Article

Not peer-reviewed version

Genetic Insights into Breast Cancer in Northeastern Mexico: Unveiling Gene-Environment Interactions and Their Links to Obesity and Metabolic Diseases

Hugo Leonid Gallardo-Blanco , María de Loures Garza-Rodríguez , Diana Cristina Pérez-Ibave , Víctor M. Salinas-Torres , Moisés González-Escamilla , Rafael Piñeiro-Retif , Ricardo M. Cerda-Flores , Oscar Vidal Gutierrez , Celia N. Sanchez-Dominguez *

Posted Date: 23 January 2025

doi: 10.20944/preprints202501.1758.v1

Keywords: breast cancer; genetic risk variants; obesity; overweight; insulin resistance; inflammation; menopause; Gene-Environment Interactions; MMP8; FTO

Article

Genetic Insights into Breast Cancer in Northeastern Mexico: Unveiling Gene-Environment Interactions and Their Links to Obesity and Metabolic Diseases

Hugo Leonid Gallardo-Blanco [1], María de Loures Garza-Rodríguez [1], Diana Cristina Pérez Ibave [1], Carlos Horacio Burciaga-Flores [1], Víctor M. Salinas-Torres [2], Moisés González-Escamilla [1], Rafael Piñeiro-Retif [1], Ricardo M. Cerda-Flores [3], Oscar Vidal-Gutiérrez [1] and Celia N. Sanchez-Dominguez [4,*]

[1] Servicio de Oncología, Centro Universitario Contra el Cáncer (CUCC), Hospital Universitario "Dr. José Eleuterio González", Universidad Autónoma de Nuevo León,Monterrey 66451, Nuevo León, Mexico
[2] Hospital General de Culiacán, Culiacán Rosales 80030, Mexico
[3] School of Nursing, Universidad Autónoma de Nuevo León, Monterrey, Nuevo León, CP 64460
[4] Departamento de Bioquímica y Medicina Molecular, Facultad de Medicina, Universidad Autónoma de Nuevo León, Monterrey 64460, Mexico
* Correspondence: celia.sanchezdm@uanl.edu.mx

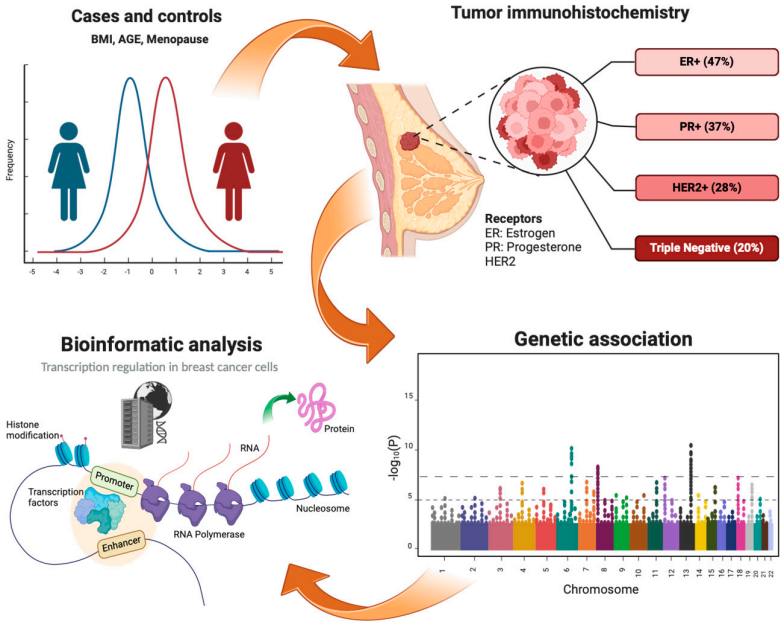
Simple Summary: Breast cancer is a common disease, and its occurrence in Mexico has been rising over the last decade. This study investigates specific genetic DNA variants in women from Northeastern Mexico to understand how genes and age, body mass index, and menopause contribute to breast cancer risk. Studying interactions between gene variants related to insulin resistance, obesity, and inflammation may clarify breast cancer development. This knowledge improves our understanding of breast cancer risk factors and finding biomarkers for personalized prevention strategies in Mexican mestizo women.

Abstract: Breast cancer (BC), one of the most common cancers, has increased in Mexico during the past decade, along with other chronic and metabolic diseases. Herein, we analyzed 121 SNPs (92 previously associated SNPs to BC, type 2 diabetes, obesity, overweight, insulin resistance, inflammation, and 36 informative ancestry markers) in 91 confirmed BC cases and 126 unaffected BC women from Northeastern Mexico. The relationship of these 121 SNPs with BC, considering BMI, menopause status, and age as cofactors, was explored by a gene-environment (G × E) interaction multi-locus model. Twelve gene variants were significantly associated with BC: rs3856806 PPARG, rs12792229 *MMP8*, and rs5218 *KCNJ11-ABCC8*. Whereas rs3917542 *PON1*; rs3750804 and rs3750805 *TCF7L2*; rs1121980 and rs3751812 *FTO*; rs12946618 *RPTOR*; rs2833483 *SCAF4*; rs11652805 *AMZ2P1-GNA13*; and rs1800955 *SCT-DEAF1-DRD4* were located in non-coding regions, involved in the accelerated decay of the mRNA transcripts, regulatory, and flanking regions. This study corroborated the association of BC along with menopause, age (above 45), obesity, and overweight with gene variants implicated in diabetes mellitus, obesity, insulin resistance, inflammation, and remodeling of the extracellular matrix.

Keywords: breast cancer; genetic risk variants; obesity; overweight; insulin resistance; inflammation; menopause; Gene-Environment Interactions; MMP8; FTO

 doi:10.20944/preprints202501.1758.v1

Graphical Abstract



Created in BioRender. Gallardo, H. (2024) https://BioRender.com/y71n320

1. Introduction

Breast cancer (BC) is the most frequent cancer in women worldwide (incidence), with 2,295,686 new cases and 665,684 deaths in women in 2022 [1]. BC represents 23.8% one-third of the total new cases of cancer diagnosed in women. In Mexico, there were 31,043 new cases of BC, 8,195 deaths, and 102,223 prevalent cases (5-year) in 2022 [2].

BC is the leading cause of death associated with neoplastic diseases in women and with an increasing number of cases due to risk factors such as obesity and age (Makowski et al., 2013).

Obesity is linked to BC by associated mediators such as endocrine, inflammatory, and oxidative stress that modify the tissue microenvironment [4–7]. There is evidence of an increased risk of BC recurrence and mortality in both pre and postmenopausal women with obesity compared with normal body mass index (BMI) women [8]. Also, age is linked to BC pathogenesis because cell aging increases the risk of developing cancer [9,10].

Estrogen exposure is one of the main contributors to BC, as there is a direct correlation between estrogen exposure and BC risk (early menarche and or delayed menopause has an increased risk).). On the other hand, pregnancy and lactation, due to the decreased levels of estrogen in these periods, reduce BC risk [11].

BC is considered a complex disease. Environmental and genetic factors contribute to BC carcinogenesis. Among the genetic risk factors involved in sporadic cancer, Single-Nucleotide Polymorphisms (SNPs) are the most common type of genomic variant (over 10 million) [obj], and some have been demonstrated to contribute to BC carcinogenesis [12].

Genome-wide association studies (GWAS) have identified over 1771 SNPs associated with a risk of BC (NHGRI-EBI GWAS catalog) on ClinVar and 1887 on 21 May 2024 in ENSEMBL [13]. Most GWAS studies are based on European origin populations; the Latin-Americans is underrepresented [14]. Latinos population is characterized by a broad admixture between European, indigenous American, and African ancestry [15,16]. For this reason, there is a need for a better approach for this population, including the employment of multi-locus models (multidimensional) in GWAS analysis to recover a model closer to reality.

SNPs often explain normal variation between individuals and frequently have minimal functional impact. BC-associated SNPs have low individual risk with an accumulative risk [17].

 doi:10.20944/preprints202501.1758.v1

Due to BC's multifactorial origin, we propose to explore the relationship between a set of SNPs previously reported as risk factors for BC, including metabolic features and informative ancestry markers from Northeastern Mexico. The present study considers BMI, menopause status, and age [5] cofactors by a gene-environment (G × E) interaction multi-locus statistical model.

This work investigated the association of specific genetic variants with breast cancer (BC) risk, menopause, and body mass index (BMI) in a population from Northeast Mexico. We found that SNPs located in regulatory regions may contribute to cancer development.

The findings emphasize the importance of considering factors such as age distribution and obesity prevalence, which may contribute to increased BC risk in this population. Furthermore, the study highlighted the significance of including ancestry-informative markers (AIMs) to minimize population stratification bias, thus ensuring an accurate representation of the mestizo population.

The results underscore the need for further research to validate these associations and explore their functional impacts on BC progression. This work contributes to understanding genetic risk factors in diverse populations, supporting the relevance of personalized medicine approaches tailored to specific genetic profiles for better cancer prevention and treatment strategies.

2. Materials and Methods

2.1. Scientific and Ethics Committees Approval

The study was conducted following the Declaration of Helsinki, and the protocol was approved by the Institutional Ethics Committee of the University Hospital, with the registration codes BI10–002 and GN13-001. All subjects (patients and controls) were invited to participate in the research project. An interview was performed, and once the patients agreed to participate, they signed an informed consent letter. Afterward, clinical and epidemiological information was collected, and blood samples were taken.

2.2. Study Design and Population

The study was retrospective and nested case-control. All the participants were Mexican women whose four generations were born in Northeast Mexico (Nuevo León, San Luis Potosi, Zacatecas, Coahuila, and Tamaulipas). Women with a family history of BC and pregnant women were excluded.

2.2. Patients Group

A total of 91 BC patients were selected among women receiving chemotherapy at two institutions in Monterrey City, Mexico: the Centro Universitario Contra el Cáncer (Hospital Universitario "Dr. José E. González" (HU), of the Universidad Autónoma de Nuevo León (UANL) and the Hospital de Especialidades #25 (Instituto Mexicano del Seguro Social (IMSS). Both are Mexican Northeast referral centers.

We included only patients with sporadic breast cancer. All BC cases were confirmed by histopathologic analysis.

2.2. Controls Group

Controls were women older than 44 years of age, without a personal history of cancer, and a BI-RADS 1–2 mammogram. Women with a family history of BC and who were pregnant were excluded.

2.3. Nucleic Acid Extraction

DNA was extracted from 200 µl EDTA-treated whole blood samples using the QIAamp® Blood Mini Kit on the automated QIAcube system (Qiagen, Hilden, Germany) according to manufacturer instructions. Purified DNA was collected in a final volume of 150 µl and stored at -20°C until use. Samples with low DNA quality extracted (1.7 ratios at 260/280 nm) were excluded from the study [18].

 doi:10.20944/preprints202501.1758.v1

2.4. SNP Selection

We selected 121 SNPs divided into three categories (Table S1): Risk factors for the BC (n=35), glucose-associated metabolic pathways (n=50), and ancestry markers (n=36). SNPs from 48 genes includes: *ANKK1, CAPN10, CCL, CD36, CDKAL1, ENPP1, ESR1, FABP2, FAS, FTO, GRIK4, HHEX, HNF1B, HNF4A*, hsa-miR-196a2, *IGF2BP2, IL10, IL1B, IL6, IL8, INHBB, KCNQ1, LEPR, LGALS3, LTA, MC4R, MMP2, MMP7, MMP8, MTHFR, NEGR1, NEUROD1, PCK1, PON1, PPARG, RPTOR, SCIN, SFRS15, SLC30A8, TCF7L2, TNF, UCP1, UCP2, UCP3, WFS1, WNT5B, ZNF365*, and *ZNF703*. We also included six intergenic regions from: CDKN2B-AS1: DMRTA1, CDC123: CAMK1D, SCT: DRD4, KCNJ11: ABCC8, USF1: TSTD1: ARHGAP30, and VLDLR: FLJ35024. Also, AIM (Ancestry-Informative Marker) was included to determine the population structure [19,20].

2.5. Genotype Analysis

Molecular analysis was performed using TaqMan® assays (Applied Biosystems, Carlsbad, CA) and analyzed in an OpenArray® NT Genotyping System (Applied Biosystems, Carlsbad, CA) (Table S2) [18]. We calculated the accuracy of the genotyping by comparison with concordance calls generated for ten samples genotyped three times. Samples with low-quality genotypes (Sample call rate ≥90%; SNPs per sample/the total number of SNPs in the dataset) were excluded from the study (Table S3) [18].

2.5. Statistical Methods

We calculated genotype statistics by marker and sample. Samples with inbreeding and a call rate <90% and markers with a call rate <0.88% were excluded. Hardy–Weinberg Equilibrium (HWE) was analyzed for each SNP (P <0.01).

For the association study, a multi-locus mixed model (MLMM) was performed [21–23] using SNP & Variation Suite v8 (Wang et al., 2016) (Golden Helix, Inc., Bozeman, MT). We used demographic data and ten genetic principal component analyses (PCA) as covariants [23] (Figure 1).

