gecp
lung cancer
research

16th CONGRESS Lung ON CANCER

BARCELONA
27 / 28
NOVEMBER 2025

Analyzing mutagenic events in never smokers lung cancer

Marcos Díaz Gay
Centro Nacional de Investigaciones Oncológicas (CNIO)

CONFLICTO DE INTERESES

No tengo conflictos de intereses que declarar.

En cumplimiento con el Código de Buenas Prácticas de la Industria Farmacéutica (Farmaindustria), se informa que toda la información compartida durante esta reunión científico-profesional es estrictamente confidencial, privilegiada y destinada únicamente al destinatario previsto. Queda expresamente prohibida la difusión, directa o indirecta, a través de redes sociales, canales de comunicación o medios externos, así como cualquier uso no autorizado, incluida la divulgación o distribución del contenido.

La información presentada no debe ser utilizada con fines promocionales, ni constituye asesoramiento médico o actividades promocionales. Además, contiene propuestas preliminares, planes, estrategias y opiniones que no representan posiciones finales ni garantías de desempeño futuro. En caso excepcional de que desee compartir algún contenido, deberá contar con la autorización previa, expresa y por escrito de **GECP.**

16th
CONGRESS
Lung
ON
CANCER

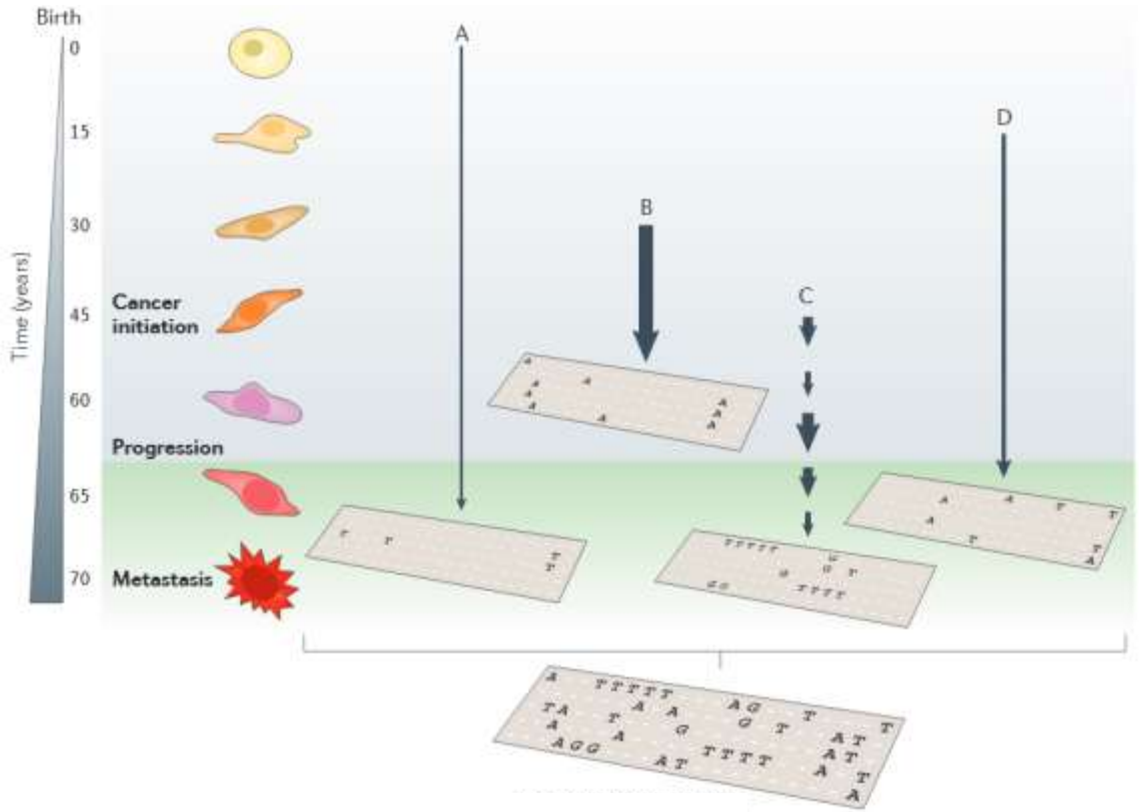
BARCELONA
27 / 28
NOVEMBER 2025

Mutational signatures framework

GeCP
lung cancer
research

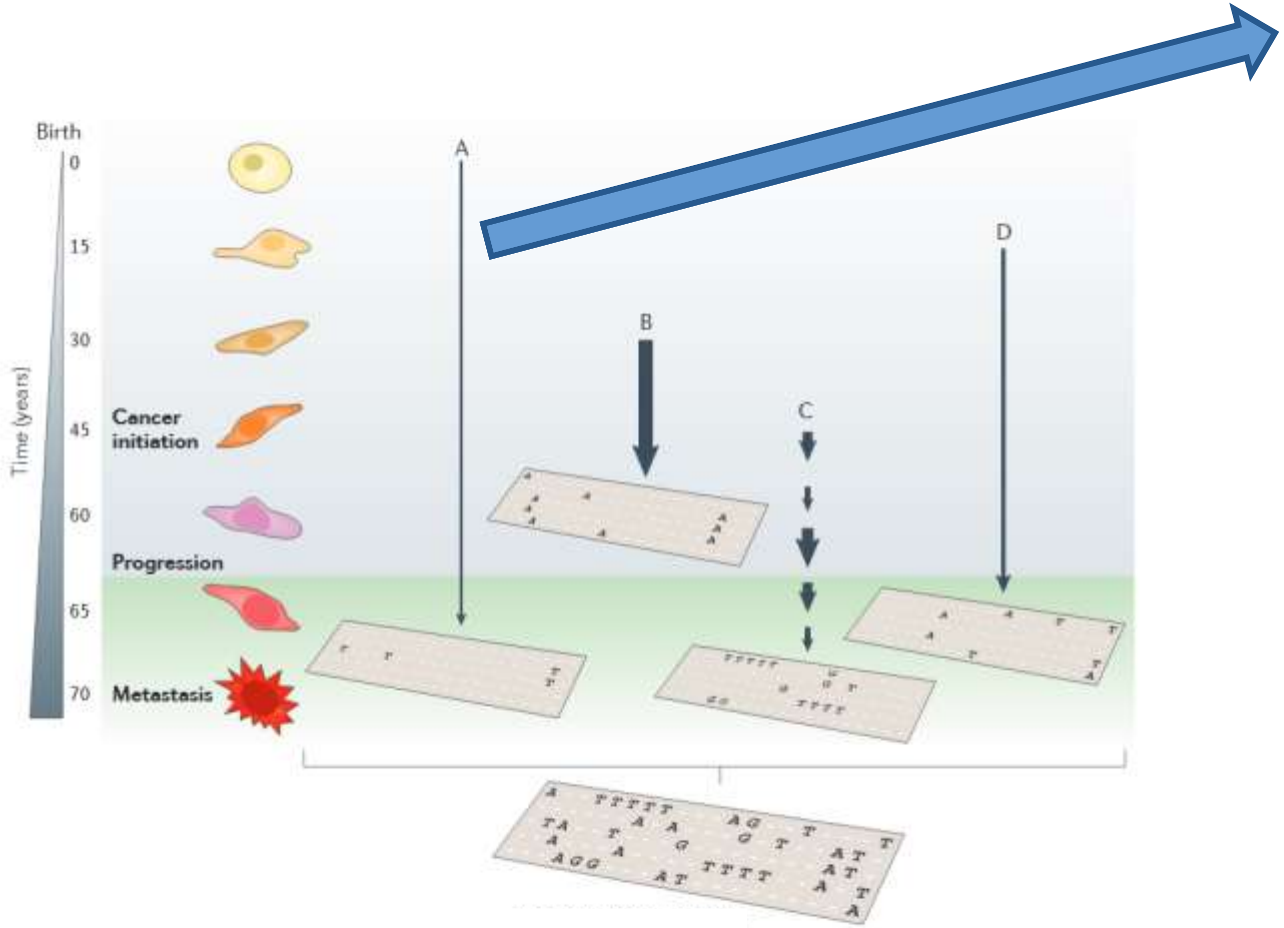
The mutational record of a cancer patient is a mix of different mutational processes characterized by specific signatures

The mutational record of a cancer patient is a mix of different mutational processes characterized by specific signatures



Helleday *et al.* 2014 *Nat Rev Genet*

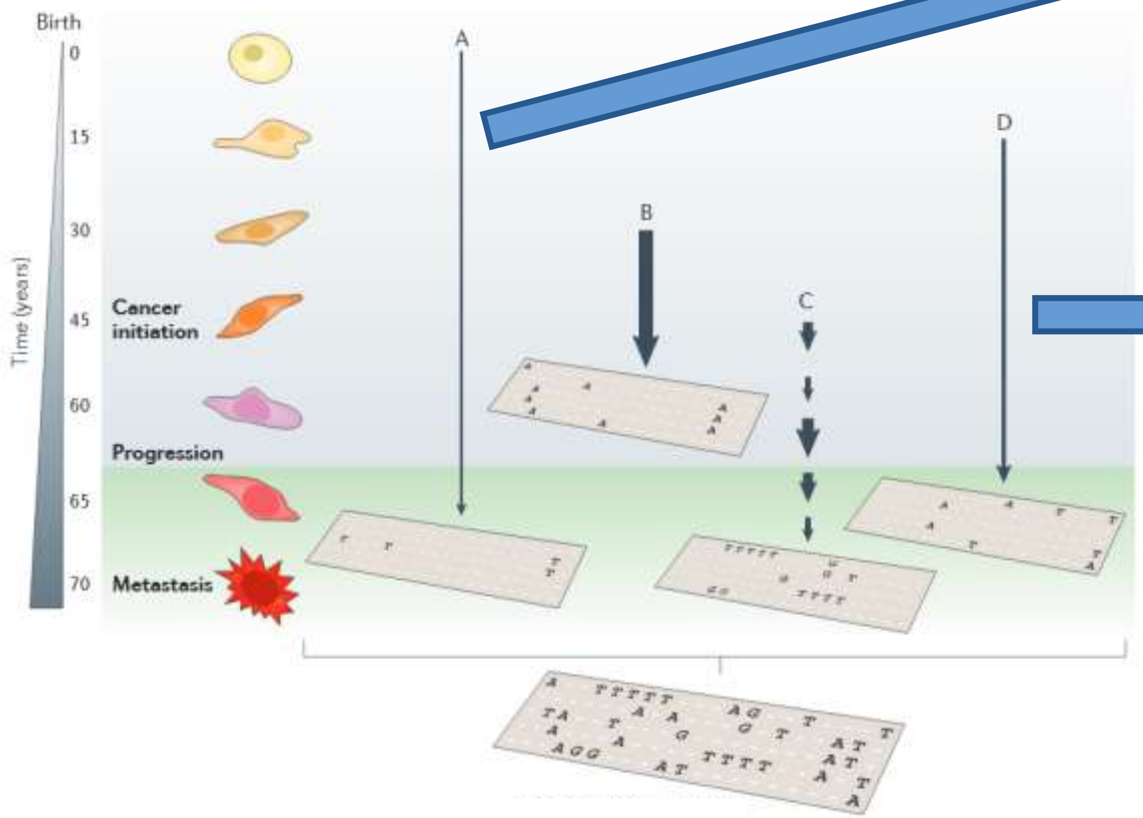
The mutational record of a cancer patient is a mix of different mutational processes characterized by specific signatures



Helleday *et al.* 2014 *Nat Rev Genet*

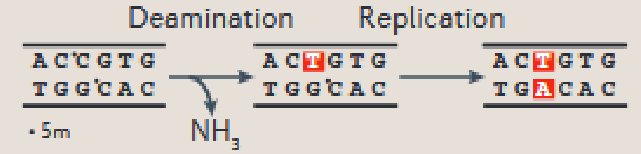
Normal cellular activities
Spontaneous deamination
of methylated cytosines

The mutational record of a cancer patient is a mix of different mutational processes characterized by specific signatures

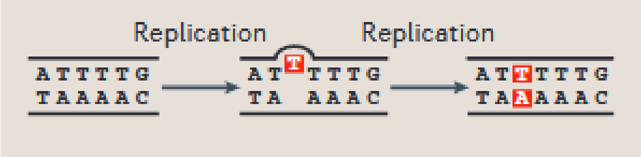


Helleday *et al.* 2014 *Nat Rev Genet*

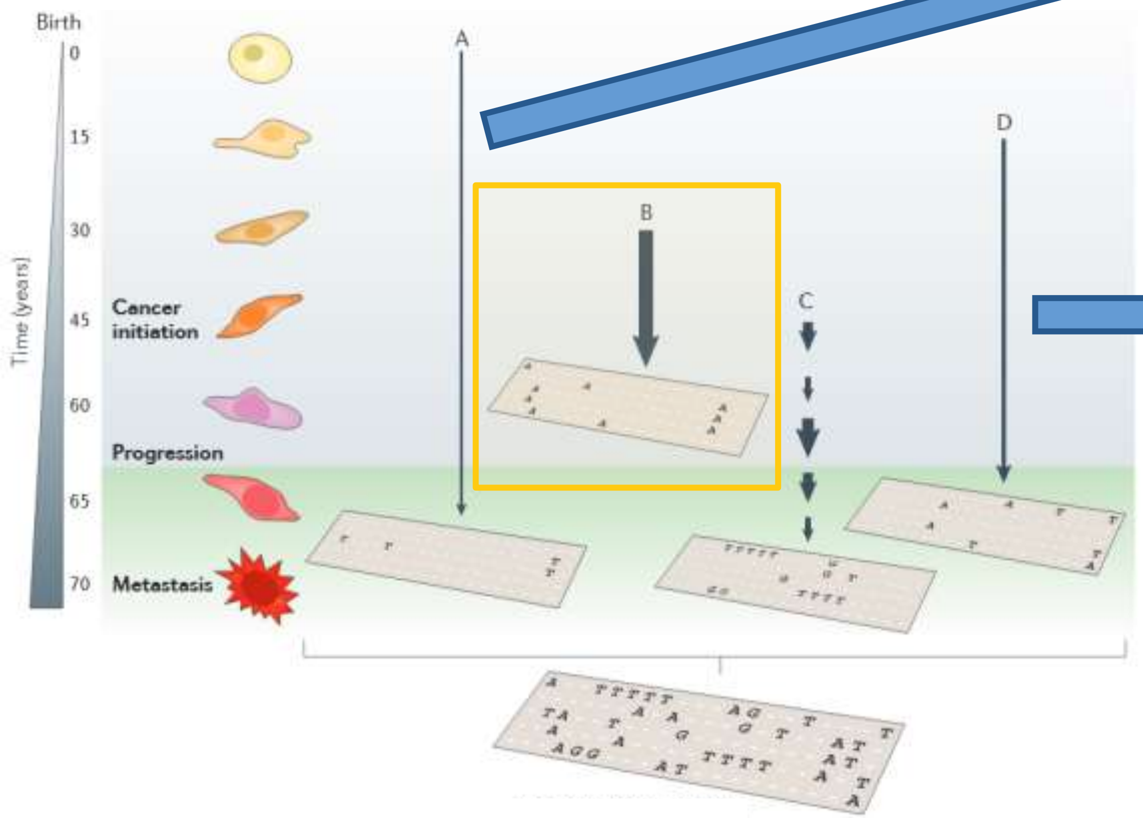
Normal cellular activities
Spontaneous deamination
of methylated cytosines



Failure in DNA repair
Aberrant mismatch repair pathway



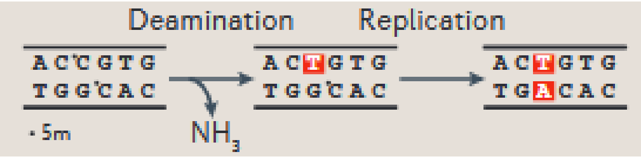
The mutational record of a cancer patient is a mix of different mutational processes characterized by specific signatures



