

RESEARCH

Open Access



Early changes in gene expression profiles in AML patients during induction chemotherapy

Ingrid Jakobsen^{1,2†} , Max Sundkvist^{1†}, Niclas Björn¹ , Henrik Gréen^{1,3}  and Kourosh Lotfi^{1,4*} 

Abstract

Background: Elucidation of the genetic mechanisms underlying treatment response to standard induction chemotherapy in AML patients is warranted, in order to aid in risk-adapted treatment decisions as novel treatments are emerging. In this pilot study, we explored the treatment-induced expression patterns in a small cohort of AML patients by analyzing differential gene expression (DGE) over the first 2 days of induction chemotherapy.

Methods: Blood samples were collected from ten AML patients at baseline (before treatment initiation) and during the first 2 days of treatment (Day 1; approximately 24 h, and Day 2; approximately 48 h after treatment initiation, respectively) and RNA was extracted for subsequent RNA sequencing. DGE between time points were assessed by pairwise analysis using the R package edgeR version 3.18.1 in all patients as well as in relation to treatment response (complete remission, CR, vs non-complete remission, nCR). Ingenuity Pathway Analysis (Qiagen) software was used for pathway analysis and visualization.

Results: After initial data quality control, two patients were excluded from further analysis, resulting in a final cohort of eight patients with data from all three timepoints. DGE analysis demonstrated activation of pathways with genes directly or indirectly associated with NF- κ B signaling. Significant activation of the NF- κ B pathway was seen in 50% of the patients 2 days after treatment start, while iNOS pathway effects could be identified already after 1 day. nCR patients displayed activation of pathways associated with cell cycle progression, oncogenesis and anti-apoptotic behavior, including the STAT3 pathway and Salvage pathways of pyrimidine ribonucleotides. Notably, a significant induction of cytidine deaminase, an enzyme responsible for the deamination of Ara-C, could be observed between baseline and Day 2 in the nCR patients but not in patients achieving CR.

Conclusions: In conclusion, we show that time-course analysis of gene expression represents a feasible approach to identify relevant pathways affected by standard induction chemotherapy in AML patients. This poses as a potential method for elucidating new drug targets and biomarkers for categorizing disease aggressiveness and evaluating treatment response. However, more studies on larger cohorts are warranted to elucidate the transcriptional basis for drug response.

Keywords: AML, RNA sequencing, Gene expression, Time course analysis

Background

The standard induction chemotherapy in acute myeloid leukemia (AML) have remained roughly the same over the last decades, with treatment consisting of the nucleoside analog cytarabine (Ara-C) in combination with an anthracycline, most commonly daunorubicin or idarubicin. Additional drugs have been used together with the cornerstone drugs in different clinical trials, and in

[†]Ingrid Jakobsen and Max Sundkvist share first authorship.

*Correspondence: Kourosh.lotfi@liu.se

⁴Department of Hematology, Linköping University Hospital, 581 85 Linköping, Sweden

Full list of author information is available at the end of the article



specific cases such as that of acute promyelocytic leukemia, treatments targeting specific chromosomal rearrangements are available [1]. However, AML patients still suffer from a significant degree of therapy-related mortality and treatment failure. More knowledge is needed of the genetic mechanisms underlying poor treatment response in AML, in order to achieve better risk stratification and individual adaption of treatment regimens as newer drugs are emerging in clinical trials. Genetic variation in genes encoding drug-metabolizing enzymes, transporters, DNA-repair systems or in genes regulating cell cycle progression and apoptosis is likely influencing cellular drug response both in leukemic and non-leukemic cells, although the clinical utility of such markers in relation to AML treatment decisions remain controversial [2]. In leukemia, gene expression profiling of diagnostic samples has been implicated as a tool able to discriminate between different types (e.g ALL and AML) [3] as well as molecular and prognostic subclasses of AML [4–6]. It has also been shown that acute lymphoblastic leukemia cells of different molecular subtypes share common but treatment-specific acute gene expression changes induced by mono- or combination therapy, as measured by gene expression arrays of samples collected before and 1 day after treatment [7]. In addition, a recent study of normal p53-regulated apoptosis, induced by DNA damage, demonstrated that the extent of the apoptotic response was attributable to the rapidity of the transcriptional response rather than to a specific set of target genes [8]. This shows that time series analysis of gene expression could provide new insights to the mechanisms behind drug response. The early drug-induced gene expression changes in AML patients undergoing standard induction chemotherapy has, to our knowledge, not been investigated previously. In this pilot study, we aimed to explore the expression patterns in a small cohort of AML patients and analyzing differential gene expression over the first 2 days of induction chemotherapy. Using RNA sequencing data, the changes in gene expression over time was analyzed overall to identify common patterns, but also compared between patients achieving complete remission and those who did not, to identify potential genes and pathways that might be used to predict or understand the response.

Materials and methods

Patients and sample collection

Ten patients with AML scheduled for induction chemotherapy at the hematology clinic of Linköping University Hospital were included in the study. The study was performed in compliance with the Declaration of Helsinki under the approval by the local ethics committee (Linköping, Sweden), and all patients gave informed

