BMS3031

Lecture 1.1

1. Understand the basic features of recombinant expression systems used in modern research.
2. Be able to describe why different systems are chosen for different proteins and provide examples.

ONE

Recombinant properties:

- Can't get it from endogenous subjects
- Used for purification
- Quality control for large batches
- Optimize activity/efficacy of protein

Bacteria: first heterologous host for recombinant protein production. Others include:

- Yeast
- Plants
- Baculovirus: infects insect cells then the insect cells make the proteins
- Mammalian cells
- Animals

Bacterial Expression:

Pro	Cons
- Widely used	- Insoluble proteins
- Easy manipulation	- No post-translational mods
- Rapid growth	- High endotoxin content: requires removal which becomes time and cost expensive
- Cheap	

Many tags can be added to enhance purification

Common tags:

- Glutathione S-transferase
- Maltose binding protein: maltose can be used in E. coli as it folds in a way that is compatible to the bacteria and enables formation of a soluble protein.
- His-tag (on column cleavage)

Fusion proteins: proteins that are bound to a tag which mask their STOP codon. The tag is attached to the gene and then the tag and gene of interest is all transcribed and translated as a whole unit.

1. Target gene is fused with tag
2. Fused gene is transcribed and translated into a fused protein (sometimes the tag binds to an affinity column)
3. Wash
4. Elution: process of extracting one material from another by washing it in solvent
5. Tag cleaved → protein purified

Thrombin is an elusion factor that cleaves the tag.

TWO

Fusion Protein:

Pros:	Cons:
Improves yield	Still low yield
Prevents proteolysis	Alters biological activity
Enables refolding	Cleavage of tag might require other proteases (therefore more time and cost expensive)
Increase solubility	A reducing environment limits secretion capacity
Easy to purify	

When making recombinant proteins, it is important to identify the internal environment of the organism you are using. For example:

- E. coli environment is reducing. Therefore, S-S bonds are hard to make
- Its periplasm offers a more oxidising environment enabling proteins to fold more easily. Therefore, when making a fusion protein you must use one that is signalled to be targeted into the periplasm.

Inclusion bodies: misfolded proteins

- Contain insoluble proteins
- Formation depends on synthesis rate and growth conditions
- Often in mammalian proteins

To limit inclusion bodies, proteins are fused to proteins that are targeted to the periplasm, where they are more likely to fold correctly.

Fusion proteins are a favourable alternate recombination protein as they can be purified easily by cleaving the tag.

INSULIN

- From: Beta cells of islets of Langerhans in pancreas
- Stimulated by glucose, amino acids, fatty acids
- Single gene trafficked from ER→cis golgi→trans golgi→ secreted
- Proteolytically processed: post-translational modifications
- A and B chain, C chain is cleaved off and A and B form disulphide bonds

Tried to used yeast to make insulin, but yeast lacks enzymes required for removal of the C-peptide. So they created a fusion protein containing a signal sequence and insulin precursor (with Lys-Arg sequence). The signal sequence directions the fusion protein to the secretory pathway of the cell which exposes the protein to Lys-Arg cleaving enzyme (trypsin cleaves after Arg, carboxypeptidase cleaves Arg and Lys from B chain) = recombinant insulin.

E. coli was difficult to use as it inhibited S-S bonding and had no secretory pathway for post-translational modifications

YEAST

- Cheap
- Simple and easily manipulated
- Good for S-S bonds, bad for glycosylation as it lacks human glycosylation patterns
- Must have yeast specific plasmids
- Has secretory pathways but lacks enzymes to remove C-peptide

ERYTHROPOIETIN (EPO)

- Synthesized in kidney
- Haemopoiesis: production of blood cells/platelets from bone marrow
- Expensive because they take a long time to replicate in mammalian systems
- 3 N and 1 O-glycosylation site (glycosylation fundamental for EPO action)

Glycosylation:

- Most proteins synthesised in rough ER glycosylates by adding an N-linked oligosaccharide
- 50% of eukaryote proteins are glycosylated involving 700 different enzymes
- DISADVANTAGE: autoglycosylation can result in autoimmune diseases
- ADVANTAGE: facilitates trafficking, folding, stability and function

Culturing:

- Conventional media: Dulbecco's Modified Eagle's Media
- Extra factors: foetal calf serum (growth hormones)
- Red colour: phenol red used as a pH indicator. Colour change to yellow indicates presence of cell from metabolic by products
- Lines used are immortal

CHO cells: Chinese hamster ovary cells

- Certified cells
- Serum-free media developed
- Large yield (10 g/L)
- Can be used large scale
- Problems: high doubling time

Comparison of therapeutic proteins

Characteristic	Bacteria	Yeast	Mammalian
Cell growth	Rapid (30 min)	Rapid (90 min)	Slow (24h)
Cost	Low	Low	High
Expression	High	medium-high	low-medium
Protein folding	Not reliable	Usually reliable	Very reliable
S-S	Poor	Good (secreted)	Good (secreted)
Glycosylation	None	Simple N-linked	Complex
Phosphorylation	No	Yes	Yes

Atryn: anti-coagulant that is produced in goats milk via transgenic expression

- Renominated anti-thrombin alpha
- Done via beta casein promotor (found in goats milk)
- Process:
  1. Human gene (that produces antithrombin) is inserted into goat DNA
  2. Transformed gene is inserted into an egg and implanted into goat

