

Ancestry-associated differences in somatic mutation rates from tumor profiling data of a pan-cancer cohort of 100,000 patients

Brooke Rhead¹, Yannick Pouliot¹, Justin Guinney¹, Francisco M. De La Vega¹

¹Tempus Labs, Chicago, IL

INTRODUCTION

Racial and ethnic disparities span the continuum of cancer care and adversely impact screening, detection, diagnosis, treatment, and outcomes of cancer. These disparities are driven by a complex interplay among social, psychosocial, lifestyle, environmental, healthcare, and biological determinants of health.

A challenge is that the social constructs of race and ethnicity may not fully capture shared genetic background. Genomics can capture ancestry in a more precise way, allowing genetic influences to be teased apart from the impact of social and environmental factors.

Here, we searched for associations between continental genetic ancestry and somatic mutations in cancer genes in a pan-cancer cohort of 100,000 patients.

METHODS

We inferred genetic ancestry from 106,924 de-identified records from cancer patients of diverse histology who underwent tumor genomic profiling with the 648-gene Tempus xT next-generation sequencing (NGS) assay. We used 654 ancestry-informative markers selected to overlap the target regions of the assay to infer global ancestry proportions at the continental level: Africa (AFR), Americas (AMR), Europe (EUR), East Asia (EAS), and South Asia (SAS).

Inclusion Criteria:

- Cancer types with at least 500 patients
- Genes for which at least 1% of patients, and minimum 10, harbored a somatic mutation
- Patient records with known gender and age

Statistical Analysis:

- Likelihood ratio (LR) tests of logistic regression models were used to test for associations between continental ancestry proportions and presence of somatic mutations, controlling each ancestry association for assay version, gender, age, and the other 4 ancestries.
- LR test p-values were adjusted for multiple testing by the Benjamini-Hochberg method to control the false discovery rate at 5%.
- Different sets of somatic mutations were tested: (1) short nonsynonymous variants, and (2) cancer driver mutations predicted by the boostDM¹ algorithm.

References

1. Muiños, F., Martínez-Jiménez, F., Pich, O., Gonzalez-Perez, A. & Lopez-Bigas, N. *In silico saturation mutagenesis of cancer genes*. Nature 596, 428–432 (2021).

RESULTS

Ancestry associations with SNV/indels and nonsynonymous somatic mutations in cancer genes

