

## Article

# Germline Variants in Cancer Genes from Young Breast Cancer Mexican Patients

Liliana Gómez-Flores-Ramos <sup>1,†</sup> , Angélica Leticia Barraza-Arellano <sup>2,†</sup>, Alejandro Mohar <sup>3</sup>, Miguel Trujillo-Martínez <sup>4</sup> , Lizbeth Grimaldo <sup>1</sup>, Rocío Ortiz-Lopez <sup>2,5</sup>  and Víctor Treviño <sup>2,5,\*</sup> 

<sup>1</sup> CONACYT/Center for Population Health Research, National Institute of Public Health, Universidad No. 655, Cuernavaca 62100, Morelos, Mexico; liliana.gomez@insp.mx (L.G.-F.-R.); cisp73@insp.mx (L.G.)

<sup>2</sup> School of Medicine, Tecnológico de Monterrey, Morones Prieto Av 3000, Los Doctores, Monterrey 64710, Nuevo Leon, Mexico; barraza.leticia@exatec.tec.mx (A.L.B.-A.); rortizl@tec.mx (R.O.-L.)

<sup>3</sup> Unidad de Investigación Biomédica en Cáncer, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, Dirección de Investigación, Instituto Nacional de Cancerología, Av. San Fernando #22, Col. Sección XVI, Delegación Tlalpan, Mexico City 14080, Mexico; mohar@iibiomedicas.unam.mx

<sup>4</sup> Instituto Mexicano del Seguro Social, Hospital General de Zona con Medicina Familiar No. 7, Cuautla 62780, Morelos, Mexico; miguel.trujillom@imss.gob.mx

<sup>5</sup> The Institute for Obesity Research, Tecnológico de Monterrey, Eugenio Garza Sada Av 2501, Monterrey 64849, Nuevo Leon, Mexico

\* Correspondence: vtrevino@tec.mx

† These authors contributed equally to this paper.

**Simple Summary:** Young, Mexican women are more susceptible to breast cancer compared to other populations. However, studies on germline young, Mexican women are scarce and limited to a panel of 143 genes. This is partially due to the lack of gene annotation tools and difficulties in determining the causal genes in understudied populations. Here, we used whole exome sequencing combined with a powerful annotation tool to analyze 862 cancer genes in 115 young, Mexican women diagnosed with breast cancer. Our results showed well-known genes and many barely reported variants in our population. Moreover, we were able to assign candidate causal genes to 34% of patients, surpassing previous studies. These results suggest that deeper bioinformatic analyses could inform medical decisions to improve diagnosis, treatment, and life expectancy in young women with breast cancer.

**Abstract:** Breast cancer (BC) is one of the most frequent cancer types in women worldwide. About 7% is diagnosed in young women (YBC) less than 40 years old. In Mexico, however, YBC reaches 15% suggesting a higher genetic susceptibility. There have been some reports of germline variants in YBC across the world. However, there is only one report from a Mexican population, which is not restricted by age and limited to a panel of 143 genes resulting in 15% of patients carrying putatively pathogenic variants. Nevertheless, expanding the analysis to whole exome involves using more complex tools to determine which genes and variants could be pathogenic. We used germline whole exome sequencing combined with the PeCanPie tool to analyze exome variants in 115 YBC patients. Our results showed that we were able to identify 49 high likely pathogenic variants involving 40 genes on 34% of patients. We noted many genes already reported in BC and YBC worldwide, such as *BRCA1*, *BRCA2*, *ATM*, *CHEK2*, *PALB2*, and *POLQ*, but also others not commonly reported in YBC in Latin America, such as *CLTCL1*, *DDX3X*, *ERCC6*, *FANCE*, and *NFKBIE*. We show further supporting and controversial evidence for some of these genes. We conclude that exome sequencing combined with robust annotation tools and further analysis, can identify more genes and more patients affected by germline mutations in cancer.

**Keywords:** breast cancer; variant annotation; variant pathogenicity; cancer predisposition; bioinformatic pipeline; cancer genes



**Citation:** Gómez-Flores-Ramos, L.; Barraza-Arellano, A.L.; Mohar, A.; Trujillo-Martínez, M.; Grimaldo, L.; Ortiz-Lopez, R.; Treviño, V. Germline Variants in Cancer Genes from Young Breast Cancer Mexican Patients.

*Cancers* **2022**, *14*, 1647. <https://doi.org/10.3390/cancers14071647>

Academic Editor: José Antonio López-Guerrero

Received: 30 January 2022

Accepted: 18 March 2022

Published: 24 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Breast cancer (BC) is one of the most frequent types of cancer among women; 2.3 million cases were reported in 2020, causing more than 680,000 deaths globally [1]. The BC incidence increases with age. The majority of the diagnosed patients are 40 years and older, reaching 89%: 4% between 39 and 35, 5% less than 35, and 1.75% in women under 30 (<https://gco.iarc.fr/> (accessed on 1 December 2021), [1]). In Mexico, BC is the first cause of death among women diagnosed with a malignant neoplasia [1], with an estimate of 26% of cancer-related deaths. Notably, the incidence of breast cancer in young women (YBC) in Mexico reached 15% in 2020, which is considerably higher than worldwide [1,2].

It has been observed that clinical outcomes and tumor biology in young patients are different from older women [3,4]. Tumors in young women are more likely to be of higher histological grade and are usually classified as estrogen receptor (ER) and progesterone negative receptors [5]. These patients typically present local recurrences to be diagnosed at an advanced stage and have inferior 5-year survival compared to the older premenopausal counterparts [5]. Moreover, YBC tends to present a more aggressive subtype such as triple-negative or HER2+ [6].

In some populations, up to 24% of the hereditary BC is linked to germline mutations in *BRCA1/2* specific genes [7], and the prevalence may reach 14% in patients not showing familial history [7]. Mutations in these genes have a lifetime risk of developing BC up to 65% [8]. Other genes with germline mutations had been reported to confer moderate cancer risk, such as *ATM*, *CHEK2*, and *PALB2*, which are also associated with other cancers [8–10].

In addition to *BRCA1*, *BRCA2*, *ATM*, *CHEK2*, and *PALB2*, germline mutations in Latin America have been reported for *CDH1*, *NBN*, *NF1*, *TP53*, *MLH1*, *BRIP1*, *MSH2*, *MSH6*, and *PMS2* in populations from Chile, Brazil, Colombia, and México [11]. In Mexico, a study in hereditary breast and ovarian cancer, which used a panel of 143 genes, found 21 germline mutations in *BRCA1* and *BRCA2*, while another 19 genes showed 1 or 2 mutations, including the *FANC(I/B/C/L/M)* gene family accounting for 6 mutations [12]. Importantly, the above study found only 15% of patients (46 of 300) showing a germline mutation in at least one of the 143 panel genes. The low detection rate highlights the need to interrogate more genes in the population known to carry familial susceptibility.

The sequencing of gene panels has been a cost-efficient tool to determine the prevalence of specific genes among cancer populations [13–16]. However, more genes need to be studied to determine pathogenicity in most patients, for example, by whole exome sequencing. One issue is that pathogenicity is challenging to assess in not well-known or reported genes without further functional assays. Fortunately, extensive sequencing efforts, such as the 1000 genomes project [17] and the gnomAD (accessed on 1 December 2021) [18], provide databases of human variation among a relatively healthy population that help remove many common variants. In addition, other databases, such as ClinVar (accessed on 1 December 2021), help to mark pathogenic variants [19]. Still, massive sequencing generates several unseen variants increasing the need to explore methodologies to identify possible causal variants. In this context, Pediatric Cancer Variant Pathogenicity Information Exchange (PeCanPIE, accessed on 1 December 2021) is a web-based tool that integrates many sources of information supported by the guidelines from the American College of Medical Genetics and Genomics (ACMG) that are useful to identify interesting candidate variants quickly [20]. PeCanPIE uses a variant categorization for putative pathogenicity based on a “medal ceremony” concept of four levels. The *Gold* category is assigned for highly likely pathogenic truncating or splicing variants, whose genes are already found in pathogenic databases and whose variants are rare among healthy populations. *Silver* variants are in-frame, indels, and truncations in non-cancer genes or predicted to be damaging and matching pathogenic databases. *Bronze* variants are those whose effects are predicted to be tolerated. An *unknown* label is assigned otherwise.

To determine the possible etiology of BC in young women, in this study, we performed whole exome sequencing followed by the PeCanPie bioinformatic analysis of 115 young patients diagnosed with BC and focused on known 862 cancer genes where pathology

has already been associated with diseases. Briefly, we obtained 49 variants classified as *Gold* resulting in more than 20 genes that have not been reported previously in LA BC. In addition, 106 variants were classified as *Silver*.

## 2. Materials and Methods

### 2.1. Patients

We included 115 patients from the National Cancer Institute of Mexico (INCAN, Instituto Nacional de Cancerología) diagnosed at 40 years old or younger with histopathological confirmation of BC. Patients were recruited between September 2015 and December 2017. Medical records and electronic files with detailed clinical and sociodemographic information were obtained from the INCAN. Trained nurses obtained blood samples after patients received genetic counseling and signed the informed consent.

### 2.2. Ethical Considerations

Regulatory approval was obtained from the INCAN Research and Ethics Committee (Approval ID CEI/1123/16). Genetic counseling and information about potential germline findings were provided to patients in addition to assurance of patient confidentiality and relevant information concerning the project, sample management, and DNA shipment for analysis to the National Cancer Institute (NCI), USA. Based on the recommendations provided by the American College of Medical Genetics, hereditary cancer diagnostic variants in their DNA were reported to the patients.

### 2.3. Sample and Panel Library Preparation for Sequencing

The DNA was extracted from peripheral leukocytes using the Wizard Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA), following the manufacturer's instructions. The resulting DNA was purified using Agencourt AMPure XP reagent (Beckman Coulter Inc., Brea, CA, USA) following the manufacturer's protocol. In addition, an adapter-ligated library was prepared using the KAPA HyperPlus Kit (KAPA Biosystems, Wilmington, MA, USA) with NEXTflex™ DNA Barcoded Adapters (Perkin Elmer Waltham, MA, USA), according to the KAPA-provided protocol.