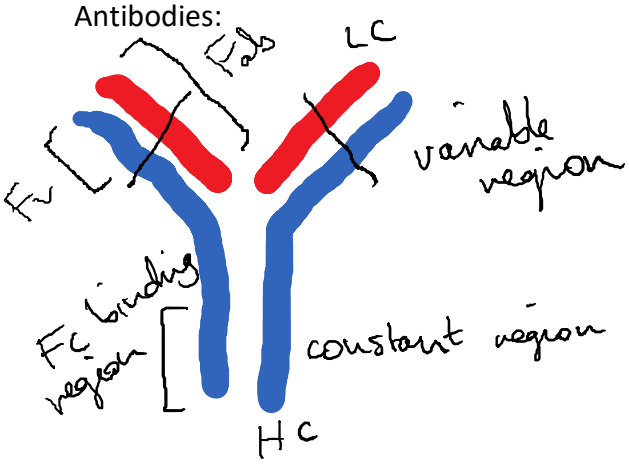
3. Resulting offspring are tests for anti-thrombin in their milk
4. Milk is filtered/purified (each goat produces as much anti-thrombin as 90,000 human blood domations/year)

Lecture 1.2

1. Be able to describe the structure and function of antibodies and engineered forms.
2. Understand the basics of how monoclonal antibodies are developed.
3. Understand the how and why monoclonal antibodies are "humanised".
4. Be able to use examples to describe the importance of antibodies as therapeutic agents.

ONE

Antibodies:



In the Fv section of the CDR loops, componentry defining region.

Componentry: parts of a machine considered collectively.

IgG: main serum antibody and most stable

CDC: complement dependent cytotoxicity

ADCP: antibody dependent cellular phagocytosis

ADCC: antibody dependent cell mediated cytotoxicity

Moore's Law: ideal decrease in cost that indicates the product is worth the investment and further development

Therapy Mechanism:

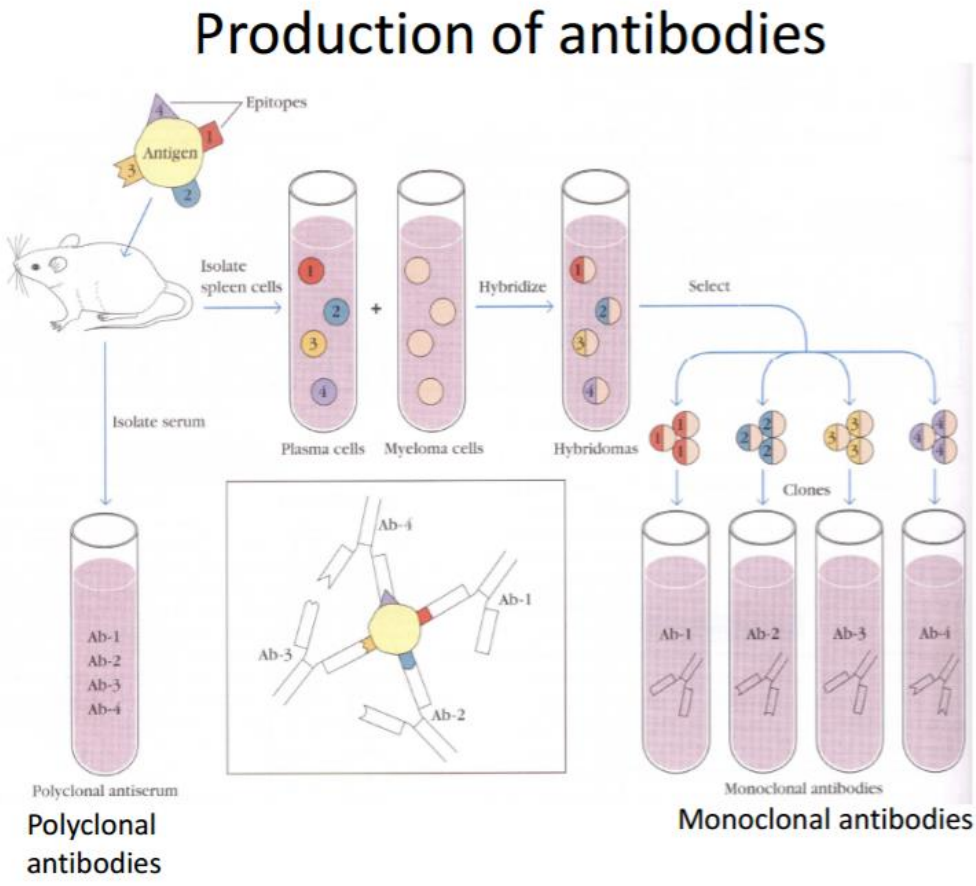
Used for autoimmunity and inflammation

- Ligand blockade
- Receptor blockage
- Receptor downregulation
    - Uses endocytosis to suck receptor back into cell
- Depletion
    - CD20 (B-lymphocyte antigen) is an activated glycosylated antigen on the surface of B-cells and is responsible for enabling immune function. Antibodies bind to CD20 antigen on B-cell surface which causes complement activation. MAC (membrane attack complex) and an effector cell bound to the antibody initiate cell destruction such that an autoimmune response does not arise.
- Signalling induction

- In the absence of a ligand, an antibody can initiate altered cell function via signal transduction.
- Example: T-cell receptor CD3 complex can induce T-cell receptor mediated signals that alter T-cell function and differentiation

