# Molecular pathology for specialty training in histopathology

A review matched to the 2021 Royal College of Pathologists curriculum

Matthew Evans
11/04/2024

Matthew Evans (matthew.evans7@nhs.net)
11/04/2024

# Contents

Overview ....... 7
2021 Histopathology Syllabus ....... 8
  Integrated Cellular Pathology Training Syllabus ....... 8
  Histopathology Higher Specialty Training Syllabus ....... 9
The structure of DNA ....... 10
The structure of chromosomes ....... 11
The structure of RNA ....... 12
The structure of genes ....... 12
Gene expression ....... 12
  Transcription ....... 12
  Post-transcriptional modification ....... 13
  Translation ....... 13
  Factors affecting gene expression ....... 14
    DNA methylation ....... 15
    Histone acetylation ....... 15
    Enhancers and silencers ....... 15
    MicroRNA ....... 15
Variation in the genome ....... 15
Heredity ....... 16
  Autosomal inheritance ....... 16
  X-linked inheritance ....... 16
  Y-linked inheritance ....... 17
Molecular alterations ....... 17
  Variants ....... 17
    Small variants ....... 17
    Structural variants ....... 18
    Copy number variants (CNVs) ....... 19
  Methylation ....... 19
  Variant allele frequency (VAF) ....... 19
Germline and somatic variants ....... 19
Oncogenes and tumour suppressor genes ....... 20
  Oncogenes ....... 20
  Tumour suppressor genes ....... 21
Driver and passenger variants ....... 21
Cancer and the immune system ....... 21
Nomenclature ....... 22
  Small variants ....... 22

Structural rearrangements ... 23
Definitions ... 24
Molecular pathology ... 24
Genomics ... 24
The purpose of molecular pathology ... 24
Diagnostic markers ... 24
Prognostic markers ... 25
Predictive markers ... 25
Markers predicting response to targeted therapy ... 25
Markers predicting response to immunotherapy ... 26
Markers predicting response to chemotherapy ... 26
In situ techniques ... 27
Pre-analytics ... 27
Immunohistochemistry ... 28
PD-L1 immunohistochemistry ... 29
MMR immunohistochemistry ... 29
In situ hybridisation ... 32
Using in situ hybridisation to detect the presence of nucleic acids ... 34
Using in situ hybridisation to assess for copy number variants ... 34
Using in situ hybridisation to assess for gene rearrangements ... 35
Nucleic acid techniques ... 37
Pre-analytics ... 37
Tissue processing ... 37
Tumour assessment ... 40
Section preparation ... 42
Nucleic acid extraction ... 43
Sequencing ... 44
Direct/Sanger sequencing ... 44
Pyrosequencing ... 45
Next-generation sequencing ... 46
Data analysis ... 52
PCR-based techniques ... 53
Real-time PCR ... 53
Reverse transcription PCR ... 56
Methylation analysis ... 58
Clonality analysis ... 61
Microsatellite analysis ... 65
Microarrays ... 66

Selecting the appropriate technique ..... 69
Techniques for assessing for small variants ..... 69
Scenarios ..... 70
Techniques for assessing for structural variants ..... 71
Scenarios ..... 72
Techniques for assessing for copy number variants ..... 72
Turnaround times ..... 74
Failure rates ..... 75
External quality assurance (EQA) ..... 76
Historical provision of molecular pathology services in England ..... 76
Current/future provision of molecular pathology services in England ..... 77
Genomic testing services ..... 77
Using Genomic Laboratory Hubs ..... 78
The advantages of GLHs ..... 78
The disadvantages of GLHs ..... 78
Molecular tests used in multiple tumour types ..... 79
*BRAF* small variant testing ..... 79
*BRCA1*/*BRCA2* variant and HRD testing ..... 80
*EGFR* small variant testing ..... 81
*FGFR* alterations ..... 81
*HER2* copy number variant testing ..... 82
*KIT* small variant testing ..... 82
*KRAS*, *HRAS* and *NRAS* small variant testing ..... 83
*NTRK* structural variant testing ..... 83
PD-L1 immunohistochemistry ..... 84
*RET* structural variant testing ..... 85
*RET* small variant testing ..... 85
*TERT* promoter variant testing ..... 86
*TP53* small variant and copy number variant testing ..... 86
Molecular testing in colorectal cancer ..... 86
MMR immunohistochemistry and MSI analysis ..... 86
*KRAS* and *NRAS* small variant testing ..... 87
*BRAF* small variant testing ..... 87
*HER2* copy number variant testing ..... 87
Molecular testing in ovarian cancer ..... 88
*BRCA1*/*BRCA2* variant testing ..... 88
HRD testing ..... 88
*SMARCA4* small variant testing ..... 89

*KRAS* small variant testing in ovarian mucinous neoplasms ... 89
Molecular testing in endometrial cancer ... 89
MMR immunohistochemistry ... 89
*POLE* variant testing ... 90
*HER2* amplification testing in endometrial serous carcinoma ... 91
Molecular testing in breast cancer ... 91
*HER2* amplification testing ... 91
Oncotype DX testing ... 91
*PIK3CA* small variant testing ... 91
*NTRK* structural variant testing in secretory carcinoma ... 92
PD-L1 immunohistochemistry in triple negative breast cancer ... 92
Germline *BRCA1/BRCA2* variant testing ... 93
Molecular testing in non-small cell lung cancer ... 93
*EGFR* small variant testing ... 93
*KRAS* small variant testing ... 93
*BRAF* small variant testing ... 94
*ALK* structural variant testing ... 94
*ROS1* structural variant testing ... 94
*RET* structural variant testing ... 95
*MET* exon 14 skipping variant testing ... 95
PD-L1 immunohistochemistry ... 96
*MET* copy number variant testing ... 96
*HER2* small variant testing ... 97
Molecular testing in small cell lung cancer ... 97
*RB1* alteration testing ... 97
Molecular testing in mesothelioma ... 98
*CDKN2A* copy number variant testing ... 98
Molecular testing in salivary gland neoplasms ... 98
*MAML2* structural variant testing in mucoepidermoid carcinoma ... 98
*MYB*::*NFIB* fusion testing in adenoid cystic carcinoma ... 98
*NTRK* structural variant testing in secretory carcinoma ... 98
Molecular testing in melanoma ... 99
*BRAF* small variant testing ... 99
*KIT* small variant testing ... 99
*NRAS* small variant testing ... 100
Diagnostic copy number variant testing ... 100
*TERT* promoter variant testing ... 100
Structural variant testing in spitzoid neoplasms ... 100

PD-L1 immunohistochemistry ........ 101
Molecular testing in uveal melanoma ........ 101
Cytogenic abnormalities ........ 101
Molecular testing in gastrointestinal stromal tumours (GISTs) ........ 101
*KIT* small variant testing ........ 101
*PDGFRA* small variant testing ........ 102
Molecular testing in thyroid neoplasms ........ 102
Molecular testing in papillary thyroid carcinoma ........ 102
*BRAF* small variant testing ........ 102
*RAS* small variant testing ........ 103
*TERT* promoter variant testing ........ 103
*RET* structural variant testing ........ 103
Molecular testing in follicular thyroid carcinoma ........ 103
*RAS* small variant testing ........ 103
*RET* structural variant testing ........ 103
Molecular testing in anaplastic thyroid carcinoma ........ 103
*TP53* small variant and copy number variant testing ........ 103
*BRAF* small variant testing ........ 104
*ALK*, *RET* and *NTRK* structural variant testing ........ 104
Molecular testing in medullary thyroid carcinoma ........ 104
*RET* small variant testing ........ 104
Molecular testing in thyroid Hurtle cell carcinoma ........ 104
*RET* structural variant testing ........ 104
Molecular testing in non-invasive follicular thyroid neoplasm with papillary-like nuclei (NIFTP) ........ 105
*BRAF* and *RAS* small variant testing ........ 105
Molecular testing in phaeochromocytoma ........ 105
*RET* small variant testing ........ 105
Molecular testing in adrenal cortical carcinoma ........ 105
*TP53* small variant and copy number variant testing ........ 105
Molecular testing in head and neck squamous cell carcinoma ........ 105
*CDK2NA* small variant and copy number variant testing ........ 105
*EGFR* small variant testing ........ 106
*TP53* small variant and copy number variant testing ........ 106
*RET* structural variant testing ........ 106
PD-L1 immunohistochemistry ........ 106
Molecular testing in renal cell carcinoma ........ 106
Molecular defined renal cell carcinomas ........ 106

*VHL* small variant and chromosome 3 copy number variant testing........................................ 107
*TSC1/TSC2* small variant testing........................................ 107
*MET* small variant and chromosome 7 copy number variant testing........................................ 108
Chromosome 17 copy number variant testing........................................ 108
*BRAF* small variant testing........................................ 108
Molecular testing in urothelial carcinoma........................................ 108
*FGFR2* and *FGFR3* alteration testing........................................ 108
PD-L1 immunohistochemistry........................................ 108
Molecular testing in prostate cancer........................................ 109
*BRCA1* and *BRCA2* variant testing........................................ 109
*TMPRSS2::ERG* fusion testing........................................ 109
Molecular testing in pancreatic cancer........................................ 109
*BRCA1* and *BRCA2* variant testing........................................ 109
MMR immunohistochemistry and MSI analysis........................................ 110
Molecular testing in biliary cancers........................................ 110
*FGFR2* structural variant testing in cholangiocarcinoma........................................ 110
*IDH1* small variant testing in cholangiocarcinoma........................................ 110
MMR immunohistochemistry and MSI analysis........................................ 111
Molecular testing in gastro-oesophageal carcinomas........................................ 111
*HER2* amplification testing........................................ 111
MMR immunohistochemistry and MSI analysis........................................ 111
PD-L1 immunohistochemistry........................................ 111
Molecular testing in thymic carcinoma........................................ 112
*KIT* small variant testing........................................ 112
Molecular testing in cervical cancer........................................ 112
PD-L1 immunohistochemistry........................................ 112
Molecular testing in sebaceous neoplasms........................................ 113
MMR immunohistochemistry and MSI analysis........................................ 113
Other tumour types........................................ 114
High-grade mature B-cell lymphomas........................................ 114
Diagnostic alterations in haematolymphoid neoplasia........................................ 114
Diagnostic alterations in bone and soft tissue neoplasia........................................ 114
Diagnostic alterations in CNS tumours........................................ 116

# Overview

This document provides an overview of molecular pathology (or, more broadly, biomarker pathology) which is directly targeted at trainees following the 2021 Curriculum for Specialty Training in Histopathology from the Royal College of Pathologists.

It is based on the requirements set out in the College's 2021 Histopathology Syllabus.

It is cautioned that this is intended to be an overview of the knowledge and skills stipulated in the curriculum and is in no way intended to replace the knowledge and skills gained through everyday exposure to clinical practice.

# 2021 Histopathology Syllabus

The syllabus is divided into two parts:

- The Integrated Cellular Pathology Training Syllabus applies to trainees undertaking any of the cellular pathology training programmes (years 1 and 2).
- The Histopathology Higher Specialty Training Syllabus applies only to those trainees undertaking histopathology training (i.e. not diagnostic neuropathology, paediatric and perinatal pathology, and forensic histopathology) (years 3, 4 and 5).

## Integrated Cellular Pathology Training Syllabus

| | **Knowledge** | **Skills** |
|---|---|---|
| **Fundamentals of molecular biology** | Understands principles of common molecular pathology techniques.<br><br>Demonstrates understanding of and identifies the origins and consequences of germ-line variation and somatic mutations, including DNA methylation and gene expression changes<br><br>Demonstrates knowledge of basic molecular databases<br>Demonstrates knowledge of how histological samples are taken, prepared and of how nucleic acids are extracted from them<br><br>Understands the principles of the most up-to-date molecular methods<br><br>Demonstrates knowledge of molecular tests currently performed on histological samples, including the limitations of these tests and of tests that are anticipated in the near future | Demonstrates ability to understand origins of, and justifications for, molecular tests<br><br>Demonstrates ability to retrieve relevant data from public sources<br><br>Demonstrates ability to undertake the appropriate sample collection, retrieval and preparation for the common molecular tests, whether performed on extracted nucleic acid or in situ<br><br>Demonstrates knowledge of sequencing, PCR, microarrays (DNA and RNA), in situ hybridisation and mutation detection<br><br>Demonstrates ability to assess the demand for molecular tests and the modes of supply |
| **Fundamentals of genetics** | Identifies the structure of genes including translation and transcription, factors affecting gene expression and inheritance patterns | Recognises the factors affecting transcription and translation |
| **Molecular techniques** | Identifies molecular techniques | Demonstrates awareness of principles, practical knowledge of sequencing, PCR, microarrays (DNA and RNA), in situ hybridisation and mutation detection |

| | | |
|---|---|---|
| **Molecular tests** | Describes molecular tests currently performed on histological samples | Interprets the common molecular tests |

Items from this portion of the syllabus are marked:

## Histopathology Higher Specialty Training Syllabus

| | Knowledge | Skills |
|---|---|---|
| **General molecular pathology** | Describes the origins and consequences of germline variation and somatic mutations, including DNA methylation and gene expression changes | Demonstrates the origins of and justifications for molecular tests |
| **Fundamentals of databases and bioinformatics** | Demonstrates ability to recall the basic molecular databases | Summarises the use of data and identify relevant data from public sources |
| **Use of histology samples** | Describes how histological samples are taken and prepared, and how nucleic acids are extracted from them | Demonstrates practical understanding of undertaking the appropriate sample collection, retrieval and preparation for the common molecular tests, whether performed on extracted nucleic acid or in situ |
| **Technology** | Outlines the principles and limitations of the most up-to-date molecular methods | Demonstrates practical knowledge of sequencing, PCR, microarrays (DNA and RNA), in situ hybridisation, mutation detection |
| **Molecular tests** | Describes molecular tests currently performed on histological samples, including the limitations of these tests and of tests that are anticipated in the near future | Demonstrates the demand for molecular tests and the modes of supply<br><br>Describes and explains common molecular tests including some of the common pitfalls and how to avoid them<br><br>Illustrates the significance of common molecular tests |

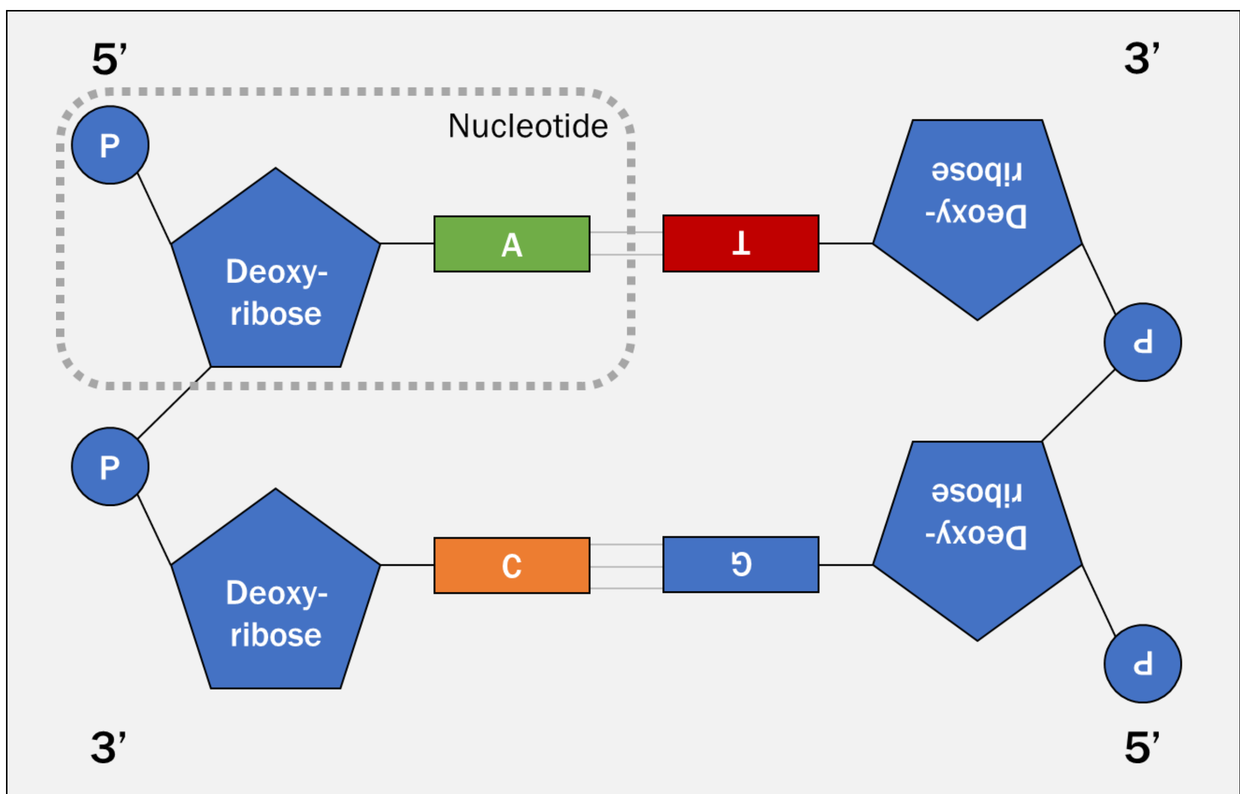
Items from this portion of the syllabus are marked:

Matthew Evans (matthew.evans7@nhs.net)
11/04/2024

# Fundamentals of molecular biology and genetics

## The structure of DNA

- Human genetic material is stored as deoxyribonucleic acid (DNA)
- DNA comprises:
  - A sugar-phosphate backbone, which comprises alternating deoxyribose and phosphate units
  - Bases
- There are four DNA bases:
  - Adenine (A) and guanine (G) are purines
  - Thymine (T) and cytosine (C) are pyrimidines
- A single unit of a base, a deoxyribose sugar and a phosphate unit is called a nucleotide
- Chains of DNA bind together through hydrogen bonds between the bases to form a double strand
- Binding of bases (and therefore nucleotides) is complementary:
  - Adenine binds thymine
  - Cytosine binds guanine
- Owing to the presence of phosphate units, DNA overall has a negative charge
- The vast majority of DNA in human cells resides in the nucleus, with a small amount being present in mitochondria (mitochondrial DNA or mtDNA)

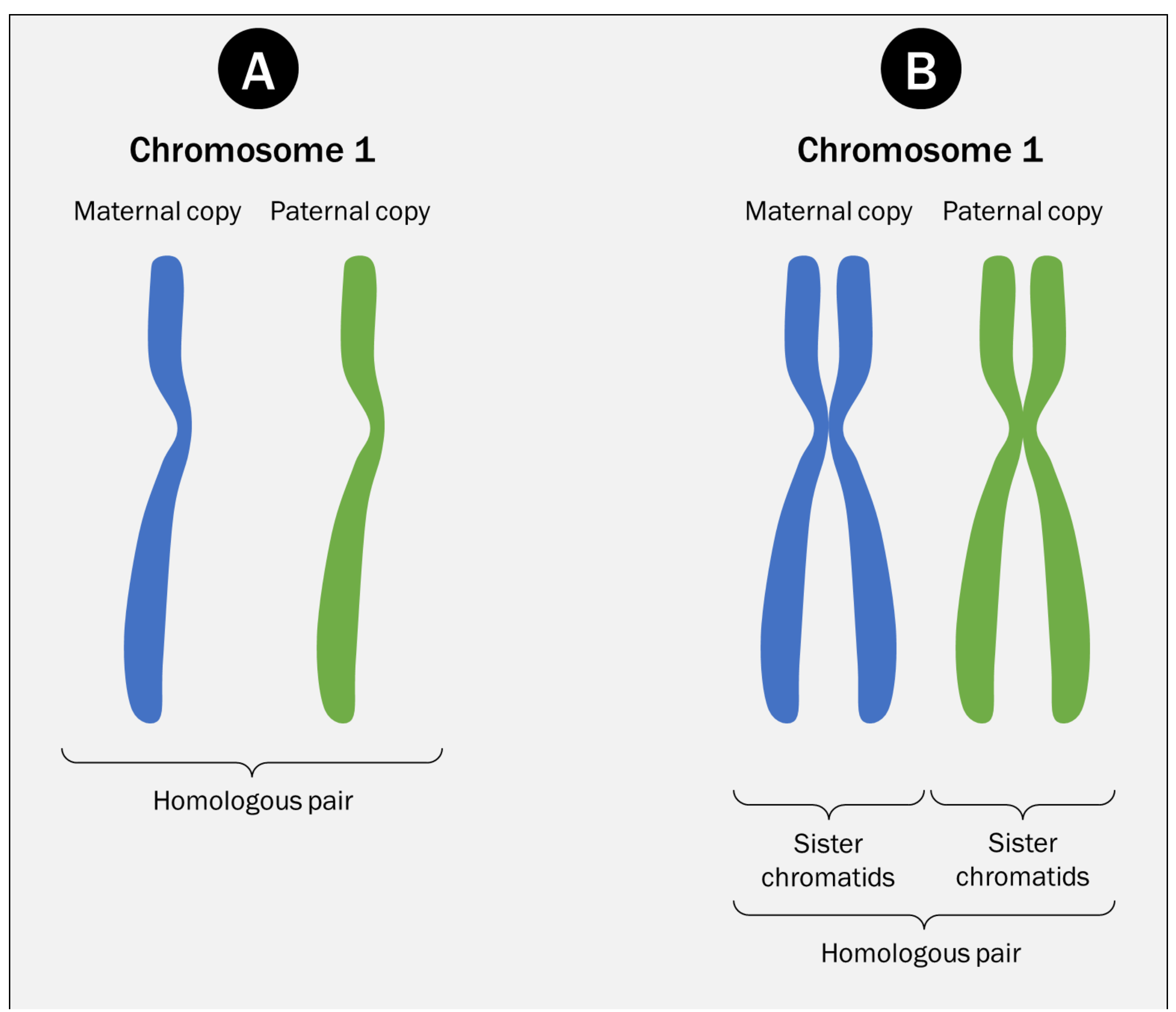



*DNA is a polymer formed of nucleotides, which themselves comprise a deoxyribose sugar, a phosphate unit and a base. It is double-stranded, with two complementary DNA molecules binding together through hydrogen bonds between bases. Each DNA strand has a 5' (5 prime) and 3' (3 prime) end.*

Matthew Evans (matthew.evans7@nhs.net)
11/04/2024

## The structure of chromosomes

- DNA is packaged into large structures called chromosomes
- Chromosomes contain histone proteins which condense the DNA and give chromosomes their structure
- Normal human somatic cells contain 23 pairs of chromosomes (46 chromosomes in total):
  - In each pair, one chromosome is inherited from each of the individual's parents
  - The two copies of each chromosome are known as homologous chromosomes (i.e. both copies of chromosome 1 are homologous chromosomes, one each of which are inherited from each of the individual's parents)
  - Every normal somatic cell in humans contains 22 pairs of autosomes and one pair of allosomes (biological females have two copies of the X-chromosome, and biological males have one X-chromosome and one Y-chromosome)
  - The state of having two copies of each chromosome is described as 'diploid'
- In contrast, gametes are 'haploid':
  - They contain only one copy of each of the autosomes
  - They contain only one allosome (either an X- or Y-chromosome, which determines the biological sex of the individual)
- During cell division, each chromosome is replicated to give rise to two identical sister chromatids
- Aneuploidy is the state in which cells have abnormal numbers of chromosomes



*A: In normal somatic cells, there are two copies of each autosomal chromosome, one inherited from each of the individual's parents. These form a homologous pair of chromosomes.*

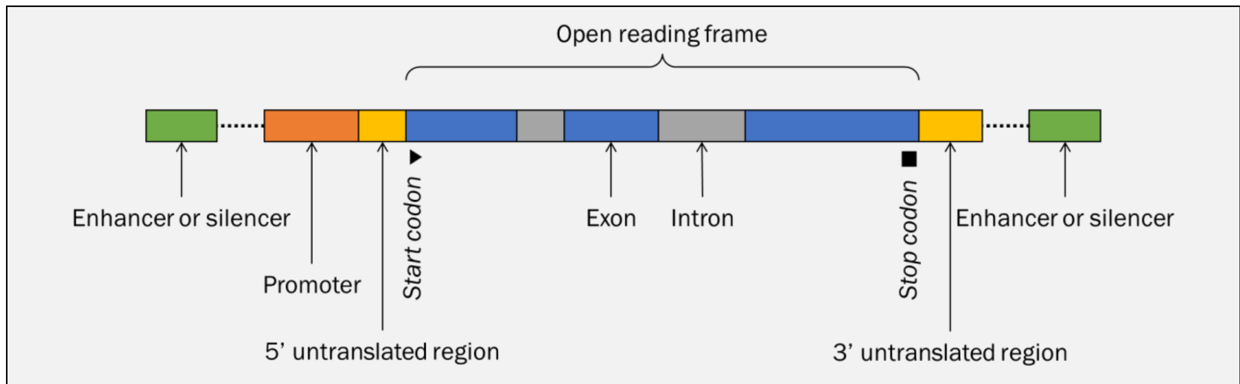
*B: During cell division each chromosome replicates, two form two pairs of sister chromatids for each chromosome. For this brief period, the cell contains four copies of each autosomal chromosome.*

## The structure of RNA

- RNA, like DNA, is a nucleic acid
- It differs from DNA:
  - It is single stranded
  - Its sugar is ribose rather than deoxyribose
  - Rather than containing thymine, it contains uracil
- RNA exists as messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA)
- RNA is inherently less stable than DNA

## The structure of genes

- A gene is a sequence of DNA which encodes the synthesis of a gene product, which can be either RNA or a protein




- Genes contain introns and exons:
  - Exons comprise DNA sequences which ultimately determine the amino acid sequence of the protein
  - Introns have a wide variety of functions, but do not provide the template for the amino acid sequence of the protein
  - Introns are often long and contain repetitive sequences

## Gene expression

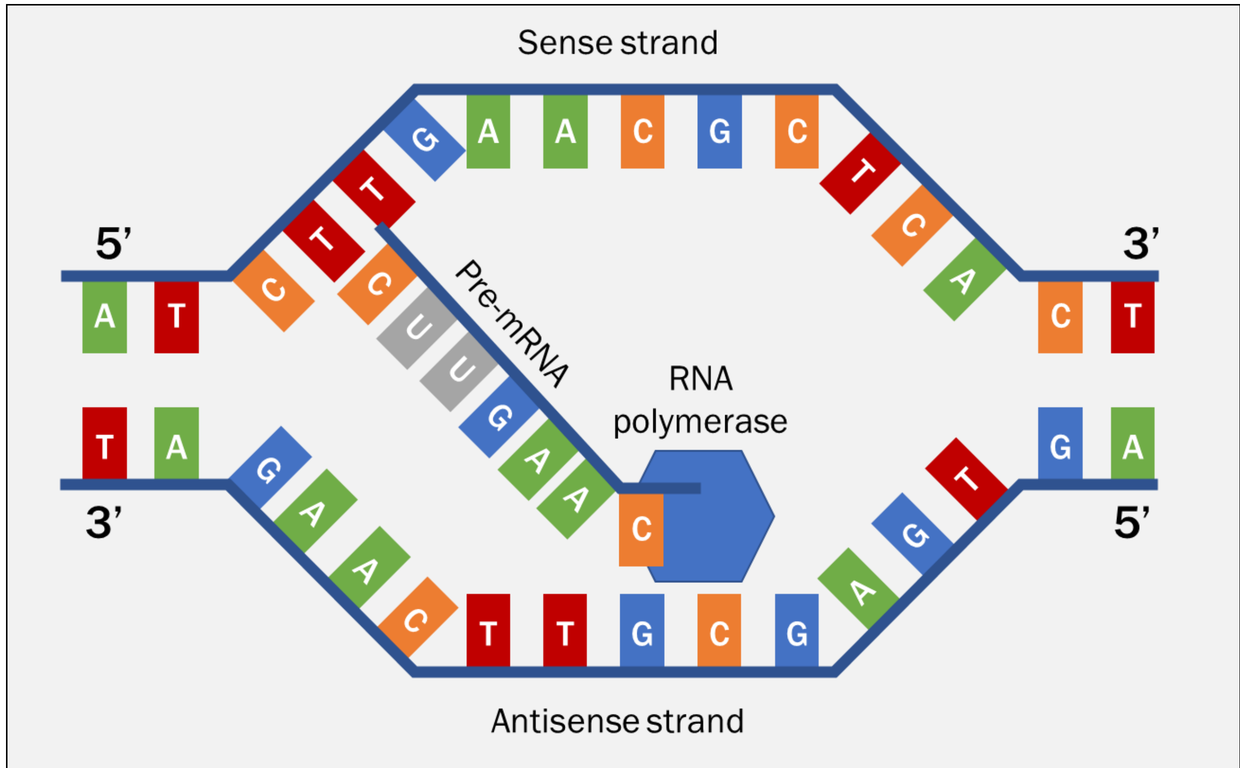
**I** Identifies the structure of genes including translation and transcription, factors affecting gene expression and inheritance patterns

- The process of a gene giving rise to its product is known as gene expression

### Transcription

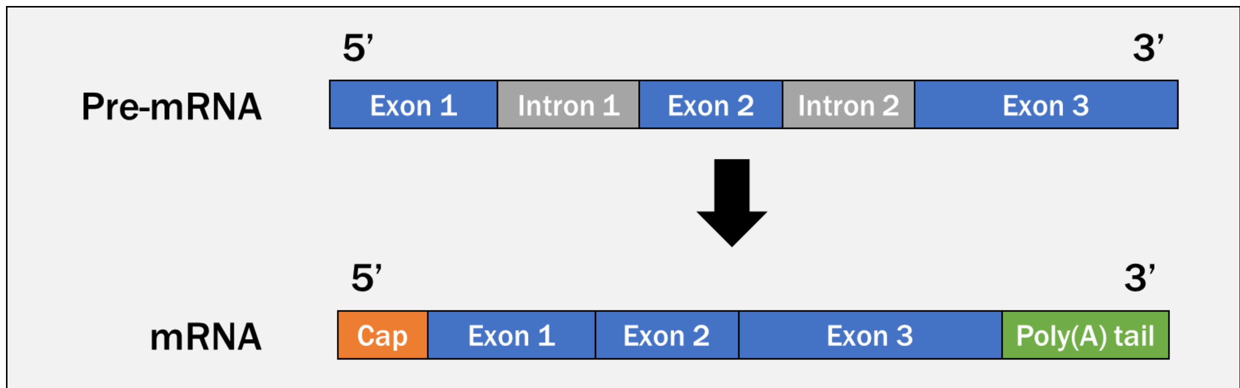
- Helicase breaks the hydrogen bonds between the DNA strands, causing it to unwind and separate into single strands
- One strand is the sense strand, and the other is the antisense strand
- RNA polymerase binds the antisense strand at the promoter region of the gene to be transcribed

- One-by-one, RNA polymerase forms an RNA strand by one-by-one incorporating nucleotides which are complementary to each nucleotide on the template strand (incorporating uracil where it encounters adenine in the template sequence)
- Transcription proceeds from the promoter at the 3' end of the antisense strand to the 5' end of the gene
- RNA polymerase has a proofreading function which prevent non-complementary nucleotides from remaining in the newly-formed RNA
- This results in pre-mRNA which has a sequence identical to that in the coding strand (apart from thymine being replaced by uracil)




## Post-transcriptional modification

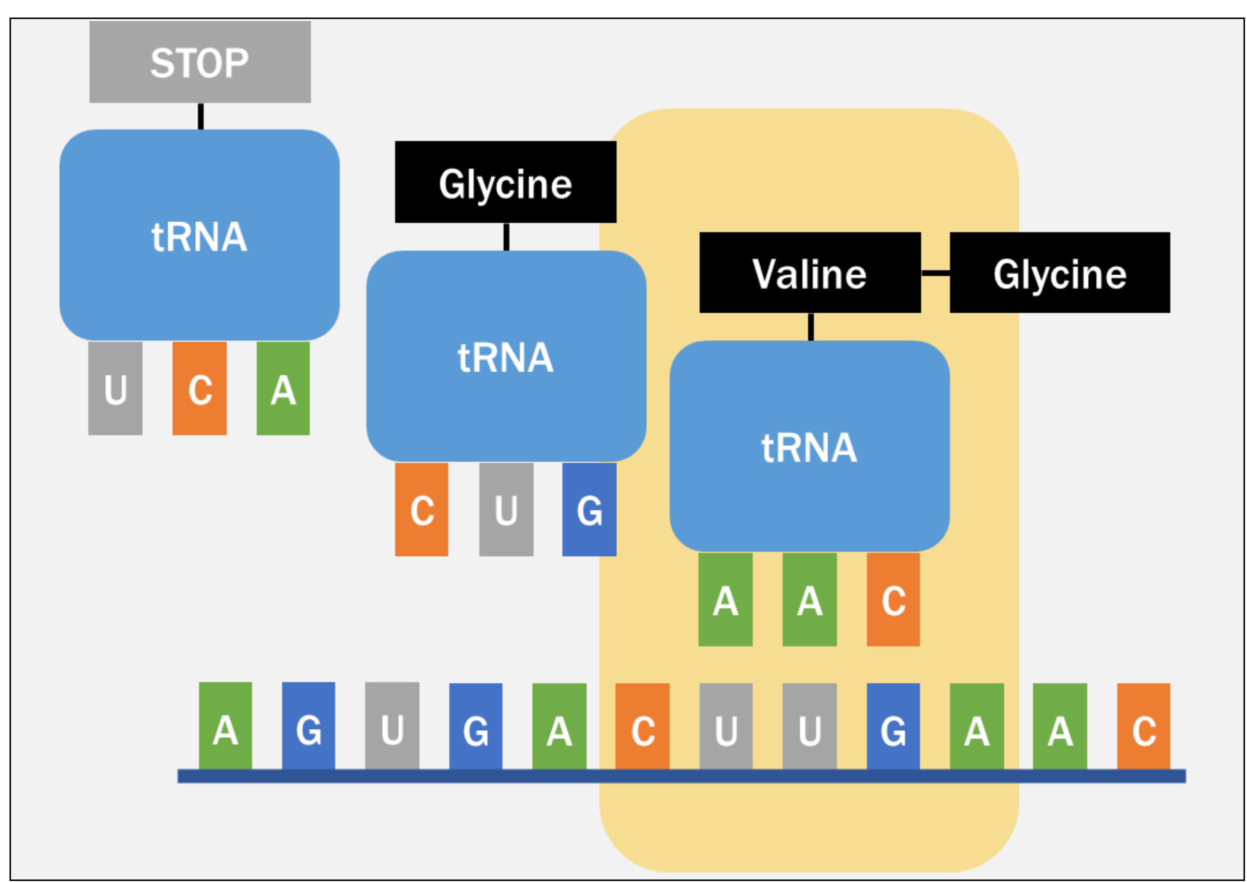
- The pre-mRNA undergoes a number of modifications, including RNA splicing
  - A spliceosome removes introns from the pre-mRNA
  - This leaves only exons
- The end result is mRNA



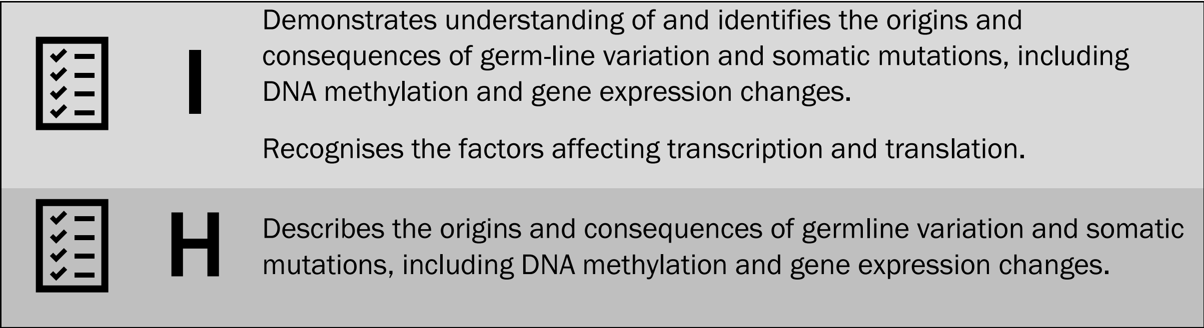

## Translation

- In the cytoplasm, mRNA encounters ribosomes (which are made of rRNA)
- The mRNA strand gradually moves through the ribosome

- The ribosome uses the sequence of the mRNA as a template to form a strand of amino acids
- Each three-nucleotide sequence in the RNA (a codon) is complementary to the sequence on a tRNA molecule
- Each tRNA molecule carries an amino acid, with each codon corresponding to a particular amino acid
- As each codon moves through the ribosome, a tRNA is recruited, and a new amino acid is added to the growing polypeptide
- Translation begins when a start codon (AUG) is detected by the ribosome, and stops ends when an end codon (UGA, UAG, UAA) is encountered
- The genetic code is degenerate:
  - Multiple similar codons encode the same amino acid
  - This means that variants resulting in a single nucleotide change may well not change the amino acid
  - This reduces the likelihood of DNA damage resulting in functional impact




## Factors affecting gene expression



### DNA methylation

- Transcription begins when transcription factors and RNA polymerase bind to the promoter region of the gene
- DNA is methylated by methyltransferase enzymes, which add methyl residues to the cytosine resides in CpG dinucleotide sequences:
  - These are sequences which comprise multiple repetitive units of cytosine-guanine
  - A region of the genome with large numbers of CpG dinucleotides is known as a CpG island
- Many gene promoters contain CpG islands, and therefore gene promoter regions are prone to methylation
- Gene promoter methylation prevents transcription factor and RNA polymerase binding, and therefore suppresses expression of (or silences) the gene
- This is an example of epigenetic regulation

### Histone acetylation

- Transcription cannot occur if DNA is tightly bound in chromosomes to histone proteins
- Acetylation of the histone proteins releases DNA, and promotes expression of the genes in that region

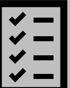
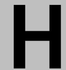
### Enhancers and silencers

- General transcription factors bind to gene promoter regions and initiate transcription
- Enhancers and silencers are sequences which are distant from the genes they control
- When transcription factors bind to them, they dramatically promote or repress expression of the gene, respectively

### MicroRNA

- miRNA are short RNA strands
- They bind complementary sequences in mRNA
- This prevents the mRNA sequence from being translated, and so prevents protein expression

## Variation in the genome

| | |
|---|---|
| | Demonstrates understanding of and identifies the origins and consequences of germ-line variation and somatic mutations, including DNA methylation and gene expression changes. |
|   | Describes the origins and consequences of germline variation and somatic mutations, including DNA methylation and gene expression changes. |

- The genome is the sum of all genetic information in an individual
- There is variation in genomes between individuals
- Variation generally arises because of either:
  - Changes in the DNA sequence introduced during DNA replication, which go unrepaired
  - Environmental factors which change the DNA sequence (e.g. ionising radiation)
- There is also variation within the genome of a single individual:
  - Each normal somatic cell contains two versions of each chromosome (one from each of the individual's parents)
  - Therefore, for genes on autosomal chromosomes, each individual has two copies (again, from each of the individual's parents)
  - Different versions of the same gene are called alleles

  - If both alleles of a given gene are the same in an individual, that individual is considered homozygous with respect to that gene
  - If the alleles are different, the individual is considered heterozygous with respect to that gene
- Variants can range from a change in a single nucleotide up to enormous changes of the DNA sequence
- Variants vary in their functional impact:
  - Variants which do not cause disease are considered benign
  - Variants which can cause disease are considered pathogenic
  - In some cases, it is not clear whether a variant causes disease or not (variants of uncertain significance, VUS)
- In general, variants which are present in a large proportion of the population are likely to be benign
- It is difficult to decide what actually constitutes a 'variant':
  - Given normal variation in the genomes of individuals, there is no single 'normal' genome
  - The Genome Reference Consortium produces a reference human genome based on sequencing of volunteers from the US
  - The definition of a variant is based on this reference genome
  - Non-Caucasian individuals have higher numbers of 'variants' purely because the reference genome is based largely on Caucasian populations
- The 'typical' (or at least most prevalent) allele is referred to as being wild-type:
  - Anything differing from the wild-type allele is considered a variant or mutation
  - 'Variant' or 'mutation' should not necessarily be considered pejorative – it simply represents a difference from what is considered 'typical'

## Heredity

- Definitions:
  - Genotype: the genetic underpinning of an organism or trait
  - Phenotype: the observable characteristics
- Genotype can impact on phenotype, but does not entirely determine it since environmental factors also have an impact
- When one allele of a gene can override the effect of a second allele, it is considered dominant and the second allele considered recessive
- Harbouring a given allele does not necessarily mean that an individual will express the phenotype:
  - An allele with complete penetrance will always be expressed in the individual's phenotype (even autosomal dominant alleles do not necessarily have complete penetrance)
  - An allele with incomplete penetrance is not expressed in the phenotype of all individuals with the allele
- Traits (or alleles) can show autosomal or sex-linked inheritance

### Autosomal inheritance

- This is the pattern of inheritance shown by genes on the autosomal chromosomes
- Heterozygotes will express only the dominant allele as their phenotype
- In order to express the phenotype of a recessive allele, the individual must be homozygous with respect to that all

### X-linked inheritance

- The encoding gene is located on the X chromosome

- X-linked recessive alleles are expressed in females only when they are homozygous (i.e. have two copies of the same allele)
- X-linked dominant alleles can be expressed in females who are heterozygous or homozygous (i.e. have one or two copies of the same allele)
- Because males have only one X chromosome, if they inherit an X-linked allele it will be expressed phenotypically (be it recessive or dominant)
- Sons cannot inherit X-linked alleles from their fathers, because fathers by definition only pass on their Y chromosomes to their sons

### Y-linked inheritance

- The encoding gene is located on the Y chromosome
- Daughters of an affected male will not inherit the allele or express the trait
- Males of an affected male will always inherit the allele, because their Y chromosome must be inherited from their father

