

## RESEARCH ARTICLE

# A comparative proteomic study of plasma in Colombian childhood acute lymphoblastic leukemia

Sandra Isabel Calderon-Rodríguez<sup>1</sup>, María Carolina Sanabria-Salas<sup>1,2</sup>, Adriana Umaña-Pérez<sup>1\*</sup>

**1** Universidad Nacional de Colombia—Sede Bogotá, Facultad de Ciencias, Departamento de Química, Grupo de investigación en Hormonas, Bogotá, D.C., Colombia, **2** Subdirección de Investigaciones, Instituto Nacional de Cancerología de Colombia, Bogotá, D.C., Colombia

\* [yaumanap@unal.edu.co](mailto:yaumanap@unal.edu.co)



## OPEN ACCESS

**Citation:** Calderon-Rodríguez SI, Sanabria-Salas MC, Umaña-Pérez A (2019) A comparative proteomic study of plasma in Colombian childhood acute lymphoblastic leukemia. PLoS ONE 14(8): e0221509. <https://doi.org/10.1371/journal.pone.0221509>

**Editor:** Jon M. Jacobs, Pacific Northwest National Laboratory, UNITED STATES

**Received:** November 9, 2018

**Accepted:** August 8, 2019

**Published:** August 22, 2019

**Copyright:** © 2019 Calderon-Rodríguez et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** This study was funded by the program Colombia Corre por la investigación en cáncer (Contract number 2012-15116), the National University of Colombia Research Council (DIB 15116) and Colciencias (Founding 2012-567).

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Acute lymphoblastic leukemia (ALL) is the most common childhood cancer. Owing to the incorporation of risk-adapted therapy and the arrival of new directed agents, the cure rate and survival of patients with ALL have improved dramatically, get near to 90%. In Latin American countries, the mortality rates of ALL are high, for example in Colombia, during the last decade, ALL has been the most prevalent cancer among children between 0–14 years of age. In the face of this public health problem and coupled with the fact that the knowledge of the proteome of the child population is little, our investigation proposes the study of the plasma proteome of Colombian children diagnosed with B-cell ALL (B-ALL) to determine potential disease markers that could reflect processes altered by the presence of the disease or in response to it. A proteomic study by LC-MS/MS and quantification by label-free methods were performed in search of proteins differentially expressed between healthy children and those diagnosed with B-ALL. We quantified a total of 472 proteins in depleted blood plasma, and 25 of these proteins were differentially expressed (fold change >2, Bonferroni-adjusted P-values <0.05). Plasma Aggrecan core protein, alpha-2-HS-glycoprotein, coagulation factor XIII A chain and gelsolin protein were examined by ELISA assay and compared to shotgun proteomics results. Our data provide new information on the plasma proteome of Colombian children. Additionally, these proteins may also have certain potential as illness markers or as therapeutic targets in subsequent investigations.

## Introduction

Pediatric acute leukemia is a genetic and phenotypically heterogeneous disease. Accordingly to the phenotype, acute lymphoblastic leukemia (ALL) may involve B-cell (B-ALL) or T-cell progenitors (T-ALL) [1]. ALL is the most common childhood malignancy, accounting for 25% of all childhood cancers [2]. Of these cases, 80 to 85% are classified as B-lineage ALL [3,4]. The incidence rate of new cases of childhood ALL increases every year; fortunately, strides have

been made in the management of childhood ALL over the past 50 years; this management has resulted in the improvement of cure rates from 10% to 90% in developed countries [2]. However, of the approximately 160,000 children and teenagers diagnosed with cancer every year worldwide, 80% live in low- and middle-income countries (LMICs) where access to quality care is limited and the chances of being cured are low [5]. In addition to the difficulties in diagnosing and treating the disease, as well as inadequate medical facilities and other socio-economic factors affecting the incidence and survival of patients with ALL in poor countries. The incidence rates of ALL in several Latin American countries are the highest in the world [6]. In Colombia, it is estimated that the mortality rate exceeds 50% of the cases diagnosed with pediatric ALL [7].

Currently, the definitive diagnosis of ALL is carried out by cytogenetic and immunophenotype analysis in a bone marrow sample, in a stage where the disease has clearly manifested [8]. So far there is no early diagnostic test for ALL, which would be very useful in countries where access to health may be limited. Thus, we are interested in having a biomarker in peripheral blood against a bone marrow aspirate is a highly desirable objective and with great potential in the design of screening or screening tests.

The proteomics now provides a reliable alternative for the discovery of new biomarkers of diagnosis [9], the body fluids have great emerging potential in biomarker studies, especially those that can be collected by non-invasive or minimally invasive means (9–11). For example, the blood is considered a complex fluid tissue that encompasses cells and extracellular fluid where the variety of possible candidate biomarkers is significant (10,12).

The aim of this study was to examine the B-ALL plasma proteome, since it is the leukemia subtype with the highest incidence in Colombia, focused on the search for potential biomarkers to facilitate early diagnosis of ALL, improving the opportunity of treatment.

## Materials and methods

### Patients and samples

Plasma samples were collected from six pediatric patients with B-ALL before induction therapy. None of the subjects had received any prior treatment (Table 1). The diagnosis for each patient was based on morphological, immunophenotypic, and genetic tests. The plasma samples used as controls were obtained from six healthy children in the same age range. Blood samples were collected into Vacutainer K3 EDTA tubes. Plasma was separated by centrifugation and subsequently aliquoted into 0.5-mL tubes for cryopreservation at  $-80^{\circ}\text{C}$ . Plasma protein concentrations were determined using a BCA protein assay (Thermo Scientific Rockford, IL, USA). The study was conducted with the approval of the Research Ethics Committee at the National University of Colombia and the National Institute of Cancer of Colombia (Protocol

**Table 1. Clinical characteristics of B-ALL patients.**

Clinical features	Control	B-ALL
Age (years)	6 (3–9)	4.5 (2–9)
Sex (male/female)	6/0	4/2
White Blood Cell (WBC) count $\times 10^9/\text{L}$	-	3.89 (0.47–112.8)
Hemoglobin (g/dL)	-	8.3 (5.02–10.9)
%Blast Bone marrow	-	61 (35–80)

Data are presented as median (minimum and maximum range) unless the sex.

<https://doi.org/10.1371/journal.pone.0221509.t001>

Number INC GT00035). Informed consents were obtained from the patients before the samples were collected for analysis.

### Depletion of high-abundance proteins from plasma

To reduce albumin and IgG, the most abundant plasma proteins, 30  $\mu$ L of plasma was depleted with ProteoPrep (Immunoaffinity Albumin and IgG Depletion kit, Sigma-Aldrich, St. Louis, MO, USA). The depletion column was equilibrated using the equilibration buffer provided in the kit, and plasma samples were depleted according to the instructions by the manufacturer. The amount of plasma protein collected was determined using a BCA protein assay (Thermo Scientific Rockford, IL, USA). The immunodepleted proteins were subjected to total protein precipitation through the addition of cold TCA to obtain a concentration of 10%. The protein solutions were then mixed and stored for 1 hour at  $-10^{\circ}\text{C}$  and subsequently high-speed centrifuged at 8500 g and  $4^{\circ}\text{C}$  for 30 min. The obtained pellet was washed three times with an ice-cold 90:10 acetone-water mixture. The samples were then centrifuged at 11,000 g for 5 minutes at  $4^{\circ}\text{C}$ , and the protein pellet was air-dried and stored at  $-70^{\circ}\text{C}$  until analysis.

