Analysis of Tumor Suppressor Mutations and Their Impact in Breast Cancer

Group Number- 6

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY
DELHI

Problem Statement

This project aims to analyze mutations in key tumor suppressor genes associated with breast cancer (eg- BRCA2) across different populations and evaluate their functional impacts. Using publicly available cancer genomics databases, we will identify mutation patterns in these genes, compare their prevalence across diverse human populations, and determine how these genetic alterations affect protein structure and function. The study will also investigate the differential effects of these mutations across cancer types. This comprehensive approach will enhance our understanding of how tumor suppressor gene mutations contribute to breast cancer development.

Breast Cancer Overview

How Breast Cancer Forms?

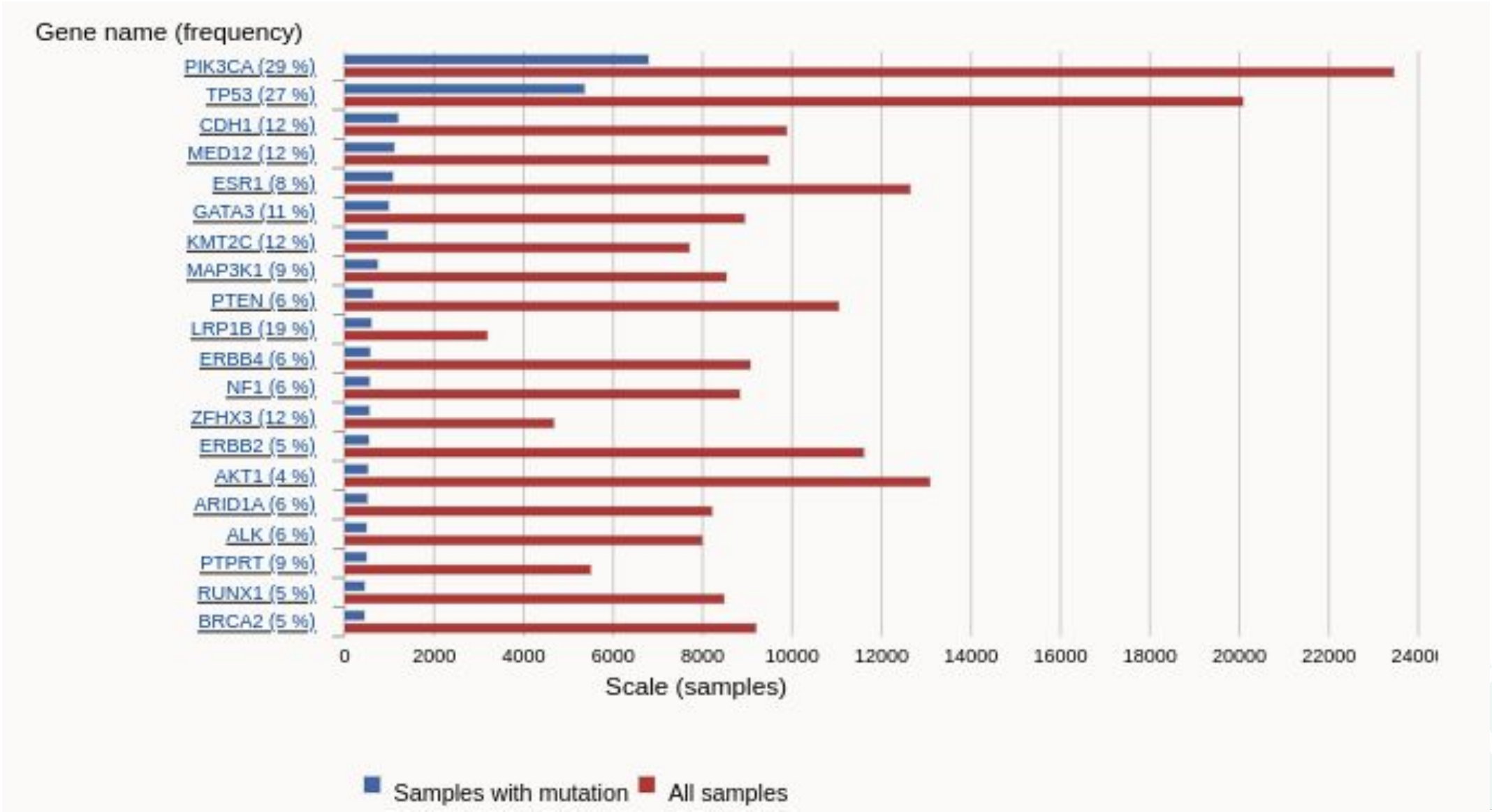
- Invasive breast cancer starts with hyperplasia (cell population increase)
- Next is atypical hyperplasia (cells show abnormalities)
- Non-invasive stage called carcinoma in situ (not spread to nearby tissue) either in milk ducts or lobules (DCIS or LCIS)
- Further genetic changes evolve them into invasive carcinomas which can spread to other tissues via blood or lymph. (This is the primary cause of death)

Somatic Mutations

Impacts of such mutations

- Less than 10% of all breast cancer cases can be accounted by genetic factors. Primary academic focus is on heritable (germline) mutations
- However critical somatic mutations in genes like BRCA1, BRCA2 and TP53 have been linked to increased risk
- We found that if we plot mutation frequency vs position, we **do not** get a normal distribution, rather mutations in some specific positions are much much more frequent than others.
- We suspect these are loss / gain of function mutations

Mutation Distribution (All Types)



Samples collected from COSMIC

Positive Data

This section shows an exportable table of positive mutation data for the selected gene or sample.
Note: The number of mutated samples displayed may not appear to match the number shown in the distribution pie charts/tables. The number of samples tested on this page include samples from the targeted screens only.

This section shows an exportable table of the postive data (mutated samples) for the selected gene or sample. You can see more information on the help pages.

