







Communication

Rare Variants in Primary Immunodeficiency Genes and Their Functional Partners in Severe COVID-19

Maryam B. Khadzhieva ^{1,2,3}, Dmitry S. Kolobkov ² , Darya A. Kashatnikova ² , Alesya S. Gracheva ^{1,4} , Ivan V. Redkin ⁵ , Artem N. Kuzovlev ¹  and Lyubov E. Salnikova ^{1,2,3,*} 

- ¹ The Laboratory of Clinical Pathophysiology of Critical Conditions, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, 107031 Moscow, Russia; m.had@mail.ru (M.B.K.); palesa@yandex.ru (A.S.G.); artem_kuzovlev@mail.ru (A.N.K.)
- ² The Laboratory of Ecological Genetics, Vavilov Institute of General Genetics, Russian Academy of Sciences, 119991 Moscow, Russia; dmitry.s.kolobkov@gmail.com (D.S.K.); daria_sv11@mail.ru (D.A.K.)
- ³ The Laboratory of Molecular Immunology, National Research Center of Pediatric Hematology, Oncology and Immunology, 117997 Moscow, Russia
- ⁴ The Department of Population Genetics, Vavilov Institute of General Genetics, Russian Academy of Sciences, 119991 Moscow, Russia
- ⁵ Competence Center for the Development of AI Technology, Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology, 107031 Moscow, Russia; iredkin@fncrr.ru
- * Correspondence: salnikovalyubov@gmail.com

Abstract: The development of severe COVID-19, which is a complex multisystem disease, is thought to be associated with many genes whose action is modulated by numerous environmental and genetic factors. In this study, we focused on the ideas of the omnigenic model of heritability of complex traits, which assumes that a small number of core genes and a large pool of peripheral genes expressed in disease-relevant tissues contribute to the genetics of complex traits through interconnected networks. We hypothesized that primary immunodeficiency disease (PID) genes may be considered as core genes in severe COVID-19, and their functional partners (FPs) from protein–protein interaction networks may be considered as peripheral near-core genes. We used whole-exome sequencing data from patients aged ≤ 45 years with severe ($n = 9$) and non-severe COVID-19 ($n = 11$), and assessed the cumulative contribution of rare high-impact variants to disease severity. In patients with severe COVID-19, an excess of rare high-impact variants was observed at the whole-exome level, but maximal association signals were detected for PID + FP gene subsets among the genes intolerant to LoF variants, haploinsufficient and essential. Our exploratory study may serve as a model for new directions in the research of host genetics in severe COVID-19.

Keywords: severe COVID-19; primary immunodeficiency (PID) genes; protein-protein interaction (PPI); functional partners (FPs) of PID genes; omnigenic model; core and peripheral genes; whole-exome sequencing; rare high-impact (HI) variants



Citation: Khadzhieva, M.B.; Kolobkov, D.S.; Kashatnikova, D.A.; Gracheva, A.S.; Redkin, I.V.; Kuzovlev, A.N.; Salnikova, L.E. Rare Variants in Primary Immunodeficiency Genes and Their Functional Partners in Severe COVID-19. *Biomolecules* **2023**, *13*, 1380. <https://doi.org/10.3390/biom13091380>

Academic Editor: Domenico Lio

Received: 25 July 2023

Revised: 4 September 2023

Accepted: 6 September 2023

Published: 12 September 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

COVID-19, caused by the SARS-CoV-2 virus, remains a global health problem. COVID-19 has a variety of clinical manifestations, ranging from asymptomatic infection to fatal respiratory or multiorgan failure. The main risk factors for the severe course of COVID-19 are older age [1], male gender [2], the presence of specific comorbidities and overall multimorbidity [3,4], race/ethnicity and social position [5]. An increased risk of developing severe COVID-19 outcomes is also associated with unhealthy lifestyle factors and an unfavorable socioeconomic status. Low socioeconomic status is often the cause of a lack of timely, quality medical care and unhealthy behaviors. The latter include an unbalanced high-calorie diet, reduced physical activity, increased alcohol and tobacco use, and sleep disturbances. All of these behaviors can weaken a person's immune status [6,7]. The severity of the course of COVID-19 is also influenced by genetically mediated immune

system functionality. Genetic and non-genetic factors can interact to orchestrate a complex cascade of immune signals that lead to the recovery or failure of critical organ systems and death [8–12].

Rare disease findings can help identify causative genes and mechanisms that explain the predisposition and course of some common diseases, and in the case of COVID-19, such rare genetic diseases may be primary immunodeficiencies (PIDs). A recent review including data on 459 PID patients with COVID-19 showed higher mortality, hospitalization rates, and a higher frequency of oxygen supplementation use in PID patients than in the general population [13], but not all PID patients have a severe disease course [14,15]. This is linked to other risk factors, the same as in the general population [15], as well as the severity of the causative variant [16] and the overall genetic background. Based on ideas from the omnigenic model of the heritability of complex traits, we recently showed that the cumulative effect of rare high-impact (HI) genetic variations across the exome is associated with the severity of COVID-19 [17]. The omnigenic hypothesis suggests that genes with regulatory variants in at least one disease-associated tissue may influence the overall risk of disease development [18]. Due to the interconnectedness of gene regulatory networks, association signals from so-called peripheral genes are transmitted to core genes, which have a direct effect on the phenotype. In our previous study, the total contribution of rare potentially pathogenic variants to the phenotype of severe COVID-19 was greater for PID genes than for other analyzed groups of genes potentially important in the context of severe infection development. We hypothesized that PID genes are enriched for core genes for a phenotype defined as severe COVID-19 [17].

Phenotype determination is one of the “open” problems in the genetics of complex human traits because there is often a lot of noise in the phenotyping process (mostly regarding binary phenotypes) [19]. This problem becomes even more relevant when COVID-19 comparison groups are formed according to disease severity, as disease severity is a continuum [20] and it is not always possible to distinguish between moderate and severe forms of disease. Because of the important role of the patient’s immune responses in the pathophysiology of COVID-19, various immunologic tests have been proposed to identify patients at high risk of progression or death caused by COVID-19 [21,22]. One promising immunological test is to measure the levels of T-cell receptor excision circles (TRECs) and kappa-removal recombination excision circles (KRECs), which are non-replicating DNA fragments that are formed during the maturation of T- and B-lymphocytes, respectively, and are retained in cells. Therefore, TRECs and KRECs are considered proxy markers of the emergence of new T- and B-lymphocytes, and their age-adjusted levels can be used to assess the efficacy or dysfunction of the respective branches of cellular immunity. The interest in evaluating TREC and KREC levels in severe COVID-19 is bidirectional. On the one hand, given the demonstrated association between TREC and, to a lesser extent, KREC levels and the development of critical conditions and adverse outcomes in patients with COVID-19 [23–26] and non-COVID-19 pneumonia [27], TREC/KREC measurements may be useful in order to confirm the classification of COVID-19 severity. On the other hand, given the importance of selecting immunologic tests with high prognostic value, the additional evaluation of TREC/KREC levels in severe COVID-19 may contribute to the selection and subsequent clinical implementation of prognostic tools for patients with COVID-19.

In the present study, we aimed to further address the hypothesis of a possible role of PID genes as potential core genes for severe COVID-19. According to the omnigenic model, the influence of a peripheral gene on the phenotype is realized through a peripheral network [28]. It can be assumed that the closer a peripheral gene is to the core gene in the tissue regulatory network, the more influence on the phenotype it can exert through the core gene. Under this assumption, genes whose products participate in networks of protein–protein interaction (PPI) with PID gene products may contribute to PID gene association signals, with the magnitude of the effect depending on the confidence of the interaction. Thus, our first objective was to create and analyze sets of genes encoding products involved

in interactions with PID gene products with varying degrees of confidence to then assess the cumulative effect of rare HI variants in these genes in severe COVID-19. Since severe COVID-19 is associated with older age and age-related diseases in the general population, for this study, to reduce the possible influence of confounders, young adults (20 individuals) were selected from our previous sample [17] for whom TREC/KREC analysis was performed. Therefore, our second objective was to investigate the levels and prognostic efficacy of TREC/KREC counts in patients with different severities of COVID-19, and to compare TREC and KREC levels with the number of rare potentially pathogenic variants in genes controlling PPI with PID genes.

