Medical Research Archives

European Society of Medicine

OPEN ACCESS

Published: November 30, 2023

Citation: MBAYE, F., et al., 2023. Implication of c-myc mutations in the tumorogenesis of breast cancer in senegalese females. Medical Research Archives, [online] 11(11). https://doi.org/10.18103/mra.v11i11.4800

DOI: https://doi.org/10.18103/mra.v11i11.4800

ISSN: 2375-1924

RESEARCH ARTICLE

Implication of C-MYC Mutations in the Tumorogenesis of Breast Cancer in Senegalese Females

Fatimata MBAYE[1]*, Nada AGENDEY[1] & Mbacké SEMBENE[1]

[1]Genomics Laboratory, Department of Animal Biology, Faculty of Science and Technology, Cheikh Anta Diop University of Dakar, Dakar, Senegal, BP: 5005.

*fatimata2.mbaye@ucad.edu.sn

ABSTRACT

Breast tumors are a frequent cause of medical consultations. Although these tumors are mainly benign, they can become malignant or cancerous. This study aimed to elucidate the involvement of genetic alterations in the *C-MYC* oncogene in breast tumorigenesis in Senegalese females. After PCR, the epidemiological and molecular profiles of 45 samples, including 19 controls (C) from healthy individuals and 11 benign (BT) and 15 cancerous (MT) samples from patients with tumors, were determined. Mutations were determined using Mutation Surveyor software, and their pathogenicity was assessed using SIFT, Polyphen-2, Mutpred2, SNAP2, PANTER-PSEP, PROVEAN, PhD-SNP, SNP&GO, MUpro, and I-mutant prediction tools. At the epidemiological level, the average ages of the BT and MT groups were 21 and 49.76 years, respectively, and the average ages at menarche were 14.14 and 14.58 years, respectively, with a high frequency of adenofibromas (53.85%) in the BT group and only infiltrating ductal carcinomas (100%) found in the MT group. The stages (III and IV) and grade of SBR were specific to the MT group 76.47% and 58.82%, respectively. At the molecular level, four mutations were identified, all of which were heterozygous and novel, and three of which were non-synonymous. One of the mutations (c.115 T > TC; p.Tyr39His) was recurrent (frequency = 26.67%, 4/15 in the MT group; and 10.53%, 2/19 in the C group) and the other two (c.113 T > TC; p.Phe38Ser and c.117 C > CT) were exclusive to the C group, with the same frequency of 5.26% (1/19). No mutations were found in the BT group. The p.Phe38Ser and p.Tyr39His mutations were described as deleterious and can cause cancer according to the prediction tools. Overall, these mutations can be considered as variants of interest and are the subject of PCR screening for breast cancer prevention.

Keywords: Mutations, oncogene, *C-MYC*, tumors, breast.

Introduction

The breast is the site of several tumors in females, most of which are benign, with occurrence rates of 57% to 70%[1]. Breast cancer, or malignant breast tumor, is the most common cancer affecting females worldwide in terms of incidence (24.5%) and mortality (15.5%)[2]. However, the etiology of breast cancer remains poorly understood. Numerous risk factors for the occurrence of breast cancer have been identified, including hormonal, environmental, genetic, and hereditary factors (5–10% of cases) or lifestyle and nutritional factors[3]. Breast cancer often occurs at a relatively advanced age (75% of cases occurring after the age of 50 years)[4]. In most sub-Saharan African countries, such as Senegal, breast cancer ranks second among female cancers, following cervical cancer.[5] In fact, breast cancer accounts for 25.1% of all cancers diagnosed in Senegalese females, with a mortality rate of 52.39%.[2] Due to its high mortality and morbidity, breast cancer is a major public health problem[6]. According to a study conducted at the Senology Unit of the Centre Hospitalier Universitaire Aristide-Le-Dantec in Dakar, breast cancer accounts for 58.2% of breast tumor pathologies in Senegal[7]. Although hormonal factors appear to contribute to the development of breast tumors, their pathophysiology remains poorly understood. Breast tumors are mainly found in young, reproductive, and active females aged 20–50 years[8]. These benign tumors (BTs) are characterized by a generally favorable evolution; however, in some cases, they are responsible for functional disorders and have esthetic impacts[7]. Despite having a better prognosis than malignant tumors, some benign breast tumors can progress to cancer[9]. Recent advances in molecular biology have led to a better understanding of tumorigenesis and its evolution. Tumor cells are characterized by genomic instability, which can be observed at the gene and chromosome levels. Genetic (mutations) and epigenetic alterations affecting tumor suppressor genes and oncogenes are considered the main molecular events that provide a selective growth advantage and enable clonal expansion during tumor transformation[10]. The C-MYC oncogene, which plays a key role in regulating cell growth and division, is deregulated in approximately 75% of human tumors, including prostate, breast, colon, lung, and Burkitt's lymphoma[11]. This oncogene is amplified and overexpressed in approximately 15% of breast cancers and in some cases, is associated with poorer prognosis and more aggressive clinical forms[12]. This study aimed to elucidate the genetic mechanisms governing tumorigenesis. In particular, it aimed to evaluate the involvement of C-MYC oncogene mutations in breast tumors in Senegalese females.