Gene BRCA2
Primary site Breast
Mutation type Substitution nonsense
Mutated Samples 49
Samples Tested 6513

Show 10 entries Export: CSV TSV Search:

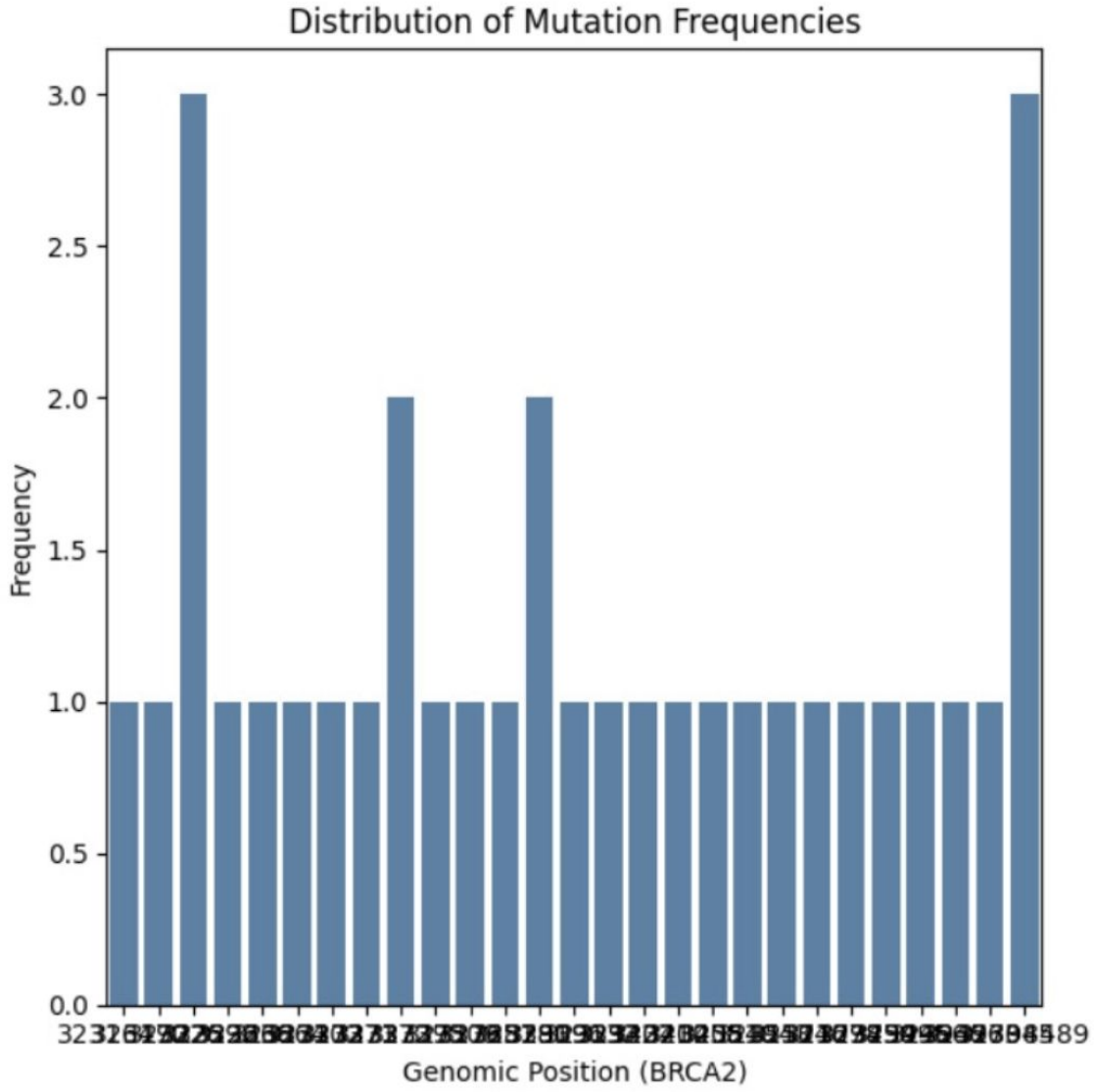
Gene Name	Transcript	Census Gene	Sample Name	Sample ID	AA Mutation	CDS Mutation	Primary Tissue	Tissue Subtype 1	Tissue Subtype 2	
BRCA2	ENST00000380152.7	Yes	P-0016773-T01-IM6	2830871	p.G4*	c.10G>T	Breast	NS	NS	C
BRCA2	ENST00000380152.7	Yes	P-0005900-T01-IM5	2829918	p.Q73*	c.217C>T	Breast	NS	NS	C
BRCA2	ENST00000380152.7	Yes	P-0005900-T01-IM5	2722282	p.Q73*	c.217C>T	Breast	NS	NS	C
BRCA2	ENST00000380152.7	Yes	MBC_45	2662748	p.E654*	c.1960G>T	Breast	NS	NS	C

