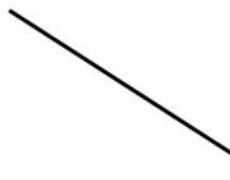


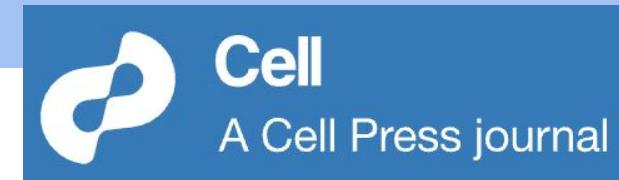
# Genomics and Next-Generation Sequencing: Project Group 5

Aishwarya, Aman, Asta  
18.01.26

NOTE: formatted slightly differently to enable reading content on slides with animations



# Original paper



## Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma

Willy Hugo,<sup>1,6,9</sup> Jesse M. Zaretsky,<sup>2,6,9</sup> Lu Sun,<sup>1,6</sup> Chunying Song,<sup>1,6</sup> Blanca Homet Moreno,<sup>3</sup> Siwen Hu-Lieskovian,<sup>3</sup> Beata Berent-Maoz,<sup>3</sup> Jia Pang,<sup>3</sup> Bartosz Chmielowski,<sup>3</sup> Grace Cherry,<sup>3</sup> Elizabeth Seja,<sup>3</sup> Shirley Lomeli,<sup>1,6</sup> Xiangju Kong,<sup>1,6</sup> Mark C. Kelley,<sup>7</sup> Jeffrey A. Sosman,<sup>8</sup> Douglas B. Johnson,<sup>8</sup> Antoni Ribas,<sup>2,3,4,5,6</sup> and Roger S. Lo<sup>1,2,5,6,\*</sup>

<sup>1</sup>Division of Dermatology, Department of Medicine

<sup>2</sup>Department of Molecular and Medical Pharmacology

<sup>3</sup>Division of Hematology & Oncology, Department of Medicine

<sup>4</sup>Division of Surgical Oncology, Department of Surgery

<sup>5</sup>Jonsson Comprehensive Cancer Center

<sup>6</sup>David Geffen School of Medicine

University of California, Los Angeles, CA 90095-1662, USA

<sup>7</sup>Department of Surgery, Vanderbilt-Ingram Cancer Center, Nashville, TN 37232, USA

<sup>8</sup>Department of Medicine, Vanderbilt-Ingram Cancer Center, Nashville, TN 37232, USA

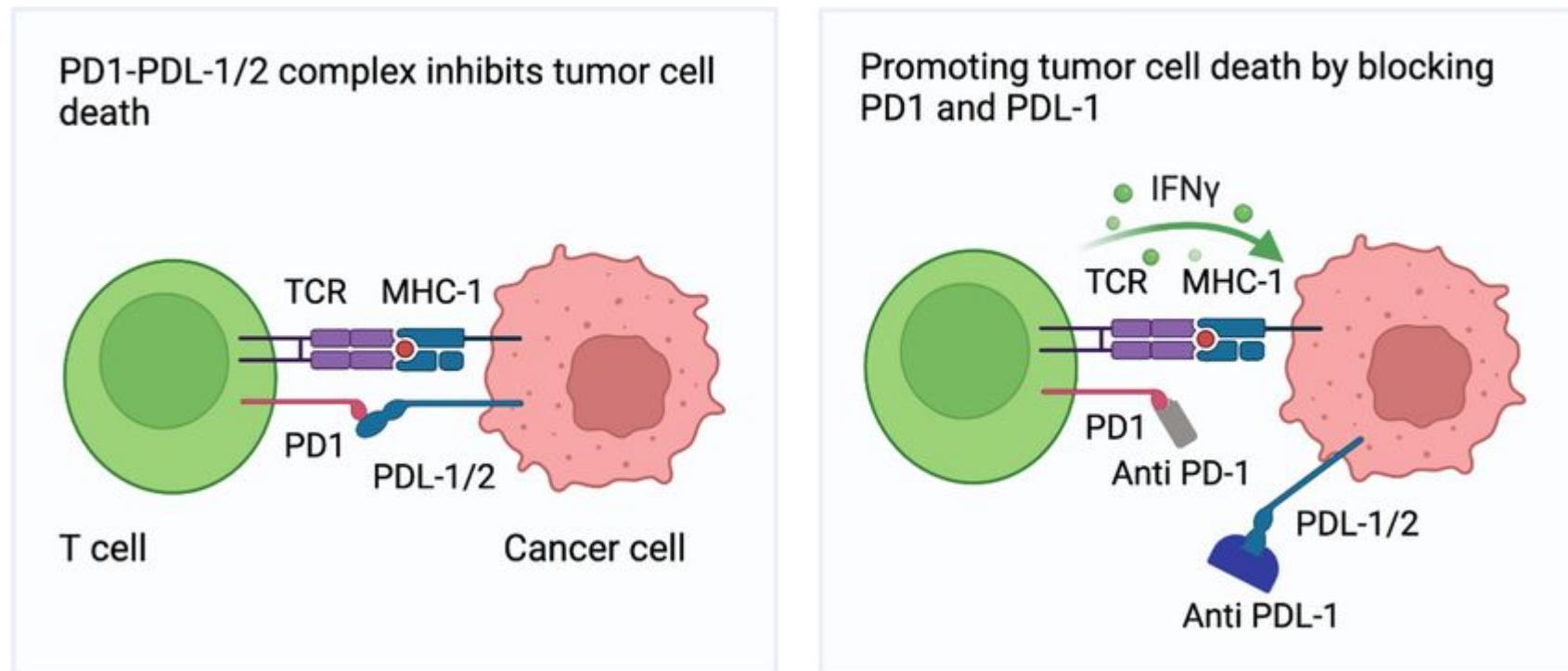
<sup>9</sup>Co-first authors

\*Correspondence: rlo@mednet.ucla.edu

<http://dx.doi.org/10.1016/j.cell.2016.02.065>



# Investigating the impact of genomic and transcriptomic features on anti-PD1 treatment success



# Research question

**What factor impact the success of anti-PD1 treatment in melanoma?**

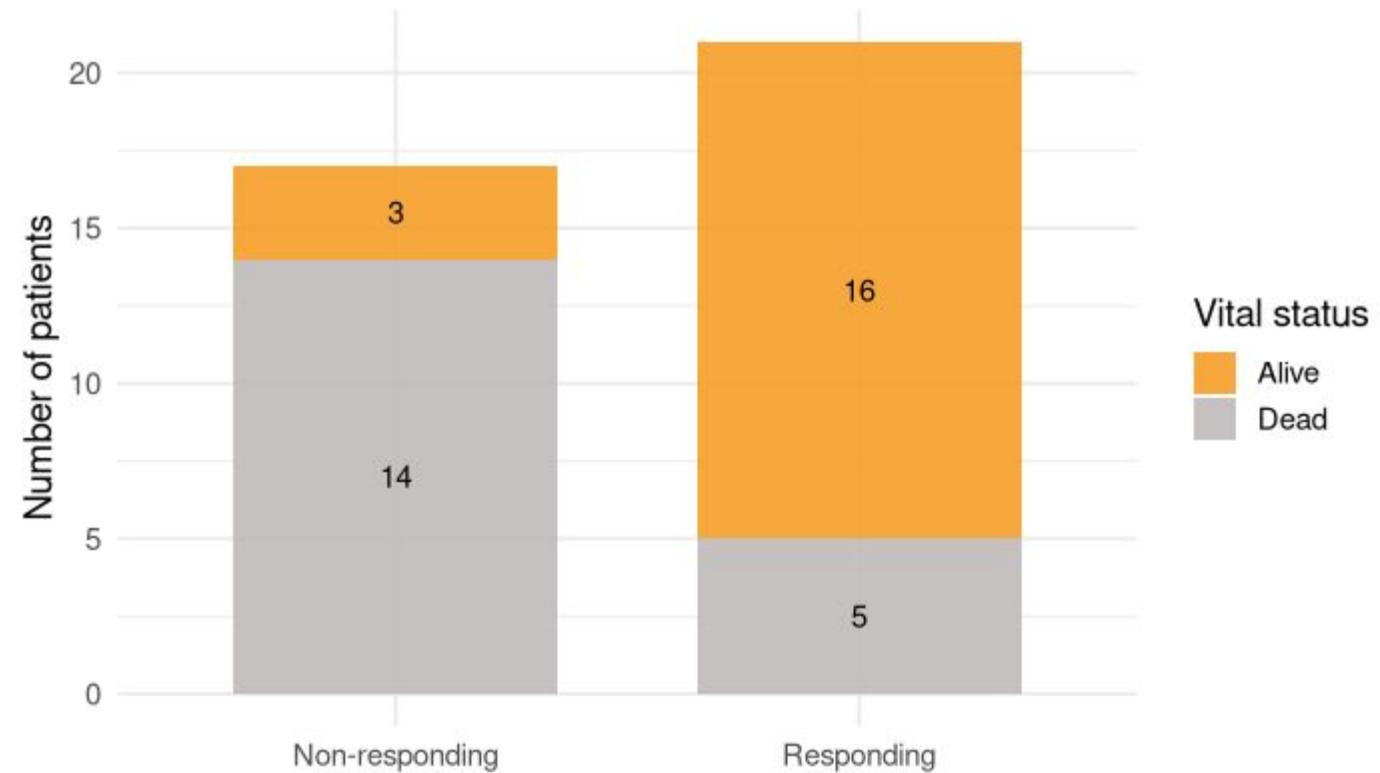
- Identifying immune-related biomarkers & signatures that distinguish responders vs non-responders (to PD-1 inhibitor) from WES and transcriptomics



# Dataset

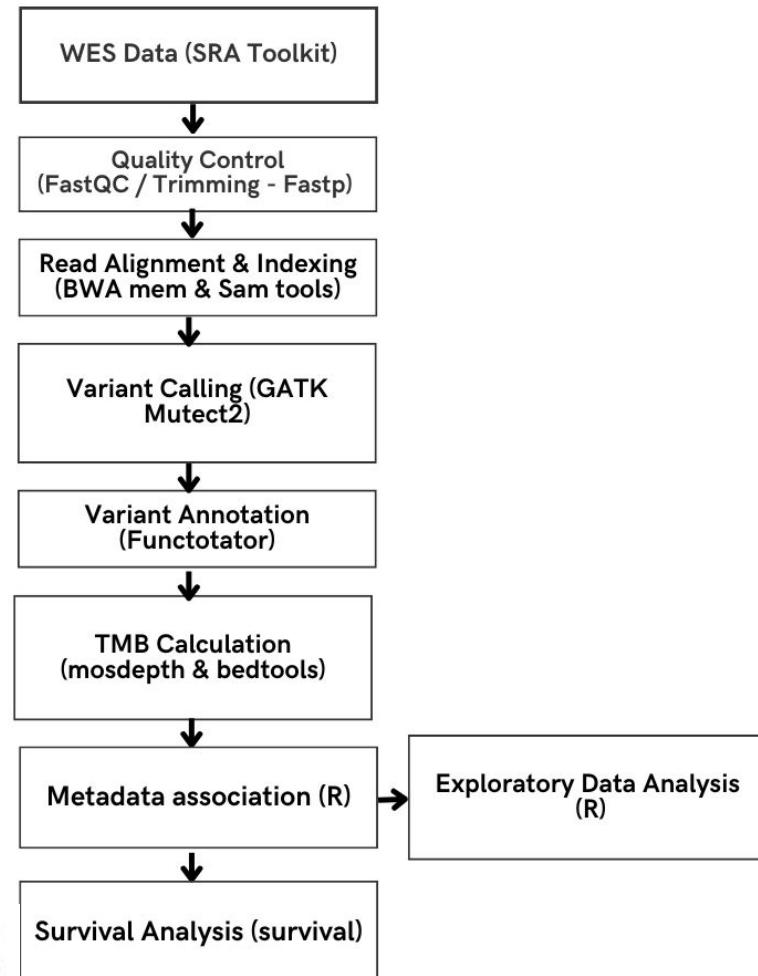
- 38 WES samples
- 28 **matched** with RNAseq samples.
- **GEO:** GSE78220 (RNA Seq Data)
- **SRA:** SRP067938 (UCLA samples)  
and SRP090294 (Vanderbilt samples)

