In Person quiz

Manipulating Nucleotides I

What is a mutation?

A mutation is an alteration in the genome of an organism

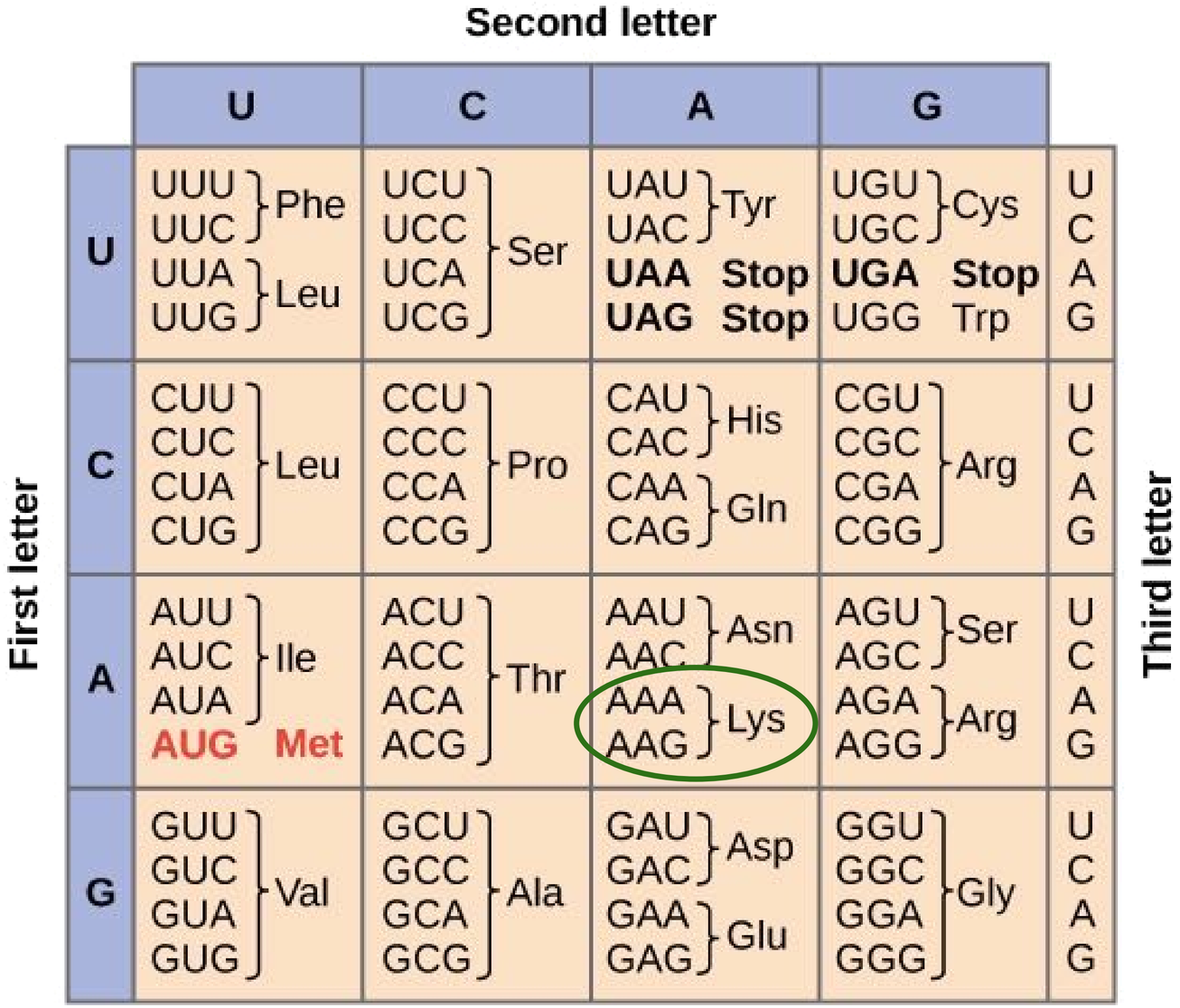
- Chromosomal scale – polyploidy, duplications, inversions, deletions, translocations.

- Nucleotide scale (one or more nucleotides) – insertions, deletions, substitutions.

Effects on gene products

- **Point mutation** – single nucleotide altered
- Effect can be neutral e.g. synonymous substitution

The Genetic code



Synonymous substitution

‘Silent substituion’

GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Codon Usage Bias

Some codons are more favoured than others

- Some specific codons are used more often than other **synonymous** codons during translation of genes
- Can vary **within** a species
- Can vary **between** species.
- Use of ‘non-optimal’ codons can **reduce** gene expression
- Selection of the optimal codon usage for the host species during **gene synthesis** is a good idea when moving genes between species

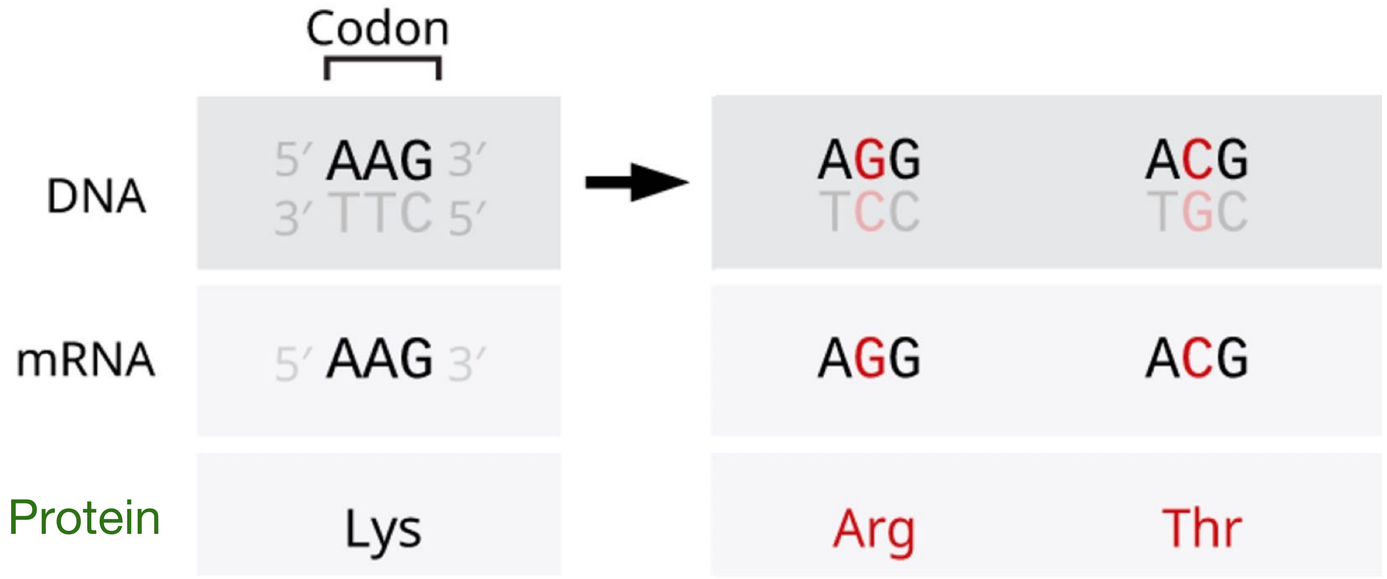
Amino acid	Codon	*A. thaliana*	*S. cerevisiae*	*S. pombe*	*H. Sapiens*
Cys	UGC	+	-	+	+
	UGU	-	+	-	-
Glu	GAA	-	+	-	-
	GAG	+	-	+	+
Phe	UUC	+	+	+	+
	UUU	-	-	-	-
His	CAC	+	+	+	+
	CAU	-	-	-	-
Lys	AAA	-	-	-	-
	AAG	+	+	+	+
Asn	AAC	+	+	+	+
	AAU	-	-	-	-
Gln	CAA	-	-	+	-
	CAG	+	+	-	+
Tyr	UAC	+	+	+	+
	UAU	-	-	-	-

Effects on gene products

- Point mutation – single nucleotide altered
- Effect can be neutral e.g. synonymous substitution
- Change codon to encode a **different amino acid**.

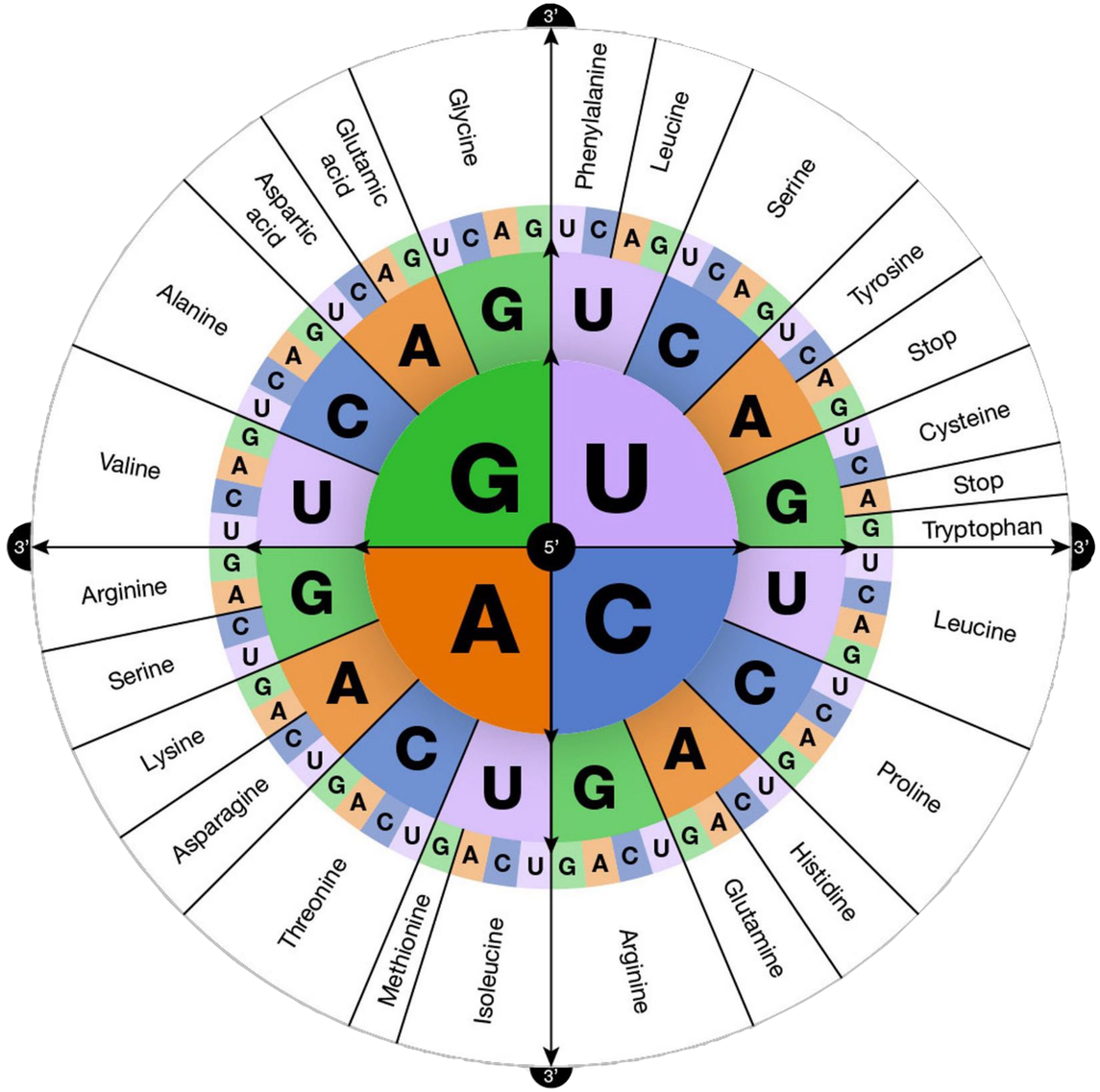
Nonsynonymous mutation

Change amino acid



GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

The Genetic code



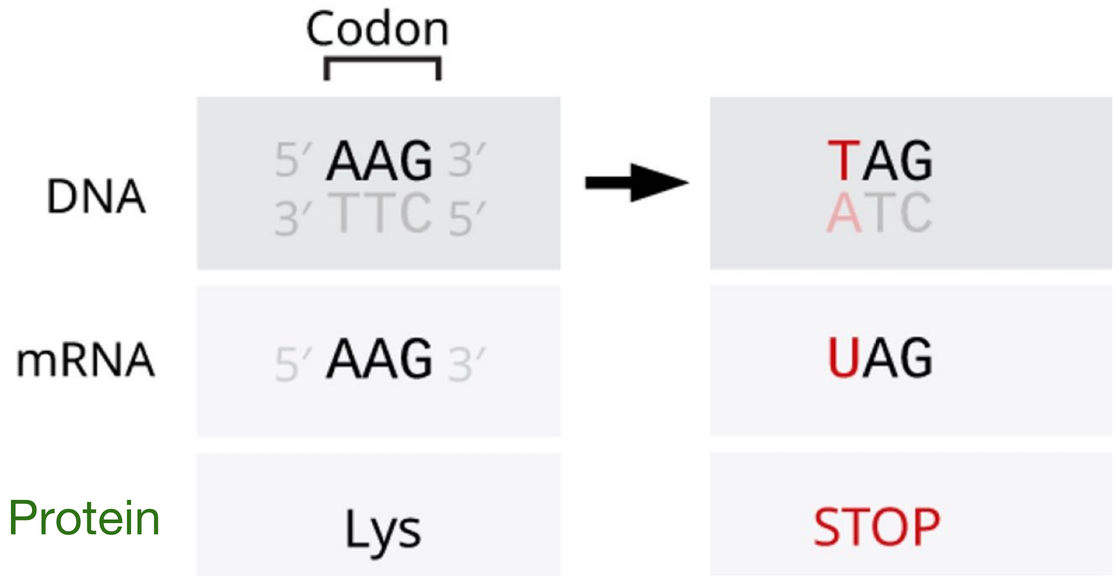
GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Effects on gene products

- Point mutation – single nucleotide altered
- Effect can be neutral e.g. synonymous substitution
- Change codon to another amino acid.
- Introduce a **new stop codon** i.e. ‘premature stop’ or ‘nonsense’ mutation

Synonymous substitution

‘nonsense mutation’