2. Materials and Methods

2.1. Study Design

Our study comprised two phases: theoretical and experimental. In the theoretical phase, we compiled and analyzed lists of genes associated with primary immunodeficiencies (PIDs) and their functional partners from PPI networks (hereinafter referred to as FPs). In the experimental phase, we used the results of the whole-exome sequencing analysis in COVID-19 patients with different disease severities to analyze the distribution of rare HI and potentially pathogenic missense variants in the gene sets generated in the previous theoretical phase of the study and based on the findings of this previous phase.

2.2. Theoretical Phase: Gene List Construction

The list of primary immunodeficiency genes (PIDs) was taken from the International Union of Immunological Societies (IUIS) Committee of Experts 2022 Updated Classification [29]. The phenotypes and associated genes are presented in this paper in 10 tables, of which Tables I through IX include the phenotype categories associated with germline variants, and Table X includes “Phenocopies.” This latter category was excluded from our analysis. The database search tool (STRING: Search Tool for the Retrieval of Interacting Genes) version 11.5 [30] was used to generate lists of FPs of PID genes. We generated three gene sets: (I) consisting of PID genes, and (II and III) consisting of FPs of PID genes. Sets II and III were generated based on STRING combined scores, which provide an assessment of STRING’s confidence (measured from zero to one) that the putative association between proteins is biologically significant, given all the contributing evidence from different sources. We hypothesized that the magnitude of the effect of FPs would depend on the confidence of their interaction with PID genes, with the greater the confidence, the greater the effect. In set II, FPs had a combined interaction score with PID genes ≥ 0.9 , and in set III, the combined interaction score with PID genes ranged from 0.4 to 0.89. STRING was also used for subsequent enrichment analysis. We next compared these three sets of genes for the number of genes that may be biologically important in relation to the development and course of acute infection. As biologically important, we considered the following gene groups: haploinsufficient [31], essential for life [32], intolerant to loss-of-function variants and intolerant to missense variants (https://storage.googleapis.com/gnomadpublic/release/2.1.1/constraint/gnomad.v2.1.1.lof_metrics.by_gene.txt.bgz (accessed on 21 June 2022)), linked to SARS-CoV-2 infection and/or COVID-19 disease from the GENCODE project (https://www.genecodegenes.org/human/covid19_genes.html# (accessed on 12 May 2023)), and immune tissue-specific [33]. Other tissue-specific or non-specific genes were included in the analysis for the comparison with immune tissue-specific genes.

Haploinsufficiency is a prediction of sensitivity to a reduced dose of a gene. DECIPHER presents haploinsufficiency scores that are based on the predicted probability of haploinsufficiency. Scores in the 0–10% range indicate a higher probability that the gene is haploinsufficient (<https://www.deciphergenomics.org/about/downloads/data> (accessed on 15 September 2022)). Genes intolerant to a loss of function, i.e., HI variants (pLI > 0.9) and missense variants (missense Z score > 3.09), were established based on constraint metrics from gnomAD v2.1 (<https://storage.googleapis.com/gnomadpublic/release/2.1.1>

[1/constraint/gnomad.v2.1.1.lof_metrics.by_gene.txt.bgz](#) (accessed on 21 June 2022)). The GENCODE project is currently re-annotating genes encoding human proteins associated with SARS-CoV-2 infection and/or COVID-19 disease. From the list of genes under consideration, we have selected those marked as updated. Compiling a list of immune-tissue-specific genes has been described elsewhere [33]. Briefly, we used the TissueEnrich R package v. 1.10.1 [34] to select tissue-enriched, group-enriched and tissue-enhanced genes from the set of RNA HPA tissue gene data (<https://www.proteinatlas.org/about/download> (accessed on 19 September 2022)). Immune-system-related tissues included appendix, B-cells, bone marrow, dendritic cells, granulocytes, lymph node, monocytes, NK-cells, spleen, T-cells, and tonsil tissues [35].

2.3. Experimental Phase

2.3.1. Patients and Clinical Data

Twenty patients with severe ($n = 9$) or mild/moderate COVID-19 ($n = 11$) under 45 years of age from our cohort of 86 patients with whole-exome sequencing data were selected for the study [17]. Patients with COVID-19 were recruited from the M.F. Vladimirovsky Moscow Regional Scientific Clinical Institute, the Moscow Clinical Center for Infectious Diseases at Voronovsky, and the V.P. Demikhov City Clinical Hospital of the Moscow Health Department in 2020 during the period preceding the start of vaccination. In addition to age (no older than 45 years), a confirmed SARS-CoV-2 infection test and TREC and KREC data [24] were required for inclusion in the study. Since age-specific TREC and KREC levels differ in patients with immunodeficiency, those taking immunomodulators and opioids, and those with changes in their sex hormone levels [36–38], the exclusion criteria in our study were selected to exclude the possible influence of comorbidities, medications, and some health-related conditions (e.g., pregnancy) on the course of COVID-19 and TREC and KREC counts. The exclusion criteria were as follows: patients with incurable terminal illness, primary or acquired immunodeficiency, long-term corticosteroid use, pregnancy, alcoholism, drug addiction, and HIV/AIDS. Although extremely unfavorable socioeconomic conditions and lifestyle were not among the exclusion criteria, such patients were not encountered in our study. For information on clinical diagnosis, see [17].

The Ethics Committee of the Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology approved the study; all included patients or their legal representatives signed an informed consent form.

2.3.2. Sample Analysis

The TREC/KREC levels were assayed in whole blood samples as described previously [24,27]. In brief, DNA was isolated from 200 μ L of venous blood via isopropanol precipitation. The RT-qPCR reactions were performed in a final volume of 25 μ L of reaction mixture containing 200 ng of DNA, primers and probes, which were designed for the specific amplification of the δ TREC- ψ J α T-cell receptor, kappa-deleting joint and human albumin (reference gene). The PCR conditions were as follows: 7 min at 95 $^{\circ}$ C, followed by 45 cycles of 30 s at 93 $^{\circ}$ C and 1 min at 59 $^{\circ}$ C (CFX96, Bio-Rad, USA; manufactured in Singapore). Standard curves for the accurate quantification of TREC, KREC, and albumin were obtained by constructing a calibration curve from sequentially diluted genetic constructs containing the TREC/KREC junction region and the corresponding albumin gene region. TREC and KREC copies were calculated and expressed as copies per 100,000 nucleated cells using the following formula: $[\text{mean TREC(KREC) value}/(\text{mean albumin value}/2)] \times 100,000$.

DNA isolation and sequencing is described elsewhere [17]. Briefly, DNA was isolated from blood using the Qiagen DNA blood mini kit DNA. The Swift 2S[®] Turbo DNA Library Kit was used for fragmentation and barcoding. Enrichment was performed with the Twist HumanCoreExome (<https://www.twistbioscience.com/products/ngs/fixd-panels/human-core-exome> (accessed on 29 January 2021)). Sequencing was executed on an Illumina HiSeq X Ten platform with 150 bp paired-end reads. Reads were aligned to

the GRCh38 reference genome using BWA MEM [39]. Duplicate reads were marked and excluded using the MarkDuplicates program. Variant calling was carried out using the HaplotypeCaller program of the GATK package. Variants were required to pass GATK's standard variant quality score recalibration (VQSR) threshold along with additional filters [17], and to have at least 10× coverage.

