

Original article

Transcriptomic Profile and Expression of Immune Checkpoints in Peripheral Blood Mononuclear Cells of Women with Breast Fibroadenoma

Olga V. Kurmyshkina, Tatyana M. Kulikovskaya, Tatyana O. Volkova

Petrozavodsk State University, Petrozavodsk, Russia

Received 5 May 2025, Revised 25 September 2025, Accepted 20 November 2025

© 2025, Russian Open Medical Journal

Abstract: *Background* — Fibroepithelial tumors of the mammary gland known as fibroadenoma (FA), being one of the most common variants, represent a rather heterogeneous group of diseases, the pathogenetic mechanisms of which remain largely unknown. This study aimed at analyzing two categories of putative changes in the cell-mediated immunity in women diagnosed with FA: (a) changes affecting the transcriptomic profile of the peripheral blood mononuclear cell (PBMC) fraction and the spectrum of activated/ dysregulated signaling pathways that form a specific immune and stromal environment of the tumor; and (b) changes in (co-)expression of immune checkpoint markers — CD279/PD-1, CD274/PD-L1, CD366/TIM3 and CD223/LAG3 — in the CD4 and CD8 subsets of circulating T lymphocytes.

Methods — Transcriptome profiling was performed using high-throughput RNA sequencing on the Illumina next-generation sequencing platform, followed by bioinformatic analysis of the resulting datasets including differential gene expression and signaling pathway enrichment analysis. Immune checkpoint marker expression patterns were identified using multicolor flow cytometry.

Results — Principal component analysis and hierarchical clustering of the transcriptome datasets revealed three clusters, with the FA cluster clearly distinguishable from the control group. DESeq2 software package detected a total of 334 genes with differential expression between PBMCs from FA patients and healthy controls (with an adjusted p -value <0.05 and fold change of expression (FC) >2.0). Gene Ontology (GO) and KEGG enrichment analyses revealed that differentially expressed genes (DEGs) were enriched in over 500 GO terms of the Biological Processes sub-ontology with a false discovery rate (FDR) set to <0.01 and more than 20 KEGG signaling pathways (FDR <0.25). The most frequently occurring GO terms included Th1 activation processes, monocyte maturation, and inflammatory responses mediated by interferon- γ (IFN- γ) and IL-12 signaling pathways, as well as IL-1, IL-6, IL-8, and TNF α -dependent pathways. Genes enriched in the negative regulation of mononuclear cell migration and T cell differentiation (via IL-4- and IL-10-dependent mechanisms) were downregulated. Another group of overrepresented terms was related to the cellular response to stress including the antioxidant defense responses mediated by metallothionein MT1E and SOD1. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed the involvement of various functional activities of antigen-presenting cells (macrophages, dendritic cells). Gene set enrichment analysis (GSEA) was performed using molecular signatures obtained from the MSigDB database (viz., the C4 and C7 ImmuneSigDB, and Hallmark subcollections) revealed that transcriptomic changes detected in the FA datasets were linked to specific metaprograms and gene networks associated with cancer, such as epithelial-mesenchymal transition (EMT) or myofibroblast-specific metaprogram (FDR <0.05 ; normalized enrichment score: |NES| >1.5). GSEA also demonstrated the involvement of pathways responsible for the differentiation and functional activity of regulatory T cells (Treg) (FDR <0.05 , |NES| >1.8) and altered PD1 expression status (PD1^{low}/PD1^{high}), indicating immune exhaustion. Flow cytometry analysis demonstrated an increase ($p<0.05$, U test) in the percentage of CD4/CD8 T cells co-expressing three checkpoint markers (PD1⁺PD-L1⁺TIM3⁺ or PD1⁺PD-L1⁺LAG3⁺ cells) in PBMCs from FA patients vs. the control group, suggesting enhanced immune checkpoint signaling and expansion of exhausted T cells.

Conclusion — To our knowledge, this is the first study to examine the transcriptomic features of circulating lymphocytes and monocytes in women diagnosed with breast FA and the first to analyze the co-expression pattern of immune checkpoint markers. Our results demonstrate the involvement of specific immune system molecular pathways in the pathogenesis of FA and the intensification of immune exhaustion mechanisms.

Keywords: transcriptome, breast fibroadenoma, signaling pathways, immune response, peripheral blood lymphocytes.

Cite as Kurmyshkina OV, Kulikovskaya TM, Volkova TO. Transcriptomic profile and expression of immune checkpoints in peripheral blood mononuclear cells of women with breast fibroadenoma. *Russian Open Medical Journal* 2026; 15: e0102.

Correspondence to Tatyana O. Volkova. Address: 33 Lenin St., Petrozavodsk 185910, Russia. Phone: +78142784697. E-mail: VolkovaTO@yandex.ru.

Introduction

Fibroadenoma of the breast (FA) is the most frequently diagnosed type of benign (non-cancerous) fibroepithelial tumors characterized by a biphasic structure demonstrating proliferative

changes in both epithelial and stromal (connective tissue) components in varying proportions. The main histological variants of FA are well known, and in most cases this diagnosis implies conservative treatment or, depending on specific medical indications, surgical removal of the tumor [1]. However, upon

closer examination, fibroepithelial tumors represent a heterogeneous group of diseases with highly variable and largely overlapping morphological features, which can complicate differential diagnosis and treatment decisions, and may increase the likelihood of overestimation or, conversely, underestimation of disease severity and prognosis, as well as the risk of developing a malignancy [2, 3]. Besides FA, a variety of histopathological changes have been described, affecting not only stromal but also epithelial compartments, sometimes including features of atypical hyperplasia and carcinoma in situ [4].

Clearly, molecular genetic causes must underlie this morphological and clinical variability rooted in dysregulated signaling mechanisms. However, the mechanisms underlying the pathogenesis of FA remain largely unknown [5] due to a rather paradoxical situation: FA is rarely chosen as the primary and independent object of molecular genetic research due to its favorable clinical prognosis in most cases; women with this diagnosis usually constitute a comparison group (or control group) used to study various aspects of the molecular pathogenesis of malignant breast tumors, and therefore FA is considered a normal variant. Changes in immune status in patients with FA are also poorly understood due to the prevailing opinion that fibroepithelial diseases have a (dys)hormonal etiology. Very little is known about possible changes in the immune microenvironment or the profile of circulating immune-related biomarkers in FA. However, based on current knowledge of bidirectional immunoendocrine interactions (including the mutual influence of the estrogen and progesterone, and also between lymphocytes, monocytes and granulocytes), it is possible to hypothesize the involvement of the immune system in the development of FA.

Some experimental data directly or indirectly indicate that immunoregulatory mechanisms may be involved in the pathogenesis of FA. First, despite its benign nature, the development of FA is probably associated with mutational processes. Large-scale genomic screening of fibroepithelial tumors (fibroadenomas and phyllodes tumors) has recently revealed a specific landscape of recurrent genetic alterations in key proto-oncogenes and tumor suppressor genes [6], with MED12, KMT2D, and RARA being the most frequently mutated in ordinary FAs. Next-generation sequencing detected mutations in genes encoding various components of the DNA repair system [7, 8], implying the emergence of neoantigens and, consequently, the existence of immune surveillance of these antigens, including the so-called immune checkpoints [9, 10]. Importantly, this mutational landscape was linked with the tumor growth pattern and some other morphological characteristics of FAs [1]. With regard to the genome-wide methylation profile, fibroepithelial tumors (FAs and phyllodes tumors) form a distinct cluster that differs from both normal breast tissue and cancer tissue. As for the pattern of copy number variations (CNVs), abnormalities commonly observed in malignant pathology cannot be ruled out [11]. Clearly, the structural genomic and epigenomic aberrations observed in FA may manifest themselves in alterations of the activation of various signaling pathways, including those mediating relationships with immune mechanisms [7]. Various changes in the secreted proteome were identified in FA samples: specifically, the content of FA-derived extracellular vesicles differed much from that of normal tissue, but was largely consistent with breast cancer in terms of signaling pathway profile [12], contributing to the formation of a specific local immune environment in FA. Indeed, very few published transcriptomic studies focusing on the issue of

benign tumors have revealed the involvement of immune-related signaling pathways [13, 14]. Some studies reported increased expression of some specific markers of epithelial-mesenchymal transition (EMT) and extracellular matrix remodeling known for their immunomodulatory activity [15]. Apart from the molecular genetics data, there is some direct evidence demonstrating local changes in the distribution of different populations of immune cells in FA, which most likely reflects dysregulations in the immunosurveillance. For instance, some studies described the changes in the levels of CD4/CD8 T lymphocytes and antigen-presenting cells infiltrating FA vs. the normal breast tissue [16, 17]. Terzeman et al. showed that the connective tissue compartment in FA was infiltrated by mast cells producing vasoactive mediators and proinflammatory immune factors [18].

Local molecular genetic alterations can extend to the systemic level and can be detected in the peripheral circulation as abnormalities in the phenotypic characteristics of peripheral leukocytes or immunological parameters of blood plasma. For example, proteomic analysis of plasma samples from patients with FA or other benign breast lesions revealed systemic alterations associated with innate immune signaling pathways [19]. In another study, whole blood transcriptome analysis allowed the characterization of specific gene expression profiles encompassing adaptive immune mechanisms in a number of high-risk breast pathologies, including FA [20]. There are also data demonstrating altered levels of some circulating microRNAs known to be involved in immune regulation and inflammation mechanisms [21]. Regarding the variations in immunophenotypes of different lymphocyte populations, there is increasing attention paid to the expression of immune checkpoint markers and the possibilities of their clinical application in various breast neoplasms, since these markers are believed to reflect not only the state of cellular immunity activation but also its exhaustion (immunosuppression), which may affect the prognosis of the disease and the effectiveness of a particular therapeutic approach. In FA, a marked increase in the expression of the PD-1 checkpoint marker in the circulating regulatory T cell (Treg) population and a trend towards increased PD-L1 expression in regulatory B cells without changes in cell number was demonstrated [22]; however, the expression/co-expression patterns of other immune checkpoint molecules have not yet been investigated. Similarly, there are no published studies examining transcriptomic changes in various peripheral blood cell subsets, primarily lymphocytes/monocytes, in patients with FA.

The aim of this study was to analyze putative changes in cellular immunity in women diagnosed with FA in two aspects: (a) changes affecting the transcriptomic profile of the peripheral blood mononuclear cell (PBMC) fraction and the spectrum of activated/dysregulated signaling pathways that form a specific immune and stromal environment of the tumor; (b) changes in (co-)expression of immune checkpoint markers – CD279/PD-1, CD274/PD-L1, CD366/TIM3 and CD223/LAG3 – in the CD4 and CD8 subsets of circulating T lymphocytes. Data were collected using RNA sequencing technology, followed by bioinformatic analysis and flow cytometry.

Material and Methods

Patients and samples

Peripheral blood samples were collected from women diagnosed with breast FA (23-54 years of age, mean age 35.8 years, n=16) undergoing surgery at the private medical institution,

RZhD-Medicine Clinical Hospital (Petrozavodsk, Russia), as well as from healthy women of the control group (22-52 years of age, mean age 38.6 years; n=16; no breast pathology and no acute infectious, inflammatory or chronic diseases). Venous blood was collected immediately before surgical removal of the breast tumor; samples were transported to the laboratory at a temperature of +4 °C and processed no later than 2 h after blood collection. In all cases, the primary diagnosis was confirmed by histopathological examination. Patient exclusion criteria were as follows: signs of a malignant neoplasm of the breast, a history of oncological diseases, pregnancy or lactation. All women were informed about the purposes of blood collection and provided written voluntary consent. The study was approved by the Medical Ethics Committee of the Institute of Medicine, Petrozavodsk State University and the Republic of Karelia Ministry of Healthcare (Protocol No. 3, approval date: October 21, 2024) and was conducted in accordance with the Declaration of Helsinki and good clinical practice guidelines.

RNA sequencing (RNA isolation and cDNA library creation).

