Validation of Transcription Factors Underlying the Control of CD8+ T Cell Differentiation Along the Paths of Exhaustion and Senescence in Acute Myeloid Leukemia Using an Artificial Neural Network Model

RESEARCH ARTICLE
HEMATOLOGY

Diseases & Research

Validation of Transcription Factors Underlying the Control of $CD8^+$ T Cell Differentiation Along the Paths of Exhaustion and Senescence in Acute Myeloid Leukemia Using an Artificial Neural Network Model

Yoshiaki Nishiyama[1,*], Akitaka Higashi[2], Fumikazu Miwakeichi[3], Nobuaki Nishiyama[1,2,#]

[1] Division of Mathematical and Physical Sciences, Graduate School of Natural Science and Technology, Kanazawa University, Ishikawa 9201192, Japan
[2] Emerging Media Initiative, Kanazawa University, Kanazawa 9201192, Ishikawa, Japan
[3] Department of Statistical Modeling, Institute of Statistical Mathematics, Tokyo 1908562, Japan

Background: The impact of $CD8^+$ T cell differentiation along the paths of exhaustion and senescence states on clinical outcomes of patients with Acute myeloid leukemia (AML) after induction therapy has been characterized by immunophenotyping, gene expression, and functional studies. The results of the preceding studies show that the relative frequencies of naïve T cells, early memory T cells, terminally differentiated effector T cells (Term), exhausted cells and senescent-like cells are correlated with CR and non-response after induction therapy.
Methods: We assumed that distinct profiles of T cell differentiation states before and after induction therapy for patients with AML, that is, the cellular states at initial diagnosis and complete remission are based on distinct dynamics of transcription factor network motifs. We selected seven transcription factors and PD-1 as key components of network motif controlling $CD8^+$ T differentiation by literature review. Selected seven transcription factors (BATF, IRF4, NFATC1, TCF-1, EOMES, T-bet, TOX) and PD-1 were validated by examining whether the expressions of these encoding genes are effective for the classification between newly diagnosed patients and CR patients with AML by use of an artificial neural network model.
Results: The expressions of the corresponding eight genes as input features were effective for classification between newly diagnosed patients and CR patients after induction therapy with high accuracy.
Conclusion: The high accuracy of classification between newly diagnosed and complete remission (CR) patients after induction therapy suggests that these two disease states exhibit distinct expression patterns of the selected eight genes. The corresponding transcription factors may play a role in regulating $CD8^+$ T cell differentiation in response to chronic antigen stimulation and its interruption by treatment in AML.
Keywords: Acute myeloid leukemia; Transcription factors; Exhaustion; Senescence; Artificial neural network model.

Introduction

Although a majority of diagnosed patients with acute myeloid leukemia (AML) attain complete remission (CR) with first-line high dose chemotherapy following consolidation chemotherapy, the relapse rate is still high. The application of immunotherapy such as immune checkpoint blockade (ICB) agents[1], bispecific antibodies (bsAb)[2] and genetically engineered T cells expressing chimeric antigen receptors (CAR)[3] has also shown limited effects on achieving durable CR in AML.

There is accumulating evidence that the development of cytotoxic T cell dysfunction, especially exhaustion and senescence contribute to the disease progression through persistence of antigen stimulation by AML blast[4]. T cell differentiation and the resulting extent of intratumoral T cell heterogeneity influenced by mutational burdens of AML blasts have been suggested to be relevant in eliciting clinical outcomes, that is, sustained responses or non-responses after chemotherapy and immunotherapy. T cell activity after antigen recognition is regulated by co-stimulatory or co-inhibitory ligand/receptor interactions known as immune checkpoints. While the expression of multiple co-inhibitory molecules such as the programmed cell death 1 (PD-1), T cell immunoglobulin and mucin protein 3 (Tim-3) and 2B4 on T cells is crucial for maintaining self-tolerance and minimizing inflammatory damage under physiological conditions, sustained overexpression of them is frequently observed as a response to chronic antigen exposure in cancer including AML and recognized as markers for exhausted T cells[4]. Exhausted T cells display an impaired ability to secrete effector cytokines such as interleukin-2 (IL-2), tumour necrosis factor (TNF) and interferon-γ (IFNγ) and to exert cytotoxic functions. While senescent T cells undergo cell cycle arrest and they are still metabolically active and can produce cytokines including TNF-a and IFN-γ, the production of cytotoxic factors such as GZMB is repressed[5,6]. Mazziotta et al. reported the relationship between T-cell senescence and clinical outcomes of patients with AML after induction therapy[7]. Bone marrow (BM) $CD8^+$ T cells were sorted from diagnosed patients

