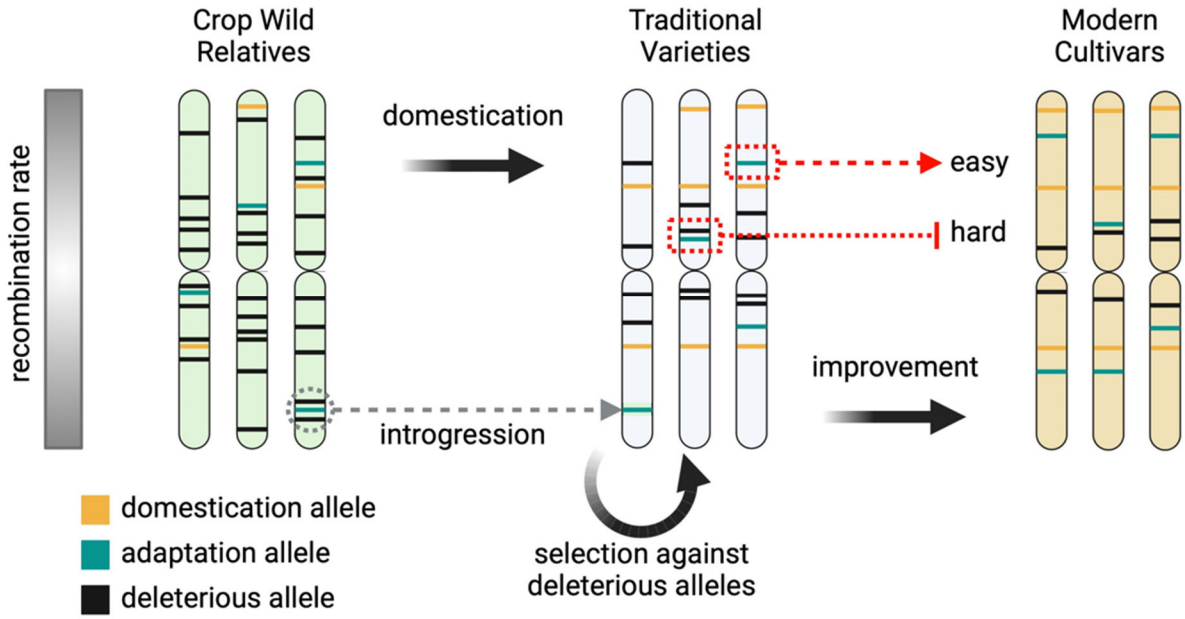
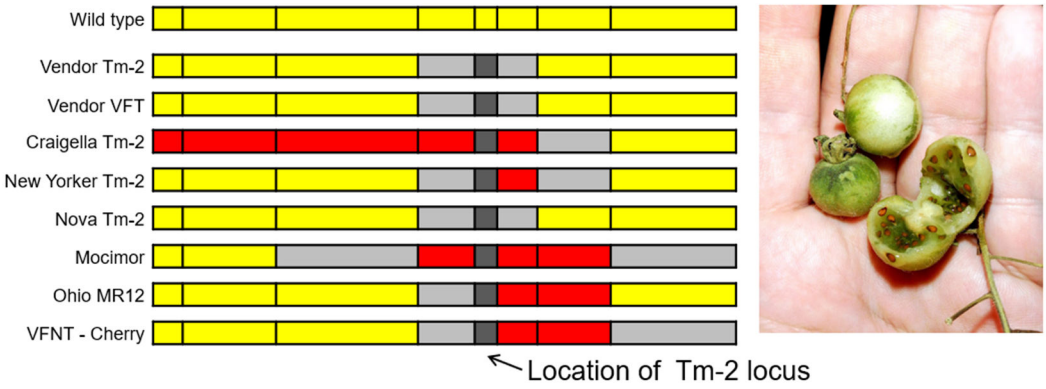
Breeding and linkage

Jeffrey Ross-Ibarra
@jrossibarra



The need to break linkages

Young and Tanksley, 1989



Red blocks illustrate DNA segments introgressed from the wild species donor of the Tm-2 locus. Lots of linkage drag is present. Targeted recombination could help avoid this linkage drag.

Significance of CO and its manipulation

Wijnker & de Jong, 2008

- Assemble new (un) desirable combinations
- Break old (un) desirable combinations

} balanced in the bivalent

- A low frequency of crossing over provides for stability. Favorable gene combinations are preserved, and help a population maintain its adaptation to a particular environment.

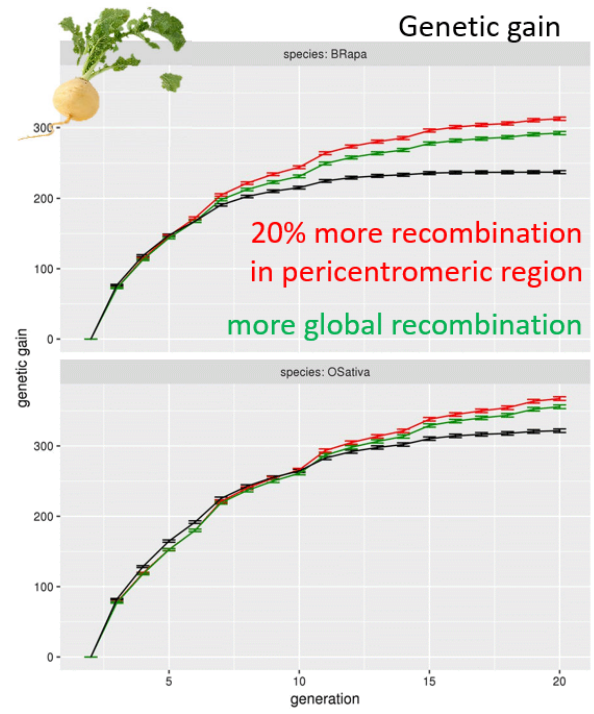
- A high frequency of recombination provides more flexibility by continuously creating new gene combinations, allowing a population to adapt to a new environment.

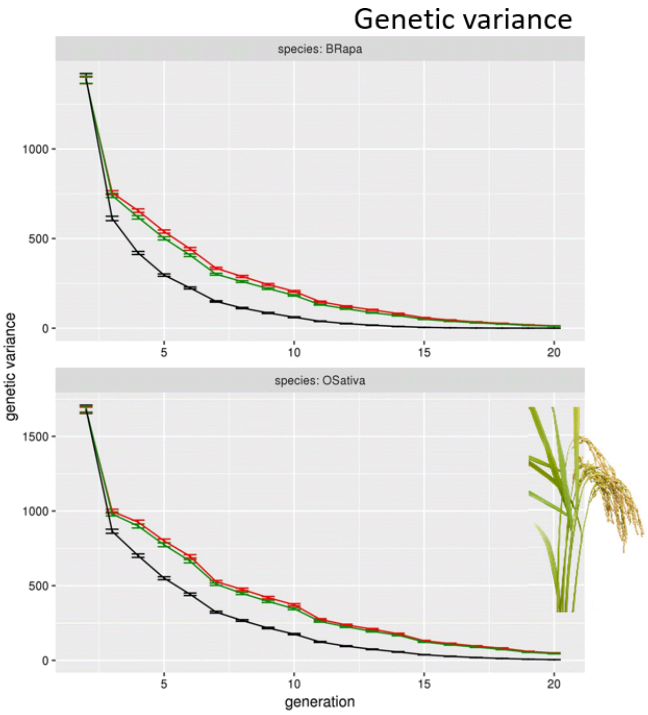
- Information on crossing over is useful to break/maintain linkages during a breeding program

Increased recombination

Tourette et al, 2019

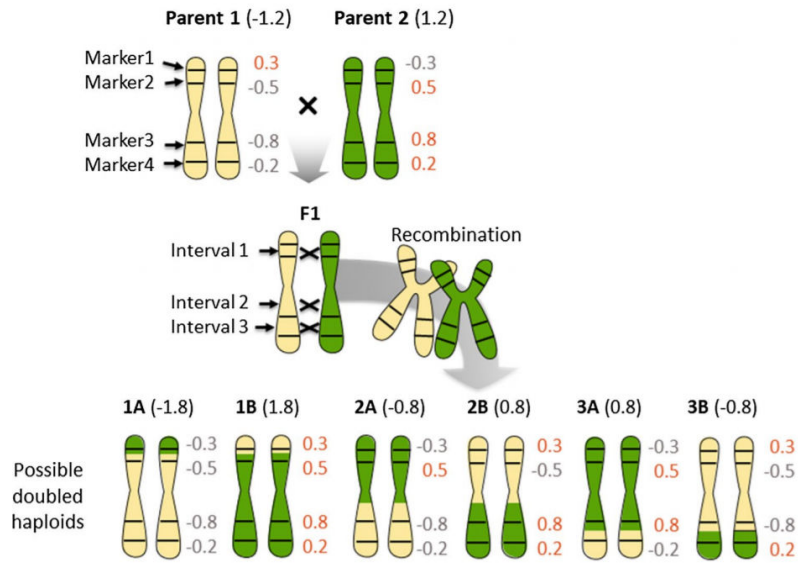
- Point is that increased recombination should permit more genetic gain while preserving diversity





Targeted recombination

Ru & Bernardo, 2019



In this example P1 has 1 favorable allele linked to 3 unfavorable alleles.

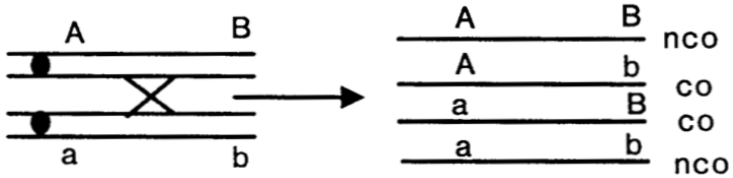
P2 is the opposite. The desired recombination would move all favorable alleles to one chromosome

RG(x#1) = 211% = gain from 1 targeted recombination

RG(x#2) = 243% = gain from 2 targeted recombinations

Recombination & genetic maps

Background information: When 100% single crossover occurs between a pair of genes, 50% of the gametes will be recombinant:



- The ratio of 1 crossover per ½ gamete (2 crossovers per recombinant gamete = 50% recombination rate) is a basic feature of recombination, and the basis for measuring the distance between genes.
- I.e., the frequency of recombinant gametes reflects one half the number of recombination events that occur. Recombination frequency = ½(crossover) frequency.
  - I.e., 2(recombination %) = CO

  - E.g., a 9.3% frequency of recombination comes from an 18.6% frequency of crossing over.
- Thus 50% normally defines the upper limit of recombination
  - 2 genes with 50% recombination are said to be unlinked
- 1% recombination = 1 map unit = 1cM. 50 map units = segment of chromosome where 1 CO occurs per meiosis
- Based on this, the total map length can be determined as the number of chiasmata × 50.

Figured out by Sturtevant, while he was an undergraduate student at Columbia University

Alfred Sturtevant

"In the latter part of 1911, in conversation with MORGAN . . ., I suddenly realized that the variations in strength of linkage, already attributed by MORGAN to differences in the spatial separation of the genes, offered the possibility of determining sequences in the linear dimension of a chromosome. I went home and spent most of the night (to the neglect of my undergraduate homework) in producing the first chromosome map."

Alfred Sturtevant,
http://www.nap.edu/readingroom/books/biomems/asturtevant.html

3-point test cross

Data from **Hutchison, 1922**

Parents: C/C sh/sh Wx/Wx ✕ c/c Sh/Sh wx/wx
(purple; shrunken; starchy) (colorless; full; waxy)

F_1 x tester: C/c sh/Sh Wx/wx ✕ c/c sh/sh wx/wx
(purple; full; starchy) (colorless; shrunken; waxy)

Figure 1. Nannas & Dawe, 2015

Measuring the distance

Here is the F2 data, and then sorted by class:

Kernel Phenotype	F_1 gamete genotype	Number
Purple, shrunken, starchy		2538
Purple, shrunken, waxy		601
Purple, full, waxy		116
Purple, full, starchy		4
colorless, full, waxy		2708
colorless, full, starchy		626
colorless, shrunken, starchy		113
colorless, shrunken, waxy		2
Total		**6708**

Kernel Phenotype	F_1 gamete genotype	Number
Purple, shrunken, starchy		2538
colorless, full, waxy		2708
Purple, shrunken, waxy		601
colorless, full, starchy		626
Purple, full, waxy		116
colorless, shrunken, starchy		113
Purple, full, starchy		4
colorless, shrunken, waxy		2
Total		**6708**

Next, determine genotype of F1 gametes for each gamete & assign categories. The least frequent class is the DCO

Kernel Phenotype	F_1 gamete genotype	Number	
Purple, shrunken, starchy	C – sh – Wx	2538	5246 Parental types
colorless, full, waxy	c – Sh – wx	2708	
Purple, shrunken, waxy	C – sh – Wx	601	1227 **SCO**: sh – Wx
colorless, full, starchy	c – Sh – Wx	626	
Purple, full, waxy	C – Sh – wx	116	229 **SCO**: c - Sh
colorless, shrunken, starchy	c – sh – Wx	113	
Purple, full, starchy	C – Sh – Wx	4	6 Double cross over
colorless, shrunken, waxy	c – sh – wx	2	
Total		**6708**	

Next, calculate gene-pair distances:

Parents: C/C sh/sh Wx/Wx × c/c Sh/Sh wx/wx
(purple; shrunken; starchy) (colorless; full; waxy)

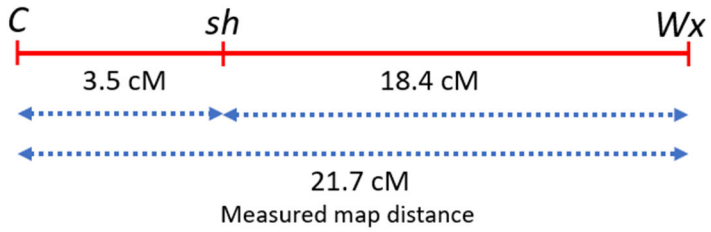
sh – Wx $\frac{1227+6}{6708} x\ 100 = 18.4\%$

C – sh $\frac{229+6}{6708} x\ 100 = 3.5\%$

C – Wx $\frac{229+1227}{6708} x\ 100 = 21.7\%$

It is now possible to determine gene order:

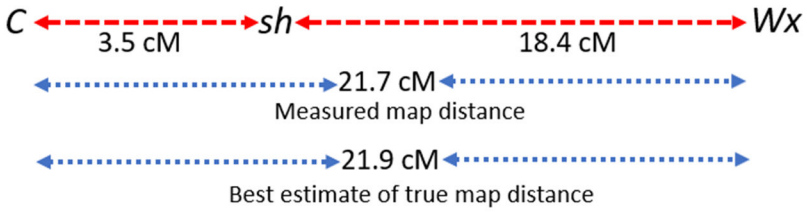
- Longest = furthest apart = C & W
- Means *sh* must be in the middle



Accounting for unseen DCOs

- There could be DCOs that are undetected, particularly in the longer Sh-Wx interval.