Helleday *et al.* 2014 *Nat Rev Genet*

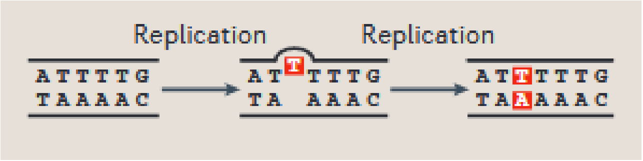
Normal cellular activities

Spontaneous deamination of methylated cytosines



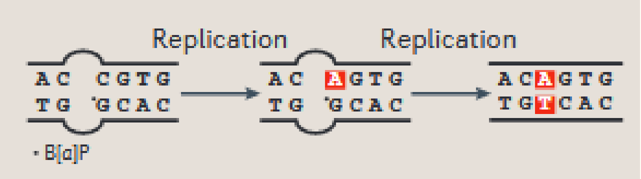
Failure in DNA repair

Aberrant mismatch repair pathway



Environmental exposures

Tobacco smoking



Mutational signatures are the genomic footprints of mutational processes

Mutational signatures are the genomic footprints of mutational processes

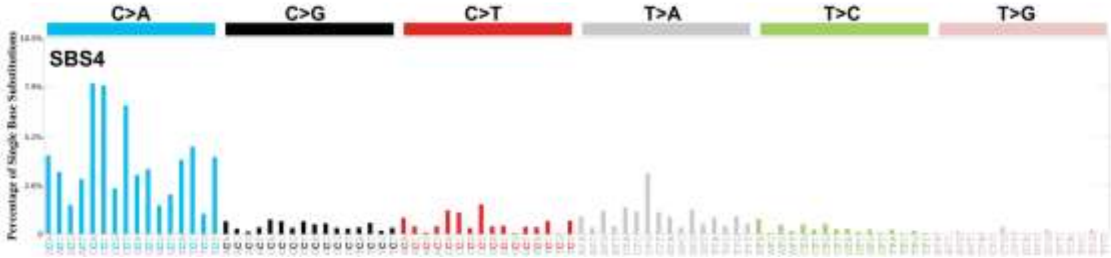
Environmental exposures

Tobacco smoking

Mutational signatures are the genomic footprints of mutational processes

Environmental exposures

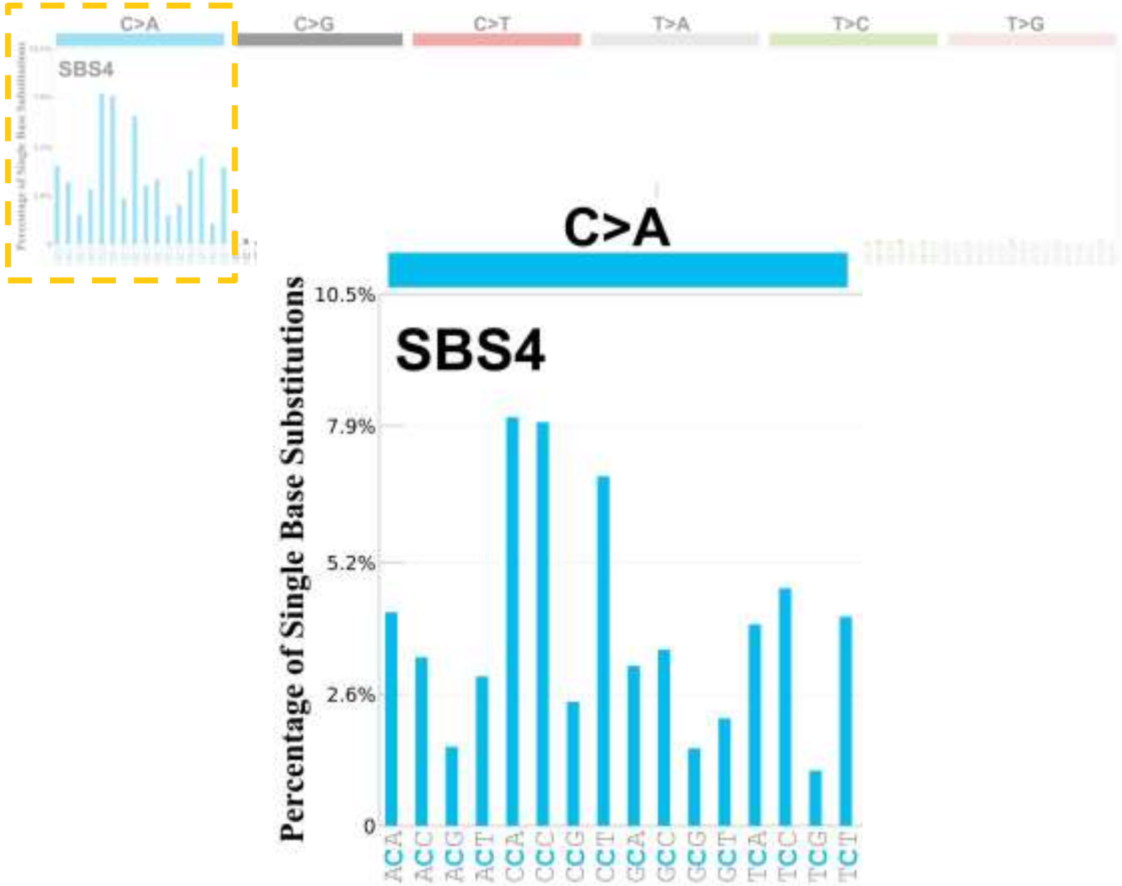
Tobacco smoking



Mutational signatures are the genomic footprints of mutational processes

Environmental exposures

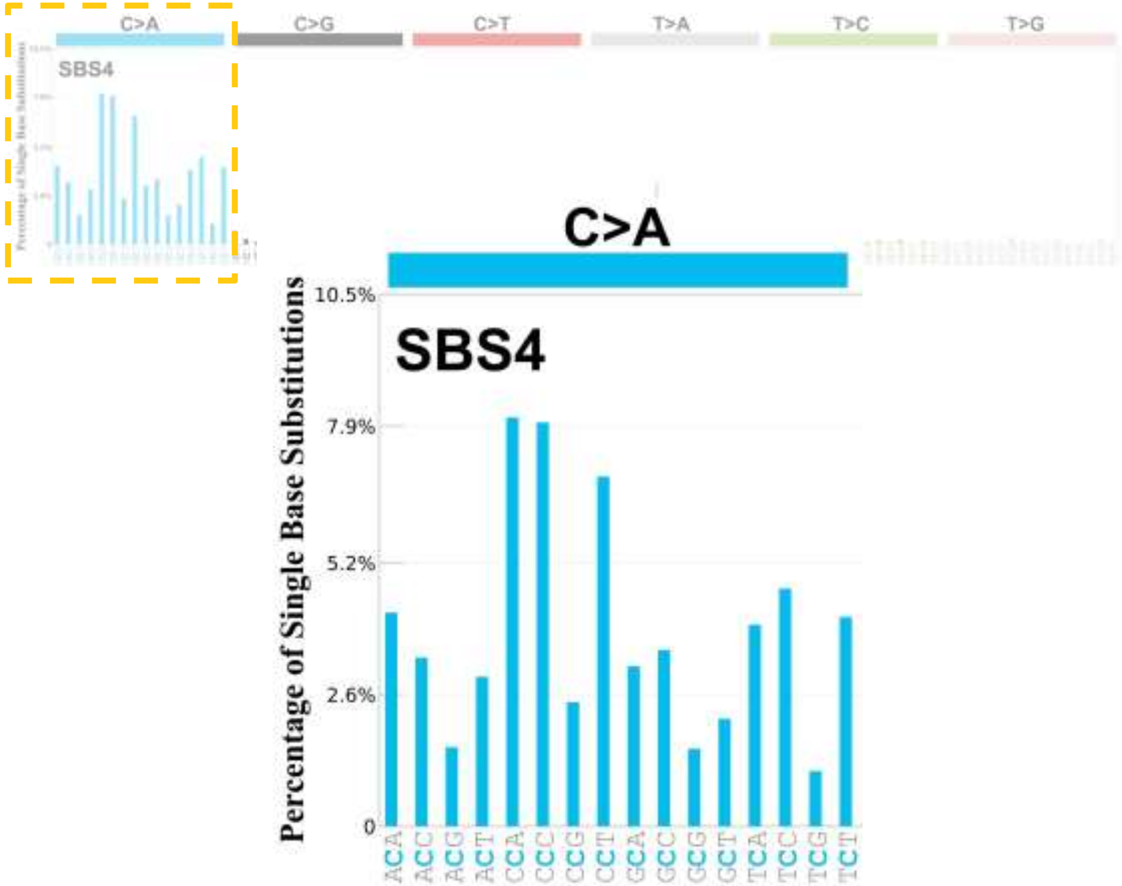
Tobacco smoking



Mutational signatures are the genomic footprints of mutational processes

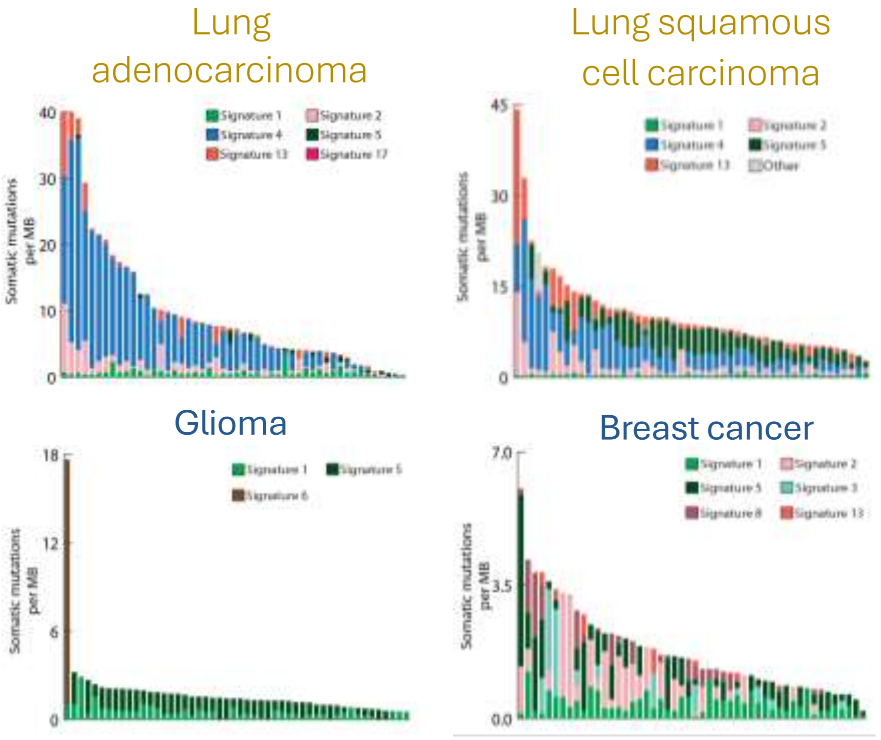
GeCP
lung cancer research

Environmental exposures
Tobacco smoking



Tumor types
associated with smoking

Tumor types
not associated with smoking


Lung adenocarcinoma

Lung squamous cell carcinoma

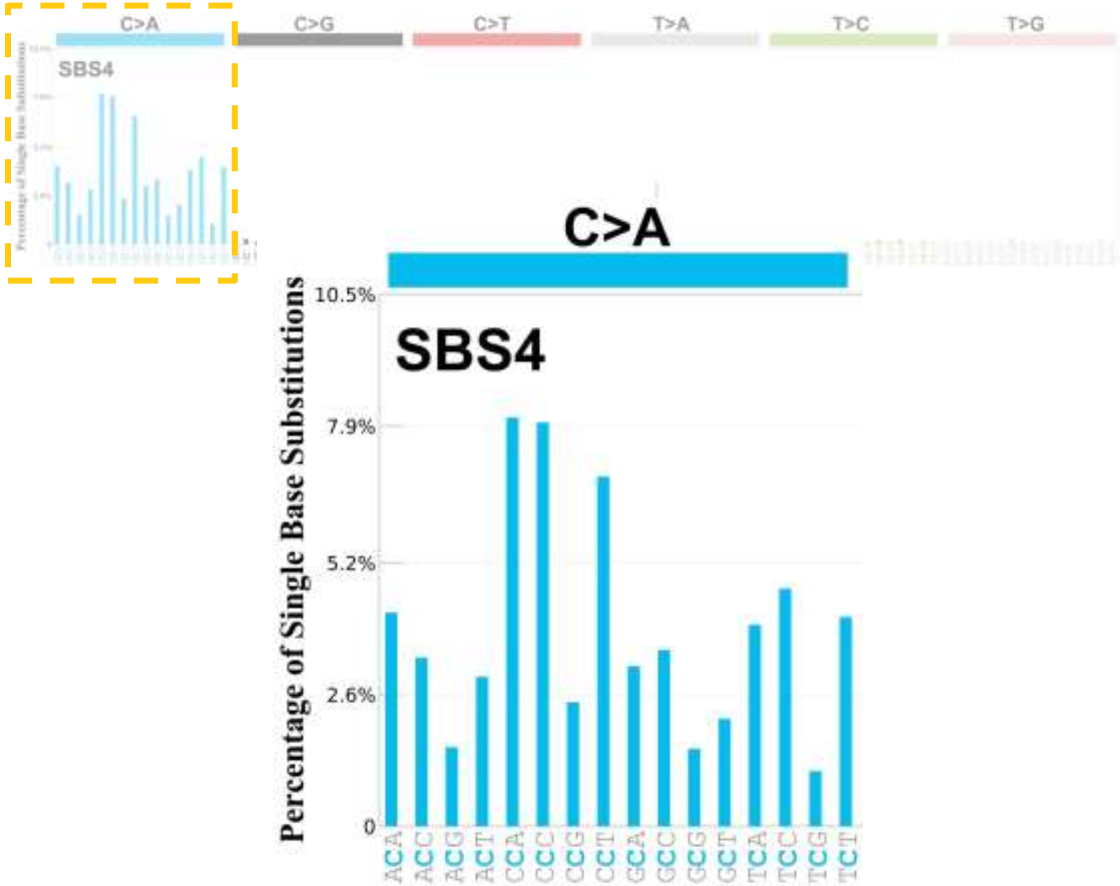
Glioma

Breast cancer

Alexandrov *et al*. 2013 *Nature*

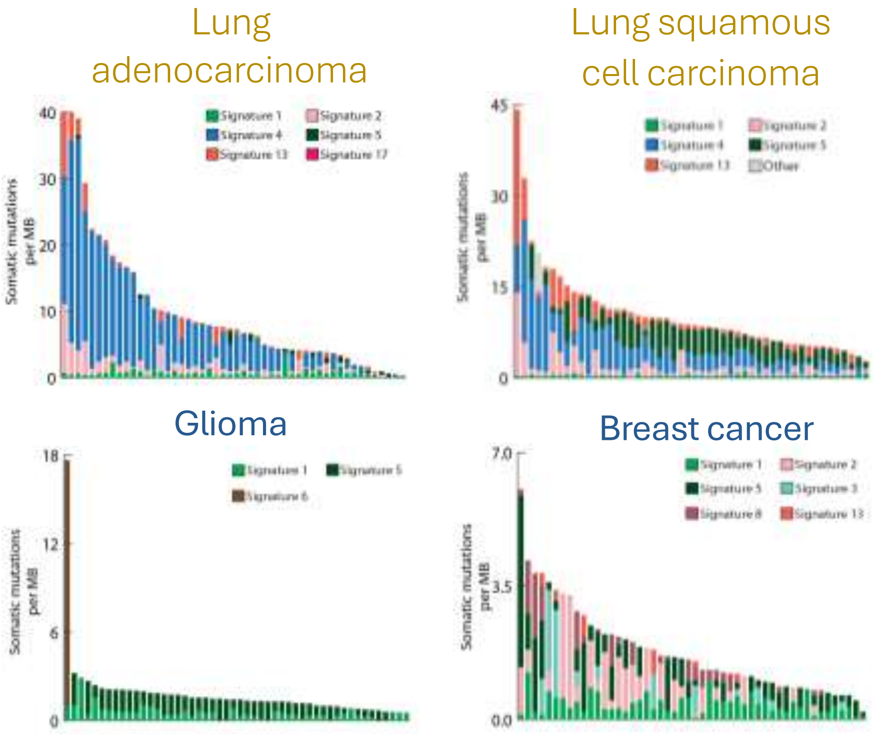
Mutational signatures are the genomic footprints of mutational processes

Environmental exposures
Tobacco smoking



Tumor types **associated with smoking**

Tumor types **<u>not</u> associated with smoking**


Lung adenocarcinoma

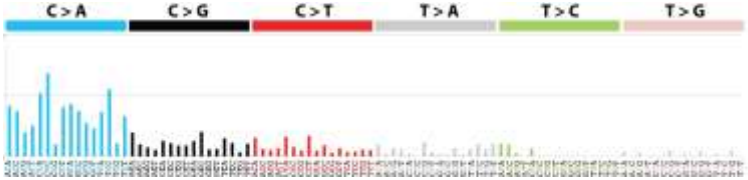
Lung squamous cell carcinoma

Glioma

Breast cancer

Alexandrov *et al.* 2013 *Nature*

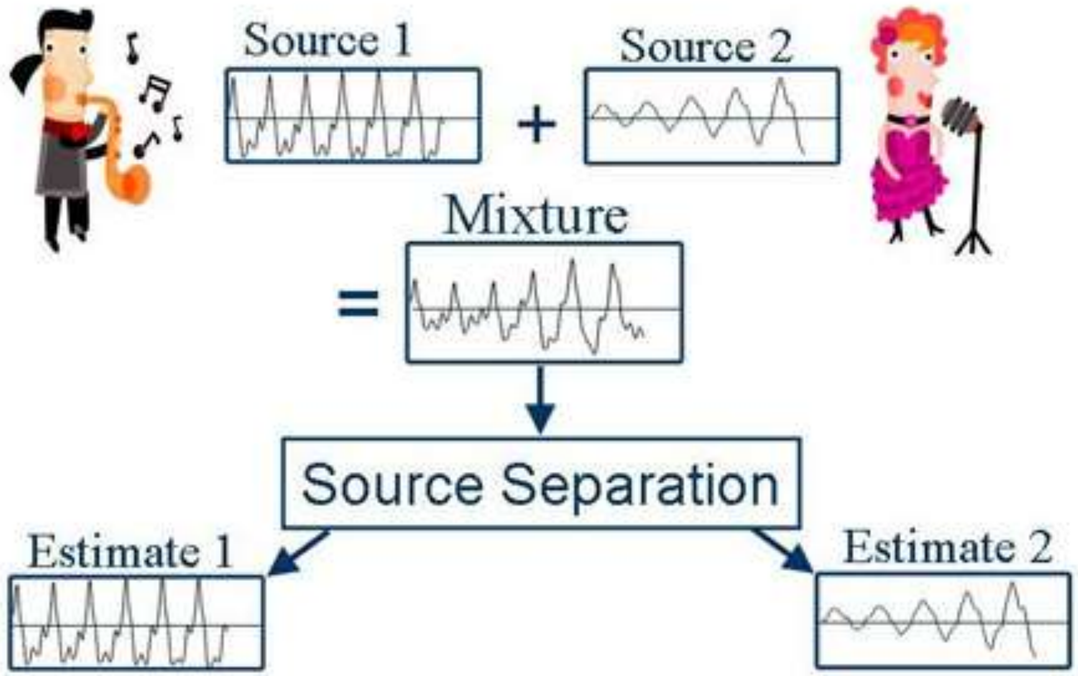
Experimental validation
Mutational signature of benzo[a]pyrene exposure



Nik-Zainal *et al.* 2015 *Mutagenesis*

IA machine learning algorithms allow the identification of mutational signatures

IA machine learning algorithms allow the identification of mutational signatures



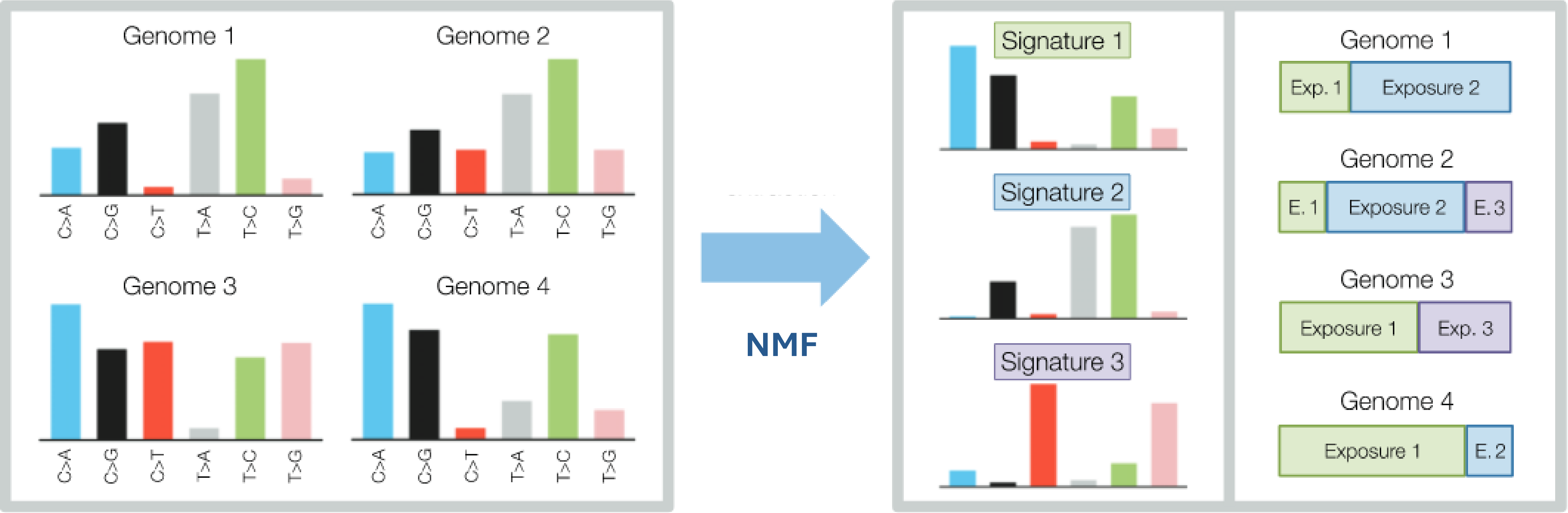
Interactive Audio Lab
Northwestern University

Non-negative matrix factorization (NMF)

IA machine learning algorithms allow the identification of mutational signatures

Tumors sequenced by NGS

Mutational signatures and activities



Baez-Ortega *et al*. 2019 *Briefings in Bioinformatics*

In the last decade >150 reference mutational signatures have been extracted from tens of thousands of sequenced human tumors