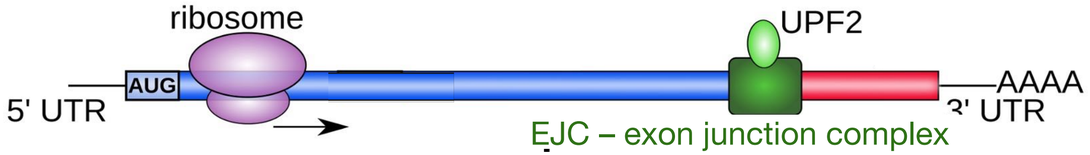


GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Nonsense-mediated decay

NMD

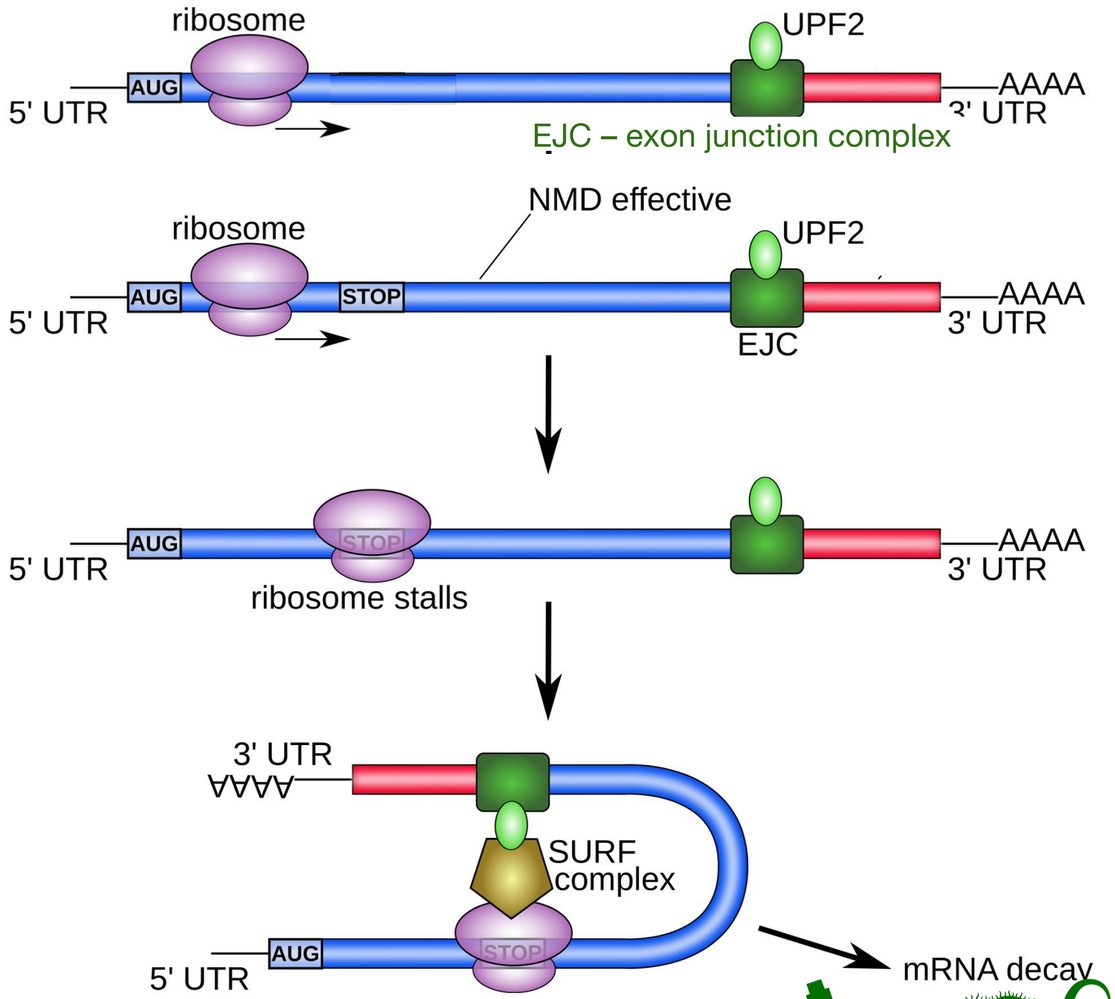
- NMD is a **surveillance mechanism** for errors in gene expression
- **Degrades mRNA transcripts** that contain premature stop codons
- After splicing, exonic-junction complexes (EJCs) remain 20–24 nucleotides upstream of every exon junction and can bind UPF2 proteins.
- If a ribosome encounters a premature stop codon, **it stalls**, allowing a complex with a downstream EJC/UPF2 promoting **mRNA decay** through the SURF complex.



Nonsense-mediated decay

NMD

- NMD is a **surveillance mechanism** for errors in gene expression
- **Degrades mRNA transcripts** that contain premature stop codons
- After splicing, exonic-junction complexes (EJCs) remain 20–24 nucleotides upstream of every exon junction and can bind UPF2 proteins.
- If a ribosome encounters a premature stop codon, **it stalls**, allowing a complex with a downstream EJC/UPF2 promoting **mRNA decay** through the SURF complex.



Effects on gene products

- Point mutation – single nucleotide altered
- Effect can be neutral e.g. synonymous substitution
- Change codon to a different amino acid.
- Introduce a new stop codon
- **Frameshift** mutation

Frameshift mutation

- Insertion or deletion of nucleotides that are **not in multiples of 3**
- Changes the reading '**frame**' and thus the protein sequence that will be produced.
- Often results in a premature stop codon.

TAT TGG CTA CTA CAT
Tyr Trp Leu Val His

TAT T**C**G GCT AGT ACA T..
Tyr **Ser Ala Ser Thr**

In Person quiz

‘Natural’ Sources of mutations

Pass the sunblock

- **Spontaneous** mutations – chemical changes to nucleotides
- **Replication** errors
- **DNA repair** errors
- Environmental **mutagens** – **radiation** e.g. Ultraviolet light, ionising radiation e.g. gamma radiation, **chemicals** that cand cause oxidative or other damage.
- Median **somatic** mutation rate per base pair is 2.8×10^{-7} per generation for humans. Order of magnitude **lower in germline cells.**

Effects of mutations

Not usually superpowers

- **Loss-of-function (LOF)** mutations - gene product function reduced (**partial LOF, hypomorph**) or no function (**null allele**, **amorph**). If no protein produced at all, called ‘protein null’.

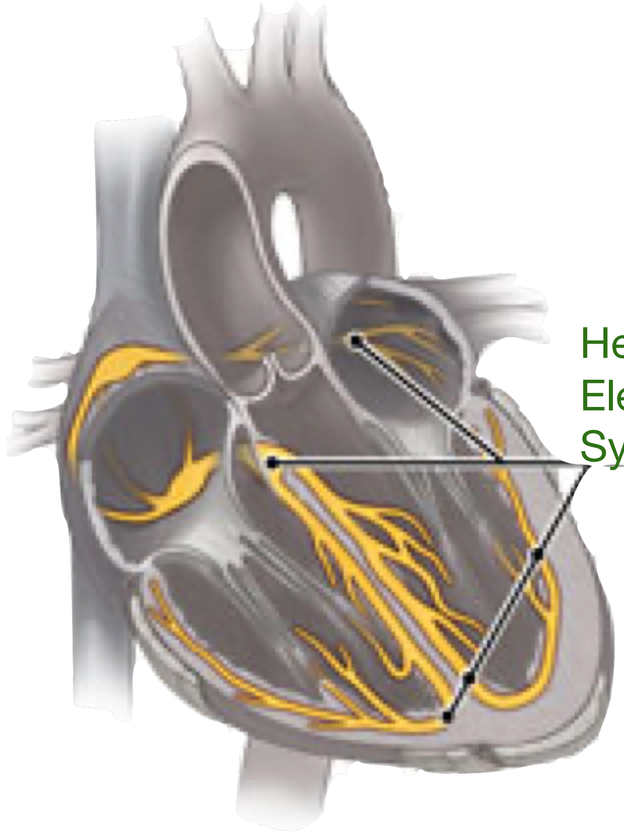
- Often **recessive** (i.e. both alleles must to be mutant to observe a phenotype) but can be **haploinsufficient** (loss of one allele can cause a phenotype)

- Human disease examples – **Recessive** e.g. Cystic fibrosis. **Haploinsufficient** e.g. DiGeorge syndrome (22q11 deletion)

GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

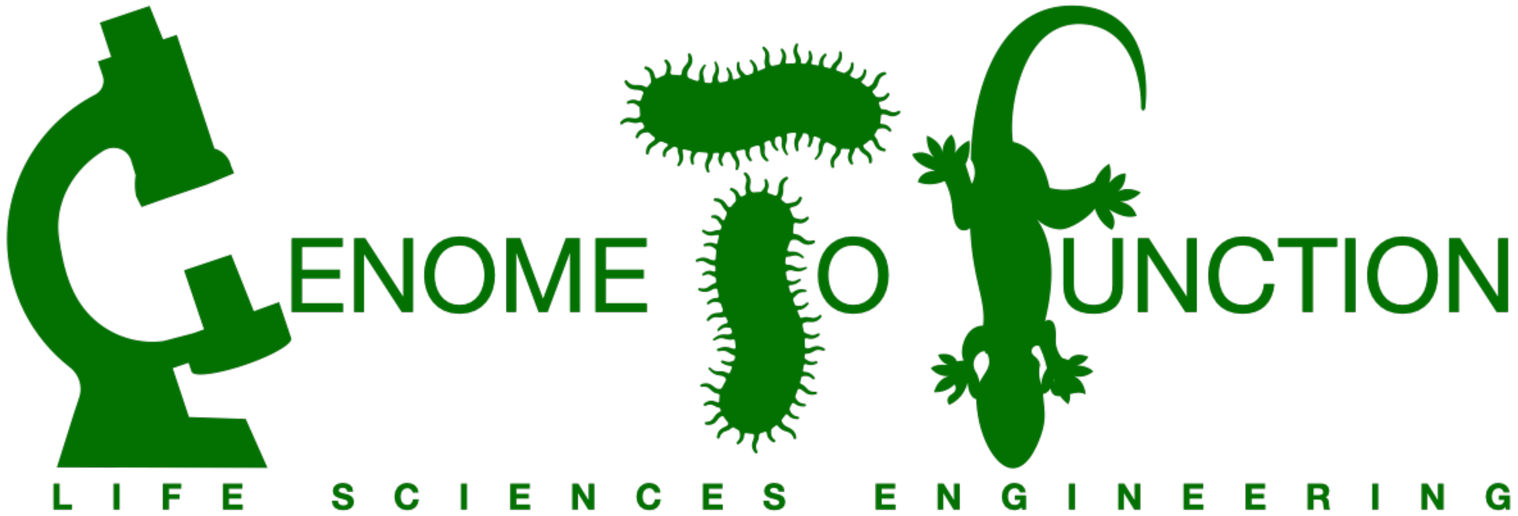
Effects of mutations

- **Gain-of-function (GOF)** mutations (**hypermorphs**).
- Gene product activity is **increased.** Examples include mutations that increase protein expression or increase protein activity.
- GOF mutations often produce **dominant phenotypes** (i.e. one mutated allele is sufficient to produce a phenotype).
- Related are **Neomorphs –** mutations causing **novel gene product functions** (e.g. expression outside of normal tissues, interaction with novel proteins).
- Human disease example - **Brugada syndrome** (GOF mutants in KCNE3) can cause sudden cardiac death in young people



Effects of mutations

- **Dominant Negative (DN)** mutations **(antimorphs)** produce altered gene products that act antagonistically to inhibit to the normal gene product.
- Mutations usually have an altered molecular function (commonly reduced activity)
- DN mutations often produce **dominant phenotypes** (i.e. one mutated allele is sufficient to produce a phenotype).
- Human disease example - **Marfan syndrome** (Dominant negative mutants in FBN1), affects connective tissue.



Genetic Model Organisms

Whistle-stop tour of genetic model organisms

Saccharomyces cerevisiae (Bakers yeast)

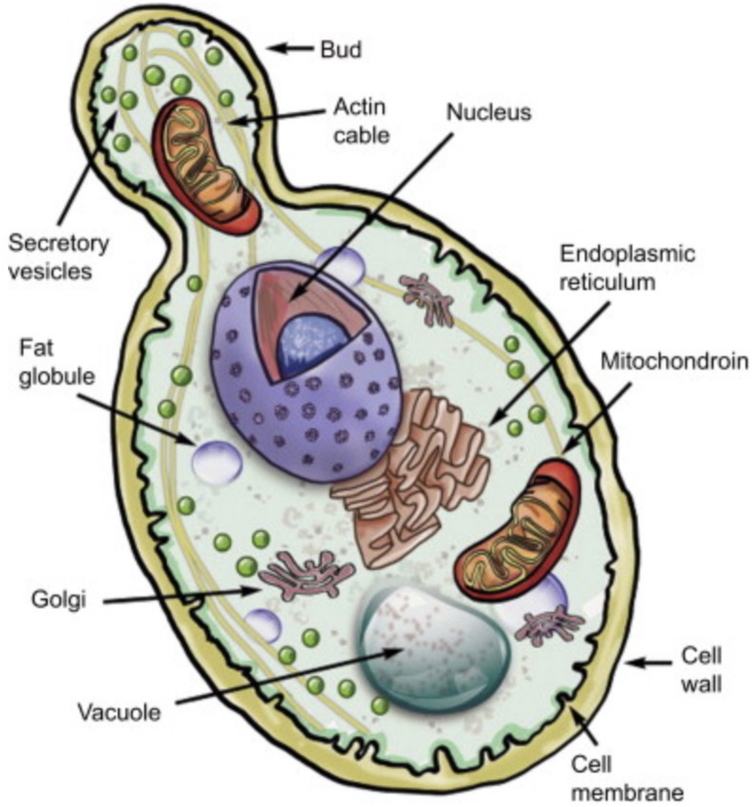
Single cell model organism

Easy to culture, doubling time at 30°C of ~90 minutes

Genome size 12 Mb, 16 chromosomes.

