Title

Genetic variants and non-genetic factors predict circulating vitamin D levels in Hispanic and non-Hispanic White women: the Breast Cancer Health Disparities Study.

Journal

International Journal of Molecular Epidemiology and Genetics, 5(1)

ISSN

1948-1756

Authors

Wang, Wei
Ingles, Sue Ann
Torres-Mejía, Gabriela
et al.

Publication Date

2014

Peer reviewed

Int J Mol Epidemiol Genet 2014;5(1):31-46
www.ijmeg.org /ISSN:1948-1756/IJMEG1311005

Original Article

Genetic variants and non-genetic factors predict circulating vitamin D levels in Hispanic and non-Hispanic White women: the Breast Cancer Health Disparities Study

Wei Wang[1,2], Sue Ann Ingles[4,5], Gabriela Torres-Mejía[6], Mariana C Stern[4], Frank Z Stanczyk[5], Gary G Schwartz[7], David O Nelson[1], Laura Fejerman[8], Roger K Wolff[9], Martha L Slattery[9], Esther M John[1,3]

[1]Cancer Prevention Institute of California, Fremont, CA 94538, USA; [2]Biomedical Informatics Training Program, Stanford University, Stanford, CA 94305, USA; [3]Division of Epidemiology, Department of Health Research and Policy, Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94503, USA; [4]Department of Preventive Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA; [5]Department of Obstetrics and Gynecology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; [6]Instituto Nacional de Salud Pública, Cuernavaca Morelos CP 62100, México; [7]Department of Cancer Biology, School of Medicine, Wake Forest University, Medical Center Boulevard, Winston-Salem, NC 27157, USA; [8]Division of General Internal Medicine, Department of Medicine, Institute for Human Genetics and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA 94158, USA; [9]Department of Medicine, University of Utah, Salt Lake City, UT 84132, USA

Received November 15, 2013; Accepted December 30, 2013; Epub February 17, 2014; Published March 1, 2014

Abstract: Genome-wide association studies (GWAS) have identified common polymorphisms in or near *GC*, *CYP2R1*, *CYP24A1*, and *NADSYN1/DHCR7* genes to be associated with circulating levels of 25-hydroxyvitamin D [25(OH)D] in European populations. To replicate these GWAS findings, we examined six selected polymorphisms from these regions and their relation with circulating 25(OH)D levels in 1,605 Hispanic women (629 U.S. Hispanics and 976 Mexicans) and 354 non-Hispanic White (NHW) women. We also assessed the potential interactions between these variants and known non-genetic predictors of 25(OH)D levels, including body mass index (BMI), sunlight exposure and vitamin D intake from diet and supplements. The minor alleles of the two *GC* polymorphisms (rs7041 and rs2282679) were significantly associated with lower 25(OH)D levels in both Hispanic and NHW women. The *CYP2R1* polymorphism, rs2060793, also was significantly associated with 25(OH)D levels in both groups. We found no significant associations for the polymorphisms in the *CYP24A1*. In Hispanic controls, 25(OH)D levels were significantly associated with the rs12785878T and rs1790349G haplotype in the *NADSYN1/DHCR7* region. Significant interactions between *GC* rs2282679 and BMI and between rs12785878 and time spent in outdoor activities were observed. These results provide further support for the contribution of common genetic variants to individual variability in circulating 25(OH)D levels. The observed interactions between SNPs and non-genetic factors warrant confirmation.

Keywords: Circulating levels, Hispanics, genetic polymorphisms, SNPs, genotype-phenotype correlation, vitamin D

Introduction

Vitamin D deficiency is a common health problem worldwide [1] and has been implicated in a wide range of diseases beyond bone disease, e.g., diabetes, cardiovascular disease, multiple sclerosis, and some cancer types, with the evidence being most consistent for colorectal cancer ([2] and reviewed in [1, 3, 4]).

Known determinants of an individual's vitamin D status, which is measured by circulating 25-hydroxyvitamin D (25(OH)D) levels, include sunlight exposure, and vitamin D from diet and supplements. Twin studies and family-based studies suggest that genetic factors also contribute to a significant proportion of the inter-individual variability in vitamin D status [5-7].

There have been three published genome-wide association studies (GWAS) of vitamin D status (circulating 25(OH)D levels and vitamin D insufficiency) in adults (http://www.genome.gov/gwastudies/). The two GWAS conducted in populations of European ancestry found significant associations with several single nucleotide polymorphisms (SNPs) located within or near vitamin D metabolic pathway genes [8, 9]. These genes include the *GC* gene (encoding vitamin D binding protein, formerly known as group-specific component), the *CYP2R1* gene (encoding 25-hydroxylase, an enzyme that catalyzes the conversion of vitamin D to 25(OH)D), the *CYP24A1* gene (encoding 24-hydroxylase, a key enzyme in vitamin D inactivation), and the *NADSYN1/DHCR7* gene region (encoding nicotinamide adenine dinucleotide synthetase and 7-dehydrocholesterol reductase, respectively. The latter converts 7-dehydrocholesterol to cholesterol and thus modulates the availability of 7-dehydrocholesterol in the skin). In the third GWAS, conducted in U.S. Hispanics, polymorphisms in these four genes or genomic regions were not associated with 25(OH)D levels at a genome-wide significance level [10]. However, that study had a much smaller sample size and used a family-based design, which typically has reduced power due to overmatching on genotype [11, 12].

The associations between polymorphisms in the *GC, CYP24A1, and CYP2R1* genes and circulating 25(OH)D have been supported by several candidate gene studies [13-18]. However, with a few exceptions [16-19], those studies have focused primarily on populations of European descent. The *NADSYN1/DHCR7* region was replicated in a candidate gene study of a Chinese population [18]. Further replication is needed in other populations, including Hispanics who are at a greater risk of vitamin D deficiency than non-Hispanic Whites [20].

The aim of this study was to replicate the GWAS-identified polymorphisms in Mexican women, and U.S. Hispanic and non-Hispanic White (NHW) women living in California. Hereafter, the Mexican and U.S. Hispanic women are referred to as Hispanics. In addition, in a subset of study participants, we assessed the interactions between the SNPs and non-genetic determinants of circulating 25(OH)D.

Materials and methods

Study population

This analysis from the Breast Cancer Health Disparities Study [21] includes participants from two population-based case-control studies who completed an in-person interview and provided a blood or mouthwash sample. All study participants signed written informed consent. The study was approved by the institutional review board of the Cancer Prevention Institute of California and each participating institution in Mexico.

The San Francisco Bay Area Breast Cancer Study (SFBCS) [22, 23] included women aged 35-79 years from the San Francisco Bay Area diagnosed with a first primary histologically confirmed invasive breast cancer between April 1997 and April 2002; controls were identified by random-digit dialing and frequency-matched to cases based on the expected race/ethnicity and 5-year age distribution. A total of 1,105 cases and 1,318 controls completed the interview and 93% of cases (750 Hispanics and 276 NHWs) and 92% of controls (916 Hispanics and 298 NHWs) provided a blood or mouthwash sample.

The Mexico Breast Cancer Study (MBCS) [24, 25] included women aged 28-74 years from Mexico. Eligible cases were diagnosed with either a new histologically confirmed *in situ* or invasive breast cancer between January 2004 and December 2007 at 12 participating hospitals from three main health care systems; controls were randomly selected from the catchment area as the cases and frequency matched to cases based on 5-year age distribution, membership in health care institution, and place of residence. A total of 1,000 cases and 1,074 controls completed the interview. Blood was collected from 85% of cases and 93% of controls.

Data collection

Both studies administered similar structured questionnaires in English or Spanish to collect information on demographic background, lifestyle factors, menstrual and reproductive history, supplement use, and other breast cancer risk factors. The interview also included measurements of weight, height, hip and waist circumferences.

Additionally, the SFBCS collected information on sunlight exposure, including lifetime residential history; lifetime history of time spent in outdoor activities (occupational or recreational physical activities, outdoor chores, walking and bicycling for transportation); as well as sun avoidance behaviors. Usual dietary vitamin D intake in the reference year (defined as the calendar year before diagnosis for cases or before selection into the study for controls) was assessed using a modified Block Food Frequency Questionnaire [26, 27]. Annual average ultraviolet (UV) exposure at the residence at the time of diagnosis (cases) or selection into the study (controls) was estimated using the GIS (geographical information systems)-based ANUSPLIN model, as described elsewhere [28]. Several summary variables of residential UV exposure were created: mean at the point residence; and mean, median, minimum and maximum, and standard deviation (SD) within a 20 km radius of the residence. Skin pigmentation was measured using a Minolta Chromameter, with measures ranging from 0 (perfect black) to 100 (perfect white), and constitutive pigmentation was calculated as the average of two measurements taken at the upper inner arm, an area usually not exposed to the sun [23]. For 206 Hispanic cases, no skin pigmentation measurements were taken.

