádiumban (pl. az embrionális sejtek szelekciójának megtörténte után) a Cre vagy FLP gént hordozó plazmid átmeneti bevitelével a szelekciós gént vagy géneket eltávolítjuk, így azok nem szólnak bele sem a *splicing* folyamatába, sem az organizmus fenotípusának kialakításába. Ugyanez érhető el, ha a megfelelő mutáns egerünket egy olyan törzsbéli egérrel **kereszteljük**, amely már a korai embrionális fejlődés során is expresszálja a Cre- vagy FLP-enzimet (*deleter strain*).

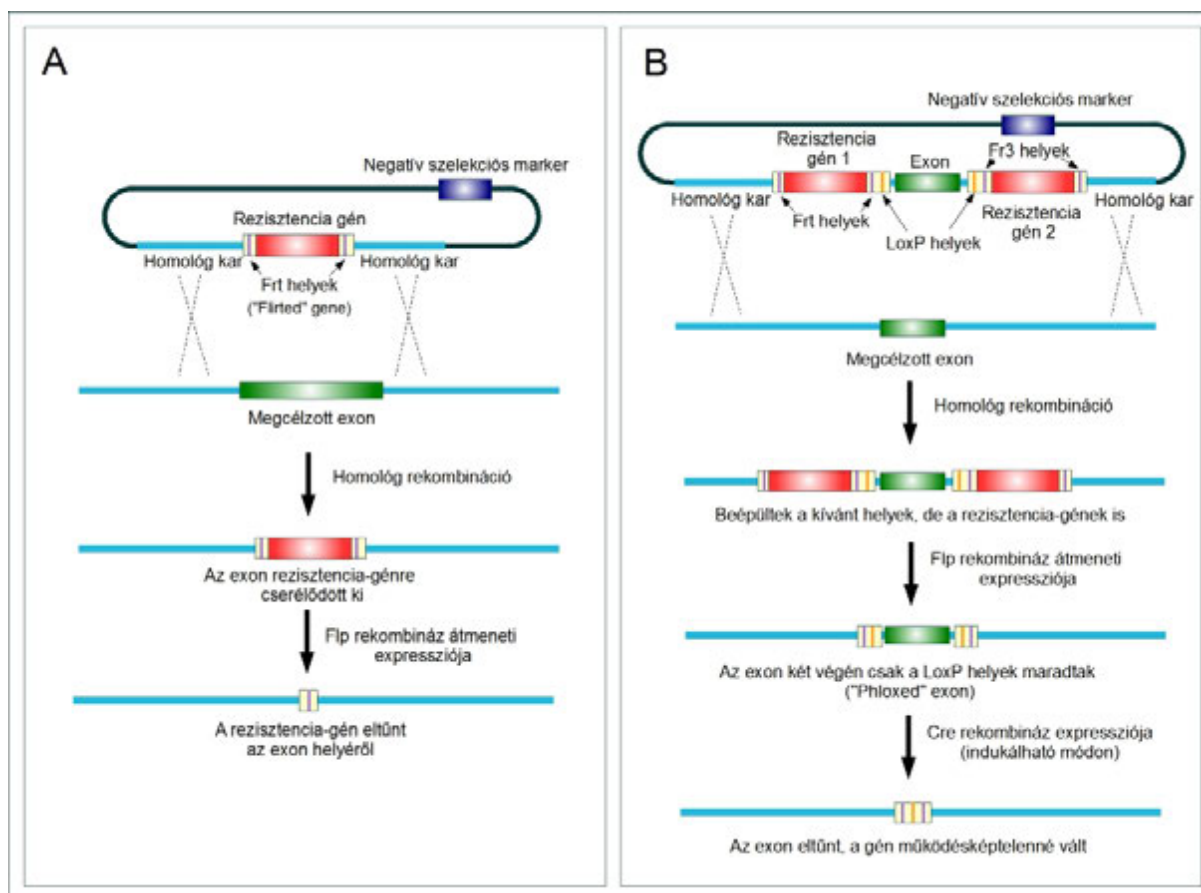
Természetesen vannak olyan helyzetek is, amikor kifejezetten szeretnénk beavatkozni az eredeti gén működésébe: például ha új genetikai elemeket viszünk be a kromoszomális DNS-be (**knock-in**). Ez kombinálható a génkiütéssel is: ha a targeting vektorunk nemcsak a rezisztencia gént tartalmazza, hanem előtte egy másik, **riporter-gént** is, akkor az megfelelő beépülés esetén átirrává válik. Ilyet eredményez, ha a riporter-gént úgy alakítjuk ki, hogy képes legyen *splicing* révén **fúziós fehérjét** alkotni az eredeti fehérje elejével – ez a "**génscapda**" (**gene trapping**) (**ld. 14.12. ábra**). Ugyanígy, a riporter (pl. GFP, béta-galaktózidáz) beépíthető az első exon helyére is (promóter trapping), sőt, ha van saját promótere, akár az eredeti promóter helyére is (*enhancer trapping*). Ezek a konstrukciók nemcsak a célzás sikerességéről adnak egyértelmű visszajelzést, de még az eredeti gén expressziós mintázatát is nyomon követhetjük velük. Ha a két bevitt (capda és rezisztencia) gén méretét szeretnénk minimalizálni, a második gén átirása saját promóter nélkül is megoldható, IRES elemekkel (bicisztronos eukarióta konstrukció).

Amikor az egértörzsünk genomjában egy adott helyen már van egy alkalmas rekombinációs helyeket (pl. egy Frt/Fr3 párt) tartalmazó kazetta, ezt **célzott génbevitelre** (**knock-in**) is felhasználhatjuk. Ilyenkor a meglévő kazettát egy transziensen bevitt plazmid új kazettájára lehet **kicserélni** (feltéve hogy ugyanazokat a rekombinációs helyeket tartalmazza, azonos orientációban, mint a kromoszóma). A Cre vagy Flt rekombinált embrionális összejektékből pedig a szokásos módon lehet kimérákat majd genetikailag tiszta egereket létrehozni (*Recombinase-mediated cassette exchange*).

14.5.4. Feltételes génkiütés, génjavítás

A homing endonukleázok vezérlése lehetővé teszi a rekombinációs események idő- és térbeli szabályozását is. Ezt ki lehet használni akkor, amikor csak a fejlődés adott pontján szeretnénk elrontani a megcélzott gént, vagy esetleg szövetspecifikus genetikai módosítást akarunk végrehajtani. Erre szolgálnak a **feltételes génkiütési** ("*conditional knock-out*") technikák. Ilyenkor olyan organizmusokat (pl. egértörzseket) használunk, amelyek genomjába a Cre- vagy FLP-rekombinázt előzőleg **indukálható promóterrel** építettük be. A gyakorlatban az egyik legnépszerűbb a **tetraciklin-indukálható** rendszer - ilyenkor a Cre gén promóterét egy **Tet-represszor** kötése tartja kikapcsolva (aminek génjét szintén beépítettük a rekombinázzal mellett). A rekombináció a **tet rendszerben** az egerek **doxiciklinnel** történő etetése után valósul csak meg. Természetesen a feltételes génkiütés esetén úgy kell megterveznünk a targeting vektort, hogy a gén a rekombináció nélkül még tökéletesen működjön. Ezért ilyenkor legalább **két kazettát** kell beépítenünk (két külön antibiotikum-rezisztenciával), méghozzá a megcélzott exon(ok) két oldalára. A kazettáknak szigorúan az **intronikus régióba** kell kerülniük, úgy, hogy a *splicing* folyamatát ne zavarják (150-200 nukleotidra az exon-intron határoktól). A rezisztencia-gének két oldalára egy-egy kompatibilis felismerő hely-pár kerül: az egyiket pl. Frt helyekkel, a másikat Fr3 helyekkel vesszük körül. Melléjük pedig egy-egy LoxP helyet teszünk. A hármas (kétszeresen pozitív, egyszeresen negatív) szelekció megtörténte után a két kazettát a transziensen bevitt FLP-rekombinázzal eltávolítjuk: **csak a LoxP helyek maradnak** meg sértetlenül a genomban.

Az ilyen genotípusú egér általában megkülönböztethetetlen lesz az eredetitől: egész addig, amíg a Cre-Tet törzssel nem **kereszteljük**, és az utódokat elkezdjük **doxiciklinnel** kezelni. Ekkor meg fog történni a rekombináció, és a kiszemelt gén működésképtelenné válik. Ez különösen akkor hasznos technika, ha tudjuk hogy az illető gén az embrionális fejlődésben más szerepet tölt be, mint felnőttben (korai kiütése esetleg letális). A Cre-Tet rendszer azonban csak egy a sok lehetséges változat közül: lehetőség van **sejttípus-specifikus** rekombinázzal promóterek előállítására is: ezek csak az adott, differenciált sejttípusban fogják végrehajtani a rekombinációt. Még arra is van mód, hogy a génkiütés ellentétét végrehajtsuk, és szabályozott módon **megjavítsuk** a kérdéses gént. Ez úgy lehetséges, hogy az előzőleg az intronba bevitt, a gént elrontó (pl. STOP kodont tartalmazó vagy egyéb módon gén-csapdázó) kazettát a rekombinázzal utólag kivágjuk, gyakorlatilag tökéletesen helyreállítva az eredeti viszonyokat. A feltételes génkiütés sémáját, összevetve az egyszerű génkiütéssel, a **14.13. ábrán** szemléltetjük.



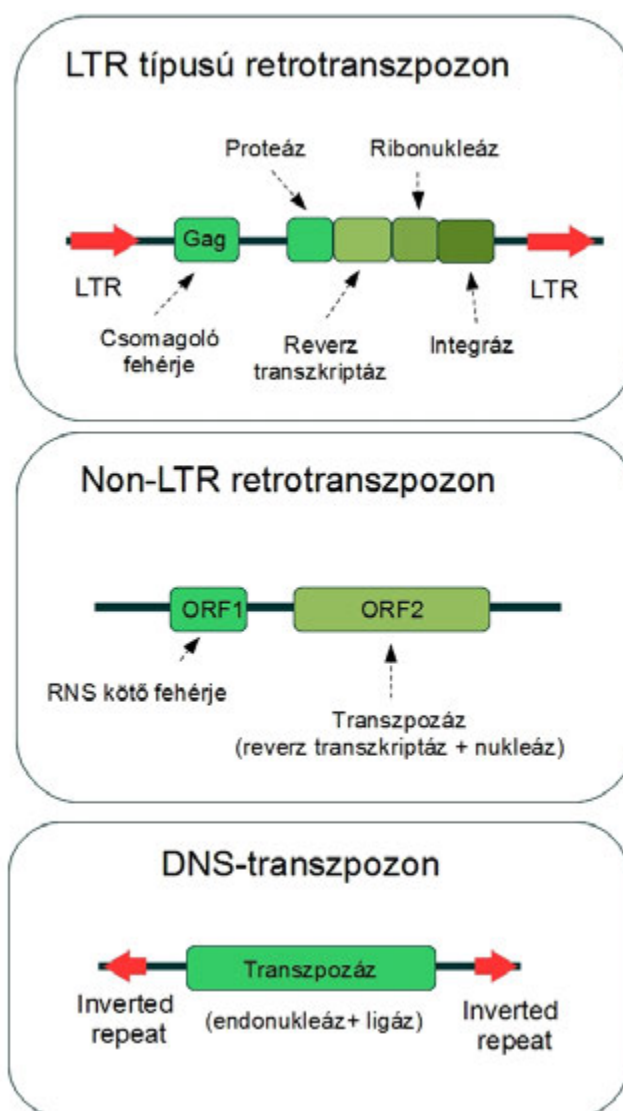
14.13. ábra: Cre/Lox és FRT/FRT rekombináza-rendszerek használata génkiütéses élőlények előállítására

14.5.5. Genetikai elemek beépítése transzpozonok segítségével

14.5.5.1. A transzpozonok molekuláris biológiája

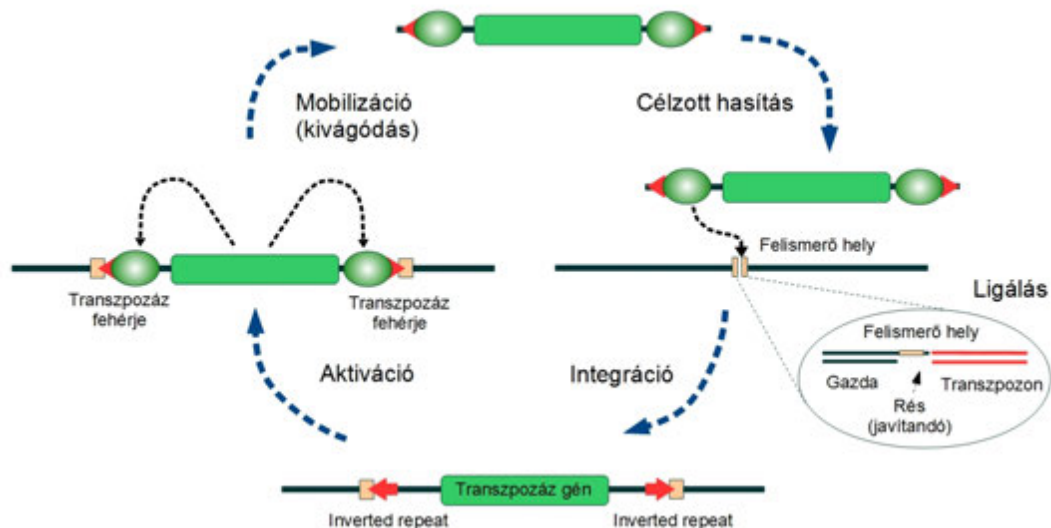
Gének - többé-kevésbé célzott - bevitelére vagy módosítására nemcsak a homológ rekombinációt vagy a *homing* endonukleázokat lehet használni. Erre a célra az olyan természetes genetikai paraziták, mint a vírusok és transzpozonok is sokszor megfelelnek. A mobilis genetikai elemeknek nevezett transzpozonoknak számos fajtája létezik: a "szaporodási" (transzpozíciós) mechanizmusuk alapján **két fő típus** különíthető el: az életciklusuk során RNS-re kiíródó majd DNS-re visszamásolódo **retrotranszpozonok** (I. típus), valamint a DNS-darabok kivágásával és átvitelével operáló **DNS-transzpozonok** (II. típus). A további, ritka és komplex típusok tárgyalásától (III - "politrónok", IV - "helitrónok") itt most eltekintünk. A transzpozonok fontosabb típusait a **14.14. ábra** mutatja be:

A retrotranszpozonok viszonylag komplexebb genetikai felépítéssel rendelkeznek, és minden esetben másolás-beillesztés (*copy-paste*) módszerrel mozognak a genomban. Egy részük esetében (LTR: *Long terminal repeat*, retrotranszpozonok) az mRNS visszairása DNS-re, majd annak integrációja a gazda-genomba (a transzpozon-DNS két végén található LTR szekvenciákat felismerő integrázok által) nagyon hasonlít a vírusokéra. A határ a retrovírusok felé semmiképpen sem éles: különösen, mivel egyes retrotranszpozonok genomja burokfehérjéket is kódol. Más részük esetében (non-LTR retrotranszpozonok) a reverz transzkripció és az integráció egyszerre következik be: Ilyen például az emberi genomban is megtalálható LINE1 retrotranszpozon. A LINE1 eredeti (aktív) formája összesen két gént tartalmaz. Az egyikük fehérje-terméke a transzpozon RNS-éhez kötődik, és annak a magba történő visszaszállítását irányítja. A másik fehérje pedig egyszerre rendelkezik DNS-endonukleáz és RNS-függő DNS-polimeráz aktivitással. Így a gazda DNS adott helyen történő hasítása után képes azt a fehérjéhez kapcsolódó RNS-szárlól kiegészíteni, bemásolva a transzpozont a képzett részbe.



14.14. ábra: A géntechnológiában alkalmazott transzpozonok legfontosabb típusai

A DNS-transzpozonok szaporodási ciklusa ennél is egyszerűbb. Ezek mindig a kivágás-beillesztés módszerrel mozognak. Ilyenkor a transzpozon által kódolt egyetlen fehérje homing endonukleáz aktivitásának köszönhetően képes felismerni és kivágni a transzpozon két, speciális szekvenciákkal (*inverted repeat*) rendelkező végét. Ezen enzimek a kivágott darab végeihez kapcsolódva a gazda DNS-sét is képesek bizonyos helyeken hasítani (ilyenkor a specificitásuk jóval kisebb, mint a transzpozon kivágásakor). Ráadásul ligáz aktivitással is rendelkeznek, így a transzpozon és a gazda genom törvégeit mindkét oldalon egyesíteni tudják, legalábbis az egyik szálon. A másik szálat a gazda saját javító rendszere fogja kipótolni és bezárni, gyakran megkettőzve ezzel a transzpozon felismerő helyét (ld. 14.15. ábra).



14.15. ábra: DNS-transzpozonok szaporodási ciklusa

14.5.5.2. Genetikai módosítás transzpozonokkal

A transzpozonok jól alkalmazhatóak idegen genetikai elemek célszervezetbe történő bevitelére is, ehhez azonban a következő néhány szempontot kell szem előtt tartanunk. Először is, a transzpozonok **beépülése** meglehetősen **specifikus**. A felismerő szekvenciák - mind az RNS, mind a DNS-alapú transzpozonok esetében - általában rövidek, lehetővé téve hogy a transzpozon potenciálisan nagyszámú helyre épülhessen be. Így - bár az integráció nem teljesen véletlenszerű - a bevitt konstrukció előre nehezen megjósolható hely(ek)re fog beépülni. Másrészt, a transzpozíciós aktivitást valahogyan irányítani kell. Az endogén transzpozonok aktivitása általában alacsony - mivel máskülönben a gazdaszervezetet veszélyeztetnék - és sok a csonkolódott vagy másképpen inaktiválódott kópia a genomban. A gyakorlatban ezért inkább olyan transzpozonokat használunk, amelyek nincsenek jelen a célszervezetben: ilyenkor mód van a transzpozáz aktivitás mesterséges megnövelésére, irányított mutációk révén. A **vektorunk csak a transzpozícióhoz szükséges felismerő végeket** tartalmazza (nem önálló transzpozon). Az enzim(ek)et kódoló (**helper**) **plazmidot** külön juttatjuk be a sejtbe, hogy csak átmenetileg legyenek jelen. Petesejtek genetikai módosítása esetén mód van a **transzpozáz** mRNS-ének mikroinjekciójára is. A transzpozon csak addig mozoghat, amíg az őt mozgató fehérjék jelen vannak, utána stabilan rögzül a genomban - ezzel elkerülhetőek a későbbi, nem kívánatos rekombinációk. De az elem irányított módon bármikor **újra mobilizálható**, a megfelelő, transzpozáz kódoló plazmid átmeneti bejuttatásával

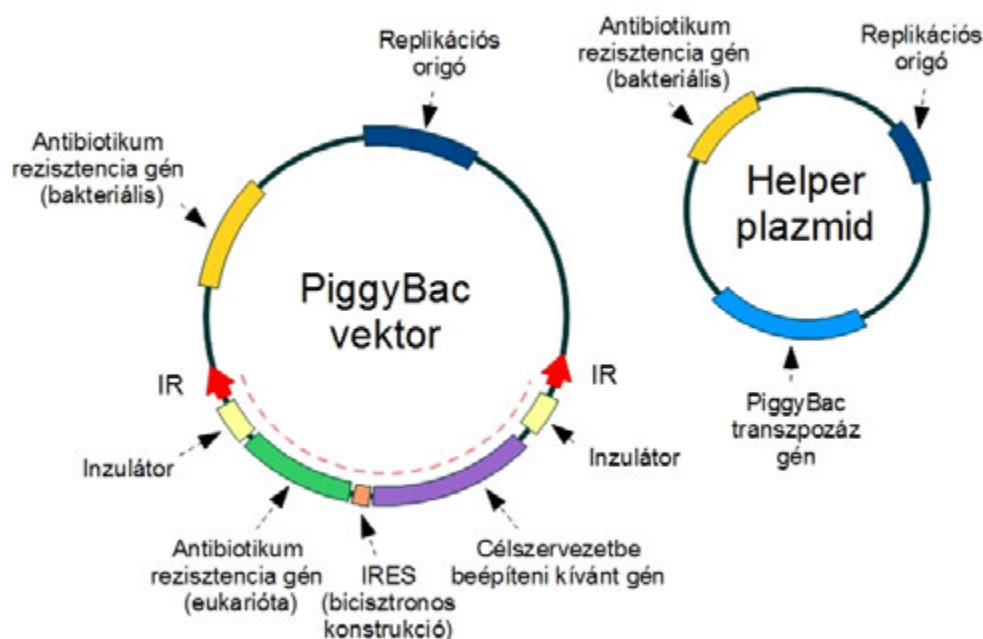
14.5.5.3. Transzpozonok géntechnológiai alkalmazásai

Új genetikai elemek beépítésére retrotranszpozonok is használhatóak: ilyen például a növényi géntechnológiában alkalmazott **Tnt1** (LTR típusú) transzpozon, amely a dohányból (*Nicotiana tabacum*) ered, és több, magasabb rendű növény genomi módosítására alkalmas. A véletlen beépülése során okozott mutációkkal (**inszerációs mutagenézis**) például feltérképezhető számos gén működése. De a DNS-transzpozonok egyszerűségüknek köszönhetően nagyobb népszerűségnek örvendenek. Ilyenek például a **Tol2**, a **PiggyBac** és a **Sleeping Beauty** elemek.

A **Tol2** transzpozon, amely a japán rizshalból (*Oryza latipes*) származik, nem rendelkezik semmiféle szekvencia-preferenciával, ami a beépülést illeti. Viszont előszeretettel integrálódik gerincesek **működő géneinek** közelébe vagy belsejébe, ezért jól alkalmazható **funkcióvesztéses mutánsok** létrehozására (géncsapdázó módszerekkel). A petébe injektált, riporter konstrukciót (pl. GFP-t) tartalmazó plazmid és Tol2 mRNS segítségével génkiűtött állatok állíthatók elő: sőt, még az eredeti gén expressziós mintázata is követhető. Az eltalált gén azonban sokszor csak szekvenálás segítségével azonosítható.

