

Genomic and Evolutionary Classification of Lung Cancers in Never Smokers from the *Sherlock-Lung* Study

MARIA TERESA LANDI, M.D., PH.D.

DCEG, NCI, NIH

Lung cancer in never smokers (LCINS):

2

- Accounts for 15-25% of all lung cancer cases worldwide
- Is the 7th cause of cancer mortality in the US
- Has predominant adenocarcinoma subtype
- Only a few risk factors are known, accounting for a small proportion of the cases

The Sherlock-*Lung* study

3

To use tumor genomic changes as “footprints” to infer etiological processes and evolutionary trajectories of LCINS tumorigenesis



Sherlock-Lung

To mine large electronic health records (EHR) to uncover associations of LCINS with other medical conditions or long-term medication use



Radon
Coal
Asbestos
Pollution
IPF...

Special exposures
population (n~300)

Design



4

General population
(n~1700)



2000 tumor samples/
2000 blood samples

1000 tumor samples/
1000 normal lung tissues

Whole genome sequencing

RNA seq + methylation arrays

- Mutational Signatures
- Mutational burden
- Driver genes
- Telomere length
- Neoepitopes
- Viral remnants
- Copy numbers
- Structural variants
- Germline variants

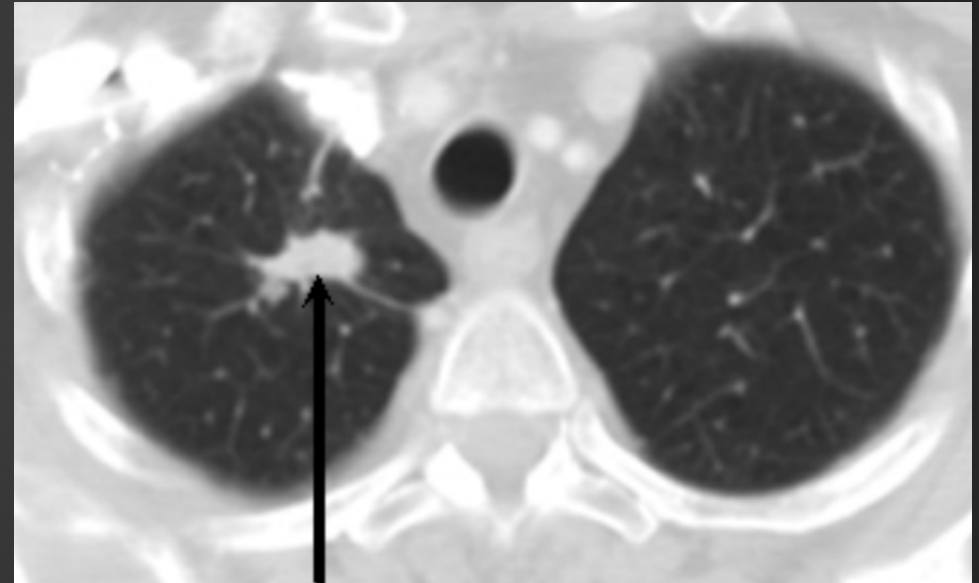
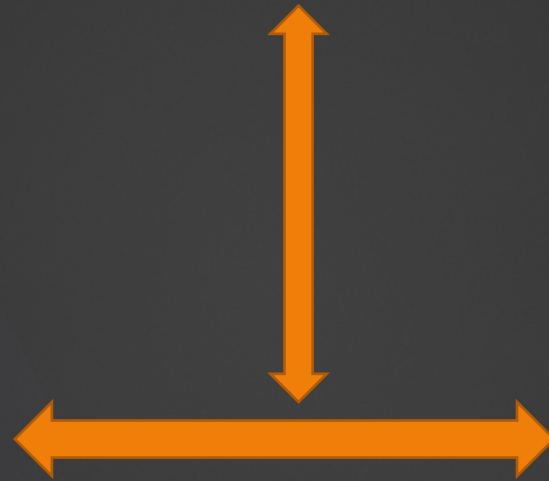
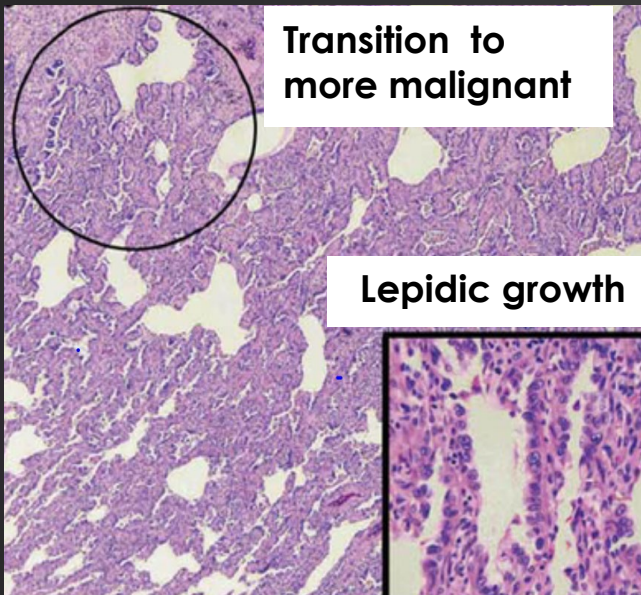
Aim 1

Characterize the
genomic and
evolutionary
landscape of LCINS

- Gene expression
- Inferred mutations
- Fusions
- Structural changes
- Methylation patterns
- CIMP
- Methylation signatures
- Immune cell subtypes
- 16S RNA – microbiome

Aim 2: Develop an integrated molecular, histological and radiological classification of lung cancer in never smokers

