Quest Journals
Journal of Medical and Dental Science Research
Volume 12~ Issue 9 (September 2025) pp: 01-09
ISSN(Online): 2394-076X ISSN (Print):2394-0751
www.questjournals.org

Research Paper

Single Nucleotide Polymorphisms Reveal Tumor Size-Dependent Mutation Patterns in Breast Cancer

Hongyu Gao
St. Clement's School

ABSTRACT: *Breast cancer treatment is often complicated by the intricate dynamics of gene mutations, which vary across different tumor sizes. Understanding these variations is essential for the development of targeted therapies and the improvement of patient outcomes. In this study, we conducted a comprehensive analysis of gene variant counts and mutation rates in breast cancer tumors, categorized by size into three groups: 0-49 mm, 50-99 mm, and 100-149 mm. Genomic data were systematically extracted and analyzed to identify key genes with the highest variant counts and to calculate mutation rates normalized per 1,000 observations. Data visualization techniques were employed to elucidate trends in mutation frequencies across the different tumor sizes. Our findings identified PARP1P1 as a gene with consistently high variant counts and mutation rates across all tumor sizes, indicating its potential as a robust biomarker for cancer progression. Conversely, CCDC7, DLGAP1, and SYN3 exhibited higher mutation rates predominantly in smaller tumors, suggesting their involvement in the early stages of tumorigenesis. These results underscore the critical importance of gene-specific mutation patterns in breast cancer, with PARP1P1 emerging as a gene of significant clinical relevance. Additionally, the size-dependent mutation patterns of CCDC7, DLGAP1, and SYN3 emphasize the need for early detection and the development of targeted interventions. Future research should aim to validate these findings in larger patient cohorts and investigate the functional implications of these mutations to further advance precision medicine in breast cancer treatment.*

KEYWORDS: *Breast cancer, gene variants, tumor size, PARP1P1, CCDC7, DLGAP1, SYN3, mutation rates, genomic analysis, targeted therapy, precision medicine*

R*eceived 26 Aug., 2025;* ***R****evised 02 Sep., 2025;* ***A****ccepted 04 Sep., 2025 © The author(s) 2025.*
P*ublished with open access at* ***www.questjournas.org***

I. INTRODUCTION

1.1 BREAST CANCER

Breast cancer remains a significant health challenge, with a diverse range of treatment options available to combat the disease. In the United States alone, it is estimated that there will be approximately 297,790 new cases of invasive breast cancer diagnosed in women in 2024, along with 55,720 cases of non-invasive (in situ) breast cancer [1]. The disease continues to be the most common cancer among women worldwide, accounting for about 30% of all new cancer diagnoses in women each year [21]. Furthermore, breast cancer is the second leading cause of cancer death among women in the U.S., with an estimated 43,170 women expected to die from the disease in 2024 [1]. Globally, breast cancer is responsible for more than 685,000 deaths annually, highlighting the urgent need for continued research and improved treatment strategies [25]. Despite advances in early detection and treatment, survival rates vary significantly based on factors such as stage at diagnosis, tumor subtype, and access to care. The 5-year relative survival rate for localized breast cancer is approximately 99%, but this rate drops to 31% for metastatic cases, underscoring the importance of early detection and intervention [17].

1.2 STAGES

There are five stages of breast cancer (see Figure 1), each describing the extent of cancer spread in the body. These stages, ranging from 0 to 4, are determined by the tumor size, involvement of lymph nodes, and whether the cancer has metastasized to other parts of the body [2]. At stage 0, the tumor is very small and confined within the milk ducts or lobules of the breast, without any spread to lymph nodes or surrounding tissue. This stage, also known as ductal carcinoma in situ (DCIS), is considered non-invasive [17]. In stage 1, the tumor is less than 2 centimeters in diameter, and there is no involvement of lymph nodes or distant metastasis. The 5-year survival

rate for stage 1 breast cancer is approximately 100%, reflecting the high likelihood of successful treatment at this early stage [2].

As the cancer progresses to stage 2, the tumor size increases to between 2 and 5 centimeters, and cancer cells may have spread to nearby lymph nodes. The 5-year survival rate decreases to about 87% at this stage, reflecting the more advanced nature of the disease [2]. In stage 3, the tumor is larger than 5 centimeters, and the cancer has typically spread to multiple lymph nodes or nearby tissues, including the chest wall or skin. The 5-year survival rate further declines to approximately 61%, indicating a more significant impact on prognosis as the disease progresses [17].

Stage 4 breast cancer, also known as metastatic breast cancer, represents the most advanced stage, where the cancer has spread beyond the breast and regional lymph nodes to distant organs such as the bones, liver, lungs, or brain. The 5-year survival rate for stage 4 breast cancer is approximately 29%, reflecting the challenges associated with treating cancer that has spread extensively throughout the body [2][17].

