Advances in Modern Genetics

Dr Katja Vogt

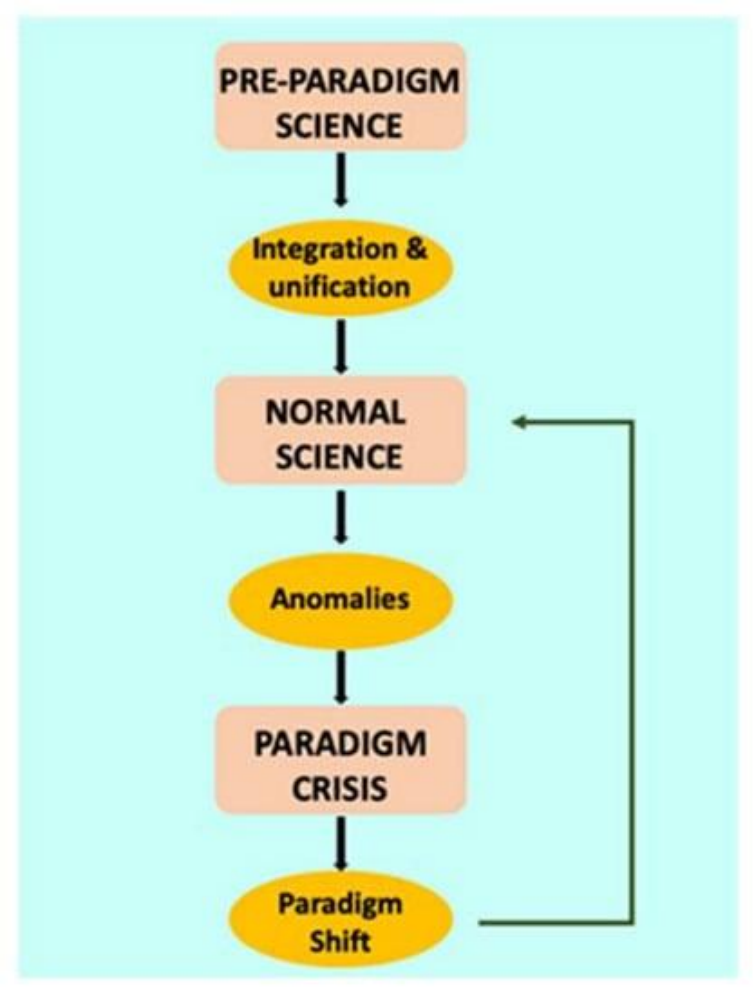
Slides based on material from Dr Emyr Bakker

Today we are going to…

- … discover and understand the concept of a paradigm shift and how this relates to advances in medicine
- … engage with areas of increasing complexity within genetics, including splicing and noncoding RNAs
- … explain the Human Genome Project
- … identify laboratory methods
- … explore advances in genetic technologies that can impact on medical research and, ultimately, clinical practice

PART 1: PARADIGM SHIFTS AND EXAMPLES OF GENETIC COMPLEXITY

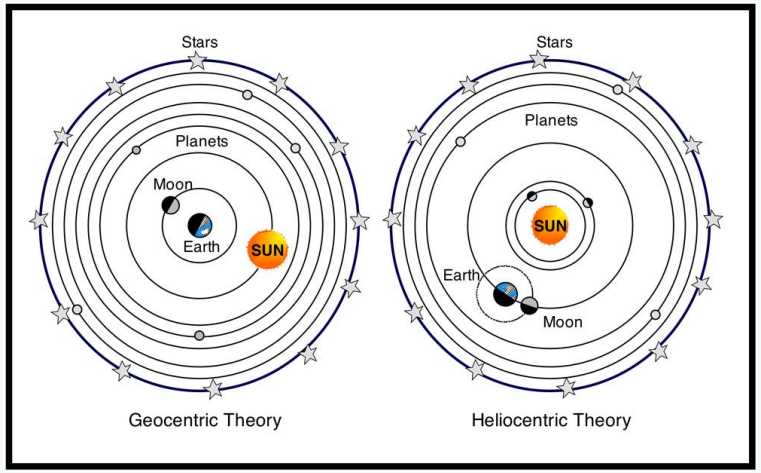
Key Terminology: Paradigm shifts



- Taken from González-Márquez et al (https://www.mdpi.com/2071-1050/12/7/2802)

- Paradigm shifts can be summarised as anomalies being observed in 'routine' science, leading to a paradigm crisis and that resolution of these anomalies leads to a shift in the underlying assumptions of a field.

- Examples: geocentricism vs heliocentrism



Central dogma of molecular biology

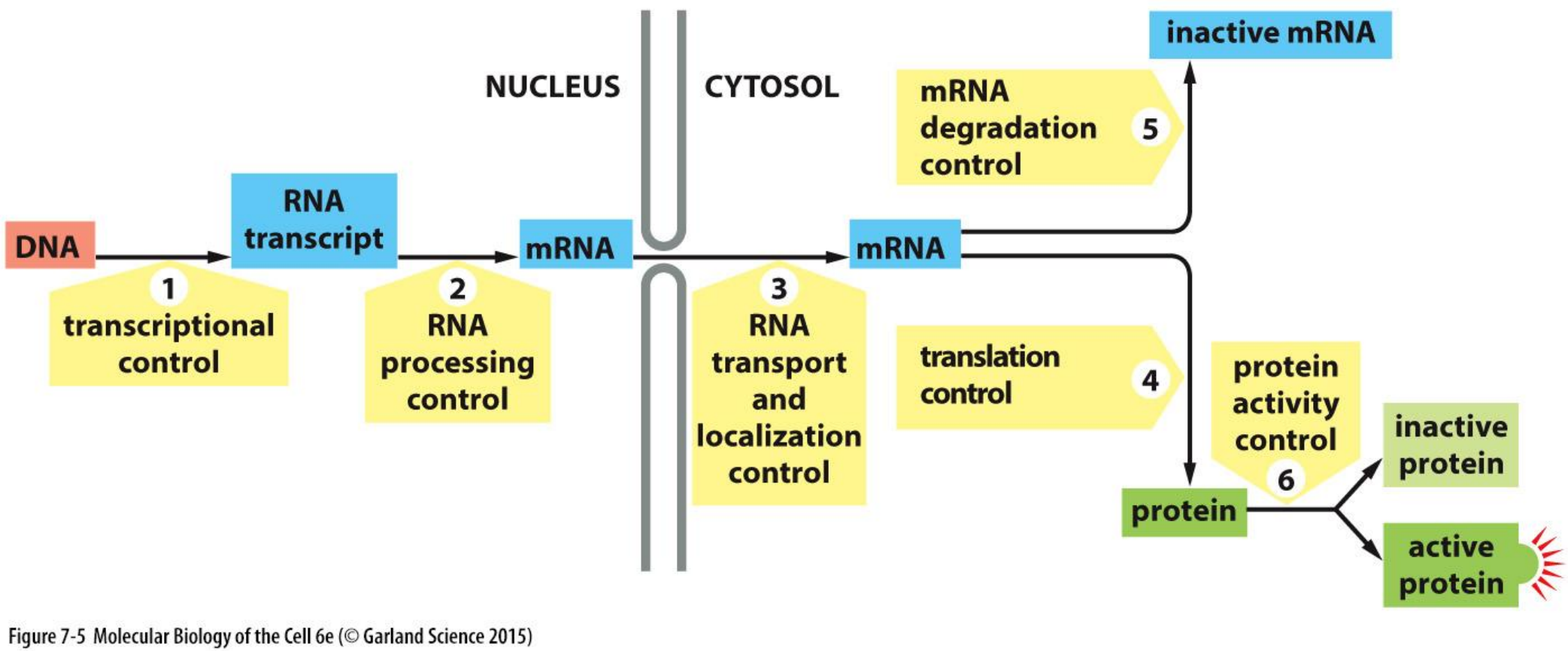


Figure 7-5 Molecular Biology of the Cell 6e (© Garland Science 2015)

Example of genetic complexity RNA splicing

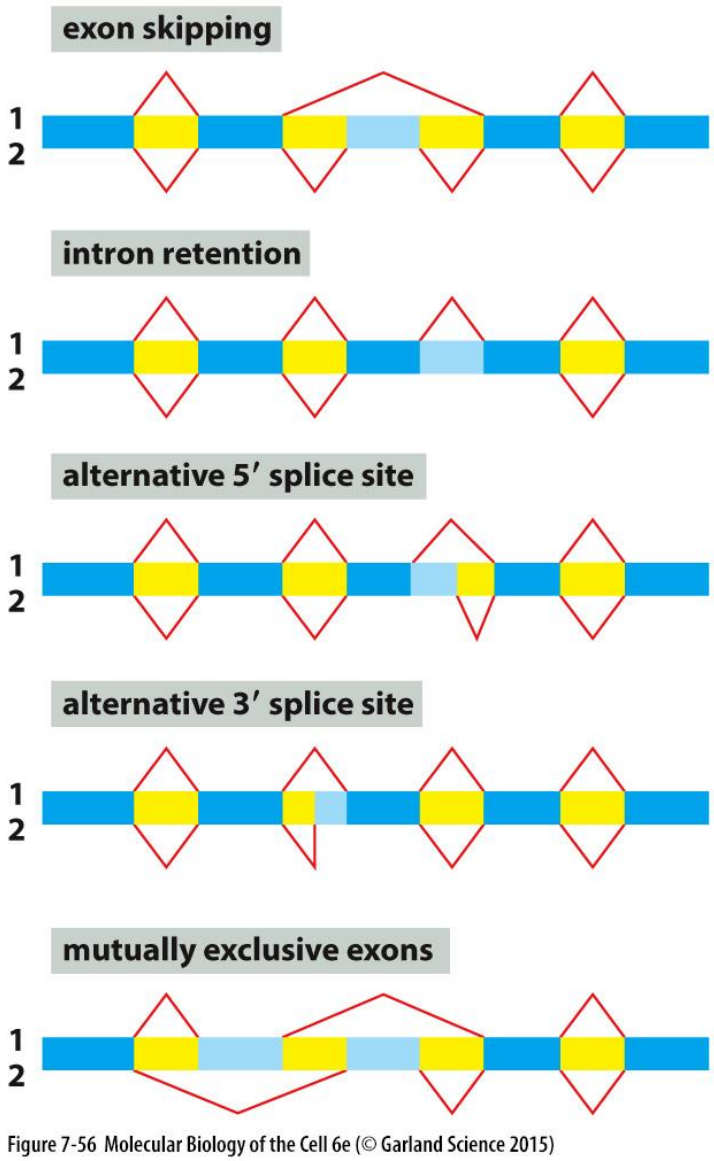


Figure 7-56 Molecular Biology of the Cell 6e (© Garland Science 2015)

- Splicing is a process that removes introns from the mRNA precursor, leaving exons
- Cells can splice RNA transcripts differently and thereby make multiple polypeptides from the same gene—this is **alternative RNA splicing**
- In the figure on the left, five patterns of alternative RNA splicing are shown
  - In each case, 1 and 2 represent different splicing patterns (follow the red line)
  - Dark blue represents exons retained in both splice variants
  - Light blue represents exons retained only in one splice variant
  - Yellow boxes represent introns

RNA splicing - example

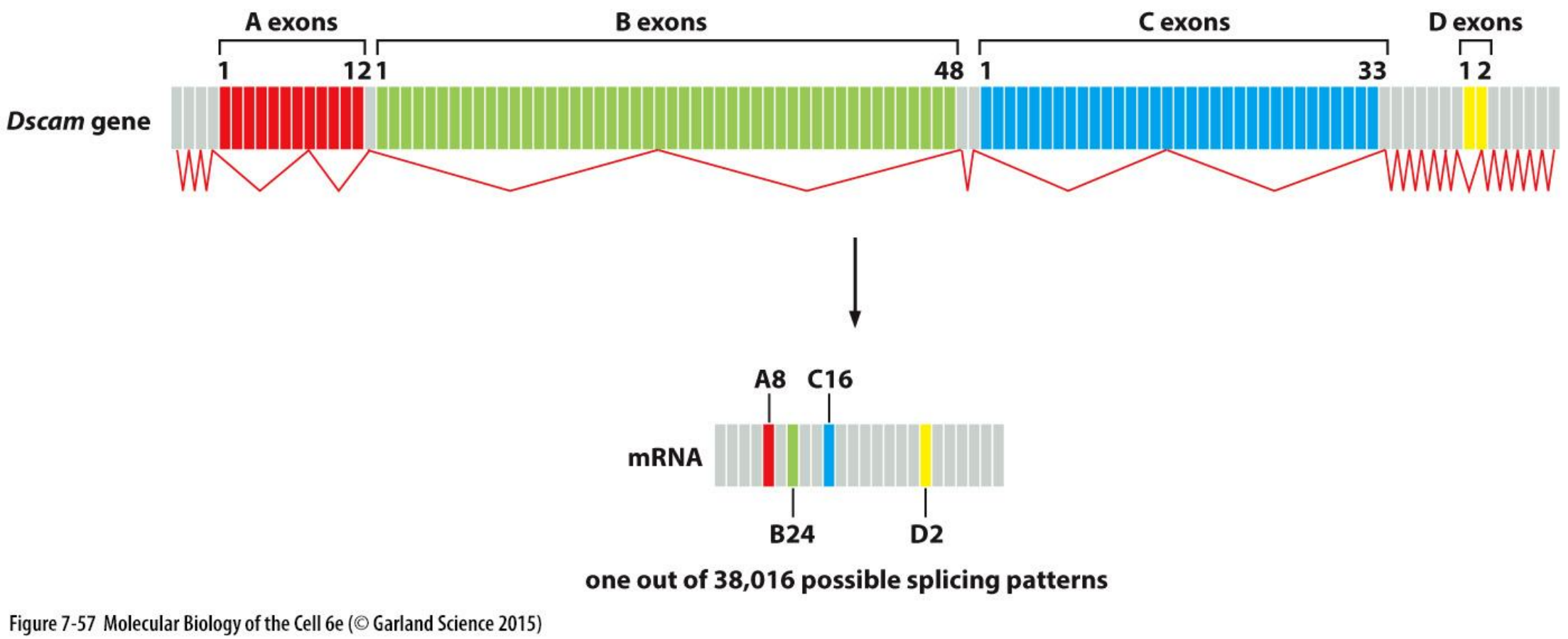
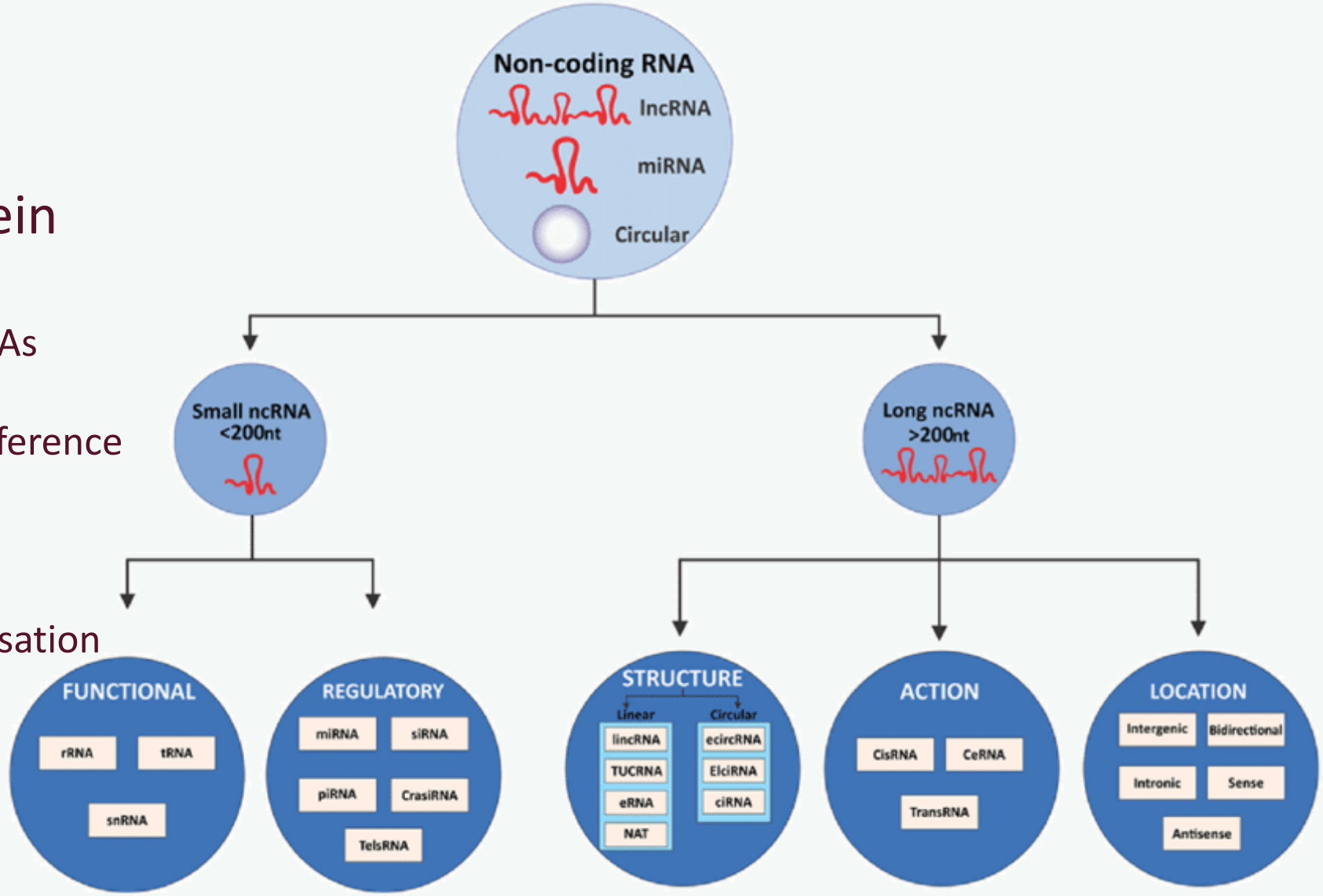