A *Drosophilában* már régóta ismert **P-elem** hasonlóképpen használható rovar genomok mutagenézisére, genetikai térképezésére (inszerációs mutagenézis, *enhancer trapping* módszerekkel). A **Sleeping Beauty** ("Csipkerózsika") transzpozon viszont egy teljesen mesterségesen létrehozott rendszer. Lazacfélék genomjában található, fosszilis

Tcl1/mariner típusú genetikai elemek (kihalt, egykori transzpozonok) alapján lett rekonstruálva, majd óvatos mutációkkal nagy aktivitására növelve. Jól használható különféle gerinces genomokba történő (jórészt célzás nélküli) **génbevitelre**, így a génterápiás próbálkozásokban is egyre gyakrabban alkalmazzák. Indukált őssejtek létrehozására is használható. Rövid célszekvenciája (TA dinukleotid) miatt szinte bárhol beépülhet a genomba, de a Tol2-től eltérően csak ritkán kerül be működő génekbe. A **PiggyBac** transzpozon eredetileg molylepkékből ered, de a legtöbb állati genom módosítására alkalmas. A felismerő helye (TTAA) rövid, de az **integráció** mindig **precíz** (nincs célhely-duplikáció), ellentétben a legtöbb egyéb transzpozonnal. Ezért a PiggyBac rendszerrel beépített elem nyom nélkül vágódik ki a genomból a transzpozáz hatására: így felhasználható **reverzibilis génmódosításra** is. Hátránya viszont, hogy a PiggyBac rendszer szintén hajlamos a működő gének közelébe beépülni. Hogy az általunk bevitt és az endogén gének interakcióit minimalizáljuk (ha ezt szeretnénk elkerülni), a génbevitelre szánt transzpozon vektorokat újabban **inzulátor elemekkel** is ellátják. Az inzulátorok gátat vetnek a kromatin módosításának, így az egymáshoz túl közel eső promóterek nem fogják zavarni egymást (ld. 14.16. ábra).



14.16. ábra: Példa a transzpozon alapú vektorokra: új gének bevitelére szánt PiggyBac vektorrendszer

14.6. Géncsendesítés (*knock-down*) RNS interferenciával

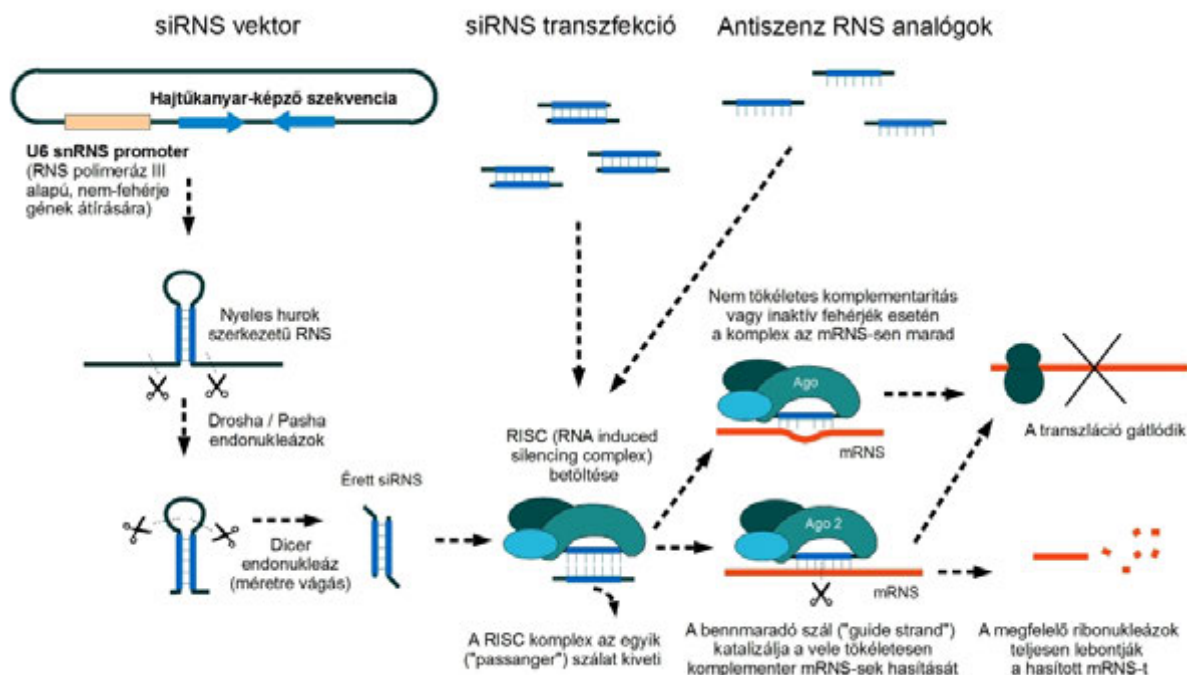
Bár a génkiütés hatékony módszer egy-egy gén működésének befolyásolására, ma már ennél finomabb módszerek is léteznek, amelyek ráadásul kevésbé munkaigényesek. Ezek általában nem genomiális, DNS-szintű módosításokat, hanem az RNS szintjén történő beavatkozást jelentik. A módszereket összefoglaló nevükön **RNS interferenciának** (RNAi) nevezik. Itt egy természetes jelenséget használunk ki. Már több évtizede ismert, hogy egy-egy fehérjekódoló gént RNS-szinten ki lehet kapcsolni, ha a megfelelő RNS szál komplementerét (**antiszensz oligonukleotidot**) bevisszük a sejtbe. Rövid, kettős szálú RNS-darabok (**siRNS: short interfering RNA**) segítségével még nagyobb hatás érhető el. A pontos mechanizmusok feltárása ma már lehetővé teszi a célzott géncsendesítést (**knock-down**) is. A legtöbb sejt számára a stabil **kétszálú RNS** teljesen idegen molekula: az őket megfertőző vírusokban, illetve azok életciklusa során viszont gyakran előfordul. A kétszálú RNS-ek lebontására eukariótákban ezért egy speciális rendszer szolgál. Ez alkalmasint lebontja a **szervezetben belül** keletkező kétszálú RNS-eket is. Az egyes RNS-ekben található hosszú (20-30 nukleotidos) hajtókanyarok (*stem-loop structures*) hurkon belüli hasításával normálisan is keletkezhetnek kettős szálú RNS-ek (**miRNS: micro-RNA**), amelyeknek a természetes génexpressziós szabályozásban van szerepük. A szervezetben a kétszálú RNS-ek feldarabolódnak, és a keletkezett egyszálú darabkák (katalitikus módon) segítik a velük komplementer **mRNS-ek hasítását**, vagy a **transzláció gátlását**.

Növények rendelkeznek olyan RNS-függő RNS-polimerázokkal is, amelyek képesek a hasított darabkákat kiegészíteni. Így a felismerő templát szaporodhat, és a plazmodezmákon keresztül az egész szervezetben szétterjedhet, **szisztémás RNAi választ** hozva létre. Szisztémás RNS-interferencia megvalósulhat a *Caenorhabditis elegans* nevű fonálféregnél is. Mivel rendelkezik a hosszú, kettős szálú RNS-ek (>400 bázispár) felvételére szolgáló receptorokkal és a szállításukhoz szükséges fehérjékkel, az RNAi válasz az állat kettős szálú **RNS-sel való etetésével** is kiváltható. Gerincesekben azonban nincs ehhez hasonló szisztémás RNS interferencia: a kétszálú RNS hatása az egyes sejtekre korlátozódik. Emlős sejtekbe természetesen nem juttathatunk be **hosszú antiszensz** nukleotidokat következmények nélkül: a 25-30 nukleotidnál hosszabb kétszálú RNS-ek ugyanis a természetes immunitás **antivirális rendszerét** is aktiválják, és a transzláció teljes blokádjához vezetnek. Ezért a mikro-RNS-eknek megfelelő, rövid nukleotidokkal kell a géncsendesítést végrehajtani.

A megfelelő RNS-ek bejuttatására több mód van: transzformálhatunk olyan **plazmiddal**, amibe erős promóter (pl. U6 splicing alegység snRNS) mögé építjük be a hajtókanyar-képző szekvenciát. A középső, hurok régió mindig azonos maradhat: az önmagával párosodó szár szekvenciáját kell a cél RNS-hez szabni. Bár az ilyen konstrukciók szekvenálása gyakran problémás (a hurok a DNS-ben is megjelenhet a PCR közben és elakad rajta a DNS-polimeráz), ezeket a rövid hajtú RNS-seket (shRNS: **short hairpin RNA**) kódoló plazmidokat széles körben alkalmazzák géncsendesítésre (**gene silencing**). A kis hajtókanyart többsejtű állatokban a **Drosha** és **Pasha** endonukleázok ismerik fel, és vágják ki a környező, hosszabb RNS-szakaszból, akárcsak a természetes miRNS gének esetén. Ezután a **Dicer** endonukleáz lehasítja a hurkot, így az siRNS valódi dimerré válik. A méretre vágott (nagyjából 21 nukleotid hosszú) siRNS-darabka végül a **RISC komplexbe** (*RNA induced silencing complex*) kerül. A komplex magját alkotó Argonaute endonukleáz a két siRNS szál egyikét (*guide strand*) stabilan **megkötö**, a másik szálát (*passanger strand*) pedig **kiveti**. A komplex ezután a megkötött siRNS-szállal párosítható, hosszabb, egyszálú RNS molekulákat keres. Tökéletes komplementaritás esetén az Argonaute endonukleáz az mRNA szálát középen **elhasítja**. Ezzel **megakadályozza a transzlációt**, és további, sejten belüli ribonukleázok szubsztrátjává teszi, amelyek előbb-utóbb teljesen **lebontják** a hasított mRNA-t. Abban az esetben, ha a vezető siRNS szál és az mRNA közötti bázispárosodás nem tökéletes (a két szál a közepén nem teljesen komplementer) vagy inaktív fehérjét tartalmazó komplexről van szó (az emlős Argonaute fehérjék közül például egyedül az Ago2-nek van mérhető enzimaktivitása), a hasítás nem megy végbe. A **transzláció** azonban ilyenkor is **gátlást szenved**. Az siRNS-ek számos eukarióta élőlényben szabályozzák a transzkripciót (RITS komplex), a transzpozonok expresszióját (PIWI-kötő RNS-ek, piRNA), sőt a kromatin módosítását is. Ezen folyamatok molekuláris szintű részleteit azonban ma még csak kevésbé ismerjük.

Az RNS-interferenciára egy másik lehetőség, hogy **szintetikusan** előállított siRNS-eket jutattunk a sejtekbe. A **tervezéskor** csak arra kell ügyelni, hogy az oligonukleotid valóban **egyedi** legyen (a teljes transzkriptom ellenőrzésével). Gyakorlatilag a célgén bármelyik szakaszát megcélozhatjuk (nem csak a kódoló szakaszt, hanem az 5'- vagy 3'-UTR végeket is). Erre automata **tervezőprogramok** is rendelkezésre állnak. Bár optimális géncsendesítés néha elérhető akár egyetlen fajta kétszálú RNS-darab bejuttatásával is, általában érdemes legalább 3-4 **különböző** szekvenciát kipróbálni, hogy a "szerencse kisasszonyt" kiiktassuk a kísérletből. A géncsendesítés sémáját a **14.17. ábrán** mutatjuk be.

A géncsendesítés a gyógyszeripar részéről is hatalmas érdeklődésre tart számot: a cél az RNS interferencia alapú **terápiák** megvalósítása. Ilyenkor nem természetes nukleotidokat, hanem például foszfotioát kötést tartalmazó (általában egyszálú) **nukleotid analógokat** használnak, amelyek ellenállóak a ribonukleázokkal szemben. Bár a az **egyszálú RNS**-sekkel nehezebb a RISC komplexet megcélozni (mivel az Argonaute fehérje kötéséhez előbb 5' pozícióban foszforilálódnia kell a nukleotid analógnak), ma már több antiszensz RNS-származékokat használnak sikerrel különféle betegségek kezelésére. Ilyen például a régóta forgalmazott fomiversen (a citomegalovírus IE2 génje ellen, antivirális szerként) vagy a mipomersen (apolipoprotein-B ellen, hiperlipidémia kezelésére). Az RNAi alapú terápiák fejlesztése azonban nem könnyű feladat, mivel a megfelelő nukleotid-analógok gyakran instabilak, továbbá hajlamosak gyulladást és egyéb nem kívánatos (*off-target*) válaszokat okozni. De az új RNS-analóg gyógyszerek előállításának legnagyobb nehézsége, hogy a testi sejtekbe egyelőre nem tudjuk igazán hatékonyan bejuttatni a megfelelő nukleotid-analógokat.



14.17. ábra: A géncsendesítés sejtenyészetben, RNS interferencia alapú módszerekkel

14.7. Stabil episzomális rendszerek

Jegyzetünkben eddig csak stabil genetikai elemekről beszéltünk, és csak a szervezet saját kromoszómáiról esett szó. Mióta azonban tudjuk hogy mi kell egy kromoszóma fenntartásához, létre tudunk hozni **mesterséges kromoszómákat** is. A plazmidokkal ellentétben, amelyeknek nincsenek öröklést segítő elemei (centromérái), de legtöbbször még eukarióta replikációs origói sem, a mesterséges kromoszómák genetikailag stabilnak tekinthetők. Élesztő esetén a technológiát már régóta használják (YAC), de ma már léteznek mesterséges emlős, sőt akár **humán kromoszómák** is (HAC). Ez lehetővé teszi **nagyméretű gének** sejtekbe történő stabil bevitelét, **genomiális integráció nélkül**. A technológia egyelőre még éretlen, de a módszer megoldást kínálhat sok, **reverzibilis génmódosítást** kívánó feladatra is. Példának említjük az **indukált őssejtek** előállítását. Ezek előállításához a ma létező összes protokoll megköveteli három vagy négy gén bevitelét a pluripotencia indukálására (pl. *Sox2*, *Klf4*, *Oct-4* és *Myc*). Ha viszont ezek az érett szövetben továbbra is expresszálnak, problémákat okozhatnak (pl. a *Myc* gén magas aktivitása daganatokra hajlamosít). Megoldást kínálhat, ha ezeket a géneket egy külön mesterséges kromoszómán visszük be, ahol a gének aktivitását könnyen szabályozhatjuk, akár blokkban is. De ha ez a minikromoszóma **szelektíven eltávolítható** (erre egyelőre csak a konvencionális negatív szelekciós technikák, pl TK/ganciklovir módszer állnak rendelkezésre), akkor megvalósítható az őssejtek programozás bármilyen genetikai elem tartós beépítése nélkül, genetikailag tiszta rendszerben, stabil episzomális gének révén.

14.8. További olvasnivaló a fejezethez

Alexandra Joyner (1999) Gene targeting: A Practical Approach. 2nd edition. Oxford University Press, ISBN: 978-0-199-63792-8

Recombineering 2012: <http://redrecombineering.ncifcrf.gov/> (the Court lab official recombineering website)

Hafez, M, Hausner, G (2012) Homing endonucleases: DNA scissors on a mission. Genome 55(8): 553-69.

Zayed H, Izsvák Z, Walisko O, Ivics Z. (2004) Development of Hyperactive Sleeping Beauty Transposon Vectors by Mutational Analysis. Molecular Therapy, 9: 292-304.

Hutvagner G & Simard MJ (2008) Argonaute proteins: key players in siRNA silencing. Nature Reviews, Molecular and Cell Biology 9(1): 22-33.

15. fejezet - Transzgenikus élőlények. Génterápia

A **genetikailag módosított élőlények (GMO: Genetically Modified Organism)** magukba foglalják az összes olyan szervezetet, melyek genetikai anyagát módosították valamely géntechnológiai módszerrel. A **transzgenikus élőlény** pedig egy olyan **GMO, melybe egy másik szervezetből izolált gént juttatunk be**, amit eredetileg az nem tartalmazott. Az így „kívülről” bejuttatott, exogén eredetű transzgén öröklődik az utódokban, ha az a csíravonalba került be. Transzgenikus élőlények előállítása sokféle módszerrel történhet, amelyeket az előző fejezet (**14. fejezet**) részletesen tárgyal.

Az idegen gén (transzgén) termékét a gazdaszervezet ugyanúgy termeli, ahogy a saját fehérjéit. Az expresszió mértéke a génnel együtt bevitt egyéb DNS szakaszoktól függően (promóter régió, 5' *upstream* elemek, 3' UTR) változhat, illetve szövetspecifikussága is ettől függ. Célunk A transzgén expresszállása mellett a gazdaszervezet génextpressziós mintázatának megváltoztatása is cél lehet. Tágabb értelemben transzgénnek nevezhetünk minden olyan, nem feltétlenül gént kódoló DNS szakaszt, vagy mesterségesen előállított gént, melyeket olyan élőlényekbe vagy vektorokba juttatunk be, ahol azelőtt nem fordultak elő. Praktikussági szempontból elkülönítjük a genomális transzgéneket a cDNS transzgénektől, mely utóbbiak nem tartalmaznak intronokat és a gén körüli szabályozó régiókat. A kifejezetten nagyméretű klónok tárolására/manipulálására kifejlesztett BAC, YAC és kozmid könyvtárak (ld. 7. fejezet) előállításával a genomális transzgének szabályozó elemeivel együtt történő klónozása könnyen megvalósítható. A transzgenikus élőlények előállítása új távlatokat nyitott a kutatásban, gyógyszeriparban, mezőgazdaságban és számos egyéb iparágban.

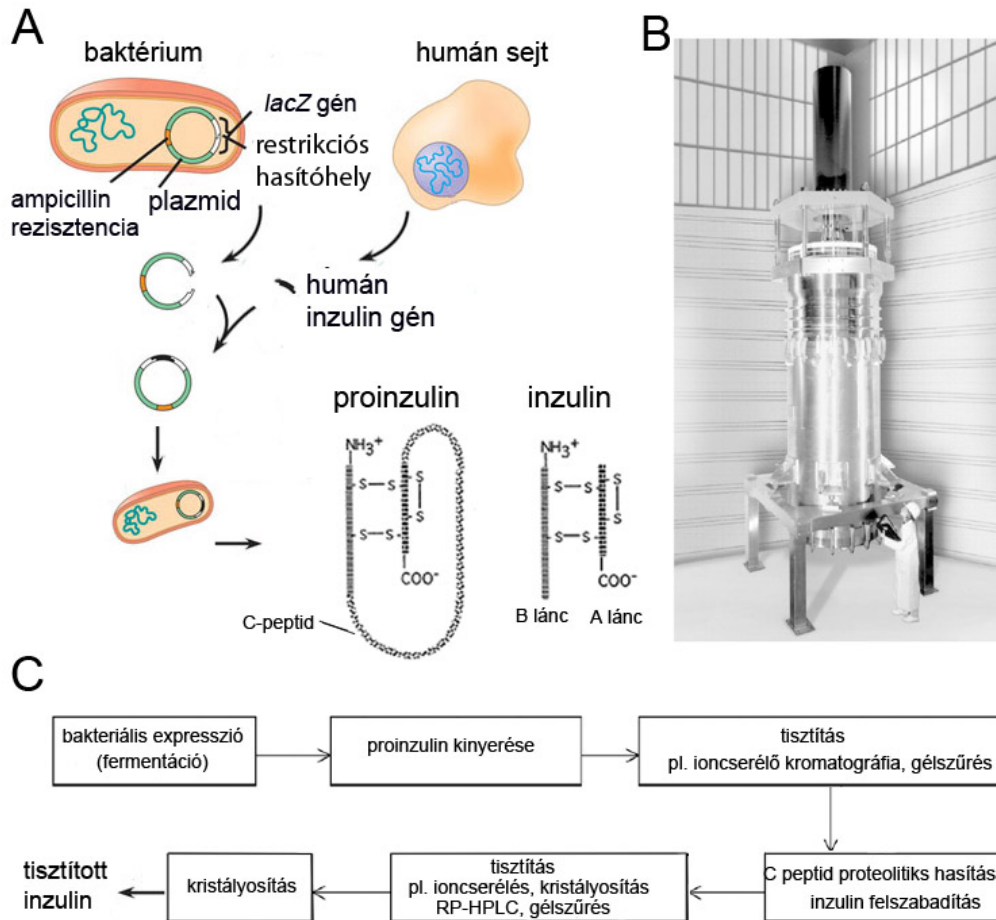
A fejezetben belül először példákon keresztül bemutatunk különböző transzgenikus élőlényeket (növényeket, állatokat, mikrobákat) és a legfontosabb felhasználási területeiket. Röviden kitérünk a GMO-kal kapcsolatos tudományetikai és környezetbiztonsági (*biosafety*) kérdésekre.

15.1. A transzgenikus élőlények típusai és felhasználásuk

A transzgén gazda és a felhasználásuk lehetőségei alapján célszerű külön tárgyalni a transzgenikus növényeket, állatokat és mikrobákat. A transzgenikus növények fő alkalmazási területe értelemszerűen a mezőgazdaság és élelmiszeripar, bár megjegyzendő, hogy ma már gyógyszerek előállítására is be lehet őket vonni. Transzgenikus baktériumokat használnak gyógyszerfehérjék előállítására, a bioremediációban és a környezeti biotechnológia több más területén is. Röviden összefoglaljuk a transzgenikus állatok szerepét az alapkutatásban, a gyógyszeriparban és a gyógyászatban is.

15.1.1. Transzgenikus mikrobák

A baktériumok voltak az első olyan szervezetek, melyek genetikai anyagát módosították laboratóriumi körülmények között. A transzgenikus mikrobák szerepe ma már a **bioremediációban** nélkülözhetetlen. A higanyszennyezések eltakarításában például egy olyan *E. coli* törzs bizonyult hatékonynak, ami egér metallotionein (mt-1) és *E. coli* polifoszfát-kináz (ppk) géneket tartalmaz dohány 16S riboszomális RNS gén konstitutív promótere mögött. Ezek az enzimek higany rezisztenciát és annak felhalmozódását eredményezik a baktériumban. Egy másik alkalmazásban *E. coli* DH5α törzs lizátumát használták bioriporiterként arzént tartalmazó víz kimutatására. A baktérium plazmidja egy arzén-indukált promótert tartalmaz, amely mögé a luciferáz enzim génjét klónozták be. Arzén jelenlétében a luciferáz szubsztrátja újratermelődik. A baktériumkultúrából szuszpenziót készítettek, amit csövekben liofilizáltak. A teszt kit egy-egy csövéhez vizet adva luminométerrel detektálták a minta lumineszcenciájának, így arzéntartalmának a mértékét.



