

## Article

# The Genetic Architecture of Vitamin D Deficiency among an Elderly Lebanese Middle Eastern Population: An Exome-Wide Association Study

Nagham Nafiz Hendi <sup>1</sup>, Marlene Chakhtoura <sup>2</sup>, Yasser Al-Sarraj <sup>3</sup>, Dania Saleh Basha <sup>2</sup>, Omar Albagha <sup>4</sup>, Ghada El-Hajj Fuleihan <sup>2,†</sup> and Georges Nemer <sup>4,5,\*</sup>

- <sup>1</sup> Division of Biological and Biomedical Sciences, College of Health and Life Sciences, Hamad Bin Khalifa University, Doha P.O. Box 34110, Qatar; nhindi@hbku.edu.qa
- <sup>2</sup> Calcium Metabolism & Osteoporosis Program, American University of Beirut Medical Center, Beirut P.O. Box 11-0236, Lebanon; mc39@aub.edu.lb (M.C.); dms29@mail.aub.edu (D.S.B.); gf01@aub.edu.lb (G.E.-H.F.)
- <sup>3</sup> Qatar Genome Program (QGP), Qatar Foundation Research, Development and Innovation, Qatar Foundation (QF), Doha P.O. Box 5825, Qatar; yalsarraj@qf.org.qa
- <sup>4</sup> Division of Genomics and Translational Biomedicine, College of Health and Life Sciences, Hamad Bin Khalifa University, Doha P.O. Box 34110, Qatar; oalbagha@hbku.edu.qa
- <sup>5</sup> Department of Biochemistry and Molecular Genetics, American University of Beirut, Beirut P.O. Box 11-0236, Lebanon
- \* Correspondence: gnemer@hbku.edu.qa
- † Those authors contributed equally to the work.

**Abstract:** The Middle East region experiences a high prevalence of vitamin D deficiency, yet most genetic studies on vitamin D have focused on European populations. Furthermore, there is a lack of research on the genomic risk factors affecting elderly people, who are more susceptible to health burdens. We investigated the genetic determinants of 25-hydroxyvitamin D concentrations in elderly Lebanese individuals ( $n = 199$ ) through a whole-exome-based genome-wide association study. Novel genomic loci displaying suggestive evidence of association with 25-hydroxyvitamin D levels were identified in our study, including rs141064014 in the *MGAM* ( $p$ -value of  $4.40 \times 10^{-6}$ ) and rs7036592 in *PHF2* ( $p$ -value of  $8.43 \times 10^{-6}$ ). A meta-analysis of the Lebanese data and the largest European genome-wide association study confirmed consistency replication of numerous variants, including rs2725405 in *SLC38A10* ( $p$ -value of  $3.73 \times 10^{-8}$ ). Although the polygenic risk score model derived from European populations exhibited lower performance than European estimations, it still effectively predicted vitamin D deficiency among our cohort. Our discoveries offer novel perspectives on the genetic mechanisms underlying vitamin D deficiency among elderly Middle Eastern populations, facilitating the development of personalized approaches for more effective management of vitamin D deficiency. Additionally, we demonstrated that whole-exome-based genome-wide association study is an effective method for identifying genetic components associated with phenotypes.

**Keywords:** exome-wide association study; vitamin D deficiency; genetic determinants; polygenic risk score; Middle Eastern population



**Citation:** Hendi, N.N.; Chakhtoura, M.; Al-Sarraj, Y.; Basha, D.S.; Albagha, O.; Fuleihan, G.E.-H.; Nemer, G. The Genetic Architecture of Vitamin D Deficiency among an Elderly Lebanese Middle Eastern Population: An Exome-Wide Association Study. *Nutrients* **2023**, *15*, 3216. <https://doi.org/10.3390/nu15143216>

Academic Editor: Ewa Jablonska

Received: 1 June 2023

Revised: 28 June 2023

Accepted: 18 July 2023

Published: 20 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Vitamin D is an essential nutrient that maintains bone and overall health. A deficiency of vitamin D is widely prevalent worldwide, which is identified when serum concentrations of 25-hydroxyvitamin D (25(OH)D) fall below 50 nmol/L (20 ng/mL). This deficiency has been associated with musculoskeletal disorders, like osteoporosis and rickets in children, as well as cardiovascular diseases and cancer [1,2]. Vulnerable populations, particularly the elderly and obese individuals, are at a higher risk due to reduced sunlight exposure, inadequate intake of vitamin D-rich foods, and intestinal malabsorption [3]. To address

this issue, several studies have explored the effectiveness of vitamin D supplementation in managing this deficiency [4]. Additionally, emerging evidence suggests that vitamin D supplementation may have a beneficial effect on weight loss outcomes, potentially through its interaction with receptors involved in energy metabolism [5]. These findings highlight the potential influence of vitamin D on vitamin D deficiency and its associated health issues, underscoring the need for further investigations in this area. However, genetic variations can influence an individual's predisposition to vitamin D deficiency, vitamin D biosynthesis, and their response to vitamin D supplementation [6–8].

Genetic studies have reported heritability estimates of genetic variations in 25(OH)D levels around 80%, underscoring the substantial impact of genetic markers on variations among individuals [9]. Through genome-wide association studies (GWAS), specific genetic variants correlated with 25(OH)D concentrations have been detected in genes responsible for the synthesis and transportation of vitamin D. Notable genes include *CYP2R1* (Cytochrome P450 Family 2R1), *CYP24A1* (Cytochrome P450 Family 24A1), *DHCR7* (7-Dehydrocholesterol Reductase), and *GC* (Vitamin D Binding Protein) [10–13]. Furthermore, studies have also determined a contribution of genetic variants, including *CYP2R1* and *DHCR7*, to the variability in response to vitamin D supplements [14].

Despite the abundant sunshine in the Middle East, vitamin D deficiency is surprisingly widespread, with rates reaching up to 74% [15,16], and approximately 63% in Lebanon [17]. However, studies investigating the genetic aspects of vitamin D have primarily centered on individuals of European descent, leaving a knowledge gap in the impact of genetic variables in Middle Eastern populations. We performed the first whole-exome-based GWAS (ExWAS) in elderly Middle Eastern individuals for vitamin D. The ExWAS approach provides an enhanced ability to detect rare variants within protein-coding genes, thus enabling a more effective analysis of the genetic architecture underlying vitamin D deficiency in our population [18]. In order to uncover additional novel and prevalent genetic factors influencing 25(OH)D levels, we conducted a meta-analysis of our cohort with the largest European GWAS ( $n = 417,580$  individuals) [11]. Additionally, we assessed the effectiveness of polygenic risk scores (PRS) derived from Europeans along with the correlation of genetic factors and vitamin D deficiency in the elderly Lebanese individuals.

## 2. Methods

### 2.1. Participants' Characteristics

The present ExWAS study utilized data from three healthcare centers in Lebanon: Rafic Hariri University Hospital, the American University of Beirut-Medical Center, and Hotel Dieu de France. A total of 199 participants aged 65 years or older, with a body mass index (BMI) between 25 and 29 kg/m<sup>2</sup>, were enrolled in this cross-sectional study. Participants with 25(OH)D concentrations between the validated value (10–30 ng/mL) were included. This study, conducted in Beirut, Lebanon from 2011 to 2013, was a randomized controlled trial that evaluated the impacts of two different vitamin D doses on indices of bone mineral density (NCT01315366) [3]. The study received approval from the Institutional Review Board of the American University of Beirut, Lebanon (IM-GEHF-20) on 24 January 2011, and informed written consent was obtained from all participants prior to participation. The identification code assigned to the clinical trial on ClinicalTrials.gov is NCT01315366.