2.3.3. Annotation of Variants

Variants were annotated using SnpSift [40], SnpEff [41], and FAVOR [42], as well as the population databases Genome Aggregation Database (GnomAD) [43], 1000G [44], and TopMed [45]. Variants assessed by Ensembl as having serious consequences for protein structure and function (acceptor splice variants, donor splice variants, stop-loss, frameshift, stop-loss, and start-loss) were classified as HI. To classify a missense variant as harmful, we used the rare exome variant ensemble learner (REVEL) tool with a recommended threshold > 0.5 [46]. Our analysis focused on rare variants with alternative allele frequency (AF) < 0.001 or no AF data (missing) in the GnomAD, 1000G, and TopMed population resources (hereinafter rare variants).

2.4. Data Analysis

Calculations were performed using the R software (version 3.4.1). In the theoretical phase of this work, we compared the proportions of biologically important genes in PID genes (set I) and their FPs with combined interaction scores with PID gene products ≥ 0.9 (set II) and 0.4 to 0.89 (set III). Because the gene sets overlapped, for statistical analysis, we excluded all genes of set I from set II and all genes of sets I + II from set III. Thus, as designed, we categorized genes according to their proximity to the phenotype in terms of causality. The analyzed sets did not contain common genes, and we used Pearson's chi-square test with Yates' correction for continuity. In the experimental phase, the calculation of the area under the receiver operating characteristic (ROC) curve (AUC) was performed using the approach of DeLong et al. [47]. The AUC was considered according to Metz [48]; AUC > 0.9 means that the diagnostic performance of the classifier is excellent. Given the sample size, we used the unadjusted two-sided Cochran–Mantel–Haenszel (CMH) test to analyze the sequencing data. The pooled analysis of rare variants was carried out with the dominant inheritance model. Graphs were plotted using <https://www.biinformatics.com.cn/en> (accessed on 23 June 2023), a free online platform for data analysis and visualization.

3. Results

3.1. Theoretical Phase: The Analysis of PID Genes and Their Functional Partners

The list of PID phenotypes and associated genes is constantly growing; for example, in 2004, the IUIS Committee update included 57 gene–phenotype pairs [49], and by 2022, this number had increased to 450 [29]. The IUIS reports categorize phenotypes into tables, which in turn consist of subtables with overlapping phenotypes. Most of the new genes, as exemplified by those established since 2004, encode proteins that are FPs of previously known proteins, so similar phenotypes in the subtables are associated with genes that form PPI networks (Supplementary Table S1). Although some other genes that will be identified as PID genes in the future can be expected to also encode proteins involved in PPI networks with previously known FPs, we are interested in finding differences between PID and FP genes today, when NGS technologies are already used, including for patients with PID [50,51].

The three sets of genes we analyzed in this study included (I) 450 PID genes [29], (II) 4580 FPs of PID genes with a combined interaction score of ≥ 0.9 , and (III) 6445 FPs of PID genes with a combined interaction score of 0.4 to 0.89 (Supplementary Table S2). Gene ontology and KEGG pathway enrichment analyses were performed for PID genes and their FPs representing similar phenotypes, i.e., belonging to the same IUIS subtables. Because the gene sets overlapped (shown in the legend panel in Figure 1), we considered sets I, I + II, and I + II + III (Figure 1A,B; Supplementary Tables S3 and S4). Although

PID genes accounted for a smaller proportion of the genes in sets II and III, the results of the enrichment analyses in these three sets were consistent with each other. These results reflected the predominance of different types of regulatory interactions (Figure 1C,D).

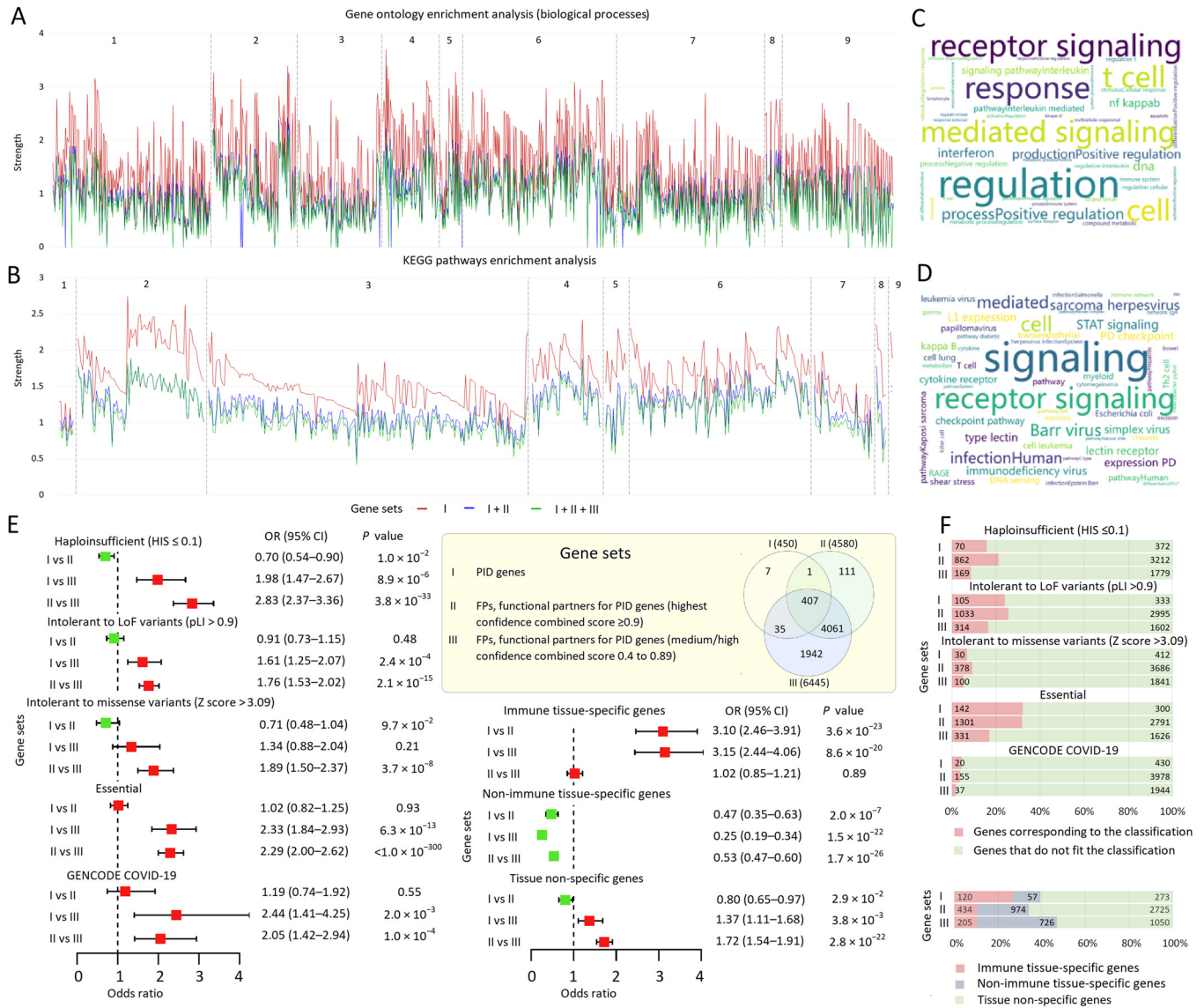


Figure 1. PID genes and their functional partners from the STRING database. (A) STRING Gene Ontology and (B) KEGG pathway enrichment analysis for sets of genes: (I) PID genes, (II and III) PID genes and their functional partners (FPs) with a combined interaction score ≥ 0.9 and 0.4 to 0.89, respectively. The legend panel for gene sets includes a built-in Venn diagram showing the overlap between these sets and the total number of genes in the sets (indicated in parentheses near each element of the Venn diagram). The analysis was performed for genes (and their partners) representing similar phenotypes, i.e., belonging to the same subtables from the IUIS tables. A list of subtables is provided in Supplementary Table S1. For sets II and III, only the terms encountered for the PID genes are indicated. Dotted lines separate enriched terms for genes (and their partners) from each of the IUIS categories, which are signed using Arabic numerals above the line graphs. See Supplementary Tables S3 and S4 for more information. (C,D) Wordclouds displaying sets of terms, related to enriched biological processes (C) and metabolic pathways (D); the higher the frequency, the larger the word in the wordcloud. (E,F) A comparison of the proportion of biologically important and tissue-specific genes in gene sets under consideration. Because the gene sets overlapped, we excluded PID genes

from set II and PID and set II genes from set 3. (E) Odds ratios and horizontal bars denoting 95% confidence intervals are shown. The experiment-wise p -value threshold corresponds to 0.0021 to account for multiple testing (0.05/24 comparisons). (F) The number of genes from the considered sets is indicated using a 100% stacked bar chart.