Methods

SAMPLES

This study comprised 73 Senegalese patients with breast tumors, of whom 23 had BTs and 50 had cancerous tumors (malignant tumors, MTs). A total of 28 undiagnosed cases of breast tumors were included in the study as controls (Cs). Patients were recruited from the Joliot-Curie Institute at Aristide-Le-Dantec Hospital. Fresh surgical specimens were obtained from each patient who underwent tumor surgery. For Cs, blood was extracted

and placed in ethylenediaminetetraacetic acid tubes. After collection, the samples were sent to the Animal Biology DeparMTent's genomics laboratory and fixed in 96% alcohol (ethanol). The tumor and blood samples were stored at −20° C. Information on the demographic and clinicopathological characteristics of the patients was obtained from clinical survey forms after obtaining informed consent. The following groups were analyzed in this study: Group 1, "C," control blood samples; Group 2, "BT," samples of benign tumors; and, Group 3, "MT," samples of cancers or malignant tumors.

EPIDEMIOLOGICAL AND CLINICAL CHARACTERISTICS

Epidemiological data obtained from patients with BT and MT were entered into an Excel (2010) sheet. Descriptive analysis was performed in Rstudio (version 2022.02.2 Build 485) to determine the characteristics of the BT and MT groups. The C group was not included in the epidemiological analyses as we sought to determine the clinicopathological differences between benign breast tumors and cancers in Senegalese females. The following parameters were compared: age, date of first menstrual period (DFMP), tumor histology, cancer stage and grade, and marital status. The mean, standard deviation, and minimum and maximum values of the quantitative variables (age and DFMP) were determined. The means between the two groups were compared to determine whether the observed differences were significant (Student's *t*-test or Wilcoxon test), depending on the normality of the data (Shapiro-Wilk). For qualitative variables (tumor histology, marital status, cancer stage, and grade), the numbers and frequencies were determined. As benign breast tumors and cancers do not evolve in the same manner, only the "Marital status" variable could be compared between the groups and was carried out using Fisher's exact test. Statistical significance was set at *p*-value < 0.05.

Genetic Characteristics

DNA EXTRACTION

Total DNA from breast tumor tissues (benign and malignant) was extracted using the Qiagen DNeasy Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The same protocol was used to extract the blood samples. However, as the blood already comprised individualized cells, after the removal of 200 µL of the biological fluid, 20 µL of proteinase K, and 200 µL of AL buffer were added to the blood samples.

POLYMERASE CHAIN REACTION (PCR) AMPLIFICATION AND SEQUENCING OF THE C-MYC GENE

In this study, exon 2 of the C-MYC gene was amplified. For the Cs and BTs, PCR was performed in a 25 µL reaction volume which comprised 2 µL of DNA, 10.375 µL of water, 5 µL of buffer (5X), 0.5 µL of dNTPs, 2.5 µL of One Taq G/C, 2.5 µL of High G/C, 1 µL of $MgCl_2$, 0.125 µL of Taq polymerase, and 0.5 µL each of the following primers: forward primer (C-MYC F 5' TCC CCC TTG CCG TCC CAA 3') and reverse primer (C-MYC R 5' CGT GCA AGT CAC AGA CTT 3'). For the breast cancer samples, PCR was performed in a 25 µL reaction volume, which comprised 4 µL of DNA, 8.375 µL of water, 5 µL of buffer (5X), 0.5 µL of dNTP, 2.5 µL of One Taq G/C, 2.5 µL of High G/C, 1 µL of $MgCl_2$, 0.125 µL of Taq polymerase, and 0.5 µL of each primer. PCR

was performed on an Eppendorf thermal cycler under the following cycling conditions: initial denaturation at 95° C for 2 min; 35 amplification cycles of DNA denaturation at 93° C for 1 min and primer hybridization and complementary strand elongation at 72 °C for 2 min; and elongation at 72° C for 8 min. The resulting PCR products were stored in the thermal cycler at 10° C. To confirm ligation of the primers, electrophoretic migration was carried out on an agarose gel. Briefly, 5 µL of the PCR product was mixed with 2 µL of bromophenol blue (loading blue) and loaded onto a 2% agarose gel in the presence of Safeview. Electrophoresis was conducted at 100 V for 30 min, and the bands were visualized under UV light. Sanger sequencing was performed using an ABI 3730XL DNA analyzer (Thermo Fisher Scientific, Waltham, MA, USA).

