Leukemia Research 78 (2019) 3–11

Contents lists available at ScienceDirect

Leukemia Research

journal homepage: www.elsevier.com/locate/leukres

Research paper

Computational modeling of early T-cell precursor acute lymphoblastic leukemia (ETP-ALL) to identify personalized therapy using genomics

Ansu Kumar[b], Leylah M. Drusbosky[a,1], Amy Meacham[a], Madeleine Turcotte[a], Priyanka Bhargav[b], Sumanth Vasista[b], Shahabuddin Usmani[b], Anusha Pampana[b], Kabya Basu[b], Anuj Tyagi[b], Deepak Lala[b], Swaminathan Rajagopalan[b], Shivgonda C. Birajdar[b], Aftab Alam[b], Kunal Ghosh Roy[b], Taher Abbasi[c], Shireen Vali[c], Manju Sengar[d], Girish Chinnaswamy[d], Bijal D. Shah[e], Christopher R. Cogle[a,*]

[a] Department of Medicine/Division of Hematology Oncology, University of Florida, Gainesville, FL, USA
[b] Cellworks Research India Pvt. Ltd., Bangalore, India
[c] Cellworks Group Inc., San Jose, CA, USA
[d] Department of Medical Oncology, Tata Memorial Centre, Mumbai, India
[e] Moffitt Cancer Center, Tampa, FL, USA

ARTICLE INFO

Keywords:
ETP-ALL
Genomics
Drug response
Computational-modeling
Leukemia
Biomarkers

ABSTRACT

Early T-cell precursor acute lymphoblastic leukemia (ETP-ALL) is an aggressive hematological malignancy for which optimal therapeutic approaches are poorly characterized. Using computational biology modeling (CBM) in conjunction with genomic data from cell lines and individual patients, we generated disease-specific protein network maps that were used to identify unique characteristics associated with the mutational profiles of ETP-ALL compared to non-ETP-ALL (T-ALL) cases and simulated cellular responses to a digital library of FDA-approved and investigational agents. Genomics-based classification of ETP-ALL patients using CBM had a prediction sensitivity and specificity of 93% and 87%, respectively. This analysis identified key genomic and pathway characteristics that are distinct in ETP-ALL including deletion of nucleophosmin-1 (*NPM1*), mutations of which are used to direct therapeutic decisions in acute myeloid leukemia. Computational simulations based on mutational profiles of 62 ETP-ALL patient models identified 87 unique targeted combination therapies in 56 of the 62 patients despite actionable mutations being present in only 37% of ETP-ALL patients. Shortlisted two-drug combinations were predicted to be synergistic in 11 profiles and were validated by *in vitro* chemosensitivity assays. In conclusion, computational modeling was able to identify unique biomarkers and pathways for ETP-ALL, and identify new drug combinations for potential clinical testing.

1. Introduction

Early T-cell precursor acute lymphoblastic leukemia (ETP-ALL), an orphan disease, is a distinct subtype of T-ALL with a poor prognosis and high risk of relapse following standard chemotherapy [1–3]. A genetically heterogeneous disease, ETP-ALL is associated with a complex karyotype and genetic mutations typically seen in myeloid neoplasia, suggesting a hematopoietic stem cell (HSC) origin [4,5]. Cases commonly display a high burden of chromosomal copy number variations (CNVs) and harbor genomic abnormalities affecting RAS signaling, HSC and progenitor cell maturation and differentiation, and histone modification [4,6,7]. Many of the identified gene mutations associated with ETP-ALL are not directly actionable by currently available targeted therapies, and empirical trials of chemotherapy have not shown

Abbreviations: ETP-ALL, early T-cell precursor acute lymphoblastic Leukemia; GEO, geneexpression omnibus; CCND1, cyclin D1; AKT, AKT serine/threonine kinase; BIRC5, baculoviral IAP repeat containing 5; ETV6, ETS variant 6; LEF-1, lymphoid enhancer binding factor 1; JAK3j, Janus kinase 3; NPM1, nucleophosmin 1; mTORC1, mammalian target of rapamycin complex 1; FOXM1, forkhead box M1; CEBPA, CCAAT enhancer binding protein alpha; HIF1A, hypoxia inducible factor 1 subunit alpha; TET2, Tet methylcytosine dioxygenase 2; FLT3, Fms related tyrosine kinase 3; CDKN2A, cyclin dependent kinase inhibitor 2A; CDKN2B, cyclin dependent kinase inhibitor 2B; TSC1, TSC complex subunit

* Corresponding author at: 1600 SW Archer Road, Box 100278, USA.
E-mail address: christopher.cogle@medicine.ufl.edu (C.R. Cogle).
[1] Author contributed equally to this project.

https://doi.org/10.1016/j.leukres.2019.01.003
Received 15 November 2018; Received in revised form 4 January 2019; Accepted 5 January 2019
Available online 07 January 2019

significant improvements in patient outcomes [3].

One approach to meeting the challenge of ETP-ALL treatment beyond first line involves the rational design of therapy for individual patients based on the panoply of unique molecular abnormalities observed in ETP-ALL [4,8,9]. We previously demonstrated the feasibility of using computational biology modeling (CBM) to create intracellular protein network maps for individual patients with myeloid malignancies (acute myeloid leukemia [AML], myelodysplastic syndromes [MDS]) based on their unique genomic mutations, and to predict drug response [10]. This method, with its ability to simulate multiple gene-drug interactions, accurately predicted clinical response in 80–100% of patients enrolled in three clinical trials. CBM thus has the potential to improve upon empiric prescription of chemotherapy and conventional methods to identify responders and non-responders among leukemia patients with complex karyotypes, including ETP-ALL.

To enhance the understanding of ETP-ALL, we hypothesized that ETP-ALL is comprised of mutational profiles and corresponding protein network disturbances that are distinct from other T-ALL (Non-ETP-ALL) cases. We further hypothesized that a CBM approach can be used to identify key aspects of the molecular signatures and protein network maps of patients with ETP-ALL, with a view toward identifying new therapeutic strategies for patients with higher-risk disease lacking actionable mutations.

2. Materials and methods

2.1. Study population and data sets

The University of Florida Institutional Review Board approved this study (IRB201600284). Genomic data including whole exome sequencing (WES) or whole genome sequencing (WGS), available array comparative genomic hybridization (aCGH; Agilent Technologies, MA, United States) and cytogenetic data of ETP-ALL and Non-ETP-ALL patients were used. All available genomic data for each profile were entered into CBM software (Cellworks Group, San Jose, CA, United States), which generated a disease-specific protein network map summarized in Fig. 1. To compare the mutational profiles and protein networks of predominantly newly diagnosed ETP-ALL and Non-ETP-ALL cases, we divided data from 69 patients into a training set of 40 patients and a unique validation set of 29 patients (Fig. 2A). To characterize the molecular signatures and protein biomarkers of patients with a view toward predicting novel therapeutic strategies, genomic data from 62 predominantly newly diagnosed ETP-ALL patients were used (Fig. 2B).