PBMCs were isolated by density gradient centrifugation in Ficoll-Paque PLUS density gradient medium (density: 1.077 g/mL), GE Healthcare, USA. The total RNA fraction was purified from PBMCs by phenol-chloroform extraction using TriZOL reagent (Invitrogen, USA) according to the manufacturer's instructions. RNA samples were then treated with DNase I (New England Biolabs, USA) to remove contaminating genomic DNA, followed by the spin column-based purification of the DNase reaction products with the CleanRNA Standard kit (Eurogen, Russia). Total RNA concentration was determined fluorometrically using a selective RNA dye and the Equalbit RNA BR Assay kit (Vazyme, China). After DNase treatment and purification, the integrity of the total RNA samples was monitored in a Fragment Analyzer automated capillary electrophoresis system (Advanced Analytical Technologies, Agilent, USA) using the HS RNA reagent kit (Agilent Technologies, USA). RNA quality assessment was based on the 28S:18S rRNA ratio. The presence of residual contaminants in RNA samples was determined based on the absorbance ratio at 260/280 nm and 260/230 nm using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). Solely RNA

samples that met quality requirements were selected for subsequent cDNA library preparation and sequencing.

cDNA libraries were created using the TruSeq Stranded Total RNA with RiboZero single index reagent kit (Illumina, USA) supplemented with SuperScript III reverse transcriptase (Invitrogen, USA) and Agencourt RNAClean XP or Agencourt AMPure XP magnetic beads for purification of RNA fragments or cDNA and PCR products, respectively (Beckman Coulter, USA). The amount of RNA used for cDNA library preparation was 500-1,000 ng. Quantification of the synthesized cDNA libraries was performed by fluorometric analysis based on an intercalating dye using the QuDye BR reagent kit (Lumiprobe, Russia). Correspondence of the obtained DNA fragments to the theoretically expected size of 260 bp was checked using the Fragment Analyzer system and the HS NGS reagent kit (Agilent Technologies, USA). After normalization, denaturation, and dilution steps, the amplified cDNA libraries were loaded into a flow cell; paired-end reads of 75 bp in length were obtained using the MiSeq v3 reagent kit (150 cycles) on the MiSeq sequencing platform (Illumina, USA).

Bioinformatic analysis

Bioinformatic analysis was performed on the Galaxy web portal (<https://usegalaxy.eu>) [23]. The quality of the raw sequence data was assessed using the FastQC quality control tool. The reads were then aligned to the human reference genome (assembly version: GRCh38/p13), and BAM files were generated using RNA STAR mapper and the GTF genome annotation file downloaded from the Gencode server (<https://www.encodegenes.org/human>). A summary of the read alignment statistics is presented in [Table 1](#). In addition, the distribution of mapped reads across the genome and its main structural elements, viz., exons (including 5'- and 3'-untranslated regions), intronic and intergenic regions, was examined using the Read Distribution tool from the RSeQC package based on the resulting BAM files. As expected, the majority of sequenced reads (%) was mapped to exonic regions, while the proportion of reads aligned to non-coding intergenic regions indicated the amount of residual contamination of genomic DNA.

Table 1. Summary statistics of RNA STAR read alignment results

Sample ID	Total number of paired-end reads	Alignment metrics				
		Uniquely mapped reads, %	Reads mapped to multiple loci, %	Unmapped reads, %	Reads mapped to exonic regions, %*	Reads mapped to intergenic regions, %
295c**	11,085,669	74	24.3	1.6	46.6	17.9
297c	9,161,075	77.3	16.6	5.9	42	13.4
300c	7,827,316	68.6	27.9	3.3	41.6	21.6
301c	8,019,941	74.6	23.4	1.8	49.2	17
302c	11,305,551	73.7	23.2	2.3	49.7	16
305c	5,948,259	81.8	14.8	3.1	46	10.1
308c	6,062,547	82.1	14.9	2.8	52.1	10.2
1FA**	7,600,015	72.9	18.6	8.2	38.9	13
2FA	7,663,658	71.9	25.5	2.3	41.3	19.5
3FA	8,020,970	69.8	25.9	4	42.7	20.1
7FA	8,302,694	78	20.6	1.1	43.6	15.4
8FA	5,601,314	72.4	25.2	2.2	49.8	18.3
10FA	7,573,448	68.7	25.3	5.8	44.2	19.5
13FA	4,996,908	79.1	18.9	1.8	50.3	13.1

* Including 5'- and 3'-untranslated regions (5'- and 3'-UTR); ** Hereinafter, the indices c and FA opposite the sample ID number denote the control group and the group of patients with fibroadenoma, respectively.

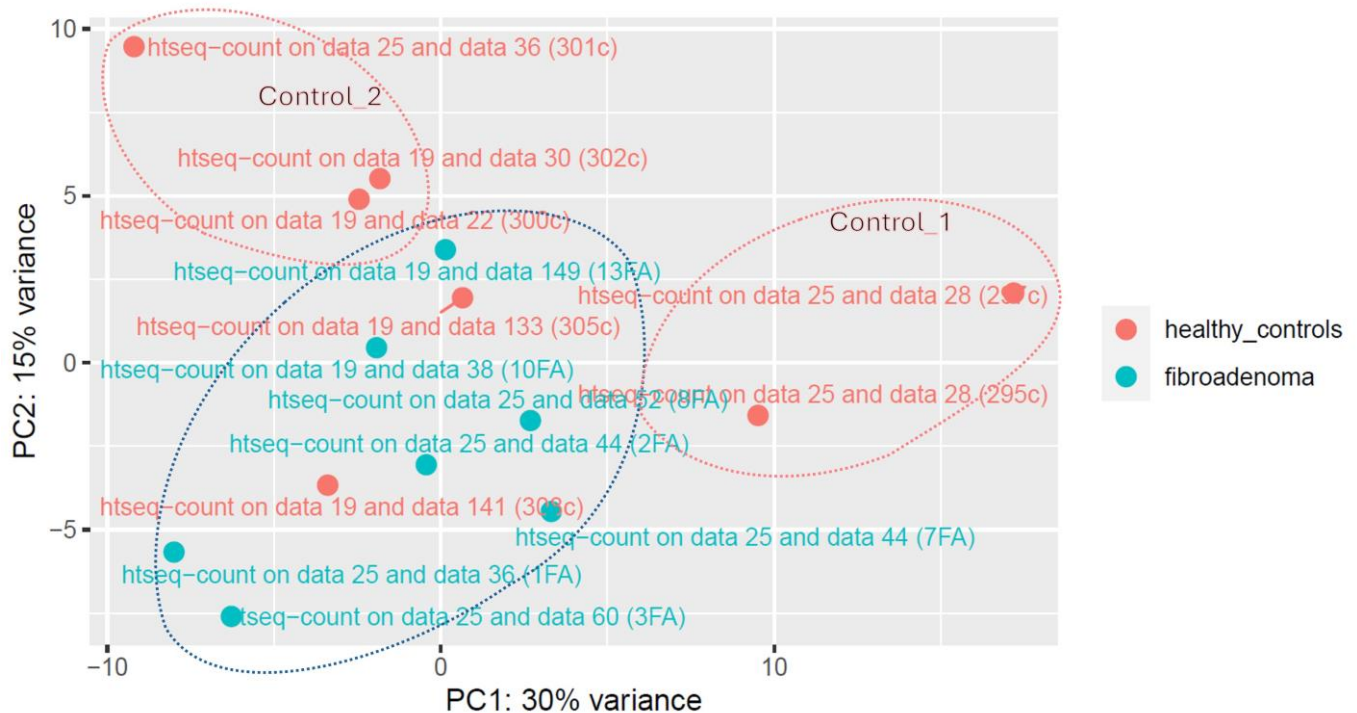


Figure 1. Principal component analysis (PCA) of global RNA expression profiles of peripheral blood mononuclear cells (PBMCs) obtained from women with fibroadenoma (FA) and healthy controls (c). The two-dimensional PCA scatterplot shows the distribution of sample transcriptomic profiles along two principal components (PC1 and PC2), which together explain 45% of the data variability; each dot represents one sample, the color of which corresponds to one of the two study groups. As can be seen in the plot, the FA samples form a cluster (indicated by the blue dashed line) that is variably separated from the transcriptomic datasets of the two control clusters (indicated by the red lines and labeled Control_1 and Control_2; the two samples are grouped together with the FA samples).

Table 2. Characteristics of fluorescently labeled monoclonal antibodies used for phenotyping peripheral blood lymphocytes

Antigen	Fluorophore	Clone	Manufacturer
CD3	VioBlue	REA613	Miltenyi Biotec, Germany
CD4	FITC	RPA-T4	
CD8	FITC	SK1	
PD-1/CD279	PE	EH12.2H7	
PD-L1/CD274/B7-H1	PerCP-Cy5.5	29E.2A3	BioLegend, USA
TIM-3/CD366	APC	F38-2E2	
LAG-3/CD223	APC	7H2C65	

Quantification of relative transcript abundance (gene counts) was performed using the HTSeq package. DESeq2 was then applied to the gene count matrix to identify genes differentially expressed between PBMCs from FA patients and healthy controls, and to obtain normalized expression values for the entire panel of samples. Genes that exhibited $|\log_2FC|$ values >1.0 , where FC is the fold change in gene expression and the adjusted p -value ($adj\ p\text{-val}$) $<$ the threshold, were considered significant differentially expressed genes (DEGs); the choice of significance threshold (0.05 or 0.25) is specified in the Results section. One of the DESeq2 outputs (a PCA scatterplot illustrating the results of a bivariate principal component analysis) was used to assess the degree of similarity/dissimilarity in global gene expression profiles between samples, as well as to identify datasets for comparison and identify DEGs or differentially activated signaling pathways between them. To better visualize differences in expression patterns between the compared sample groups, heatmaps were constructed using unsupervised hierarchical clustering based on Z-score normalized

gene expression values, as well as Volcano plots using Galaxy's built-in tools (heatmap2). Ensembl gene IDs were similarly converted to corresponding gene symbols using Galaxy's built-in function for annotating DESeq2 output tables.

Enrichment analysis of DEGs by the Gene Ontology (GO) terms categorized into three categories – Biological Process (BP), Molecular Function (MF), and Cellular Component (CC) – as well as by KEGG terms was performed in the Goseq package using the Benjamini-Hochberg correction method to control for the false discovery rate (FDR) in multiple comparisons. Terms with an FDR-corrected p -value <0.05 or <0.25 were considered significantly enriched for various comparison groups, as described in the Results section. To study the differences in the spectrum of activated signaling pathways for DEGs, gene set enrichment analysis (GSEA) was applied using the built-in FGSEA (fast preranked GSEA) package (with a minimum gene set size of 20, a maximum gene set size of 500, and a number of permutations of 1,000). The following gene sets were downloaded from the curated collections of MSigDB (Molecular Signature Database, <https://www.gsea-msigdb.org/gsea/msigdb>): “C2 subcollection CGP – chemical and genetic perturbations”, “C4 cancer-oriented genes”, “C6 oncogenic signature gene sets”, “C7 immunologic signature gene sets (subcollection ImmuneSigDB)”, and “H collection (hallmark gene sets)”. Before GSEA analysis, all genes in the DESeq2 output table were ranked by the sign of their FC and statistical significance values according to the formula: $(-\log_{10}(adj\ p\text{-val})) \times \text{sign}(FC)$. GSEA scores were considered statistically significant at $FDR < 0.05$.