15.1. ábra: **Humán inzulin termelése *E. coli* baktériumban.** (A) Baktériumból izolált plazmidot és a humán inzulin génjét (izolált vagy szintetikus) ugyanazokkal a restrikciós enzimekkel hasítjuk, majd a ragadós végeikkel a humán inzulin génjét a plazmidba ligáljuk. Szelekció után a plazmidot a megfelelő *E. coli* törzsekbe transzformáljuk, ahol folyamatos a β -galaktozidáz promótere utáni proinzulin transzkripciója és transzlációja. Különböző tisztítási lépések után történik a C-peptid tripszines hasítása, ami az inzulin végleges formáját eredményezi. (B) Az inzulin tisztításához használt HPLC oszlop. (C) Az inzulin termelés folyamata. (a módosított ábra forrása: Walsh, Gary. "Therapeutic insulins and their large-scale manufacture." *Applied microbiology and biotechnology* 67.2 (2005): 151-159.)

A gyógyászatban is alkalmaznak transzgenikus mikrobákat, **rekombináns fehérjék termelésére**. Az inzulin az egyik leggyakrabban használt gyógyászati célú fehérje, olcsó és nagy mennyiségű termelése rendkívül fontos. Humán inzulin nagy mennyiségben először *Escherichia coli* törzsek felhasználásával állítottak elő (a *humulin* fantázianévű rekombináns inzulin a Genentech cég fejlesztette ki 1978-ban, gyógyszerként az Ely Lilly kezdte forgalmazni 1982-ban). A humán inzulin A és B láncát külön törzs termelte szintetikus nukleotid szekvenciákról, ezeket aztán külön tisztították, majd oxidatív környezetben együtt inkubálták, így a két szál között kialakult két diszulfid-híd kötéssel jött létre az intakt inzulin. Az egyes inzulin láncokat egy *lac* konstitutív promóter mögé klónozták, a β -galaktozidáz génjével együtt. A pozitív klónokra kék-fehér szelekcióval és ampicillin antibiotikummal szelektáltak. Transzlációt követően a β -galaktozidáz C-terminálisához egy metioninnal kapcsolt inzulin cianogén-bromidos kezeléssel szabadították fel, elhasítva a metionin utáni peptid kötést. Egy másik módszernél a humán proinzulin egyetlen szintetikus szárlól íródik át, a polipeptidből pedig tripszinnel hasítják ki a C-peptidet (**ld. 15.1. ábra**). Rekombináns humán inzulin termelésére élesztőt (*Saccharomyces cerevisiae*) is használnak. Helyspecifikus mutagenézissel módosított inzulin analógokkal mára már számos eltérő hatóidejű és stabilitású inzulin van forgalomban.

Genetikailag módosított baktériumokkal véralvadási faktorokat, és növekedési hormonokat is előállítanak (**ld. 15.1. táblázat**). A baktériumok által termeltetett fehérjék jóval biztonságosabbak, mint a korábban alkalmazott módszer, amikor is kadaverekből vonták ki a még le nem bomlott fehérjéket (pl. hasnyálmirigyből inzulin, mely

egy cukorbeteg két-három napig lát el). Egyrészt így a szerben jelenlevő kórokozók is továbbjuthattak (AIDS, hepatitis-C, Creutzfeldt-Jakob szindróma), másrészt a megnövekedett igényeket ezekből nem lehetett kielégíteni.

Egyes országok katonai hadviselésének részét képezi a mikrobiális biológiai fegyverek előállítás és esetleges használata. Ezeket a nemzetközi szerződésekben tiltott biofegyverek egy részét transzgenikus baktériumok formájában állítják elő.

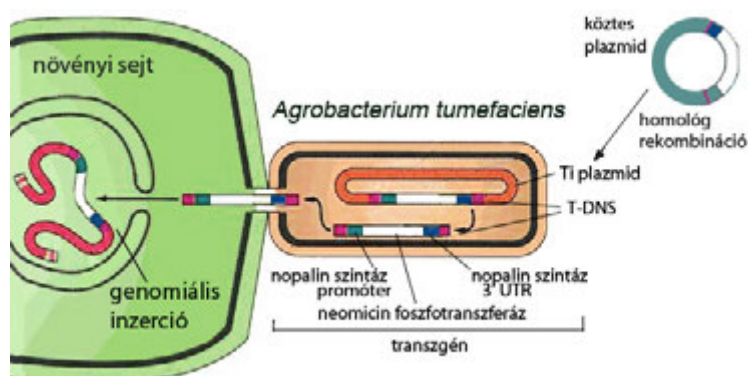
15.1.2. Transzgenikus növények

A mezőgazdaságban termesztett transzgenikus növények az egyik leggyakrabban emlegetett példái a genetikailag módosított (GM) szervezeteknek. Az olyan génmódosítások, melyek nagyobb terméshozamot, kevesebb rovarirtószer használatát, nagyobb tápértékű termést vagy betegségeknek ellenállóbb fajtát eredményeznek, számos előnyt jelentenek a mezőgazdaságban. Növények, melyek gyorsabban beérnek, tolerálják a különböző talajszennyezéseket, sőt, szárazságot, fagyot vagy egyéb szélsőséges időjárást olyan vidéken is meg tudnak élni, ahol előtte nem, illetve ahol előtte növénytermesztésre nem volt lehetőség. Földünk egyre növekvő emberi populációjának ezek a tulajdonságok a jövőben létfontosságúak lesznek.

Hagyományos növénytermesztők kereszteztek először olyan növényeket, melyek természetes szaporodással nem jöhettek volna létre. Ilyen volt az 1875-ben előállított búza-rozs hibrid, a tritikálé. Hamar rájöttek, hogy a hibrid tartalmazza mindkét faj hasznos tulajdonságait, amire szelektálni lehetett. Az első generációs GM növényeket mutációs szaporítással állították elő, ahol a növények sugárzásnak, illetve vegyszereknek voltak kitéve. Ilyen körülményekkel növelték a mutációs rátájukat, és utódaik közül válogatni lehetett a különböző, stabilan megmaradó és hasznos tulajdonságokra. Ehhez a nem specifikus módszerhez képest a transzgének bevitel egy sokkal pontosabb és szabályozhatóbb módja a kívánt tulajdonságok elérésének.

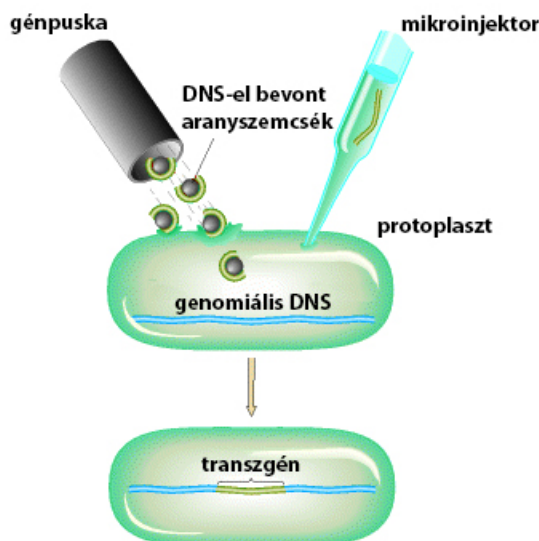
15.1.2.1. Transzgenikus növények előállítása

Az első transzgenikus növényt 1983-ban állították elő, amikor az aminoglikozid típusú antibiotikumok (pl. kanamicin, neomicin, geneticin) rezisztencia génjét petúnia és dohány növényekbe juttatták be [1]. A kiméra gén a neomicin I-es vagy II-es típusú foszfortranszferáz enzim génjét illetve a nopalín-szintáz gén szabályozó régióit tartalmazta (ld. 15.2. ábra). A neomicin-foszfortranszferáz foszforilációval inaktiválja az aminoglikozid antibiotikumokat, míg a nopalín-szintáz konstitutív promótere biztosítja az állandó expressziót a növényi sejtekben. A kiméra génszakaszt először egy **köztes plazmidba** klónozták, amiről rekombinációval került be a transzgén egy **tumort indukáló plazmidba (Ti plazmid)**, melyet *Agrobacterium tumefaciens* talajbaktériumba transzformáltak (ld. 14.4.4. fejezet). Ez a talajbaktérium a természetben is sokféle Ti plazmidot tartalmaz. A baktériummal fertőzve a növényi sejteket, a Ti plazmidból a **növényi sejtek genomjába** épül be egy DNS-darab (**T-DNS**), mely tartalmazza a transzgént is. A beépülés célzottan történik, a T-DNS két végén lévő ismétlődő szekvenciák segítségével. A megfelelő plazmiddal transzformált növényi sejtek antibiotikum jelenléte mellett szelektálhatóak.



15.2. ábra: *Agrobacterium*-közvetített génátvitel. A transzgenikus növények előállításának egyik típusa az *Agrobacterium tumefaciens* talajbaktériummal történő fertőzés, mely segítségével a növényi genomba beépülő T-DNS-ről történik a transzgén átírása. A módszer jól alkalmazható olyan növényeknél, amit az *Agrobacterium* hatékonyan képes fertőzni (pl. dohány, paradicsom, burgonya). Szemléltető videó itt nézhető meg.

Az olyan növényeknél, amik nem érzékenyek az *Agrobacterium* által közvetített folyamatokra, gyakran alkalmazott génbeviteli módszerek a **génpuska**, a direkt **génátvitel protoplasztokba (pl. elektroporációval)** és a **mikroinjektációs** génbevitel. A **protoplaszt** (izolált növényi sejt sejtfal nélkül) transzformálása nagyon hatékonyan működik bármely felsorolt fizikai folyamaton alapuló technikával, hiszen nem szükséges biológiai vektor a transzgén közvetítéséhez, így elkerülhetőek a nem-kompatibilis gazda problémák. Homológ rekombinációval pedig specifikusan építhető be a transzgén az ismert növényi genomba. A protoplaszt transzformálásával az egyetlen probléma, hogy a növény regenerációja protoplasztból egy rendkívül kényes folyamat, ami sok problémával jár. A **génpuska** (ld. 15.3. ábra) számos előnnyel rendelkezik (egyszerű használat, egy lövésből sok transzformált sejt, a legkülönbözőbb szövetek/sejtek transzformálása, stb.), és akkor is hasznosnak bizonyul, amikor más módszer nem működik egy transzgenikus növény előállításánál.



15.3. ábra: **Növények transzformálása fizikai módszerekkel.** Protoplasztok, vagy sejtfallal rendelkező növényi sejtek transzformálására is alkalmas a génpuska, vagy a mikroinjektor, ami a DNS-t nagy nyomással juttatja át a sejtfalon és a sejtmembránra.

Az első transzgenikus szója létrehozásánál a génpuska sokkal sikeresebbnek bizonyult, mint az *Agrobacterium*-közvetített transzformálás. A módszer hátránya, hogy a DNS mennyisége nem szabályozható, illetve a transzformált növényi vonalak kis részénél történik csupán genomális integráció, a citoszolikus plazmid-transzgének pedig hosszútávon nem fenntarthatóak. Protoplasztok illetve embriók **mikroinjektálása** növényeknél is alkalmazható, azonban a technológia nehezebb, és kevésbé hatékony, mint a génpuska.

15.1.2.2. Transzgenikus növények a mezőgazdaságban és az élelmiszeriparban

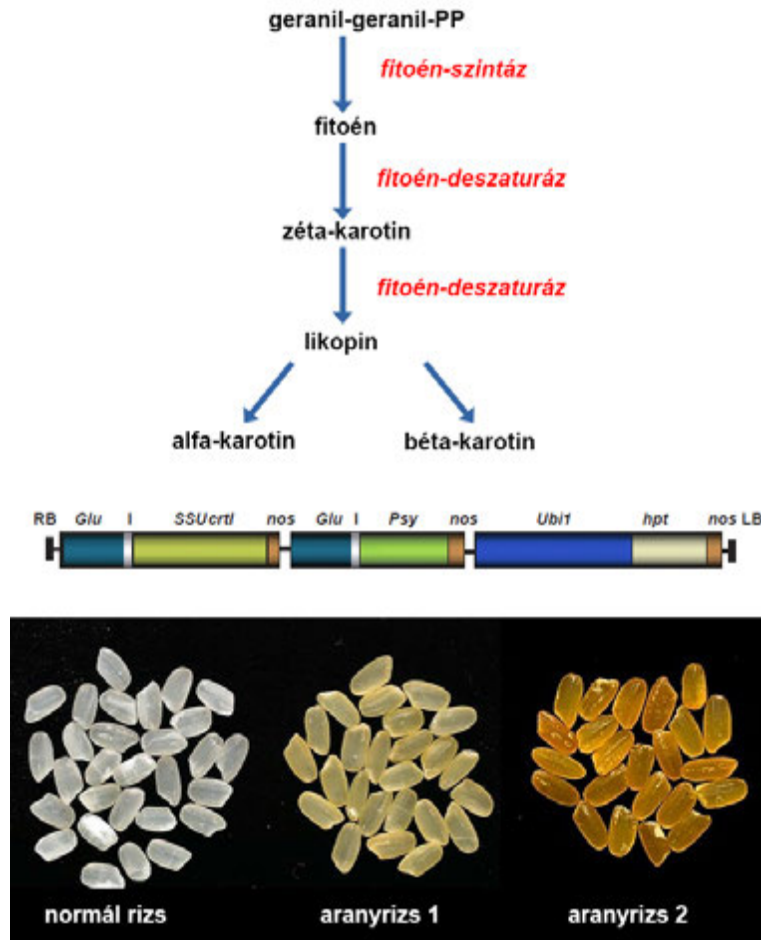
Transzgenikus növények első szabadföldi termesztése Franciaországban és az Egyesült Államokban 1986-ban kezdődött, növényirtószerekre rezisztens dohány. 1987-ben alapították az első céget, mely *Bacillus thuringiensis* baktériumból izolált, rovarra toxikus fehérjét termelő transzgént tartalmazó dohánynövényt állított elő (**Bt dohány**), szintén *Agrobacterium*-közvetített technológiával. A dohánymoly lárváját öld háttérként *Bacillus thuringiensis berliner* baktériumban termelődő endotoxin (*bt2*) génjének a bevitelével érték el. A baktérium citoszolikus plazmidjából izolált *bt2* vagy *cry* gén (*crystal* protein) terméke egy protoxin, mely kristályként jut be a rovar emésztőrendszerébe, itt oldatba kerül, majd proteázok kihasítják az aktív toxint belőle. Később ezen baktériumok plazmidjaiból többféle *cry* gént izoláltak, így mára a biológiai rovarölő szerek széles spektrumával rendelkezünk (pl. *Lepidoptera*, *Diptera*, *Coleoptera* vagy *Hymenoptera* fajok ellen). A Bt növények alkalmazhatósága megosztja a kutatókat. Mivel az ilyen növényeket nemcsak a kórokozók, hanem a közelben élő rovarpopulációk is elfogyaszthatják, a toxin ezek populációit is pusztíthatja. Egyesek szerint a Bt kukorica (amelyet ma már nagy területen termesztnek) pollenje áll királylepke populációk, illetve méh kolóniák nagyfokú pusztulásának háttérben, azonban egyikre sincs egyértelmű bizonyíték. Számos kutatás vizsgálja a vad típusú és a Bt kukorica

génállományának keveredését is, a transzgének átjutását a vad típusba. Aggodalomra adhat okot az is, hogy egy Bt kukorica a növényvédőszerekben megengedett maximum toxin mennyiség ezerszeresét termelheti.

Az első GM termény az Egyesült Államokban 1994-ben került forgalomba, egy hosszabb eltarthatósági idejű **paradicsom** formájában (FlavrSavr néven). A génmódosított paradicsomot antiszensz RNS technológiával állították elő. A poligalakturonáz enzim génjét antiszensz RNS szállal csendesítették. A poligalakturonáz a paradicsom érésekor a sejtfalban lévő pektint bontja le, így a paradicsom érés közben puhul, amitől azonban sérülékenyebb lesz, és jobban kitett a fertőzéseknek. Ezért a hagyományos paradicsomot éretlenül szedik le, mivel addig szállítható, amíg kemény és kevésbé sérülékeny. Ez a génmódosított paradicsom azonban nem hozta meg a várt eredményeket. Eltarthatósága megnövekedett ugyan, de éretten mégsem volt elég kemény a szállításhoz, így ugyanúgy zölden kellett leszedni, mint a hagyományosat, ezért vissza is vonták a forgalomból. Szintén egy példa a fogyasztói igények kielégítésére a nem-barnuló alma, mely jelenleg piacra kerülés alatt áll Kanadában és az Egyesült Államokban. Ez a génmódosított alma a barnulást okozó polifenol-oxidáz enzimet csak kis mértékben termeli.

Számos egyéb transzgenikus növényt hoztak forgalomba (paradicsom, krumpli, dohány, kukorica, búza, szója, papaja, citrusfélék, gyapot, repce), melyek ellenállnak a bromoxinil vagy glifozát tartalmú gyomirtóknak, különböző rovarirtóknak, vírusrezisztensek vagy rovarölő hatásuk van.

2000-ben állították elő az első olyan transzgenikus növényt, az **aranyrizst**, melynek a tápértékét sikerült javítani. Nevét sárgás színéről kapta, mely β -karotin termelésének köszönhető. A cél a főleg Afrikát, Indiát és Ázsia déli részét érintő A-vitamin hiány miatti nagyfokú elhalálozás (és vakság) visszaszorítása volt. Habár humanitárius célokat szolgált, erőteljes negatív visszhangot váltott ki környezetvédők és anti-globalizációs aktivisták körében, elsősorban transzgenikus volta miatt. A β -karotin előállításához, ami az A-vitamin prekursora, két enzim hiányzik a rizs terméséből (azonban a leveleiben megvan, így az termel is β -karotint): a fitoén-szintáz és a fitoén-deszaturáz (**ld. 15.4. ábra**).



15.4. ábra: **Az aranyrizs előállítása.** A rizs endospermiumban a béta-karotin szintézis útvonala a geranyl-geranyl-pirofoszfátnál megáll. A pirossal jelölt két enzim bevitelével azonban biztosítja a teljes bioszintetikus út lefolyását a β -karotinig. A transzgén konstrukció: RB (*right border*): T-DNS szekvencia; Glu: rizs endospermium specifikus glutein promóter; SSUcrtI: borsó ribulóz bisz-foszfát-karboxiláz kis alegység transzipeptid, a kloroplasztisban való lokalizáció biztosítására; nos: nopal-szintáz terminátor; Psy: fitoén-szintáz; Ubi1: kukorica poliubiquitin promóter; hpt: szelekciós marker gén; LB (*left border*): T-DNS szekvencia. Az ábra alsó részén a vad típusú, az első és második generációs aranyrizs fotója (az alsó ábrarész forrása: www.goldenrice.org/Content2-How/how1_sci.php)

Az előbbi génjét (*psy*) a *Narcissus pseudonarcissus* növényből, az utóbbit (*crtI*) az *Erwinia uredovora* talajbaktériumból izolálták. A két transzgént egy plazmidba klónozták, egy konstitutív, és a magban expresszálandó glutein promóter mögé. A plazmidot higromicin antibiotikum rezisztencia markergénnel történő szelekció után *Agrobacterium* transzformálták, így juttatták a transzgéneket rizs embriókba. Kutatások igazolták, hogy az emberi szervezet képes a rizsszel bevitt β -karotinból az A-vitamin hatékony előállítására. Pár évvel később egy új fajtáját állították elő, az aranyrizs 2-öt, ami 23-szor annyi β -karotint termelt, mint elődje. Ebben a változatban az eredeti aranyrizs *Narcissus psy* génjét a kukorica *psy* génjére cserélték, mely enzimaktivitása jóval hatékonyabbnak bizonyult. Mindezeket még egy sor aranyrizs változat követte, melyek E vitamint, vas-, cink-iont tartalmaztak. Annak ellenére, hogy jelenlegi kutatások szerint nem okoz semmilyen allergiát, nem hozható kapcsolatba betegségekkel, és egy bögre rizs fedezi a napi A-vitamin szükséglet 50%-át, még nem engedélyezett a piacra kerülése. Egyes kutatók szerint azonban aggályos, hogy nem tesztelték állatokon, így nem jósolható hosszú távú hatása a szervezet egészére, mivel a β -karotinból (amit a rizs túltermel) származó retinoidok bizonyítottan károsak az egészségre. β -karotin azonban természetes A-vitamin forrásként számos más élelmiszerben is van.

15.1.2.3. Transzgenikus növények a gyógyszeriparban, vegyiparban, a környezeti iparban

A transzgenikus növények szerepe a gyógyszerkutatásban is növekvő tendenciát mutat. Géntechnológiai eszközökkel olyan banánt állítottak elő, melyben hepatitisz-B antigént lehet termelni, illetve olyan dohányt, mely monoklonális ellenanyagot termel HIV ellen. Egyik termék sincs még forgalomban, de a GMO növényekben előállított rekombináns fehérjegyógyszereknek szerepe lehet a jövő személyre szabott gyógyászatában.

Ezen kívül a szerves vegyipar is nagy potenciált lát a transzgenikus növényekben. Sok esetben hatékonyabb a termelés, illetve a növények által előállított anyagoknak kedvezőbbek lehetnek a tulajdonságai, mint a kémiai szintézissel előállítottak. Számos laboratórium foglalkozik azzal, hogy növényekkel biopolimereket termeltesen, amilyenek különböző műanyagok, a természetes gumi vagy a poliaminok. Szintén hasznos lehet a kémiai szintézisekhez szükséges prekursorok előállítása növényekkel. A rostfehérjék (selyem, kollagén, elasztin, keratin, glutenin vagy rezilin) repetitív DNS szakaszainak kombinációjával, vagy akár szintetikus szekvenciákkal végtelen sokféle, különböző tulajdonságú biopolimer állítható elő.

A bioremediációban is fontos szerepük van a főleg bakteriális transzgéneket tartalmazó növényeknek. Az *Arabidopsis thaliana* olyan transzgenikus változatát állították elő, amely bakteriális enzimjei segítségével a robbanóanyagokból származó talajszennyezést képes lebontani. A TNT (2,4,6-trinitrotoluén) vagy RDX (hexahidro-1,3,5-trinitro-1,3,5-triazin) olyan mérgező xenobiotikumok, melyek a nem tökéletes robbantások után nagy mennyiségben, hosszú ideig jelen vannak a talajban, bejutva az ivóvizekbe is. Az Egyesült Államokban ez a típusú szennyezés igen nagy probléma. Nem transzgenikus növények is képesek a fitoremediációra, főlvé a szennyezőanyagot, vakuólumokban vagy a sejtfalukban tárolják azt (*Vicia faba*, *Zea mays*, *Catharanthus roseus*), ám lebontani nem képesek. Bioremediáció mikrobákkal is működhet, azonban számos esetben egy adott vegyület lebontásához szükségesek a növényi enzimek, így a két szervezet ötvözése bizonyult eddig a leghatékonyabbnak.