6604 protein coding genes, ~2000 non-coding genes,

Mutants in every gene available



GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Whistle-stop tour of genetic model organisms

Caenorhabditis elegans (nematodes)

1000 somatic cells in adults plus 1000-2,000 germ cells

Generation time 3 days at 25°C

Genome size 100 Mb, 12 chromosomes

20,470 protein coding genes, ~1300 non-coding genes, mutants in most genes available

Whistle-stop tour of genetic model organisms

Drosophila melanogaster (fruit flies)

>million cells

Generation time 10 days at 25°C

Genome size 180 Mb, 4 chromosomes

13,968 protein coding genes, 4,044 non-coding genes, mutants in most genes available

https://drosophila.epfl.ch

Whistle-stop tour of genetic model organisms

Danio rerio (Zebrafish)

millions of cells

Generation time 3 months

Genome size 1.4 Gb, 25 chromosomes

25,545 protein coding genes, 6,599 non-coding genes, mutants in some genes available

Whistle-stop tour of genetic model organisms

Mus Musculus (mice)

Millions of cells

Generation time 10 weeks

Genome size 2.6 Gb, 20 chromosomes

22,213 protein coding genes, 17,398 non-coding genes, mutants in some genes available

Whistle-stop tour of genetic model organisms

Arabidopsis thaliana (Thale cress)

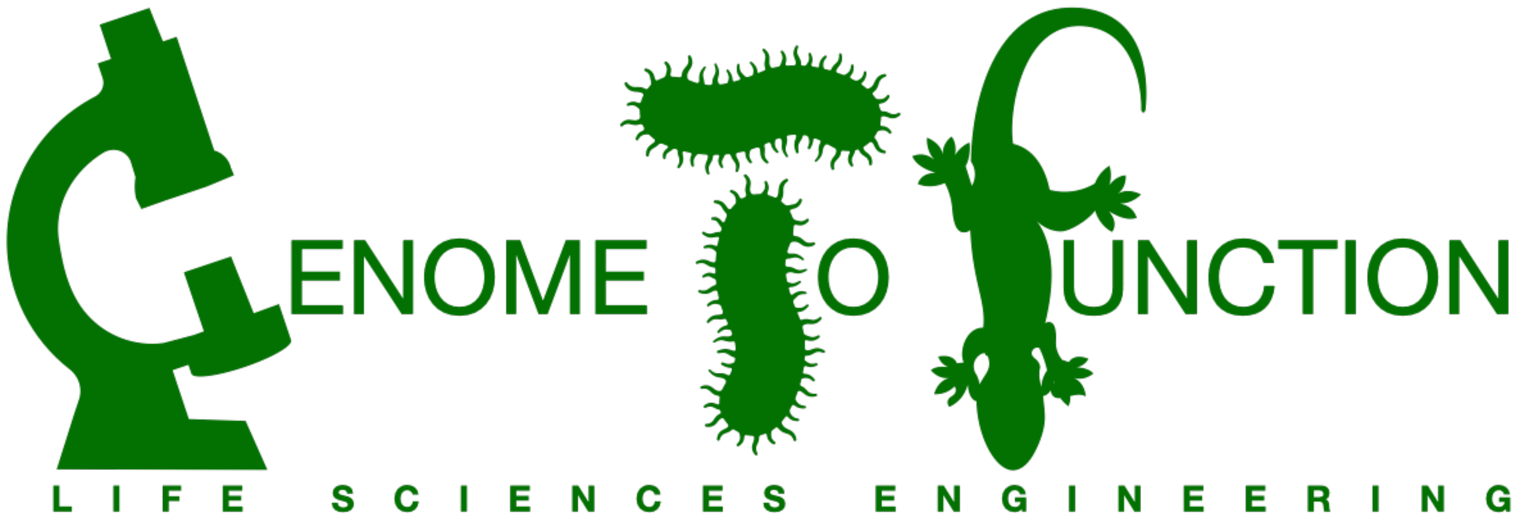
millions of cells

Generation time 6 weeks

Genome size 1.35 Gb, 5 chromosomes

27,655 protein coding genes, 5,178 non-coding genes, mutants in some genes available

In Person quiz



Transgenesis

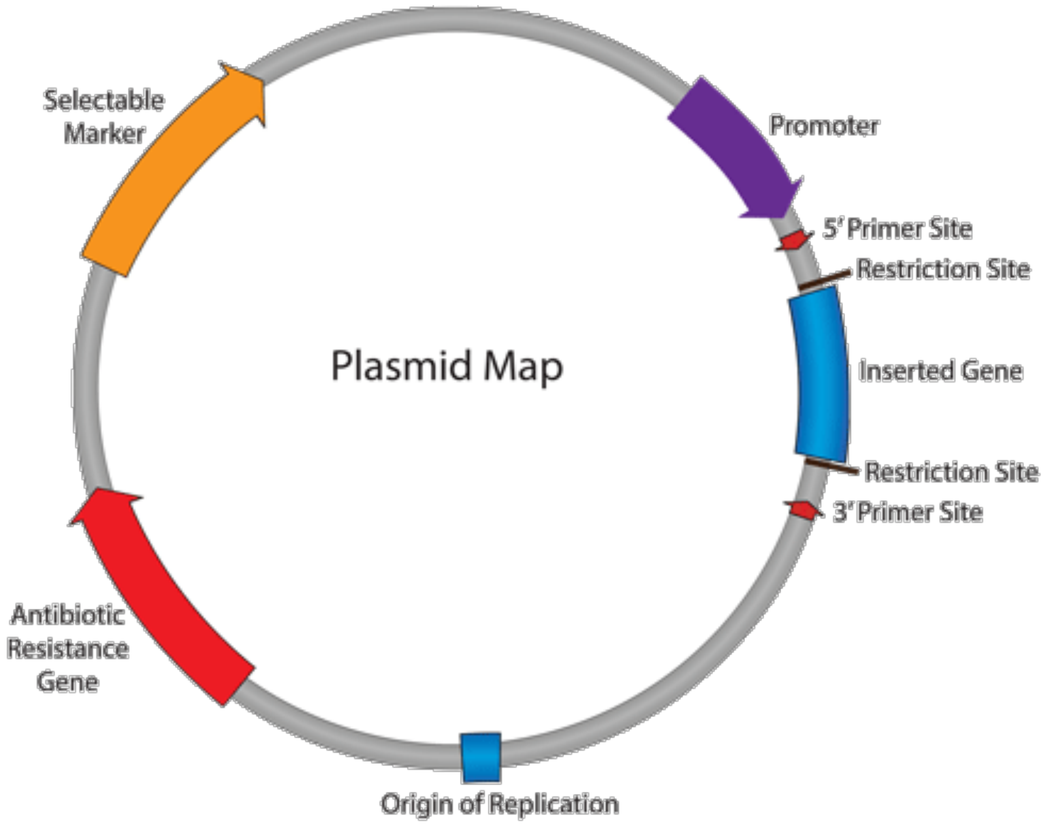
What is Trangenesis?

- The process of introduction of a transgene into an organism
- A transgene is any exogenous genetic sequence either derived from the same species (e.g. an extra copy of a gene), a different species or an artificial sequence.
- Most transgenic animals are generated for research but transgenic animals (GMO – genetically modified organisms) are used in agriculture.

Transgene Donor constructs

Plasmids

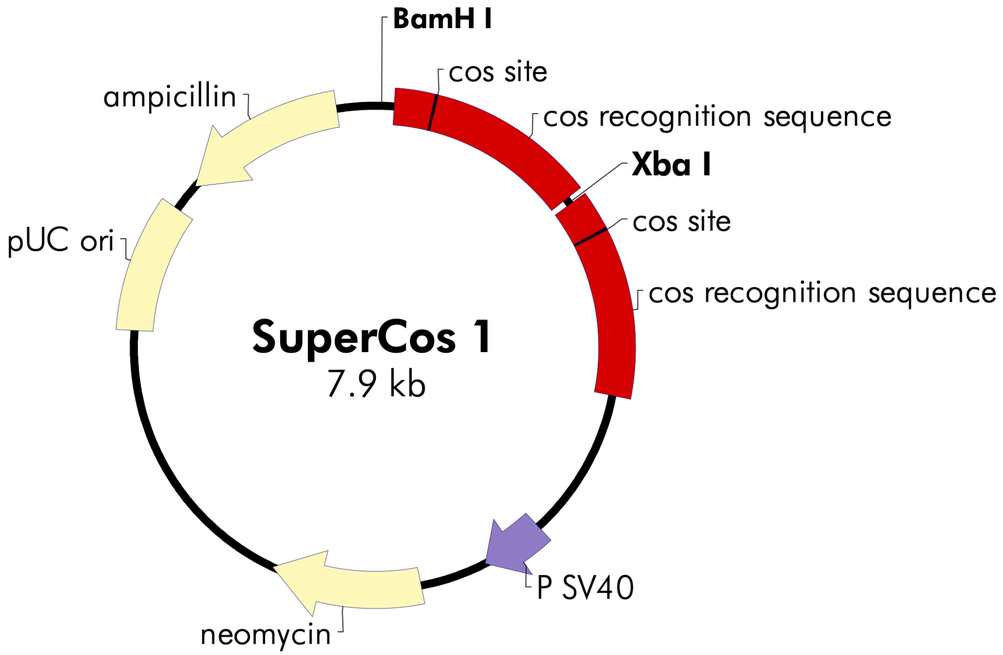
- Easy to work with - Convenient size (generally 1,000-20,000 bp), grow in bacteria (often e.coli).
- Self-replicating - endless number of copies.
- Stable – store in freezer or even dried.
- Useful for lots of things, not species limited, many types of sequences proteins, RNA’s etc. etc.
- Larger DNA sequences are difficult.



Transgene Donor constructs

Cosmids

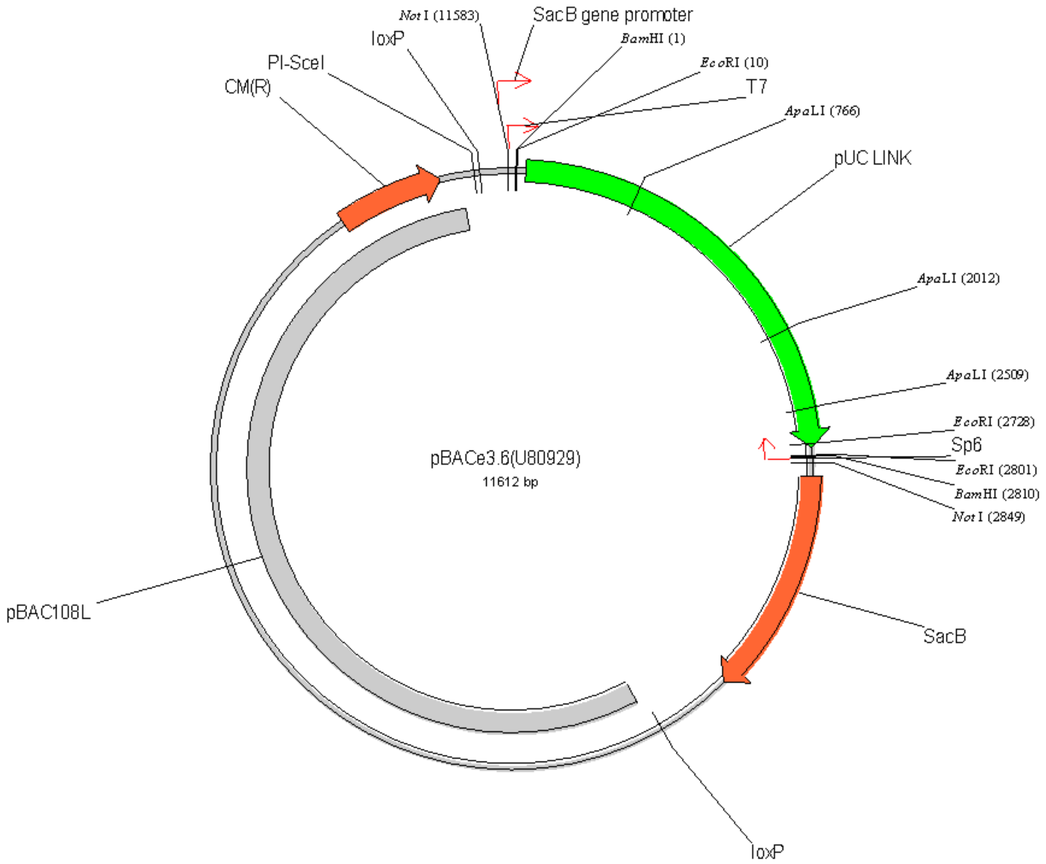
- **Cos** sequence containing plas**mid**
- Can replicate like a plasmid but unlike plasmids can be packaged in phages.
- Can accommodate larger DNA fragments ranging in size from 30 to 45 kb.
- DNA fragments have to be introduced by restriction digestion.
- Most useful for larger genomic DNA fragments, less so for engineered constructs.