		Map distances
sh – Wx	$\frac{1227}{6708} x\ 100 = 18.291\%$	18.291 + 0.089 = 18.38%
C – sh	$\frac{229}{6708} x\ 100 = 3.414\%$	3.414 + 0.089 = 3.503%
C – Wx	$\frac{6}{6708} x\ 100 = 0.089\%$	18.291 + 3.414 + 0.089 + 0.089 = 21.883%



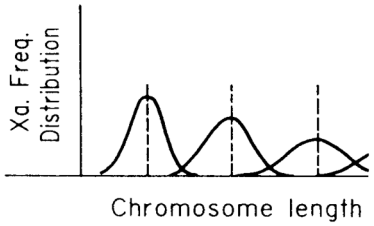
Interference

Sturtevant, 1913; Muller, 1916; McPeek and Speed, 1995

Chiasma, position, or chromosome interference:

- The number and position of crossovers are not independent of the numbers and locations of crossovers in other regions

- First observed by **Sturtevant,1913**
  -Called it crossover homeostasis as lengths more regular than expected by chance.
- **Muller 1916** [Amer Nat 50:193-434], coined 'interference'
  - **Mather, 1937**, termed the process as seriation:
- The process can be quantified as follows:



Interference = 1 - coefficient of coincidence, where

$$\frac{Observed}{Expected} = coefficient\ of\ coincidence$$

So back to the data of Hutchinson, 1922:

- Observed DCOs = 6
- Expected DCOs = 3.503% x 18.38% = 0.0059
  - 0.0059 x 6708 = 42.93

$$\frac{Observed}{Expected} = \frac{6}{42.93} = 0.139 = coefficient\ of\ coincidence$$

Interference = 1 - CC = 1 - 0.139 = 0.86 = 86%

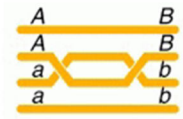
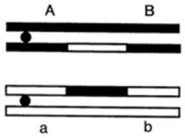
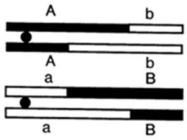
- Coincidence coefficients < 1 = positive interference, as 1 crossover decreases the occurrence of another.
- The magnitude of interference varies both within and between bivalents.
- Interference is not found in all species.
- Mapping functions factor in interference

2, 3, and 4-strand DCOs

Chromatid interference:

- A crossover between a pair of chromatids inhibits the occurrence of an adjacent crossover between this same pair of chromatids
- Does not happen often and is difficult to observe, requiring a pair of linked genes to detect it
- Expected chromatid crossovers in the absence of positive interference:

A 2-plane model used for clarity (not to endorse the 2 plane model!)

		Actual	Observed
2-strand DCO 		NCO	NCO
		DCO	
		DCO	
		NCO	
3-strand DCO 		NCO	50% CO
		SCO	
		DCO	
		SCO	
3-strand DCO 		SCO	50% CO
		DCO	
		SCO	
		NCO	
4-strand DCO 		SCO	100% CO
		SCO	
		SCO	
		SCO	

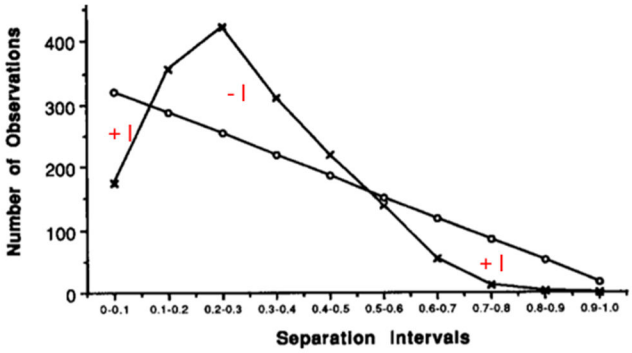
DCO graphics from http://www.mun.ca/biology/desmid/brian/BIOL2250/Week_Five/CXORF.jpg

- As long as the frequency of 2-strand DCO = that of 4-strand DCO, recombination = 50%
- If 2-strand DCOs are more frequent, it would give the appearance of less crossover
- If 4-strand DCOs are more frequent, it would
    - Give the appearance of negative chromatid interference
- It is the only time that cross over frequency will be greater than 50%
    - Coincidence coefficients > 1

Interference vs chromosome position

Sherman and Stack, 1995

The graph shows the predicted vs the observed double crossovers in the long arms of tomato chromosomes.



- Positive interference is occurring in the ≤ 0.1 and the 0.5 - 1.0 intervals (ie, below the line)
- and/or negative interference in the 0.1 - 0.5 interval (ie, above the line)
- means that 4-strand DCO's can happen in the middle of long chromosomes.

Xmata vs recombination maps

Sherman and Stack, 1995

In addition to other problems, note discrepancies in maps calculated in different ways. Data from tomato:

Chromosome	Length based on xma	Length based on classical map	Length based on molecular map
1	124.0	161	131.5
2	104.0	74	124.2
3	105.0	111	126.1
4	94.5	89	124.6
5	83.5	55	97.4
6	86.5	113	101.9
7	88.5	71	91.1
8	84.0	67	96.9
9	79.0	62	111.0
10	83.0	132	90.1
11	83.0	97	88.0
12	79.5	31	93.1
Total	1094.5	1063	1275.9

NOTE: Predicted maps based on Xma frequency are usually longer than classical maps, which do not have enough markers to detect all the CO that have taken place.

Darlington, 1934 Predicted that classical and xma maps would converge as more markers became available, permitting all DCOs to be detected.

Nilsson et al., 1993

Chrom-osome of maize	Genetic length (cM) estimated from					
	Chiasma Counts	Linkage data				
		1934	1950	1976	1990 classical	1990 molecular
1	187	102	156	161	176	238
2	163	58	128	155	155	229
3	150	92	121	128	167	194
4	148	80	111	143	137	174
5	148	44	72	87	107	235
6	110	52	64	68	78	169
7	123	50	96	112	112	131
8	123	20	28	28	42	173
9	110	52	71	138	140	132
10	98	68	57	99	95	115

Note that the molecular map is longer than predicted. This has happened for other crops as well.

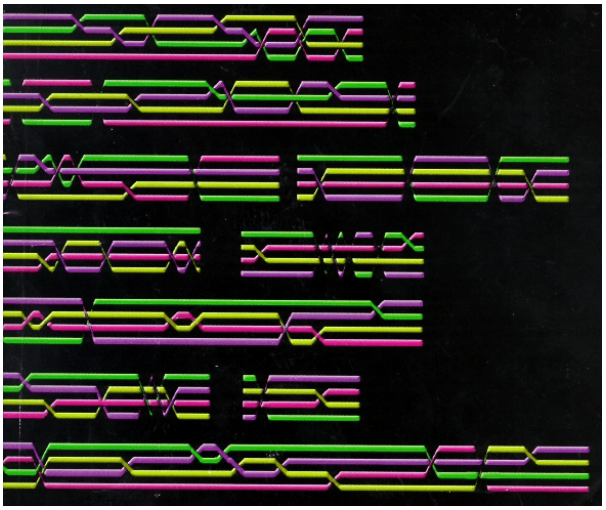
Why are molecular maps longer than chiasma maps?

Species	2n	Estimated CO/meiosis	
		xma counts	molecular maps
Brassica campestris	20	10.0-18.5	37.0
Brassica oleracea	18	12.8-14.8	22.2
Hordeum vulgare	14	13.5-15.6	22.7
Lactuca sativa	18	14.6-20.7	28.1
Lycopersicon esculentum	24	16.2-17.0	25.5
Oryza sativa	24	18.9-27.6	36.7
Pisum sativum	14	10.3-18.1	29.3
Solanum tuberosum	12/24	13.2-14.1	20.7
Zea mays	20	17.4-25.0	35.8

- Xma counts based on % meiosis only, while molecular maps based on both % and & meiosis
- In tomato, & map is 18% longer than % map, so the combined map should be 9% longer, yet it is 30% longer

- Errors in counting xma, scoring markers
- Sampling of different genotype
- Low number of meioses sampled
- Mapping functions assume undetected double COs exist, and lengthen map distances accordingly
- Mapping functions were designed for species with multiple crossovers, such as the grasshopper at right, or yeast, below

CO's assumed in mapping functions



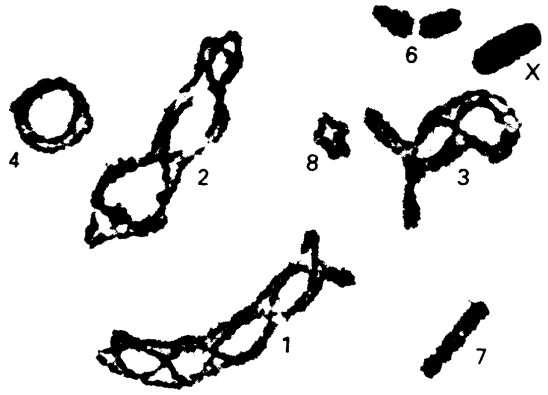
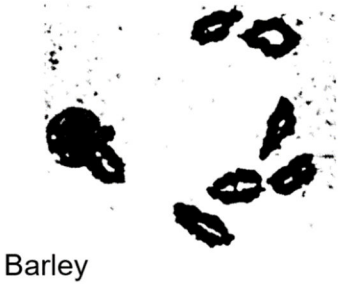
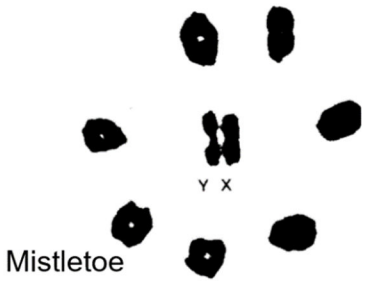


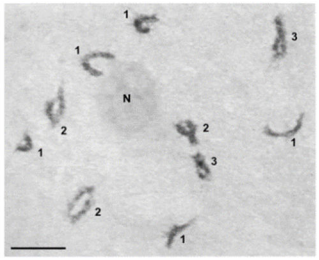
Figure 2. Diplotene in male Meadow grasshopper, Chorthippus parallelus

Figure 3. Crossovers in yeast, as determined by allele detection on a DNA chip

How common are DCOs in plants?



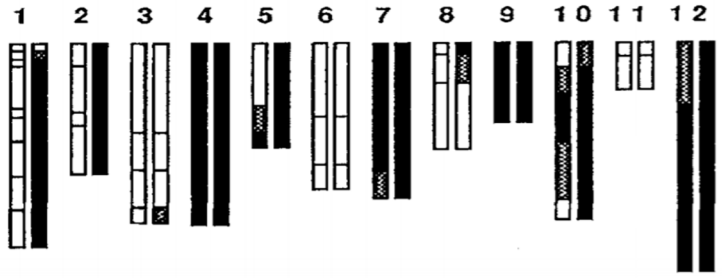




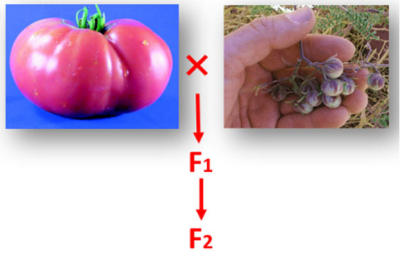
Maize- Anderson et al. 2003

- Based on above photos, the maximum appears to be ~ 1 CO per arm
- Note that various xmta are present in the grasshopper chromosomes, while mostly terminal ones are present in the plant chromosomes.
- Corroborated by molecular F2 maps:

Tanksley et al., 1989



Figu



Sherman and Stack, 1995

Counted xmta in 5228 paired chromosomes of tomato

- 40% had 1 CO per chromosome
- 49% had 2 CO per chromosome, ie, 1 per arm
- 10% had 3 CO per chromosome, ie, 1 arm with 1 CO, and 1 arm with 2 CO
- 1% had 4 CO per chromosome, ie, 2 per arm

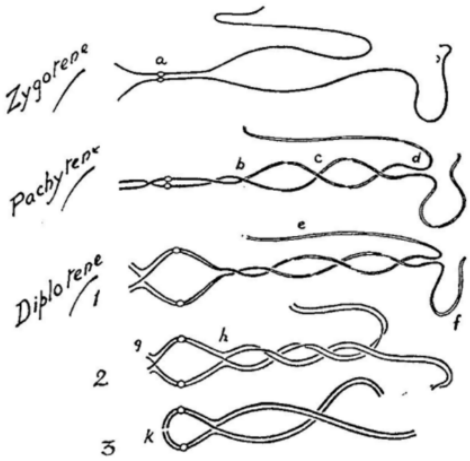
} DCO = 11%

Terminalization

Darlington 1935

In the 1930's, Darlington decided multiple CO's were present but got terminalized in order to explain discrepancy.

He was wrong.

