blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

An IDO1-related immune gene signature predicts overall survival in acute myeloid leukemia

Tracking no: ADV-2021-004878R1

Simone Ragaini (Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale, Università degli Studi di Bologna, Italy) Sarah Wagner (Nottingham Trent University, United Kingdom) Giovanni Marconi (University of Bologna, Italy) Sarah Parisi (Institute of Hematology and Medical Oncology "L. & A. Seràgnoli", Italy) Chiara Sartor (Università degli Studi di Bologna, Italy) Jacopo Nanni (Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna and Institute of Hematology and Medical Oncology "L. and A. Seràgnoli", Italy) Gianluca Cristiano (Istituto di Ematologia "Seràgnoli, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy, Italy) Annalisa Talami (Azienda Ospedaliero-Universitaria di Modena, Italy) Matteo Olivi (Università degli Studi di Bologna, Italy) Darina Ocadlikova (IRCCS Azienda Ospedaliero-Universitaria di Bologna, Istituto di Ematologia "Seràgnoli", Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale, Università di Bologna, Italy) Marilena Ciciarello (Institute of Hematology and Medical Oncology, Italy) Giulia Corradi (Institute of Hematology and Medical Oncology, Italy) Emanuela Ottaviani (Institute of Hematology and Medical Oncology "L. A. Seragnoli"-University of Bologna, Italy) Cristina Papayannidis (IRCCS, Azienda Ospedaliero-Universitaria di Bologna, Italy) Stefania Paolini (IRCCS Azienda Ospedaliero-Universitaria di Bologna, Istituto di Ematologia "Seràgnoli", Italy) Jayakumar Vadakekolathu (Nottingham Trent University, United Kingdom) Michele Cavo (IRCCS Azienda Ospedaliero-Universitaria di Bologna, Istituto di Ematologia, Italy) Sergio Rutella (Nottingham Trent University, United Kingdom) Antonio Curti (Institute of Hematology, Italy)

Abstract:
The contribution of the bone marrow (BM) immune microenvironment (TME) to acute myeloid leukemia (AML) development is well-known, but its prognostic significance is still elusive. Indoleamine 2,3-dioxygenase 1 (IDO1), which is negatively regulated by the BIN1 proto-oncogene, is an interferon (IFN)-□-inducible mediator of immune tolerance. With the aim to develop a prognostic IDO1-based immune gene signature, biological and clinical data of 732 patients with newly diagnosed, non-promyelocytic AML were retrieved from public datasets and analyzed using established computational pipelines. Targeted transcriptomic profiles of 24 diagnostic BM samples were analyzed using the NanoString's nCounter platform. BIN1 and IDO1 were inversely correlated and individually predicted overall survival. PLXNC1, a semaphorin receptor involved in inflammation and immune response, was the IDO1-interacting gene retaining the strongest prognostic value. The incorporation of PLXNC1 into the 2-gene IDO1-BIN1 score gave rise to a powerful immune gene signature predicting survival, especially in patients receiving chemotherapy. The top differentially expressed genes between IDO1low and IDO-1high and between PLXNC1low and PLXNC1 high cases further improved the prognostic value of IDO1 providing a 7 and 10-gene immune signature, highly predictive of survival and correlating with AML mutational status at diagnosis. Taken together, our data indicate that IDO1 is pivotal for the construction of an immune gene signature predictive of survival in AML patients. Given the emerging role of immunotherapies for AML, our findings support the incorporation of immune biomarkers into current AML classification and prognostication algorithms.

Conflict of interest: COI declared - see note

COI notes: Michele Cavo: Jannsen, BMS, Celgene, Sanofi, GlaxoSmithKline, Takeda, Amgen, Oncopeptides, AbbVie, Karyopharm, Adaptive: Consultancy, Honoraria. Sergio Rutella: NanoString Technologies, Inc.: Research Funding; MacroGenics, Inc.: Research Funding; Kura Oncology: Research Funding. Antonio Curti: Novartis, Pfizer, Abbvie: Consultancy, Honoraria, Research Funding. The other authors declare that they have not conflicts of interest

Preprint server: No;

Author contributions and disclosures: S. Ragaini, S.W., S. Rutella and A.C. designed the project and supervised the study. S. Ragaini, S.W., J.V. and S. Rutella processed primary patient samples, performed in silico and bioinformatic analyses and interpreted gene expression data. S. Parisi, G.M., C.S., J.N., G. Cristiano, A.T., M.O., D.O, M. Ciciarello, G. Corradi, E.O., C.P., S. Paolini, J.V. contributed to data collection and reviewed the manuscript. S. Ragaini, S. Rutella and A. Curti wrote the paper. M. Cavo supervised the study and reviewed the manuscript. All authors provided final approval of the manuscript. S. Rutella and A.C. share the senior authorship of this study.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: The results shown in this paper are in part based upon data retrieved from the Cancer Genome Atlas (TCGA) profiling project and the HOVON (E-MTAB-3444) dataset. The data referring to the validation cohort are included in the Gene Expression Omnibus database (accession number GSE146204).

Clinical trial registration information (if any):

Title:

An *IDO1*-related immune gene signature predicts overall survival in acute myeloid leukemia

Running title:

A prognostic *IDO1*-related immune signature for AML

Authors:

Simone Ragaini[1,2], Sarah Wagner[3], Giovanni Marconi[1], Sarah Parisi[1], Chiara Sartor[1], Jacopo Nanni[1], Gianluca Cristiano[1], Annalisa Talami[1], Matteo Olivi[1,2], Darina Ocadlikova[1], Marilena Ciciarello[4], Giulia Corradi[1], Emanuela Ottaviani[4], Cristina Papayannidis[4], Stefania Paolini[4], Jayakumar Vadakekolathu[3], Michele Cavo[1,4], Sergio Rutella[3,5] and Antonio Curti[4]

# S. Rutella and A.C. contributed equally to this study

Affiliations

1. Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale, Università degli Studi, Bologna, Italy
2. Division of Hematology, Department of Molecular Biotechnology and Health Sciences, University of Torino, A.O.U., Città della Salute e della Scienza di Torino, Italy
3. John van Geest Cancer Research Centre, School of Science and Technology, Nottingham Trent University, Nottingham, United Kingdom
4. IRCCS Azienda Ospedaliero-Universitaria di Bologna, Istituto di Ematologia "Seràgnoli", Bologna, Italy
5. Centre for Health, Ageing and Understanding Disease (CHAUD), Nottingham Trent University, Nottingham, United Kingdom

Corresponding author's information

Michele Cavo

Istituto di Ematologia "Seràgnoli"

Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale,

Università degli Studi di Bologna,

IRCCS Azienda Ospedaliero-Universitaria di Bologna;

Mail: michele.cavo@unibo.it

Text word count 3836

Abstract word count 236

Number of figures 4

Number of tables 5

Number of references 47

Key points

- The semaphorin receptor *PLXNC1* is an *IDO1*-interacting gene and a strong predictor of survival in acute myeloid leukemia (AML)
- An *IDO1*-related immune gene signature, including *PLXNC1*, predicts survival in AML

Abstract

The contribution of the bone marrow (BM) immune microenvironment (TME) to acute myeloid leukemia (AML) development is well-known, but its prognostic significance is still elusive. Indoleamine 2,3-dioxygenase 1 (IDO1), which is negatively regulated by the *BIN1* proto-oncogene, is an interferon (IFN)-γ-inducible mediator of immune tolerance. With the aim to develop a prognostic *IDO1*-based immune gene signature, biological and clinical data of 732 patients with newly diagnosed, non-promyelocytic AML were retrieved from public datasets and analyzed using established computational pipelines. Targeted transcriptomic profiles of 24 diagnostic BM samples were analyzed using the NanoString's nCounter platform. *BIN1* and *IDO1* were inversely correlated and individually predicted overall survival. *PLXNC1*, a semaphorin receptor involved in inflammation and immune response, was the *IDO1*-interacting gene retaining the strongest prognostic value. The incorporation of *PLXNC1* into the 2-gene *IDO1-BIN1* score gave rise to a powerful immune gene signature predicting survival, especially in patients receiving chemotherapy. The top differentially expressed genes between $IDO1^{low}$ *and* $IDO\text{-}1^{high}$ and between $PLXNC1^{low}$ *and* $PLXNC1^{high}$ cases further improved the prognostic value of *IDO1* providing a 7 and 10-gene immune signature, highly predictive of survival and correlating with AML mutational status at diagnosis. Taken together, our data indicate that *IDO1* is pivotal for the construction of an immune gene signature predictive of survival in AML patients. Given the emerging role of immunotherapies for AML, our findings support the incorporation of immune biomarkers into current AML classification and prognostication algorithms.