MOLECULAR ANALYSIS

Sequences of exon 2 of C-MYC from healthy, benign, and cancerous tissues were thoroughly checked, corrected, and aligned using BioEdit software (version 7.0.5.3)[13] and the Clustal W algorithm[14] to highlight sequence similarities and perform further genetic analyses. Similarly, to identify the exact domain for the analysis, we compared our sequences with the reference sequence NM_002467.6 downloaded from the National Center for Biotechnology Information GenBank database.

IDENTIFICATION OF MUTATIONS

To identify DNA mutations and their positions relative to the *C-MYC* gene, raw sequencing data were analyzed using Mutation Surveyor software (version 5.1, www.softgenetics.com). The mutations were then searched in the online databases Single Nucleotide Polymorphism Database (dbSNP), COSMIC, ClinVar, and Ensembl to distinguish new mutations from those already referenced in the databases or described in the literature.

PREDICTING THE PATHOGENICITY OF MUTATIONS

The impact of non-synonymous mutations on the function of the encoded proteins was determined using SIFT (Sorting Intolerant From Tolerant)[15], Polyphen-2[16], PROVEAN (Protein Variation Effect Analyzer, version 1.1.3)[17], PANTHER-PSEP[18], SNAP2[19] and Mutpred2[20]. To determine the changes in protein stability, the mutations were subjected to two prediction tools: I-Mutant 3.0[21] and MUpro[22]. To determine whether the mutations were associated with disease, two human disease prediction tools were used: PhD-SNP (Predictor of human Deleterious Single Nucleotide Polymorphisms)[23] and SNPs&GO[24]. The ConSurf prediction tool[25] was used to assess the likelihood of mutations affecting the functional regions of the protein.

ASSOCIATION BETWEEN MUTATION STATUS AND CLINICOPATHOLOGICAL PARAMETERS

The association between the mutational status of exon 2 of the C-MYC gene and clinicopathological characteristics (age, DFMP, and marital status) was analyzed using multivariate logistic regression and Rstudio software (version 2022.02.2 Build 485). The degree of association is indicated by odds ratios (ORs) and *p*-values. The significance of these ORs was assessed using Pearson's Chi-square test. A non-significant *p*-value or an OR of 1 indicates no association between the mutation and parameters studied. If the *p*-value is significant and the OR > 1, the

parameter is a risk factor; however, if the OR < 1, the parameter is a protective factor. Statistical significance was set at *p*-value < 0.05.

Results

Epidemiological and clinicopathological characterizations were performed using samples from 30 females. Of these females, 13 had BT and 17 had MT. The mean age of patients with breast BT was 21 ± 5.61 years, with extremes of 13 and 35 years, whereas that of patients with MT was 49.76 ± 11.69 years, with extremes of 30 and 73 years (significant *p*-value; $p < 0.05$). Females aged between 13 and 23 years were found to be more affected by BT (75% of cases), while those aged between 41 and 56 years were mainly affected by MT. The average age of onset of menarche was 14.14 ± 2.11 years and 14.58 ± 1.08 years (non-signficant *p*-value) for the BT and MT groups, respectively, with extremes of 11 and 17 years for BT and 13 and 17 years for MT. In terms of marital status, most females with BT were single (92.31%), whereas those with MT were predominantly married (70.6%) (significant *p*-value). Based on histology, adenofibromas were predominant in the BT group (53.85%), followed by fibrocystic conditions (15.38%). The MT group exclusively had infiltrated ductal carcinomas (100%). The rates of advanced stages (III and IV) and Scarff-Bloom-Richardson (SBR) grade 2 specific to MT were 76.47% and 58.82%, respectively (Table 1).

Table 1: Clinicopathological characteristics of breast tumors

Variables	BT (n = 13)			MT (n = 17)		
		Number	Frequencies		Number	Frequencies
Histology	AF AF + AL Fib NP	7 1 2 3	53.85% 7.69% 15.38% 23.08%	ICC	17	100%
Grade SBR				Grade 1 Grade 2 Grade 3 NP	3 10 1 3	17.65% 58.82% 5.88% 17.65%
Stade				(I +II) (III + IV)	4 13	23.53% 76.47%

BT, benign tumor; MT, malignant tumor; NP, not specified; AF, adenofibroma; Fib, fibrocystic state; ICC, invasive ductal carcinoma; LA, associated lesions (e.g., epithelial hyperplasia).

Genetic Characterization

Exon 2 of the C-MYC gene was sequenced in 45 samples, including 19 C, 11 BT, and 15 MT samples. Sequencing failure occurred in four samples. After correction and alignment, the obtained sequences were found to be 453 bp long and located between codons 25 and 175 of the C-MYC gene.