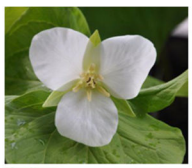


Localized Chiasmata

Darlington was looking at Localized chiasmata

- In some species, chiasmata occur preferentially in one part of the chromosome, such as next to the centromere, as in the examples below (Table from John, 1990):

PMC of *Paeonia lutea*, 2n = 10 with distal chiasmata. John, 1990



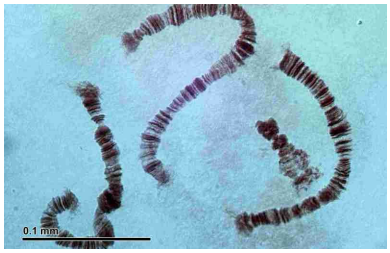
Species:	Chiasmata Site:			% Proximal Site
	Proximal	Interstitial	Distal	
Trillium kamtschaticum				
Clone 1	466	5	2	98
Clone 2	407	33	20	89
Allium fistulosum	2056	41	40	96

- Most frequently in crops, localized chiasmata occur at the end distal to the centromere, as in peony example.

Recombination is variable

Data from Bridges, from the Drosophila sex chromosome
Order of genes = y - pr - w - rst - fc – centromere

Gene pair:	# Bands:	Map distance:	%CO / band
y – pr	57	0.8	0.014
pr – w	18	0.7	0.038
w – rst	2	0.2	0.100
rst – fc	2	1.3	0.650

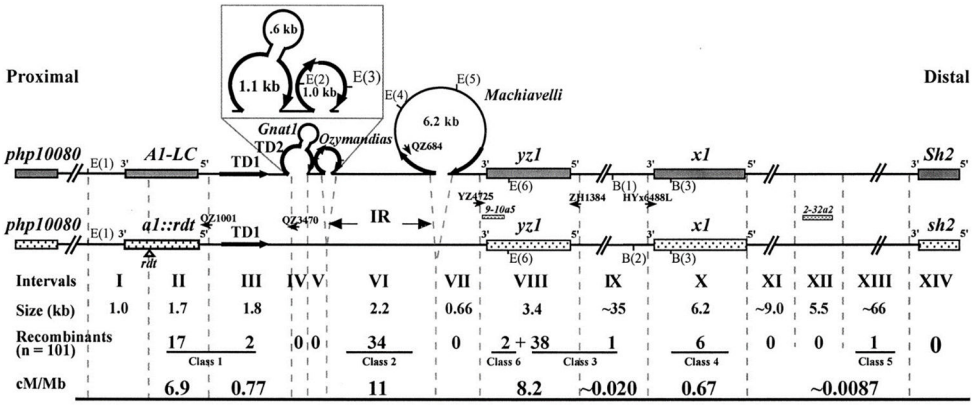


Recombination hotspots

The point is that some spots on the chromosome are particularly prone to recombination, including within genes.

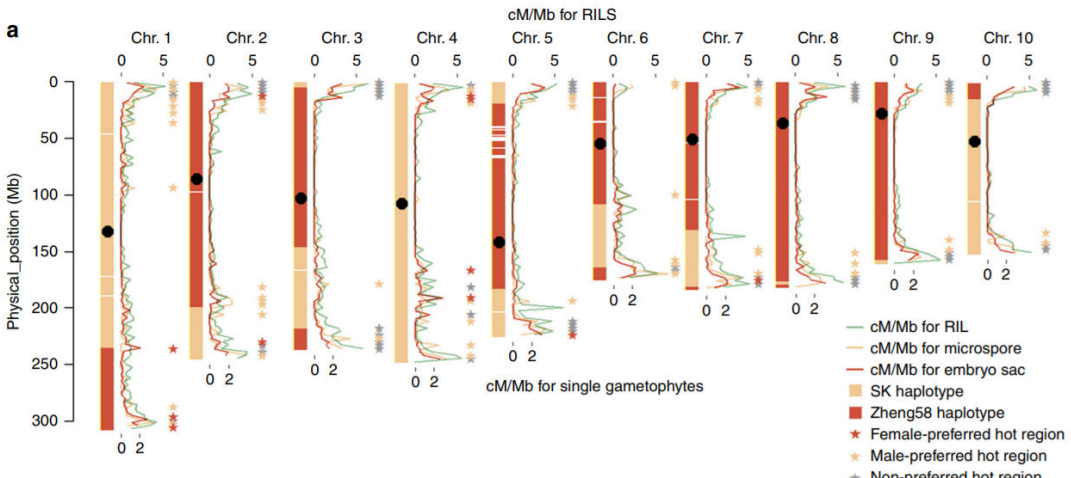
Yao et al., 2002

Example across 5 loci



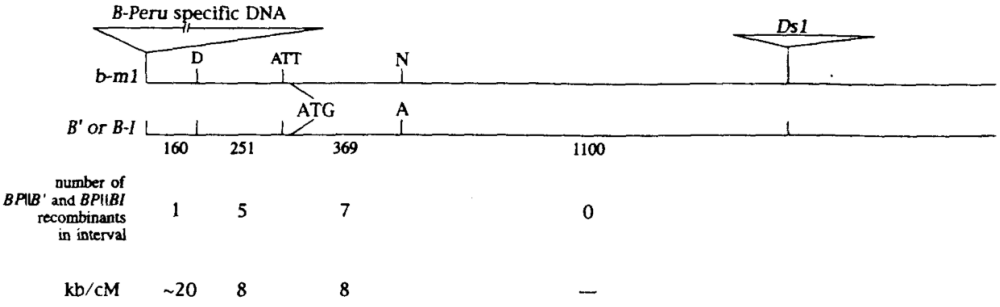
Luo et al., 2019

Hot spots are in the gene-rich ends of the chromosome, and can vary between % and &



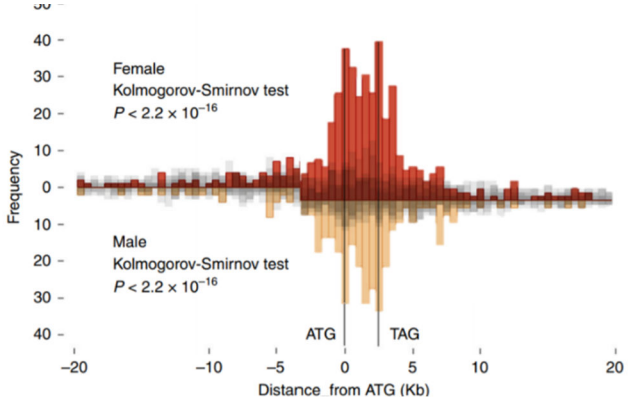
Recombination can vary within a gene, with hotspot near ATG

Patterson et al., 1995



Luo et al., 2019

This observation has since been generalized, and there is a tendency of hotspots to be near start and stop codons:

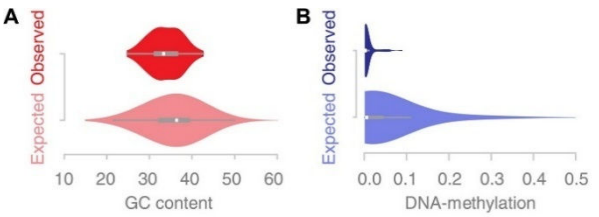


Preferred recombination spots

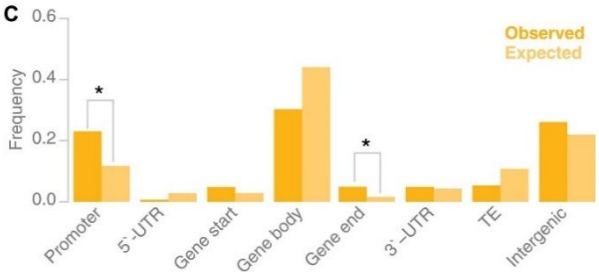
Wijnker et al., 2023

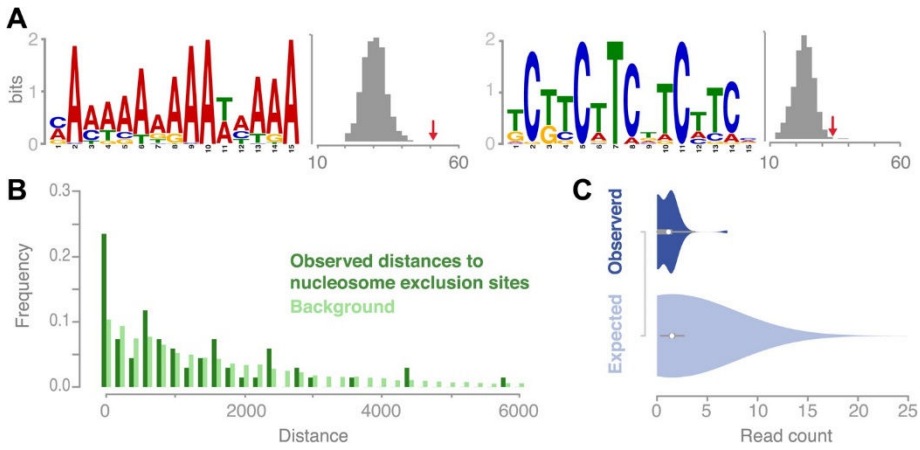
Recombination hot spots do not occur at random. Preferred sites have

- GC content 30-50%
- Undermethylated
- Promoters/terminators



As shown below, there is also a preferred sequence, and they are associated with nucleosome exclusion sites.



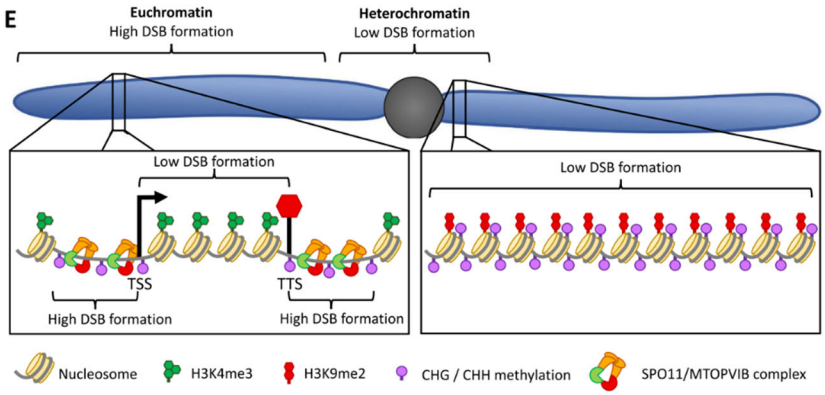


Nucleosome exclusion sites are defined as $(A)_{10}$ and $((GC)_3NN)_3$

Model of CO locations

Lloyd, 2022

"Crossovers are primarily located in open chromatin associated with gene promoters and terminators with low nucleosome occupancy"

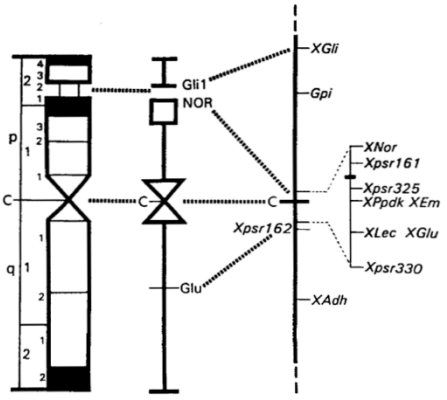


Yelina et al, 2015. CG methylation suppresses recombination. Is it methylation as such, or the conformational changes it induces - ie, heterochromatin?

Physical vs genetic maps

Because of all interferences with crossing over, physical maps do not equal linkage maps:

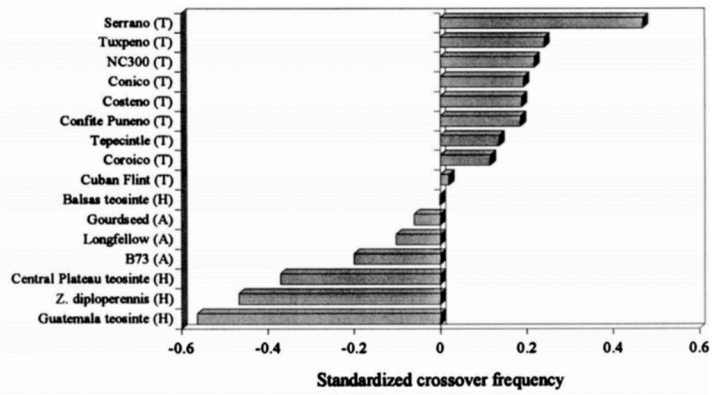
- Again, note that favorable alleles cannot be recombined together in the areas with no recombination
- I.e., these areas are not amenable to breeding

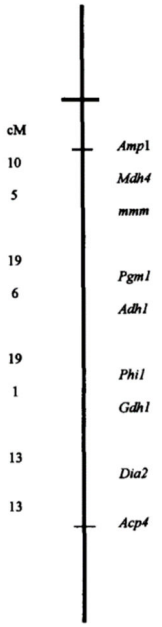


Heslop-Harrison, 1991

Genotype

Williams, Goodman, and Stuber, 1995





Studied chromosome 1L of maize

- Looked at individual maps vs composite maps
- Found wide differences in CO frequency

- Notice that composite maps are derived from formulae which assume no changes in recombination frequency
- Composite maps are inaccurate for estimating target gene proximity in cloning projects
- Supposedly, 1 cM in maize = 1460 kb.
  - With this ratio *amp1* to *mdh4* distance would range from 8720 to 24,800 kb
- Similar genotypic effects also seen in other species:
  - Arabidopsis (**Sanchez-Moran et al, 2004**)
  - Rye (**Rees, 1961**)
  - Barley (**Säll, 1990**)

The adjacent table compares this in more detail

- The genotype effect probably includes some effects from all other factors.

