



OPEN

Vitamin D-related polymorphisms and vitamin D levels as risk biomarkers of COVID-19 disease severity

Ana Teresa Freitas^{2,3}, Conceição Calhau^{4,5}, Gonçalo Antunes², Beatriz Araújo¹¹, Matilde Bandeira^{6,15}, Sofia Barreira^{6,15}, Filipa Bazenga¹¹, Sandra Braz⁷, Daniel Caldeira^{1,10}, Susana Constantino Rosa Santos¹, Ana Faria⁴, Daniel Faria^{3,13}, Marta Fraga⁹, Beatriz Nogueira-Garcia¹⁰, Lúcia Gonçalves², Pavlo Kovalchuk², Luísa Lacerda¹¹, Hugo Lopes², Daniel Luís², Fábio Medeiros⁸, Ana M. P. Melo^{13,14}, José Melo-Cristino⁹, Ana Miranda⁹, Clara Pereira², Ana Teresa Pinto¹, João Pinto¹¹, Helena Proença⁹, Angélica Ramos^{11,12}, João P. R. Rato^{13,14}, Filipe Rocha¹, Júlio César Rocha^{4,5}, André Moreira-Rosário^{4,5}, Helena Vazão², Yuliya Volovetska⁹, João-Tiago Guimarães^{11,12} & Fausto J. Pinto^{1,10}✉

Vitamin D is a fundamental regulator of host defences by activating genes related to innate and adaptive immunity. Previous research shows a correlation between the levels of vitamin D in patients infected with SARS-CoV-2 and the degree of disease severity. This work investigates the impact of the genetic background related to vitamin D pathways on COVID-19 severity. For the first time, the Portuguese population was characterized regarding the prevalence of high impact variants in genes associated with the vitamin D pathways. This study enrolled 517 patients admitted to two tertiary Portuguese hospitals. The serum concentration of 25 (OH)D, was measured in the hospital at the time of patient admission. Genetic variants, 18 variants, in the genes *AMDHD1*, *CYP2R1*, *CYP24A1*, *DHCR7*, *GC*, *SEC23A*, and *VDR* were analysed. The results show that polymorphisms in the vitamin D binding protein encoded by the *GC* gene are related to the infection severity ($p = 0.005$). There is an association between vitamin D polygenic risk score and the serum concentration of 25 (OH)D ($p = 0.04$). There is an association between 25 (OH)D levels and the survival and fatal outcomes ($p = 1.5e-4$). The Portuguese population has a higher prevalence of the *DHCR7* RS12785878 variant when compared with its prevalence in the European population (19% versus 10%). This study shows a genetic susceptibility for vitamin D deficiency that might explain higher severity degrees in COVID-19 patients. These results reinforce the relevance of personalized strategies in the context of viral diseases.

Trial registration: NCT04370808.

¹Centro Cardiovascular da Universidade de Lisboa (CCUL), CAML, Faculdade de Medicina, Universidade de Lisboa, 1649-028 Lisboa, Portugal. ²HeartGenetics, Genetics and Biotechnology SA, Biocant Park, Cantanhede, Portugal. ³INESC-ID, Instituto Superior Técnico, University of Lisbon, Lisboa, Portugal. ⁴Nutrition and Metabolism, NOVA Medical School, Universidade Nova de Lisboa, Lisboa, Portugal. ⁵CINTESIS - Center for Health Technology and Services Research, Porto, Portugal. ⁶Rheumatology Department, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, Centro Académico de Medicina de Lisboa, Lisboa, Portugal. ⁷Internal Medicine Department, Santa Maria Hospital, Centro Hospitalar Universitário Lisboa Norte, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal. ⁸Infectious Disease Department, Santa Maria Hospital - CHULN, Lisboa, Portugal. ⁹Clinical Pathology Department, Hospital Santa Maria, Centro Hospitalar Universitário Lisboa Norte, EPE, Lisbon, Portugal. ¹⁰Cardiology Department, Hospital Universitário de Santa Maria - CHULN, Lisboa, Portugal. ¹¹Clinical Pathology Department, Centro Hospitalar Universitário de São João, Porto, Portugal. ¹²Biomedicine Department, Faculdade de Medicina, EPIUnit, Instituto de Saúde Pública, Universidade do Porto, Porto, Portugal. ¹³Portuguese Infrastructure of Biological Data – BioData.pt, Lisbon, Portugal. ¹⁴Instituto Gulbenkian de Ciência, Oeiras, Portugal. ¹⁵Rheumatology Research Unit, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, CAML, Lisboa, Portugal. ✉email: faustopinto@medicina.ulisboa.pt

From the end of 2019 until now, a new coronavirus named SARS-CoV-2 has been causing a worldwide pandemic leading to many deaths due to an acute respiratory disease named COVID-19. The lack of knowledge about the pathogenesis of this disease and the reasons underlying the different clinical outcomes have been pushing society as a whole to take preventive measures and to accelerate research. During the last year, several research groups and healthcare institutions have made available evidence that indicates that an uncontrolled inflammatory response is a major cause for the occurrence of an acute respiratory distress syndrome¹.