### 2.2. Quantification of 25(OH)D and Related Covariates

Physical measurements, like height and body weight, were taken for all participants by the Seca 284 stadiometer and balance. BMI was calculated by dividing the weight (kg) by the square of height (m<sup>2</sup>). Additionally, a standardized questionnaire was administered to gather information regarding lifestyle factors and medical history. The concentrations of 25(OH)D were assessed with liquid chromatography mass spectroscopy at Mayo Clinic Laboratories in Rochester, MN, USA. Prior to the statistical analyses, the 25(OH)D concentrations for all individuals underwent a rank-based inverse normal transformation using R version 4.1.3. This method entails arranging the values in ascending order, and

subsequently, converting the ranks to a normal distribution by utilizing the inverse of the cumulative distribution function. The purpose of this approach is to account for outliers and skewness present in traits that deviate from a normal distribution [19].

### 2.3. Whole-Exome Sequencing and Bioinformatics Analyses

Peripheral blood cells were collected from participants through venipuncture during similar seasons of the year and processed for DNA extraction using a DNA extraction kit (the Qiagen QIAamp blood midi kit, catalog number: 51,185) according to the manufacturer's guidelines. The resulting samples were assessed for quantity by a NanoDrop (Thermo) at the molecular core facilities of the American University of Beirut, and subsequently stored at a temperature of  $-80$  degrees Celsius. Subsequently, genomic DNA of the 199 samples were transferred to the Macrogen Inc. (Seoul, Republic of Korea) to create DNA sequencing libraries using 101-base-pair paired-end reads on the SureSelectXT Library Prep Kit on Illumina HiSeq 4000 platform. Each sample yielded paired-end reads with a range of 68,807,342–91,844,864 reads and a total of 6.9–9.2 G base-pair reads. Of these, 95.92–96.67% of the reads passed Q30 (a phred quality score of over 30).

Sample-specific FASTQ files, representing the HiSeq reads for that sample, were aligned with BWA MEM to the GRCh37 reference genome [20]. Metric statistics were captured for all samples to evaluate genome capture, variant alignment, and calling quality using SnpEff version 4.2 [21], bcftools [22] with dbsnp version 138 [23], and ClinVar as databases [24]. SNV genotypes with read depth less than 20 and mapping quality less than 40 were excluded. After the DP genotype filtering, duplicate reads followed by SNVs and InDels genotypes were removed using SureCall 2.0 SNPPET algorithm.

To record variant calling for Lebanese samples, we utilized the HaplotypeCaller from GATK (Genome Analysis Toolkit version 3.4, <https://software.broadinstitute.org/gatk/documentation/article?id=3238>, accessed on 1 February 2023). All the individual intermediate genomic variant call files (gVCF) were used in a joint calling process to create a joint multisample VCF file for all the samples. This process had two steps. In the joint calling process, all sample intermediate genomic variant call files (gVCF) were merged to create a joint multisample VCF file through a two-step process. First, the regions for all the samples were combined using GenomicsDB. Then, GenotypeGVCFs was used to merge all regions, while also applying SNP/Indel recalibration. After applying the VQSR filtration from GATK, only variants that met the specified criteria were retained for further analysis.

We performed a thorough quality control examination via PLINK (version 2.0) to ensure the reliability and integrity of the data [25]. We excluded SNPs that had a MAF  $< 1\%$ , genotyping call rate  $< 90\%$ , and Hardy–Weinberg equilibrium  $p < 1 \times 10^{-6}$ . Furthermore, we eliminated individuals that had a call rate below 95% ( $n = 5$ ) and conducted checks for gender ambiguity, excess heterozygosity, and duplicates. Multidimensional scaling analysis was also performed using PLINK to detect individuals who deviated from the expected population ancestry. We generated a matrix of pairwise identities-by-state (IBS) using a pruned set of independent autosomal variants, 14,893 SNPs, a window size of 200 variants, and an LD threshold of  $r^2 = 0.05$  to study the genetic relatedness among the participants in our study. Outliers of population were removed when deviated from the average by four standard deviation units or more ( $\pm 4$  SD) of the first two mds components. A dataset comprising 481,395 SNPs of 194 participants was used on the genome-wide association analysis.

### 2.4. Exome-Wide Association Analysis

We conducted an ExWAS to investigate the relationship between each genetic marker and concentrations of 25(OH)D. We utilized the generalized mixed model implemented in SAIGE [26], which is a linear method based on variance components. This approach corrects for genetic substructure and relatedness by incorporating the genomic kinship matrix. To account for population mixtures, we analyzed the population principal component (PC) through PLINK [25]. The first four principal components (PCs) were included as covariates

in the regression model to adjust for population structure along with the adjustment of age and gender. The finding of interest cutoff was set at  $p$ -value  $< 5 \times 10^{-5}$  with suggestive significance level of  $p$ -value below 0.05. R was utilized to create the Manhattan plot and quantile–quantile plot, and to calculate the genomic inflation factor. The linear model of SAIGE was employed to estimate the heritability, quantifying the proportion of variation in levels of 25(OH)D attributed to genetic differences among samples.

The pairwise linkage disequilibrium (LD) among top-associated SNPs was examined using the LD clumping analysis of PLINK software (version 1.9) [25]. We applied an  $r^2$  threshold of 0.2 within a window of 250 kb to identify SNPs exhibiting strong LD through our ExWAS findings. Subsequently, we utilized LocusZoom software [27] to visualize the LD patterns, which generated regional plots based on the SAIGE summary statistics. These plots effectively highlighted clusters of SNPs demonstrating high LD.

### 2.5. Meta-Analysis

A meta-analysis for the suggestively associated loci was performed by combining the results of our Lebanese ExWAS and a recent large GWAS (GCST90000616) by Revez JA et al. from the UK Biobank ( $n = 417,580$  European ancestry individuals) [12]. The GWAS model of GCST90000616 for vitamin D was conducted using similar methods of our analysis, with correction of genotype PCs, age, and gender [12]. The NHGRI-EBI GWAS Catalog [28] provided summary statistics for study GCST90000616 taken in December 2022. The alleles (A1 and A2) in our dataset were compared and validated against the alternate and reference alleles in the referenced dataset (GCST90000616). The statistics of association in our study were canonicalized following the alternate allele as specified in the reference genome. Additionally, both studies used rank-based transformation to inversely normalize 25(OH)D measurements. We performed a meta-analysis using the inverse-variance weighted method and assessed the heterogeneity using PLINK (version 1.9) [25].

### 2.6. Validation of Replicated Loci Previously Associated with 25(OH)D

Our ExWAS results were compared to the European GWAS study for 25(OH)D levels (GCST90000616) [12], to validate the variant replication and correlation of our findings in terms of allele effect value and frequency. We examined the presence of SNPs identified in the meta-analysis within the GWAS catalog (EFO\_0004631) for 25(OH)D levels, using the November 2022 release [28]. Our focus was on SNPs that demonstrated significance levels below  $p$ -value  $< 5.0 \times 10^{-8}$ . Initially, we looked for the genetic regions linked to vitamin D in our cohort and compared it to those present in the UK Biobank dataset. Markers in a region of 250 kb surrounding the GWAS catalog signals were examined, to detect potential novel variants as well as replicated associations with vitamin D.

### 2.7. Analysis of Polygenic Risk Scores

We assessed the performance of a polygenic risk score (PRS) derived from a European population in estimating the genetic predisposition for 25(OH)D levels in the Lebanese population by PLINK (version 1.9) [25]. The PRS was computed by aggregating SNPs based on their risk allele frequency, using the effect weights estimated from GWAS. We obtained the PRS files from the most comprehensive vitamin D GWAS in Europeans (PGS000882:  $n = 417,580$  individuals and 1,094,650 variants) [12], through the Polygenic Score Catalog (<https://www.PGSCatalog.org>, accessed on 15 January 2023) [29].

We then calculated Pearson's correlation (R) of the PRS derived from Europeans and baseline levels of 25(OH)D, adjusted for the first four principal components, gender, and age through R software. The performance of the derived polygenic score model in detecting individuals having vitamin D deficiency, characterized by concentrations of 25(OH)D less than 20 ng/mL, was assessed through the area under the receiver operating characteristic (ROC) curve (AUC). The range of AUC is from 0.5 (no distinction) to 1 (complete distinction), providing a measure of the PRS model in effectively detecting individuals who are at risk of vitamin D deficiency.