Introduction

Acute myeloid leukemia (AML) is a molecularly and clinically heterogeneous hematologic malignancy that progresses rapidly and originates from a rare population of leukemic stem cells. Despite intensive chemotherapy and stem cell transplantation, the outcome of AML has not changed substantially in the last decades. The estimated 5-year overall survival (OS) rate is approximately 30% and death rates have remained stable[1]. In this scenario, innovative strategies and tools are urgently needed to improve outcomes for AML patients.

In many cancers, such as diffuse large B-cell lymphoma[2], primary breast cancer[3,4], melanoma[5], gastric cancer[6] and lung adenocarcinoma[7,8], the tumor microenvironment (TME) has been shown to retain a prognostic value that can be asserted by immune-specific gene expression patterns.[9,10] However, the current risk classification[11] of AML is exclusively focused on leukemic cell-intrinsic cytogenetic and molecular alterations, which have historically been known to impact on response to conventional chemotherapy and risk of relapse. Compelling preclinical data clearly demonstrate the impact of tolerogenic mechanisms played by TME in dysregulating patients' immune response to AML cells. Very recently, the prognostic significance of immune landscape has been addressed in AML, revealing that immune-related genes may predict response to therapy and survival[12].

Indoleamine 2,3-dioxygenase 1 (IDO1) catalyzes the rate-limiting step in tryptophan metabolism along the kynurenine pathway. In tumors, *IDO1* is negatively controlled by the *BIN1* tumor suppressor[13], which in turn is regulated by the RBM25 splicing factor generating a dominant-negative *BIN1* isoform that is unable to repress MYC activity[14]. We and others have

shown that *IDO1* is expressed in a significant proportion of AML patients at disease onset[15] where it promotes the establishment of an immunosuppressive TME through the induction of T regulatory cells (Tregs)[16–19]. Previous studies investigated the impact of *IDO1* on AML survival. In particular, *IDO1* mRNA expression in the bone marrow (BM), evaluated by gene expression profiling or quantitative real-time polymerase chain reaction (qRT-PCR) was a predictor of shorter survival[20,21]. Recently, an immunohistochemical score based on IDO1 expression was shown to predict early mortality in AML[22]. Herein, we interrogated public AML transcriptomic datasets and profiled primary BM samples from newly diagnosed AML patients with the aim to identify an *IDO1*-related immune gene signature that may further refine our ability to predict survival.

Materials and methods

Data sources

For the *in-silico* generation of a prognostic *IDO1*-associated gene signature, three publicly available independent gene expression datasets were used. Biological and clinical data of 982 patients with newly diagnosed non-promyelocytic AML patients, including complete cytogenetic, immunophenotypic, and clinical annotations, were retrieved from The Cancer Genome Atlas (TCGA) profiling project, the HOVON (E-MTAB-3444) dataset[23] and the GSE106291 dataset (available through Gene Expression Omnibus, GEO)[24]. The TCGA series consisted of RNA-sequencing data (Illumina HiSeq 2000) available through cBioPortal for Cancer Genomics at https://www.cbioportal.org/[25,26] and Xena Platform at http://xena.ucsc.edu/[27]. In the TCGA-AML dataset, only patients treated with curative intent on a "7+3" chemotherapy backbone (n=123) were considered for survival analyses.

The HOVON series (available through Array Express; E-MTAB-3444) consisted of gene array data from AML patients treated according to Dutch-Belgian Hemato-Oncology Cooperative

Group and the Swiss Group for Clinical Cancer Research (HOVON/SAKK) AML-04, -04A, -29, -32, -42, -42A, -43 and -92 protocols (available at http://www.hovon.nl). In particular, clinical and pathological data were available for 609 patients[28] HOVON patients aged more than 65 years were excluded from survival analyses. Patients' characteristics are summarized in **Table 1 and 2**.

The GSE106291 dataset (available through Gene Expression Omnibus, GEO) consists of gene expression data by high throughput sequencing from AML patients treated in the AMLCG-2008 (NCT01382147, n=210) and AMLG-1999 trials (NCT00266136, n=40). Survival data were available in 248/250 patients in the GSE106291 series. The results shown in this paper are in part based upon data retrieved from the Cancer Genome Atlas (TCGA) profiling project and the HOVON (E-MTAB-3444) dataset. The data referring to the validation cohort are included in the Gene Expression Omnibus database (accession number GSE146204).

IDO1-related gene normalization and co-expression analysis

The analysis of genes with a coordinated expression pattern in TCGA-AML was performed through the cBioPortal platform[25,26]. The top 5 genes correlated with *IDO1* were, then, selected and their impact on survival was evaluated using the Kaplan-Meier method.

Validation cohort

Twenty-four BM samples from patients with non-promyelocytic AML were collected at diagnosis and were used to validate the correlation between *IDO1* and *PLXNC1* expression. Median age at diagnosis was 55 years; 18 patients were male and 6 were female (patients' clinical data are summarized in **Table 3**). Patients provided written informed consent. The investigations were conducted in accordance with the Declaration of Helsinki. Approval was obtained from Bologna Hospital Ethics Committee (94/201/O/Tess).

Cox regression analysis and computation of gene prognostic signatures

Gene expression was min-max normalized to a value between 0 and 1. The association between gene expression and survival time was evaluated using Cox regression analysis. Beta coefficients from Cox PH models were used to assign a prognostic weight to each individual gene in a given signature. According to a previously published formula[29], we developed a gene signature calculated as the linear combination of mRNA expression weighted by the regression coefficient (β) derived from multivariate Cox regression analysis with overall survival as a dependent variable. Regarding the *IDO1-BIN1-PLXNC1* signature, patients were stratified into 3 groups (low score group, intermediate score group and high score group) using the 25° and 75° percentiles of the score as cut-off. A summary of the key signatures used in this study is reported in **Supplementary Table 1.**

Probe ID selection

In the HOVON dataset, the following probe IDs were used for the analysis of mRNA expression: 210029_at (*IDO1*), 206470_at (*PLXNC1*), 210202_s_at (*BIN1*), 212028_at (*RBM25*), 209341_s_at (*IKBK3*), 204420_at (*FOSL1*), 223903_at (*TLR9*), 210321_at (*GZMH*), 205495_s_at (*GNLY*), 217502_at (*IFIT2*), and 204747_at (*IFIT3*).

Real-time polymerase chain reaction

Total RNA was extracted from mononuclear cells isolated from the BM of our validation cohort (n=24) using the Qiagen RNeasy kit according to the manufacturer's protocol. Quality control of the isolated RNA was performed using an Agilent bioanalyzer and NanoDrop 8000 (Thermo Scientific, Waltham MA, USA). An $OD_{260/280}$ ratio between 1.8 and 2.2 and a RIN value above 9.0 was considered for further processing. For cDNA synthesis, 1 μg of denatured total RNA was reverse transcribed using an Improm II kit and random hexamers (both from

Promega, Madison WI, USA) in a 20 μl final volume according to the manufacturer's instruction. qRT-PCR was performed in a 96-well Optical Reaction Plate using the ABI-PRISM 7900 Sequence Detection System (Applied Biosystems, Foster City CA, USA). Threshold cycle (C_t) values were determined automatically. Relative quantification was calculated using ΔCt comparative method[30]. Primer probes for PLXNC1 Hs00194968_m1, IDO1 Hs00158027_m1 and glyceraldehyde 3-phosphate dehydrogenase (GAPDH), Hs00266705_g1 were purchased from Applied Biosystems.

NanoString nCounter™ platform

Gene expression analysis was performed on the nCounter™ platform (NanoString Technologies Inc., Seattle, WA)[31] using the PanCancer IO 360™ Gene Expression Panel[32]. 150 ng of total RNA from 24 primary BM samples was used in each reaction. Hybridization of probes was carried out at 65°C for 20 hours. Post-hybridization samples were purified using a NanoString Prep Station and immobilized onto a cartridge. Raw data were acquired using the nCounter® FLEX Analysis System with a scanning resolution of 555 FOV using the probe annotation file NS_IO360_V1.0. Quality controls, data normalization and differential expression analysis were performed using the nSolver advanced analysis module (version 2.0.115) according to the manufacturer's instructions. The Benjamini-Yekutieli (B-Y) method was used to generate an adjusted p-value (BY.p.value).

Pan-cancer analysis

Survival meta-analyses of TCGA cancer datasets were performed using the GEPIA2 web-server at http://gepia2.cancer-pku.cn/ [33].

