

**wellcome
connecting
science**

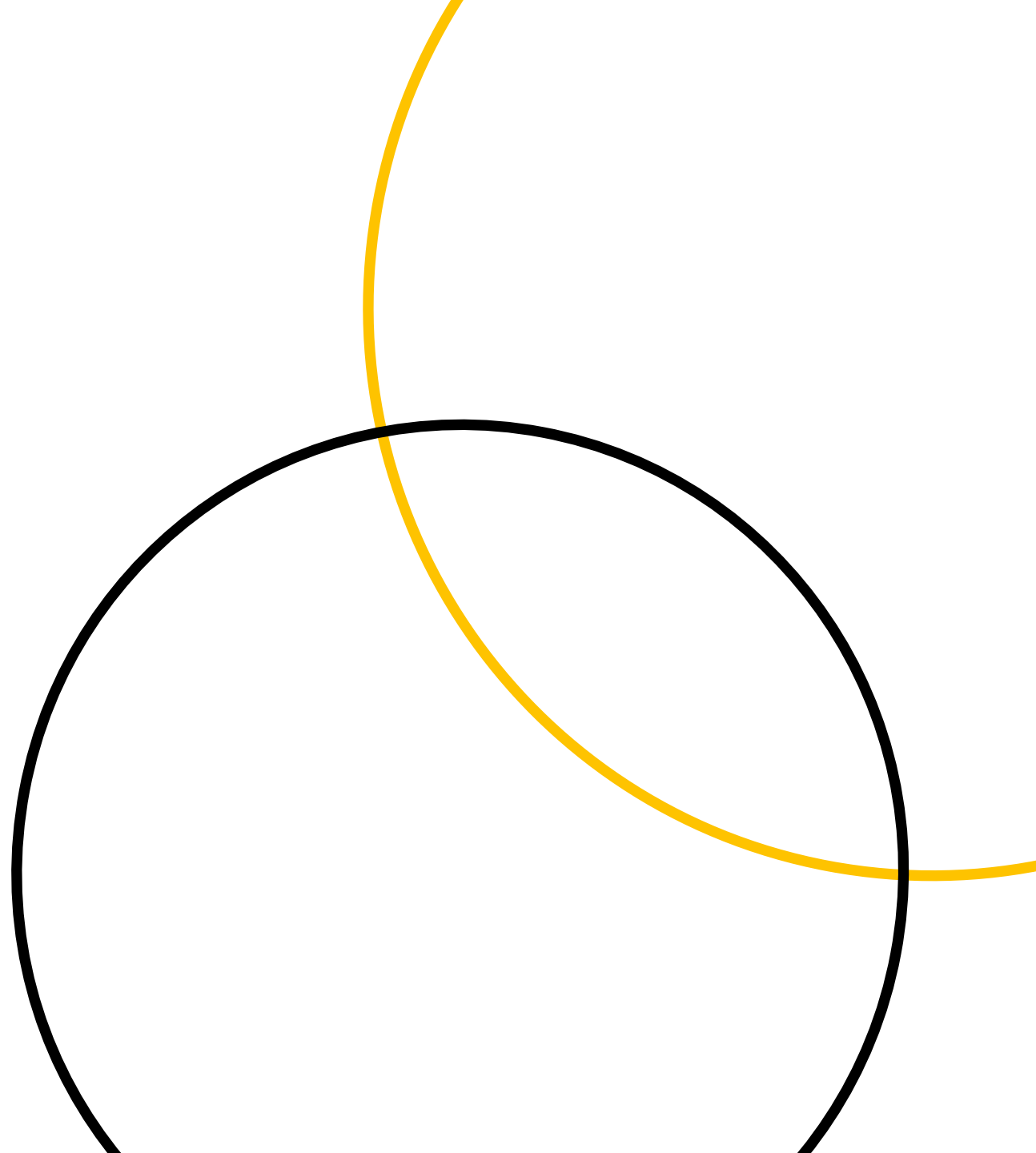
connecting people with science

Dr Jia-Wern Pan
Cancer Research Malaysia

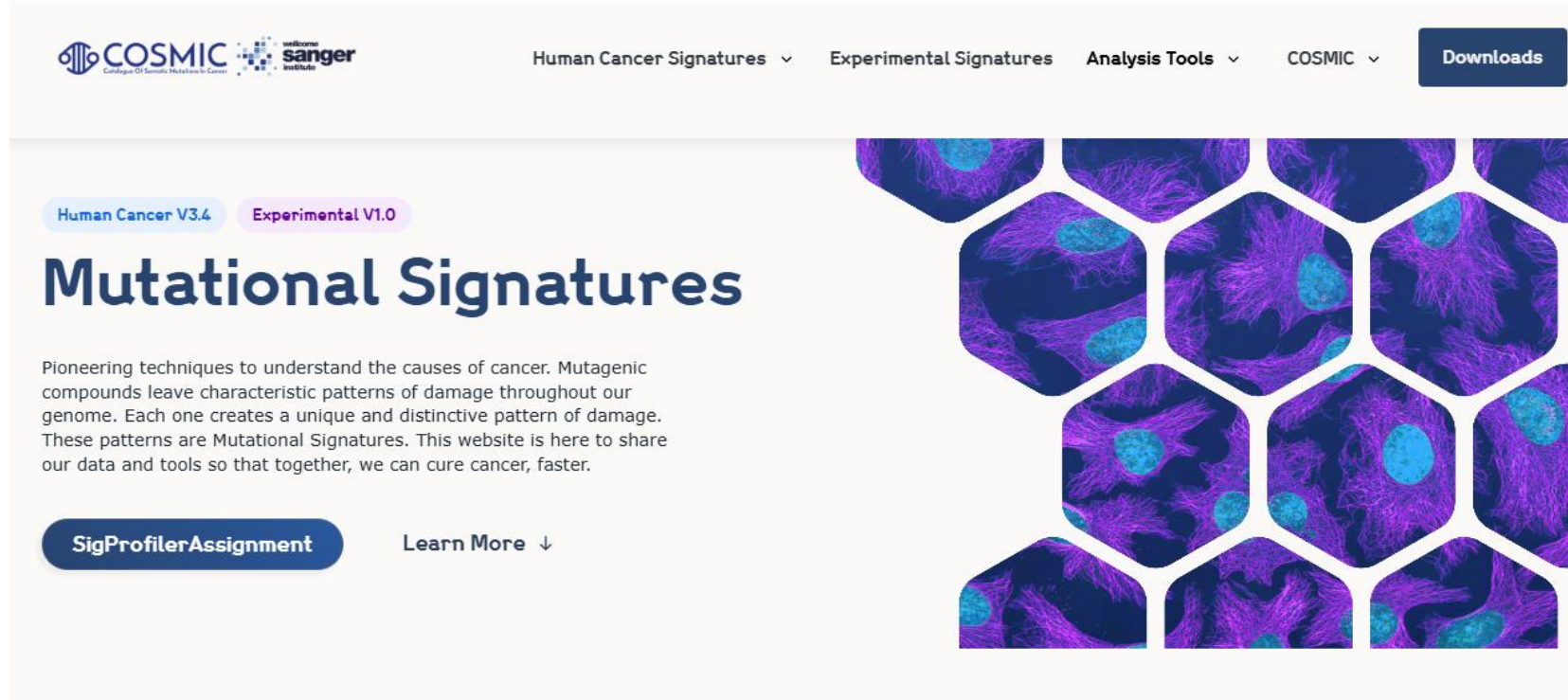
Mutational signatures:
Databases, aetiology, and clinical utility

(adapted from Dr Marcos Díaz Gay and
Dr Mariano Golubicki, WCS CGA 2023)

April 2025



COSMIC Mutational Signatures Database



October v3.4 (June 2023)

Genomics

- 99 SBS signatures
- 20 DBS signatures
- 23 ID signatures
- 25 CN signatures
- 10 SV signatures
- 5 RNA-SBS signatures

Experimental

- 140 signatures

The COSMIC mutational signature database

Mini-practical 1:

Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS
- ID
- CNV
- SV
- RNA-SBS

The COSMIC mutational signature database

Mini-practical 1:

Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS (78 classes of strand-agnostic doublet base substitutions)
- ID
- CNV
- SV
- RNA-SBS

