

Genomica Vegetale

2020-21

Simone Ferrari

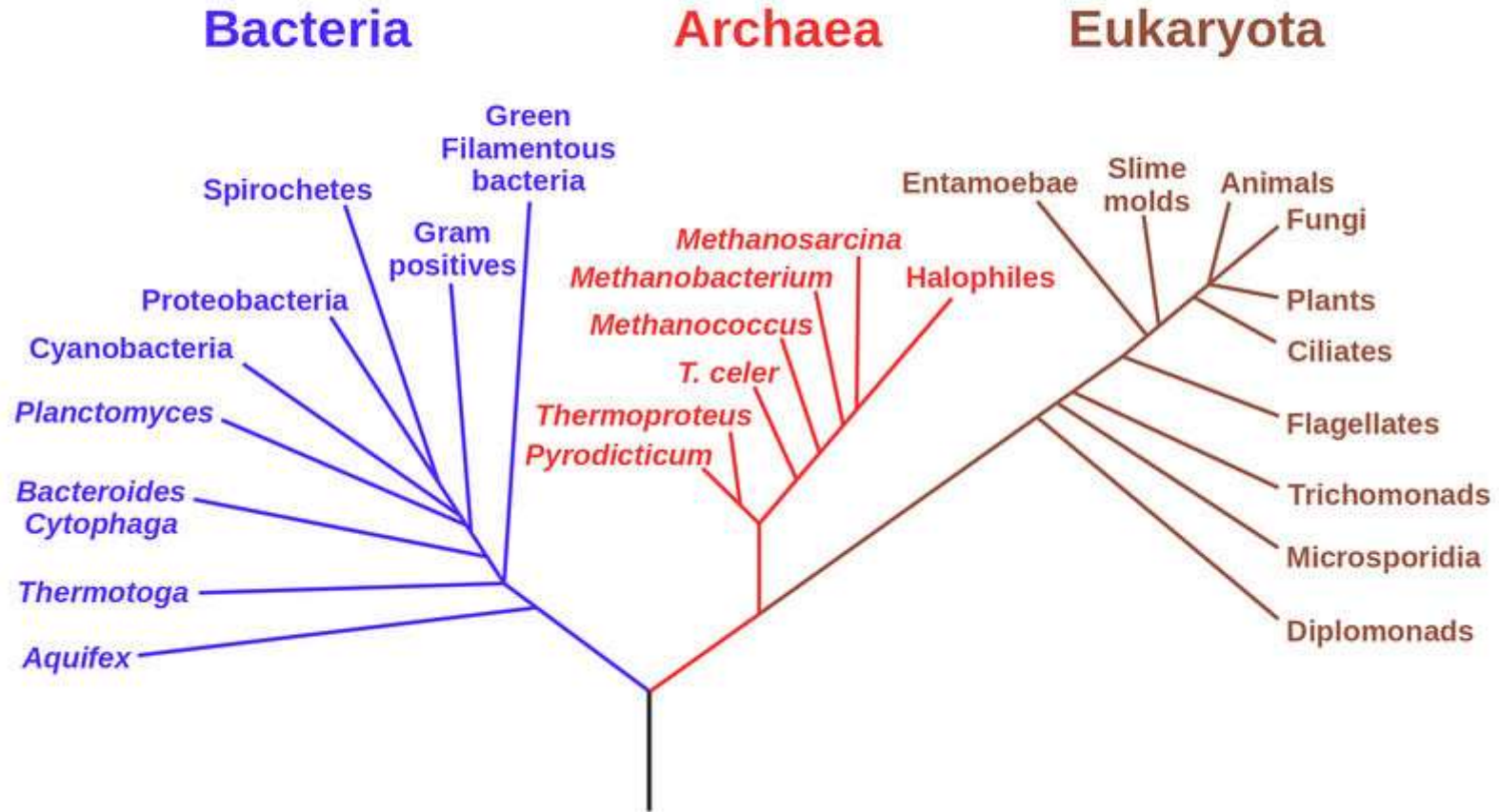
Simone.ferrari@uniroma1.it

Federica Taddei

federica.taddei@crea.gov.it

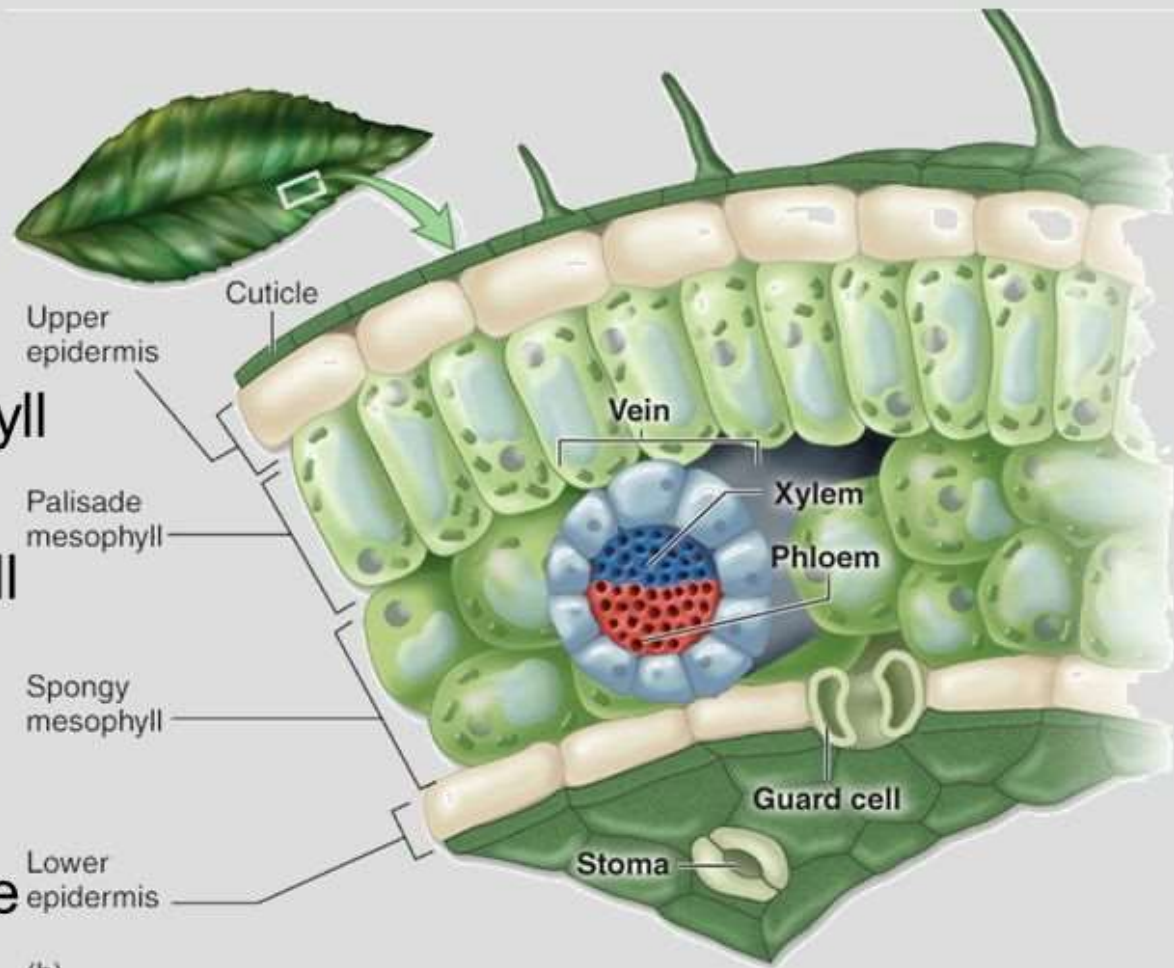
<https://elearning.uniroma1.it/mod/page/view.php?id=371427>