### 2.8. Functional Insights and Annotation of Variants

The annotation of variants associated with 25(OH)D levels were performed using the Ensembl Variant Effect Predictor release 108 (VEP, <https://grch37.ensembl.org/index.html>, accessed on 15 March 2023) [30]. We utilized the gnomAD (Genome Aggregation Database, <https://gnomad.broadinstitute.org>, accessed on 16 March 2023) and the ALFA (Allele Frequency Aggregator, [www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/](http://www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/), accessed on 16 March 2023) to compare the frequencies of variants detected in our population to those in global populations. Expressions of quantitative trait loci (eQTLs) of that are associated with the expression levels of novel genes in healthy human tissues were analyzed by ARCHS4 (RNA-seq and CHIP-seq sample and signature search [31]) and GTEx (genotype-tissue expression, <http://commonfund.nih.gov/GTEx/>, accessed on 18 March 2023, ARCHS4) database.

## 3. Results

### 3.1. Study Description

This study utilizes whole-exome sequence data from elderly Lebanese participants, with an average age of 71 ( $\pm 4.8$ ) years (interquartile range: 65 to 91 years) and 46.4% female participants ( $n = 104$ ). Notably, the mean of 25(OH)D baseline levels was 20.12 ( $\pm 7.2$ ) ng/mL (interquartile range: 6 to 44 ng/mL). The average of BMI (in kg/m<sup>2</sup>) was almost equivalent between males and females, with a value of  $30.2 \pm 4.6$  (Table 1). No significant linkages were observed between the 25(OH)D levels and related covariates, such as age, gender, or BMI. Participant enrolment was distributed across all seasons. At enrolment, 19% of participants were on calcium supplementation, and only 0.5% were receiving vitamin D supplements.

**Table 1.** Mean characteristics of the enrolled subjects.

Characteristic	Male	Female	Total
Age (years)	72.7 ( $\pm 5.50$ )	69.77 ( $\pm 3.51$ )	71.13 ( $\pm 4.82$ )
BMI (kg/m <sup>2</sup> )	28.69 ( $\pm 3.35$ )	31.60 ( $\pm 5.10$ )	30.24 ( $\pm 4.60$ )
Serum 25(OH)D (ng/mL)	19.36 ( $\pm 6.25$ )	20.83 ( $\pm 7.93$ )	20.12 ( $\pm 7.22$ )
Sample size	90 (46.4%)	104 (53.6%)	194

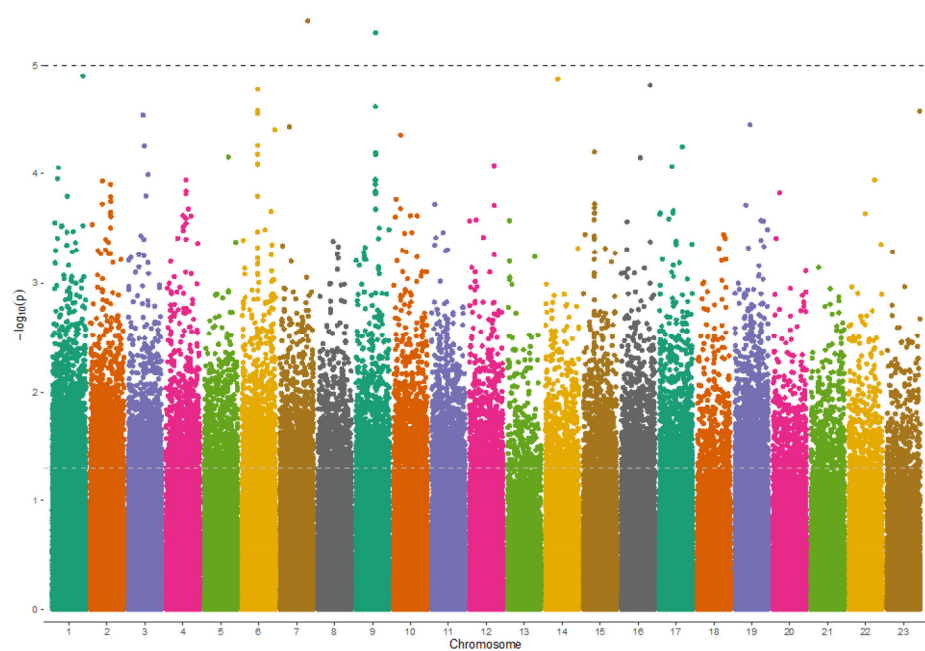
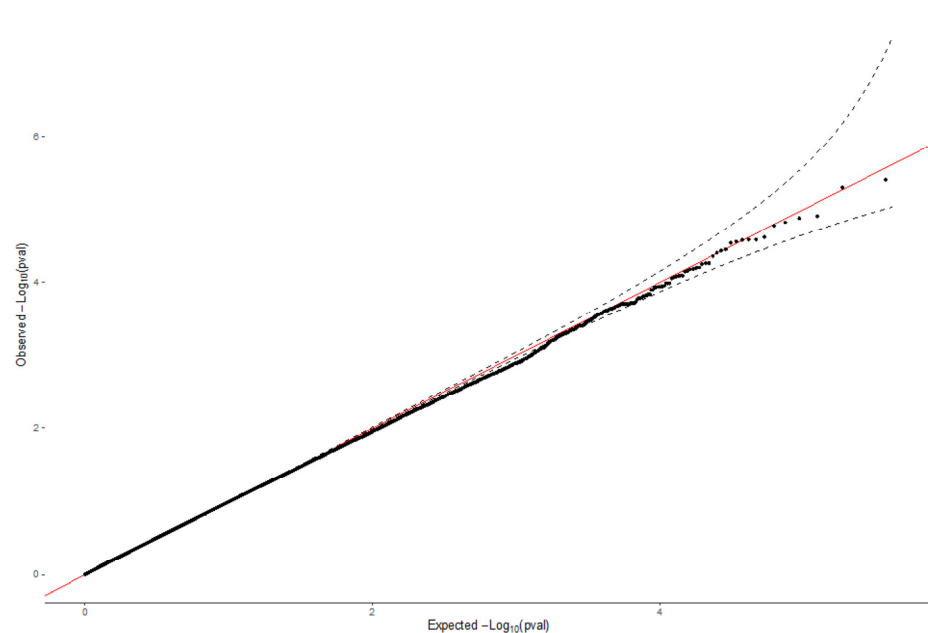
Descriptive statistics are presented as the mean ( $\pm$ SD) or as number (percentages, %) for Lebanese participant characteristics. Abbreviations: SD, standard deviation; BMI, body mass index.

### 3.2. Exome-Wide Association Study on 25(OH)D

The generalized linear mixed model was utilized to detect the genetic basis and potential causal genes of 25(OH)D levels in samples of the elderly Lebanese population who passed quality control criteria ( $n = 194$  participants). In the model, we adjusted age, gender, and population PCs to account for any population stratification and relatedness, focusing on risk alleles with common frequency ( $MAF > 1\%$ ;  $n = 481,395$ ). The Manhattan plot of the ExWAS revealed top SNP signals on chromosome 7 and chromosome 9 as potential risk loci for the circulating 25(OH)D levels at a  $p$ -value  $\leq 1.0 \times 10^{-5}$  (Figure 1A). These variants map to novel loci not previously associated with 25(OH)D levels. Our results showed no evidence of widespread inflation or significant population stratification ( $\lambda_{CC} = 1.002$ , standard error (SE) =  $4.9 \times 10^{-6}$ ), as shown in the Q-Q plot (Figure 1B).

