Genomic landscape of inflammatory breast cancer by whole-genome sequencing

Xiaotong Li

May 17, 2018

Background

- **Inflammatory Breast Cancer (IBC)**
  - IBC is a clinical diagnosis, no molecular marker
  - Rare (<5%) & Aggressive
  - 5-year survival rate: limited to 40%[1]
  - Affects young patients population
  - characterized by a highly metastatic phenotype
  - NO known risk factors specified to IBC
  - failed to identify, recurrent, IBC-specific gene expression or DNA copy number alterations

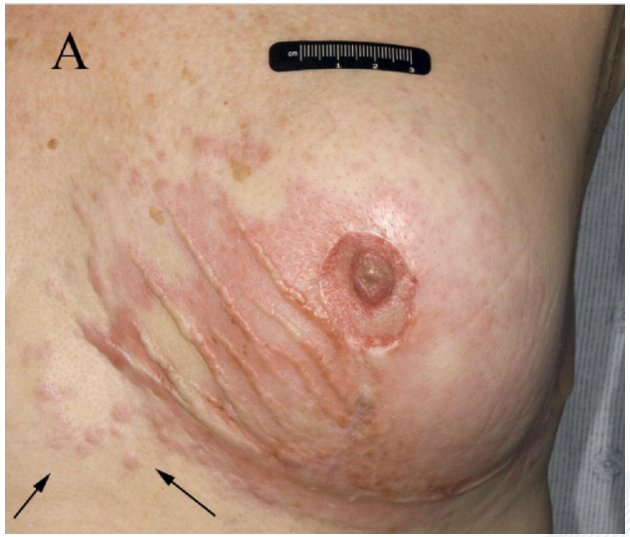


Fig1 (A) Clinical and pathologic signs of IBC. Erythema of the breast with tumor nodules extending to the opposite breast (arrows).[2]

[1] Bertucci, Francois, et al. Cancer research 64.23 (2004): 8558-8565.
[2] Woodward, Wendy A., et al. Seminars in radiation oncology (2009).

Goals

- **Is there a unique, highly recurrent DNA sequence alteration that defines IBC (as CDH1 deletion defines ILC) in the coding or non-coding regions of the whole genome?**

- **What are the genomic differences between IBC and Non-IBC**
  - Single nucleotide variants (SNV), insertions/deletions (indel)
  - Large structural variations (SV)
  - Copy number variations (CNV)
  - Germ-line polymorphisms (SNP)
  - Mutation signatures
  - Clonal composition
  - Bacteria or non-human genome
  - Canonical cancer pathway-level alterations

Materials

- **N=20 IBC biopsies from The Morgan Welch IBC Clinic/Research Program of MDACC**
  - DNA from snap frozen, newly diagnosed IBC
  - DNA from matched blood samples
  - individually reviewed for accuracy of diagnosis
  - tumor cellularity estimated by pathologists
  - N= 9 ER+/HER2-, 6 ER-/HER2-, 5 HER2+

- **N=23 Non-IBC breast cancer data from the TCGA for comparison**
  - matched ER, PR and HER2 status, age, race
  - N= 10 ER+/HER2-, 11 ER-/HER2-, 2 HER2+

Methods

- **Illumina pair-end whole-genome sequencing at Yale Center for Genome Analysis**
  - median coverage: 60X (cancer) and 40X (normal)
  - percent of mapped reads 99.3% (cancer) and 99.2% (normal)

- Germline and somatic variants, INDELs as well as large scale structural variants for both IBC and non-IBC cohorts were identified using the same pipeline

- FunSeq2[1] and PredictSNP2[2] were used to annotate variants and estimate functional impact

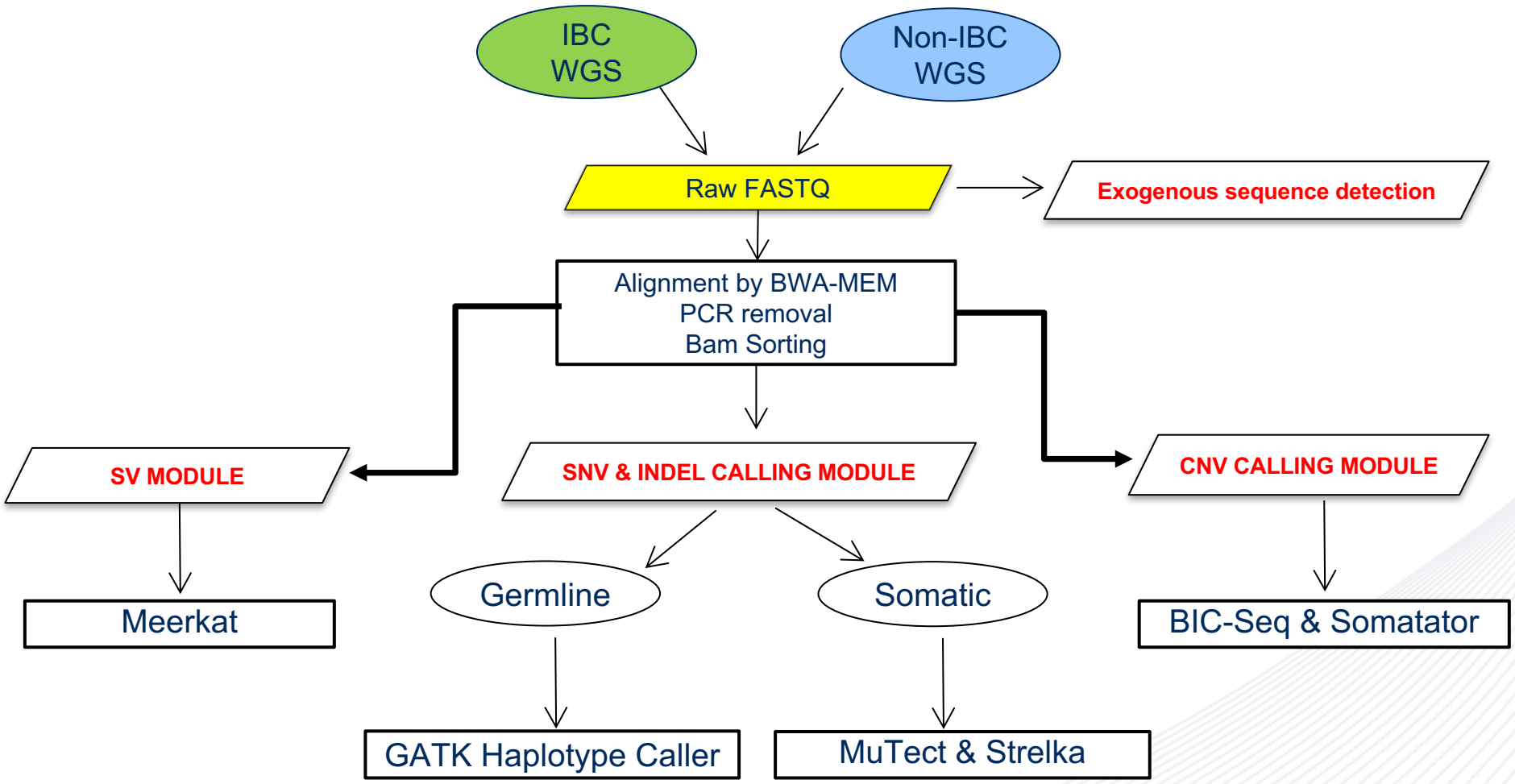
- DeconstructSigs[3] was used to determine mutational signatures

- Non-human sequences were detected using the exceRpt small RNA-seq pipeline[4].

- Clonal architecture and tumor evolution analysis were implemented by SciClone[5].

[1] Fu, Yao, et al. Genome biology 15.10 (2014): 480. [2] Bendl, Jaroslav, et al. PLoS Comput Biol 12.5 (2016): e1004962. [3] Rosenthal, Rachel, et al. Genome biology 17.1 (2016): 31. [4] The Genboree Workbench [http://www.genboree.org/] [5] Miller, Christopher A., et al. PLoS Comput Biol 10.8 (2014): e1003665.

Calling pipeline for the multiple types of genomic variants

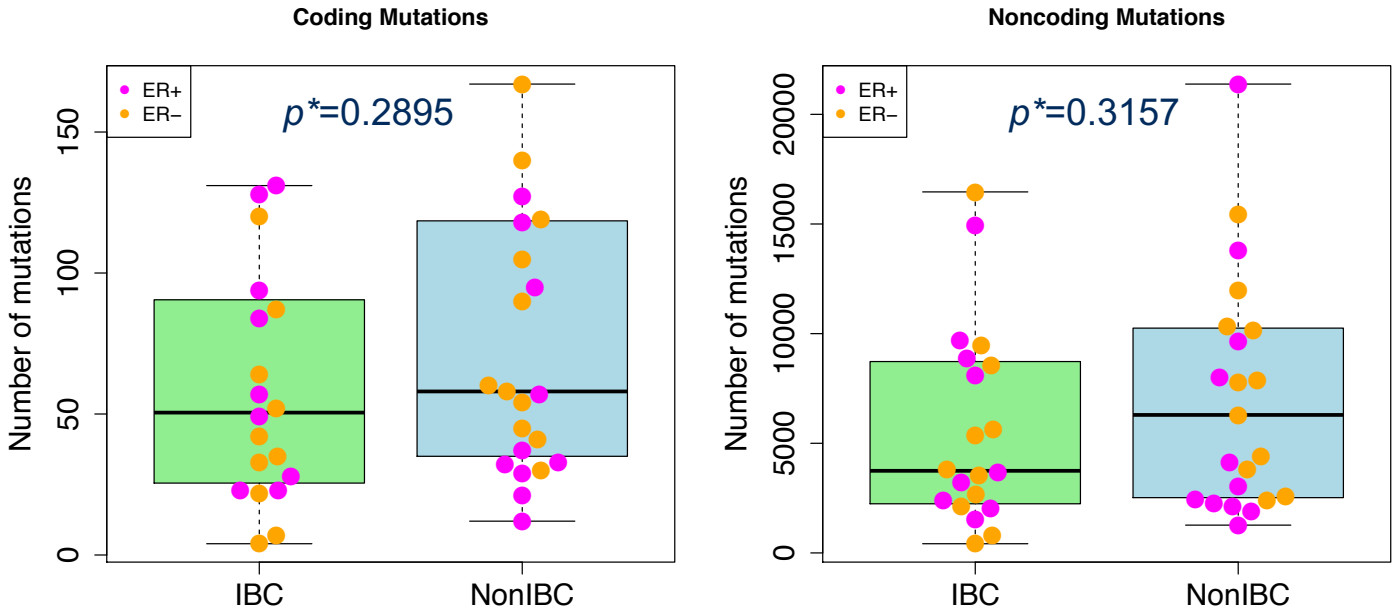


Results

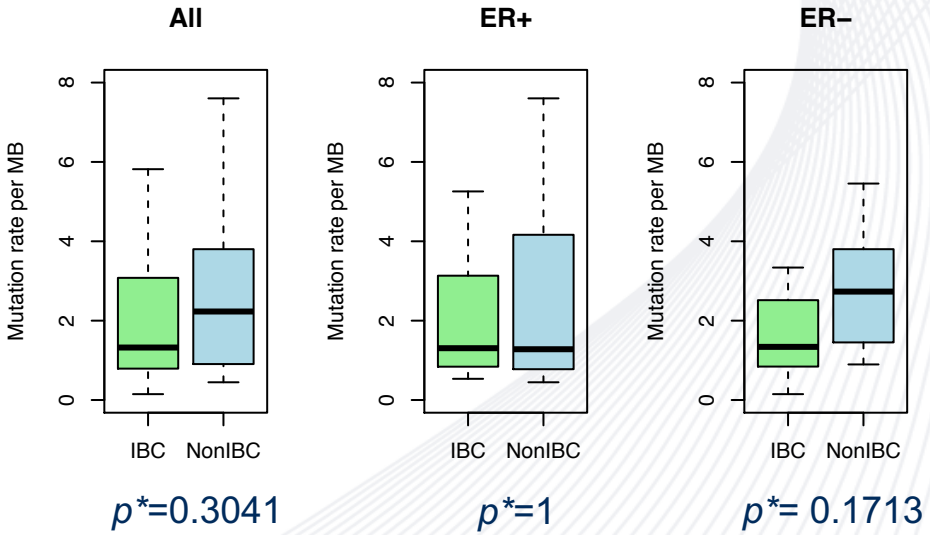
Somatic Single Nucleotide Variants (SNVs)

