Where plants rank

Based on **Arumuganathan and Earle, 1991**

- The following table is a partial listing. Non-plant values have been added to it.
- 1 pg = 965 million base pairs

Species	Common name	Family	DNA Mbp/1C	DNA pg/2C
Escherichia coli			4.6	**0.04**
Genlinsea tuberosa	Corkscrew plant	Lentibulariaceae	61	0.13
Drosophila melanogaster	**Fruit fly**		130	**0.24**
Arabidopsis thaliana	Mouse-ear cress	Cruciferaeae	145	0.30
Oryza sativa	Rice	Gramineae	415	0.86
Phaseolus vulgaris	Bean	Leguminoseae	637	1.32
Musca domestica	**House fly**			**1.78**
Glycine max	Soybean	Leguminoseae	1115	2.31
Zea mays	Maize	Gramineae	2292	4.75
Homo sapiens	**Human**			**6.00**
Pisum sativum	Pea	Leguminoseae	3947	8.18
Hordeum vulgare	Barley	Gramineae	4873	10.10
Gryllus domestica	**Crickett**			**12.00**
Bufo bufo	**Toad**			**12.00**
Allium cepa	Onion	Amaryllidaceae	15,290	31.69
Triticum aestivum	Wheat	Gramineae	15,966	33.09
Amphiuma means	**Congo eel**			**170.00**
Fritillaria platyptera+++		Liliaceae	84,150	174.40
Paris japonica	Japanese canopy plant	Melanthiaceae	149,000	304.46

+++ True diploid (2n = 2x = 24) with largest genome

C-value paradox

Thomas, 1971

Angiosperm DNA C-values Database

Pellicer & Leitch, 2020

- Plant DNA C values can be found at https://cvalues.science.kew.org/
- 1 pg = 978 million base pairs (value from Delezal et al, 2003)

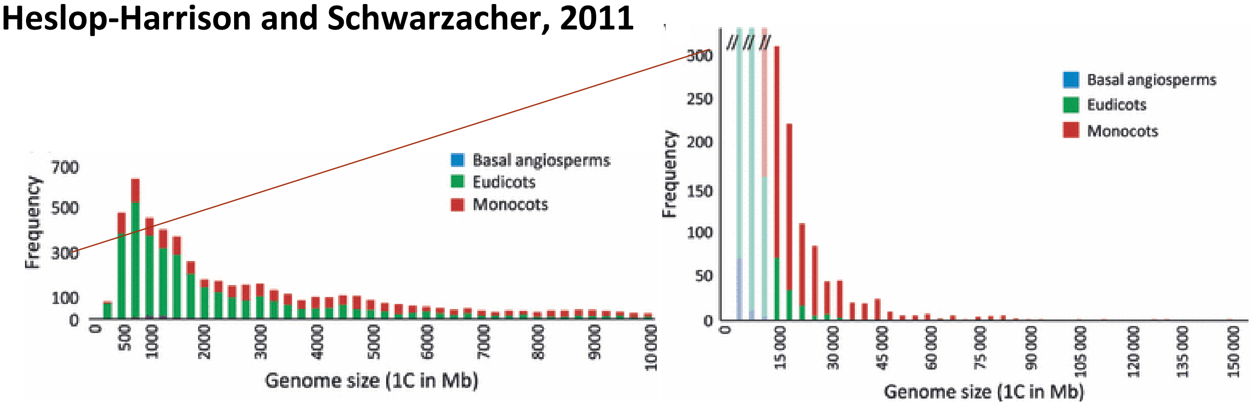
Plant DNA C-values Database

Release 7.1, April 2019. Leitch IJ, Johnston E, Pellicer J, Hidalgo O, Bennett MD

https://cvalues.science.kew.org/

Range in angiosperm DNA content

Heslop-Harrison and Schwarzacher, 2011



Smallest genome size

Heslop-Harrison and Schwarzacher, 2011
Fleischmann et al, 2014

Genlinsea aurea

DNA Mbp/1C = 63
DNA pg/2C = 0.13

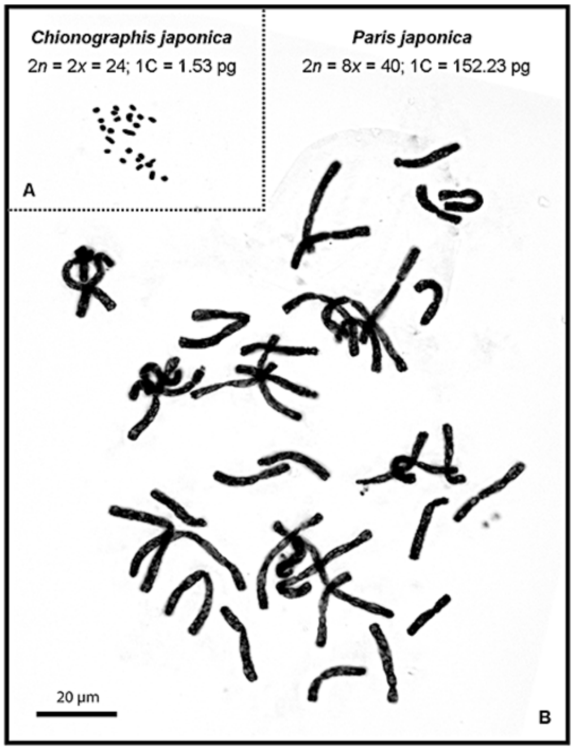
Genlinsea tuberosa

DNA Mbp/1C = 61
DNA pg/2C = 0.126

http://cpphotofinder.com/genlisea-aurea-516.html

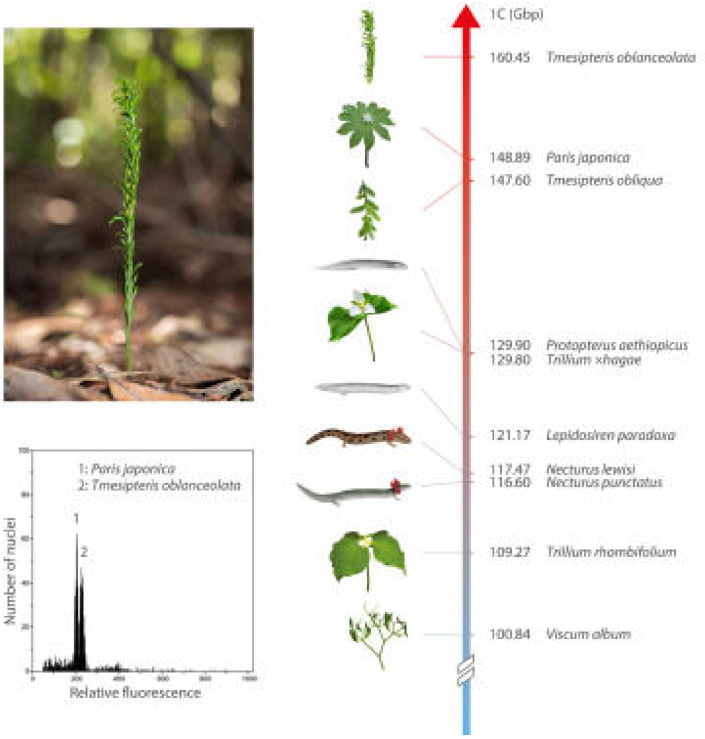
Largest plant genome

Pellicer et al., 2010



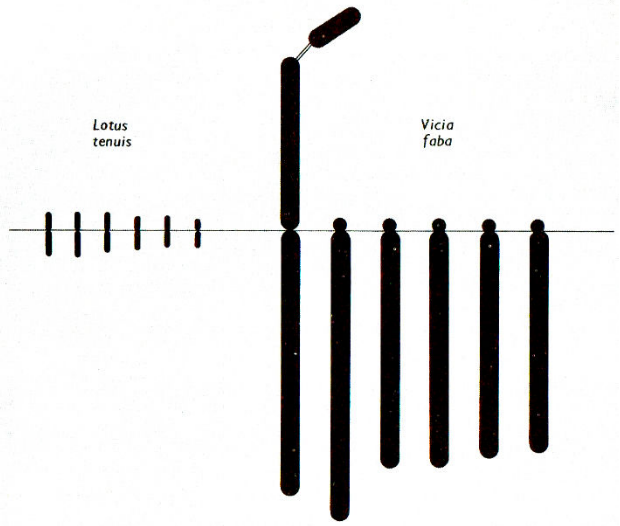
Paris japonica, 2n = 40, Wikipedia

Fernández et al, 2024 (https://doi.org/10.1016/j.isci.2024.109889)