Molecular landscape





6

**General population
(n~1700)**

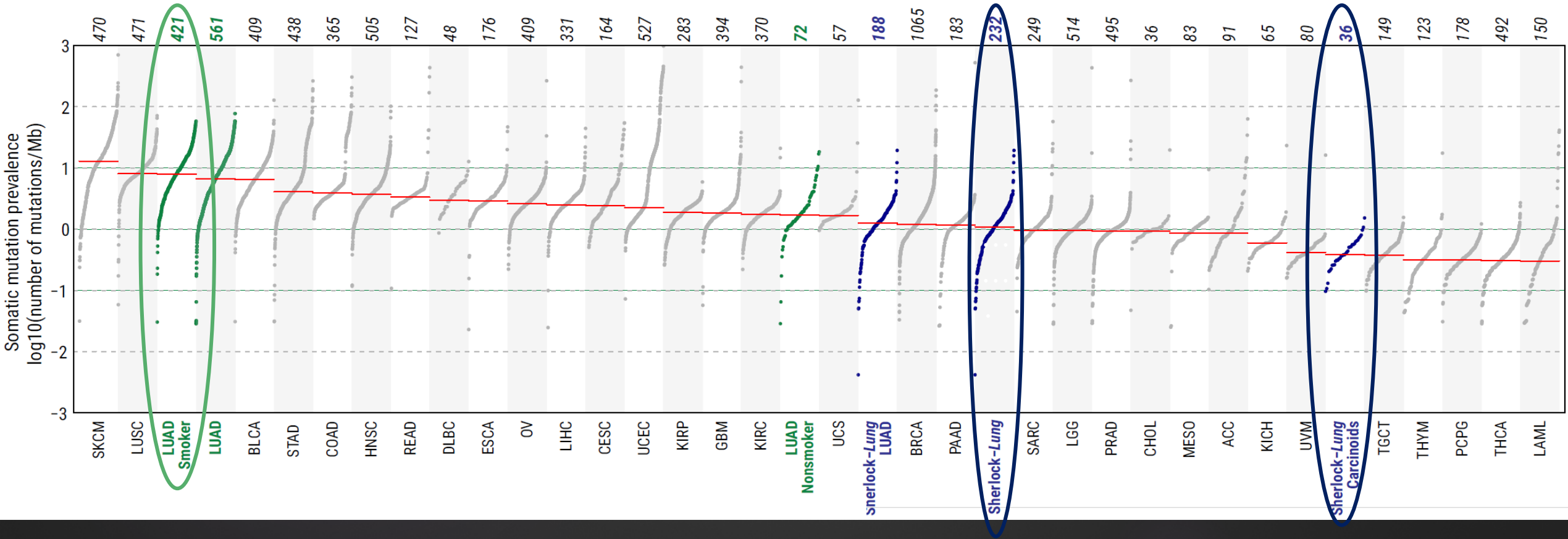


232 cases
97% EUR
189 LUAD
36 carcinoids
7 others

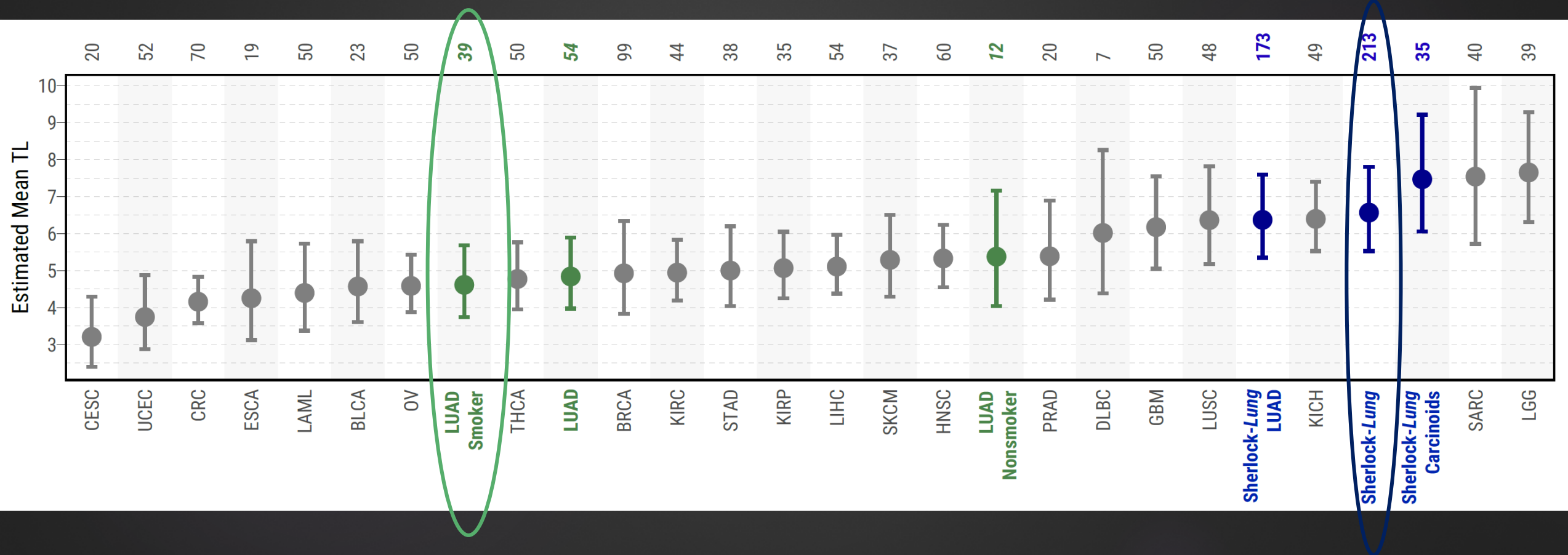
WGS
T=85x, B=35x



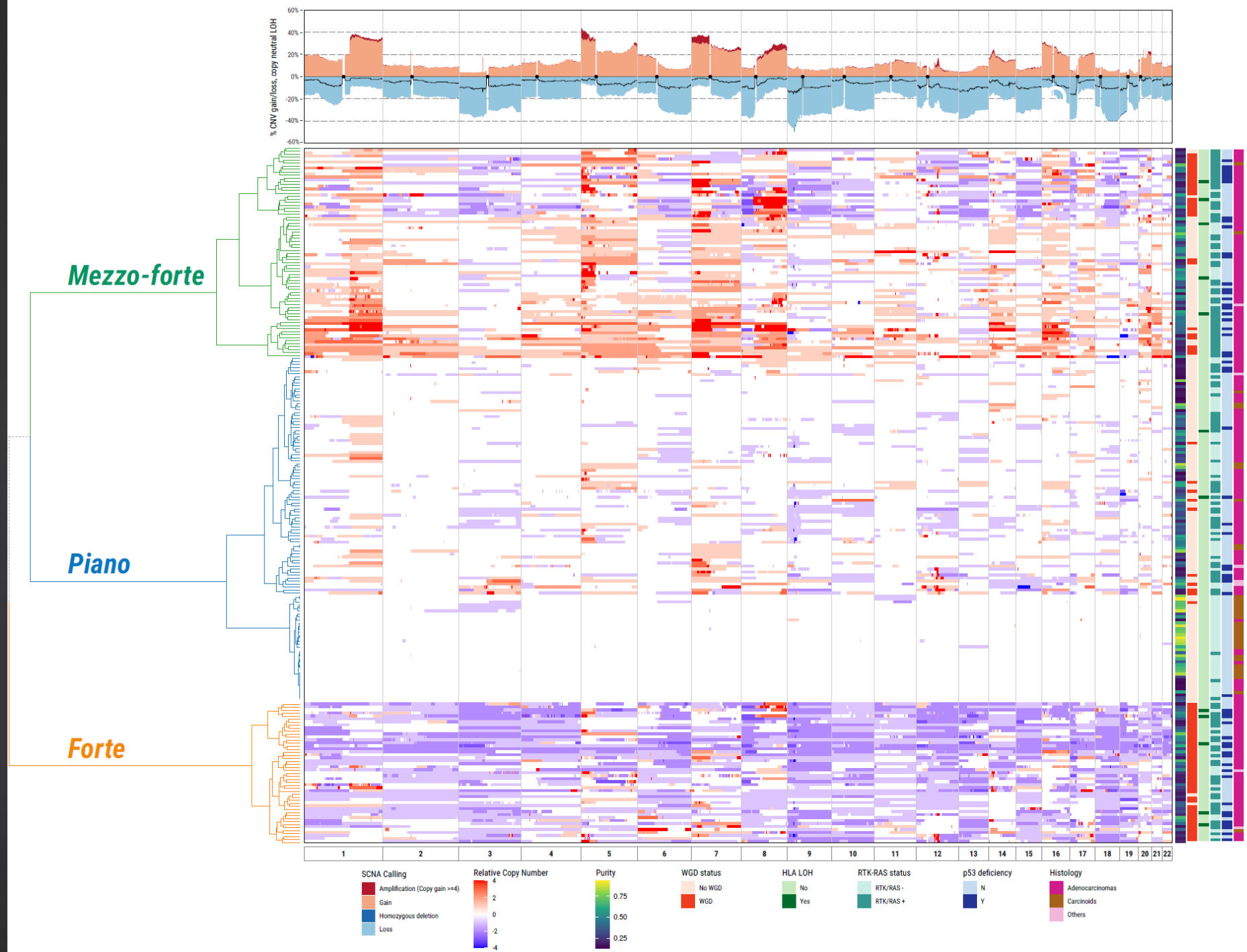
Tumor mutational burden across cancer types