### 2.4. Sequencing

The deep whole exome sequencing was performed at the Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics (NCI, Rockville, MD, USA). The GRCh37 (hg19) genome assembly was used for genome mapping reference. The DNA sequencing was performed using the Illumina HiSeq 2000 sequencer for 2 × 100-pb paired-end cartridge (Illumina, San Diego, CA, USA). The sequencing included regulatory, splicing, and 3' and 5' UTR regions.

### 2.5. Variant Calling

The reads were aligned using Novoalign software (v3.00.05, Novocraft Technologies Sdn Bhd, Petaling Jaya Selangor, Malaysia). Duplicate reads were removed using MarkDuplicates from Picard Software (v1.126, Broad Institute, Cambridge, MA, USA). For variant calling, the pipeline involved RealignerTargetCreator, IndelRealigner, BaseRecalibrator, UnifiedGenotyper, and HaplotypeCaller tools from GATK (v4.1.3.0, Broad Institute, Cambridge, MA, USA). Variants that failed to pass the pipeline control metrics (CScorefilter) had a read depth minor to 10, ABHet (reference to alternate reads ratio) <0.2 or >0.8 were excluded for the analysis. In addition, all variants were filtered using popmaxfreq <0.01.

Before variant categorization, variants were filtered to remove those that do not show at least 10 reads in the alternate allele or that the minor allelic fraction was lower than 0.25.

### 2.6. Variant Categorization

The obtained variant calling files were analyzed using PeCanPIE, which classified the variants into three categories: *Gold*, *Silver*, and *Bronze* [20]. The variants were assigned as

*Gold* if it is a truncating variant, including splicing, or a loss-of-function associated with a reported disease documented in ClinVar with at least two starts, and at least one of the following databases: IARC: Tp53, ASU: TERT, ARUP: Ret, BIC, PCFP, and COSMIC. The variants assigned as *Silver* are considered as in-frame indels, truncation events in non-tumor suppressor genes. Variants had to be reported in at least one of the following databases: UMD, LOVD, RB1, or ALSoD [20]. The variant is assigned as *Bronze* if predicted to be tolerant by in silico algorithms.

For PeCanPie, we used the union of their internal set of cancer predisposition genes and the Cancer Gene Census v94 from COSMIC [21] containing 712 and 723 genes, respectively, generating a list of 862 genes.

Only gold and silver variants were considered and studied separately for further analysis and filtering. The variants with fewer than 20 total reads or variant allele frequency (VAF) less than 0.25 were discarded for further analysis.

### 2.7. Statistical Analysis

We performed a descriptive analysis to calculate central tendency and dispersion measures for quantitative variables and absolute/relative frequencies for categorical variables. We also constructed logistic regression models to analyze the predictors of *Gold* variants. The final model included age < 30 years at diagnosis, family history of cancer, immunohistochemical subtype, advanced stage (IIB–IV), and continuous BMI. To evaluate disease-free and overall survival, we performed a Kaplan–Meier analysis, stratified by *Gold* variant, and considered time (in months) from diagnosis to the date of first recurrence or death. To evaluate differences by mutational status, we performed a Log-rank test. For all tests, we assumed a *p*-value < 0.05 to determine statistical significance. We used STATA 14<sup>®</sup> software to perform all statistical analysis tests.

## 3. Results

### 3.1. Clinical Population

A summary of our cohort characteristics, including 115 patients, is shown in Table 1 (raw data in Table S1). The median age was 33.9 years at diagnosis, and 18.3% were younger than 30. Patients diagnosed at advanced clinical stages (IIB–IV) corresponded to 65% of the cohort. Ductal histology (85%) and Luminal-B (50%) subtype were the most common in this group. A family history of cancer was present in 32.2% of these young women. All participants were premenopausal. The mean age at menarche was 12.34 years, 28.7% of participants were nulliparous, and the average number of children among parous was 2.1. Breastfeeding was practiced in 68.7% of parous women. Overweight and obesity was a condition present in 61.7% of the participants.

### 3.2. Variant Categorization

From the 1,189,705 raw variants, quality filtering generated 350,546 variants used for PeCanPIE annotation. PeCanPIE detected 6496 variants within the 862 cancer genes selected. After further filtering (for the number of reads described in the Methods section), 49 were categorized as *Gold* and 106 as *Silver*, corresponding to 40 and 87 genes, respectively (Figure 1). Only 39 patients (34%) showed high confidence *Gold* variants, while 74 (64%) showed *Silver* variants. Overall, 88 patients showed *Gold* or *Silver* variants.

### 3.3. High Confident Germline Variants in Cancer Genes

The criteria for *Gold* variants involve known pathogenic variants, a strong alteration variant (nonsense, frameshift, truncation) in a known pathogenic gene, and low allele frequency in public non-cancer databases [20]. From the 52 *Gold* variants obtained before manual filtering, two were removed (chr19, positions 34945343 and 34945354, *UBA2* gene) because six patients showed alternative and varied genotypes at those positions, always with fewer reads in the alternate genotype suggesting mapping and sequencing artifacts. Additionally, one variant in (chr16, position 72991715, *ZFH3* gene) was also removed

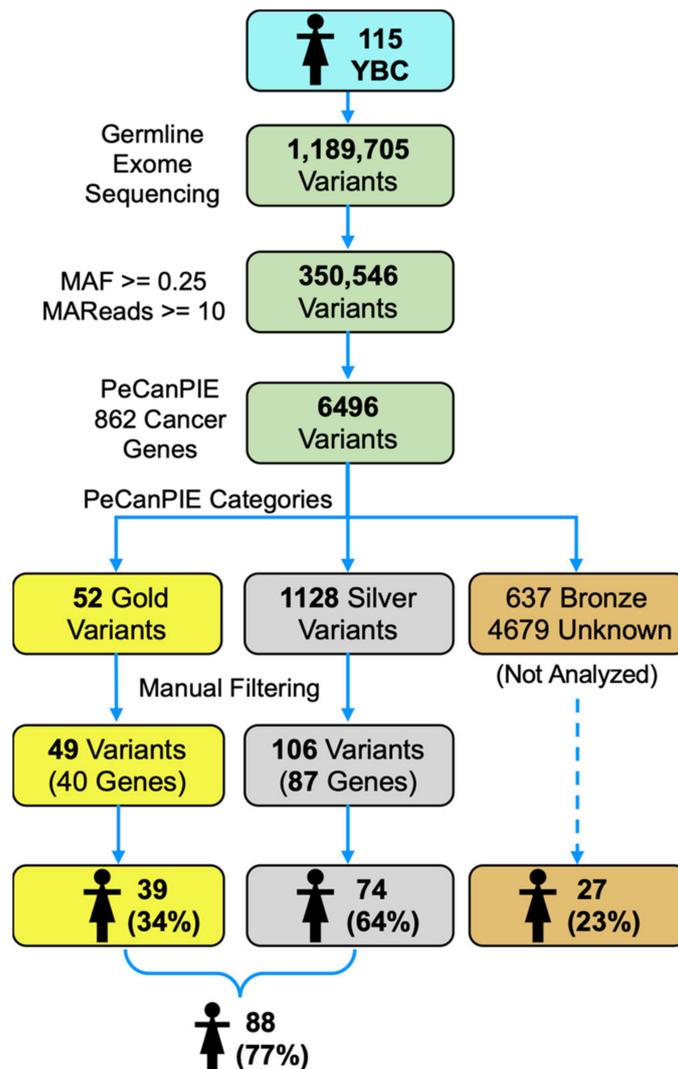
because it was present in all patients suggesting a common variant. Thus, 49 variants involving 40 genes and 39 patients were finally designated as *Gold* (Figure 2 and Table 2). All variants were heterozygous. Except for one variant present in two patients (*RBM8A*), all *Gold* variants were observed only in one patient. Of these variants, we noted 20 splicing, 17 frameshift, eight nonsense, three missense, and one '5-UTR. Most patients showed only one *Gold* variant, but 11 patients of 39 (28%) showed two. The well-known *BRCA2* gene was the most recurrently altered, observed in five patients, showing four frameshifts and one missense. The following most frequent alterations were observed in *CHEK2*, *PALB2*, *POLQ*, *DDX3X*, and *FLG* affecting two patients. From these genes observed in two or more patients, the variants observed in *BRCA2*, *PALB2*, and *CHEK2* were also found in ClinVar with documented association to BC (Table S2). However, variants in *POLQ*, *DDX3X*, *FLG*, and *RBM8A* are barely reported in BC and will be further described.