Phylogenetic Tree of Life

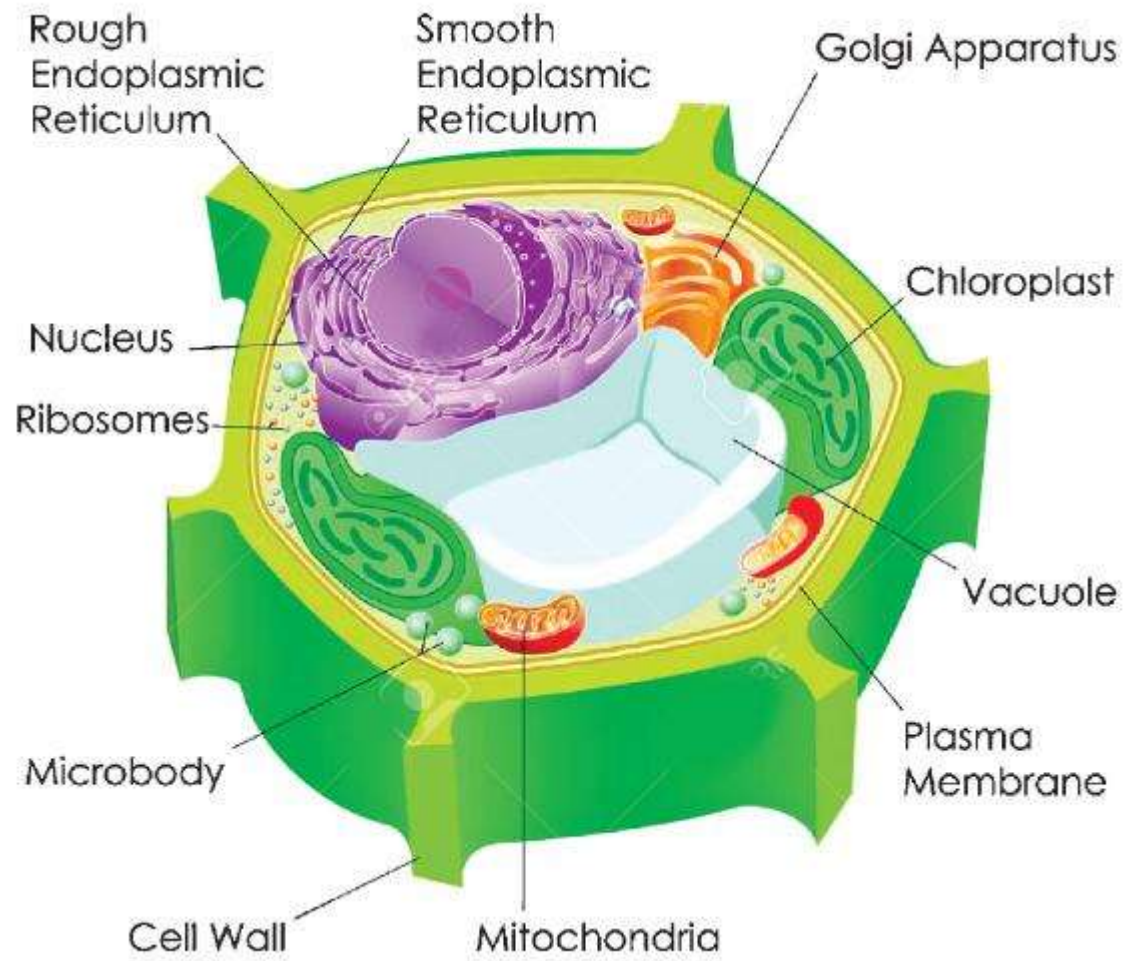


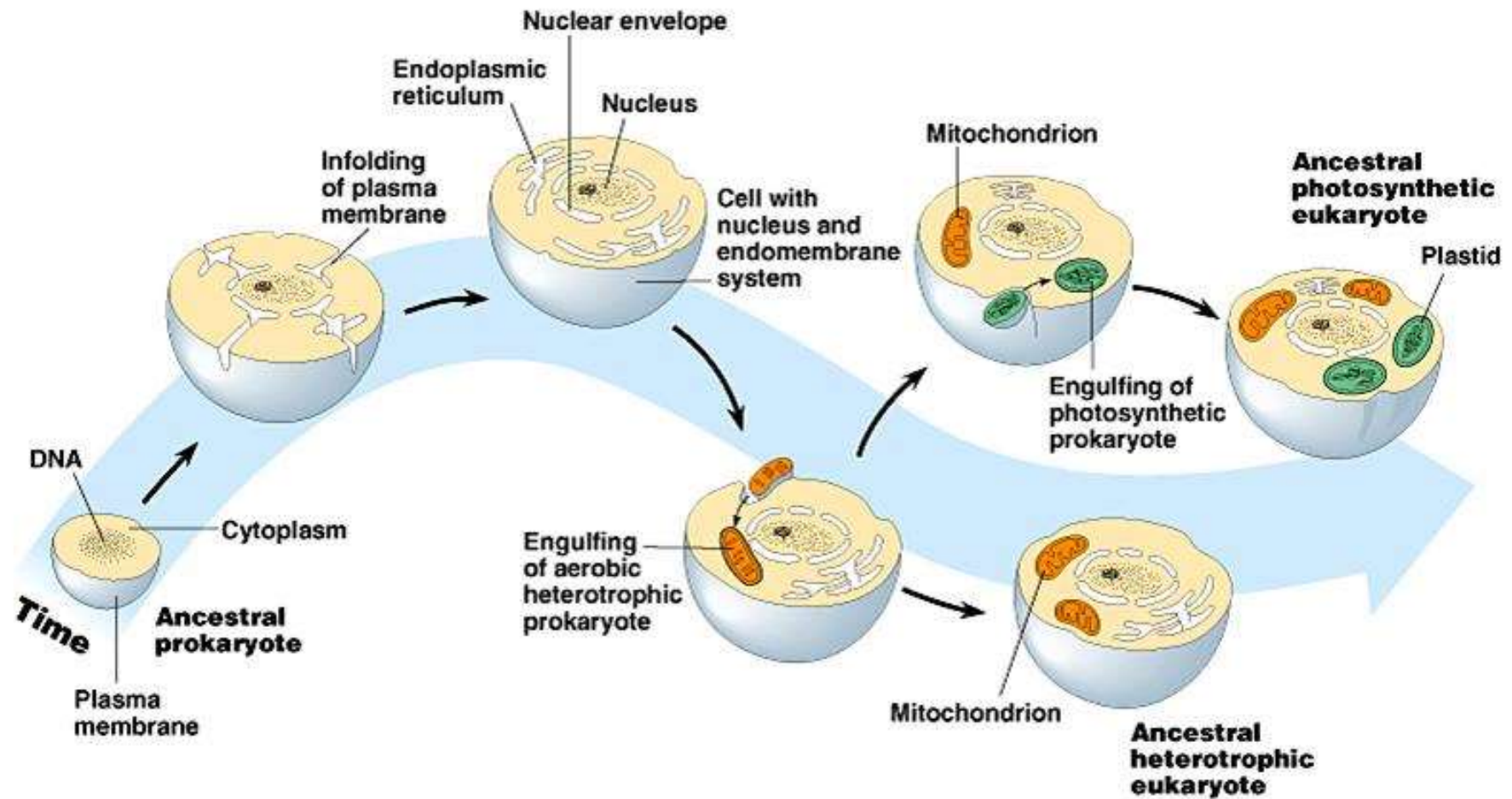
Leaves – Structure and Development

- Most dicots have 2 types of mesophyll
 - Palisade mesophyll
high photosynthesis
 - Spongy mesophyll
air spaces for gas & water exchange
- Monocot leaves have 1 type of mesophyll



(b)