consent. Patient characteristics and treatment details are summarized in Table 1. Blood samples were collected at baseline (before treatment initiation) and during the first 2 days of treatment (Day 1; approximately 24 h after treatment initiation, and Day 2; approximately 48 h after treatment initiation). All patients received Ara-C either as 100 mg/m² as a continuous infusion (Patient 2) and Patient 3) or as 1 g/m² twice daily as 2 h infusions before the Day 1 sample collection; before the Day 2 sample collection all patients except Patient 2 and Patient 3 had also received anthracycline treatment either as idarubicin (10 mg/m² as 1 h infusion; Patient 1 and Patient 4) or daunorubicin (60 mg/m² as 8 h infusion). Patient 4 was also part of a randomized phase 2 trial investigating the addition of cladribine, as a 5 mg/m² 1 h infusion two times, to Ara-C and idarubicin [9]. Patient 5 was initially suspected to have acute promyelocyte leukemia and received treatment with all-trans retinoic acid (ATRA) before the final AML diagnosis and inclusion in this study. No cases of secondary AML were included. In addition to chemotherapy, the patients received standard supportive care with common co-medications including dexamethasone eye drops, allopurinol and antiemetics. RNA was extracted from EDTA blood samples using the Qiagen Blood RNA mini kit (Qiagen), or from blood samples stabilized in PAXgene blood RNA tubes (PreAnalytiX) using the PAXgene blood RNA kit (PreAnalytiX), according to the manufacturers' instructions. The RNA concentration was measured on a NanoDrop (Thermo Scientific) and the RNA integrity number (RIN) was assessed using an Agilent 2100 BioAnalyzer (Agilent) with the Agilent RNA 6000 nano kit (Agilent) according to the manufacturer's instructions.

Sequencing and data preprocessing

All samples with a RIN > 7.5 were sent to the National Genomics Infrastructure (NGI, Stockholm, Sweden) hosted by the Science for Life Laboratory, for RNA sequencing. Libraries were prepared with the TruSeq RNA Library Prep Kit v2 (Illumina) with poly-A selection and paired-end (2x100bp) sequencing was performed on the NovaSeq 6000 Sequencing System (Illumina) using an s1 flow cell. Two samples, Patient 6 Day 2 and Patient 10 Day 2, failed library preparation and due to the low amount of extracted RNA no further attempts could be made. Preprocessing of the raw sequencing data was performed by NGI according to the nf-core/rnaseq pipeline (see <https://github.com/nf-core/rnaseq> for details). Briefly, the pipeline processes the raw fastq data using FastQC for read quality control and TrimGalore! for adapter removal and quality trimming. Alignment of the reads to the human reference genome GRCh38 was made using STAR version 2.7.2b [10]. The read distributions

Table 1 AML patient characteristics

Patient ID	Year of diagnosis	Age	Induction treatment regime	Blast count at baseline	Leukocyte particle count at baseline ($\times 10^9/L$)	Leukocyte particle count Day 2 (48 h) ($\times 10^9/L$)	Cytogenetics	FLT3-ITD	NPM1 mutation	CEBPA mutation	Morphological complete remission after induction ^c
1	2001	67	Idarubicin and cytarabine	80% (BM), 60% (PB)	60,8	58	Aberrant expression of T-cell markers	NA	NA	NA	No
2	2001	81	Daunorubicin and cytarabine ^a	NA	93	75,9	NA	NA	NA	NA	Yes
3	2003	39	Daunorubicin and cytarabine ^a	20% (PB)	51,7	31,6	Normal karyotype	No	No	NA	Yes
4	2001	67	Cladribine, cytarabine and idarubicin	90% (BM)	72	1,9*	Trisomy 8	NA	NA	NA	Yes
5	2003	66	Daunorubicin and low-dose cytarabine ^b	40% (BM)	35,8	11,2	1-40dmin [22] /46XY [3]	Yes	No	NA	No
6	2015	64	Daunorubicin and cytarabine	78% (BM), 57% (PB)	1,6	0,7	Normal karyotype	No	Yes	Heterozygous	Yes
7	2015	74	Daunorubicin and cytarabine	88% (BM)	164	2,6	Not possible due to no mitoses after culturing	NA	NA	NA	No
8	2015	73	Daunorubicin and cytarabine	60% (BM), 77% (PB)	135	37	Normal karyotype	Yes	No	No	No
9	2016	21	Daunorubicin and cytarabine	75% (BM), 44% (PB)	68	5,2	47XY, inv.(16) (p13;q22), +mar [23]	No	No	NA	Yes
10	2016	46	Daunorubicin and cytarabine	32% (BM), 2% (PB)	7,4	3,4	Complex karyo-type	No	No	No	No

^aInitially only cytarabine; no daunorubicin during study sample collection

^bAfter one previous day of all-trans-retinoic acid (ATRA) treatment

^cBone marrow evaluated at day 15–28 after the start of induction chemotherapy

*Leukocyte particle count available also from Day 1 (24 h): $32 \times 10^9/L$

over genomic features were summarized using featureCounts from the Subread package version 1.6.2 [11, 12], and additional quality-control of the results was performed using DupRadar [13] for duplication rate quality control, edgeR [14, 15] for sample similarity, Preseq [16] for library complexity estimation, and RSeQC [17] for read distribution, duplication and gene body coverage. Finally, the quality measures were summarized using MultiQC version 1.7 [18].

Differential gene expression analysis

To explore changes in gene expression in response to treatment, R version 3.6.1 [19] and the R package edgeR version 3.18.1 [14, 15] was used. The data was structured to contain read counts, ensemble gene ID, patient ID and sample time point, for all genes with read counts > 1 count per million (CPM) in ≥ 3 samples. The data was then normalized using the calcNormFactors to scale the raw library sizes for each sample. Differential gene expression (DGE) compared to baseline levels were investigated by pairwise comparisons. In the analysis of DGE between time points, where replicates were lacking for individual patients, the biological coefficient of variation (BCV) was defined as 0.4 as recommended in the edgeR user guide. Negative binomial generalized linear models (GLMs) were fitted for each patient using the function glmFit with the dispersion defined as the BCV squared. Likelihood ratio tests were then performed with the function glmLRT to determine which genes were differentially expressed between the time points. For the overall analysis, the function estimateDisp was used to determine the trended, tagwise and common dispersion for the samples. These were then fitted using glmQLFit and a quasi-likelihood F-test was performed using glmQLFTest. Gene expression changes were also investigated in the clinical response groups, morphological complete remission (CR; final $n=4$) and non-complete remission (nCR; final $n=4$) as assessed by bone marrow sampling between day 15 and day 28 after treatment initiation.