A statistical analysis of COVID-19 pandemic data from healthcare centers across China, France, Germany, Italy, Iran, South Korea, Spain, Switzerland, the United Kingdom and the United States, performed by Northwestern University, shows a strong correlation between severe vitamin D deficiency and mortality rates².

Vitamin D has many relevant and documented roles in general health maintenance, being its deficiency particularly associated with severe impacts on the functional integrity of the immune system, such as influencing cytokine production³. It has been hypothesized that vitamin D could terminate the “cytokine storm” and oxidative stress with possible antiviral activity through its classical precursors and novel CYP11A1-derived hydroxyderivatives⁴. Vitamin D deficiency has also been linked to hypertension, autoimmune, infectious and cardiovascular diseases, also known risk factors for severe COVID-19^{5,6}.

Previous research shows that genetics contribute to up to 28% of inter-individual variability in serum 25(OH) D concentrations, while season and vitamin D intake explain another 24% of the variability⁷. Large-scale genome-wide association studies (GWAS), considering 79,366 individuals with European ancestry⁸, or considering imputed genotypes from 401,460 white British UK Biobank participants⁹, have identified relatively common single nucleotide polymorphisms which play an important biological role in vitamin D metabolism, transport, degradation and downstream pathways, to evaluate their impact on circulating 25(OH) D concentrations. Vitamin D polymorphisms research shows that not all genomes will respond to vitamin D supplementation due to the loss of function of different genes¹⁰.

In a normal scenario, it is known that vitamin D deficiency is common in Europe and the Middle East. It occurs in <20% of the population in Northern Europe, in 30–60% in Western, Southern and Eastern Europe and up to 80% in Middle East countries¹¹. Data from Portugal shows that 66% of adults present vitamin D insufficiency/deficiency¹². This high percentage in a population living in a sunny country, motivated the characterization that was performed regarding the prevalence of high impact alleles in genes associated with the vitamin D pathways in the Portuguese population when compared with the European population.

The main objective of this study was to understand if an association exists between polymorphisms in vitamin D-related genes and vitamin D levels, and between these variables and the COVID-19 severity. Secondly, we aimed to compare the frequency of the genetic variants under analysis with the observed frequency for the European population. To achieve this objective, we have assessed the genetic variants in vitamin D-related genes in hospitalized patients and the vitamin D serum levels, and evaluated the association between these data and the severity of the disease.

Results and discussion

A total of 491 patients with a laboratory confirmed positive COVID-19 test, 371 (75.6%) from Santa Maria hospital and 120 (24.4%) from São João hospital, were considered. There were 217 female and 266 male patients with mean \pm SD age of 69.7 ± 15.8 years (see supplemental data for demographic, clinical and phenotypic data).

In COVID-19 positive patients, 120 (24.4%), 311 (63.3%), and 59 (12.0%) patients had deficient, insufficient, and sufficient levels of vitamin D, respectively, using the Endocrine Society cutoff (see supplemental data). The prevalence of vitamin D deficiency was 61.7% and 68.3% in Santa Maria and São João hospitals, respectively.

Dead, severe and moderate disease were observed in 18.5%, 21.8% and 59.7% of patients, respectively. From a total of 311 patients with vitamin D deficiency, 68 died (13.8% of all patients), 69 (14% of all patients) had a severe response and 174 (35.4% of all patients) had a moderate response to COVID-19. Disease severity for each patient was defined using the World Health Organization (WHO) discrete clinical progression scale from 0 to 10¹³. In this study patients' severity levels were between 4 and 10.

Identification of the vitamin D polymorphisms as risk biomarkers. From previous large-scale GWAS, several single nucleotide polymorphisms that play an important biological role in vitamin D metabolism, transport, degradation, and downstream pathways, have been identified as having an impact on circulating 25(OH) D concentrations^{8,9}. To understand if an association exists between the polymorphisms in the vitamin D-related genes and the disease severity, four polygenic risk scores (PRSs) were defined. These scores considered contributions from different genes and were identified as: Synthesis (*DHCR7*; *CYP2R1*); Metabolism (*GC*; *CYP24A1*); Pathway (*DHCR7*; *CYP2R1*; *GC*; *CYP24A1*) and Vitamin D total (*DHCR7*; *CYP2R1*; *GC*; *CYP24A1*; *AMDHD1*; *SEC23A*).

Table 1 represents the correlations between the PRSs and COVID-19 disease severity, and the PRSs and vitamin D levels. The results show a significant positive correlation between the Metabolism score and COVID-19 disease severity, and a significant negative correlation between the Vitamin D total score and vitamin D levels.

The relation between the Vitamin D total score values and vitamin D levels can be clearly verified in Fig. 1. Higher PRSs values are associated with patients with deficient levels of vitamin D, with a higher aggregation of points at the median/high level of the PRSs in vitamin D deficient patients.