**Table 1.** Clinical data of 115 BC young patients.

Clinical Data		<i>n</i>	Frequency (%)
Patients	<i>n</i>	115	
Age	Mean	33.9	
	Interquartile range	31–38	
Age at menarche	Mean	12.34	
	Interquartile range	11–13	
Parity	Nulliparous	33	28.7
	1 child	27	23.2
	2 children	32	23.5
	>3 children	23	20.0
Breastfeeding	No	37	32.2
	Yes	78	67.8
	Luminal B (Her2-positive)	19	13.2
	Luminal B (Her2-negative)	31	28.3
	Her2-positive (non-luminal)	8	6.6
	Triple Negative	30	28.3
Histology	Ductal	98	85.2
	Lobular	6	5.2
	Mixed	9	7.8
	Others	2	1.7
Clinical Stage	I	15	13.0
	II	45	39.1
	III	49	42.6
	IV	6	5.2
Consumption of Hormonal contraceptives	No	50	43.5
	Yes	65	56.55
First-grade family history of cancer	No	90	78.3
	Yes	25	21.7
Second-grade family history of cancer	No	78	67.8
	Yes	37	32.2
Body Mass Index	Normal (BMI < 25)	44	38.3
	Overweight	45	39.1
	Obesity (BMI > 30)	26	22.6

*DDX3X*, located in Xp11.4, encodes for an RNA helicase linked to somatic mutations in medulloblastoma [22]. It is also X chromosome inactivated in ovarian cancer [23]. Germline mutations have been reported in female brain development and disability, whose variants were observed on the *Helicase ATP-binding* and the *Helicase C-terminal* domains [24,25]. The

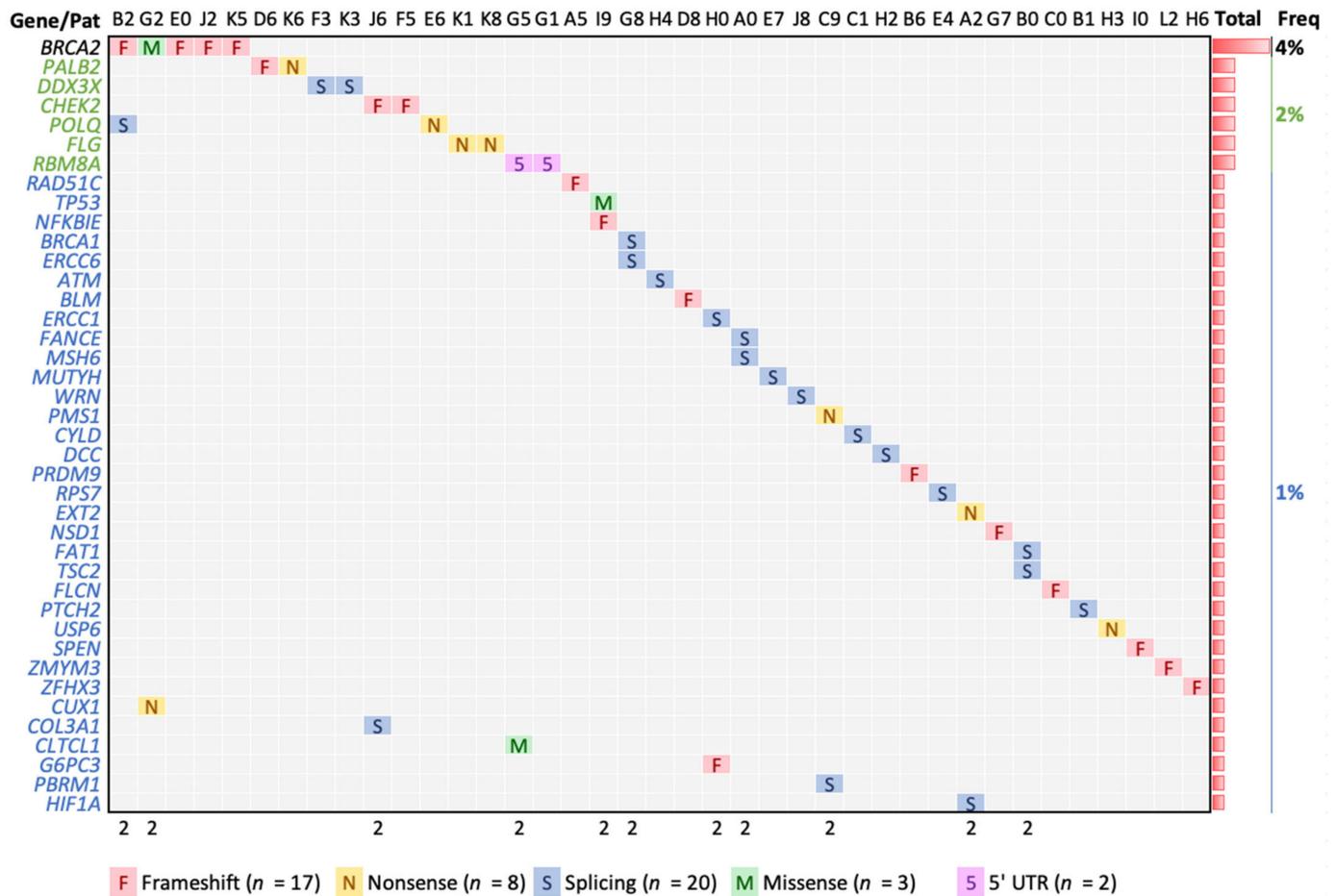
two observed T→C heterozygous variants in our cohort affect the exon-intron splicing sites located in the *Helicase C-terminal* region (G539 and S590) responsible for the interaction with the nuclear mRNA export receptor TAP [26]. Moreover, DDX3X plays a role in DNA damage response [27]. There is no evidence in the clinical record for mental disability in these patients. These pieces of evidence suggest that DDX3X is potentially a predisposing gene in young BC patients.



**Figure 1.** The diagram summarizes all the obtained variants with their exclusion criteria. YBC young Mexican breast cancer patients.

*POLQ* encodes a polymerase involved in DNA repair [28,29]. Germline mutations in BC have been reported mainly in non BRCA1/2 carriers [30,31]. In our data, we observed one splicing and one nonsense variant in I2385V and L1430\*, respectively. Consistent to the above studies, these *POLQ* positive patients were not carriers of BRCA1/2 or TP53 variants.

*FLG* encodes for filaggrin that aggregates keratin intermediate filaments in the mammalian epidermis. *FLG* germline variants have been recently reported in around 16% of Taiwanese BC patients [28] and 17% of hepatocellular carcinomas from Thailand [29]. In addition, it has been found somatically mutated in 10% of ER + BC patients [30]. We observed two G→A nonsense mutations (R788\* and R501\*), each one affecting a single patient. One of these patients carrying a *FLG* germline variant had a second primary contralateral breast cancer (K1).



**Figure 2.** Gold variants detected. The top seven genes show two or more patients affected (black and green), while the rest show one patient only (blue). The 11 patients (in columns) marked with “2” show two Gold variants.

We observed a promoter variant in *RBM8A* affecting two patients. According to ClinVar (ID 30464), this recessive variant causes a decrease in gene expression [31], which is crucial when combined with another severe variant affecting patients with Radial aplasia-thrombocytopenia syndrome. *RBM8A* differential expression has been noted in several cancers types but observed more expressed in the tumor than in the normal tissue [32], which seems counterintuitive. The allelic fraction reported in gnomAD is 1.8%, similar to the 1.7% observed in our sample, suggesting that promoter variants may be random. Like reports in ClinVar, it may need to be combined with another unknown alteration. Overall, the evidence is unclear, suggesting that the promoter variant in *RBM8A* is a variant of unknown significance (VUS).

From the genes present in only one patient, 12 genes have strong support shown by their reports in ClinVar for pathogenic or likely pathogenic variants in *ATM*, *BLM*, *BRCA1*, *CLTCL1*, *ERCC6*, *FANCE*, *G6PC3*, *MSH6*, *MUTYH*, *TP53*, and *TSC2*. Moreover, some of these genes also show germline mutations in BC patients. For example, *ERCC6* has also been reported in Brazilian YBC patients [33] and Lebanon familial BC [34]; *BLM* in Russian YBC patients [35] and USA patients [36]; *TSC2* in Italian patients [37]; and *ATM* is also well known in BC [11].

Table 2. Gold variants.