Pathway analysis

The function TopTags in edgeR was used to extract differentially expressed (DE) genes and their \log_2 (fold change), p -value and false discovery rate adjusted p -value. DE genes with p -values ≤ 0.05 were used as input for Ingenuity Pathway Analysis (IPA) (QIAGEN) software that analyses and visualizes DE genes to identify activated or inhibited pathways [20]. Canonical pathways extracted from IPA with a $-\log(p\text{-value}) \geq 1.3$ (corresponding to a p -value < 0.05) were marked as inhibited if the Z-score was ≤ -2 and activated if it was ≥ 2 . All pathway images in the manuscript were generated using IPA (QIAGEN).

Molecules are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). All edges are supported by at least one literature reference or canonical information stored in the IPA knowledge base. The intensity of the node color indicates the degree of up- or down-regulation in red and green, respectively. Nodes are displayed using various shapes that represent the functional class of the gene product. Edges are displayed with various labels that describe the nature of the relationship between the nodes (e.g., P for phosphorylation, T for transcription).

Results

Sequencing data quality assessment

Based on high duplication rate, skewed GC distribution and large portion of multimapping reads, Patient 6 (baseline and 24h samples) and Patient 10 (baseline and 24h samples) were excluded from further analysis. The duplication rates in Patient 6 and 10 were 82,9% and 66,8% (at baseline) and 71,9% and 61,2% (at 24h), respectively. For the remaining eight samples the overall duplication rate was $46,1 \pm 2,48\%$; at baseline it was $47,4 \pm 5,23\%$, at 24h $45,2 \pm 3,84\%$ and at 48h $45,6 \pm 3,42\%$ (mean \pm 95%CI). The skewed GC distribution (see Supplementary Fig. 1) and high levels of overrepresented sequences (see Supplementary Fig. 2) are indicative of ribosomal, bacterial or viral RNA contamination, or biased selection during sample preparation. It could be noted that both these patients had extremely high expression of one single gene – *HBB* – encoding the hemoglobin subunit β -hemoglobin, with feature counts of several million. However, due to the overall quality metrics divergences, the decision to exclude these two patients remained. This resulted in a final cohort of eight patients with data from all three time points available. Summary statistics from MultiQC are shown in Additional file 3.

Common changes in gene expression

Pairwise comparisons with the assessment of overlap between individual patients, and analysis of DGE overall between the time points of the whole cohort demonstrated significant activation of pathways with genes directly or indirectly associated with NF- κ B signaling. This includes processes regulating immune responses, including inflammatory processes, proteasomal processing, lymphogenesis and cell survival as well as proliferation [21]. In addition, iNOS production is associated with NF- κ B signaling, and this could be observed in our data set as DGE between time points in the iNOS pathway (Figs. 1 and 2 illustrate the significant transcriptional changes observed in the NF- κ B pathway and iNOS pathway, respectively). The entire list of affected pathways with Z-scores can be found in Additional file 1, Table S1 (individual patient data) and Additional file 2,