The strongest association signal is at SNP rs141064014 (chromosome 7q34: 141736273: G>A) at a  $p$ -value of  $4.40 \times 10^{-6}$  (Table 2), in the intronic region of the maltase-glucoamylase (MGAM) gene. The pairwise LD assessment defined that the rs141064014 is not significantly correlated ( $r^2 < 0.2$ ) with any other SNPs within a genomic region spanning 250 Mb (Figure 2A). This indicates that rs141064014 is independent and shows no strong genetic correlation with nearby genetic markers over a considerable genomic distance. Another strong association is SNP rs7036592 (chromosome 9: 96425777: C>T) at a  $p$ -value of  $8.43 \times 10^{-6}$  in the plant homeodomain (PHD) finger 2 (PHF2) gene (Table 2), which is in LD with several other nearby SNPs in the same locus (Figure 2B). Variants suggestively associated with 25(OH)D levels with a  $p$ -value below  $5 \times 10^{-4}$  are shown in Table S1. Furthermore, we

investigated the contribution of ExWAS SNPs to the variation in 25(OH)D levels within the Lebanese population. Through this analysis, we estimated that the heritability of 25(OH)D was approximately 29%, as determined by all the filtered SNPs.

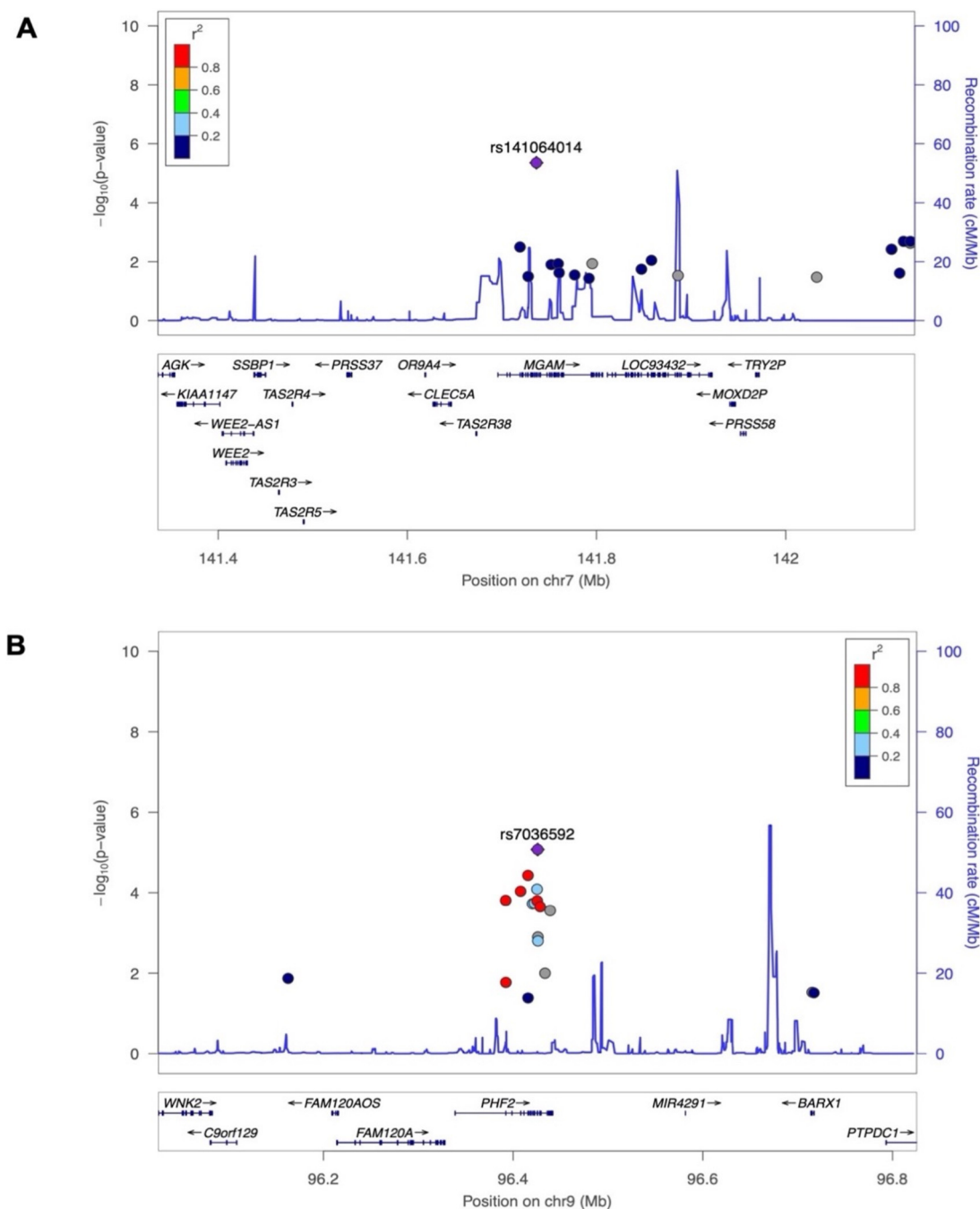
**A****B**

**Figure 1.** Manhattan plot and Q-Q plot of the ExWAS results for 25(OH)D levels. **(A)** Manhattan plot displays chromosomal positions of genetic variants (*x*-axis), and the  $-\log_{10} p$ -value (*y*-axis) with a horizontal grey line represents the top signals of  $p$ -value  $< 1 \times 10^{-5}$ . The plot shows novel genomic regions on chromosome 7 and 9 that exceed the significance threshold of vitamin D ExWAS ( $n = 481,395$  variants). **(B)** Q-Q plot displays a good fit between the observed  $-\log_{10} p$ -values (*y*-axis) with expected  $-\log_{10} p$ -values (*x*-axis), indicating that the ExWAS results are not biased and are consistent with the null hypothesis. The red line depicts the expected distribution of the  $p$ -values, while the black line represents the observed distribution. The black dashed lines represent the lower and upper limits of the 95% confidence interval.

**Table 2.** Variants of exome-wide association study for 25-hydroxyVitamin D levels.

SNP	Gene	HGVS ID	CHR	Position	A1	A2	Beta	SE (Beta)	p-Value
rs141064014	MGAM	NC_000007.13:g.141736273G>A	7	141736273	G	A	−2.38	0.52	$4.4 \times 10^{-6}$
rs7036592	PHF2	NC_000009.11:g.96425777C>T	9	96425777	C	T	−0.54	0.12	$8.4 \times 10^{-6}$

Statistical summary of the Lebanese ExWAS analysis using linear mixed models' correction of principal population components, sex, and age with a  $p$ -value  $< 1 \times 10^{-5}$ . We used GRCh37/hg19 genome reference. Columns are SNP, single-nucleotide polymorphism rs ID; Gene, mapped genes affected by a variant from ANNOVAR; HGVS ID, Human Genome Variation Society sequence variant nomenclature descriptions from Ensembl; CHR, chromosome; A1, reference allele; BP, base pair; A2, alternative allele; Beta, variant effect size determined based on the allele A1; SE, the standard error for Beta; MAF, minor allele frequency determined based on the allele A1;  $p$ -value,  $p$ -value of ExWAS analysis. Abbreviations: *MGAM*, Maltase-glucoamylase; *PHF2*, PHD finger 2.



**Figure 2.** Genetic regions surrounding the top novel genetic variant that is associated with 25(OH)D levels using LocusZoom visualization tool. (A) the *MGAM* rs141064014 on chromosome 9; (B) the

*PHF2* rs7036592 on chromosome 7 (top variant—shown in purple diamonds). On the Manhattan plot, the *p*-values are represented on the left vertical axis in log 10 scale, while the recombination rates are displayed as a blue line on the right vertical axis. The horizontal axis represents the chromosomal positions. In the bottom panel, the names and locations of genes are provided. The genes within the region are annotated and represented as arrows. The linkage disequilibrium relationships, indicated by the  $r^2$  values, between each SNP and the lead SNP are depicted using different colors.

### 3.3. Evaluation of Common Loci Replication

To evaluate the extent of replication, we performed a comparison between our findings to the largest vitamin D GWAS from the UK Biobank (GCST90000616). The GCST90000616 study, equivalent to our study, used an association analysis approach and rank-based inverse normalization to analyze 25(OH)D levels [12], enabling the comparison of effect weights for the identified loci in both cohorts. In the Lebanese cohort, we observed the replication of 7151 loci that showed significant association in the GCST90000616 study. Specifically, we replicated 58 variants that exhibited suggestive significance in the GCST90000616 study (Table S2).