TWO

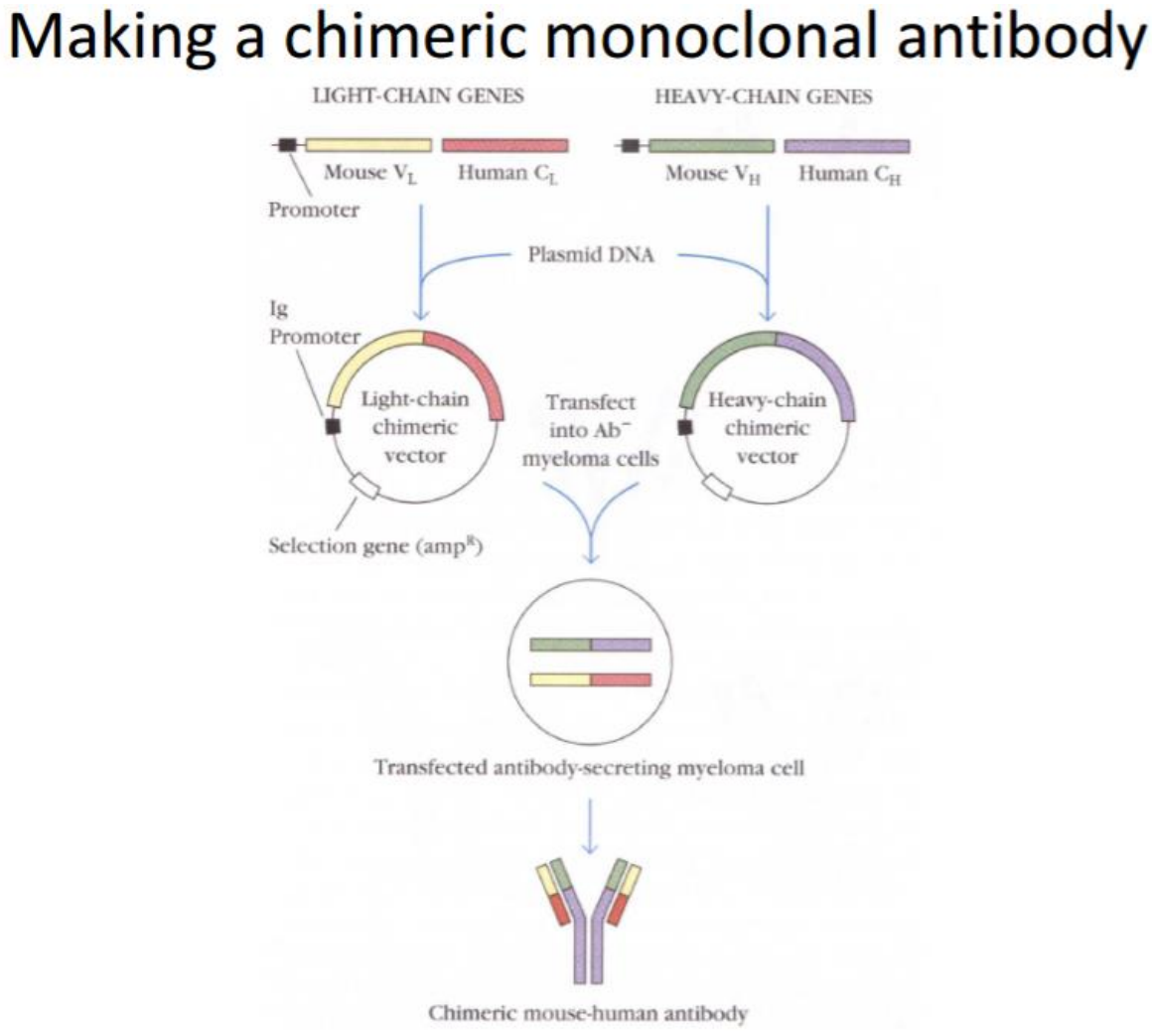
Steps:



1. Animal is injected with virus (attenuated or killed or live depending on its effect)
2. Animal makes polyclonal antibodies
3. Booster is given to increase number of antibodies in serum
4. Serum is harvested and antibodies purified. However, there are thousands of isotypes (all have different receptors)
5. Monoclonal production:
    a. Activated B cells are fused with myeloma cells to form a hybridoma.
    b. Myeloma lack enzymes that enable them to synthesise nucleotides via the salvage pathway.
    c. Culture in a HAT medium that selects for the positive antibody of interest.
        i. To select for the antibody of interest, add antigen and observe binding.
        ii. This is repeated a few times so only high affinity binding is left
    d. To remove the myeloma:
        i. Turn off enzymes in the myeloma that allow them to make nucleotides the normal way. Now that they are unable to synthesis nucleotides, this kills the cell leaving the antibody of interest behind.
    e. Monoclonal antibodies are harvested.

Antibody Engineering:

- Genetic engineering
- Gene cloning
- CDR grafting
- Eukaryotic expression



THREE

Steps:

1. Variable light mouse chain gene is bound to a constant human light chain gene.
2. Variable heavy mouse chain gene is bound to a constant human heavy chain gene.
3. These genes are inserted into two separate plasmids and transfected into an Ab myeloma cell.
4. The antibody secreting cells are then selected for which produce an antibody with a human constant region (region which is detected by human T-cell receptors) and a mouse variable region (regions that come into contact with the foreign antigen)

E.g. Rituximab treats non-Hodgkin's B-cell

- It is a chimeric anti-human CD20 Mab. It binds to the CD20 receptors on B-cells, signalling compliment to activate NK cells and macrophages to kill the overactive B-cell.