Cohort overview: response and survival

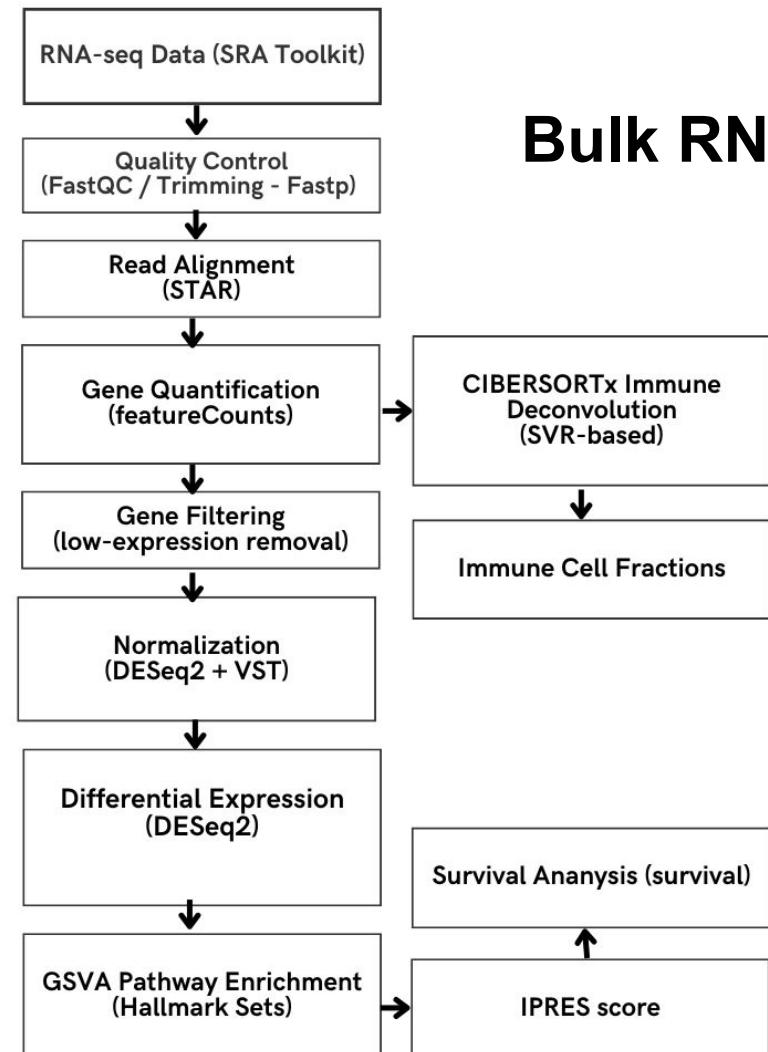


# Analysis Method (Computational Infrastructure)

**WES**



**Bulk RNAseq**



# RESULTS - WES

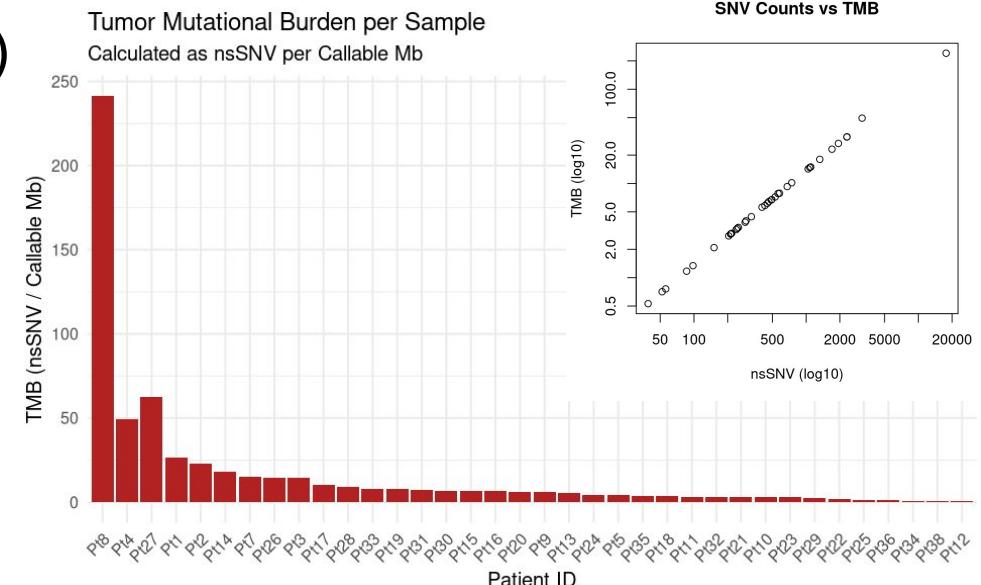
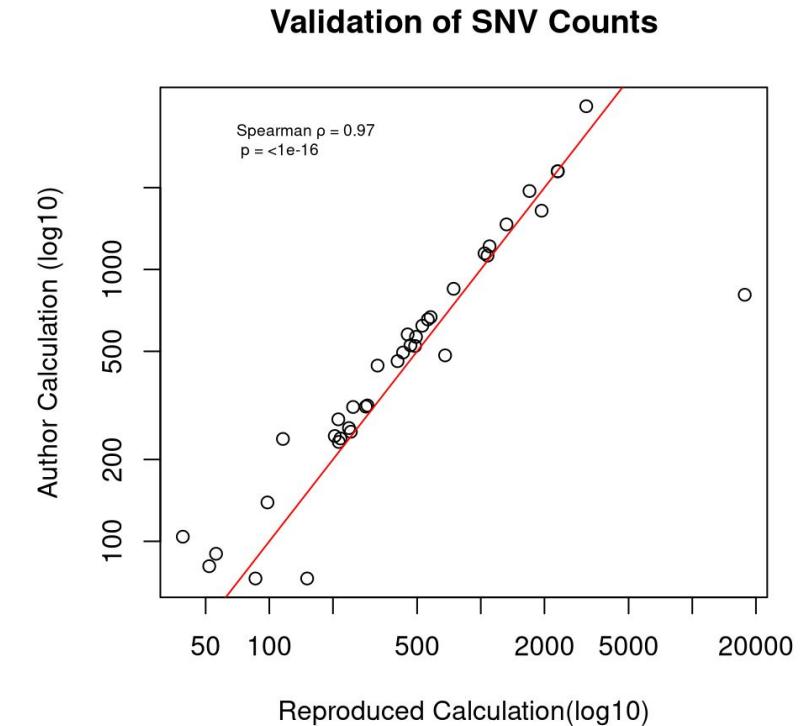


UNIVERSITY OF  
BIRMINGHAM | DUBAI  
دبي

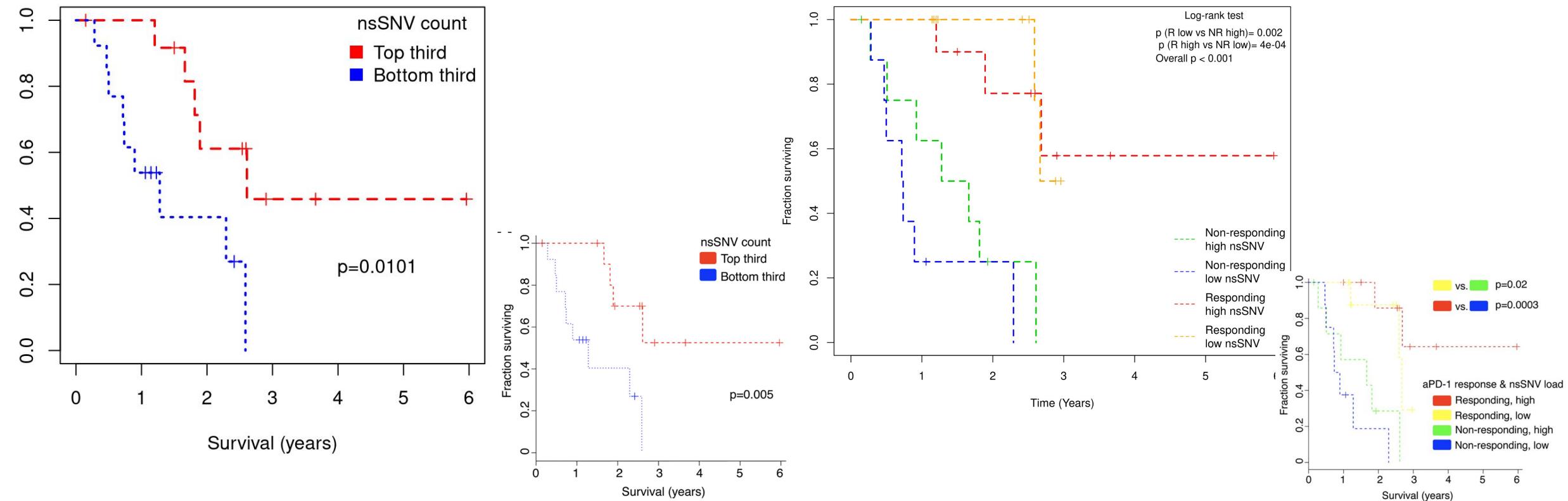
# Results - WES

## Investigation of **non-synonymous SNVs** (nsSNV)

- counts in nsSNVs correlate well with the counts found by the authors
- one outlier in our data → excluded it
- Calculated TMB (non-silent variants/callable bases) → high correlation with SNV counts.



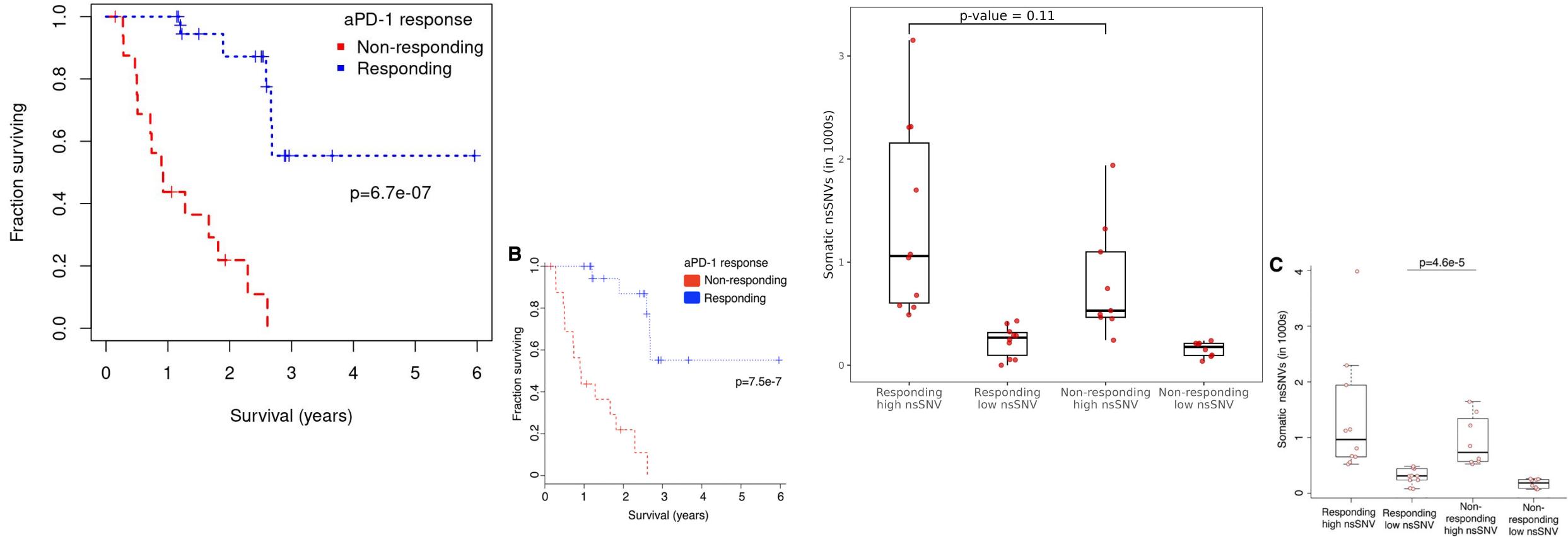
# Is higher mutational burden associated with better survival?



Patients with responding tumors and low mutational loads still outlived those with non-responding tumors and high mutational loads.  
→ Survival is not driven ONLY by mutational load.