Figure 7-57 Molecular Biology of the Cell 6e (© Garland Science 2015)

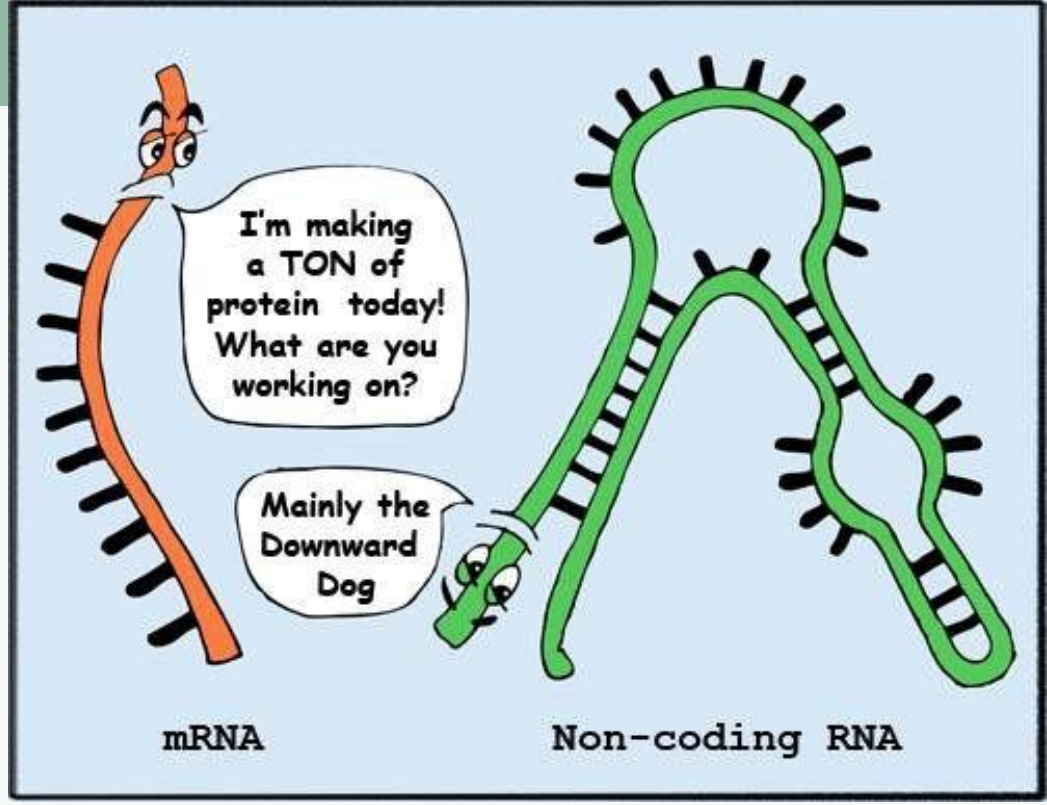
Non – coding RNAs

- Do not code for a protein
- Multiple types
  - Short RNAs such as microRNAs (miRNAs)
    - These perform RNA interference (RNAi)
  - Long noncoding RNAs
    - Variety of functions e.g. scaffolding, protein localisation



Small non-coding RNAs

- There are over 1000 different microRNAs produced from the human genome
- MicroRNAs appear to regulate at least one-third of all human protein-coding genes



- siRNA has had use in research, for example to silence genes during experiment
  - This aids in understanding the role of genes, though as highlighted through gene ontology & functional annotation analysis, these can be very complex!
- There are cases to use RNAi in clinical practice. As many diseases arise from aberrant gene expression, the potential to “turn them on or off” is very powerful
  - Caution must **always** be exercised however—otherwise we can have tragic outcomes

Interference RNA

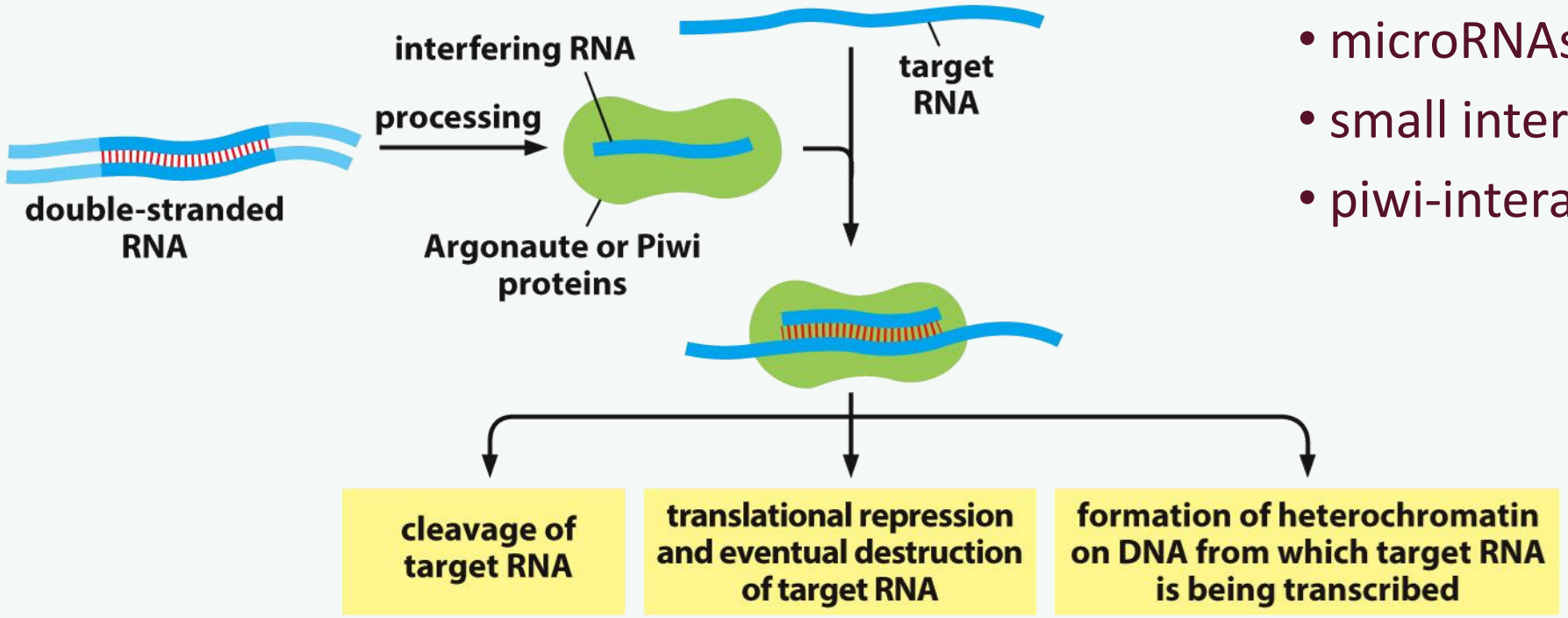


Figure 7-74 Molecular Biology of the Cell 6e (© Garland Science 2015)

- microRNAs (miRNAs)
- small interfering RNAs (siRNAs)
- piwi-interacting RNAs (piRNAs)

Long non-coding RNAs (lncRNA) ROLES OF lncRNAs

- Scaffolds: bringing together proteins that recognise parts of the lncRNA
- Localise proteins to specific RNA or DNA sequences

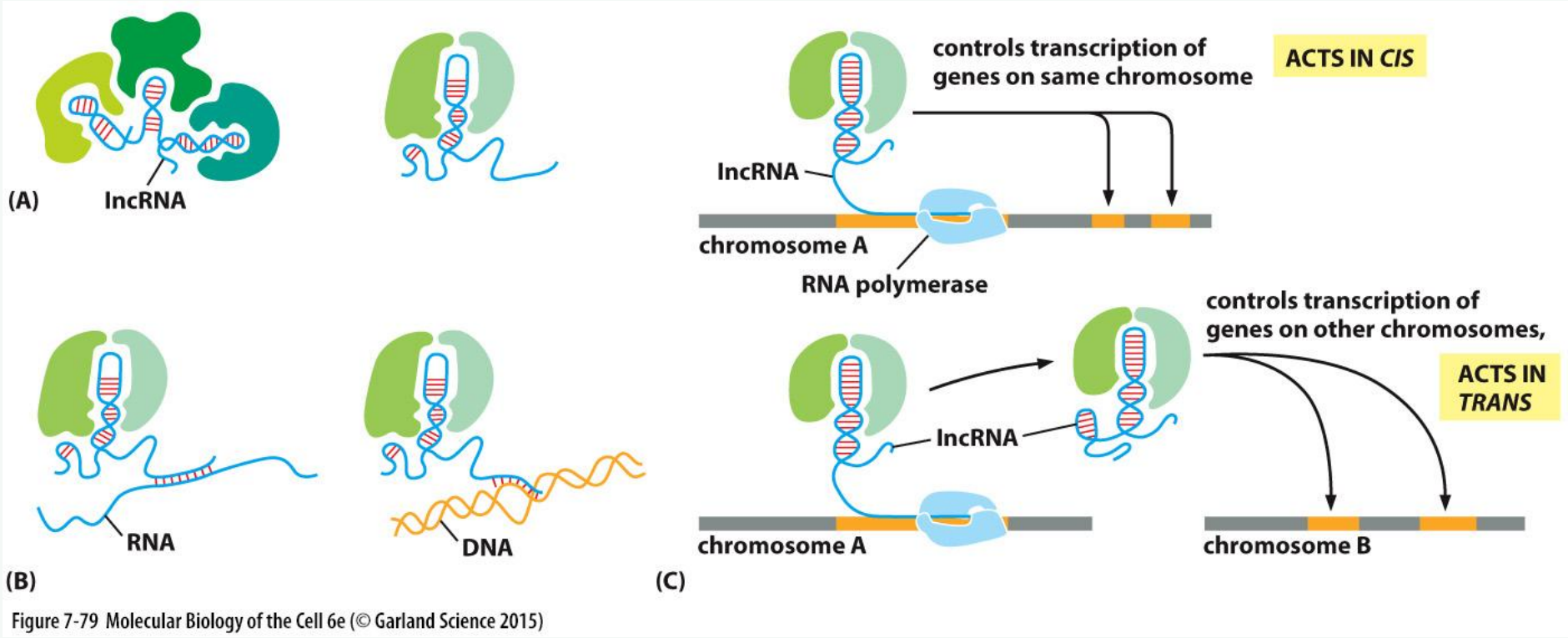
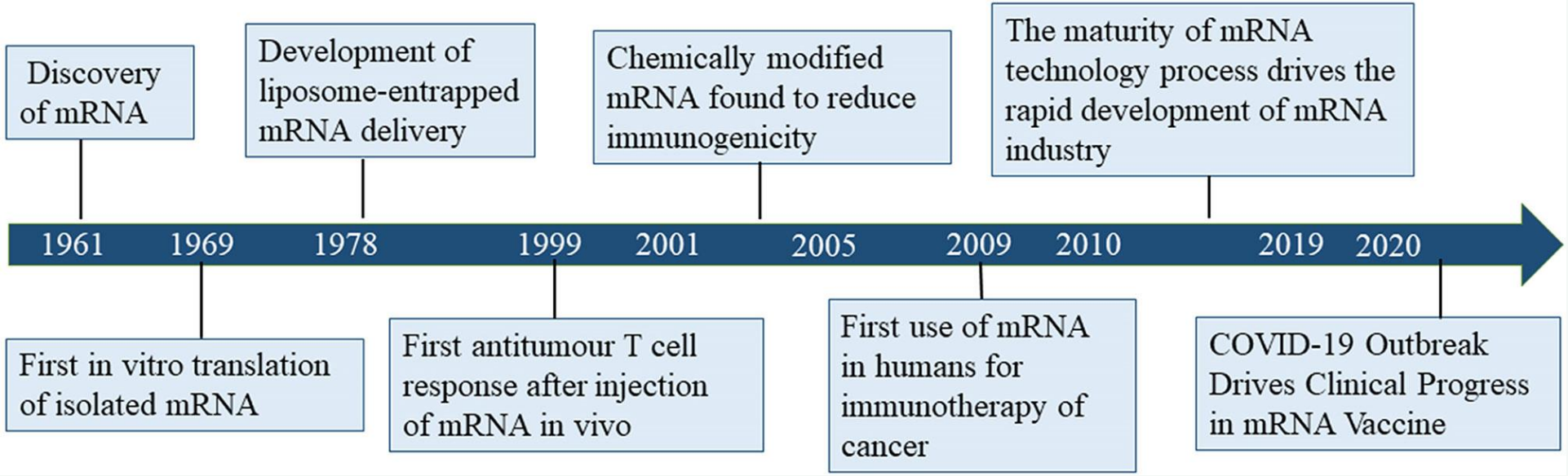