* Current Affiliation: Clinical Research Support Center Kyushu, Fukuoka 8128582, Japan.
# Correspondence: Nobuaki Nishiyama. E-mail: nnishiya@staff.kanazawa-u.ac.jp. Address: Division of Mathematical and Physical Sciences, Graduate School of Natural Science and Technology, Kanazawa University, Ishikawa 9201192, Japan.
Received 20 March 2025; Revised 9 June 2025; Accepted 24 July 2025; Available Online 6 August 2025
DOI: 10.54457/DR.202503003

and induction therapy treated patients. They found that the frequencies of naïve T cells and early memory T cells (CXCR3^{+} stem cell memory T cells (Tscm)) in responders to the treatment were significantly increased compared with those in non-responders and conversely the frequencies of terminally differentiated effector T cells (Term) and senescent-like cells (SenL) with significantly expressed senescence-associated genes like killer cell lectin like receptor F1 (KLRF1) increased compared with those in responders[7]. In addition, the high ratio of naïve T cells and early memory T cells (Tscm) / Term cells and SenL cells was found in patients with mutated NPM1 of AML blast and low ratio in patients with mutated TP53, ASXL1, and RUNX1, suggesting an association between gene mutations of AML blasts and T-cell senescence which is in line with the finding of AML cells directly inducing their senescence[8]. The impact of CD8^{+} T cell differentiation along the paths of exhaustion and senescence states on clinical outcomes after induction therapy has been also characterized by studies examining gene expression profiles[8,9]. Knaus et al. found that peripheral blood (PB) CD8^{+} T cells from initial diagnosed AML patients before induction chemotherapy consist of two populations of the senescent cells with high production of proinflammatory cytokine and low PD-1 expression and the exhausted cells with low production of proinflammatory cytokine and high PD-1 expression[8]. Next, they found that PB CD8^{+} T cells of patients who achieved CR after induction therapy display similar profiles of the immune-related gene expression with those from healthy controls and PB CD8^{+} T cells of non-responders to induction therapy did not have similar profiles, suggesting that transient reduction of AML cell frequency by cytotoxic effects of induction chemotherapy could induce the change of gene expression profile underlying the control of CD8^{+} T cell differentiation toward that of healthy controls and explaining the finding of AML cells directly inducing their senescence[8]. In addition, they found that the frequencies of CD8^{+} Term cells of newly diagnosed AML patients have lower levels when CR was achieved and the frequencies of CD8^{+} Term cells of non-responders maintain higher levels, consistent with the result of Mazziotta et al[7]. In the setting of the treatment with bispecific antibodies (bsAb) for AML patients, Kazerani et al. assessed BM CD8^{+} cells based on the immune-related gene expression profiles at initial diagnosis, CR and first relapse after the treatment and also found that the dysfunction of BM CD8^{+} cells in AML consists of senescense and exhaustion, whereas BM CD8^{+} cells senescence is prominent from newly diagnosed patients and BM CD8^{+} cells from non-responders after treatment show gene expression profiles of exhaustion[9]. In addition, they found that a higher frequency of Term cells was observed at initial diagnosis compared to CR and conversely a higher frequency of naïve T cells was observed at CR than at initial diagnosis, consistent with the result of Mazziotta et al[7]. The results of these previous studies[7-9] regarding CD8^{+} T cell frequencies at initial diagnosis, complete remission, and relapse are summarized in Table 1.

Table 1. Summary of CD8^{+} T cell frequency at initial diagnosis, complete remission (CR) and relapse after induction therapy, based on previously published studies[7–9].

Patients	Cell Frequency				
Response Group	Naïve T Cells	Early Memory T Cells	Exhausted T Cells	Terminally Differentiated Effector T Cells	Senescence-like T Cells
Responders (CR)	high[7]	high[7]		low[7]	low[7]
After	high[9]			low[8]	
Induction Therapy				low[9]	
Non-Responders					
After	low[7]	low[7]	high[9]	high[7]	high[7]
Induction Therapy				high[8]	
Initial Diagnosis					
Before	low[9]		high[8]	high[9]	high[8]
Induction Therapy					high[9]

Taken together, deciphering network motifs underlying the control of CD8^{+} T cell differentiation along the paths of exhaustion and senescence is essential for confining cellular state of AML to durable complete remission after induction therapy. In this study, we focused on gene expression of transcription factors relating to exhaustion and senescence[10,11] which could be components of network motifs underlying the control of CD8^{+} T cell differentiation.

