ELECTRONIC SUPPLEMENTARY MATERIAL 5

Key factors of the deranged antiviral response in elderly patients with COVID-19: a machine-learning analysis

Tamara Postigo[1], Alicia Ortega[1,2], Alejandro Álvaro-Meca[3,4], Daniel Vélez-Serrano[5], Nadia García-Mateo[1,2], Ana P. Tedim[1,2], Raquel Almansa[1,2,6], Jose María Eiros[7], David de Gonzalo-Calvo[2,8], Anna Moncusí-Moix[2,8], Clara Gort-Paniello[2,8], Manel Perez-Pons[2,8], Marta Molinero[2,8], Marta Dominguez-Gil[7], Amanda de la Fuente[1,2], Laura González-González[1,2], Tania Luis-García[1,2], Fátima Rodríguez-Jara[2,8], Noelia Jorge[1,2], Jessica González[8], Gerard Torres[2,8], Oliver Norberto Gutiérrez-Pérez[9], Maria José Villegas[9], Sonia Campo[9], Eva Ayllon[9], Tomás Ruiz Albi[10], Julio de Frutos Arribas[10,11], Ainhoa Arroyo Domingo[10], Jesica Abadia-Otero[12], Julia Gómez Barquero[12], Wysali Trapiello[13], Luis Javier Garcia Frade[9], Luis Inglada[12], Felix del Campo[10,11,14], Ferran Barbé[2,8,15], Antoni Torres[2,15,16], Raúl López-Izquierdo[17,†], Jesús F. Bermejo-Martin[1,2,18 ,†,*], CIBERES-UCI-COVID group.

† These senior authors contributed equally to this article.

Institutions:

[1]Group for Biomedical Research in Sepsis (BioSepsis). Instituto de Investigación Biomédica de Salamanca, (IBSAL), Gerencia Regional de Salud de Castilla y León, Paseo de San Vicente, 58-182, 37007 Salamanca, Spain.

[2]Centro de Investigación Biomédica en Red en Enfermedades Respiratorias (CIBERES), Instituto de Salud Carlos III, Avenida de Monforte de Lemos, 3-5, 28029 Madrid, Spain.

[3]Department of Preventive Medicine and Public Health, Faculty of Health Science, Universidad Rey Juan Carlos, Madrid, Spain.

[4]Centro de Investigación Biomédica en Red en Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Avenida de Monforte de Lemos, 3-5, 28029, Madrid, Spain.

[5]Department of Statistics and Operations Research, Universidad Complutense de Madrid, Madrid, Spain.

[6]Department of Cellular Biology, Hystology and Pharmacology. University of Valladolid, Av. Ramón y Cajal, 7, 47005 Valladolid, Spain.

[7]Microbiology Service, Hospital Universitario Rio Hortega, Gerencia Regional de Salud de Castilla y León, Calle Dulzaina, 2, 47012 Valladolid, Spain.

[8]Translational Research in Respiratory Medicine, University Hospital Arnau de Vilanova and Santa Maria, IRB Lleida, Av. Alcalde Rovira Roure, 80, 25198 Lleida Spain.

[9]Hematology Service. Hospital Universitario Río Hortega, Gerencia Regional de Salud de Castilla y León, Calle Dulzaina, 2, 47012 Valladolid, Spain.

[10]Pneumology Service, Hospital Universitario Río Hortega, Gerencia Regional de Salud de Castilla y León, Calle Dulzaina, 2, 47012 Valladolid, Spain.

[11]Deparment of Medicine, Dermatology and Toxicology, School of Medicine, University of Valladolid, Av. Ramón y Cajal, 7, 47005 Valladolid, Spain.

[12]Internal Medicine Service, Hospital Universitario Río Hortega, Gerencia Regional de Salud de Castilla y León, Calle Dulzaina, 2, 47012 Valladolid, Spain.

[13]Clinical Analysis Service. Hospital Clínico Universitario de Valladolid, Gerencia Regional de Salud de Castilla y León, Av. Ramón y Cajal, 3, 47003 Valladolid, Spain.

[14]Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Instituto de Salud Carlos III, Av. de Monforte de Lemos, 3-5, 28029 Madrid, Spain.

[15]Catalan Institution for Research and Advanced Studies, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain.

[16]Department of Pulmonology, Hospital Clinic de Barcelona, Universidad de Barcelona, Institut D investigacions August Pi I Sunyer (IDIBAPS), Carrer del Rosselló, 149, 08036 Barcelona, Spain.

[17]Emergency Department, Hospital Universitario Rio Hortega, Gerencia Regional de Salud de Castilla y León, Calle Dulzaina, 2, 47012 Valladolid, Spain.