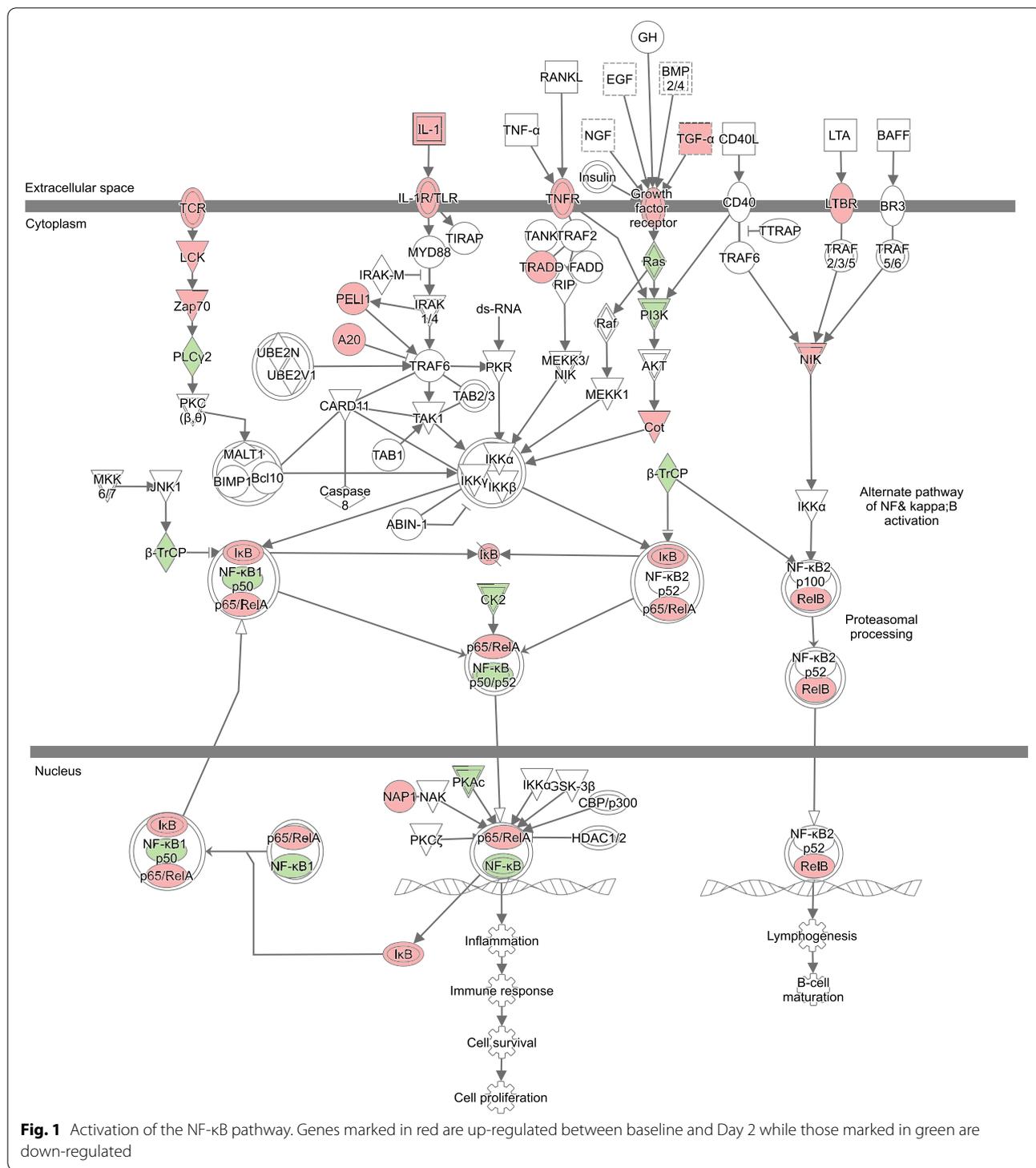
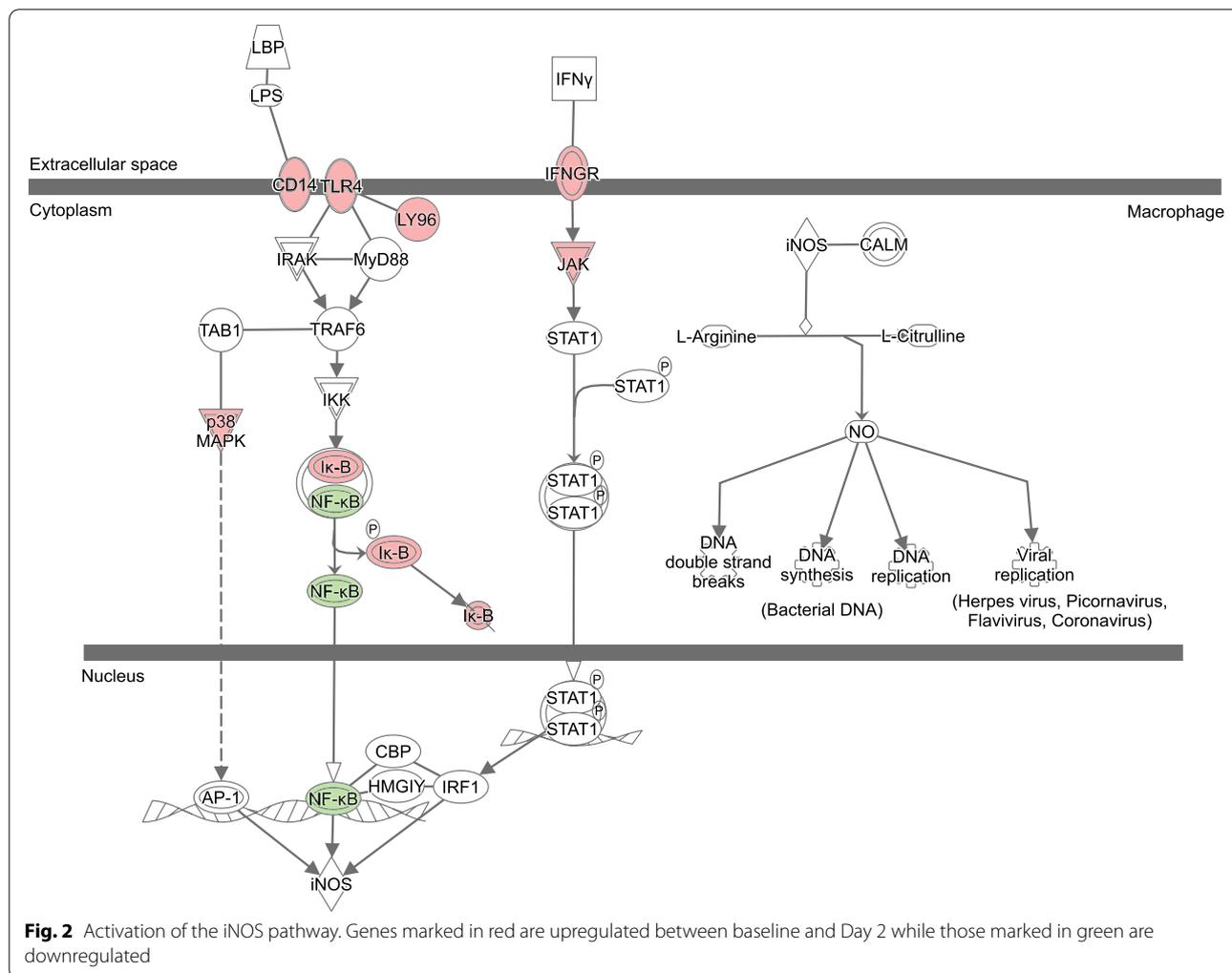


Table S2 (response groups and Overall data). Significant activation of the NF-κB pathway was seen in 50% of the patients (patient 1, 4, 5 and 7) 2 days after treatment start, while iNOS pathway effects could be identified already after 1 day; however this effect was evident only in the grouped analysis.