Chr	Position	Ref *	Alt *	Depths	Gene	LA <sup>+</sup>	Type	AA Chg	Pat	pLI	AF Lat <sup>§</sup>	ProtSize
11	108175579	G	A	21,19	<i>ATM</i>		Splice	E1892_E37	H4	0	$1 \times 10^{-4}$	3056
15	91341429	A	AG	15,25	<i>BLM</i>		Frameshift	T1074fs	D8	0	-	1417
17	41258472	C	A	54,35	<i>BRCA1</i>		Splice	R71_E4	G8	0	$3 \times 10^{-5}$	1863
13	32914033	CA	C	18,20	<b><i>BRCA2</i></b>		Frameshift	S1848fs	E0	0	$3 \times 10^{-5}$	3418
13	32914122	AC	A	52,30	<b><i>BRCA2</i></b>		Frameshift	N1877fs	J2	0	$3 \times 10^{-5}$	3418
13	32930653	C	CA	15,17	<b><i>BRCA2</i></b>		Frameshift	S2509fs	B2	0	-	3418
13	32937507	A	G	44,30	<b><i>BRCA2</i></b>		Missense	D2723G	G2	0	A:C 0	3418
13	32954260	CG	C	23,17	<b><i>BRCA2</i></b>		Frameshift	V3079fs	K5	0	$2 \times 10^{-4}$	3418
22	29091718	TA	T	29,23	<b><i>CHEK2</i></b>		Frameshift	L413fs	J6	0	-	543
22	29115401	A	ATGAT	28,16	<b><i>CHEK2</i></b>		Frameshift	M222fs	F5	0	$1 \times 10^{-4}$	543
22	19222211	C	T	33,36	<i>CLTCL1</i>	No	Missense	E330K	G5	0	$9 \times 10^{-4}$	1640
2	189868190	T	A	35,26	<i>COL3A1</i>	No	Splice	P869P	J6	1	$9 \times 10^{-5}$	1466
7	101926377	C	T	57,41	<i>CUX1</i>	No	Nonsense	Q678 *	G2	1	$3 \times 10^{-3}$	678
16	50818362	T	A	30,39	<i>CYLD</i>	No	Splice	I650_E13	C1	1	$3 \times 10^{-5}$	956
18	50734187	G	A	28,23	<i>DCC</i>	No	Splice	V621_E11	H2	0.99	$2 \times 10^{-4}$	1447
X	41206109	T	C	46,39	<b><i>DDX3X</i></b>	No	Splice	G539_E15	K3	1	$4 \times 10^{-5}$	662
X	41206562	T	C	44,41	<b><i>DDX3X</i></b>	No	Splice	S590_E16	F3	1	$2 \times 10^{-3}$	662
19	45917294	T	C	21,22	<i>ERCC1</i>		Splice	V235_E7	H0	0	$6 \times 10^{-5}$	297
10	50680422	C	T	16,25	<i>ERCC6</i>	No	Splice	R975_E16	G8	0	C:A $3 \times 10^{-5}$	1493
11	44228353	G	A	33,34	<i>EXT2</i>	No	Nonsense	W535 *	A2	0	$1 \times 10^{-4}$	718
6	35425330	C	T	25,34	<i>FANCE</i>	No	Splice	D286_E3	A0	0	$3 \times 10^{-4}$	536
4	187530955	C	G	34,21	<i>FAT1</i>	No	Splice	T3356T	B0	0	0	4588
17	17117000	CG	C	45,40	<i>FLCN</i>	No	Frameshift	R570fs	C0	0.79	C:T $3 \times 10^{-5}$	579
1	152285000	G	A	46,43	<b><i>FLG</i></b>	No	Nonsense	R788 *	K8	0	$6 \times 10^{-4}$	4061
1	152285861	G	A	45,34	<b><i>FLG</i></b>	No	Nonsense	R501 *	K1	0	$4 \times 10^{-3}$	4061
17	42148542	TC	T	11,14	<i>G6PC3</i>	No	Frameshift	I70fs	H0	0	$8 \times 10^{-4}$	346
14	62203827	G	A	20,20	<i>HIF1A</i>	No	Splice	D417_E9	A2	0	$2 \times 10^{-4}$	826
2	48033791	GT(26)	G	15,12	<i>MSH6</i>		Splice	R1334_E9	A0	0	$1 \times 10^{-4}$ &	1360
1	45797228	C	T	25,23	<i>MUTYH</i>		Splice	G396_E13	E7	0	$3 \times 10^{-3}$	546
6	44233331	G	GC	21,12	<i>NFKBIE</i>	No	Frameshift	A57fs	I9	0.77	$3 \times 10^{-4}$	500
5	176722446	TC(6)	T	27,25	<i>NSD1</i>	No	Frameshift	S2424fs	G7	1	-	2696
16	23641062	CAG	C	25,39	<b><i>PALB2</i></b>		Frameshift	S804fs	D6	0	$1 \times 10^{-4}$	1186
16	23641139	G	C	24,36	<b><i>PALB2</i></b>		Nonsense	S779 *	K6	0	$9 \times 10^{-5}$	1186
3	52620706	TG	T	15,12	<i>PBRM1</i>	No	Splice	E1017_E21	C9	1	0 &	1689
2	190728500	C	T	26,29	<i>PMS1</i>	No	Nonsense	R630 *	C9	0	$1 \times 10^{-4}$	932
3	121168273	T	C	19,22	<b><i>POLQ</i></b>		Splice	I2385V	B2	0	-	2590
3	121207489	A	T	20,15	<b><i>POLQ</i></b>		Nonsense	L1430 *	E6	0	$9 \times 10^{-5}$	2590
5	23527845	CA	C	31,42	<i>PRDM9</i>	No	Frameshift	T883fs	B6	0	$6 \times 10^{-5}$	894
1	45294985	C	T	14,13	<i>PTCH2</i>	No	Splice	L406_E10	B1	0	$3 \times 10^{-5}$	1203
17	56774167	C	CT	47,59	<i>RAD51C</i>		Frameshift	A173fs	A5	0	-	376
1	145507646	G	A	15,27	<b><i>RBM8A</i></b>	No	UTR_5	E1_UTR_5	G1	0.57	$1 \times 10^{-2}$	174
1	145507646	G	A	18,24	<b><i>RBM8A</i></b>	No	UTR_5	E1_UTR_5	G5	0.57	$1 \times 10^{-2}$	174
2	3623181	G	A	51,58	<i>RPS7</i>	No	Splice		E4	0.95	-	194
1	16262459	G	GC(27)	27,17	<i>SPEN</i>	No	Frameshift	A3242fs	I0	1	$9 \times 10^{-4}$ &	3664
17	7578406	C	T	22,29	<i>TP53</i>		Missense	R175H	I9	0.53	0	393
16	2124201	C	T	44,31	<i>TSC2</i>	No	Splice	R786C	B0	1	0	1807
17	5074956	T	A	83,62	<i>USP6</i>	No	Nonsense	Y1343 *	H3	0	$9 \times 10^{-5}$	1406
8	31014882	A	G	13,13	<i>WRN</i>		Splice	K1274_E33	J8	0	$1 \times 10^{-4}$	1432
16	72991713	C	CC(9)	20,14	<i>ZFXH3</i>	No	Frameshift	A778fs	H6	1	0	3703
X	70466308	GTGGT	G	28,11	<i>ZMYM3</i>	No	Frameshift	P821fs	L2	1	-	1370

\* Numbers in parenthesis represent the total length. <sup>+</sup> Represent whether the gene has been reported in Latin-American BC patients in the Urbina-Lara et al. analysis [11]. <sup>§</sup> Allele frequency in Latino population from GnomAD website (<https://gnomad.broadinstitute.org/>, accessed on 1 December 2021). Variants in GnomAD slightly different to those found are explicitly indicated or marked with &. A total of 50 variants is shown, 49 unique (RBM8A is present in two patients). Genomic positions correspond to hg19. Ref and Alt refer to reference and alternate alleles respectively. Depths refers to Ref and Alt alleles respectively. AA Chg refers to aminoacid change. Pat refers to patient. pLI refers to the probability of LoF intolerance. ProtSize refers to canonical transcript protein size in aminoacids. Bold genes mark those found more than once. AF Lat = Allele Frequency in Latin Americans.

The remaining *Gold* variants genes are, by definition in PeCanPIE, associated with diseases in databases; inquiringly, they do not show clear evidence of pathogenicity, specifically in ClinVar. All these variants carry protein truncations in genes where the loss of

function (LoF) mutations is a known mechanism of disease (*BLM*, *COL3A1*, *CUX1*, *CYLD*, *ERCC1*, *EXT2*, *FAT1*, *FLCN*, *NFKBIE*, *NSD1*, *PBRM1*, *PMS1*, *PRDM9*, *PTCH2*, *RAD51C*, *RPS7*, *USP6*, *WRN*, and *ZFH3*). Many of these genes already show evidence of variants in BC in other populations such as *BLM* in Slavic [38] and Brazilians [39]; *ERCC1* in Brazilians [40]; *PMS1* in men [41]; *WRN* in Latins [42], or in other cancers such as *CYLD* in head and neck [43]; *EXT2* in osteochondromas and chondrosarcomas [44,45]; *FAT1* in retinoblastomas [46]; *NSD1* and *PBRM1* in renal cell carcinoma [47,48]; *PRDM9* in acute lymphoblastic leukemia [49]; *PTCH2* in rhabdomyosarcoma [50]; *RAD51C* also in breast cancer [51]; *RPS7* in hypocellular bone marrow failure [52]; *SPEN*, *USP6*, and *ZFH3* in pancreatic adenocarcinoma [53]; *TICAM1* in thyroid cancer [54]; *ZFH3* also in endometrial cancer [55]; and *DCC* in gallbladder cancer [56].

We also noted seven genes marked with “caution” in PeCanPIE showing the truncation close to the C-terminal (*CUX1*, *FLCN*, *FLG*, *MSH6*, *NSD1*, *PRDM9*, and *USP6*), questioning its functional effects in the cancer context. To provide additional support for these variants, we considered the probability of LoF intolerance (pLI) provided by gnomAD [18]. Natural selection purifies highly deleterious variants, therefore, genes showing fewer than expected LoF variants in a large healthy population are seen as highly LoF intolerant, proposing association to disease when observed in an individual. Thus, pLI close to 0.9 and up to 1 are a strong indicator of LoF intolerance as recommended by gnomAD. We noted pLI = 1 in *NSD1*, strongly suggesting some contribution to disease consistent with previous evidence [47]. We also noted pLI = 1 for *CUX1*, but the stop codon gain is shown at the last protein amino acid (Q678), marked in gnomAD as ‘low confidence loss of function’. In between, we observed pLI = 0.79 in *FLCN* and pLI = 0.77 for *NFKBIE*. Contrary, we noted pLI = 0 for *PRDM9* and *USP6*; thus, although categorized as *Gold*, these variants are less likely to be pathogenic. We also noted pLI > 0.9 for *COL3A1*, *DCC*, *DDX3X*, *PBRM1*, *RPS7*, *SPEN*, *TSC2*, *ZFH3*, and *ZMYM3*, in which a considerable proportion of the protein is altered by a frameshift, splice, or nonsense variant. We noted that one of the patients carrying TP53 and NFKBIE presented a second primary glioblastoma (I9).

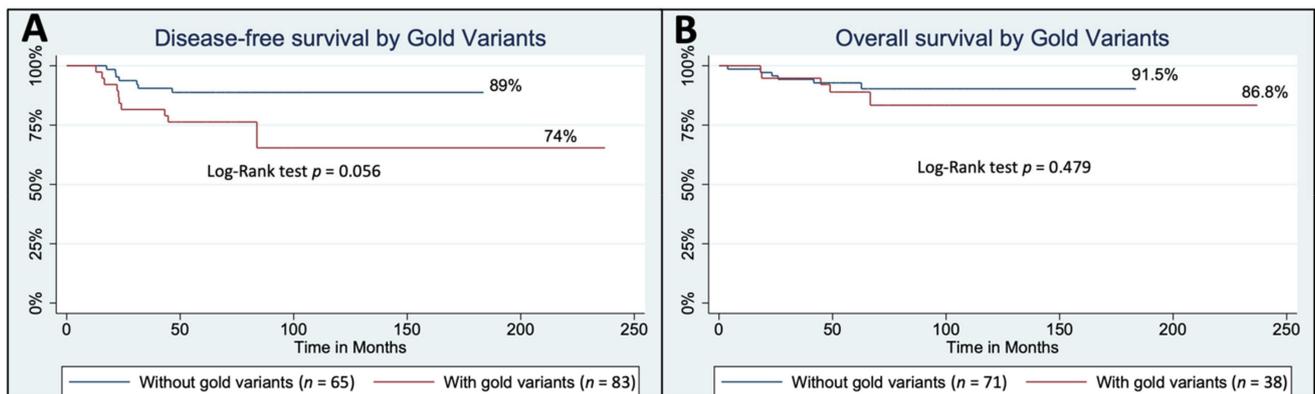
We observed that the variant allelic fraction in Latino populations, is in general, low (Table 2), validating the PeCanPie filtering. Nevertheless, few variants in the order or few per a thousand ( $10^{-3}$ ) could indicate a random finding due to our sample size.