2.2. Computational biology modeling

CBM has been described in detail elsewhere [10,11]. Briefly, the system is based on > 50,000 PubMed references and online sources, and includes more than 3000 genes, 2500 unique biomarkers and 85,000 functional interactions associated with signaling pathways important in cancer. We used cytogenetic profiling by spectral karyotyping [10] to identify chromosomal aberrations, array comparative genomic hybridization (aCGH) to assess copy number variations (CNV), and whole exome sequencing (WES) for identifying genetic variants to interpret the genomic signature of each patient's disease. The resulting data were compiled to create a list of genes with mutations and CNVs in the patient's genome. The genes found on loci of the affected regions of chromosomes were extracted from the human reference genome (ENSEMBL), with the complete gene list matched with the Cancer CBM to determine the subset to be represented in the model.

2.3. General ETP-ALL model creation

CBM was customized to create a generalized disease model of ETP-ALL based on key pathways and processes involved in pathogenesis from peer-reviewed literature, experimental data, and genotyping.

2.4. Creation of individualized ETP-ALL patient models

Genomic sequencing, gene expression, and/or cytogenetic data for 62 ETP-ALL patients were collected from a GEO database of pediatric ETP-ALL patients (GSE28703; N = 11), five adult and one pediatric cases provided by Moffitt Cancer Center (IRB201600284; N = 6), prior published reports of pediatric T-ALL (N = 43) [1,4], and the LOUCY & PEER cell line (N = 2) [12]. Genomic aberrations were interpreted for phenotypic implications (i.e. gain vs. loss of function) using PubMed to create 62 models customized to each patient's individual genomics. Protein network maps were created based on each patient's unique mutational profile, incorporating intersecting protein networks when multiple genomic abnormalities were detected.

PEER cell line underwent whole exome sequencing (Agilent SureSelect V5; Otogenetics, Atlanta, GA) and CNV extraction (DNAnexus CNVkit), and interpretation (Farsight Genome Systems, Sunnyvale, CA). All data were entered into CBM to generate a protein network map.

2.5. Digital drug models

A digital drug library of 68 FDA-approved and investigational agents was created for CBM by programming each agent's mechanism of action (MOA), as well as effects on specific protein targets and pathways determined from published literature.

2.6. Simulation and drug combination prediction

Virtual applications of all drugs within the digital library were applied individually and in combination in a dose-respondent manner on 62 ETP-ALL models. Each drug or drug combination was mapped to each patient's genomic profile, supported by peer-reviewed data providing a MOA relevant to disease biology. For patients with multiple predicted-to-be successful combination therapies, the top 3 combination therapies were shortlisted for further analysis. For each patient, the efficacy of each therapeutic protocol was measured as a function of disease inhibition score (DIS) – the degree to which crucial cancer phenotypes, such as proliferation and viability, were repressed. The proliferation index is an average function of the active CDK-cyclin complexes that define cell cycle checkpoints, and is determined by calculating permutations in the biomarkers CDK4-CCND1, CDK2-CCNE, CDK2-CCNA, and CDK1-CCNB complex. The viability index is assessed based on survival and apoptosis phenotypes. The biomarkers constituting the survival index include AKT1, BCL2, MCL1, BIRC5, BIRC2, and XIAP, while the apoptosis index includes BAX, CASP3, NOXA, and CASP8. The overall viability index is calculated as a ratio of survival index/apoptosis index. DIS is calculated according to the following formula:

$$DIS = Proliferation + Viability = Proliferation + \frac{Survival}{Apoptosis}$$

Unique combinations that reduced the disease endpoints by interacting with patient-specific disease biomarkers were selected. Synergistic effects of drug combinations were calculated using the coefficient of drug interaction (CDI) formula. For the most efficacious combinations, CDI was calculated to categorize the interaction between constituent drugs as follows:

$$CDI = \frac{A \times B}{AB}$$

where A or B = 1 – DIS% of single agent groups, and AB = 1 – (A + B). A CDI ≥ 1.2 indicates synergistic effect; 0.8 < CDI < 1.2 indicates additive effect; and CDI ≤ 0.8 indicates antagonistic effect.

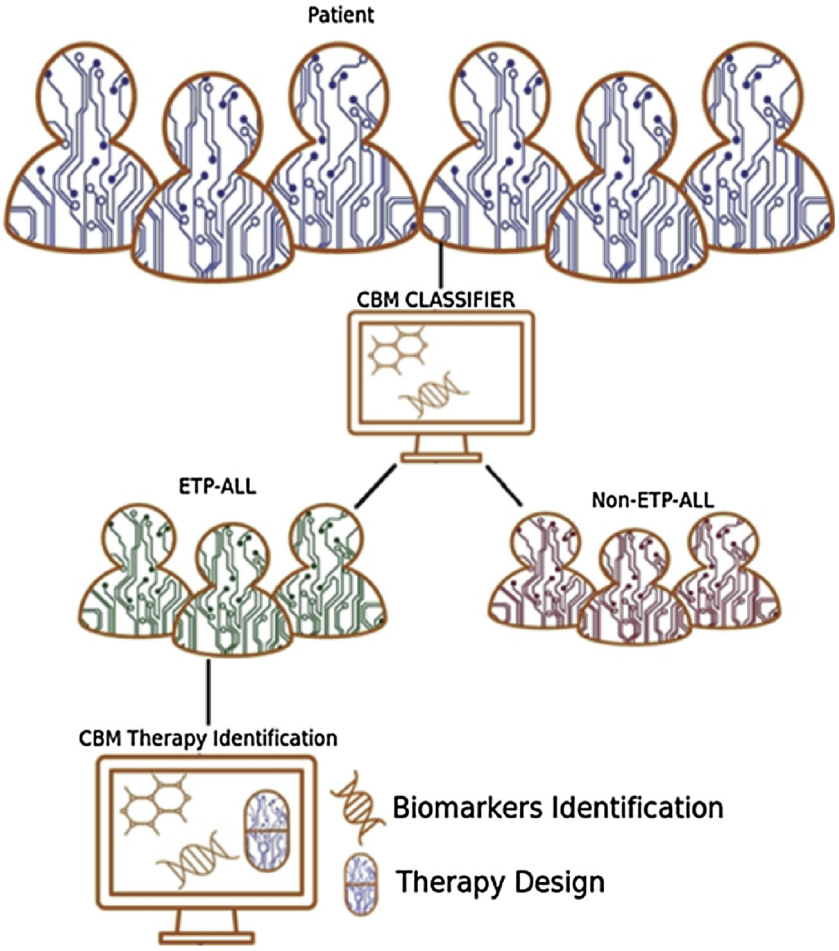

Fig. 1. Individual patient genomic data are entered into CBM software, which generates a disease-specific protein network map used to distinguish between ETP-ALL and non-ETP-ALL cases, identify unique ETP-ALL patient characteristics associated with the mutanome and corresponding dysfunctional protein networks, and predict novel drug combinations based on actionable mutations.

2.7. Classification of ETP-ALL versus Non-ETP-ALL

A computational biology methodology classifier (Cellworks Group) was applied using genomics with the purposes of: (1) distinguishing ETP-ALL patients from non-ETP-ALL patients, and (2) identifying unique biomarkers and pathways for ETP-ALL.