The COSMIC mutational signature database

Mini-practical 1:

Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS
- ID (83 classes based on size, nucleotides affected and presence on repetitive and/or microhomology regions)
- CNV
- SV
- RNA-SBS

The COSMIC mutational signature database

Mini-practical 1:

Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS
- ID
- CNV (48 classes based on loss-of-heterozygosity status, total copy number state, and segment length of allele-specific copy number segments)
- SV
- RNA-SBS

April 2025

The COSMIC mutational signature database

Mini-practical 1:

Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS
- ID
- CNV
- SV (32 classes based on deletions, inversions, and tandem duplications, size, and distance between adjacent SVs)
- RNA-SBS

The COSMIC mutational signature database

Mini-practical 1:

Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS
- ID
- CNV
- SV
- RNA-SBS (192 stranded trinucleotide classes)

The COSMIC mutational signature database

Mini-practical 1:

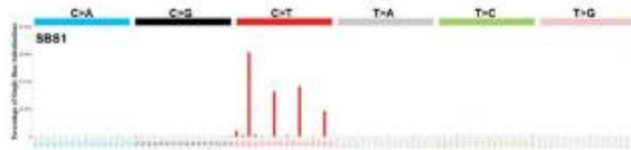
Go to the COSMIC website. The input matrix for SBS consists of 96 classes based on the trinucleotide context of single base substitutions. What are the input matrix classes for:

- DBS
- ID
- CNV
- SV
- RNA-SBS (192 stranded trinucleotide classes)

Linking mutational signatures to specific causes

Many signatures in COSMIC have a specific proposed aetiology, i.e. SBS 4 is linked with tobacco smoking.

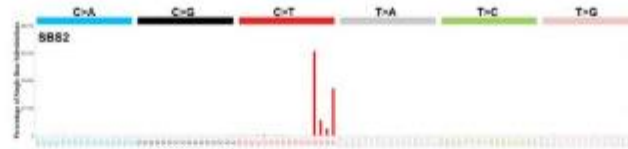
How do we know this? What is the evidence?



SBS1

Proposed Aetiology

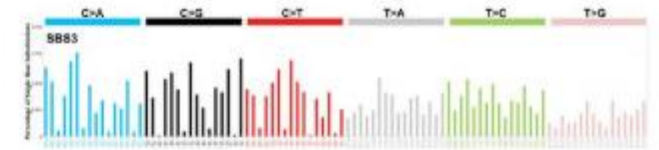
Spontaneous deamination of 5-methylcytosine
(clock-like signature)



SBS2

Proposed Aetiology

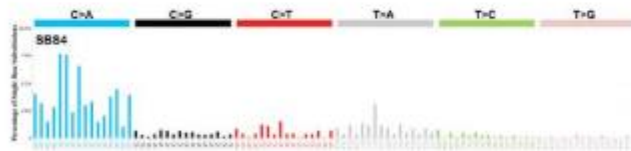
Activity of APOBEC family of cytidine deaminases



SBS3

Proposed Aetiology

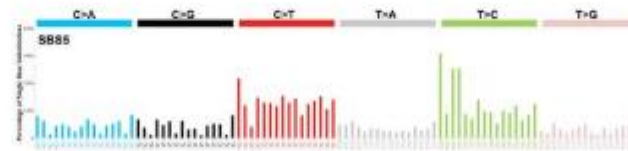
Defective homologous recombination DNA
damage repair



SBS4

Proposed Aetiology

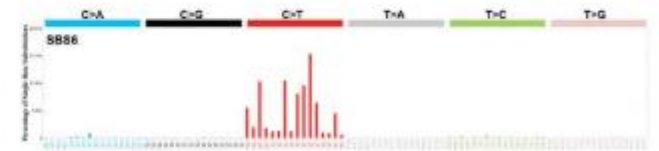
Tobacco smoking



SBS5

Proposed Aetiology

Unknown (clock-like signature)