Deconvolution analysis

Deconvolution analysis of AML-TCGA and Genotype-Tissue Expression (GTEx) datasets was performed using the GEPIA2021[34] web-server (http://gepia2021.cancer-pku.cn/) through the EPIC method[35] to estimate the proportion of immune and stromal cells from bulk gene expression data.

Statistical methods

Differences between sets of data were considered statistically significant for P-values <0.05. Statistical analyses were performed using the IBM *SPSS Statistics* (version 25), GraphPad Prism software packages *(*version 7) and R (version 4.0.4).

Results

A two-gene BIN1-IDO1 signature predicts OS in AML

In the attempt to develop an *IDO1*-related immune signature which predicts clinical outcome in AML, we initially focused on *BIN1*, a master regulator of *IDO1* in solid tumors[13], and we found that *IDO1* and *BIN1* mRNA expressions were anti-correlated (r = - 0.41, P<0.0001, **Figure 1A**). Of note, *RMB25,* known to be a key modulator of *BIN1* expression[14], correlated both with *BIN1* (Pearson R = - 0.29, P<0.0001, **Supplementary Figure 1A)** and *IDO1* (Pearson R = 0.46, P<0.0001, **Supplementary Figure 1B**). We next investigated the impact of *IDO1* and *BIN1* expression on AML survival. To minimize any bias due to differences in treatment approaches (intensive *versus* non-intensive, curative *versus* palliative) and to increase comparability among patient groups, we stratified HOVON cases based on patient age. We analyzed only patients aged less than 65 years, thus reducing the likelihood of including patients who were not treated with curative intent. We plotted the normalized *IDO1* and *BIN1* mRNA expression in a single score using a previously reported formula[29] and split the HOVON cohort into three groups according to the score quartile. OS was significantly different among the three score groups (P<0.01, **Figure 1B**). With a median follow up of 8.1 years [95% CI, 7.0 - 9.2], patients in the low and intermediate score groups showed significantly longer OS than those in the high-risk score (low score group: median OS = 1.9 years [95% CI, 0.2 - 3.5],

intermediate score group: median OS = 1.8 years [95% CI, 1.3 - 2.3], high score group: median OS = 1.1 years [95% CI, 0.8 - 1.4]). In particular, the comparison among high, intermediate, and low score groups showed a hazard ratio (HR) of 1.5 (95% CI, 1.1- 2.0, P<0.01). These data therefore indicate that an *IDO1*-centered gene signature may predict OS in AML.

PLXNC1 correlates with *IDO1* and affects OS both in TCGA- and HOVON-AML cases

To identify other genes in the network that could be incorporated into the prognostic score based on *IDO1-BIN1* signature, we performed a co-expression analysis on TCGA cases[25,26] (**Table 4**). This approach identified *IDO2*, *CD1C, CD1E, XCR1, PLXNC1* as the top 5 co-expressed genes. The median expression of each gene was selected as a cut-off to split TCGA-AML patients into two groups (high and low) and was then correlated with patient survival. Interestingly, among the top 5 *IDO1*-correlated genes, only *IDO2* and *PLXNC1* showed a significant impact on AML survival in univariate analyses (log-rank P<0.05, **Figures 1C and 1D**). The correlations between *IDO1*, *IDO2* and *PLXNC1* (*IDO1 vs IDO2*: r = - 0.27, 95% C.I. [0.19 to 0.33], P<0.0001, **Figure 1E**; *IDO1 vs PLXNC1*: r = - 0.25, 95% C.I. [- 0.33 to - 0.18], P<0.0001, **Figure 1F**) as well as their impact on AML survival were independently validated in the HOVON dataset (*IDO2:* log-rank P<0.05, **Figure 1G** and *PLXNC1*: log-rank P<0.001, **Figure 1H**). As shown by Cox regression analyses, the *IDO1-PLXNC1* signature was the only predictor of survival (*IDO1*: β=0.68, HR=1.99 and P<0.05; *PLXNC1*: β=0.84, HR=2.32 and P<0.01*)*. Taken together, these data indicate that *PLXNC1* is a novel *IDO1*-related gene that stratifies survival.

Incorporation of PLXNC1 into the IDO1-BIN1 score improves the predictive power of the gene signature

Given the established interactions between *IDO1*, *PLXNC1* and *BIN1* and their potential impact on AML survival, we plotted the three genes together in a 3-gene signature to develop a new

score, which was then tested for its prognostic value. The resulting *IDO1-BIN1-PLXNC1* signature was predictive of overall survival (*IDO1*: β=1.032, HR=2.81, P<0.01; *BIN1*: β=0.758, HR=2.13, P<0.05 and *PLXNC1*: β=0.820, HR 2.27, P<0.01, **Table 5**), which prompted us to split patients into three groups using score quartiles as cut-off.

Kaplan Meier analysis showed a significantly different OS for the three score groups (P<0.001). In particular, among the 572 HOVON patients, the highest score predicted the shortest survival. With a median follow up of 8.1 years [95% CI, 7.0 – 9.2], low and intermediate score groups showed a median OS of 2.9 years [95% CI, 0.0 – 6.0] and 1.6 years [95% CI, 1.2 – 2.1], respectively, whereas high score group correlated with a median OS of 1.1 years [95% CI, 0.8– 1.5], (**Figure 2A**). Intermediate and high score groups versus low score one showed hazard ratios (HR) of 1.2 (95% CI 1.0 - 1.6, P=NS) and 1.8 (95% C.I. 1.4 - 2.4, P<0.01), respectively. However, the *IDO1-BIN1-PLXNC1* signature did not significantly stratify patients according to ELN cytogenetic risk groups (**Supplementary Figure 2A-D)**. In *FLT3*-wild type patients score values resulted significantly higher than those of *FLT3*-mutated ones (P<0.01, **Supplementary Figure 2E**). Among *FLT3*-wild type patients, the score remained statistically significant (P<0.0001, **Supplementary Figure 2F**). Of note, *IDO1-BIN1-PLXNC1*-based score was capable of predicting OS both in patients treated only with chemotherapy (**Figure 2B**) and in those who received allogeneic hematopoietic stem transplantation (HSCT) (**Figure 2C**). To validate these results, the score was implemented in the TCGA data set by using the same Cox regression coefficients derived from the analysis performed on the HOVON cases. Overall, the predictive ability of the score remained highly significant (P<0.01, **Figure 2D**). However, the score impacted on OS of patients who received only chemotherapy (P<0.0001, **Figure 2E**), whereas no statistically significant difference was observed in patients undergoing allogeneic HSCT (**Figure 2F**). We further confirmed the three different survival groups identified by *IDO1-BIN1-PLXNC1* signature in the GSE106291 dataset (P<0.05, **Supplementary Figure 3).**

Taken together, these data suggest that *IDO1-BIN1-PLXNC1* gene signature may predict OS in AML patients, especially when treated only with chemotherapy.

Targeted immune transcriptomic profiling of IDO1 and PLXNC1 high/low patients uncovers non-overlapping pathways and identifies a parsimonious gene set predicting AML outcome.

Firstly, a deconvolution analysis was applied to the AML-TCGA and Genotype-Tissue Expression (GTEx) datasets through the EPIC method[35] to explore *IDO1* and *PLXNC1* gene sub-expression in single cell-types. Interestingly, both *IDO1* and *PLXNC1* expression by B-cells, T cells and macrophages resulted to be higher in AML samples when compared to healthy donors bone marrow samples (**Supplementary Figure 4A and 4B)**.

We next profiled unfractionated BM samples from a cohort of 24 patients with newly diagnosed AML using the nCounter platform, which allows the quantitative measurement of mRNA species without RNA amplification. To identify transcriptional patterns associated with changes in *IDO1* and *PLXNC1* expression, the patient cohort was split according to the median value of *IDO1* and *PLXNC1* as detected by RT-PCR. The top 20 differentially expressed (DE) genes between *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ samples (value threshold of 0.01; log_2 fold-change threshold of 1.4, **Supplementary Table 2 and 3**) showed negligible overlap (**Figure 3A**), suggesting that *PLXNC1* and *IDO1* expression may reflect non-redundant biological processes in AML. We next showed that the expression of the top 20 DE genes between *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ samples was higher in TCGA-AML cases compared with blood samples from healthy donors available through the Genotype-Tissue Expression (GTEx) project (**Figure 3B**). Pathway enrichment analysis using the DE genes between *PLXNC1*$^{high/low}$ samples as an input showed that viral infection, apoptosis regulation, glucose metabolism and *c-met* signaling were among the most significantly enriched pathways in samples with high expression of *PLXNC1* (**Figure 3C**). In contrast, samples with high *IDO1* expression were significantly enriched in IFN signaling, cytokine-cytokine receptor interaction and IFN-γ

response pathways (**Figure 3D**). Notably, the top 20 DE genes between *PLXNC1*$^{high/low}$ samples were able to stratify survival in TCGA-cases. In particular, patients with higher than median gene expression showed significantly worse survival estimates than patients with lower than median gene expression (HR=2.0; log-rank *P*=0.002; **Figure 4A**). In addition, the DE genes between *IDO1*$^{high/low}$ samples were able to assist outcome prediction in TCGA-AML (HR=2.0; log-rank *P*=0.002; **Figure 4B)**.