For training of the classifier, a blinded cohort of 40 patients was randomly selected, including 22 patients diagnosed with ETP-ALL and a subset of 18 Non-ETP-ALL patients (Fig. 2a). Genomic data were used from 11 pediatric ETP-ALL patients whose data were deposited in the NCBI Gene Expression Omnibus database (GSE28703) along with 11 pediatric ETP-ALL patients described in a prior genomics report [4]. Genomic data from the 18 Non-ETP-ALL pediatric patients were sourced from a GEO database (GSE28703). Model inputs included genomic profiling data from the training cohort, and phenotypic information from public data resources (PubMed). Using CBM, model inputs were used to translate each patient's genomic data into individual protein network pathways.

Evaluation of the candidate model was conducted with a blinded, independent testing cohort of 29 patients, including 15 ETP-ALL profiles and 14 Non-ETP-ALL profiles. Genomic data from the 15 ETP-ALL patients were distinct from the training set and were variously derived from a GEO database of pediatric patients (GSE28703; N = 1), Moffitt Cancer Center provided five adult cases and one pediatric case (Tampa, FL, United States; IRB201600284; N = 6), a prior genomics report of pediatric patients (N = 7), [4] and the LOUCY cell line with known genomic aberrations similar to ETP-ALL (N = 1) [12]. Genomic data from the 14 Non-ETP-ALL patients distinct from the training set were sourced from a GEO database (GSE28703). Genomic profiling data from all patients were applied to the candidate model whose ability to accurately predict ETP-ALL diagnoses on the unseen sample was assessed. Separation plots were generated for both training and test cohorts using pattern matching scores and relative inclusion/exclusion scores to evaluate the candidate model.

Pattern Matching Score

$$\textbf{Fractional Similarity Score} = \begin{array}{c}\textit{genomic similarity}\\ \textit{between ETP} - \textit{ALL}\\ \textit{and Non} - \textit{ETP} - \textit{ALL}\\ \textit{patients}\end{array} + \textbf{Response Score} \begin{array}{c}1 = \textit{ETP} - \textit{ALL}\\ -1 = \textit{Non} - \textit{ETP} - \textit{ALL}\end{array}$$

Relative Inclusion/Exclusion Score

$$\textbf{Pathway Dominance Score} = \begin{array}{c}\textit{sum of abberated genes}\\ \textit{per pathway}\\ + \textit{pathway dominance}\end{array} + \textbf{Correlation Score} \begin{array}{c}\textit{positive or negative}\\ \textit{correlation between}\\ \textit{gene and ETP}-\\ \textit{ALL diagnosis}\end{array}$$

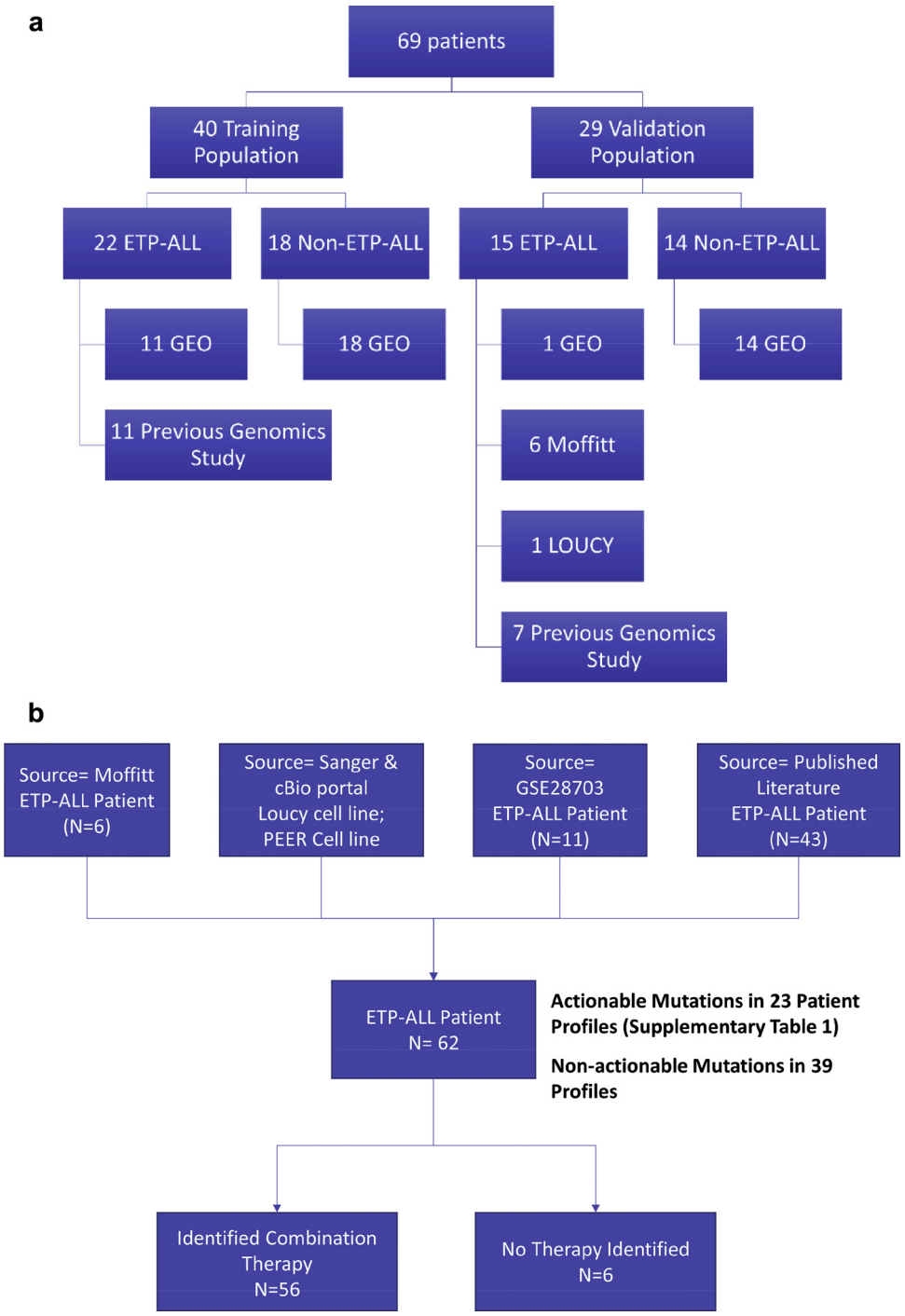

Fig. 2. Schematic summarizing patient and data sources used to compare ETP-ALL and non-ETP-ALL cases (A) and characterize the molecular signatures and protein biomarkers of ETP-ALL patients to predict novel therapeutic strategies (B).