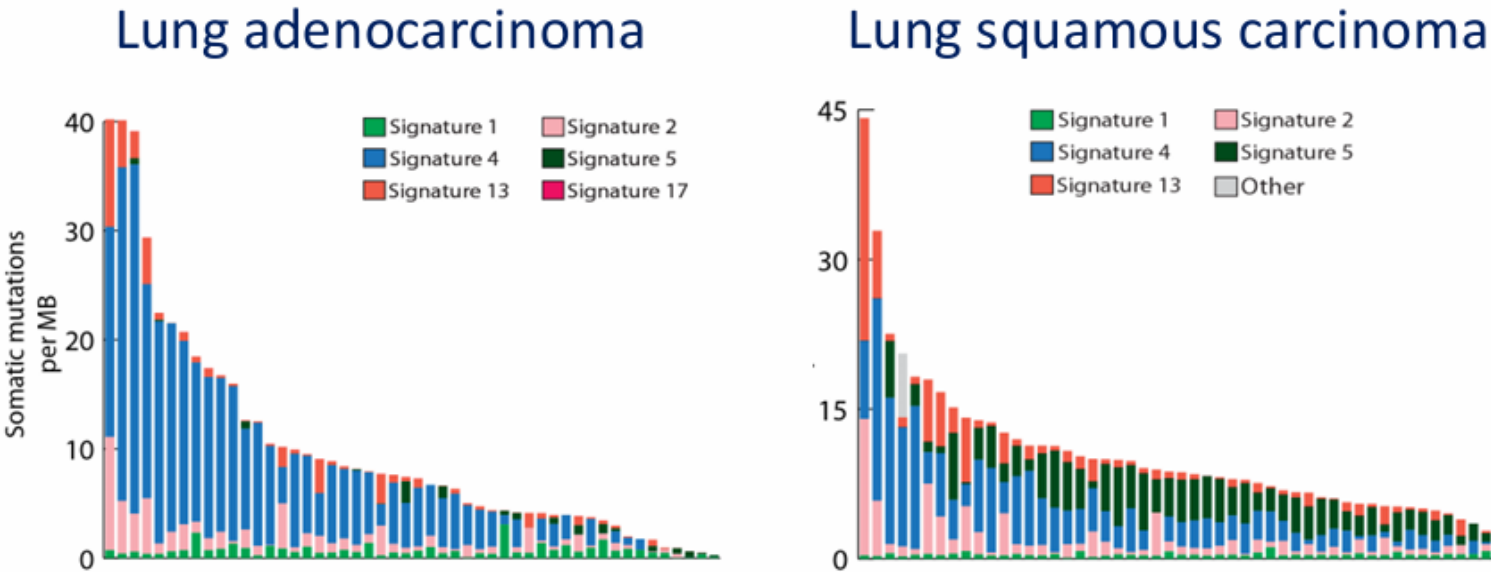
SBS6

Proposed Aetiology

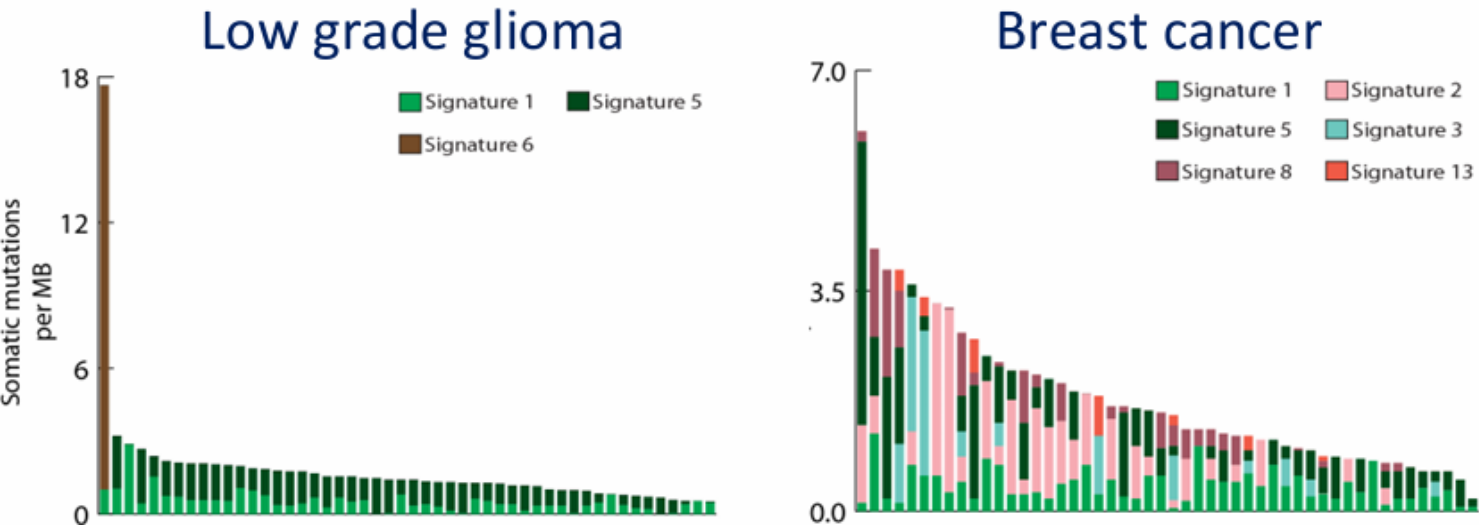
Defective DNA mismatch repair

Signature 4 is associated with smoking induced cancer types

Smoking induced cancer types

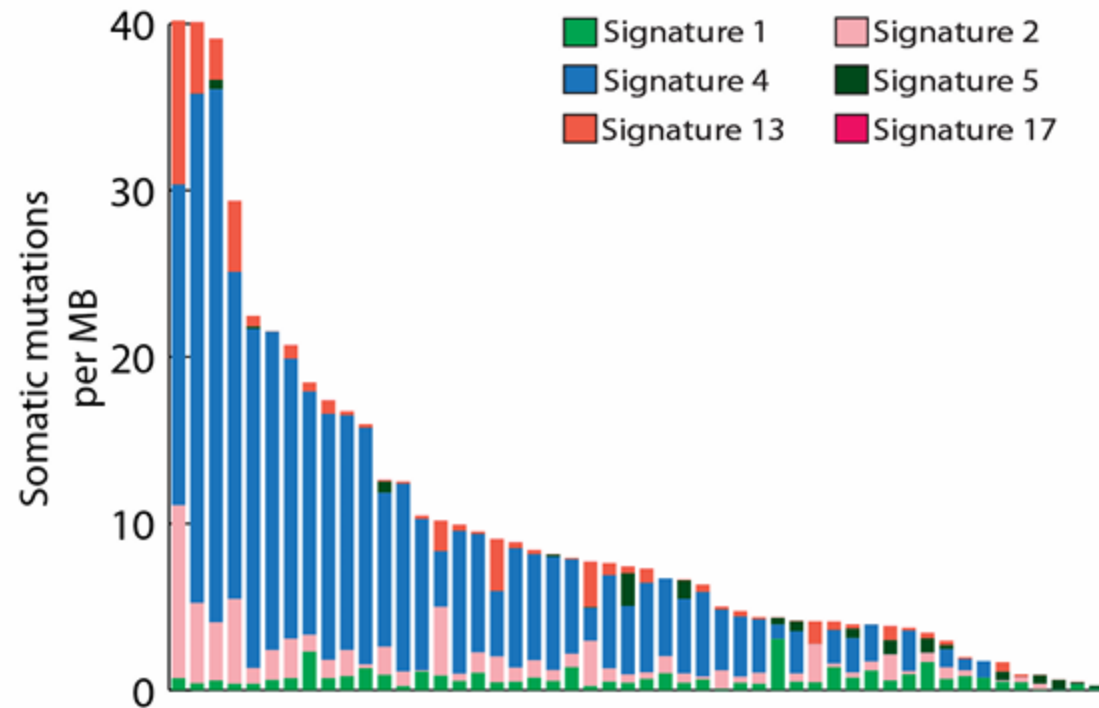


Non-smoking induced cancer types

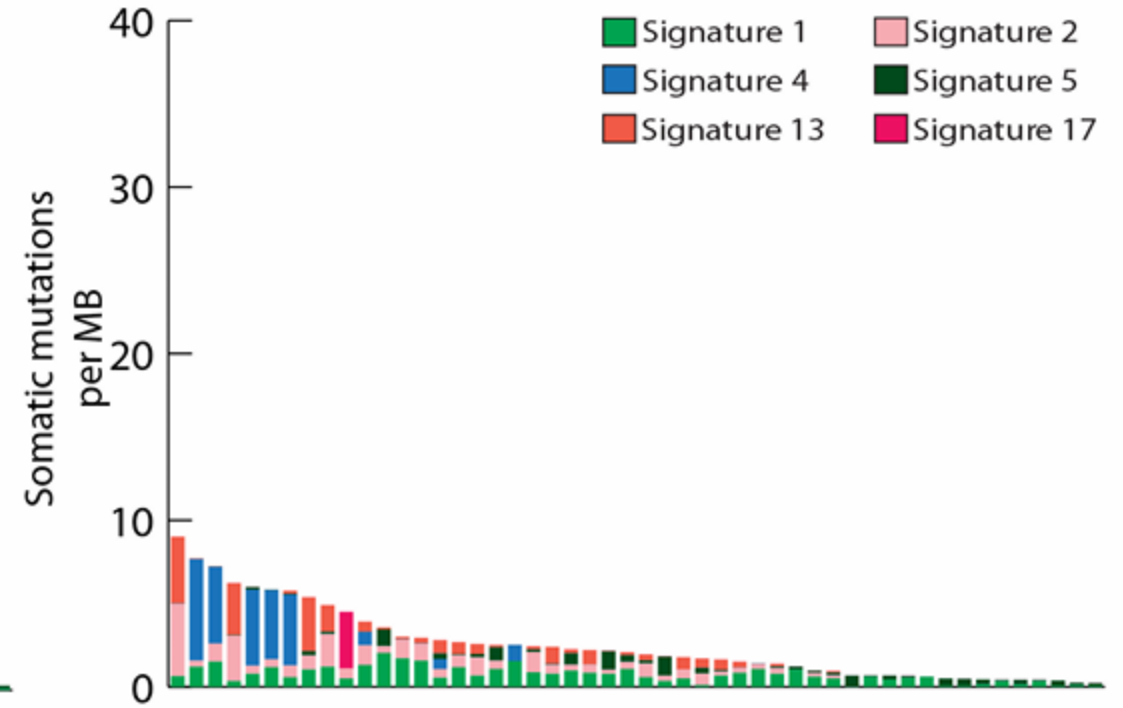


Within lung cancer, SBS 4 is much more prevalent in tobacco smokers compared to non-smokers

Tobacco smokers



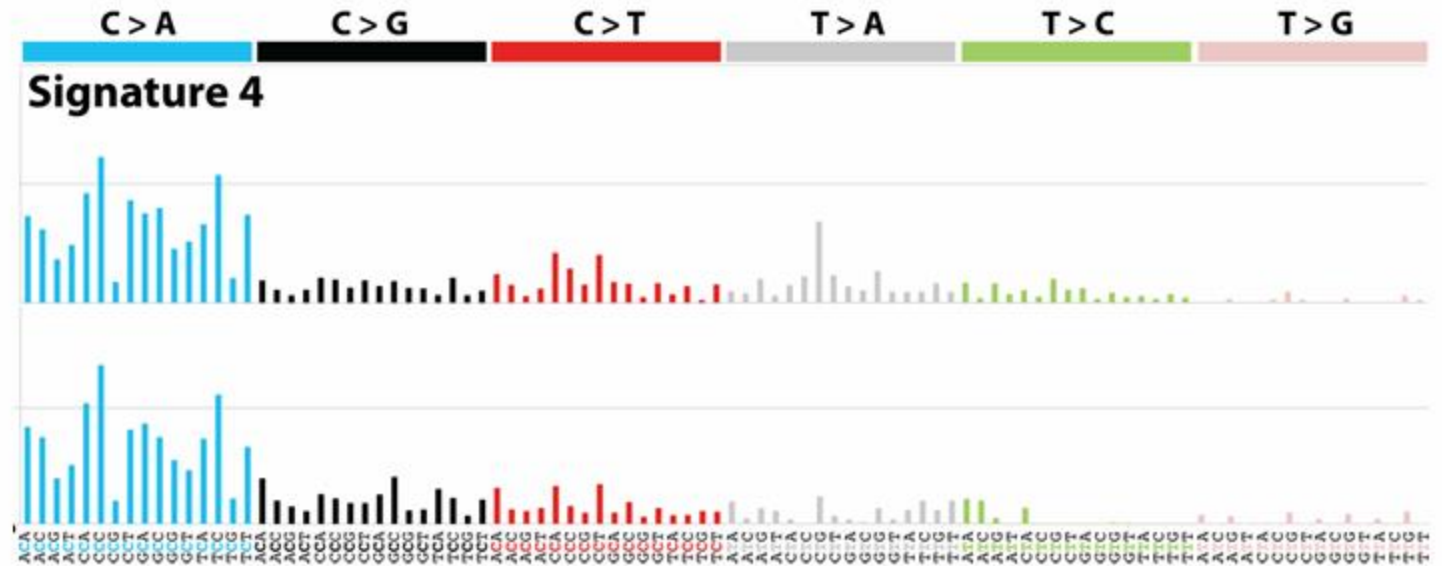
Life-long non-smokers



SBS 4 matches the signature of benzo[a]pyrene exposure, a common chemical component of cigarette smoke, in cell lines and mouse models