Telomere length across cancer types



Landscape of copy number alterations



Driver gene mutations

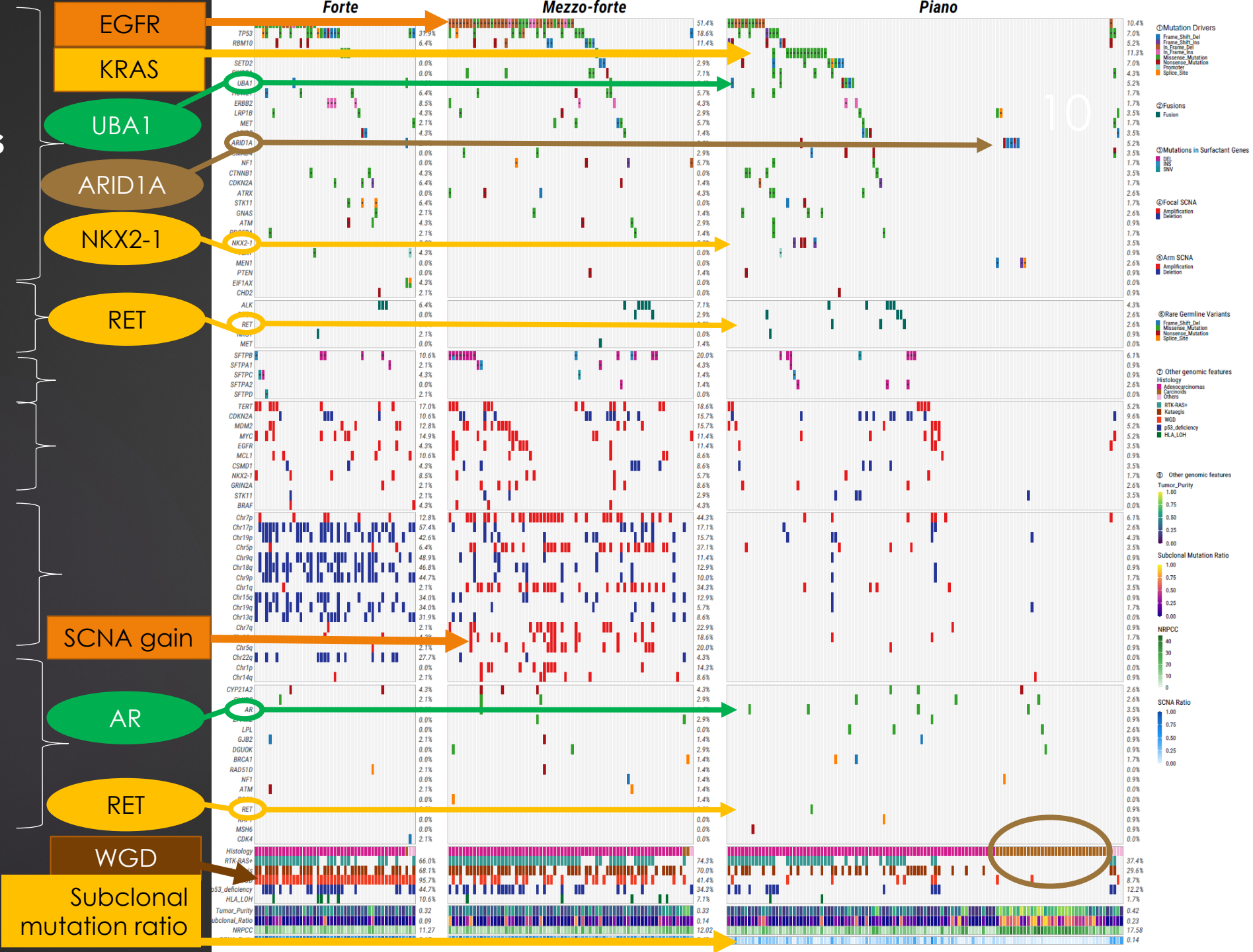
Driver gene fusions

Surfactant genes

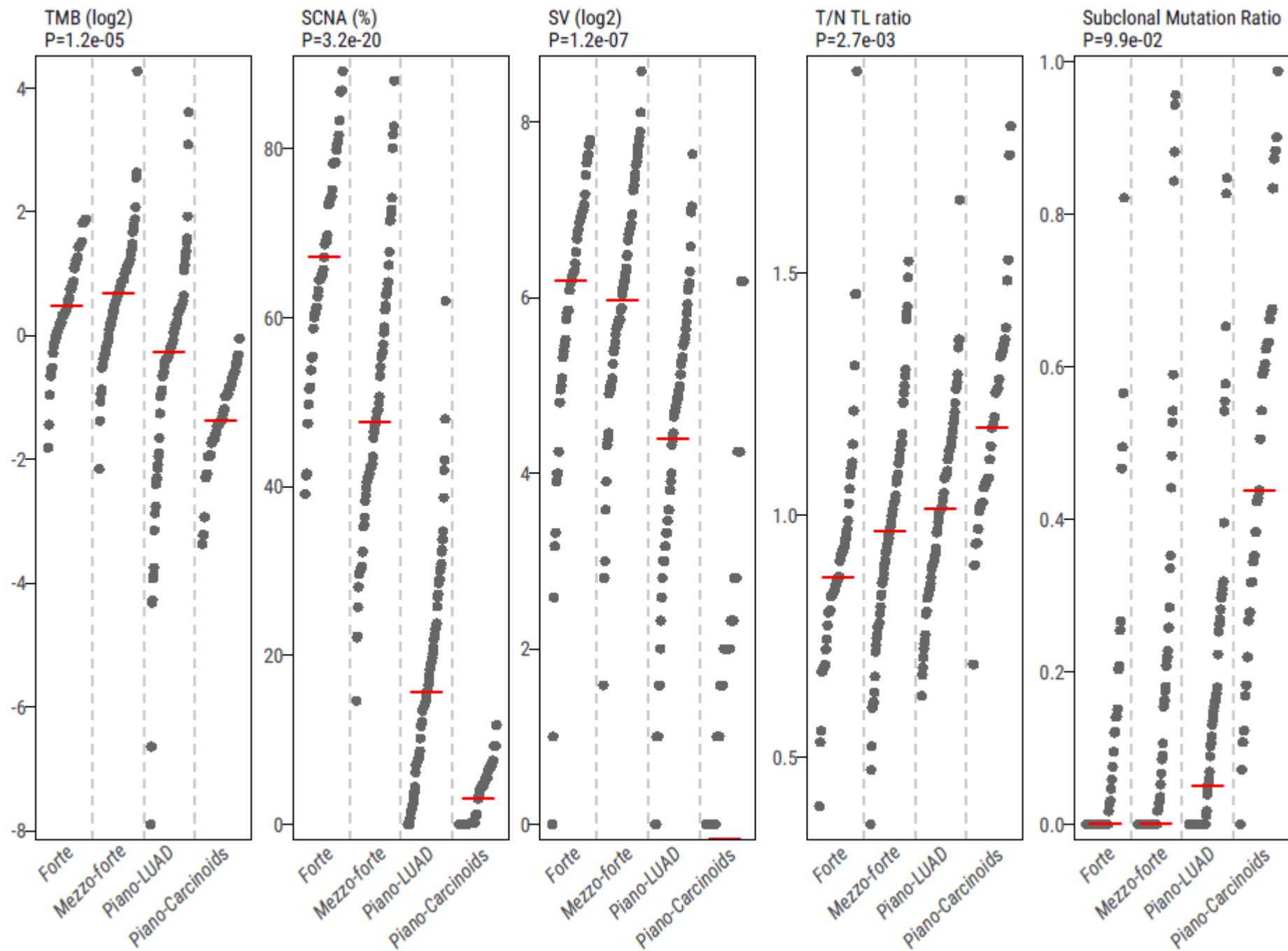
Focal SCNAs

Arm level SCNAs

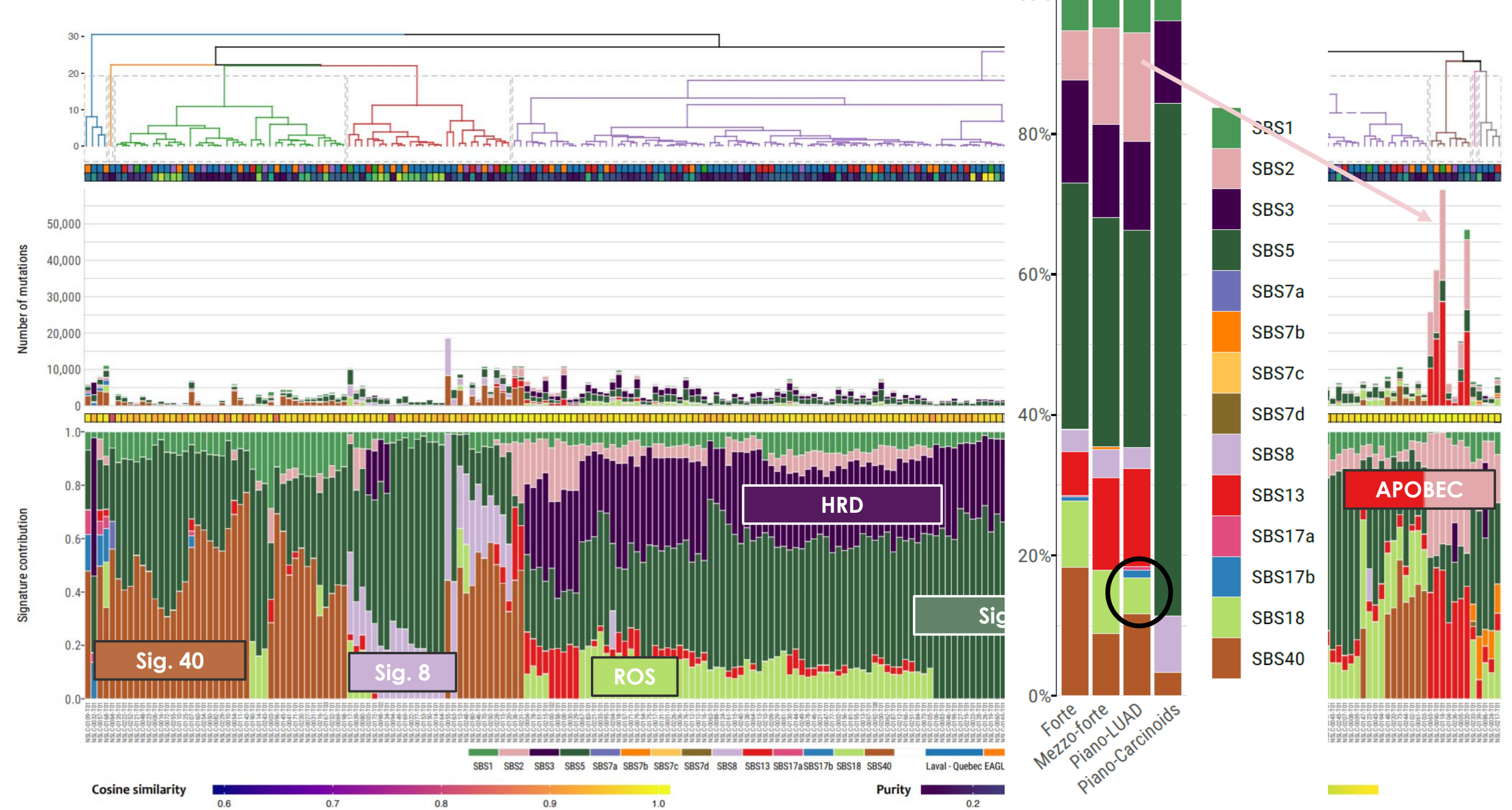