We also compared the proportion of biologically important genes (see Section 2.2) in the three sets under consideration and found that sets I and II did not differ from each other, but differed from set III in the proportion of genes that are haploinsufficient, intolerant to LoF and missense variants, essential, and annotated by GENCODE-COVID-19 (Figure 1E,F; Supplementary Table S2). The distribution of immune tissue-specific genes did not correspond to this pattern (Figure 1E,F; Supplementary Table S2). The proportion of immune tissue-specific genes was higher in set I than in sets II and III. Set I was depleted in non-immune tissue-specific genes compared to sets II and III. The highest proportion of tissue non-specific genes was reported in set II compared to set III, while the other results were non-significant. Thus, pronounced differences in tissue specificity were observed for PID genes and their FPs.

As follows from the theoretical phase of this work, among the FP genes there is a large share of genes involved in the same biological processes and metabolic pathways as the PID genes themselves. In addition, among the FPs (mostly in set II), as well as among the PID genes themselves, there is a high proportion of biologically important genes whose variability may influence the resistance of the organism in response to acute infection. This hypothesis was tested in the experimental phase of this work.

3.2. Experimental Phase; Whole-Exome Sequencing of Twenty Patients with COVID-19

3.2.1. Demographic and Clinical Characteristics of Patients

Twenty unrelated patients with COVID-19 aged ≤ 45 years were included in the present study. Key demographic and clinical data are presented in Supplementary Table S5. The severe COVID-19 subgroup included nine patients with a severe or extremely severe course (mean age \pm SD, 38.33 ± 5.72), and the non-severe COVID-19 subgroup included eleven patients with mild or moderate COVID-19 (35.18 ± 5.91). The subgroups did not differ in demographic characteristics and previous diseases.

3.2.2. TREC and KREC Levels in Severe and Non-Severe COVID-19

A strong difference was found in the number of TRECs (copies/ 10^5 cells) in patients with severe (Median; Q1–Q3: 11.41; 3.74–23.50) versus non-severe (96.63; 71.42–192.41) COVID-19; Mann–Whitney U test p -value 0.00063 (Figure 2A, Supplementary Table S5). The TREC levels did not differ between men and women for severe COVID-19 and were higher in women than in men for non-severe COVID-19, although the results were not significant after adjustment for multiplicity. The KREC levels did not differ between men and women for either severe or non-severe COVID-19 (Supplementary Table S5). Other immunologic markers, such as the leukocyte, lymphocyte, neutrophil, and monocyte counts, as well as the neutrophil-to-lymphocyte ratio (NLR) and lymphocyte-to-monocyte ratio (LMR) at admission and at the last measurement before discharge/death, did not differ significantly between patients with different COVID-19 severities (Supplementary Table S5).

In the ROC analysis, the TREC assay demonstrated excellent diagnostic performance for severe COVID-19 (AUC 0.96, 95% CI 0.765 to 1.000, $p < 0.001$, Youden index J 0.818) (Figure 2B). The results for KRECs were non-significant (Figure 2A,B; Supplementary Table S5). According to the literature, SARS-CoV-2 infection affects both T and B cell generation, but this effect is less pronounced in the B cell compartment [26]; our sample size was apparently insufficient to detect effects at the level of KRECs. The pronounced effects found for TRECs may be related to the fact that SARS-CoV-2 can directly invade the thymus and alter the gene expression patterns of the thymic epithelium [26].