Identified 114,563 somatic SNVs in the IBC cohort, median 3,789 range: 424 - 16,662 including 1,282 variants (1.12%) in the coding regions

Number of mutations in IBC vs non-IBC



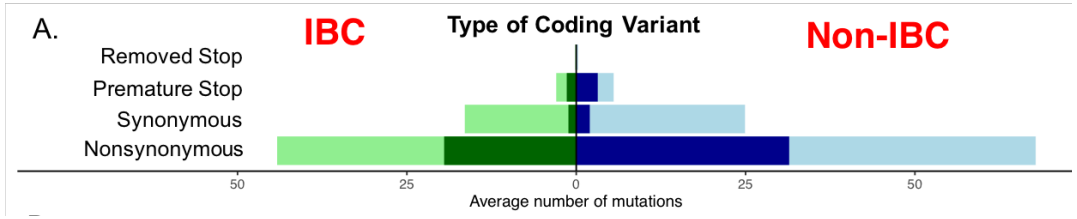
Overall mutation rate by ER status in IBC and non-IBC

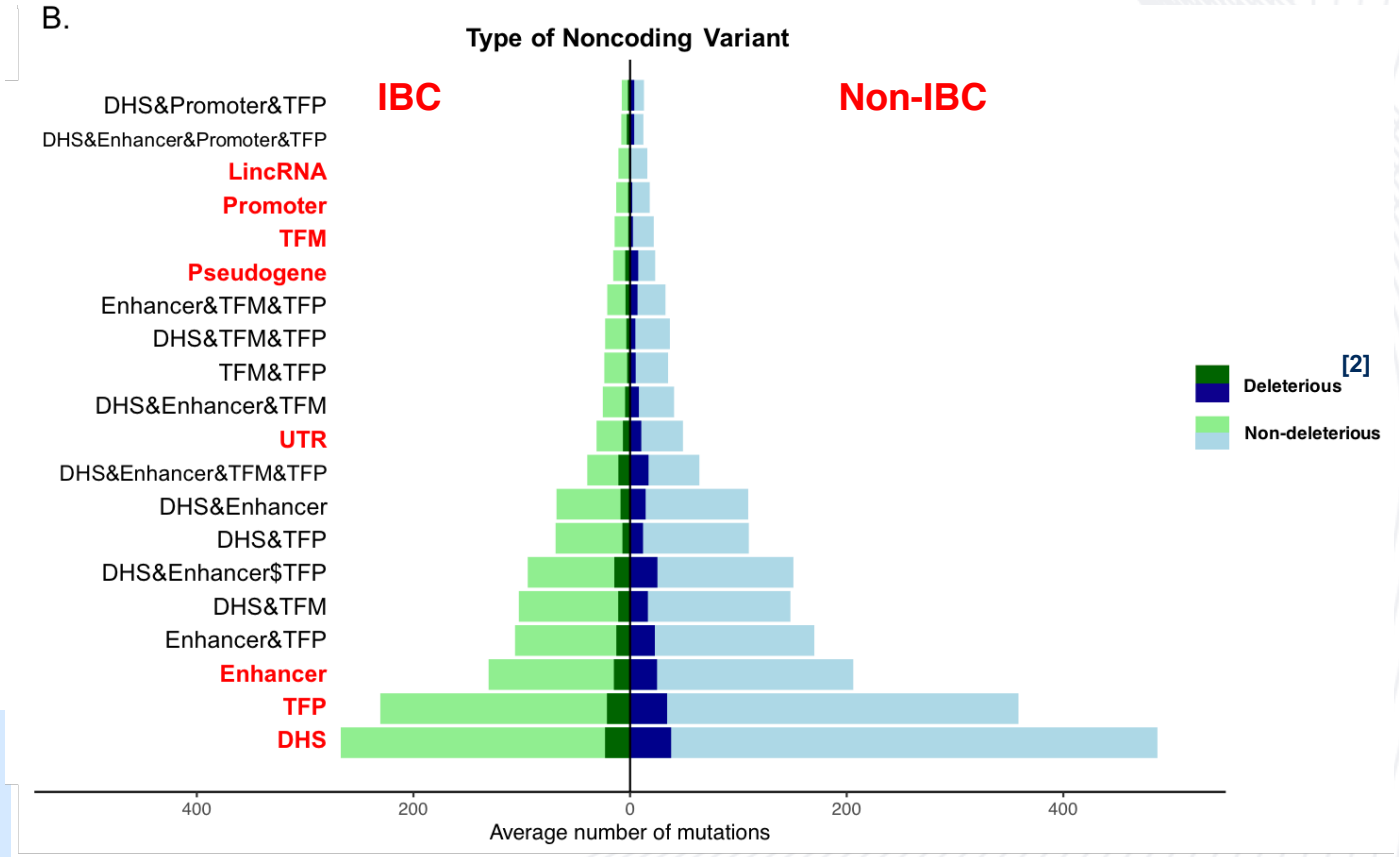


Number of mutations and overall mutational rate are similar between IBC and Non-IBC

* P-values calculated by Wilcoxon rank sum test.

Functional Annotations[1] of Somatic SNVs

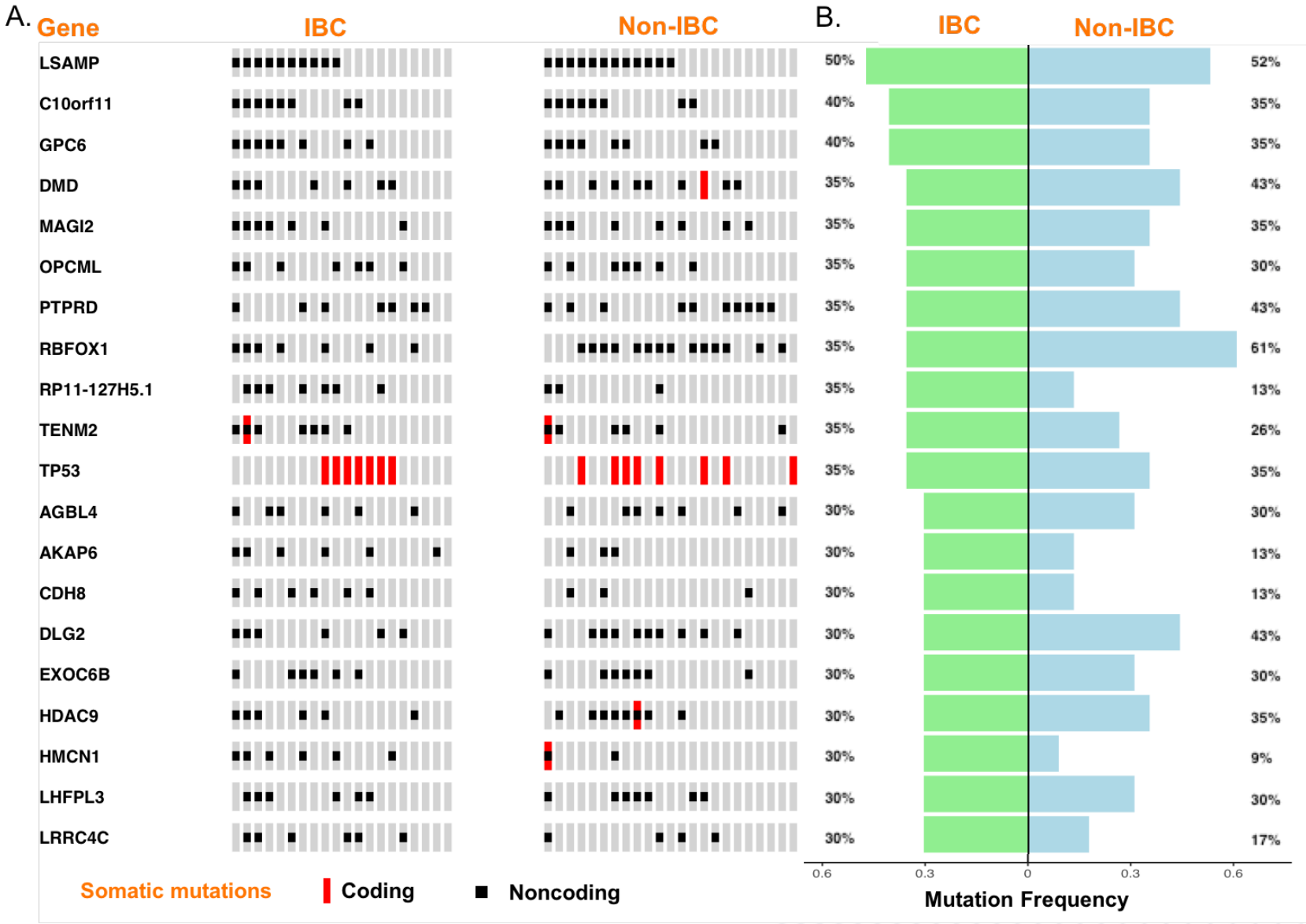




The numbers of coding (A) or noncoding variants (B) within each annotation category are similar between IBC and non-IBC cohorts

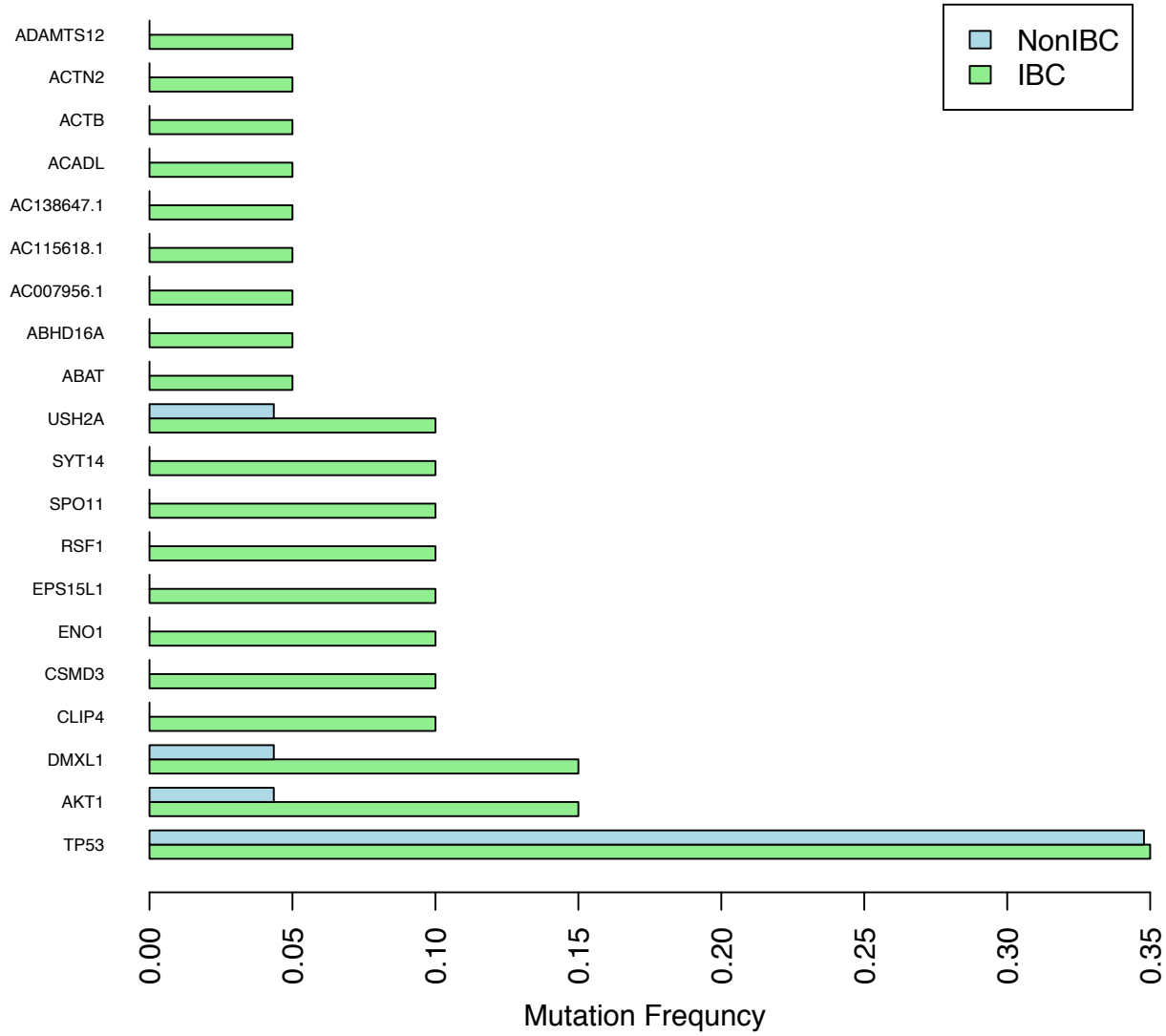
[1] Fu, Yao, et al. Genome biology 15.10 (2014): 480.
[2] Bendl, Jaroslav, et al. PLoS Comput Biol 12.5 (2016)

