 doi:10.20944/preprints202209.0360.v1

Article

Synergistic Interaction of A Healthy Lifestyle Index and Antioxidant Genes on Breast Cancer Risk

Liliana Gómez-Flores-Ramos[1,2], Luisa María Sánchez-Zamorano*[2], Angélica Ángeles-Llerenas[2], Rocío Rodríguez-Valentín[2] and Gabriela Torres-Mejía[2]

[1]National Council of Science and Technology (CONACYT)

[2]Population Health Research Center, National Institute of Public Health, Cuernavaca, Morelos, Mexico;

* Correspondence: Author: *Luisa María Sánchez-Zamorano. Centro de Investigacion en Salud Poblacional, Instituto Nacional de Salud Pública. Avenida Universidad 655 Colonia Santa Maria Ahuacatitlan, Cuernavaca, Morelos, C. P. 62508, Mexico. Email: szamoran@insp.mx

Abstract: Lifestyle has been associated with breast cancer risk through different pathways, including oxidative stress. Antioxidant enzymes are endogenous defense mechanisms against oxidative stress damage, and this response might be modulated by the genetic variation in these enzyme-codifying genes. This study aimed to analyze the synergistic effect of an antioxidant Healthy Lifestyle Index (HeLiX) composed of principal components of Western dietary pattern, alcohol consumption, smoking and physical activity, and genetic polymorphisms in the first-line antioxidant response family genes *SOD*, *GPX*, and *CAT* on breast cancer risk. We included 176 SNPs, and only *CAT rs554576* remained significant after correction for multiple comparisons. Breast cancer odds were reduced at the highest (T3) and medium (T2) tertiles of the HeLiX. When stratified by HeLiX, we observed a reduced risk of breast cancer with at least one T-allele, and the effect increased in a dose-dependent manner. Compared to the reference category (HeLiX T1 and AA genotype), women at the HeLiX T3 with AT and TT genotypes in postmenopausal women showed an OR = 0.15 (95% CI 0.07–0.32). For HeLiX T2 and AT genotype OR = 0.26 (95% CI 0.13–0.54); for TT genotype OR = 0.24 (95% CI 0.12–0.45). For premenopausal women, at the HeLiX T3 and AT genotype OR = 0.29 (95% CI 0.13–0.62); for the TT genotype OR = 0.21 (95% CI 0.08–0.51). We also observed an inverse association for HeLiX T2 and TT genotype (OR 0.39 95% CI 0.17–0.87). Our study shows a significant synergistic gene-environment interaction on an additive scale, contributing to understanding pathways involved in breast cancer etiology and prevention.

Keywords: Breast cancer; Oxidative stress; Healthy lifestyle; Catalase; Gene–environment interaction

Introduction

Breast cancer (BC) is the most common cancer and the leading cause of cancer death for women worldwide, but BC incidence and mortality burden vary by geographical region[1].

On the molecular level, BC is a complex and heterogeneous disease that has been associated with an unhealthy lifestyle, diet, smoking, and alcohol consumption, which are important exogenous sources of reactive oxygen species (ROS) that

 doi:10.20944/preprints202209.0360.v1

cause oxidative stress (OS)[2]. ROS are generated as byproducts of aerobic cellular respiration, metabolism, and the mitochondrial electron transport chain[3]. These oxidants act as second messengers in cell proliferation and apoptosis and are essential for redox homeostasis and immune response. However, when the balance between oxidants and antioxidants is lost in favor of free radicals, OS occurs[4].

This OS causes DNA base modification and mutation, structural damage, genomic instability, and methylation changes. Furthermore, it causes protein conformational changes and lipid peroxidation[5]. OS promotes BC by activating signaling pathways of cell survival, proliferation, apoptosis evasion, angiogenesis, and metastasis, such as Src, Wnt/-catenin, Ras/Raf/Mek/ERK, EGFR, and PI3K/Akt[6–9].

Antioxidant enzymes are endogenous cellular defense mechanisms that continuously work to neutralize any molecule with the potential to become a free radical that causes damage. Three enzymes act as the first line of defense: superoxide dismutase (SOD) converts two molecules of superoxide anion to hydrogen peroxide; glutathione peroxidase (GPx) catalyzes the reduction of different hydroperoxides; and catalase (CAT) transforms hydrogen peroxide into water and oxygen. The cellular locations of these enzymes are the mitochondria, cytoplasm, and peroxisomes, respectively[10]. Moreover, some single nucleotide polymorphisms (SNPs) in these antioxidant enzyme-codifying genes have been associated with BC risk[11].

Unhealthy lifestyles have been associated with more ROS production and a higher risk of cancer, while healthier lifestyles are associated with less ROS and a decreased risk of cancer[7]. Furthermore, some studies suggest that the effect of a healthy lifestyle might be potentiated by genes for an antioxidant response[12–14].

 doi:10.20944/preprints202209.0360.v1

In a previous study, we found that the Healthy Lifestyle Index (HeLiX), which included a Western dietary pattern, physical activity, alcohol consumption, and tobacco smoking, was associated with the risk of BC and that the size of the effect differed by menopausal status [15]. To better understand the interaction between lifestyle and the endogenous antioxidant response, this study aimed to analyze the synergistic effect of HeLiX and genetic polymorphisms in the first-line antioxidant response genes *SOD*, *GPX*, and *CAT* on the risk of BC.

Materials and Methods

Study design and population

We analyzed data from the CAMA study, a population-based case-control study performed from January 2004 to December 2007, which has been described previously[16]. In brief, the study was carried out in three ethnically and culturally diverse Mexican cities: Monterrey (northern), Veracruz (east coast), and Mexico City (center). Women between 35 and 69 who had been residents for the last five years at one of the three sites were eligible to participate in the study.

Cases (n = 1000) were selected if the participant (i) had a recent histopathological BC diagnosis (median = 3 days; range = 0–6 days), regardless of the stage of disease, (ii) had not been previously treated with chemotherapy, radiotherapy, or estrogen antagonist in the past six months, and (iii) were not pregnant. Controls (n = 1074) were frequency matched to cases based on 5-year age groups, place of residence, and healthcare system using a probabilistic multistage approach. Study personnel visited the selected households and determined their willingness to participate in the study. Trained interviewers gathered information for both cases and controls.

 doi:10.20944/preprints202209.0360.v1

The Institutional Research and Ethics Committees of the National Institute of Public Health, Mexico (Number CI-349), and the research committees at the collaborating hospitals authorized this study. All research participants provided written informed consent for all procedures and measurements described in this study.

The study subjects provided information about their health, physical activity, and diet. The health questionnaire collected data on sociodemographic characteristics, reproductive factors, use of hormones, and other health characteristics. We obtained information on food consumption the year before the onset of symptoms for cases and a regular week for controls, using a Food Frequency Questionnaire adapted from Willet[17]. A semi-structured interview that estimates an individual's time spent in physical activities (sleep or light, moderate, or vigorous intensity) during a regular week was applied to assess physical activity. For cases, diet and physical activity information was obtained preceding any BC symptoms to reduce the possibility of reverse causation bias[16].

Subsample

For the present study, out of 1000 cases and 1074 controls, we included 711 incident cases and 706 controls with genetic information from GWAS genotyping. Individuals with missing data for calculating the HeLiX were excluded. The final analytical sub-sample consisted of 636 cases and 679 controls (Fig. 1).

Menopause

Menopause was defined as a woman's self-report of experiencing natural amenorrhea for at least 12 months. Induced menopause was considered if women reported a bilateral oophorectomy or hysterectomy. If the status was unknown, women > 48 years were considered postmenopausal, since the mean age of menopause in Mexico for the study period

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

was 48 years[18]. Analyses were stratified by menopausal status (Fig.1), as previous analyses have shown that the effect of lifestyle and OS is different in pre- and postmenopausal women[16,19–21].