A second short questionnaire was administered at the time of blood collection (except for 206 Hispanic cases), with a median time between interview and blood draw of 10 months for cases and 9 months for controls. This questionnaire updated information on current weight, menstrual and reproductive history, and outdoor activities during six months prior to blood collection.

Measurement of circulating 25(OH)D

In the MBCS, blood samples were processed within 1 hour of blood draw and serum was stored at -70°C until 25(OH)D measurement by liquid chromatography-tandem mass spectrometry at Quest Diagnostics Nichols Institute (San Juan Capistrano, CA). The intra-assay and inter-assay coefficients of variation (CVs) were 6.4-12.6% and 8.0-14.4%, respectively. The assay sensitivity was 4 ng/mL (1 ng/mL=2.496 nmol/L). 25(OH)D measurements were available for a random sample of 574 cases and 649 controls.

In the SFBCS, blood samples were transported to the laboratory on ice, processed within 3 hours of blood draw, and stored at -70°C until 25(OH)D analysis. Serum total 25(OH)D in controls was measured using chemiluminescent immunoassay from Nichols Institute Diagnostics (San Clemente, CA). The intra-assay and inter-assay CVs were 3.0-4.5% and 6.4-14.5%, respectively. The assay sensitivity was 4 ng/mL. For cases, plasma 25(OH)D level was measured using an iodinated RIA kit from DiaSorin (Stillwater, MN) with a two-step procedure [29]. The intra-assay and inter-assay CVs were 8.6-12.5% and 8.2-11.0%, respectively. The assay sensitivity was 1.5 ng/mL. 25(OH)D measurements were available for 670 cases (440 Hispanics and 230 NHWs) and a subset of 358 controls (223 Hispanics and 135 NHWs).

Genotyping

DNA was extracted from either blood or mouthwash samples. Whole Genome Amplification (WGA) was applied to the mouthwash-derived DNA samples prior to genotyping. Six SNPs were selected from among those that passed the genome-wide significance level ($p<10^{-8}$) in the two prior GWAS (described above). Redundant SNPs (r^2 of at least 0.7 in HapMap-MEX panels) were excluded, with priority given to SNPs that were identified by both GWAS (rs2282679 and rs7041 in *GC*; rs12785878 and rs1790349 in *DHCR7/NADSYN1*; rs2060793 in *CYP2R1*; rs6013897 in *CYP24A1)* (**Table 1**). The six SNPs and 104 Ancestral Informative Markers (AIMs) were genotyped using a custom Illumina GoldenGate 1536 SNP OPA (Illumina Inc, San Diego, CA) [21]. Overall, a genotyping call rate of 99.93% was attained (99.65% for whole-genome amplified samples). For the 132 duplicate samples that were included as an internal quality control, the concordance rate was 99.996% as determined by 193,297 matching genotypes among sample pairs [21].

Statistical analysis

Because different measurement methods were used between the MBCS and SFBCS and between cases and controls in the SFBCS, circulating 25(OH)D was standardized according to the mean and SD in each of the following groups: MBCS controls, MBCS cases, SFBCS Hispanic controls, SFBCS Hispanic cases,

Table 1. Linkage Disequilibrium (Measured by r^2) Between the GWAS-identified SNPs and Those Selected in the Study

Nearby gene	SNPs[1]	*p* value reported in GWAS	LD in Hapmap-CEU panel	LD in Hapmap-MEX panel
GC	**rs2282679**	$2.0x10^{-30}$ [9]; $1.9x10^{-109}$ [8]		
	rs7041	$4.1x10^{-22}$ [9]; $6.31x10^{-59}$ [8]	r^2=0.6 with rs2282679	r^2=0.2 with rs2282679
	rs3755967	$2.4x10^{-75}$ [8]	r^2=1.0 with rs2282679	r^2=1.0 with rs2282679
	rs17467825	$6.7x10^{-74}$ [8]	r^2=1.0 with rs2282679	Not available
	rs1155563	$3.8x10^{-25}$ [9]; $2.4x10^{-73}$ [8]	r^2=0.8 with rs2282679	r^2=0.7 with rs2282679
	rs2298850	$2.0x10^{-71}$ [8]	r^2=0.95 with rs2282679	r^2=0.8 with rs2282679
DHCR7/NADSYN1	**rs12785878**	$2.1x10^{-27}$ [8]		
	rs7944926	$9.0x10^{-16}$ [8]	r^2=1.0 with rs12785878	r^2=1.0 with rs12785878
	rs12800438	$2.5x10^{-15}$ [8]	r^2=1.0 with rs12785878	r^2=0.96 with rs12785878
	rs3794060	$3.4x10^{-15}$ [8]	r^2=1.0 with rs12785878	r^2=1.0 with rs12785878
	rs4945008	$4.6x10^{-15}$ [8]	r^2=0.95 with rs12785878	Not available
	rs4944957	$8.7x10^{-15}$ [8]	r^2=1.0 with rs12785878	r^2=0.96 with rs12785878
	rs1790349	$1.8x10^{-6}$ [9]	r^2=0.5 with rs12785878	r^2=0.1 with rs12785878
	rs3829251	$8.8x10^{-7}$; $3.4x10^{-9}$ in combined meta-analysis [9]	r^2=0.9 with rs1790349	r^2=0.7 with 1790349
	rs11234027	$3.4x10^{-3}$; $3.4x10^{-9}$ in combined meta-analysis [9]	r2=0.9 with rs1790349	r^2=0.7 with 1790349
CYP2R1	**rs2060793**	$2.9x10^{-5}$; $2.9x10^{-17}$ in combined meta-analysis [9]; $1.7x10^{-11}$ [8]		
	rs10741657	$3.3x10^{-20}$ [8]	r^2=1.0 with rs2060793	Not available
	rs1993116	$2.9x10^{-17}$; $2.9x10^{-17}$ in combined meta-analysis [9]	r^2=0.98 with rs2060793	r^2=0.9 with rs2060793
	rs12794714	$1.8x10^{-9}$ [8]	r^2=0.5 rs2060793	r^2=0.7 with rs2060793
	rs10500804	$2.7x10^{-9}$ [8]	r^2=0.5 with rs2060793	r^2=0.9 with rs2060793
	rs7116978	$5.0x10^{-9}$ [8]	r^2=0.8 with rs2060793	r^2=0.9 with rs2060793
CYP24A1	**rs6013897**	$7.2x10^{-10}$ [8]		

[1]SNPs that were selected for genotyping in the present study are highlighted in bold.

SFBCS NHW controls and SFBCS NHW cases, with standardized 25(OH)D calculated as [measured 25(OH)D minus group mean] divided by group SD.

Hardy-Weinberg equilibrium (HWE) was tested for each control group using the Pearson goodness-of-fit statistic. Multiple testing-adjusted HWE *p*-value was obtained using the Benjamini and Hochberg procedure in SAS MULTTEST [30]. Multiple linear regression models were fit to examine associations between each SNP and standardized 25(OH)D levels. SNPs were analyzed assuming a co-dominant model. An additive model was also assessed. Individual ancestry was computed using the STRUCTURE program [31, 32] based on 104 AIMs [21]. Percent Native American ancestry according to a two-founding population model was included as a continuous covariate in all regression models, in addition to age at blood draw (continuous), season at blood draw (spring: March 21-June 20; summer: June 21-September 20; fall: September 21-December 20; December 21-March 20), and body mass index (BMI) [weight (kg)/height^2 (m), continuous]. Associations between SNPs and 25(OH)D levels were analyzed separately in Hispanic controls, Hispanic cases, NHW controls and NHW cases. No significant differences in associations by disease status were detected in either Hispanics or NHWs, and we therefore pooled cases and controls within each ethnic group. Secondary analyses were performed on the non-standardized levels of 25(OH)D within each of the following groups: Mexican controls, Mexican cases, U.S. Hispanic controls, U.S. Hispanic cases, NHW controls and NHW cases. A generalized linear model framework with natural log link and a gamma distribution was applied. Results on the genotype-25(OH)D associations were similar (data not shown).

Haplotype analysis was performed for the two *GC* SNPs and the two SNPs in the *NADSYN1/DHCR7* region. Estimated haplotype frequencies and individual haplotype probabilities were obtained using the TagSNP software [33]. Gene-gene interactions between polymorphisms in or near the *GC, CYP2R1 and CYP24A1* genes and the *NADSYN1/DHCR7* region were examined by stratified analysis. Statistical interactions were assessed by including cross-product terms in regression models.

Associations between 25(OH)D levels and BMI, total vitamin D intake, residential UV exposure, outdoor activity during the 6 months prior to blood draw, and constitutive skin pigmentation were analyzed in the SFBCS only. Gene-environment interactions were examined between the significant non-genetic predictors of 25(OH)D levels and the SNPs.