We explored associations between patients carrying *Gold* variants and those not, along with clinical co-variables. The adjusted model showed an association of *Gold* variants with first- and second-family history of cancer (OR 3.21; CI 95% 1.15–8.95) and age < 30 years (OR 3.74; CI 95% 1.20–11.70). None of the tumor subtypes was associated with carrying a *Golden* variant.

We also observed that an increase in one unit of continuous BMI raises the odds of detecting a *Gold* variant in young women with breast cancer (OR 1.19; CI 95% 1.07–1.34), suggesting that BMI might be a modifier in women with *Gold* germline variants, and might reduce the age of breast cancer presentation. This phenomenon has been described previously for BRCA1/2 carriers, but our results suggest BMI could be a modifier for other genes as well [57–59].

We noticed that women with a *Gold* variant were diagnosed in advanced stages: IIB-IV (OR 3.21; CI 95% 1.21–8.98), suggesting that a *Gold* variant might increase breast cancer aggressiveness. It has been described that BRCA1/2 carriers have a higher risk of lymph node involvement at diagnosis [60].

The Median follow-up of this cohort was 62 months (48–73). Although not significant at  $p < 0.05$  using a Log-Rank test probably due to the small sample size, we found that disease-free survival could be higher in women without *Gold* variants (89%) versus women with *Gold* variants (74%) (Figure 3A) after adjusting for clinical prognostic factors. We observed fewer differences for overall survival (OS), with 91.5% of OS in women without carrying a *Gold* variant and 86.8% OS in women with a *Gold* variant (Log-Rank test  $p = 0.48$ , Figure 3B). These results are consistent with many previous studies involving hereditary breast cancer and survival [61,62].



**Figure 3.** Survival analyses of YBC patients carrying *Gold* variants. (A) Disease free survival. (B) Overall survival.

### 3.4. Modest Confidence Germline Variants

The criteria for *Silver* are in-frame indels, truncation events in non-tumor suppressor genes but associated to diseases, variants predicted to be damaging by in silico algorithms, and matches to databases such as ClinVar with fewer than two stars, BRCA Share, ALSOD, LOVD, and a locus-specific database for *APC*, *MSH2*, and *RB1*. In addition, we filtered for the 862 genes associated with cancer. Under these conditions, PeCanPIE detected 1128 *Silver* variants. To further filter *Silver* variants to choose those more likely pathogenic, we reasoned that if a variant is the same as a somatic mutation found in a cancer patient, specifically in the tumor biopsy, it could indicate a higher degree of confidence. Thus, to further explore potential pathogenic variants, we only considered *Silver* variants categorized as *Gold* or *Silver* in the PeCanPIE somatic category. This additional category considers somatic databases such as COSMIC and PCCP [63].

Thus, 106 *Silver* variants were obtained distributed in 87 genes (Table S2). Of 115 patients, 74 (64%) presented one or more silver variants, all heterozygous. Of these, 18 were missense, four splice site, one frameshift, one nonsense, six protein deletions, eight protein insertions, and 78 splice regions. Of these genes, the most frequent were *AKAP9* and *ATM* in four patients, followed by *KMT2D*, *MGA*, *COL3A1*, *NCOR2*, *ERBB2*, and *MLLT3* in three patients each. We noted the following gene families: *BRCA1/2*, *CDK 4/N2A*, *ERBB 2/3*, *ERCC 1/2/3*, *FANC A/D2/E/M*, *MRE 11/11A*, *NOTCH 1/2*, and *SMARC A4/B1/E1*. Interestingly, we detected the same missense mutation in *ERBB2* (R896H) in two patients. The gene position has been reported to activate HER2 function (R896C) [64], suggesting that R896H could affect normal function and potentially contribute to tumorigenesis.

## 4. Discussion

In the current work, we report an exome analysis of 115 young Mexican BC patients using the pipeline PeCanPIE focused on well-defined evidence of pathogenicity following ACMG guidelines. To our knowledge, this is the first effort for Mexican patients covering the germline whole exome. Previous efforts in Mexico and Latin America have focused on gene panels from 20 to 140 genes irrespective of the age of diagnosis [12,65]. Similar approaches showed a prevalence of 10.2% of pathogenic variants in BC in the USA [66]. Nevertheless, age at diagnosis is important because it may indicate an accelerated tumorigenesis process supported by recent reports showing an increase in BC incidence in young women [1,2,67–70]. Therefore, we focused on extreme phenotypes patients, where BC was diagnosed at 40 years old or less. We found that 39 patients (34%) showed a likely pathogenic *Gold* variant in 40 genes. This finding is higher than recent prevalence estimations in Latin American countries (13–25%) [42], which is likely due to our increased analysis in over the more than 800 genes by whole exome, and a higher genetic risk background of the younger population.

Comparing our *Gold* variants results with those of Quezada-Urban et al. in Mexican's BC, where more than 53% were older than 40 years old and using a panel of 143 genes [12], only six genes were overlapping (*BRCA1/2*, *ATM*, *WRN*, *RAD51C*, and *CHEK2*). We did not find variants in 15 of the 21 genes reported (*MSR1*, *ERCC3* [1 Silver], *LIG4*, *PDE11A*, *ATR*, *FANCI/B/C/L/M*, *RECQL4*, *SDHB*, *MLH1*, *NBN*, and *PTEN*), and we found 36 other genes which were not present in the Quezada-Urban et al. study. Some genes, however, showed similar gene families, such as *ERCC1/6* and *FANCE*. We also noted that many genes were not reported in BC for Latin-American countries (Table 2).

In our sample, we noted that 58% of the patients were overweight at the time of diagnosis, which is consistent with the 60% reported by Quezada-Urban et al. [12] and other reports in Mexico [2]. Weight loss in the young woman has been associated with lower cancer risk in *BRCA1* carriers [71]. Thus, we explored possible associations of known variants (*Gold*) to BMI. We observed a small but significant increase in BMI among *Gold* variant carriers. These results should be confirmed in larger cohorts.

We noted many *Gold* variants in genes not previously reported in Latin American cohorts but reported in BC in the USA, Europe, Asia, or other cancer types and gene families. For instance, Fanconi anemia and excision repair genes (*FANCI* and *ERCC* genes) have been reported in Latin American BC cases [11]. These findings highlight the use of a broader set of genes combined with powerful analysis tools, to expand the results.

PeCanPIE uses a pipeline considering the observed variant frequency among 'healthy' populations deposited in databases such as ExAC, which is primarily based on Caucasian populations [72]. In the *Silver* category, we noted few variants present in many individuals in our cohort even with low allelic fraction in ExAC, confirming that estimations of Latin America variations are needed to identify common variants in this population. Although PeCanPIE was initially conceived for pediatric cancers, they included several gene databases from other cancers reaching 712 genes. To complement this, we added 723 genes from COSMIC v94. Thus, our analysis was limited to the unified 862 genes from these two sources. If we extend the analysis to the whole exome and focus on *Gold* variants, besides the two common variants that would need to be removed, six genes are added (*C8B*, *DMD*, *HBB*, *IRF8*, *KCNQ1*, and *MYBPC3*, of which three are splices, two frameshifts, and one nonsense). Nevertheless, these were not considered in our analysis, since we focus on more likely genes with a stronger background in the cancer context.

We focused on *Gold* variants because pathogenicity is theoretically the highest provided by PeCanPIE. Nevertheless, *Silver* variants may also show interesting results, such as that mentioned for *HER2* (*ERBB2* R896H) and other gene families such as *CDK*, *BRCA1/2*, *ERBB*, *ERCC*, and *FANCI*. However, more careful revision is needed for *Silver* variants. For example, we noted a missense mutation in five patients in *WRN* (R834C) that has been shown to abolish important *WRN* function [73]. Although the variant was filtered out because of quality criteria (fewer than 10 reads), this polymorphism is frequent in the Mexican and is also unlikely to be pathogenic [74]. Thus, *Silver* variants should be handled more thoroughly. This evidence also highlights that annotation tools are crucial to facilitate interpretation but that results must be revised, and tools should be continuously updated.

Overall, our study provides candidate pathogenic variants in Mexican YBC, a barely studied population. Some variants need more careful analyses; for example, those regarding splice site variants and those in *RBM8A*. In addition, recent evidence rises questions even for well-known breast cancer genes [75]. Thus, confirmatory information may be needed either by specific experimental assays or analyses of large cohorts to potentially translate our findings into clinical practice.

## 5. Conclusions

We conclude that using whole exome sequencing to analyze an extended set of cancer genes, and a rigorous bioinformatic pipeline that includes PeCanPIE, we were able to identify candidate pathogenic genes for a more extensive set in young, Mexican breast cancer patients.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14071647/s1>, Table S1: Clinical Data, Table S2: Gold Variants.