A significant association was observed between the allele frequencies of the common SNPs, which exhibited a Pearson's coefficient (*R*) of 0.62 (95% CI = 0.26 to 0.83) at a *p*-value of 0.0025. In addition, the correlation analysis of effect weights and directions for the common variants showed consistent directionality, with larger effect weights than the findings of the GCST90000616 ( $n = 12$ ,  $R = 0.92$ , 95% CI = 0.74 to 0.98, *p*-value < 0.0001). The other variants ( $n = 46$  variants) exhibited an opposite association direction as compared to GCST90000616, which could be influenced by genetic diversity within the populations, study design, and environmental factors that contribute to 25(OH)D levels.

### 3.4. Meta-Analysis of Vitamin D GWAS

We then conducted a meta-analysis to uncover potential new SNPs demonstrating significant genome-wide associations in the Lebanese dataset. This analysis involved combining our ExWAS results with summary statistics obtained from the largest European GWAS conducted by Revez JA et al. for vitamin D, consisting of 417,580 individuals. The replication GCST90000616 study has been previously described in detail [12]. Through our meta-analysis, we validated the replication of a missense variant rs2725405 (chromosome 17:79220224 C>G) at a *p*-value of  $3.73 \times 10^{-8}$ , Beta = 0.0109. This variant is in the solute carrier family 38A10 (*SLC38A10*) gene that had been previously reported in the GWAS Catalog (*p*-value =  $2 \times 10^{-8}$ , Beta = 0.0114) [12]. We further identified several variants in known loci ( $n = 70$  variants), as presented in Table S3.

### 3.5. Analysis of Functional Variant Expression and Frequency

We compared the allele frequencies of the significant SNPs in the Lebanese samples with healthy individuals from the ALFA and gnomAD databases. The Lebanese elderly population exhibited a higher frequency of rs141064014, whereas the frequencies of rs7036592 and rs2725405 were lower compared to the European population (Table 3).

We performed a targeted tissue enrichment analysis using the GTEx map tool to gain insights into the potential functions of the top linked SNPs involved in vitamin D. We discovered notable correlations between particular alleles and decreased gene expression in the digestive tract and skin tissues when compared to other tissues. The homozygous risk allele "CC" and heterozygous "TC" of rs141064014 were significantly associated with reduced *MGAM* gene expression in the small intestine, with a *p*-value of  $1.9 \times 10^{-21}$  (Figure 3A). The presence of the "CC" risk allele of rs2725405 was associated with a significant reduction in *SLC38A10* gene expression in the colon (Figure 3C) and small intestine (Figure 3D), with *p*-values of  $1.8 \times 10^{-48}$  and  $1.5 \times 10^{-23}$ , respectively. Finally, an enriched expression of the *PHF2* gene was detected in multiple tissues through the RNA-seq public resource ARCHS4, including chondrocytes of bone marrow and osteoblasts in

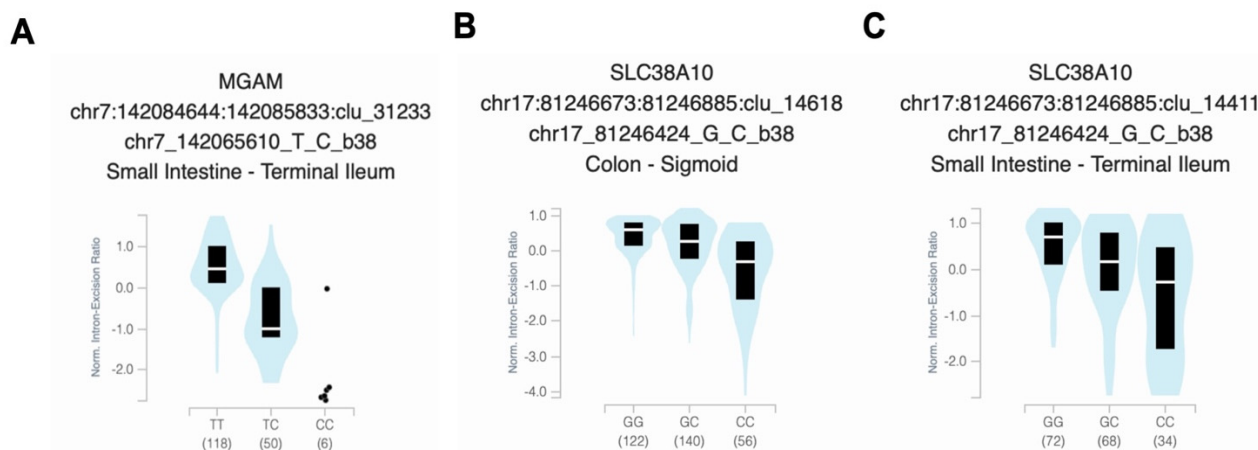


bone tissues. The normalized gene expression, measured as transcript per million (nTPM), was approximately 9.7 (Figure S1).

**Table 3.** Allele frequencies of the significant SNPs across different populations.

Populations	Frequency for rs141064014 in <i>MGAM</i>	Frequency for rs7036592 in <i>PHF2</i>	Frequency for rs2725405 in <i>SLC38A10</i>
Lebanese elderly population	0.0103	0.2408	0.4845
European population of ALFA	0.00794	0.39533	0.5729
Controls of gnomAD populations			
European	0.00634	0.3829	0.5468
East Asian	0.001	0.2136	0.3467
African	0.001	0.2783	0.9234
All populations	0.00553	0.3216	0.5084

Abbreviations: gnomAD, Genome Aggregation Database; ALFA, Allele Frequency Aggregator.

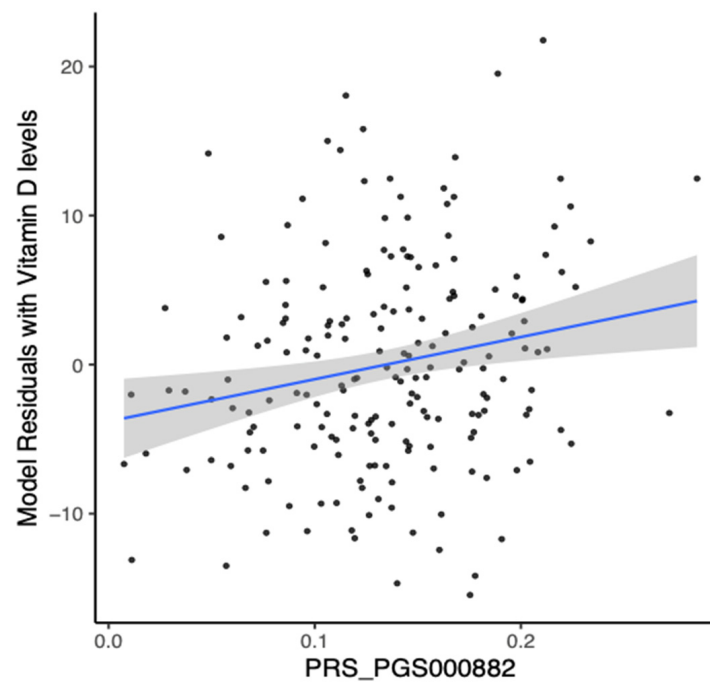


**Figure 3.** Relationship between the genotypes of vitamin D-associated SNPs and the gene expression enrichment. The bean plots display the normalized intron–excision ratio and their median (indicated by a white horizontal line) and interquartile range (represented by a black box) for (A) expression of *MGAM* rs141064014 in the small intestine with a significant  $p$ -value of  $1.9 \times 10^{-21}$ ; the expression of *SLC38A10* rs2725405 in (B) the colon with a significant  $p$ -value of  $1.8 \times 10^{-48}$ ; and (C) the small intestine with a significant association with  $p$ -value of  $1.5 \times 10^{-23}$ . The data presented are derived from the GTEx database.