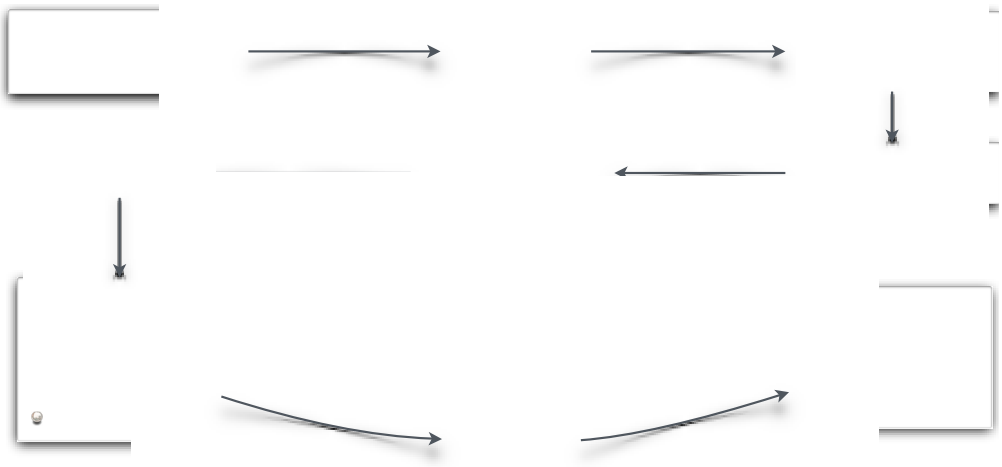


Figure 1. Schematic diagram of statistical design.

The quantile-quantile (Q-Q) plots with false discovery rate (FDR) corrected p-values were analyzed to avoid statistical overestimation. Additive, dominant, and recessive models were tested. The Composite haplotype method (CHM) was used for the Q-Q and linkage disequilibrium (LD) plots with the SNP & Variation Suite v8 (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com).

2.6. In Silico Variant Effect Predictor Analysis

 doi:10.20944/preprints202501.1758.v1

The Ensembl-VEP bioinformatics platform was used to classify regulatory region variants, nonsense-mediated decay transcript, and missense variants[24]. The analysis was based on the human genome annotation GRCh37/hg19.

2.7. In Silico Protein-Protein Interaction Analysis

Using the String database, a protein-protein interaction network, including physical, functional, and biological processes associated with genes co-segregating with DRD4, was created with a confidence score of 0.7 [25].

2.8. In Silico Related Pathways Analysis

The ToppGene Suite (Cincinnati, OH, USA) was used to identify pathways based on identifying related GO functional categories (biological processes and pathways) from significant genes associated with variants.

3. Results

3.1. Patients Data

Our cohort consisted of 217 women, including 91 breast cancer (BC) cases and 126 unaffected controls. The age range of patients and unaffected controls was between 45 and 72 years old. There were no statistically significant differences between the groups regarding age, height, and age at menarche (p>0.01). BMI was similar in both groups, with overweight being the most common status across the entire cohort and more prevalent in the BC group, and was statistically significant (p=0.011). The estrogen exposure was observed by age at menarche (p=0.231), menopausal status (p=0.004), age at menopause (p=0.025), use of oral contraceptives (p=0.110), number of children (p=0.187), and BMI (p=0.009). The anthropometric and physiological data of the participants are detailed in Table 1.

Table 1. Anthropometric and physiological characteristics of cases and controls.

Characteristics	Controls (n = 126)	Cases (n = 92)	p-value[2]
Age (years)[1]	54.93 ± 7.091	58.45 ± 8.811	0.023
Height (meters)[1]	1.578 ± 0.068	1.558 ± 0.080	0.048
BMI[1]	28.03 ± 4.950	29.92 ± 5.713	0.011
Menarche (Age, years)[1]	12.63 ± 1.543	12.88 ± 1.503	0.231
Menopause (Age, years)[1]	47.37 ± 5.630	45.29 ± 5.977	0.025
Menopause confirmed[3]	83 (65.87%)	76 (82.61%)	0.004
Oral contraceptives[3]	20 (15.87%)	23 (25.00%)	0.110
Children (number) [3]	111 (88.10%)	82 (89.13%)	0.187

[1]Results are presented as the mean ± standard deviation/standard error. [2]Unpaired t-test with Welch's correction p-value. [3]Number and percentage values. [4]Unpaired t-test with Welch's correction.

The cancer group exhibited more prolonged estrogen exposure than the non-cancer group, with BMI, early menopause, and menopausal status being statistically significant factors in BC patients. The immunohistochemistry analysis of the cancer patients reported the following frequencies: ER positive (46%), PR positive (36.96%), Her-2 positive (28.26%), and triple-negative (19.57%) (Table 2).

Table 2. Breast cancer tumor immunohistochemistry (n = 92), age range 45-79 years.

Immunohistochemistry	ER status n (%)	PR status n (%)	Her-2 status n (%)	TNBC status n (%)
Positive	43 (46.87%)	34 (36.96%)	26 (28.26%)	18 (19.57%)
Negative	49 (53.26%)	58 (63.04%)	66 (71.74%)	74 (80.47%)

ER Estrogen receptor; PR Progesterone receptor; TNBC Triple Negative Breast Cancer.

 doi:10.20944/preprints202501.1758.v1

3.2. Genotype Analysis

From the 128 SNPs considered, 121 SNPs were eligible for statistical analysis, and seven SNPs were excluded due to low quality. No significant genomic inflation (expresses the deviation of the distribution results observed compared to the distribution of the expected test statistic) [26,27] was detected between the observed and expected P-values in the MLMM (polygenic inheritance) under additive (Table S4), dominant (Table S5), and recessive (Table S6) genetic models (Fig. 2 and 3, Table 2) [26]. The genotype frequencies and HWE were reported in 12 BC risk-associated SNPs as there were statistically significant ones (Table 3).

Table 3. Genotype frequencies and Hardy-Weinberg equilibrium for variants associated in cases and controls with age range 45-79 years.

Chr[1]	Gene[2]	Variant	DD Frequency Cases/Controls[3]	Dd Frequency Cases/Controls[3]	dd Frequency Cases/Controls[3]	HWE *p* (Cases)[4]	HWE *p* (Controls)[5]
3	*PPARG*	rs3856806	TT (0.000/0.024)	TC (0.242/0.206)	CC (0.758/0.770)	0.190	0.437
7	*PON1*	rs3917542	TT (0.088/0.064)	TC (0.385/0.376)	CC (0.527/0.560)	0.657	0.977
10	*TCF7L2*	rs3750804	TT (0.033/0.074)	TC (0.300/0.262)	CC (0.667/0.664)	0.986	0.031
		rs3750805	TT (0.011/0.016)	TA (0.154/0.167)	AA (0.835/0.817)	0.698	0.449
11	*MMP8*	rs12792229	TT (0.000/0.000)	TG (0.011/0.000)	GG (0.989/1.000)	0.958	1.000
	SCT; DEAF1; DRD4	rs1800955	CC (0.156/0.135)	CT (0.344/0.423)	TT (0.500/0.441)	0.038	0.490
	KCNJ11; ABCC8	rs5218	AA (0.033/0.016)	AG (0.253/0.230)	GG (0.714/0.754)	0.589	0.900
16	*FTO*	rs1121980	AA (0.088/0.119)	AG (0.418/0.381)	GG (0.495/0.500)	0.996	0.222
		rs3751812	TT (0.044/0.040)	TG (0.352/0.360)	GG (0.604/0.600)	0.809	0.584
17	*AMZ2P1; GNA13*	rs11652805	CC (0.044/0.048)	CT (0.319/0.298)	TT (0.637/0.653)	0.877	0.511
	RPTOR	rs12946618	AA (0.012/0.008)	AG (0.256/0.182)	GG (0.733/0.810)	0.545	0.847
21	*SCAF4*	rs2833483	CC (0.144/0.158)	CT (0.422/0.433)	TT (0.433/0.408)	0.456	0.408

[1]Chromosome. [2]Gene or nearest genes. [3]Genotype frequencies. [4]Hardy-Weinberg Equilibrium for Cases. [5]Hardy-Weinberg Equilibrium for controls.

3.2. MLMM Testing

A significant association between BC and 12 gene variants (Table 4) was identified in the following genes: rs3856806 (PPARG), rs3917542 (PON1), rs3750804 and rs3750805 (TCF7L2), rs12792229 (MMP8), rs1121980 and rs3751812 (FTO), rs12946618 (RPTOR), rs2833483 (SCAF4), rs1800955 (SCT-DEAF1-DRD4), rs5218 (KCNJ11-ABCC8), and rs11652805 (AMZ2P1-GNA13).

 doi:10.20944/preprints202501.1758.v1

Table 4. Breast cancer variants associated with multi-locus mixed-model analysis, age 45-79 years.

Chr[1]	Gen[2]	Marker	Genetic Model[3]	P-Value[4]	FDR[5]	Regression Beta	Beta Standard Error	Prop. Var. Expl.[6]
3	*PPARG*	rs3856806	3	9.39E-04	2.84E-02	-1.014	0.302	7.59E-03
7	*PON1*	rs3917542	3	9.17E-04	3.70E-02	0.660	0.196	1.23E-03
10	*TCF7L2*	rs3750804	1	1.4E-05	5.6E-04	0.704	0.158	5.1E-03
		rs3750805	1	1.8E-05	5.3E-04	0.885	0.201	5.1E-03
11	*MMP8*	rs12792229	2	8.56E-09	1.04E-06	2.002	0.333	6.81E-03
	SCT;DEAF1;DRD4	rs1800955	1	4.6E-04	9.2E-03	0.395	0.111	5.1E-03
	KCNJ11;ABCC8	rs5218	3	1.78E-11	2.16E-09	1.859	0.261	1.58E-03
16	*FTO*	rs1121980	1	6.7E-08	8.1E-06	1.191	0.212	1.6E-02
		rs3751812	1	1.8E-06	1.1E-04	1.109	0.225	1.6E-02
17	*AMZ2P1;GNA13*	rs11652805	1	6.4E-05	1.6E-03	0.650	0.159	1.6E-02
	RPTOR	rs12946618	2	2.75E-07	1.66E-05	-0.898	0.169	7.97E-03
21	*SCAF4*	rs2833483	2	1.18E-06	4.77E-05	0.826	0.165	7.97E-03

[1]Chromosome. [2]Gene or nearest genes. [3]Genetic Model; 1: additive; 2: Dominant; 3: Recessive. [4]MLMM P-value adjusted for age, BMI, and menopause status and ten genetic PCA. [5]*p*-value False Discovery Rate (FDR) corrected. [6]Proportion of Variance Explained.

Figure 2 illustrates the –log10 of FDR *p*-values under the Q-Q´s additive, dominant, and recessive models. Figure 2 A plots the cluster of samples based on the similarity of two main genetic principal components. Figure 2B shows Q-Q plots for additive genetic model results (step 117 selected by Multiple PPA and Multiple Bonferroni). Figure 2 C shows the dominant genetic model results (step 117, chosen by Multiple and Bonferroni). Figure 2 D shows the recessive genetic model results (step 115 selected by Modified Bonferroni and Extended Bonferroni). TCF7L2 and FTO variants were associated in the additive genetic model, RPTOR and MMP8 in the dominant genetic model, and KCNJ11, RPTOR, PON1, and PPARG variants in the recessive genetic model. Figure 3 shows the results of the Manhattan plots by relative chromosome position of the association study in additive, dominant, and recessive models with MLMM.

 doi:10.20944/preprints202501.1758.v1

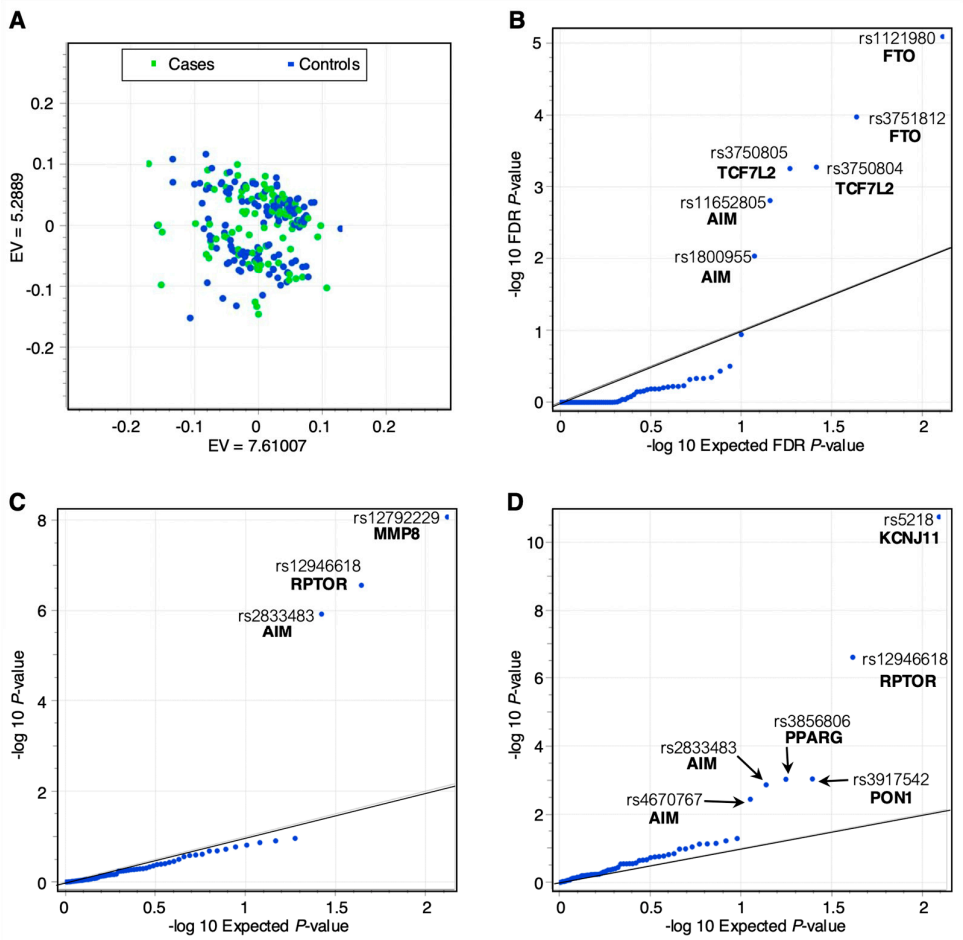


Figure 2. (A) Principal component analysis plot, expected vs. observed P-values in a P-P plot. (B) Additive genetic model step 117 selected by Multiple PPA and Multiple Bonferroni. (C) The dominant genetic model was chosen by Multiple and Bonferroni step 117. (D) Recessive genetic model step 115 selected by Modified Bonferroni and Extended Bonferroni.