Transgene Donor constructs

BACs & YACs

- **B**acterial **A**rtificial **C**hromosome
- Artificial circular chromosomes based on E.coli F-plasmids
- Can accommodate larger DNA fragments ranging in size up to 300 kb.
- YACs are similar but grow in Yeast not bacteria. Can accommodate up to 2000 Kb
- DNA fragments have to be introduced by restriction digestion.
- Useful for large genomic DNA fragments, often used as donors for mouse genetic engineering
- Not useful for small constructs

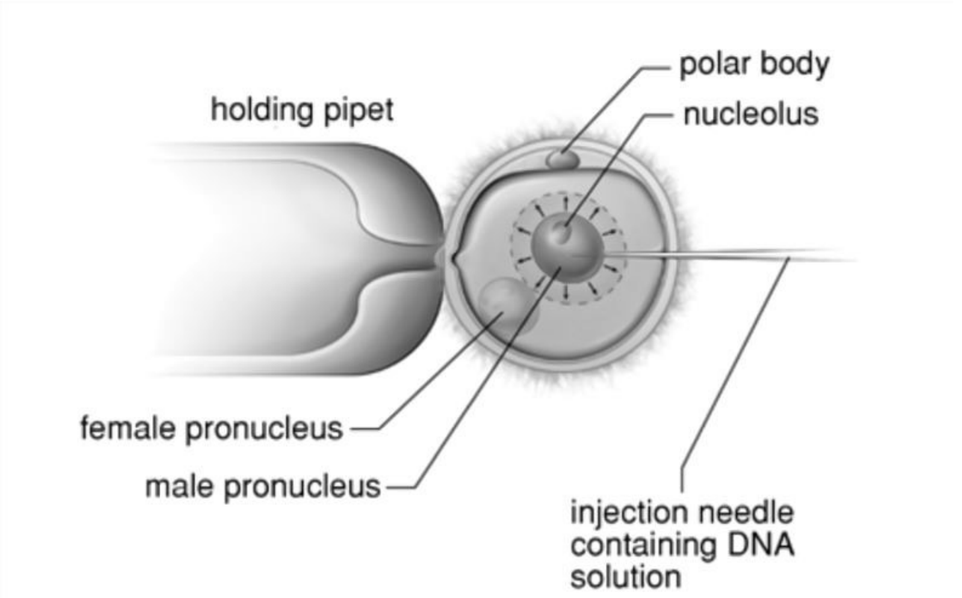


GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Transgenic Techniques - mammals

DNA microinjection

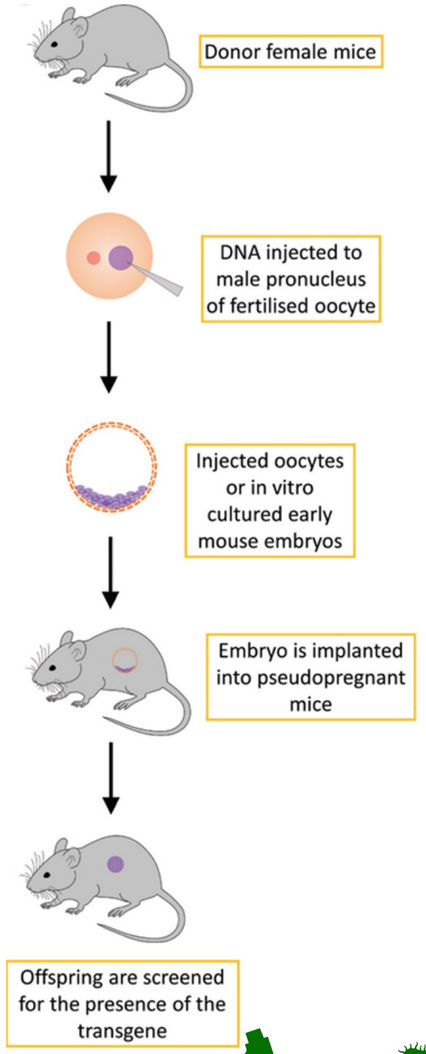
- A plasmid, cosmid or BAC is microinjected
- DNA is introduced into the pronucleus of a developing zygote
- Eggs that survive the injections are transferred into the oviduct of a foster mother.



Transgenic Techniques - mammals

DNA microinjection

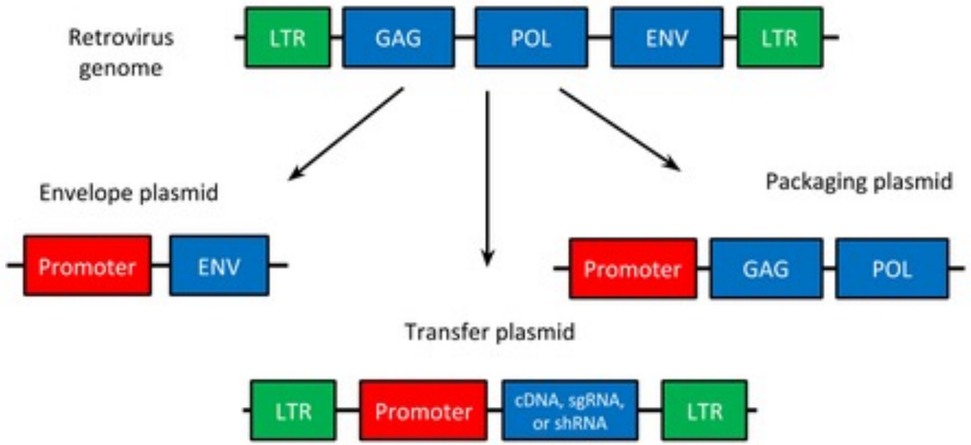
- Widely used
- The amount of DNA delivered per cell is not limited
- Broadly useful across mammalian hosts
- Low success rate
- Mosaic founders
- **Random integration**



Transgenic Techniques - mammals

DNA microinjection with retrovirus

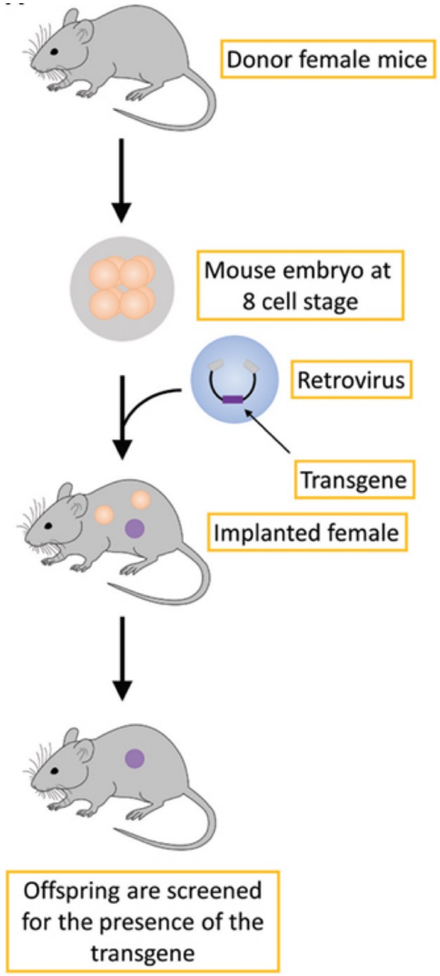
- Commonly use γ-retroviruses (gamma-retroviruses)
- Mostly derived from MoMLV (Moloney Murine Leukemia Virus) or MSCV (Murine Stem Cell Virus) sequences
- Sequences between and including the LTRs is integrated into the host genome upon viral transduction



Transgenic Techniques - mammals

DNA microinjection with retrovirus

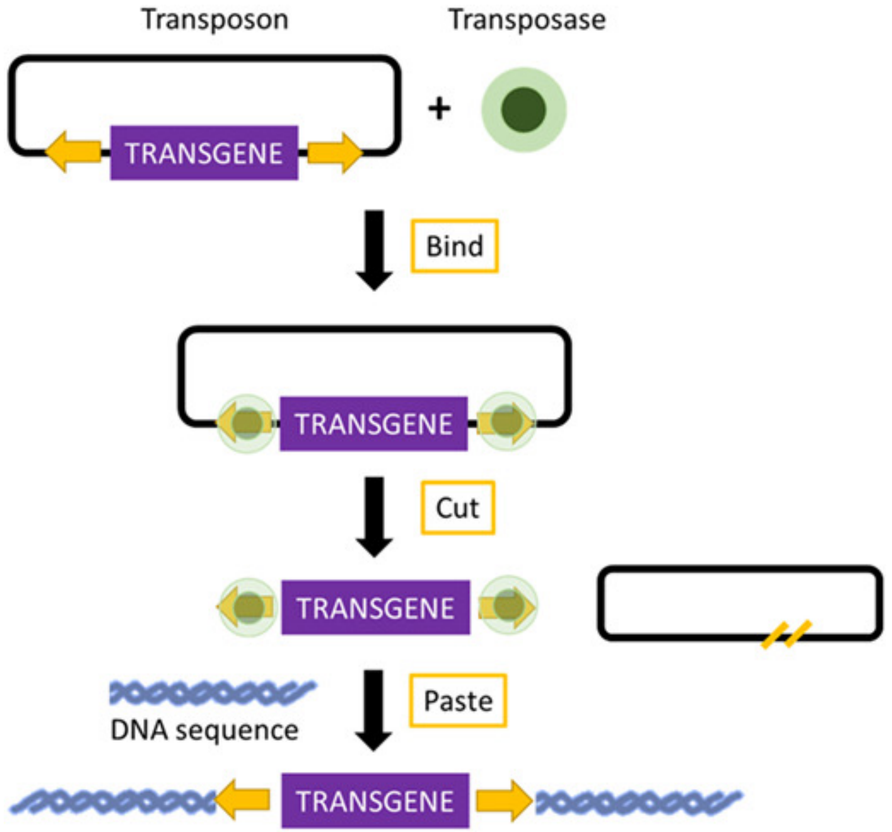
- Virus integration is very efficient
- The amount of DNA delivered per cell is **limited** by virus size max 8.5Kb, but ~3Kb is better
- **Random** integration
- Transgene can be silenced by DNA methylation



Transgenic Techniques - transposons

Transposons

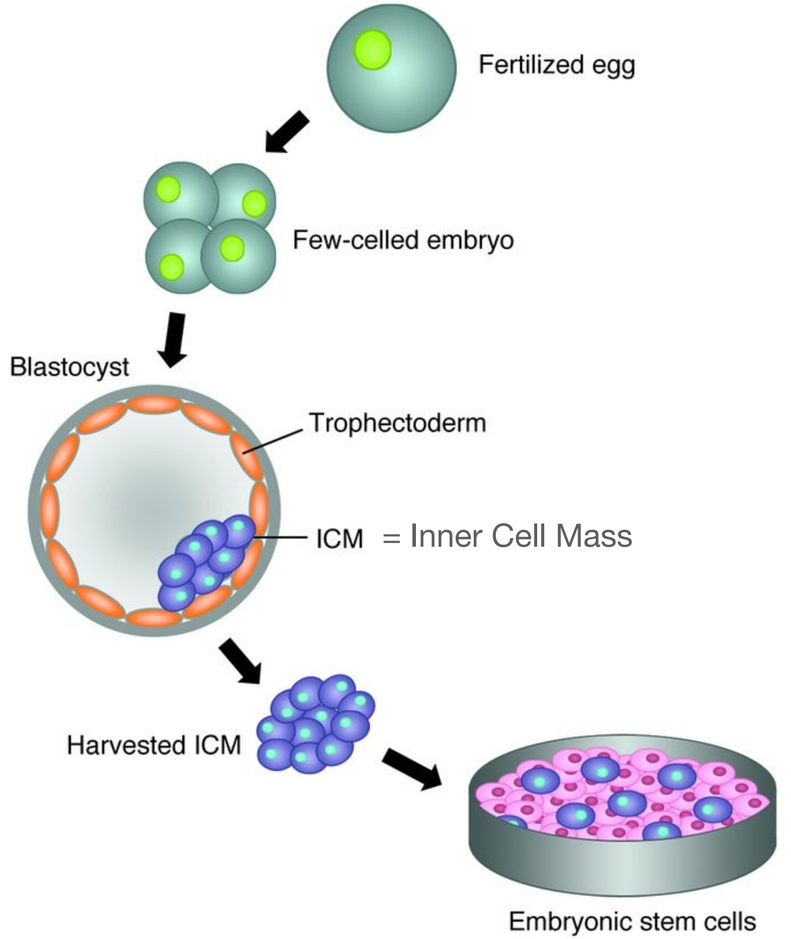
- Transposons, ‘jumping genes’, are genetic elements that can translocate within the genome – Selfish genetic elements
- Transposition process requires sequences at the ends of the transposon and a specific transposase protein
- Can be engineered to replace internal sequences with transgenes
- Have specific sequence preferences e.g. PiggyBaC - TTAA
- Random insertion



Transgenic Techniques - mammals

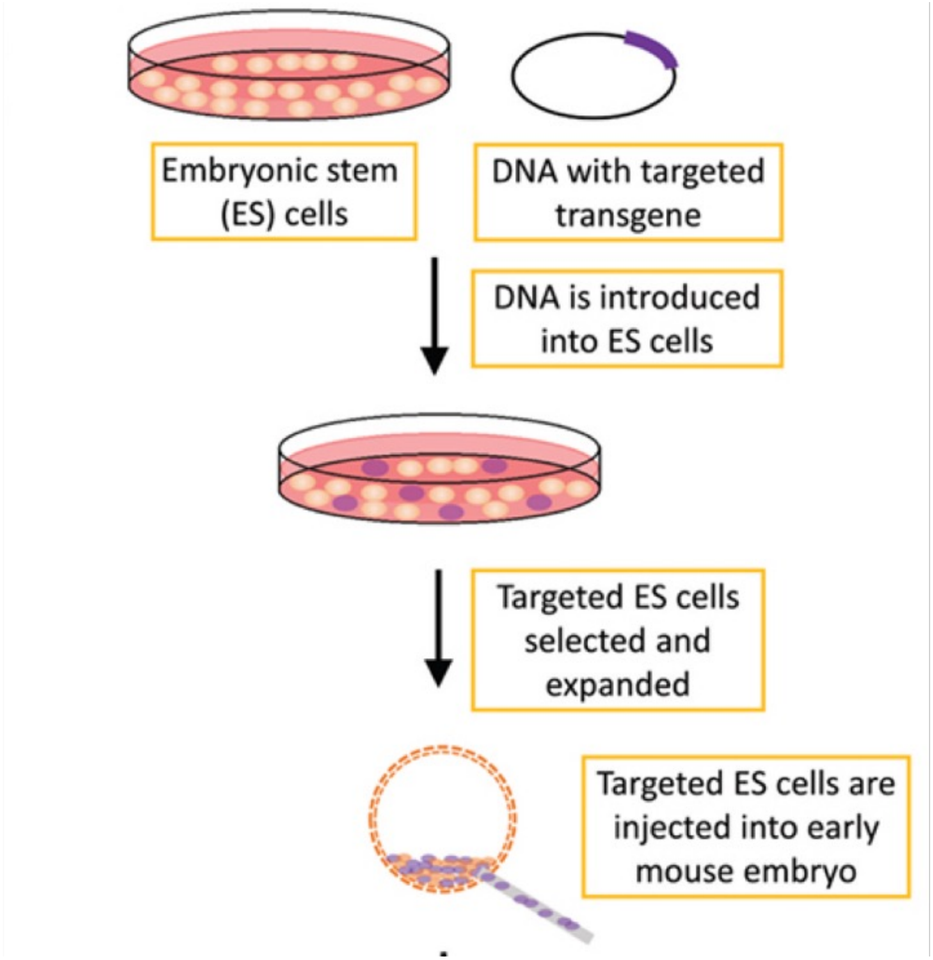
Embryonic Stem Cells (ES cells)