Pathways affected in relation to treatment outcome

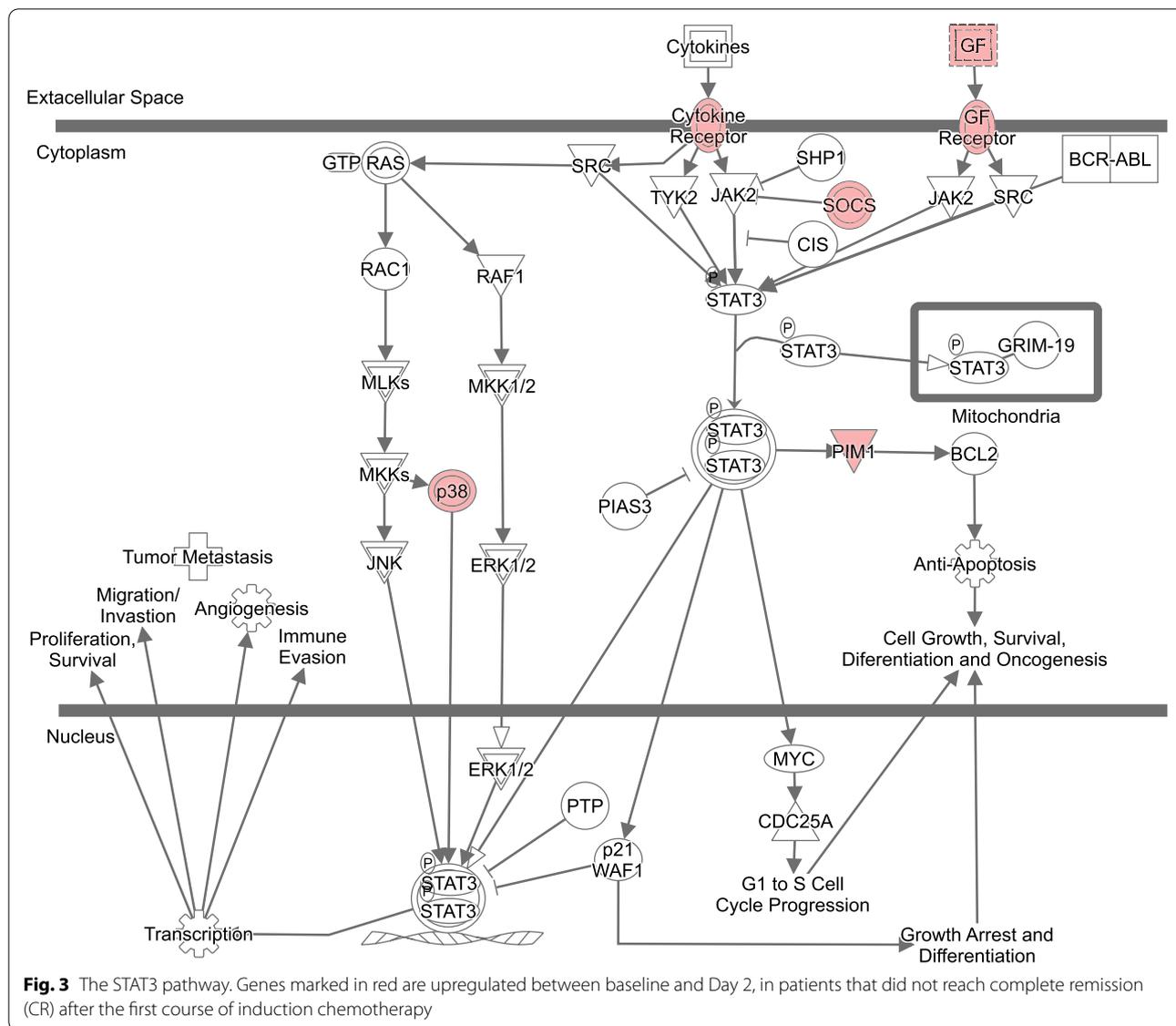
We also investigated which pathways were affected during the treatment in relation to patient outcome; CR or nCR, based on day 15 or day 28 bone marrow evaluation. In patients reaching CR, the pathway



Oxidative phosphorylation was activated already after the first treatment day, while the pathways Purine de novo biosynthesis II, tRNA charging, and Cyclins and cell cycle regulation were inhibited. These pathways were not significant in patients that did not reach CR. In contrast, nCR patients displayed activation of pathways associated with cell cycle progression, oncogenesis and anti-apoptotic behavior, including activation of the STAT3 pathway (Fig. 3) and Salvage pathways of pyrimidine ribonucleotides (Fig. 4). Notably, a significant induction of cytidine deaminase, an enzyme responsible for the deamination of Ara-C, could be observed between baseline and Day 2 in the nCR patients but not in the patients achieving CR. All pathways affected with their corresponding z-scores in the two response groups are listed in Additional file 2, Table S2, with selected relevant pathways summarized below in Tables 2 and 3.

Discussion

In this pilot study, we analyzed gene expression changes over the first 2 days of standard induction chemotherapy in eight AML patients. We could show induction of relevant pathways related to immune response, cell proliferation and apoptosis, but also including genes of importance for drug pharmacokinetics. Some of these changes were detectable already after the first day of treatment. Overall, the NF-κB pathway was significantly activated after 2 days of treatment compared to baseline. This was also reflected by expression alterations in the iNOS pathway, which is directly linked to NF-κB signaling. Significant activation of the NF-κB pathway was seen in four patients 2 days after treatment start, while iNOS pathway effects could be identified already after 1 day; however this effect was evident only in the grouped analysis. This raises the question why we see this time delay between the activation of



these two parallel pathways. The reason cannot easily be determined based on our present data. One possible explanation could be intrinsic relationships in the release and interactions between different cytokines and their respective downstream effects, such as IFN γ preceding the release of IL-1 and TGF- α . Due to the heterogeneous and limited number of samples in our study, further in-depth, high resolution studies of a larger number of samples would be needed in order to confirm and explore these specific time relationships.

The canonical NF- κ B pathway is normally activated by pro-inflammatory cytokines, Toll-like receptor binding or T-cell receptor activation, subsequently leading to proteasomal degradation of the NF- κ B inhibitor I κ B.