2.8. In vitro validation

LOUCY cell line was purchased from ATCC (Manassas, VA) and cultured in RPMI-1640 (ATCC) supplemented with 10% fetal bovine serum (FBS) (Atlanta Biologicals, Flowery Branch, GA) and 1X penicillin-streptomycin-glutamine (PSQ) (Gibco- ThermoFisher Scientific, Waltham, MA). Cells were incubated at 37C and 5% CO_2. PEER cell line was purchase from DSMZ (Braunschweig, Germany) and cultured in RPMI-1640 (ATCC) supplemented with 20% FBS (Atlanta Biologicals) and 1X PSQ (Gibco). Cells were incubated at 37C and 5% CO_2. KG1 cells were purchased from ATCC (Manassas, VA) and cultured in IMDM (ATCC) supplemented with 20% FBS (Atlanta Biologicals) and 1X PSQ (Gibco). Cells were incubated at 37C and 5% CO_2. Jurkat Cells were purchased from ATCC and cultured in RPMI-1640 (ATCC) supplemented with 10% FBS (Atlanta Biologicals) and 1X PSQ (Gibco). Cells were incubated at 37C and 5% CO_2.

Loucy and PEER cell lines were selected as representatives of ETP-ALL disease model. Loucy was first established from the peripheral blood of a patient with T-ALL and has ETP-ALL biomarkers seen in the patient cohort [32–34]. PEER cell line was established from the peripheral blood of a 4-year old girl with T-ALL. This cell line also has biomarkers seen in ETP-ALL patients [33]

JURKAT and KG1 are selected as representatives of non ETP-ALL disease models from lymphoid and myeloid origins respectively.

JURKAT is a human T lymphoblastoid cell line derived from an acute T cell leukemia [33]. The KG-1 cell line, established from bone marrow cells of a patient with acute myelogenous leukemia (AML) [35].

2.9. Drug sensitivity assay

The following agents were purchased from Selleckchem (USA) and used for drug sensitivity assays: nelfinavir, palbociclib, cytarabine, bortezomib, nilotinib, and idarubicin. Each drug was dissolved in dimethyl sulfoxide (DMSO) at a concentration of 10 mM, and used as a working stock. Serial dilutions were performed for drug sensitivity assays.

LOUCY and PEER cells were plated in 96-well plates at a density of 10 [5] cells/mL, in their respective media. Single drugs were added to the wells at seven concentrations ranging between 0.0001–500 μM. Cells were treated for 24 h and analyzed using XTT (ATCC) on a Bio-Tek Synergy HT plate reader, per manufacturer's protocol. Each drug's IC_{50} value was calculated using GraphPad Prism 7 software (La Jolla, CA). Each drug treatment was performed in triplicate and normalized to a DMSO treated control.

Combination drug treatments were performed with multiple concentrations ranging from 0.0001 to 300 μM on LOUCY, PEER, JURKAT, and KG1 cells. Cells were treated for 24 h and analyzed by XTT. Each drug treatment was performed in triplicate and normalized to a DMSO treated control. CompuSyn software was used to calculate the combination index (CI) to identify synergistic vs. additive effects of combination treatments, as well as the dose reduction index (DRI).

2.10. Statistical analyses

Statistical parameters of accuracy, positive predicted value (PPV), negative predicted value (NPV), sensitivity and specificity were calculated for the CBM-based classification model. Paired student's *t*-test was used in drug sensitivity assays. $p < 0.05$ was considered statistically significant.

3. Results

3.1. Identification of unique ETP-ALL characteristics through CBM-based classification

40 profiles (22 ETP-ALL and 18 non-ETP-ALL) as described in methods (Fig. 2) were used in the CBM training set for classification based on genomic inputs. This yielded a prediction accuracy in the blinded validation set (N = 29; 15 ETP-ALL and 14 non-ETP-ALL, Fig. 2) of 90% based on a PPV of 87%, NPV of 93%, and a sensitivity and specificity of 93% and 87%, respectively (Table 1). Key genomic characteristics of the 2 false positive and 1 false negative profile are mentioned in supplementary table 5. Among genes known to be mutated in T-ALL, [4,13,14] activating mutations of *NRAS* were found at a higher frequency in ETP-ALL vs. non-ETP-ALL cases (32% vs. 9%), whereas *LEF-1* mutations (0% vs. 22%) and deletions in chromosome 9p (del[9p]; 0% vs. 63%) were notable for their absence. In addition, del (9q) occurred at a lower frequency in ETP-ALL vs. non-ETP-ALL cases (26% vs. 63%). [5] Of genomic aberrations previously reported in ETP-ALL cases, deletion of *ETV6* [4,9] was observed in 42% vs. 9% of ETP-ALL and non-ETP-ALL cases, respectively, del(12p) occurred in 32% vs. 9% [1], *JAK3* mutations were observed in 23% vs. 0%, [4] del(5q) occurred in 26% vs. 3% [1,3], and del(17p) occurred in 13% vs. 3% [1]. We further report a novel genomic aberration associated with ETP-ALL cases that involved deletion of nucleophosmin 1 (*NPM1*; 26% vs. 3% in non-ETP-ALL cases), mutations of which are commonly found in patients with AML, frequently co-occurring with *NRAS* and *FLT3* mutations, and clinically impacting therapeutic decision making in AML (Table 2A). [15,16,17,18].

Table 2A
Frequency of genomic aberrations found in ETP-ALL and Non-ETP-ALL cases Del, deletion; Mut, mutation.

Genomic Aberration	ETP-ALL	Non-ETP-ALL	Student T Test (p Value)
ETV6 –Del	42%	9 %	0.0063*
ETV6 – MUT	16%	9%	0.58
NRAS – MUT	32%	9%	0.059
Del 5q	26%	3%	0.061
NPM1 – Del	26%	3%	0.061
Del 12p	32%	9%	0.059
JAK3 – MUT	23%	0%	0.062
Del 17p	13%	3.%	0.42
Del 9p	0%	63%	< 0.00001*
LEF1 – MUT	0%	22%	0.072

* $p < 0.05$ indicates significance.
Del, deletion; Mut, mutation.

NPM1 gene is located on chromosome 5q 35.1. CNV loss of NPM1 identified along with deletion of 5q chromosome arm across 26% of ETP-ALL population. 5q deletion is also very frequent in AML population. There is one patient identified with NPM1-exon 9 rearrangement, considered as a deleterious mutation.

3.2. Key differences between ETP-ALL and non-ETP-ALL cases

We used the CBM simulation model and the combined training and validation sets to identify biomarkers present at a higher frequency in 37 ETP-ALL profiles compared with 32 non-ETP-ALL profiles. Among these, simulation-based selection of 6 protein biomarkers predicted increased activity in ETP-ALL profiles vs. non-ETP-ALL cases. Runt-related transcription factor-1 (RUNX1) was found in 41% of ETP-ALL cases vs. 10% of non-ETP-ALL cases together with forkhead box protein M1 (FOXM1; 56% vs. 19%), mammalian target of rapamycin complex 1 (mTORC1; 31% vs. 6%), signal transducer and activator of transcription-3 (STAT3; 38% vs. 6%), the leucine zipper CCAAT-enhancer binding protein A(CEBPA; 23% vs. 3%), and hypoxia-inducible factor 1A (HIF1A; 6% vs. 0%) (Table 2B). In addition to these protein biomarkers, CBM also identified that folate synthesis pathway, FOXO

Table 1
Test cohort evaluation of CBM classification of ETP-ALL versus Non-ETP-ALL.