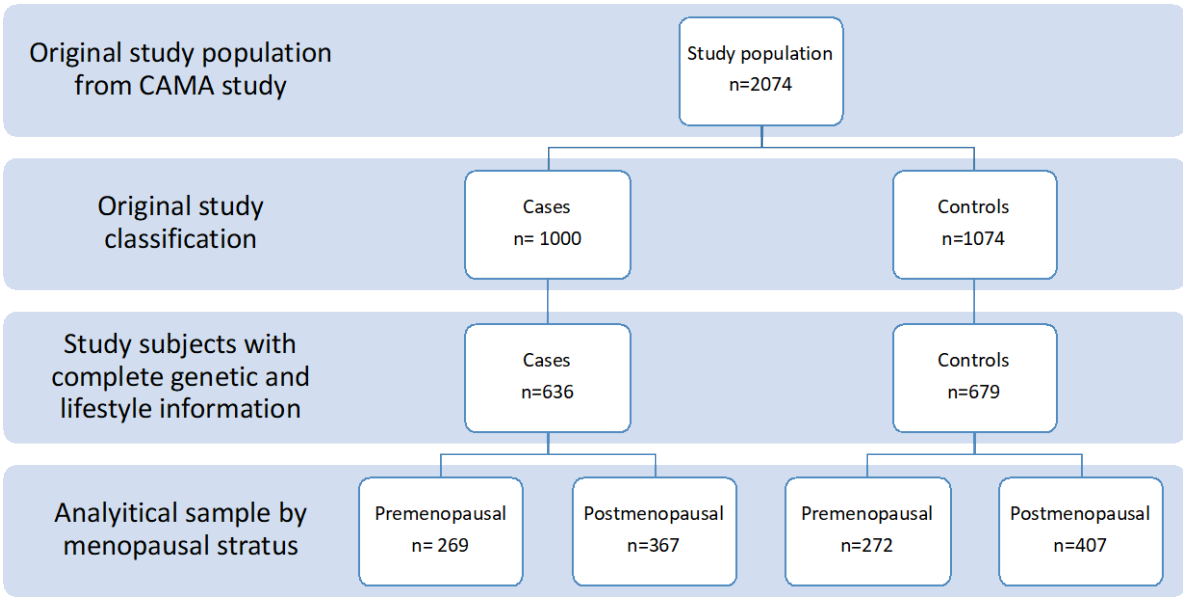

Figure 1. Study population

Healthy Lifestyle Index (HeLiX)

We constructed a HeLiX for the subsample included in this study by means of principal components using a dietary pattern, physical activity, alcohol consumption, and tobacco smoking as the main drivers of exogenous OS. The index composition was based on previously published research by Sánchez-Zamorano[15].

Diet assessment

 doi:10.20944/preprints202209.0360.v1

For the diet assessment, we applied a semiquantitative food frequency questionnaire (FFQ) adapted from Willett[17] and validated for the Mexican population[22]. The FFQ contained 104 items and 10 multiple-choice frequency categories of consumption. We used the USDA National Nutrient Database[23] and the National Institute of Nutrition database for standard reference food composition tables in Mexico[24].

Western dietary pattern. We grouped the 104 food items into 40 food groups following the recommendations of Knol and VanderWeele[25] (Appendix 1); the Western dietary pattern was constructed by principal factor analysis (Appendix 2). The Western dietary pattern has been associated with several diseases, being overweight, and obesity[26,27], so its lowest tertile was considered the least risk. Tertiles were generated based on the distribution among the controls.

Alcohol consumption. We measured alcohol in grams per day (g/d), and we created three categories: (i) did not consume, (ii) consumed less than 1 g/d, and (iii) consumed ≥ 1 g/d. Women who did not drink alcohol were considered to have the lowest risk.

Tobacco smoking. We created a dichotomous variable for tobacco smoking on whether women had smoked > 100 cigarettes in their lifetime (yes/no). Women who never smoked or smoked < 100 cigarettes in their lifetime were considered the low-risk group.

Physical activity. We applied a 7-day recall questionnaire based on a semi-structured interview[28,29] to assess physical activity within the previous 12 months; it included time spent on a physical recreational activity, household work, and sleep. Physical activity was categorized as (i) light-intensity physical activities [1.1–2.9 metabolic equivalents of energy expenditure (METS)], (ii) moderate-intensity physical activities (3.0–5.9 METS), and (iii) vigorous-intensity physical

 doi:10.20944/preprints202209.0360.v1

activities (6 or more METS)[30]. We used only moderate- and vigorous-intensity physical activities (hours per week) for this analysis.

Genetic analysis

DNA was extracted from whole blood, and GWAS genotyping was conducted as part of the Breast Health Disparities Study[22]. We measured European and Native American ancestry based on 104 informative ancestry markers. We chose a pathway-driven approach to select the genes included in the analysis[31]. We evaluated single nucleotide polymorphisms (SNPs) in 12 genes that codify for the first-line response enzymes to OS: *CAT* (43 SNPs); *SOD1 (17 SNPs), SOD2* (39 SNPs), and *SOD3* (17 SNPs); and *GPX 1* (5 SNPs), *GPX2* (7 SNPs), *GPX3* (23 SNPs), *GPX4* (13 SNPs), *GPX5* (5 SNPs), *GPX6* (6 SNPs), *GPX7* (5 SNPs), and *GPX8* (1 SNPs). We selected the SNPs based on a minor allele frequency (MAF) > 1% in the 1000 Genomes database and previous publications obtained from the NCBI dbSNP variation database[32]. Supplementary Table 1 describes the 176 SNPs and their distribution in our population in detail.

Direct counting was used to determine allele and genotype frequencies. Hardy–Weinberg equilibrium (HWE)[33] was tested by the exact test among controls, and SNPs in deviation of HWE were excluded. We also excluded SNPs with less than 6 individuals per genotype. We used logistic regression adjusted by ancestry (% Native American) to compare genotype frequencies between cases and controls; *P*-values were obtained using likelihood ratio tests. A *P*-value < 0.05 was considered statistically significant. As we hypothesized a synergistic effect, only SNPs that showed an association with BC (n = 12) were selected for interaction analyses.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

Linkage disequilibrium was evaluated for the variants in HWE and remained associated after adjustment. We used the SNPStats[34] to assess linkage disequilibrium within the same gene. To assess linkage disequilibrium at a larger scale between the SNPs associated with HeLiX (Arnold et al., 2015)(Arnold et al., 2015) (Arnold et al., 2015) and explore the functional annotation of the variants of interest, we used the free software SNiPA[35]. To analyze the genotype–phenotype correlations of the SNPs of interest, we used Genotype-Tissue Expression (GTEx) Project information ([CSL STYLE ERROR: reference with no printed form.]) [37].

Statistical analyses

We conducted a descriptive analysis of the study population for pre- and postmenopausal status in cases and controls. We used the mean and standard deviation for continuous variables and percentages for categorical variables.

Western dietary patterns. We used the ROTATE = VARIMAX function to rotate the loading matrix using an orthogonal transformation[38]. A factor loading of ≥ 0.58 was used to identify the primary factor in which the items were loaded. We calculated factor scores for the Western dietary pattern for each woman[39]. The final factor scores were calculated by weighting each food category proportionately to its participation in each subject's daily food consumption. The Western dietary pattern index was classified by tertiles, with the lowest tertile having the least risk.

Healthy lifestyle index. Diet, physical activity, alcohol, and tobacco consumption were used to develop HeLiX. We considered it healthy to engage in moderate- and vigorous-intensity physical activity, belong to the lowest tertile of the Western dietary pattern, have never smoked or consumed less than 100 cigarettes, and have never consumed alcohol. The HeLiX was constructed by means of principal components. Given that the variables were ordinal, a polychoric correlation was performed. After obtaining the polychoric matrix, a primary factor analysis was constructed using a regression model with only one component. The Kaiser–Meyer–Olkin test[40] was used to diagnose the statistical model.

 doi:10.20944/preprints202209.0360.v1

We compared cases vs. controls' socioeconomic status (Low, medium, high), history of diabetes, BC family history, smoking ≥100 cigarettes, alcohol consumption per day (never, <1g/d, ≥1g/d), and Healthy lifestyle index (Low, medium and high) using– X^2 test (Table 1). Age (years), breastfeeding, waist-hip ratio (%), height (cm), dietary folate intake, energy intake (kilocalories per day), Native American ancestry (%), moderate-vigorous physical activity, and calories consumed by the Western diet (kilocalories per day) were compared using Kruskal–Wallis rank sum tests. To investigate the effect of the interaction between HeLiX (exposure of interest A) and antioxidant enzyme gene variants (exposure of interest B) on the risk of BC (outcome), we followed the next steps[25,41,42].

First, we estimated the odds ratios (ORs) and 95% confidence intervals (CIs) for the lowest HeLiX tertile (T1 = low) compared to T2 = medium and T3 = high within genotypes (wild type homozygous, heterozygous, and alternative homozygous) and genotypes within strata of HeLiX (Tertiles) by menopausal status. Second, we estimated the joint effect of the lowest tertile of the HeLiX and the highest reference genotype category versus the highest tertile of the HeLiX and the lowest risk genotype category using conditional logistic regression models. T1 and the alternative (less frequent) homozygous were the references for both analyses.

Finally, we calculated the relative excess risk resulting from the interaction (RERI) to estimate the joint effect on the additive scale[43] and the odds ratios (ORs) to estimate the joint effect on the multiplicative scale[41]. We calculated the 95% confidence intervals (95% CI) for additive and multiplicative interactions, applying the delta method implemented in Stata 14 software[40,44,45].