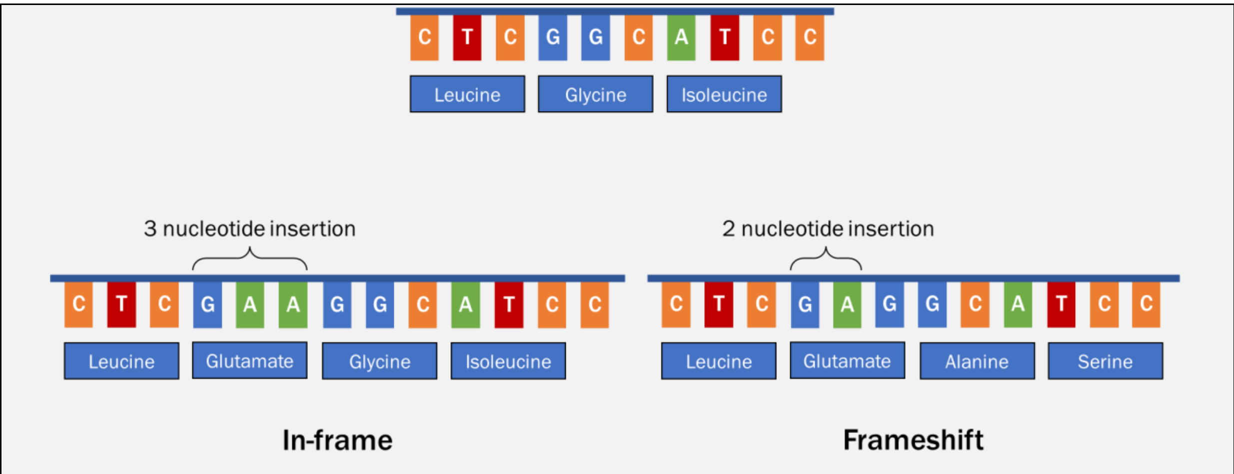
## Molecular alterations

### Variants

- Variants are alterations in the sequence of DNA
- Variants are broadly divided into small variants, structural variants and copy number variants (CNVs)

#### Small variants

- Small variants are also known as point mutations
- They affect one or only a few nucleotides
- Substitutions:
  - One nucleotide is changed to another
  - Can be transitions (a purine exchanged for a purine or a pyrimidine for a pyrimidine) or transversions (a purine exchanged for a pyrimidine, or vice versa)
  - Because the genetic code is degenerate, a change in the nucleotide may not necessarily result in a change in the amino acid (a synonymous mutation); if the change does result in a change in the amino acid, it is a non-synonymous variant
  - If a non-synonymous variant results in a different amino acid, it is considered missense; if it introduces and early stop codon, it is considered nonsense
- Insertions and deletions can collectively be described as indels:
  - Deletions: where up to a few nucleotides are lost
  - Insertions: where up to a few nucleotides are added
- Indels may be in-frame or may cause a frameshift:
  - If the change in the number of nucleotides is a multiple of three, the sequence after the affected codon (and therefore the amino acid sequence) will be unaffected
  - If the change is not a multiple of three, there will likely be a significant effect on the downstream sequence (i.e. a frameshift)
  - Variants resulting in a frameshift are more likely to have a functional impact

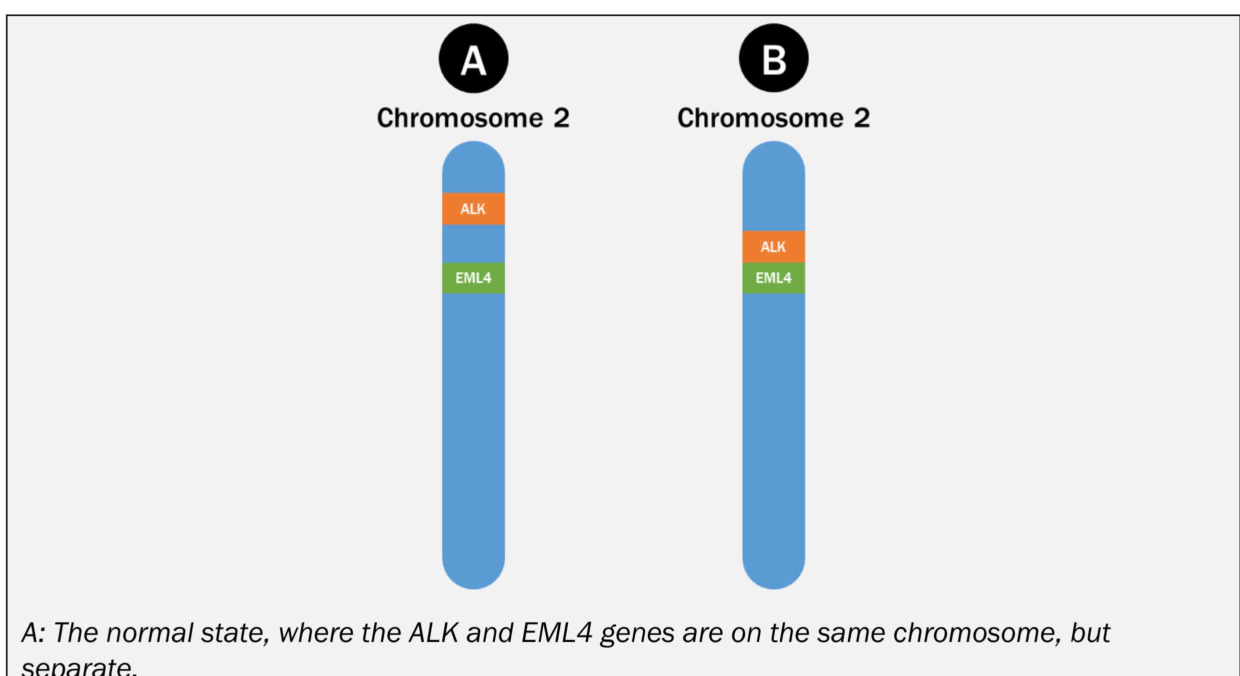



*In an in-frame insertion, the number of inserted nucleotides is a multiple of three. In this example, the insertion has added a glutamate to the amino acid sequence, but because the insertion is in-frame, all subsequent amino acids remain unchanged. The same principle applies for deletions.*

*In a frameshift insertion, the number of inserted nucleotides is not a multiple of three. This changes the reading frame, such that the amino acids after the insertion may well be changed. The same principle applies for deletions. Variants which result in frameshifts often have more dramatic biological effects than those which are in-frame.*

### Structural variants

- These affect large-scale chromosome structure
- Also known as rearrangements
- Also includes large insertions/deletions
- Inversion: when part of a chromosome flips around
- Translocations: where genetic material swaps between two separate locations in the genome
- Structural variants may bring genes which are normally separate into contact with each other – this is a gene fusion




*A: The normal state, where the ALK and EML4 genes are on the same chromosome, but separate.*

*B: A translocation has brought ALK and EML4 next to each other. This has resulted in an EML4::ALK* fusion.

### Copy number variants (CNVs)

- When there are abnormal numbers of copies of a stretch of DNA:
  - Normal somatic cells usually contain two copies of each stretch of DNA on the autosomal chromosomes (including each gene)
- Gain: when there are more copies than normal
- Loss: when there are fewer copies than usual
- There is a distinction based on the mechanism of gain:
  - Amplification: where the increase in copy number results from an increase in the number of copies of the gene itself
  - Polysomy: where the increase in copy number results from an increase in the number of copies of the chromosome on which the gene is located
- The distinction between amplification and polysomy may or may not be relevant clinically, depending on the setting

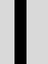
### Methylation

- See DNA methylation, above
- Methylation of a gene's promoter region reduces expression of the gene, and therefore reduces its function

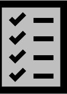
### Variant allele frequency (VAF)

- Every normal somatic cell has two alleles with respect to each gene on autosomal chromosomes
- The percentage of all alleles in a sample which are variant is described as the variant allele frequency (VAF)
- The higher the VAF, the greater the proportion of alleles in the sample which are variant

## Germline and somatic variants



Demonstrates understanding of and identifies the origins and consequences of germ-line variation and somatic mutations, including DNA methylation and gene expression changes.



Describes the origins and consequences of germline variation and somatic mutations, including DNA methylation and gene expression changes.

- Germline variants have the potential to be passed down to an individual's offspring:
  - Germline variants must begin in the germ cells of an individual (or in the zygote of their offspring)
  - If that individual's offspring inherit the variant, it will be present in all somatic cells (and around half of the gametes) in that individual's body
  - That variant may then be passed on further down the generations
  - Homozygous germline variants typically have a VAF around 100%
  - Heterozygous germline variants typically have a VAF around 50%
- Somatic variants cannot usually be passed down to an individual's offspring:
  - Somatic variants are not present in germ cells
  - The variant is acquired in a cell (e.g. from the effect of ionising radiation)
  - If that cell proliferates, the variant can end up being present in a potentially large amount of tissue, but it will not be present in all cells in the individual
- Distinguishing between a somatic and germline variant in cancer:

- A variant detected in tumour may be either somatic or germline
- A high VAF raises the possibility that a variant is germline, but this is by no means definitive
- To identify definitively whether a variant is germline, non-tumour tissue should be sampled (e.g. blood, saliva, skin biopsy) and tested
- Detection of the variant in non-tumour tissue with a high VAF indicates that the variant is germline
- Circulating tumour DNA may result in the detection of a variant in blood, but this will invariably be present at low VAF

# Oncogenes and tumour suppressor genes

## Oncogenes

- Genes which have the potential to cause cancer
- They typically encode proteins which have pro-survival, pro-growth and pro-proliferation effects
- Their functions are normally tightly regulated, and they are switched on only when needed, and only in a time-limited fashion
- Oncogenes can cause cancer through:
  - Activating small variants which cause the protein to be constitutively active
  - Small variants which prevent the protein from being degraded
  - Amplifications which cause the protein to be overexpressed
  - Structural variants which cause the protein to be overexpressed and/or constitutively active
- Many are receptor tyrosine kinases (e.g. EGFR, HER2):
  - Cell surface receptors
  - Binding of ligand (e.g. growth factor) to the extracellular domain causes dimerization
  - The intracellular kinase domain is activated, resulting in autophosphorylation
  - This activates intracellular signalling cascades which typically promote cell growth and proliferation
  - Activating variants can cause the protein to be active even in the absence of ligand binding, resulting in uncontrolled cell growth
  - A random variant in an oncogene is unlikely to have an activating effect – these variants must be of very specific types at very specific locations
  - This means that the number of clinically relevant activating variants in an oncogene is usually quite small, making targeted testing methods feasible
  - Some variants prevent the protein from being degraded, resulting in it accumulating and causing the cell to be highly sensitive to activation
- Some are downstream mediators of receptor tyrosine kinases:
  - Serine/threonine-specific protein kinases, e.g. BRAF
  - GTPases, e.g. KRAS, NRAS
  - Like receptor tyrosine kinases, activating variants in the genes encoding the proteins result in constitutive activity, even without activation of the upstream receptor tyrosine kinase
- Some are transcription factors (e.g. MYC):
  - They promote expression of genes promoting cell growth and proliferation
  - Amplification of the gene encoding the transcription factor causes overexpression of pro-growth/survival proteins
  - Structural variants which bring the gene encoding the transcription factor under the control of a constitutively expressed gene result in overexpression of the transcription factor, with the same effect

### Tumour suppressor genes

- Genes whose natural function prevents cancer
- They typically encode proteins which:
  - Promote DNA repair
  - Halt cell cycle progression
  - Induce apoptosis
- Inactivating variants in tumour suppressor genes can cause cancer:
  - Loss-of-function variants which impair the function of the protein
  - Whereas activating variants in oncogenes occurs only at a small number of stereotyped locations, inactivating variants in tumour suppressor genes can occur in a much wider range of locations in the gene
  - E.g. inactivating variants in *BRCA1*/*BRCA2* or *TP53*
- Methylation of the promoter region of a tumour suppressor gene can cause cancer:
  - Methylation prevents transcription from being initiated, resulting in reduced expression of the protein
  - E.g. methylation of the *MLH1* promoter region

## Driver and passenger variants

- Tumour cells are heavily mutated
- They acquire large numbers of variants because of exposure to carcinogens and impaired DNA repair mechanisms
- Driver variants:
  - Variants which give the 'clone' (i.e. a subset of cells with that variant) a selective advantage over other clones
  - This causes clonal expansion: the cells with the variant proliferate and come to represent a greater proportion of all tumour cells present
  - Targeting driver variants may be an effective cancer treatment, since they are 'driving' cancer growth
  - Tumours which are reliant on a single activated oncogene are described as being 'oncogene addicted' – targeting of the oncogene may very well be an effective treatment
- Passenger variants:
  - Randomly-acquired variants which do not confer a survival advantage to the affected cell
  - They are unlikely to see clonal expansion
  - Targeting passenger variants is unlikely to be a clinically useful strategy
  - Passenger variants often have low VAF
- Generally, passenger variants are commoner than driver variants

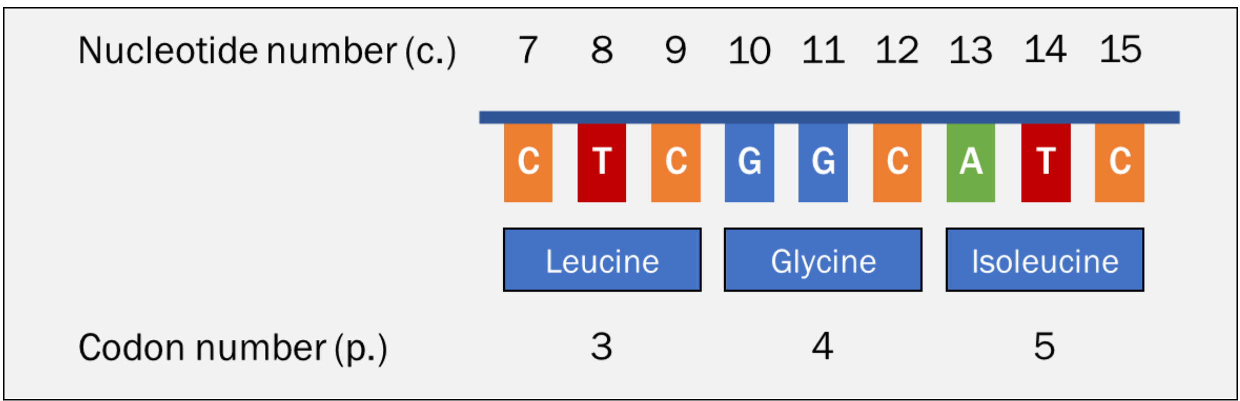
## Cancer and the immune system

- Cancer cells can incite an anti-tumour immune response:
  - Variants in the genome result in the expression of abnormal proteins (neoantigens) which can be recognised as foreign by the immune system
  - Cancer cells are generally heavily mutated
  - They therefore express large numbers of neoantigens
  - The immune system is therefore primed to attack cancer cells
- Cancer cells are able to evade immune-mediated destruction by expressing immune checkpoints:
  - Immune checkpoints expressed on tumour cells bind to immune checkpoints expressed on immune cells (e.g. cytotoxic T-cells)
  - This triggers anergy or apoptosis of the immune cells

- Expression of immune checkpoints therefore allows tumour cells to evade immune-mediated destruction

## Nomenclature

- Gene names are customarily written in italics, whereas the encoded protein name is not (e.g. the *KRAS* gene encodes the KRAS protein)
- The position of a particular nucleotide can be expressed in two ways:
  - In terms of nucleotide number, which is simply the count of nucleotides from the start codon of that gene
  - In terms of codon number, which is the set of three continuous nucleotides counted from the beginning of the sequence of interest in which the nucleotide falls (so base number 7 would be codon number 3)
- In general, mutations should be describe in terms of the nucleotide change (prefixed with “c.” for coding DNA)
- They can also be described in terms of the amino acid change (prefixed with “p.” for protein):
  - Amino acids can be described using their full names (e.g. alanine), a three-letter abbreviation (e.g. Ala) or a single-letter abbreviation (e.g. A)




### Small variants

- Substitutions: c.50G>A means that the normal guanine at nucleotide position 50 has been replaced by adenine
- Insertions: c.50_51insA means that an adenine has been inserted between nucleotide positions 50 and 51
- Deletions: c.50delG means that the normal guanine at nucleotide position 50 has been deleted
- The variants can also be expressed in terms of the amino acid change:
  - p.(Lys50Cys) means that lysine in codon 50 has been replaced by cysteine
  - p.(K50C) or simply K50C is a shorthand way of saying the same thing, which is often favoured by clinicians
- Rather than resulting in a change in amino acid, some variants introduce stop codons (nonsense variants):
  - This can be expressed as “Ter” or “*”
  - p.(Asp2Ter) or p.(Asp2*) mean that the aspartic acid in codon 2 has been replaced by a stop codon
- Some techniques will group together similar variants and therefore will not identify the precise amino acid change:
  - EGFR p.(Gly719Xaa) or p.(G719X) means that the glycine at codon 719 has been replaced by a different amino acid, but the technique does not identify the exactly amino acid

- This is usually where multiple similar variants with the same clinical implications are grouped together in real-time PCR

## Structural rearrangements

- Translocations can be expressed in terms of the gene fusion produced:
  - *CD74*::*ROS1* means that there is a translocation which has resulted in a fusion between the *ROS1* and *CD74* genes
  - More complex terminology may be used: *TMPRSS2*[NM_005656.3](Ex10)::*RSPO2*[NM_178565.4](Ex3) means that a fusion has occurred between exon 10 of *TPMRSS2* and exon 3 of *RSPO2* (the 'NM_' numbers refer to the reference sequence for each of the genes)
- Translocations can be expressed in terms of the chromosomes involved:
  - t(9;22) means that there is a translocation which has resulted in a fusion between a gene on chromosome 9 and a gene on chromosome 22
  - More complex terminology may be used: t(X;4)(p21.2;q35)(c.857+101_857+102) means that the translocation involves the short arm (p) of chromosome X and the long arm (q) of chromosome 4, with the coordinates of the exact breakpoint

# Molecular pathology

## Definitions

### Molecular pathology

- The use of molecular markers to study disease
- Applies to all pathology disciplines
- In histopathology, generally used to refer to the use of molecular markers to personalise cancer management:
  - Sometimes used only to refer to nucleic acid-based testing
  - Sometimes used more broadly to include protein-based testing
- It supplements and enriches the information provided by conventional histopathological examination

### Genomics

- Genomics is the study of the genome, which is the sum of all an organism's genetic information
- From a practical clinical point of view, genomics is the comprehensive examination of genetic information, often by sequencing
- It strictly does not look at proteins

## The purpose of molecular pathology

- In general, molecular markers serve three purposes
- There is considerable overlap, with many markers having roles in more than one category

### Diagnostic markers

- These markers help to arrive at a diagnosis
- Most provide additional information which helps to push the balance in favour of a particular diagnosis:
  - E.g. detection of an *EWSR1* rearrangement in a small round blue cell tumour would push towards a diagnosis of Ewing sarcoma, but failure to detect an *EWSR1* rearrangement would not exclude the possibility
  - E.g. detection of TCR clonality in a T-cell infiltrate would provide evidence of T-cell lymphoma, but TCR clonality can be seen in reactive lymphoid proliferations (and T-cell lymphomas do not necessarily always show clonality)
  - They are used in much the same way as diagnostic immunohistochemistry – they are neither entirely necessary nor entirely sufficient for the diagnosis in question
  - These markers should not override a histological diagnosis without good reason
  - They are particularly common in soft tissue neoplasms and lymphoid neoplasms
- Increasing numbers of molecularly-defined diagnoses are arising:
  - In order to make the diagnosis, a particular molecular marker must be detected
  - This is because the presence of the molecular marker is so strongly associated with a particular prognosis and treatment response that it is given its own diagnostic category in the WHO classification
  - E.g. glioblastoma is divided into '*IDH*-mutant' and '*IDH*-wildtype' subtypes based on *IDH* status (tumours with *IDH* variants being associated with much better survival)
  - E.g. there is a category of renal cell carcinomas known as MiT family translocation renal cell carcinomas, which are characterised by the presence of translocations involving the *TFE3* and *TFEB* genes

  - These tests should only be used in appropriate histological settings (e.g. detection of a *TFE3* translocation does not necessarily make a carcinoma of unknown primary an MiT family translocation renal cell carcinoma)
- Some markers help to diagnose underlying familial cancer syndromes:
  - As a general rule, testing of tumour tissue alone cannot definitively allow a diagnosis of an underlying familial cancer syndrome to be made – testing of non-tumour tissue (e.g. blood) is required
  - However, identification of alterations in the tumour tissue can give an indication of the likelihood of a familial cancer syndrome
  - E.g. identification of loss of MSH2 and MSH6 on MMR immunohistochemistry in a colorectal cancer indicates a high likelihood that the patient has underlying Lynch syndrome
  - E.g. identification of a pathogenic *BRCA1* variant in an ovarian carcinoma raises the possibility of a germline *BRCA1* variant, which requires testing of non-tumour tissue for confirmation

## Prognostic markers

- These markers provide extra independent information about the aggressiveness of the tumour, and therefore the patient's prognosis
- In order to be useful, they must provide extra prognostic information beyond that provided by routine histological assessment
- Although many molecular markers carry some prognostic information, it is usually not clinically useful because there are so many confounding factors
- In some cases, detection of a prognostic marker could affect the oncologist's decision to give adjuvant therapy
- E.g. detection of a high recurrence risk on Oncotype DX demonstrates a high risk of recurrence following surgery in breast cancers, and will usually prompt the oncologist to offer adjuvant chemotherapy – for patients with low risk of recurrence, it is likely that the risks of adjuvant chemotherapy would outweigh the benefits
- E.g. detection of an MMR defect in a resected colorectal cancer which is at borderline risk of recurrence will usually prompt the oncologist to offer adjuvant chemotherapy

## Predictive markers

- These markers predict which treatments to which the tumour will likely respond

### Markers predicting response to targeted therapy

- By targeting particular molecular alterations in tumour cells, these drugs have a potent effect on tumour cells, while leaving non-tumour cells fairly unscathed
- There are:
  - Very effective, often with very rapid and dramatic effects
  - Associated with a low rate of side effects, which are usually mild (very low rates of neutropenia)
  - Often take orally
  - Often suitable for even very ill patients who would not be fit for any other treatment
  - Expensive, but often cost-effective when taking into account the reduced time patients spend in hospital
- They include:
  - Tyrosine kinase inhibitors
  - Serine/threonine kinase inhibitors
  - GTPase inhibitors
  - Monoclonal antibodies

  - Inhibitors of DNA repair mechanisms, e.g. PARP inhibitors
- These drugs generally do not cure cancers, but instead keep them under control
- Tumours eventually become resistant (usually after a year or two):
  - Mechanisms may be on-target: tumour cells can develop secondary mutations in the same target gene, or can develop other alterations in the same target gene (e.g. amplification)
  - Mechanisms may be off-target: tumour cells can develop activating alterations in other targets in the same signalling pathway, or can develop alterations in other pathways which bypass the target

## Markers predicting response to immunotherapy

- Immunotherapy drugs manipulate the immune system so that it targets tumour cells
- Immune checkpoint inhibitors are a type of immunotherapy which works by preventing immune checkpoints from working:
  - Tumour cells are unable to suppress tumour-directed immune responses
  - The immune system attacks the tumour
- Immune checkpoint inhibitors generally only keep tumours under control, but there are reports of them occasionally (apparently) curing metastatic cancer
- Immune checkpoint inhibitors are:
  - Better tolerated than chemotherapy, but not as well tolerated as targeted therapy
  - Associated with immune-related side effects (e.g. hypothyroidism, colitis, hepatitis, pneumonitis), but these are usually not life-threatening
  - Given as IV infusions, often alongside chemotherapy
- There is some evidence to suggest that treatment can be stopped after a couple of years, and that the effect continues to persist
- There are various markers which predict response to immune checkpoint inhibitors:
  - Some tumours respond so reliably that no testing is required (e.g. Hodgkin lymphoma, renal cell carcinoma)
  - PD-L1 expression by immunohistochemistry (e.g. NSCLC, head and neck squamous cell carcinoma, urothelial carcinoma, triple negative breast cancer, oesophageal carcinoma)
  - MMR defects (e.g. colorectal cancer, endometrial cancer, upper gastrointestinal tract cancers, pancreatobiliary cancers)
  - Tumour mutation burden (TMB), which is a measure of how mutated the tumour genome is

## Markers predicting response to chemotherapy

- There are only a few markers which predict how well a tumour is likely to respond to types of chemotherapy
  - E.g. MMR defects in colorectal cancer predict lack of benefit from chemotherapy regimens containing 5-fluorouracil
  - E.g. *MGMT* promoter methylation in glioblastoma predicts benefit from temozolomide, compared to glioblastoma without *MGMT* promoter methylation
- Germline *DPYD* variant testing is used to predict adverse reactions to 5-fluorouracile and capecitabine
  - Patients may receive reduced doses or alternative regimens
  - This testing is usually performed on blood

Matthew Evans (matthew.evans7@nhs.net)
11/04/2024

# Molecular pathology techniques

Understands principles of common molecular pathology techniques.

Understands the principles of the most up-to-date molecular methods.

Demonstrates ability to undertake the appropriate sample collection, retrieval and preparation for the common molecular tests, whether performed on extracted nucleic acid or in situ.

Demonstrates knowledge of sequencing, PCR, microarrays (DNA and RNA), in situ hybridisation and mutation detection.

Identifies molecular techniques.

Demonstrates awareness of principles, practical knowledge of sequencing, PCR, microarrays (DNA and RNA), in situ hybridisation and mutation detection.

Demonstrates practical understanding of undertaking the appropriate sample collection, retrieval and preparation for the common molecular tests, whether performed on extracted nucleic acid or in situ.

Outlines the principles and limitations of the most up-to-date molecular methods.

Demonstrates practical knowledge of sequencing, PCR, microarrays (DNA and RNA), in situ hybridisation, mutation detection.

## In situ techniques

- In situ techniques involve cutting sections from formalin-fixed paraffin-embedded tissue and directly visualising molecules of interest in the tissue section

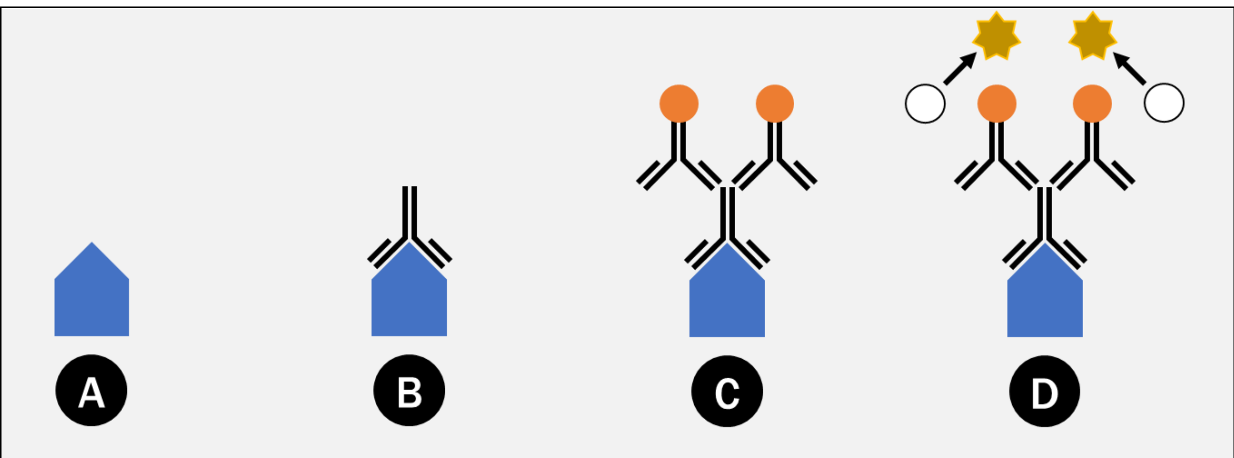
### Pre-analytics

- In situ techniques are generally validated for use only in formalin-fixed paraffin-embedded tissue
- Prolonged ischaemic time results in protein and nucleic acid degradation and may negatively impact the results of testing (see nucleic acid techniques, below)
- Particularly for immunohistochemistry, inadequate fixation reduces the intensity of staining which can impact on results:
  - Long duration of fixation is generally less of a problem
  - Results are often poorer with resections than with biopsies because of poorer formalin permeation
  - See nucleic acid techniques, below
- In situ techniques are generally not validated for use in decalcified tissue and decalcification may reduce the intensity of staining
- Proteins and nucleic acids degrade slowly in paraffin blocks, and more rapidly in cut sections:
  - Negative results from old blocks should be treated with caution
  - Cut sections should only be used if there is no alternative
- Whether cytology samples can be used depends on their preparation:
  - Cytology preparations themselves generally should not be used, since the techniques have not been validated

- If the sample has been processed in formalin as a cell block/clot, it should be acceptable to use it as if it were a biopsy
- Manufacturers' instructions with regards to section thickness should be followed:
  - Use of thin sections may result in less intense staining, which may impact on results
  - This is another reason to avoid using pre-cut sections

## Immunohistochemistry

- IHC is a means of detecting expression of antigens in tissue sections:
  - A primary antibody which is complementary to the antigen is added
  - A peroxidase-bound secondary antibody which is complementary to the primary antibody is added
  - A colourless substrate is added, which is converted to a coloured product by peroxidase
  - Counterstain is used to highlight background architecture
- Use of secondary antibodies increases the signal and therefore the sensitivity of immunostaining
- Staining is almost always automated nowadays
- Different clones of the same antibody bind different epitopes within the same target and may produce different results
- IHC can be used:
  - To test for intrinsic protein biomarkers (e.g. PD-L1)
  - To test for protein proxy markers of underlying small variants (e.g. *BRAF* V600E mutation)
  - To test for protein proxy markers of underlying structural rearrangements (e.g. *ALK* fusions)
  - To test for protein proxy markers of underlying copy number variants (e.g. *HER2* amplification)
- The reliability of protein proxy markers depends on the nature of the abnormality being assessed
  - ALK IHC is extremely reliable for detecting underlying *ALK* fusions
  - ROS1 IHC has high sensitivity but low specificity for detecting *ROS1* fusions
  - HER2 IHC is reliably for detecting clear lack or presence of *HER2* amplification, but is unreliable for intermediate levels of amplification




*A: The target of interest contains multiple epitopes against which antibodies can bind.*

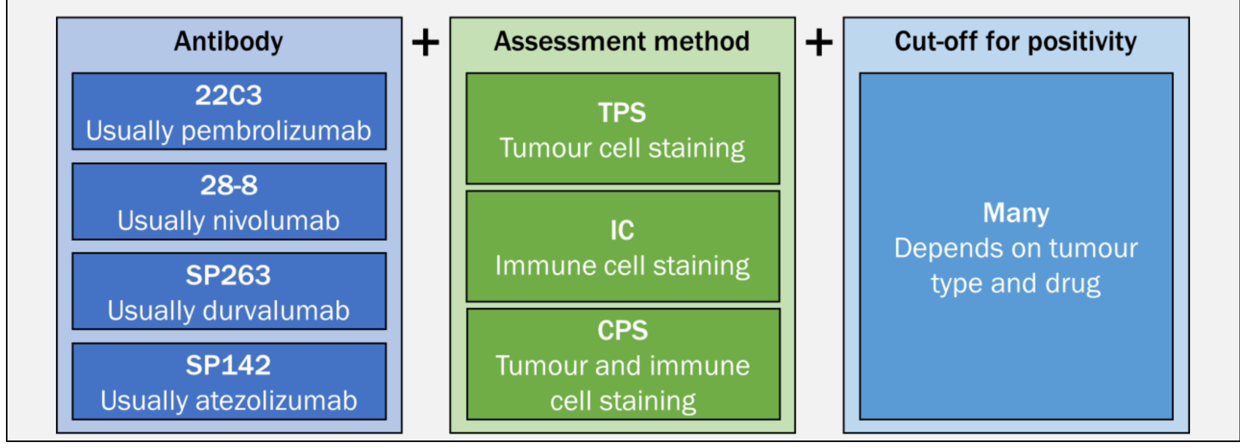
*B: A primary antibody which is complementary to an epitope on the target of interest is added.*

*C: A secondary antibody which is complementary to the primary antibody is added. The purpose is to amplify the signal, since multiple secondary antibodies bind to a single primary antibody. The secondary antibody has bound to it an enzyme (often horseradish peroxidase).*

*D: A colourless substance (often biotin) is added to the section which, when acted upon by the enzyme, is converted to a coloured product. In this way, the presence of the antigen in the section is demonstrated by the presence of the coloured product.*

## PD-L1 immunohistochemistry

- PD-L1 IHC is used to predict response to immune checkpoint inhibitors in certain situations:
  - For some tumour types and some drugs, PD-L1 IHC is mandatory prior to prescription
  - For some tumour types and some drugs, PD-L1 IHC is advisable but not mandatory prior to prescription
  - For some tumour types and some drugs, an alternative marker is required prior to prescription (e.g. MMR IHC)
  - For some tumour types and some drugs, no marker is required prior to prescription
- 'PD-L1' is not a single test:
  - There are multiple commercial PD-L1 antibodies which produce different staining
  - There are multiple methods of assessing PD-L1 staining, which involve assessing tumour cells, inflammatory cells or a combination of both
  - There are multiple definitions of positivity, which depend on the tumour type and drug to be prescribed



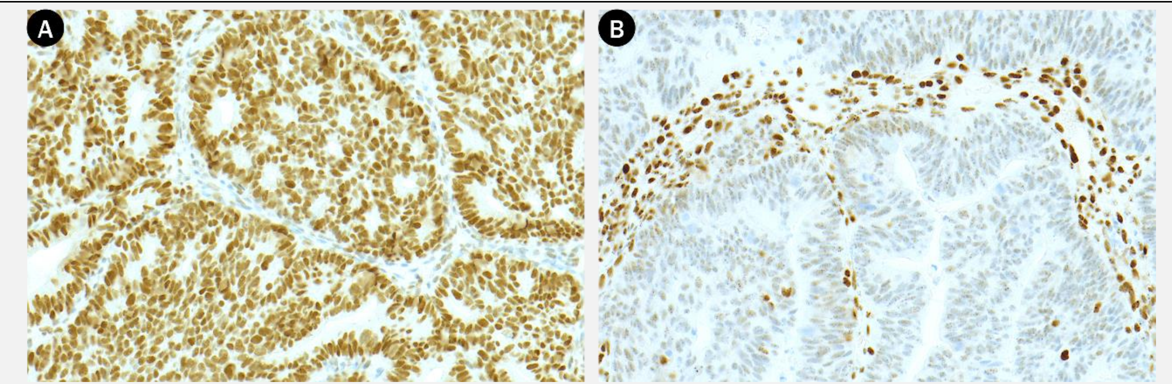

- The PD-L1 staining will be scored based on the extent of staining
- When PD-L1 is requested, therefore, it is essential to know the tumour type and the drug which is to be prescribed
- CPS and IC scoring can only be performed on solid tissue – cytological samples (even as cell blocks/clots) cannot be used

## MMR immunohistochemistry

- MMR IHC is used as a screening tool to infer the presence of molecular alterations involving the MMR system
- The expression of four proteins – MLH1, PMS2, MSH2 and MSH6 – is used
- Therefore, MMR IHC requests will result in four slides – one for each of the proteins
- Under normal circumstances, both neoplastic and non-neoplastic cells will show preserved nuclear expression for all four markers:

- Staining in the tumour cell nuclei should be as strong or stronger as the staining seen in internal controls
- Internal controls are background non-neoplastic cells (e.g. inflammatory cells, stromal cells, benign epithelial cells)

- Loss of staining is when staining in tumour cell nuclei is significantly weaker than is seen in immediately adjacent internal controls
    - Staining is usually completely lost
    - Sometimes it may just be significantly weaker
    - It is important to compare staining to that seen in immediately adjacent internal controls to ensure that the lack of staining is not simply artefactual – technical artefact would result in loss of staining in both neoplastic cells and internal controls
- Complete loss of staining in both neoplastic and non-neoplastic cells almost always indicates a technical failure, but very rare familial cancer syndromes can result in this staining (usually in paediatric cases)

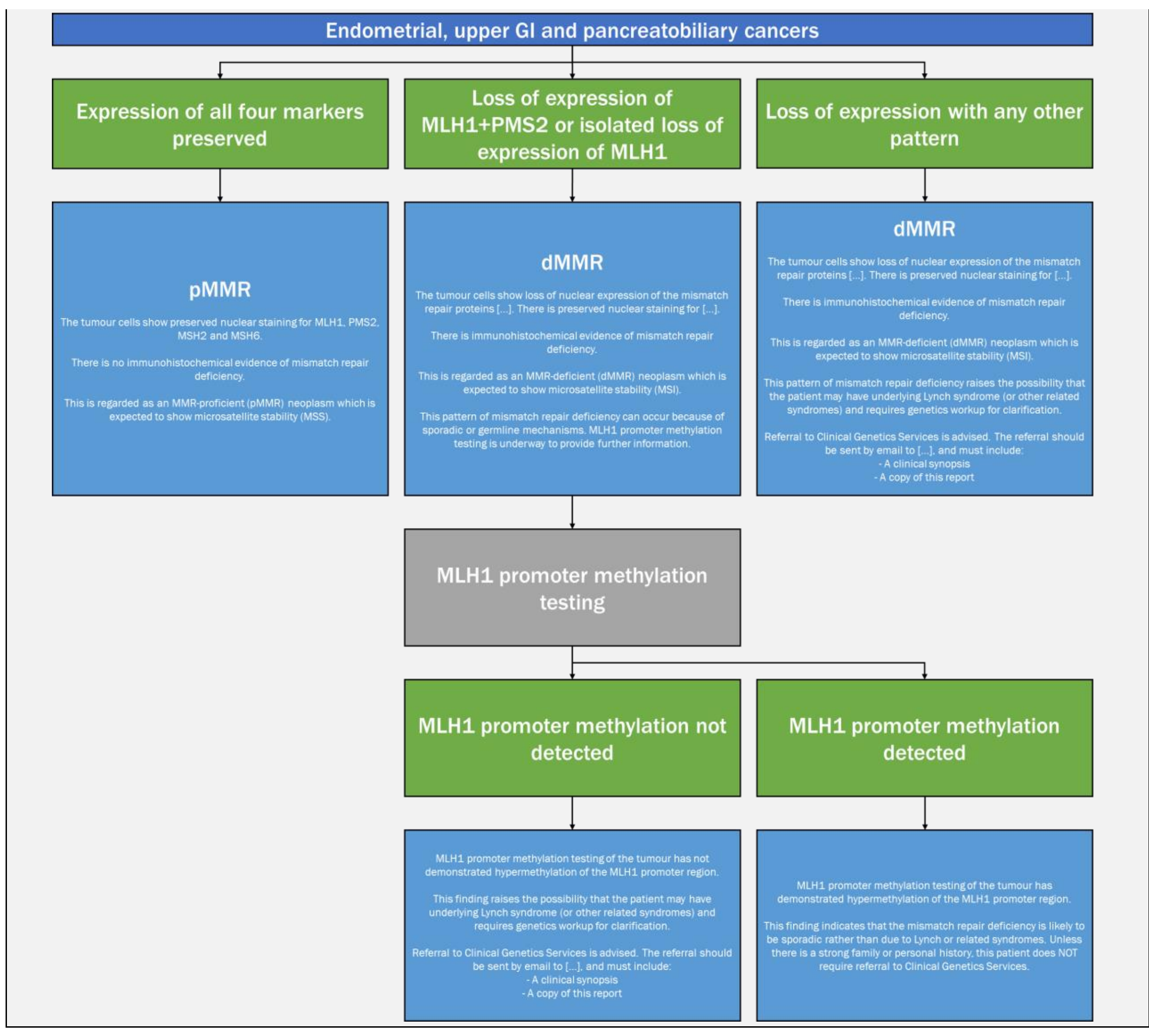



*A: The tumour cells show staining which is much stronger than the internal controls (the intervening inflammatory and stromal cells). Therefore, there is preserved expression for this marker. Assuming that all four markers show preserved expression, the tumour would be considered MMR-proficient (pMMR).*

*B: The tumour cells show faint nuclear staining in places. However, this is significantly weaker than the staining which is seen in the internal controls. Although there is some tumour cell nuclear staining, therefore, there is loss of nuclear expression of this marker. Irrespective of the other three markers, this tumour is considered MMR-deficient (dMMR).*

- If there is preserved normal staining with all four markers, the tumour is considered MMR-proficient (pMMR)
- If there is loss of expression of any of the markers, the tumour is considered MMR-deficient (dMMR)
- The pattern of loss of expression gives an idea of likelihood of underlying Lynch (or related) syndromes

Colorectal cancer
Expression of all four markers preserved
Loss of expression of MLH1+PMS2 or isolated loss of expression of MLH1
Loss of expression with any other pattern
pMMR
The tumour cells show preserved nuclear staining for MLH1, PMS2, MSH2 and MSH6.
There is no immunohistochemical evidence of mismatch repair deficiency.
This is regarded as an MMR-proficient (pMMR) neoplasm which is expected to show microsatellite stability (MSS).
dMMR
The tumour cells show loss of nuclear expression of the mismatch repair proteins [...]. There is preserved nuclear staining for [...].
There is immunohistochemical evidence of mismatch repair deficiency.
This is regarded as an MMR-deficient (dMMR) neoplasm which is expected to show microsatellite stability (MSI).
This pattern of mismatch repair deficiency can occur because of sporadic or germline mechanisms. BRAF mutation testing +/- MLH1 promoter methylation testing is underway to provide further information.
dMMR
The tumour cells show loss of nuclear expression of the mismatch repair proteins [...]. There is preserved nuclear staining for [...].
There is immunohistochemical evidence of mismatch repair deficiency.
This is regarded as an MMR-deficient (dMMR) neoplasm which is expected to show microsatellite stability (MSI).
This pattern of mismatch repair deficiency raises the possibility that the patient may have underlying Lynch syndrome (or other related syndromes) and requires genetics workup for clarification.
Referral to Clinical Genetics Services is advised. The referral should be sent by email to [...], and must include:
- A clinical synopsis
- A copy of this report
BRAF mutation testing
BRAF V600E mutation not detected
BRAF V600E mutation present
MLH1 promoter methylation testing
BRAF mutation testing of the tumour has identified a V600E mutation.
This finding indicates that the mismatch repair deficiency is likely to be sporadic rather than due to Lynch or related syndromes. Unless there is a strong family or personal history, this patient does NOT require referral to Clinical Genetics Services.
MLH1 promoter methylation not detected
MLH1 promoter methylation detected
BRAF mutation testing of the tumour has not identified a V600E mutation, and MLH1 promoter methylation testing of the tumour has not demonstrated hypermethylation of the MLH1 promoter region.
This finding raises the possibility that the patient may have underlying Lynch syndrome (or other related syndromes) and requires genetics workup for clarification.
Referral to Clinical Genetics Services is advised. The referral should be sent by email to [...], and must include:
- A clinical synopsis
- A copy of this report
MLH1 promoter methylation testing of the tumour has demonstrated hypermethylation of the MLH1 promoter region.
This finding indicates that the mismatch repair deficiency is likely to be sporadic rather than due to Lynch or related syndromes. Unless there is a strong family or personal history, this patient does NOT require referral to Clinical Genetics Services.