### In-solution tryptic digestion

The resulting protein pellet was solubilized in 100  $\mu$ L of 6 M urea in 50 mM  $\text{NH}_4\text{HCO}_3$ . Dithiothreitol (DTT) was added to a concentration of 5 mM and the samples were incubated for 30 min at  $37^{\circ}\text{C}$ . Next, 20 mM iodoacetamide (IAA) was added to an in-solution concentration of 15 mM, and the samples were incubated for 30 min at room temperature; this incubation was followed by the addition of 20  $\mu$ L of 15 mM DTT over 10 min, and the samples were held at room temperature for 4 h. The samples were then diluted to a urea concentration of  $<1$  M by the addition of 550  $\mu$ L of 25 mM  $\text{NH}_4\text{HCO}_3$ . Trypsin (Promega, USA) was next added in a 1:25 enzyme-to-protein ratio and the samples were incubated overnight at  $37^{\circ}\text{C}$ . The following day, the samples were desalted using C18 macro spin columns (Nest Group) and dried by vacuum centrifugation.

### Nano-LC-MS/MS analysis

LC separation was done on a nano Proxeon Easy-nLCTM II HPLC (Thermo Scientific, Waltham, MA) with a Proxeon nanospray source. The digested peptides were reconstituted in 2% acetonitrile/0.1% trifluoroacetic acid and roughly 3  $\mu$ g of each sample was loaded onto a 100-micron x 25 mm Magic C18 100 $\text{\AA}$  5U reverse phase trap where they were desalted online before being separated on a 75-micron x 150 mm Magic C18 200 $\text{\AA}$  3U reverse phase column. Each sample was injected three times. Peptides were eluted using a gradient of 0.1% formic acid (A) and 100% acetonitrile (B) with a flow rate of 300 nL/min. A 90-minute gradient was run with 5% to 35% B over 70 minutes, 35% to 80% B over 8 minutes, 80% B for 1 minute, 80% to 5% B over 1 minute, and finally held at 5% B for 10 minutes. Each of the gradients was followed by a 1h column wash. Mass spectra were collected at the Davis Proteomics Core, University of California (<http://proteomics.ucdavis.edu>), on an Orbitrap Q Exactive Plus mass spectrometer (Thermo Fisher Scientific) in a data-dependent mode with one MS precursor scan followed by 15 MS/MS scans. A dynamic exclusion of 15 seconds was used. MS spectra were acquired with a resolution of 70,000 and a target of  $1 \times 10^6$  ions or a maximum injection time of 30ms. MS/MS spectra were acquired with a resolution of 17,500 and a target of  $5 \times 10^4$  ions or a maximum injection time of 50ms. Peptide fragmentation was performed using higher-energy collision dissociation (HCD) with normalized collision energy (NCE) value of 27. Unassigned charge states, as well as +1 and ions  $>+5$ , were excluded from MS/MS fragmentation.

## Protein identification and quantification

The raw data generated from the LC-MS/MS analysis were examined using database searches performed with Proteome Discoverer version 1.4 (Thermo Fisher Scientific), cross-referenced with results from UniProtKB/SwissProt (2016\_05). The search parameters were defined as a precursor mass tolerance of 20 ppm and a fragment mass tolerance of 0.6 Da. The enzyme used was trypsin and one missed cleavage site was allowed. The carbamidomethylating of cysteines was defined as a fixed modification, while protein N-terminal acetylation and methionine oxidation were defined as variable modifications for database searching. The percolator program implemented in Proteome Discoverer was used to calculate the false discovery rate (FDR) of the identified peptides, and only peptides with FDR <0.01 were considered.

Progenesis QI for proteomics (v 3.0, Nonlinear Dynamics, New Castle, UK) was used for the performance of ion intensity-based label-free quantification. The retention times of eluting peptides from all of the samples in the experiment were aligned to a selected reference run, and only MS/MS peaks with a charge of 2+ to 5+ were considered for the total number of features (signal at one particular retention time and m/z). Only the five most intense spectra per feature were included. After the alignment and feature filtering, the raw abundances of all features were normalized against the total intensity to correct for the experimental variations, and a comparison of features between groups was performed by one-way analysis of variance (ANOVA,  $P < 0.05$  for statistical significance). Type-I errors were controlled by FDR with the q value set at 0.02. The associated unique peptide ion intensities for a specific protein were then summed to generate an abundance value. Search results from Proteome Discoverer were imported into Progenesis LC-MS to combine peptide quantification and identification. Only unique peptides for a corresponding protein were used for quantification. The dataset from the Progenesis analysis is shown in Supporting Information (S1 and S2 Tables).

Statistical analysis was performed using the MSStats and ROTS packages in R statistics [10–12]. Samples were annotated into two respective conditions (Control and B-ALL). Each sample was run three times in LC-MS/MS. For more details, see Supporting Information (S3 and S4 Tables).

## Bioinformatic analysis

To further understand the biological relevance of the differentially expressed proteins, we performed functional enrichment analysis using ClueGO (13). ClueGO, a widely used Cytoscape plugin, facilitates the visualization of functionally related genes displayed as a clustered network and chart. UniProt IDs of differential proteins were analyzed using the default parameters, which specify via a right-sided hypergeometric test an enrichment correction method using a Bonferroni step-down. The “Function” analysis mode, the gene cluster list for *Homo sapiens*, and a kappa score of 0.4 were used; the evidence codes were set to “All”, and the networking specificity was set to medium (GO levels 3 to 8).

## Determination of plasma concentrations by ELISA

Plasma aggrecan core protein (PGCA), alpha-2-HS-glycoprotein (FETUA), coagulation factor XIII A chain (F13A) and gelsolin (GELS) protein concentrations were determined by enzyme-linked immunosorbent assay (ELISA) analysis and compared 10 new samples of both control and B-ALL. Detailed procedures were performed according to manufacturer instructions to measure PGCA (Aviva Systems Biology Cat.OKEH00570), FETUA (MyBiosource Cat. MBS175929), F13A (Aviva Systems Biology Cat. OKEH02738) and GELS (Aviscera Bioscience, INC. Cat. SK00384-01) levels to generate measurement standards. Standard curves were created with protein concentration on the y-axis and average absorbance on the x-axis. The

