

wellcome
connecting
science



มหาวิทยาลัยขอนแก่น
KHON KAEN UNIVERSITY

Introduction to the course

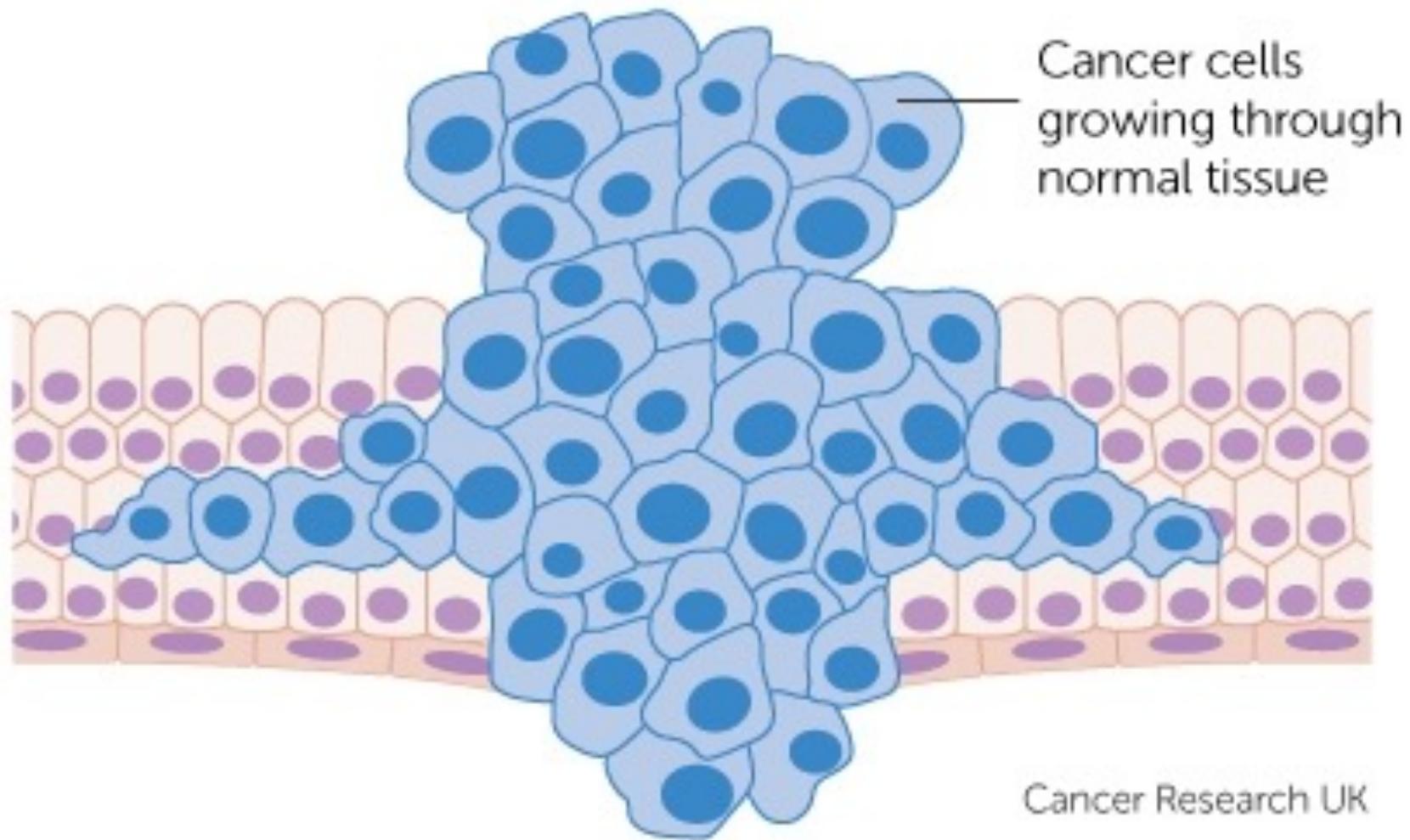


Cancer Genome Analysis - Asia

26–31 October 2025

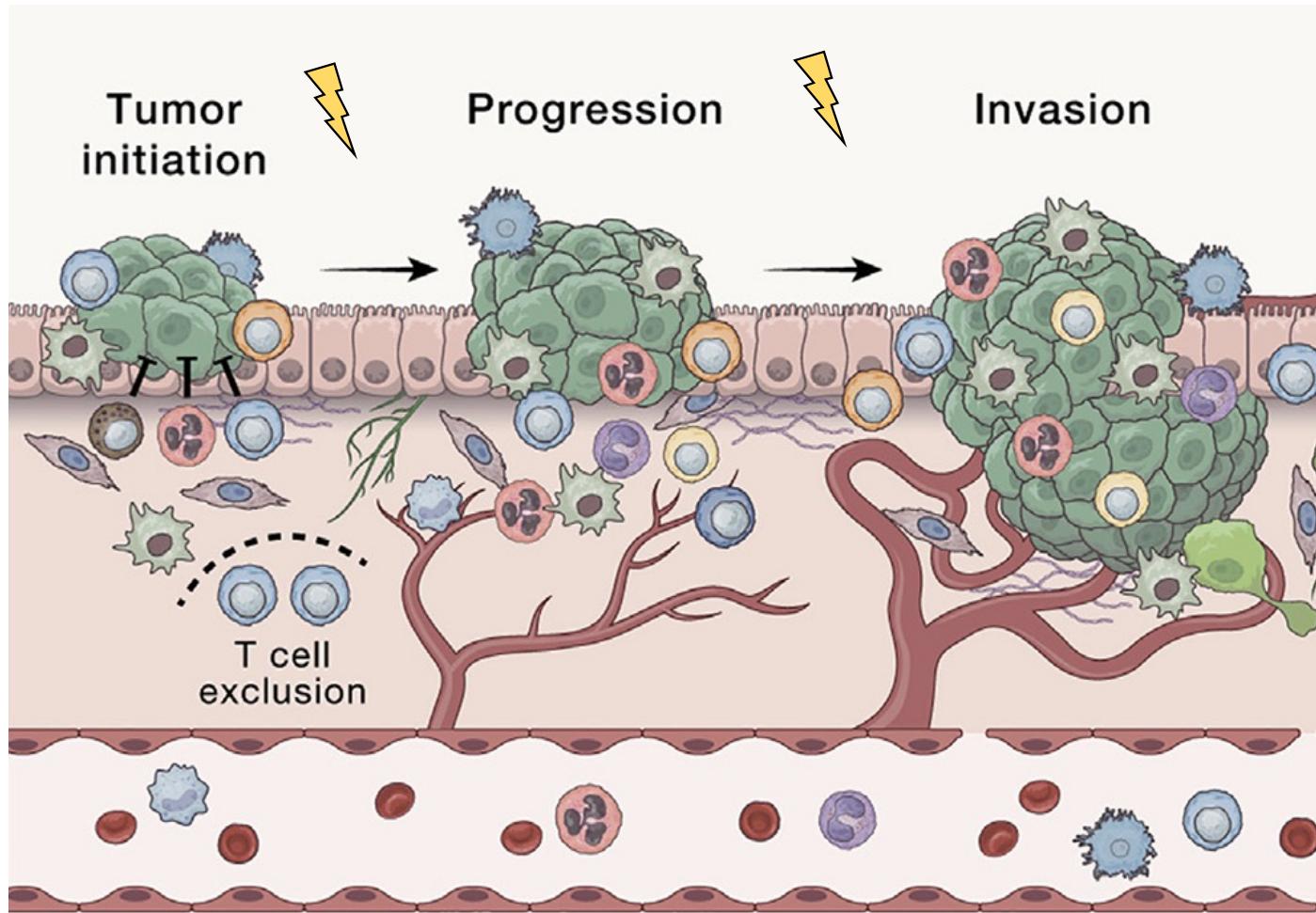
Khon Kaen University, Thailand

What is cancer?