Table 3. Results of Gene Ontology (GO) functional enrichment analysis for differentially expressed genes (DEGs) between the PBMC-FA and Control_1 sample datasets; showing the top 10 significantly enriched terms (ranked by enrichment score) in the Molecular Function and Cellular Component categories (FDR: false discovery rate)

FDR	Number of DEGs	Total number of genes	Fold enrichment	GO term
Category: Molecular Function				
1.3E-08	6	8	53.6	MHC class Ib protein complex binding
9.7E-03	2	3	47.7	Arachidonate 5-lipoxygenase activity
1.7E-04	4	9	31.8	IgG binding
1.1E-11	13	47	19.8	Inhibitory MHC class I receptor activity
8.4E-03	3	11	19.5	Complement receptor activity
9.4E-05	6	28	15.3	Pattern recognition receptor activity
5.2E-03	4	22	13	Sialic acid binding
4.1E-05	8	54	10.6	MHC class II receptor activity
7.2E-04	6	42	10.2	Protein phosphatase 1 binding
8.0E-04	6	43	10	SH2 domain binding
Category: Cellular Component				
1.6E-11	16	100	11.4	MHC class II protein complex
1.7E-08	12	81	10.6	Ficolin-1-rich granule membrane
1.2E-09	15	114	9.4	Clathrin-coated endocytic vesicle membrane
6.7E-09	14	110	9.1	Integral component of endoplasmic reticulum
3.5E-08	13	106	8.8	Tertiary granule membrane
7.6E-11	18	149	8.6	Trans-Golgi network membrane
8.7E-19	36	358	7.2	Secretory granule membrane
2.6E-08	16	169	6.8	COPII-coated ER to Golgi transport vesicle
1.6E-11	25	294	6.1	Endocytic vesicle membrane
3.7E-16	47	757	4.4	Side of membrane

Table 4. The set of most significantly enriched GO terms from the Molecular Function and Cellular Component functional categories in differentially expressed genes (DEGs) between the PBMC-FA and Control_2 datasets; (FDR: false discovery rate)

Category	Number of DEGs	Total number of genes	Term	FDR
Category: Molecular Function				
GO:0003823	15	159	Antigen binding	0
GO:0034987	6	69	Immunoglobulin receptor binding	5.51e-06
GO:0005125	7	236	Cytokine activity	0.000438
GO:0005126	7	273	Cytokine receptor binding	0.001045
GO:0008009	4	48	Chemokine activity	0.001670
GO:0042379	4	70	Chemokine receptor binding	0.006597
GO:0030546	7	495	Signaling receptor activator activity	0.026178
GO:0030545	7	530	Signaling receptor regulator activity	0.036645
Category: Cellular Component				
GO:0019814	17	150	Immunoglobulin complex	0
GO:0042571	6	66	Immunoglobulin complex, circulating	8.05e-07
GO:0072562	6	145	Blood microparticle	0.000159
GO:0005886	26	5831	Plasma membrane	0.002874
GO:0071944	27	6312	Cell periphery	0.003082
GO:0009897	7	449	External side of plasma membrane	0.007701

Flow cytometry

Venous blood samples were collected in special vacuum tubes containing 3.8% sodium citrate immediately before surgery. Erythrocytes were removed using osmotic shock by treating whole blood with ammonium chloride buffer containing 0.15 M NH₄Cl, 0.01 M KHCO₃, and 0.0001 M Na₂EDTA (RBC lysis buffer). Leukocytes were then washed and resuspended in phosphate-buffered saline. To analyze the expression of CD markers on the cell surface, the latter were treated with fluorochrome-conjugated monoclonal antibodies, the characteristics of which are presented in [Table 2](#). The optimal working dilution was determined empirically for each antibody clone. The incubation time was 30 min at ambient temperature in the dark according to the manufacturer's specifications. Measurements were performed using a MACSQuant flow cytometer (Miltenyi Biotec, Germany). The data were further analyzed using MACSQuantify

v.2.11 software (Miltenyi Biotec, USA). Statistical analysis of the flow cytometry results, and construction of violin plots were performed using the RStudio environment (version R-4.3.2). The lymphocyte population was gated based on forward and side scatter (FSC/SSC) parameters. To more accurately determine the gating boundaries for positively/negatively stained cell populations, 'fluorescence minus one' controls were prepared for each of the fluorophores used. In the case of PD-L1 marker expression, in addition to the percentage of positively stained cells, the level of PDL1 antibody binding was calculated as the median increase in fluorescence intensity (Δ MFI) relative to the autofluorescence level of unlabeled cells in the corresponding detector channel. At least 100,000 events were measured for each probe. The statistical significance of the differences between the FA group and the control group was determined by nonparametric Wilcoxon-Mann-Whitney test (U test; $p < 0.05$ was considered statistically significant).

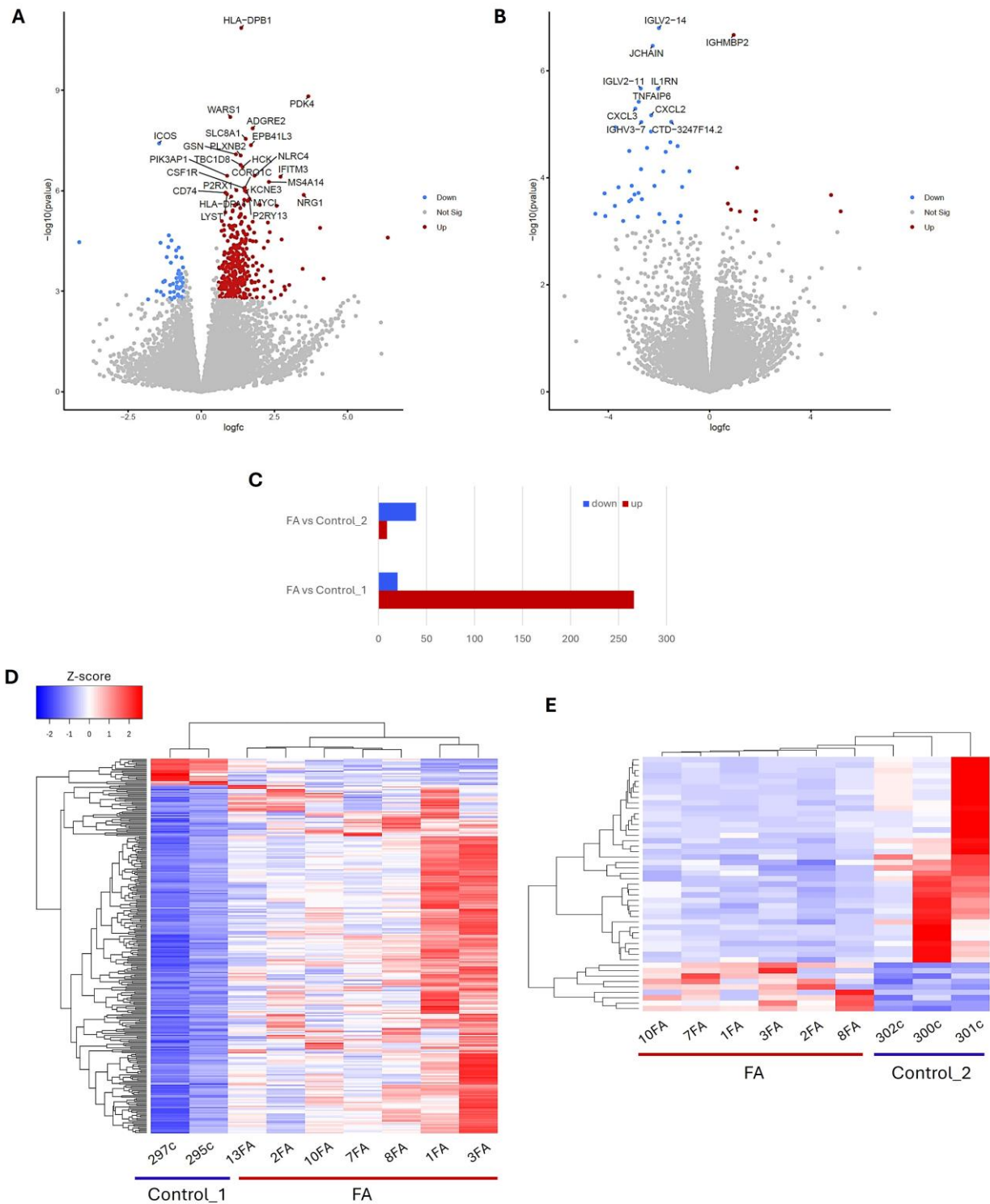


Figure 2. Analysis of differentially expressed genes (DEGs) between the transcriptomes of peripheral blood mononuclear cells (PBMCs) from women with fibroadenoma (FA) and the control group. (A) and (B) Volcano plot showing the distribution of differentially expressed genes (DEGs) according to statistical significance and magnitude of change. The y-axis represents the $-\log_{10}$ transformed adjusted p -values ($adj\ p\text{-val}$); the x-axis represents the $\log_2(\text{FC})$ values, where FC denotes the fold change of expression obtained by comparing the FA datasets to the Control_1 subset (A) or to Control_2 (B). Individual genes are depicted as dots colored according to the significance threshold and FC of expression threshold: red dots indicate DEGs that showed significantly upregulated expression in PBMCs of the FA patient group (Up); blue dots indicate genes with downregulated expression, and gray dots represent non-DEGs. The plot also shows the gene names for the 25 (A) and 10 (B) DEGs that showed the largest expression change in the FA datasets compared to Control_1 and Control_2, respectively. (C) Total numbers of up- and downregulated DEGs identified via comparing the PBMC sample groups: FA vs. Control_1 and FA vs. Control_2; the full list of genes is available upon request. (D) and (E) Results of hierarchical cluster analysis of DEGs identified between the FA and control datasets; heatmap was generated based on the standard score (or Z-score) values; red color corresponds to upregulated genes, while blue color indicates downregulated genes.

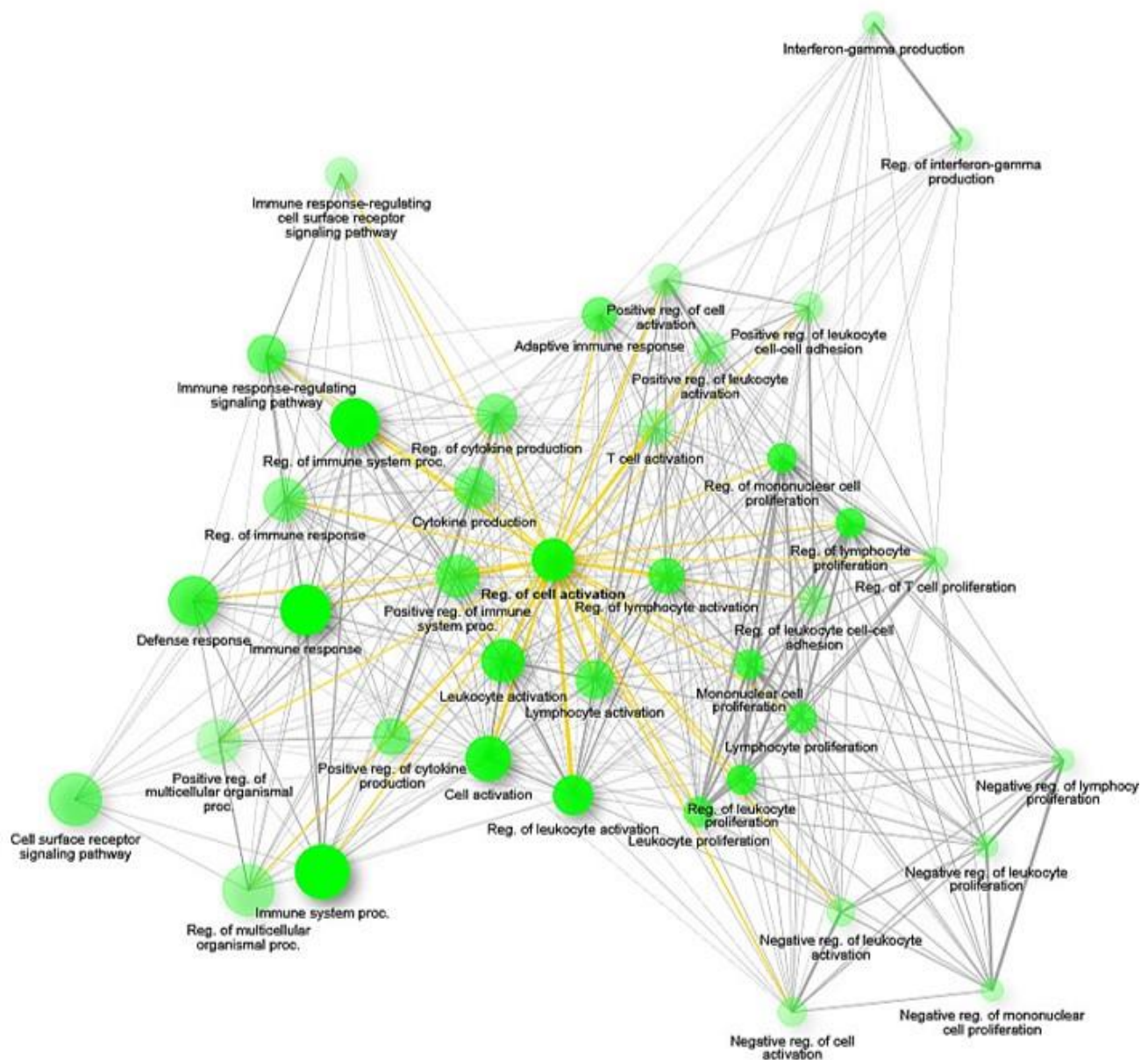


Figure 3. Network diagram illustrating the results of Gene Ontology (GO) functional enrichment analysis of differentially expressed genes between the PBMC-FA and Control_1 datasets; significantly enriched terms from the Biological Process (BP) category were visualized using the ShinyGO online tool [24]. To create the diagram, GO-BP terms meeting the FDR<0.01 criterion were sorted by fold enrichment, and the 40 most represented terms were defined (the full list of terms is available upon request). Connecting lines indicate the degree of relatedness between terms (>20% overlapping genes), and their thickness indicates the proportion of shared genes. The brightness of the nodes indicates the level of statistical significance; the size of the node indicates the number of genes annotated with a specific term.