Of the 2,251 women with measured 25(OH)D levels, 184 (153 from the MBCS and 31 from the SFBCS) were excluded due to missing genotype data. We also excluded 108 women with missing data on covariates that were adjusted for in the analysis. The final analysis included 976 women from the MBCS and 983 from the SFBCS, for a total of 1,605 Hispanics and 354 NHWs.

Results

Characteristics of study participants

Participant characteristics are summarized in **Table 2**. Overall, vitamin D deficiency (defined as ≤20 ng/mL as in [34]) was present in 44% of the study participants and was more prevalent in breast cancer cases than in controls in both the MBCS and SFBCS. Hispanic cases and controls from the SFBCS both had a higher prevalence of vitamin D deficiency than their NHW counterparts. In the SFBCS, Hispanic women had darker constitutive skin pigmentation and spent less time in outdoor activities than NHW women. The SFBCS Hispanics had higher dietary vitamin D intake than NHWs, but were less likely to take vitamin D supplements and therefore had lower total vitamin D intake.

The genes, corresponding SNPs and their minor allele frequencies (MAFs) in controls are presented in **Table 3**. No significant departure of observed genotype frequencies from HWE expectations was found for any of the SNPs [false discovery rate (FDR) *p*-value>0.05], except for rs12785878 in Mexican controls.

Associations between circulating 25(OH)D levels and SNPs

The associations between the SNPs and standardized 25(OH)D levels are shown in **Table 4** for Hispanic and in **Table 5** for NHW controls and cases. In both ethnic groups, associations did not differ significantly between cases and controls for any of the six SNPs, thus cases and

Table 2. Circulating 25(OH)D Levels and Covariates by Study, Ethnicity and Case-control Status[1]

	Mexico Breast Cancer Study (MBCS)		San Francisco Bay Area Breast Cancer Study (SFBCS)				
	Controls N=564	Cases N=412	Hispanic controls N=215	NHW controls N=130	Hispanic cases N=414	NHW cases N=224	*P* value[5]
Vitamin D deficient (<20 ng/mL), %	27.3	47.3	52.0	27.7	63.5	48.2	<0.01
Age at blood draw (years), mean (SD)	50.3 (8.9)	52.6 (9.7)	52.8 (11.6)	56.8 (14.0)	59.5 (11.2)	61.1 (12.3)	<0.01
Body mass index at blood draw (kg/m^2), mean (SD)[2]	30.7 (5.4)	29.8 (5.9)	29.7 (5.5)	26.8 (6.3)	29.5 (8.8)	27.2 (5.1)	<0.01
Percent Native American ancestry, mean (SD)	0.71 (0.2)	0.66 (0.2)	0.46 (0.2)	0.05 (0.1)	0.42 (0.2)	0.05 (0.1)	<0.01
Time between interview and blood draw (months), median			10	7	18	5	
Dietary vitamin D intake (IU/day), mean (SD)			283.0 (216.0)	248.3 (154.6)	254.0 (150.7)	247.3 (146.7)	0.09
Total vitamin D intake (IU/day), mean (SD)			426.0 (385.1)	509.8 (346.6)	341.4 (252.5)	503.2 (370.6)	<0.01
Outdoor activities (hours/week), mean (SD)[3]			3.5 (4.6)	5.0 (4.8)	3.2 (4.6)	4.0 (3.7)	<0.01
Constitutive skin pigmentation, mean (SD)[4]			34.7 (4.3)	39.6 (4.2)	34.5 (3.9)	38.6 (3.4)	<0.01

[1]N=1959 women with measured 25(OH)D levels, after excluding those with missing covariates, including season at blood draw, age at blood draw, body mass index (BMI), and genetic ancestry. [2]For 190 SFBCS Hispanic cases who had missing BMI at blood draw, BMI in the reference year was used instead. [3]For 190 SFBCS Hispanic cases, no information was collected on outdoor activities within the 6 months prior to blood collection. [4]Skin pigmentation measured at the upper inner arm, not taken for the 190 SFBCS Hispanic cases and missing for another 13 women. [5]*P* values are for chi square tests for categorical variables and for ANOVA tests comparing mean values for continuous variables.

Table 3. Genes, SNPs, and Allele Frequencies in Controls, by Study and Ethnicity

Nearby gene	SNP	Chromosome location	Region	Major/Minor Allele	MAF[1] in MBCS[2] controls	HWE *p*-value[3]	MAF in SFBCS[4] Hispanic controls	HWE *p*-value[3]	MAF in SFBCS NHW controls	HWE *p*-value[3]
GC	rs7041	4q12-q13	coding	G/T	0.50	0.04 (0.21)	0.47	0.78 (0.93)	0.44	0.58 (0.87)
GC	rs2282679	4q12-q13	intron	A/C	0.19	0.50 (0.84)	0.21	0.05 (0.21)	0.30	0.68 (0.89)
NADSYN1	rs12785878	11q13.4	intron	T/G	0.56	<0.001 (0.01)	0.48	0.24 (0.47)	0.31	0.89 (0.94)
NADSYN1/DHCR7	rs1790349	11q13.4	Intergenic	A/G	0.17	0.13 (0.34)	0.20	0.04 (0.21)	0.15	0.19 (0.44)
CYP2R1	rs2060793	11p15.2	5' upstream	G/A	0.36	0.69 (0.89)	0.38	0.84 (0.94)	0.38	0.51 (0.84)
CYP24A1	rs6013897	20q13	intergenic	TT/A	0.42	0.08 (0.23)	0.34	0.97 (0.97)	0.21	0.07 (0.23)

[1]Minor allele frequency (MSF). [2]Mexico Breast Cancer Study. [3]Hardy-Weinberg Equilibrium (HWE) *p*-value from a 1df chi-square test (false discovery rate *p*-value in parenthesis). [4]San Francisco Bay Area Breast Cancer Study.

Table 4. Associations Between Circulating 25(OH)D Levels and Genotypes in Hispanic Cases and Controls from the Mexico Breast Cancer Study and the San Francisco Bay Area Breast Cancer Study

	Cases and Controls Combined				Controls				Cases				
Genotypes	N	25(OH)D (ng/mL) Mean (SD)	Adjusted beta (SE)[1]	Adjusted *P* value[1]	N	25(OH)D (ng/mL) Mean (SD)	Adjusted Beta (SE)[1]	Adjusted *P* value[1]	N	25(OH)D (ng/mL) Mean (SD)	Adjusted Beta (SE)[1]	Adjusted *P* value[1]	$P_{interaction}$[2]
rs7041													
GG	419	21.8 (7.7)			216	23.6 (7.0)			203	19.8 (7.9)			
GT	764	20.6 (7.3)	-0.15 (0.06)	0.01	362	22.2 (7.3)	-0.20 (0.08)	0.02	402	19.1 (7.0)	-0.09 (0.09)	0.29	
TT	422	19.4 (7.2)	-0.33 (0.07)	$<10^{-4}$	201	21.3 (7.3)	-0.34 (0.10)	0.0004	221	17.8 (6.7)	-0.32 (0.10)	0.001	
Per copy of T allele			-0.17 (0.03)	$p_{trd}<10^{-4}$			-0.17 (0.05)	$p_{trd}=4x10^{-4}$			-0.16 (0.05)	$p_{trd}=0.001$	0.73
rs2282679													
AA	1005	21.4 (7.7)			516	23.1 (7.5)			489	19.6 (7.4)			
AC	529	19.5 (6.8)	-0.23 (0.05)	$<10^{-4}$	227	21.4 (6.2)	-0.28 (0.08)	0.0002	302	18.2 (6.9)	-0.18 (0.07)	0.01	
CC	71	16.7 (6.4)	-0.56 (0.12)	$<10^{-4}$	36	17.7 (7.4)	-0.65 (0.17)	0.0001	35	15.8 (5.0)	-0.43 (0.18)	0.02	
Per copy of C allele			-0.25 (0.04)	$p_{trd}<10^{-4}$			-0.30 (0.06)	$p_{trd}<10^{-4}$			-0.19 (0.06)	$p_{trd}=0.002$	0.21
rs12785878					185	22.3 (6.6)			189	19.4 (7.0)			
TT	374	20.8 (6.9)			353	22.6 (7.8)	0.09 (0.09)	0.33	410	18.9 (7.4)	-0.07 (0.09)	0.42	
TG	763	20.6 (7.8)	0.01 (0.06)	0.87	241	22.0 (7.0)	-0.05 (0.10)	0.59	227	18.6 (6.9)	-0.10 (0.10)	0.31	
GG	468	20.4 (7.1)	-0.07 (0.07)	0.29	185	22.3 (6.6)			189	19.4 (7.0)			
Per copy of G allele			-0.04 (0.03)	$p_{trd}=0.26$			-0.03 (0.05)	$p_{trd}=0.50$			-0.05 (0.05)	$p_{trd}=0.32$	0.86
rs1790349													
AA	1100	20.8 (7.5)			527	22.6 (7.2)			573	19.2 (7.3)			
AG	454	20.0 (7.3)	-0.11 (0.06)	0.05	227	21.7 (7.4)	-0.07 (0.08)	0.34	227	18.3 (6.8)	-0.14 (0.08)	0.07	
GG	51	20.6 (6.9)	-0.07 (0.14)	0.63	25	23.3 (6.2)	0.01 (0.20)	0.95	26	18.0 (6.6)	-0.13 (0.20)	0.52	
Per copy of G allele			-0.08 (0.05)	$p_{trd}=0.08$			-0.05 (0.07)	$p_{trd}=0.49$			-0.11 (0.07)	$p_{trd}=0.08$	0.39
rs2060793													
GG	618	19.9 (6.8)			315	21.5 (6.4)			303	18.3 (6.9)			
GA	747	21.0 (7.8)	0.20 (0.05)	0.0002	356	23.0 (7.7)	0.25 (0.08)	0.001	391	19.3 (7.4)	0.15 (0.08)	0.05	
AA	240	20.8 (7.6)	0.21 (0.08)	0.005	108	22.6 (7.9)	0.21 (0.11)	0.05	132	19.4 (7.2)	0.20 (0.10)	0.05	
Per copy of A allele			0.13 (0.04)	$p_{trd}=0.004$			0.14 (0.05)	$p_{trd}=0.005$			0.11 (0.05)	$p_{trd}=0.03$	0.72
rs6013897													
TT	614	20.7 (7.9)			294	22.4 (7.7)			320	19.1 (7.8)			
TA	736	20.7 (7.3)	0.005 (0.05)	0.93	351	22.4 (7.3)	0.02 (0.08)	0.84	385	19.0 (7.0)	-0.02 (0.08)	0.77	
AA	255	20.1 (6.4)	-0.14 (0.07)	0.06	134	21.9 (6.1)	-0.10 (0.10)	0.35	121	18.2 (6.1)	-0.19 (0.11)	0.09	
Per copy of A allele			-0.06 (0.04)	$p_{trd}=0.11$			-0.04 (0.05)	$p_{trd}=0.46$			-0.08 (0.05)	$p_{trd}=0.14$	0.66