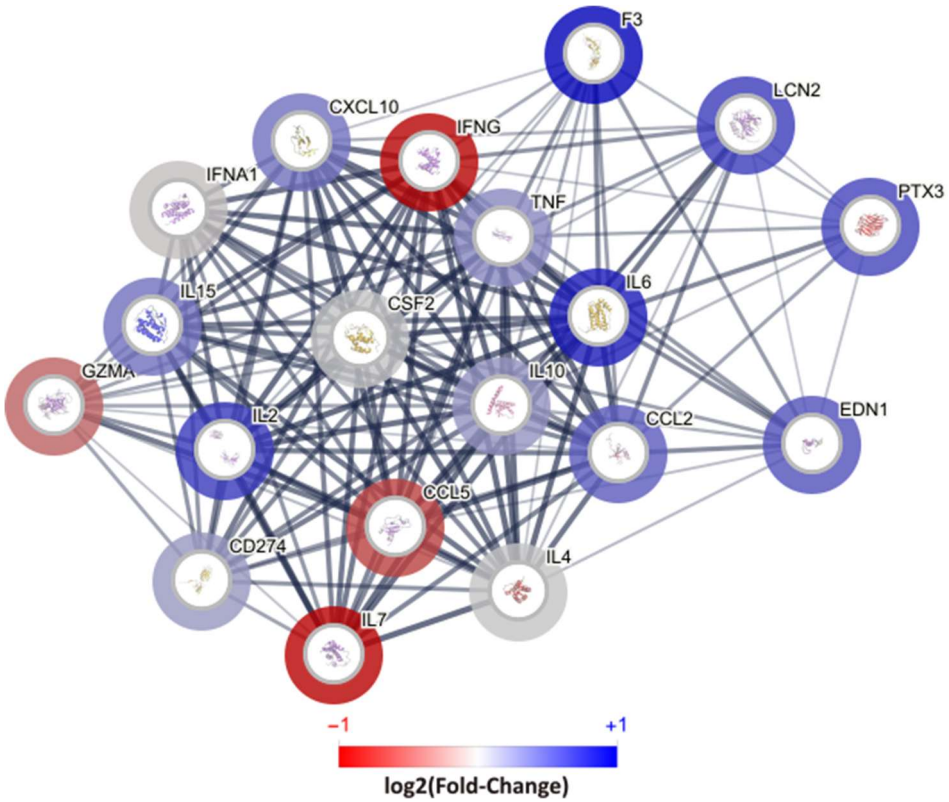
[18]Department of Medicine, Faculty of Medicine, University of Salamanca, Salamanca, Spain.

Corresponding Author: Dr. Jesús F Bermejo-Martin, Department of Medicine, Faculty of Medicine, University of Salamanca, C. Alfonso X El Sabio, s/n, 37007 Salamanca, Spain, jfbermejo@usal.es, +34 923 294 500 Ext. 1885.

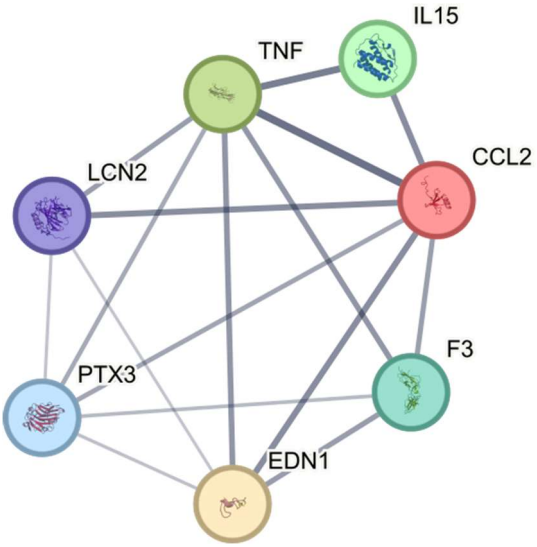
ELECTRONIC SUPPLEMENTARY MATERIAL 5

RESULTS:

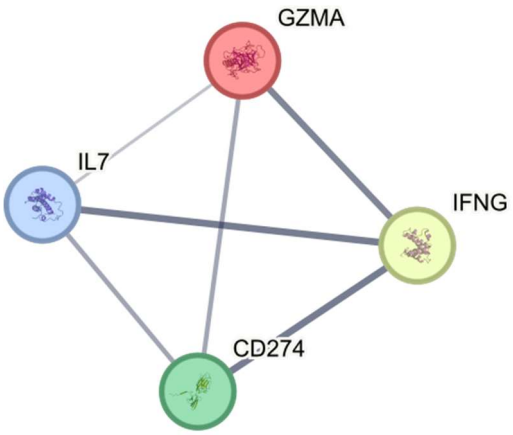
Protein–protein interaction network and functional pathway enrichment analysis