Germline variants



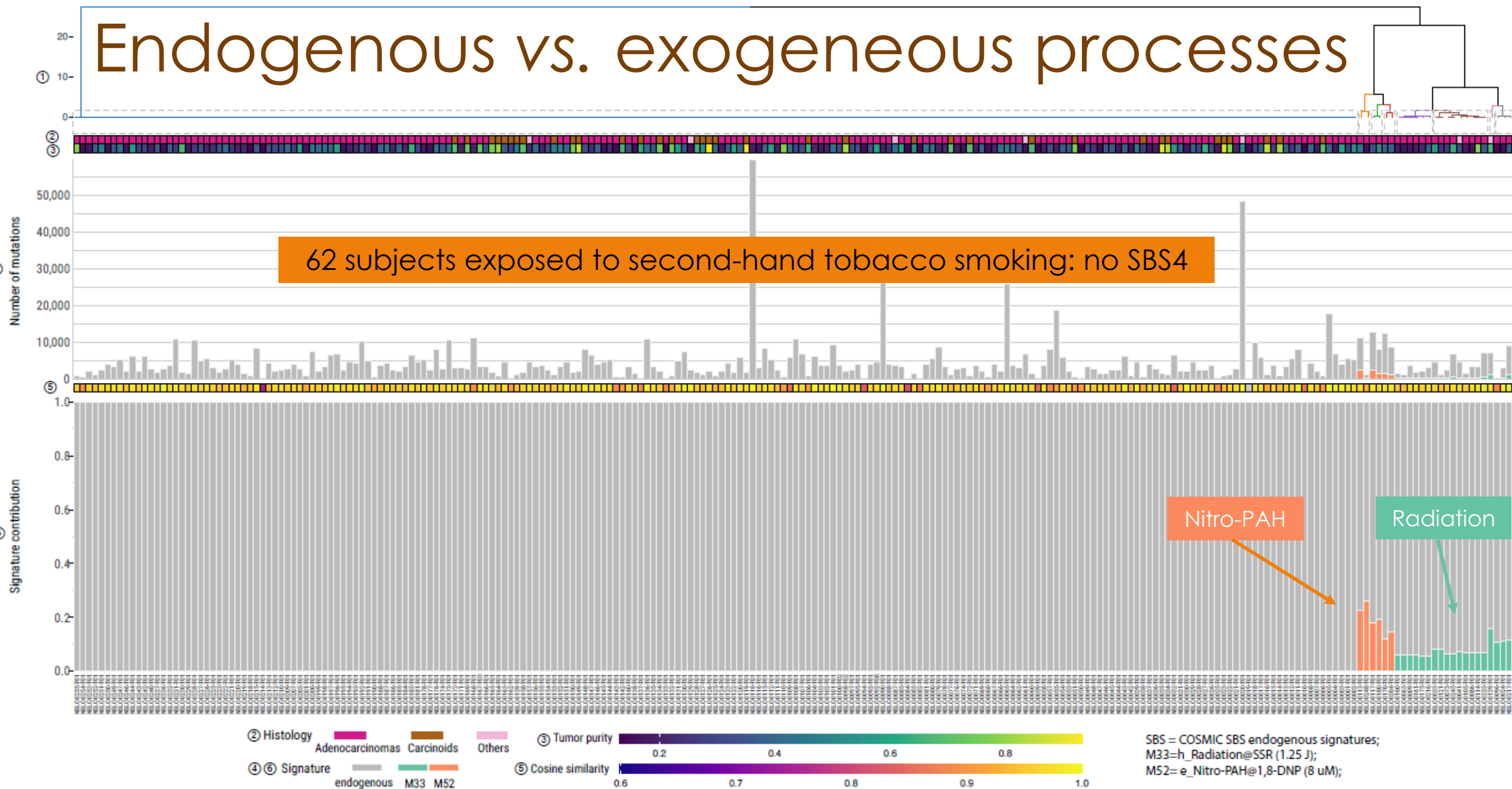
Genomic features by subtype



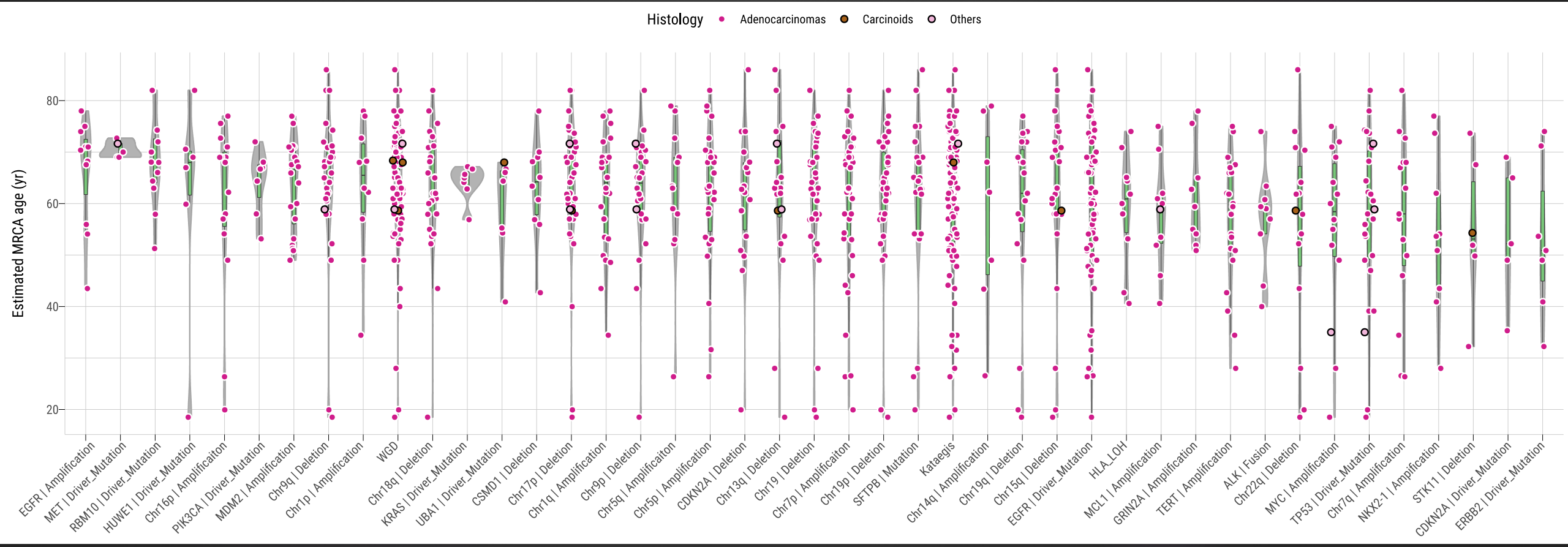
Mutational signatures



Endogenous vs. exogenous processes

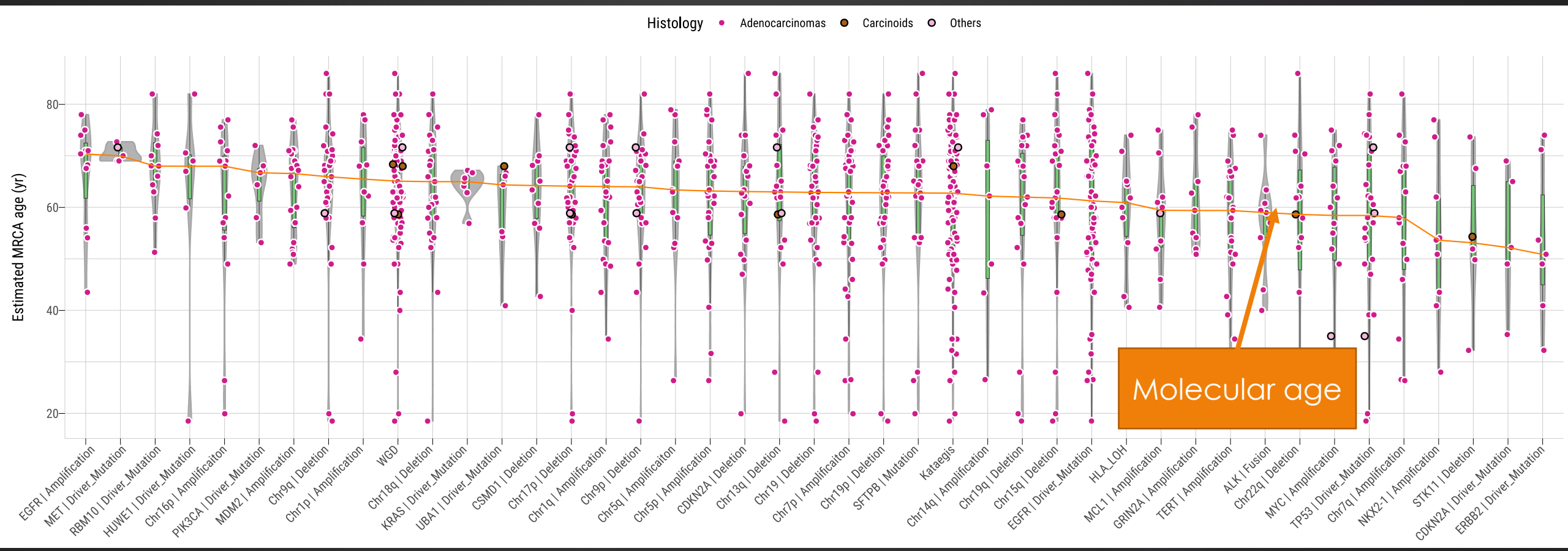


Timing tumors' evolutionary history



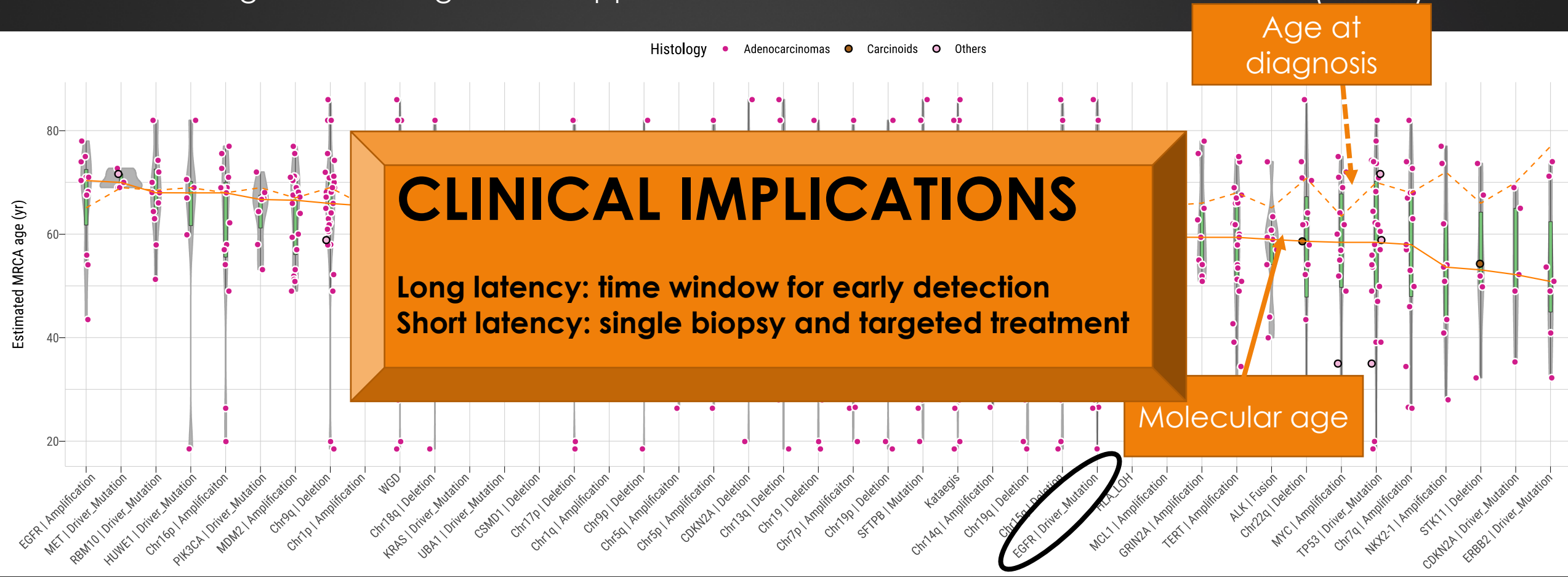