Antibodies and phages exist in a library setting.

- Libraries are created by amplifying (via PCR) different antibody chains separately. In order to create a minimal binding unit, only variable chains are amplified. The chains are then paired at random multiple times to create a library. These pairs are then amplified in E. coli and presented on the surface of the cell = scFv

Minimal binding unit

1. Antibodies are large and this causes problems when they are secreted or when they move across membranes from bacteria to bacteriophage. Therefore, small units containing only the variable regions of heavy and light chains can be used to overcome this problem.

Antibodies and phage display:

1. Plasmids containing heavy chain fragment gene and light chain fragment gene are inserted into a viral vector.
2. The bacteriophage secrets the plasmid into a bacterium where the chains are created using bacteria machinery.
3. The chains are secreted across the membrane (hence why not whole antibodies are used, only their fundamental parts) and reattach to the surface of the bacteriophage.
    a. Single chain Fab/Fv molecules are expressed on the surface of phage
4. The cycle repeats.
    a. Phage is easily removed by lowering pH.

Library features:

- Allow creation of synthetic immune systems
- Eliminate use of animals to generate antibodies
- Circumvent immune tolerance
- Can identify multiple binders to synthesis superior molecules

Tumour Necrosis Factor (super family of cytokines and causes apoptosis of infected cell)

- Selection of phage antibodies is done by TNF affinity.
- Phage uses bacterium to create antibodies.
- TNF affinity selection occurs again until only very high affinity antibodies remain

Generation of Human MAb (monocolonal antibody) against Human TNF: Adalimumab

- MAK195 (mouse MAb against TNF) is used to guide selection of human Mab
1. Heavy chain of MAK195 is paired with multiple human light chains
    a. This creates a phage library
2. At the same time, mouse MAk195 light chain is paired with multiple human heavy chain
    a. This creates a second phage library

3. Antigen (human TNF) is then used to select Abs from libraries
4. Remaining human heavy/light chains pair together and tested for:
    a. Antigen specificity
    b. Function

Successful completion of this project produced Adalimumab (humira) which targets TNFalpha to treat autoimmune diseases.

- Chrohn's disease
- Rheumatoid arthritis

OTHER APPROACHES

Antibody Fragments:

- Involves identifying variable domain bindings sites
- These regions can block toxin interaction and can be made in tandem to be bispecific
- Done in vitro, they can be made using phages and screened (as mentioned above via affinity columns)
- Advantage: don't need to elicit immune response in animals

Engineering IgGs

Process	Result
Mutate variable domain using IgG library	Alters specificity/affinity
Mutate Fc region→ select IgG isotype	Varies ADCC/ADCP/CDC and half life
	Lacking Fc fragment decreases CDC/ADCC/ADCP and decreases half life
Glycosylate by:	
- Aglycosylation	
- Bisecting N-acetylglucosamine	
- Non-fucosylation	

FOUR

Example: Zmapp and Ebola

Combination of 3 humanized Mabs in glycoprotein 1 in the Ebola virus

- Produced in plants
- Lacks N-glycans containing mammalian enzymes
- Lacks core fucose = more effective
- 2/3 patients recovered

Lecture 1.3

1. Give examples of different therapeutic approaches used for personalised medicine
2. Understand how genes can be edited
3. Understand the general differences between ZF-nucleases, TALENs and CRISPR/Cas9
4. Be able to use examples to describe how genome-editing approaches are being used for research & current/future therapies
5. Gain an appreciation of the hurdles involved in genome engineering technologies for therapy

ONE

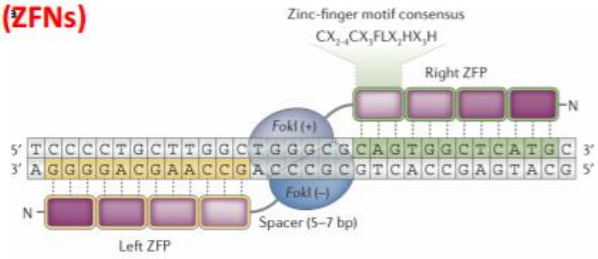
Personal Medicine: diagnoses based on what works best for the individual patient

- 20%+ of new molecular entities

Genomic Medicine: uses genomic information as part of clinical care

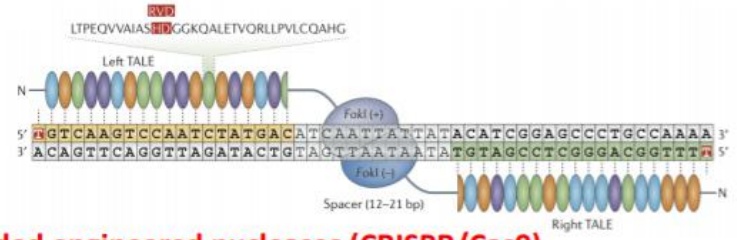
THREE

- **Zinc finger nucleases (ZFNs)**



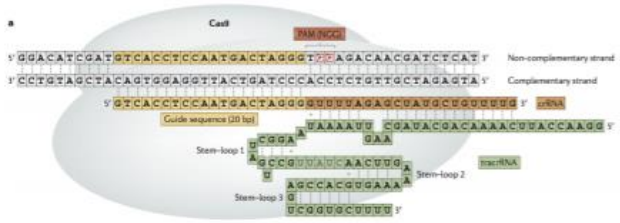
Series of protein modules that bind 3 nucleotides bringing Fok1 nuclease to site (dimer becomes active)

- **Transcription activator-like effector nucleases (TALENs)**



Series of protein modules that bind individual nucleotides bringing Fok1 nuclease to site (dimer becomes active)

- **RNA-guided engineered nucleases (CRISPR/Cas9)**



RNA molecule (CRISPR) designed that binds nucleotides and targets Cas9 nuclease to site

Zinc Finger Domains: first eukaryote transcription factors discovered

ZFNs are fused with Fok1 to form a fusion protein. Fok1 is complementary for a specific, small sequence of DNA (5'-GGATG-3'). Once two Fok1s have bound on both strands of DNA, they dimerise and this activates ZFNs to cleave the DNA with a double-stranded break.