Table 5. The KEGG signaling pathway enrichment analysis of the differentially expressed genes (DEGs) in peripheral blood mononuclear cells (FA: fibroadenoma; FDR: false discovery rate)

Category	Signaling pathway	FDR	Category	Signaling pathway	FDR
	FA vs. Control_1			FA vs. Control_2	
4380	Osteoclast differentiation	7.98e-07	4060	Cytokine-cytokine receptor interaction	2.0637e-06
4145	Phagosome	0.0129	5323	Rheumatoid arthritis	0.0006
4142	Lysosome	0.0136	4062	Chemokine signaling pathway	0.0074
5140	Leishmaniasis	0.0136	5140	Leishmaniasis	0.0074
5150	Staphylococcus aureus infection	0.0324	5332	Graft-versus-host disease	0.0571
5310	Asthma	0.0362	4940	Type I diabetes mellitus	0.0594
5323	Rheumatoid arthritis	0.0668	4621	NOD-like receptor signaling pathway	0.1152
5330	Allograft rejection	0.1059	4010	MAPK signaling pathway	0.1733
4672	Intestinal immune network for IgA production	0.1085	4640	Hematopoietic cell lineage	0.1733
4612	Antigen processing and presentation	0.2006	4210	Apoptosis	0.1733

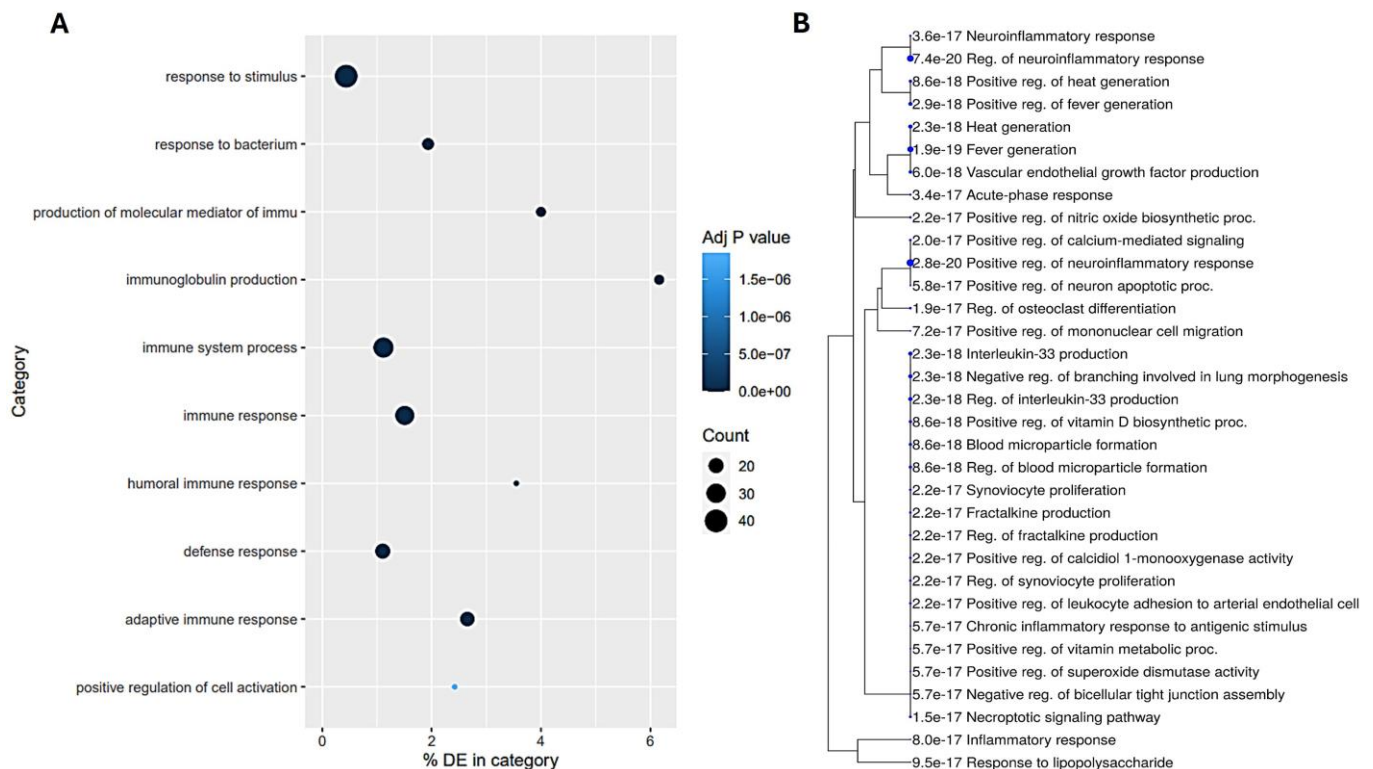


Figure 4. Gene Ontology (GO) functional enrichment analysis of differentially expressed genes between the PBMC-FA and Control_2 datasets, with terms belonging to the Biological Process (BP) sub-ontology. (A) The bubble chart shows the 10 terms with the minimum *adj p*-val (FDR). (B) The hierarchical clustering dendrogram demonstrates the relationships between statistically significant terms: those sharing a higher number of genes in common are grouped together; the size of the blue dots at the branch ends is determined by the significance level (*adj p*-val) indicated alongside. To construct the dendrogram, GO-BP terms satisfying the FDR<0.01 criterion were sorted by fold enrichment, and a set of the 40 most significantly enriched terms is shown (the full list of terms is available upon request). The ShinyGO online tool [24] was employed to visualize the GO analysis results.

Results

Transcriptome analysis

As a first step, complete transcriptome profiles of PBMCs isolated from seven FA patient samples (PBMC-FA) and seven control samples were analyzed using principal component analysis (PCA) to assess the distance between sequenced datasets and identify groups of samples for comparison. The similarity/dissimilarity of global gene expression profiles is illustrated by the PCA plot in [Figure 1](#). As can be seen in the plot, the FA datasets form a cluster separated from the two control datasets (Control_1) along the first principal component axis (PC1) and from the three control datasets (Control_2) along the second principal component axis (PC2). Pairwise comparisons of the resulting clusters were then performed, and genes differentially expressed in the FA patient PBMCs were identified using DESeq2 analysis.

Comparative transcriptome analysis of PBMC-FA vs. Control_1 identified 286 DEGs meeting the significance criteria of *adj p*-val<0.05 and $|\log_2FC|>1$, most of which (266) were upregulated in PBMC-FA group compared to Control group, while only 20 DEGs were significantly downregulated in FA datasets ([Figure 2A, C](#)). Comparison of PBMC-FA datasets with Control_2 resulted in the identification of 48 differentially expressed genes (with *adj p*-val<0.25 and $|FC|>1.5$), most of which (39 DEGs) were downregulated in PBMC-FA, while 9 DEGs were found upregulated at the given significance level ([Figure 2B, C](#)). Further hierarchical

clustering analysis of DEGs confirmed a clear division of the datasets into subgroups of FA patients and controls ([Figure 2D, E](#)). The resulting heatmaps indicate, on the one hand, a certain degree of heterogeneity in the PBMC transcriptomic profiles across all compared groups, and on the other hand, the existence of distinct expression patterns of up- and downregulated genes.

We also noted that the list of upregulated genes in the FA datasets included, on the one hand, a large number of regulators of inflammatory responses and modulators of innate immune response mechanisms, including those mediated by natural killer (NK) cells, macrophages, and dendritic cells. On the other hand, some of the upregulated genes indicate possible dysfunction in a T cell subset (e.g., activation of apoptosis, response to glucocorticoids and oxidative stress, differentiation of Treg lymphocytes). Meanwhile, among the genes showing reduced expression in the FA datasets, we identified B cell markers and various stimulators of T cell function (such as activation, expansion/proliferation, differentiation, or migration), as well as negative modulators of inflammation. These data suggest that abnormalities in the cellular immune compartment, presumably accompanying FA development, may arise from various molecular mechanisms underlying its pathogenesis. To elucidate which biological functions may be affected by the pathological process, we additionally performed GO and KEGG functional enrichment analysis for DEGs, followed by comparative signaling pathway activation pattern analysis via GSEA method.

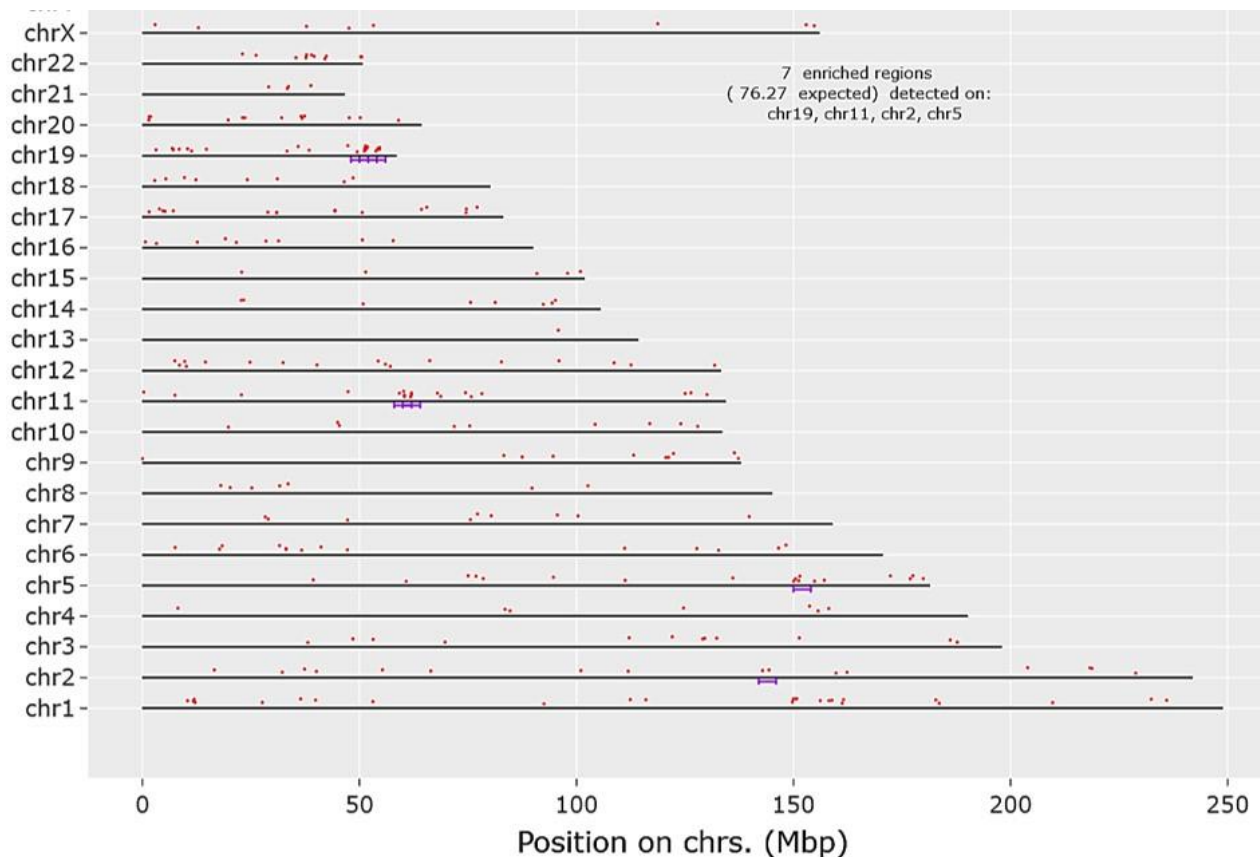


Figure 5. Distribution of differentially expressed genes (DEGs) across chromosomal loci. DEG positions are marked with red dots. Purple segments indicate genomic regions that are statistically significantly enriched in consistently overexpressed DEGs compared to the average gene density in the human genome. Genome screening was performed using the sliding window technique; the window size was set to 4 Mb, and the number of steps in each window was 2. The ShinyGO online tool [24] was employed to visualize the scanning results.