In the attempt to refine our 20-gene signatures, we focused on genes that were individually associated with significant differences in OS in TCGA cases. Interestingly, *IKBKB, FOSL1* and *TLR9* in the *PLXNC1*$^{high/low}$ signature and *GZMH*, *GNLY*, *IFIT2* and *IFIT3* in the *IDO1*$^{high/low}$ signature stratified patient survival (HR=2.4 and log-rank *P*<0.0001; HR=1.6 and log-rank *P*=0.03, respectively; **Figures 4C** and **4D**). The ability to predict outcomes was improved by combining the 7 DE genes in the *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ signatures (HR=1.6 and log-rank *P*=0.029; **Figure 4E**). As next step, we asked whether deregulated expression of the 7 DE genes correlated with specific molecular features in TCGA cases. As shown in **Figure 4F**, no mutations of the 7 DE genes were documented in TCGA patients. In contrast, abnormalities in the 7 genes utilized in the query (by default, non-synonymous mutations, fusions, amplifications and deep deletions) were detected in 28% of TCGA cases (**Figure 4G**) and were significantly enriched in patients with adverse-risk molecular features, including *TP53* and *KRAS* mutations [36] (*P=0.016* and *P=0.019*, respectively). Further analyses of mutual exclusivity and co-occurrence patterns indicated that *CEBPA* mutations, which correlate with more favorable prognosis [36], were negatively associated with abnormalities in the 7 DE genes (*P=0.0159*). Moreover, we added the *IDO1*, *BIN1* and *PLXNC1* genes to the previously discovered 7 DE genes. The combined 10-gene signature was highly predictive of AML outcome (HR=2.6 and log-rank *P*<0.0001; **Figure 4H**). The 7- and 10-gene signatures also predicted survival in the HOVON dataset (respectively P<0.05, **Supplementary Figure 5A** and P<0.0001, **Supplementary Figure 5B**). When patients were stratified by cytogenetic risk,

no statistically significant differences in survival emerged according to the 10-gene signature (**Supplementary Figure 5C-E**). Furthermore, regarding *FLT3* mutational status, the 10-gene signature enabled survival prediction only in *FLT3* wild-type patients (P<0.001, **Supplementary Figure 5F**).

Finally, we performed a pan-cancer analysis of TCGA solid tumor types. The expression of genes in the *PLXNC1*-derived and *IDO1*-derived signatures in matched tumor samples and adjacent normal tissues is shown in **Supplementary Figures 6 and 7**. This analysis indicated that individual genes in the *PLXNC1*-derived signature may retain prognostic relevance also in selected solid tumor cell types, including low-grade glioma, and hepatocellular, lung and adrenocortical carcinoma (**Supplementary Figure 8A**). In contrast, the prognostic power of genes in the *IDO1*-derived signature was restricted to low-grade glioma, thymoma and uveal melanoma (**Supplementary Figure 8B)**. Taken together, these data indicate that *IDO1* and *PLXNC1* are implicated in non-overlapping biological mechanisms and may refine the accuracy of survival prediction in AML.

Discussion

Our data indicate that *IDO1* is pivotal for the construction of an immune gene signature predictive of survival in AML patients. Based on its well-known immunologic properties, *IDO1* was used as a key input gene to further explore the immunogenomic AML landscape and to identify genes that could be incorporated into a novel prognostic signature for newly diagnosed AML. We identified a previously unexplored correlation between *IDO1* and *BIN1* in AML and we demonstrated that this 2-gene score predicts OS. Indeed, inclusion of the *IDO1*-interacting gene *PLXNC1* improved the predictive ability of the *IDO1-BIN1* signature, especially in patients who received chemotherapy. This observation prompted us to explore the *IDO1*-related gene network in depth, leading to the identification of highly predictive 7 and 10-gene immunological signatures.

In solid tumors, *BIN1* negatively regulates *IDO1* expression at the level of IFN-γ-related transcription program via STAT1- and NF-κB[13]. Herein, we report for the first time that a similar negative correlation between *IDO1* and *BIN1* mRNA expression exists in AML. Moreover, we showed that a low expression of the key *BIN1*-regulator *RBM25* [14] correlates with high *BIN1* and low *IDO1* levels in AML, suggesting a common molecular pathway in *IDO1* gene regulation, which may be shared across tumors of different histological types and cell of origin. In line with a previous report[13], our data highlight that *IDO1* expression is mainly associated to IFN-related pathways. Of note, our study identified a novel set of *IDO1*-interacting genes, among which *PLXNC1* emerged as a crucial and master one. In myeloid precursors, *PLXNC1* expression is restrained by the RUNX1 transcription factor and dependent on KIT signaling[37]. Although its function is still elusive, Plexin C1 as receptor for semaphorin 7A[38] dampens the acute inflammatory response through the regulation of dendritic cell (DC) activity and migration[38,39]. Our data expand to the AML setting the characterization of *PLXNC1*. In particular, *PLXNC1* was co-regulated with genes involved in T-cell differentiation, lymphocyte proliferation and activation, consistently with the above-mentioned preferential activity of Plexin C1 in the activation of T-cell immune response via DCs. Overall, this analysis revealed an enrichment in pathways correlated with immune response, thus supporting our hypothesis that an *IDO1*-centered gene signature may be a useful tool to prognostically dissect AML immunological landscape.

By first moving from the negative correlation between *IDO1* and *BIN1*, we demonstrated that these two genes constitute a molecular signature, which may predict OS. The addition of *PLXNC1* to *IDO1* and *BIN1* resulted in a more powerful gene immune signature predicting survival, which was further implemented leading to the identification of a highly predictive 7- and 10-gene immunological signature. The robustness of the proposed immune signatures is confirmed by the fact that they retain their predictive value when applied to independent AML datasets. In both HOVON and TCGA-AML, the ability of the *IDO1-BIN1-PLXCN1* gene

signature to stratify survival was highly significant especially when we analyzed patients who received chemotherapy. Although this finding warrants further investigation, it may suggest that, along with established leukemic cell-intrinsic chromosomal translocations and genetic mutations, cell-autonomous and immune-related factors, such as those deriving from the immune TME, may contribute to regulate response to conventional chemotherapy. Indeed, an increasing body of evidence has highlighted the immunomodulatory effects of some antineoplastic agents, especially anthracyclines, which may act as adjuvants of the immune system along with inducing anti-proliferative effects on tumor cells[40]. In AML, we reported that chemotherapy is capable to reshape leukemic microenvironment by activating immune effector T cells and, concomitantly, inducing IDO1-expressing tolerogenic DCs and Tregs[41]. These data support the notion that the immune composition of BM microenvironment may influence the response to chemotherapy.

Of note, a close association between the immunological TME and cancer-intrinsic genomic alterations has been recently highlighted[42] and correlated to the response to chemotherapy. In AML, a specific TME-related immunogenomic profile correlates with increased chemoresistance and with response to immunotherapy[43] and a correlation between *TP53* mutations and an immunosuppressive TME has been recently established[44,45]. Of note, our targeted immune transcriptomic profiling revealed that abnormalities in immune-related genes in the *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ signatures were more frequently documented in patients with adverse-risk molecular features, including *TP53* and *KRAS* mutations, whereas they were negatively correlated with *CEBPA* mutations, known to confer favorable prognosis and better response to chemotherapy. Consistently, higher levels of *PLXNC1* were observed in AML patients with cytogenetic abnormalities, whereas lower mRNA levels were reported in patients with *CEBPA* mutations and with inv(16) or t(8;21)[46,47].

Regarding patients who received chemotherapy and allogeneic transplantation, the predictive ability of the *IDO1-BIN1-PLXCN1* gene signature is retained in the HOVON dataset whereas it

is not confirmed in the TCGA dataset. Although a formal demonstration was not the main focus of our work, we have argued that the differences in the predicting value of our score among patients who underwent allogeneic transplantation and those who received only chemotherapy may rely on the impact that allogeneic transplantation could exert on tumor immunologic microenvironment.