MAP	*amp1* - *mdh4*	*mdh4* - *pgm1*	*pgm1* - *phi1*	*phi1* - *dia2*	*dia2* - *acp4*	*Total cM ×1000*
Wendel (1989)	**15.0**	**19.0**	**25.0**	**14.0**	**13.0**	**86.0**
Composite	**9.7**	**15.9**	**17.4**	**14.4**	**12.7**	**70.2**
Composite US genotypes	**12.4**	**19.3**	**15.0**	**9.3**	**10.5**	**66.5**
B73	8.5	16.2	11.7	11.6	9.4	57.4
Gourdseed	6.9	17.5	19.2	11.1	12.3	66.9
Composite Exotic genotypes	**11.8**	**19.1**	**17.6**	**15.7**	**12.9**	**77.2**
NC300	12.1	17.4	18.9	14.7	15.0	78.1
Serrano	13.0	25.1	20.6	22.8	13.0	94.5
Tepecintle	11.0	17.9	12.6	20.9	12.3	74.7
Tuxpeño	12.5	21.2	21.6	8.0	15.1	78.3
Confite Puneño	16.7	23.0	18.7	7.1	11.5	76.9
Cónico	6.3	17.7	11.1	19.1	20.4	74.6
Coroico	15.4	13.2	15.6	19.9	10.8	74.9
Costeño	10.4	24.0	22.2	15.0	9.0	80.6
Cuban Flint	8.8	13.9	17.0	17.8	12.3	69.6
Composite High	**12.5**	**21.2**	**20.3**	**14.9**	**14.3**	**83.2**
Composite maize × teosinte	**4.8**	**6.3**	**18.6**	**12.7**	**13.5**	**56.0**
Balsas teosinte	14.0	9.3	24.5	12.2	9.8	69.8
Zea diploperennis	0	3.9	11.8	15.2	18.1	49.0
Central Plateau teosinte	0	5.6	20.7	10.6	12.7	49.7

Crossover sites can be altered

Jones, 1967

Crossed *Secale dighoricum* with *Secale turkestanicum*, and got an F2 segregant with an altered pattern of chiasma formation:

Chiasmata	Parental			F_2	
Distribution:	Observed	Expected		Observed	Expected
Distal	546	192		199	134
Interstitial	30	192		132	134
Proximal	0	192		71	134
Total	576			402	

- Rye normally has one crossover per chromosome arm, and it is localized distally to the centromere
- Segregants varied for location of crossover (notice random number for interstitial region), multiple crossovers, and non-serial crossovers.

Factors Affecting Recombination

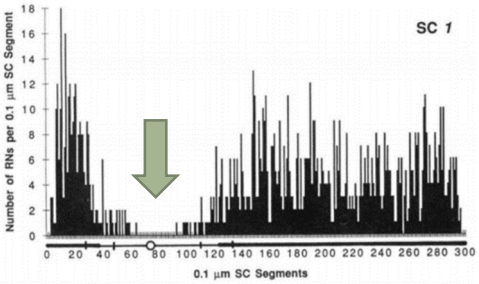
Landmarks/Chromosome Structure

a. The centromere

Sherman and Stack, 1995

The centromere suppresses crossing over in the immediate vicinity

- Found only 7 out of 9562 crossovers occurred within 10% of the physical distance from the centromere. An example for chromosome 1 is shown



Location of COs in chromosome 1 of tomato

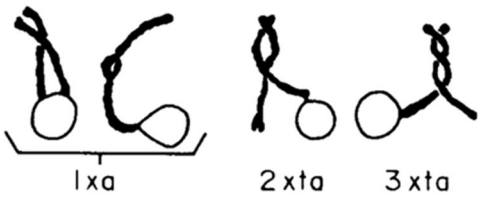
b. Heterochromatin

Crossing over in euchromatin and heterochromatin of tomato (Sherman & Stack, 1995):

		Euchromatin			Heterochromatin		
C-some	Length (µm)	Length (µm)	Ave no. RNs / II	RNs / µm	Length (µm)	Ave no. RNs / II	RNs / µm
1	30.0	22.5	2.44	0.112	7.5	0.04	0.005
2	21.3	17.1	2.05	0.122	4.2	0.03	0.007
3	23.1	16.0	2.07	0.131	7.1	0.03	0.004
4	20.8	13.7	1.87	0.138	7.1	0.02	0.003
5	16.2	9.5	1.63	0.168	6.7	0.04	0.006
6	18.5	12.9	1.69	0.132	5.6	0.04	0.007
7	18.5	11.6	1.75	0.155	6.9	0.02	0.003
8	18.5	11.9	1.66	0.143	6.6	0.02	0.003
9	16.2	10.0	1.58	0.160	6.2	0.004	0.0006
10	16.2	10.0	1.64	0.160	6.2	0.02	0.003
11	16.2	9.7	1.63	0.165	6.5	0.03	0.005
12	14.0	8.2	1.54	0.183	5.8	0.05	0.009

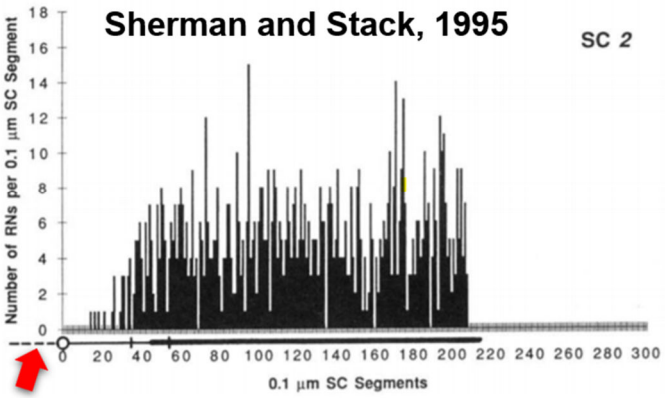
Heterochromatin suppresses crossing over. Crossing over rarely occurs in heterochromatin (i.e., repetitive DNA) itself.

c. The NOR



COs in *Eremurus spectabilis*

Upcott, 1936 via **John, 1990**

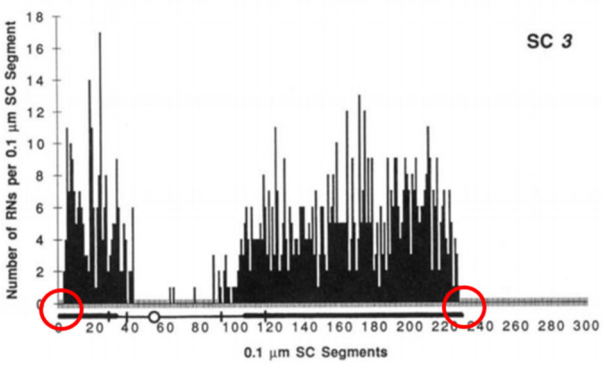


The NOR also suppresses crossing over in its vicinity.
The figure above shows this phenomenon in *Eremurus spectabilis* (after Upcott, 1936)

- Note: This phenomenon has not been found to be universal– eg, Anderson et al., 2003
- The short arm of chromosome 2 (red arrow above) does not pair, and if it pairs, it does not recombine.
    - It is where the NOR is, so there might be a causal relationship.

d. Telomeres

Sherman and Stack, 1995; Barton et al, 2008

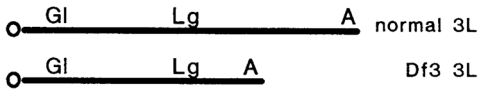


- Subtelomeric heterochromatin suppresses recombination
- Adjacent euchromatin recombines at 2x overall rate
- Also note telomere in tomato above, showing no CO in the telomere of the long arm of chromosome 2

e. Interchromosomal effects

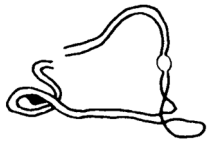
Rhoades, 1958

- Effects of insertions/deletions



Cross	*Gl – lg*	*Lg – A*
N3L / N3L	28.0	30.6
N3L / N3L*	29.0	28.2
N3L / D3L	35.1	12.8

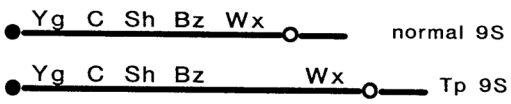
Marcus Rhoades
1903 - 1991



An indel during pairing

Notice that recombination is increased in the Gl-Lg region, but decreased in the Lg-A region.

- * In this case, a duplication in chromosome 9 elsewhere in the cell affects crossing over in chromosome 3!



Poor pairing in the heterozygote reduces recombination across the chromosome

Cross	*Yg – Sh*	*Sh – Wx*
N9 / N9	21.0	17.0
Tp9 / N9	2.0	2.0
Tp9 / Tp9	29.0	18.0

Also, cryptic structural differentiation (CSD) is present.

- Deters introgression and increases linkage drag
- Today called structural variation

f. Chromosome length

Rees & Durant, 1986; Sherman & Stack, 1995; Anderson et al., 2003

Table below: rows show the within-species effect, depending on chromosomal DNA content, or in other words, length: Longer chromosomes have more DCOs

- The following data are from tomato (***Sherman and Stack, 1995***):

Chromosome	Mean Length	CO/Chromosome	Unit Length per CO
1	30.0	2.48	12.10
2	21.3	2.08	10.24
3	23.1	2.10	11.00
4	20.8	1.89	11.00
5	16.2	1.67	9.70
6	18.5	1.73	10.69
7	18.5	1.77	10.45
8	18.5	1.68	11.01
9	16.2	1.58	10.25
10	16.2	1.66	9.75
11	16.2	1.66	9.76
12	14.0	1.59	8.80

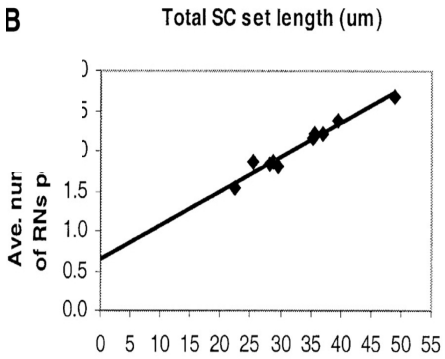


Figure 5. Anderson et al, 2003

- Longer chromosomes have more CO, but not enough to compensate for the longer length.
- Therefore, CO/bp of DNA must vary between chromosomes of different lengths
  - With longer chromosomes having less COs per unit length.
- Within a species, chromosome length explains 96% of differences in CO
  - This is shown in the graph at above right, from Anderson et al., 2003.

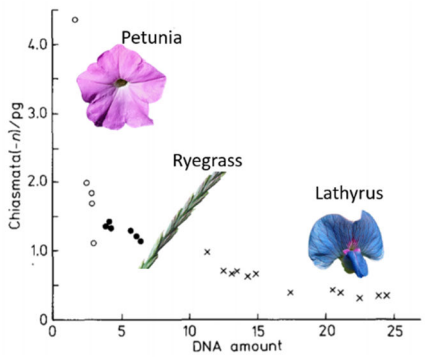


Table from John, 1990: Species	2C DNA (pg)	DNA content (pg) of II with means of:		
		1.5 xta	2 xta	3 xta
Lathyrus clymenum	13.43	1.0	1.4	2.2
L. cicero	14.64	1.3	1.8	2.8
L. sativus	16.78	2.1	2.7	3.8
L. tinitanus	22.08	2.4	3.1	4.4
Lolium perenne	4.16	0.36	0.8	--
L. temulentum	6.23	0.51	1.2	--

Rees & Durrant, 1986

- In general, as DNA content increases, the CO/pg decreases
- Minimum CO is always 1 (unless achiasmate meiosis is involved, as in chromosome 4 of Drosophila)

But, while the longest chromosome has the most COs, the increased CO in longer chromosomes is not enough to compensate for longer lengths

- Therefore: Long chromosomes increase linkage within a species
- This holds across species

g. Arm length

Sherman and Stack, 1995

In subacrocentrics, COs occur preferentially on long arms

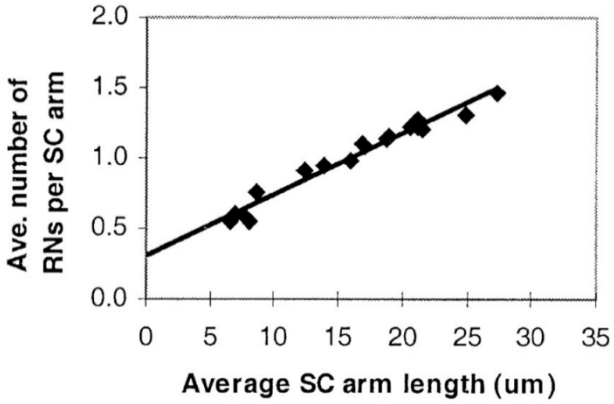
- If there is only 1 CO, then it is always in the long arm
- Eg, chromosome 1 of tomato. 1S is 17% of length of 1L
    - Thus, expect single CO to occur on 1S 17% of time ➞ not the case

Submetacentrics: Single COs were sometimes on the short arm, but not proportionately

Metacentrics: There is preference for one arm over the other

- If a second crossover occurs, 2nd CO takes place in 2nd arm more often than is expected based on length.
- Occurrence of a CO in an arm drives the 2nd one into the other arm
- Synapsis always begins in long arms or given arm of metacentrics

Bottom line: short arms have tighter linkage

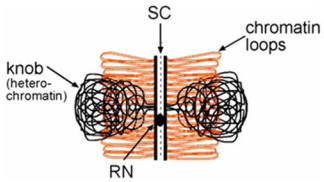


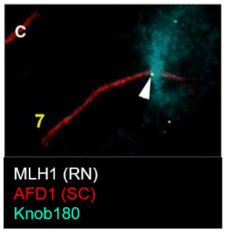
Anderson et al., 2003. The effects of chromosome arm length on xmata in maize

h. Knobs

Naranjo and Lacadena, 1980; Stack et al., 2017

- Durum wheat × rye: AABB × RR ➞ number of xma on 1R depends on telomeric heterochromatin
- Note: this phenomenon is not universal (eg, **Anderson et al., 2003** in maize. See eg from **Stack et al., 2017** below, who looked at recombination nodules that formed in a knob)



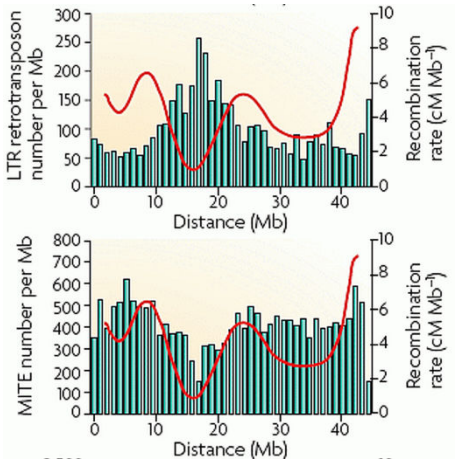


SC	SCs obs	No. MLH1 foci			No. of MLH1 foci per µm of SC length		
		SC	Near knob	In knob	SC	Near knob	In knob
5	89	139	7	4	0.042	0.039	0.026
6	72	96	8	3	0.043	0.056	0.052
7	94	132	10	9	0.046	0.053	0.056
Total	255	367	25	16			

i. MITEs (skip)

Gaut et al., 2007

NOTE- since mites insert preferentially into euchromatin, this effect may reflect recombination in euchromatin vs heterochromatin than an actual MITE effect per se.