Figure 7-79 Molecular Biology of the Cell 6e (© Garland Science 2015)

Some lncRNAs act in **cis** manner (stay tethered to RNA polymerase and control genes on the same chromosome) whilst others act in a **trans** manner (diffuse from synthesis site and control transcription of other sites)

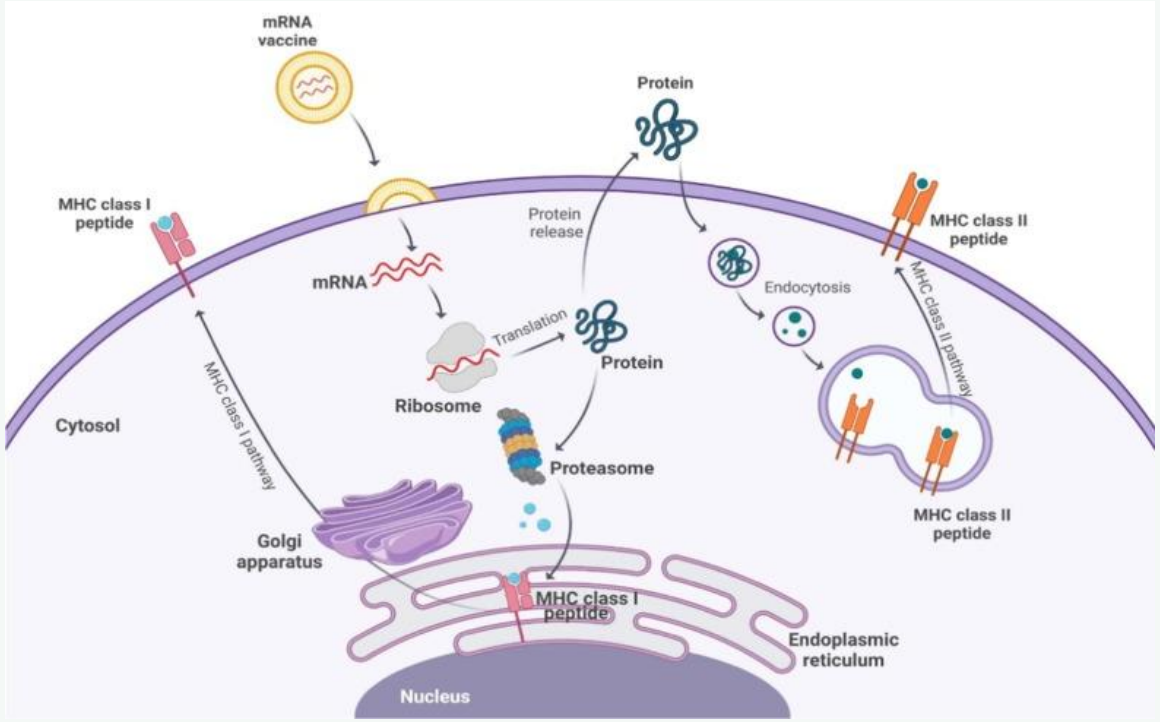
History of mRNA vaccines



- Historically RNA has been tricky to study due to it being much less stable than DNA
- mRNA vaccines were first trialled in a cancer context through TSAs and TAAs in 2009
  - **Of TSAs and TAAs, which do you think is better?**
- There was a long period of development for mRNA vaccines before their use in COVID

Figure 1, Deng et al 2022
(https://www.frontiersin.org/articles/10.3389/fimmu.2022.887125/full)

mRNA VACCINES—MECHANISM OF ACTION



- mRNA encoding a viral protein enters the cell where it is translated into protein by the ribosome.
- The resulting protein is broken down into peptides by the proteasome or transported by the Golgi apparatus to the outside of the cell.
- The remaining fragments in the cell are presented as a complex.
- Additionally, protein outside of the cell can be taken up by various immune cells and fragmented into smaller pieces by the endosome.

Figure 3, Jain et al 2021 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8502079/)

PART 2: THE HUMAN GENOME PROJECT AND LABORATORY METHODS

The human genome

- The human haploid genome is 3,234.83 Mb (Megabases) in size
- Thus, part of the challenge with our genome is one of sheer scale

- But approximately there are only 20,000 protein-coding genes
- This 20,000 is an estimate – originally we estimated 100,000 protein-coding genes!

- Understanding how these bases are linked and how they function helps us to understand ourselves
- But the fun part is something else…

The human geneome project

An international government project

Had multiple aims, including:

- Identifying all genes present
- Work out the base-pair order of DNA
- Develop faster DNA sequencing technologies

A thirteen-year effort, led by the Department of Energy and National Institutes of Health in the US

For once, a project finished ahead of schedule – two years in this case!

Benefits of the human genome project

- Disease Understanding and Diagnosis
- Personalised Medicine
- Biomedical Research
- Prenatal and Reproductive health
- Global scientific collaboration
- Evolutionary and population studies

Genomic medicine

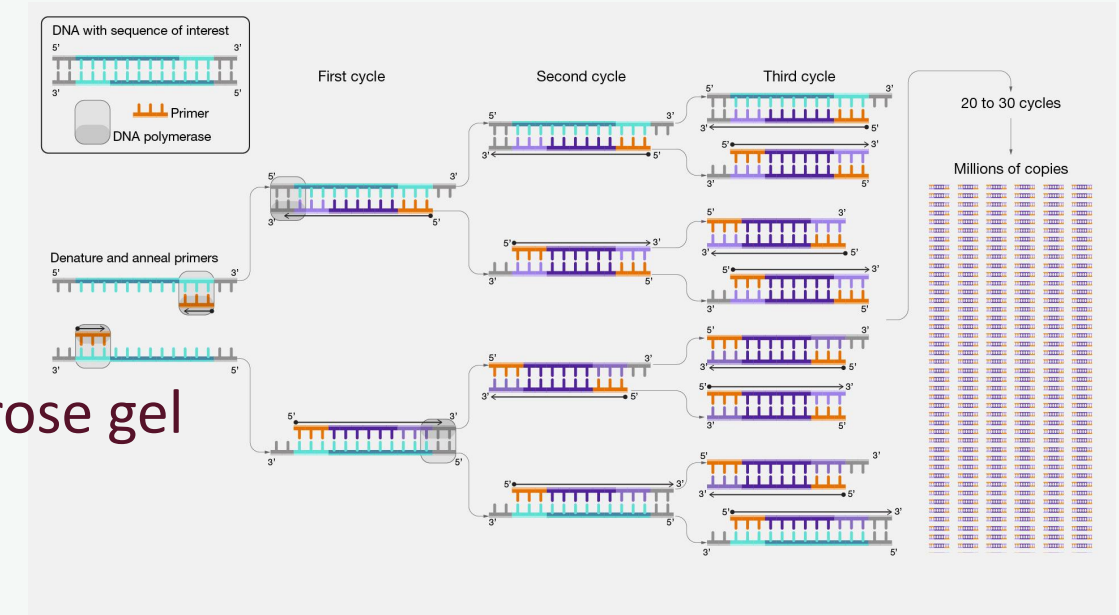
An emerging medical discipline which aims to use patient genetic data as part of their clinical care

- Influences Diagnostics and Therapeutics
- Utilises the concept of patient stratification
- Can relate to disorder, partly caused by genetics, such as cancer, diseases with incomplete penetrance and expressivity

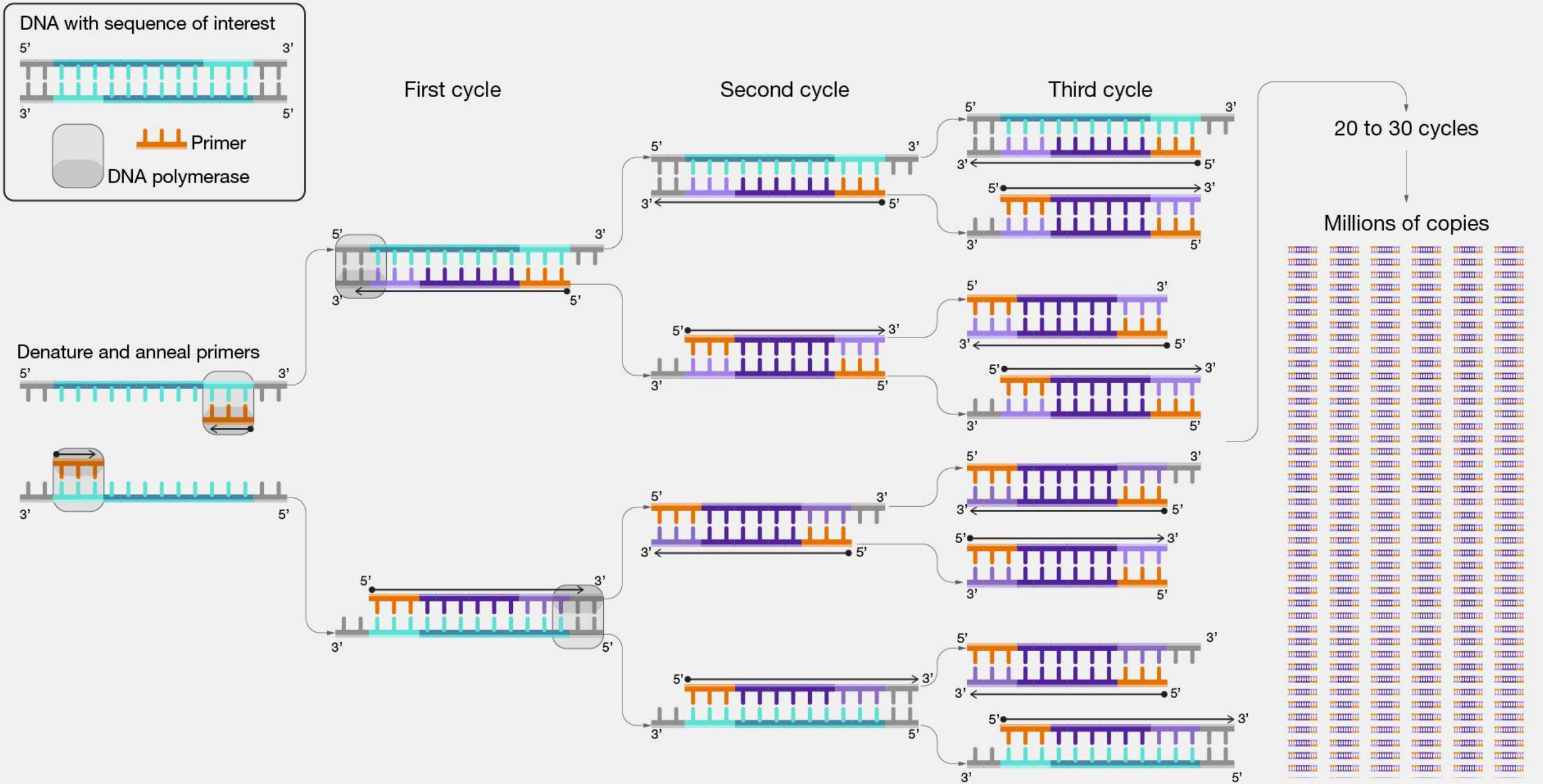
DNA DETECTION METHODS

DNA detection - PCR

- Polymerase chain reaction
- Amplifies (smallish) pieces of DNA
- Can be detected as the product → Agarose gel
- Can be detected live → qPCR



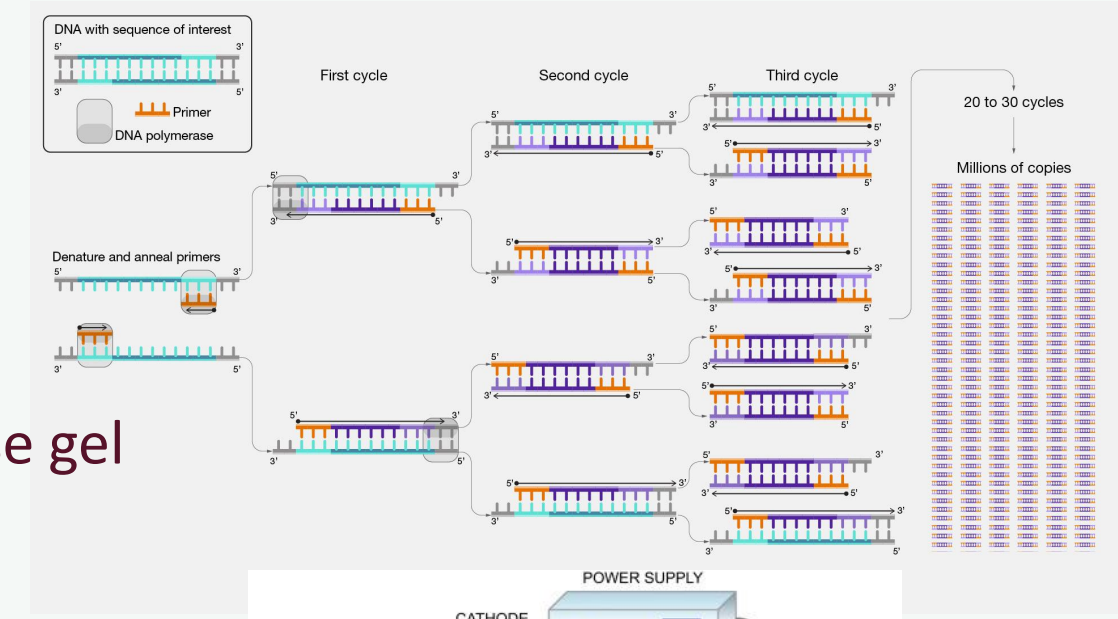
- Use:
  - Identification of known mutations
  - Detection of pathogens (especially qPCR)

