*Mathematical Biology and Bioinformatics*
*2022. V. 17. № 2. P. 174–187. doi: 10.17537/2022.17.174*

========== BIOINFORMATICS ==========

# Additional Pathogenic Pathways in RBCK1 Deficiency

**Demicheva E.I.**[1,2], **Shinwari K.**[2], **Ushenin K.S.**[1,2], **Bolkov M.A.***[1,2]

[1]*Institute of Immunology and Physiology, Ekaterinburg, Russia*
[2]*Ural Federal University, Ekaterinburg, Russia*

***Abstract.*** RBCK1 deficiency is a rare congenital autoinflammatory disease that causes inflammatory disruption on the molecular level. This deficiency has three major clinical manifestations: increased sensitivity to bacterial infections, autoinflammation syndrome, and the accumulation of amylopectin in skeletal muscle. The amylopectinosis causes myopathy and cardiomyopathy. The pathogenesis of the disease is poorly investigated and may include unnoticed relationships. We performed gene expression analysis on patients with RBCK1 deficiency and three other autoinflammatory diseases. The identification of differentially expressed genes revealed a large number of downregulated genes that are involved in the activation of essential metabolic and immune pathways, including NF-kB and Pi3k-Akt-mTOR. Signaling pathways were analysed using the KEGG (Kyoto Encyclopedia of Genes and Genomes) and Gene Ontology resource. Predicted protein-protein interactions were retrieved from the STRING (Search Tool for the Retrieval of Interacting proteins database). Besides the primary involvement of RBCK1 in disease pathology, several downregulated pathways aggravate symptoms of myopathy, cardiomyopathy, and bacterial disease. The studied pathways may serve as new targets for the development of compensatory therapies for patients with RBCK1 deficiency.

***Key words:*** *gene expression profiling, autoinflammatory diseases, RBCK1, signaling pathways, protein-protein interactions.*

## INTRODUCTION

Autoinflammatory diseases are a broad class of human pathology related to innate immune errors and associated with defects in inflammatory mechanisms [1]. This class of diseases was discovered relatively recently [2], but at least 49 autoinflammatory diseases are already known from the Infevers data [3]. The main difference among these diseases is the uncontrolled autoinflammation in the absence of autoantibodies. However, mechanisms of autoinflammation have been identified in many long-known diseases, such as systemic-onset juvenile idiopathic arthritis, Crohn's disease, Behcet's disease [4].

RBCK1 (RanBP-Type And C3HC4-Type Zinc Finger-Containing Protein 1) deficiency is an autoinflammatory syndrome characterized by increased susceptibility to infections, and disorders of glycogen metabolism, leading to glycogen accumulation (amylopectinosis). The accumulation of amylopectin leads to myopathy and cardiomyopathy development. Similar cases were reported in patients with RBCK1 deficiency, who suffered from cardiomyopathy and/or skeletal myopathy.

RBCK1 protein, along with HOIL-1L interacting protein and SHANK-associated RH domain interactor, is part of the linear ubiquitination assembly complex (LUBAC). A damage in any component from the LUBAC complex results in autoinflammatory diseases. Mutations in the RBCK1 gene cause an excessive increase in the number of leukocytes, which explains

*mbolkov@iip.uran.ru

the autoinflammatory nature of the disease. In addition, the production of cytokines increases in response to TNF and IL-1b, and the innate immune response deteriorates, whereas the production of proinflammatory cytokines is decreasing. Such reactions explain the increased sustainability to infections in patients with RBCK1 deficiency.

RBCK1 protein may negatively affect NF-kB activation by inhibiting IkappaB kinase. Defects in NF-kB activation lead to abnormalities in glucose metabolism and in the immune system. In the case of RBCK1 deficiency, the normal storage of glycogen is being depleted due to the abnormal accumulation of amylopectin, which cannot be used by the cell. One of the recent hypotheses suggests that the polyubiquitination of glycogen, initiated by RBCK1, may lead to its removal by glucophagy [5].

The major pathogenesis reasons for RBCK1 deficiency are described above. However, RBCK1 participates in a wide area of intracellular interactions that possibly cause a dysregulation of other cellular functions. Such changes aggravate the main symptoms and provide an additional negative effect on patients' health conditions. This article reevaluates differential gene expression data from patients with RBCK1 deficiency. Results of the analysis reveal several downregulated pathways that aggravate symptoms of myopathy, cardiomyopathy, and bacterial diseases. The studied pathways may serve as new targets for the development of compensatory therapies for patients with RBCK1 deficiency.

## METHODS

### Experimental dataset

The GSE40561 dataset, obtained from the Illumina Human HT-Beadchips microchip was used for this research. It contains a transcription of whole blood from patients with RBCK1 deficiency and a group of other comparable autoinflammatory diseases, namely, chronic infantile neurological cutaneous and articular (CINCA) syndrome, mevalonate kinase deficiency (MVK), and Muckle-Wells syndrome (MWS). The chosen dataset includes 48.803 genes spread among 51 samples, including 1 for RBCK1 deficiency, 5 for MWS, 2 for MVK, 2 for CINCA, and 41 for healthy people. The NCBI Gene Expression Omnibus (GEO) portal provided the series matrix and related annotation file. Low-expression probes, and samples with missing gene IDs, were eliminated from the analysis.

### Data processing and analysis

In the R language program, the dataset collected from NCBI GEO was first normalized, and quality control was performed. Quantile normalization was used to limit the impact of any non-biological influences that may occur during RNA sequencing. Bolstad BM [6] developed a normalization approach for the study, which is based on the assumption that the majority of genes in the study did not change their expression values while comparing samples. As a result, expression values should be recalculated as follows: expression values of each gene are replaced with the gene's mean value across all samples. As our study involved a large number of genes, we have determined that this normalization method was enough.