15.1.3. Transzgenikus állatok az alap kutatásban és a gyógyászatban

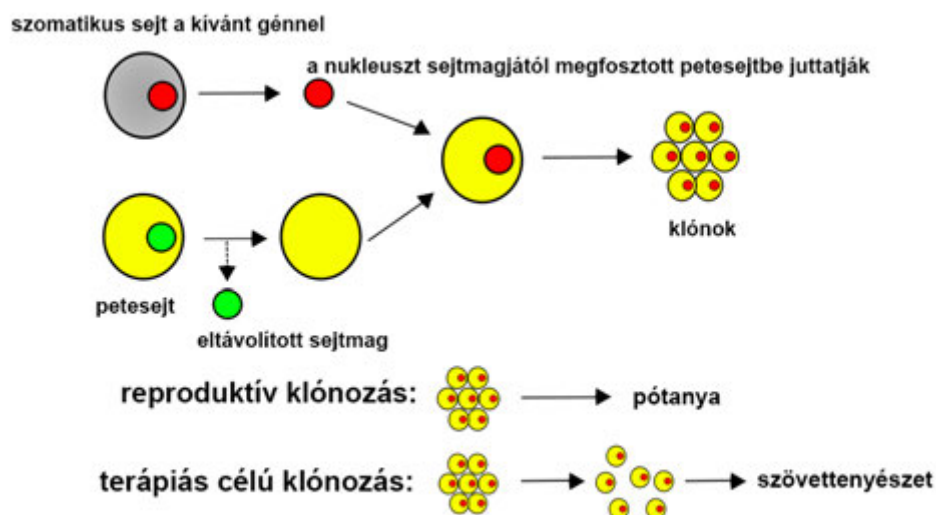
Az állatok genetikai módosítása egy viszonylag lassú, aprólékos és drága folyamat, azonban az újabb és újabb módszerek megjelenése egyre egyszerűbbé és pontosabbá teszi a technológiát. A genetikailag módosított állatok egy részét kutatási célokból hozzák létre, hogy a specifikusan megváltoztatott gének fenotípusra kifejtett hatását elemezzék. Így még ismeretlen gének funkciója, vagy azok mutációinak hatása, géninterakciók, vagy a génexpresszió helyei feltérképezhetőek. Genetikai módosítással olyan élőlényeket is létrehozhatunk, melyek érzékenyebbek bizonyos vegyületekre, ezt használják ki a gyógyszerkutatásban, hatóanyagok tesztelésénél. Transzgenikus állatok kiválóan alkalmazhatóak modellrendszerként emberi betegségek, illetve azok gyógyításának kutatására is. Más állatokat úgy módosítottak, hogy bizonyos tulajdonságaikat a gyógyászat vagy az ipar hasznosíthassa.

Az első transzgenikus állatot Brinster és Palmiter 1982-ben állították elő, amikor is egy egér megtermékenyített petesejtjének pronukleuszába mikroinjektálással patkány növekedési hormon (GH) transzgént juttattak (**ld. 1.5.2. fejezet**). A híres egér ikerpár a *Nature* folyóirat címlapjára is rákerült (**ld. 15.5. ábra**). 1985-re nyúl, bárány illetve sertés megtermékenyített petesejtjébe is sikerült mikroinjektálással transzgént bevinni. Megjegyzendő, hogy a patkány GH transzgént tartalmazó egér nem azért lett nagyobb, mert a patkány GH az állat méretét szabja meg, hanem az ún. dózishatás miatt, azaz több patkány transzgén épült be véletlenszerűen az egér genomba. A transzgént a nehézfémekkel szabályozott metallotionein gén promótere mögé építették be a rekombináns DNS konstrukcióba.



15.5. ábra: A patkány növekedési hormon transzgént tartalmazó egér (bal) és vad típusú testvére (jobb) a Nature folyóirat címlapján (Vol. 300, 1982. december 16 ; a Nature Publishing Group engedélyével)

A későbbiekben, a szomatikus nukleáris géntranszfer alkalmazásával több, mint 10 transzgenikus fajt állítottak elő (ld. 15.6. ábra). Ezzel a módszerrel született 1996-ban az első olyan bárányklón is (Dolly), amely felnőtt szomatikus sejt genomját örökölte, habár transzgént nem tartalmazott. Ezzel bizonyítást nyert, hogy bármely testi sejtünk genetikai anyagából a teljes szervezet klónozható.

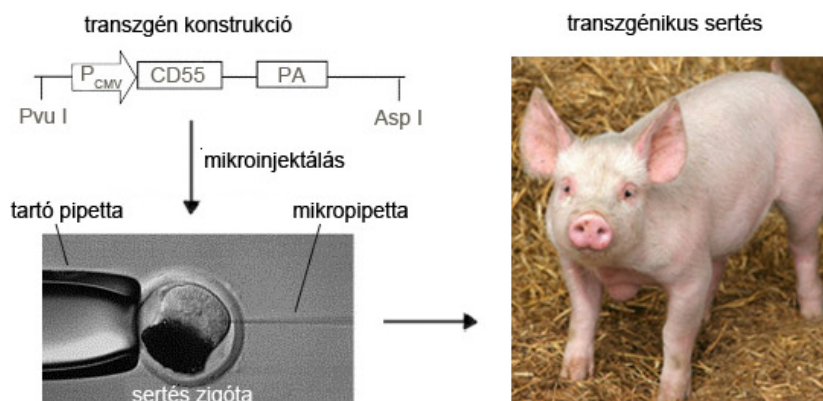


15.6. ábra: **Szomatikus nukleáris géntranszfer.** A módszer segítségével módosított genetikai anyagú donorsejtek szaporíthatóak. (Forrás: wikipedia.org, GNU Free Documentation License, szerző: Quelle).

15.1.3.1. Xenotranszplantáció és xenogén sejterápia a gyógyászatban

A szervtranszplantációra várók száma többszöröse a felhasználható humán szerveknek, és szomorú tény, hogy minden évben több ezer beteg hal meg megfelelő szervek hiányában. Ezt az igényt próbálja kielégíteni a **xenotranszplantáció**, amikor egy **nem humán élőlény szervét** ültetik be a **betegbe**. A sertés jó választásnak tűnt a kutatók számára, mivel szerveinek mérete hasonlít az emberéhez, az anatómiája és fiziológiája sem túl különböző, hatékonyan és egyszerűen szaporítható és fenntartható, illetve a technológia adott az immungenetikájának a megváltoztatására. Transzgenikus sertések előállításával a beültetett szerv azonnali kilökődését (HAR: *hyperacute rejection response*) és a napokkal későbbi akut vaszkuláris kilökődését (AVR: *acute vascular rejection*) szeretnék elkerülni. A hosszútávú kilökődés elkerülésére egyenlőre csupán az egész életen át tartó immunszuppresszánsok

szedése jelent megoldást. Bízató eredményeket mutat a HAR folyamán aktiválódott humán komplement rendszert szabályozó fehérjék termelése az átültetett sertés szervben. Az ember által termelt antitestek felismerik a sertés antigéneket, és az ellenanyag-antigén komplex aktiválja a komplement kaskádát, ami a membránkárosító komplex (MAC: *membrane attack complex*) létrejöttéhez vezet. Azonban a komplement kaskádát szabályozó fehérjék, mint a CD55 (DAF: *decay accelerating factor*) vagy CD46 (MCP: *membrane cofactor protein*) meggátolják a MAC kialakulását. Embriók mikroinjektálásával előállítottak DAF és MCP fehérjéket expresszáló transzgenikus sertést, mely sikeresen meggátolta az átültetett sertésvese kilökődését főemlősökben (15.7. ábra). A kilökődés megakadályozására egy másik stratégia, hogy meggátolják a sejtfelszíni antigének kifejeződését. Az antigének főleg az 1,3- α -galaktoziltranszferáz aktivitása révén létrejövő 1,3- α -gal-epitópok. Az α -gal allélt homológ rekombinációval ültették ki sertés embrionális fibroblaszt sejtekből, ami aztán nukleáris transzferrel került embriósejtekbe, és erre a génre nézve egészséges KO (*knock-out*) malacokat eredményezett.



15.7. ábra: **Xenotranszplantációra használt transzgenikus sertés előállítása.** A CD55 fehérjét termelő génszakasz előtt egy citomegalovírus promótere áll (P_{CMV}), ami biztosítja a gén transzkripcióját már az élet korai szakaszában. A poliadenilációs szakasz (PA) után, és a promóter szakasz előtt találhatóak a restrikciós enzimek hasítóhelyei (Pvu I, Asp I). (a módosított ábra forrása: Niemann & Kues, 2003, *Application of transgenesis in livestock for agriculture and biomedicine. Animal Reproduction Science*, 79(3), 291-317)

A szervtranszplantáció mellett a **xenogén sejterápia** is bízató eredményekkel szolgál a gyógyászat terén. A xenotranszplantációnál megismert, megfelelő transzgenikus állatokból (sertés, marha) történő sejtizolálás és beültetés megoldást nyújthat a humán őssejterápia mellett a különböző betegségek, rendellenességek vagy sérülések miatti elhalt szövetek regenerációjára. A xenogén sejterápia egyes betegségek esetében már klinikai tesztelés előtti fázisban van. Transzgenikus sertés embrionális idegsejteket Parkinson illetve Huntington kórbán szenvedő betegek agyába transzplantáltak, ahol azok dopaminerg neuronokká és glia sejtekké alakultak. A sertés eredetű neuronok a beültetés helyéről kiindulva kiterjedt axonhálózatot alakítottak ki a páciens agyában. Számos más példa létezik arra, hogy idegen eredetű sejtek képesek visszaadni a gazdaszervezet egyes szöveteinek funkcióját. Transzgenikus sertés májsejtek transzplantáció után a máj szöveteibe vándoroltak, és ott integrálódva működő LDL receptorokat expresszáltak, ezzel csökkentve egy magas koleszterinszintű nyúlörzsből a szérum koleszterinszintjét 30-60%-kal. A használt nyúlörzs a homozigóta öröklődésű hiperkoleszterinémia állati modelljeként szolgál. Ezzel el is érkeztünk a **transzgenikus állatok** következő rendkívül fontos felhasználásához, **az emberi betegségek állati modelleken** keresztül történő kutatásához. A donorsejtek sikeres genetikai módosítása, a xenogén sejterápia alkalmazása emberi betegségeket modellező állatörzseken sikeres terápiás megoldásokhoz vezethet a humán gyógyászatban.

15.1.3.2. Emberi betegségek állatmodelljei: KO (*knock-out*) egerek

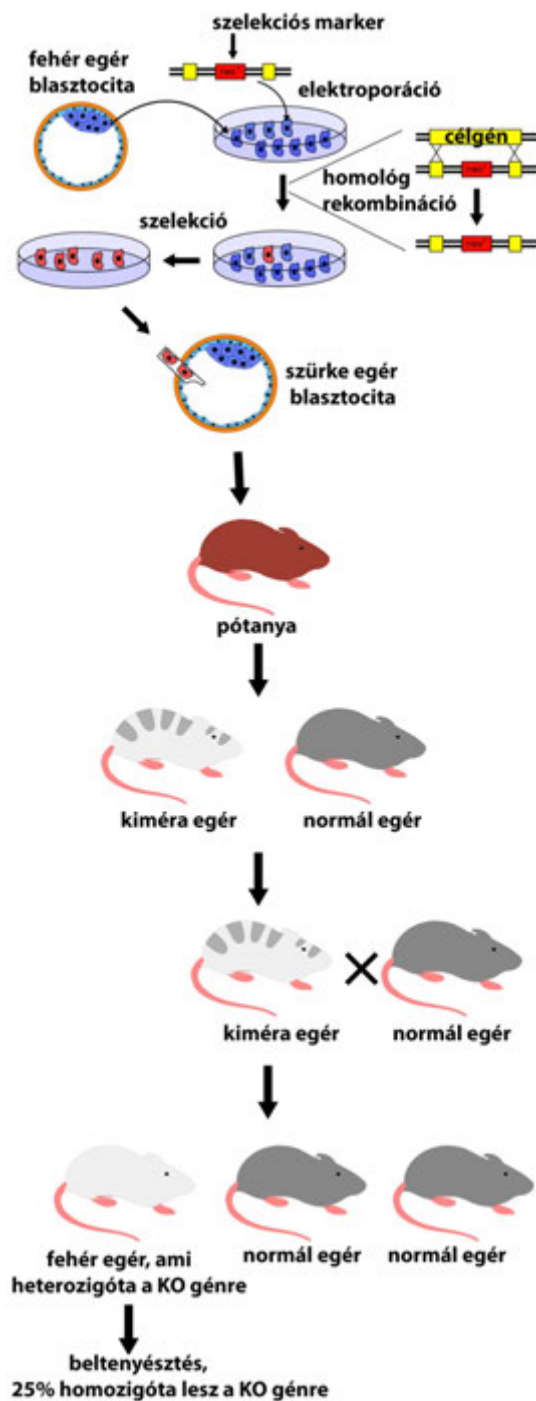
Knock-out (KO, génkiütött) egerek előállítása mára már általános módszerré vált, mellyel egy gén részét/egészét távolítják el vagy helyettesítik egyéb DNS szekvenciával (ld. 14.5.3. fejezet). A gén kiütésével vizsgálni lehet annak funkcióját, fenotípusát, illetve „utánozni” lehet funkcióvesztéses mutációkon alapuló emberi betegségeket. Az ember és egér génjei között nagyfokú a hasonlóság, ezért a különböző KO egerek jellegzetességeit vizsgálva, illetve a kialakult tünetegyüttesre kezeléseket kidolgozva közelebb kerülhetünk számos genetikai hátterű betegség gyógyításához. A legkülönbözőbb ráktípusok, túlsúlyosság, szívelégtelenség, cukorbetegség, köszvény, Parkinson kór, drogfüggőség, öregedés mind olyan területek, ahol a különböző KO egerek nagy szolgálatot tettek a kutatóknak.

Az első KO egeret 1989-ben állították elő Capecchi, Evans és Smithies (**ld. 1.5.2. fejezet**). Mára már több ezer KO vagy transzgenikus egértörzs létezik (MKMD: *Mouse Knock-out and Mutation Database*), és évente több millió egeret használnak fel különböző kutatásokhoz.

A célzott génkiütés, mint gén targeting módszer alapjait a **14. fejezetben** részleteztük. Ebben a fejezetben a KO egerek létrehozásának a lépéseit foglaljuk össze (**15.8. ábra**):

- a. A **kiütni kívánt gént (célgén)** génkönyvtárból izolálják, majd a gént módosítják (delécio, inszercio, specifikus/random mutagenesis). A génszakaszhoz egy **markergént** is fuzionáltatnak, ami általában egy antibiotikum vagy toxin rezisztencia génje, de lehet egyéb detektálható változással járó géntermék is (szín, fluoreszcencia). A legegyszerűbb, ha homológ rekombinációval kicserélik a markergént a kiütni kívánt génnel.
- b. Eger (fehér) **blasztocitából összejtet (ESC: embryonic stem cell) izolálnak**, amit *in vitro* tenyésztnek.
- c. A módosított génszakaszt **elektroporációval** juttatják be az összejtetekbe, ami **homológ rekombinációval** épül be a sejtek genomjába. A sikeres rekombinációs esemény igen ritka, ezért annak az esélye, hogy az összejt anyai és apai kromoszómája is tartalmazza a módosított gént, igen csekély. Így a legtöbb sejt, ahol a rekombináció megtörtént, heterozigóta lesz erre a génre nézve.
- d. Azokra a sejtekre, amelyek genomja tartalmazza a módosított gént, **markergénnel szelektálnak** antibiotikum/toxin jelenlétében, illetve a fluoreszcens jel alapján különböztetik meg őket.
- e. A **KO gént tartalmazó sejteket egy másik eger (fekete) blasztocitájába visszajuttatják**. Így a blasztocita kétféle sejtet tartalmaz: a módosított fehér egérből származót, és a nem módosított fekete eger eredetűt. A blasztocitát ezután beültetik egy álvemhes eger pótanya uterusába, ahol kifejlődhet a kiméra utód. Az utód bundáján valóban látható, hogy kiméra, a fehér eger eredetű sejtek utódai fehér bundát eredményeznek, míg a feketéé feketét.
- f. Néhány kiméra gonádja is lehet fehér eger leszármazott, így az azok által termelt ivarsejtek tartalmazzák a KO gént. Ezért ha a **kiméra egereket pároztatják vad típussal**, lesz olyan utód, ahol egy kópiában ugyan, de meglesz a KO gén. Ezek az egerek fehérek, de heterozigóták a KO génre nézve.
- g. Már csak **keresztezni kell a heterozigóta egyedeket egymás között**, hogy kapjunk homozigóta egereket a KO génre nézve, ahol egyáltalán nincs funkcionális géntermék.

Ezzel a módszerrel készült egy örökletes törpeséggel járó betegség, a Laron szindróma állatmodellje is. A Laron szindróma hátterében a növekedési hormon receptorának (GHR: *growth hormone receptor*) defektusa áll, aminek következtében a növekedési hormonokra nem reagál a páciens. A GHR funkcióvesztésének okát emberben nehéz vizsgálni, aminek etikai és gyakorlati okai is vannak. Homológ rekombináción alapuló célzott génkiütéssel ezért létrehozták a Laron egereket, melyekben a GHR és GHR-kötő fehérje génjét üttették ki. Az egerek fenotípusa tipikus Laron szindróma jellegzetességeket produkált (születés utáni növekedési hibák, majd törpeség, a GHR és kötőfehérjéjének hiánya, csökkent koncentrációjú inzulin-szerű növekedési faktor I és megnövekedett GH szint). Ezt a modellállatot azóta az öregedés kutatásához is használják, mivel azt figyelték meg, hogy a GH rezisztens állatok életideje meghosszabbodott.



15.8. ábra: **Knock-out egerek előállítása.** (Forrás: wikipedia.org. GNU Free Documentation licence; szerző: Kjaergaard)

Egy gén kiütése nem minden esetben eredményez életképes utódokat, a KO egérgének 15%-a letális, ezért legjobb esetben is csak embrionális korban tanulmányozható a hatás. Előfordulhat az is, hogy hiába nagyon hasonló az egér genomja az emberéhez, egy gén funkcióvesztése más fenotípust eredményez. Számos betegség modelljét egér helyett patkányra dolgozták ki, mely szaporítás terén lényegében azonos tulajdonságokkal rendelkezik, azonban jóval nagyobb, ami miatt könnyebb operálni vagy vért venni tőle. Fontos megjegyezni, hogy fiziológiai szempontból is ideálisabb választás a patkány. Pulzusa például csaknem a miénkkel azonos, míg az egér szíve 5-10-szer gyorsabban ver. Detoxifikáló enzimek is jóval közelebb állnak hozzánk, ezért a potenciális gyógyszerek farmakológiai hatását és toxikusságát jobb patkányon vizsgálni. Mindezek ellenére az egérmodellek is rendkívül

fontosak, hisz több rágszáló modell törzs birtokában jobban meg tudjuk becsülni, hogy az észlelt fenotípus rágszáló-specifikus, vagy általános az emlősökben.

Az első **transzgenikus patkányt** azonban csak 2008-ban sikerült létrehozni, melynek fő oka, hogy a KO egerek előállításához használt technológia patkánynál nem működött. Mivel a rekombinációs események gyakorisága igen kicsi, a kutatók egyszerűen nem találtak olyan embrionális őssejteket, ahol ez megtörtént volna. Mivel a szomatikus nukleáris transzfer módszer sem eredményezett KO patkányokat, maradt a kémiai mutagenézis, amikor is N-etil-N-nitrozourea (ENU) injekálásával random mutációkat indukáltak. Az ENU etil-csoportja csoporttranszfer reakcióval a DNS oxigén és nitrogén gyökeire kerül, ami *mismatch* bázispárosodást eredményez. Az ENU-val injekált himet vad típusú nőténnyel pároztatták, hogy mutáns utódokat kapjanak. Az ENU mutagenézis körülbelül egy mutációt okoz bármely génre nézve minden 200-700-adik ivarsejtben. Annak ellenére, hogy az ENU mutagenetikus hatása nagy, mivel kicsi a penetranciája, így csupán körülbelül 500 gén mutálódik el egy himben, és az összmutációk csak nagyon kis hányadának van fenotípusos hatása. Az ENU módszerrel előállított mutációk feltérképezése igen nehézkes, pénz- és időigényes, ezért ma már nem nagyon használják. Az áttörést a transzgenikus patkányok előállításában a Zn-ujj nukleázok (ZFN: *zinc finger nucleases*) és a transzkripció aktivátor-szerű effektor nukleázok (TALEN: *transcription activator-like effector nucleases*) jelentették, amik a DNS kettős szálán szekvenciaspecifikus hasításra képesek, így segítik a genom célzott módosítását (**ld. 14.5.2. fejezet**). A hasítással aktiválják a DNS-javító mechanizmusokat, a homológ rekombinációt és a törött végek közvetlen összeragasztását. A nem homológian alapuló közvetlen ligálás a törött végeknél általában mutációkhoz vezet.

A mobilis genetikai elemekkel (retrotranspozonok, transzpozonok) előállított mutagenézis bár random és nem célzott, mégis alapvető eszköze a transzgenikus emlősök előállításának. Előnye, hogy a gének teljes kiesését eredményező mutagenézis egyenesen az ivarsejteketben történik, a bevitt mutáció pedig stabilan fenntartható. A mutációk továbbá könnyen feltérképezhetők, így a KO patkányokból történő könyvtárak létrehozása viszonylag egyszerű feladat. KO vonalak és szövetspecifikus CRE-rekombinázt expresszáló patkányok pároztatásával feltételes (kondicionális) mutációkat is létre lehet hozni (további részleteket **ld. 14.5.4. fejezet**).

A rágszálók fiziológiája, anatómiája, életideje sok esetben jelentősen eltér az emberétől, ami miatt számos betegség modellezéséhez nem megfelelőek. Lassú lefolyású ráktípusokhoz, illetve neurodegeneratív betegségekhez a sertés, bárány vagy marha jobb modell, mivel velük a betegségek hosszabb ideig tanulmányozhatók.