IDENTIFICATION OF MUTATIONS

Chromatogram analysis using the reference sequence (NM_002467.6) revealed the presence of four types of heterozygous mutations in blood samples of Cs and patients with breast cancer (MT). BTs had no mutations in the C-MYC gene. A recurrent mutation (c.115 T > TC), which was equally present in C and MT, was identified at a frequency of 26.67% in MT and 10.53% in C. The other three mutations were found exclusively in C with the same frequency (5.26%). All identified mutations were new. In fact, these variants were not found in the COSMIC, ClinVar, Ensembl, or dbSNP database. Approximately, 3/4 of the mutations were non-synonymous (missense), two of which (c.115 T > TC and c.117 C > CT) induced the replacement of tyrosine by histidine at position 39 (p.Tyr39His) and had two protein consequences: (p.Tyr39His) and (p.Tyr39Tyr). In contrast, the mutation (c.113 T > TC) results in a change of phenylalanine to serine at position 38 (p.Phe38Ser). The remaining mutations were found to be synonymous (Table 2).

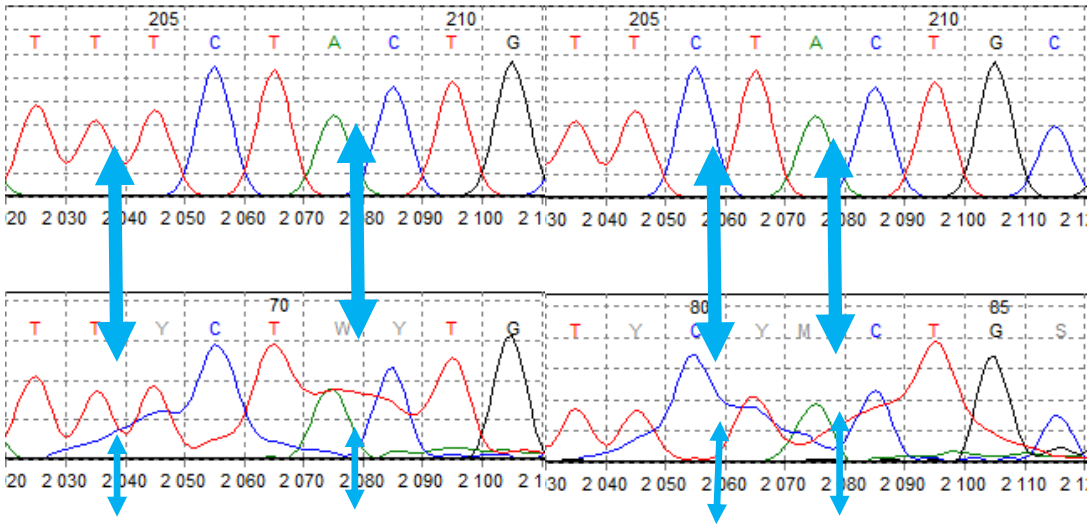
Table 2: Summary of the C-MYC exon 2 mutations found in breast tumors based on the Mutation Surveyor software

Genomique Pos.	Nucleotide Pos.	cDNA Mutation Pos.	Amino acid	Mutation type	Tissues	Freq.
8:127738330	3650	c.113 T > TC	Phe38Ser	Non synonymous	C20	1/19 C (5.26%)
8:127738332	3652	c.115 T > TC	Tyr39His	Non synonymous	MT12 MT13 MT14 MT15	4/15 MT (26.67%)
					C15 C19	2/19 C (10.53%)
8:127738334	3654	c.117 C > CT	Tyr39His	Non synonymous	C19	1/19 C (5.26%)
8:127738334	3654	c.117 C > CT	Tyr39Tyr	Synonymous	C20	1/19 C (5.26%)

Pos, position; Nt, nucleotide; cDNA, coding DNA; Freq, frequency: Ser, serine; Phe, phenylalanine; Tyr, tyrosine; His, histidine; C, cytosine; T, thymine

The positions of the heterozygous mutations in the chromatograms are indicated by blue arrows (Figure 1). The mutations c.113 T > TC and c.115 T > TC comprised the replacement of the T nucleotide with TC nucleotides at positions 113 (individual C20) and 115 (individual C19), respectively, in the DNA encoding the C-MYC gene. For the c.117 C > CT mutation, C was replaced by CT at position 117 in the DNA encoding the same gene for the two individuals, namely C20 and C19.

Figure 1: Nature and position of C-MYC exon 2 mutations determined using Mutation Surveyor software



PREDICTION OF THE PATHOGENICITY OF THE MUTATIONS

Prediction analyses were performed for two mutations: p.Phe38Ser, found in the blood of a single control and p.Tyr39His, found in the blood samples of specific control individuals and patients with breast cancer. According to six prediction tools (SIFT, Polyphen-2, PROVEAN, PANTHER-PSEP, SNAP2, and Mutpred2), the two mutations, p.Phe38Ser and p.Tyr39His, affect the function of the C-MYC protein (Table 3). These proteins were predicted to be damaging by all prediction tools.

