

Article

Clinically Relevant Germline Mutations in a Cohort of Young Women with Breast Cancer: A Comprehensive Analysis of Hereditary-Cancer Genes from Whole-Exome Sequencing

Liliana Gómez-Flores-Ramos ^{1*}, Michael Dean², Kristine Jones ³, Mingyi Wang ⁴, Cynthia Villarreal-Garza ⁵, Rosa María Álvarez-Gómez ⁶, Talia Wegman-Ostrosky ⁷, Nancy Reynoso-Noverón ⁸, Verónica Fragoso-Ontiveros ⁹, Luisa Sánchez-Zamorano ¹⁰, Miguel Trujillo-Martínez ¹¹, Lizbeth Grimaldo ¹², Patricia Gallegos-Arreola ¹³, Patricia Ostrosky-Wegman ¹⁴, Alejandro Mohar-Betancourt ¹⁵.

¹ Cátedras-CONACYT, Mexico; Population Health Research Center, National Institute of Public Health; Cuernavaca, Mexico; liliana.gomez@insp.mx

² Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Gaithersburg, MD, United States; deanm@mail.nih.gov

³ Cancer Genetics Research Laboratory, Division of Cancer Epidemiology and Genetics, Leidos Biomedical Research Inc.; Frederick National Laboratory for Cancer Research, Gaithersburg, MD, United States; kristine.jones@nih.gov

⁴ Cancer Genetics Research Laboratory, Division of Cancer Epidemiology and Genetics, Leidos Biomedical Research Inc.; Frederick National Laboratory for Cancer Research, Gaithersburg, MD, United States; mingyi.wang@nih.gov

⁵ Breast Cancer Center, Hospital Zambrano Hellion, Tecnológico de Monterrey; Monterrey, Mexico; cynthiavg@gmail.com

⁶ Research division, National Cancer Institute; Mexico City, Mexico; rosamag2@hotmail.com

⁷ Research division, National Cancer Institute; Mexico City, Mexico; taliaw@gmail.com

⁸ Research division, National Cancer Institute; Mexico City, Mexico; nrn231002@yahoo.com.mx

⁹ Research division, National Cancer Institute; Mexico City, Mexico; ontiverosfvero@gmail.com

¹⁰ Population Health Research Center, National Institute of Public Health; Cuernavaca, Mexico; szamorano@insp.mx

¹¹ General Hospital No.7. Cuautla, Morelos. Mexican Institute of Social Security, Mexico; Population Health Research Center, National Institute of Public Health; Cuernavaca, Mexico; csmet@hotmail.com

¹² Department of Cell Biology, Centre for Research and Advanced Studies of the National Polytechnic Institute, Mexico City, Mexico; lizbeth.grimaldo@cinvestav.mx

¹³ Western Biomedical Research Center, Guadalajara, Jalisco. Mexican Institute of Social Security, Mexico; marthapatriciagallegos08@gmail.com

¹⁴ Institute of Biomedical Research, National Autonomous University of Mexico, Mexico City, Mexico; patricia.ostrosky@gmail.com

¹⁵ Institute of Biomedical Research, National Autonomous University of Mexico, Mexico City, Mexico; Research division, National Cancer Institute; Mexico City, Mexico; alejandromohar@gmail.com

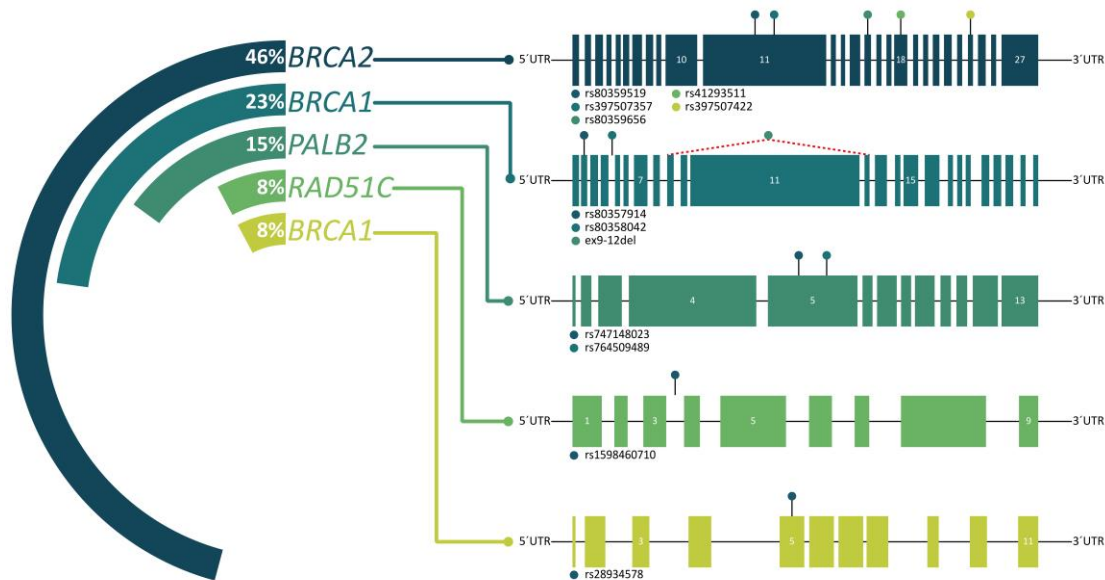
* Correspondence: liliana.gomez@insp.mx; tel.: +52 55 54 63 43 02

Abstract: Young women with breast cancer represent 15% of cancer cases in Latin America. Genomic studies have found that early-onset breast-cancer cases exhibit a higher genetic susceptibility and a specific genomic signature as compared to their older counterparts. The aim of this study was to describe clinically relevant germline mutations in a cohort of young women with breast cancer. To achieve this, we analyzed hereditary-cancer genes from whole-exome

sequencing data in 108 unrelated women with an extreme phenotype of breast cancer (≤ 40 years of age), diagnosed and treated at the National Cancer Institute of Mexico; 11% of the patients carried a pathogenic variant. *BRCA2* comprised 46% of the mutations, followed by *BRCA1* with 23%; *PALB2* with 15%; and *TP53* and *RAD51C* with 8 % each. This article describes a novel pathogenic mutation in *RAD51C* c.519dupT. The median age at diagnosis was 35 years overall; however, it was six years younger in patients with mutations. Age at diagnosis ($OR=0.82$, $CI\ 95\% \ 0.71-0.94$; $P=0.008$) and first-degree family history of cancer ($OR=8.26$, $CI95\% \ 1.35-50$; $P=0.022$) were the only epidemiological variables associated with mutational status. We found no differences in disease-free survival ($p=0.403$) or overall survival ($p=0.735$) among mutational status subgroups.