Egy ritka szembetegség (PR: *retinitis pigmentosa*) sertésmoделljét, melyre a fotoreceptorok korai elhalása jellemző, mikroinjekálásal hozták létre. A transzgenikus sertésmoделl mutáns rodopszint expresszál, ami hasonló fenotípust eredményez az embernél megfigyelthez. A létrehozott moделltörzs segítségével lehetővé vált a betegség kezelésének folyamatos fejlesztése.

Míg a transzgenikus sertések igen elterjedtek az egyes humán betegségek modellezésében, addig a transzgenikus kutya, macska vagy főemlősök előállítása csak nemrégiben vált lehetővé. Az első transzgenikus kutya 2009-ben született, retrovírus-mediált géntanszferrel. A transzgén egy tengeri szellőrózsa piros fluoreszcens fehérje volt, mely random integrálódott a kutya fibroblaszt sejtjeibe. Ezután a fibroblaszt sejtekből szomatikus nukleáris transzferrel a transzgént petesejtbe juttatták, amit osztódás után ültettek be pótanyákba. A kutatók most azon dolgoznak, hogy a kívánt inszerciót specifikus helyre lehessen bevinni. Ha ez sikerül, akkor következő lépésként ösztrogén receptor KO kutyákon tanulmányozná a hormon termékenységre gyakorolt hatását. Ugyanebben az évben, mikor az első transzgenikus kutya is született, japán kutatók bejelentették, hogy vírus-közvetített géntanszferrel előállították az első stabil transzgenikus főemlős törzset. A transzgenikus vonal sikeresen szaporodik, és utódai is tartalmazzák a transzgént, ami egy zöld fluoreszcens fehérje (GFP: *green fluorescent protein*). Az újonnan előállított főemlősökben az egéren nem tanulmányozható betegségeket szeretnék vizsgálni, mint amilyen a Huntington kór vagy az infarktus. 2011-ben amerikai-japán kutatók sikeresen hoztak létre GFP transzgént tartalmazó macskát, mely a későbbiekben moделljé lehet az emberi HIV okozta AIDS betegségre, mivel a macska immundeficiencia vírus (FIV: *feline immunodeficiency virus*) a HIV rokona.

15.1.3.3. Transzgenikus állatok az élelmiszeriparban

Annak ellenére, hogy a transzgenikus növények jóval elterjedtebbek az élelmiszeriparban, számos példát találunk a GM állatok felhasználására is. Az ilyen, géntechnológiai módszerekkel módosított állati termékek egy része arra szolgál, hogy az adott termék emberi egészségre gyakorolt pozitív hatását fokozzák. Ilyen a csökkentett koleszterin tartalom, magas vérnyomást csökkentő angiotenzin-konvertáló enzim inhibitorok vagy immunrendszert stimuláló peptidok jelenléte az ételben. A húsipar nagy felfedezettje egy olyan transzgenikus sertésvonal, ami nagyobb

arányban termel telítetlen zsírsavakat, mint telítetteket, így húsa egészségesebb, és fogyasztásával csökkenthető az infarktus vagy a szívproblémák veszélye. Transzgenikus marhával a tejtermékek legkülönbözőbb változatait lehet előállítani: zsírszegény tej (lipidanyagcsere enzim modulációjával), hipoallergén tej (β -laktalbumin gén kiütésével), laktózmentes tej (α -laktalbumin gén kiütésével), vagy csecsemőtej (laktoferrin túltermeléssel). A tejben található egyik fontos összetevő, a kazein különböző típusai határozzák meg a tej fizikokémiai tulajdonságait. Ezért a kazein a tejminőség javításának az egyik fő célpontja. Az egyes összetevők változtatásának hatása, illetve mellékhatásainak mértéke egérmodelleken jól követhető.

15.1.3.4. Transzgenikus állatok a gyógyszeriparban: humán fehérjék előállítása

A terápiára használt humán fehérjék emberi vérből vagy szövetekből történő előállítása kis hatékonyságú, drága, időigényes és magában rejt a humán kórokozók terjesztésének veszélyét. Rekombináns baktérium vagy sejt kultúrák felhasználásával egyszerűbb a gyógyszerfehérjék termelése. Hátrányuk, hogy a baktériumok csak viszonylag egyszerű fehérjék termelésére képesek, a poszt-transzlációs módosításokra alkalmas mechanizmusok pedig limitáltak. Megfelelőnek bizonyul erre a célra többféle transzgenikus állat, a bárány, kecske, marha, sertés, sőt a nyúl emlőmirigye, mellyel sok fehérje előállítható. A rekombináns fehérje az állatból a tejjel távozik, így tisztítása sem jelent különösebb problémát. Ezen állatokban a glikozilációs mintázatok és a poszt-transzlációs módosítások is megfelelőek. Rekombináns fehérjék szekréciójára a tej mellett a vért, a vizeletet illetve egyéb testnedveket is felhasználják. Ígéretes lehetőség gyengén immunogén antigének elleni ellenanyag termeltetésére neonatális Fc-receptorra transzgenikus állatok (egér, nyúl) létrehozása (az ImmunoGenes cég fejlesztése).

Számos rekombináns fehérjét termelnek emlőmirigy specifikus promóterrel rendelkező génkonstrukciókkal. Az így termelt α -antitripszin, antitrombin-III vagy szöveti plazminogén aktivátor (TPA) jelenleg klinikai kipróbálás alatt áll. Transzgenikus nyulak emlőmirigye által termelt α -glükózidázt sikeresen alkalmazták a Pompe-kór kezelésére. Ez a betegség a glikogénraktározás rendellenességeihez vezet és halálos két éves kor alatt, amennyiben a beteg nem kap α -glükózidázt. Egy másik esetben biológiailag aktív laktoferrint állítottak elő transzgenikus szarvasmarha tejében, amivel profilaxist illetve számos fertőző betegséget lehet kezelni. A géntechnológia úton (baktériumban, élesztőben, növényi és állati sejtekben) előállított terápiás célú fehérjék száma gyorsan növekszik. A **15.1. táblázatban** összefoglaltuk a legfontosabbakat (részben az FDA által jóváhagyott, részben klinikai kipróbálás, részben még kísérleti stádiumban vannak)

A különböző baktériumok növekvő antibiotikum rezisztenciájával párhuzamosan egyre nő az újfajta antimikrobiális anyagok iránti igény. A kationos antimikrobiális peptidek (AMP) tulajdonságai igen sok elvárásnak megfelelnek: széles hatásspektrummal bírnak, mind a Gram+, mind a Gram- baktériumokat gyorsan elpusztítják, a klasszikus rezisztencia génnek nem hatnak rájuk, illetve különböző állatmodellekben is aktívnak bizonyultak. Az AMP specifikusan a bakteriális sejtmembránnal kerül közvetlen kapcsolatba, amiért a rengeteg anionos foszfolipid és az alacsony membránpotenciál a felelős. Úgy gondolják, hogy az AMP a membránt fizikailag roncsolja, illetve intracelluláris molekulákkal is kapcsolatba lép. Az élővilágban sokféle AMP-t fedeztek már fel, azonban ezek emberre akár toxikusak is lehetnek (pl. rovar AMP).

A bemutatott sikerek ellenére azt is fontos megjegyezni, hogy jelenleg nem minden fehérjét lehet transzgenikus állatok felhasználásával előállítani. Az eritropoetin termeltetésére irányuló erőfeszítések például mind szarvasmarhában, mind nyúlban kudarcot vallottak, utóbbi esetben a géntermék toxicitása miatt. A példából látszik, hogy a technológiának jelenleg még vannak korlátai.

15.1. táblázat: Rekombináns fehérjék terápiás használatra

<i>fehérje neve</i>	<i>funkció és/vagy alkalmazás</i>	<i>rekombináns fehérje forrása</i>
Szöveti plazminogén aktivátor (TPA)	plazmin aktiváció, vérrögök feloldása; szívinfarktus kezelése	tr-kecske
VIII-as véralvadási faktor	véralvadáshoz szükséges; hemofiliás egyének kezelésére	tr-juh, CHO és HEK (humán) sejt kultúra,
granulocita kolónia stimuláló faktor (G-CSF)	fehérvérsejt növekedési faktor, immundeficienciák fertőzések gyógyszere	<i>E. coli</i> , CHO sejt kultúra
antitripszin	elasztáz inhibitor; emfizéma kezelése	tr-juh,
CFTR	cisztás fibrózis transzmembrán regulátor; CF betegek kezelése	tr-juh
hemoglobin	mesterséges vérhez	tr-sertés
protein C	antikoaguláns; vérrög feloldása	tr-sertés
inzulin	Diabetes kezelés	<i>E. coli</i> , élesztő

Egy érdekes új fejlesztés a transzkromoszomális állatok létrehozása. A humán immunglobulin teljes nehéz és könnyűlánc szekvenciáját tartalmazó humán mesterséges kromoszómát (HAC: *human artificial chromosome*) marha fibroblaszt sejtekbe juttatták, amely sejteket ezután nukleáris transzferre használtak. A transzkromoszomális utódok a vérükben humán immunglobulint termeltek. Egy ilyen rendszer nagy előrelépés lehetne a humán poliklonális ellenanyagok termelésében. Szintén ígéretes génterápiás kísérletek folynak a Duchenne izomdisztrofia (MDM) HAC-vektorral történő „génjavítására”. Az még a jövő kérdése, hogy az extra kromoszóma mennyire marad fenn a következő generációkban, és mennyire stabil a róla történő expresszió.

15.1.3.5. Transzgenikus állatok a környezetvédelem szolgálatában

Az Enviropig néven elhíresült transzgenikus sertést környezetvédelmi okokból hozták létre. A gabonafélékben található foszfort a sertés nem képes lebontani, így az nagy mennyiségben jut a talajba, illetve ivóvizekbe az ürülékkel együtt távozva. A gazdák eddig foszfort bontó fitáz enzimet is tartalmazó takarmány vásárlására voltak kötelezve, hogy csökkentsék környezetük foszfát szennyezését. Az Enviropig tartalmazza a fitáz gént, ami az állat nyálában termelődik. Az eredetileg *E. coli* baktériumból származó fitáz gént egy egerből származó konstitutív promóter mögé klónozták be, amit a sertés kromoszómájába mikroinjektálással juttattak be. A módszerhez nem volt szükség virális géntranszferre vagy marker génre, a transzgén a genom egy specifikus helyére épült be, ahol stabilan fennmaradt.

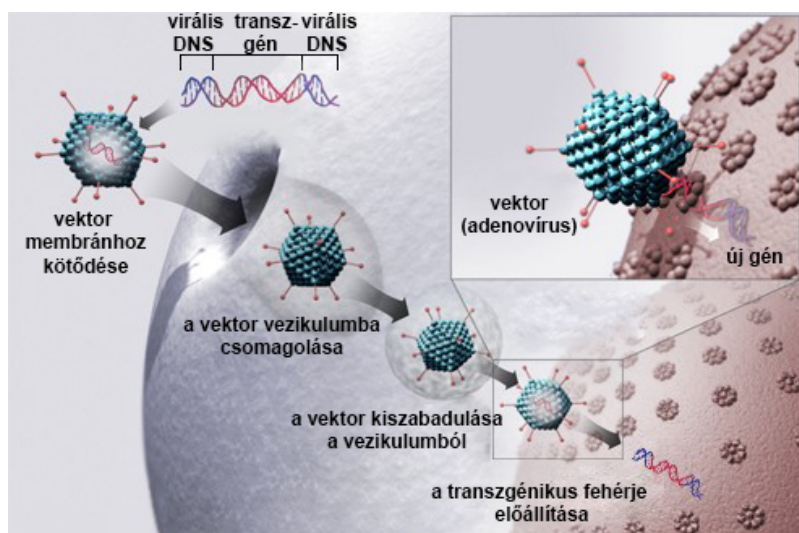
15.2. Générápia

A générápia célja, hogy genetikailag hibás testi sejtek genomját módosítsa emberben, a helyes gén bevitelével hozva helyre az „elromlott” gén szerepét. Az emberi zigóta, vagy az ivarsejtek genetikai módosítása etikai normákat sért, ezért a générápia kifejezetten a nem öröklődő, szomatikus sejtek genetikai állományát célozza meg. Kétféle szomatikus sejtípust használhat a générápia: őssejteket és differenciált, nem osztódó sejteket. Az őssejtek esetében a transzgént hordozó vektor RNS vírus (retro- vagy lentivírus) az, ami beépül az őssejt genomjába, így osztódást követően is jelen van az utódsejtekben. Générápiára jelenleg kétféle forrásból származó őssejtet használnak, az egyik a csontvelőből származó hematopoetikus őssejt, a másik egy, a bőr alsó rétegében található őssejtípus. Ezek az őssejtek viszonylag könnyen hozzáférhetőek, és *ex vivo* is fenntarthatók. A générápia másik célpontja - differenciált, nem osztódó sejtek - esetében, a vektor nem kell, hogy beépüljön a genomba, így itt **DNS vírusokat** használnak (**adenovírus vagy adeno-asszociált vírus**) **transzgén hordozóként**. A legújabb irányzat, hogy vírus-alapú génbejuttatás helyett a transzgént egy halak genomjában megtalált és „felélesztett” **transzpozon (ugráló gén)** viszi be a gazdasejt genomjába (ld. 14.5.5. fejezet). Egy „alvó” hal transzpozonból készült a ***Sleeping Beauty*** (Csipkerózsika) névre hallgató, humán sejtek genomjában „ugráló” transzpozon, amivel transzgének vihetők be humán sejtekbe (két magyar kutató, Izsvák Zsuzsa és Ivics Zoltán eredményei).

A générápia alkalmazhatósága sok akadályba ütközik. A technológia meg kell, hogy küzdjön a transzgén bejuttatásának gyenge hatékonyságával, az alacsony szintű fehérje expresszióval, az integrációval járó mutagenézis veszélyével és a vektor vagy a fehérje által indukált immunválasszal. A virális vektorok használata is veszélyt

rejteth magában. Bár a vektor nem tartalmaz patogénitási faktorokat, egy élő szervezetbe kerülve mindig fennállhat annak a lehetősége, hogy a betegséget okozó tulajdonságok visszatérnek.

Az **első génterápiás kísérletet 1990-ben** végezték az Egyesült Államokban, egy adenzin-deamináz- és súlyos kombinált immundeficiencia (SCID: *severe combined immunodeficiency*) betegségekben szenvedő páciens esetében. Azóta szerte a világon pár ezer génterápiát alkalmazó klinikai tesztől számoltak be, különböző betegségek sikeres gyógyításával kapcsolatban (X kromoszóma kapcsolt SCID, adrenoleukodisztrófia, krónikus limfoid leukémia, Parkinson kór). 2012-ben az Európai Bizottság elfogadta az első génterápián alapuló kezelést hivatalos gyógymódként, mely 2013 végétől már széleskörűen alkalmazható lesz az Európai Unióban és az Egyesült Államokban egyaránt. A Glybera márkanéven forgalomba kerülő génterápiás szer a súlyos hasnyálmirigy gyulladását okozó lipoprotein-lipáz deficienciában szenvedőknek nyújt gyógyulást. A humán lipoprotein-lipáz (LPL) génjét szállító vektor az adeno-asszociált vírus 1-es szerotípusa (AAV1), amely a vázizom sejteket célozza meg (**ld.15.9. ábra**). A vázizom sejtek az LPL fő termelési helyei. A kezelést követő 3-12 hétben a vérben lévő zsírok koncentrációja jelentősen lecsökken. Az AAV1 alkalmazásának nagy előnye, hogy hiányoznak a patogénitási faktori és nem épül be a genomba, így nem okoz random mutagenézist, ami a rák kockázatát növelné. Az AAV1 nagyon alacsony immunválaszt indukál, ami általában semlegesítő antitestek termelésében merül ki, és alacsony a citotoxikusságát okozó válaszreakció is.



15.9. ábra: **Adenovírus-közvetített génterápia.** A vírusvektor specifikusan felismeri az izomsejteket, amikbe receptor mediált endocitózissal jut be. A sejt belsejében az endocitotikus vezikulum lebomlik, majd a vírusvektor kiszabadulása után a sejtmag pórusaihoz asszociálódik. A sejtmagba a transzgén a vírus DNS-sel együtt jut be. (Forrás: US National Library of Medicine).

A mai normák szerint a humán génterápia nem érintheti az ivarsejteket, illetve a transzgén generációról generációra nem öröklődhet. A génterápia használatakor alkalmazott alapvető gondolkodásmód, a Weismann-elv értelmében az információ mindig az ivarsejtek felől áramlik a belőlük kialakult szomatikus sejtek felé, és soha nem fordítva. Azonban ez nem mindig történik így. Az RNS alapú retrovírusok reverz transzkriptáz segítségével épülnek be a genomba, és aktivitásuk révén gyakran szállítanak géneket egyik sejtbe a másikba, vagy akár egyik fajból a másikba (horizontális géntranszfer). Ha egy szomatikus gént szállító retrovírus egy ivarsejtet fertőz meg, a szomatikus génnel együtt beépülhet annak genomjába. Fontos megjegyezni, hogy az emberi szervezetben számos endogén retrovírus létezik, mely igen fontos lehet a génterápia jövője szempontjából.

15.3. A transzgenikus élőlények felhasználásának tudományetikai és környezetbiztonsági kérdései

A géntechnológiai módszerekkel történő transzgenikus élőlények előállítása szigorú szabályozás alatt áll, mely kiterjed a kísérletes munkát megelőző engedélyeztetési eljárásra, a laboratóriumban felhasználásra kerülő

protokollokra illetve a transzgenikus élőlény későbbi sorsára. Ezek a szabályok országról országra változhatnak, kiváltképp az emberi fogyasztásra szánt transzgenikus élelmiszerek felhasználásának és piacra kerülésének tekintetében, ahol a legszigorúbb szabályozás az Európai Unióban van. Az EU-ban – éles ellentétben például a GMO-t sokkal inkább támogató Egyesült Államokkal – minden genetikailag módosított élelmiszer új élelmiszernek minősül, ebből adódóan egyedi engedélyeztetési eljárásra kötelezett. Magyarországon az alkotmány tiltja a GM növények szabadtéri termesztését.

A transzgenikus élelmiszerek termesztésének és fogyasztásának a kérdése éles vitákat vált ki világszerte, ahol a tudományos érvek mellett sokszor az érzelmek és a politika is jelentős szerephez jut. A világ növekvő emberi populációjának az élelmezése komoly probléma, melyet az egyre szélsőséesebb és kiszámíthatatlanabb időjárás tovább súlyosbít. Fontos továbbá, hogy ha fizikai értelemben képesek is vagyunk az emberiség számára szükséges mennyiségű élelmiszer megtermelésére, azt olyan költséggel és hatékonysággal kell tennünk, hogy az mindenki számára hozzáférhető és megfizethető legyen. Amíg egy átlagos észak-amerikai lakos a jövedelmének 12%-át költi csupán élelmiszerre, addig ez a szám Magyarországon 27%, de Afrikában és Ázsiában akár 60-80% is lehet. Ha az élelmiszer árak tovább nőnek a fejlődő országok emberei egész egyszerűen nem lesznek képesek élelmet vásárolni. Ilyen körülmények között a GM-növények előnyei, mint a magasabb terméshozam vagy a kiszámíthatóbb termelés felértékelődnek.

Nem szabad azonban megfeledkezni a GM élelmiszerek termesztésének és fogyasztásának környezeti és egészségügyi kockázatairól sem. A GM-élelmiszerek fogyasztásával kapcsolatosan nincs olyan dokumentált eset, ami azt bizonyítaná, hogy a génmódosítás, mint technológia bármiféle egészségügyi kockázatot jelentene. Ettől függetlenül a transzgén által kódolt fehérje kiválthat allergiás reakciókat az arra érzékeny emberekben, amivel számolni kell.

A transzgenikus emlősök, kiváltképp a főemlősök kísérleti illetve gyógyszeripari felhasználása is számos vita forrása, melyben az ellenzők az állatok „felesleges” kínzása ellen tiltakoznak. Nem szabad azonban megfeledkeznünk arról, hogy az állatkísérletek mással nem helyettesíthetők, így szükségesek. Az állati áldozatok minimalizálása illetve a velük való kíméletes bánásmód minden kutató törvényi és erkölcsi kötelessége.

A humán génterápia - korunk nagy lehetőségekkel kecsegtető gyógyászati alkalmazása - bár még gyerekcipőben jár, mégis rengeteg etikai kérdést vet fel. Hogyan lehet különbséget tenni a génterápia „jó” és „rossz” felhasználása között? Ki dönti el, hogy mely tulajdonságok normálisak, és mik a rendellenesek? Megengedhető-e az alapvető képességeink génterápiával való fokozása, úgy mint fizikai teljesítő képesség vagy intelligencia? Bár a jelenlegi szabályozás szerint a humán génterápia nem érintheti az ivarsejteket, illetve a transzgén generációról generációra nem öröklődhet, nem tudhatjuk, hogy e tekintetben mit hoz a holnap.

15.4. További olvasnivaló a fejezethez

Fraley, RT, et al. (1983) Expression of bacterial genes in plant cells. *Proceedings of the National Academy of Sciences* 80: 4803-4807.

Walsh, G (2005) Therapeutic insulins and their large-scale manufacture. *Applied microbiology and biotechnology* 67: 151-159.

Kues, WA., and Niemann, H (2004) The contribution of farm animals to human health. *TRENDS in Biotechnology* 22: 286-294.

Yang, Y-G, and Sykes, M (2007) Xenotransplantation: current status and a perspective on the future. *Nature Reviews Immunology* 7: 519-531.

Sheridan, C. (2011). Gene therapy finds its niche. *Nature Biotechnology* 29: 121–128.

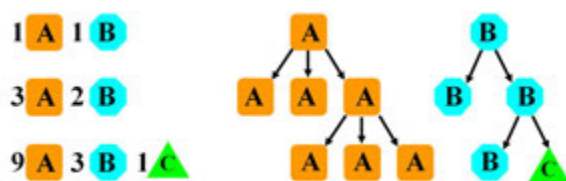
Prickett M, Jain M. (2013) Gene therapy in cystic fibrosis. *Transl Res.* 161:255-64.

Kay MA (2011) State-of-the-art gene-based therapies: the road ahead. *Nat Rev Genet.* 12(5):316-28.