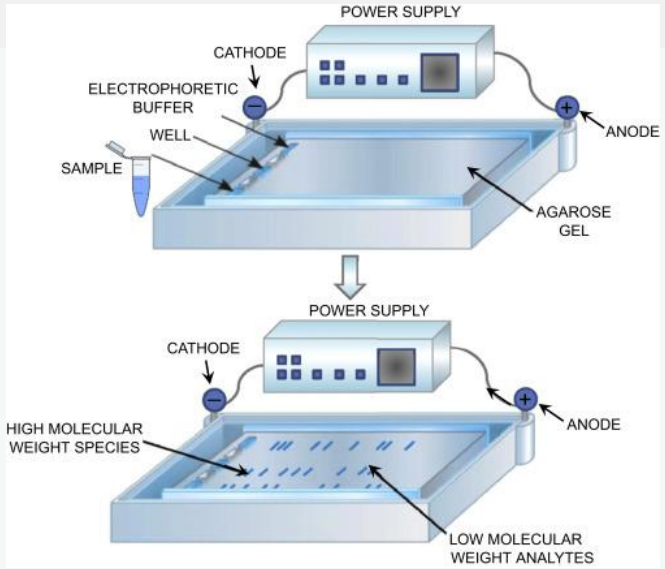

DNA with sequence of interest
5'
3'
3'
5'
Primer
DNA polymerase
First cycle
Second cycle
Third cycle
20 to 30 cycles
Millions of copies
Denature and anneal primers

DNA detection - PCR

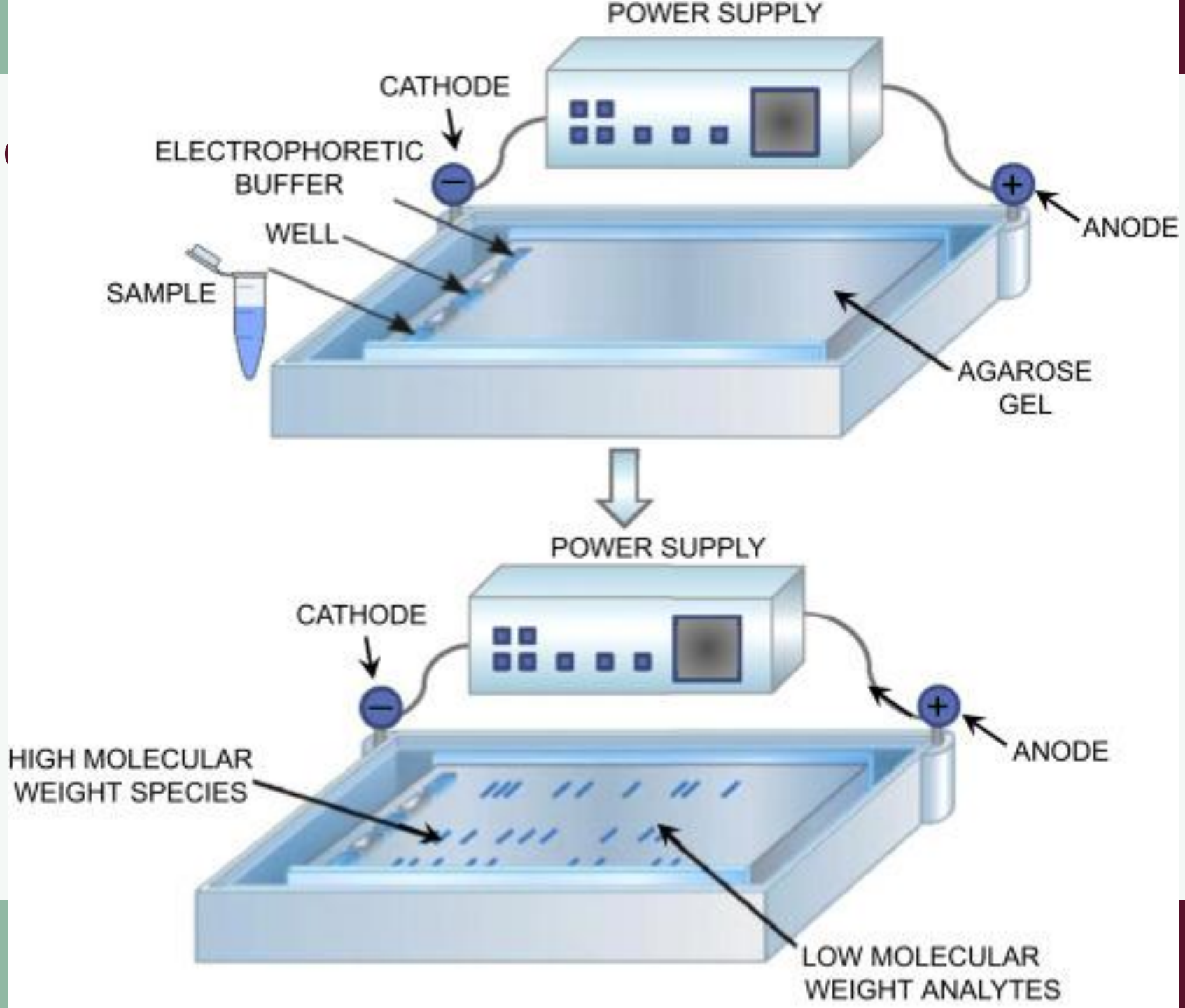
- Polymerase chain reaction
- Amplifies (smallish) pieces of DNA
- Can be detected as the product → Agarose gel
- Can be detected live → qPCR

- Use:
  - Identification of known mutations
  - Detection of pathogens (especially qPCR)

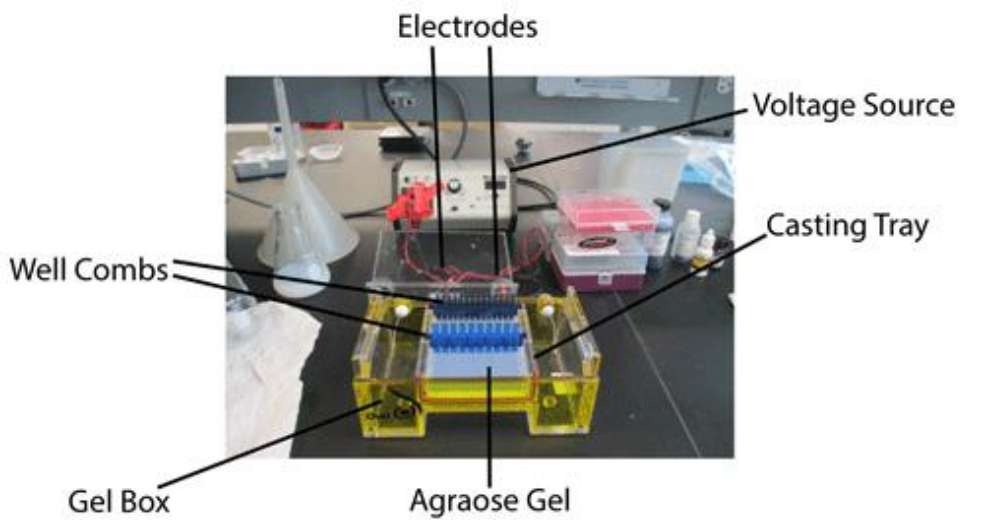




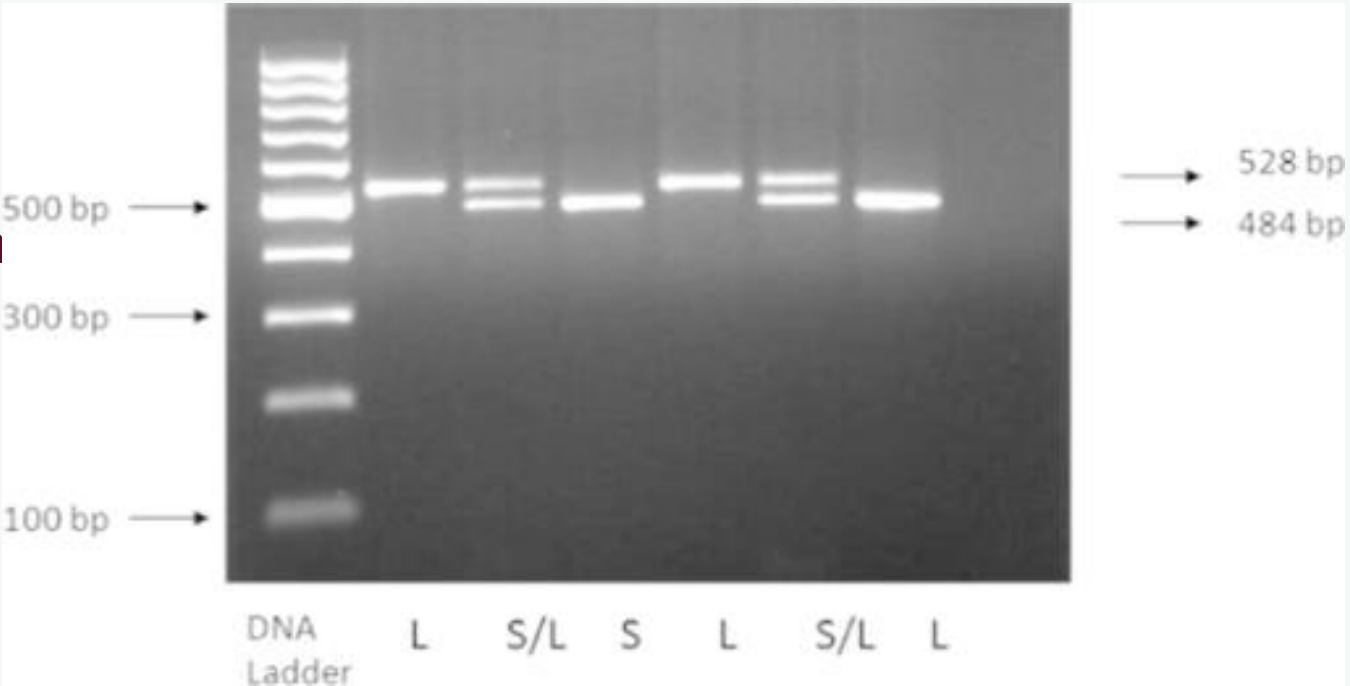
Agaros


POWER SUPPLY
CATHODE
ELECTROPHORETIC BUFFER
WELL
SAMPLE
ANODE
AGAROSE GEL
POWER SUPPLY
CATHODE
ANODE
HIGH MOLECULAR WEIGHT SPECIES
LOW MOLECULAR WEIGHT ANALYTES

PCR data analysis ANALYSIS



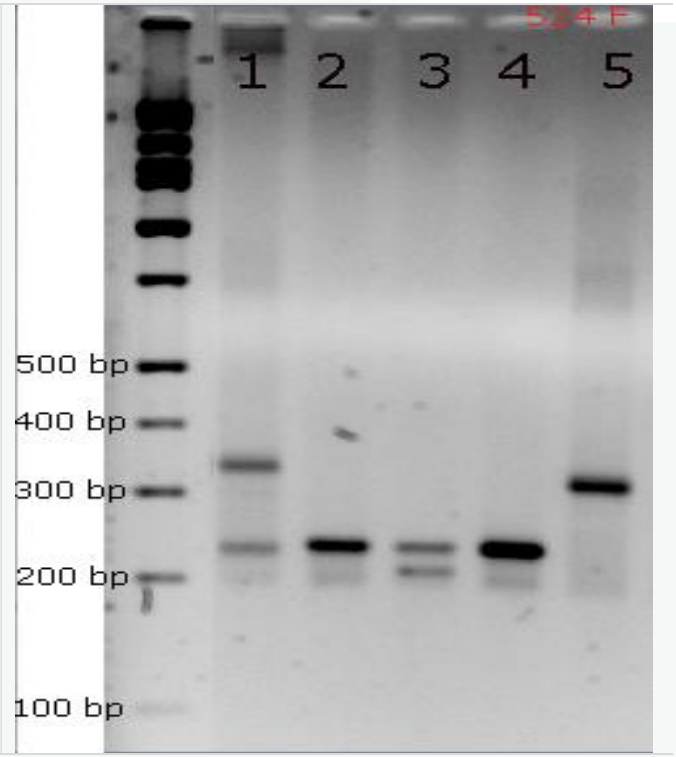
m



https://www.addgene.org/protocols/gel-electrophoresis/

Rozak, N. I. A., Ahmad, I., Gan, S. H., & Bakar, R. A. (2014). Lack of association between the serotonin transporter (5-htt) and serotonin receptor (5-ht2a) gene polymorphisms with smoking behavior among malaysian malays. Scientia pharmaceutica, 82(3), 631-642.