Keywords: early-onset breast cancer, hereditary cancer, whole-exome sequencing, young women

Graphical Abstract



Graphical abstract. Frequency and distribution of clinically relevant germline mutations in young Mexican women with breast cancer

1. Introduction

Breast cancer accounts for one out of four cancer cases among women. Aside from being the most frequently diagnosed cancer worldwide, it is the leading cause of death by neoplasia [1]. Breast cancer commonly occurs during the sixth decade of life. However, between 6% and 10% of cases are observed in women <40 years of age [2]. In low- and middle-income countries, the incidence and mortality rates for breast cancer in young women are twice as high as those for high-income countries (14% vs. 7%, respectively) [3]. Early-onset breast cancer is a growing burden in Latin America, and it is the leading cause of death from cancer in Mexican women, accounting for 15% of cancer-related deaths [3]. Additionally, it is the second leading cause of death among women aged 30-54.

Breast cancer in young women, defined as cases in women equal or under the age of 40, poses a challenge to public health, given its implications as regards the considerable number of years (45 on average) of life lost, late diagnosis, and poor prognosis [4].

In breast-cancer patients without age stratification, 10%–30% of cases are attributed to family history, while 5%-10% have an important inherited component explained by autosomal-dominant,

highly penetrant pathogenic mutations [5]. To diagnose these hereditary breast cancers, international guidelines suggest that patients <50 years of age receive genetic counseling. The identification of clinically relevant germline mutations in this group of patients can have a substantial impact on treatment, secondary prevention, and medical follow-up. Other family members can benefit from early screening and prevention as well [6].

Clinical presentation at a young age, with family aggregation, can lead to a suspicion of hereditary breast cancer. However, thus far, the contribution of pathogenic variants has only been partially characterized in young women.

The aim of this study was to determine the prevalence, clinical impact and risk factors associated with hereditary cancer mutations in young women with breast cancer, using whole-exome analysis.

2. Results

2.1. Description of breast cancer cases

This study included 108 cases of early-onset breast cancer. The median age at diagnosis was 35 years (interquartile range: 32-38); however, in women with a positive mutation, it was 28.5 (26.5-38) years. Almost 50% of the cases were diagnosed at an advanced clinical stage (III-IV). Ductal carcinoma represented 83% of the tumors, whereas 58% had reached a high histological grade. Almost 65% of the tumors tested positive for estrogen and/or progesterone receptors. In addition, 20% were Her2Neu positive and 27% were Triple-Negative. No statistical differences emerged among mutational status subgroups. More details on the clinical and pathological characteristics of patients are described in Table 1.

Table 1. Clinical characteristics of young women with breast cancer

Clinical characteristics of patients	Patients without mutation n=96 (89%)		Patients with mutation n=12 (11%)		p*	Total (n=108)	
Age at diagnosis	35	(32, 38)	28.5	(26.5, 38)	0.107	35	(30, 38)
Tumor size	3.5	(2.4, 6.0)	3.3	(1.9, 5)	0.50	3.5	(2.1, 5.8)
Clinical stage at diagnosis	n	%	n	%		n	%
I	14	14.6	1	8.3	0.760	15	13.9
II	37	38.5	4	33.3		41	38.0
III	39	40.6	7	58.3		46	42.5
IV	6	6.3	0	0.0		6	5.6
Histology							
Ductal	80	83.3	10	83.3	0.768	90	83.4
Lobular	4	4.2	1	8.3		5	4.6
Mixed	8	8.3	1	8.3		9	8.3
Others	4	4.2	0	0		4	3.7
Histologic grade							
Low	8	8.3	1	8.3	0.895	9	8.3
Intermediate	32	33.3	5	41.7		37	34.3
High	56	58.3	6	50.0		62	57.4
Immunohistochemical biomarkers							
Estrogen receptor	62	64.6	8	66.7	0.580	70	64.8
Progesterone receptor	60	62.5	7	58.3	0.506	67	62.0

HER2 neu	21	22.9	1	8.3	0.248	22	20.4
Ki67 (n=105)							
<20%	28	29.8	4	36.4	0.445	32	30.5
≥20%	66	70.2	7	63.6		73	69.5
Breast cancer subtype							
Luminal A	22	22.9	4	33.3	0.865	26	24.0
Luminal B							
(Her2-negative)	28	29.2	3	25.0		31	28.7
Luminal B							
(Her2-positive)	13	13.5	1	8.3		14	13.0
HER2-positive							
(non-luminal)	8	8.3	0	0.0		8	7.4
Triple-negative	25	26.0	4	33.3		29	26.9
Laterality							
Right	52	50	6	50	0.516	58	53.7
Left	41	41.67	5	41.7		46	42.6
Billateral	3	8.33	1	8.3		4	3.7
Treatment							
Surgical treatment	89	92.7	11	91.7	0.623	100	92.6
Breast-conserving							
surgery	40	44.9	3	27.3	0.345	43	43.0
Mastectomy	49	55.1	8	72.7		57	57.0
Chemotherapy	87	90.6	11	91.7	0.693	98	90.7
Adjuvant	38	43.7	3	27.3	0.592	41	41.8
Neoadjuvant	45	51.7	8	72.7		53	54.1
Palliative	4	4.6	0	0.0		4	4.1
Radiotherapy	69	71.9	9	75.0	0.561	78	72.2
Hormonal therapy	63	65.6	5	41.7	0.098	68	63.0
Deaths **	10	10.4	1	8.3	0.648	11	10.2
Recurrences**	20	20.8	4	33.3	0.259	24	22.2

*p value corresponding to the Wilcoxon Test for continuous variables and Fisher's Exact Test for categorical variables. **After a 48-month (30-54) follow-up. Age (yrs; median, interquartile range) at diagnosis and tumor size (cm; median, interquartile range).

At the time of diagnosis, 60% of the women analyzed were overweight or obese, and 19% exhibited altered plasma glucose. In addition, 4% were pregnant and 25% were nulliparous. In the group with positive mutations, 33% had a first-degree history of cancer, as compared to 15% among the patients without positive mutations (see Table 2). The majority of patients came from central Mexico (76.9%); 65.7% were married, and 36% had attained higher education (see Supplementary Table 1).