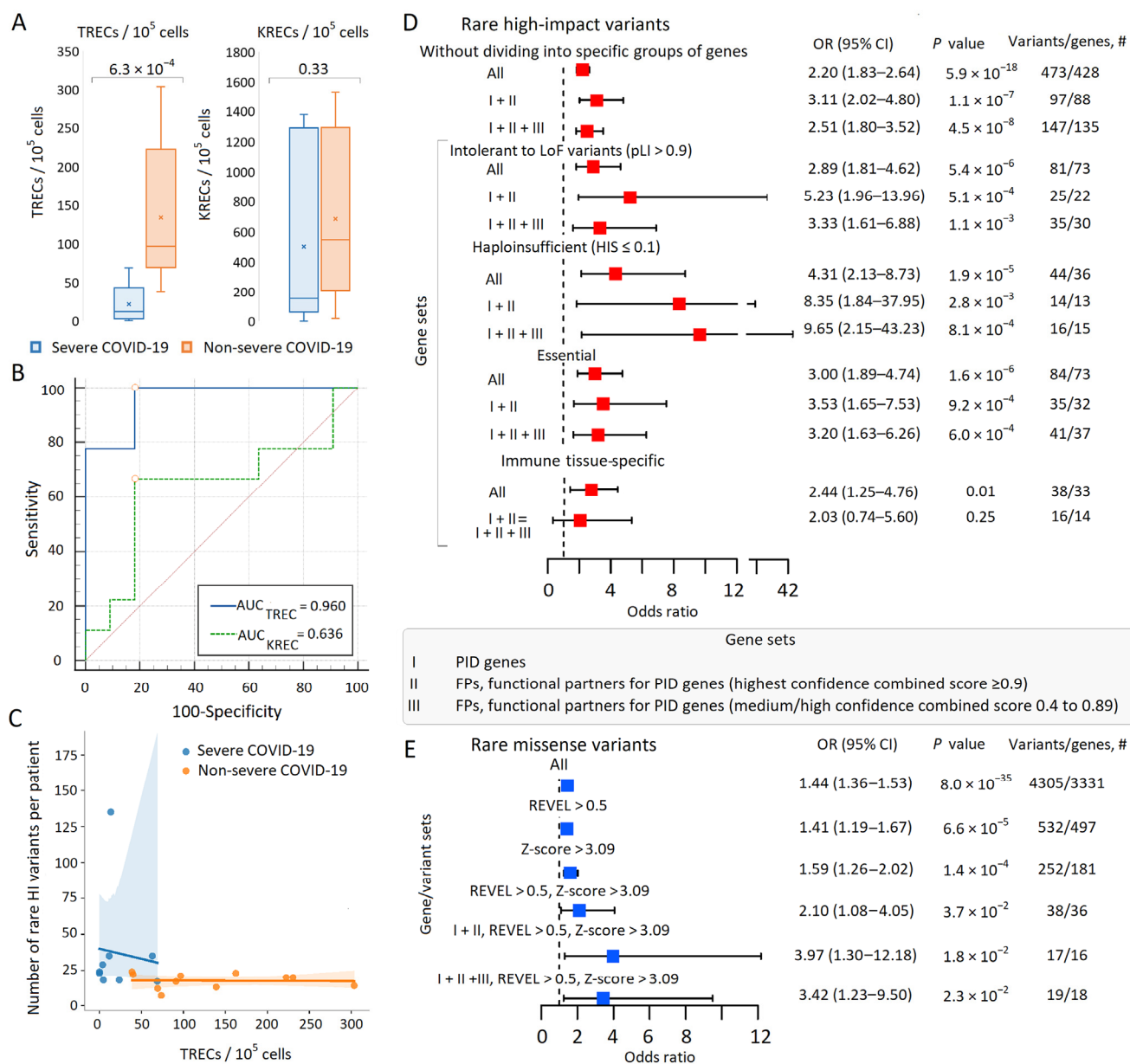


Figure 2. Immunological and genetic variability in patients with severe and non-severe COVID-19. (A) Boxplots depicting differences between TREC and KREC levels in patients with severe and non-severe COVID-19. (B) Receiver operating characteristic (ROC) curves for TREC and KREC levels in predicting the severity of COVID-19. (C) Multi-group scatter plots showing the inverse correlation between TREC levels and the number of rare HI variants in patients with severe and non-severe COVID-19. (D,E) Burden of rare HI variants (D) and missense variants (E) in gene sets under study. Odds ratios and horizontal bars indicating 95% confidence intervals are provided. The right column indicates the number of rare variants ((D): HI variants; (E): missense variants) and the number of genes with these variants in the entire sample of 20 patients. The *p*-value threshold corresponds to 0.0025 to account for multiple testing (0.05/20 comparisons). #: number.

Given the association demonstrated in the literature between TREC levels and the development of critical conditions and adverse outcomes in patients with COVID-19 and non-COVID-19 pneumonia [23–27], we believe that the results of ROC analysis support the accuracy of phenotype determination in severe and non-severe COVID-19.

3.2.3. Correlation between TREC and KREC Levels and the Number of Rare High-Impact Variants at the Whole-Exome Level

A significant inverse correlation was observed between the number of TRECs (copies/ 10^5 cells) and the number of rare HI variants in the study cohort (Figure 2C). The Spearman rank correlation coefficient r_s was -0.53 (two-tailed p -value was 0.017).

3.2.4. Rare Variant Burden in PID Genes and Their Functional Partners in Severe COVID-19

In the entire sample of 20 patients, the total number of rare HI variants in PID genes was small (eight variants in seven genes, Supplementary Table S6), so our primary analysis (without dividing into specific gene groups) involved comparing the burden of rare HI variants in all genes and in PID + FP genes (sets I + II and I + II + III) between patients with severe and non-severe COVID-19 (Figure 2D). In our sample, the burden of rare HI variants measured according to OR (95% CI) in severe versus non-severe COVID-19 decreased in the sets $I + II > I + II + III > \text{all genes}$, but the results were highly significant for all sets. We then performed a similar analysis for genes intolerant to LoF variants, haploinsufficient, essential, and immune tissue-specific. For genes intolerant to LoF variants and essential genes, the distribution of association signals was consistent with that obtained earlier for all genes ($I + II > I + II + III > \text{all genes}$); for haploinsufficient genes, the greatest effect was observed in the $I + II + III$ set. The association results for immune-tissue-specific genes results did not remain significant after adjusting for multiple comparisons, so the groups of non-immune-tissue-specific or tissue-non-specific genes were not considered (Figure 2D). GENCODE-COVID-19 genes were not considered because they included only six rare HI variants.

For missense variants, the effect had the same direction of association but was weaker than for HI variants, and in the smaller sets ($\text{REVEL} > 0.5$, $Z\text{-score} > 3.09$, and $\text{REVEL} > 0.5$, $Z\text{-score} > 3.09$ in the PID + FP gene sets) the results were non-significant when corrected for multiple comparisons (Figure 2E).

Summarizing these results, we can conclude that there was a burden of rare potentially pathogenic variants at the whole-exome level in patients with severe compared to non-severe COVID-19. The association signals were stronger when considering certain groups of genes, particularly genes intolerant to LoF variants, haploinsufficient and essential genes. The maximum effect sizes were observed in PID + FP gene subsets selected within these gene groups. Expanding the sample (from $I + II$ to $I + II + III$) led to a decrease in the variant burden in genes intolerant to LoF and essential genes. Notably, among the PID + FP subsets with the largest effect sizes, PID genes accounted for only 13.6% (3/22), 6.7% (1/15 genes), and 15.6% (5/32 genes) among the genes intolerant to LoF variants, haploinsufficient, and essential, respectively (Figure 2D, Supplementary Table S6).

4. Discussion

In this study, we turned to the omnigenic hypothesis to delineate the role of PID genes as core genes in severe COVID-19 and the FPs of PID genes at the level of PPI networks as near-core peripheral genes. In the theoretical phase, we compared PID genes (set I) and their FPs with combined interaction scores ≥ 0.9 and 0.4 to 0.89 (sets II and III, respectively) according to several characteristics. We found similarities between PID genes and FP genes with respect to their involvement in biological processes and metabolic pathways, and with respect to FPs from set II in the proportion of biologically important genes in the context of overall organismal resilience and resistance to various stressors. The main differences between PID genes and their FPs were related to tissue specificity. In the experimental phase of this work, we combined the TREC/KREC immunologic analysis in COVID-19 patients with the whole-exome sequencing results. Using a cohort of 20 patients under 45 years of age with different severities of COVID-19, we showed the excellent ability of the TREC count (copies/ 10^5 cells) to help diagnose severe COVID-19 (AUC = 0.96), as well as an inverse correlation between the number of rare HI variants and the number of TRECs.

In this small cohort, the results of the whole-exome sequencing analysis were consistent with those previously described [17], i.e., patients with severe COVID-19 had an excess of rare potentially pathogenic variants at the whole-exome level. The association effects were stronger when specific groups of genes were considered, particularly genes intolerant to LoF variants, haploinsufficient and essential genes, but not immune tissue-specific genes. In most cases, the effects decreased in the following series: set I + II > set I + II + III > all genes in the group or at the whole-exome level.

The main question linked to the omnigenic model and discussed in the literature is the identification of core and peripheral genes. While core genes are expected to have largely consistent effects, the effects of peripheral genes are mediated through interactions with environmental and genetic factors and thus have inter- and intrapopulation variability [28]. We extended this hypothesis by considering that the magnitude of the effect of peripheral genes may be influenced by their proximity to core genes in regulatory networks due to fewer interactions during associative signal transduction. The results are consistent with our assumption, as the effects of PID + FP genes, among which FP genes strongly dominated, were higher than the total effects of all genes considered, especially within groups of biologically important genes. The possibility of the gradual delineation of peripheral genes has already been mentioned in the literature [52], and our work illustrates an approach to such delineation and the isolation of a near-core “layer” of peripheral genes.

One of the main differences between core and peripheral genes in our study was the significantly lower number of tissue-specific genes among peripheral genes. It has been shown that disease-causing genes are often tissue-specific and, in a healthy state, are expressed at a higher level in those tissues that are affected in pathology [33,53,54]. In this context, one of the possible explanations for why biologically important peripheral genes from the near-core FPs remain peripheral is their low tissue specificity. FP genes are mainly genes that control various types of interactions; among them there are few genes specific to immune tissue, and they are enriched with non-specific genes that control the processes in various tissues.

We showed the utility of TREC analysis in differentiating patients at increased risk of severe COVID-19 and an inverse correlation between TREC level and the number of rare HI variants. This finding may simply reflect the correlation of disease severity with both variables, the number of TRECs and the number of genetic variants. But it is also known that an excess of rare variants in large specific gene groups is associated with the severity of the course, and sometimes, the age of manifestation of a number of psychiatric [55,56], neurologic [57,58], immune [59], and cardiovascular [60] diseases. Given that the immune system is involved to some degree in most pathologic phenotypes [61,62], their more frequent and severe manifestations may lead to earlier immunosenescence, which is reflected in TREC levels.