A deeper analysis of the positive correlation between the Metabolism score and COVID-19 disease severity showed that the polymorphism *GC* RS2282679 in the vitamin D binding protein encoded by the *GC* gene could explain most of the interesting correlation observed. The association of this polymorphism with the severity presents a $\rho = 0.13$ (p -value = 0.005).

Score	Severity (ρ)	Severity (p -value)	Vitamin D level (ρ)	Vitamin D level (p -value)
Synthesis	-0.04	0.39	-0.07	0.14
Metabolism	0.14	0.002	-0.07	0.12
Pathway	0.08	0.08	-0.09	0.053
Vitamin D total	0.07	0.12	-0.09	0.04

Table 1. PRSs correlation with COVID-19 disease severity and vitamin D level (Spearman correlation). The bold means that the p -value is below 0.05 which is the standard cut-off to be considered statistically significant. The significance level for a study is chosen before data collection and is typically set to 5%.

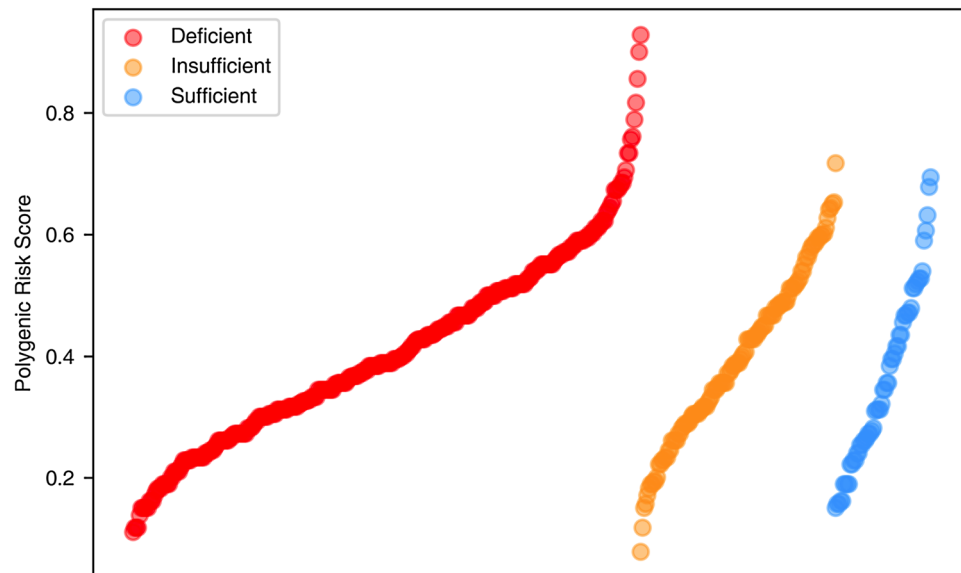


Figure 1. PRSs are represented as a continuous value. Patients were divided in three groups (Deficient—red; Insufficient—yellow; Sufficient—blue) considering its vitamin D level (a continuous variable). In each group, patients are sorted in ascending order of PRSs.

This finding was particularly relevant because only one previous GWAS meta-analysis from the COVID-19 host genetics initiative considering hospitalized vs. not hospitalized COVID-19 patients (<https://www.covid19hg.org/results/>) briefly highlighted this fact. In this meta-analysis, the GC RS2282679 correlated significantly with COVID-19 disease severity (p -value = 0.002).

Regarding the association of the GC RS2282679 polymorphism with the disease severity, it might be explained by an additional role played by the vitamin D binding protein other than the transport of vitamin D in the bloodstream. Available data indicates that this protein may also act as a neutrophil chemotactic factor and a macrophage activator, therefore actively participating in the inflammation process^{14–16}. Also, vitamin D binding protein is an extracellular scavenger for actin released from damaged/dead cells. When in excess, actin can cause intravascular coagulation resulting in multi-organ dysfunction and cardiac arrest¹⁷.

The functional consequence of this polymorphism in this pathway needs to be further explored in future studies.

Relation between vitamin D level and COVID-19 disease severity. It was observed a trend towards an increased proportion of patients with deficient levels of vitamin D (<20 ng/ml) and an increase of severity from moderate to severe, and dead. The proportion of patients with deficient levels of vitamin D was higher in the group that died (76%) when compared with the two other groups (moderate—59% and severe—64% disease). The combined proportion of patients with insufficient and sufficient vitamin D levels was 40% in the group with moderate severity, compared with 24% in the group of deceased patients (see supplemental material).

This evidence was also confirmed when a binomial variable of survival and fatal outcome was used (Fig. 2). Vitamin D levels were significantly lower in patients that died than in those that survived (Median = 11.70 [Q1 = 8.67, Q3 = 19.67] vs Median = 17.40 [Q1 = 11.00, Q3 = 24.60], respectively (p -value = 1.5×10^{-4} , Mann-Whitney U test). These results corroborate what was already observed by other studies across the world.