Timing tumors' evolutionary history

Molecular age: median age of the appearance of the most recent common ancestor (MRCA)



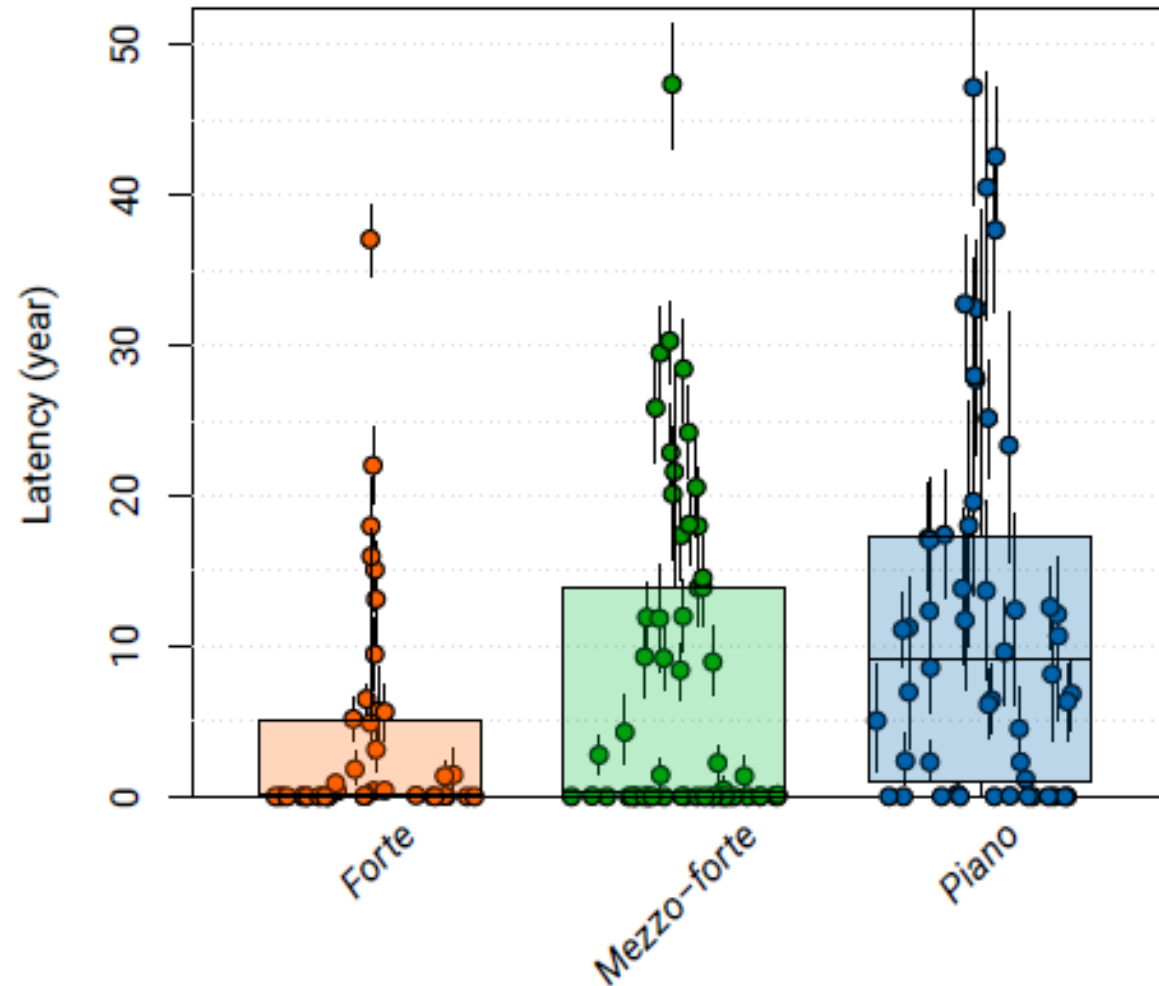
Timing tumors' evolutionary history

Molecular age: median age of the appearance of the most recent common ancestor (MRCA)