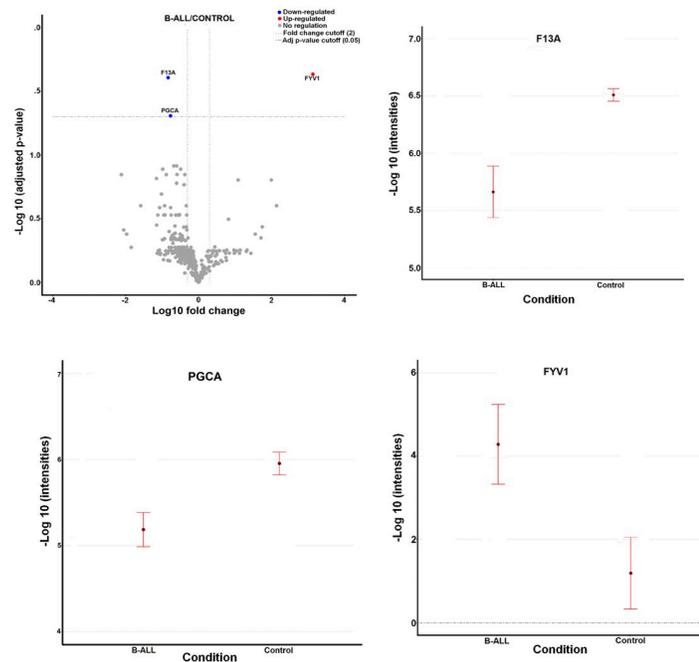
results were then calculated using the standard curves and multiplied by the dilution factor. Student's t-test was used to analyze differences in protein levels between the control and cancer groups. A P-value <0.05 was considered statistically significant.

## Results

### Twenty-five differently expressed proteins were identified in B-ALL patients

In this study, we carried out a label-free proteome analysis using samples collected from six B-ALL patients and six healthy controls. A total of 472 proteins were quantified in the label-free analysis. MSStats was used to detect variations in protein abundance between the studied groups from the peptide measurements database. The analysis was designed as a comparative experiment using the “Group Comparison” function, setting the conditions to “Control” and “B-ALL”. A unique identifier was assigned to each biological replicate and each run. The intensity value was defined as the total area of each feature without transformation (data shown in supporting information S3 and S4 Tables). The significant changes in protein abundance were found using a linear mixed model, and the raw P-values were adjusted by the Benjamin and Hochberg method [12]. Significantly differentially expressed proteins in B-ALL patients compared to controls (those with an adjusted P-value <0.05, a fold change >2 and an FDR of 5%) are shown in a volcano plot (x = Log<sub>10</sub>FC, y = -Log<sub>10</sub> adjusted P-value) (Fig 1A). The results showed that F13A and PGCA were down-regulated, and FYV1 was up-regulated (Table 2); the corresponding condition plot graphs are shown in Fig 1B, 1C and 1D.

ROTS classifies proteins according to evidence by means of examining the differential of expression in two-group comparisons [11,13]. For this analysis, the database “export protein



**Fig 1. MSStats statistics.** A. Volcano plot of the comparison B-ALL vs Control. The dashed line represents the false discovery rate (FDR) cutoff 5% and fold-change cutoff = 2.0. The up-regulated protein is shown in red dots and down-regulated in blue dots. Condition plots for FYV1 (B), PGCA (C) and F13A (D) proteins. Dots indicate the mean of log ratio and error bars have confidence intervals with 0.95 significant level for each condition.

<https://doi.org/10.1371/journal.pone.0221509.g001>

**Table 2. Differentially expressed proteins between B-ALL and control groups.**

Accession number	Gene name	Protein name	MSStats Log <sub>10</sub> Fold change LLA-B vs. Control	MSStats P BH	ROTS Log <sub>2</sub> Fold change Control vs. LLA-B	ROTS P BONF
Q9Y2I7	<i>FYV1</i>	1-phosphatidylinositol 3-phosphate 5-kinase	3.134	0.0232		
P16112	<i>PGCA</i>	Aggrecan core protein	-0.7718	0.0487	2.359	0.001
P07333	<i>CSF1R</i>	Macrophage colony-stimulating factor 1 receptor			2.021	0.001
P00488	<i>F13A</i>	Coagulation factor XIII A chain	-0.8375	0.0244	1.961	< 0.001
Q9HCU4	<i>CELR2</i>	Cadherin EGF LAG seven-pass G-type receptor			1.864	< 0.001
P54108	<i>CRIS3</i>	Cysteine-rich secretory protein 3			1.817	< 0.001
P23142	<i>FBLN1</i>	Fibulin-1			1.802	< 0.001
P04278	<i>SHBG</i>	Sex hormone-binding globulin			1.780	< 0.001
O00533	<i>NCHL1</i>	Neural cell adhesion molecule L1-like protein			1.770	< 0.001
P02679	<i>FIBG</i>	Fibrinogen gamma chain			1.605	< 0.001
Q9UJV3	<i>TRIM1</i>	Probable E3 ubiquitin-protein ligase MID2			1.604	< 0.001
P55290	<i>CAD13</i>	Cadherin-13			1.574	0.0175
P02675	<i>FIBB</i>	Fibrinogen beta chain			1.576	< 0.001
P06396	<i>GELS</i>	Gelsolin			1.405	< 0.001
P06276	<i>CHLE</i>	Cholinesterase			1.378	< 0.001
P05160	<i>F13B</i>	Coagulation factor XIII B chain			1.359	< 0.001
P02671	<i>FIBA</i>	Fibrinogen alpha chain			1.333	< 0.001
P51884	<i>LUM</i>	Lumican			1.256	< 0.001
P11597	<i>CETP</i>	Cholesteryl ester transfer protein			1.247	0.0095
P01717	<i>LV403</i>	Ig lambda chain V-IV region Hil			1.244	< 0.001
P04196	<i>HRG</i>	Histidine-rich glycoprotein			1.239	< 0.001
Q7Z7M0	<i>MEGF8</i>	Multiple epidermal growth factor-like domains protein 8			1.150	0.009
P02765	<i>FETUA</i>	Alpha-2-HS-glycoprotein			1.140	< 0.001
Q6UXB8	<i>PI16</i>	Peptidase inhibitor 16			1.135	0.029
O00187	<i>MASP2</i>	Mannan-binding lectin serine protease 2			1.112	0.0085

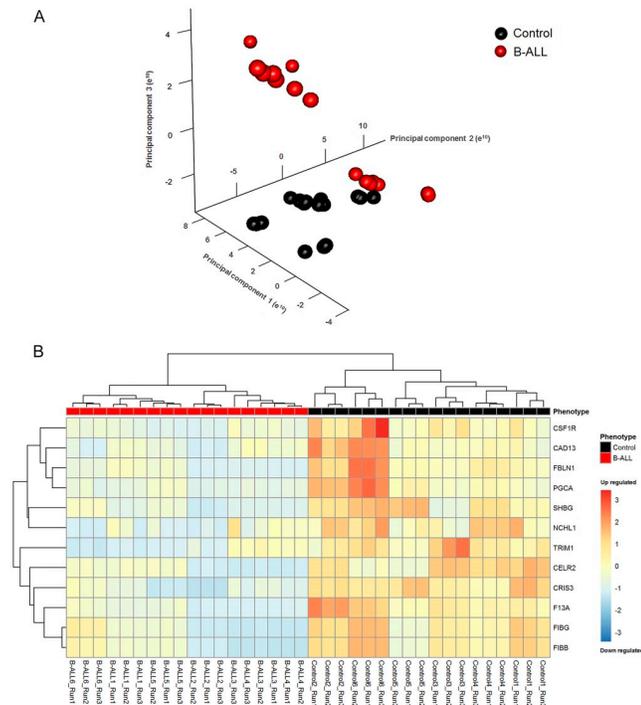
FDR: False Discovery Rate, P BONF: P adjusted Bonferroni, P BH: P adjusted Benjamini-Hochberg. Fold Change represents a log value

<https://doi.org/10.1371/journal.pone.0221509.t002>

measurements” was used and a 1000-repetition bootstrapping process was performed [11,14]. With the resulting enriched database, a multivariate analysis was performed using principal component analysis (PCA) to reveal larger variations (Fig 2A). A heatmap corresponding to differentially expressed proteins which had a fold change >2 and a Bonferroni-adjusted P-value <0.05 is shown in Supporting Information (data in S1 Fig). The 24 proteins identified provided a good separation between the groups, although the sample corresponding to control 5 did not match the overall control group. Fig 2B shows the heatmap corresponding to differentially expressed proteins with a Bonferroni-adjusted P-value <0.05 and a fold change >3. The number of proteins identified through these parameters was reduced from 24 to 12, but a complete separation between the groups was achieved.