Differences between impact genotypes' frequency in Portuguese and European populations. The genotypes' frequency of the genetic variants under analysis was compared with the observed frequency for the European population in the 1000 Genomes Project, Fig. 3. The variant *DHCR7* RS12785878 shows a significant deviation in the impact genotypes' frequency (G/GT), 18.9% for the Portuguese population com-

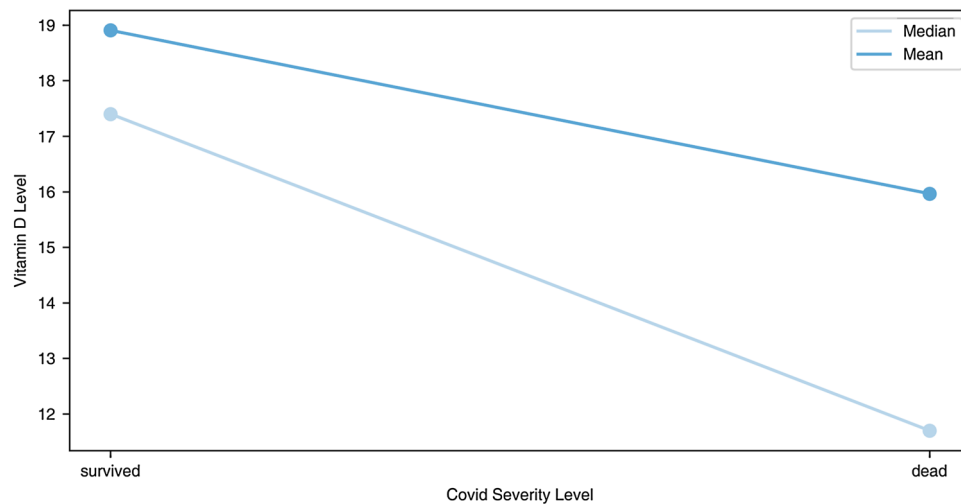


Figure 2. Vitamin D level (ng/mL) (continuous) and COVID-19 disease severity (binomial: survived, dead).

pared to 9.7% for the European population. The prevalence of the impact polymorphism *DHCR7* RS12785878 is similar in the Portuguese population in the HeartGenetics's research database with more than 8000 Portuguese individuals (data not shown), where GG genotype has a prevalence of 18.8%.

In this study, and for the first time, it is shown that the Portuguese population presents a genetic makeup for a higher predisposition to vitamin D deficiency when compared to the European population, therefore increasing the potential severity of the COVID-19 response, consequently impacting on patient outcomes. Furthermore, these results also explain, in part, the higher vitamin D deficiency in the Portuguese population in different previous publications¹². It is important to know that this metabolic vulnerability must be considered for a vitamin D supplementation clinical decision. This data proves that it is wrong to assume that sunny countries would not have an issue with vitamin D deficiency, showing that genetic characterization and vitamin D monitoring at a population level should be put in place in order to define guidelines for vitamin D intake.

One year after this pandemic scenario, it has been described that vitamin D deficiency may be a risk factor for mortality in COVID-19 patients¹⁰. On top of this, evidence revealed that supplementation with high-dose vitamin D3 booster therapy and reduced the risk of mortality¹⁸. This highlights the importance of vitamin D on COVID-19¹⁹ and reinforces the hypothesis of vitamin D as a terminator of the “cytokine storm”⁴. Moreover, it was reported an increased risk of infection with COVID-19 in individuals presenting lower levels of 25(OH) D^{20,21}, which could be in line with the suggested anti-SARS-CoV-2 activity of a range of vitamin D3-related compounds, including 7-dehydrocholesterol and lumisterol hydroxyderivatives²². Whereas vitamin D intake, sun exposure, demographics, and, specially, genes have been identified as being crucial determinants of vitamin D status, the impact of these factors is expected to be different across populations. To improve current prevention and treatment strategies, it is essential to propose novel diagnostic tools for personalization.

These results reinforce the role of vitamin D polymorphisms, in particular the *GC* RS2282679 polymorphism, and vitamin D levels as biomarkers for COVID-19 disease severity and emphasize the relevance for personalized strategies in the context of viral diseases. Whether these relationships are causal or consequential is unknown. Nevertheless, they represent targets for diagnostic surveillance, or for intervention studies.

Methods

Subjects. Between August 2020 and January 2021, 517 patients were enrolled in this project when admitted to Hospital de Santa Maria in Lisbon and Hospital de São João in Oporto. Eligibility criteria included patients aged 18 years and above, patients admitted with COVID-19 disease requiring hospitalization, with a positive test for SARS-CoV-2 by nasopharyngeal swabs using quantitative RT-PCR performed in national reference laboratories and in accordance with recommendations from the National Directorate of Health.

