1. DNA library preparation: fragmentation of the DNA by nebulization in 300-500 bp fragments (each fragment is bonded with two double strand adaptors (A and B) in 3' and 5' ends); selection of A/B fragments

2. emPCR in which sstDNA is the template

3. Sequencing reaction in which all the beads containing the sstDNA amplified library are mixed up with other smaller beads that contain all the enzyme for pyrosequencing reaction

4. Addition of polymerase and the other enzymes for the reaction

454

Solid

Illumina

For short DN

Most used

1. Break of DNA by
2. Addition of adap
3. Amplification by
4. Original templat
5. Attachment to a complementary
6. PCR with P7 pri fragments on th
7. Annealing of the becomes double
8. Generation of D include around 1
9. Addition of 4 NT

GENOME ANALYSES

Arabidopsis Genome Sequence

- Very easy to handle because it needs little space to grow
- Very adaptable in vary environments
- Different seeds for each pollination
- A lot of synteny
- Genome of 130 Mb, 5 chromosomes

Vectors

- BAC
- YAC

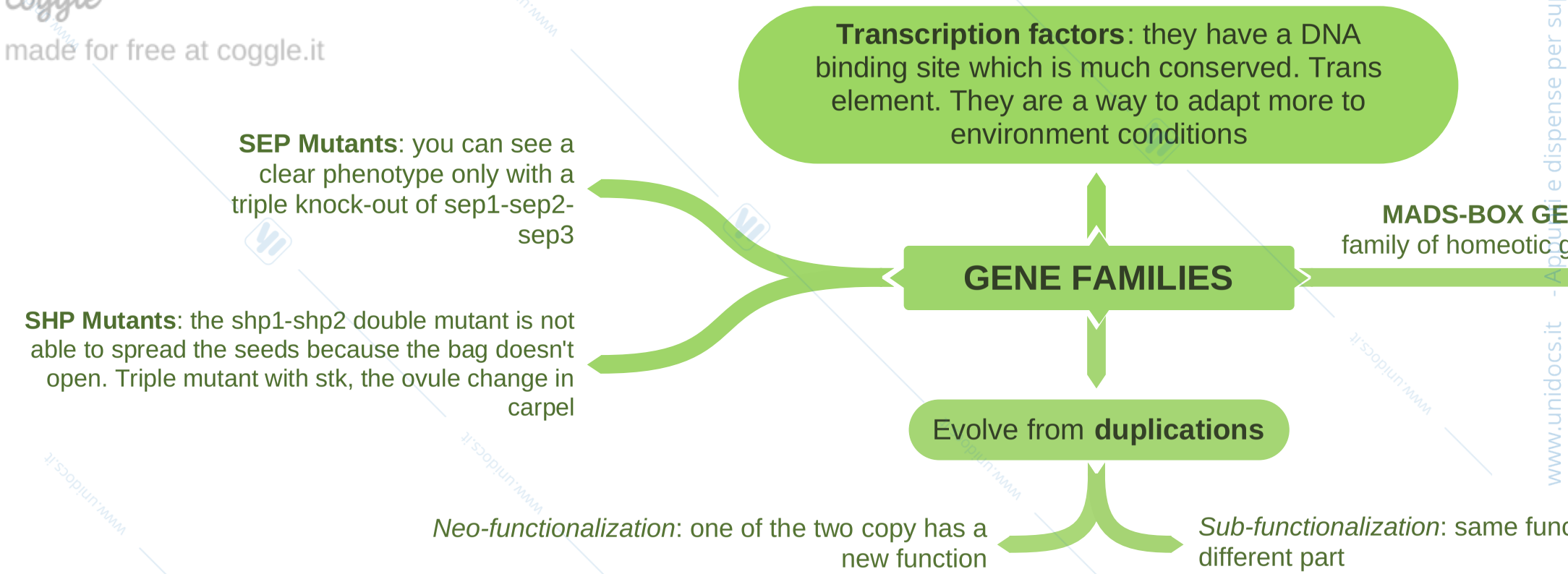
Approaches of sequencing

- **Hierarchical Shotgun**
- **Whole Genome Shotgun**

Exome Sequencing: sequencing o regions that contain genes

- **Limitations**: not mitochondrial ge sequence, triplet

The separation can be done in *liquid* or *solid* state



Mouse

Knock-out

1. Isolate the gene by PCR and then made a construction with neomycin as selection marker
2. Add an extra part that confers sensitivity to ganciclovir
3. Take stem cells from a brown mouse and introduce our DNA construction in it. We can have 3 situations: homologous recombination, random insertion and wild-type.
4. Take cells with homologous recombination and put them in a blastocyst, obtaining a *chimeric mouse*
5. Tet-on/Tet-off system; lac system; cre-lox system

Disadvantages: different in cognitive function, behaviour and gene expression; difficult to monitor development in utero; diploid vs haploid; a large genome that makes it harder to target and work with than organisms with more compact genomes

Inbred strain

Genome just a bit smaller than human, but with quite the same number of genes

Spontaneous mutations, induced mutations in known genes, addition of exogenous/modified genes, cell type-specific or temporal specific mutations, random mutagenesis, toxins, diet, radiation

Plants

Transposons

Used when you don't have efficient transformation

A **disadvantage** is that transposon is an unstable element, so it can move to another gene, and you lose your phenotype.