### GO terms of differentially expressed proteins

Overall, 69 GO terms were significantly enriched; these terms were categorized into seven GO groups as represented in Fig 3. The main GO categories were protein activation cascade, platelet degranulation, blood coagulation-fibrin clot formation, extracellular matrix organization, cellular component morphogenesis, and cell morphogenesis involved in the differentiation and positive regulation of cell-substrate adhesion.



**Fig 2. ROTS statistics.** A. Unsupervised principal component score plot for controls and B-ALL patients. Red and black bubbles represent B-ALL and control plasma samples, respectively. The analysis resulted in a good separation between groups according to the principal component 1 (PC1) with 64,56% of explained variance, principal component 2 (PC2) 18,39% and principal component 3 (PC3) with 11,92%. B. Heat map of unsupervised clustering of the patients (columns) across the 12 proteins detected as differentially expressed by ROTS (rows), with a Bonferroni-adjusted P-value <0.05 and a fold change >3. Red and black bars represent B-ALL and control plasma samples, respectively.

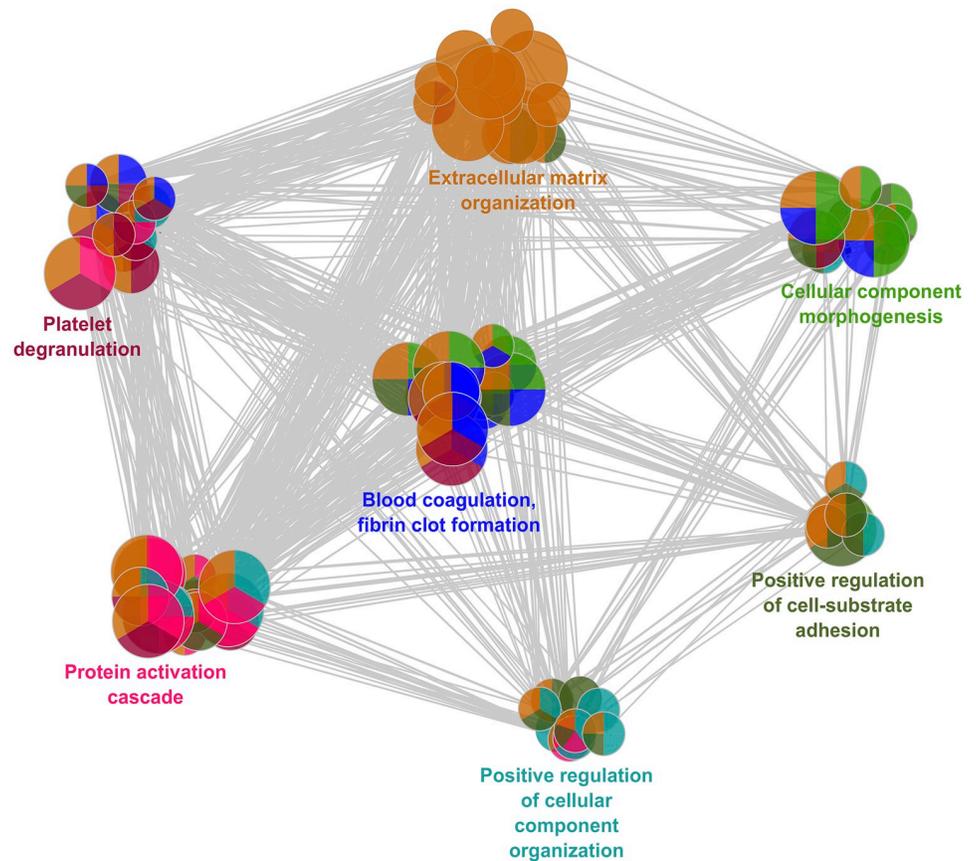
<https://doi.org/10.1371/journal.pone.0221509.g002>

### Plasma concentrations of PGCA, FETUA, F13A, and GELS proteins

The plasma levels of PGCA, FETUA, F13A, and GELS were measured in plasma samples from individuals with B-ALL and healthy controls by ELISA (n = 10). The results showed that B-ALL patients had lower plasma concentrations of these proteins compared to those of the control group (Fig 4): (B-ALL patients vs controls) PGCA 106.7 ± 4.844 ng/mL vs 127.6 ± 4.985 ng/mL, FETUA 184.7 ± 3.111 µg/mL vs 245.2 ± 11.14 µg/mL, F13A 10.68 ± 0.4860 µg/mL vs 12.48 ± 0.3921 µg/mL and gelsolin 148.4 ± 3.005 µg/mL vs 193.7 ± 5.673 µg/mL (data shown in supporting information S5 Table). These results were consistent with the label-free quantitation measurements. Although there was a difference in the magnitude of the change between groups probably associated with the difference in sensitivity and the discriminating capacity of the immunoassay, the trends were in agreement in the two techniques.

### Discussion

Blood plasma is a biological fluid that is highly attractive for use in screening tests and biomarker discovery due to its ease of collection and its reflection of physiological changes produced by or in response to disease [15]. In this study, a proteomic approach was used to identify differentially expressed proteins in the blood plasma of patients diagnosed with B-ALL versus healthy controls. The blood plasma was depleted for two of the most abundant