- Multiple Sequence Alignments
- First gene-specific eukaryotic transcription factor discovered
- Regulates transcription by RNA polymerase III
- Binds to internal control region of 5S rRNA gene (DNA)
- Also binds to the 5S rRNA itself
- Contains 9 zinc finger domains

Has positively charged lysine residues:

- Allows the protein to interact with the +ve charged DNA

Can determine DNA bound structures:

- Zinc Fingers 1-3 bind to 5S RNA gene (DNA)
- Each finger contains one Zn2+ ion
- TFIIIA: independent units that are stabilised by Zn^{2+}
    1. Homologous regions across humans and African claw frog
    2. Hence these conserved regions typically have important structural o functional benefits

TALENs:

-

CRISPR-Cas: clustered regularly interspaced short palindromic repeats (situated next to Cas-protein coding genes)

- Used to genetically modify monkeys.
- Done by using CRISPR to locate Cas encoding exogenous genetic regions and cutting at the desired location. The existence gene can then be removed or replaced by another gene with specific endonuclease enzymes
- CRISPR-Cas system: provides a form of acquired immunity

Advantages vs. Disadvantages

	ZFN	TALENs (redundant)	CRISPR/Cas9
Targeting specificity	Zinc finger proteins	Transcription activators	RNA (gRNA is much easier/cheaper and quicker to use than TALENs)
Target constraints	Difficult to target non G-rich sequences	5' target base must be a T	Position must precede PAM
Ease of engineering	Difficult, requires extensive protein engineering	Moderate: requires complex cloning methods	Easy: uses standard cloning
East of in vivo delivery	Easy	Difficult to due large size of TALENs = unwanted recombination in lentivirus	Moderate: if using smaller orthologues. Harder to use Cas9 because of its large size
Nucleases	Fok1	Fok1	Cas9
Success	Low	High	High
Mutation rate	Low	High	High
Target site length	18-36bp	30-40bp	22bp
Design Density	One/100bp	1+/bp	1/8bp
Off target effects	Moderate	Low	Medium due to small recognition sequence
Cytotoxicity	High	Low	Low
Size	1 kb x2	3 kb x 2	4.2 kb + 0.1 k

FOUR

Personalised med is transformed by genomic med. Cost of genome medicine has greatly reduced

Rubraca: used to treat patients with BRCA-mutated advanced ovarian cancer.

- Poly ADP-ribose polymerase inhibitor
- Trial in individuals with two or more chemotherapy regimens

Epclusa: treatment of Hep C infection

- Without cirrhosis or compensated cirrhosis
- Contains 2 chains of viral protein found in Hep C virus

Atezolizumab: Humanized MAb (from phage) binds to PD-1 (receptor on T-cell) and blocks interactions between B7.1 receptor. Blocked interactions inhibits glycosylation sites and allows T-cell to initiate immune response

- PD-L1 is a receptor that is expressed on the surface of a tumour cell. It binds to PD-1 on a T-cell and inhibits T-cell activity. Meanwhile, tumour cell expresses an antigen that is detected by T-cell receptor.
- Humanized Mab binds to PD1 and blocks interaction with PD-L1 which enables T-cell to kill tumour cell.
- PD-L1: checkpoint protein and biological marker for target treatment

Exondys 51: used to treat Duchenne's muscular dystrophy (muscle wasting disease). Life expectancy: 25 years.

Dystrophin:

- 2.4 Mb gene
- 79 exons
- 427 kDa (protein size)
- X-linked disease (in males)

DNA modification paves the way for gene therapy. It usually involves a viral mediator which introduces corrective gene into defective cells. ADA (adenosine deaminase) defect in SCID is corrected by retroviral gene therapy, but the process was stopped in 2002.

- Stopped because patients being treated with gene therapy were developing leukaemia.
    - This is because a retrovirus was being used and injecting corrective DNA randomly into cell genome. If injected next to an oncogene, upregulation results in tumour formation.
- Costs $1 million per treatment
- Involves 60 injections in a single series
- Treats lipoprotein lipase deficiency: milky blood (unable to process fats)

TWO

Gene editing

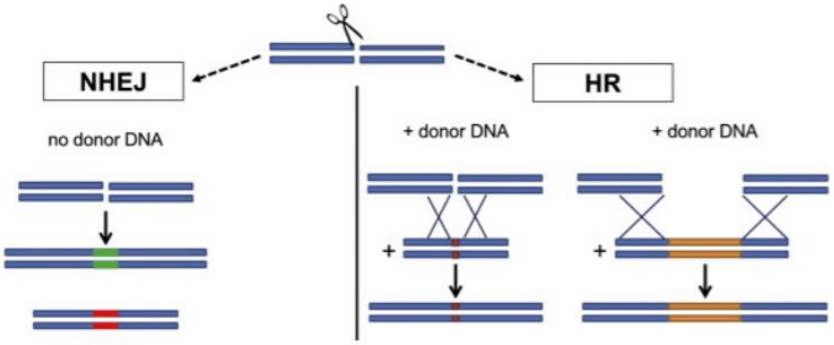
Sense (5'-3' direction)

Anti-sense (3'-5' complementary to 5-3 strand)

Exon-skipping by anti-sense oligonucleotides which can block recognition of intron-exon splice sites.

