

Növényi genomika

- Bevezetés - mi is a genomika?
- Növényi genomok szerveződése
- Növényi genomprogramok
- Növényi genomszekvenciák néhány alkalmazása (összehasonlító genomika, génizolálás, translational biology - from models to crops)
- Genetikailag módosított növények

Genomika - bevezetés I.

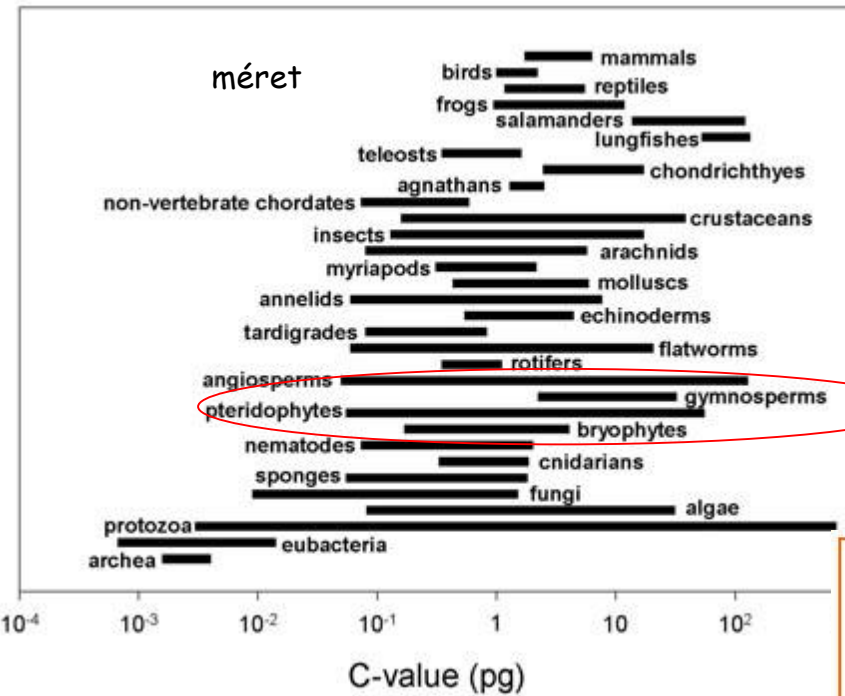
- **Genom:** - a haploid kromoszómaállomány géntartalma (Hans Winker német botanikus, 1920)
 - egy organizmus teljes öröklődő információja
- **Genomika:** - "mapping, sequencing and characterizing genomes" (Thomas Roderick, 1986)
 - egy organizmus teljes örökítő anyagának (nukleáris, mitokondriális, kloroplasztisz DNS-ének), genomszerkezetének és változatosságának vizsgálata és elemzése
- **Funkcionális genomika** - a gének összességének, vagy egy-egy csoportjának a vizsgálata nagy áteresztőképességű kísérletekkel, amelyek vagy a genomszekvencia alapján lettek megtervezve (microarray, SNP), vagy amelyekhez szekvenálást használnak (ChIP-seq, RNA-seq, sRNAseq, methylome, stb.)
 - genom léptékű biológia (nagyobb léptékű DNS, expressziós vagy fehérje vizsgálat) <-> klasszikus molekuláris biológia
- Vizsgálatok **DNS** (strukturális genomika), **RNS** (funkcionális genomika), fehérje (proteomika) és anyagcsere termékek (metabolomika) szintjén - nem teljesen pontos (Isd. pl. SNP analízis - funkcionális genomika)

Genomika - bevezetés II.

- **Újszerű megközelítés:** a teljes genomok ismerete és vizsgálata nagy mennyiségű biológiai információt eredményez
- Ennek az információnak a kiaknázása más szemléletet és új módszereket igényel
- Paradigmaváltás a kutatásban:
 - Klasszikus tudományos kutatási módszer hipotézis alapú
 - Genomika - adatalapú kutatás
 - Kísérletek megtervezése a fenotípus alapján
 - adatgenerálás
 - adatfeldolgozás, szisztematikus analízis
 - új hipotézisek felállítása
 - hipotézis igazolása kísérlettel
- **Alapok:**
 - Metodikai háttér kibővülése: új generációs szekvenálási módszerek, nagy volumenű vizsgálati metódusok, eljárások (pl. microarray/gén-chip)
 - Genomszekvenálási projektek (több száz élőlény teljes genomja ismert, és a lista folyamatosan bővül)
 - Bioinformatika
 - nagy mennyiségű adat kezelése
 - új eredmények kinyerése és megjelenítése
 - egyelőre lemaradásban a metodikai fejlesztések mögött (szűk keresztmetszet)

Növényi genomok szerveződése I.

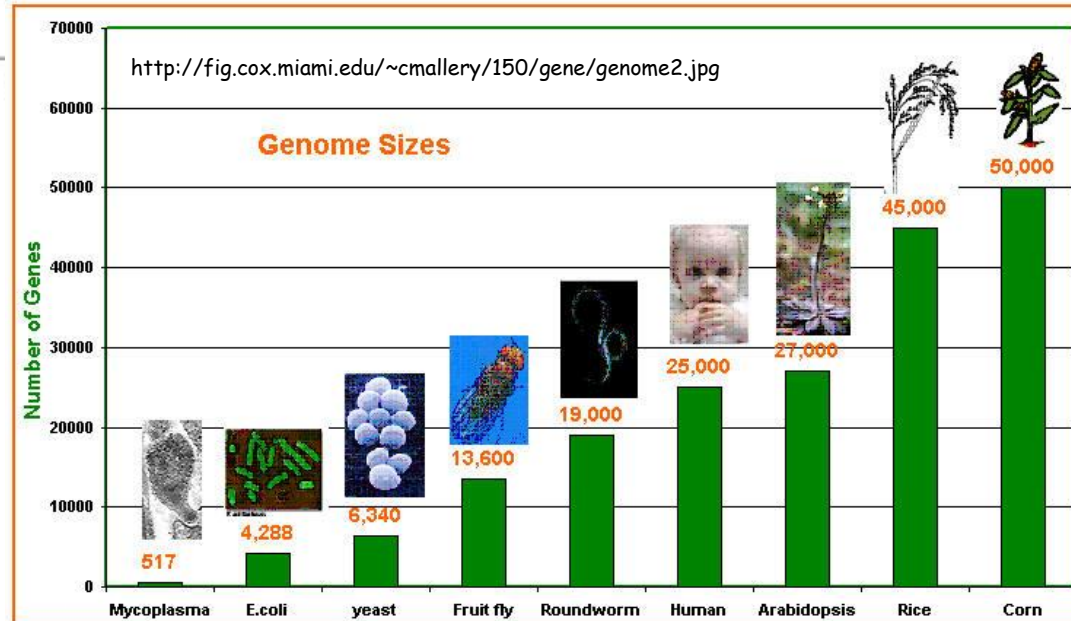
- eukarióta genomok szerveződése komplexebb a prokarióta genomokénál



gének száma növekszik az organizmus fejlettségével (haploid)

	génszám	haploid genom (Mb)
<i>E. coli</i>	4289	4,2
élesztő	6300	12,1
fonalféreg	19000	100
muslica	14000	180
<i>Arabidopsis</i>	26000	125
kukorica	~58000	2500
emlős	~35000	3000

3,2x (between *E. coli* and élesztő)
 2,5x (between élesztő and fonalféreg)
 43x (between élesztő and 100 Mb)
 16,6x (between 100 Mb and 180 Mb)

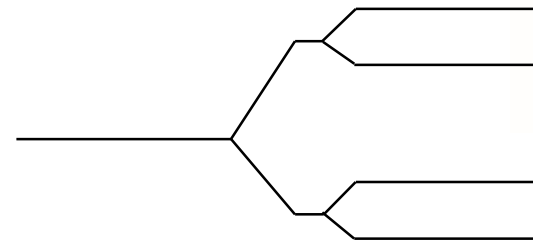


Növényi genomok szerveződése II.

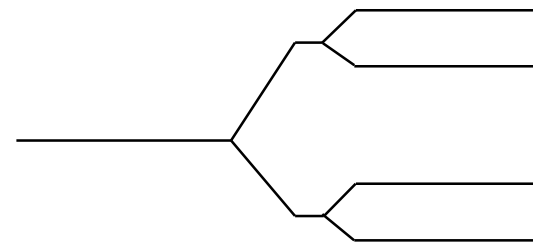
szerveződés:

- kromoszómákba „darabolva” (4 - 224)
relatív vagy ténylegesen alacsony géndenzitás
intronok (törzsfajlódással párhuzamosan számuk és méretük/sejt növekszik)
duplikációk (WGD, géncsaládok, pszeudogének, nagy géncsaládok, kromoszóma szegmentek)
- poliploidia
- (retro)transzpozonok (gypsy/copia-like)
- repetitív szekvenciák, mini/mikro szatellit szekv.
- mtDNS (200 - 2500 kb növényekben)

~ 75 M év



~ 150 M év



Növényi genomok szerveződése III.

- növényi genomok összetettebbek (3 genom)

- kloroplaszt genom (10 - >100 kópia/sejt, 60-100 génszám/mol., 120-160 kb)
- genomméretek (Mb)

	<u>sejtmag (Mb)</u>	<u>kloroplaszt</u>	<u>mitokondrium</u>
Arabidopsis thaliana	125	0,156	0,370
Kukorica (<i>Zea mays</i>)	~ 2500	0,136	0,570
Lóbab (<i>Vicia faba</i>)	~ 14000	0,120	0,290
Humán	~ 3000	-	0,016

- Génduplikációk/géncsaládok (tandem duplikáció, áthelyeződés, szegmentális duplikáció, teljes genom duplikáció (WGD))

Proportion of genes in different organisms present as either singletons or in paralogous families.

Species	No of genes*	Unique gene families containing					
		1	2	3	4	5	> 5 member
<i>H. influenzae</i>	1,587	88.8%	6.8%	2.3%	0.7%	0.0%	1.4%
<i>S. cerevisiae</i>	5,105	71.4%	13.8%	3.5%	2.2%	0.7%	8.4%
<i>D. melanogaster</i>	10,736	72.5%	8.5%	3.4%	1.9%	1.6%	12.1%
<i>C. elegans</i>	14,177	55.2%	12.8%	4.5%	2.7%	1.6%	24.0%
Arabidopsis	11,601	35.0%	12.5%	7.0%	4.4%	3.6%	37.4%
<i>G.max</i> (soybean)	201	35.0%	25.0%	10.4%	10.4%	4.5%	15%

* The number of genes in the genomes of *Haemophilus influenzae*, *S. cerevisiae*, *Drosophila*, *C. elegans*, *Arabidopsis* and *Glycine max* that were present either as singletons or in gene families with two or more members were listed. To be grouped in a gene family, two genes had to show similarity exceeding a BLASTP value E,10⁻²⁰ and a FASTA alignment over at least 80% of the protein length or hybridize at 65 C and 0.1 M ionic strength (Tm - 25). In column 1, the number of genes that were unique plus the number of gene families were listed. Columns 2 to 6 give the percentage of genes present as singletons or in gene families of n members (from TAIR 2000).

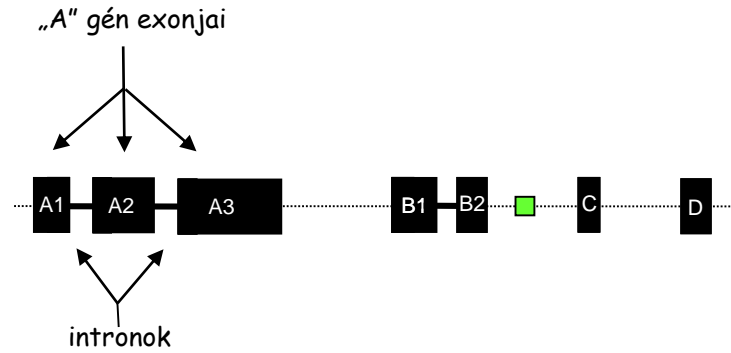
Shopinski *et al. Plant Methods* 2006 2:20 doi:10.1186/1746-4811-2-20

- Poliploidizáció

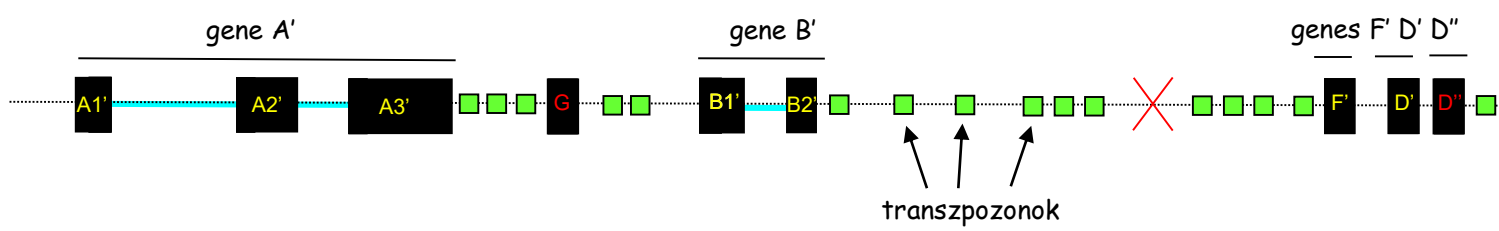
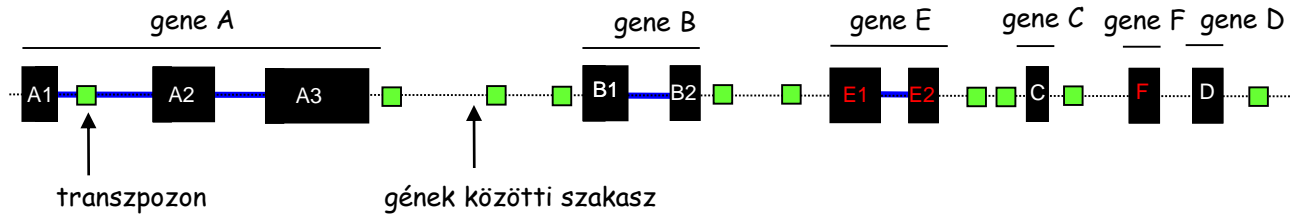
Növényi genomok szerveződése IV.

- Fiktív példa genoméret növekedésre

Arabidopsis (125 Mb, ~ 26 000 gén)

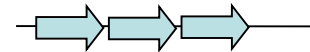


kukorica (2500 Mb, ~ 58 000 gén)

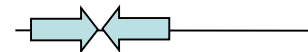


- repetitív szekvenciák

tandem ismétlődés

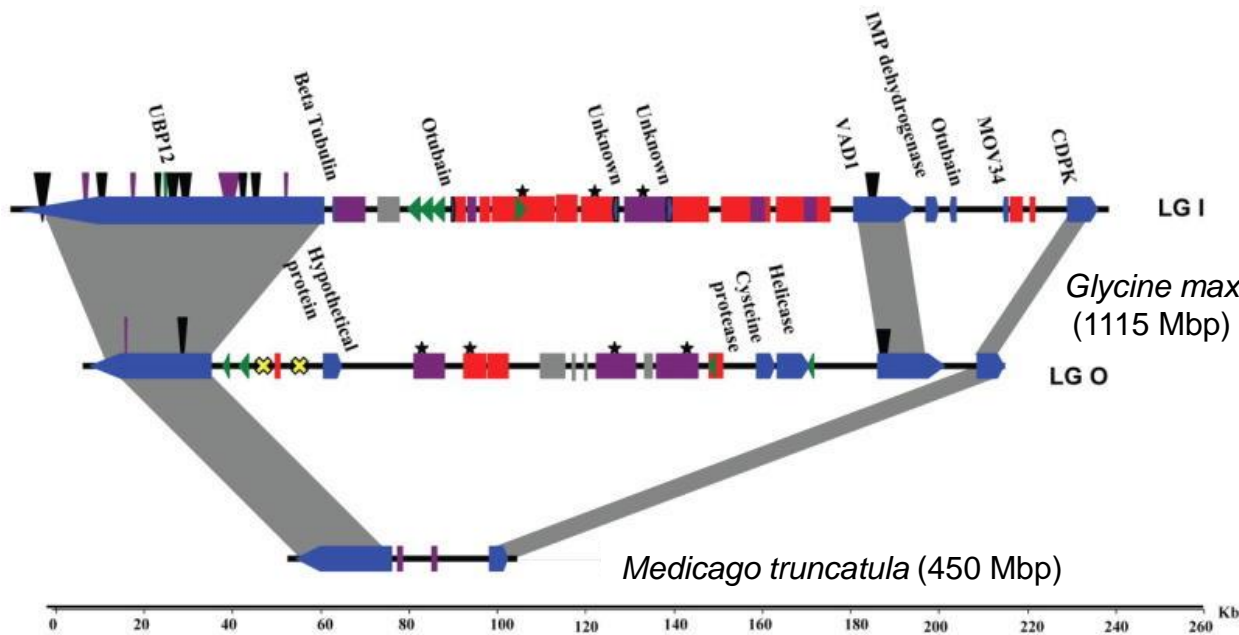


fordított ismétlődés

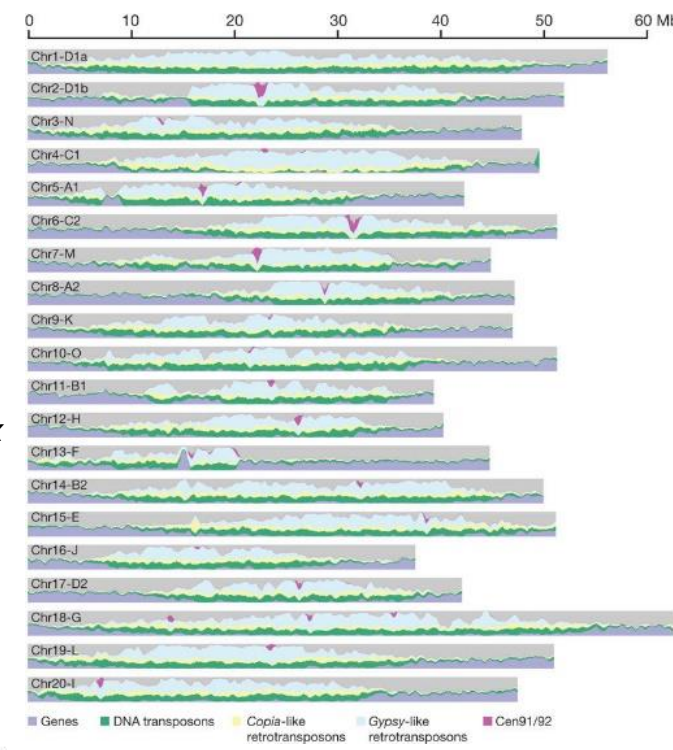


Növényi genomok szerveződése V.

- Konkrét példa genoméret növekedésre



- Copia elements
- Gypsy elements
- Non-LTR elements
- Unknown class elements
- Genes
- solo LTRs



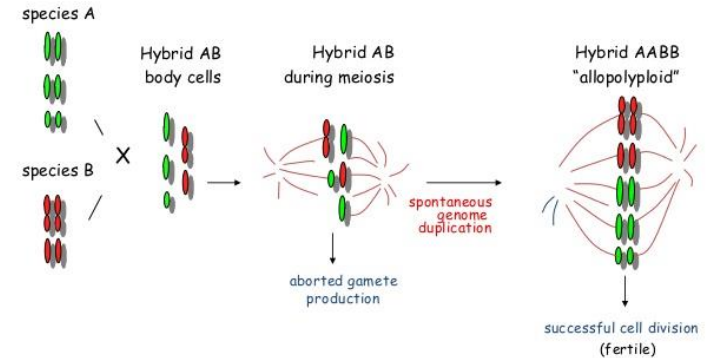
Schmutz et al. 2010 Nature

Joseph et al. 2009 The Plant Genome

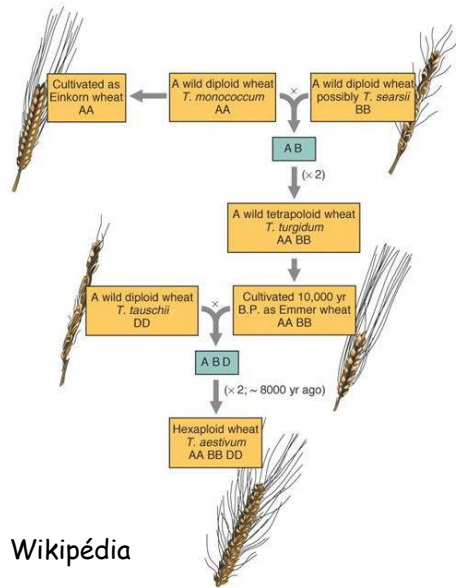
Növényi genomok szerveződése V. Kromoszómakészlet többszöröződése (poliploidizáció)

- alloplidia: genetikailag elhatárolható kromoszómák,
pl. búza allohexaploidia (AABBDD) vagy Brassica

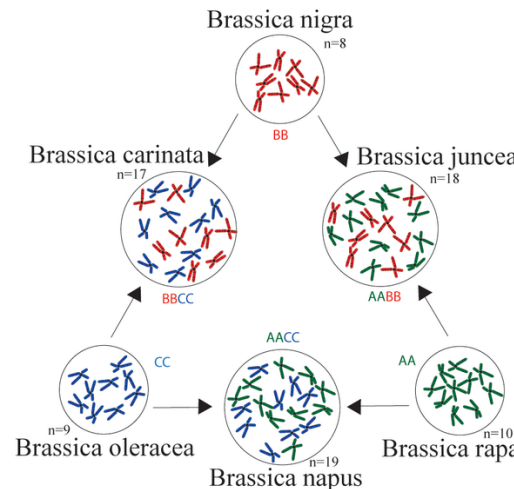
Allopolyploidy arises from hybridization plus genome duplication



Duplicated genomes are fertile !!
Botanical term: Allopolyploids



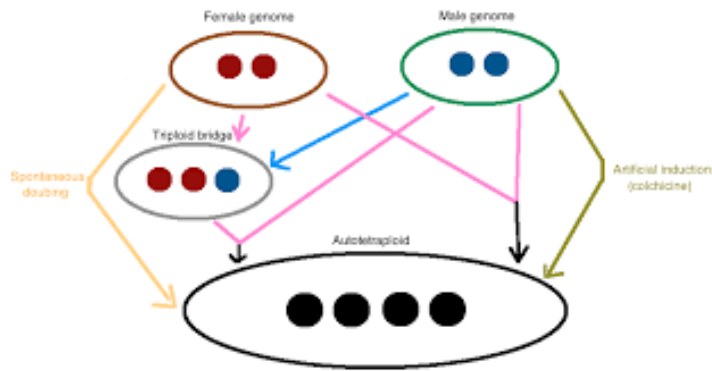
Forrás: Wikipédia



Brassica nigra - fekete mustár
B. carinata - etióp mustár
B. juncea - barna v. indiai mustár
B. oleracea - vadkáposzta
 (káposzta, brokkoli, karfiol, karalábé)
B. napus - repce
B. rapa - tarlórépa

Növényi genomok szerveződése VI. Poliploidizáció

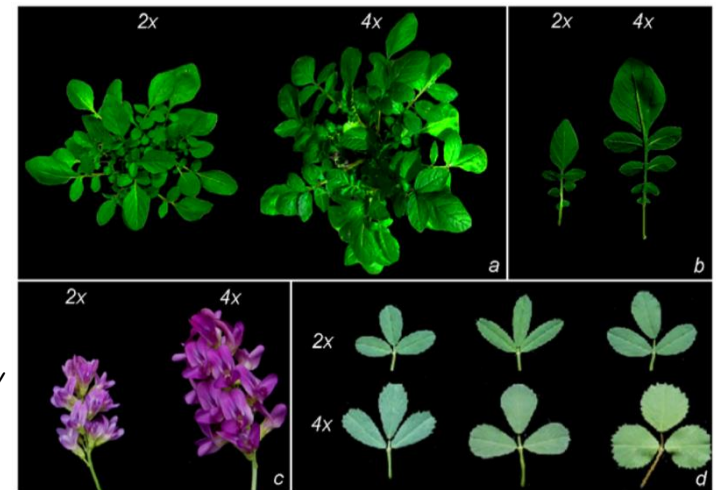
- autopoloidia: egy alap-kromoszómakészlet sokszorozódása, pl. tetraploid lucerna és burgonya



Solanum commersonii

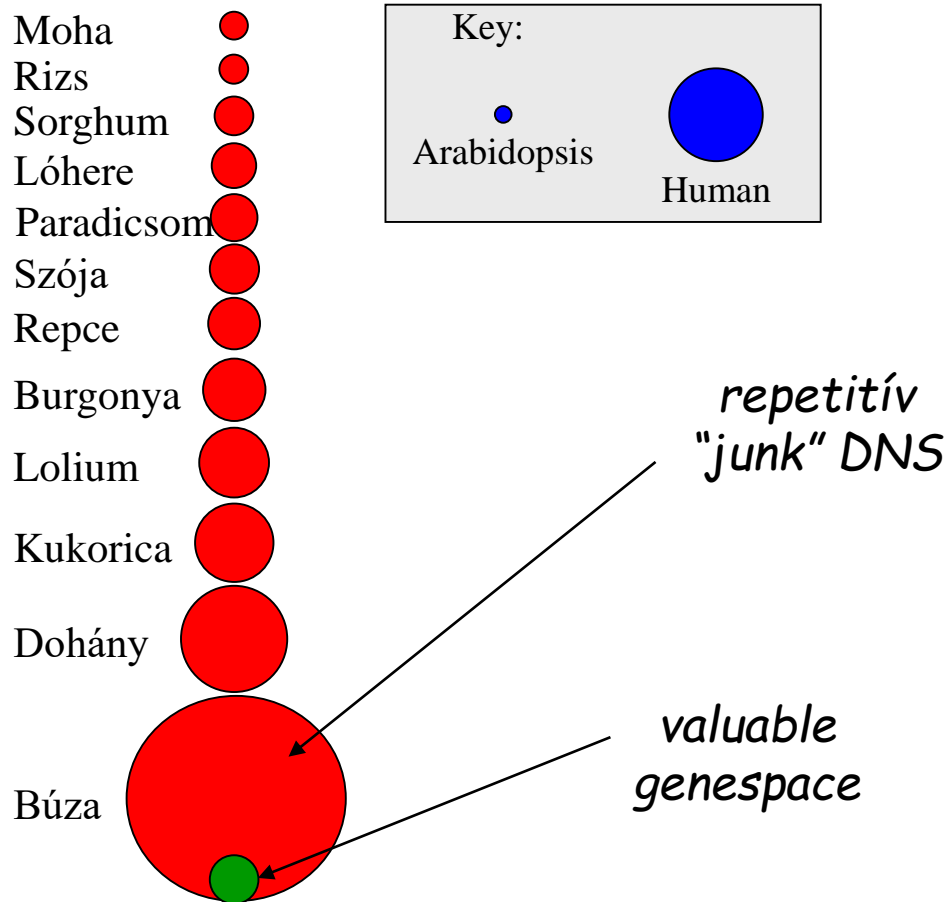
Medicago sativa ssp. *coerulea*/
M. sativa ssp. *sativa*

Aversano et al. 2012



Növényi genomok szerveződése VII.

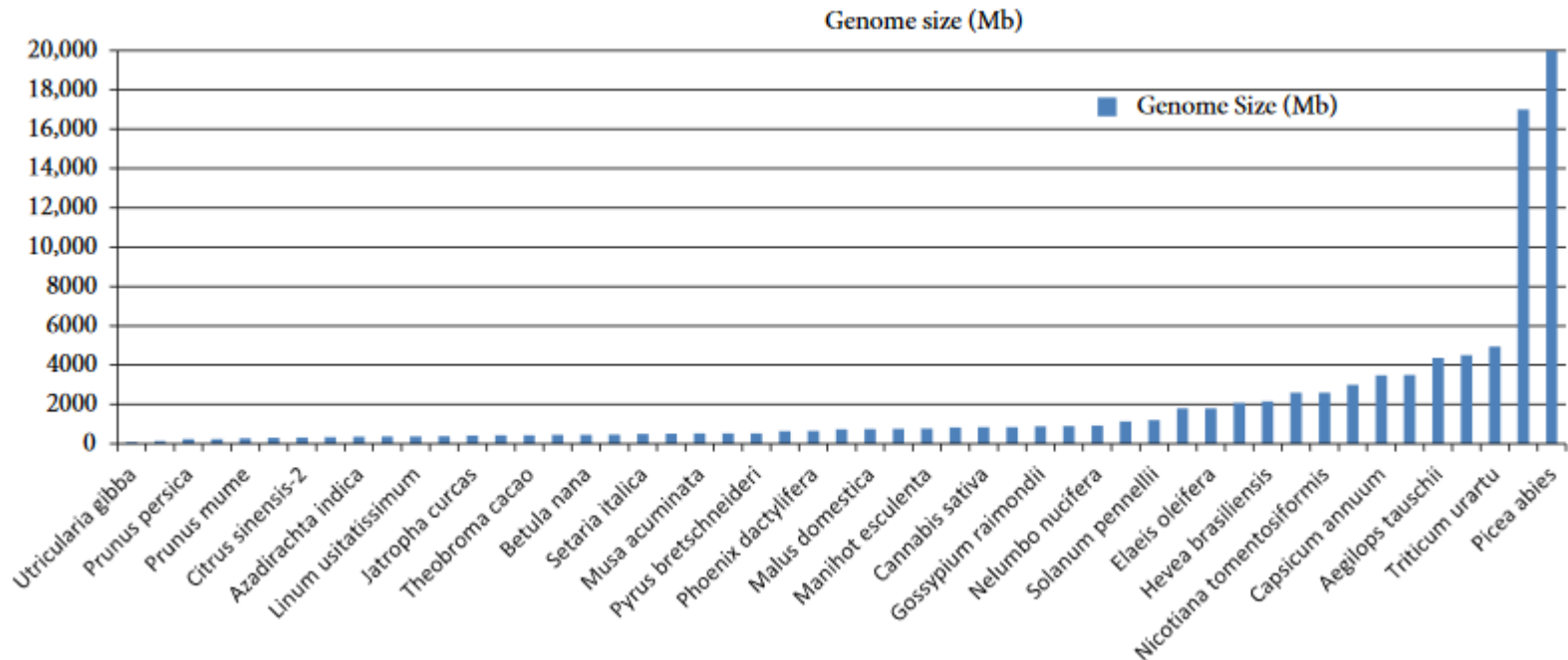
Relatív genomméret



- Legkisebb genom: *Genlisea aurea* (rovarevő) 63 Mbp
- Legnagyobb genom: *Paris japonica* 148000 Mbp, $8n=40$ (négy faj allopoliploid hibridje)
- Legkisebb publikált genomszekvencia: rovarevő renceféle (*Utricularia gibba*) 82 Mbp,
- Legnagyobb publikált genomszekvencia: norvég lucfenyő (*Picea abies*) 19 600 Mbp
- Egyéb nagyméretű publikált genomszekvencia: kukorica - 2300 Mbp
- Paprika - 3300 Mbp
- Átlagos növényi genomméret: 480 Mbp

Növényi genomok szerveződése VIII.

- Legkisebb publikált genomszekvencia: rovarrevő renceféle (*Utricularia gibba*) 82 Mbp,
- Legnagyobb publikált genomszekvencia: norvég lucfenyő (*Picea abies*) 19 600 Mbp



A genomanalízis főbb lépései

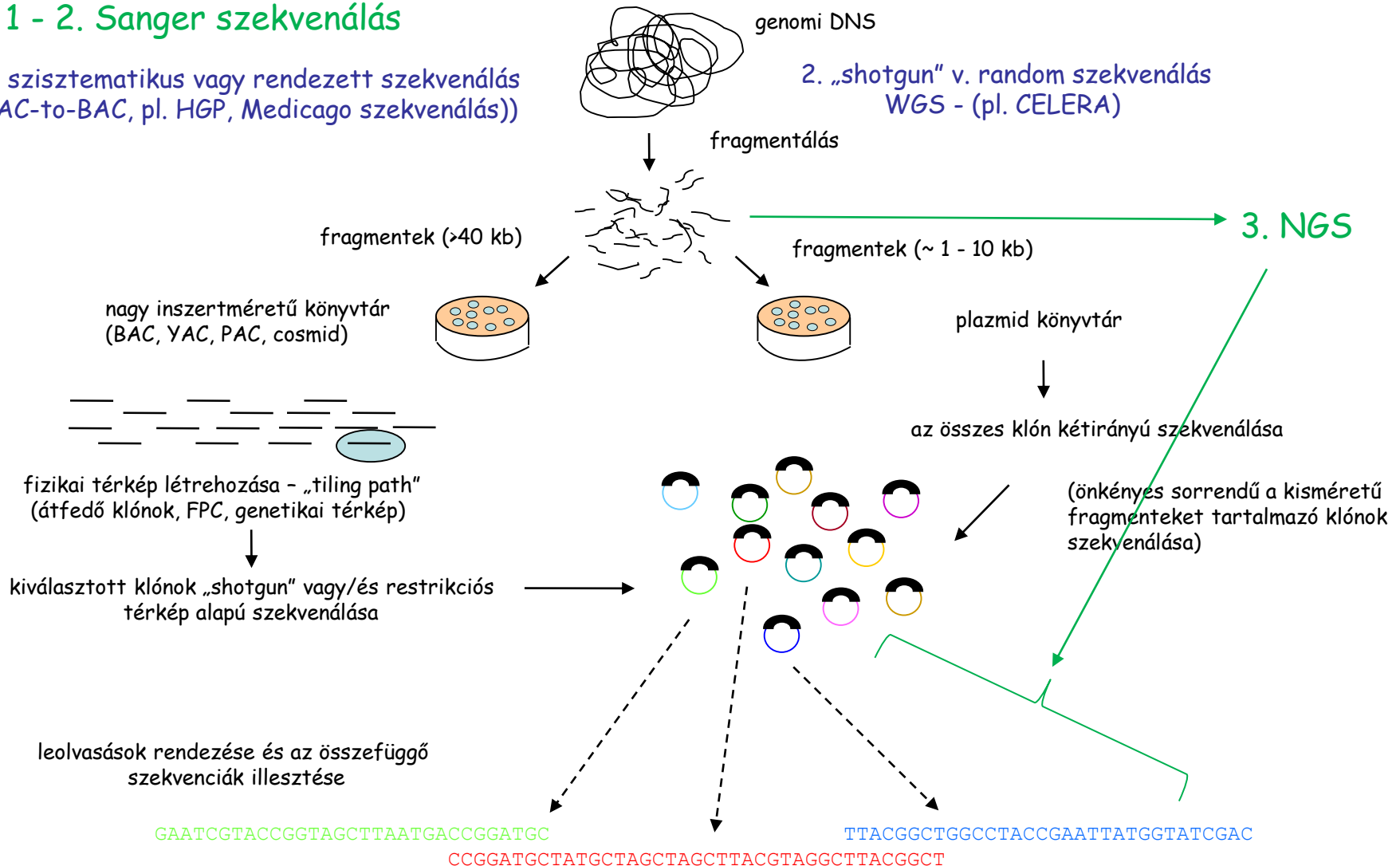
- **Növényi genomprogramok** („mint az állatok esetében csak kicsit máshogyan és lassabban”)
 - növényi alapanyag: referencia faj/vonal, beltenyésztett vonalak, mutánsgyűjtemény, természetes ésvad változatok, stb.
 - DNS-szintű vizsgálatok: genetikai térképezés
fizikai térképezés
genomszekvencia meghatározása, összeszerelése
genomszekvencia analízise (kódoló és nem kódoló gének, génannotáció, ismétlődő szekvenciák, SNP, összehasonlító térképezés, duplikációk azonosítása, stb.)
 - RNS-szintű vizsgálatok: EST szekvenálás, teljes cDNS szekv./RNAseq, alternatív splicing, microarray/gén-chip, géntlasz
 - fehérjeszintű vizsgálatok: lokalizáció, ChIP, fehérje-fehérje interakciók vizsg., teljes fehérje összetétel, fehérje komplexek vizsg., poszttranszlációs módosulások
 - funkcionális vizsgálatok: loss-of-function technikák (mutáció, TILLING, RNAi)
episztázis vizsgálatok (gének ill. géntermékek hierarchiája), szabályozás vizsgálata
gain-of-function technikák (ha van rá lehetőség)
 - metabolika: nagyléptékű anyagcsere vizsgálatok, kémiai „fingerprint”
 - funkcionális génannotáció: ismeretlen gének funkciójának azonosítása
 - evolúciós modellezés

Genomszekvenálási stratégiák I.