Conditional logistic regression models were adjusted for breastfeeding, age, waist-to-hip ratio, height, folate intake, energy intake, Native American ancestry, socioeconomic status, history of diabetes, and first-grade family history of BC. Given that we performed several statistical tests simultaneously, the *P*-value for each interaction was corrected for

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

multiple comparisons according to the methods described by Bland and Altman[46] ($P_{interaction} < 0.336$). Analyses were performed using STATA v14 software (StataCorp LP, College Station, TX, USA).

Results

The distribution of cases and control characteristics by menopausal status is shown in Table 1. Overall, compared to controls, we observed that more cases reported less breastfeeding ($P < 0.001$), had a higher dietary folate intake ($P = 0.006$), were less physically active (moderate-vigorous physical activity) ($P < 0.001$), consumed more calories from a Western dietary pattern ($P = 0.005$), had the highest socioeconomic level ($P < 0.001$), smoked more than 100 cigarettes ($P = 0.001$), and consumed more alcohol ($P = 0.001$).

Table 1. Characteristics of study controls and BC cases stratified by menopausal status

Characteristics	Premenopausal (n=541)			Postmenopausal (n=774)		
	Controls	Cases	*P*-value	Controls	Cases	*P*-value
Continuous variables						
Age	43.82	43.70	0.800	57.46	58.41	0.046
Breastfeeding (%)	19.95	17.35	0.175	40.51	25.56	<0.001
Waist-hip ratio (%)	91.07	88.66	0.005	92.08	91.24	0.253
Height (cm)	152.76	154.38	0.001	150.85	151.96	0.025
Dietary folate intake (mg/d)	338.54	394.40	<0.001	340.39	370.70	0.006
Energy intake (Kcal/d)	1951.57	2298.39	<0.001	1865.53	2160.20	<0.001
Native American ancestry (%)	65.06	61.40	0.032	64.13	57.98	<0.001
Moderate-vigorous physical activity	17.25	15.62	0.258	15.21	8.87	<0.001
Energy intake by Western dietary pattern (Kcal/d)	1177.75	1368.64	<0.001	1058.57	1197.18	<0.001
Discrete variables						
Socioeconomic status						
Low (T1)	28.47	31.62	0.009	34.95	29.52	<0.001
Medium (T2)	36.86	25.09		34.95	25.24	
High (T3)	34.67	43.30		30.09	45.24	
History of diabetes						
Yes	8.76	9.97	0.623	21.53	30.55	0.003
Family history of breast cancer						
Yes	2.19	2.06	0.916	1.85	5.24	0.007
≥100 cigarettes smoked over lifetime						
Yes	23.36	23.02	0.925	17.59	26.90	0.001
Alcohol consumption per day						
Never	42.70	26.12	ND	46.53	34.37	0.001
<1g/d	57.30	72.16		52.08	63.48	
≥1g/d	0.00	1.72		1.39	2.15	
Healthy lifestyle index						
Low (T1)	34.31	45.86	0.001	34.03	57.18	<0.001
Medium (T2)	29.20	31.03		34.95	27.27	
High (T3)	36.50	23.10		31.02	15.55	

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

In this study, we included 176 SNPs from the family genes *SOD*, *GPX*, and *CAT* (Supplementary Table 1). Twelve variants of the *CAT* gene remained significant for the association with BC after adjustment for Native American Ancestry (*rs475043, rs494024, rs511895, rs533425, rs554576, rs560807, rs2073058, rs2179625, rs2300181, rs7104301, rs7933285, rs12270780*); the allele and genotype distribution and comparison between cases and controls are shown in Table 2.

Table 2. Description of the CAT gene SNPs included for analysis in cases and controls

			Control (n=706)						Cases (n=711)							
		Alleles	Genotype frequency			Allele Frequency		HWE	Genotype frequency			Allele Frequency		Cases versus controls		
Gene	SNP rs	Major/Minor	Homozygous major	Heterozygous	Homozygous minor	Major	Minor	*P*-value	Homozygous major	Heterozygous	Homozygous minor	Major	Minor	*P*-value for allelic frequency	*P*-value for genotipic frequency	*P*-value for genotypic frequency adjusted by ancestry
CAT	rs704724	C/T	698	8	0	0.99	0.01	1.000	703	8	0	0.99	0.01	0.989	0.99	0.965
	rs769217	C/T	367	293	46	0.73	0.27	0.250	354	294	63	0.70	0.30	0.180	0.24	0.360
	rs1049982	T/C	235	350	121	0.58	0.42	0.700	243	323	145	0.57	0.43	0.525	0.19	0.225
	rs475043	T/C	521	167	18	0.86	0.14	0.290	460	223	28	0.80	0.20	<0.001	<0.001	**0.012***
	rs480496	G/A	299	322	85	0.65	0.35	0.930	268	323	120	0.60	0.40	0.009	0.021	0.088
	rs480575	G/A	186	357	163	0.52	0.48	0.760	179	339	193	0.49	0.51	0.164	0.21	0.393
	rs494024	C/T	522	165	19	0.86	0.14	0.170	463	220	28	0.81	0.19	<0.001	0.001	**0.015***
	rs511895	C/T	521	167	18	0.86	0.14	0.290	459	224	28	0.80	0.20	<0.001	<0.001	**0.012***
	rs524154	G/A	190	356	160	0.52	0.48	0.820	184	340	187	0.50	0.50	0.214	0.28	0.466
	rs525938	C/T	190	356	160	0.52	0.48	0.820	182	341	188	0.50	0.50	0.175	0.26	0.460
	rs533425	G/A	522	165	19	0.86	0.14	0.170	463	220	28	0.81	0.19	<0.001	0.001	**0.015***
	rs554576	T/A	232	350	124	0.58	0.42	0.760	191	329	191	0.50	0.50	<0.001	<0.001	**0.001***
	rs560807	T/A	522	165	19	0.86	0.14	0.170	463	220	28	0.81	0.19	<0.001	0.001	**0.015***
	rs564250	C/T	364	287	55	0.72	0.28	0.930	366	282	63	0.71	0.29	0.734	0.750	0.972
	rs566979	C/A	233	350	123	0.58	0.42	0.700	190	331	190	0.50	0.50	<0.001	<0.001	<0.001
	rs769214	G/A	235	350	121	0.58	0.42	0.700	243	323	145	0.57	0.43	0.525	0.19	0.225
	rs769218	G/A	366	293	47	0.73	0.27	0.300	353	294	64	0.70	0.30	0.181	0.24	0.415
	rs1001179	C/T	620	85	**1**	0.94	0.06	0.510	587	112	12	0.90	0.10	0.001	<0.001	**0.029**
	rs1408034	T/C	610	93	**3**	0.93	0.07	1.000	578	125	8	0.90	0.10	0.005	0.019	**0.047**
	rs2073058	A/G	342	304	60	0.70	0.30	0.590	383	260	68	0.72	0.28	0.201	0.044	**0.027***
	rs2179625	A/C	344	302	60	0.70	0.30	0.650	387	257	67	0.73	0.27	0.160	0.038	**0.031***
	rs2284365	T/C	48	297	361	0.72	0.28	0.220	64	303	344	0.70	0.30	0.147	0.250	0.367
	rs2284367	T/C	46	293	367	0.73	0.27	0.250	62	296	353	0.70	0.30	0.180	0.270	0.395
	rs2284369	A/G	366	292	48	0.73	0.27	0.350	352	292	67	0.70	0.30	0.145	0.180	0.358
	rs2300181	C/T	342	304	60	0.70	0.30	0.590	383	260	68	0.72	0.28	0.201	0.044	**0.027***
	rs4755374	A/C	444	235	27	0.80	0.20	0.640	461	217	33	0.80	0.20	0.708	0.440	0.633
	rs4756146	T/C	27	235	444	0.80	0.20	0.640	34	216	461	0.80	0.20	0.743	0.390	0.560
	rs7104301	A/G	340	305	61	0.70	0.30	0.590	385	260	66	0.72	0.28	0.116	0.038	**0.024***
	rs7933285	C/T	342	304	60	0.70	0.30	0.590	383	260	68	0.72	0.28	0.201	0.044	**0.027***
	rs7943316	A/T	240	347	119	0.59	0.41	0.760	245	325	141	0.57	0.43	0.498	0.270	0.328
	rs7947841	G/A	612	92	**2**	0.93	0.07	0.760	583	120	8	0.90	0.10	0.007	0.016	0.046
	rs9282626	T/C	676	29	1	0.98	0.02	0.290	688	23	0	0.98	0.02	0.260	0.340	0.313
	rs10488736	C/T	312	313	81	0.66	0.34	0.870	353	279	79	0.69	0.31	0.097	0.110	0.059
	rs10836235	C/T	611	92	3	0.93	0.07	1.000	580	123	8	0.90	0.10	0.006	0.022	0.045
	rs10836244	G/T	367	293	46	0.73	0.27	0.250	354	294	63	0.70	0.30	0.180	0.240	0.360
	rs11032699	T/G	124	350	232	0.58	0.42	0.760	146	327	238	0.56	0.44	0.526	0.270	0.337
	rs11032700	C/A	125	354	227	0.57	0.43	0.540	151	327	233	0.56	0.44	0.434	0.170	0.243
	rs11032702	C/T	659	45	2	0.97	0.03	0.200	670	41	0	0.97	0.03	0.373	0.220	0.400
	rs11032703	C/T	671	34	1	0.97	0.03	0.370	669	42	0	0.97	0.03	0.511	0.330	0.806
	rs12269988	A/G	676	29	1	0.98	0.02	0.290	683	28	0	0.98	0.02	0.673	0.490	0.758
	rs12270780	G/A	68	308	330	0.69	0.31	0.790	72	267	372	0.71	0.29	0.141	0.063	**0.050***
	rs12273124	A/G	659	46	1	0.97	0.03	0.560	670	41	0	0.97	0.03	0.431	0.490	0.346
	rs12803540	T/C	619	85	2	0.94	0.06	1.000	588	116	7	0.91	0.09	0.005	0.014	0.041