We used correlation analysis, which included hierarchical clustering, to see the distribution of gene expression across samples and locate groupings with comparable data. This method is widely regarded as the most important statistical method for observing clusters of genes with similar expression patterns and displaying them as a binary tree. Euclidean distance was used to calculate the distance between each pair of genes. A singular value decomposition of the centered data matrix of gene expression, which was scaled previously by dividing by the standard deviation, was used to perform the calculations. Genes with low gene expression have lower mRNA copy numbers and are more susceptible to technical noise. Hence, future investigations should focus on differentially expressed genes (DEGs).

DEGs were selected using an empirical Bayes moderated t-test [7], with the following cut-off criteria for genes: P-value $< 0.05$, $|\text{logFC}| > 1$. The DEGs were then separated into two groups: those that were upregulated and downregulated. A co-expression analysis is another technique to detect DEGs. It is a method of clustering data into numerous modules of genes with similar functions. This method is implemented in cemiTool [8] and uses inverse gamma distribution to provide unsupervised approaches for gene filtration and correlation. The selection of soft-thresholding power, which is automatically performed by cemiTool, determines the gene's similarity. The Pearson method was used for all pairwise correlations, and the following cut-off criteria was used to select significant genes: P-value < 0.05.

**Pathway and gene ontology analysis**

The next step was to perform an over-enrichment analysis (ORA) [9] and a Gene Ontology (GO) pathway analysis of chosen DEGs. ORA involves the findings of the most enriched pathways by using hypergeometric testing [10], in which the likelihood of a group of genes from the input list being involved in a specific genetic route is determined. This analysis is suitable for establishing that the chosen DEGs have similar functions and were not randomly chosen. The following formula was used to compute the aforementioned probability:

$$p = 1 - \sum_{i=0}^{k-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

$N$ is the total number of genes in the background distribution, $M$ is the number of genes within that distribution that are annotated (either directly or indirectly) to the gene set of interest, $n$ is the size of the list with genes of interest, and $k$ is the number of genes within that list that are annotated to the gene set. For each category investigated, the findings were calculated as adjusted P-values (after multiple test corrections) and presented as barplots. To illustrate the P-value ratio among biological pathways, we used a red-yellow color scheme, with the most significant pathways being reddish in hue. The Kyoto Encyclopedia of Genes and Genomes (KEGG) [11] and the Gene Ontology (GO) databases [12] were used to find the significant pathways. KEGG contains comprehensive pathway maps, chemical processes, and tiny molecules for various species, whereas GO correlates gene sets to biological functions. Biological process, cellular component, and molecular function were the three categories in which enriched pathways from GO were searched for. In addition, ORA was performed on the WikiPathways database's whole set of curated human pathways [13]. WikiPathways is a valuable resource that, like KEGG, is frequently used to investigate disease models. It also incorporates data from other route databases, such as KEGG and Reactome. The following cut-off criteria was used to identify relevant pathways in the KEGG and WikiPathways databases: P-value < 0.05.

**Protein-protein interaction (PPI)**

The development of protein-protein interaction (PPI) graphs, which aid in discovering the functional relationships between selected proteins, is frequently used in conjunction with gene expression research. A Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) [14] was used to create PPIs. STRING's algorithm for predicting interactions is implemented by computing scores, which describe the level of confidence for each interaction - that is, how probable STRING considers a connection to be realistic based on the evidence. STRING ranks all genes from 0 to 1 using Spearman's rank correlation [15] (where 1 represents the highest confidence). Genomic characteristics, scientific literature, and other databases, such as KEGG, are used to import interactions into STRING. STRING creates networks with nodes representing proteins and edges representing interactions between them.

### Software

For biological investigations on high-throughput genomic data, we employed R, one of the most efficient and frequently used programming languages in the bioinformatics industry, and its repository of packages, Bioconductor. The GEOquery [16] package was used to download both the gene expression matrix and the annotation file. The limma package [17], which provides methods for working with data derived from microarray or RNA-seq technologies, was used to perform the quality analysis, which included data normalization and linear modeling. Limma uses multiple regression, but it fits a model for each gene separately. During the pairwise comparison, an empirical Bayes approach was utilized, which moderates the standard errors of the calculated log-fold changes. The DEGs identification findings were observed using the EnhancedVolcano package. The following signaling pathway search was provided using clusterProfiler [18] and the KEGG database. The ggplot2 and Rgraphviz packages were used to visualize the found pathways. The STRING database was used to create known and projected protein-protein interactions, and to visualize the resulting graphs. CemiTool was used for the co-expression analysis.

## RESULTS

### DEGs and patients

Out of 48.803 input genes, 380 were found to be DEGs: 229 were upregulated, and 151 were downregulated. The comparison between healthy samples and those with RBCK1 deficiency appeared to have the biggest number of DEGs (119 downregulated genes) (Table 1), in addition to a comparison of RBCK1 and MWS samples (142 upregulated genes). All non-expressed genes were removed, and a new PCA plot was created: RBCK1 and healthy samples were clearly segregated into two clusters in a group of downregulated genes (Supplementary data, Fig. 2), demonstrating the difference between healthy and disease conditions. The data distribution before and after normalisation can be observed in the histograms and boxplots in Figure 2. No difference was found between the healthy and diseased patients, as revealed in principal component analysis (PCA) plot - healthy and MWS samples may be regarded to be split into discrete clusters, which could not be observed in the remaining groups, according to the PCA plot (Supplementary data, Fig. 1). From a biological standpoint, this result may point to some similarities in gene expression patterns in CINCA and MVK diseases. Each column in the heatmap of gene expression data (Supplementary data, Fig. 3) shows the gene expression values in one sample, and each row indicates the gene expression of one specific gene in all samples. The red-green color scheme was used to distinguish between genes with high and low expression levels: red genes have high expression levels, whereas green genes have low expression levels. Contrary to our expectations, the majority of genes were not differentially expressed.

All sample comparisons were done pairwisely, and samples with the same disease type were compared as a single group. Volcano plots (Supplementary data, Figs. 4 and 5), which depict statistical significance (P-value) versus logarithmic magnitude of change (log fold change) of gene expression, show the results of the analysis of differentially expressed genes. An over-representation analysis revealed 91 enriched pathways. Upregulated and downregulated genes from the pairwise comparison, and genes from cemiTool clusters, were included in the input gene list for this phase. According to the findings, some of our genes are involved in the same pathways, preferably metabolic. Even though the number of such genes is generally small (between 1 and 5 in each pathway) (Table 2, Supplementary data, Figs. 6–8).