5. Conclusions

The main limitation of the study is the small sample size, which may be responsible for several types of bias. This applies not only to the genetic part of the study, but also to the interpretation of the results of the TREC/KREC analysis due to their high variability and dependence on many demographic, environmental and patient health factors. Therefore, a larger and more diverse cohort is needed to increase the statistical power of the results and to provide a more complete understanding of the observed patterns and their significance for clinical outcomes. Thus, we consider our study to be a preliminary one, which nevertheless has methodological and biological significance. In a sample of young patients with severe and non-severe COVID-19, we confirmed our previous findings regarding the cumulative effect of rare HI variants at the whole-exome level on individuals' susceptibility to severe COVID-19 [17]. COVID-19 is a complex multisystem disease associated with multiple genes, most of which are peripheral genes whose action is modulated by environmental, epigenetic and genetic factors [28,63]. Focusing on the ideas of the omnigenic model, which divides genes by their proximity in terms of causality to the phenotype, we

used PPI networks to identify the sets of FP genes that are closest to PID genes, which were treated as core genes in this study. We hypothesized that the effects of near-core FP genes might be more pronounced and more stable than those of other genes expressed in disease-relevant tissues. Consistent with this assumption, top association signals for the severe form of COVID-19 were obtained for the subsets of PID + FP genes within the groups of biologically important genes (intolerant to LoF variants, haploinsufficient and essential for life). In most cases, the largest effects of rare variants were obtained for the set of PID genes and their FPs with a combined interaction score ≥ 0.9 (set I + II), providing indirect evidence for the role of proximity to core genes in PPI networks. Our work may serve as a model for systematic studies and new directions in the study of host genetics in severe COVID-19.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom13091380/s1>, Table S1: Year-to-year trends in PPI network statistics for genes with overlapping phenotypes according to subcategories in the IUIS report (2022) [29]. Table S2: Characteristics of primary immunodeficiency (PID) genes and their functional partners (FPs) at the protein level. For each gene, the full STRING network was queried for interactors with the maximal available size cutoff of 50 and combined interaction score of ≥ 0.4 . The names of the categories and subcategories for PID genes are shown in Table S1. Table S3: STRING Gene Ontology enrichment analysis for primary immunodeficiency (PID) gene set products with overlapping phenotypes according to IUIS subcategories compared to similar data for PID gene functional partners (FPs). For the FP sets, only those biological processes that occurred for the PID sets are indicated. I—PID genes; II—FPs (combined interaction score ≥ 0.9); III—FPs (combined interaction score 0.4 to 0.89). Terms missing for set I are not shown, but the enrichment analysis results (strength and false discovery rate) are given according to the STRING enrichment analysis for the full set of associated terms. Table S4: STRING analysis of KEGG pathways for primary immunodeficiency (PID) gene set products with overlapping phenotypes according to IUIS subcategories compared to similar data for PID gene functional partners (FPs). For the FP sets, only those biological processes that occurred for the PID sets are indicated. I—PID genes; II—FPs (combined interaction score ≥ 0.9); III—FPs (combined interaction score 0.4 to 0.89). Terms missing for set I are not shown, but the enrichment analysis results (strength and false discovery rate) are given according to the STRING enrichment analysis for the full set of associated terms. Table S5: Demographic and clinical characteristics of the COVID-19 patients. * The exact two-tailed Fisher test unless otherwise indicated in parentheses. LR, the likelihood ratio χ^2 test; MWU, the Mann–Whitney U test. Significant results are highlighted in bold. (*p*-value threshold corresponds to 0.0015 to account for multiple testing; 0.05/33 comparisons). ** Median (Q1–Q3). Table S6: Gene annotations for the entire gene set with variants identified by whole exome sequencing in the 20-patient sample.

Author Contributions: L.E.S. conceived the idea. M.B.K., A.S.G. and I.V.R. recruited participants and collected data. M.B.K., D.S.K. and D.A.K. helped with data curation, analysis, and plotting of the data. L.E.S. wrote the manuscript. I.V.R. and A.N.K. edited the manuscript regarding clinical healthcare practice. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the state assignment of the Ministry of Education and Science of Russia (No. FGWS-2022-0001 “Multiple organ dysfunction and failure in critical conditions: leading mechanisms of development, new methods of diagnosis and treatment” and No. 0092-2022-0001 “Genetic processes in microorganisms, plants, animals, and humans”).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki. The study was approved by the Ethics Committee of the Federal Research and Clinical Center of Intensive Care Medicine and Rehabilitology (No. AE 2.1.18).

Informed Consent Statement: Informed consent was obtained from all patients or their legal representatives.

Data Availability Statement: All raw sequencing data have been submitted to the NCBI BioProject database (<https://www.ncbi.nlm.nih.gov/bioproject/> (accessed on 29 June 2023)) under accession number PRJNA947511.