According to the GSeq analysis of genes differentially expressed between the PBMC-FA and Control_1 datasets, 486 BP GO terms (with $FDR < 0.01$), 24 MF GO terms ($FDR < 0.05$), 68 CC GO terms ($FDR < 0.01$), and 16 KEGG signaling pathways ($FDR < 0.25$) were overrepresented in the PBMC-FA datasets. Given the hierarchical modular architecture of the GO dictionary and its redundancy (overlapping synonymous terms), the ShinyGO v.0.81 graphical tool (<https://bioinformatics.sdstate.edu/go> [24]) was used to reduce the dimensionality of the GO-BP enrichment data, visualize it, and highlight the most representative terms. [Figure 3](#) shows a large cluster of terms covering processes regulating T cell activation and proliferation (primarily Th1), monocyte maturation, and inflammatory responses. A closer inspection revealed that these processes were enriched primarily in IFN- γ and IL-12 signaling pathways, while enrichment in IL-1, IL-6, IL-8, and TNF α -dependent pathways indicates activation of inflammatory responses. Another cluster of BP terms represents the processes of mononuclear cell adhesion and directed cell migration. Notably, downregulated genes (representing the minority of DEGs) were associated with negative regulation of mononuclear cell migration, as well as with processes of T cell selection, activation, and differentiation, including IL-4 and IL-10-dependent mechanisms, apoptosis, and Treg functions. Another group of overrepresented terms was associated with various functions of antigen-presenting cell populations (macrophages, dendritic cells), including phagocytosis, cytokine production, reactive oxygen/nitrogen generation, and pattern recognition receptor signaling. Key

enrichment results for the MF and CC categories are presented in [Table 3](#). The protein expression products of DEGs involved in the above-mentioned processes are primarily components of the major histocompatibility complex (MHC) and members of signaling cascades triggered by innate immune receptors, antigen recognition receptors, and cytokine receptors, including signaling kinases/phosphatases and adapter proteins/signal transducers. Their subcellular localization relates primarily to cellular structures involved in secretion or, conversely, receptor-mediated endocytosis, or to membrane-bound receptor complexes, as well as to the membrane structures responsible for cell adhesion/movement ([Table 3](#)).

Comparison of the PBMC-FA and Control_2 datasets using GSeq revealed functional enrichment for 81 GO terms of the BP category ($FDR < 0.01$), 10 GO terms of the MF category ($FDR < 0.05$), and 8 GO terms of the CC category ($FDR < 0.01$), as well as for 12 KEGG signaling pathways ($FDR < 0.25$). The most statistically significantly enriched GO BP categories were humoral regulatory mechanisms mediated by immunoglobulins or some other soluble factors of innate immunity ([Figure 4A, B](#)). In terms of fold enrichment score, the Th2-like inflammatory profile, pyrogenic/thermogenic responses, and the associated mechanisms of NO generation, chemotaxis, and phagocytosis were the most over-represented. Another group of terms reflects the general cellular response to stress, including apoptotic processes and antioxidant defense responses, which is interesting since among DEGs, we found the glucocorticoid-inducible gene encoding

the zinc-binding protein metallothionein MT1E, as well as the SOD1 gene, the role of which in immune responses is becoming the subject of active research [25]. The enrichment results for the MF category suggest that the DEGs involved in the above-mentioned processes are components of cytokine/chemokine or immunoglobulin receptor complexes (Table 4). KEGG pathway analysis reveals that different functional activities of antigen-presenting cells (macrophages), cytotoxic T/NK cell responses to auto-/alloantigens, pattern recognition receptor-mediated inflammatory responses, and a spectrum of specific cytokines/chemokines may also contribute to the transcriptional landscape of PBMC-FA (Table 5).

We also investigated whether the identified DEGs could be associated with specific chromosomal regions (or genomic loci). As

shown in Figure 5, highly enriched regions were located on chromosomes 2, 5, 11, and 19. This is expected given that chromosome 19, for example, contains a large collection of immunoglobulin superfamily genes, which are expressed in leukocytes and encode receptors that recognize different classes of antigenic molecules or molecular patterns, as well as genes encoding cytochrome and serine protease subfamilies involved in inflammatory responses [26]. Chromosome 5 contains the interleukin gene cluster encoding several hematopoietic cytokines [27], and chromosome 2 includes the IL-1 gene family, which are inflammatory mediators [28]. Thus, it can be concluded that the distribution of genes differentially expressed in the PBMCs of patients with FA is non-random.

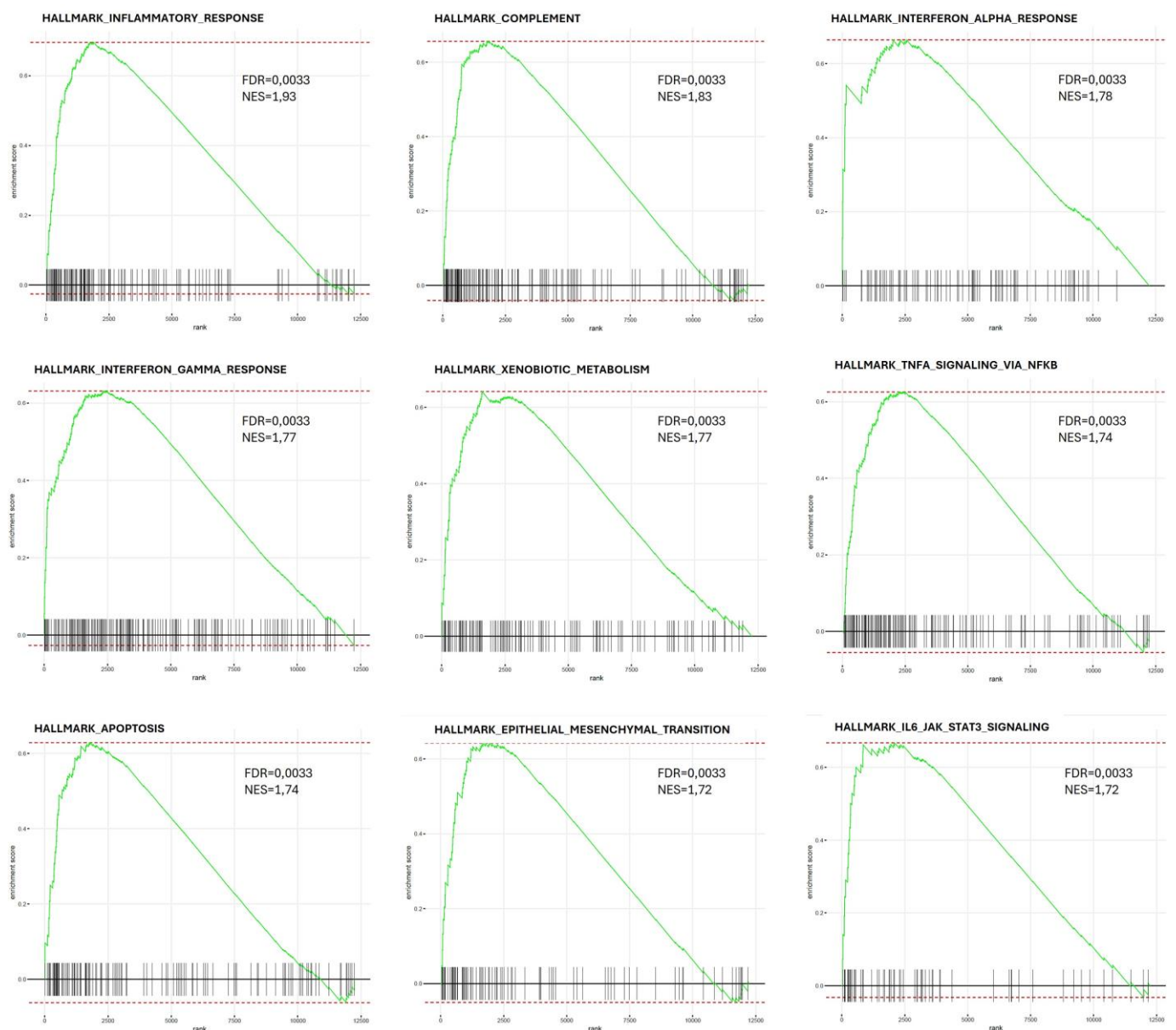


Figure 6. Plotted results of gene set enrichment analysis (GSEA) comparing the transcriptomic profiles of PMBC-FA vs. the control group (Control_1) using gene sets obtained from the Hallmark MSigBD collection. Nine upregulated gene sets are shown, ranked by the highest enrichment score, normalized enrichment score (NES), and the lowest false discovery rate (FDR). The black lines at the bottom of each plot correspond to the position of the given gene set in the ranked list (ranking was performed as described in the Materials and Methods section); the green lines represent the enrichment score (y-axis) for each position in the ranked list.

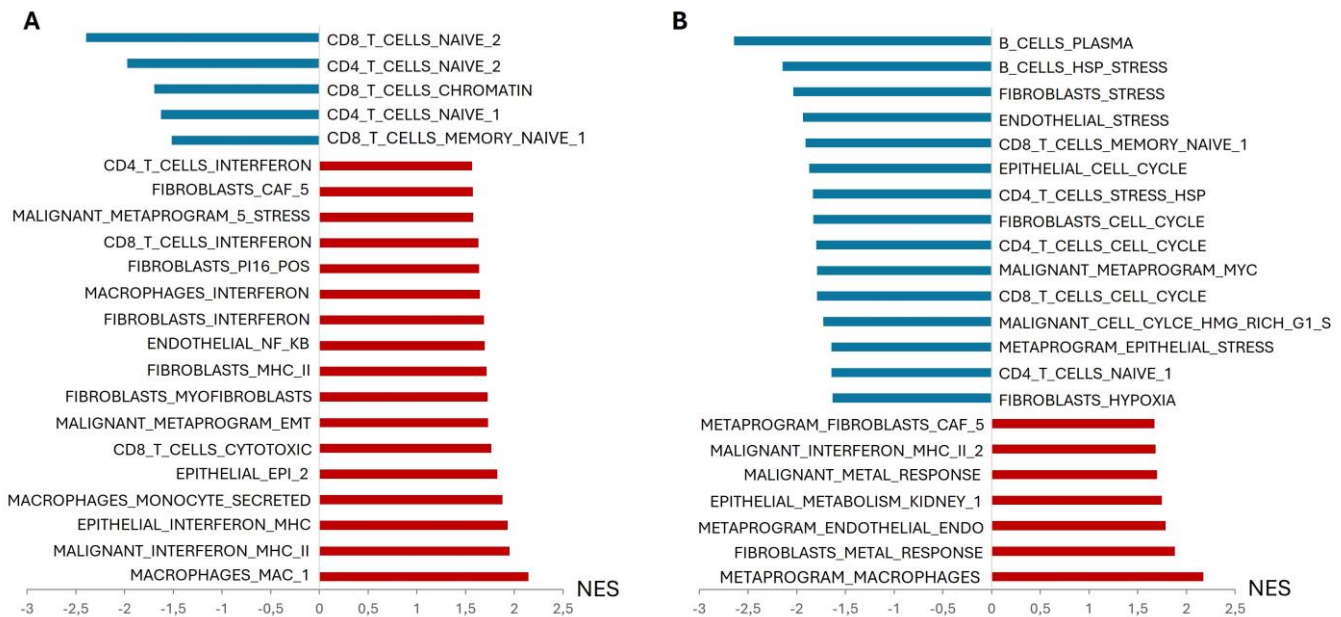


Figure 7. Gene set enrichment analysis (GSEA) results for gene sets obtained from the MSigDB C4 collection (3CA: Curated Cancer Cell Atlas and CGN: Cancer gene neighborhoods). The PBMC-FA transcriptome was compared with the Control_1 (A) or Control_2 (B) transcriptome, false discovery rate (FDR)<0.05; the horizontal axis shows the normalized enrichment score (NES).