16. fejezet - Irányított evolúciós technikák

A molekuláris biológiai módszerek fejlődése lehetővé tette a gének klónozását és az irányított mutagenézist, azaz egyes aminosavak célzott kicserélését, minek köszönhetően ma már a fehérjék aminosav sorrendje szabadon változtatható. A kutatók mutagenézisen alapuló technikák sorát fejlesztették ki. A fehérjék vizsgálatának **klasszikus, fehérjemérnöki megközelítése egyedi fehérje-mutánsok vizsgálatán alapul**. Az ilyen vizsgálatok során a fehérje egy vagy néhány pozíciójában megváltoztatják az eredeti aminosavat, majd ezeknek a mutáns fehérjéknek a tulajdonságait hasonlítják össze a természetben is előforduló, úgynevezett vad típusú fehérjével. A mutációk szerkezetre és funkcióra gyakorolt hatásán keresztül igyekeznek a vad típusú aminosavak szerepét feltárni. Hamar kiderült azonban, hogy a lehetséges változatok nagy száma miatt az összes variánst nem lehet előállítani. A fehérje minden pozíciót érintő teljes feltérképezésének gyakorlati akadályai vannak, már egyedi mutánsok vizsgálata esetén is, a mutációk kombinációjáról nem is beszélve. A sok egyenkénti mutáns előállítása, tisztítása és vizsgálata igen idő- és költségigényes. A vizsgálatok eredményei alapján felmerülő újabb kérdéseket megválaszolni, a hipotéziseket alátámasztani csak egy, de általában több újabb mutáns előállításával és vizsgálatával van mód.

Áttörést jelentett az irányított evolúciós módszerek megjelenése. Ezek a módszerek olyan esetekben alkalmazhatók egyszerűen, amikor a vizsgálat tárgya valamely adott molekulához szelektíven kötődő fehérje. Az evolúció működésbe lépésének mindössze néhány alapvető feltétele van: a szaporodás képessége, az öröklődés (az utód hasonlít a szülő(k)re) és a változatosság (az utód nem tökéletes másolata a szülő(k)nek). Abban a rendszerben, amelyben a feltételek teljesülnek, evolúciós folyamatok indulnak be abban az esetben, ha a tagok eltérő tulajdonságai befolyásolják azok szaporodási sikerét, azaz a rátermettséget. Egy ilyen rendszerben az adott környezeti feltételek között legjobban teljesítők szaporodnak a legnagyobb mértékben, utódaik aránya a következő generációban nagyobb lesz (**ld. 16.1. ábra**). Az ábrán bemutatott rendszerben a kiinduló helyzethez képest az A egyedek aránya megnőtt, és B mutációjával megjelent a C típus.



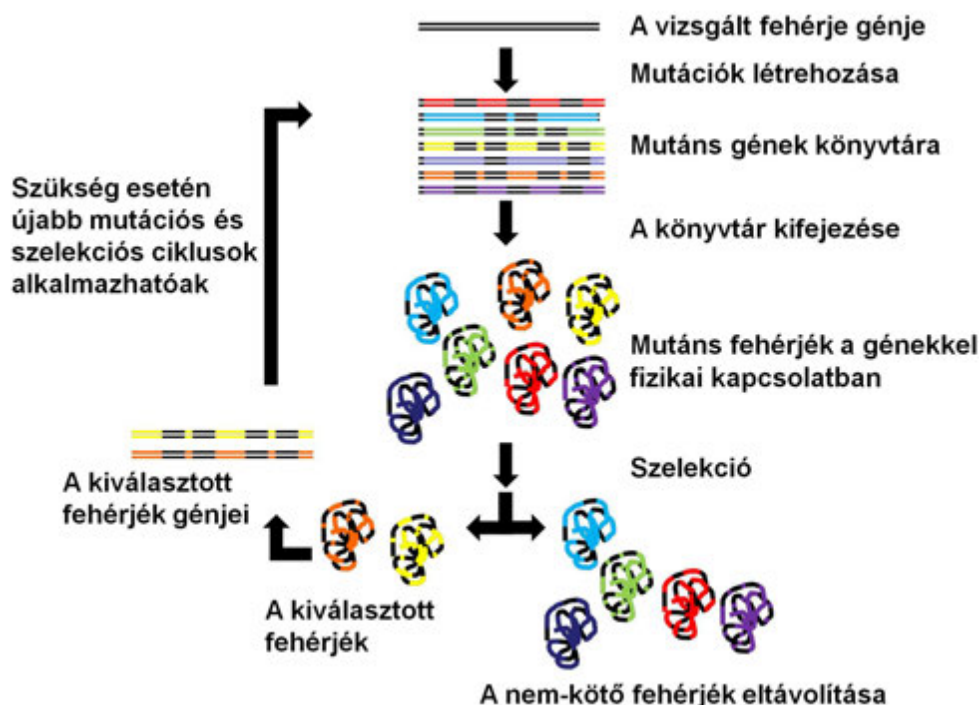
16.1. ábra: Az evolválódó rendszerek általános ábrázolása

A folyamat, ha hosszú idő (illetve elegendően sok szaporodási ciklus) áll rendelkezésre, a változatok bámulatos sokféleségét képes létrehozni, mely változatok rendelkeznek az adott környezetben való fennmaradáshoz és a tovább szaporodáshoz szükséges adaptációkkal. Erre a földi élőlények diverzitása a látványos bizonyíték. Az irányított evolúciós módszerek az evolúcióhoz hasonlóan működnek, és a természetben előforduló folyamatokat utánozva segítik a kutatót. Működésük a fehérje nagyszámú egyedi mutánsának, a változatosságnak (**variánsok könyvtára**) előállításán és az így létrehozott sokaságból a szelektív kötésen alapuló **szelekción** nyugszanak. A kutató dönti el, hogy a fehérje mely részeinek mekkora változatosságát hozza létre, és, hogy milyen tulajdonságra szelektál, innen az irányított jelző.

16.1. Az irányított evolúciós megközelítésről általában

Az irányított evolúciós technikák nagy előnye, hogy a mutánsokat nem kell egyenként előállítani, tisztítani és vizsgálni. A génkönyvtárat különböző, PCR-en és/vagy degenerált szintetikus oligonukleotidok használatán alapuló módszerek segítségével, egy reakcióban állítják elő. **A génkönyvtárat úgy fejezik ki fehérjék formájában, hogy azok a génjeikkel fizikai kapcsolatban maradnak.** A szelekció során a kötőpartnerrel összehozva kiválaszthatók a sokaságból azok a fehérje változatok, melyek teljesítik a szelekciós feltételeket. Egyedi mutánsok vizsgálata

helyett ezen, funkcióra szelektált gének szekvenciájának analízisével deríthető ki, hogy a vizsgált fehérjén milyen aminosavak szükségesek vagy alkalmasak a funkció ellátására. A **könyvtárkészítés és a szelekció során tulajdonképpen egy térben zajlik sok millió egyedi kísérlet**, melyeket egyenként képtelenség volna elvégezni. Így sokkal több mutáns megvizsgálására van lehetőség, mint egyedi változatok előállítása esetén. A módszer általános sémáját a **16.2. ábra** ismerteti.



16.2. ábra: Az irányított evolúciós eljárások általános ábrázolása

A könyvtárkészítés során a mutációk aránya és típusa bizonyos határok között szabadon, illetve a célokra megfelelően választható meg. A szelekciós feltételeket, a szelekció körülményeit szintén a vizsgálat céljától függően lehet beállítani. Ennek köszönhetően lehetőség van nagyobb affinitású kölcsönható partnerek, vagy enzimek esetén nagyobb katalitikus aktivitású változatok előállítására éppúgy, mint az eredetitől eltérő partner kötésére, illetve átalakítására képes fehérjék kifejlesztésére. Az így előállított, korábban nem létező fehérjék, illetve enzimek egy részét ipari folyamatokban vagy terápiás célokkal használják, ami a megközelítés óriási gyakorlati jelentőségét is mutatja.

Az irányított evolúciós módszereknek is vannak ugyanakkor korlátai. A könyvtárkészítési és szelekciós eljárások csak bizonyos varianciájú könyvtárak áteresztésére képesek, nagyobb könyvtár alkalmazása esetén könnyen elveszhetnek véletlenszerűen olyan változatok is, melyek teljesíthetnék a szelekciós feltételeket. Olyan könyvtár létrehozásának, ami egy adott hosszúságú polipeptid összes lehetséges variánsát tartalmazza, már elvi akadályai is lehetnek a lehetséges variánsok számának exponenciális növekedése miatt. Egy 60 aminosavból álló polipeptid esetében a lehetséges variánsok száma 20^{60} , ami meghaladja az univerzumban létező elemi részecskék számát. Ezért ezen módszerek alkalmazásakor sincs lehetőség a fehérje összes aminosav pozíciójának teljes körű vizsgálatára. Vagy az egyes pozíciókban előforduló aminosavak számát, vagy a vizsgált pozíciók számát korlátok közé kell szorítani. A módszerek alkalmazhatóságát korlátozza az is, hogy segítségükkel csak más molekulához szelektíven kötő fehérjék vizsgálhatók, melyeknek a kötőpartnere is ismert.

16.2. Könyvtárkészítési technikák

16.2.1. PCR alapú könyvtárkészítési eljárások

16.2.1.1. Hibát ejtő (*error-prone*) PCR

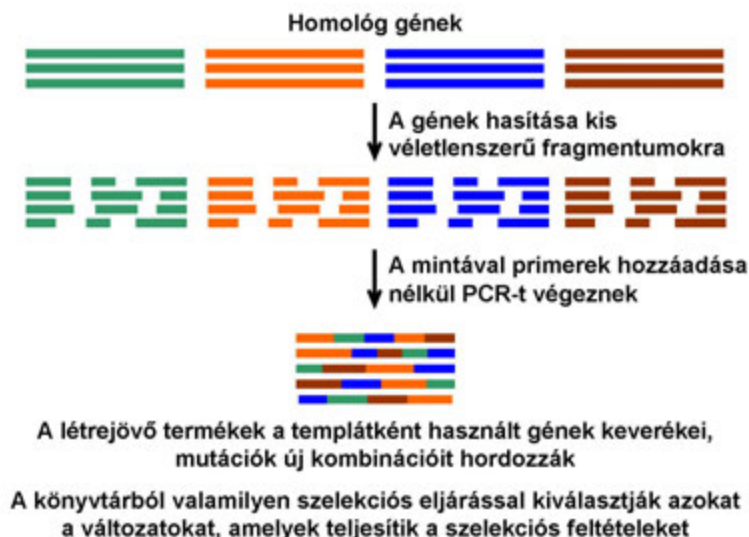
A PCR alapú könyvtárkészítés legegyszerűbb módja az úgynevezett hibát ejtő (*error prone*) PCR alkalmazása, amelyet részletesen ismertettünk a **11.1.2. fejezetben**. A hibát ejtő PCR olyan eljárás, melynek során **gyakrabban hibázó DNS-polimeráz enzim** használatával, a reakciókörülmények szuboptimális, hibára hajlamosító beállításával, vagy éppen egyenlőtlen arányban adagolt nukleotidokkal a vizsgálni kívánt fehérje génjének PCR-rel történő felszaporítása során véletlenszerű pontmutációkat tartalmazó kópiákat tartalmazó génkönyvtárat állítanak elő. Ezt követően a mutánsokat is tartalmazó könyvtárat fehérje formában kifejezik, és egy megfelelő szelekciós módszerrel kiválaszthatják a sokaságból a megváltozott funkciójú fehérjéket. A kiválasztott fehérjék ezután tovább vizsgálhatók. Megállapítható a mutáció vagy mutációk pontos helye és típusa, valamint a funkcióra gyakorolt hatása. A módszer alkalmazásában nehézséget okoz például a helyes mutációs ráta megválasztása, ugyanis a mutációk között a káros, például teljes funkcióvesztést okozó mutációk sokkal gyakoribbak, mint a hasznosak, amelyek a funkció kimutatható megváltozásával járnak. Magas mutációs ráta esetén a sok káros, például a térszerkezetet torzító mutáció elfedheti a vizsgálat szempontjából hasznos, megváltozott funkciót eredményező mutációk hatását, míg túl alacsony mutációs ráta esetén a vad típus dominanciája miatt nehezebb a mutánsokat megtalálni. A pontmutáción alapuló technika előnyösen megváltozott specifitású fehérjék előállítására kevésbé alkalmas, ugyanis a mutációk véletlenszerűen jelennek meg, és egy lényeges funkcionális ugrás eléréséhez valószínűleg több aminosav együttes megváltozása van szükség.

Bár nem PCR technika használatán alapul, de szintén random mutációkat tartalmazó könyvtár előállítására alkalmas módszer a hibára hajlamosított *E. coli* sejtek használata (**ld. 11.1.1. fejezet**). Ezek a sejtek a DNS hibajavítás bizonyos fehérjéi, vagy *proofreading* aktivitással rendelkező DNS-polimerázok mutáns változatait hordozzák, ezért ezekben a sejtekben a DNS replikáció sok hibával zajlik.

16.2.1.2. „Szexuális” PCR

A fejlettebb PCR alapú eljárások a könyvtárkészítés során a homológ rekombinációval analóg folyamatokat használnak ki. Számítógépes szimulációk igazolják, hogy a rekombináció képességével rendelkező rendszerek gyorsabban evolválódnak és „robosztusabbak”, vagyis külső behatásokra (esetünkben káros mutációkra) ellenállóbbak, mint azok, melyek rekombinációra képtelenek. A rekombináció biztosítja a különböző egyedekben létrejött mutációk keverését. A rekombináció és szelekció együttes megvalósulása lehetővé teszi a kedvező mutáció kombinációk kialakulását, valamint megakadályozza a káros mutációk elszaporodását. A fent leírt hibára hajlamos PCR, valamint a később bemutatásra kerülő helyspecifikus mutagenézisen alapuló technikák nem rendelkeznek a rekombináció képességével, ellentétben a következőkben ismertetett „szexuális PCR” módszerrel (**ld. 16.3. ábra**).

A szexuális PCR vagy DNS keverés technika a természetes vagy mesterséges változatosság homológ rekombinációval történő keverését imitáló *in vitro* módszer, ami kombinálható az így létrejövő változatok szelekciójával. A vizsgált fehérje génjének a természetben előforduló vagy például hibára hajlamos PCR segítségével előállított különböző változatait DN-áz I enzim segítségével véletlenszerűen fragmentumokra hasítják, majd az így kapott DNS töredékeket tartalmazó mintán hozzáadott primerek nélkül PCR-t végeznek. A PCR során a gén különböző változataiból származó töredékek a komplementer szekvenciájú szakaszoknál párosodhatnak egymással, így kvázi primerként szolgálva a másoláshoz. Az eredetivel megegyező hosszúságú szekvenciák épülnek fel a PCR reakció során, amelyek több kiindulási szekvenciára jellemző szakaszból állnak. Az újra felépült szekvenciák a változatok olyan kombinációit hordozzák, melyek a szülői szekvenciák között nem fordultak elő, azaz a változatok között rekombináció következik be. A rekombinációnak köszönhetően a módszer alkalmas a kezdeti variáció kombinatorikus keverésére, segítségével kombinatorikus könyvtár állítható elő. A rekombináció előnye tehát a többi mutagenézis technikával szemben, hogy több cikluson keresztül zajló molekuláris evolúcióra ad lehetőséget.

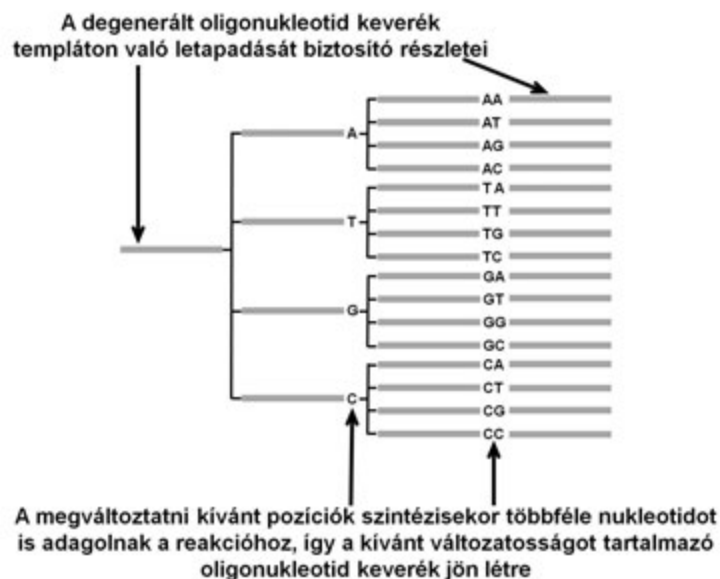


16.3. ábra: „Szexuális” PCR

A módszer másik nagy előnye a rugalmasságában rejlik, ugyanis a körülmények beállításával a véletlenszerűen jelentkező pontmutációk gyakorisága a variáció további növelése érdekében változtatható. A szülői szekvenciák hozzáadásával mintegy visszakeresztézést (*back cross*) is el lehet végezni, amivel a káros, illetve az esetleges csendes, a funkciót nem befolyásoló mutációk mennyisége alacsonyan tartható. A szexuális PCR alkalmazható például enzimek fejlesztésére is egy olyan eljárásban, ami hasonlít az ellenanyagok képződésére a szervezetben. A kutatók négy különböző b-laktamáz enzim génjét keverték össze, majd több cikluson keresztül engedtek további rekombinációt és pontmutációt. A ciklusok között a könyvtárat hordozó sejteket moxolaktám rezisztenciára szelektálták. A leghatékonyabb előállított kiméra enzim mind a 4 kiindulási fehérjétől legalább 100 aminosav pozícióban különbözött, és számos újonnan megjelent pontmutációt is tartalmazott.

16.2.2. Helyspecifikus mutagenézisen alapuló eljárások

A DNS szintetizálás módszereinek fejlődése lehetővé teszi **degenerált oligonukleotidok** szintézisét, melyek segítségével - oligonukleotiddal irányított helyspecifikus mutagenézis technikákkal - **kombinatorikus könyvtárak** állíthatók elő. A degenerált oligonukleotidok úgy készülnek, hogy a szintézis során a meghatározott pozícióknál a reakcióhoz többféle nukleotidot adagolnak. A többféle nukleotid véletlenszerűen beépül az oligonukleotidokba. Így az oligonukleotidoknak egy olyan keveréke keletkezik, melyek az összes lehetséges kombinációban tartalmazzák a tervezett mutációkat (**ld. 16.4. ábra**). A degenerált oligonukleotidokkal végzett helyspecifikus mutagenézis segítségével kombinatorikus, azaz a tervezett mutációkat az oligonukleotidokhoz hasonlóan minden lehetséges kombinációban tartalmazó fehérjekönyvtár állítható elő. Az egyes mutációk gyakorisága a degenerált oligonukleotid szintézise során a különböző nukleotidok adagolásakor, azok arányának helyes megválasztásával állítható be. A helyspecifikus mutagenézis technikákon alapuló könyvtárkészítési eljárások előnye, hogy a mutációk pozíciója és az egyes pozíciókban létrejövő változatosság pontosan tervezhető.



16.4. ábra: Degenerált oligonukleotid szintézis

16.1. táblázat: IUB nukleotid szimbólumok

Szimbólum	Jelentés (bázisok)	A szimbólum jelentése, eredete	Komplementer szimbólum
A	A	Adenin	T/U
G	G	Guanin	C
C	C	Citozin	G
T	T	Timin	A
U	U	Uracil	A
R	G vagy A	puRin	Y
Y	T/U vagy C	pYrimidin	R
M	A vagy C	aMino	K
K	G vagy T/U	Keto	M
S	G vagy C	erős (Strong) kölcsönhatás	W
W	A vagy T/U	gyenge (Weak) kölcsönhatás	S
B	G vagy C vagy T/U	nem A	V
D	A vagy G vagy T/U	nem C	H
H	A vagy C vagy T/U	nem G	D
V	A vagy G vagy C	nem T/U	B
N	A vagy G vagy C vagy T/U	bármelyik (aNy)	N

16.2.2.1. Telítési mutagenézis

Telítési mutagenézis során a vizsgált pozíciókban **mind a 20 aminosav előfordulása megengedett**. Ilyen könyvtárak esetében a randomizálandó pozíciókban a degenerált oligonukleotidok általában NNK vagy NNS kodonokat tartalmaznak, ahol N a négyféle bázist jelenti, K = G/T és S = G/C (**ld. 16.4. ábra és 16.2. táblázat**). Az NNK és NNS kodonok 32 kodon keverékét eredményezik. A 32 lehetséges kodon kódolja mind a 20 aminosavat, és mindössze egyféle stop kodon megjelenését teszi lehetővé. Ezzel elérhető, hogy a könyvtárban az egyes aminosavak kiindulási aránya egyenletesebb legyen, mint NNN kodonok alkalmazása esetén, valamint a stop kodon előfordulási gyakorisága is kisebb.

Telítési mutagenézissel egyszerre kevés pozíció vizsgálható, mert a könyvtár elméleti mérete, azaz a lehetséges változatok száma 6-7 pozíció vizsgálata esetén eléri a leggyakrabban alkalmazott szelekciós technikák által vizsgálható maximális méretet. Az ezzel a módszerrel készített könyvtárak legfontosabb alkalmazási területei általában lokalizált régiók feltárását célozzák meg. Ilyen feladatok pl. lineáris peptid motívumot felismerő fehérjék preferenciáinak feltárására random peptid könyvtárak alkalmazásával, vagy az ellenanyag variábilis hurok régiók vizsgálata. A telítési mutagenézis a legjobb választás akkor is, ha rövid összefüggő szakasz irányított evolúciós vizsgálatát célozzuk meg.

16.2.2.2. Mutációkkal tűzdelt (*spiked*) oligonukleotidok használata

Az egy könyvtárban vizsgált pozíciók számának kiterjesztésére alkalmas módszer a mutációkkal tűzdelt (*spiked*) oligonukleotidok használata. Ilyen könyvtárak esetén a degenerált oligonukleotid sereg tagjai alapvetően az összes pozícióban a vad típusnak megfelelő nukleotidot tartalmazzák. A szintézis során célzottan, a megváltoztatni kívánt pozíciókban adnak a vad típusú mellett kis mennyiségben más nukleotidot vagy nukleotidokat is. A mutagenézis során létrejövő könyvtár a legtöbb pozícióban vad típusú géneket fog tartalmazni, amelyek a vad típusútól eltérő nukleotidok adagolásának függvényében fognak mutációkat hordozni. Ezt a módszert használják például, ha egy kölcsönhatás paraméterein szeretnének javítani.