- ES cells are derived from cells of the early embryos (late blastocyst stage)
- Have the capacity to self-renew indefinitely.
- Pluripotent i.e. can differentiate into all cell types in the body including germ cells



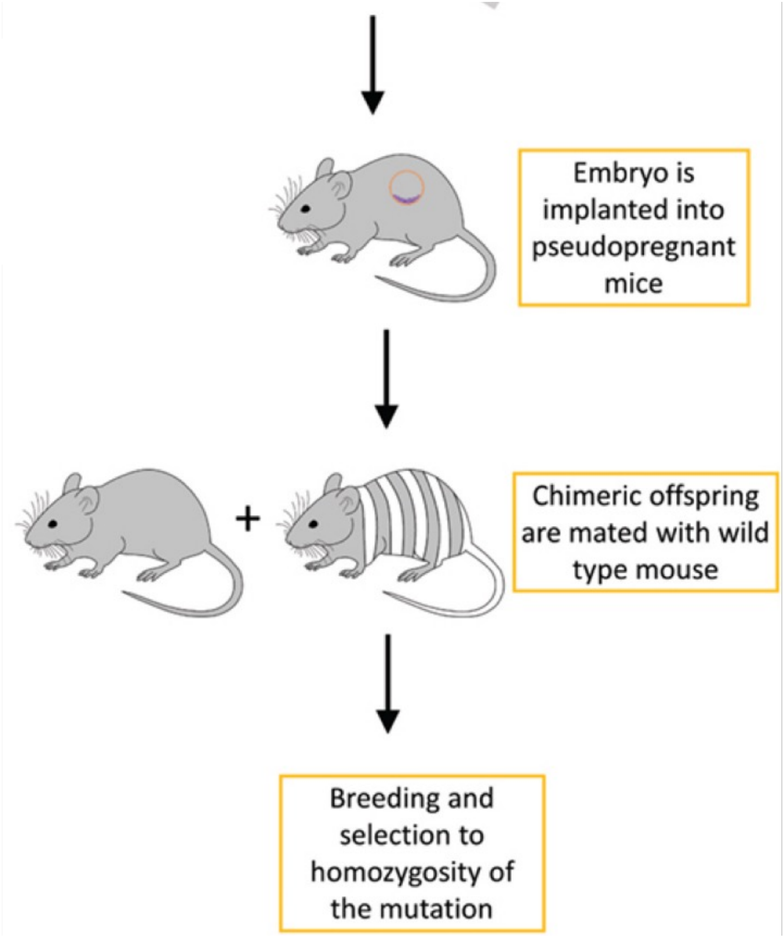
Transgenic Techniques - mammals

Introducing transgenes using ES cells



Transgenic Techniques - mammals

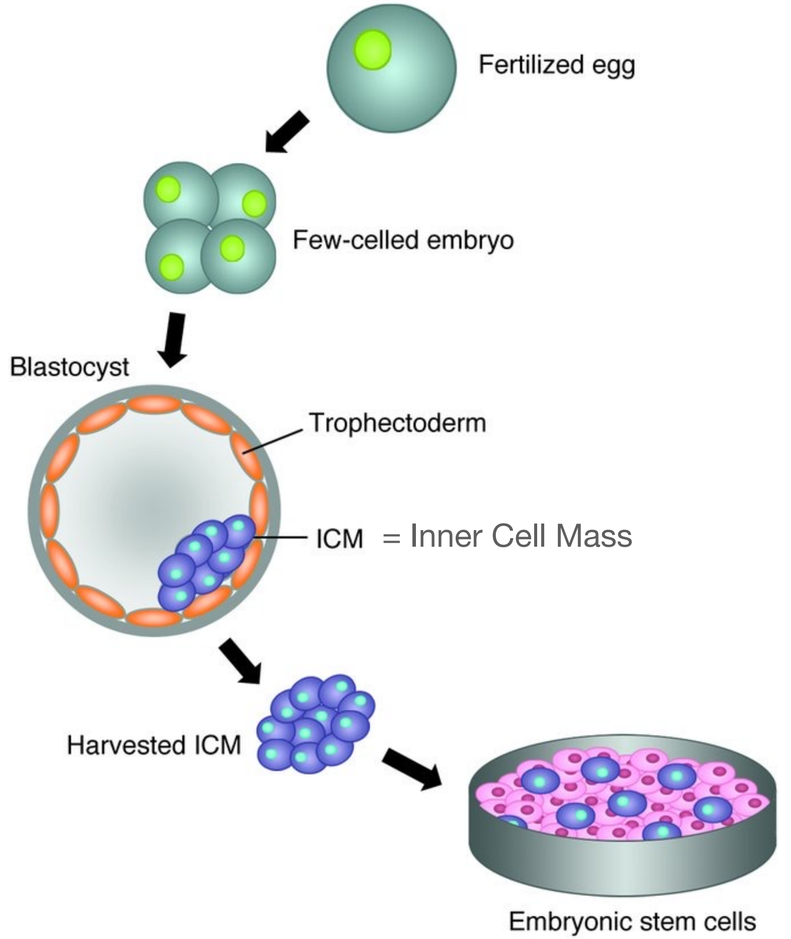
Introducing transgenes using ES cells



Transgenic Techniques - mammals

Introducing transgenes using ES cells

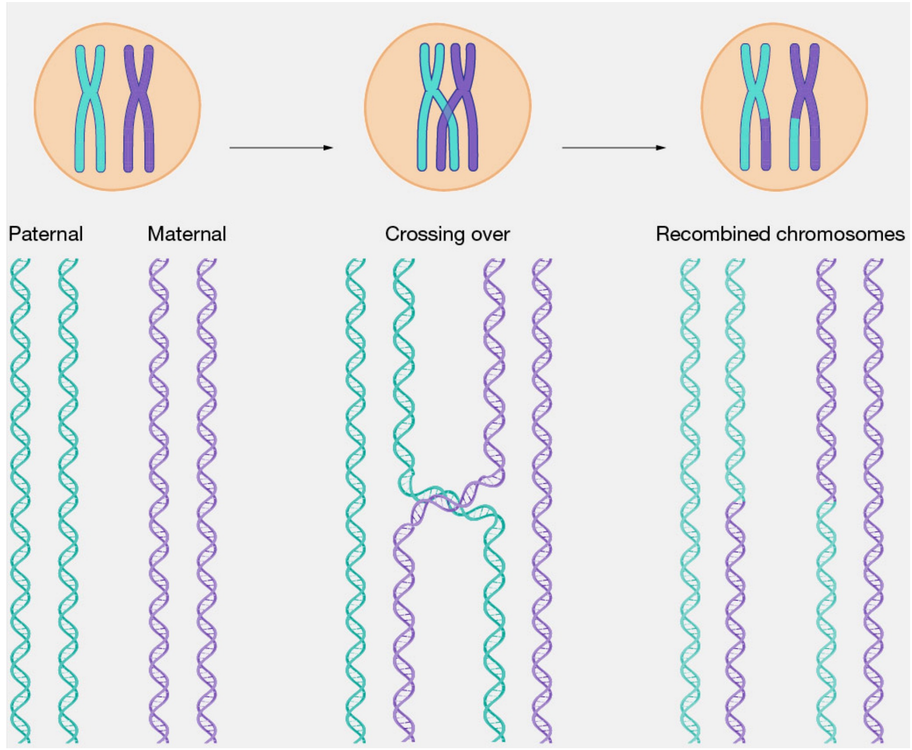
- Efficient through use of cell selectable markers
- Transgene location can be confirmed in ES cells
- Takes longer to have completely transgenic animals.



Transgenic Techniques - mammals

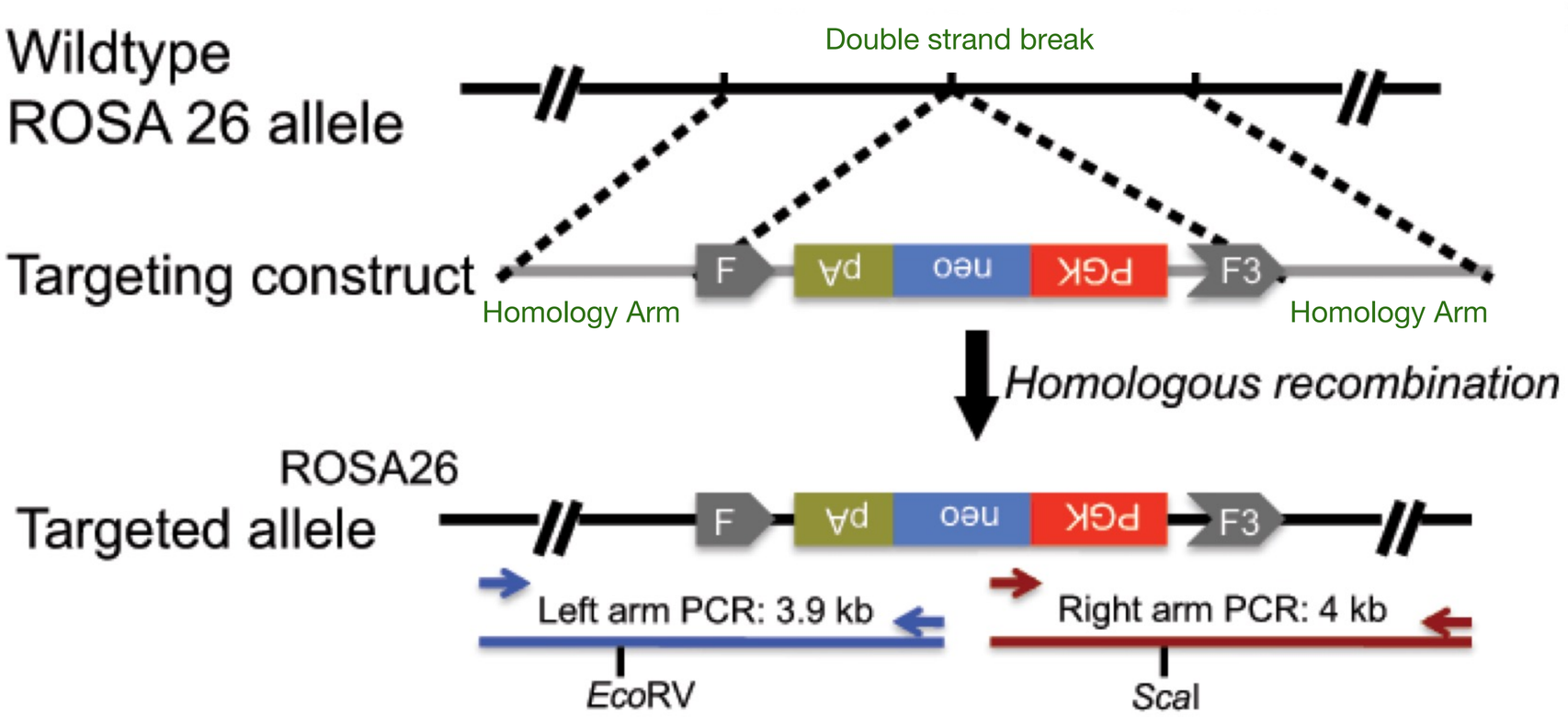
Targeting transgenes to specific locations

- Homologous recombination is a series of processes that enable the repair of DNA and allow interstrand crosslinks.
- essential to exploit the redundancy of genetic information that exists in the form of sister chromatids or homologous chromosomes
- Very important when both strands of the DNA double helix are compromised (double-strand breaks).
- Used during DNA replication somatic cells and during meiosis.



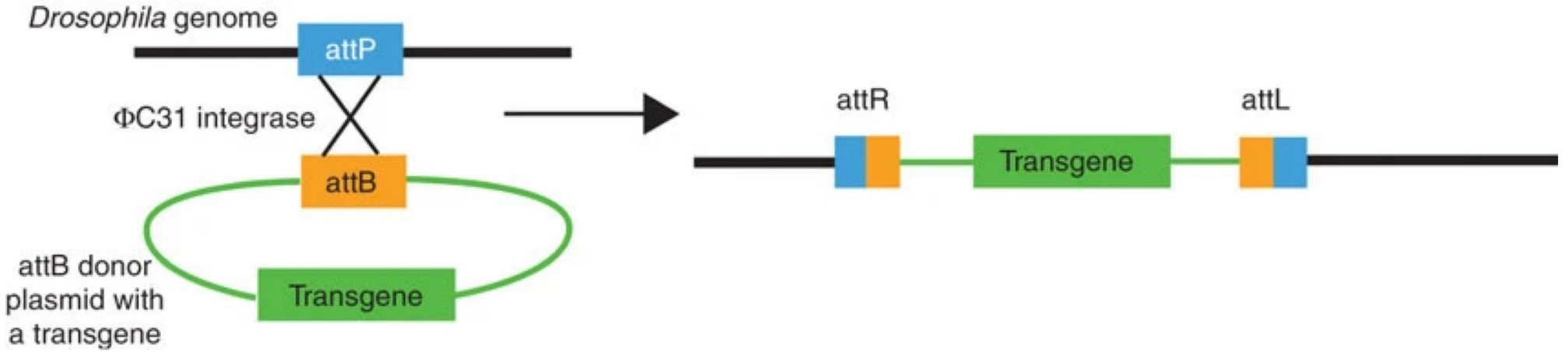
Transgenic Techniques - mammals

Targeting transgenes to specific locations



Transgenic Techniques - recombinase

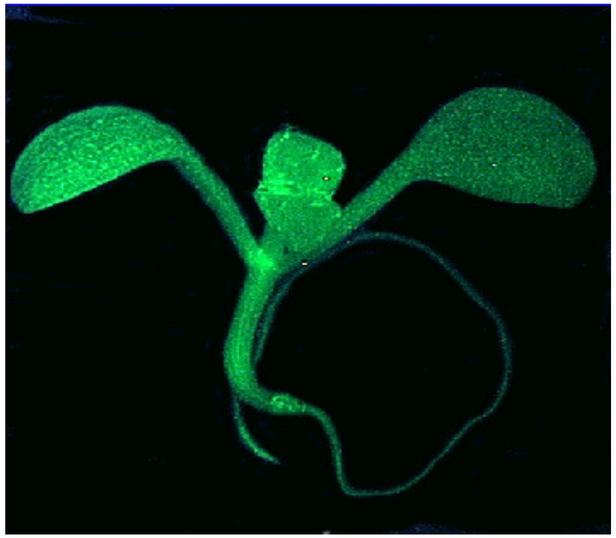
Drosophila - Targeting transgenes phiC31



- A preplaced (with transposon) attP sequence (attP) acts as the recipient site in the Drosophila genome. These sequences are derived from phages.
- An attB plasmid containing both a transgene and donor sequence (attB) is injected togthere with ϕC31 integrase mRNA into attP-containing recipient embryos
- This results in the site-specific insertion of the transgene into the attP site.