Since GO analysis can only extract information about the functions of overrepresented genes but does not allow us to infer the contribution of entire signaling pathways to the observed differences, we performed GSEA of the identified PBMC clusters using gene sets curated by the molecular signature database (MSigDB) [29] to better understand the biological processes that may be involved in the development of FA at the level of cellular immunity. Of primary interest was, of course, the ImmuneSigDB collection of immune-associated gene signatures (subset C7) [30]. Among these gene sets, the following pathways were most enriched in PBMC-FA samples vs. Control_1 (*adj p-val*/FDR=0.006, NES>1.8, where NES is the normalized enrichment score): (1) innate proinflammatory pathways activated in monocytes/macrophages and dendritic cells in response to stimulation of pattern recognition receptors specific for exo/endogenous ligands (such as NOD2, TLR2, TLR1, TLR4/CD14, TLR7/TLR8, TLR9, TREM1 and lectins) and mediated by MYD88, DUSP1/p38, WPB7/MLL4, CSF2RA and other regulators; (2) pathways responsible for the suppression of Treg differentiation (including LEF1/TCF10-mediated pathways); (3) signaling networks controlling IL-6/STAT3-dependent Th17 cell differentiation, SAP (SH2D1A)-dependent Tfh cell differentiation, and CD4 effector and memory T cell differentiation (Teff_mem, T_cent_mem) in general; (4) pathways associated with inhibition of CD8 T cell activation (downstream of type I and type II interferons, CD3/CD28 costimulatory receptors, IL-2, IL-21, STAT4, and STAT5), CD8 T cell energy, and CD8 memory T cell formation. The signaling pathways identified by comparing PBMC-FA with Control_2 (FDR<0.2, |NES|>1.4) are likely alternative pathways that can cause inflammatory immune dysregulation. These include: (1) PPARG- and STAT6-dependent mechanisms that trigger the polarization of M2 macrophages (exhibiting profibrotic and proliferative activity) [31] and activate the type 2 immune response in dendritic cells (a recognized factor in persistent inflammation) [32]; (2) mechanisms

that control B cell activation and the induction of T cell energy (mediated by IRF4 and NFATC1); and (3) mechanisms that induce the differentiation of or Tregs (including adipose tissue-resident Tregs) and are controlled by PPARG, DNMT1, FOXP3, and STAT5 [33].

Of particular interest are the differences found in CD8 subsets characterized by the PD1^{low}/PD1^{high} phenotype. PD1 expression status (low/high) was shown to correlate with the state of T cell activation/exhaustion [34]. In line with this, GSEA data indicate a general upregulation of PD-1-dependent processes and a shift in the PD-1 status of CD8 T cells during FA development. Further evidence is provided by analysis of the Hallmark (H) collection of key biological processes and physiological states [29]: among the most enriched gene sets, inflammation and T cell activation processes involving interferons, TNF α , and IL6, as well as induction of apoptosis, showed the highest enrichment score (with NES>1.7 and FDR=0.033) (Figure 6).

Given that the immune microenvironment is an integral part of malignant tumors, a similar role for the immune compartment can be hypothesized in benign tumors (such as FA). Although our transcriptome analysis was limited to peripheral blood and did not include the tumor compartment, we decided to investigate whether the transcriptomic differences identified in PBMCs could be linked to any cancer-related metaprograms or gene networks using the C4 collections from MSigDB (3CA: Curated Cancer Cell Atlas and CGN: cancer gene neighborhoods). When comparing the transcriptional profiles of the PBMC-FA and Control_1 datasets using tumor-specific metaprograms, the highest enrichment score (|NES|>1.5, FDR<0.05) was obtained for inflammatory signaling pathways initiated by macrophages through interferon secretion and upregulation of proteasomal degradation and antigen presentation mechanisms (Figure 7A), which is consistent with the results of the C7 ImmuneSig and Hallmark analyses. It is noteworthy that metaprograms specific to epithelial cells and

fibroblasts (in particular, myofibroblasts) were also identified, as well as an EMT metaprogram, which is most likely directly related to the pathology under study, viz., breast fibroepithelial tumors. It is also interesting that FGSEA revealed an increase in interferon-associated antigen presentation activity in epithelial and stromal cells. Furthermore, FGSEA revealed a positive gene signature of PI16, which is recognized as a marker of a proinflammatory fibroblast population capable of attracting immune cells [35] and a negative modulator of Treg differentiation [36]. Enrichment of the signature of cancer-associated fibroblasts is also noteworthy. Only naïve CD4 and CD8 T cell-specific gene sets were negatively enriched (Figure 7A). The enrichment profile of the 3CA gene pathways obtained by comparing the PBMC-FA transcriptome with Control_2 largely overlapped with that when comparing PBMC-FA with Control_1. However, some specific and interesting mechanisms were identified that deserve mentioning: (1) mechanisms underlying the impaired cellular response to stress; (2) mechanisms of cell cycle regulation, including those specific to

fibroblasts and epithelial cells; (3) signaling pathways directing the activation of the cellular response to metal ions mediated by class I metallothioneins (MT1) with Zn/Mg binding activity. The involvement of the latter gene family seems quite interesting given the fact that, in addition to their metabolic and antioxidant roles, MT1 are currently considered important regulators of immune responses (including the formation of an immunosuppressive environment) that can occur in response to a wide range of proinflammatory factors [25], as well as members of regulatory networks of carcinogenesis relevant to the biology of both benign and malignant breast tumors [37] (Figure 7B). Regarding gene clusters coregulated by any of the 380 tumor-associated genes in the CGN collection (MSigDB's C4 gene pool), FGSEA identified gene networks organized around key mediators of monocyte/macrophage-initiated inflammatory responses (CASP1, CARD15/NOD2, HCK, TNFRSF1B, S100A4, CD14, MYD88, TNFSF10/TRAIL, MCL1/BCL2L3, SPI1), some of which also act as regulators of apoptosis in lymphocytes/monocytes.

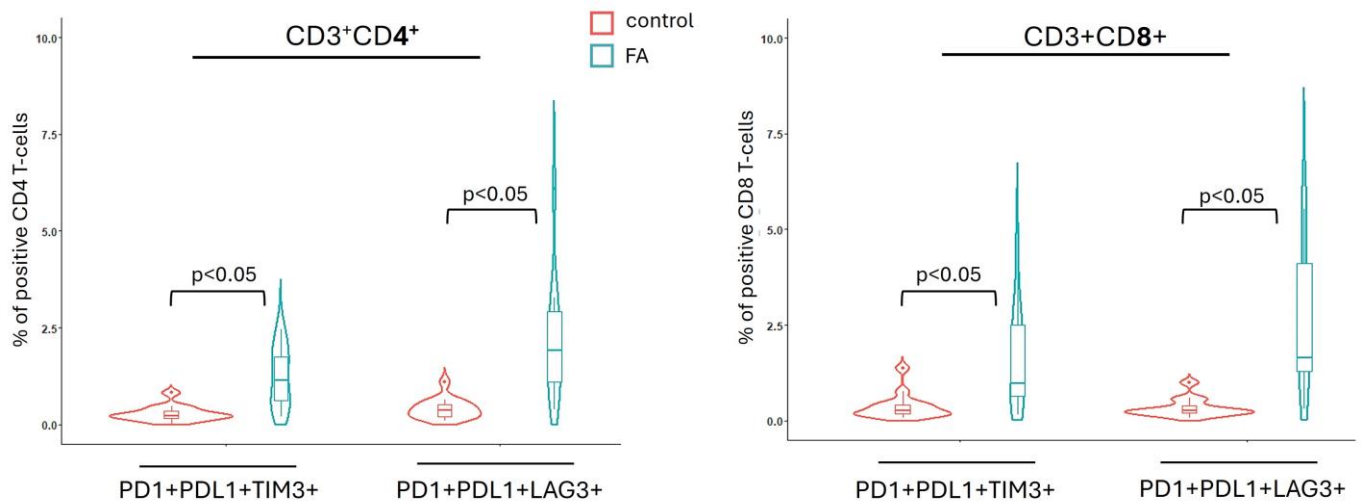


Figure 8. Co-expression levels of three immune checkpoint markers in peripheral blood T lymphocytes obtained from women with fibroadenoma (FA) and healthy controls, measured by flow cytometry. Hereinafter, the group median and interquartile range are shown in box and whisker plots overlaid onto violin plots.

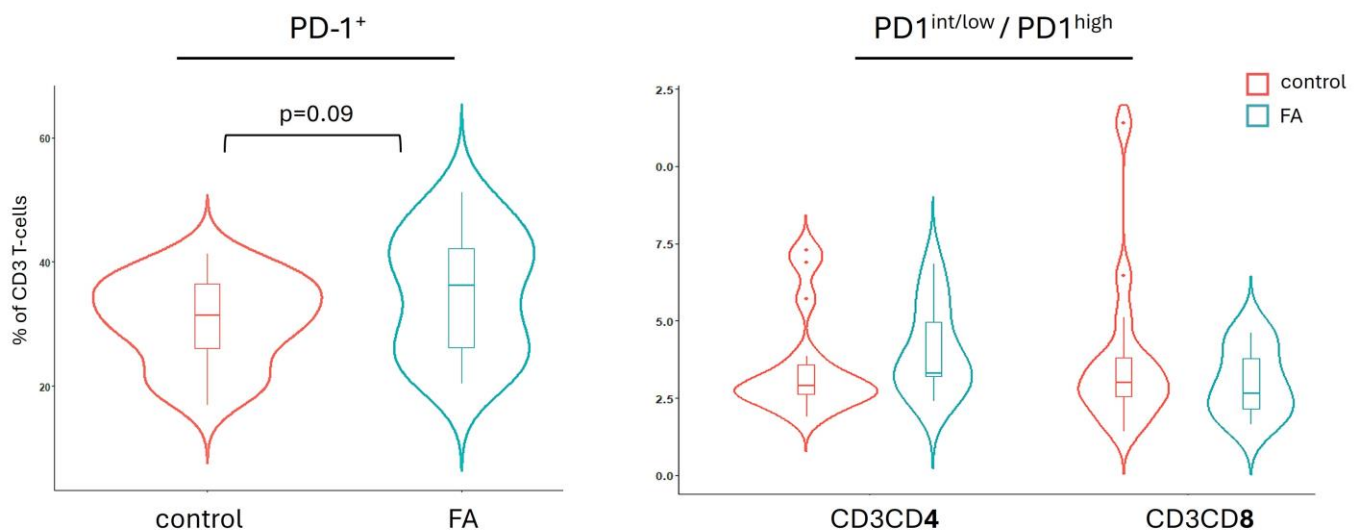


Figure 9. Expression levels of the PD-1 marker in peripheral blood T lymphocytes from women with fibroadenoma (FA) vs. the control (based on flow cytometry data).

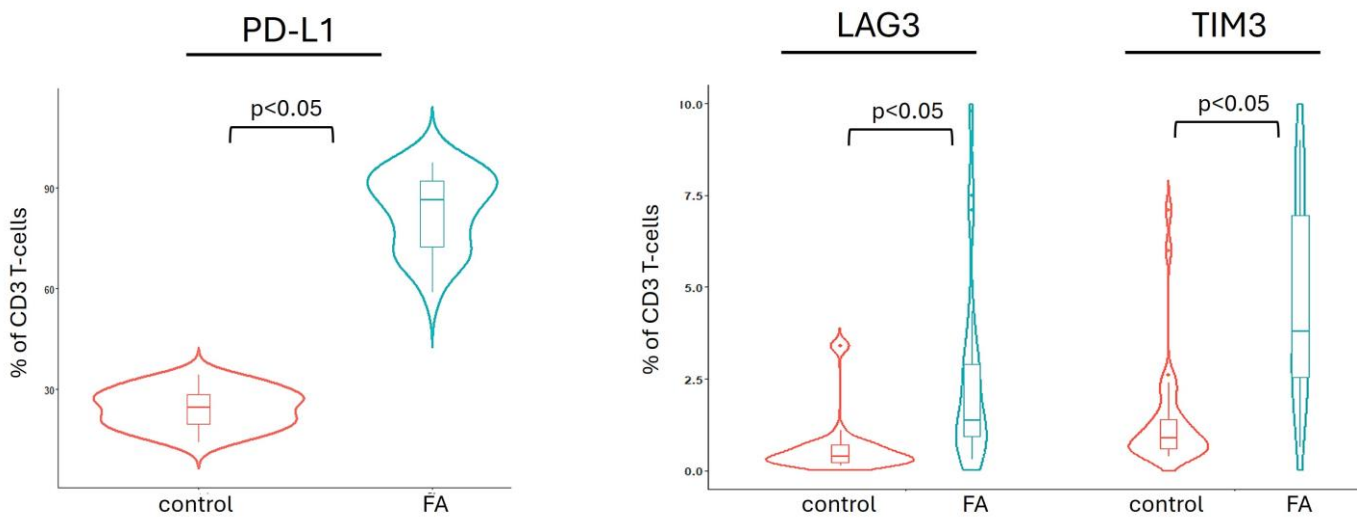


Figure 10. Changes in the expression levels of PD-L1, LAG3, and TIM3 markers in peripheral blood T lymphocytes from women with fibroadenoma (FA) vs. the control (based on flow cytometry data).