16.2.2.3. Megszabott kodonokat hordozó (*tailored*) oligonukleotidok használata

Ez a stratégia úgy terjeszti ki a vizsgálatba bevonható pozíciók számát, hogy az egyes pozíciókban csak a lehetséges 20 aminosav egy részének megjelenését engedi meg (ld. 16.2. táblázat).

16.2. táblázat: Gyakran használt aminosav készletek kódolása degenerált kodonokkal az IUB nevezéktan szerint

Degenerált kodon	Jellemzők	Kódolt aminosavak	Stop kodonok	Kodonok száma
NNN	mind a 20 aminosav	mind a 20 aminosav	TAA, TAG, TGA	64
NNK	mind a 20 aminosav	mind a 20 aminosav	TAG	32
NNS	mind a 20 aminosav	mind a 20 aminosav	TAG	32
NNC	15 aminosav	A,C,D,F,G,H,I,L,N,P,R,S,T,V,Y	nincs	16
NWW	töltött vagy hidrofób aminosavak	D,E,F,H,I,K,L,N,Q,V,Y	TAA	16
RVK	töltött vagy hidrofil aminosavak	A,D,E,G,H,K,N,R,S,T	nincs	12
DVT	hidrofil	A,C,D,G,N,S,T,Y	nincs	9
NVT	töltött vagy hidrofil aminosavak	C,D,G,H,N,P,R,S,T,Y	nincs	12
NNT	kevert	A,D,G,H,I,L,N,P,R,S,T,V	nincs	16
VVC	hidrofil aminosavak	A,D,G,H,N,P,R,S,T	nincs	9
NTT	hidrofób aminosavak	F,I,L,V	nincs	4
RST	kis oldalláncú aminosavak	A,G,S,T	nincs	4
TDK	hidrofób aminosavak	C,F,L,W,Y	TAG	6

Ahhoz, hogy ezt a stratégiát eredményesen lehessen alkalmazni, szükségesek előzetes ismeretek a vizsgált interakcióról. Ezek alapján lehet kiválasztani, hogy az egyes pozíciókban milyen aminosavak fordulhassanak elő. A tervezésben segíthet, ha több homológ fehérje szekvenciája ismert, ha már vannak előzetes irányított evolúciós eredmények, azaz lehet tudni, hogy milyen karakterű aminosavak képesek ellátni az interakció által megkövetelt funkciót, vagy rendelkezésre áll a vizsgált molekula háromdimenziós szerkezete.

16.3. Szelekciós eljárások

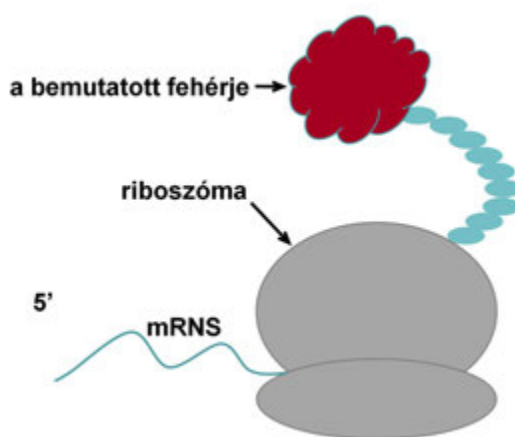
16.3.1. *In vitro* szelekciós módszerek

In vitro könyvtárkészítési és szelekciós módszerek alkalmazásával van lehetőség a legnagyobb, akár 10^{13} különböző változatot tartalmazó könyvtárak előállítására. Ennek az oka az, hogy a folyamat során a DNS könyvtárat *in vitro* translációs technikákkal fejezik ki fehérje formában. Ennek köszönhetően nem kell alacsony hatékonyságú módszerekkel a DNS-t sejtekbe juttatni, így a diverzitást csak az *in vitro* transláció során a rendszerben jelen lévő riboszómák és információhordozó molekulák mennyisége korlátozza.

16.3.1.1. Riboszóma-bemutató

A riboszóma bemutatással szelektálható könyvtárat alkotó gének egy T7 transzkripciós promóterből, egy riboszóma-kötőhelyből, a könyvtárfehérje génjéből és egy polipeptid linkert kódoló szakaszból állnak (**ld. 16.5. ábra**). A gén nem tartalmaz stop kodont. A transzkripciót *in vitro* végzik T7 RNS-polimeráz segítségével, riboszómák és aminoacil-tRNS-ek, GTP valamint a szükséges iniciációs és elongációs faktorok jelenlétében. A könyvtárfehérje és az azt követő linker átíródik, a riboszóma az mRNS végére ér, de a stop kodon hiánya miatt a disszociáció nem következik be. Az átíródott peptid linkernek köszönhetően a könyvtárfehérje el tud távolodni annyira a riboszómától, hogy a feltekeredése zavartalanul végbemehessen és a szelekció során a kötőpartner hozzáférjen.

A reakció során keletkező, a riboszómából, az mRNS-ből és a könyvtárfehérjéből álló hármas komplex 4°C -on stabil. Az így létrehozott fehérje könyvtárból a vizsgált fehérje kötőpartnerével kiválaszthatóak azok a klónok, amelyek a mutációk ellenére megőrizték a partner kötésének képességét, vagy a mutációk következtében a vadtypusnál erősebben kötik azt. A szelekciós követelményeket teljesítő fehérjéket kódoló mRNS-eket reverz transzkripció után fel lehet szaporítani, ezáltal lehetőség nyílik egyrészt ismételt szelekcióra, másrészt a gének szekvenálására is.



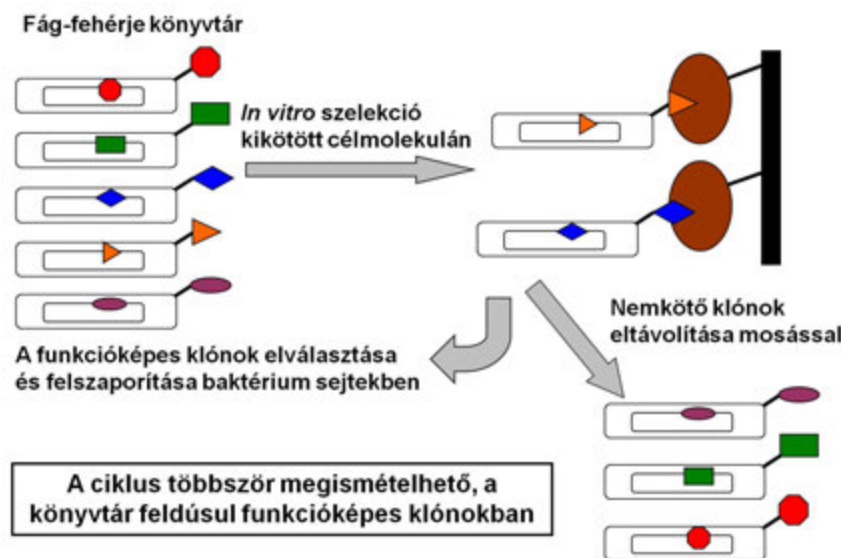
16.5. ábra: Riboszóma bemutatás

16.3.1.2. mRNS-bemutató

Az mRNS-bemutató a riboszóma-bemutató továbbfejlesztett változata. A fehérjét kódoló mRNS 3' végéhez puromicin molekulát kapcsolnak kovalensen, egy rövid DNS linkeren keresztül. Az *in vitro* transláció során a riboszóma megáll az RNS-DNS határon. Ekkor a puromicin belép a riboszóma peptidil-transzferáz helyére, és stabil amid kötést alakít ki a keletkező fehérje C-terminálisával. A keletkező komplex stabil, a kifejezett könyvtár a riboszóma-bemutatóhoz hasonlóan szelektálható.

16.3.2. A fág-bemutató

A fág-bemutató az irányított evolúciós módszerek legkorábban kifejlesztett képviselője. Egyszerűségének, rugalmasságának és hatékonyságának köszönhetően mindmáig a fág-bemutató a **leggyakrabban alkalmazott irányított evolúciós eljárás** (ld. 16.6. ábra).



16.6. ábra: A fág bemutatás alapvonásai

A fág-bemutató során az **evolálni kívánt fehérje génjét** egy **bakteriofág** egyik **burokfehérjéje génjéhez** illesztik. A gén terméke fúziós fehérje, ami a burokfehérjéből és a könyvtárfehérjéből áll, melyeket általában flexibilis peptid linker köt össze. A fúziós fehérje beépül a bakteriofág burkába, így a vizsgált fehérje megjelenik a fág felszínén. A keletkező fágok a kötőpartner számára hozzáférhető formában fejezik ki a vizsgált fehérje egy variánsát, és hordozzák annak génjét, azaz megvalósul a funkcióért felelős fenotípus és az azt kódoló genotípus fizikai kapcsolata. A könyvtárát általában helyspecifikus mutagenézisen alapuló eljárással hozzák létre, de van lehetőség pl. szexuális PCR alkalmazására is. A bakteriofág felszínén kifejezett könyvtárból kiválaszthatók a szelektív kötési képességével rendelkező változatok például szilárd felszínhez kötött specifikus kötőpartner segítségével. A kiválasztás folyamata tulajdonképpen affinitás kromatográfiás eljárással történik.

Első bemutatása óta a fág-bemutató a nagyszámú egyedi változatot tartalmazó kombinatorikus peptid ill. fehérje könyvtárak szelekciójának leghatékonyabb módszerévé vált. Leggyakrabban a biológiailag jól feltérképezett, nem-litikus fonális M13 fág felszínén fejezik ki a könyvtárat, de litikus fágokon (l, T4, T7) megjelenített könyvtárak is használatosak.

T7-bemutató esetén lírikus, kétszálú DNS-t hordozó, *E. coli*-t fertőző T7 vírus felszínén jelenítik meg a fúziós fehérjéket. A DNS az ikozaéder formájú fejbe pakolódik. A fejet a 10-es gén két, eltérő méretű terméke alkotja hatszögű és ötszögű lapokba rendeződve. A kis peptidet tartalmazó könyvtárakat a nagyobb burokfehérje C-terminálisához kapcsolva fejezik ki nagy kópiaszámban. Nagyobb fehérjék kifejezésére csak kisebb kópiaszámban van lehetőség. A lírikus fág használatának egyik hátránya az, hogy a fágok elszaporodása során egy sejtet több fág is megfertőzhet, így létrejöhet olyan vírus is, ami több könyvtárat is kifejez, de közülük csak az egyik génjét hordozza.

Az **M13 fág** használatán alapuló könyvtárak elterjedtebbek az egyszerűbb kezelhetőség, valamint a genotípus és a fenotípus közti fizikai kapcsolat könnyű biztosíthatósága miatt. A fonális fágok kb. 10 gént tartalmazó vírusok, amelyek Gram-negatív baktériumokat fertőznek. A virion egy 6,5 nm átmérőjű, 800-2000 nm hosszú flexibilis pálcá, melynek hossza a virionba pakolt DNS mennyiségétől, méretétől függ. A virion 5-féle fehérjéből áll. A fő burokfehérje (protein 8, P8) több ezer példánya alkotja a fág burkát. A virion lezárásáért a sejtől elsőként távozó végén két burokfehérje (P7, P9) öt-öt példánya, míg a másik végén két másik burokfehérje (P3, P6) szintén öt-öt példánya felel. A burok belsejében található a fág genomja, egy, a törzstől függően 5-8000 bázis méretű, cirkuláris, egyszálú DNS (ssDNS) molekula. Az M13 bakteriofág F-faktort hordozó, azaz konjugatív F-pilussal rendelkező

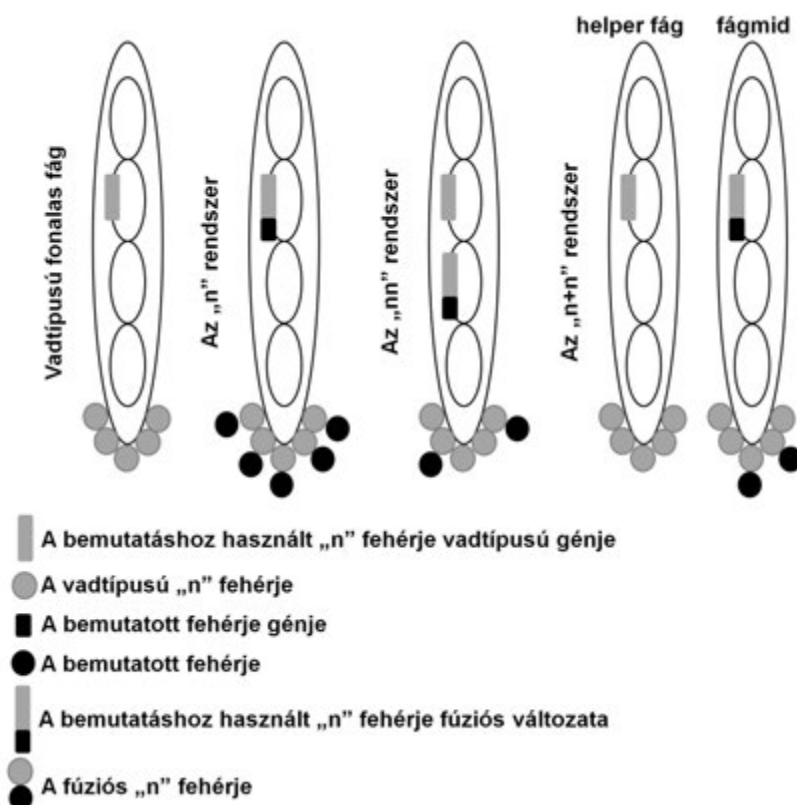
E. coli sejteket fertőz. A fertőzés során a vírus ssDNS-e bejut a baktérium sejtbe és a sejt különböző fehérjéi és enzimeit által kétszálú, replikatív dsDNS-sé alakul. A kétszálú formáról átíródnak a vírus fehérjéi. Ha a virális fehérjék koncentrációja meghalad egy küszöböt, akkor kezdetét veszi a vírus genom ún. gördülő kör (*rolling-circle*) mechanizmus szerint történő replikációja, amelynek során a genom virionba pakolódo szála sokszorozódik meg. A virion összeszereléséhez szükséges fehérjék a baktérium sejt belső membránjába kerülnek, és integráns membránfehérjeként játszanak szerepet az összeszerelésben.

Az M13 bakteriofág felszínén megjelenített könyvtárak általában a **P3**, a **P6** vagy a **P8 fehérjéhez kapcsoltn** fejezik ki a könyvtárfehérjét. A fág-bemutatós során alkalmazható vektorok típusokba oszthatók, az alapján, hogy a fúzióban érintett burokfehérje génjét milyen formában és hány kópiában hordozzák.

Az „n” (itt az n az érintett burokfehérje gén száma) típusú fág-genom vektorok a fág összes génjének, így a könyvtárfehérjét bemutató fehérje génjének (n) is egyetlen kópiáját hordozzák. A könyvtárfehérjét kódoló szekvenciát az n génhez illesztik. Az ilyen vektorról termelődő fágok virionja az n gén termékének kizárólag fúziós példányát tartalmazza. A „3”-típusú vektorral létrehozott fág csak fúziós P3-at kódol és virionja elvileg csak fúziós P3-at tartalmaz, de bakteriális enzimek a könyvtárfehérjét lehasíthatják. A „8”-típusú vektor a fő burokfehérje minden példányán hordozza a könyvtárfehérjét, ami ebben az esetben nem fehérje, hanem legfeljebb 8-10 aminosavból álló peptid, ugyanis hosszabb peptid kifejezése a virion összeszerelését akadályozná. A könyvtárfehérjék több, „8”-as típus esetében több ezer példányban fejeződnék ki. Ez jelentős aviditást okoz: az egy fágon bemutatott fehérje több példány egyszerre kapcsolódhat egy-egy kötőpartner molekulához, így a fág látszólagos affinitása a kötőpartnerhez az önálló könyvtár affinitásának sokszorosa is lehet. Az „n” típusú könyvtárak esetében nem szabályozható a bemutatás valenciája, azaz az, hogy a fág hány példányban mutatja be a fúziós peptidet (**ld. 16.7. ábra**).

Az „nn”-típusú fág-genom vektorok az egyik burokfehérje génjének két példányát tartalmazzák, egy vad típusút és egy fúziósat. Az ilyen vektorról termelődő fágok mozaikosak, az egyik burokfehérjéből fúziósat és vad típusút is tartalmaznak. A fúziós gén elé szabályozható promótert helyezve lehetőség van a fúziós/vadtípus arányt, azaz a bemutatás valenciáját szabályozni. Az ilyen vektorok alkalmasak nagyobb peptidek vagy akár egész fehérjék kifejezésére is. Végül az „n+n”-típusú vektorok fágmid konstrukciók használatán alapulnak.

A fág-bemutatós fontos eleme az, hogy a kötőpartnerhez kapcsolódni képes variánsok képesek baktériumokban szaporodni. Ennek köszönhetően rövid idő alatt több szelekciós kör alkalmazható, melyek végére a szelektált könyvtár feldúsul a funkcionálisan hatékony klónokban. A fág-bemutatós nagy sikerrel alkalmazható kötőfelszín energetikai viszonyainak felderítését és biológiai folyamatokban fontos szerepet játszó kötőfelszín kismolekulájú, specifikus ligandumai fejlesztését célzó kutatásokban.



16.7. ábra: A fág-bemutatáshoz használható vektor rendszerek

16.3.3. Rekombináns fehérje könyvtár bemutatása és szelekciója élő sejtek felszínén

A fág-bemutatás egyik gyakran alkalmazott alternatívája a sejt felszínén való bemutatás. Ezekkel a technológiákkal a rekombináns fehérje könyvtár egyes tagjait élő sejtek felszínén jelenítik meg. A könyvtárat bemutató sejtek közül kiválasztják a partnerhez kötődő variánsokat bemutató sejteket, ezeket a többitől elválasztják, majd az általuk hordozott géneket elemzik. Bizonyos fehérjéket nem lehet más irányított evolúciós eljárásokkal vizsgálni vagy fejleszteni, ezért különösen fontos, hogy ma már lehetőség van sejt felszínén történő bemutatást alkalmazni. Az évek során a kutatók kifejlesztettek bakteriális, és élesztő alapú rendszereket is.

16.3.3.1. Fehérjék irányított evolúciója baktérium sejtek felszínén

A baktérium sejtek felszínén történő irányított evolúciót lehetővé tevő rendszerek a legtöbb ponton hasonlítanak a fág-bemutatás rendszerekre. A sejtek egyetlen könyvtárag génjét tartalmazzák, ezáltal csak ez a fehérje jelenik meg a felszínükön, azaz megvalósul a genotípus és a fenotípus klonális fizikai kapcsolata. Fontos eltérés viszont, hogy a rendszerben lévő részecskék, azaz maguk a sejtek sokkal nagyobbak a fágoknál, valamint a bemutatott fehérje minden esetben sok példányban megjelenik a sejt felszínén. Eltérés az eddig bemutatott rendszerektől az is, hogy **a rekombináns fehérje könyvtárat megjelenítő sejtek szelekcióját általában nem affinitás kromatográfiás módszerrel, hanem fluoreszcenciával aktivált sejtszortírozó berendezés (FACS) segítségével végzik el.**

A könyvtárból való szelekciónak legegyszerűbb módja a fluoreszcensen jelölt partnerrel való inkubálás oldat fázisban, majd a jelölt sejtek elválasztása a jelöletlenektől FACS rendszerben. A könyvtáragokat bemutató sejteket a könyvtárban vizsgált fehérje fluoreszcensen jelölt kötőpartnerével inkubálják oldat fázisban. A partnert kötni képes változatokat bemutató sejtek így fluoreszcensen jelöltté válnak. A multivalens bemutatásnak köszönhetően a funkcióképes változatokat bemutató sejtek elegendő fluoreszcensen jelölt partner molekulát kötnek meg ahhoz, hogy a FACS berendezés segítségével el lehessen választani a jelölt és a jelöletlen sejteket egymástól. A FACS készüléken be lehet állítani, hogy milyen intenzitású fluoreszcens jelet adó sejteket gyűjtsön külön a többitől. A

sejtek által megkötött jelölt fehérjék mennyisége, azaz a jel intenzitása arányos a bemutatott változat affinitásával, azaz az ún. küszöbérték helyes megválasztásával kiválaszthatóak a könyvtárból a legnagyobb affinitású változatok.

Ezzel a módszerrel lehetőség van bizonyos **antigénekre specifikus ellenanyagok fejlesztésére** is. Erre a fág-bemutatás is alkalmas, viszont a sejt felszínén való bemutatás és FACS rendszerrel történő szelekciónak új lehetőségeket is biztosít. Az ellenanyagok fontos tulajdonsága a kereszt-reaktivitás, azaz az, hogy a célponthoz hasonló antigéneket felismer-e az ellenanyag. A sejt felszínén történő bemutatással megoldható akár kereszt-reaktív, akár nagyon specifikus ellenanyagok kifejlesztése. Ilyen esetekben a szelekció során többféle antigént alkalmaznak, amelyeket különböző jelöléssel látnak el. Ha kereszt-reaktív ellenanyag kifejlesztése a cél, akkor a FACS készüléket úgy programozzák, hogy azokat a sejteket gyűjtse külön, amelyek többféle fluoreszcens jelölést is hordoznak. Ha specifikus ellenanyagot keresnek a könyvtárban, akkor a sejtek válogatása során kizárják a gyűjtött mintából a többféle fluorofórral jelölt sejteket.

Az elmúlt években sikeres irányított evolúciós vizsgálatokat végeztek különböző ellenanyag fragmentumok, teljes ellenanyagok, rövid peptidek és más globuláris fehérjék baktérium sejtek felszínén bemutatott és szelektált könyvtárakkal.

16.3.3.2. Fehérjék irányított evolúciója élesztő sejtek felszínén

A rekombináns fehérje könyvtárak élesztő sejtek felszínén való megjelenítése és szelekciója dinamikusan fejlődő terület. Az élesztő felszíni bemutatás egyedülálló lehetőséget kínál számos olyan fehérje vizsgálatára, amelyeket más irányított evolúciós rendszerben nem lehet kifejezni. Az élesztő felszíni bemutatás rendszerekben **olyan fúziós gént** hoznak létre, **amelyről a könyvtár fehérje az élesztő sejtfalába ágyazott fehérjék egyikével összekapcsoltnak termelődik**, és ezáltal megjelenik az élesztő sejt felszínén. Ezekben a rendszerekben a könyvtárfehérjét glikozil-foszfatidil-inozitol (GPI) horgonnyal ellátott fehérjékkel, vagy azok megfelelő részletével fúzióban állítják elő. Ezek a fehérjék a sejten kívül a sejtfa komponenseivel vagy más membránfehérjékkel kovalens kötással is kapcsolódhatnak.