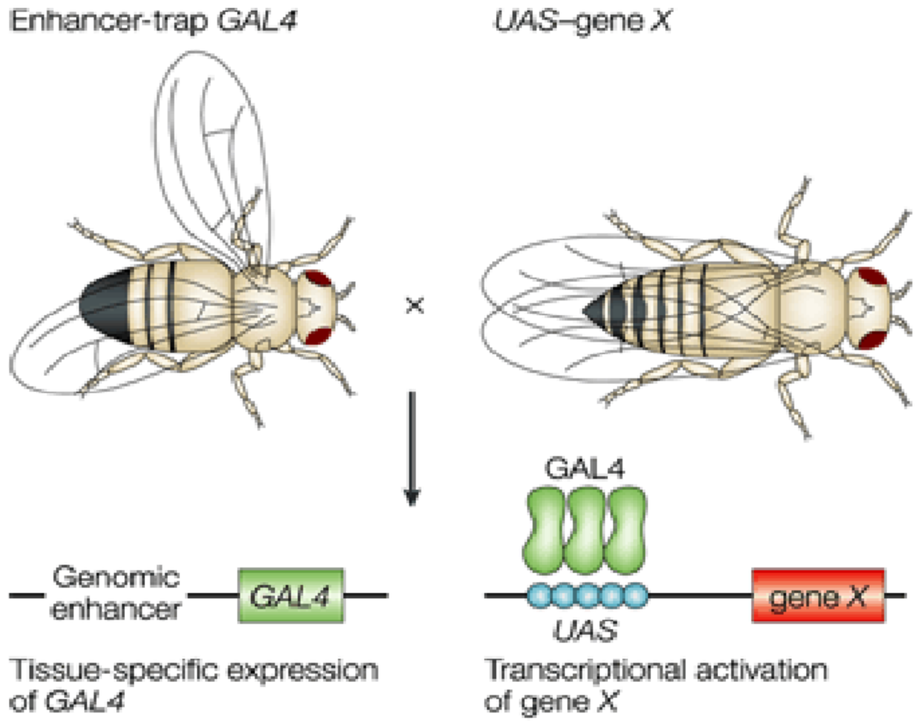
How do I know my animal is transgenic?

- Molecular confirmation e.g. PCR, commonly used in mice
- Selectable marker e.g. GFP
- Rescue of a mutant phenotype e.g. coat colour, eye colour



Controlling transgene expression - space

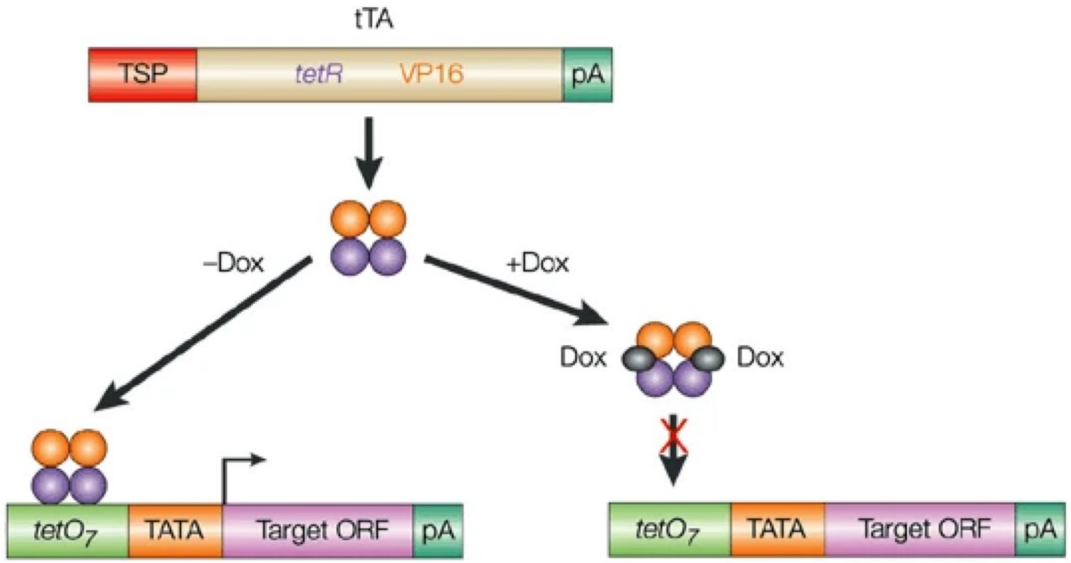
- Binary gene expression system
- Yeast derived Gal4 transcription factor is expressed under control on tissue specific enhancer sequence
- Transgene has Gal4 binding UAS (upstream activating sequence) before gene to be expressed. Without Gal4 – no expression
- In tissues that express Gal4, gene is expressed. Many UAS regulated genes can be expressed simultaneously.



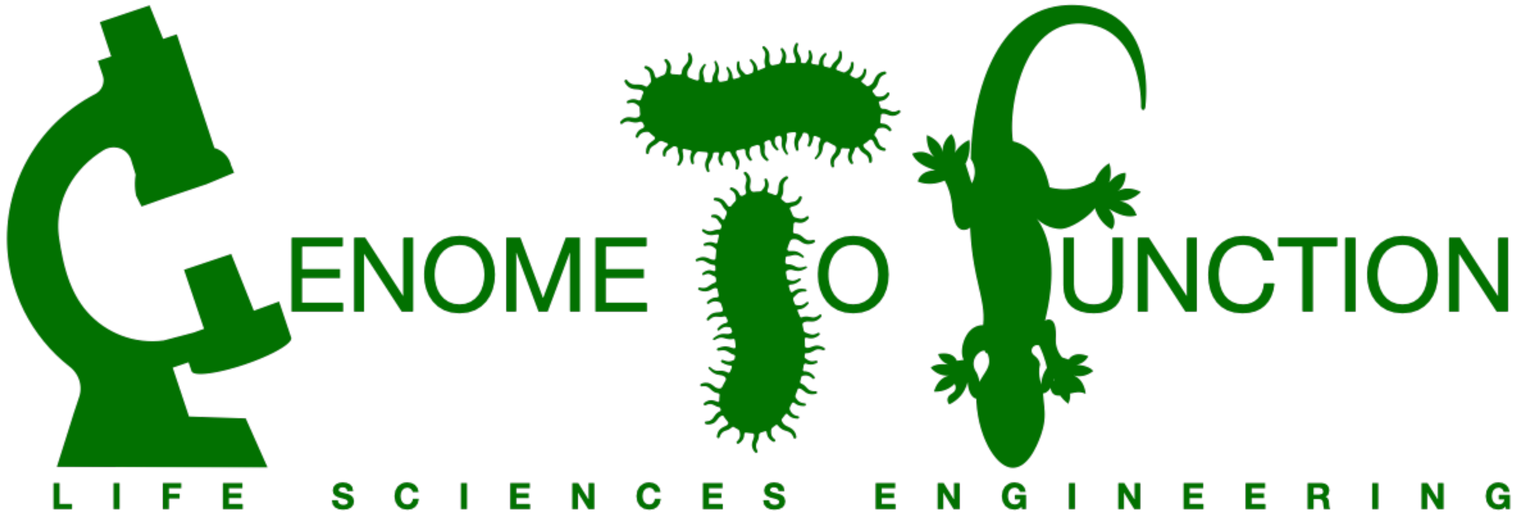
GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Controlling transgene expression - time

- **Tet-Off**
- Tetracycline-controlled transactivator (tTA) of transcription regulates gene expression.
- In the absence of the drug doxycycline (Dox), tTA dimers specifically bind to tetO sequences, activating transcription of the target transgene
- When Dox is provided , tTA undergoes a conformational change and cannot bind tetO sequences.
- Tet-On and other variants available



In Person quiz



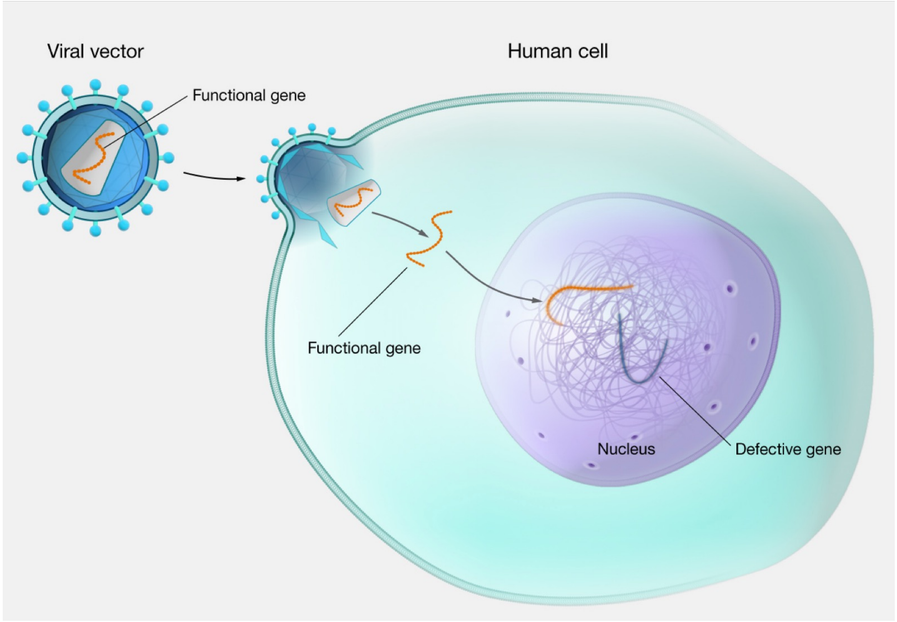
Gene Therapy

Gene Therapy

Treatng humans with transgenes

- Gene therapy uses genetic technology to treat, prevent or cure a disease or medical disorder.

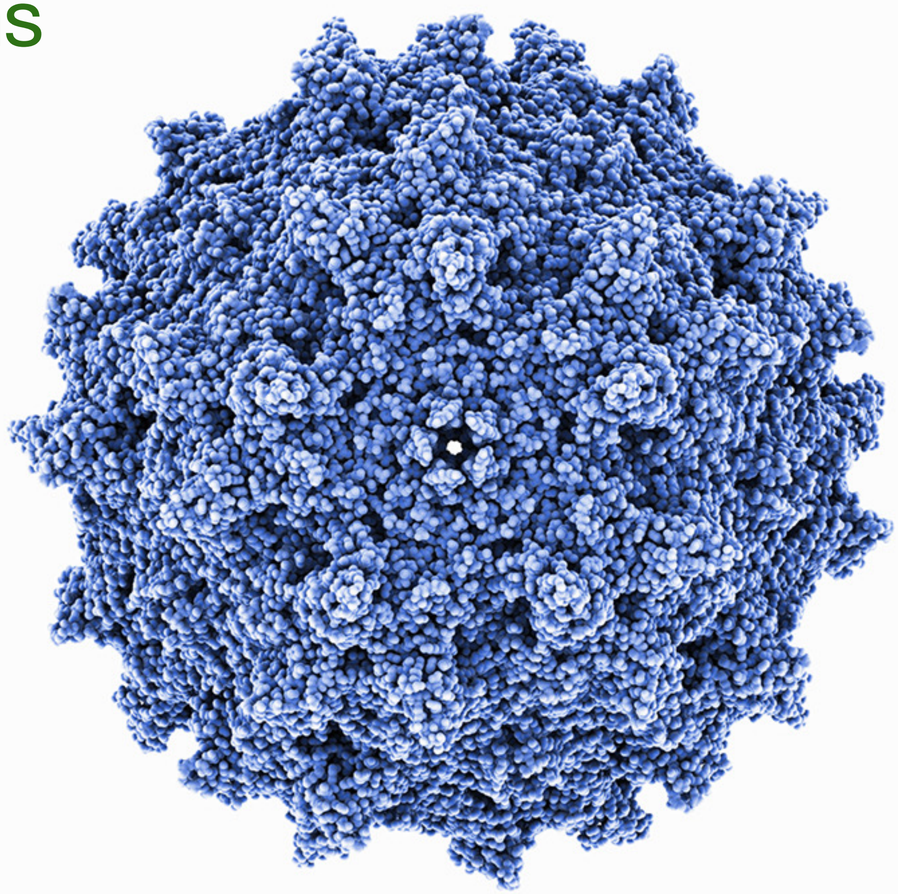
- Commonly utilises additional new copies of a gene to compensate for a defective or absent gene in a patient