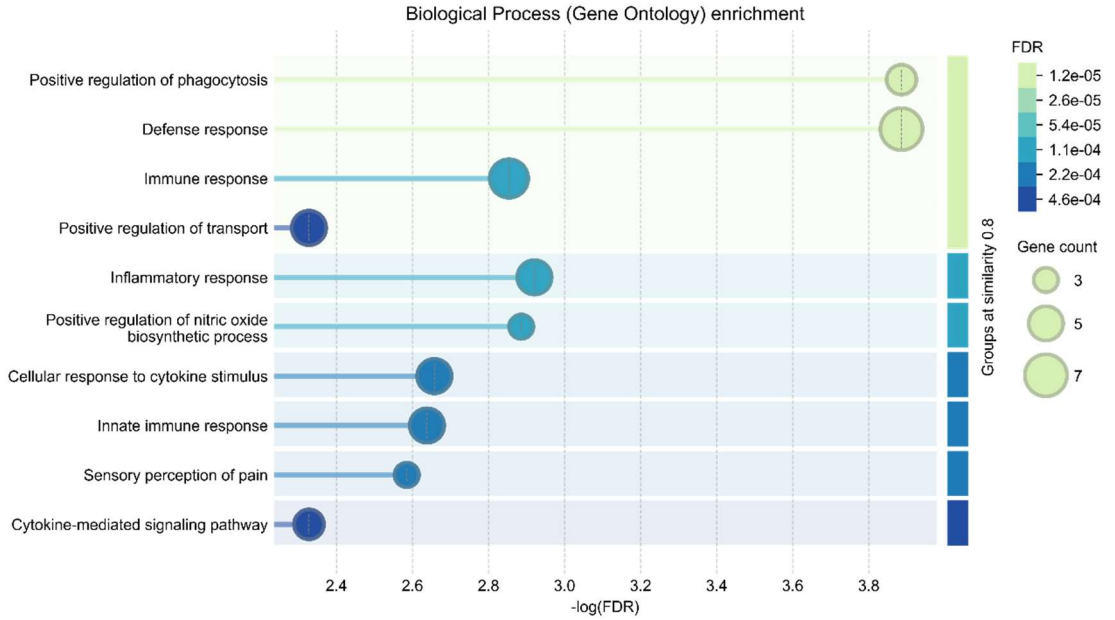
Supplementary Figure S3. Network plot representing protein-protein interactions and the magnitude and direction of change in plasma concentrations of the protein biomarkers studied between the elderly and adult groups of patients. The points representing each protein were colored according to log2-transformed fold change value, with negative values in colored in red and the positive ones in blue. Edges reflected evidence from experiments, curated databases, text mining, and co-expression that probe protein associations (the thicker the edges the greater association). (Interleukin-2: IL2, Granulocyte-macrophage colony-stimulating factor: CSF2, C-C motif chemokine ligand 2: CCL2, Granzyme A: GZMA, Interferon-gamma: IFNG, Interleukin-4: IL4, Tumor necrosis factor alpha: TNF, Programmed death-ligand 1: CD274, Endothelin-1: EDN1, Interferon alpha 1: IFNA1, Interleukin-10: IL10, Interleukin-7: IL7, Interleukin-15: IL15, C-X-C motif chemokine ligand 10: CXCL10, Interleukin-6: IL6, Tissue Factor: F3, C-C motif chemokine ligand 5: CCL5, Pentraxin-3: PTX3, Lipocalin-2: LCN2). Tissue Factor (F3) was used as an upward indicator of the procoagulant state reflected by the elevation of D-dimer in functional enrichment analyses.



Supplementary Figure S4. Network plot representing protein-protein interactions of the biomarkers that showed the greatest importance and the highest **positive** SHAP values in the final plot obtained after the training with the machine learning algorithm, XGBoost. Edges reflected evidence from experiments, curated databases, text mining, and co-expression that probe protein associations (the thicker the edges the greater association). (C-C motif chemokine ligand 2: CCL2, Tumor necrosis factor alpha: TNF, Endothelin-1: EDN1, Interleukin-15: IL15, Tissue Factor: F3, Pentraxin-3: PTX3, Lipocalin-2: LCN2). Tissue Factor (F3) was used as an upward indicator of the procoagulant state reflected by the elevation of D-dimer in functional enrichment analyses.