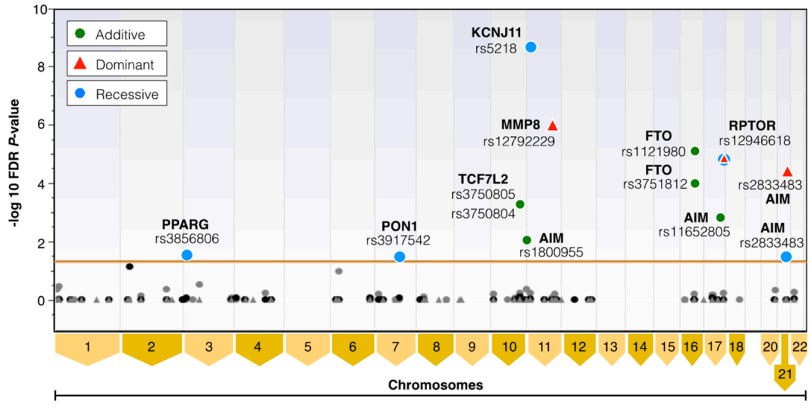


Figure 3. Manhattan plot results of the association study with MLMM in cases and controls age 45-79 years - log of FDR adjusted P-values. The horizontal orange line marks the significant p-value cut-off.

Of the twelve highly BC-associated SNPs, ten are located in non-coding regions (Table 5) involved in the decay of mRNA transcripts, enhancers, promoters, and regulatory regions (nonsense-mediated decay transcript variants (NMDTV), promoter flanking region (PFR), and regulatory region variants (RRV)).

 doi:10.20944/preprints202501.1758.v1

Table 5. Nonsense-mediated decay transcript variants (NMDTV), promoter flanking region (PFR), and regulatory region variants (RRV) are associated with breast cancer.

Chr[1]	Genes	Marker	HGVS nomenclature	Consequence details[2]
3	*PPARG*	rs3856806	NM_015869.4(PPARG):c.1431C>T (p.His477=) NM_138712.3:c.1347C>T	Synonymous variant; 3 prime UTR variant; NMDTV
7	*PON1*	rs3917542	NC_000007.14:g.95307380C>T NM_000446.5:c.698+631G>A	RRV; NMDTV
10	*TCF7L2*	rs3750804	NC_000010.11:g.113074091C>T	PFR, RRV
		rs3750805	NC_000010.10:g.114847143A>T	PFR; RRV
11	*SCT;DEAF1; DRD4*	rs1800955	NC_000011.10:g.636784T>C	RRV; PFR
	MMP8	rs12792229	NC_000011.10:g.102718512G>T XM_005271556.1:c.617C>A XP_011541137.1:p.Ser206Tyr	Missense variant; SIFT: deleterious; PolyPhen: possibly damaging; NMDTV; Downstream gene variant
	KCNJ11; ABCC8; NCR3LG1	rs5218	NC_000011.10:g.17387522G>A NP_001159762.1:p.Ala103=	Synonymous variant, Upstream gene variant Downstream gene variant, Regulatory region variant; a CTCF binding site variant
16	*FTO*	rs1121980	NC_000016.10:g.53775335G>A NM_001080432.2:c.46-34805G>A	RRV; NMDTV
		rs3751812	NC_000016.10:g.53784548G>T NM_001080432.2:c.46-25592G>T	RRV
17	*AMZ2P1; GNA13*	rs11652805	NC_000017.11:g.64991033C>T	RRV
	RPTOR	rs12946618	NC_000017.11:g.80603368G>A NM_001163034.1:c.163-22323G>A	NMDTV; Upstream gene variant
21	*SCAF4*	rs2833483	NC_000021.9:g.31703091T>C NM_001145444.1:c.276+674A>G	Upstream gene variant

NMDTV, Nonsense Mediated Decay transcript variant; UTR, untranslated region; PFR, promoter flanking region; RRV, regulatory region variant; SIFT, Sorting Tolerant From Intolerant; PolyPhen, Polymorphism Phenotyping; HGVS, Human Genome Variation Society Nomenclature.

Variants in the MMP8 (rs12792229) and the PPARG (rs3856806) genes have a dual effect in coding and non-coding regions about the alternative transcripts of the genes [28] (Table 6).

Table 6. Exome variants associated with breast cancer.

Marker	Chr[1]	Genes	HGVS nomenclature	Consequence details[2]
rs3856806	3	PPARG	NM_015869.4(PPARG):c.1431C>T (p.His477=) NM_138712.3:c.1347C>T	Synonymous variant 3 prime UTR variant, NMDTV

 doi:10.20944/preprints202501.1758.v1

rs12792229	11	MMP8	NC_000011.10:g.102718512G>T XM_005271556.1:c.617C>A XP_011541137.1:p.Ser206Tyr	Missense variant. SIFT: deleterious PolyPhen: possibly damaging. NMDTV: Downstream gene variant
rs5218	11	KCNJ11; ABCC8; NCR3LG1	NC_000011.10:g.17387522G>A NP_001159762.1:p.Ala103=	Synonymous variant, Upstream gene variant Downstream gene variant, Regulatory region variant; a CTCF binding site variant

[1]Chromosome. [2]Consequence details extracted from ENSEMBL.

3.3. Linkage Disequilibrium

Four SNPs were found to have LD: two for TCF7L2 (rs3750804 and 3750805) (CHM, r2=0.4359, D'=1) and two for FTO (rs1121980 and rs375181) (CHM, r2=0.5967, D'= 0.9598). The full LD data for PON1, TCF7L2, KCNJ11, MMP8, FTO, RPTOR, and SCAF4 are reported in Table 7.

Table 7. LD for the statistically significant associated SNPs.

Chr	Genes	Marker 1	Marker 2	CHM - R Squared	CHM - D Prime
7	PON1	rs3917542[1]	rs662	0.290	0.866
10	TCF7L2	rs3750804[1]	rs3750805[1]	0.436	1.000
11	KCNJ11;ABCC8	rs5210	rs5218[1]	0.363	0.960
		rs5218[1]	rs5219	0.074	0.800
		rs5210	rs5219	0.274	0.963
	MMP8;MMP27	rs35866072	rs12792229[1]	0.000	1.000
16	FTO	rs1121980[1]	rs3751812[1]	0.597	0.960
		rs8050136	rs3751812[1]	0.825	0.970
		rs17817449	rs3751812[1]	0.869	0.970
		rs1121980[1]	rs8050136	0.746	1.000
		rs1121980[1]	rs17817449	0.696	0.995
17	RPTOR	rs12946618[1]	rs12946115	0.790	0.900
		rs12946618[1]	rs12950541	0.727	0.853
21	SCAF4	rs2833479	rs2833483[1]	0.898	0.962

[1]MLMM significant associated SNPs.

4. Discussion

The current study investigated the association between breast cancer and age, estrogen exposure, and biochemical and clinical parameters in a Case and Control study. Although there were no significant differences in age at menarche, oral contraceptive use, and children number in the two groups, several factors emerged as potential contributors to breast cancer susceptibility.

 doi:10.20944/preprints202501.1758.v1

One of the most notable findings was the difference in BMI distribution. While the cohort as a whole presented a predominance of overweight status, this trend was more pronounced in the breast cancer group, revealing a statistically significant association between increased BMI and breast cancer (p=0.011). Our results are consistent with the consensus suggesting that elevated BMI is associated with increased estrogen levels due to adipose tissue functioning as an additional source of estrogen production. This link becomes especially pertinent in postmenopausal women, where excess adipose tissue can lead to prolonged estrogen exposure, thus contributing to cancer development.

Mexican demographics directly impact the BC risk (age and BMI), as population prediction in the following decades tends to increase in the population between 50-60 years [29]; also the fact that Mexico is considered one of the countries with high overweight and obesity rates (World Obesity Federation, World Obesity Atlas 2023. https://data.worldobesity.org/publications/?cat=19), especially in women with more than 70% having overweight or obesity [3], reflects a critical health issue that impacts in BC incidence. Other risk-related factors not influenced by ethnicity include menarche, menopause status, age of first pregnancy, and number of pregnancies [5].

In the Northeast region of Mexico (Nuevo León, Tamaulipas, and Coahuila states), the distribution of breast cancer subtypes is generally consistent with broader national trends. Still, there can be regional variations could be considered. Based on available studies and data [30–33], the approximate distribution of breast cancer subtypes in this region is as follows: Triple-Negative Breast Cancer (TNBC) ~10-20% frequency (present study 19.57%); this subtype is prevalent in the Northeast region of Mexico, similar to other parts of Mexico, and is particularly significant among younger women. HER2-Enriched subtype ~10-15%, the prevalence of this subtype is also notable in the Northeast region (present study 5.43%). This distribution of breast cancer subtype statistics is a general overview, but precise distribution may vary slightly depending on specific studies and population groups within the region. Efforts for local cancer registries, hospital data, and regional epidemiological studies would offer the most accurate and detailed information for the Northeast of Mexico.

Previous genome-wide association studies (GWAS) have identified SNPs associated with the risk of breast cancer and obesity [13,34]. Our study analyzed 121 SNPs and found 12 highly associated with an increased risk of developing BC. Statistical analysis for SNP classification was designed to minimize false positives and eliminate confusion factors, considering population, kinship, and cryptic relation.

One of the common problems in genome-wide studies is that the population structure needs to be considered, especially in ethnically diverse and complex populations [35,36]. Including AIM in this study ensures that a mestizo population like ours is well-represented and that population bias is excluded. The present study included 36 AIM, and there was no evidence of substructure in the populations of the cases and control groups.

Several projects and articles were reported in GWAS studies of breast cancer for women of European and African ancestry [37–39]. However, a low proportion of projects and articles reported breast cancer GWAS in other populations.

We found an association of 12 SNPs with BC; two are located in coding and ten in non-coding regions, as expected, because most SNPs and CNAs (Copy Number Alterations) reside in non-coding sequences regulating oncogenes and tumor-suppressors in the carcinogenesis process [40]. Products of the associated genes in this work involve various functions such as transcriptional factors and regulators, lipid metabolism, remodeling and modifying DNA, cell growth, RNA polymerase modification, ionic channels, and addiction-related receptors. The effects of these SNPs are listed in Tables 5 and 6.

To address the associated variants in the present study, we individually discuss each variant and corresponding gene and their implications for breast cancer.

 doi:10.20944/preprints202501.1758.v1

4.1. The rs11652805 (AMZ2P1; GNA13 Intergenic Variant, Regulatory Region Variant)

The intergenic variant located between AMZ2P1 and the GNA13 and contiguous to a regulatory region (enhancer) is part of the sequence of a transcription factor binding site (ENSEMBL database) [13]. This regulatory variant could be involved in the transcriptional regulation of AMZ2P1 and GNA13; evidence supports the gene expression correlations between rs11652805 and AMZ2P1 (ENSG00000214174) and GNA13 (ENSG00000120063) transcript expression [13].

4.1.1. AMZ2P1 (Archaelysin Family Metallopeptidase 2 Pseudogene 1)

The AMZ2P1 (Archaelysin Family Metallopeptidase 2 Pseudogene 1) is classified as a pseudogene. However, ENSEMBL reported an unprocessed mRNA transcript and 19 lncRNA biotypes could be functionally relevant [13]. The lncRNA ENST00000397713.5 related to rs11652805 of AMZ2P1 was analyzed for Protein–RNA interaction predictions in the RNAct database (http://rnact.crg.eu/) [41], as a result (*in silico*) predicted interactions by protein–RNA interaction prediction algorithm with 100 protein coded genes. The primary function classification (Toppgene, https://toppgene.cchmc.org/) of the 100 genes that interact is Molecular Function: ATP hydrolysis activity (17/100), ATP-dependent activity (20/100), and Biological Process; cell cycle process (24/100), chromatin remodeling (16/100) [42].