*SNPs selected according to a P-value< 0.05 when adjusted by Native American ancestry, genotype frequencies in controls in Hardy Weingerg Equilibrium (HWE), and no less than 6 individuals in the category.

When we analyzed the interaction among these 12 *CAT* SNPs and HeLiX, the *CAT rs554576* SNP remained significant ($P_{interaction}$ = 0.200; Table 3) after Bland and Altman's correction for multiple comparisons. Interaction analysis showed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

a synergic reduction in the odds of having BC progressively according to the highest tertiles of the HeLiX and the number of protective alleles T per genotype (AT and TT) for both pre and postmenopausal women. The reference category was a low HeLiX (T1) and AA genotype (Supplementary Table 2, Figure 2).

Table 3. *P*-value for interaction of selected single nucleotide polymorphisms of gene CAT and Healthy lifestyle index on breast cancer risk.

	P-value for interaction*	
***CAT* SNPs**	**Premenopausal**	**Postmenopausal**
rs475043	0.999	0.983
rs4940024	0.999	0.988
rs511895	0.999	0.973
rs533425	0.999	0.988
rs554576	0.814	**0.200***
rs560807	0.999	0.988
rs2073058	0.999	0.994
rs2179625	0.999	0.994
rs2300181	0.999	0.994
rs2073058	0.999	0.993
rs2179625	0.999	0.994
rs12270780	0.999	0.999

*** *P*-value for interaction corrected for multiple comparisons according to Bald y Altman (*P*<0.336). Adjusted by age, breastfeeding, waist-hip ratio, folate and calorie consumption, Native American ancestry, physical activity, socio-economical level, diabetes and family history of breast cancer.**

The effect of a synergistic interaction was more evident in postmenopausal than premenopausal women. Compared to women in the reference category, those women in the highest HeLiX tertile (T3) and TT or AT genotypes showed 0.15 times the odds of BC (OR = 0.15, 95% CI 0.07–0.32). Compared to women in the reference category, those women in the medium tertile (T2) and TT or *AT* genotypes also showed a protective effect (OR = 0.26 95% CI 0.13–0.54; OR = 0.24, 95% CI 0.12–0.45, respectively; Table 4 Fig. 2). For premenopausal women, compared with those women in the lowest tertile of HeLiX (T1) and AA genotype (reference category), we observed an odds reduction at the highest tertile

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

of HeLiX (T3) in women with the TT (OR 0.21, 95% CI 0.08–0.51) and AT genotypes (OR = 0.29, 95% CI 0.13–0.62). For the medium tertile of HeLiX (T2), we also observed an inverse association for women with the TT genotype (OR = 0.39 95% CI 0.17–0.87) compared with those women in the reference category (Supplementary Table 2, Fig. 2).

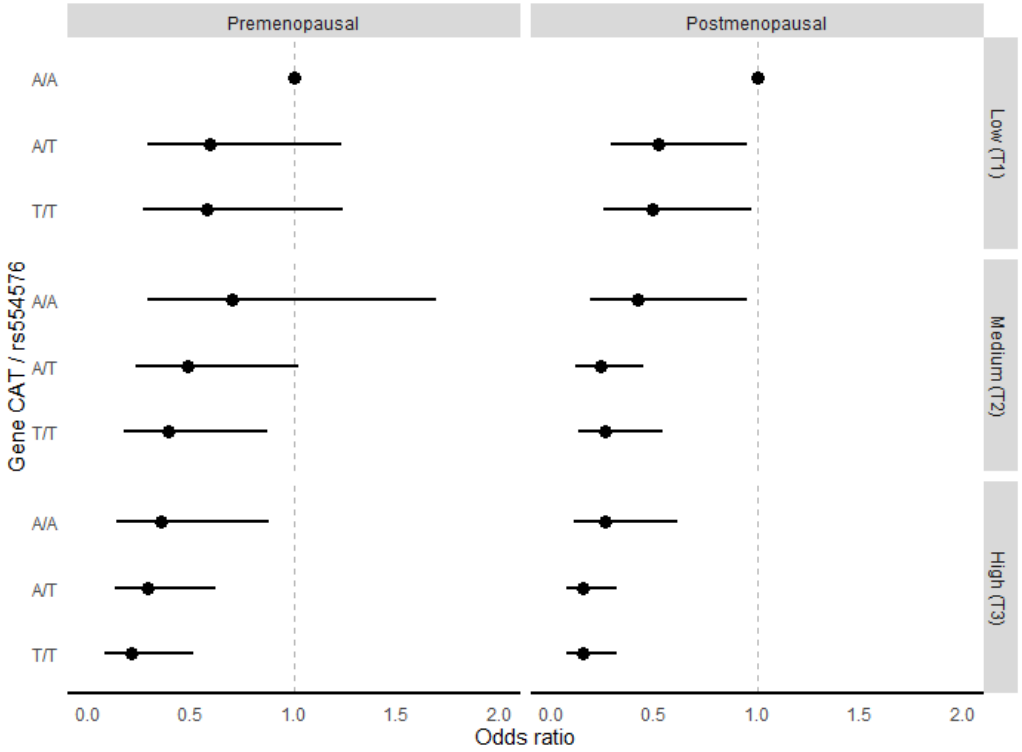

Figure 2. Synergistic interaction between HeLiX tertiles and *CAT rs554576* genotypes among pre- and postmenopausal women.

There was a significant interaction of the HeLiX and *CAT rs554576* on an additive scale, but not on the multiplicative scale in postmenopausal women (Table 4). The results stratified by genotype showed a significant association between the medium tertile (T2) of the HeLiX and the AA genotype (OR = 0.53, 95% CI 0.29–0.97), compared with those women in the lowest tertile of HeLiX (T1) and AA genotype. When stratified by the HeLiX, we observed a reduced risk of BC in each tertile of the index when having at least one *T* allele at *CAT rs554576*, and the effect increased in a dose-dependent

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

manner (Table 4). Our results showed that the estimated combined effect was greater than the product of the estimated effects alone, even at the lowest tertile of the HeLiX (T1).

Table 4. Evaluation of the type of interaction of the CAT rs554576 SNP and the Healthy Lifestyle Index on the risk of breast cancer in postmenopausal women

Gene *CAT / rs554576*	Healthy Lifestyle Index Tertiles														ORs (95% CI) for Healthy lifestyle index (Medium and High versus Low) among *CAT rs554576* genotype strata					
	Low (T1)					Medium (T2)					High (T3)					Medium (T2)			High (T3)		
Genotype	Cases	Controls	OR	95% CI	*P* -value	Cases	Controls	OR	95% CI	*P* -value	Cases	Controls	OR	95% CI	*P* -value	OR	95% CI	*P* -value	OR	95% CI	*P* -value
A/A	25	64	1.00			22	29	0.42	0.19 - 0.95	0.038	25	20	0.26	0.11 - 0.61	0.002	0.53	0.29 - 0.97	0.040	0.51	0.26 - 1.01	0.056
A/T	81	123	0.52	0.29 - 0.95	0.034	80	51	0.24	0.12 - 0.45	<0.001	60	27	0.15	0.07 - 0.32	<0.001	0.54	0.25 - 1.13	0.106	0.56	0.25 - 1.26	0.165
T/T	41	52	0.49	0.25 - 0.97	0.041	49	34	0.26	0.13 - 0.54	<0.001	49	18	0.15	0.07 - 0.32	<0.001	0.58	0.25 - 1.32	0.200	0.56	0.23 - 1.36	0.201
														P -value for interaction	0.111						
ORs (95% CI) for genotype (T/T, A/T vs. A/A) among Healthy Lifestyle Index strata																					
A/T			0.35	0.14 - 0.85	0.021			0.46	0.27 - 0.77	0.003			0.48	0.025 - 0.93	0.030						
T/T			0.23	0.09 - 0.61	0.003			0.29	0.15 - 0.54	<0.001			0.28	0.13 - 0.58	0.001						
Measure of interaction on additive scale: RERI (95% CI)								0.33	-0.13 - 0.80	0.167			0.33	-0.16 - 0.82	0.187						
								0.38	-0.01 - 0.79	0.059			0.32	-0.15 - 0.79	0.186						
Measure of interaction on multiplicative scale: ratio of ORs (95% CI)								1.06	0.06 - 2.07	0.897			1.25	0.03 - 2.55	0.696						
								1.12	0.004 - 2.24	0.829			1.13	0.10 - 2.37	0.831						