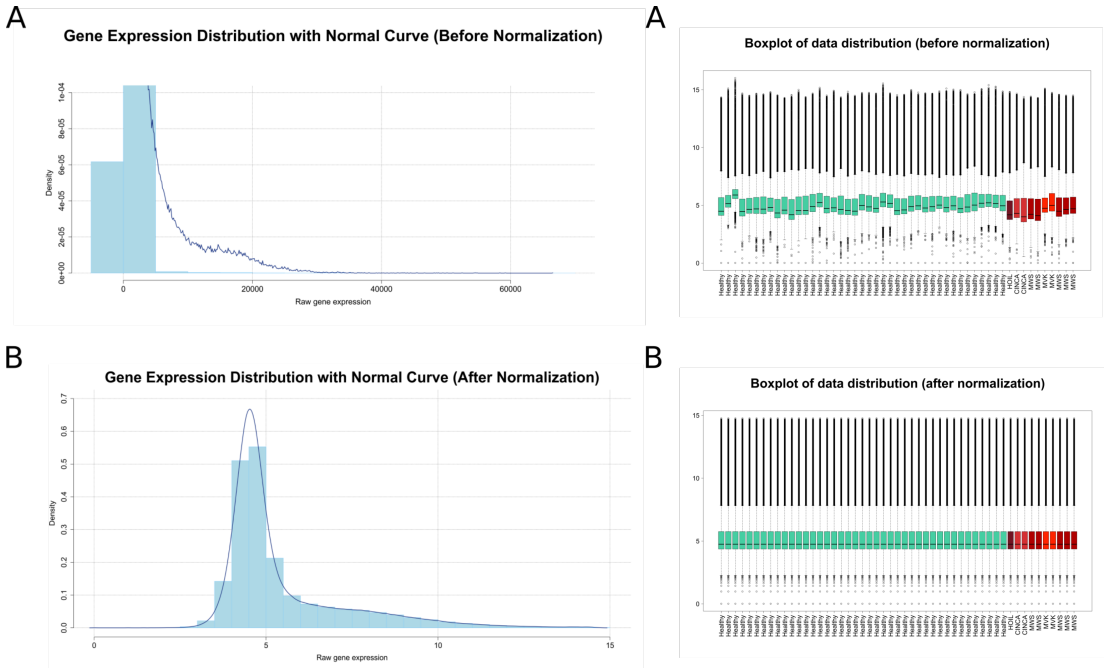


**Fig. 1.** Raw expression data distribution. Histograms and boxplots of raw expression data distribution A: Before normalization. B: After normalization.

**Table 1.** Top 10 downregulated and upregulated DEGs from the comparison between healthy samples and those with RBCK1 deficiency

| **Downregulated DEGs** | | | | **Upregulated DEGs** | | | |
|---|---|---|---|---|---|---|---|
| | Gene Name | P-value | logFC | | Gene Name | P-value | logFC |
| 1 | $CISD2$ | $7.537936e-18$ | $-0.5969954$ | 1 | $HS.551123$ | $6.458791e-13$ | 3.1687011 |
| 2 | $EPB41$ | $4.108343e-16$ | $-0.6256048$ | 2 | $HS.552143$ | $1.583777e-07$ | 1.2937618 |
| 3 | $LOC253012$ | $5.739984e-16$ | $-0.6810695$ | 3 | $FLJ00312$ | $3.640115e-07$ | 1.6176159 |
| 4 | $FAM83A$ | $1.703811e-13$ | $-0.7337758$ | 4 | $HS.19339$ | $6.034732e-07$ | 0.4643016 |
| 5 | $NUP98$ | $5.965986e-13$ | $-0.4364548$ | 5 | $ANKMY2$ | $6.517125e-06$ | 0.4189648 |
| 6 | $CHD2$ | $6.279939e-12$ | $-0.4504437$ | 6 | $RPS29$ | $1.978019e-05$ | 0.4123195 |
| 7 | $RAP1GAP$ | $1.338901e-11$ | $-0.9886351$ | 7 | $HS.531457$ | $7.585137e-05$ | 0.4194266 |
| 8 | $HS.563750$ | $1.539916e-11$ | $-0.5622475$ | 8 | $HS.542923$ | 0.000385 | 1.1559277 |
| 9 | $ABCC13$ | $1.908913e-11$ | $-0.5476611$ | 9 | $HIST1H2BI$ | 0.000278 | 0.9167271 |
| 10 | $MAOA$ | $3.456357e-11$ | $-0.7114969$ | 10 | $PLA2R1$ | 0.000193 | 0.8924394 |

## Involvement of DEGs in key pathways and gene ontologies

The immunological response, inflammatory response, and protein phosphorylation pathways in the biological process category were found to be overrepresented in the GO and pathways derived from co-expressed gene clusters (Table 2, Figure 2). Protein binding and RNA binding pathways were the most enriched in the molecular function category. Finally, the cytoplasm, cytosol, and nucleus were the most enriched routes in the cellular component

category. Supplementary data, Figs. 9–19 shows the most representative GO pathways for co-expression clusters. After loading downregulated genes IDs from the comparison between healthy samples and the ones with RBCK1 deficiency into the WikiPathways database, 425 pathways were collected. The three most remarkable pathways were connected with SARS-CoV-2 (COVID-19): WP5115, WP5039, WP5098, among which were 4 genes from our set: FAM83A, IFI27, NUP98, and TSC1. Furthermore, HP gene, which was in a group of down-regulated genes in comparison between healthy and MWS samples, was found to be involved in a pathway connected with COVID-19. All additional tables are presented in the Supplementary data.