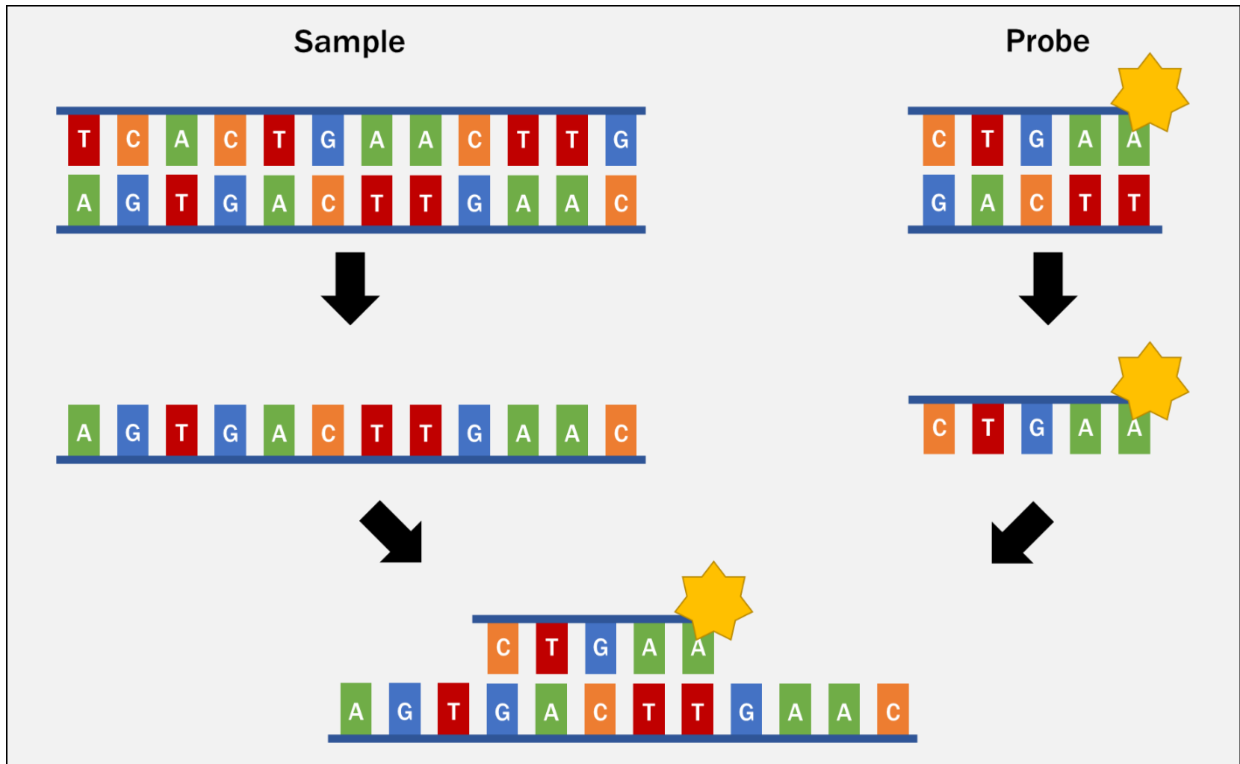



- Microsatellite analysis (MSI) may be used as an alternative method for assessing the integrity of the MMR system
- Across colorectal and endometrial carcinoma, MMR IHC and MSI testing show very high concordance
- Reasons to favour MMR IHC include:
  - Those working in a histopathology department generally have much easier access to IHC than to MSI testing
  - It is unaffected by low neoplastic cell percentage
  - A reliable result can often be obtained even with very few cells
  - It is less affected by over-fixation
  - The pattern of loss gives a sense of the likelihood of underlying Lynch syndrome, and avoids the necessity for some (*BRAF* variant and) *MLH1* promoter methylation tests
- Reasons to favour MSI include:
  - Those working in a genomics lab usually have no access to IHC
  - It is less affected by under-fixation (particularly a problem with resections)
- See microsatellite analysis, below

## In situ hybridisation

- ISH is similar to IHC, but is used to detect nucleic acid sequences rather than proteins:

  - The tissue section is permeabilised
  - The nucleic acids in the section are denatured
  - One or more probes are hybridised against the nucleic acid sequence(s) of interest
- The probe comprises a nucleic acid sequence which is complementary to the sequence of interest:
  - Attached to it is some means of visualising the probe
  - This can be fluorescent, coloured (i.e. a chromogen, like those used in IHC) or silver (appearing black)
- Fluorescence in situ hybridisation (FISH):
  - Uses fluorescently labelled probes
  - Requires a fluorescence microscope for interpretation which is expansive
  - Needs to be undertaken in a dark room
  - Fluorescence fades fairly rapidly, even if slides are stored in the dark
- Chromogenic in situ hybridisation (CISH) and silver-enhanced in situ hybridisation (SISH):
  - These are often described as brightfield techniques, since they use light rather than fluorescence microscopy
  - CISH uses probes which can be detected through colour (as in IHC)
  - SISH uses probes which carry silver, and so appear black
  - DDISH (dual-chromogen/dual-hapten ISH) involves using two methods (e.g. one CISH and one SISH probe)
  - The main advantage is that this does not require expensive fluorescence microscopy
  - The slides fade only very slowly (especially with SISH)
  - Under some circumstances, CISH/SISH may have lower sensitivity than FISH
  - There is less variety of commercial CISH/SISH probes than FISH probes available
- FISH preparation is normally undertaken manually
- CISH/SISH hybridisation is often automated on standard IHC stainers
- Some IHC stainers can automate FISH preparation, but this is still fairly uncommon



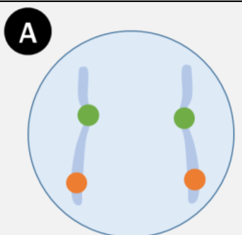
*ISH involves using a probe which is complementary to the nucleic acid sequence of interest, and which is labelled (e.g. with a fluorophore). Both the sample DNA on the slide and the probe are denatured to separate them into single strands. The probe is then hybridised against the sample. If the sequence of interest is present, it will be labelled (e.g. by fluorescence).*

### Using in situ hybridisation to detect the presence of nucleic acids

- ISH can be used simply to determine whether particular nucleic acids are present or not
- E.g. high-risk HPV ISH:
  - Probes against viral DNA or RNA sequences
  - Can be used to detect infection by high-risk HPV
  - Can be used as an alternative to high-risk HPV PCR
- E.g. immunoglobulin ISH:
  - Kappa and lambda light chain IHC can be used to test for light chain restriction
  - However, because light chains are present ubiquitously in body fluids, there is a large amount of background staining which makes assessment difficult
  - ISH for light chain RNA sequences highlights only cells which are actively expressing light chains – much cleaner staining
  - Disadvantage is that ISH is more expensive than IHC

### Using in situ hybridisation to assess for copy number variants

- FISH or CISH/SISH can be used to assess for gene amplification, using enumeration probes
- Generally involves the use of two probes, each differently labelled (e.g. with different fluorescence):
  - A probe which hybridises against the gene of interest
  - A reference probe which hybridises against a different area on the same chromosome
- Allows gene amplification to be distinguished from polysomy (extra copies of the chromosome)
- Interpreter counts the number of signals corresponding to the gene of interest and the reference probe across a pre-determine minimum number of tumour cell nuclei
- Generally returns two values:
  - Mean copy number, which is the average absolute number of copies of the gene of interest in each cell
  - The ratio between the gene of interest and the reference signals
- The combination of these determines whether there is gene amplification
- Sometimes there is such enormous gene amplification that the total number of gene signals cannot be accurately enumerated, but there is obvious amplification
- E.g. *HER2* amplification:
  - *HER2* is the gene of interest, CEP17 is the reference probe (which hybridises against the centromeric region of the chromosome on which *HER2* is located)
  - A sufficiently raised *HER2*:CEP17 ratio constitutes amplification
  - A raised *HER2* copy number in the absence of raised *HER2*:CEP17 ratio generally means polysomy without evidence of *HER2* gene amplification
  - However, once the *HER2* copy number is sufficiently raised (irrespective of *HER2*:CEP17 ratio), *HER2* gene amplification can be considered present
- There is generally good concordance between FISH and CISH/SISH/DDISH, although FISH is generally believed to be more sensitive for detection of low levels of amplification

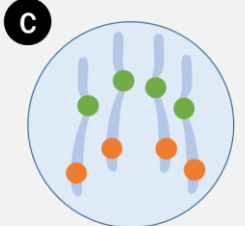


A

HER2 signals: 2
CEP17 signals: 2
HER2 copy number: 2
HER2:CEP17: 1
**Negative**

B

HER2 signals: 8
CEP17 signals: 2
HER2 copy number: 8
HER2:CEP17: 4
**Positive**



C

HER2 signals: 4
CEP17 signals: 4
HER2 copy number: 4
HER2:CEP17: 1
**Negative**

D

HER2 signals: 7
CEP17 signals: 7
HER2 copy number: 7
HER2:CEP17: 1
**Positive**

*A: This is the normal situation. The nucleus contains two* HER2 *and two CEP17 signals (representing the maternal and paternal alleles). The tumour is considered negative on grounds of both* HER2 *copy number and* HER2*:CEP17.*

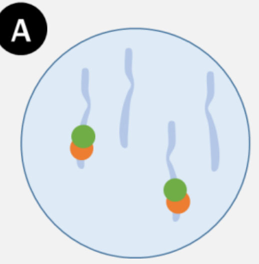
*B: Here, there are increased numbers of* HER2 *relative to CEP17 signals, consistent with* HER2 *amplification.*

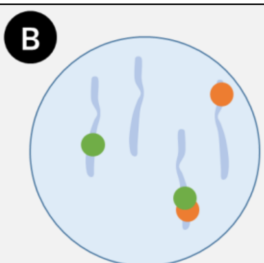
*C: The number of* HER2 *signals is increased, but in proportion with the number of CEP17 signals. This is consistent with chromosome 17 polysomy. Although the* HER2 *copy number is raised, the* HER2*:CEP17 ratio remains 1 and so this case is negative for* HER2 *amplification.*

*D: This is similar to (C). There are increased numbers of* HER2 *signals which remain in proportion to the CEP17 signals, consistent with polysomy. However, scoring in breast cancer allows for a case to be considered positive if the copy number exceeds 6, regardless of* HER2*:CEP17. This case is therefore considered positive.*

### Using in situ hybridisation to assess for gene rearrangements

- Generally only FISH is used to assess for gene rearrangements
- Dual-colour break-apart probes:
  - Used to determine whether a gene is involved in a rearrangement, irrespective of its fusion partner
  - One probe is hybridised against the sequence of interest, and a second probe is hybridised against a neighbouring sequence
  - Under normal circumstances, the two probes will hybridise next to each other so both signals (usually orange and green) will be next to each other (fused)
  - Fused signals manifest as a single yellow spot or as very close orange and green signals
  - If the gene of interest has undergone a rearrangement, the two signals will be separated (split)
  - This indicates that the gene has undergone a rearrangement, but does not identify the partner gene

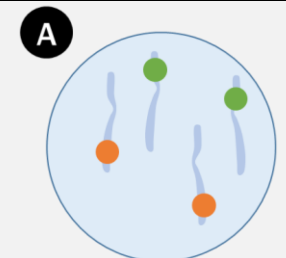


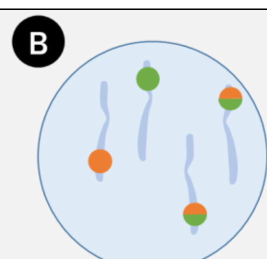




*A: For a dual-colour break-apart probe, the normal situation is to have two fused signals (either yellow, or an orange and green signal immediately next to each other).*

*B: Normal somatic cells contain two alleles of each gene; a rearrangement would only be expected to affect one allele of the gene of interest. If there is a rearrangement, there will remain one fused signal (since one allele will not undergo the rearrangement). However, there will be two split signals representing the allele which is participating in the rearrangement. Note that this only tells you that the gene is involved in the rearrangement; it gives no indication of the nature of that rearrangement (i.e. the partner gene).*

- Dual-colour dual-fusion probes:
  - Used to determine the exact nature of a rearrangement
  - One probe each is hybridised against each of the genes which are suspected to be involved in the rearrangement
  - Under normal circumstances, the signals from each of the probes should be separate, meaning that there is not a rearrangement involving the two genes
  - If both genes are involved in a rearrangement, the signals will be fused (i.e. yellow, or a green and orange signal next to each other)
  - Importantly, this will only detect a rearrangement involving both genes
  - It will miss a rearrangement involving only one of the genes




*A: One probe is hybridised against each of the genes of interest (bearing in mind that there are two copies of each gene). Normally, these genes should be remote from each other, and so there will be four split signals.*

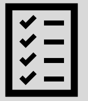
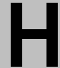
*B: If there is a rearrangement involving both genes of interest, part of the signal of one of the genes will fuse with part of the signal from the other gene. There will therefore be two fused and two split signals. Note that if only one of the genes of interest is involved in a rearrangement, a dual-colour dual-fusion probe will fail to detect a rearrangement at all.*

- In many cases, the precise gene fusion is not important:
  - For *ALK*, *ROS1* and *RET* fusions in NSCLC, all that matters is that there is rearrangement involving the gene
  - Dual-colour break-apart probes are therefore sufficient
- If the precise nature of the fusion matters, it is generally advisable to use a dual-colour break-apart probe initially:
  - This confirms that a translocation involving the gene of interest is present

  - Dual-colour dual-fusion probes can then be used to work out what the fusion partner is
  - Multiple reactions would need to be undertaken until the precise fusion is identified, taking a long time and consuming a good deal of tissue
  - It may therefore be more efficient to use reverse transcription PCR or next-generation sequencing
- FISH may give rise to false negative results:
  - Small-scale rearrangements where the gene has moved a short distance (e.g. intrachromosomal rearrangements) may be missed by FISH and give a false negative result
- FISH may give rise to ‘false positive’ results:
  - A rearrangement may genuinely be present at the DNA level, but this may not result in expression of an mRNA fusion transcript or a functional fusion protein
  - Therefore FISH will be positive but the patient will be unlikely to respond to targeted treatment
  - RNA-based methods are less likely to detect non-biologically significant rearrangements
- A particular difficulty with FISH is that because of the lack of morphology, it may be extremely difficult to distinguish between neoplastic and non-neoplastic cells

# Nucleic acid techniques

## Pre-analytics

| | | |
|---|---|---|
|  | I | Demonstrates knowledge of how histological samples are taken, prepared and of how nucleic acids are extracted from them. |
| |  | Describes how histological samples are taken and prepared, and how nucleic acids are extracted from them. |

### Tissue processing

- Following loss of blood supply, nucleic acids begin to significantly degrade after a few hours
- Release of enzymes and putrefaction break down nucleic acids
- Cooling is known to slow this process but will not prevent it
- There are three approaches to halting this:
  - Fixation
  - Immediate nucleic acid extraction
  - Freezing

#### *Fixation*

- Fixation is important for histological processing because:
  - It prevents autolysis and putrefaction, and so maintains tissue morphology
  - It increases the rigidity and stability of tissues, allowing for more accurate sectioning and assessment of morphology
  - It destroys infectious agents
- Fixation is normally by immersion in 10% neutral buffered formalin
- Formalin permeates tissue at a rate of only a few millimetres per hour:
  - Autolysis/putrefaction in a biopsy will be stopped after a few hours

  - For a large resection specimen, it may take days or weeks for full formalin permeation and therefore for autolysis/putrefaction to stop, by which time the proteins and nucleic acids will be extremely degraded
  - It is therefore extremely important that resection specimens are sliced/opened/inflated as promptly as possible
- Formalin is an extremely good option for preserving morphology and allowing for ancillary techniques like IHC
- However, formalin fixation damages nucleic acids:
  - It causes cross-links between nucleic acids and protein which, when broken, result in fragmentation of nucleic acids
  - It can cause deamination of cytosine, converting it to thymine
- RNA is generally more severely affected than DNA, so RNA-based testing techniques generally have higher failure rates than DNA techniques
- The degree of nucleic acid damage increases with the duration of fixation
- Some fixatives have less deleterious effects on nucleic acids, but:
  - They are generally more expensive and less readily available
  - They produce different microscopic appearances which make diagnosis more challenging
  - They may well jeopardise in situ techniques like IHC and ISH
- Nucleic acids gradually degrade in tissue blocks (and more rapidly in cut sections), and so testing on very old blocks may fail

#### *Immediate nucleic acid extraction*

- Tissue can be sent fresh and nucleic acid extraction immediately undertaken before autolysis/putrefaction begins
- The big advantage is that this preserves nucleic acids and avoids the need to use liquid nitrogen for freezing
- In reality, this is generally not undertaken because of the inevitable delays which happen in transporting fresh tissue to the site of nucleic acid extraction
- Where non-fixed tissue is required, it is usually sent to the pathology department fresh and then frozen before nucleic acid extraction

#### *Freezing*

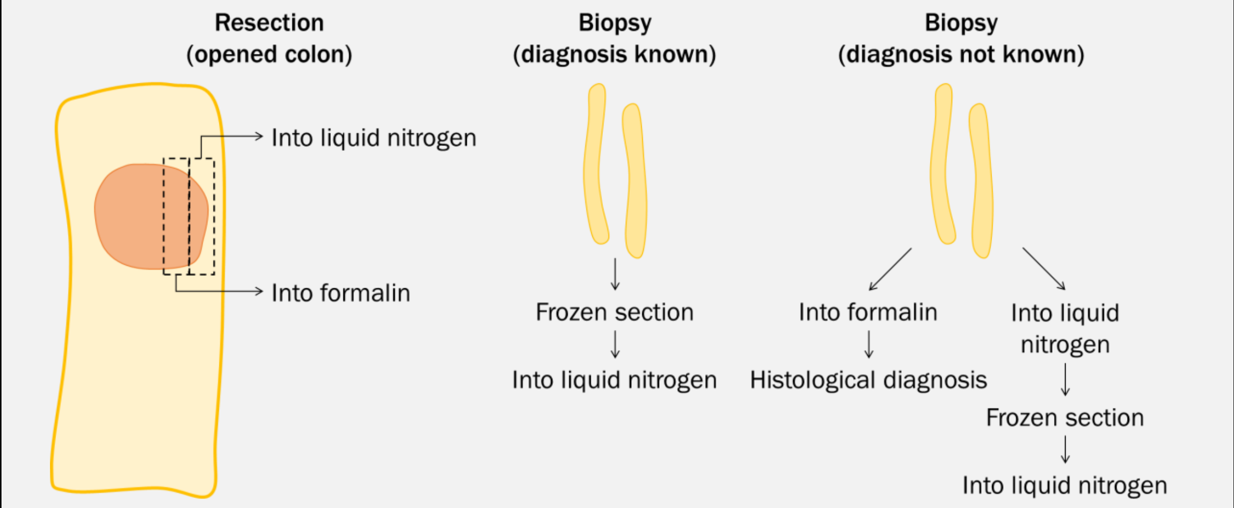
- Freezing fresh tissue in liquid nitrogen has several advantages:
  - It preserves the integrity of nucleic acids
  - It halts autolysis/putrefaction
- Logistically, however, it poses challenges:
  - The laboratory must have access to liquid nitrogen, and must be ready to place samples in it soon after being received
  - The laboratory must have an extremely low temperature freezer
  - There must be capability to transport the frozen tissue in a temperature-controlled fashion
- For a resection specimen:
  - The specimen must be promptly sent to the pathology laboratory without anybody unknowingly immersing it in formalin
  - Fresh tumour tissue must be promptly sampled from the tumour before being rapidly frozen in liquid nitrogen
  - The next slice from the sampled tumour is processed in formalin as a ‘mirror block’ and the appearances of this block used to infer the contents of the tissue sampled for nucleic acid extraction
- For a biopsy specimen where the diagnosis is already known (i.e. where further biopsies have been taken purely for molecular profiling):

- The specimen must be promptly sent to the pathology laboratory without anybody unknowingly immersing it in formalin
- Frozen section examination is undertaken on the biopsy as soon as it arrives in the laboratory
- The remainder of the tissue is rapidly frozen in liquid nitrogen
- The frozen section is used to infer the tumour cell percentage, and to confirm that the diagnosis is as expected

▪ For a biopsy specimen where the diagnosis is not already known:
  - Making a diagnosis is prioritised
  - If there are multiple biopsies, one/some can be processed in formalin and the remainder processed as above
  - It is essential that tissue is not sent for nucleic acid extraction until it is known that the formalin-fixed tissue is adequate to make the diagnosis
  - If a diagnosis cannot be made on the formalin-fixed tissue, more of the frozen tissue should be fixed in formalin in the hope that it will allow a diagnosis to be made

▪ In general, freezing of tissue is required only where whole-genome sequencing is being undertaken

▪ Use of fresh/frozen tissue certainly reduces the likelihood of failed testing due to nucleic degradation (assuming the tissue has been handled appropriately), but is generally not necessary in the vast majority of situations




*Fresh/frozen tissue sampling for a resection is fairly straightforward. The specimen is opened and a small amount of the tumour is removed to be placed in liquid nitrogen. A slice is taken from the tumour immediately adjacent to this which is placed into formalin. This block will be used to infer the contents of the tumour which was placed in nitrogen.*

*Sometimes, a patient will already have a histological diagnosis and repeat biopsy has been undertaken purely for molecular testing. On receipt, a frozen section is taken from the tissue to determine the contents of the sample. The tissue is then placed in liquid nitrogen.*

*It may be that at a patient's diagnostic biopsy, tissue is set aside for molecular testing. Here, the priority must always be the histological diagnosis. Some of the tissue is placed in formalin and a histological diagnosis is reached. One biopsy is placed into liquid nitrogen. Assuming that a malignant diagnosis is reached, this biopsy can then undergo frozen section to determine its contents and can then be used for molecular testing. This biopsy will have to be examined*

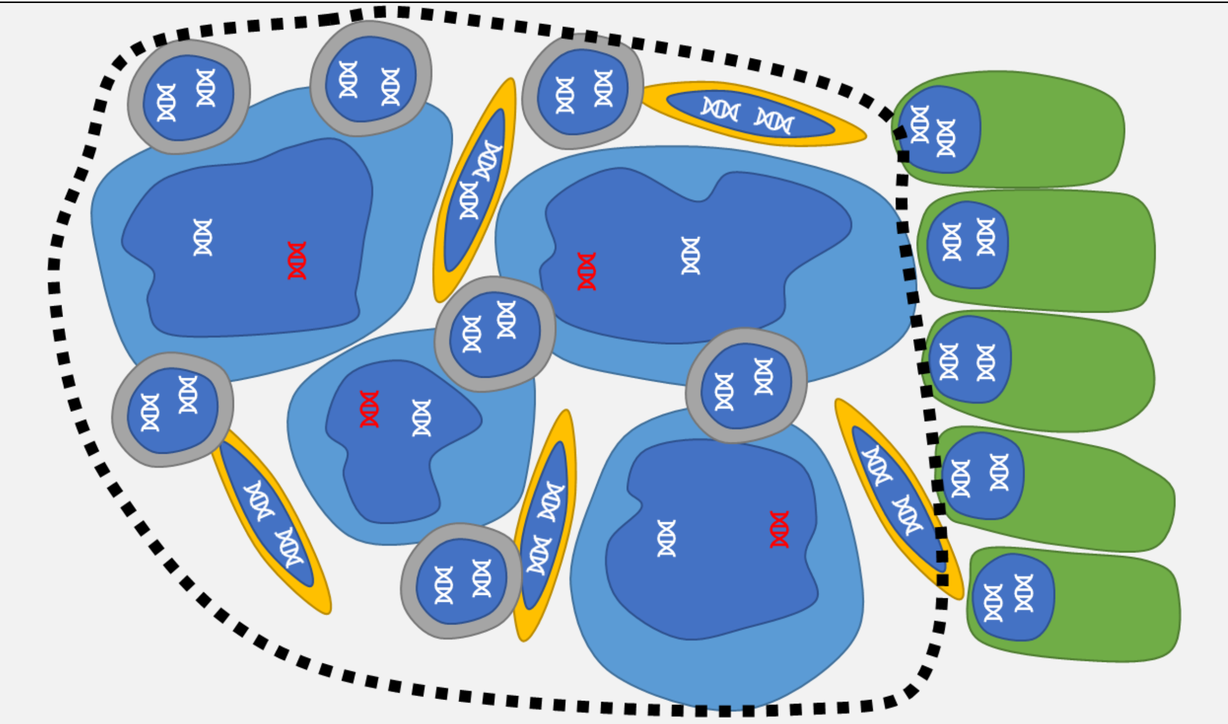
*histologically (following formalin fixation) if a diagnosis could not be reached using the rest of the tissue.*

## Tumour assessment

- What matters for nucleic acid-based molecular testing is:
  - Neoplastic cell percentage
  - Absolute number of nucleated cells (total cellularity)
  - Percentage area necrosis
- Before undertaking nucleic acid-based molecular tests, it is important that an H&E-stained section is assessed
- This gives an indication of the contents of any sections to be used for molecular testing and is essential to accurate interpretation of the results of testing

### *Tumour cell percentage*

- Limit of detection:
  - All molecular tests have a limit of detection
  - This is the minimum proportion of variant alleles which can reliably be detected in a sample containing wild-type alleles
  - E.g. if a test has a limit of detection of 10%, it may fail to detect a variant which is constitutes only 5% of alleles in a sample
  - The proportion of alleles in a sample which are a given variant of interest is the variant allelic frequency (VAF)



*This is a sample of tumour in which the tumour cells contain a variant (in red). Because each normal somatic cell contains two alleles of each gene, only half the alleles in the tumour cells will be variant (VAF = 50%). However, this tumour sample does not contain only tumour cells: it also contains stromal cells (yellow), inflammatory cells (grey) and benign epithelial cells (green). Each of these non-tumour cells carries two wild-type alleles. Once all these alleles are taken into account, the VAF is 9%. If the limit of detection of the technique being used is 10%, failure to detect a mutation from this sample could represent a false negative. However, it is possible to macrodissect out the benign epithelium (the inflammatory and stromal cells cannot be macrodissected out, since there are scattered amongst the tumour cells); if this is done, twelve*

*wild-type alleles will be removed and the VAF will increase to 13%. Assuming that our technique has a limit of detection of 10%, a negative result would be reliable.*

- Limit of detection is related to minimum neoplastic cell percentage:
  - All molecular tests have a minimum tumour cell percentage requirement
  - If the neoplastic cell percentage of a sample is below the minimum neoplastic cell percentage of the technique, failure to detect a mutation could represent a false negative result
- The presence of inflammatory cells, stromal cells and background benign epithelial cells effectively 'dilutes' any mutant alleles present in the tumour cells, and increases the likelihood of them not being detected
- If possible, areas of non-tumour cells should be removed from the sections to be used for testing by macrodissection:
  - The goal is to increase the neoplastic cell percentage to reduce the likelihood of a false negative result
  - Sometimes, it may be necessary to macrodissect out areas of tumour in order to increase the overall neoplastic cell percentage (e.g. if areas of tumour are heavily inflamed)

#### *Absolute number of nucleated cells*

- This is an estimate of the total number of nucleated cells in the tissue section
- This gives an idea of the starting amount of nucleic acid
- Samples with very small numbers of nucleated cells run the risk of failing outright (compared to samples with low neoplastic cell percentage, where the risk is not of failure, but of false negative results)
- This is generally categorised in brackets: very low (less than 700), low (700-4,000), medium (4,000-10,000), high (10,000-50,000) and very high (more than 50,000)
- As a general rule, low neoplastic cell percentage is more of a problem than low overall cellularity:
  - Modern molecular techniques generally require little starting nucleic acid
  - However, if the nucleic acid in the sample is of low quality (prolonged ischaemic time, prolonged fixation, decalcification), the overall cellularity requirement will increase

#### *Percentage area necrosis*

- This is the percentage area of the tissue which is occupied by necrosis
- Necrosis may impair testing

#### *Assessment*

- When undertaking molecular testing, a marked H&E section should be provided:
  - This should indicate the tumour area to be macrodissected out
  - If macrodissection is not possible or will not be helpful, the fact that the entire section should be used should be clearly indicated
- The following information is also required:
  - The absolute number of nucleated cells (in both the marked area and whole section)
  - The neoplastic cell percentage (in both the marked area and whole section)
  - The percentage area necrosis (in both the marked area and whole section)

*This slide has been marked to show the tumour area. An assessment of neoplastic cell percentage has been made of that area and of the entire section. The marked tumour area provides a guide for macrodissection. The request form should also state the nucleated cell content and the percentage area occupied by necrosis.*

- During macrodissection, the marked H&E-stained section is used as a guide to scrape the unwanted tissue from an unstained section:
  - It is difficult to do this accurately
  - It is therefore important not to mark areas which are small or complex – doing so runs the risk of the high-quality areas of tissue accidentally being scraped off the slide
- Some tissues have a dramatically negative impact on neoplastic cell percentage and should be removed by macrodissection if possible:
  - Lymph nodes contain huge numbers of non-neoplastic cells, even if it is only the uninvolved rim of node around a metastasis
  - Areas of inflamed tumour contain large numbers of non-neoplastic cells
  - Cellular desmoplastic stroma contains a high density of non-neoplastic cells
- Some tissues have rather little impact on neoplastic cell percentage:
  - Because adipocytes are so large, even a large area of fat contains fairly few cells, meaning that the neoplastic cell percentage is fairly unchanged
  - Normal lung parenchyma contains largely empty space and so the number of non-neoplastic cells contributed is small (this does not apply if the parenchyma is inflamed)

## Section preparation

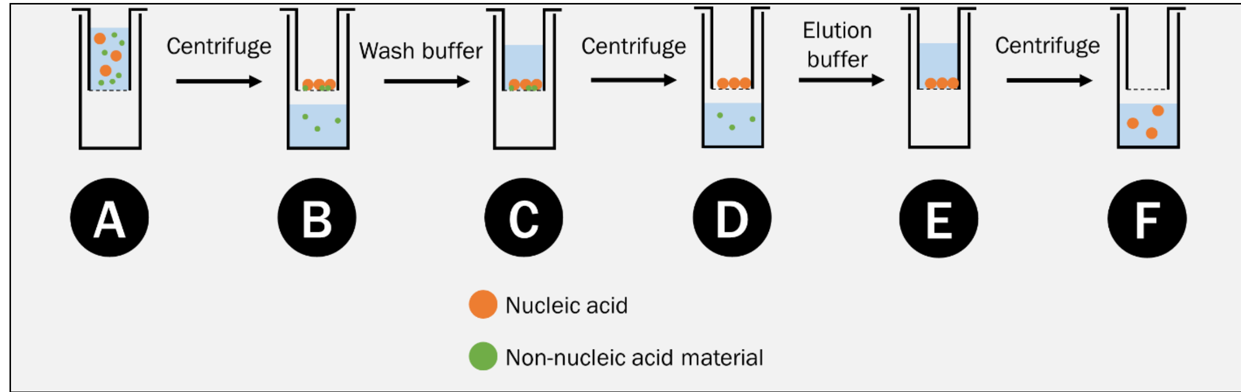
- Sections should be prepared in a sterile fashion and to the thickness determined by the technique
- Molecular techniques are validated for use with a particular section thickness:
  - Using a different section thickness may affect the extraction of nucleic acids
  - This introduces uncertainty into the validity of any results obtained
- Different techniques have different tolerances for contamination from nucleic acids from other specimens, but in general it is important to minimise the likelihood of cross-contamination:

- New, clean gloves must always be used for each specimen
- A fresh microtome blade must always be used for each specimen
- The microtome and surrounding workspace must be decontaminated with alcohol/bleach between specimens
- Instruments (e.g. forceps) must be thoroughly cleaned between specimens
- Ideally, a dedicated microtome in a clearly-delineated area of the laboratory should be used to prepare sections for molecular testing
- Paraffin curls should be placed in sterilised tubes
- Sections should be mounted on sterilised slides and stored in sterilised slide holders
- A water bath should be avoided by using curls wherever possible
- If slide-mounted sections must be used, the water bath should be emptied and cleaned between cases

▪ If a case requires macrodissection, slide-mounted sections must be prepared

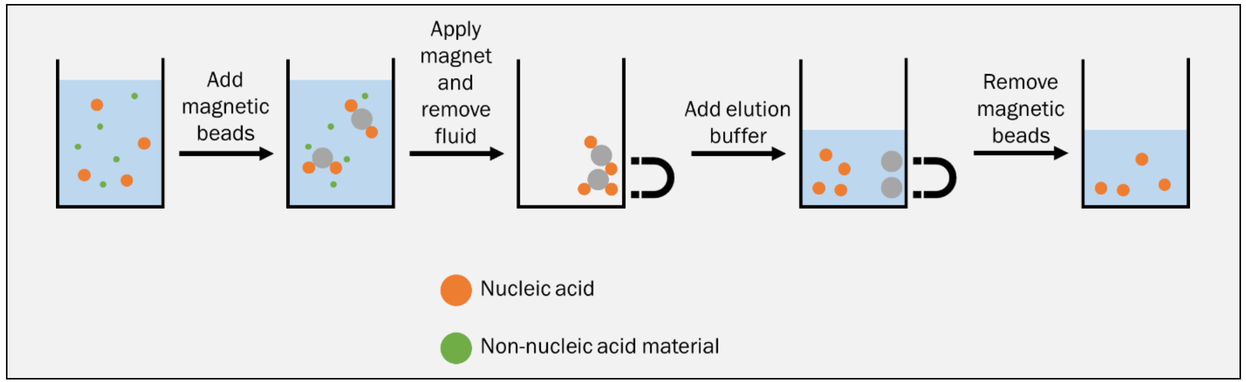
## Nucleic acid extraction

▪ There are many techniques for extracting nucleic acids
▪ Different methods are used to extract DNA and RNA
▪ They generally begin with:
  - Initial deparaffinisation and rehydration
  - Digestion (e.g. heating with proteinase K) to rupture cell and nuclear membranes
  - Heating to denature proteinases and to reverse formalin-induced crosslinks
▪ Spin column purification techniques:
  - The sample is placed in an inner tube, within a larger outer tube
  - At the base of the inner tube is a semi-permeable membrane
  - The tubes are centrifuged repeatedly with wash buffer
  - This pushes non-nucleic acid molecules through the semi-permeable membrane and into the larger outer tube, from which they are discarded
  - Finally, an elution buffer is added to the inner tube and centrifuged again
  - The elution buffer releases nucleic acids from the semi-permeable membrane
  - The purified nucleic acids are collected in the outer tube




▪ Magnetic bead isolation techniques:
  - The sample is placed in a tube and magnetic beads are added
  - The beads are incubated in the sample, such that the nucleic acids bind to their surfaces
  - A magnet is applied to the tube, causing the beads (with attached nucleic acid) to be pulled to one side of the tube
  - The remainder of the fluid is then removed and the beads washed
  - Elution buffer is then added to remove nucleic acid from the surfaces of the beads

- The magnetic beads are removed, leaving purified nucleic acid



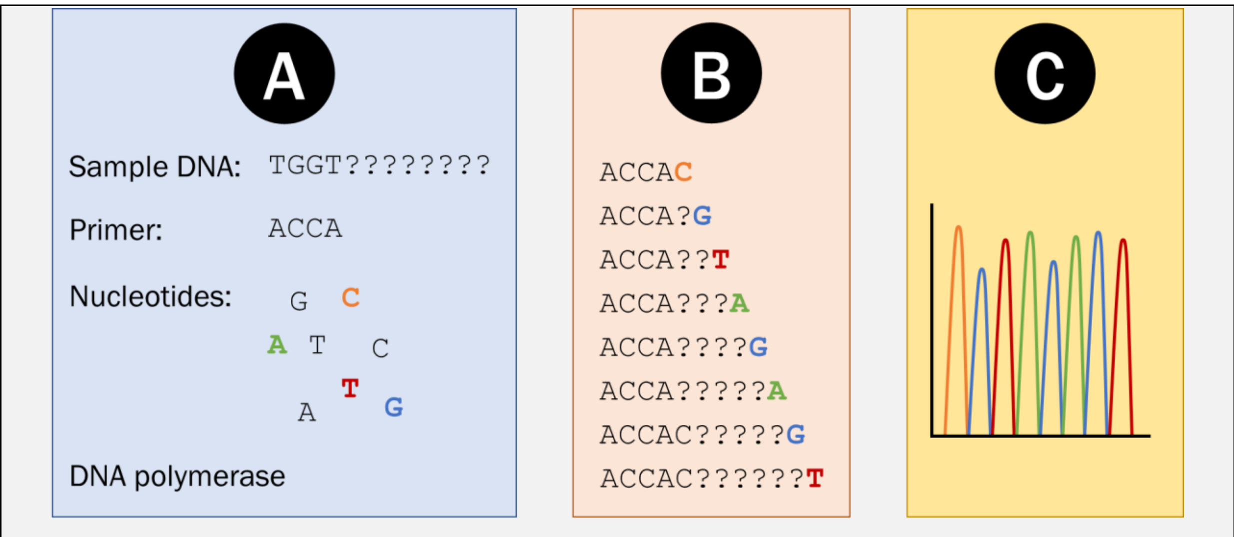

- Extracted and purification nucleic acids can then be used for testing
- It is typical for any excess nucleic acid to be frozen and stored, in case additional testing is required in future

## Sequencing

- Sequencing involves determining the sequence of bases in a nucleic acid
- This derived sequence is then compared to the reference sequence to identify variants

### Direct/Sanger sequencing

- Direct/Sanger sequencing is a chain termination method
- The idea is that:
  - The template strand of the sample DNA is amplified and used as a basis for synthesising new chains of DNA
  - This involves the incorporation of nucleotides by a DNA polymerase
  - There are also dideoxynucleotide triphosphates (ddNTPs) which differ from typical nucleotides in that they terminate chain elongation when they are incorporated
  - The ddNTPs are each fluorescently labelled, with a different fluorescence depending on the base
  - The newly formed chains are ordered by their length by capillary electrophoresis and their fluorescence read off
  - By looking at the fluorescence in each of the size-ordered chains, it is possible to infer the sequence of the original DNA



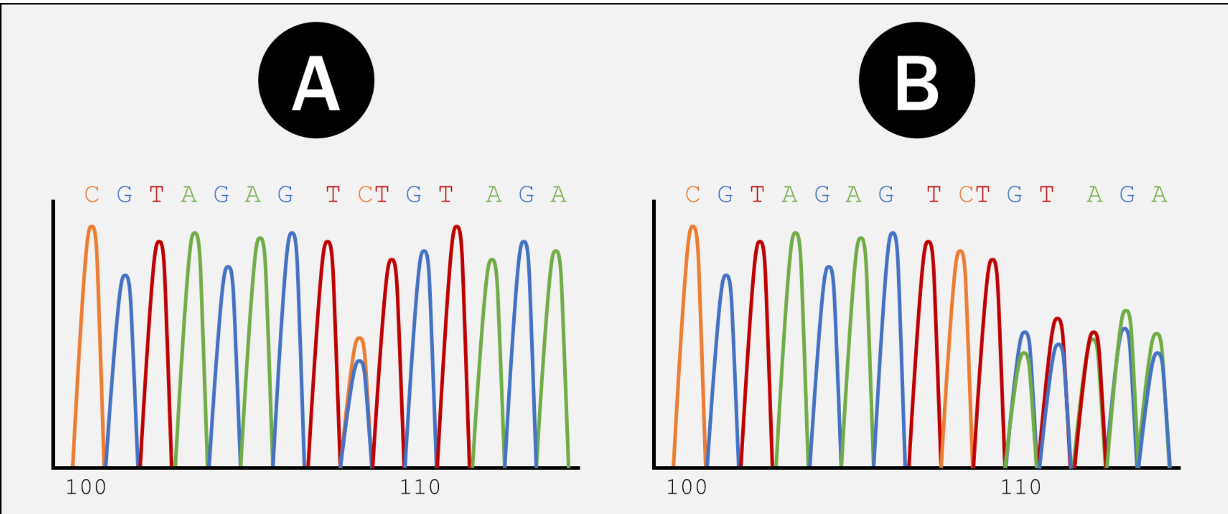

*A: The DNA extracted from the sample is amplified and denatured. We want to know the sequence of the DNA downstream of a short sequence which we already know (TGGT). We add a*