Methods

We assumed that distinct profiles of T cell differentiation states before and after induction therapy for patients with AML, that is, the cellular states at initial diagnosis and complete remission are based on distinct dynamics of transcription factor network motifs.

Given that the studies described in Introduction section suggest that the relative frequencies of naïve T cells, early memory T cells, terminally differentiated effector T cells (Term), exhausted cells and senescent-like cells are correlated with CR and non-response of patients with AML after induction therapy, genes involved in the bifurcation point between effector T cells, memory T cells and exhausted T cells could be potential targets. In settings of persistent antigen exposure, T cell factor 1 (TCF-1)$^{+}$ CD8^{+} T cells are located at the bifurcation point, where precursor exhausted T cells (Tpex) can differentiate into stem-like early memory T (Tscm) cells or CX3CR1+ cytotoxic effector T cells which can further progress toward terminally differentiated effector cells (Term) or senescent-like (SenL) phenotypes. A subset of CD8+ T cells expressing CX3CR1 has been reported to be associated with terminal effector (Term) and senescent-like (SenL) phenotypes in AML[7]. Recent work by Chen et al. demonstrated that the basic leucine zipper ATF-like TF (BATF) facilitates the differentiation of TCF-1+ progenitors, including precursor exhausted T cells (Tpex), into CX3CR1+ effector cells by binding to regulatory elements near the *TBX21* locus, which encodes T-bet[12]. Broomfield et al. showed that Tpex cells can reversibly transition into Tscm cells in an antigen load–dependent manner[13]. These findings suggest that Tpex may represent a physiological precursor state of Tscm and that this plasticity is disrupted upon antigen clearance[13]. These findings provide insight into how transcription factor networks regulate the balance between Tpex and Tscm maintenance versus progression toward T cell exhaustion and senescence, depending on antigen load and consequent TCR engagement frequency.

Under conditions of persistent antigen stimulation, a network of transcription factors responsive to TCR signaling plays a pivotal role in both the differentiation of effector CD8+ T cells and the progression toward CD8+ T cell exhaustion. This network includes interferon regulatory factor 4 (IRF4), nuclear factor of activated T cells cytoplasmic 1 (NFATC1), BATF, and thymocyte selection-associated high mobility group box protein (TOX), which collectively contribute to the upregulation of inhibitory receptors such as programmed cell death protein 1 (PD-1)[14]. Man et al. showed that NFATC1 promotes the gene expression of IRF4, and the resulting transcription factor IRF4 and BAFT cooperate to enhance the gene expression of NFATC1. As a result, NFATC1 and IRF4 form a positive loop with BAFT[15], which may function as a bistable switch[16] and play a role in decoding antigen load and consequent TCR engagement frequency.

There is growing evidence that PD-1 and TOX are not solely associated with T cell exhaustion, but also play critical roles in maintaining progenitor-like (Tpex) and stem cell-like (Tscm) CD8+ T cell populations. PD-1 has been shown to upregulate BATF, which in turn enhances TCF-1 expression. Conversely, BATF may also function downstream of TCF-1 within the Tpex compartment, implying a potential bidirectional regulatory loop between PD-1 and TCF-1 mediated by BATF[12,17,18] that controls the balance between the persistence of Tpex and Tscm cells and exhaustion progression. Moreover, TOX has been reported to interact directly with TCF-1[19] and is essential for the establishment of progenitor-like CD8+ T cells during chronic infection[20]. These findings suggest that TOX also contributes to the long-term maintenance of the TCF-1+ stem-like CD8+ T cell pool under sustained antigen exposure. The critical role of this network consisting of TCF-1, BATF, TOX and PD-1 for regulating the balance of the persistence of Tpex and Tscm cells, exhaustion progression and the differentiation into terminal effector (Term) cells and senescent-like (SenL) cells is likely to need the regulation of the T-bet to EOMES transition. In Tpex cells, TCF-1 mediates the transition by promoting the expression of EOMES through binding its regulatory gene and in turn opposes the expression of T-bet. T-bet induces the differentiation of Tpex cells toward the CX3CR1+ effector cells by BATF binding to regulatory elements near the *TBX21* locus described above[12] and by inhibition of the transcription of *PDCD1*, which encodes PD-1[21]. In addition, the finding that TOX may have a direct inhibitory effect on the expression of T-bet[22] suggests the existence of mutual inhibition motif regulating the balance of the persistence of Tpex and Tscm cells, exhaustion progression and the differentiation into terminally differentiated effector T cells (Term) and senescent-like (SenL) cells.