The top 20 most frequently affected genes by deleterious somatic SNVs in IBC

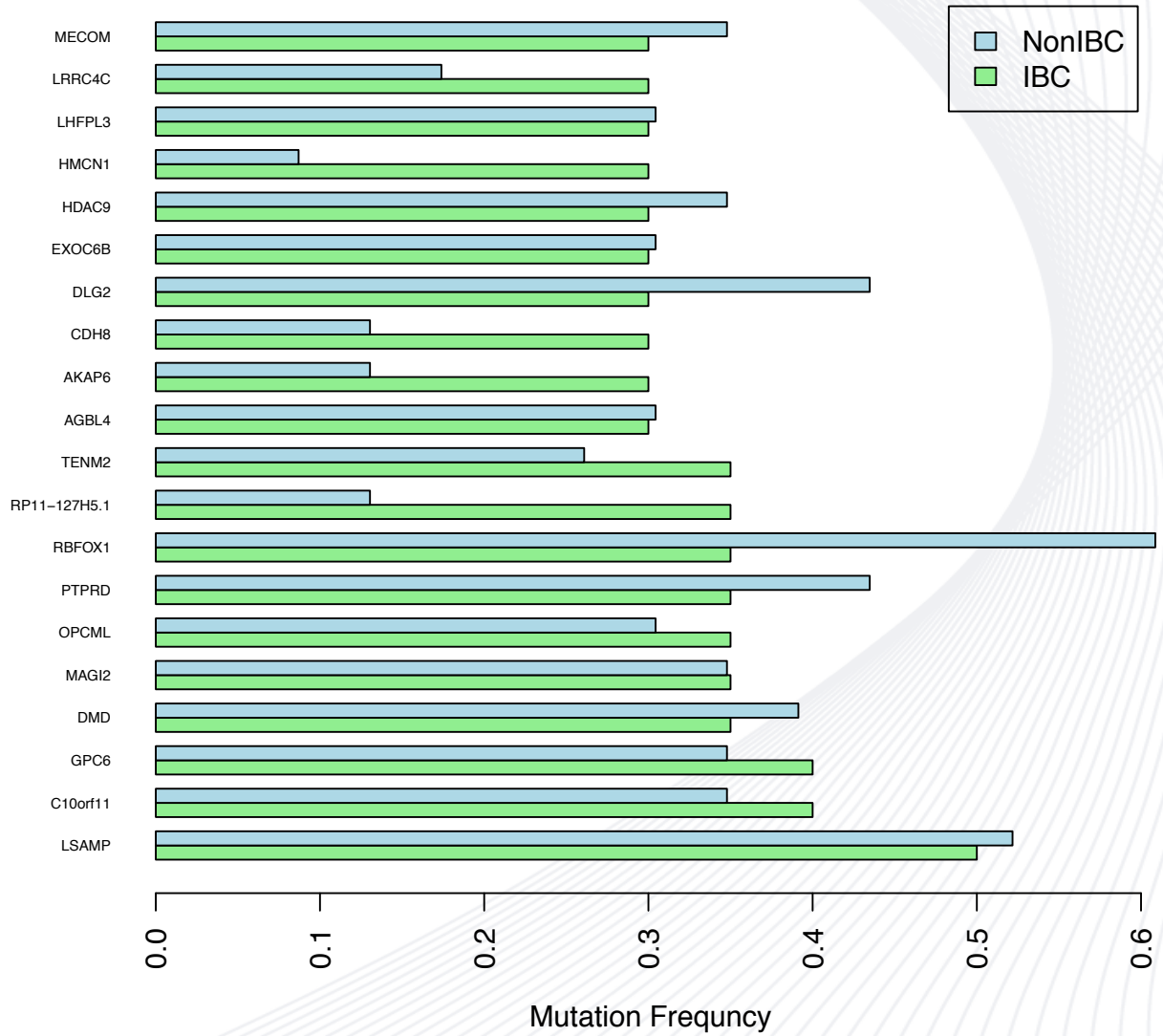


p>0.1
Fisher exact test

Top 20 affected genes in IBC cohort by coding SNVs



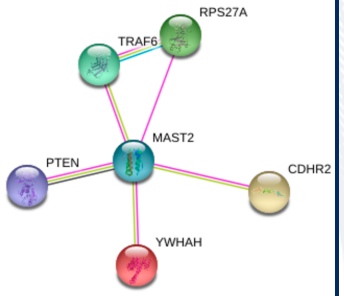
Top 20 affected genes in IBC cohort by noncoding SNVs

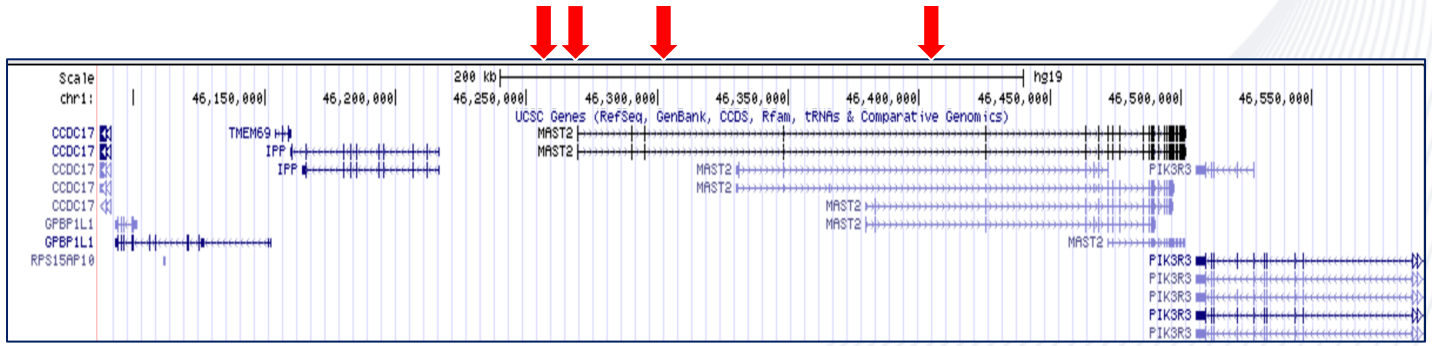


Significantly differentially mutated HFI genes between IBC and Non-IBC

Gene	Cases with mutations in IBC	Cases with mutations in Non-IBC	P-values (before adjustment)
RGL1	0	7	0.010
TENM3	0	7	0.010
HPSE2	2	10	0.019
FAM49B	0	6	0.023
ELP4	1	8	0.024
LDB2	1	8	0.024
MAST2	4	0	0.039
AFF2	2	9	0.039

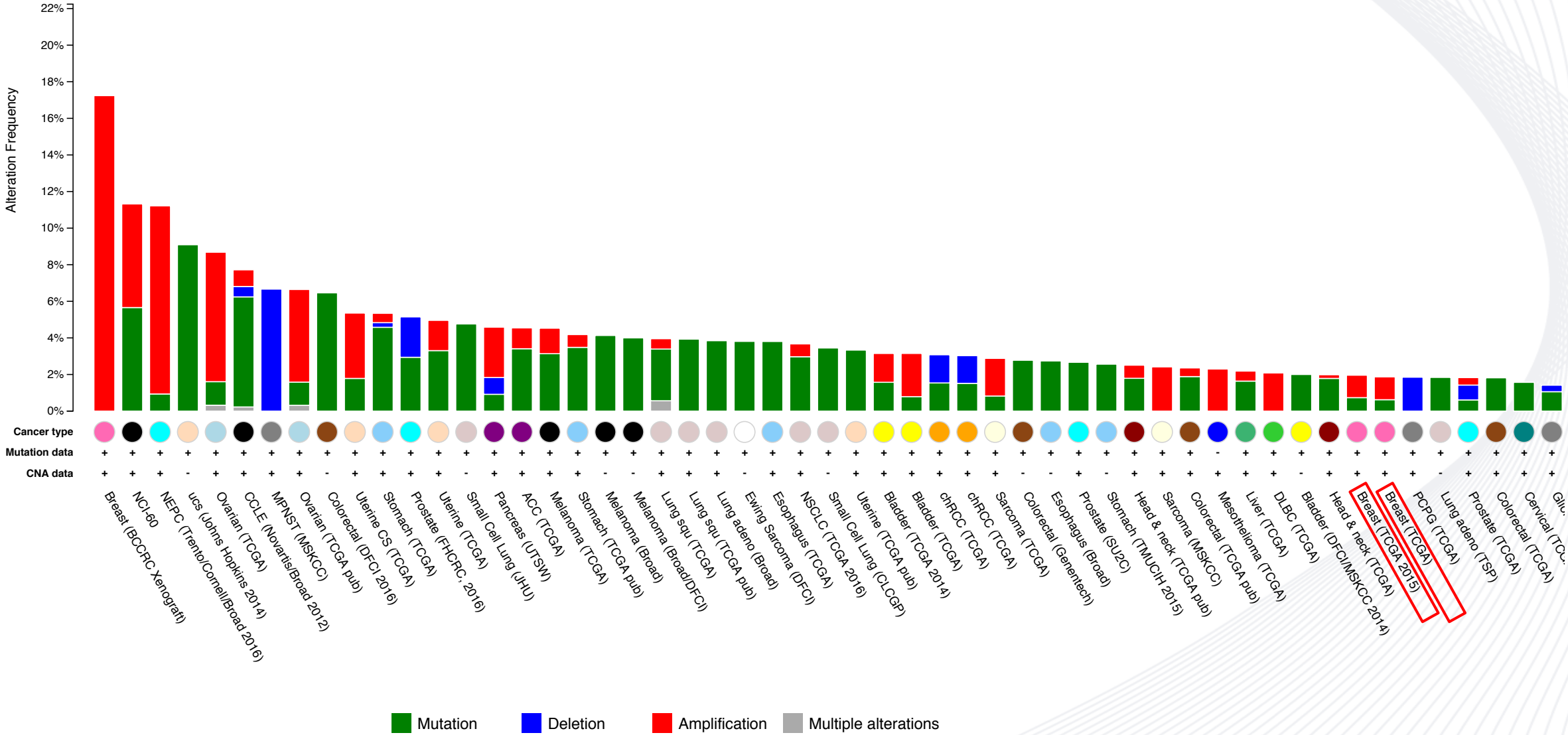
MAST2 = Microtubule Associated Serine/Threonine Kinase 2 MAST2 can interact with Protocadherin LKC and is a new candidate for a tumor suppressor of colon and liver cancers associated with contact inhibition of cell proliferation





The only significantly more frequently mutated gene in IBC was MAST2 in the non-coding region affecting 4/20 (20%) of cases. Each case had a different variant, each predicted to be deleterious.