# Is the mutational load a predictor for response to aPD-1 treatment?

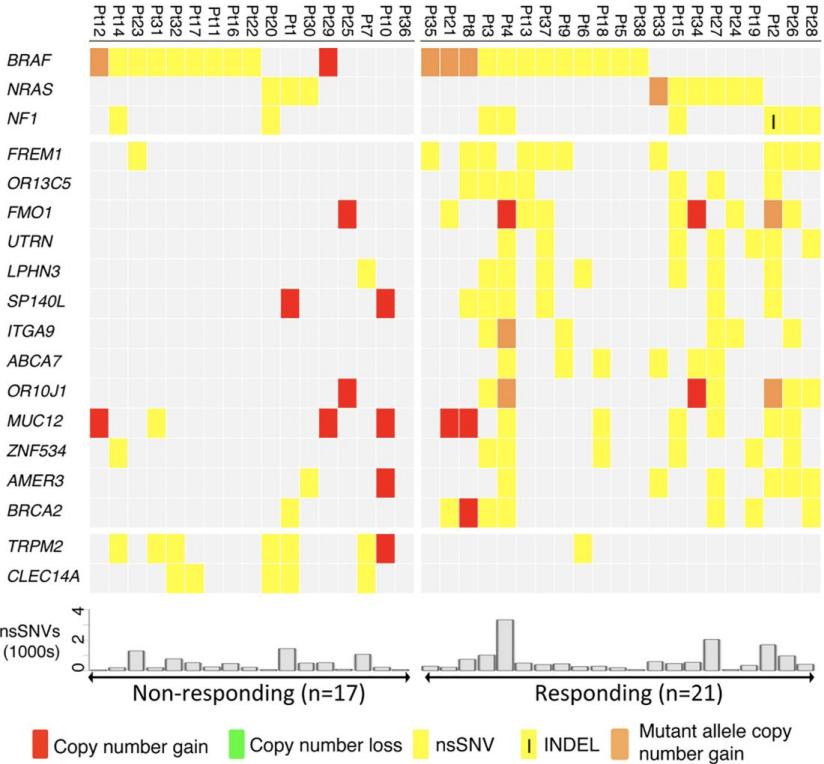
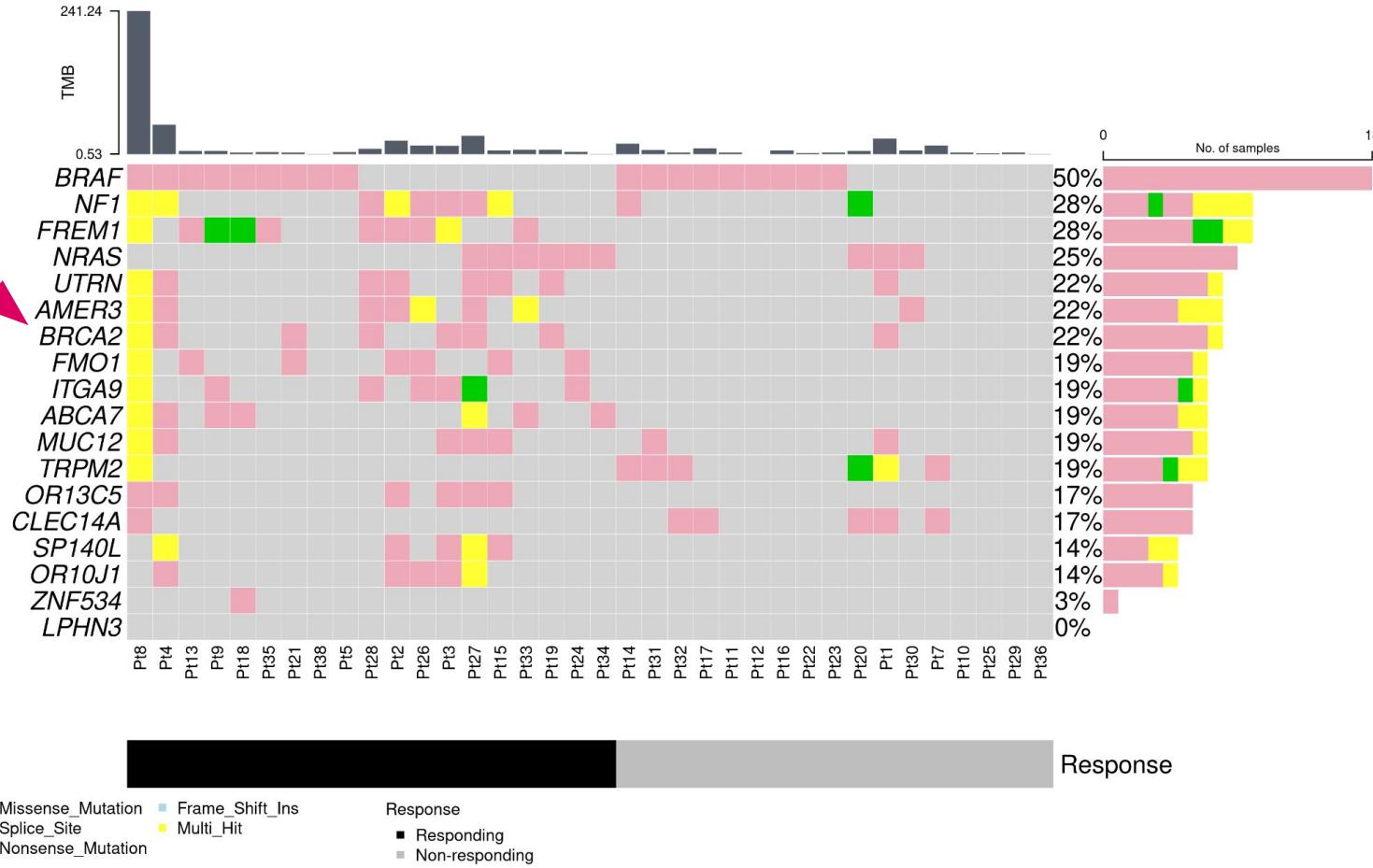


→ Mutational load is **not** a reliable predictor for the response to treatment.



# Recurrent variants in responding versus non-responding patients

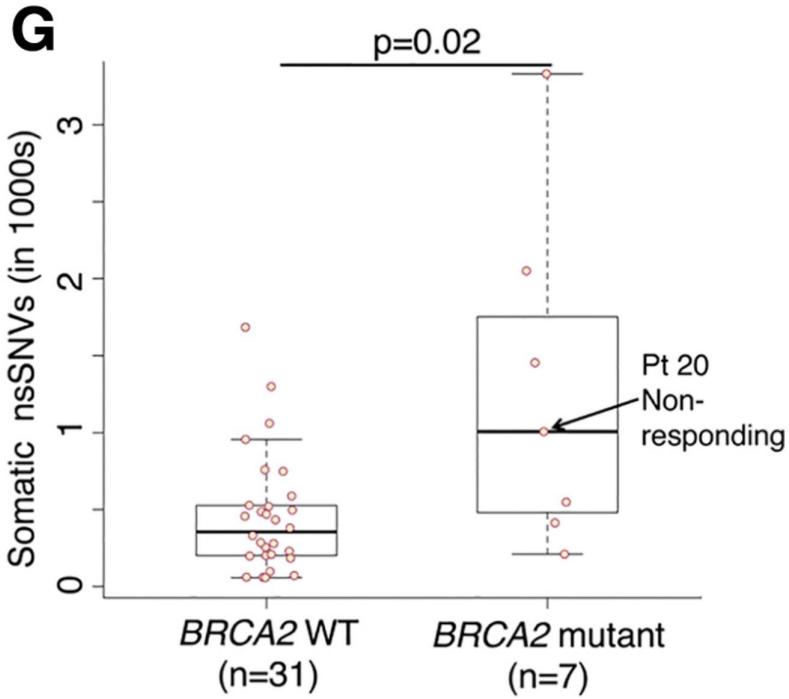
Altered in 32 (88.89%) of 36 samples.



UNIVERSITY OF  
BIRMINGHAM

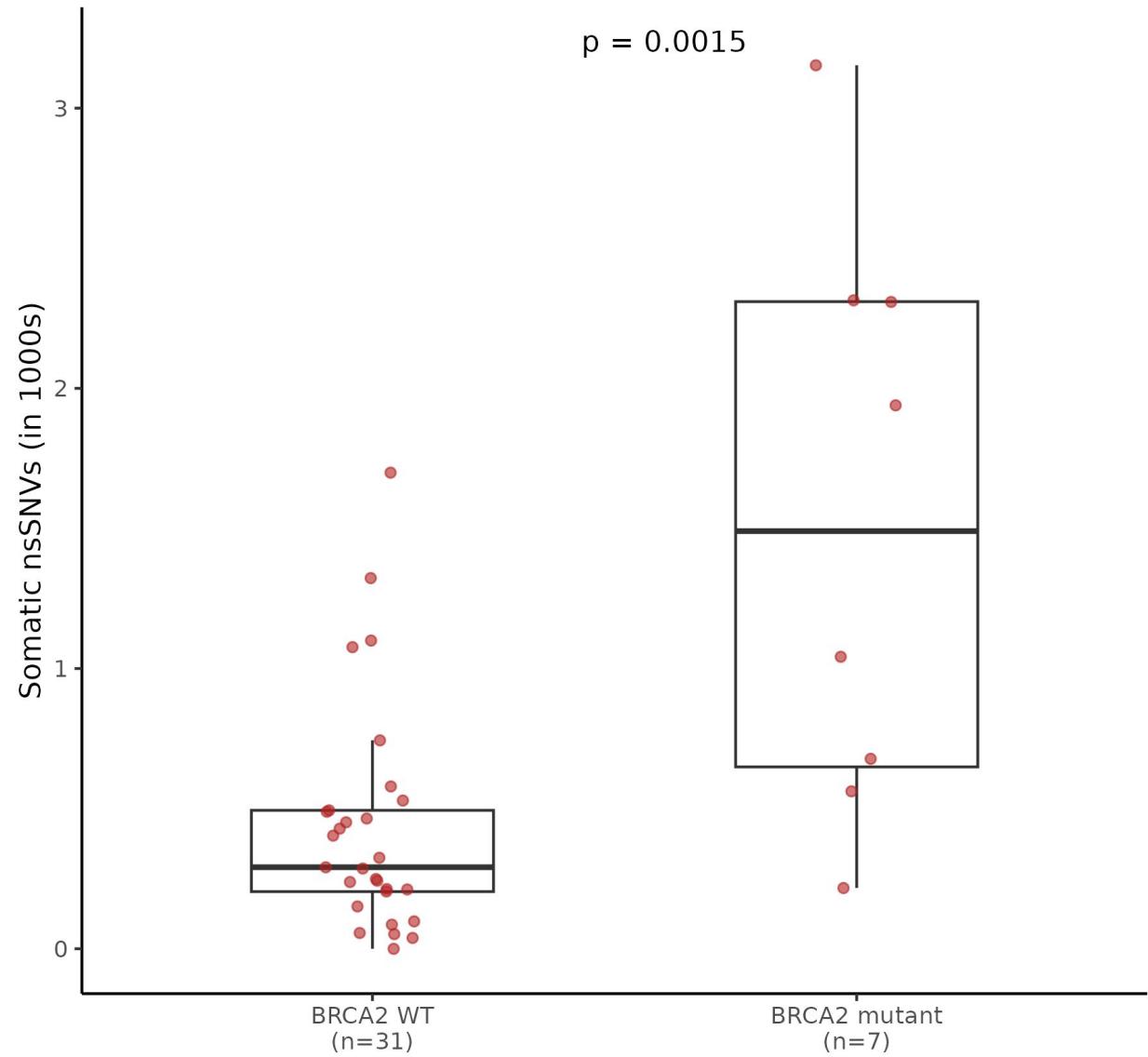
DUBAI

# Mutational load in BRCA2 mutants



BRCA2 mutations lead to more genomic instability:

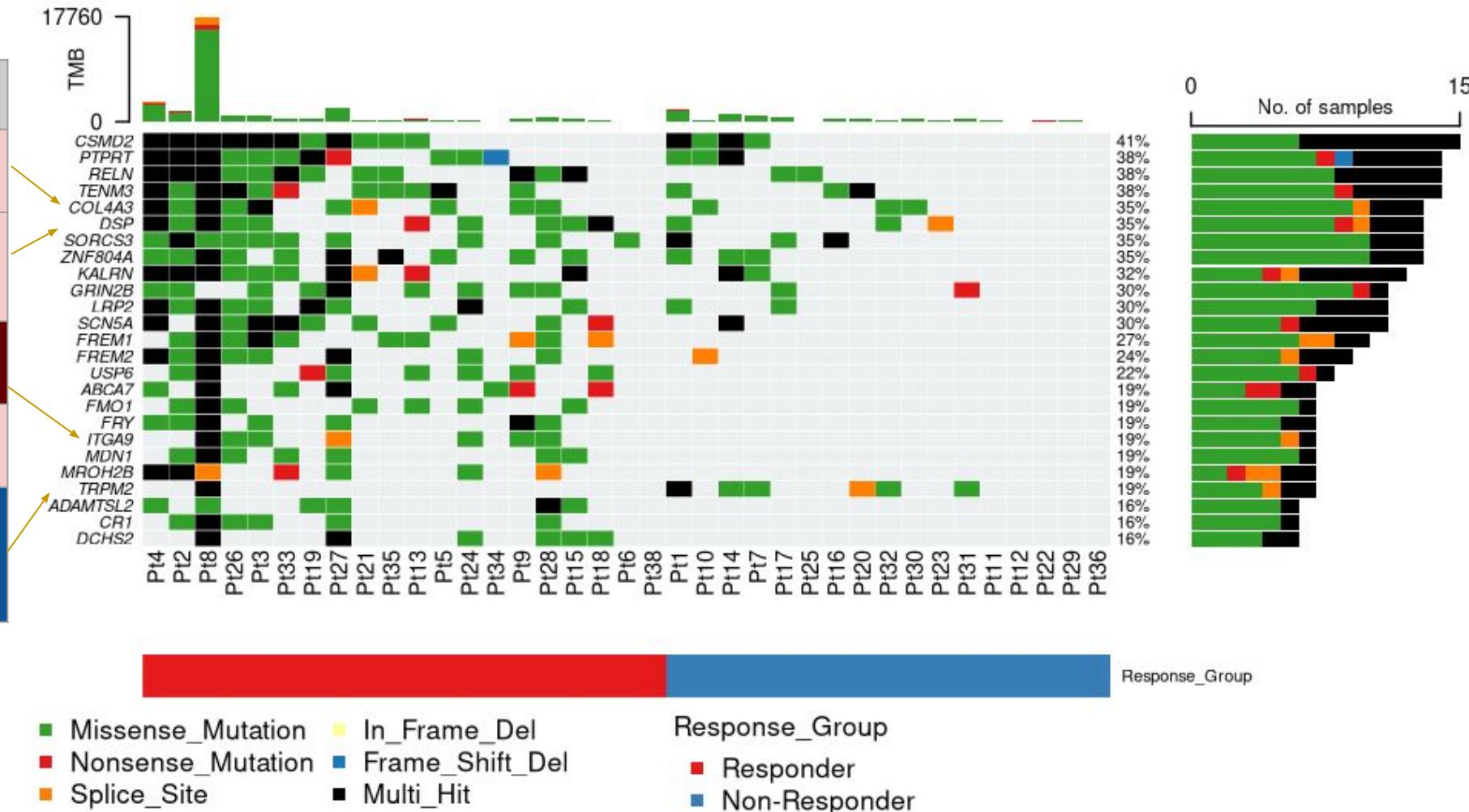
→ Accumulation of more mutations



# Exploratory search of recurrent variants

Altered in 31 (83.78%) of 37 samples.