Table 2. Clinical and reproductive features of young women with breast cancer

Clinical characteristics of patients	Patients without mutation		Patients with mutation		p	Total	
	n= 96 (89%)		n=12 (11%)			(n=108)	
	n	%	n	%		n	%
BMI (kg/m2)	26.3	(23.7, 29.2)	26.47	(25.3, 28.8)	0.629	26.4	(23.8, 29.2)
Normal (18.6-24.9)	38	39.6	3	25.0	0.671	41	38.0
Overweight (25-29.9)	38	39.6	6	50.0		44	40.7

Obese (≥ 30)	20	20.8	3	25.0		23	21.3
Plasma glucose levels							
<100 mg/gL	78	81.3	9	75.0	0.600	87	80.6
100-125 mg/dL	15	15.6	3	25.0		18	16.7
≥ 126 mg/dL	3	3.1	0	0		3	2.8
Reproductive risk factors							
Age at menarche, years	12	(11, 14)	13	(12, 13.5)	0.274	12	(11.5, 14)
<11	26	27.1	1	8.3	0.354	27	25.0
11 to 13	45	46.8	8	66.7		53	49.1
>13	25	26.1	3	25		28	25.9
Pregnancy	74	77.1	10	83.3	0.474	84	77.8
Pregnant at diagnosis	3	3.1	1	8.3	0.380	4	3.7
Age at first pregnancy, years	22	(19, 26)	21.5	(20, 28)	0.937	22	(19, 26)
Age at last pregnancy, years	29	(25, 33)	27.5	(24, 33)	0.666	29	(25, 33)
Parity							
Nulliparous	25	26.1	2	16.7	0.706	27	25.0
1 child	21	21.8	3	25		24	22.2
2 children	26	27.1	5	41.7		31	28.7
≥ 3 children	24	25	2	16.7		26	24.1
Breastfeeding (positive)	67	69.8	10	83.3	0.27	77	71.3
< 12 months	26	38.2	1	10.0	0.071	27	35.1
≥ 12 months	41	61.2	9	90.0		50	64.9
Use of hormonal contraceptive							
	59	61.5	4	33.3	0.211	63	58.3
< 5 years	47	79.6	2	50.0		49	77.8
≥ 5 years	12	20.3	2	50.0		14	22.2
Other risk factors							
Alcohol consumption	20	20.8	3	25	0.493	23	21.3
Smoking	26	27.1	2	16.7	0.35	28	25.9
Family history of cancer (first and second degree)							
First degree	15	15.6	4	33.3	0.133	19	17.6
Second degree	55	57.4	6	50	0.428	61	56.5

*p value corresponding to the Wilcoxon Test for continuous variables and Fisher's Exact Test for categorical variables. **After a 48-month (30-54) follow-up. Age at menarche (median, p25 p75), age at first pregnancy (median, p25 p75) and age at last pregnancy (median, p25 p75).

2.1. Description of young women with clinically relevant mutations

We found pathogenic mutations in 12 patients (11%), none of them recurrent. The most prevalent mutated gene was *BRCA2*, representing 46% of the hereditary-cancer cases; followed by *BRCA1*, representing 23%. Only one pathogenic mutation corresponded to the Mexican founder mutation (*BRCA1* ex9-12del), a large deletion described in women from the center region of the Country. Two patients (15%) had a mutation in *PALB2*, one patient (8%) in *TP53*, and another (8%) carried a pathogenic mutation in *RAD51C*. The mutation found in *RAD51C* C.519dupT (rs1598460710) had not been reported previously. This pathogenic variant produces a frameshift [7] (see Table 3).

All patients in the mutation subgroup were diagnosed with either a Luminal or a Triple-Negative immunohistochemical subtype, and most had reached an intermediate or high histologic

grade. Half of the tumors had Ki67>20%, 70% were located in the right breast and one was bilateral (see Supplementary Table 2). All but two patients had undergone mastectomies, and four patients had experienced recurrence (see Supplementary Table 3).

Table 3. Characteristics of germline mutations

Patient ID	Gene	Location	dbSNP ID	HGVS	ClinVar classification	Molecular consequence
I	BRCA1	Intron 4	rs80358042	NM_007294.4(BRCA1):c.212+1G>T	Pathogenic	Splice donor
II	BRCA1	Exon 2	rs80357914	NM_007294.4(BRCA1):c.66_67AG[1] (p.Glu23fs)	Pathogenic	Frameshift
III	BRCA1			ex9-12del	Pathogenic	Stop codon
IV	BRCA2	Exon 11	rs80359519	NM_000059.3(BRCA2):c.5542del (p.Ser1848fs)	Pathogenic	Frameshift
V	BRCA2	Exon 18	rs41293511	NM_000059.3(BRCA2):c.8168A>G (p.Asp2723Gly)	Pathogenic	Missense
VI	BRCA2	Exon 15	rs80359656	NM_000059.3(BRCA2):c.7525dup (p.Ser2509fs)	Pathogenic	Frameshift
VII	BRCA2	Exon 11	rs397507357	NM_000059.3(BRCA2):c.5631del (p.Asn1877fs)	Pathogenic	Frameshift
VIII	BRCA2	Exon 24	rs397507422	NM_000059.3(BRCA2):c.9235del (p.Val3079fs)	Pathogenic	Frameshift
IX	PALB2	Exon 5	rs747148023	NM_024675.3(PALB2):c.2411_2412delCT	Pathogenic	Frameshift
X	PALB2	Exon 5	rs764509489	NM_024675.3(PALB2):c.2336C>G (p.Ser779Ter)	Pathogenic	Nonsense
XI	TP53	Exon 5	rs28934578	NM_000546.6(TP53):c.524G>A (p.Arg175His)	Pathogenic	Missense
XII	RAD51C	Intron 3	rs1598460710	NM_058216.2:c.519dup, T/TT	Pathogenic	Frameshift

*This variant is reported for the first time in our study. It is not included in the ClinVar database.

The youngest patient in the group, 19 years old, presented with a *TP53* p.Arg175His mutation. Half of the mutations analyzed affected very young women (<35 years old). In patients <30 years of age, we observed a higher frequency of a first-degree family history of cancer (See Supplementary Table 4).

Patients came from variety of states, all situated in central and southern Mexico (Figure 1, Supplementary Table 4); 67% of patients were married and 83% had given birth to at least one child. One woman was pregnant at the time of diagnosis. All patients with children had breastfed for at least 12 months, and 42% had been using hormonal contraceptives for 50 months on average (See Supplementary Table 5).