Changes in DNA amount & chromosome size

Review by **Walbot & Cullis, 1985**

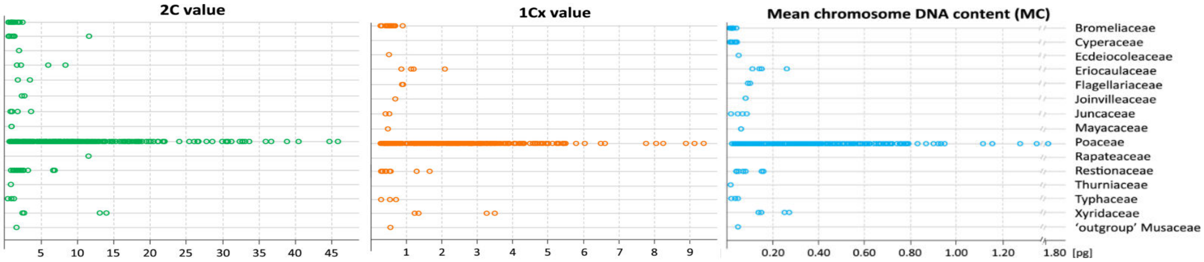


Variation in chromosome size in 2 leguminous species with the same chromosome number. ***Stebbins 1971***

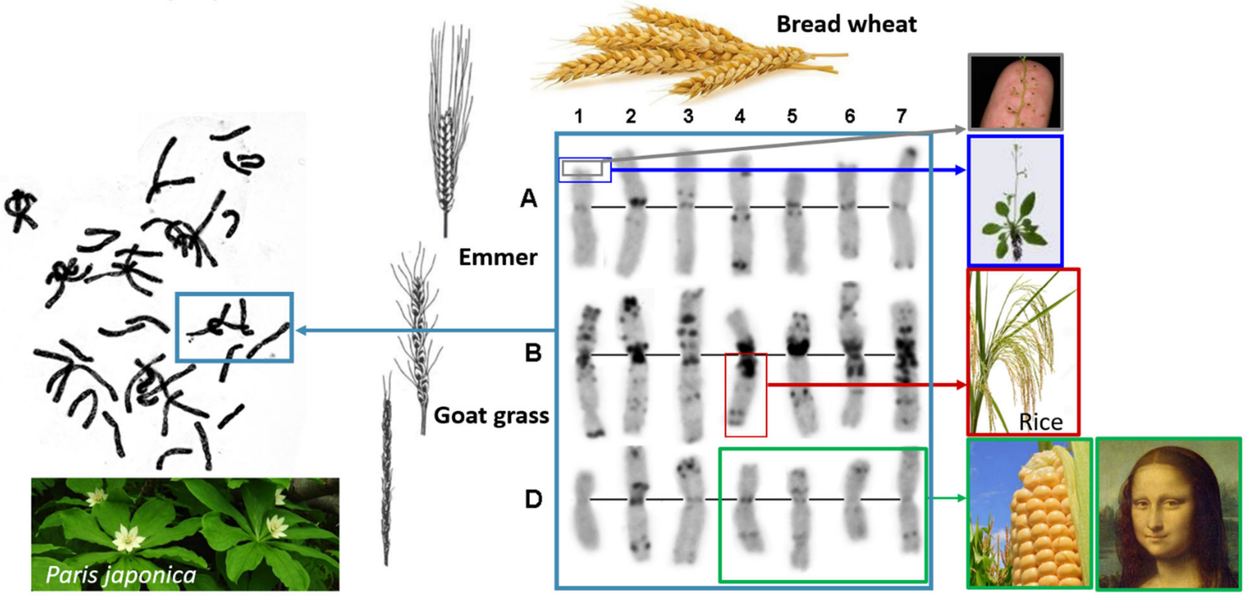
Family	Species	*n*	Chromosome length (µm)
Ranunculaceae	*Isopyrum fumarioides*	7	1.3
	Anemone hepatica	7	7.0
Oxalidaceae	*Oxalis cuneata*	6	1.5
	Oxalis dispar	6	15.1
Leguminoseae	*Lotus tenuis*	6	1.8
	Vicia faba	6	14.8
Droseraceae	*Drosera rotundifolia*	10	0.8
	Drosophyllum lusitanicum	6	14.0
Compositae	*Agroseris heterophylla*	9	2.4
	Chaetadelpha wheeleri	9	6.4
Liliaceae	*Tofieldia nuda*	15	1.3
	Lilium pardalinum	12	20.0
Gramineae	*Chloris barbata*	10	1.5
	Secale cereale	7	7.2

Genome size variation in Poales vs WGDs

Winterfeld et al, 2025

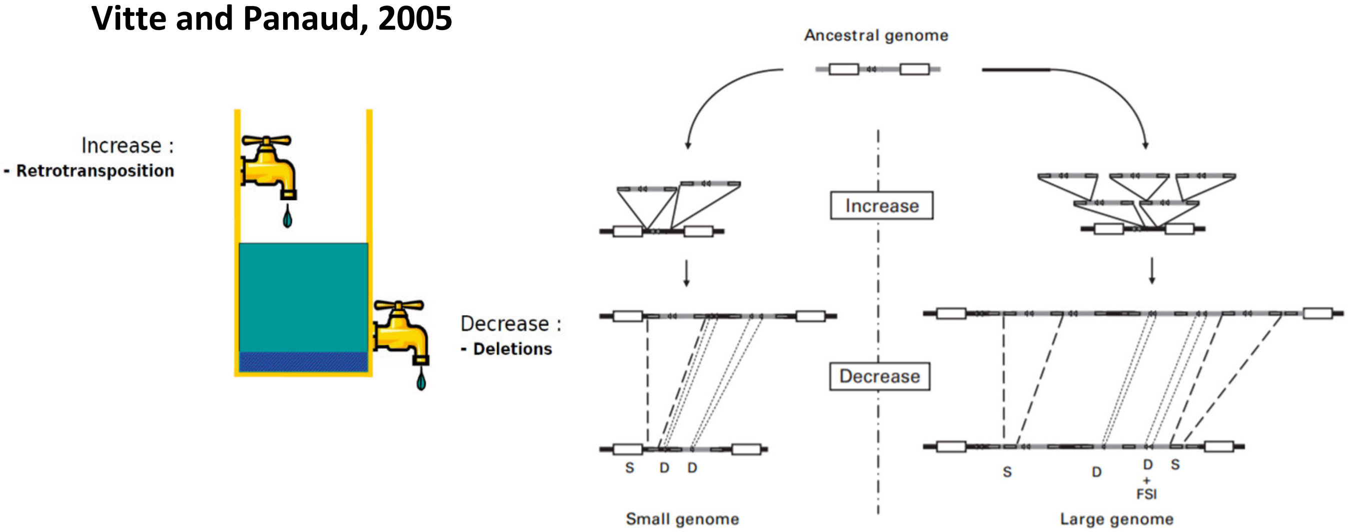


Slide concept by Catherine Feuillet



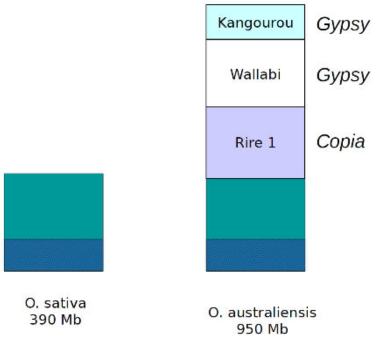
The increase/decrease model

Vitte and Panaud, 2005

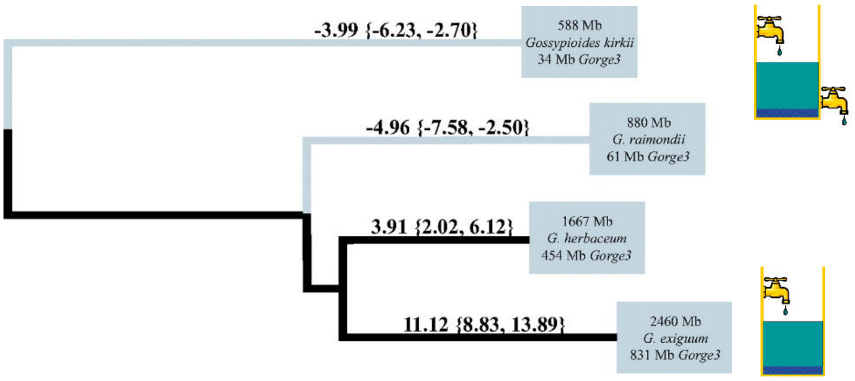


Piegu et al, 2006 (graphics by O Panaud)