Frequency of MAST2 alterations across cancer types in the TCGA[1]

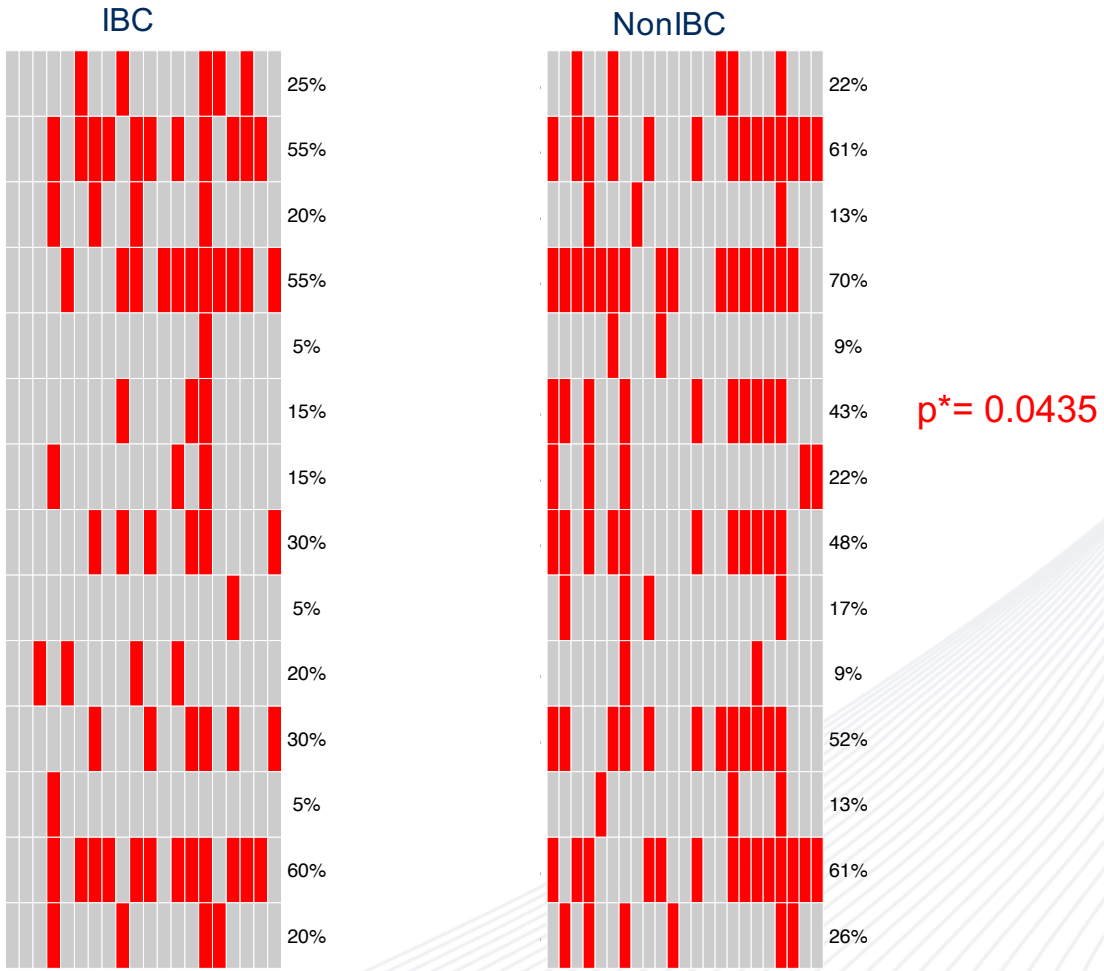


[1] Gao et al. Sci. Signal. 2013 & Cerami et al. Cancer Discov. 2012 & www.cbioportal.org

Canonical cancer pathway alteration by deleterious somatic SNVs

- 14 essential cancer pathways including sets of tumor suppressor genes and oncogenes[1]

- Cell cycle
- Chromatic remodeling
- Differentiation and development
- DNA damage
- Immune regulation
- MAPK & PI3K pathway
- Metabolism
- PI3K pathway
- RAS pathway
- RNA metabolism
- RTK pathway
- TGFB pathway
- Transcription regulation
- WNT signaling



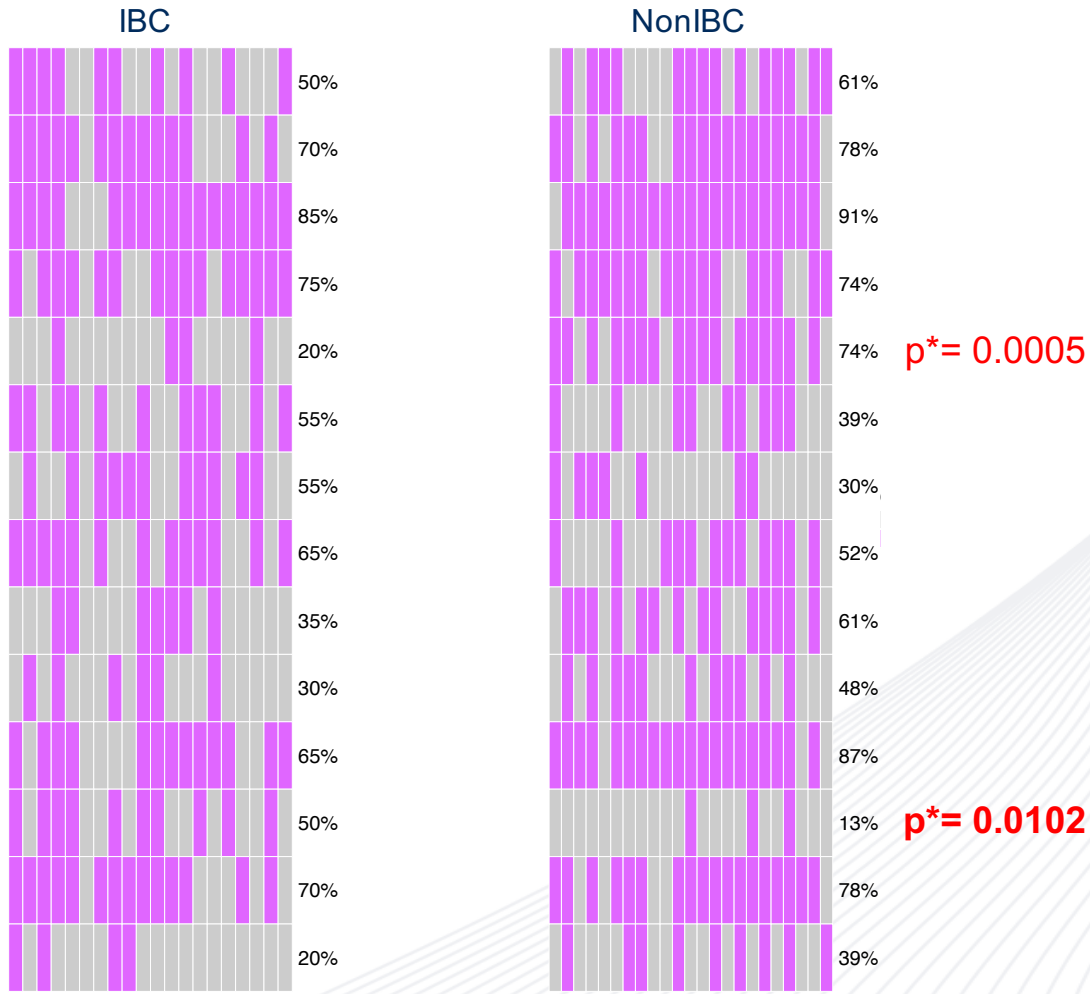
p*= 0.0435

[1] Davoli, Teresa, et al. Science 355.6322 (2017): eaaf8399.
* Fisher exact test (one-side)

Canonical cancer pathway alteration by deleterious germline SNVs

- 14 essential cancer pathways including sets of tumor suppressor genes and oncogenes[1]

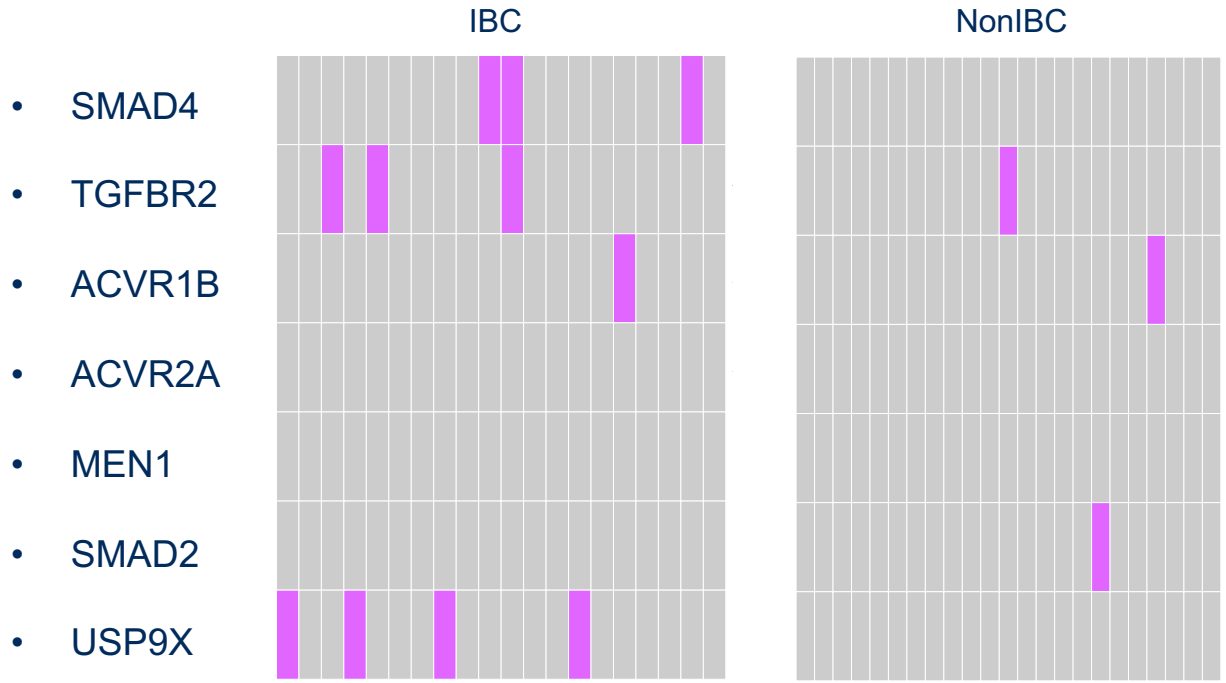
  - Cell cycle
  - Chromatic remodeling
  - Differentiation and development
  - DNA damage
  - Immune regulation
  - MAPK & PI3K pathway
  - Metabolism
  - PI3K pathway
  - RAS pathway
  - RNA metabolism
  - RTK pathway
  - TGFB pathway
  - Transcription regulation
  - WNT signaling



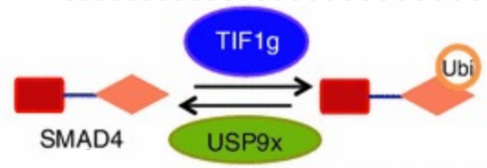
[1] Davoli, Teresa, et al. Science 355.6322 (2017): eaaf8399.
* Fisher exact test (one-side)

TGFβ pathway alterations by deleterious germline SNVs

- **TGFβ signaling pathway**
  - The transforming growth factor-beta (TGF-beta) family members are structurally related secreted cytokines
  - A wide spectrum of cellular functions such as proliferation, apoptosis, differentiation and migration
  - Localized and reversible TGFβ signaling switches breast cancer cells from cohesive to single cell motility[1]
  - TGF-β signaling pathway is suppressed in IBC carcinoma tissues compared to non-IBC[2].
    - Attenuation of TGF-β signaling pathway may contribute to tumor emboli formation and lymphatic invasion of IBC carcinoma cells[2].