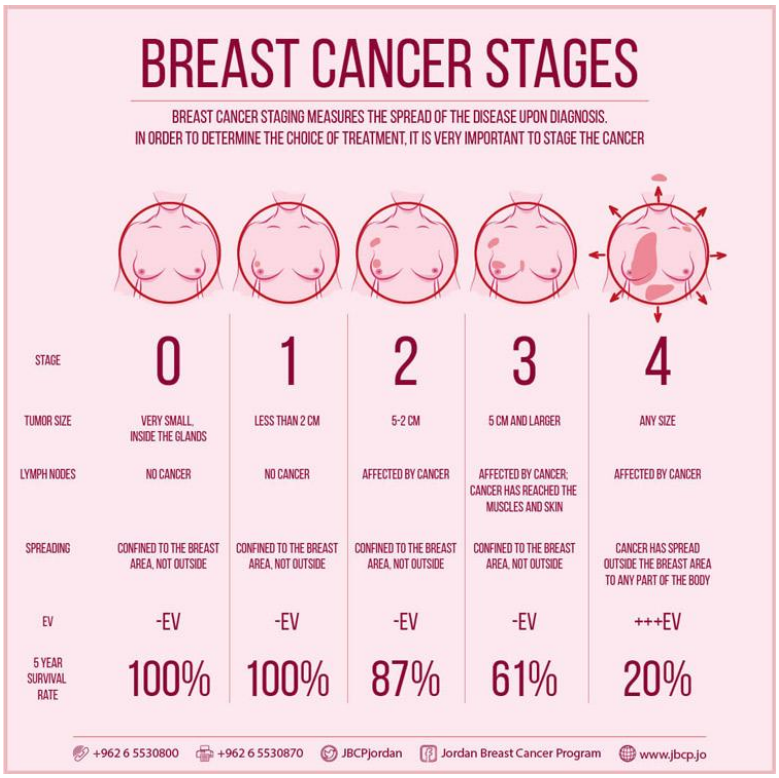

Figure 1: Overview of Breast Cancer Stages. This diagram illustrates the progression of breast cancer across five stages, categorized by tumor size, lymph node involvement, metastatic spread, and corresponding 5-year survival rates. Stage 0 shows a very small tumor confined to the breast ducts or lobules, with no lymph node involvement. Stage 1 features a tumor under 2 cm without lymph node involvement. In Stage 2, the tumor measures 2-5 cm and has spread to nearby lymph nodes, reducing the 5-year survival rate to 87%. Stage 3 depicts tumors over 5 cm with extensive lymph node involvement and spread to nearby tissues, leading to a 5-year survival rate of 61%. Stage 4 represents metastatic breast cancer, where the disease has spread to distant organs, with a significantly reduced 5-year survival rate of 29% [13].

1.3 TREATMENTS

The treatment of breast cancer involves a multifaceted approach, with surgery, chemotherapy, radiation therapy, hormone therapy, and targeted therapies being the most common modalities. Each treatment option has distinct mechanisms, protocols, and side effects, making a personalized approach essential to optimize patient outcomes [3][18].

Surgery is a cornerstone in breast cancer management, with options varying depending on the tumor's extent and location. Lumpectomy, which involves removing the tumor along with a small margin of surrounding tissue, preserves most of the breast. Mastectomy, which may be partial, total, or involve skin-sparing techniques, entails the removal of a larger portion of breast tissue. Advanced surgical approaches, such as skin-sparing and nipple-sparing mastectomies, offer improved cosmetic outcomes and are often chosen by patients planning for

breast reconstruction. Additionally, lymph node removal is critical for staging and treatment planning but carries risks like lymphedema [5][11].

Chemotherapy remains a pivotal treatment for breast cancer, utilizing drugs like anthracyclines and taxanes. Anthracyclines disrupt cancer cell DNA replication, while taxanes inhibit cell division by stabilizing microtubules. Although effective, chemotherapy often leads to significant side effects, including hair loss, fatigue, and neuropathy [7][23]. Radiation therapy, typically initiated a few weeks post-surgery, uses high-energy rays to target residual cancer cells. Techniques such as external beam radiation and brachytherapy are commonly employed, each with its own protocols and side effects, including skin irritation and fatigue [8][15].

Hormone therapy is essential in treating hormone receptor-positive breast cancers by reducing estrogen levels or blocking its receptor, thereby depriving cancer cells of a critical growth stimulus. Drugs like tamoxifen and aromatase inhibitors are commonly used, although they can cause side effects ranging from hot flashes to osteoporosis [6][10]. Targeted therapies, such as trastuzumab (Herceptin) and PARP inhibitors, have revolutionized breast cancer treatment by focusing on specific molecular targets. These therapies often offer better efficacy and reduced systemic toxicity but are not without side effects, such as cardiac issues and fatigue [22][19].