- Doubling of rice EE genome size

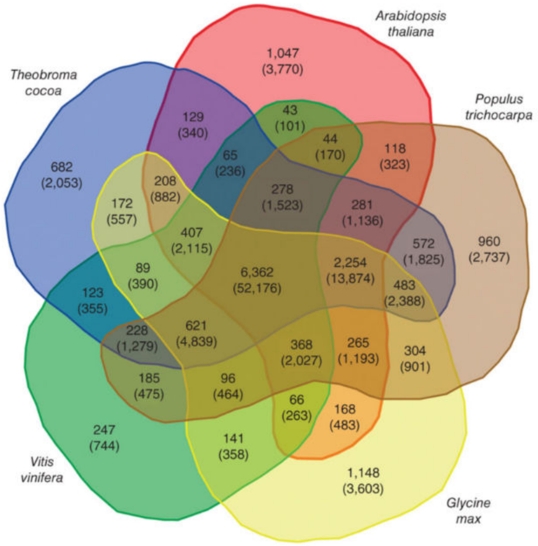


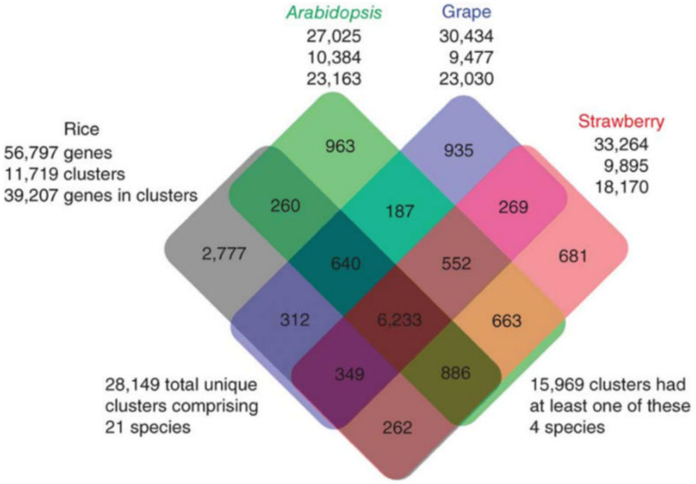
Hawkins et al., 2009



Number of genes to make an angiosperm?

Argout et al, 2011; Shulaev et al, 2011





Number of genes per monoploid genome

Goldberg, 1986

Edwards et al, 2017

Gibson and Sommerville, 1993

Arabidopsis sequence

Wang et al, 2022

Sequenced angiosperm genomes

- CoGepedia - https://genomevolution.org/wiki/index.php/Sequenced_plant_genomes
- Wikipedia - https://en.wikipedia.org/wiki/List_of_sequenced_plant_genomes
    - Rosids
        - 18,397 marula tree – 73.013 for poplar
    - Asterids
        - 14,474 for milkweed – 52,232 for sunflower
    - Grasses
        - 25,225 for moso bamboo - 107,891 for wheat
    - Other monocots
        - 19,623 for duckweed – 52,342 for *Daemonorops jenkinsiana*

Besides protein-coding genes

Cacao - **Argout et al, 2011**

- 28,798 protein-coding genes
- 6 rRNA genes (many missing due to sequencing method!)
- 473 tRNA genes
- 83 miRNA genes
- 67,575 transposable elements (25.7% of assembly)

Strawberry - **Shulaev et al, 2011**

- 34,809 protein-coding genes
- TE's = 22% of genome
- 569 transfer RNAs (tRNAs)
- 177 ribosomal RNA (rRNA)
- 111 spliceosomal RNAs
- 168 small nucleolar RNAs
- 76 micro RNAs
- 24 other RNAs

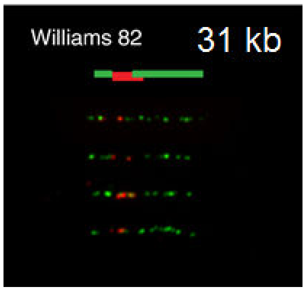
Bladderwort – **Ibarra-Laclette et al, 2013**

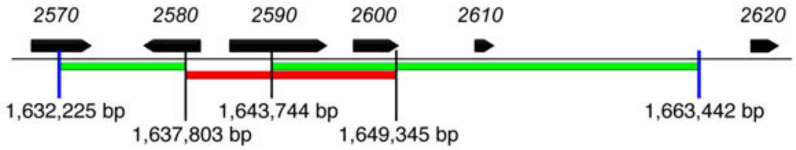
- 28,500 protein-coding genes – but genome is only 77 Mbp
  - Shorter promoters and introns; fewer introns
- Only 3% of DNA is repetitive
- Only 579 transposable elements
  - 375 are retroelements
  - 95 intact (and supposedly functional)
- All miRNA genes are there

Copy Number Variation (CNV)

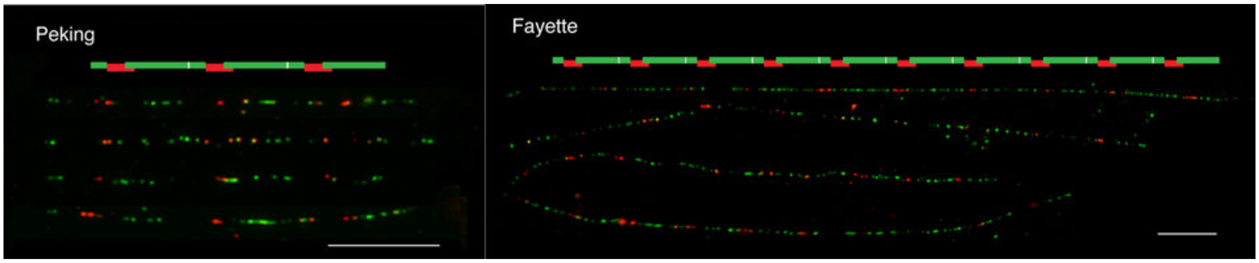
Eg, Cyst nematode resistance in soybean

Cook et al, 2012



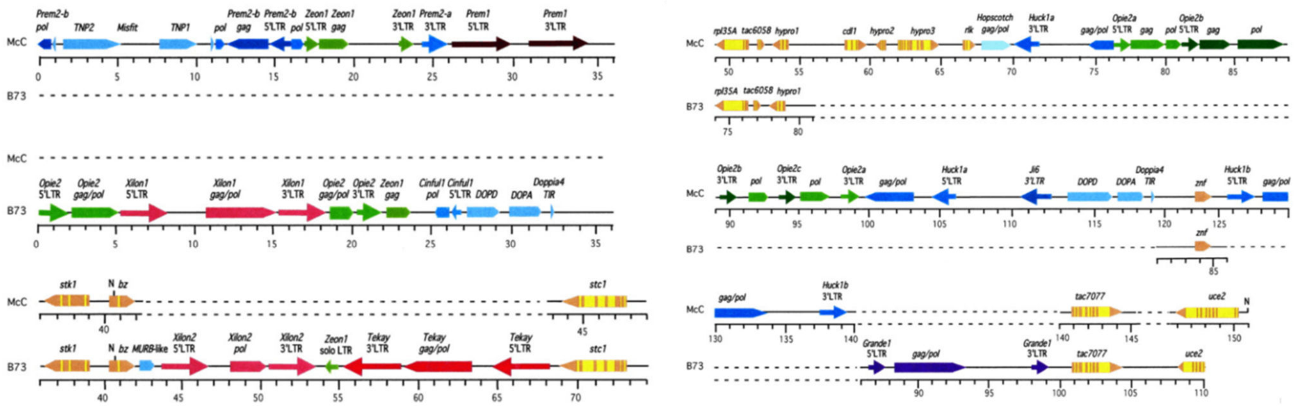


Presence



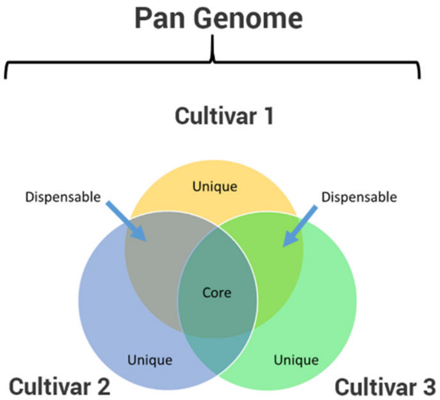
Absence Variation (PAV)