**Author Contributions:** Conceptualization, L.G.-F.-R. and A.M.; methodology, L.G.-F.-R., A.M., M.T.-M., L.G., R.O.-L. and V.T.; software, V.T.; validation, L.G.-F.-R.; formal analysis, L.G.-F.-R., A.L.B.-A., M.T.-M. and V.T.; investigation, L.G.-F.-R. and V.T.; data curation, A.L.B.-A., L.G., R.O.-L. and V.T.; writing—original draft preparation, A.L.B.-A. and V.T.; writing—review and editing, L.G.-F.-R., A.L.B.-A. and V.T.; visualization, A.L.B.-A., V.T. and M.T.-M.; supervision, A.M. and V.T.; project administration, L.G.-F.-R. and A.M.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by the Mexican National Council of Science and Technology (CONACyT, grant number S0008-2016-1/272823) and supported by the program of *Cátedras-CONACyT* and the Intramural Program of the NIH. A.L.B.-A. has a CONACyT fellowship 662328. The APC was funded by Tecnológico de Monterrey.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Research Ethics Committee of National Cancer Institute of Mexico (INCAN, Instituto Nacional de Cancerología) (protocol code CEI/1123/16 approved in 2016).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study may be available on request from the corresponding author. The data are not publicly available due to embargo period.

**Acknowledgments:** The authors would like to express our sincere gratitude to Rosa María Álvarez-Gómez, Yulianna Sánchez, and Paulina Álvarez and Talia Wegman, from the Hereditary Cancer Clinic at the National Cancer Institute in Mexico, for providing the facilities to recruit the patients for this study. The authors thank Michael Dean for facilitating whole exome sequencing at the NCI. Verónica Guzmán for her support in administering this project.

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

## References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. Villarreal-Garza, C.; Platas, A.; Miaja, M.; Fonseca, A.; Mesa-Chavez, F.; Garcia-Garcia, M.; Chapman, J.A.; Lopez-Martinez, E.A.; Pineda, C.; Mohar, A.; et al. Young women with breast cancer in Mexico: Results of the pilot phase of the Joven & Fuerte prospective cohort. *J. Glob. Oncol.* **2020**, *6*, 395–406. [[CrossRef](#)]
3. Gómez-Flores-ramos, L.; Castro-Sánchez, A.; Peña-Curiel, O.; Mohar-Betancourt, A. Molecular biology in young women with breast cancer: From tumor gene Expression to DNA mutations. *Rev. Invest. Clin.* **2017**, *69*, 181–192. [[CrossRef](#)] [[PubMed](#)]
4. Kudela, E.; Samec, M.; Kubatka, P.; Nachajova, M.; Laucekova, Z.; Liskova, A.; Dokus, K.; Biringer, K.; Simova, D.; Gabonova, E.; et al. Breast cancer in young women: Status quo and advanced disease management by a predictive, preventive, and personalized approach. *Cancers* **2019**, *11*, 1791. [[CrossRef](#)] [[PubMed](#)]
5. Gabriel, C.A.; Domchek, S.M. Breast cancer in young women. *Breast Cancer Res.* **2010**, *12*, 212. [[CrossRef](#)] [[PubMed](#)]
6. Assi, H.; Khoury, K.; Dbouk, H.; Khalil, L.; Mouhieddine, T.; El Saghir, N. Epidemiology and prognosis of breast cancer in young women. *J. Thorac. Dis.* **2013**, *5* (Suppl. S1), S2–S8. [[CrossRef](#)] [[PubMed](#)]
7. Kast, K.; Rhiem, K.; Wappenschmidt, B.; Hahnen, E.; Hauke, J.; Bluemcke, B.; Zarghooni, V.; Herold, N.; Ditsch, N.; Kiechle, M.; et al. Prevalence of BRCA1/2 germline mutations in 21 401 families with breast and ovarian cancer. *J. Med. Genet.* **2016**, *53*, 465–471. [[CrossRef](#)]
8. Nielsen, F.C.; Van Overeem Hansen, T.; Sørensen, C.S. Hereditary breast and ovarian cancer: New genes in confined pathways. *Nat. Rev. Cancer* **2016**, *16*, 599–612. [[CrossRef](#)]
9. Wendt, C.; Margolin, S. Identifying breast cancer susceptibility genes—A review of the genetic background in familial breast cancer. *Acta Oncol.* **2019**, *58*, 135–146. [[CrossRef](#)]
10. Economopoulou, P.; Dimitriadis, G.; Psyrris, A. Beyond BRCA: New hereditary breast cancer susceptibility genes. *Cancer Treat. Rev.* **2015**, *41*, 1–8. [[CrossRef](#)]
11. Urbina-Jara, L.K.; Rojas-Martinez, A.; Martinez-Ledesma, E.; Aguilar, D.; Villarreal-Garza, C.; Ortiz-Lopez, R. Landscape of germline mutations in dna repair genes for breast cancer in latin america: Opportunities for parp-like inhibitors and immunotherapy. *Genes* **2019**, *10*, 786. [[CrossRef](#)] [[PubMed](#)]

12. Urban, R.Q.; Velásquez, C.E.D.; Gitler, R.; Castillo, M.P.R.; Toporek, M.S.; Morales, A.F.; García, O.M.; Esquivel, L.G.; Mejía, G.T.; Dean, M.; et al. Comprehensive Analysis of Germline Variants in Mexican Patients with Hereditary Breast and Ovarian Cancer Susceptibility. *Cancers* **2018**, *10*, 361. [[CrossRef](#)] [[PubMed](#)]
13. Fountzilas, C.; Kaklamani, V.G. Multi-gene Panel Testing in Breast Cancer Management. *Optim. Breast Cancer Manag.* **2018**, *173*, 121–140. [[CrossRef](#)]
14. Spurdle, A.B.; Bowman, M.A.; Shamsani, J.; Kirk, J. Endometrial cancer gene panels: Clinical diagnostic vs research germline DNA testing. *Mod. Pathol.* **2017**, *30*, 1048–1068. [[CrossRef](#)] [[PubMed](#)]
15. Smolle, M.A.; Kashofer, K.; Riedl, J.M.; Stotz, M.; Gerger, A. Genetic Analysis Using a Gene Panel in 87 Caucasian Patients With Colorectal Cancer: Own Results and Review of Literature. *Anticancer Res.* **2019**, *39*, 847–852. [[CrossRef](#)]
16. Yatabe, Y.; Sunami, K.; Goto, K.; Nishio, K.; Aragane, N.; Ikeda, S.; Inoue, A.; Kinoshita, I.; Kimura, H.; Sakamoto, T.; et al. Multiplex gene-panel testing for lung cancer patients. *Pathol. Int.* **2020**, *70*, 921–931. [[CrossRef](#)]
17. Auton, A.; Abecasis, G.R.; Altshuler, D.M.; Durbin, R.M.; Bentley, D.R.; Chakravarti, A.; Clark, A.G.; Donnelly, P.; Eichler, E.E.; Flück, P.; et al. A global reference for human genetic variation. *Nature* **2015**, *68*, 526.
18. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)]
19. Landrum, M.J.; Chitipiralla, S.; Brown, G.R.; Chen, C.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; Kaur, K.; Liu, C.; et al. ClinVar: Improvements to accessing data. *Nucleic Acids Res.* **2020**, *48*, D835–D844. [[CrossRef](#)]
20. Edmonson, M.N.; Patel, A.N.; Hedges, D.J.; Wang, Z.; Rampersaud, E.; Kesserwan, C.A.; Zhou, X.; Liu, Y.; Newman, S.; Rusch, M.C.; et al. Pediatric Cancer Variant Pathogenicity Information Exchange (PeCanPIE): A cloud-based platform for curating and classifying germline variants. *Genome Res.* **2019**, *29*, 1555–1565. [[CrossRef](#)]
21. Sondka, Z.; Bamford, S.; Cole, C.G.; Ward, S.A.; Dunham, I.; Forbes, S.A. The COSMIC Cancer Gene Census: Describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* **2018**, *18*, 696–705. [[CrossRef](#)] [[PubMed](#)]
22. Robinson, G.; Parker, M.; Kranenburg, T.A.; Lu, C.; Chen, X.; Ding, L.; Phoenix, T.N.; Hedlund, E.; Wei, L.; Zhu, X.; et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature* **2012**, *488*, 43–48. [[CrossRef](#)] [[PubMed](#)]
23. Winham, S.J.; Larson, N.B.; Armasu, S.M.; Fogarty, Z.C.; Larson, M.C.; McCauley, B.M.; Wang, C.; Lawrenson, K.; Gayther, S.; Cunningham, J.M.; et al. Molecular signatures of X chromosome inactivation and associations with clinical outcomes in epithelial ovarian cancer. *Hum. Mol. Genet.* **2019**, *28*, 1331–1342. [[CrossRef](#)] [[PubMed](#)]
24. Scala, M.; Torella, A.; Severino, M.; Morana, G.; Castello, R.; Accogli, A.; Verrico, A.; Vari, M.S.; Cappuccio, G.; Pinelli, M.; et al. Three de novo DDX3X variants associated with distinctive brain developmental abnormalities and brain tumor in intellectually disabled females. *Eur. J. Hum. Genet.* **2019**, *27*, 1254–1259. [[CrossRef](#)] [[PubMed](#)]
25. Blok, L.S.; Madsen, E.; Juusola, J.; Gilissen, C.; Baralle, D.; Reijnders, M.R.F.; Venselaar, H.; Helmsmoortel, C.; Cho, M.T.; Hoischen, A.; et al. Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am. J. Hum. Genet.* **2015**, *97*, 343–352. [[CrossRef](#)]
26. Lai, M.-C.; Lee, Y.-H.W.; Tarn, W.-Y. The DEAD-Box RNA Helicase DDX3 Associates with Export Messenger Ribonucleoproteins as well as Tip-associated Protein and Participates in Translational Control. *Mol. Biol. Cell* **2008**, *19*, 3847–3858. [[CrossRef](#)]
27. Cargill, M.J.; Morales, A.; Ravishanker, S.; Warren, E.H. RNA helicase, DDX3X, is actively recruited to sites of DNA damage in live cells. *DNA Repair* **2021**, *103*, 103137. [[CrossRef](#)]
28. Midha, M.K.; Huang, Y.-F.; Yang, H.-H.; Fan, T.-C.; Chang, N.-C.; Chen, T.-H.; Wang, Y.-T.; Kuo, W.-H.; Chang, K.-J.; Shen, C.-Y.; et al. Comprehensive Cohort Analysis of Mutational Spectrum in Early Onset Breast Cancer Patients. *Cancers* **2020**, *12*, 2089. [[CrossRef](#)]
29. Kunadirek, P.; Chuaypen, N.; Jenjaroenpun, P.; Wongsurawat, T.; Pinjaroen, N.; Sirichindakul, P.; Nookaew, I.; Tangkijvanich, P. Cell-Free DNA Analysis by Whole-Exome Sequencing for Hepatocellular Carcinoma: A Pilot Study in Thailand. *Cancers* **2021**, *13*, 2229. [[CrossRef](#)]
30. Gellert, P.; Segal, C.V.; Gao, Q.; López-Knowles, E.; Martin, L.A.; Dodson, A.; Li, T.; Miller, C.A.; Lu, C.; Mardis, E.R.; et al. Impact of mutational profiles on response of primary oestrogen receptor-positive breast cancers to oestrogen deprivation. *Nat. Commun.* **2016**, *7*, 13294. [[CrossRef](#)]
31. Albers, C.A.; Paul, D.S.; Schulze, H.; Freson, K.; Stephens, J.C.; Smethurst, P.A.; Jolley, J.D.; Cvejic, A.; Kostadima, M.; Bertone, P.; et al. Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. *Nat. Genet.* **2012**, *44*, 435–439. [[CrossRef](#)] [[PubMed](#)]
32. Mei, N.; Chen, H.; Zhao, N.; Yi, Y.; Li, C. A Comprehensive Pan-Cancer Analysis of RBM8A Based on Data Mining. *J. Oncol.* **2021**, *2021*, 9983354. [[CrossRef](#)] [[PubMed](#)]
33. Encinas, G.; Sabelnykova, V.Y.; De Lyra, E.C.; Katayama, M.L.H.; Maistro, S.; Valle, P.W.M.D.V.; Pereira, G.F.D.L.; Rodrigues, L.M.; Serio, P.A.D.M.P.; De Gouvêa, A.C.R.C.; et al. Somatic mutations in early onset luminal breast cancer. *Oncotarget* **2018**, *9*, 22460–22479. [[CrossRef](#)] [[PubMed](#)]
34. Jalkh, N.; Chouery, E.; Haidar, Z.; Khater, C.; Atallah, D.; Ali, H.; Marafie, M.J.; Al-Mulla, M.R.; Al-Mulla, F.; Megarbane, A. Next-generation sequencing in familial breast cancer patients from Lebanon. *BMC Med. Genom.* **2017**, *10*, 8. [[CrossRef](#)]