Table 3: Impact of non-synonymous mutations on protein function

Substitution	SIFT		Polyphen-2 (HumDiv)		PROVEAN		PANTHER-PSEP		SNAP2		Mutpred2	
	Pred	Score	Pred	Score	Pred	Score	Pred	Score	Pred	Score	Pred	Score
Phe38Ser	DEL	0	PRO DOM	0.99	DEL	- 5.79	PRO DOM	1036	effect	71	PAT	0.88
Tyr39His	DEL	0	POT DOM	0.91	DEL	- 3.39	PRO DOM	1036	effect	75	PAT	0.88

Pred, prediction; DEL, deleterious; PRO DOM, probably damaging; POT DOM, potentially damaging; PAT, pathogenic

According to the I-Mutant 3.0 and MUpro prediction tools (Table 4), the p.Phe38Ser and p.Tyr39His mutations tended to decrease the stability of the C-MYC protein structure, with the decrease being greater with the p.Phe38Ser mutation [DDG (p.Phe38Ser) < DDG (p.Tyr39His)].

Table 4: Impact of non-synonymous mutations on protein stability

Substitution	I-mutant 3.0			MUpro	
	Prediction	DDG (Kcal/mol)	RI	Prediction	DDG (Kcal/mol)
Phe38Ser	Decrease	- 1.87	9	Decrease	- 2.22
Tyr39His	Decrease	- 1.31	6	Decrease	- 1.61

According to the PhD-SNP and SNPs&GO predictions (Table 5), the p.Phe38Ser and p.Tyr39His mutations could be associated with disease development. The p.Phe38Ser mutation was associated with the greatest risk [RI (p.Phe38Ser) > RI (p.Tyr39His]).

Table 5: Impact of non-synonymous mutations on disease development

Substitution	PHD-SNP		SNPs&GO	
	Prediction	RI	Prediction	RI
Phe38Ser	Desease	8	Desease	7
Tyr39His	Desease	3	Desease	4

According to the ConSurf results (Table 6), the Phe38 residue was highly conserved during evolution (score of 9), and no amino acid variation occurred at position 38 in the C-MYC protein. The Tyr39 residue was relatively well conserved.

Table 6: Conservation status of the protein surfaces affected by mutations

Substitution	Prediction	Score
Phe38Ser	Phe38, a highly conserved and buried residue, predicted to have a functional role	9
Tyr39His	Tyr39, buried residue	6

RELATIONSHIP BETWEEN MUTATIONS AND CLINICOPATHOLOGICAL CHARACTERISTICS

Logistic regression analysis revealed no significant association between the presence of at least one mutation and the clinicopathological parameters investigated (age, DFMP, and marital status; $p > 0.05$; Table7). The same result was found for the p.Phe38Ser and p.Tyr39His mutations.

Table 7: Association between mutational status of the C-MYC gene and some clinicopathological parameters

Parameters	Mutations		Phe38Ser		Tyr39His	
	OR	*P-value*	OR	*P-value*	OR	*P-value*
Age	1.02	*0.55*	$2.82.10^{-07}$	*0.99*	1.05	*0.22*
DFMP	$9.48.10^{-01}$	*0.71*	$8.81.10^{-05}$	*0.99*	$9.9.10^{-01}$	*0.94*
Marital status	$3.3.10^{+07}$	*0.99*	$1.83.10^{+45}$	*0.99*	$4.4.10^{+07}$	*0.99*

DFMP, date of first menstrual period

Discussion

This study aimed to determine the involvement of the c-MYC oncogene in breast tumors in Senegalese females. Analysis of the two types of breast tumors in relation to age revealed a predominance of the benign form in females with an average age of 21 years (representative age range: 11–23 years) and a high frequency of cancers in females with an average age of 49.76 years (representative age range: 41–56 years). Most of these cancers were high-grade II (58.82%) and were diagnosed at more advanced stages III and IV (76.47%). Our results regarding BTs were comparable to those of Gueye et al.[7], who found a greater frequency (69.9% of cases) of BTs in females with an average age of 24 years (between 11 and 30 years) in Senegal. Ibrahim et al.[26] found that in Nigeria, the most represented age range for BTs was 21–30 years, with an average age of 29. Consequently, benign breast tumors may be preserved in adolescent and premenopausal females.Additionally, this finding is supported by several studies[8]. In the present study, the age range obtained for MTs were significantly lower than that in the literature from industrialized countries but were consistent with that from studies in sub-Saharan Africa[27]. Breast cancer occurs predominantly in females, with an age range of 42 to 53 years, in a highly aggressive form (high-grade, advanced stage), which reduces the chances of cure and survival of patients[28]. The diagnosis of breast cancer at an advanced stage in sub-Saharan Africa could be explained by several factors: inadequate health education (particularly on cancer), lack of specialized cancer care facilities, shortage of qualified health personnel, precarious socioeconomic conditions of the populations and the costly treatment, neglect of the first symptoms of the disease, recourse to traditional medicine as first-line treatment, and perceptions of the disease (mystical, bewitchment, and evil eye)[29].