GECP lung cancer research

COSMIC Mutational Signatures

TCGA

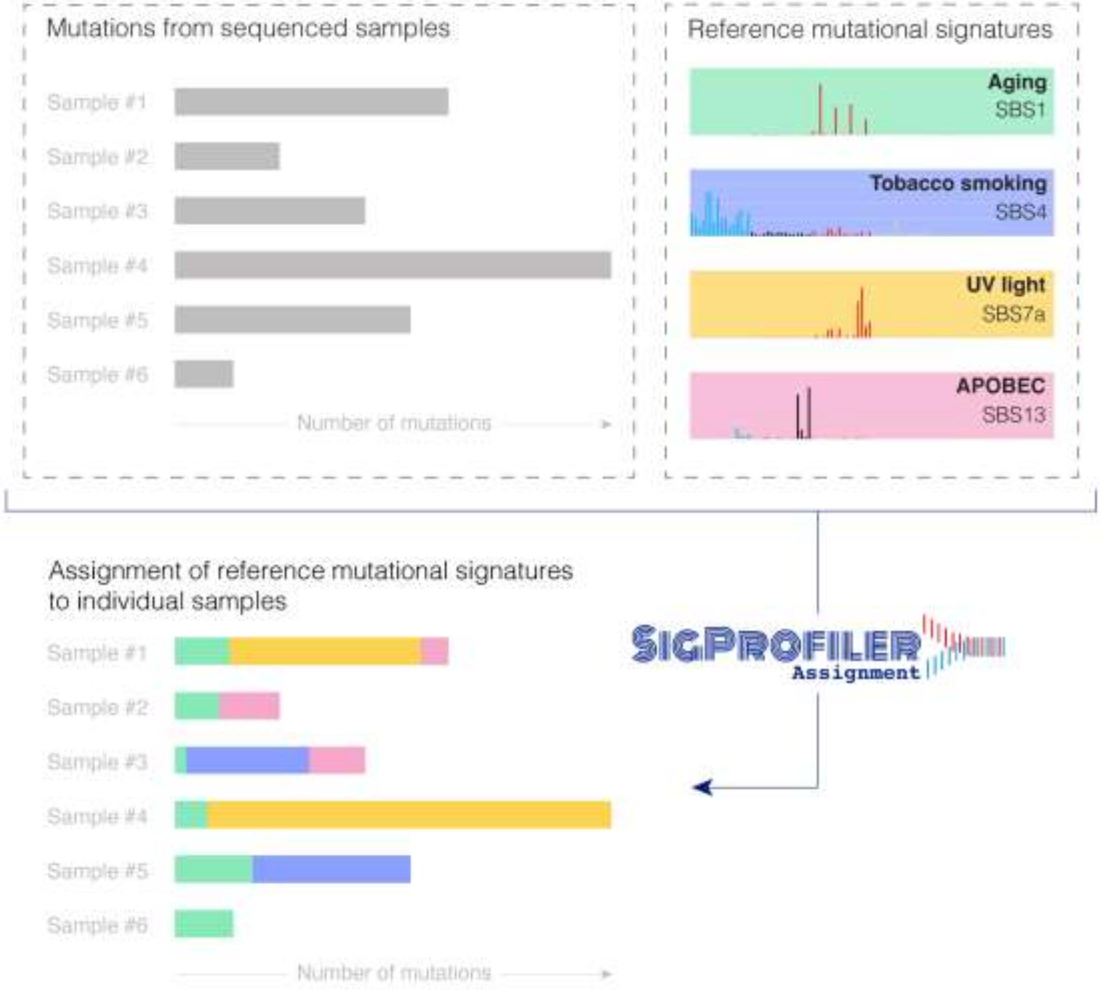
PCAWG PanCancer Analysis OF WHOLE GENOMES

Genomics england

Alexandrov *et al.* 2013 *Nature* | Alexandrov *et al.* 2020 *Nature* | Otlu, Díaz-Gay *et al.* 2023 *Cell Reports* | Everall *et al.* 2023 *medRxiv* | Sondka *et al.* 2024 *Nucleic Acids Research*

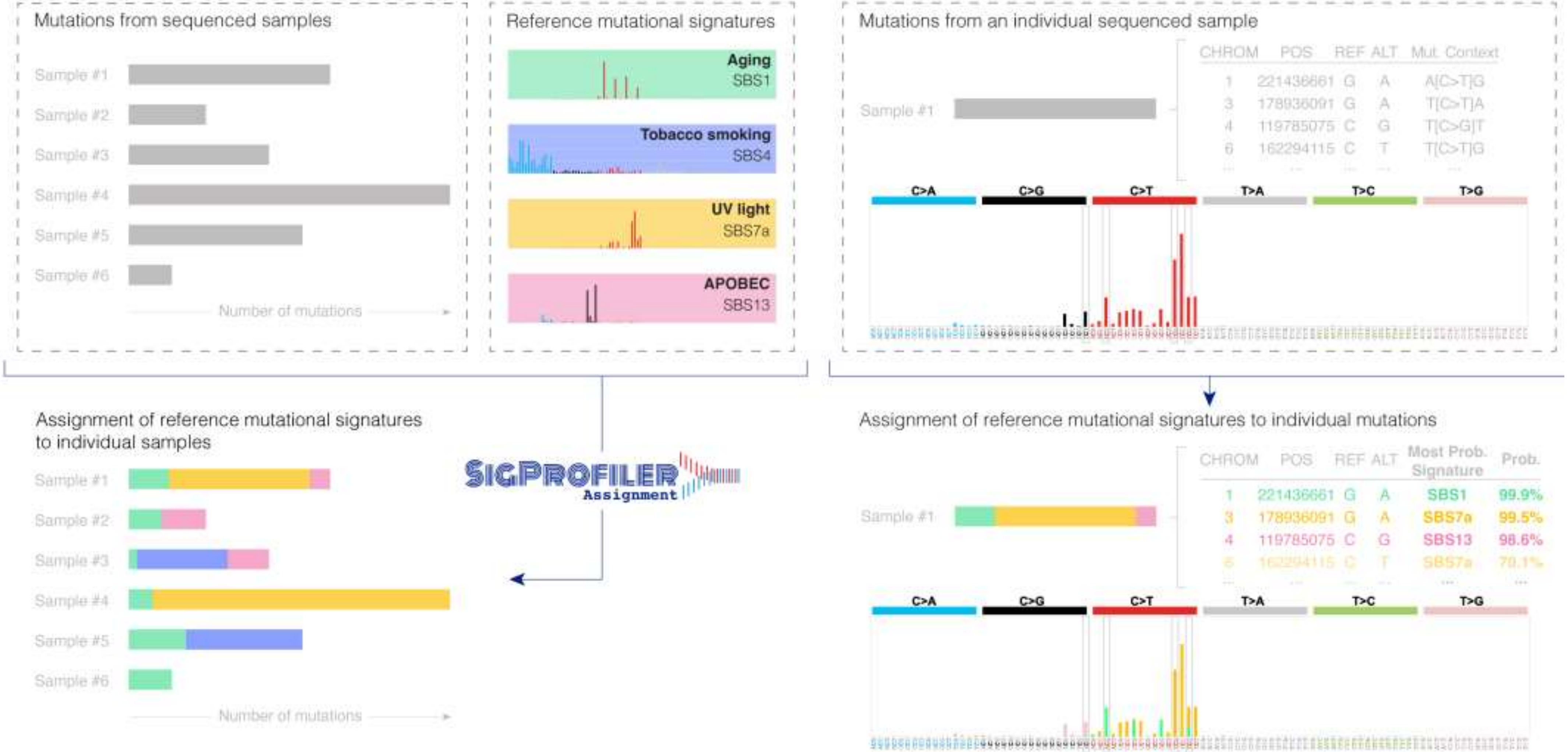
Reference mutational signatures can be assigned to individual tumors

Reference mutational signatures can be assigned to individual tumors



Díaz-Gay *et al.* 2023 *Bioinformatics*

Reference mutational signatures can be assigned to individual tumors and individual somatic mutations

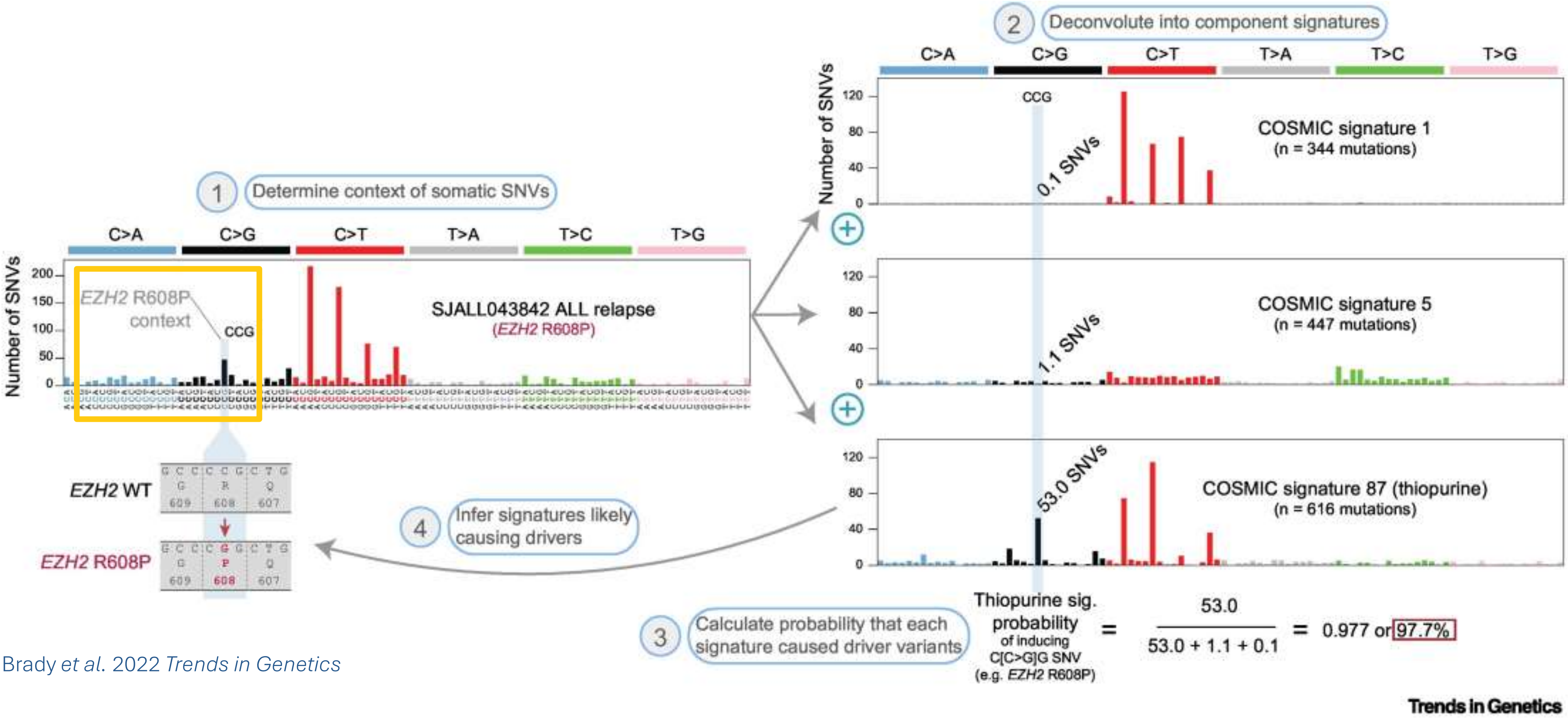


CHROM	POS	REF	ALT	Mut. Context
1	221436661	G	A	A[C>T]G
3	178936091	G	A	T[C>T]A
4	119785075	C	G	T[C>G]T
6	162294115	C	T	T[C>T]G
...	...	...	...	...

CHROM	POS	REF	ALT	Most Prob. Signature	Prob.
1	221436661	G	A	SBS1	99.9%
3	178936091	G	A	SBS7a	99.5%
4	119785075	C	G	SBS13	98.6%
6	162294115	C	T	SBS7a	70.1%
...	...	...	...	...	...

Díaz-Gay *et al.* 2023 *Bioinformatics*

Reference mutational signatures can be assigned to individual tumors and individual somatic mutations



Brady *et al.* 2022 *Trends in Genetics*

SigProfiler suite of bioinformatics tools for mutational signature analysis

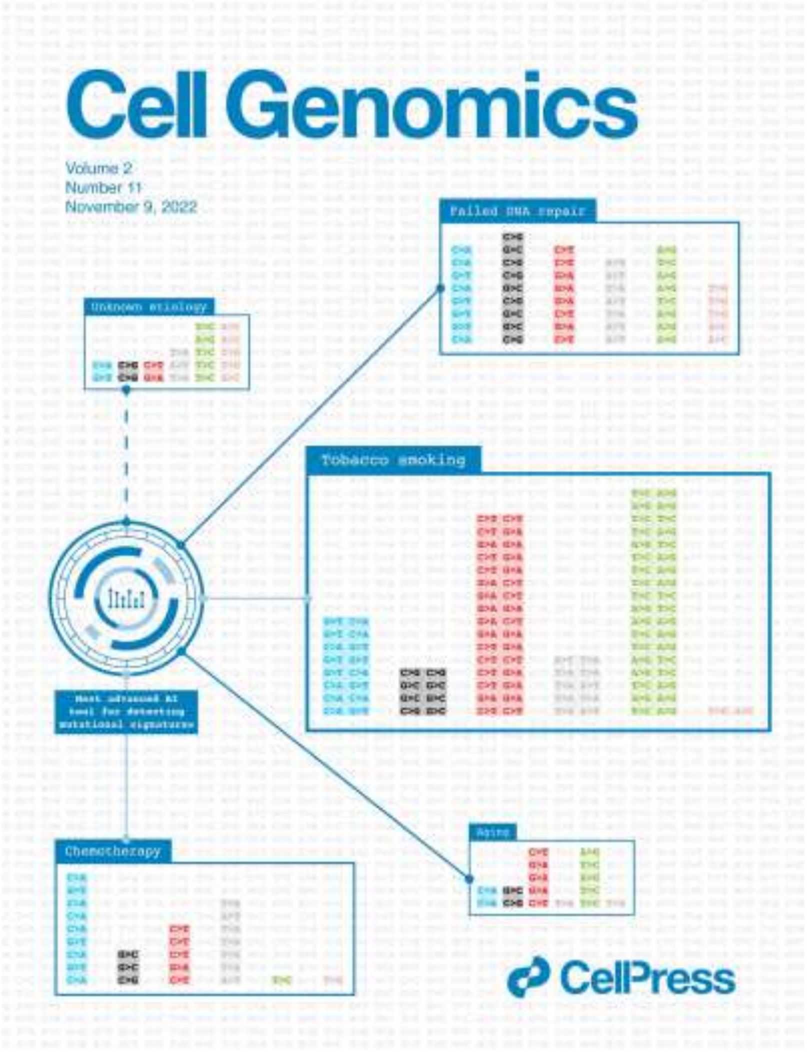
SigProfiler suite of bioinformatics tools for mutational signature analysis