1 - 2. Sanger szekvenálás

1. szisztematikus vagy rendezett szekvenálás (BAC-to-BAC, pl. HGP, Medicago szekvenálás))

2. „shotgun” v. random szekvenálás WGS - (pl. CELERA)



Genomszekvenálási stratégiák II.

```

<-- GAATCGTACCGGTAGCTTAATGACCGGATGC          <-- TTACGGCTGGCCTACCGAATTATGGTATCGAC
GAATCGTACCGGTAGCTTAATGACCGGATGC -->          TTACGGCTGGCCTACCGAATTATGGTATCGAC -->
          CCGGATGCTATGCTAGCTAGCTTACGTAGGCTTAGGGCT -->
<-- CCGGATGCTATGCTAGCTAGCTTACGTAGGCTTAGGGCT
          <-- TTACGTAGGCTTACGGCTGGCCTACCGAA
          TTACGTAGGCTTACGGCTGGCCTACCGAA -->
          <-- GCTGGCCTACCGAATTATGGTATCGACCTGACC
          GCTGGCCTACCGAATTATGGTATCGACCTGACC -->

GAATCGTACCGGTAGCTTAATGACCGGATGCTATGCTAGCTAGCTTACGTAGGCTTACGGCTGGCCTACCGAATTATGGTATCGACCTGACC
  
```

kész szekvencia - minden szál kb. 4-szer szekvenálva, 8X lefedettség

Szisztematikus szekvenálás

Random szekvenálás

Előnyök

- minden egyes lépésnél tudjuk hol járunk a genomban
- kevesebb hiba az összeszerelés során, összeszerelés viszonylag egyértelmű (ismert térkép és kontig pozíció)
- lehetőség a génekben gazdag régiók szekvenálására

- a szekvencia összeszerelése későbbi fázisában derül ki, hogy a genom mely részét szekvenáljuk

- nagyobb esély a hibás összeszerelésre (pl. duplikációk, repetitív szekvenciák - gabonafélék)
- automatizálható
- kisebb genomok esetén hatékonyabb, nagyobbak esetén problémásabb

Hátrányok

- költségesebb és munkaigényesebb
- lassabb

- kevésbé költséges és munkaigényes
- gyorsabb
- hatalmas klónkönyvtárat igényel
- nagy számítógép-kapacitást igényel

Genomszekvenálási stratégiák III.

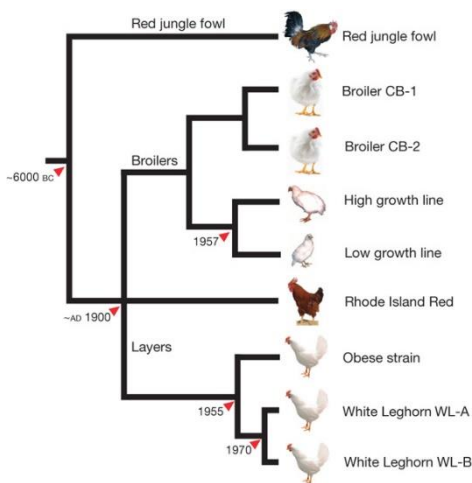
Alternatív megközelítések:

- nem teljes genomszekvenálás

- klónozás DNS frakcionálást követően (a gyakori fragmentek mennyiségének csökkentése RE emésztést követően)
- egyes kromoszómák izolálása
- gén gazdag régiók feldúsítását követően (kevésbé metilált alacsony kópiás régiók, CoT analízis)

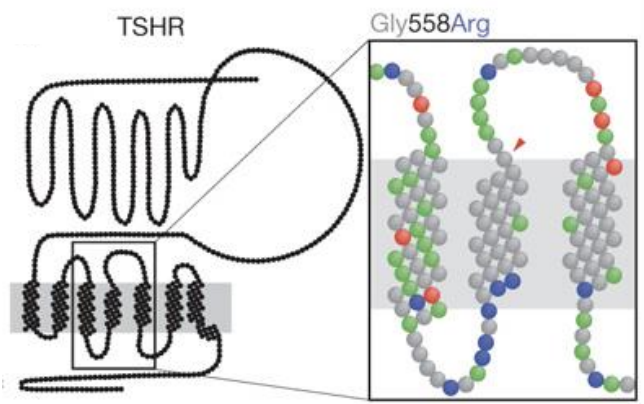
- Re-sequencing <-> de novo sequencing

- genetikai eltérések keresése és variációk azonosítása (SNPs, INDELS) egész genomban vagy genom részletekben
- referencia szekvencia szükséges
- új generációs szekvenálási módszerek
- humán genom SNP map (Nature 2000), C. Venter és J. Watson (PlosBiol 2007 and Nature 2008)
- 1000 humán genom diverzitásának meghatározása (Science 2008)
- Arabidopsis Col-0, Bur-0 és Tsu-1 - 2000 hiba a referencia genomban, 800 000 SNP és 80 000 INDEL azonosítása
- SNP azonosítás rizs és *Medicago truncatula* genotípusokban (30 *M. truncatula* genotípus szekvencia - 2010)
- 25 vad és 25 termesztett rizs alacsony lefedettségű genomszekvenálása - eltérés génekben és fenotípusban
 - Mezőgazdasági Biotechnológiai Kutatóközpont, Gödöllő - mangalica - házisertés (Duroc)
- a tyúkháziasítás genetikai hátterének felderítése genomikai módszerekkel (SOLID; -7 millió SNP, 1300 INDEL)





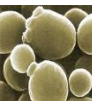










thyroid stimulating hormone receptor - a szaporodás fotoperiódusos szabályozása

MLGGWLS	SCFL	LALL	PLVGV	SSYS	KVSI	CL	
-----	-----	-----	-----	-----	-----	-----	R
-V--	VC	-----	I	-----	A	-----	-----
-V--	VC	-----	I	-----	A	-----	-----
-A--	V	-----	M	-----	I	-----	A
-V--	VC	-----	I	-----	A	-----	-----
-A--	VC	-----	I	-----	A	-----	-----
-V--	VC	-----	I	-----	A	-----	-----
-L--	VC	-----	I	-----	A	-----	-----
-L--	AF	I	-----	I	V	-----	-----
-L--	VF	I	-----	V	Q	-----	-----
-L--	I	F	L	-----	V	Q	-----
-L--	F	L	-----	V	Q	-----	-----
-LV-	I	LS	L	-----	M	VL	V
-LV-	I	LS	L	-----	M	VL	V
-A--	FSV	V	-----	GM	Y	V	S



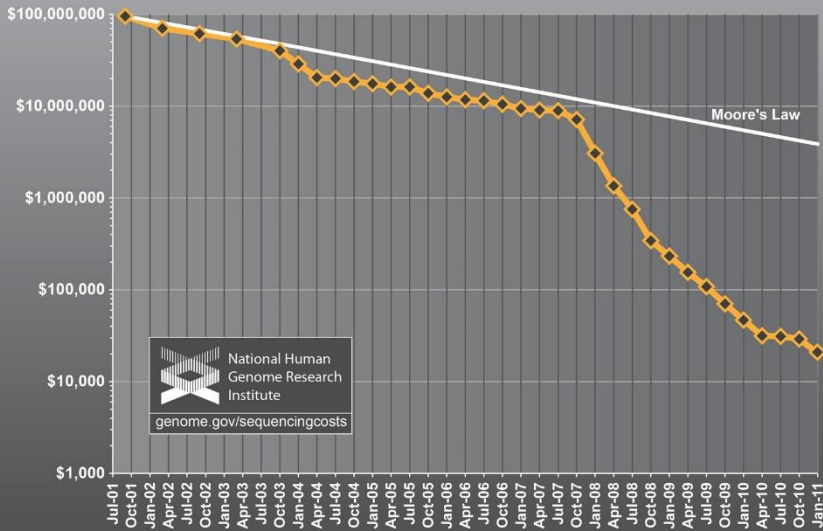
Mérőkövek a genomszekvenálásban

Organizmus/genom	Haploid genomméret (Mb)	Haploid kromoszóma-szám	Kromoszóma-szám	Gének száma	Genomszekvenancia			Megjegyzés
					közlés	labor	módszer	
 φX174 bakteriofág	0,005	cirkuláris		11	1977	Sanger et al.		első megszekvenált organizmus
Humán mtDNS	 0,016	1 cirkuláris		37	1981	Sanger et al.		
 <i>Haemophilus influenzae</i>	1,8	1 cirkuláris		1 743	1995	TIGR	shotgun	első megszekvenált baktérium
 <i>Escheria coli</i> K-12	4,2	1 cirkuláris		4 290	1997	U. Wisconsin, Madison	shotgun	
 sörélesztő (<i>Saccharomyces cerevisiae</i>)	12,1	16	2n=32	5 770	1996	konzorcium	térképezés	első teljesen megszekvenált eukarióta genom
 fonálféreg (<i>Caenorhabditis elegans</i>)	100	6	2n=12	18 424	1998	konzorcium	térképezés	első teljesen megszekvenált többsejtű genom
 ecetmuclica (<i>Drosophila melanogaster</i>)	180	4	2n=8	13 601	2000	UC Berkley Celera Genomics	shotgun/BAC térkép	
 human (<i>Homo sapiens</i>)	3 000	23	2n=46	~ 30 000	2000	Human Genome Project & Celera Genomics	shotgun/térképezés	
 lúdfű (<i>Arabidopsis thaliana</i>)	125	5	2n=10	~25 000	2000	konzorcium	shotgun/BAC térkép	első megszekvenált növényi genom
 rizs (<i>Oryza sativa</i>)	389	12	2n=24	~ 28 000	2005	konzorcium	shotgun/BAC térkép	első megszekvenált egyszikű
új generációs szekvenálási módszerek megjelenése					2005			
 nyárfa (<i>Populus trichocarpa</i>)	485	19	2n=38	~ 45 000	2006 (draft seq)	konzorcium	shotgun/BAC térkép	első megszekvenált fa
 kukorica (<i>Zea mays</i>)	2 500	10	2n=20	~ 32 000	2009 (draft seq)	konzorcium	shotgun/BAC térkép	
 <i>Physcomitrella patens</i> (moss sp.)	520	27		~ 36 000	2008	konzorcium	shotgun	
grandiózus szekvenálási projektek meghirdetése					2010	humán 1000 genomprojekt, 100 Solanaceae projekt, 1000 növényi és állati genom projekt, stb.		

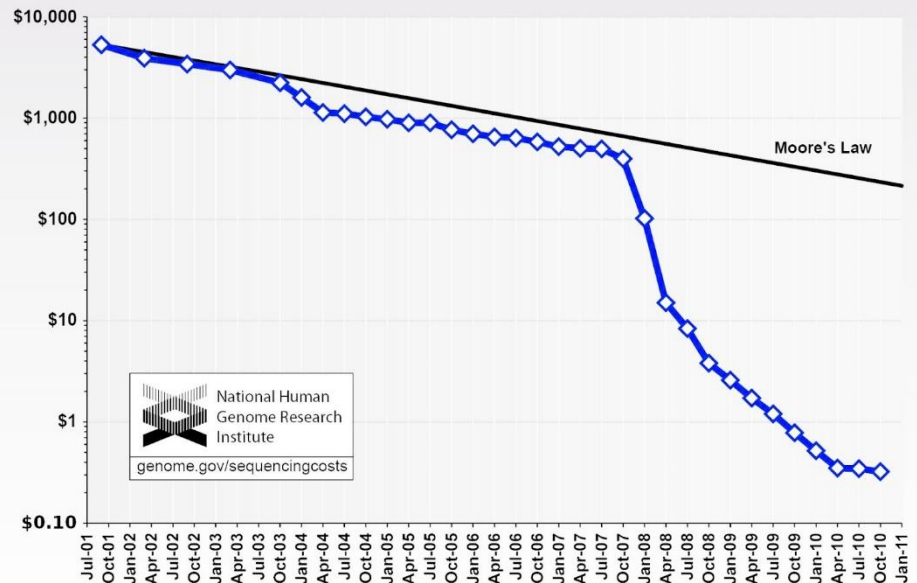
Növényi genomszekvenálási projektek áttekintése I. (2013)

- az utóbbi 10 évben a szekvenálási költségek kb. 10 000-szeresére csökkentek → hamarosan nagy komplex genomok előzetes genomszekvenciájának meghatározására is lehetőség nyílik (*Triticea* 5-17Gb, *Pinus* spp. 18-40 Gb)

Cost per Genome



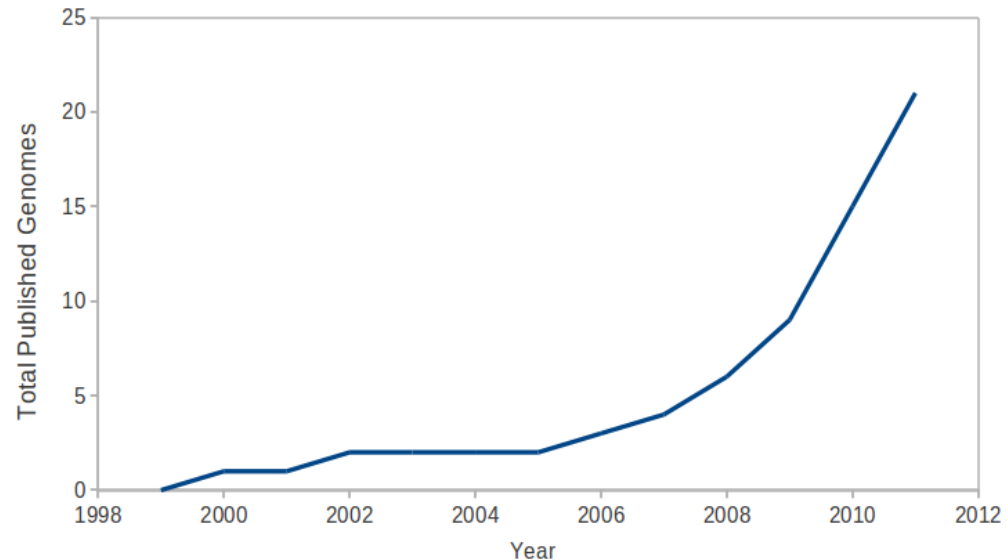
Cost per Megabase of DNA Sequence



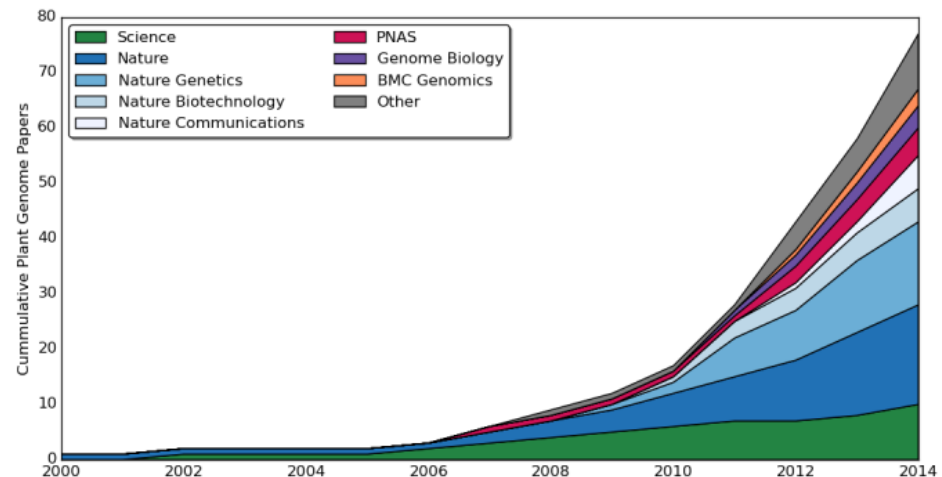
Növényi genomszekvenálási projektek áttekintése II. (2017)

- 19 növényfaj 25 genomszekvenciája publikus, 15 folyamatban vagy pedig nem publikus
(http://genomevolution.org/wiki/index.php/Sequenced_plant_genomes, TIPS 2011, Plant Journal 2012)
- 2017 - 115 növényfaj genomja ismert teljesen vagy majdnem teljesen

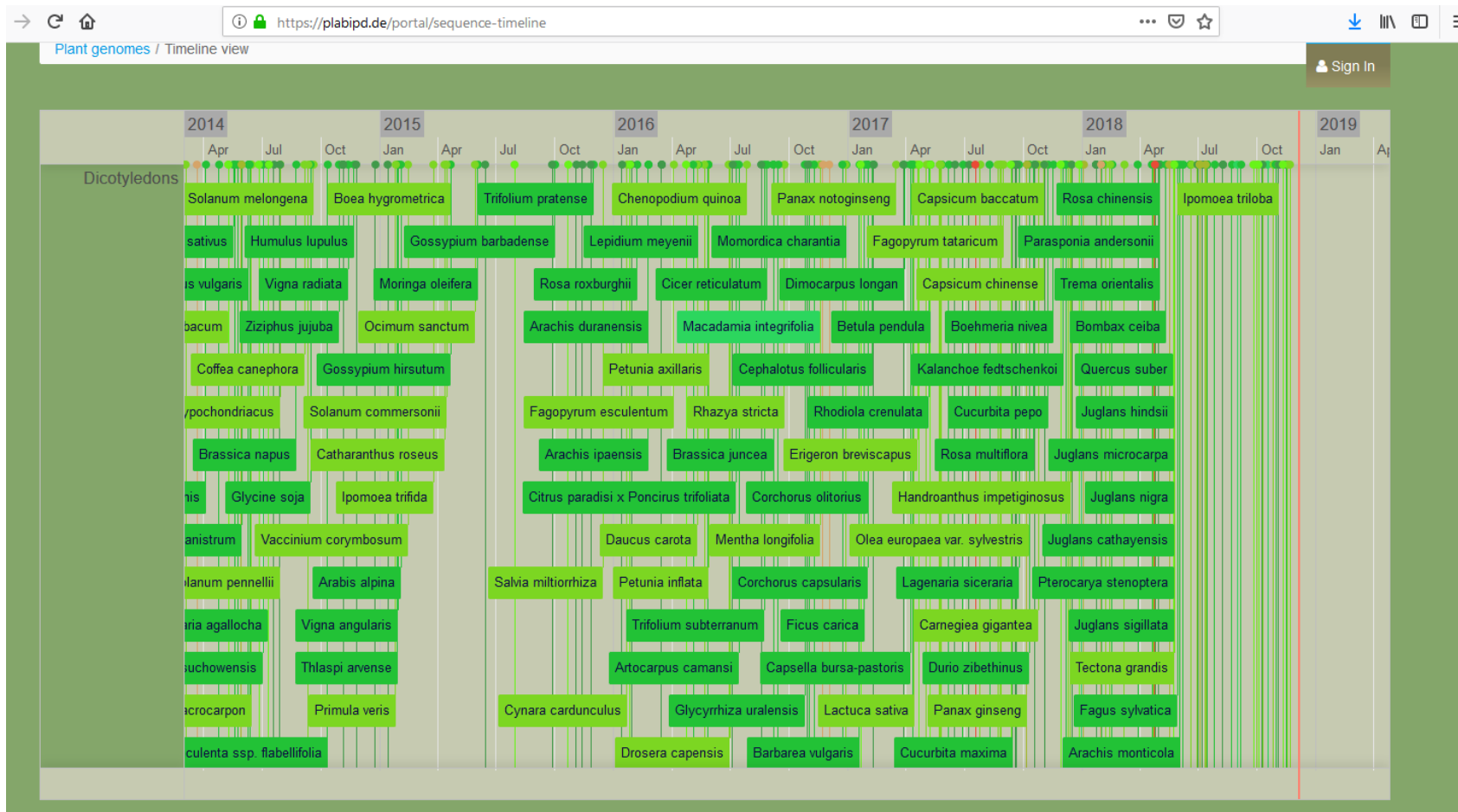
Growth in Plant Genomes



- Algák, mohák, páfrányok - 23 faj
- Magasabb rendű növények:
 - zárwatermők: > 67 kétszikű; > 22 egyszikű
 - nyitwatermők: legalább 3 faj

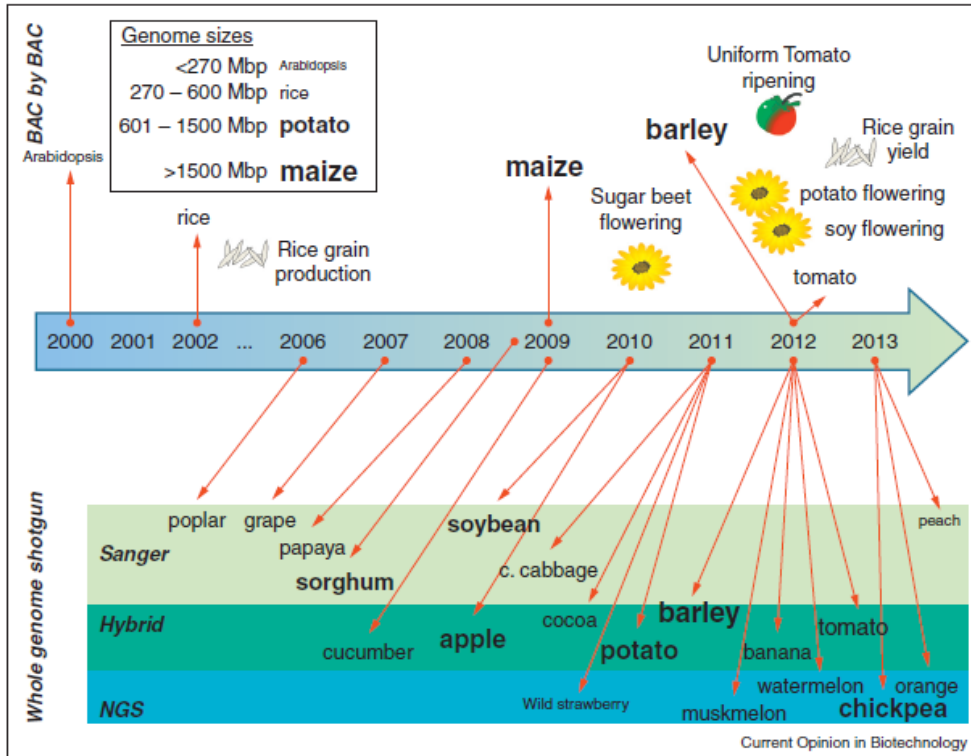


Növényi genomszekvenálási projektek áttekintése III.

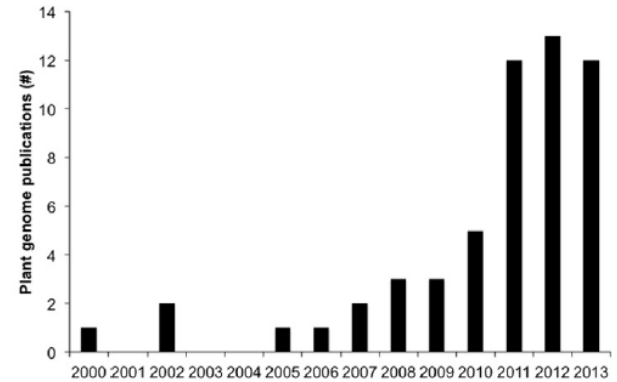


<https://plabipd.de/portal/sequence-timeline>

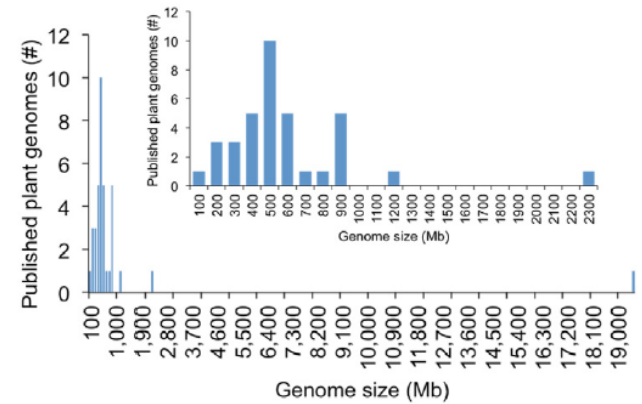
Növényi genomszekvenálási projektek fontosabb állomásai



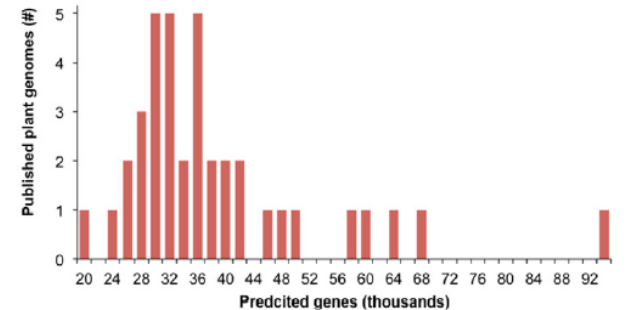
A



B



C

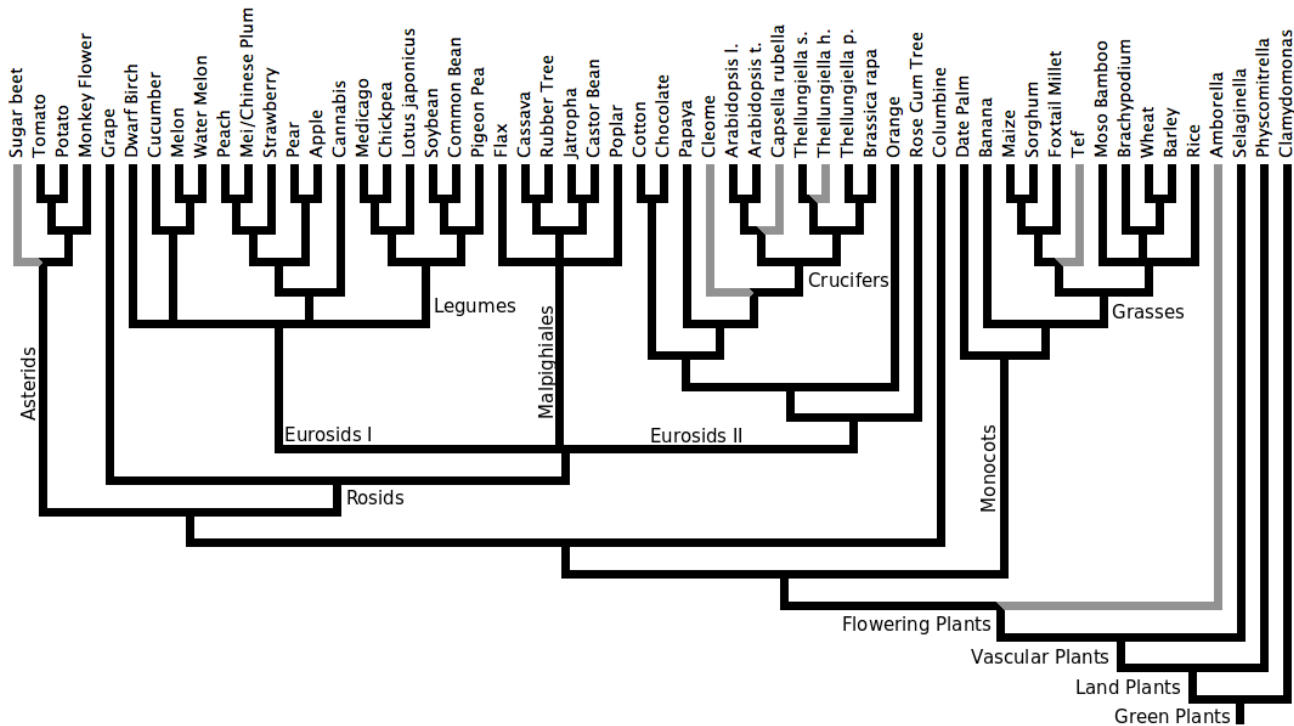


T.P Michael and S. Jackson 2013.

Bolger et al. 2014.

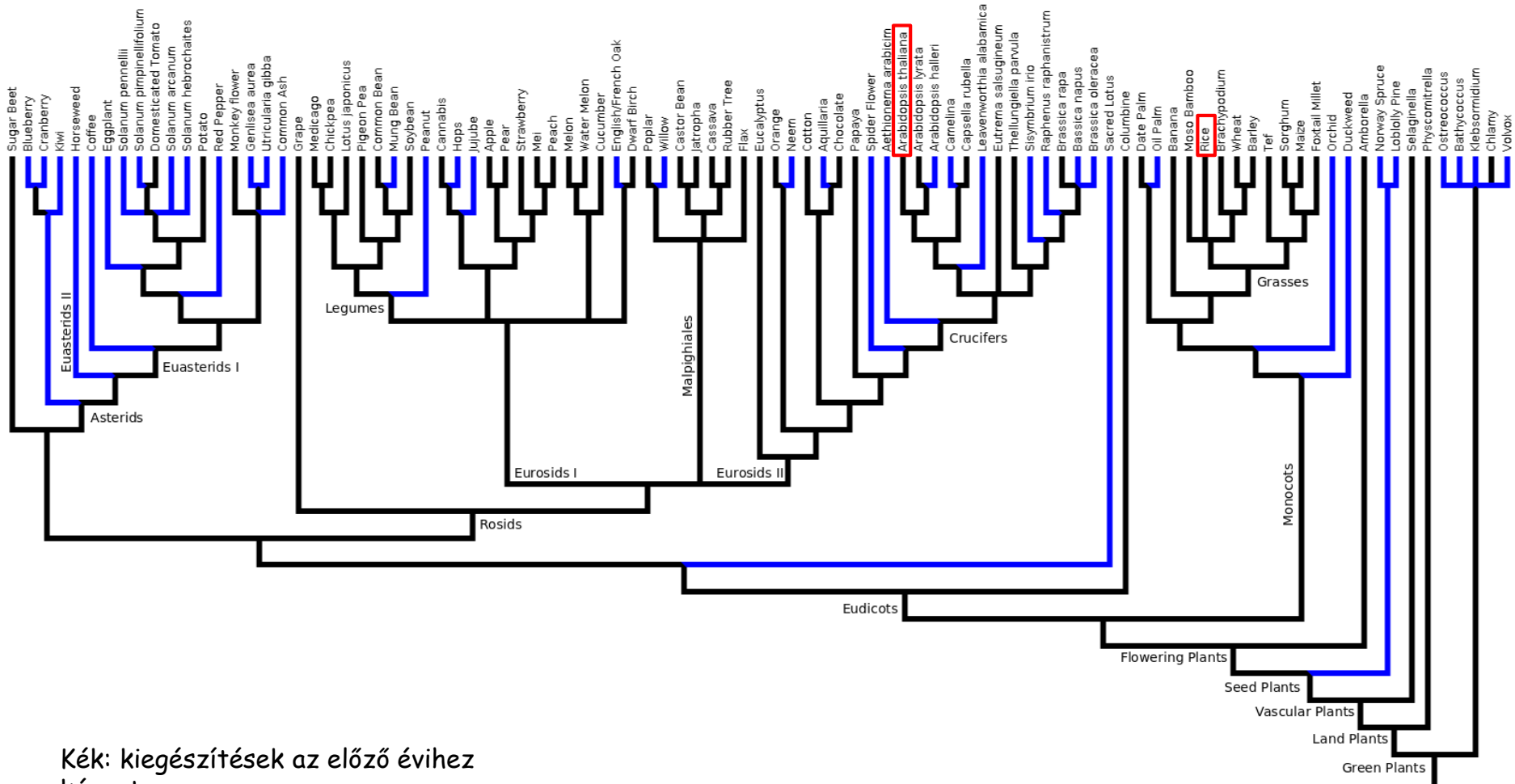
Növényi genomszekvenálási projektek áttekintése III.

- főbb genomszekvenálási szempontok:
 - növénybiológiai modellnövény (*A. thaliana*, *Brachypodium distachyon*, *Medicago truncatula*)
 - haszonnövény (rizs, szója, kukorica, szőlő, stb.)
 - evolúciós vizsg. (pl. *Amborella trichopoda* - legkorábbi zárvatermő)
 - molekuláris eszköz a növénynevelésben
- csak a rizs genomszekvenciájának minősége éri el az *Arabidopsis* szekvenciáét



Fekete: publikált genomszekvencia
Szürke: nem elérhető vagy
kezdetleges állapotú genomszekv.

