Identification of proteins regulating phenotype-associated genes of M2 macrophages: a bioinformatic analysis

E.A. Antropova (iD) (✉), I.V. Yatsyk (iD), P.S. Demenkov (iD), T.V. Ivanisenko (iD), V.A. Ivanisenko (iD)

Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
(✉) nzhenia@bionet.nsc.ru

Abstract. Macrophages are immune system cells that perform various, often opposing, functions in the organism depending on the incoming microenvironment signals. This is possible due to the plasticity of macrophages, which allows them to radically alter their phenotypic characteristics and gene expression profiles, as well as return to their original, non-activated state. Depending on the inductors acting on the cell, macrophages are activated into various functional states. There are five main phenotypes of activated macrophages: M1, M2a, M2b, M2c, and M2d. Although the amount of genome-wide transcriptomic and proteomic data showing differences between major macrophage phenotypes and non-activated macrophages (M0) is rapidly growing, questions regarding the mechanisms regulating gene and protein expression profiles in macrophages of different phenotypes still remain. We compiled lists of proteins associated with the macrophage phenotypes M1, M2a, M2b, M2c, and M2d (phenotype-associated proteins) and analyzed the data on potential mediators of macrophage polarization. Furthermore, using the computational system ANDSystem, we conducted a search and analysis of the relationships between potential regulatory proteins and the genes encoding the proteins associated with the M2 group phenotypes, obtaining estimates of the statistical significance of these relationships. The results indicate that the differences in the M2a, M2b, M2c, and M2d macrophage phenotypes may be attributed to the regulatory effects of the proteins JUN, IL8, NFAC2, CCND1, and YAP1. The expression levels of these proteins vary among the M2 group phenotypes, which in turn leads to different levels of gene expression associated with specific phenotypes.

Key words: macrophage phenotypes; expression regulation; proteomes; ANDSystem; automated text analysis

Выявление белков, регулирующих фенотип-ассоциированные гены макрофагов группы M2: биоинформатический анализ

Е.А. Антропова (iD) (✉), И.В. Яцык (iD), П.С. Деменков (iD), Т.В. Иванисенко (iD), В.А. Иванисенко (iD)

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
(✉) nzhenia@bionet.nsc.ru

Аннотация. Макрофаги – клетки иммунной системы, выполняющие в организме различные, часто противоположные функции в зависимости от поступающих сигналов микроокружения. Это возможно благодаря пластичности макрофагов, позволяющей кардинально менять фенотипические признаки и профили экспрессии генов, а также возвращаться в исходное, неактивированное состояние. В зависимости от действующих на клетку индукторов макрофаги поляризуются в различные функциональные состояния. Принято выделять пять основных фенотипов активированных макрофагов: M1, M2a, M2b, M2c и M2d. Хотя количество полногеномных транскриптомных и протеомных данных, показывающих различия между основными фенотипами макрофагов и неактивированными макрофагами (M0), растет стремительно, все еще остаются вопросы, касающиеся механизмов регуляции профилей экспрессии генов и белков у макрофагов разных фенотипов. Нами были составлены списки белков, ассоциированных с фенотипами макрофагов M1, M2a, M2b, M2c, M2d (фенотип-ассоциированные белки), проанализированы данные о возможных посредниках поляризации макрофагов. Далее с использованием компьютерной системы ANDSystem проведен поиск и анализ связей между потенциальными регуляторными белками и генами, кодирующими белки, ассоциированные с фенотипами группы M2, получены оценки стати-

стической значимости этих связей. Результаты указывают на то, что различия в фенотипах макрофагов M2a, M2b, M2c, M2d могут быть обусловлены регуляторными действиями белков JUN, IL8, NFAC2, CCND1 и YAP1. Уровень их экспрессии варьируется в зависимости от фенотипов группы М2, что в свою очередь приводит к различным уровням экспрессии генов, связанных с конкретными фенотипами.

Ключевые слова: фенотипы макрофагов; регуляция экспрессии; протеомы; система ANDSystem; автоматический анализ текстов

Introduction

Macrophages are immune system cells that play a key role in processes such as: maintaining body homeostasis (Mosser et al., 2021), defense against infections (Zhang M., Wang, 2014), proinflammatory and anti-inflammatory responses (Xu et al., 2013), tissue regeneration with concomitant stimulation of proliferation (Wynn, Vannella, 2016), and many others. The ability of macrophages to exhibit different functions through polarization (changing their functional state depending on signals from the microenvironment) is associated with their unique plasticity (Mills, 2012; Gurvich et al., 2020). Polarization leads to macrophages acquiring various phenotypes – functional states characterized by unique morphological, molecular and functional features, depending on the polarization inducers: proteins, peptides, polysaccharides, etc.

Each macrophage phenotype is characterized by a group of proteins (Martinez et al., 2008). These groups overlap, but different macrophage phenotypes can have radically different functions. For example, the M1 phenotype corresponds to proinflammatory macrophages, essential for the body's response to infections. M2a macrophages promote wound healing and clear the body of apoptotic cells (Murray et al., 2014). M2b macrophages are called regulatory for their ability to regulate T-helper cells, which leads to a switch in the immune response from proinflammatory to anti-inflammatory. M2c macrophages are necessary for tissue remodeling and the phagocytosis of apoptotic cells. M2d macrophages are called tumor-associated macrophages because they accompany tumor tissues (Zhang Q., Sioud, 2023).

In several studies, a link has been demonstrated between specific macrophage phenotypes and certain pathologies, as well as an association of disease outcomes with particular macrophage phenotypes. For example, patients with ovarian cancer exhibited a pronounced predominance of M1 phenotype macrophages over M2, which was associated with improved survival (Zhang M. et al., 2014). Additionally, the shift of macrophages from the M2 phenotype to M1 suppressed tumor metastasis (Yuan et al., 2017). Research on juvenile idiopathic arthritis in remission showed that the M2 macrophage group predominantly consisted of M2b and M2c, while the number of M2a macrophages was significantly reduced (Feng et al., 2021). In contrast, children with active juvenile idiopathic arthritis had a predominance of M2a and M2b macrophages, while the presence of M2c was decreased. The study of differences between macrophage phenotypes holds significant fundamental importance and also represents substantial practical interest for early disease diagnosis, prognosis, and management of disease progression (Zhang M. et al., 2014; Lampiasi, 2023).

It should be noted that there is conflicting information in the literature regarding the proteins and genes characterizing different macrophage phenotypes. For example, the fractalkine receptor (CX3CR1) is designated as a marker of the M2a phenotype in one publication (Joerink et al., 2011), while in another publication (Chhor et al., 2013), this protein is identified as a marker of the M1 phenotype. Metalloproteinase MMP12 is highlighted as a marker of the M1 phenotype (Hirani et al. 2021), but the article (Lee et al. 2014) shows that this protein is characteristic of the proteomes of the M2 phenotype and dendritic cells. The chemokine CXCL13 is described as an M1 marker in the study (Martinez et al. 2006), while in the work (van der Lans et al. 2015) it is noted as a marker of M2.