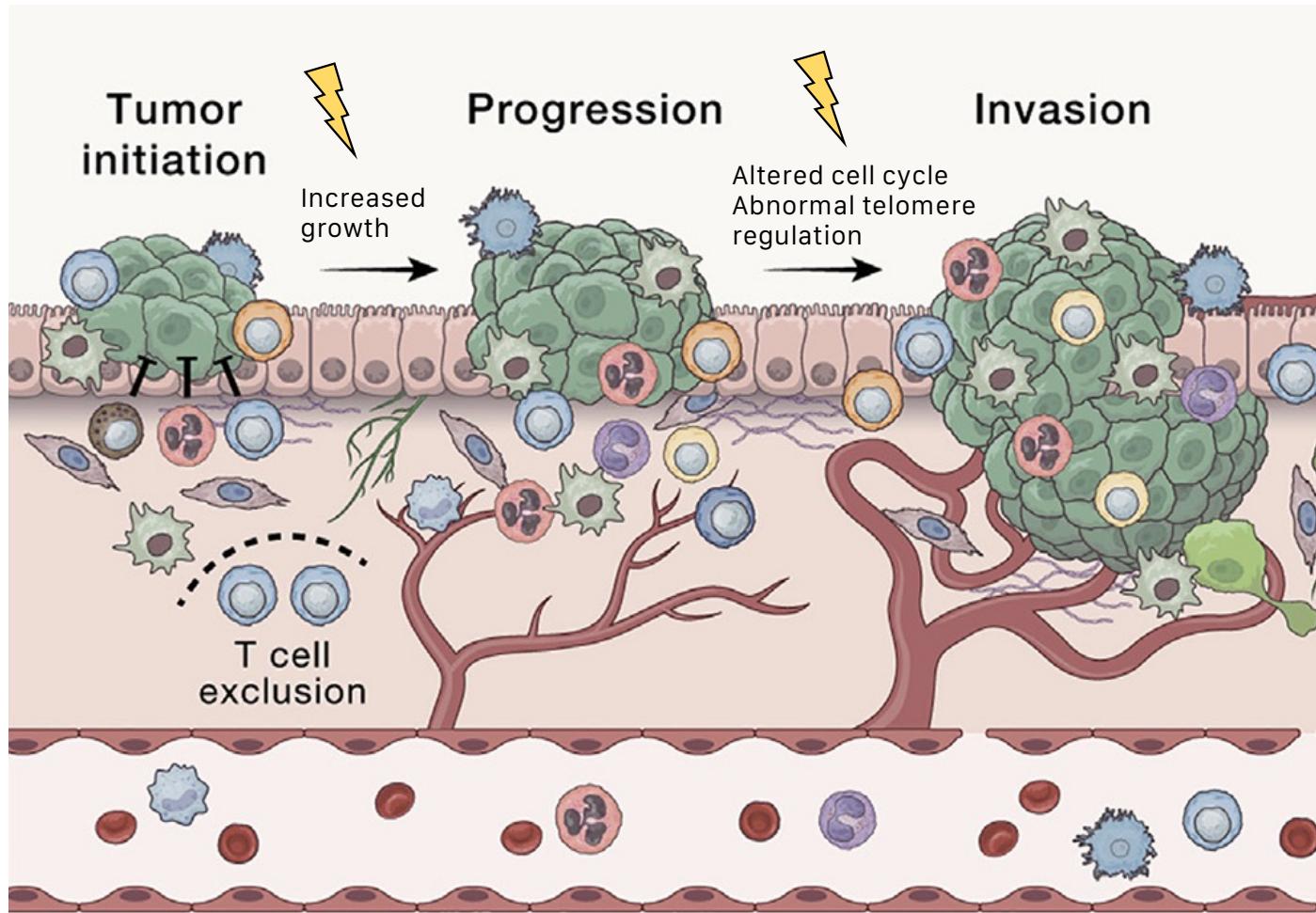
Cancer Research UK

Cancer develops as cells acquire mutations



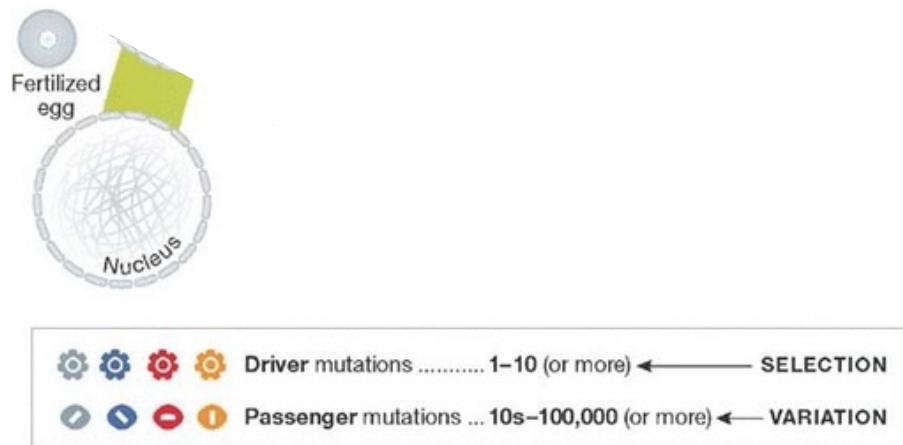
De Visser and Joyce (2023)

Cancer develops as cells acquire mutations



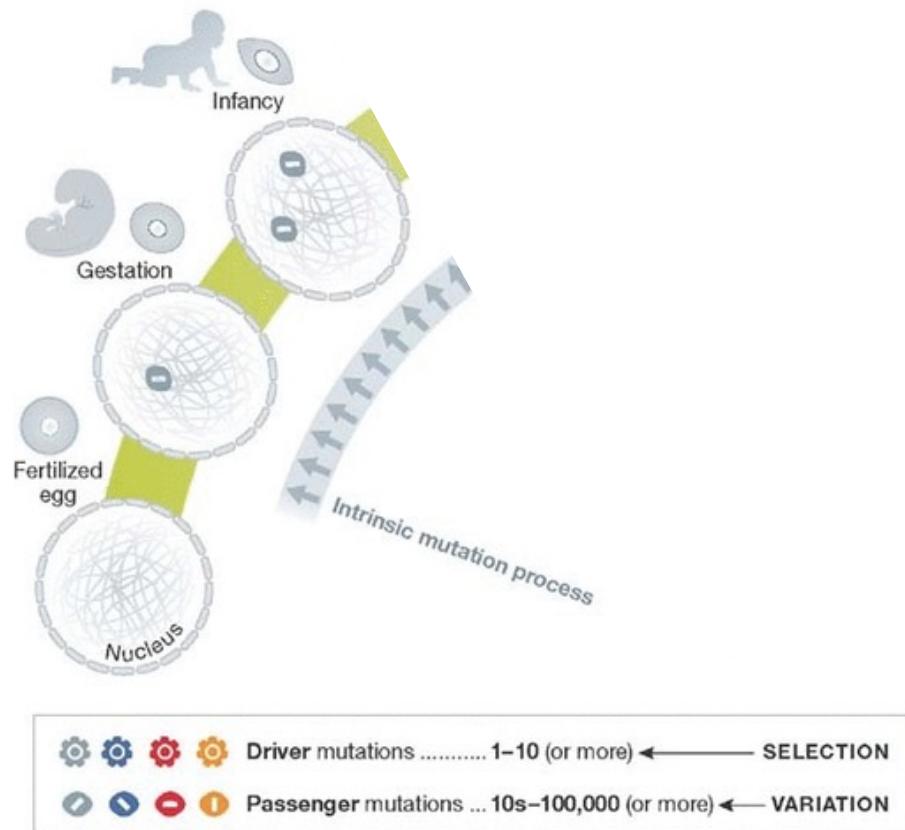
De Visser and Joyce (2023)