Signature 4 extracted from
human cancers

Signature of benzo[a]pyrene exposure *in vitro*



Evidence for the aetiology of SBS 4

- Identified only in cancer types epidemiologically known to be caused by tobacco smoking
- Highly enriched in tobacco smokers compared to non-smokers
- The pattern of SBS 4 matches *in vitro* experiments exposing cell lines to known tobacco carcinogens

Summary of evidence for SBS 4

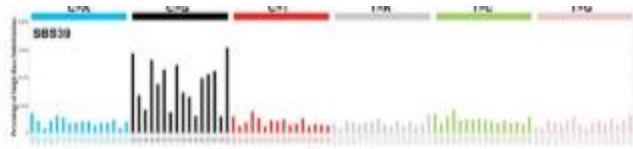
Background	Identification study		First included in COSMIC
	Alexandrov <i>et al.</i> 2013 Nature		v1
Identification	NGS technique	Different variant callers	Multiple sequencing centres
	<div><div></div> WES & WGS</div>	<div><div></div> Yes</div>	<div><div></div> Yes</div>
Technical validation	Validated in orthogonal techniques	Replicated in additional studies	Extended context enrichment
	<div><div></div> Yes</div>	<div><div></div> Yes</div>	-
Proposed aetiology	Mutational process		Support
	<div><div></div> Tobacco smoking</div>		<div><div></div> Experimental confirmation</div>
Experimental validation	Experimental study		Species
	<div><div></div> Nik-Zainal <i>et al.</i> 2015 Mutagenesis</div>		<div><div></div> Mouse</div>

Summary of evidence for SBS 92

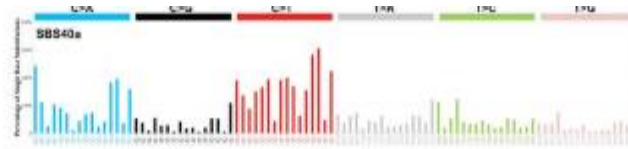
Background	Identification study		First included in COSMIC
	Lawson et al. 2020 Science		v3.2
Identification	NGS technique	Different variant callers	Multiple sequencing centres
	● WGS	● Yes	● Yes
Technical validation	Validated in orthogonal techniques	Replicated in additional studies	Extended context enrichment
	● Yes	● Yes	-
Proposed aetiology	Mutational process		Support
	● Tobacco smoking		● Statistical association
Experimental validation	Experimental study		Species
	-		-

Also associated with tobacco smoking, but only in bladder cancer and no experimental validation

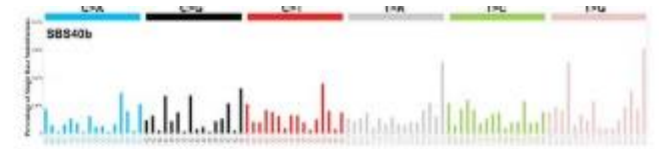
Many signatures still have unknown aetiology



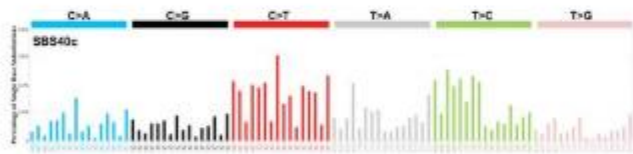
SBS39
Proposed Aetiology
Unknown



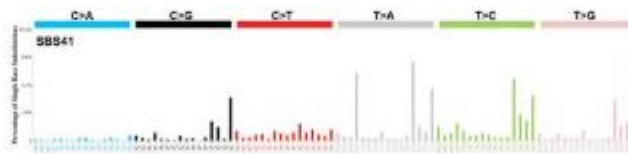
SBS40a
Proposed Aetiology
Unknown



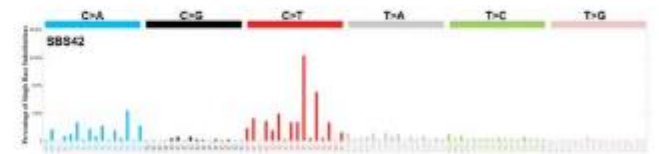
SBS40b
Proposed Aetiology
Unknown



SBS40c
Proposed Aetiology
Unknown



SBS41
Proposed Aetiology
Unknown



SBS42
Proposed Aetiology
Haloalkane exposure

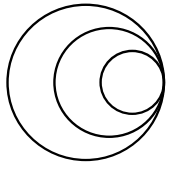
Caveats to consider

Discuss: How might the COSMIC database be biased in terms of mutational signatures included and also their proposed aetiologies?

Caveats to consider

Discuss: How might the COSMIC database be biased in terms of mutational signatures included and also their proposed aetiologies?

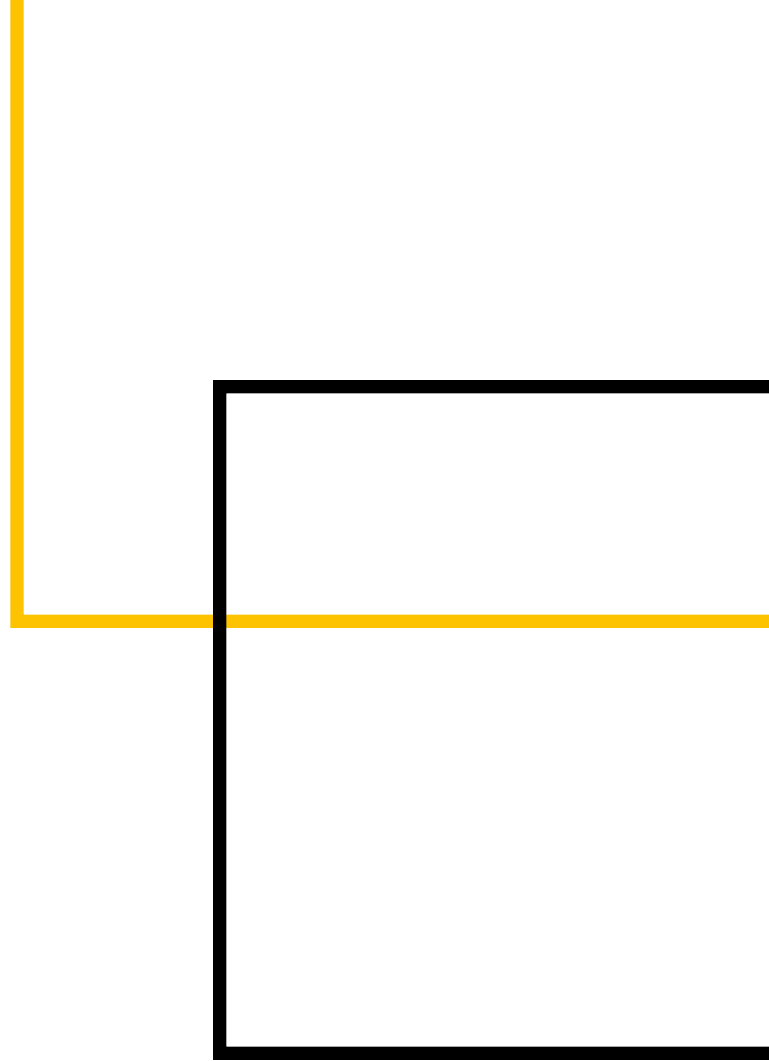
- Underlying cohorts primarily Western
- Relies mostly on short-read sequencing
- Biased towards common cancer types and common risk factors