NONVASCULAR BRYOPHYTES

VASCULAR SEEDLESS PLANTS

VASCULAR SEED PLANTS



Hornworts

Liverworts

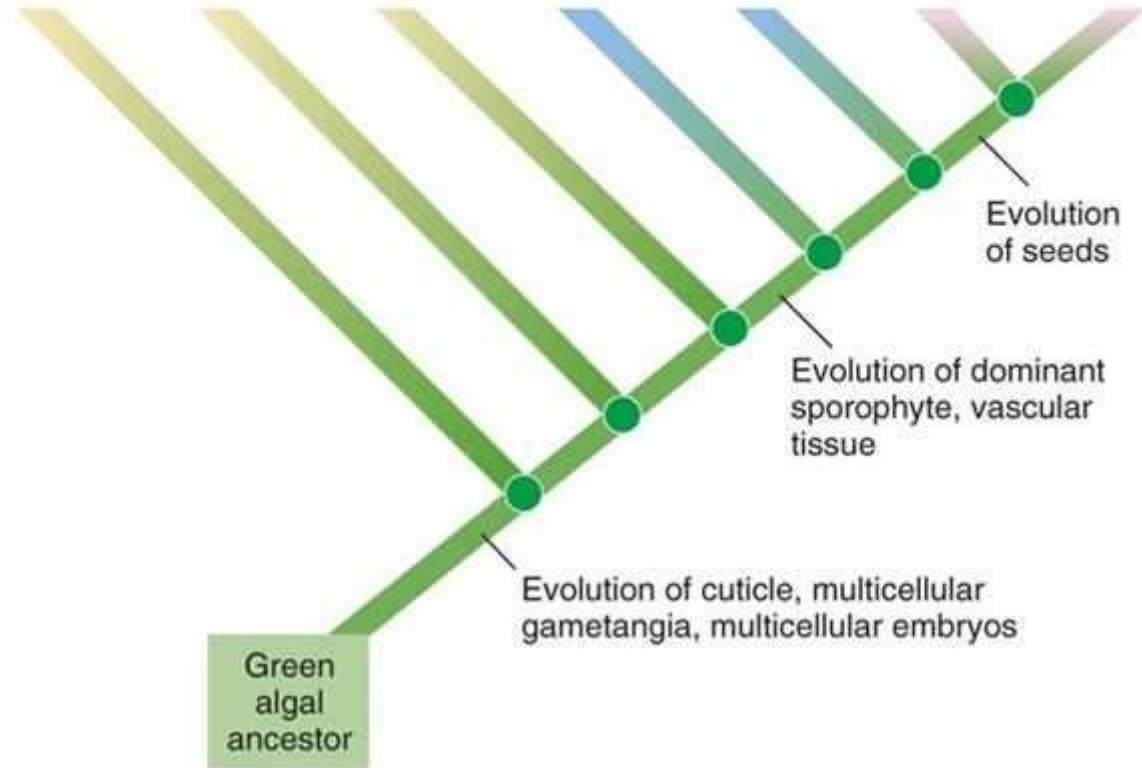
Mosses

Club mosses

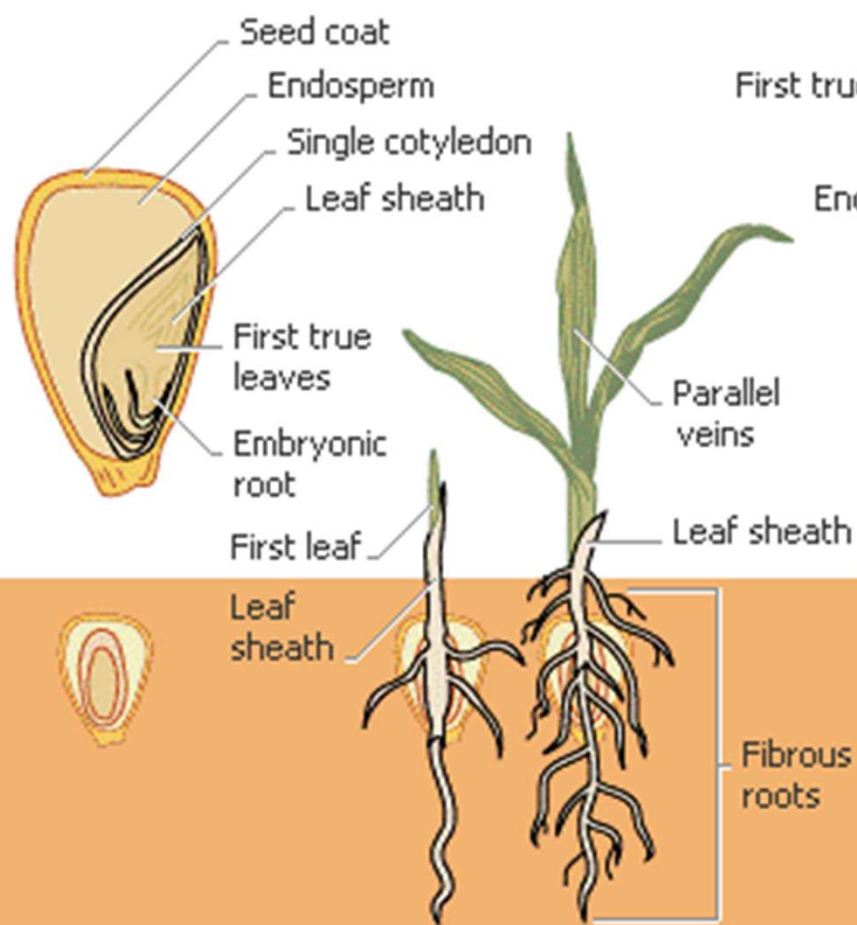
Ferns

Gymnosperms

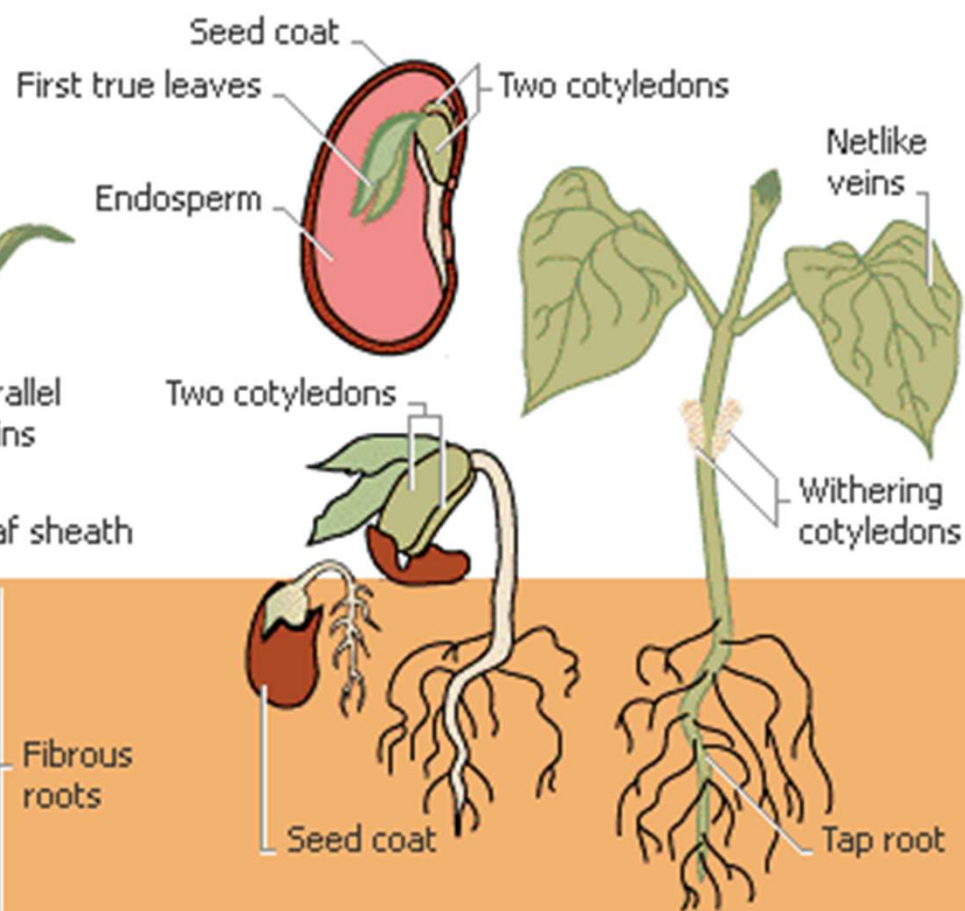
Angiosperms



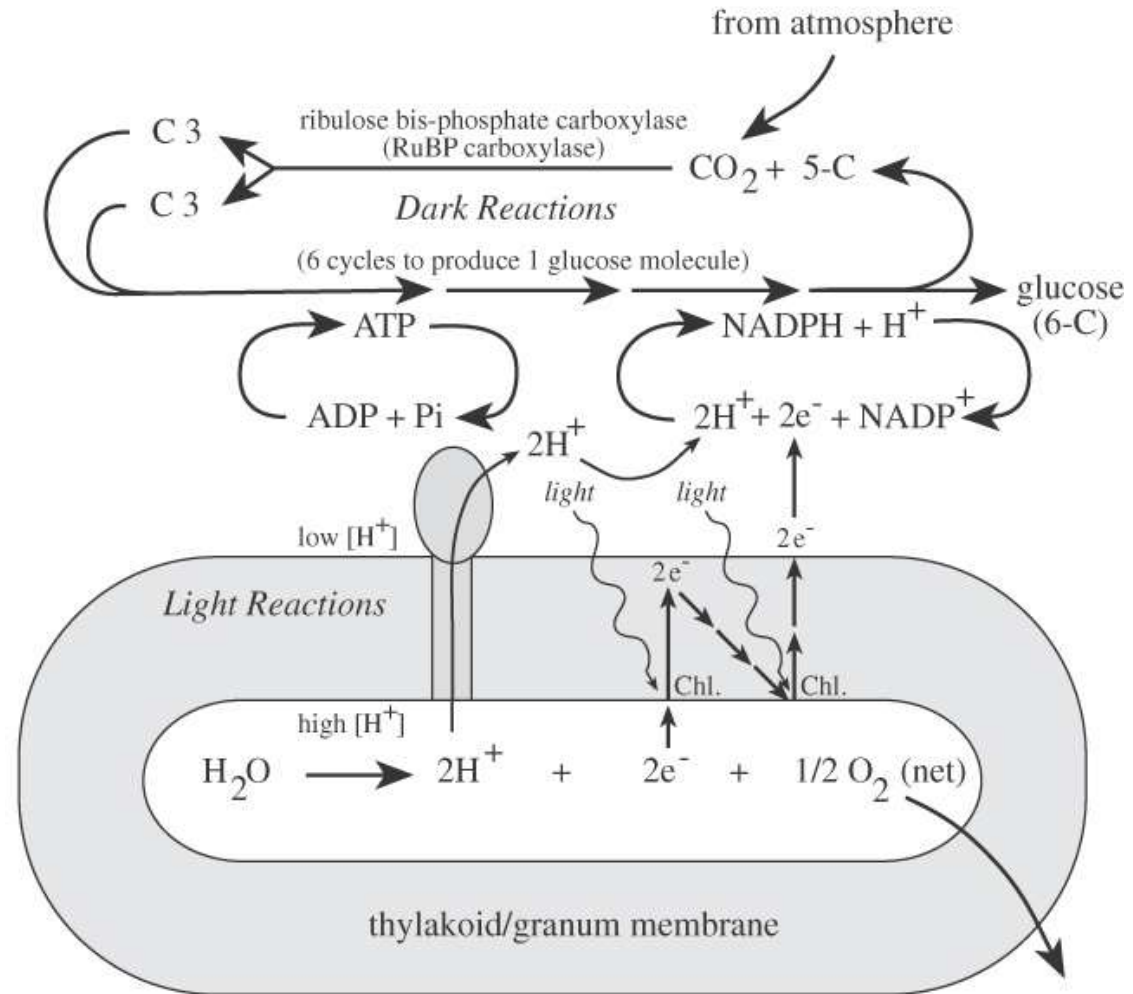
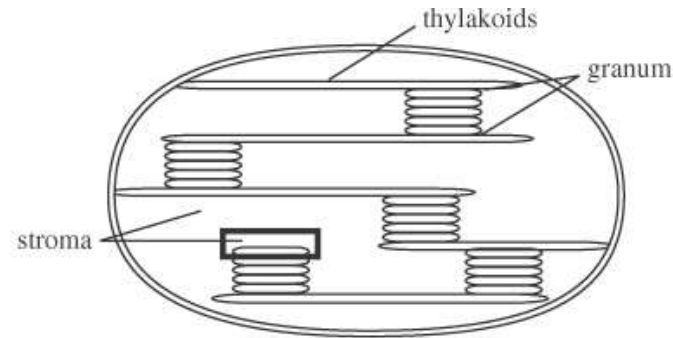
Monocotyledon (corn)



Dicotyledon (bean)



PHOTOSYNTHESIS



Le piante hanno organi di riproduzione pluricellulari detti gametangi (anteridi i maschili, archegoni quelli femminili).

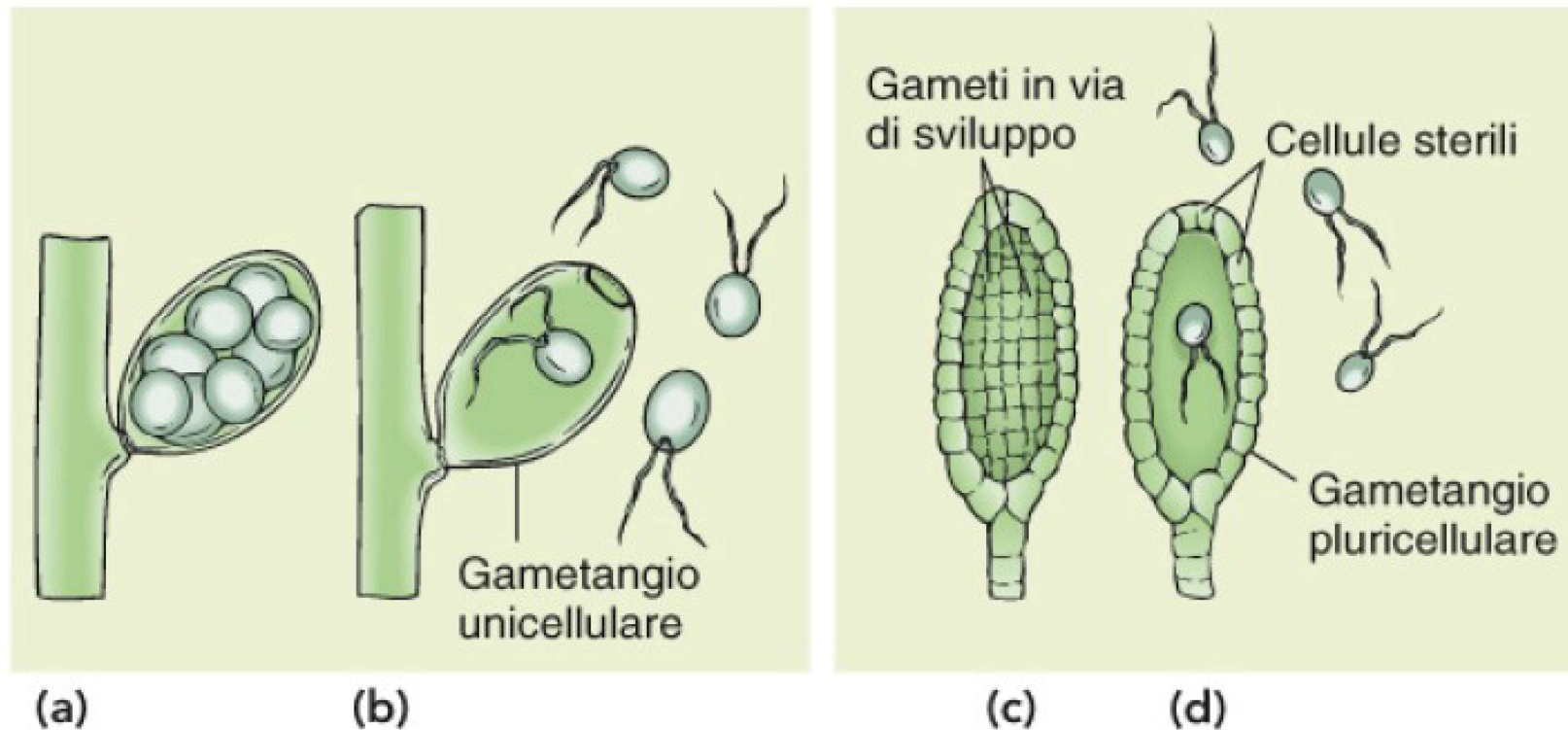
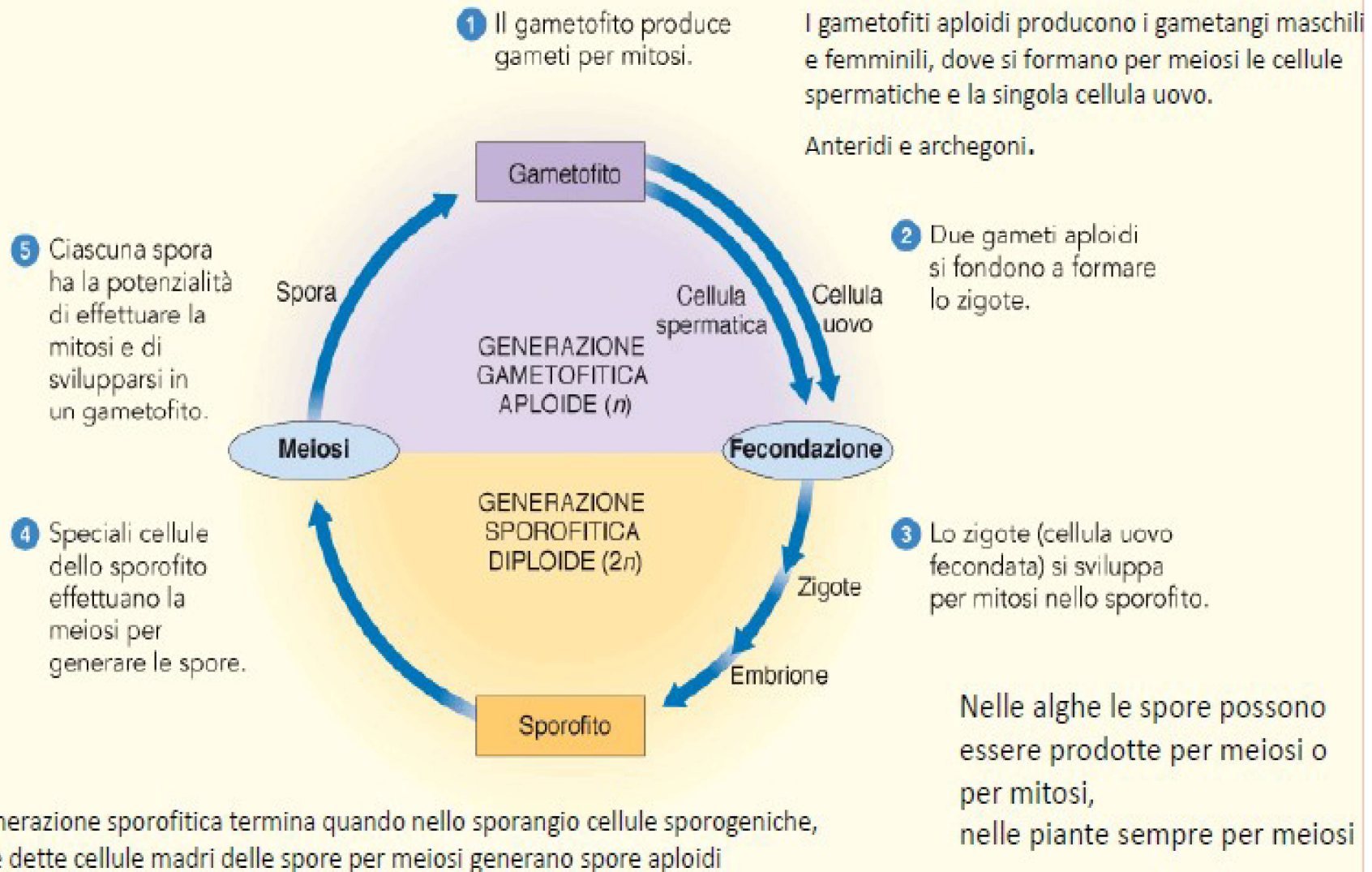


