ACTA ABP BIOCHIMICA POLONICA

Vol. 67, No 4/2020
475–483
https://doi.org/10.18388/abp.2020_5094

Regular paper

Interaction network of proteins associated with unfavorable prognosis in acute myeloid leukemia

Juan Jose Rendon-Rodriguez, Luisa Fernanda Restrepo-Rodriguez and Sarah Röthlisberger✉

Biomedical Innovation and Research Group, Faculty of Applied and Exact Sciences, Instituto Tecnologico Metropolitano ITM, 050034 Medellin, Colombia

Acute myeloid leukemia (AML) is a malignant disorder of hematopoietic stem and progenitor cells, characterized by accumulation of immature blasts in the bone marrow and peripheral blood of affected patients. Standard induction therapy leads to complete remission in approximately 50% to 75% of patients. In spite of favorable primary response rates, only 20% to 30% of patients enjoy long-term disease free survival. Identifying proteins involved in prognosis is important for proposing biomarkers that can aid in the clinical management of the disease. The aim of this study was to construct a protein-protein interaction (PPI) network based on serum proteins associated with unfavorable prognosis of AML, and analyze the biological pathways underlying molecular complexes in the network. We identified 16 candidate serum proteins associated with unfavorable prognosis (in terms of poor response to treatment, poor overall survival, short complete remission, and relapse) in AML *via* a search in the literature: IL2RA, FTL, HSP90AA1, D2HGDH, PLAU, CCL18A1, FGF19, SPP1, FGA, PF4, NME1, TNF, ANGPT2, B2M, CD274, LGALS3. The PPI network was constructed with Cytoscape using association networks from String and BioGRID, and Gene Ontology enrichment analysis using the ClueGo pluggin was performed. The central protein in the network was found to be PTPN11 which is involved in modulating the RAS-ERK, PI3K-AKT and JAK-STAT pathways, as well as in hematopoiesis, and in the regulation of apoptotic genes. Therefore, a dysregulation of this protein and/or of the proteins connected to it in the network leads to the defective activation of these signaling pathways and to a reduction in apoptosis. Together, this could cause an increase in the frequency of leukemic cells and a resistance to apoptosis in response to treatment.

Key words: acute myeloid leukemia, interaction network, prognosis, protein

Received: 06 December, 2019; revised: 01 April, 2020; accepted: 13 October, 2020; available on-line: 07 December, 2020

✉e-mail: saritaruth@gmail.com
Acknowledgements of Financial Support: The authors gratefully acknowledge the support of the Instituto Tecnologico Metropolitano (ITM) of Medellin, Colombia, and Hospital Manuel Uribe Angel, through Grant No. P17215 awarded to SR.
Abbreviations: AML, acute myeloid leukemia; BC, betweenness centrality; CC, Ccoseness centrality; GO, Gene Ontology; k, connectivity degree; SP, shortest path; PPI, protein-protein interaction

INTRODUCTION

Acute myeloid leukemia (AML) is a malignant disorder of hematopoietic stem and progenitor cells, characterized by accumulation of immature blasts in the bone marrow and peripheral blood of affected patients. Response to chemotherapy treatment in patients with AML is wide-ranging, and there are no adequate biomarkers to predict their clinical outcome (Bienz *et al*., 2005; Lazarevic *et al*., 2015; Slovak *et al*., 2014). Standard induction therapy, based on cytarabine and anthracycline, leads to complete remission in approximately 50% to 75% of patients, depending on prognostic factors, such as age or the presence of certain gene or chromosomal changes (Mrożek *et al*., 2012). In spite of favorable primary response rates, only approximately 20% to 30% of the patients enjoy long-term disease survival. This heterogeneity is related to acquired mutations, and deregulation in the expression of genes and non-coding RNA (miRNA) (Liao *et al*., 2017; Walker & Marcucci, 2012). It is clear that genetic studies are very valuable, but when isolated from a context in which thousands of proteins mediate cellular function, this information cannot be interpreted properly and without bias. Protein-protein interaction (PPI) networks seek to characterize this flow of information within the cell and the organism in order to understand the functional relevance of expressed proteins (Končarević *et al*., 2014). Analysis of PPI networks can help understand mechanisms involved in diseased states, and orient research strategies into biomarkers or therapeutic targets. Identifying proteins involved in response to treatment is important for proposing biomarkers that can aid in the clinical management of AML. The aim of this study was to construct a PPI network with key proteins identified in the literature as associated with chemotherapy resistance in AML, and analyze the biological pathways underlying molecular complexes in the network. This approach recognizes that many pathways are involved in the pathogenesis of AML, and thus a multi-marker strategy will almost certainly be necessary, as a single biomarker is unlikely to be sensitive and specific enough.

MATERIALS AND METHODS

Seed proteins. We systematically searched PubMed for proteomic studies that analyzed prognosis of AML patients, with the criteria that blood or serum was used as a biological sample (Acute Myeloid Leukemia AND prognosis AND serum OR blood AND protein OR proteomics). Based on these criteria, and after manual curation, we identified 16 candidate proteins.

Construction of a protein-protein interaction network. The PPI network was constructed using the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) web source (Szklarczyk *et al*., 2017) and the Biological General Repository for Interaction Datasets (BioGRID) database (Chatr-Aryamontri *et al*., 2017). The parameters of confidence for STRING were restricted in

order to reduce the amount of data while maintaining the most reliable interactions. The active prediction methods taken into account for STRING predictions were: experiments, co-expression, neighborhood and databases; a confidence score > 0.4 – medium confidence; 1st and 2nd shell no more than 20 interactors and no more than 5 interactors, respectively. We did not consider the direction of each protein interaction, and the duplicate edges and self-interactions were removed from the results.