Signature extraction

SigProfiler suite of bioinformatics tools for mutational signature analysis

SigProfiler Extractor

Signature extraction



Islam, Díaz-Gay *et al.* 2022 *Cell Genomics*

SigProfiler suite of bioinformatics tools for mutational signature analysis

SIGPROFILER
Extractor

Signature extraction

SigProfiler suite of bioinformatics tools for mutational signature analysis

SIGPROFILER
MatrixGenerator

SIGPROFILER
Plotting

Pre-processing

SIGPROFILER
Extractor

Signature extraction

SigProfiler suite of bioinformatics tools for mutational signature analysis

GeCP
lung cancer
research

SIGPROFILER
MatrixGenerator

SIGPROFILER
Plotting

Pre-processing

Bergstrom *et al.* 2019 *BMC Genomics*
Khandekar *et al.* 2023 *BMC Genomics*

SigProfiler suite of bioinformatics tools for mutational signature analysis

GECP lung cancer research

SIGPROFILER MatrixGenerator

SIGPROFILER Plotting

Pre-processing

SIGPROFILER Extractor

Signature extraction

SigProfiler suite of bioinformatics tools for mutational signature analysis

SIGPROFILER
MatrixGenerator

SIGPROFILER
Plotting

Pre-processing

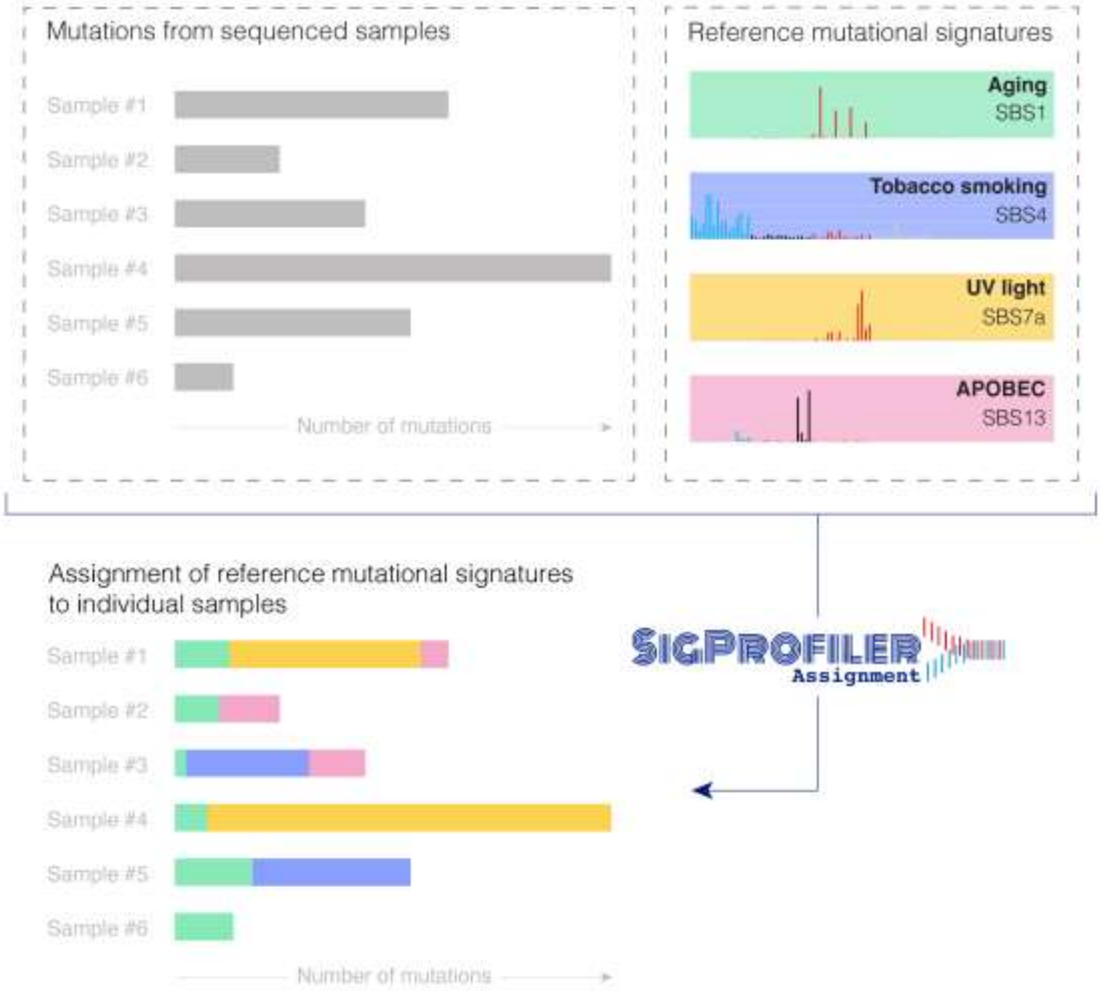
SIGPROFILER
Extractor

Signature extraction

SIGPROFILER
Assignment

Post-processing

SigProfiler suite of bioinformatics tools for mutational signature analysis



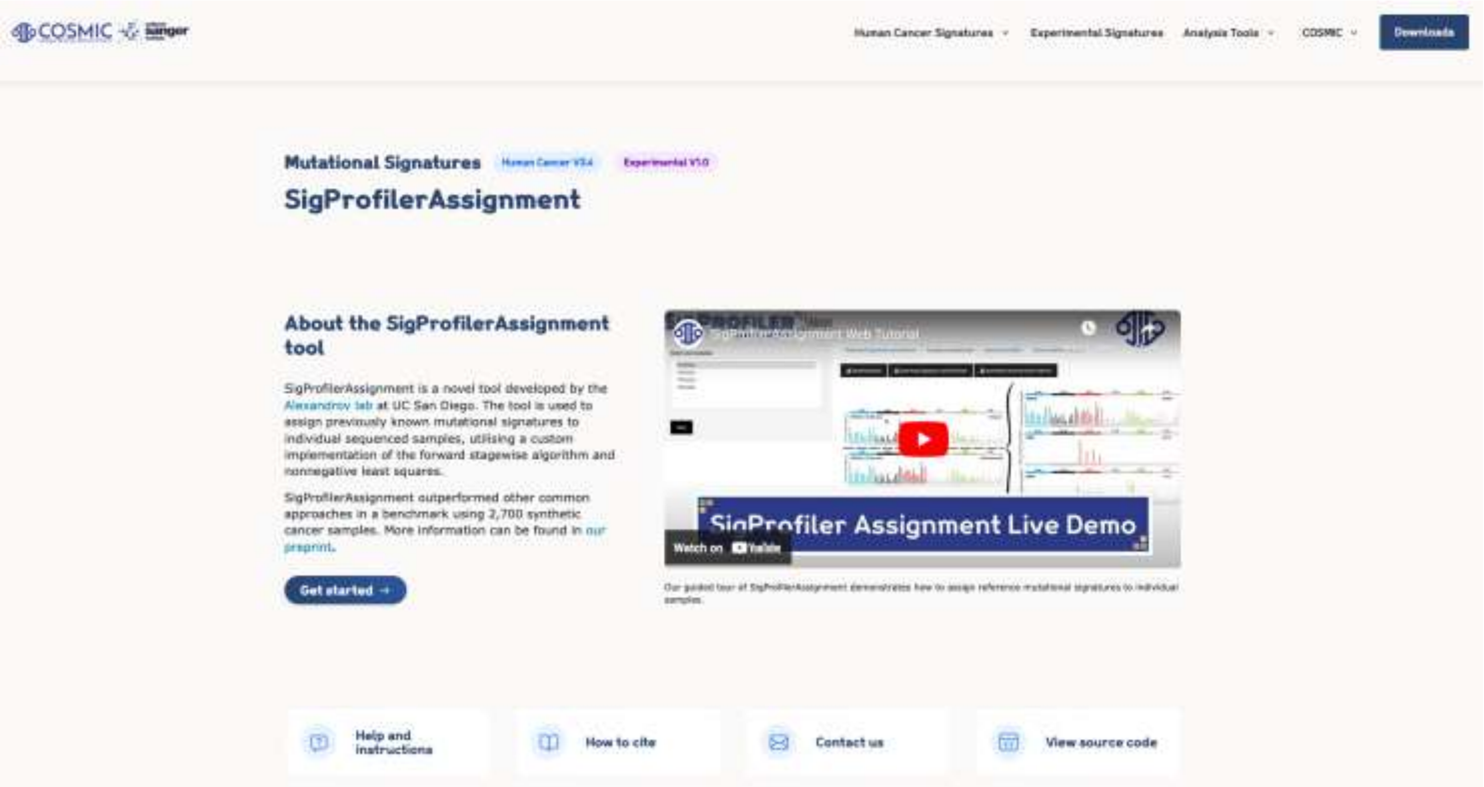
Díaz-Gay *et al.* 2023 *Bioinformatics*

SIGPROFILER Assignment

Post-processing

SigProfiler suite of bioinformatics tools for mutational signature analysis

We implemented an **online tool** for non-bioinformatic experts as part of **COSMIC Mutational Signatures**
cancer.sanger.ac.uk/signatures/assignment/



SigProfilerAssignment

About the SigProfilerAssignment tool

SigProfilerAssignment is a novel tool developed by the Alexandrov lab at UC San Diego. The tool is used to assign previously known mutational signatures to individual sequenced samples, utilising a custom implementation of the forward stagewise algorithm and nonnegative least squares.

SigProfilerAssignment outperformed other common approaches in a benchmark using 2,700 synthetic cancer samples. More information can be found in our preprint.

Our guided tour of SigProfilerAssignment demonstrates how to assign reference mutational signatures to individual samples.

Díaz-Gay *et al.* 2023 *Bioinformatics*

SIGPROFILER Assignment

Post-processing

16th CONGRESS Lung ON CANCER

BARCELONA
27 / 28
NOVEMBER 2025

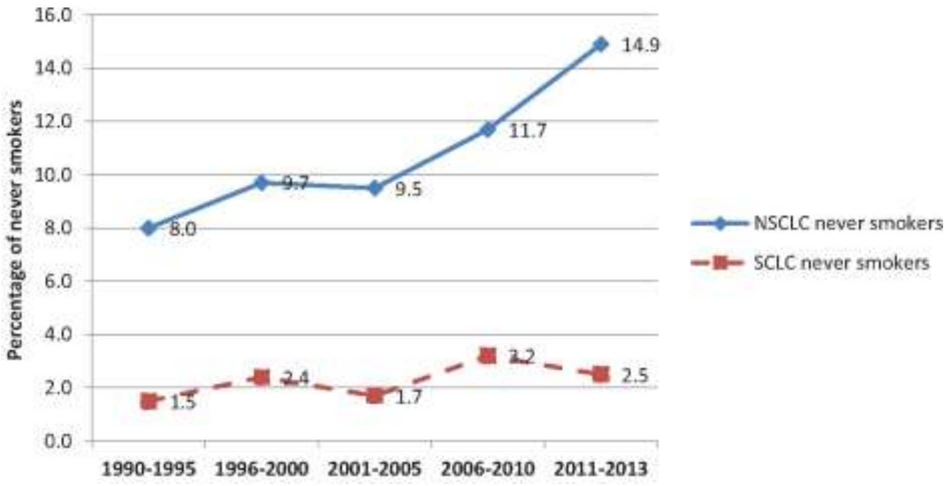
Global genomic characterization of lung cancer in never smokers

GECP
lung cancer research

Lung cancer in never smokers

Lung cancer in never smokers

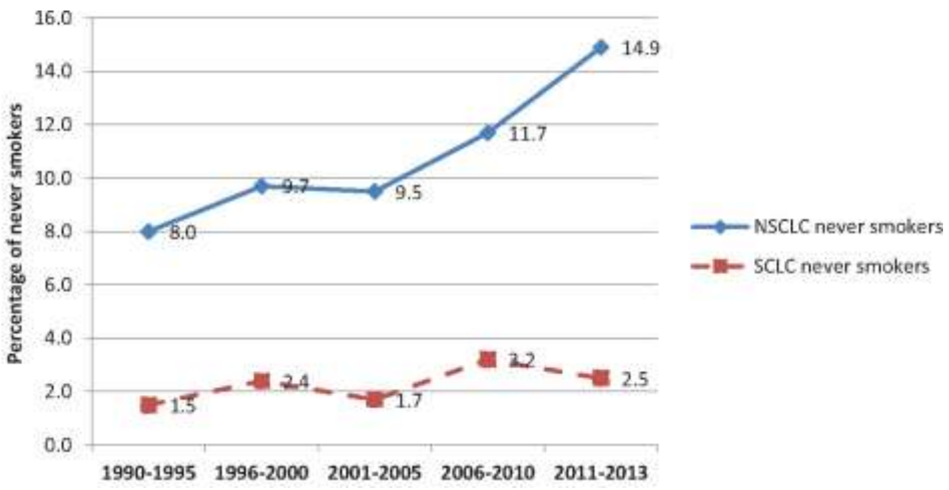
Increase in the proportion of NSCLC
who are never smokers



Pelosof *et al*. 2017 *JNCI*

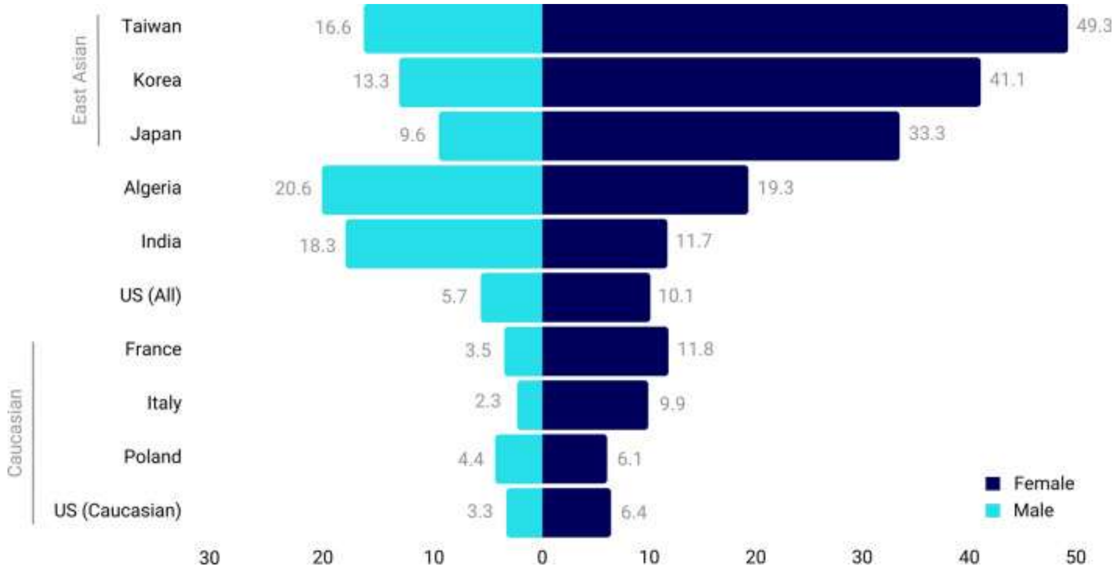
Lung cancer in never smokers

Increase in the proportion of NSCLC who are never smokers



Pelosof *et al.* 2017 *JNCI*

Enrichment of the prevalence of LCINS in females and East Asian countries



Wang *et al.* 2023 *Journal of Translational Medicine*

Tracing lung cancer mutational processes in never smokers: the Sherlock-*Lung* study

Tracing lung cancer mutational processes in never smokers: the Sherlock-*Lung* study

1st phase: 232 samples (188 LUAD / 226 European ancestry)

Tracing lung cancer mutational processes in never smokers: the Sherlock-*Lung* study

1st phase: 232 samples (188 LUAD / 226 European ancestry)

ARTICLES

https://doi.org/10.1038/s41588-021-00920-0

nature genetics

Check for updates

Genomic and evolutionary classification of lung cancer in never smokers

Tongwu Zhang[1], Philippe Joubert[2], Naser Ansari-Pour[3], Wei Zhao[1], Phuc H. Hoang[1], Rachel Lokanga[4], Aaron L. Moye[5], Jennifer Rosenbaum[6], Abel Gonzalez-Perez[7], Francisco Martínez-Jiménez[7], Andrea Castro[8], Lucia Anna Muscarella[9], Paul Hofman[10], Dario Consonni[11], Angela C. Pesatori[11,12], Michael Kebede[1], Mengying Li[1], Bonnie E. Gould Rothberg[13,14], Iliana Peneva[15,16], Matthew B. Schabath[17], Maria Luana Poeta[18], Manuela Costantini[19], Daniela Hirsch[4], Kerstin Heselmeyer-Haddad[4], Amy Hutchinson[1,20], Mary Olanich[1,20], Scott M. Lawrence[1,20], Petra Lenz[1,20], Maire Duggan[21], Praphulla M. S. Bhawsar[1], Jian Sang[1], Jung Kim[1], Laura Mendoza[1], Natalie Saini[22], Leszek J. Klimczak[23], S. M. Ashiqul Islam[24], Burcak Otlu[24], Azhar Khandekar[24], Nathan Cole[1,20], Douglas R. Stewart[1], Jiyeon Choi[1], Kevin M. Brown[1], Neil E. Caporaso[1], Samuel H. Wilson[22], Yves Pommier[25], Qing Lan[1], Nathaniel Rothman[1], Jonas S. Almeida[1], Hannah Carter[8], Thomas Ried[4], Carla F. Kim[5,26], Nuria Lopez-Bigas[7,27], Montserrat Garcia-Closas[1], Jianxin Shi[1], Yohan Bossé[2,28], Bin Zhu[1], Dmitry A. Gordenin[22], Ludmil B. Alexandrov[24], Stephen J. Chanock[1], David C. Wedge[3,29] and Maria Teresa Landi[1] ✉

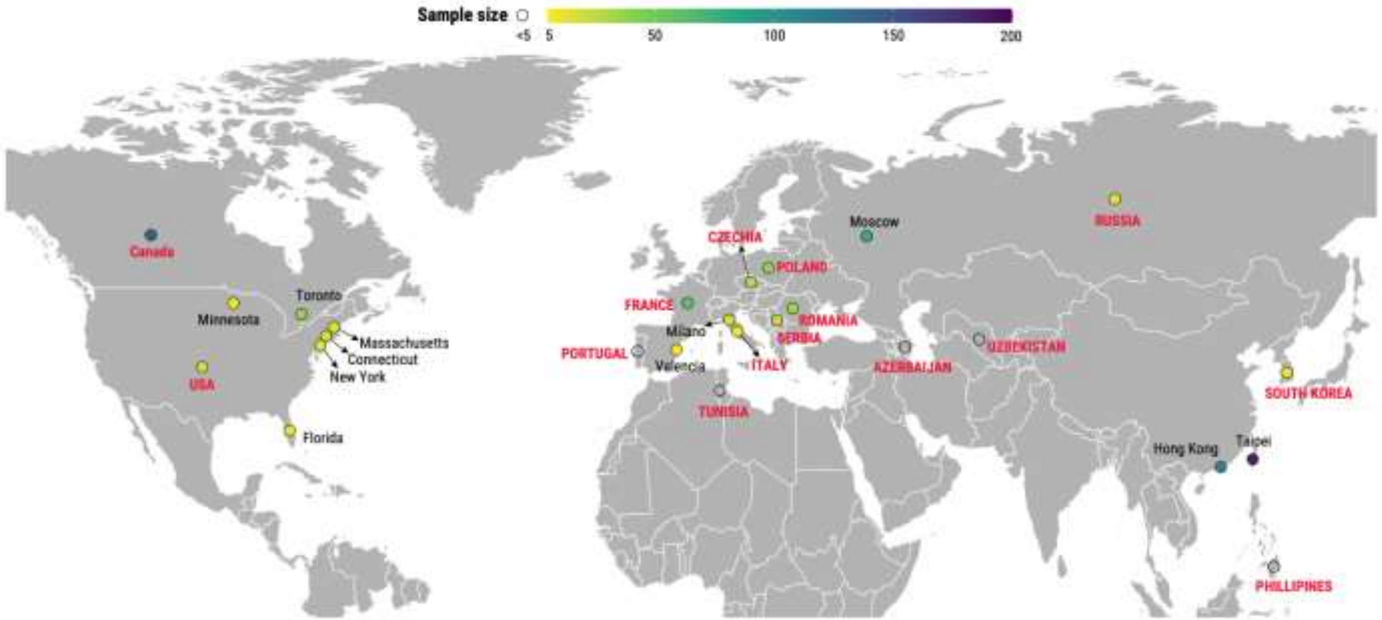
Zhang *et al*. 2021 *Nature Genetics*

Tracing lung cancer mutational processes in never smokers: the Sherlock-*Lung* study

2nd phase: 871 samples (28 regions across four continents)

Tracing lung cancer mutational processes in never smokers: the Sherlock-*Lung* study

2nd phase: 871 samples (28 regions across four continents)



Article

The mutagenic forces shaping the genomes of lung cancer in never smokers

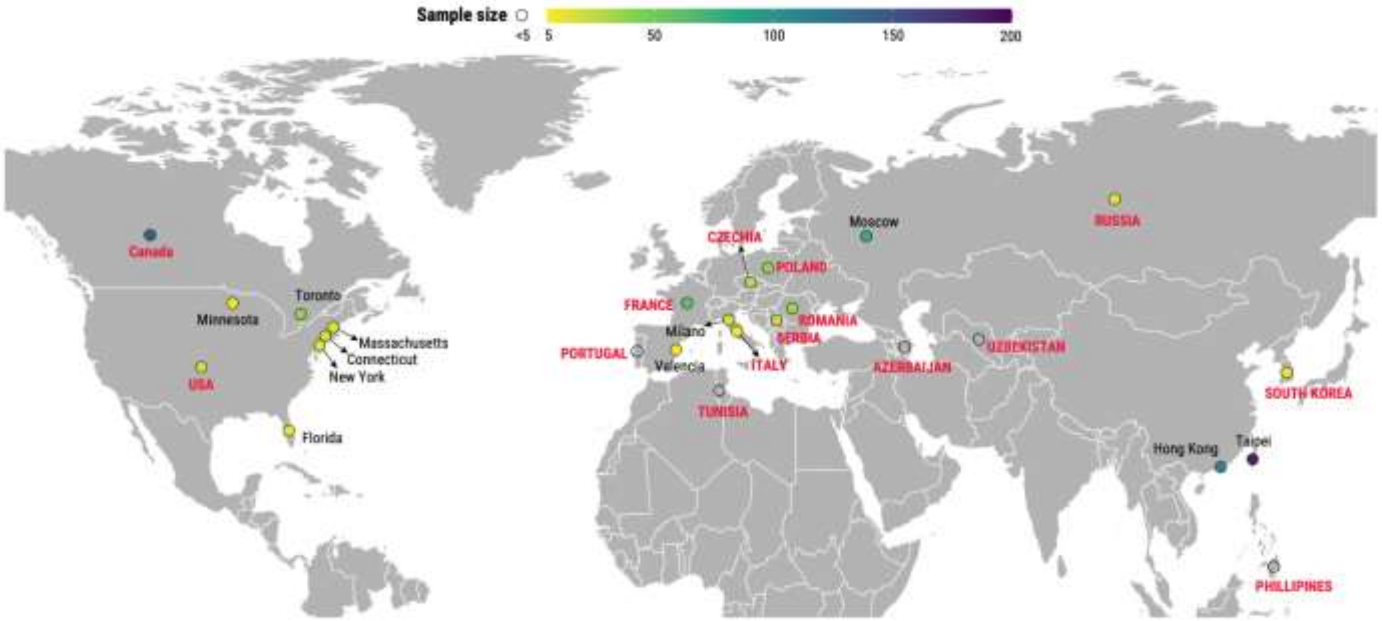
https://doi.org/10.1038/s41586-025-09219-0
Received: 16 May 2024
Accepted: 30 May 2025