GNA13 (Guanine nucleotide binding protein alpha 13) is predicted to enable D5 dopamine receptor binding activity, G-protein beta/gamma-subunit complex binding activity, and GTPase activity [43]. GNA13 promotes tumor cell invasion and metastasis by activating the RhoA/ROCK signaling pathway [44–46] and the GPCR, breast cancer regulation by Stathmin1, and Estrogen Pathway [43]. GNA13 expression in breast cancer cells is regulated by post-transcriptional mechanisms involving miR-31 [47]. MiR-31 partially controls breast cancer cell invasion by targeting GNA13. The loss of miR-31 and increased GNA13 expression may serve as biomarkers for breast cancer progression [47]. Increased GNA13 expression has been observed in metastatic breast cancer cells. In breast cancer, enhanced GNA13 signaling represses KLK gene expression [48]. We found the association of the rs11652805 and BC; this SNP has no previous record of association with cancer risk. However, the proximity localization to a regulatory region may be implicated in regulating gene expression of contiguous genes, probably AMZ2P1 and GNA13, as context; the chromatin stretch enhancer states can drive the cell-specific gene regulation in human disease risk variants [49,50].

The GNA13 RNA transcript ENST00000439174.6 was analyzed in the RNAct database [41], as a result (*in silico*) predicted interactions by protein–RNA interaction prediction algorithm with 100 protein-coded genes. The primary function classification (Toppgene) of the 100 genes that interact is Molecular Function: chromatin binding (16/100), histone binding (8/100), and Biological Process; chromatin organization (20/100), protein-DNA complex organization (20/100), and chromatin remodeling (16/100) [42]

The GNA13 and AMZ2P1 RNA transcripts interact with protein-coded genes related to chromatin remodeling.

4.1.2. DRD4 (Dopamine Receptor D4).

The gene encodes a D4 subtype dopamine receptor coupled to G protein. *DRD4 is* responsible for neuronal signaling in the brain's mesolimbic system, inhibiting adenylyl cyclase [43]. In one study, DRD4 was overexpressed in BC cells compared to normal tissue. In the same study, MCF-7 and MDA-MB-231 cells, only DRD1 and DRD4 proteins were detected [51]. With decades of evidence suggesting a link between dopamine (DA) receptors and cancer, the DA pathway has recently emerged as a potential target in antitumor therapies [52]. The dopamine receptors (DAR) are associated with tumor cell death, proliferation, invasion, and migration. The DAR could regulate several ways of tumor cell death (apoptosis, autophagy-induced death, and ferroptosis) [53]. We found an association between rs1800955 and BC risk not previously reported; based on the relation of DRD4 with cancer biology, this DMA variant rs1800955 association is probably not casual and

 doi:10.20944/preprints202501.1758.v1

maybe could be helpful to identify breast cancer patients and pharmacology strategies recommendations in further studies. Another possible impact in BC is that the rs1800955 SNP is an intergenic flanking promotor variant in a transcription factor binding site at the 5′ of the DRD4 gene [13]. Genes regulating DRD4 transcript expression are involved in the cell cycle process (25/100) and chromatin remodeling (14/100) [42].

4.2. FTO (Fat Mass and Obesity-Associated Protein)

FTO gene is part of the AlkB-related non-heme iron and 2-oxoglutarate-dependent oxygenase superfamily. Its gene product has RNA demethylase activity and regulates adipose and energy homeostasis [43]. FTO is related to obesity and diabetes predisposition [54]. Some studies have related FTO SNPs with various types of cancer [55], including breast cancer, where the overexpression of FTO promotes cell glycolysis and impacts the PI3K/Akt signaling pathway [56]. There has been multiple SNP related to cancer (rs9939609, rs8050136, rs1477196, rs6499640, rs1121980, rs17817449, rs11075995, rs8047395, and rs7206790, rs3751812)[55]. The most studied is rs9939609, associated with lung cancer, renal cancer, breast cancer, prostate cancer, pancreatic cancer, endometrial [57–61]. In a Chinese cohort of women with BMI<24 kg/m2, rs1477196 AA genotype and TAC haplotype for rs9939609, rs1477196, and rs1121980 have a protective effect for BC, and SNP rs16953002 AA genotype has an increased BC risk [62]. A meta-analysis found an association between rs11075995 and ER-negative BC in the European population and rs17817449 with a generally increased risk for BC in the African population [63]. In our study, we found rs3751812 and rs1121980 associated with BC as seen in other populations, and there was no association for the rs10938397, rs17817449, rs2815752, and rs8050136 in our population regardless of the BMI. The rs1121980 and rs3751812 were classified as regulatory region variants, RRV, each variant was located in a different regulatory feature, the rs1121980 located in the Regulatory Feature ENSR00001001829, and the rs3751812 in the Regulatory Feature ENSR00000537825 (Ensembl database) [13]. The DNA variants linked to two Regulatory regions were gene expression correlations (Ensembl database) to genes functionally implicated in the mRNA N6-methyladenosine dioxygenase activity (Toppgene database) [42], which plays an important role in mediating fundamental pathological and physiological metabolic processes, processing, translation, and stability of RNA [64], thromboxane A2 receptor binding, thromboxane A2 (TXA2) is a potent lipid mediator released by platelets and inflammatory cells and is capable of inducing vasoconstriction [65], tRNA demethylase activity. Additionally, the rs1121980 is a Nonsense Mediated Decay transcript variant that NMD surveys newly synthesized mRNAs and degrades those that harbor a premature termination codon (PTC), thereby preventing the production of truncated proteins that could result in disease in humans [66].

4.3. KCNJ11 (Potassium Voltage-Gated Channel Subfamily J Member)

This gene is a member of the potassium channel gene family, controlled by G proteins. KCNJ11 is associated with diabetes and cardiac conduction defects [43]. KCNJ11 gene has more than 200 polymorphisms; six of them, rs5215, rs5210, rs5218, rs5219, rs886288, and rs2285676, are related to diabetes [67]. The rs5219 has been previously associated with colon cancer risk [68][69]. The high KCNJ11 gene expression is associated with favorable prognostic and life expectancy in renal cancer [70] (data available from The Human Protein Atlas, version 23.0). Our work found an association of rs5218 (A190A) and BC. This SNP has been associated with Type II Diabetes Mellitus [71], but no previous evidence exists with BC risk. The rs5218 is a synonymous variant for KCNJ11 and the regulatory region variant, a CTCF binding site variant for KCNJ11, ABCC8, and NCR3LG1 (ENSR00000263063) [13]. NCR3LG1 is tumor overexpressed and is associated with fatal disease progression of various cancers [72]. ABCC8 (SUR1) sulfonylurea receptor 1 is a tumor-enhancer in non-small cell lung carcinoma (NSCLC) [73].

The rs5218 SNP is a regulatory region variant in a transcription factor binding site (Regulatory feature ENSR00000263063) [13]. The rs1800955 has a gene expression correlation (p-value < 0.001) in the ENSEMBL database (https://www.ensembl.org) with 25 genes in different tissues [13]. The

 doi:10.20944/preprints202501.1758.v1

analysis of the 25 genes expressed in Toppgene results in a molecular function identity. The study of the 25 genes expressed in Toppgene results in a molecular function identity in a voltage-gated potassium channel activity (KCNC1, ABCC8, and KCNJ11) and a WikiPathways VITAMIN B12 METABOLISM, and FOLATE METABOLISM [42].

4.4. MMP8 (Matrix Metalloproteinase 8)

MMPs are members of a large multigene family of zinc-dependent endopeptidases involved in remodeling the extracellular matrix protein degradation of proteins to allow ductal progression across the basement membrane [74–76]. MMP8 is also known as collagenase-2 or neutrophil collagenase. Tumorigenic and antitumorigenic properties have been attributed to MMP8 activity in cases of head and neck squamous carcinoma cells [77]. In BC, MMP8 gene variation may influence prognosis and could have an inhibitory effect on cancer metastasis; minor allele (T) of the promoter SNP (rs11225395) has been linked to a better prognosis, such as reduced susceptibility to lymph node metastasis [78], reduced cancer relapse[79], and higher survival [79]. In this work, we found the rs12792229 of MMP8 associated with BC. There are no reports of clinical significance about this SNP; further studies are required to understand its role in BC development.

4.5. PON1 (Human Serum Paraoxonase 1 Enzyme)

This gene is a member of the paraoxonase family of enzymes with lactonase and ester hydrolase activity. PON1 is involved in HDL regulation [43,80,81]. PON1 SNPs have been related to multiple inflammatory diseases, such as atherosclerosis, diabetes, and some cancer types, including lung, multiple myeloma, papillary thyroid cancer, prostate, breast, and ovarian [82]. PON1 rs854555 was also related to an increased risk of BC in U.S. postmenopausal women [83]. Another study found that rs854560, rs662, rs705379, and PON1_304A/G were associated with an increased risk of developing cancer in Caucasian and Asian populations [84]. We included rs3917542, rs854555, and rs662, but only the rs3917542 was associated with BC. PON1 rs3917542 has been associated with cardiovascular diseases but not cancer [85]. PON1 polymorphism association studies in cancer differ among populations; there are no studies on the Mexican population [82]. Increased expression of PON1 was correlated with higher survival probability in liver cancer [70] (data available from The Human Protein Atlas, version 23.0).

The rs3917542 has a gene expression correlation (p-value < 0.05) in the ENSEMBL database (https://www.ensembl.org) with 14 genes in different tissues, breast mammary tissue included [13]. The analysis of the 14 genes expressed in Toppgene results in a molecular function identity. The Toppgene results in a molecular function identity in arylesterase, and lactonohydrolase activity [42].

4.6. PPARG (Peroxisome Proliferator-Activated Receptor Gamma)

The PPARG is a nuclear hormone receptor of the family of transcriptional regulators activated by ligands and involved in regulating adipogenesis and glucose [86]. The PPARG high expression in renal and urothelial cancer was associated with favorable life expectancy and unfavorable life expectancy in liver cancer (https://www.proteinatlas.org/ENSG00000132170-PPARG/pathology). PPARG SNPs have been associated with cancer risk, in particular for BC, ovarian carcinoma, follicular lymphoma, and colorectal cancer [86–88]. We analyzed rs3856806 and rs1801282 for this gene and found that rs3856806 was associated with BC development. Our results are concordant with another study in the Turkish population, where they found that both SNPs were in linkage disequilibrium and related to increased BC risk [89].

Conversely, a meta-analysis study found that rs3856806 was not associated with BC, and rs1801282 has a protective factor in Caucasian women [90]. PPARG expression as a prognostic effect depends on metastasis localization in advanced colorectal cancer patients. Increased expression of PPARG high expression increased the malignancy-associated traits such as proliferation in colorectal cancer cell lines and increased sensitivity towards the chemotherapeutic agent 5-FU [91]. The combination therapy of PPARG agonists and 5-FU-based chemotherapy is a promising strategy [91].

 doi:10.20944/preprints202501.1758.v1

The rs3856806 has a gene expression correlation (p-value < 0.05) in the ENSEMBL database (https://www.ensembl.org) with 15 genes in different tissues, breast mammary tissue included [13]. The analysis of the 15 genes expressed in Toppgene does not find statistically significant results [42].

4.7. RPTOR (Regulatory Associated Protein of MTOR Complex 1, RPTOR, Also Named RAPTOR)

RPTOR encodes a protein part of the PI3K–AKT–mTOR pathway, regulating negatively the mTOR kinase [43]. The PI3K pathway is frequently dysregulated in cancer, with multiple targeted therapies available and more in development [92]. The mammalian target of rapamycin complex 1 (mTORC1) is evolutionally conserved and commonly activated in various tumors [93]. Akt/Raptor signaling upregulation is associated with rapamycin resistance of breast cancer cells [94]. Hypomethylation of CpG sites in RPTOR in blood had been associated with BC, suggesting that it might serve as a biomarker [95]. RAPTOR gene polymorphisms (rs11653499 and rs7212142) were significantly related to the risk of urothelial cancer [96]. Cheng, T.Y et al. report that rs9900506 and rs3817293 have been linked to a protective effect for BC in European-American and African-American women, respectively [97]. We found rs12946618 associated with BC. To our knowledge, no previous reports of this SNP and BC risk exist. The rs12946618 is a regulatory region variant (Regulatory feature ENSR00001345149) correlated with gene expression of 33 genes in different tissues (p-value < 0.05) in the ENSEMBL database [13]. The Toppgene analysis of the 33 genes expressed correlated results in a molecular function in a single-stranded RNA binding, methylated histone binding, and cellular component PRC1 complex, a protein complex that includes a ubiquitin-protein ligase and enables ubiquitin protein ligase activity (EMBL-EBI database, GO:0000151), the PcG protein complex; transcriptional repressors that regulate several crucial developmental and physiological processes [98]. The related pathways are Reactome regulation of PTEN gene transcription and Reactome sumoylation of DNA methylation proteins. In the present study, we find LD with the rs12946115 and rs12950541, located in the RPTOR gene. VEP identified the rs12946618, rs12946115, and rs12950541 as Nonsense-mediated decay transcript variants (NMDTV) in the ENST00000574767.5 (name RPTOR-208) transcript, which codifies to a peptide of 95 amino acids; the canonical protein has 1335 amino acids (Transcript ENST00000306801.8; name RPTOR-201). The RPTOR rs12946618, rs12946115, and rs12950541 were classified as NMDTV; these variants can regulate RPTOR gene expression. RPTOR rs12950541 is an SNV (single nucleotide variant) intronic variant that impacts alternative splicing, interfering with splice site recognition, affecting the synthesis of the protein, and thus interfering with its function. RPTOR is involved in nutrient signaling, mitochondrial oxygen consumption, and oxidative capacity [99].