Based on the findings of the aforementioned studies, we hypothesized a regulatory network involving transcription factors that govern CD8+ T cell differentiation under chronic TCR stimulation, as illustrated in Fig. 1.

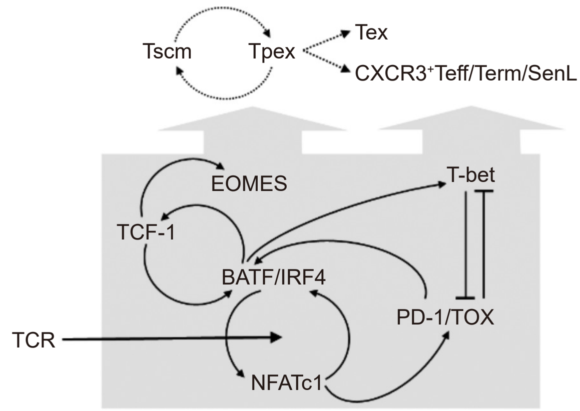

Fig. 1. Hypothetical regulatory network comprising transcription factors proposed to control CD8+ T cell differentiation under chronic TCR stimulation. The processes of activation and inhibition are indicated by solid lines with arrowheads and blunt bars, respectively.

To assess the separability of initial diagnosis and complete remission groups, t-distributed stochastic neighbor embedding (t-SNE) was performed based on the gene expression profiles of the seven transcription factors (BATF, IRF4, NFATC1, TCF-1, EOMES, T-bet, TOX) and PD-1 as shown in Fig. 1. The t-SNE analysis was conducted using the tsne function which is a part of the Statistics and Machine Learning Toolbox in MATLAB R2022a, with the number of dimensions set to 2, perplexity set to 30, and data standardization enabled.

The seven transcription factors (BATF, IRF4, NFATC1, TCF-1, EOMES, T-bet, TOX) and PD-1 as shown in Fig. 1 were validated by examining whether the expressions of these encoding genes are effective for the classification between initial diagnosed patients and CR patients with AML by using an artificial neural network model. The gene expression profiles of GSE134589 datasets were obtained from Gene Expression Omnibus (GEO)[23]. We fitted an artificial neural network model to the relationship between expression data of the eight genes (*BATF*, *IRF4*, *NFATC1*, *PDCD1*, *TCF7*, *EOMES*, *TBX21*, and *TOX*) and two disease states, that is, newly diagnosed state and CR state. The input features to the neural network model were obtained as $\log_2$-transformed and normalized gene expression values. The fitting was performed by using patternnet function in MATLAB 2022a neural network toolbox. Fitting artificial neural networks to data is to find model parameters including weights that attain the best classification performance. A fully connected network was used and the node weigths were updated using Levenberg-Marquardt algorithm based back propogation method, and it has seven hidden layers as shown in Fig. 2.

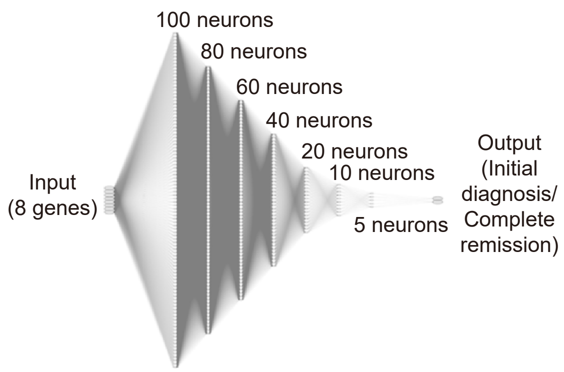

Fig. 2. Architecture of the fully connected neural network model with seven hidden layers.

The input layer consists of eight features corresponding to log2 normalized count of the RNA from eight genes (*BATF*, *IRF4*, *NFATC1*, *PDCD1*, *TCF7*, *EOMES*, *TBX21*, *TOX*). The first, second, third, fourth, fifth, sixth, and seventh hidden layers include 100, 80, 60, 40, 20, 10, 5 neurons, respectively. The output layer is composed of two neurons representing the two target classes (Initial diagnosis encoded as [1 0], Complete remission encoded as [0 1]).