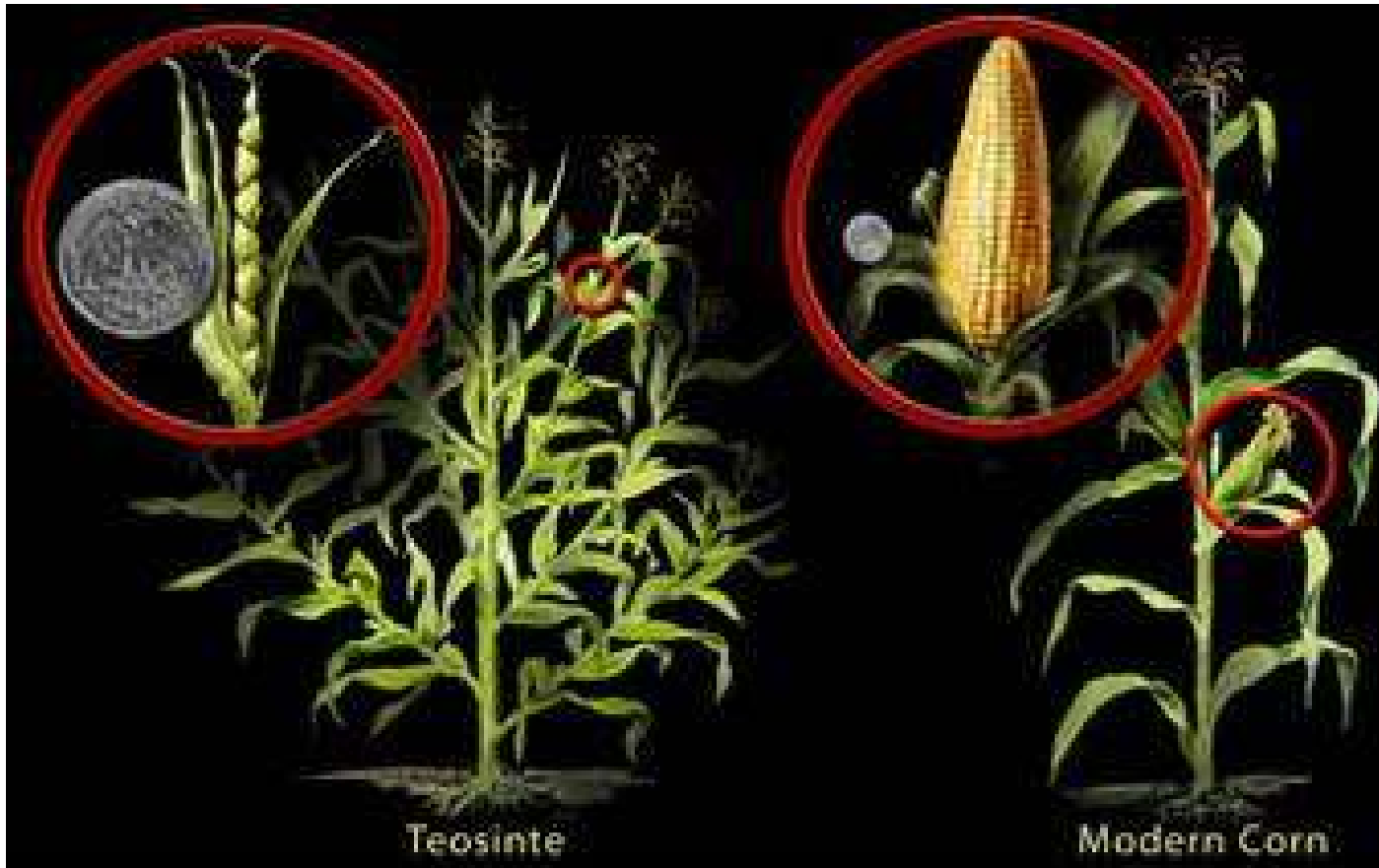
FIGURA 27-1 Strutture riproduttive generalizzate di alghe e piante

(a, b) Nelle alghe, i gametangi sono generalmente unicellulari. Con il rilascio dei gameti, della cellula originale rimane solo la parete cellulare. (c, d) Nelle piante, i gametangi sono invece pluricellulari, ma solo le cellule interne diventano gameti. I gameti sono circondati da uno strato protettivo di cellule sterili (non riproduttive).

Le piante sono caratterizzate da un'alternanza di generazioni, gametofitica aploide e sporofitica diploide.



Plant domestication



PLANT GENOMICS

- Identification of genes important for agronomic traits
- Evolution of plants

Why do we need a plant model?

- Plants constitute over 90% of the world's biomass; 250,000 species of flowering plant
- Important biological processes are plant specific; photosynthesis that fixes carbon and produces oxygen
- Plants are economically important; in agriculture or in secondary metabolites as medicines and in nutrition
- Plants evolved multicellularity independently, and use different mechanisms of cell to cell communication
- Plants represent important genetic model systems; transposons and gene silencing were first identified in plants.

MODEL PLANTS

Arabidopsis thaliana

Dicot (*Brassicaceae*)

Small diploid genome (C1 = 125 Mbp)

Easy to transform

5 chromosomes

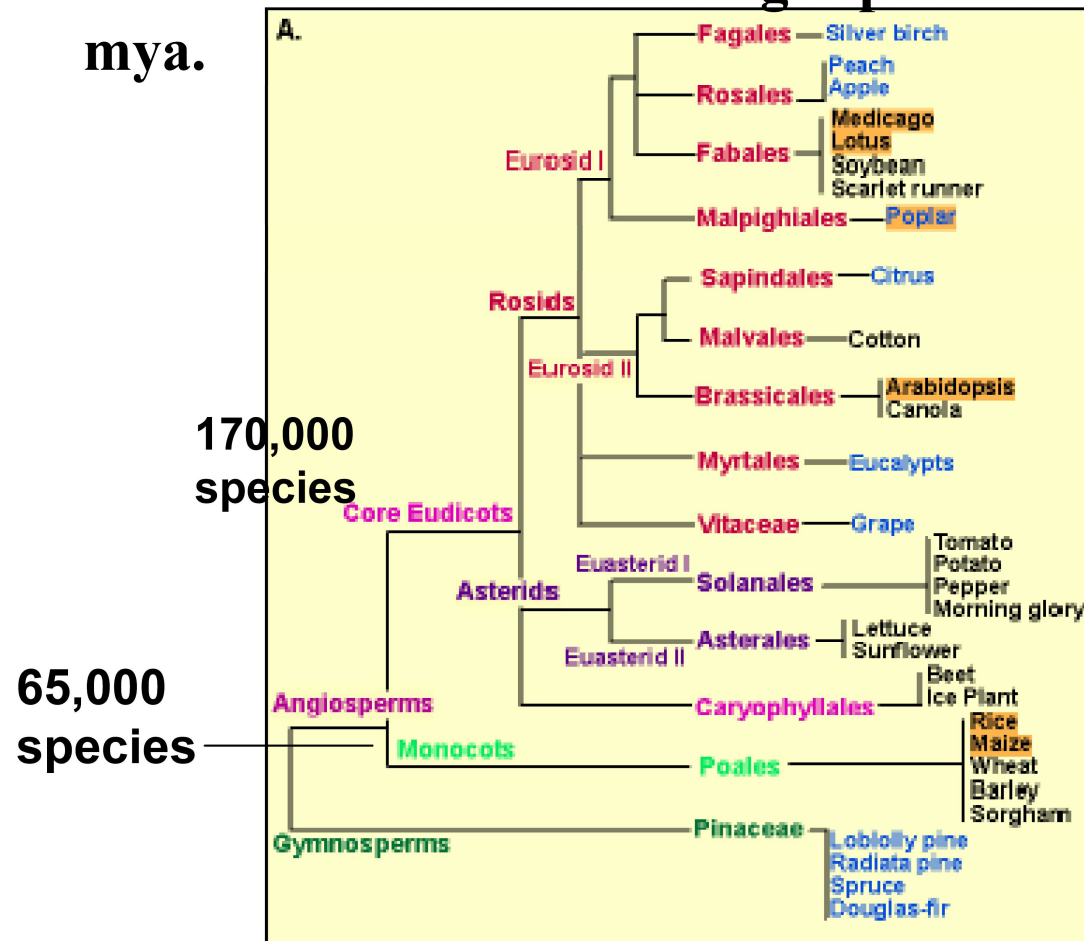
Small size

Short life cycle (2 months)