[1]Using standardized 25(OH)D levels as the dependent variable; multiple regression models adjusted for age at blood draw (continuous), body mass index (BMI, at blood draw if available; if not, BMI at interview for controls and BMI in the reference year for cases, continuous), season of blood draw (4 categories), and genetic ancestry (continuous). [2]Interaction between the genotypes and case-control status.

Table 5. Associations Between Circulating 25(OH)D Levels and Genotypes in non-Hispanic White Cases and Controls from the San Francisco Bay Area Breast Cancer Study

	Cases and Controls Combined				Controls				Cases				
Genotypes	N	25(OH)D (ng/mL) Mean (SD)	Adjusted beta (SE)[1]	Adjusted *P* value[1]	N	25(OH)D (ng/mL) Mean (SD)	Adjusted beta (SE)[1]	Adjusted *P* value[1]	N	25(OH)D (ng/mL) Mean (SD)	Adjusted beta (SE)[1]	Adjusted *P* value[1]	$P_{interaction}$[2]
rs7041													
GG	115	24.5 (10.5)			42	26.8 (11.0)			73	23.2 (10.1)			
GT	177	22.5 (8.4)	-0.27 (0.12)	0.02	61	25.0 (10.1)	-0.22 (0.19)	0.24	116	21.2 (7.1)	-0.29 (0.15)	0.05	
TT	62	23.1 (8.5)	-0.30 (0.15)	0.05	27	27.5 (7.6)	-0.15 (0.25)	0.56	35	19.8 (7.7)	-0.42 (0.20)	0.04	
Per copy of T allele			-0.17 (0.07)	p_{trd}=0.02			-0.09 (0.12)	p_{trd}=0.45			-0.23 (0.10)	p_{trd}=0.02	0.32
rs2282679													
AA	182	23.9 (9.9)			62	26.4 (10.6)			120	22.6 (9.4)			
AC	144	22.6 (8.5)	-0.28 (0.11)	0.01	57	26.1 (9.6)	-0.22 (0.18)	0.22	87	20.3 (6.9)	-0.32 (0.14)	0.02	
CC	28	22.6 (7.0)	-0.24 (0.20)	0.22	11	23.9 (7.8)	-0.58 (0.32)	0.07	17	21.8 (6.5)	-0.06 (0.25)	0.81	
Per copy of C allele			-0.19 (0.08)	p_{trd}=0.02			-0.26 (0.13)	p_{trd}=0.06			-0.16 (0.10)	p_{trd}=0.12	0.62
rs12785878													
TT	185	23.2 (9.0)			62	26.0 (9.5)			123	21.8 (8.4)			
TG	145	23.0 (9.1)	-0.06 (0.11)	0.58	55	25.0 (9.7)	-0.14 (0.18)	0.42	90	21.8 (8.5)	0.003 (0.14)	0.98	
GG	24	25.0 (11.2)	0.02 (0.21)	0.92	13	30.7 (11.8)	0.56 (0.30)	0.06	11	18.3 (5.5)	-0.49 (0.31)	0.11	
Per copy of G allele			-0.02 (0.08)	p_{trd}=0.77			0.11 (0.13)	p_{trd}=0.39			-0.10 (0.11)	p_{trd}=0.36	0.17
rs1790349													
AA	263	23.4 (9.2)			95	26.0 (10.0)			168	21.9 (8.3)			
AG	82	22.8 (9.4)	-0.04 (0.12)	0.72	30	26.0 (9.9)	0.07 (0.21)	0.75	52	20.9 (8.6)	-0.11 (0.16)	0.48	
GG	9	24.8 (8.3)	-0.01 (0.33)	0.98	5	28.1 (9.5)	0.16 (0.44)	0.72	4	20.6 (4.6)	-0.21 (0.50)	0.68	
Per copy of G allele			-0.03 (0.10)	p_{trd}=0.76			0.07 (0.16)	p_{trd}=0.65			-0.11 (0.14)	p_{trd}=0.43	0.37
rs2060793													
GG	140	22.5 (9.1)			51	26.9 (10.2)			89	19.9 (7.3)			
GA	162	23.3 (8.8)	0.13 (0.11)	0.22	58	24.8 (9.0)	-0.17 (0.18)	0.62	104	22.5 (8.7)	0.30 (0.14)	0.03	
AA	52	25.3 (10.2)	0.38 (0.16)	0.02	21	27.7 (11.6)	0.12 (0.25)	0.34	31	23.7 (9.0)	0.52 (0.21)	0.01	
Per copy of A allele			0.18 (0.07)	p_{trd}=0.02			0.01 (0.12)	p_{trd}=0.91			0.27 (0.10)	p_{trd}=0.005	0.09
rs6013897													
TT	227	23.4 (9.3)			85	25.7 (10.1)			142	22.1 (8.5)			
TA	112	23.2 (8.8)	0.001 (0.11)	0.99	36	28.8 (7.8)	0.23 (0.19)	0.23	76	20.5 (8.1)	-0.12 (0.14)	0.39	
AA	15	21.9 (10.7)	-0.21 (0.26)	0.42	9	19.2 (12.7)	-0.50 (0.33)	0.14	6	26.0 (5.2)	0.45 (0.42)	0.29	
Per copy of A allele			-0.04 (0.09)	p_{trd}=0.63			-0.05 (0.14)	p_{trd}=0.74			-0.02 (0.13)	p_{trd}=0.85	0.92

[1]Using standardized 25(OH)D levels as the dependent variable; multiple regression models adjusted for age at blood draw (continuous), body mass index (BMI, at blood draw if available; if not, BMI at interview for controls and BMI in the reference year for cases, continuous), season of blood draw (4 categories), and genetic ancestry (continuous). [2]Interaction between the genotypes and case-control status.

controls were combined. In both groups, the less common alleles at the two *GC* loci (rs7041 and rs2282679) were associated with significantly lower 25(OH)D levels, and the A allele of the *CYP2R1* polymorphism (rs2060793) was significantly associated with higher 25(OH)D levels. The magnitudes of the associations were similar for Hispanic and NHW women. Overall, the three SNPs individually accounted for 0.6-3.5% of the inter-individual variability in circulating 25(OH)D levels, with the R^2 varying between Hispanic and NHW women and between women with or without breast cancer (data not shown).

We did not observe significant associations between 25(OH)D levels and SNPs rs12785878 and rs1790349 in the *NADSYN1/DCRH7* gene region or SNP rs6013897 near *CYP24A1*, for either Hispanic women (**Table 4**) or NHW women (**Table 5**).