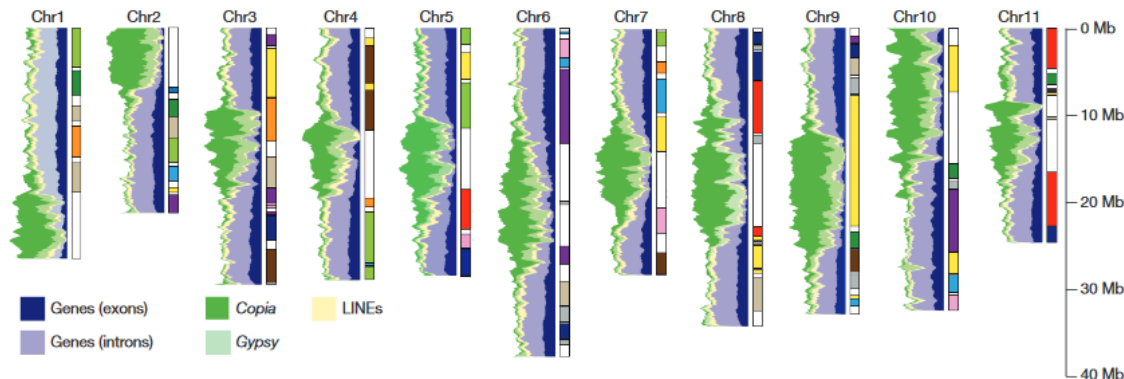
Növényi genomszekvenálási projektek áttekintése IV.



Kék: kiegészítések az előző évihez képest
2014 ős

Növényi genomszekvenálási projektek áttekintése IV.

- a növényi genomok számos tulajdonsága komoly kihívást jelent a genomszekvenálás során
 - Méret: fenyőfélék átlag 7,9 Gb (legkisebb is akkora mint a kukorica genomja)
 - Komplexitás: ploidia ($2n - 8n$), repetitív elemek ($\sim 10\%$ Arabidopsis, $>80\%$ búza), heterozigótáság (szőlő, lucerna)
- két legnagyobb, sok repetitív szekvenciában dús genom szekvenciája is elkészült (kukorica 2,5 Gb, szója 1,1 Gb), ezek egyelőre nem teljesek (80-85%) (BAC-to-BAC szekvenálás kukoricában)
- problémák a valódi poliploidok genomszekvenciájának összeszerelésében (banán, burgonya, gyapot, búza, cukorrépa) → *Mussa acuminata* genom szekvencia



- N50 - a genom-összeszerelés minőségét jelző szám: a kontigok és scaffoldok hossza 50%-os genom összeállítás esetén (legjobb N50 Sorghum 62,4 Mb, *Brachypodium distachyon* 59,3 Mb, szója 47,8 Mb, köles 47,3 Mb)

Arabidopsis thaliana (lúdfű) genom projekt I.



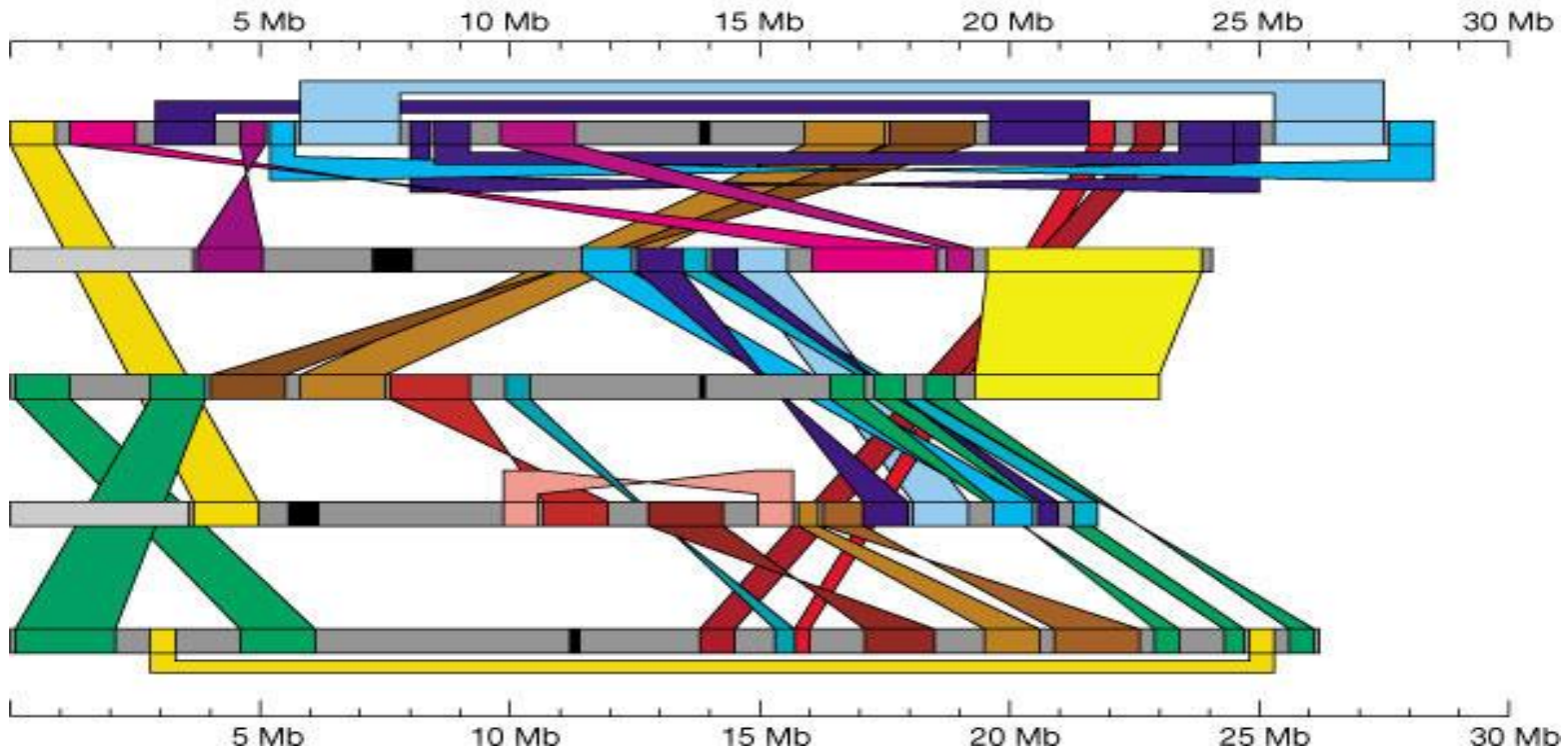
- kistermetű (20-25 cm), kétszikű gyomnövény (Európa, Ázsia, ÉNy Afrika)
- életsiklusa 6 hét (csírázástól magérésig); genomméret 125 Mb
- Johannes Thal (-> *thaliana*), (*Pilosella siliquosa*) XVI szd. Harz-hegység
- F. Laibach javasolta először modell növénynek 1943; ($2n=10$)
- Rédei György (1921-2008); U. Missouri, 1957-ben kezdett el foglalkozni komolyabban *Arabidopsis*-szal; nem volt elismert kezdetben (1969 NSF támogatás felfüggesztése) - mutációs vizsgálatok alapján a *A. thaliana* gének számát 27 813-ra becsülte
- Marteen Korneef - Wageningen University, 1976 (részletes genetikai térkép 1983)
- 1990 - genomprojekt indulása
- 1996 - *Arabidopsis* Genome Initiative (BAC-by-BAC)
- 2000 - genomszekvencia kész
nagy szegmentális duplikációk
- 2001 - NSF2010 Program - gének funkciójának azonosítása
- 2008 - 1001 genom projekt - SNP; (HAPMAP, teljes genomváltozatok) - 1001 *Arabidopsis* változat szekvenálása - 88 már elérhető (2010), befejezést 2012-re ígérték

www.arabidopsis.org (pl. At1g01480)

www.weigelworld.org/resources/microarray/AtGenExpress/

Arabidopsis egy ősi tetraploid genom

hasonlóan a növények többségéhez



A genom 60 % of (67.9 Mb) duplikálódott kromoszómaregiókból áll
A poliploidizáció jelentősen hozzájárult a genom formálódásához és fontos szerepet játszott a növények evolúciójában
A növények kb. 30-80%-a poliploid

Arabidopsis thaliana genom projekt II.



Summary of Gene Structure Annotation (TAIR9 release, 2011)

- 27,379 protein coding genes
- 4827 pseudogenes or transposable elements and 1312 ncRNAs
- 33,518 genes in all, 39,640 gene models

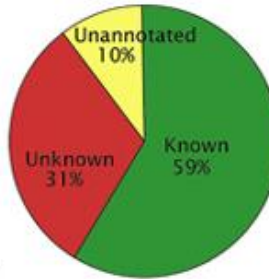
TAIR9 Genome Statistics										
Chr	Protein coding	pre-tRNA	rRNA	snRNA	snoRNA	miRNA	Other RNA	Pseudo gene	TE gene	Total
1	7,054	240	0	2	18	52	107	242	683	8,398
2	4,237	96	2	0	15	29	75	218	825	5,497
3	5,436	93	2	7	15	29	67	201	878	6,728
4	4,214	79	0	0	11	28	48	121	711	5,122
5	6,318	123	0	4	12	36	53	144	804	7,494
Total	27,169	631	4	13	71	174	350	926	3,901	33,239
C	88	37	8	0	0	0	0	0	0	133
M	122	21	3	0	0	0	0	0	0	146
Total	27,379	689	15	13	71	174	350	926	3,901	33,518

Arabidopsis thaliana genom projekt III.

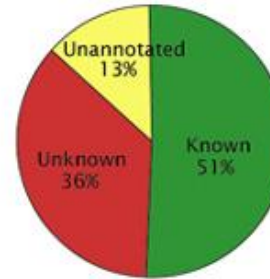


Summary of Gene Functional Annotation

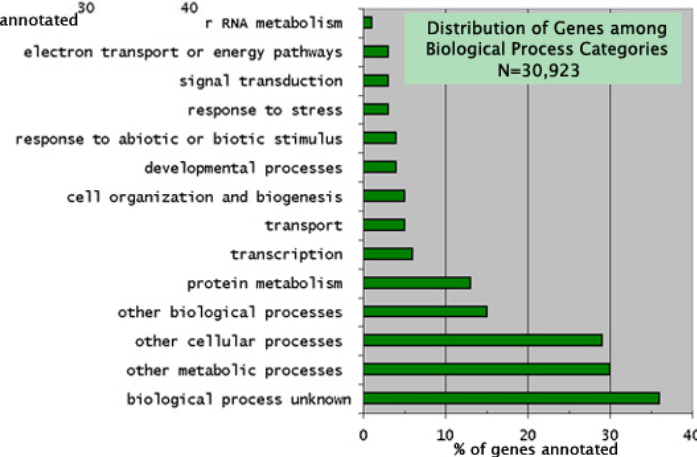
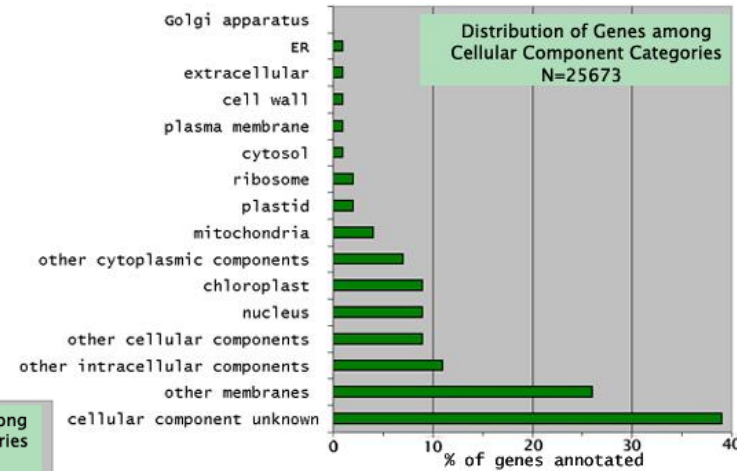
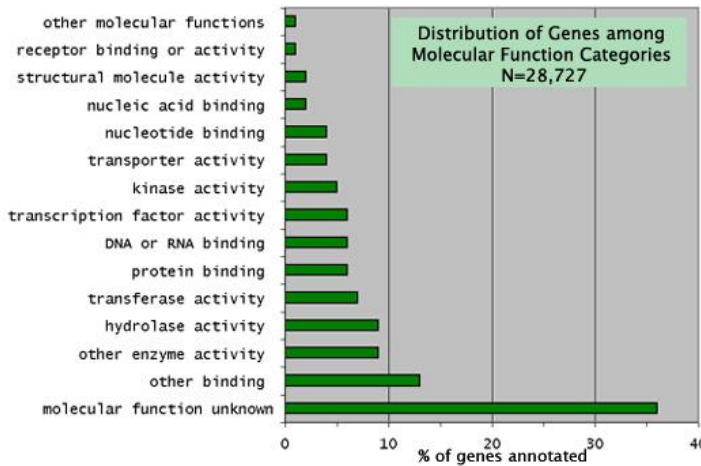
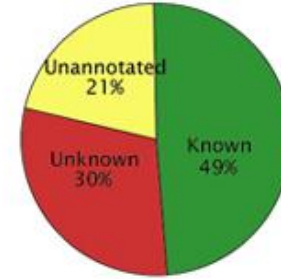
Molecular Function



Biological Process



Cellular Component



Arabidopsis thaliana genom projekt - egyedi gének

- „orphan genes” - gének (kódoló szekvenciák), melyeknek nincsenek homológjai más fajokban vagy rendszertani kategóriában
- „rossz annotáció” vagy valódi jelenség?
- faj specifikus metabolikus útvonalak vagy szabályozó utak génjei, gyakran valamilyen speciális környezeti körülménnyel kapcsolatosak
- NGS; minden organizmusban előfordulnak
- általában a gének 5-15%-a tartozik ebbe a kategóriába
- *Arabidopsis thaliana* protein kódoló gének törzsrétegződése
 - homológia azonosítás
 - legkorábbi rendszertani kategóriába sorolás, ahol van homológ

	Clade-specific genes (Number)	(%)	Age (Ma)
Cenozoic	<i>Arabidopsis thaliana</i>	1084 (4.0%)	0
	<i>Arabidopsis</i>	216 (0.8%)	13
	Camelineae	172 (0.6%)	23
	Brassicaceae	900 (3.3%)	31
Mesozoic	Brassicales	32 (0.1%)	71
	Malvids	146 (0.5%)	96
	Rosids	218 (0.8%)	104
	Core eudicotyledons	510 (1.9%)	125
Phanerozoic	Magnoliophyta (flowering)	971 (3.6%)	170
	Spermatophyta (seed bearing)	946 (3.5%)	329
	Tracheophyta (vascular)	475 (1.7%)	496
Precambrian	Embryophyta (land)	3840 (14.1%)	547
	Viridiplantae (green plants)	1031 (3.8%)	968
	Eukaryota	8229 (30.3%)	1628
	Cellular organisms	8433 (31.0%)	2520
	Total: 27203		

Rizs genom projekt I.



- egyszikű, diploid ($2n=24$), kis genomméret a fűfélék között (389 Mb)
- a rizs a kukoricát követően második a termelt mennyiséget tekintve
- 1997 - International Rice Genome Sequencing Project (IRGSP) - *Oryza sativa* ssp. *japonica* cv. Nipponbare (ragadós, rövidszemű rizs)
- 2002 chr1 és 4
- 2003 chr10
- 2004 teljes genom szekvencia - 370 Mb
- 37 544 gén modell (2,859 nem mutat hasonlóságot *Arabidopsis* génekhez)
- *Oryza sativa indica* (hosszúszemű) genomszekvencia meghatározása szintén befejeződött
(2002 - draft szekv.)

<http://rgp.dna.affrc.go.jp/IRGSP/>

<http://rice.plantbiology.msu.edu/>

<http://rice.genomics.org.cn/rice/index2.jsp>

Rizs genom projekt II.



MSU Rice Genome Annotation Project Release 7 (10/31/2011)

Features	TE-related genes/Loci	Non-TE-related genes/Loci
Gene Number	16 941	39 045
Average Gene Size (bp)	3 223	2 853
Average Exon/Gene	4.2	4.9
Average Exon Length (bp)	512	318
Average Intron/Gene	3.2	3.9
Average Intron Length (bp)	346	418

(TE-related: Transposable-element related genes and gene models)

Genes from MSU Rice Genome Annotation, Release 7

Functional Classification	Genes/Loci	Genes/Loci (%)	Gene Models	Gene Models (%)
Putative	22 896	40.9%	31 047	46.8%
Expressed	13 816	24.7%	15 686	23.6%
Conserved Hypothetical	133	0.2%	133	0.2%
Hypothetical	2 200	3.9%	2 200	3.3%
TE-related	16 941	30.3%	17 272	26.0%
Total	55 986	100%	66 338	100%

Rizs genom projekt III.

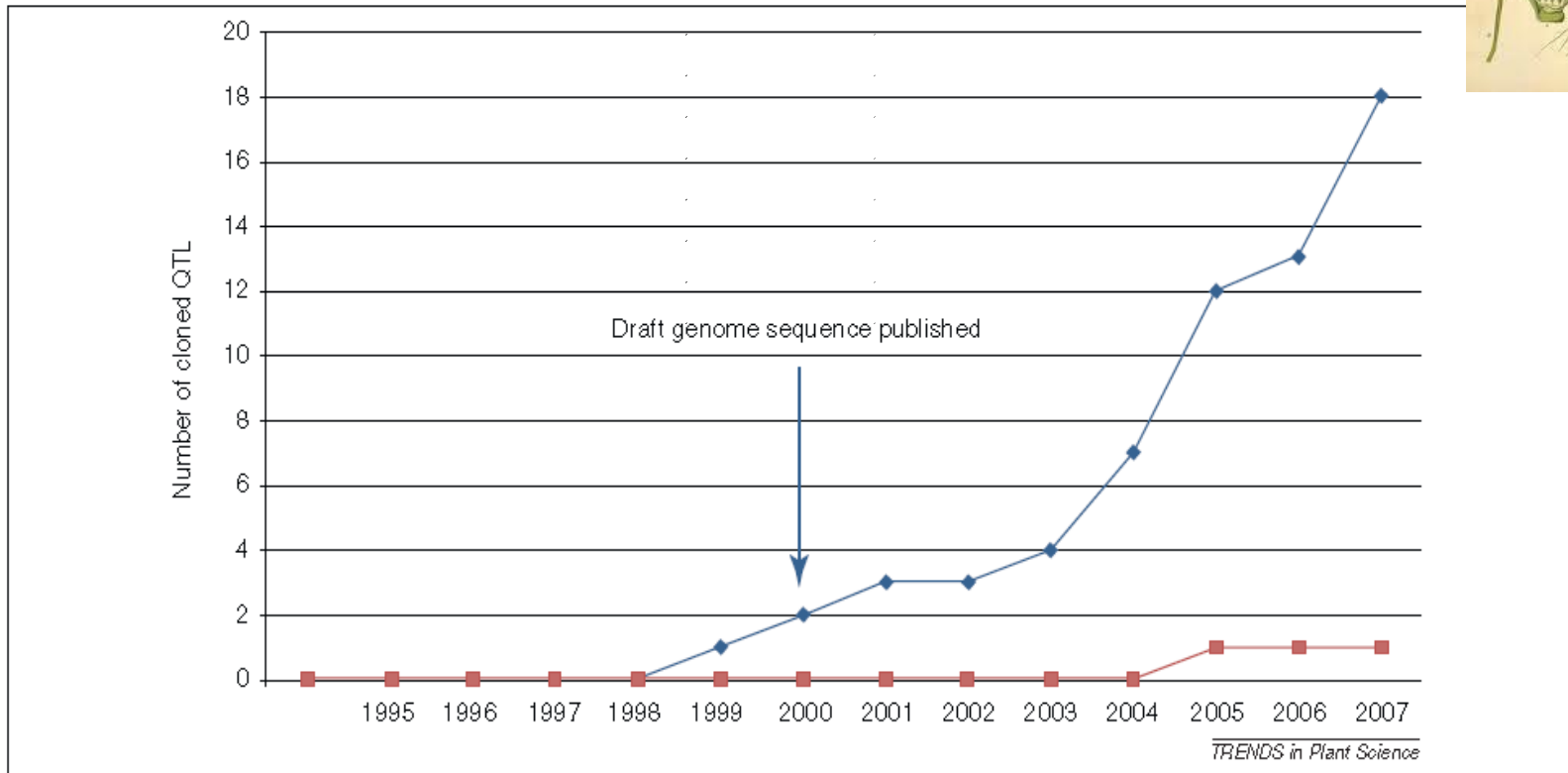
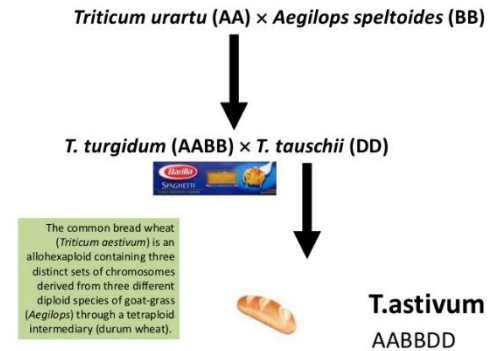


Figure 2. Number of QTL cloned in rice (blue) and wheat (red) since 1995. The blue arrow indicates the year in which the rice genome sequence became available and spurred the number of cloned genes and QTL (published source: NCBI and [42,61]). The Y axis represents the number of cloned QTL.

Wheat genome project

- wheat (*Triticum aestivum* L.) genome: $2n = 6x = 42$, AABBDD, estimated genome size 15,5 -17 Gbp
- Ancestor species: *Aegilops umbellata* ($2n=14$) és *A. tauschii* ($2n=14$, D genome) and *Triticum urartu* ($2n=14$, A genome)
- International Wheat Genome Sequencing Consortium (2005): 200 scientists from 73 research institutions in 20 countries
- result of 13 years of collaborative international research
- DNA sequence of bread wheat variety Chinese Spring ordered along the 21 wheat chromosomes resulting in the highest-quality genome sequence of wheat
- 107,891 high-confidence genes, 14,5 Gb



Reference: Consortium (IWGSC), T. I. W. G. S., Investigators, I. R. principal, Appels, R., Eversole, K., Feuillet, C., Keller, B., ... Uauy, C. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403), eaar7191. <https://doi.org/10.1126/science.aar7191>

Wheat genome project I.

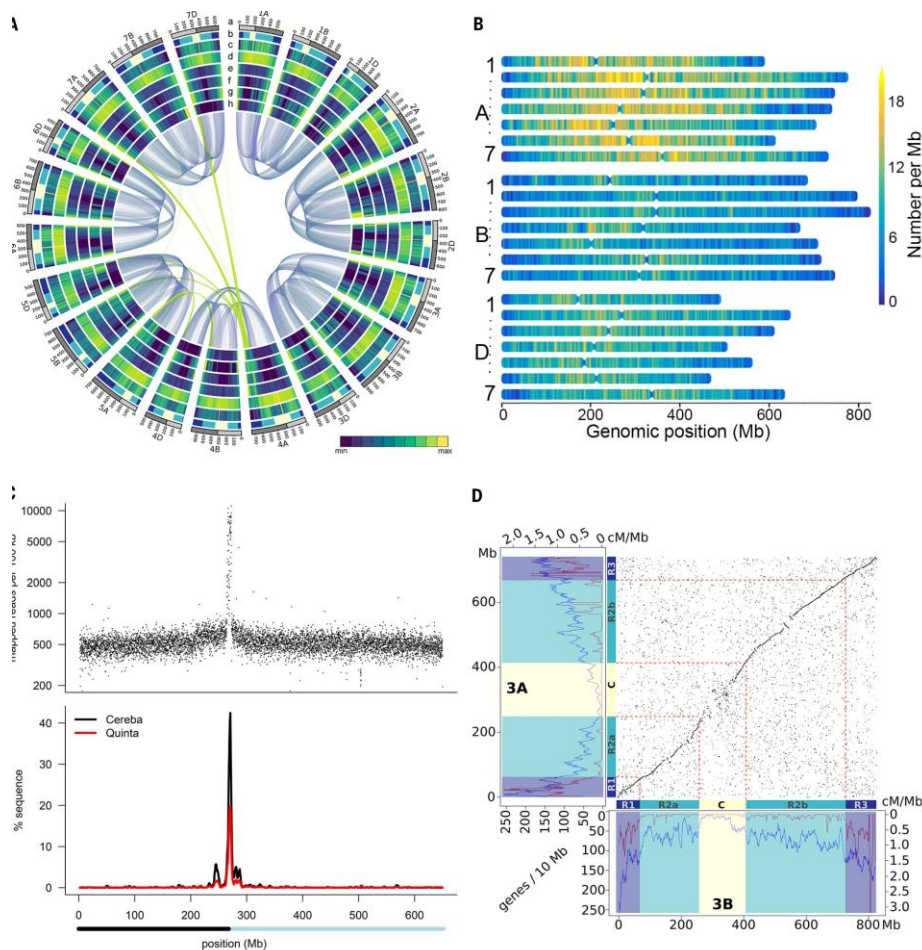
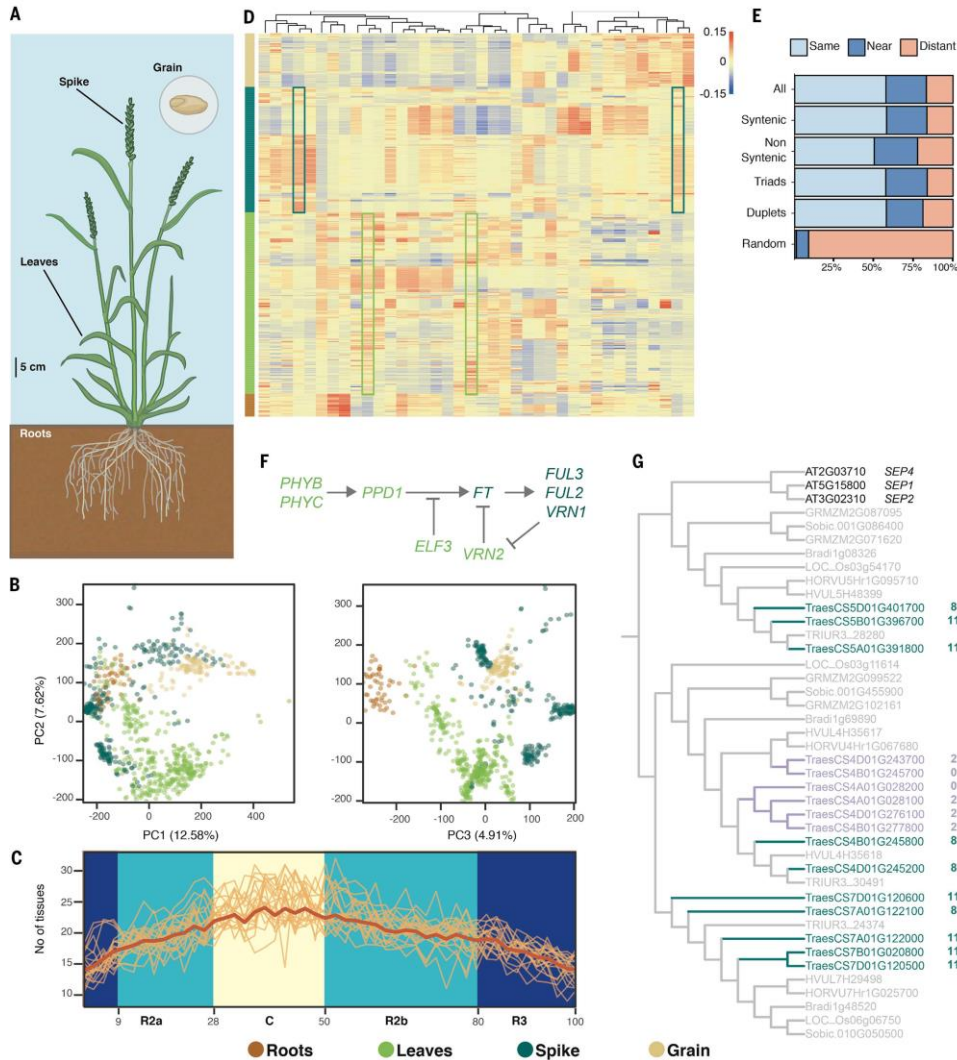


Table 2. Relative proportions of the major elements of the wheat genome. Proportions of TEs are given as the percentage of sequences assigned to each superfamily relative to genome size. Abbreviations in parentheses under the headings "Class 1" and "Class 2" indicate transposon types.

Major elements	Wheat subgenome			
	AA	BB	DD	Total
Assembled sequence assigned to chromosomes (Gb)	4.935	5.180	3.951	14.066
Size of TE-related sequences (Gb)	4.240	4.388	3.285	11.913
TEs (%)	85.9	84.7	83.1	84.7
Class 1				
LTR-retrotransposons				
Gypsy (RLG)	50.8	46.8	41.4	46.7
Copia (RLC)	17.4	16.2	16.3	16.7
Unclassified LTR-retrotransposons (RLX)	2.6	3.5	3.7	3.2
Non-LTR-retrotransposons				
Long interspersed nuclear elements (RIX)	0.81	0.96	0.93	0.90
Short interspersed nuclear elements (SIX)	0.01	0.01	0.01	0.01
Class 2				
DNA transposons				
CACTA (DTC)	12.8	15.5	19.0	15.5
Mutator (DTM)	0.30	0.38	0.48	0.38
Unclassified with terminal inverted repeats	0.21	0.20	0.22	0.21
Harbinger (DTH)	0.15	0.16	0.18	0.16
Mariner (DTT)	0.14	0.16	0.17	0.16
Unclassified class 2	0.05	0.08	0.05	0.06
hAT (DTA)	0.01	0.01	0.01	0.01
Helitrons (DHH)	0.0046	0.0044	0.0036	0.0042
Unclassified repeats	0.55	0.85	0.63	0.68
Coding DNA	0.89	0.89	1.11	0.95
Unannotated DNA	13.2	14.4	15.7	14.4
(Pre)-microRNAs	0.039	0.057	0.046	0.047
tRNAs	0.0056	0.0050	0.0068	0.0057

Reference: Consortium (IWGSC), T. I. W. G. S., Investigators, I. R. principal, Appels, R., Eversole, K., Feuillet, C., Keller, B., ... Uauy, C. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403), eaar7191. <https://doi.org/10.1126/science.aar7191>

Wheat genome project II.



B. Principal component (PC) analysis plots for similarity of overall transcription, with samples colored according to their high-level tissue of origin.

C. Chromosomal distribution of the average expression breadth [number of tissues in which genes are expressed (total number of tissues, n=32).

D. Heatmap illustrating the expression of a representative gene (eigengene) for the 38 co-expression modules defined by WGCNA.

E. Bar plot of module assignment (same, near, or distant) of homeologous triads and duplets in the WGCNA network.

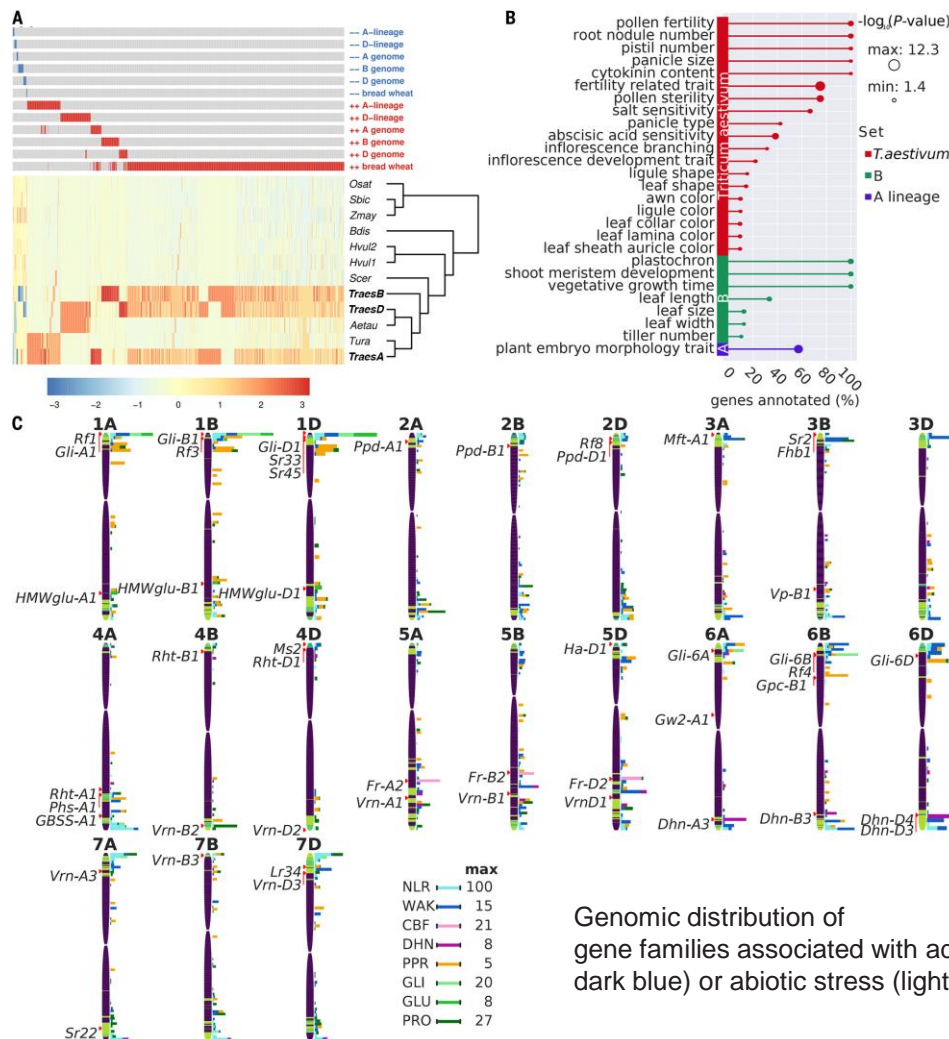
F. Simplified flowering pathway in polyploid wheat. Genes are colored according to their assignment to leaf (light green)– or spike (dark green)–correlated modules.

G. Excerpt from phylogenetic tree for MADS transcription factors, including known Arabidopsis flowering regulators SEP1, SEP2, and SEP4 (black).

Reference: Consortium (IWGSC), T. I. W. G. S., Investigators, I. R. principal, Appels, R., Eversole, K., Feuillet, C., Keller, B., ... Uauy, C. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403), eaar7191. <https://doi.org/10.1126/science.aar7191>

Wheat genome project III.