**wellcome
connecting
science**

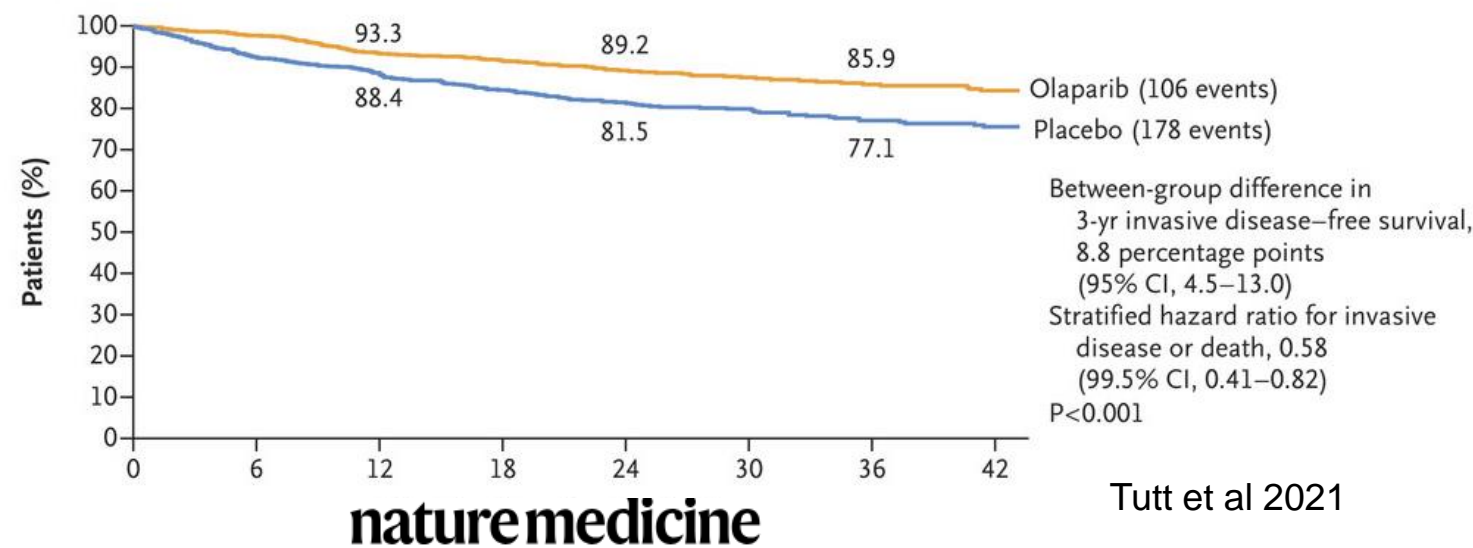
Clinical and research applications of mutational signatures

April 2025



Mutational signatures as predictive biomarkers for targeted therapy

A Invasive Disease-free Survival



Homologous recombination repair deficiency (HRD) is a common many cancer types

Patients with HRD may be more likely to respond to platinum chemotherapy and PARP inhibitor targeted therapy

Mutational signatures associated with HRD may be able to identify patients who should receive PARP inhibitors, particularly in breast and ovarian cancer

No. at Risk		
Olaparib	921	820
Placebo	915	807

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

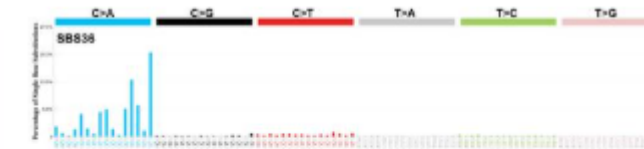
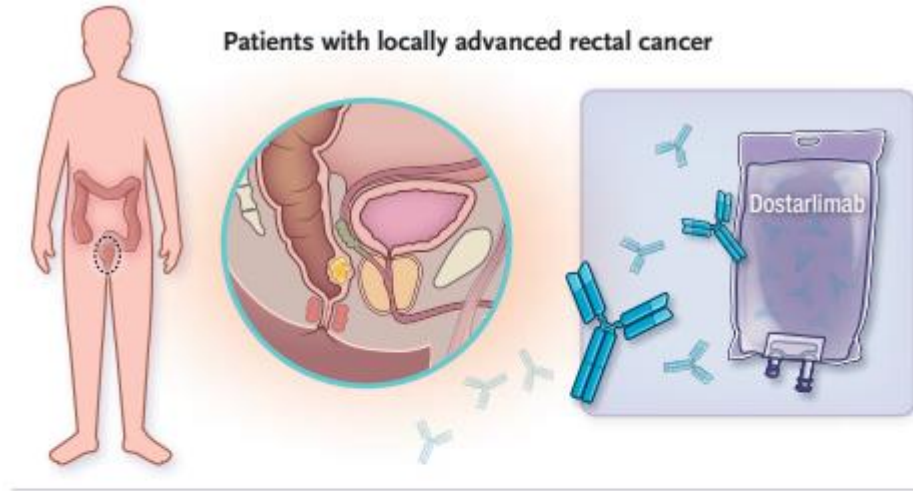
[nature](#) > [nature medicine](#) > [analyses](#) > article

Analysis | Published: 13 March 2017

HRDetect is a predictor of *BRCA1* and *BRCA2* deficiency based on mutational signatures

[Helen Davies](#), [Dominik Glodzik](#), [Sandro Morganella](#), [Lucy R Yates](#), [Johan Staaf](#), [Xueqing Zou](#), [Manasa Ramakrishna](#), [Sancha Martin](#), [Sandrine Boyault](#), [Anieta M Sieuwerts](#), [Peter T Simpson](#), [Tari A King](#), [Keiran Raine](#), [Jorunn E Eyfjord](#), [Gu Kong](#), [Åke Borg](#), [Ewan Birney](#), [Hendrik G Stunnenberg](#), [Marc J van de Vijver](#),

Mutational signatures as predictive biomarkers for immunotherapy

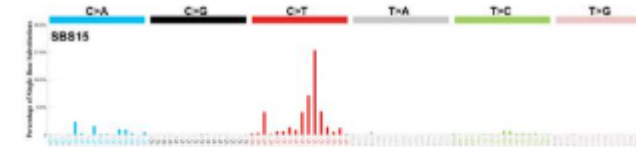
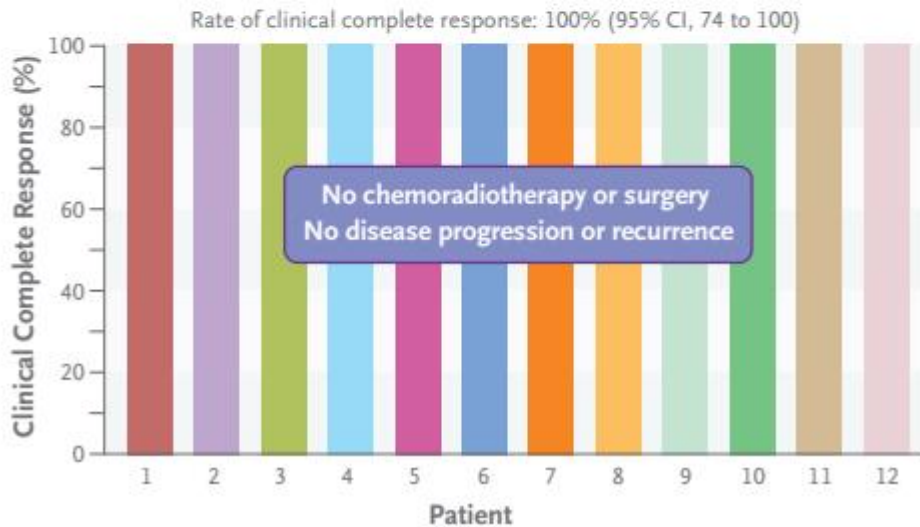