The acquisition of mutations is a life-long process

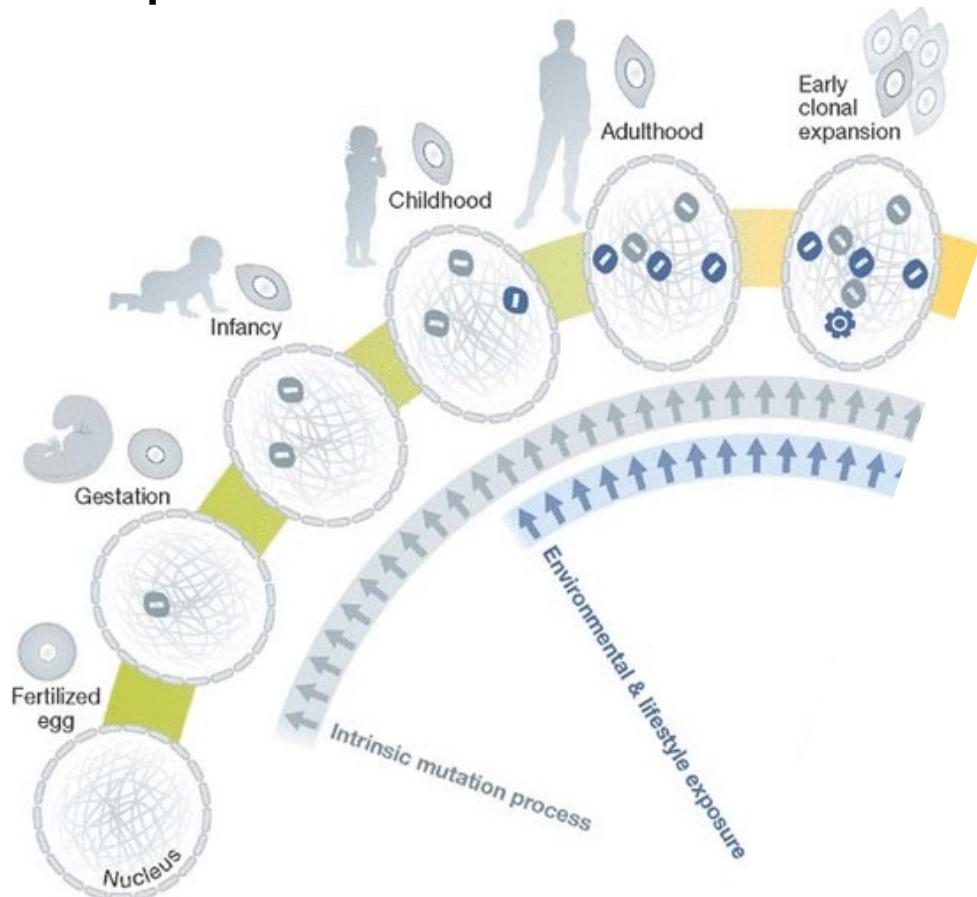


Stratton (2013)

The acquisition of mutations is a life-long process



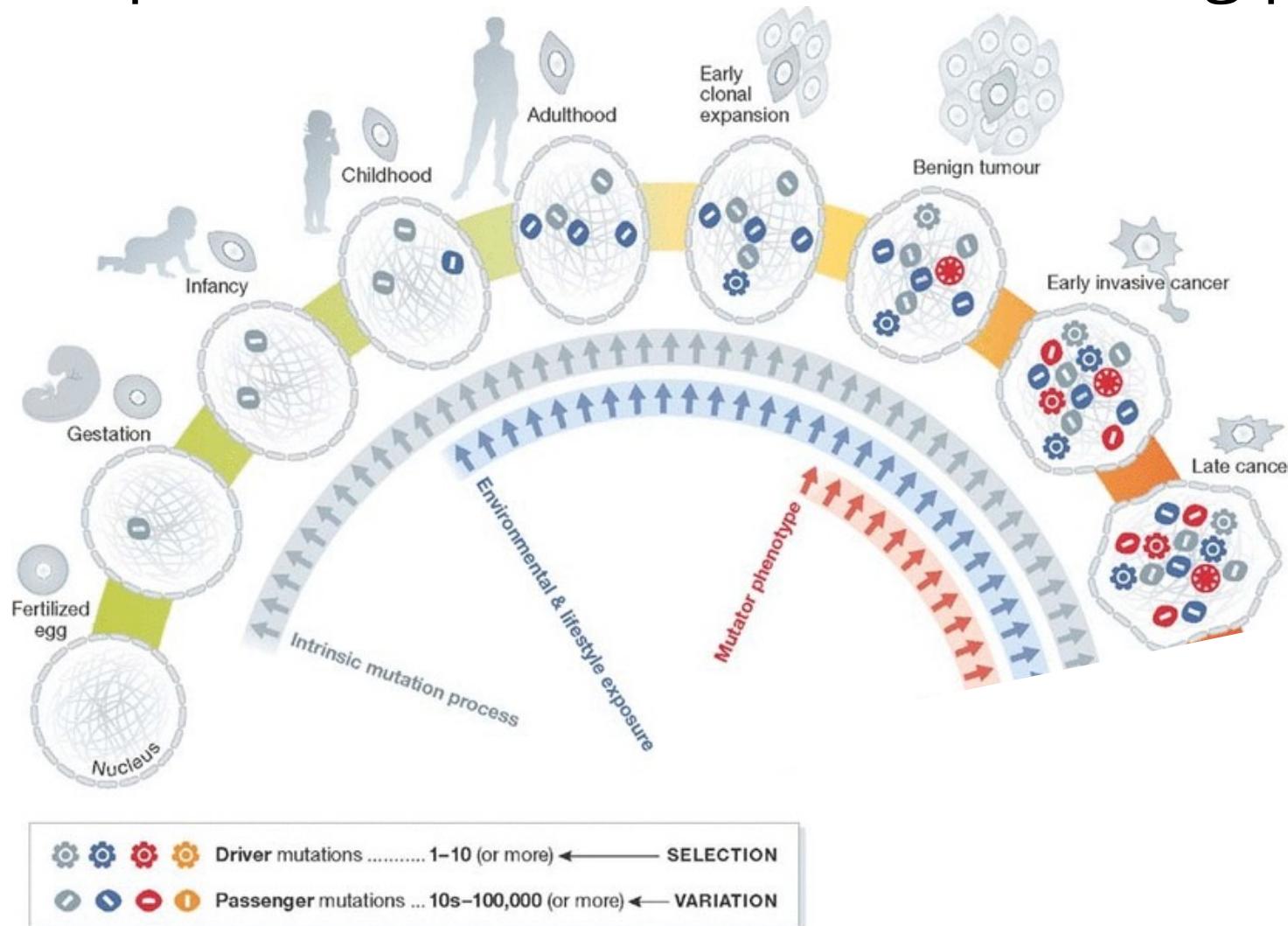
The acquisition of mutations is a life-long process



Driver mutations 1-10 (or more) ← SELECTION
Passenger mutations ... 10s-100,000 (or more) ← VARIATION

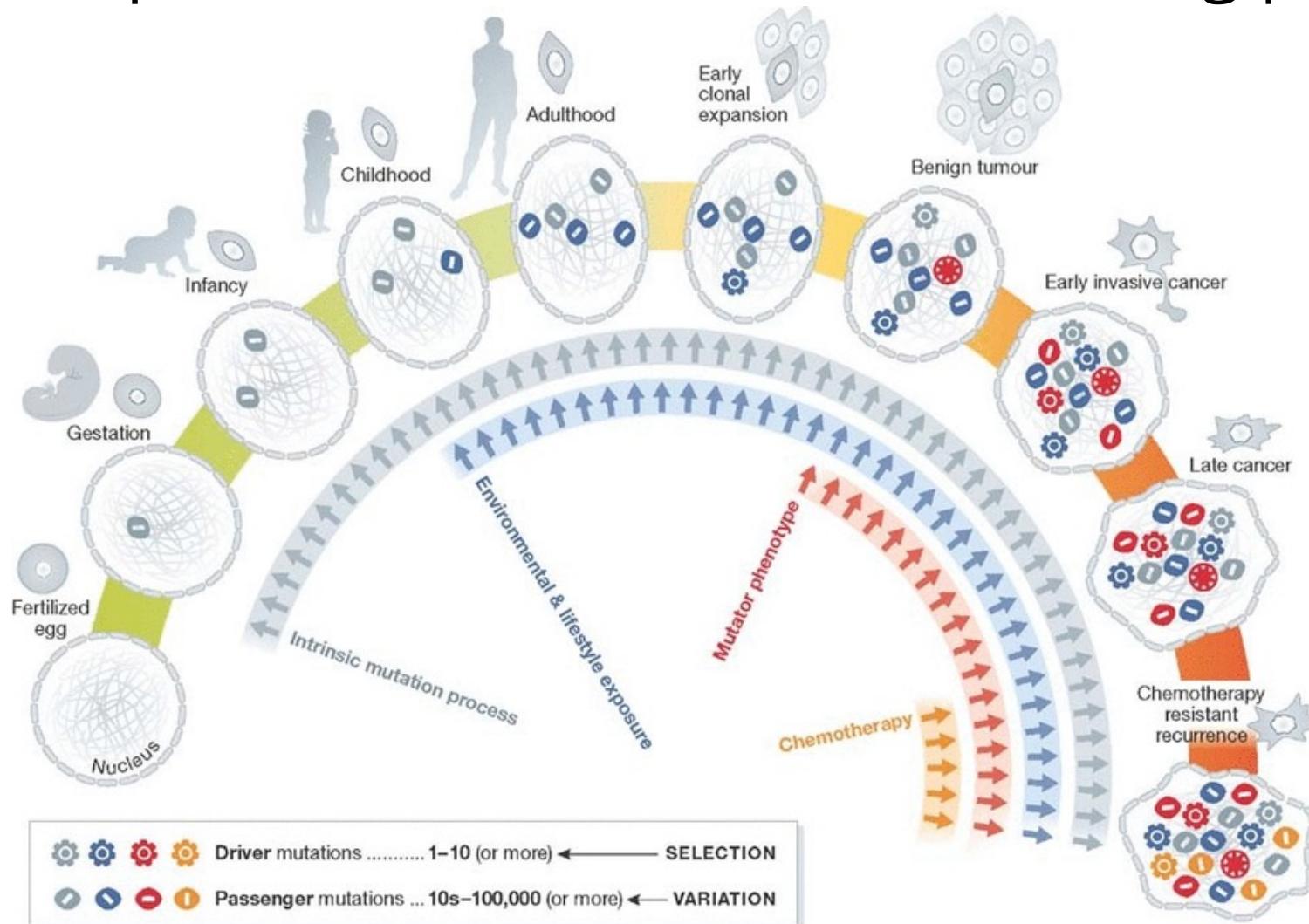
Stratton (2013)