Fu & Dooner, 2002



The pangenome concept

Review by **Della Coletta et al, 2021**

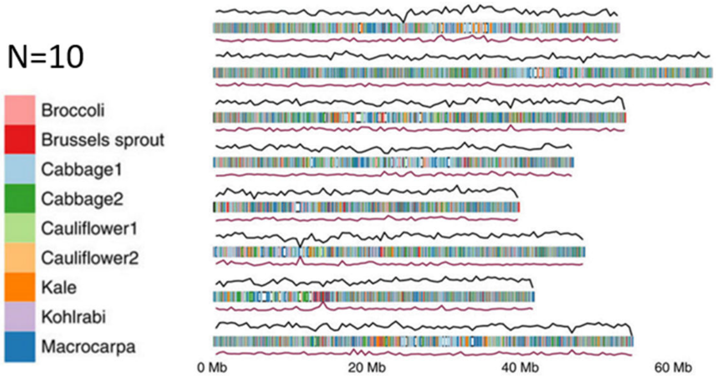


http://www.10wheatgenomes.com

- **Pan genome**
- **Core genome**
- **Dispensable genome**

PAV in *Brassica oleracea*

Agnieska et al, 2016

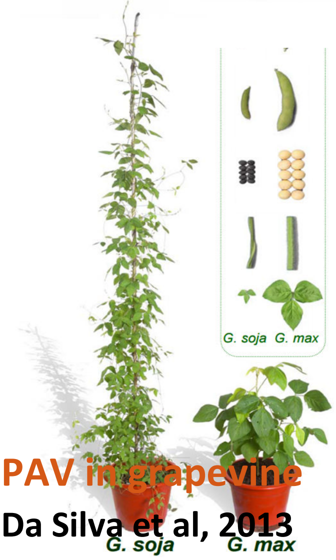


- 81.3% of genes always present (49,895) → Core genome
- 18.7% are in variable numbers (11,484) → Dispensable, variable, or shell genome
- 02.2% are present in only 1 line (1,322)

} Pan Genome

PAV in *Glycine*

Li et al, 2010



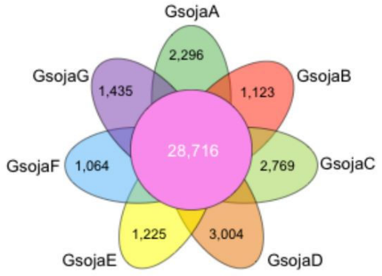
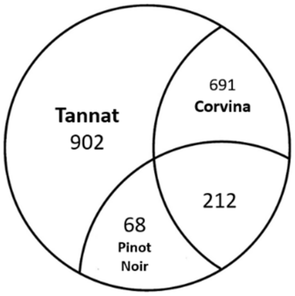


Figure 1. Gene families unique to each genotype

PAV in grapevine

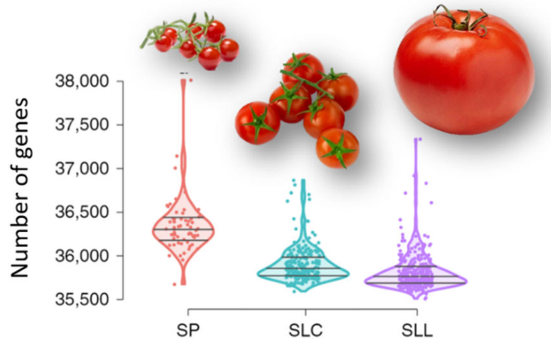
Da Silva et al, 2013

"Aromas and flavors of black berry custard tart, chocolate gelato, and praline with a fruity body and a long cranberry chutney and floral honey finish"



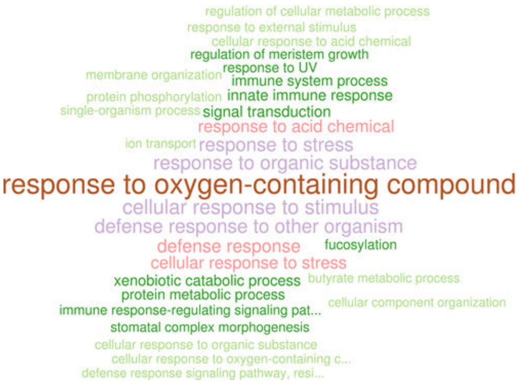
PAV in tomato

Gao et al, 2013



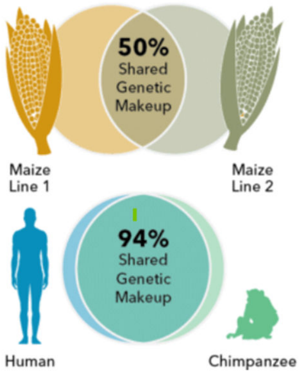
PAV in wheat

Montenegro et al, 2017



PAV in maize

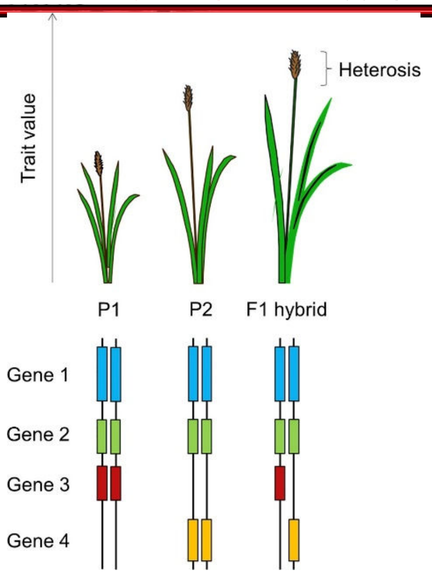
Springer et al, 2018



https://www.pacb.com/blog/sequencing101pangenome

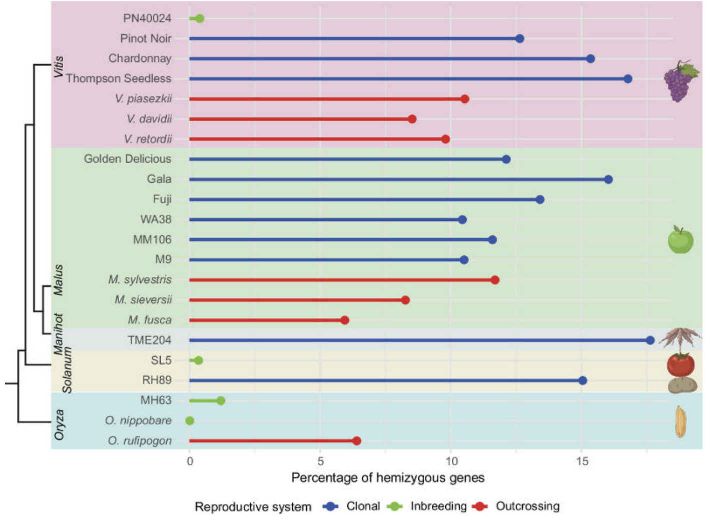
Pan genome hypothesis for heterosis

Tao et al, 2019



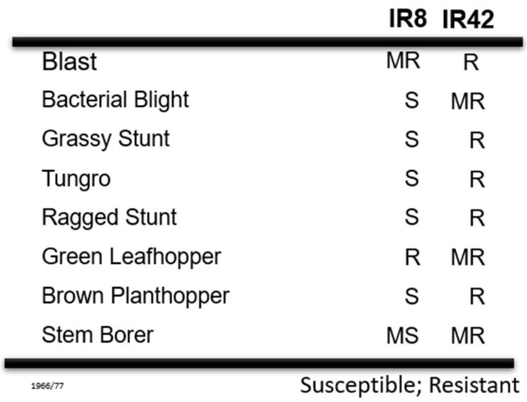
Hemizygosity

Peng et al, 2018



Trait accumulation in conventional plant breeding

IRRI 1980 Annual Report

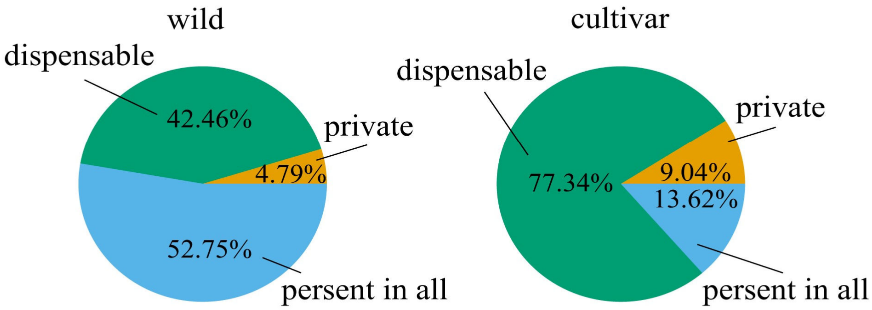


	IR8	IR42
Blast	MR	R
Bacterial Blight	S	MR
Grassy Stunt	S	R
Tungro	S	R
Ragged Stunt	S	R
Green Leafhopper	R	MR
Brown Planthopper	S	R
Stem Borer	MS	MR