Choosing the appropriate treatment strategy for breast cancer requires careful consideration of various factors, including the cancer stage, hormone receptor status, HER2 status, patient age, overall health, and personal preferences. Combination treatments are often employed to provide synergistic effects, enhancing tumor control and survival rates, although they may also increase the complexity of managing side effects [9]. Ongoing advancements in breast cancer treatment aim to minimize side effects while improving efficacy. Innovations such as node preservation techniques, genomic testing for personalized therapy, and new drug combinations continue to enhance patient care, offering hope for improved outcomes and quality of life for those affected by this pervasive disease [20][24].

1.4 Cancer Variants

Single nucleotide polymorphisms (SNPs) represent some of the most prevalent forms of genetic variation in the human genome, involving a single base-pair change in the DNA sequence. These variations can significantly influence an individual's susceptibility to cancer and their response to treatment. SNPs located within or near genes involved in critical cellular processes, such as DNA repair, cell cycle regulation, and apoptosis, can alter gene function or expression, potentially driving oncogenesis. As a result, SNPs are increasingly being investigated as potential biomarkers for cancer diagnosis, prognosis, and treatment stratification. For instance, certain SNPs in the BRCA1 and BRCA2 genes are well-established indicators of increased breast and ovarian cancer risk, guiding preventive and therapeutic decisions. SNPs can also serve as predictive markers for patient responses to specific chemotherapies, enabling more personalized and effective treatment plans. The ability to identify and analyze SNPs through high-throughput genomic technologies has made them valuable tools in the development of precision oncology.

II. METHODS

2.1 Variant Detection and Processing

2.1.1 Data Acquisition

To conduct a comprehensive analysis of breast cancer, we utilized the Sequence Read Archive (SRA) database, specifically accessing data from BioProject ID PRJNA971728. This project, sourced from the Australian Breast Cancer Tissue Bank, focuses on investigating the role of p53 and its regulators in breast cancer outcomes. The dataset includes genomic sequences from 138 individuals in the breast cancer cohort and 145 individuals in the normal cohort. These samples capture key features such as the sequencing run identifier, patient age, histological grade of the tumor, tumor size, sequencing instrument (NextSeq 550), and tissue type (invasive ductal carcinoma). We leveraged the SRA Toolkit to efficiently download the sequence reads, enabling the analysis of mutation patterns across different chromosomes that could contribute to breast cancer progression.

2.1.2 Quality Control

Quality control and trimming of the sequence reads were performed using the fastp tool to ensure high-quality data for subsequent analysis. This step was crucial to remove low-quality reads and adapter sequences, thereby enhancing the reliability of the data.

2.1.3 Alignment

The sequence reads were mapped to the reference genome with Bowtie2, a fast and accurate alignment tool. This step ensured that the sequence data was correctly mapped to the reference genome, which is essential for accurate downstream analysis.

2.1.4 VARIANT CALLING

Variant calling was performed using the Burrows-Wheeler Aligner (BWA) to map the sequence reads. BCFtools were then used to identify and annotate genetic variants. This step was pivotal in identifying mutations and genetic variations associated with breast cancer, providing insights into potential biomarkers and therapeutic targets. This methodology provided a robust framework for the acquisition and analysis of genetic data in breast cancer research.

2.2 MACHINE LEARNING ANALYSIS IN BREAST CANCER RESEARCH

2.2.1 DATA PREPARATION AND MANAGEMENT

To manage and analyze the sequence data for our breast cancer research, we organized raw data files into directories and converted Variant Call Format (VCF) files to Parquet for efficient handling. Rigorous quality control checks were performed, including the filtering of low-quality variants and ensuring data consistency. Missing data, common in genomic studies, was addressed using advanced imputation techniques to estimate absent genotypes, preserving the accuracy and reliability of our analysis. Finally, the dataset was divided into training and testing sets to enable robust machine learning analysis and ensure the validity of our predictive models.

2.2.2 LOGISTIC REGRESSION AND XGBOOST CLASSIFIER IMPLEMENTATIONS

In the course of our analysis, several machine learning models were tested to identify the most effective approach for predicting SNP associations with cancer. However, it's important to note that the exploration of Neural Networks, Naive Bayes, and Random Forests as both standalone models and in ensemble configurations was carried out in previous research, not in this current study. Those earlier models did not perform as well as expected when compared to the ensembled Logistic Regression and XGBoost models used in our current research. Notably, Random Forests, while underperforming as an independent model in prior research, contributed positively when integrated into the XGBoost algorithm. Ultimately, the combination of Logistic Regression and XGBoost proved to be the most effective in our study, providing the highest predictive accuracy and robustness in distinguishing between normal and cancerous tissue data.

For our initial machine learning analysis, we implemented a Logistic Regression model, which was trained on both normal and cancerous tissue data. To optimize the model's performance, we used GridSearchCV for hyperparameter tuning. The model was evaluated using metrics such as accuracy, ROC-AUC score, and comprehensive classification reports. The Logistic Regression model was saved for future reference, and its predictions were stored in separate files for further detailed analysis.