The acquisition of mutations is a life-long process



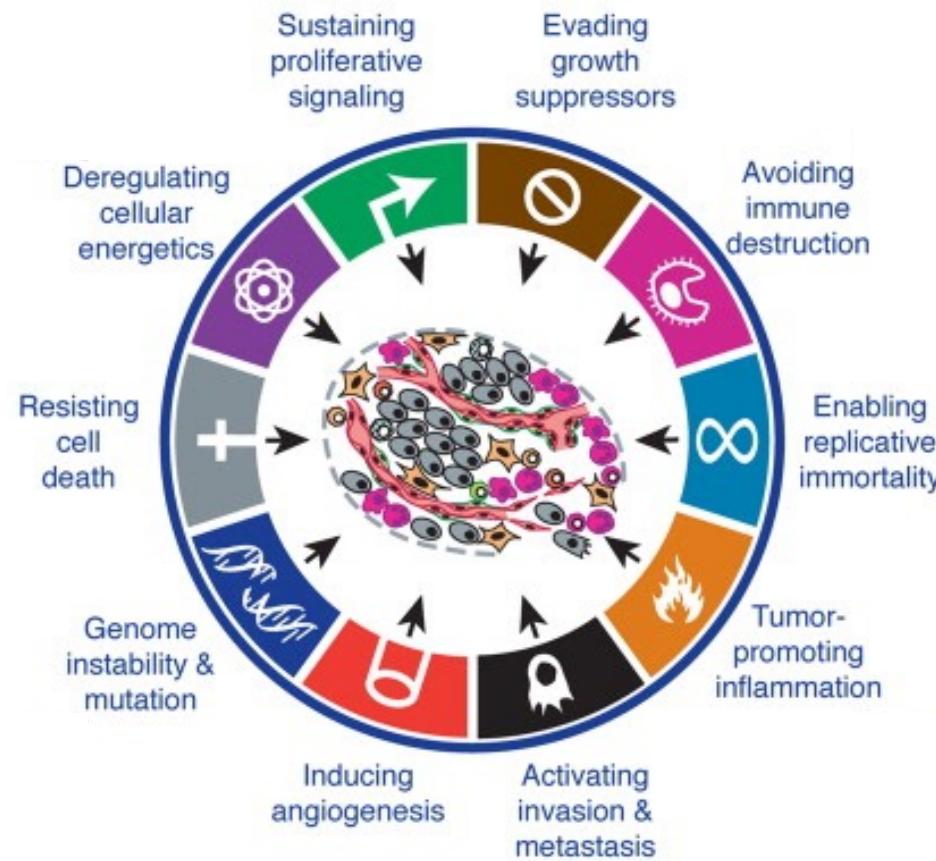
Stratton (2013)

The acquisition of mutations is a life-long process



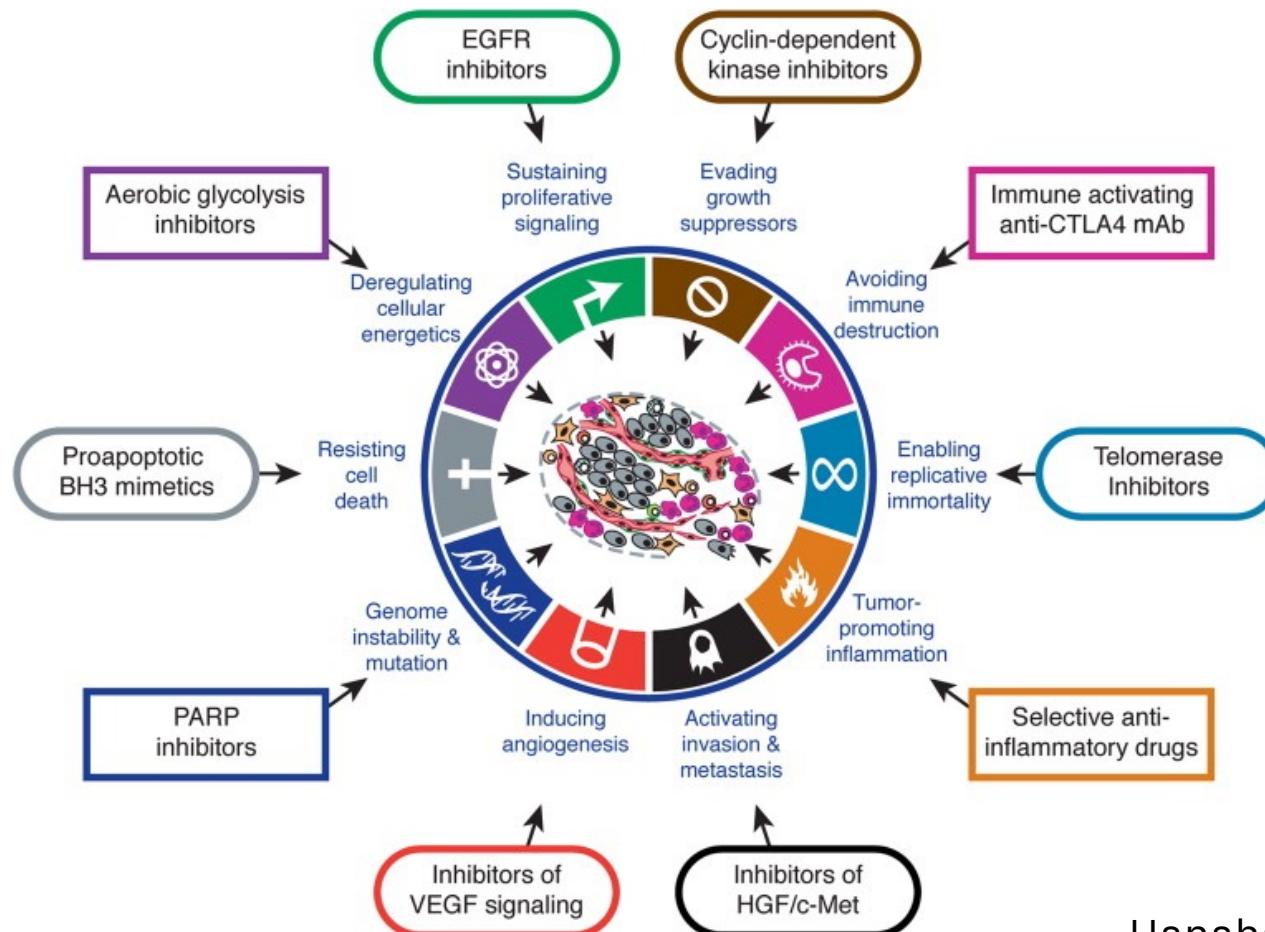
Stratton (2013)

Mutations facilitate the acquisition of cancer hallmarks



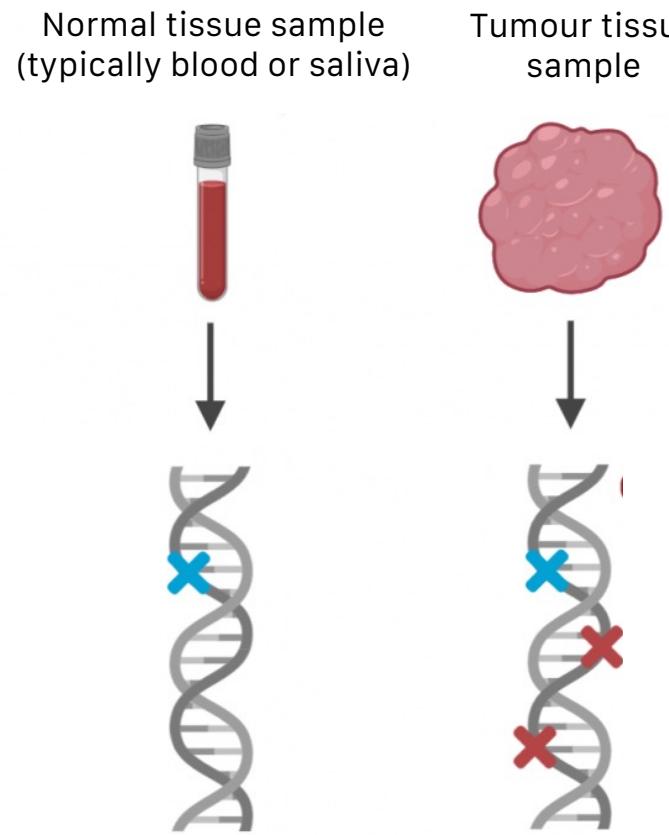
Hanahan and Weinberg (2011)