The two *GC* SNPs were in weak LD in Hispanics (r^2=0.24) and remained statistically significant after mutual adjustment (β=-0.19 for the rs7041 TT genotype, p for trend=0.02; β=-0.44 for the rs2282679 CC genotype, p for trend<10^{-4}), suggesting independent associations with 25(OH)D levels for the two SNPs. In NHW women, a stronger LD (r^2=0.50) between the two SNPs and a much smaller sample size precluded the examination of the independent associations for the two SNPs. The *GC* haplotype analysis did not provide better prediction than the single SNP analysis in either Hispanics or NHWs (data not shown).

The two SNPs in the *NADSYN1/DHCR7* region had r^2=0.14 and 0.19 in Hispanic and NHW women, respectively. Compared to the most common rs12785878T_rs1790349A haplotype, the rare rs12785878T_rs1790349G haplotype was significantly associated with 25(OH)D levels in Hispanic, but not NHW, control women. The difference in associations between Hispanic and NHW controls was statistically significant (p=0.02). Associations did not differ significantly by case-control status (**Table 6**).

No significant difference by self-reported ethnicity was observed for the associations between 25(OH)D levels and any of the six SNPs, nor did the associations differ significantly by estimated genetic ancestry (data not shown).

Hispanics and NHWs were combined in the analyses of gene-gene interactions to increase statistical power. No significant interactions between the *GC* SNPs and the *CYP2R1* rs2060793 were found, nor did they interact significantly with rs6013897 near *CYP24A1* or the SNPs in the *NADSYN1/DHCR7* region (data not shown).

Non-genetic predictors of circulating 25(OH)D levels

Associations between 25(OH)D levels and potential non-genetic predictors in women from the SFBCS are shown in **Table 7**. Various summary measures of residential UV exposure were examined individually and all provided similar model prediction [based on Aikaike Information Criterion (AIC)]. Mean UV levels within a 20 km radius of the residence was included in the final model. Models with BMI or weight also had similar AIC scores and therefore only associations with BMI are presented. Among cases and controls combined, we observed a significant negative association for BMI (p for trend=0.001), and positive associations for total vitamin D intake (p=0.001) and recent outdoor activities (p=0.001). Results were similar in the analysis of standardized 25(OH) levels (data not shown). Associations did not differ significantly between cases and controls, although the association for outdoor activities was significant only in controls. In cases, neither stage at diagnosis or time interval between diagnosis and blood collection were significant predictors of 25(OH)D levels. Adding them in the model did not result in discernible changes in the associations for other predictors (data not shown). Lighter constitutive skin pigmentation was significantly associated with higher 25(OH)D levels in Hispanic controls, but lower levels in NHW controls. Constitutive skin pigmentation was not a significant predictor of standardized 25(OH)D levels (data not shown).

Overall, the significant non-genetic factors (BMI, total vitamin D intake, outdoor activity, and constitutive skin pigmentation) accounted for 15% of the variability in circulating 25(OH)D levels (22% in controls and 13% in cases), after adjusting for season and age at blood draw. Each of the three significant SNPs accounted for approximately an additional 1% of the variability.

Table 6. Associations Between Circulating 25(OH)D Levels and Haplotypes in the NADSYN1/DHCR7 Region, by Ethnicity and Case-control Status

		Cases and Controls Combined			Controls			Cases			
Ethnicity	Haplotype	Frequency	Adjusted beta (SE)[1]	Adjusted *P* value[1]	Frequency	Adjusted beta (SE)[1]	Adjusted *P* value[1]	Frequency	Adjusted beta (SE)[1]	Adjusted *P* value[1]	$P_{interaction}$[2]
Hispanic	T_A	0.47	ref		0.46	ref		0.48			
	G_A	0.36	-0.02 (0.04)	0.58	0.36	-0.03 (0.05)	0.52	0.35	-0.004 (0.06)	0.95	0.37
	G_G	0.17	-0.08 (0.05)	0.1	0.18	-0.04 (0.07)	0.53	0.17	-0.13 (0.08)	0.09	0.64
	T_G	0.001	-1.41 (0.70)	0.05	0.002	-1.52 (0.70)	0.03	0.0004	7.76 (23.72)	0.74	0.71
NHW	T_A	0.72			0.68	ref		0.74			
	G_A	0.14	0.01 (0.11)	0.90	0.17	0.13 (0.15)	0.38	0.13	-0.06 (0.15)	0.69	0.39
	G_G	0.13	-0.07 (0.11)	0.55	0.14	0.05 (0.18)	0.80	0.12	-0.14 (0.15)	0.33	0.30
	T_G	0.01	0.34 (0.37)	0.36	0.009	0.69 (0.70)	0.33	0.01	0.10 (0.45)	0.83	0.48

[1]Using standardized 25(OH)D levels as the dependent variable; multiple regression models adjusted for age at blood draw (continuous), body mass index (BMI, at blood draw if available; if not, BMI at interview for controls and BMI in the reference year for cases, continuous), season of blood draw (4 categories), and genetic ancestry (continuous). [2]Interaction between the haplotypes and case-control status.

Table 7. Associations Between Circulating 25(OH)D Levels and Non-genetic Factors in Cases and Controls from the San Francisco Bay Area Breast Cancer Study

	Cases and Controls Combined				Controls				Cases				
Exposure	N	25(OH)D (ng/mL) Mean (SD)	Adjusted beta[1,2]	Adjusted *P* value[1,2]	N	25(OH)D (ng/mL) Mean (SD)	Adjusted Beta[1]	Adjusted *P* value[1]	N	25(OH)D (ng/mL) Mean (SD)	Adjusted Beta[3]	Adjusted *P* value[3]	$P_{interaction}$[4]
Body mass index (kg/m^2)													
<25	244	23.7 (9.6)			103	25.8 (10.2)			141	22.2 (8.8)			
25-29.9	254	20.3 (8.1)	-0.09 (0.04)	0.02	116	21.7 (8.7)	-0.04 (0.06)	0.58	138	19.2 (7.4)	-0.10 (0.05)	0.04	
≥30	278	17.9 (7.8)	-0.18 (0.04)	$<10^{-4}$	119	18.5 (8.3)	-0.18 (0.06)	0.003	159	17.5 (7.4)	-0.17 (0.05)	$2x10^{-4}$	
				$p_{trd}<10^{-4}$				$p_{trd}=0.002$				$p_{trd}=2x10^{-4}$	0.57
Total vitamin D intake from diet and supplements (IU/day)													
<215	242	18.7 (8.6)			106	19.1 (8.9)			136	18.4 (8.3)			
215-499.9	236	19.7 (8.2)	0.05 (0.04)	0.15	99	21.4 (9.5)	0.15 (0.06)	0.01	137	18.4 (6.9)	-0.003 (0.05)	0.96	
≥500	298	22.7 (9.0)	0.14 (0.04)	$5x10^{-4}$	133	24.2 (9.4)	0.17 (0.06)	0.003	165	21.4 (8.6)	0.12 (0.05)	0.01	
				$p_{trd}=0.001$				$p_{trd}=0.04$				$p_{trd}=0.01$	0.43
Residential UV exposure (Wh/m^2)[5]													
<4757.0	270	20.8 (9.3)			135	22.3 (10.4)			135	19.3 (7.7)			
4757.1-4827.9	245	21.2 (8.6)	-0.01 (0.04)	0.74	85	22.8 (9.7)	-0.04 (0.06)	0.50	160	20.3 (7.8)	0.01 (0.05)	0.78	
≥4828.0	261	19.6 (8.4)	0.001 (0.04)	0.97	118	20.5 (8.0)	-0.003 (0.06)	0.96	143	18.9 (8.8)	0.02 (0.05)	0.65	
				$p_{trd}=0.98$				$p_{trd}=0.93$				$p_{trd}=0.65$	0.40
Recent outdoor activities (hrs/wk)													
0	211	18.0 (8.0)			94	18.4 (8.9)			117	17.7 (7.2)			
0.1-3	156	20.2 (8.8)	0.07 (0.04)	0.09	53	22.2 (10.2)	0.11 (0.08)	0.14	103	19.1 (7.8)	0.03 (0.05)	0.63	
≥3	409	22.0 (8.9)	0.12 (0.04)	0.001	191	23.4 (9.1)	0.19 (0.06)	$8x10^{-4}$	218	20.7 (8.5)	0.07 (0.05)	0.14	
				$p_{trd}=0.001$				$p_{trd}=0.001$				$p_{trd}=0.13$	0.09
Constitutive skin pigmentation (NHWs only)													
Dark	46	24.6 (9.7)			13	34.4 (9.6)			33	20.7 (6.5)			
Intermediate	103	23.7 (8.2)	-0.06 (0.07)	0.39	39	25.3 (8.7)	-0.34 (0.13)	0.01	64	22.6 (7.8)	0.06 (0.08)	0.43	
Light	190	22.7 (9.5)	-0.09 (0.06)	0.17	72	24.8 (9.6)	-0.36 (0.12)	0.002	118	21.4 (9.2)	0.01 (0.07)	0.86	
				$p_{trd}=0.17$				$p_{trd}=0.01$				$p_{trd}=0.85$	0.11
Constitutive skin pigmentation (Hispanics only)													
Dark	213	17.7 (7.9)			104	18.4 (8.7)			109	17.1 (7.0)			
Intermediate	156	18.6 (7.3)	0.07 (0.05)	0.13	74	19.8 (7.4)	0.10 (0.07)	0.15	82	17.4 (7.1)	0.04 (0.06)	0.54	
Light	68	20.3 (9.1)	0.14 (0.06)	0.02	36	21.4 (9.8)	0.16 (0.09)	0.07	32	18.9 (8.1)	0.11 (0.09)	0.20	
				$p_{trd}=0.01$				$p_{trd}=0.04$				$p_{trd}=0.21$	0.71