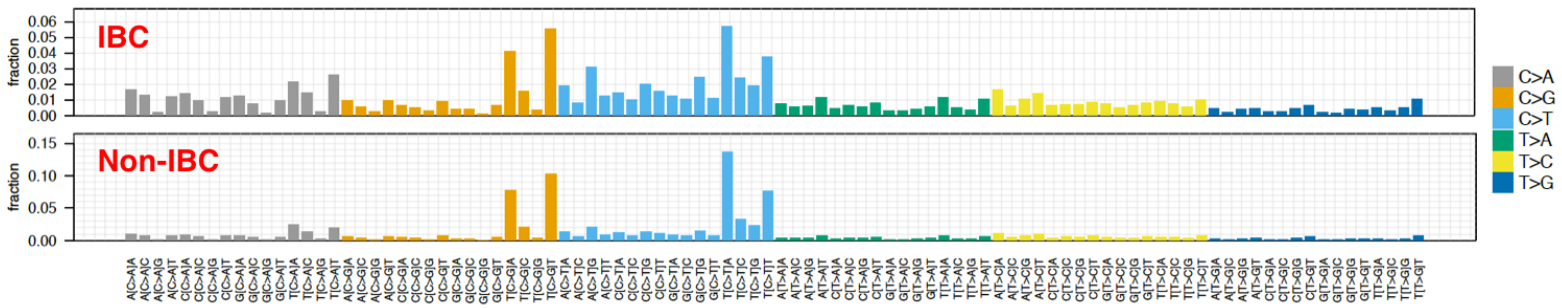
- **SMAD4**
  - **Three IBC samples have different deleterious variants**
  - A member of the Smad family of signal transduction proteins
  - SMAD4 was blocked by SMAD6, which was a repressor of TGFβ signaling and significantly upregulated in IBC[3]
- **USP9X**
  - **Four IBC samples reported 5 deleterious variants, with two of them have identical location**
  - A deubiquitinating enzyme essential for TGFβ signaling
  - Controls SMAD4 mono-ubiquitination[4]



[1] Giampieri, Silvia, et al. *Nature cell biology* 11.11 (2009): 1287. [2] Van Laere S.J., Ueno N.T., et al. Clin Cancer Res. 2013;19(17):4685–4696.
[3] Robertson, F. M. et al. *J. Clin. Exp. Pathol*. **2**, 119 (2012). [4] Zhang, Juan, et al. *Protein & cell* 5.7 (2014): 503-517.

Mutational Signature Analysis

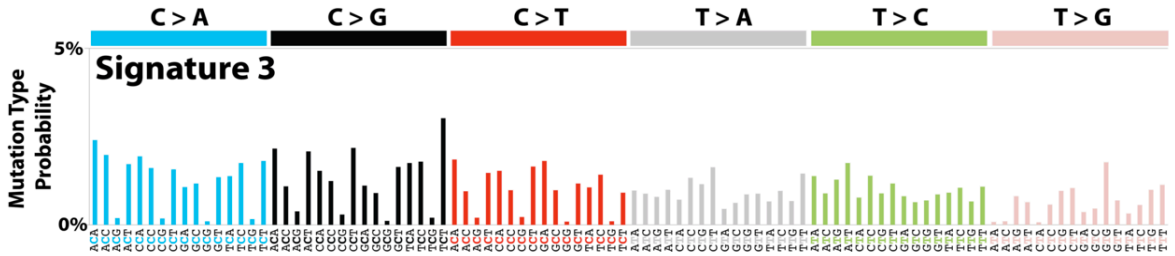
Mutational spectrum of IBC vs. Non-IBC cohorts



- The top four dominant mutation types (C>G and C>T) are shared by two cohorts.
- Mutations in IBC were more broadly distributed.

Different mutational processes generate unique combinations of mutation types, termed "**Mutational Signatures**"

- **COSMIC**[1] delivered **30** validated signatures



- **Found** in breast, ovarian, and pancreatic cancers.
- **Associated** w/ germline and somatic BRCA1/2 mutations
  - failure of DNA double-strand break-repair by HR

Decomposition of mutational spectrum by validated mutational signatures

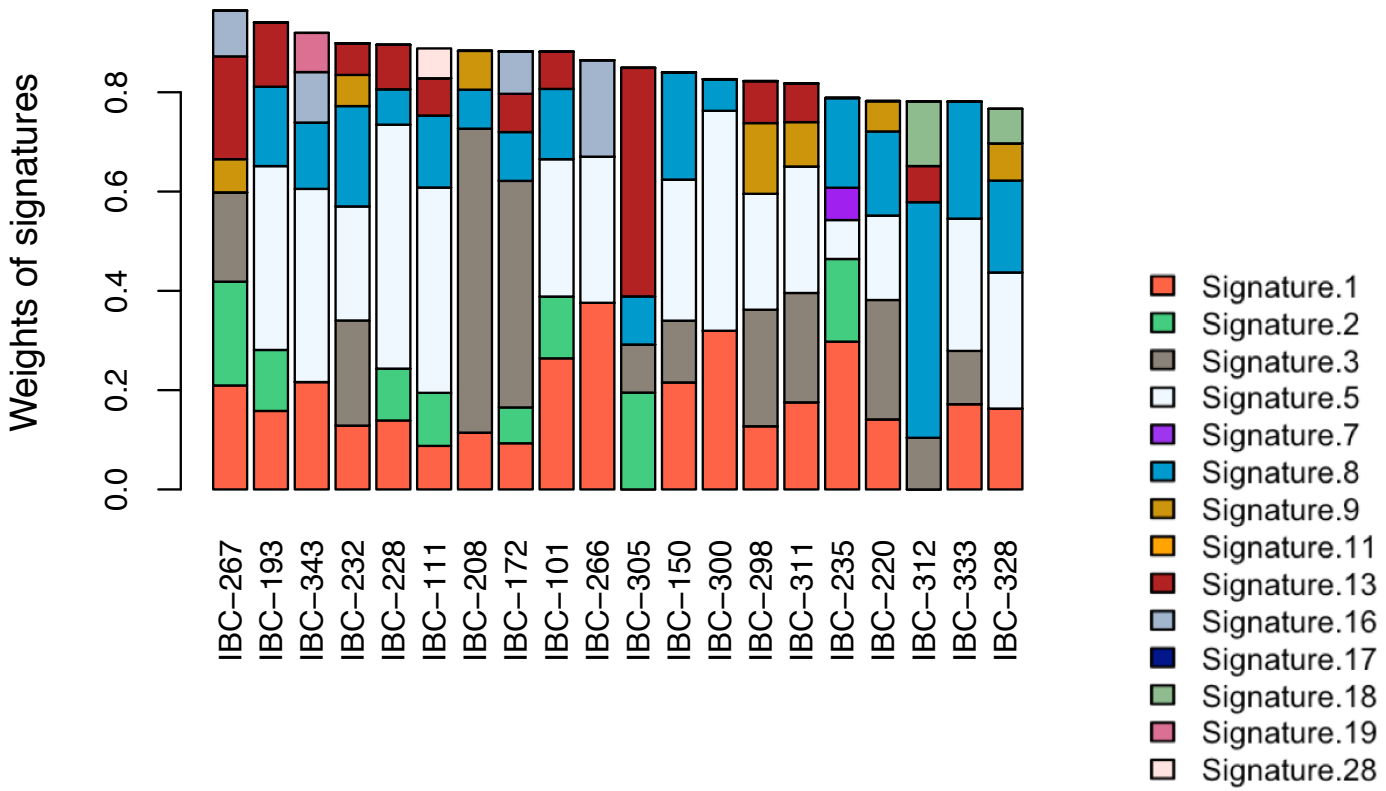
$$M = S \times E$$

Mutation Spectrum — **Validated Signatures** — **Weights/Contributions of each signature**

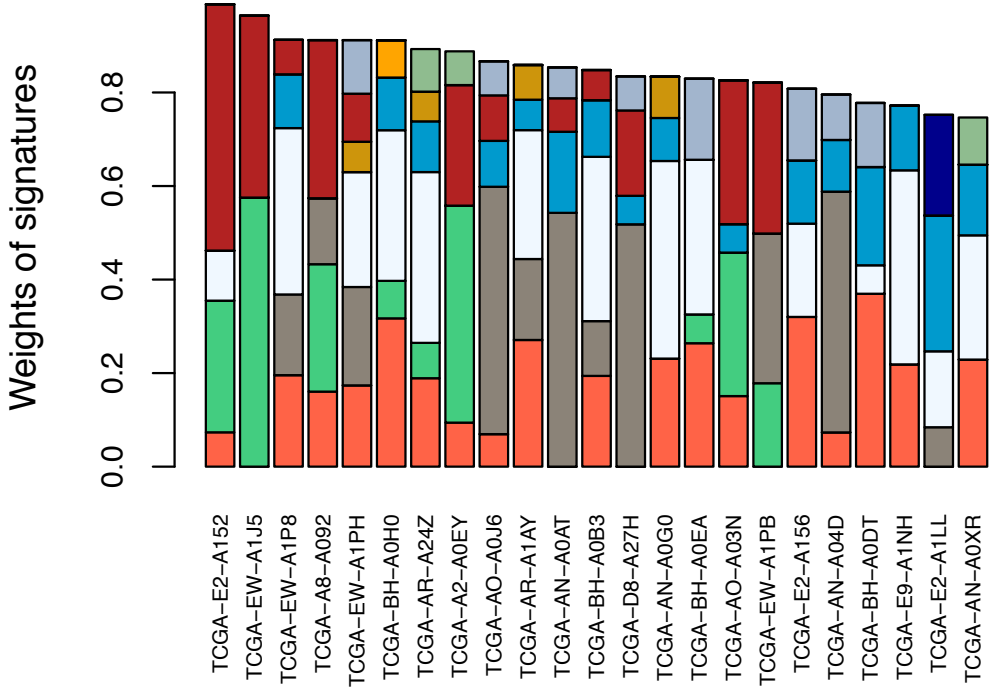
- **Goal:** find optimal E while M and S are known
- **Methods:**
  - Generalized linear model[2]
  - Linear programming/optimization

[1] COSMIC : Catalogue of somatic mutations in cancer
[2] Rosenthal, Rachel, et al. Genome biology 17.1 (2016): 31.