How do proteomes intersect in macrophages of different phenotypes to achieve significant functional differences? What molecular and genetic regulatory mechanisms underlie macrophage polarization? Despite the rapid accumulation of genome-wide transcriptomic and proteomic data characterizing the differences between the major macrophage phenotypes and their differences from non-activated macrophages (M0) (Gurvich et al., 2020; Oates et al., 2023), questions about how gene and protein expression profiles are regulated in macrophages of different phenotypes remain open.

The aim of this study was to identify mediator proteins that control the activity of phenotype-associated genes in different phenotypes of M2 macrophages. For this purpose, we used the ANDSystem information system, which is based on machine learning and artificial intelligence methods, including graph neural networks (Ivanisenko V.A. et al., 2015; Ivanisenko T.V. et al., 2024). ANDSystem provides automated analysis of scientific publication texts and factographic databases in the medical and biological domains. Currently, the ANDSystem knowledge base contains knowledge and facts extracted from more than 40 million scientific publications and patents, as well as factual databases, including information on molecular and genetic objects and processes that are important for the functioning of gene networks and their basic components: metabolic networks, signal transduction pathways, DNA-protein and protein-protein interaction networks. The effectiveness of ANDSystem has been demonstrated in a wide range of studies: reconstruction of molecular genetic mechanisms of asthma and hypertension comorbidity (Zolotareva et al., 2019), analysis of the plasma metabolome of patients with postoperative delirium (Ivanisenko V.A. et al., 2023), reconstruction of the hypermethylation regulatory network affecting the development of hepatocellular carcinoma in hepatitis C virus disease (Antropova et al., 2023).

In this work, the following tasks were addressed: 1) formation of phenotype-associated protein lists in macrophages of
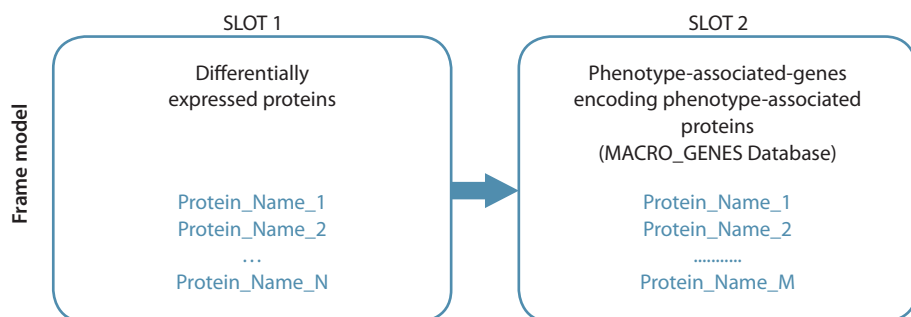
Fig. 1. Schematic diagram of a frame model for searching for regulatory links between differentially expressed proteins and phenotype-associated genes.

the main phenotypes (M1, M2a, M2b, M2c, M2d); 2) analysis of differential protein expression data in the M2 phenotype group as potential mediators of macrophage polarization; 3) analysis of regulatory relationships from mediator proteins to genes encoding phenotype-associated proteins using ANDSystem.

Materials and methods

Proteomic data on macrophages of different phenotypes. Two types of information about proteins in different macrophage phenotypes were used in the work:

1) Our curated database MACRO_GENES, containing lists of genes and proteins associated with macrophage phenotypes (Table S1)[1]. It was formed through manual analysis of scientific publications describing characteristic proteins that allow differentiation of macrophage phenotypes M1, M2a, M2b, M2c, M2d. Only those proteins, the presence of which in macrophages of certain phenotypes was confirmed by experimental data, were included in the MACRO_GENES database.

2) Proteomic data on differentially expressed proteins in M2a, M2b, M2c, and M2d macrophage phenotypes were obtained from the work by P. Li and colleagues (2022): approximately 200 proteins for each phenotype under consideration. Hereafter, such proteins will be referred to as regulatory proteins or differentially expressed proteins.

Search for potential regulators influencing the activity of phenotype-associated genes. The search for potential regulatory proteins influencing the activity of phenotype-associated genes was carried out using the knowledge base of the ANDSystem software and the ANDVisio software module included in this system (Demenkov et al., 2012; Ivanisenko V.A. et al., 2015; Ivanisenko T.V. et al., 2024). The ANDSystem knowledge base includes information on interactions between molecular biological objects (genes, proteins, metabolites, biological processes, etc.), obtained through automated analysis of over 40 million scientific publications and patents, as well as a large number of biomedical factual databases. The current version of this knowledge base contains information on over 36 million proteins from various organisms and approximately the same number of genes, 76 thousand metabolites, 100 million interactions, 21 thousand diseases, and more.

[1] Supplementary Tables S1–S3 and Figures S1–S6 are available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Antropova_Engl_29_7.pdf

To search for connections between regulatory proteins and phenotype-associated genes, the frame model software of the ANDSystem was used (Fig. 1). Step 1: The first slot of the frame was filled based on proteomic analysis data (Li et al., 2022) with a list of differentially expressed proteins for each phenotype (M2a, M2b, M2c, and M2d). Step 2: The second slot of the frame was filled with a list of phenotype-associated genes for the same phenotype from our curated MACRO_GENES database (Table S1). Step 3: Using the ANDVisio software module with the filled frame, regulatory connections described in the ANDSystem knowledge base were searched for the studied macrophage phenotype.

The analysis resulted in graphs of regulatory processes, in which nodes corresponded to differentially expressed proteins from the paper (Li et al., 2022) and phenotype-associated macrophage genes from the MACRO_GENES database. Edges connecting graph nodes corresponded to regulatory relationships between them.

Search for functionally significant regulatory proteins of phenotype-associated macrophage genes in regulatory process graphs. A key step in analyzing regulatory processes, associated with macrophage phenotype-associated genes and identified using frame models, is the search for functionally significant regulatory proteins (also called central nodes). Central nodes play a key role in signaling and coordinating regulatory processes. A wide range of methods have been developed to assess centrality (Ghasemi et al., 2014; Jalili et al., 2016; Ivanisenko V.A. et al., 2019). In our study, node centrality was assessed based on the number of interactions of the protein in question with phenotype-associated genes of the corresponding phenotype.