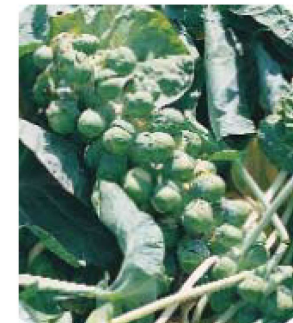


Arabidopsis is a model for Eudicot seed plants

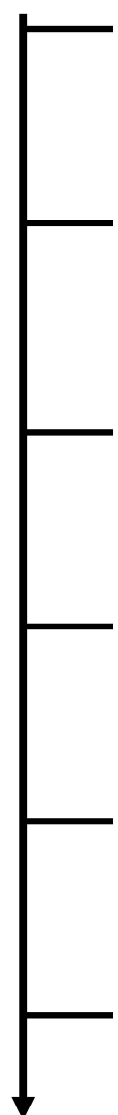
Flowering plants originated relatively recently and the last common ancestor of all Angiosperms lived around 180 mya.



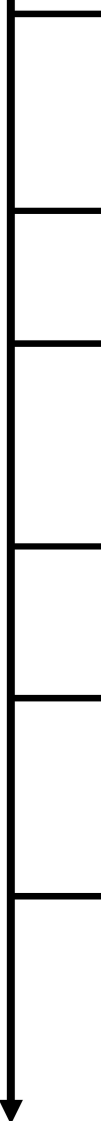
Close relatives of Arabidopsis



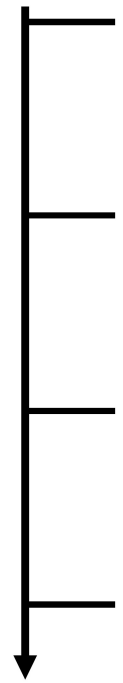
Timeline – important advances in Arabidopsis research (1)

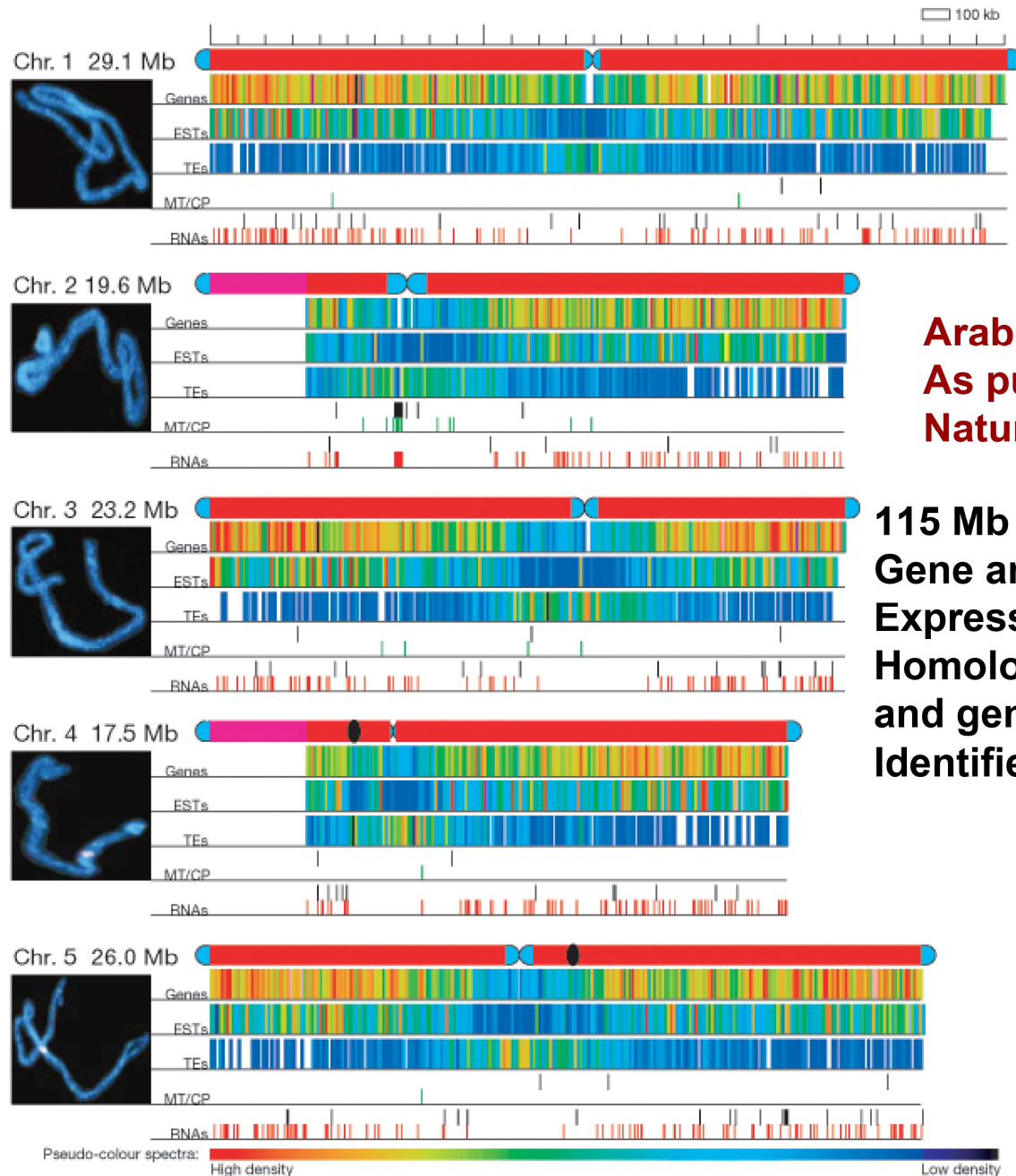
- 
- 1907. Arabidopsis first used by Laibach for cytology. Showed 5 chromosomes during his PhD in Bonn.**
 - 1943. Laibach described usefulness of Arabidopsis for studying genetics of phenotypes such as variation in flowering time, while Prof. of Botany in Frankfurt.**
 - 1947. Laibach's student, Erna Reinholz, isolated first mutants of Arabidopsis using X-rays.**
 - 1965. First Arabidopsis conference held in Göttingen, 25 people attended.**
 - 1983. First genetic map of Arabidopsis with genetic linkage groups covering all five chromosomes made by Maarten Koornneef.**
 - 1984. Arabidopsis DNA characterized using lambda libraries. Genome size estimated at 70 Mb, by Elliot Meyerowitz.**

Timeline – important advances in Arabidopsis research (2)

- 
- 1986. Transgenic Arabidopsis plants generated. Regeneration of transformed plants from roots most widely used method.**
 - 1988. First restriction fragment length polymorphism map made.**
 - 1989. Cloning of first gene by insertional mutagenesis. T-DNA of *Agrobacterium tumefaciens* as the mutagen.**
 - 1992. First Arabidopsis genes isolated by positional cloning.**
 - 1993. High efficiency transformation established by vacuum infiltration of *Agrobacterium* cultures into plant tissues.**
 - 1997. Physical map of Arabidopsis genome completed. Whole genome in overlapping bacterial artificial chromosomes or yeast artificial chromosomes.**

Timeline – important advances in Arabidopsis research (3)

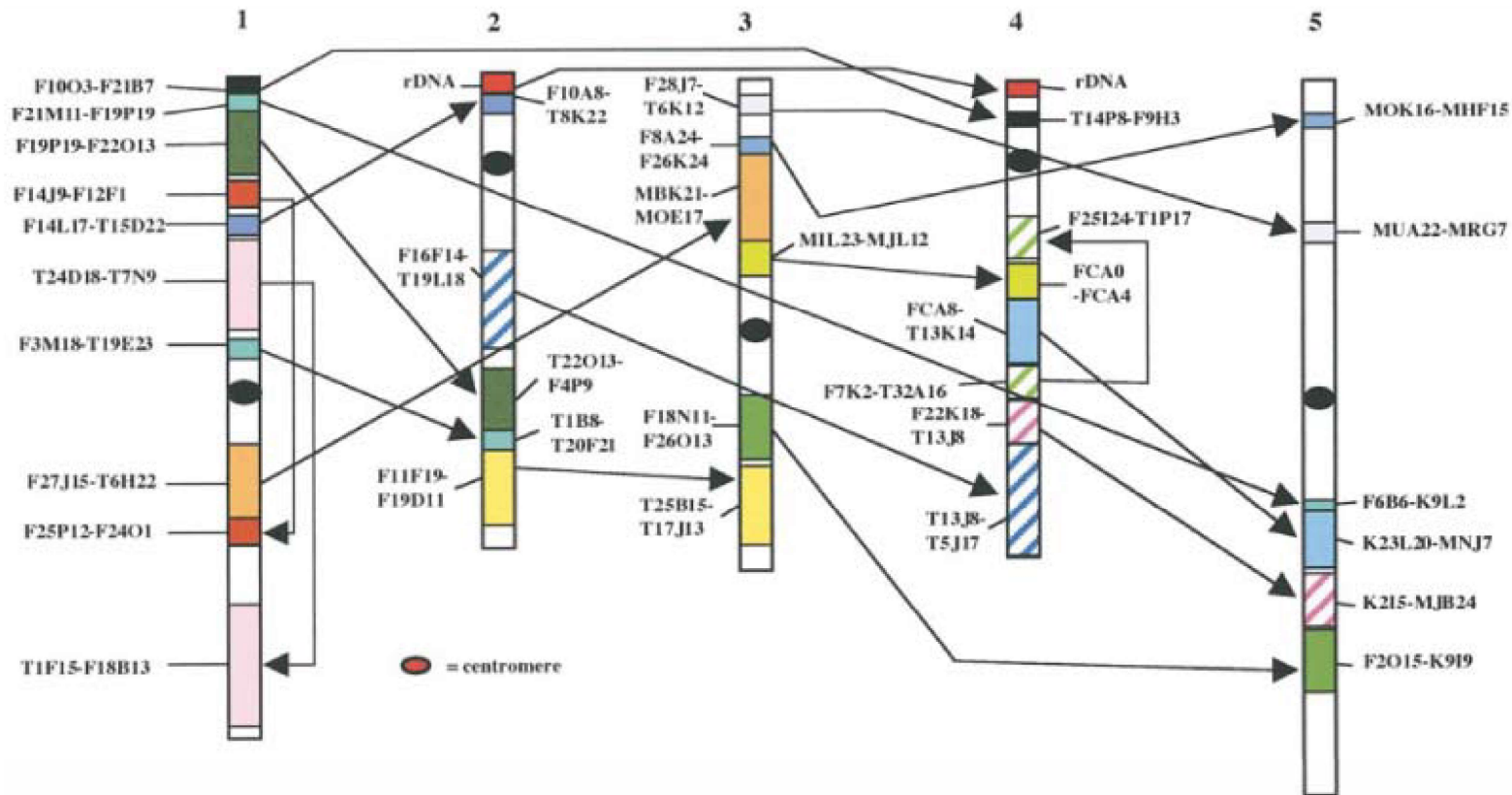
- 
- 2000. Paper describing completion of main phase of sequencing the Arabidopsis genome appears in *Nature*.**
 - 2002. Availability of Affymetrix microarrays allowing the Simultaneous analysis of all known Arabidopsis genes.**
 - 2003. Availability of over 330,000 insertions at precisely sequenced locations. Provides genome-wide resources for reverse genetics with insertions in 90% of genes.**
 - 2004. 15th International Arabidopsis conference held in Berlin. 1100 people attended.**