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

References

1. Brodin, P. Immune determinants of COVID-19 disease presentation and severity. *Nat. Med.* **2021**, *27*, 28–33. [[CrossRef](#)]
2. Peckham, H.; de Gruijter, N.M.; Raine, C.; Radziszewska, A.; Ciurtin, C.; Wedderburn, L.R.; Rosser, E.C.; Webb, K.; Deakin, C.T. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nat. Commun.* **2020**, *11*, 6317. [[CrossRef](#)]
3. Clark, A.; Jit, M.; Warren-Gash, C.; Guthrie, B.; Wang, H.H.X.; Mercer, S.W.; Sanderson, C.; McKee, M.; Troeger, C.; Ong, K.L.; et al. Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: A modelling study. *Lancet Glob. Health* **2020**, *8*, e1003–e1017. [[CrossRef](#)] [[PubMed](#)]
4. Russell, C.D.; Lone, N.I.; Baillie, J.K. Comorbidities, multimorbidity and COVID-19. *Nat. Med.* **2023**, *29*, 334–343. [[CrossRef](#)]
5. Webb Hooper, M.; Nápoles, A.M.; Pérez-Stable, E.J. COVID-19 and Racial/Ethnic Disparities. *JAMA* **2020**, *323*, 2466–2467. [[CrossRef](#)] [[PubMed](#)]
6. Lange, K.W.; Nakamura, Y. Lifestyle factors in the prevention of COVID-19. *Glob. Health J.* **2020**, *4*, 146–152. [[CrossRef](#)] [[PubMed](#)]
7. Foster, H.M.E.; Ho, F.K.; Mair, F.S.; Jani, B.D.; Sattar, N.; Katikireddi, S.V.; Pell, J.P.; Niedzwiedz, C.L.; Hastie, C.E.; Anderson, J.J.; et al. The association between a lifestyle score, socioeconomic status, and COVID-19 outcomes within the UK Biobank cohort. *BMC Infect. Dis.* **2022**, *22*, 273. [[CrossRef](#)]
8. Severe COVID-19 GWAS Group; Ellinghaus, D.; Degenhardt, F.; Bujanda, L.; Buti, M.; Albillos, A.; Invernizzi, P.; Fernández, J.; Prati, D.; Baselli, G.; et al. Genomewide Association Study of Severe COVID-19 with Respiratory Failure. *N. Engl. J. Med.* **2020**, *383*, 1522–1534.
9. Zhang, Q.; Bastard, P.; Liu, Z.; Le Pen, J.; Moncada-Velez, M.; Chen, J.; Ogishi, M.; Sabli, I.K.D.; Hodeib, S.; Korol, C.; et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **2020**, *370*, eabd4570. [[CrossRef](#)]
10. Pairo-Castineira, E.; Clohisey, S.; Klaric, L.; Bretherick, A.D.; Rawlik, K.; Pasko, D.; Walker, S.; Parkinson, N.; Fourman, M.H.; Russell, C.D.; et al. Genetic mechanisms of critical illness in COVID-19. *Nature* **2021**, *591*, 92–98. [[CrossRef](#)]
11. COVID-19 Host Genetics Initiative. A first update on mapping the human genetic architecture of COVID-19. *Nature* **2022**, *608*, E1–E10. [[CrossRef](#)] [[PubMed](#)]
12. van der Poll, T.; Shankar-Hari, M.; Wiersinga, W.J. The immunology of sepsis. *Immunity* **2021**, *54*, 2450–2464. [[CrossRef](#)]
13. Drzymalla, E.; Green, R.F.; Knuth, M.; Khoury, M.J.; Dotson, W.D.; Gundlapalli, A. COVID-19-related health outcomes in people with primary immunodeficiency: A systematic review. *Clin. Immunol.* **2022**, *243*, 109097. [[CrossRef](#)] [[PubMed](#)]
14. Marcus, N.; Frizinsky, S.; Hagin, D.; Ovadia, A.; Hanna, S.; Farkash, M.; Maoz-Segal, R.; Agmon-Levin, N.; Broides, A.; Nahum, A.; et al. Minor Clinical Impact of COVID-19 Pandemic on Patients with Primary Immunodeficiency in Israel. *Front. Immunol.* **2021**, *11*, 614086. [[CrossRef](#)] [[PubMed](#)]
15. Meyts, I.; Buccioli, G.; Quinti, I.; Neven, B.; Fischer, A.; Seoane, E.; Lopez-Granados, E.; Gianelli, C.; Robles-Marhuenda, A.; Jeandel, P.Y.; et al. Coronavirus disease 2019 in patients with inborn errors of immunity: An international study. *J. Allergy Clin. Immunol.* **2021**, *147*, 520–531. [[CrossRef](#)]
16. Gruber, C.; Bogunovic, D. Incomplete penetrance in primary immunodeficiency: A skeleton in the closet. *Hum. Genet.* **2020**, *139*, 745–757. [[CrossRef](#)]
17. Khadzhiyeva, M.B.; Gracheva, A.S.; Belopolskaya, O.B.; Kolobkov, D.S.; Kashatnikova, D.A.; Redkin, I.V.; Kuzovlev, A.N.; Grechko, A.V.; Salnikova, L.E. COVID-19 severity: Does the genetic landscape of rare variants matter? *Front. Genet.* **2023**, *14*, 1152768. [[CrossRef](#)]
18. Boyle, E.A.; Li, Y.I.; Pritchard, J.K. An expanded view of complex traits: From polygenic to omnigenic. *Cell* **2017**, *169*, 1177–1186. [[CrossRef](#)]
19. Brandes, N.; Weissbrod, O.; Linial, M. Open problems in human trait genetics. *Genome Biol.* **2022**, *23*, 131. [[CrossRef](#)]
20. Guérin, P.J.; McLean, A.R.D.; Rahan, S.; Lawal, A.; Watson, J.A.; Strub-Wourgaft, N.; White, N.J. Definitions matter: Heterogeneity of COVID-19 disease severity criteria and incomplete reporting compromise meta-analysis. *PLoS Glob. Public Health* **2022**, *2*, e0000561. [[CrossRef](#)]
21. Mueller, Y.M.; Schrama, T.J.; Ruijten, R.; Schreurs, M.W.J.; Grashof, D.G.B.; van de Werken, H.J.G.; Lasinio, G.J.; Álvarez-Sierra, D.; Kiernan, C.H.; Castro Eiro, M.D.; et al. Stratification of hospitalized COVID-19 patients into clinical severity progression groups by immuno-phenotyping and machine learning. *Nat. Commun.* **2022**, *13*, 915. [[CrossRef](#)] [[PubMed](#)]
22. Sánchez-Montalvá, A.; Álvarez-Sierra, D.; Martínez-Gallo, M.; Perurena-Prieto, J.; Arrese-Muñoz, I.; Ruiz-Rodríguez, J.C.; Espinosa-Pereiro, J.; Bosch-Nicolau, P.; Martínez-Gómez, X.; Antón, A.; et al. Exposing and Overcoming Limitations of Clinical Laboratory Tests in COVID-19 by Adding Immunological Parameters; A Retrospective Cohort Analysis and Pilot Study. *Front. Immunol.* **2022**, *13*, 902837. [[CrossRef](#)] [[PubMed](#)]
23. Cuvelier, P.; Roux, H.; Couëdel-Courteille, A.; Dutrieux, J.; Naudin, C.; Charmeteau de Muylder, B.; Cheyrier, R.; Squara, P.; Marullo, S. Protective reactive thymus hyperplasia in COVID-19 acute respiratory distress syndrome. *Crit. Care* **2021**, *25*, 4. [[CrossRef](#)]
24. Khadzhiyeva, M.B.; Kalinina, E.V.; Larin, S.S.; Sviridova, D.A.; Gracheva, A.S.; Chursinova, J.V.; Stepanov, V.A.; Redkin, I.V.; Avdeikina, L.S.; Rumyantsev, A.G.; et al. TREC/KREC Levels in Young COVID-19 Patients. *Diagnostics* **2021**, *11*, 1486. [[CrossRef](#)] [[PubMed](#)]