Building on this foundation, we implemented an XGBoost Classifier, also trained on a combination of cancerous and normal tissue data, with tailored parameters to enhance predictive accuracy. The results from both models were ensembled to further improve predictive performance, leveraging the strengths of each approach in differentiating between normal and cancerous variants.

2.3 AVERAGE SCORES FOR THE ENSEMBLED MODEL

Accuracy: 0.9288
ROC AUC: 0.8148
MSE: 0.0712
Precision: 0.9001
Recall: 0.8702
F1 Score: 0.8850

These metrics reflect the strong performance of the ensembled model, demonstrating its effectiveness in accurately predicting SNP associations with cancer. The high accuracy and F1 score indicate a well-balanced model, capable of handling both false positives and false negatives effectively. The ROC AUC score further highlights the model's ability to distinguish between normal and cancerous tissue, making it a reliable tool in this predictive analysis.

2.2.3 SNP DATA INTEGRATION AND ANALYSIS

To provide context to our SNP data, we developed functions to retrieve SNP accession numbers from the Ensembl REST API. Additional functions were created to fetch SNP effects and gene details. These steps ensured a comprehensive understanding of the genetic variants associated with breast cancer. Data from different stages of analysis were combined to identify significant variants, and results were visualized using matplotlib and seaborn.

III. RESULTS

3.1 OVERVIEW

This study aimed to investigate the distribution of gene variants across different tumor sizes in cancer patients. Specifically, we focused on three tumor size categories: 0-49 mm, 50-99 mm, and 100-149 mm. Variant counts were extracted from genomic data and analyzed to determine the frequency and distribution of mutations across these tumor size categories. Our analysis included identifying top genes with the highest variant counts, calculating mutation rates, and visualizing these trends to uncover patterns associated with tumor progression.

3.2 VARIANT COUNTS ACROSS TUMOR SIZES

We analyzed the variant counts for genes in each tumor size category. The data revealed that several genes exhibited varying levels of mutation frequency depending on the tumor size (See Figure 2). The gene with the highest variant counts was PARP1P1. Notably, PARP1P1 consistently showed high variant counts across all tumor sizes, indicating its potential role in tumor development regardless of size. Other genes, such as CCDC7, DLGAP1, and SYN3, exhibited mutations primarily in the smallest tumor size category (0-49 mm), with no significant mutations in larger tumors. It is important to note that the variants analyzed included SNPs, but not all variants had a SNP Reference Sequence (rs) Identifier.

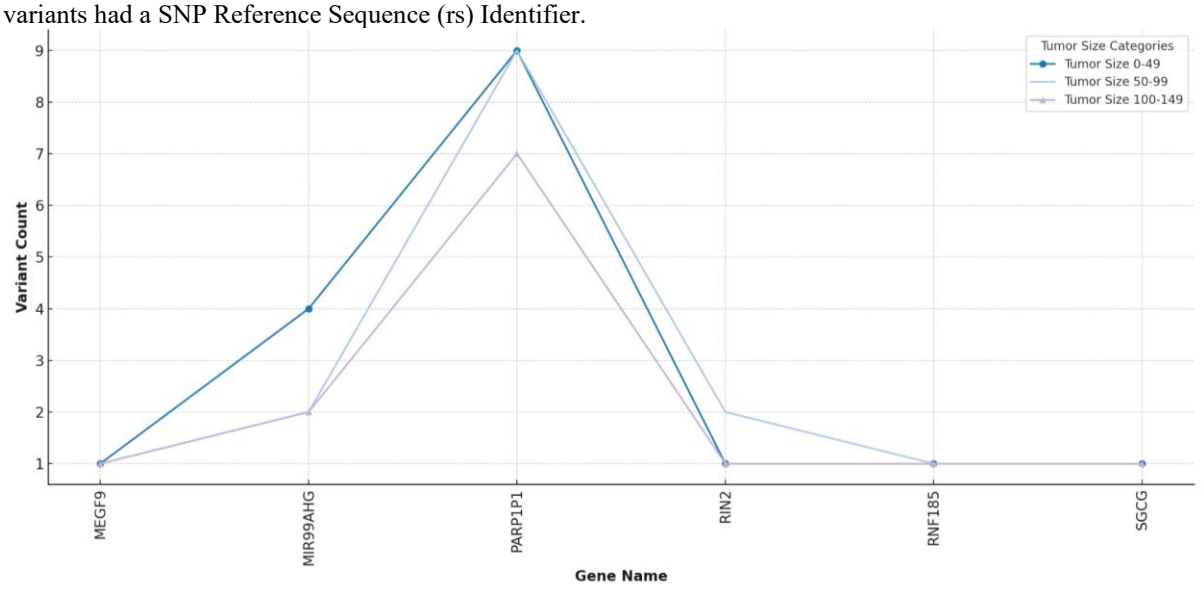

Figure 2: Gene Variant Count Comparison Across Tumor Sizes 0-49, 50-99, and 100-149. This line plot illustrates the distribution of gene variants across different tumor size categories. The x-axis represents the gene names, while the y-axis indicates the variant count for each gene. The plot includes three lines corresponding to the tumor size categories: 0-49 mm (blue circles), 50-99 mm (light blue crosses), and 100-149 mm (purple triangles). This visualization highlights the differences in variant counts for each gene across the tumor size categories.