Topological analysis of the protein interaction network. The network analysis includes three fundamental parameters that allow for nodes in a network to be evaluated: Connectivity degree (k), betweenness centrality (BC) and closeness centrality (CC). The most basic characteristic of a node in a network is its degree (k), which represents the number of interactions (links) the node has to other nodes (Barabási & Oltvai, 2004). Nodes with a higher k value are called hubs and therefore are the principal agents in the interaction network, affecting the network´s function and stability (Patil *et al*., 2010). The BC value is an indicator of a node's centrality in the network. It is the fraction of the number of non-redundant shortest paths (SP) that pass through each node, which measures how often the node is located on the shortest path between other nodes. The SP refers to the path with the smallest number of links between the selected nodes in a network (Raman, 2010). Nodes with higher BC are called bottlenecks and indicate that a large number of SP in the network passes through them. The CC of a node is defined as the inverse of the average length of the SP to/from all the other nodes in the graph. The node with the highest CC value is usually the topological center of the network (Ran *et al*., 2013). In the present study, a network analyzer Cytoscape 3.6.1 (Shannon *et al*., 2003; Assenov *et al*., 2008) was used to compute the properties of the whole network. In order to classify the hub and bottleneck proteins, we divided all of the proteins into four categories, as proposed in the literature (Yu *et al*., 2007): (1) nonhub–nonbottlenecks (small k and low BC); (2) hub–nonbottlenecks (large k but low BC); (3) nonhub–bottlenecks (small k but high BC); and (4) hub–bottlenecks (large k and high BC).

Construction of the backbone network of the AML PPI network. In order to construct a backbone network, we selected proteins from the giant network within the top 10% BC values, excluding those not within the main network. Based on graph theory, the protein bottlenecks are nodes with SP, therefore these control communication among other nodes in the giant network (Yu *et al*., 2007). This information can give us an approximation about the shortest pathway from the giant network that could be activated in chemotherapy resistance in AML patients.

Construction of a subnetwork consisting of all the shortest paths between the seed proteins. In order to construct a subnetwork in which the 16 seed proteins are connected directly or indirectly with the minimum number of connections, we found the SP between seed proteins using the PesCa 3.0.8 plug-in for Cytoscape (Scardoni *et al*., 2015). The subnetwork was constructed using the SP that interconnects seed proteins with a size of less than 6 nodes, thus helping to determine the principal pathways and biological processes between seed proteins related to chemotherapy resistance in AML patients.

Gene ontology and pathway analysis. We performed Gene Ontology (GO) enrichment analysis and

Table 1. List of the 16 seed proteins used to construct the PPI network

Uniprot	Symbol	Protein Name	References
P01589	IL2RA	Interleukin-2 receptor subunit alpha (aka CD25)	(Yabushita *et al*., 2018; Nakase *et al*., 2018; Allan *et al*., 2018)
P02792	FTL	Serum ferritin light chain	(Tachibana *et al*., 2018; Bertoli *et al*., 2019)
P07900	HSP90AA1	Heat shock protein HSP 90-alpha	(Fredly *et al*., 2012; Ma *et al*., 2017)
Q8N465	D2HGDH	D-2-hydroxyglutarate dehydrogenase	(Wang *et al*., 2013; Janin *et al*., 2014; Balss *et al*., 2016)
P00749	PLAU	Urokinase-type plasminogen activator	(Mustjoki *et al*., 2000; Li *et al*., 2014)
P39060	COL18A1	Collagen alpha-1(XVIII) chain - (cleaved to endostatin)	(Aref *et al*., 2008)
O95750	FGF19	Fibroblast growth factor 19	(Su *et al*., 2015)
P10451	SPP1	Osteopontin	(Liersch *et al*., 2012; Chen *et al*., 2017)
P02671	FGA	Isoform 1 of fibrinogen alpha chain precursor	(Krug *et al*., 2010)
P02776	PF4	Platelet factor 4	(He *et al*., 2010)
P15531	NME1 (NM23-H1)	Nucleoside diphosphate kinase A	(Okabe-Kado *et al*., 2009a, 2009b; Lilly *et al*., 2011)
P01375	TNF	Tumor necrosis factor	(Hu *et al*., 2019)
O15123	ANGPT2	Angiopoietin-2	(Kümpers *et al*., 2008)
P61769	B2M	Beta 2 microglobulin	(Tsimberidou *et al*., 2008; Melillo *et al*., 1992)
Q9NZQ7	CD274 (PD-L1)	Programmed cell death 1 ligand 1	(Ma *et al*., 2017)
P17931	LGALS3	Galectin-3	(Gao *et al*., 2017a, 2017b)

Table 2. Network topology measures.

Description	Giant network	Backbone network	SP network
Number of nodes (N)	340	28	81
Number of edges (E)	2223	73	409
Average degree (k)	13.076	24.678	20.160
Average BC (BC)	0.0112	0.0935	0.0419
Average CC (CC)	0.2738	0.2672	0.2547

pathway annotation of the networks with ClueGO (Plug-in for Cytoscape) (Bindea *et al.*, 2009) and pathway annotation with Reactome (Croft *et al.*, 2011).

RESULTS AND DISCUSSION

Seed proteins

The literature search yielded 16 candidate serum proteins associated with unfavorable prognosis (in terms of poor response to treatment, poor overall survival, short complete remission, and relapse): interleukin-2 receptor subunit alpha (IL2RA), serum ferritin light chain (FTL), heat shock protein HSP 90-alpha (HSP90AA1), D-2-hydroxyglutarate dehydrogenase (D2HGDH), urokinase-type plasminogen activator (PLAU), collagen alpha-1(XVIII) chain (COL18A1), fibroblast growth factor 19 (FGF19), osteopontin (SPP1), isoform 1 of fibrinogen alpha chain precursor (FGA), platelet factor 4 (PF4), nucleoside diphosphate kinase A (NME1), tumor necrosis factor (TNF), angiopoietin-2 (ANGPT2), beta 2 microglobulin (B2M), programmed cell death 1 ligand 1 (CD274), and galectin-3 (LGALS3). These seed proteins are listed in Table 1 and were used to construct the giant protein–protein interaction (PPI) network.

Giant network

The PPI network was constructed with 16 seed proteins associated with unfavorable prognosis in Acute Myeloid Leukemia, and was found to have 340 nodes connected by 2223 edges (Fig. 1 and Table 2). Each of the nodes represents a protein, while the edges between nodes represent interactions between proteins. As can be seen in Fig. 1, there is one main network and two smaller ones, with D2HGDH and FTL as seed proteins that are not connected to the main network.

The proteins in the network were classified into four categories according to their k and BC values, as described in the Methods section. This analysis revealed that 184 nodes were nonhub-nonbottleneck (low k and low BC), 122 nodes were hub-nonbottlenecks (high k and low BC), 13 nodes were nonhub-bottlenecks (low k and high BC), and 55 nodes were hub-bottlenecks (high k and high BC), the latter being of the most interest as they are the most central and well-connected nodes in the network.