1966/77

Susceptible; Resistant

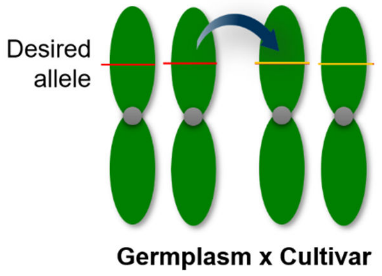
Long et al, 2024



Reevaluating the backcross

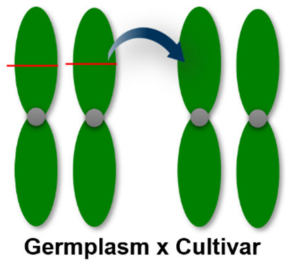
Traditional view

- Backcrossing replaces an existing allele with a desired allele



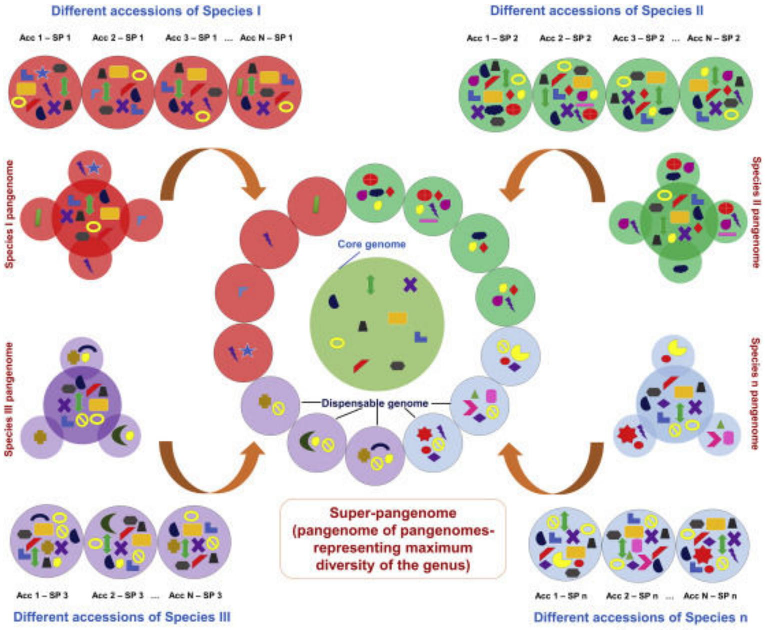
Emerging view

- Besides allele replacement, backcrossing can introduce a PAV gene
- Can also eliminate genes that were there



Super pan genome

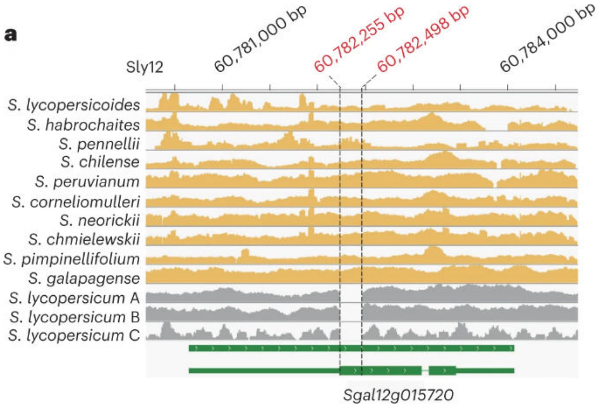
Khan et al, 2020



Trends in Plant Science

Using a super pangenome

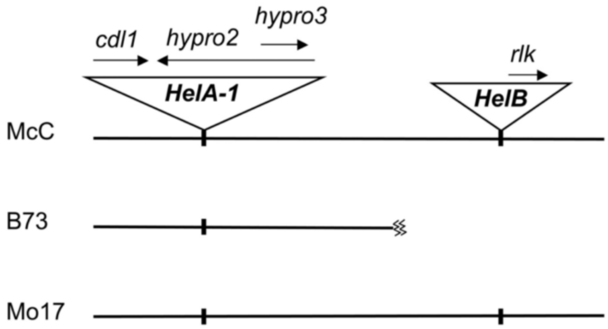
Li et al, 2023 (https://www.nature.com/articles/s41588-023-01340-y)



Origins of PAV

Helitrons

Lai et al., 2005



Fractionation/differential gene loss

Modified from **Vision, 2005**

Town et al, 2006; Eichten et al, 2011; Schnable et al, 2011

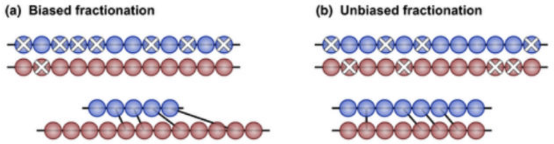
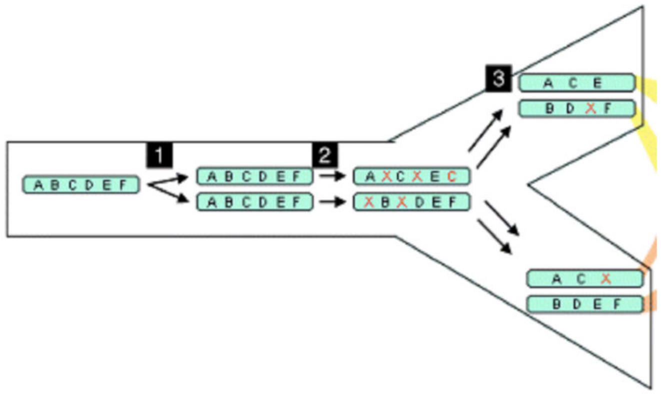
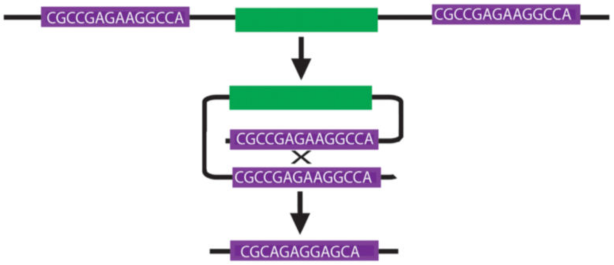


Figure 2. Bird et al, 2018 (Edger lab)