**Table 2.** Example of five GO-terms with the lowest false discovery rate

| GO-terms | Description | Gene ratio | False discovery |
|---|---|---|---|
| GO:0071799 | cellular response to prostaglandin D stimulus | 2/5 | 0.0125 |
| GO:0021796 | cerebral cortex regionalization | 2/7 | 0.0162 |
| GO:0030656 | regulation of vitamin metabolic process | 2/12 | 0.0349 |
| GO:0051712 | positive regulation of killing of cells of other organism | 2/13 | 0.0376 |
| GO:0001829 | trophectodermal cell differentiation | 2/15 | 0.0456 |

### Protein-protein interactions results

Based on clusters formed using cemiTool and DEGs identified earlier, 30 distinct protein-protein interaction graphs were displayed. Only one graph, which depicts an interaction between 54 proteins from one of cemiTool's 14 clusters, showed statistical significance. The lectin-like receptor subfamilies KLRD1, KLRC1, KIR2DL1, KIR2DL2, KIR2DL3, KIR2DL4, KIR3DL2, and KIR3DL3 had the tightest association on the PPI graph. In comparison to healthy people and people with CINCA syndrome, several genes were downregulated. In addition, the genes CXCL8 and CXCL10 were of particular relevance, because they interact with the chemokine receptor, and a deficiency in their activity can lead to crucial errors in cell function or death. The created network's P-value for PPI enrichment was $1.0e-16$.

## DISCUSSION

RBCK1 is a major part of the LUBAC complex and an important part of all intracellular systems of protein ubiquitination. Patients with RBCK1 deficiency have high sensitivity to bacterial infections [1], autoinflammation syndrome, and accumulation of amylopectin in skeletal muscles and myocardium. The main reasons for the observed pathological conditions are related to the direct function of ubiquitination process in response to bacterial infections and the production of cytokines [19]. A recent study proposes the biochemical process of RBCK1 participation in glycogenesis [5]. All of these are three primary reasons that cause clinical manifestations of the disease. In this study, we performed some bioinformatics analysis aiming to find second-order effects that also contribute to the health condition of patients with RBCK1 deficiency.

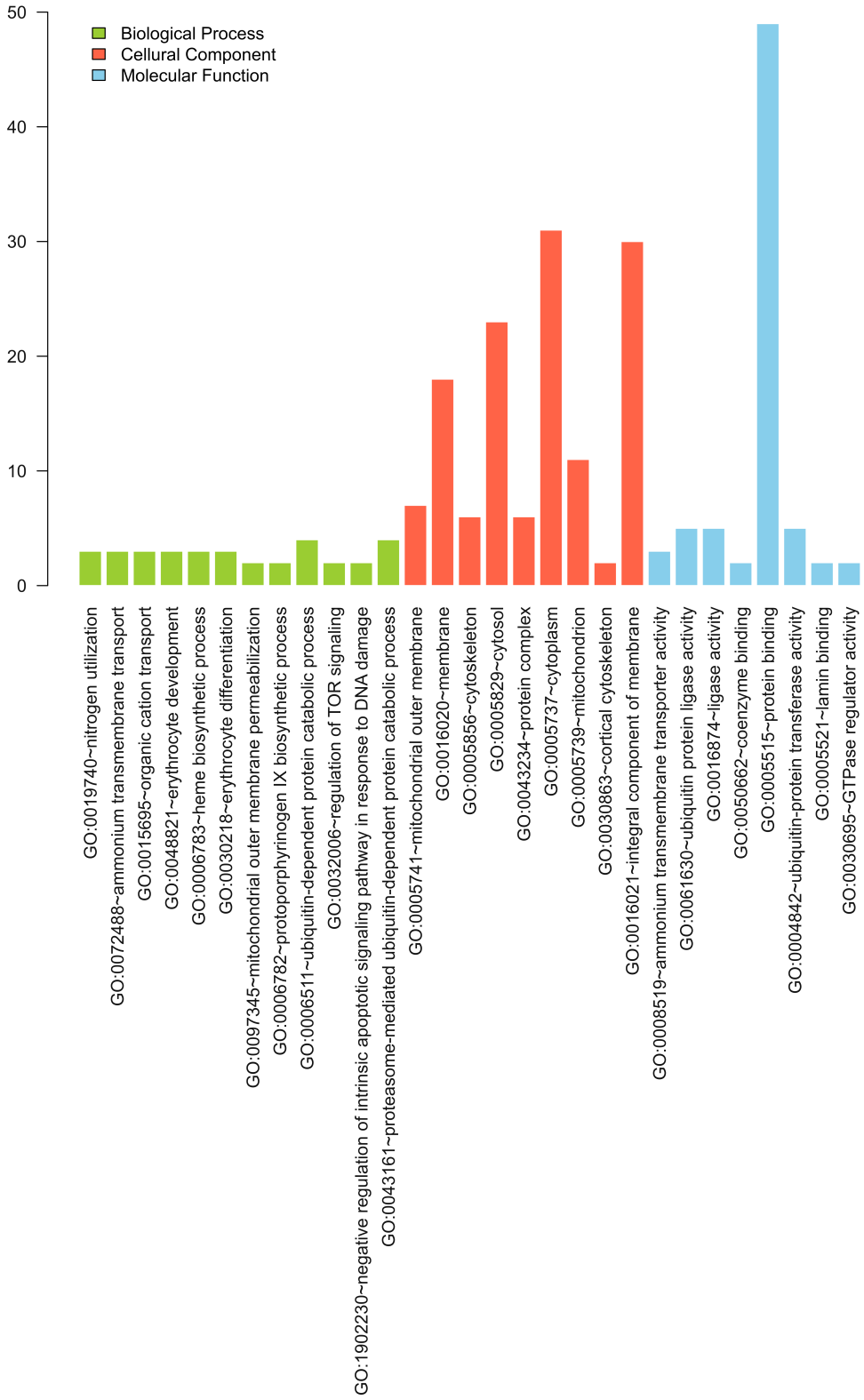


**Fig. 2.** Gene Ontology (GO) barplots representing over-enriched pathways among downregulated genes from the comparison between healthy and RBCK1-deficient patients. GO terms belong to molecular function (MF), cellular component (CC), and biological process (BP) categories.