In order to identify the most central node in the network, we compared these k and BC values and found two proteins of interest with the highest k and BC: PTPN11 (BC 0.186; k 37) and UBC (BC 0.130; k 44). As PTPN11 (also called Shp2) has the highest BC it was selected as the central node in the network.

Backbone network

We retrieved PTPN11, HSP90AA1, and the other 26 proteins within the top 10% largest degree (k) or highest BC and considered them as the hubs or bottlenecks and constituted the backbone of the giant network (Fig. 2). Of the 28 nodes comprising the backbone network, 6 are original seed proteins (PLAU, FGA, COL18A1, HSP90AA1, LGALS3, PF4). PTPN11 had the highest BC of the network (Table 3) meaning that it is the main node controlling the flow of information through

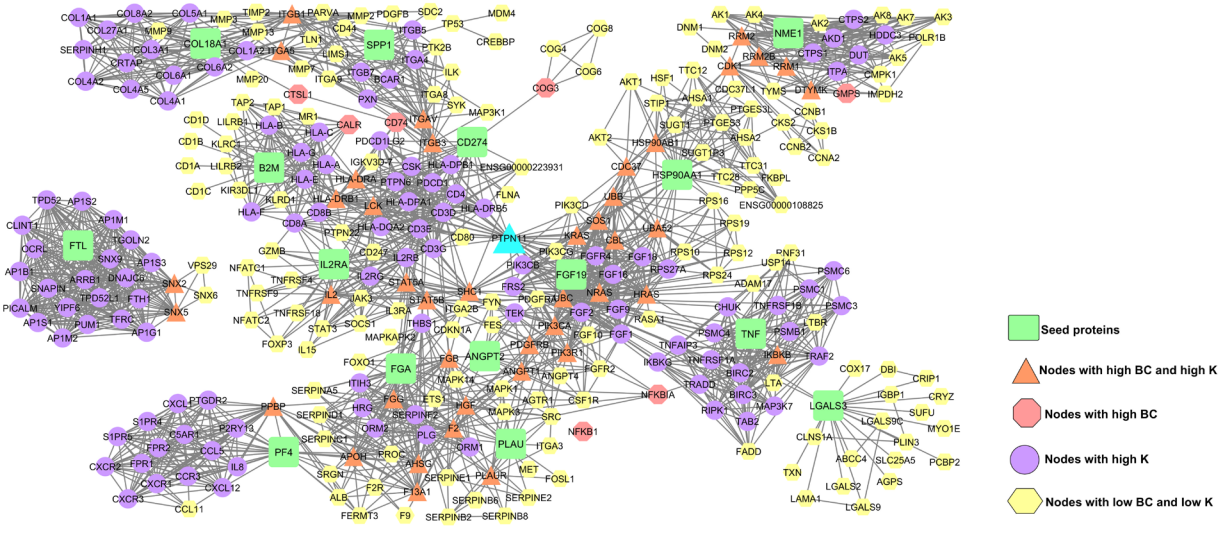

Figure 1. Protein-protein interaction network of proteins associated with unfavorable prognosis in Acute Myeloid Leukemia. PTPN11 is the central node in the network, represented by a cyan triangle. Betweenness centrality (BC), connectivity degree (k).

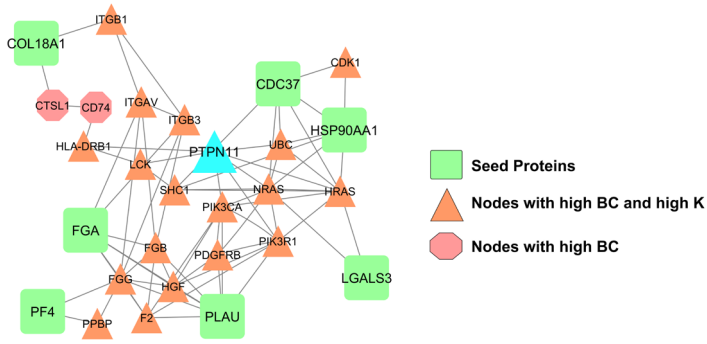

Figure 2. Backbone network.
The central node in the network is represented by a cyan triangle.

Table 3. Proteins in the backbone network.

N°	Protein	BC	k	CC	N°	Protein	BC	k	CC
1	PTPN11	0.187	37	0.321	18	CDC37	0.079	16	0.283
2	CDK1	0.164	17	0.234	19	CTSL1	0.077	2	0.211
3	**HSP90AA1**	0.161	32	0.287	20	**PLAU**	0.065	20	0.289
4	PIK3CA	0.137	33	0.319	21	LCK	0.062	35	0.287
5	F2	0.133	27	0.292	22	PDGFRB	0.052	17	0.307
6	UBC	0.131	44	0.293	23	**PF4**	0.051	24	0.214
7	NRAS	0.129	33	0.326	24	HLA-DRB1	0.050	30	0.278
8	PIK3R1	0.122	31	0.319	25	FGB	0.050	26	0.258
9	HGF	0.113	23	0.290	26	FGA	0.050	26	0.258
10	FGG	0.112	31	0.263	27	PPBP	0.046	23	0.214
11	**LGALS3**	0.112	20	0.253	28	ITGB1	0.044	26	0.199
12	ITGAV	0.087	22	0.223					
13	**COL18A1**	0.083	22	0.197					
14	ITGB3	0.083	22	0.223					
15	CD74	0.081	7	0.234					
16	HRAS	0.080	26	0.321					
17	SHC1	0.079	19	0.289					

Seed proteins are in bold

the network, followed by CDK1 and HSP90AA1, while UBC had the highest k.

Gene Ontology (Go) analysis was performed on the Backbone network to identify which GO terms (biological process and molecular function) were over or under-represented in the network (Table 4).

Shortest path network

The subnetwork of the shortest paths between the seed proteins was made up of 81 nodes and 409 edges. In Fig. 3, it can be observed that the 16 seed proteins are related to each other through intermediate nodes and there is a shorter pathway through which these proteins are related, which suggests that there are common signaling pathways between these proteins that could explain the biological context associated with an unfavorable prognosis in patients with AML. Just like in the Backbone network, in the SP network PTPN11 had the highest BC (0.186) while UBC had the highest k (44), and both values are well above the average. Topological analysis of this network is summarized in Table 2.