We extracted a total of 432 gene expression profiles of AML patients from GSE134589 datasets[23]. Among them, 368 samples were obtained from newly diagnosed patients before induction therapy, 29 samples from patients who had achieved complete remission (CR) after induction therapy, and 43 samples from relapsed patients. For the present study, we used only the 397 samples obtained from patients in the newly diagnosed and CR state. Due to the considerable class imbalance between these two groups, we performed stratified random sampling based on class labels to divide the dataset into the training data (70%, n = 247) and the testing data (30%, n = 150), while ensuring that the original class distribution was preserved in both data subsets. To further mitigate the class imbalance within the training data, we applied random oversampling to the minority class (CR group). Samples from the CR group were duplicated to match the number of samples in the newly diagnosed group, resulting in a balanced training dataset. This oversampling was applied only to the training set, while the testing set remained untouched to preserve its representativeness. The model was trained on the training data. During training using MATLAB's patternnet function, the 70% training data is further automatically subdivided into training, validation, and internal test sets in a 70:15:15 ratio. The validation set is used for purposes such as early stopping to prevent overfitting. For the final model evaluation, we used the 30% testing data that was completely independent and not involved in any part of the training process. Early stopping was employed based on the validation loss, with a maximum of 6 validation failures. In addition, we applied L2 regularization by setting the parameter net.performParam.regularization = 0.15 in the patternnet model, while dropout was not applied in this implementation.

The training and testing with the same ratio of data division were repeated 20 times with randomly chosen values of parameters. Classification performance into two classes of newly diagnosed state and CR state was evaluated using the accuracy from a confusion matrix.

For data visualization, the figures were created using Python 3.12 via the Spyder IDE (Anaconda distribution), employing matplotlib (v3.9.2) and seaborn (v0.13.2) libraries.

Results

As shown in Fig. 3, the t-SNE plot demonstrates a clear separation between samples from the initial diagnosis and complete remission groups, suggesting the distinct gene expression patterns of the seven transcription factors (BATF, IRF4, NFATC1, TCF-1, EOMES, T-bet, TOX) and PD-1 between the two groups.

As artificial neural networks are a class of non-linear regression methods which can be used to classify via regression, we validated the relationship between the selected eight gene expressions and two distinct disease states of initial diagnosis and complete remission by classification accuracy for two classes, that is, the output layer of the artificial neural network model described in Method section. The classification accuracy as a performance measure of the neural network model was used via a confusion matrix. As a representative example among the 20 independent runs, one trial yielded the accuracies of 99.2% for training and 95.8% for testing. The corresponding confusion matrices are shown in Fig. 4 (training) and Fig. 5 (testing) and the receiver op-

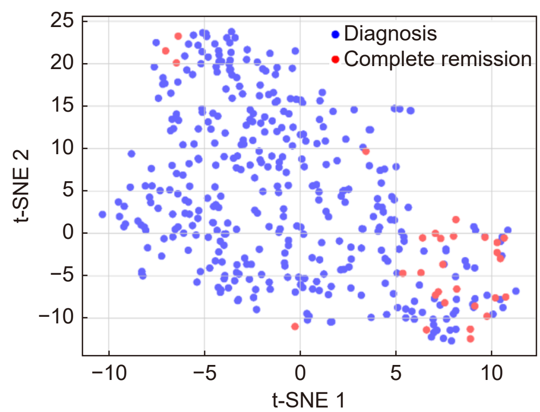

Fig. 3. t-SNE plot of samples from initial diagnosis and complete remission groups based on the expression of eight genes.

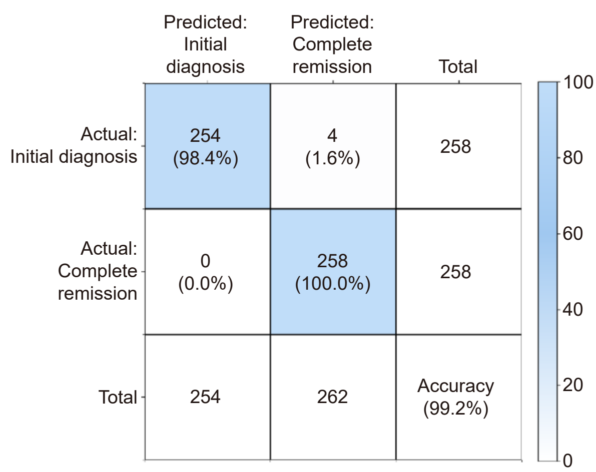

Fig. 4. Confusion matrix for the training set in a representative run (accuracy = 99.2%).

erating characteristic (ROC) curve is shown in Fig. 6. In the confusion matrices, the counts of correct predictions are represented in the main diagonal cells and incorrect predictions are in the non-diagonal cells. Classification accuracy is obtained by considering the count of correctly classified data to the ratio of the total count of training or testing data. As described in Method section, the training and testing with the same ratio of data division were repeated 20 times with randomly chosen values of parameters. The present neural network model achieved the average accuracies 99.2 ± 0.4% of training and 93.8 ± 2.3% of testing (mean ± SD over 20 runs). This result of highly accurate classification suggests the distinct expression patterns of selected eight genes are underlying in initial diagnosis and complete remission.