Gene loss through intra-strand recombination

Woodhouse et al., 2010

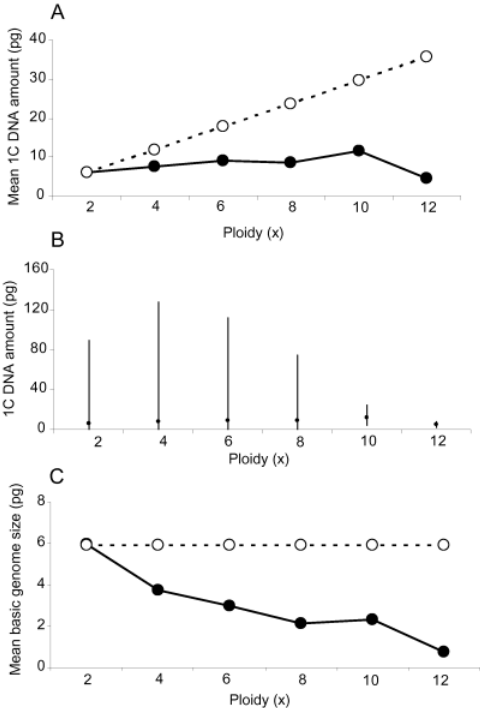


Creation of new genes

Fan et al., 2008

Genome downsizing

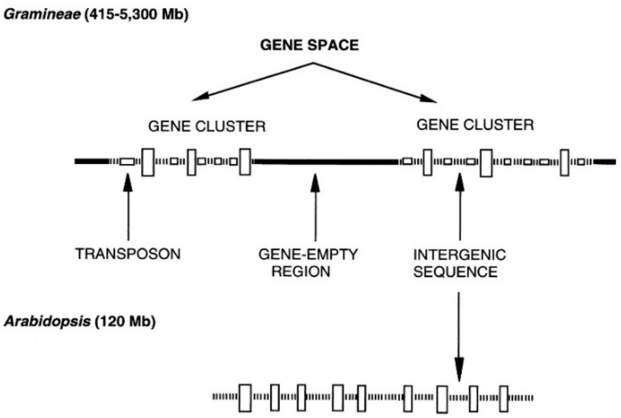
Leitch & Bennett, 2004



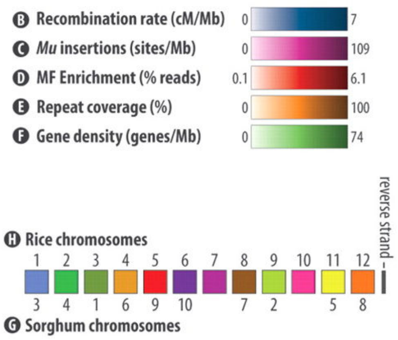
Genome sequence summary & caveats

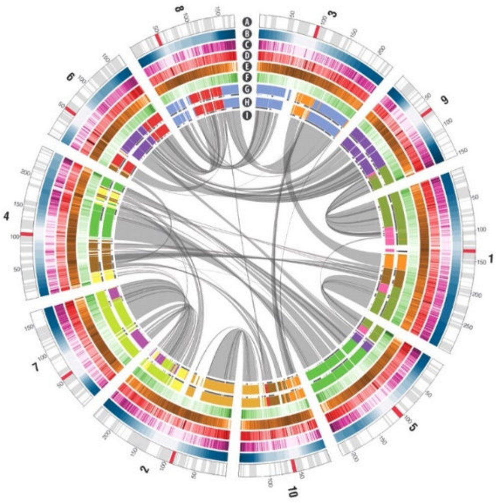
Gene space in the plant genome

Barakat, Matassi, & Bernardi, 1998

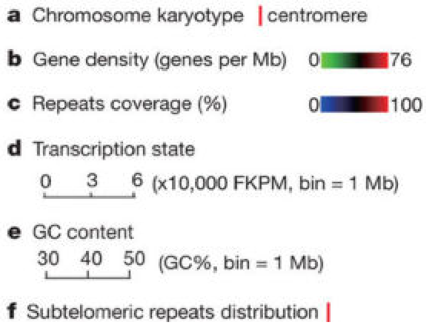


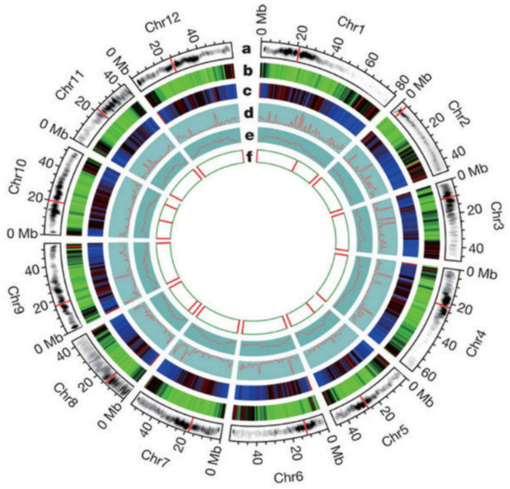
Schnable et al, 2009





The Potato Genome Sequencing Consortium, 2011

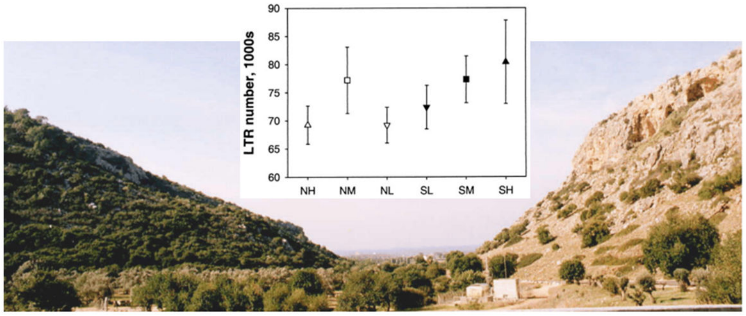




Adaptive Value of DNA Content

Environmental effects on genome size

Kalendar et al, 2000



Cell size vs DNA content

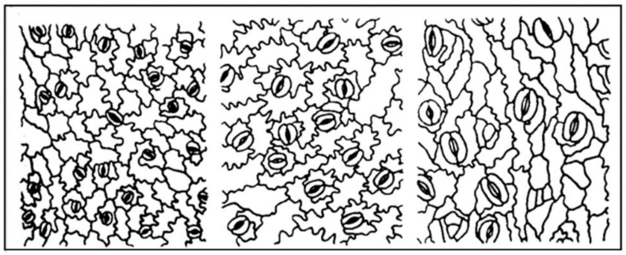
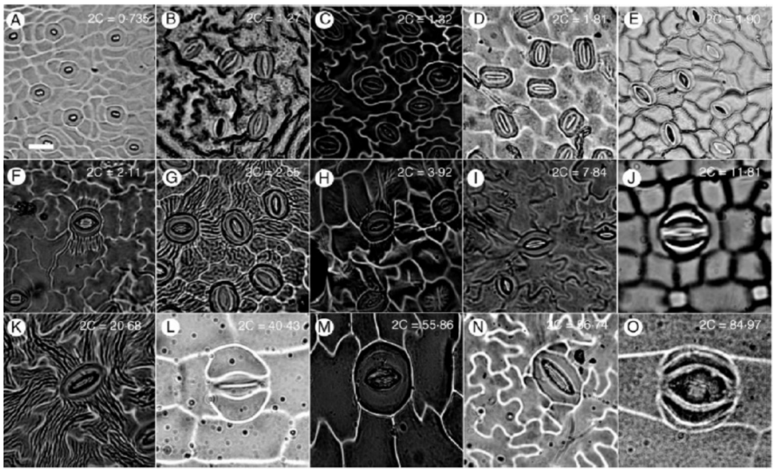
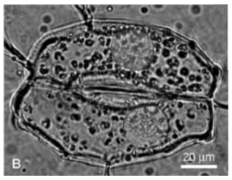


Figure 3. Epidermal cells of 2x, 4x and 8x Nicotiana. Greenleaf, 1938

Knight & Beaulieu, 2008; Pellicer et al, 2010



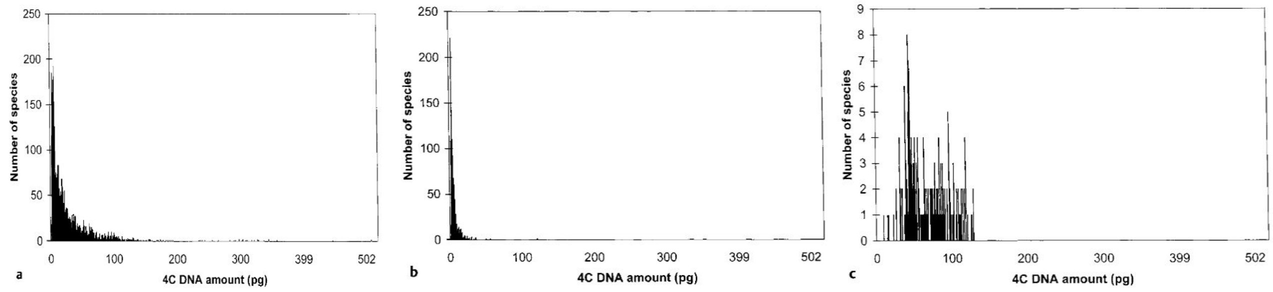


Paris japonica
2C = 304.46

White scale bar in (A) = 20 µm. (A) *Citrus aurantifolia*, (B) *Rosa acicularis*, (C) *Origanum vulgare*, (D) *Quercus robur*, (E) *Fraxinus excelsior*, (F) *Ligustrum vulgare*, (G) *Platanus orientalis*, (H) *Cypripedium irapeanum*, (I) *Capsicum annuum*, (J) *Zingiber officinale*, (K) *Sambucus racemosa*, (L) *Tradescantia paludosa*, (M) *Cypripedium formosanum*, (N) *Paeonia clusii*, (O) *Tradescantia virginiana*.