The GO analysis was performed on the SP network to identify which GO terms (biological process and molecular function) were over or under-represented in the network (Table 5). In terms of signaling pathways, the main pathways represented were signaling by EGFR in cancer, MET activated PI3K/AKT signaling, adaptive immune system, signaling by receptor tyrosine kinases, cytokine signaling in Immune system, diseases of signal transduction, hemostasis, PI3K-Akt signaling pathway, MAPK family signaling cascades, signaling by Interleukins, pathways in cancer, proteoglycans in cancer, cell surface interactions at the vascular wall, platelet activation, signaling and aggregation, acute myeloid leukemia, and chronic myeloid leukemia.

The backbone and SP networks were analyzed with Reactome and KEGG Pathway databases, and the

Table 4. Gene ontology analysis of the backbone network.

Biological Process	Associated Proteins
Platelet activation	F2, FGA, FGB, FGG, ITGB3, LCK, PF4, PIK3CA, PIK3R1, PTPN11
Positive regulation of protein kinase B signaling	CD74, HGF, HSP90AA1, ITGB1, LCK, PDGFRB, PIK3CA, PIK3R1, PTPN11
Heterotypic cell-cell adhesion	CD74, FGA, FGB, FGG, ITGAV, ITGB1, ITGB3, LCK
Negative regulation of extrinsic apoptotic signaling pathway	FGA, FGB, FGG, HGF, ITGAV, LGALS3, PF4
Platelet degranulation	FGA, FGB, FGG, HGF, ITGB3, PF4, PPBP
Extrinsic apoptotic signaling pathway via death domain receptors	FGA, FGB, FGG, HGF, LGALS3, PIK3R1
Fibrinolysis	F2, FGA, FGB, FGG, PLAU
Plasminogen activation	FGA, FGB, FGG, PLAU
ERBB2 signaling pathway	HSP90AA1, PIK3CA, PIK3R1, SHC1
Regulation of blood coagulation	F2, FGA, FGB, FGG, PLAU
Regulation of hemostasis	F2, FGA, FGB, FGG, PLAU
Molecular Function	**Associated Proteins**
Fibronectin binding	CTSL, ITGAV, ITGB1, ITGB3
Phosphotyrosine residue binding	LCK, PIK3R1, PTPN11, SHC1
Protein phosphorylated amino acid binding	LCK, PIK3R1, PTPN11, SHC1
C-X3-C chemokine binding	ITGAV, ITGB1, ITGB3
CD4 receptor binding	CD74, HLA-DRB1, LCK
Insulin receptor substrate binding	PIK3CA, PIK3R1, PTPN11
MHC protein complex binding	CD74, HLA-DRB1, HSP90AA1
Phosphatidylinositol 3-kinase binding	LCK, PDGFRB, PIK3R1
Co-receptor activity	ITGAV, ITGB1, ITGB3

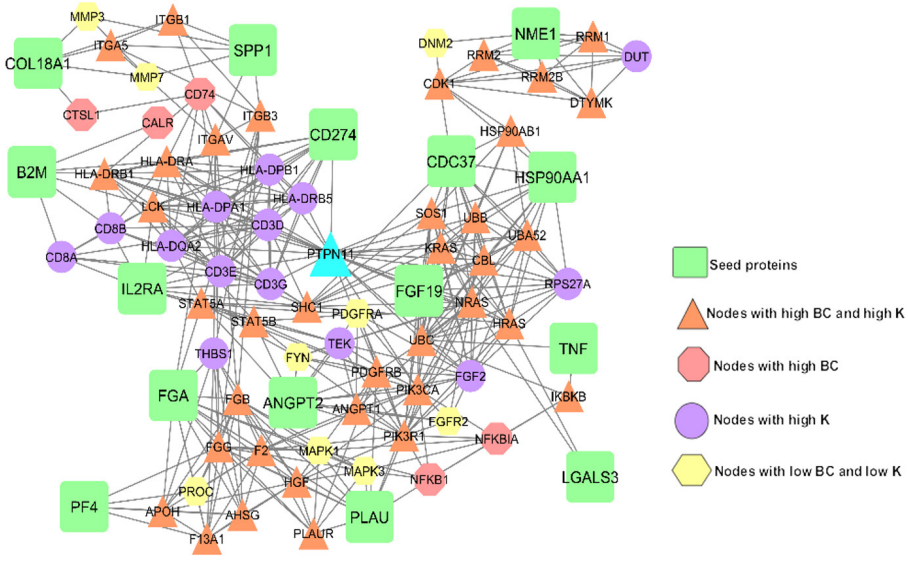

Figure 3. Shortest path network.
The central node in the network is represented by a cyan triangle.

main pathways represented were signaling by EGFRvIII in cancer, diseases of signal transduction, cell surface interactions at the vascular wall, platelet activation, signaling and aggregation, MAPK family signaling cascades, proteoglycans in cancer, MET activates PI3K/AKT signaling, acute myeloid leukemia, and chronic myeloid leukemia (Table 6).

Importance of PTPN11

PTPN11 encodes the Shp2 non-receptor protein-tyrosine that is involved in cytokine receptor and receptor tyrosine kinase signaling (Rehman *et al.*, 2018). This protein is required for the complete activation of the RAS-ERK pathway in response to growth factors and cytokines, in addition to modulating the PI3K-AKT and

Table 5. Gene ontology analysis of the shortest path network.