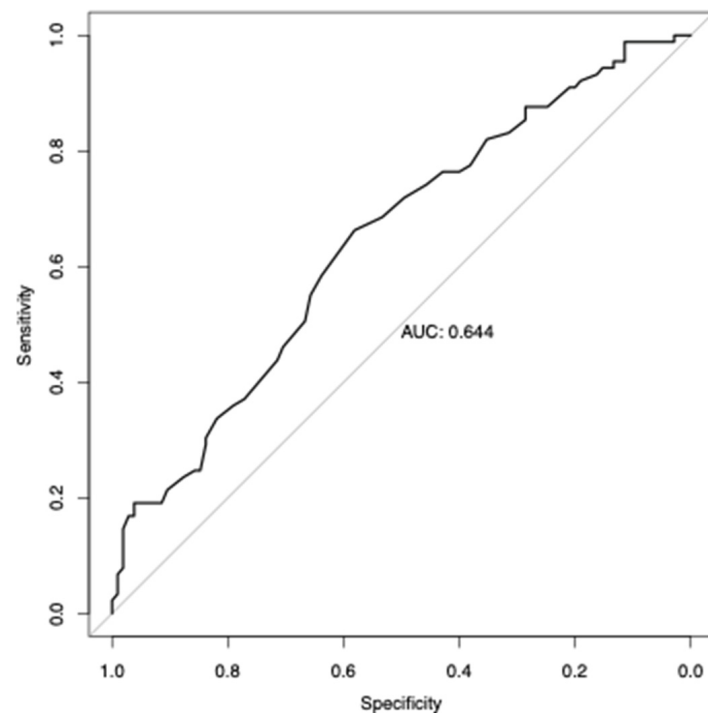
### 3.6. Analysis of Polygenic Risk Score

We then assessed the performance of a European-derived PRS from panel PGS000882 [12]. Out of the 1,094,650 variants in this panel, we detected 41,736 in our whole-exome sequence data. The performance of these polygenic risk scores in predicting the levels of 25(OH)D in the Lebanese individuals is shown in Figure 4.

The predictive performance of the European-derived PRS demonstrated lower results on the Lebanese cohort ( $R = 0.2033$ , 95% CI = 0.0643 to 0.3346,  $p$ -value = 0.0044) compared to the reported  $R$  values of 0.31 in the PGS000882 study. Despite this, the derived-PRS was able to efficiently predict the risk of vitamin D deficiency in Lebanese individuals, with an AUC of 0.644 ( $p$ -value = 0.0192, Odds Ratio = 1.6, 95% CI = 1.3 to 13.7) (Figure 5).



**Figure 4.** Performance of polygenic risk score derived from European population on the Lebanese individuals. Linear regression of the baseline levels of 25(OH)D and European-derived polygenic risk scores (PRS) (PGS000882:  $R = 0.2033$ ,  $p$ -value = 0.0044). The blue line corresponds to the linear regression analysis, representing the best fit of the data.



**Figure 5.** Performance of vitamin D deficiency prediction using European-derived PRS. The receiver operating characteristic (ROC) curve of the European-derived PRS was applied to evaluate the performance of predicting vitamin D deficiency (25(OH)D < 20 ng/mL) in Lebanese individuals.

#### 4. Discussion

Our study represents the first known attempt to detect common and new genetic SNPs that are linked to 25(OH)D levels in elderly individuals from the Middle East. Through a whole-exome analysis of 194 elderly Lebanese individuals, we discovered essential loci with suggestive evidence of an association with levels of vitamin D. The main findings of our study deviate from previous GWAS, where the most significant associated loci are located within four major genes, namely *GC*, *CYP2R1*, *CYP24A1*, and *DHCR7* [32]. Our research sheds light on previously unidentified genes that potentially contribute to the process of intestinal absorption of vitamin D. Importantly, this observation may be influenced by the specific characteristics of our cohort, which primarily consists of elderly individuals. Further validation of our results in independent populations, with an expanded sample size and utilizing a GWAS approach, would be advantageous. Indeed, such replication is the most reliable method to validate our findings.

In our study, the top genetic predisposition was the novel *MGAM* locus, rs141064014, with no indication of LD. The *MGAM* encodes the maltase-glucoamylase protein belonging to the glycoside hydrolase family 31 that breaks down complex carbohydrates in the small intestine [33]. In elderly individuals, reduced *MGAM* activity may exacerbate age-related changes in the digestive system, leading to reduced intestinal and hepatic enzymes, intestinal motility, and malabsorption of nutrients [34]. Our functional analyses showed that the “C” risk allele significantly reduces the intestinal expression of *MGAM*, which is highly elevated in our regional populations compared to other populations. These observations suggest a potential regulatory role in vitamin D absorption in the intestine among Middle Easterners, which requires further investigation.

We also identified a novel suggestive SNP, rs7036592, in the *PHF2* gene belonging to the Jumonji C family. *PHF2* is essential in several biological processes, such as the regulation of gene expression, different tissue functions, metabolism, and adipogenesis [35]. Interestingly, there is evidence that *PHF2* might be involved in regulating vitamin D metabolism and signaling [36]. *PHF2* has been shown to physically interact with and regulate the activity of *CYP27B1*, thereby affecting the synthesis of 1,25(OH)<sub>2</sub>D [37]. On the other hand, *PHF2* was found to involve in and enhance the transcriptional activity exhibited by the vitamin D receptor in osteoblasts, suggesting a potential role in regulating vitamin D-dependent gene expression [38]. Recent studies indicate that *PHF2* plays a crucial role in controlling the methylation pattern and subsequent expression of genes responsible for osteoblast differentiation in mice. Moreover, deleting *PHF2* in mice results in inadequate bone formation [39]. The high expression of *PHF2* in bone tissues suggests direct links with vitamin D in regulating osteoblast differentiation, which requires further investigation.

The heritability of vitamin D among Middle Easterners is currently unknown and requires further investigation. Previous GWAS has approximated the heritability of vitamin D to range from 7.5% to 16% among Europeans [10,13]. We found that the SNP-based heritability of vitamin D in the Lebanese group surpassed the estimation observed in UK Biobank participants, with a higher estimate of approximately 29%. This observation may be due to several factors, including study design, genetic diversity, and environmental exposures, which lead to increased heritability estimates [13]. Further investigation is needed to examine the underlying mechanisms driving these differences.

To ensure the validity and reliability of our findings, we examined data from the UK Biobank since the frequency and impact of alleles may differ across populations. Our analysis of this dataset enabled us to confirm several vitamin D-related variants identified in the GCST90000616 study, indicating the consistency of our results. Nonetheless, we acknowledge that differences in the genetic backgrounds of study populations and environmental exposures may have led to variants with opposing effect sizes. In order to better understand the reasons for opposing effect sizes and insufficient statistical power, further investigation may be needed, including studies in larger and more diverse populations.

To enhance the statistical significance of our observations, we merged our findings from Lebanese samples with data from the largest European GWAS [12]. Our meta-

analysis has revealed additional SNPs related to 25(OH)D that were not considered in the GWAS catalog. Furthermore, we have confirmed the replication of multiple SNPs from the GCST90000616 study, including a missense variant, rs2725405, located in the *SLC38A10* gene. This gene is responsible for regulating protein transportation, synthesis, and cellular stress responses. In some cases, SLC38A10 protein can also mediate the intestinal transportation of some vitamins into the blood and lipid metabolism [40]. Notably, the expression of *SLC38A10* decreases in the presence of the “C” risk allele in the intestine and skin. The enrichment of *SLC38A10* in intestinal and skin tissues and regional populations, suggesting possible mechanisms in vitamin D absorption and metabolism. However, more research is needed to understand how it may impact vitamin D status.