SBS36

Proposed Aetiology

Defective DNA base excision repair due to MUTYH mutations

Overall Response to Dostarlimab in 12 Patients



SBS15

Proposed Aetiology

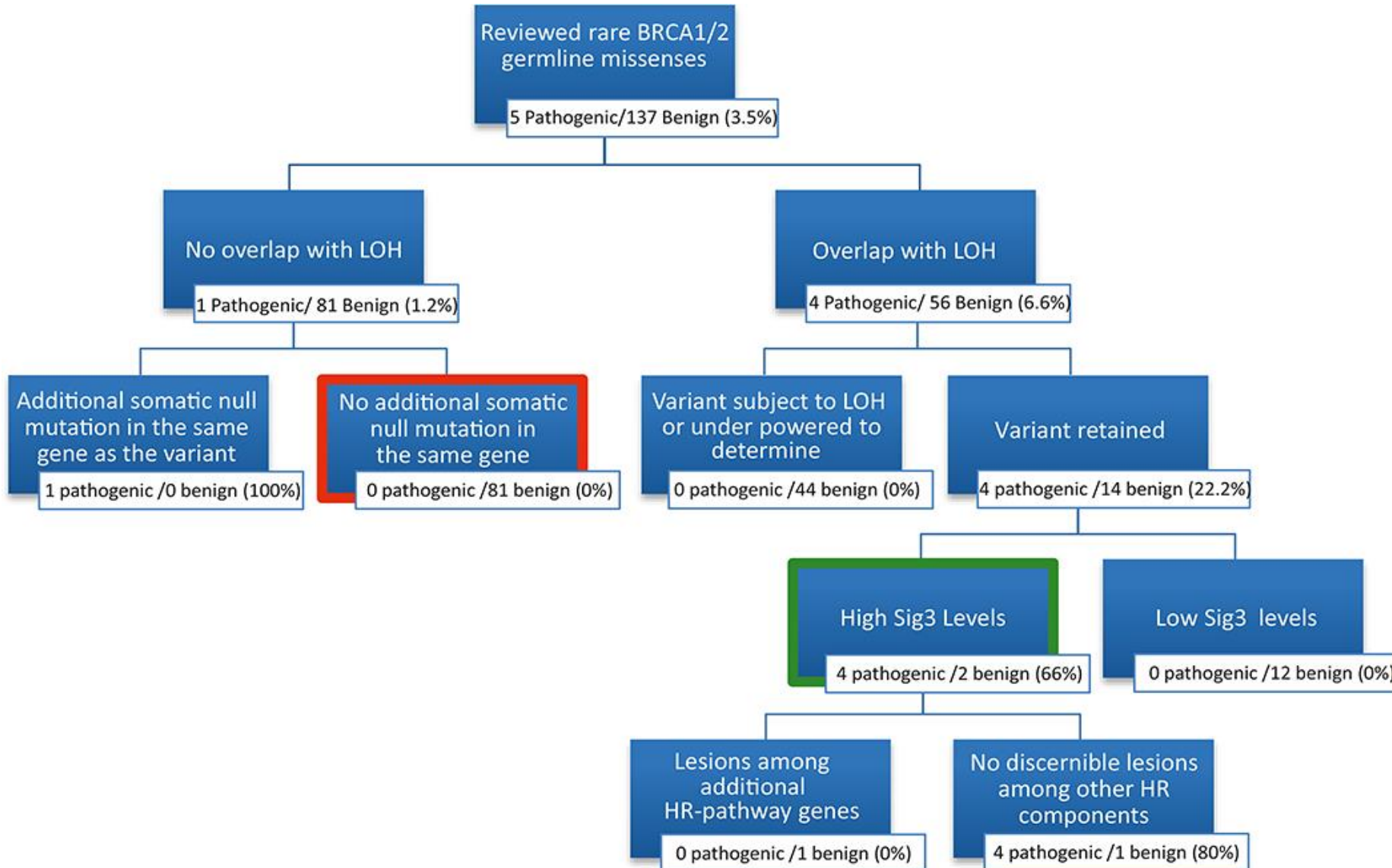
Defective DNA mismatch repair

- Microsatellite instability (MSI) due to deficiencies in mismatch repair (dMMR) is a common in colorectal cancer
- Patients with MSI are more likely to respond to checkpoint immunotherapy
- Mutational signatures associated with MSI (SBS15, SBS36, etc.) may be able to identify colorectal patients who should be treated with immunotherapy

Mutational signatures as predictive biomarkers for targeted therapy

A Mutational signatures useful in analysis						B Underlying mutational process	C Relevant genes	D Predisposition syndrome	E Proposed therapy choice
CS-3	CS-8	MH-Indels	RS-3	RS-5	HRD Index	Homologous Recombination Repair Deficiency	<i>BRCA1, BRCA2, RAD51C, PALB2</i>	Hereditary Breast and Ovarian Cancer Syndrome	PARP inhibition ³²⁻³⁴ , Platinum-based chemotherapy ³⁵⁻³⁷
CS-6	CS-15	CS-20	CS-26	STR-Indels		Mismatch Repair Deficiency	<i>MLH1, MSH2, MSH6, PMS1, PMS2</i>	Lynch, CMMRD, BMMR-D, HNPCC	PD1-immunotherapy ^{38-49,52}
CS-5	CS-8	TSB-sign				Nucleotide Excision Repair Deficiency	<i>ERCC1, ERCC2, XPC</i>	Xeroderma Pigmentosum	Cisplatin ⁶³⁻⁶⁵
CS-18	CS-30	TSB-sign	C>A*	G>T*	C>T*	Base excision Repair Deficiency	<i>MUTYH, OGG1</i> <i>NTHL1, SMUG1</i>	MAP NAP	
CS-10	STR-Indels					Deficient DNA polymerase proofreading activity	<i>POLE, POLD1</i>	PPAP	PD1-immunotherapy ^{48-49,52}
?						Non-Homologous End Joining Deficiency		Nijmegen Breakage Syndrome	
CS-2	CS-13	Kataegis				APOBEC Over-activity	<i>APOBEC1, APOBEC3A, APOBEC3B</i>		Tamoxifen Resistance ^{70,71}

Mutational signatures as a tool to classify VUS



- Genes such as *BRCA1/2* are associated with specific mutational signatures such as HRD-associated SBS3
- Rare variants in these genes are often classified as VUS
- The presence of HRD mutational signatures can be used as evidence for whether a VUS is benign or deleterious

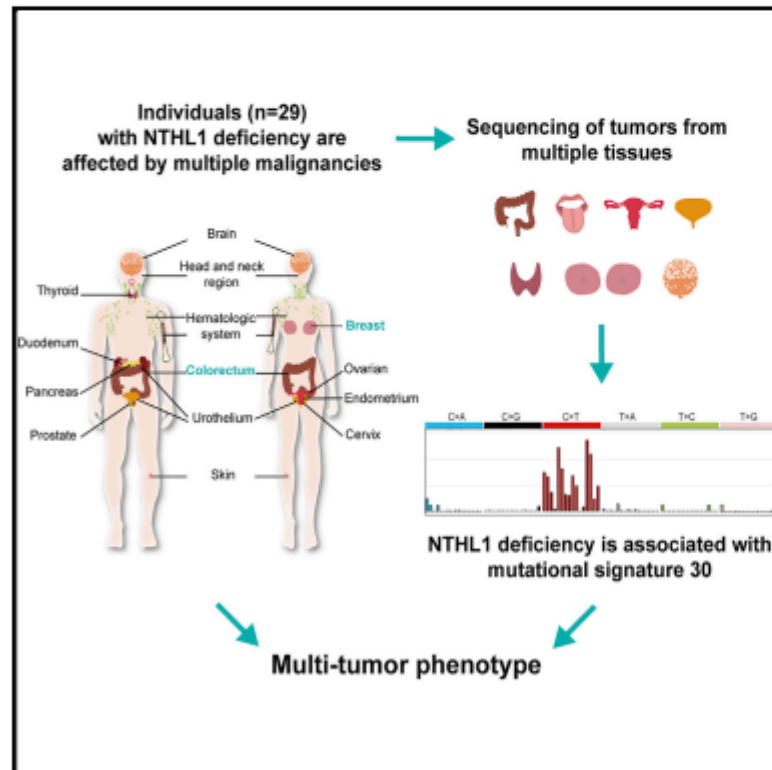
Mutational signatures as a tool to quantify cancer risk in individuals with *NTHL1* mutations

Article

Cancer Cell

Mutational Signature Analysis Reveals *NTHL1* Deficiency to Cause a Multi-tumor Phenotype

Graphical Abstract



Authors

Judith E. Grolleman,
Richarda M. de Voer,
Fadwa A. Elsayed, ..., Tom van Wezel,
Nicoline Hoogerbrugge,
Roland P. Kuiper

Correspondence

richarda.devoer@radboudumc.nl
(R.M.d.V.),
r.kuiper@
prinsesmaximacentrum.nl (R.P.K.)