The RT-PCR results for the SARS-CoV-2 were obtained through one of the following tests: LightCycler® Multiplex RNA Virus Master (Roche Life Science, Penzberg, Germany) at a LightCycler® 480 Instrument II (Roche Life Science, Penzberg, Germany), Cobas® SARS-CoV-2 Test (Roche® Diagnostics) at a Cobas® 6800 (Roche® Diagnostics), Allplex™ SARS-CoV-2 Master Assay (Seegene Inc.) at a CFX96® Real-Time (Bio-Rad®), GeneFinder™ COVID-19 PLUS RealAmp Kit (OSANG Healthcare Co, Ltd.) at ELITE InGenius® (ELITechGroup®) or QuantGene 9600 (Hangzhou Bioer Technology Co. Ltd.), Xpert®Xpress SARS-CoV-2 (Cepheid®) at a GeneXpert® (Cepheid®).

All eligible patients were followed till the closure of case i.e. either discharge or mortality.

Serum 25-hydroxyvitamin D. Vitamin D status was determined by analyzing 25(OH) D concentrations in the serum samples collected at the time of patient admission. Serum 25(OH) D was determined in the local pathology laboratories by a competitive electrochemiluminescence protein-binding immunoassay using a Cobas® e411

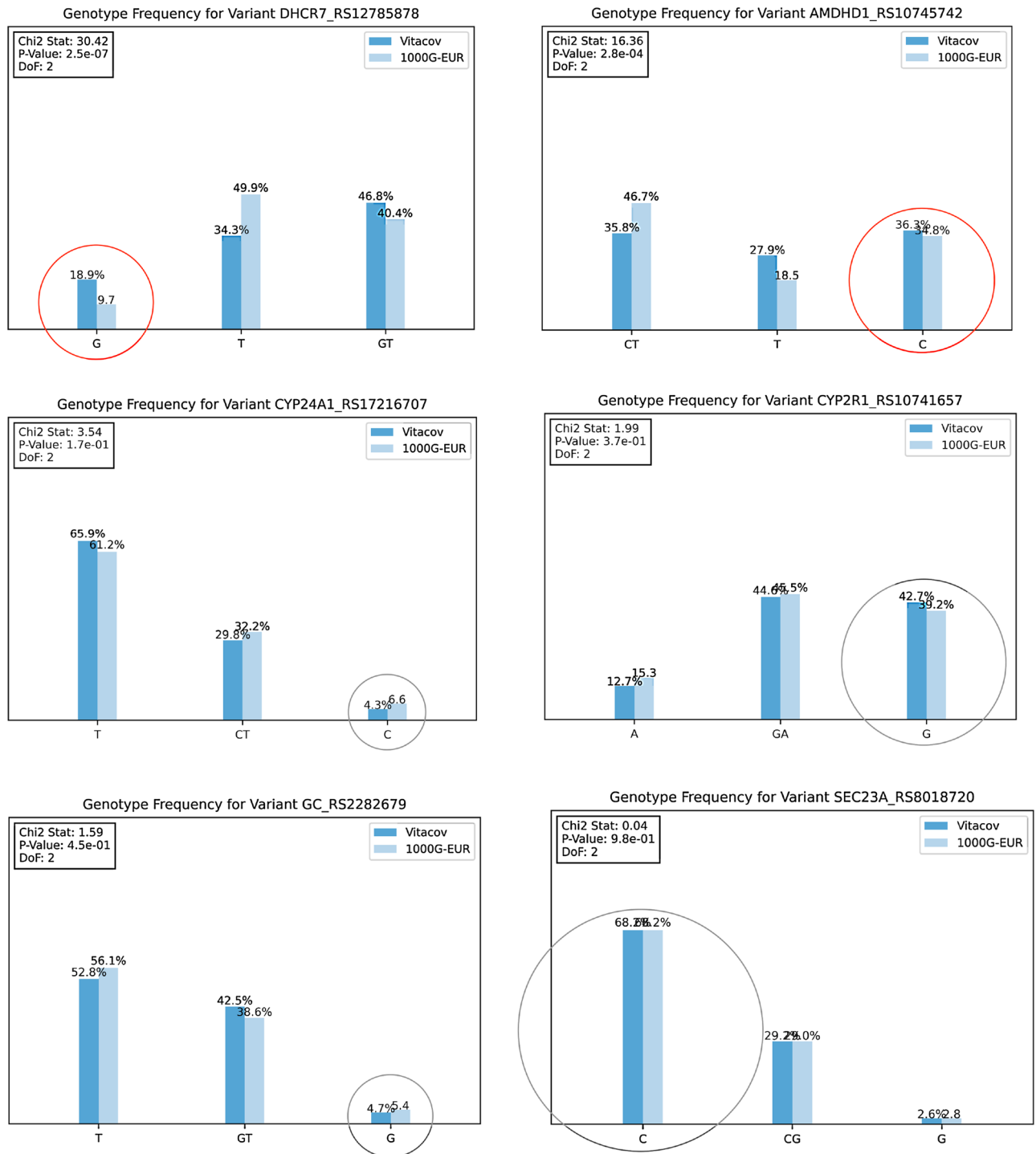


Figure 3. Comparison of the impact genotypes’ frequency in the Portuguese and European populations. Circles highlight the risk genotype. The red circles show the risk genotypes from variants with a statistically significant difference between these two populations (p -value = $2.5e-7$ for *DHCR7* RS12785878, and p -value = $2.8e-4$ for *AMDHD1* RS10745742).