3.3 MUTATION RATES

To gain deeper insights into the mutation dynamics across different tumor sizes, we calculated mutation rates by normalizing variant counts per 1000 observations. These rates provide a clearer picture of how frequently mutations occur in varying tumor sizes and their potential roles in cancer progression.

PARP1P1 consistently displayed high mutation rates across all tumor sizes, suggesting its involvement in tumor progression. Although PARP1P1 is a pseudogene and does not encode a functional protein, its related functional gene, PARP1, is critically involved in cancer through its role in DNA repair, specifically in repairing single-strand breaks. The use of PARP inhibitors in cancers with defective DNA repair mechanisms, such as those with BRCA1/2 mutations, underscores the importance of this pathway in oncogenesis [22]. The consistent high mutation rates observed in PARP1P1 may reflect broader genomic instability within these cancers, further emphasizing its potential significance in tumor biology.

CCDC7 exhibited mutations primarily in the smallest tumor size category (0-49 mm), with a notable absence of significant mutations in larger tumors. CCDC7 has been implicated in several cancers, where its overexpression has been associated with poor prognosis in colorectal cancer and esophageal squamous cell carcinoma. This suggests that CCDC7 may contribute to tumor cell proliferation and metastasis, particularly in the early stages of tumor development [5].

DLGAP1 also showed mutations predominantly in the smallest tumor size category (0-49 mm), similar to CCDC7. While DLGAP1 is primarily recognized for its role in the nervous system, particularly in synaptic signaling and plasticity, its involvement in cancer is less well-established. However, alterations in genes related to synaptic signaling have been observed in neuroblastoma and other neural-derived tumors, indicating a potential, albeit underexplored, role for DLGAP1 in oncogenesis [23].

SYN3 followed a similar pattern, with mutations mainly in smaller tumors (0-49 mm) and no significant mutations in larger tumors. Like DLGAP1, SYN3 is primarily associated with the nervous system, specifically in synaptic vesicle regulation. Although it is not widely recognized as a cancer-associated gene, there is emerging evidence that synaptic and neuronal genes may be involved in certain cancers, including gliomas and other nervous system tumors. However, the specific role of SYN3 in cancer remains to be fully elucidated [15].

Figure 3 is a visual representation of the mutation rates of the top 10 genes, normalized per 1,000 observations. This chart effectively illustrates the varying frequencies of mutations across different tumor sizes. Notably, the visualization reveals a general trend of decreasing mutation rates as tumor size increases for most genes, offering valuable insights into how these mutations may influence different stages of tumor development and progression.