In Brief

In addition to the known colorectal tumors, Grolleman et al. find tumors in 13 tissue types, including a high breast cancer incidence, among 29 carriers of biallelic germline *NTHL1* mutations and identify a mutation signature across tumor types, which may facilitate the identification and management of new cases.

Mutational signature analyses of tumours in people with a rare germline mutation in *NTHL1* reveals that they are at risk to develop multiple different types of tumours

Mutational signatures can identify individuals with pathogenic germline variants

nature genetics

Explore content ▾

About the journal ▾

Publish with us ▾

Subscribe

[nature](#) > [nature genetics](#) > [articles](#) > article

Article | Published: 21 August 2017

A mutational signature reveals alterations underlying deficient homologous recombination repair in breast cancer

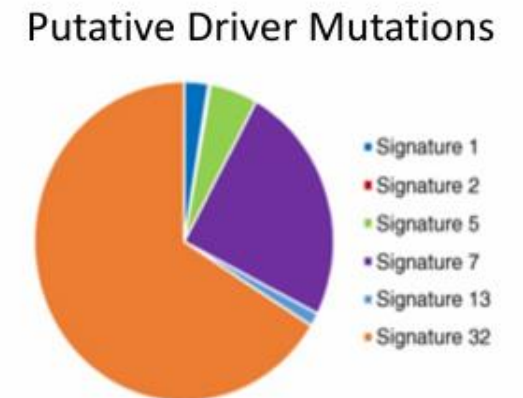
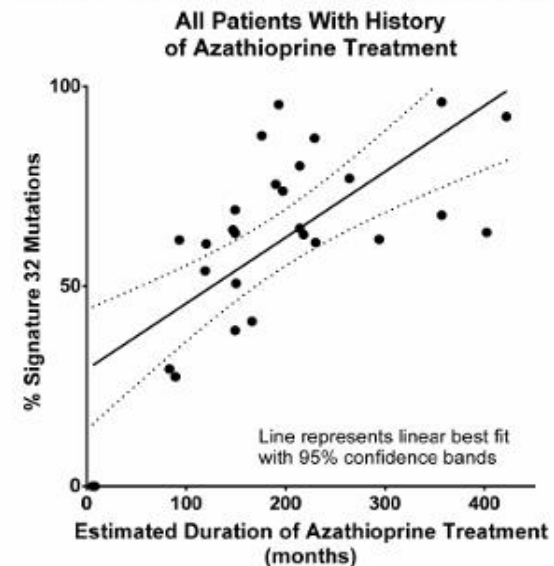
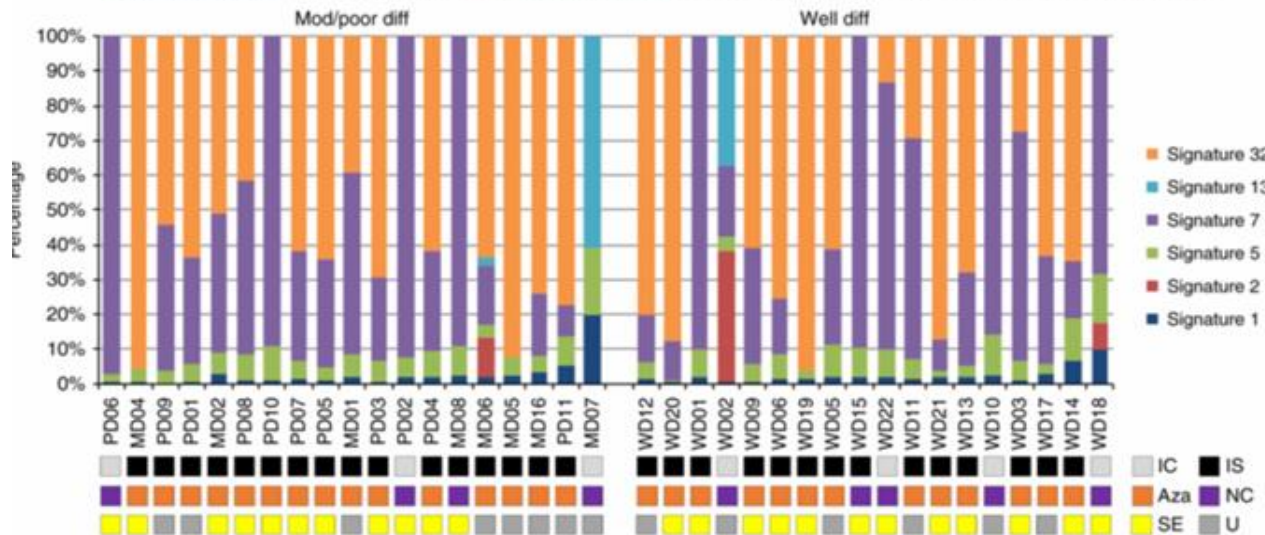
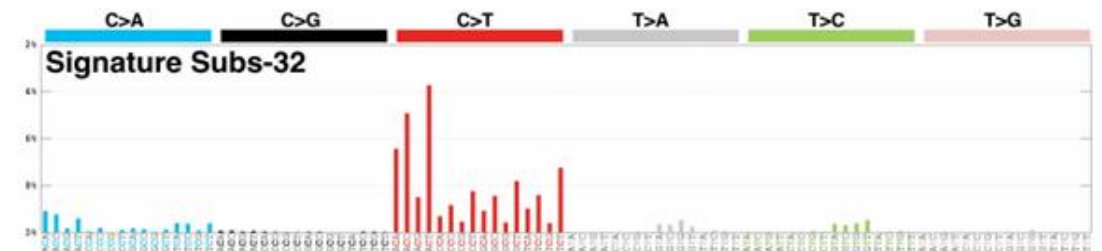
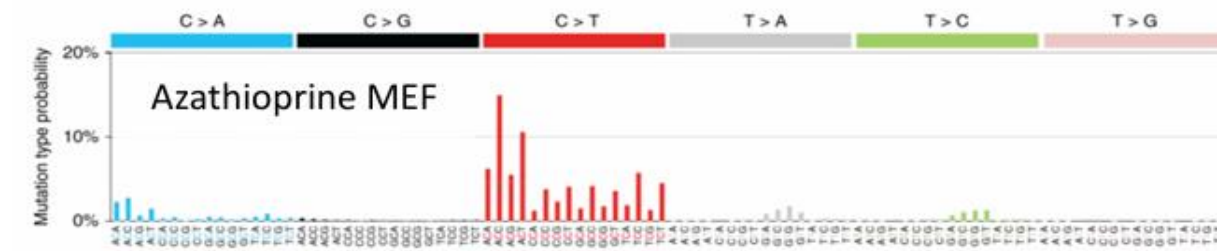
[Paz Polak](#), [Jaegil Kim](#), [Lior Z Braunstein](#), [Rosa Karlic](#), [Nicholas J Haradhavala](#), [Grace Tiao](#), [Daniel Rosebrock](#), [Dimitri Livitz](#), [Kirsten Kübler](#), [Kent W Mouw](#), [Atanas Kamburov](#), [Yosef E Maruvka](#), [Ignaty Leshchiner](#), [Eric S Lander](#), [Todd R Golub](#), [Aviad Zick](#), [Alexandre Orthwein](#), [Michael S Lawrence](#), [Rajbir N Batra](#), [Carlos Caldas](#), [Daniel A Haber](#), [Peter W Laird](#), [Hui Shen](#), [Leif W Ellisen](#), ... [Gad Getz](#)  [+ Show authors](#)