Application of PCR beyond pure research – example of trinucleotide repeat disorders



Use primers the region covered the expanded region for PCR amplification

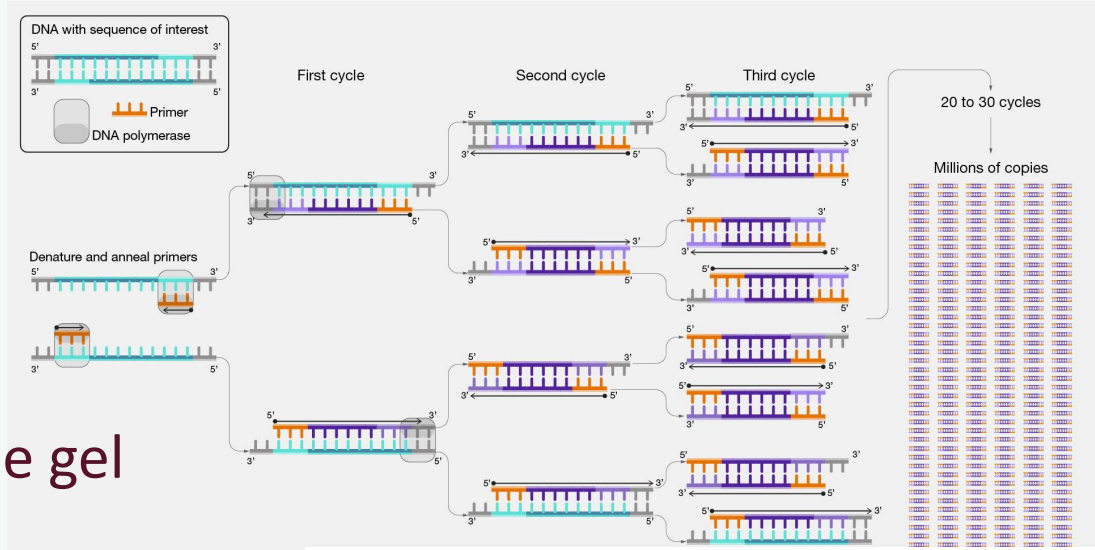
Lane 1: DNA from affected female
Lane 2: Normal gDNA (220, 240bp)
Lane 3: Normal gDNA (220, 240bp)
Lane 4: 18 CAG genemer control DNA (240bp)
Lane 5: 44 CAG genemer control DNA (318bp)

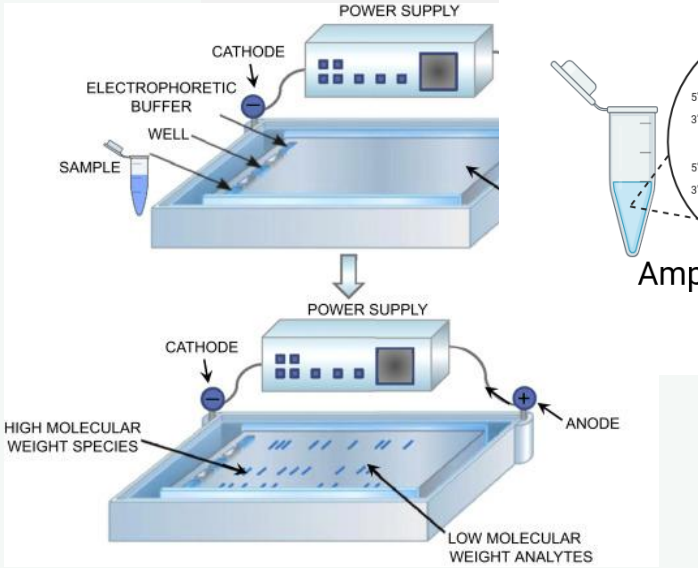
Huntington's Disease GScanTM Ver2 Kits

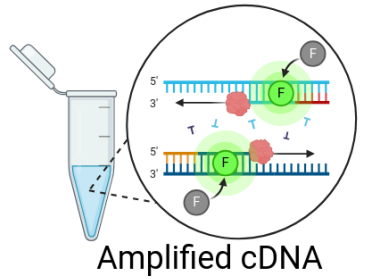
DNA detection - PCR

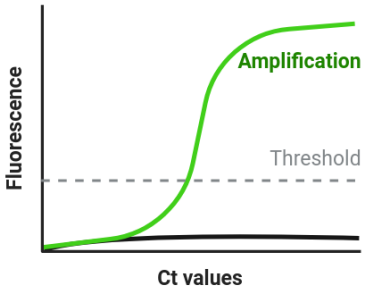
- Polymerase chain reaction
- Amplifies (smallish) pieces of DNA
- Can be detected as the product → Agarose gel
- Can be detected live → qPCR

- Use:
  - Identification of known mutations
  - Detection of pathogens (especially qPCR)

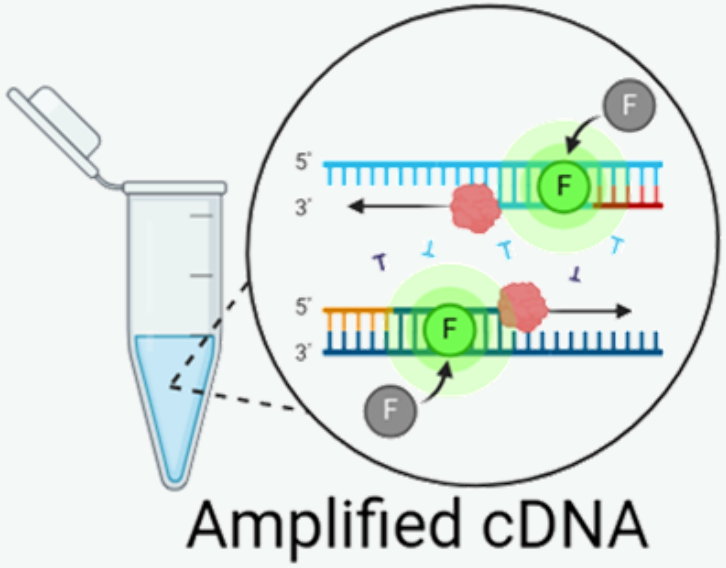




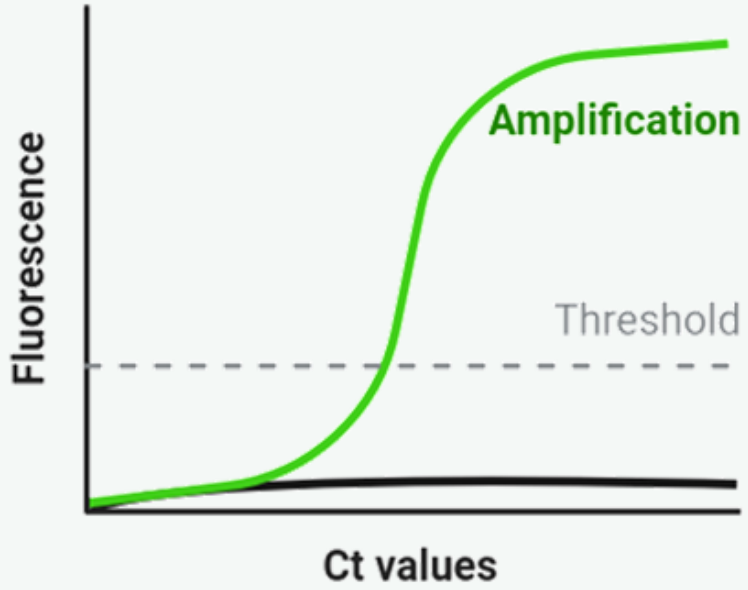




Quantitative PCR = qPCR



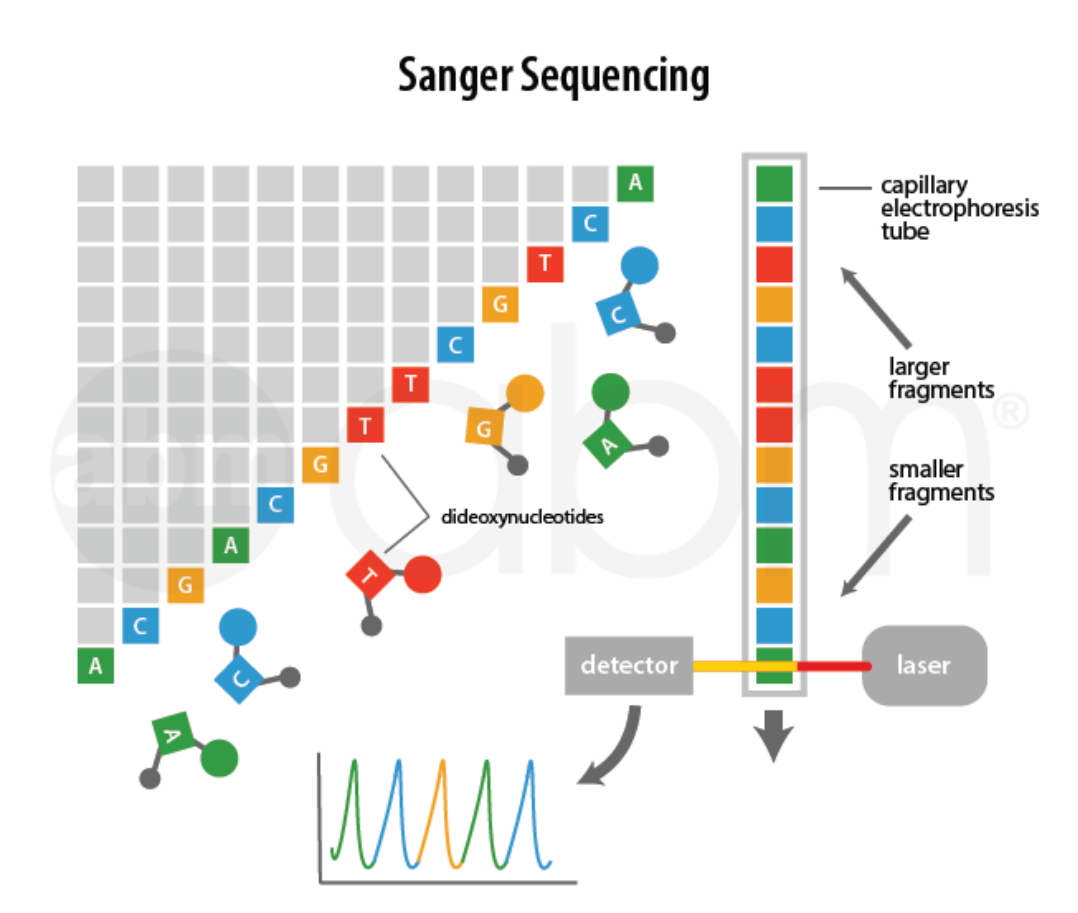
- Whilst useful – particularly for presence or absence of the DNA – this approach is limited as it does not provide quantitative information



- Scientists studying gene expression will typically extract RNA and reverse transcribe it to DNA, and then do qPCR on the DNA
  - This is known as qRT-PCR (quantitative reverse transcriptase PCR)
  - **Why would scientists use RNA instead of DNA?**

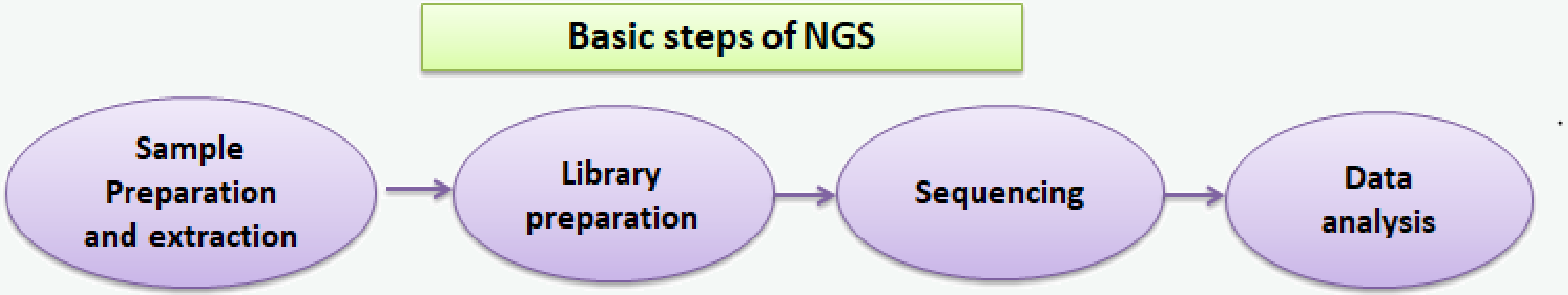
DNA Sequencing – Sanger Dideoxy sequencing

- Fluorescent dideoxy chain termination method
  - DNA template
  - Oligonucleotide primer
  - DNA polymerase (thermostable)
  - Deoxynucleotide phosphates dNTPs (dATP, dCTP, dGTP, dTTP).
  - Dideoxynucleotide triphosphates ddNTPs (ddATP, ddCTP, ddGTP, ddTTP) – fluorescentl labelled.
  - ddNTPs lack 3'-OH on deoxyribose sugar for addition of next nucleotide.



New sequencing technologies

- Sequencing technologies developed after Sanger are collectively referred to “next generation” sequencing (NGS), also called massively parallel sequencing
  - Very high throughput
  - Can sequence ‘exome’ or whole genome.
- Increasing drive for single cell methods



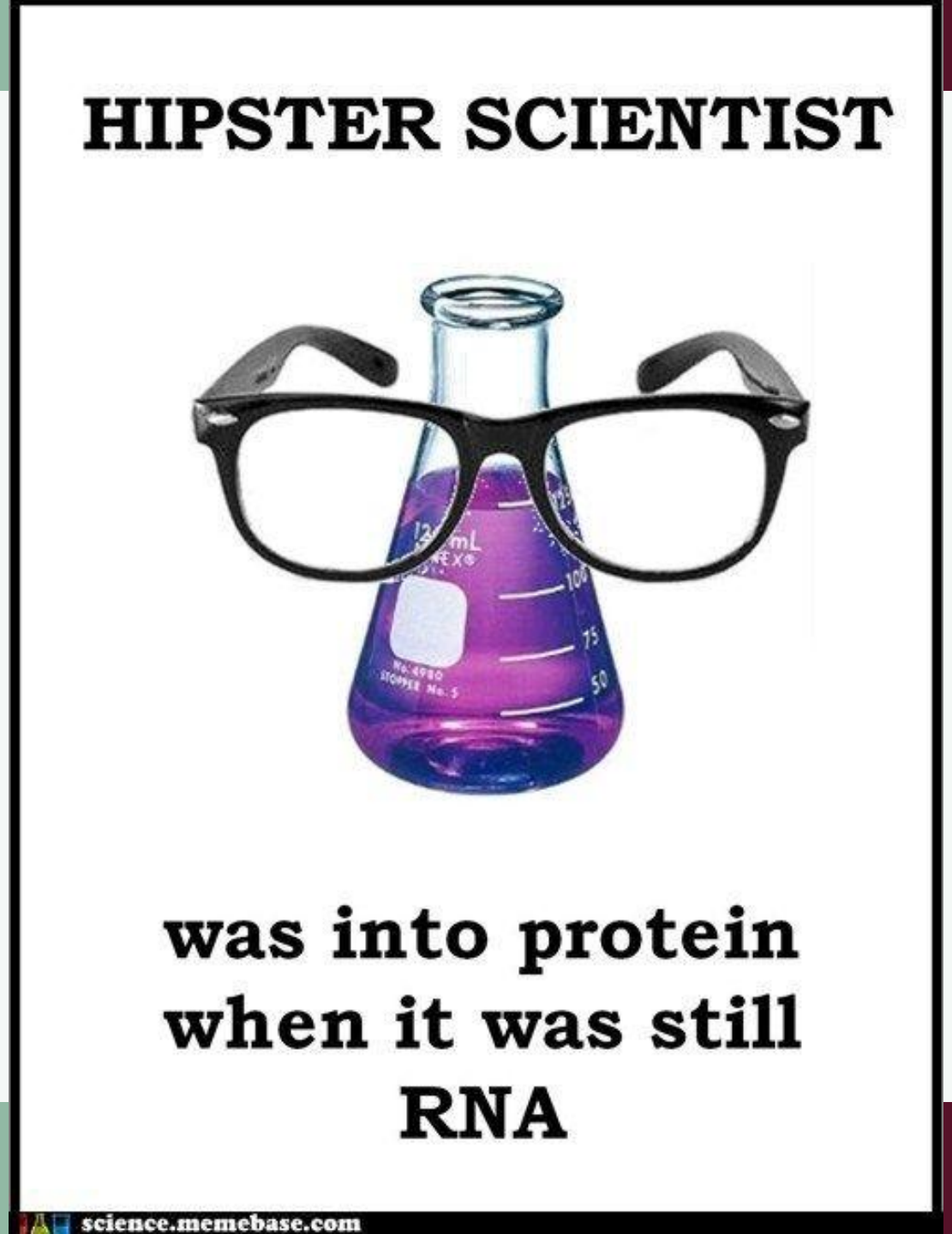
Applications of NGS

- **Oncology**: NGS is widely used to identify mutations, gene fusions, and other alterations in cancer-related genes. This helps in diagnosing specific cancer types, predicting prognosis, and selecting targeted therapies. It also enables monitoring of minimal residual disease and recurrence through liquid biopsies
- **Inherited and Rare Diseases**: Whole-exome sequencing (WES) and whole-genome sequencing (WGS) allow clinicians to detect genetic variants responsible for inherited disorders. This is particularly useful for diagnosing rare diseases where traditional methods fail
- **Infectious Diseases**: NGS can identify pathogens directly from clinical samples, including bacteria, viruses, and fungi. It helps in outbreak tracking, antimicrobial resistance profiling, and understanding pathogen evolution
- **Prenatal and Neonatal Screening**: Non-invasive prenatal testing (NIPT) uses NGS to detect chromosomal abnormalities like trisomy 21 (Down syndrome) from maternal blood. It's also used in newborn screening for metabolic and genetic disorders
- **Pharmacogenomics**: NGS helps determine how a patient's genetic makeup affects their response to drugs, guiding personalized treatment plans and reducing adverse drug reactions
- **Precision Medicine**: By integrating genomic data with clinical information, NGS supports tailored treatment strategies for individual patients, improving outcomes and reducing unnecessary interventions
- **Multi-Omics Integration**: NGS is increasingly used alongside transcriptomics, proteomics, and epigenomics to understand complex disease mechanisms and identify novel biomarkers