A high degree of centrality can be observed as a result of functional innovations between genes and proteins, as well as due to random factors. To distinguish between these situations, the statistical significance of the observed degree of centrality was assessed using the hypergeometric test. In this context, the hypergeometric test is used to measure the number of connections between a given protein and randomly determined phenotype-associated genes.

Here: M is the total number of genes represented in the ANDSystem knowledge base; N is the total number of genes with which a specific protein interacts in the ANDSystem knowledge base; n is the number of phenotype-associated

genes for a specific phenotype in the MACRO_GENES database; x is the observed number of interactions of the protein in question with phenotype-associated genes for a specific phenotype. Then, under the null hypothesis of a random distribution of interactions, the value of X obeys the hypergeometric distribution law:

$$X \sim Hypergeom(M, N, n),$$

and the p-value for the right-tailed (enrichment) test was calculated using the formula:

$$p = P(X \geq x) = \sum_{k=x}^{\min(N,n)} \frac{\binom{N}{k}\binom{M-N}{n-k}}{\binom{M}{n}}.$$

p-values were calculated using the SciPy Python library (scipy.stats.hypergeom). A Bonferroni correction was used to correct for multiple testing. At $p < 0.05$, the observed degree of centrality was considered statistically significant, and the corresponding protein was considered as a functionally significant regulatory protein controlling the expression of phenotype-associated genes.

Results and discussion
Our work aimed to identify regulatory proteins that influence genes, the expression of which differs between the M2a, M2b, M2c, and M2d macrophage phenotypes. Understanding the regulatory mechanisms that determine differences between macrophage phenotypes is not only of fundamental importance but also holds promise for applications in medicine and pharmacology, as the prevalence of a certain macrophage phenotype has been shown to be associated with the development and outcome of a number of pathologies (Zhang M. et al., 2014; Yuan et al., 2017; Feng et al., 2021).

General characteristics of phenotype-associated genes and proteins of macrophages M1, M2a, M2b, M2c, M2d
Table 1 presents a summary of our curated database, MACRO_GENES, of phenotype-associated genes encoding phenotype-associated proteins, i. e., proteins specific to macrophages of each of the phenotypes under consideration: M1, M2a, M2b, M2c, and M2d. The presence of proteins in specific phenotypes was confirmed by experimental data presented in the relevant publications. A detailed description of the gene information in MACRO_GENES is given in Table S1.

Some phenotype-associated proteins are used in experimental studies as specific markers for distinguishing macrophage phenotypes. In Table 1, the genes encoding such proteins are highlighted in green. If a protein is characteristic of a specific phenotype but is also considered a specific marker for another phenotype, the gene encoding it is highlighted in orange (Table 1). For example, the CCL2 protein is considered a marker for the M1 phenotype, but some publications indicate that it is also characteristic of the M2a and M2d phenotypes. Table 1 illustrates the complex pattern of marker intersections between different macrophage phenotypes.

Figure 2 shows a Venn diagram demonstrating the distribution of genes encoding phenotype-associated proteins across five macrophage types (M1, M2a, M2b, M2c, M2d). The diagram is constructed based on the information provided in Table 1. Note that the M1 and M2d phenotypes have the most matching proteins (17). The M2a/M2c and M1/M2c phenotype pairs have 15 and 13 common proteins, respectively. The M2b and M1 phenotypes have 11 matching proteins. A relatively small number of matching proteins (8) can be noted for the M2c and M2d phenotype pair. M2b has the fewest overlaps (6 proteins) with M2a.

Table 1. Lists of genes encoding phenotype-associated proteins of macrophages M1, M2a, M2b, M2c and M2d presented in the MACRO_GENES database

Macrophage phenotype	Genes encoding phenotype-associated proteins*
M1	*ARGI2, CAHM6, CCL2, CCL3, CCL4, CCL5, CCL7, CCL8, CCL15, CCL19, CCL20, CCR2, CCR7, CD38, CD80, CD86, CSF2, CXCL2, CXCL5, CXCL9, CXCL10, CXL11, FCG2B, FCG3A, FCGR1, GBP2, GBP5, GROA, HLA-DRA, HMGB1, IDO1, IFNA1, IFNB, IFNG, IL1A, IL1B, IL1R1, IL2RA, IL3, IL6, IL7RA, IL8, IL12B, IL15, IL15RA, IL17, IL18, IL23A, IRF1, IRF4, IRF7, ISG20, ITGAX, KCNN2, LAG3, MARCO, MET, MIF, MMP13, NAMPT, NFKB1, NOS2, PGH2, SOCS3, STAT1, TIMP1, TLR2, TLR4, TNFA, TNR5, TSP1, UBD, VEGFA*
M2a	*ALOX15, ARG1, CCL1, CCL2, CCL7, CCL8, CCL13, CCL14, CCL15, CCL17, CCL18, CCL22, CCL23, CCL24, CCL26, CCR2, CD200R1, CD209, CD274, CDH1, CDK11B, CLEC4A, CLEC7A, CLEC10A, CSF1R, CXCR1, CXCR2, EDN1, EGR2, FCER2, FGF2, FLT1, FN1, HAVCR2, HLA-DPA1, HLA-DPB1, HLA-DRA, HLA-DRB1, HLA-DRB3, HRH1, IGF1, IL1R1, IL1R2, IL4, IL10, IL27, IRF4, ITGAX, KLF2, LGALS3, MMP9, MMP14, MRC1, MRC2, MYC, PDCD1LG2, PGF, PPARG, PTGS1, RAMP1, SCN3A, SOCS1, TGFB1, TGM2, TREM2, VEGFA*
M2b	*CCL1, CCL4, CCL20, CD86, CD163, COX2, CXCL3, HLA-DRA, IFNA1, IL1B, IL1R1, IL6, IL10, MRC1, NOS2, PTPRC, SIGLEC10, SPHK1, TNFA, TNFSF14, VEGFA*
M2c	*ARG1, BCL3, C1QA, CCL8, CCL16, CCL18, CCL23, CCR2, CCR3, CCR10, CD14, CD163, CD300E, CDK11A, CDK11B, CSF1R, CX3CR1, CXCL12, CXCL13, CXCR4, EPAS1, F5, FCRLA, FPR1, GAS6, GXYLT2, HIF1A, IL1B, IL1R1, IL4R, IL10, IL21R, IRF3, IRF5, IRF8, ITGAX, JAK3, LIN7A, MAF, MARCO, MCTP2, MERTK, MMP2, MMP8, MMP14, MRC1, MRC2, MSR1, NOS2, PCOLCE2, PGF, PLOD2, SELENOP, SERPINA1, SH3PXD2B, SLAMF1, SOCS3, SPP1, SRPX2, STAT1, STAT3, STAT6, TGFB1, THBS1, TIMP1, TLR1, TLR2, TLR4, TLR5, TLR8, VCAN, VTCN1*
M2d	*ADORA2A, AIF1, C1QA, C1QC, CCL2, CCL3, CCL4, CCL5, CCL7, CD81, CD274, COX2, CSF3R, IL8, CXCL9, CXCL10, CXCL16, EGF, FCRL2, FGF2, FGFR1, GDF15, HLA-DMA, ID3, IDO1, IDO2, IL1A, IL1B, IL6, IL10, IRF7, LILRB4, MIF, MMP2, MMP9, MRC1, MSR1, NCAM1, NOS2, PDCD1LG2, PDGFB, TBX6, TGFBI, TNFA, VEGFA*