35. Sokolenko, A.P.; Preobrazhenskaya, E.V.; Aleksakhina, S.N.; Iyevleva, A.G.; Mitiushkina, N.V.; Zaitseva, O.A.; Yatsuk, O.S.; Tiurin, V.I.; Strelkova, T.N.; Togo, A.V.; et al. Candidate gene analysis of BRCA1/2 mutation-negative high-risk Russian breast cancer patients. *Cancer Lett.* **2015**, *359*, 259–261. [[CrossRef](#)]
36. Ellingson, M.S.; Hart, S.N.; Kalari, K.R.; Suman, V.; Schahl, K.A.; Dockter, T.J.; Felten, S.J.; Sinnwell, J.P.; Thompson, K.J.; Tang, X.; et al. Exome sequencing reveals frequent deleterious germline variants in cancer susceptibility genes in women with invasive breast cancer undergoing neoadjuvant chemotherapy. *Breast Cancer Res. Treat.* **2015**, *153*, 435–443. [[CrossRef](#)]
37. Tedaldi, G.; Tebaldi, M.; Zampiga, V.; Danesi, R.; Arcangeli, V.; Ravegnani, M.; Cangini, I.; Pirini, F.; Petracci, E.; Rocca, A.; et al. Multiple-gene panel analysis in a case series of 255 women with hereditary breast and ovarian cancer. *Oncotarget* **2017**, *8*, 47064–47075. [[CrossRef](#)]
38. Thompson, E.R.; Doyle, M.A.; Ryland, G.L.; Rowley, S.M.; Choong, D.Y.H.; Tothill, R.W.; Thorne, H.; Barnes, D.R.; Li, J.; Ellul, J.; et al. Exome Sequencing Identifies Rare Deleterious Mutations in DNA Repair Genes FANCC and BLM as Potential Breast Cancer Susceptibility Alleles. *PLoS Genet.* **2012**, *8*, e1002894. [[CrossRef](#)]
39. Sokolenko, A.P.; Bogdanova, N.; Kluzniak, W.; Preobrazhenskaya, E.V.; Kuligina, E.S.; Iyevleva, A.G.; Aleksakhina, S.N.; Mitiushkina, N.V.; Gorodnova, T.V.; Bessonov, A.A.; et al. Double heterozygotes among breast cancer patients analyzed for BRCA1, CHEK2, ATM, NBN/NBS1, and BLM germ-line mutations. *Breast Cancer Res. Treat.* **2014**, *145*, 553–562. [[CrossRef](#)]
40. Torrezan, G.T.; de Almeida, F.G.d.S.R.; Figueiredo, M.C.P.; Barros, B.D.d.F.; de Paula, C.A.A.; Valieris, R.; de Souza, J.E.S.; Ramalho, R.F.; da Silva, F.C.C.; Ferreira, E.N.; et al. Complex Landscape of Germline Variants in Brazilian Patients With Hereditary and Early Onset Breast Cancer. *Front. Genet.* **2018**, *9*, 161. [[CrossRef](#)]
41. Scarpitta, R.; Zanna, I.; Aretini, P.; Gambino, G.; Scatena, C.; Mei, B.; Ghilli, M.; Rossetti, E.; Roncella, M.; Congregati, C.; et al. Germline investigation in male breast cancer of DNA repair genes by next-generation sequencing. *Breast Cancer Res. Treat.* **2019**, *178*, 557–564. [[CrossRef](#)] [[PubMed](#)]
42. Oliver, J.; Quezada Urban, R.; Franco Cortés, C.A.; Díaz Velásquez, C.E.; Montealegre Paez, A.L.; Pacheco-Orozco, R.A.; Castro Rojas, C.; García-Robles, R.; López Rivera, J.J.; Gaitán Chaparro, S.; et al. Latin American Study of Hereditary Breast and Ovarian Cancer LACAM: A Genomic Epidemiology Approach. *Front. Oncol.* **2019**, *9*, 1429. [[CrossRef](#)] [[PubMed](#)]
43. Verhoeft, K.R.; Ngan, H.L.; Lui, V.W.Y. The cylindromatosis (CYLD) gene and head and neck tumorigenesis. *Cancers Head Neck* **2016**, *1*, 10. [[CrossRef](#)] [[PubMed](#)]
44. Heddar, A.; Fermey, P.; Coutant, S.; Angot, E.; Sabourin, J.-C.; Michelin, P.; Parodi, N.; Charbonnier, F.; Vezain, M.; Bougeard, G.; et al. Familial solitary chondrosarcoma resulting from germline EXT2 mutation. *Genes Chromosom. Cancer* **2017**, *56*, 128–134. [[CrossRef](#)]
45. Santos, S.C.L.; Rizzo, I.M.P.O.; Takata, R.I.; Speck-Martins, C.E.; Brum, J.M.; Sollaci, C. Analysis of mutations in EXT1 and EXT2 in Brazilian patients with multiple osteochondromas. *Mol. Genet. Genom. Med.* **2018**, *6*, 382–392. [[CrossRef](#)]
46. Afshar, A.R.; Pekmezci, M.; Bloomer, M.M.; Cadenas, N.J.; Stevers, M.; Banerjee, A.; Roy, R.; Olshen, A.B.; Van Ziffle, J.; Onodera, C.; et al. Next-Generation Sequencing of Retinoblastoma Identifies Pathogenic Alterations beyond RB1 Inactivation That Correlate with Aggressive Histopathologic Features. *Ophthalmology* **2020**, *127*, 804–813. [[CrossRef](#)]
47. Park, S.; Supek, F.; Lehner, B. Systematic discovery of germline cancer predisposition genes through the identification of somatic second hits. *Nat. Commun.* **2018**, *9*, 2601. [[CrossRef](#)]
48. Benusiglio, P.R.; Couvé, S.; Gilbert-Dussardier, B.; Deveaux, S.; Le Jeune, H.; Da Costa, M.; Fromont, G.; Memeteau, F.; Yacoub, M.; Coupier, I.; et al. A germline mutation in PBRM1 predisposes to renal cell carcinoma. *J. Med. Genet.* **2015**, *52*, 426–430. [[CrossRef](#)]
49. Spinella, J.-F.; Healy, J.; Saillour, V.; Richer, C.; Cassart, P.; Ouimet, M.; Sinnett, D. Whole-exome sequencing of a rare case of familial childhood acute lymphoblastic leukemia reveals putative predisposing mutations in Fanconi anemia genes. *BMC Cancer* **2015**, *15*, 539. [[CrossRef](#)]
50. Taeubner, J.; Brozou, T.; Qin, N.; Bartl, J.; Ginzl, S.; Schaper, J.; Felsberg, J.; Fulda, S.; Vokuhl, C.; Borkhardt, A.; et al. Congenital embryonal rhabdomyosarcoma caused by heterozygous concomitant PTCH1 and PTCH2 germline mutations. *Eur. J. Hum. Genet.* **2018**, *26*, 137–142. [[CrossRef](#)]
51. Li, N.; McInerney, S.; Zethoven, M.; Cheasley, D.; Lim, B.W.X.; Rowley, S.M.; Devereux, L.; Grewal, N.; Ahmadloo, S.; Byrne, D.; et al. Combined Tumor Sequencing and Case-Control Analyses of RAD51C in Breast Cancer. *J. Natl. Cancer Inst.* **2019**, *111*, 1332–1338. [[CrossRef](#)] [[PubMed](#)]
52. Blombery, P.; Fox, L.C.; Ryland, G.L.; Thompson, E.R.; Lickiss, J.; McBean, M.; Yerneni, S.; Hughes, D.; Greenway, A.; Mechinaud, F.; et al. Utility of clinical comprehensive genomic characterization for diagnostic categorization in patients presenting with hypocellular bone marrow failure syndromes. *Haematologica* **2021**, *106*, 64–73. [[CrossRef](#)] [[PubMed](#)]
53. Ma, H.; Song, B.; Guo, S.; Li, G.; Jin, G. Identification of germline and somatic mutations in pancreatic adenosquamous carcinoma using whole exome sequencing. *Cancer Biomark.* **2020**, *27*, 389–397. [[CrossRef](#)]