- Block recognition splice sites
- SSO and splicing factors are then released with an extra exon. Frame shift mutations cause disease. Deleting an exon can bring the nucleotide sequence back into the correct reading frame.

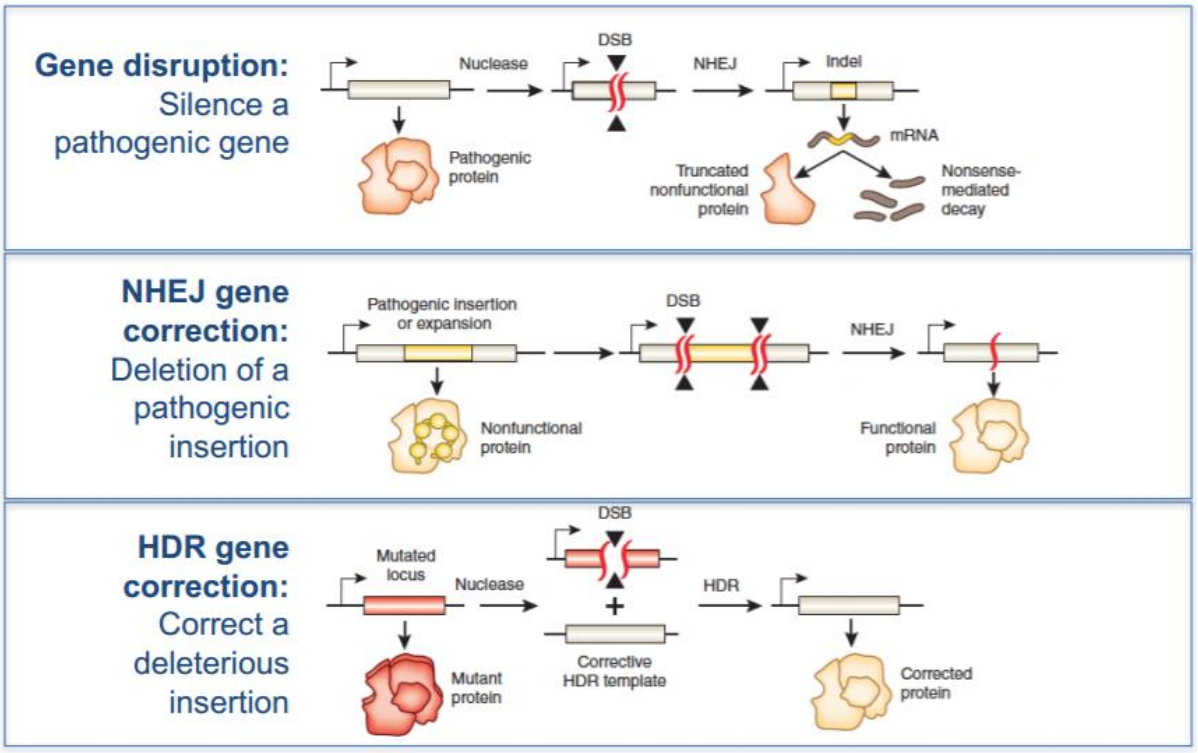
Modifying genes following the introduction of double strand breaks



NON HOMOLOGOUS END JOINING
Intrinsically error prone & results in formation of **indels** *(small insertions/deletions of typically 5- 20 nucleotides).*
Can disrupt reading frame of genes if targeted appropriately

HOMOLOGOUS RECOMBINATION
Requires the addition of template DNA with homologous arms but can have variable sequences (mutations, insertions) in between.