Bulk mRNA analysis was performed on cells that passed treatment with IL-1b and TNF cytocines for repeat conditions of the acute autoinflammatory syndrome. Patients that provided material for the analysis could have current biological infections. An indirect evidence for this is the increased expression of Histone H2B type 1-C/E/F/G/I (HIST1H2BI). According to UniProt:P62807, this histone is involved in several biological processes: antibacterial humoral response, antimicrobial humoral immune response mediated by an antimicrobial peptide, defense response against Gram-positive bacteria, and innate immune response in the mucosa.

RBCK1 is part of the LUBAC complex. It is involved in cellular energy metabolism at the level of protein ubiquitination and autophagy, which is an evolutionarily ancient pathway of intracellular homeostasis [20]. In addition, RBCK1 is known to be closely associated with Nuclear factor-kB (NF-kB) activation, thus participating in pattern recognition receptor signaling. Regarding NF-kB pathways, RBCK1 has three effects on the expression of this protein. RBCK1 kills NF-kB inhibitor, and RBCK1 increases IkB kinase activity by enhancing the breakdown of NF-kB inhibitor, RBCK1 provides the processing of NF-kB from precursors. Thus, RBCK1 deficiency results in a defect of the NF-kB pathway. NF-kB is the main pro-inflammatory transcription factor; it is activated by descending signals from pattern recognition receptors, and the increased susceptibility to bacterial infections is observed due to the absence of adequate immune system response to Toll-like receptors [19]. Bacterial infection in patients with RBCK1 deficiency cause the overexpression of cytokines.

The effect of RBCK1 on glucose metabolism is mostly unclear. Histological studies show that patients with RBCK1 deficiency accumulate amylopectin in addition to normal glycogen [21]. Amylopectin is a polysaccharide that is less branched than glycogen. Amylopectin cannot be transformed back to glycose by human cells. Amylopectin over-accumulation causes cell death. Because of poor solubility in water, amylopectin aggregates cannot be removed from the tissue and strongly affect their function. Patients with RBCK1 deficiency usually have syndromes of myopathy and cardiomyopathy. According to [22], RBCK1 may perform the ubiquitination of polysaccharides inside lysosomes. RBCK1 deficiency causes poor branching of polysaccharides and the accumulation of amylopectin instead of glucogen.

Cells of patients with RBCK1 deficiency try to convert all glycogen and amylopectin that are stored inside lysosomes into glucose. This suggestion is based on observed changes in mammalian target of rapamycin (mTOR), Pi3k, and AKT pathways. The transformation of glucose into glycogen is controlled by various enzymes, but insulin is the one that initiates it. The insulin receptor downregulates the activity of the protein phosphatase 2A via the Pi3k-Akt-mTOR and mitogen-activated protein kinases signaling pathways [23]. The transformation of glycogen in glucose is regulated by glucagon. Glucagon-related pathways are also affected by protein phosphatase 2A, but this effect is contrary to the effect of insulin. The transformation of glycogen into glucose occures in the cytosol and autophagosomes. The second metabolic process requires a Ca+2, which works as a signaling ion.

Patients with RBCK1 deficiency show a decreased expression of family with sequence similarity 83 (FAM83), family with sequence similarity 83 member A (FAM83A), and phosphatidylinositol-4-phosphate 5-kinase type 1 beta (PIP5K1B) genes. According to [24], all of them are involved in the activation of the PI3K/AKT/TOR signaling cascade. Furthermore, a synergistic decrease in the production of Intelectin 1 occurs, which activates the AKT cascade in the absence of insulin [25]. The increasing level of Ca+2 into autophagosomes is related to a decreased expression of CDGSH iron sulfur domain 2 (CISD2). This protein is related to intracellular calcium and takes participants in the transformation of glycogen into glucose in autophagosomes and the transformation of free glucose in lysosomes into glucose 6-phosphate [26].

We suppose that the excessive conversion of glycogen into glucose may be a compensatory

response to the excessive accumulation of polysaccharides. This response protects muscular fibroblast from cellular death caused by an excessive amount of lysosomes and autophagosomes. However, amylopectin cannot be utilized inside the cell, and the storage of normal glycogen becomes depleted. This made it impossible to provide glucose into muscle cells during physical activity. Thus, a downregulation of the Pi3k-Akt-mTOR pathway may be a secondary reason for myopathy and cardiomyopathy in addition to amylopectin accumulation in patients with RBCK1 deficiency.

According to the observed gene expression, the mechanisms of apoptosis must have been changed in the cells. The expression of E3 ubiquitin-protein ligase rififylin was decreased. This protein negatively regulates cysteine-type endopeptidase activity, which is involved in the execution phase of apoptosis, extrinsic apoptotic signaling pathway via death domain receptors, signal transduction by p53 class mediator, and tumor necrosis factor-mediated signaling pathway. Thus, decreasing E3 ubiquitin-protein ligase rififylin may activate all these apoptotic pathways. By contrast, a reduced expression was found in several pro-apoptotic factors and proteins involved in the apoptosis, such as TNF-related apoptosis-inducing ligand, interferon alpha inducible protein 27, protein phosphatase 1A, and adenovirus E1B 19 kDa protein-interacting protein 3-like.

Patients with RBCK1 deficiency do not show strong symptoms of reduced lymphocytes motility and adhesion. However, they exhibit many changes in the expression of genes related to lymphocytes' cytoskeleton, adhesion, integrin-related pathways, and microtubule-based hair-like organelles (cilia) in peripheral mononuclear blood cells.

The overexpression of ANKMY2 (ankyrin repeat And MYND domain-containing 2) leads to the increased transport of signaling proteins in cilia. The expression levels of C14ORF45 (basal body orientation factor 1) responsible for orientation, and tetratricopeptide repeat domain 25 responsible for dynein arm attachment in cilia are reduced. Moreover, we observed a decreased expression of proteins responsible for mitochondrial transport along microtubules (RHOT1, ras homolog family member T1) and microtubule assembly (TBCEL, tubulin folding cofactor E like), as well as decreased expression of kinesin for mitochondrial transport along microtubules (KLC3, kinesin light chain 3).