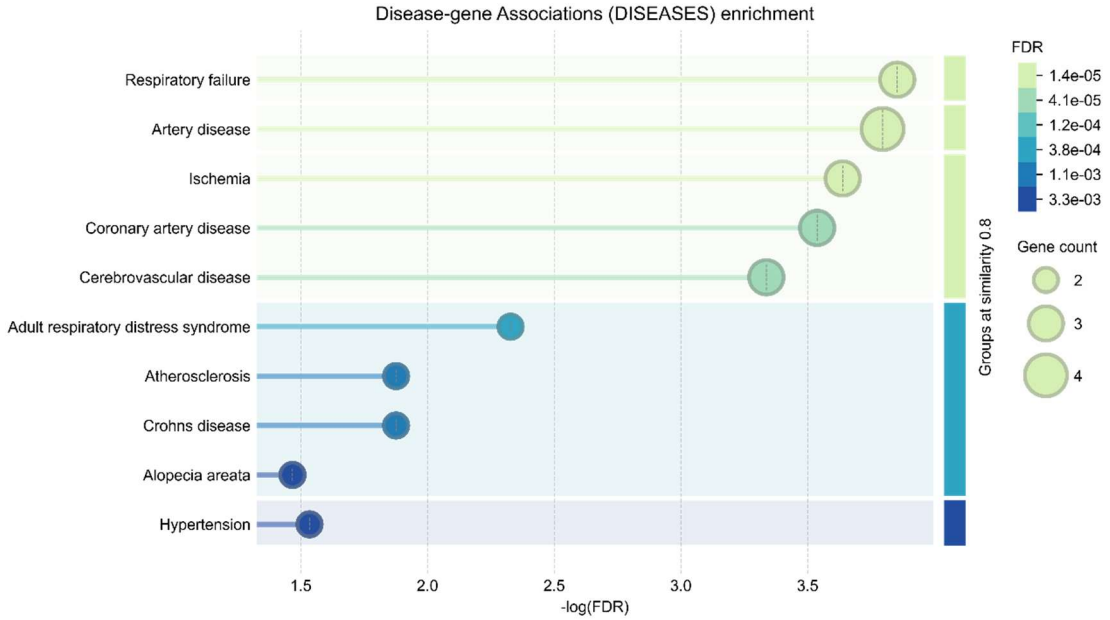
Supplementary Figure S5. Representation of the relationships between the protein biomarkers that showed the greatest importance and the highest **negative** SHAP values in the final plot obtained after the training with the machine learning algorithm, XGBoost. Edges reflected evidence from experiments, curated databases, text mining, and co-expression that probe protein associations (the thicker the edges the greater association). (Granzyme A: GZMA, Interferon-gamma: IFNG, Programmed death-ligand 1: CD274, Interleukin-7: IL7).



Supplementary Figure S6. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Representation of top 10 significantly enriched pathways in the 'Gene Ontology (GO) Biological Process' database identified by String according to their -log(FDR). False discovery rate (FDR).