* Genes encoding proteins that are markers of various macrophages phenotypes are highlighted in green. Genes that are expressed in macrophages of a particular phenotype, according to some sources, but are markers of macrophages of a different phenotype, according to other sources, are highlighted in orange.
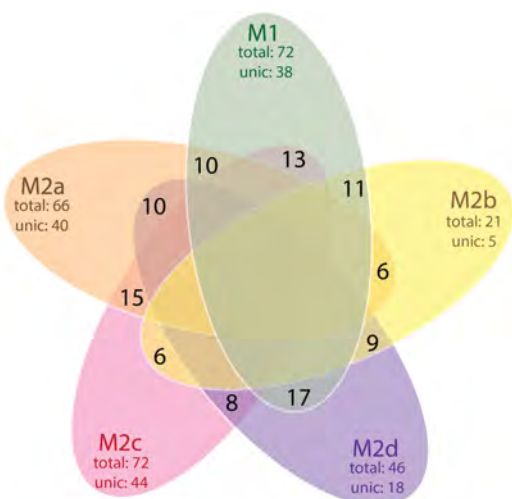
Fig. 2. Venn diagram for comparison of macrophage phenotypes M1, M2a, M2b, M2c, M2d according to the gene lists presented in the MACRO_GENES database.

General characteristics of differentially expressed proteins of the M2 macrophage group

To search for mediator proteins that transmit signals from macrophage polarization inducers to phenotype-associated genes, we used lists of differentially expressed proteins (compared to non-activated macrophages) from P. Li et al. (2022). The published data, summarized in Tables S2 and S3, indicate that the distribution of differentially expressed proteins across the four macrophage phenotypes (M2a, M2b, M2c, M2d) is characterized by significant overlap, i. e., there is no one-to-one correspondence between differentially expressed proteins and macrophage phenotypes. Therefore, to identify regulatory pathways that mediate macrophage polarization into different phenotypes, we required bioinformatic analysis of large volumes of molecular genetic data, conducted using the ANDSystem computer system.

Search for regulatory links from differentially expressed proteins to phenotype-associated genes of macrophages based on frame models

To analyze large volumes of published data on various macrophage phenotypes, we used the methods and information resources of computer-aided knowledge engineering implemented in the ANDSystem. Using the framework-based approach realized in this system, we searched for regulatory links between differentially expressed proteins and phenotype-associated genes in macrophages.

Regulatory process graphs were reconstructed, with nodes corresponding to differentially expressed proteins from the paper (Li et al., 2022) and phenotype-associated macrophage genes from the MACRO_GENES database. Edges connecting graph nodes corresponded to regulatory relationships between them. Figure 3 shows an example of a graph of potential regulatory relationships between differentially expressed proteins and phenotype-associated macrophage genes in the M2b phenotype.

Figure 3 shows that most phenotype-associated genes are regulated by more than one protein. Furthermore, most of the proteins shown in the figure are involved in the regulation of multiple genes. Similar regulatory relationship diagrams for M2a, M2c, and M2d macrophages are presented in the Supplementary Materials (Fig. S1–S3).

Identification of statistically significant regulators of phenotype-associated genes

Quantitative characteristics of regulatory links between differentially expressed proteins and phenotype-associated genes identified using frame models are shown in Section A of Table 2.

In the second stage of the analysis, centrality metrics characterizing the functional significance of differentially expressed proteins for the regulation of phenotype-associated genes were assessed. Centrality assessments allowed us to identify proteins regulating phenotype-associated genes with a Bonferroni-corrected statistical significance threshold of $p < 0.05$ (Table 2B). Accounting for the statistical significance of differentially expressed proteins based on centrality metrics led to a significant reduction in the number of nodes corresponding to phenotype-associated genes and the number of edges corresponding to regulatory events. For example, for the M2a phenotype, of the 40 differentially expressed proteins associated with phenotype-associated genes, 16 were statistically significantly associated with these genes (Table 2B). For the M2b phenotype, it was 6 out of 12 proteins. Accordingly, the number of regulatory events and phenotype-associated genes in the reconstructed graphs decreased.

Figure 4 shows the lists of differentially expressed proteins statistically significantly associated with the phenotype-associated genes of macrophages M2a, M2b, M2c and M2d. Proteins, the levels of which are elevated in specific macrophage phenotypes according to the study (Li et al., 2022), are highlighted in red. Proteins, the levels of which are decreased compared to non-activated macrophages are highlighted in blue. Green lines connect proteins with oppositely changing expression levels in macrophages of different phenotypes.

Figure 5 shows examples of schemes of statistically significant regulatory interactions between differentially expressed proteins of M2a macrophages and phenotype-associated genes reconstructed using the ANDSystem. The corresponding schemes for macrophages of the M2b, M2c, and M2d phenotypes are presented in Figures S4–S6. Figure 5 demonstrates regulatory connections using two alternative options for controlling molecular genetic processes in the same M2a macrophage phenotype: through an increase (Fig. 5*A*) and a decrease (Fig. 5*B*) in the levels of regulatory proteins. A description of the reconstructed connections obtained using frame models is given in Table 3.