*primer which is complementary to this region in the template strand (sequence ACCA). We then add two sets of nucleotides. The first are unlabelled deoxynucleotide triphosphates (dNTPs); the second are dideoxynucleotide triphosphates (ddNTPs), each of which is labelled with a different fluorescence. The dNTPs allow the strand to continue to elongate after they have bound, whereas the ddNTPs terminate the chain. Finally, a DNA polymerase is added to incorporate the nucleotides into the growing chain.*

*B: The dNTPs bind to the sample DNA. As ddNTPs are incorporated, they stop chain elongation. Since the ddNTPs are fluorescently labelled, it is possible to determine which nucleotide is present at the end of each of the chains formed.*

*C: Capillary electrophoresis is used to order the chains by size. By reading off the fluorescence associated with each of these terminated chains in order of their sizes, it is possible to infer which nucleotide is present at each position in the sequence. The graph shows that the sequence (after the initial already know sequence of TGGT) is CGTAGAGT.*

- The DNA sequence obtained through sequencing is then compared to the reference human genome sequence



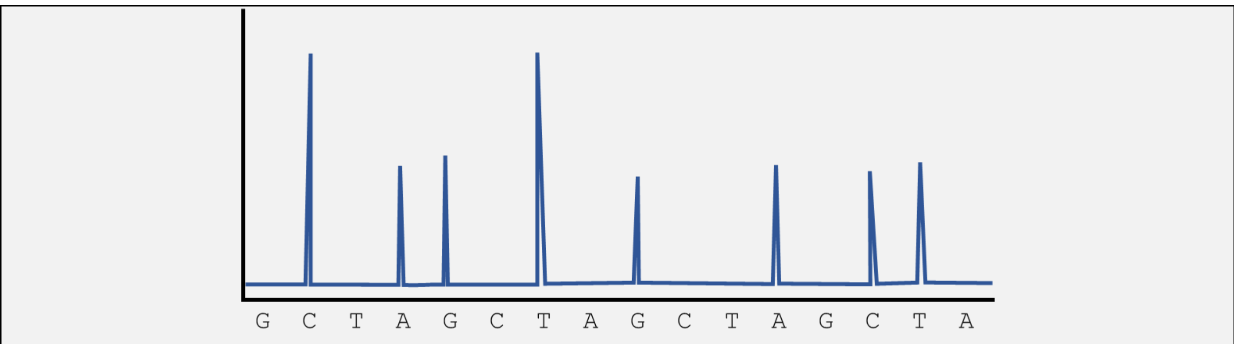

*A: The derived sequence matches what is expected from the reference sequence provided. However, at position 108 there are two fluorescence signals: one from the cytosine which would be expected, but also one from guanine. This indicates that there is a C>G substitution variant (c.108C>G). Because there will inevitably be wild-type alleles in the sample (both from the wild-type allele in the neoplastic cells and from background non-neoplastic cells) there will still be a cytosine signal, but its height is reduced because there is now also a variant guanine signal.*

*B: Here, the derived sequence matches the reference sequence until position 110, where there is a guanine signal as well as the expected adenine signal. It is tempting to say that this is a c.110A>G substitution, but all the signals following this also seem to show evidence of variants. This actually shows an insertion of an adenine between positions 109 and 110. This has resulted in a frameshift, which explains why there is an impact on the subsequent signals. This variant would be described as c.109_110insA.*

### Pyrosequencing

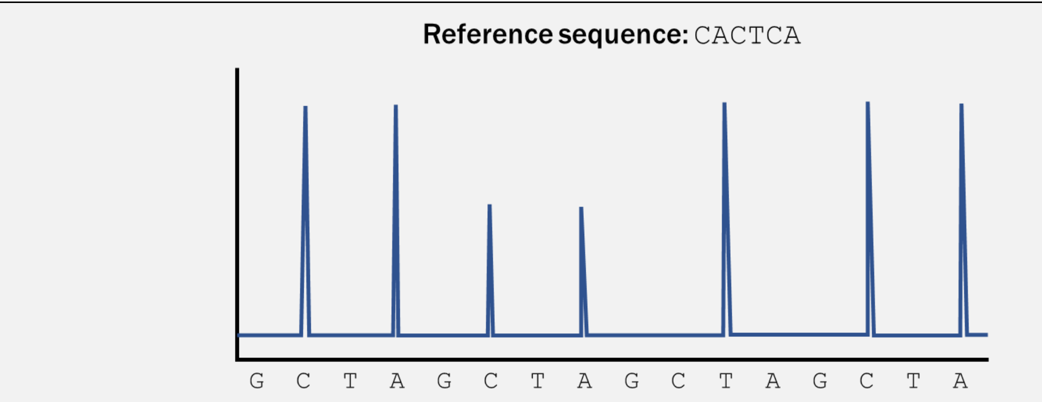
- This involves:
  - Sequentially adding dNTPs to the amplified sample DNA
  - Binding of a dNTP causes release of phosphate
  - The action of sulphurylase uses this phosphate to produce ATP
  - Luciferase uses the ATP to generate light

- Therefore, each dNTP is sequentially added to the sample DNA and emission of light demonstrates that the dNTP has bound
- The amount of light generated is proportional to the number of dNTP which bind




*In pyrosequencing, each nucleotide is added one-by-one; this is demonstrated on the x-axis: G-C-T-A are added sequentially over and over again, and the emitted light detected. When the first guanine is added, there is no light detected, and so the first nucleotide is not a guanine. However, when cytosine is added, there is a very large peak which is around twice the height of most of the peaks; this indicates that the sequence begins CC. The next light emission is at addition of adenine, and then guanine. The sequence so far is therefore CCAG. Next, there is a double-height peak for thymine. Overall, the sequence of this DNA is CCAGTTGACT.*

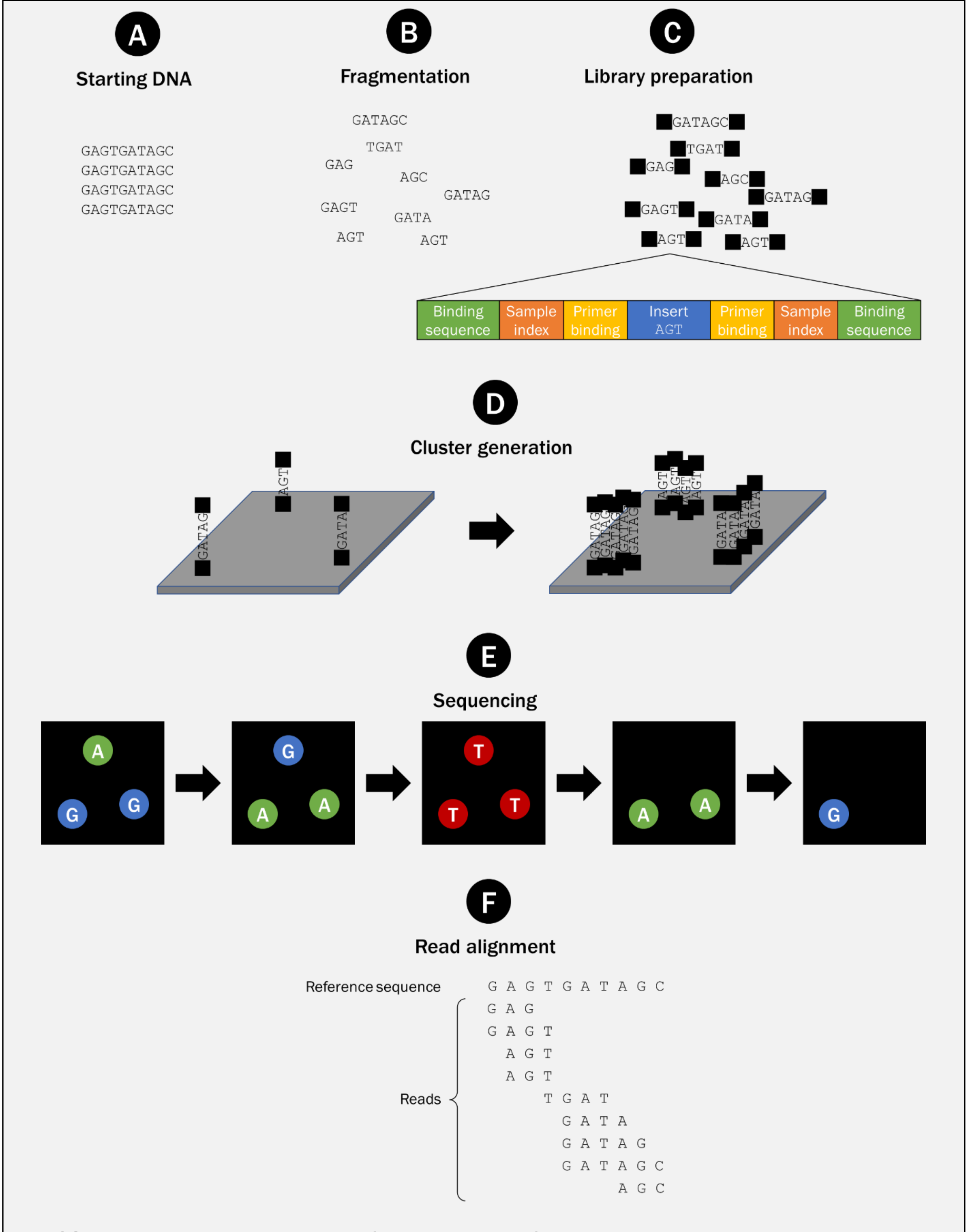
- Identification of a variant relies on comparison between the derived sequence and the reference sequence, and comparison of the heights of the peaks




*The first peaks are for cytosine and adenine, giving an initial CA sequence which is in agreement with the reference sequence. However, there is then a half-height peak for C followed by a half-height peak for A; the reference sequence says that it should be a cytosine. The explanation is that there is a C>A substitution at this position (the half-height cytosine signal reflects the presence of the wild-type allele).*

### Next-generation sequencing

- Next-generation sequencing (NGS) is a form of high throughput or massively parallel sequencing
- It allows multiple regions of the genome to be sequenced (from multiple samples) simultaneously
- Anything from a single gene, to large panels of genes, to whole exomes, to whole genomes may be sequenced
- There are many types of NGS, all based on different principles




*A: NGS begins with multiple copies of the sequence of interest.*

*B: The starting DNA is randomly fragmented into short pieces.*

*C: Library preparation involves adding adapters to the ends of all the short fragments. These are sequences which include (1) a binding sequence which allows the fragment to bind to the flow cell, (2) a sample index, and (3) a primer binding sequence, which allows sequencing to occur. The sample index is essentially a barcode which indicates from which sample the fragment has*

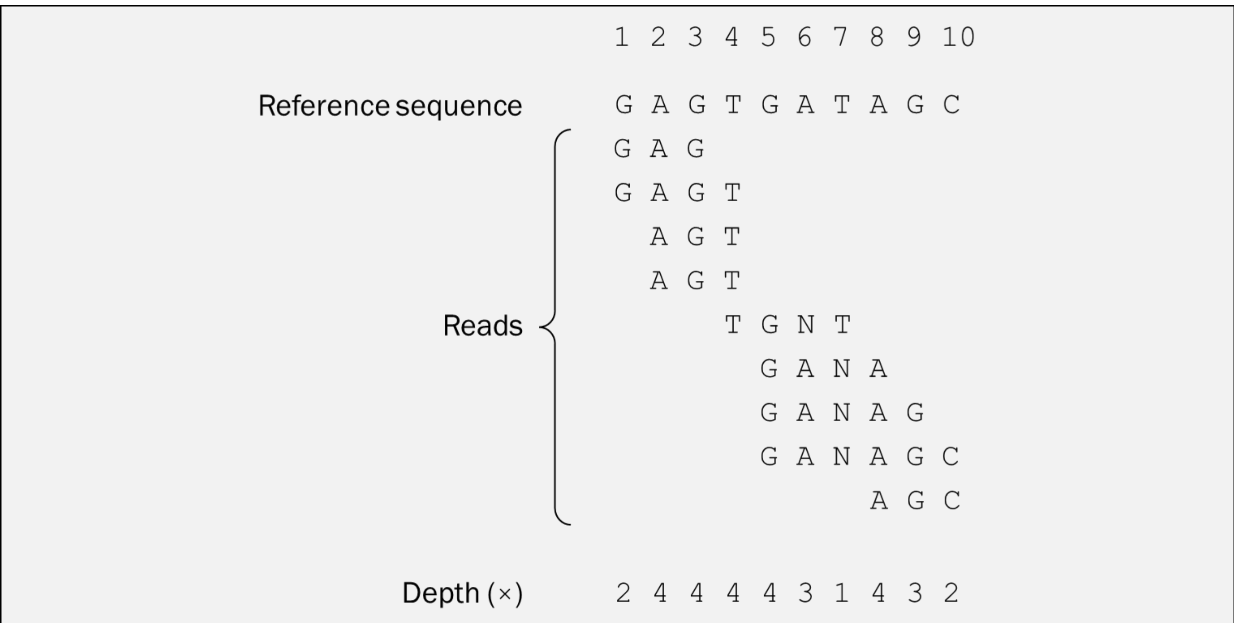
*come. Once library preparation is complete, DNA from multiple different samples can be sequenced together; the sample index makes it possible to work out which sample that fragment came from.*

*D. The fragments with their attached adapters are attached to the surface of a flow cell. Cluster generation is used to replicate the fragments. By the end of this process, the surface of the flow cell is covered with clusters made up of multiple copies of the same DNA sequence.*

*E. A detector looks at the flow cell from above. Fluorescently labelled nucleotides are added to the flow cell. As they bind to the fragments, the fluorescence from each of the clusters is detected. In this example, in the lower-left cluster, the binding of nucleotides follows the order: guanine-adenine-thymine-adenine-guanine. This is the sequence of this cluster.*

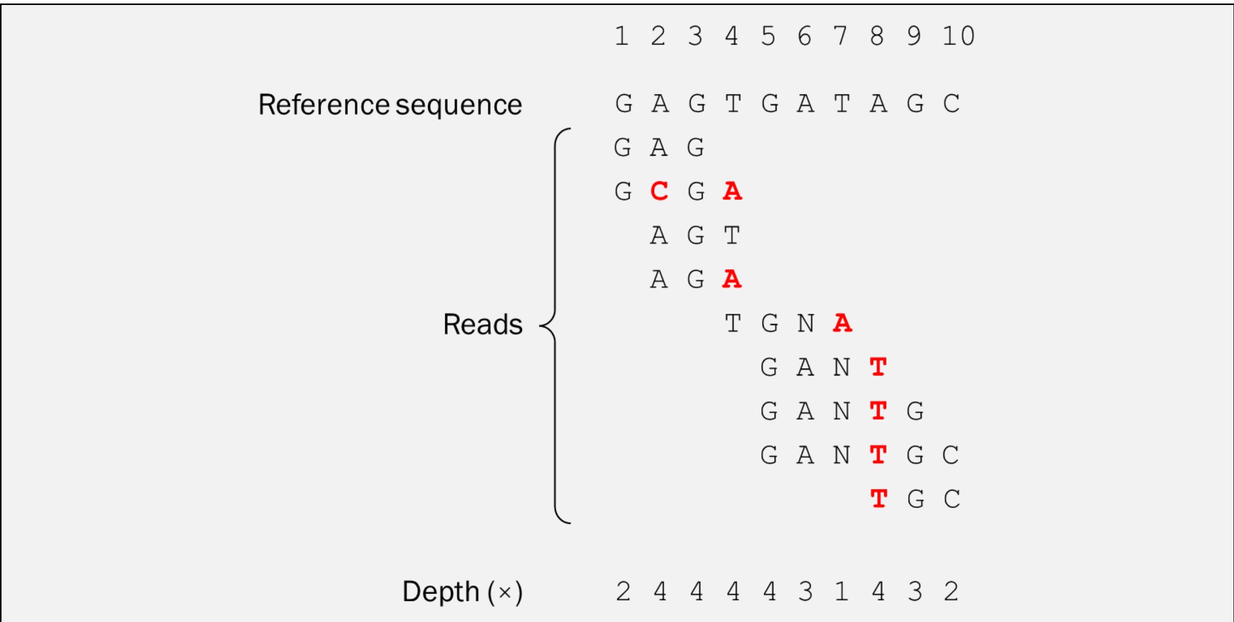
*F. Once the sequences of all the clusters (called reads) has been determined, computer software aligns them. This involves looking at areas of overlapping between the reads and using the reference sequence as a guide. This is called read alignment. From this process, the sequence of the starting DNA can be determined.*

- Coverage/depth of sequencing is an indication of the confidence with which a particular nucleotide can be called




*Coverage/depth is the number of unique reads which include a given base position. The greater the coverage/depth, the greater the confidence with which the nucleotide can be called. Here, the nucleotide at position 2 has a depth of 4×, whereas the nucleotide at position 10 has a depth of only 2×; therefore, the certainty that position 10 is cytosine is lower than the certainty that position 2 is an adenine. At position 6, the sequencer has failed to call the nucleotide in one of the reads, and therefore the depth here is 3×. There was a problem at position 7, and it failed to be called in three of the reads; depth is therefore only 1× and the confidence of this being a thymine will be extremely low. Note that depth/coverage is usually much greater than is shown in this example.*

- NGS can be used to detect small variants

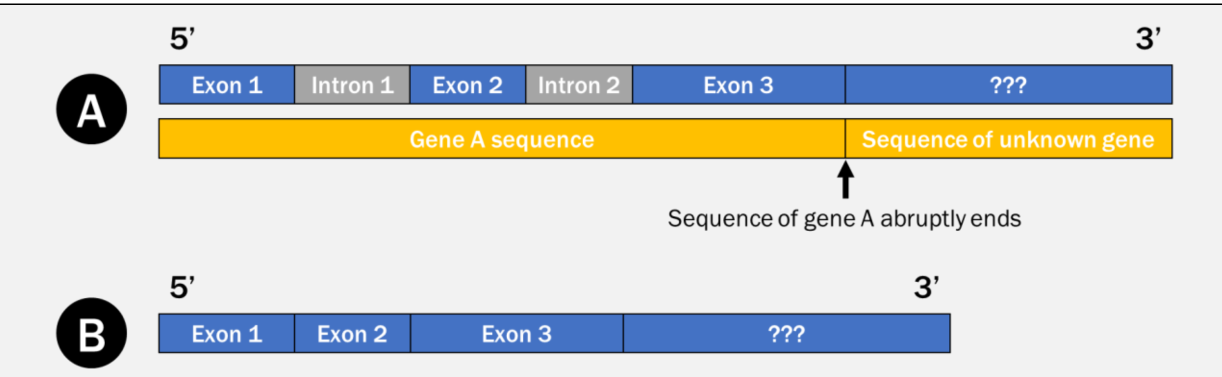



*Note that coverage/depth here is extremely low and in real life would be too low to draw any conclusions.*

*At position 8, the reference sequence shows that there should be an adenine, but in fact thymine has been called in all four reads. This represents a VAF of 100%, raising the possibility that this could be a germline variant (homozygous). At position 4, half the reads contain thymine (as in the reference) and the other half contain adenine; this is a VAF of 50% which is rather high. It is possible that this could represent a somatic variant in a pure sample of tumour, or could represent a germline variant (heterozygous). At position 2, only one of the reads shows an unexpected nucleotide (cytosine). This could be a genuine variant (especially if the neoplastic cell percentage of the sample was low), but it is present in so few reads that it could simply be background noise – a variant cannot be called with confidence. At position 7, the sequencer has failed to call a nucleotide in three of the four reads; the only nucleotide called is adenine, when thymine is expected. Technically, this has a VAF of 100%, but because the coverage/depth is only 1×, it is impossible to determine whether this is a genuine variant or simply noise.*

*It is worth emphasising that the distinction between a germline and somatic variant cannot be made with confidence on sequencing of tumour tissue. A high VAF is concerning for a variant being germline, but the only way to be sure is by sequencing non-tumour tissue.*

- NGS can also be used to detect structural variants
  - Sequencing of DNA may be used, but there is an increased risk of missing rearrangements which are present
  - RNA sequencing is generally used for testing for translocations: RNA is extracted from the sample and converted to cDNA by reverse transcriptase – the cDNA is then sequenced

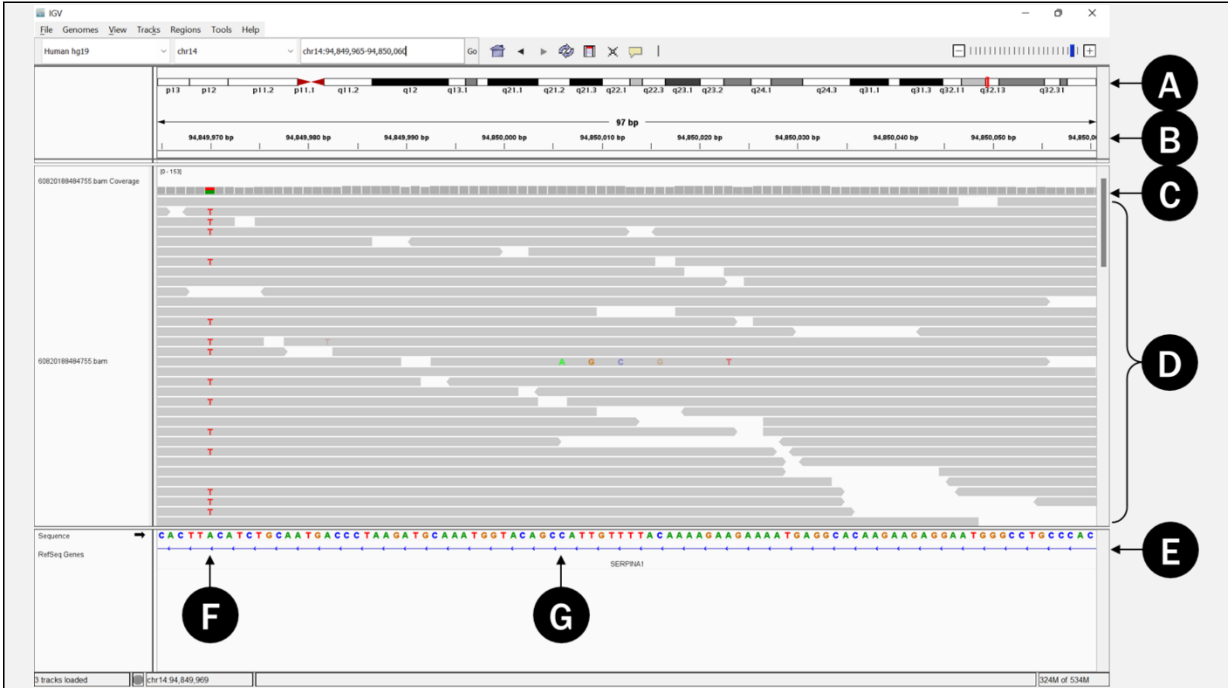



*A: The principle of sequencing for gene rearrangements in DNA. The gene of interest (gene A) is sequenced. The derived sequence matches the reference sequence until exon 3, when suddenly the sequence no longer matches that expected for gene A. The sequence of the unknown gene is derived and, by searching online databases, it is possible to determine which gene this sequence belongs to, and therefore to identify the nature of the translocation. The problem with sequencing DNA for rearrangements is that the fusion points frequently occur in introns, which are often long, repetitive and therefore difficult to sequence. There is therefore a considerable risk of missing a rearrangement which is present. A second problem is that a rearrangement present at the DNA level may not result in the production of a functional fusion RNA transcript or a functional fusion protein; giving these patients targeted treatment is unlikely to be beneficial for the patient.*

*B: RNA sequencing is generally considered to be a reliable method for detecting rearrangements than DNA sequencing. mRNA has intronic sequences spliced out, resulting in only exons; sequencing a gene from its mRNA therefore avoids having to sequence the long, repetitive intronic sequences which are what make DNA sequencing challenging. In addition, a fusion which is detected at the mRNA level is more likely to be functionally relevant than one present only at the DNA level. There are disadvantages to RNA sequencing, however. RNA is generally less stable than DNA and so RNA sequencing is significantly more likely to fail. RNA sequencing requires a separate nucleic acid extraction process and a separate NGS panel to that used for DNA sequencing.*

- NGS can be, and is, used to detect copy number variants:
  - It is generally not possible to distinguish between polysomy and focal amplification
  - FISH may be required as a second-line technique where this distinction is important clinically
- NGS can be used to look for as little or as much as is needed (depending on the quality of the sample):
  - A single region of a single gene
  - Multiple regions of a single gene
  - The entirety of a single gene
  - Multiple regions of multiple genes
  - The entirety of multiple genes
  - The entire exome (i.e. all exons in the genome)
  - All coding DNA
  - Whole genome sequencing
- The most common approach is to use NGS panels:
  - Look at key regions of multiple genes of interest
  - For example, a solid tumour panel will look at key regions of genes which are known to be targetable across the full range of solid cancers

  - Panels may vary in size from a few dozen genes to several hundred genes
  - The provider may report on all mutations detected, or may only report on detected mutations which are known to be relevant for the exact tumour type in question
- Depending on the tumour type, it is common to run two sequencing operations in parallel:
  - A DNA NGS panel to assess for small variants
  - An RNA NGS panel to assess for structural variants
- There are many types of NGS:
  - They differ in their chemistry, the extent of sequencing, and the range of alterations which they can detect
  - This results in substantial differences in turnaround times, failure rates and sample requirements




*Real-life output data from NGS.*

*A: A map of the chromosome being viewed, in this case is it chromosome 14, around the centromeric region (at the junction between the short – p – and long – q – arms).*

*B: The base position.*

*C: A plot of coverage/depth of sequencing at each base position. Higher bars indicate greater coverage/depth. Coverage/depth is fairly even across all base positions in this region.*

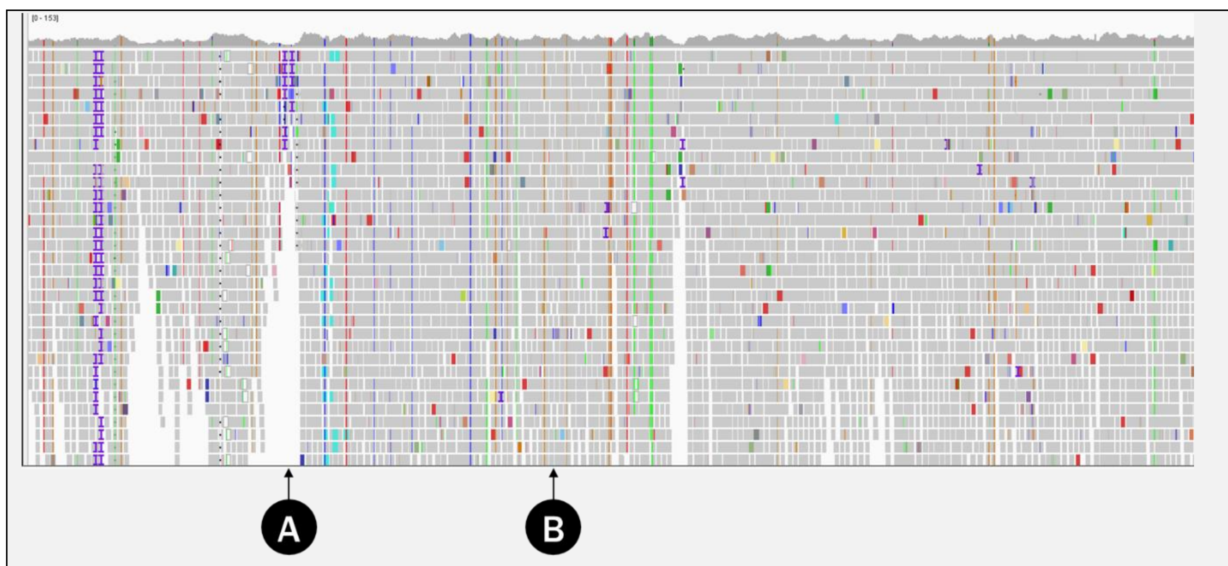
*D: The alignment track. Each grey bar represents a single read. The reads have been aligned against each other and against the reference sequence. Where the nucleotide called in the reads agrees with what would be expected from the reference sequence, no nucleotide is shown in the reads; only where the nucleotide called differs from that in the reference sequence is the nucleotide shown in the read.*

*E: The reference sequence. In the top left, the 'Human hg19' sequence has been selected as the reference.*

*F: Here, around half the reads contain a thymine and the other half contain adenine (as expected in the reference sequence). Coverage/depth here is good (39×) so this is likely genuine. If this represents sequencing of tumour tissue, it could either represent a somatic*

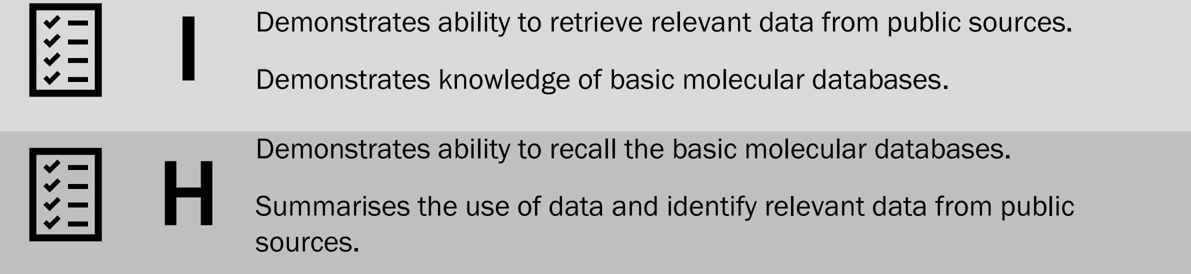
*variant in an extremely pure sample of neoplastic cells, or it could be a heterozygous germline variant. If this represents sampling of non-tumour tissue, it is a heterozygous germline variant.*

*G: Here, coverage/depth is good (44×). 43 of the 44 reads contain cytosine, which is expected from the reference sequence. Only one read contains an adenine. It is extremely unlikely that this represents a genuine variant – it is almost certainly noise. Interestingly, this read contains several other similar calls, suggesting that there may have been a problem with this read.*




A zoomed out image of the previous NGS output. Here, it is much easier to appreciate the variation in coverage/depth across the genome. At position (A), there is dramatic drop in coverage/depth of sequencing, as indicated by the histogram above the reads. It is possible that this represents a highly repetitive region of the genome where it has been very difficult to align the reads. Any variants identified here are likely to be called with very low confidence. Contrast this to position (B), where the reads have aligned nicely, resulting in high coverage/depth of sequencing.

## Data analysis



| | | |
|---|---|---|
| | I | Demonstrates ability to retrieve relevant data from public sources.<br>Demonstrates knowledge of basic molecular databases. |
| | H | Demonstrates ability to recall the basic molecular databases.<br>Summarises the use of data and identify relevant data from public sources. |

- Regardless of the technology, sequencing produces a list of variants (i.e. differences in the DNA sequence between the sample and the reference sequence)
- The next step is to determine the significance of the data
- Processing of this data is referred to as bioinformatics
- Are the detected variants genuine or background noise/artefact?
  - This is largely determined by the depth/coverage and by the number of reads in which the 'variants' have been called
  - Software algorithms can identify other artefacts (including formalin-induced artefacts)

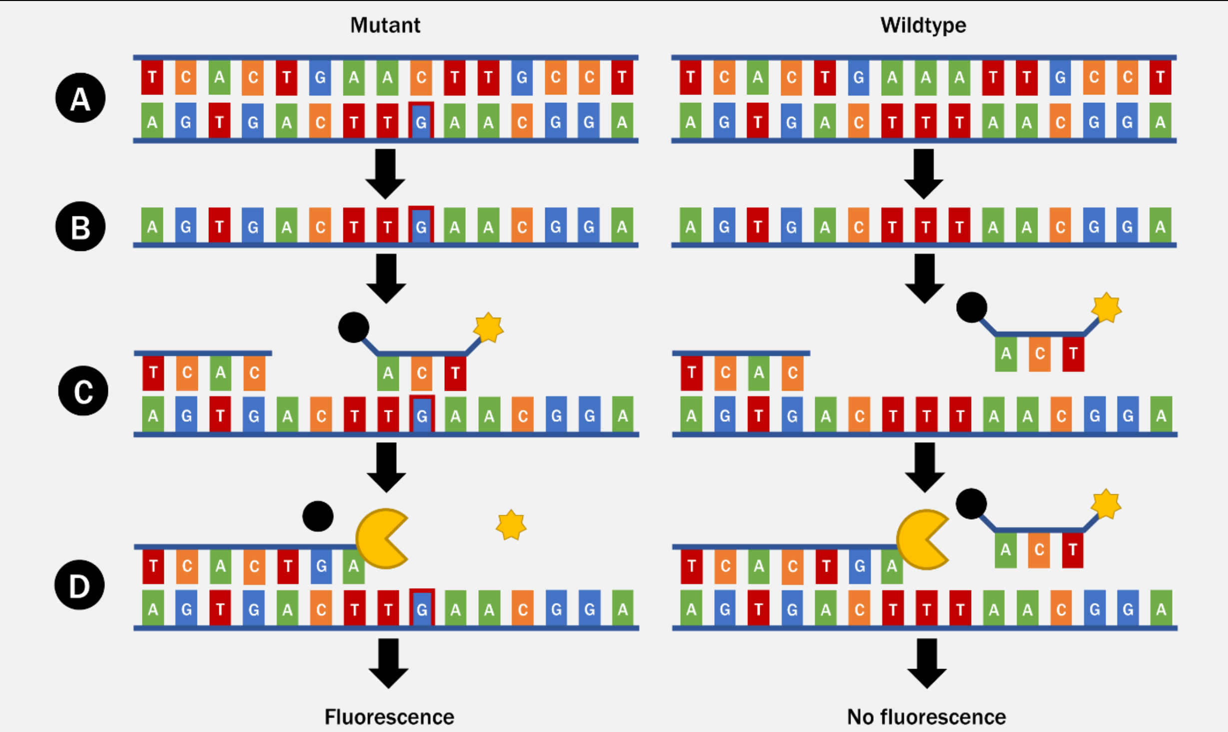
  - Much of this is automated and 'variants' which are unlikely to be genuine can be filtered out
- Are the detected variants pathogenic?
  - In some cases, online databases will give an indication of whether variants are likely pathogenic or benign, based on reports from previous cases
  - ClinVar compiles this information for germline and somatic variants
  - COSMIC compiles this information specifically for cancer
  - If this information is not available, the population prevalence of the variant can give an idea of its pathogenicity – prevalent variants are unlikely to be pathogenic
  - In some cases, in silico modelling can be used to predict whether the variant is likely to be pathogenic, by predicting whether there will be a change in amino acid and, if so, whether this is likely to change the chemical properties of the protein
- Are the detected variants relevant?
  - The first thing worth checking is whether the variant has an associated NICE-recommended treatment – if a treatment is recommended by NICE, it will be funded by NHS England (as long as the criteria stipulated by the recommendation are met)
  - If not recommended by NICE, there may be a treatment available via the Cancer Drugs Fund, which means that it is funded by NHS England on a temporary basis while more evidence is accrued to determine whether it should be NICE-recommended
  - If not NICE-recommended, the next thing to check is whether the variant can be targeted by a drug which has a license in the UK – licensed drugs which are not NICE-recommended may be prescribed but are not routinely funded by the NHS (the patient could access them privately or the drug company may provide them on a compassionate use basis)
  - If not yet licensed, drugs may be accessible under the Early Access to Medicines Scheme (EAMS) – this scheme allows patients to access promising new drugs prior to them receiving a license
  - If there is no licensed therapy, the next option would be clinical trial enrolment – clinical trials can be found at clinicaltrials.gov
  - Finally, if there is no clinical trial opportunity, but there is reason to believe that a drug may be effective against a detected variant, the patient and clinician may decide to pursue off-label prescription – this is when a drug is not used in its approved setting, and in doing so the prescribing clinician may accept more responsibility to adverse outcomes
- Certain variants detected in tumour tissue may raise concern that they may be germline:
  - Because the detected variant is one which is frequently seen in germline
  - Because the VAF is unusually high (especially if it is disproportionate to the neoplastic cell content)
  - Ultimately, the only way to determine whether a variant is germline it to sequence non-tumour tissue – this may happen via referral to clinical genetics or may be instigated by the clinician

## PCR-based techniques

### Real-time PCR

- Real-time PCR is a technique which monitors the amplification of target DNA as the reaction proceeds
- It is either fully or semi-quantitative

- There are variants of real-time PCR, but the most commonly used involves using fluorescently labelled probes which are complementary to pre-specified variants of interest



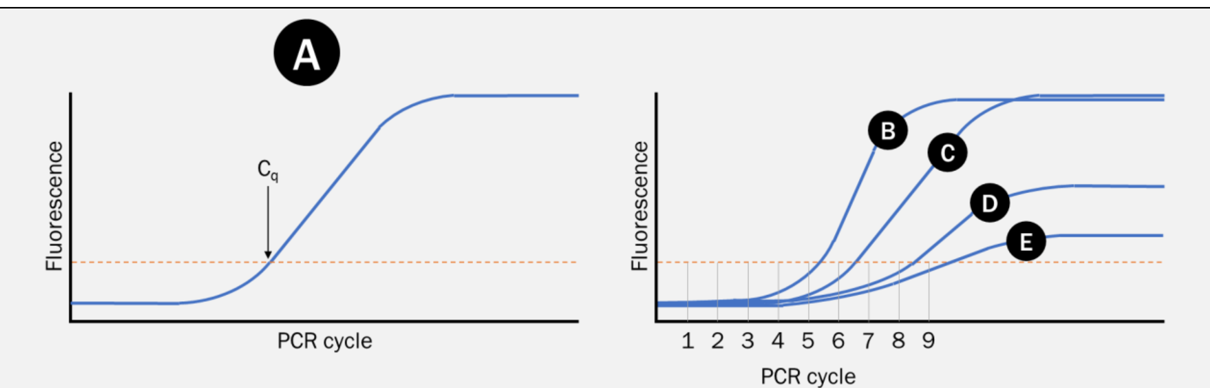

*A: There are two samples. One is wild-type. The other harbours a C>G substitution.*

*B: The DNA is denatured to produce single-stranded DNA.*

*C: A probe which is complementary to the variant of interest is added. The probe has bound to it a fluorophore and a quencher; as long as the fluorophore and quencher remain in close contact, attached to the probe, no fluorescence is release. The probe binds to the variant sequence but not to the wild-type sequence. A primer is added.*

*D: Taq polymerase sequentially incorporates nucleotides from the primer. As the Taq polymerase reaches the bound probe in the variant sequence, the probe is degraded. This results in the fluorophore and quencher separating. Therefore, if there is a variant against which the probe is complementary, fluorescence will be released. In contrast, if there is no variant against which the probe is complementary, no fluorescence will be released.*

- Multiple probes can be used in each real-time PCR reaction to detect multiple different variants
- However, because the probes will identify only pre-specified mutations, if there is a variant for which no probe has been used, it will not be detected
- By labelling the probes with different fluorescence, it is possible to determine the exact variant present
  - There is a limit to the number of variants which can be distinguished from one another, since there is a limit to the different types of fluorescence which can be discriminated from each other



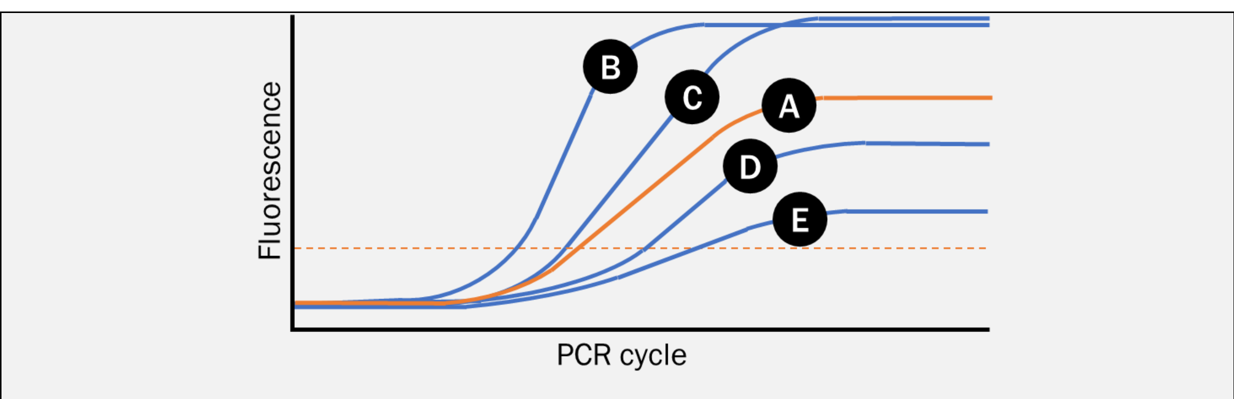

*A: The fluorescence detected from real-time PCR adopts a sigmoid curve. There is initially a slow take-off; then, as the reaction gains pace, fluorescence rapidly intensifies; finally, as substrates are consumed and products accumulate, the reaction slows to a plateau. The level of fluorescence must reach a particular threshold within a prespecified number of PCR cycles in order for the target to be considered present. The PCR cycle at which the curve crosses the threshold is the $C_q$ value.*

*B-E: These are amplification curves for a variety of targets. In this reaction, the amplification curve must cross the threshold within 9 PCR cycles in order to be considered valid.*

*B-D: These amplification curves cross the threshold within 9 PCR cycles, and so the target has been detected in all three cases. The $C_q$ is approximately 5.5, 6.5 and 8.5 for B, C and D, respectively. The fact that D crossed the threshold later than B (i.e. it has a higher $C_q$) suggests that there is more of target D present than there is target B.*