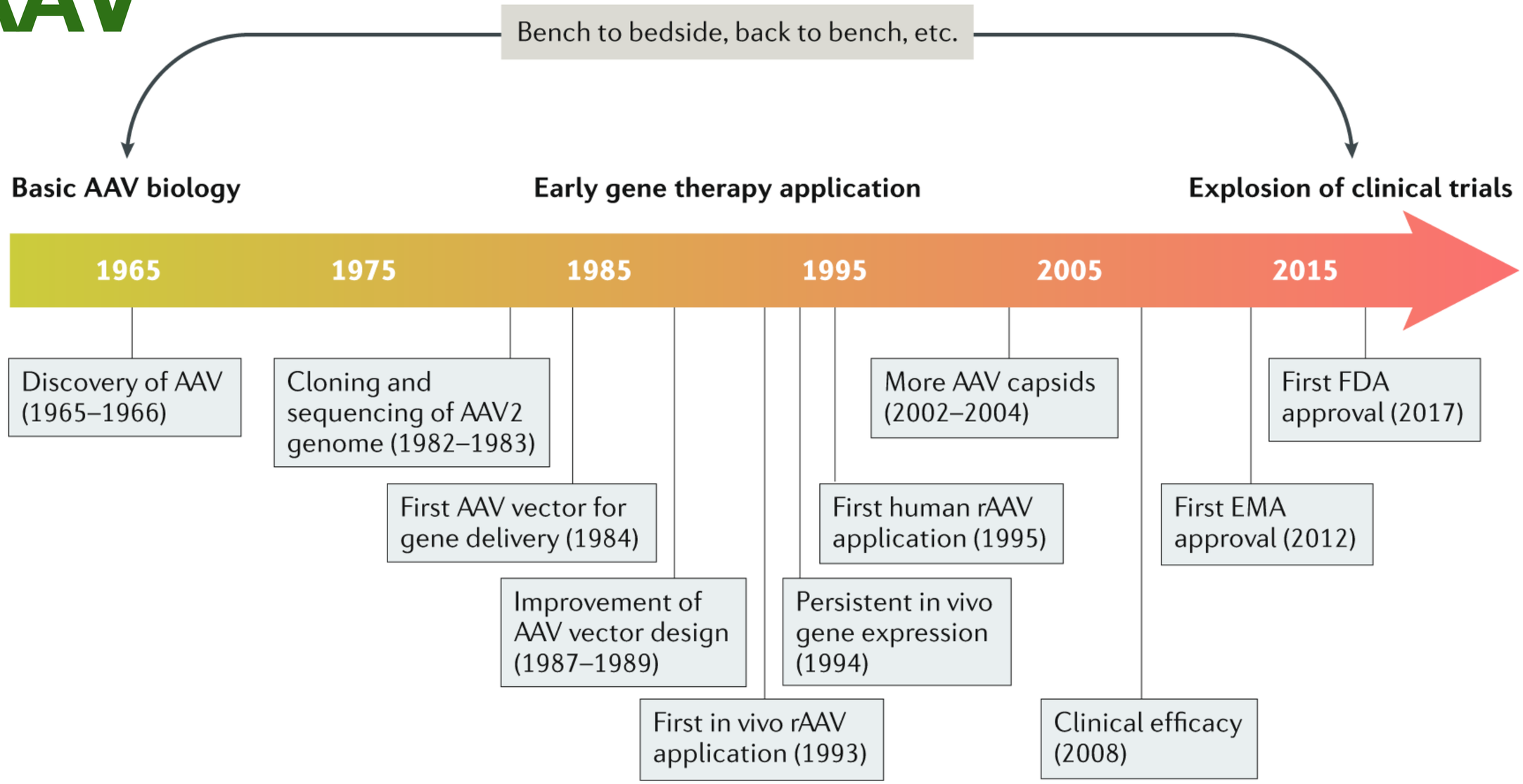
AAV

Targeting transgenes to specific locations

- Adeno-associated viruses (AAV) viruses that infect humans and some other primates.
- Are not thought to cause disease but can elicit a mild immune response.
- Single-stranded DNA genome of 4.8 Kb
- **Does not integrate into genome**

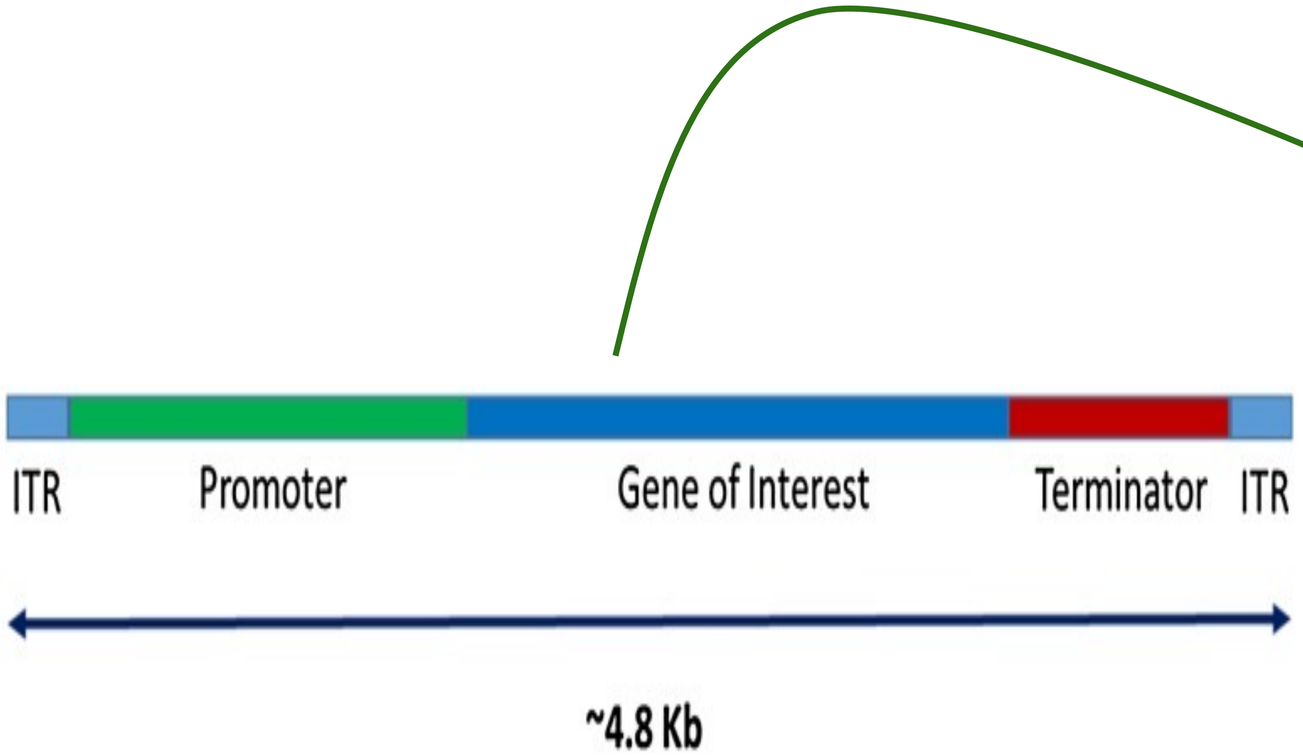


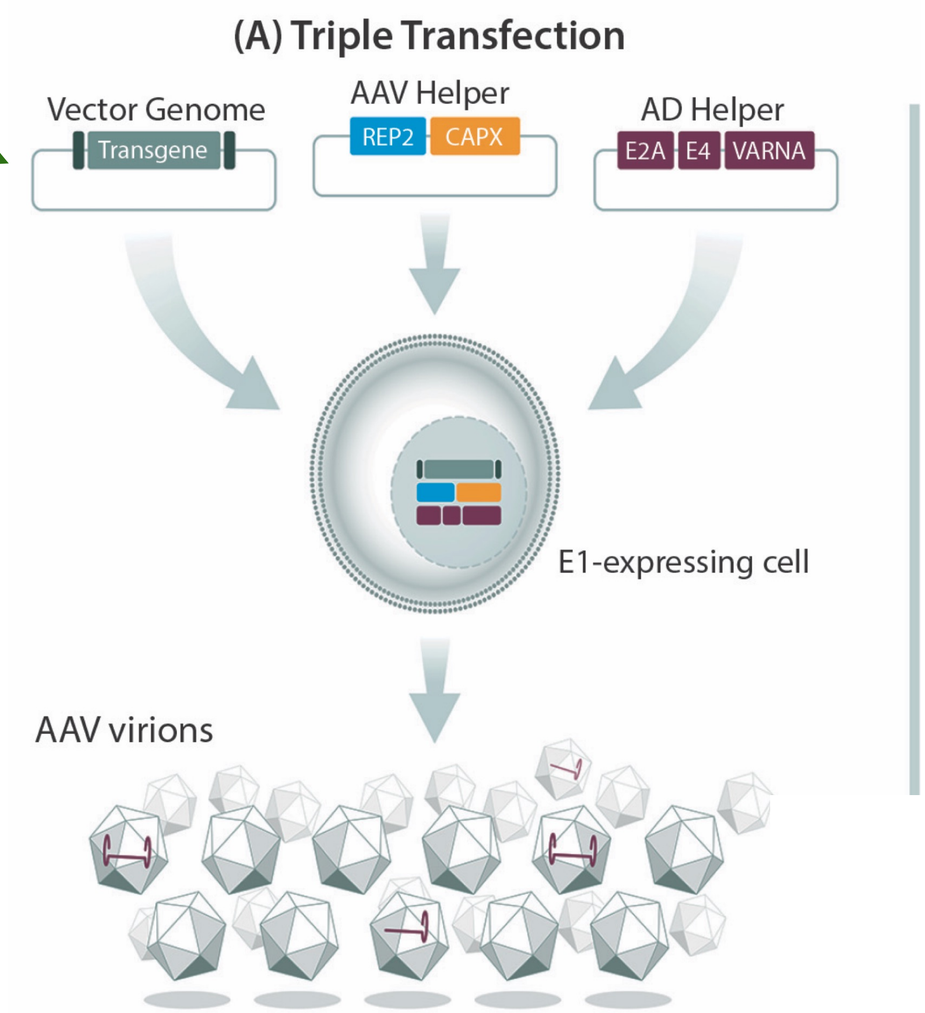
AAV



GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

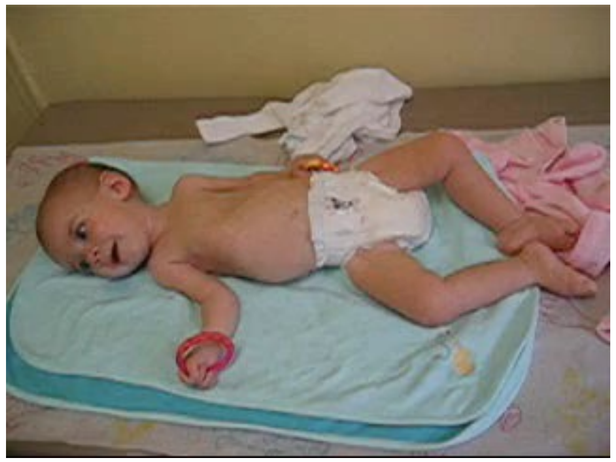
AAV





GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

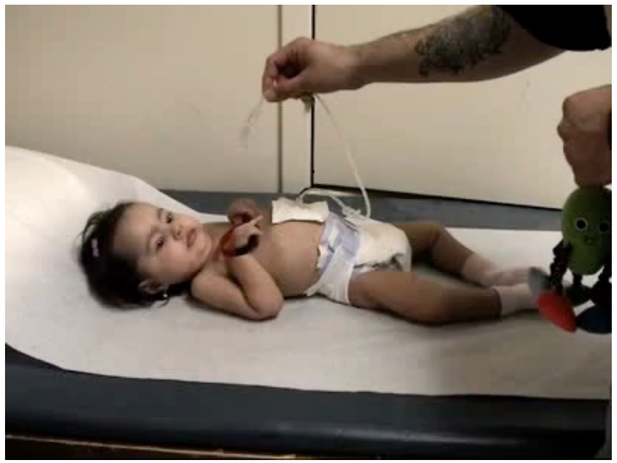
Spinal Muscular Atrophy (SMA)



SMA Type I

Severe form
Never sit
Limited life expectancy
Respiratory failure

Birth Prevalence 60%



SMA Type II

Intermediate form
Sitting or standing
Life expectancy shortened
Skeletal deformities

Birth Prevalence 27%

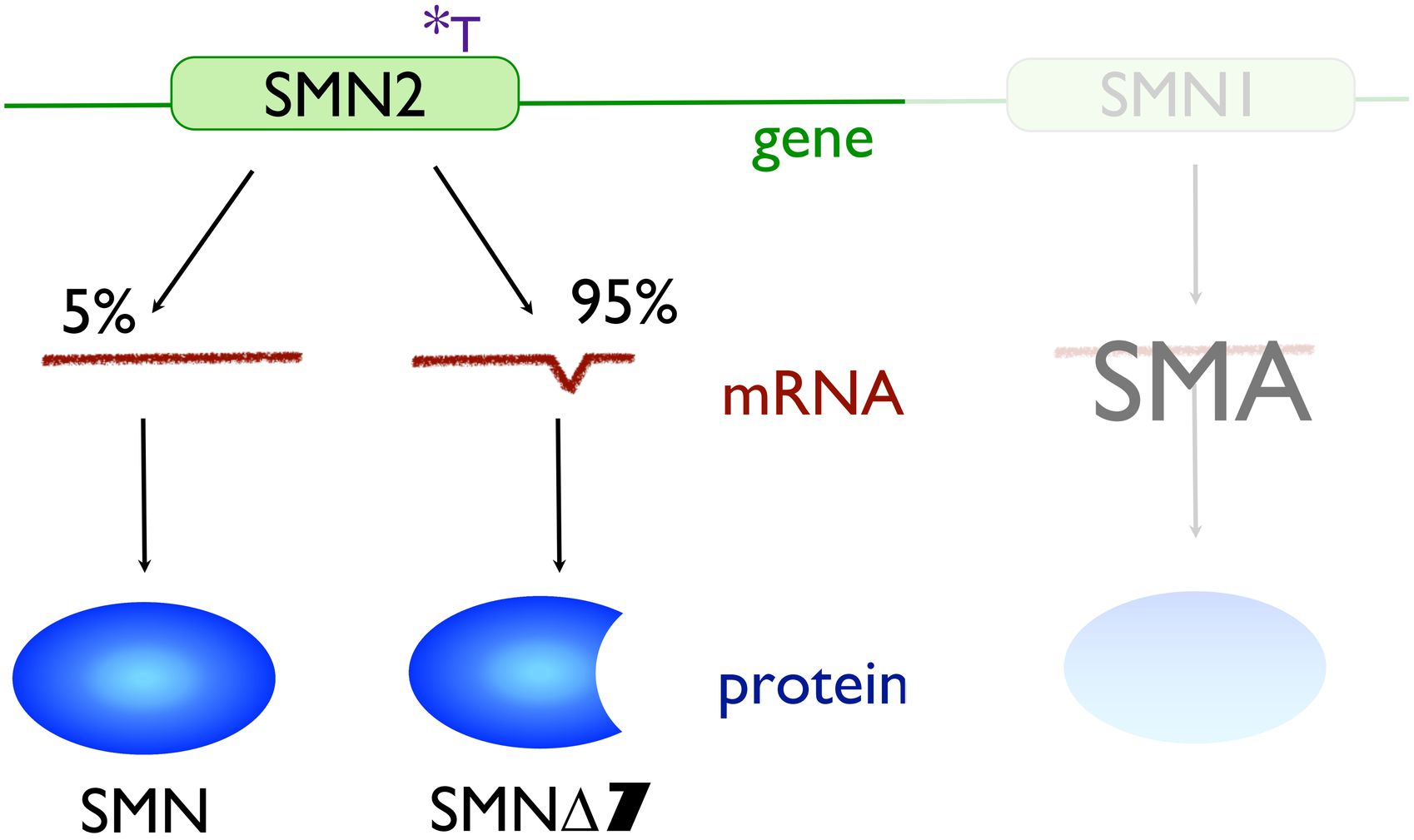
SMA Type III

Mild form
Walkers at some point
Life expectancy (nearly) normal
Proximal weakness prominent