RPTOR is involved in insulin receptor signaling and the AMPK, mTORC1, and metformin pathways. Complex 1 of mTOR (mTORC1) controls cell proliferation, growth, and metabolism in various cell types through a complex signaling network. To compensate for hyperglycemia, pancreatic β cells become hyperplasic hypertrophic, insulin secretion is increased [100], and finally, β cells begin to fail, and hypoinsulinemia appears. During T2D progression, complex 1 of the mammalian target of rapamycin (mTORC1) is hyperactivated, with a benefit due to an increase in β cell proliferation. Nevertheless, the chronic mTORC1 hyperactivation increases endoplasmic reticulum (ER) and Golgi apparatus stress and β cell failure [100,101].

With the in silico analysis of the RPTOR the rs12946618, rs12946115, and rs12950541 variants, we found that the degradation of mRNA RPTOR via the NMD process, and as a consequence, a deficient expression level of RPTOR protein in homozygous and heterozygous carriers, can be the factor associated with T2D development because RPTOR is required for maintaining a β-cell identity as well as repressing β-cell to α-cell transdifferentiation [102]. Furthermore, in a high caloric diet, mainly a high carbohydrate diet, mTORC1 is hyperactivated, and the rs12950541 variant can play a role in insulin resistance, obesity, and T2D development by loss of maintaining β-cell identity [102].

 doi:10.20944/preprints202501.1758.v1

4.8. KCNJ11 Pathway and RPTOR

Raptor/mTORC1 is essential for b-cell function, mTORC1 and mTORC2 regulate insulin secretion through Akt in INS-1 cells [103].

4.9. SCAF4 (SR-Related c-Terminal Domain-Associated Factor 4)

SCAF4 encodes a member of the arginine/serine-rich splicing factor family. It is predicted to bind to the serine-phosphorylated c-terminal domain of POLR2A during RNA transcription and to be involved in RNA splicing [43]. Increased expression of SCAF4 is related to a poor survival probability in BC (https://www.proteinatlas.org/ENSG00000156304-SCAF4/pathology/breast+cancer). In another study, a subtype of signature composed of ELOA and SCAF4 was identified, and a subtype diagnostic and prognostic model was constructed for therapeutic targeting in esophageal cancer [104]. We found rs2833483 associated with BC. Based on recent studies, SCAF4 is a target and biomarker of interest in cancer. The rs2833483 has a gene expression correlation (p-value < 0.05) in the ENSEMBL database (https://www.ensembl.org) with 17 genes in different tissues [13], and a negative gene expression correlation with ENSG00000273091 (lincRNA gene), the ENSG00000273091 has a interaction with AEBP2 protein in breast mammary tissue [41], AEBP2 protein acts as an accessory subunit for the core PRC2 (Polycomb repressive complex 2), which mediates histone H3K27 (H3K27me3) trimethylation on chromatin leading to transcriptional repression [43].

4.10. TCF7L2 (Transcription Factor 7 Like 2)

The TCF7L2 gene is involved in the Wnt signaling pathway. TCF7L2 is implicated in angiogenesis, cell proliferation, apoptosis, transcription regulation, and glucose homeostasis [13,43]. TCF7L2 has been associated with type 2 diabetes (T2D), chronic renal disease [105], and cancer, including BC and hepatocarcinoma [43,106,107]. In 2012, Connor et al. found in a Hispanic cohort, including Mexican women with BC, that rs7903146, rs3750805, rs7900150, and rs1225404 were associated with a higher risk of BC regarding ancestry [108]. Another study in the Egyptian population did not find an association between BC and rs12255372 [109]. Also, the TT genotype of the rs3750804 increased the BC risk in women with a history of diabetes but did not demonstrate significant interactions by ethnicity or genetic admixture [108]. We included in our study the TCF7L2 rs10885390, rs11196175, rs7903146, rs10885406, rs7900150, rs12255372, rs3750804, rs3750805, rs290487 and rs1225404, and we found that the rs3750804 and rs3750805 were associated to BC for ages 45 to 79. The DNA variants rs3750804 and rs3750805 are classified as promoter flanking region (PFR) and regulatory region variants (RRV). rs3750804 and rs3750805 could be an impact in TCF7L2 expression.

GWAS studies should include population structure analysis with geographical data in mestizo populations. Mexican native populations are ethnically and genetically diverse [36]. The DNA variants reported in this study could be validated in further studies and can be useful as biomarkers for response to chemotherapeutic treatment in BC patients. Mexico is a developing country where genetic tests are not routine, and SNP detection could help to support the potential benefits of personalized medicine for the Mexican population. Our results reinforce that individual studies by populations should be done as there are ethnic differences in the expression and effect of these SNPs.

However, the association between SNP alleles and phenotypic impact can be confounding due to linkage disequilibrium, which segregates driver DNA variants with passenger DNA variants (passively inherited but regulatory neutral) within populations [96]. Because of this, it is necessary to analyze the SNPs and genes implicated and screen other gene variants that could participate in the disease's manifestation. Consequently, GWAS should be validated by extensive and meticulous analysis to determine which of the associated SNPs are the actual driver factors of the disease [40].

 doi:10.20944/preprints202501.1758.v1

4.11. GNA13 and RAPTOR, TCF7L2, SCAF4, KCNJ11, and FTO.

With data from 60 Invasive Breast Carcinoma cell lines of the DepMap portal (Expression Public 24Q2), a linear regression was obtained for RPTOR and GNA13 (p-value 7.81E-05) [110]. Similar statistically significant linear regression results were obtained for GNA13 compared to TCF7L2 (p-value 7.98E-04), SCAF4 (p-value 1.22E-3), KCNJ11 (p-value 1.22E-3), FTO (p-value 2.27E-05): However, no significant linear regression with the other genes was reported in this study.

4.12. GNA13 (Gα13) and RPTOR (mTORC1) Pathway Interaction

Germinal center B cell-like diffuse large B-cell lymphoma (GCB-DLBCL) is characterized by the downregulation of PTEN and the activation of the PI3K signaling pathway. When PTEN is lost or repressed, phosphatidylinositol-3-phosphate accumulates, activating AKT and mTORC1, which promotes cell survival, proliferation, and growth. GCB-DLBCL is also associated with the loss of S1PR2 (sphingosine-1-phosphate receptor-2) and Gα13 (G-protein alpha 13, GNA13 gene) signaling, negatively modulating GC B-cell migration and PI3K signaling. The constant activation of the PI3K/AKT/mTORC1 pathway in GCB-DLBCL is due to the loss or inactivation of PTEN and S1PR2-Gα13 signaling, which usually suppress cell survival, proliferation, and growth [111].

5. Conclusions

We found the association of 12 genetic variants located in the DRD4, PPARG, PON1, TCF7L2, MMP8, FTO, RPTOR, and SCAF4 genes and the non-coding regions in SCT, DEAF1, KCNJ11, ABCC8, AMZ2P1, and GNA13, with risk of BC, menopause, and BMI in Northeast Mexico. Most have been previously associated with DM, obesity, insulin resistance, inflammation, and extracellular matrix remodeling. Of the 12 variants, three are located in the exome, and 10 are classified as non-coding, involved in the accelerated decay of the transcript, enhancer, promoter, and regulatory regions (Nonsense-mediated decay transcript variants (NMDTV), promoter flanking region (PFR) and regulatory region variants (RRV). DNA Variants in MMP8 (rs12792229) and PPARG (rs3856806) have a dual effect on exome and regulatory regions.

This study's findings indicate a significant association between 12 specific SNPs and BC risk in the studied population. Many of these SNPs were found in regulatory regions, highlighting their potential impact on gene expression and regulation, which may contribute to carcinogenesis.

The characteristics of the Mexican population, including age distribution and high prevalence of obesity, are influential factors that may contribute to an increased risk of BC, particularly in women aged 45-64 years.

Our analysis of 121 SNPs identified 12 SNPs strongly associated with BC risk. These variants are located in both the coding and non-coding regions of genes such as GNA13, DRD4, FTO, KCNJ11, MMP8, PON1, PPARG, RPTOR, SCAF4, and TCF7L2. Many of these SNPs were found in regulatory regions, highlighting their potential impact on gene expression and regulation, which may contribute to carcinogenesis.

Including ancestry informative markers (AIMs) in our study permits verifying that no population stratification bias was absent, ensuring that the mestizo population is well-represented and allowing for more accurate identification of SNPs relevant to the Mexican population. SNPs in regulatory regions, such as enhancer sites, suggest that these variants influence gene expression, potentially affecting pathways involved in cell cycle regulation, transcriptional activity, and oncogenic signaling.

Notably, the study underscores the need for additional research to validate these findings and better understand the functional impact of the identified SNPs on BC development and progression. The results also highlight the relevance of personalized medicine approaches tailored to the genetic profile of Mexican and other underrepresented populations, which can enhance the effectiveness of treatment and preventive strategies.

 doi:10.20944/preprints202501.1758.v1

This work contributes to the growing body of evidence that supports the inclusion of diverse populations in genomic studies to uncover unique genetic risk factors and improve the applicability of genetic tests for cancer risk assessment.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: title; Table S1: Genetic markers list. Table S2: Openarray TaqMan probes. Table S3: Genotypes and Clinical Data. Table S4: Full MLMM association results under additive genetic model. Table S5: Full MLMM association results under dominant genetic model. Table S6 Full MLMM association results under recessive genetic model.

Author Contributions: Conceptualization, C.N.S.D. and H.L.G.B.; methodology, C.N.S.D. and H.L.G.B.; software, H.L.G.B. and R.M.C.F.; validation, R.M.C.F., H.L.G.B., M.L.G.R. and C.N.S.D.; formal analysis, C.N.S.D. and H.L.G.B.; investigation, C.N.S.D. and H.L.G.B.; resources, O.V.G., C.N.S.D. and H.L.G.B.; data curation, R.M.C.F., H.L.G.B., M.L.G.R.; writing—original draft preparation, C.N.S.D., H.L.G.B., R.M.C.F., D.C.P.I., M.L.G.R., C.H.B.F., V.M.S.T., M.G.E., R.P.R., O.V.G.; writing—review and editing, C.N.S.D., H.L.G.B.; supervision, C.N.S.D., and H.L.G.B.; project administration, C.N.S.D., and H.L.G.B.; funding acquisition, C.N.S.D., and O.V.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by SEP-PRODEP, and UANL-PAICYT awarded to C.N.S.D.

Institutional Review Board Statement: The study was conducted following the Declaration of Helsinki, and the protocol was approved by the Institutional Ethics Committee of the University Hospital, with the registration codes BI10–002 and GN13-001. All subjects (patients and controls) were invited to participate in the research project. An interview was performed, and once the patients agreed to participate, they signed an informed consent letter. Afterward, clinical and epidemiological information was collected, and blood samples were taken.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.
Data Availability Statement: We encourage all authors of articles published in MDPI journals to share their research data. In this section, please provide details regarding where data supporting reported results can be found, including links to publicly archived datasets analyzed or generated during the study. Where no new data were created, or where data is unavailable due to privacy or ethical restrictions, a statement is still required. Suggested Data Availability Statements are available in section "MDPI Research Data Policies" at https://www.mdpi.com/ethics.

Acknowledgments: We are grateful for the support provided by the SEP-PRODEP, and UANL-PAICYT.