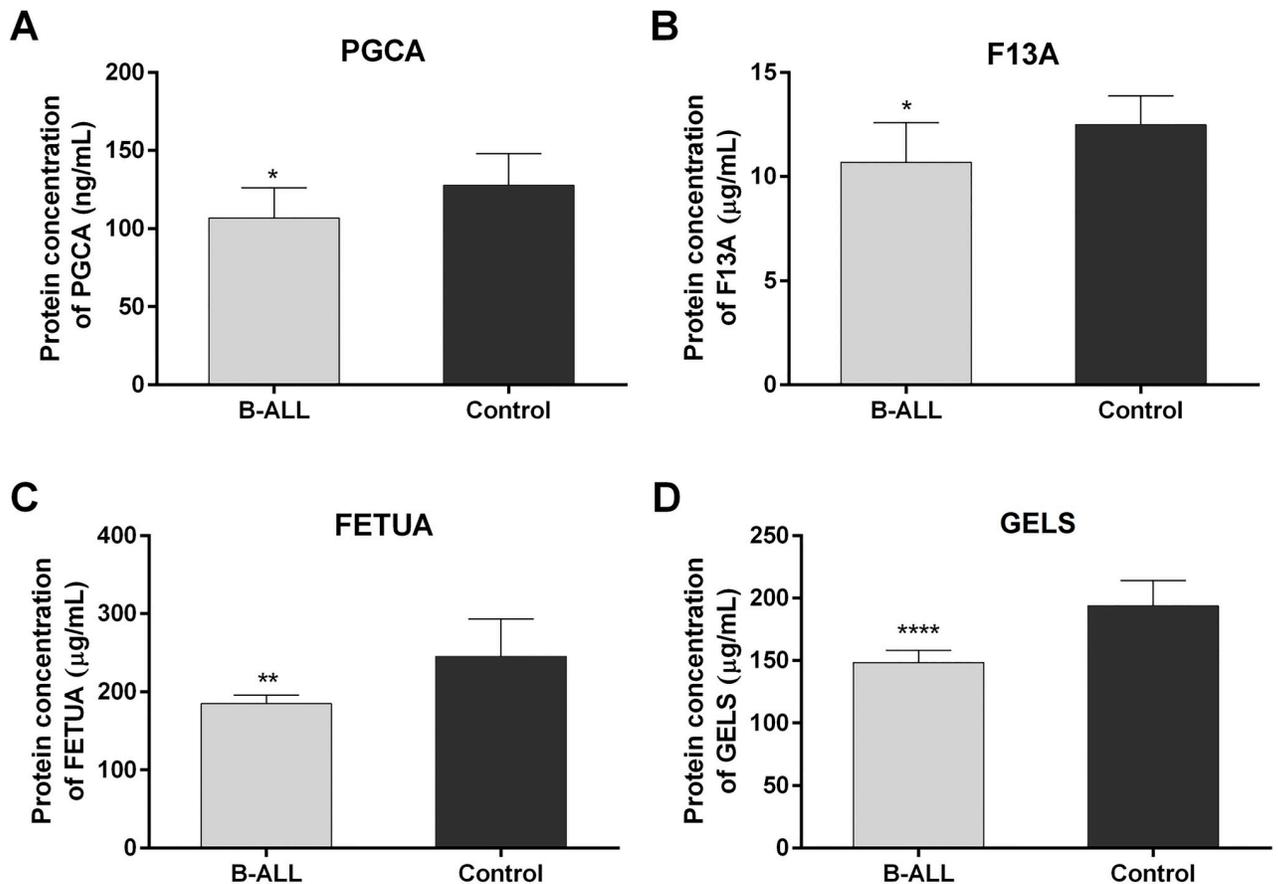


**Fig 3. Grouping of networks based on functionally enriched GO terms.** The functionally grouped network of enriched categories was generated for the differential proteins using ClueGO. GO terms are represented as nodes, the edges connecting the nodes are based on the kappa statistic that measures the overlap of shared genes between terms (less to 0.4). The node size indicates the number of proteins mapped to each term and the color represents the term enrichment significance. Only the most significant term in the group is labeled. Functionally related groups partially overlap. Visualization has been carried out using Cytoscape 3.4.0.

<https://doi.org/10.1371/journal.pone.0221509.g003>

proteins, albumin, and IgG. Next, the plasma was analyzed by LC-MS/MS and quantified by label-free measurement. After statistical analysis, 25 proteins were identified as being differentially expressed. A decreased expression level was observed for 24 of these proteins, while the protein lipidic kinase 1-phosphatidylinositol-3-phosphate 5-kinase (PIKfyve or FYV1) was expressed at higher levels in patients with B-ALL than in healthy controls. This type-III PIP kinase participates in endosomal processes, autophagy regulation and exosome liberation [16]. PIKfyve controls phosphatidylinositol (PtdIns) (3,5) P2 levels to trigger the action of the SAC3 phosphatase to regenerate PtdIns3P or of the MTMR3 phosphatase to generate PtdIns5P. PtdIns5P is an intermediate that regulates the PI3K/Akt pathway and has been postulated as a second messenger in cellular migration, which suggests a role in oncogenesis [17]. Oppelt *et al* showed that FYV1 and MTMR3 are expressed in most cancerous cells and that decreased levels of these proteins produce an alteration of cell migration and are involved in the invasive behavior of cancer cells [18].

The hierarchical grouping produced in our analysis showed a group separation among the decreased differentially expressed proteins. The 12 proteins that presented a Bonferroni-adjusted P-value <0.05 and a fold >3 were interestingly grouped, given that they allowed a



**Fig 4. Plasma levels of PGCA, FETUA, F13A, and GELS.** Plasma levels of PGCA (A), FETUA (B), F13A (C) and GELS (D) was detected by ELISA in B-ALL patients (n = 10) and control groups (n = 10). Statistical significance was determined by two-tailed non-paired Student's t-test. Data as mean ± SEM. The analysis was performed using GraphPad Prism 7.0 \*\*P < 0.01 is significant, \*\*\*\*P < 0.0001 is highly significant.

<https://doi.org/10.1371/journal.pone.0221509.g004>

complete separation of the patients and the controls. We postulated that quantification of these proteins may be applied to achieve identification of B-ALL patients.

The functional annotation analysis showed that several of the identified proteins are involved in the blood coagulation process and have been reported as down-regulated in leukemia cases. For example, the coagulation factor F13A is a plasma transglutaminase that increases the coagulum resistance to fibrinolysis by forming covalent bonds between adjacent fibrin monomers and interweaving plasmin with fibrin inhibitors. Our proteomic approach showed that the F13A concentration was decreased in the blood plasma of patients with B-ALL. This observation has also been documented in whole plasma by ELISA assay in a study by Shi and Wang, in which a decrease of F13A in patients with acute myeloid and lymphoblastic leukemia and in solid tumors was reported [19]. Normal lymphoid cells do not synthesize or contain any F13 subunits, while lymphoid cells in bone marrow samples derived from patients with acute leukemia F13A do. This indicates F13A as a possible intracytoplasmic marker that can be used in the identification of aberrant phenotypes, which could provide a valuable diagnostic tool for the identification of acute leukemia from myeloid and B-lymphoid lineages [20].

Fibrinogen, another important member of the coagulation cascade, undergoes altered expression in different types of cancer, and it has been reported that the expression of FGA is increased in serum from ALL patients [21]. However, we observed a decreased level of all three-polypeptide chains ( $\alpha$ ,  $\beta$  and  $\gamma$ ) probably due to a common thrombocytopenia stage in the leukemic patients.

A second identified group of proteins is involved in adhesion, extracellular matrix organization, and cell morphogenesis processes. Of these, fibulin 1, gelsolin, histidine-rich glycoprotein, cadherin EGF LAG seven-pass G-type receptor 2 and cadherin-13 are highlighted and discussed below.