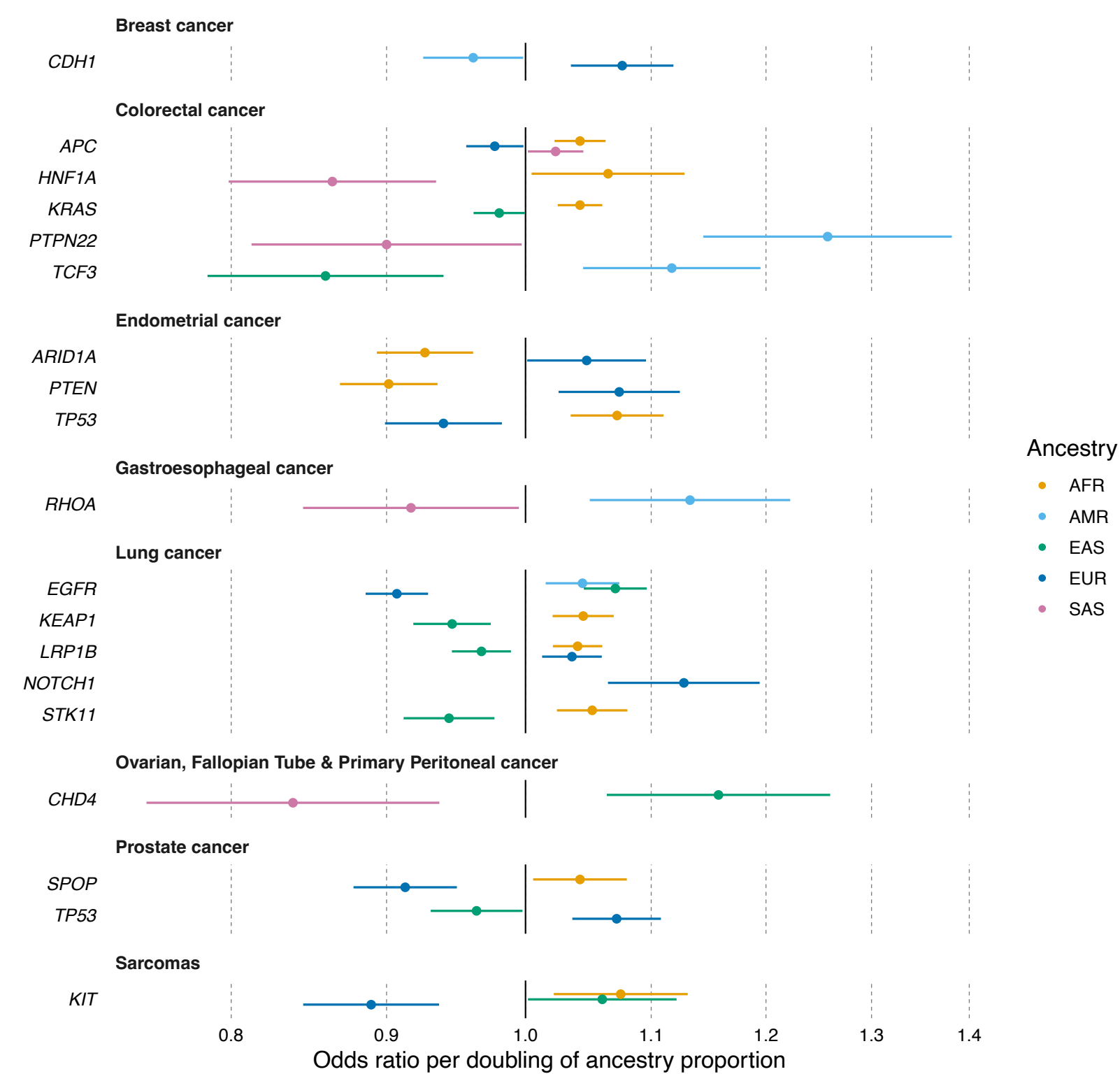


Figure 1. We performed association tests between proportions of each continental ancestry and presence of somatic mutations (SNVs and small indels) for the cancer genes and cancer types that met our criteria (see Methods).

- The forest plots show the increased/decreased odds of having a variant in a gene for every doubling in ancestry proportion.
- Some of our analyses replicate previous findings (e.g., increased somatic mutations in *EGFR* associated with East Asian and Native American ancestry in lung cancer), but not others.
- Furthermore, we find several new associations with both small variants and actionable variants across several cancer types.

SUMMARY

- Genetic ancestry inference by ancestry-informative markers in tumor profiling data permits the direct study of the **association between ancestry and somatic mutations** and overcomes lack and ambiguity of race/ethnicity labels.
- We identified several **associations** between **continental ancestry** and presence of **somatic mutations** in cancer genes, replicating previous observations and adding new findings that warrant further study.

Ancestry associations with predicted cancer driver somatic mutations in cancer genes