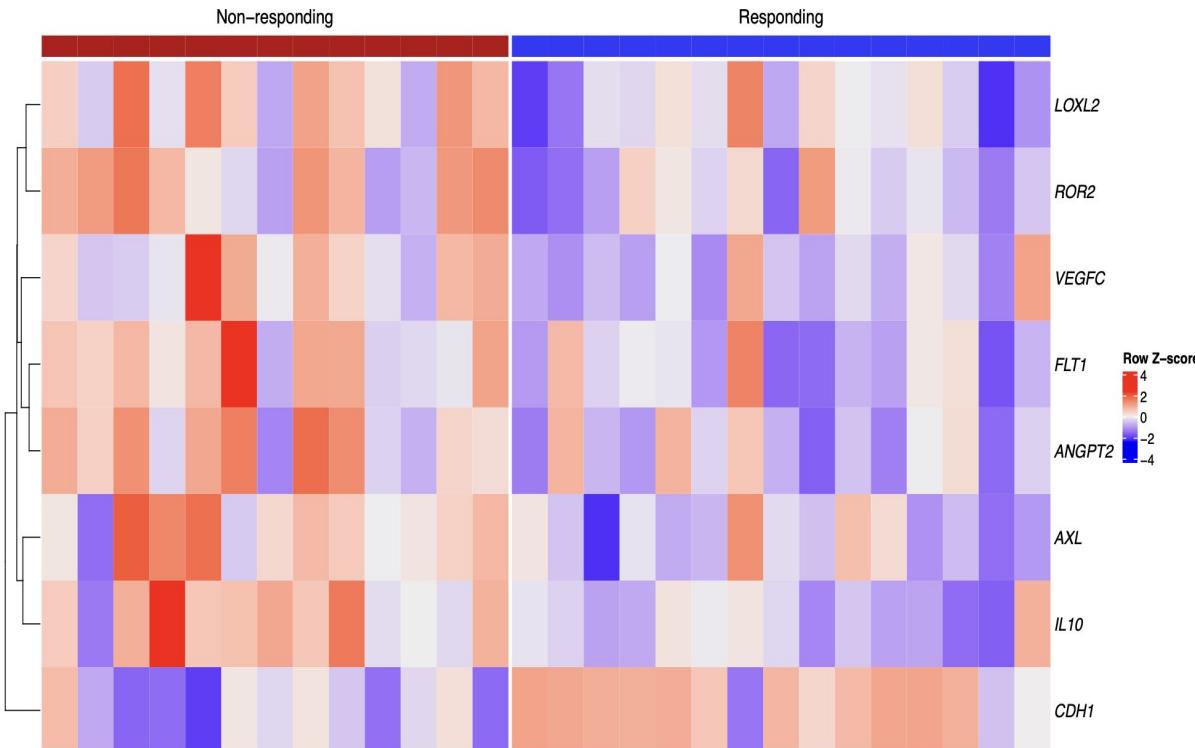
Gene	Function
<b>COL4A3</b>	basement membrane collagen, barrier to immune infiltration
<b>DSP (Desmoplakin)</b>	Central to cell–cell junction integrity, related to tissue disorganization and inflammation
<b>ITGA9</b>	cell adhesion and immune cell interactions
<b>CSMD2, CR1</b>	Immune interaction and complement-related
<b>TRPM2</b>	Stress and redox-sensitive calcium channel, Linked to adaptation to chronic inflammatory stress



# RESULTS - RNA-seq



# Non-responders exhibit an immune-excluded transcriptional program

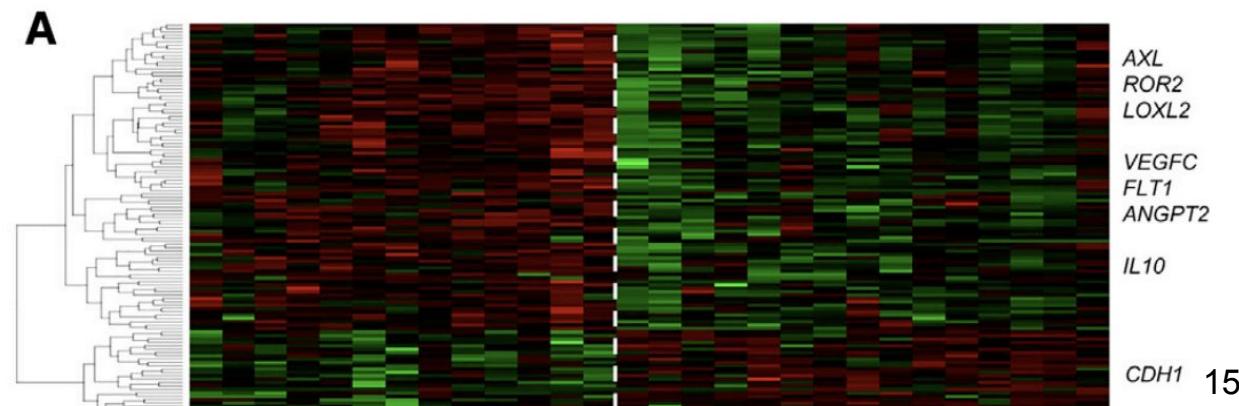


Differential expression analysis highlights genes linked to:

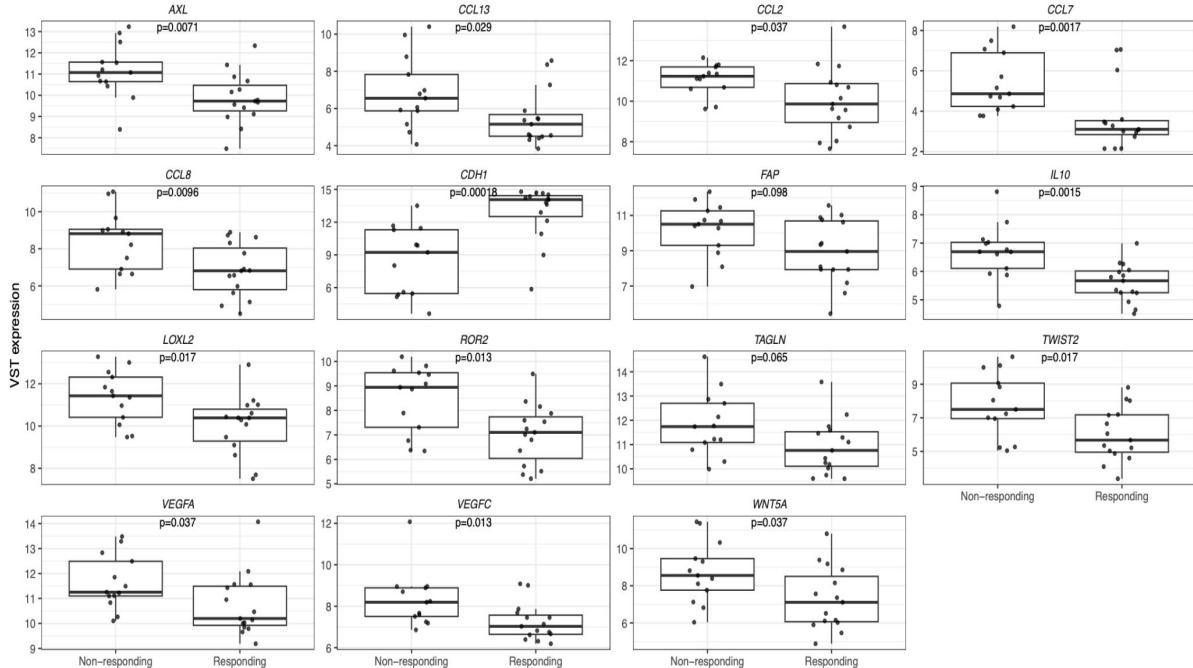
- Angiogenesis (VEGFC, FLT1, ANGPT2)
- EMT / invasion (LOXL2, ROR2, AXL)
- Immunosuppression (IL10)

These genes are consistently **upregulated in non-responders**

Suggests a **T-cell excluded, therapy-resistant tumor phenotype**



# Non-responders display an immune-excluded transcriptional program

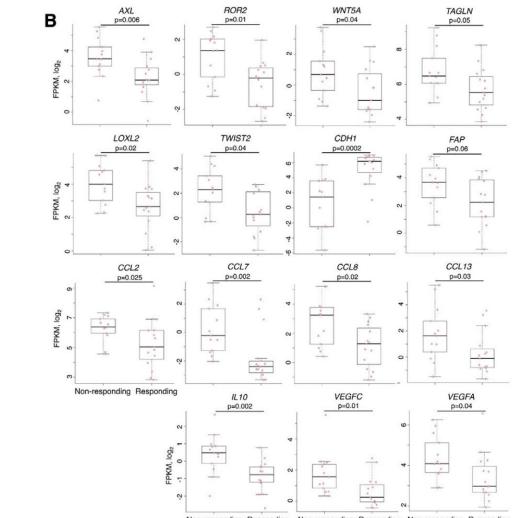


Differential expression analysis reveals consistent upregulation of:

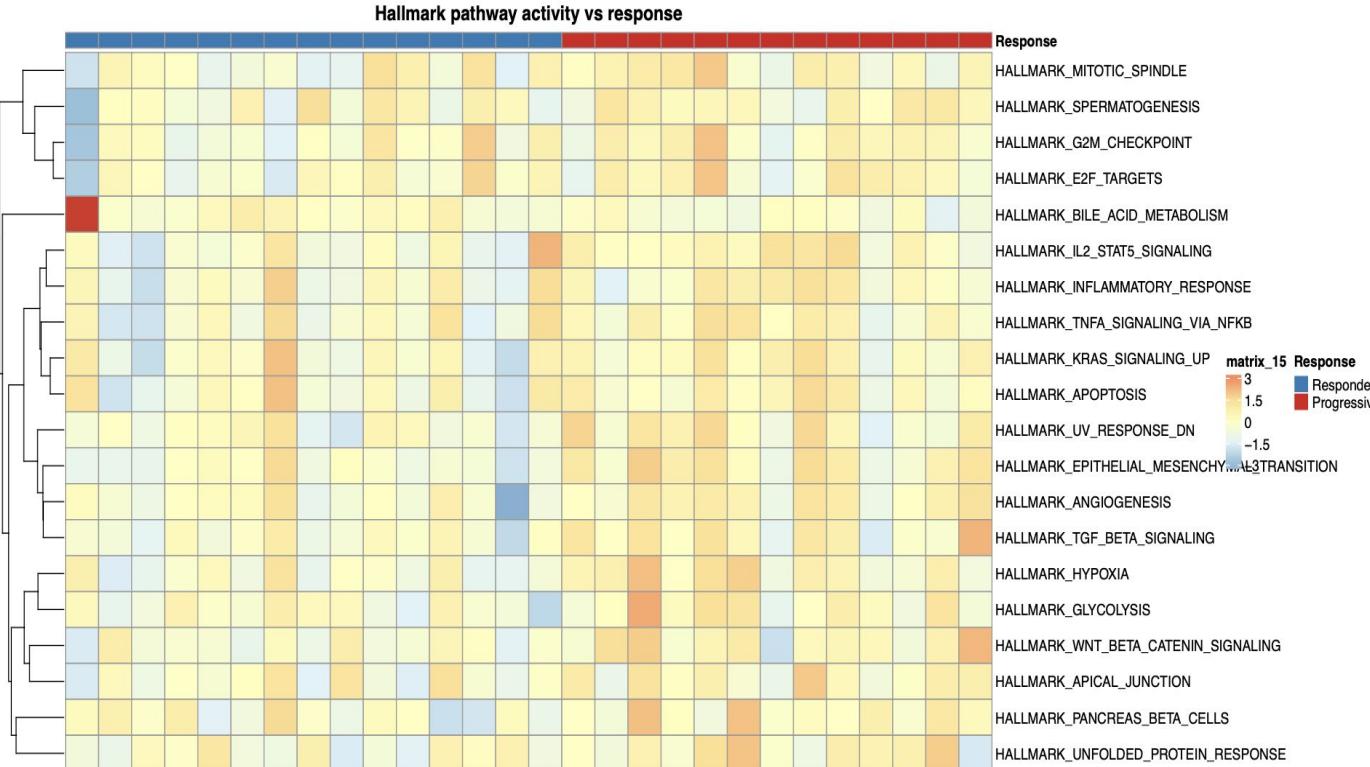
- angiogenesis-related genes (VEGFA, VEGFC)
- EMT markers (LOXL2, TWIST2, ROR2)
- immunosuppressive mediators (IL10, AXL)

These pathways are **significantly elevated in non-responders**

Suggests a tumor microenvironment hostile to immune infiltration



# GSVA reveals immune-excluded pathway activation in non-responders



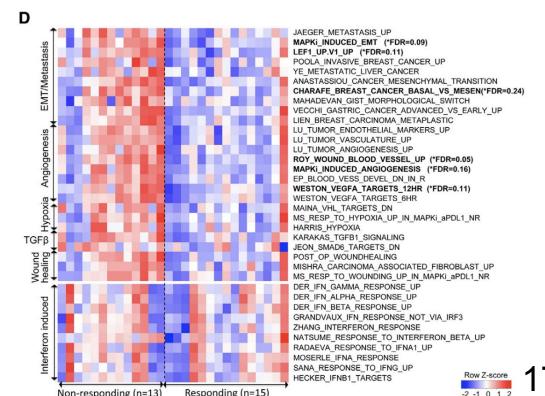
GSVA was applied to Hallmark gene sets across all samples.