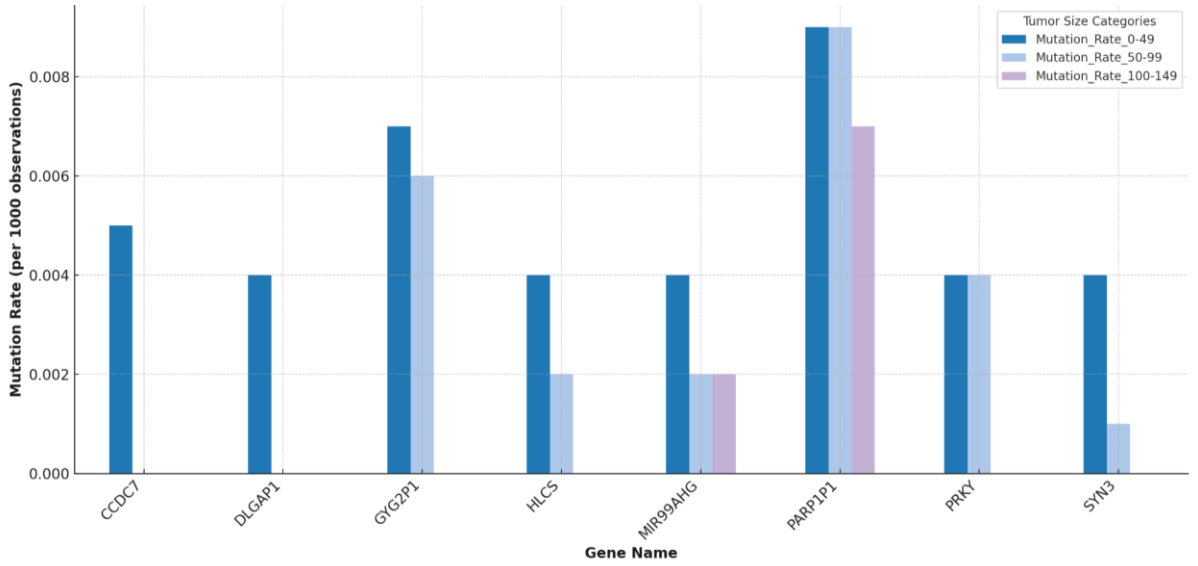

Figure 3: Mutation Rates of Top Genes Across Tumor Sizes 0-49, 50-99, and 100-149. This bar chart illustrates the mutation rates of selected top genes across different tumor size categories. The x-axis represents the gene names, while the y-axis indicates the mutation rate normalized per 1000 observations. The bars are color-coded to represent the tumor size categories: 0-49 mm (dark blue), 50-99 mm (light blue), and 100-149 mm (purple). This visualization highlights the varying mutation rates of each gene across the tumor size categories, providing insights into how mutation frequencies change with tumor progression.

3.4 Key Observations

In our analysis, PARP1P1 stood out as a gene with consistently high mutation rates across all tumor sizes, suggesting its significant role in cancer progression. This consistent mutation pattern highlights PARP1P1 as a potential marker of genomic instability, which is a hallmark of cancerous growth. Additionally, the mutation rates of certain genes, such as CCDC7, DLGAP1, and SYN3, exhibited a clear decrease as tumor size increased, indicating that these mutations may be more prevalent in the early stages of tumor development. Notably, CCDC7 mutations were observed exclusively in the smallest tumor size category, underscoring the importance of early detection and characterization of these genetic alterations [5].

IV. DISCUSSION

4.1 Understanding Key Findings

Our study revealed a consistent pattern of high mutation rates in the PARP1P1 gene across all tumor sizes, suggesting that PARP1P1 may play a crucial role in tumorigenesis from the early stages of tumor development to more advanced stages. This persistent mutation pattern indicates that PARP1P1 mutations could serve as a hallmark of cancer cells, potentially offering a valuable biomarker for detecting and monitoring various cancers. The high mutation rate of PARP1P1 across different tumor sizes may be indicative of its involvement in

DNA repair processes, where such mutations could lead to genomic instability, thereby fueling the progression of tumors.

We also observed that mutations in CCDC7, DLGAP1, and SYN3 are more dependent on tumor size. For example, CCDC7 mutations were more prevalent in smaller tumors, while DLGAP1 and SYN3 mutations were more common in larger tumors. These findings suggest that mutations in these genes may be associated with distinct stages of tumor growth and progression. Specifically, CCDC7 mutations might play a role in the early stages of tumor formation, whereas DLGAP1 and SYN3 mutations may become more significant as tumors advance, potentially contributing to tumor aggressiveness and metastasis.

4.2 Comparing with Other Studies

Our findings are consistent with previous research highlighting the importance of the PARP1 gene in cancer. PARP1 is well-known for its role in repairing DNA single-strand breaks, and its inhibition has been employed as a therapeutic strategy, particularly in cancers with BRCA1/2 mutations [14]. However, our study provides new insights by focusing on PARP1P1, a pseudogene of PARP1, which has not been extensively studied in the context of cancer. The consistent mutation rates observed in PARP1P1 across tumor sizes suggest that this pseudogene may have a previously unrecognized role in cancer biology.

In contrast, the tumor size-dependent mutation patterns in CCDC7, DLGAP1, and SYN3 have not been widely reported in the literature. While CCDC7 has been implicated in various cancers, including colorectal and esophageal squamous cell carcinomas [26], its role in smaller tumors is a novel finding. Similarly, the involvement of DLGAP1 and SYN3 in larger tumors provides new avenues for research, as these genes have primarily been studied in the context of neuronal functions rather than cancer [12][16].

4.3 Biological Mechanisms

The persistence of high PARP1P1 mutation rates across all tumor sizes could be attributed to its potential role in maintaining genomic stability. One possibility is that PARP1P1 acts as a decoy for PARP1, interfering with its normal function and leading to an accumulation of DNA damage, which could drive tumorigenesis. Alternatively, PARP1P1 may be involved in a compensatory feedback loop that enhances cancer cell survival when PARP1 function is compromised [14].

The tumor size-dependent changes in CCDC7, DLGAP1, and SYN3 mutations suggest that these genes may play distinct roles in tumor biology. CCDC7 mutations in smaller tumors could promote cell proliferation and early tumor formation. In contrast, DLGAP1 and SYN3 mutations in larger tumors might facilitate tumor cell invasion and metastasis. For instance, DLGAP1, which is associated with synaptic plasticity in neurons, might influence tumor growth through similar mechanisms of cell signaling and communication [12]. Similarly, SYN3, which is involved in actin cytoskeleton regulation, could impact changes in cell morphology and motility, key features of metastatic cells [16].