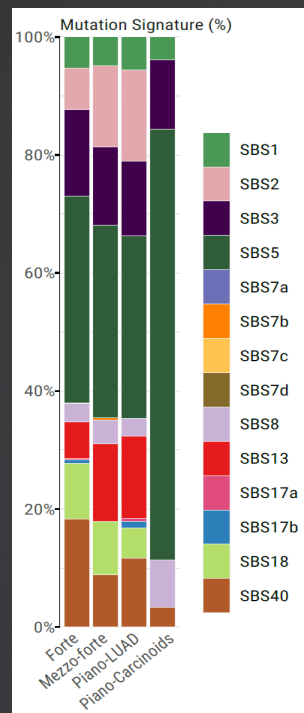
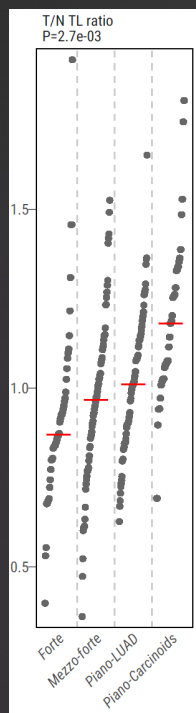
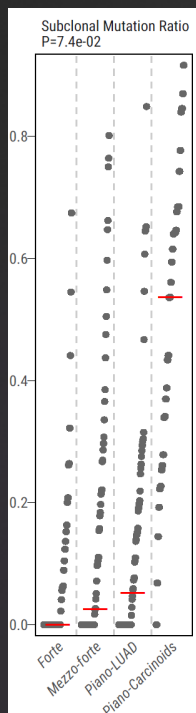
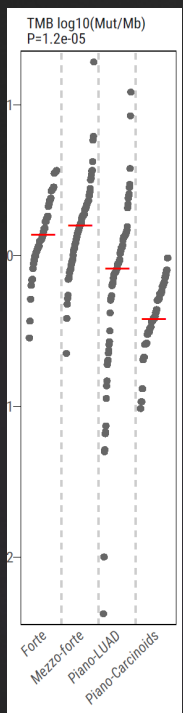
Timing tumors' evolutionary history by subtype

18

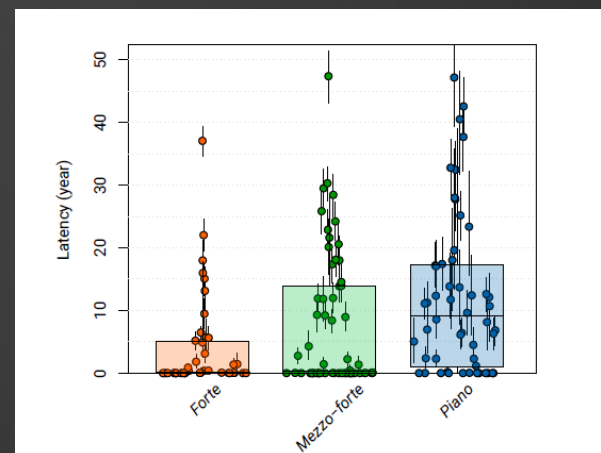


Piano tumor features vs. other subtypes

19



↑ KRAS
UBA1, RET, NKX2-1, ARID1A
only in piano
mutually exclusive



Stem cells?