Contributions of signatures in IBC cohort



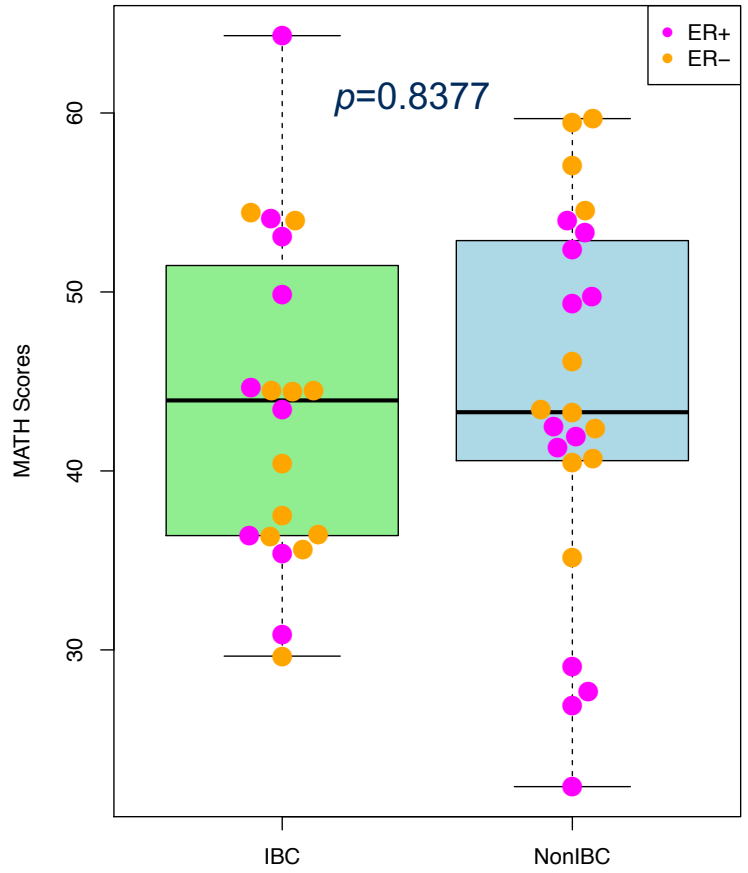
Contributions of signatures in Non–IBC cohort



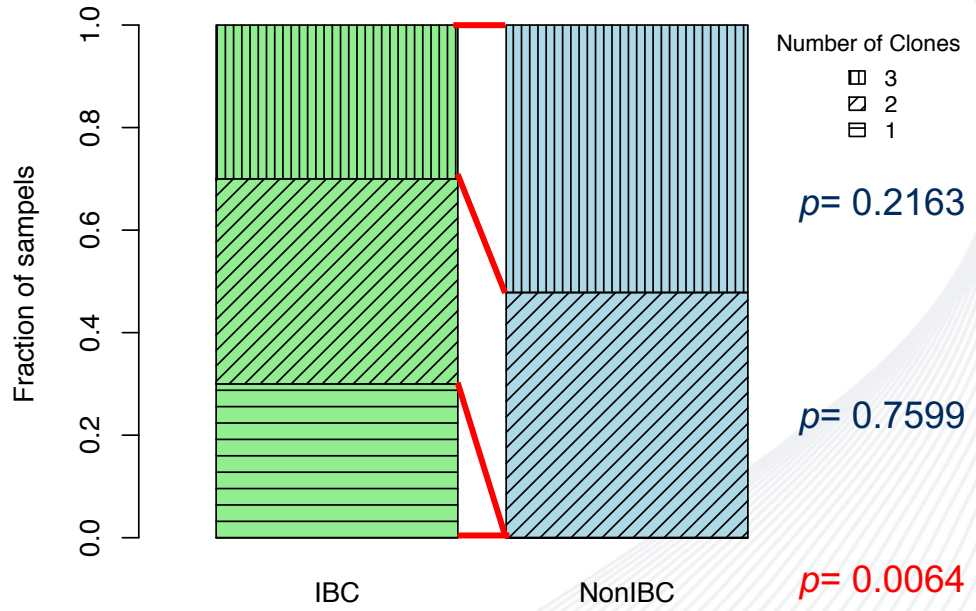
There is no substantial difference in the distribution of mutation signatures in IBC compared to Non-IBC

Clonal Architecture

Mutant-Allele Tumor Heterogeneity (MATH[1]) score of IBC and Non-IBC



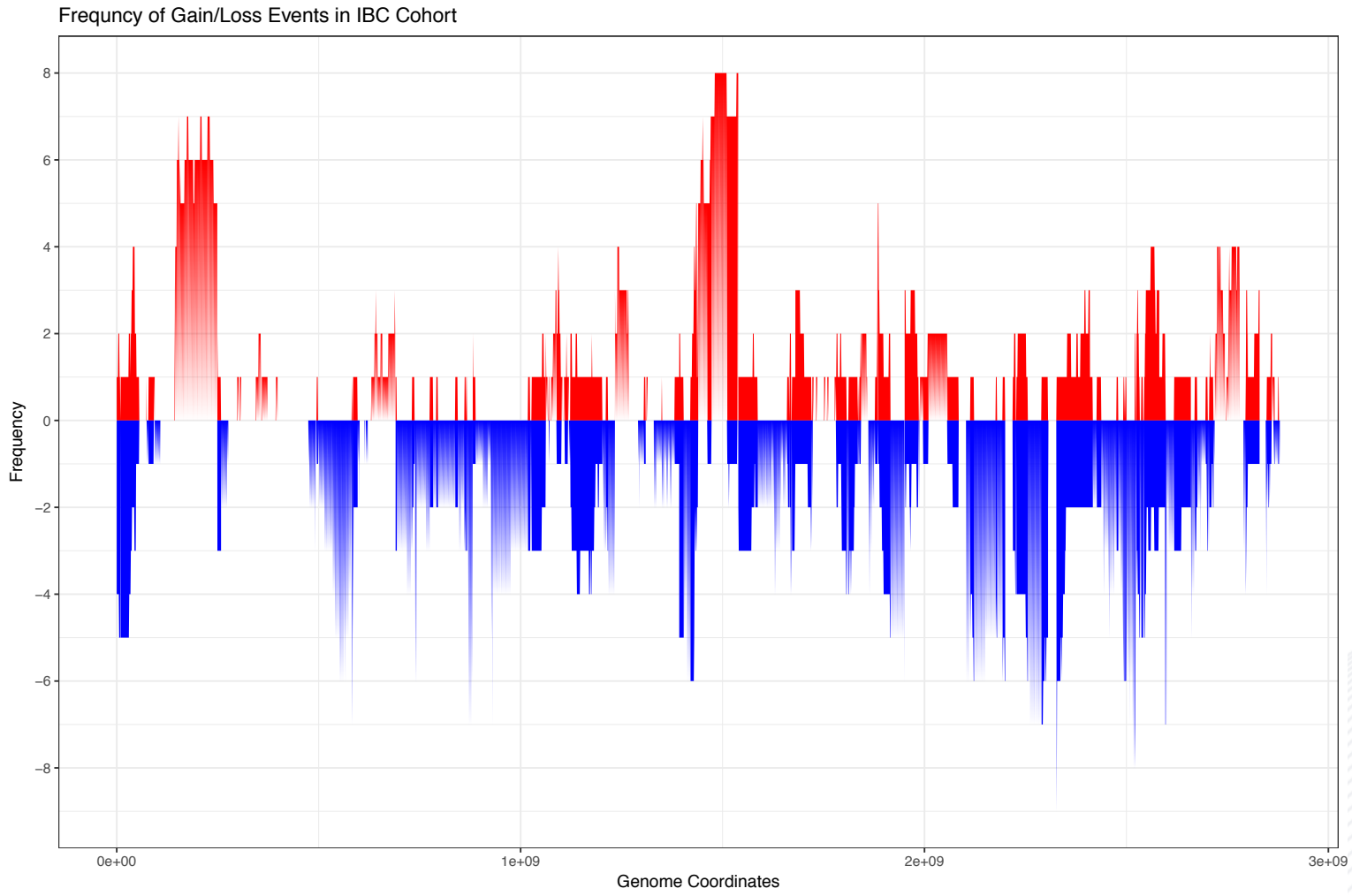
Number of cell cones estimated by SciClone[2] in IBC and non-IBC



There is no difference in overall mutational heterogeneity, but there seems to be a lower clonality in IBC

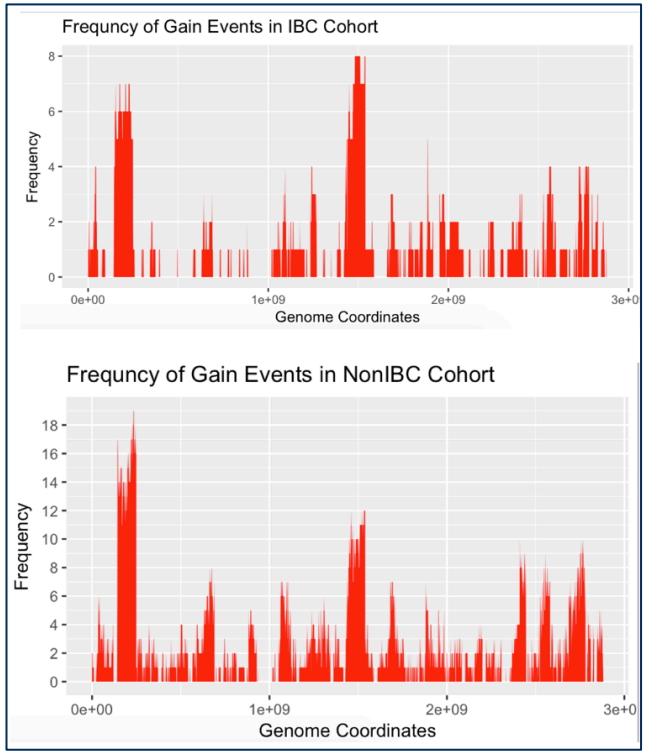
[1] Mroz, Edmund A., and James W. Rocco. Oral oncology 49.3 (2013): 211-215.
[2] Miller, Christopher A., et al. PLoS Comput Biol 10.8 (2014): e1003665.

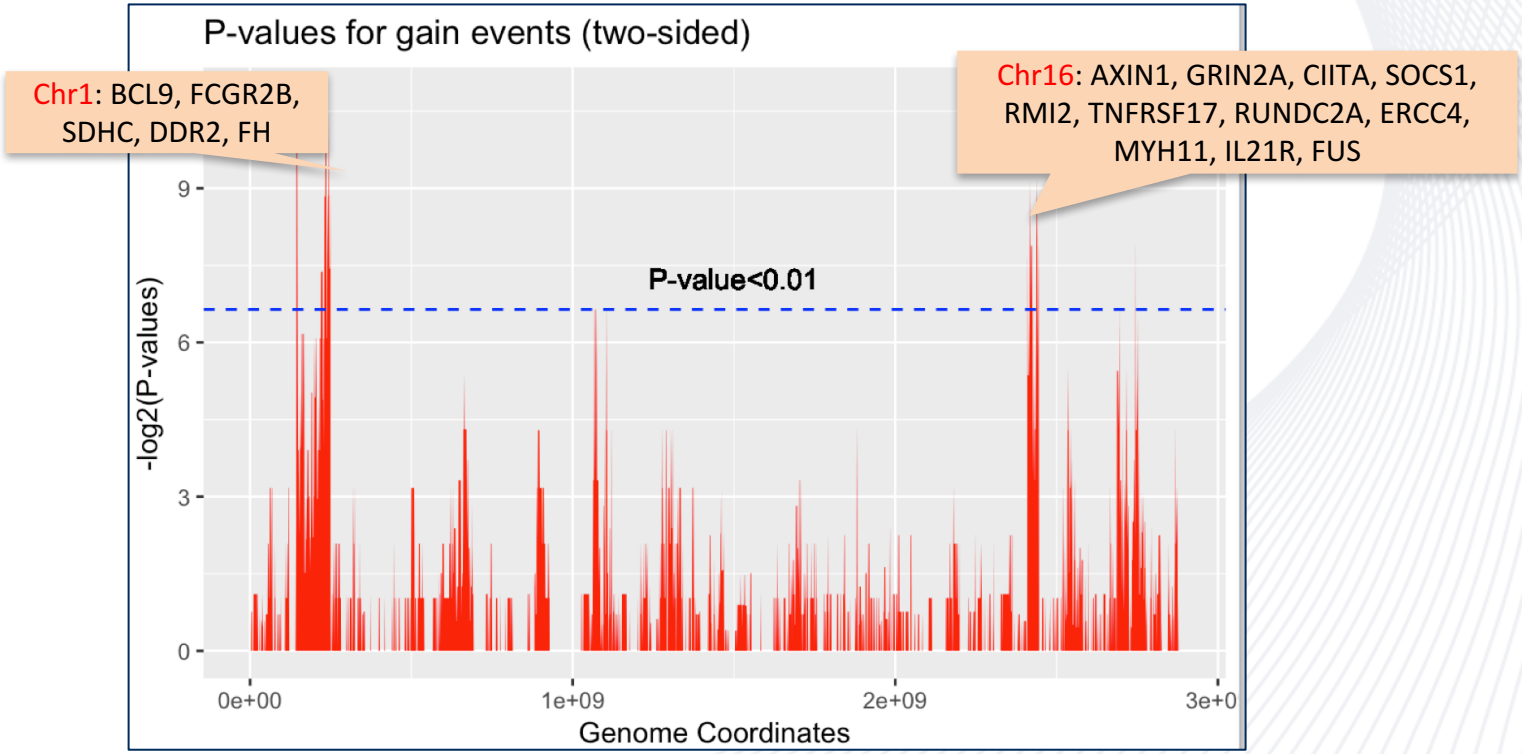
Profile of somatic CNV events in IBC cohort