	ETP-ALL	Non-ETP-ALL	Predictive value
Predicted ETP-ALL	True Positive 13	False Positive 2	Positive: 86.67% (95% CI: 63.96%–95.97%)
Predicted Non-ETP-ALL	False Negative 1	True Negative 13	Negative: 92.86% (95% CI: 66.06%–98.86%)
Sensitivity/ Specificity	Sensitivity 92.86% (95% CI: 66.13%– 99.82%)	Specificity 86.67% (95% CI: 59.54%–98.34%)	Accuracy: 89.66%
Likelihood Ratio	Negative Likelihood Ratio 0.08 (95% CI: 0.01-0.55)	Positive Likelihood Ratio 6.98 (95% CI: 1.90-25.51)	

Table 2B
Biomarkers with predicted increased activity in ETP-ALL vs. Non-ETP-ALL cases.

Biomarker	ETP-ALL	Non-ETP-ALL	Student T Test (p Value)
STAT3	38%	6%	0.0052
RUNX3	41%	10%	0.0067
HIF1A	6%	0%	0.32
FOXM1	56%	19%	0.0014
CEBPA	25%	3%	0.0415
mTORC1 Complex	31%	6%	0.0240

* $p < 0.05$ indicates high significance.

signaling, cell cycle and JAK-STAT pathways were found to be dominant within the ETP-ALL cohort relative to non-ETP-ALL cohorts.

3.3. Predictive modeling for targeted treatment

Simulation avatars were created for 62 ETP-ALL profiles selected for CBM to identify targeted treatment options, from which disease-specific biomarkers were identified. Of these, only 23 (37%) harbored actionable mutations, including *NRAS*, *TET2*, *FLT3* and various deletions such as del(5q) or deletions in tumor suppressor genes such as *CDKN2 A/2B*, *TSC1* and others (**Supplementary Table 1**). 39 of 62 (63%) avatars lacked known clinically actionable mutations based on currently available drugs, but harbored other somatic variants or variants of unknown significance with predicted deleterious protein effects (data not shown). Due to its focus on identifying dysregulated protein pathways unique to the individual rather than a population-driven approach to treating disease, the integrated genomics and CBM workflow identified a total of 87 unique targeted combination therapies compatible with the dominant protein pathways of 56 of 62 (90%) cases (**Supplementary Table 2**). As calculated by CDI, shortlisted 2-drug combinations were predicted to be synergistic in 11 ETP-ALL patients and had a significant impact on their unique biomarkers from the classification analysis (**Supplementary Table 3**). Five of these synergistic combinations included the anthracycline, idarubicin, which is used in combination with cytarabine as first-line therapy for AML [19]. The anti-myeloma drugs lenalidomide and bortezomib were represented in four and two of these synergistic combinations, respectively, and the protease inhibitor, nelfinavir, and the poly ADP ribose polymerase (PARP) inhibitor, olaparib, were synergistic in two combinations each. With a CDI in one patient of 3.34, lenalidomide/nelfinavir demonstrated the most synergy followed by idarubicin/olaparib (2.62), and idarubicin/melphalan (2.01); (**Supplementary Table 3**). All other synergistic drug combinations had a CDI < 2. The remaining 2-drug combinations shortlisted for other patients on the basis of their individual protein maps were predicted to be mostly additive and a few were antagonistic despite a high DIS (**Supplementary Table 2**). Key genomic characteristics of the 6 patients where combination drug therapy was not selected are described in supplementary table 6.

3.4. Laboratory validation of computational biology modeling

ETP-ALL cell lines, LOUCY and PEER, were used to confirm the synergistic drug combinations predicted by CBM. KG1 cells, an acute myelogenous leukemia (AML) cell line, and Jurkat cells, an acute T-cell leukemia cell line (T-ALL), were used as non-ETP-ALL comparisons. Genomic data was acquired for the LOUCY cell line [12] and modeled to predict drugs with potential therapeutic activity, by simulating the drug's effect on the protein network map and were predicted to respond to a combination of palbociclib + nelfinavir **(**Fig. 3A,B**)** and nelfinavir + everolimus (Fig. 3C). An XTT assay was performed in the presence of increasing drug doses, both as single agents and in combination, confirming synergistic activity in three of five treatment conditions (Fig. 3**A**,**B Supplementary Table 4**). LOUCY cells showed enhanced sensitivity to palbociclib + nelfinavir as compared to non-ETP-ALL cell lines, KGI and Jurkat (Fig. 3B). We also performed XTT assay of single drug components nelfinavir, everolimus and palbociclib and confirmed higher efficacy on LOUCY cell line (Fig. 3**D, E and F**). Predictive dose response plots of nelfinavir, everolimus, palbociclib and combination of palbociclib + nelfinavir tested on CBM generated JURKAT, KG1 and LOUCY cell lines are included in Supplementary Fig. 1. The CBM predicted enhanced effect of the selected therapy on LOUCY cell line as validated experimentally (Fig. 3B).

Similarly, PEER cells underwent whole exome sequencing, CNV analysis, and computational modeling to predict drugs with potential therapeutic activity. CBM predicted PEER cells to respond to three drug combinations: cytarabine + bortezomib, cytarabine + nilotinib, and bortezomib + idarubicin. XTT assays were performed in the presence of increasing drug doses, both as single agents and in combination, confirming synergistic activity in 16 of 18 treatment conditions. Synergy was observed in each treatment condition involving cytarabine + bortezomib and cytarabine + nilotinib. However, in the bortezomib + idarubicin conditions, the lowest drug concentrations were antagonistic, and the highest drug concentrations were additive (Fig. 4 **A-C**, **Supplementary Table 4**). PEER cells were significantly more sensitive to each drug combination as compared to KG1 and Jurkat cells (Fig. 4 **d-F**). Predictive dose response plots of the selected combinations of cytarabine + bortezomib, cytarabine + nilotinib and cytarabine + idarubicin tested on CBM generated JURKAT, KG1 and PEER cell lines are included in Supplementary Fig. 2. The CBM predicted enhanced effect of the selected therapies on PEER cell line correlated with the in vitro experimental results (Fig. 4A-C).

4. Discussion

Clinical outcomes of patients with ETP-ALL are mixed, with early reports suggesting a high-risk subset of T-ALL characterized by an immature immunophenotype, increased genomic instability, distinct gene expression profile, poor early response to conventional chemotherapy, and high risk of relapse. [1–3] This characterization has been tempered by recent reports suggesting similar clinical outcomes for patients with ETP-ALL and non-ETP-ALL phenotypes [20,22,23], a discrepancy which may in part reflect the implementation of early response-based intensification strategies for patients with immunophenotypically defined high-risk disease, and targeting treatment to the myeloid lineage and HSC compartment [20,21]. Indeed, the observation of genetic and transcriptional features that overlap with AML as well as T-ALL has led to suspicion that ETP-ALL has a HSC origin and therefore necessitates the use of therapeutic strategies appropriately tailored to the genomic landscape of the disease [4,5]. Current clinical practice is to treat such patients through the empiric administration of chemotherapy. In this study, we tested the hypothesis that it is possible to project the genomic abnormalities of ETP-ALL patient profiles into dysregulated protein network maps that can be used not only to identify optimal disease-specific treatment strategies for the patient population, but also to simulate drug treatments appropriate for the individual.