Biological Process	Associated Proteins
Tie signaling pathway	ANGPT1, ANGPT2, TEK
Peptide antigen assembly with MHC protein complex	CALR, HLA-DRA, HLA-DRB1
MHC protein complex assembly	CALR, HLA-DRA, HLA-DRB1
Interleukin-2-mediated signaling pathway	IL2RA, SHC1, STAT5A, STAT5B
Regulation of tau-protein kinase activity	HGF, HSP90AA1, HSP90AB1, SOS1
Positive regulation of heterotypic cell-cell adhesion	CD74, FGA, FGB, FGG, LCK, TNF
Response to interleukin-2	IL2RA, SHC1, STAT5A, STAT5B
Fibrinolysis	APOH, F2, FGA, FGB, FGG, PLAU, PLAUR, THBS1
Plasminogen activation	APOH, FGA, FGB, FGG, PLAU, THBS1
ERBB2 signaling pathway	EGFR, ERBB2, HSP90AA1, PIK3CA, PIK3R1, SHC1, SOS1
Regulation of heterotypic cell-cell adhesion	CD74, FGA, FGB, FGG, LCK, TNF
Blood coagulation, fibrin clot formation	APOH, F13A1, F2, FGA, FGB, FGG
Molecular Function	**Associated Proteins**
Oxidoreductase activity, acting on CH or CH2 groups, disulfide as acceptor	RRM1, RRM2, RRM2B
Ribonucleoside-diphosphate reductase activity	RRM1, RRM2, RRM2B
C-X3-C chemokine binding	ITGAV, ITGB1, ITGB3
CD4 receptor binding	CD74, FYN, HLA-DRB1, LCK
T cell receptor binding	CD3E, CD3G, FYN, LCK
Insulin receptor substrate binding	PIK3CA, PIK3R1, PTPN11
Platelet-derived growth factor receptor binding	ITGA5, ITGB3, PDGFRA, PDGFRB
Nitric-oxide synthase regulator activity	EGFR, HSP90AA1, HSP90AB1
MHC protein complex binding	CD74, HLA-DRA, HLA-DRB1, HSP90AA1, HSP90AB1
Phosphatidylinositol bisphosphate kinase activity	EGFR, ERBB2, FGF19, FGF2, FGFR2, FYN, HGF, LCK, PDGFRA, PDGFRB, PIK-3CA, PIK3R1, PTPN11, SOS1, STAT5A
Fibronectin binding	CTSL, ITGAV, ITGB1, ITGB3, THBS1

JAK-STAT pathways (Grossmann *et al.*, 2010). Mutations in PTPN11 occur in approximately 6.6% of patients with AML (Chen *et al.*, 2015) and lead to alterations in signaling pathways associated with the cell differentiation and growth. As this protein plays critical roles in hematopoiesis and leukemogenesis, myeloid and erythroid differentiation is affected in embryonic stem cells that express mutated Ptpn11 (Qu *et al.*, 1997). It has been reported that PTPN11 mutations cause an increase in the frequency of leukemic cells in both, humans and murine models (Chen *et al.*, 2015; Deng *et al.*, 2018). Shp2 also regulates apoptotic genes, and it has been reported that it increases the expression of Bcl2 and Mcl1. Patients who have mutations in Ptpn11, therefore, prove to be resistant to anti-Mcl1 drugs (Chen *et al.*, 2015).

There are several *in vitro* and *in vivo* studies in the literature that have shown that Ptpn11 is a potential target for cancer treatment, specifically when there is drug resistance. The potential for cancer treatment is observed in a study with transgenic mice containing a doxycycline (Dox)-inducible PTP-defective Shp2 mutant; when the Shp2 activity is inhibited in these mice, this results in suppressed EGFR signaling and fewer/smaller hyper proliferative lesions (Schneeberger *et al.*, 2015). In terms of drug resistance, an interesting study, published in 2015 (Prahallad *et al.*, 2015), revealed that when Ptpn11 was knocked-down, BRAF mutant colon cancer cells that were previously resistant to treatment with selective BRAF inhibitors became sensitive to these drugs.

There are currently three Ptpn11 (Shp2) inhibitors that have been developed for patients with advanced solid tumors that have failed, are intolerant to (drug resistance), or are considered ineligible for standard treatments. These function in a similar manner, by binding to and inhibiting Shp2 signaling, which in turn inhibits the Ras-MAPK pathway that is often hyperactivated in cancer cells. These are: JAB-3068 (Jacobio Pharmaceuticals Co.), RMC-4630 (Revolution Medicines, Inc. & Sanofi), and TNO155 (Novartis). For each of these there are two registered clinical trials: JAB-3068 (ClinicalTrials.gov identifier: NCT03565003 and NCT03518554), RMC-4630 (ClinicalTrials.gov identifier: NCT03634982 and NCT03989115), TNO155 (ClinicalTrials.gov identifier: NCT03114319 and NCT04000529). All of these are phase 1/2a clinical trials that are currently recruiting participants, with the aim of determining the maximum tolerated dose, as well as characterizing the safety, tolerability, and pharmacokinetics profile of these drugs.

Presently, there are no Ptpn11 (Shp2) inhibitors specifically aimed towards AML. However, there is a trial (ClinicalTrials.gov identifier: NCT03311815) sponsored by the PETHEMA Foundation, in which bone mar-

Table 6. Main pathways represented in the backbone and shortest pathway network

Only in Backbone	Only in SP	In both
	Adaptive Immune System	Signaling by EGFRvIII in Cancer
	Signaling by Receptor Tyrosine Kinases	Diseases of signal transduction
	Cytokine Signaling in Immune system	Cell surface interactions at the vascular wall
	Hemostasis	Platelet activation, signaling and aggregation
	PI3K-Akt signaling pathway	MAPK family signaling cascades
	Signaling by Interleukins	Proteoglycans in cancer
	Pathways in cancer	MET activates PI3K/AKT signaling
		Acute myeloid leukemia
		Chronic myeloid leukemia
		RAF/MAP kinase cascade

row and peripheral blood samples from 500 AML patients will be taken at diagnosis and at resistance or first and subsequent relapses. These samples will be analyzed by Next Generation Sequencing (NGS) in order to sequence 26 consensus genes recurrently mutated in AML (ASXL1, HADH, CBL, CEBPA, DNMT3A, EZH2, FLT3, GATA2, IDH1, IDH2, JAK2, KIT, KRAS, MPL, MLL, NPM1, NRAS, PTPN11, RUNX1, SETBP1, SF3B1, SRSF2, TET2, TP53, U2AF1, WT1). With this data, it will be possible to determine which gene mutations can be classified as the driver or passenger mutations, and establish a diagnostic platform for rapid molecular diagnosis of the disease. As samples will be taken at resistance and relapse, this will also provide information regarding which genes/proteins are associated with unfavorable prognosis, which could be useful for determining prognosis at time of diagnosis, thus informing treatment options.

CONCLUSIONS

This *in silico* analysis revealed 28 proteins that could be considered potential biomarkers of poor prognosis in AML, with PTPN11 as the main node controlling the flow of information through the network.

One of the biggest challenges in biomarker research is that more often than not, a single biomarker is shared by several pathologies; so rather than a single protein biomarker, a panel of biomarkers is required in order to achieve the overall level of specificity needed. Therefore, this *in silico* approach is highly useful for informing which proteins could be included in such a panel, and which of these contribute significantly to the overall specificity and sensitivity.