Lower TMB

High ITH

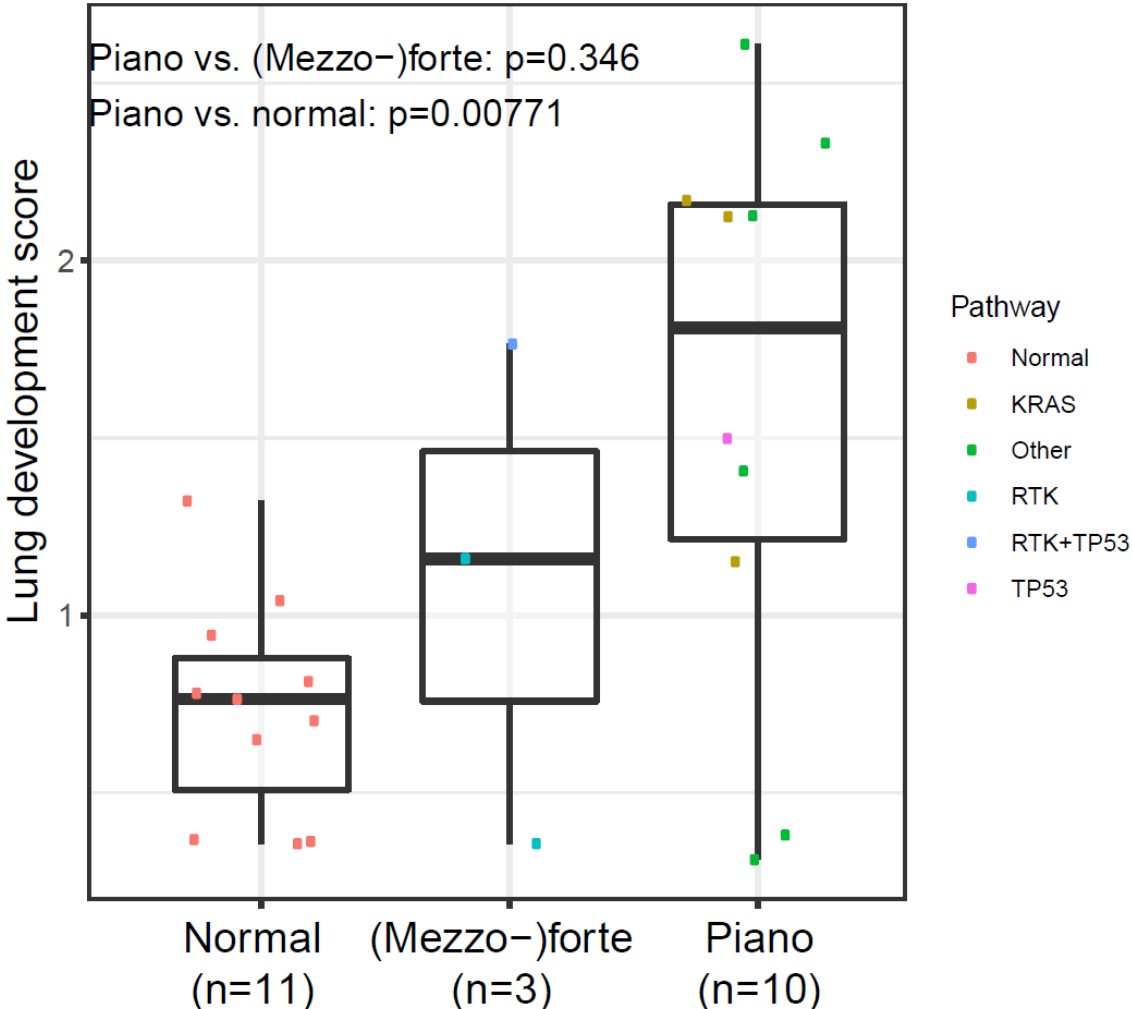
Longer telomeres

Low/absent SBS18

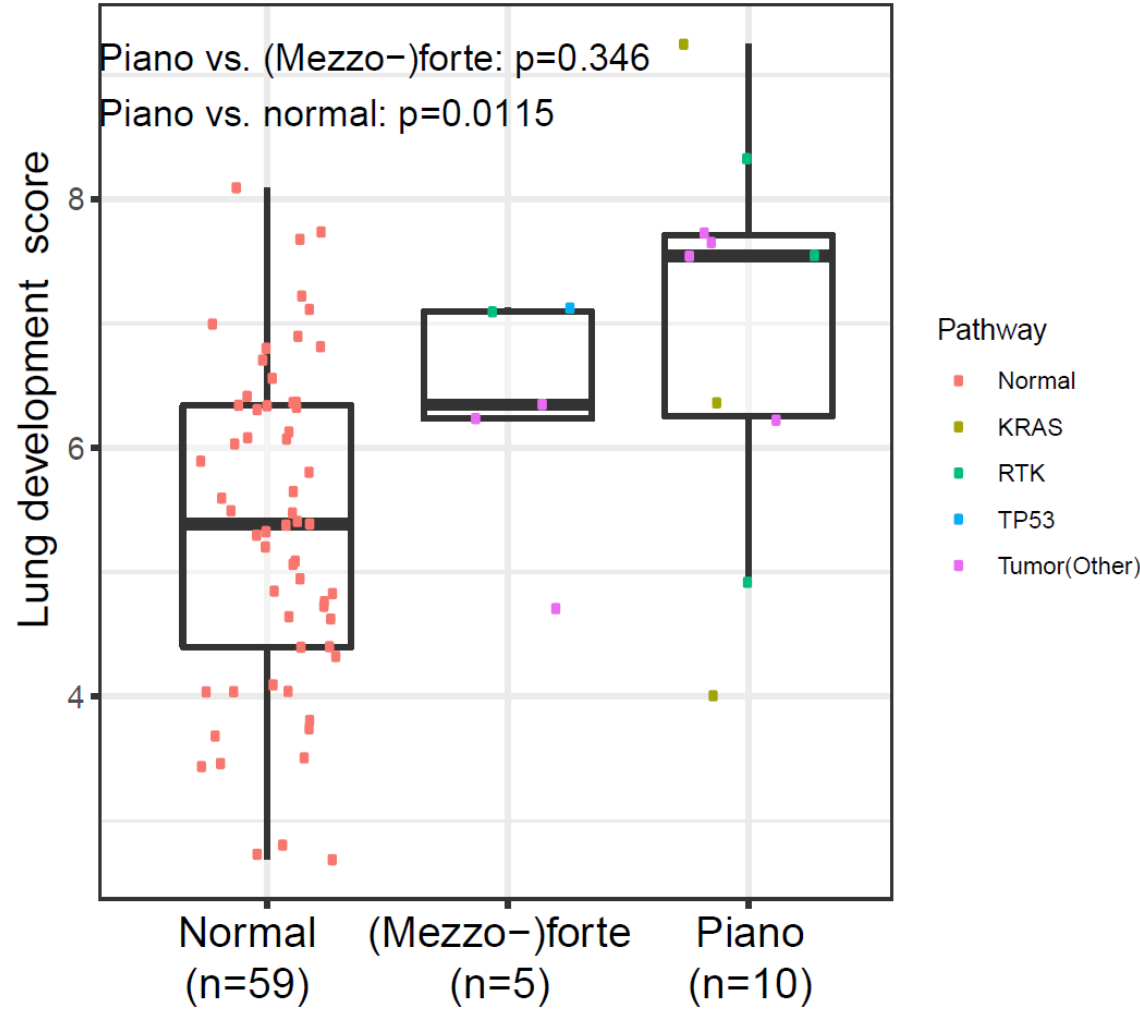
Slow growth rate

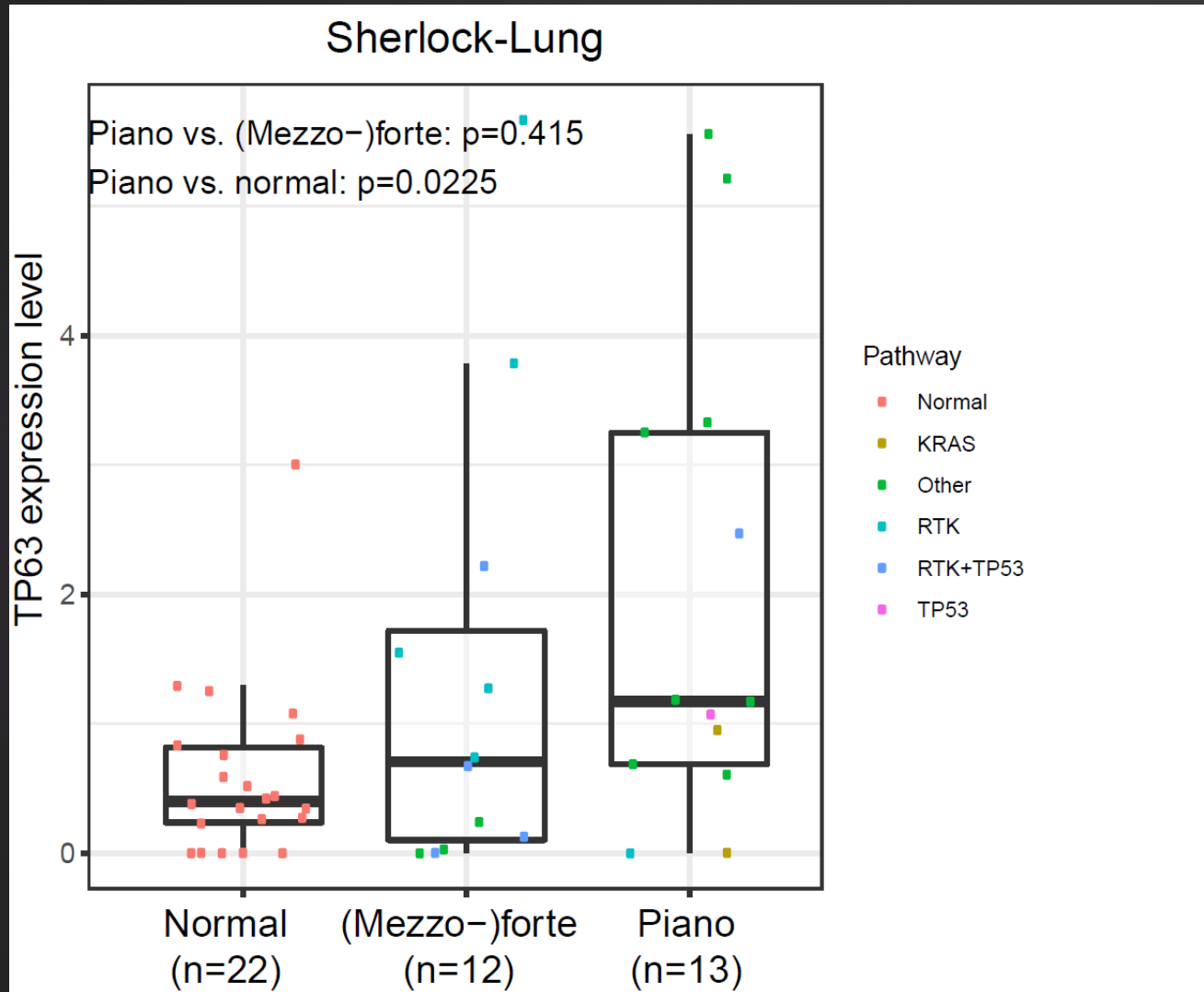
Development score: SOX2, SOX9, HMGA2

Sherlock-Lung



TCGA-LUAD

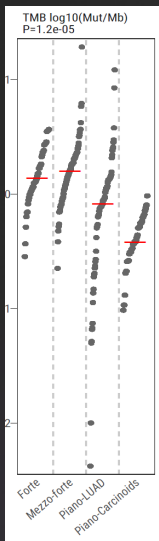




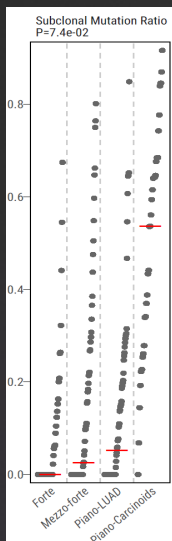
Pilot studies to verify cell-of-origin and evolutionary trajectories:

- snRNA-seq
- Single molecule DNA-seq

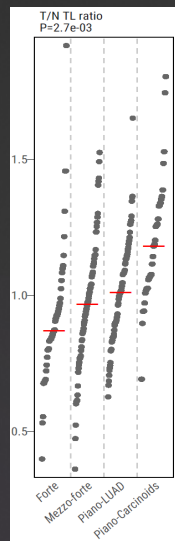
Piano tumor stem cell-like features



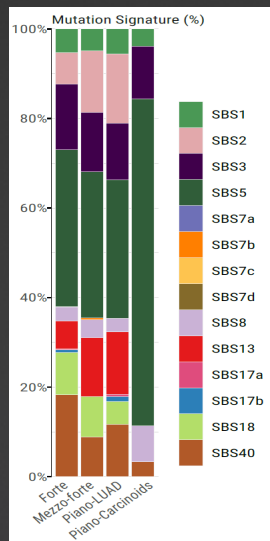
Lower TMB



High ITH

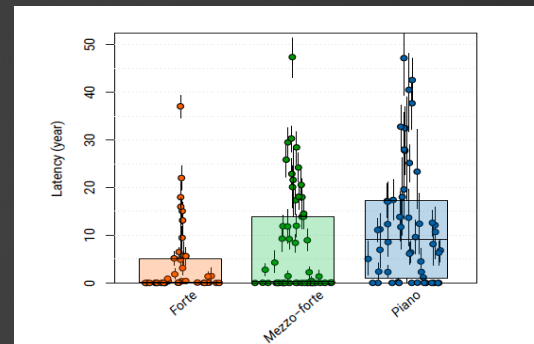


Longer telomeres



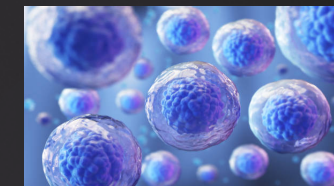
Low/absent SBS18

↑ KRAS
UBA1, RET, NKX2,
ARID1A



Slow growth rate

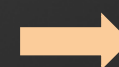
Stem cells?