Non-responders show enrichment of:

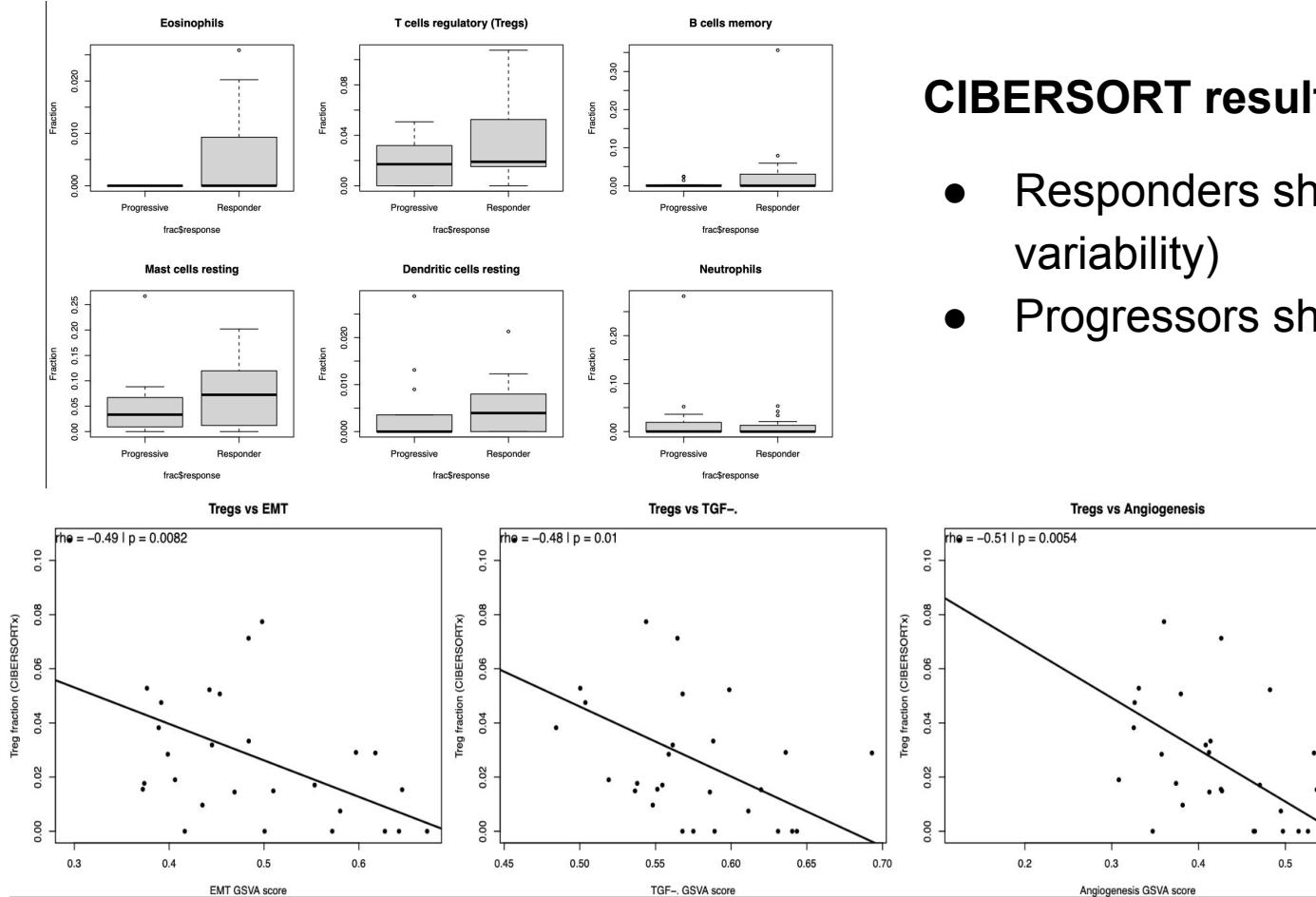
- EMT and TGF- $\beta$  signaling
- Angiogenesis and hypoxia
- Metabolic reprogramming

Responders lack activation of these resistance-associated programs

Suggests **transcriptional immune exclusion** as a key resistance mechanism



# Immune Exclusion Rather Than Immune Suppression Drives Resistance



## CIBERSORT results

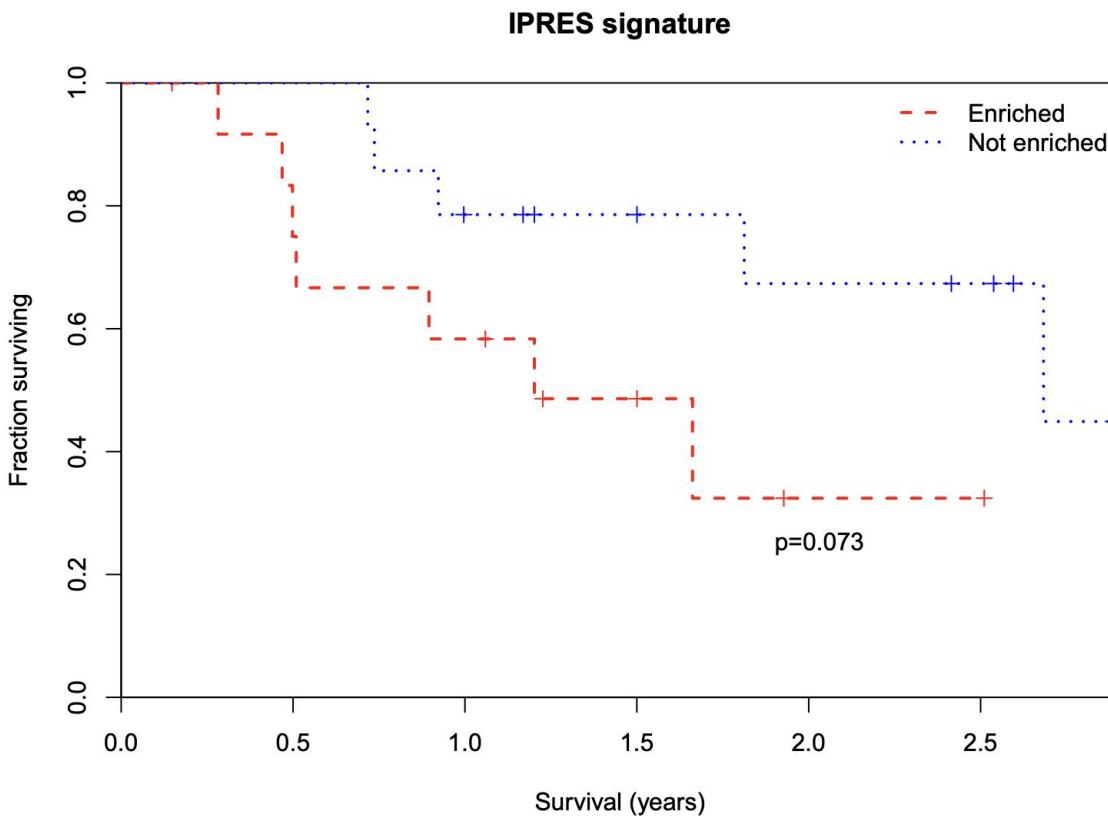
- Responders show higher Treg and B-cell fractions (higher variability)
- Progressors show altered innate immune composition

## GSVA correlations

- Tregs inversely correlated with:
  - EMT
  - TGF- $\beta$  signaling
  - Angiogenesis
- Indicates immune exclusion rather than immune absence



# IPRES enrichment is associated with poorer survival

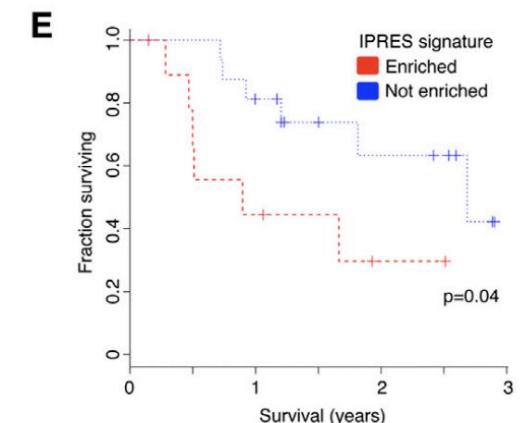


Patients with high IPRES activity show reduced overall survival

Trend consistent with immune-excluded tumor phenotype

Paper's IPRES definition more optimized, hence lower p value (0.04)

Supports transcriptomic evidence of resistance mechanisms



# Conclusion

- Mutational load plays a role in survival but is not the only factor.
  - it does NOT predict response to treatment by itself.
- Further analysis of genes with enriched mutations in responder vs. non-responder needs to be done.
  - Through literature review, we found interesting genes involved in extracellular remodeling, cell adhesion, immune regulation.
- RNA-seq analysis identified an immune-excluded transcriptional state driven by EMT and stromal pathways, consistent with IPRES-mediated resistance to immunotherapy.

# DEMO



UNIVERSITY OF  
BIRMINGHAM | DUBAI


[Home](#) | [Submit](#) ▾ | [Search](#) ▾ | [About](#) ▾ | [Support](#) ▾

## Study: SRP067938

Only tumours that cannot upregulate programmed death ligand 1 (PD-L1) upon interferon exposure should be considered negative for this ligand, and would be unlikely to respond to PD-1 blockade therapy. Here we show that genetic mutations in the Janus kinases (JAK) controlling signalling downstream of the interferon receptor prevent PD-L1 upregulation on melanoma cells upon interferon exposure and result in innate resistance to PD-1 blockade in patients. Among 50 human melanoma cell lines tested for response to interferon gamma by PD-L1 surface expression, three were negative. Two of them had loss of function mutations in JAK1 or JAK2, leading to lack of downstream signalling from the interferon receptor. Whole exome sequencing of biopsies from 23 patients with metastatic melanoma treated with the anti-PD1 antibody pembrolizumab revealed a homozygous JAK1 missense mutation in the biopsy of the patient with the highest mutational load among the 9 patients without a clinical response to anti-PD-1 therapy, and no interferon receptor pathway homozygous loss of function mutations in the biopsies of 14 patients with response to therapy. A cell line established from this biopsy showed lack of sustained signalling response to interferon alpha, beta or gamma leading to markedly decreased reactive PD-L1 expression. Analysis of whole exome sequencing of 16 cases of colon cancer with mismatch repair deficiency also showed a homozygous JAK1 loss of function mutation in a patient without a tumour response to PD-1 blockade. Analysis of 905 cell lines from the

[Show More](#)

[Search](#)

Examples: histone, BN000065

[View](#)

Examples: Taxon:9606, BN000065, PRJEB402

### General

[View:](#) XML

[Download:](#) XML

[Cross References:](#) Show

[Publications:](#) Show

[Related ENA Records:](#) Show

### Data

[Read Files:](#) Hide

### Tags

xref

EuropePMC

<b>Organism:</b>	Homo sapiens (human)
<b>Study Accession:</b>	PRJNA307199
<b>Center Name:</b>	UCLA
<b>Study Name:</b>	Homo sapiens
<b>ENA-FIRST-PUBLIC:</b>	2016-09-27
<b>ENA-LAST-UPDATE:</b>	2016-09-27

# 1 - Downloading data with SRA ToolKits

```

1#!/bin/bash
2#SBATCH --job-name=rnaseq_download
3#SBATCH --cpus-per-task=8
4#SBATCH --mem=32G
5#SBATCH --time=24:00:00
6#SBATCH --output=rnaseq_%j.out
7#SBATCH --error=rnaseq_%j.err
8
9# load SRA Toolkit module on BlueBEAR
10module purge
11module load bear-apps/2023a
12module load SRA-Toolkit/3.0.10-gompi-2023a
13
14# go to your project directory
15cd /rds/projects/e/elhamsak-genomics-ngs/group_5/main_project/rnaseq
16
17# list your SRA accessions here
18SRR_LIST=(
19"SRR3184279"
20"SRR3184280"
21"SRR3184281"
22"SRR3184282"
23"SRR3184283"
24"SRR3184284"
25"SRR3184285"
26"SRR3184286"
27"SRR3184287"
28"SRR3184288"
29"SRR3184289"
30"SRR3184290"
31"SRR3184291"
32"SRR3184292"
33"SRR3184293"
34"SRR3184294"
35"SRR3184295"
36"SRR3184296"
37"SRR3184297"
38"SRR3184298"
39"SRR3184299"
40"SRR3184300"
41"SRR3184301"
42"SRR3184302"
43"SRR3184303"
44"SRR3184304"
45"SRR3184305"
46"SRR3184306"
47)
48
49for SRR in "${SRR_LIST[@]}"; do
50    echo "Processing $SRR ..."
51
52    prefetch $SRR
53
54    fasterq-dump $SRR --split-files --threads 8
55
56    gzip ${SRR}_1.fastq
57    gzip ${SRR}_2.fastq
58
59    echo "$SRR done!"
60    echo "-----"
61done
62
63echo "All SRR accessions processed."