Ac (Activator) encode for a transposase; *Ds (Dissociation)* is an element that can jump but not by itself, because it needs for Ac

so in the human genome

rmed with
vectors

T-DNA insertion

We can have 3 different possibilities: intron, exon or promoter

enesis: ethyl-methan-sulfonate; we have
om GC to AT bases since it allows ethyl-
guanine pairs with thymine

It adds an ethyl group

dine

cs method

CODDLE

ILLING (Targeting Induced Local Lesions IN Genomes)

1. Amplification by PCR of a certain region of a genome you are interested in.
2. Mutagenesis on the seeds, then let them grow and to do that you can use NMU (N-nitroso-N-methyl-urea
3. The analysis is done to M2 level
4. PCR on the region you are interested in (**tilled fragment**)
5. Treat DNA with Cel1
6. Denature and see what i obtained by running and separating fragments by electrophoresis on agarose gel.
7. If you don't have a mismatch, you don't have cuts. If you have a mismatch there will be two bands on the gel

NA;
and
sion

Examples in DEVELOPMENT

miR156 and **miR172**

- mi156 promotes degradation; mi172 controls the flowering
- Target in the 3' UTR of SPL
- mi156 are high in the juvenile phase

lin-14 and **lin-4**

- In wt lin-14 is regulated by lin-4
- Post-transcriptional kind of regulation
- Lin-14 has different binding sites for lin-4

- With miRNA the analysis of functional genomics are complicated
- Encoded by specific genes in the genome of eukaryotes (*miR genes*). In plants they are involved in developmental and physiological processes.
- It comes from the dicer-mediated cleavage of pre-miRNA
- Specific DICER enzymes for miRNA

Examples in PHYSIOLOGY

miR399

- It is able to move from the shoot to the root
- E2 is an ubiquitination factor highly complementary to miR399
- Target a pathway of E2, involved on phosphate accumulation factors
- Its expression is induced by phosphate starvation
- Pho2 is the E2 loss-of-function mutant
- miR399 also partial anneals with IPS1

Cloning Technology: uses the lambda
n system to facilitate transfer of heterologous
nces (flanked by modified att sites) between
vectors.

- *attP* (phage I) and *attB* (E.coli) recombination sites

ing the error-prone
ent the formation of

e derived from ZFNs by
of the FokI domain in one
atalytically inactive but
en they form an
t the target site.

nt (*Asp10Ala*) form of
egligible NHEJ-driven

Non-integrating viral vectors (*IDLVs*, *adenoviruses* and *AAVs*) can deliver genes encoding programmable nucleases both in vitro and in vivo.

Delivered in various forms: *plasmid DNA*, *in vitro transcribed mRNA*, *viral vectors* or *purified proteins*.

mable nucleases

y of ZFNs and TALENs
dulated by changing the
f zinc fingers and TALE
modules.

ZFNs and TALENs that contain more modules are generally believed to recognize longer DNA sequence and thus be more specific than those containing fewer modules.

BUT nucleases with too many modules target many additional sites owing to partial interactions that involve only a portion of the modules

All 3 nucleases can cause off-target mutations.

ZFNs are cytotoxic, instead TALENs and RGENs not

: use different type of enzymes that can produce
instead of 9 bp -> longer are the tags sequence
he result

short sequence tags and the number of times it
statistical methods can be applied to tag and
m different samples, in order to determine which
re highly expressed -> example of normal
e corresponding tumour

ver genes not yet characterized

ides sequence (no more than 9bp) are able to
2 thousand transcripts. this technique is based

NG A GENE REGULATED BY A NSCRIPTIONAL FACTOR

ChIP

It's an immunoprecipitation, based on finding the binding site

1. We take the tissue and incubate it, proteins are crosslinked to the DNA.
2. We isolate the chromatin and shear it in pieces, the proteins are crosslinked thus they remain linked.
3. We then use antibodies for our protein of interest.
4. We immunoprecipitate and pull it down.
5. We then use a protease and purify the DNA
6. We could then take primers and do a PCR. By RT-PCR we can quantify the amount of the wanted DNA.

Criteria to choose the genes: i have seen induction or decreasing, some variation on expression; it must have putative binding site; expression profile.

ChIP-SEQ

EMSA

Footprinting in vitro

How to find binding regions for transcriptional factors?

Method used to study protein-DNA or protein-RNA interactions

(LMD) is a
recision
ion

EMSA (Electrophoretic Mobility Shift Assay)

FACS (Fluorescence Activated Cell Sorting)

Particular form of *flow cytometry* that enables a mixture of different cells to be sorted one by one into one or more containers

The cells are sorted according their specific light scattering and fluorescent characteristics

Cell sorting

1. Individual cells are "interrogated" by the laser as in a normal flow cytometer
2. The machine is set up so that each individual cell then enters a single droplet as it leaves the nozzle tip. This drop is given an electronic charge, depending on the fluorescence of the cell inside the drop.
3. Deflection plates attract or repel the cells accordingly into collection tubes.
4. Sorted cell populations are then analysed to ensure successful cell sorting
5. Sorted cells can then be cultured

estigating the
cificity of DNA
ns in vitro

ctivity and fluorescence