Category	Term name	Description	p-value	FDR value	Overlap proteins
GO Biological Process	GO:0006952	Defense response	9.05E-9	1.3E-4	[CCL2, PTX3, IL15, F3, LCN2, EDN1, TNF]
GO Biological Process	GO:0050766	Positive regulation of phagocytosis	8.3E-9	1.3E-4	[CCL2, PTX3, IL15, TNF]
GO Biological Process	GO:0006954	Inflammatory response	3.13E-7	0.0012	[CCL2, PTX3, IL15, F3, TNF]
GO Biological Process	GO:0045429	Positive regulation of nitric oxide biosynthetic process	4.14E-7	0.0013	[PTX3, EDN1, TNF]
GO Biological Process	GO:0006955	Immune response	6.09E-7	0.0014	[CCL2, PTX3, IL15, LCN2, EDN1, TNF]
GO Biological Process	GO:0071345	Cellular response to cytokine stimulus	1.23E-6	0.0022	[CCL2, IL15, F3, EDN1, TNF]
GO Biological Process	GO:0045087	Innate immune response	1.65E-6	0.0023	[CCL2, PTX3, LCN2, EDN1, TNF]
GO Biological Process	GO:0019233	Sensory perception of pain	1.99E-6	0.0026	[CCL2, EDN1, TNF]
GO Biological Process	GO:0019221	Cytokine-mediated signaling pathway	4.23E-6	0.0047	[CCL2, IL15, F3, TNF]
GO Biological Process	GO:0051050	Positive regulation of transport	4.26E-6	0.0047	[CCL2, PTX3, IL15, EDN1, TNF]
GO Biological Process	GO:0032101	Regulation of response to external stimulus	5.5E-6	0.0051	[CCL2, IL15, F3, EDN1, TNF]
GO Biological Process	GO:0071346	Cellular response to interferon-gamma	5.51E-6	0.0051	[CCL2, EDN1, TNF]
GO Biological Process	GO:0009893	Positive regulation of metabolic process	1.09E-5	0.0074	[CCL2, PTX3, IL15, F3, LCN2, EDN1, TNF]
GO Biological Process	GO:0032103	Positive regulation of response to external stimulus	9.45E-6	0.0074	[IL15, F3, EDN1, TNF]
GO Biological Process	GO:0061041	Regulation of wound healing	1.03E-5	0.0074	[F3, EDN1, TNF]
GO Biological Process	GO:0001936	Regulation of endothelial cell proliferation	1.23E-5	0.0077	[CCL2, F3, TNF]
GO Biological Process	GO:0048660	Regulation of smooth muscle cell proliferation	1.34E-5	0.0081	[IL15, EDN1, TNF]
GO Biological Process	GO:0071396	Cellular response to lipid	1.41E-5	0.0082	[CCL2, IL15, EDN1, TNF]
GO Biological Process	GO:0071407	Cellular response to organic cyclic compound	1.48E-5	0.0083	[CCL2, IL15, EDN1, TNF]
GO Biological Process	GO:0009605	Response to external stimulus	1.85E-5	0.01	[CCL2, PTX3, IL15, LCN2, EDN1, TNF]
GO Biological Process	GO:0071356	Cellular response to tumor necrosis factor	2.47E-5	0.0121	[CCL2, EDN1, TNF]
GO Biological Process	GO:0010469	Regulation of signaling receptor activity	2.64E-5	0.0125	[CCL2, EDN1, TNF]
GO Biological Process	GO:0051769	Regulation of nitric-oxide synthase biosynthetic process	2.73E-5	0.0125	[CCL2, EDN1]
GO Biological Process	GO:0009617	Response to bacterium	4.2E-5	0.0169	[CCL2, LCN2, EDN1, TNF]
GO Biological Process	GO:0000165	MAPK cascade	4.78E-5	0.0183	[CCL2, EDN1, TNF]

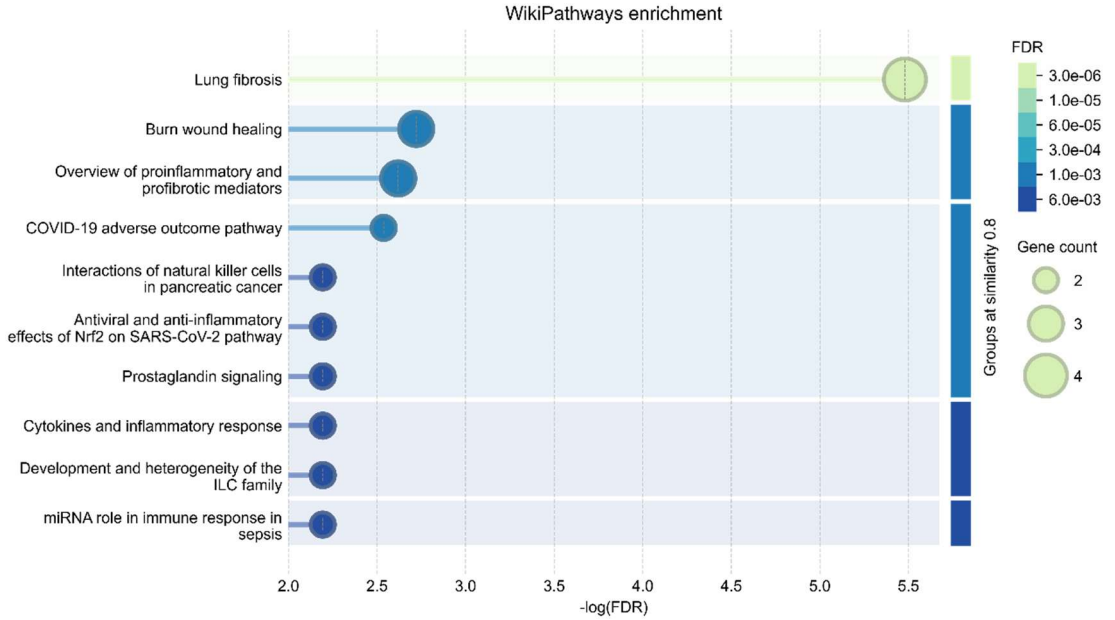
Supplementary Table S5. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Top 25 (total = 68) significantly enriched pathways in the 'Gene Ontology (GO) Biological Process' database identified by String. Each row reports the category (database), the term name and description (code and name of the pathway in the database), the corresponding raw p-value and the False Discovery Rate (FDR), a Benjamini-Hochberg adjusted p-value for multiple hypothesis testing. (C-C motif chemokine ligand 2: CCL2, Tumor necrosis factor alpha: TNF, Endothelin-1: EDN1, Interleukin-15: IL15, Tissue Factor: F3, Pentraxin-3: PTX3, Lipocalin-2: LCN2).