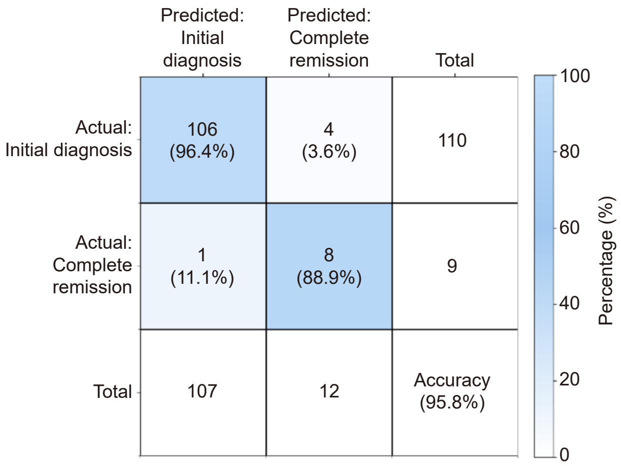

Fig. 5. Confusion matrix for the testing set in a representative run (accuracy = 95.8%).

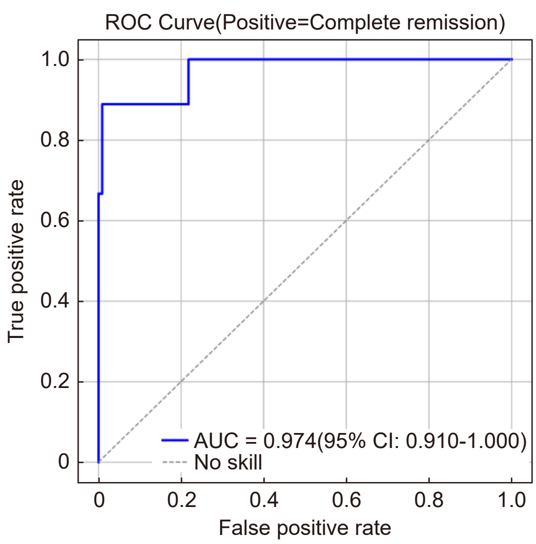

Fig. 6. ROC curve for the testing set in a representative run.

Discussion

T cell is a critical component of the anti-cancer immune microenvironment in AML. The trajectories of $CD8^+$ T cell differentiation from naïve cells to effector cells, memory cells, exhausted cells and senescent cells are influenced by the diversity and abundance of neoantigens originating from gene mutations in AML blasts, and the duration of the antigen exposure. In our present study, we assumed that transient interruptions of AML antigen exposure by cytotoxic effects of induction therapy could induce the change of transcription factors profile underlying the control of $CD8^+$ cell differentiation leading to CR and we validated eight encoding genes using an artificial neural network model. On the other hand, the impact of AML gene mutations on achieving CR with induction therapy was identified and evaluated by use of machine learning models including neural network models. Eckardt et al. reported that gene expression profiles of BM or PB samples from AML patients before induction therapy were obtained and genes were selected as input features applied to nine machine learning classifiers including an artificial neural network model by use of five selection methods[24]. They identified 14 mutated genes including inversion, translocation and deletion as predictors of CR by validation using the machine learning classifiers[24]. Besides validation of mutated genes, machine learning technics were applied for prediction of CR achievement in AML. Gal et al. analyzed gene expression patterns obtained from BM of pediatric patients with AML before induction therapy and identified the genes which were differentially expressed between responder achieving CR and non-responder after treatments. Using the selected top 50 genes as input features, a machine learning technique, the k-nearest neighbors algorithm (K-NN) model was applied to the classification between patients achieving CR and non-responders after induction therapy[25]. Given the preceding studies showing the usefulness of machine learning technics to predicting the likelihood of achieving CR, we focused on transcription factors underlying T cell differentiation as input features for an artificial neural network model. In the present study, the classification models effectively discriminated between patients with newly diagnosed disease and those in complete remission (CR), indicating that these clinical states are associated with distinct transcriptional patterns of the selected eight genes. These findings imply a potential regulatory role of the corresponding transcription factors in $CD8^+$ T cell differentiation under continuous versus interrupted antigen stimulation in AML. However, the limitations of our study should be acknowledged. First, the patient data was derived from a single public dataset that lacks information on gene mutations in AML blasts. As described in the Introduction, the extent of modulation in the cell frequency ratio between naïve T cells and stem-like early memory T cells (Tscm) versus terminally differentiated effector T cells (Term) and senescent-like (SenL) T cells induced by induction therapy was reported to depend on the specific gene mutations present in AML blasts[7]. Furthermore, as mentioned above, gene mutations of AML blasts were identified as predictive features in machine learning models for complete remission and 2-year overall survival[24]. In our further work, it will be important to investigate whether the distinct difference in gene expression profiles between initial diagnosis and complete remission are attributable to specific gene mutations in AML blasts and/or to changes in antigenic stimulation caused by treatment. Second, although the neural network model achieved high classification performance, the validation with separate datasets was not performed. External validation should be performed in our further work. Despite these limitations, our findings provide a basis for future investigation into transcriptional signatures associated with disease state and treatment response in AML.