In the present study, all females (100%) with breast cancer experienced their first menarche after the age of 12 years. The mean age of patients at menarche was 14.58 years. Our results align with those of Mbaye et al.[30], who found a very high predominance (94.1% of cases) of breast cancer in females who had their first menarche after the age of 12 years in Senegal. However, our findings contrast with those reported previously in the literature; most females with breast cancer had their first menstrual period before 12 years old. Our results do not support the important influence of age at menarche on breast cancer development in Senegalese females, which could be due to certain

internal genetic factors specific to Senegalese females or dietary factors. Benign breast tumors are common during early menarche (between 11 and 17 years of age) and adolescence, suggesting a hormonal influence linked to puberty on the excessive proliferation and abnormal involution of breast tissue during this period in Senegalese females.

Among breast BTs, adenofibroma proved to be the most frequent histological variant (53.85%), followed by fibrocystic states (15.38%). Adenofibroma was also the predominant histological type found by Doupa et al.[31] (76.57%), Gueye et al.[7] (86.3%), and Ibrahim et al.[26] (47.1%). Our results disagree with those obtained in a study in Sudan[32], wherein fibrocystic states were reported as the most frequent benign lesions (39.3%), followed by adenofibroma (30.1%). According to Ageep[32], this difference could be explained by differences in social habits, diet, and the environment. Herein, MTs were represented by a single histological type: infiltrating ductal carcinoma (100% of cases). ICC is the most dominant malignant breast tumor pathology, with frequencies ranging from 40% to 75% depending on the study[33] and may be explained by the small sample size (15).

Regarding marital status, most females with BT were single (92.31%), whereas most females with MT were married (70.60%). Our results align with those of Lutula[1] who found a high frequency of single females with BT (48.15%) and married females with MT (70.31%). The results regarding MT oppose those of Li et al.[34], who observed a higher frequency of celibacy among females with MTs in China. According to Li et al.[34], single females are at a greater risk of developing the disease because they are, for the most part, less exposed to certain protective factors against breast cancer, such as parity, breastfeeding, and early motherhood.

The search for mutations via chromatogram analysis revealed that 3/4 of the newly identified heterozygous mutations were found exclusively in the C group. The remaining 1/4 was recurrent and found in the C and MT groups. Furthermore, these mutations were absent in the BT group. Our findings indicate a very low C-MYC polymorphism rate in breast tumors and are consistent with those of Chakravorty et al.[35], who reported a low polymorphism of C-MYC. All nucleotide substitutions were of the transition type. These nucleotide substitutions induce less significant changes in the DNA structure and are less likely to cause amino acid substitutions[36]. However, 3/4 of the transitions found in our study led to missense (non-synonymous) mutations, thereby altering the structure and properties of the C-MYC protein. Therefore, these substitutions are suspicious and do not affect breast tumorigenesis.

An analysis of other non-synonymous mutations in the protein encoded by the C-MYC oncogene using predictive tools revealed that they could be detrimental to the proper functioning of the protein. Indeed, the recurrent p.Tyr39His mutation, localized to both MT and C, as well as the C-specific p.Phe38Ser mutation, led to modification of the amino acid sequence of the C-MYC protein. These mutations were predicted to be detrimental to the function and structure of the resulting protein using all prediction tools (SIFT, Polyphen-2, PROVEAN, PANTHER-PSEP, SNAP2, Mutpred2, I-Mutant 3.0, and MUpro). In addition, these mutations are associated with the development of a disease,

which may be tumor pathology, according to the PhD-SNPs, SNPs, and GO prediction tools. As a result, non-synonymous (missense) mutations have deleterious effects on the protein (e.g., alterations in protein folding, protein stability, functional domains, and sites of interaction with other proteins) and influence disease susceptibility[37]. Based on the difference between the two mutations, p.Phe38Ser appeared to have more far-reaching effects than p.Tyr39His. The p.Phe38Ser substitution causes the replacement of the amino acid phenylalanine, which is strongly conserved during evolution according to ConSurf (score of 9), with another amino acid, serine. Phenylalanine was also predicted by ConSurf to be a residue buried within the C-MYC protein and play a functional role. This alteration is a radical amino acid change because serine is a small polar molecule while phenylalanine is a large hydrophobic molecule[38]. Meanwhile, the p.Tyr39His substitution resulted in a change from tyrosine, which is predicted to be moderately well-conserved in evolution (ConSurf score of 6), to histidine.