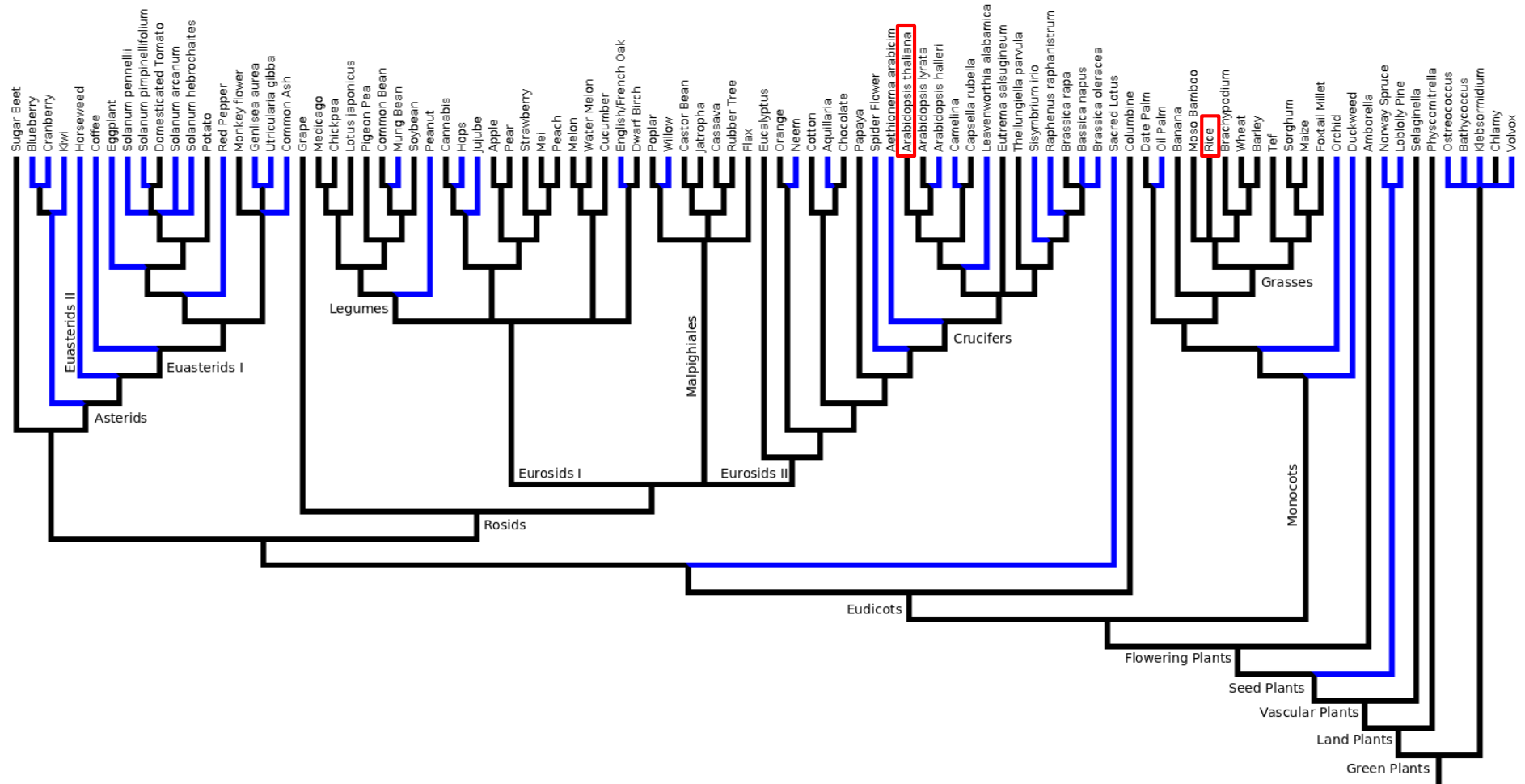
Heatmap of expanded and contracted gene families.



Genomic distribution of gene families associated with adaptation to biotic (light and dark blue) or abiotic stress (light and dark pink),

Reference: Consortium (IWGSC), T. I. W. G. S., Investigators, I. R. principal, Appels, R., Eversole, K., Feuillet, C., Keller, B., ... Uauy, C. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361(6403), eaar7191. <https://doi.org/10.1126/science.aar7191>

További növényi genomszekvenálási projektek I.



fekete: publikált szekvencia
 kék: a közelmúltban meghatározott genomszekvenca
 (törzsfák ágainak hossza nem arányos az elválás idejével)

További növényi genom szekvenálási projektek II.

- Algák - 18 faj; *Chlamydomonas* - egysejtű modell szervezet, amelynek fejlődése kb. 1 milliárd éve vált el a szárazföldi növényektől
flagellum - a növények és az állatok közös őseitől; 112 Mb, ~ 17700 gén (2007)
- *Physcomitrella patens*: moha (virág és szállítószövet hiánya); 480 Mb (2008),
36 000 gén modell
- csipkeharaszt (*Selaginella moellendorffii*): virág hiánya (2011)
egyik legkisebb szekvenált növényi genom (~110 Mb)
- nyitvatermők: lucfenyő (20 Gb - ~ 28 300 gén), cukorsüvegfenyő (20,8 Gb - 56 000 gén) mindkettő 2013, tömjénfenyő (20 Gb - 50 000 gén - 2014)
- zárvatermők: 82 faj
 - 67 valódi kétszikű: 50 Rosidae alosztály, 12 Asteridae, 5 egyéb
 - 22 egyszikű: 16 fűféle, 6 nem fűféle

További növényi genom szekvenálási projektek III.

- harangláb (*Aquilegia formosa*): Ranunculales rend; ~ 130 M éve vált el a rosid és asterid ágától, szekvencia meghatározás az utolsó fázisban

Asterids

- paradicsom: 950 Mb, 2n=24; 35000 gén



nyers genomszekv. (2012); *S. lycopersicum* és *S. pimpinellifolium* (legközelebbi vad rokonfaj;

0,6% SNP) - www.solgenomics.net

- burgonya: 844 Mbp; 4n=48; 2011, *S. commersonii* 2015, 838 Mbp

- paprika: 3,3 Gbp Nat. Genetics 2014 CM334, PNAS 2014 Zunla és vadfajok szekv.

- *Nicotiana benthamiana* (~ 3 Gbp, 2012), *N. sylvestris* - dohány (2,6 Gbp, 2013)

- *Petunia* 2011

Rosids

- szőlő: ~ 500 Mbp, 19 kromoszóma; 2007 - pinot noir, 2093 metakontig, a genom ~ 94,6%-a

Eurosids I

- uborka: 7 kromoszóma; „Chinese long 9930” ~ 350 Mbp; Gy14 inbred line 203 Mb

- nyárfa: 3. megszekvenált növényi genom (2006), 510 Mbp, 19 kromoszóma, nagyfokú heterozigotáság

- ricinus: 10 kromoszóma, ~ 320 Mbp, első genomszekvencia 2010

- cassava: manioka v. tapióka; jelenlegi verzió 416 Mbp (Sanger és Roche 454 szekv.; 2012)

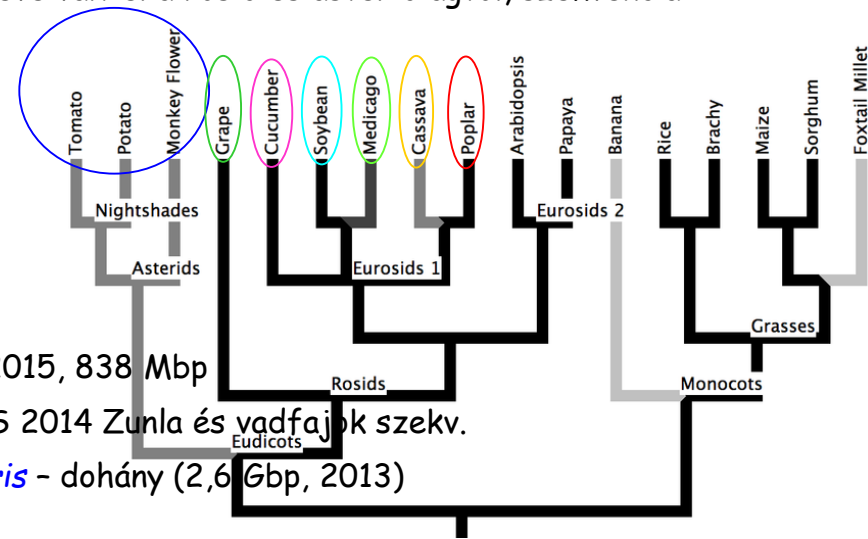
- vadkender: 400 Mb genom - Medicinal Genomics (Amsterdam) 2011; 131 000 Mb szekvencia - 327x lefedettség, nyers adatok publikálva - összeszerelésre várva

- alma: 17 kromoszóma, ~740 Mbp becsült genom; ~ 600 Mb (2010), 57 400 gén (genom duplikáció); körte

- őszibarack: 8 kromoszóma, 265 Mbp (2013), ~ 28 000 gén, mandula, cseresznye,

- citrusfélék (mandarin, narancs)

- földieper 240 Mbp (2010)



További növényi genom szekvenálási projektek IV.

- Rosids, Eurosids I

- *Medicago truncatula* : 8 kromoszóma, verzió 3.5 (2010), ~450 Mb becsült, szekv.: 240 Mb kromoszómákhoz kapcsolt, 17 Mb nem kapcs.

- *Lotus japonicus* : 6 kromoszóma, verzió 1.0 (2009) ~470 Mb becsült megszekvenált > 90%

- szója: 20 kromoszóma, ~1115 Mb becsült, 950 Mb ismert (2010) 46400 felt. kódoló gén (két teljes genom dupl. 59 és 13 M éve)

- pigeonpea (*Cajanus cajan*): Phaseoleae, $2n=2x=22$, 833 Mb félsivatagi mg-i kultúrnövény, Illumina nyers genomszekvencia (2012) 606 Mb (73%), 48 690 gén predikció nincs egész genomra kiterjedő duplikáció

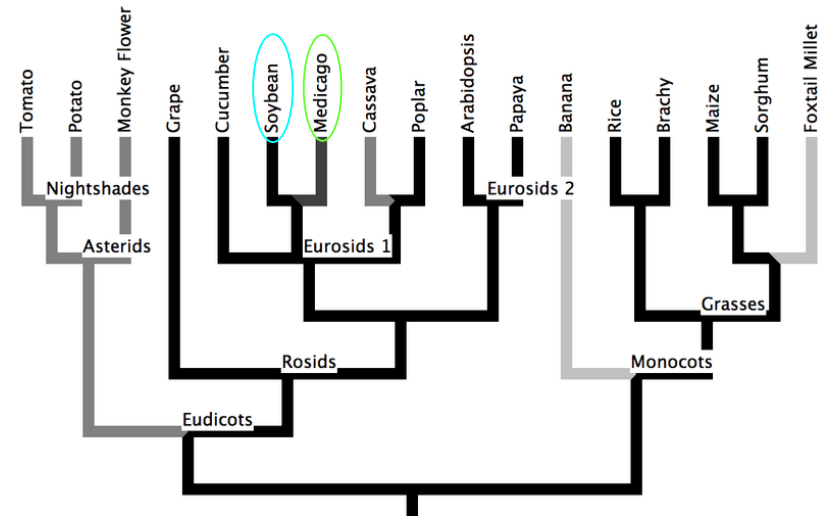
- *Cicer arietinum* 2013

- *Phaseolus vulgaris* 520 Mbp, 2013

- földimogyoró - 2016 Nat Genet

„A” genom *A. duranensis*

„B” genom *A. ipaensis*



További növényi genom szekvenálási projektek V.

Eurosid2 II - papaya : az első genetikailag módosított termesztett növény (papaya ringspot vírus rezisztencia) az Arabidopsishoz legközelebbi megszekvenált genom (2009), amely nem mutat nem teljes genom duplikációt; 9 kromoszóma, 372 Mb

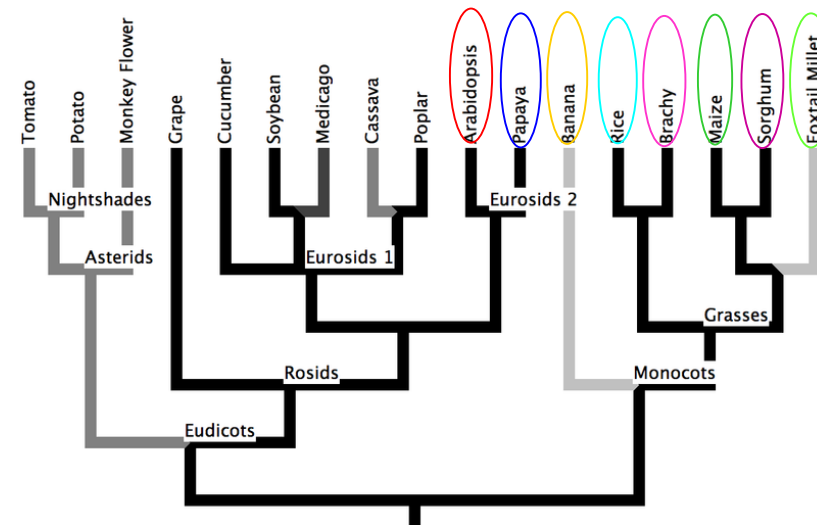
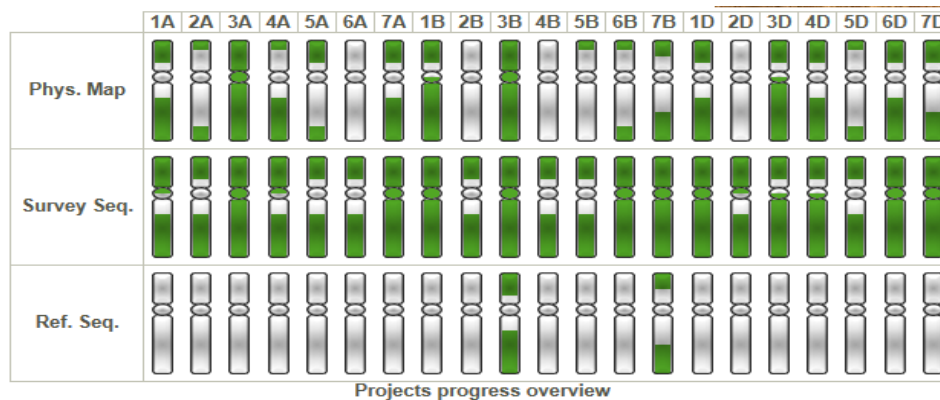
- *Arabidopsis thaliana*, *Arabidopsis lyrata* és egyéb *Arabidopsis* fajok - összehasonlító genomika

Egyszikűek

- datolyapálma: 658 Mbp (2011)
- banán: (2013), *Musa acuminata* - A-genom 523 Mbp, *M. balbisiana* - B-genom 438 Mbp

Fűfélék: - rizs - 5 féle genom köztük vadfajok is

- *Brachypodium*: búza, árpa zab és rozs modell növénye, energianövény?; a Földközi-tenger vidékén és a Közel-Keleten honos évelő fűfélé, $2n$, $4n$ és $6n$, 272 Mbp (Nature, 2010), 5 kromoszóma
- kukorica: 2500-3000 Mbp becsült, $2n=20$, 2300 Mbp (2009), >40 000 gén, teljes genom dupl. és 2xtranszpozon felszap.
- sorghum: kukorica közeli rokona (C_4 fotoszint.), ~730 Mbp szekv., 35 000 gén (2009)
- *Setaria italica* : rókafarkú köles v. olasz muhar, C_4 fotoszint. modellnövény (2007)
- búza: $2n = 6x = 42$, AABBDD, 15500-17000 Mb becsült; fizikai térképezési fázisban
Aegilops umbellata ($2n=14$) és *A. tauschii* ($2n=14$, D genom)
Triticum urartu ($2n=14$) 5 Gbp, búza A genom



További növényi genomszekvenálási projektek

- napraforgó: 2010-ben kezdődött (~ 3000Mb; 2n=34), egyik legnagyobb megszekvenált genom lesz
- gyapot: szekvenálás 2009-ben indult, 2013?
- szamóca: első verzió 2011
- kakaó: MARS finanszírozza (Roche 454) -2013
- Brassica fajok: *B. oleracea* (káposzta; CC), *B. rapa* (tarlórépa; AA), *B. napus* (repce; AACC) - Bayer CropScience - nem elérhető, BGI, Shenzhen és IVF, Beijing - Illumina
- olajpálma
- kávé: *Coffea arabica* - allotetraploid; ~1160 Mb, 2014 - N-methyltransferáz gének vizsgálata: bizonyítékok a teobromin/koffein bioszintézis konvergens evolúciójára
- Arabidopsis rokon fajok: *Boechera stricta*, *Thellungiella halophila* (sótűrő), *Capsella rubella*, *A. halleri*, *A. arenosa*, *Boechera divericarpa*
- *Eucalyptus grandis*
- *Zostera marina* (tengerifű), *Panicum virgatum* és *P. hallii* (díszfüvek)
- *Pinus taeda* (terpentin fenyő)
- *Miscanthus giganteus* (elefántfű)
- bab
- *Spirodela polyrhiza* - békalencse: 158 Mbp, 19 600 gén
- *Salix pupurea* (fűzféle)
- körte, cseresznye
- kristályvirág (ice plant) - CAM (crassulacean acid metabolism) fotoszintézis modell növénye

Célzott növényi genom projektek

- kukorica domesztikációja: kukorica - sorghum
- ivari kromoszóma vizsg.: papaya, nyárfa (chr XIX)
- termés íz és tápanyag-utak vizsg. : paradicsom, kakaó, stb.
- magfejlődés: szója, bab
- virágfejlődés: oroszlánszáj, harangláb, *Amborella trichopoda*, avokádó, *Magnolia* fajok, kaliforniai mák, *Zamia fischeri* (pálmafaj), *Welwitschia mirabilis*, spárga, gerbera, stb.
- CAM (Crassulacean acid metabolism) fotoszintézist végző növény *Kalanchoe* - 2013?
- kromatinszerkezet és génexpresszió vizsg.: petúnia
- vadfajok: rizs, szója, *Solanum* sp.
- energianövények: elefántfű és díszfüvek, nyárfa, fűzfa, olajpálma
- 1000 Plant Genomes Project meghirdetése (2008)
 - kb. 390 000 növényfaj létezik
 - ~ 125 000 fajról (2012) van DNS információ (95%-ban csak pár gén szekvencia)
 - Beijing Genomics Institute (BGI . Shenzhen, China) - Illumina szekvenálás
 - transzkriptom szekvenálás (teljes génkészlet szekvenálás <-> EST projektek)
 - biotechnológiai alkalmazás az előtérben - a mintákat az egyes fajok esetében a biotechnológiai szempontból fontos szövetekből veszik (másodlagos anyagcseretermékek, olajtartalom, orvosi vonatkozás, stb.)



Fontosabb genomszekvenálási projektek áttekintése

Table 1. Major crops for food, feed and non-food uses

	Crop name	Amount produced ^a	Order/Family ^b	Genome size (Mb)	Sequencing status
Food crops					
Cereals	Maize ^c	791.7	Poales/Poaceae	2500	Yes
	Rice (paddy)	659.5	Poales/Poaceae	389	Yes
	Wheat	605.9	Poales/Poaceae	17 000	Project
	Barley	133.4	Poales/Poaceae	5000	Project
	Sorghum	63.3	Poales/Poaceae	736	Yes
	Millet ^c	33.9	Poales/Poaceae	515	Project
	Oats	24.8	Poales/Poaceae	13 000	No
	Rye	14.7	Poales/Poaceae	8000	No
	Potatoes	309.3	Solanales/Solanaceae	840	Project
	Cassava ^c	214.5	Malpighiales/Euphorbiaceae	770	project
Vegetables and melon	Sweet potatoes	107.6	Solanales/Convolvulaceae	1500	No
	Tomatoes	129.9	Solanales/Solanaceae	950	Project
	Watermelons/melons	97.4	Cucurbitales/Cucurbitaceae	480	Project
	Cabbages and other brassicas	68.9	Brassicales/Brassicaceae	530	Project
	Onions, dry	66.0	Asparagales/Alliaceae	16 000	No
	Cucumbers and gherkins	44.2	Cucurbitales/Cucurbitaceae	367	Yes
	Eggplants (aubergines)	32.2	Solanales/Solanaceae	1100	No
	Carrots and turnips	27.2	Apiales/Apiaceae	473	No
	Lettuce and chicory	23.3	Asterales/Asteraceae	2500	No
	Pumpkins, squash and gourds	21.0	Cucurbitales/Cucurbitaceae	500	No
Oil crops	Cauliflowers and broccoli	17.7	Brassicales/Brassicaceae	880	No
	Garlic	15.8	Asparagales/Alliaceae	11 100	No
	Spinach	14.0	Caryophyllales/Amaranthaceae	990	No
	Soybeans ^c	220.5	Fabales/Fabaceae	1115	Yes
	Oil palm ^c	192.6	Arecales/Arecaceae	1800	Yes
	Seed cotton	73.6	Malvales/Malvaceae	2250	No
	Coconuts	61.5	Arecales/Arecaceae	2150	No
	Rapeseed ^c	50.6	Brassicales/Brassicaceae	600	Yes/project
	Groundnuts, with shell	37.1	Diverse order/families	800	No
	Sunflower seed ^c	26.8	Asterales/Asteraceae	3000	Project
Fruits	Olives	17.4	Lamiales/Oleaceae	1500	No
	Bananas/plantains	119.8	Zingiberales/Musaceae	600	Project
	Oranges/clementines/lemons	104.3	Sapindales/Rutaceae	367	Project
	Grapes	67.2	Vitales/Vitaceae	500	Yes
	Apples	66.0	Rosales/Rosaceae	750	Yes
	Mangoes, mangosteens and guavas	33.4	Diverse order/families	400	No
	Pineapples	20.9	Poales/Bromeliaceae	440	No
	Pears	20.6	Rosales/Rosaceae	500	No
	Peaches and nectarines	17.4	Rosales/Rosaceae	250	Yes
	Papayas	7.2	Brassicales/Caricaceae	372	Yes
Non-food crops					
Energy crops ^c	Poplar	NA	Malpighiales/Salicaceae	485	Yes
	Eucalyptus	NA	Myrtales/Myrtaceae	600	Project
	<i>Jatropha</i>	NA	Malpighiales/Euphorbiaceae	400	Yes
	Switchgrass	NA	Poales/Poaceae	480	Project
	Castor bean	NA	Malpighiales/Euphorbiaceae	400	Yes
	<i>Miscanthus</i>	NA	Poales/Poaceae	3300	Project
	Sugarcane	NA	Poales/Poaceae	2300	Project

Abbreviation: NA, not available.

^aCrops are sorted according to their relative importance in terms of production in 2007 (million tons yr⁻¹; source: <http://faostat.fao.org/>).

^bThe order and family names are from the Angiosperm Phylogeny website (<http://www.mobot.org/MOBOT/research/APWeb/welcome.html>).

^cSpecies used as both food and energy crops.

Fontosabb genom szekvenálási projektek áttekintése (2012)

Table 2. Overview of higher plant genome sequencing projects for which sequences are publicly available

Species and genotype	Genome size (Mb)	Ploidy	Sequencing strategy	Coverage	Refs
Finished genome sequences^a					
<i>Arabidopsis thaliana</i> (thale cress, cv. Columbia)	125	Diploid	BAC-by-BAC	~15x	[1]
<i>Oryza sativa</i> ssp. <i>japonica</i> (rice, cv. Nipponbare)	389	Diploid	BAC-by-BAC	10x	[38]
Improved high-quality draft genome sequences					
<i>Sorghum bicolor</i> (sorghum, cv. BTx623)	770	Diploid	WGS	8.5x	[72]
<i>Vitis vinifera</i> (grapevine, Pinot noir cv. PN40024)	487	Dihaploid	WGS	8.4x	[22]
<i>Brachypodium distachyon</i> (purple false brome, line Bd21)	300	Diploid	WGS	8x	[86]
<i>Glycine max</i> (soybean, cv. Williams 82)	1100	Diploid	WGS	8x	[6]
<i>Populus trichocarpa</i> (black cottonwood, poplar, cv. Nisqually-1)	485	Diploid	WGS	7.5x	[25]
<i>Zea mays</i> ssp. <i>mays</i> (maize, cv. B73) (gene space)	2600	Diploid	BAC-by-BAC	6x	[5] http://maizesequence.org/
<i>Cucumis sativus</i> (cucumber, IL 9930)	367	Diploid	WGS	72.2x (NGS)	[16]
High-quality draft genome sequences					
<i>Oryza sativa</i> ssp. <i>japonica</i> (rice, cv. Nipponbare)	433	Diploid	WGS	6x	[87]
<i>Oryza sativa</i> ssp. <i>indica</i> (rice, cv. 93-11)	466	Diploid	WGS	6.3	[40]
<i>Oryza sativa</i> ssp. <i>japonica</i> (rice, cv. Nipponbare)	399	Diploid	BAC-by-BAC	5x	[39]
<i>Vitis vinifera</i> (grapevine, Pinot noir cv. ENTAV 115)	505	Diploid	WGS and SBS	6.5/4.2x	[23]
<i>Carica papaya</i> (transgenic papaya, 'Scv.unUp')	372	Diploid	WGS	3x	[88]
<i>Prunus persica</i> (Peach, cv. Lovell)	220	Dihaploid	WGS	7.7x	http://www.rosaceae.org/peach/genome
<i>Malus x domestica</i> Borkh (Apple, cv. Golden delicious)	742	Dihaploid	WGS	16.9x	[89]
Standard draft genome sequences					
<i>Medicago truncatula</i> (barrel medic, cv. Jemalong A17)	500	Diploid	BAC-by-BAC	ND ^b	http://www.medicago.org/
<i>Zea mays</i> (popcorn, cv. Palomero Toluqueno)	2100	Diploid	WGS	3.2x	[90]
<i>Lotus japonicus</i> (trefoil, cv. Miyakojima MG-20)	472	Diploid	BAC-by-BAC	8.4x	http://www.kazusa.or.jp/lotus/
<i>Mimulus guttatus</i> (common monkeyflower)	430	Diploid	WGS	7x	http://www.phytozome.net/mimulus
<i>Ricinus communis</i> (castor bean, cv. Hale)	400	Diploid	WGS	4x	[91]
<i>Solanum tuberosum</i> (potato, DM1-3-516 R44)	840	Dihaploid	WGS	70x (NGS)	http://www.potatogenome.net
<i>Manihot esculenta</i> (cassava, cv AM560-2) (gene space 416 Mb)	770	Amphiploid	WGS	ND	http://www.phytozome.net/cassava
<i>Solanum lycopersicum</i> (common tomato, cv esculentum x pennellii)	950	Diploid	BAC-by-BAC		http://sgn.cornell.edu/about/tomato_project_overview.pl
			WGS	22 x	http://mips.helmholtz-muenchen.de/plant/tomato/index.jsp
<i>Arabidopsis lyrata</i> (rock cress)	230	Diploid	WGS	ND	http://www.phytozome.net/alyrata

Abbreviations: ND, no data; SBS, sequencing by synthesis.

^aThe sequences have been classified according to the criteria proposed by Chain *et al.* [8].

Fontosabb genom szekvenálási projektek áttekintése (2012)

Table 3. Overview of higher plant genome sequencing projects that are either underway or not yet publicly available

Species and genotype	Genome size (Mb)	Ploidy	Sequencing strategy	Refs
<i>Aquilegia formosa</i> (Western columbine)	400	Diploid	WGS	http://www.jgi.doe.gov/genome-projects/
<i>Brassica oleracea</i>	600	Diploid	WGS	http://www.genomesonline.org
<i>Brassica rapa</i> ssp. <i>pekinensis</i> (Chinese cabbage, cv. Chifu 401-42)	529	Diploid	BAC-by-BAC	http://www.brassica.info/resource/sequencing.php
<i>Brassica rapa</i> ssp. <i>pekinensis</i> (Chinese cabbage, cv. Chifu 401-42)	492	Diploid	WGS	http://www.intl-pag.org/18/abstracts/W14_PAGXVIII_104.html
<i>Brassica rapa</i> (B3)	530	Diploid	WGS	http://www.jgi.doe.gov/genome-projects/pages/projects.jsf?kingdom5Plant
<i>Brassica napus</i> (rapeseed) (<i>oleracea/rapa</i>)	1100	Tetraploid	WGS	http://www.bayercropscience.com/bcsweb/cropprotection.nsf/id/EN_20091009?open&I5EN&ccm5500020
<i>Capsella rubella</i> (pink shepherds purse)	250	Diploid	WGS	http://www.jgi.doe.gov/genome-projects/
<i>Citrus sinensis</i> (sweet orange, cv. Ridge pineapple)	382	Diploid	WGS	http://www.citrusgenome.ucr.edu/
<i>Elaeis guineensis</i> (oil palm, cv. <i>tenera</i> × <i>dura</i>)	1800	Diploid	BAC pools and WGS	http://www.syntheticgenomics.com/media/press/52108.html
<i>Eucalyptus grandis</i> (BRASUZ1)	600	Diploid	WGS	http://www.jgi.doe.gov/genome-projects/
<i>Gossypium raimondii</i> (cotton)	880	Diploid	WGS	http://www.jgi.doe.gov/genome-projects/
<i>Jatropha curcas</i> (synthetic genomics)	400	Diploid	Unpublished	http://www.syntheticgenomics.com/media/press/52009.html
<i>Nicotiana tabacum</i> (tobacco, cv. Hicks broadleaf)	4500	Tetraploid	Methyl filtration	http://www.tobaccogenome.org/
<i>Setaria italica</i> (foxtail millet, Yugu1)	515	Diploid	WGS	http://www.jgi.doe.gov/genome-projects/
<i>Solanum tuberosum</i> (potato, RH89-039-16)	840	Diploid	BAC-by-BAC	http://www.potatogenome.net [92]
<i>Vigna unguiculata</i> (cowpea)	620	Diploid	Methyl filtration	[93]
<i>Zea mays</i> (maize, cv. Mo17)	2100	Diploid	WGS	http://www.maizegdb.org/sequencing_project.php

Fontosabb genomszekvenálási projektek áttekintése (2012)

Table 3 List of published plant genome sequences

Plant species (common name)	Family	Approach	References
<i>Arabidopsis lyrata</i>	Brassicaceae	WGS	Hu <i>et al.</i> (2011)
<i>Arabidopsis thaliana</i> (mouse ear cress)	Brassicaceae	BAC-by-BAC	Arabidopsis Genome Initiative, (2000)
<i>Brachypodium distachyon</i>	Poaceae	WGS	International Brachypodium Initiative, (2010)
<i>Brassica rapa</i> (Chinese cabbage)	Brassicaceae	WGS	<i>Brassica rapa</i> Genome Sequencing Project Consortium, (2011)
<i>Cajanus cajan</i> (pigeonpea)	Fabaceae	WGS	Varshney <i>et al.</i> (2011)
<i>Carica papaya</i> (papaya)	Caricaceae	WGS	Ming <i>et al.</i> (2008)
<i>Cucumis sativus</i> (cucumber)	Cucurbitaceae	WGS	Huang <i>et al.</i> (2009a,b) and Woycicki <i>et al.</i> (2011)
<i>Fragaria vesca</i> (woodland strawberry)	Roseaceae	WGS	Shulaev <i>et al.</i> (2011)
<i>Glycine max</i> (soybean)	Fabaceae	WGS	Schmutz <i>et al.</i> (2010)
<i>Glycine soja</i>	Fabaceae	WGS	Kim <i>et al.</i> (2010)
<i>Medicago truncatula</i> (barrel medic)	Fabaceae	WGS and BAC-by-BAC	Young <i>et al.</i> (2011)
<i>Malus × domestica</i> (apple)	Roseaceae	WGS	Velasco <i>et al.</i> (2010)
<i>Oryza sativa</i> (rice, indica)	Poaceae	WGS	Yu <i>et al.</i> (2002)
<i>Oryza sativa</i> (rice, japonica)	Poaceae	WGS and BAC-by-BAC	Goff <i>et al.</i> (2002); International Rice Genome Sequencing Project (2005)
<i>Phoenix dactylifera</i> (date palm)	Arecaceae	WGS	Al-Dous <i>et al.</i> (2011)
<i>Physcomitrella patens</i>	Funariaceae (Bryophyta)	WGS	Rensing <i>et al.</i> (2008)
<i>Populus trichocarpa</i> (poplar)	Salicaceae	WGS	Tuskan <i>et al.</i> (2006)
<i>Ricinus communis</i> (castor bean)	Euphorbiaceae	WGS	Chan <i>et al.</i> (2010)
<i>Selaginella moellendorffii</i>	Selaginellaceae (Lycopodiophyta)	WGS	Banks <i>et al.</i> (2011)
<i>Solanum tuberosum</i> (potato)	Solanaceae	WGS	Potato Genome Sequencing Consortium, (2011)
<i>Sorghum bicolor</i> (sorghum)	Poaceae	WGS	Paterson <i>et al.</i> (2009)
<i>Thellungiella parvula</i>	Brassicaceae	WGS	Dassanayake <i>et al.</i> (2011)
<i>Theobroma cacao</i> (cacao)	Malvaceae	WGS	Argout <i>et al.</i> (2011)
<i>Vitis vinifera</i> (grapevine)	Vitaceae	WGS	Jaillon <i>et al.</i> (2007)
<i>Zea mays</i> (maize)	Poaceae	BAC-by-BAC	Schnable <i>et al.</i> (2009)

Növényi genomprogramok weboldalai

<http://arabidopsis.org/>

<http://rice.genomics.org.cn/rice/index2.jsp>

<http://compbio.dfci.harvard.edu/tgi/plant.html>

http://solgenomics.net/genomes/Solanum_lycopersicum/index.pl

http://www.potatogenome.net/index.php/Main_Page

<http://maizesequence.org>

<http://www.medicago.org/genome/>

<http://www.brassica.info/resource/sequencing.php>

<http://www.wheatgenome.org/>

http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html

www.ncbi.nlm.nih.gov/

www.jgi.doe.gov/

<http://www.phytozome.net/>

www.sanger.ac.uk/

www.expasy.ch/

www.arabidopsis.org/

www.kazusa.or.jp/

A genetikai térképezés és a genomszekvencia néhány alkalmazása

- a genetikai térképek biztos alapot jelentenek genomikai projektnél -
genomszerveződés vizsgálata (pl. genomszekvenálás, fizikai térképezés,
genetikai és citológiai térkép összevetése, stb.)
- összehasonlító genetikai térképezés (genom evolúció)
- gyakorlati szempontból fontos génekhez vagy lókuszekhez kapcsolt
markerek azonosítása
- génklónozás

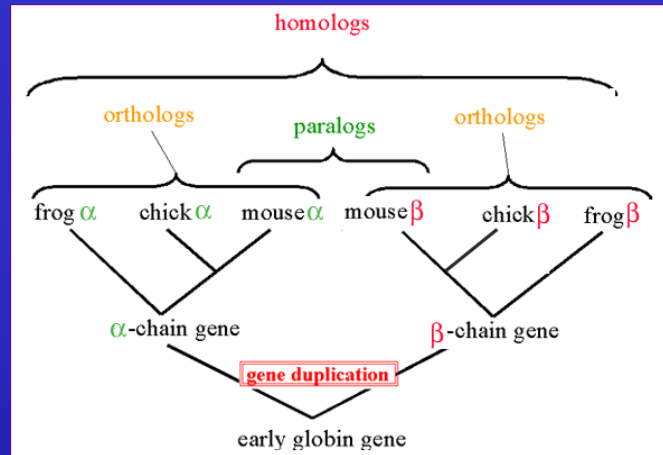
Összehasonlító genetikai térképezés I. (alapfogalmak)

- Kapcsolt genetikai markerek (kapcsoltság) vizsgálata különböző fajokban vagy rendszertani kategóriák között
- Homológia = hasonlóság; *közös evolúciós eredet* (madár és denevér szárny; hasonló szerkezet és pozíció)
- Analógia = hasonlóság, *közös evolúciós eredet nélkül* (pl. légy és madár szárny)
- Homológia - ortológia és paralógia
 - két gén ortológ, ha két különböző fajban találhatóak, és egy közös ősgénből származnak, mely a két faj közös ősében jelen volt. Ugyanazt a funkciót szolgálják, a két fajban.
 - két gén paralóg, ha ugyanabban az organizmusban találhatóak, és egy közös ősgénből génduplikáció és azt követő divergens evolúció útján alakultak ki. Többnyire különböző, de egymással összefüggésben lévő funkciójuk van.