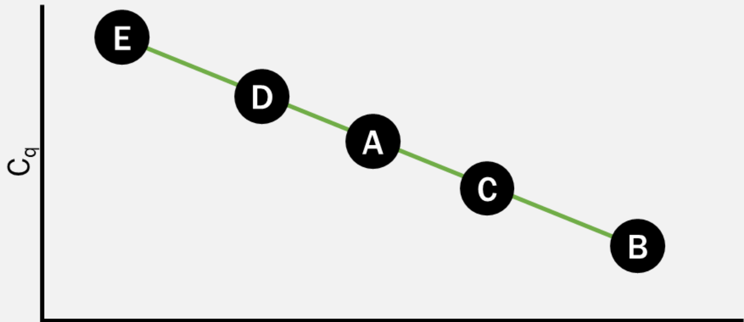
*E: Although the amplification curve has crossed the threshold, it has only done so after 9 PCR cycles. The target cannot, therefore, be regarded as having been detected. If this reaction was looking for a specific variant, that variant cannot be considered to have been detected. However, if the tumour cell percentage of the sample was lower than the minimum requirement for this technique, the fact that it crossed the threshold only just after 9 cycles raises the possibility of this being a false negative result. In this situation, it would certainly be worth suggesting repeat testing if more tissue becomes available.*

- Real-time PCR can be used to quantify the amount of target in the sample
    - This is very seldom required in solid tumour pathology, but is quite frequently used in haemato-oncology




*A: This amplification curve belongs to the test sample. We want to know how much of the target DNA is present in this sample.*

*B-E: Alongside the test sample, we also undertake real-time PCR for the same target in multiple control samples in which we know how much of the target DNA is present. B has the largest amount of the target, and E has the smallest amount.*



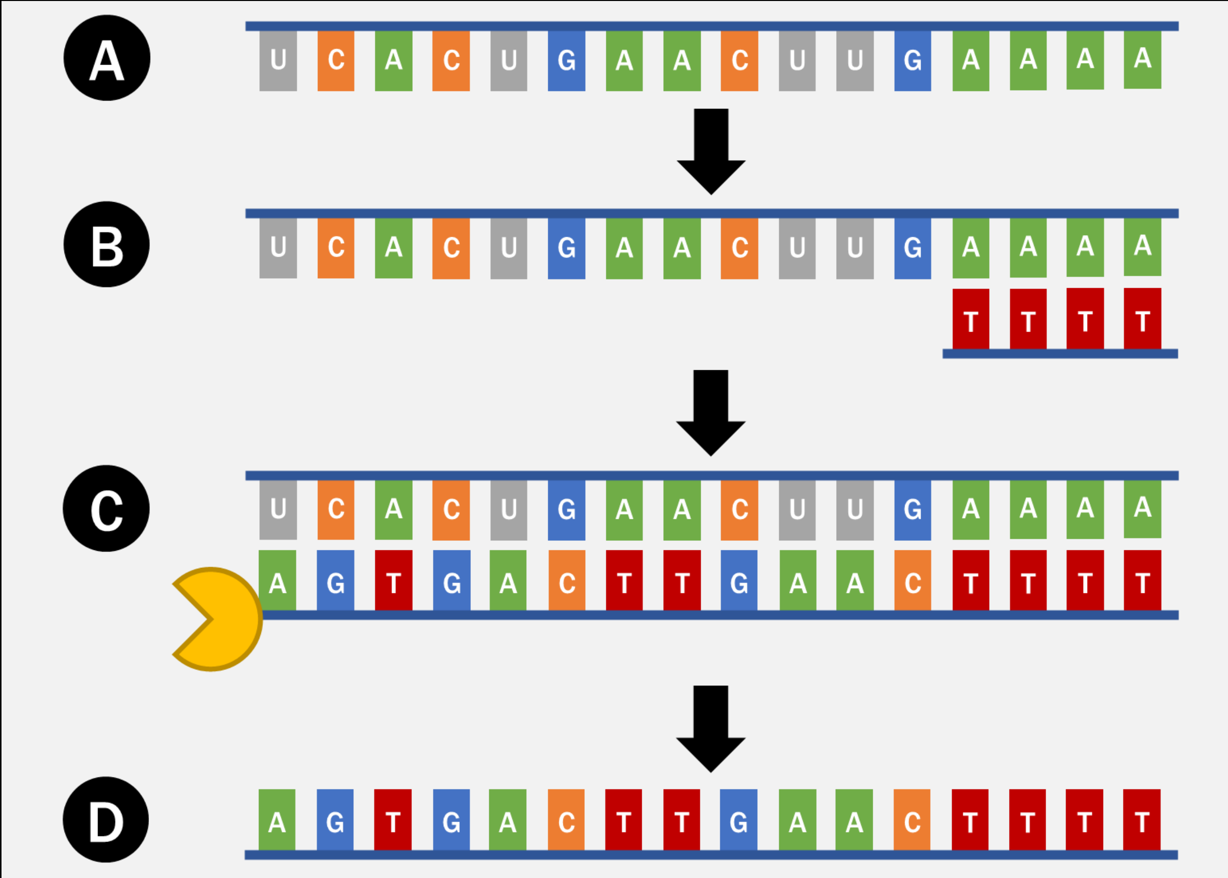

*B-E: We plot the $C_q$ values from the above amplification curves against the logarithm of the starting amount of target DNA (which we already know for these samples, since they are controls). This forms a straight line.*

*A: Using the $C_q$ value from the amplification curve, we place our test sample on the graph and read off the log[starting amount] from the x-axis. Calculate $e^{(x\text{-axis value})}$ to determine the amount of starting DNA in the sample.*

- A single real-time PCR assay can often detect up to a few dozen targets:
  - These may be within a single gene
  - These may be across multiple genes
- Major advantages of real-time PCR are that:
  - It is generally fast
  - It generates data output which is relatively easy to analyse
  - It has a low failure rate in formalin-fixed tissue
  - It is able to reliably detect variants in samples with low neoplastic cell content

### Reverse transcription PCR

- Reverse transcription PCR (RT-PCR) involves two steps:
  - RNA is converted to cDNA
  - PCR is undertaken on the cDNA
- The PCR component of RT-PCR is very often real-time PCR:
  - Reverse transcription PCR is abbreviated to RT-PCR
  - Real-time PCR may be abbreviated to qPCR
  - When RT-PCR is combined with real-time PCR, it can be described as qRT-PCR



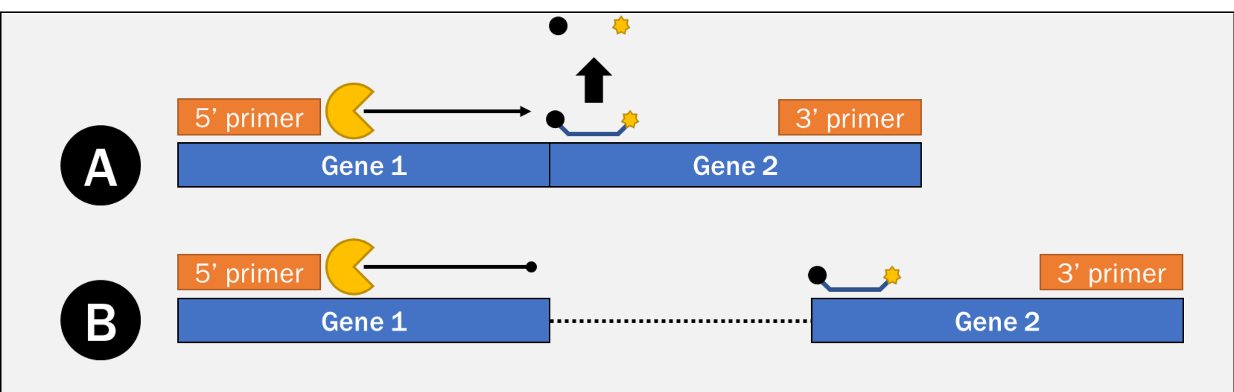

*A: The starting mRNA sequence, which bears a poly(A) tail at its 3' end.*

*B: A primer complementary to the poly(A) tail is added.*

*C: RNA polymerase sequentially incorporates complementary DNA nucleotides, extending from the primer.*

*D: The result is a strand of cDNA. The sequence of this cDNA will be identical to the DNA from which the mRNA was transcribed.*

- RT-PCR is typically used to:
  - Quantify gene expression, by assessing levels of mRNA sequences
  - Infer the presence of gene fusions
  - [Detect the presence of RNA viruses, e.g. SARS-CoV-2]

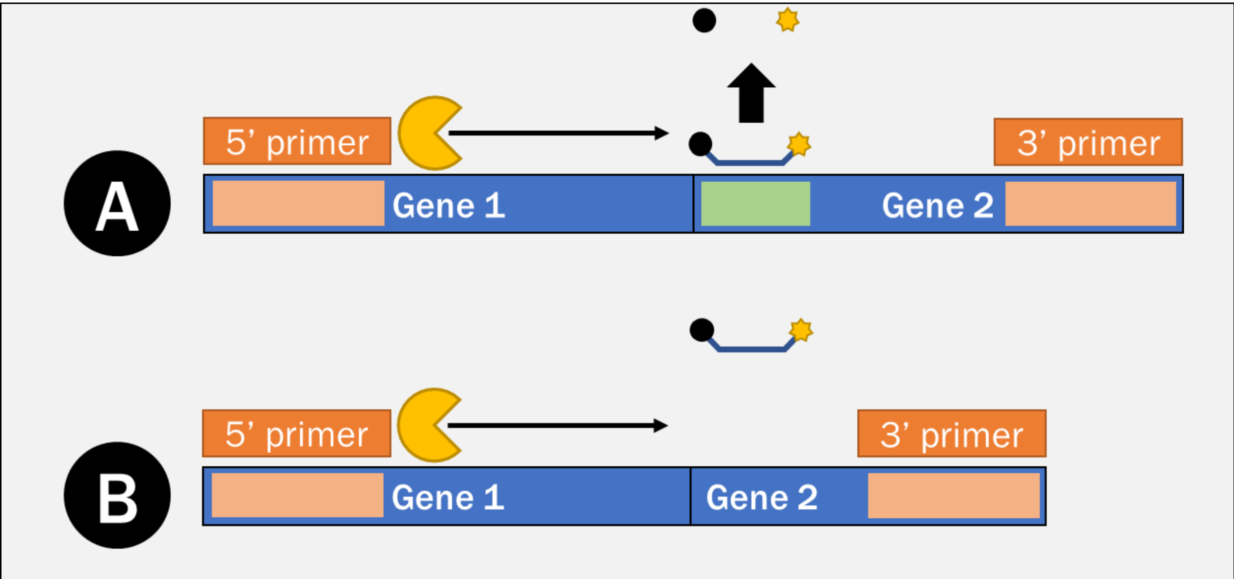



*A: We want to know whether there is a fusion involving genes 1 and 2. We have already used RT-PCR to generate cDNA from the mRNA of gene 1. We now use real-time PCR. We use primers*

*complementary to a sequence in gene 1 (the 5' side of the fusion we want to look for) and to a sequence in gene 2 (the 3' side). We use a probe which is complementary to a sequence between the two probes. Taq polymerase incorporates nucleotides from the 5' primer. Because there is a fusion, gene 1 and gene 2 are next to each other. They are very close and so Taq polymerase is able to produce a DNA strand which reaches the probe. When Taq polymerase reaches the probe, it causes it to break down. The fluorophore and quencher separate, and fluorescence is emitted. The detection of fluorescence therefore indicates that the fusion of interest is present.*

*B. Here, there is no fusion. The primers and probe bind as above. However, now genes 1 and 2 are widely separated (or perhaps even on separate chromosomes). They are therefore too far apart for Taq polymerase to bridge the gap. The probe therefore does not break down. The quencher and fluorophore remain adjacent to each other and no fluorescence is emitted.*

- As for when real-time PCR is used to detect small variants, when it is used to look for translocations it will only detect specific gene fusions which are being tested for:
  - It will only detect fusions between specific genes: if the included primers and probes only bind *ROS1* and *CD74*, an *EZR*::*ROS1* fusion will be missed
  - It will only detect fusions of particular types:

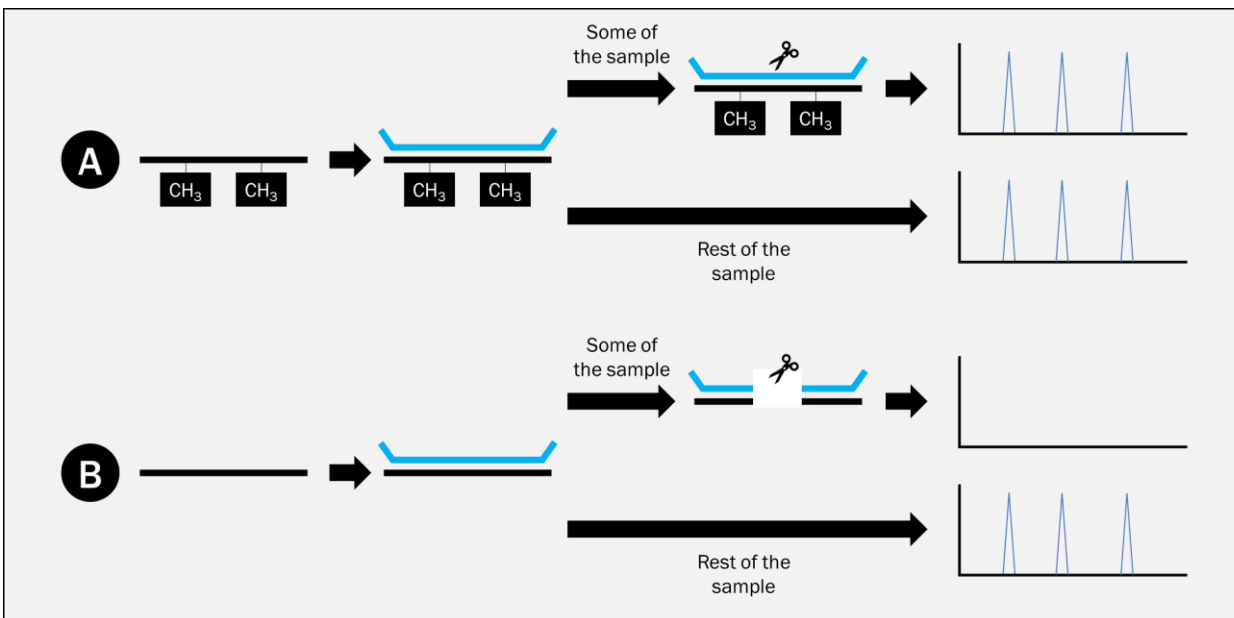



*A: This fusion has preserved both primer the probe sequences. Therefore, primers and the probe can bind appropriately and the fusion will be detected.*

*B: In this example, the fusion has resulted in loss of the probe binding region (green). The probe therefore cannot bind and so, although there is a fusion, there will be no detectable fluorescence and the fusion will not be detected. A second probe could be included to allow for this fusion to be detected. Including sufficient primers and probes to cover all possible fusions between these two genes is, however, impossible (and that is only for two genes). Real-time PCR will therefore invariably miss fusions.*

## Methylation analysis

- There are multiple techniques for assessing the degree of DNA methylation
- Methylation-specific multiplex-ligation dependent probe amplification (MS-MLPA) and methylation-specific PCR are two common techniques

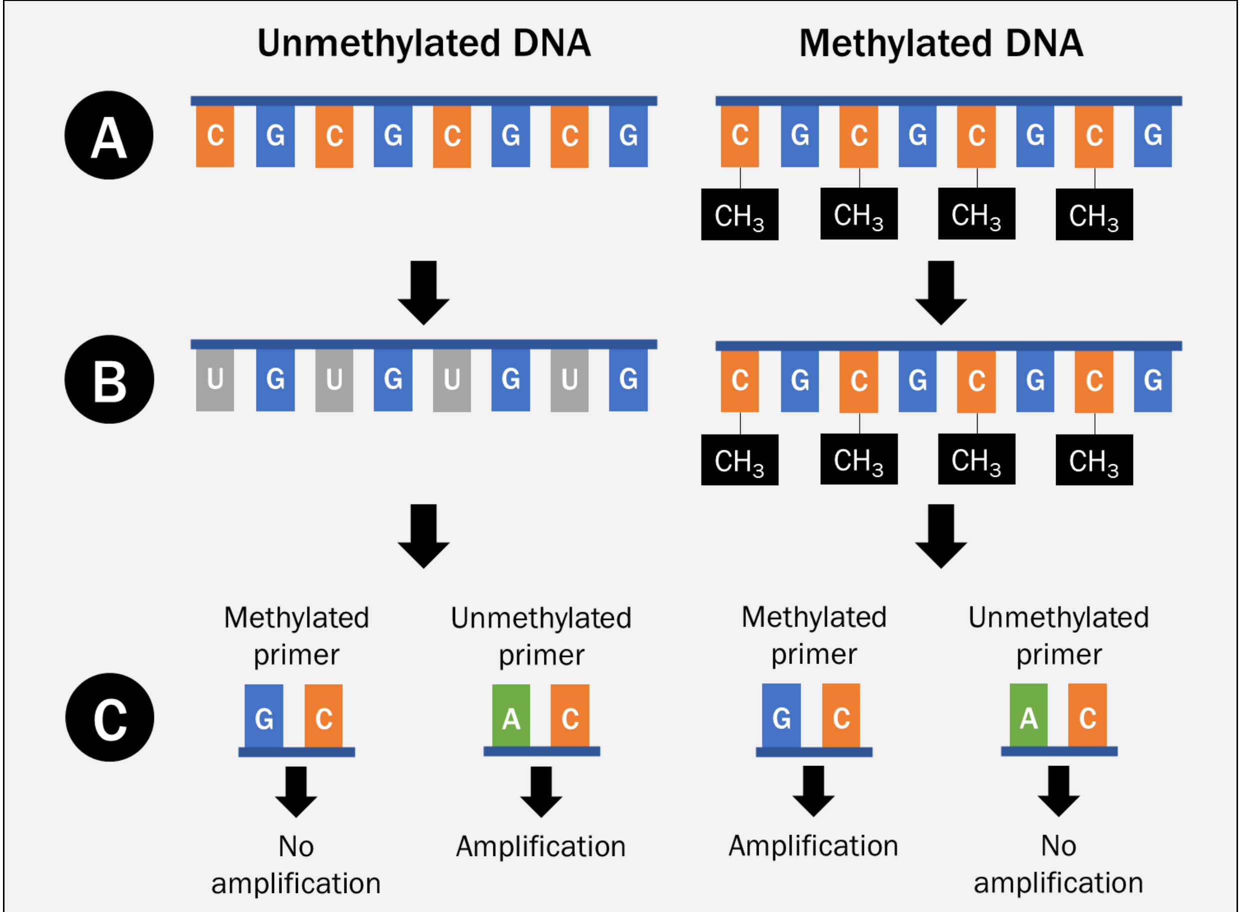



*MS-MLPA*

*A: A methylated DNA sample is tested using MS-MLPA. The DNA is denatured and hybridised with a probe which is complementary to the DNA region whose methylation is to be assessed. The sample is then split. Some of the sample is incubated with a methylation-sensitive restriction enzyme; this digests the probe and DNA if the DNA is unmethylated, but leaves the probe and DNA intact if the DNA is methylated. In this case, it leaves the probe and DNA intact because it is methylated. The rest of the sample is not incubated with the enzyme. The probes from both samples are then amplified. Because the probe is intact in both samples, they both show amplification. This indicates that the starting DNA was methylated.*

*B. An unmethylated DNA sample undergoes testing. The DNA is denatured and hybridised with the probe. Again, the sample is split. Some is incubated with the same methylation-sensitive restriction enzyme; here, because the DNA is not methylated, both the DNA and probe are digested. Meanwhile, the rest of the sample is set aside and not treated with the enzyme (therefore, the DNA and probe remain intact). The probes from both samples are amplified. Because the probe was digested, there is no amplification from the first part of the sample; conversely, there is amplification from the second part of the sample. This finding indicates that the reaction has worked appropriately, and that the sample was not methylated.*

*In summary, methylated DNA gives rise to amplification but unmethylated DNA does not.*




*Methylation-specific PCR*

*A: We start out with two DNA samples, one methylated and one unmethylated. CpG islands (as here) are more prone to methylation.*

*B: The first step is bisulphate conversion. This converts unmethylated cytosine residues to uracil, but leaves methylated cytosine as cytosine.*

*C: PCR is then undertaken using two primers. One primer (the methylated primer) binds only to the methylated sequence of interest; the other primer (the unmethylated primer) binds only to the unmethylated sequence of interest. Here, the methylated primer cannot bind the unmethylated sequence because it contains uracil residues; however, the unmethylated primer amplifies, indicating that the DNA sample is not methylated. Conversely, for the methylated sample, only the methylated primer can bind and therefore only the methylated primer is amplified.*

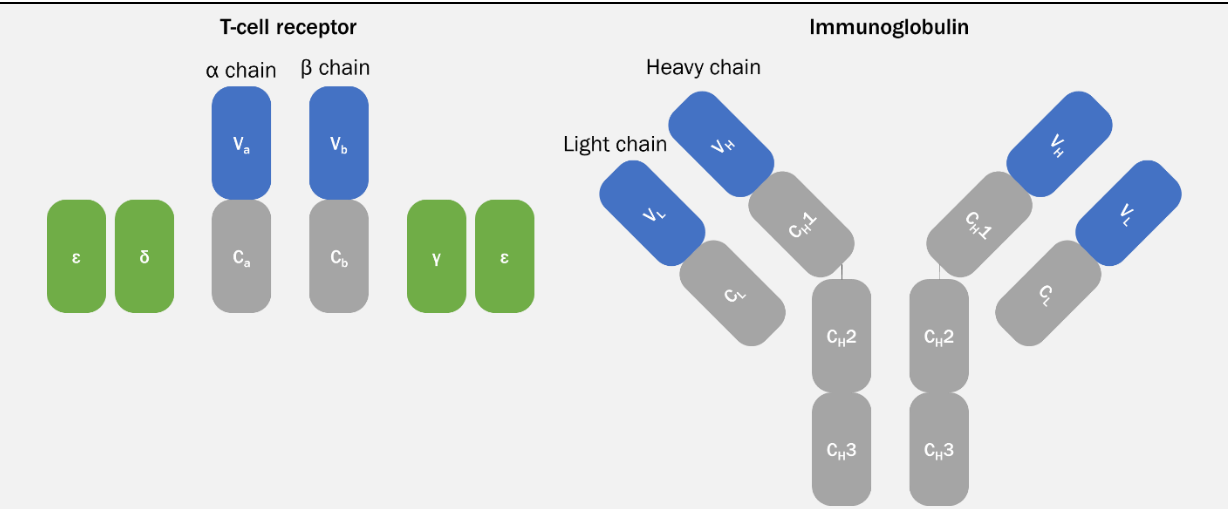
*Any sample, of course, will contain both methylated and unmethylated DNA. By comparing the amplification of both the methylated and unmethylated primers, it is possible to infer the percentage methylation of the sequence of interest.*

- In cases of cancers showing abnormal MMR immunohistochemical staining or microsatellite instability, the finding of *MLH1* promoter methylation almost always indicates that the abnormality is somatic rather than germline (i.e. Lynch syndrome):
  - Very rarely, it may be because of constitutional *MLH1* promoter methylation (i.e. Lynch syndrome)
  - It is therefore prudent to undertake *MLH1* promoter methylation testing in parallel on both tumour and non-tumour tissue

  - Detection of methylation in tumour but not in non-tumour tissue indicates that the methylation is somatic
  - Detection of methylation in both tumour and non-tumour tissue indicates that the methylation is likely germline, and that the patient has Lynch syndrome

## Clonality analysis

- Clonality testing is a tool to help in determining whether a lymphoid proliferation is reactive or neoplastic
- B-cells express B-cell receptors/immunoglobulins which are encoded by *IG* genes
- T-cells express T-cell receptors which are encoded by *TCR* genes
- It relies on the fact that lymphoid cells randomly rearrange their *IG/TGR* genes, such that each cell's *IG/TGR* genes should differ in length and sequence from all other cells' *IG/TGR* genes
- Therefore:
  - All cells in a reactive proliferation should have different *IG/TGR* gene structure
  - All cells in a neoplastic proliferation should have the same *IG/TGR* gene structure
- By looking at how similar the *IG/TGR* genes are to each other in a sample, it is possible to provide evidence to suggest how likely a lymphoid proliferation is to be reactive versus neoplastic

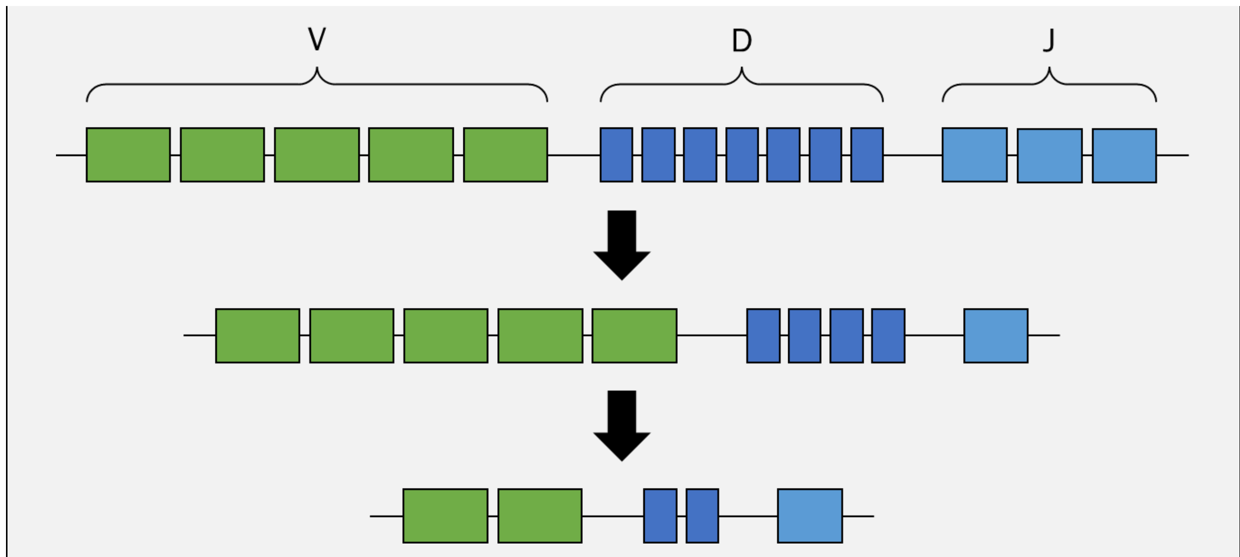



*The T-cell receptor and immunoglobulin molecules have similar structures. The T-cell receptor comprises an α and a β chain, each of which comprises a constant and a variable domain (although in a minority of T-cells, the receptor comprises a γ and δ chain). It is also associated with accessory molecules, which comprise γ, δ and ε chains.*

*Immunoglobulins comprise a heavy chain and a light chain, each of which comprises a constant and a variable domain. The constant domain does not vary, but the variable domain undergoes rearrangement. The latter is what allows T-cell receptors and immunoglobulins to bind to the full diversity of antigens.*

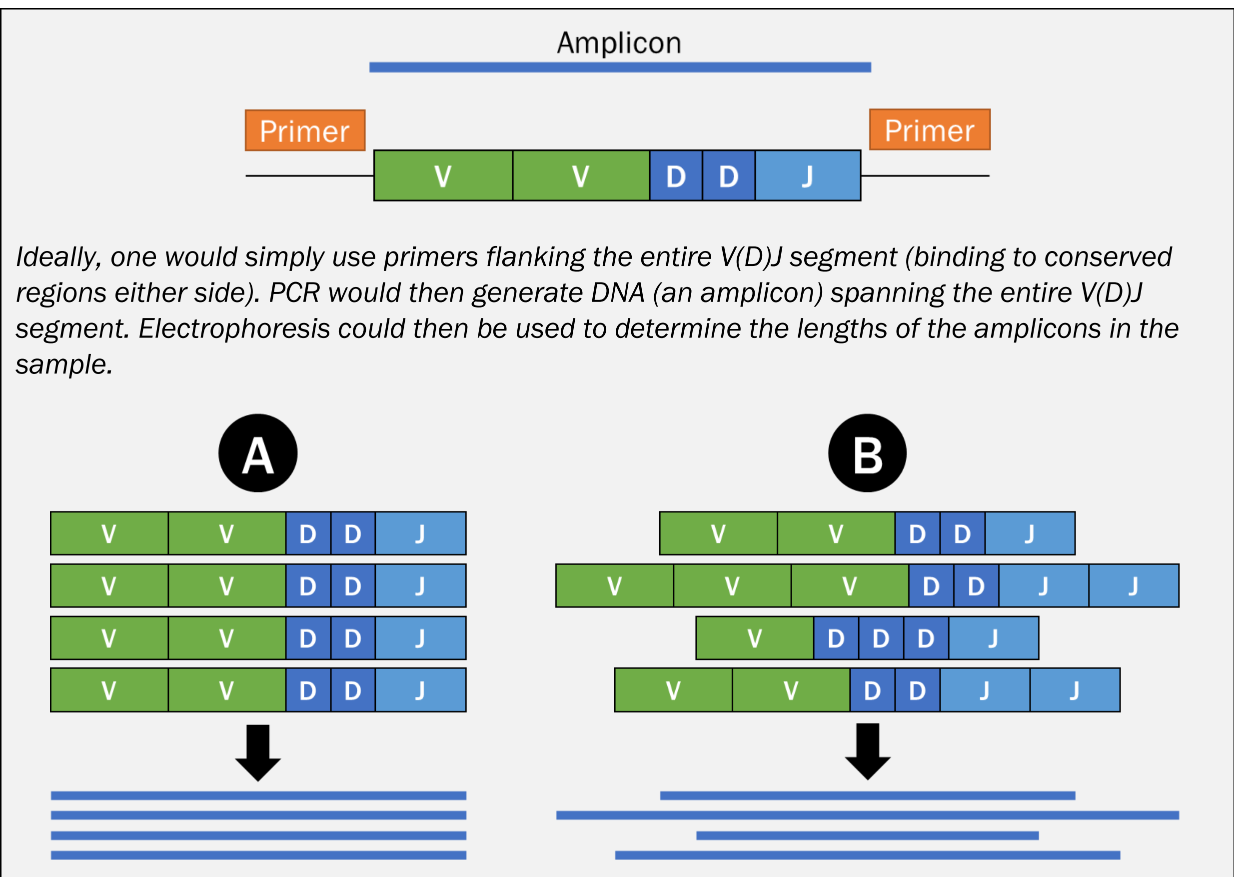
*The immunoglobulin genes are encoded in three loci: the heavy locus (*IGH*) which encodes the heavy chain; the kappa locus (*IGK*) which encodes the kappa light chain; and the lambda locus (*IGL*) which encodes the lambda light chain.*

*The T-cell receptor genes are encoded by four genes.* TCRA, TCRB, TCRG *and* TCRD *encode the α, β, γ and δ chains, respectively.*




*The* TCR *and* IG *genes undergo somatic recombination. During T-/B-cell development, parts of the genes are removed randomly to give rise to* TCR/IG *genes (and T-cell receptors/immunoglobulins) with different structures. This is also known as V(D)J recombination. The variable, diversity and joining segments are randomly deleted (although some chains lack diversity segments).*

- PCR is used to determine how similar in length the *IG/TCR* genes are in lymphoid populations



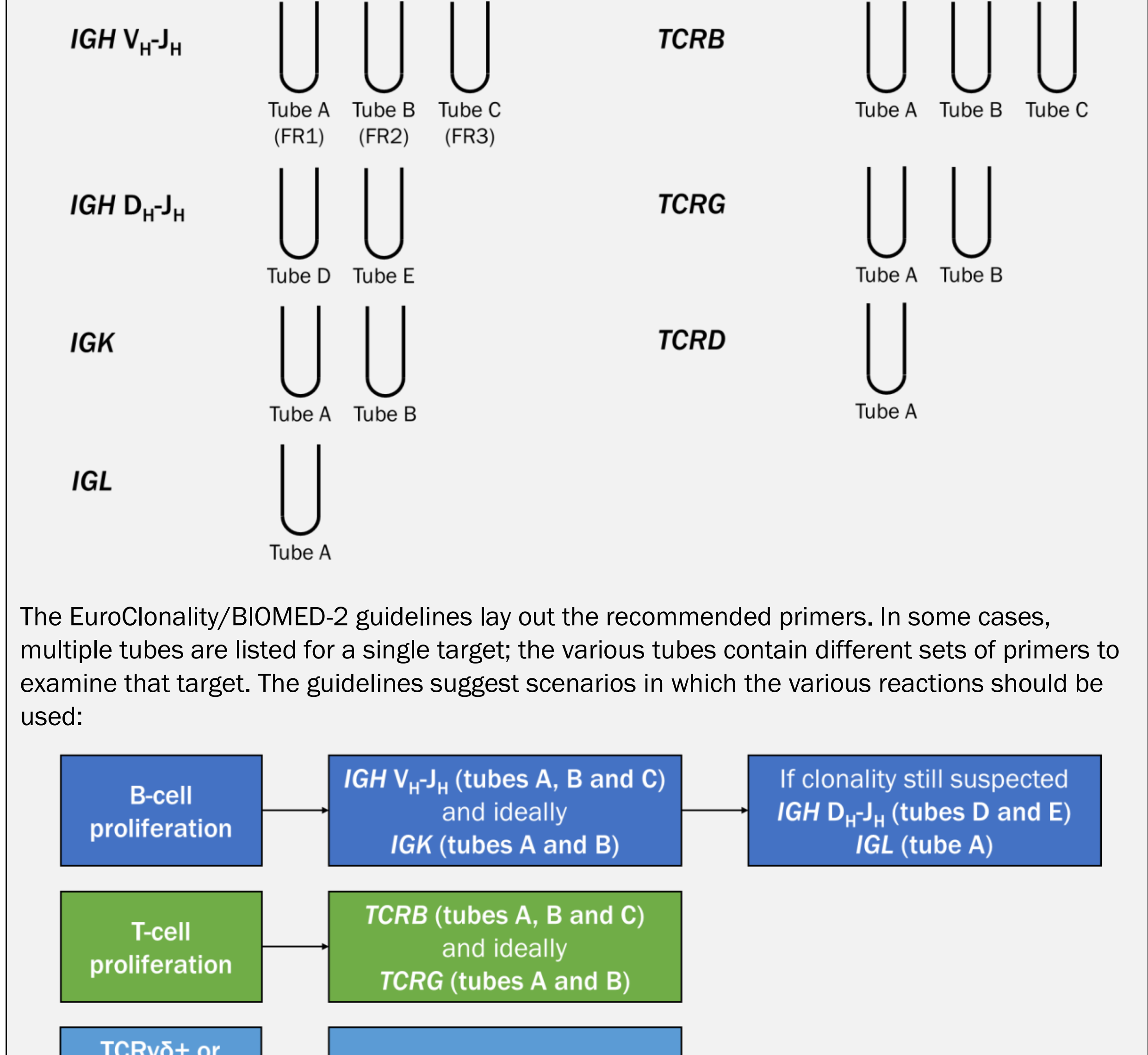

*Ideally, one would simply use primers flanking the entire V(D)J segment (binding to conserved regions either side). PCR would then generate DNA (an amplicon) spanning the entire V(D)J segment. Electrophoresis could then be used to determine the lengths of the amplicons in the sample.*

*A: All the V(D)J segments have rearranged in exactly the same way. Therefore PCR has generated amplicons which are all of the same length. This indicates that there is likely to be a clonal (i.e. neoplastic) lymphoid proliferation.*

*B: All the V(D)J segments have rearranged in different ways, giving rise to amplicons of different lengths. This is a polyclonal result, and supports the lymphoid proliferation being reactive.*

*Depending on the suspected nature of the lymphoid proliferation, clonality testing can be undertaken on different genes. If it is a B-cell proliferation,* IGH, IGK *and* IGL *can be assessed. If it is a T-cell proliferation,* TCRB, TCRG *and* TCRD *can be assessed.*

- The problem is that it is not possible to cover the entire V(D)J segment with primers either side because it is so large
- Multiple sets of primers must be used, each covering a short part of the entire region
- The EuroClonality/BIOMED-2 guidelines lay out the recommended primers to be used

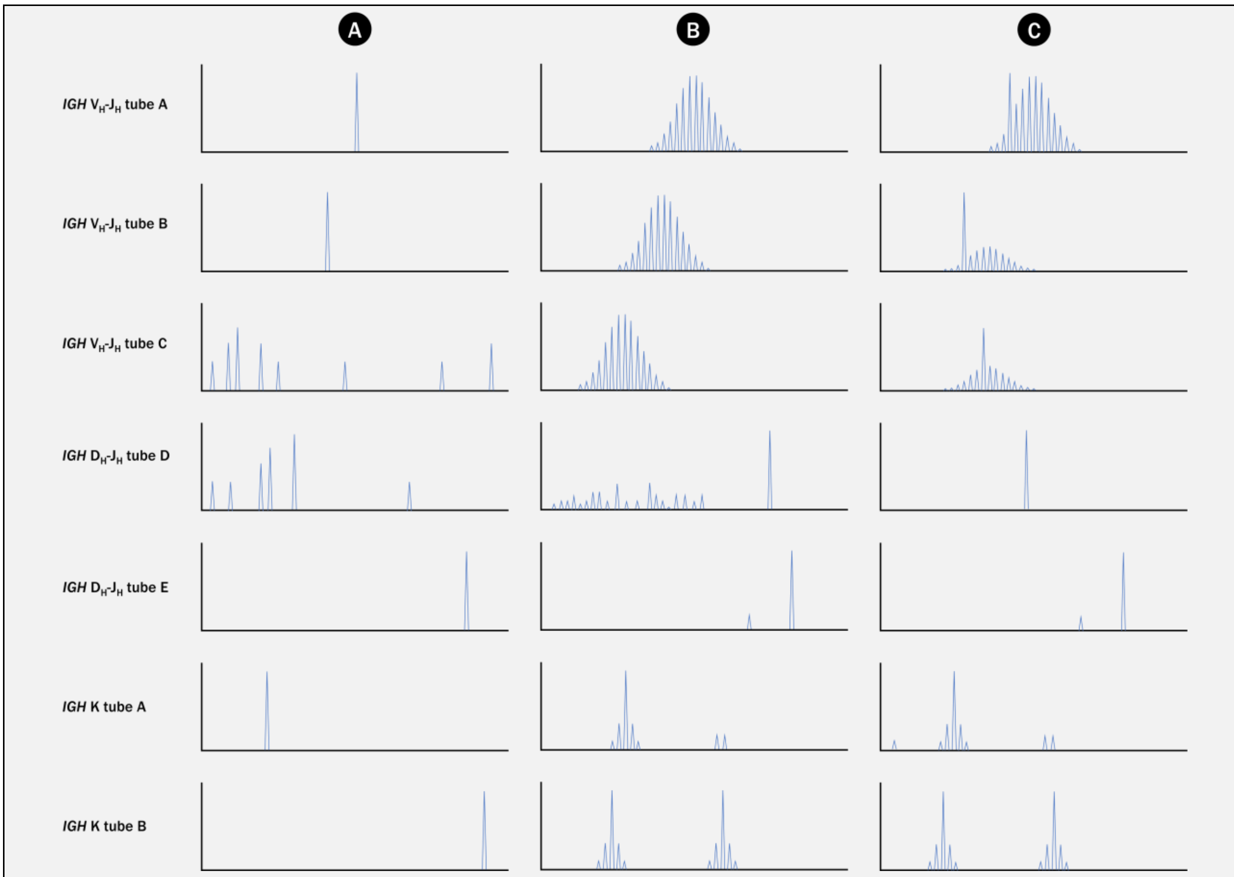



The EuroClonality/BIOMED-2 guidelines lay out the recommended primers. In some cases, multiple tubes are listed for a single target; the various tubes contain different sets of primers to examine that target. The guidelines suggest scenarios in which the various reactions should be used:



- The data output from clonality testing comprises a graph for each tube:
  - X-axis is amplicon length
  - Y-axis is intensity
- A polyclonal result will manifest as an (approximately) normal distribution:

  - The length of the V(D)J segment will vary, but most will be clustered around a median length with a long tail of shorter and longer segments
- A clonal result will manifest as a high peak
- A clonal result within a polyclonal background will manifest as a typical polyclonal pattern with a high peak:
  - This is seen when there is a neoplastic lymphoid proliferation within a reactive lymphoid proliferation (especially in T-cell lymphomas)




*A: This shows evidence of a B-cell clonal population. There is a high peak in* IGH *tubes A and B, and in* IGK *tubes A and B. There is what looks like a clonal peak in* IGH *tube E, but this is a recognised non-specific peak (which is listed in the manufacturer's documentation).*

*B: This shows no evidence of clonality. Most of the tubes show smooth normal distributions, consistent with a polyclonal lymphoid proliferation. Again, there are peaks in* IGH *tubes D and E, but these are listed by the manufacturer as being non-specific.*