Birth Prevalence 12%

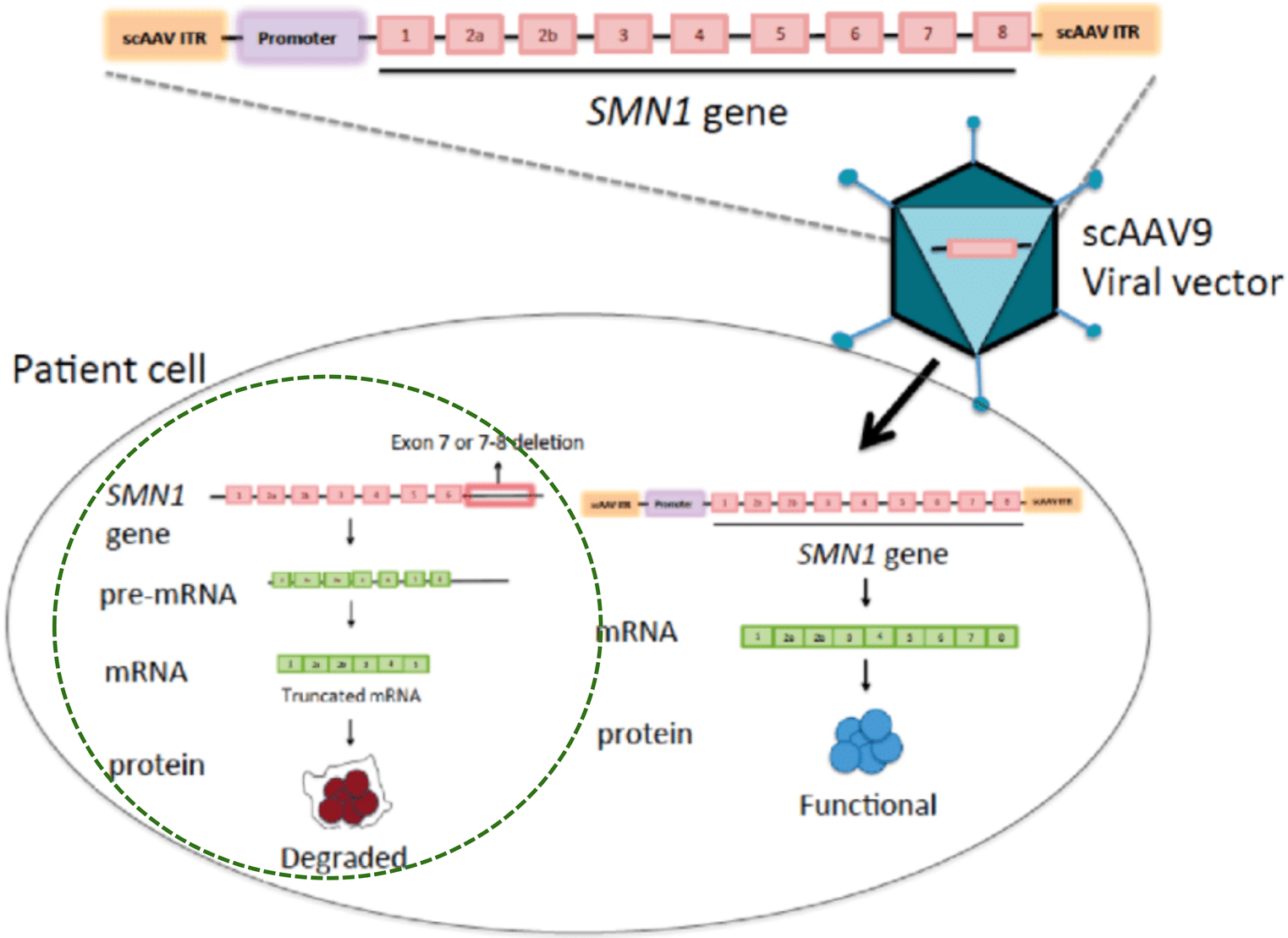
GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Reduction of SMN (Survival of Motor Neuron)causes SMA

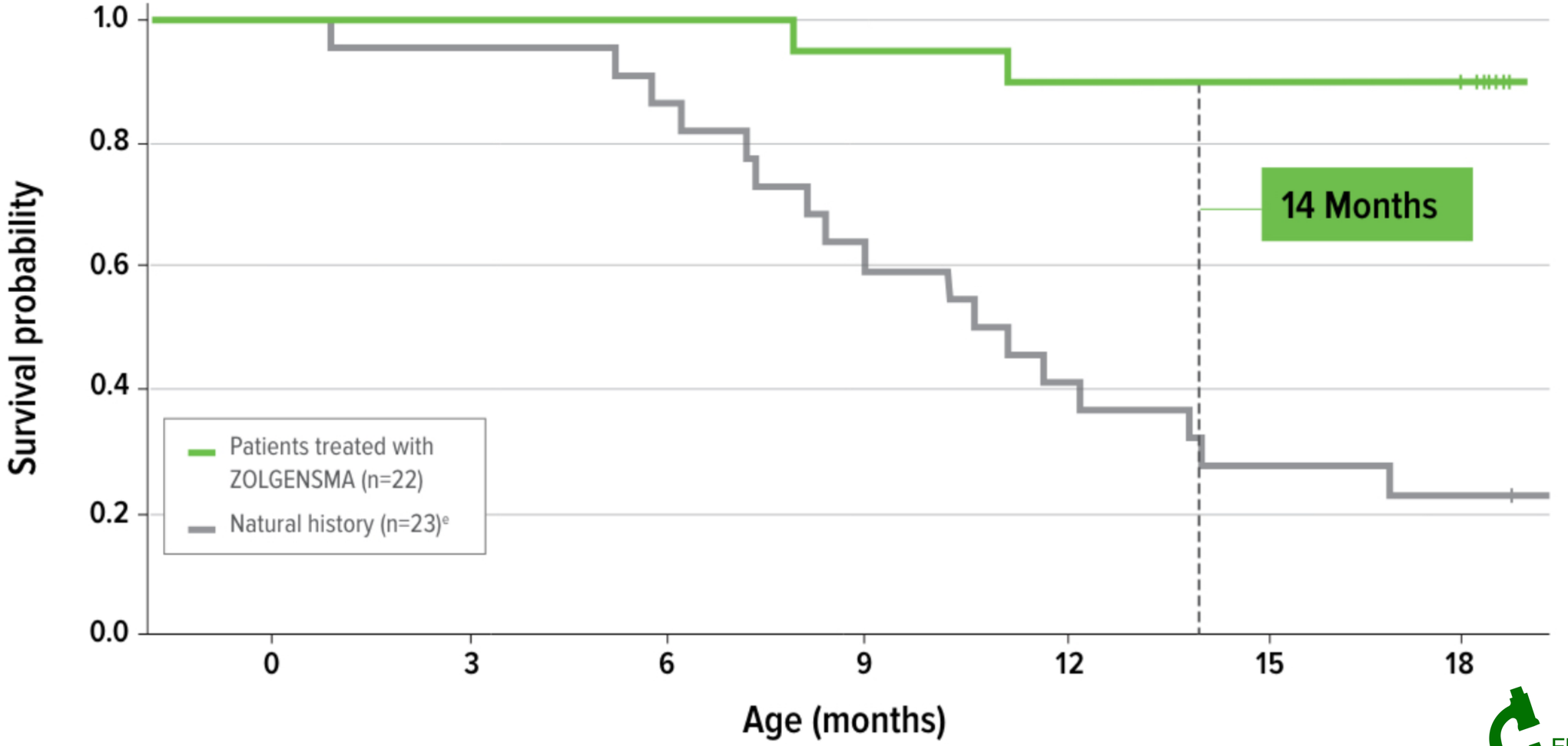


GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

Gene Therapy for SMA



Gene Therapy for SMA

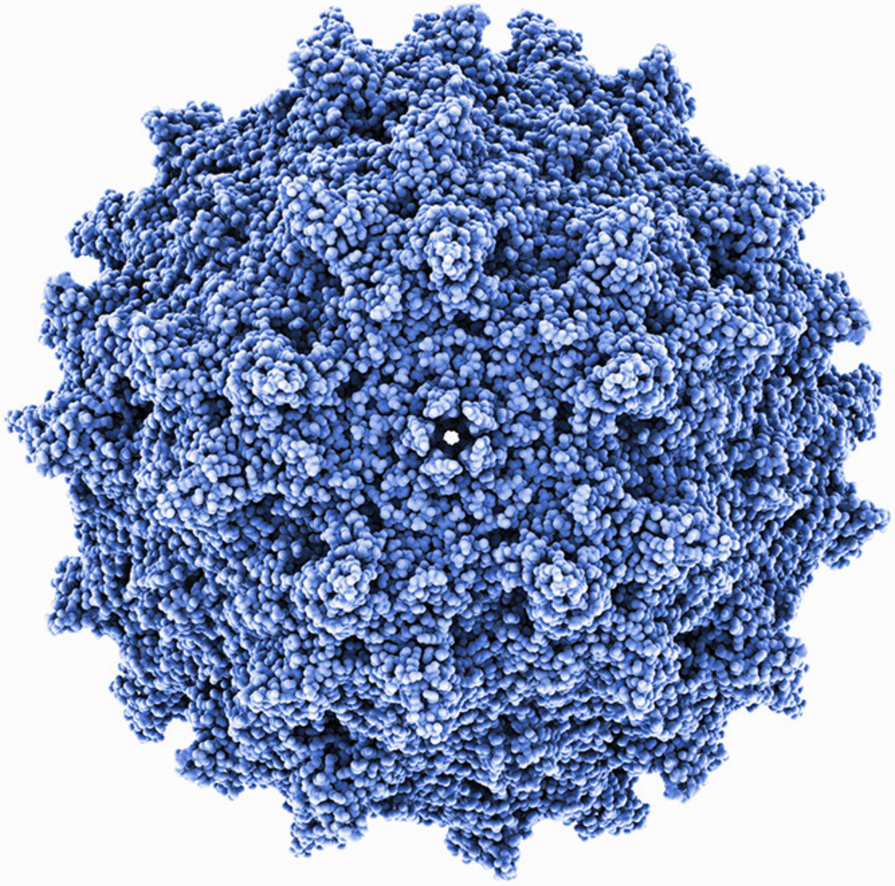


GENOME TO FUNCTION
LIFE SCIENCES ENGINEERING

AAV gene therapy

Advantages

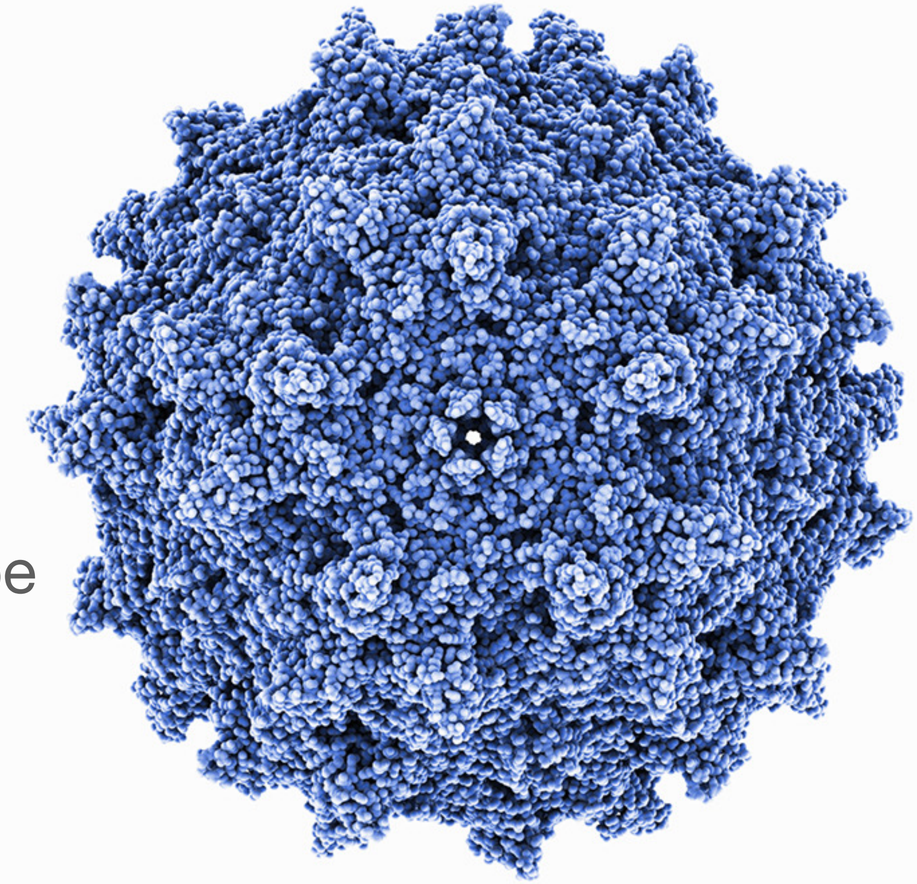
- Single administration
- Does not integrate – could cause mutation.
- Different types of manipulations can be achieved e.g. overexpression or reduction
- Can target cells hard to access with proteins e.g. brain
- May be faster to move from animal model to therapy – i.e. no slow drug development



AAV gene therapy

Disadvantages

- Small insert size
- Viral vectors have tropism for some cells.
- Hard to target large numbers of cells (e.g. entire brain in adults)
- Expression levels and pattern of gene may not be correct
- Non-integrating virus like AAV will be lost eventually to cell division.
- Expensive – SMA treatment was originally $2M!



Thank You & Questions