As an example, let us consider the binding of the regulatory protein TGM2 (the lower protein in Figure 5*A*, marked with a blue asterisk). Proteomic data (Li et al., 2022) show that the level of this protein is elevated in the M2a phenotype compared to non-activated macrophages. According to information from the ANDSystem knowledge base obtained
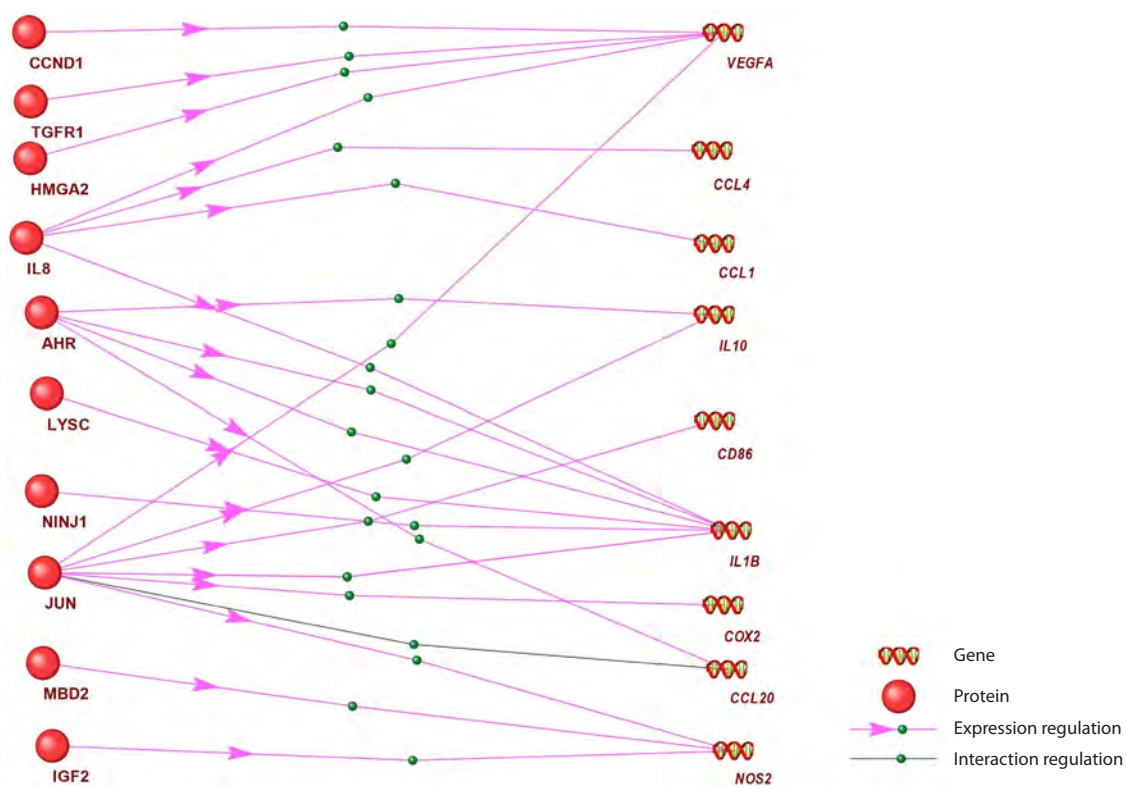
Fig. 3. A graph of potential regulatory links between differentially expressed proteins (left) and phenotype-associated genes (right) in M2b macrophages, presented in the ANDSystem interface. Green balls on the arrows in the interactive ANDSystem interface allow users to obtain additional information about specific regulatory links.

Table 2. Quantitative characteristics of potential regulatory links identified based on frame models

Components of frame models	M2a	M2b	M2c	M2d
A. Results of the first stage of analysis of the regulatory interactions reconstructed graph				
Differentially expressed proteins	40	12	41	43
Regulatory events	127	51	216	252
Phenotype-associated genes	26	12	41	31
B. Results of the second stage of the analysis (taking into account statistical estimates of the centrality of regulatory proteins, $p < 0.05$)				
Differentially expressed proteins	16	6	10	15
Regulatory events	85	19	89	133
Phenotype-associated genes	23	8	28	29

through its interface, in M2a macrophages, the TGM2 protein has an activating effect on the expression of the M2a phenotype-associated genes *CD274* and *FN1* (Liu et al., 2021; Sun et al., 2021), which is consistent with the data presented in Table 1 (the mentioned genes are marked with black asterisks in Figure 5*A*). TGM2 also has a suppressive effect on the *PPARG* gene (Maiuri et al., 2008), which is inconsistent with the data in Table 1 and indicates that the expression

of this gene is also activated by some other factors, such as AHR, CDK4, CCL5, the level of which is increased in this phenotype (Fig. 5*A*).

Among the proteins with reduced levels (compared to non-activated macrophages) in M2a macrophages, as an example we consider the CBP protein, which regulates the phenotype-associated genes *CCL2*, *CD274*, and *CDH1* (Fig. 5*B*, marked with asterisks). According to information from the

M2a	M2b	M2c	M2d
ITAM	IL8	JUN	IL1B
AHR	JUN	IL8	CXL10
TGM2	AHR	AHR	NAMPT
CCL5	TGFR1	MBD2	IL8
HMGA2	COX8A	PLMN	AHR
TGFR1	NINJ1	IGF2	MBD2
CCND1	LYSC	TGFR1	GDF15
CXL10	YLAT1	ZEB2	IL1A
NFAC2	IGF2	TRFL	SRC
CCL3		CCND1	HMGA2
CDK4		COX8A	CCL20
FOSL2			YAP1
FLT3			TIGAR
CD38			TGFR1
PLMN			ITB3
CBP			
COX8A			

Fig. 4. Differentially expressed proteins statistically significantly ($p < 0.05$) associated with the phenotype-associated genes of macrophages M2a, M2b, M2c, and M2d. Proteins, the levels of which, according to (Li et al. 2022), are elevated in a particular phenotype are highlighted in red, while those, the levels of which are decreased compared to non-activated macrophages, are highlighted in blue. Green lines connect proteins with oppositely expressed changes in macrophages of different phenotypes.

ANDSystem knowledgebase, when CBP is suppressed in the M2a phenotype, *CCL2* expression increases (Huang et al., 2021), which is consistent with the data presented in Table 1 (MACRO_GENES database). At the same time, the CBP protein positively influences the expression of the phenotype-

associated genes *CD274* and *CDH1* (Liu et al., 2020; Heng et al., 2021). It can be hypothesized that other regulators have a greater influence on the activity of these genes. Figure 5*A* shows that such regulators for the *CD274* gene may include the proteins AHR, CCL5, TGM2, and CDK4, the levels of which are elevated in the M2a phenotype (Fig. 5*A*, double green asterisks).

All statistically significant regulatory interactions identified in M2 macrophages between differentially expressed proteins and phenotype-associated genes are presented in Table 3. For M2a macrophages, these were interactions of nine regulatory proteins with increased levels (compared to non-activated macrophages), marked with arrows (↑), and four proteins with decreased levels (↓), regulating 23 phenotype-associated genes. For M2b, these were four upregulated and two down-regulated proteins regulating eight phenotype-associated genes (see also Figure S4). For M2c, two upregulated and eight downregulated proteins regulating 28 genes were identified (see also Figure S5). For M2d, 13 upregulated and two downregulated proteins regulating 29 genes were found (see also Figure S6).