In conclusion, our data shed light into the biological significance of immune-related gene networks in AML prognostication. In this scenario, *IDO1* emerged as pivotal and paramount for the construction of powerful predictive immune gene signatures. In an era of emerging novel approaches targeting the immune system, our results highlight the need to integrate immunogenomic biomarkers into current AML prognostic classification system.

Data sharing statement: The results shown in this paper are in part based upon data retrieved from the Cancer Genome Atlas (TCGA) profiling project and the HOVON (E-MTAB-3444) dataset. The data referring to the validation cohort are included in the Gene Expression Omnibus database (accession number GSE146204).

Prior presentation: This study was presented in abstract form at the 62nd Annual Meeting of the American Society of Hematology (2020).

Acknowledgements

The research was supported by FATRO/Foundation Corrado and Bruno Maria Zaini-Bologna, Fabbri1905 and Bologna AIL (Associazione Italiana contro le Leucemie)/Section of Bologna, Associazione Italiana per la Ricerca sul Cancro (IG20654 to AC), the John and Lucille van Geest Foundation (to S. Rutella). MC was supported by the University of Bologna (Alma Idea Junior Grant 2017). We would like to thank Marie-Paule Vedrine, and Bologna AIL (Associazione Italiana contro le Leucemie)/Section of Bologna for the kind support.

Author Contributions

S. Ragaini, S.W., S. Rutella and A.C. designed the project and supervised the study.

S. Ragaini, S.W., J.V. and S. Rutella processed primary patient samples, performed *in silico* and bioinformatic analyses and interpreted gene expression data.

S. Parisi, G.M., C.S., J.N., G. Cristiano, A.T., M.O., D.O, M. Ciciarello, G. Corradi, E.O., C.P., S. Paolini, J.V. contributed to data collection and reviewed the manuscript.

S. Ragaini, S. Rutella and A. Curti wrote the paper.

M. Cavo supervised the study and reviewed the manuscript.

All authors provided final approval of the manuscript.

S. Rutella and A.C. share the senior authorship of this study.

Competing Interests statement

Michele Cavo: Jannsen, BMS, Celgene, Sanofi, GlaxoSmithKline, Takeda, Amgen, Oncopeptides, AbbVie, Karyopharm, Adaptive: Consultancy, Honoraria.

Sergio Rutella: NanoString Technologies, Inc.: Research Funding; MacroGenics, Inc.: Research Funding; Kura Oncology: Research Funding**.**

Antonio Curti: Novartis, Pfizer, Abbvie: Consultancy, Honoraria, Research Funding.

The other authors declare that they have not conflicts of interest

References

1. Surveillance, epidemiology and end results (SEER) program cancer stat facts. Leukemia - acute myeloid leukemia (AML). [Internet].
2. Alizadeh AA, Gentles AJ, Alencar AJ, Liu CL, Kohrt HE, Houot R, et al. Prediction of survival in diffuse large B-cell lymphoma based on the expression of 2 genes reflecting tumor and microenvironment. Blood. 2011;118(5):1350–8.
3. Perez EA, Thompson EA, Ballman K V, Anderson SK, Asmann YW, Kalari KR, et al. Genomic analysis reveals that immune function genes are strongly linked to clinical outcome in the North Central Cancer Treatment Group n9831 Adjuvant Trastuzumab Trial. J Clin Oncol. 2015;33(7):701–8.
4. Stoll G, Enot D, Mlecnik B, Galon J, Zitvogel L, Kroemer G. Immune-related gene signatures predict the outcome of neoadjuvant chemotherapy. Oncoimmunology. 2014;3(1):e27884.
5. Chen P-L, Roh W, Reuben A, Cooper ZA, Spencer CN, Prieto PA, et al. Analysis of Immune Signatures in Longitudinal Tumor Samples Yields Insight into Biomarkers of Response and Mechanisms of Resistance to Immune Checkpoint Blockade. Cancer Discov. 2016;6(8):827–37.
6. Park C, Cho J, Lee J, Kang SY, An JY, Choi MG, et al. Host immune response index in gastric cancer identified by comprehensive analyses of tumor immunity. Oncoimmunology. 2017;6(11):e1356150–e1356150.
7. Li B, Cui Y, Diehn M, Li R. Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer. JAMA Oncol. 2017;3(11):1529–37.
8. Song Q, Shang J, Yang Z, Zhang L, Zhang C, Chen J, et al. Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. J Transl Med. 2019;17(1):70.
9. Chifman J, Pullikuth A, Chou JW, Bedognetti D, Miller LD. Conservation of immune gene signatures in solid tumors and prognostic implications. BMC Cancer. 2016;16(1):911.
10. Gentles AJ, Newman AM, Liu CL, Bratman S V, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015/07/20. 2015;21(8):938–45.
11. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424 LP – 447.
12. Wang Y, Cai Y-Y, Herold T, Nie R-C, Zhang Y, Gale RP, et al. An Immune Risk Score Predicts Survival of Patients with Acute Myeloid Leukemia Receiving Chemotherapy. Clin Cancer Res. 2021;27(1):255-266.
13. Muller AJ, DuHadaway JB, Donover PS, Sutanto-Ward E, Prendergast GC. Inhibition of indoleamine 2,3-dioxygenase, an immunoregulatory target of the cancer suppression gene Bin1, potentiates cancer chemotherapy. Nat Med. 2005;11(3):312–9.
14. Ge Y, Schuster MB, Pundhir S, Rapin N, Bagger FO, Sidiropoulos N, et al. The splicing factor RBM25 controls MYC activity in acute myeloid leukemia. Nat Commun. 2019;10(1):172.
15. Curti A, Aluigi M, Pandolfi S, Ferri E, Isidori A, Salvestrini V, et al. Acute myeloid leukemia cells constitutively express the immunoregulatory enzyme indoleamine 2,3-dioxygenase. Vol. 21, Leukemia. England; 2007. p. 353–5.
16. Frumento G, Rotondo R, Tonetti M, Damonte G, Benatti U, Ferrara GB. Tryptophan-derived catabolites are responsible for inhibition of T and natural killer cell proliferation induced by indoleamine 2,3-dioxygenase. J Exp Med. 2002;196(4):459–68.
17. Fallarino F, Grohmann U, Vacca C, Orabona C, Spreca A, Fioretti MC, et al. T cell apoptosis

by kynurenines. Adv Exp Med Biol. 2003;527:183–90.
18. Curti A, Pandolfi S, Valzasina B, Aluigi M, Isidori A, Ferri E, et al. Modulation of tryptophan catabolism by human leukemic cells results in the conversion of CD25- into CD25+ T regulatory cells. Blood. 2007;109(7):2871–7.
19. Curti A, Trabanelli S, Salvestrini V, Baccarani M, Lemoli RM. The role of indoleamine 2,3-dioxygenase in the induction of immune tolerance: focus on hematology. Blood. 2009;113(11):2394 LP – 2401.
20. Chamuleau MED, van de Loosdrecht AA, Hess CJ, Janssen JJWM, Zevenbergen A, Delwel R, et al. High INDO (indoleamine 2,3-dioxygenase) mRNA level in blasts of acute myeloid leukemic patients predicts poor clinical outcome. Haematologica. 2008;93(12):1894–8.
21. Fukuno K, Hara T, Tsurumi H, Shibata Y, Mabuchi R, Nakamura N, et al. Expression of indoleamine 2,3-dioxygenase in leukemic cells indicates an unfavorable prognosis in acute myeloid leukemia patients with intermediate-risk cytogenetics. Leuk Lymphoma. 2015;56(5):1398–405.
22. Mangaonkar A, Mondal AK, Fulzule S, Pundkar C, Park EJ, Jillella A, et al. A novel immunohistochemical score to predict early mortality in acute myeloid leukemia patients based on indoleamine 2,3 dioxygenase expression. Sci Rep. 2017;7(1):12892.
23. Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, Robertson AG, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059–74.
24. Tobias Herold, Vindi Jurinovic, Aarif M. N. Batcha, Stefanos A. Bamopoulos, Maja Rothenberg-Thurley, Bianka Ksienzyk, et al. A 29-gene and cytogenetic score for the prediction of resistance to induction treatment in acute myeloid leukemia. Haematologica. 2018;103(3 SE-Articles):456–65.
25. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discov. 2012;2(5):401 LP – 404.
26. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(269):pl1.
27. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol. 2020;38(6):675–8.
28. Valk PJM, Verhaak RGW, Beijen MA, Erpelinck CAJ, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med. 2004;350(16):1617–28.
29. Ng SWK, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature. 2016;540:433.
30. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402–8.
31. Kulkarni MM. Digital multiplexed gene expression analysis using the NanoString nCounter system. Curr Protoc Mol Biol. 2011;Chapter 25:Unit25B.10.
32. Cesano A. nCounter((R)) PanCancer Immune Profiling Panel (NanoString Technologies, Inc., Seattle, WA). J Immunother cancer. 2015;3:42.
33. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. 2019;47(W1):W556–60.
34. Li C, Tang Z, Zhang W, Ye Z, Liu F. GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA. Nucleic Acids Res. 2021;
35. Racle J, Gfeller D. EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data. Methods Mol Biol. 2020;2120:233–48.
36. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med.