Gupta and Verma, 2019

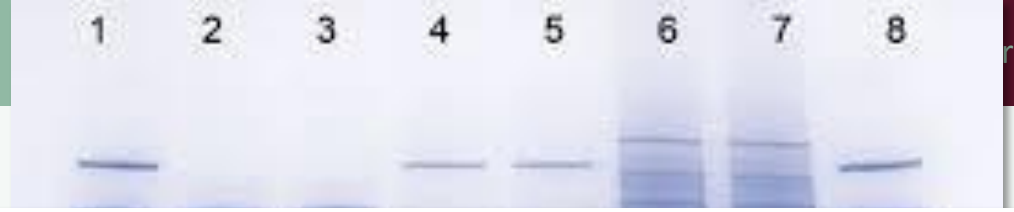
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7122948/

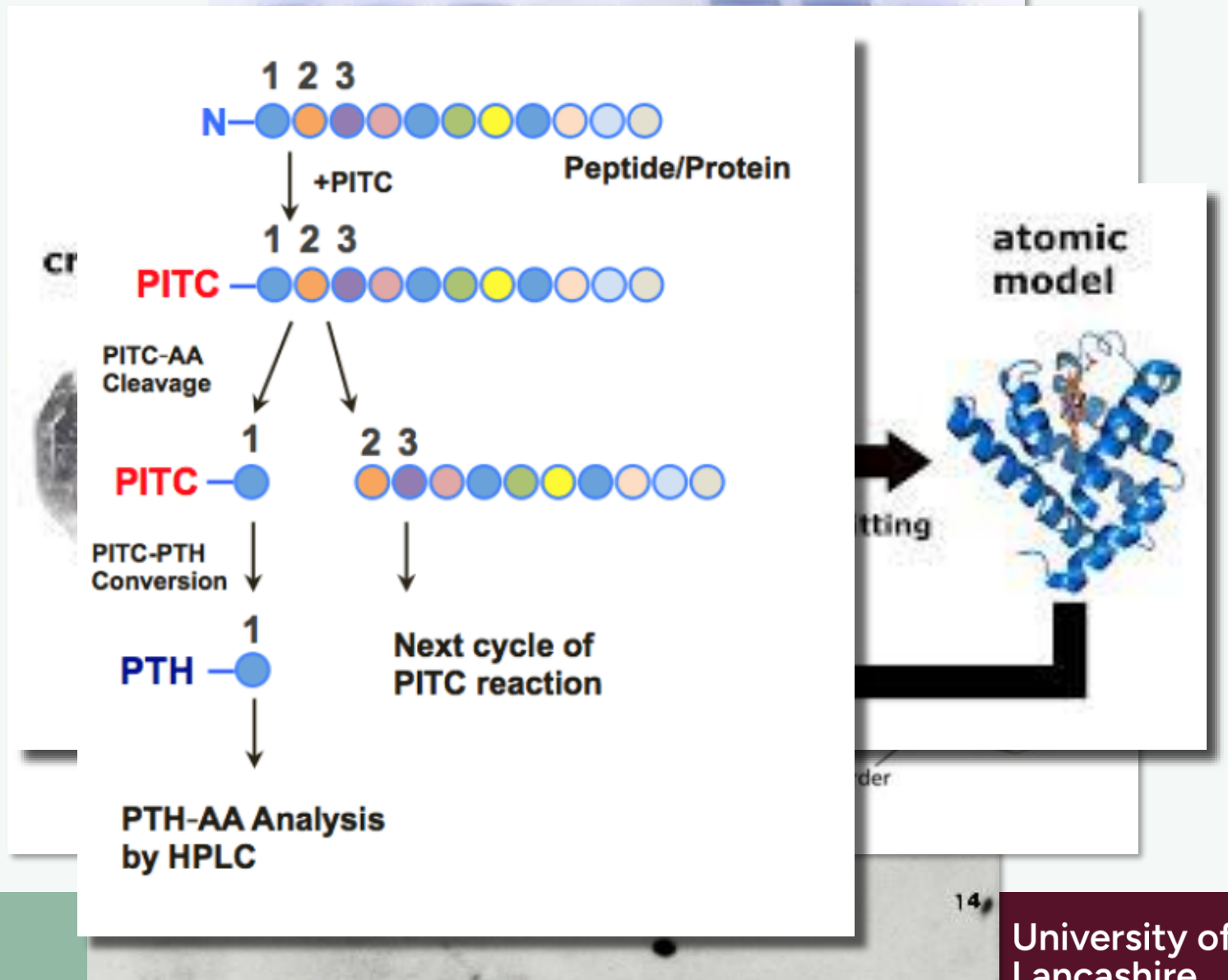
Protein detection



Protein detection

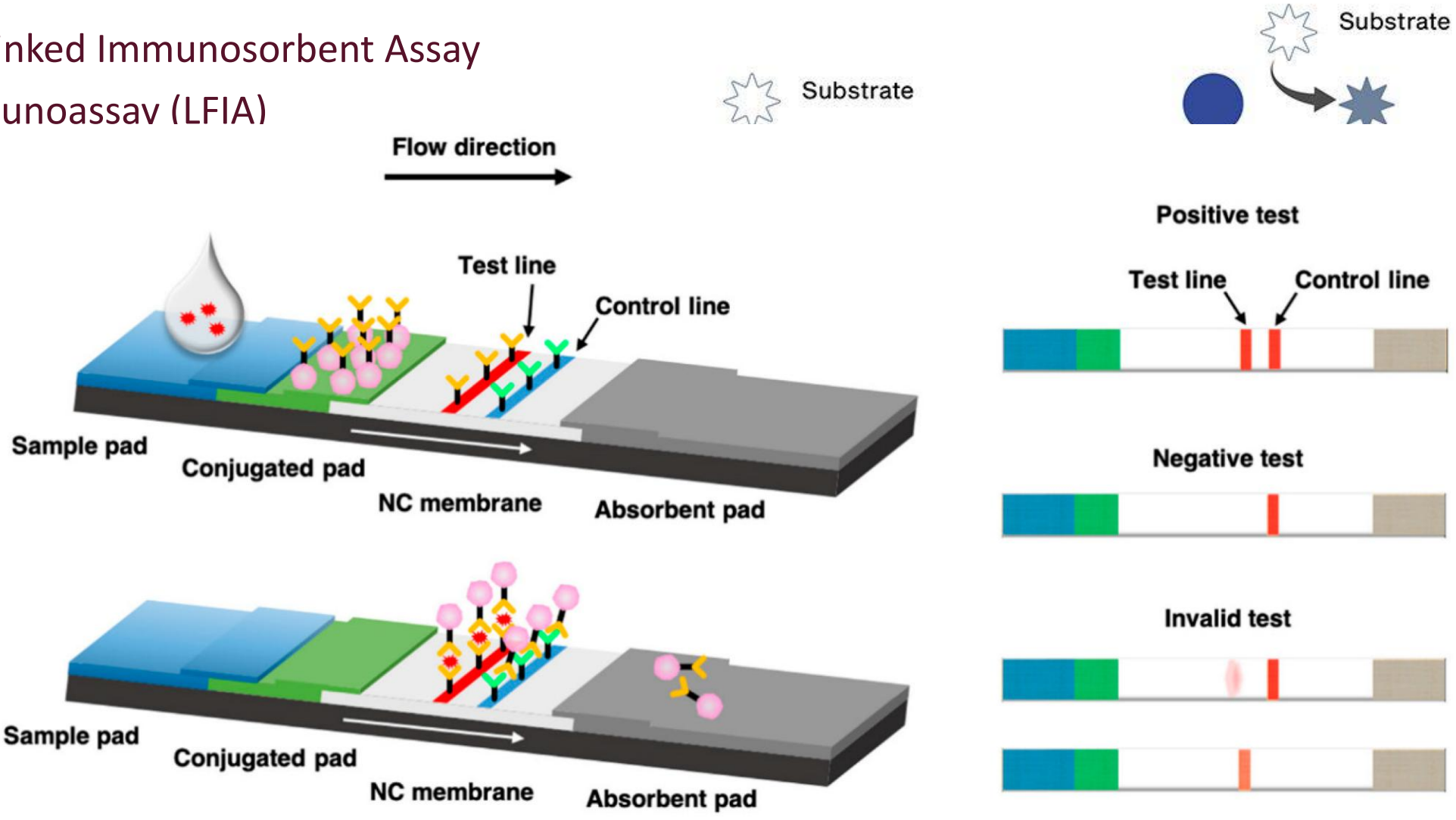
- SDS Page gels – 1D and 2D
  - Staining techniques
  - Immunoblotting
- Mass spectrometry
- Crystallography
- Edman sequencing





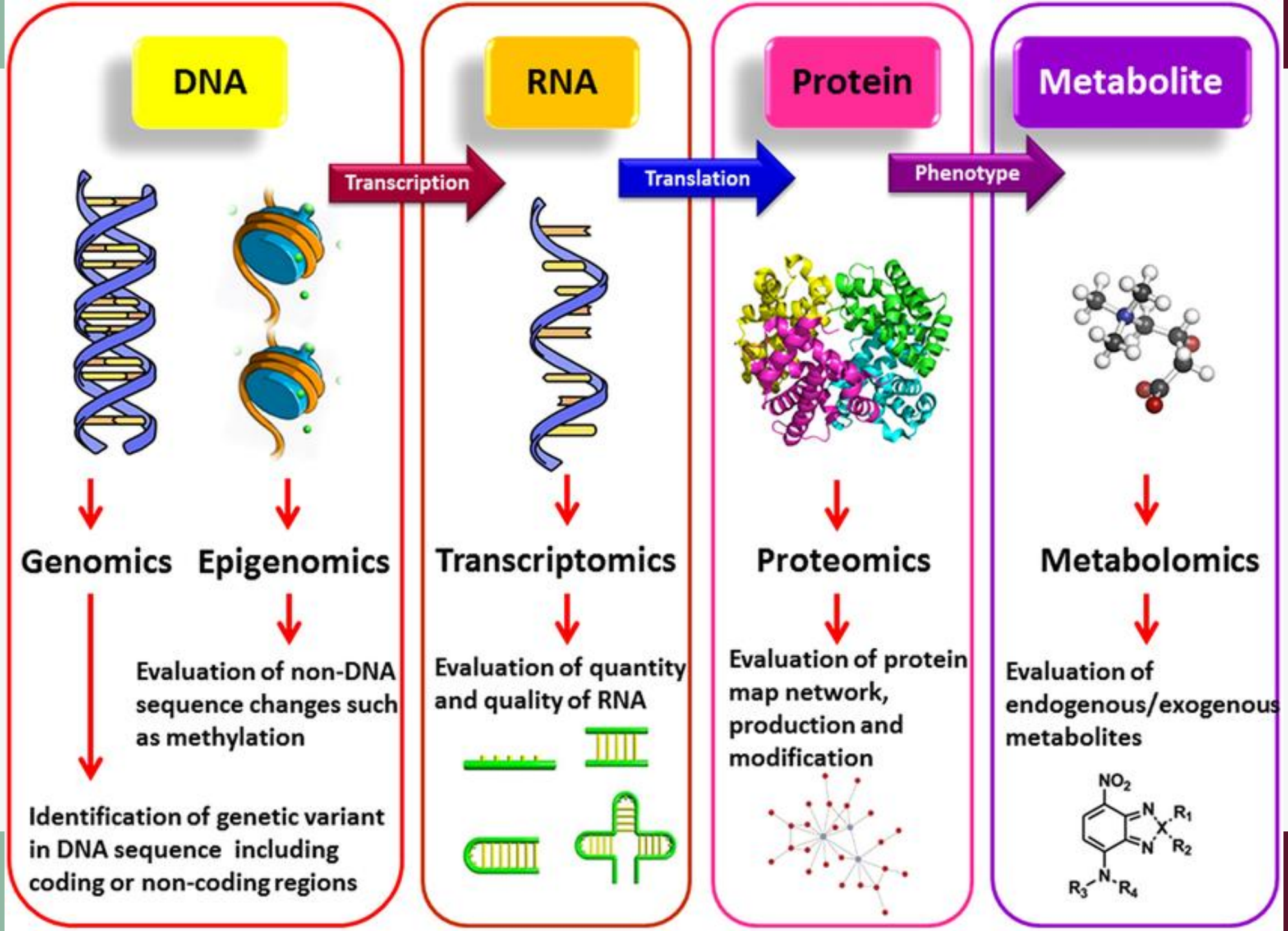
Immuno-based detection

- ELISA - Enzyme-Linked Immunosorbent Assay
- Lateral Flow Immunoassay (LFIA)
- Western Blot



PART 3: OMICS TECHNOLOGIES AND EXAMPLES OF MEDICAL RESEARCH USING OMICS-TYPE DATA

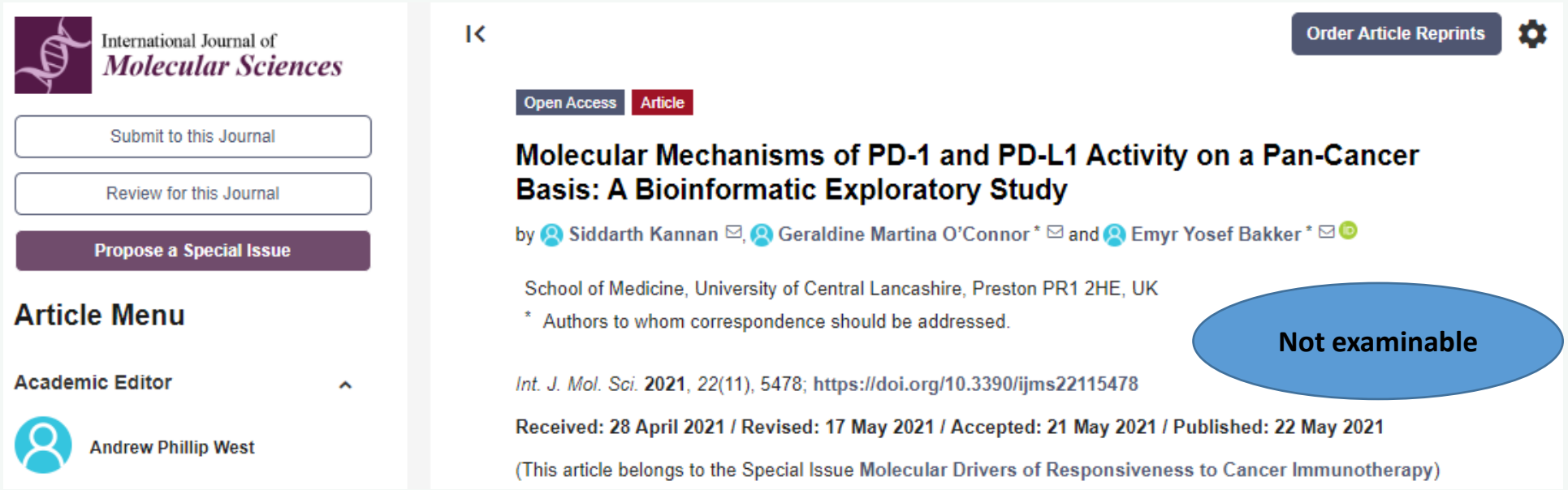
OMICS!