* Interaction *P* -value significance according to Bald y Altman methodology for multiple comparissons *P* <0.336 when at least one category is different and Bonferroni adjusment with reference *P* = 0.00625. Model adjusted by age, hip-waist ratio, Native American ancestry, history of diabetes and family history of breast cancer.

Discussion

The study findings support our hypothesis that healthy lifestyles and antioxidant defense genes interact synergistically to reduce BC risk. The HeLiX was created by combining the effects of physical activity[47], dietary pattern[48], tobacco smoking[49,50], and alcohol consumption[51], epidemiologically and experimentally proven exogenous sources of free radicals, OS, and risk factors for BC[52–55]. The genes were selected according to the role they play in the endogenous capacity of the body to respond to exogenous OS: first antioxidant response *SOD, GPX*, and *CAT genes*.

Of the 12 tested genes, only *CAT* showed an association with BC, consistent with previous reports[56–60]. In this study, we observed that catalase and not superoxide dismutase or glutathione peroxidase act synergistically with lifestyle. This might be partially explained by the reactive function of catalase,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

which counteracts a high-fat diet, and its activity seems to be responsive to nutrition and weight improvements[61].

Moreover, there was a synergistic effect with lifestyle, as shown by a decrease in the odds of BC as the tertiles of the HeLiX and the number of protective alleles increased (*CAT rs554576* T-allele). The *CAT* gene is expressed in most human body cells; however, it has higher expression in adipose and breast tissues[36]. Variant *CAT rs554576* is a polymorphism located at intron 9[62], and it is described as a modifier variant[63]. This variant is an expression quantitative trait loci (eQTL) with a cis-regulatory effect on *CAT* itself and the upstream gene *ABTB2*[37]. For both genes, the AA genotype has a higher normalized expression in blood than the AT and TT genotypes[64]. As *rs554576* is linked to *ABTB2*, the effects we observed might also be associated with *ABTB2* function, which mediates protein–protein interactions and interplays with E3 ubiquitin ligases RBX1 and CUL3, which regulate antioxidant response by NFE2-related factor 2. *ABTB2* has been linked as a predictor of neoadjuvant chemotherapy with the epirubicin response in Luminal-A BC[65]. It is described as a gene susceptible to regulation by noncoding variants in cancer genomes[66].

CAT rs554576 is in linkage disequilibrium ($r^2 > 80$) with nearby catalase variants, *rs2103711* (r2 = 0.81) *and rs2103710* (r2 = 0.81), which are modifier variants with cis-regulatory effects on *CAT* and *ABTB2* genes and have a Combined Annotation Dependent Depletion (CADD) score of 9 and 4.5, respectively. The variants are associated with serum levels of phosphatidylcholine (PC) aa C34:3 ($P = 3.84 \times 10^{-5}$)[63]. PCs have been associated with BC risk[67] and mammographic density. However, in a study of premenopausal Mexican women, this association did not remain significant after correction for multiple testing[68]. The rest of the variants in linkage disequilibrium, *rs11032727* ($r^2 = 0.87$), *rs4756154* (r^2

 doi:10.20944/preprints202209.0360.v1

= 0.87), and *rs3781710* (r^2 = 0.84), are also modifier variants, with a discrete cis-regulatory effect on *CAT* and *ABTB2* in the blood and thyroid[63] .

The strength of the synergistic association found in this study was greater in postmenopausal women, which might be explained by mechanisms related to estrogen depletion and aging. The overproduction of free radicals occurs during the aging process, making the human body vulnerable to various age-related pathologies, such as cancer[20]. The critical reduction in estrogen levels during menopause has been shown to increase OS levels in the body; at low concentrations, estrogen has pro-oxidant-like effects, mainly when its chemical structure contains catechol[69]. Additionally, menopause can induce differential expression of antioxidant genes, with a higher expression of the *CAT* gene[21]. Consumption of antioxidants, such as vitamin C, vitamin E, polyphenols, and lycopene, might benefit women in the peri- and postmenopausal phases [70,71].

In this study, higher HeLiX tertiles were related to daily half-hour vigorous- or moderate-intensity physical activity, reduced/no consumption of alcohol, fat, processed foods, refined cereals, complex sugars, and no tobacco smoking. As body mass index (BMI) is not a component of a healthy lifestyle but rather a result of a lifestyle, we did not include it in the HeLiX[15]. However, we adjusted for the waist-to-hip ratio in our final multiple models because it measures adiposity better than BMI[72].

There is consistent epidemiological evidence about the protective effects of physical activity against pre-and postmenopausal BC. In the metanalysis by Xuyu Chen et al., BC's overall relative risk (ORR) was 0.87 (CI95% 0.84–0.90) in women who practiced physical activity. The risk reduction was similar among individuals with a BMI < 25 kg/m2 (0.88; CI 95% 0.83–0.93) compared to a BMI ≥ 25 kg/m2 (0.87; CI 95% 0.77–0.97). The ORR was reduced by 3% (95%

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

CI 0.95–0.99) for every 10 METS per week increment in recreational physical activity and by 2% (CI 95% 0.97–0.99) for every 10 METS per week increment in total physical activity[52]. Similar results have been found with regard to the protective effects of physical activity for BC in Mexican women; Ángeles-Llerenas et al. observed a decreased risk of BC in pre- and postmenopausal women (OR = 0.96; CI 95% 0.92–0.99; OR = 0.90; CI 95% 0.86–0.93, respectively) for every 3 hours per week of moderate-intensity physical activity[16].

The biological mechanisms of how physical activity reduces BC risk may include adiposity, regulation of sex hormones, insulin sensitivity, immune system regulation, inflammation (Mc et al., 2017)(Mc et al., 2017) (Mc et al., 2017), and (Mc et al., 2017)(Mc et al., 2017) (Mc et al., 2017) OS[73]. In the short term, exercise increases ROS production; however, when practiced regularly, it enhances the body's antioxidant capacity, protecting against subsequent oxidative stress. There is evidence of a synergic effect between the antioxidant gene *CAT* and physical activity. Data from the Long Island Breast Cancer Study Project found that women with ≥ 1 variant allele in *CAT rs4756146* had a 23% reduced risk of postmenopausal breast cancer than women with the common TT genotype (OR = 0.77; CI 95% 0.59–0.99)[12].

Researchers from the Sister Study found significant inverse associations between F2-isoprostane, a biomarker of OS, and physical activity, consumption of vegetables, fruit, and antioxidant vitamins[74].

The Western dietary pattern is high in pro-oxidant foods, such as processed meats, dairy, fat, and sugary foods and drinks, contrasting with low amounts of fruits, vegetables, and fiber[75]. This dietary pattern is associated with increased BC risk; Yunjun Xiao et al. described an overall 14% increased risk (RR 1.14, CI 95% 1.02–1.28), and the association was consistent across different ethnicities[76], including Hispanic women[77].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 September 2022 doi:10.20944/preprints202209.0360.v1

Unique metabolite profiles linked to OS have been described in comparative studies of dietary intake[75]. The oxidative potential of the Western dietary pattern has been related to its high glycemic index and glycemic load in association with the concentration of oxidative biomarkers phosphatidylethanolamine plasmalogens[75] and urinary F2-IsoP and 15-F2t-IsoP-M[78]. Saturated fat intake also increases ROS generation, p47phox expression, and plasma concentration of thiobarbituric acid-reactive species, markers of lipid peroxidation[79]. A modest reduction in caloric intake has been shown to rapidly reduce OS biomarkers[80–83] and DNA damage[84].