**Arabidopsis genome sequence.
As published in 2000.
Nature 408, 796.**

**115 Mb of 125 Mb genome.
Gene annotation using
Expressed sequence tags (ESTs)
Homology with cloned plant genes
and genes of other organisms
Identified 25,500 genes.**

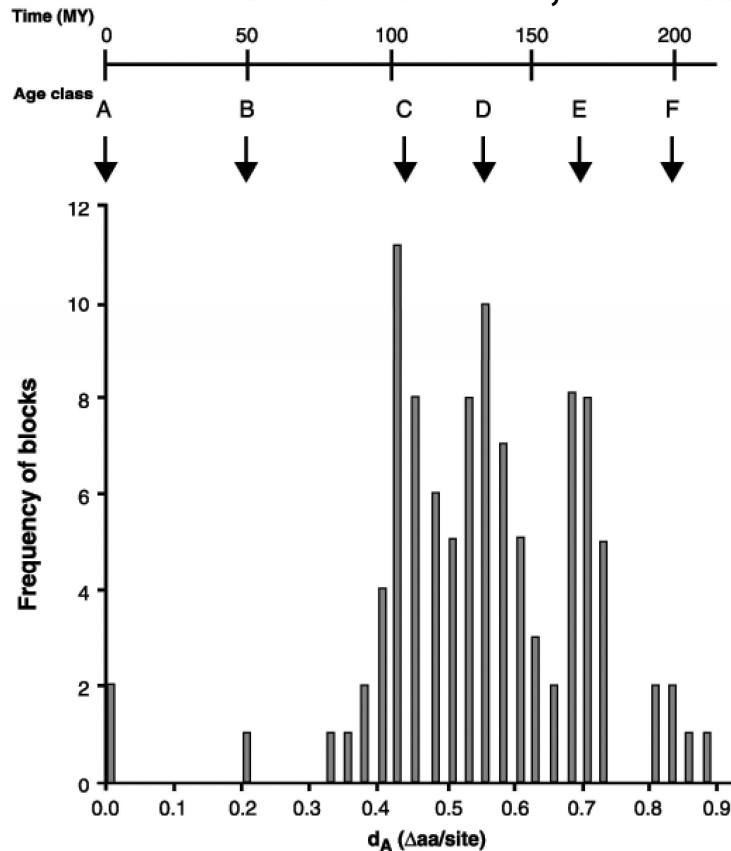
Large segments of the Arabidopsis genome are duplicated



Origin of genomic duplications in Arabidopsis

- 103 duplicated blocks containing 7 or more genes
- Over 81% of ORFs fall within the bounds of a block, but only 28% of genes are present in duplicate due to extensive deletions of genes

Number of duplicated genes, suggests that the whole genome may have been duplicated, then expect all duplicated genes to have diverged to a similar extent. However, fall into three major age classes:



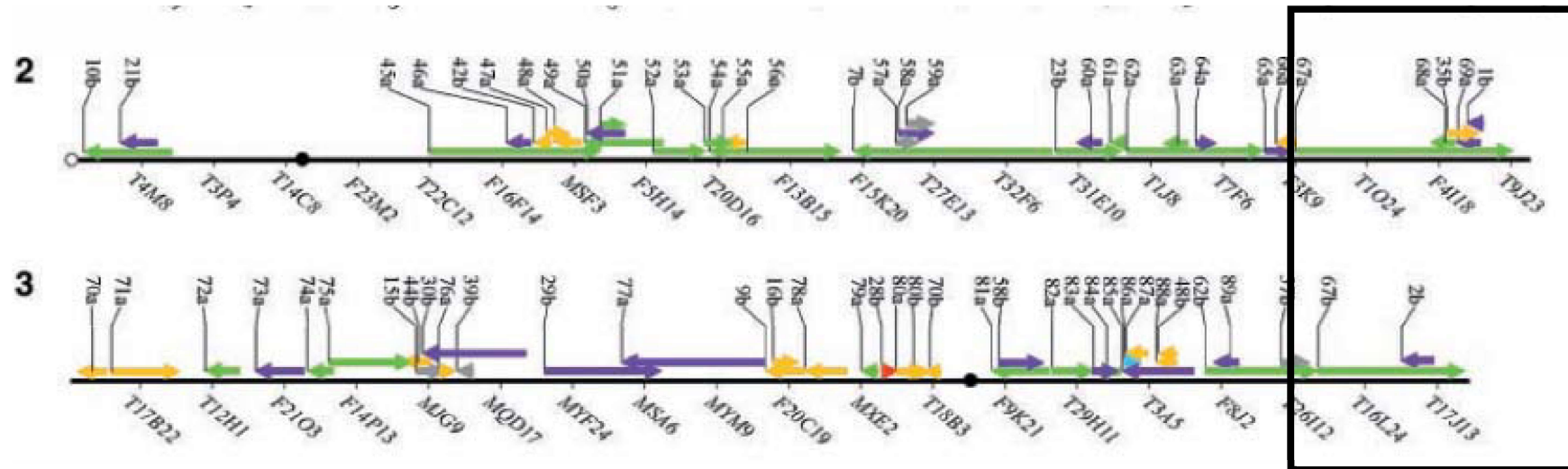
C – 48% of genes bounded;
probably represents whole
genome polyploidization.

D – 39% genes

E – 11% genes

F - 3% genes

Genetic redundancy can exist between genes in duplicated blocks



A duplicated block of genes exists on chromosomes 2 and 3.

One of the duplicated genes encodes a MADS box transcription factor, and the proteins encoded by the two genes are 87% identical at the amino acid level.



SHATTERPROOF 1

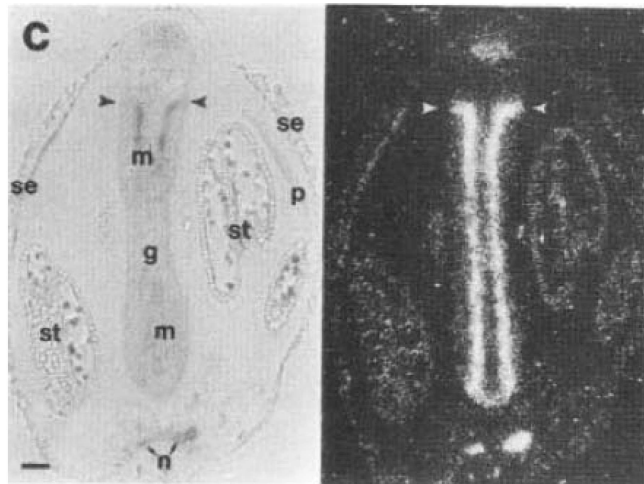
100% identical in MADS DNA binding domain



SHATTERPROOF2

SHP1 and SHP2 are expressed in similar patterns in the the developing Arabidopsis fruit