Targeting these hallmarks can lead to effective treatments

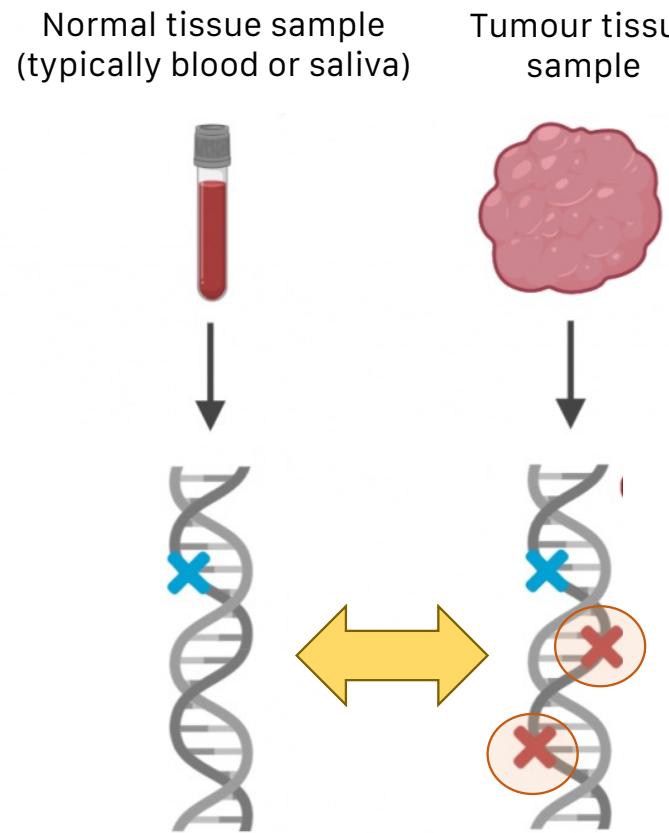


Hanahan and Weinberg (2011)

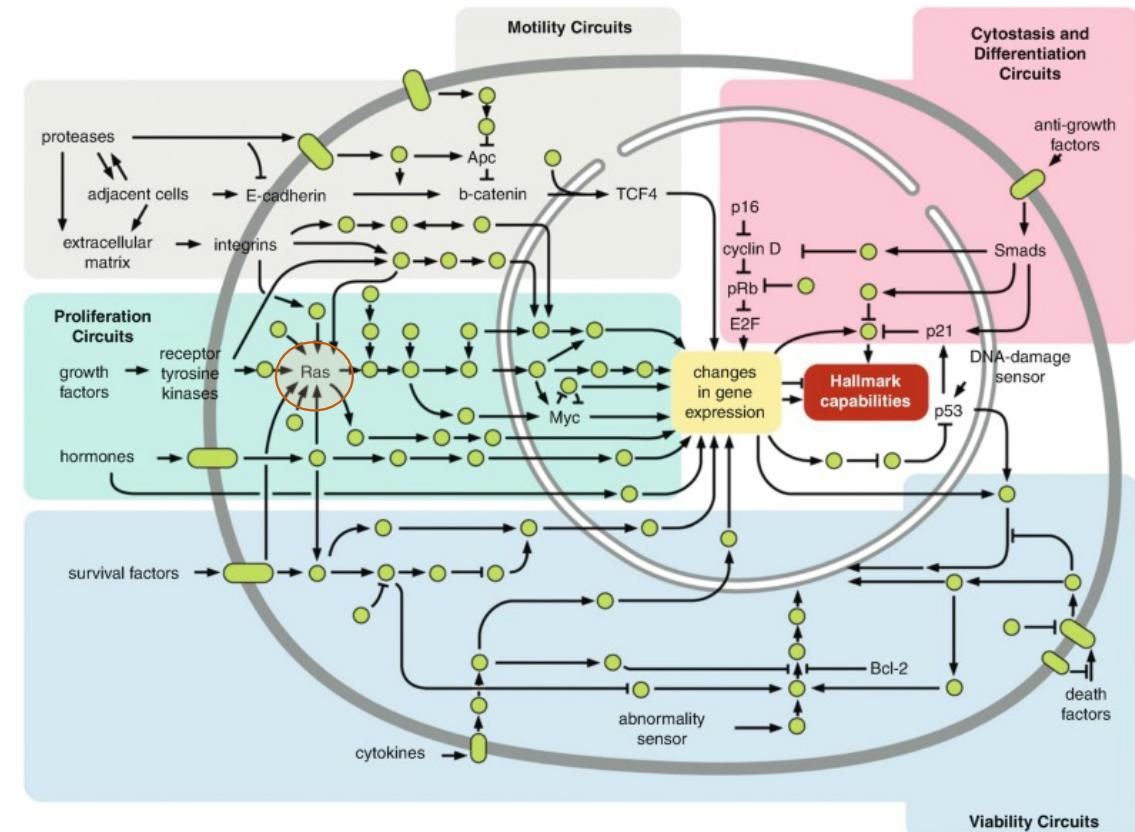
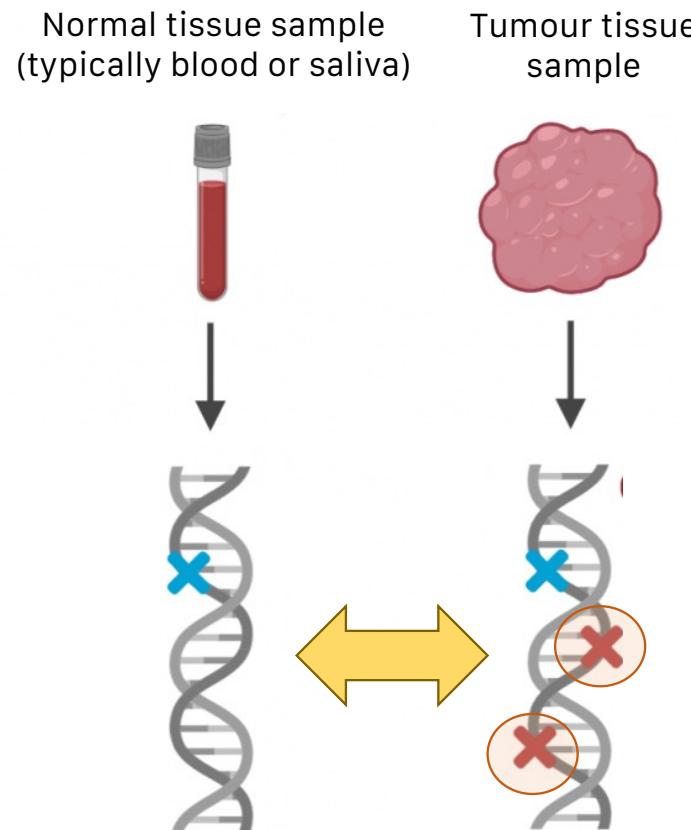
The study of genomes can help us identify the processes that fuel cancer development



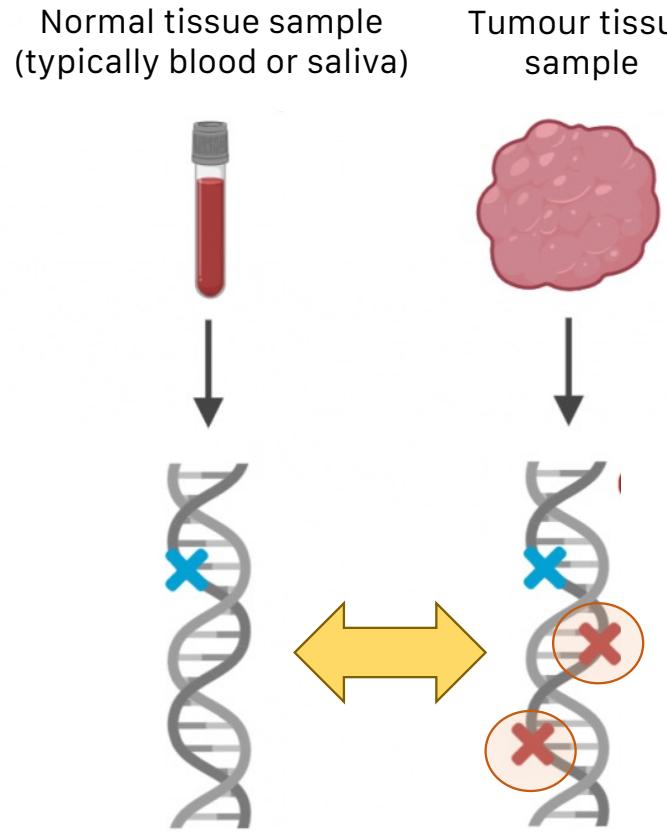
The study of genomes can help us identify the processes that fuel cancer development