Age at diagnosis (OR = 0.82, CI95% 0.71-0.94; P= 0.008) and a first-degree family history of cancer (OR = 8.26, CI95% 1.35-50; P= 0.022) were the only epidemiological variables associated with mutational status, after adjusting for immunohistochemical subtype, breastfeeding, hormonal status and body mass index (BMI). In these cases, we found no association among clinical stage, histologic type, hormone receptor and mutational status. Analysis by individual gene revealed no association between mutations and clinical or pathological characteristics.

After a median cohort follow-up of 1410 days, we found no differences in disease-free survival ($p=0.403$) or overall survival ($p=0.735$) among the mutational status subgroups (Supplementary Figures 1, 2, 3 and 4).

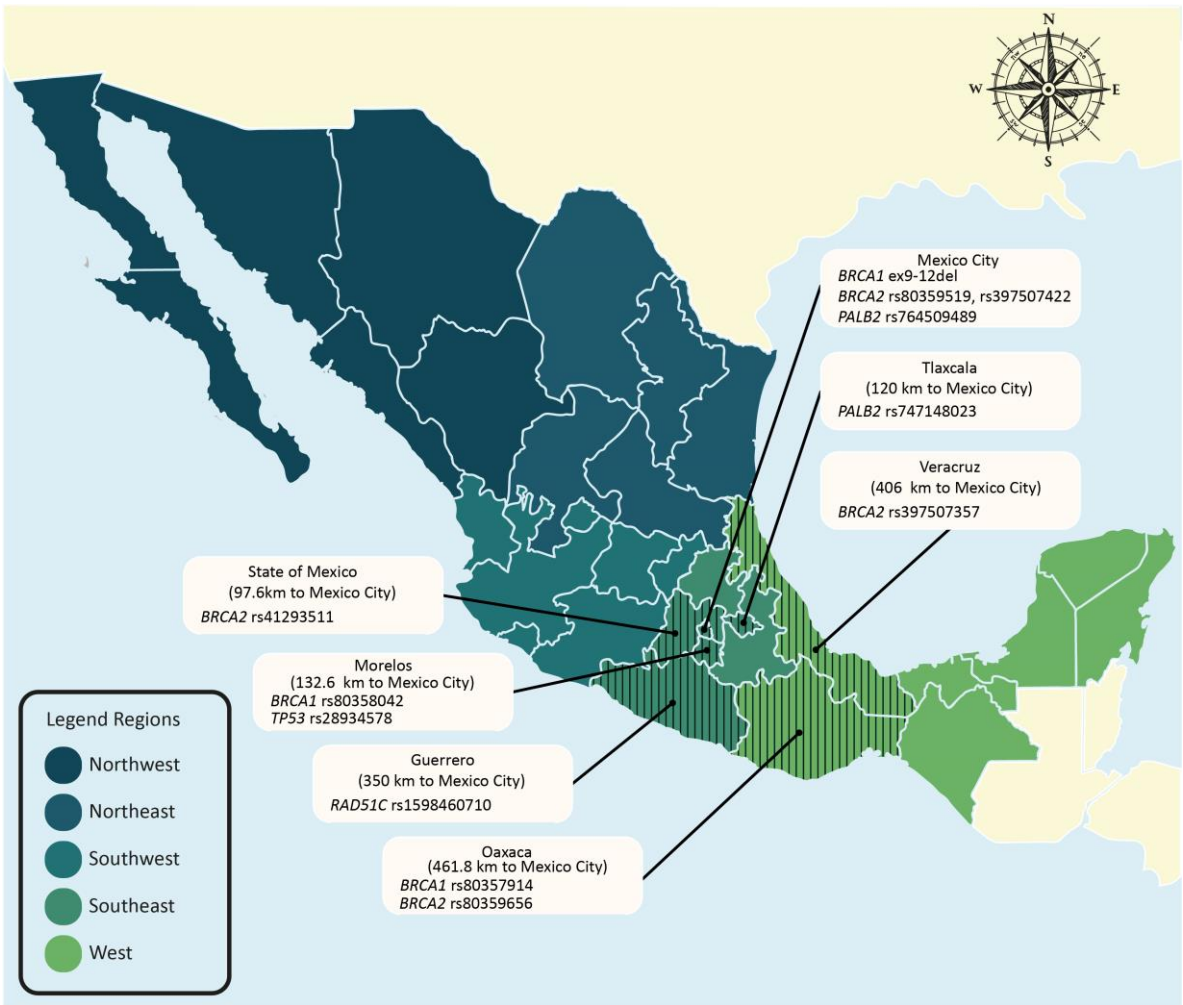


Figure 1. Geographical distribution of clinically relevant germline mutations in young Mexican women with breast cancer

2.3 Variants of uncertain significance (VUS)

In the analyzed genes, we found 30 VUS at the following genes: 8 at *ATM*, 2 at *BRCA1*, 3 in *BRCA2*, 5 in *CHECK2*, 2 in *FANCA*, *MRE11A*, *NF1*, and *PMS2* and one on the following genes: *NBN*, *RAD51C*, *RECQL* and *XRCC2*. The details of this VUS can be consulted in Supplementary Table 6.

3. Discussion

3.1 Characteristics of young women with breast cancer

We used an extreme phenotype sampling strategy based on early-onset breast cancer among participants. Overall, the median age at diagnosis was 35 years; however, it was six years younger (29) in the group with mutations. Breast cancer in young women, a well-described risk factor for hereditary breast cancer [8], can vary among populations [9]. Half of the patients in our study were very young (<35 years old) and therefore at greater risk for hereditary cancer. We found that age at diagnosis was a significant clinical factor ($P= 0.01$) for predicting hereditary breast cancer, after adjusting for other relevant variables such as a family history of cancer, immunohistochemical subtype, breastfeeding, hormonal status and body mass index (BMI). Not surprisingly, the youngest patient in the cohort was diagnosed with Li-Fraumeni syndrome and indicated a positive

family history of cancer. Patients in the cohort diagnosed at an age <30 years showed a higher risk for hereditary breast cancer.

Genetic mutations have been found to be most common among Triple-Negative subtypes [10]. Greenup *et al.* [11] described ethnic differences in the frequency of *BRCA* mutations, with the lowest frequency (20%) pertaining to the Hispanic population. Triple-Negative represented one-third of the cases and half of the *BRCA1/2* mutations in our cohort. However, this did not prove statically different in the negative- and positive-mutation subgroups.

Consistent with reports on populations similar to ours [12], we found that seven out of ten patients analyzed exhibited hormone-receptor-positive and 20% Her2-positive tumors. No statistical differences emerged as regards immunohistochemical subtypes among the mutation and negative mutation groups.