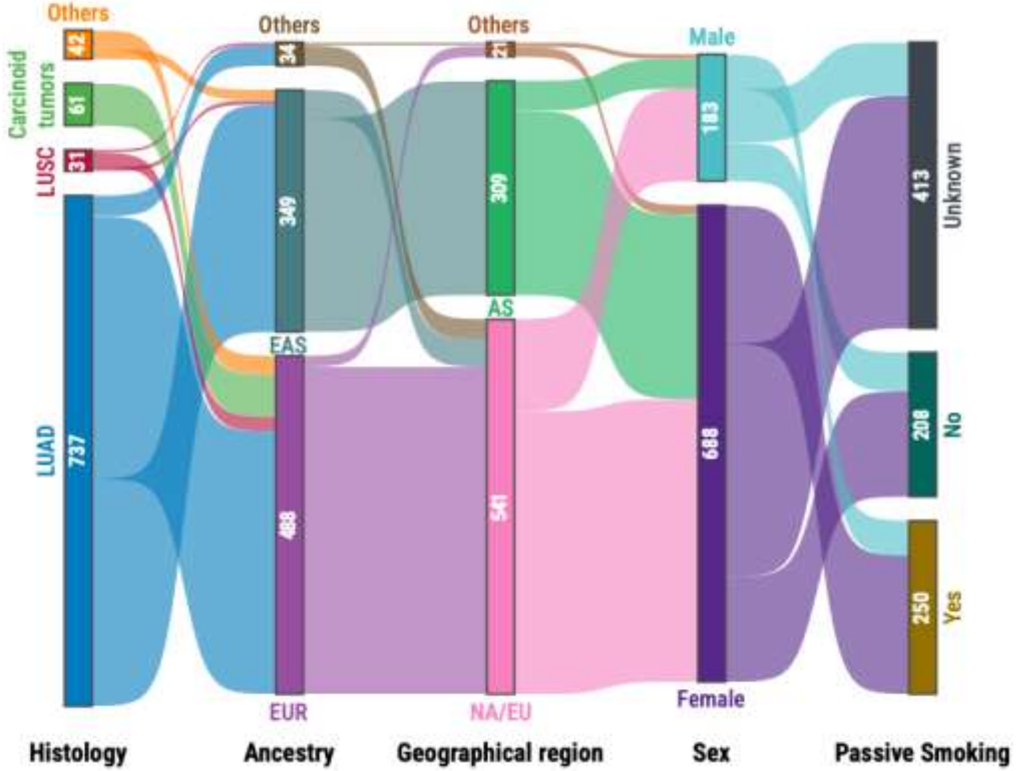
Marcos Díaz-Gay, Tongwu Zhang, Phuc H. Hoang, Charles Leduc, Marina K. Baine, William D. Travis, Lynette M. Sholl, Philippe Joubert, Azhar Khandekar, Wei Zhao, Christopher D. Steele, Burçak Otlu, Shuvro P. Nandi, Raviteja Vangara, Erik N. Bergstrom, Mariya Kazachkova, Oriol Pich, Charles Swanton, Chao Agnes Hsiung, I-Shou Chang, Maria Pik Wong, Kin Chung Leung, Jian Sang, John P. McElderry, Caleb Hartman, Frank J. Colón-Matos, Mona Miraftab, Monjoy Saha, Olivia W. Lee, Kristine M. Jones, Pilar Gallego-García, Yang Yang, Xiaoming Zhong, Eric S. Edell, Jacobo Martínez Santamaría, Matthew B. Schabath, Sai S. Yendamuri, Marta Manczuk, Jolanta Lissowska, Beata Świątkowska, Anush Mukeria, Oxana Shangina, David Zaridze, Ivana Holcatova, Dana Mates, Sasa Milosavljevic, Milica Kontic, Yohan Bossé, Bonnie E. Gould Rothberg, David C. Christiani, Valerie Gaborieau, Paul Brennan, Geoffrey Liu, Paul Hofman, Lixing Yang, Martin A. Nowak, Jianxin Shi, Nathaniel Rothman, David C. Wedge, Robert Homer, Soo-Ryum Yang, Angela C. Pesatori, Dario Consonni, Qing Lan, Bin Zhu, Stephen J. Chanock, Jiyeon Choi, Ludmil B. Alexandrov & Maria Teresa Landi

Díaz-Gay *et al.* 2025 *Nature*

Tracing lung cancer mutational processes in never smokers: the Sherlock-*Lung* study

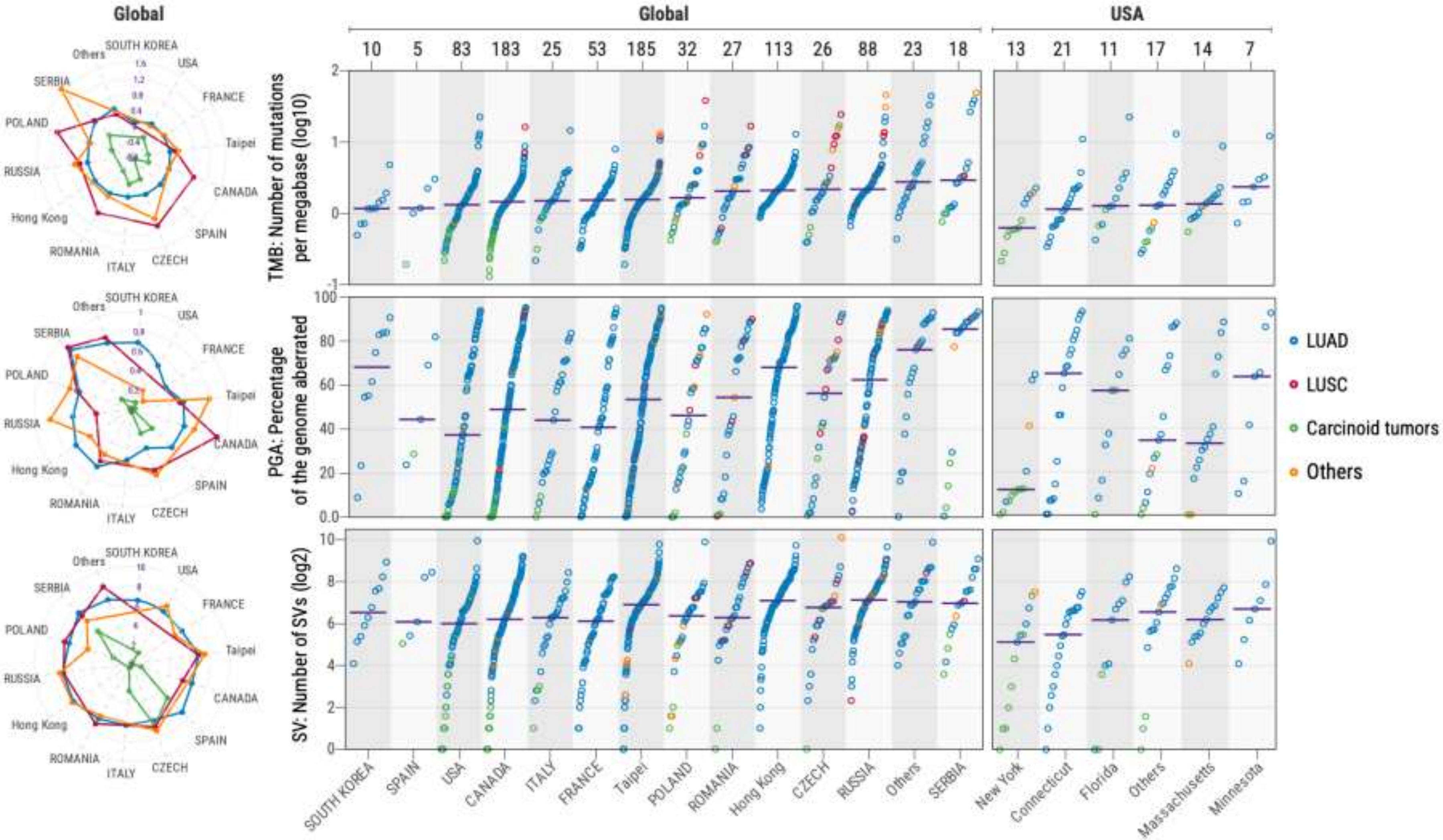
2nd phase: 871 samples (28 regions across four continents)



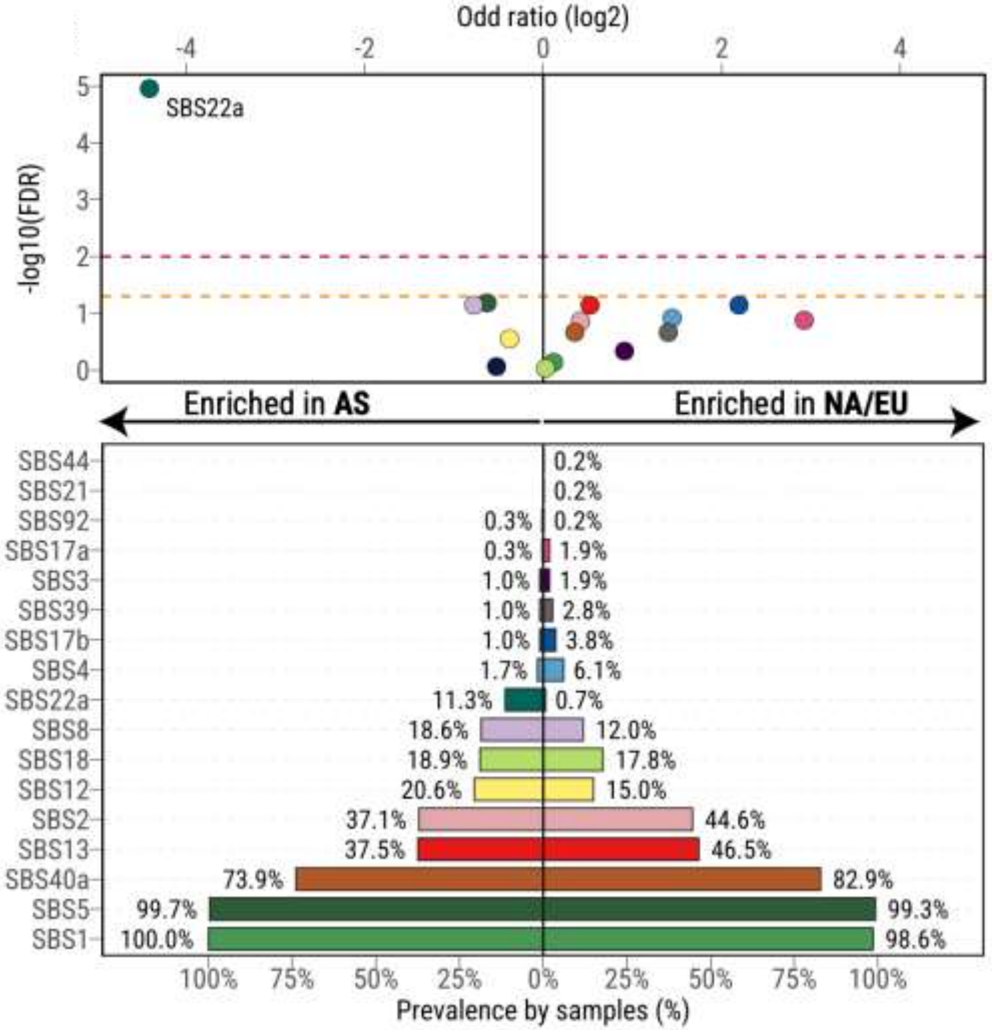


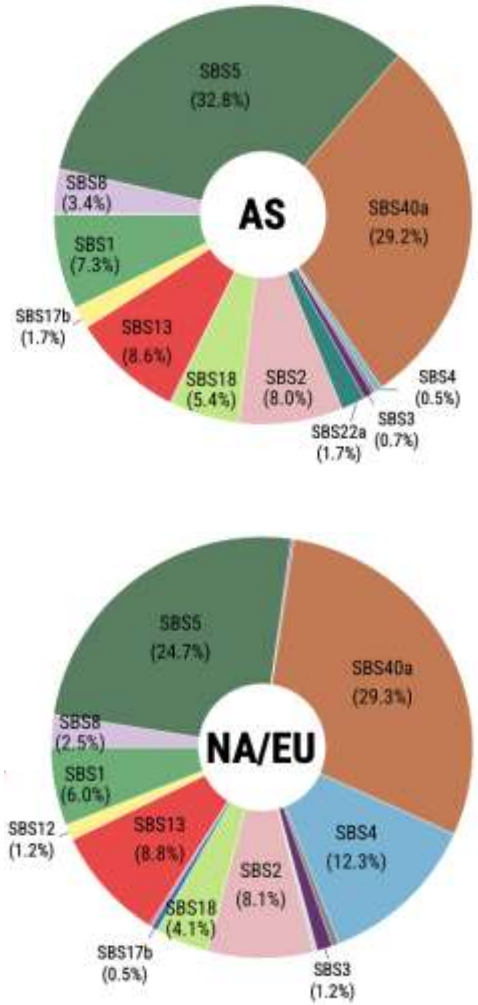
Díaz-Gay *et al.* 2025 *Nature*

We found geographical and histological differences across small and large mutational events

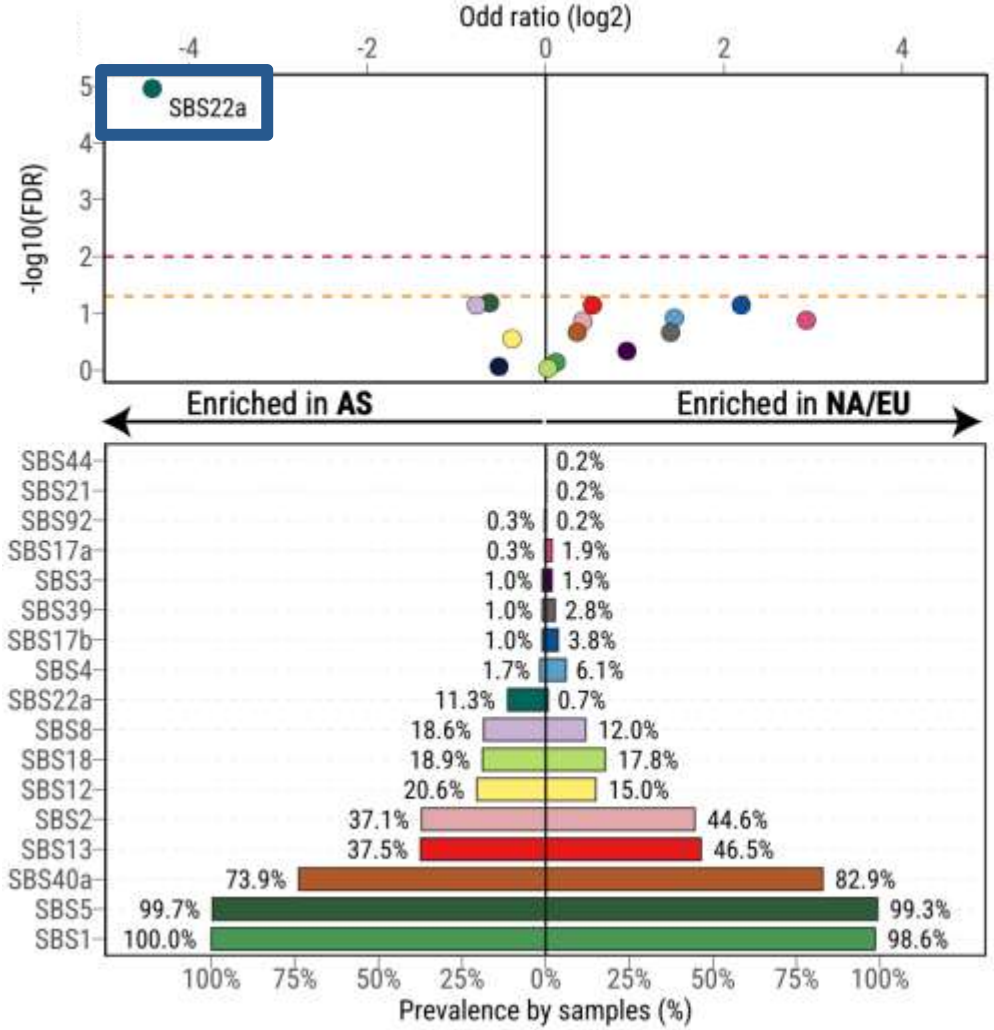


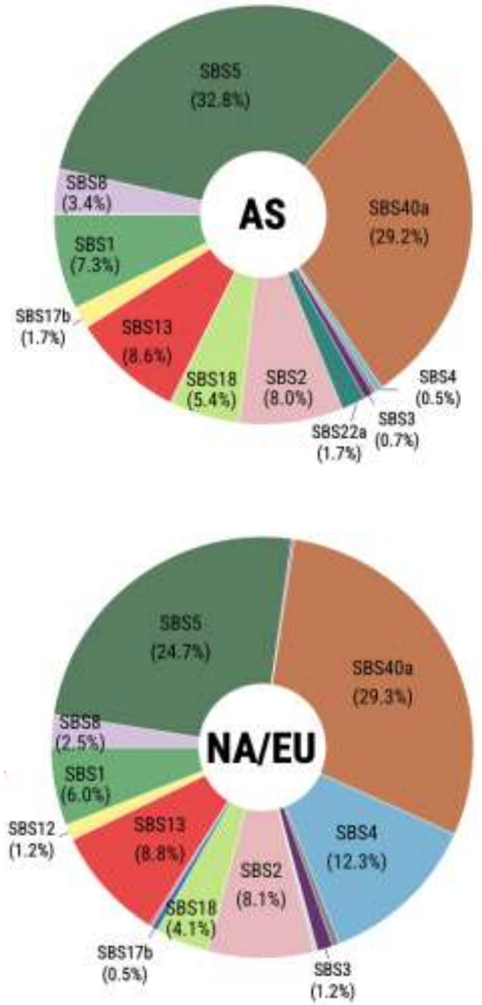
Genomic landscape of LUAD of never smokers



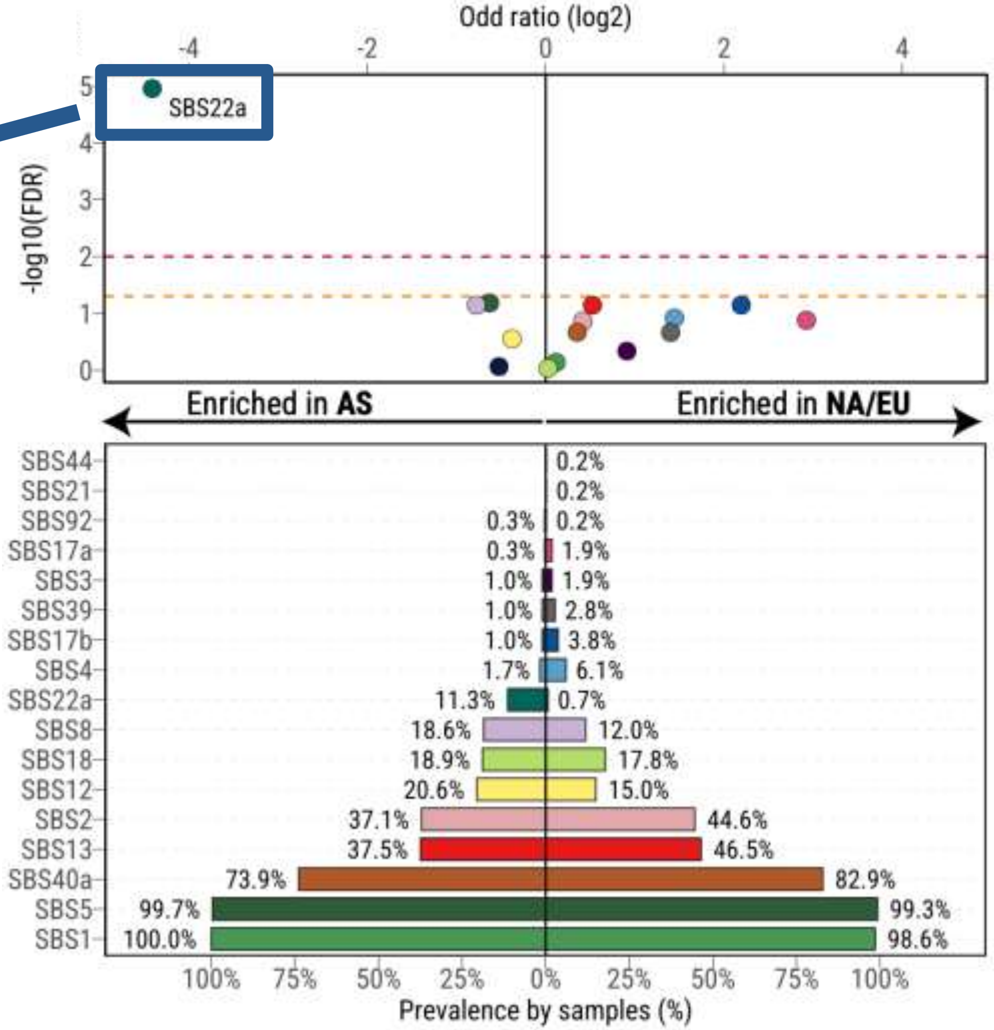


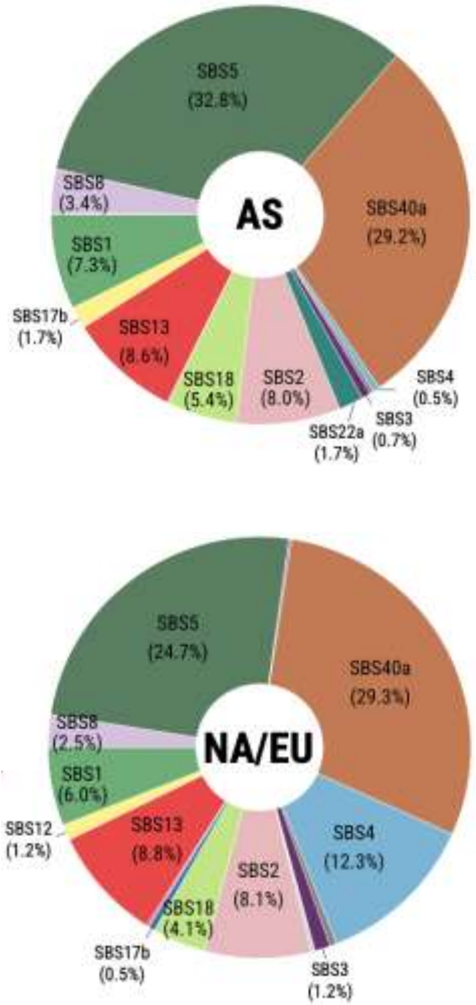
Genomic landscape of LUAD of never smokers