While general recommendations of vitamin D deficiency management provide a foundation [4], personalized approaches hold value in identifying individuals at higher risk of vitamin D deficiency [7]. Various factors, such as genetic variations and environmental factors, can contribute to individual variations in vitamin D status and response to supplementation [6]. Therefore, personalized approaches can optimize interventions and address the specific needs of individuals of vitamin D deficiency. In our study, we derived PRS for predicting 25(OH)D levels in Lebanese individuals from European populations. The effectiveness of European-derived PRS in the Lebanese population was lower than that estimated in Europeans. This variation in performance might be due to factors such as variations in variant effect weights and frequencies of causal alleles across ethnicities, genetic heterogeneity, and the number of variants and participants utilized in the study [41]. These findings emphasize the need for a larger genome-wide association study tailored specifically for the Middle Eastern population to improve the performance of PRS estimation. Nevertheless, our polygenic risk score model for vitamin D demonstrated predictive ability in estimating vitamin D deficiency in the Lebanese cohort. The modest performance of the PRS in predicting 25(OH)D levels in individuals of European and Lebanese descent, with R values of 0.31 and 0.2033, respectively, may suggest the influence of non-genetic factors related to vitamin D deficiency, such as inadequate sunlight exposure and lifestyle/environmental factors. Therefore, the development of a more effective risk score tool for 25(OH)D levels may require the incorporation of such factors.

## 5. Conclusions

Overall, we explored the first suggestive evidence of an association between several loci and levels of 25(OH)D in an elderly Middle Eastern population through our ExWAS. Our study showed that ExWAS can more easily identify the genes and biological pathways associated with Mendelian phenotypes. The results of our PRS model may provide a new avenue for guiding the personalized treatment of vitamin D deficiency. However, further replications with increased sample sizes are necessary to confirm the potential of these findings and advance their application in the development of personalized medicine.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/nu15143216/s1>, Table S1: Single nucleotide polymorphisms identified in genome-wide analyses for circulating 25-hydroxyvitamin D concentrations in Lebanese samples; Table S2: Replication of loci identified by European GWAS (Revez JA et al., 2020) for vitamin D in the Lebanese GWAS; Table S3: Variants Discovered from a Meta-analysis of the Lebanese GWAS and GCST90000616 European study for Vitamin D Levels; Figure S1: Tissue Expression of PHF2 mRNA.

**Author Contributions:** G.N., G.E.-H.F. and M.C. conceived the work and secured funding. D.S.B., M.C. and G.E.-H.F. carried out the clinical analysis. N.N.H. carried out the ExWAS analysis, meta-analysis, and PRS analysis, generated all the results, and authored the initial draft of the manuscript. N.N.H., Y.A.-S. and O.A. wrote the codes for the computational analyses. All authors participated in analyzing the final data and the write-up of the current version of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the American University of Beirut (IM-GEHF-20 on 24 January 2011).

**Informed Consent Statement:** Informed consent was obtained from all study participants prior to their involvement.

**Data Availability Statement:** All data produced during the study have been examined and incorporated into this published article or documented in the referenced data repositories. Publicly available software tools were used for all analyses, as indicated in the primary body of the text and Methods sections.

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

## References

1. Ruiz-Garcia, A.; Pallares-Carratala, V.; Turegano-Yedro, M.; Torres, F.; Sapena, V.; Martin-Gorgojo, A.; Martin-Moreno, J.M. Vitamin D Supplementation and Its Impact on Mortality and Cardiovascular Outcomes: Systematic Review and Meta-Analysis of 80 Randomized Clinical Trials. *Nutrients* **2023**, *15*, 1810. [[CrossRef](#)] [[PubMed](#)]
2. Zhang, X.; Niu, W. Meta-analysis of randomized controlled trials on vitamin D supplement and cancer incidence and mortality. *Biosci. Rep.* **2019**, *39*, BSR20190369. [[CrossRef](#)] [[PubMed](#)]
3. Rahme, M.; Sharara, S.L.; Baddoura, R.; Habib, R.H.; Halaby, G.; Arabi, A.; Singh, R.J.; Kassem, M.; Mahfoud, Z.; Hoteit, M.; et al. Impact of Calcium and Two Doses of Vitamin D on Bone Metabolism in the Elderly: A Randomized Controlled Trial. *J. Bone Miner. Res.* **2017**, *32*, 1486–1495. [[CrossRef](#)] [[PubMed](#)]
4. Cashman, K.D.; Ritz, C.; Kiely, M.; Odin, C. Improved Dietary Guidelines for Vitamin D: Application of Individual Participant Data (IPD)-Level Meta-Regression Analyses. *Nutrients* **2017**, *9*, 469. [[CrossRef](#)]
5. Perna, S. Is Vitamin D Supplementation Useful for Weight Loss Programs? A systematic review and meta-analysis of randomized controlled trials. *Medicina* **2019**, *55*, 368.
6. Mitchell, B.L.; Zhu, G.; Medland, S.E.; Renteria, M.E.; Eyles, D.W.; Grasby, K.L.; McGrath, J.J.; Martin, N.G. Half the Genetic Variance in Vitamin D Concentration is Shared with Skin Colour and Sun Exposure Genes. *Behav. Genet.* **2019**, *49*, 386–398. [[CrossRef](#)]
7. Ammar, M.; Heni, S.; Tira, M.S.; Khalij, Y.; Hamdouni, H.; Amor, D.; Ksibi, S.; Omezzine, A.; Bouslama, A. Variability in response to vitamin D supplementation according to vitamin D metabolism related gene polymorphisms in healthy adults. *Eur. J. Clin. Nutr.* **2023**, *77*, 189–194. [[CrossRef](#)]
8. Datta, P.; Philipsen, P.A.; Olsen, P.; Petersen, B.; Andersen, J.D.; Morling, N.; Wulf, H.C. Pigment genes not skin pigmentation affect UVB-induced vitamin D. *Photochem. Photobiol. Sci.* **2019**, *18*, 448–458. [[CrossRef](#)]
9. Wjst, M.; Altmuller, J.; Braig, C.; Bahnweg, M.; Andre, E. A genome-wide linkage scan for 25-OH-D(3) and 1,25-(OH)<sub>2</sub>-D<sub>3</sub> serum levels in asthma families. *J. Steroid Biochem. Mol. Biol.* **2007**, *103*, 799–802. [[CrossRef](#)]
10. Manousaki, D.; Mitchell, R.; Dudding, T.; Haworth, S.; Harroud, A.; Forgetta, V.; Shah, R.L.; Luan, J.; Langenberg, C.; Timpson, N.J.; et al. Genome-wide Association Study for Vitamin D Levels Reveals 69 Independent Loci. *Am. J. Hum. Genet.* **2020**, *106*, 327–337. [[CrossRef](#)]
11. Sinnott-Armstrong, N.; Tanigawa, Y.; Amar, D.; Mars, N.; Benner, C.; Aguirre, M.; Venkataraman, G.R.; Wainberg, M.; Ollila, H.M.; Kiiskinen, T.; et al. Author Correction: Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat. Genet.* **2021**, *53*, 1622. [[CrossRef](#)]
12. Revez, J.A.; Lin, T.; Qiao, Z.; Xue, A.; Holtz, Y.; Zhu, Z.; Zeng, J.; Wang, H.; Sidorenko, J.; Kemper, K.E.; et al. Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration. *Nat. Commun.* **2020**, *11*, 1647. [[CrossRef](#)] [[PubMed](#)]
13. Jiang, X.; O'Reilly, P.F.; Aschard, H.; Hsu, Y.-H.; Richards, J.B.; Dupuis, J.; Ingelsson, E.; Karasik, D.; Pilz, S.; Berry, D.; et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat. Commun.* **2018**, *9*, 260. [[CrossRef](#)] [[PubMed](#)]
14. Autier, P.; Mullie, P.; Macacu, A.; Dragomir, M.; Boniol, M.; Coppens, K.; Pizot, C.; Boniol, M. Effect of vitamin D supplementation on non-skeletal disorders: A systematic review of meta-analyses and randomised trials. *Lancet Diabetes Endocrinol.* **2017**, *5*, 986–1004. [[CrossRef](#)]
15. Lips, P.; Cashman, K.D.; Lamberg-Allardt, C.; Bischoff-Ferrari, H.A.; Obermayer-Pietsch, B.; Bianchi, M.L.; Stepan, J.; El-Hajj Fuleihan, G.; Bouillon, R. Current vitamin D status in European and Middle East countries and strategies to prevent vitamin D deficiency: A position statement of the European Calcified Tissue Society. *Eur. J. Endocrinol.* **2019**, *180*, 23–54. [[CrossRef](#)]
16. Chakhtoura, M.; Rahme, M.; Chamoun, N.; Fuleihan, G.E.-H. Vitamin D in the Middle East and North Africa. *Bone Rep.* **2018**, *8*, 135–146. [[CrossRef](#)]
17. Salman, S.; Khouzami, M.; Harb, M.; Saleh, B.; Boushnak, M.O.; Moussa, M.K.; Mohsen, Z.H. Prevalence and Predictors of Vitamin D Inadequacy: A Sample of 2,547 Patients in a Mediterranean Country. *Cureus* **2021**, *13*, e14881. [[CrossRef](#)]
18. McMahon, A.; Lewis, E.; Buniello, A.; Cerezo, M.; Hall, P.; Sollis, E.; Parkinson, H.; Hindorf, L.A.; Harris, L.W.; MacArthur, J.A. Sequencing-based genome-wide association studies reporting standards. *Cell Genom.* **2021**, *1*, 100005. [[CrossRef](#)] [[PubMed](#)]