A family history of cancer has been identified as the most important determinant of risk for hereditary cancer syndromes. However, in young women, two factors can modify this association: 1. the younger the woman at the time of diagnosis, the lower the likelihood of cancer occurring in other family members, and 2. by contrast, a younger age at diagnosis denotes a greater likelihood of a high-penetrance gene segregating in the family. The overall family history of cancer in our study was 26%, lower than that reported for other series of unselected cases in Mexico [13, 14]; however, the family history of cancer in women diagnosed with hereditary breast cancer came to 46%. All patients in our study participated in a detailed interview and were provided with genetic counseling by a geneticist at the Hereditary Cancer Clinic. To determine a family history of cancer, we obtained a pathological confirmation of diagnosis for their family members, as well as information on their specific type of cancer. A first- and second-degree family history was found to be twice as frequent in patients with positive mutations compared to those in the negative-mutation subgroup; this difference proved statistically significant.

3.2 Pathogenic variants in young women with breast cancer

Our study evaluated clinically relevant mutations in young women with breast cancer, including a pathogenic mutation described for the first time in *RAD51C* (*c.519dup, T/TT*). We found that 11% of our patients carried pathogenic mutations: 75% of them in *BRCA1* and *BRCA2* genes. This proportion is the same as that reported in the literature for women who experienced the onset of breast cancer at ≤ 40 years of age. Our research group published a review of genetic mutations in young women with breast cancer, indicating that mutations in *BRCA* amounted to 74.4% of mutations. For *TP53*, the proportion found in the present study was 8%, similar to that reported in the literature (10.7%). In contrast to these two sets of genes, *PALB2* (15%) and *RAD51C* (8%) showed significantly higher proportions than previous findings for young women with breast cancer carrying positive mutations (1%) [8]. Recent studies have reported that *PALB2* represents 1%-4% [15-17] of total mutations in early-onset breast cancer. Our study found that, overall, *PALB2* accounted for 2% and *RAD51C* for 1% of the germline pathogenic variants.

A previous study of unselected Mexican women with breast cancer found a frequency of 4.3% for pathogenic mutations [18], while our research, focusing on women ≤ 40 years of age, revealed that the frequency doubled (8%). Significantly, 25% of mutations in our participants occurred in *BRCA1* and 50% in *BRCA2*. These findings differ from those reported by Villarreal-Garza *et al.* [19] who found that *BRCA1* mutations represented 97.7% of total mutations in Mexican women <50 years of age with Triple-Negative breast cancer. Millan-Catalan *et al.* [20] reported that *BRCA1* mutations corresponded to 66% of the total, similar figures to those published by Dean *et al.* for US Hispanics (57%) [21]. However, other groups have associated *BRCA2* mutations with early-onset breast cancer [15]. The Mexican founder mutation *BRCA1* (ex9-11del) accounted for one-third of

BRCA1 mutations found, similar to what Villareal-Garza *et al.* reported for patients <40 years of age [19-22].

The National Cancer Institute of Mexico (NCIMx), the cancer-referral center where we conducted the study, provides diagnoses and treatment for women throughout Mexico. A study of women with breast cancer in northern Mexico [14] reported a 14% frequency of *BRCA* mutations, higher than the 8% figure we obtained for young women from the central, southern, and western regions. The patients in our cohort represented women treated at a single hospital; therefore, the prevalence/distribution of mutations among cancer patients in other geographical regions of Mexico may differ due to heterogenic genetic ancestry.

3.3 High penetrance genes

Hereditary Breast and Ovarian Cancer (HBOC)

HBOC, the most common hereditary-cancer syndrome among the young patients in our study, is caused by pathogenic mutations in *BRCA* genes. Women with *BRCA1* mutations run a lifetime risk for breast cancer as high as 80%. The rate of mutations in this gene is particularly high among women diagnosed at younger ages [23]. Malone *et al.* observed that breast-cancer patients <35 years old with a negative family history of breast cancer had a 9.4% probability of carrying a germline mutation in *BRCA1/2* [24].

The prevalence of *BRCA1/2* mutations varies among populations: it ranges from 0.2% to 0.3% in the general population, but reaches 3% in unselected women with breast cancer, 6% in women with early-onset breast cancer, 10% in women with ovarian cancer, and 20% in those from high-risk families [25]. Our study determined that 3% of early-onset breast cancers were attributable to *BRCA1* and 6% to *BRCA2*.

Tumors associated with *BRCA1* tend to be Triple-Negative; however, *BRCA2*-related tumors resemble sporadic breast tumors and often express hormone receptors [25]. Our study found that most of *BRCA*-related tumors expressed hormonal receptors and were highly proliferative.

Li-Fraumeni syndrome

Li-Fraumeni syndrome is a hereditary-cancer syndrome associated with a diversity of cancers. Early-onset breast cancer is the most common tumor among women with germline mutations in *TP53*. The propensity for primary cancers beginning from a young age is explained by the crucial function of the tumor protein p53. This protein responds to DNA damage by cell cycle arrest, apoptosis, senescence, DNA repair and changes in cell metabolism [26].

In women with Li-Fraumeni syndrome, the lifetime risk for any type of cancer is greater than 90% and up to 50% for breast cancer [27]. The reported frequency of *TP53* mutations in women <35 diagnosed with breast cancer ranges from <1% to 7% [28-30]; this frequency can reach up to 30% in patients diagnosed at <30 years of age [31]. In our population, *TP53* was positive in 1% of the entire cohort and was found in the youngest patient.

3.4 Moderate-penetrance genes

Women <40 with heterozygous germline mutations in *PALB2* are nine to ten times more likely to develop breast cancer than the general population. The absolute breast-cancer risk for *PALB2* female mutation carriers by the age of 70 ranges from 33% for those with no family history of breast cancer to 58% for those with two or more first-degree relatives with this condition [32]. The contribution of *PALB2* mutations to hereditary, early-onset breast cancers negative for *BRCA* mutations is 1%-2% according to studies of various ethnic groups and early-onset breast cancer [33];

this corresponds exactly to the frequency found in this study. It is worth noting that the *BRCA1-PALB2-BRCA2* complex is crucial for DNA double-strand break repair. All three proteins are known to play a significant role in both checkpoint activation and checkpoint maintenance [34]. This study, which included extreme-phenotype breast cancer, found that 15% of mutations were in *PALB2*. Given its importance as an integral component of the *BRCA* complex, some authors have postulated that *PALB2* can be considered a high-penetrance breast-cancer predisposing gene [35].