Patients with RBCK1 deficiency have reduced activity of the Rho pathway, which involves the following downregulated genes: catenin alpha like 1, proline-rich 5, and KANK2 (KN motif and ankyrin repeat domains 2). Proline-rich 5 gene activates the Rho pathway via the mTORC2 cascade, and KANK2, which plays part in the Rho pathway regulation [27, 28]. This pathway is responsible for increasing actin polymerization and the contractile function of actin filaments [29, 30]; it is directly related to cell motility and integrins. A decrease in the Rho pathway activity can lead to changes in the cytoskeleton, impaired actin polymerization, and changes in the filament activity. We suppose that the downregulation of the Rho pathway and cilia may be a secondary cause of immunological dysfunction due to impaired leukocyte motility and adhesion.

RBCK1 deficiency modifies the expression of some genes that function in the pathogenesis of other diseases are well known. Dilated cardiomyopathy in RBCK1 deficiency is described by the accumulation of polyglucose bodies in the myocardium. At the same time, the analysis of the presented dataset revealed an increased expression of FOXO1 (Forkhead Box O1) gene in fibroblasts. The ability of the FOXO1 gene to aggravate glucophagy in diabetic cardiomyopathy, and the role of FOXO1 in dilated cardiomyopathy in type 2 diabetes mellitus have been reported [23]. FOXO1 activation is dependent on the AKT signaling pathway, which is downregulated in RBCK1 deficiency. Thus, reinforced glycophagy in RBCK1 deficiency may be a secondary reason for dilated cardiomyopathy.

RBCK1 deficiency has some similarities with diseases such as Friedreich's ataxia. Despite

the neurodegenerative nature of Friedreich's ataxia, it also causes hypertrophic cardiomyopathy and changes in the fibroblast cytoskeleton structure [31]. Friedreich's Ataxia negatively affects the level of PIP5K1B. PIP5K1B downregulation results in a decreased synthesis of frataxin, which is involved in iron biochemistry [32]. Notably, changes have been observed in biochemical pathways with the participation of iron. The research on the regulation of cellular iron homeostasis in Drosophila showed that it depends on glycogen branching enzyme (AGBE) [33]. According to previous research on human cell culture, it is known that iron regulatory protein 2 is a modulator of iron metabolism, and it depends on RBCK1 functions [34]. We suppose that intracellular iron metabolism is also related to glycogen biosynthesis and the clinical manifestations of the disease.

Concerning the antiviral activity of the immune response in RBCK1-deficient patients, increased susceptibility is a rare case, according to clinical data. Our study has demonstrated the involvement of the signaling pathway responsible for the response to coronavirus infection. However, no evidence has shown that RBCK1-deficient patients are at risk for COVID-19. It is known that overexpression of RBCK1 is linked with SARS-CoV infection [35]. This has an explanation related to the ubiquitination of interferon regulatory factor 3, a significant signaling molecule linked with Toll/IL-1R domain-containing adaptor inducing IFN and Toll-like receptor 3.

In 2021, the relationship between toll-like receptor 3, interferon regulatory factor 3, and RBCK1 was demonstrated [36, 35]. Therefore, one should not expect increased susceptibility to SARS-CoV infection in this category of patients in the reverse situation. Our own observations of various patients with primary immunodeficiency in the Sverdlovsk oblast region of the Russian Federation are consistent with international data, which confirm that most patients with inborn immunity errors tolerate COVID-19 no more severely, and sometimes more easily than patients with normally functioning immunity. The only exception is patients with a deficiency of IFN-I-type signaling pathway molecules.

## CONCLUSIONS

In this study, we performed the bioinformatics analysis to find additional pathways that affect the health condition of patients with RBCK1 deficiency.

The identification of DEGs revealed a three threefold higher number of downregulated genes than upregulated ones in the comparison between healthy and RBCK1-deficient individuals. Some highly enriched pathways, which were modified by altered gene expression were mTOR, PI3K/AKT, Rho, and Nf-kB signaling pathways. The following group of genes were downregulated: CISD2, EPB41, FAM83A, NUP98, CHD2, and RAP1GAP. Among upregulated genes were: ANKMY2, FLJ00312, RPS29, HIST1H2BI, and PLA2R1. The subsequent PPI construction revealed tight connections between killer cell lectin-like receptor subfamilies (KLRC3, KIR2DL1, KLRC1, KLRD1, and KIR3DL1-KIR3DL4), which play a big role in innate immunity, and between NK cells.

We suppose that RBCK1 deficiency causes extensive glucophagy in some types of cells. This occurs as a compensatory mechanism that protects cells from death because excessive amounts of stored polysaccharides negatively affect cellular function. However, amylopectin cannot be used by the cell, and the storage of normal glycogen becomes depleted. The depletion of normal glycogen storage causes secondary effects that aggravate myopathy and dilated cardiomyopathy in patients with RBCK1 deficiency.

We suppose that patients with RBCK1 deficiency also have decreased motility and adhesion of lymphocytes. Observed changes lead to the secondary effect that aggravates sensitivity to bacterial infections in patients with RBCK1 deficiency.