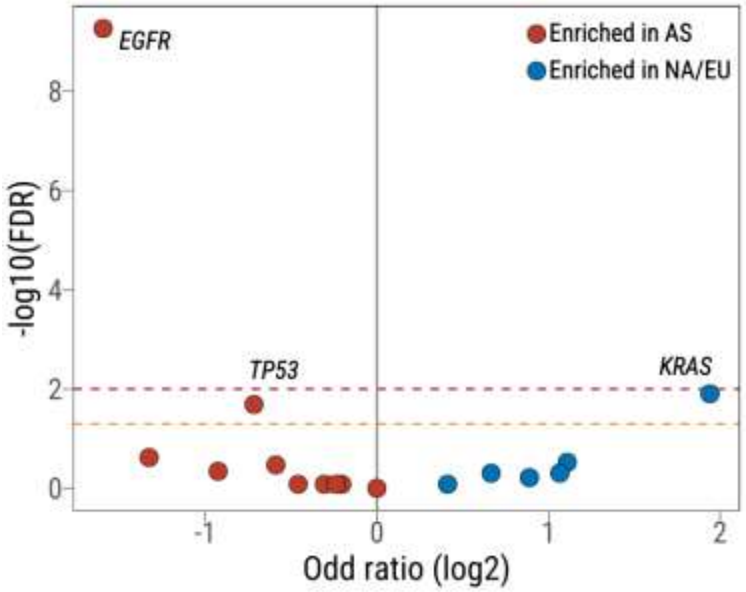


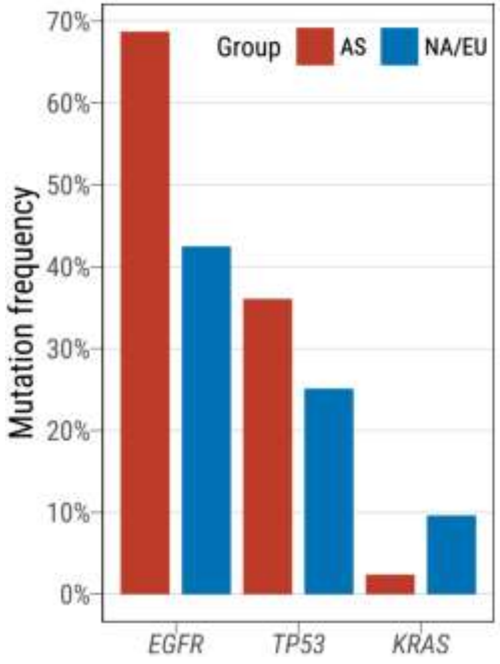
Genomic landscape of LUAD of never smokers

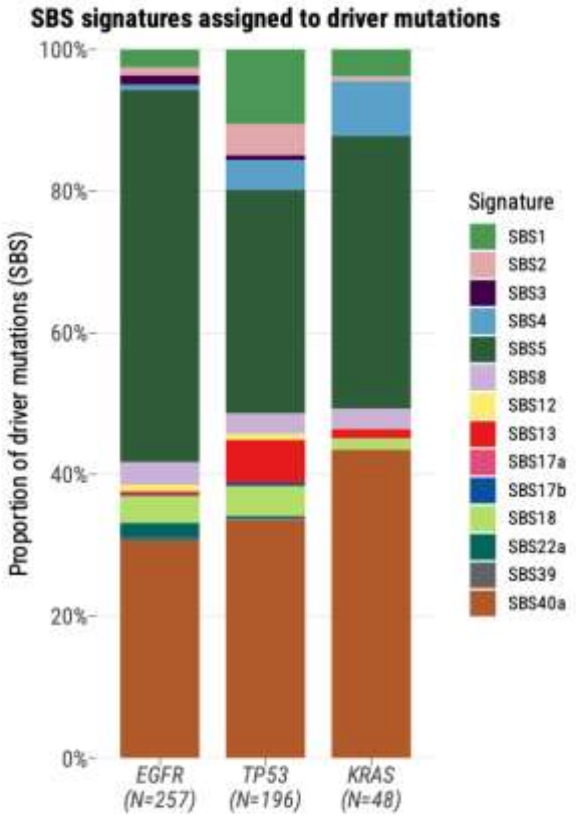




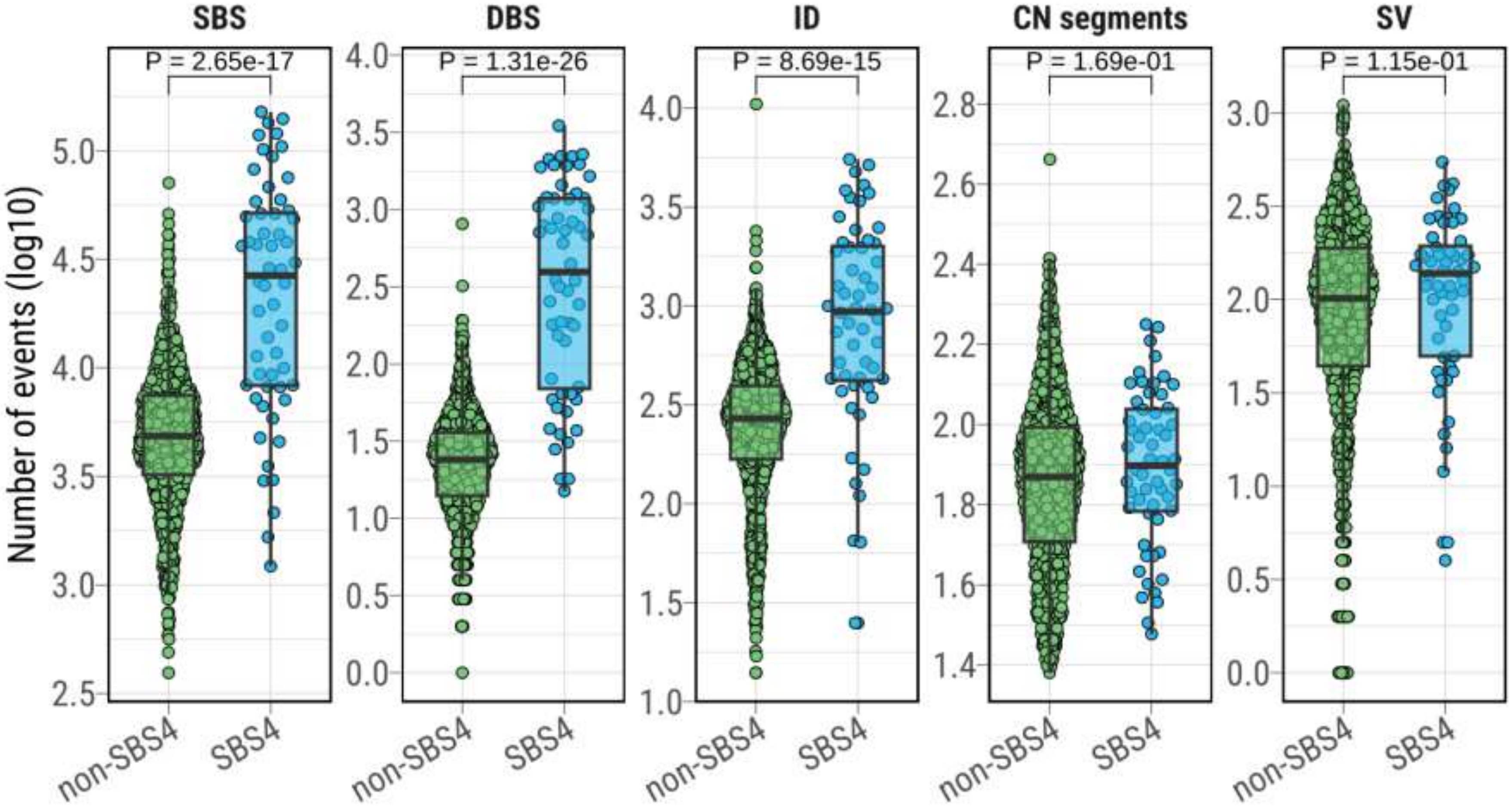
Genomic landscape of LUAD of never smokers



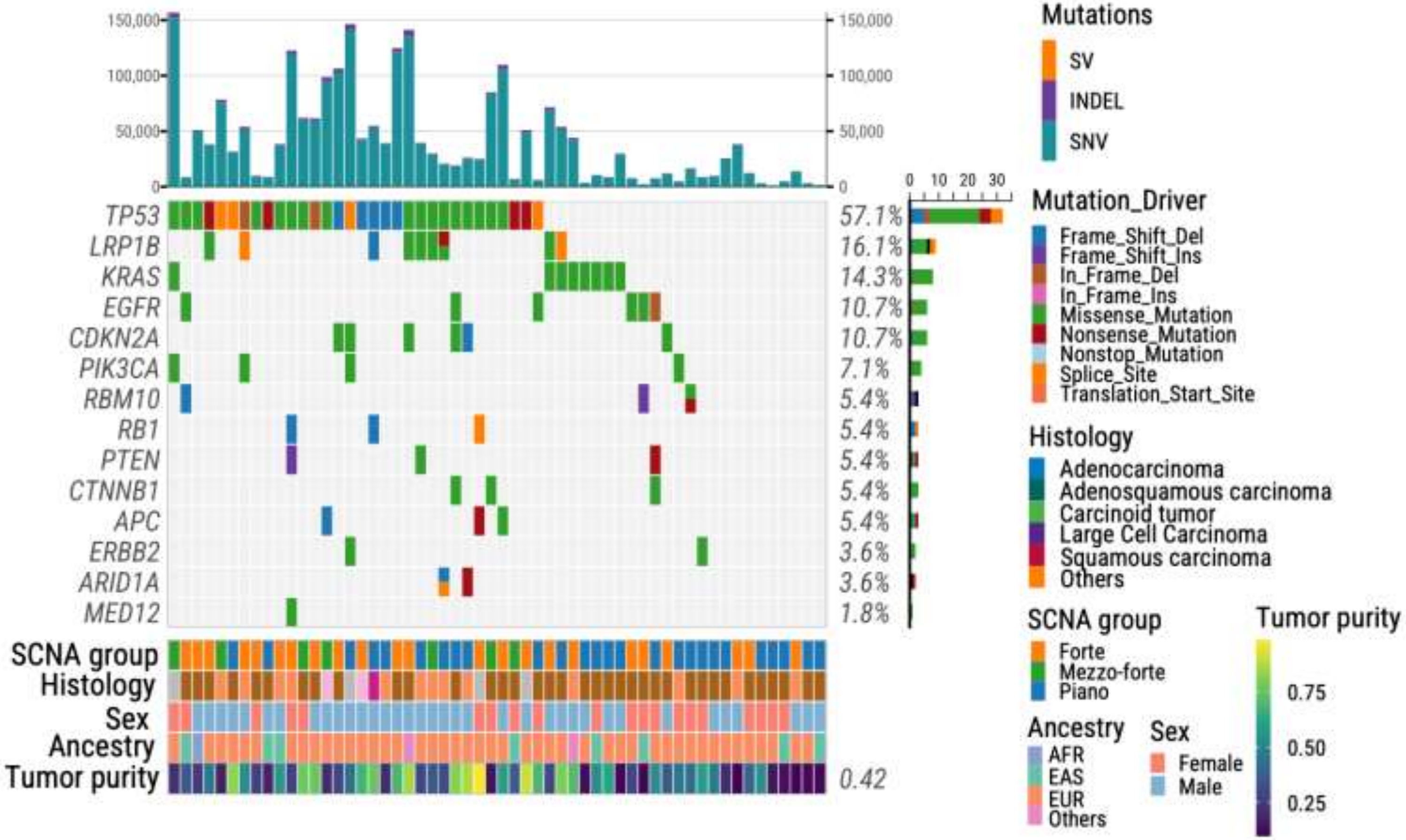



SBS signatures assigned to driver mutations

SBS4+ never-smoker cases have elevated mutation burden



SBS4+ never-smoker cases exhibited different driver mutations compared to smokers



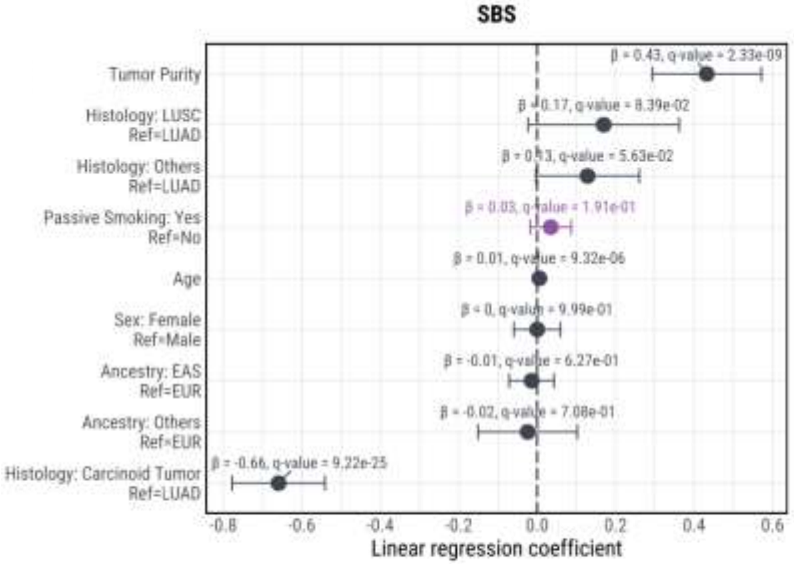
Driver gene comparison

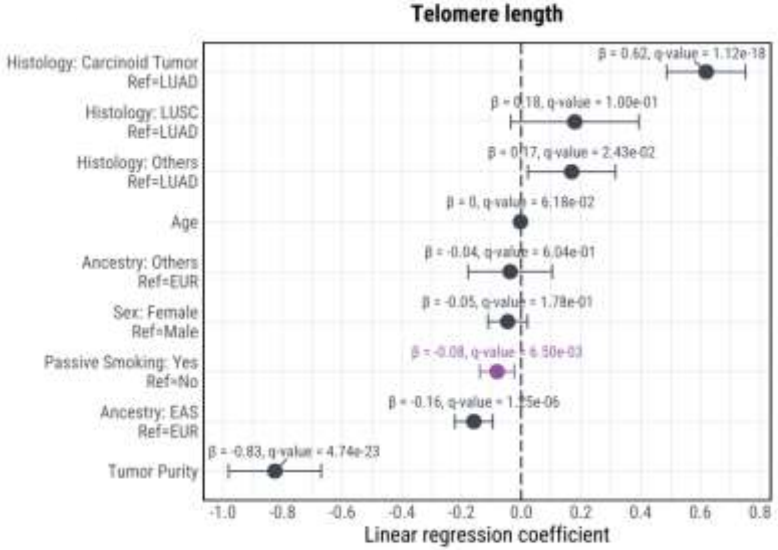
never smokers vs. smokers

- Lower KRAS mutations (OR=0.41, p=0.021)
- **Higher EGFR L858R hotspot mutation (OR=29.4, p=4.1×10^{-4})**