Expression of SHP1 in young flower bud in the developing fruit

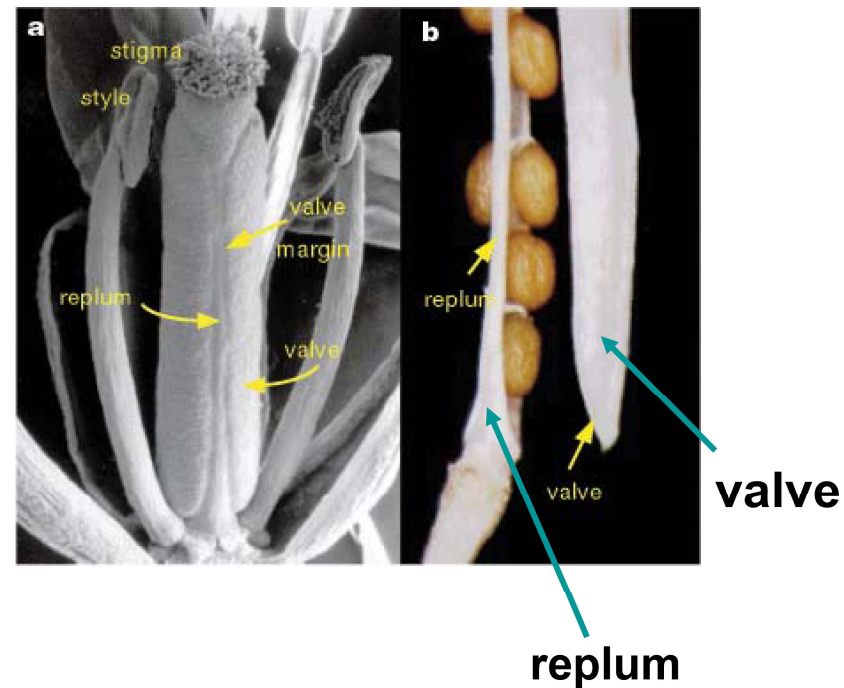


Bright field

In situ hybridization

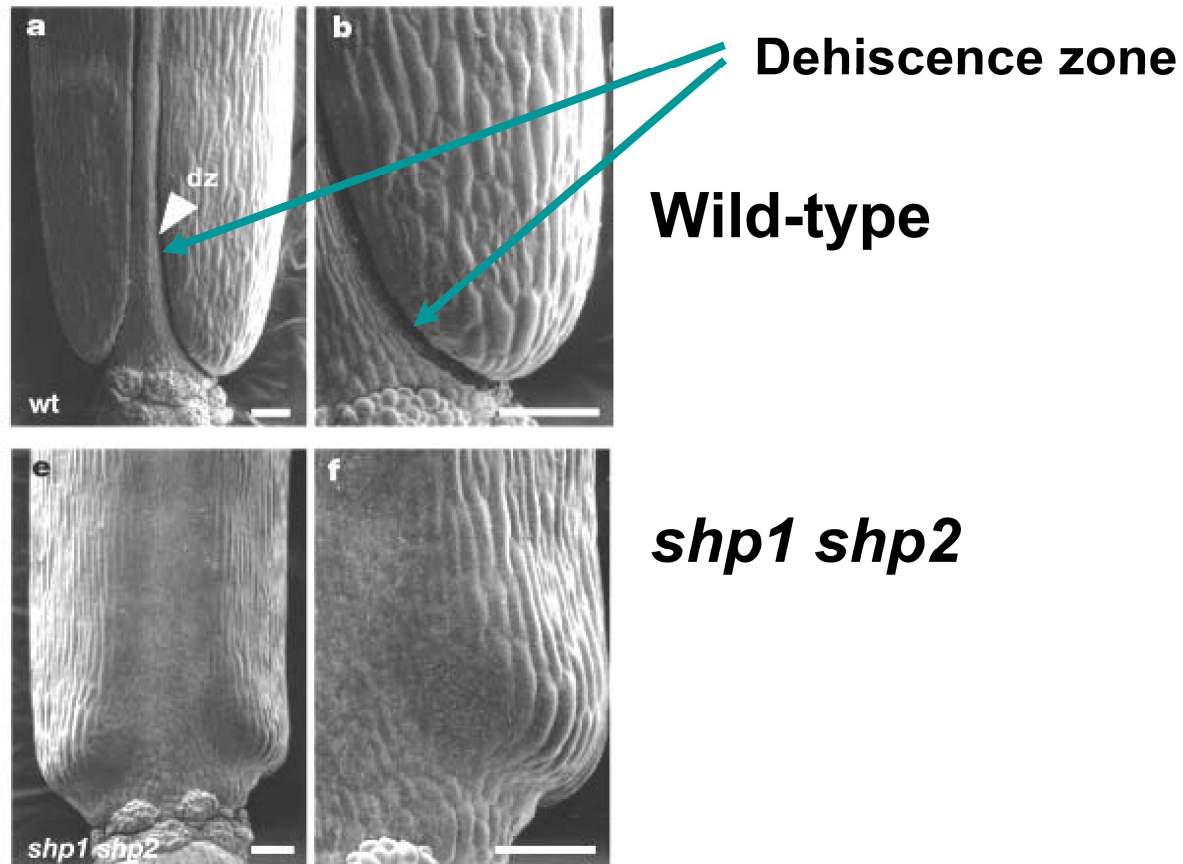
SHP2 expressed in a similar pattern

Wild-type flower



Valve separates from replum to release the seeds

SHP1 and *SHP2* are genetically redundant

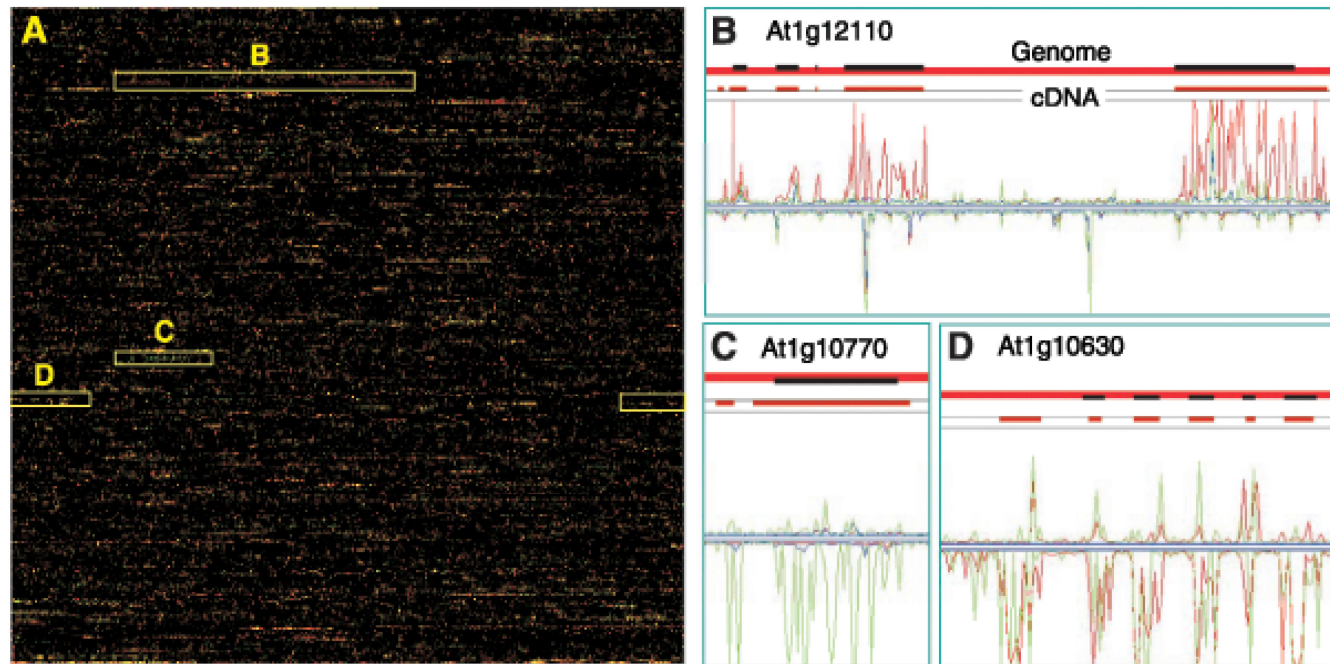


shp1 or shp2 single mutants show no phenotype, but the fruit of the double mutant is impaired in dehiscence.

Improved annotation of the Arabidopsis genome sequence

Initial sequence analysis relied heavily on expressed sequence tags (ESTs) and gene prediction programmes. Interpretation of genome sequence constantly improved.

Isolation of full-length cDNAs and use of whole genome oligonucleotide tiling arrays greatly improved the annotation. (32% of predicted gene models incorrect).



Arabidopsis genome represented on 12 oligonucleotide arrays.

Each array 834,000 25-mer oligonucleotides.

Is hybridized with cDNA made from 4 RNA populations; seedlings, roots, flowers, cultured cells. Sequences that hybridize are mapped directly onto genome sequence.


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

The screenshot shows a web browser window displaying the TAIR (The Arabidopsis Information Resource) homepage. The browser's address bar shows the URL <http://www.arabidopsis.org/>. The page features a navigation menu with links for Home, Help, Contact, About Us, and Login/Register. Below this is a search bar with a dropdown menu set to 'Gene' and a 'Search' button. A secondary navigation bar includes links for Search, Browse, Tools, Portals, Download, Submit, News, and ABRC Stocks. The main content area is titled 'The Arabidopsis Information Resource' and contains several paragraphs of text describing the database and its resources. On the right side, there are social media links for RSS, Twitter, and Facebook, followed by a 'Breaking News' section with two news items: 'ABRC tool for adding comments about stocks' and 'Notice from ABRC for Chinese Researchers'. The page also features logos for Carnegie Institution for Science and a small sun icon.

TAIR - Home Page

File Modifica Visualizza Preferiti Strumenti ?

Autenticazione Google Ap... [S] Scheda attività didattica S... MIUR - loginmiur.cineca Google blast Basic Local Alignme... Home - PubMed - NCBI Raccolta Web Slice


 Home Help Contact About Us Login/Register

Search Browse Tools Portals Download Submit News ABRC Stocks


The Arabidopsis Information Resource

The Arabidopsis Information Resource (TAIR) maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*. Data available from TAIR includes the complete genome sequence along with gene structure, gene product information, gene expression, DNA and seed stocks, genome maps, genetic and physical markers, publications, and information about the Arabidopsis research community. Gene product function data is updated every week from the latest published research literature and community data submissions. TAIR also provides extensive linkouts from our data pages to other Arabidopsis resources.

The Arabidopsis Biological Resource Center at The Ohio State University collects, reproduces, preserves and distributes seed and DNA resources of *Arabidopsis thaliana* and related species. Stock information and ordering for the ABRC are fully integrated into TAIR.

 TAIR is located at the Carnegie Institution for Science Department of Plant Biology and funded by the National Science Foundation.

Beginning on October 1, 2013, access to TAIR by researchers at companies or for-profit institutions requires a subscription. Academic / nonprofit access to TAIR will require a subscription after April 1, 2014. ABRC catalog, stock and ordering pages hosted at TAIR will remain accessible without a subscription. Please see our subscription page for further details.



Subscribe to news feed

Follow our Twitter feed

Join our Facebook group

Breaking News

ABRC tool for adding comments about stocks
[January 29, 2014]
We are encouraging users to add comments about stocks that they have ordered. We are collecting information such as validation of insert/sequence, and also problems such as poor germination, lethality, incorrect insert, etc.

Notice from ABRC for Chinese Researchers
[January 15, 2014]
In an effort to improve successful delivery of orders to China, we will include your address in Chinese on all

Improved annotation of the Arabidopsis genome:

Reportoire of gene families in Arabidopsis (11,000 – 15,000) similar to other sequenced multicellular eukaryotes.

However, gene number in Arabidopsis surprisingly high:

Arabidopsis – 30,700 genes. (Version 5 annotation).

Drosophila melanogaster – 13, 676 genes (Release 3).

Some of these extra genes are due to genome duplications, and other plants also seem to have high gene numbers. Rice the second plant genome to be sequenced is estimated to contain around 40,000 genes in 15,000 families.

But there appear to be many genes that are unique to plants and not found in animals:

8,000 (25%) of Arabidopsis genes have homologues in the rice genome, but not in Drosophila, C.elegans or yeast.

What is unique about plants that can be inferred from the Arabidopsis genome?....

Enzymes involved in secondary metabolism

Arabidopsis genome contains many classes of enzymes involved in secondary metabolism that are required for the synthesis of specialized compounds.

An example, is the family of genes encoding the Cytochrome P450 monooxygenase enzymes.

Mammals, C.elegans, Drosophila – 80 – 105 genes.

Arabidopsis – 246 genes.

In plants these enzymes are required for the synthesis of compounds such as growth regulators (gibberellic acid, Brassinosteroid), carotenoids (protect cell from oxidative damage) and phenylpropanoids that are present in plant cell walls.

Transcription factors

Arabidopsis contains around 1500 genes encoding transcription factors (aprox. 5%)

Drosophila contains around 640 genes encoding transcription factors, around 4.5%.

Many important animal transcription factor families are absent in plants, such as nuclear steroid receptors, NHR zinc finger proteins (252 in C. Elegans) and Fork head transcription factors (18 in Drosophila, 15 in C.elegans).

Each eukaryotic lineage has its own set of transcription factor families.