Mutation Frequency Matrix

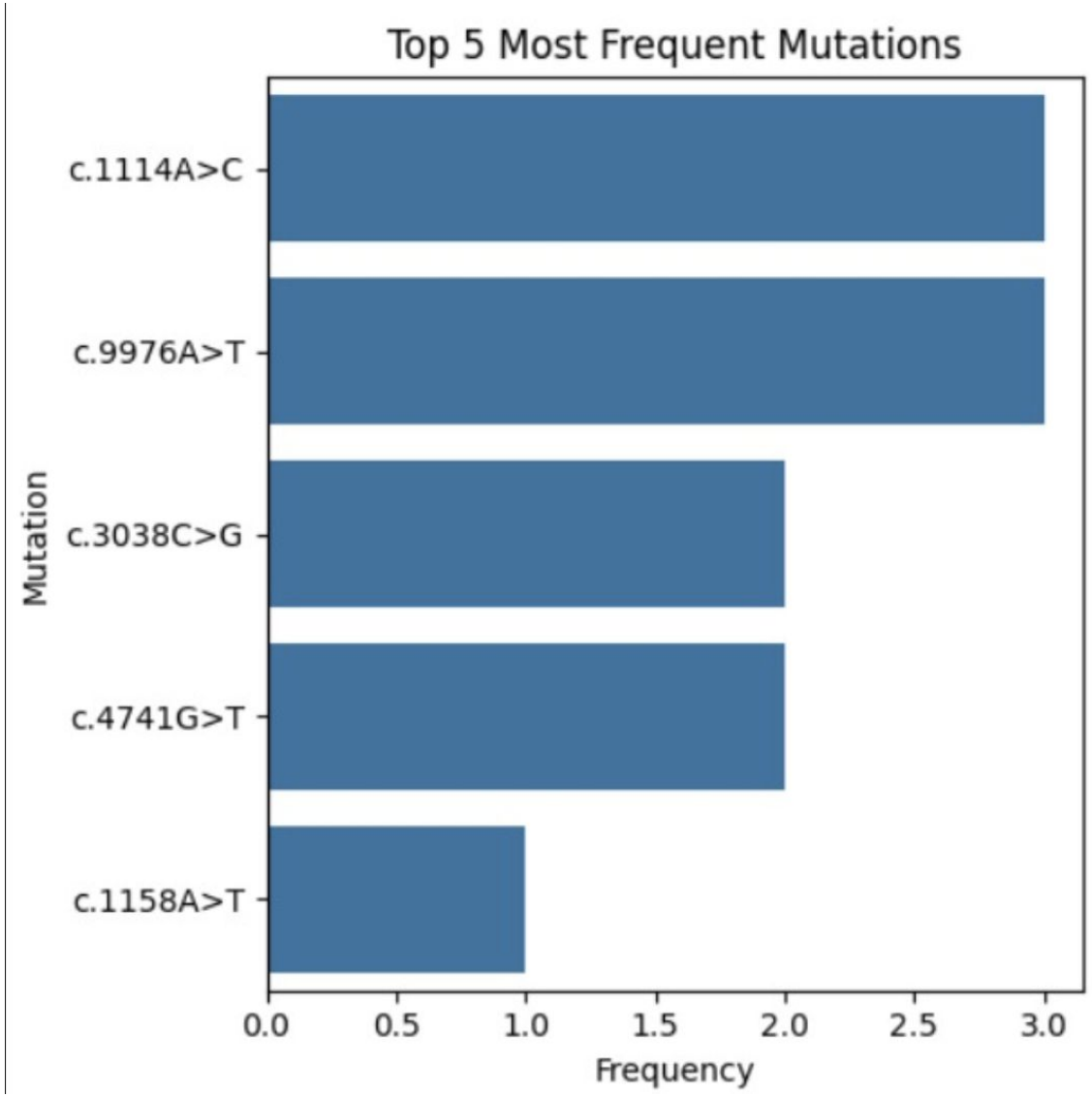
=== Mutation Frequency Matrix ===

Mutation	P-0016773-T01-IM6	2974352	2687075	P-0007685-T01-IM5	P-0004824-T01-IM5	PMSNH1002	D
32316470c.10G>Tp.G4*	1	0	0	0	0	0	
32319226c.217C>Tp.Q73*	0	0	0	0	0	0	
32332592c.1114A>Cp.N372H	0	0	0	0	0	0	
32332636c.1158A>Tp.E386D	1	0	0	0	0	0	
32336864c.2509_2512delinsCAATp.E837_K838delinsQ*	0	0	0	1	0	0	
32337302c.2947C>Ap.P983T	1	0	0	0	0	0	
32337371c.3016G>Tp.G1006*	0	0	0	0	0	1	
32337372c.3017G>Tp.G1006V	0	0	0	0	0	1	
32337393c.3038C>Gp.S1013*	0	0	0	0	0	0	
32337509c.3154G>Ap.A1052T	1	0	0	0	0	0	
32337651c.3296C>Ap.S1099*	0	0	0	0	0	0	
32338781c.4426G>Cp.D1476H	0	0	0	1	0	0	
32339096c.4741G>Tp.E1581*	0	0	0	0	0	0	
32339132c.4777G>Tp.E1593*	0	0	0	0	0	0	
32339327c.4972C>Tp.Q1658*	0	0	0	0	0	0	
32340210c.5855T>Ap.L1952*	1	0	0	0	0	0	
32340258c.5903C>Gp.S1968*	0	0	0	0	0	0	
32340510c.6155C>Ap.S2052*	0	0	0	0	1	0	
32354943c.7090G>Tp.E2364*	0	1	0	0	0	0	
32355046c.7193C>Ap.T2398N	1	0	0	0	0	0	
32371078c.8610G>Tp.Q2870H	1	0	0	0	0	0	
32379450c.8888C>Tp.S2963L	0	0	0	0	0	0	
32379495c.8933C>Ap.S2978*	0	0	0	0	0	0	
32394666c.9257-23C>Gp.?	0	0	0	0	0	0	
32394763c.9331G>Tp.E3111*	0	0	1	0	0	0	
32397045c.9648+1G>Ap.?	1	0	0	0	0	0	
32398489c.9976A>Tp.K3326*	0	0	0	0	0	0	