2016;374(23):2209–21.
37. Lebedev TD, Vagapova ER, Popenko VI, Leonova OG, Spirin P V, Prassolov VS. Two Receptors, Two Isoforms, Two Cancers: Comprehensive Analysis of KIT and TrkA Expression in Neuroblastoma and Acute Myeloid Leukemia. Front Oncol. 2019;9:1046.
38. Worzfeld T, Offermanns S. Semaphorins and plexins as therapeutic targets. Nat Rev Drug Discov. 2014;13:603.
39. König K, Marth L, Roissant J, Granja T, Jennewein C, Devanathan V, et al. The plexin C1 receptor promotes acute inflammation. Eur J Immunol. 44(9):2648–58.
40. Galluzzi L, Buqué A, Kepp O, Zitvogel L, Kroemer G. Immunogenic cell death in cancer and infectious disease. Nat Rev Immunol. 2017;17(2):97–111.
41. Lecciso M, Ocadlikova D, Sangaletti S, Trabanelli S, De Marchi E, Orioli E, et al. ATP Release from Chemotherapy-Treated Dying Leukemia Cells Elicits an Immune Suppressive Effect by Increasing Regulatory T Cells and Tolerogenic Dendritic Cells. Front Immunol. 2017;8:1918.
42. Dufva O, Pölönen P, Brück O, Keränen MAI, Klievink J, Mehtonen J, et al. Immunogenomic Landscape of Hematological Malignancies. Cancer Cell. 2020;38(3):380-399.e13.
43. Vadakekolathu J, Minden MD, Hood T, Church SE, Reeder S, Altmann H, et al. Immune landscapes predict chemotherapy resistance and immunotherapy response in acute myeloid leukemia. Sci Transl Med. 2020;12(546):eaaz0463.
44. Vadakekolathu J, Lai C, Reeder S, Church SE, Hood T, Lourdusamy A, et al. TP53 abnormalities correlate with immune infiltration and associate with response to flotetuzumab immunotherapy in AML. Blood Adv. 2020;4(20):5011–24.
45. Sallman DA, McLemore AF, Aldrich AL, Komrokji RS, McGraw KL, Dhawan A, et al. TP53 mutations in myelodysplastic syndromes and secondary AML confer an immunosuppressive phenotype. Blood. 2020;
46. Walzer T, Galibert L, De Smedt T. Dendritic cell function in mice lacking Plexin C1. Int Immunol. 2005;17(7):943–50.
47. Behrens K, Triviai I, Schwieger M, Tekin N, Alawi M, Spohn M, et al. Runx1 downregulates stem cell and megakaryocytic transcription programs that support niche interactions. Blood. 2016;127(26):3369–81.

Tables

Table 1 *Overview of biological and clinical data referring to patients in the HOVON and TCGA datasets.*

	HOVON	TCGA	GSE106291
Sex			
Female	**294**	**56**	**129**
Male	**315**	**67**	**119**
Fab			
AML-notM	**1**	**0**	**N.A.**
M0	**26**	**12**	**N.A.**
M1	**134**	**38**	**N.A.**
M2	**154**	**28**	**N.A.**
M4	**111**	**29**	**N.A.**
M5	**139**	**12**	**N.A.**
M6	**9**	**2**	**N.A.**
M7	**0**	**1**	**N.A.**
RAEB	**5**	**0**	**N.A.**
RAEB-t	**19**	**0**	**N.A.**
Unknown	**11**	**1**	**N.A.**
BM blasts abundance at diagnosis			
Median blasts (%)	**67 (0-98)**	**74 (30-100)**	**73 (6-100)**
Cytogenetic risk			
Adverse	**123**	**26**	**N.A.**
Favorable	**204**	**17**	**N.A.**
Intermediate	**280**	**78**	**N.A.**
Not evaluable risk	**2**	**2**	**N.A.**
Treatment			
Allogeneic HSCT	**196**	**64**	**N.A.**
Autologous HSCT	**91**	**6**	**N.A.**
Chemotherapy	**320**	**53**	**N.A.**
Unknown	**2**	**0**	**N.A.**

Table 2 Overview of mutational data referring to patients in the HOVON and the TCGA datasets.

	HOVON	TCGA	GSE106291
***NPM1* status**			
Wild-type	**422**	**83**	**N.A.**
Mutated	**183**	**40**	**N.A.**
Unknown	**4**	**0**	**250**
***FLT3* status**			
Wild-type	**441**	**88**	**N.A.**
Mutated	**165**	**35**	**N.A.**
Unknown	**3**	**0**	**250**
***NRAS* status**			
Wild-type	**509**	**112**	**N.A.**
Mutated	**60**	**11**	**N.A.**
Unknown	**40**	**0**	**250**
***KRAS* status**			
Wild-type	**483**	**120**	**N.A.**
Mutated	**5**	**3**	**N.A.**
Unknown	**121**	**0**	**250**
***KIT* status**			
Wild-type	**427**	**107**	**N.A.**
Mutated	**20**	**6**	**N.A.**
Unknown	**162**	**10**	**250**
***ASXL1* status**			
Wild-type	**573**	**121**	**N.A.**
Mutated	**31**	**2**	**N.A.**
Unknown	**5**	**0**	**250**
***IDH1* status**			
Wild-type	**528**	**108**	**N.A.**
Mutated	**42**	**15**	**N.A.**
Unknown	**39**	**0**	**250**
***IDH2* status**			
Wild-type	**510**	**110**	**N.A.**
Mutated	**60**	**13**	**N.A.**
Unknown	**39**	**0**	**250**

Table 3 Overview of clinical data referring to validation cohort patients

Sample type	Sex	Median age	ELN risk class	BM blast abundance (%)	Assay
BM = 24	**Female = 6** **Male = 18**	**55**	**High = 7** **Intermediate = 7** **Low = 6** **N.A. = 4**	**70 (20-90)**	**NanoString**

BM., bone marrow; N.A., not available.

Table 4 Gene list resulting from *IDO1*-focused co-expression analyses of RNA-sequencing AML-TCGA data

Correlated Gene	Cytoband	Spearman's Correlation	p-Value	q-Value
IDO2	**8p11,21**	**0.45**	**P<0.001**	**P<0.001**
CD1C	**1q23,1**	**0.39**	**P<0.001**	**P<0.01**
CD1E	**1q23,1**	**0.38**	**P<0.001**	**P<0.01**
XCR1	**3p21,31**	**0.38**	**P<0.001**	**P<0.01**
PLXNC1	**12q22**	**0.37**	**P<0.001**	**P<0.01**

Table 5 Results of cox regression analysis including *IDO1*, *BIN1* and *PLXNC1* genes in the HOVON dataset.

Genes	Significance	HR [95,0% CI]
BIN1	**P<0.05**	**2.13 [1.17 – 3.90]**
IDO1	**P<0.01**	**2.81 [1.44 – 5.47]**
PLXNC1	**P<0.01**	**2.27 [1.35 – 3.81]**

Figure legends

***Figure 1 IDO1-BIN1* signature predicts AML survival and may be refined by adding *IDO1*-interacting genes**

(A) Correlation between *IDO1* and *BIN1* gene expression values in the HOVON cases (r= -0.41, P<0.0001). **(B)** Kaplan-Meier estimates of OS in the HOVON cases according to the *IDO1-BIN1* score (P<0.01). Patients were split into three different groups according to score quartiles. **(C)** Kaplan-Meier estimates of OS according to *IDO2* expression in the TCGA-AML dataset (*IDO2* median expression value used as cut-off, P<0.05). **(D)** Kaplan-Meier estimates of OS according to *PLXNC1* expression in the TCGA-AML dataset (*PLXNC1* median expression value used as cut-off, P<0.05). **(E)** Correlation between *IDO1* and *IDO2* gene expression values in the HOVON cases (r= 0.27, P<0.0001). **(F)** Correlation between *IDO1* and *PLXNC1* gene expression values in the HOVON cohort of patients (r= -0.25, P<0.0001). **(G)** Kaplan-Meier estimates of OS according to *IDO2* expression in the HOVON dataset (*IDO2* median expression value used as cut-off, P<0.05). **(H)** Kaplan-Meier estimates of OS according to *PLXNC1* expression in the HOVON dataset (*PLXNC1* median expression value used as cut-off, P<0.001).