Constraints on cell size

Ohri, 2008



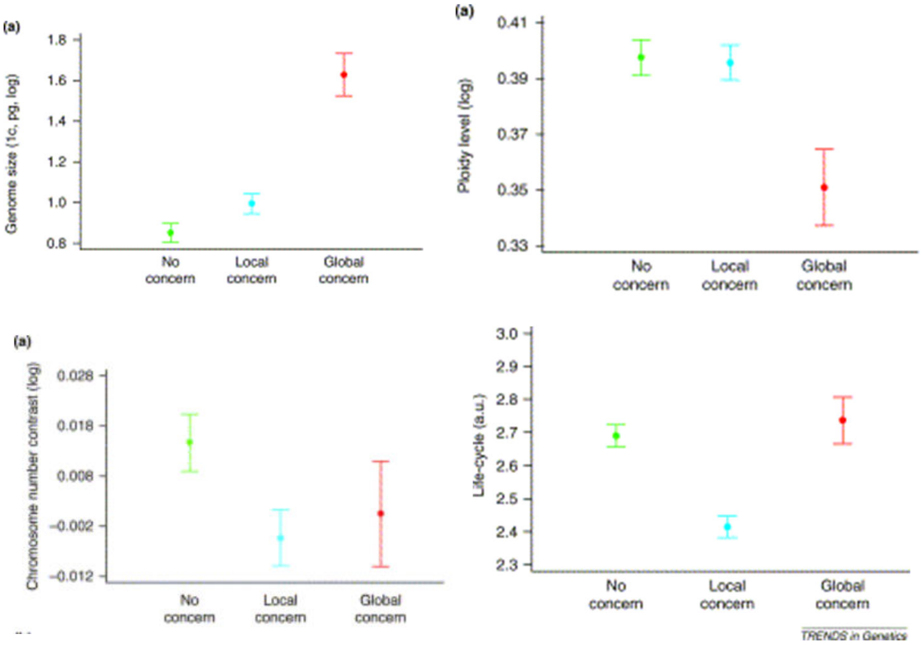
3103 herbaceous angiosperms

733 woody angiosperms

181 woody gymnosperms

DNA content vs fitness

Vinogradov 2003

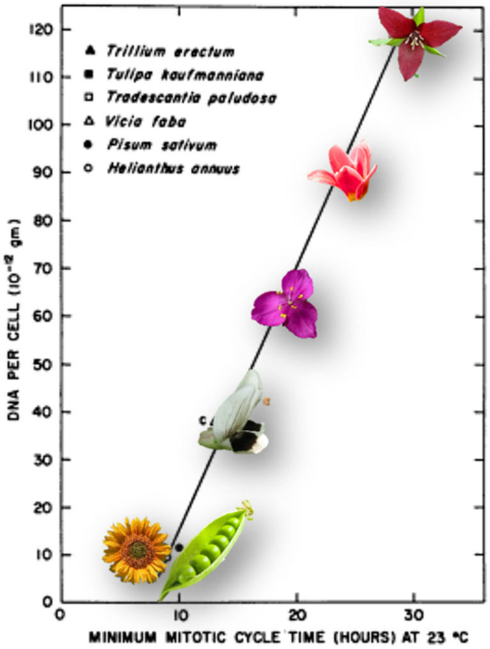


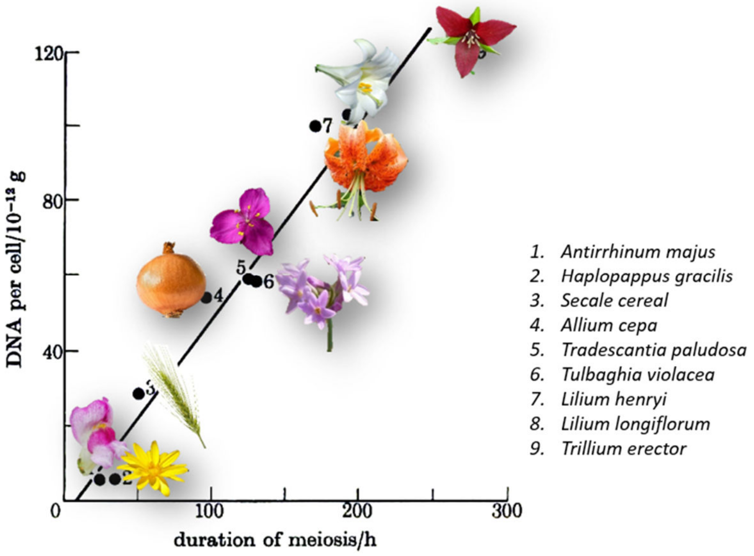
The nucleotype

Bennett, 1972

DNA content vs cycle time

Van't Hopf and Sparrow, 1963; Bennett, 1971

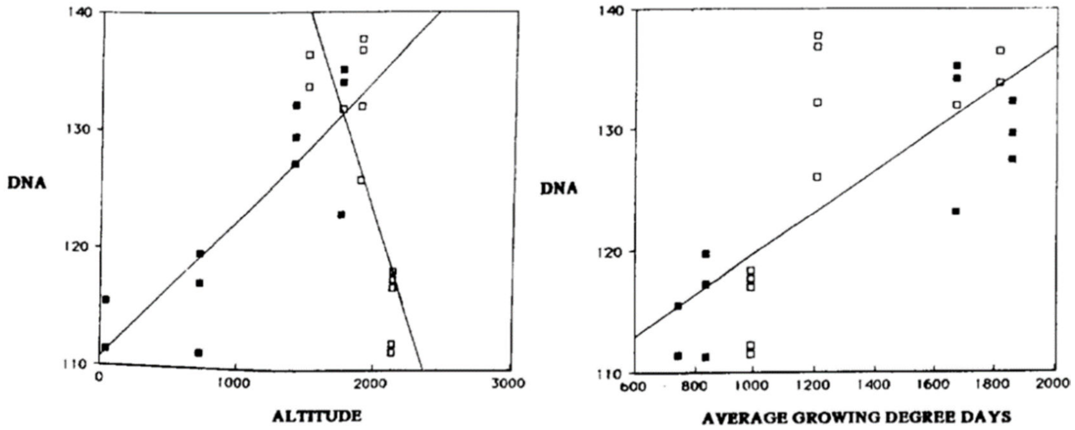




DNA content vs growth

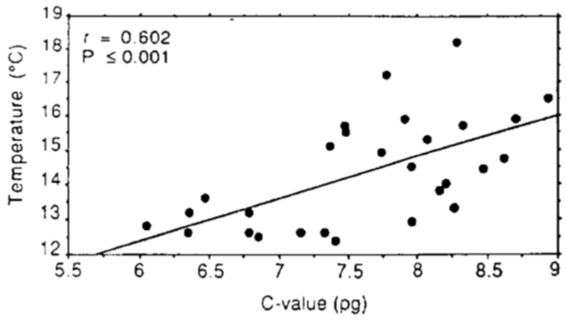
Corn vs altitude

Bullock and Rayburn, 1991; **McMurphy and Rayburn, 1992**



Tall fescue in Italy

Ceccarelli et al, 1992



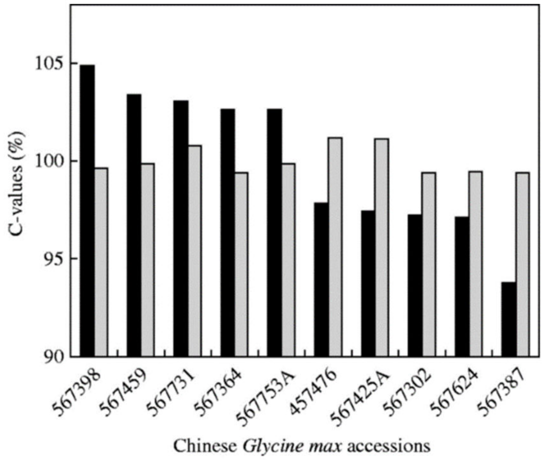
Knight et al., 2005

	# Studies Showing		
Variable	**+ Correlation to 2C DNA Content**	**- Correlation to 2C DNA Content**	**Non-linear or Non-significant Correlation**
Altitude	8	9	7
Latitude	5	8	6
Temperature	4	9	2
Precipitation	4	3	1
Seed Mass	11	0	1
Generation Time	12	3	4
Relative Growth Rate	5	4	2

Measurement errors

Greilhuber 2005

Figure 4. Greilhuber 2005 data in grey, super-imposed on Rayburn data (black)



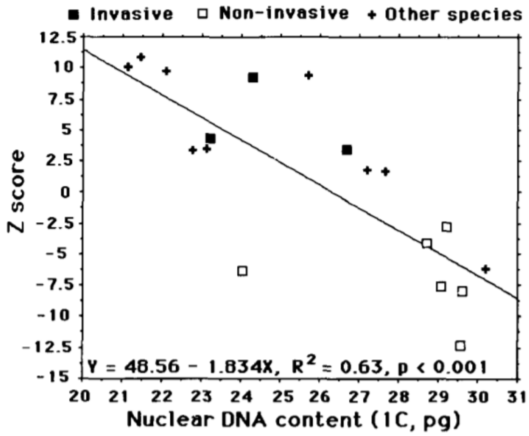
Rayburn et al, 2004

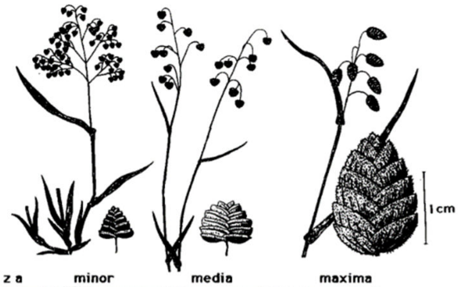
Line	# Plants	Relative mean	Standard deviation
PI 266085	22	2.847	0.093
PI 227324	22	2.825	0.064
PI 437088	22	2.814	0.071
PI 253666	22	2.814	0.076