The study of genomes can help us identify the processes that fuel cancer development

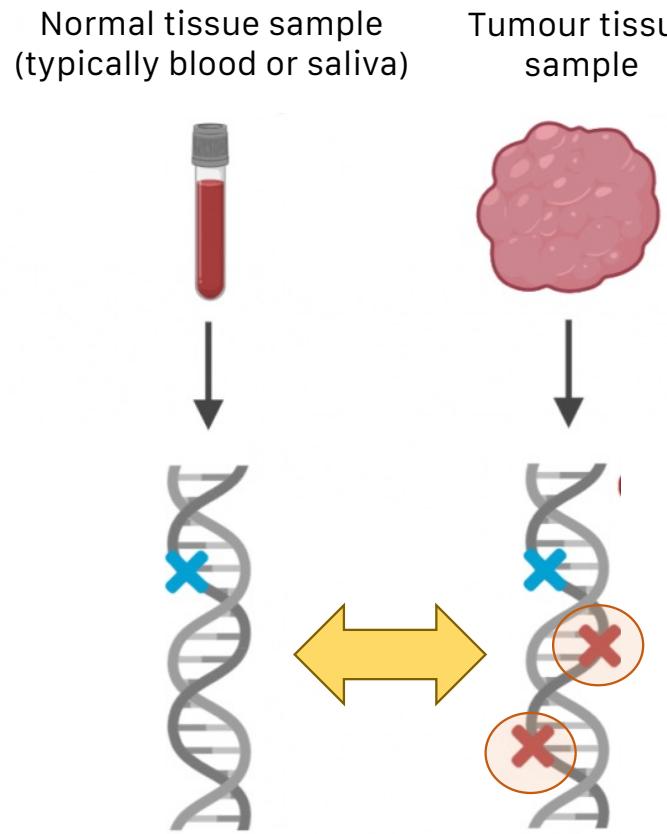


Studying the genome



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGCTTCCTGGCCCCACCATTCCAGGAAACATCTTGTCA
+
3IIIIIIIIIIII>1IIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAGATAGTCTTATCTTTTAAAAGAAAT
+
IIIIIGII.AIII=?I9G-/II=-+I=4?761BA2C9I+5A711+&>1$/I
```

Studying the genome



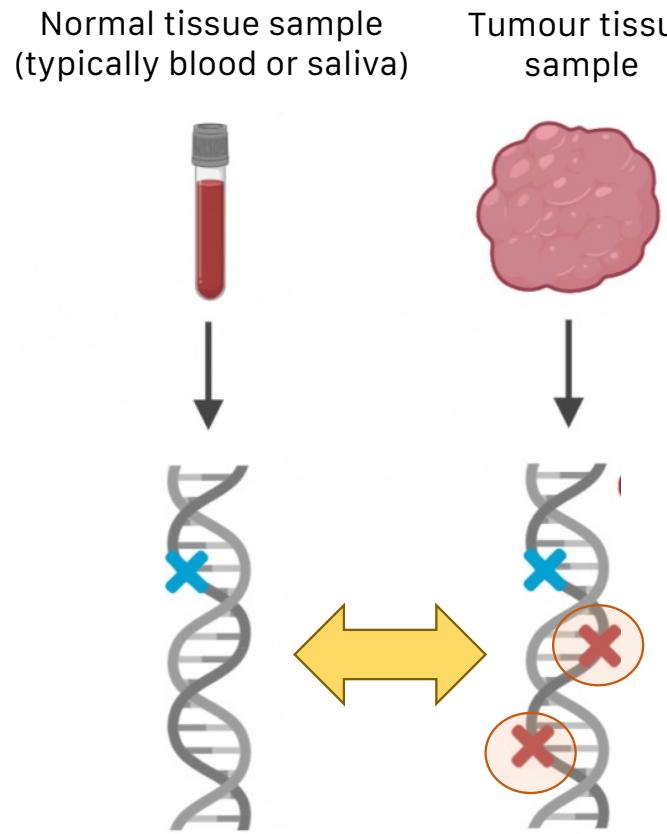
```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTTCTCCCTGCTTCCTGGCCCCACCATTCCAGGAAACATCTTGTCA
+
3IIIIIIIIIIII>IIIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAGATAGTCTTATCTTTTAAAAGAAAT
+
IIIIIGII.AIIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```



ACGT T TAGCAT
ACGT T CAGCAT

The text shows two lines of DNA sequence. The first line is "ACGT T TAGCAT". The second line is identical except for the fourth base pair, which is highlighted in yellow as "CAGCAT". This visual comparison highlights a mutation or change in the DNA sequence between the normal and tumour samples.

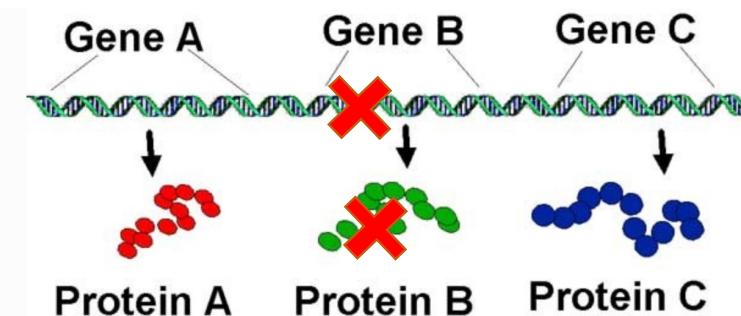
Studying the genome



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTTCTCCCTGCTTCCTGGCCCCACCATTTCAGGAAACATCTTGTCA
+
3IIIIIIIIIIII>1IIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCAATTAAGATAAGTCTTATCTTTTAAAGAAAT
+
IIIIIGII.AIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```

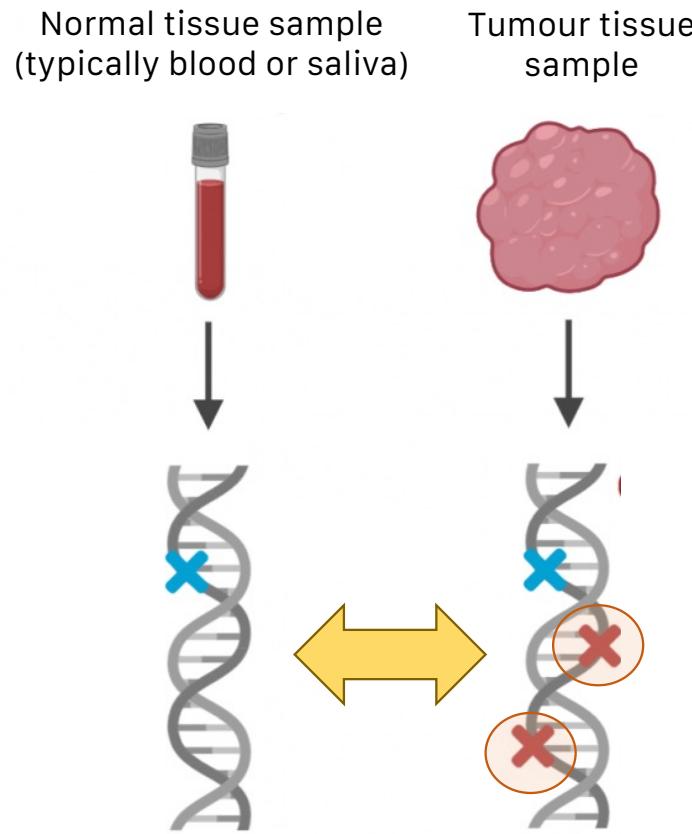


ACGT T TAGCAT
ACGT T CAGCAT



- Actionable mutations
- Driver genes
- Patterns of mutations

Studying the genome



Normal tissue sample (typically blood or saliva) → DNA (represented by a DNA double helix with a blue 'X' on one strand) → Comparison (indicated by a double-headed yellow arrow) → Tumour tissue sample (represented by a pink tissue sample with a blue 'X' on one strand and red 'X's on both DNA strands).

Illumina MiSeq (sequencing machine) and **laptop** (data analysis).