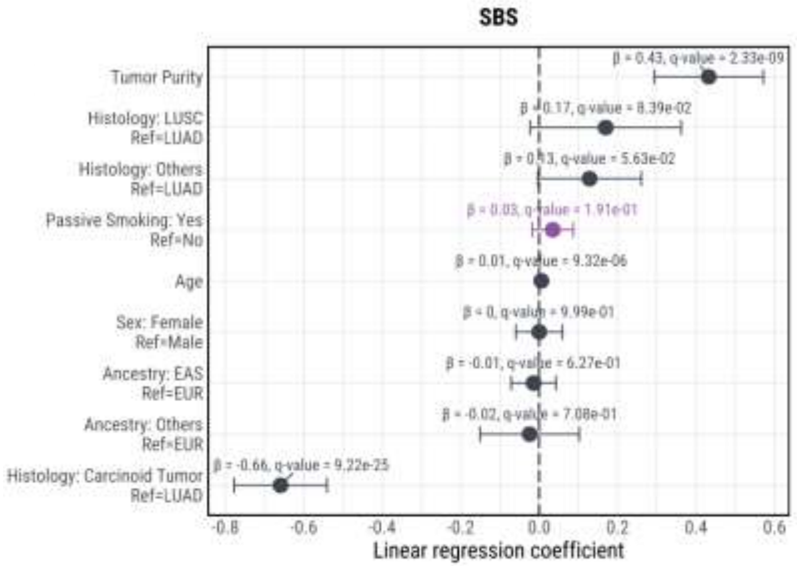
Passive smoking showed low mutagenicity

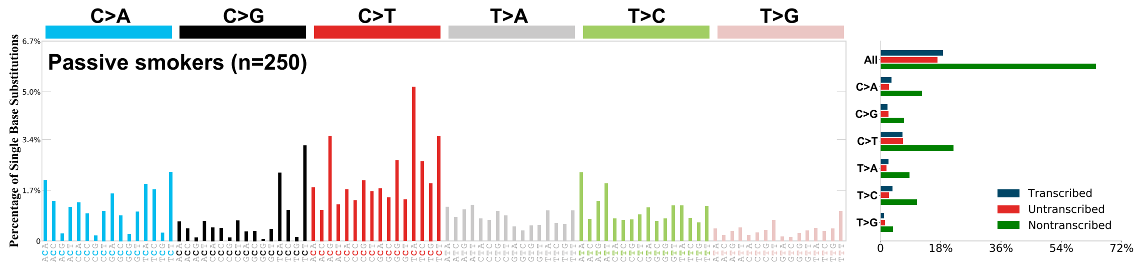
Passive smoking showed low mutagenicity

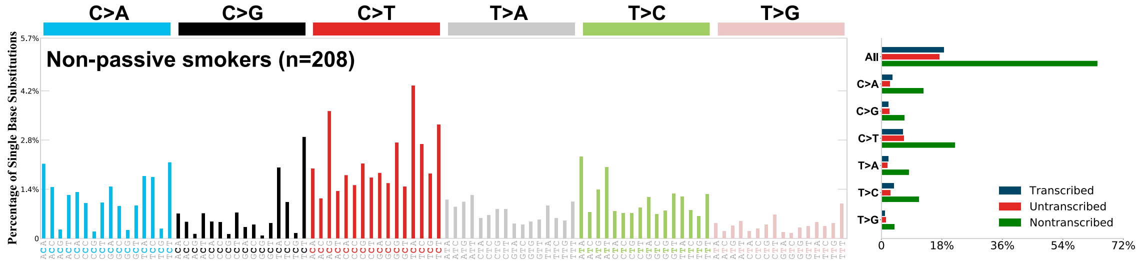


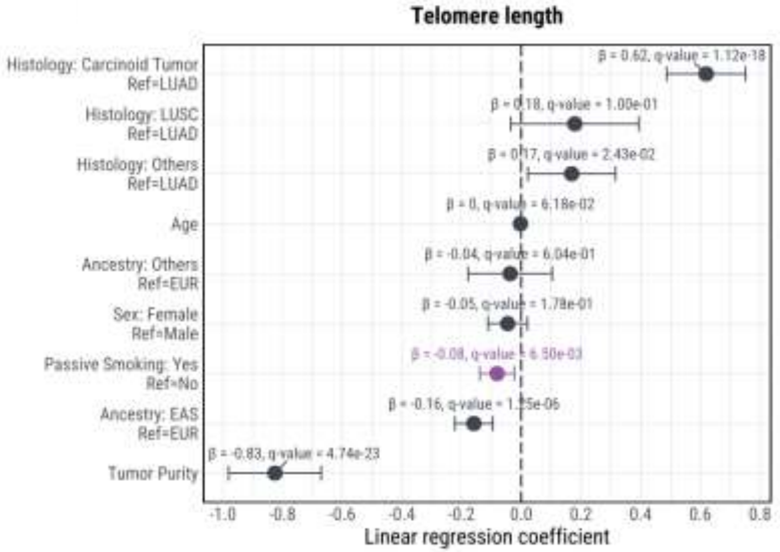


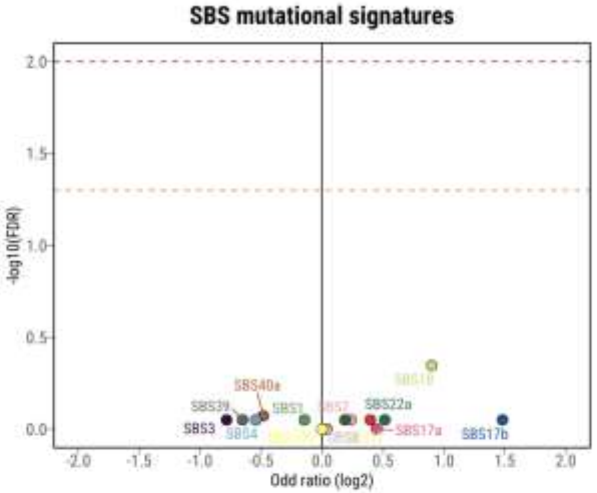
Passive smoking showed low mutagenicity

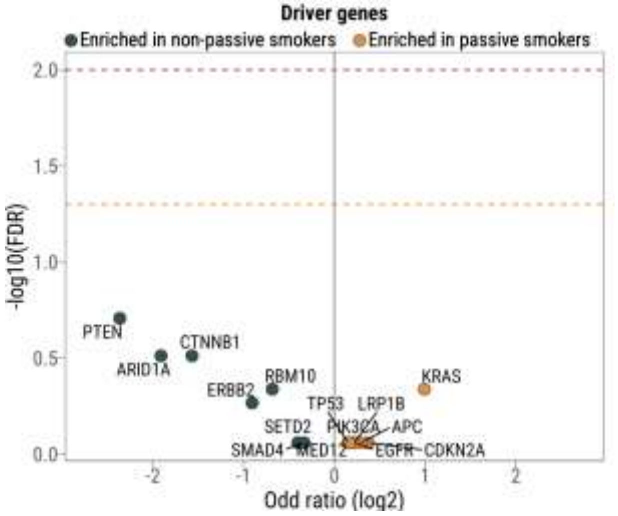






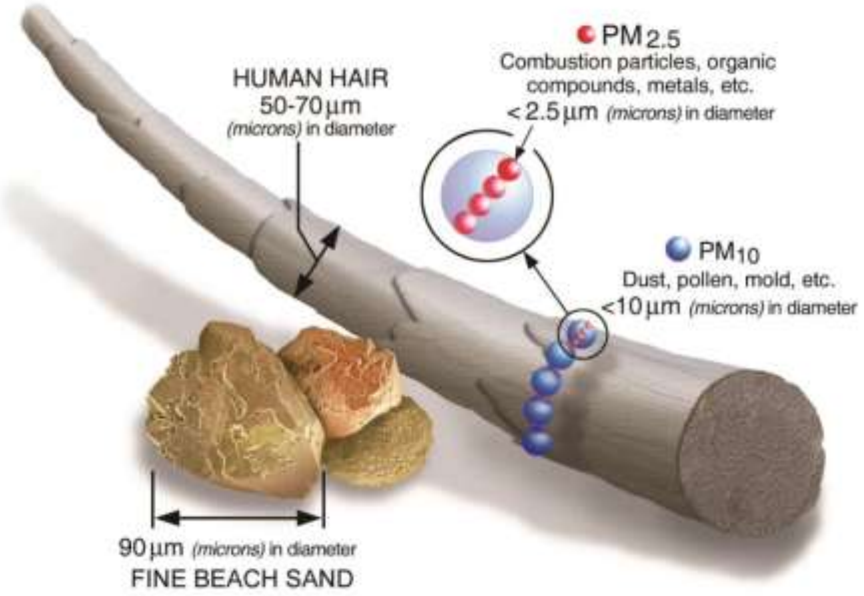




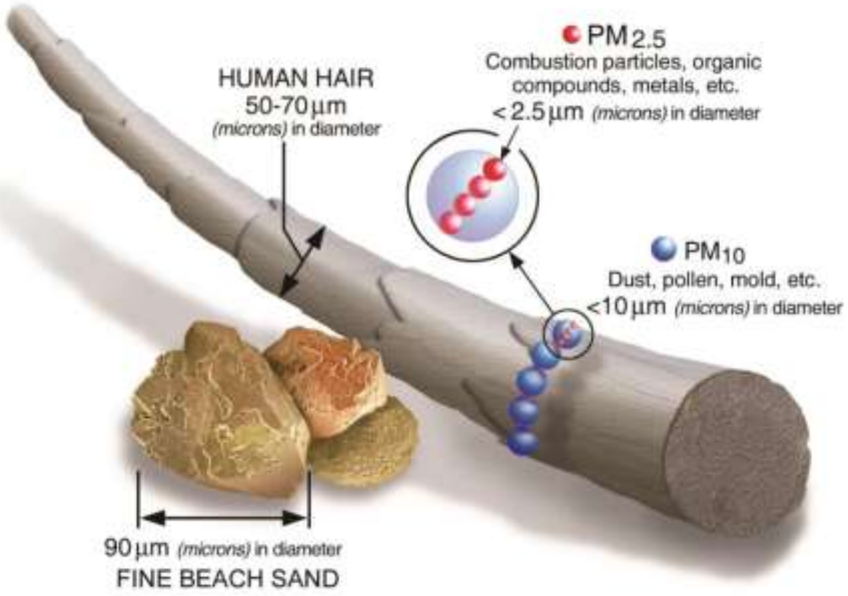


Evaluated the effect of outdoor air pollution estimated by the environmental particulate matter <2.5 µm ($PM_{2.5}$)

Evaluated the effect of outdoor air pollution estimated by the environmental particulate matter <2.5 µm ($PM_{2.5}$)



Evaluated the effect of outdoor air pollution estimated by the environmental particulate matter <2.5 µm ($PM_{2.5}$)



Atmospheric Composition Analysis Group

Surface PM2.5

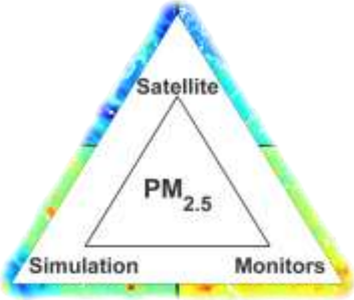
Contents:
V5.GL.03 is recommended for all regions
V4.NA.03 is available for compositional use over North America

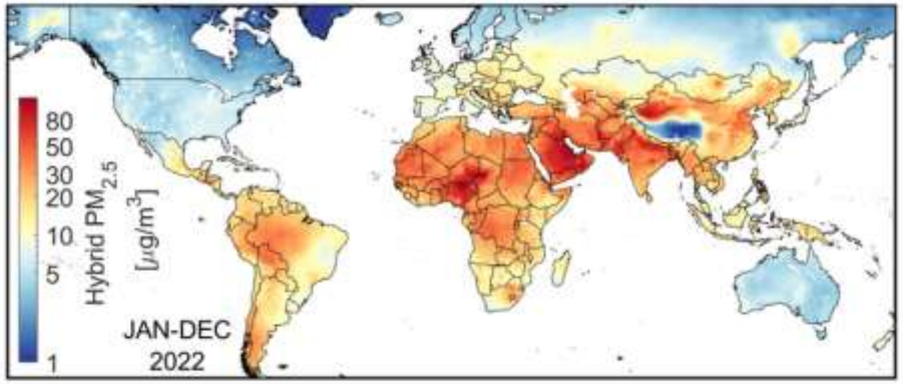
Previous versions, including V4.NA.02.MAPLE, are also available.

Additions tools developed by users for accessing these datasets are available here

Global/Regional Estimates (V5.GL.03):

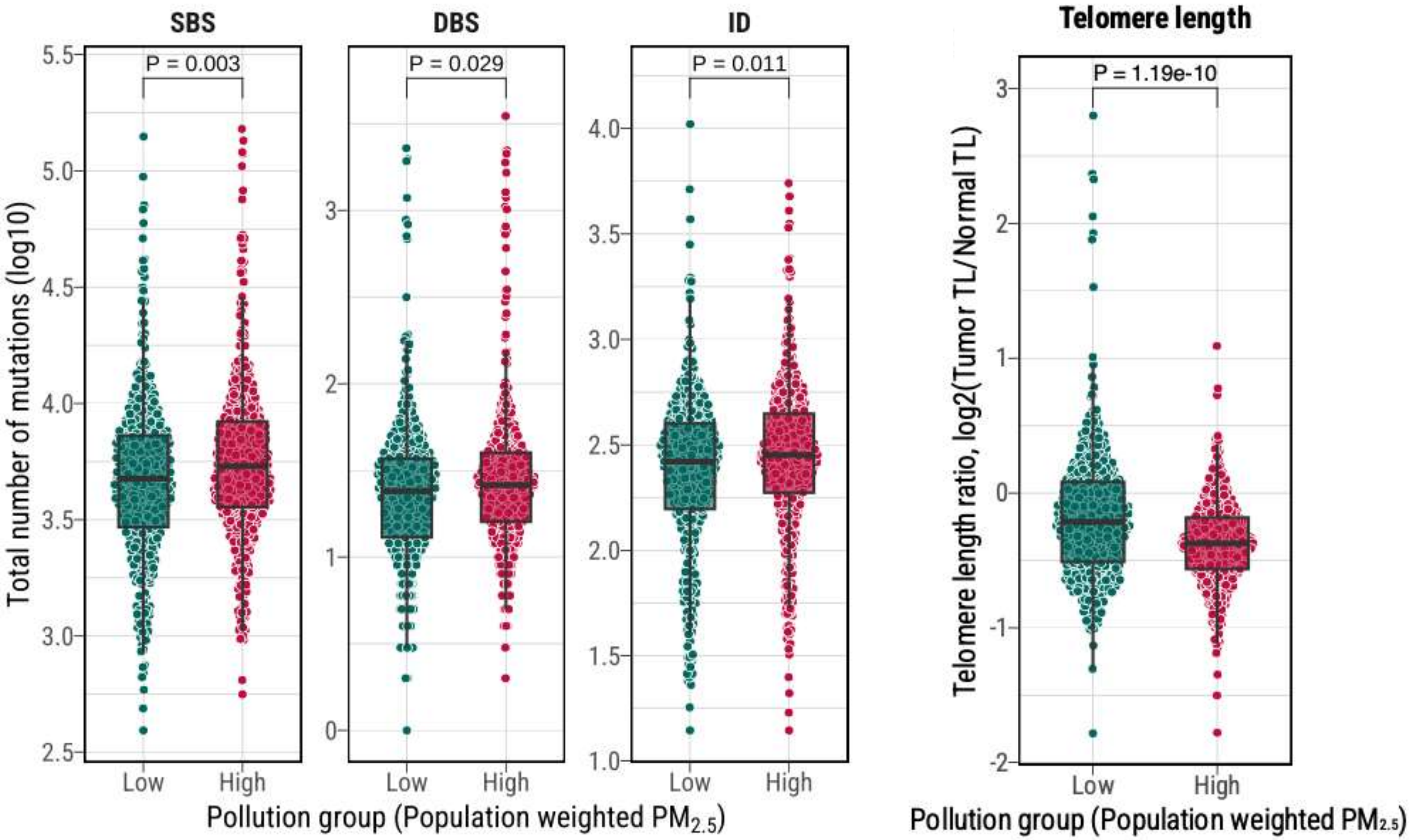
We estimate annual and monthly ground-level fine particulate matter ($PM_{2.5}$) for 1998-2021 by combining Aerosol Optical Depth (AOD) retrievals from the NASA MODIS, MISR, and SeaWiFS instruments with the GEOS-Chem chemical transport model, and subsequently calibrating to global ground-based observations using a Geographically Weighted Regression (GWR), as detailed in the below reference for V5.GL.01. V5.GL.03 follows the methodology of V5.GL.01, but updates the ground-based observations used to calibrate the geophysical PM2.5 estimates for the entire time series, and extends temporal coverage through 2021.



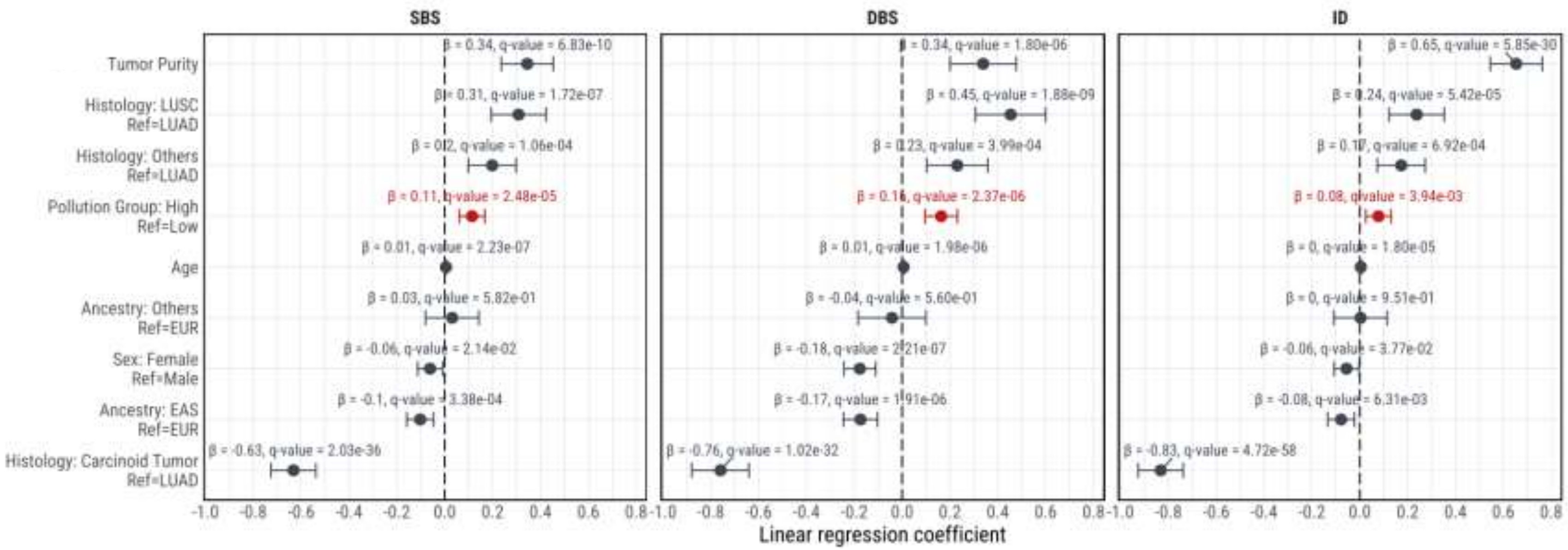


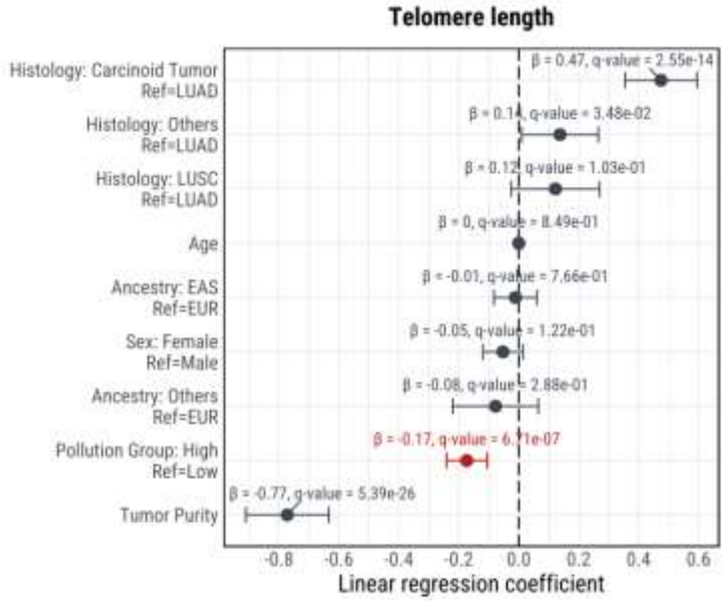
Van Donkelaar *et al.* 2021 *Environmental Science & Technology*

High $PM_{2.5}$ exposure corresponds to an increase in genomic alterations and shorter telomeres



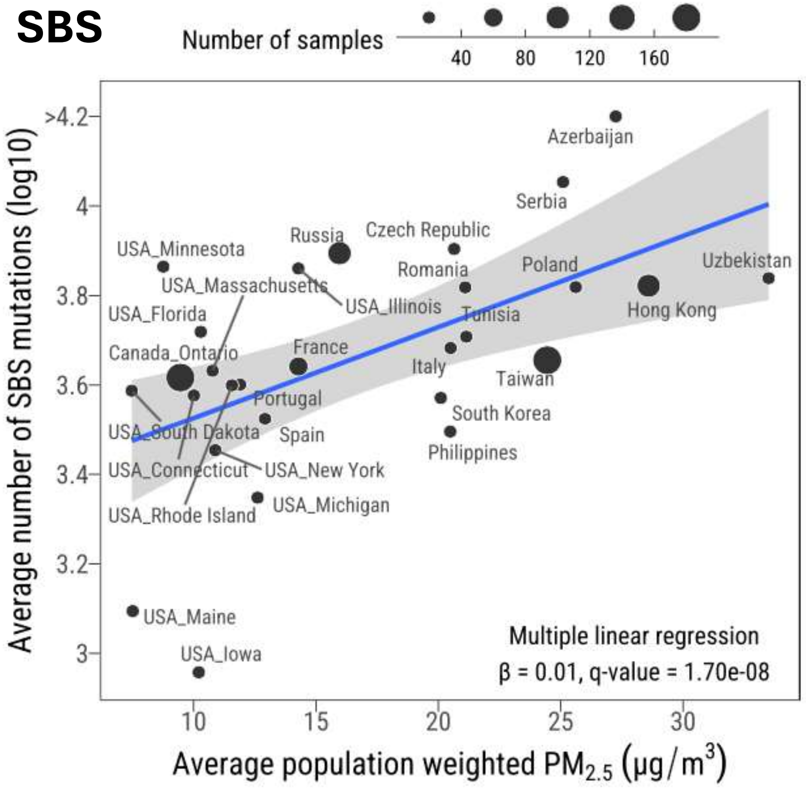
High PM$_{2.5}$ exposure corresponds to an increase in genomic alterations and shorter telomeres