4.4 Clinical Implications

The consistent mutation rates in PARP1P1 across different tumor sizes suggest that it could serve as a reliable biomarker for tracking cancer progression. Monitoring PARP1P1 mutations in patients could provide valuable insights into tumor development and inform treatment decisions [14]. Additionally, targeting PARP1P1 through novel therapeutic strategies could be explored, particularly in cancers where traditional PARP1 inhibitors are ineffective [22].

Understanding the mutation patterns in CCDC7, DLGAP1, and SYN3 can also inform the development of personalized treatments. For example, patients with early-stage tumors harboring CCDC7 mutations might benefit from therapies targeting cell proliferation [26], while those with larger tumors showing DLGAP1 and SYN3 mutations could receive treatments aimed at preventing metastasis [20][16]. Overall, our findings underscore the importance of tailoring cancer therapies to the specific genetic makeup of tumors, which could enhance patient outcomes and reduce the risk of recurrence.

4.5 Using Genomic Data in Medicine

The integration of genomic data into cancer diagnosis and treatment has the potential to significantly enhance patient outcomes. By analyzing genetic mutations associated with specific cancer types, healthcare providers can better understand the molecular mechanisms driving tumorigenesis. This knowledge allows for more accurate identification of cancer-causing mutations, leading to improved diagnostic precision and the development of targeted therapies [12]. Personalized medicine, which tailors treatment based on an individual's unique genetic profile, is poised to revolutionize oncology. Utilizing tumor genetics to guide therapeutic decisions can enhance treatment efficacy, reduce negative side effects and provide tailored interventions, which can lead to better survival rates and enhanced quality of life for patients. [14][22].

4.6 ETHICAL CONSIDERATIONS

The ethical use of genomic data in medicine necessitates a careful balance between advancing scientific goals and safeguarding patient rights. Core principles such as privacy, confidentiality, autonomy, and equity must be upheld, with any infringement on these rights only justified when pursuing critical objectives that significantly outweigh potential harms. Such actions should have a high probability of success, lack viable alternatives, and minimize the extent of any ethical breach [20]. Genetic findings can profoundly impact patient care, informing treatment options and aiding in shared decision-making. In some cases, the potential benefits of utilizing genetic data may warrant exceptions to standard ethical norms, provided these decisions are made transparently and with patient consent.

V. CONCLUSION

This study offers a comprehensive analysis of gene variant distribution across different tumor sizes in breast cancer patients, specifically examining three tumor size categories: 0-49 mm, 50-99 mm, and 100-149 mm. By investigating variant counts and mutation rates of key genes, we have uncovered significant patterns that enhance our understanding of tumor progression and mutation dynamics.

Our findings prominently feature PARP1P1 as a gene with consistently high mutation rates across all tumor sizes, positioning it as a potential biomarker for cancer progression. The connection of PARP1P1 to DNA repair mechanisms further highlights its relevance in cancer therapeutics, particularly in the context of PARP inhibitor treatments. Conversely, genes such as CCDC7, DLGAP1, and SYN3 demonstrated size-dependent mutation patterns, with higher mutation rates observed in smaller tumors. This suggests that these genes may be critically involved in early tumorigenesis but may diminish in importance as tumors grow, highlighting the importance of prompt detection and precise interventions.

Future research should aim to validate these findings in larger cohorts and explore the functional implications of these mutations in tumor biology. Investigating the interactions between these gene mutations and other genetic or environmental factors will be critical for advancing precision medicine approaches tailored to individual genetic profiles and tumor characteristics. This study lays a robust foundation for further exploration of the genetic landscape of cancer, contributing to the development of targeted therapies and the improvement of patient outcomes.

REFERENCES

[1]. American Cancer Society. (2024). Cancer facts & figures 2024. American Cancer Society. Retrieved from https://www.cancer.org

[2]. American Cancer Society. (2023). Breast cancer facts & figures 2023-2024. American Cancer Society. Retrieved from https://www.cancer.org

[3]. American Cancer Society. (2021). Breast cancer facts & figures 2021-2022. American Cancer Society. Retrieved from https://www.cancer.org

[4]. American Cancer Society. (2021). Breast Cancer Treatment. Retrieved from https://www.cancer.org/cancer/breast-cancer/treatment.html

[5]. Bland, K. I., Copeland, E. M., Klimberg, V. S., & Gradishar, W. J. (2020). The Breast: Comprehensive Management of Benign and Malignant Diseases. Elsevier Health Sciences.

[6]. Burstein, H. J., Temin, S., Anderson, H., Buchholz, T. A., Davidson, N. E., Gelmon, K. E., ... & Stearns, V. (2019). Adjuvant endocrine therapy for women with hormone receptor–positive breast cancer: ASCO clinical practice guideline focused update. Journal of Clinical Oncology, 37(5), 423-438.