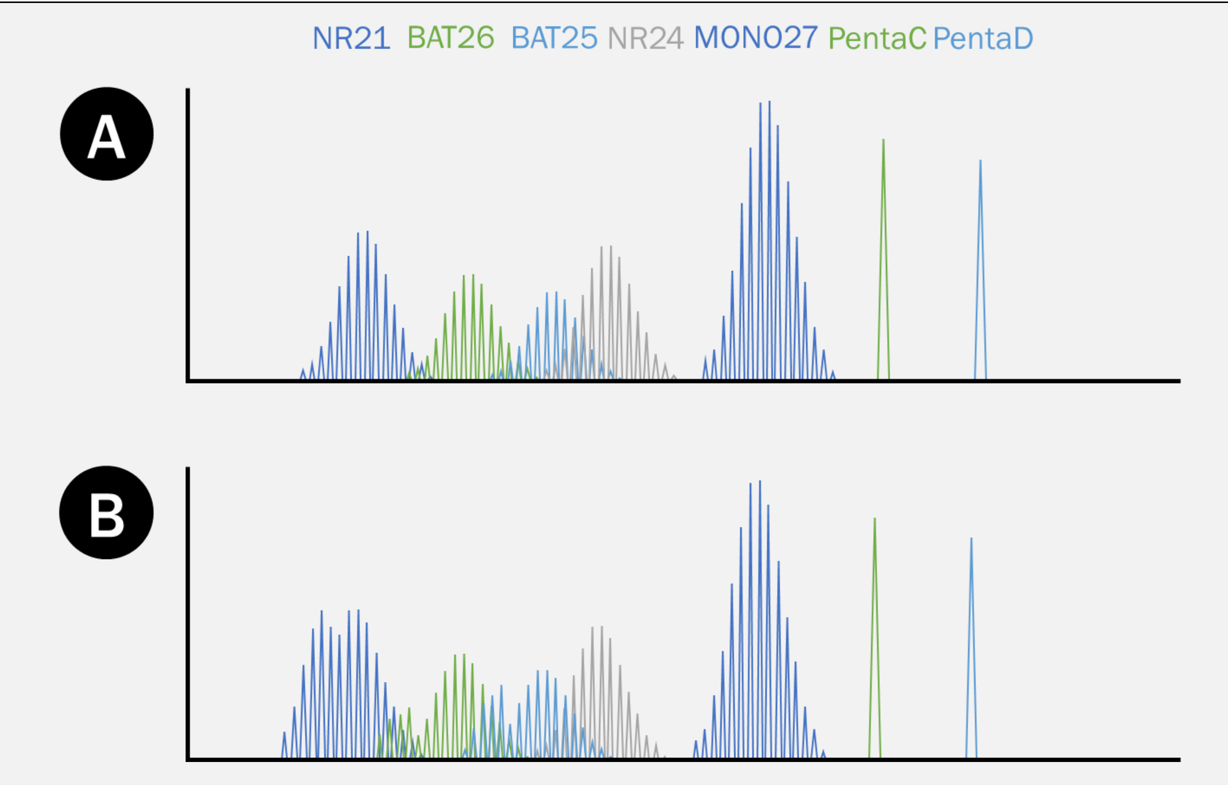
*C: This shows evidence of clonality in a polyclonal background. In* IGH *tubes A, B and C, there are convincing normal distributions containing high peaks. Again, there are non-specific peaks in* IGH *tubes D and E.*

- Clonality testing should be requested only where it has not been possible to determine whether a lymphoid proliferation is reactive or neoplastic on clinical/histological grounds
- The results should be interpreted with caution, and should never outright overrule a confident diagnosis made on clinical/histological grounds
- False negative results are possible:
  - If the neoplastic cell percentage is very low, the clonal signal may be missed (for this reason, clonality testing should never be used to explore the possibility of Hodgkin lymphoma)

  - Although the primers used have very good coverage, some rearrangements will invariably be missed
- False positive results are possible:
  - If the sample contains small numbers of T-/B-cells, falsely clonal results can be obtained ('pseudoclonality')
  - Some reactive proliferations can contain genuinely clonal lymphoid proliferations
- Clonality results must be interpreted in light of all available clinical and histological information

## Microsatellite analysis

- Microsatellites are repetitive units of a short sequence of DNA found normally throughout the genome
- Because they are repetitive, they are difficult to replicate accurately during DNA replication:
  - The replication machinery can 'slip' either introducing extra repeats or removing existing repeats
  - The mismatch repair (MMR) system normally identifies these slips and keeps the microsatellite length stable
  - If the MMR system is not functioning normally, the microsatellites are prone to expansion or contraction (microsatellite instability)
  - Microsatellite stability can therefore be used as a proxy marker for inferring the status of the MMR system
- All normal somatic cells harbour two alleles of each microsatellite (one on the maternal and one on the paternal chromosome):
  - They should normally be of the same length
  - In microsatellite instability (MSI), the lengths of the two alleles differ
- Ideally, one would assess the lengths of all microsatellites in the genome in order to determine whether MSI is present or not:
  - This is clearly not practical
  - Small panels of microsatellites have been found to be accurate markers of MSI more generally
- The Bethesda panel includes BAT25, BAT26, D2S123, D5S346 and D17S250
- The Promega panel includes BAT25, BAT26, NR21, NR24, MONO27, Penta C and Penta D
- There are many methods for assessing microsatellite length
- Most commonly, testing is through PCR-based methods:
  - Primers flanking the microsatellites are used
  - PCR is undertaken to generate amplicons which reflect the lengths of the microsatellites
  - The lengths of the amplicons are determined by electrophoresis
- The lengths of the amplicons are plotted graphically
  - Microsatellite stability manifests as a single peak
  - Microsatellite instability manifests as multiple peaks




*A: There is a single peak corresponding to each of the microsatellites assessed. There are no double peaks. There is therefore no evidence of microsatellite instability.*

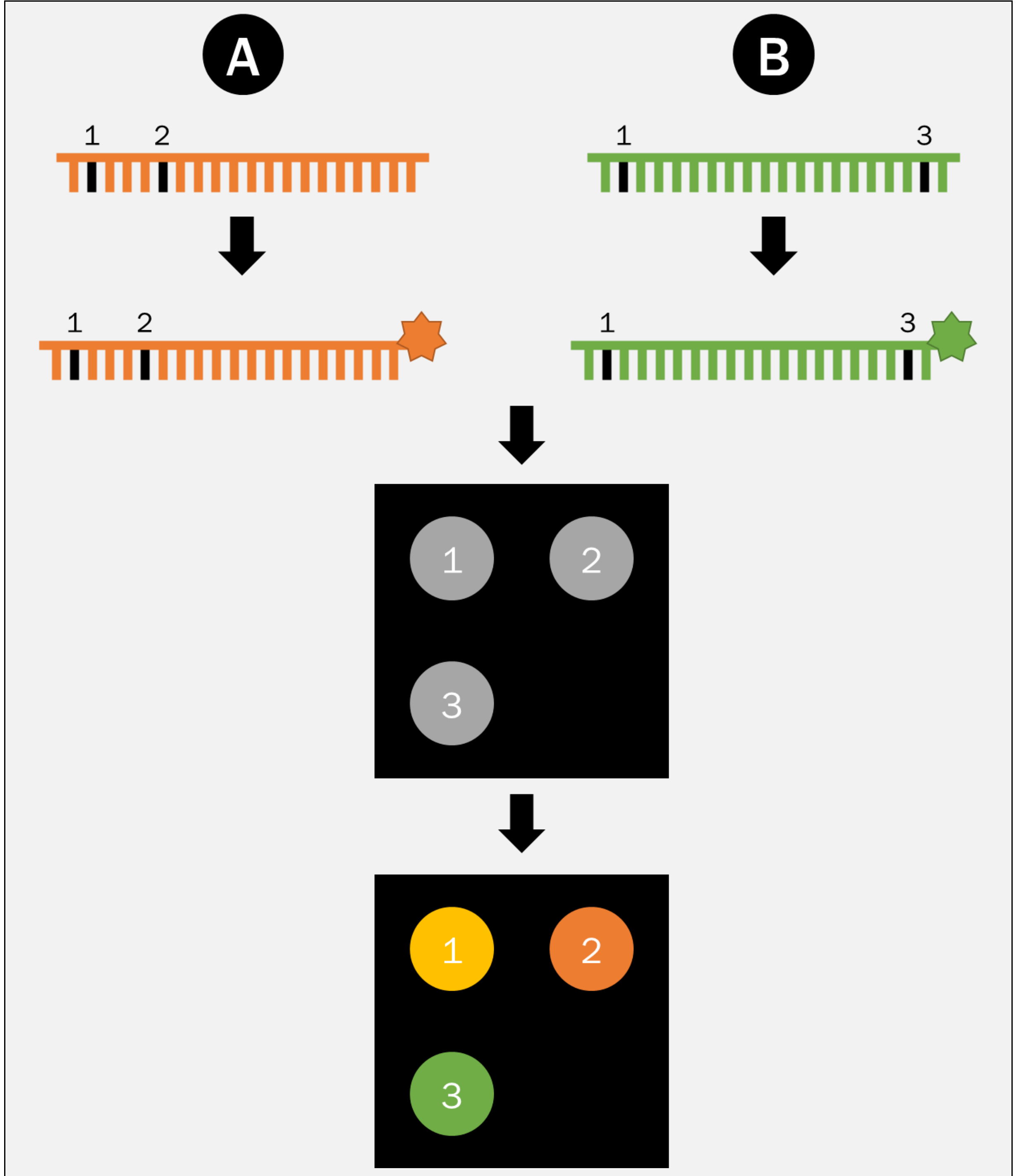
*B: NR24, MONO27, PentaC and PentaD are associated with single peaks and are therefore considered stable. NR21, BAT26 and BAT25, however, show double peaks indicating that they are unstable. Overall, this tumour is regarded as showing microsatellite instability (MSI).*

- Results are interpreted as follows:
  - The tissue is microsatellite stable (MSS) if none of the markers show instability
  - The tissue is considered microsatellite instability high (MSI-H) if at least two markers show instability
  - The tissue is considered microsatellite instability low (MSI-L) if one marker shows instability
- Whether MSI-L cases should be managed as MSS or MSI-H is unclear
- See MMR immunohistochemistry, above

## Microarrays

- Microarrays are glass chips:
  - Their surfaces are covered with spots, each of which contains a different DNA probe
  - Nucleic acids from the sample are extracted and hybridised against the DNA probes
  - Fluorescence is used to detect whether the nucleic acid has hybridised
  - The surface of the microarray is scanned and the fluorescence pattern analysed
- Microarrays serve two main purposes
- DNA microarrays are used to determine whether variants of interest are present:
  - By using many probes on the surface of the microarray, it is possible to assess for the presence of huge numbers of variants (across multiple genes if required)
  - The balance of different colours of fluorescence determines whether each variant is present or not

- Of course, this will only detect variants for which probes have been included on the microarray




*We are interested in knowing whether a sample contains variants 1, 2 and 3.*

*A: Nucleic acid is extracted from non-tumour control tissue. This contains variants 1 and 2, but not variant 3.*

*B: Nucleic acid is extracted from tumour tissue. This contains variants 1 and 3, but not variant 2.*

*The non-tumour and tumour samples are labelled with different types of fluorescence.*

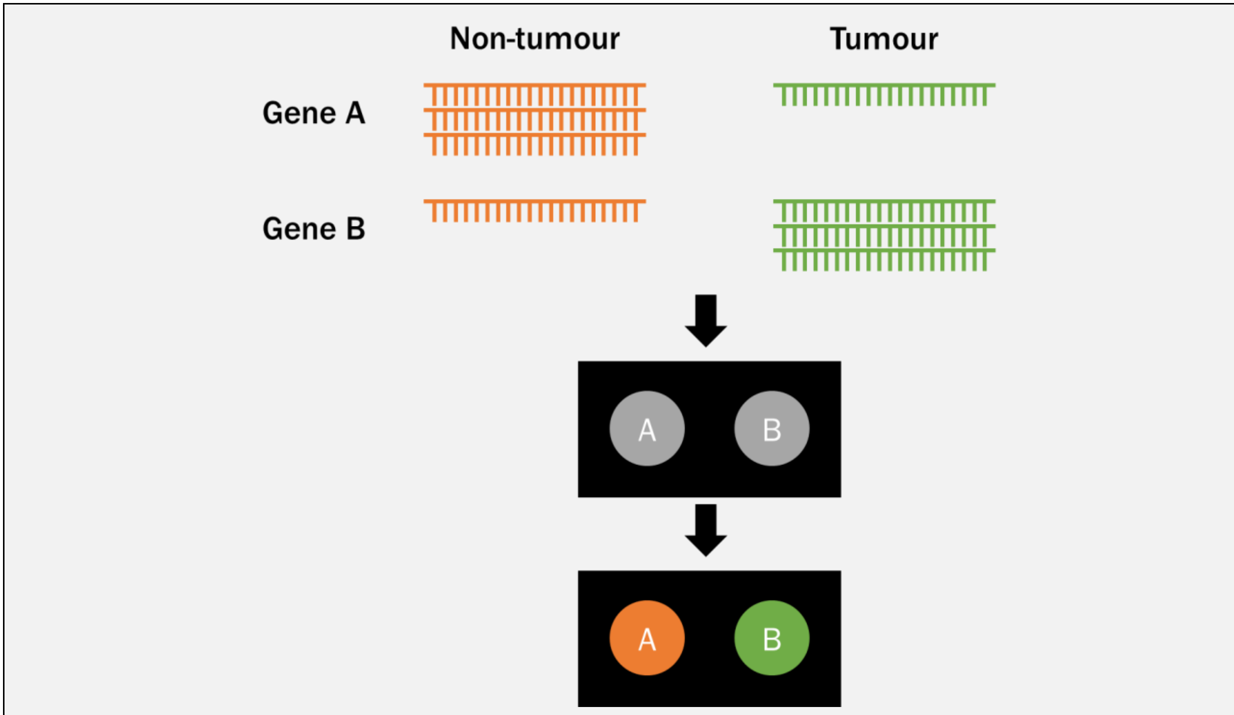
*The microarray contains three spots, one for each of the variants of interest.*

*Because variant 1 is present in both the non-tumour and tumour, nucleic acid from both samples hybridises against the probes. There is therefore both green and orange fluorescence, which results in a yellow signal. This means that the variant is present in both samples.*

*Because variant 2 is present only in the non-tumour tissue, only non-tumour nucleic acid hybridises against the probe. There is therefore only orange fluorescence at this spot.*

*Because variant 3 is present only in the tumour tissue, only tumour nucleic acid hybridises against the probe. There is therefore only green fluorescence at this spot.*

- RNA microarrays are used to determine levels of gene expression:
  - Instead of starting with DNA, this starts with extracted RNA which is converted to cDNA
  - The microarray bears probes which are complementary to the sequences of the genes whose level of expression needs to be determined
  - The type of fluorescence detected determines whether the gene is more highly expressed in non-tumour tissue, tumour tissue, or whether it is equally expressed in both




*RNA is extracted from both non-tumour tissue (control, orange) and tumour tissue (green), and is converted to cDNA. We want to know how highly genes A and B are expressed in tumour, relative to non-tumour tissue.*

*The microarray bears two spots, containing probes complementary to the sequences of genes A and B respectively.*

*The non-tumour (orange) expresses gene A more highly than the tumour (green). Therefore the spot for gene A fluoresces orange.*

*The tumour (green) expresses gene B more highly than the non-tumour (orange). Therefore the spot for gene B fluoresces green.*

*If expression of a gene is equivalent in both samples, fluorescence would be yellow.*

- Nowadays, microarrays are seldom used in solid tumour pathology in routine practice:

- DNA microarrays have largely been superseded by NGS
- RNA microarrays have largely been superseded by qRT-PCR

## Selecting the appropriate technique

- There will almost always be multiple different ways of assessing a given marker, and there is rarely a 'best' way of undertaking testing
- The best technique depends on:
  - The nature of the alteration
  - The quality of the sample
  - The required turnaround time
  - Whether other targets are required
- As a general rule, as the number of required markers increases, the argument for testing with NGS becomes stronger
  - With targeted testing, as the number of required targets increases, the number of individual tests required (and, depending on the tests, the amount of tissue required) increases
  - With NGS, all the required targets can usually be covered by one or two panels

### Techniques for assessing for small variants

| | Advantages | Disadvantages |
|---|---|---|
| **Direct/Sanger sequencing** | At least in theory, can detect any alteration present in the region sequenced<br>Fast if short region needs to be sequenced (slow if multiple regions)<br>Relatively cheap | Generally can only sequence one region at a time<br>Slow and expensive if multiple regions need to be sequenced<br>Generally needs a fairly pure tumour sample (> 40%)<br>Requires relatively large amounts of tissue if multiple targets need to be assessed |
| **Pyrosequencing** | At least in theory, can detect any alteration present in the region sequenced<br>Fast if short region needs to be sequenced (slow if multiple regions)<br>Relatively cheap<br>Fairly low tissue requirements (> 10% neoplastic cell percentage) | Generally can only sequence one region at a time<br>Slow and expensive if multiple regions need to be sequenced<br>Short read length can make it difficult to complete sequencing, especially in repetitive regions<br>Requires relatively large amounts of tissue if multiple targets need to be assessed |
| **Next-generation sequencing** | At least in theory, can detect any alteration present in the region sequenced<br>Easy to sequence multiple regions simultaneously<br>Relatively cheap and fast if multiple regions need to be sequenced<br>Cost-effective if high volumes of work are being sequenced<br>Can also be used to test for large-scale alterations<br>Intermediate tissue requirements (> 20% tumour cell percentage)<br>Consumes relatively little tissue if multiple targets need to be assessed | Requires high initial capital input<br>Requires considerable expertise<br>Relatively expensive and slow if only one/few regions need to be sequenced<br>Very expensive and required batching (slow) if low volumes of work are to be sequenced |

| **Real-time PCR** | Can detect alterations across multiple genes<br>Often very fast (hours)<br>Fairly low tissue requirements (> 10% neoplastic cell percentage)<br>Relatively little expertise needed<br>Relatively low initial capital input | Will only detect specific pre-determined alterations (with a limit to the total number of alterations which can be detected)<br>Becomes expensive if large volumes of work<br>Requires relatively large amounts of tissue if multiple targets need to be assessed |
|---|---|---|

- IHC may occasionally be useful for detecting small variants:
  - The stains are usually specific to particular variants
  - For example, BRAF V600E IHC will only detect the protein product of *BRAF* V600E, but will not detect other *BRAF* V600 variants
  - For example, IDH1 R132H IHC will only detect the protein product of *IDH1* R132H, but will not detect other *IDH1* variants

## Scenarios

1. I need to know whether there is a pathogenic variant in *EGFR* as quickly as possible.
   - *EGFR* is an oncogene and so has a fairly small number of possible activating variants in a small region of the gene.
   - Any of direct/Sanger sequencing, pyrosequencing, NGS or real-time PCR could be used.
   - If speed is what matters most, real-time PCR is probably the best option.
   - However, real-time PCR will inevitably miss a small number of variants which sequencing would detect.
   - If multiple targets needed to be assessed quickly, it would be more difficult to decide between real-time PCR and NGS.
2. I need to know whether there is a pathogenic variant in *EGFR*, *BRAF*, *KRAS* or *HER2*.
   - These are all oncogenes and have a fairly small number of possible activating variants.
   - Any of the above techniques could be used, in theory.
   - However, if direct/Sanger sequencing, pyrosequencing or real-time PCR are used, multiple reactions will likely be needed to cover all the genes (and for real-time PCR, a small proportion of variants will be missed).
   - Conversely, NGS will cover all these genes and so save tissue (and possibly money), but will probably take longer.
   - The default option here should probably be NGS unless there is a strong reason not to use it (e.g. need for a very rapid result).
3. I need to know whether this colorectal cancer has a variant in *KRAS*, *NRAS* or *BRAF* as quickly as possible.
   - This is more difficult.
   - These are all oncogenes and have a fairly small number of possible activating variants.
   - Any of direct/Sanger sequencing, pyrosequencing, NGS or real-time PCR could be used.
   - The need for a fast turnaround favours real-time PCR.
   - The fact that multiple genes need to be assessed favours NGS.
   - The decision should probably be based on the real-time PCR assays available and the amount of tissue available.
   - If there is very little tissue available and there is a real-time PCR assay which can assess more than one of these genes at once, real-time PCR is probably the better option.

   - Otherwise, NGS is likely the better option.
4. I need to know whether this tumour harbours a *BRCA1/BRCA2* variant.
   - *BRCA1/BRCA2* are tumour suppressor genes, and variants are diverse and scattered throughout the genes, rather than being limited to a small number of hotspots.
   - Real-time PCR is not appropriate because it would require a huge number of probes which cannot be incorporated into a single test.
   - Pyrosequencing is not ideal since it has short read lengths.
   - Direct/Sanger or NGS could be used, but NGS is favoured since it would allow multiple regions of both genes to be sequenced in parallel.
5. I need to know whether this NSCLC sample has a small variant in *EGFR*, *BRAF* or *KRAS*, or a structural variant involving *ALK*, *ROS1*, *RET* or *NTRK*.
   - There are generally two approaches.
   - A DNA and RNA NGS panel could be used to assess for all alterations in parallel.
   - A mixture of real-time PCR could be used to detect the small variants, and a mixture of IHC/FISH could be used to detect the structural variants.
   - Using NGS would generally consume less tissue, but would take longer and would require higher neoplastic cell percentage.
   - Using multiple single-gene tests (real-time PCR, IHC, etc.) would generally be faster and require neoplastic cell percentage, but would consume more tissue and miss a small proportion of alterations.
   - The decision depends the setting and clinical priorities.
6. My patient has exhausted standard treatment options and I want to know whether I can get access to off-label treatment or clinical trials.
   - The only feasible option here is NGS using a large gene panel or whole genome sequencing (WGS).
   - Arguably, multiple real-time PCR tests could be used to target multiple genes, but this will be very expensive and will still miss most variants present.
   - Direct/Sanger or pyrosequencing could theoretically be used, but a huge number of reactions would be needed to cover the equivalent of a large NGS panel.

## Techniques for assessing for structural variants

| | **Advantages** | **Disadvantages** |
|---|---|---|
| **IHC** | Very fast<br>Requires very little tissue<br>Requires very few tumour cells<br>Generally fairly cheap<br>Very easy to implement in a histopathology laboratory | Only works for certain genes<br>Depending on the gene, may be associated with false positive or false negative results (some cases may need confirmation by an alternative technique)<br>Can generally only be used to assess one gene at a time |
| **Reverse transcription PCR** | Fairly fast<br>Requires moderate amounts of tissue and moderate neoplastic cell percentage | May produce false negative results<br>Can generally only be used to assess one gene at a time (although some assays assess multiple genes at once) |
| **FISH** | Fairly fast<br>Requires moderate amounts of tissue | Requires expensive fluorescence microscope<br>Requires considerable expertise<br>May produce either false positive or false negative results<br>Can generally only be used to assess one gene at a time |

| **Next-generation sequencing** | Requires relatively little tissue if multiple genes need to be assessed<br>Thought to be reliable for most genes<br>Tissue-efficient if multiple genes are being assessed | Fairly slow<br>Requires fairly high neoplastic cell percentage (> 20%)<br>High capital input required<br>Requires considerable expertise<br>DNA panel: higher chance of false negative results<br>RNA panel: requires separate RNA extraction if small variants also being assessed |
|---|---|---|

- IHC can be used only for certain genes:
  - For *ALK* in NSCLC it is exceptionally reliable. Neither negative nor positive results require confirmation.
  - For *ROS1* in NSCLC a negative result essentially excludes the possibility of a gene rearrangement. However, it is poorly specific and positive results require confirmation.
  - For *NTRK* in NSCLC, pan-TRK IHC is available but its reliability remains uncertain.
  - For *RET* in NSCLC, IHC is available but its reliability remains uncertain.

## Scenarios

1. I need to know whether an NSCLC has an *ALK* rearrangement.
   - Only one gene needs to be examined.
   - RT-PCR, FISH or NGS could therefore be used.
   - IHC is extremely reliable for detecting ALK rearrangements in NSCLC.
   - If *ALK* rearrangements alone are required, IHC is probably the best option – it is fast, accurate and requires little tissue.
   - RT-PCR or FISH are also suitable possibilities.
   - FISH is probably best if neoplastic cell percentage is very low.
   - NGS is a good option, but is likely to be slowest and is less efficient if only one target is needed.
2. I need to know whether my tumour has an *ALK*, *ROS1*, *RET* or *NTRK* rearrangement.
   - FISH would not be a good option: it would require six separate FISH assessments (included one for each *NTRK* gene). This would require a large amount of time and tissue.
   - RT-PCR could be a reasonable option, particularly if a single assay can cover more than one of these genes. It would probably be faster than NGS but it would miss rare fusions.
   - IHC would be a good option for *ALK* and a reasonable option for *ROS1*, but it is unclear how useful it would be for *RET* and *NTRK*.
   - Because multiple targets are needed, NGS is the best option. An RNA panel would be superior to DNA for detecting fusions.
   - If a very fast result is needed or if the neoplastic cell content is very low, some combination of IHC, FISH and/or RT-PCR may be chosen instead of NGS,

## Techniques for assessing for copy number variants

- The decision is generally between ISH and next-generation sequencing
- At present, ISH tends to be favoured for CNV detection over NGS:
  - Most NGS assays in routine clinical use are believed not to be entirely reliable for the detection of CNVs
  - In some cases, guidelines explicitly advise against NGS (e.g. *HER2* amplification in breast cancer)
- Immunohistochemistry may be used in some cases:

- HER2 IHC is a reliable surrogate for *HER2* amplification testing at the extremes, but with intermediate levels of staining confirmatory ISH is required

# Provision of molecular pathology services

## Turnaround times

- For many cancers, systemic treatment cannot begin until the results of molecular testing are known:
  - A small proportion of patients will die while waiting for results
  - A small proportion of patients will deteriorate while waiting for results to the point where they are no longer fit for treatment
  - Patients may be booked into a clinic to discuss treatment options, only to find that results are not available – this wastes patient and clinician time if the appointment goes ahead, and wastes administrative time if the appointment needs to be deferred
  - The wait to find out about prognosis and treatment is agonising for patients
- There is therefore huge pressure to improve turnaround times for molecular testing
- In-house testing is almost always faster than outsourced testing:
  - Time is taken preparing tissue for dispatch, physically transporting the tissue, booking it into the receiving department, and getting the report onto the referring department's LIMS
- Batch testing technologies:
  - Many technologies have defined 'runs'
  - Cases are placed on a single run, and then run in a batch
  - For these technologies, cost per case is extremely high if few cases are included on a run, and drops dramatically as the number of cases per run increases
  - Therefore, a minimum workload is required for these technologies to make them economical
  - Too low a workload means either: (1) runs are not used optimally and costs are extremely high, or (2) cases are 'saved up' until there are sufficient numbers to make a run economical, which causes substantial delays
  - Therefore, with batch testing approaches, some degree of centralisation of testing is usually required to ensure the workload is sufficient for testing to be economical and fast
  - In-house testing with batch testing technologies may not deliver time savings compared to outsourced testing if batching is necessary
- Continuous testing technologies:
  - With some technologies, cases are added to the testing process one-by-one
  - Here, there is no real advantage to having centralised testing and testing could be instituted in-house
  - There still needs to be a minimum workload to ensure appropriate expertise and to justify the time/cost of validation, verification, SOP writing, etc.
- Molecular testing comes at the end of a long process of tissue acquisition and diagnosis, and improved turnaround times at any point in the process may speed the provision of molecular results:
  - Prioritising in the histopathology laboratory specimen types which are likely to require molecular testing (e.g. breast and lung biopsies)
  - For specimen types likely to require diagnostic IHC, cutting a few spare sections at the time of cutting the initial H&E to avoid delays in waiting for IHC sections later (although this will waste tissue in cases where diagnostic IHC is not needed)

- Requesting molecular testing automatically on diagnosis of particular entities, without waiting for the oncologist or MDT request ('reflex' testing)
- Automatically sending the results of molecular testing to the referring centre by email, rather than sending out paper reports by post
- Extended day and weekend working minimise laboratory downtime

- Reflex testing:
  - For cancer types where there is a high likelihood of molecular results being needed and where those results are needed quickly, it may be prudent for the pathologist to request testing as soon as they make the diagnosis
  - This is often done for diagnoses of non-small cell lung cancer, where patients often present at high stage (and so need systemic therapy) and where they rapidly deteriorate (and so need results quickly)
  - However, a proportion of results obtained through reflex testing will never be useful (e.g. early stage cancers which are resected, patients who are not fit for treatment)
  - It is for local services to balance the benefits from improved turnaround times against wasted resources from unneeded testing, and to decide whether reflex testing is appropriate
- Different technologies clearly have different intrinsic turnaround times:
  - Targeted testing is generally faster
  - NGS is generally slower, particularly with large panels
  - One approach is to use NGS as the default testing technique, and to supplement this with targeted testing when a rapid result is required (either alongside or instead of NGS)

## Failure rates

- Failed testing can occur for two reasons:
  - Insufficient quantity of tumour tissue to complete testing
  - Insufficient quality of proteins or nucleic acids to complete testing
- Failed testing is clearly problematic:
  - If the failed sample is the only sample available, it precludes the possibility of molecular profiling unless a repeat sample can be obtained (with the associated risks, delays and costs)
  - Even if repeat testing can be performed on a different sample, the need to repeat testing introduces a considerable delay
- The sooner it becomes clear that testing will fail, the better:
  - Pathologists can alert clinicians to the fact that a sample is suboptimal for testing before testing even starts – clinicians can therefore put a plan in place in case additional tissue is required
  - In some settings, extracted nucleic acids undergo quality assessment before testing starts – failures therefore become clear early in the process, and repeat sampling can be instigated more quickly
- Preserving tissue for molecular testing is essential:
  - Where multiple biopsies are received for a tissue type frequently requiring molecular testing (e.g. breast, lung), they should ideally each be processed in a separate cassette (or at least split over a couple of cassettes) – this means that one block can be used for diagnostic IHC, and another held in reserve for molecular testing
  - For specimen types likely to require diagnostic IHC, cutting a few spare sections at the time of cutting the initial H&E to avoid delays in waiting for IHC sections later (although this will waste tissue in cases where diagnostic IHC is not needed)

  - Routine cutting of levels should be avoided in specimen types which are likely to require molecular testing, unless this is likely to be required to make a prompt diagnosis
  - Diagnostic IHC should be used judiciously – this is easier if clinicians provide comprehensive clinical information
- Preserving the integrity of proteins and nucleic acids is essential:
  - Ischaemic time needs to be minimised – this is easier for biopsies which simply need to be placed in formalin, but for resections prompt and thorough slicing/opening/inflation is essential to halt autolysis and putrefaction (this may not happen for specimens arriving late in the day, especially on a Friday)
  - There needs to be enough formalin fixation to preserve proteins but not so much as to degrade nucleic acids – for a biopsy, a minimum of 6 and a maximum of 12-24 hours is often considered optimal
  - Extended day working reduces the time specimens spend in formalin – if a biopsy reaches the laboratory at 4:30 PM, it will almost certainly have to sit overnight in formalin, but if the laboratory stays open until 7 PM, it may make it onto a processor the same day
  - Weekend working reduces the time specimens spend in formalin – if a biopsy reaches the laboratory at 4:30 PM on a Friday, it is highly likely that it will sit in formalin until Monday morning (around 60 hours)
  - Decalcification should be undertaken using nucleic acid-kind solutions and only strictly for the duration required – if a bony cancer resection is received, it may be wise to remove the soft tumour to be processed without decalcification, and to place the remainder of the specimen in decalcification (assuming that this would not hamper histological assessment)
- The use of fixatives other than formalin could be considered:
  - The ideal fixative will preserve morphology, protein and nucleic acids
  - Most such fixatives are, however, extremely expensive
  - The use of non-formalin fixatives will generally also need to be validated

## External quality assurance (EQA)

- EQA for non-genomic tests is provided by UKNEQAS
- EQA for genomic tests is provided by GenQA

## Historical provision of molecular pathology services in England

- Testing was provided in a variety of locations:
  - Some departments performed some/all testing in-house
  - Some departments sent work away to other histopathology departments (sometimes in other regions) for some/all testing
  - Some departments sent work away to genetics departments for some/all testing
- Testing was provided using a variety of technologies:
  - Targeted testing using IHC, direct/Sanger sequencing, pyrosequencing, real-time PCR, ISH, etc.
  - NGS panels
- The costs of testing were borne either by pathology services or by local cancer services
- With the consolidation of pathology services driven by NHS Improvement, there was a drive to bring molecular testing in-house to reduce turnaround times and costs

# Current/future provision of molecular pathology services in England

- Provision of non-genomic tests will continue as at present:
  - HER2 amplification by either IHC or ISH
  - MMR protein expression
  - PD-L1 expression
  - IHC surrogates of genomic alterations (e.g. ALK and ROS1 IHC)
- These tests can continue to be performed wherever the local pathology department decides (i.e. in-house or outsourced) and the costs borne as at present

## Genomic testing services

- All other (genomic) tests are now commissioned centrally, i.e. NHS England pays for them directly rather than the funding coming from local budgets
- The tests which NHS England will fund are listed in the National Genomic Test Directory for Cancer which also stipulates:
  - The technology/technologies by which testing can be performed
  - The situations in which testing can be requested
- The Test Directory is delivered by seven Genomic Laboratory Hubs (GLHs) which cover England:
  - Central and South Genomic Laboratory Hub led by Birmingham Women's and Children's NHS Foundation Trust
  - East Genomic Laboratory Hub led by Cambridge University Hospitals NHS Foundation Trust
  - North West Genomic Laboratory Hub led by Manchester University NHS Foundation Trust
  - North Thames Genomic Laboratory Hub led by Great Ormond Street Hospital for Children NHS Foundation Trust
  - South East Genomic Laboratory Hub led by Guy's and St Thomas' NHS Foundation Trust
  - South West Genomic Laboratory Hub led by North Bristol NHS Trust
  - North East and Yorkshire Genomic Laboratory Hub led by The Newcastle upon Tyne Hospitals NHS Foundation Trust
- The general thrust of the Test Directory is that:
  - Testing should be delivered by NGS panels wherever possible
  - Targeted testing methods are available for some tests but this should be used only where necessary
- Testing is usually performed in centralised genetics laboratories to achieve economies of scale, although the GLHs may subcontract other laboratories to test on their behalf (including histopathology laboratories) – these are Local Genomic Laboratories (LGLs)
- NHS England will not fund:
  - Tests not listed on the Test Directory
  - Tests requested outside the indications stipulated in the Test Directory
  - Tests listed on the Test Directory, but performed outside the GLH structure
- Commons tests which generally be provided by all GLHs, but some tests are so niche that they may be offered only by one or a few GLHs:
  - Material can still be sent to the local GLH, which will usually extract nucleic acids and forward this to a GLH which is able to provide testing
  - The entire Test Directory may not be provided by every GLH, but the idea is that the GLH system as a whole should be able to deliver the whole Test Directory
- The Test Directory is updated annually, and anybody can propose new additions

### Using Genomic Laboratory Hubs

- Histopathology departments will need to send material to their local GLHs for testing
- Sections will need to be cut in-house according to the GLH's requirements
- A local pathologist will need to assess an H&E-stained section for some combination of neoplastic cell percentage, overall cellularity and percentage area necrosis
- The sections and marked H&E-stained section will be sent to the GLH, with a completed request form
- The GLH will return a report

### The advantages of GLHs

- All GLHs must deliver the same Test Directory, so all patients across England have access to the same tests
- Every patient in England now has access to testing which is funded by NHS England:
  - This means that no matter where the testing is initiated, it will be funded by the same source
  - This contrasts to the historical situation: when a case was referred to an outside centre which arranged testing, there would be complex billing arrangements to make sure that the patient's referring centre picked up the bill – this is no longer necessary
- By centralising testing:
  - They concentrate expertise
  - They result in heavy workloads, making NGS feasible, with minimal need to batch testing (which harms turnaround times)
  - They standardise testing

### The disadvantages of GLHs

- Sending work away is inevitably slower than performing testing in-house
- They remove (to a large extent) the freedom for pathologists to decide the best place and technology for testing
- Histopathology departments are now being asked to prepare and assess sections but are not funded for this work (in the past it was often possible to pass these costs on to cancer services)
- If histopathology laboratories would like to maintain in-house capabilities for single-gene testing, they will need to self-fund (unless they can be subcontracted by a GLH)
- The different GLHs have a good deal of autonomy which has resulted in the development of quite different approaches across England – this brings into question the stated aim of the system to provide equity of access

# Specific molecular tests in solid cancers

| | | |
|---|---|---|
| | | Describes molecular tests currently performed on histological samples.<br><br>Interprets the common molecular tests.<br><br>Demonstrates knowledge of molecular tests currently performed on histological samples, including the limitations of these tests and of tests that are anticipated in the near future.<br><br>Demonstrates ability to understand origins of, and justifications for, molecular tests. |
| | | Describes and explains common molecular tests including some of the common pitfalls and how to avoid them.<br><br>Illustrates the significance of common molecular tests.<br><br>Describes molecular tests currently performed on histological samples, including the limitations of these tests and of tests that are anticipated in the near future.<br><br>Demonstrates the origins of and justifications for molecular tests. |

## Molecular tests used in multiple tumour types

### *BRAF* small variant testing

- *BRAF* encodes a protein which transduces signals from receptor tyrosine kinases (e.g. EGFR) as part of the MAPK/ERK pathway
- Activating variants cause the protein to be constitutively active and result in uncontrolled cell proliferation (i.e. it is an oncogene)
- The vast majority of variants occur in codon 600, with the substitution c.1799T>A (resulting in V600E) being by far the commonest:
  - Less common variants are also seen in codon 600 (e.g. V600K, V600R, V600D)
  - Uncommonly, variants are seen outside codon 600 – generally, their significance is uncertain
- The range of clinically relevant *BRAF* variants is so small that targeted testing methods are entirely adequate for their detection, but sequencing can also be used
- Immunohistochemistry can be used to detect the BRAF V600E protein:
  - It manifests as cytoplasmic staining in cells which express the protein product of the mutated gene
  - It will be negative if there is no *BRAF* variant, or if there is a *BRAF* variant other than V600E
  - It is therefore useful if positive, but negative results need confirmation with an alterative technique to avoid missing non-V600E variants (e.g. in melanoma, where around a quarter of *BRAF* variants are not of V600E type)
- Targeted methods will invariably miss rare variants, but there are so few clinically relevant variants in *BRAF* that this is not a significant problem:
  - Real-time PCR
- Sequencing approaches should reduce the (already tiny) risk of missing variants, but will often have longer turnaround times:

  - Direct/Sanger sequencing
  - Pyrosequencing
  - Next-generation sequencing (NGS)
- NGS becomes a more attractive prospect if multiple gene need to be assessed in parallel

## *BRCA1/BRCA2* variant and HRD testing

- Double-strand DNA breaks are amongst the most harmful forms of DNA damage
- They are repaired through homologous recombination which involves a large number of proteins, including BRCA1 and BRCA2
- Homologous recombination deficiency (HRD) is the state in which homologous recombination does not occur properly:
  - Double-strand DNA breaks accumulate
  - Severe DNA damage rapidly accumulates
  - There is a high risk of cancers developing
- There are many causes of HRD:
  - Inactivating variants in *BRCA1* or *BRCA2*
  - Epigenetic inactivation of *BRCA1* or *BRCA2* (e.g. promoter methylation)
  - Alterations affecting the genes encoding any other proteins involved in repair of double-strand DNA breaks
- HRD results in a characteristic pattern of abnormalities in the genome ('HRD signature'), including:
  - Telomeric allelic imbalance
  - Large scale state transitions
  - Loss of heterozygosity
- There are therefore two ways of determining whether HRD is present:
  - One can test for *BRCA1* and *BRCA2* variants, as causes of HRD
  - One can test for the effects of HRD, and therefore infer the presence of HRD
- The problem with looking for *BRCA1* and *BRCA2* variants to determine whether HRD is present is that it will miss cases of HRD (i.e. where it has arisen through epigenetic alterations or through abnormalities of non-*BRCA* genes)
- Some tumours which show HRD respond well to PARP inhibitors:
  - PARP repairs single-strand DNA breaks
  - Single-strand DNA breaks on their own increase the likelihood of developing cancer
  - However, when single-strand and double-strand DNA breaks accumulate together, the cell sustains so much DNA damage that it cannot survive ('synthetic lethality')
  - This is why tumours with HRD generally respond to PARP inhibition
- *BRCA* variants may be somatic or germline:
  - A variant detected in tumour tissue may be either somatic or germline
  - A variant detected in non-tumour tissue (usually blood) will generally be germline
  - Tumour and germline *BRCA* testing should either be done in parallel or, if a variant is detected in tumour, it should be followed up with germline testing
- *BRCA1* and *BRCA2* are tumour-suppressor genes, in that they normally oppose the development of cancers:
  - Pathogenic variants are therefore fairly diverse and distributed quite widely through the genes
  - Targeted testing approaches will miss too many variants
  - Nowadays, testing is likely performed using next-generation sequencing
- HRD testing is invariably undertaken using next-generation sequencing

## *EGFR* small variant testing

- *EGFR* encodes a receptor tyrosine kinase
- Growth factor binding activates EGFR
- Activation results in dimerisation and autophosphorylation, which triggers signalling through the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- Activating small variants cause constitutive EGFR activation in the absence of ligand binding, causing uncontrolled cell growth and proliferation (i.e. it is an oncogene)
- Activating *EGFR* small variants are clustered in small hotspots in exons 18-21
- Although substitutions are common, more complex indels are not infrequently seen
- Exon 19 deletions and L858R are the commonest variants
- Targeted methods will invariably miss rare variants, but because the clinically relevant *EGFR* variants are fairly few in number this is generally not a big issue:
    - Real-time PCR
- Sequencing approaches should reduce the risk of missing variants, but will often have longer turnaround times:
    - Direct/Sanger sequencing
    - Pyrosequencing
    - Next-generation sequencing (NGS)

## *FGFR* alterations

- The *FGFR* genes constitute a family which encodes related receptor tyrosine kinases
- On growth factor binding, the receptor dimerises and autophosphorylates, resulting in activation of the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- A variety of activating small variants, structural variants and copy number variants can affect the *FGFR* genes
- Small variants cause constitutive activity in the MAPK/ERK pathway:
    - The small variants are clustered into a few small hotspots
    - Targeted approaches may be appropriate
    - However, if variants involving more than one gene are important (and especially if structural variants and copy number variants are important), next-generation sequencing will generally be the most appropriate technique
- Structural variants involve various fusion partners, which differ depending on the tumour type:
    - This results in overexpression of the *FGFR* gene, and results in the production of an FGFR protein which is constitutively active
    - This leads to intense signalling through the MAPK/ERK pathway, resulting in uncontrolled cell proliferation
    - FISH may be used to detect structural variants, but particularly *FGFR2* rearrangements are small-scale, meaning that there is a risk of false negative results with FISH
    - Reverse transcription PCR can be used but will miss some variants
    - Next-generation sequencing is generally preferred
- Increases in *FGFR* copy number result in increased expression of the FGFR protein, which makes the cell more sensitive to pro-growth signalling:
    - FISH or next-generation sequencing may be used
- Immunohistochemistry is generally not regarded as a reliable surrogate marker for *FGFR* alterations