High $PM_{2.5}$ exposure corresponds to an increase in genomic alterations and shorter telomeres

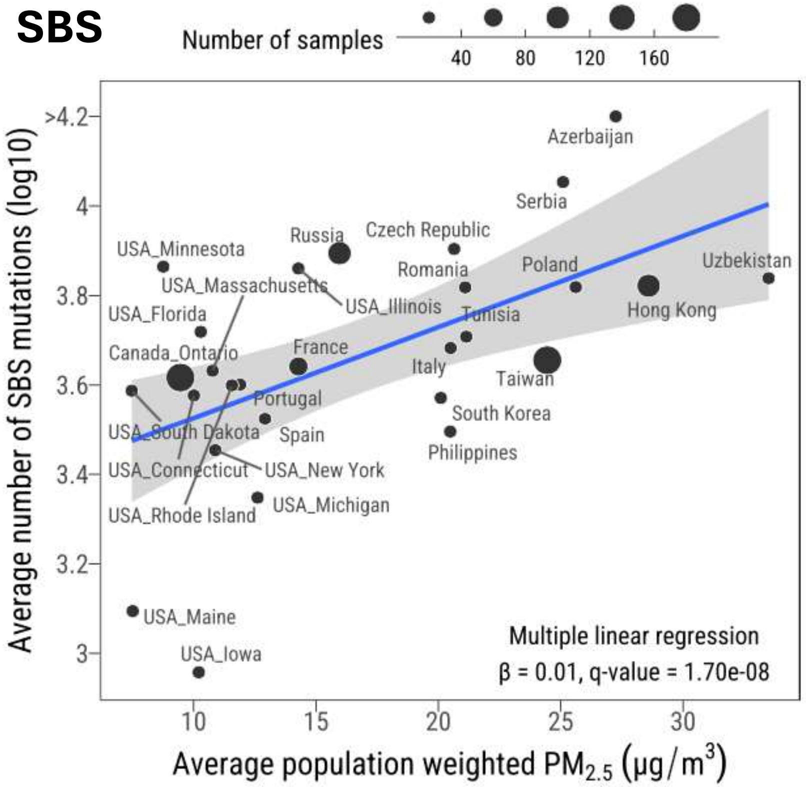
SBS



Regression model adjusted by age + sex + genetic ancestry + histology + tumor purity

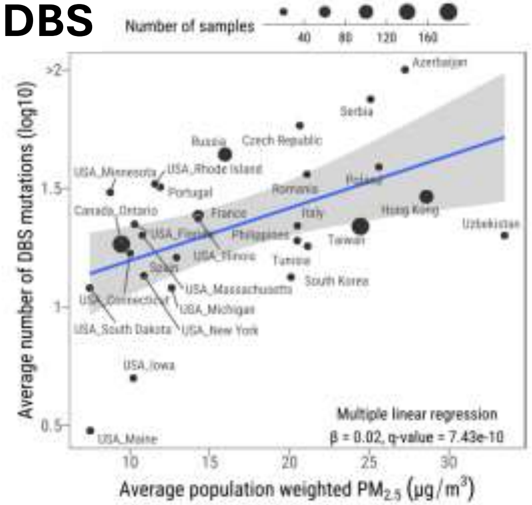
High $PM_{2.5}$ exposure corresponds to an increase in genomic alterations and shorter telomeres

SBS

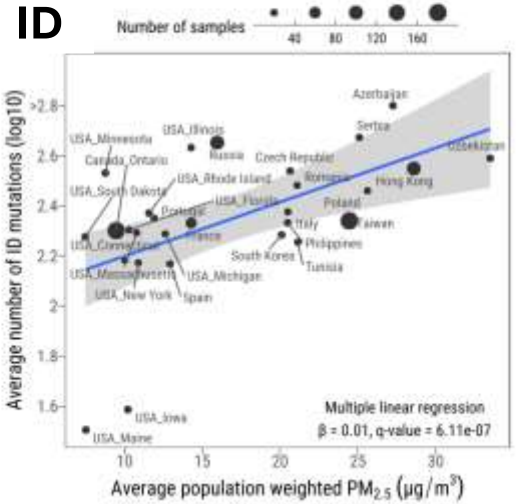


Regression model adjusted by age + sex + genetic ancestry + histology + tumor purity

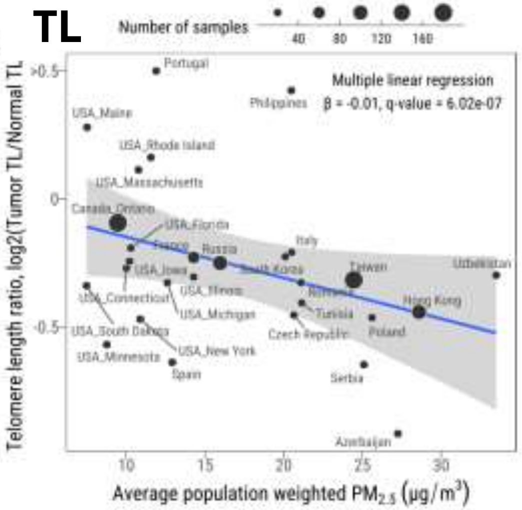
DBS



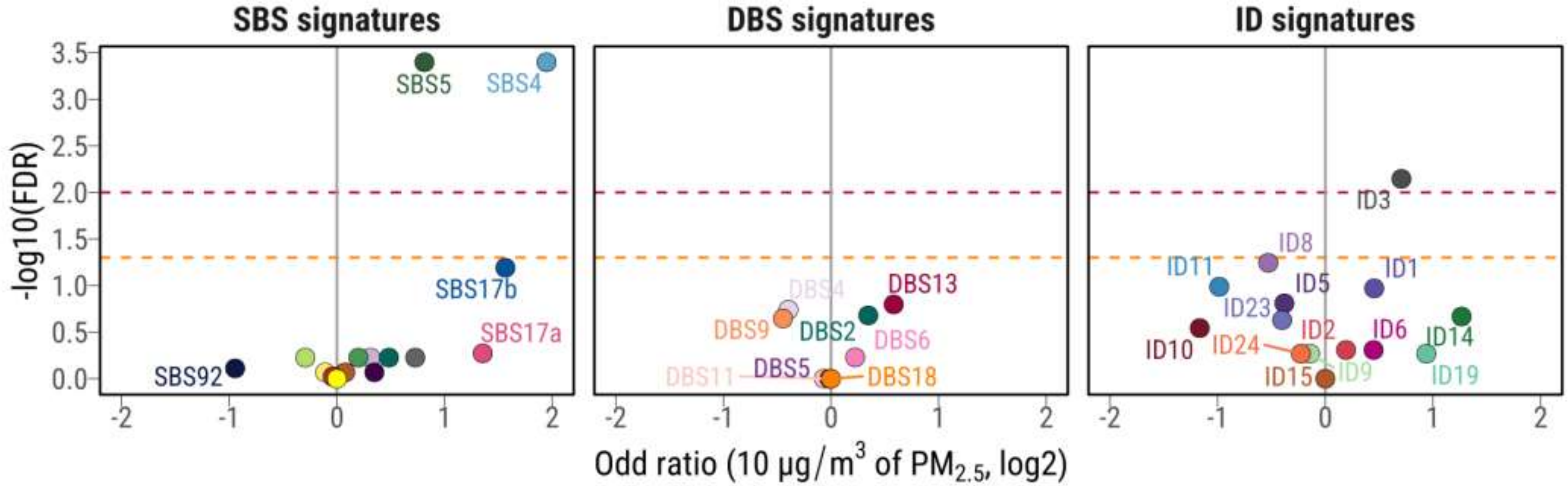
ID



TL

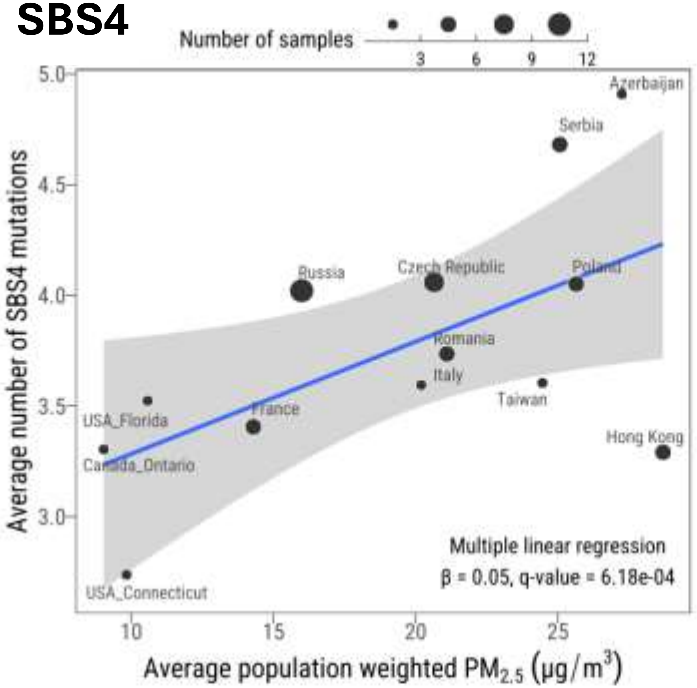


High PM$_{2.5}$ exposure corresponds to an increase in specific mutational signatures

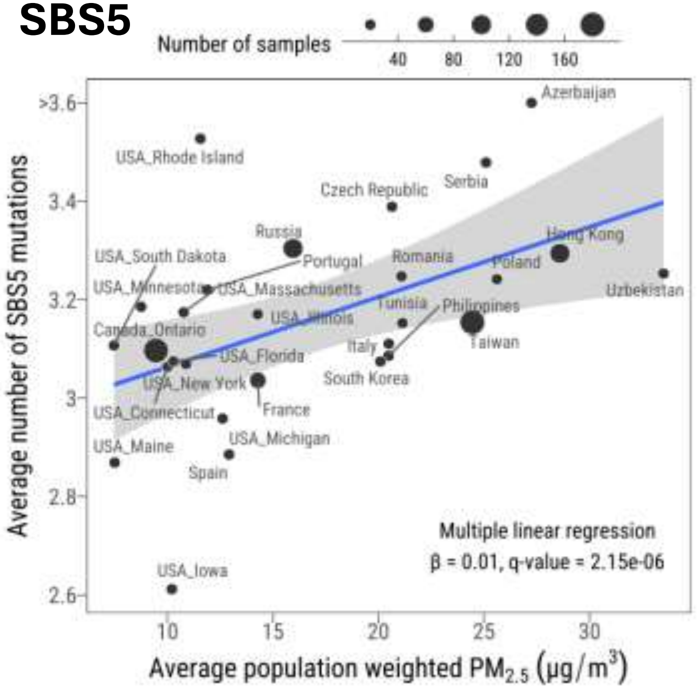


High PM$_{2.5}$ exposure corresponds to an increase in specific mutational signatures

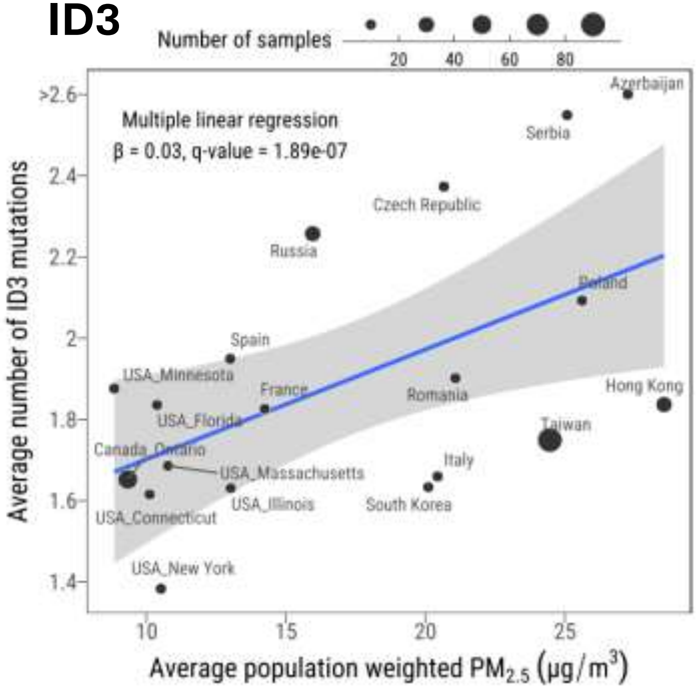
SBS4



SBS5

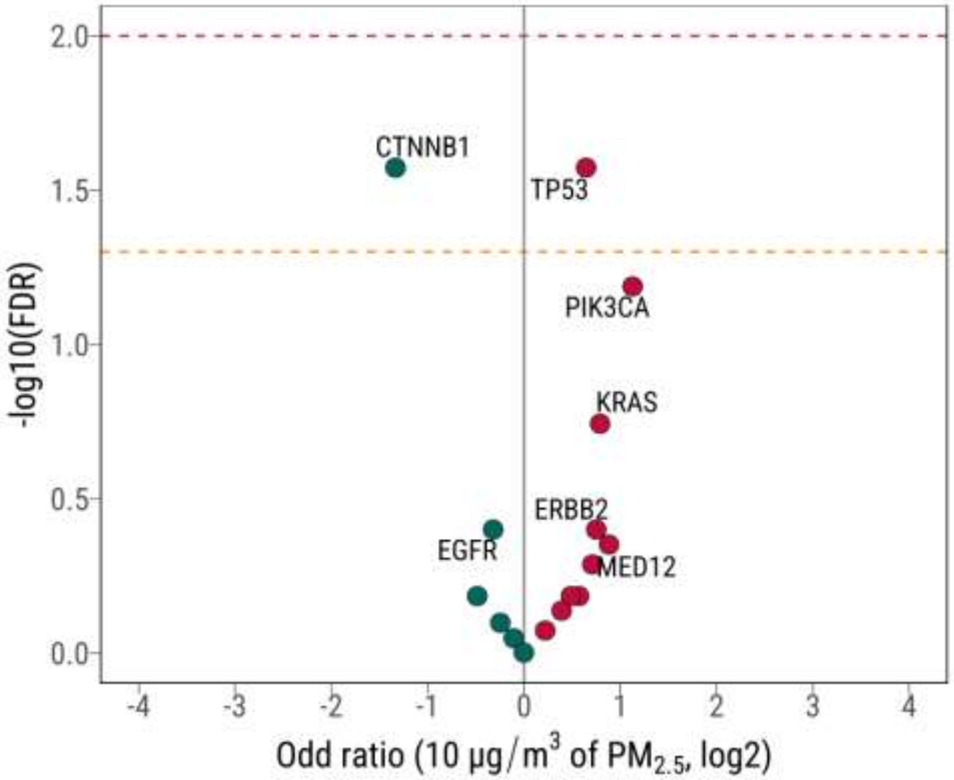


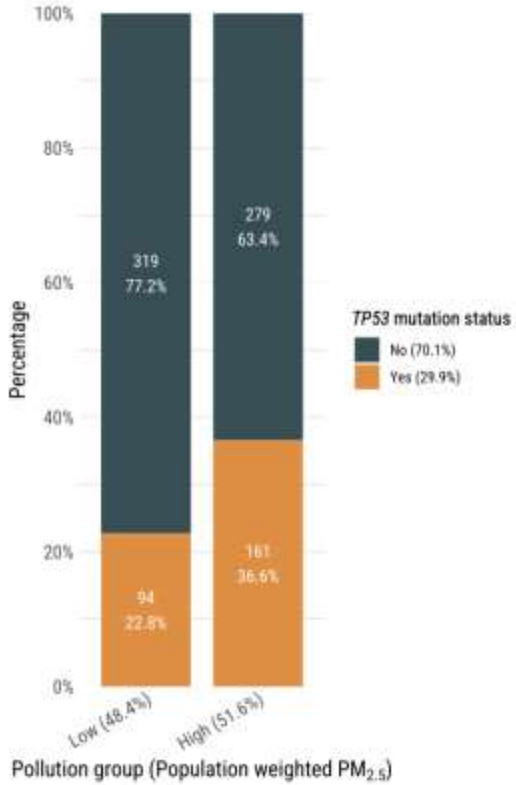
ID3

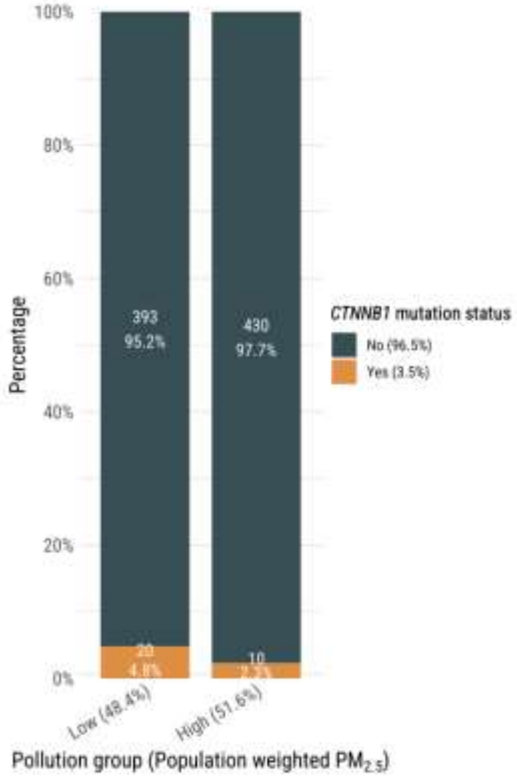


Regression models adjusted by age + sex + genetic ancestry + histology + tumor purity

High $PM_{2.5}$ exposure corresponds to an enrichment in specific driver mutations







Take-home messages

Mutational signatures framework

- Mutational signature analysis allows the identification of the endogenous and exogenous biological processes generating somatic mutations in human cancer

Global genomics characterization of lung cancer in never smokers

- Tumors from highly polluted regions have shorter telomeres and higher mutagenesis, linked to specific signatures & TP53 muts.
- Passive smoking had a low mutation burden and no associated signature or drivers, whereas regional differences were found.

Acknowledgements

cnio SPANISH NATIONAL CANCER RESEARCH CENTRE

Digital Genomics Group

Pilar Gallego García
José Córdoba Caballero
Sandra Ben Romdhane

NIH NATIONAL CANCER INSTITUTE

Tongwu Zhang
Phuc Hoang
Azhar Khandekar
Stephen Chanock
Maria Teresa Landi

COSMIC Mutational Signatures

Tom Mutimer
Stephen Duke

UC San Diego

S M Ashiqul Islam
Raviteja Vangara
Mariya Kazachkova
Ammal Abbasi
Jessica Au

The Alexandrov Lab

Mark Barnes
Erik Bergstrom
Chris Steele
Burçak Otlu
Ludmil Alexandrov

wellcome sanger institute

Sarah Moody
Mike Stratton

International Agency for Research on Cancer
World Health Organization

Wellington dos Santos
Paul Brennan

GOBIERNO DE ESPAÑA MINISTERIO PARA LA TRANSFORMACIÓN DIGITAL Y DE LA FUNCIÓN PÚBLICA

Comunidad de Madrid

GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Cofinanciado por la Unión Europea

AGENCIA ESTATAL DE INVESTIGACIÓN

THANK YOU

GECP
lung cancer
research

GECP
lung cancer
research

16th CONGRESS Lung ON CANCER

BARCELONA
27 / 28
NOVEMBER 2025

Analyzing mutagenic events in never smokers lung cancer

Marcos Díaz Gay

Centro Nacional de Investigaciones Oncológicas (CNIO)