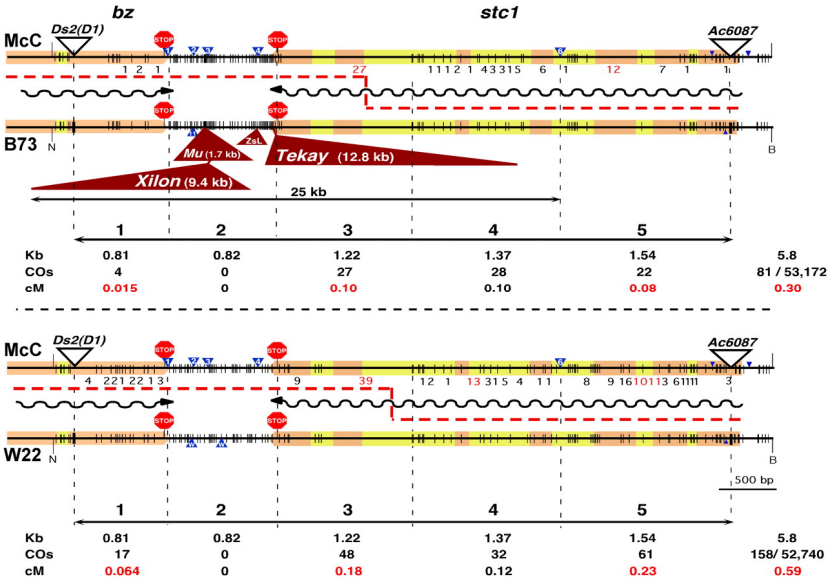
j. Retrotransposons

Dooner & He, 2008

Key thing to notice is the cM distance in the 5 intervals encompassing bz and stc1 differ in McC x B73 crosses compared to McC x W22.

Heterozygous retrotransposon insertions halves the genetic distance between bz1 & stc1

bz1 & stc1 are at half the genetic distance in B73, due to reduced CO



Basically, an indel effect (ie, intrachromosomal effect)

k. Gene density

Fengler et al., 2007

In general, recombination outside of the centromeric area, is correlated with gene density.

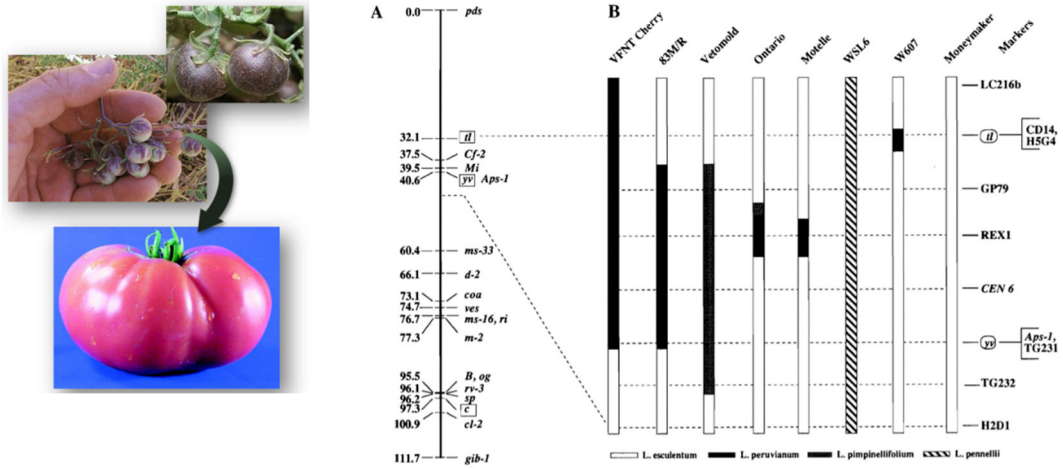
Maize Chromosome	Genic sequences	Genetic length (cM)	Correlation between genes and recombination
1	3357	1137.9	0.96
2	2619	725.3	0.95
3	2478	829.9	0.90
4	2286	750.2	0.83
5	2618	676.7	0.93
6	1721	548.7	0.92
7	1757	618.4	0.85
8	1970	632.0	0.77
9	1608	638.7	0.93
10	1432	533.7	0.78
total/Ave	21,846	7090.0	0.87

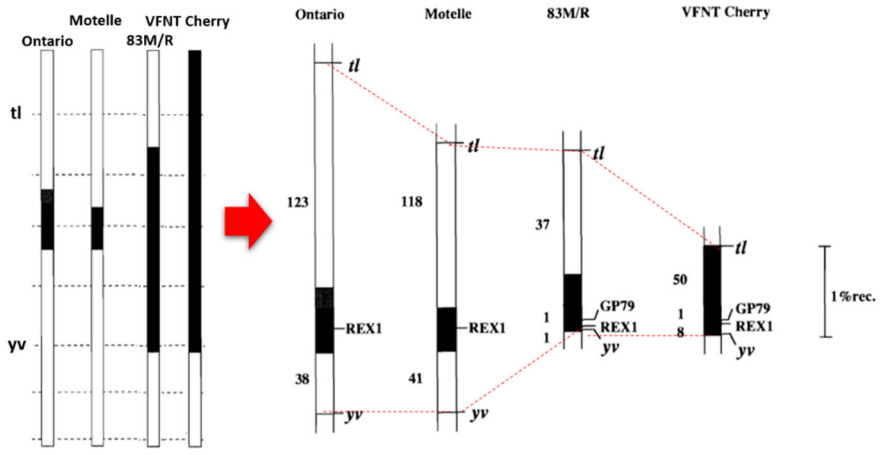
I. Alien introgressions

Liharska et al., 1996

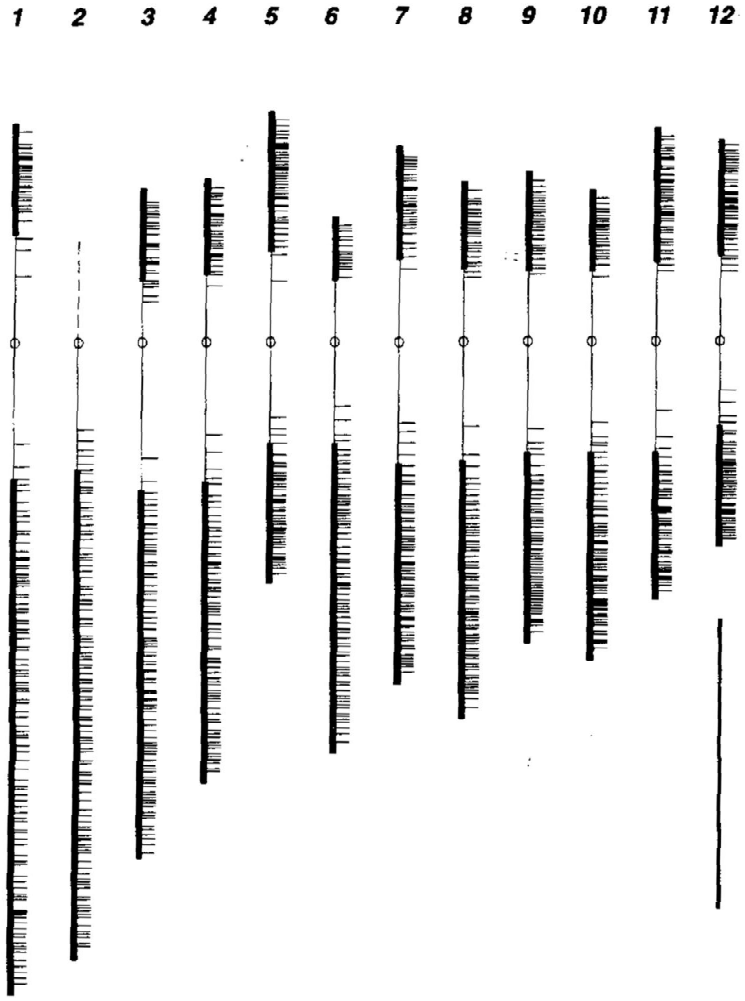
Looked at recombination around Mi gene introgression between tl and yv loci.

- The largest introgressed segments had 6x lower recombination





Effect of CO variability on linkage maps



Each pair of horizontal lines represents 1 cM

Basic take-home is that the physical distance represented by 1 cM can vary widely

Sherman & Stack, 1995

Demarly, 1979

Concept that the unit of inheritance in plants is not a single gene

Instead it is a block of linked genes, he called a linkat

Nowadays called a haplotype

Factors Affecting Recombination

Genetics

A. Gender - Heterochiasmy

John 1991; review by Girault et al, 2011

Crossing over is absent in % Drosophila, & silkmoth (ie, XO, so no pairing partner!)

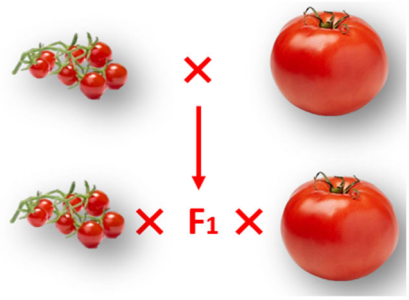
- The heterogametic sex tends to have less crossing over
- Differences in crossover frequency have been found between the eggs and the pollen from the same plants of *Listeria*, *Pisum*, *Primula*, *Zea* & *Arabidopsis*

	Xma freq.		Dist'n	
	♀	♂	♀	♂
Fritillaria martagon	41.0	36.3		
F. meleagris	37.8	24.8		
F. longiflorum	31.4	27.3		
Allium nigrum	16.9	21.9	R*	R
A. consanguineum	17.5	21.9	R	R
A. cepa	17.9	22.4	R	R
A. kachrooi	15.0	12.9	P*	R

* R = random; P = proximal

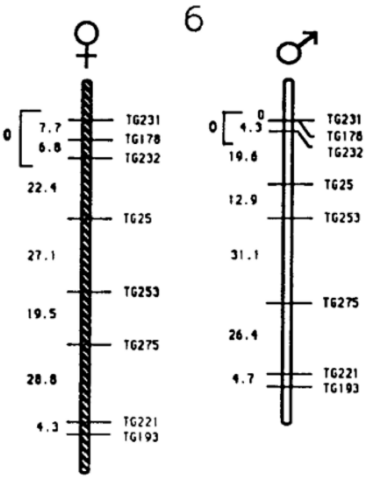
de Vicente & Tanksley, 1991

The first to actually document heterochiasmy in plants, using RFLP markers.

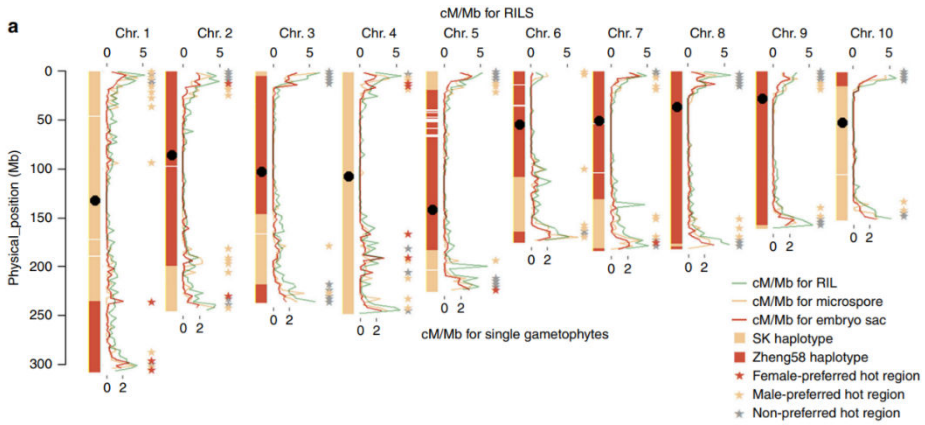


- Map distance reflects recombination in each parent

- The % parent always had less recombination for all chromosomes (0[= centromere)



Li et al, 2019

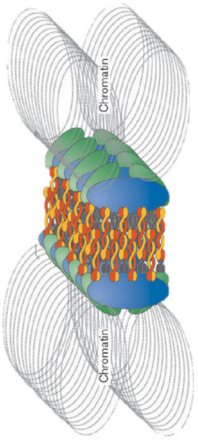


Single cell maize gametophyte sequencing
Male = 19.3 CO/µspore
Female = 12.4 CO/megaspore

Zickler & Kleckner, 1999; Koul & Nagel, 2002; Giraut et al, 2011; Lloyd, 2022

- "Thus, along euchromatic portions of genomes, loop lengths (in kb) may be relatively constant."
- "25–30 loops per chromatid per µm of chromosome axis."
    - Varies 2x

1. "Changing loop size changes length of synaptonemal complex"



Popa. 2011

DNA loops on bumblebee pachytene chromosomes. Zickler & Kleckner, 1999

Koul and Nagel, 2002:

- most legumes: increased xmata in male
- most grasses: increased xmata in female
    - these also differed in their distribution pattern
- not all spp showed difference
    - eg, fava beans, beans, and rye

Säll & Nilsson, 1994

Found no sex differences in barley:

♂ = 10.8 ± 10.7

♀ = 12.1 ± 13.8 ns

So the phenomenon is not universal

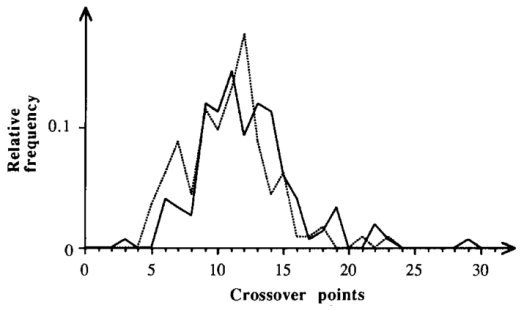


Figure 6. Comparison of CO in ♀ (dotted) and ♂ (solid) meiosis of barley

B. B chromosomes

John, 1990

Tp9/Tp9	*Yg-C*	*C-Wx*
+ 0 B	28.8	17.7
+ 1 B	12-13	37.0
+ 2 B	12-13	40.4
+ 3 B	12-13	42.0

- B chromosomes also influenced the location of the crossovers
- B chromosomes had no effect in N9/N9 stocks
- Wheat × *Aegilops speltoides* or *mutica* ⟶ B chromosome from *Aegilops* can normalize pairing between homologues, similar to Ph1 effect

- Same is true for *Lolium-Festuca* hybrids for *Lolium*-derived B chromosome, and wheat × rye

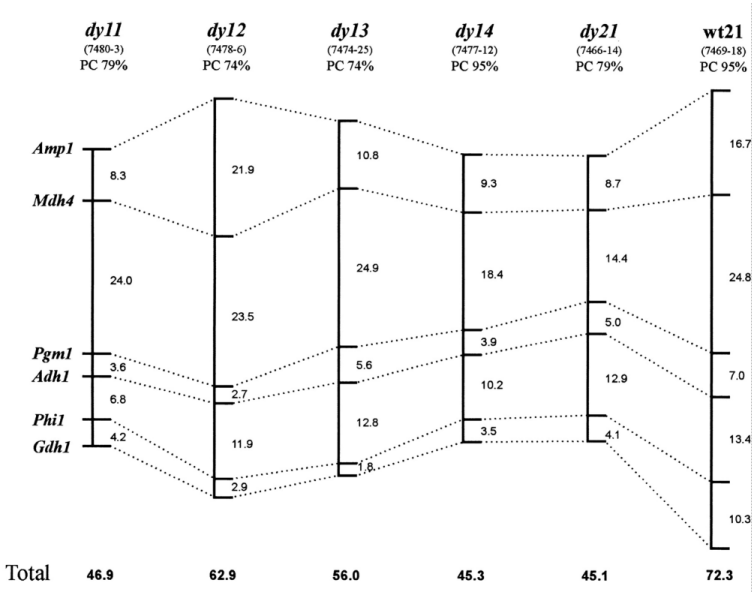
	Mean chiasmata/cell									
Species:	0B	1B	2B	3B	4B	5B	6B	7B	8B	10B
Crepis capillaris	↑ 4.08	3.90	4.65	4.68	5.33					
Puschkinia libanotica	↑ 9.07	10.82	11.70	11.62	13.80					
Lolium perenne	↓ 11.93	11.21	10.00							
Festuca mairei	↑ 18.87	20.95	22.85	26.25	28.00					
Zea mays	↑↓ 18.50		19.50	19.70	19.80		22.53		23.80	19.80
Secale cereale -inbred	↓ 14.85	13.43	13.22	13.29	12.64	12.86	13.20	12.64	13.03	
- wild	↑ 13.20	15.12	16.57	17.78	18.40					
Listeria ovata - PMC	↑ 26.9	28.9	28.2	30.3	29.1					
- EMC	↑ 30.3	32.6	31.3	32.5	32.7					

C. Desynaptic mutants

Ji et al., 1999

Desynaptic mutants reduce recombination. Not surprising as homologs separate before completion of the COs

-Looking at progeny of *dy* plants, found one homozygote (dy14) where recombination was reduced, without any effect on chiasma formation.



D. Zygosity

Robbins et al., 1995

Case I: homozygosity increases CO

-Eg, petunia, (n=7)

- 135 petunia wide-hybrid transgenics - 47 T-DNAs linked to *Hfl* or *Fl*
- 19 were 1 cM from *Hfl* or *Fl*
  - nonrandom insertions? ➞ no evidence in any species
  - lack of recombination in wide hybrids?
- To distinguish, measured recombination in inbreds
  - recombination around Hfl increased 3x
  - recombination around Fl increased 12 x
- same is true in wheat

Rees & Thompson, 1956

Case II: homozygosity decreases CO

… rye → reduced chiasma frequency (right)
RI pea map shorter than F2 pea map (below)

Note that pea, petunia and wheat are normally selfers, while rye is normally an out crosser.

Note that the zygosity effect appears to be chromosome-specific in some species

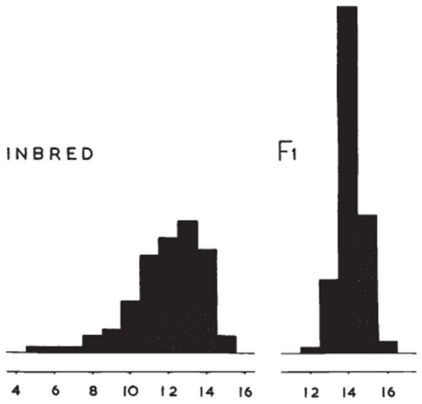
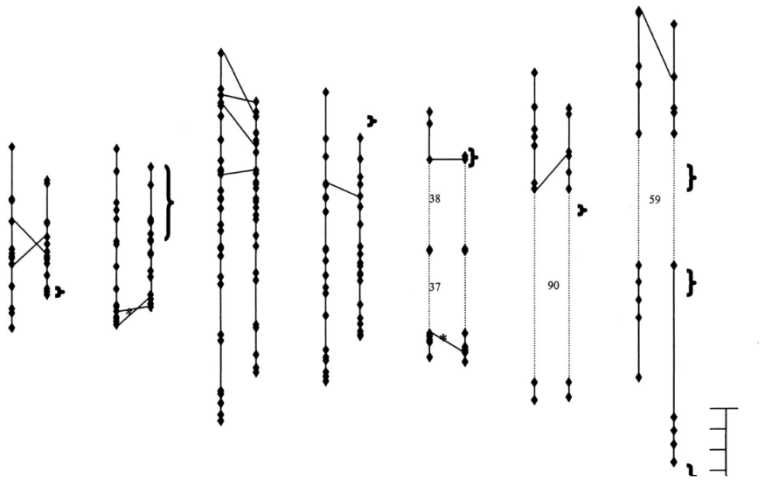


Figure 7. Chiasmata/cell in rye

Knox & Ellis, 2002

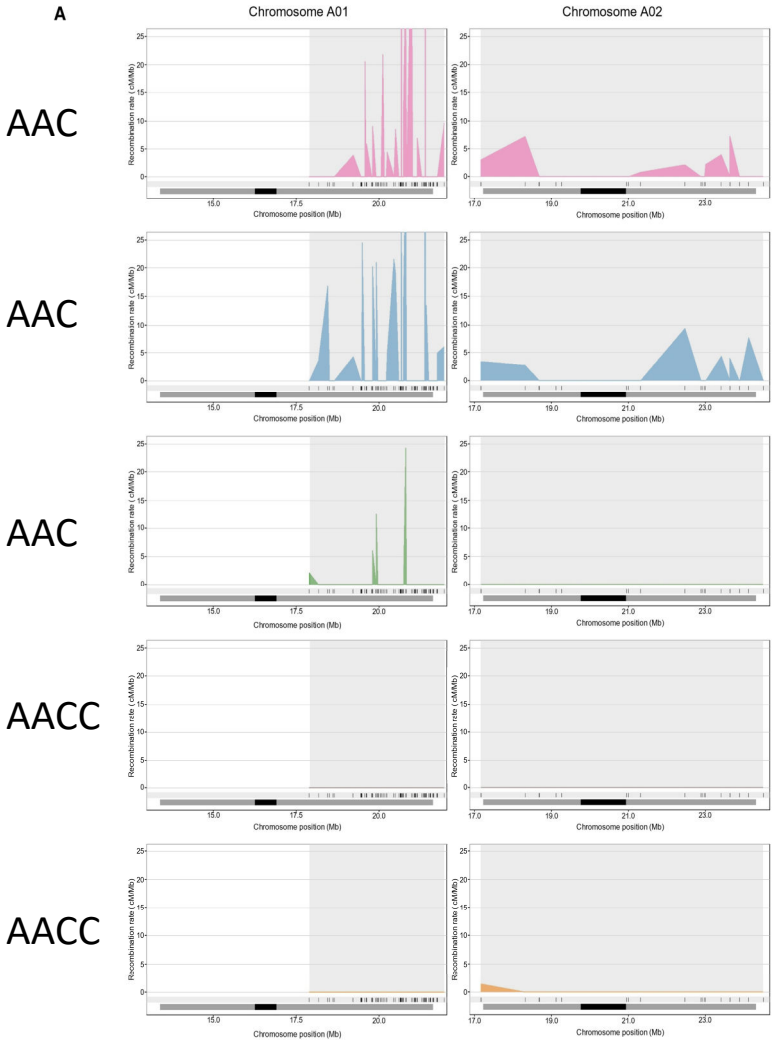


Comparison of the length of maps for F2 and RI populations of the same cross in pea

Boideau et al, 2024

E. Allotriploidy

- Increased recombination in pericentric region
- Reduced interference
- Effect disappears in allo4x plants



- Gray box = heterozygous area
- X axis = chromosome
- Black = centromere
- Dark gray = pericentric heterochromatin
- Vertical bars = genetic markers

F. Species

Price et al., 1993

Represents the sum total of genetic & structural differences that accumulate between species

Interval	Distance (cM)	
	Tomato	Pepper
CT268 - TG273	23.0	5.7
TG197 - TG158	52.2	34.6
r45S - TG31	2.8	5.0
TG48 - CD30A	15.6	6.0
CT128A - CT166	28.9	20.3
TG366 - TG244	86.3	75.3
TG264 - TG574	15.5	15.4
TG363 - CD64	18.3	23.2
TG623 - TG379	49.0	0.0
TG232 - TG253	46.3	51.9
TG20 - TG499	10.6	9.2
TG201 - TG496	10.6	9.2
TG47 - TG400	6.4	13.7
TG618 - TG296	39.9	0.0
TOTAL:	441.2	282.0

G. Genes

Robert, Farcy, & Cornu, 1991

	rm1/rm1	*Rm1/rm1*	*Rm1/Rm1*
Hf1-Lg1	15.3%	26.9%	
An2-Rt	0-0.5%	6%	26%

- *Rm1* gene on chromosome II of Petunia, an incompletely dominant gene that tends to increase recombination, especially for tightly linked loci.
  - Increases recombination for chromosomes 1, 2, 5, 6, and 7
  - Decreases it for chromosome 3; does not affect chromosome 4.

- For the following gene pairs on chromosome 1:
  Hf1-Lg1, where *Hf1* = purple vs. salmon corolla & *Lg1* = green vs. light green leaves
  An2-Rt, where *An2* = High vs. low anthocyanin content & *Rt* = purple vs. salmon corolla

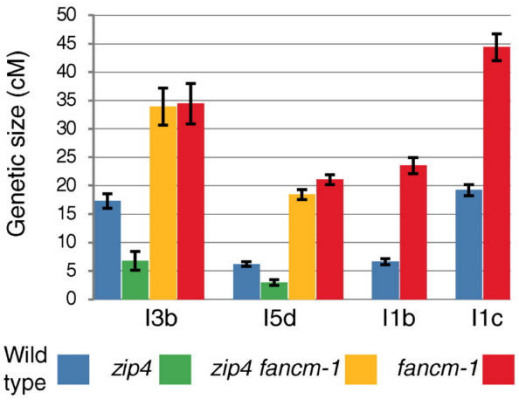
I. Helicases – FANC & Re4

Crismani et al., 2012

The FANCM gene in arabidopsis

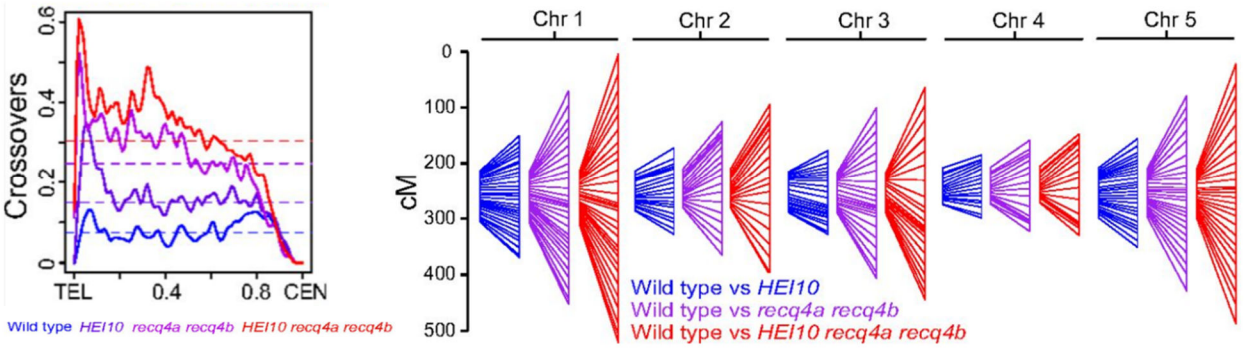
- Searched for recombination mutants by finding mutants that could complement *zip4* mutants. Mutants have 100-85% fewer COs.
- Used fluorescent protein transgenes linked together to measure recombination in different chromosome intervals