Összehasonlító genetikai térképezés II. (alapfogalmak)

ortológ-paralóg



Synteny („egybeesés”) = ortológ lókuszok két különböző faj azonos kromoszómáján helyezkednek el

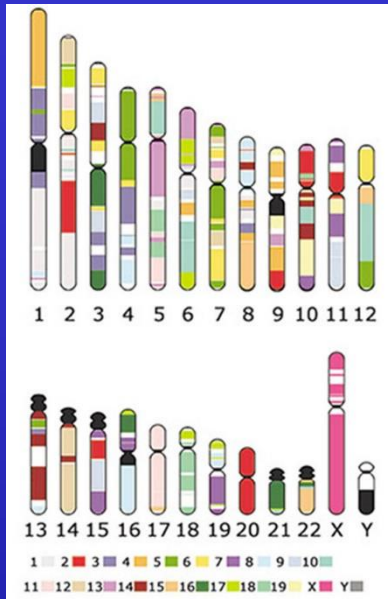
Colinear = a lókuszok két különböző faj azonos kromoszómáján konzervált sorrendben helyezkednek el

(nem helyesen manapság a kettőt azonos értelemben használják!)

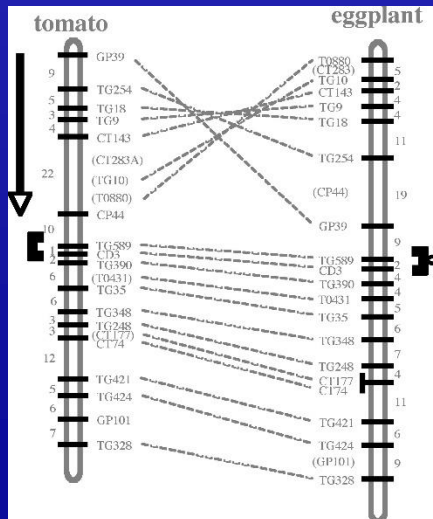
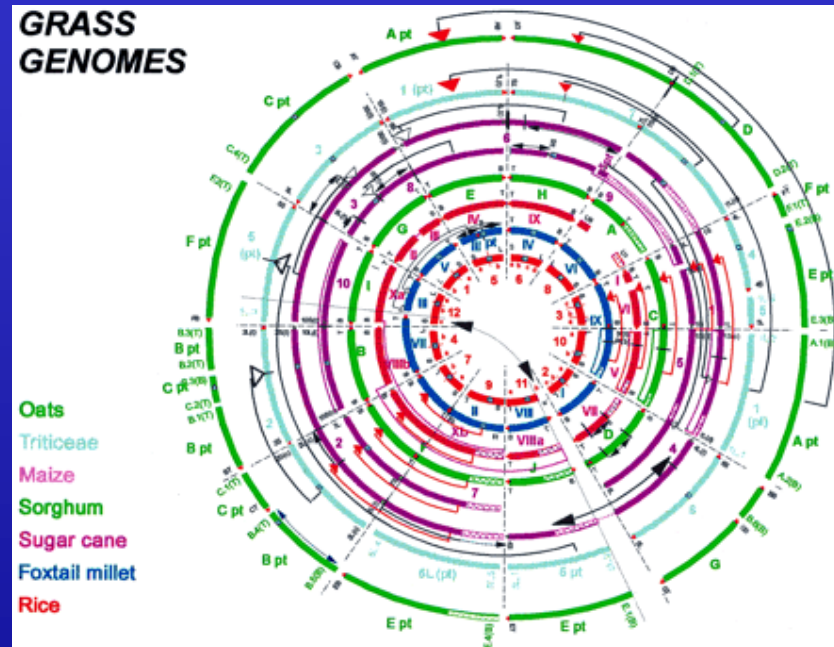
Macrosynteny = markerek sorrendjének konzerváltsága kromoszómaszinten

Microsynteny = markerek sorrendjének konzerváltsága kisebb genomi szakaszon

Összehasonlító genetikai térképezés III.

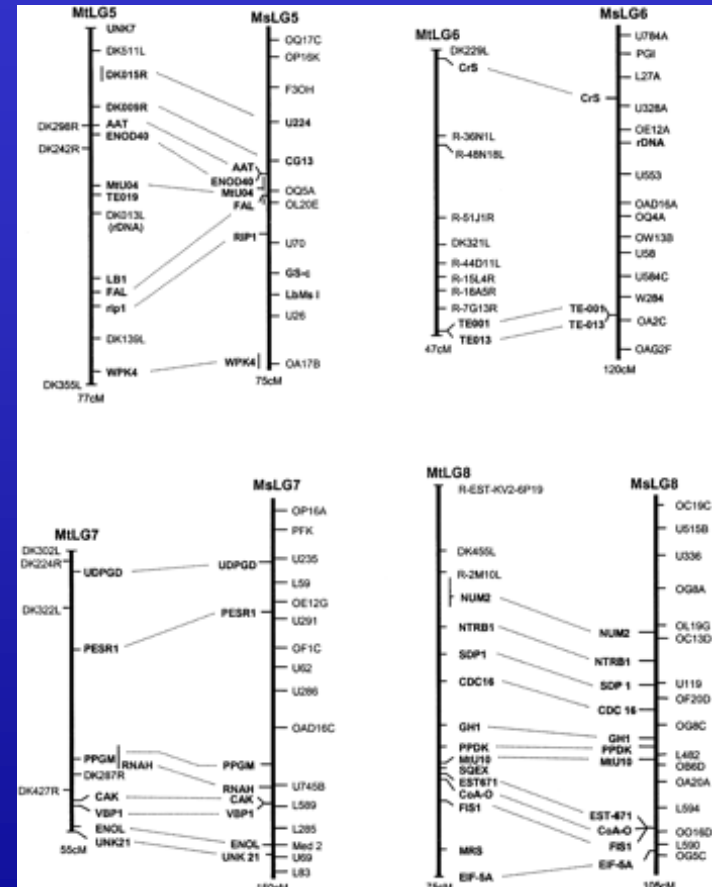
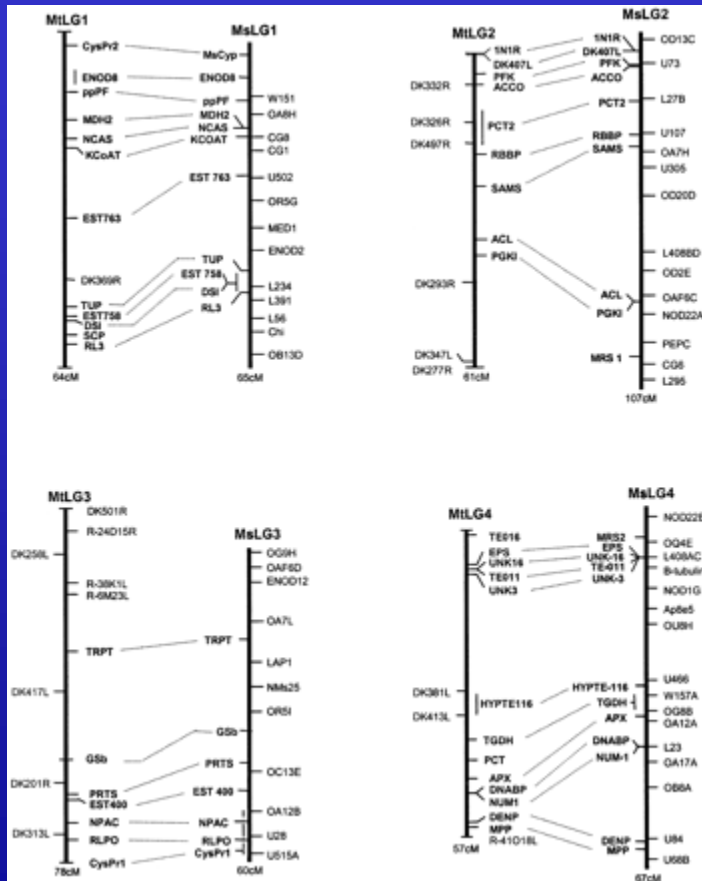


Macrosynteny humán és egér genom között

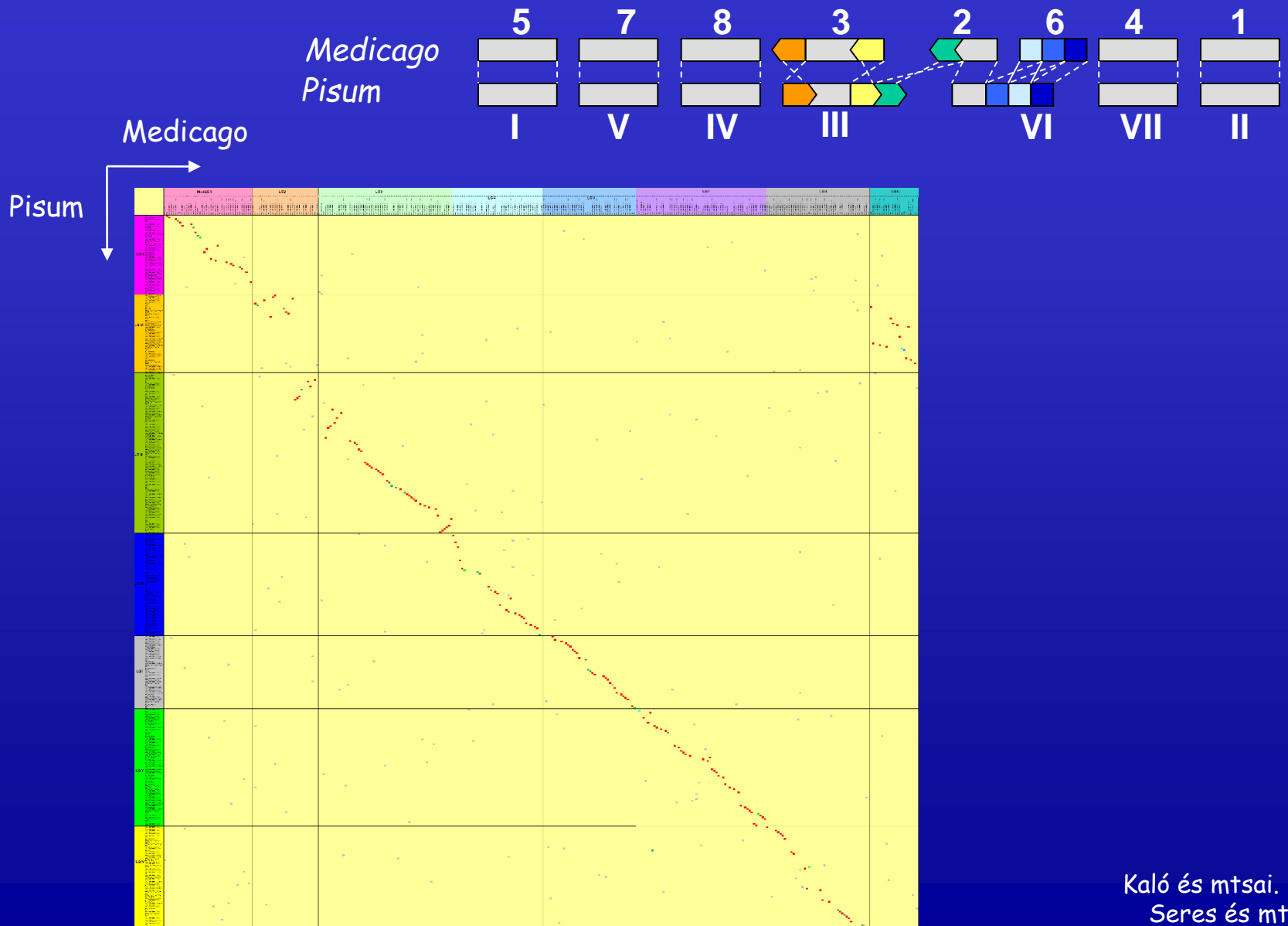


Inverzió egy paradicsom és padlizsán kromoszóma esetében

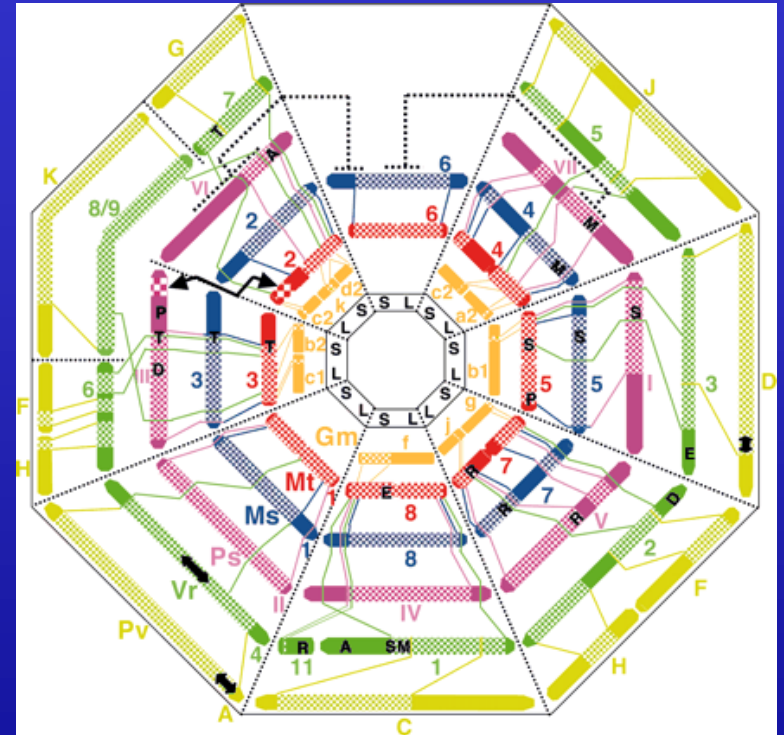
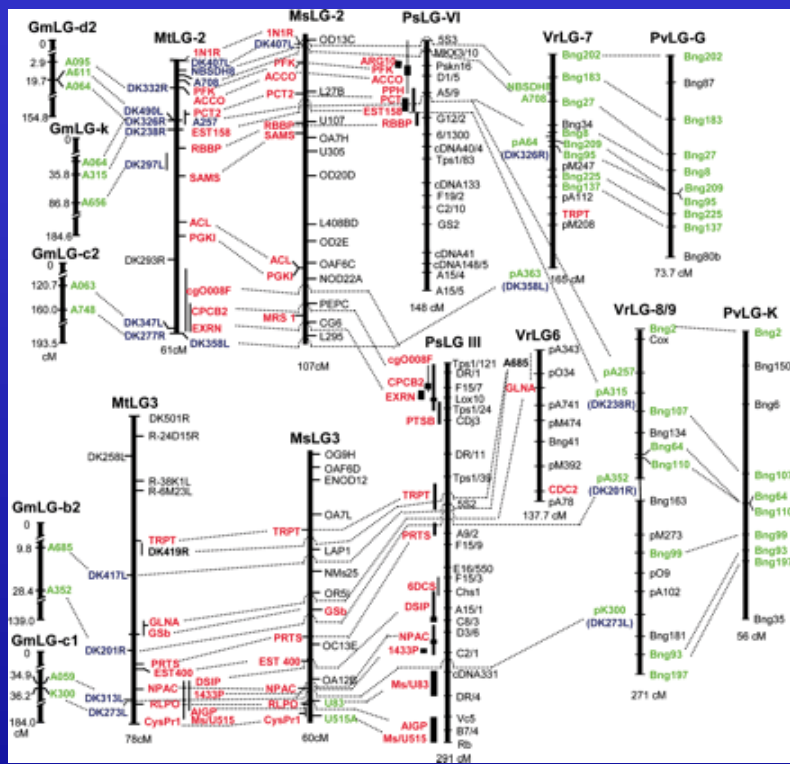
Összehasonlító genetikai térképezés IV. (*Medicago sativa* és *M. truncatula*)



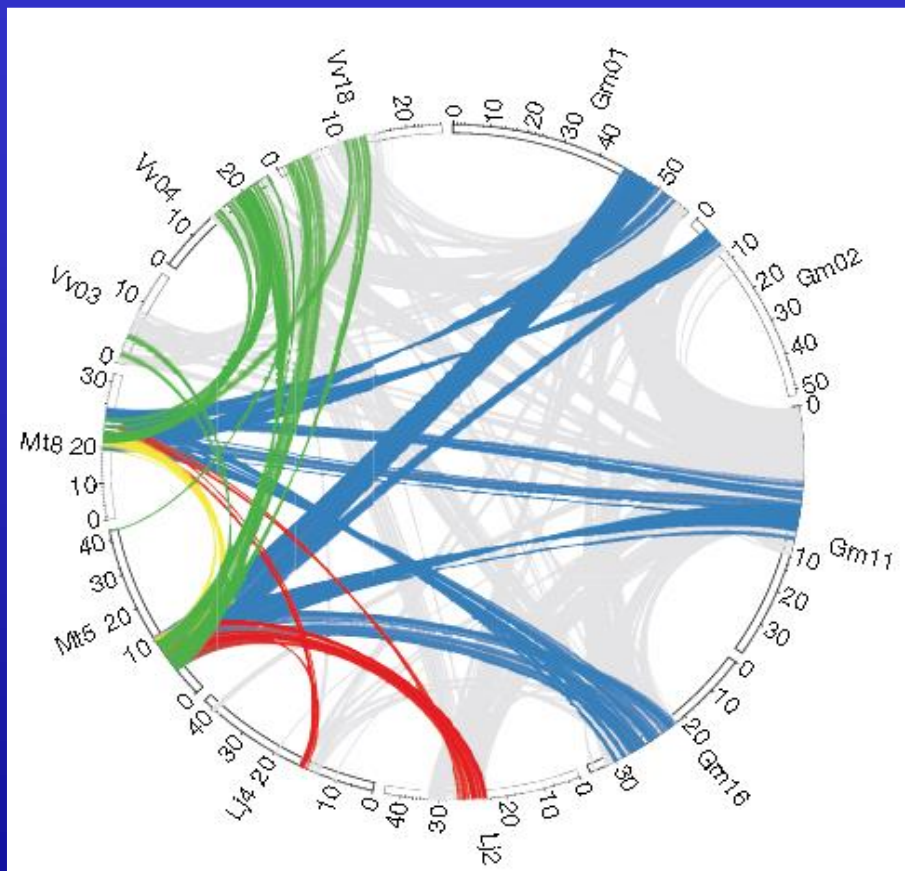
Összehasonlító genetikai térképezés V. (Medicago - borsó)



Összehasonlító genetikai térképezés VI. (pillangósok)



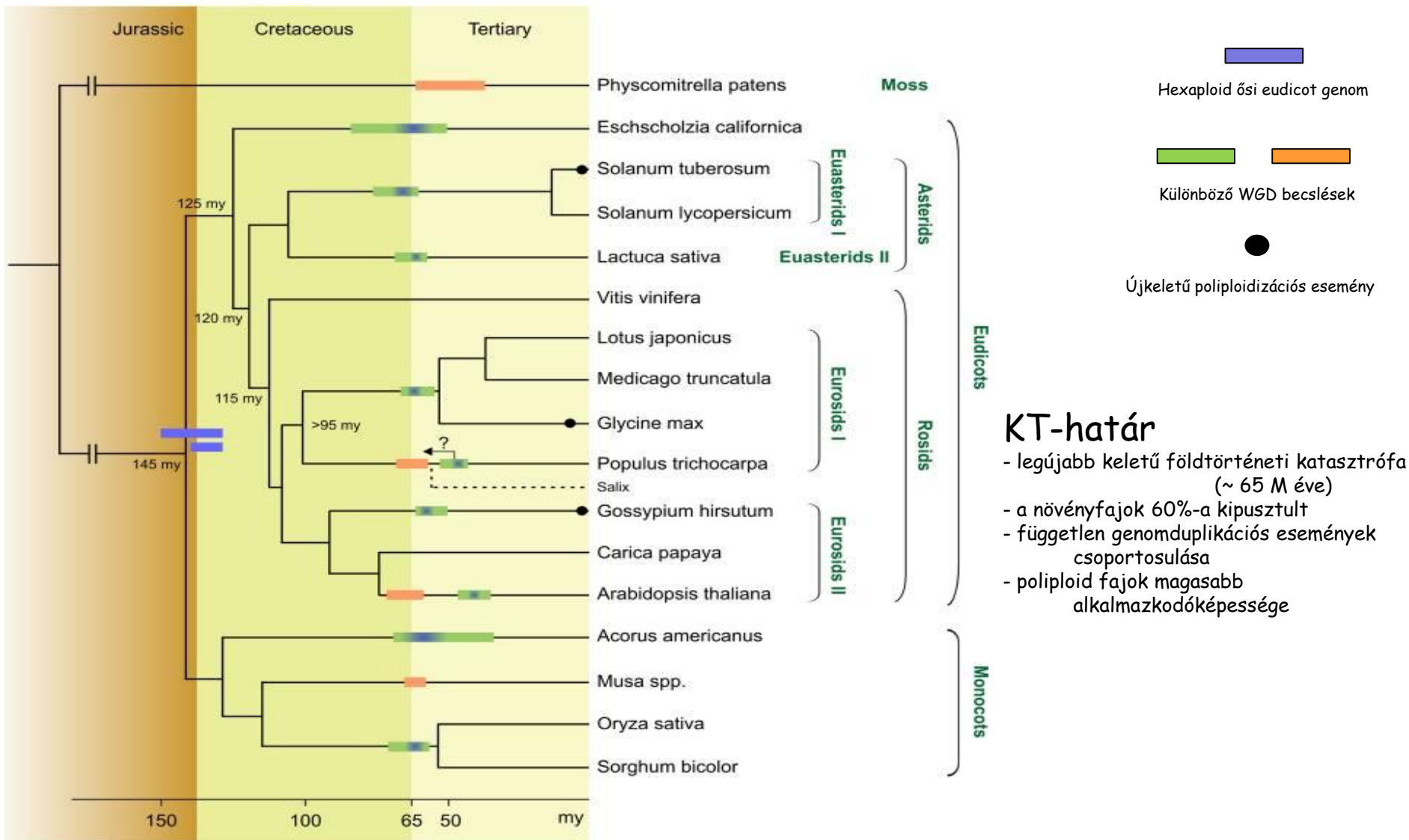
Genom összehasonlítás teljes genomszekvencia felhasználásával



Medicago truncatula chr5 és chr8
(Young et al. 2011, Nature)

sárga - Mt5 és Mt8 szintenikus régió
piros - *Lotus japonicus*
kék - szója
zöld - szőlő

Poliploidizáció a zárvatermők evolúciója során

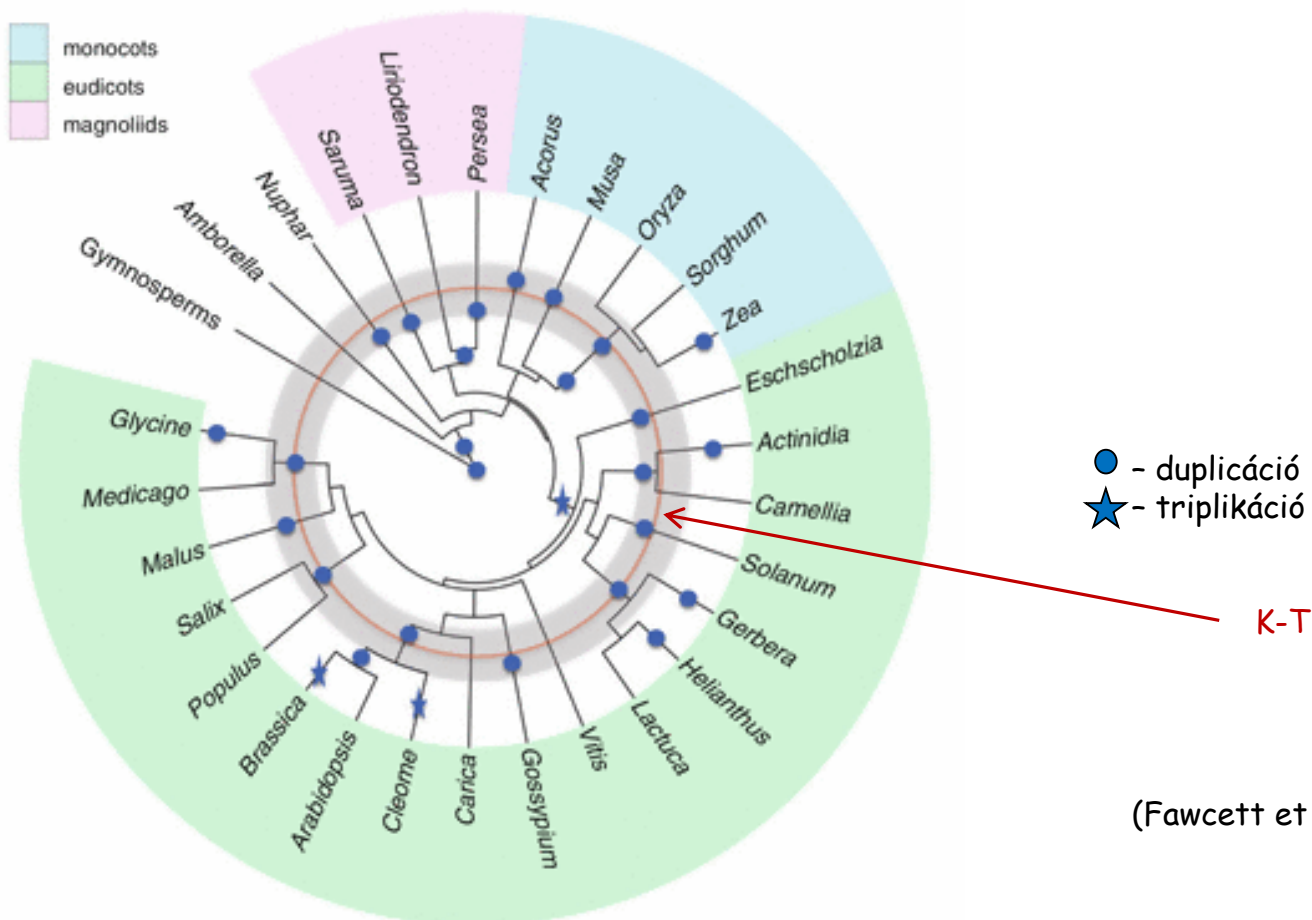


KT-határ

- legújabb keletű földtörténeti katasztrófa (~ 65 M éve)
- a növényfajok 60%-a kipusztult
- független genomduplikációs események csoportosulása
- poliploid fajok magasabb alkalmazkodóképessége

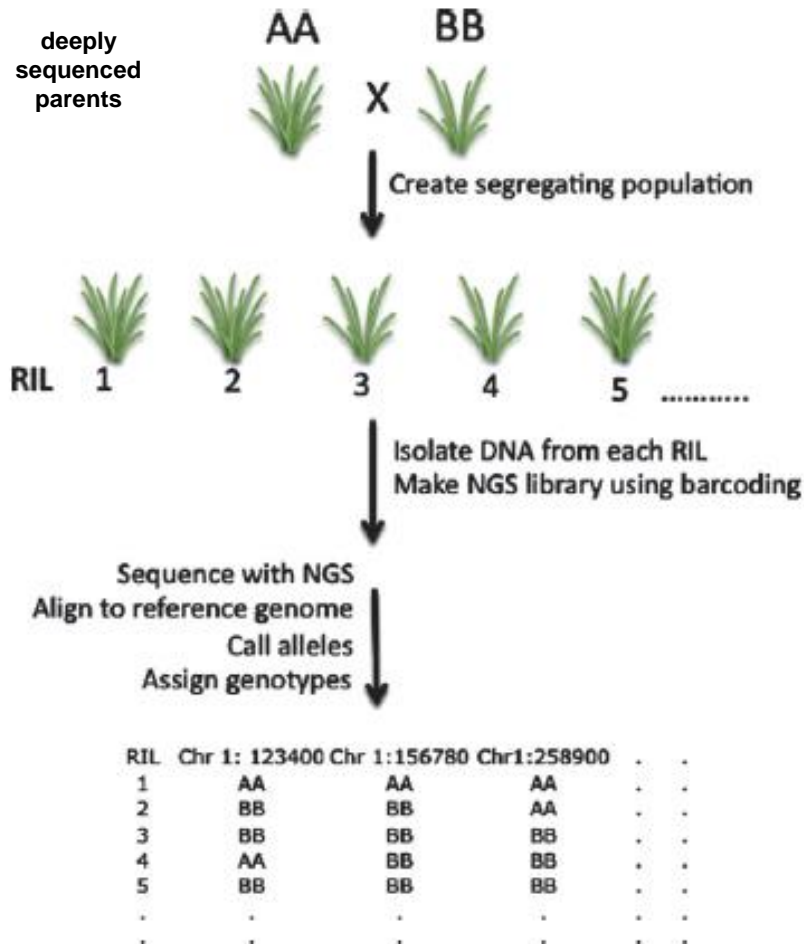
Poliploidizáció a növények evolúciója során

- a fajok 35%-a neopoliploid
- a legtöbb faj az evolúció során többször átment poliploidizációs eseményen
- életképes aneuploid variánsok
(allopoliploidizációt követően gyakran - hexaploid búza)
stabil búza vonalak hiányzó homeológ kromoszómákkal

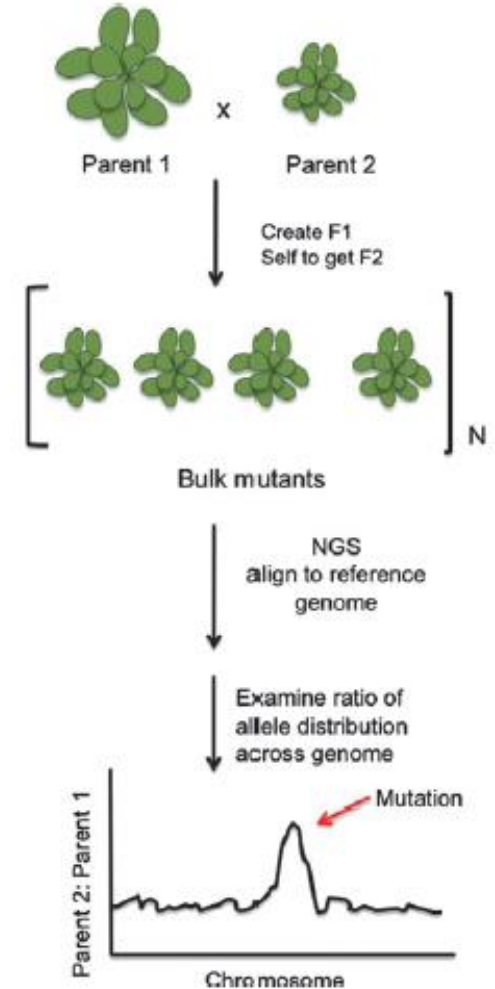


NGS alkalmazása a növényi biotechnológiában

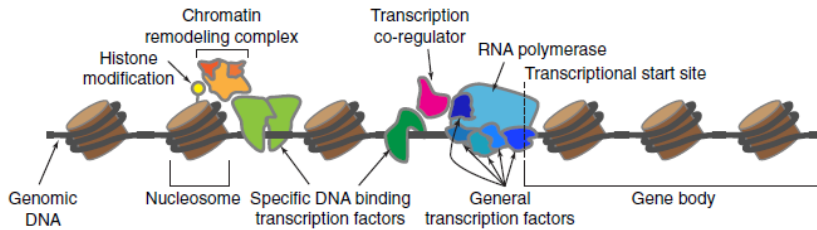
Genetikai térképezés szekvenálással



Mutációt okozó allél azonosítása egy genetikai screenben

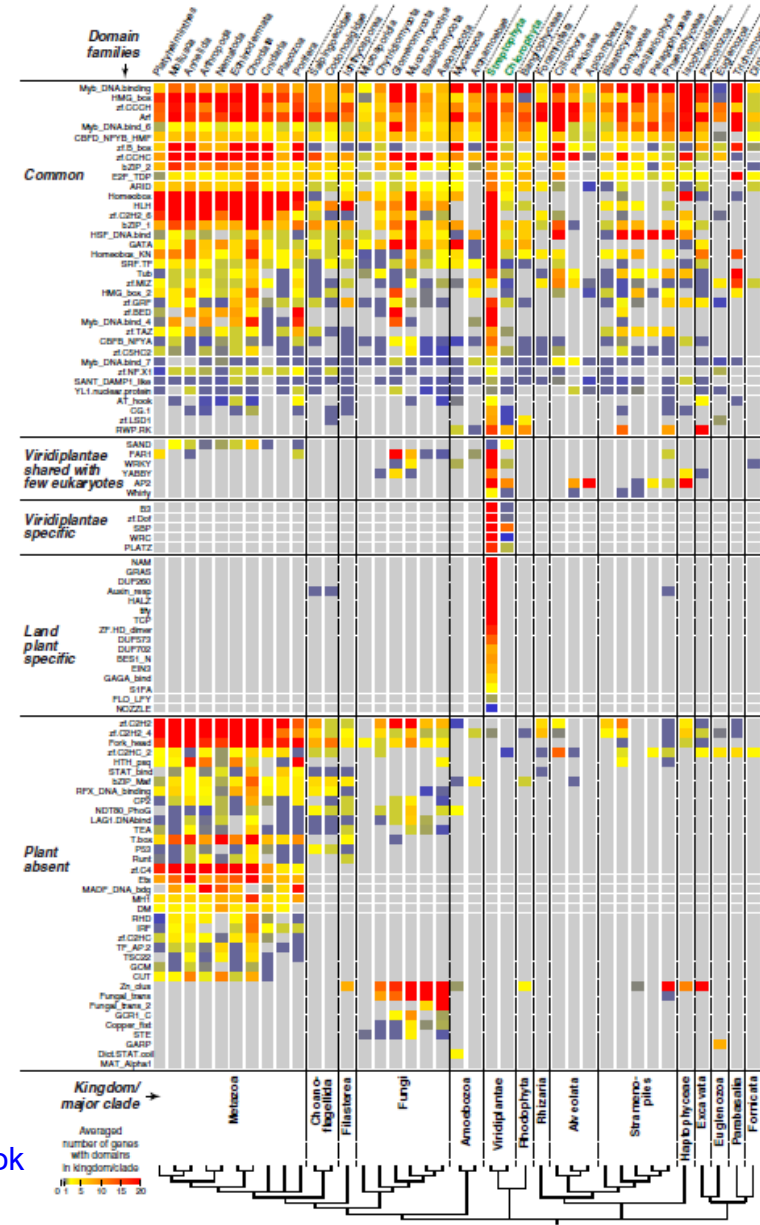


Transzkripciós faktorok (TF) genomszintű vizsgálata I.