The release from I κ B allows NF- κ B to relocate from the cytoplasm to the nucleus where it regulates a wide variety of genes. In oncogenesis, this includes the transcription of genes promoting cell proliferation and survival as well as angiogenesis, metastasis, and suppression of the tumorsuppressors p53 and FOXO3a [21]. In AML, a growing body of evidence has demonstrated dysregulated NF- κ B [22] and opened up the possibility of targeting this pathway; here, proteasome inhibitors are constituting a promising group of drugs with broad application to many types of cancers. In addition, in vitro experiments have shown that treatment with cytarabine induces NF- κ B expression in the HL-60 cell line [23]. Our results in contrast show relative up-regulation of the I κ B gene

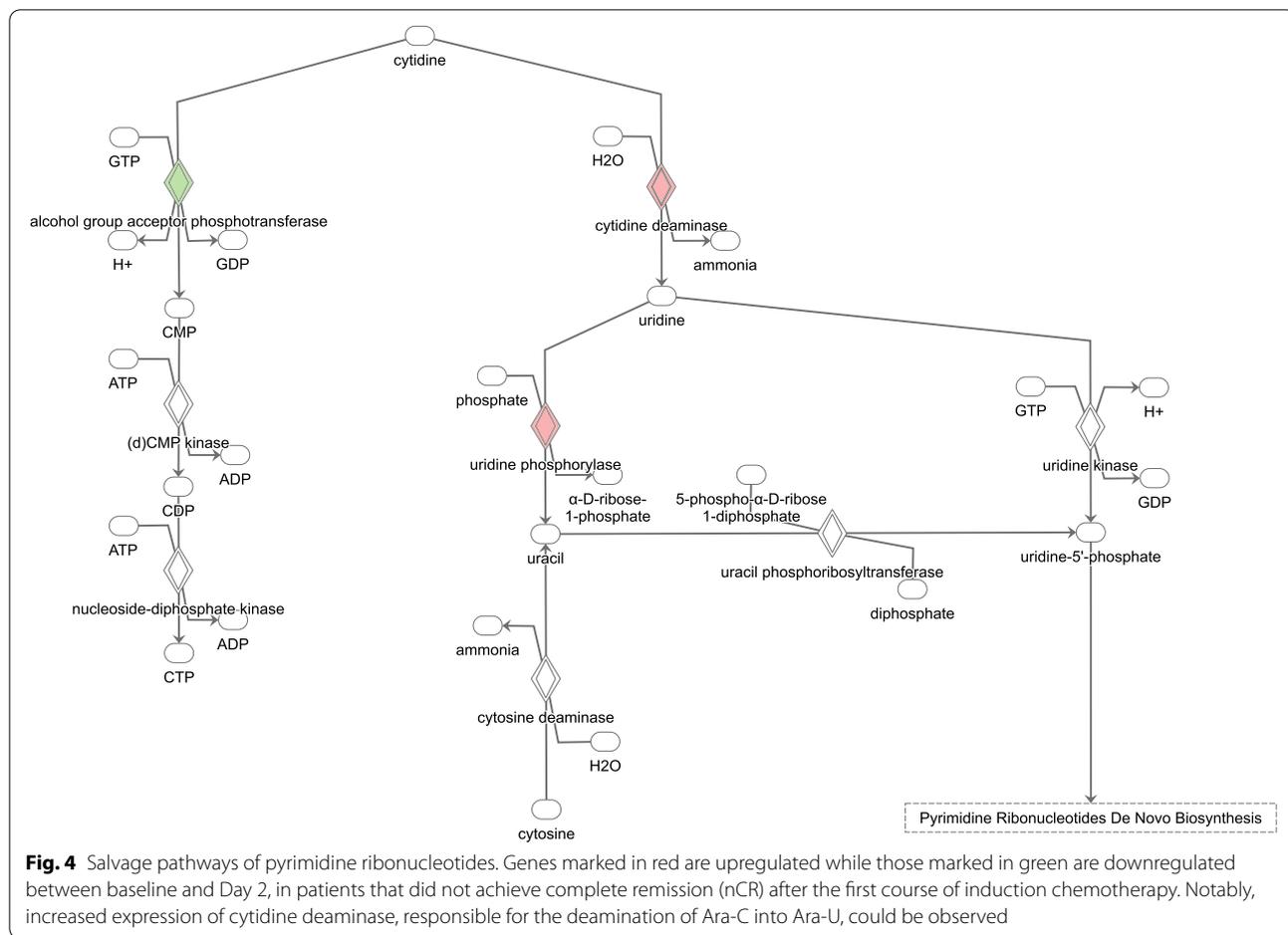


Table 2 Pathways significantly affected in the CR patients

Pathway	24 h vs baseline	48 h vs 24 h	48 h vs baseline
Purine Nucleotides De Novo Biosynthesis II	–	Inhibited	Inhibited
tRNA Charging	–	Inhibited	Inhibited
Oxidative Phosphorylation	Activated	–	Activated
Cyclins and Cell Cycle Regulation	–	–	Inhibited

---: No significant difference

Table 3 Pathways significantly affected in the nCR patients

Pathway	24 h vs baseline	48 h vs 24 h	48 h vs baseline
STAT3 Pathway	Activated	Activated	Activated
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	Activated	Activated	Activated
IL-6 Signaling	Activated	Activated	Activated
Toll-like Receptor Signaling	Activated	Activated	Activated
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	Activated	–	Activated
Salvage Pathways of Pyrimidine Ribonucleotides	–	Activated	Activated

---: No significant difference

and correspondingly down-regulation of the NF- κ B gene at 24 h (as part of the iNOS pathway; Fig. 2), and at 48 h (as part of the NF- κ B pathway; Fig. 1) compared to baseline. This is indicative of an overall positive effect of the chemotherapy treatment, inhibiting cell proliferation and survival; however, we cannot determine whether these transcriptional changes are present in all cell populations (malignant as well as healthy) or to which extent these relative changes are reflecting a decreasing proportion of malignant cells (with assumed dysregulated NF- κ B signaling). Interestingly, among the four patients with activation of the NF κ B pathway as described above, three were non-responders after first induction (patient 1, 5 and 7) while patient 4 reached CR, as assessed by bone marrow sampling between day 15 and day 28 after treatment initiation. Patient 5 and 7 later reached morphological remission after additional treatment cycles; however patient 5 shortly thereafter relapsed with soft-tissue engagement of their AML, indicating residual disease after induction.