## REFERENCES


1. de Jesus A.A., Goldbach-Mansky R. Genetically defined autoinflammatory diseases. *Oral Diseases.* 2016. V. 22. № 7. P. 591–604. doi: 10.1111/odi.12448
2. McDermott F., Aksentijevich I., Galon J., McDermott E.M., Ogunkolade B.W., Centola M., Mansfield E., Gadina M., Karenko L., Pettersson T. et al. Germline mutations in the extracellular domains of the 55 kda tnf receptor, tnfr1, define a family of dominantly inherited autoinflammatory syndromes. *Cell.* 1999. V. 97. № 1. P. 133–144. doi: 10.1016/s0092-8674(00)80721-7
3. Touitou I., Lesage S., McDermott M., Cuisset L., Hoffman H., Dode C., Shoham N., Aganna E., Hugot J.-P., Wise C. et al. Infevers: an evolving mutation database for autoinflammatory syndromes. *Hum. Mutat.* 2004. V. 24. № 3. P. 194–198. doi: 10.1002/humu.20080
4. Perazzio S.F., Allenspach E.J., Eklund K.K., Varjosalo M., Shinohara M.M., Torgerson T.R., Seppänen M.R.J. Behçet disease (bd) and bd-like clinical phenotypes: Nf-b pathway in mucosal ulcerating diseases. *Scand. J. Immunol.* 2020. V. 92. № 5. P. e12973. doi: 10.1111/sji.12973
5. Kelsall I., McCrory E.H., Xu Y., Scudamore C., Nanda S.K., Mancebo-Gamella P., Wood N.T., Knebel A., Matthews S.J., Cohen P. *Hoil-1-catalysed ubiquitylation of unbranched glucosaccharides and its activation by ubiquitinoligomers*: preprint of bioRxiv. 2021.
6. Bolstad M., Irizarry R.A., Astrand M., Speed T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* 2003. V. 19. № 2. P. 185–193. doi: 10.1093/bioinformatics/19.2.185
7. Efron B., Tibshirani R., Storey J.D., Tusher V. Empirical bayes analysis of a microarray experiment. *Journal of the American Statistical Association.* 2001. V. 96. № 456. P. 1151–1160. doi: 10.1198/016214501753382129
8. Russo P., Ferreira G.R., Cardozo L.E., Bürger M.C., Arias-Carrasco R., Maruyama S.R., Hirata T.D.C., Lima D.S., Passos F.M., Fukutani K.F., Lever M., Silva J.S., Maracaja-Coutinho V., Nakaya H.I. Cemitool: a bioconductor package for performing comprehensive modular co-expression analyses. *BMC Bioinformatics.* 2018. V. 19. № 1. doi: 10.1186/s12859-018-2053-1
9. Khatri P., Sirota M., Butte A.J. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput. Biol.* 2012. V. 8. № 2. P. e1002375. doi: 10.1371/journal.pcbi.1002375
10. Falcon S., Gentleman R. Hypergeometric Testing Used for Gene Set Enrichment Analysis. In: *Bioconductor Case Studies. Use R!*. New York: Springer, 2008. P. 207–220. doi: 10.1007/978-0-387-77240-0_14
11. Kanehisa M., Goto S. Kegg: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research.* 2000. V. 28. № 1. P. 27–30. doi: 10.1093/nar/28.1.27
12. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong *Nucleic Acids Research*. 2019. V. 47. P. D330–D338. doi: 10.1093/nar/gky1055
13. Martens M., Ammar A., Riutta A., Waagmeester A., Slenter D.N., Hanspers K., Miller R.A., Digles D., Lopes E.N., Ehrhart F. et al. WikiPathways: connecting communities. *Nucleic Acids Research.* 2021. V. 49. № D1. P. D613–D621. doi: 10.1093/nar/gkaa1024
14. Szklarczyk D., Gable A.L., Lyon D., Junge A., Wyder S., Huerta-Cepas J., Simonovic M., Doncheva N.T., Morris J.H., Bork P. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research.* 2018. V. 47. № D1. P. D607–D613. doi: 10.1093/nar/gky1131
15. Zar H. Spearman rank correlation. *Encyclopedia of Biostatistics*. 2005. doi: 10.1002/0470011815.b2a15150

16. Davis S., Meltzer P.S. Geoquery: a bridge between the gene expression omnibus (geo) and bioconductor. *Bioinformatics.* 2007. V. 23. № 14. P. 1846–1847. doi: 10.1093/bioinformatics/btm254
17. Ritchie M., Phipson B., Wu D., Hu Y., Law C.W., Shi W., Smyth G.K. *limma* powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Research.* 2015. V. 43. № 7. P. e47. doi: 10.1093/nar/gkv007
18. Yu G., Wang L.G., Han Y., He Q.Y. clusterProfiler: an r package for comparing biological themes among gene clusters. *OMICSMay.* 2012. V. 16. № 5. P. 284–287. doi: 10.1089/omi.2011.0118
19. Liu T., Zhang L., Joo D., Sun. Nf-b signaling in inflammation. *Sig. Transduct. Target. Ther.* 2017. V. 2. P. 17023. doi: 10.1038/sigtrans.2017.23
20. Iwai K. Lubac-mediated linear ubiquitination: a crucial regulator of immune signaling. *Proceedings of the Japan Academy Series B, Physical and biological sciences.* 2021. V. 97. № 3. P. 120–133. doi: 10.2183/pjab.97.007
21. Thomsen C., Malfatti E., Jovanovic A., Roberts M., Kalev O., Lindberg C., Oldfors A. Proteomic characterization of polyglucosan bodies in skeletal muscle in rbck1 deficiency. *Neuropathology and Applied Neurobiology.* 2021. V. 10. doi: 10.1111/nan.12761
22. Nilsson J., Schoser B., Laforet P., Kalev O., Lindberg C., Romero N.B., López M.D., Akman H.O., Wahbi K., Iglseder S. et al. Polyglucosan body myopathy caused by defective ubiquitin ligase rbck1. *Ann. Neurol.* 2013. V. 74. № 6. P. 914–919. doi: 10.1002/ana.23963
23. Li X., Wan T., Li Y. Role of foxo1 in regulating autophagy in type 2 diabetes mellitus (review). *Experimental and Therapeutic Medicine.* 2021. V. 22. № 707. doi: 10.3892/etm.2021.10139
24. Porciello N., Kunkl M., Viola A., Tuosto L. Phosphatidylinositol 4-phosphate 5-kinases in the regulation of t cell activation. *Frontiers in Immunology.* 2016. V. 7. P. 186. doi: 10.3389/fimmu.2016.00186
25. Yang R.Z., Lee M.J., Hu H., Pray J., Wu H.-B., Hansen B.C., Shuldiner A.R., Fried S.K., McLenithan J.C., Gong D.-W. et al. Identification of omentin as a novel depot-specific adipokine in human adipose tissue: possible role in modulating insulin action. *Am. J. Physiol. Endocrinol. Metab.* 2006. V. 290. № 6. P. E1253–E1261. doi: 10.1152/ajpendo.00572.2004
26. Zhao H., Tang M., Liu M., Chen L. Glycophagy: An emerging target in pathology. *Clin. Chim. Acta.* 2018. V. 484. P. 298–303. doi: 10.1016/j.cca.2018.06.014
27. Zhu Y., Kakinuma N., Wang Y., Kiyama R. Kank proteins a new family of ankyrin-repeat domain-containing proteins. *Biochim. Biophys. Acta.* 2008. V. 1780. № 2. P. 128–133. doi: 10.1016/j.bbagen.2007.09.017
28. Gee H. Y., Zhang F., Ashraf S., Kohl S., Sadowski C.E., Vega-Warner V., Zhou W., Lovric S., Fang H., Nettleton M. et al. KANK deficiency leads to podocyte dysfunction and nephrotic syndrome. *J. Clin. Invest.* 2015. V. 125. № 6. P. 2375–2384. doi: 10.1172/JCI79504
29. Thumkeo D., Watanabe S., Narumiya S. Physiological roles of rho and rho effectors in mammals. *Eur. J. Cell Biol.* 2013. V. 92 № 10–11. P. 303–315. doi: 10.1016/j.ejcb.2013.09.002
30. Dai Y., Luo W., Chang J. Rho kinase signaling and cardiac physiology. *Curr. Opin. Physiol.* 2018. V. 1. P. 14–20. doi: 10.1016/j.cophys.2017.07.005
31. Bayot A., Reichman S., Lebon S., Csaba Z., Aubry L., Sterkers G., Husson I., Rak M., Rustin P. Cis-silencing of PIP5K1B evidenced in friedreich's ataxia patient cells results in cytoskeleton anomalies. *Hum. Mol. Genet.* 2013. V. 22. № 14. P. 2894–2904. doi: 10.1093/hmg/ddt144
32. Hanson E., Sheldon M., Pacheco B., Alkubeysi M., Raizada V. Heart disease in friedreich's