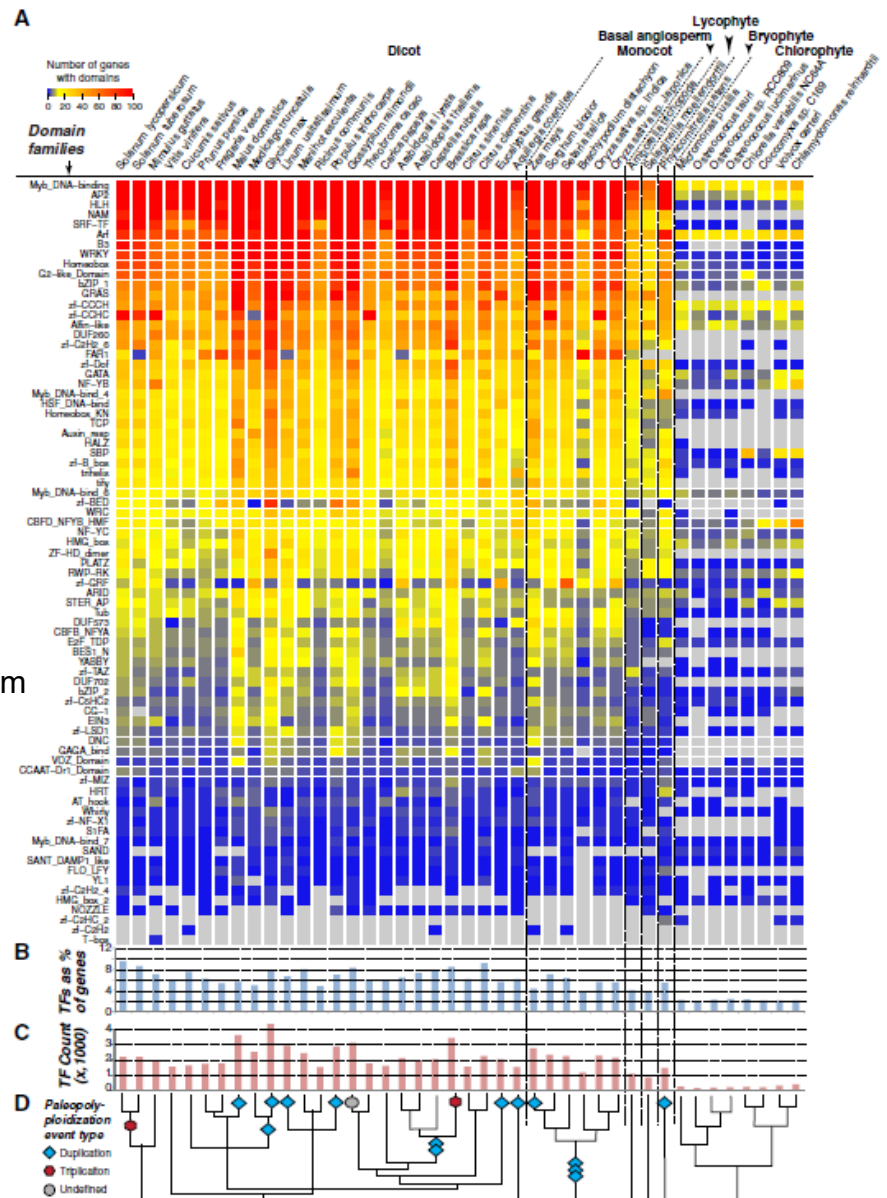
- TF – DNS-kötő domén (DBD); gén expresszió szabályozása
- az élőlények fejlődési, morfológia, fiziológiai funkcióit szabályozzák
- növények - helyhez kötött életmódot folytatnak
- 1500-2000 gén/növényi genom (151 *Ostreococcus tauri*; > 4000 *Glycine m.*)
- > 50 géncsalád
- génduplikáció - alkalmazkodás, új funkciók, diverzitás
- elegendő mennyiségű és minőségű genomszekvencia
- endophyta szennyezés; ha a csoport fajainak több mint 20%-ban előfordul
- 99 TF DBD géncsalád
- a növényi TF-ok közül
 - 36 közös a legtöbb eukarióta élőlényvel
 - 6 közös a Viridiplantae és egy-két nagyobb eukarióta csoport között
 - 5 specifikus a Viridiplantae fajokra
 - 16 specifikus a szárazföldi növényekre

növény-specifikus TF családok



Transzkripciós faktorok (TF) genom szintű vizsgálata II.

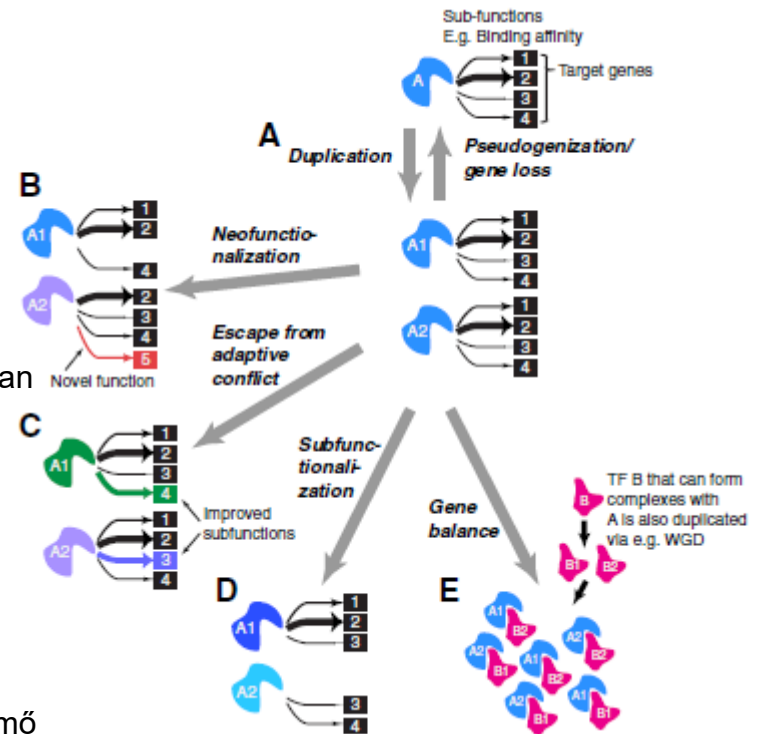
- paleopolyploidizáció, WGD → több gén, több TF
- szárazföldi növények esetén jelentő növekedés génszámban
- NAC TF család elterjedtebb a kétszikűekben, míg WRKY család kiterjedtebb egyszikűekben
- génduplikáció - WGD, tandem, transzpozon közvetített
 - Arabidopsis esetében 2x WGD eredetű, mint a tandem duplikált génszám
- génduplikáció
 - elvesztés (teljes deléció vagy pszeudogénné válás)
 - megmaradás – „biztonsági tartalék”, új funkció



Transzkripciós faktorok (TF) genomszintű vizsgálata III.

Duplikálódott gének sorsa

- génduplikáció
 - elvesztés (teljes deléció vagy pszeudogénné válás)
 - megmaradás
 - „biztonsági tartalék”
 - új funkció (virágszerkezet evolúciója)
 - eltérő alkalmazkodás
 - alfunkció kialakulása (eltérő mértékű funkcióvesztés/nyerés)
 - heteromer komplex, melynek tagjai duplikálódnak - „gén szimmetria” a sztöchiometirai egyensúly megtartására (gyakran nagyfokú expressziós hasonlóság; Arabidopsis gyökérben kifejeződő TF-ok)
- új tulajdonságok kialakulása/evolúciója, domesztikáció
 - MADS-box TF – virágstruktúra evolúciója - **ABCE** modell (BCE MADS-box)
 - A+E csészelevél, A+B+E szíromlevél, B+C+E porzó, C+E termő
 - domesztikáció
 - *teosinte branched1 (tb1)* TCP TF, erősebb expresszió kukoricában egy TE inszerció következtében (enhancer) nőivarú virágzat fejlődését segíti elő
 - a mag kipergését elindító TF gén deléciója (sorghum, rizs, stb.)



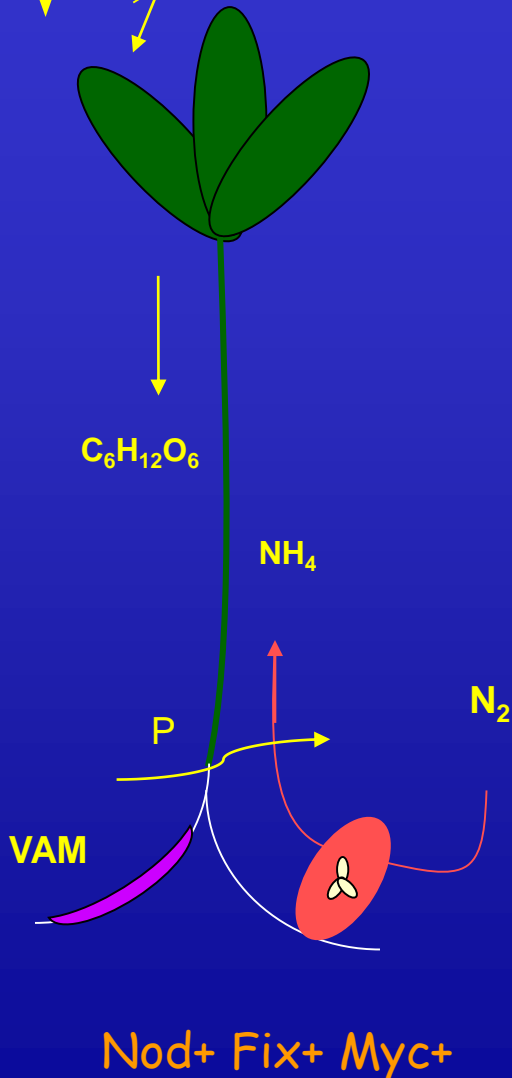
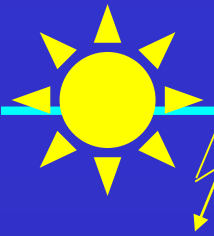
Génizolálás

- Kérdés: hogyan azonosítanánk egy gént egy komplex szerveződésű eukarióta genomban?

Arabidopsis	125 Mb/1C	~25 000 gén
rizs	430 Mb/1C	~ 50 000 gén
humán	3400 Mbp/1C	30 000 - 35 000 gén
búza	16000Mbp/1C	~ 25 000 gén
gyapot	2,200 Mb/1C	50 000 - 80 000 gén
<i>Medicago truncatula</i>	450 Mb/1C	~ 30 000 gén

- pl. mutagenézissel (transzpozonokkal egyből klónozzuk) de..
 - kicsi az egyedszám (nem tudunk elég utódot átnézni, mint mikrobáknál)
 - letális hatások
 - több helyre ugrik a transzpozon
 - szekvenálással megoldható
 - csak a mi génünkbe nem ugrik a transzpozon.. stb. MIT CSINÁLJUNK??
- kapcsoltsági analízissel
 - mihez legyen kapcsoltság? - **MARKEREK**
 - hogyan határozzuk meg a kapcsoltságot? **GENETIKAI ANALÍZIS**

Nitrogénkötő szimbiotikus kapcsolat



mutációk

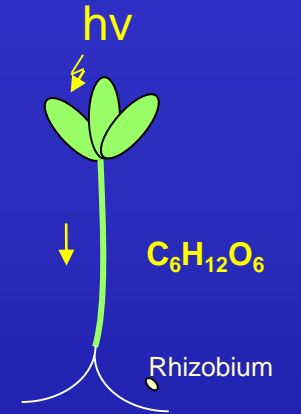


Ms-MN1008

Mt dmi1, dmi2, dmi3, nsp2, nsp2, nfp, lin, ipd3, stb.

Ps sym10; Ps sym19, stb.

Lj sym1; Lj sym2; sym5, stb.

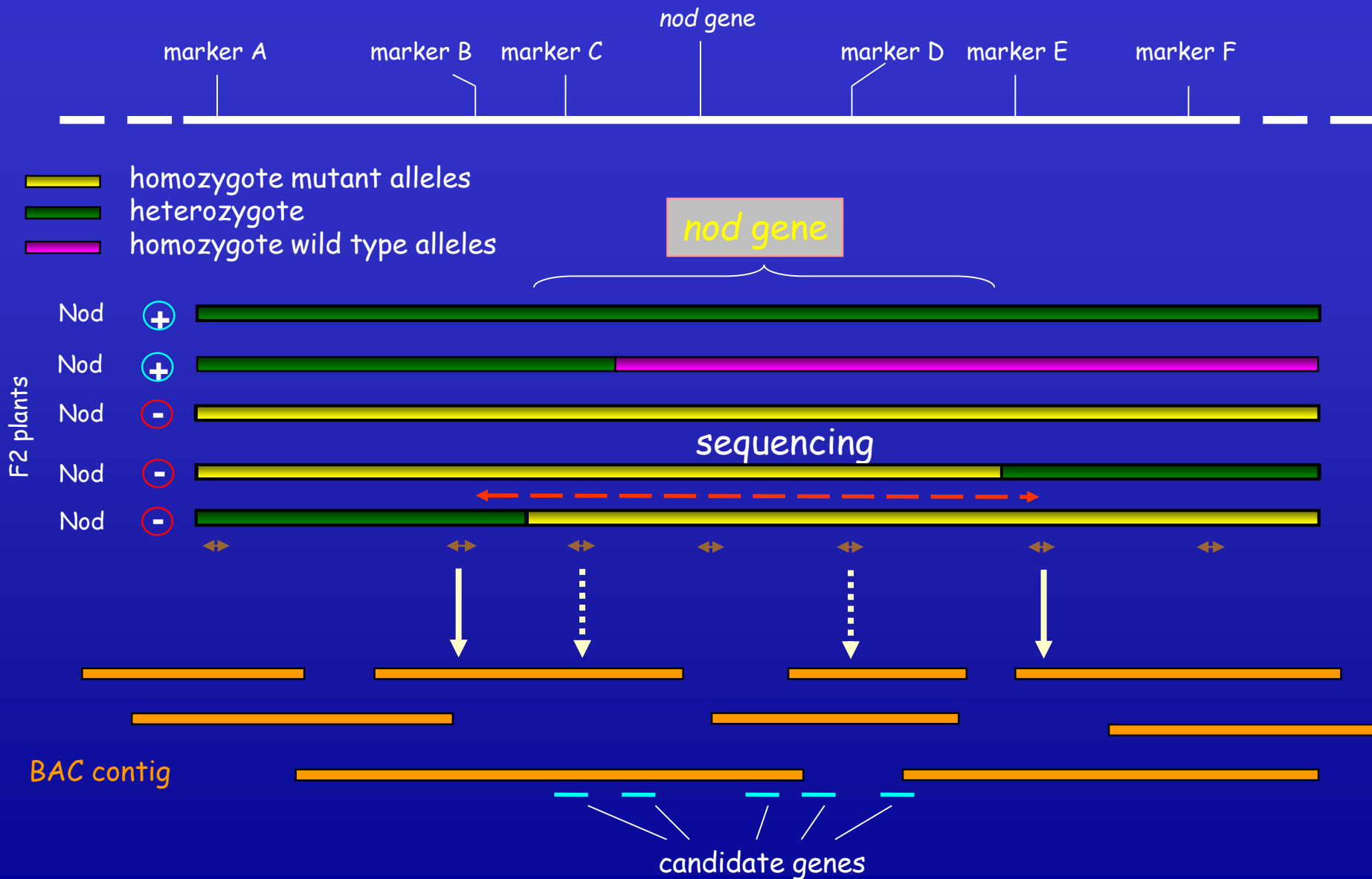


Nod- Myc-/+

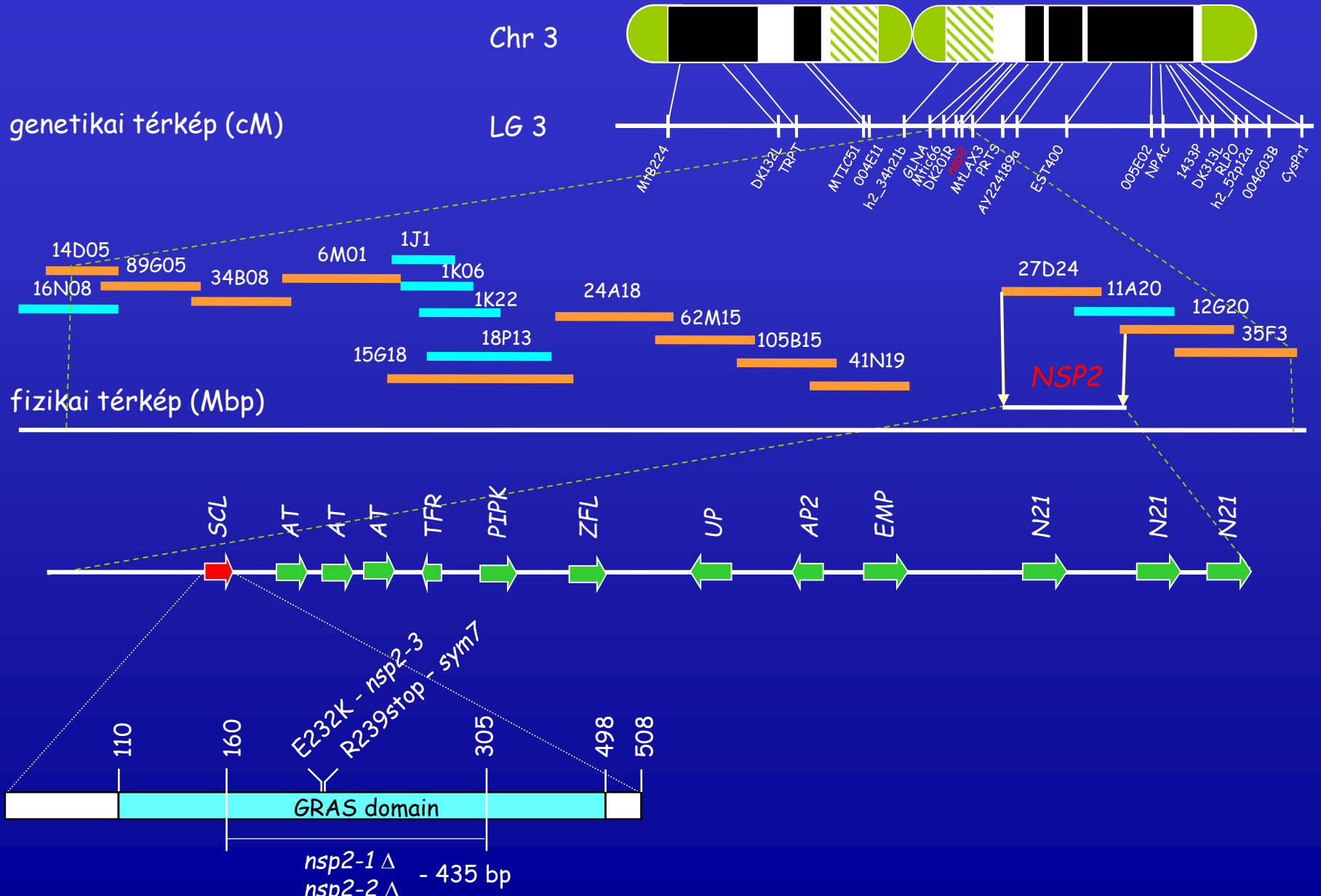
Nod+ Fix- Myc-/+

cél: a szimbiotikus nitrogénkötés biológiai folyamatának megismerése

Génklónozás - térképezésen alapuló génizolálás stratégiája



Az MtNSP2 gén térképezésén alapuló izolálása



Az MtNSP2 gén környezete a *M. truncatula* GenomeBrowser-en megjelenítve I.

<http://medicago.jcvi.org/medicago/index.php>

Available Tracks

- filter tracks
- A. MTGD Data 26
 - select all from category
 - Annotation (Mt4.0v2) 4
 - select all from category
 - Mt4.0v2 - Gene Loci
 - Mt4.0v2 - Protein Coding Genes
 - Mt4.0v2 - Transposable Element (TE) Genes
 - Mt4.0v2 - tRNA models
 - Assembly (Mt4.0) 4
 - select all from category
 - Assembly Golden Path
 - BAC Mapping
 - Gaps
 - Reference Sequence
 - Evidence 17
 - select all from category
 - ab initio Gene Predictions 2
 - select all from category
 - Augustus
 - FGeneSH
 - EST and Protein Alignments 3
 - select all from category
 - Sanger/454 ESTs (Sep 2012)
 - Tentative Consensus (TCs, MtGI 11.0)
 - UniRef90 non-Mt Plant Proteins (Sep 2012)
 - Orthologs (BLAT Alignment) 3
 - select all from category
 - Arabidopsis thaliana
 - Glycine max
 - Lotus japonicus
 - RNA-seq Mapping Coverage 9
 - select all from category
 - Blade (4 week)
 - Bud (4 week)
 - Nodule
 - Open Flower
 - Root (4 week)
 - Root (Control)
 - Root (NF-treated)
 - Root (Rhizobium Infected 36hr)
 - SeedPod

Medicago truncatula v4.0

chr3:32666101..32733350 (67.25)

32,675,000 32,687,500 32,700,000 32,712,500 32,725,000

Mt4.0v2 - Protein Coding Genes

- IL/p24 family protein
- YbaK/proline-tRNA ligase associated domain protein
- Medtr3g072620.1
- Medtr3g072630.2 zinc finger, C3HC4 type (RING finger) protein
- Medtr3g072630.1 zinc finger, C3HC4 type (RING finger) protein
- Medtr3g072610.1 AP2 domain class transcription factor
- Medtr3g072650.2 phosphatidylinositol-4-phosphate 5-kinase
- Medtr3g072650.1 phosphatidylinositol-4-phosphate 5-kinase
- Medtr3g072660.2 transcription factor
- Medtr3g072660.1 transcription factor
- Medtr3g072660.6 transcription factor
- Medtr3g072660.4 transcription factor
- Medtr3g072660.3 transcription factor
- Medtr3g072660.5 transcription factor
- Medtr3g072660.8 transcription factor
- Medtr3g072660.7 transcription factor
- Medtr3g072680.1 aspartate aminotransferase family protein
- Medtr3g072690.1 aspartate aminotransferase family protein
- Medtr3g072710.1 GRAS family transcription factor

BAC Mapping

CR538722.1

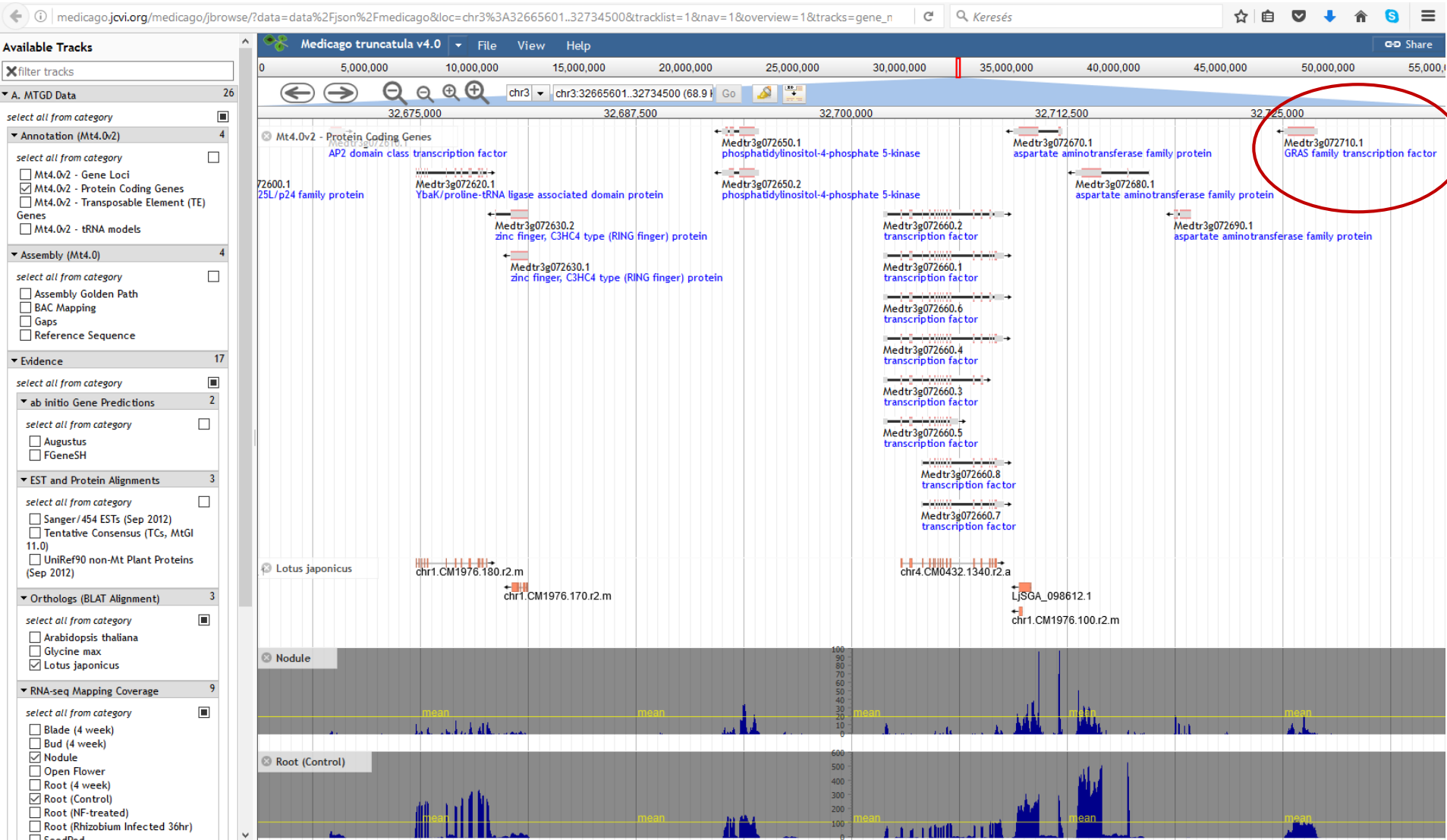
Sanger/454 ESTs (Sep 2012)

Tentative Consensus (TCs, MtGI 11.0)

TC145763 TC159224 TC151559 TC164439 TC150374 TC142190

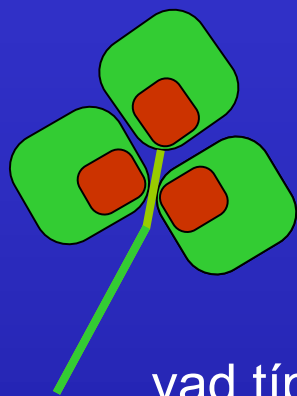
Az MtNSP2 gén környezete a *M. truncatula* GenomeBrowser-en megjelenítve II.

<http://medicago.jcvi.org/medicago/index.php>

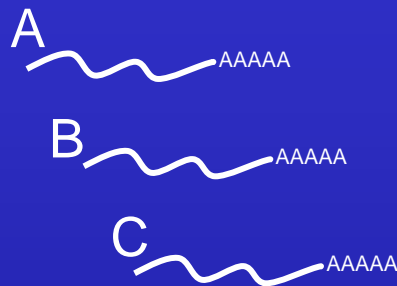


Deléciók genetikai térképezése 1.

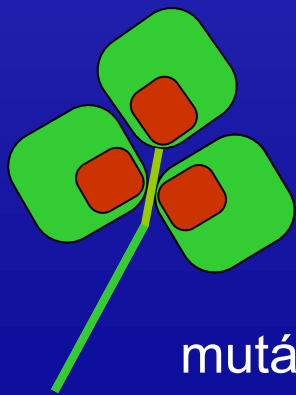
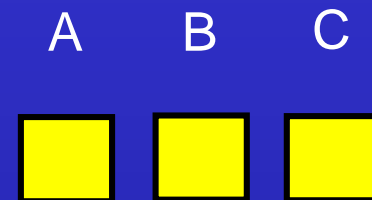
Transzkripció mintázat alapján történő klónozás (Affymetrix gene chip)



vad típus



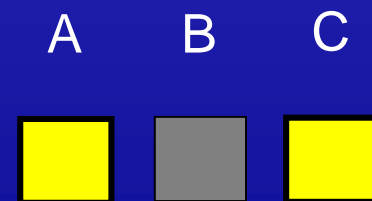
CHIP 1



mutáns
(fast neutron)

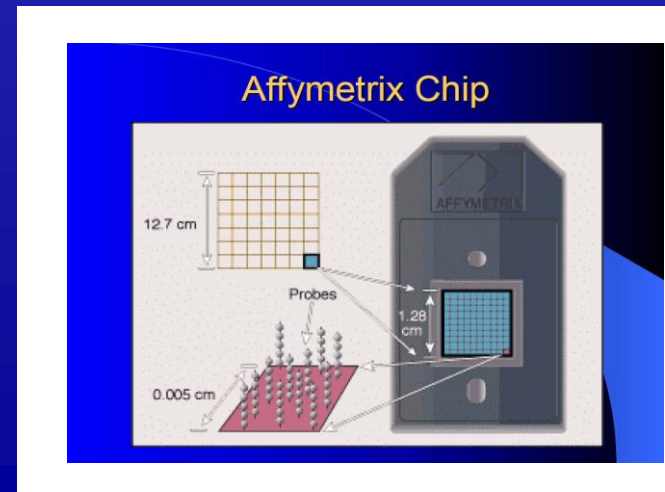


CHIP 2

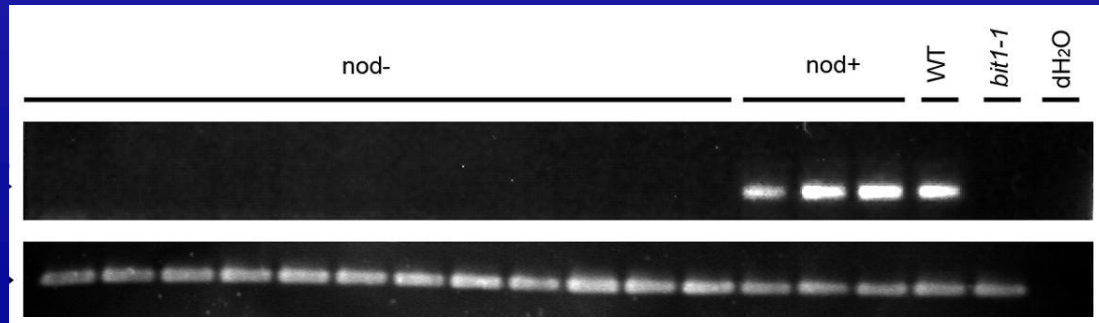
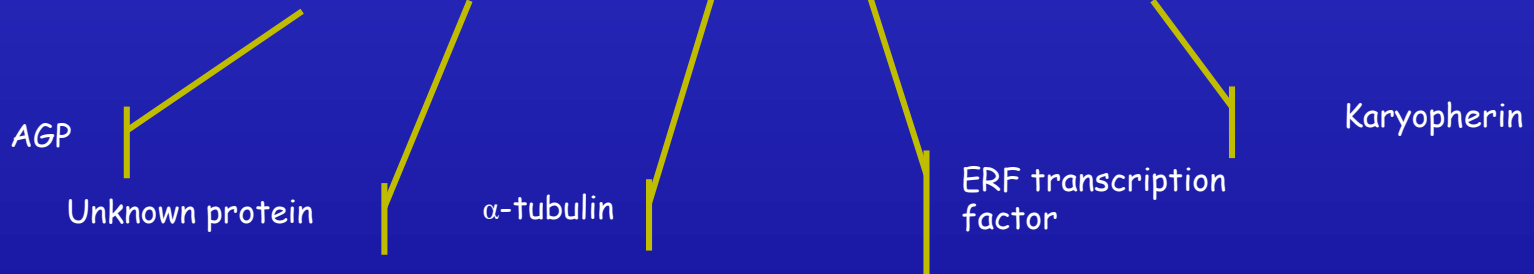
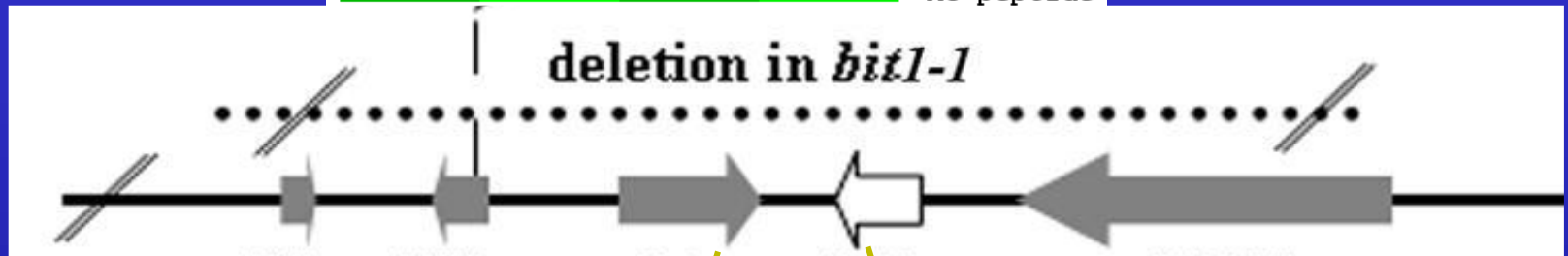
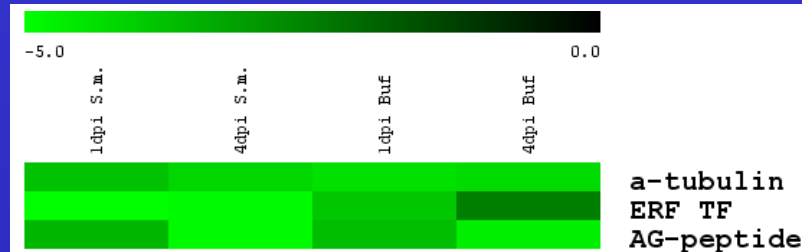