However, we should remember that this is a study on only eight patients.

When analyzing the patients based on their first response evaluation (bone marrow analysis at day 15 or 28), we were able to identify pathways that were significantly affected in one group but not the other. These differences could potentially be indicative of differences in tumor cell sensitivity to drug treatment. As such, the pathways Purine de novo biosynthesis II, tRNA charging, and Cyclins and cell cycle regulation were inhibited in patients reaching CR already after the first course of induction therapy, which might demonstrate reduced tumor cell proliferation in this group. These pathways were not significantly affected in the group of patients that did not achieve complete remission; however, two patients did so after additional treatment. In contrast to the CR group, nCR patients displayed significant activation of pathways associated with cell cycle progression, anti-apoptotic signaling and malignant transformation. Persistent activation of Stat3, as well as other Stat proteins, has been implicated as an oncogene promoting malignant transformation in various experimental settings [24, 25]. Stat3 target genes include Bcl-X and Bcl-2, both encoding anti-apoptotic proteins, and cooperation with c-Jun constitute yet another mechanism of interference with tumor cell apoptosis. In addition, the targets Cyclin D1 and Myc promote proliferation; in solid tumors an influence of Stat3 on VEGF expression could also further promote tumor growth and metastasis/tissue invasion through increased angiogenesis, a process of lower relevance for blood cancers. Many other target genes are likely involved in the oncogenic potential of Stat3, and activation of this protein does not necessarily lead to malignant transformation. However, suppression of Stat signaling has demonstrated

tumor growth arrest, induction of apoptosis in cancer cells and inhibition of malignant transformation, indicating that it has potential as a future drug target [24, 26–28]. Our results suggest that patients displaying rapid activation of the Stat3 pathway could be harboring more resistant or highly proliferative leukemic clones. Interestingly, another pathway identified to be significantly activated in patients with poor initial treatment response was Salvage pathways of pyrimidine ribonucleotides. This pathway represents an important system for the maintenance of ribonucleotide pools in all proliferating cells, but also have implications for drug response [29, 30]. Notably, the induction of the enzyme cytidine deaminase (CDA) could result in a more treatment-resistant phenotype in these patients, as the pyrimidine analog Ara-C is subject to degradation into Ara-U by CDA. CDA induction as a mechanism of resistance to Ara-C has been demonstrated *in vitro* [31, 32], and our results support also an *in vivo* relevance.

Some limitations of this study need to be addressed and emphasize that the results needs to be kept in perspective. Our study population is small and the influence of interindividual variation may occlude common changes in gene expression related to the treatment, It should also be noted that these patient's are heterogeneous in their molecular diagnostics and treatment. We are partly compensating for this by comparing the different time point data to that from the baseline sample for each patient. However, the bulk RNA sequencing data represent the gene expression in both healthy and cancerous blood cells, and the leukemic blast proportion in blood samples from AML patients varies. Throughout the treatment, blood leukocyte counts are rapidly decreasing in many patients, and in our approach, we are not discriminating between cancer cells and healthy cells being eradicated. Changes in the relative abundance of different cell populations are likely to be reflected in the expression data, and unfortunately only leukocyte particle counts and not detailed peripheral blast counts for each sampling time point was documented in the patient journals. As an example, we cannot determine whether a decreased gene expression at day two compared to baseline means that the expression has been inhibited, or if it means that a specific cell population responsible for the baseline expression was eradicated, leaving only cell populations with another pattern. In addition, our patient samples were collected over a long period, and variations in treatment protocols and sampling procedures may have influenced the result. Initial principal component analysis indicated that the data tended to partly cluster depending on inclusion period (defined as before or after the introduction of national treatment guidelines, data not shown) although there was an overlap between these clusters.

In conclusion, we show that time-course analysis of gene expression represents a feasible approach to identify relevant pathways affected by standard induction chemotherapy in AML patients. This poses as a potential method for elucidating new drug targets and biomarkers for categorizing disease aggressiveness and evaluating treatment response. However, more studies on larger cohorts are warranted to elucidate the transcriptional basis for drug response, also assessing differences that are likely attributable to baseline characteristics such as cytogenetics and morphology. As single-cell RNA sequencing methodologies are developed and becoming more common, this would likely represent an option to further enhance the resolution also in time course studies like this one, although single-cell approaches are associated with other practical challenges in both sample management and analysis compared to conventional bulk RNA sequencing. Preferably, future studies would include more time points as well as follow-up analysis in the event of relapse.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-08960-4>.

Additional file 1.
Additional file 2.
Additional file 3.
Additional file 4.
Additional file 5.
Additional file 6.
Additional file 7.

Acknowledgements

We gratefully acknowledge the invaluable help from lab technicians Lucia Pelle (Ph.D), Kerstin Willander (Ph.D) and Naomi Yamada (research engineer) in the sample collection process. The authors also wish to acknowledge NBS (National Bioinformatics Infrastructure Sweden). The authors acknowledge sequencing support from the National Genomics Infrastructure in Genomics Production Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council. Parts of the sequencing data computations and handling were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) partially funded by the Swedish Research Council through grant agreement no. 2018-05973.

Authors' contributions

HG: conception and study design, sample preparation and analysis; JJ: sample preparation, clinical data compilation, results interpretation, and manuscript preparation; MS and NB: bioinformatics and statistical analysis, manuscript preparation; KL: Patient recruitment, sample collection and clinical data collection. All authors critically reviewed the manuscript and approved the final version for publication.

Funding

Open access funding provided by Linköping University. This study was funded by grants from The Swedish Cancer Society (HG), Lions research fund for non-communicable diseases (JJ), and Clas Groschinsky Memorial Fund (HG).