or e801 automated analyzer (Roche Diagnostics GmbH). This assay uses a vitamin D-binding protein as capture protein, which binds to both forms of vitamin D: 25(OH) D2 and 25(OH) D3. Thus, in this report, the terms ‘vitamin D’ and ‘25(OH) D’ refer to both forms: 25(OH) D2 and 25(OH) D3. The specific terms for vitamin D2 and D3 are used to refer to the corresponding individual form. The 25(OH) D cut-offs established by the Clinical Guidelines Subcommittee of the Endocrine Society were followed, considering the absence of a global consensus on the 25(OH) D concentration that defines vitamin D deficiency and “adequate” 25(OH) D levels for extra-skeletal functions²³. Thus, vitamin D adequacy was classified according to the following 25(OH) D cut-off levels:

deficiency, <20 ng/mL; insufficiency, 20–29 ng/mL; and sufficiency, ≥ 30 ng/mL²⁴. The season of blood sample collection was also considered, as it may affect the patient's vitamin D concentrations²⁵.

Genotyping. All patients signed an informed consent to perform a genetic test for the analysis of 18 genetic variants that play a role in vitamin D metabolism, transport, degradation and downstream pathways. See supplemental data for details about the genetic panel, Table ii. The genetic test was performed using the iPLEX® MassARRAY® system at the HeartGenetics' certified laboratory in Portugal. DNA samples were prepared from blood samples by HeartGenetics, CCUL and São João hospital laboratories.

Data management. Clinical history, genotypic and phenotypic data, stored at the e-CRF, were collected and managed using REDCap electronic data capture tools²⁶ hosted by BioData.pt (<https://biodata.pt/>), the Portuguese distributed infrastructure for biological data, at INESC-ID research institute. All datasets are pseudo-anonymous and only one of them has a key that connects to the patient. In accordance with the GDPR, only the PI of the project (Prof. Fausto Pinto) and the clinicians responsible for data acquisition have access to this key. All project participants have controlled data access. The clinical history and the patient informed consent were supervised by the Cardiology Service at Santa Maria hospital and by the Pathology Service at São João hospital (see supplemental data for e-CRF content details).

Data analysis. Clinical history, genotypic and phenotypic data was evaluated using statistical, machine learning and polygenic risk scores methodologies. Vitamin D polymorphisms prevalence in different populations was obtained from 1000 Genomes database (<https://www.internationalgenome.org/>) and from HeartGenetics's research database for the Portuguese population, with more than 8,000 Portuguese individuals.

Regarding the methodological approach, the following steps were undertaken:

1. **Data cleaning and validation:** All variables were analysed for outliers and missing values. Some discrepancies, such as different units of measure and data entry errors, were identified and fixed. No imputation was made. Regarding data transformation, both disease severity and vitamin D levels were categorized in different levels, and the risk genotypes were aggregated in PRSs.
2. **Descriptive analysis:** A complete, graphical descriptive analysis of the data was created for all variables of interest as univariate analysis. Data are presented as numbers or percentages for categorical variables, while continuous variables are shown as mean and standard deviation, and median and interquartile range (25th percentile–75th percentile).
3. **Analysis of data distribution:** The data normality was accessed using Shapiro–Wilk test and D'Agostino Pearson's test. Statistical normality testing is relevant in order to set up the category of statistical methods (parametric or non-parametric) used in further analysis. Results showed that most parameters do not follow a normal distribution, thus for further analysis it was considered only non-parametric statistical tests that do not assume any particular data distribution.
4. **Identification of the vitamin D polymorphisms as risk biomarkers:** Several statistical tests were used, namely Mann–Whitney and Kruskal–Wallis Tests. Spearman rank correlation coefficient was also calculated. Four PRSs have been computed, focused on the vitamin D metabolism, transport and degradation pathways, based on an additive weighted model, having values in the interval [0, 1]. In this interval, 0 corresponds to a lower risk of having low vitamin D levels due to genetics, and 1 corresponds to a higher risk of having low vitamin D levels due to genetics (see supplemental material for details about the PRSs). The four different scores considered the following genetic variants.

$$(1) \text{ Synthesis score} = DHCR7 \text{ RS12785878} + CYP2R1 \text{ RS10741657}$$

$$(2) \text{ Metabolism score} = GC \text{ RS 2,282,679} + CYP24A1 \text{ RS17216707}$$

$$(3) \text{ Pathway score} = DHCR7 \text{ RS12785878} + CYP2R1 \text{ RS10741657} + GC \text{ RS 2,282,679} + CYP24A1 \text{ RS17216707}$$

$$(4) \text{ Vitamin D total score} = DHCR7 \text{ RS12785878} + CYP2R1 \text{ RS10741657} + GC \text{ RS 2,282,679} + CYP24A1 \text{ RS17216707} + AMDHD1 \text{ RS10745742} + SEC23A \text{ RS8018720}$$

The PRSs did not model other genetic variants that have been tested, since their impact has not been obtained by the same GWAS studies, which could introduce a bias in its relative impact.