Transzkripció mintázat alapján történő klónozás (Affymetrix gene chip)

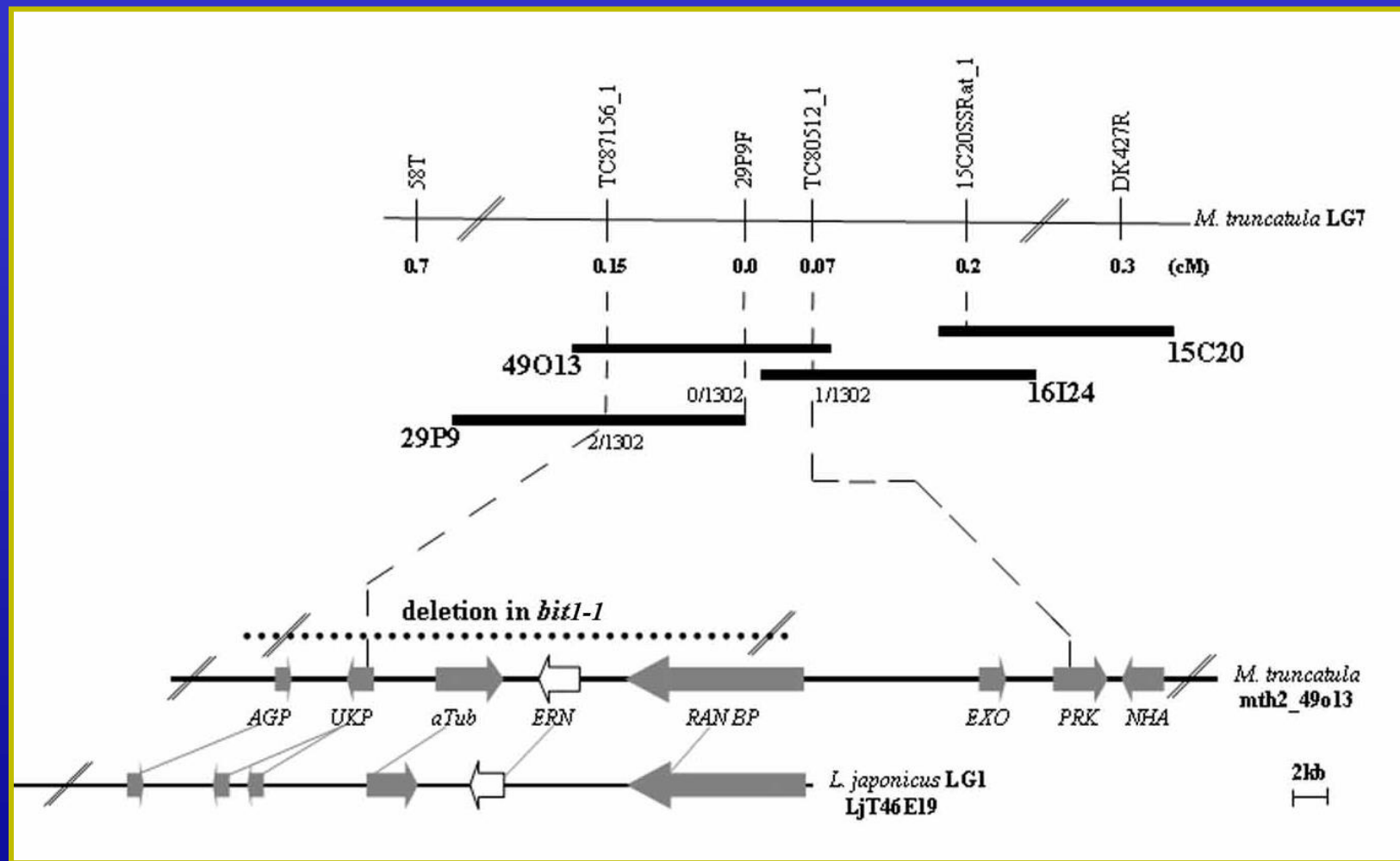
- minden egyes gént 16-20 szintetizált próbapár képvisel
- egy-egy próbapár egy tökéletesen párosodó (PM) és egy „mismatch”-et (MM) tartalmazó oligonukleotidból áll
- egy vizsgált génre a PM és a MM oligonukleotidok jelerősségei közötti átlagos különbségek adják az eredményt
- a Medicago Affymetrix gene chip 61 200 probe set („gént”) tartalmaz
 - 32 167 *M. truncatula* EST/mRNA
 - további 18 733 prediktált gént
 - 1 896 lucerna (*M. sativa*) EST
 - 8 305 *Sinorhizobium meliloti* gént ill. prediktált gén



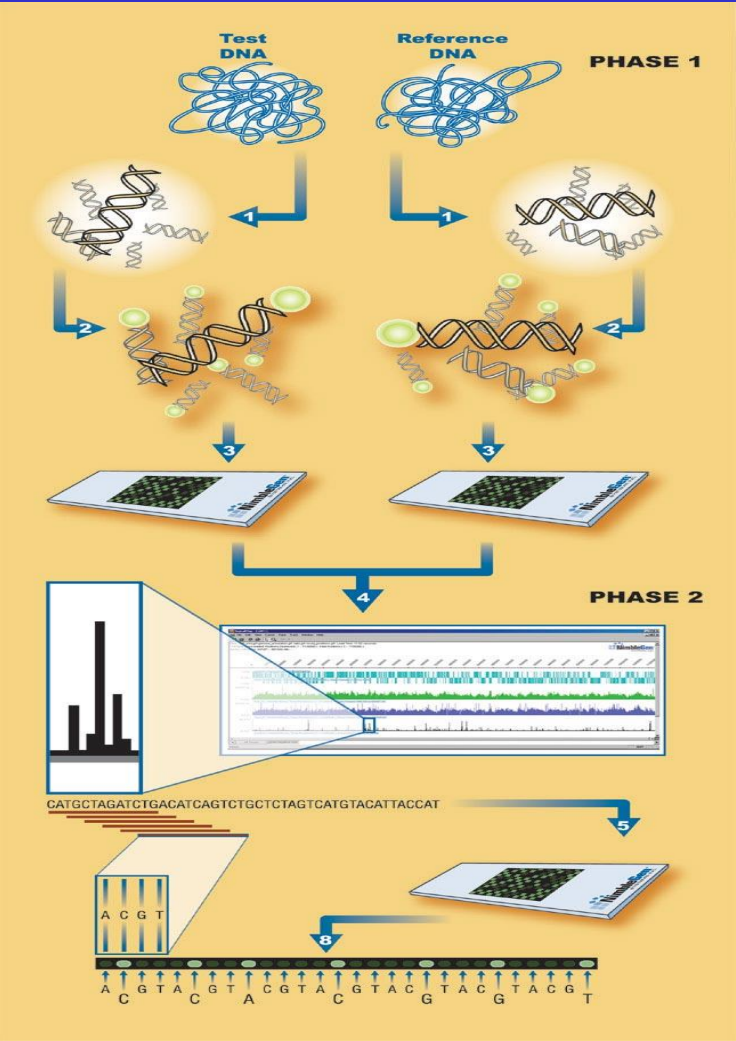
Az *MtERN1* (*bit1-1* mutáns) gén transzkripció mintázat alapján történő azonosítása



Az *MtERN1* (*pdl* mutáns) gén térképezésén alapuló klónozása



Deléciók genetikai térképezése 2. Oligonucleotide Array Comparative Genomic Hybridization (oaCGH)

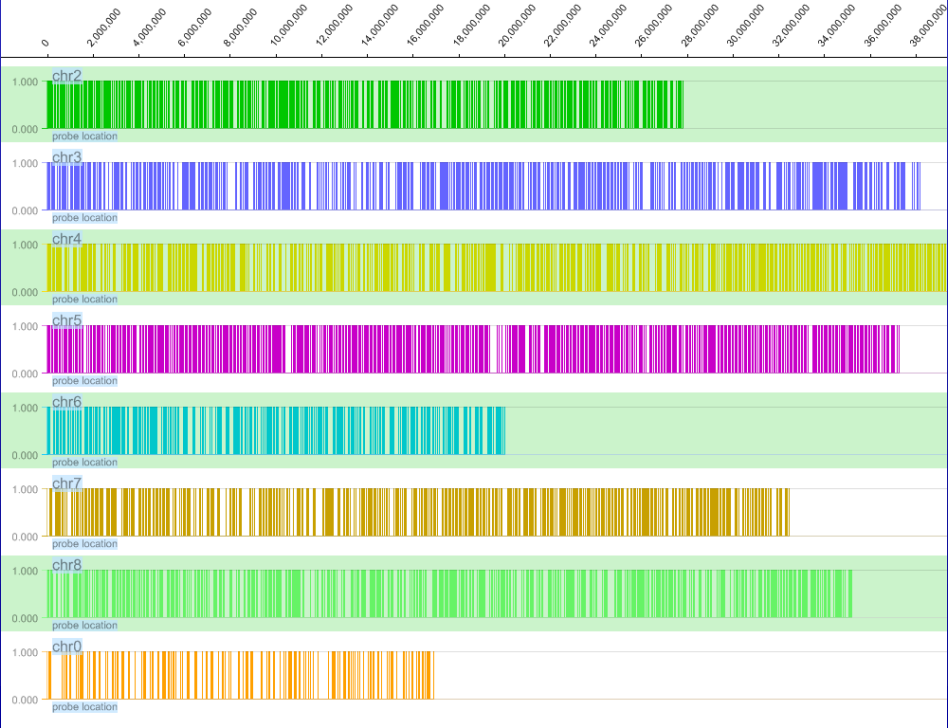


CGS - Comparative Genome Sequencing

Medicago truncatula - Oligonukleotidok

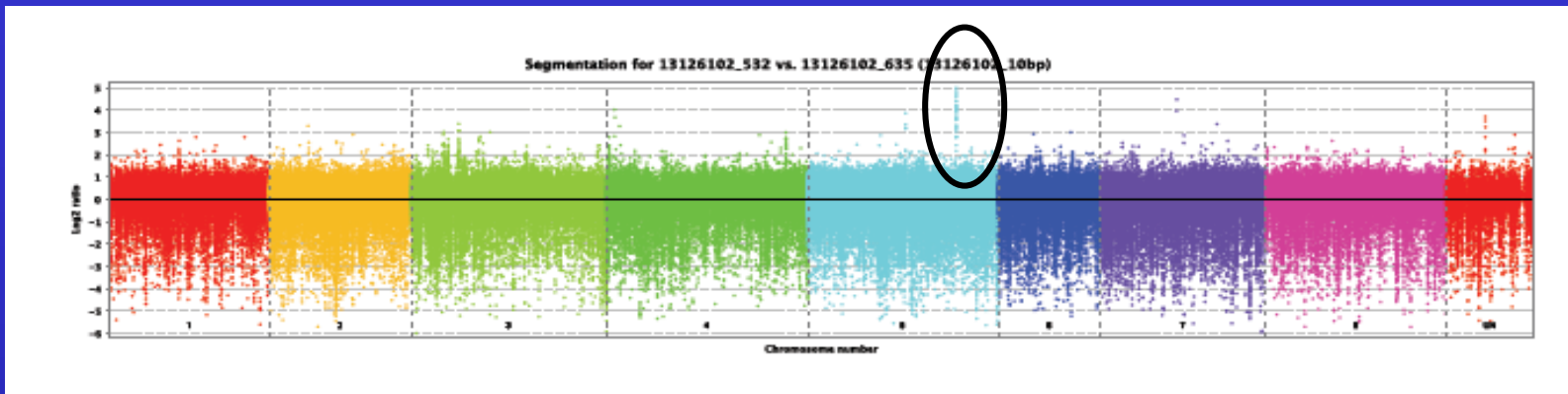
- exon 150bp
- intron/utr 300bp
- unigenes ≥ 1000 bp 300bp

Deléciók kimutatása
Egy és több kópiás amplifikáció kimutatása
Transzlokáció kimutatása

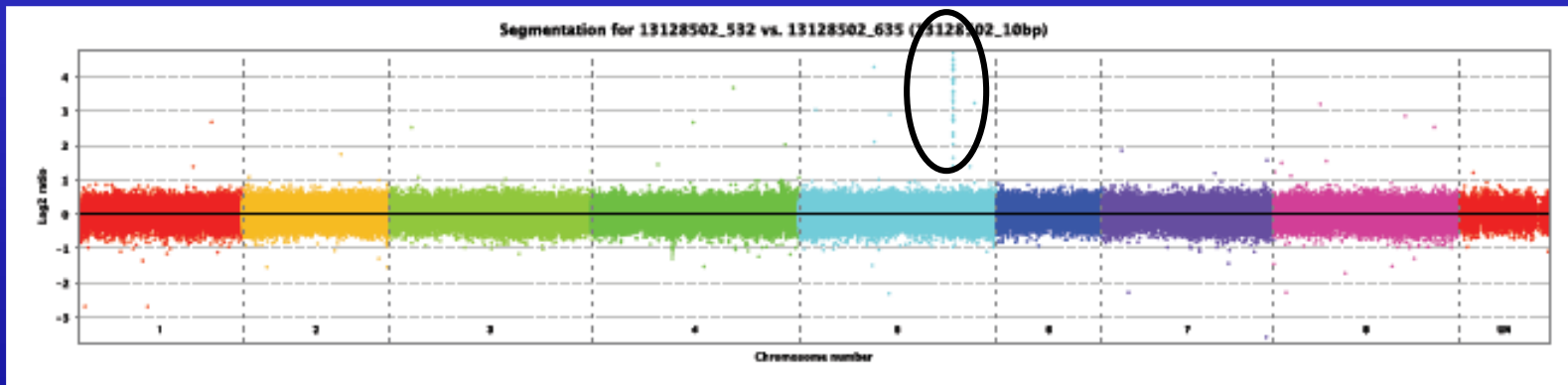


oaCGH vizsgálat *M. truncatula* növényeken

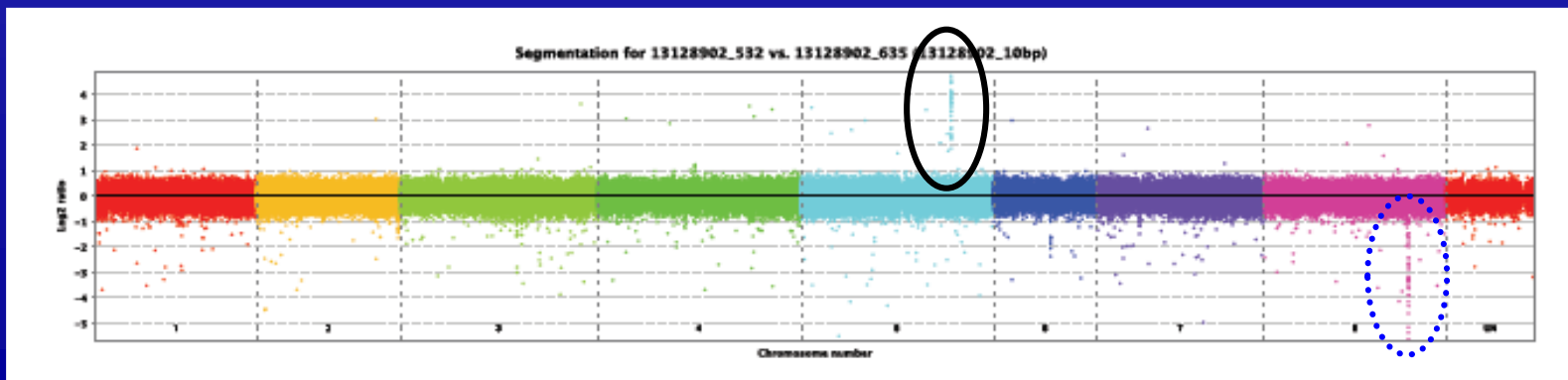
R108



PIN2



FNB#4



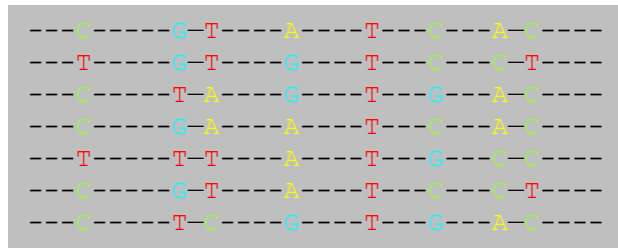
GWAS (Genome-Wide Association Studies)

vagy WGA study (whole genome association study)

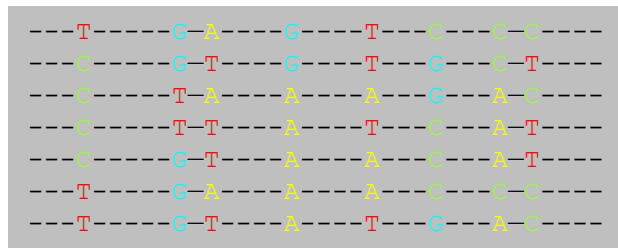
Mi a GWAS? - A genetikai változatosság és egy tulajdonság kapcsolatának vizsgálata egy genomon belül.

SNP-ék a genomban

vizsgált
tulajdonságra
kiemelt egyedek



kontrol



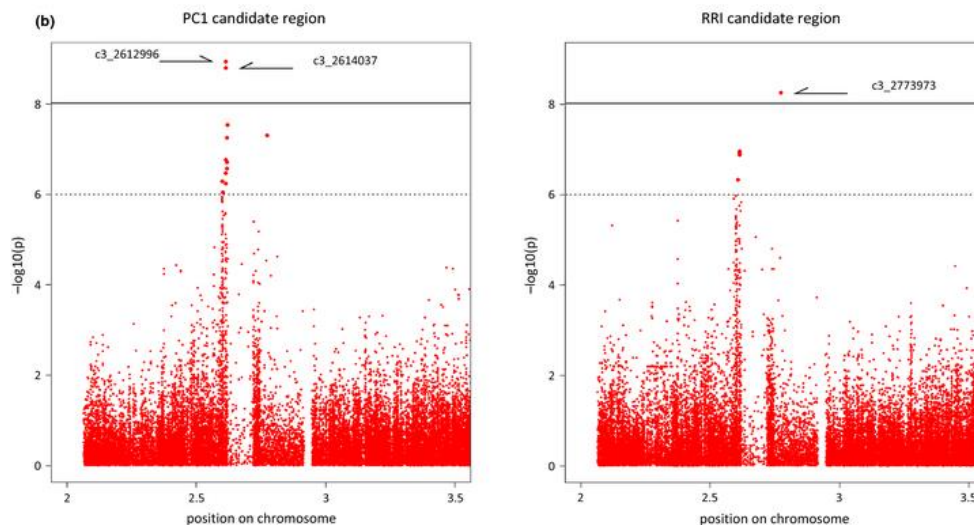
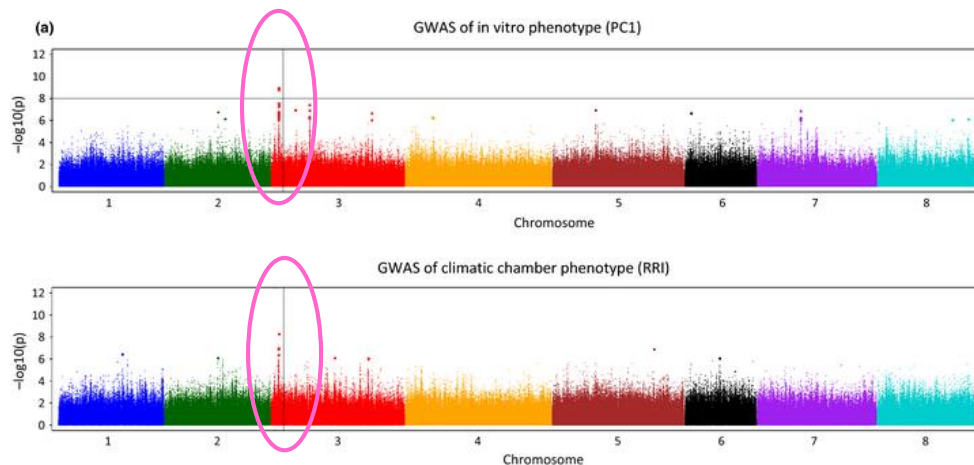
- Populációs csoportok kialakítása
- Teljes genom genotipizálása
- GWAS
- Fine mapping
- Jelölt gének keresése.
- Polimorfizmus azonosítás.
- Fenotípusért felelős SNP

GWAS - *Aphanomyces euteiches* elleni rezisztenciát biztosító génvariánsok azonosítása *Medicago truncatula*-ban I.

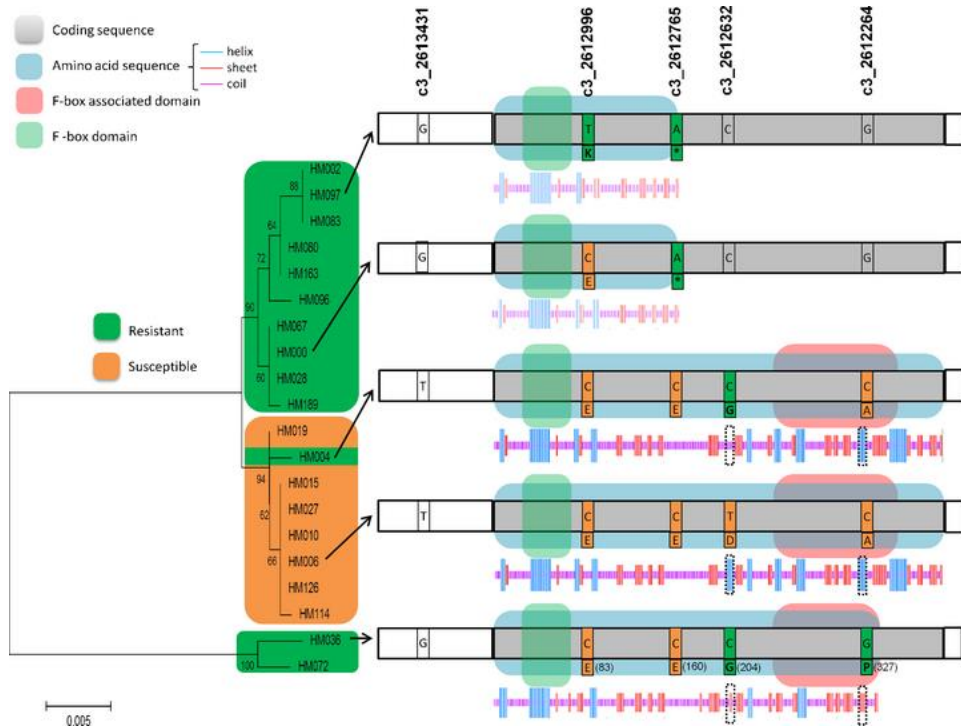


Bonhomme et al. 2014 New Phyt

- Két fenotipizálási módszer (*in vitro* és növénynevelő kamra)
- *M. truncatula* HapMap Project - 288 ökotípus (www.medicagohapmap.org)
- 16 515 723 SNP vizsgálat
- 2 SNP-t azonosítottak GWA vizsgálattal egy F-box proteint kódoló gén kódoló régiójában és promóterében



GWAS - *Aphanomyces euteiches* elleni rezisztenciát biztosító génvariánsok azonosítása *Medicago truncatula*-ban II.



Filogenetikai fa az *F-box* gén 1185 kb-os genomi szekvenciájában azonosított 65 SNP alapján

- Két fenotipizálási módszer (*in vitro* és növénynevelő kamra)
- *M. truncatula* HapMap Project - 288 ökotípus (www.medicago-hapmap.org)
- 16 515 723 SNP vizsgálat
- 2 SNP-t azonosítottak *GWA* vizsgálatnál egy *F-box* proteint kódoló gén kódoló régiójában és promóterében (chr 3)
- Nem azonosítottak a szenzitív és rezisztens vonalak között expressziós különbséget az *F-box* gén esetében
- a nem működő *F-box*-ot hordozó allélek a rezisztensek
- *F-box* protein a rezisztencia negatív regulátora

KÖSZÖNÖM A FIGYELMET!

**#Istand
withCEU
#aCEUval
vagyok**

GMO I.

- **GM vagy GMO:** genetically modified organism - génmódosított, genetikailag módosított vagy „génkezelt” élőlény
(hagyományos mutagének, poliploidizáció, keresztezések? genomszerkesztés?)
génmanipulált?
- **Géntechnológia alkalmazásával előállított módosított genetikai állományú élőlény**
 - transzgenikus: nem rokon fajtól származó genetikai információ
 - ciszgenikus: azonos vagy rokon (pl. vad) fajtól származó genetikai információ
- **A transzgenikus növények nélkülözhetetlenek a gének szerepének kutatásában**
- **Mikroorganizmusok:** az iparban elterjedt a használatuk - vakcinák, gyógyszerek, táplálékkiegészítők, vitaminok, enzimek, adalékanyagok, stb., pl. B2 és C vitamin, citromsav, lizozim, α -amiláz, (keményítő \rightarrow cukor), kimozin (sajtgyártás), inzulin, növekedési hormon, véralvadási faktorok
- **Állatok:**
 - omega-3 disznó (omega-3 zsírsav termeltetése egy nematóda gén beültetésével, 2006)
 - ENVIROPIG - jobb foszforhasznosítás (Canada, 1999)
 - megváltozott összetételű ill. szerkezetű tejet adó kecske
 - növekedési hormon termeltetése halakban (lazac, ponty, stb.)

GMO II.

- Növények:

- **Első generációs GMO:** termelőnek/fajtatulajdonosnak kedvező tulajdonság bevitele [herbicidrezisztencia (pl. glifozát), kórokozó elleni rezisztencia (Bt toxin, vírusrezisztencia), ipari szempontból kedvező tulajdonságot hordozó növény - szinte kizárólag ilyenek vannak köztermesztésben



Bt toxint
hordozó kukorica



glifozát rezisztens
szója

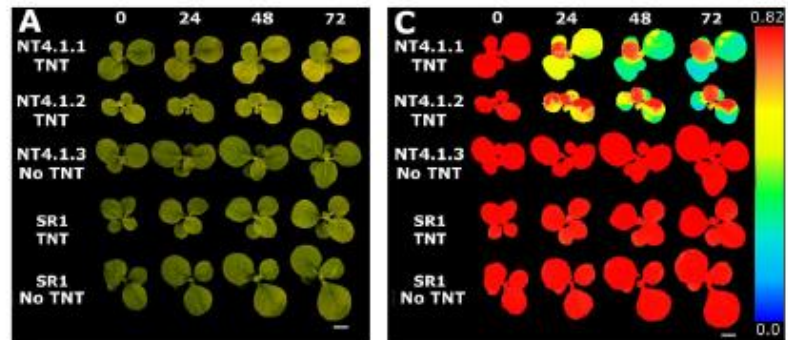
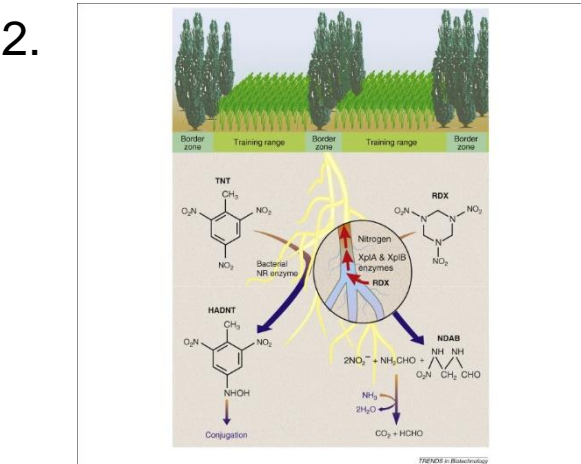
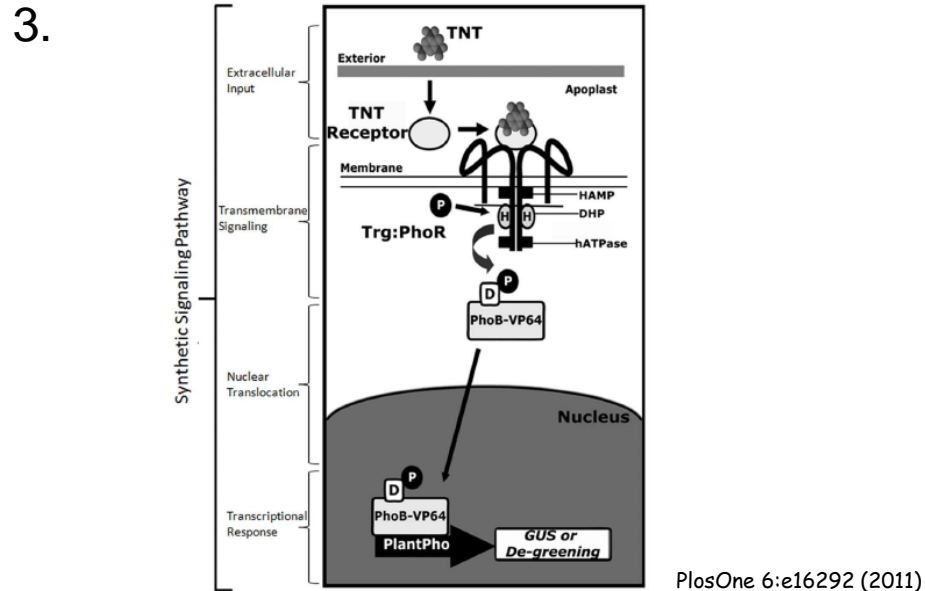
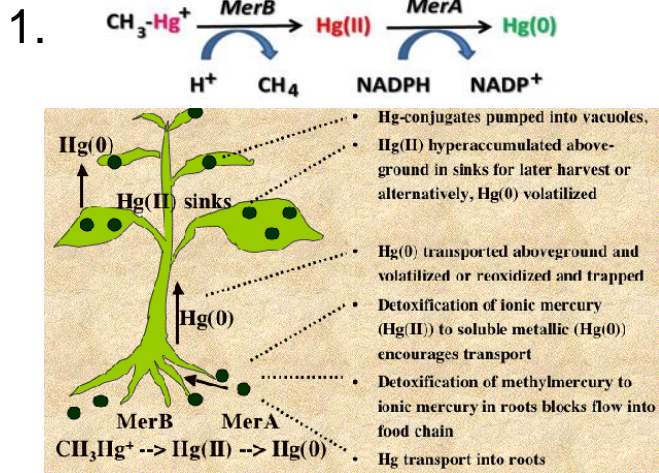
- **Második generációs GMO:** előnyös táplálkozási vagy élettani tulajdonsággal bíró növények (fogyasztó számára kedvező tulajdonság, pl. Lys/Trp összetétel megváltoztatása, Golden Rice)



GMO III.

• Növények:

- Harmadik generációs GMO: hideg-, és szárazságtűrő fajták
- Negyedik generációs GMO: gyógyszer-, és enzimtermelő fajták (bioreaktorok), fitoremedáció (1., 2.) , bioüzemanyag, növényi szenzorok (TNT)



GMO IV.

A kezdet - 1994, USA

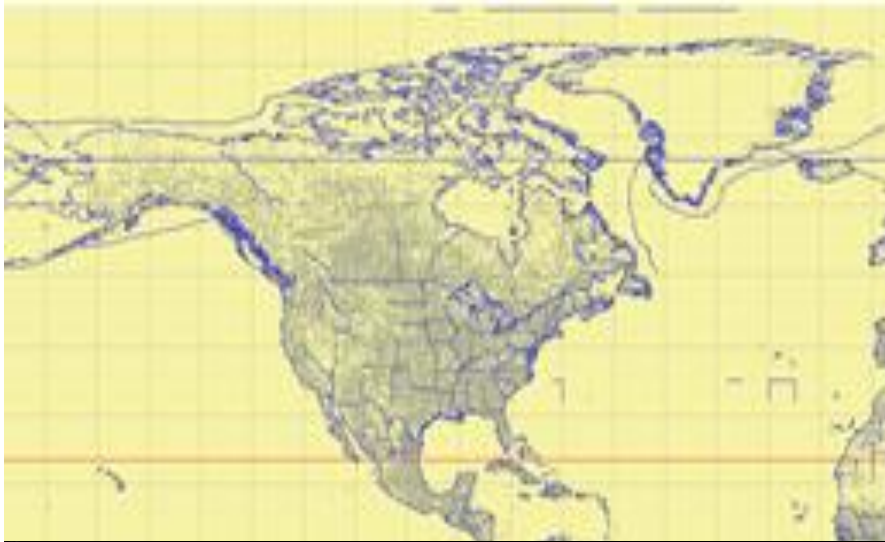
Az első szabadföldi termesztésbe és kereskedelmi forgalomba kerülő genetikailag módosított növény - UC Davis egyetem és a Calgene cég, későbbi tulajdonos - Monsanto



A Flavr Savr egy olyan paradicsom, amely a genetikai módosítás hatására tovább megőrizte frissességét, az érett termés lassabban puhult meg (sejtfal pektinjének lassított lebomlása). Termesztését 1997-ben megszüntették.

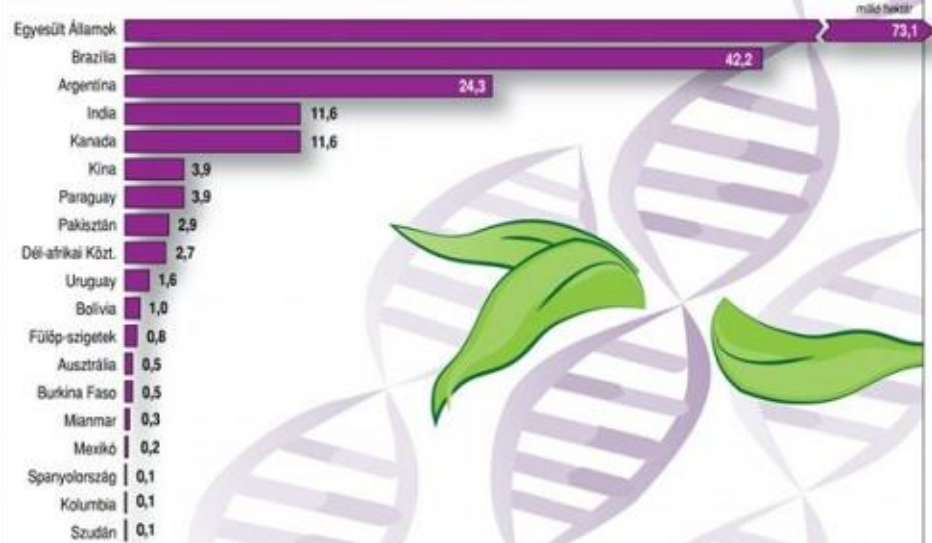
Zeneca - csökkentett víztartalmú, pürének alkalmas módosított paradicsom. 1996 és 1999 között 1.8 M tonna pürét adtak el az Egyesült Királyságban.