These two amino acids are large molecules that share common features (hydrophobicity and aromaticity)[38]. Crucial amino acids, which are essential for the structure and functionality of a protein, are located in conserved regions of the protein[39]. Consequently, we can infer that the p.Phe38Ser substitution has markedly more deleterious effects on the structural and functional properties of the C-MYC protein than the p.Tyr39His substitution. Cs with p.Tyr39His and p.Phe38Ser substitutions should be monitored to ensure that they do not develop neoplasia. Nevertheless, these two substitutions could constitute molecular markers of breast cancer and contribute to early tumor diagnosis. In contrast, the non-synonymous transversion mutations found in our study may have been induced by genetic or dietary factors linked to the environment or lifestyle. Transversions are eliminated by natural selection because they are more likely to modify amino acids.

Tumors are considered diseases of multifactorial origin that result from the interaction between numerous factors (genetic, environmental, hormonal, and dietary habits) that induce genetic deregulation (mutations). In this study, we assessed the association of three clinicopathological parameters (age, DFMP, and marital status) with the identified mutations. Based on our results, the p.Phe38Ser and p.Tyr39His mutations were not significantly associated with these parameters. However, our sample size was only modest and other parameters were not included in the analysis. Thus, conclusions regarding the association between these mutations and the parameters investigated could not be drawn.

Conclusion

The pathology of female breast tumors is a frequent context of medical consultations. These tumors affect females of all ages from puberty to menopause. BTs account for most cases; however, despite their favorable prognosis, BTs remain a concern for females and practitioners due to fear of breast cancer development.

The main objective of the present study was to provide a better understanding of the molecular profile of breast tumors caused by

mutations in the C-MYC gene in Senegalese females. Based on the results of our epidemiological analyses, our study population was relatively younger than that of studies in Western countries. Benign forms of tumors are more prevalent during adolescence and premenopause, are dominated by adenofibromas, and are more common in single females. Conversely, malignant forms are more frequently encountered in older married females and are mainly of the aggressive infiltrating ductal carcinoma type. The results of our genetic analyses revealed a low variability in exon 2 of C-MYC, with mutations in both healthy and cancerous tissues.

Despite the recurrence of tyrosine replacement with histidine (p.Tyr39His), and the low frequency of phenylalanine conversion to serine (p.Phe38Ser), the penetrance of mutations in this gene remains incomplete. These two predicted damaging mutations, with p.Phe38Ser having more far-reaching effects, are likely involved in the onset of tumorpathology. Therefore, these mutations could be used as molecular markers for breast cancer and may contribute to early diagnosis. However, the risk factors associated with these mutations have not been identified.

Conflicts of Interest Statement:

None

Acknowledgements Statement:

None

Funding Statement:

None

References:

1. Lutula S. Etude épidémiologique, clinique et morphologique des tumeurs du sein au Mali, Thèse de doctorate en Medecine. Bamako: Université de Bamako, 2008; 120p.

2. GLOBOCAN. Cancer Today. [en ligne]. In: The Global Cancer Observatory (GCO) – International Agency for Research on Cancer (IARC); Lyon, France. 2020. Disponible sur: « https://gco.iarc.fr/today/online-analysis-pie?v=2020&mode=population » (Consulté le 7/04/ 2021)

3. Sun YS, Zhao Z, Yang ZN, et al. Risk Factors and Preventions of Breast Cancer. *Int J Biol Sci*. 2017; 13(11):1387–1397.

4. Coyne E, Wollin J, Creedy DK. Exploration of the family's role and strengths after a young woman is diagnosed with breast cancer: Views of women and their families. *Eur J Oncol Nurs*. 2012; 16(2):124–130.

5. Trimble EL. Breast Cancer in Sub-Saharan Africa. *J Glob Oncol*. 2017; *3*(3):187–188.

6. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin*. 2021; 71(3):209–249.

7. Gueye SMK, Gueye M, Coulibaly MT, Mahtouk D, Moreau JC. Tumeurs bénignes du sein à l'unité de sénologie du Centre Hospitalier Universitaire Aristide Le Dantec de Dakar (Sénégal). *Pan Afr Med J*. 2017; 27(251):7p.

8. Stachs A, Stubert J, Reimer T, HarMTann S. Benign Breast Disease in Women. *Deutsches Ärztebl Int*. 2019; 116(33-34):565–574.

9. Sasaki J, Geletzkea A, Kass RB, Klimberg VS, Copeland III EM, Bland KI. Section III: Benign and Premalignant Lesions – 5: Etiology and Management of Benign Breast Disease. *In*: Bland K.I., Copeland III E.M., Klimberg V.S. et Gradishar W.J. (Eds), The Breast: Comprehensive Management of Benign and Malignant Diseases. Fifth Edition, Amsterdam (Pays-Bas): Elsevier. 2018;79 – 92.e5.

10. Polyak K, Haviv I, Campbell IG. Co-evolution of tumor cells and their microenvironment. *Trends Genet*. 2009; 25(1) :30–38.

11. Carabet LA, Rennie PS, Cherkasov A. Therapeutic Inhibition of Myc in Cancer. Structural Bases and Computer-Aided Drug Discovery Approaches. *Int J Mol Sci*. 2018; 20(1):120.