RAD51, a protein family that responds to DNA damage signaling for break repair, is pivotal for homologous recombination. *RAD51C* interacts with *BRCA1*, *BRCA2*, *PALB2* and *p53* [36]. Mutations in *RAD51* have been previously described in breast and ovarian cancer with familial aggregation [37]. We found a pathogenic frameshift mutation, *RAD51* rs1598460710, an intronic T duplication. This is the first study to report such a mutation in women with breast cancer.

The most recent studies suggest that moderate-penetrance genes explain a proportion of hereditary breast cancers. Germline mutations in moderate-penetrance genes are significantly more often reported in older as compared to younger women [8]; however, this could be due to a bias in genetic testing. The observed disparities may change as genomic testing becomes increasingly available.

Clinical guidelines for patients with mutations in genes such as *BRCA1/2*, *CDH1*, *PTEN*, *STK11*, and *TP53* are available. However, they are limited for moderate-penetrance genes. This represents a challenge for clinicians, who make medical management recommendations based on allele-specific risk information and individual clinical history.

3.5 Opportunities

The recent revolution in genetic sequencing technology opens the door for a rapid increase in knowledge concerning the genetic basis for breast cancer in young women. Although treatment for early-onset breast cancer is similar to that for other age groups, the clinical management of these patients must meet specific needs [38]. Having adequate genetic counseling empowers young women to make optimal decisions about their treatment and follow-up, as well as enabling oncologists to offer patients more targeted treatment and additional prophylactic options. Likewise, a greater opportunity exists to implement cancer prevention programs for other members of the family [39]. The poly adenosine diphosphate-ribose polymerase inhibitors (PARPi) permit crafting a targeted approach for those women with germline *BRCA1/2* mutations who present with advanced-stage breast cancer [40].

3.5 Limitations

Whole-exome sequencing allows for the detection of thousands of changes in DNA. However, it is not optimal for the detection of large deletions. We may therefore have missed the aforementioned type of mutations in this set of patients.

4. Materials and Methods

4.1 Selection of cases and data collection

All procedures employed in this study were performed in strict accordance with the relevant local and international guidelines and regulations. The study, conducted at the National Cancer Institute of Mexico (NCIMx), included 108 women with histopathological breast cancer diagnosed at ≤ 40 years of age. Patients were recruited between September 2015 and December 2017.

We reviewed medical records and electronic files to obtain detailed clinical and sociodemographic information. Data concerning a family history of cancer were collected during genetic counseling sessions at the Hereditary Cancer Clinic of the NCIMx. Having completed our

database, we reviewed the contents at the epidemiology department of Mexico's National Institute of Public Health.

4.2 Ethical considerations

We obtained regulatory approval from the Research Ethics Committee of the National Cancer Institute of Mexico (NCIMx) (Approval ID CEI/1123/16). At study enrollment, we provided patients with genetic counseling and information about potential germline findings. Written consent included assurance of patient confidentiality as well as relevant information concerning the project, sample management and DNA shipment for analysis to the National Cancer Institute (NCI), USA. Based on the recommendations of the American College of Medical Genetics, we expressed our commitment to provide patients with the findings on clinically relevant variants in their DNA. Nurses at the Hereditary Cancer Clinic took blood samples from participants only after obtaining their informed consent.

4.3 Sample preparation and whole-exome sequencing

DNA preparation. According to the manufacturer's instructions, genomic DNA was extracted from peripheral leukocytes using the Wizard Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA). DNA was purified using Agencourt AMPure XP Reagent (Beckman Coulter Inc, Brea, CA, USA) based on the manufacturer's protocol. In addition, an adapter-ligated library was prepared using the KAPA HyperPlus Kit (KAPA Biosystems, Wilmington, MA) with Bioo Scientific NEXTflex™ DNA Barcoded Adapters (Bioo Scientific, Austin, TX, USA), in conformity with the KAPA-provided protocol.

Exome capture and sequencing were conducted at the Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, at the NCI, in Rockville, Maryland, USA. The Genome Reference Consortium assembly GRCh37 was used as the genome reference. For the purpose of exome-sequence capture, the SeqCAP EZ Human Exome Library v3.0 (Roche NimbleGen, Madison, WI) was used. Meanwhile the captured DNA was sequenced using the Illumina HiSeq2000 sequencer for 2×100 -bp sequencing of paired-ends (Illumina, San Diego, CA). Variants that failed to pass our pipeline quality control metric (CScorefilter), had a read depth < 10 , ABHet < 0.2 or > 0.8 , or could not pass other quality control filters were excluded from the analysis. Additionally, all variants were filtered using popmaxfreq < 0.01 .

4.4 Bioinformatic analysis

Alignment and variant calling

Subsequently, reads were aligned to the hg19 reference genome using Novoalign software (v3.00.05) [41]. After further analysis, duplicate reads were removed using the MarkDuplicates module of Picard software (v1.126) [42]. High-quality alignments for each individual were further refined using the RealignerTargetCreator and IndelRealigner modules from the Genome Analysis Toolkit [43] (GATK v 4.1.3.0). Bam file level recalibration was also performed using the BaseRecalibrator module from GATK. Variant discovery and genotype calling of multi-allelic substitutions, insertions and deletions were performed on all individuals using the UnifiedGenotyper and HaplotypeCaller modules from the Genome Analysis Toolkit (GATK v 4.1.3.0) [43].

Filtering variants

This analysis incorporated genes associated with hereditary-cancer syndromes including those related to hereditary breast cancer such as *BRCA1/2*, *PALB2*, *CHEK2*, *PTEN* and *TP53*. For the purpose of analysis, we focused on rare exonic variants, and therefore excluded synonymous,

intronic and untranslated regions (UTRs) and variants with an allele frequency >0.1% in the 1000 Genomes Project 2.

Classification of variants

We classified variants as frameshift, stop-gain, splicing, in-frame deletion/insertion or nonsynonymous and missense. Frameshift and stop-gain variants were defined as loss-of-function variants, since they are deleterious and expected to be protein truncating. We examined all potential mutations in the known genes in ClinVar [44] and added appropriate annotations.

4.5 Mutation Validation

After analyzing the variants identified through exome sequencing, we designed specific primers to amplify the region of interest. We then implemented the design utilizing the PrimerQuest Tool [45] and synthesized the DNA sequences using Integrated DNA Technologies. Subsequently, we used Exonuclease 1 to excise the degraded and unincorporated oligonucleotides. Following manufacturer instructions, we produced a dideoxynucleotide incorporation reaction utilizing the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems™). We then removed terminators and unincorporated salts. Terminator sequences were verified through capillary electrophoresis on the 3500 Genetic Analyzer (Applied Biosystems™). Finally, we verified the sequences according to the eBioX program before contrasting them with the reference genome deposited at the base of the NCBI Reference Sequence. We used the MLPA technique to identify the mutated *BRCA1* ex9-12del gene in all the patients included.