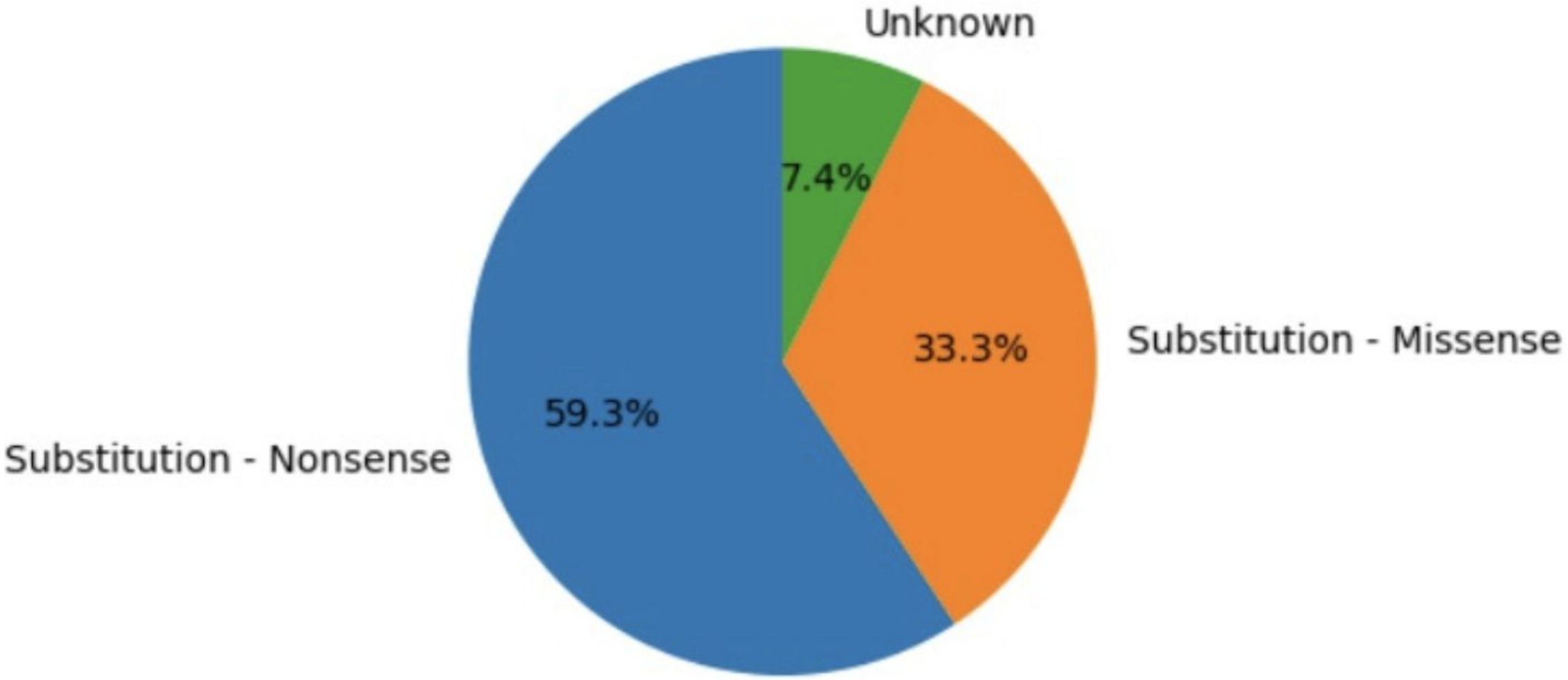
Mutation vs Frequency – Histogram



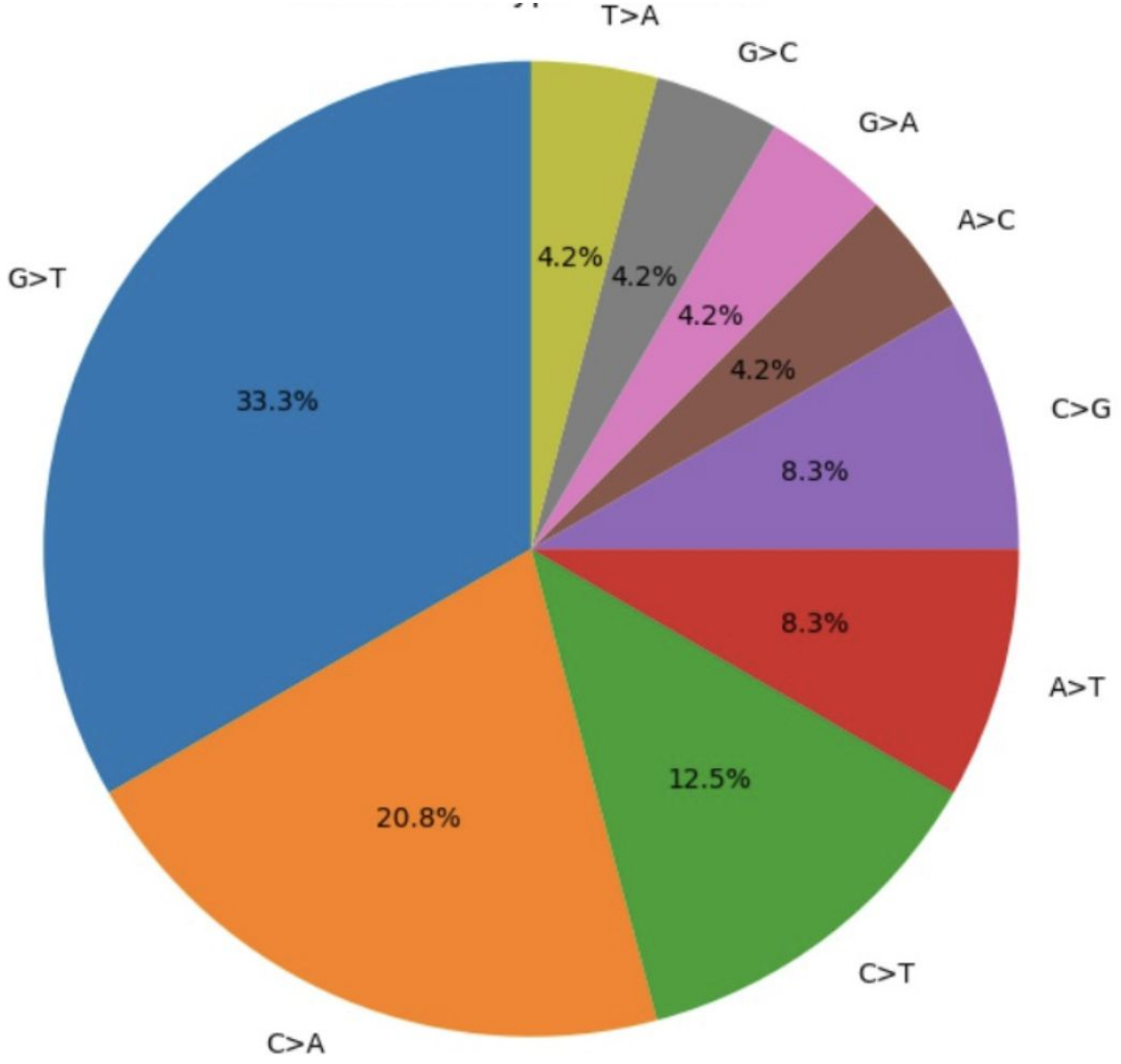
Top 5 Most Frequent Mutations



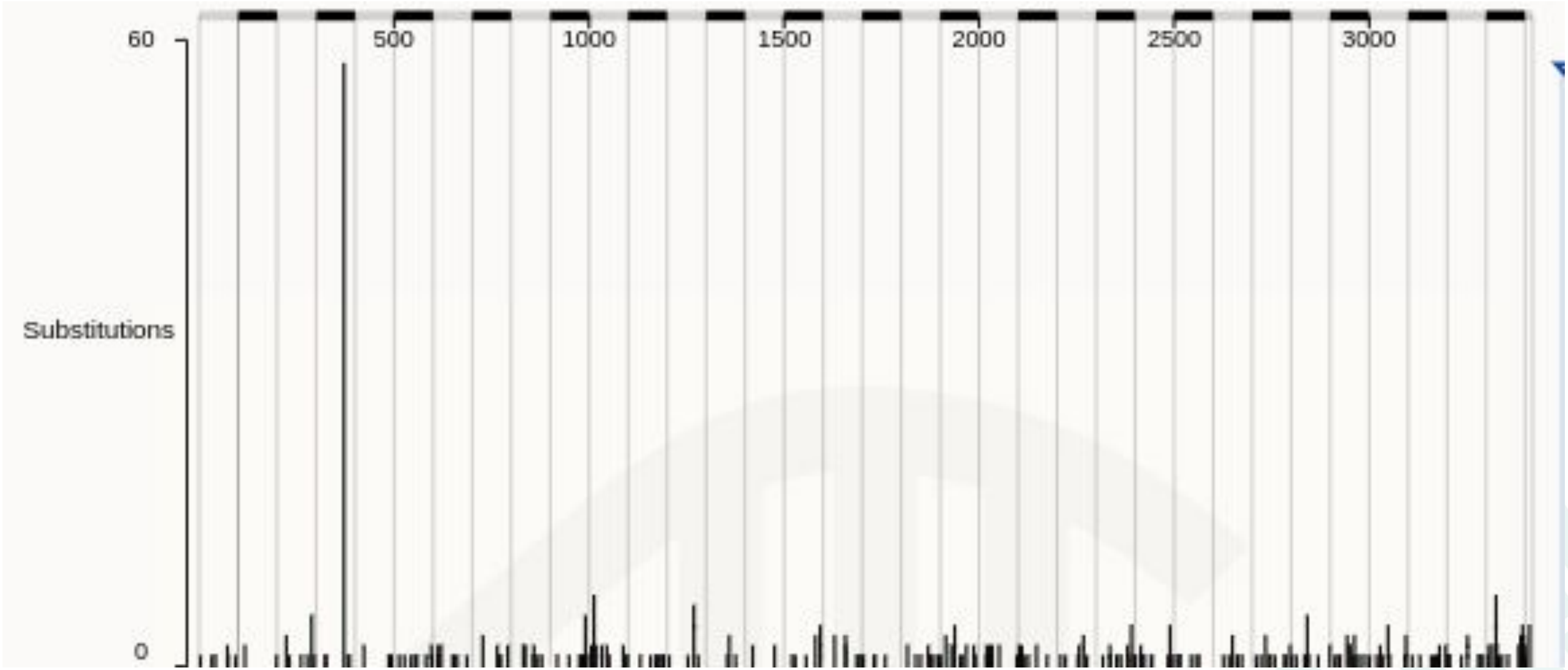
Mutation Type Distribution



Substitution Type Distribution



Mutation Distribution (BRCA2)



Diving Deeper into BRCA2

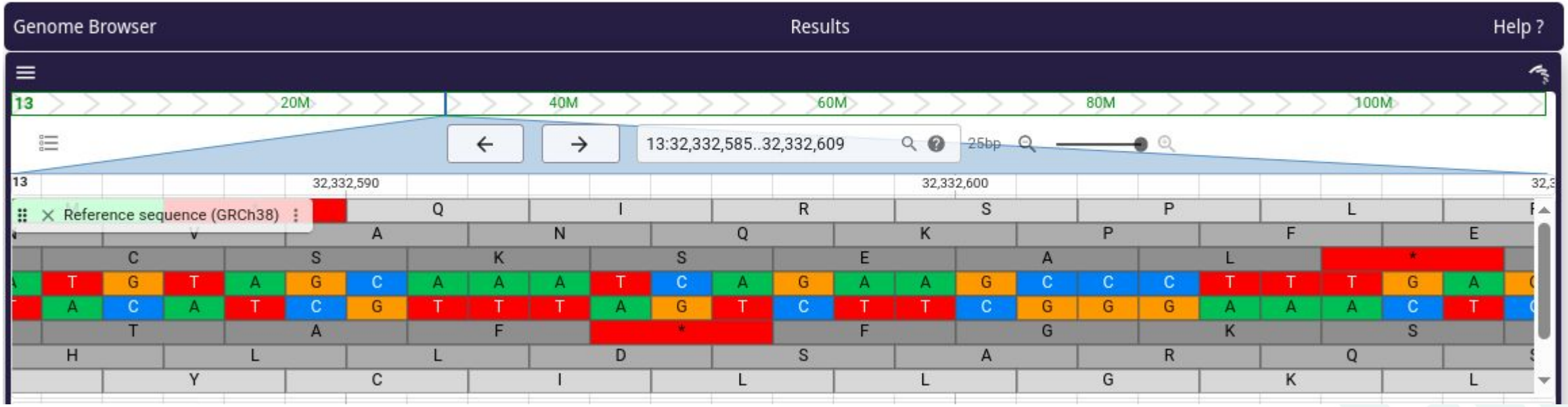
BReast CAncer gene 2

- Tumour Suppressor Gene. It helps prevent cells from growing and dividing too rapidly
- It performs DNA damage repair, helping maintain genomic stability - preventing cells from turning cancerous.
- We found that if we plot mutation frequency vs position, we **do not** get a normal distribution, rather mutations in some specific positions are much much more frequent than others.
- We suspect these are loss / gain of function mutations