Conclusion

We assumed that transient interruptions of AML antigen exposure by cytotoxic effects of induction therapy could induce the change of transcription factors profile underlying the control of CD8$^+$ T cell differentiation leading to CR. We selected transcription factors as key components of network motif controlling CD8$^+$ T differentiation by literature review. Using an artificial neural network model, the expressions of the corresponding eight genes as input features were effective for classification between newly diagnosed patients and CR patients after induction therapy with high accuracy, suggesting that the two states have distinct expression patterns of the eight genes and their possible involvement in the modulation of CD8$^+$ T cell differentiation in response to therapy. These insights may inform future targeting and therapeutic strategies in AML.

Abbreviations

2B4, T cell surface receptor, also known as CD244, member of the SLAM family; AML, Acute Myeloid Leukemia; ASXL1, Additional Sex Combs-like 1; BATF, Basic Leucine Zipper ATF-Like Transcription Factor; BM, Bone Marrow; bsAb, Bispecific Antibodies; CAR, Chimeric Antigen Receptors; CD8, Cluster of Differentiation 8; CR, Complete Remission; CXCR3, C-X-C Chemokine Receptor 3; EOMES, Eomesodermin; GEO, Gene Expression Omnibus; GZMB, Granzyme B; ICB, Immune Checkpoint Blockade; IFN-γ, interferon-γ; IL-2, Interleukin-2; IRF4, Interferon Regulatory Factor 4; KLRF1, Killer Cell Lectin like Receptor F1; K-NN, k-Nearest Neighbors Algorithm; NFATC1, Nuclear Factor of Activated T cells Cytoplasmic 1; NPM1, Nucleophosmin 1; PB, Peripheral Blood; PD-1, Programmed Cell Death Protein 1; PDCD1, Programmed Cell Death 1; ROC, Receiver Operating Characteristic; RUNX1, Runt-related Transcription Factor 1; SenL, Senescent-like Cells; T-bet, T-box Expressed in T cells; TBX21, T-box Transcription Factor 21; TCF-1, T cell factor 1; TCR, T-cell Receptor; Term, Terminally Differentiated Effector T cells; Tim-3, T cells Immunoglobulin and Mucin Protein 3; TNF, Tumour Necrosis Factor; TOX, Thymocyte Selection-associated High Mobility Group Box Protein; TP53, Tumor Protein p53; Tpex, Precursor Exhausted T cells; Tscm, Stem Cell Memory T cells; t-SNE, t-distributed Stochastic Neighbor Embedding.

Conflicts of interest

The authors declare no conflicts of interest.

Ethics approval and consent to participate

This study used only publicly available and de-identified data from the Gene Expression Omnibus (GEO) database (accession number: GSE134589). The original study from which the data were derived had obtained ethical approval and informed consent from all participants, as stated in the original publication[23]. Therefore, no additional ethical approval was required for this secondary analysis.

Authors' contribution

YN: Methodology, Validation, Visualization, Writing–original draft, Writing–review & editing, AH: Methodology, Validation, Visualization, Writing–review & editing, FM: Methodology, Validation, Writing–review & editing, NN: Conceptualization, Methodology, Validation, Visualization, Writing–review & editing.

References

[1] Daver N. Immune checkpoint inhibitors in acute myeloid leukemia. Best Pract Res Clin Haematol, 2021, 34(1): 101247.

[2] Hoseini SS, Cheung NK. Acute myeloid leukemia targets for bispecific antibodies. Blood Cancer J, 2017, 7(2): e522.

[3] Daver N, Alotaibi AS, Bücklein V, et al. T-cell-based immunotherapy of acute myeloid leukemia: current concepts and future developments. Leukemia, 2021, 35(7): 1843-1863.