[Nature Genetics](#) **49**, 1476–1486 (2017) | [Cite this article](#)

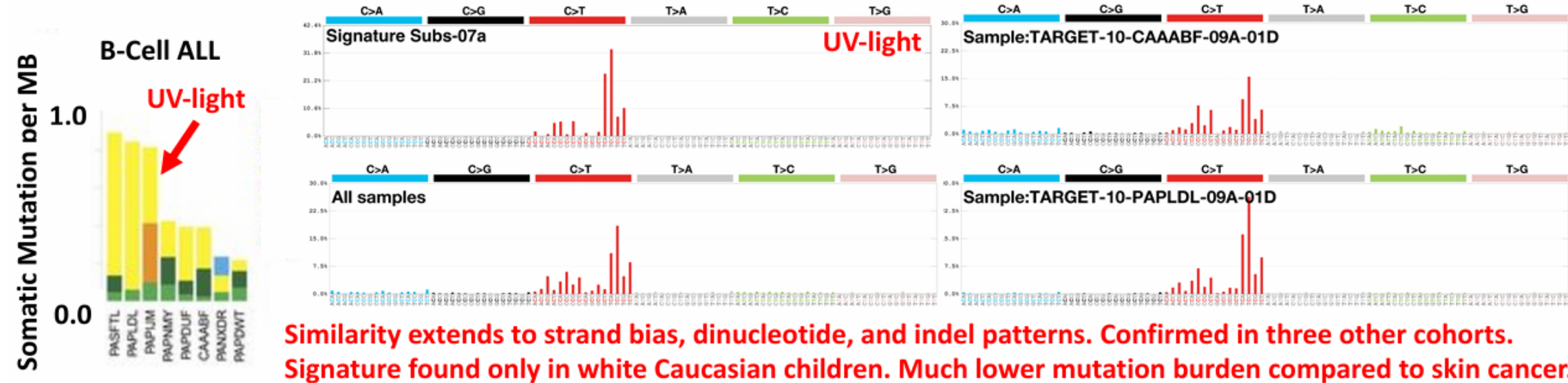
Quantification of HRD-associated mutational signatures in a patient's tumour can help doctors and patients decide if the patient needs genetic testing and/or genetic counselling.

Mutational signatures as a tool to discover/validate new carcinogens

Mutational signature analysis provides strong evidence that azathioprine, an immunosuppressive medication, is the carcinogen responsible for some cutaneous squamous cell carcinomas



Mutational signatures as a tool to discover known carcinogens in unexpected settings



Similarity extends to strand bias, dinucleotide, and indel patterns. Confirmed in three other cohorts. Signature found only in white Caucasian children. Much lower mutation burden compared to skin cancer.



[Cancer Causes & Control](#)
 October 2017, Volume 28, [Issue 10](#), pp 1075–1083 | [Cite as](#)

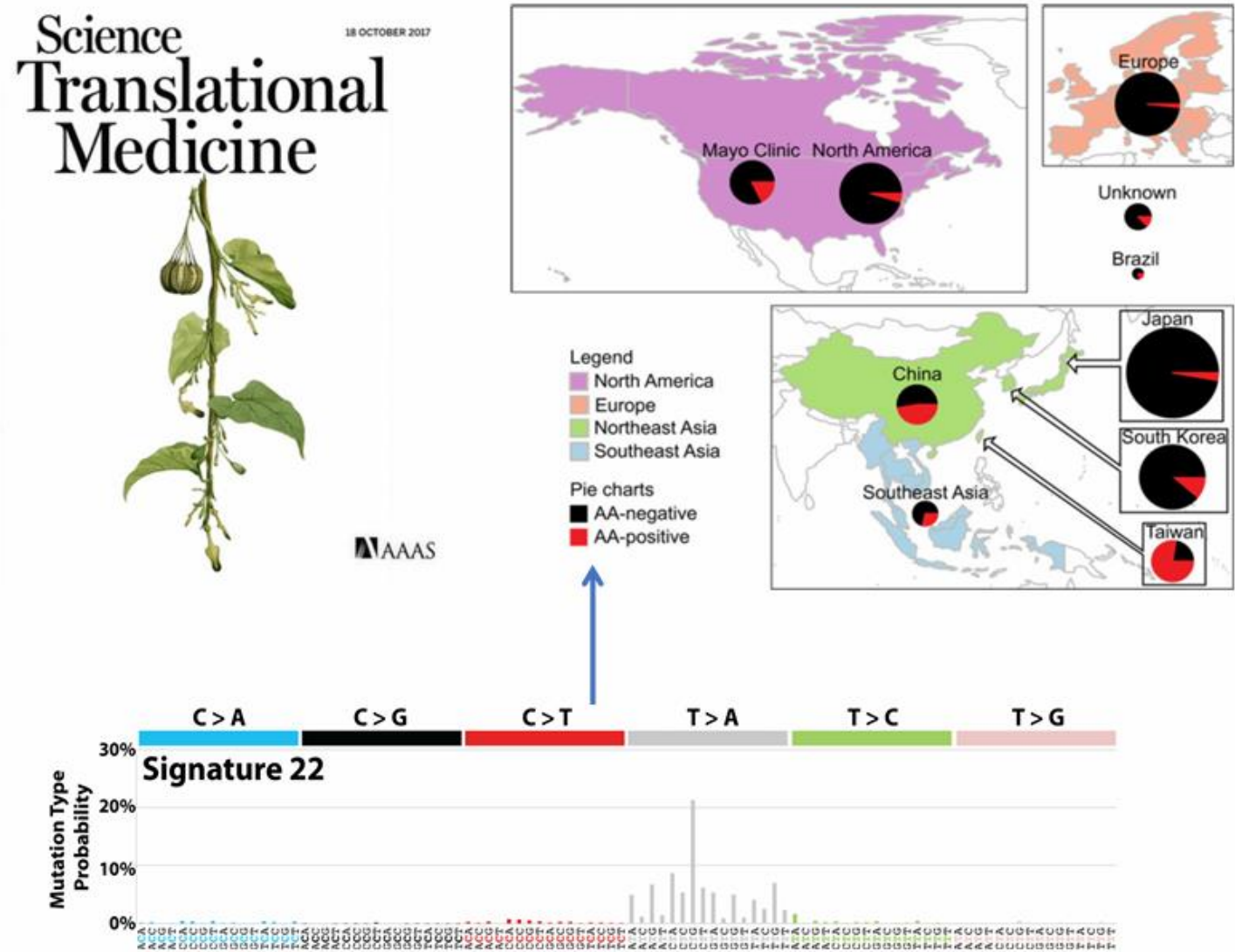
Residential exposure to ultraviolet light and risk of precursor B-cell acute lymphoblastic leukemia: assessing the role of individual risk factors, the ESCALE and ESTELLE studies

Authors [Authors and affiliations](#)

Astrid Coste , Denis Hémon, Laurent Orsi, Mathieu Boniol, Jean-François Doré, Laure Faure, Jacqueline Clavel, Stéphanie Goujon

Mutational signature analysis revealed the surprising association between UV light and some childhood leukemias (B-cell ALL), identifying UV light exposure as a new potential risk factor for B-cell ALL.

Mutational signatures as a tool to quantify the relative importance of different carcinogens in specific cancer types



Quantification of the aristolochic acid mutational signature 22 in Asian liver tumours provide strong evidence that some traditional medicines are strong carcinogens and important drivers of liver cancer in East Asia

Clinical and research applications

Discuss: How else might you use mutational signatures in your research?

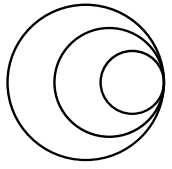
Clinical and research applications

Discuss: How else might you use mutational signatures in your research?

- Track carcinogenic exposure over time
- Compare carcinogenic exposure between different populations
- Orthogonal validation of self-reported data

Summary and takeaways

- The COSMIC database compiles validated mutational signatures across several classes of mutations
- Mutational signatures in COSMIC are linked with specific aetiologies through association studies and experimental validation, but many signatures still have unknown aetiologies
- Mutational signatures do not yet have widespread clinical utility, but they are valuable research tools in cancer biology and cancer epidemiology



**wellcome
connecting
science**

thanks!

Please contact wellcomeconnectingscience.org
for more information.