Alcohol consumption increases the burden of BC, causing 8.7% of its incidence (Shield et al., 2016)(Shield et al., 2016) (Shield et al., 2016)[85]. There is a causal link between alcohol intake and BC, mediated through biological pathways: regulation of estrogen levels and associated signaling pathways, the metabolism of ethanol resulting in carcinogens, such as acetaldehyde, the inhibition of one-carbon metabolism, and OS[86].

Alcohol-metabolizing enzymes are expressed in breast cells, indicating that alcohol can be metabolized in mammary tissue. These reactions produce oxidative species, including superoxide anions, hypochlorite ions, hydroxyl radicals, and hydrogen peroxide[87]. Alcohol is mainly cleared by aldehyde dehydrogenase (ALDH), alcohol dehydrogenase (ADH), cytochrome P450 (CYP2E1), and catalase[88]. Catalase, located in peroxisomes, can oxidize ethanol *in vitro* in the presence of hydrogen peroxide. There is a high correlation between the frequency of alcohol drinking and the activity of catalase and ALDH ($r = 0.68$, $p < 0.0001$) in animal models[89]. Never drinking alcohol was associated with a BC risk reduction of 44% (0.56; CI 95% 0.35–0.91) and 21% (0.79; CI 95% O.40–1.00) in postmenopausal and premenopausal Mexican women, respectively[15]. The combination of long-term cigarette smoking and alcohol consumption appears to contribute to severe oxidative imbalances and DNA damage[90].

 doi:10.20944/preprints202209.0360.v1

Growing epidemiological evidence suggests that smoking increases BC risk by 10–40%[54]. The multiethnic cohort described that smoking increased the risk by 35% (HR = 1.35, CI95% 1.13-1.63). There was no difference between ethnic groups[91] or tumor receptor status[92]. The Generations Study cohort found a smaller risk (HR = 1.14, CI95% 1.03–1.25, P = 0.010); however, this risk increased when smoking started at younger ages (HR = 1.24, CI95% 1.08-1.43) and in women with a family history of BC (HR 1.35, CI 95% 1.12-1.62)[93]. Carcinogen products of tobacco metabolism might act as estrogen disruptors. There seems to be a higher effect of cigarette smoking during breast development, reinforcing the importance of smoking prevention, especially in early adolescence[94].

A healthy lifestyle is protective, even when there is a high genetic risk of cancer. A study described a joint effect of genetic and lifestyle factors on overall cancer risk (HR = 2.38; CI 95% 2.05–2.76) for women with high genetic risk and unhealthy lifestyles compared to those with low genetic risk and healthy lifestyles. Among women at high genetic risk, the standardized 5-year cancer incidence was significantly reduced from 5.77- 3.69% for women with a healthy lifestyle[95].

Finally, our study has design strengths. First, control subjects were selected from the same base population where the cases occurred, such that if they were cases, they would have been selected for our study. Second, anthropometric measurements, blood samples, and interviews were conducted by trained nurses according to standardized protocols. Third, the women in this study belonged to three diverse subpopulations in Mexico, and genetic ancestry was considered for adjustment. Finally, we constructed an index based on dietary patterns and healthy habits rather than on consumption of micronutrients, which makes recommendations easier for the general population to understand.

Regarding the limitations, in postmenopausal women, we observe an interaction of the HELiX and *CAT rs554576*. However, this interaction does not reach statistical significance; these results might be due to a smaller sample size

 doi:10.20944/preprints202209.0360.v1

compared to postmenopausal women. For the HeLiX, composed by lifestyle, there might be a recall bias, an inherent limitation of case–control studies. To address this bias, we only include incident BC. This study only included SNPs; we did not include Copy Number Variants, so we do not discard that this type of genetic variation has an effect on antioxidant response.

Conclusions

Our study shows evidence for the interplay of lifestyle and genetic factors and contributes to the understanding of the pathways involved in BC etiology. The results show that following a healthy lifestyle can interact synergistically with the endogenous antioxidant capacity of the body to reduce BC risk.

Rapid genomic advances make it easier to identify individuals at high risk of BC to implement modifiable interventions for primary and secondary prevention. Genomic profiling, an emerging technology in epidemiology, may offer insights into the biological mechanisms linking dietary patterns with oxidative stress as underlying mechanisms for BC incidence and survival.

Personal genetic makeup can modulate our capacity to perform oxidative detoxification. Following a healthy lifestyle by reducing calorie ingestion from the Western diet, practicing moderate physical activity 30 minutes five times a week, and avoiding cigarette smoking and alcohol consumption reduces BC risk. This study highlights the significance of putting governmental policies into action to reduce sedentary behavior, unhealthy diet, and alcohol and tobacco consumption. These measures might be more useful in population groups with specific molecular makeup.

 doi:10.20944/preprints202209.0360.v1

Acknowledgments

We thank the health workers in the participating hospitals for their support: Germán Castelazo (IMSS, Mexico City, DF), Sinhué Barroso Bravo (IMSS, Mexico City, DF), Joaquín Zarco Méndez (ISSSTE, Mexico City, DF), Edelmiro Pérez Rodríguez (University Hospital, Monterrey, Nuevo León), Jesús Pablo Esparza Cano (IMSS, Monterrey, Nuevo León), Heriberto Fabela (IMSS, Monterrey, Nuevo León), Fausto Hernéndez Morales (ISSSTE, Veracruz, Veracruz), Pedro Coronel Brizio (CECAN SS, Xalapa, Veracruz), Vicente A. Saldaña Quiroz (IMSS, Veracruz, Veracruz), and Fernando Mainero-Ratchelous (IMSS, Mexico City). We also thank María Felix Lazcano López, Silvia Cardoso Muñoz, Libia Zulema Franco Velázquez, and Jenny Tejeda Espinoza.

Author contribution

LGFR: conceptualization, writing of the manuscript, genetic analysis, and interpretation of data; LMSZ: conceptualization, formal analysis, data interpretation, critical review of the manuscript; AAL, RRV: critical reading and review of the manuscript; GTM: conceptualization, acquisition of data, and funding, critical review, and supervision.

Financing

This work was financed by the Mexican National Council of Science and Technology (SALUD-2002-C01-7462). The authors declare no conflicts of interest.

Bibliography

1. DeSantis, C. E. *et al.* International Variation in Female Breast Cancer Incidence and Mortality Rates. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 24, 1495–1506 (2015).

2. Lee, J. D., Cai, Q., Shu, X. O. & Nechuta, S. J. The Role of Biomarkers of Oxidative Stress in Breast Cancer Risk and Prognosis: A Systematic Review of the Epidemiologic Literature. *J. Womens Health 2002* 26, 467–482 (2017).
3. Dröge, W. Free radicals in the physiological control of cell function. *Physiol. Rev.* 82, 47–95 (2002).
4. Alkadi, H. A Review on Free Radicals and Antioxidants. *Infect. Disord. Drug Targets* 20, 16–26 (2020).
5. Sharma, G., Gupta, G. & Sharma, P. A Comprehensive Review of Free Radicals, Antioxidants, and Their Relationship with Human Ailments. *Critical reviews in eukaryotic gene expression* https://pubmed.ncbi.nlm.nih.gov/30055541/ (2018) doi:10.1615/CritRevEukaryotGeneExpr.2018022258.
6. Matés, J. M., Segura, J. A., Alonso, F. J. & Márquez, J. Intracellular redox status and oxidative stress: implications for cell proliferation, apoptosis, and carcinogenesis. *Arch. Toxicol.* 82, 273–299 (2008).
7. Ziech, D., Franco, R., Pappa, A. & Panayiotidis, M. I. Reactive Oxygen Species (ROS)—Induced genetic and epigenetic alterations in human carcinogenesis. *Mutat. Res. Mol. Mech. Mutagen.* 711, 167–173 (2011).
8. Truong, T. H. & Carroll, K. S. Redox Regulation of Epidermal Growth Factor Receptor Signaling through Cysteine Oxidation. *Biochemistry* 51, 9954–9965 (2012).
9. Devi, G. R., Allensworth, J. L., Evans, M. K. & Sauer, S. J. Chapter 1 - The Role of Oxidative Stress in Breast Cancer. in *Cancer* (ed. Preedy, V.) 3–14 (Academic Press, 2014). doi:10.1016/B978-0-12-405205-5.00001-5.
10. Ighodaro, O. M. & Akinloye, O. A. First line defence antioxidants-superoxide dismutase (SOD), catalase (CAT) and glutathione peroxidase (GPX): Their fundamental role in the entire antioxidant defence grid. *Alex. J. Med.* 54, 287–293 (2018).
11. Gallegos-Arreola, M. P., Ramírez-Hernández, M. A., Figuera, L. E., Zúñiga-González, G. M. & Puebla-Pérez, A. M. The rs2234694 and 50 bp Insertion/Deletion polymorphisms of the SOD1 gene are associated with breast cancer risk in a Mexican population. *Eur. Rev. Med. Pharmacol. Sci.* 24, 8017–8027 (2020).
12. McCullough, L. *et al.* Polymorphisms in oxidative stress genes, physical activity, and breast cancer risk. *Cancer Causes Control CCC* 23, (2012).
13. Mahasneh, A. A., Zhang, Y., Zhao, H., Ambrosone, C. B. & Hong, C.-C. Lifestyle predictors of oxidant and antioxidant enzyme activities and total antioxidant capacity in healthy women: a cross-sectional study. *J. Physiol. Biochem.* 72, 745–762 (2016).
14. Li, Y. *et al.* Oxidative stress-related genotypes, fruit and vegetable consumption and breast cancer risk. *Carcinogenesis* 30, 777–784 (2009).