Somatic gain events profiles: IBC vs. NonIBC

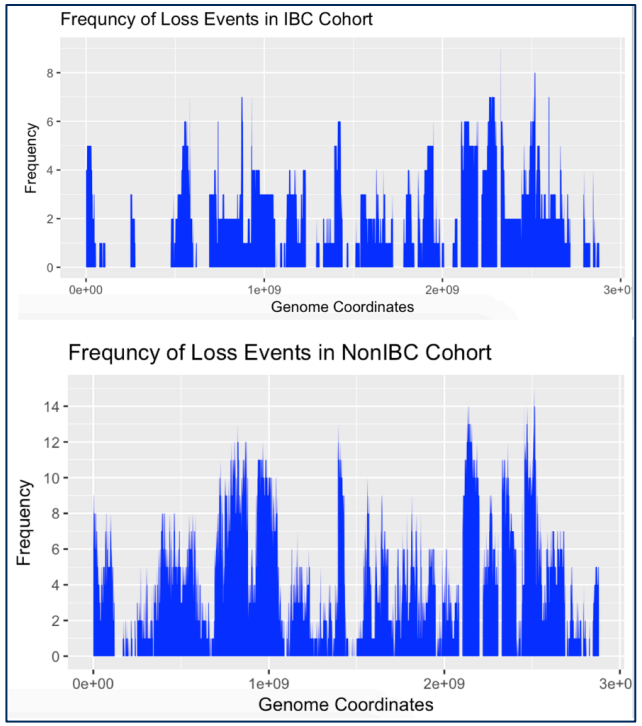
- For each 1Mb bin across the entire genome, compare the frequency of gain events in two cohorts; implement the fisher's exact test

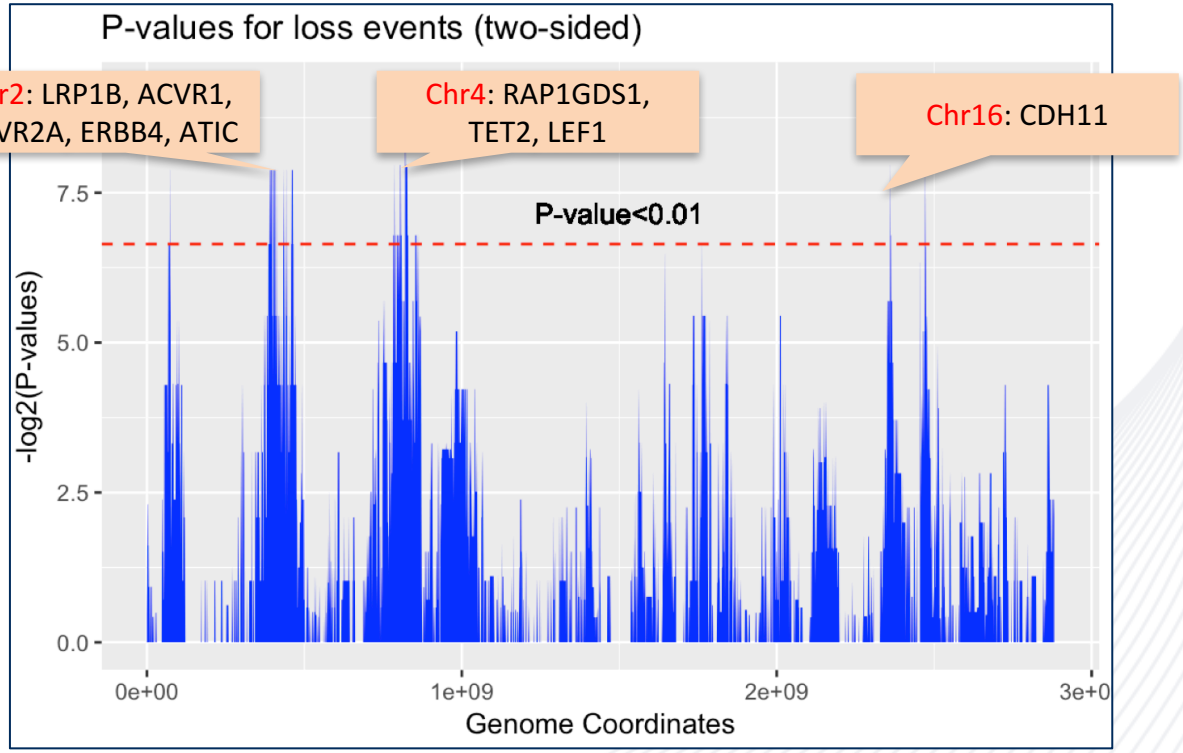




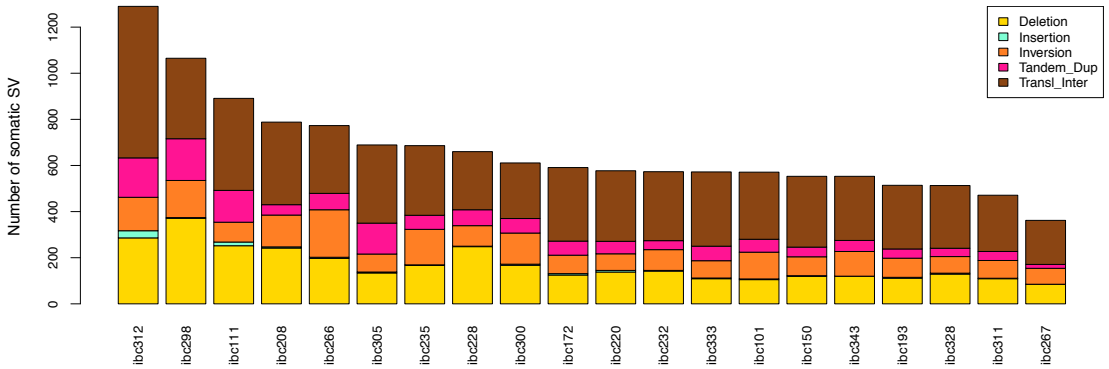
Somatic loss events profiles: IBC vs. NonIBC

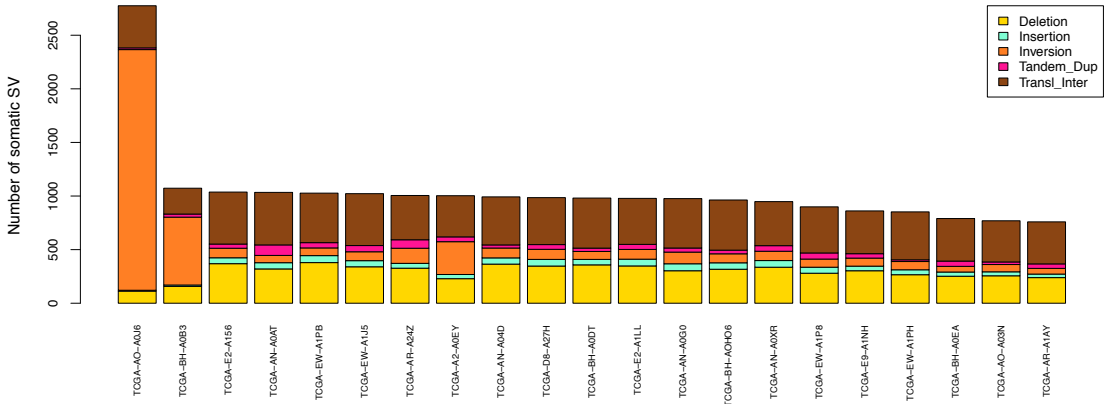
- For each 1Mb bin across the entire genome, compare the frequency of loss events in two cohorts; implement the fisher's exact test

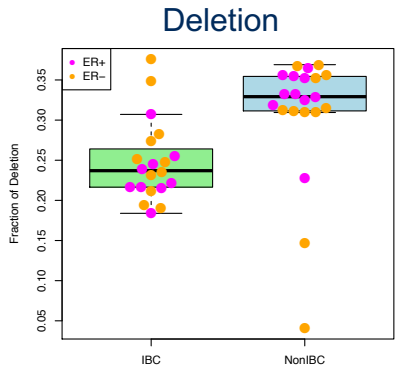


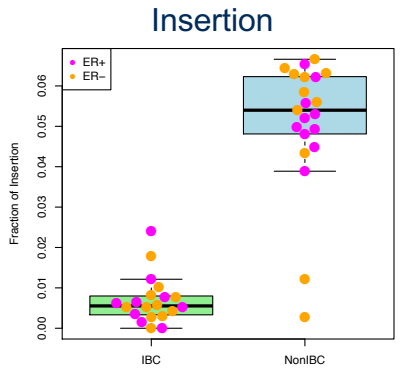


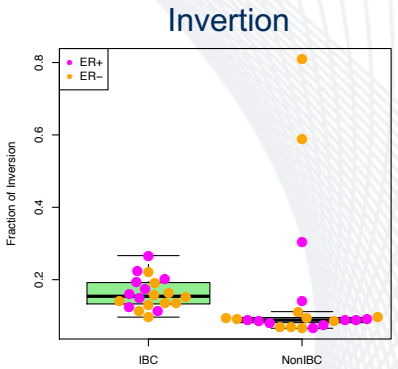
Types of somatic structural variants: IBC vs. NonIBC