Diving Deeper into BRCA2

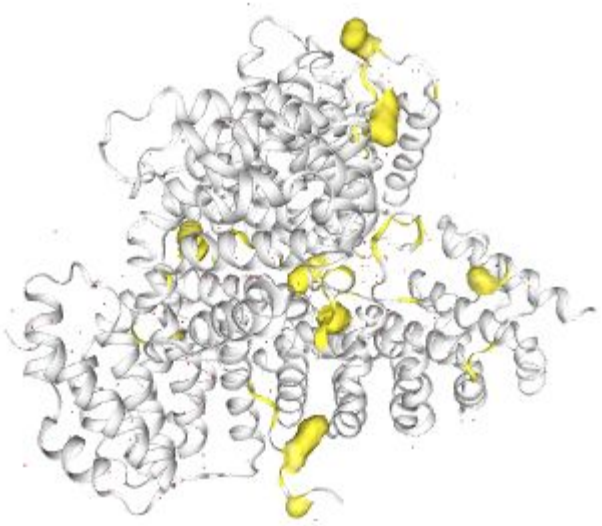
For BRCA2, the most frequent mutation is a missense mutation, two nucleotides (TT) are deleted at positions 7679 and 7680 in the coding DNA sequence, resulting in a frameshift mutation.

This frameshift leads to a change in the amino acid sequence at position 2560 and results in a premature **stop codon!** downstream

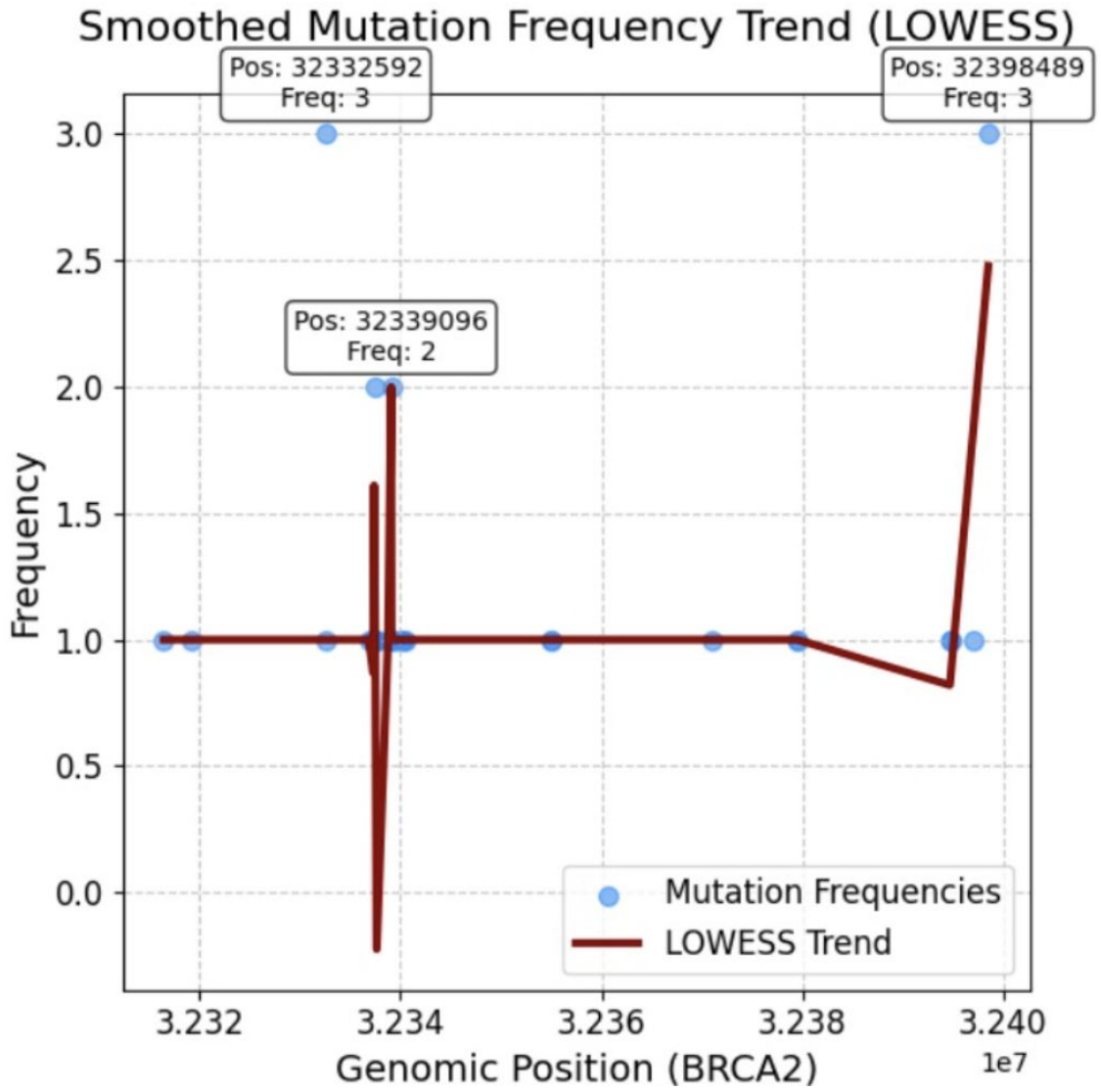


The Protein Encoded

There are 13 recognized proteins from BRCA2 available on UniProt. Studying : 7BDX (first multimeric 3D Structure)