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

References

1. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J Clinicians* **2024**, *74*, 229–263, doi:10.3322/caac.21834.
2. Ferlay, J.; Ervik, M.; Lam, F.; Laversanne, M.; Colombet, M.; Mery, L.; Piñeros, M.; Znaor, A.; Soerjomataram, I.; Bray, F. *Global Cancer Observatory: Cancer Today.*; Global Cancer Observatory: Cancer Today; IARC, WHO: Lyon, France, 2022;
3. Shveid Gerson, D.; Gerson-Cwilich, R.; Lara Torres, C.O.; Chousleb De Kalach, A.; Ventura Gallegos, J.L.; Badillo-Garcia, L.E.; Bargalló Rocha, J.E.; Maffuz-Aziz, A.; Sánchez Forgach, E.R.; Castorena Roji, G.; et al. Establishment of Triple-Negative Breast Cancer Cells Based on BMI: A Novel Model in the Correlation between Obesity and Breast Cancer. *Front. Oncol.* **2022**, *12*, 988968, doi:10.3389/fonc.2022.988968.
4. Chen, M.J.; Wu, W.Y.; Yen, A.M.; Fann, J.C.; Chen, S.L.; Chiu, S.Y.; Chen, H.H.; Chiou, S.T. Body Mass Index and Breast Cancer: Analysis of a Nation-Wide Population-Based Prospective Cohort Study on 1 393 985 Taiwanese Women. *Int J Obes (Lond)* **2016**, *40*, 524–530, doi:10.1038/ijo.2015.205.
5. Kawai, M.; Tomotaki, A.; Miyata, H.; Iwamoto, T.; Niikura, N.; Anan, K.; Hayashi, N.; Aogi, K.; Ishida, T.; Masuoka, H.; et al. Body Mass Index and Survival after Diagnosis of Invasive Breast Cancer: A Study Based on the Japanese National Clinical Database-Breast Cancer Registry. *Cancer Med* **2016**, *5*, 1328–1340, doi:10.1002/cam4.678.
6. Samavat, H.; Kurzer, M.S. Estrogen Metabolism and Breast Cancer. *Cancer Letters* **2015**, *356*, 231–243, doi:10.1016/j.canlet.2014.04.018.
7. Zhang, X.; Eliassen, A.H.; Tamimi, R.M.; Hazra, A.; Beck, A.H.; Brown, M.; Collins, L.C.; Rosner, B.;

Hankinson, S.E. Adult Body Size and Physical Activity in Relation to Risk of Breast Cancer According to Tumor Androgen Receptor Status. *Cancer Epidemiology Biomarkers & Prevention* **2015**, *24*, 962–968, doi:10.1158/1055-9965.EPI-14-1429.

8. Surakasula, A.; Nagarjunapu, G.; Raghavaiah, K. A Comparative Study of Pre- and Post-Menopausal Breast Cancer: Risk Factors, Presentation, Characteristics and Management. *J Res Pharm Pract* **2014**, *3*, 12, doi:10.4103/2279-042X.132704.
9. Chlebowski, R.T.; Anderson, G.L. Menopausal Hormone Therapy and Breast Cancer Mortality: Clinical Implications. *Therapeutic Advances in Drug Safety* **2015**, *6*, 45–56, doi:10.1177/2042098614568300.
10. Millikan, R.C.; Newman, B.; Tse, C.-K.; Moorman, P.G.; Conway, K.; Smith, L.V.; Labbok, M.H.; Geradts, J.; Bensen, J.T.; Jackson, S.; et al. Epidemiology of Basal-like Breast Cancer. *Breast Cancer Res Treat* **2008**, *109*, 123–139, doi:10.1007/s10549-007-9632-6.
11. Pomerantz, M.M.; Freedman, M.L. The Genetics of Cancer Risk. *The Cancer Journal* **2011**, *17*, 416–422, doi:10.1097/PPO.0b013e31823e5387.
12. 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 Research* **2020**, *48*, D835–D844, doi:10.1093/nar/gkz972.
13. Harrison, P.W.; Amode, M.R.; Austine-Orimoloye, O.; Azov, A.G.; Barba, M.; Barnes, I.; Becker, A.; Bennett, R.; Berry, A.; Bhai, J.; et al. Ensembl 2024. *Nucleic Acids Research* **2024**, *52*, D891–D899, doi:10.1093/nar/gkad1049.
14. Shieh, Y.; Fejerman, L.; Lott, P.C.; Marker, K.; Sawyer, S.D.; Hu, D.; Huntsman, S.; Torres, J.; Echeverry, M.; Bohórquez, M.E.; et al. A Polygenic Risk Score for Breast Cancer in US Latinas and Latin American Women. *JNCI: Journal of the National Cancer Institute* **2020**, *112*, 590–598, doi:10.1093/jnci/djz174.
15. COLUMBUS Consortium; Hoffman, J.; Fejerman, L.; Hu, D.; Huntsman, S.; Li, M.; John, E.M.; Torres-Mejia, G.; Kushi, L.; Ding, Y.C.; et al. Identification of Novel Common Breast Cancer Risk Variants at the 6q25 Locus among Latinas. *Breast Cancer Res* **2019**, *21*, 3, doi:10.1186/s13058-018-1085-9.
16. Fejerman, L.; Stern, M.C.; John, E.M.; Torres-Mejía, G.; Hines, L.M.; Wolff, R.K.; Baumgartner, K.B.; Giuliano, A.R.; Ziv, E.; Pérez-Stable, E.J.; et al. Interaction between Common Breast Cancer Susceptibility Variants, Genetic Ancestry, and Nongenetic Risk Factors in Hispanic Women. *Cancer Epidemiology, Biomarkers & Prevention* **2015**, *24*, 1731–1738, doi:10.1158/1055-9965.EPI-15-0392.
17. Roberts, E.; Howell, S.; Evans, D.G. Polygenic Risk Scores and Breast Cancer Risk Prediction. *The Breast* **2023**, *67*, 71–77, doi:10.1016/j.breast.2023.01.003.
18. Gallardo-Blanco, H.L.; Villarreal-Perez, J.Z.; Cerda-Flores, R.M.; Figueroa, A.; Sanchez-Dominguez, C.N.; Gutierrez-Valverde, J.M.; Torres-Muñoz, I.C.; Lavalle-Gonzalez, F.J.; Gallegos-Cabriales, E.C.; Martinez-Garza, L.E. Genetic Variants in KCNJ11, TCF7L2 and HNF4A Are Associated with Type 2 Diabetes, BMI and Dyslipidemia in Families of Northeastern Mexico: A Pilot Study. *Experimental and Therapeutic Medicine* **2017**, *13*, 523–529, doi:10.3892/etm.2016.3990.
19. Kosoy, R.; Nassir, R.; Tian, C.; White, P.A.; Butler, L.M.; Silva, G.; Kittles, R.; Alarcon-Riquelme, M.E.; Gregersen, P.K.; Belmont, J.W.; et al. Ancestry Informative Marker Sets for Determining Continental Origin and Admixture Proportions in Common Populations in America. *Hum. Mutat.* **2009**, *30*, 69–78, doi:10.1002/humu.20822.
20. Nassir, R.; Kosoy, R.; Tian, C.; White, P.A.; Butler, L.M.; Silva, G.; Kittles, R.; Alarcon-Riquelme, M.E.; Gregersen, P.K.; Belmont, J.W.; et al. An Ancestry Informative Marker Set for Determining Continental Origin: Validation and Extension Using Human Genome Diversity Panels. *BMC Genet* **2009**, *10*, 39, doi:10.1186/1471-2156-10-39.
21. Dandine-Roulland, C.; Bellenguez, C.; Debette, S.; Amouyel, P.; Génin, E.; Perdry, H. Accuracy of Heritability Estimations in Presence of Hidden Population Stratification. *Sci Rep* **2016**, *6*, 26471, doi:10.1038/srep26471.
22. Liu, X.; Huang, M.; Fan, B.; Buckler, E.S.; Zhang, Z. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genet* **2016**, *12*, e1005767, doi:10.1371/journal.pgen.1005767.
23. Wang, S.-B.; Feng, J.-Y.; Ren, W.-L.; Huang, B.; Zhou, L.; Wen, Y.-J.; Zhang, J.; Dunwell, J.M.; Xu, S.; Zhang,

Y.-M. Improving Power and Accuracy of Genome-Wide Association Studies via a Multi-Locus Mixed Linear Model Methodology. *Sci Rep* **2016**, *6*, 19444, doi:10.1038/srep19444.

24. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol* **2016**, *17*, 122, doi:10.1186/s13059-016-0974-4.
25. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.; Roth, A.; Bork, P.; et al. The STRING Database in 2017: Quality-Controlled Protein–Protein Association Networks, Made Broadly Accessible. *Nucleic Acids Res* **2017**, *45*, D362–D368, doi:10.1093/nar/gkw937.
26. Segura, V.; Vilhjálmsson, B.J.; Platt, A.; Korte, A.; Seren, Ü.; Long, Q.; Nordborg, M. An Efficient Multi-Locus Mixed-Model Approach for Genome-Wide Association Studies in Structured Populations. *Nat Genet* **2012**, *44*, 825–830, doi:10.1038/ng.2314.
27. Van Den Berg, S.; Vandenplas, J.; Van Eeuwijk, F.A.; Lopes, M.S.; Veerkamp, R.F. Significance Testing and Genomic Inflation Factor Using High-density Genotypes or Whole-genome Sequence Data. *J Animal Breeding Genetics* **2019**, *136*, 418–429, doi:10.1111/jbg.12419.
28. Reyes, A.; Huber, W. Alternative Start and Termination Sites of Transcription Drive Most Transcript Isoform Differences across Human Tissues. *Nucleic Acids Research* **2018**, *46*, 582–592, doi:10.1093/nar/gkx1165.
29. Kamińska, M.; Ciszewski, T.; Łopacka-Szatan, K.; Miotła, P.; Starosławska, E. Breast Cancer Risk Factors. *pm* **2015**, *3*, 196–202, doi:10.5114/pm.2015.54346.
30. De Almeida, L.M.; Cortés, S.; Vilensky, M.; Valenzuela, O.; Cortes-Sanabria, L.; De Souza, M.; Barbeito, R.A.; Abdelhay, E.; Artagaveytia, N.; Daneri-Navarro, A.; et al. Socioeconomic, Clinical, and Molecular Features of Breast Cancer Influence Overall Survival of Latin American Women. *Front. Oncol.* **2022**, *12*, 845527, doi:10.3389/fonc.2022.845527.
31. Gallegos-Arreola, M.P.; Briseño-Zuno, C.J.; Figuera, L.E.; Sánchez-López, J.Y.; Zúñiga-González, G.M.; Puebla-Pérez, A.M.; Gómez-Meda, B.C.; Montoya-Fuentes, H.; Delgado-Saucedo, J.I. The Rs1008562, Rs2234671, and Rs3138060 Polymorphisms of the CXCR1 Gene Are Associated with Breast Cancer Risk in a Mexican Population. *European Review for Medical and Pharmacological Sciences* **2020**, *24*, 9990–10002, doi:10.26355/eurrev_202010_23212.
32. Gutiérrez-Rubio, S.A.; Quintero-Ramos, A.; Durán-Cárdenas, A.; Franco-Topete, R.A.; Castro-Cervantes, J.M.; Oceguera-Villanueva, A.; Jiménez-Pérez, L.M.; Balderas-Peña, L.M.A.; Morgan-Villela, G.; Del-Toro-Arreola, A.; et al. 1236 C/T and 3435 C/T Polymorphisms of the ABCB1 Gene in Mexican Breast Cancer Patients. *Genet. Mol. Res.* **2015**, *14*, 1250–1259, doi:10.4238/2015.February.13.3.
33. Pérez-Rodríguez, G.; Aranda-Moreno, C.; Olivares-Corichi, I.M.; Garcia-Sanchez, J.R. The Association of Subtypes of Breast Cancer with Tumour Characteristics and Reproductive Factors in 1326 Mexican Women. *wo* **2015**, *6*, 462–466, doi:10.5114/wo.2015.56652.
34. Park, S. Interaction of Polygenic Variants Specific for Abdominal Obesity Risk with Energy Metabolism in Large Korean Cohorts. *Nutrition Bulletin* **2022**, *47*, 307–321, doi:10.1111/nbu.12569.
35. Azim, H.A.; Partridge, A.H. Biology of Breast Cancer in Young Women. *Breast Cancer Res* **2014**, *16*, 427, doi:10.1186/s13058-014-0427-5.
36. Moreno-Estrada, A.; Gignoux, C.R.; Fernandez-Lopez, J.C.; Zakharia, F.; Sikora, M.; Contreras, A.V.; Acuna-Alonzo, V.; Sandoval, K.; Eng, C.; Romero-Hidalgo, S.; et al. The Genetics of Mexico Recapitulates Native American Substructure and Affects Biomedical Traits. *Science* **2014**, *344*, 1280–1285, doi:10.1126/science.1251688.
37. Adedokun, B.; Du, Z.; Gao, G.; Ahearn, T.U.; Lunetta, K.L.; Zirpoli, G.; Figueroa, J.; John, E.M.; Bernstein, L.; Zheng, W.; et al. Cross-Ancestry GWAS Meta-Analysis Identifies Six Breast Cancer Loci in African and European Ancestry Women. *Nat Commun* **2021**, *12*, 4198, doi:10.1038/s41467-021-24327-x.
38. Lindström, S.; Wang, L.; Feng, H.; Majumdar, A.; Huo, S.; Macdonald, J.; Harrison, T.; Turman, C.; Chen, H.; Mancuso, N.; et al. Genome-Wide Analyses Characterize Shared Heritability among Cancers and Identify Novel Cancer Susceptibility Regions. *JNCI: Journal of the National Cancer Institute* **2023**, *115*, 712–732, doi:10.1093/jnci/djad043.
39. Middha, P.; Wang, X.; Behrens, S.; Bolla, M.K.; Wang, Q.; Dennis, J.; Michailidou, K.; Ahearn, T.U.; Andrulis, I.L.; Anton-Culver, H.; et al. A Genome-Wide Gene-Environment Interaction Study of Breast

 doi:10.20944/preprints202501.1758.v1

Cancer Risk for Women of European Ancestry. *Breast Cancer Res* **2023**, *25*, 93, doi:10.1186/s13058-023-01691-8.