Supplementary Figure S7. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Representation of top 10 significantly enriched pathways in the 'Diseases-gene associations (DISEASES)' database identified by String according to their -log(FDR). False discovery rate (FDR).

Category	Term name	Description	p-value	FDR value	Overlap proteins
DISEASES	DOID:11162	Respiratory failure	3.12E-8	1.4E-4	[CCL2, F3, TNF]
DISEASES	DOID:0050828	Artery disease	7.04E-8	1.6E-4	[CCL2, F3, EDN1, TNF]
DISEASES	DOID:326	Ischemia	1.49E-7	2.3E-4	[F3, EDN1, TNF]
DISEASES	DOID:3393	Coronary artery disease	2.49E-7	2.9E-4	[F3, EDN1, TNF]
DISEASES	DOID:6713	Cerebrovascular disease	5.02E-7	4.6E-4	[F3, EDN1, TNF]
DISEASES	DOID:11394	Adult respiratory distress syndrome	7.13E-6	0.0047	[CCL2, TNF]
DISEASES	DOID:1936	Atherosclerosis	2.73E-5	0.0133	[CCL2, TNF]
DISEASES	DOID:8778	Crohns disease	3.5E-5	0.0133	[CCL2, TNF]
DISEASES	DOID:10763	Hypertension	9.7E-5	0.0292	[EDN1, TNF]
DISEASES	DOID:986	Alopecia areata	1.3E-4	0.0341	[CCL2, TNF]
DISEASES	DOID:417	Autoimmune disease	1.4E-4	0.0361	[CCL2, F3, TNF]

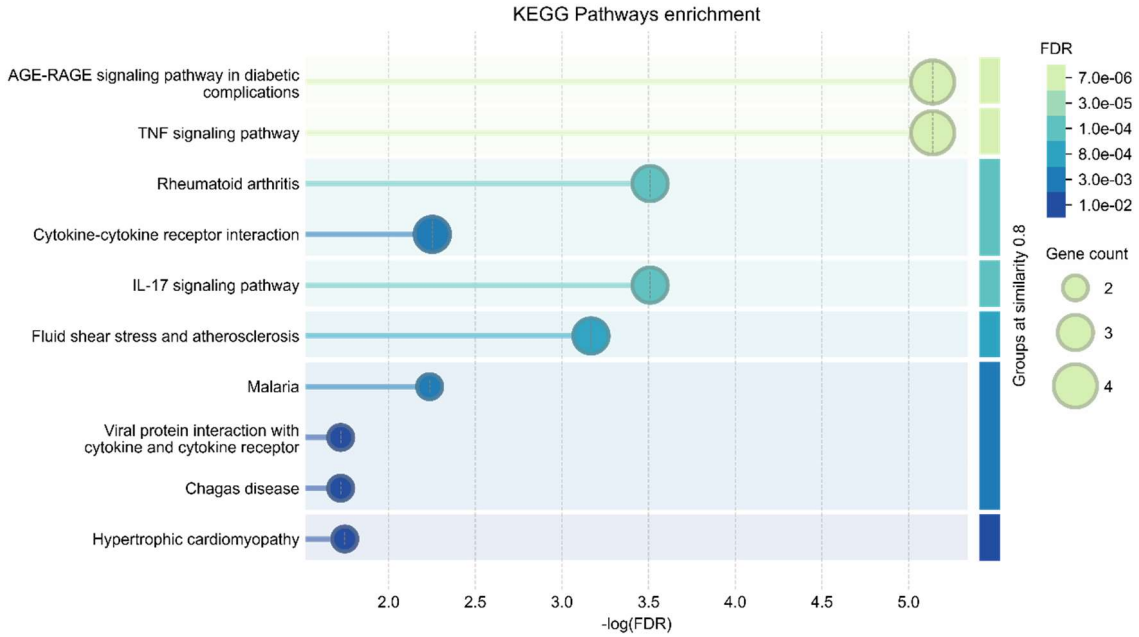
Supplementary Table S6. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Significantly enriched pathways in the 'Diseases-gene associations (DISEASES)' database identified by String. Each row reports the category (database), the term name and description (code and name of the pathway in the database), the corresponding raw p-value and the False Discovery Rate (FDR), a Benjamini-Hochberg adjusted p-value for multiple hypothesis testing. (C-C motif chemokine ligand 2: CCL2, Tumor necrosis factor alpha: TNF, Endothelin-1: EDN1, Interleukin-15: IL15, Tissue Factor: F3, Pentraxin-3: PTX3, Lipocalin-2: LCN2).