Smoothed Mutation Frequency Trend – LOWESS



Smoothed Mutation Frequency Trend – LOWESS

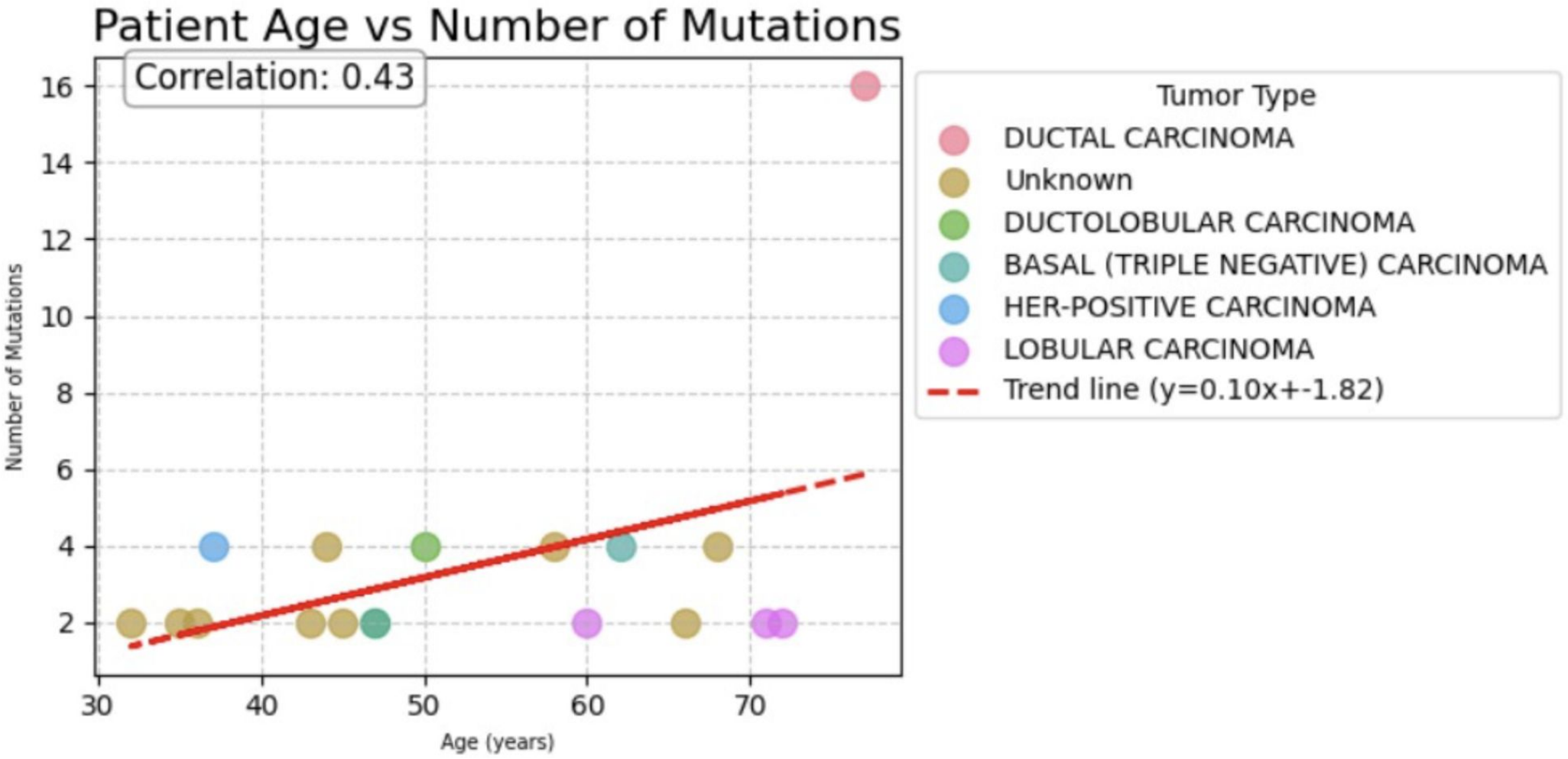
Key Observations:

Mutation frequencies are generally low, ranging from 1 to 3.
Slight peaks are observed at positions 32323592 and 32339849 (both with frequency = 3).
LOWESS trend is largely flat, indicating no strong directional change in mutation frequency across these genomic positions.

Conclusion:

Frequencies are low and relatively consistent.

Patient Age vs Number of Mutations



Patient Age vs Number of Mutations

Key Observations:

Moderate positive correlation (r = 0.43) between age and mutation count.

Regression line shows older patients have slightly more mutations.

Wide spread suggests other influencing factors.

Conclusion:

Age is linked to mutation load, supporting somatic accumulation with age.

Other factors like tumor type and genetics also play a role.

Patient Age vs Mutation By Ethnicity

Patient Age vs Mutation By Ethnicity

Key Observations:

Overall correlation remains r = 0.43.

Red group: flat trend; Blue group: slight increase with age.

Black dashed line shows general upward trend.

Conclusion:

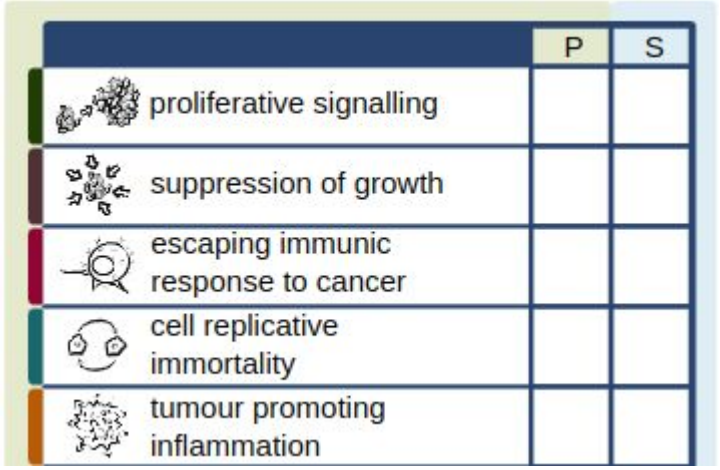
Age-related mutation accumulation differs by ethnicity.

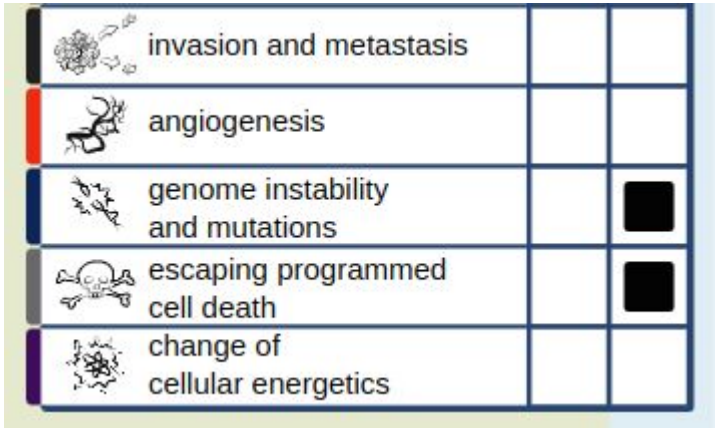
Ethnic-specific genetic or environmental factors may influence mutation patterns.