Availability of data and materials

The sequencing datasets generated and/or analysed during the current study are not publicly available due to restraints of the ethical permit, but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was performed in compliance with the Declaration of Helsinki under the approval by the local ethics committee (Linköping University registration numbers 98319, 02–459 and 2017–397–31), and all patients gave informed consent to study participation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests. The research grant providers had no role in the study design, collection, analyses or interpretation of data, writing of the manuscript, or in the decision to publish the results.

Author details

¹Division of Clinical Chemistry and Pharmacology, Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Linköping, Sweden. ²Department of Laboratory Medicine, School of Medical Sciences, Faculty of Medicine and Health, Örebro University, Örebro, Sweden. ³Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden. ⁴Department of Hematology, Linköping University Hospital, 581 85 Linköping, Sweden.

Received: 23 May 2022 Accepted: 20 October 2022

Published online: 14 November 2022

References

- Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424–47.
- Megias-Vericat JE, Montesinos P, Herrero MJ, Boso V, Martinez-Cuadron D, Poveda JL, et al. Pharmacogenomics and the treatment of acute myeloid leukemia. *Pharmacogenomics*. 2016;17(11):1245–72.
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*. 1999;286(5439):531–7.
- Bullinger L, Dohner K, Bair E, Frohling S, Schlenk RF, Tibshirani R, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med*. 2004;350(16):1605–16.
- Lindvall C, Furge K, Bjorkholm M, Guo X, Haab B, Blennow E, et al. Combined genetic and transcriptional profiling of acute myeloid leukemia with normal and complex karyotypes. *Haematologica*. 2004;89(9):1072–81.
- Schoch C, Kohlmann A, Schnittger S, Brors B, Dugas M, Mergenthaler S, et al. Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proc Natl Acad Sci U S A*. 2002;99(15):10008–13.
- Cheok MH, Yang W, Pui CH, Downing JR, Cheng C, Naeve CW, et al. Treatment-specific changes in gene expression discriminate in vivo drug response in human leukemia cells. *Nat Genet*. 2003;34(1):85–90.
- Bainer RO, Trendowski MR, Cheng C, Pei D, Yang W, Paugh SW, et al. A p53-regulated apoptotic gene signature predicts treatment response and outcome in pediatric acute lymphoblastic leukemia. *Cancer Manag Res*. 2017;9:397–410.
- Juliusson G, Hoglund M, Karlsson K, Lofgren C, Mollgard L, Paul C, et al. Increased remissions from one course for intermediate-dose cytosine arabinoside and idarubicin in elderly acute myeloid leukaemia when combined with cladribine. A randomized population-based phase II study. *Br J Haematol*. 2003;123(5):810–8.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.

11. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923–30.
12. Liao Y, Smyth GK, Shi W. The subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 2013;41(10):e108.
13. Sayols S, Scherzinger D, Klein H. dupRadar: a bioconductor package for the assessment of PCR artifacts in RNA-Seq data. *BMC Bioinformatics*. 2016;17(1):428.
14. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40.
15. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multi-factor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. 2012;40(10):4288–97.
16. Daley T, Smith AD. Predicting the molecular complexity of sequencing libraries. *Nat Methods*. 2013;10(4):325–7.
17. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics*. 2012;28(16):2184–5.
18. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047–8.
19. RCoreTeam. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021.
20. Kramer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*. 2014;30(4):523–30.
21. Wong ET, Tergaonkar V. Roles of NF-kappaB in health and disease: mechanisms and therapeutic potential. *Clin Sci (Lond)*. 2009;116(6):451–65.
22. Zhou J, Ching YQ, Chng WJ. Aberrant nuclear factor-kappa B activity in acute myeloid leukemia: from molecular pathogenesis to therapeutic target. *Oncotarget*. 2015;6(8):5490–500.
23. Strum JC, Small GW, Pauig SB, Daniel LW. 1-beta-D-Arabinofuranosylcytosine stimulates ceramide and diglyceride formation in HL-60 cells. *J Biol Chem*. 1994;269(22):15493–7.
24. Bromberg J. Stat proteins and oncogenesis. *J Clin Invest*. 2002;109(9):1139–42.
25. Levy DE, Lee CK. What does Stat3 do? *J Clin Invest*. 2002;109(9):1143–8.
26. Kube D, Holtick U, Vockerodt M, Ahmadi T, Haier B, Behrmann I, et al. STAT3 is constitutively activated in Hodgkin cell lines. *Blood*. 2001;98(3):762–70.
27. Epling-Burnette PK, Liu JH, Catlett-Falcone R, Turkson J, Oshiro M, Kothapalli R, et al. Inhibition of STAT3 signaling leads to apoptosis of leukemic large granular lymphocytes and decreased Mcl-1 expression. *J Clin Invest*. 2001;107(3):351–62.
28. Turkson J, Ryan D, Kim JS, Zhang Y, Chen Z, Haura E, et al. Phosphotyrosyl peptides block Stat3-mediated DNA binding activity, gene regulation, and cell transformation. *J Biol Chem*. 2001;276(48):45443–55.
29. Frances A, Cordelier P. The emerging role of Cytidine Deaminase in human diseases: a new opportunity for therapy? *Mol Ther*. 2020;28(2):357–66.
30. Nygaard P. On the role of cytidine deaminase in cellular metabolism. *Adv Exp Med Biol*. 1986;195 Pt B:415–20.
31. Neff T, Blau CA. Forced expression of cytidine deaminase confers resistance to cytosine arabinoside and gemcitabine. *Exp Hematol*. 1996;24(11):1340–6.
32. Schroder JK, Kirch C, Flasshove M, Kalweit H, Seidelmann M, Hilger R, et al. Constitutive overexpression of the cytidine deaminase gene confers resistance to cytosine arabinoside in vitro. *Leukemia*. 1996;10(12):1919–24.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