***Figure 2 IDO1-BIN1-PLXNC1* score predicts AML survival**

(A) *PLXNC1* mRNA expression value was added to *IDO1* and *BIN1* mRNA expression values to generate a new signature. The figure shows Kaplan-Meier estimates of OS according to *IDO1-BIN1-PLXNC1* score quartiles in the HOVON cohort of patients (P<0.0001). **(B)** Kaplan-Meier estimates of OS according to *IDO1-BIN1-PLXNC1* score quartiles in patients of the HOVON cohort who received chemotherapy alone (P<0.001) or **(C)** received chemotherapy and allogeneic transplantation (P<0.05). **(D)** Kaplan-Meier estimates of OS according to *IDO1-BIN1-PLXNC1* score quartiles in the TCGA-AML dataset (P<0.01). **(E)** Kaplan-Meier

estimates of OS according to *IDO1-BIN1-PLXNC1* score quartiles in patients of the TCGA-AML dataset who received chemotherapy alone (P<0.01) or **(F)** who received chemotherapy and allogeneic transplantation (P = not significant).

Figure 3 ***IDO1*** **and** ***PLXNC1*** **are overexpressed and may reflect independent biological processes in AML**

(A) The expression of the top 20 DE genes between *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ samples was higher in TCGA-AML cases compared with blood samples from healthy donors available through the Genotype-Tissue Expression (GTEx) project **(B)** Representation of the top 20 differentially expressed (DE) genes between *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ samples (P value threshold of 0.01; $\log_2$ fold-change threshold of 1.4). **(C)** Enrichment analysis showing the top significant pathways associated with DE genes between *PLXNC1*$^{high/low}$ samples. **(D)** Enrichment analysis showing the top significant pathways associated with DE genes between *IDO1*$^{high/low}$ samples.

Figure 4 **New immune signatures emerge from differently expressed genes between** ***PLXNC1*****$^{high/low}$ and** ***IDO1*****$^{high/low}$ samples**

(A) Kaplan-Meier estimates of OS according to the signature composed by the top 20 differentially expressed (DE) expressed genes between *PLXNC1*$^{high/low}$ samples in the TCGA-AML cases (median used as cut-off, P<0.05). **(B)** Kaplan-Meier estimates of OS according to the signature composed by the top 20 DE genes between *IDO1*$^{high/low}$ samples in the TCGA-AML cases (median used as cut-off, P<0.01). **(C)** Kaplan-Meier estimates of OS according to the signature composed by the top 3 DE genes (*IKBKB*, *FOSL1* and *TLR9*) between *PLXNC1*$^{high/low}$ samples in the TCGA-AML cases (median used as cut-off, P<0.0001). **(D)** Kaplan-Meier estimates of OS according to the signature composed by the top 4 DE genes (*GZMH*, *GNLY*, *IFIT2* and *IFIT3*) between *IDO1*$^{high/low}$ samples in TCGA-AML cases (median used as cut-off, P<0.05). **(E)** Kaplan-Meier estimates of OS according to the signature

composed by the top 3 DE genes from the *PLXNC1*$^{high/low}$ signature (*IKBKB, FOSL1* and *TLR9*) and the top 4 DE genes from *IDO1*$^{high/low}$ signature (*GZMH, GNLY, IFIT2* and *IFIT3*) in the TCGA-AML dataset (median used as cut-off, P<0.05). **(F)** Representation of genetic alterations of the 7 DE genes deriving from the *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ signatures *(IKBKB, FOSL1, TLR9, GZMH, GNLY, IFIT2* and *IFIT3*) in the TCGA-AML dataset. **(G)** Comparison of frequency of mutations between samples with abnormalities (mRNA high/low) *vs* without abnormalities of the 7 DE genes derived from the *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ signatures *(IKBKB, FOSL1, TLR9, GZMH, GNLY, IFIT2* and *IFIT3*). **(H)** Kaplan-Meier estimates of OS according to the signature composed by the integration of the 7 DE genes derived from the *PLXNC1*$^{high/low}$ and *IDO1*$^{high/low}$ signatures *(IKBKB, FOSL1, TLR9, GZMH, GNLY, IFIT2* and *IFIT3*) with the *IDO1, BIN1* and *PLXNC1* genes in the TCGA-AML dataset (median used as cut-off, P<0.0001).

Figure 1

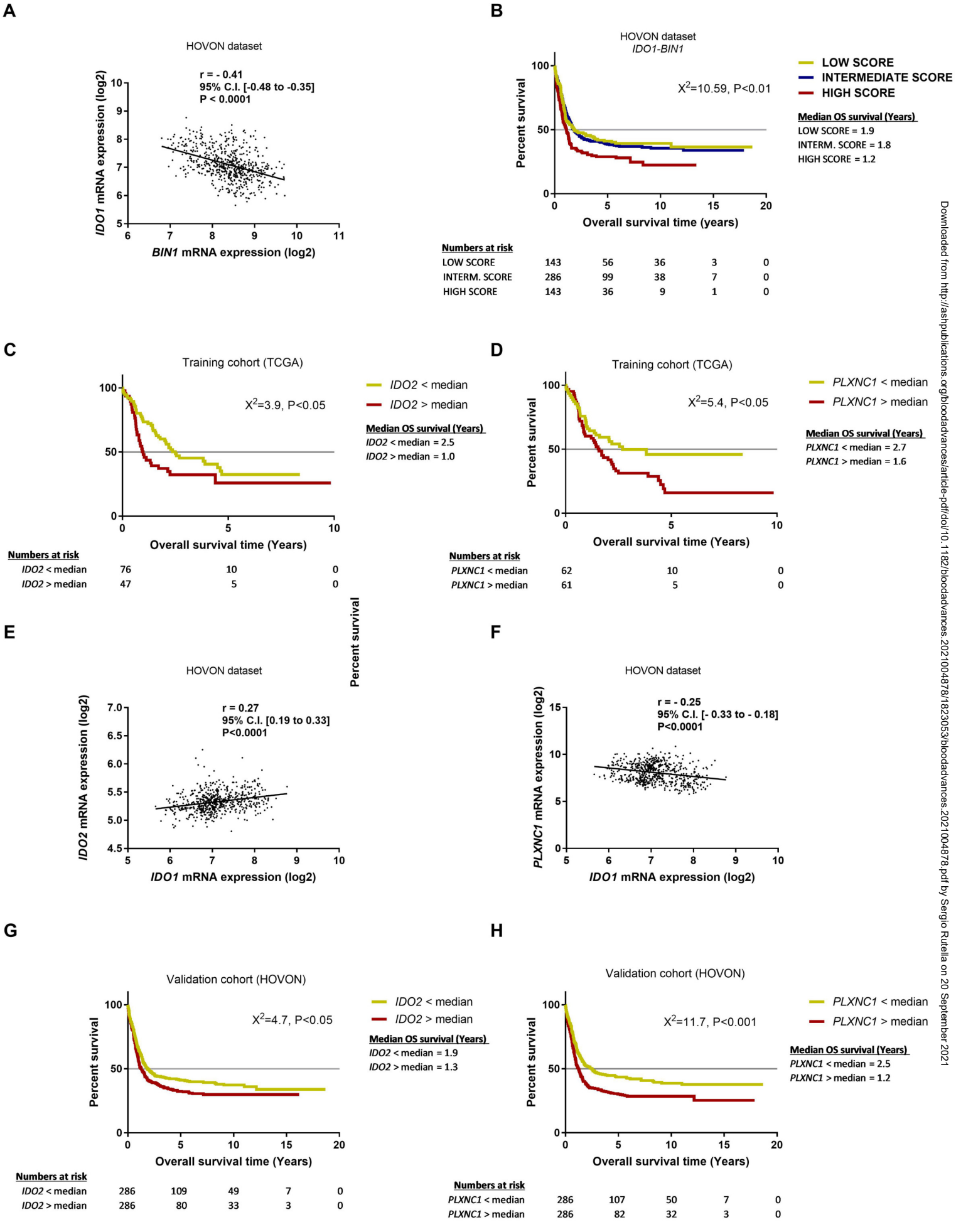
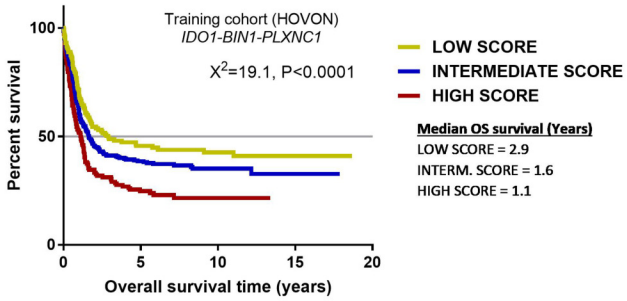