[1]Plasma 25(OH)D levels on the original scale were analyzed using a generalized linear model with natural log link and a gamma distribution, adjusting for season and age at blood draw, ethnicity, in addition to mutual adjustment for covariates listed in the table; interaction between ethnicity and constitutive skin pigmentation was also adjusted. [2]With further adjustment for case-control status. [3]Additional adjustment for stage and time since diagnosis did not alter the results. [4]Interaction between the exposure of interest and case-control status. [5]Within a 20 km radius of the residence at diagnosis (cases) or selection into study (controls).

Table 8. Interactions Between Genotypes and Non-genetic Factors in Cases and Controls from the San Francisco Bay Area Breast Cancer Study

	Genotypes	N	25(OH)D (ng/mL) Mean (SD)	Adjusted beta[1]	Adjusted *P* value[1]	$P_{interaction}$ [2]
	rs2282679					
Body mass index <30 kg/m^2	AA	348	22.3 (9.4)			
	AC	246	20.8 (8.1)	-0.29 (0.08)	0.0006	
	CC	34	20.3 (8.3)	-0.45 (0.18)	0.01	
					$p_{trd}<10^{-4}$	
Body mass index ≥30 kg/m^2	AA	199	17.5 (7.7)			
	AC	130	17.3 (7.4)	-0.02 (0.10)	0.83	
	CC	26	16.0 (6.1)	-0.11 (0.19)	0.55	
					p_{trd}=0.59	0.02
	rs12785878					
Outdoor activities <2 hours/week[3]	TT	102	18.6 (7.6)			
	TG	140	18.5 (8.8)	0.16 (0.13)	0.22	
	GG	51	19.7 (9.1)	0.39 (0.17)	0.02	
					p_{trd}=0.03	
Outdoor activities ≥2 hours/week[3]	TT	198	23.0 (8.7)			
	TG	225	21.7 (8.8)	-0.05 (0.10)	0.57	
	GG	77	18.2 (8.5)	-0.31 (0.14)	0.03	
					P_{trd}=0.05	0.005

[1]Using standardized 25(OH)D levels as the dependent variable; multiple regression models adjusted for age at blood draw (continuous), body mass index (BMI, at blood draw if available; if not, BMI at interview for controls and BMI in the reference year for cases, continuous), season of blood draw (4 categories), and genetic ancestry (continuous). [2]Interaction between the genotypes and non-genetic factors. [3]Outdoor activities within the 6 months prior to blood collection.

Statistical interactions between SNPs and non-genetic predictors of circulating 25(OH)D levels

Statistically significant interactions were observed between *GC* SNP rs2282679 and BMI (p for interaction=0.02, **Table 8**). Specifically, 25(OH)D level was inversely associated with the number of rs2282679_C alleles only in women who had a BMI <30 kg/m^2. Significant interaction was also found between NADSYN1/DHCR7 SNP rs12785878 and outdoor activity (p for interaction=0.005). The rs12785878_G allele was associated with increased 25(OH)D levels in women who spent less than 2 hours per week in outdoor activities, but reduced 25(OH)D levels in those who spent more time in outdoor activities. No statistically significant interaction was observed between any of the SNPs and total vitamin D intake (data not shown).

Discussion

In this study of 1,605 Hispanic and 354 NHW women, several common SNPs identified by GWAS as predictors of 25(OH)D levels in populations of European descent were also significantly associated with 25(OH)D levels in Hispanics. Specifically, the minor alleles of rs7041 and rs2282679 in the *GC* gene were associated with lower levels of 25(OH)D, whereas the minor allele of the *CYP2R1* rs2060793 was associated with higher 25(OH)D levels. Associations did not differ significantly between Hispanics and NHWs. In Hispanic controls, circulating 25(OH)D levels were significantly associated with the haplotypes of SNPs rs12785878 and rs1790349 in the *NADSYN1/DHCR7* region. Our study confirms associations between 25(OH)D levels and known non-genetic predictors, including BMI, total vitamin D intake, outdoor activity, and constituent skin pigmentation. Statistically significant interactions were observed between the *GC* rs2282679 SNP and BMI, and between NADSYN1/DHCR7 SNP rs12785878 and outdoor activity.

GC is the major binding protein of vitamin D metabolites in the circulation and responsible for their transportation to target tissue [35]. In the *GC* gene, two non-synonymous SNPs,

rs7041 (Glu416Asp) and rs4588 (Thr420Lys, which we did not examine) have been associated with circulating 25(OH)D levels [13]. In experimental studies, these two SNPs have been linked with altered binding affinity [36] and metabolic turnover of the GC protein [37]. However, in two GWAS [8, 9], a non-coding SNP, rs228679, was an even stronger predictor of 25(OH)D levels, compared to SNP rs7041. It has been speculated that the association of rs228679 with 25(OH)D may be due to its tight LD with rs4588 (r^2=1.0 in HapMap-CEU panel and 0.7 in HapMap-MEX panel) [16, 18]. The exact mechanism linking changes in GC affinity and circulating levels to a change in circulating 25(OH)D levels is not yet established, although an influence on 25(OH)D clearance from the circulation, reabsorption in renal tubules, and/or further metabolism into 1,25-$(OH)_2$ have been suggested [8, 13, 38].

Our finding of a statistically significant interaction between the *GC* SNP rs2282679 and BMI is intriguing. Vitamin D is fat soluble and studies in animals and humans have shown that adipose tissue is the major storage site of 25(OH)D [39, 40]. A strong inverse association between BMI or body fat content and circulating 25(OH)D levels has been repeatedly observed [41-43], consistent with our findings. It is biologically plausible that variants in the *GC* gene that affect the efficiency of transporting vitamin D to adipose tissue interact with the overall body fat content to influence circulating vitamin D levels. The same interaction with BMI was not observed for the *GC* SNP rs7401. Different effects of the two SNPs (or the true functional variants they are in strong LD with) on the GC protein could be a potential explanation.

Although a previous small study suggested that vitamin D supplementation may result in differential changes in serum 25(OH)D levels according to the rs4588 genotypes [44], we observed no significant interactions between rs2282679 (in higher LD with rs4588) and self-reported vitamin D intake.

Overall, the role of the GC protein as the major transporter of vitamin D in circulation and the repeatedly observed associations between common *GC* SNPs and 25(OH)D levels across diverse populations, which include NHW ([8, 9, 14, 15] and references in [13]), African-Americans [19], Asians [16, 18], and Hispanics (in the present study and in [19]), provide compelling evidence for the contribution of the *GC* gene to individual variability of circulating 25(OH)D levels.

Compared to the *GC* gene, *CYP2R1*, has been less well studied as a potential contributor to variation of 25(OH)D levels [13, 15, 45-47]. It encodes 25-hydroxylase which catalyzes the conversion of vitamin D to 25(OH)D. Experimental data on the functional impact of the common polymorphisms in this gene are still lacking. Consistent with our results, SNP rs206793, located around 2 kilobases upstream of the transcription start site, was associated with 25(OH)D levels in the GWAS by Ahn et al. [9] and in a candidate gene study of primarily NHW participants [15]. However, in the Chinese study by Lu et al. [18], this SNP only showed a marginally significant association in subgroup analysis. Several SNPs in or near *CYP2R1* that are in various degrees of LD with rs206793 have also been associated with 25(OH)D levels in some other studies [14, 46, 48]. Taken together, these findings suggest that multiple yet-to-identified "causal" variants may exist in the *CYP2R1* gene or the chromosomal region.