Thus, based on a computer analysis of differences in the proteomes of different macrophage phenotypes, as well as the use of large volumes of information accumulated in the ANDSystem knowledge base, some regulatory proteins were identified that mediate the action of macrophage polarization inducers on phenotype-associated macrophage genes. Future research is planned using frame models containing more slots reflecting the intermediate stages of action of macrophage polarization inducers on phenotype-associated macrophage genes. This will enable the identification of more subtle features of the regulatory pathways running from macrophage polarization inducers to phenotype-associated genes through the action of intermediary proteins.
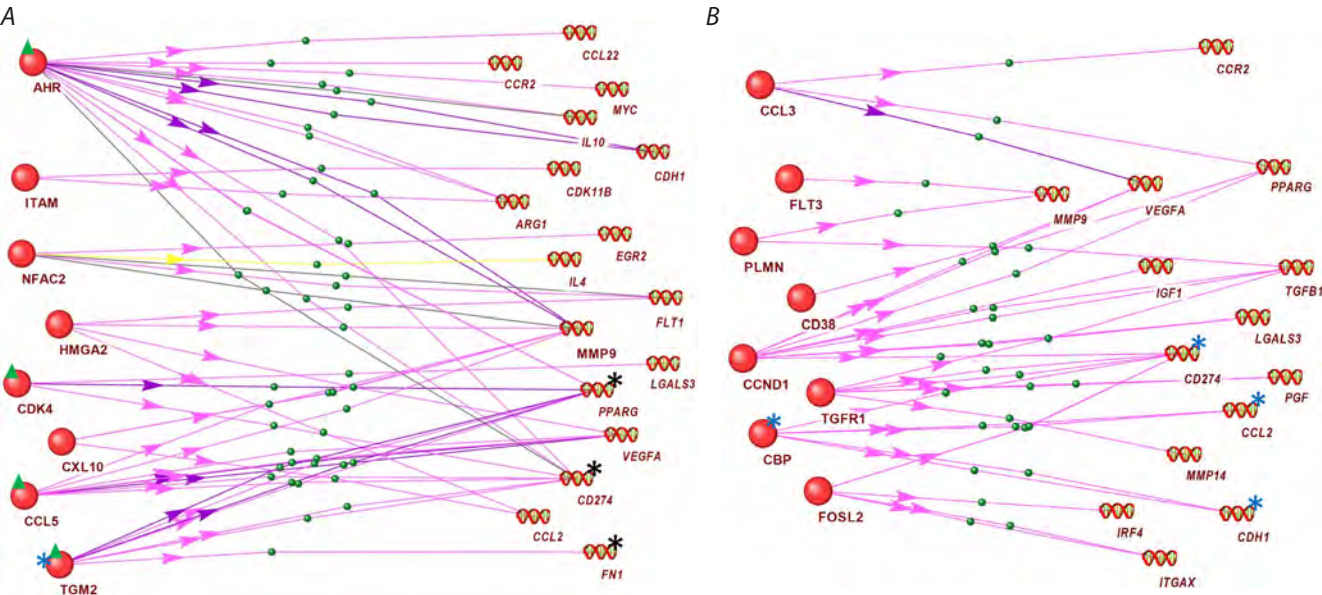

Fig. 5. Graph of the regulation of phenotype-associated gene expression (from the MACRO_GENES database) in M2a macrophages by differentially expressed regulatory proteins from the article (Li et al., 2022), statistically significantly associated with these genes: (*A*) through an increase and (*B*) through a decrease in the level of regulatory proteins in this phenotype. The blue asterisk in (*A*) indicates the TGM2 protein discussed in the text, black asterisks indicate its target genes; green triangles indicate the discussed proteins AHR, CDK4, CCL2, TGM2. Blue asterisks in (*B*) indicate the CBP protein discussed in the text and its target genes.

Table 3. Relationships between functionally significant differentially expressed regulatory proteins*
and the phenotype-associated genes they regulate in M2 group macrophages, identified using frame models

M2a		M2b		M2c		M2d	
Protein	Target genes	Protein	Target genes	Protein	Target genes	Protein	Target genes
AHR (↑)**	*ARG1, MMP9, MYC, PPARG, CCL22, CCR2, CDH1, CD274, IL10*	AHR (↑)	*IL1B, CCL20, IL10*	AHR (↑)	*ARG1, CXCL12, STAT3, SOCS3, CCR2, HIF1A, TLR2, IL1B, IL10*	AHR (↑)	*IDO1, IDO2, IL1B, IL8,MMP9, NCAM1, CD274, IL1A, IL10*
TGM2 (↑)	*PPARG, CD274, FN1, MMP14*	IGF2 (↑)	*NOS2*	MBD2 (↑)	*CXCL12, NOS2, SOCS3*	IL1B (↑)	*FGF2, FGFR1, CXCL10, IL8, MMP9, VEGFA, TGFB1, CCL2, CCL3, CCL4, CCL5, NOS2, COX2, IL10, IL1A, TNF, PDCD1LG2, MIF*
CCL5 (↑)	*VEGFA, CD274, MMP9*	NINJ1 (↑)	*IL1B*				
NFAC2 (↑)	*EGR2, IL4, FLT1, MMP9*	IL8 (↑)	*CCL1, VEGFA*	TGFR1 (↓)	*MMP14, TGFB1, HIF1A, PGF*	MBD2 (↑)	*CXCL9, CXCL10, CD274, CCL5, MMP2, NOS2*
HMGA2 (↑)	*FLT1, MMP9, VEGFA, CCL2*	TGFR1 (↓)	*VEGFA*	ZEB2 (↓)	*MAF, MMP14, TGFB1,*	IL8 (↑)	*IL1B, FGF2, CXCL10, MMP9, CCL2, CD274, MMP2, VEGFA*
ITAM (↑)	*ARG1, CDK11B*	JUN (↓)	*CD86, COX2, NOS2, IL1B, IL10, CCL20, VEGFA*	CCND1 (↓)	*TGFB1, STAT3, HIF1A*	YAP1 (↑)	*IRF7, FGFR1, PDGFB, CD274*
CDK4 (↑)	*PPARG, LGALS3, CD274*			IL8 (↓)	*MMP14, MMP2, MMP8, TLR5, CXCL12, ARG1, CCR2, HIF1A, IL1B, CDK11B*	NAMPT (↑)	*IL1B, FGF2, IL8, MMP9, CCL2, CD274, MMP2*
CXL10 (↑)	*MMP9, CCL2*						
CCND1 (↓)	*VEGFA, TGFB1, CD274*			IGF2 (↓)	*CDK11B, MMP2, HIF1A, NOS2*	IL1A (↑)	*IL6, IL1B, VEGFA, CD274*
						CCL20 (↑)	*MMP9, CD274*
FLT3 (↓) ***	*MMP9*			PLMN (↓)	*TGFB1, CXCL12, CXCR4*	CXL10 (↑)	*IL1B, MMP9, CCL2, MMP2*
CCL3 (↓)	*CCR2, VEGFA, PPARG*			TRFL (↓)	*HIF1A, TLR4*	TIGAR (↑)	*MMP9, MMP2*
PLMN (↓)	*MMP9, TGFB1*			JUN (↓)	*TGFB1, CXCL12, MMP14, MMP2, ARG1, IL1B, NOS2, HIF1A, TLR2, TLR4, IL10, CX3CR1, THBS1, SERPINA1, BCL3, CCL18, ITGAX*	HMGA2 (↑)	*MMP9, CCL2, MMP2, VEFGA*
CD38 (↓)	*VEGFA*					SRC (↑)	*CXCL10, MMP2, VEFGA, NOS2*
TGFR1 (↓)	*TGFB1, CD274, MMP14, PGF*					ITB3 (↑)	*MMP2*
CBP (↓)	*CD274, CCL2, CDH1*					TGFR1 (↓)	*VEGFA, TGFB1, CD274*
FOSL2 (↓)	*CD274, IRF4, ITGAX*					GDF15 (↓)	*MMP9, CCL2, GDF15, CD274, MMP2*