DNA
RNA
Protein
Metabolite
Transcription
Translation
Phenotype
Genomics
Epigenomics
Transcriptomics
Proteomics
Metabolomics
Evaluation of non-DNA sequence changes such as methylation
Identification of genetic variant in DNA sequence including coding or non-coding regions
Evaluation of quantity and quality of RNA
Evaluation of protein map network, production and modification
Evaluation of endogenous/exogenous metabolites

OMICS – advantages and Considerations

Advantages	Considerations
Provides information on thousands of factors at once—classical microarray (transcriptomic) technology gave information on ~54,000 probes!	What do you do with all of this information? How do you know where to **focus**? FDR such as Benjamini-Hochberg and FC can be useful to narrow these down
Data can be deposited in online repositories (e.g. GEO) which allow other people to access it and analyse themselves in a bespoke manner	The large sets of data generated from omics approaches require often different skills and **computational expertise**
The technology is getting cheaper and more accessible	**False positives**! Any potential targets from an omics result need to be validated individually in the laboratory using traditional methods (PCR, Western blot, etc)
	Provides information on **only a single level of biology**—for instance a change in mRNA doesn't necessarily correlate with a change in protein This is where multi-omics analysis can come in useful, but this is an emerging field

THE POWER OF MODERN GENETICS – FROM UCLan's OWN RESEARCH!



International Journal of *Molecular Sciences*

Submit to this Journal

Review for this Journal

Propose a Special Issue

Article Menu

Academic Editor

Andrew Phillip West

Order Article Reprints

Open Access | Article

Molecular Mechanisms of PD-1 and PD-L1 Activity on a Pan-Cancer Basis: A Bioinformatic Exploratory Study

by Siddarth Kannan, Geraldine Martina O'Connor * and Emyr Yosef Bakker *

School of Medicine, University of Central Lancashire, Preston PR1 2HE, UK

* Authors to whom correspondence should be addressed.

Not examinable

Int. J. Mol. Sci. **2021**, *22*(11), 5478; https://doi.org/10.3390/ijms22115478

Received: 28 April 2021 / Revised: 17 May 2021 / Accepted: 21 May 2021 / Published: 22 May 2021

(This article belongs to the Special Issue Molecular Drivers of Responsiveness to Cancer Immunotherapy)

https://www.mdpi.com/1422-0067/22/11/5478

An analysis of the **transcriptomic** data of over **7800 patients** across **fifteen different cancer types**

Identification of **cancer-specific** survival relevance of a **42-gene network** and **one hundred and thirty significant survival associations for 46 microRNAs**

MODERN GENETICS AND CANCER-SPECIFIC SURVIVAL DIFFERENCES

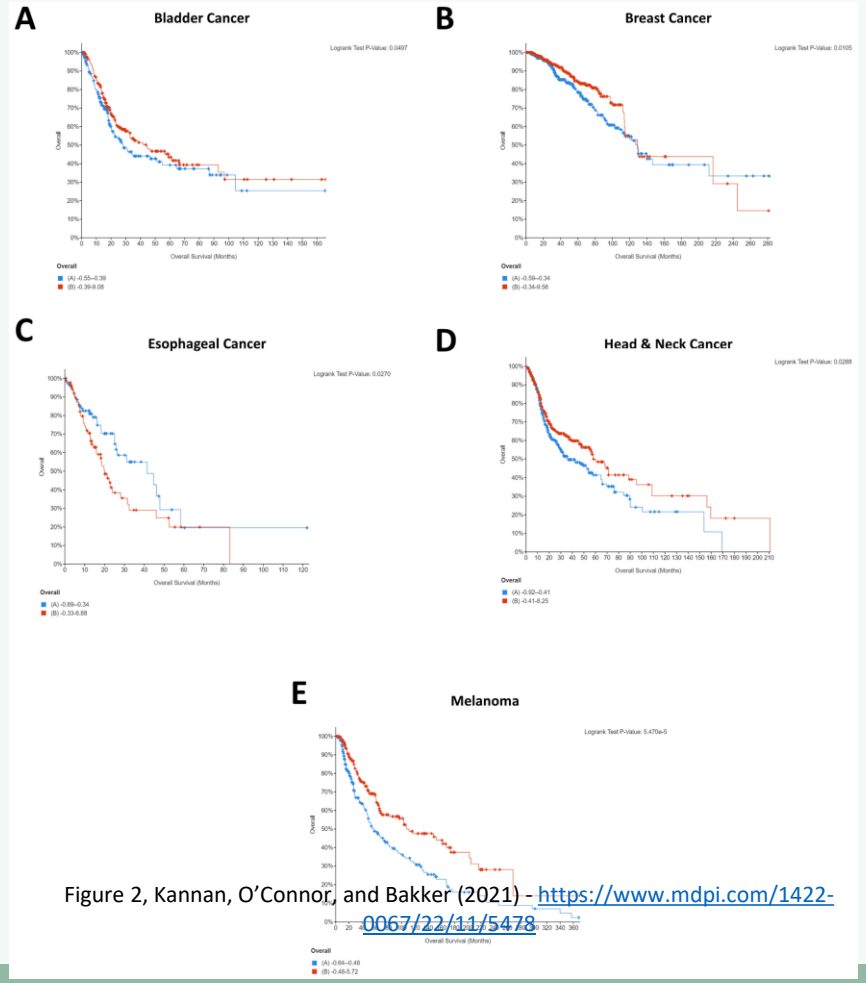


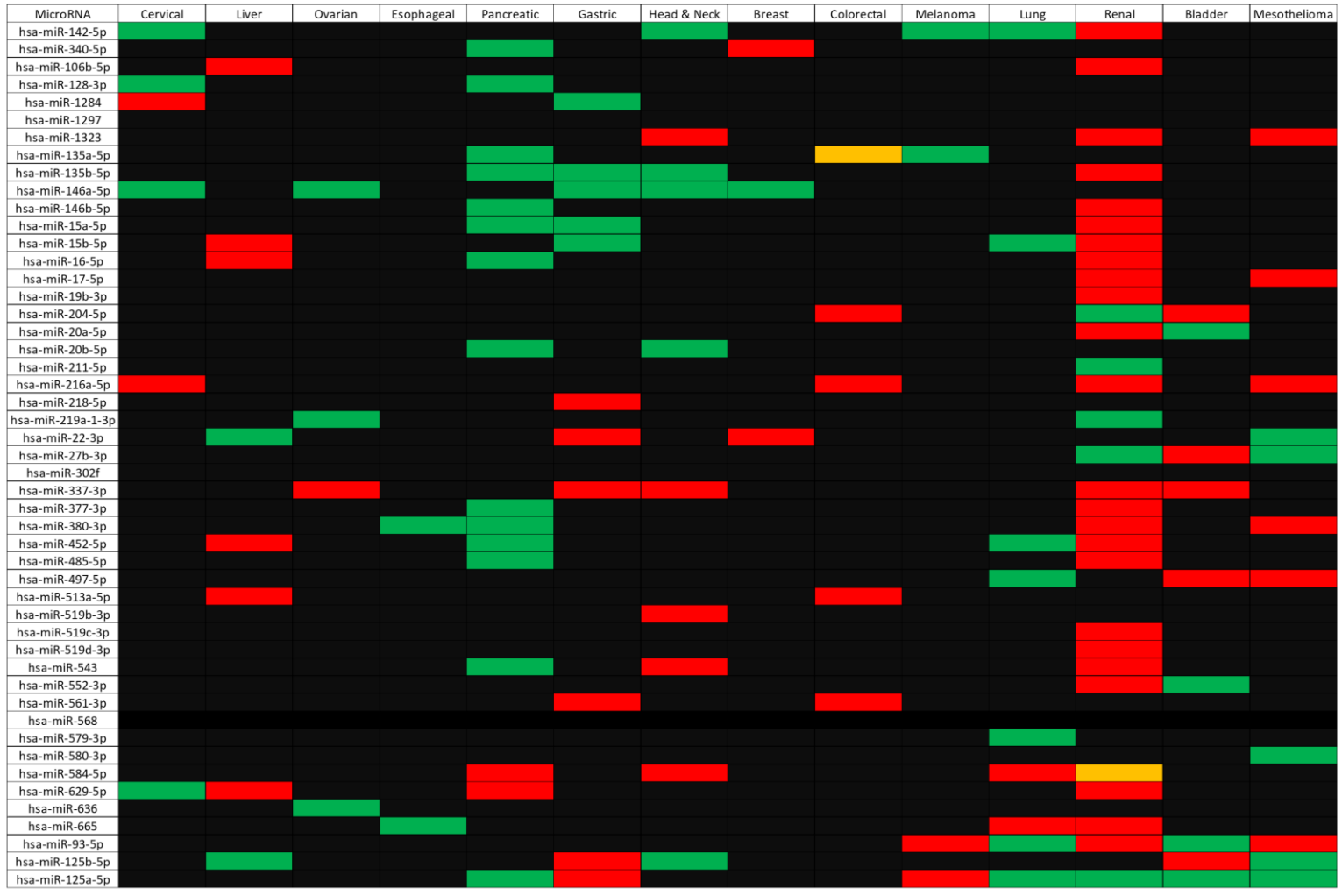
Figure 2, Kannan, O'Connor, and Bakker (2021) - https://www.mdpi.com/1422-0067/22/11/5478

- Kaplan-Meier (survival) analysis based on patient survival time and grouping patients by expression of particular genes

- Despite the same gene being used in A-E, notice that oesophageal cancer has a unique flipped survival profile

Not examinable

FURTHER EXAMPLE OF MODERN GENETICS & CLINICAL STRATIFICATION



- Survival analysis of microRNAs across fourteen cancer types
- Note the unique profile of renal cancer
- Hints at different underlying molecular developments in this cancer type, suggesting a need for altered treatment approaches

Not examinable

Figure 3, Kannan, O'Connor, and Bakker (2021) - https://www.mdpi.com/1422-0067/22/11/5478

FURTHER EXAMPLE OF MODERN GENETICS & CLINICAL STRATIFICATION

		Indication		
Drug	Target	Cancer	Other	Cancer Trial
Cytarabine	JAK2	Leukemia		Yes
Pyrimethamine	JAK2	No	Toxoplasmosis, acute malaria	Yes
Fluorouracil	JAK2	Multiple (including colon, esophageal, gastric, breast, stomach, head and neck, cervical, pancreas, renal cell)		Yes
Sunitinib	JAK2, MAP4K1	Renal cell carcinoma; gastrointestinal stromal tumor		Yes
Azathioprine	JAK2	No	Rheumatoid arthritis, transplant rejection	Yes
Floxuridine	JAK2	Liver cancer and metastases		Yes
Cladribine	JAK2	Leukemia, lymphoma		
Erlotinib	JAK2	Non-small cell lung cancer, pancreatic cancer		Yes
Albendazole	JAK2	No	Anthelmintic	
Triamterene	JAK2	No	Edema	
Podofilox	JAK2	No	Genital warts	
Dasatinib	JAK2, MAP4K1	Chronic myelogenous leukemia, acute lymphoblastic leukemia		Yes
Astemizole	JAK2	No	Allergy	
Trifluridine	JAK2	Colorectal	Keratoconjunctivitis and recurrent epithelial keratitis due to herpes simplex virus	Yes
Disulfiram	CCR5, CXCR6	No	Chronic alcoholism	Yes
Terfenadine	CCR5	No	Allergic rhinitis, hay fever, and allergic skin disorders	
Maraviroc	CCR5	No	HIV-1	Yes
Clioquinol	CXCR6	No	Antifungal	Terminated (Phase 1)
Chloroxine	CXCR6	No	Dandruff and seborrheic dermatitis	
Oxyphenbutazone	CXCR6	No		
Etanercept	LTA	No	Rheumatoid arthritis, plaque psoriasis, polyarticular idiopathic arthritis, psoriatic arthritis, ankylosing spondylitis	
Nilotinib	MAP4K1	Leukemia		Yes
Sorafenib	MAP4K1	Liver, renal		Yes

Table 5, Kannan, O'Connor, and Bakker (2021) - https://www.mdpi.com/1422-0067/22/11/5478

- Through drug repurposing we identified 23 candidate drugs that seem to target different genes of the 42-gene network described earlier

- Note that some were not approved for cancer, which opens up new treatment possibilities