Given that, according to GO and GSEA, certain signs of T lymphocyte exhaustion were observed at the transcriptome level (probably caused by their activation), we decided to use flow cytometry for evaluating possible changes in the expression patterns of four markers of immune exhaustion: immune checkpoint molecules PD-1, PD-L1, TIM-3 and LAG-3 in CD4 and CD8 T cell subsets.

Expression of immune checkpoints in circulating lymphocytes (based on flow cytometry)

Under conditions of chronic stimulation (and, in particular, chronic inflammation), the simultaneous (co-)expression of several immune checkpoints is more important and representative than the expression of any of the immune checkpoint markers individually, since it more deeply characterizes the degree of lymphocyte exhaustion [38]. That is why we decided to test whether it was possible to detect changes in the number of T cells simultaneously positive for three key checkpoint markers (PD1+PD-L1+TIM3+ or PD1+PD-L1+LAG3+) in patients with FA. As can be seen in [Figure 8](#), in the group of patients with FA, there was an average increase in the fraction of cells with the specified phenotype, while no significant differences between the CD4 and CD8 subsets were detected. In the patient group, a relatively wide range of variations in cell counts was also noted, whereas in the control group, cells with the PD1+PD-L1+TIM3+ or PD1+PD-L1+LAG3+ phenotypes were virtually undetectable.

When analyzing the mono-expression of each of the selected markers in the total T lymphocyte population or in its CD4/CD8 subsets separately, a trend towards an increased share of PD-1+ T cells in the FA sample set was observed ([Figure 9](#)), although it did not reach statistical significance. Taking into account the results of the signaling pathway analysis described above, as well as the published data reporting heterogeneity of the PD-1+ T cell population in terms of PD-1 expression level (high, low/intermediate) [34, 39, 40], we decided to investigate whether a shift in the intensity of PD-1 expression could be observed in FA group vs. the control. CD4+PD-1+ or CD8+PD-1+ T cell populations were divided into PD-1^{high} and PD-1^{low/int} expression groups (high or low/intermediate expression, respectively), and the ratio of

%PD-1^{low/int} to %PD-1^{high} cells was measured ([Figure 9](#)). No statistically significant differences were found between the FA group and the control group, even though there was a clear trend towards an increased relative proportion of CD8+PD-1^{high} T cells. The latter were reported to exhibit more pronounced immune exhaustion [34, 39, 40]. At the same time, the results for the CD4 subset indicated a potential increase in the relative frequency of the PD-1^{low/int} phenotype.

When PD-L1 mono-expression was assessed, a significant increase in the percentage of PD-L1+ T cells was detected, as opposed to PD-1 expression: in fact, most T cells in the FA group were PD-L1-positive ([Figure 10](#)). However, it is important to point out that the fluorescent signal intensity was very weak, thereby implying a low overall cell surface expression density of this checkpoint marker in both study groups. TIM3 and LAG3 checkpoints also exhibited marked increases in expression levels, with a similar pattern of changes observed in the CD4 and CD8 subsets.

Hence, in our study, the results of transcriptome analysis and flow cytometry suggest that changes in the number of T cells (co-)expressing immune checkpoint markers that accompany the development of FA potentially reflect the disruptions in signaling pathways and gene expression patterns described above.

Discussion

Transcriptome analysis of the PBMC fraction is currently considered a promising approach to studying the pathogenetic mechanisms underlying the formation of benign and malignant breast lesions, helping to decipher the heterogeneity of this group of diseases and, consequently, expand diagnostic capabilities and treatment monitoring [41]. Transcriptomic studies of circulating blood cell populations can provide a deeper understanding of the mechanisms of the body's immune response to the onset and progression of malignant neoplastic processes at the systemic level. However, as already mentioned, although such studies are actively conducted in relation to breast cancer, information on benign breast pathologies is virtually nonexistent (experimental research data presented in the literature concern mainly malignant phyllodes tumors). In this study, we attempted to

characterize the changes in the transcriptomic profile of PBMCs obtained from patients with FA as a proof-of-concept study, and for the above reasons, we can compare our results exclusively with the data obtained for invasive breast cancer, of course, taking into account the fundamental differences between malignant neoplasms and benign proliferative lesions of the breast. For example, by sequencing the PBMC transcriptome, Ming et al. demonstrated the existence of two independent immune subtypes of breast cancer that do not overlap with the clinical classification of breast cancer, which is based on the immunohistochemical expression of certain receptors by tumor cells [42]. In some cases, the PBMC transcriptome in breast cancer shows a strong signature of inflammatory gene expression [41]. In our study, we did not aim to identify distinct FA immune subtypes and define their specific gene signatures; however, we were able to demonstrate major intrinsic heterogeneity in the PBMC-FA transcriptome compared to the transcriptomic profile of control PBMCs, which likely reflects different modes of interaction between FA tumors and the immune system. The consistency of these changes is supported by the fact that, following bioinformatic analysis, we were able to identify not only differentially expressed genes but also entire signaling pathways. In this regard, Ming et al. reported an interesting finding that the major transcriptome differences they found between the two immune subtypes of breast cancer PBMCs were primarily associated with cytokine-dependent activation mechanisms, in particular, TNF α -dependent monocyte/macrophage-mediated inflammatory pathways [42]. Several transcriptomic studies of peripheral blood mononuclear cells (PBMCs) in breast cancer revealed an important role of innate immune response regulators, such as Toll-like receptors (TLRs) [41], and reported the possibility of detecting an inhibitory T cell phenotype that can be defined by altered immune checkpoint-dependent signaling. Consistent with this, we also noted a significant impact of inflammation-related mechanisms and innate immune responses on the PBMC-FA transcriptome, along with evidence of inhibition of T cell functions.

Changes in the transcriptomic profile of PBMCs reflect changes in the phenotypic composition of circulating lymphocytes/monocytes. The expression of immune checkpoint molecules in different PBMC populations is currently of great interest, not only because of their potential as targets for cancer immunotherapy, but also because the analysis of immune checkpoint molecules allows assessing the balance between the states of lymphocyte activation and exhaustion as a natural mechanism for controlling the immune response. However, even for breast cancer, changes in immune checkpoint expression in PBMCs are not yet well understood [43], let alone for benign breast diseases. We found an increased frequency of circulating PD-L1⁺ T cells in a set of FA samples, which have a tolerogenic effect on tumor cells and an immunosuppressive effect on effector T cells [44]. Recent studies also showed that increased expression of the PD-L1 molecule on CD4 T cells may be a consequence of their activation upon stimulation by proinflammatory factors and may interfere with the cytotoxic T cell response [45]. In experimental (organoid) models of breast cancer, LAG3 was shown to have an immunosuppressive effect as well, in particular on CD8 T cells [46]. It can be speculated that this effect may be extrapolated to the case of FA, given the fact that we found not only an increase in LAG3 expression but also an increase in its co-expression with PD-1/L1, which may be an indicator of an exhausted T cell phenotype (as in the case of triple-negative breast

cancer [47]). Similar observations were made for TIM3 expression levels. It is widely believed that while the expression of a single immune checkpoint molecule is a marker of transient T cell activation, co-expression of multiple checkpoints is a sign of T cell exhaustion, which progresses in the context of chronic inflammation [48]. Indeed, according to our flow cytometry results, peripheral blood immune cells from patients with FA exhibited an increased frequency of simultaneous expression of three checkpoint molecules on CD4 and CD8 T cells and also displayed features of chronic inflammation at the PBMC transcriptome level.

Unlike malignant tumors, benign breast lesions do not pose a serious clinical problem and are not directly associated with immune dysfunction. Nevertheless, studying the immune profile (including the transcriptome landscape) of women diagnosed with FA may provide insight into impairment of putative immunoendocrine pathways which, in turn, may lead to various disruptions in the mechanisms that control epithelial and stromal cell proliferation, tissue architecture, and the normal processes of development and involution of glandular breast tissue. As follows from our results, transcriptomic and phenotypic changes in peripheral blood lymphocytes/monocytes in women with FA coincide to a certain extent with oncogenesis-associated features and metaprograms of cancer. Therefore, further research on this issue will expand our understanding of oncogenic risk factors for early cancer detection and help in the development of differentiated approaches to the clinical management of FA per se.

Limitations of the study

The authors fully acknowledge the limited sample size analyzed at this stage of the study and the need for further expansion, given the particularly high interpatient variability of the peripheral blood immune profile, which is typically influenced by numerous intrinsic and extrinsic (environmental) factors. Another important point is the need for long-term follow-up of patients in the postoperative period to identify correlations between molecular profiling, immunophenotyping, and clinicopathological parameters. In the flow cytometry portion of our study, we did not analyze the expression of immune checkpoints in Treg, which is a goal for future research.

Conclusion

To our knowledge, this is the first study to examine the transcriptomic features of circulating lymphocytes and monocytes in women diagnosed with breast FA and the first to analyze the co-expression pattern of immune checkpoint markers (PD-1, PD-L1, TIM-3, and LAG-3). Our results convincingly demonstrate the involvement of specific immune molecular pathways in the pathogenesis of FA, including inflammatory mechanisms, regulation of innate immunity, humoral response, and antioxidant defense. All of these pathways converge on the immune checkpoint regulatory network that controls the activation/suppression of lymphocytes. For the first time, we demonstrated signs of T cell exhaustion determined by changes in the expression of these immune checkpoint molecules.

Conflict of interest

The authors declare that they have no conflicts of interest.

Funding

The study received no external funding.

Acknowledgments

The authors state that they did not use AI or AI-assisted technologies in writing the article.

Ethical approval

All procedures involving human participants were conducted in accordance with the ethical standards and with the 1964 Declaration of Helsinki and its subsequent amendments. The study was approved by the Committee on Medical Ethics at the Institute of Medicine, Petrozavodsk State University, Republic of Karelia Ministry of Healthcare (protocol No.3, approval date: 21 Oct 2024), and was completed in accordance with the good clinical practice guidelines.