Types of modification



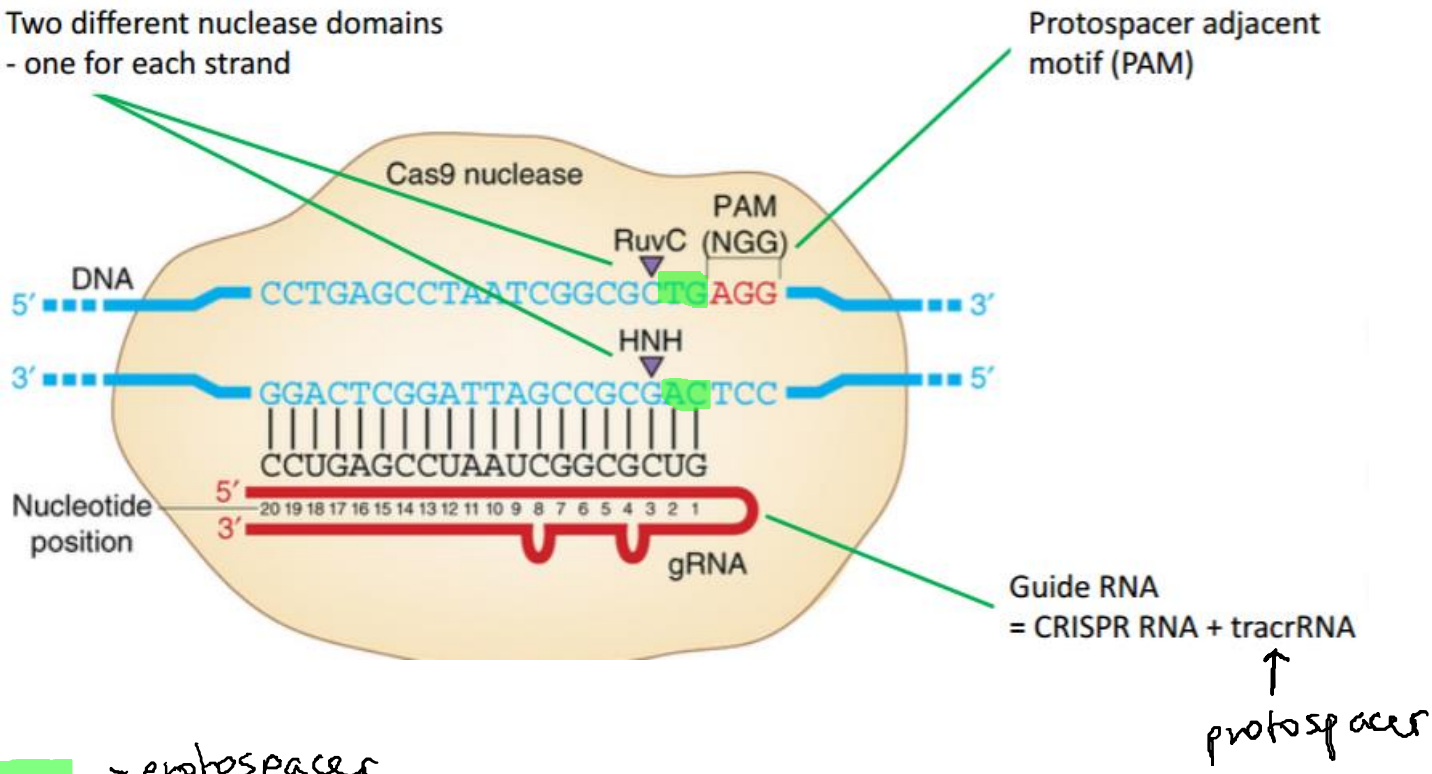
FIVE

Clinical trial controversy: Strong campaign by families of anecdotal improvements accelerated approval by FDA. It was found that the drug did not improve overall health any more than the placebo.

Lecture 1.4

CRISPR/Cas9

Complex:



- protospacer

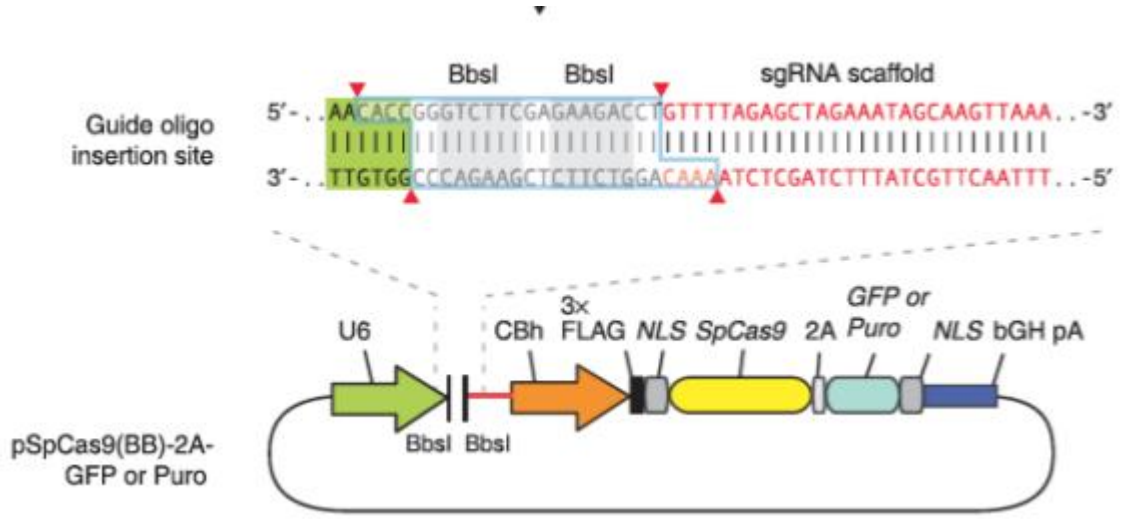
Cleavage:

- Occurs 2-3 bp upstream of PAM
- PAM, not included in cleavage

To target genes in a cell, you must express Cas9 and gRNA.

- The Cas9 gene must be expressed to encode the Cas9 protein
- The gRNA gene must be encoded to produce the complementary DNA recognition (gRNA) strand that targets the gene and acts with Cas9 to cleave the DNA.

In order to enable expression of these two genes, a plasmid is made:



Cells are transfected with plasmid with gRNA sequence, Cas9 sequence and sometimes a GFP label so the successfully transfected cells can be identified. gRNA sequence is annealed (forms dsDNA) via binding of sticky ends to a promotor and a sgRNA scaffold.