It would be of great interest to perform a dependency analysis on this proposed panel of 28 proteins, in order to determine which nodes have a positive or negative influence on other nodes, thus identifying activators and inhibitors of the network. Perturbation experiments can also help identify which nodes are essential to the network, by eliminating them and observing how the behavior of the network changes. This optimization of the panel will ensure prognostic precision, while keeping costs down by avoiding unnecessary testing of biomarkers that do not significantly contribute. Wet bench research is enhanced when computational analysis is incorporated.

REFERENCES

Allan JN, Roboz GJ, Askin G, Ritchie E, Scandura J, Christos P, Hassane DC, Guzman ML (2018) CD25 expression and outcomes in older patients with acute myelogenous leukemia treated with plerixafor and decitabine. *Leuk. Lymphoma* **59**: 821–828. https://doi.org/10.1080/10428194.2017.1352089

Aref S, El-Sherbiny M, Azmy E, Goda T, Selim T, El-Refaie M, Twafik E (2008) Elevated serum endostatin levels are associated with favorable outcome in acute myeloid leukemia. *Hematology* **13**: 95–100. https://doi.org/10.1179/102453308X315898

Assenov Y, Ramirez F, Schelhorn S-E, Lengauer T, Albrecht M (2008) Computing topological parameters of biological networks. *Bioinformatics* **24**: 282–284. https://doi.org/10.1093/bioinformatics/btm554

Balss J, Thiede C, Bochtler T, Okun JG, Saadati M, Benner A, Pusch S, Ehninger G, Schaich M, Ho AD, Von Deimling A, Krmer A, Heilig CE (2016) Pretreatment d-2-hydroxyglutarate serum levels negatively impact on outcome in IDH1-mutated acute myeloid leukemia. *Leukemia* **30**: 782–788. https://doi.org/10.1038/leu.2015.317

Barabási A-L, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* **5**: 101–113. https://doi.org/10.1038/nrg1272

Bertoli S, Paubelle E, Bérard E, Saland E, Thomas X, Tavitian S, Larcher MV, Vergez F, Delabesse E, Sarry A, Huguet F, Larrue C, Bosc C, Farge T, Sarry JE, Michallet M, Récher C (2019) Ferritin heavy/light chain (FTH1/FTL) expression, serum ferritin levels, and their functional as well as prognostic roles in acute myeloid leukemia. *Eur. J. Haematol.* **102**: 131–142. https://doi.org/10.1111/ejh.13183

Bienz M, Ludwig M, Mueller BU, Leibundgut EO, Ratschiller D, Solenthaler M, Fey MF, Pabst T (2005) Risk assessment in patients with acute myeloid leukemia and a normal karyotype. *Clin. Cancer Res.* **11**: 1416–1424. https://doi.org/10.1158/1078-0432.CCR-04-1552

Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J (2009) ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**: 1091–1093. https://doi.org/10.1093/bioinformatics/btp101

Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, O'Donnell L, Oster S, Theesfeld C, Sellam A, Stark C, Breitkreutz B-J, Dolinski K, Tyers M (2017) The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* **45**: D369–D379. https://doi.org/10.1093/nar/gkw1102

Chen L, Chen W, Mysliwski M, Serio J, Ropa J, Abulwerdi FA, Chan RJ, Patel JP, Tallman MS, Paietta E, Melnick A, Levine RL, Abdel-Wahab O, Nikolovska-Coleska Z, Muntean AG (2015) Mutated Ptpn11 alters leukemic stem cell frequency and reduces the sensitivity of acute myeloid leukemia cells to Mcl1 inhibition. *Leukemia* **29**: 1290–1300. https://doi.org/10.1038/leu.2015.18

Chen Y, Ren S, Li S, Du Z (2017) Prognostic significance of osteopontin in acute myeloid leukemia: A meta-analysis. *Mol. Clin. Oncol.* https://doi.org/10.3892/mco.2017.1302

Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalatskaya I, May-Mahajan S, May B, Ndegwa N, Schmidt E, Shamovsky V, Yung C, Birney E, Hermjakob H, D'Eustachio P, Stein L (2011) Reactome: A database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**: 691–697. https://doi.org/10.1093/nar/gkq1018

Deng L, Richine BM, Virts EL, Jideonwo-Auman VN, Chan RJ, Kapur R (2018) Rapid development of myeloproliferative neoplasm in mice with PTPN11D61Y mutation and haploinsufficient for Dnmt3a. *Oncotarget* **9**: 6055–6061. https://doi.org/10.18632/oncotarget.23680

Fredly H, Reikvam H, Gjertsen BT, Bruserud Ø (2012) Disease-stabilizing treatment with all-trans retinoic acid and valproic acid in acute myeloid leukemia: Serum hsp70 and hsp90 levels and serum cytokine profiles are determined by the disease, patient age, and anti-leukemic treatment. *Am. J. Hematol.* **87**: 368–376. https://doi.org/10.1002/ajh.23116

Gao N, Wang XX, Sun JR, Yu WZ, Li XZ (2017a) Clinical impact of galectin-3 in newly diagnosed t (15;17)(q22;q21)/PML-RARa acute promyelocytic leukemia treated with all-trans retinoic acid and arsenic trioxide-based regimens. *Ann. Hematol.* **96**: 711–718. https://doi.org/10.1007/s00277-017-2948-3

Gao N, Yu WZ, Guo NJ, Wang XX, Sun JR (2017b) Clinical significance of galectin-3 in patients with adult acute myeloid leukemia: a retrospective cohort study with long-term follow-up and formulation of risk scoring system. *Leuk. Lymphoma* **58**: 1394–1402. https://doi.org/10.1080/10428194.2016.1243677

Grossmann KS, Rosário M, Birchmeier C, Birchmeier W (2010) The tyrosine phosphatase Shp2 in development and cancer. *Adv. Cancer Res.* **106**: 53–89. https://doi.org/10.1016/S0065-230X(10)06002-1

He AL, Bai J, Huang C, Zhang WG, Yang J, Wang JL, Meng X, Tian W (2010) [Preliminarily screening of serum characteristic markers in acute myeloid leukemia and clinical significance]. *Zhongguo Shi Yan Xue Ye Xue Za Zhi* **18**: 1132–1137.