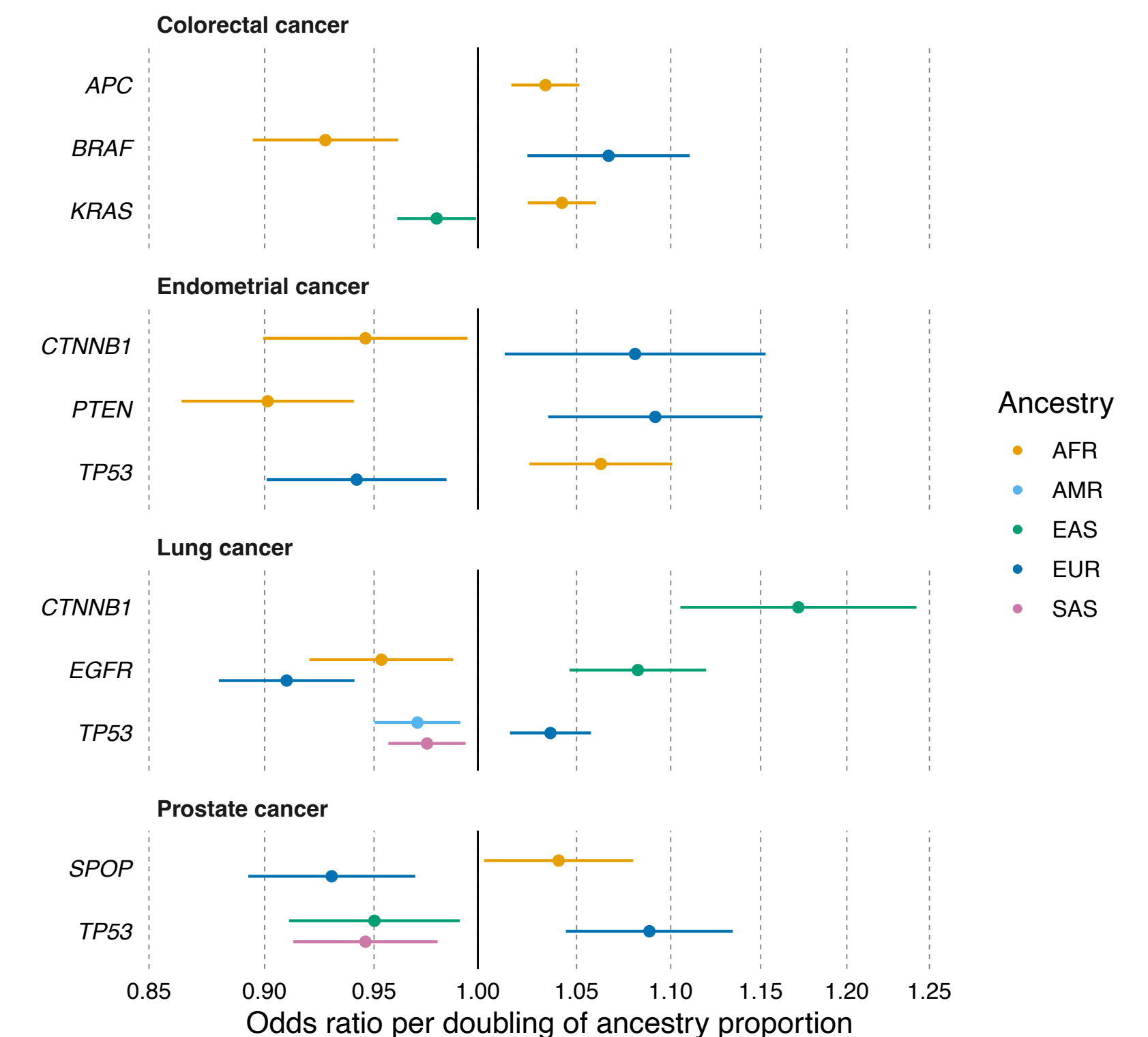


Figure 2. We performed association tests between proportions of each continental ancestry and presence of predicted driver somatic mutations for the cancer genes and cancer types that met our criteria (see Methods).

- The association between EAS and *CTNNB1* in lung cancer is noteworthy since activating mutations in this gene have been suggested as biomarkers for TTK inhibitor effectiveness.
- We found that 7% of lung patients of East Asian ancestry in our cohort carry driver mutations in this gene as compared to <1.7% in other groups of patients.