ANNACHIARA KORCHMAROS
PEIJUE ZHANG
CHEN GUO
MENGNAN ZHANG

REPORT

October 25, 2016

OVERVIEW

Scope

We analyzed the variants for BRCA1 gene in individual NA12878 to see if she has risk to develop breast cancer. The main reason why we chose to focus on BRCA1 gene is that 55 – 65 % of women who inherit a harmful BRCA1 mutation will develop breast cancer by age 70 years (1,2).

Results

No mutations were identified. This means that there are no pathogenic or likely pathogenic genetic variants associated with an increased risk of breast cancer in the gene tested. In particular, we found an overall of 34 variants. 32 of them are no pathogenic and 2 of them, at position *1196821* and *41256089*, do not have clinical annotations. Since *119682* has depth below 50, we ignore it. On the another, *41256089* has depth 86, therefore we suggest further analysis on this region.

Methodology

- The variants were called by our pipeline. For further details, refer to NA12878_variants.vcf in the supplementary material.
- In order to create highly accurate variant set, we run GATK genome recalibrator on the variants file. For further details, refer to GATK Genome Recalibrator in the supplementary material.
- Match the variant set found in the previous step with BRCA Exchange database which contains the clinical variants information for BRCA1. For further details, refer to BRCA Exchange database BRCA1 in the supplementary material.

- Calculate the variants coverage. For further details, refer to Fig1. in the supplementary material.
- Plot the distribution of depths across BRCA1 and choose depth threshold. For further details, refer to Fig2. in the supplementary material.

Limitations

In Table 1 we show the coverage for only 5 variants. The reason for this incomplete information is in our pipeline step. We used only some data provided by Genomic in a Bottle Project.

SUPPLEMENTARY MATERIAL

- Table 1: Results

CHROM	POSITION	REF	ALT	PATHOGENICITY	Depth	Coverage
chr17	41196408	G	A	Benign	57	59
chr17	41196821	CTTT	C	-	12	6
chr17	41197274	C	A	Benign	69	78
chr17	41199913	T	C	Benign	5	NA
chr17	41203325	T	A	Benign	7	NA
chr17	41215825	C	T	Benign	18	NA
chr17	41219560	C	T	Benign	43	NA
chr17	41219780	T	C	Benign	39	NA
chr17	41219804	T	C	Benign	19	NA
chr17	41223094	T	C	Benign	93	103
chr17	41224833	G	C	Benign	2	NA
chr17	41226601	G	C	Benign	29	NA
chr17	41226675	A	T	Benign	12	NA
chr17	41231221	A	C	Benign	44	NA
chr17	41231516	C	T	Benign	134	NA
chr17	41231902	G	A	Benign	3	NA
chr17	41232344	G	C	Benign	3	NA
chr17	41234470	A	G	Benign	106	119
chr17	41243190	T	G	Benign	28	NA
chr17	41244000	T	C	Benign	107	NA
chr17	41244435	T	C	Benign	68	NA
chr17	41244936	G	A	Benign	61	NA
chr17	41245237	A	G	Benign	56	NA
chr17	41245466	G	A	Benign	92	NA
chr17	41248164	C	T	Benign	4	NA
chr17	41249363	TA	T	Benign	23	NA
chr17	41251646	T	A	Benign	9	NA
chr17	41254486	T	G	Benign	4	NA
chr17	41256089	AAAAAAAAAGAAAAG	A	-	86	NA
chr17	41257134	T	C	Benign	2	NA
chr17	41270277	C	T	Benign	2	NA
chr17	41276247	A	G	Benign	18	NA
chr17	41276348	T	C	Benign	10	NA
chr17	41277187	G	C	Benign	5	NA

- NA12878_variants.vcf
  http://vannberglab.biology.gatech.edu/data/ahcg2016/vcf/NA12878_variants.vcf

- GATK Genome Recalibrator:

  recalibrating SNPs in exome data
  https://software.broadinstitute.org/gatk/gatkdocs/org_broadinstitute_gatk_tools_walkers_variantrecalibration_VariantRecalibrator.php

  filtering variants

  https://software.broadinstitute.org/gatk/gatkdocs/org_broadinstitute_gatk_tools_walkers_variantrecalibration_VariantRecalibrator.php

- BRCA Exchange database BRCA1

  http://vannberg.biology.gatech.edu/data/ahcg2016/BRCA/BRCA1_brca_exchange_variants.csv

- Figures

  Fig1: Coverage depth across BRCA1
  The vertical lines represent the 5 variants whose coverage is known.