- The thing to notice is how *fancm-1* restores crossing over in *zip4* mutants and supersedes WT CO levels.
- Overall, there are far more Double Strand Breaks than there are COs
  - Eg, arabidopsis has ~230 DSBs but only ~10 Cos
- Helicases appear to play a role in keeping COs down
  - Ie, are anti-CO factors
- Silencing of helicases could increase CO frequency in crops to facilitate breakage of old linkages



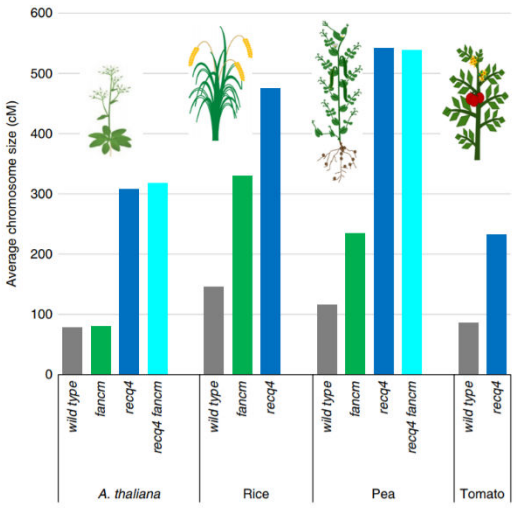
Serra et al, 2018

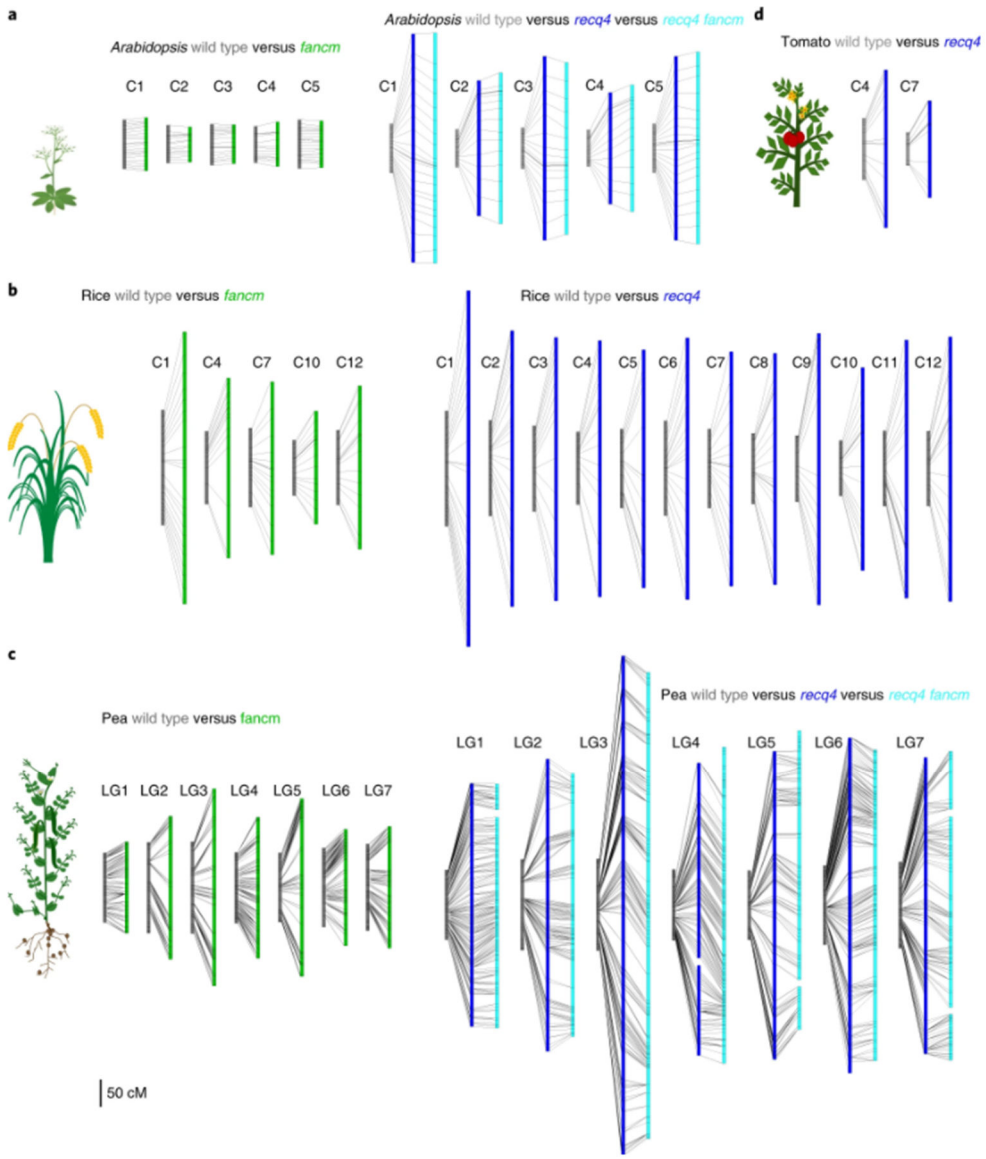
- Hel10 - procrossover E3 ligase gene
- Req4A & B – affect interfering and non-interfering crossovers
- Adding Hel10 increases CO coincidence (ie, decreases interference)
- Removing the Req4 genes increases recombination
- Hel10 – Req4 is synergistic



Mieulet et al., 2018

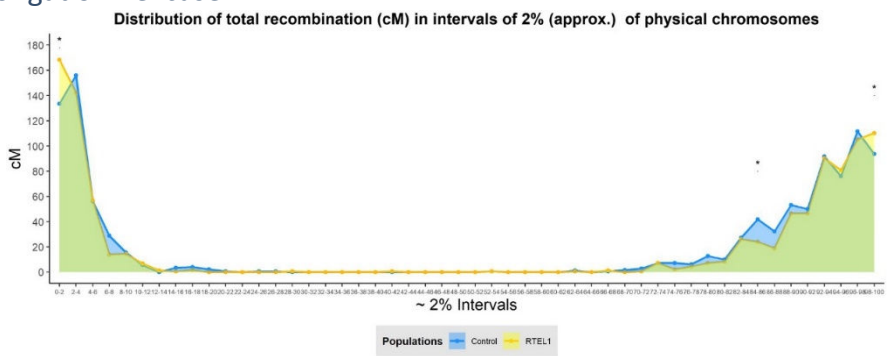
KOs of FANC are somewhat synergistic with KOs of req4 across species





Barakate et al, 2021

- Regulator of telomere elongation helicase



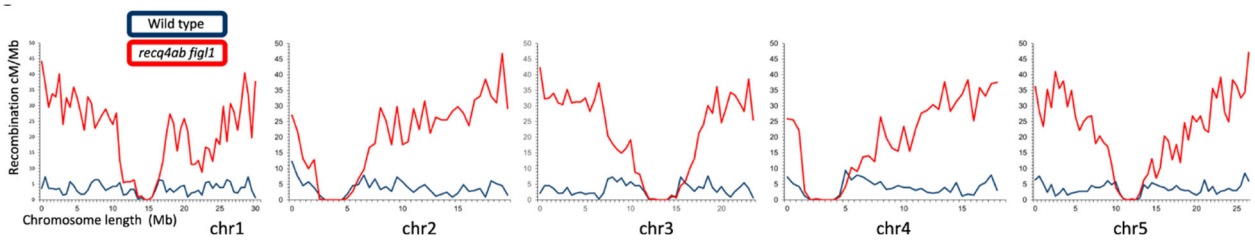
Capilla-Pérez et al, 2024

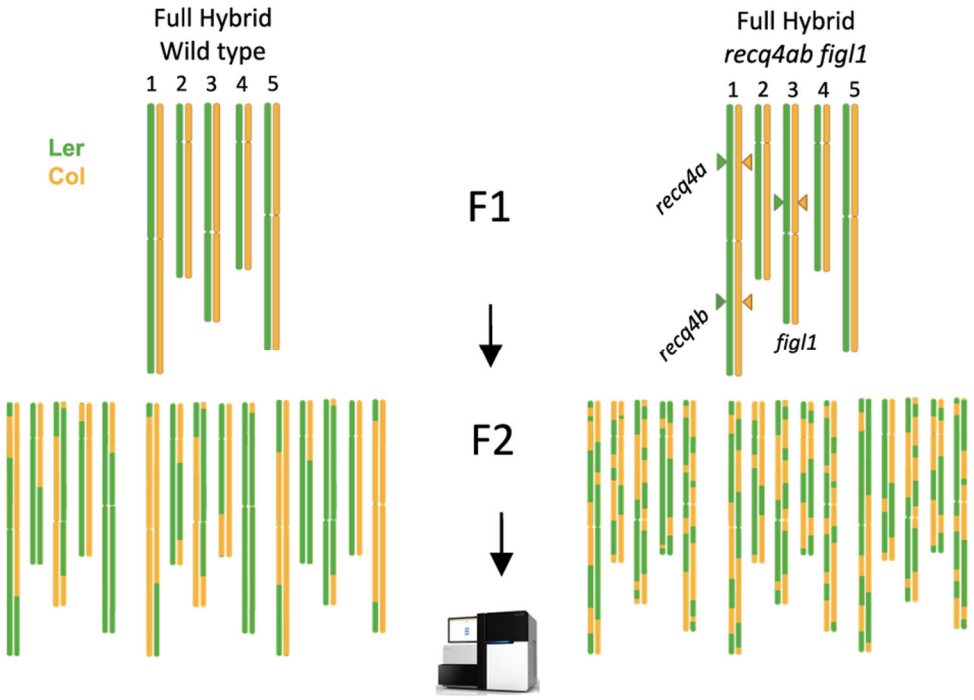
There are two types of crossovers

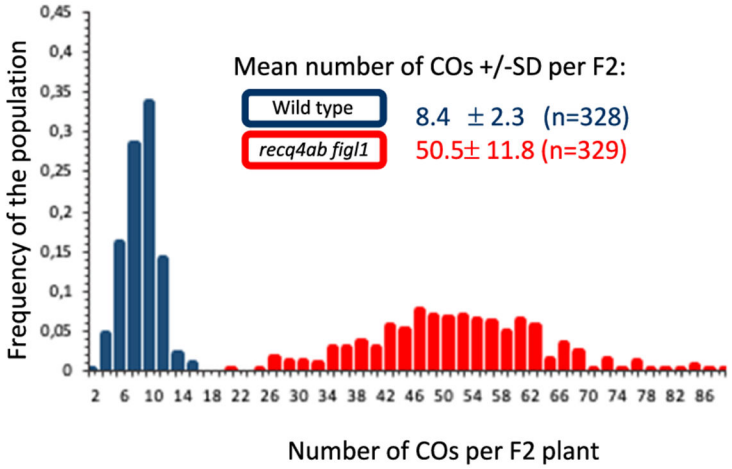
- Class I
  - The most common
  - Subject to interference
  - Limited by dosage of
    - HEI10
    - Phosphatase X1
    - Synaptonemal proteins ZYP1/SCEP1/SCEP2

- Class II
  - Limited by 3 protein complexes
    - TOP3/RECQ4AB/RMI1
    - FANCM
    - FIGL1/FLIP

- Got a 6x increase in recombination → the highest rate yet achieved




Full Hybrid
Wild type
1 2 3 4 5
Ler
Col
F1
F2
Full Hybrid
recq4ab figl1
1 2 3 4 5
recq4a
recq4b
figl1


Mean number of COs +/-SD per F2:
Wild type
8.4 ± 2.3 (n=328)
recq4ab figl1
50.5± 11.8 (n=329)
Frequency of the population
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
2 6 10 14 18 22 26 30 34 38 42 46 50 54 58 62 66 70 74 78 82 86
Number of COs per F2 plant

Factors Affecting Recombination

Environment

1. Age

Bridges, 1915

- 2nd brood of progeny from a female had lower recombination rates

2. Temperature

Dowrick, 1957

°C	Time (h)	% interstitial xmata	Total xmata (100 PMC)	°C	Time	% interstitial xmata	Total xmata (100 PMC)
control	--	2.3	743	control	--	2.2	757
20	12	4.8	773	34	12	8.8	832
	24	6.2	836		18	18.9	984
	48	11.9	847		24	21.6	1006
	96	15.3	842		48	12.1	860
	192	13.7	883		96	0.7	431
control	--	2.0	707	control	--	2.3	743
27	12	7.4	757	39	18	16.8	841
	24	15.8	845		24	41.2	1152
	48	17.4	849		27	29.4	961
	144	17.3	851		30	19.3	787
					48	0	72

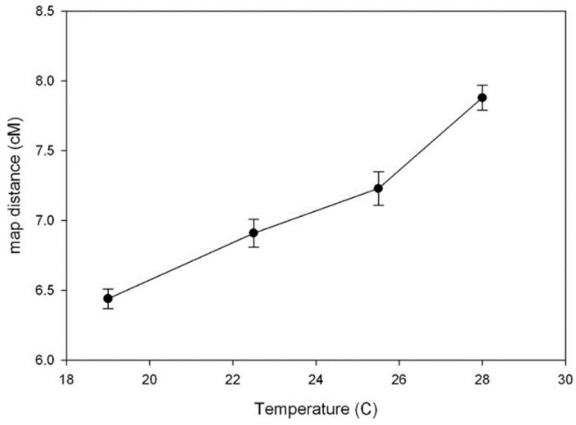
First reported by Plough (1917) on Drosophila chromosome # 2 [J. exp. Zool. 32:147-209].