We used WES, CNV and cytogenetics to accurately classify profiles of ETP-ALL and Non-ETP-ALL patients into discrete molecular subsets for subsequent CBM simulations to identify potentially therapeutic drug regimens. We further demonstrated using CBM-based classification that ETP-ALL cases were distinct from other T-ALL cases with respect to genomic abnormalities and projected protein network abnormalities. Consistent with previous reports in ETP-ALL, we found an increased frequency of aberrations in genes known to be involved in regulating cytokine receptor and RAS signaling (*NRAS*, *JAK3*), [4] and hematopoietic development and leukemogenesis (*ETV6*), [4,9] as well as a high prevalence of 5q, 12p and 17p chromosomal deletions [1,3]. Conversely, ETP-ALL cases lacked del(9p), which delete the *CDKN2 A/B* tumor suppressor gene in most T-ALL cases [5], and mutations of *LEF-1* which is involved in thymocyte differentiation [14]. A novel finding

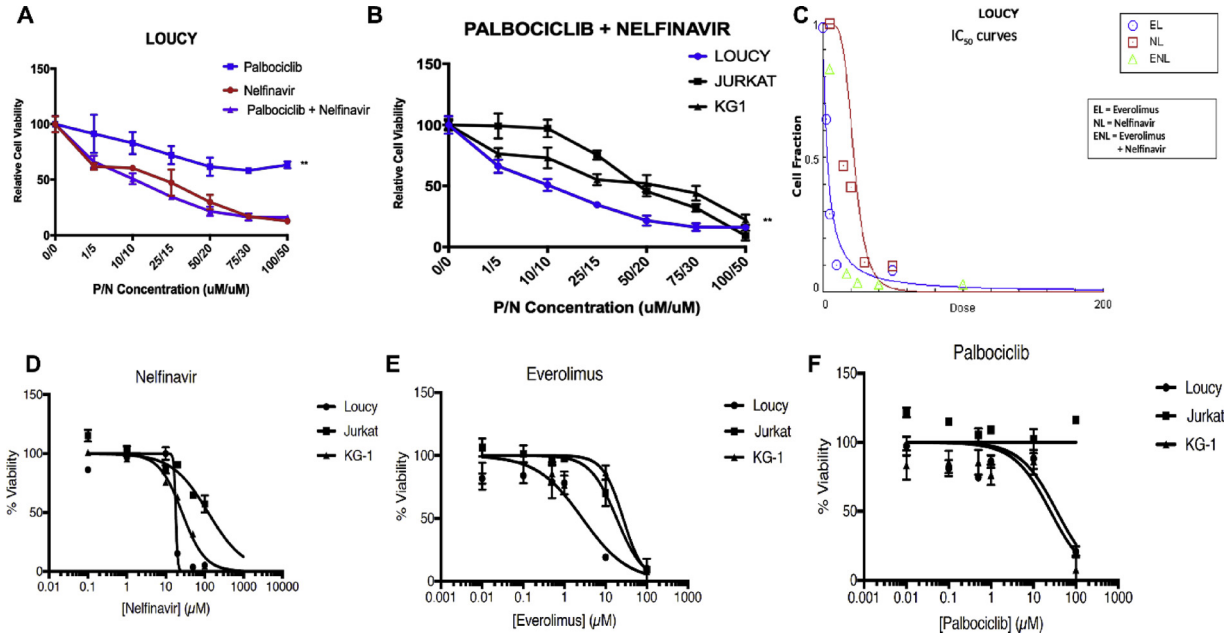

Fig. 3. XTT drug sensitivity data of LOUCY cells treated with various doses of palbociclib, nelfinavir, and palbociclib + nelfinavir (A), treatment of palbociclib + nelfinavir combination on LOUCY cells compared to KG1 and Jurkat (B), various doses of everolimus, nelfinavir, and everolimus + nelfinavir (C), treatment of nelfinavir alone on LOUCY cells compared to KG1 and Jurkat (D), treatment of nelfinavir everolimus alone on LOUCY cells compared to KG1 and Jurkat (E), treatment of palbociclib alone on LOUCY cells compared to KG1 and Jurkat (F). $p < 0.05$ is considered statistically significant (**Supplementary Table 7**).

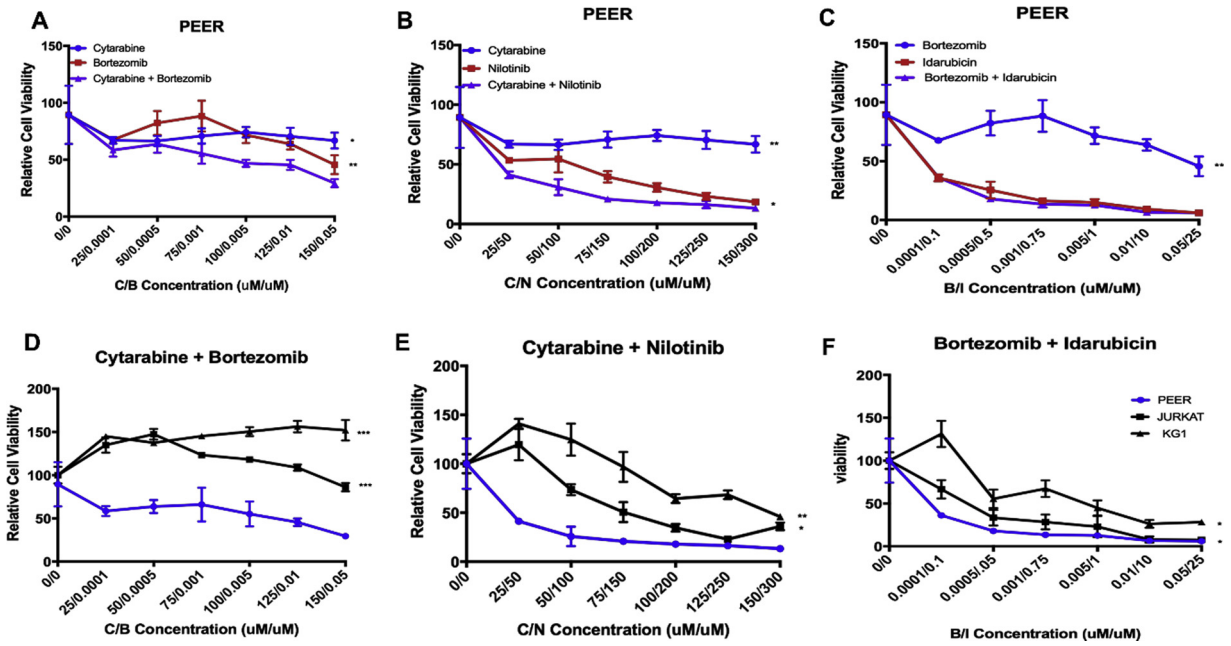

Fig. 4. XTT drug sensitivity data of PEER cells treated with various doses of cytarabine, bortezomib, and cytarabine + bortezomib (A); cytarabine, nilotinib, and cytarabine + nilotinib (B); bortezomib, idarubicin, and bortezomib + idarubicin (C); and treatment of each combination on PEER cells compared to KG1 and Jurkat (C–D). $p < 0.05$ is considered statistically significant (**Supplementary Table 7**).