DHCR7 catalyzes the UV radiation-induced synthesis of vitamin D_3 (cholecalciferol) from 7-dehydroxy cholesterol in the skin after sun exposure. SNPs in *NADSYN1/DHCR7* region were first associated with 25(OH)D levels in the two GWAS [8, 9] and later replicated in Chinese by Lu et al. [18]. In our study, SNPs rs12785878 (an intergenic SNP in the region) and rs1790349 (an intronic SNP in *DHCR7*) were not significantly associated with 25(OH)D levels. However, a statistically significant association was observed for the rs12785878T_rs1790349G haplotype in Hispanic controls, supporting a role of these SNPs as markers of functional variants in this region. Although the exact reason is unclear, the lack of replication in NHW women could be potentially explained by different LD structure and a smaller sample size in this group. Interestingly, in the controls and cases from the SFBCS, we observed a significant interaction between the rs12785878 SNP and time spent in outdoor activities, a measure of sunlight exposure which is required for cutaneous vitamin D synthesis. Given the role of DHCR7 in modulating the availability of 7-dehy-

drocholesterol (provitamin-D_3), in the skin, it is reasonable to postulate that functional variants in *DHCR7* may have different effects on vitamin D synthesis in people with different sunlight exposure levels. However, because the rs12785878 genotype frequencies in the Mexican control group showed a significant departure from the HWE, our results for this SNP should be interpreted with caution.

The main strengths of this study include the large sample size of Hispanic women, and the availability of measures of known environmental determinants of 25(OH)D levels, which allowed us to examine important gene-environment interactions. One limitation of our study is that circulating 25(OH)D was measured with different assays between the MBCS and SFBCS, and between cases and controls in the SFBCS. To minimize the variability introduced by the different measurement methods, we standardized 25(OH)D levels according to group means and standard deviations and used this variable in the regression analysis. A caveat of this approach is that we could not evaluate the effects of factors such as ethnicity, skin pigmentation and disease status on 25(OH)D levels and their interactions with the SNPs in the overall study population. Nevertheless, similar conclusions regarding the SNPs and 25(OH)D associations could be drawn from the pooled analysis on standardized 25(OH)D levels and from the stratified analysis on the non-standardized 25(OH)D levels.

Our study provides further support for the contributions of common polymorphisms in the vitamin D metabolic pathway to variability in circulating 25(OH)D levels and risk of vitamin D deficiency among individuals. However, in contrast to known non-genetic predictors, these common variants accounted for less than 4% of the inter-individual variability in 25(OH)D levels. Additional variability may be explained by rare variants in these genomic regions, by variants in genes that regulate the vitamin D pathway, and by interactions between genetic and non-genetic factors.

Acknowledgements

This work was supported by the Breast Cancer Health Disparities Study funded by grant CA14002 from the National Cancer Institute to Dr. Slattery, and by grant 161B-0039 to Dr. Wang from the California Breast Cancer Research Program. The San Francisco Bay Area Breast Cancer Study was supported by grants CA63446 and CA77305 from the National Cancer Institute, grant DAMD17-96-1-6071 from the U.S. Department of Defense and grant 7PB-0068 from the California Breast Cancer Research Program. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000036C awarded to the Cancer Prevention Institute of California; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement #1U58 DP000807-01 awarded to the Public Health Institute. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official view of the National Cancer Institute or endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors. The Mexico Breast Cancer Study was funded by Consejo Nacional de Ciencia y Tecnología (CONACyT) (SALUD-2002-C01-7462). We would also like to acknowledge the contributions of Dr. Myles Cockburn for providing the solar UV data; Jennifer Herrick and Sandra Edwards for data harmonization oversight and management; Erica Wolff and Michael Hoffman for laboratory support; Carolyn Ortega for data management for the Mexico Breast Cancer Study; and Jocelyn Koo for data management for the San Francisco Bay Area Breast Cancer Study.

Abbreviations

AIMs, ancestral informative markers; BMI, body mass index; CV, coefficients of variation; FDR, false discovery rate; HWE, Hardy-Weinberg equilibrium; GWAS, genome-wide association studies; LD, linkage disequilibrium; MAF, minor allele frequency; MBCS, Mexican Breast Cancer Study; NHW, non-Hispanic White; SD, standard deviation; SFBCS, San Francisco Bay Area Breast Cancer Study; SNP, single nucleotide polymorphisms; 25(OH)D, 25-hydroxyvitamin D.

Address correspondence to: Dr. Esther M John, Cancer Prevention Institute of California, 2201 Walnut Avenue, Suite 300, Fremont, CA 94538. Tel: 510-608-5007; Fax: 510-608-5085; E-mail: esther.john@cpic.org

References

[1] Holick MF. Vitamin D deficiency in 2010: health benefits of vitamin D and sunlight: a D-bate. Nat Rev Endocrinol 2011; 7: 73-75.

[2] Byers T. Anticancer vitamins du Jour--The ABCED's so far. Am J Epidemiol 2010; 172: 1-3.

[3] Rosen CJ, Adams JS, Bikle DD, Black DM, Demay MB, Manson JE, Murad MH and Kovacs CS. The nonskeletal effects of vitamin d: an endocrine society scientific statement. Endocr Rev 2012; 33: 456-492.

[4] Holick MF and Chen TC. Vitamin D deficiency: a worldwide problem with health consequences. Am J Clin Nutr 2008; 87: 1080S-1086S.

[5] Shea MK, Benjamin EJ, Dupuis J, Massaro JM, Jacques PF, D'Agostino RB Sr, Ordovas JM, O'Donnell CJ, Dawson-Hughes B, Vasan RS and Booth SL. Genetic and non-genetic correlates of vitamins K and D. Eur J Clin Nutr 2009; 63: 458-464.

[6] Hunter D, De Lange M, Snieder H, MacGregor AJ, Swaminathan R, Thakker RV and Spector TD. Genetic contribution to bone metabolism, calcium excretion, and vitamin D and parathyroid hormone regulation. J Bone Miner Res 2001; 16: 371-378.

[7] Snellman G, Melhus H, Gedeborg R, Olofsson S, Wolk A, Pedersen NL and Michaelsson K. Seasonal genetic influence on serum 25-hydroxyvitamin D levels: a twin study. PLoS One 2009; 4: e7747.

[8] Wang TJ, Zhang F, Richards JB, Kestenbaum B, van Meurs JB, Berry D, Kiel DP, Streeten EA, Ohlsson C, Koller DL, Peltonen L, Cooper JD, O'Reilly PF, Houston DK, Glazer NL, Vandenput L, Peacock M, Shi J, Rivadeneira F, McCarthy MI, Anneli P, de Boer IH, Mangino M, Kato B, Smyth DJ, Booth SL, Jacques PF, Burke GL, Goodarzi M, Cheung CL, Wolf M, Rice K, Goltzman D, Hidiroglou N, Ladouceur M, Wareham NJ, Hocking LJ, Hart D, Arden NK, Cooper C, Malik S, Fraser WD, Hartikainen AL, Zhai G, Macdonald HM, Forouhi NG, Loos RJ, Reid DM, Hakim A, Dennison E, Liu Y, Power C, Stevens HE, Jaana L, Vasan RS, Soranzo N, Bojunga J, Psaty BM, Lorentzon M, Foroud T, Harris TB, Hofman A, Jansson JO, Cauley JA, Uitterlinden AG, Gibson Q, Jarvelin MR, Karasik D, Siscovick DS, Econs MJ, Kritchevsky SB, Florez JC, Todd JA, Dupuis J, Hypponen E and Spector TD. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet 2010; 376: 180-188.

[9] Ahn J, Yu K, Stolzenberg-Solomon R, Simon KC, McCullough ML, Gallicchio L, Jacobs EJ, Ascherio A, Helzlsouer K, Jacobs KB, Li Q, Weinstein SJ, Purdue M, Virtamo J, Horst R, Wheeler W, Chanock S, Hunter DJ, Hayes RB, Kraft P and Albanes D. Genome-wide association study of circulating vitamin D levels. Hum Mol Genet 2010; 19: 2739-2745.

[10] Engelman CD, Meyers KJ, Ziegler JT, Taylor KD, Palmer ND, Haffner SM, Fingerlin TE, Wagenknecht LE, Rotter JI, Bowden DW, Langefeld CD and Norris JM. Genome-wide association study of vitamin D concentrations in Hispanic Americans: the IRAS family study. J Steroid Biochem Mol Biol 2010; 122: 186-192.

[11] Thomas DC, Haile RW and Duggan D. Recent developments in genomewide association scans: a workshop summary and review. Am J Hum Genet 2005; 77: 337-345.

[12] McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP and Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008; 9: 356-369.

[13] McGrath JJ, Saha S, Burne TH and Eyles DW. A systematic review of the association between common single nucleotide polymorphisms and 25-hydroxyvitamin D concentrations. J Steroid Biochem Mol Biol 2010; 121: 471-477.