Lehetőség van a könyvtárfehérje nem-kovalens bemutatására is. Ilyen rendszerekben a fúziós partner olyan élesztő fehérje, ami a sejtől kijutva a sejtfa mannán láncához kötődik másodlagos kölcsönhatásokon keresztül. Eukarióta transzmembrán fehérjék horgonyként szolgáló fúziós partner alkalmazása nélkül is kifejezhetőek élesztő felszínén.

Az élesztő felszínén megjelenített könyvtárak **szelekciója** a baktérium sejt felszínén kifejezett könyvtárakhoz hasonlóan történhet szilárd felszínhez rögzített **partner molekulákhoz való kötődés alapján**, vagy fluoreszcensen jelölt partner és **FACS alkalmazásával**.

Az élesztő felszíni bemutatást leggyakrabban más molekulákhoz kötni képes fehérjék kötőfelszínének feltérképezésére, vagy a vad típusú molekulánál nagyobb affinitású, esetleg megváltozott specifitású változatok kifejlesztésére használják. Ezzel a módszerrel számos specifikus és nagy affinitású ellenanyagot sikerül kifejleszteni különböző antigénekkal szemben.

16.4. További olvasnivaló a fejezethez

Sidhu S. S. (2005) (szerk) Phage display in biotechnology and drug discovery. Taylor&Francis

Smith G. P. és Petrenko V. A. (1997) Phage Display. Chemical Reviews 97, 391–410.

Minshull J. és Stemmer W. P. (1999) Protein Evolution by Molecular Breeding. Current Opinion in Chemical Biology 3, 284–290.

Diaz J. E. és mtsai. (2003) Exploring Biochemistry and Cellular Biology with Protein Libraries. Current Issues in Molecular Biology 5, 129–145.

Pepper, L. R. és mtsai. (2008) A Decade of Yeast Surface Display Technology: Where Are We Now? Combinatorial Chemistry & High Throughput Screening 11, 127–134.

Löfblom, J. Bacterial Display in Combinatorial Protein Engineering. Biotechnology Journal 6, 1115–1129.

17. fejezet - Kölcsönhatások kimutatása géntechnológiai módszerekkel

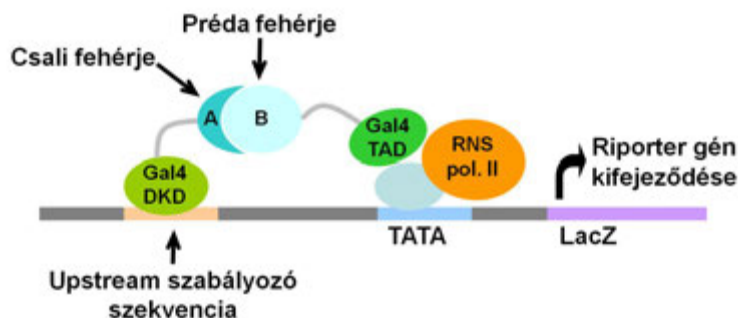
A fehérjék kölcsönhatásainak kimutatása az 1990-es évek elejéig gyakran nehezen leküzdhető akadályokat gördített a kutatók útjába. A géntechnológia eszköztárának fejlődése és a biológiai folyamatok molekuláris szinten való egyre részletesebb megismerése együtt tette lehetővé a kölcsönhatások kimutatására szolgáló technikák kifejlesztését az 1980-as évek második felétől. A fehérjék más fehérjékkel, DNS-sel vagy kis molekulákkal az élő sejten kialakított kölcsönhatásainak kimutatására alkalmas *in vivo* módszerekre nagy igény mutatkozott az alap- és az alkalmazott kutatásban egyaránt. Az **élesztő két-hibrid** (*yeast two-hybrid*) eljárás volt az egyik első, erre alkalmas eljárás.

17.1. Az élesztő két-hibrid rendszer bemutatása

Az **élesztő két-hibrid** rendszer és annak különböző változatai a fehérjék más molekulákkal kialakított interakcióinak kimutatására szolgáló, teljes egészében ***in vivo* rendszerek**. A módszercsalád különböző változatainak alkalmazásával lehetőség van fehérje-fehérje, fehérje-DNS, fehérje-kismolekula kölcsönhatások kimutatására. A módszerek alkalmasak az irányított evolúciós technikáknál alkalmazott könyvtárakból történő szelekcióra is. Ezekben az esetekben egy fehérje rengeteg változata közül választják ki azokat, amelyek képesek kötni a vizsgálatban alkalmazott partnerhez. Az élesztő két-hibrid módszert azonban gyakrabban alkalmazzák arra is, hogy nagyszámú különböző fehérjét tartalmazó könyvtárból kimutassák azokat a fehérjéket, amelyek kötnek az alkalmazott partnerhez vagy partnerekhez. A fentiek értelmében az élesztő két-hibrid rendszer alkalmazása során nem egyetlen interakció részletes feltérképezését végzik el, hanem a sejten belül létrejövő, fiziológiás vagy patológias folyamatokban releváns kölcsönhatásokat keresnek.

Az elsőként leírt élesztő két-hibrid technika elnevezése arra utal, hogy ebben a rendszerben **kétféle** hibrid, azaz **fúziós fehérjét** alkalmaznak a kölcsönhatás kimutatására. Az eljárás azon alapul, hogy a génátírást serkentő **transzkripció faktorok** moduláris felépítésű, több doménből álló fehérjék. Bizonyos domének felelősek a megfelelő DNS szakaszok (*upstream* szabályozó elemek) specifikus megkötéséért (**DKD:DNS-kötő domén**). Más domének a transzkripció iniciációjáért felelős fehérjékkel alakítanak ki kölcsönhatást és a kölcsönhatásokon keresztül serkentik a gén transzkripcióját (**transzkripciót aktiváló domén, TAD**). A génátírás serkentéséhez a két doménnek egy időben kell jelen lennie a gén promóterénél, de nem feltétlenül kell egyetlen polipeptidláncban belül, kovalensen összekötve lenniük, a kölcsönhatásuk lehet közvetett. A két önálló fehérjeként előállított DKD és TAD domének nem ismerik fel egymást, nem lépnek egymással kölcsönhatásba. Azonban ha az egyiket egy „A” fehérjéhez, a másikat pedig egy „B” fehérjéhez fuzionáljuk, akkor amennyiben az „A” és „B” fehérjék egymást felismerik, úgy a kötődésük miatt megvalósul a DKD és TAD domének térbeni közelsége. A DKD révén a komplex a megfelelő helyen lokalizálódik és a TAD képes serkenteni azoknak a géneknek a kifejeződését, amelyek szabályozó régiója tartalmazza az alkalmazott DKD által felismert szekvenciát. Az élesztő két-hibrid rendszerekben az „A” és „B” fehérjék kölcsönhatása egy riporter gén átíródását eredményezi. A **riporter gén** kifejeződése alapján választják ki azokat a klónokat, amelyekben kialakult az „A” és „B” fehérje közti kapcsolat.

Az eredeti élesztő két-hibrid rendszerben a *S. cerevisiae* **Gal4 transzkripció faktor DKD és TAD doménjét** használták. A Gal4 transzkripció faktor szükséges a **LacZ riporter gén** aktiválásához. A LacZ gén terméke, a β -galaktozidáz enzim. Azok a sejtek kék színűek lesznek X-Gal kromogén szubsztrát analógot tartalmazó közegben, amelyekben a riporter gén kifejeződik. A vizsgálatok során olyan élesztő törzseket alkalmaznak, amelyekben nem termelődik a riporter gén terméke. Ennek köszönhetően a riporter fehérje csak akkor jelenhet meg, ha az „A” és „B” fehérje kölcsönhatásba lép egymással (**ld. 17.1. ábra**)



17.1. ábra: Az élesztő két-hibrid rendszer sémája

Az élesztő két-hibrid rendszerben a **DKD doménhez** fúzionálják az úgynevezett **csali fehérjét (bait)**. Egy vizsgálatban egyféle csali fehérjét alkalmaznak, a vizsgálat célja a csali fehérje kötőpartnereinek azonosítása. A **TAD doménhez** fúzionálják az úgynevezett **préda fehérjét (prey)** vagy **fehérjéket**. A préda lehet egyetlen fehérje könyvtára, de gyakrabban alkalmaznak sok ismert vagy ismeretlen fehérjét tartalmazó cDNS könyvtárakat. A kétféle plazmiddal ezután élesztő sejteket transzformálnak, majd kiválasztják a rendszerből azokat a sejteket, amelyekben kifejeződött a riporter gén a csali és a préda fehérje kölcsönhatásának következtében. A sejtek **szelekciójára** többféle lehetőség van. A fent bemutatott **LacZ riporter** gén használata során azok a sejtek **kék színűek** lesznek **X-Gal jelenlétében**, amelyekben a riporter gén kifejeződik. Ebben az esetben azok a sejtek is életképesek, amelyekben nincs kölcsönhatás a csali és a préda fehérje között. A szelekciós eljárások másik típusa során a **13.1.2. fejezetben** bemutatott **auxotróf törzset** használnak, a riporter gén pedig az, amelynek mutációja a törzset auxotróffá teszi. **HIS3** szelekciós marker esetén az egyik plazmid tartalmazza a hisztidin bioszintézishez szükséges, az alkalmazott törzsben hibás enzim génjének jó kópiáját. A gén olyan promóterrel van ellátva, amiről csak akkor íródhat át a fehérje, ha a csali-préda interakció biztosítja a TAD jelenlétét a promóteren szerveződő fehérje komplexben. A csali plazmidot és a préda-könyvtár egy tagját tartalmazó sejteket ezután **hisztidin mentes tápközegben** növesztik. Csak azok a sejtek élnek túl és képesek növekedni, amelyekben a csali és a préda kölcsönhatásba lép egymással és átíródik a hisztidin bioszintézishez nélkülözhetetlen fehérje.

17.2. Az élesztő két-hibrid rendszer gyengeségei

Az élesztő két-hibrid rendszer vitathatatlan előnye, hogy élő sejteken belül, *in vivo* lehet a segítségével fehérjék kölcsönhatásait kimutatni. Sajnos azonban a rendszernek számos olyan korlátja is van, amelyek ismeretében az élesztő két-hibrid vizsgálatok során szerzett információkat, az azonosított kölcsönható partnereket fenntartásokkal kell kezelni, illetve további tesztekben igazolni kell azok valóságát.

Az élesztő két-hibrid vizsgálatok nagy arányban eredményeznek **hamis pozitív találatokat**, azaz az azonosított kötőpartner az élő rendszerben nem partnere a csali fehérjének. Elmondható az is, hogy az élesztő két-hibrid kísérletek nem mutatnak ki számos olyan kölcsönhatást, amelyek egyébként megvalósulnak és relevánsak, azaz a **hamis negatív eredmények** aránya is jelentős. Hamis pozitív találatokat eredményez a fúziós fehérjék intenzív expressziója. Ennek köszönhetően azok nagy koncentrációban vannak jelen, ezért gyenge, a fehérjék normális koncentrációja mellett ki sem alakul kölcsönhatások is létrejöhetnek. Amikor az élesztő sejtben egyszerre expresszálódnak, akkor olyan fehérjék között is kialakulhat kölcsönhatás, amelyek a valóságban soha nem fordulnak elő egy időben ugyanannak a sejtnek ugyanabban a kompartmentumban. Hamis negatív találatokat, azaz releváns interakciók kimutatásának elmaradását okozhatja, ha a fehérjék interakciós felszíne a fúziós fehérjében nem hozzáférhető, mert azt a fúziós partner eltakarja. Más élőlényből származó fehérjék helyes feltekeredése vagy a működéshez szükséges poszttranszlációs módosítása nem biztos, hogy megfelelően végbemegy az élesztő sejtben, ami szintén azt eredményezi, hogy nem alakul ki a kölcsönhatás a fehérjék között. Az élesztő két-hibrid rendszerben a kölcsönhatás helye a sejt mag. Elképzelhető, hogy két, más kompartmentumban kölcsönható fehérje a sejt magban uralkodó körülmények között nem ismeri fel egymást.

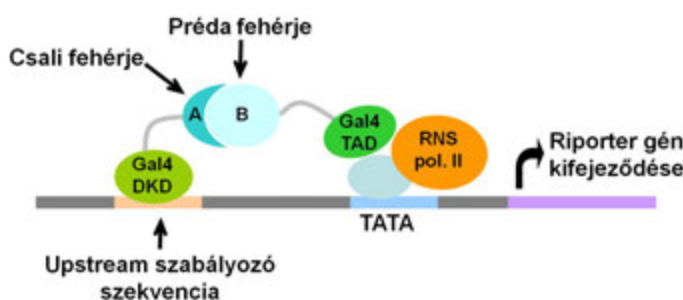
A fentiek miatt az élesztő két-hibrid kísérletek során azonosított **interakciós partnereket** minden esetben **igazolni kell** más, megbízhatóbb módszerrel is. A módszer korlátait ellensúlyozza az egyszerűsége és gyorsasága. A mai napig széles körben használják a vizsgált fehérjék potenciális partnereinek gyors azonosítására.

17.3. Hasonló alapokon működő technikák

Az élesztő két-hibrid rendszerhez hasonló technikák sorát fejlesztették ki az elmúlt két évtizedben. Ezek egyrészt olyan kölcsönhatások vizsgálatára alkalmasak, amelyekre az eredeti rendszer alkalmazásával nincs lehetőség, másrészt az eredeti rendszer hiányosságainak kiküszöbölésére kínálnak megoldást. A következőkben bemutatott technikák az irányított evolúciós rendszerekhez hasonlóan egy adott kölcsönhatás feltérképezésére, vagy új kölcsönható molekulák kifejlesztésére alkalmazhatóak.

17.3.1. Fordított két-hibrid rendszerek

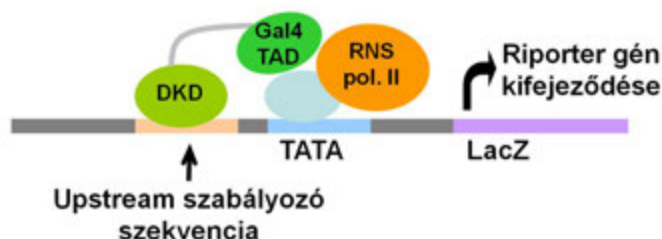
A **fordított (reverz) két-hibrid rendszerek** alkalmasak egy adott kölcsönhatás kialakulása szempontjából kritikus aminosavak feltérképezésére. Ezek a rendszerek olyan riporter géneket alkalmaznak, amelyek terméke a sejt számára toxikus. Ennek következtében azok a sejtek maradnak életben, amelyekben a csali és a préda interakciója a csali fehérje mutációja következtében nem jön létre. A túlélő sejtek által kifejezett csali fehérjék génjeinek vizsgálatával megállapítható, hogy milyen mutáció vagy mutációk voltak felelősek a két fehérje közti kölcsönhatás kialakulásának megakadályozásáért (ld. 17.2. ábra).



17.2. ábra: Reverz élesztő két-hibrid rendszer

17.3.2. Egy-hibrid rendszerek

Az egy-hibrid rendszerekben **egyetlen fúziós fehérjét** alkalmaznak. A leggyakrabban adott DNS szekvenciát felismerő DKD kifejlesztésére, vagy adott DKD által felismert DNS szekvencia megkeresésére használják. Előbbi esetben a riporter fehérje génjének szabályozó régiójába építik az adott DNS szekvenciát. A kísérlet során alkalmazott fúziós fehérje egy DKD-hez fuzionált TAD konstrukció. A DKD DNS kötésért felelős aminosav pozícióit randomizálják, így DKD könyvtárakat hoznak létre. A szelekció során azokban a sejtekben fejeződik ki a riporter fehérje, amelyekben a DKD variáns felismeri a riporter gén szabályozó régiójába épített DNS szekvenciát. Az utóbbi esetben a riporter gén szabályozó régiójába nem egy konkrét DNS szekvenciát építenek be, hanem DNS könyvtárakat hoznak létre. A könyvtárból aztán egy DKD-TAD fúziós fehérjével kiválasztják azokat a változatokat, amelyekhez a DKD kötni képes (ld. 17.3. ábra).

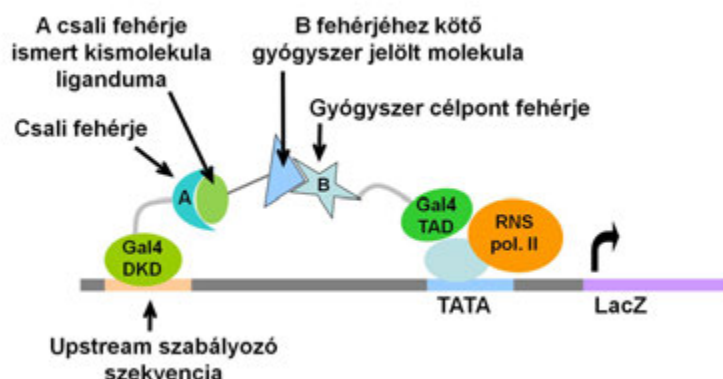


17.3. ábra: **Az élesztő egy-hibrid rendszer.** Az élesztő egy-hibrid rendszer alkalmazható egy adott DKD által felismert DNS szekvencia megkeresésére vagy adott DNS szekvenciát felismerő DKD kifejeztetésére.

17.3.3. Három-hibrid rendszerek

A három-hibrid rendszerek **három fúziós fehérje vagy hibrid molekula** alkalmazásával mutatják ki a keresett interakciót. Különböző változatai alkalmasak **fehérje-RNS** és **fehérje-kismolekulakölcsönhatások**, valamint **enzimatis aktivitás kimutatására** is.

A fehérje-kismolekula interakció kimutatására szolgáló vizsgálatok során általában kismolekula könyvtárakból válogatják ki azokat a gyógyszer hatóanyag jelölt molekulákat, amelyek kötnek egy adott gyógyszer célpont fehérjéhez. Az ilyen vizsgálatok során a DKD doménhez egy olyan fehérjét fuzionálnak, amelynek ismerik erős és specifikus kismolekula partnerét. A TAD doménhez fuzionálják a gyógyszer célpont fehérjét. A vizsgálat során olyan hibrid kismolekula könyvtárát használnak, amelynek tagjai a tesztelt gyógyszer hatóanyag jelölt csoportok mellett tartalmazzák a DKD doménhez kapcsolt fehérje kismolekula ligandumát is. A kismolekula könyvtár tagjai az ismert ligandumon keresztül kapcsolódnak a DKD fúziós fehérjéhez. A vizsgálat során sejtenyésző lemezek mintahelyein élesztő sejteket növesztenek, és minden mintahelyhez hozzáadják a hibrid kismolekula könyvtár egy tagját. Ha az adott könyvtár tag gyógyszer hatóanyag jelölt csoportja kapcsolódik a gyógyszer-célpont fehérjéhez, akkor kifejeződik a riporter fehérje. A riporter gén kifejeződése alapján feltételezhetően a gyógyszer célpont fehérjéhez kötő kismolekulákat később megbízhatóbb módszerekkel is tesztelik (ld. 17.4. ábra).



17.4. ábra: Gyógyszer hatóanyag jelölt molekulák kiválasztása kismolekula könyvtárból élesztő három-hibrid módszerrel

17.4. További olvasnivaló a fejezethez

Fields S. and Song O. (1989) A Novel Genetic System to Detect Protein-protein Interactions. *Nature* 340 (6230): 245–246.

Fields S. (2009) Interactive Learning: Lessons from Two Hybrids over Two Decades. *Proteomics* 9 (23): 5209–5213.

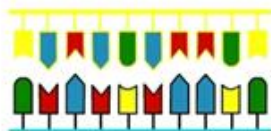
Hamdi A. and Colas P. (2012) Yeast Two-hybrid Methods and Their Applications in Drug Discovery. *Trends in Pharmacological Sciences* 33 (2): 109–118.

Stynen B. Tournu H. Tavernier J. and Van Dijck P. (2012). Diversity in Genetic in Vivo Methods for Protein-protein Interaction Studies: From the Yeast Two-hybrid System to the Mammalian Split-luciferase System. *Microbiology and Molecular Biology Reviews* 76 (2): 331–382.

18. fejezet - Függelék

Néhány géntechnológiai módszer animációja:

Hibridizáció (forrás: <http://www.youtube.com/watch?v=76eKnmCtCTU>)



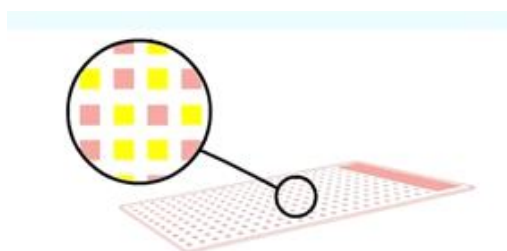
18.1. animáció: Hibridizáció

Fluorescens *in situ* hibridizáció (FISH; forrás: <http://www.youtube.com/watch?v=BBQWWi6cFXU>)



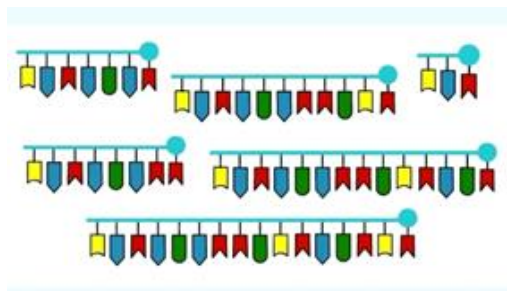
18.2. animáció: Fluoreszcens *in situ* hibridizáció

DNS-chip (*microarray*; forrás: http://www.youtube.com/watch?v=3jX_08zdYCE)



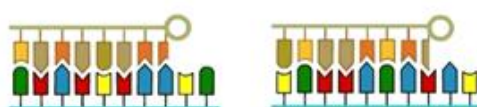
18.3. animáció: DNS-chip (*microarray*)

DNS szekvenálás láncterminációs módszerrel (forrás: <http://www.youtube.com/watch?v=oYpllbI0qF8>)



18.4. animáció: DNS szekvenálás láncterminációs módszerrel

Polimeráz láncreakció (PCR; forrás: <http://www.youtube.com/watch?v=QaWLJVGEFi8>)



18.5. animáció: Polimeráz láncreakció