Sequence data (FASTQ format):

```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGCTTCCTGGCCCCACATTCCAGGAAACATCTTGTCA
+
3IIIIIIIIIIII>1IIIFF9BG08E00I%IG+&?(4)%00646.C1#&
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAGATAGTCTTATCTTTTAAAAGAAA
+
IIIIIGII.AIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/
```

ACGT T TAGCAT
ACGT T CAGCAT

Genes: Gene A, Gene B, Gene C

Proteins: Protein A, Protein B, Protein C

D Lee

- Actionable mutations
- Driver genes
- Patterns of mutations

This course!

Topics we will cover

- Data formats and QC



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGTTCTGGCCCCACATTCCAGGGAACATCTTGTCA
+
3IIIIIIIIII>1IIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAAGATAGTCTTATCTTTAAAAGAAAT
+
IIIIIGII.AIII=I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```

- How do I interpret sequencing and mutation files?
- How do I know if my data looks good?

Topics we will cover

- Data formats and QC
- Mutation calling



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGTTCTGGCCCCACATTCCAGGGAACATCTTGTAT
+
3IIIIIIIIII>1IIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAAGATAGTCCTATCTTTAAAAGAAAT
+
IIIIIGII.IIIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```



ACGTTTAGCAT
ACGTTCAGCAT****

- How do I identify mutations in my sequencing data?
- How do I know this is a good quality set of mutations?

Topics we will cover

- Data formats and QC
- Mutation calling
- Identifying gene drivers



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGTTCTGGCCCCACATTCCAGGGAACATCTTGTAT
+
3IIIIIIIIIIII>IIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAAGATAGTCCTATCTTTAAAAGAAAT
+
IIIIIGII.IIIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```

ACGTT TAGCAT
ACGTT CAGCAT

4AS PTPRB PAPC1 MEOM MED12 FAM46C MYD88
STK11 NFEA SPOP PBRM1 ELF3 MAP3K1 B1
SMARCA4 CTNNB1 FAT1 FBXW7 NOTCH1
MAD4 LRP1B APC KMT2C N
iFR KMT2D TP53 KRAS A
STAG2 IDH1 BRAF PIK3CA PTE
BP RB1 NRAS TBX3 FAT4 CDKN2A
IT3 NFE2L2 ERBB2 EP300 KEAP1
3RCA2 HRAS FGFR3 RBM10 ZFHX3 NCOR1
D AMER1 DNMT3A CDKN1A KIT ALK PI
GTF2I TSC1 HLA-B U2AF1 TNFRSF14

- How do I identify which genes are driving the cancer I'm studying?

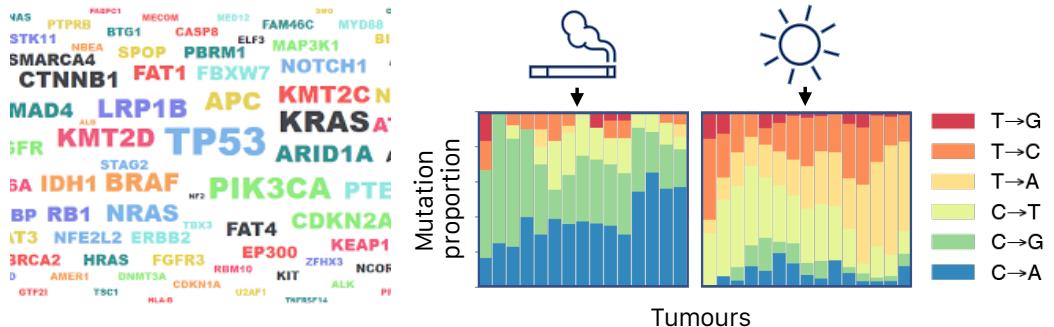
Topics we will cover

- Data formats and QC
- Mutation calling
- Identifying gene drivers
- Mutational signature analysis



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGTTCTGGCCCCACATTCCAGGGAACATCTTGTAT
+
3IIIIIIIIIIII>IIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAAGATAGTCCTATCTTTAAAAGAAAT
+
IIIIIGII.IIIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```

↓
ACGTT TAGCAT
ACGTT CAGCAT ↗



- How do I identify patterns of mutations in my samples?
- How do I associate them to etiological factors?

Topics we will cover

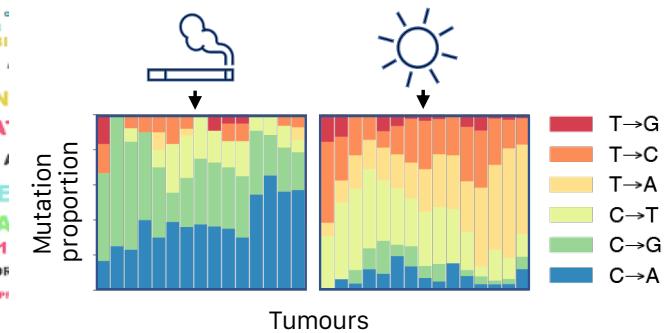
- Data formats and QC
- Mutation calling
- Identifying gene drivers
- Mutational signature analysis
- Special sessions: Clinical challenges and the role of genomics, tumour board



```
@ERR000589.41 EAS139_45:5:1:2:111/1
CTTCCTCCCTGTTCTGGCCCCACATTCCAGGGAACATCTTGTAT
+
3IIIIIIIIIIII>IIIIFF9BG08E00I%IG+&?(4)%00646.C1#&(
@ERR000589.42 EAS139_45:5:1:2:1293/1
AGTTGTTAAATCCAAGCCAATTAAAGATAGTCCTATCTTTAAAAGAAAT
+
IIIIIGII.IIIII=?I9G-/II=+I=4?761BA2C9I+5A711+&>1$/I
```

↓
ACGTT TAGCAT
ACGTT CAGCAT ↗

4AS PTPRB PAPC1 MECOM MED12 FAM46C MYD88
STK11 NFEA SP0P PBRM1 ELF3 MAP3K1 B1
SMARCA4 CTNNB1 FAT1 FBXW7 NOTCH1
MAD4 LRP1B APC KMT2C N
iFR KMT2D TP53 KRAS A
STAG2 BRAF PIK3CA PTE
5A IDH1 NRAS TBX3 FAT4 CDKN2A
BP RB1 ERBB2 EP300 KEAP1
IT3 NFE2L2 FGFR3 ZFHX3 NCOB
3RCA2 HRAS CDKN1A KIT ALK PI
D AMER1 DNMT3A RBM10 U2AF1 TNFRSF14
GTF2I TSC1 HLA-B CDKN1U U2AF1 PI



- How is genomics used to guide clinical decisions?

Cancer develops as cells acquire mutations



Mutation

DAY 2:

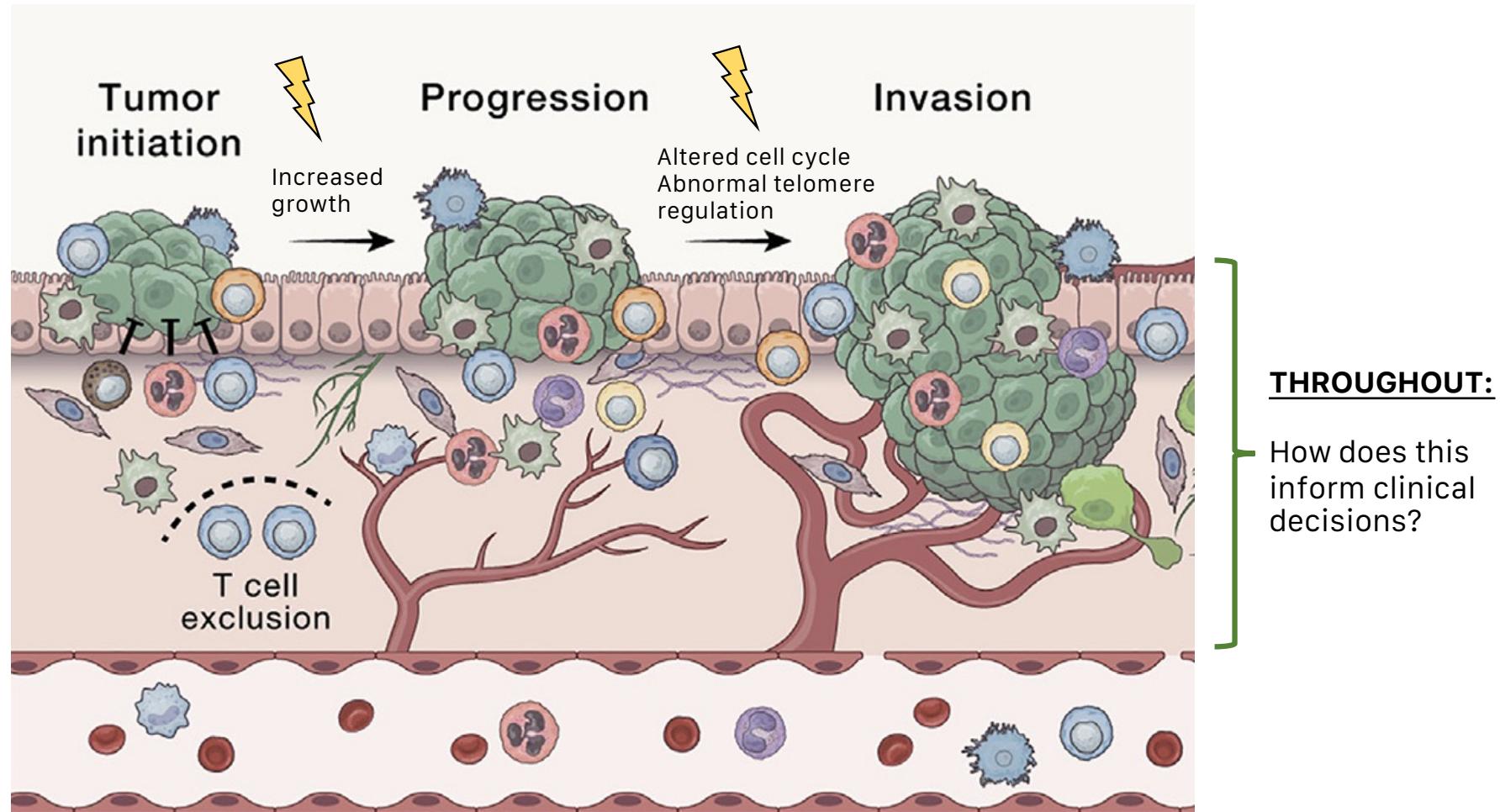
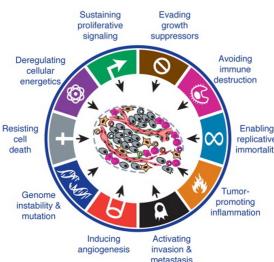
- Identification of mutations