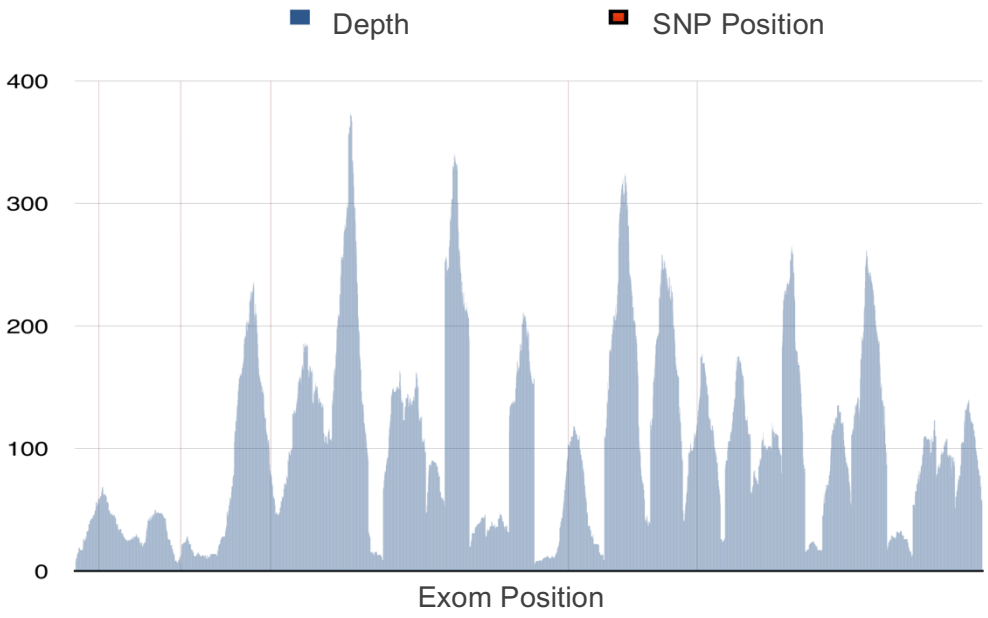
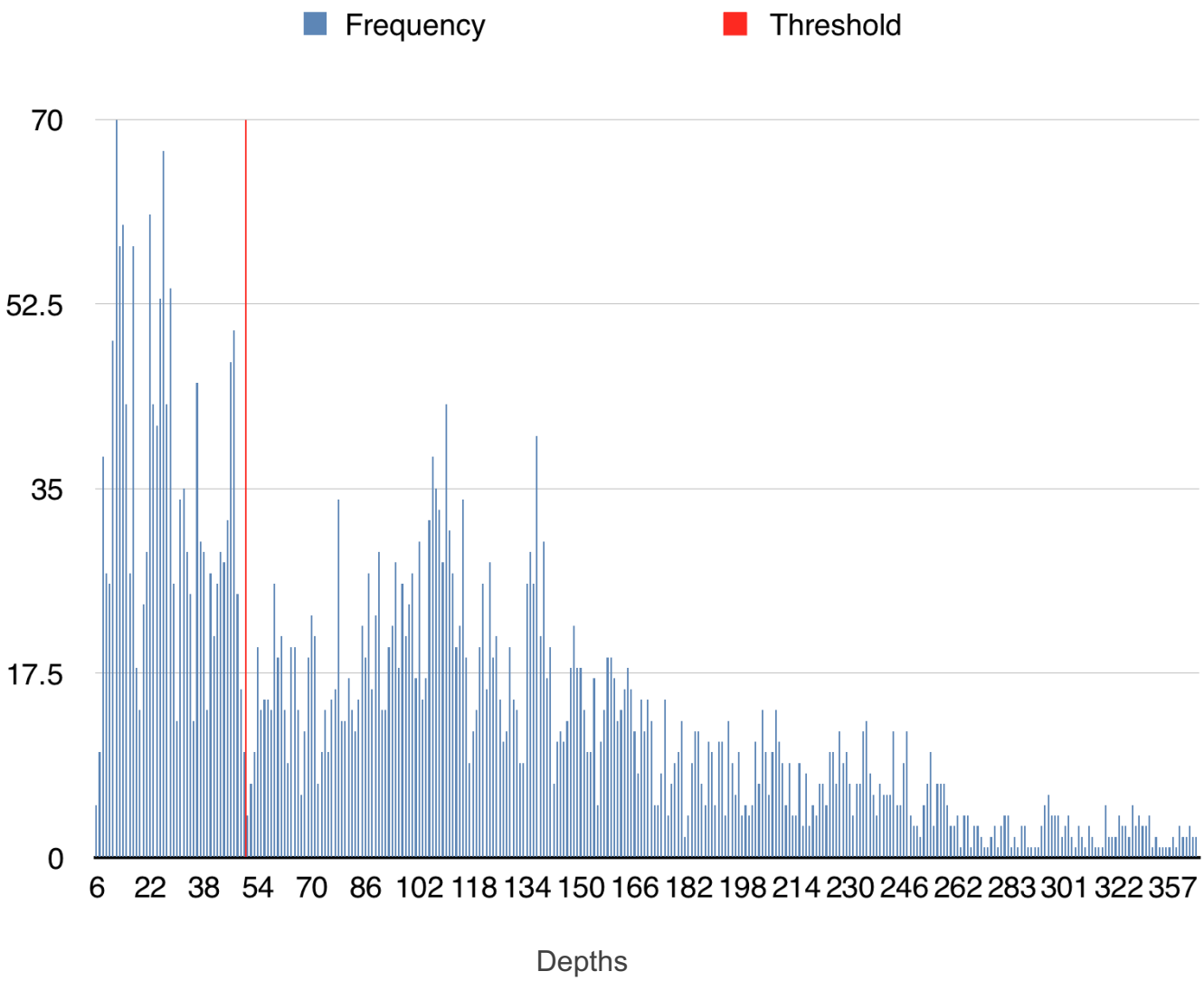


Fig2: Distribution of depths across BRCA1

The red line represents the threshold (depth = 50). We chose this threshold, because the depths smaller than 50 do not follow a normal distribution as expected. You can also use depth=37 as cutoff, where 37 is the first quartile. However, this does not change our results.



REFERENCE

1. Antoniou A, Pharoah PD, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: A combined analysis of 22 studies. *American Journal of Human Genetics* 2003; 72(5):1117–1130.

2. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. *Journal of Clinical Oncology* 2007; 25(11):1329–1333.