40. Lange, M.; Begolli, R.; Giakountis, A. Non-Coding Variants in Cancer: Mechanistic Insights and Clinical Potential for Personalized Medicine. *ncRNA* **2021**, *7*, 47, doi:10.3390/ncrna7030047.
41. Lang, B.; Armaos, A.; Tartaglia, G.G. RNAct: Protein–RNA Interaction Predictions for Model Organisms with Supporting Experimental Data. *Nucleic Acids Research* **2019**, *47*, D601–D606, doi:10.1093/nar/gky967.
42. Chen, J.; Bardes, E.E.; Aronow, B.J.; Jegga, A.G. ToppGene Suite for Gene List Enrichment Analysis and Candidate Gene Prioritization. *Nucleic Acids Research* **2009**, *37*, W305–W311, doi:10.1093/nar/gkp427.
43. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T.I.; Nudel, R.; Lieder, I.; Mazor, Y.; et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics* **2016**, *54*, doi:10.1002/cpbi.5.
44. Kelly, P.; Stemmle, L.N.; Madden, J.F.; Fields, T.A.; Daaka, Y.; Casey, P.J. A Role for the G12 Family of Heterotrimeric G Proteins in Prostate Cancer Invasion. *J Biol Chem* **2006**, *281*, 26483–26490, doi:10.1074/jbc.M604376200.
45. Kelly, P.; Moeller, B.J.; Juneja, J.; Booden, M.A.; Der, C.J.; Daaka, Y.; Dewhirst, M.W.; Fields, T.A.; Casey, P.J. The G12 Family of Heterotrimeric G Proteins Promotes Breast Cancer Invasion and Metastasis. *Proc Natl Acad Sci U S A* **2006**, *103*, 8173–8178, doi:10.1073/pnas.0510254103.
46. Yuan, B.; Cui, J.; Wang, W.; Deng, K. Gα12/13 Signaling Promotes Cervical Cancer Invasion through the RhoA/ROCK-JNK Signaling Axis. *Biochem Biophys Res Commun* **2016**, *473*, 1240–1246, doi:10.1016/j.bbrc.2016.04.048.
47. Rasheed, S.A.K.; Teo, C.R.; Beillard, E.J.; Voorhoeve, P.M.; Zhou, W.; Ghosh, S.; Casey, P.J. MicroRNA-31 Controls G Protein Alpha-13 (GNA13) Expression and Cell Invasion in Breast Cancer Cells. *Mol Cancer* **2015**, *14*, 67, doi:10.1186/s12943-015-0337-x.
48. Teo, C.R.; Casey, P.J.; Rasheed, S.A.K. The GNA13-RhoA Signaling Axis Suppresses Expression of Tumor Protective Kallikreins. *Cellular Signalling* **2016**, *28*, 1479–1488, doi:10.1016/j.cellsig.2016.07.001.
49. Armendariz, D.A.; Sundarrajan, A.; Hon, G.C. Breaking Enhancers to Gain Insights into Developmental Defects. *eLife* **2023**, *12*, e88187, doi:10.7554/eLife.88187.
50. Parker, S.C.J.; Stitzel, M.L.; Taylor, D.L.; Orozco, J.M.; Erdos, M.R.; Akiyama, J.A.; Van Bueren, K.L.; Chines, P.S.; Narisu, N.; NISC Comparative Sequencing Program; et al. Chromatin Stretch Enhancer States Drive Cell-Specific Gene Regulation and Harbor Human Disease Risk Variants. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 17921–17926, doi:10.1073/pnas.1317023110.
51. Rosas-Cruz, A.; Salinas-Jazmín, N.; Valdés-Rives, A.; Velasco-Velázquez, M.A. DRD1 and DRD4 Are Differentially Expressed in Breast Tumors and Breast Cancer Stem Cells: Pharmacological Implications. *Transl Cancer Res TCR* **2022**, *11*, 3941–3950, doi:10.21037/tcr-22-783.
52. Grant, C.E.; Flis, A.L.; Ryan, B.M. Understanding the Role of Dopamine in Cancer: Past, Present and Future. *Carcinogenesis* **2022**, *43*, 517–527, doi:10.1093/carcin/bgac045.
53. Wang, X.; Wang, Z.-B.; Luo, C.; Mao, X.-Y.; Li, X.; Yin, J.-Y.; Zhang, W.; Zhou, H.-H.; Liu, Z.-Q. The Prospective Value of Dopamine Receptors on Bio-Behavior of Tumor. *J. Cancer* **2019**, *10*, 1622–1632, doi:10.7150/jca.27780.
54. Bañales-Luna, M.; Figueroa-Vega, N.; Marín-Aragón, C.I.; Perez-Luque, E.; Ibarra-Reynoso, L.; Gallardo-Blanco, H.L.; López-Aguilar, I.; Malacara, J.M. Associations of Nicotidamide-N-Methyltransferase, FTO, and IRX3 Genetic Variants with Body Mass Index and Resting Energy Expenditure in Mexican Subjects. *Sci Rep* **2020**, *10*, 11478, doi:10.1038/s41598-020-67832-7.
55. Lan, N.; Lu, Y.; Zhang, Y.; Pu, S.; Xi, H.; Nie, X.; Liu, J.; Yuan, W. FTO – A Common Genetic Basis for Obesity and Cancer. *Front. Genet.* **2020**, *11*, 559138, doi:10.3389/fgene.2020.559138.
56. Liu, Y.; Wang, R.; Zhang, L.; Li, J.; Lou, K.; Shi, B. The Lipid Metabolism Gene FTO Influences Breast Cancer Cell Energy Metabolism via the PI3K/AKT Signaling Pathway. *Oncology Letters* **2017**, *13*, 4685–4690, doi:10.3892/ol.2017.6038.
57. Abdollahi, S.; Hasanpour Ardekanizadeh, N.; Poorhosseini, S.M.; Gholamalizadeh, M.; Roumi, Z.; Goodarzi, M.O.; Doaei, S. Unraveling the Complex Interactions between the Fat Mass and Obesity-Associated (FTO) Gene, Lifestyle, and Cancer. *Advances in Nutrition* **2022**, *13*, 2406–2419,

doi:10.1093/advances/nmac101.

58. Delahanty, R.J.; Beeghly-Fadiel, A.; Xiang, Y.-B.; Long, J.; Cai, Q.; Wen, W.; Xu, W.-H.; Cai, H.; He, J.; Gao, Y.-T.; et al. Association of Obesity-Related Genetic Variants With Endometrial Cancer Risk: A Report From the Shanghai Endometrial Cancer Genetics Study. *American Journal of Epidemiology* **2011**, *174*, 1115–1126, doi:10.1093/aje/kwr233.
59. Huang, X.; Zhao, J.; Yang, M.; Li, M.; Zheng, J. Association between FTO Gene Polymorphism (Rs9939609 T/A) and Cancer Risk: A Meta-Analysis. *Eur J Cancer Care* **2017**, *26*, e12464, doi:10.1111/ecc.12464.
60. Kaklamani, V.; Yi, N.; Sadim, M.; Siziopikou, K.; Zhang, K.; Xu, Y.; Tofilon, S.; Agarwal, S.; Pasche, B.; Mantzoros, C. The Role of the Fat Mass and Obesity Associated Gene (FTO) in Breast Cancer Risk. *BMC Med Genet* **2011**, *12*, 52, doi:10.1186/1471-2350-12-52.
61. Lin, Y.; Ueda, J.; Yagyu, K.; Ishii, H.; Ueno, M.; Egawa, N.; Nakao, H.; Mori, M.; Matsuo, K.; Kikuchi, S. Association between Variations in the Fat Mass and Obesity-Associated Gene and Pancreatic Cancer Risk: A Case–Control Study in Japan. *BMC Cancer* **2013**, *13*, 337, doi:10.1186/1471-2407-13-337.
62. Zeng, X.; Ban, Z.; Cao, J.; Zhang, W.; Chu, T.; Lei, D.; Du, Y. Association of FTO Mutations with Risk and Survival of Breast Cancer in a Chinese Population. *Disease Markers* **2015**, *2015*, 1–6, doi:10.1155/2015/101032.
63. Long, J.; Zhang, B.; Signorello, L.B.; Cai, Q.; Deming-Halverson, S.; Shrubsole, M.J.; Sanderson, M.; Dennis, J.; Michailiou, K.; Easton, D.F.; et al. Evaluating Genome-Wide Association Study-Identified Breast Cancer Risk Variants in African-American Women. *PLoS ONE* **2013**, *8*, e58350, doi:10.1371/journal.pone.0058350.
64. Pan, J.; Huang, T.; Deng, Z.; Zou, C. Roles and Therapeutic Implications of m6A Modification in Cancer Immunotherapy. *Front. Immunol.* **2023**, *14*, 1132601, doi:10.3389/fimmu.2023.1132601.
65. Allen, I.C.; Hartney, J.M.; Coffman, T.M.; Penn, R.B.; Wess, J.; Koller, B.H. Thromboxane A $_2$ Induces Airway Constriction through an M $_3$ Muscarinic Acetylcholine Receptor-Dependent Mechanism. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **2006**, *290*, L526–L533, doi:10.1152/ajplung.00340.2005.
66. Kurosaki, T.; Maquat, L.E. Nonsense-Mediated mRNA Decay in Humans at a Glance. *J Cell Sci* **2016**, *129*, 461–467, doi:10.1242/jcs.181008.
67. Haghvirdizadeh, P.; Mohamed, Z.; Abdullah, N.A.; Haghvirdizadeh, P.; Haerian, M.S.; Haerian, B.S. KCNJ11: Genetic Polymorphisms and Risk of Diabetes Mellitus. *Journal of Diabetes Research* **2015**, *2015*, 1–9, doi:10.1155/2015/908152.
68. Cheng, I.; Caberto, C.P.; Lum-Jones, A.; Seifried, A.; Wilkens, L.R.; Schumacher, F.R.; Monroe, K.R.; Lim, U.; Tiirikainen, M.; Kolonel, L.N.; et al. Type 2 Diabetes Risk Variants and Colorectal Cancer Risk: The Multiethnic Cohort and PAGE Studies. *Gut* **2011**, *60*, 1703–1711, doi:10.1136/gut.2011.237727.
69. Shuhua, W.; Chenbo, S.; Yangyang, L.; Xiangqian, G.; Shuang, H.; Tangyue, L.; Dong, T. Autophagy-Related Genes Raptor, Rictor, and Beclin1 Expression and Relationship with Multidrug Resistance in Colorectal Carcinoma. *Human Pathology* **2015**, *46*, 1752–1759, doi:10.1016/j.humpath.2015.07.016.
70. Uhlen, M.; Zhang, C.; Lee, S.; Sjöstedt, E.; Fagerberg, L.; Bidkhori, G.; Benfeitas, R.; Arif, M.; Liu, Z.; Edfors, F.; et al. A Pathology Atlas of the Human Cancer Transcriptome. *Science* **2017**, *357*, eaan2507, doi:10.1126/science.aan2507.
71. Zhang, Z.; Zhang, S.; Wong, M.; Wareham, N.J.; Sha, Q. An Ensemble Learning Approach Jointly Modeling Main and Interaction Effects in Genetic Association Studies. *Genetic Epidemiology* **2008**, *32*, 285–300, doi:10.1002/gepi.20304.
72. Zhao, L.; Lian, T.; Li, J.; Wei, S.; Li, H.; Li, C.; Wang, H. NCR3LG1 (B7-H6) Is a Potential Prognostic Factor for Bladder Cancer Patients. *Biomarkers* **2021**, *26*, 260–267, doi:10.1080/1354750X.2021.1883110.
73. Chen, H.; Zhao, L.; Meng, Y.; Qian, X.; Fan, Y.; Zhang, Q.; Wang, C.; Lin, F.; Chen, B.; Xu, L.; et al. Sulfonylurea Receptor 1-Expressing Cancer Cells Induce Cancer-Associated Fibroblasts to Promote Non-Small Cell Lung Cancer Progression. *Cancer Letters* **2022**, *536*, 215611, doi:10.1016/j.canlet.2022.215611.
74. Davies, K.J. The Complex Interaction of Matrix Metalloproteinases in the Migration of Cancer Cells through Breast Tissue Stroma. *International Journal of Breast Cancer* **2014**, *2014*, 1–5, doi:10.1155/2014/839094.
75. Radisky, E.S. Matrix Metalloproteinases as Drivers and Therapeutic Targets in Breast Cancer. *Front Biosci* **2015**, *20*, 1144–1163, doi:10.2741/4364.
76. Walsh, L.; Roy, D. Candidate Prognostic Markers in Breast Cancer: Focus on Extracellular Proteases and

Their Inhibitors. *BCTT* **2014**, 81, doi:10.2147/BCTT.S46020.