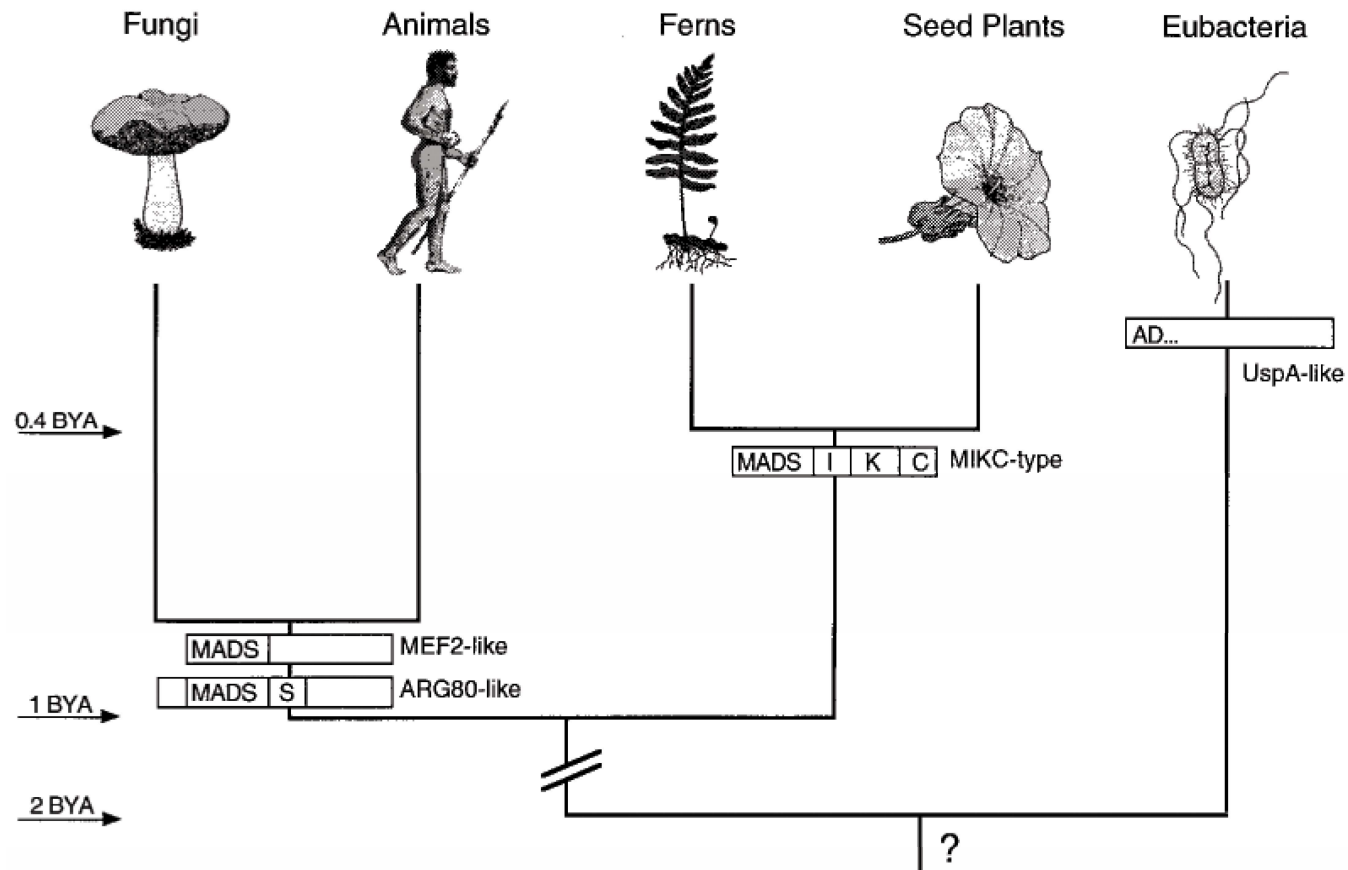
MADS box transcription factors are named after proteins found in yeast, humans and plants

M : MCM1 yeast

A : Agamous

D : Deficiens, Antirrhinum B function gene

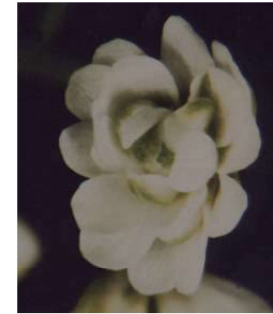
S : serum response factor, humans



MADS box TFs have been amplified in the plant lineage

| | MADS |
|-------------|------|
| Arabidopsis | 107 |
| Drosophila | 2 |
| C. Elegans | 2 |
| Yeast | 4 |

agamous
mutant



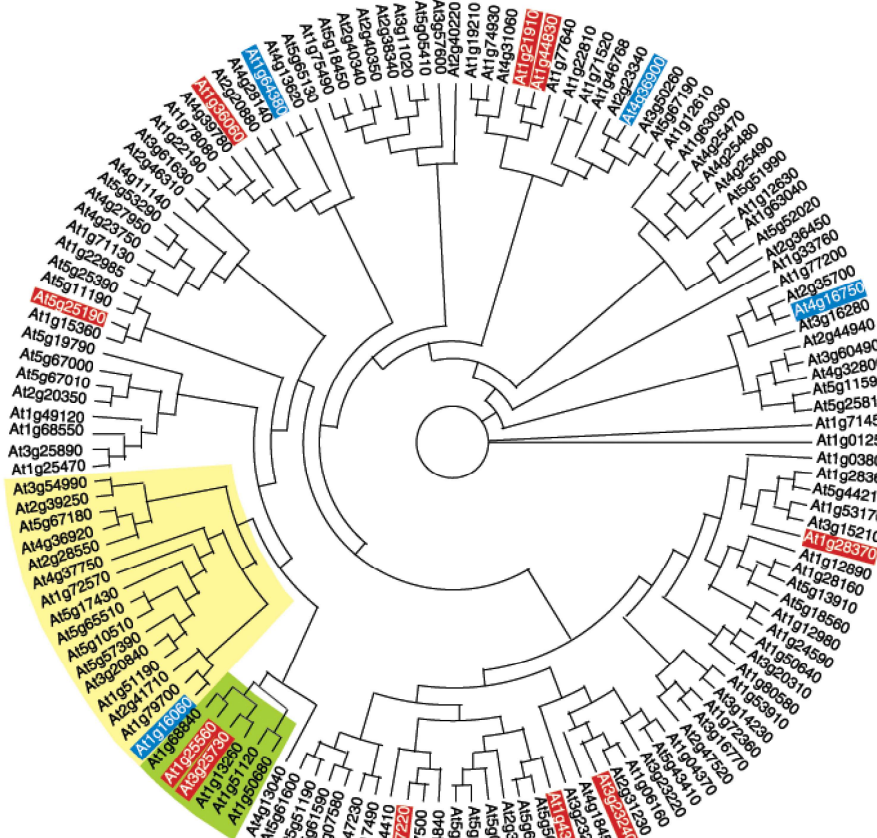
WT



Some have well defined roles in flower development, like AGAMOUS, but 84% are of unknown function

APETALA 2-like transcription factors are unique to plants

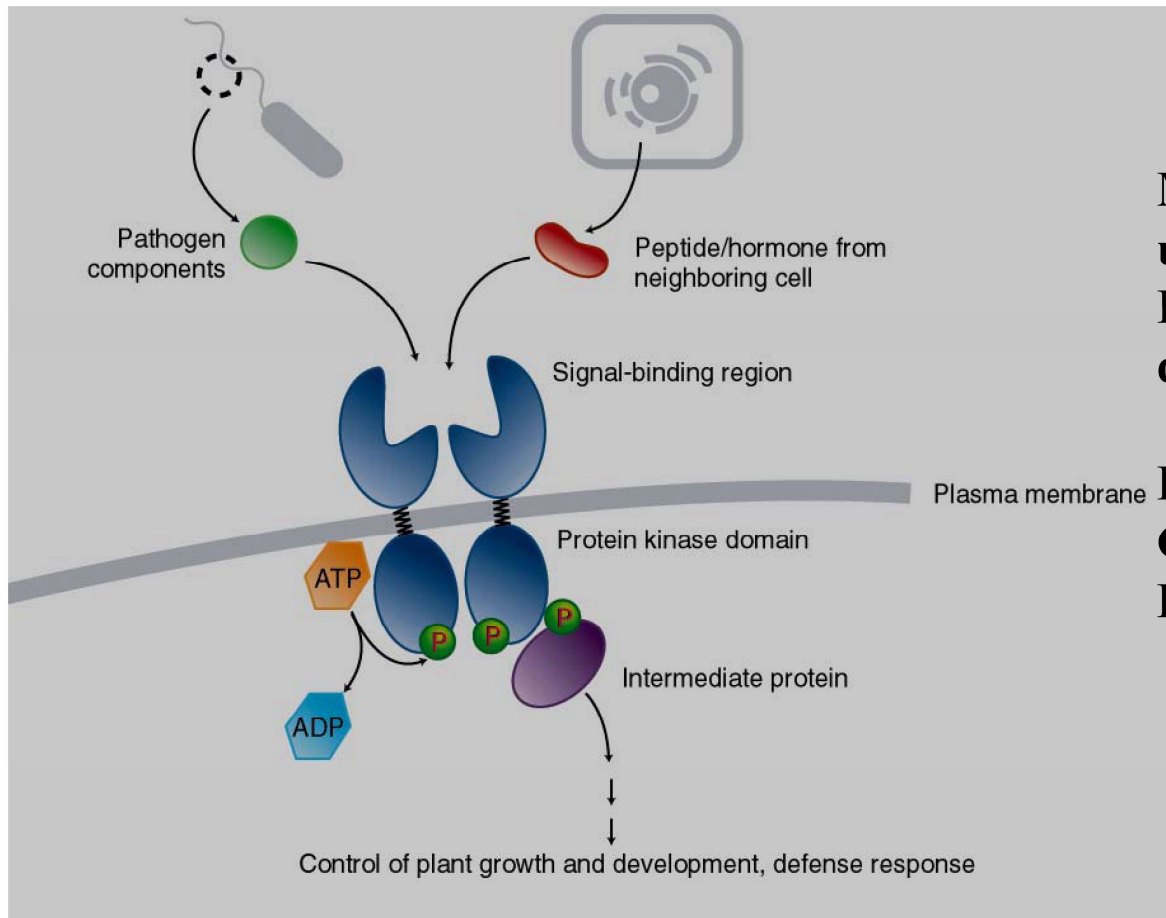
This is one of several families of transcription factor that are only found in plants. May have evolved from bacterial and bacteriophage endonucleases



141 AP2-domain TFs in Arabidopsis

- Some related to APETALA2, Involved in flower development (yellow)
- Some involved in ethylene response and their transcription is changed by ethylene (red/blue)
- Others are involved in stress responses to drought

Some Receptor Like Kinases have important functions, but for most their function is unknown



Many RLKs are of unknown function. However, some have defined roles:

**Brassinosteroid receptor
Clavata 1
Resistance to pathogens**

Other plant-specific processes

Hundreds of genes involved in photosynthesis

- light harvesting**
- chlorophyll biosynthesis**
- carbon dioxide fixation**
- energy generating photosystems**

Transporters

- plants mainly use proton-type ATPases, whereas in animals transport is usually coupled with sodium ions via sodium-type ATPases.**

**No major histocompatibility complex, however 100s of genes
Encoding nucleotide binding site leucine rich repeat proteins
Involved in pathogen resistance.**