Genetikailag módosított növények termesztése



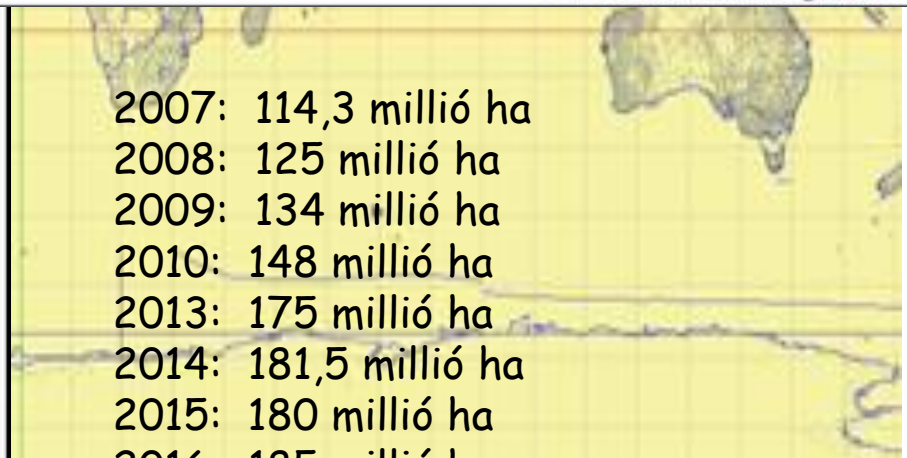
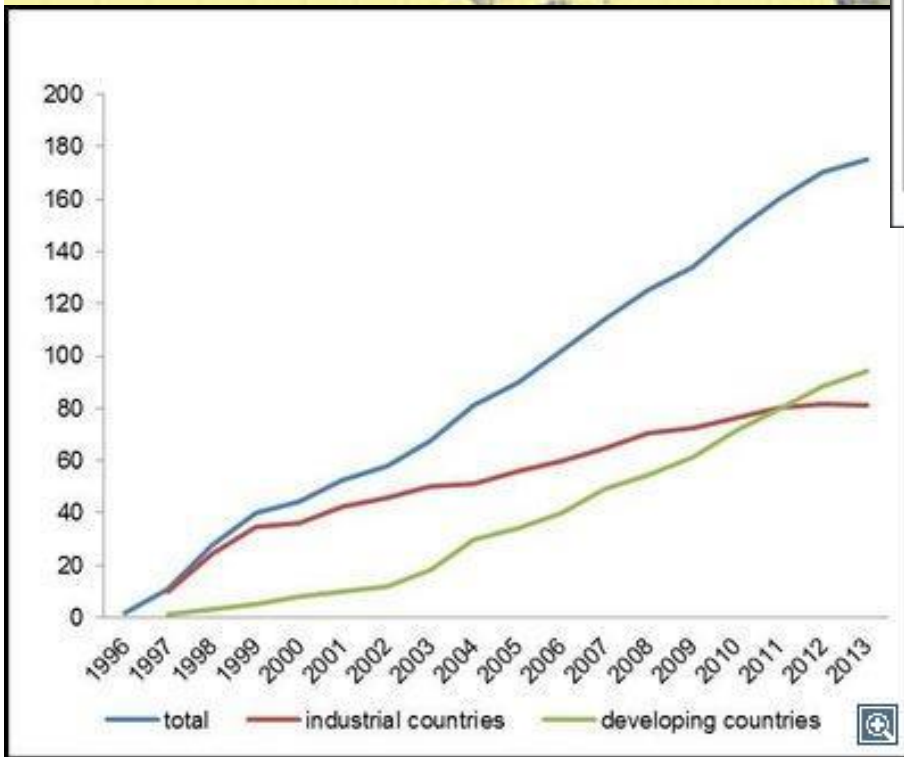
A genetikailag módosított növények vetésterülete a világon, 2014

A világon 28 országban termesztnek genetikailag módosított (gm) növényeket. 2014-ben a gm-növények vetésterülete összesen **181,5 millió hektár**, ebből:



0,1 millió hektár alatti vetésterület: Banglades, Chile, Costa Rica, Csehország, Honduras, Kuba, Portugália, Románia, Szlovákia

Forrás: ISAAA, MTVA Sajtó- és Foltérchivum / MTI



2007: 114,3 millió ha
 2008: 125 millió ha
 2009: 134 millió ha
 2010: 148 millió ha
 2013: 175 millió ha
 2014: 181,5 millió ha
 2015: 180 millió ha
 2016: 185 millió ha

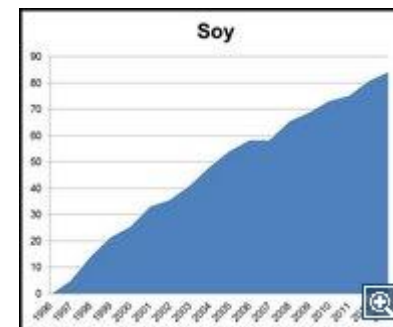
GM növények termesztése 2013

Country	Area (millions of hectare)	GMO crop
USA	70,1	S,M,C,R, SB,squash, papaya
Brazil	40,3	S,M,C
Argentina	24,4	S,M,C
India	11,0	C
Canada	10,8	R,M,S, SB
China	4,2	C, poplar, papayas, tomato, sweet pepper
Paraguay	3,6	S
South Africa	2,9	M,S,C
Uruguay	1,5	S,M
Bolivia	1,0	S
Philippines	0,8	M
Australia	0,6	C,R
Burkina Faso	0,5	C
Mexico	0,2	C,S
Chile	<0.1	M,S,R
Colombia	<0.1	C,carnation
Honduras	<0.1	M
Costa Rica	<0.1	C,S
EU (six member states)	0,1	M

Field areas for GM plants according to country in 2013; field are in millions of hectares

S = Soybeans, M= Maize, R = Rapeseed, C = Cotton, SB=sugarbeet

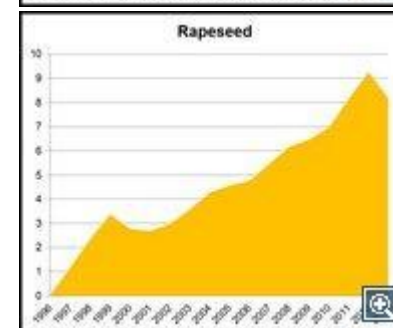
Source: ISAAA Brief No 46-2013



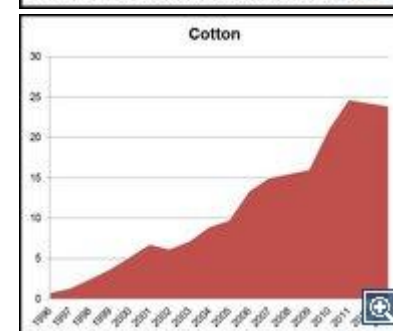
84,5 millió ha
79%-a GM



57,4 millió ha
32%-a GM



8,2 millió ha
24%-a GM



23,9 millió ha
70%-a GM

GM növények termesztése világszerte 26 országban történt ~ 180 millió hektáron (2015/2016)

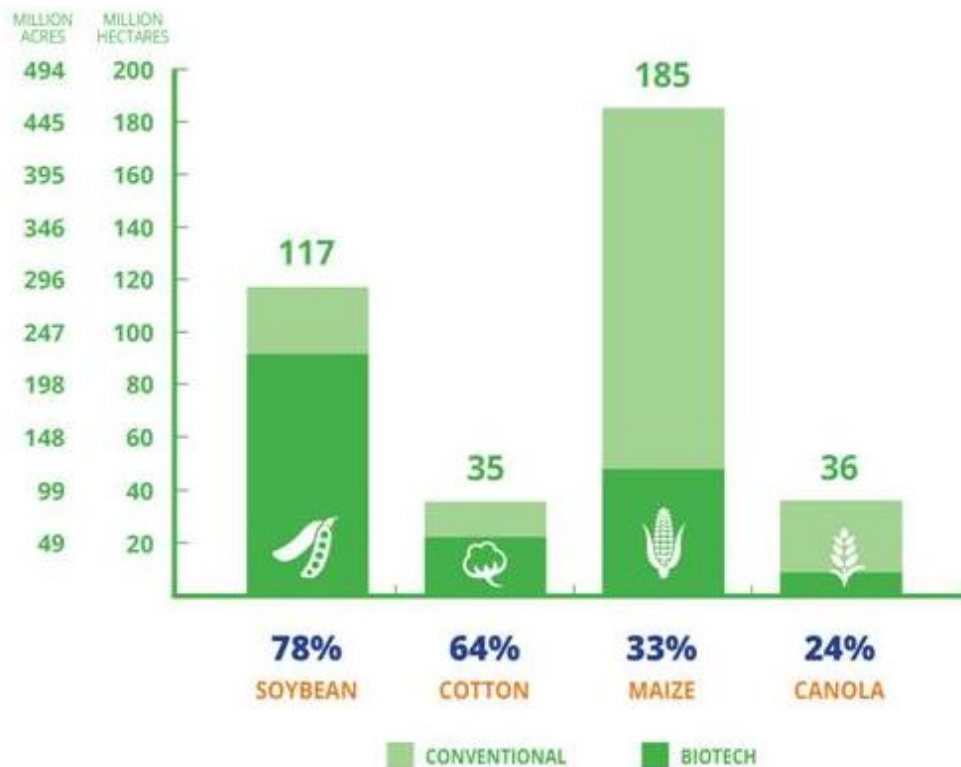
Rank	Country	2015	2016
1	USA*	70.9	72.9
2	Brazil*	44.2	49.1
3	Argentina*	24.5	23.8
4	Canada*	11.0	11.6
5	India*	11.6	10.8
6	Paraguay*	3.6	3.6
7	Pakistan*	2.9	2.9
8	China*	3.7	2.8
9	South Africa*	2.3	2.7
10	Uruguay*	1.4	1.3
11	Bolivia*	1.1	1.2
12	Australia*	0.7	0.9
13	Philippines*	0.7	0.8
14	Myanmar*	0.3	0.3
15	Spain*	0.1	0.1
16	Sudan*	0.1	0.1
17	Mexico*	0.1	0.1
18	Colombia*	0.1	0.1
19	Vietnam	<0.1	<0.1
20	Honduras	<0.1	<0.1
21	Chile	<0.1	<0.1
22	Portugal	<0.1	<0.1
23	Bangladesh	<0.1	<0.1
24	Costa Rica	<0.1	<0.1
25	Slovakia	<0.1	<0.1
26	Czech Republic	<0.1	<0.1
	Total	181.5	179.7

*Biotech mega-countries which grew more than 50,000 hectares, or more.

**Rounded-off to the nearest hundred thousand.

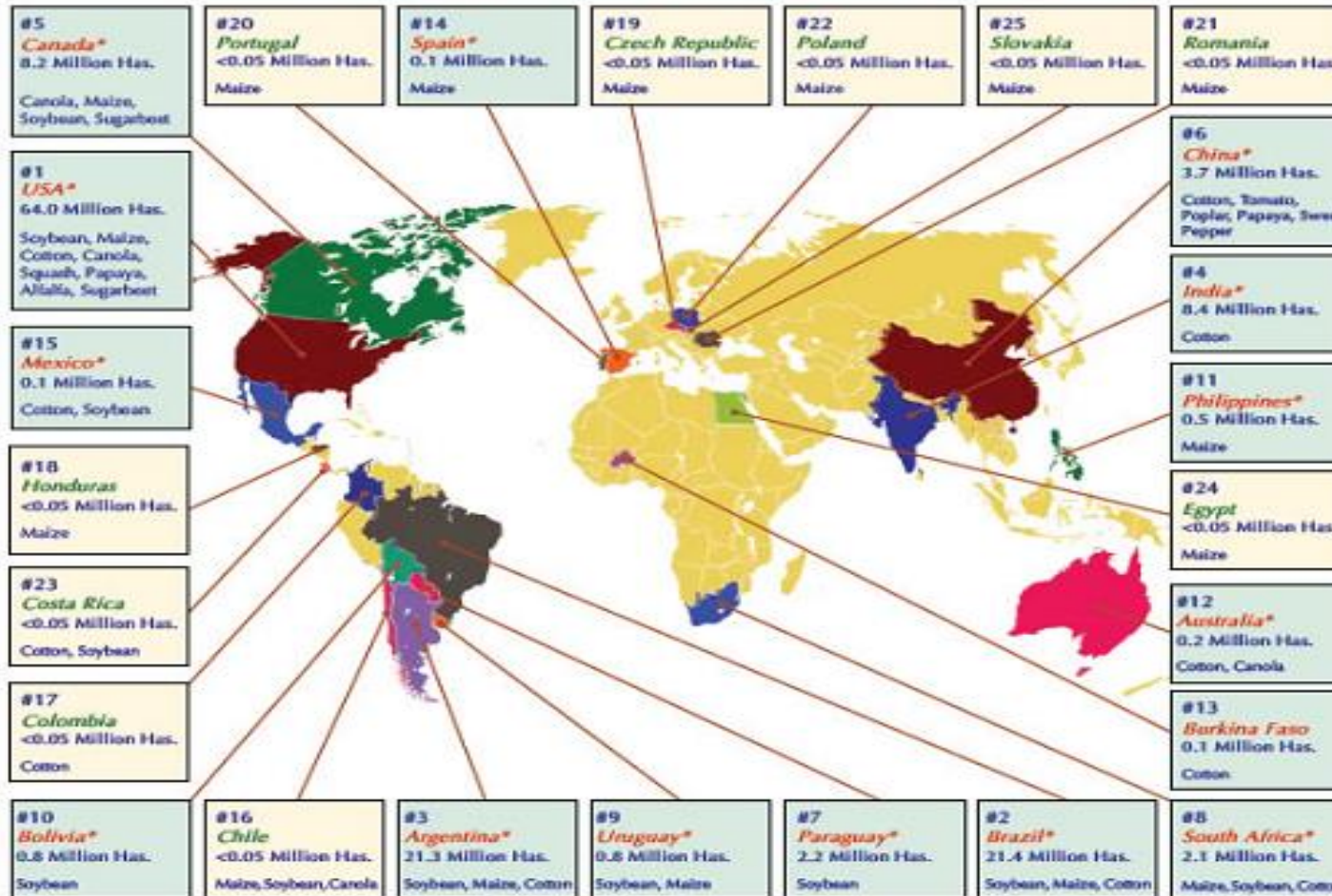
Source: ISAAA, 2016.

USA 72.9 M ha
 Brazília 49.1 M ha
 Argentina 23.8 M ha
 Canada 11.6 M ha
 India - 10.8 M ha



Európában GM növények termesztése 2009-ben 6 országban, 2010-ben 8 országban történt

Biotech Crop Countries and Mega-Countries*, 2009



* 15 biotech mega-countries growing 50,000 hectares, or more, of biotech crops.

Source: Clive James, 2009.

2009 - 94750 hektáron

kukoricamoly rezisztens
Mon810 GM kukorica
2013 - Spanyolo., Port.

2010 - Amflora burgonya (csak amilopektin)
Svédország, Németország
Cseh köztársaság
2011 óta nem termesztik



Magyarországon köztermesztésben nincs GM növény

Egyéb növények - 2013

Cukorrépa: USA-ban 2007 óta, 2013 - 460000 ha, a termelt mennyiség 95%-a herbicid rezisztens;
Kanada 15000 ha; 96%
Cukkini: 2000 ha
Papaya: Hawaii - 2000 ha (60% GM), Kína - 6275 ha
Nyárfa: Kína 450 ha
Kisebb mennyiségben GM paradicsom, paprika, petúnia

EU 2013

- Mon 810 és az Amflora rendelkezik termesztési engedéllyel, Pioneer 1507- toxintermelő és glufozinátrezisztens, 2013
- MON810 - Spanyolország és Portugália; 30% ill. 10%-a a kukorica mennyiségnek GM (137000 ha, 9300 ha)
- 2013 - Az Európai Bizottság döntése nyomán egy harmadik GMO növény is közel került a szabadföldi termesztéshez: a Pioneer 1507 kukorica egyszerre toxintermelő és glufozinátrezisztens
- élelmezési és takarmányozási célra 49 féle GMO növényt lehet felhasználni az Európai Unióban
- ezek több mint fele, 27 valamilyen génmódosított kukorica, a többi olajosmagvak, szója- és gyapotféle

Magyarország 2013

- vetési moratórium mindkét uniós szinten engedélyezett GM növényre (Amflora GM burgonya, MON810 GM kukorica)
- Magyarországon kísérleti céllal engedélyezett GMO kibocsátások adatbázisa - <http://biosafety.abc.hu>

GMO szabályozás

USA: 1986 – minden laboratórium kívüli felhasználás hatósági engedélyhez kötött; 1997 óta csak bejelentési kötelezettség

EU: 2004-ben oldotta fel a tilalmat, de több ország korlátozást vezetett be

A fogyasztónak joga van választani, ezért az EU törvényhozás szabályozta a GMO-k kereskedelmi forgalomba kerülését és jelölését az Európai Parlament és a Tanács 1829/2003/1830/2003 rendelete alapján:

jelölési kötelezettség van minden olyan alapanyagra és termékre amelyben az EU-ban engedélyezett GM növényekre nézve a GMO tartalom 0,9% felett van

2018 - Court of Justice of the European Union (ECJ): Gene-edited crops should be subject to the same stringent regulations as conventional genetically modified (GM) organisms (new inventions such as CRISPR–Cas9 crops need to go through the lengthy approval process of the European Union)

„Organisms obtained by mutagenesis are GMOs”

<https://www.nature.com/magazine-assets/d41586-018-05814-6/d41586-018-05814-6.pdf>

GMO helyzete Magyarországon

Magyarország: Géntechnológiai tevékenységről szóló törvény (XXVII/1998) 1999.01.01-től - minden GMO-val kapcsolatos tevékenység engedély köteles

Engedélykérelem: Géntechnológiai Eljárásokat Véleményező Bizottság

Magyarország alaptörvénye szerint Magyarország a testi és lelki egészséghez való jog érvényesülését - többek között - genetikailag módosított élőlényektől mentes mezőgazdasággal is elősegíti.

Magyarország takarmány fehérjeigénye:

- 580 e t/év

- ennek kb. 50%-át szójaból fedezzük, aminek 90%-a import

- a világ 350 M t/év szójatermelésének 96%-a GMO

- Mo. szója termelése 60-70 e Ha (GMO-mentes), ennek 85%-a export (kb 100 USD

felár/t)

- a magyar szójaigényt kb. 230 e Ha-ról lehetne kielégíteni („tudásintenzív” növény)

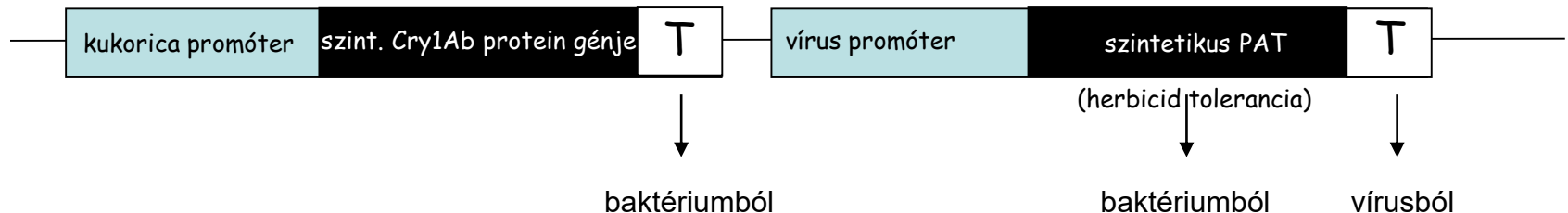
- fehérjeprogramok

GMO példák I.

- Példák:

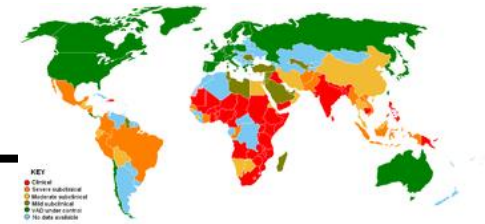
1. MON810 jelű, rovarrezisztens génmódosított kukorica; *Bacillus thuringiensis* (Bt) talajbaktérium toxintermelő génjét építették be (kukoricamoly ellen; Mo-on ritka kártevő); USA - 1996 óta, világon 21, EU-ban 6 országban termesztik; a kivont toxint inszekticidként alkalmazzák (!!)

MON810 génkonstrukció (Heszky, Agrofórum 2008):



phosphinothricin acetyltransferase (PAT)

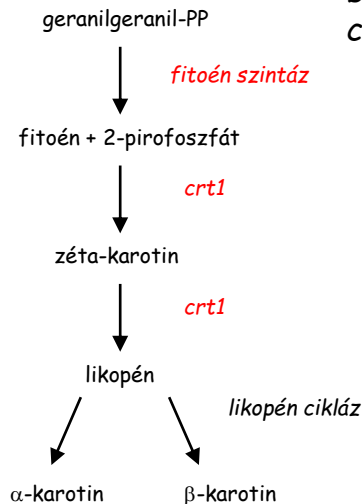
GMO példák II.



Példák:

2. **Golden Rice:** arany rizs - Ingo Potrykus (Swiss Federal Institute of Technology) és Peter Beyer (Univ. of Freiburg); Ye et al. 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. Science 287 (5451): 303-305.

- A vitamin hiány vakságot és egyéb fejlődési rendellenességet okoz (egyoldalú táplálkozás), ~ 250-500 E eset /év
- β -karotint (A vitamin prekuzort) termelő rizs előállítására endospermiumban két bioszintézisben résztvevő gén transzformációjával: *psy* (phytoene synthase) nárciszból és *crt1* *Erwinia uredovara* baktériumból



- létrehozásakor nagy reményeket fűztek hozzá (nem profitszerzés a cél; a szabadság tulajdonosok lemondtak minden jogokról)
- Fülöp-szigeteken és Tajvanon szabadföldi termesztés során 4-5x több mint β -karotint termelt, mint üvegházban
- köztermesztése nagy ellenállásba ütközött (környezetvédelmi és antiglobalista aktivisták; változatosabb étrend?)

Golden Rice 2 (2005) - 23x több β -karotint termel mint az előző verzió (kukorica fitoén szintáz génnel)

Bill és Melinda Gates (2005) - támogatás a rizs további minőségi javítására (fehérje tartalom, E vitamin, vas, cink)

3. magasabb Ca tartalmú sárgarépa előállítás (2008) - 15 kg/napi fogyasztás lenne szükséges

GMO példák III.

- További példák:

4. szója - herbicidrezisztens fajták, allergén fehérjék csökkentése

5. kukorica - Lys termelő, nagy vas, C-vitamin, folsavtartalmú fajták fejlesztése folyik

6. gyapot - herbicid- és rovarrezisztens fajták

7. repce - herbicidrezisztens fajták, nagy mennyiségű laurinsav termelés

8. papaya - gyűrűfoltosodást okozó vírus (PRSV) rezisztens változat

9. cukornád - rovarrezisztens és magasabb cukortartalmú (2. gen.) fajták

10. cukorrépa - herbicidtoleráns fajták.

11. lucerna - herbicidtoleráns fajták (glifozát)

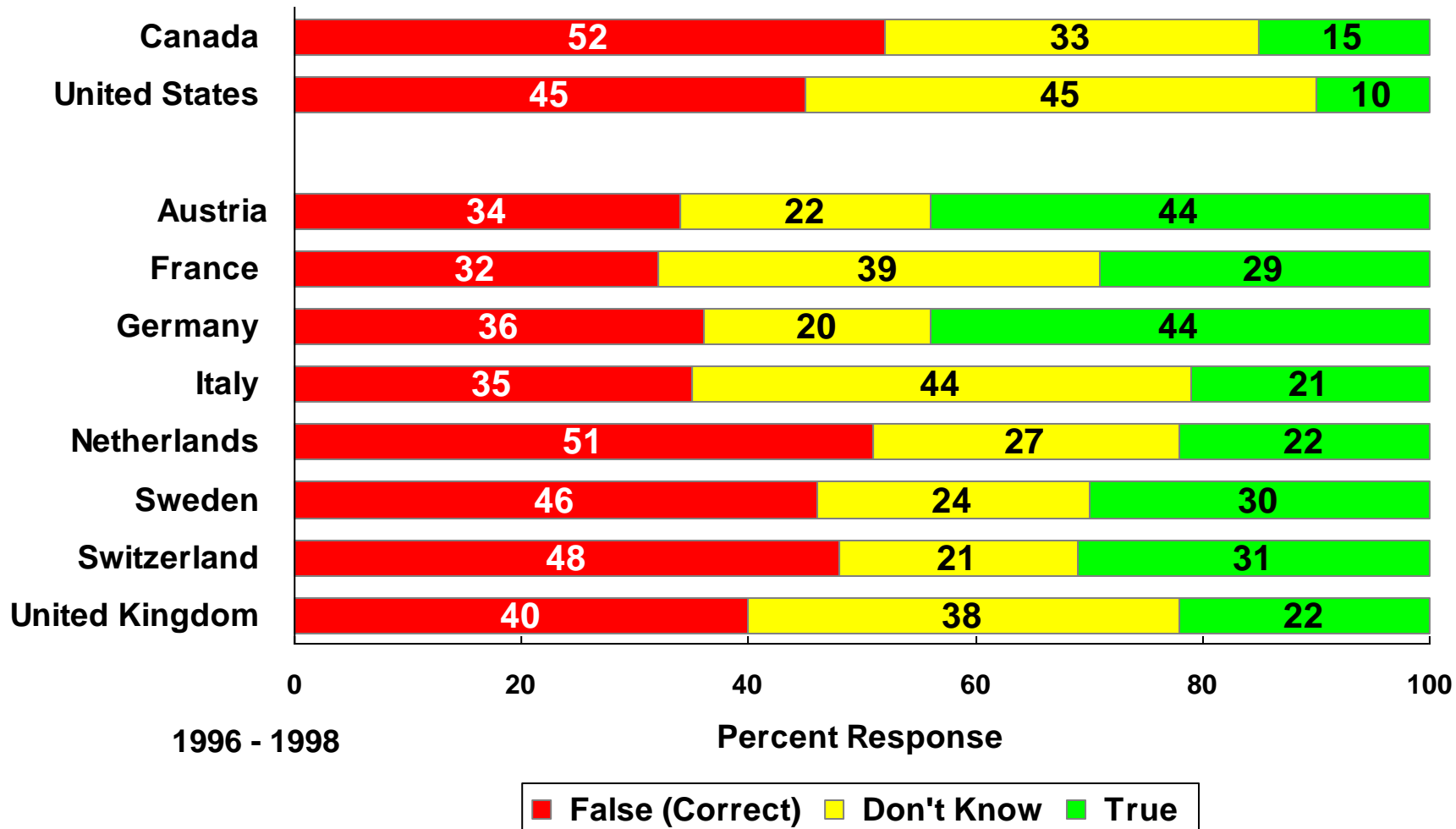
12. burgonya - ('Amflora' változat - BASF Plant Science) - majdnem tiszta amilopektint termel (amilóz tartalom minimális); csak az előállító biotechnológiai céggel szerződött partnerek számára elérhető, nincs közpiacon, az EU-ban ipari felhasználásra termesztik 2010-től

13. paradicsom - a puhulás nélkül érő Flavr Savr az első kereskedelmi célra termesztett génmódosított élelmiszer volt, forgalmazása nem termelt nyereséget, kivonták a piacról (1997). Azóta több zöldség és gyümölcsféle késleltetett érésű/puhulás nélkül érő változatát is előállították, jelentőségüket a meghosszabbodott eltarthatósági idő adja.

GMO - mezőgazdasági és társadalmi vonatkozások

- probléma: 2010 - 6,8 milliárd a világ népessége
2050 - 9,1 milliárd
- cél: elegendő mennyiségű élelmiszer előállítás, terméshozamok növelése
termőterület növelése
hatékonyság növelése (öntözés kiterjesztése, műtrágyázás, vegyszeres védekezés, betakarítási veszteség és a fejlett világban az élelmiszerfogyasztás csökkentése - a nyugati világban a gabona 60%-át állatok etetésére használják, stb.)
 - ezek korlátozott lehetőségek (?)
 - nem a nyugati világban kellene növelni a hatékonyságot
- GMO a megoldás? Forradalmasítja a mezőgazdasági termelést?
- Előnyök: ezekről már esett szó - jelenleg > 100 millió ha-on természetnek GM növényeket
- Hátrányok: etikai - egészségügyi vonatkozások
génszökés/génáramlás problémája (2010. aug.) - hímsteril fajták
 - Brassica napus (Bt toxin) - Brassica rapa - vad rokon faj:géntranszfer?
antibiotikum-rezisztencia kialakulása/terjedése
környezeti károk, ökológiai veszély (toxintermelő fajták)
élelmiszerbiztonság, keveredés, allergének, (0,9% alatt nem kell jelölni)
helyi termelők és multinacionális cégek érdeke ütközhet (mag- és vegyszerköltség)
szabadalmi költségek - magyar fajta?
- Eredendő rossz? - nukleáris energia? Nem a technológia a probléma!
- Tudáshiány! kutatók feladata- lehetséges veszélyek -> biztonsági rendszerek kifejlesztése
a társadalom tájékozottsága és ismerete alacsony szintű

"Ordinary tomatoes do not contain genes, while genetically modified ones do"



KÖSZÖNÖM A FIGYELMET!

**#Istand
withCEU
#aCEUval
vagyok**

Triticum urartu (AA) × *Aegilops speltoides* (BB)



T. turgidum (AABB) × *T. tauschii* (DD)



The common bread wheat (*Triticum aestivum*) is an allohexaploid containing three distinct sets of chromosomes derived from three different diploid species of goat-grass (*Aegilops*) through a tetraploid intermediary (durum wheat).



T.astivum
AABBDD

Table 3: Approved “plants with novel traits” versus GM crops/foods on the market

This table consolidates the CFIA’s list of approved “Plants with Novel Traits” (PNTs), with additional information from Health Canada’s list of “Novel Foods”, to explain which of these PNTs are GMOs, and which are on the market.

	CROP	GMO	MARKET (COMMERCIAL) STATUS	GROWN IN CANADA	IMPORTED TO CANADA	GM TRAIT(S)
1	Canola	✓	Grown in Canada	✓	✓	Herbicide tolerant
2	Corn	✓	Grown in Canada	✓	✓	Insect resistant Herbicide tolerant
3	Soy	✓	Grown in Canada	✓	✓	Herbicide tolerant
4	Sugar Beet	✓	Grown in Canada	✓	✓	Herbicide tolerant
5	Papaya	✓	Grown in the US and China	✗	✓	Virus resistant
6	Squash	✓	Grown in the US	✗	✓	Virus resistant
7	Cotton	✓	Grown in the US, India, China, and others	✗	✓	Insect resistant
8	Alfalfa	✓	Grown in the US	✗	Imported as animal feed	Herbicide tolerant Low lignin
9	Apple	✓	Approved March 20, 2015	?	?	Non-browning
10	Potato	✓	Not grown anywhere in the world	✗	✗	Insect resistant
11	Rice	✓	Not grown anywhere in the world	✗	✗	Herbicide tolerant
12	Flax	✓	Deregistered in Canada. Not grown anywhere in the world	✗	✗	Herbicide tolerant
13	Tomato	✓	Not grown anywhere in the world	✗	✗	Delayed ripening Insect resistant
14	Lentils	✗*				Herbicide tolerant
15	Sunflower	✗**				Herbicide tolerant
16	Wheat	✗*				Herbicide tolerant

Grown in Canada

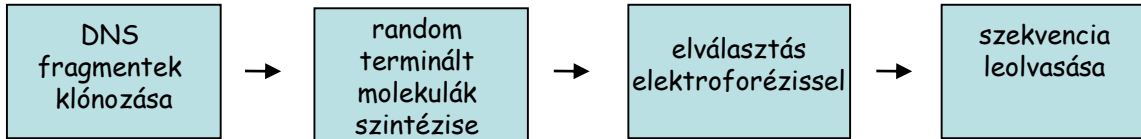
GMO On the market

* Product of chemically induced seed mutagenesis
** Product of conventional plant breeding

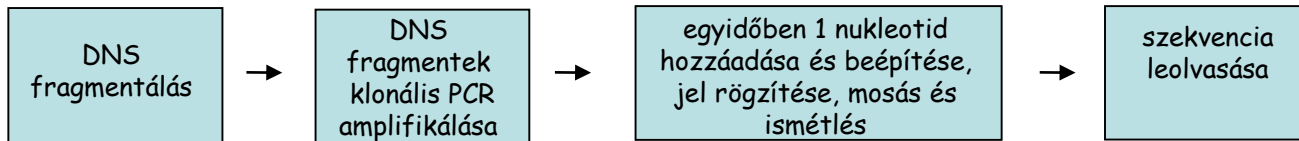
Genomszekvenálási módszerek

A klasszikus és a leggyakoribb nagy átteresztőképességű szekvenálási módszerek folyamatábrája

klasszikus Sanger-féle szekvenálás



szekvenálás szintézissel (454, Illumina, Solid, stb.)



Nagy átteresztőképességű vagy új generációs szekvenálási módszerek (milliónyi nukleotid leolvasás zajlik egy futtatás során; párhuzamos szekvenálás)

- Roche 454 technológia (Life Sciences)- amplifikáció víz-olaj emulzióban, < 450 bp leolvasási hossz
- Illumina (Solexa) - „bridge-amplification”, „reversible dye-terminator”; 70-80 bp leolvasási hossz
- SOLiD (Applied Biosystems) - ligálással, ~ 50 bp leolvasási hossz
- Ion Torrent Systems Inc. (Life Technologies, 2010 február) - hagyományos szekvenálás új, pH változáson alapuló detektálási módszerrel (a DNS polimerizáció során keletkező proton detektálása félvezető szenzorral), leolvasási hossz kb. 100 bp