## *HER2* copy number variant testing

- *HER2* (or *ERBB2*) is located on chromosome 17 and encodes the receptor tyrosine kinase HER2/neu
- Unlike other receptor tyrosine kinases, it has no known ligand and instead appears to be activated by other receptor tyrosine kinases
- Activation results in dimerisation and autophosphorylation, which triggers signalling through the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- Increases in *HER2* copy number result in increased expression of the HER2 protein, which makes the cell more sensitive to pro-growth signalling (i.e. it is an oncogene)
- In general, *HER2* copy number can be assessed directly by counting the number of copies of the gene in the DNA, or indirectly by assessing the level of expression of the HER2 protein
- Direct assessment of *HER2* copy number in DNA is usually undertaken with ISH (generally FISH or ddISH):
  - One probe is hybridised against *HER2* and another against *CEP17* (at the centromeric region of chromosome 17, on which *HER2* is located)
  - The mean *HER2* copy number is the average number of copies of *HER2* in each cell, which should normally be two
  - *HER2*:*CEP17* ratio is the ratio of copies of *HER2* to *CEP17* across all the cells assessed
  - Assessment of both indices allows *HER2* gene amplification to be distinguished from chromosome 17 polysomy
  - The bar for calling a tumour *HER2*-positive is generally higher when *HER2* gains result from polysomy than from when they result from amplification
- ISH is, however, relatively costly and time-consuming:
  - FISH in particularly requires expensive fluorescence microscopy and a dark room
- Indirect assessment of HER2 protein expression by IHC can be used as a surrogate marker for *HER2* gene copy number variants:
  - High levels of *HER2* gain invariably result in HER2 protein overexpression (3+ staining)
  - Lack of *HER2* gain results in no or slight HER2 protein expression (0 or 1+ staining)
  - Intermediate degrees of staining (2+) are equivocal and require confirmatory ISH
  - Scoring of HER2 IHC varies by tumour type (and specimen type, in some cases) but is usually based on the proportion of cells showing staining, the intensity of staining, and the degree of completeness of membranous staining
  - Using HER2 IHC as a surrogate marker helps to reduce the number of cases requiring ISH
- Next-generation sequencing may be used to assess for copy number variants, but most commonly used assays are not considered fully reliable – NGS is generally not recommended for assessment for *HER2* copy number variants

## *KIT* small variant testing

- *KIT* encodes a receptor tyrosine kinase
- Growth factor binding activates KIT
- Activation results in dimerisation and autophosphorylation, which triggers signalling through the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- Activating small variants cause constitutive KIT activation in the absence of ligand binding, causing uncontrolled cell growth and proliferation (i.e. it is an oncogene)

- Variants are clustered in exons 11, 9, 13 and 17
- Unlike other clinically relevant oncogenes, many of the variants in these hotspots are varied and complex
- Given the diversity of complex variants, targeted techniques would miss an unacceptable proportion of variants
- Sequencing approaches are the only testing paradigms commonly used:
  - Direct/Sanger sequencing
  - Pyrosequencing
  - Next-generation sequencing (NGS)
- Immunohistochemistry for KIT is not a reliable surrogate marker for underlying *KIT* variants

## *KRAS*, *HRAS* and *NRAS* small variant testing

- The *RAS* gene family includes *KRAS*, *HRAS* and *NRAS*; they all have similar structure and function
- All three genes encode a protein which transduces signals from receptor tyrosine kinases (e.g. EGFR) as part of the MAPK/ERK pathway
- Activation of the protein drives cell survival, growth and proliferation
- Activating variants result in uncontrolled cell proliferation (i.e. they are oncogenes)
- Activating small variants occur in a small number of well-defined hotspots:
  - *KRAS*: most commonly codons 12, 13, 61 and 146
  - *NRAS*: most commonly codons 61, 12 and 13
  - *HRAS*: most commonly codons 61, 12 and 13
- The vast majority of activating small variants are substitutions
- Because there is a limited number of relevant variants and they are generally easy to detect, targeted testing or sequencing approaches may be used
- Targeted methods will invariably miss rare variants, but the odds of this are extremely low with these genes:
  - Real-time PCR
- Sequencing approaches should reduce the (already small) risk of missing variants, but will often have longer turnaround times:
  - Direct/Sanger sequencing
  - Pyrosequencing
  - Next-generation sequencing (NGS)
- NGS becomes a more attractive prospect if multiple gene need to be assessed in parallel

## *NTRK* structural variant testing

- *NTRK1*, *NTRK2* and *NTRK3* encode TrkA, TrkB and TrkC which are receptor tyrosine kinases
- On growth factor binding, the receptor dimerises and autophosphorylates, resulting in activation of the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- The three *NTRK* genes form fusions with different genes:
  - They result in overexpression of the *NTRK* gene, and result in the production of a Trk protein which is constitutively active
  - This leads to intense signalling through the MAPK/ERK pathway, resulting in uncontrolled cell proliferation (i.e. they are oncogenes)
- Different fusions are more common in different tumour types:
  - They result in overexpression of the *NTRK* gene, and results in the production of a Trk protein which is constitutively active
  - This leads to intense signalling through the MAPK/ERK pathway, resulting in uncontrolled cell proliferation

- Regardless of histological tumour type, a patient with an advanced/unresectable solid tumour type harbouring an *NTRK* rearrangement is eligible for TKI therapy, assuming they have no suitable treatment option
- *NTRK* rearrangements occur in most cases of a small group of rare tumours:
  - Secretory carcinoma of breast
  - Secretory carcinoma of salivary gland
  - Infantile fibrosarcoma
  - Congenital mesoblastic nephroma
- In these high-frequency tumour types, *NTRK* may service a diagnostic purpose
- *NTRK* rearrangements occur with fair frequency (25-75% of cases) in a small group of uncommon tumours:
  - GIST which are negative for *KIT* and *PDGFRA* small variants
  - Spitzoid neoplasms
  - Papillary thyroid carcinomas, especially paediatric
  - Paediatric gliomas
- *NTRK* rearrangements are rare in other tumour types
- In common carcinomas (e.g. lung, bowel, breast, prostate), they are generally seen in less than 1% of cases
- Multiple methods may be used to assess for *NTRK* rearrangements:
  - Immunohistochemistry
  - FISH
  - RT-PCR
  - NGS (usually RNA-based)
- In general, to access treatment, the precise fusion must be known – this means that RT-PCR or NGS are generally needed
- Coverage of all common fusions across all three genes in an RT-PCR assay is challenging, and therefore RT-PCR is likely to miss uncommon variants
- One of the issues with *NTRK* rearrangement testing is that three genes must be examined, therefore assessment for *NTRK* with FISH requires three separate FISH assessments – this is expensive, time-consuming and requires a lot of tissue
- Pan-Trk immunohistochemistry is available:
  - It detects overexpression of all three of the Trk proteins and may be used to infer structural variants involving any of the three *NTRK* genes
  - However, its reliability is not certain
  - Positive staining certainly requires confirmation with an alternative technique
  - Negative staining makes an underlying rearrangement unlikely, but there is some evidence that IHC may miss a significant minority of *NTRK3* rearrangements
- RNA-based NGS is generally the favoured technique

## PD-L1 immunohistochemistry

- For the background theory, see PD-L1 immunohistochemistry, above
- PD-L1 is a quantitative immunohistochemical assay used to determine eligibility for immune checkpoint inhibitors
- PD-L1 is not a single test
- The tumour type, the drug to be prescribed and (in a few situations) the line of therapy determine:
  - The PD-L1 assay required
  - The assessment method required
  - The cut-off for positivity
- Each PD-L1 assay was developed alongside a particular immune checkpoint inhibitor:
  - Agilent 22C3 assay with pembrolizumab
  - Agilent 28-8 assay with nivolumab

- VENTANA SP263 assay with durvalumab
- VENATANA SP142 assay with atezolizumab

- In most cases, these assays generate different results when staining the same tissue and are not interchangeable (with a few exceptions)
- There are three assessment methods:
  - Tumour proportion score (TPS) is based on expression in tumour cells alone
  - Immune cell (IC) score is based on the percentage area of tumour occupied by positive tumour-associated inflammatory cells
  - Combined positive score (CPS) is based on expression in both tumour and tumour-associated inflammatory cells (CPS is capped at 100)
- Because TPS and IC scoring are based on tumour-associated inflammatory cells, cytology generally cannot be used:
  - In cytology it is not possible to determine which inflammatory cells are tumour-associated
- The definition of positivity is a numerical value and the definition of positivity varies depending on the situation
- TPS and CPS require a minimum of 100 viable tumour cells, and IC requires a minimum of 50 viable tumour cells

## *RET* structural variant testing

- *RET* encodes a receptor tyrosine kinase
- On growth factor binding, the receptor dimerises and autophosphorylates, resulting in activation of the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- The rearrangements involve various fusion partners, which differ depending on the tumour type:
  - This results in overexpression of the *RET* gene, and results in the production of an RET protein which is constitutively active
  - This leads to intense signalling through the MAPK/ERK pathway, resulting in uncontrolled cell proliferation (i.e. it is an oncogene)
- *RET* structural variants can be detected by FISH
- *RET* structural variants can be detected by RT-PCR, although this will miss uncommon fusions
- RNA-based NGS may be used
- Immunohistochemistry is currently not considered a reliable tool for detecting *RET* structural variants

## *RET* small variant testing

- As described above, *RET* encodes a receptor tyrosine kinase:
  - Activating small variants result in constitutive activity
  - This drives uncontrolled cell growth
- By far the commonest variant is M918T, with uncommon variants seen in other hotspots
- Targeted methods will invariably miss rare variants, but because the clinically relevant *EGFR* variants are fairly few in number this is generally not a big issue:
  - Real-time PCR
- Sequencing approaches should reduce the risk of missing variants, but will often have longer turnaround times:
  - Direct/Sanger sequencing
  - Pyrosequencing
  - Next-generation sequencing (NGS)

### *TERT* promoter variant testing

- *TERT* encodes telomerase:
  - Telomeres are repetitive sequences at the ends of chromosomes which progressively shorten with each round of replication
  - This places a limit on the number of rounds of cell division
  - Telomerase maintains telomere length
  - Normally, telomerase is only active in germ cells
  - In other cells, *TERT* expression is reduced through epigenetic mechanisms
- Variants in the *TERT* promoter region generate novel transcription factor binding sites, which allow telomerase to be expressed
  - This opposes the normal senescence seen in cells which have undergone multiple rounds of replication
- Two well-characterised hotspot variants account for the vast majority of *TERT* promoter variants
- In view of the small number of variants, targeted methods (e.g. real-time PCR) would be a good option, but in reality are not generally used
- Next-generation sequencing is generally used
- TERT immunohistochemistry has been reported as not being reliable

### *TP53* small variant and copy number variant testing

- *TP53* encodes p53
- *TP53* is a tumour suppressor gene:
  - In settings of cell stress, p53 accumulates in the cell
  - It tiggers DNA repair mechanisms, halts the cell cycle, and if the issue cannot be resolved triggers apoptosis
- Loss of p53 function by either inactivating small variants or by loss of the gene (on chromosome 17) are extremely common across cancers
- Inactivating small variants are diverse and distributed widely across the gene, and so only sequencing methods are feasible – usually next-generation sequencing nowadays
- Copy number variants can be assessed using either FISH (for 17p loss) or next-generation sequencing
- p53 immunohistochemistry may be used as a surrogate marker for *TP53* inactivating variants under certain situations:
  - The staining seen in the absence of an underlying *TP53* inactivating variant is described as 'wild-type' staining: there is variable intensity of staining of tumour cell nuclei from complete negativity to strong staining
  - Patterns of staining suggestive of an underlying variant include overexpression, loss of expression and aberrant cytoplasmic staining
  - Wild-type staining may occasionally seen even in the presence of an underlying inactivating *TP53* variant – it is therefore not entirely sensitive

## Molecular testing in colorectal cancer

- According to NICE guidance, all newly diagnosed colorectal cancers must undergo either MMR IHC or MSI testing to screen for Lynch syndrome
- *KRAS*, *NRAS* and *BRAF* small variant testing is required in locally-advanced or metastatic disease – testing is usually performed at oncologist request
- *NTRK* structural variant and *HER2* amplification testing may be performed at oncologist request for patients who are running out of options

### MMR immunohistochemistry and MSI analysis

- Around 15% of colorectal cancers are dMMR or show MSI

- MMR defects and MSI are seen more commonly in:
  - Right-sided cancers
  - Poorly-differentiated cancers
  - Mucinous cancers
  - Tumours with large numbers of tumour-associated inflammatory cells
  - Tumour with a syncytial appearance and/or pushing border
- The commonest abnormal MMR immunohistochemical staining pattern is loss of expression of MLH1 and PMS2
- The majority of cases of dMMR/MSI are somatic, rather than reflecting underlying Lynch syndrome
- In colorectal cancer, MMR IHC and MSI testing are essentially 100% concordant
- The main reason that NICE recommends screening with MMR IHC/MSI testing is to screen for Lynch syndrome
- It also has other uses:
  - dMMR/MSI-H colon cancers show more indolent behaviour
  - Patients with dMMR/MSI-H colon cancer do not appear to benefit from chemotherapy regimens containing 5-fluorouracil
  - Patients with dMMR/MSI-H advanced/metastatic colon cancer are eligible for immune checkpoint inhibitors
- See MMR immunohistochemistry, above
- See microsatellite analysis, above

### *KRAS* and *NRAS* small variant testing

- *KRAS* small variants are seen in around 40-50% colon cancers, and *NRAS* in around 5%
- The presence of a pathogenic variant in either gene predicts resistance to anti-EGFR monoclonal antibody therapy:
  - If a tumour is found to harbour a pathogenic variant, the patient will receive chemotherapy alone
  - In general, the exact nature of the variant does not matter, as long as it is pathogenic
- They are also associated with more aggressive behaviour, but this is not normally relevant for clinical decision-making
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

### *BRAF* small variant testing

- *BRAF* small variants are seen in around 10% of colon cancers as a whole
- The interpretation of a *BRAF* variant depends on the MMR/MSI status of the tumour
- In a pMMR/MSS colon cancer with a *BRAF* variant:
  - Behaviour is likely to be significantly more aggressive
- In a dMMR/MSI-H colon cancer with a *BRAF* variant:
  - The MMR defect or MSI is considered somatic rather than associated with Lynch syndrome
  - Behaviour is likely to be indolent (because of the MMR defect/MSI)
- Regardless of MMR/MSI status, colon cancers with *BRAF* variants:
  - Are generally considered not to benefit from the addition of anti-EGFR monoclonal antibodies to chemotherapy (like *KRAS* and *NRAS*)
  - Are eligible for a combination of BRAF inhibitor with anti-EGFR monoclonal antibody therapy
- See *BRAF* small variant testing, above

### *HER2* copy number variant testing

- *HER2* amplification is seen in less than 5% of colon cancers

- Testing is normally done using (IHC and) ISH:
  - The current recommendation is that the upper GI assessment method be used
- Testing is not routinely undertaken:
  - There is evidence that *HER2*-amplified colon cancers are resistant to anti-EGFR monoclonal antibody therapy
  - Patients may be able to access HER2-targeted therapy, although this is an unlicensed indication
- See *HER2* copy number variant testing, above

## Molecular testing in ovarian cancer

- High-grade ovarian carcinomas are eligible for *BRCA1/BRCA2* variant testing and homologous recombination deficiency (HRD) testing
  - Patients need to be explicitly consented prior to either test
  - Testing may be performed in a reflex fashion if consent is communicated to the pathologist when the sample is sent to pathology
  - Otherwise, testing is performed at oncologist request
- *KRAS* small variant testing may occasionally be undertaken in mucinous ovarian neoplasms
- Because of the need for *BRCA1/BRCA2* variant and/or HRD testing, it is recommended in the UK that a minimal diagnostic IHC panel be undertaken on biopsies where the clinical picture would fit for ovarian cancer and where the morphology is suggestive of high-grade serous carcinoma:
  - PAX8: expected to be positive
  - ER: usually positive
  - WT1: expected to be positive
  - p53: expected to show aberrant staining patterns
  - If the results of this small panel are inconclusive, referral to a cancer centre for further work up is advised
- Whenever tumour *BRCA1/BRCA2* variant testing or HRD testing is undertaken on tumour tissue, it is also advisable to undertake *BRCA1/BRCA2* variant testing on a germline sample (usually blood)

### *BRCA1/BRCA2* variant testing

- Around 10% of ovarian high-grade serous carcinomas harbour somatic and a further 15% harbour germline variants in *BRCA1/BRCA2*
- Detection of a pathogenic variant (tumour or germline) predicts response to PARP inhibitors
- Detection of a pathogenic variant in tumour tissue may also raise the possibility of a germline variant
- The UK recommendation is for tumour and germline testing to be performed in parallel:
  - A tumour or germline variant makes the patient eligible for PARP inhibitors – tumour DNA is degraded by formalin and variants may end up being missed, so testing of a germline sample provides a backup
  - Comparing the results of germline and tumour tissue indicates whether any detecting variant is somatic or germline
- See BRCA1/BRCA2 variant and HRD testing, above

### HRD testing

- HRD testing has two components:
  - It tests for pathogenic *BRCA1/BRCA2* variants in the tumour (i.e. two potential causes of HRD)

  - It tests for a Genomic Instability Score (GIS), which is based on the presence of loss of heterozygosity, large-scale state transitions and telomeric allelic imbalance (these are downstream 'scars' of HRD)
- A tumour can be considered to show HRD if there is a pathogenic *BRCA1/BRCA2* variant and/or it has a high GIS
- Approximately 50% of ovarian high-grade serous carcinomas show HRD
- Detection of HRD predicts response to PARP inhibitors
- If not done in parallel, detection of either a pathogenic *BRCA1/BRCA2* variant or a high GIS should prompt germline *BRCA1/BRCA2* variant testing
- See *BRCA1/BRCA2* variant and HRD testing, above

### *SMARCA4* small variant testing

- *SMARCA4* encodes a transcription factor which regulates the expression of a variety of genes
- Almost all cases of small cell carcinoma of hypercalcaemic type harbour small variants in *SMARCA4*
- Detection of a variant in a compatible tumour is highly suggestive of the diagnosis, and failure to identify a variant should prompt reconsideration of the diagnosis
- Variants are distributed widely across the gene and so sequencing is the only viable approach
- A proportion of these variants will be germline, and so screening for germline variants is advisable

### *KRAS* small variant testing in ovarian mucinous neoplasms

- Around half of ovarian mucinous neoplasms harbour *KRAS* small variants
- In patients who are exhausting standard treatment options, *KRAS* small variant testing may be requested by analogy with colorectal cancer – detection of a *KRAS* small variant would prompt prescription of chemotherapy without anti-EGFR monoclonal antibodies
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

## Molecular testing in endometrial cancer

- Current UK guidance is that all newly diagnosed endometrial carcinomas (of any histological subtype, including carcinosarcoma), undergo reflex IHC:
  - MMR IHC
  - p53
  - ER
- *POLE* variant testing should also be requested in certain situations in a reflex fashion by pathologists, or at oncologist request
- For colorectal cancer, either MMR IHC or MSI testing can be undertaken at diagnosis, but for endometrial cancer, only MMR IHC is recommended:
  - MSI testing may return false negative results in some tumours with *MSH6* abnormalities
  - MSI testing should not be used as a screening tool for Lynch syndrome

### MMR immunohistochemistry

- Around 20% of endometrial carcinomas are dMMR
- MMR defects are seen more commonly in:
  - Endometrioid, undifferentiated, dedifferentiated or mixed carcinomas
  - Tumours with large numbers of tumour-associated inflammatory cells
  - Tumour with a syncytial appearance and/or pushing border
  - Tumours centred on the lower uterine segment

- The commonest abnormal MMR immunohistochemical staining pattern is loss of expression of MLH1 and PMS2
- The majority of cases are somatic, rather than reflecting underlying Lynch syndrome
- Concordance between MMR IHC and MSI testing is very high in endometrial cancer, but not perfect – *MSH6* alterations in particular are less reliably detected by MSI
- The principle reason that NICE recommends screening with MMR IHC/MSI testing is to screen for Lynch syndrome
- It also has other uses:
  - It may be useful diagnostically – dMMR supports a diagnosis of endometrioid, undifferentiated, dedifferentiated or mixed carcinoma
  - dMMR endometrial carcinomas generally show more indolent behaviour (despite sometimes high-grade histological appearances)
- See MMR immunohistochemistry, above

## *POLE* variant testing

- Around 10% of endometrial carcinomas harbour pathogenic *POLE* variants
- *POLE* variants are more common with:
  - Endometrioid carcinomas
  - High-grade morphology
  - Large numbers of tumour-infiltrating lymphocytes
- POLE's function is to correct errors introduced during DNA replication (i.e. it is a tumour-suppressor gene) – the genomes of *POLE* variant-containing tumours are therefore heavily mutated
- Because of this, POLE inactivation may cause *TP53* variants (therefore resulting in abnormal p53 staining) and may cause variants in the MMR genes (therefore resulting in abnormal MMR immunohistochemistry)
- Only 11 POLE variants are proven to have clinical significance:
  - Targeted testing could in theory be used to capture most cases
  - In reality, testing is essentially universally performed by next-generation sequencing
- The importance of *POLE* variant status is that although *POLE* mutated endometrial carcinomas often have high-grade histology, they show indolent behaviour
- Pathologists should instigate *POLE* variant testing in biopsies if:
  - The tumour shows abnormal MMR staining – the idea being that the MMR defect may be secondary to impaired *POLE* function
  - The tumour shows abnormal p53 staining – the idea being that any underlying *TP53* variant may be secondary to impaired *POLE* function
- Pathologists show instigate *POLE* variant testing in resections if:
  - The final diagnosis is a FIGO stage I or II non-endometrioid carcinoma
  - The final diagnosis is a FIGO stage IB or II endometrioid carcinoma
  - The final diagnosis is an ER-negative endometrioid carcinoma
  - The final diagnosis is a FIGO grade 3, FIGO stage IA endometrioid carcinoma showing no or only focal LVSI
  - The final diagnosis is a FIGO stage IA endometrioid carcinoma with substantial LVSI
- This of course requires that all endometrial carcinomas undergo MMR, p53 and ER IHC
- The idea is that an endometrial carcinoma with a pathogenic *POLE* variant – irrespective of other features – is likely to show less aggressive behaviour than expected and so may not require such aggressive treatment:
  - Detection of a pathogenic *POLE* variant may result in the decision not to give adjuvant radiotherapy

### *HER2* amplification testing in endometrial serous carcinoma

- *HER2* amplification is seen in around 25% of endometrial serous carcinomas
- Testing is normally done using (IHC and) ISH:
  - It is not entirely clear how IHC staining should be scored
- Testing is not routinely undertaken:
  - Patients may be able to access HER2-targeted therapy, although this is an unlicensed indication
- See *HER2* copy number variant testing, above

## Molecular testing in breast cancer

- All newly diagnosed breast cancers require, at a minimum, ER IHC and *HER2* amplification testing (and usually also PR IHC)
- *PIK3CA* small variant testing is typically undertaken at oncologist request
- Oncotype DX testing is only performed after surgery in cases with an intermediate risk of recurrence, and is not initiated by pathology
- PD-L1 IHC is usually only undertaken at oncologist request

### *HER2* amplification testing

- *HER2* amplification is seen in around 20% of breast cancers
- It is more commonly seen in:
  - Invasive ductal carcinoma
  - High grade disease
- *HER2* amplification predicts sensitivity to anti-HER2 targeted therapy in the neo-adjuvant, adjuvant and metastatic setting
- It also predicts more aggressive disease, although the availability of anti-HER2 targeted therapy means that there is now no significant prognostic difference
- *HER2* amplification testing is possible on cytological samples from metastatic sites (e.g. nodal metastases, pleural fluids):
  - Should not be performed on breast FNAs since it is not possibility to determine whether any malignant cells represent invasive carcinoma or DCIS
- See *HER2* copy number variant testing, above

### Oncotype DX testing

- A commercial assay which assesses the expression level of a set of 21 genes including *HER2* and the genes encoding ER and PR
- Cases need to be sent to the US for testing
- It returns a recurrence score, which estimates the likelihood of the cancer recurring following surgery
- To be used in ER-positive, HER2-negative, early-stage breast cancer which is at intermediate risk of recurrence after surgery:
  - A high recurrence score may shift the balance in favour of giving adjuvant chemotherapy
  - A low recurrence score will generally discourage the use of adjuvant chemotherapy
- Oncotype DX returns a *HER2* gene expression score which (although there is correlation with *HER2* gene amplification) should not be used to guide decisions about anti-HER2 therapy

### *PIK3CA* small variant testing

- *PIK3CA* encodes part of a PI 3-kinase (PI3K)

- PI3Ks are activated by binding of growth factor to receptor tyrosine kinases, triggering signalling through the PI3K/AKT/mTOR, ultimately promoting cell survival, growth and proliferation
- Activating small variants cause constitutive PI3K activation in the absence of ligand binding, causing uncontrolled cell growth and proliferation (i.e. it is an oncogene)
- Just under half of ER-positive, *HER2*-negative breast cancers harbour *PIK3CA* small variants
- Activating *PIK3CA* small variants are clustered in a small number of hotspots, predominantly in exons 9 and 20, but a substantial minority of variants fall outside this small group of common variants
- Activating variants in *PIK3CA* in ER-positive *HER2*-negative advanced breast cancer are associated with responses to PI3K inhibitors
- Whether there is a prognostic association with *PIK3CA* variants is unclear, but the balance of evidence seems to suggest that it is associated with more favourable outcomes
- Targeted methods will invariably miss uncommon variants, albeit infrequently:
    - Real-time PCR
- Sequencing approaches should reduce the risk of missing variants, but will often have longer turnaround times:
    - Direct/Sanger sequencing
    - Pyrosequencing
    - Next-generation sequencing (NGS)

## *NTRK* structural variant testing in secretory carcinoma

- Like its counterpart in salivary gland, a large majority of cases of secretory carcinoma of breast harbour *NTRK* fusions, particularly *ETV6*::*NTRK3*
- These tumours are eligible for TKI therapy if the patient has no other suitable treatment options (although these tumours generally show fairly indolent behaviour)
- This can also be used to support the diagnosis, although the morphology is generally quite distinctive
- See *NTRK* structural variant testing, above
- The value of pan-Trk IHC in detecting *NTRK* fusions for predictive purposes is uncertain:
- However, if a diagnosis of secretory carcinoma is already suspected, pan-Trk IHC is extremely useful for confirming the diagnosis:
    - Positive staining effectively confirms the diagnosis
    - Negative staining should prompt *NTRK* fusion testing through alternative techniques (especially NGS)

## PD-L1 immunohistochemistry in triple negative breast cancer

- Atezolizumab with chemotherapy is licensed for use in advanced triple negative breast cancer in first line:
    - VENTANA SP142 assay
    - IC assessment method
    - Positivity is IC $\geq$ 1%
    - A minimum of 50 viable tumour cells is needed
    - Testing is generally not possible on cytological samples
- Pembrolizumab with chemotherapy is licensed for use in advanced triple negative breast cancer in first line, if the tumour is negative for PD-L1 using the above assay:
    - Agilent 22C3 assay
    - CPS assessment method
    - Positivity is CPS $\geq$ 10
    - A minimum of 100 viable tumour cells is needed

  - Testing is generally not possible on cytological samples
- See PD-L1 immunohistochemistry, above

### Germline *BRCA1/BRCA2* variant testing

- Detection of a pathogenic germline makes the patient eligible for adjuvant PARP inhibitors therapy in high-risk, *HER2*-negative breast cancer
- Testing is typically initiated by clinicians and undertaken on blood; pathology is not normally involved
- See BRCA1/BRCA2 variant and HRD testing, above

## Molecular testing in non-small cell lung cancer

- Given its poor prognosis, the norm is for all currently-actionable molecular tests to be requested as diagnosis of non-squamous NSCLC in a reflex fashion
- The situation for squamous cell carcinoma varies:
  - Some centres treat squamous cell carcinoma like any other NSCLC and request all tests
  - Some centres request only PD-L1 in squamous cell carcinoma
  - Some centres use a combination of approaches
- *MET* copy number and *HER2* small variant testing is not routine at present and is essentially only undertaken at oncologist request

### *EGFR* small variant testing

- In Caucasian populations, *EGFR* small variants are seen in around 10% of NSCLC, and around 15% of adenocarcinoma
- They are seen more commonly in:
  - Adenocarcinomas
  - Lepidic growth pattern
  - Female patients
  - Never or light smokers
  - Individuals of East Asian heritage (especially South East Asian)
- Most variants predict response to anti-EGFR tyrosine kinase inhibitors, especially the exon 19 deletions and L858R
- Some variants predict primary resistance to TKIs (e.g. exon 20 insertions)
- Patients inevitably progress on anti-EGFR TKIs, usually after around 1-2 years
  - If the patient has been treated with a first- or second-generation TKI, in 50% of cases the reason for progression will be acquisition of a secondary *EGFR* T790M variant
  - Patients progressing on first- or second-generation TKI require repeat tumour biopsy or plasma testing to assess for the presence of T790M
  - Acquisition of T790M can usually be successfully managed by changing the patient to a third-generation TKI (osimertinib)
- Patients who have intermediate-stage resected NSCLC harbouring an *EGFR* exon 19 deletion or L858R variant are eligible for adjuvant TKI therapy
- For a patient with surgically resectable disease to be eligible for neoadjuvant immune checkpoint inhibitors, their tumour must not harbour an *EGFR* exon 19 deletion or L858R variant (as well as an *ALK* structural variant)
- See *EGFR* small variant testing, above

### *KRAS* small variant testing

- Currently, the only *KRAS* variant which is actionable is the G12C variant
- It is seen in around 10-15% of adenocarcinomas
- It is seen more commonly in:

  - Adenocarcinomas
  - Males
  - Smokers
- The presence of a G12C variant predicts response to a KRAS G12C inhibitor
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

### *BRAF* small variant testing

- *BRAF* V600 variants are seen in around 1-2% of adenocarcinomas
- Associations are not entirely clear, but they are:
  - More common in adenocarcinomas
  - Not seen particularly more frequently in never/light smokers
- The presence of a V600 variant predicts response to combination MEK and BRAF inhibition
- See *BRAF* small variant testing, above

### *ALK* structural variant testing

- *ALK* encodes a receptor tyrosine kinase
- On growth factor binding, the receptor dimerises and autophosphorylates, resulting in activation of the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- The vast majority of rearrangements are inversions which generate *EML4*::*ALK* fusions:
  - This results in overexpression of the *ALK* gene, and results in the production of an ALK protein which is constitutively active
  - This leads to intense signalling through the MAPK/ERK pathway, resulting in uncontrolled cell proliferation (i.e. it is an oncogene)
- *ALK* structural variants are seen in around 2-3% of adenocarcinomas
- They are seen more commonly in:
  - Adenocarcinomas
  - Tumours with solid growth or with mucinous features
  - Never or light smokers
  - Younger patients
- The presence of an *ALK* structural variant predicts response to TKIs
- For a patient with surgically resectable disease to be eligible for neoadjuvant immune checkpoint inhibitors, their tumour must not harbour an *ALK* structural variant (as well as an *EGFR* exon 19 deletion or L858R variant)
- Multiple techniques can be used to assess for fusions:
  - Immunohistochemistry
  - FISH
  - RT-PCR
  - NGS (usually RNA-based)
- In general, immunohistochemistry and RNA-based NGS are considered the best options
- Immunohistochemistry using the VENTANA D5F3 assay:
  - This particular assay is extremely sensitive and specific (other ALK IHC assays are not)
  - Strong granular cytoplasmic staining in any proportion of tumour cells makes the tumour positive for *ALK* rearrangement

### *ROS1* structural variant testing

- *ROS1* encodes a receptor tyrosine kinase
- On growth factor binding, the receptor dimerises and autophosphorylates, resulting in activation of the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation

- *ROS1* forms fusions with a variety of genes, especially *CD74*:
  - This results in overexpression of the *ROS1* gene, and results in the production of a ROS1 protein which is constitutively active
  - This leads to intense signalling through the MAPK/ERK pathway, resulting in uncontrolled cell proliferation (i.e. it is an oncogene)
- *ROS1* structural variants are seen in around 1-2% of adenocarcinomas
- They are seen more commonly in:
  - Adenocarcinomas
  - Tumours with solid growth or with mucinous features
  - Never or light smokers
  - Younger patients
- The presence of a *ROS1* structural variant predicts response to TKIs
- Multiple techniques can be used to assess for fusions:
  - Immunohistochemistry
  - FISH
  - RT-PCR
  - NGS (usually RNA-based)
- In general, immunohistochemistry and RNA-based NGS are considered the best options
- Immunohistochemistry using either of the two commercial assays:
  - They are highly sensitive but have lower specificity
  - A negative result effectively rules out a *ROS1* rearrangement, and may be used as a convenient way of quickly screening out negative cases
  - Positive results do not, however, necessarily mean that a rearrangement is present, and confirmatory techniques are required
  - There is no formal definition of positivity, but staining is cytoplasmic
  - Staining is normally seen in reactive pneumocytes, which act as a useful internal control but may cause confusion in well differentiated adenocarcinomas

## *RET* structural variant testing

- *RET* structural variants are seen in around 1-2% of adenocarcinomas
- They are seen more commonly in:
  - Adenocarcinomas
  - Tumours with solid growth or with mucinous features
  - Never or light smokers
  - Younger patients
- The presence of a *RET* structural variant predicts response to TKIs
- In NSCLC, the commonest fusion partner is *KIF5B*, although various fusions are seen
- See *RET* structural variant testing, above

## *MET* exon 14 skipping variant testing

- *MET* encodes a receptor tyrosine kinase
- On growth factor binding, the receptor dimerises and autophosphorylates, resulting in activation of the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- Exon 14 of MET is involved in marking the protein for degradation:
  - *MET* exon 14 variants result in aberrant splicing
  - Gives rise to a MET protein which lacks exon 14
  - Lack of exon 14 prevents the MET protein from being degraded
  - MET accumulates and uncontrolled cell proliferation results (i.e. it is an oncogene)
- *MET* exon 14 skipping is seen in less than 5% of adenocarcinomas
- It is seen more commonly in:

  - Sarcomatoid carcinomas and adenosquamous carcinomas
  - Smokers
  - Older patients
- The presence of *MET* exon 14 skipping predicts response to TKIs
- The main approaches for testing are RT-PCR, and RNA- and DNA-based NGS
- RNA approaches are considered superior to DNA:
  - DNA approaches assess for the presence of the variants which cause skipping
  - RNA approaches assess for the presence of skipping itself
  - DNA may be used but it is generally less comprehensive than RNA approaches
- MET immunohistochemistry is generally not considered to be a reliable surrogate marker for *MET* exon 14 skipping variants
- Note that this is entirely different to *MET* copy number variant testing

## PD-L1 immunohistochemistry

- Pembrolizumab is licensed for advanced NSCLC in first line:
  - Agilent 22C3 assay
  - TPS assessment method
  - If TPS ≥ 50%, the patient can receive pembrolizumab monotherapy
  - Otherwise, the patient can receive it only in combination with chemotherapy
  - There must be no *EGFR* small variant or *ALK* structural variant
  - A minimum of 100 viable tumour cells is needed
  - There is reasonable evidence to show that the Agilent 22C3 and VENTANA SP263 assays are concordant in this setting
- Atezolizumab is licensed for advanced NSCLC in first line:
  - VENTANA SP142 assay
  - TPS ≥ 50% or IC ≥ 10% is positive
  - There must be no *EGFR* small variant or *ALK* structural variant
- Durvalumab is licensed for advanced NSCLC in second line:
  - VENTANA SP263 assay
  - TPS assessment method
  - TPS ≥ 1% is positive
  - A minimum of 100 viable tumour cells is needed
  - There is reasonable evidence to show that the Agilent 22C3 and VENTANA SP263 assays are concordant in this setting
- Atezolizumab is licensed as adjuvant therapy for resected intermediate-stage NSCLC:
  - VENTANA SP263 assay
  - TPS assessment method
  - TPS ≥ 50% is positive
  - A minimum of 100 viable tumour cells is needed
  - There is reasonable evidence to show that the Agilent 22C3 and VENTANA SP263 assays are concordant in this setting
- Nivolumab with chemotherapy is licensed as neoadjuvant therapy for patients with resectable NSCLC:
  - No PD-L1 testing is required
  - The tumour must not harbour an *ALK* structural variant, or an *EGFR* exon 19 deletion or L858R variant
- See PD-L1 immunohistochemistry, above

## *MET* copy number variant testing

- *MET* amplification is seen as a *de novo* phenomenon in up to around 5% of NSCLC
- It is a known mechanism of resistance to targeted therapy
- It is currently not actionable in either setting

- MET immunohistochemistry is generally not considered to be a reliable surrogate marker for *MET* amplification
- FISH is generally preferred
- NGS may be used but most routinely-used assays are not currently considered entirely reliable for detecting copy number variants
- Note that this is entirely different to *MET* exon 14 skipping variant testing

### *HER2* small variant testing

- *HER2* (or *ERBB2*) encodes the receptor tyrosine kinase HER2/neu
- Unlike other receptor tyrosine kinases, it has no known ligand and instead appears to be activated by other receptor tyrosine kinases
- Activation results in dimerisation and autophosphorylation, which triggers signalling through the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- Activating small variants cause constitutive HER2 activation in the absence of ligand binding, causing uncontrolled cell growth and proliferation
- *HER2* small variants are seen in around 1-3% of adenocarcinomas
- The vast majority are small insertions in exon 20
- They are more commonly seen in:
  - Adenocarcinomas
  - Females
  - Never or light smokers
- The presence of activating *HER2* small variants predicts response to anti-HER2 targeted therapy, although this is not licensed in the UK
- Targeted methods will invariably miss rare variants, but because the clinically relevant *EGFR* variants are fairly few in number this is generally not a big issue:
  - Real-time PCR
- Sequencing approaches should reduce the risk of missing variants, but will often have longer turnaround times:
  - Direct/Sanger sequencing
  - Pyrosequencing
  - Next-generation sequencing (NGS)
- It must be emphasised that *HER2* small variants are different to *HER2* copy number variants
- However, there are conflicting reports that *HER2* small variants are associated with the presence of *HER2* gene amplification
- HER2 immunohistochemistry is generally considered not to be a suitable surrogate marker for *HER2* small variant testing

## Molecular testing in small cell lung cancer

- At present, molecular testing is generally not required in small cell lung cancer
- It may be used as a diagnostic tool in rare cases

### *RB1* alteration testing

- *RB1* encodes retinoblastoma protein, which halts the cell cycle
- It is therefore a tumour suppressor gene – loss/inactivation of retinoblastoma protein allows cells to proliferative uncontrollably
- *RB1* small variants and copy number variants (deletions) are extremely common in small cell carcinoma
- Potentially useful diagnostically in challenging cases of small cell carcinoma
- Detection of an *RB1* alteration would support a diagnosis of small cell carcinoma
- However:

  - The diagnosis is usually straightforward histologically and, even when it is not, the clinical picture tends to be suggestive
  - Patients deteriorate very rapidly, so there often is not time to wait for molecular testing
- Pathogenic variants are therefore fairly diverse and distributed quite widely through the genes
  - Targeted testing approaches will miss too many variants
  - Nowadays, testing is performed using next-generation sequencing

# Molecular testing in mesothelioma

- At present, molecular testing is generally not required in mesothelioma
- It may be used as a diagnostic tool in rare cases

## *CDKN2A* copy number variant testing

- *CDKN2A* is located on chromosome 9, and encodes p16
- p16 halts cell cycle progression in response to cellular stress, and therefore *CDKN2A* is a tumour suppressor gene
- Loss of *CDKN2A* therefore results in uncontrolled cell proliferation
- Distinction between reactive mesothelial proliferations and mesothelioma can be challenging, particularly in cytology samples and superficial biopsies
- More than half of mesotheliomas are reported as showing homozygous deletion of *CDKN2A*
- This applies to both epithelioid and sarcomatoid subtypes
- Detection of homozygous deletion of *CDKN2A* therefore supports a diagnosis of mesothelioma (assuming, of course, that the clinical picture fits)
- Lack of homozygous deletion of *CDKN2A* does not, however, exclude the diagnosis
- Testing is generally undertaken with FISH or NGS

# Molecular testing in salivary gland neoplasms

- Molecular testing is useful for diagnostic purposes in select salivary gland tumours
- In a few cases, testing may be undertaken for predictive purposes

## *MAML2* structural variant testing in mucoepidermoid carcinoma

- Particularly in poorly differentiated cases, the diagnosis of mucoepidermoid carcinoma can be very challenging (especially difficult to distinguish from squamous cell carcinoma)
- Around 80% of mucoepidermoid carcinomas harbour *CRTC1*::*MAML2* fusions
- There is some suggestion that fusion-associated cancers have an improved prognosis, but this remains controversial
- Testing is invariably by FISH or NGS

## *MYB*::*NFIB* fusion testing in adenoid cystic carcinoma

- Adenoid cystic carcinoma is generally a straightforward diagnosis on morphology
- *MYB*::*NFIB* fusions have been reported in around three quarters of cases of adenoid cystic carcinoma
- Fusion testing may therefore be helpful in uncommon challenging cases
- Testing is invariably by FISH or NGS