- Extensively studied by Dowrick, 1957, in *Tradescantia*
  - Note that effect depends on temperature
  - Increase in number & location of chiasmata
  -

Francis et al., 2007

Effect confirmed in arabidopsis by measuring recombination between linked transgenes

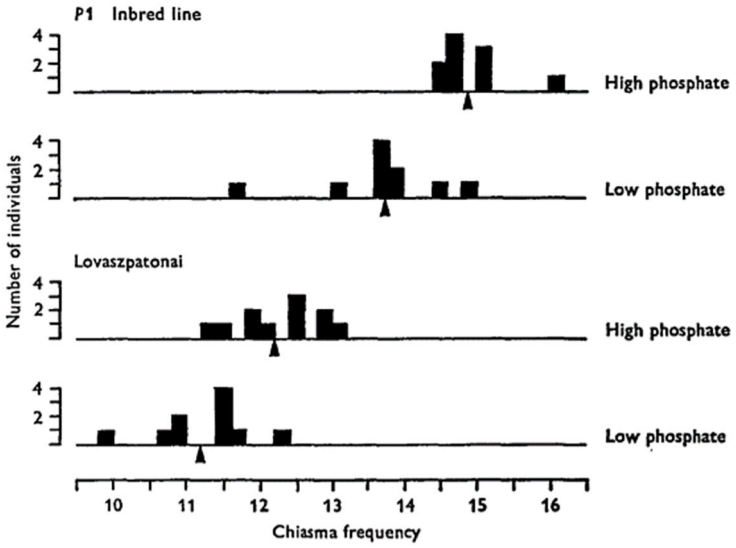
- Also seen in barley [Powell & Nilan, 2000. Crop Sci 3:11-13] and
- Faba bean [Berkemeir & Linnert. 1987. Biol. Zentralbl. 106:219-230.]



3. Stress

Bennett & Rees, 1970

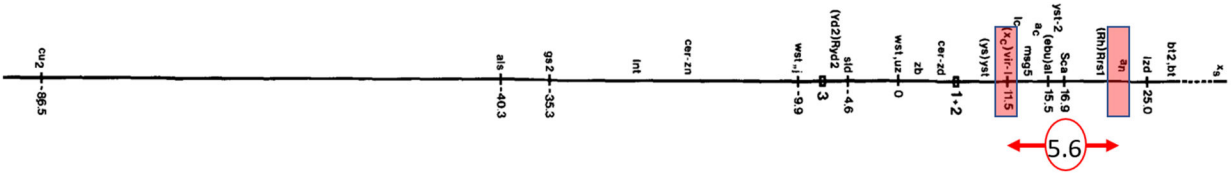
Starvation, metal chelates, antibiotics and radiation have all been found to increase recombination, while low calcium concentrations decrease it. Here are some data on the effects of phosphorus.



Sinha & Helgason, 1969

Lethal genes *Xantha* (*xc*) and *albino* (*an*) linked in repulsion near centromere.

- Thus F2 segregates 9:3:3:1 (9:7) for lethality



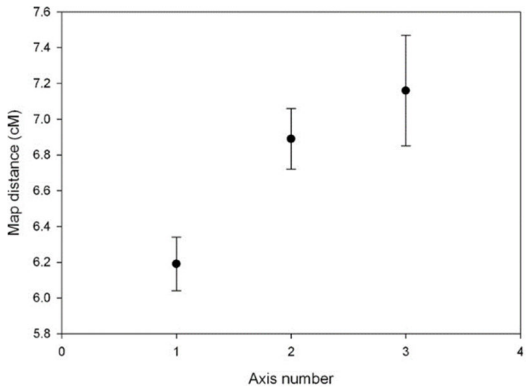
- The use of
  - Actinomycin D
  - Diepoxybutane
- Increases recombination distance to 15.1%

4. Position on flower head

Francis et al., 2007

Flowers on the main stem exhibited less recombination than those on the first or second inflorescence branches.

Recombination was measured between a pair of transgenes.



A NOTE OF CAUTION

Barth et al, 2000

Looked at effect of sex, temperature and phosphate on recombination in arabidopsis

- Found an effect, but effect was not genome-wide
- Effect limited to specific chromosome regions

Manipulating Crossovers

Yanagira et al., 1992

Gene order: *Wx – C – S-5* on chromosome 6
Note: today *javonica* is called tropical *japonica*

Where

- *wx* = glutinous endosperm
- *c* = chromogen
- *S-5* = compatibility locus

S-5 has several alleles

- $S\text{-}5^{j}$ is found in *japonica* type rices
- $S\text{-}5^{i}$ is found in *indica* type rices
- $S\text{-}5^{n}$ is found in the *javonica* type rices, & permits crossing to *indica* and *japonica* types.

The *Wx-C* distance = 22.9% in *japonica* × *japonica* ($S\text{-}5^{j} \times S\text{-}5^{j}$)

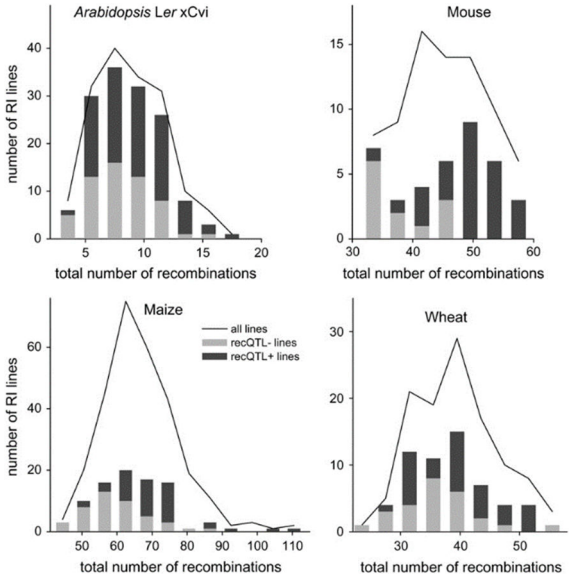
- 16.5% in *japonica* × *javonica* ($S\text{-}5^{j} \times S\text{-}5^{n}$)
- 30.3% in *javonica* × *indica* ($S\text{-}5^{n} \times S\text{-}5^{i}$)

When the *javonica* allele is backcrossed into *japonica (japonica*), japonica* behaves like *javonica*

- 16.1% in *japonica* × *japonica** ($S\text{-}5^{j} \times S\text{-}5^{n}$)
- 30.3% in *japonica** × *indica* ($S\text{-}5^{n} \times S\text{-}5^{i}$)

Esch et al., 2007

This concept has been verified in wheat, arabidopsis and maize, by studying RILs with different recombination frequencies



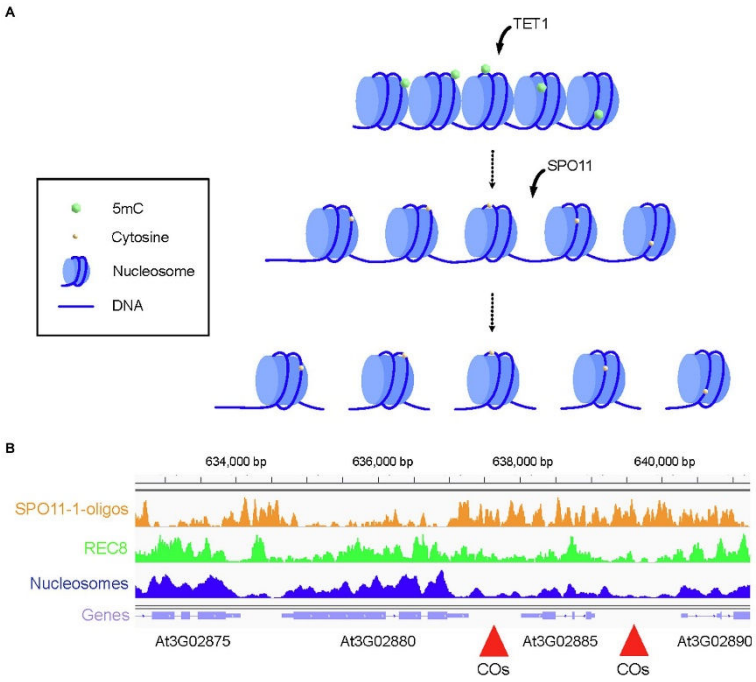
- 1 QTL on chromosome 1 in arabidopsis
- 2 QTLs on chromosome 3 in maize
- 1 QTL on chromosome 3B in wheat
- QTLs could explain up to 15% of observed variability in CO frequency

Engineering recombination hotspots

Kuo et al, 2021

Strategies to remodel the crossover rate locally.

A cold spot region is enriched in nucleosome and silencing epigenetic marks, such as 5mCs methylation in all three contexts (CG, CHG, and CHH).

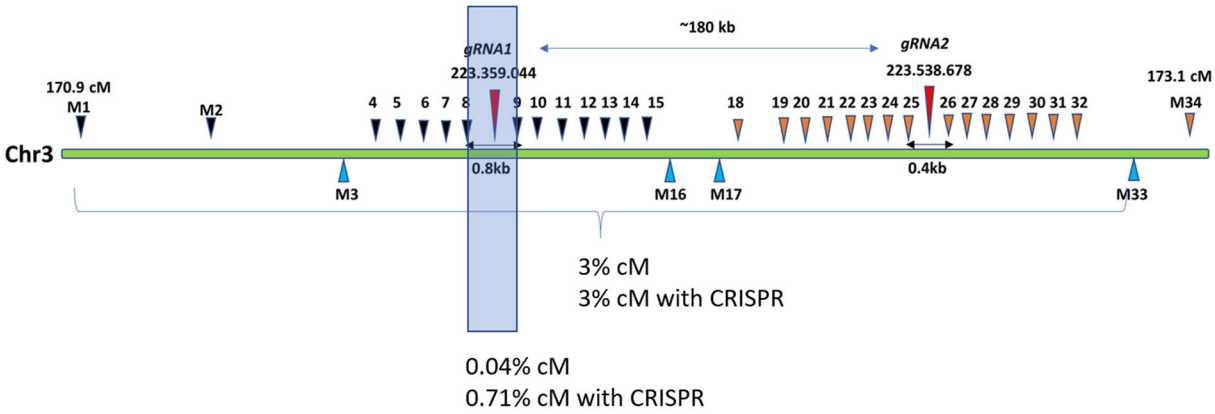


- Targeted recruitment of TET1 would remove silencing epigenetic marks, leading to decompaction of chromatin
- Targeted recruitment of SPO11 produces DSBs.
- Meiotic DSBs are repaired by the homologous recombination pathway leading to the formation of NCO, NCO with gene conversion, or CO.

(B) Genome browser view of the crossover hotspot 3a (red arrowheads) on chromosome 3 of arabidopsis. SPO11-1-oligo (orange), REC8 ChIP-seq (green), nucleosome (MNase, dark blue) & gene organization (purple).

Kouranov et al, 2022

Mitosis. Engineer DSBs in the same spot in the F1. CO frequency is increased, but still get chimeras, as CRISPR not always gets the job done in the engineered cell, but after the cell has divided.



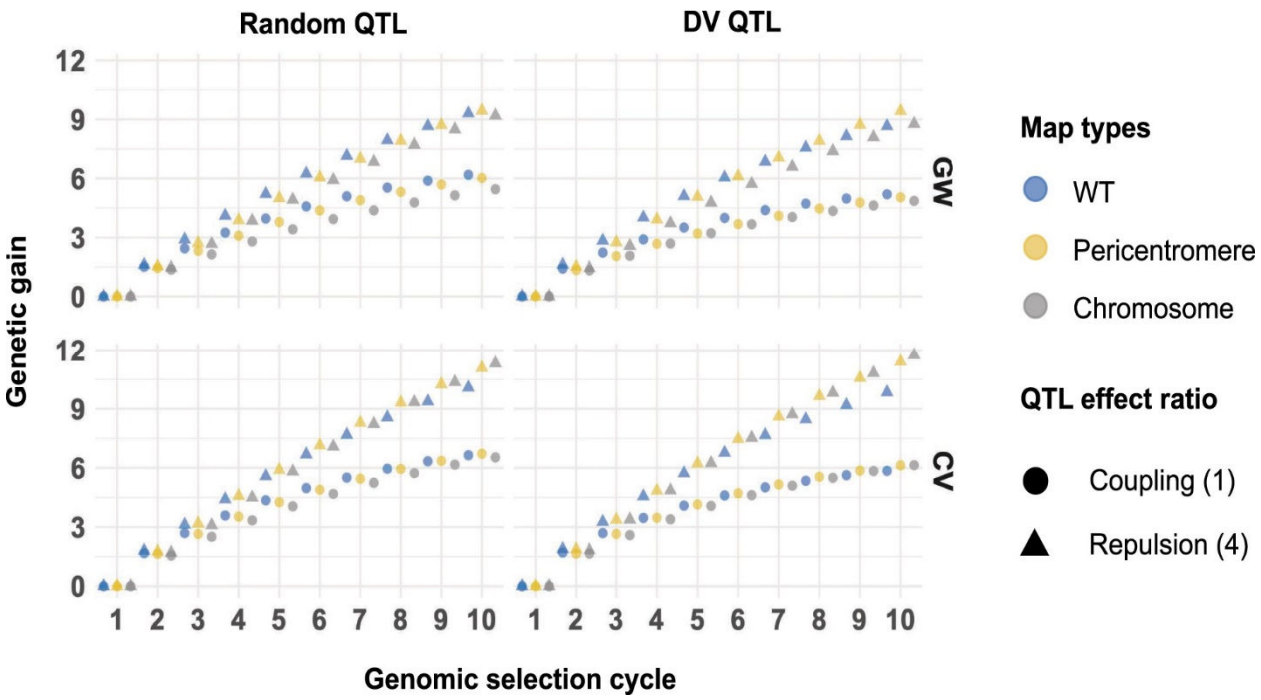
Limitations

Taagen et al., 2022

Compared

- WT recombination, vs increased recombination in the peri-centric area or increased recombination in the whole chromosome
- Random markers vs markers for deleterious alleles

- Markers in coupling vs repulsion



- Increased recombination is more effective for traits linked in repulsion than for those linked in coupling
- "Increased recombination is not beneficial for oligogenic traits"
- "The variance generated from increased recombination may not increase genetic gain"
- "The efficiency of increased recombination may depend on knowledge of QTL locations"