DAY 4:

- Identification of potential etiological factors

DAY 3:

- Identification of driver genes



De Visser and Joyce (2023)

Topics we will cover

Sunday October 26		Cancer Genome Analysis 26 October - 31 October 2025, Khon Kaen University, Thailand					Monday October 27		Tuesday October 28		Wednesday October 29		Thursday October 30		Friday October 31	
8:00			Bus to venue		Bus to venue		Bus to venue		Bus to venue		Bus to venue		Bus to venue		8:00	
8:30		Unraveling Cholangiocarcinoma: Clinical Challenges and the Role of Genomics Sirinya Sitthirak												Multidisciplinary Tumor Board: Integrating Genomics into Clinical Practice Thanit Immekamon Apinya Jusakul	8:30	
9:00				Mutation Calling Daniela Robles, Eric Dawson, Sirinya Sitthirak			Driver gene & oncoplets Nivedita Mukherjee, Eric Dawson				Mutational signature analysis Eric Dawson, Jia Wern Pan, Daniela Robles			9:00		
9:30		Data formats & QC Daniela Robles & Arporn (Koi) Wangwiwatsin												Group Projects Eric Dawson Support: All Training Team	9:30	
10:00															10:00	
10:30		Break		Break		Break		Break		Break		Break		10:30		
11:00														11:00		
11:30		Data formats & QC Daniela Robles & Arporn (Koi) Wangwiwatsin		Mutation Calling Daniela Robles, Eric Dawson, Sirinya Sitthirak			Driver gene & oncoplets Nivedita Mukherjee, Eric Dawson			Mutational signature analysis Eric Dawson, Jia Wern Pan, Daniela Robles			Group Projects Eric Dawson Support: All Training Team	11:30		
12:00	Registration and Lunch		Cis-regulatory effect of HPV Integrations in cervical cancer Sabarinathan Radhakrishnan (virtual)	Chornobyl to Tahiti: uncovering the effects of ionizing radiation in the genome using accessible accelerated computing designs - Eric Dawson			Genome-wide association studies in Cancer Onnappa Kongphan			Genomics-driven precision immunotherapy for Malaysian breast cancer Jia Wern Pan			12:00			
12:30														12:30		
13:00	Introductions WCS, KKU, Training Team, Participants		Lunch		Lunch		Lunch		Lunch		Lunch		13:00			
13:30														13:30		
14:00														14:00		
14:30	Intro to the course Daniela Robles		Data formats & QC Daniela Robles & Arporn (Koi) Wangwiwatsin	Mutation Calling Daniela Robles, Eric Dawson, Sirinya Sitthirak		Driver gene & oncoplets Nivedita Mukherjee / Eric Dawson			Mutational signature analysis Eric Dawson, Jia Wern Pan, Daniela Robles		Project Presentations		14:30			
15:00		Break											Wrap-up and close	15:00		
15:30	Advanced Learning Skills (ALS) Martin Aslett, Isabela Malta		Break		Break		Break		Break		Bus to hotel		15:30			
16:00			Discussions: Data sharing and data management Jia Wern Pan		Mutation Calling Daniela Robles, Eric Dawson, Sirinya Sitthirak		Germline mutations calling / UK Biobank Eric Dawson		Mutational Signatures and Risk Factors: Decoding the Origins of Cholangiocarcinoma Sirinya Sitthirak				16:00			
16:30		Intro to Cancer Genomics Arporn (Koi) Wangwiwatsin												16:30		
17:00	Cancer genomics & research in Asia (Online Databases) Apinya Jusakul		Discussion: Ethical consideration when retrieving samples from patients. Daniela Robles	Mutation Calling in Cholangiocarcinoma: Identifying Key Genetic Alterations for Clinical Decision-Making - Sirinya Sitthirak		Driver Genes in Cholangiocarcinoma: Connecting Genomic Alterations to Clinical Outcomes - Sirinya Sitthirak			Group Projects Eric Dawson, Daniela Robles Support: All Training Team				17:00			
17:30	Networking and Dinner		Day wrap-up		Day wrap-up		Day wrap-up		Day wrap-up					17:30		
18:00			Bus to hotel		Bus to hotel		Bus to restaurant		Bus to hotel		Dinner at restaurant			18:00		
								Bus to hotel								

Seminars on cancer genomics and related work in Latin America

- Sabarinathan Radhakrishnan (India)
- Eric Dawson (USA)
- Onnapa Kongphan (Thailand)
- Jia Wern Pan (Malaysia)

Training on Advanced Learning

- Martin Aslett and Isabela Malta (WGCAC, UK)

Group projects



- Melanoma
- Lung
- Pancreatic
- Liver

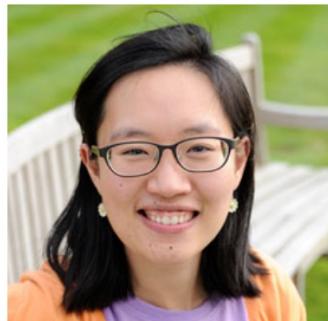


Marcos Díaz Gay

Training team



[Daniela Robles-Espinoza](#)
LIIGH-UNAM, Mexico



[Arporn \(Koi\) Wangwiwatsin](#)
Khon Kaen University, Thailand



[Eric Dawson](#)
EarthFrame, Corp.
Phileal, LLC



[Apinya Jusakul](#)
Khon Kaen University, Thailand



[Sirinya Sitthirak](#)
Khon Kaen University, Thailand



[Sabarinathan Radhakrishnan](#)
The National Centre for Biological Sciences, India



[Onnapa Kongphan](#)
Khon Kaen University, Thailand



[Jia Wern Pan](#)
Cancer Research Malaysia, Malaysia



[Haslina Makmur](#)
Cancer Research Malaysia, Malaysia



[Nivedita Mukherjee](#)
The National Centre for Biological Sciences, India

Organisers



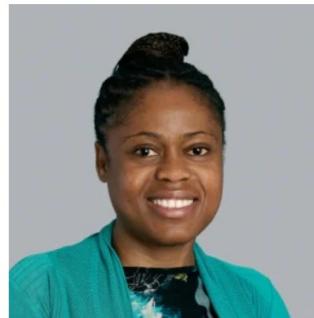
Monica Abrudan



[Martin Aslett](#)



[Isabela Malta](#)



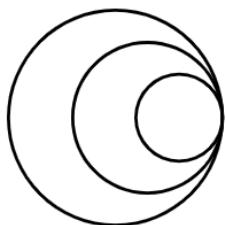
[Alice Matimba](#)



Karon Chappell



[Vaishnavi Gangadhar](#)



wellcome
connecting
science



มหาวิทยาลัยขอนแก่น
KHON KAEN UNIVERSITY