12. Green AR, Aleskandarany MA, Agarwal D, et al. MYC functions are specific in biological subtypes of breast cancer and confers resistance to endocrine therapy in luminal tumours. *Br J Cancer*. 2016; 114:917–928.

13. Hall TA. BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symp Ser*. 1999;41;95–98.

14. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; 22(22):4673–4680.

15. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Gen Res*. 2001; 11(5):863–874.

16. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7(4):248–249.

17. Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino

acid substitutions and indels. *Bioinformatics*. 2015; 31(16):2745–2747.

18. Tang H, Thomas PD. PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics*. 2016; 32(14) :2230–2232.

19. Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genomics*. 2015; 16 Suppl 8(8):S1.

20. Pejaver V, Urresti J, Lugo-Martinez J, et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat Commun*. 2020; 11:5918.

21. Kumar MD, Bava KA, Gromiha MM, et al. ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. *Nucleic Acids Res*. 2006; *34*(Database issue);D204–D206.

22. Cheng J, Randall A, Baldi P. Prediction of Protein Stability Changes for Single-Site Mutations Using Support Vector Machines. *Proteins*. 2006; 62(4):1125–1132.

23. Capriotti E, Calabrese R, Casadio R. Predicting the insurgence of human genetic diseases associated to single point protein mutations with Support Vector Machines and evolutionary information. *Bioinformatics*. 2006; 22:2729–2734.

24. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat*. 2009; 30:1237–1244.

25. Ashkenazy H, Abadi S, Martz E, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules, *Nucleic Acid Res. 2026;*44: W344–W350.

26. Ibrahim IM, Iliyasu Y, Mohammed AZ. Histopathological review of breast tumors in Kano, Northern Nigeria. *Sub-Saharan Afr J Med*. 2015; 2(1):47–51.

27. Brinton LA, Figueroa JD, Awuah B, et al. Breast cancer in Sub-Saharan Africa: opportunities for prevention. *Breast Cancer Res Treat*. 2014; 144(3):467–478.

28. Ly M, Antoine M, André F, Callard P, Bernaudin J-F, Diallo DA. Le cancer du sein chez la femme de l'Afrique sub-saharienne: état actuel des connaissances – Breast cancer in Sub-Saharan African women: review. *Bull Cancer*. 2011; 98(7):797–806.

29. Kantelhardt EJ, Muluken G, Sefonias G, et al. A Review on Breast Cancer Care in Africa. *Breast Care*. 2015; 10(6):364–370.

30. Mbaye F, Dem A, Fall M, et al. Correlation of Clinical and Pathological Parameters with the Diversity and Genetic Evolution of Breast Cancer in Senegalese Women. *Br J Med Res*. 2014;4(14):2627-2638.

31. Doupa D, Badji MN, Mbaye F, et al. Benign Breast Tumors among Senegalese Women: Diversity and Genetic Evolution of D-Loop. *Open Access Libr J*. 2015; 2:e1758.

32. Ageep AK. Benign breast tumors in Red Sea State, Sudan. *J Cancer Res Exp Oncol*. 2011; 3(7):84–87.

33. DucaMTan BS, Wang HH. Chapter 9 – Breast. In: Cibas E.S. et DucaMTan B.S. (Eds), Cytology, Diagnostic Principles and Clinical Correlates. Fourth Edition, Etats-Unis d'Amérique: Elsevier (Saunders). 2014;233–265.

34. Li M, Han M, Chen Z, et al. Does marital status correlate with the female breast cancer risk? A systematic review and meta-analysis of

observational studies. *PloS ONE*. 2020; 15(3):e0229899.

35. Chakravorty D, Jana T, Das Mandal S, Seth A, Bhattacharya A, Saha S. MYCbase: a database of functional sites and biochemical properties of Myc in both normal and cancer cells. *BMC Bioinformatics*. 2017; 18(1):224.

36. Guo C, McDowell IC, Nodzenski M, et al. Transversions Have Larger Regulatory Effects than Transitions. *BMC Genomics*. 2017; 18(1):394.

37. Stefl S, Nishi H, Petukh M, Panchenko AR, Alexov E. Molecular mechanisms of disease-causing missense mutations. *J Mol Biol*. 2013; 425(21):3919–3936.

38. Exarchos KP, Exarchos TP, Rigas G, et al. Extraction of consensus protein patterns in regions containing non-proline cis peptide bonds and their functional assessment. *BMC bioinform*. 2011; 12(142):12 p.

39. Risso VA, Manssour-Triedo F, Delgado-Delgado A, et al. Mutational Studies on Resurrected Ancestral Proteins Reveal Conservation of Site-Specific Amino Acid Preferences throughout Evolutionary History. *Mol Biol Evol*. 2015; 32(2):440–455.