was the high frequency of cases involving deletion of *NPM1*, which is typically associated with patients with MDS or AML [15–18]. Simulation-based selection was used to identify proteins that were commonly dysregulated in ETP-ALL cases. Among these, RUNX1, FOXM1 and CEBPA are key transcription factors involved in hematopoiesis and myeloid cell differentiation [4,5,24,25], STAT3, a transcription activator whose phosphorylation is associated with the majority of ETP-ALL cases [6], and mTORC1, whose signaling pathway is often activated in ALL and is involved in regulating cell growth and metabolism [26]. In addition, HIF1□, absent altogether from Non-ETP-ALL cases, was present in a small proportion of ETP-ALL cases, suggesting a role in tumor aggression and progression [27].

In addition to differentiating between T-ALL subsets, the potency of CBM resides in its capacity to simulate response to chemotherapy and rapidly identify near-limitless iterations of single-agent and combination treatment options in individual patients. As a molecularly complicated disease, the use of computer analytics affords the opportunity to track the large number of molecular abnormalities in T-ALL and model their complex interactions within dysregulated protein networks as we have previously shown for MDS [10,11]. Herein, CBM was used

to predict appropriate treatment in 54 ETP-ALL patients, and two ETP-ALL cell lines, whose protein maps were projected from their mutational profiles, including those associated with dysregulated protein pathways involving RUNX1, FOXM1, STAT3 and CEBPA. Dysregulated protein maps involving *NRAS* mutations predicted treatment with the mitogen-activated kinase (MEK) inhibitor, trametinib, in nine cases, whereas the presence of del(5q) and deletions of *CDKN2A/B/C* predicted treatment with lenalidomide and the CDK inhibitor, palbociclib, in five and four cases, respectively. The majority of patient avatars involved deleterious mutations that were not directly actionable using a single-drug approach. Despite this, CBM identified 87 unique targeted combination therapies compatible with the dominant dysregulated protein networks in 89% of cases overall, with the selected 2-drug combinations predicted to be synergistic in eleven cases.

Among the synergistic 2-drug combinations, three have been evaluated in AML patients including bortezomib/idarubicin [28], lenalidomide/cytarabine [29] and daunorubicin/midostaurin [30]. This is interesting in the present context given the putative HSC origins of ETP-ALL, but it is noteworthy that none of these combinations have been evaluated in clinical trials of ETP-ALL patients. Although new drug combinations were identified by CBM and validated with in vitro chemosensitivity assays, follow-up prospective clinical trials are needed to validate the clinical utility of these new drug combinations in ETP-ALL. Preliminary data from a prospective clinical trial of MDS and AML patients showed that CBM had high predictive values of protein network identification and predicting clinical outcomes in 80 patients who received standard of care chemotherapy (NCT02435550). [31] Importantly, the method uncovered molecular reasons for treatment failure and highlighted resistance pathways that could be targeted to recover chemosensitivity. Molecular profiling of tumor samples using NGS is now a feasible option being widely used in clinical settings. Therefore, in relatively rare diseases such as ETP-ALL, determining optimal treatment options based on the molecular profile and characteristics of the tumor using CBM analysis can be applied to establishing eligibility for precision enrolment of patients in clinical trials.

In conclusion, these findings show that ETP-ALL is distinct from other T-ALL subtypes with respect to genomic and protein network abnormalities, with CBM identifying unique biomarkers that accurately distinguish between ETP-ALL and non-ETP-ALL cases. CBM was further used to identify new drug combinations for future clinical trial testing. As a predictive technology, CBM has the ability to improve the utility of genetic profiling and may enhance clinical decision-making to identify treatments for ETP-ALL patients who would otherwise have limited options.

Conflict of interest

AK, PB, SV, SU, AP, KB, AT, DL, SR, SB, AA, KR, TA, SV are employees of Cellworks Group, Inc.

Acknowledgements

CRC and LMD were supported by a grant from Gateway for Cancer Research Foundation (G-15-700) and the Harry T. Mangurian Foundation (F022243). CRC was supported by a Scholar in Clinical Research award from the Leukemia & Lymphoma Society (0725-14) and a Pierre Chagnon Professorship. Medical writer Howard Christian assisted in manuscript preparation. Cellworks Research India Pvt Ltd received a grant from TATA Trust for this project.

AK, LMD, CRC, SV, and TA designed the study. LMD, AK, AM, and MT performed experiments, analyzed data, created figures, and assisted in manuscript writing. AK, PB, SV, SU, AP, KB, AT, DL, SR, SB, AA, KR, TA and SV analyzed data and created tables. MS, GC, and BDS contributed patient data.