5. **Analysis of the correlation between hypovitaminosis D and the disease severity:** Different statistical tests were employed, namely Mann–Whitney and Kruskal–Wallis Tests, depending on the type of categorization under analysis. Spearman rank correlation coefficient was also calculated in order to analyse not only an eventual association but also to quantify it and observe its direction.
6. **Genotypes frequency comparison:** For this comparison the 1000 Genomes (<https://www.ensembl.org/index.html>) and the HeartGenetics's research database with more than 8000 Portuguese individuals were used.

Concerning the different statistical tests performed, a p -value < 0.05 was considered statistically significant.

Study approval. This project was approved by the Ethics Committee of Hospital de Santa Maria in Lisbon and Hospital de São João in Oporto. The study was conducted in accordance with the ethical principles of the Dec-

laration of Helsinki and followed the Good Clinical Practice guidelines. Written informed consent was obtained from all study participants prior to their inclusion in the study.

Received: 30 May 2021; Accepted: 23 September 2021

Published online: 21 October 2021

References

- Coperchini, F. *et al.* The cytokine storm in COVID-19: An overview of the involvement of the chemokine/chemokine-receptor system. *Cytokine Growth Factor Rev.* **53**, 25–32. <https://doi.org/10.1016/j.cytogfr.2020.05.003> (2020).
- Daneshkhan, A. *et al.* Evidence for possible association of vitamin D status with cytokine storm and unregulated inflammation in COVID-19 patients. *Aging Clin. Exp. Res.* **32**(10), 2141–2158. <https://doi.org/10.1007/s40520-020-01677-y> (2020).
- Aranow, C. Vitamin D and the immune system. *J. Invest. Med.* **59**(6), 881–886. <https://doi.org/10.2310/JIM.0b013e31821b8755> (2011).
- Slominski, R. *et al.* COVID-19 and Vitamin D: A lesson from the skin. *Exp. Dermatol.* **29**(9), 885–890. <https://doi.org/10.1111/exd.14170> (2020).
- Afzal, S. *et al.* Vitamin D, hypertension, and ischemic stroke in 116 655 individuals from the general population: A genetic study. *Hypertens* **70**(3), 499–507. <https://doi.org/10.1161/HYPERTENSIONAHA.117.09411> (2017).
- Liu, H. *et al.* Comorbid chronic diseases are strongly correlated with disease severity among COVID-19 patients: A systematic review and meta-analysis. *Aging Dis.* **11**(3), 668. <https://doi.org/10.14336/AD.2020.0502> (2020).
- Lips, P. *et al.* 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.* **180**(4), 23–54. <https://doi.org/10.1530/EJE-18-0736> (2019).
- Jiang, X. *et al.* Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. *Nat. Commun.* **9**, 260. <https://doi.org/10.1038/s41467-017-02662-2> (2018).
- Manousaki, D. *et al.* Genome-wide association study for vitamin D levels reveals 69 independent Loci. *Am. J. Hum. Genet.* **106**(3), 327–337. <https://doi.org/10.1016/j.ajhg.2020.01.017> (2020).
- Yao, P. *et al.* Effects of genetic and nongenetic factors on total and bioavailable 25(OH)D responses to vitamin D supplementation. *J. Clin. Endocrinol. Metab.* **102**(1), 100–110. <https://doi.org/10.1210/jc.2016-2930> (2017).
- Rhodes, J. M. *et al.* Editorial: Low population mortality from COVID-19 in countries south of latitude 35 degrees North supports vitamin D as a factor determining severity. *Aliment. Pharmacol. Ther.* **52**(12), 1434–1437. <https://doi.org/10.1111/apt.15777> (2020).
- Duarte, C. *et al.* Prevalence of vitamin D deficiency and its predictors in the Portuguese population: A nationwide population-based study. *Arch Osteoporos.* **15**, 36. <https://doi.org/10.1007/s11657-020-0695-x> (2020).
- WHO Working Group on the Clinical Characterisation and Management of COVID-19 infection. A minimal common outcome measure set for COVID-19 clinical research. *Lancet Infect. Dis.* **20**(8) e192–e197 [https://doi.org/10.1016/S1473-3099\(20\)30483-7](https://doi.org/10.1016/S1473-3099(20)30483-7) (2020).
- Wood, A. M. *et al.* Vitamin D-binding protein contributes to COPD by activation of alveolar macrophages. *Thorax* **66**(3), 205–210. <https://doi.org/10.1136/thx.2010.140921> (2011).
- Chun, R. F. New perspectives on the vitamin D binding protein. *Cell Biochem. Funct.* **30**(6), 445–456. <https://doi.org/10.1002/cbf.2835> (2012).
- Delanghe, J. R. *et al.* Behind the scenes of vitamin D binding protein: More than vitamin D binding. *Best Pract. Res. Clin. Endocrinol. Metab.* **29**(5), 773–786. <https://doi.org/10.1016/j.beem.2015.06.006> (2015).
- Kew, R. R. The vitamin D binding protein and inflammatory injury: A mediator or sentinel of tissue damage?. *Front. Endocrinol.* **10**, 470. <https://doi.org/10.3389/fendo.2019.00470> (2019).
- Xu, Y. *et al.* The importance of vitamin D metabolism as a potential prophylactic, immunoregulatory and neuroprotective treatment for COVID-19. *J. Transl. Med.* **18**, 322. <https://doi.org/10.1186/s12967-020-02488-5> (2020).
- Verdoia, M. *et al.* Potential role of hypovitaminosis D and vitamin D supplementation during COVID-19 pandemic. *QJM* **114**(1), 3–10. <https://doi.org/10.1093/qjmed/hcaa234> (2021).
- Cozier, Y. C. *et al.* Lower serum 25(OH)D levels associated with higher risk of COVID-19 infection in U.S. Black women. *PLoS ONE* **16**(7), e0255132. <https://doi.org/10.1371/journal.pone.0255132> (2021).
- Hf, H. *et al.* The interplay between vitamin D and COVID-19: protective or bystander?. *Eur. Ver. Med. Pharmacol. Sci.* **25**(4), 2131–2145. https://doi.org/10.26355/eurrev_202102_25119 (2021).
- Qayyum, S. *et al.* Vitamin D and lumisterol novel metabolites can inhibit SARS-CoV-2 replication machinery enzymes. *Am. J. Physiol. Endocrinol. Metab.* **321**(2), E246–E251. <https://doi.org/10.1152/ajpendo.00174.2021> (2021).
- Holick, M. F. The vitamin D deficiency pandemic: Approaches for diagnosis, treatment and prevention. *Rev. Endocr. Metab. Disord.* **18**(2), 153–165. <https://doi.org/10.1007/s11154-017-9424-1> (2017).
- Holick, M. F. *et al.* Evaluation, treatment, and prevention of vitamin D deficiency: An Endocrine Society clinical practice guideline. *J. Clin. Endocrinol. Metab.* **96**(7), 1911–1930. <https://doi.org/10.1210/jc.2011-0385> (2011).
- Stagi, S. *et al.* Vitamin D levels and effects of vitamin D replacement in children with periodic fever, aphthous stomatitis, pharyngitis, and cervical adenitis (PFAPA) syndrome. *Int. J. Pediatr. Otorhinolaryngol.* **78**(6), 964–968. <https://doi.org/10.1016/j.ijporl.2014.03.026> (2014).
- Harris, P. A. *et al.* The REDCap consortium: Building an international community of software platform partners. *J. Biomed. Inform.* **95**, 103–208. <https://doi.org/10.1016/j.jbi.2019.103208> (2019).