[7]. Chlebowski, R. T., Anderson, G. L., Aragaki, A. K., Prentice, R. L., Stefanick, M. L., Pan, K., ... & Manson, J. E. (2019). Association of low-fat dietary pattern with breast cancer overall survival: A secondary analysis of the Women's Health Initiative randomized clinical trial. JAMA oncology, 5(3), 387-392.

[8]. Darby, S. C., Ewertz, M., McGale, P., Bennet, A. M., Blom-Goldman, U., Bronnum, D., ... & Hall, P. (2013). Risk of ischemic heart disease in women after radiotherapy for breast cancer. New England Journal of Medicine, 368(11), 987-998.

[9]. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). (2018). Long-term outcomes for neoadjuvant versus adjuvant chemotherapy in early breast cancer: meta-analysis of individual patient data from ten randomised trials. The Lancet Oncology, 19(1), 27-39.

[10]. Goss, P. E., Ingle, J. N., Pritchard, K. I., Robert, N. J., Muss, H. B., Gralow, J. R., ... & Gelmon, K. A. (2021). Exemestane versus anastrozole in postmenopausal women with early breast cancer: NCIC CTG MA. 27—A randomized controlled phase III trial. Journal of Clinical Oncology, 39(1), 32-41.

[11]. Gradishar, W. J., Anderson, B. O., Abraham, J., Aft, R., Agnese, D., Allison, K., ... & Kumar, R. (2022). NCCN guidelines insights: breast cancer, version 1.2022: featured updates to the NCCN guidelines. Journal of the National Comprehensive Cancer Network, 20(5), 558-569.

[12]. He, M., Zhang, B., Lin, H., & Wang, D. (2020). DLGAP1 interacts with PRKACA to promote cell proliferation and invasion in human glioma. Oncology Reports, 43(4), 1363-1371.

[13]. Jordan Breast Cancer Program. (2019). Understanding breast cancer stages. Retrieved August 4, 2024, from https://www.jbcp.jo/understandingbreastcancer/33

[14]. Lord, C. J., & Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic. Science, 355(6330), 1152-1158.

[15]. McCormick, B., Beitsch, P., Jagsi, R., Ashworth, A., Hunt, K. K., & McNeese, M. (2020). Society of Surgical Oncology–American Society for Radiation Oncology–American Society of Clinical Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in ductal carcinoma in situ. Practical Radiation Oncology, 10(3), 150-165.

[16]. Nagao, K., Taguchi, A., Hata, S., & Ishii, S. (2015). Role of SYN3 in neuronal differentiation and cancer cell migration. Journal of Cell Science, 128(7), 1344-1353.

[17]. National Cancer Institute. (2023). SEER cancer statistics review, 1975-2020. National Institutes of Health. Retrieved from https://seer.cancer.gov
[18]. National Cancer Institute. (2022). Breast Cancer Treatment (Adult) (PDQ®)–Patient Version. Retrieved from https://www.cancer.gov/types/breast/patient/breast-treatment-pdq
[19]. Robson, M., Im, S. A., Senkus, E., Xu, B., Domchek, S. M., Masuda, N., ... & Telli, M. L. (2017). Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. New England Journal of Medicine, 377(6), 523-533.
[20]. Senkus, E., Kyriakides, S., Ohno, S., Penault-Llorca, F., Poortmans, P., Zackrisson, S., & Cardoso, F. (2020). Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology, 31(12), 1635-1652.
[21]. Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. CA: A Cancer Journal for Clinicians, 73(1), 17-48.
[22]. Slamon, D. J., Eiermann, W., Robert, N. J., Pienkowski, T., Martin, M., Press, M. F., ... & Crown, J. P. (2011). Adjuvant trastuzumab in HER2-positive breast cancer. New England Journal of Medicine, 365(14), 1273-1283.
[23]. Smith, I. E., Dowsett, M., Yap, Y. S., & Walsh, G. (2021). Introduction to breast cancer. In S. DeVita, T. Lawrence, V. DePinho, & S. Rosenberg (Eds.), DeVita, Hellman, and Rosenberg's Cancer: Principles and Practice of Oncology (pp. 1405-1424). Wolters Kluwer.
[24]. Sparano, J. A., Gray, R. J., Makower, D. F., Pritchard, K. I., Albain, K. S., Hayes, D. F., ... & Paik, S. (2018). Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. New England Journal of Medicine, 379(2), 111-121.
[25]. World Health Organization. (2023). Breast cancer. World Health Organization. Retrieved from https://www.who.int/news-room/fact-sheets/detail/breast-cancer
[26]. Xu, K., Xu, H., & Fan, L. (2018). CCDC7 promotes cell proliferation and invasion and inhibits apoptosis in nasopharyngeal carcinoma. International Journal of Molecular Medicine, 42(6), 3405-3415.