- Cells are screened and gene of interest is sequenced to confirm abnormality/indels
- Indel: insertion or deletion mutation

<u>Uses</u>:

- Can insert missing genes (gene therapy)
- Can insert tags to create fusion proteins
- Can insert fluorescent markers
- Can create a condition allele (only activated/deactivated if in contact with another compound. This enables researchers to observe the effects of a lethal gene in later life)
- Can knockout genes (allow for random HEJ/NHEJ) and allow researchers to observe the effects
- Can currently purchase knockout mice using this method for $5,000 (very cheap in comparison to when it first was being used)

West Nile Virus

Causes massive neuronal cell death. It wasn't understood how downstream death was occurring and which cells were keeping WNV alive.

1. 20,000 different genes involved in neuronal death and survival were collected from an sgRA (single guide) library.
2. They were then all placed in separate plasmids and packed into a lentivirus
3. The lentivirus transfected a human culture

4. The WNV infected the culture and the cells that survived indicated which genes were involved in WNV survival
5. Once the genes are identified, treatment of WNV can be made by targeting these pathways facilitated by these genes.

Somatic Cell Editing:

- Done ex vivo in a human culture
- Process:
    1. Cells are removed from patient and cultured
    2. These cells are then genetically altered via nucleases
    3. GM cells are put back into the patient
        a. Systematic reintroduction (blood/muscles)
            i. Harder to go into muscles because it affects smooth and cardiac muscle too
        b. Targeted reintroduction (specific organ)

General Limitations	Ex vivo limitations	In vivo limitations
Non dividing cells do not employ homologous directed repair.	Isolation is difficult	Potential problems when trying to target specific organs
Sometimes the amount of successful cells may not be enough to provide therapeutic benefits	Cell maintenance for specific cells is difficult	Life span of cells might not be long enough to provide therapeutic benefits
New cells might have decreased fitness	Potential for cells to mutate in culture	

These limitations impact the efficacy of therapeutic cell use.

Possible treatable diseases examples:

- Haemophilia B (lack of clotting factors) (ZFN)- HDR
- HIV (ZFN/CRISPR)- NHEJ
    - CCR5 (chemokine receptor→ T-cell activation) is a lymphocyte antigen that binds to HIV and allows entry of HIV into the cell.
    - An African population missing 32 AA in the CCR5 gene are resistant to HIV
    - Hence, removing the 32AA = potential HIV immunity.
        - Found to be very successful using ZFN activity with 25% efficiency (as that is what was available back then)
        - Success suspected to be attributed to inhibition of HIV to entre cell, enabling modified, functional T-cell to proliferate.
- Duchenne muscular dystrophy (CRISPR/TALEN)- NHEJ
- Hep B (TALEN/CRISPR)- NHEJ
- SCID (ZFN)- HDR
- Cataracts (CRISPR)- HDR
- Cystic fibrosis (CRISPR)- HDR
- B-cell acute lymphoblastic leukaemia:
    - T-cell engineered to express chimeric antigen receptor against B-cell CD19 antigen.
    - Done via TALENs
    - UCART cells (universal cells taken from another human) persisted beyond recovery. Their self-markers were modified such that it did not alter the immune system while they acted to kill overactive B-cells

THE FUTURE:

- Approval is slow because many factors must be overcome
- Immunogenic (what Tony Tiganis is working on)
- Issues:
    - Patent wars
    - May lead to designer babies
    - Incorrect reporting by the media

Notes:

- Non-viable zygotes: zygotes that are trinuclear

<u>Mitochondrial Disease</u>

Disease mostly caused by defects in OXPHOS pathway

Results from mutations in more than 120 nDNA and 35/37 mtDNA

Heteroplasmy: cell contains mixed populations of normal and mutant DNA

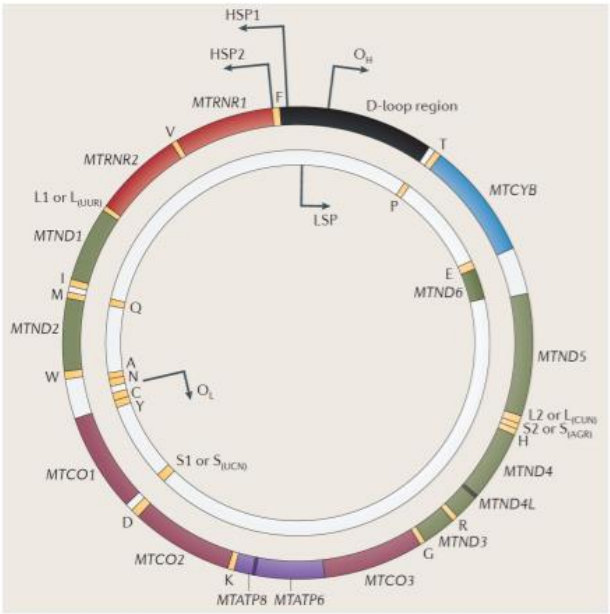
Mitochondria are implicated in aging and neurodegenerative disorders:

- Parkinson's
- Dementia
- Cancer

Mitochondria contain ~1,200 but mtDNA only encodes 13 proteins. Most proteins come from nDNA and are imported into the mitochondria.

As mtDNA is only inherited from the mum, inheritance modes (Mendalian/non-mendalian) play a role in acquisition of disease/intensity.

mtDNA



Features

- 37 genes
- Maternally-inherited
- 100s-100s copies/cell
- mtDNA repair mechanisms less efficient that nDNA
- Cells can contain mixed populations of normal and mutant mtDNA = heteroplasmy

Inherited mtDNA disorders

- Birth to old age

Somatic mutations

- Accumulate with age