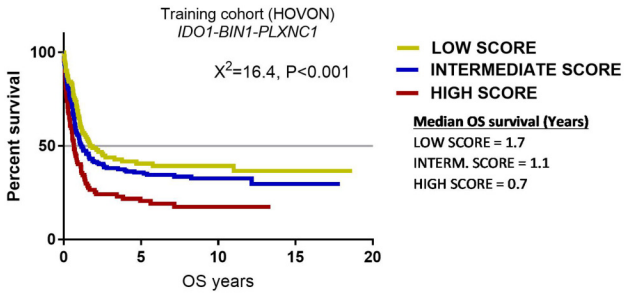
Figure 2

A



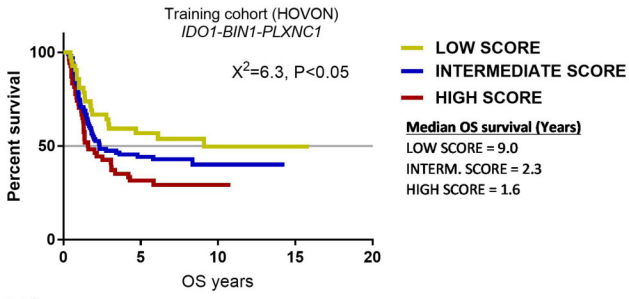
Numbers at risk					
LOW SCORE	143	60	36	5	0
INTERM. SCORE	286	96	38	5	0
HIGH SCORE	143	34	9	1	0

B



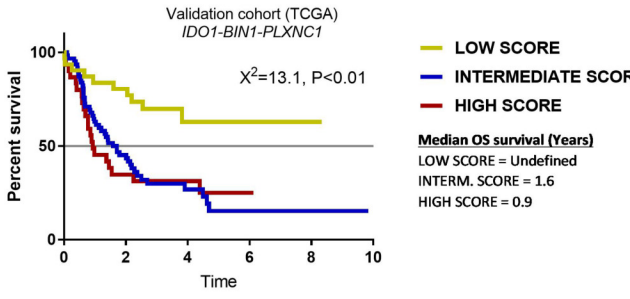
Numbers at risk					
LOW SCORE	102	37	25	4	0
INTERM. SCORE	186	58	28	5	0
HIGH SCORE	89	18	7	1	0

C



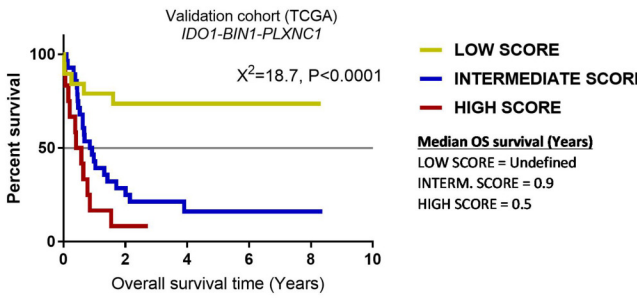
Numbers at risk					
LOW SCORE	42	24	12	2	1
INTERM. SCORE	99	39	11	1	1
HIGH SCORE	54	17	3	1	1

D



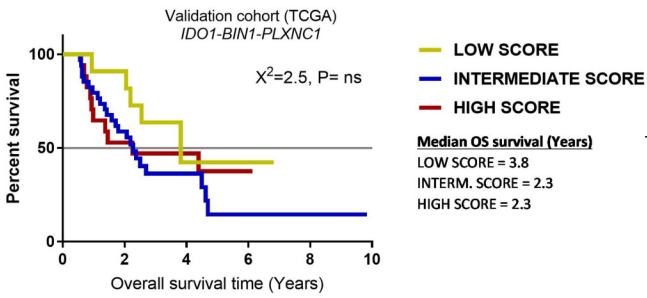
Numbers at risk						
LOW SCORE	31	25	10	5	2	0
INTERM. SCORE	62	28	9	4	3	0
HIGH SCORE	30	11	6	2	1	0

E



Numbers at risk						
LOW SCORE	19	15	8	4	2	0
INTERM. SCORE	28	9	3	2	2	0
HIGH SCORE	12	2	1	1	1	0

F



Numbers at risk						
LOW SCORE	12	11	3	2	1	0
INTERM. SCORE	34	20	7	3	2	0
HIGH SCORE	18	10	6	2	1	0

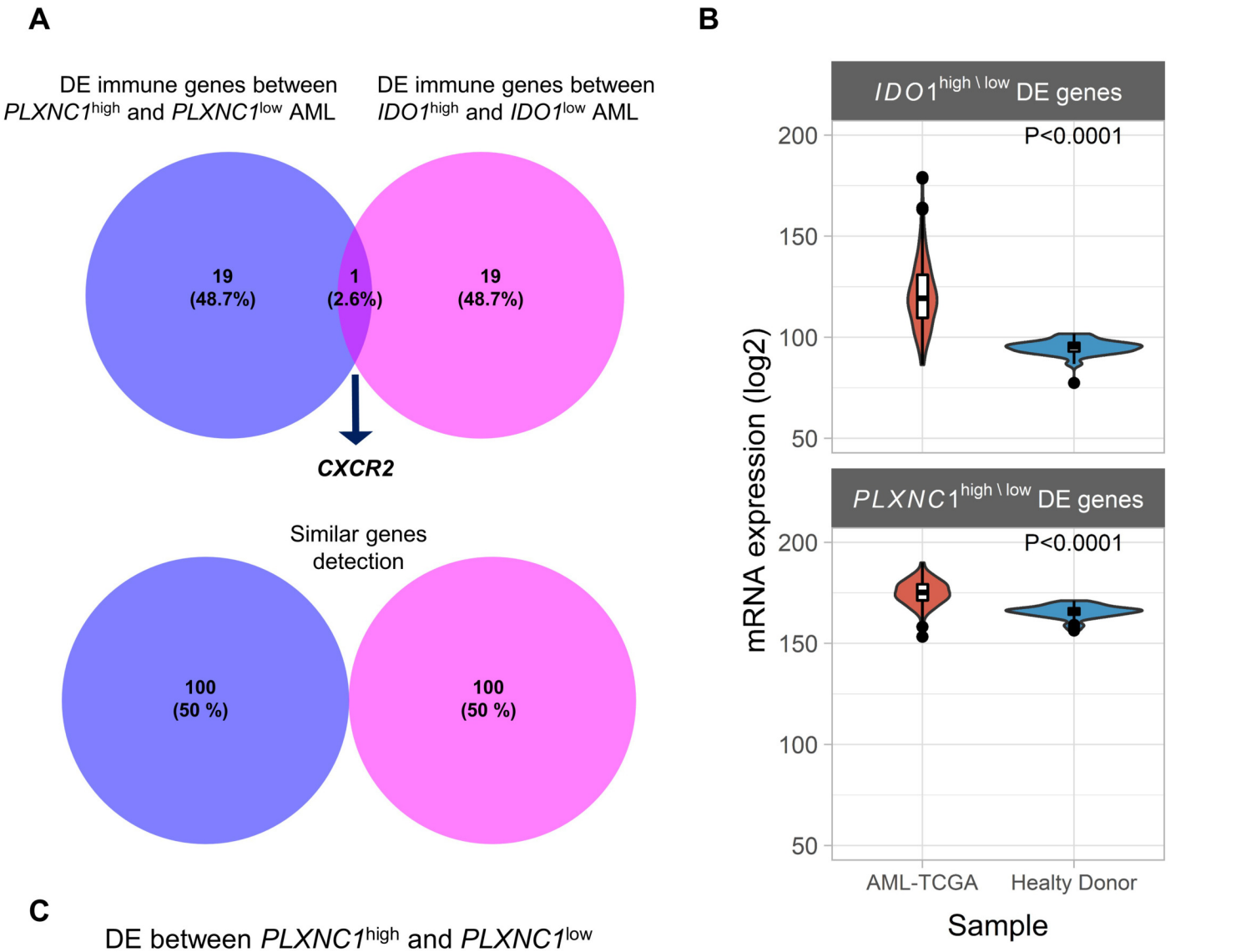


Figure 4