15. Sánchez-Zamorano, L. M. *et al.* Healthy lifestyle on the risk of breast cancer. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 20, 912–922 (2011).
16. Angeles-Llerenas, A. *et al.* Moderate physical activity and breast cancer risk: the effect of menopausal status. *Cancer Causes Control CCC* 21, (2010).
17. Willett, W. *Nutritional Epidemiology*. (Oxford University Press, 2012).
18. Becerra-Fernández, A. *La edad de la menopausia*. (Ediciones Díaz de Santos, 2003).
19. Sánchez-Zamorano, L. M. *et al.* Healthy lifestyle on the risk of breast cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* https://pubmed.ncbi.nlm.nih.gov/21335508/ (2011) doi:10.1158/1055-9965.EPI-10-1036.
20. Doshi, S. B. & Agarwal, A. The role of oxidative stress in menopause. *J. -Life Health* 4, 140–146 (2013).
21. Rangel-Zuñiga, O. A. *et al.* Differential menopause- versus aging-induced changes in oxidative stress and circadian rhythm gene markers. *Mech. Ageing Dev.* 164, 41–48 (2017).
22. Hernández-Avila, M. *et al.* Validity and reproducibility of a food frequency questionnaire to assess dietary intake of women living in Mexico City. *Salud publica de Mexico* https://pubmed.ncbi.nlm.nih.gov/9617194/ (1998) doi:10.1590/s0036-36341998000200005.
23. USDA National Nutrient Database for Standard Reference, Legacy Release | Ag Data Commons. https://data.nal.usda.gov/dataset/usda-national-nutrient-database-standard-reference-legacy-release.
24. Morales de León J BV, Bourges H, Camacho PM. Tables of composition of Mexcian foods. (2000).
25. Knol, M. J. & VanderWeele, T. J. Recommendations for presenting analyses of effect modification and interaction. *Int. J. Epidemiol.* 41, 514–520 (2012).
26. Wiseman, M. The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *The Proceedings of the Nutrition Society* https://pubmed.ncbi.nlm.nih.gov/18452640/ (2008) doi:10.1017/S002966510800712X.
27. Kopp, W. How Western Diet And Lifestyle Drive The Pandemic Of Obesity And Civilization Diseases. *Diabetes, metabolic syndrome and obesity : targets and therapy* https://pubmed.ncbi.nlm.nih.gov/31695465/ (2019) doi:10.2147/DMSO.S216791.
28. Sallis, J. *et al.* Physical activity assessment methodology in the Five-City Project. *American journal of epidemiology* https://pubmed.ncbi.nlm.nih.gov/3964995/ (1985) doi:10.1093/oxfordjournals.aje.a113987.

29. Pereira, M. *et al.* A collection of Physical Activity Questionnaires for health-related research. *Medicine and science in sports and exercise* https://pubmed.ncbi.nlm.nih.gov/9243481/ (1997).
30. Ainsworth, B. *et al.* Compendium of physical activities: an update of activity codes and MET intensities. *Med. Sci. Sports Exerc.* 32, (2000).
31. Thomas, D. Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Annual review of public health* https://pubmed.ncbi.nlm.nih.gov/20070199/ (2010) doi:10.1146/annurev.publhealth.012809.103619.
32. Home - SNP - NCBI. https://www.ncbi.nlm.nih.gov/snp/.
33. Namipashaki, A., Razaghi-Moghadam, Z. & Ansari-Pour, N. The Essentiality of Reporting Hardy-Weinberg Equilibrium Calculations in Population-Based Genetic Association Studies. *Cell journal* https://pubmed.ncbi.nlm.nih.gov/26199897/ (2015) doi:10.22074/cellj.2016.3711.
34. Solé, X., Guinó, E., Valls, J., Iniesta, R. & Moreno, V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics* 22, 1928–1929 (2006).
35. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* 31, 1334–1336 (2015).
36. GTEx Portal. https://gtexportal.org/home/gene/ENSG00000121691.4.
37. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (New York, N.Y.)* https://pubmed.ncbi.nlm.nih.gov/25954001/ (2015) doi:10.1126/science.1262110.
38. Fung, T. *et al.* Dietary patterns and the risk of postmenopausal breast cancer. *International journal of cancer* https://pubmed.ncbi.nlm.nih.gov/15756679/ (2005) doi:10.1002/ijc.20999.
39. Sieri, S. *et al.* Dietary patterns and risk of breast cancer in the ORDET cohort. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* https://pubmed.ncbi.nlm.nih.gov/15066921/ (2004).
40. Stata. *Multion Consulting* https://multion.com/producto/stata.
41. VanderWeele, T. J. & Knol, M. J. A Tutorial on Interaction. *Epidemiol. Methods* 3, (2014).
42. Vandenbroucke, J. *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS medicine* https://pubmed.ncbi.nlm.nih.gov/17941715/ (2007) doi:10.1371/journal.pmed.0040297.

43. Kenneth J. Rothman. *Modern Epidemiology*. (1986).
44. Hosmer, D. W. & Lemeshow, S. Confidence Interval Estimation of Interaction. *Epidemiology* 3, 452–456 (1992).
45. Oehlert, G. W. A Note on the Delta Method. *Am. Stat.* 46, 27–29 (1992).
46. Bland, J. M. & Altman, D. G. Multiple significance tests: the Bonferroni method. *BMJ* 310, 170 (1995).
47. Boccatonda, A., Tripaldi, R., Davì, G. & Santilli, F. Oxidative Stress Modulation Through Habitual Physical Activity. *Current pharmaceutical design* https://pubmed.ncbi.nlm.nih.gov/27072166/ (2016) doi:10.2174/1381612822666160413123806.
48. Ávila-Escalante, M., Coop-Gamas, F., Cervantes-Rodríguez, M., Méndez-Iturbide, D. & Aranda-González, I. The effect of diet on oxidative stress and metabolic diseases-Clinically controlled trials. *Journal of food biochemistry* https://pubmed.ncbi.nlm.nih.gov/32160647/ (2020) doi:10.1111/jfbc.13191.
49. Dikalov, S. *et al.* Tobacco smoking induces cardiovascular mitochondrial oxidative stress, promotes endothelial dysfunction, and enhances hypertension. *American journal of physiology. Heart and circulatory physiology* https://pubmed.ncbi.nlm.nih.gov/30608177/ (2019) doi:10.1152/ajpheart.00595.2018.
50. Cao, C. *et al.* Smoking-promoted oxidative DNA damage response is highly correlated to lung carcinogenesis. *Oncotarget* https://pubmed.ncbi.nlm.nih.gov/26942876/ (2016) doi:10.18632/oncotarget.7810.
51. Wu, D. & Cederbaum, A. Alcohol, oxidative stress, and free radical damage. *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism* https://pubmed.ncbi.nlm.nih.gov/15540798/ (2003).
52. Chen, X., Wang, Q., Zhang, Y., Xie, Q. & Tan, X. Physical Activity and Risk of Breast Cancer: A Meta-Analysis of 38 Cohort Studies in 45 Study Reports. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* https://pubmed.ncbi.nlm.nih.gov/30661625/ (2019) doi:10.1016/j.jval.2018.06.020.
53. Xiao, Y. *et al.* Associations between dietary patterns and the risk of breast cancer: a systematic review and meta-analysis of observational studies. *Breast cancer research : BCR* https://pubmed.ncbi.nlm.nih.gov/30696460/ (2019) doi:10.1186/s13058-019-1096-1.
54. Macacu, A., Autier, P., Boniol, M. & Boyle, P. Active and passive smoking and risk of breast cancer: a meta-analysis. *Breast cancer research and treatment* https://pubmed.ncbi.nlm.nih.gov/26546245/ (2015) doi:10.1007/s10549-015-3628-4.