```

# 2 - Trimming seqs with fastp

```

1#!/bin/bash
2#SBATCH --job-name=fastp_trim
3#SBATCH --cpus-per-task=8
4#SBATCH --mem=16G
5#SBATCH --time=24:00:00
6#SBATCH --output=logs/fastp_%j.out
7#SBATCH --error=logs/fastp_%j.err
8
9module purge
10module load bear-apps/2023a
11module load fastp/0.24.0-GCC-12.3.0
12
13INPUT_DIR="/rds/projects/e/elhamsak-group5/main_project/rnaseq"
14OUT_DIR="/rds/projects/e/elhamsak-group5/main_project/trimmed_files/rnaseq"
15
16
17SRR_LIST=(
18"SRR3184280"
19"SRR3184281"
20"SRR3184282"
21"SRR3184283"
22"SRR3184284"
23"SRR3184285"
24"SRR3184286"
25"SRR3184287"
26"SRR3184288"
27"SRR3184289"
28"SRR3184290"
29"SRR3184291"
30"SRR3184292"
31"SRR3184293"
32"SRR3184294"
33"SRR3184295"
34"SRR3184296"
35"SRR3184297"
36"SRR3184298"
37"SRR3184299"
38"SRR3184300"
39"SRR3184301"
40"SRR3184302"
41"SRR3184303"
42"SRR3184304"
43"SRR3184305"
44"SRR3184306"
45)
46
47for SRR in "${SRR_LIST[@]}"; do
48    echo "-----"
49    echo "Running fastp for $SRR"
50    echo "-----"
51
52    R1="${INPUT_DIR}/${SRR}_1.fastq.gz"
53    R2="${INPUT_DIR}/${SRR}_2.fastq.gz"
54
55    OUT_R1="${OUT_DIR}/${SRR}_1.trimmed.fasta.gz"
56    OUT_R2="${OUT_DIR}/${SRR}_2.trimmed.fasta.gz"
57
58    HTML="${OUT_DIR}/${SRR}.fastp.html"
59    JSON="${OUT_DIR}/${SRR}.fastp.json"
60
61    fastp \
62        -i "$R1" -I "$R2" \
63        -o "$OUT_R1" -O "$OUT_R2" \
64        -h "$HTML" -j "$JSON" \
65        -w 8
66
67    echo "Done: $SRR"
68done
69
70echo "All fastp jobs completed."

```

# 4 - Alignment with STAR

```

1#!/bin/bash
2#SBATCH --job-name=STAR_index
3#SBATCH --cpus-per-task=8
4#SBATCH --mem=64G
5#SBATCH --time=12:00:00
6#SBATCH --output=star_index_%j.out
7#SBATCH --error=star_index_%j.err
8
9cd /rds/projects/e/elhamsak-group5/main_project/reference_genome/STAR_index
10
11module purge
12module load bear-apps/2024a
13module load STAR/2.7.11b-GCC-13.3.0
14
15STAR \
16    --runMode genomeGenerate \
17    --runThreadN 8 \
18    --genomeDir /rds/projects/e/elhamsak-group5/main_project/reference_genome/STAR_index \
19    --genomeFastaafiles /rds/projects/e/elhamsak-group5/main_project/reference_genome/hg38.fa \
20    --sjdbGTFfile /rds/projects/e/elhamsak-group5/main_project/reference_genome/STAR_index/gencode.v49.annotation.nochr.gtf \
21    --sjdbOverhang 99
22
23#!/bin/bash
24#SBATCH --job-name=featureCounts
25#SBATCH --cpus-per-task=8
26#SBATCH --mem=16G
27#SBATCH --time=04:00:00
28#SBATCH --output=logs/featureCounts_%j.out
29#SBATCH --error=logs/featureCounts_%j.err
30
31set -eu pipefail
32
33mkdir -p logs
34
35cd /rds/projects/e/elhamsak-group5/main_project/
36
37module purge
38module load bear-apps/2024a
39module load STAR/2.7.11b-GCC-13.3.0
40
41# Paths
42GENOME_DIR="/rds/projects/e/elhamsak-group5/main_project/reference_genome/STAR_index"
43INPUT_DIR="/rds/projects/e/elhamsak-group5/main_project/trimmed_files/rnaseq"
44OUT_DIR="/rds/projects/e/elhamsak-group5/main_project/aligned_files/rnaseq"
45
46
47# Your paired-end SRR list
48SRR_LIST=(
49"SRR3184280"
50"SRR3184281"
51"SRR3184282"
52"SRR3184283"
53"SRR3184284"
54"SRR3184285"
55"SRR3184286"
56"SRR3184287"
57"SRR3184288"
58"SRR3184289"
59"SRR3184290"
60"SRR3184291"
61"SRR3184292"
62"SRR3184293"
63"SRR3184294"
64"SRR3184295"
65"SRR3184296"
66"SRR3184297"
67"SRR3184298"
68"SRR3184299"
69"SRR3184300"
70"SRR3184301"
71"SRR3184302"
72"SRR3184303"
73"SRR3184304"
74"SRR3184305"
75"SRR3184306"
76)
77
78for SRR in "${SRR_LIST[@]}"; do
79    echo "-----"
80    echo "Aligning $SRR with STAR"
81    echo "-----"
82
83    R1="${INPUT_DIR}/${SRR}_1.trimmed.fasta.gz"
84    R2="${INPUT_DIR}/${SRR}_2.trimmed.fasta.gz"
85    PREFixX="${OUT_DIR}/${SRR}"
86
87    if [[ ! -f "$R1" || ! -f "$R2" ]]; then
88        echo "ERROR: Missing trimmed FASTQs for $SRR"
89        echo "$SRR"
90        exit 1
91    fi
92
93    STAR \
94        --runThreadN "${SLURM_CPUS_PER_TASK}" \
95        --genomeDir "$GENOME_DIR" \
96        --readFilesIn "$R1 $R2" \
97        --readOrder random \
98        --outFileNamePrefix "$PREFIXX" \
99        --outSAMtype BAM SortedByCoordinate \
100        --quantMode GeneCounts \
101        --tqposMode Basic \
102        --outMetabolicLine ID:$SRR SM:$SRR PL:ILLUMINA \
103        --outSAMattrField introMotif
104
105    echo "Done: $SRR"
106done
107
108echo "All STAR alignments completed."

```

# 3 - Indexing with STAR

# 5 - Count data with featurecounts

https://portal.bear.bham.ac.uk/rnode/bear-pg0201u25a.bear.cluster/27192/

File Edit Code View Plots Session Build Debug Profile Tools Help

deseq2\_final.Rmd x Knit on Save Knit Run Publish Outline

```
1 ---  
2 title: "Differential Expression & Pathway Analysis (DESeq2)"  
3 authors: "Aishwarya, Asta, & Aman"  
4 date: ``r format(Sys.Date())``  
5 output:  
6   html_document:  
7     toc: true  
8     toc_depth: 3  
9     number_sections: true  
10    df_print: paged  
11 ---  
12  
13 `r setup, include=FALSE}  
14 knitr::opts_chunk$set(echo = TRUE, message = FALSE, warning = FALSE)  
15 set.seed(123)  
16  
17 # File locations  
18 counts_path <- "/rds/projects/e/elhamsak-group5/main_project/data/rnaseq_counts/gene_counts.txt"  
19 meta_path <- "/rds/projects/e/elhamsak-group5/main_project/data/meta_data/SraRunTable.csv"  
20 ciber_path <- "/rds/projects/e/elhamsak-group5/main_project/data/cibersortx/CIBERSORTx_Results.csv"  
21  
22 # Output folder  
23 out_dir <- "/rds/projects/e/elhamsak-group5/main_project/figures/rnaseq"  
24 out_dir_results <- "/rds/projects/e/elhamsak-group5/main_project/results/rnaseq"  
25  
26  
27 # Analysis parameters  
28 min_count <- 10  
29 min_prop_samples <- 0.10  
30 alpha_fdr <- 0.10  
31  
32 check_pkg <- function(pkg) {  
33   if (!requireNamespace(pkg, quietly = TRUE)) {  
34     stop(  
35       "Missing package: ", pkg, "\n",
```

30:18 Chunk 1: setup R Markdown

Environment History Connection

1.95 GiB Global Environment

Data

- dds Large DESeqDataS...
- pca Large prcomp (5 ...
- vsd Large DESeqTrans...
- vst\_mat Large matrix (81...

Files Plots Packages Help

deseq2\_final.Rmd deseq2.Rmd

Console

axp1383 Project: (None)

# Sources

**Original Paper:**

**Hugo et al., Cell, 2016**

*Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma*

**Image sources:**

Barot, Shrikant, et al. "Recent advancement in targeted therapy and role of emerging technologies to treat cancer." *Medical Oncology* 40.11 (2023): 324.