[14] Bu FX, Armas L, Lappe J, Zhou Y, Gao G, Wang HW, Recker R and Zhao LJ. Comprehensive association analysis of nine candidate genes with serum 25-hydroxy vitamin D levels among healthy Caucasian subjects. Hum Genet 2010; 128: 549-556.

[15] Shui IM, Mucci LA, Kraft P, Tamimi RM, Lindstrom S, Penney KL, Nimptsch K, Hollis BW, Dupre N, Platz EA, Stampfer MJ and Giovannucci E. Vitamin D-related genetic variation, plasma vitamin D, and risk of lethal prostate cancer: a prospective nested case-control study. J Natl Cancer Inst 2012; 104: 690-699.

[16] Gozdzik A, Zhu J, Wong BY, Fu L, Cole DE and Parra EJ. Association of vitamin D binding protein (VDBP) polymorphisms and serum 25(OH) D concentrations in a sample of young Canadian adults of different ancestry. J Steroid Biochem Mol Biol 2011; 127: 405-412.

[17] Signorello LB, Shi J, Cai Q, Zheng W, Williams SM, Long J, Cohen SS, Li G, Hollis BW, Smith JR and Blot WJ. Common variation in vitamin D pathway genes predicts circulating 25-hydroxyvitamin D Levels among African Americans. PLoS One 2011; 6: e28623.

[18] Lu L, Sheng H, Li H, Gan W, Liu C, Zhu J, Loos RJ and Lin X. Associations between common variants in GC and DHCR7/NADSYN1 and vita-

min D concentration in Chinese Hans. Hum Genet 2012; 131: 505-512.
[19] Engelman CD, Fingerlin TE, Langefeld CD, Hicks PJ, Rich SS, Wagenknecht LE, Bowden DW and Norris JM. Genetic and environmental determinants of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D levels in Hispanic and African Americans. J Clin Endocrinol Metab 2008; 93: 3381-3388.
[20] Yetley EA. Assessing the vitamin D status of the US population. Am J Clin Nutr 2008; 88: 558S-564S.
[21] Slattery M, John E, Torres-Mejia G, Lundgreen A, Herrick J, Baumgartner K, Hines L, Stern M and Wolff R. Genetic variation in genes involved in hormones, inflammation, and energetic factors and breast cancer risk in an admixed population. Carcinogenesis 2012 Aug; 33: 1512-21.
[22] John EM, Phipps AI, Davis A and Koo J. Migration history, acculturation, and breast cancer risk in Hispanic women. Cancer Epidemiol Biomarkers Prev 2005; 14: 2905-2913.
[23] John EM, Schwartz GG, Koo J, Wang W and Ingles SA. Sun exposure, vitamin D receptor gene polymorphisms, and breast cancer risk in a multiethnic population. Am J Epidemiol 2007; 166: 1409-1419.
[24] Angeles-Llerenas A, Ortega-Olvera C, Perez-Rodriguez E, Esparza-Cano JP, Lazcano-Ponce E, Romieu I and Torres-Mejia G. Moderate physical activity and breast cancer risk: the effect of menopausal status. Cancer Causes Control 2010; 21: 577-586.
[25] Fejerman L, Romieu I, John EM, Lazcano-Ponce E, Huntsman S, Beckman KB, Perez-Stable EJ, Gonzalez Burchard E, Ziv E and Torres-Mejia G. European ancestry is positively associated with breast cancer risk in Mexican women. Cancer Epidemiol Biomarkers Prev 2010; 19: 1074-1082.
[26] Wang J, John EM, Horn-Ross PL and Ingles SA. Dietary fat, cooking fat, and breast cancer risk in a multiethnic population. Nutr Cancer 2008; 60: 492-504.
[27] Horn-Ross PL, John EM, Lee M, Stewart SL, Koo J, Sakoda LC, Shiau AC, Goldstein J, Davis P and Perez-Stable EJ. Phytoestrogen consumption and breast cancer risk in a multiethnic population: the Bay Area Breast Cancer Study. Am J Epidemiol 2001; 154: 434-441.
[28] Tatalovich Z, Wilson JP, Mack T, Yan Y and Cockburn M. The objective assessment of lifetime cumulative ultraviolet exposure for determining melanoma risk. J Photochem Photobiol B 2006; 85: 198-204.
[29] Cauley JA, Lacroix AZ, Wu L, Horwitz M, Danielson ME, Bauer DC, Lee JS, Jackson RD, Robbins JA, Wu C, Stanczyk FZ, LeBoff MS, Wactawski-Wende J, Sarto G, Ockene J and Cummings SR. Serum 25-hydroxyvitamin D concentrations and risk for hip fractures. Ann Intern Med 2008; 149: 242-250.
[30] Benjamini Y and Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Genetics 2005 Oct; 171: 783-90.
[31] Falush D, Stephens M and Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003; 164: 1567-1587.
[32] Pritchard JK, Stephens M and Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000; 155: 945-959.
[33] Stram DO, Leigh Pearce C, Bretsky P, Freedman M, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE and Thomas DC. Modeling and E-M estimation of haplotype-specific relative risks from genotype data for a case-control study of unrelated individuals. Hum Hered 2003; 55: 179-190.
[34] Holick MF. Vitamin D deficiency. N Engl J Med 2007; 357: 266-281.
[35] Speeckaert M, Huang G, Delanghe JR and Taes YE. Biological and clinical aspects of the vitamin D binding protein (Gc-globulin) and its polymorphism. Clin Chim Acta 2006; 372: 33-42.
[36] Arnaud J and Constans J. Affinity differences for vitamin D metabolites associated with the genetic isoforms of the human serum carrier protein (DBP). Hum Genet 1993; 92: 183-188.
[37] Lauridsen AL, Vestergaard P and Nexo E. Mean serum concentration of vitamin D-binding protein (Gc globulin) is related to the Gc phenotype in women. Clin Chem 2001; 47: 753-756.
[38] Weinstein SJ, Stolzenberg-Solomon RZ, Kopp W, Rager H, Virtamo J and Albanes D. Impact of Circulating Vitamin D Binding Protein Levels on the Association between 25-Hydroxyvitamin D and Pancreatic Cancer Risk: A Nested Case-Control Study. Cancer Res 2012; 72: 1190-1198.
[39] Rosenstreich SJ, Rich C and Volwiler W. Deposition in and release of vitamin D3 from body fat: evidence for a storage site in the rat. J Clin Invest 1971; 50: 679-687.
[40] Mawer EB, Backhouse J, Holman CA, Lumb GA and Stanbury SW. The distribution and storage of vitamin D and its metabolites in human tissues. Clin Sci 1972; 43: 413-431.
[41] Arunabh S, Pollack S, Yeh J and Aloia JF. Body fat content and 25-hydroxyvitamin D levels in healthy women. J Clin Endocrinol Metab 2003; 88: 157-161.
[42] Young KA, Engelman CD, Langefeld CD, Hairston KG, Haffner SM, Bryer-Ash M and Norris JM. Association of plasma vitamin D levels with

adiposity in Hispanic and African Americans. J Clin Endocrinol Metab 2009; 94: 3306-3313.
[43] Wortsman J, Matsuoka LY, Chen TC, Lu Z and Holick MF. Decreased bioavailability of vitamin D in obesity. Am J Clin Nutr 2000; 72: 690-693.
[44] Fu L, Yun F, Oczak M, Wong BY, Vieth R and Cole DE. Common genetic variants of the vitamin D binding protein (DBP) predict differences in response of serum 25-hydroxyvitamin D [25(OH)D] to vitamin D supplementation. Clin Biochem 2009; 42: 1174-1177.
[45] Wjst M, Altmuller J, Faus-Kessler T, Braig C, Bahnweg M and Andre E. Asthma families show transmission disequilibrium of gene variants in the vitamin D metabolism and signalling pathway. Respir Res 2006; 7: 60.
[46] Ramos-Lopez E, Bruck P, Jansen T, Herwig J and Badenhoop K. CYP2R1 (vitamin D 25-hydroxylase) gene is associated with susceptibility to type 1 diabetes and vitamin D levels in Germans. Diabetes Metab Res Rev 2007; 23: 631-636.
[47] Ramos-Lopez E, Kahles H, Weber S, Kukic A, Penna-Martinez M, Badenhoop K and Louwen F. Gestational diabetes mellitus and vitamin D deficiency: genetic contribution of CYP27B1 and CYP2R1 polymorphisms. Diabetes Obes Metab 2008; 10: 683-685.
[48] Wjst M, Altmuller J, Braig C, Bahnweg M and Andre E. A genome-wide linkage scan for 25-OH-D(3) and 1,25-(OH)2-D3 serum levels in asthma families. J Steroid Biochem Mol Biol 2007; 103: 799-802.