Hu J, Sun Q, Fang W, Wang Q (2019) Effect of combination of all-trans retinoic acid and arsenic trioxide on apoptosis of acute promyelocytic leukemia cells. *Cell. Mol. Biol. (Noisy-le-grand)* **65**: 97–100. https://doi.org/10.14715/cmb/2019.65.4.16

Janin M, Mylonas E, Saada V, Micol JB, Renneville A, Quivoron C, Koscielny S, Scourzic L, Forget S, Pautas C, Caillot D, Preudhomme C, Dombret H, Berthon C, Barouki R, Rabier D, Auger N, Griscelli F, Chachaty E, Leclercq E, Courtier M-H, Bennaceur-Griscelli A, Solary E, Bernard OA, Penard-Lacronique V, Ottolenghi C, de Botton S (2014) Serum 2-hydroxyglutarate production in IDH1- And IDH2-mutated de novo acute myeloid leukemia: A study by the acute leukemia french association group. *J. Clin. Oncol.* **32**: 297–305. https://doi.org/10.1200/JCO.2013.50.2047

Končarević S, Lößner C, Kuhn K, Prinz T, Pike I, Zucht H-D (2014) In-Depth Profiling of the Peripheral Blood Mononuclear Cells Proteome for Clinical Blood Proteomics. *Int. J. Proteomics* **2014**: 1–9. https://doi.org/10.1155/2014/129259

Krug U, Röllig C, Koschmieder A, Heinecke A, Sauerland MC, Schaich M, Thiede C, Kramer M, Braess J, Spiekermann K, Haferlach T, Haferlach C, Koschmieder S, Rohde C, Serve H, Wörmann B, Hiddemann W, Ehninger G, Berdel WE, Büchner T, Müller-Tidow C, for the German Acute Myeloid Leukaemia Cooperative Group an (2010) Complete remission and early death after intensive chemotherapy in patients aged 60 years or older with acute myeloid leukaemia: A web-based application for prediction of outcomes. *Lancet* **376**: 2000–2008. https://doi.org/10.1016/S0140-6736(10)62105-8

Kümpers P, Koeneeke C, Hecker H, Hellpap J, Horn R, Verhagen W, Buchholz S, Hertenstein B, Krauter J, Eder M, David S, Göhring G, Haller H, Ganser A (2008) Angiopoietin-2 predicts disease-free survival after allogeneic stem cell transplantation in patients with high-risk myeloid malignancies. *Blood* **112**: 2139–2148. https://doi.org/10.1182/blood-2007-12-130021

Lazarevic V, Hörstedt AS, Johansson B, Antunovic P, Billström R, Derolf Å, Lehmann S, Möllgård L, Peterson S, Stockelberg D, Uggla B, Vennström L, Wahlin A, Höglund M, Juliusson G (2015) Failure matters: Unsuccessful cytogenetics and unperformed cytogenetics are associated with a poor prognosis in a population-based series of acute myeloid leukaemia. *Eur. J. Haematol.* **94**: 419–423. https://doi.org/10.1111/ejh.12446

Li Q, Li M, Wu YH, Zhu XJ, Zeng C, Zou P, Chen ZC (2014) Detection of cytokine expression patterns in the peripheral blood of patients with acute leukemia by antibody microarray analysis. *J. Huazhong Univ. Sci. Technol. – Med. Sci.* **34**: 176–180. https://doi.org/10.1007/s11596-014-1254-8

Liao Q, Wang B, Li X, Jiang G (2017) miRNAs in acute myeloid leukemia. *Oncotarget* **8**: 3666–3682. https://doi.org/10.18632/oncotarget.12343

Liersch R, Gerss J, Schliemann C, Bayer M, Schwöppe C, Biermann C, Appelmann I, Kessler T, Löwenberg B, Büchner T, Hiddemann W, Müller-Tidow C, Berdel WE, Mesters R (2012) Osteopontin is a prognostic factor for survival of acute myeloid leukemia patients. *Blood* **119**: 5215–5220. https://doi.org/10.1182/blood-2011-11-389692

Lilly AJ, Khanim FL, Hayden RE, Luong QT, Drayson MT, Bunce CM (2011) Nm23-H1 indirectly promotes the survival of acute myeloid leukemia blast cells by binding to more mature components of the leukemic clone. *Cancer Res.* **71**: 1177–1186. https://doi.org/10.1158/0008-5472.CAN-10-1704

Ma XJ, Zhang FH, Sun L (2017) Expression and significance of PD-L1, HSP90 and HSP90α in serum of patients with acute leukemia. *Zhongguo shi yan xue ye xue za zhi* **25**: 1384–1389. https://doi.org/10.7534/j.issn.1009-2137.2017.05.018

Melillo L, Cascavilla N, Lombardi G, Carotenuto M, Musto P (1992) Prognostic relevance of serum beta 2-microglobulin in acute myeloid leukemia. *Leukemia* **6**: 1076–1078

Mrożek K, Marcucci G, Nicolet D, Maharry KS, Becker H, Whitman SP, Metzeler KH, Schwind S, Wu YZ, Kohlschmidt J, Pettenati MJ, Heerema NA, Block AMW, Patil SR, Baer MR, Kolitz JE, Moore JO, Carroll AJ, Stone RM, Larson RA, Bloomfield CD (2012) Prognostic significance of the European LeukemiaNet standardized system for reporting cytogenetic and molecular alterations in adults with acute myeloid leukemia. *J. Clin. Oncol.* **30**: 4515–4523. https://doi.org/10.1200/JCO.2012.43.4738

Mustjoki S, Sidenius N, Sier CFM, Blasi F, Elonen E, Alitalo R, Vaheri A (2000) Soluble urokinase receptor levels correlate with number of circulating tumor cells in acute myeloid leukemia and decrease rapidly during chemotherapy. *Cancer Res.* **60**: 7126–7132

Nakase K, Kita K, Katayama N (2018) IL-2/IL-3 interplay mediates growth of CD25 positive acute myeloid leukemia cells. *Med. Hypotheses* **115**: 5–7. https://doi.org/10.1016/j.mehy.2018.03.007

Okabe-Kado J, Kasukabe T, Honma Y, Kobayashi H, Maseki N, Kaneko Y (2009a) Extracellular NM23 protein promotes the growth and survival of primary cultured human acute myelogenous leukemia cells. *Cancer Sci.* **100**: 1885–1894. https://doi.org/10.1111/j.1349-7006.2009.01276.x