77. Juurikka; Butler; Salo; Nyberg; Åström The Role of MMP8 in Cancer: A Systematic Review. *IJMS* **2019**, *20*, 4506, doi:10.3390/ijms20184506.
78. Decock, J.; Long, J.-R.; Laxton, R.C.; Shu, X.-O.; Hodgkinson, C.; Hendrickx, W.; Pearce, E.G.; Gao, Y.-T.; Pereira, A.C.; Paridaens, R.; et al. Association of *Matrix Metalloproteinase-8* Gene Variation with Breast Cancer Prognosis. *Cancer Research* **2007**, *67*, 10214–10221, doi:10.1158/0008-5472.CAN-07-1683.
79. Decock, J.; Long, J.; Laxton, R.; Shu, X.; Hodgkinson, C.; Hendrickx, W.; Pearce, E.; Gao, Y.; Pereira, A.; Paridaens, R.; et al. Association of MMP8 Gene Variation with Breast Cancer Prognosis. *Breast Cancer Res* **2008**, *10*, P32, bcr1916, doi:10.1186/bcr1916.
80. Chistiakov, D.A.; Melnichenko, A.A.; Orekhov, A.N.; Bobryshev, Y.V. Paraoxonase and Atherosclerosis-Related Cardiovascular Diseases. *Biochimie* **2017**, *132*, 19–27, doi:10.1016/j.biochi.2016.10.010.
81. Mackness, M.; Mackness, B. Human Paraoxonase-1 (PON1): Gene Structure and Expression, Promiscuous Activities and Multiple Physiological Roles. *Gene* **2015**, *567*, 12–21, doi:10.1016/j.gene.2015.04.088.
82. Medina-Díaz, I.M.; Ponce-Ruíz, N.; Rojas-García, A.E.; Zambrano-Zargoza, J.F.; Bernal-Hernández, Y.Y.; González-Arias, C.A.; Barrón-Vivanco, B.S.; Herrera-Moreno, J.F. The Relationship between Cancer and Paraoxonase 1. *Antioxidants* **2022**, *11*, 697, doi:10.3390/antiox11040697.
83. Stevens, V.L.; Rodriguez, C.; Pavluck, A.L.; Thun, M.J.; Calle, E.E. Association of Polymorphisms in the *Paraoxonase 1* Gene with Breast Cancer Incidence in the CPS-II Nutrition Cohort. *Cancer Epidemiology, Biomarkers & Prevention* **2006**, *15*, 1226–1228, doi:10.1158/1055-9965.EPI-05-0930.
84. Yuzhalin, A.E.; Kutikhin, A.G. Common Genetic Variants in the Myeloperoxidase and Paraoxonase Genes and the Related Cancer Risk: A Review. *Journal of Environmental Science and Health, Part C* **2012**, *30*, 287–322, doi:10.1080/10590501.2012.731957.
85. Kim, D.S.; Burt, A.A.; Ranchalis, J.E.; Richter, R.J.; Marshall, J.K.; Eintracht, J.F.; Rosenthal, E.A.; Furlong, C.E.; Jarvik, G.P. Additional Common Polymorphisms in the *PON* Gene Cluster Predict PON1 Activity but Not Vascular Disease. *Journal of Lipids* **2012**, *2012*, 1–11, doi:10.1155/2012/476316.
86. Evans, R.M.; Barish, G.D.; Wang, Y.-X. PPARs and the Complex Journey to Obesity. *Nat Med* **2004**, *10*, 355–361, doi:10.1038/nm1025.
87. Oh, S.W.; Park, C.-Y.; Lee, E.S.; Yoon, Y.S.; Lee, E.S.; Park, S.S.; Kim, Y.; Sung, N.J.; Yun, Y.H.; Lee, K.S.; et al. Adipokines, Insulin Resistance, Metabolic Syndrome, and Breast Cancer Recurrence: A Cohort Study. *Breast Cancer Res* **2011**, *13*, R34, doi:10.1186/bcr2856.
88. Zhang, S.; Jiang, J.; Chen, Z.; Wang, Y.; Tang, W.; Chen, Y.; Liu, L. Relationship of PPARG, PPARGC1A, and PPARGC1B Polymorphisms with Susceptibility to Hepatocellular Carcinoma in an Eastern Chinese Han Population. *OTT* **2018**, *Volume 11*, 4651–4660, doi:10.2147/OTT.S168274.
89. Unal, E.; Aslan, E.I.; Ozturk, T.; Kurnaz Gomleksiz, O.; Kucukhuseyin, O.; Tuzuner, M.B.; Seyhan, M.F.; Ozturk, O.; Yilmaz Aydogan, H. Peroxisome Proliferator-Activated Receptor Gamma Pro12Ala/C161T Genotypes and Risky Haplotype Altering Risk of Breast Cancer: A Turkish Case–Control Study. *Biochem Genet* **2021**, *59*, 1413–1426, doi:10.1007/s10528-021-10068-5.
90. Tang, W.; Chen, Y.; Wang, Y.; Gu, H.; Chen, S.; Kang, M. Peroxisome Proliferator-Activated Receptor Gamma (PPARG) Polymorphisms and Breast Cancer Susceptibility: A Meta-Analysis. *Int J Clin Exp Med* **2015**, *8*, 12226–12238.
91. Schöckel, L.; Woischke, C.; Surendran, S.A.; Michl, M.; Schiergens, T.; Hölscher, A.; Glass, F.; Kreissl, P.; Klauschen, F.; Günther, M.; et al. PPARG Activation Promotes the Proliferation of Colorectal Cancer Cell Lines and Enhances the Antiproliferative Effect of 5-Fluorouracil. *BMC Cancer* **2024**, *24*, 234, doi:10.1186/s12885-024-11985-5.
92. Janku, F.; Yap, T.A.; Meric-Bernstam, F. Targeting the PI3K Pathway in Cancer: Are We Making Headway? *Nat Rev Clin Oncol* **2018**, *15*, 273–291, doi:10.1038/nrclinonc.2018.28.
93. Wang, T.; Zhang, W.; Wang, Z.; Wu, Z.; Du, B.; Li, L.; Chen, Y.; Yang, X.; Hao, X.; Guo, T. RAPTOR Promotes Colorectal Cancer Proliferation by Inducing mTORC1 and Upregulating Ribosome Assembly Factor URB1. *Cancer Medicine* **2020**, *9*, 1529–1543, doi:10.1002/cam4.2810.
94. Shchegolev, Y.; Sorokin, D.; Scherbakov, A.; Shunaev, A.; Andreeva, O.; Mikhaevich, E.; Gudkova, M.; Bure, I.; Berstein, L.; Nemtsova, M.; et al. Upregulation of Akt/Raptor Signaling Is Associated with

Rapamycin Resistance of Breast Cancer Cells. *Chemico-Biological Interactions* **2020**, *330*, 109243, doi:10.1016/j.cbi.2020.109243.

95. Tang, Q.; Holland-Letz, T.; Slynko, A.; Cuk, K.; Marme, F.; Schott, S.; Heil, J.; Qu, B.; Golatta, M.; Bewerunge-Hudler, M.; et al. DNA Methylation Array Analysis Identifies Breast Cancer Associated *RPTOR* , *MGRN1* and *RAPSN* Hypomethylation in Peripheral Blood DNA. *Oncotarget* **2016**, *7*, 64191–64202, doi:10.18632/oncotarget.11640.
96. Luo, H.L.; Chiang, P.H.; Lee, N.L.; Chiang, P.H. RAPTOR Gene Polymorphism Is Independently Correlated with Urothelial Cancer Susceptibility Compared with Environmental Toxin Exposure. *Urological Science* **2017**, *28*, 197–199, doi:10.1016/j.urols.2016.05.002.
97. Cheng, T.-Y.D.; Shankar, J.; Zirpoli, G.; Roberts, M.R.; Hong, C.-C.; Bandera, E.V.; Ambrosone, C.B.; Yao, S. Genetic Variants in the mTOR Pathway and Interaction with Body Size and Weight Gain on Breast Cancer Risk in African-American and European American Women. *Cancer Causes Control* **2016**, *27*, 965–976, doi:10.1007/s10552-016-0774-x.
98. Wang, W.; Qin, J.; Voruganti, S.; Nag, S.; Zhou, J.; Zhang, R. Polycomb Group (PcG) Proteins and Human Cancers: Multifaceted Functions and Therapeutic Implications. *Medicinal Research Reviews* **2015**, *35*, 1220–1267, doi:10.1002/med.21358.
99. Sengupta, S.; Peterson, T.R.; Sabatini, D.M. Regulation of the mTOR Complex 1 Pathway by Nutrients, Growth Factors, and Stress. *Molecular Cell* **2010**, *40*, 310–322, doi:10.1016/j.molcel.2010.09.026.
100. Bartolome, A.; Kimura-Koyanagi, M.; Asahara, S.-I.; Guillen, C.; Inoue, H.; Teruyama, K.; Shimizu, S.; Kanno, A.; Garcia-Aguilar, A.; Koike, M.; et al. Pancreatic -Cell Failure Mediated by mTORC1 Hyperactivity and Autophagic Impairment. *Diabetes* **2014**, *63*, 2996–3008, doi:10.2337/db13-0970.
101. Makhoul, C.; Gleeson, P.A. Regulation of mTORC1 Activity by the Golgi Apparatus. *Fac Rev* **2021**, *10*, doi:10.12703/r/10-50.
102. Yin, Q.; Ni, Q.; Wang, Y.; Zhang, H.; Li, W.; Nie, A.; Wang, S.; Gu, Y.; Wang, Q.; Ning, G. Raptor Determines β-Cell Identity and Plasticity Independent of Hyperglycemia in Mice. *Nat Commun* **2020**, *11*, 2538, doi:10.1038/s41467-020-15935-0.
103. Le Bacquer, O.; Queniat, G.; Gmyr, V.; Kerr-Conte, J.; Lefebvre, B.; Pattou, F. mTORC1 and mTORC2 Regulate Insulin Secretion through Akt in INS-1 Cells. *Journal of Endocrinology* **2013**, *216*, 21–29, doi:10.1530/JOE-12-0351.
104. Liu, W.; Xie, L.; He, Y.-H.; Wu, Z.-Y.; Liu, L.-X.; Bai, X.-F.; Deng, D.-X.; Xu, X.-E.; Liao, L.-D.; Lin, W.; et al. Large-Scale and High-Resolution Mass Spectrometry-Based Proteomics Profiling Defines Molecular Subtypes of Esophageal Cancer for Therapeutic Targeting. *Nat Commun* **2021**, *12*, 4961, doi:10.1038/s41467-021-25202-5.
105. Köttgen, A.; Hwang, S.-J.; Rampersaud, E.; Coresh, J.; North, K.E.; Pankow, J.S.; Meigs, J.B.; Florez, J.C.; Parsa, A.; Levy, D.; et al. TCF7L2 Variants Associate with CKD Progression and Renal Function in Population-Based Cohorts. *J Am Soc Nephrol* **2008**, *19*, 1989–1999, doi:10.1681/ASN.2007121291.
106. Lu, X.P.; Hu, G.N.; Du, J.Q.; Li, H.Q. TCF7L2 Gene Polymorphisms and Susceptibility to Breast Cancer: A Meta-Analysis. *Genet. Mol. Res.* **2015**, *14*, 2860–2867, doi:10.4238/2015.March.31.16.
107. Lyssenko, V. The Transcription Factor 7-like 2 Gene and Increased Risk of Type 2 Diabetes: An Update: *Current Opinion in Clinical Nutrition and Metabolic Care* **2008**, *11*, 385–392, doi:10.1097/MCO.0b013e328304d970.
108. Connor, A.E.; Baumgartner, R.N.; Baumgartner, K.B.; Kerber, R.A.; Pinkston, C.; John, E.M.; Torres-Mejia, G.; Hines, L.; Giuliano, A.; Wolff, R.K.; et al. Associations between TCF7L2 Polymorphisms and Risk of Breast Cancer among Hispanic and Non-Hispanic White Women: The Breast Cancer Health Disparities Study. *Breast Cancer Res Treat* **2012**, *136*, 593–602, doi:10.1007/s10549-012-2299-7.
109. Shaheen, K.Y.; Al Ahmed, H.A.; Shaaban, M.A.A.; Emam, Y.M. Study of The Association between TCF7L2 Gene Polymorphism (Rs12255372) and Breast Cancer. *QJM: An International Journal of Medicine* **2021**, *114*, hcab091.015, doi:10.1093/qjmed/hcab091.015.

 doi:10.20944/preprints202501.1758.v1

110. DepMap, B. DepMap 24Q2 Public 2024, 25031191226 Bytes.
111. Camicia, R.; Winkler, H.C.; Hassa, P.O. Novel Drug Targets for Personalized Precision Medicine in Relapsed/Refractory Diffuse Large B-Cell Lymphoma: A Comprehensive Review. *Mol Cancer* **2015**, *14*, 207, doi:10.1186/s12943-015-0474-2.