Fibulin 1, an extracellular matrix protein, and gelsolin, an actin-binding protein and key regulator in the assembling/disassembling of actin filaments, were found to be down-regulated in B-ALL samples. Down-regulation of these proteins has also been shown in other types of cancers, suggesting that they may function as tumor suppressors [22,23]. In leukemia U937 cells, gelsolin overexpression induced a retarded growth, an improved monocytic morphology, increased NADPH activity and enhanced superficial expression of the CD11b  $\beta$ -integrin receptor in comparison with U937 parental cells [24]. Our study showed concordant results with these previous findings. Gelsolin was decreased in plasma from B-ALL patients, a result that was confirmed by the ELISA measurement of the protein in whole plasma.

Interestingly, cadherin EGF LAG seven-pass G-type receptor (CELSR2) and cadherin-13 have not previously been reported in blood plasma, but they appear to be related to the malignancy processes of hematopoietic stem cells (HSC). These cadherins belong to the subfamily of Flamingo (Fmi) or CELSR cadherins and are involved in adhesion and cell-cell recognition processes [25]. These two cadherins intervene in the non-canonical Wnt pathway through binding with Frizzled (Fz) to regulate planar cell polarity and increase free intracellular calcium. Non-canonical Wnt signaling is necessary to maintain HSC quiescence by suppressing the  $\text{Ca}^{2+}$ -nuclear factor signaling of activated T-cell (NFAT)-gamma-interferon ( $\text{IFN}\gamma$ ) and the antagonist signaling of canonical Wnt. Fz8 and Fmi binding results in a non-canonical receptor of Wnt signaling, which response to non-canonical Wnt signaling as Wnt5a. Under stress conditions, canonical Wnt signaling activates HSCs to induce auto-renovation and differentiation. Non-canonical signaling by Wnt5a is increased in the new HSC population for a short term through the maintenance of HSCs in the quiescent phase (G0) [26,27]. The decreased expression of these two cadherins in B-ALL patients suggests an alteration of the Wnt signaling pathway to produce a change in the quiescence state of HSCs. Truncated cadherin-13 (T-cadherin) is an atypical member of the cadherin superfamily. It lacks transmembrane and cytoplasmic domains and is bound to the plasma membrane through a glycosylphosphatidylinositol anchor. T-cadherin has been suggested as a tumor suppressor, since in several types of cancer, including B-cell lymphoma and chronic myeloid leukemia (CML), it has been found to be down-regulated by allelic deletion and/or by promoter hypermethylation [28]. It is interesting to find these two related-to-non-canonical-Wnt-signaling cadherins decreased in B-ALL patients, we could speculate an alteration of this signaling pathway and thus a change of quiescence state of HSC.

Another highlighted protein is Aggrecan core protein (PGCA), a proteoglycan that is the main component of the extracellular matrix of cartilaginous tissue [29]. The extracellular matrix has been a focus of study in the examination of tumor invasion and metastasis, since these processes imply a complex series of events, including extra-cellular matrix proteolysis. PGCA is thought to be a tumor suppressor. In studies of larynx cancer, PGCA has presented a reduced expression and has been correlated with a worse prognosis and an increased cancer metastasis rate. PGCA homozygotic deletion is also related to the development of classic Hodgkin's lymphoma induced by the *Epstein-Barr* virus [30]. PGCA plasma expression levels

measured by ELISA confirmed the decreased expression of the protein in plasma from patients with B-ALL.

Lumican is a member of the small leucine-rich proteoglycan family, and it has been identified to undergo overexpression or negative regulation in different types of cancer. For instance, a decrease of 2–3 times the normal expression has been reported for lumican in breast cancer tissue. Additionally, lumican expression was further decreased in advanced disease stages than in early disease stages. Plasma proteomic analysis has demonstrated that lumican levels were significantly higher in patients with lung cancer in comparison with normal subjects [31]. Lumican is also implied in the negative regulation of the proteolytic activity associated with endothelial cell membranes, in particular, that of matrix metalloproteinases MMP-14 and MMP-9, as well as that of MMP-1, MMP-2, and MMP-13 [32]. Lumican blocks the migration and invasion of tumor cells. In general, lumican seems to be a potent agent for the inhibition of tumor progression, which is interesting given that MMP-2 and MMP-9 expression has been reported as positively altered in ALL cases [33,34].

CSF1R, also known as the receptor for the macrophage colony-stimulating factor (M-CSFR), is a member of the class-III tyrosine kinase receptor family. Along with its ligand CSF1L, it has a critical role in cell survival regulation and in the myeloid cell differentiation involved in macrophage lineage as well as bone formation. In adult hematopoiesis, CSF1R is selective and highly expressed in myeloid and dendritic cells but is not detectable in lymphocytes, including progenitor and mature B-cells. It has been suggested that there exists a connection between B-lineage development and myeloid lineages during fetal development, which may explain the chromosomal translocations of the mixed lineage leukemia (MLL) gene observed in breastfed infants diagnosed with B-ALL [35]. The CSF1R expression results are particularly interesting due to its role in a translocation observed in high-risk patients with a Philadelphia-like chromosome (Ph-like)<sup>+</sup>. Translocation t(1; 5) (q21; q33) has been reported between *CSF1R* and *MEF2D* in one ALL case [36], and the fusion of *CSF1R* with *SSBP2*, *TBL1XR1* genes has also been reported [37].

Other differentially expressed protein related to diverse cell metabolic processes was the fetuin-A, also down-regulated in B-ALL patients. Fetuin-A is a phosphorylated glycoprotein with multipotent properties which plays an anti-inflammatory role by counteracting pro-inflammatory cytokine production. Like haptoglobin, fetuin-A is an acute-phase protein, for which the serum concentration diminishes in comparison with normal levels as a response to inflammation. It is important to highlight that acute-phase proteins increase blood flow to the tumor microenvironment, and they may be relevant in tumorigenesis processes. Acute-phase proteins can also reflect a systemic inflammatory response. Fetuin-A is synthesized and secreted predominantly by the liver in adults. It is also involved in calcium homeostasis and in the decrease of insulin signaling through inhibition of the phosphorylation of insulin tyrosine-kinase receptors. Fetuin-A has been proposed as a potential tumor marker in specific malign neoplasia, including in pancreatic and breast cancer, and a decrease in Fetuin-A levels in serum from patients with chronic lymphoblastic leukemia (CLL) compared to those of healthy controls has been reported [38]. In proteomic studies, fetuin-A has been shown to be decreased in several cancer types. Tian showed a fetuin-A decrease in plasma from patients with hypopharyngeal squamous cell carcinoma (HSCC) [39]. Dowling reported decreased fetuin-A levels in serum from breast cancer patients compared to those of healthy controls [40]. Kwak also showed a fetuin-A decrease in serum from acute myeloid leukemia (AML) patients in comparison with healthy controls [41]. Our results also indicated lower fetuin-A expression levels in plasma from patients with B-ALL in comparison with those of healthy controls via both the proteomic and ELISA approaches, which further indicates the potential of fetuin-A as a possible tumor marker.

## Conclusion

This study makes public a first approximation of the alterations in the plasma proteome of Colombian B-ALL patients. We report some potential biomarkers that could be used to differentiate unhealthy patients from healthy individuals; these biomarkers would aid treatment providers in one of the most complex and difficult clinical decisions in relation to certain age groups. This exploratory work paves the way for new studies using the identified differentially expressed proteins for further clinical development. To validate the specificity and the potential applicability of our proposed markers, it is necessary to perform a further study involving a population of greater size and diversity.