54. Sigurdson, A.J.; Brenner, A.V.; Roach, J.A.; Goudeva, L.; Müller, J.A.; Nerlich, K.; Reiners, C.; Schwab, R.; Pfeiffer, L.; Waldenberger, M.; et al. Selected single-nucleotide polymorphisms in FOXE1, SERPINA5, FTO, EVPL, TICAM1 and SCARB1 are associated with papillary and follicular thyroid cancer risk: Replication study in a German population. *Carcinogenesis* **2016**, *37*, 677–684. [[CrossRef](#)] [[PubMed](#)]
55. Walker, C.J.; Miranda, M.A.; O'Hern, M.J.; McElroy, J.P.; Coombes, K.R.; Bundschuh, R.; Cohn, D.E.; Mutch, D.G.; Goodfellow, P.J. Patterns of CTCF and ZFH3 Mutation and Associated Outcomes in Endometrial Cancer. *J. Natl. Cancer Inst.* **2015**, *107*, 11. [[CrossRef](#)]
56. Rai, R.; Sharma, K.L.; Tiwari, S.; Misra, S.; Kumar, A.; Mittal, B. DCC (deleted in colorectal carcinoma) gene variants confer increased susceptibility to gallbladder cancer (Ref. No.: Gene-D-12-01446). *Gene* **2013**, *518*, 303–309. [[CrossRef](#)] [[PubMed](#)]
57. Kim, S.J.; Lubinski, J.; Huzarski, T.; Møller, P.; Armel, S.; Karlan, B.Y.; Senter, L.; Eisen, A.; Foulkes, W.D.; Singer, C.F.; et al. Weight Gain and the Risk of Ovarian Cancer in BRCA1 and BRCA2 Mutation Carriers. *Cancer Epidemiol. Biomark. Prev.* **2021**, *30*, 2038–2043. [[CrossRef](#)]
58. Qian, F.; Rookus, M.A.; Leslie, G.; Risch, H.A.; Greene, M.H.; Aalfs, C.M.; Adank, M.A.; Adlard, J.; Agnarsson, B.A.; Ahmed, M.; et al. Mendelian randomisation study of height and body mass index as modifiers of ovarian cancer risk in 22,588 BRCA1 and BRCA2 mutation carriers. *Br. J. Cancer* **2019**, *121*, 180–192. [[CrossRef](#)]
59. Qian, F.; Wang, S.; Mitchell, J.; McGuffog, L.; Barrowdale, D.; Leslie, G.; Oosterwijk, J.C.; Chung, W.K.; Evans, D.G.; Engel, C.; et al. Height and Body Mass Index as Modifiers of Breast Cancer Risk in BRCA1/2 Mutation Carriers: A Mendelian Randomization Study. *J. Natl. Cancer Inst.* **2019**, *111*, 350–364. [[CrossRef](#)]
60. Wang, Y.A.; Jian, J.-W.; Hung, C.-F.; Peng, H.-P.; Yang, C.-F.; Cheng, H.-C.S.; Yang, A.-S. Germline breast cancer susceptibility gene mutations and breast cancer outcomes. *BMC Cancer* **2018**, *18*, 315. [[CrossRef](#)]
61. Zhu, Y.; Wu, J.; Zhang, C.; Sun, S.; Zhang, J.; Liu, W.; Zhang, Z.; Huang, J. BRCA mutations and survival in breast cancer: An updated systematic review and meta-analysis. *Oncotarget* **2016**, *7*, 70113–70127. [[CrossRef](#)] [[PubMed](#)]
62. Verhoog, L.C.; Brekelmans, C.T.M.; Seynaeve, C.; Dahmen, G.; Van Geel, A.N.; Bartels, C.C.M.; Tilanus-Linthorst, M.M.A.; Wagner, A.; Devilee, P.; Halley, D.J.J.; et al. Survival in Hereditary Breast Cancer Associated with Germline Mutations of BRCA2. *J. Clin. Oncol.* **1999**, *17*, 3396–3402. [[CrossRef](#)] [[PubMed](#)]
63. Downing, J.R.; Wilson, R.K.; Zhang, J.; Mardis, E.R.; Pui, C.H.; Ding, L.; Ley, T.J.; Evans, W.E. The pediatric cancer genome project. *Nat. Genet.* **2012**, *44*, 619–622. [[CrossRef](#)] [[PubMed](#)]
64. Bose, R.; Kavuri, S.M.; Searleman, A.C.; Shen, W.; Shen, D.; Koboldt, D.C.; Monsey, J.; Goel, N.; Aronson, A.B.; Li, S.; et al. Activating HER2 Mutations in HER2 Gene Amplification Negative Breast Cancer. *Cancer Discov.* **2013**, *3*, 224–237. [[CrossRef](#)] [[PubMed](#)]
65. Da Costa E Silva Carvalho, S.; Cury, N.M.; Brotto, D.B.; De Araujo, L.F.; Rosa, R.C.A.; Texeira, L.A.; Plaça, J.R.; Marques, A.A.; Peronni, K.C.; Ruy, P.D.C.; et al. Germline variants in DNA repair genes associated with hereditary breast and ovarian cancer syndrome: Analysis of a 21 gene panel in the Brazilian population. *BMC Med. Genom.* **2020**, *13*, 21. [[CrossRef](#)] [[PubMed](#)]
66. Couch, F.J.; Shimelis, H.; Hu, C.; Hart, S.N.; Polley, E.C.; Na, J.; Hallberg, E.; Moore, R.; Thomas, A.; Lilyquist, J.; et al. Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. *JAMA Oncol.* **2017**, *3*, 1190–1196. [[CrossRef](#)] [[PubMed](#)]
67. Fazel, A.; Hasanpour-Heidari, S.; Salamat, F.; Rajaie, S.; Kazeminezhad, V.; Naeimi-Tabiei, M.; Jahangirrad, A.; Sedaghat, S.M.; Hosseinpour, R.; Ghasemi-Kebria, F.; et al. Marked increase in breast cancer incidence in young women: A 10-year study from Northern Iran, 2004–2013. *Cancer Epidemiol.* **2019**, *62*, 101573. [[CrossRef](#)]
68. Dimitrova, N.; Znaor, A.; Agius, D.; Eser, S.; Sekerija, M.; Ryzhov, A.; Primic-Žakelj, M.; Coebergh, J.W.; Agius, D.; Coza, D.; et al. Breast cancer in South-Eastern European countries since 2000: Rising incidence and decreasing mortality at young and middle ages. *Eur. J. Cancer* **2017**, *83*, 43–55. [[CrossRef](#)]
69. Rocha-Brischiliari, S.C.; De Oliveira, R.R.; Andrade, L.; Brischiliari, A.; Gravena, A.A.F.; Carvalho, M.D.D.B.; Pelloso, S.M. The Rise in Mortality from Breast Cancer in Young Women: Trend Analysis in Brazil. *PLoS ONE* **2017**, *12*, e0168950. [[CrossRef](#)]
70. Colonna, M.; Delafosse, P.; Uhry, Z.; Poncet, F.; Arveux, P.; Molinie, F.; Cherie-Challine, L.; Grosclaude, P. Is breast cancer incidence increasing among young women? An analysis of the trend in France for the period 1983–2002. *Breast* **2008**, *17*, 289–292. [[CrossRef](#)]
71. Friebel, T.M.; Domchek, S.M.; Rebbeck, T.R. Modifiers of Cancer Risk in BRCA1 and BRCA2 Mutation Carriers: Systematic Review and Meta-Analysis. *J. Natl. Cancer Inst.* **2014**, *106*, dju091. [[CrossRef](#)] [[PubMed](#)]
72. Lek, M.; Karczewski, K.; Minikel, E.; Samocha, K.; Banks, E.; Fennell, T.; O'Donnell-Luria, A.; Ware, J.; Hill, A.; Cummings, B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *bioRxiv* **2015**, *536*, 030338. [[CrossRef](#)] [[PubMed](#)]
73. Kamath-Loeb, A.S.; Welcsh, P.; Waite, M.; Adman, E.T.; Loeb, L.A. The Enzymatic Activities of the Werner Syndrome Protein Are Disabled by the Amino Acid Polymorphism R834C. *J. Biol. Chem.* **2004**, *279*, 55499–55505. [[CrossRef](#)] [[PubMed](#)]
74. Kamath-Loeb, A.S.; Zavala-van Rankin, D.G.; Flores-Morales, J.; Emond, M.J.; Sidorova, J.M.; Carnevale, A.; Del Carmen Cárdenas-Cortés, M.; Norwood, T.H.; Monnat, R.J.; Loeb, L.A.; et al. Homozygosity for the WRN Helicase-Inactivating Variant, R834C, does not confer a Werner syndrome clinical phenotype. *Sci. Rep.* **2017**, *7*, srep44081. [[CrossRef](#)]
75. Dorling, L.; Carvalho, S.; Allen, J.; González-Neira, A.; Luccarini, C.; Wahlström, C.; Pooley, K.A.; Parsons, M.T.; Fortunato, C. Breast Cancer Risk Genes—Association Analysis in More than 113,000 Women. *N. Engl. J. Med.* **2021**, *384*, 428–439. [[CrossRef](#)]