## *NTRK* structural variant testing in secretory carcinoma

- Like its counterpart in breast, over three quarters of cases of secretory carcinoma of salivary gland harbour *ETV6*::*NTRK3* fusions
- These tumours are eligible for TKI therapy if the patient has no other suitable treatment options (although these tumours generally show fairly indolent behaviour)

- This can also be used to support the diagnosis, although the morphology is generally quite distinctive
- See *NTRK* structural variant testing, above
- The value of pan-Trk IHC in detecting *NTRK* fusions for predictive purposes is uncertain:
- However, if a diagnosis of secretory carcinoma is already suspected, pan-Trk IHC is extremely useful for confirming the diagnosis:
  - Positive staining effectively confirms the diagnosis
  - Negative staining should prompt *NTRK* fusion testing through alternative techniques (especially NGS)

# Molecular testing in melanoma

- Molecular testing may be helpful diagnostically in melanocytic neoplasms:
  - Various copy number variants may help (in some cases) to distinguish between benign and malignant melanocytic neoplasms
  - Structural variant detection may support a diagnosis of a spitzoid neoplasm
- Diagnostic molecular tests are usually only undertaken once all other avenues for diagnosis have been exhausted
- *BRAF* small variant testing is generally undertaken in all advanced melanomas, and sometimes in all melanomas
- Other tests are generally undertaken only at oncologist request

## *BRAF* small variant testing

- *BRAF* small variants are seen in around half of non-acral cutaneous melanomas:
  - Far less common in acral cutaneous sites
  - Rarely, if ever seen, in mucosal or uveal melanomas
- More commonly seen in:
  - Younger patients
  - Trunk melanomas
  - The setting of less sun damage
  - Association with superficial spreading melanoma
  - Higher stage disease
- Vast majority are V600E, with smaller proportions of V600K, V600R, etc.
- Very small proportion are outside codon 600
- V600-mutated melanomas may be eligible for targeted treatment:
  - Nowadays, advanced melanomas are usually treated with combination BRAF and MEK inhibitors
  - Resected stage III melanomas are eligible for adjuvant combination BRAF and MEK inhibitors
- *BRAF* small variants are also associated with more aggressive behaviour, but since these patients have a highly effective treatment, it does not have an impact on prognosis
- Most patients progress within a year:
  - Most commonly from acquired alterations in *BRAF*
  - Currently not targetable
- See *BRAF* small variant testing, above

## *KIT* small variant testing

- *KIT* small variants are seen in up to around 5% of melanomas:
  - More common in chronically sun-damaged skin
  - Also fairly common in acral and mucosal melanomas
- Mostly in exon 11 and less so in exon 13
- Activating *KIT* variants are associated with fair responses to TKIs, but this is not licensed
- Resistance develops fairly rapidly

- See *KIT* small variant testing, above

## *NRAS* small variant testing

- *NRAS* small variants are seen in around 15-20% of melanomas:
  - More common in nodular melanoma
  - More common in acral and mucosa melanoma
  - Less common in chronically sun-exposed skin
- Presence of *NRAS* small variants is associated with more aggressive disease
- Little success so far in targeting with MEK inhibitors
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

## Diagnostic copy number variant testing

- The following have been reported more commonly in melanomas than in other differential diagnoses:
  - Gain of *RREB1* (6p25)
  - Loss of *MYB* (6q24)
  - Gain of *CCND1* (11q13)
  - Gain of *MYC* (8q24)
  - Homozygous loss of *CDKN2A* (9p21)
- The results of testing must, however, be interpreted in light of the clinical and histological picture
- Testing is mostly undertaken by FISH

## *TERT* promoter variant testing

- 50-80% of melanomas harbour *TERT* promoter variants
- They are more common in:
  - Non-acral sun-exposed skin
  - Older age
  - Thicker tumours
  - The presence of *BRAF* small variants
- Testing may be helpful – alongside clinical and histological assessment – in distinguishing between melanoma and non-malignant melanocytic proliferations:
  - Variants are reported to be seen in only around 10% of benign naevi
  - The detection of a variant favours a diagnosis of melanoma over other mimics
  - Failure to detect a variant, however, by no means excludes the diagnosis of melanoma
- In established melanoma, variants are associated with more aggressive behaviour
- See *TERT* promoter variant testing, above

## Structural variant testing in spitzoid neoplasms

- Structural variants involving *NTRK*, *ALK* and *ROS1* (as well as others) are extremely uncommon in conventional melanomas, but are seen with reasonable frequency in spitzoid neoplasms
- *NTRK* structural variants are seen in 20-30% of spitzoid melanomas
  - Patients with *NTRK* fusion-positive melanomas who have exhausted other treatment options would be eligible for TKIs, as with any solid cancers
- It is unclear whether fusions involving genes other than *NTRK* are also targetable
- The presence of structural variants may be helpful diagnostically in occasional cases:
  - If the differential diagnosis is between melanoma and a non-malignant spitzoid neoplasm, detection of a structural variant would favour the latter
  - This needs to be interpreted with considerable caution
- See *NTRK* structural variant testing, above

### PD-L1 immunohistochemistry

- Nivolumab with/without ipilimumab (another immune checkpoint inhibitor) is licensed for advanced melanoma in first line:
  - Agilent 28-8 assay
  - TPS assessment method
  - Positivity is TPS ≥ 5%
  - A minimum of 100 viable tumour cells is needed
- If PD-L1 is negative, the patient will generally receive combination
- If PD-L1 is positive, the patient will generally receive nivolumab alone
- See PD-L1 immunohistochemistry, above

## Molecular testing in uveal melanoma

- Unlike in the skin, uveal melanomas only extremely rarely harbour *BRAF* or *NRAS* small variants

### Cytogenic abnormalities

- A variety of cytogenetic abnormalities has been reported to carry prognostic information:
  - Monosomy 3 (M3) is associated with poorer prognosis
  - 8q gain is associated with poorer prognosis
  - Combined monosomy 3 and 8q gain is associated with a particularly poor prognosis
  - 6q loss is associated with poorer prognosis
  - 8p loss is associated with poorer prognosis
  - 6p gain is associated with better prognosis
  - Combined monosomy 3 and 1p loss is associated with poor prognosis
- It is recommended that uveal melanomas routinely undergo testing for at least monosomy 3
- Multiple techniques are available for testing, including:
  - Karyotyping
  - FISH
  - Multiple ligand-dependent probe amplification (MLPA)
- *BAP1* is located on chromosome 3, and so loss of BAP1 expression on immunohistochemistry may be used as a surrogate marker for monosomy 3

## Molecular testing in gastrointestinal stromal tumours (GISTs)

- Given their difficult-to-predict biological behaviour, molecular testing is often untaken in a reflex fashion in GIST
- Some centres reserve testing for high-risk tumours
- GISTs which lack small variants in *KIT* and *PDGFRA* are sometimes described as 'wild-type GISTs'
  - These tumours have a fairly high rate of *NTRK* structural variants
  - Those GISTs which lack alterations in *KIT*, *PDGFRA* and *NTRK* have a relatively high chance of being associated with a familial syndrome
- Chemotherapy is poorly effective in advanced GIST, and so identifying an actionable target is extremely important

### *KIT* small variant testing

- *KIT* small variants are seen in up to around 80% of GISTs overall:
  - 90% of extra-gastric GISTs
  - 60% of gastric GISTs
- The majority are in exon 11, with a smaller proportion in exon 9

- The presence of *KIT* small variants predicts sensitivity to TKIs
- Variants in exon 9 are associated with less TKI sensitivity, and so are usually treated with double-dose TKI
- See *KIT* small variant testing, above

### *PDGFRA* small variant testing

- *PDGFRA* encodes a receptor tyrosine kinase
- Growth factor binding activates PDGRA
- Activation results in dimerisation and autophosphorylation, which triggers signalling through the MAPK/ERK pathway, ultimately promoting cell survival, growth and proliferation
- Activating small variants cause constitutive PDGFRA activation in the absence of ligand binding, causing uncontrolled cell growth and proliferation (i.e. it is an oncogene)
- Around 80% of variants are in exon 18, around 10% in exon 12, and less than 5% in exon 14
- The vast majority of exon 18 variants are D842V
- *PDGFRA* small variants are seen in up to around 10% of GISTs overall:
    - Mostly gastric
    - Mostly epithelioid or mixed morphology
- *PDGFRA* small variants are associated with more indolent behaviour
- D842V is generally associated with resistance to TKIs, but other *PDGFRA* small variants are generally considered to confer sensitivity to TKIs
- In view of the fact that the variants are located in a small number of hotspots, targeted methods are generally considered feasible, although rare variants will invariably be missed:
    - Real-time PCR
- Sequencing approaches should reduce the risk of missing variants, but will often have longer turnaround times:
    - Direct/Sanger sequencing
    - Pyrosequencing
    - Next-generation sequencing (NGS)

# Molecular testing in thyroid neoplasms

- Molecular testing may be used for diagnostic purposes in thyroid neoplasms (include in thyroid cytology, where morphology may be less conclusive):
    - Detection of *BRAF* V600E variants tends to favour papillary thyroid carcinoma
    - Detection of *KRAS*, *HRAS* and *NRAS* variants tends to favour follicular neoplasms or follicular variant of papillary carcinoma
- In general, molecular testing for predictive purposes is usually only initiated at oncologist request
- Papillary thyroid carcinomas harbour relatively high rates of *NTRK* structural variants (around 5-10%)
- Anaplastic thyroid carcinoma has an abysmal prognosis and so oncologists are generally keen to know about any potential targets, even if they do not have approved drugs

## Molecular testing in papillary thyroid carcinoma

### *BRAF* small variant testing

- Around half of papillary thyroid carcinomas harbour *BRAF* V600E variants
- They are seen:
    - More commonly in classical and tall cell variants
    - Less commonly in follicular variant

- The presence of the variant is associated with more aggressive behaviour
- Testing may be useful diagnostically: detection of a variant favours the diagnosis of papillary thyroid carcinoma, although failure to detect a variant does not exclude the diagnosis
- See *BRAF* small variant testing, above

### *RAS* small variant testing

- *KRAS*, *NRAS* and *HRAS* variants are only rarely seen in classical papillary thyroid carcinoma
- They are seen in a third to a half of follicular variant of papillary carcinoma
- Detection of a *RAS* variant militates against a diagnosis of classical papillary thyroid carcinoma, and instead favours follicular variant of papillary carcinoma or a follicular neoplasm
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

### *TERT* promoter variant testing

- *TERT* promoter variants are seen in around 10% of papillary thyroid carcinomas
- Predict more aggressive behaviour
- Predict particularly more aggressive behaviour when seen alongside *BRAF* variants
- See *TERT* promoter variant testing, above

### *RET* structural variant testing

- *RET* structural rearrangements are seen in around 10-20% of papillary thyroid carcinoma
- They are much commoner in paediatric and radiation-associated tumours
- The presence of *RET* structural variants predicts response to TKIs
- In thyroid carcinomas, the commonest fusion partner is *CCDC6*, although various fusions are seen
- Note that this is different to *RET* small variant testing in medullary thyroid carcinoma
- See *RET* structural variant testing, above

## Molecular testing in follicular thyroid carcinoma

### *RAS* small variant testing

- *KRAS*, *NRAS* and *HRAS* small variants are seen in a third to a half of follicular thyroid carcinomas (very similar to follicular variant of papillary thyroid carcinoma)
- They are reported to be slightly less common in follicular adenomas
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

### *RET* structural variant testing

- *RET* structural variants are vanishingly rare in follicular thyroid carcinoma
- However, if detected, they are associated with response to TKIs
- Note that this is different to *RET* small variant testing in medullary thyroid carcinoma
- See *RET* structural variant testing, above

## Molecular testing in anaplastic thyroid carcinoma

### *TP53* small variant and copy number variant testing

- *TP53* small variants and deletions are seen in thyroid malignancies, with increasing frequency as the degree of differentiation declines, reaching a peak in anaplastic thyroid carcinoma (around half of cases)
- Across thyroid carcinomas a whole, the presence of *TP53* variants is associated with more aggressive behaviour
- See *TP53* small variant and copy number variant testing

### *BRAF* small variant testing

- *BRAF* V600E variants are seen in around half of anaplastic thyroid carcinomas (similar to papillary thyroid carcinoma)
- There is evidence that it is associated with poorer prognosis (although the prognosis is already so poor that this makes little practical difference)
- Combination MEK/BRAF inhibition is approved in the US for *BRAF* V600E-mutated anaplastic thyroid carcinoma, but not in the UK
- See *BRAF* small variant testing, above

### *ALK*, *RET* and *NTRK* structural variant testing

- *ALK* structural variants are uncommon in anaplastic thyroid carcinomas (less than 5% of cases):
  - There is no approved treatment
  - However, given the very poor prognosis for this cancer, off-label treatment with TKIs may well be considered
- *RET* structural variants are seen in around 5-10% of anaplastic thyroid carcinomas:
  - Their presence predicts response to TKIs
  - See *RET* structural variant testing, above
- *NTRK* structural variants have been reported in 20% of poorly differentiated thyroid carcinomas
  - Their presence makes the patient eligible for TKIs if they have no other suitable treatment options
  - This invariably is the case for anaplastic thyroid carcinomas
  - See *NTRK* structural variant testing, above

## Molecular testing in medullary thyroid carcinoma

### *RET* small variant testing

- Around a quarter of medullary thyroid carcinomas arise in the context on MEN2 syndrome:
  - These are characterised by germline *RET* small variants
- Around half of non-syndromic medullary thyroid carcinomas harbour *RET* small variants:
  - M918T is by far the commonest single variant
  - This variant is associated with more aggressive behaviour
  - Less common variants are scattered across a small number of other hotspots
- Detection of a pathogenic *RET* small variant in a medullary thyroid carcinoma has several implications:
  - It raises the possibility that the patient may have MEN2, and should prompt referral to clinical genetics (although clinicians often refer immediately on diagnosis of medullary thyroid carcinoma, without waiting for *RET* testing)
  - *RET* small variants (irrespective of whether they are somatic or germline) predict response to TKIs – much the same as *RET* structural variants in other types of thyroid carcinoma
- See *RET* small variant testing, above

## Molecular testing in thyroid Hurtle cell carcinoma

### *RET* structural variant testing

- *RET* structural variants have been described in around a third to a half of Hurtle cell carcinomas
- As with other thyroid carcinomas, the presence of *RET* structural variants predicts response to TKI
- In thyroid carcinomas, the commonest fusion partner is *CCDC6*, although various fusions are seen

- See *RET* structural variant testing, above

### Molecular testing in non-invasive follicular thyroid neoplasm with papillary-like nuclei (NIFTP)

#### *BRAF* and *RAS* small variant testing

- *BRAF* V600E variants are not seen in NIFTP
- *BRAF* non-V600E variants are seen with low frequency (around 5%) in NIFTP
- Small variants in *NRAS*, *HRAS* and *KRAS* are seen around half of NIFTP (*NRAS* > *HRAS* > *KRAS*)
- Molecular testing may therefore be helpful in making this diagnosis:
  - Detection of a *BRAF* V600E variant all but excludes the diagnosis of NIFTP, and instead suggests a diagnosis of papillary thyroid carcinoma
  - Detection of a *RAS* small variant supports the diagnosis of NIFTP, in the appropriate histological context
- See *BRAF* small variant testing, above
- See *KRAS*, *HRAS* and *NRAS* small variant testing, above

## Molecular testing in phaeochromocytoma

### *RET* small variant testing

- It is reported that up to 10% of individuals with phaeochromocytoma have underlying MEN2 syndrome, which is characterised by germline *RET* small variants
- A similar proportion of phaeochromocytoma have been reported to harbour somatic *RET* small variants
- The significance of detection of a *RET* small variant in phaeochromocytoma:
  - It should prompt consideration of germline testing
  - At present (unlike thyroid carcinoma), patients with phaeochromocytoma with *RET* small variants are not eligible for TKI therapy, although there are case reports of responses
- See *RET* small variant testing, above

## Molecular testing in adrenal cortical carcinoma

### *TP53* small variant and copy number variant testing

- Inactivating variants and homozygous deletions in *TP53* have been reported in around a quarter of adrenal cortical carcinomas
- Detection of a variant is associated with more aggressive behaviour
- Germline *TP53* variants (causing Li Fraumeni syndrome) are reported in less than 5% of adult adrenal cortical carcinomas, but are much more common in paediatric cases
- Detection of a *TP53* variant in tumour should therefore prompt consideration of germline testing
- See *TP53* small variant and copy number variant testing, above

## Molecular testing in head and neck squamous cell carcinoma

### *CDK2NA* small variant and copy number variant testing

- Around a quarter of head and neck squamous cell carcinomas are reported to show *CDKN2A* inactivating small variants or copy number loss
- At least copy number loss appears to be commoner in HPV-negative cancers
- *CDKN2A* copy number loss is associated with poorer outcomes
- The association with prognosis is less clear for cancers with *CDKN2A* small variants

### *EGFR* small variant testing

- *EGFR* small variants are reported in less than 5% of head and neck squamous cell carcinomas, although there is considerable uncertainty surrounding this figure
- There is some suggestion that head and neck squamous cell carcinomas with *EGFR* small variants may have better outcomes, but this remains uncertain
- There are reports that patients with certain variants may respond to anti-EGFR monoclonal antibody therapy
- See *EGFR* small variant testing, above

### *TP53* small variant and copy number variant testing

- *TP53* small variants and copy number loss have been described in around three quarters of head and neck squamous cell carcinomas, with rates being somewhat higher than HPV-negative cases
- There is evidence to show that the presence of *TP53* small variants is associated with poorer survival, and that the prognostic association varies with the exact variant
- See *TP53* small variant and copy number variant testing, above

### *RET* structural variant testing

- *RET* structural variants are, at best, very rare in head and neck squamous cell carcinomas
- Unlike in NSCLC and thyroid carcinomas, no targeted therapies are available for *RET*-rearranged head and neck squamous cell carcinomas
- See *RET* structural variant testing, above

### PD-L1 immunohistochemistry

- Pembrolizumab with/without chemotherapy is licensed for advanced head and neck squamous cell carcinoma in first line:
  - Agilent 22C3 assay
  - CPS assessment method
  - Positivity is CPS ≥ 1
  - A minimum of 100 viable tumour cells is needed
  - Testing is generally not possible on cytological samples
- Pembrolizumab is licensed for advanced head and neck squamous cell carcinoma in second line:
  - Agilent 22C3 assay
  - TPS assessment method
  - Positivity is TPS ≥ 50%
  - A minimum of 100 viable tumour cells is needed
- Nivolumab is licensed for advanced head and neck squamous cell carcinoma in second line, but PD-L1 testing is only advised and not mandatory:
  - Agilent 28-8 assay
  - TPS assessment method
  - Positivity is TPS ≥ 1%
  - A minimum of 100 viable tumour cells is needed
- See PD-L1 immunohistochemistry, above

## Molecular testing in renal cell carcinoma

### Molecular defined renal cell carcinomas

- The latest WHO classification lays out a considerable number of diagnostic entities which are defined by their molecular alterations:
  - *TFE3*-rearranged renal cell carcinoma

- *TFEB*-rearranged renal cell carcinoma
- *ELOC*-mutated renal cell carcinoma (formerly *TCEB1*)
- Fumarate hydratase-deficient renal cell carcinoma (*FH* gene)
- Succinate dehydrogenase-deficient renal cell carcinoma (*SDHA*, *SDHB*, *SDHC*, *SDHD* genes)
- *ALK*-rearranged renal cell carcinoma
- *SMARCB1*-deficient renal cell carcinoma

- In *ELOC*-mutated renal cell carcinoma:
  - One allele of *ELOC* is typically inactivated by a small variant
  - The second allele is typically lost by chromosome 8 loss (loss of heterozygosity)
  - Testing for *ELOC* small variants or FISH for chromosome 8 are therefore useful in making this diagnosis

## *VHL* small variant and chromosome 3 copy number variant testing

- *VHL* encodes a protein whose normal function is to promote apoptosis and oppose angiogenesis:
  - It reduces the level of hypoxia-inducible factor which helps tissues to survive hypoxia
- Inactivating variants in or loss of *VHL* lead to reduced VHL function and therefore high levels of hypoxia-inducible factor:
  - There is an increase in angiogenesis and reduction in apoptosis
  - This allows tumour cells to persist and provides them with a rich blood supply to grow and metastasise
- It has been reported that more than half of clear cell renal cell carcinomas harbour small variants in *VHL*
- Almost all cases with somatic *VHL* variants also show loss of the second *VHL* allele (loss of heterozygosity) – there will be loss on chromosome 3 FISH
- This may be useful diagnostically: detection of a small variant or loss involving *VHL* supports a diagnosis of clear cell carcinoma, but the converse is not necessarily true
- A small proportion of detected *VHL* variants will be germline (i.e. von Hippel-Lindau syndrome)
- Small variants in *VHL* are widely distributed throughout the gene and so sequencing is the only viable approach
- Chromosome 3 losses are usually detected with FISH

## *TSC1/TSC2* small variant testing

- *TSC1* and *TSC2* encode hamartin and tuberin, respectively
- Hamartin and tuberin are tumour suppressors
- Tuberous sclerosis is associated with germline *TSC1/TSC2* variants
- Renal cell carcinoma occurs with greater frequency in individuals with tuberous sclerosis:
  - These tumours harbour germline small variants in *TSC1* or *TSC2*
  - Many of these germline variants arise *de novo*, meaning that the lack of a suspicious family history does not exclude the diagnosis
  - Renal cell carcinomas arising in patients with tuberous sclerosis most commonly fit into chromophobe carcinoma and papillary carcinoma categories
- Somatic *TSC1/TSC2* small variants can also be seen in renal cell carcinoma:
  - They are reported in less than 5% of clear cell carcinomas and less than 10% of chromophobe carcinomas
  - Interestingly, they appear to be significantly more common in renal cell carcinomas with abundant eosinophilic cytoplasm but which cannot be otherwise classified

- Inactivating variants are widely distributed through both genes, and so sequencing is generally the only feasible approach for testing in tumour tissue

### *MET* small variant and chromosome 7 copy number variant testing

- *MET* is an oncogene, as described above
- The small variants seen in papillary renal cell carcinoma are different to the *MET* exon 14 skipping variants seen in NSCLC:
  - In renal cell carcinoma, a variety of different activating variants are scattered throughout the kinase domain
  - These cause constitutive activation of the MET protein, resulting in uncontrolled cell growth
- In type 1 papillary renal cell carcinomas:
  - Small variants have been reported in around 20% of cases
  - Gain (chromosome 7) has been reported in a majority of cases
- In type 2 papillary renal cell carcinomas:
  - The rate of small variants is not well defined, but is said to be lower than in type 1
  - Gain (chromosome 7) has been reported in around half of cases
- Detection of a *MET* alteration may therefore favour papillary renal cell carcinoma
- Chromosome 7 gain may also be seen as evidence for papillary renal cell carcinoma, but this phenomenon may also be seen in clear cell renal cell carcinoma
- Testing for *MET* small variants is typically by sequencing
- Testing for chromosome 7 gain is typically by FISH

### Chromosome 17 copy number variant testing

- Chromosome 17 gain is a frequent finding in type 1 papillary renal cell carcinoma
- Testing for chromosome 17 gain is typically by FISH

### *BRAF* small variant testing

- Around 90% of metanephric adenomas harbour *BRAF* V600E variants
- In contrast, they are very rare in papillary renal cell carcinomas
- Detection of a *BRAF* V600E variant therefore favours metanephric adenoma over papillary renal cell carcinoma
- See *BRAF* small variant testing, above

## Molecular testing in urothelial carcinoma

### *FGFR2* and *FGFR3* alteration testing

- *FGFR2* and *FGFR3* alterations are reported in around 20-25% of urothelial carcinomas
  - 15% have *FGFR3* small variants
  - 3% have *FGFR2* copy number variants
  - The remainder have a variety of structural variants
- Alterations are commoner in low-grade carcinomas and in the upper urinary tract
- Patients with *FGFR2* and *FGFR3* alteration-harbouring urothelial cancer are eligible for TKI therapy in the US, and treatment is being considered in the UK
- See *FGFR* alterations

### PD-L1 immunohistochemistry

- Atezolizumab is licensed for use in advanced urothelial carcinoma in first line, if the patient is not eligible for cisplatin:
  - VENTANA SP142 assay
  - IC assessment method
  - Positivity is IC $\geq$ 5%
  - A minimum of 50 viable tumour cells is needed

  - Testing is generally not possible on cytological samples
- Pembrolizumab is licensed for use in advanced urothelial carcinoma in first line, if the patient is not eligible for cisplatin:
  - Agilent 22C3 assay
  - CPS assessment method
  - Positivity is CPS ≥ 10
  - A minimum of 100 viable tumour cells is needed
  - Testing is generally not possible on cytological samples
- Nivolumab is licensed for use as adjuvant therapy in muscle-invasive urothelial carcinoma:
  - Agilent 28-8 assay
  - TPS assessment method
  - Positivity is TPS ≥ 1%
  - A minimum of 100 viable tumour cells is needed
- See PD-L1 immunohistochemistry, above

# Molecular testing in prostate cancer

## *BRCA1* and *BRCA2* variant testing

- It has been reported that between 5 and 15% of prostate cancers harbour pathogenic *BRCA1* and *BRCA2* variants:
  - *BRCA2* is significantly more common than *BRCA1*
- The presence of pathogenic variants is generally considered to be associated with more aggressive disease
- There are suggestions that *BRCA1* and *BRCA2* variants are associated with improved responses to platinum chemotherapy
- Any detected variant could potentially be germline, and so follow-up germline testing is necessary
- The presence of a variant (somatic or germline) makes a patient with metastatic castration-resistant prostate cancer eligible for PARP inhibitors
  - Tumour testing is performed in the first instance, in order to detect both somatic and germline variants
  - If tumur testing fails, testing of blood can be performed (although this will miss somatic variants)
- See BRCA1/BRCA2 variant and HRD testing, above

## *TMPRSS2*::*ERG* fusion testing

- The *TMPRSS*:*ERG* fusion is reported in around half of cases of prostate cancer
- This may be useful diagnostically in cases where prostate origin for metastatic carcinoma is suspected but routine histopathology has not provided a definitive answer
- Rearrangements can be detected by FISH or next-generation sequencing

# Molecular testing in pancreatic cancer

## *BRCA1* and *BRCA2* variant testing

- It has been reported that up to around 5% of pancreatic ductal carcinomas harbour variants in *BRCA1* or *BRCA2*:
  - *BRCA2* variants are about twice as common as *BRCA1*
- The ratio of somatic to germline variants is not clearly defined – any variant detected in tumour tissue should prompt germline testing
- PARP inhibitors are available for metastatic pancreatic cancer harbouring germline variants in *BRCA1* or *BRCA2*:

  - Because the variant must be germline, there is little reason to test tumour tissue
- See BRCA1/BRCA2 variant and HRD testing, above

### MMR immunohistochemistry and MSI analysis

- MMR defects have been reported in less than 5% of pancreatic ductal carcinomas
- These tumours are reported to have improved prognosis
- MMR defects predict response to immune checkpoint inhibitors
- See MMR immunohistochemistry, above
- See microsatellite analysis, above

## Molecular testing in biliary cancers

### *FGFR2* structural variant testing in cholangiocarcinoma

- *FGFR2* structural variants are seen in around 10-15% of intrahepatic cholangiocarcinoma
- They are rarely, if ever, seen in extrahepatic cholangiocarcinoma
- Around half of rearrangements are intrachromosomal
- The commonest fusion partner is *BICC1*
- The presence of an *FGFR2* rearrangement in a cholangiocarcinoma predicts sensitivity to TKI therapy
- See *FGFR* alterations

### *IDH1* small variant testing in cholangiocarcinoma

- *IDH1* encodes isocitrate dehydrogenase 1:
  - It is normally involved in converting citrate to α-ketoglutarate, producing NADPH and carbon dioxide in the process
- *IDH1* variants have complex effects:
  - It is now believed that the main oncogenic mechanism is that mutant IDH1 results in the production of 2-hydroxyglutarate (2HG)
  - 2HG impairs demethylation and so results in abnormal gene expression
  - It is believed that this is what promotes cancer development
  - It is also believed that low levels of α-ketoglutarate result in increased levels of hypoxia-inducible factor, which results in increased angiogenesis and cell survival
  - It also reduces the cell's ability to undertake aerobic respiration
- Inhibitors of the protein product of mutant *IDH1* reduce the production of 2HG and counteract its oncogenic effect
- It has been reported that *IDH1* small variants are seen in around 10-15% intrahepatic cholangiocarcinoma and up to around 10% extrahepatic cholangiocarcinoma
- Detection of particular *IDH1* variants makes the patient eligible for TKI therapy
- Very few activating variants are known, all at codon 132
- R132C accounts for around half of all variants
- Targeted approaches work well in view of the small number of variants:
  - Real-time PCR
- Sequencing approaches should reduce the risk of missing variants, but will often have longer turnaround times:
  - Direct/Sanger sequencing
  - Pyrosequencing
  - Next-generation sequencing (NGS)
- Immunohistochemistry is frequently used to screen CNS tumours for the commonest *IDH1* variant, R132H:
  - This is different to the commonest variant in cholangiocarcinoma
  - The immunostain will not detect the commonest variants in cholangiocarcinoma

### MMR immunohistochemistry and MSI analysis

- The rate of MMR defects in biliary cancers is highly uncertain
- Most studies have given a rate of less than 5% in cholangiocarcinoma, but some series have reported rates up to around 20%
- MMR defects predict response to immune checkpoint inhibitors
- See MMR immunohistochemistry, above
- See microsatellite analysis, above

## Molecular testing in gastro-oesophageal carcinomas

### *HER2* amplification testing

- *HER2* amplification is seen in around 10-20% of gastric and gastro-oesophageal adenocarcinomas
- It is more commonly seen in:
  - Intestinal-type cancers
  - Proximal gastric cancers
- *HER2* amplification predicts sensitivity to anti-HER2 targeted therapy
- It also predicts more aggressive disease
- See *HER2* copy number variant testing, above

### MMR immunohistochemistry and MSI analysis

- MMR defects have been reported in up to around 10% of gastric cancers and up to around 5% of gastro-oesophageal adenocarcinomas
- The rate in oesophageal squamous cell carcinoma is uncertain, but is likely very low
- It is more commonly seen in:
  - Older patients
  - The distal stomach
  - Intestinal-type adenocarcinomas
  - *HER2*-negative cancers
- These tumours are reported to have improved prognosis
- MMR defects predict response to immune checkpoint inhibitors
- See MMR immunohistochemistry, above
- See microsatellite analysis, above

### PD-L1 immunohistochemistry

- Gastric cancer:
  - Nivolumab in combination with chemotherapy is licensed for advanced *HER2*-negative gastric cancer in first line:
    - Agilent 28-8 assay
    - CPS assessment method
    - Positivity is CPS $\geq 5$
    - A minimum of 100 viable tumour cells is needed
    - Testing is generally not possible on cytological samples
  - Pembrolizumab in combination with chemotherapy is licensed for advanced *HER2*-positive or *HER2*-negative gastric cancer in first line:
    - Agilent 22C3 assay
    - CPS assessment method
    - Positivity is CPS $\geq 1$
    - A minimum of 100 viable tumour cells is needed
    - Testing is generally not possible on cytological samples
- Gastro-oesophageal junction adenocarcinoma:

- Nivolumab in combination with chemotherapy is licensed for advanced *HER2*-negative gastro-oesophageal junction adenocarcinoma in first line:
  - Agilent 28-8 assay
  - CPS assessment method
  - Positivity is CPS ≥ 5
  - A minimum of 100 viable tumour cells is needed
  - Testing is generally not possible on cytological samples
- Pembrolizumab in combination with chemotherapy is licensed for advanced *HER2*-positive or *HER2*-negative gastro-oesophageal junction adenocarcinoma in first line:
  - Agilent 22C3 assay
  - CPS assessment method
  - Positivity is CPS ≥ 1
  - A minimum of 100 viable tumour cells is needed
  - Testing is generally not possible on cytological samples

- Oesophageal cancer:
  - Pembrolizumab in combination with chemotherapy is licensed for advanced *HER2*-negative oesophageal carcinoma in first line:
    - Agilent 22C3 assay
    - CPS assessment method
    - Positivity is CPS ≥ 10
    - A minimum of 100 viable tumour cells is needed
    - Testing is generally not possible on cytological samples
  - Nivolumab in combination with chemotherapy is licensed for advanced *HER2*-negative oesophageal adenocarcinoma in first line:
    - Agilent 28-8 assay
    - CPS assessment method
    - Positivity is CPS ≥ 5
    - A minimum of 100 viable tumour cells is needed
    - Testing is generally not possible on cytological samples
  - Nivolumab in combination with ipilimumab (another immune checkpoint inhibitor) or chemotherapy is licensed for advanced oesophageal squamous cell carcinoma in first line:
    - Agilent 28-8 assay
    - TPS assessment method
    - Positivity is TPS ≥ 1%
    - A minimum of 100 viable tumour cells is needed
- See PD-L1 immunohistochemistry, above

# Molecular testing in thymic carcinoma

## *KIT* small variant testing

- KIT protein expression is seen in a majority of thymic carcinomas
- A minority harbour variants in *KIT*
- There are multiple case reports of responses to TKIs
- See *KIT* small variant testing, above

# Molecular testing in cervical cancer

## PD-L1 immunohistochemistry

- Pembrolizumab in combination with chemotherapy is licensed for persistent, recurrent or metastatic cervical carcinoma:

  - Agilent 22C3 assay
  - CPS assessment method
  - Positivity is CPS ≥ 1
  - A minimum of 100 viable tumour cells is needed
  - Testing is generally not possible on cytological samples
- See PD-L1 immunohistochemistry, above

# Molecular testing in sebaceous neoplasms

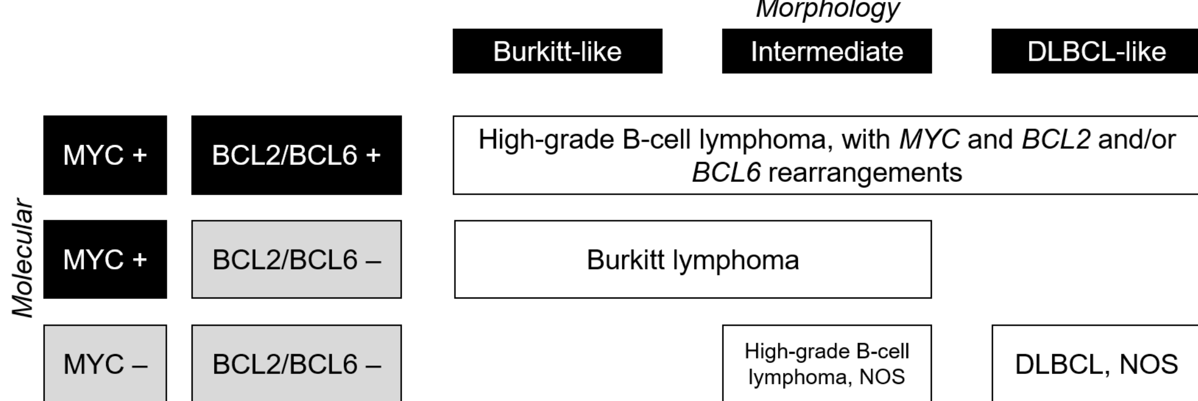
## MMR immunohistochemistry and MSI analysis

- Sebaceous adenomas, sebaceomas and sebaceous carcinomas can arise in the context of Muir-Torre syndrome (which itself is a variant of Lynch syndrome)
- There is significant uncertainty:
  - Rates of MMR defects vary enormously between studies
  - It is not clear whether *MLH1* promoter methylation reliability distinguishes between somatic and germline causes for loss of MLH1 staining
  - It is not clear what proportion of patients with MMR defects have Muir-Torre syndrome
- Different centres have different approaches:
  - Some screen all sebaceous neoplasms with MMR immunohistochemistry, others test only those in patients with suggestive histories
  - Some perform *MLH1* promoter methylation testing as per endometrial cancers and refer to clinical genetics if there is no methylation
  - Some refer all cases showing MMR immunohistochemical loss
- See MMR immunohistochemistry, above
- See microsatellite analysis, above

# Other tumour types

## High-grade mature B-cell lymphomas

- Most centres test all high-grade mature B-cell lymphomas for *MYC* rearrangements (some will only perform testing on cases showing high MYC IHC expression)
- *BCL2* and *BCL6* are often only tested if *MYC* is translocated




- 'High-grade B-cell lymphoma, with *MYC* and *BCL2* and/or *BCL6* rearrangements' are also known as double/triple-hit lymphomas, and have a very poor prognosis

## Diagnostic alterations in haematolymphoid neoplasia

| | Variants | Fusion gene |
|---|---|---|
| **Burkitt lymphoma** | t(8;14)<br>t(2;8)<br>t(8;22) | *IGH::MYM*<br>*IGK::MYC*<br>*IGL::MYC* |
| **DLBCL** | t(14;18) | *IGH::BCL2* |
| **Extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue** | t(1;14)<br>t(11;18)<br>t(11;14) | *IGH::BCL10*<br>*API2::MALT1*<br>*IGH::MALT1* |
| **Follicular lymphoma** | t(14;18) | *IGH::BCL2* |
| **Hairy cell leukaemia** | *BRAF* small variant | - |
| **Paediatric-type follicular lymphomas** | *MAP2K1* small variant | - |
| **Large B-cell lymphoma with *IRF4* rearrangement** | | *IRF4* |
| **Lymphoplasmacytic lymphoma** | t(9;14)<br>*MYD88* small variant | *IGH::PAX5*<br>- |
| **Mantle cell lymphoma** | t(11;14) | *IGH:BXL1* |
| **Langerhans cell histiocytosis** | *BRAF* small variant | - |

- In addition to these, large numbers of other molecular tests are undertaken for diagnostic, prognostic and predictive purposes in haematolymphoid neoplasms (see the WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues and the Genomic Test Directory)

## Diagnostic alterations in bone and soft tissue neoplasia

This list is taken from the RCPath Dataset for histopathological reporting of soft tissue sarcomas.

| | Rearrangement | Fusion gene |
|---|---|---|
| **Alveolar soft part sarcoma** | t(x;17) | *ASPL::TFE3* |
| **Angiomatoid fibrous histiocytoma** | t(12;22)<br>t(12;16) | *EWSR1::ATF1*<br>*FUS::ATF1* |

| | | |
|---|---|---|
| | t(2;22) | *EWSR1::CREB1* |
| **Aneurysmal bone cyst (primary)** | t(16;17) | *CDH11::USP6* |
| **Clear cell sarcoma** | t(12;22) | *EWSR1::ATF1* |
| **Clear cell sarcoma of the GI tract** | t(2;22) | *EWSR1::CREB1* |
| **Dermatofibrosarcoma protuberans** | t(17;22) | *COL1A1::PDGFB* |
| **Desmoplastic small round cell tumour** | t(11;22) | *EWSR1::WT1* |
| **Epithelioid haemangioendothelioma** | t(1;3) | *WWTR1::CAMTA1* |
| **Epithelioid sarcoma** | 22q abnormalities | *INI1* inactivation |
| **Endometrial stromal nodule and endometrial stromal sarcoma, low-grade** | t(7;17) | *JJAZ1::JAZF1* |
| **Endometrial stromal sarcoma, high-grade** | t(10;17) | *YWHAE::FAM22* |
| **Ewing sarcoma** | t(12;22)<br>t(21;22)<br>t(2;22)<br>t(7;22)<br>t(17;22)<br>inv(22) | *EWSR1::FLI1*<br>*EWSR1::ERG*<br>*EWSR1::FEV*<br>*EWSR1::ETV1*<br>*EWSR1::E1AF*<br>*EWSR1::ZSG* |
| **Extraskeletal myxoid chondrosarcoma** | t(9;22)<br>t(9;17)<br>t(9;15) | *EWSR1::NR4A3*<br>*TAF1168::NR4A3*<br>*TCF12::NR4A3* |
| **Infantile fibrosarcoma** | t(12;15) | *ETV6::NTRK3* |
| **Inflammatory myofibroblastic sarcoma** | 2p23 rearrangement | *ETV6::NTRK3* |
| **Liposarcoma, well-differentiated or dedifferentiated** | - | *MDM2* amplification |
| **Liposarcoma, myxoid** | t(12;16)<br>t(12;22) | *FUS::DDIT3*<br>*EWSR1::DDIT3* |
| **Low-grade fibromyxoid sarcoma** | t(7;16)<br>t(11;16) | *FUS::CREB3L2*<br>*FUS::CREB3L1* |
| **Malignant rhabdoid tumour** | Del(22q) | *INI1* inactivation |
| **Myoepithelial neoplasms** | t(19;22)<br>t(1;22)<br>t(6;22) | *EWSR1::ZNF444*<br>*EWSR1::PBX1*<br>*EWRS1::POU5F1* |
| **Osteosarcoma, low-grade central** | - | *MDM2* amplification |
| **Rhabdomyosarcoma, alveolar** | t(1;13)<br>t(2;13) | *PAX7::FKHR*<br>*PAX3::FKHR* |
| **Solitary fibrous tumour** | Inv(12) | *NAB2::STAT6* |
| **Synovial sarcoma** | t(x;18) | *SS18::SSX1*<br>*SS18::SSX2*<br>*SS18::SSX4* |

- In addition to these, large numbers of other molecular tests are undertaken for diagnostic purposes in bone and soft tissue neoplasms (see the WHO Classification of Soft Tissue and Bone Tumours and the Genomic Test Directory)

## Diagnostic alterations in CNS tumours

- Large numbers of molecular tests are undertaken for diagnostic purposes CNS neoplasms (see the WHO Classification of Central Nervous System Tumours and the Genomic Test Directory)