[4] Li Z, Philip M, Ferrell PB. Alterations of T-cell-mediated immunity in acute myeloid leukemia. Oncogene, 2020, 39(18): 3611-3619.

[5] Zhang J, He T, Xue L, et al. Senescent T cells: a potential biomarker and target for cancer therapy. EBioMedicine, 2021, 68: 103409.

[6] Magkouta S, Markaki E, Evangelou K, et al. Decoding T cell senescence in cancer: Is revisiting required? Semin Cancer Biol, 2025, 108: 33-47.

[7] Mazziotta F, Biavati L, Rimando J, et al. CD8$^+$ T-cell differentiation and dysfunction inform treatment response in acute myeloid leukemia. Blood, 2024, 144(11): 1168-1182.

[8] Knaus HA, Berglund S, Hackl H, et al. Signatures of CD8$^+$ T cell dysfunction in AML patients and their reversibility with response to chemotherapy. JCI Insight, 2018, 3(21): e120974.

[9] Kazerani M, Marcinek A, Philipp N, et al. Evolution of T-cell fitness through AML progression: enhanced bispecific T-cell engager-mediated function of bone marrow T cells at remission compared to initial diagnosis and relapse. Leukemia, 2024, 38(10): 2270-2275.

[10] Chen Y, Zander R, Khatun A, et al. Transcriptional and epigenetic regulation of effector and memory CD8 T cell differentiation. Front Immunol, 2018, 9: 2826.

[11] Chan JD, Lai J, Slaney CY, et al. Cellular networks controlling T cell persistence in adoptive cell therapy. Nat Rev Immunol, 2021, 21: 769-784.

[12] Chen Y, Zander RA, Wu X, et al. BATF regulates progenitor to cytolytic effector CD8$^+$ T cell transition during chronic viral infection. Nat Immunol, 2021, 22: 996-1007.

[13] Broomfield BJ, Tan CW, Qin RZ, et al. Transient inhibition of type I interferon enhances CD8$^+$ T cell stemness and vaccine protection. J Exp Med, 2025, 222(5): e20241148.

[14] Kallies A, Zehn D, Utzschneider D T. Precursor exhausted T cells: key to successful immunotherapy? Nat Rev Immunol, 2020, 20(2): 128-136.

[15] Man K, Gabriel SS, Liao Y, et al. Transcription Factor IRF4 Promotes CD8$^+$ T Cell Exhaustion and Limits the Development of Memory-like T Cells during Chronic Infection. Immunity, 2017, 47(6): 1129-1141. e5.

[16] Xiong W, Ferrell J. A positive-feedback-based bistable ‘memory

module' that governs a cell fate decision. Nature, 2003, 426: 460-465.

[17] Quigley M, Pereyra F, Nilsson B, et al . Transcriptional analysis of HIV-specific CD8+ T cells shows that PD-1 inhibits T cell function by upregulating BATF. Nat Med, 2010, 16: 1147-1151.

[18] Boi SK, Lan X, Youngblood B. BATF targets T cell exhaustion for termination. Nat Immunol, 2021, 22(8): 936-938.

[19] Ford BR, Vignali PDA, Rittenhouse NL, et al. Tumor microenvironmental signals reshape chromatin landscapes to limit the functional potential of exhausted T cells. Sci Immunol, 2022, 7(74): eabj9123.

[20] Yao C, Sun HW, Lacey NE, et al. Single-cell RNA-seq reveals TOX as a key regulator of $CD8^+$ T cell persistence in chronic infection. Nat Immunol, 2019, 20: 890-901.

[21] Tian W, Qin G, Jia M, et al. Hierarchical transcriptional network governing heterogeneous T cell exhaustion and its implications for immune checkpoint blockade. Front. Immunol, 2023, 14: 1198551.

[22] Niu H and Wang H. TOX regulates T lymphocytes differentiation and its function in tumor. Front Immunol, 2023, 14: 990419.

[23] Vadakekolathu J, Minden MD, Hood T, et al. Immune landscapes predict chemotherapy resistance and immunotherapy response in acute myeloid leukemia. Sci Transl Med, 2020, 12(546): eaaz0463.

[24] Eckardt JN, Röllig C, Metzeler K, et al. Prediction of complete remission and survival in acute myeloid leukemia using supervised machine learning. Haematologica, 2023, 108(3): 690-704.

[25] Gal O, Auslander N, Fan Y, et al. Predicting Complete Remission of Acute Myeloid Leukemia: Machine Learning Applied to Gene Expression. Cancer Inform, 2019, 18: 1176935119835544.