* – Proteins selected based on the centrality criterion ($p < 0.05$); ** ↑ – proteins with increased expression levels; ***↓ – proteins with decreased expression levels.

Conclusion

A study of published data on phenotype-associated genes and proteomes of M2 macrophages, and a subsequent search for regulatory links between them using a frame-based approach implemented in the ANDSystem computer system, made it possible to identify potential regulatory proteins that mediate differences in gene expression in M2 macrophage phenotypes.

The obtained results suggest that the differences between the M2a, M2b, M2c and M2d phenotypes may be associated, in particular, with the regulatory functions of the proteins JUN, IL8, NFAC2, CCND1 and YAP1, the level of which varies between phenotypes, leading to differences in the expression of phenotype-associated genes.

References

Antropova E.A., Khlebodarova T.M., Demenkov P.S., Volianskaia A.R., Venzel A.S., Ivanisenko N.V., Gavrilenko A.D., Ivanisenko T.V., Adamovskaya A.V., Revva P.M., Kolchanov N.A., Lavrik I.N., Ivanisenko V.A. Reconstruction of the regulatory hypermethylation network controlling hepatocellular carcinoma development during hepatitis C viral infection. *J Integr Bioinform*. 2023;20(3): 20230013. doi 10.1515/jib-2023-0013

Chhor V., Le Charpentier T., Lebon S., Oré M.V., Celador I.L., Josserand J., Degos V., Jacotot E., Hagberg H., Sävman K., Mallard C., Gressens P., Fleiss B. Characterization of phenotype markers and neuronotoxic potential of polarised primary microglia *in vitro*. *Brain Behav Immun*. 2013;32:70-85. doi 10.1016/j.bbi.2013.02.005

Demenkov P.S., Ivanisenko T.V., Kolchanov N.A., Ivanisenko V.A. ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem. *In Silico Biol*. 2012;11(3-4):149-161. doi 10.3233/ISB-2012-0449

Feng D., Huang W.Y., Niu X.L., Hao S., Zhang L.N., Hu Y.J. Significance of macrophage subtypes in the peripheral blood of children with systemic juvenile idiopathic arthritis. *Rheumatol Ther*. 2021; 8(4):1859-1870. doi 10.1007/s40744-021-00385-x

Ghasemi M., Seidkhani H., Tamimi F., Rahgozar M., Masoudi-Nejad A. Centrality Measures in Biological Networks. *Curr Bioinform*. 2014;9:426-441. doi 10.2174/15748936113086660013

Gurvich O.L., Puttonen K.A., Bailey A., Kailaanmäki A., Skirdenko V., Sivonen M., Pietikäinen S., Parker N.R., Ylä-Herttuala S., Kekarainen T. Transcriptomics uncovers substantial variability associated with alterations in manufacturing processes of macrophage cell therapy products. *Sci Rep*. 2020;10(1):14049. doi 10.1038/s41598-020-70967-2

Heng W.S., Kruyt F.A.E., Cheah S.C. Understanding lung carcinogenesis from a morphostatic perspective: prevention and therapeutic potential of phytochemicals for targeting cancer stem cells. *Int J Mol Sci*. 2021;22(11):5697. doi 10.3390/ijms22115697

Hirani D., Alvira C.M., Danopoulos S., Milla C., Donato M., Tian L., Mohr J., … Seeger W., Khatri P., Al Alam D., Dötsch J., Alejandre Alcazar M.A. Macrophage-derived IL-6 trans-signaling as a novel target in the pathogenesis of bronchopulmonary dysplasia. *Eur Respir J*. 2021;59(2):2002248. doi 10.1183/13993003.02248-2020

Huang Y.H., Cai K., Xu P.P., Wang L., Huang C.X., Fang Y., Cheng S., Sun X.J., Liu F., Huang J.Y., Ji M.M., Zhao W.L. *CREBBP/EP300* mutations promoted tumor progression in diffuse large B-cell lymphoma through altering tumor-associated macrophage polarization via FBXW7-NOTCH-CCL2/CSF1 axis. *Signal Transduct Target Ther*. 2021;6(1):10. doi 10.1038/s41392-020-00437-8

Ivanisenko T.V., Demenkov P.S., Ivanisenko V.A. An accurate and efficient approach to knowledge extraction from scientific publications using structured ontology models, graph neural networks, and large language models. *Int J Mol Sci*. 2024;25(21):11811. doi 10.3390/ijms252111811

Ivanisenko V.A., Saik O.V., Ivanisenko N.V., Tiys E.S., Ivanisenko T.V., Demenkov P.S., Kolchanov N.A. ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology. *BMC Syst Biol*. 2015;9(Suppl. 2):S2. doi 10.1186/1752-0509-9-S2-S2

Ivanisenko V.A., Demenkov P.S., Ivanisenko T.V., Mishchenko E.L., Saik O.V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression. *BMC Bioinformatics*. 2019; 20(Suppl. 1):34. doi 10.1186/s12859-018-2567-6

Ivanisenko V.A., Basov N.V., Makarova A.A., Venzel A.S., Rogachev A.D., Demenkov P.S., Ivanisenko T.V., Kleshchev M.A., Gaisler E.V., Moroz G.B., Plesko V.V., Sotnikova Y.S., Patrushev Y.V., Lomivorotov V.V., Kolchanov N.A., Pokrovsky A.G. Gene networks for use in metabolomic data analysis of blood plasma from patients with postoperative delirium. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2023;27(7):768-775. doi 10.18699/VJGB-23-89