4.6 Statistical analysis of clinical and sociodemographic data

We performed a descriptive analysis to calculate measures of central tendency and dispersion for quantitative variables and absolute/relative frequencies for categorical variables. Subsequently, we classified the patients into two strata based on the presence or absence of germline mutations in the selected genes. We evaluated the differential distribution of the variables between the strata using the Fisher's Exact Test for categorical variables and the Wilcoxon Rank Test for quantitative variables. We also constructed logistic regression models to analyze the clinical and sociodemographic variables as predictors of mutation. The final model, which included age of diagnosis as a continuous variable, was adjusted for family history of cancer, immunohistochemical subtype, breastfeeding, use of hormonal contraceptives and BMI. To evaluate disease-free and overall survival we performed a Kaplan-Meier analysis, stratified by mutational status, and considered time (in days) from the date of diagnosis to the date of first recurrence or death. To evaluate differences by mutational status, we performed a logrank test. For all tests, we assumed a p value <0.05 to determine statistical significance. We used STATA 14® software to perform all statistical analysis.

5. Conclusions

This study represents the first time that whole-exome sequencing has been performed among young Mexican women with breast cancer. We analyzed hereditary-cancer genes and obtained detailed clinical and sociodemographic information on this group of patients. In this cohort, we found an 11% prevalence of germline mutation in hereditary-cancer genes. The most prevalent mutated gene was *BRCA2*, representing half of the mutation found and accountable for 6% of breast-cancer cases in young women. We describe for the first time a pathogenic duplication in gene *RAD51C* c.519dupT in a patient with breast cancer. We found that the best predictors for hereditary-cancer mutations in this young population were age at breast-cancer diagnosis and first-degree family history of cancer. Genetic ancestry was not associated with the prevalence of hereditary-cancer mutations. We found no association between disease-free survival and overall survival among the mutation-status subgroups.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Table S1: Sociodemographic features in young women with breast cancer; Table S2: Tumor characteristics of patients with mutations; Table S3: Treatment characteristics for patients with germline mutations; Table S4: Familial and sociodemographic characteristics of patients with germline mutations; Table S5: Reproductive variables in patients with germline mutations; Figure S1: Disease-free survival by mutation adjusted; Figure S2: Disease-free survival by mutation; Figure S3: Overall survival by mutation adjusted; Figure S4: Overall survival by mutation; and Table S6: Variants of uncertain significance.

Author Contributions: Conceptualization, L.G.F.R., A.M.B. and P.O.W.; methodology, M.D., K.J., M.W. and L.G.F.R.; validation, V.F.O., L.G., C.V.G. and A.M.B.; formal analysis, L.G.F.R., N.R.N., L.S.Z. and M.T.M.; investigation, R.M.A.G., T.W.O. and P.G.A.; writing—original draft preparation, L.G.F.R., A.M.B. and L.G.; writing—review and editing, all authors; supervision, L.G.F.R., A.M.B. and P.O.W.; funding acquisition, L.G.F.R., A.M.B. and P.O.W. All authors critically reviewed and approved the final version of the manuscript.

Funding: This study was funded by the Mexican National Council of Science and Technology (Grant number S0008-2016-1/ 272823) and supported by the program of Cátedras-CONACYT and the Intramural Program of the NIH.

Acknowledgments: We would like to express our sincere gratitude to Yulianna Sánchez and Paulina Álvarez, from the Hereditary Cancer Clinic at the National Cancer Institute in Mexico, for providing the facilities to recruit the patients for this study, and María de la Luz Mejía Aguayo and Rodolfo Muñoz-Mendoza, for their help with DNA processing. We also thank Eduardo Lafuente-Flores for his collaboration in preparing the tables and his help in designing Figure 1. Finally, we take this opportunity to thank Verónica Guzmán and Geny Fierros for their support in administering this project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. *CA Cancer J Clin*, 2018. **68**(6): p. 394-424.
2. Assi, H.A., et al., *Epidemiology and prognosis of breast cancer in young women*. *J Thorac Dis*, 2013. **5 Suppl 1**(Suppl 1): p. S2-8.
3. Villarreal-Garza, C., et al., *Breast cancer in young women in Latin America: an unmet, growing burden*. *Oncologist*, 2013. **18**(12): p. 1298-306.
4. Fredriksson, I. and H. Fredholm, [Worse prognosis for young women with breast cancer]. *Lakartidningen*, 2017. **114**.
5. Rich, T.A., et al., *Hereditary breast cancer syndromes and genetic testing*. *J Surg Oncol*, 2015. **111**(1): p. 66-80.
6. Bevers, T.B., et al., *Breast Cancer Screening and Diagnosis, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology*. *J Natl Compr Canc Netw*, 2018. **16**(11): p. 1362-1389.
7. Varsome. Available from: <https://varsome.com/variant/hg19/chr17-56774167-C-CT>.
8. Gómez-Flores-Ramos, L., et al., *Breast cancer genetics in young women: What do we know?* *Mutat Res*, 2017. **774**: p. 33-45.
9. Ebrahimi, E., et al., *The NCCN Criterion "Young Age at Onset" Alone is Not an Indicator of Hereditary Breast Cancer in Iranian Population*. *Cancer Prev Res (Phila)*, 2019. **12**(11): p. 763-770.
10. Ellsworth, D.L., C.E. Turner, and R.E. Ellsworth, *A Review of the Hereditary Component of Triple Negative Breast Cancer: High- and Moderate-Penetrance Breast Cancer Genes, Low-Penetrance Loci, and the Role of Nontraditional Genetic Elements*. *J Oncol*, 2019. **2019**: p. 4382606.
11. Greenup, R., et al., *Prevalence of BRCA mutations among women with triple-negative breast cancer (TNBC) in a genetic counseling cohort*. *Ann Surg Oncol*, 2013. **20**(10): p. 3254-8.
12. Villarreal-Garza, C., et al., *Molecular Subtypes and Prognosis in Young Mexican Women With Breast Cancer*. *Clin Breast Cancer*, 2017. **17**(3): p. e95-e102.
13. Quezada Urban, R., et al., *Comprehensive Analysis of Germline Variants in Mexican Patients with Hereditary Breast and Ovarian Cancer Susceptibility*. *Cancers (Basel)*, 2018. **10**(10).
14. Zayas-Villanueva, O.A., et al., *Analysis of the pathogenic variants of BRCA1 and BRCA2 using next-generation sequencing in women with familial breast cancer: a case-control study*. *BMC Cancer*, 2019. **19**(1): p. 722.
15. Cao, A.Y., et al., *The prevalence of PALB2 germline mutations in BRCA1/BRCA2 negative Chinese women with early onset breast cancer or affected relatives*. *Breast Cancer Res Treat*, 2009. **114**(3): p. 457-62.