DNA content vs invasiveness

Rejmánek, 1996

Review by Suda et al, 2014

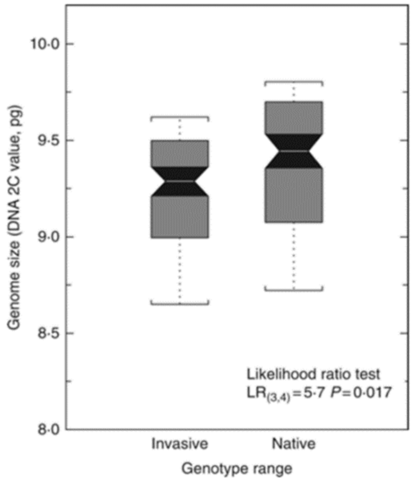




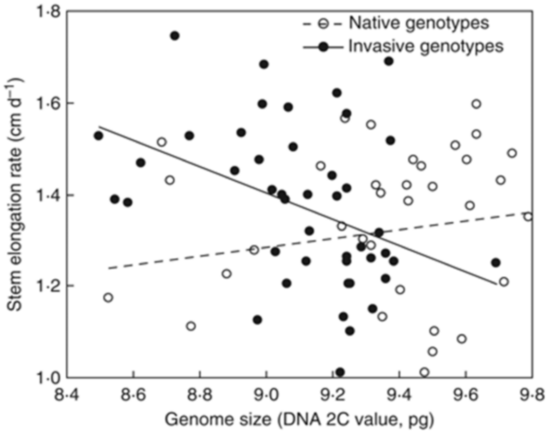
Briza	minor	media	maxima
2C DNA (pg)	14.6	16.9	21.6
Cariopsis length (mm)	0.8–1.0	1.3–1.7	2.5–3.0
Mean number of seeds/plant	329.0	215.8	93.6
No. of states	16	9	≤5

Reed canarygrass

Lavergne et al., 2010



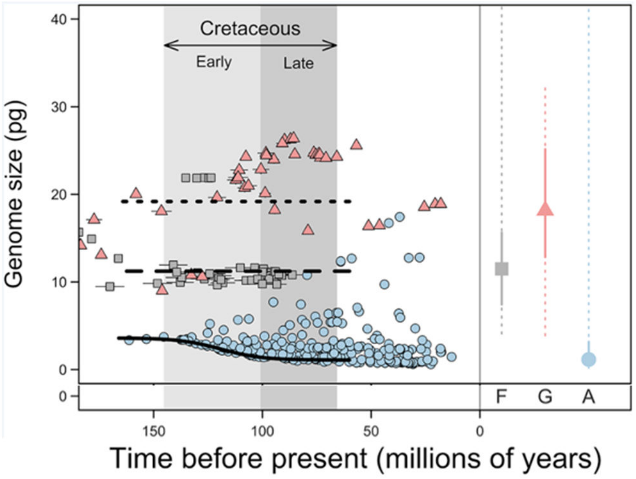
Average DNA content of invasive and non invasive genotypes

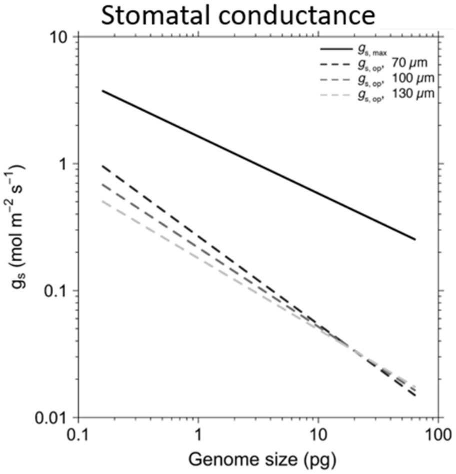


Correlation between stem growth rate and DNA content. These parameters were not significantly correlated in native (non invasive) genotypes, but very negatively correlated with invasive genotypes.

DNA content vs rise of angiosperms

Simonin & Roddy, 2018

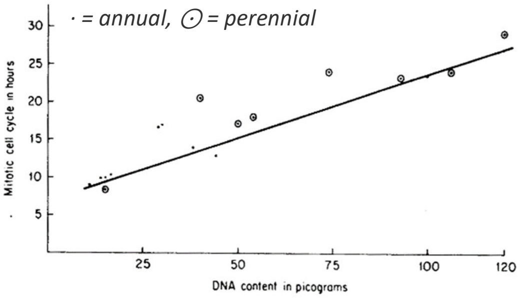




DNA content vs life cycle

Swanson, Merz, & Young, 1981

Species	pg/haploid genome	Mitosis (hours)	Meiosis (hours)	Plant habit
Crepis capillaris	1.20	10.8	--	Annual
Haplopappus gracilis	1.85	10.5	36.0	Annual
Pisum sativum	3.9, 4.8	10.8	--	Annual
Ornithogalum virens	6.43	--	96.0	Perennial
Secale cereale	8.8, 9.6	12.8	51.2	Annual
Vicia faba	13.0, 14.8	13.0	72.0	Annual
Allium cepa	14.8, 16.25	17.4	72.0	Perennial
Tradescantia paludosa	18.0	18.0	126.0	Perennial
Endymion nonscriptus	21.8	--	48.0	Perennial
Tulipa kaufmanniana	31.2	23.0	--	Perennial
Lilium longiflorum	35.3	24.0	192.0	Perennial
Trillium erectum	40.0	29.0	274.0	Perennial



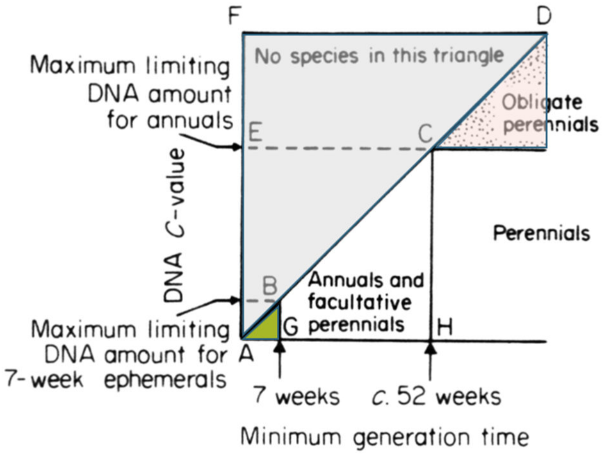
Definitions

Bennett, 1987

- **Annual**
    - Sets seed and completes life cycle within one year (52 weeks) or shorter growth season.
    - Hence, a fast growth rate is essentia
- **Perennial**
    - Does not complete life cycle within 1 year
- **Facultative perennial**
    - Sets seed within 1 year
- **Nucleotypically determined obligate perennial**
    - The DNA content slows down the cell cycles to where the plant cannot set seed or complete its life cycle within one year
- **Genotypically determined obligate perennial**
    - Long life cycle not determined by DNA content, such as juvenility in trees
- **Temperature-determined obligate perennials**
    - Extreme environments with growing seasons too short to complete a life cycle in 1 year

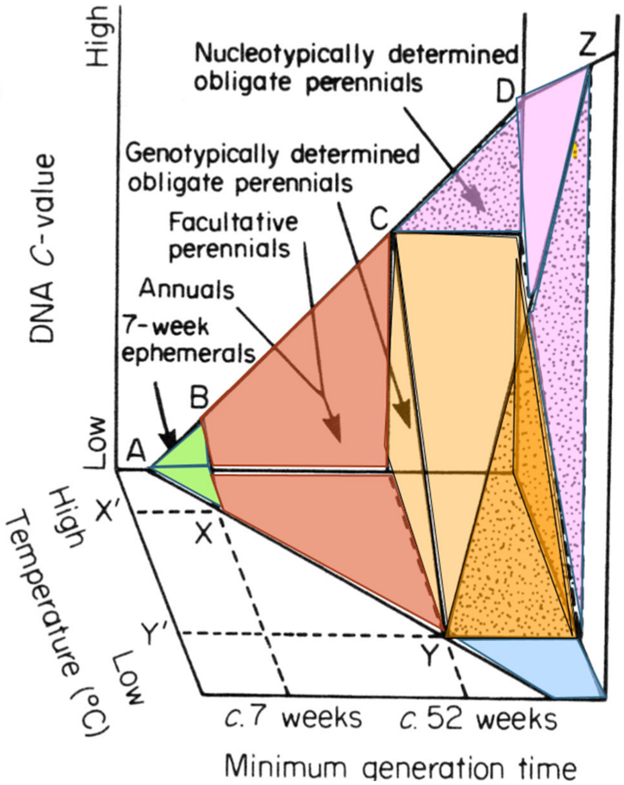
Effect of DNA @ constant temperature

Bennett, 1987



Effect of DNA considering temperature

Bennett, 1987



Category of Plant	DNA C-value is below the segment:	Minimum generation time is:	The plants can exist at temperatures:
Ephemerals	AB	≤7 weeks	>X′
Annuals and Facultative Perennials	BC	≤52 weeks but >7 weeks	>Y′
Genotypically- determined obligate perennials	C	>52 weeks	>Y′
Nucleotypically-determined obligate perennials	CD	>52 weeks	>Y′
Temperature-determined obligate perennials	can be as great as CD but tends to be lower	>52 weeks	<Y′