Acknowledgements

We thank all the study participants, who donated blood and authorized the genetic analysis and the collection of the clinical and phenotypic personal data. This project was supported by the “Fundação para a Ciência e Tecnologia”, program “Research 4 Covid-19 Apoio especial a projetos de implementação rápida para soluções inovadoras de resposta à pandemia de COVID-19”. It was also partially supported by each institution.

Author contributions

A.T.F. conceived and supervised the study, analysed the results, and wrote the manuscript with input from all the authors. M.B., S.B., S.B., D.C., S.C.R.S., M.F., B.N.G., F.M., J.M.C., A.M., A.T.P., H.P., F.R., Y.V. and F.P. were responsible for sample and data collection and interpretation from 370 patients of Hospital Santa Maria. B.A., F.B., L.L., J.P., A.R. and J.T.G. were responsible for sample and data collection and interpretation from 120 patients

of Hospital São João. H.V., L.G. and C.P. were responsible for genetic testing and sample logistics and performed the patients genotyping. G.A. supervised the study and the bioinformatics team. H.L. and P.K. performed the bioinformatics analyses. D.L. designed the genetic testing panel and analysed the results. D.F., J.P.R.R. and A.M.P.M. developed the e-CRM and performed data cleaning. A.F., J.C.R. contributed to the interpretation of the analysis. A.M.-R. contributed to statistical analysis, and data interpretation. C.C., J.T.G. and F.P. conceived and supervised the study, and analysed the results. All authors revised and approved the manuscript.

Competing interests

All authors affiliated with the company HeartGenetics, Genetics and Biotechnology SA, declare that the company developed a genetic test, the MyVitDGenes®, that was used to evaluate all the polymorphisms under analysis in this work. All other authors have declared that no conflict of interest exists.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-99952-z>.

Correspondence and requests for materials should be addressed to F.J.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021