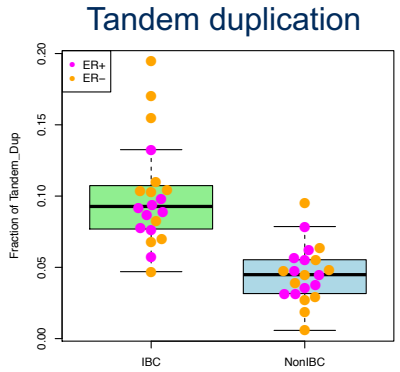


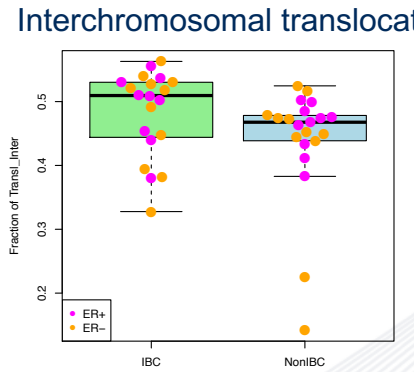








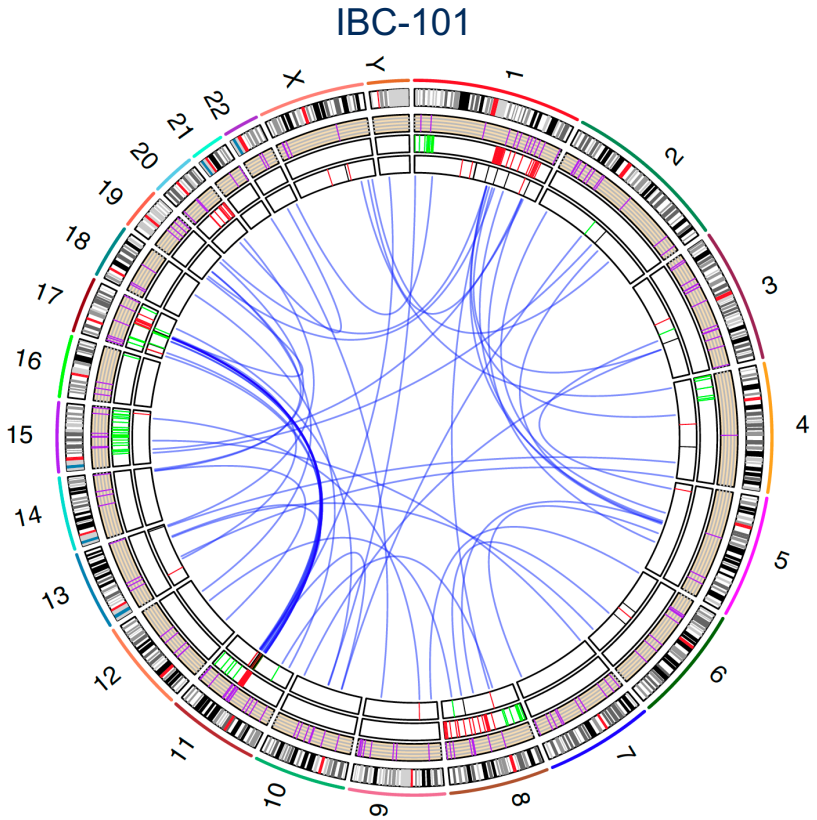


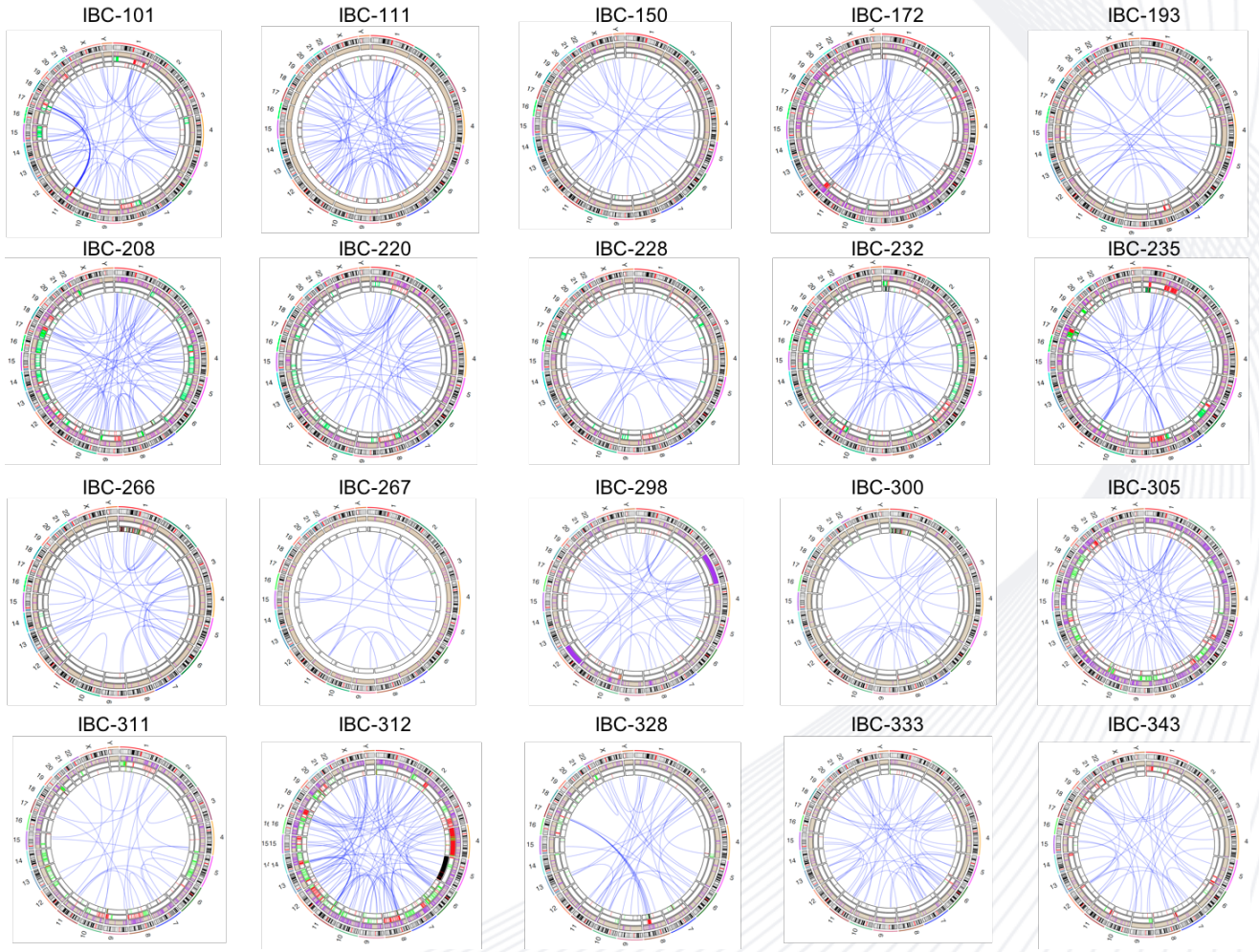


Wilcoxon test	
Deletion	7.3E-04
Insertion	6.7E-07
Inversion	1.1E-04
Tandem_Dup	1.1E-07
Transl_Inter	3.3E-02

IBC has higher fractions of complex structural variants, while lower fractions of simple structural variants.

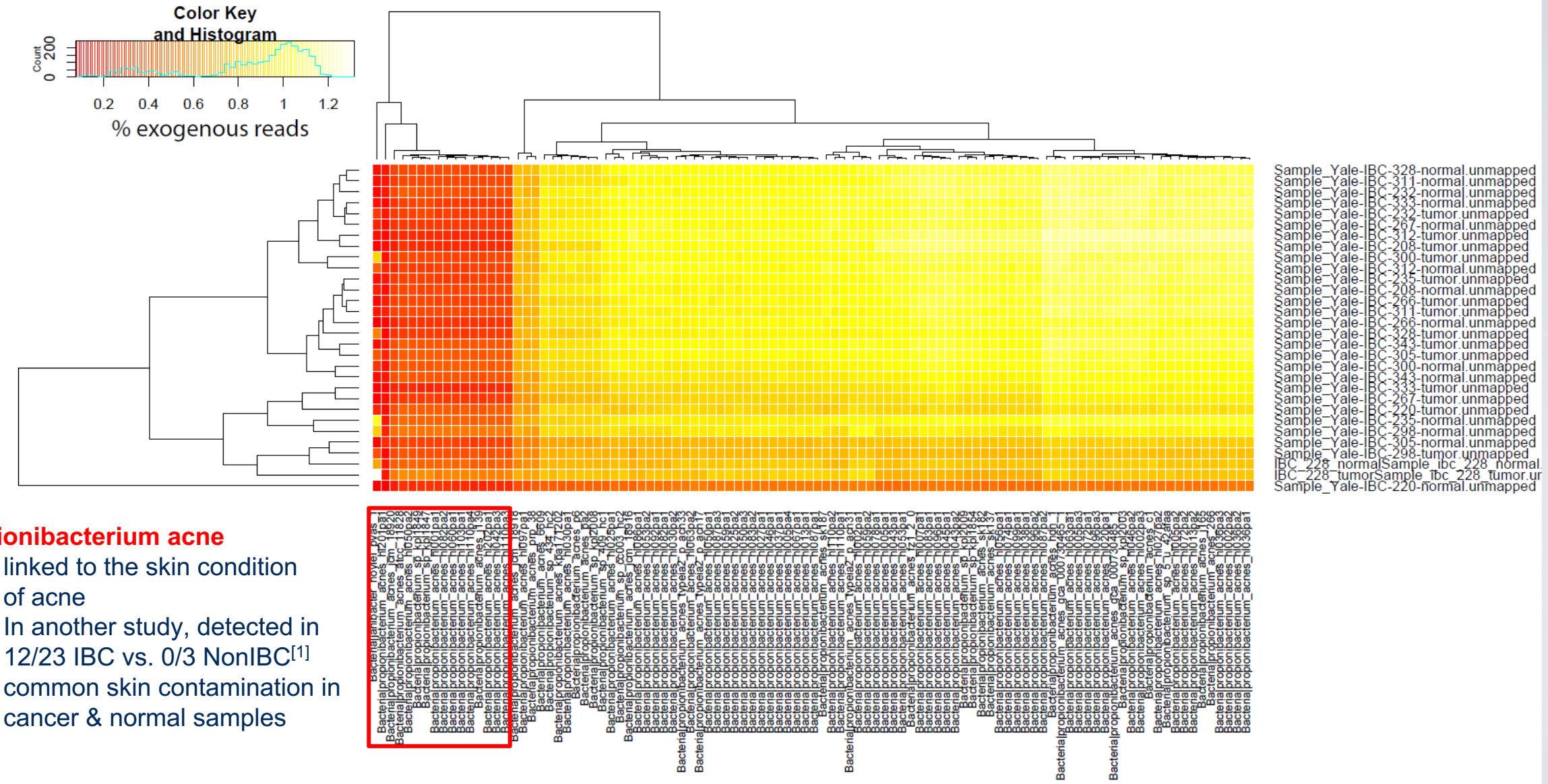
Landscape of somatic variants in IBC whole-genome sequences





- From outer to inner rings, each represents:
  - chromosome ideogram (HG19.Human.CytoBand)
  - SNV : purple
  - CNV: green->loss; red->gain
  - SV: green->deletion; red: tandem duplication; black -> Inversion; blue: Interchromosomal translocation

Top 100 most frequent non-human sequences in the IBC cancer and normal DNA



- **Propionibacterium acne**
  - linked to the skin condition of acne
  - In another study, detected in 12/23 IBC vs. 0/3 NonIBC[1]
  - common skin contamination in cancer & normal samples

[1] Fernandez, S. V., et al. (2013): Cancer Research P6-12.

Summary of genomic difference between IBC vs. NonIBC

Genomic features	Existence of significant difference
Single nucleotide variants (SNV), insertions/deletions (indel)	MAST2 is more frequently mutated in IBC
Structural variations (SV)	More complex SVs in IBC while overall loads are similar
Copy number variations (CNV)	CN loss: chr2, chr4 and chr16; CN gain: chr1 and chr16
Germ-line polymorphisms (SNP)	USP9X is more frequently mutated in IBC
Mutation signatures	No
Clonal composition	No; maybe a lower clonality in IBC
Bacteria or non-human genome	common skin contamination in IBC cancer & normal samples
Canonical cancer pathway-level alterations	TGFβ pathway is significantly more mutated in IBC

Conclusions

- The overall mutation load, genomic heterogeneity and mutation signatures are similar between IBC and non-IBC.
- However,
  - Some genes maybe more frequently mutated in IBC, e.g. our lead candidate is MAST2
  - A few of canonical cancer pathways differentially mutated in IBC vs. NonIBC.
    - TGFβ pathway is significantly more mutated in IBC by germline variants
  - There are significantly different frequency of copy number changes in IBC vs. NonIBC
    - copy number losses in Chr2, Chr4 and Chr16
    - copy number gains in Chr1 and Chr16
  - Complex structural variants more frequently appear in IBC
    - Interchromosomal translocation & Tandem duplication

- We find no plausibly pathogenic non-human infectious agents in the IBC genome. Propionibacterium acne seems to be a common skin contamination in both normal and cancer samples.

Acknowledgement

Gerstein Lab

- Dr. Mark Gerstein
- Sushant Kumar
- Arif Harmanci
- Shantao Li
- Robert Kitchen
- Yan Zhang

Pusztai Lab

- Dr. Lajos Pusztai
- Dr. Christos Hatzis
- Vikram Wali

MDACC IBC Program

- Dr. Naoto T. Ueno
- Dr. Savitri Krishnamurthy
- Dr. Sangeetha Reddy
- Dr. Wendy Woodward
- Dr. James Reuben