Inflammation



Tissue damage



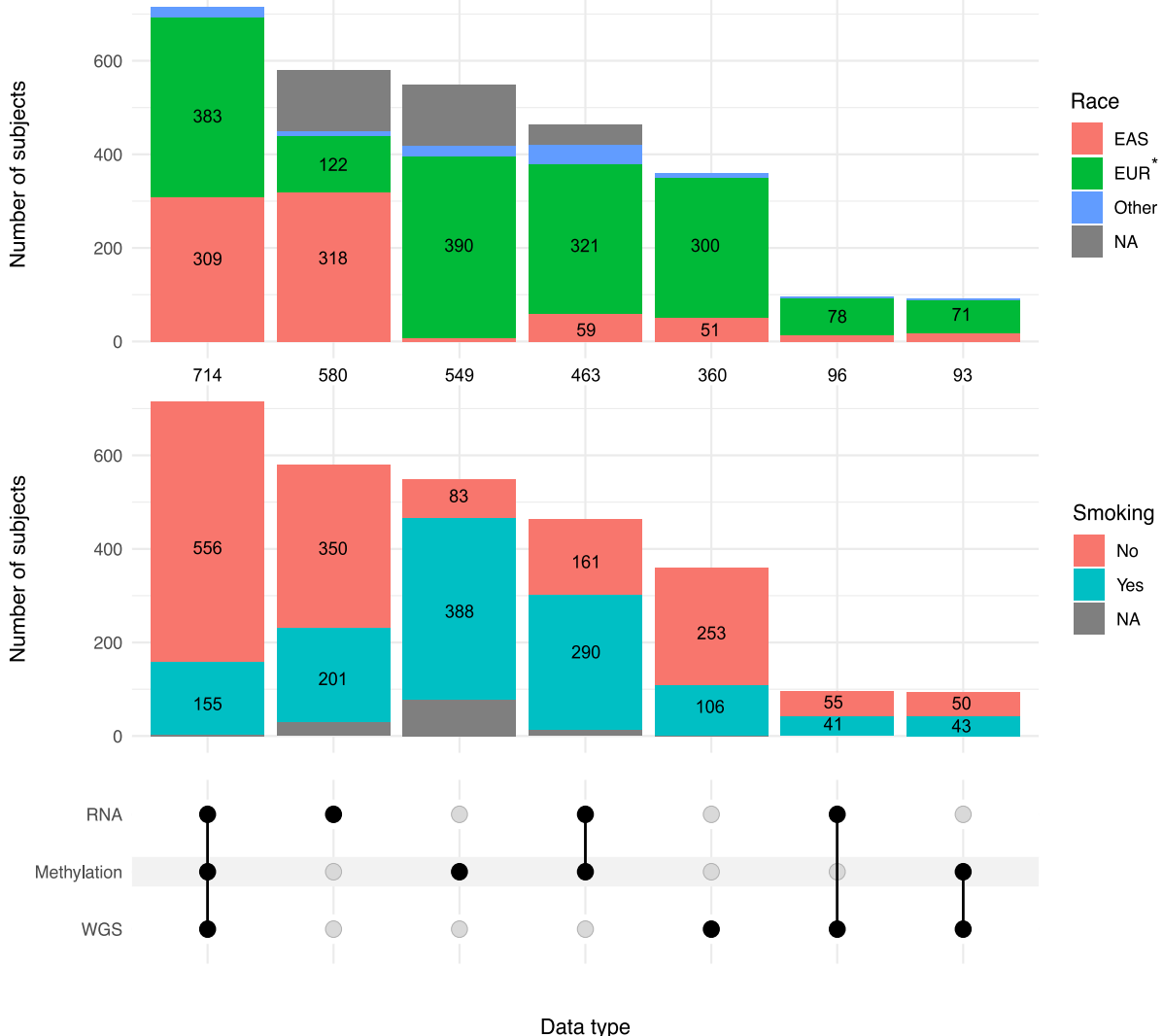
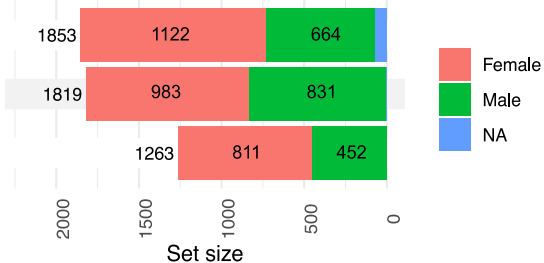
Tissue regeneration



Number of cases with T/N WGS, RNA seq and methylation by sex, ethnicity and tobacco smoking

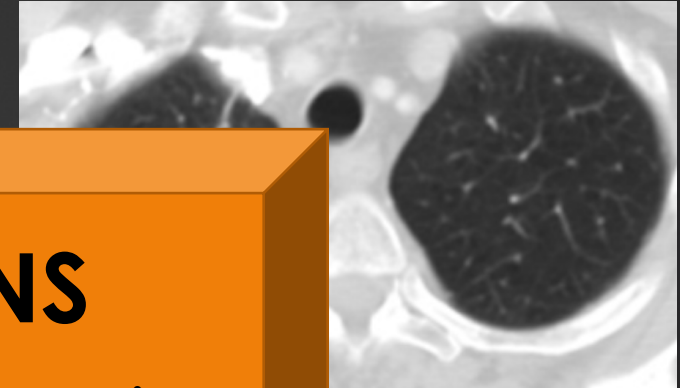
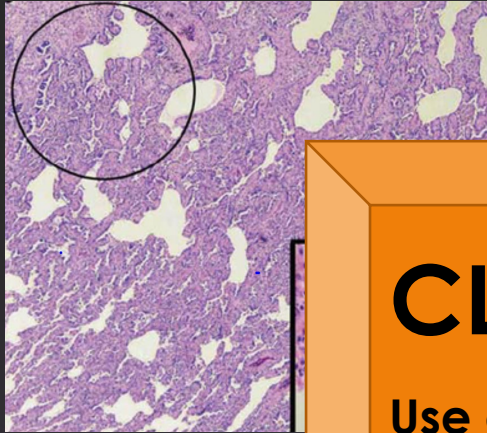
* 32 Hispanic subjects without matched WGS were included in EUR category

WGS 1,263
RNA-seq 1,853
Methylation 1,819



Aim 2: tumor classification

Molecular landscape



CLINICAL IMPLICATIONS

Use of CT-scans or biopsy slides to predict genomic features/TME of the corresponding lesions and provide crucial information for clinical decision-making

2,430 scanned
875 subjects

CT-scans
analyzed

Analyses by SCNA-subtypes, driver genes, smoking status, ethnicity, sex...

TME by H&E, immunofluorescence, RNA-seq, methylation, deep learning

Collaborators

25

DCEG major analysts: Tongwu Zhang, Wei Zhao, Phuc Hoang, Jian Sang (ITEB)

DCEG Co-Investigators:

OD: Stephen Chanock, Montse Garcia-Closas, Jonas Almeida. **ITEB:** Mustapha Abubakar, Monica D'Arcy, Mengying Li, Naoise Synnott, John McElderry, Clara Bodelon. **OEEB:** Qing Lan, Nat Rothman, Rena Jones, Neil Caporaso, Debra Silverman. **BB:** Jianxin Shi, Bin Zhu, Ruth Pfeiffer. **LTG:** Jiyeon Choi, Ludmila Prokunina. **CGB:** Jung Kim, Doug Stewart, Shashinaz Gadalla. **REB:** Amy Berrington. **CGR:** Yelena Golubeva, Belynda Hicks, Amy Hutchinson, Scott Lawrence, Petra Lenz

Epidemiological collaborators:

Christine Ambrosone, **Roswell Park Cancer Center**; Chris Amos, **Baylor University**; Yohan Bosse', **Laval University**; Paul Brennan, **IARC**; David Christiani, **Harvard University**; Dario Consonni, Angela Pesatori, **University of Milan**, Italy; Paul Hoffman, **University of Nice**, France; Chao Agnes Hsiung, **National Health Research Institutes, Zhunan**, Taiwan; Geoff Liu, **University of Toronto**, Canada; Tobias Peikert, **Mayo Clinic**; Bonnie Rothberg, **Yale University**; Matthew Schabath, **Moffitt Cancer Center**; Maria Wong, **University of Hong Kong**

Analytical collaborators:

Ludmil Alexandrov, **UCSD**; Brian Bartholmai, **Mayo Clinic**; Peter Campbell, **Sanger Institute, UK**; Chongyi Chen, **CCR, NCI**; Hannah Carter, **UCSD**; Gaddy Getz, **Broad Institute**; Dmitry Gordenin, **NIEHS**; Rob Homer, **Yale University**; Marcin Imielinski, **Cornell University**; Curtis Harris, **CCR, NCI**; Carla Kim, **Harvard University**; Nuria Lopez-Bigas, **Inst. Res. Biomedicine, Barcelona, Spain**; David Wedge, **U. Oxford/Manchester, UK**; Lixing Yang, **U. Chicago**

Advisory Group:

Matthew Meyerson, **Broad Institute**
Jon Samet, **Colorado School Public Health**
Margaret Spitz, **Baylor University**
Ronald Summers, **NIH Clinical Center**
Michael Thun, **American Cancer Society**
Bill Travis, **Memorial Sloan Kettering CC**