19. McCaw, Z.R.; Lane, J.M.; Saxena, R.; Redline, S.; Lin, X. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. *Biometrics* **2020**, *76*, 1262–1272. [[CrossRef](#)]
20. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
21. Cingolani, P.; Platts, A.; Wang, L.L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)] [[PubMed](#)]
22. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
23. Sherry, S.T.; Ward, M.; Sirotkin, K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* **1999**, *9*, 677–679. [[CrossRef](#)] [[PubMed](#)]
24. Landrum, M.J.; Lee, J.M.; Benson, M.; Brown, G.R.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **2018**, *46*, D1062–D1067. [[CrossRef](#)]
25. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [[CrossRef](#)]
26. Zhou, W.; Nielsen, J.B.; Fritsche, L.G.; Dey, R.; Gabrielsen, M.E.; Wolford, B.N.; LeFaive, J.; VandeHaar, P.; Gagliano, S.A.; Gifford, A.; et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet.* **2018**, *50*, 1335–1341. [[CrossRef](#)]
27. Pruim, R.J.; Welch, R.P.; Sanna, S.; Teslovich, T.M.; Chines, P.S.; Gliedt, T.P.; Boehnke, M.; Abecasis, G.R.; Willer, C.J. LocusZoom: Regional visualization of genome-wide association scan results. *Bioinformatics* **2010**, *26*, 2336–2337. [[CrossRef](#)] [[PubMed](#)]
28. Buniello, A.; MacArthur, J.A.L.; Cerezo, M.; Harris, L.W.; Hayhurst, J.; Malangone, C.; McMahon, A.; Morales, J.; Mountjoy, E.; Sollis, E.; et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **2019**, *47*, D1005–D1012. [[CrossRef](#)]
29. Lambert, S.A.; Gil, L.; Jupp, S.; Ritchie, S.C.; Xu, Y.; Buniello, A.; McMahon, A.; Abraham, G.; Chapman, M.; Parkinson, H.; et al. The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nat. Genet.* **2021**, *53*, 420–425. [[CrossRef](#)]
30. Hunt, S.E.; Moore, B.; Amode, R.M.; Armean, I.M.; Lemos, D.; Mushtaq, A.; Parton, A.; Schuilenburg, H.; Szpak, M.; Thormann, A.; et al. Annotating and prioritizing genomic variants using the Ensembl Variant Effect Predictor—A tutorial. *Hum. Mutat.* **2022**, *43*, 986–997. [[CrossRef](#)]
31. Lachmann, A.; Torre, D.; Keenan, A.B.; Jagodnik, K.M.; Lee, H.J.; Wang, L.; Silverstein, M.C.; Ma’ayan, A. Massive mining of publicly available RNA-seq data from human and mouse. *Nat. Commun.* **2018**, *9*, 1366. [[CrossRef](#)] [[PubMed](#)]
32. Wang, T.J.; Zhang, F.; Richards, J.B.; Kestenbaum, B.; van Meurs, J.B.; Berry, D.; Kiel, D.P.; Streeten, E.A.; Ohlsson, C.; Koller, D.L.; et al. Common genetic determinants of vitamin D insufficiency: A genome-wide association study. *Lancet* **2010**, *376*, 180–188. [[CrossRef](#)] [[PubMed](#)]
33. Zhang, E.; Chen, Y.; Bao, S.; Hou, X.; Hu, J.; Mu, O.Y.N.; Song, Y.; Shan, L. Identification of subgroups along the glycolysis-cholesterol synthesis axis and the development of an associated prognostic risk model. *Hum. Genom.* **2021**, *15*, 53. [[CrossRef](#)]
34. Chiruvella, V.; Cheema, A.; Arshad, H.M.S.; Chan, J.T.; Yap, J.E.L. Sucrase-Isomaltase Deficiency Causing Persistent Bloating and Diarrhea in an Adult Female. *Cureus* **2021**, *13*, e14349. [[CrossRef](#)] [[PubMed](#)]
35. Okuno, Y.; Ohtake, F.; Igarashi, K.; Kanno, J.; Matsumoto, T.; Takada, I.; Kato, S.; Imai, Y. Epigenetic regulation of adipogenesis by PHF2 histone demethylase. *Diabetes* **2013**, *62*, 1426–1434. [[CrossRef](#)]
36. Luo, P.; Zhang, Y.-D.; He, F.; Tong, C.-J.; Liu, K.; Liu, H.; Zhu, S.-Z.; Luo, J.-Z.; Yuan, B. HIF-1 $\alpha$ -mediated augmentation of miRNA-18b-5p facilitates proliferation and metastasis in osteosarcoma through attenuation PHF2. *Sci. Rep.* **2022**, *12*, 10398. [[CrossRef](#)] [[PubMed](#)]
37. Pereira, F.; Barbáchano, A.; Singh, P.K.; Campbell, M.J.; Muñoz, A.; Larriba, M.J. Vitamin D has wide regulatory effects on histone demethylase genes. *Cell Cycle* **2012**, *11*, 1081–1089. [[CrossRef](#)]
38. Sawatsubashi, S.; Nishimura, K.; Mori, J.; Kouzmenko, A.; Kato, S. The Function of the Vitamin D Receptor and a Possible Role of Enhancer RNA in Epigenomic Regulation of Target Genes: Implications for Bone Metabolism. *J. Bone Metab.* **2019**, *26*, 3–12. [[CrossRef](#)]
39. Kim, H.J.; Park, J.W.; Lee, K.H.; Yoon, H.; Shin, D.H.; Ju, U.I.; Seok, S.H.; Lim, S.H.; Lee, Z.H.; Kim, H.H.; et al. Plant homeodomain finger protein 2 promotes bone formation by demethylating and activating Runx2 for osteoblast differentiation. *Cell Res.* **2014**, *24*, 1231–1249. [[CrossRef](#)]
40. Tripathi, R.; Aggarwal, T.; Lindberg, F.A.; Klemm, A.H.; Fredriksson, R. SLC38A10 Regulate Glutamate Homeostasis and Modulate the AKT/TSC2/mTOR Pathway in Mouse Primary Cortex Cells. *Front Cell Dev. Biol.* **2022**, *10*, 854397. [[CrossRef](#)]
41. Patel, R.A.; Musharoff, S.A.; Spence, J.P.; Pimentel, H.; Tcheandjieu, C.; Mostafavi, H.; Sinnott-Armstrong, N.; Clarke, S.L.; Smith, C.J.; Program, V.A.M.V.; et al. Genetic interactions drive heterogeneity in causal variant effect sizes for gene expression and complex traits. *Am. J. Hum. Genet.* **2022**, *109*, 1286–1297. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.