Jalili M., Salehzadeh-Yazdi A., Gupta S., Wolkenhauer O., Yaghmaie M., Resendis-Antonio O., Alimoghaddam K. Evolution of centrality measurements for the detection of essential proteins in biological networks. *Front Physiol*. 2016;7:375. doi 10.3389/fphys.2016.00375

Joerink M., Rindsjö E., van Riel B., Alm J., Papadogiannakis N. Placental macrophage (Hofbauer cell) polarization is independent of maternal allergen-sensitization and presence of chorioamnionitis. *Placenta*. 2011;32(5):380-385. doi 10.1016/j.placenta.2011.02.003

Lampiasi N. Macrophage polarization: learning to manage it 2.0. *Int J Mol Sci*. 2023;24(24):17409. doi 10.3390/ijms242417409

Lee J.T., Pamir N., Liu N.C., Kirk E.A., Averill M.M., Becker L., Larson I., Hagman D.K., Foster-Schubert K.E., van Yserloo B., Bornfeldt K.E., LeBoeuf R.C., Kratz M., Heinecke J.W. Macrophage metalloelastase (MMP12) regulates adipose tissue expansion, insulin sensitivity, and expression of inducible nitric oxide synthase. *Endocrinology*. 2014;155(9):3409-3420. doi 10.1210/en.2014-1037

Li P., Ma C., Li J., You S., Dang L., Wu J., Hao Z., Li J., Zhi Y., Chen L., Sun S. Proteomic characterization of four subtypes of M2 macrophages derived from human THP-1 cells. *J Zhejiang Univ Sci B*. 2022;23(5):407-422. doi 10.1631/jzus.B2100930

Liu J., He D., Cheng L., Huang C., Zhang Y., Rao X., Kong Y., … Jones K., Napier D., Lee E.Y., Wang C., Liu X. p300/CBP inhibition enhances the efficacy of programmed death-ligand 1 blockade treatment in prostate cancer. *Oncogene*. 2020;39(19):3939-3951. doi 10.1038/s41388-020-1270-z

Liu J., Liu Q., Zhang X., Cui M., Li T., Zhang Y., Liao Q. Immune subtyping for pancreatic cancer with implication in clinical outcomes and improving immunotherapy. *Cancer Cell Int*. 2021;21(1):137. doi 10.1186/s12935-021-01824-z

Maiuri L., Luciani A., Giardino I., Raia V., Villella V.R., D'Apolito M., Pettoello-Mantovani M., Guido S., Ciacci C., Cimmino M., Cexus O.N., Londei M., Quarantino S. Tissue transglutaminase activation modulates inflammation in cystic fibrosis via PPARgamma down-regulation. *J Immunol*. 2008;180(11):7697-7705. doi 10.4049/jimmunol.180.11.7697

Martinez F.O., Gordon S., Locati M., Mantovani A. Transcriptional profiling of the human monocyte-to-macrophage differentiation and polarization: new molecules and patterns of gene expression. *J Immunol*. 2006;177(10):7303-7311. doi 10.4049/jimmunol.177.10.7303

Martinez F.O., Sica A., Mantovani A., Locati M. Macrophage activation and polarization. *Front Biosci*. 2008;13:453-461. doi 10.2741/2692

Mills C.D. M1 and M2 macrophages: oracles of health and disease. *Crit Rev Immunol*. 2012;32(6):463-488. doi 10.1615/critrevimmunol.v32.i6.10

Mosser D.M., Hamidzadeh K., Goncalves R. Macrophages and the maintenance of homeostasis. *Cell Mol Immunol*. 2021;18(3):579-587. doi 10.1038/s41423-020-00541-3

Murray P.J., Allen J.E., Biswas S.K., Fisher E.A., Gilroy D.W., Goerdt S., Gordon S., … Suttles J., Udalova I., van Ginderachter J.A., Vogel S.N., Wynn T.A. Macrophage activation and polarization: nomenclature and experimental guidelines. *Immunity*. 2014;41(1): 14-20. doi 10.1016/j.immuni.2014.06.008

Oates T.C., Moura P.L., Cross S., Roberts K., Baum H.E., Haydn-Smith K.L., Wilson M.C., Heesom K.J., Severn C.E., Toye A.M. Defining the proteomic landscape of cultured macrophages and their polarization continuum. *Immunol Cell Biol*. 2023;101(10):947-963. doi 10.1111/imcb.12687

Sun W., Qin Y., Wang Z., Dong W., He L., Zhang T., Zhang H. The *NEAT1_2*/miR-491 axis modulates papillary thyroid cancer invasion and metastasis through TGM2/NFκb/FN1 signaling. *Front Oncol*. 2021;11:610547. doi 10.3389/fonc.2021.610547

van der Lans A.A., Boon M.R., Haks M.C., Quinten E., Schaart G., Ottenhoff T.H., van Marken Lichtenbelt W.D. Cold acclimation affects immune composition in skeletal muscle of healthy lean subjects. *Physiol Rep*. 2015;3(7):e12394. doi 10.14814/phy2.12394

Wynn T.A., Vannella K.M. Macrophages in tissue repair, regeneration, and fibrosis. *Immunity*. 2016;44(3):450-462. doi 10.1016/j.immuni.2016.02.015

Xu W., Zhao X., Daha M.R., van Kooten C. Reversible differentiation of pro- and anti-inflammatory macrophages. *Mol Immunol*. 2013;53(3):179-86. doi 10.1016/j.molimm.2012.07.005

Yuan R., Li S., Geng H., Wang X., Guan Q., Li X., Ren C., Yuan X. Reversing the polarization of tumor-associated macrophages inhibits tumor metastasis. *Int Immunopharmacol*. 2017;49:30-37. doi 10.1016/j.intimp.2017.05.014

Zhang M., Wang C.C. Inflammatory response of macrophages in infection. *Hepatobiliary Pancreat Dis Int*. 2014;13(2):138-152. doi 10.1016/s1499-3872(14)60024-2

Zhang M., He Y., Sun X., Li Q., Wang W., Zhao A., Di W. A high M1/M2 ratio of tumor-associated macrophages is associated with extended survival in ovarian cancer patients. *J Ovarian Res*. 2014; 7:19. doi 10.1186/1757-2215-7-19

Zhang Q., Sioud M. Tumor-associated macrophage subsets: shaping polarization and targeting. *Int J Mol Sci*. 2023;24:7493. doi 10.3390/ijms24087493

Zolotareva O., Saik O.V., Königs C., Bragina E.Y., Goncharova I.A., Freidin M.B., Dosenko V.E., Ivanisenko V.A., Hofestädt R. Comorbidity of asthma and hypertension may be mediated by shared genetic dysregulation and drug side effects. *Sci Rep*. 2019;9(1):16302. doi 10.1038/s41598-019-52762-w