ataxia. *World J. Cardiol.* 2019. V. 11. № 1. P. 1–12. doi: 10.4330/wjc.v11.i1.1
33. Huynh N., Ou Q., Cox P., Lill R., King-Jones K. Glycogen branching enzyme controls cellular iron homeostasis via iron regulatory protein 1 and mitoneet. *Nature Communications.* 2019. V. 10. № 1. P. 5463. doi: 10.1038/s41467-019-13237-8
34. Yamanaka K., Ishikawa H., Megumi Y., Tokunaga F., Kanie M., Rouault T.A., Morishima I., Minato N., Ishimori K., Iwai K. Identification of the ubiquitin-protein ligase that recognizes oxidized irp2. *Nat. Cell Biol.* 2003. V. 5. № 4. P. 336–340. doi: 10.1038/ncb952
35. Zhang M., Tian Y., Wang R., Gao D., Zhang Y., Diao F.-C., Chen D.-Y., Zhai Z.-H., Shu H.-B. Negative feedback regulation of cellular antiviral signaling by rbck1-mediated degradation of irf3. *Cell Research.* 2008. V. 10. № 1038. P. 1096–1104. doi: 10.1038/cr.2008.277
36. Menezes M.C.S., Veiga A.D.M., de Lima T.M., Ariga S.K.K., Barbeiro H.V., de Lucena Moreira C., Pinto A.A.S., Brandao R.A., Marchini J.F., Alencar J.C. et al. Lower peripheral blood toll-like receptor 3 expression is associated with an unfavorable outcome in severe covid-19 patients. *Sci. Rep.* 2021. V. 11. P. 15223. doi: 10.1038/s41598-021-94624-4

Accepted 24.06.2022.
Revised 17.07.2022.
Published 09.08.2022.

*Математическая биология и биоинформатика*
*doi: 10.17537/2022.17.174*

========== **БИОИНФОРМАТИКА** ==========

# Дополнительные пути патогенеза RBCK1 дефицита

**Демичева Е.И.**[1,2], **Шинвари Х.**[2], **Ушенин К.С.**[1,2], **Болков М.А.***[1,2]

[1]*Институт иммунологии и физиологии УрО РАН, Екатеринбург, Россия*
[2]*Уральский федеральный университет, Екатеринбург, Россия*

***Аннотация.*** RBCK1 дефицит – редкое врожденное наследственное заболевание аутовоспалительной природы, которое вызывает нарушение воспалительных процессов на молекулярном уровне. Данное заболевание имеет три основных клинических проявления: повышенную чувствительность к бактериальным инфекциям, синдром аутовоспаления и накопление амилопектина в скелетных мышцах. Амилопектиноз вызывает миопатию и кардиомиопатию. Патогенез заболевания мало изучен и может включать незамеченные взаимосвязи. В данной работе был проведен анализ дифференциальной экспрессии генов у пациентов с RBCK1 дефицитом и тремя другими аутовоспалительными заболеваниями. Идентификация дифференциально экспрессированных генов выявила большое число генов с пониженной экспрессией, которые вовлечены в активацию ключевых метаболических и имунных путей, включающих Nf-kb и Pi3k-Akt-mTOR. Сигнальные пути были проанализированы с использованием KEGG (Kyoto Encyclopedia of Genes and Genomes) и ресурса Gene Ontology. Предсказанные белок-белковые взаимодействия были получены из STRING (Search Tool for the Retrieval of Interacting proteins database). Помимо первичного вовлечения в патологию RBCK1 дефицита, некоторые пути пониженной регуляции усиливают симптомы миопатии, кардиомиопатии и бактериальных заболеваний. Изученные сигнальные пути могут служить новыми мишенями в разработке компенсаторной терапии для пациентов с RBCK1 дефицитом.

***Ключевые слова:*** *профилирование генетической экспрессии, аутовоспалительные заболевания, RBCK1, сигнальные пути, белок-белковые взаимодействия.*