Supplementary Figure S8. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Representation of top 10 significantly enriched pathways in the 'WikiPathways 2024 Human' database identified by String according to their -log(FDR). False discovery rate (FDR).

Category	Term name	Description	p-value	FDR value	Overlap proteins
WikiPathways	WP3624	Lung fibrosis	4.24E-9	3.31E-6	[CCL2, PTX3, EDN1, TNF]
WikiPathways	WP5055	Burn wound healing	4.78E-6	0.0019	[CCL2, IL15, TNF]
WikiPathways	WP5095	Overview of proinflammatory and profibrotic mediators	9.41E-6	0.0024	[CCL2, IL15, TNF]
WikiPathways	WP4891	COVID-19 adverse outcome pathway	1.47E-5	0.0029	[CCL2, TNF]
WikiPathways	WP2036	TNF-related weak inducer of apoptosis (TWEAK) signaling pathway	1.0E-4	0.0064	[CCL2, TNF]
WikiPathways	WP3893	Development and heterogeneity of the ILC family	6.04E-5	0.0064	[IL15, TNF]
WikiPathways	WP4136	Fibrin complement receptor 3 signaling pathway	1.1E-4	0.0064	[CCL2, TNF]
WikiPathways	WP4329	miRNA role in immune response in sepsis	7.96E-5	0.0064	[LCN2, TNF]
WikiPathways	WP4754	IL-18 signaling pathway	8.93E-5	0.0064	[CCL2, PTX3, TNF]
WikiPathways	WP5088	Prostaglandin signaling	6.04E-5	0.0064	[CCL2, TNF]
WikiPathways	WP5092	Interactions of natural killer cells in pancreatic cancer	4.37E-5	0.0064	[CCL2, TNF]
WikiPathways	WP5113	Antiviral and anti-inflammatory effects of Nrf2 on SARS-CoV-2 pathway	5.68E-5	0.0064	[CCL2, TNF]
WikiPathways	WP5285	Immune infiltration in pancreatic cancer	8.38E-5	0.0064	[CCL2, TNF]
WikiPathways	WP530	Cytokines and inflammatory response	4.07E-5	0.0064	[IL15, TNF]
WikiPathways	WP1533	Vitamin B12 metabolism	1.4E-4	0.0074	[CCL2, TNF]
WikiPathways	WP4747	Netrin-UNC5B signaling pathway	1.5E-4	0.0075	[CCL2, TNF]
WikiPathways	WP176	Folate metabolism	2.5E-4	0.0112	[CCL2, TNF]
WikiPathways	WP5039	SARS-CoV-2 innate immunity evasion and cell-specific immune response	2.4E-4	0.0112	[CCL2, TNF]
WikiPathways	WP4341	Non-genomic actions of 1,25 dihydroxyvitamin D3	3.0E-4	0.0125	[CCL2, TNF]
WikiPathways	WP15	Selenium micronutrient network	3.9E-4	0.0149	[CCL2, TNF]
WikiPathways	WP1544	MicroRNAs in cardiomyocyte hypertrophy	3.8E-4	0.0149	[EDN1, TNF]
WikiPathways	WP4298	Acute viral myocarditis	4.0E-4	0.0149	[EDN1, TNF]
WikiPathways	WP231	TNF-alpha signaling pathway	4.7E-4	0.0151	[CCL2, TNF]
WikiPathways	WP4963	p53 transcriptional gene network	4.5E-4	0.0151	[CCL2, TNF]
WikiPathways	WP2431	Spinal cord injury	7.2E-4	0.0225	[CCL2, TNF]
WikiPathways	WP4396	Nonalcoholic fatty liver disease	0.0013	0.0383	[CCL2, TNF]

Supplementary Table S7. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Significantly enriched pathways in the 'WikiPathways 2024 Human' database identified by String. Each row reports the category (database), the term name and description (code and name of the pathway in the database), the corresponding raw p-value and the False Discovery Rate (FDR), a Benjamini-Hochberg adjusted p-value for multiple hypothesis testing. (C-C motif chemokine ligand 2: CCL2, Tumor necrosis factor alpha: TNF, Endothelin-1: EDN1, Interleukin-15: IL15, Tissue Factor: F3, Pentraxin-3: PTX3, Lipocalin-2: LCN2).