## Supporting information

**S1 Table. Quantification results using Progenesis QI: Normalized intensities by proteins.**  
(XLSX)

**S2 Table. Quantification results using Progenesis QI: Normalized intensities by peptides.**  
(XLSX)

**S3 Table. Data results of MSStats statistical package.**  
(XLSX)

**S4 Table. Data results of ROTS statistical package.**  
(XLSX)

**S5 Table. Data results of ELISA assay.**  
(XLSX)

**S1 Fig. Heat map of unsupervised clustering of the patients across the 12 proteins detected as differentially expressed by ROTS.** Bonferroni-adjusted P-value <0.05 and a fold change >2. Red and black bars represent B-ALL and control plasma samples, respectively.  
(TIF)

## Acknowledgments

We thank the ALL patients and volunteers who donated samples for this study. We thank Drs. Myriam Sanchez and Martha Serrano for helpful discussions in the early stages of this work. We thank Dr. Joaquin Abián and the members of the Proteomics Laboratory at CSIC/UAB for their technical support.

## Author Contributions

**Conceptualization:** Adriana Umaña-Perez.

**Formal analysis:** María Carolina Sanabria-Salas.

**Funding acquisition:** Adriana Umaña-Perez.

**Investigation:** Sandra Isabel Calderon-Rodríguez.

**Methodology:** Sandra Isabel Calderon-Rodríguez.

**Validation:** Sandra Isabel Calderon-Rodríguez.

**Visualization:** Sandra Isabel Calderon-Rodríguez.

**Writing – original draft:** Sandra Isabel Calderon-Rodríguez.

Writing – review & editing: Adriana Umaña-Perez.

## References

1. Arber DA, Orazi A, Hasserjian R, Borowitz MJ, Le Beau MM, Bloomfield CD, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016; 127: 2391–2406. <https://doi.org/10.1182/blood-2016-03-643544> PMID: 27069254
2. Bhojwani D, Yang JJ. Biology of Childhood Acute Lymphoblastic Leukemia. *Pediatr Clin N Am*. 2015; 62: 47–60. Available: <https://doi.org/10.1016/j.pcl.2014.09.004> PMID: 25435111
3. Kęsy J, Januszkiewicz-lewandowska D. Genes and childhood leukemia. *Postepy Hig Med Dosw (online)*. 2015; 69: 302–308.
4. Kebriaei P, Anastasi J, Larson RA. Acute lymphoblastic leukaemia: diagnosis and classification. *Best Practice & Research Clinical Haematology*. 2002; 15: 597–621. <https://doi.org/10.1053/beha.2002.0224>
5. Rodriguez-Galindo C, Friedrich P, Alcasabas P, Antillon F, Banavali S, Castillo L, et al. Toward the cure of all children with cancer through collaborative efforts: Pediatric oncology as a global challenge. *Journal of Clinical Oncology*. 2015. pp. 3065–3073. <https://doi.org/10.1200/JCO.2014.60.6376> PMID: 26304881
6. Howard SC, Metzger ML, Wilimas JA, Quintana Y, Pui CH, Robison LL, et al. Childhood Cancer Epidemiology in Low-Income Countries. *Cancer*. 2008; 112: 461–472. <https://doi.org/10.1002/cncr.23205> PMID: 18072274
7. Suarez A, Pina M, Nichols-Vinueza DX, Lopera J, Rengifo L, Mesa M, et al. A Strategy to Improve Treatment-Related Mortality and Abandonment of Therapy for Childhood ALL in a Developing Country Reveals the Impact of Treatment Delays. *Pediatr Blood Cancer*. 2015; 62: 1395–1402. <https://doi.org/10.1002/pbc.25510> PMID: 25808195
8. Rose-Inman H, Damon Kuehl. Acute leukemia. *Emerg Med Clin N Am*. 2014; 32: 579–596. Available: <http://dx.doi.org/10.1016/j.emc.2014.04.004>
9. Cunningham R, Ma D, Li L. Mass spectrometry-based proteomics and peptidomics for systems biology and biomarker discovery. *Frontiers in Biology*. 2012; 7: 313–335. <https://doi.org/10.1007/s11515-012-1218-y> PMID: 24504115
10. Team RDC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.; 2008.
11. Pursiheimo A, Vehmas AP, Afzal S, Suomi T, Chand T, Strauss L, et al. Optimization of Statistical Methods Impact on Quantitative Proteomics Data. *Journal of Proteome Research*. 2015; 14: 4118–4126. <https://doi.org/10.1021/acs.jproteome.5b00183> PMID: 26321463
12. Choi M, Chang C-Y, Clough T, Broudy D, Killeen T, MacLean B, et al. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*. 2014; 30: 2524–2526. <https://doi.org/10.1093/bioinformatics/btu305> PMID: 24794931
13. Elo LL, Hiisa J, Tuimala J, Kallio A, Korpelainen E, Aittokallio T. Optimized detection of differential expression in global profiling experiments: case studies in clinical transcriptomic and quantitative proteomic datasets. *Briefings in Bioinformatics*. 2009; 10: 547–555. <https://doi.org/10.1093/bib/bbp033> PMID: 19549804
14. Kerr MK, Churchill GA. Bootstrapping cluster analysis: Assessing the reliability of conclusions from microarray experiments. *Proceedings of the National Academy of Sciences*. 2001; 98: 8961–8965. Available: <https://doi.org/10.1073/pnas.161273698> PMID: 11470909
15. Surinova S, Schiess R, Hu R, Cerciello F. On the Development of Plasma Protein Biomarkers. *Journal of Proteome Research*. 2011; 10: 5–16. <https://doi.org/10.1021/pr1008515> PMID: 21142170
16. Hessvik NP, Øverbye A, Brech A, Torgersen ML, Jakobsen IS, Sandvig K, et al. PIKfyve inhibition increases exosome release and induces secretory autophagy. *Cellular and Molecular Life Sciences*. 2016; 73: 4717–4737. <https://doi.org/10.1007/s00018-016-2309-8> PMID: 27438886
17. Coronas S, Lagarrigue F, Ramel D, Chicanne G, Delsol G, Payrastra B, et al. Elevated levels of PtdIns5P in NPM-ALK transformed cells: Implication of PIKfyve. *Biochemical and Biophysical Research Communications*. 2008; 372: 351–355. <https://doi.org/10.1016/j.bbrc.2008.05.062> PMID: 18501703
18. Oppelt A, Haugsten EM, Zech T, Danielsen HE, Sveen A, Lobert VH, et al. PIKfyve, MTMR3 and their product PtdIns5 P regulate cancer cell migration and invasion through activation of Rac1. *Biochemical Journal*. 2014; 461: 383–390. <https://doi.org/10.1042/BJ20140132> PMID: 24840251
19. Shi D-Y, Wang S-J. Advances of Coagulation Factor XIII. *Chinese Medical Journal*. 2017; 130: 2190–223.
20. Kiss F, Simon A, Csáthy L, Hevessy Z, Katona E, Kiss C, et al. A coagulation factor becomes useful in the study of acute leukemias: studies with blood coagulation factor XIII. *Cytometry Part A: the journal of*