References

- Tan PH. Fibroepithelial lesions revisited: Implications for diagnosis and management. *Mod Pathol* 2021; 34(Suppl 1): 15-37. <https://doi.org/10.1038/s41379-020-0583-3>.
- Nassar A, Visscher DW, Degnim AC, Frank RD, Vierkant RA, Frost M, et al. Complex fibroadenoma and breast cancer risk: A Mayo Clinic Benign Breast Disease Cohort Study. *Breast Cancer Res Treat* 2015; 153(2): 397-405. <https://doi.org/10.1007/s10549-015-3535-8>.
- Fisenko EP, Ivanova AG. Categorization of fibroepithelial breast tumors according to ultrasound BI-RADS classification. *Ultrasound & Functional Diagnostics* 2023; (1): 10-22. Russian. <https://doi.org/10.24835/1607-0771-2023-1-10-22>.
- Krishnamurthy K, Alghamdi S, Gyapong S, Kaplan S, Poppiti RJ. A clinicopathological study of fibroadenomas with epithelial proliferation including lobular carcinoma in-situ, atypical ductal hyperplasia, DCIS and invasive carcinoma. *Breast Dis* 2019; 38(3-4): 97-101. <https://doi.org/10.3233/BD-180368>.
- Chen Z, Zhang Y, Li W, Gao C, Huang F, Cheng L, et al. Single cell profiling of female breast fibroadenoma reveals distinct epithelial cell compositions and therapeutic targets. *Nat Commun* 2023; 14(1): 3469. <https://doi.org/10.1038/s41467-023-39059-3>.
- Md Nasir ND, Ng CCY, Rajasegaran V, Wong SF, Liu W. International Fibroepithelial Consortium, et al. Genomic characterization of breast fibroepithelial lesions in an international cohort. *J Pathol* 2019; 249(4): 447-460. <https://doi.org/10.1002/path.5333>.
- Xu W, Ma W, Wang D, Zhou X, Wang K, Mu K. Integrated multi-omics profiling reveals a clinically relevant molecular feature and potential therapeutic target on phyllodes tumors of breast. *Transl Oncol* 2024; 46: 101998. <https://doi.org/10.1016/j.tranon.2024.101998>.
- Kwong A, Ho CYS, Leung HCM, Leung AWS, Au CH, Ma ESK. Mutation spectrum comparison between benign breast lesion cohort, unselected cancer cohort and high-risk breast cancer cohort. *Cancers (Basel)* 2024; 16(17): 3066. <https://doi.org/10.3390/cancers16173066>.
- Winham SJ, Wang C, Heinzen EP, Bhagwate A, Liu Y, McDonough SJ, et al. Somatic mutations in benign breast disease tissues and association with breast cancer risk. *BMC Med Genomics* 2021; 14(1): 185. <https://doi.org/10.1186/s12920-021-01032-8>.
- Zhou Y, Tan Y, Zhang Q, Duan Q, Chen J. MED12 mutation as a potential predictive biomarker for immune checkpoint inhibitors in pan-cancer. *Eur J Med Res* 2022; 27(1): 225. <https://doi.org/10.1186/s40001-022-00856-z>.
- Hench J, Vlajnic T, Soysal SD, Obermann EC, Frank S, Muenst S. An integrated epigenomic and genomic view on phyllodes and phyllodes-like breast tumors. *Cancers (Basel)* 2022; 14(3): 667. <https://doi.org/10.3390/cancers14030667>.
- Pane K, Quintavalle C, Nuzzo S, Ingenito F, Roscigno G, Affinito A, et al. Comparative proteomic profiling of secreted extracellular vesicles from breast fibroadenoma and malignant lesions: A Pilot study. *Int J Mol Sci* 2022; 23(7): 3989. <https://doi.org/10.3390/ijms23073989>.
- Li X, Vail E, Maluf H, Chaum M, Leong M, Lownik J, et al. Gene expression profiling of fibroepithelial lesions of the breast. *Int J Mol Sci* 2023; 24(10): 9041. <https://doi.org/10.3390/ijms24109041>.
- Yin Lee JP, Thomas AJ, Lum SK, Shamsudin NH, Hii LW, Mai CW, et al. Gene expression profiling of giant fibroadenomas of the breast. *Surg Oncol* 2021; 37: 101536. <https://doi.org/10.1016/j.suronc.2021.101536>.
- Arkhipov SA, Studenikina AA, Arkhipova VV, Proskura AV, Autenshlyus AI. Coupling of the expression of proliferation and epithelial-mesenchymal transition markers with the histidine-rich glycoprotein HRG mRNA expression in breast diseases. *Siberian Scientific Medical Journal* 2024; 44(2): 90-95. Russian. <https://doi.org/10.18699/SSMJ20240211>.
- Adhikary S, Hoskin TL, Stallings-Mann ML, Arshad M, Frost MH, Winham SJ, et al. Cytotoxic T cell depletion with increasing epithelial abnormality in women with benign breast disease. *Breast Cancer Res Treat* 2020; 180(1): 55-61. <https://doi.org/10.1007/s10549-019-05493-5>.
- Ogony J, Hoskin TL, Stallings-Mann M, Winham S, Brahmabhatt R, Arshad MA, et al. Immune cells are increased in normal breast tissues of BRCA1/2 mutation carriers. *Breast Cancer Res Treat* 2023; 197(2): 277-285. <https://doi.org/10.1007/s10549-022-06786-y>.
- Terzeman MV, Kheifets VCh, Kostoutchek IN, Sopova ES. The functional activity of breast mast cells in women with fibroadenoma and benign breast disease. *Journal of Obstetrics and Women's Diseases* 2007; 56(4): 37-40. <https://www.elibrary.ru/ijjusp>.
- Sinha I, Fogle RL, Gulfidan G, Stanley AE, Walter V, Hollenbeak CS, et al. Potential early markers for breast cancer: A Proteomic approach comparing saliva and serum samples in a pilot study. *Int J Mol Sci* 2023; 24(4): 4164. <https://doi.org/10.3390/ijms24044164>.
- Hou H, Lyu Y, Jiang J, Wang M, Zhang R, Liew CC, et al. Peripheral blood transcriptome identifies high-risk benign and malignant breast lesions. *PLoS One* 2020; 15(6): e0233713. <https://doi.org/10.1371/journal.pone.0233713>.
- Perepechaeva ML, Studenikina AA, Grishanova AYU, Glushkov AN, Polenok EG, Bajramov PV, Autenshlyus AI. Serum miR-181a and miR-25 in patients with malignant and benign breast diseases. *Biomeditsinskaya Khimiya* 2023; 69(5): 307-314. <https://doi.org/10.18097/PBMC20236905307>.
- Guan H, Wan Y, Lan J, Wang Q, Wang Z, Li Y, et al. PD-L1 is a critical mediator of regulatory B cells and T cells in invasive breast cancer. *Sci Rep* 2016; 6: 35651. <https://doi.org/10.1038/srep35651>.
- Galaxy Community. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Res* 2024; 52(W1): W83-W94. <https://doi.org/10.1093/nar/gkae410>.
- Ge SX, Jung D, Yao R. ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 2020; 36(8): 2628-2629. <https://doi.org/10.1093/bioinformatics/btz931>.
- Dai H, Wang L, Li L, Huang Z, Ye L. Metallothionein 1: A new spotlight on inflammatory diseases. *Front Immunol* 2021; 12: 739918. <https://doi.org/10.3389/fimmu.2021.739918>.
- Grimwood J, Gordon LA, Olsen A, Terry A, Schmutz J, Lamerdin J, et al. The DNA sequence and biology of human chromosome 19. *Nature* 2004; 428(6982): 529-535. <https://doi.org/10.1038/nature02399>.
- Schmutz J, Martin J, Terry A, Couronne O, Grimwood J, Lowry S, et al. The DNA sequence and comparative analysis of human chromosome 5. *Nature* 2004; 431(7006): 268-274. <https://doi.org/10.1038/nature02919>.
- Busfield SJ, Comrack CA, Yu G, Chickering TW, Smutko JS, Zhou H, et al. Identification and gene organization of three novel members of the IL-1 family on human chromosome 2. *Genomics* 2000; 66(2): 213-216. <https://doi.org/10.1006/geno.2000.6184>.
- Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0.

- Bioinformatics* 2011; 27(12): 1739-1740. <https://doi.org/10.1093/bioinformatics/btr260>.
30. Godec J, Tan Y, Liberzon A, Tamayo P, Bhattacharya S, Butte AJ, et al. Compendium of immune signatures identifies conserved and species-specific biology in response to inflammation. *Immunity* 2016; 44(1): 194-206. <https://doi.org/10.1016/j.immuni.2015.12.006>.
 31. Strizova Z, Benesova I, Bartolini R, Novyzedlak R, Cecdrlava E, Foley LK, et al. M1/M2 macrophages and their overlaps – Myth or reality? *Clin Sci (Lond)* 2023; 137(15): 1067-1093. <https://doi.org/10.1042/CS20220531>.
 32. Nobs SP, Natali S, Pohlmeier L, Okreglicka K, Schneider C, Kurrer M, et al. PPAR γ in dendritic cells and T cells drives pathogenic type-2 effector responses in lung inflammation. *J Exp Med* 2017; 214(10): 3015-3035. <https://doi.org/10.1084/jem.20162069>.
 33. Grusdat M, Brenner D. Adipose Treg cells in charge of metabolism. *Nat Immunol* 2024; 25(3): 392-393. <https://doi.org/10.1038/s41590-024-01762-8>.
 34. Fan P, Li X, Feng Y, Cai H, Dong D, Peng Y, et al. PD-1 Expression status on CD8+ tumour infiltrating lymphocytes associates with survival in cervical cancer. *Front Oncol* 2021; 11: 678758. <https://doi.org/10.3389/fonc.2021.678758>.
 35. Garrity R, Haque A, Kavelaars A, Heijnen C, Shepherd A. Fibroblast-derived PI16 drives the physiological process of immune cell recruitment and activation during inflammatory response. *J Pain* 2024; 25(4 Suppl.): 13. <https://doi.org/10.1016/j.jpain.2024.01.067>.
 36. Sun Y, Lin S, Wang H, Wang L, Qiu Y, Zhang F, et al. Regulatory role of PI16 in autoimmune arthritis and intestinal inflammation: implications for Treg cell differentiation and function. *J Transl Med* 2024; 22(1): 327. <https://doi.org/10.1186/s12967-024-05082-1>.
 37. Sampaio FA, Martins LM, Dourado CSME, Revoredo CMS, Costa-Silva DR, Oliveira VA, et al. A case-control study of Metallothionein-1 expression in breast cancer and breast fibroadenoma. *Sci Rep* 2019; 9(1): 7407. <https://doi.org/10.1038/s41598-019-43565-0>.
 38. Tarhini A, Hedges D, Tan AC, Rodriguez P, Sukrithan V, Ratan A, et al. Differences in co-expression of T cell co-inhibitory and co-stimulatory molecules with PD1 across different human cancers. *J Immunother Cancer* 2022; 10(Suppl 2): A1190. <https://doi.org/10.1136/jitc-2022-SITC2022.1147>.
 39. Solorzano-Ibarra F, Alexandre-Gonzalez AG, Ortiz-Lazareno PC, Bastidas-Ramirez BE, Zepeda-Moreno A, Tellez-Bañuelos MC, et al. Immune checkpoint expression on peripheral cytotoxic lymphocytes in cervical cancer patients: Moving beyond the PD-1/PD-L1 axis. *Clin Exp Immunol* 2021; 204(1): 78-95. <https://doi.org/10.1111/cei.13561>.
 40. Ma J, Zheng B, Goswami S, Meng L, Zhang D, Cao C, et al. PD1Hi CD8+ T cells correlate with exhausted signature and poor clinical outcome in hepatocellular carcinoma. *J Immunother Cancer* 2019; 7(1): 331. <https://doi.org/10.1186/s40425-019-0814-7>.
 41. Čelešnik H, Potočnik U. Peripheral blood transcriptome in breast cancer patients as a source of less invasive immune biomarkers for personalized medicine, and implications for triple negative breast cancer. *Cancers (Basel)* 2022; 14(3): 591. <https://doi.org/10.3390/cancers14030591>.
 42. Ming W, Xie H, Hu Z, Chen Y, Zhu Y, Bai Y, et al. Two distinct subtypes revealed in blood transcriptome of breast cancer patients with an unsupervised analysis. *Front Oncol* 2019; 9: 985. <https://doi.org/10.3389/fonc.2019.00985>.
 43. Gao H, Ouyang D, Guan X, Xu J, Chen Q, Zeng L, et al. Immune characteristics and clinical significance of peripheral blood lymphocytes in breast cancer. *BMC Cancer* 2024; 24(1): 50. <https://doi.org/10.1186/s12885-024-11815-8>.
 44. Diskin B, Adam S, Cassini MF, Sanchez G, Liria M, Aykut B, et al. PD-L1 engagement on T cells promotes self-tolerance and suppression of neighboring macrophages and effector T cells in cancer. *Nat Immunol* 2020; 21(4): 442-454. <https://doi.org/10.1038/s41590-020-0620-x>.
 45. Mazerolles F. New expression of PD-L1 on activated CD4+ T cells opens up new opportunities for cell interactions and signaling. *Hum Immunol* 2024; 85(4): 110831. <https://doi.org/10.1016/j.humimm.2024.110831>.
 46. Carrese B, Coppola L, Smaldone G, D'Aiuto M, Mossetti G, Febraro A, et al. Role of immune-checkpoint LAG3 as a biomarker finding tool in patient-derived organoid cultures of breast cancer. *Sci Rep* 2024; 14(1): 31504. <https://doi.org/10.1038/s41598-024-83061-8>.
 47. Du H, Yi Z, Wang L, Li Z, Niu B, Ren G. The co-expression characteristics of LAG3 and PD-1 on the T cells of patients with breast cancer reveal a new therapeutic strategy. *Int Immunopharmacol* 2020; 78: 106113. <https://doi.org/10.1016/j.intimp.2019.106113>.
 48. Chauhan SK, Dunn C, Andresen NK, Røssevold AH, Skorstad G, Sike A, et al. Peripheral immune cells in metastatic breast cancer patients display a systemic immunosuppressed signature consistent with chronic inflammation. *NPJ Breast Cancer* 2024; 10(1): 30. <https://doi.org/10.1038/s41523-024-00638-2>.

Authors:

Olga V. Kurmyshkina – PhD, Associate Professor, Department of Biomedical Chemistry, Immunology and Laboratory Diagnostics, Institute of Medicine, Petrozavodsk State University, Petrozavodsk, Russia. <https://orcid.org/0000-0002-0554-8038>.

Tatyana M. Kulikovskaya – MD, PhD, Associate Professor, Department of General and Faculty Surgery, Institute of Medicine, Petrozavodsk State University (PetrSU), Petrozavodsk, Russia. <https://orcid.org/0009-0001-2493-9544>.

Tatyana O. Volkova – DSc, Professor, Chair of the Department of Biomedical Chemistry, Immunology and Laboratory Diagnostics, Institute of Medicine, Petrozavodsk State University, Petrozavodsk, Russia. <https://orcid.org/0000-0002-4606-0784>.