**Computational Tools:**

**fastp:** [github.com/OpenGene/fastp](https://github.com/OpenGene/fastp)

**STAR aligner:** [github.com/alexdobin/STAR](https://github.com/alexdobin/STAR)

**featureCounts:** [rddocumentation.org/.../featureCounts](https://rddocumentation.org/.../featureCounts)

**DESeq2:** [bioconductor.org/packages/DESeq2](https://bioconductor.org/packages/DESeq2)

**GSVA:** [bioconductor.org/packages/GSVA](https://bioconductor.org/packages/GSVA)

**CIBERSORTx:** [cibersortx.stanford.edu](https://cibersortx.stanford.edu)

**survival:** [cran.r-project.org/web/packages/survival](https://cran.r-project.org/web/packages/survival)

**GATK** <https://gatk.broadinstitute.org/hc/en-us>

**bedtools** <https://bedtools.readthedocs.io/en/latest/>

**maftools** <https://bioconductor.org/packages/release/bioc/vignettes/maftools/inst/doc/maftools.html>



# QUESTIONS?

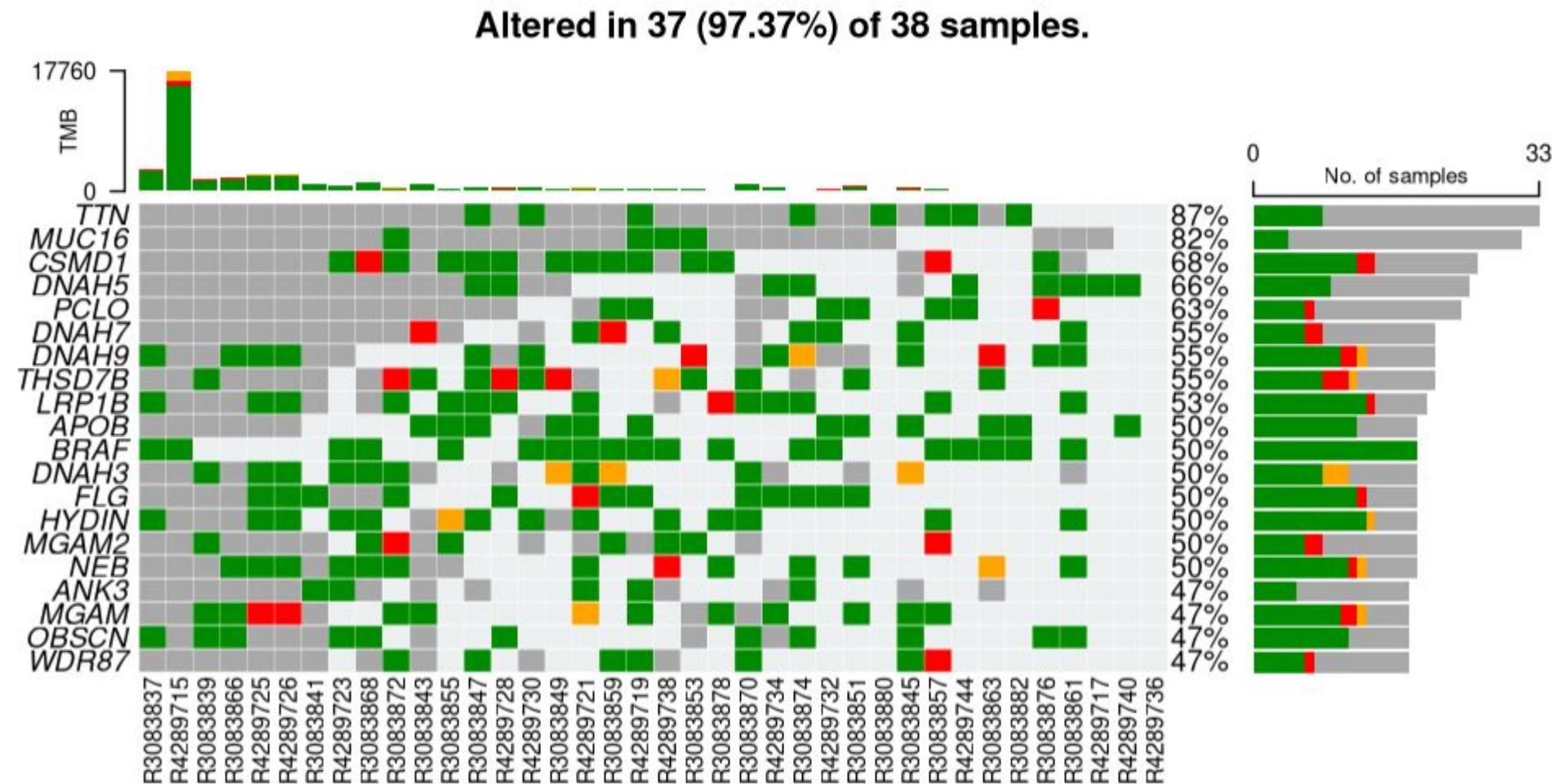


UNIVERSITY OF  
BIRMINGHAM | DUBAI

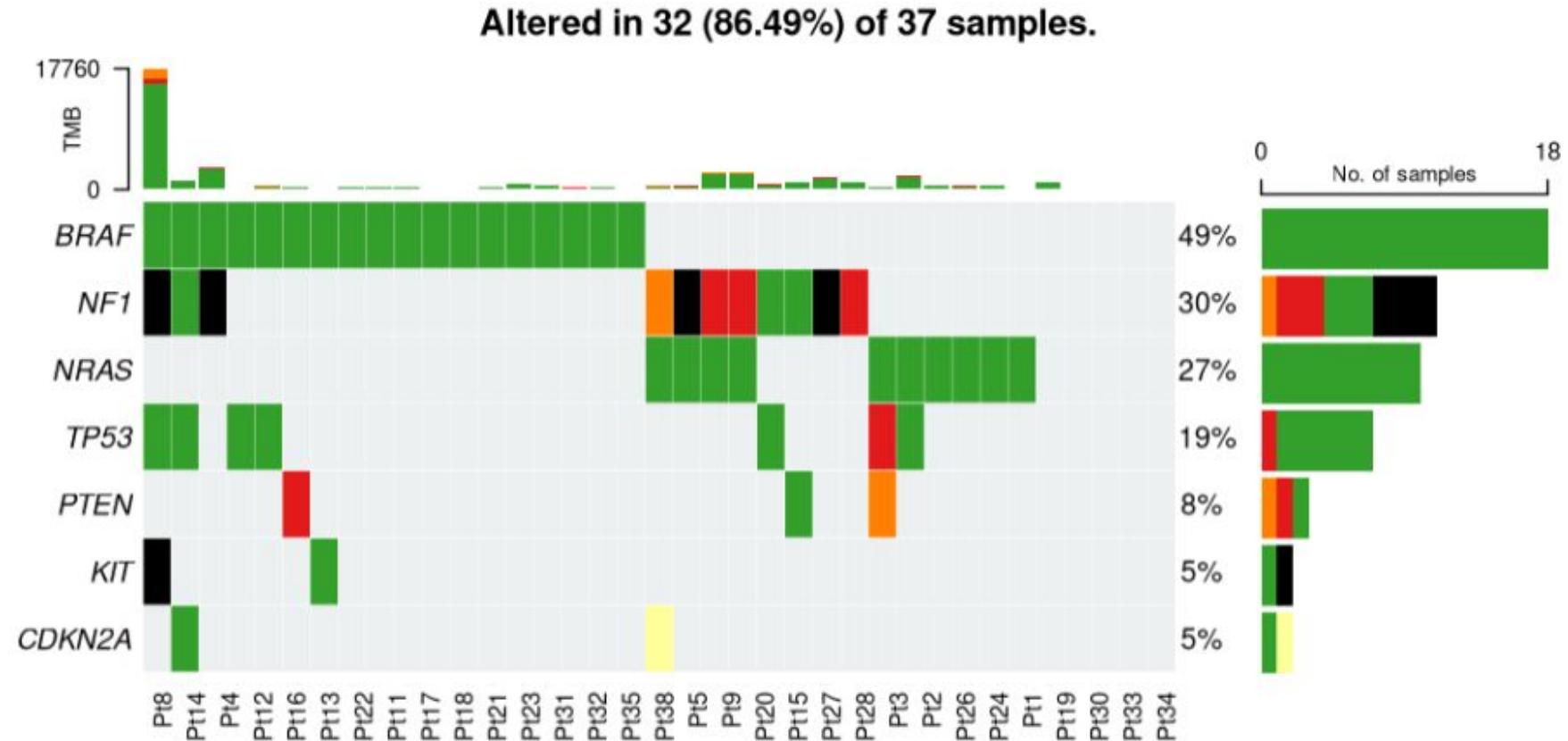
# **Extra slides**



# Results - General summary of top 20 genes



# Results - General summary of top 20 genes (plotted using a list of common driver mutations)



# Results - WES-TMB distribution and the most frequently mutated genes

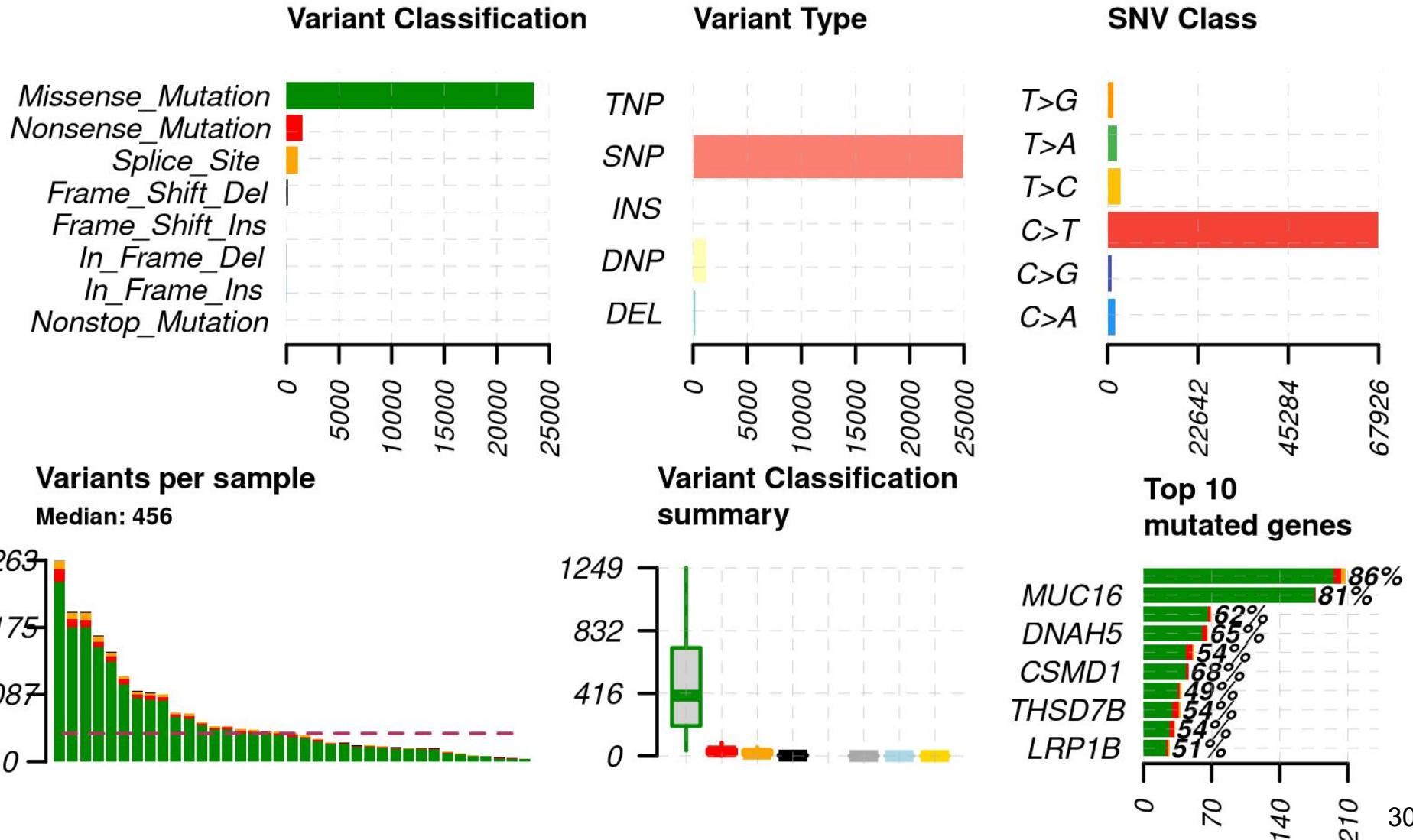
Investigation of non-synonymous SNVs (snSNV)

- counts in nsSNVs correlate well with the counts found by the authors
- one outlier in our data