Okabe-Kado J, Kasukabe T, Honma Y, Kobayashi H, Maseki N, Kaneko Y (2009b) Extracellular NM23-H1 protein inhibits the survival of primary cultured normal human peripheral blood mononuclear cells and activates the cytokine production. *Int. J. Hematol.* **90**: 143–152. https://doi.org/10.1007/s12185-009-0384-4

Patil A, Kinoshita K, Nakamura H (2010) Hub promiscuity in protein-protein interaction networks. *Int. J. Mol. Sci.* **11**: 1930–1943. https://doi.org/10.3390/ijms11041930

Prahallad A, Heynen GJJE, Germano G, Willems SM, Evers B, Vecchione L, Gambino V, Lieftink C, Beijersbergen RL, Di Nicolantonio F, Bardelli A, Bernards R (2015) PTPN11 is a central node in intrinsic and acquired resistance to targeted cancer drugs. *Cell Rep.* **12**: 1978–1985. https://doi.org/10.1016/j.celrep.2015.08.037

Qu CK, Shi ZQ, Shen R, Tsai FY, Orkin SH, Feng GS (1997) A deletion mutation in the SH2-N domain of Shp-2 severely suppresses hematopoietic cell development. *Mol. Cell. Biol.* **17**: 5499–5507. https://doi.org/10.1128/mcb.17.9.5499

Raman K (2010) Construction and analysis of protein–protein interaction networks. *Autom. Exp.* **2**: 2. https://doi.org/10.1186/1759-4499-2-2

Ran J, Li H, Fu J, Liu L, Xing Y, Li X, Shen H, Chen Y, Jiang X, Li Y, Li H (2013) Construction and analysis of the protein-protein interaction network related to essential hypertension. *BMC Syst. Biol.* **7**: 32. https://doi.org/10.1186/1752-0509-7-32

Rehman AU, Rahman MU, Khan MT, Saud S, Liu H, Song D, Sultana P, Wadood A, Chen H-F (2018) The Landscape of Protein Tyrosine Phosphatase (Shp2) and Cancer. *Curr. Pharm. Des.* **24**: 3767–3777. https://doi.org/10.2174/1381612824666181106100837

Scardoni G, Tosadori G, Pratap S, Spoto F, Laudanna C (2015) Finding the shortest path with PesCa: a tool for network reconstruction. *F1000Research*2–9. https://doi.org/10.12688/f1000research.6769.1

Schneeberger VE, Ren Y, Luetteke N, Huang Q, Chen L, Lawrence HR, Lawrence NJ, Haura EB, Koomen JM, Coppola D, Wu J (2015) Inhibition of Shp2 suppresses mutant EGFR-induced lung tumors in transgenic mouse model of lung adenocarcinoma. *Oncotarget* **6**: 6191–6202. https://doi.org/10.18632/oncotarget.3356

Shannon P, Markiel A, Owen Ozier 2, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2498–2504. https://doi.org/10.1101/gr.1239303.metabolite

Slovak ML, Kopecky KJ, Cassileth PA, Harrington DH, Theil KS, Paietta E, Willman CL, Head DR, Rowe JM, Forman SJ, Appelbaum FR, Mohamed A (2014) Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group study Karyotypic analysis predicts outcome of preremission and postremission t. 4075–4083.

Su YZ, Wang CB, Zhou Y, Sun NT (2015) Effects of changes in serum endostatin and fibroblast growth factor 19 on the chemotherapeutic sensitivity in acute myeloid leukemia patients. *Genet. Mol. Res.* **14**: 5181–5187. https://doi.org/10.4238/2015.May.18.8

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**: D362–D368. https://doi.org/10.1093/nar/gkw937

Tachibana T, Andou T, Tanaka M, Ito S, Miyazaki T, Ishii Y, Ogusa E, Koharazawa H, Takahashi H, Motohashi K, Aoki J, Nakajima

Y, Matsumoto K, Hagihara M, Hashimoto C, Taguchi J, Fujimaki K, Fujita H, Fujisawa S, Kanamori H, Nakajima H, Yokohama Cooperative Study Group for Hematology (YACHT) (2018) Clinical significance of serum ferritin at diagnosis in patients with acute myeloid leukemia: A YACHT multicenter retrospective study. *Clin. Lymphoma, Myeloma Leuk.* **18**: 415–421. https://doi.org/10.1016/j.clml.2018.03.009

Tsimberidou AM, Kantarjian HM, Wen S, O'Brien S, Cortes J, Wierda WG, Koller C, Pierce S, Brandt M, Freireich EJ, Keating MJ, Estey EH (2008) The prognostic significance of serum β2 microglobulin levels in acute myeloid leukemia and prognostic scores predicting survival: Analysis of 1,180 patients. *Clin. Cancer Res.* **14**: 721–730. https://doi.org/10.1158/1078-0432.CCR-07-2063

Walker A, Marcucci G (2012) Molecular prognostic factors in cytogenetically normal acute myeloid leukemia. *Expert Rev. Hematol.* **5**: 547–558. https://doi.org/10.1586/ehm.12.45

Wang JH, Chen WL, Li JM, Li Y, Wu SF, Chen TL, Zhu YM, Zhang WN, Li Y, Qiu YP, Zhao AH, Mi JQ, Jin J, Wang YG, Ma QL, Huang H, Wu DP, Wang QR, Yan XJ, Yan JS, Li J-Y, Wang S, Huang X-J, Wang B-S, Jia W, Shen Y, Chen Z, Chen S-J (2013) Prognostic significance of 2-hydroxyglutarate levels in acute myeloid leukemia in China. *Proc. Natl. Acad. Sci. U. S. A.* **110**: 17017–17022. https://doi.org/10.1073/pnas.1315558110

Yabushita T, Satake H, Maruoka H, Morita M, Katoh D, Shimomura Y, Yoshioka S, Morimoto T, Ishikawa T (2018) Expression of multiple leukemic stem cell markers is associated with poor prognosis in de novo acute myeloid leukemia. *Leuk. Lymphoma* **59**: 2144–2151. https://doi.org/10.1080/10428194.2017.1410888

Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M (2007) The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* **3**: 713–720. https://doi.org/10.1371/journal.pcbi.0030059