Hallmarks of Cancer in BRCA2

Role-

- BRCA2 is a tumour suppressing gene, critical for meiotic recombination
- It suppresses escaping apoptosis, genomic instability and mutations. It is critical for stabilization of stalled replication forks

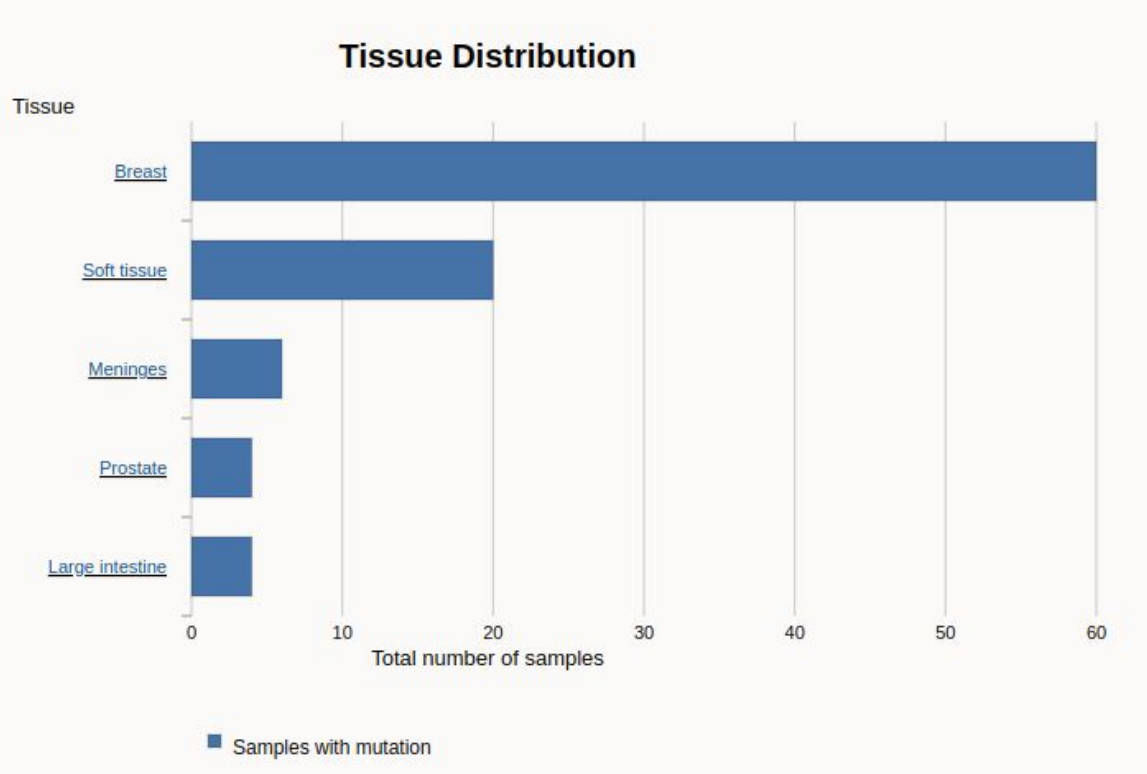




Tissue Distribution for BRCA2

Role-

- Since most breast cancers (80~85%) originate from epithelial cells of estrogen receptor rich milk ducts (ductal carcinoma), ductal carcinoma is the most common type of breast cancer



	Sample Name	Gene name	Transcript	Primary Tissue	Tissue Subtype 1	Histology	Histology Subtype 1	Pubmed ID
2	2943787	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
3	2943790	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
4	2943791	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
5	2943793	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
6	2943795	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
7	2943798	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
8	2943799	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
9	2943803	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
10	2943804	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
11	2943805	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
12	2943811	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
13	2943813	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
14	2943814	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
15	2943817	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
16	2943818	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
17	2943820	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
18	2943821	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
19	2943822	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
20	2943823	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
21	2943824	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
22	2943825	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
23	2943826	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
24	2943828	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
25	2943829	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	lobular_carcinoma
26	2943830	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
27	2943832	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
28	2943833	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
29	2943835	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
30	2943837	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
31	2943839	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
32	2943840	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
33	2943841	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
34	2943844	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
35	2943845	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
36	2943847	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
37	2943848	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
38	2943849	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
39	2943850	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
40	2943855	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
41	2943860	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
42	2943866	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
43	2943868	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
44	2943871	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma
45	2943872	BRCA2	ENST00000380152.7	breast	NS	35875117	carcinoma	ductal_carcinoma

References

1. Catalogue of Somatic Mutations in Cancer **(COSMIC)**
2. https://link.springer.com/article/10.1007/s12609-024-00567-w
3. https://pubmed.ncbi.nlm.nih.gov/14681210/
4. https://pubmed.ncbi.nlm.nih.gov/24598993/
5. https://asia.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG0000000139618;r=13:32315086-32400268
6. https://cancer.sanger.ac.uk/cosmic/mutation/overview?id=122096550
7. https://colab.research.google.com/drive/1w2IO2YbfzSYzj1d0Fbag8gl6ZSQ8cfR_?authuser=1#scrollTo=0TUXAtjn9oO9

Team Members

Aditi Aryan - 2023037

Akshat Kumar - 2023060

Asher Ul Haque - 2023151

Bhuvika Mehta - 2023172

Simar Ahi - 2023527

Sree Sravya Uppalapati - 2023534

Varsha Ganesh - 2023583

Contributions

Python code & Final Presentation Slides
(https://colab.research.google.com/drive/1w2IO2YbfzSYzj1d0Fbag8gl6ZSQ8cfR_?usp=sharing)

Aditi Aryan
Akshat Kumar
Asher Ul Haque
Bhuvika Mehta
Simar Ahi
Sree Sravya Uppalapati
Varsha Ganesh

Everyone contributed equally and were present in all meetings held.