Supplementary Figure S9. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Representation of top 10 significantly enriched pathways in the 'Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways' database identified by String according to their -log(FDR). False discovery rate (FDR).

Category	Term name	Description	p-value	FDR value	Overlap proteins
KEGG Pathways	hsa04668	TNF signaling pathway	3.8E-8	7.26E-6	[CCL2, IL15, EDN1, TNF]
KEGG Pathways	hsa04933	AGE-RAGE signaling pathway in diabetic complications	2.16E-8	7.26E-6	[CCL2, F3, EDN1, TNF]
KEGG Pathways	hsa04657	IL-17 signaling pathway	3.63E-6	3.1E-4	[CCL2, LCN2, TNF]
KEGG Pathways	hsa05323	Rheumatoid arthritis	2.77E-6	3.1E-4	[CCL2, IL15, TNF]
KEGG Pathways	hsa05418	Fluid shear stress and atherosclerosis	1.01E-5	6.8E-4	[CCL2, EDN1, TNF]
KEGG Pathways	hsa04060	Cytokine-cytokine receptor interaction	1.0E-4	0.0056	[CCL2, IL15, TNF]
KEGG Pathways	hsa05144	Malaria	1.2E-4	0.0058	[CCL2, TNF]
KEGG Pathways	hsa05410	Hypertrophic cardiomyopathy	4.3E-4	0.0179	[EDN1, TNF]
KEGG Pathways	hsa04061	Viral protein interaction with cytokine and cytokine receptor	5.1E-4	0.0189	[CCL2, TNF]
KEGG Pathways	hsa05142	Chagas disease	5.2E-4	0.0189	[CCL2, TNF]
KEGG Pathways	hsa05135	Yersinia infection	8.3E-4	0.0255	[CCL2, TNF]
KEGG Pathways	hsa05164	Influenza A	0.0014	0.0399	[CCL2, TNF]
KEGG Pathways	hsa04621	NOD-like receptor signaling pathway	0.0016	0.0413	[CCL2, TNF]

Supplementary Table S8. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **positive** SHAP values. Significantly enriched pathways in the 'Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways' database identified by String. Each row reports the category (database), the term name and description (code and name of the pathway in the database), the corresponding raw p-value and the False Discovery Rate (FDR), a Benjamini-Hochberg adjusted p-value for multiple hypothesis testing. (C-C motif chemokine ligand 2: CCL2, Tumor necrosis factor alpha: TNF, Endothelin-1: EDN1, Interleukin-15: IL15, Tissue Factor: F3, Pentraxin-3: PTX3, Lipocalin-2: LCN2).

Category	Term name	Description	p-value	FDR value	Overlap proteins
TISSUES	BTO:0000289	Cytotoxic T-lymphocyte	1.03E-8	2.49E-5	[IFNG, GZMA, CD274]
TISSUES	BTO:0004520	Regulatory T-lymphocyte	4.2E-6	0.0051	[IFNG, CD274]
WikiPathways	WP3893	Development and heterogeneity of the ILC family	1.73E-5	0.0072	[IFNG, IL7]
WikiPathways	WP4585	Cancer immunotherapy by PD-1 blockade	9.26E-6	0.0072	[IFNG, CD274]
WikiPathways	WP530	Cytokines and inflammatory response	1.17E-5	0.0072	[IFNG, IL7]
DISEASES	DOID:614	Lymphopenia	1.7E-6	0.0078	[IFNG, IL7]
TISSUES	BTO:0000914	Natural killer cell	6.21E-5	0.0376	[IFNG, GZMA]
KEGG Pathways	hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	1.2E-4	0.0404	[IFNG, CD274]
WikiPathways	WP5095	Overview of proinflammatory and profibrotic mediators	2.5E-4	0.0486	[IFNG, IL7]

Supplementary Table S9. Functional enrichment analysis for the biomarkers with the greatest importance and the highest **negative** SHAP values. Significantly enriched pathways in the 'Brenda Tissue Ontology (BTO) / Tissue Expression (TISSUES)', WikiPathways 2024 Human', Diseases-gene associations (DISEASES)', and 'Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathways' databases identified by String. Each row reports the category (database), the term name and description (code and name of the pathway in the database), the corresponding raw p-value and the False Discovery Rate (FDR), a Benjamini-Hochberg adjusted p-value for multiple hypothesis testing. (Granzyme A: GZMA, Interferon-gamma: IFNG, Programmed death-ligand 1: CD274, Interleukin-7: IL7).