→ excluded it

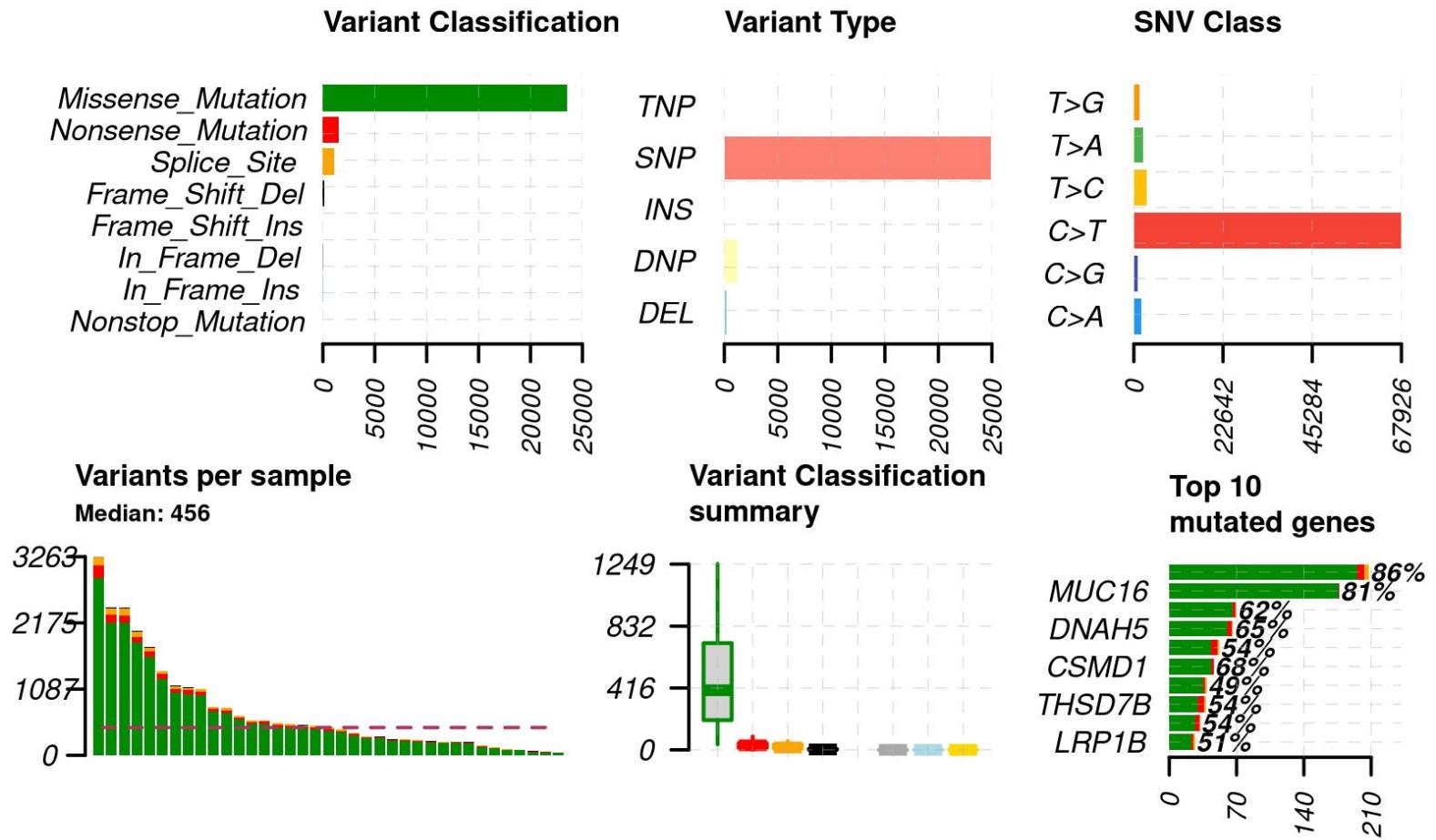


UNIVERSITY OF  
BIRMINGHAM | DUBAI

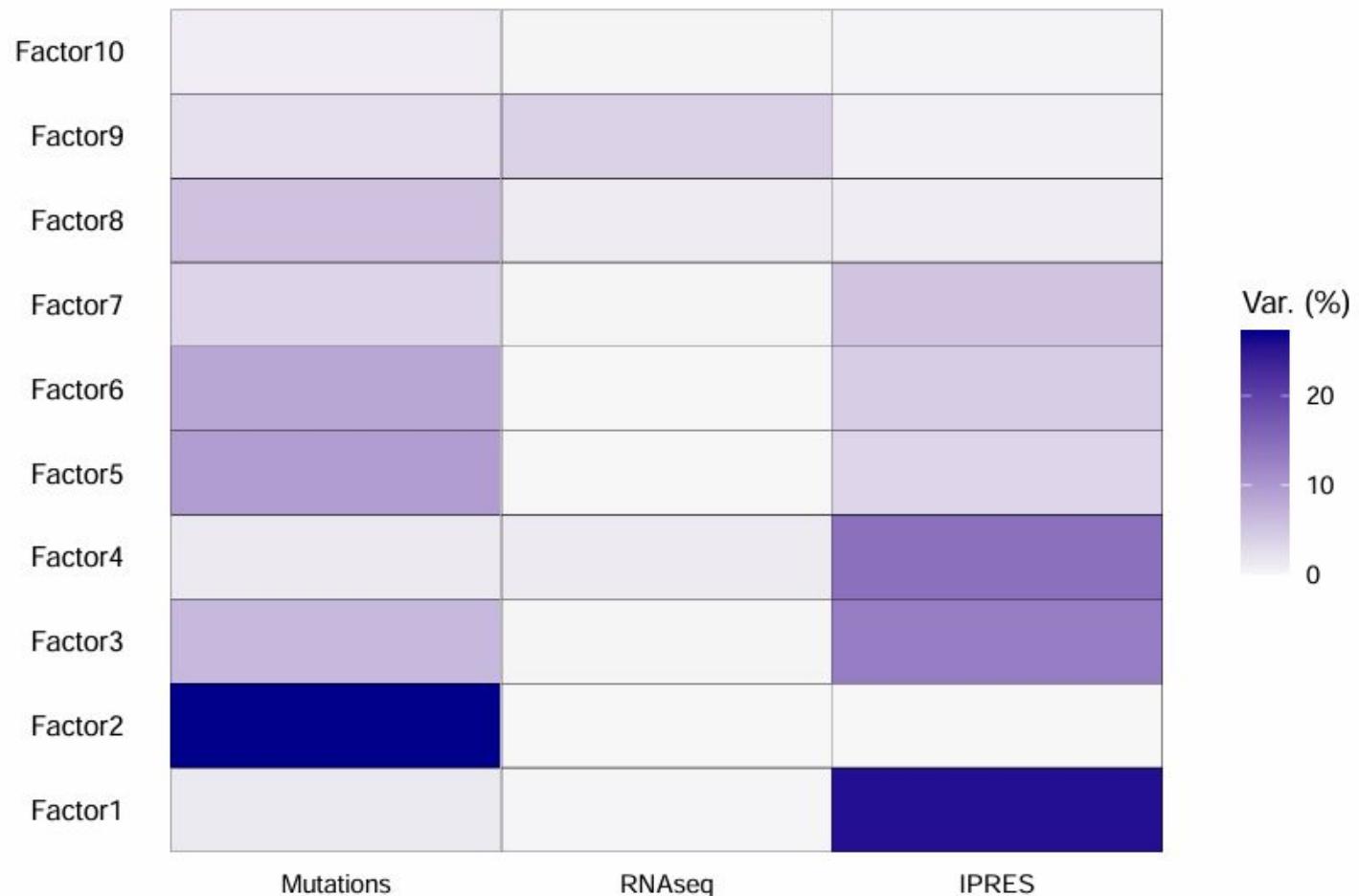


# Results - WES

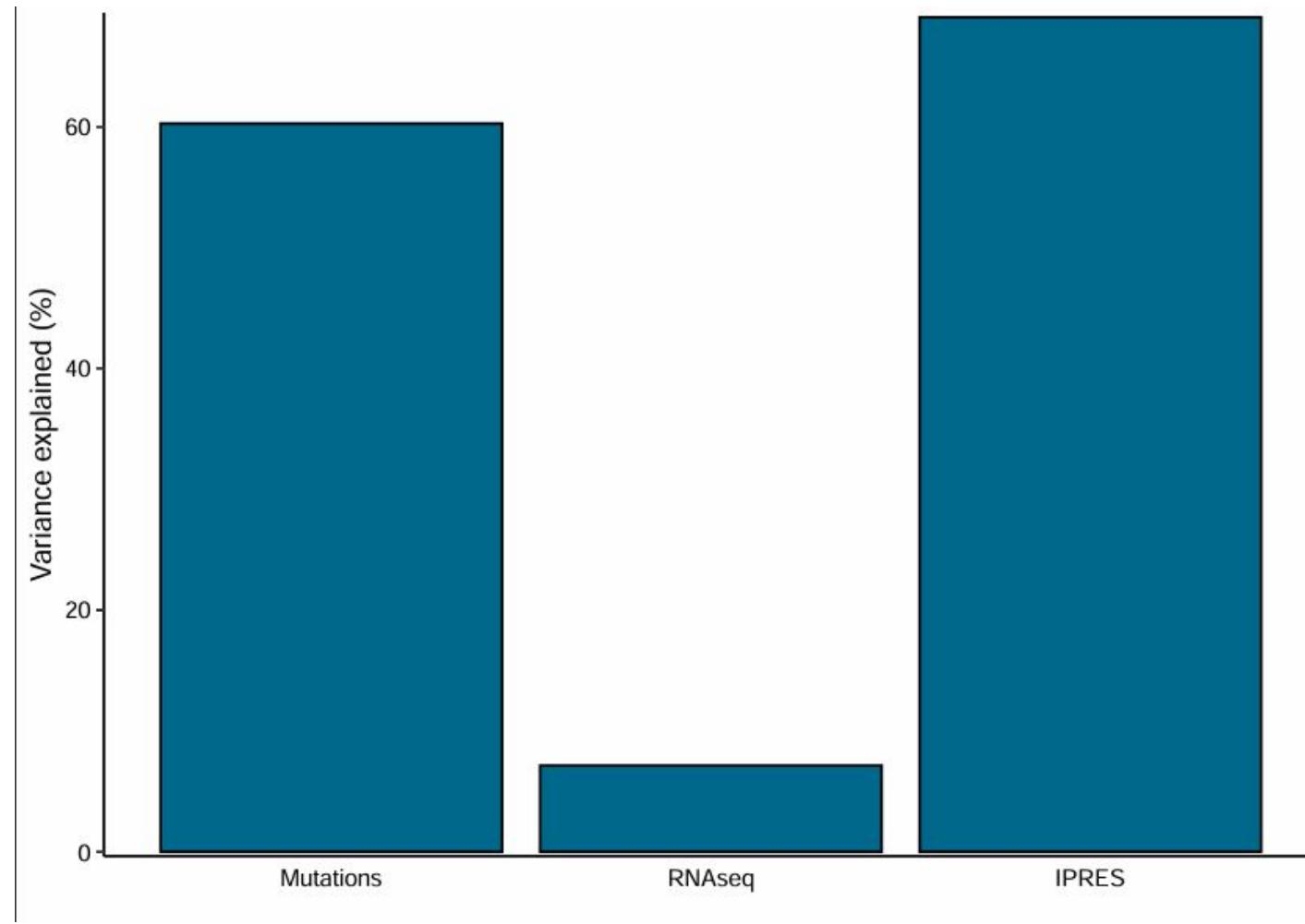
shows variant distribution and the most frequently mutated genes.



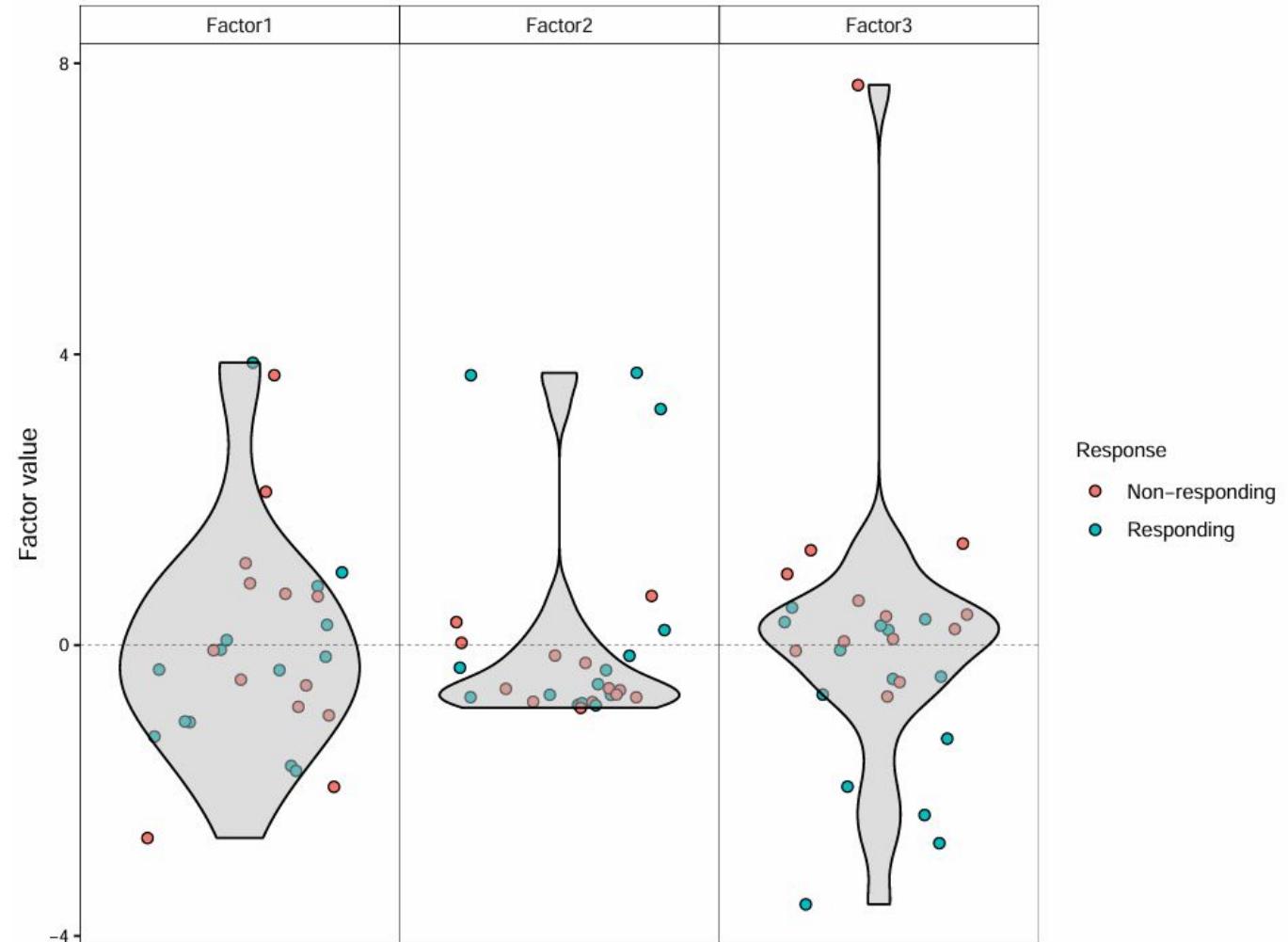
# Results - MOFA- FACTOR VARIANCE



# Results - MOFA- TOTAL VARIANCE



# Results - MOFA- TOTAL VARIANCE



# Results - WES

## TMB per sample- Outlier in Pt8