16. Cecener, G., et al., *Association of PALB2 sequence variants with the risk of early-onset breast cancer in patients from Turkey*. Mol Biol Rep, 2016. **43**(11): p. 1273-1284.
17. Rashid, M.U., et al., *Prevalence of PALB2 Germline Mutations in Early-onset and Familial Breast/Ovarian Cancer Patients from Pakistan*. Cancer Res Treat, 2019. **51**(3): p. 992-1000.
18. Torres-Mejía, G., et al., *Recurrent BRCA1 and BRCA2 mutations in Mexican women with breast cancer*. Cancer Epidemiol Biomarkers Prev, 2015. **24**(3): p. 498-505.
19. Villarreal-Garza, C., et al., *The prevalence of BRCA1 and BRCA2 mutations among young Mexican women with triple-negative breast cancer*. Breast Cancer Res Treat, 2015. **150**(2): p. 389-94.
20. Millan Catalan, O., et al., *A Multi-Center Study of BRCA1 and BRCA2 Germline Mutations in Mexican-Mestizo Breast Cancer Families Reveals Mutations Unreported in Latin American Population*. Cancers (Basel), 2019. **11**(9).
21. Dean, M., et al., *Addressing health disparities in Hispanic breast cancer: accurate and inexpensive sequencing of BRCA1 and BRCA2*. Gigascience, 2015. **4**: p. 50.
22. Villarreal-Garza, C., et al., *Significant clinical impact of recurrent BRCA1 and BRCA2 mutations in Mexico*. Cancer, 2015. **121**(3): p. 372-8.
23. Sepahi, I., et al., *Investigating the effects of additional truncating variants in DNA-repair genes on breast cancer risk in BRCA1-positive women*. BMC Cancer, 2019. **19**(1): p. 787.
24. Malone, K.E., et al., *Frequency of BRCA1/BRCA2 mutations in a population-based sample of young breast carcinoma cases*. Cancer, 2000. **88**(6): p. 1393-402.
25. Narod, S.A. and W.D. Foulkes, *BRCA1 and BRCA2: 1994 and beyond*. Nat Rev Cancer, 2004. **4**(9): p. 665-76.
26. Soussi, T. and K.G. Wiman, *TP53: an oncogene in disguise*. Cell Death Differ, 2015. **22**(8): p. 1239-49.
27. Sorrell, A.D., et al., *Tumor protein p53 (TP53) testing and Li-Fraumeni syndrome: current status of clinical applications and future directions*. Mol Diagn Ther, 2013. **17**(1): p. 31-47.
28. Masciari, S., et al., *Breast cancer phenotype in women with TP53 germline mutations: a Li-Fraumeni syndrome consortium effort*. Breast Cancer Res Treat, 2012. **133**(3): p. 1125-30.
29. Lee, D.S., et al., *Comparable frequency of BRCA1, BRCA2 and TP53 germline mutations in a multi-ethnic Asian cohort suggests TP53 screening should be offered together with BRCA1/2 screening to early-onset breast cancer patients*. Breast Cancer Res, 2012. **14**(2): p. R66.
30. Mouchawar, J., et al., *Population-based estimate of the contribution of TP53 mutations to subgroups of early-onset breast cancer: Australian Breast Cancer Family Study*. Cancer Res, 2010. **70**(12): p. 4795-800.
31. McCuaig, J.M., et al., *Routine TP53 testing for breast cancer under age 30: ready for prime time?* Fam Cancer, 2012. **11**(4): p. 607-13.
32. Antoniou, A.C., et al., *Breast-cancer risk in families with mutations in PALB2*. N Engl J Med, 2014. **371**(6): p. 497-506.
33. Foulkes, W.D., et al., *Identification of a novel truncating PALB2 mutation and analysis of its contribution to early-onset breast cancer in French-Canadian women*. Breast Cancer Res, 2007. **9**(6): p. R83.
34. Simhadri, S., et al., *PALB2 connects BRCA1 and BRCA2 in the G2/M checkpoint response*. Oncogene, 2019. **38**(10): p. 1585-1596.
35. Piffer, A., E. Luporsi, and C. Mathelin, *[PALB2, a major susceptibility gene for breast cancer]*. Gynecol Obstet Fertil Senol, 2018. **46**(10-11): p. 701-705.
36. Sullivan, M.R. and K.A. Bernstein, *RAD-ical New Insights into RAD51 Regulation*. Genes (Basel), 2018. **9**(12).
37. Sato, K., et al., *Mutation status of RAD51C, PALB2 and BRIP1 in 100 Japanese familial breast cancer cases without BRCA1 and BRCA2 mutations*. Cancer Sci, 2017. **108**(11): p. 2287-2294.
38. Villarreal-Garza, C., et al., *Medical and information needs among young women with breast cancer in Mexico*. Eur J Cancer Care (Engl), 2019. **28**(4): p. e13040.
39. Nelson, H.D., et al., *U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews, in Risk Assessment, Genetic Counseling, and Genetic Testing for BRCA1/2-Related Cancer in Women: A Systematic Review for the U.S. Preventive Services Task Force*. 2019, Agency for Healthcare Research and Quality (US): Rockville (MD).
40. Shao, N., et al., *Prospect for Application of PARP Inhibitor in Patients with HER2 Negative Breast Cancer*. Int J Biol Sci, 2019. **15**(5): p. 962-972.
41. NOVOCRAF. [cited 2019 20 September]; Available from: <http://www.novocraft.com/products/novoalign/>.
42. Picard Toolkit. Broad Institute, G.R. [cited 2019 20 September]; Available from: <http://broadinstitute.github.io/picard/>.

43. Poplin, R., et al., *Scaling accurate genetic variant discovery to tens of thousands of samples*. bioRxiv, 2018: p. 201178.
44. Landrum, M.J., et al., *ClinVar: improving access to variant interpretations and supporting evidence*. Nucleic Acids Res, 2018. **46**(D1): p. D1062-d1067.
45. IDT, I.D.t. [cited 2019 20 September]; Available from: <https://www.idtdna.com/pages/products/custom-dna-rna/dna-oligos>.