- Particular focus on disulfiram

Not examinable

OTHER GENETIC ADVANCES—FUNCTIONAL ANNOTATION

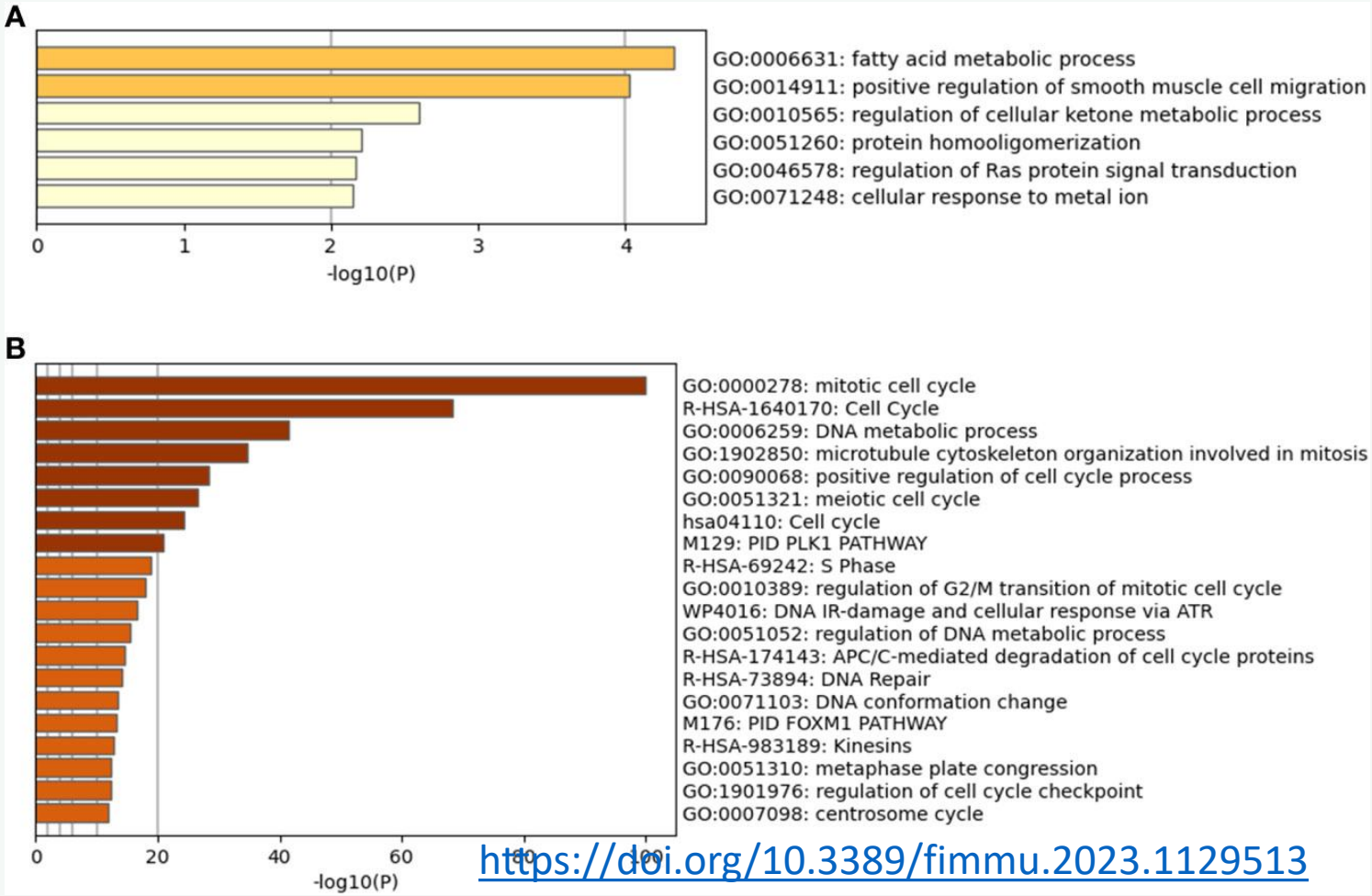
- Advances in genetic technologies have also led to extensive functional annotations for each gene – this can provide insightful information when dozens, hundreds, or thousands of genes are analysed as a group representing a particular group/condition, as you can see overrepresented biological functions

- You do not need to memorise the table below!! However, you just need to understand that functional annotation is a highly beneficial result from advances in genetic technologies and that it can inform you of the function of a gene or group of genes

Gene	Gene Ontology Terms
NR3C1 (The Glucocorticoid Receptor)	negative regulation of transcription from RNA polymerase II promoter, regulation of gluconeogenesis, chromatin organization, regulation of transcription, DNA-templated, regulation of transcription from RNA polymerase II promoter, apoptotic process, cell cycle, chromosome segregation, signal transduction, glucocorticoid metabolic process, gene expression, microglia differentiation, adrenal gland development, intracellular steroid hormone receptor signaling pathway, regulation of glucocorticoid biosynthetic process, synaptic transmission, glutamatergic, maternal behavior, glucocorticoid receptor signaling pathway, glucocorticoid mediated signaling pathway, positive regulation of neuron apoptotic process, negative regulation of transcription, DNA-templated, positive regulation of transcription from RNA polymerase II promoter, astrocyte differentiation, cell division, mammary gland duct morphogenesis, motor behavior, cellular response to steroid hormone stimulus, cellular response to glucocorticoid stimulus, cellular response to dexamethasone stimulus, cellular response to transforming growth factor beta stimulus, positive regulation of pri-miRNA transcription from RNA polymerase II promoter,

Generated on DAVID (https://david.ncifcrf.gov/summary.jsp) on 13.09.23

OTHER GENETIC ADVANCES—FUNCTIONAL ANNOTATION



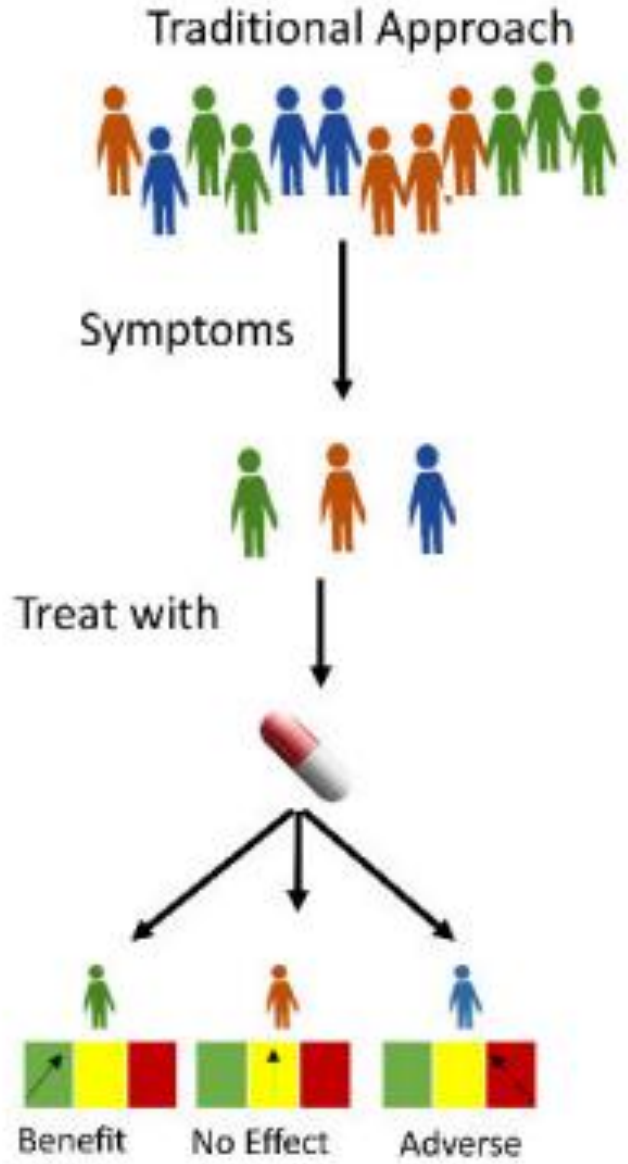
https://doi.org/10.3389/fimmu.2023.1129513

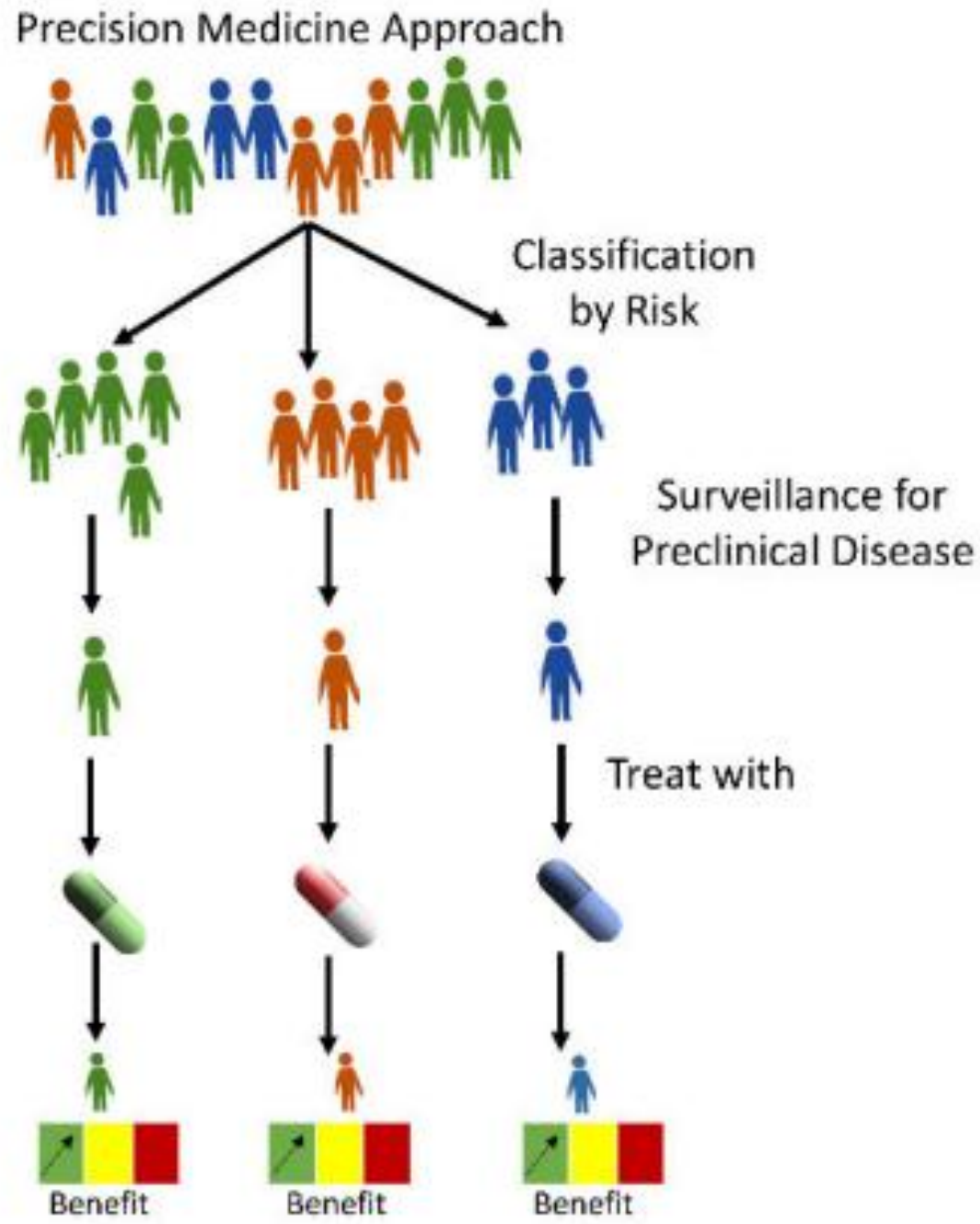
Not examinable

PART 4: PRECISION MEDICINE

The end goal!

- Moving away from current '**batch medicine'** to more **personalised and precise medicine**
- Personalised vs precision medicine





Taken from Afzal et al (http://uclab.khu.ac.kr/resources/publication/J_266.pdf)

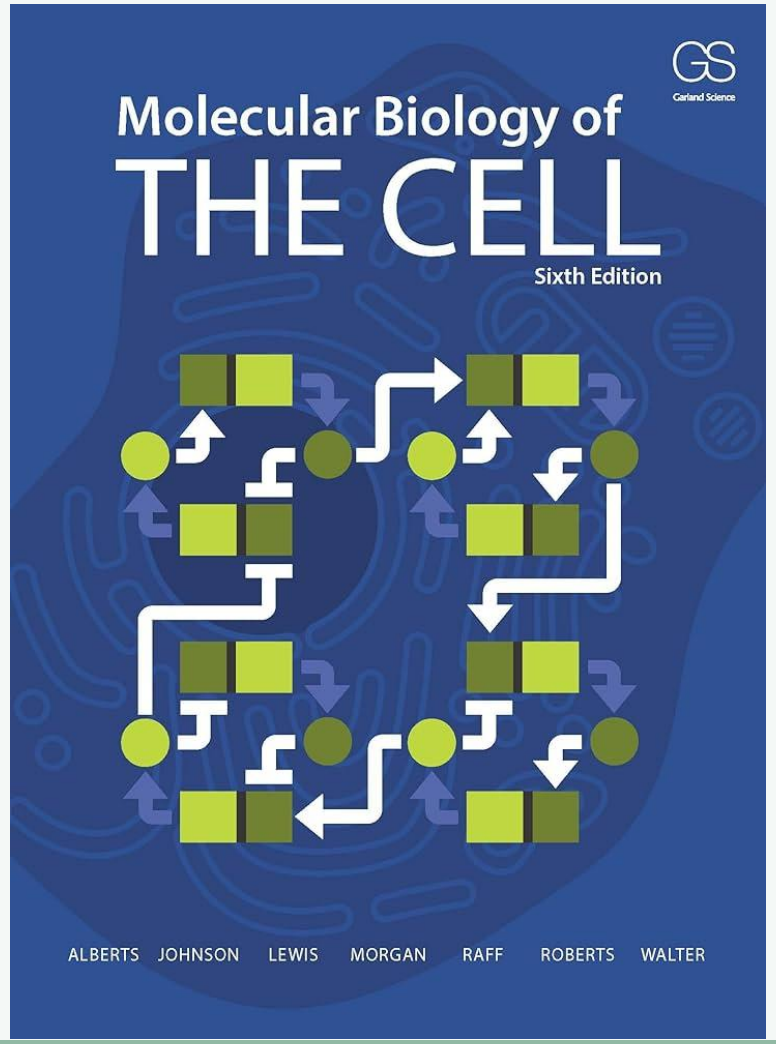
Can you…

- … understand the concept of a paradigm shift and how this relates to advances in medicine?
- … engage with areas of increasing complexity within genetics, including splicing and noncoding RNAs?
- … explain the Human Genome Project?
- … identify laboratory methods?
- … explore advances in genetic technologies that can impact on medical research and, ultimately, clinical practice?

MBBS learning outcomes

- Identify molecular genetic techniques used to identify and diagnose disease
- Describe different techniques for gene analysis and the relevance of omics in clinical medicine

Supplementary reading



- Many figures in this lecture were taken from Molecular Biology of the Cell (6e) – you can refer to Chapter 7 for more details

- You can also read the articles linked at various points through this lecture