55. Jung, S. *et al.* Alcohol consumption and breast cancer risk by estrogen receptor status: in a pooled analysis of 20 studies. *International journal of epidemiology* https://pubmed.ncbi.nlm.nih.gov/26320033/ (2016) doi:10.1093/ije/dyv156.
56. Ahn, J. *et al.* Associations between breast cancer risk and the catalase genotype, fruit and vegetable consumption, and supplement use. *Am. J. Epidemiol.* 162, 943–952 (2005).
57. Sen, S., Kawahara, B. & Chaudhuri, G. Maintenance of higher H_2O_2 levels, and its mechanism of action to induce growth in breast cancer cells: important roles of bioactive catalase and PP2A. *Free Radic. Biol. Med.* 53, 1541–1551 (2012).
58. Li, Y. *et al.* Oxidative stress-related genotypes, fruit and vegetable consumption and breast cancer risk. *Carcino-genesis* 30, 777–784 (2009).
59. Quick, S. K. *et al.* Effect modification by catalase genotype suggests a role for oxidative stress in the association of hormone replacement therapy with postmenopausal breast cancer risk. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 17, 1082–1087 (2008).
60. Kakkoura, M. G. *et al.* MnSOD and CAT polymorphisms modulate the effect of the Mediterranean diet on breast cancer risk among Greek-Cypriot women. *Eur. J. Nutr.* 55, 1535–1544 (2016).
61. Oliveras-López, M.-J. *et al.* Antioxidant activity and nutritional status in anorexia nervosa: effects of weight recovery. *Nutrients* 7, 2193–2208 (2015).
62. rs554576 RefSNP Report - dbSNP - NCBI. https://www.ncbi.nlm.nih.gov/snp/rs554576?vertical_tab=true.
63. SNiPA - a single nucleotide polymorphisms annotator and browser. https://snipa.helmholtz-muenchen.de/snipa3/in-dex.php?task=variant_browser.
64. GTEx Portal. https://gtexportal.org/home/snp/rs554576.
65. Gong, Y. *et al.* ABTB2 Regulatory Variant as Predictor of Epirubicin-Based Neoadjuvant Chemotherapy in Luminal A Breast Cancer. *Front. Oncol.* 10, 571517 (2020).
66. Liu, Y. *et al.* Discovery of regulatory noncoding variants in individual cancer genomes by using cis-X. *Nat. Genet.* 52, 811–818 (2020).
67. Yuan, B. Identification of Plasma Metabolites Associated with Breast and Ovarian Cancer and Breast Cancer Prog-nosis. https://archiv.ub.uni-heidelberg.de/volltextserver/28690/ (2020) doi:DOI:10.11588/heidok.00028690.
68. His, M. *et al.* Biomarkers of mammographic density in premenopausal women. *Breast Cancer Res. BCR* 23, (2021).

69. Wang, Z. *et al.* Redox cycling of catechol estrogens generating apurinic/apyrimidinic sites and 8-oxo-deoxyguanosine via reactive oxygen species differentiates equine and human estrogens. *Chem. Res. Toxicol.* 23, 1365–1373 (2010).
70. Fernandez-Lazaro, C. I. *et al.* Dietary Antioxidant Vitamins and Minerals and Breast Cancer Risk: Prospective Results from the SUN Cohort. *Antioxid. Basel Switz.* 10, 340 (2021).
71. Ko, S.-H. & Kim, H.-S. Menopause-Associated Lipid Metabolic Disorders and Foods Beneficial for Postmenopausal Women. *Nutrients* 12, E202 (2020).
72. Swainson, M. G., Batterham, A. M., Tsakirides, C., Rutherford, Z. H. & Hind, K. Prediction of whole-body fat percentage and visceral adipose tissue mass from five anthropometric variables. *PloS One* 12, e0177175 (2017).
73. de Boer, M., Wörner, E., Verlaan, D. & van Leeuwen, P. The Mechanisms and Effects of Physical Activity on Breast Cancer. *Clin. Breast Cancer* 17, (2017).
74. Anderson, C., Milne, G., Sandler, D. & Nichols, H. Oxidative stress in relation to diet and physical activity among premenopausal women. *Br. J. Nutr.* 116, (2016).
75. Chandler, P. D. *et al.* Metabolic signatures associated with Western and Prudent dietary patterns in women. *Am. J. Clin. Nutr.* 112, 268–283 (2020).
76. Castelló, A. *et al.* Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. *Maturitas* 103, (2017).
77. Murtaugh, M. A. *et al.* Diet patterns and breast cancer risk in Hispanic and non-Hispanic white women: the Four-Corners Breast Cancer Study. *Am. J. Clin. Nutr.* 87, 978–984 (2008).
78. Anderson, C., Milne, G. L., Park, Y.-M. M., Sandler, D. P. & Nichols, H. B. Dietary Glycemic Index and Glycemic Load Are Positively Associated with Oxidative Stress among Premenopausal Women. *J. Nutr.* 148, 125–130 (2018).
79. Biobaku, F., Ghanim, H., Batra, M. & Dandona, P. Macronutrient-Mediated Inflammation and Oxidative Stress: Relevance to Insulin Resistance, Obesity, and Atherogenesis. *J. Clin. Endocrinol. Metab.* 104, 6118–6128 (2019).
80. Buchowski, M. S. *et al.* Effect of modest caloric restriction on oxidative stress in women, a randomized trial. *PloS One* 7, e47079 (2012).
81. Johnson, J. B. *et al.* Alternate day calorie restriction improves clinical findings and reduces markers of oxidative stress and inflammation in overweight adults with moderate asthma. *Free Radic. Biol. Med.* 42, 665–674 (2007).

82. López-Torres, M. & Barja, G. [Calorie restriction, oxidative stress and longevity]. *Rev. Espanola Geriatr. Gerontol.* 43, 252–260 (2008).

83. Capó, X. *et al.* Calorie Restriction Improves Physical Performance and Modulates the Antioxidant and Inflammatory Responses to Acute Exercise. *Nutrients* 12, E930 (2020).

84. Heilbronn, L. K. *et al.* Effect of 6-month calorie restriction on biomarkers of longevity, metabolic adaptation, and oxidative stress in overweight individuals: a randomized controlled trial. *JAMA* 295, 1539–1548 (2006).

85. Shield, K. D., Soerjomataram, I. & Rehm, J. Alcohol Use and Breast Cancer: A Critical Review. *Alcohol. Clin. Exp. Res.* 40, 1166–1181 (2016).

86. Liu, Y., Nguyen, N. & Colditz, G. A. Links between alcohol consumption and breast cancer: a look at the evidence. *Womens Health Lond. Engl.* 11, 65–77 (2015).

87. Wang, Y., Xu, M., Ke, Z.-J. & Luo, J. Cellular and molecular mechanisms underlying alcohol-induced aggressiveness of breast cancer. *Pharmacol. Res.* 115, 299–308 (2017).

88. Zakhari, S. Overview: how is alcohol metabolized by the body? *Alcohol Res. Health J. Natl. Inst. Alcohol Abuse Alcohol.* 29, 245–254 (2006).

89. Gill, K., Amit, Z. & Smith, B. R. The regulation of alcohol consumption in rats: the role of alcohol-metabolizing enzymes-catalase and aldehyde dehydrogenase. *Alcohol Fayettev. N* 13, 347–353 (1996).

90. Venza, I., Visalli, M., Oteri, R., Teti, D. & Venza, M. Combined effects of cigarette smoking and alcohol consumption on antioxidant/oxidant balance in age-related macular degeneration. *Aging Clin. Exp. Res.* 24, 530–536 (2012).

91. Gram, I. T. *et al.* Smoking and Risk of Breast Cancer in a Racially/Ethnically Diverse Population of Mainly Women Who Do Not Drink Alcohol: The MEC Study. *Am. J. Epidemiol.* 182, 917–925 (2015).

92. Gram, I. T. *et al.* Smoking and breast cancer risk by race/ethnicity and oestrogen and progesterone receptor status: the Multiethnic Cohort (MEC) study. *Int. J. Epidemiol.* 48, 501–511 (2019).

93. Jones, M. E., Schoemaker, M. J., Wright, L. B., Ashworth, A. & Swerdlow, A. J. Smoking and risk of breast cancer in the Generations Study cohort. *Breast Cancer Res. BCR* 19, 118 (2017).

94. Band, P. R., Le, N. D., Fang, R. & Deschamps, M. Carcinogenic and endocrine disrupting effects of cigarette smoke and risk of breast cancer. *Lancet Lond. Engl.* 360, 1044–1049 (2002).

95. Zhu, M. *et al.* Genetic Risk for Overall Cancer and the Benefit of Adherence to a Healthy Lifestyle. *Cancer Res.* (2021) doi:10.1158/0008-5472.CAN-21-0836.

 doi:10.20944/preprints202209.0360.v1