25. Savchenko, A.A.; Tikhonova, E.; Kudryavtsev, I.; Kudlay, D.; Korsunsky, I.; Beleniuk, V.; Borisov, A. TREC/KREC Levels and T and B Lymphocyte Subpopulations in COVID-19 Patients at Different Stages of the Disease. *Viruses* **2022**, *14*, 646. [[CrossRef](#)] [[PubMed](#)]
26. Rosichini, M.; Bordoni, V.; Silvestris, D.A.; Mariotti, D.; Matusali, G.; Cardinale, A.; Zambruno, G.; Condorelli, A.G.; Flamini, S.; Genah, S.; et al. SARS-CoV-2 infection of thymus induces loss of function that correlates with disease severity. *J. Allergy Clin. Immunol.* **2023**, *151*, 911–921. [[CrossRef](#)]
27. Kashatnikova, D.A.; Khadzhieva, M.B.; Kolobkov, D.S.; Belopolskaya, O.B.; Smelaya, T.V.; Gracheva, A.S.; Kalinina, E.V.; Larin, S.S.; Kuzovlev, A.N.; Salnikova, L.E. Pneumonia and Related Conditions in Critically Ill Patients-Insights from Basic and Experimental Studies. *Int. J. Mol. Sci.* **2022**, *23*, 9896. [[CrossRef](#)]
28. Mathieson, I. The omnigenic model and polygenic prediction of complex traits. *Am. J. Hum. Genet.* **2021**, *108*, 1558–1563. [[CrossRef](#)]
29. Tangye, S.G.; Al-Herz, W.; Bousfiha, A.; Cunningham-Rundles, C.; Franco, J.L.; Holland, S.M.; Klein, C.; Morio, T.; Oksenhendler, E.; Picard, C.; et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. *J. Clin. Immunol.* **2022**, *42*, 1473–1507. [[CrossRef](#)]
30. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [[CrossRef](#)]
31. Huang, N.; Lee, I.; Marcotte, E.M.; Hurler, M.E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* **2010**, *6*, e1001154. [[CrossRef](#)] [[PubMed](#)]
32. Dickinson, M.E.; Flenniken, A.M.; Ji, X.; Teboul, L.; Wong, M.D.; White, J.K.; Meehan, T.F.; Wenginger, W.J.; Westerberg, H.; Adissu, H.; et al. High-throughput discovery of novel developmental phenotypes. *Nature* **2016**, *537*, 508–514. [[CrossRef](#)]
33. Kolobkov, D.S.; Sviridova, D.A.; Abilev, S.K.; Kuzovlev, A.N.; Salnikova, L.E. Genes and Diseases: Insights from Transcriptomics Studies. *Genes* **2022**, *13*, 1168. [[CrossRef](#)] [[PubMed](#)]
34. Jain, A.; Tuteja, G. TissueEnrich: Tissue-specific gene enrichment analysis. *Bioinformatics* **2019**, *35*, 1966–1967. [[CrossRef](#)]
35. Uhlen, M.; Karlsson, M.J.; Zhong, W.; Tebani, A.; Pou, C.; Mikes, J.; Lakshminanth, T.; Forsström, B.; Edfors, F.; Odeberg, J.; et al. A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science* **2019**, *366*, eaax9198. [[CrossRef](#)] [[PubMed](#)]
36. Olsen, N.J.; Kovacs, W.J. Evidence that androgens modulate human thymic T cell output. *J. Investig. Med.* **2011**, *59*, 32–35. [[CrossRef](#)] [[PubMed](#)]
37. Hellberg, S.; Mehta, R.B.; Forsberg, A.; Berg, G.; Brynhildsen, J.; Winqvist, O.; Jenmalm, M.C.; Ernerudh, J. Maintained thymic output of conventional and regulatory T cells during human pregnancy. *J. Allergy Clin. Immunol.* **2019**, *143*, 771–775.e7. [[CrossRef](#)]
38. Adatia, A.; Ling, L.; Chakraborty, P.; Brick, L.; Brager, R. Neonatal abstinence syndrome is a potential cause of low TREC copy number. *Allergy Asthma Clin. Immunol.* **2021**, *17*, 115. [[CrossRef](#)]
39. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
40. Cingolani, P.; Patel, V.M.; Coon, M.; Nguyen, T.; Land, S.J.; Ruden, D.M.; Lu, X. Using *Drosophila melanogaster* as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front. Genet.* **2012**, *3*, 35. [[CrossRef](#)]
41. 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](#)]
42. Zhou, H.; Arapoglou, T.; Li, X.; Li, Z.; Zheng, X.; Moore, J.; Asok, A.; Kumar, S.; Blue, E.E.; Buyske, S.; et al. FAVOR: Functional annotation of variants online resource and annotator for variation across the human genome. *Nucleic Acids Res.* **2023**, *51*, D1300–D1311. [[CrossRef](#)] [[PubMed](#)]
43. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)] [[PubMed](#)]
44. 1000 Genomes Project Consortium; Auton, A.; Brooks, L.D.; Durbin, R.M.; Garrison, E.P.; Kang, H.M.; Korbel, J.O.; Marchini, J.L.; McCarthy, S.; McVean, G.A.; et al. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74.
45. Taliun, D.; Harris, D.N.; Kessler, M.D.; Carlson, J.; Szpiech, Z.A.; Torres, R.; Taliun, S.A.G.; Corvelo, A.; Gogarten, S.M.; Kang, H.M.; et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **2021**, *590*, 290–299. [[CrossRef](#)]
46. Ioannidis, N.M.; Rothstein, J.H.; Pejaver, V.; Middha, S.; McDonnell, S.K.; Baheti, S.; Musolf, A.; Li, Q.; Holzinger, E.; Karyadi, D.; et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* **2016**, *99*, 877–885. [[CrossRef](#)]
47. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [[CrossRef](#)]
48. Metz, C.E. Basic principles of ROC analysis. *Semin. Nucl. Med.* **1978**, *8*, 283–298. [[CrossRef](#)]
49. Notarangelo, L.; Casanova, J.L.; Fischer, A.; Puck, J.; Rosen, F.; Seger, R.; Geha, R.; International Union of Immunological Societies Primary Immunodeficiency Diseases Classification Committee. Primary immunodeficiency diseases: An update. *J. Allergy Clin. Immunol.* **2004**, *114*, 677–687. [[CrossRef](#)]

50. Stray-Pedersen, A.; Sorte, H.S.; Samarakoon, P.; Gambin, T.; Chinn, I.K.; Coban Akdemir, Z.H.; Erichsen, H.C.; Forbes, L.R.; Gu, S.; Yuan, B.; et al. Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. *J. Allergy Clin. Immunol.* **2017**, *139*, 232–245. [[CrossRef](#)]
51. Vorsteveld, E.E.; Hoischen, A.; van der Made, C.I. Next-Generation Sequencing in the Field of Primary Immunodeficiencies: Current Yield, Challenges, and Future Perspectives. *Clin. Rev. Allergy Immunol.* **2021**, *61*, 212–225. [[CrossRef](#)] [[PubMed](#)]
52. Franke, B. What's in a name: The "omnigenic" model as a point of departure for polygenic psychiatric disorders. *J. Psychiatry Brain Sci.* **2017**, *2*, 5.
53. Lage, K.; Hansen, N.T.; Karlberg, E.O.; Eklund, A.C.; Roque, F.S.; Donahoe, P.K.; Szallasi, Z.; Jensen, T.S.; Brunak, S. A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20870–20875. [[CrossRef](#)]
54. Aguet, F.; Ardlie, K.G. Tissue specificity of gene expression. *Curr. Genet. Med. Rep.* **2016**, *4*, 163–169. [[CrossRef](#)]
55. Ganna, A.; Satterstrom, F.K.; Zekavat, S.M.; Das, I.; Kurki, M.I.; Churchhouse, C.; Alfoldi, J.; Martin, A.R.; Havulinna, A.S.; Byrnes, A.; et al. Quantifying the Impact of Rare and Ultra-rare Coding Variation across the Phenotypic Spectrum. *Am. J. Hum. Genet.* **2018**, *102*, 1204–1211. [[CrossRef](#)] [[PubMed](#)]
56. Zoghbi, A.W.; Dhindsa, R.S.; Goldberg, T.E.; Mehralizade, A.; Motelow, J.E.; Wang, X.; Alkelai, A.; Harms, M.B.; Lieberman, J.A.; Markx, S.; et al. High-impact rare genetic variants in severe schizophrenia. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2112560118. [[CrossRef](#)]
57. Farhan, S.M.K.; Howrigan, D.P.; Abbott, L.E.; Klim, J.R.; Topp, S.D.; Byrnes, A.E.; Churchhouse, C.; Phatnani, H.; Smith, B.N.; Rampersaud, E.; et al. Exome sequencing in amyotrophic lateral sclerosis implicates a novel gene, DNAJC7, encoding a heat-shock protein. *Nat. Neurosci.* **2019**, *22*, 1966–1974. [[CrossRef](#)]
58. Epi25Collaborative. Ultra-rare genetic variation in the epilepsies: A whole-exome sequencing study of 17,606 individuals. *Am. J. Hum. Genet.* **2019**, *105*, 267–282. [[CrossRef](#)]
59. Xu, H.; Zhen, Q.; Bai, M.; Fang, L.; Zhang, Y.; Li, B.; Ge, H.; Moon, S.; Chen, W.; Fu, W.; et al. Deep sequencing of 1320 genes reveals the landscape of protein-truncating variants and their contribution to psoriasis in 19,973 Chinese individuals. *Genome Res.* **2021**, *31*, 1150–1158. [[CrossRef](#)]
60. Lali, R.; Chong, M.; Omid, A.; Mohammadi-Shemirani, P.; Le, A.; Cui, E.; Paré, G. Calibrated rare variant genetic risk scores for complex disease prediction using large exome sequence repositories. *Nat. Commun.* **2021**, *12*, 5852. [[CrossRef](#)]
61. Sattler, S. The role of the immune system beyond the fight against infection. *Adv. Exp. Med. Biol.* **2017**, *1003*, 3–14. [[PubMed](#)]
62. Rankin, L.C.; Artis, D. Beyond host defense: Emerging functions of the immune system in regulating complex tissue physiology. *Cell* **2018**, *173*, 554–567. [[CrossRef](#)] [[PubMed](#)]
63. Zguro, K.; Fallerini, C.; Fava, F.; Furini, S.; Renieri, A. Host genetic basis of COVID-19: From methodologies to genes. *Eur. J. Hum. Genet.* **2022**, *30*, 899–907. [[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.