- the International Society for Analytical Cytology. 2008; 73: 194–201. <https://doi.org/10.1002/cyto.a.20485> PMID: 18000871
21. Bai J, He A, Huang C, Yang J, Zhang W, Wang J, et al. Serum peptidome based biomarkers searching for monitoring minimal residual disease in adult acute lymphocytic leukemia. *Proteome Science*. 2014; 12: 1–17.
  22. Cui Y, Liu J, Yin H-B, Liu Y-F, Liu J-H. Fibulin-1 functions as a prognostic factor in lung adenocarcinoma. *Japanese Journal of Clinical Oncology*. 2015; 45: 854–859. <https://doi.org/10.1093/jjco/hyv094> PMID: 26185140
  23. Xiao W. Fibulin-1 is epigenetically down-regulated and related with bladder cancer recurrence Wei. *BMC Cancer*. 2014; 14. Available: <https://doi.org/10.1186/1471-2407-14-677> PMID: 25234557
  24. Shirkoohi R, Fujita H, Darmanin S, Takimoto M. Gelsolin induces promonocytic leukemia differentiation accompanied by upregulation of p21CIP1. *Asian Pacific Journal of Cancer Prevention*. 2012; 13: 4827–4834. <https://doi.org/10.7314/apjcp.2012.13.9.4827> PMID: 23167427
  25. Gul IS, Hulpiau P, Saeys Y, van Roy F. Evolution and diversity of cadherins and catenins. *Experimental Cell Research*. 2017;
  26. Sugimura R, He XC, Venkatraman A, Arai F, Box A, Semerad C, et al. Noncanonical Wnt Signaling Maintains Hematopoietic Stem Cells in the Niche. *Cell*. 2012; 150: 351–365. <https://doi.org/10.1016/j.cell.2012.05.041> PMID: 22817897
  27. Frank J, Staal T. Wnt signalling meets epigenetics. *Stem Cell Investigation*. 2016; 3. Available: <https://doi.org/10.21037/sci.2016.08.01> PMID: 27668245
  28. Philippova M, Joshi MB, Kyriakakis E, Pfaff D, Erne P, Resink TJ. A guide and guard: The many faces of T-cadherin. *Cellular Signalling*. 2009; 21: 1035–1044. <https://doi.org/10.1016/j.cellsig.2009.01.035> PMID: 19399994
  29. Kiani C, Chen L, Wu YJ, Yee AJ, Yang BB. Structure and function of aggrecan. *Cell Research*. 2002; 12: 19–32. <https://doi.org/10.1038/sj.cr.7290106> PMID: 11942407
  30. Binder MJ, McCoombe S, Williams ED, McCulloch DR, Ward AC. The extracellular matrix in cancer progression: Role of hyaluronan proteoglycans and ADAMTS enzymes. *Cancer Letters*. 2017; 385: 55–64. <https://doi.org/10.1016/j.canlet.2016.11.001> PMID: 27838414
  31. Nikitovic D, Papoutsidakis A, Karamanos NK, Tzanakakis GN. Lumican affects tumor cell functions, tumor-ECM interactions, angiogenesis and inflammatory response. *Matrix Biology*. 2014; 35: 206–214. <https://doi.org/10.1016/j.matbio.2013.09.003> PMID: 24060754
  32. Pietraszke K, Chatron-Colliet A, Brezillon S, Perreau C, Jakubiak-Augustyn A, Krotkiewski H, et al. Lumican: A new inhibitor of matrix metalloproteinase-14 activity. *FEBS Letters*. 2014; 588: 4319–4324. <https://doi.org/10.1016/j.febslet.2014.09.040> PMID: 25304424
  33. Suminoe A, Matsuzaki A, Hattori H, Koga Y, Ishii E, Hara T. Expression of matrix metalloproteinase (MMP) and tissue inhibitor of MMP (TIMP) genes in blasts of infant acute lymphoblastic leukemia with organ involvement. *Leukemia Research*. 2007; 31: 1437–1440. <https://doi.org/10.1016/j.leukres.2007.01.015> PMID: 17350093
  34. Schneider P, Costa O, Legrand E, Bigot D, Lecleire S, Grassi V, et al. In vitro secretion of matrix metalloproteinase 9 is a prognostic marker in childhood acute lymphoblastic leukemia. *Leukemia Research*. 2010; 34: 24–31. <https://doi.org/10.1016/j.leukres.2009.07.039> PMID: 19748669
  35. Zriwil A, Böiers C, Wittmann L, Green JCA, Woll PS, Jacobsen SEW, et al. Macrophage colony-stimulating factor receptor marks and regulates a fetal myeloid-primed B-cell progenitor in mice. *Blood*. 2016; 128: 217–226. <https://doi.org/10.1182/blood-2016-01-693887> PMID: 27207794
  36. Liljebjörn H, Ågerstam H, Orsmark-Pietras C, Rissler M, Ehrencrona H, Nilsson L, et al. RNA-seq identifies clinically relevant fusion genes in leukemia including a novel MEF2D/CSF1R fusion responsive to imatinib. *Leukemia*. 2014; 28: 977–9. <https://doi.org/10.1038/leu.2013.324> PMID: 24186003
  37. Reshmi SC, Harvey RC, Roberts KG, Stonerock E, Smith A, Jenkins H, et al. Targetable kinase gene fusions in high risk B-ALL: a study from the Children's Oncology Group. *Blood*. 2017; 129: 3352–3361. <https://doi.org/10.1182/blood-2016-12-758979> PMID: 28408464
  38. Dalamaga M, Polyzos SA, Karmaniolas K, Chamberland J, Lekka A, Triantafylli M, et al. Fetuin-A levels and free leptin index are reduced in patients with chronic lymphocytic leukemia: a hospital-based case-control study. *Leukemia & Lymphoma*. 2016; 57: 577–584. <https://doi.org/10.3109/10428194.2015.1075523> PMID: 26317877
  39. Tian W-D, Li J-Z, Hu S-W, Peng X-W, Li G, Liu X, et al. Proteomic identification of alpha-2-HS-glycoprotein as a plasma biomarker of hypopharyngeal squamous cell carcinoma. *Int J Clin Exp Pathol*. 2015; 8: 9021–9031. PMID: 26464644
  40. Dowling P, Clarke C, Hennessy K, Torralbo-Lopez B, Ballot J, Crown J, et al. Analysis of acute-phase proteins, AHSG, C3, CLI, HP and SAA, reveals distinctive expression patterns associated with breast,

colorectal and lung cancer. *International Journal of Cancer*. 2012; 131: 911–923. <https://doi.org/10.1002/ijc.26462> PMID: 21953030

41. Kwak J-Y, Ma T-Z, Yoo M-J, Hee Choi B, Kim H-G, Kim S-R, et al. The comparative analysis of serum proteomes for the discovery of biomarkers for acute myeloid leukemia. *Experimental Hematology*. 2004; 32: 836–842. <https://doi.org/10.1016/j.exphem.2004.06.006> PMID: 15345285