References

[1] E. Coustan-Smith, C.G. Mullighan, M. Onciu, F.G. Behm, S.C. Raimondi, D. Pei, et al., Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia, Lancet Oncol. 10 (February (2)) (2009) 147–156.
[2] M. Dose, F. Gounari, Sleeping beauty: does ETP-ALL awaken later? Blood 118 (October (17)) (2011) 4500–4501.
[3] N. Jain, A.V. Lamb, S. O'Brien, F. Ravandi, M. Konopleva, E. Jabbour, et al., Early T-cell precursor acute lymphoblastic leukemia/lymphoma (ETP-ALL/LBL) in adolescents and adults: a high-risk subtype, Blood 127 (April (15)) (2016) 1863–1869.
[4] J. Zhang, L. Ding, L. Holmfeldt, G. Wu, S.L. Heatley, D. Payne-Turner, et al., The genetic basis of early T-cell precursor acute lymphoblastic leukaemia, Nature 481 (January (7380)) (2012) 157–163.
[5] J.E. Haydu, A.A. Ferrando, Early T-cell precursor acute lymphoblastic leukaemia, Curr. Opin. Hematol. 20 (July (4)) (2013) 369–373.
[6] E. Danis, T. Yamauchi, K. Echanique, X. Zhang, J.N. Haladyna, S.S. Riedel, et al., Ezh2 controls an early hematopoietic program and growth and survival signaling in early t cell precursor acute lymphoblastic leukemia, Cell Rep. 14 (March (8)) (2016) 1953–1965.
[7] A. Gutierrez, A. Kentsis, T. Sanda, L. Holmfeldt, S.C. Chen, J. Zhang, et al., The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia, Blood 118 (October (15)) (2011) 4169–4173.
[8] B. Knoechel, A. Bhatt, L. Pan, C.S. Pedamallu, E. Severson, A. Gutierrez, et al., Complete hematologic response of early T-cell progenitor acute lymphoblastic leukemia to the gamma-secretase inhibitor BMS-906024: genetic and epigenetic findings in an outlier case, Cold Spring Harb. Mol. Case Stud. 1 (October (1)) (2015) a000539.
[9] M. Neumann, S. Heesch, C. Schlee, S. Schwartz, N. Gokbuget, D. Hoelzer, et al., Whole-exome sequencing in adult ETP-ALL reveals a high rate of DNMT3A mutations, Blood 121 (June (23)) (2013) 4749–4752.
[10] L. Drusbosky, C. Medina, R. Martuscello, K.E. Hawkins, M. Chang, J.K. Lamba, et al., Computational drug treatment simulations on projections of dysregulated protein networks derived from the myelodysplastic mutanome match clinical response in patients, Leuk. Res. 52 (January) (2017) 1–7.
[11] L.M. Drusbosky, C.R. Cogle, Computational modeling and treatment identification in the myelodysplastic syndromes, Curr. Hematol. Malig. Rep. (September) (2017) 14.
[12] H. Ben-Bassat, Z. Shlomai, G. Kohn, M. Prokocimer, Establishment of a human T-acute lymphoblastic leukemia cell line with a (16;20) chromosome translocation, Cancer Genet. Cytogenet. 49 (Octobre (2)) (1990) 241–248.
[13] M. Kaparou, D. Choumerianou, C. Perdikogianni, G. Martimianaki, M. Kalmanti, E. Stiakaki, Enhanced levels of the apoptotic BAX/BCL-2 ratio in children with acute lymphoblastic leukemia and high-risk features, Genet. Mol. Biol. 36 (March (1)) (2013) 7–11.
[14] S. Yu, X. Zhou, F.C. Steinke, C. Liu, S.C. Chen, O. Zagorodna, et al., The TCF-1 and LEF-1 transcription factors have cooperative and opposing roles in T cell development and malignancy, Immunity 37 (November (5)) (2012) 813–826.
[15] B. Falini, C. Mecucci, E. Tiacci, M. Alcalay, R. Rosati, L. Pasqualucci, et al., Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype, N. Engl. J. Med. 352 (January (3)) (2005) 254–266.
[16] Cancer Genome Atlas Research N, T.J. Ley, C. Miller, L. Ding, B.J. Raphael, A.J. Mungall, et al., Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia, N. Engl. J. Med. 368 (May (22)) (2013) 2059–2074.
[17] E. Papaemmanuil, M. Gerstung, L. Bullinger, V.I. Gaidzik, P. Paschka, N.D. Roberts, et al., Genomic classification and prognosis in acute myeloid leukemia, N. Engl. J. Med. 374 (June (23)) (2016) 2209–2221.
[18] O.M. Dovey, J.L. Cooper, A. Mupo, C.S. Grove, C. Lynn, N. Conte, et al., Molecular synergy underlies the co-occurrence patterns and phenotype of NPM1-mutant acute myeloid leukemia, Blood (August) (2017) 23.
[19] E.M. Stein, M.S. Tallman, Remission induction in acute myeloid leukemia, Int. J. Hematol. 96 (August (2)) (2012) 164–170.
[20] J. Bond, C. Graux, L. Lhermitte, D. Lara, T. Cluzeau, T. Leguay, et al., Early response-based therapy stratification improves survival in adult early thymic precursor acute lymphoblastic leukemia: a group for research on adult acute lymphoblastic leukemia study, J. Clin. Oncol. 35 (August (23)) (2017) 2683–2691.
[21] V. Conter, M.G. Valsecchi, B. Buldini, R. Parasole, F. Locatelli, A. Colombini, et al., Early T-cell precursor acute lymphoblastic leukaemia in children treated in AIEOP centres with AIEOP-BFM protocols: a retrospective analysis, Lancet Haematol. 3 (February (2)) (2016) e80–e86.
[22] R. Fukano, S. Sunami, M. Sekimizu, T. Takimoto, T. Mori, T. Mitsui, et al., Clinical features and prognosis according to immunophenotypic subtypes including the early T-Cell precursor subtype of T-Lymphoblastic lymphoma in the japanese pediatric Leukemia/Lymphoma study group ALB-NHL03 study, J. Pediatr. Hematol. Oncol. (May) (2017) 22.
[23] D.M. Sayed, H.A.R. Sayed, H.N. Raslan, A.M. Ali, A. Zahran, R. Al-Hayek, et al., Outcome and clinical significance of immunophenotypic markers expressed in different treatment protocols of pediatric patients with T-ALL in developing countries, Clin. Lymphoma Myeloma Leuk. 17 (July (7)) (2017) 443–449.
[24] R. Avellino, R. Delwel, Expression and regulation of C/EBPalpha in normal myelopoiesis and in malignant transformation, Blood 129 (April (15)) (2017) 2083–2091.
[25] Y. Hou, W. Li, Y. Sheng, L. Li, Y. Huang, Z. Zhang, et al., The transcription factor Foxm1 is essential for the quiescence and maintenance of hematopoietic stem cells, Nat. Immunol. 16 (August (8)) (2015) 810–818.
[26] S. Daakour, L.J. Hajingabo, D. Kerselidou, A. Devresse, R. Kettmann, N. Simonis,

et al., Systematic interactome mapping of acute lymphoblastic leukemia cancer gene products reveals EXT-1 tumor suppressor as a Notch1 and FBWX7 common interactor, BMC Cancer 16 (May) (2016) 335 26.
[27] J. Zou, P. Li, F. Lu, N. Liu, J. Dai, J. Ye, et al., Notch1 is required for hypoxia-induced proliferation, invasion and chemoresistance of T-cell acute lymphoblastic leukemia cells, J. Hematol. Oncol. 05 (January (6)) (2013) 3.
[28] T.M. Horton, J.P. Perentesis, A.S. Gamis, T.A. Alonzo, R.B. Gerbing, J. Ballard, et al., A Phase 2 study of bortezomib combined with either idarubicin/cytarabine or cytarabine/etoposide in children with relapsed, refractory or secondary acute myeloid leukemia: a report from the Children's Oncology Group, Pediatr. Blood Cancer 61 (October (10)) (2014) 1754–1760.
[29] G. Visani, F. Ferrara, F. Di Raimondo, F. Loscocco, F. Fuligni, S. Paolini, et al., Low-dose lenalidomide plus cytarabine in very elderly, unfit acute myeloid leukemia patients: final result of a phase II study, Leuk. Res. 25 (September (62)) (2017) 77–83.
[30] R.M. Stone, S.J. Mandrekar, B.L. Sanford, K. Laumann, S. Geyer, C.D. Bloomfield, et al., Midostaurin plus chemotherapy for acute myeloid leukemia with a FLT3 mutation, N. Engl. J. Med. 377 (August (5)) (2017) 454–464.
[31] L.W.E. Drusbosky, S. Vali, T. Abbasi, A. Kumar, N.K. Singh, et al., A Prospective Clinical Trial to Predict Treatment Response Based on Mutanome-informed Computational Biology in Patients With AML and MDS, American Society of Hematology; 2016, San Diego, CA, 2016, p. 594 Blood.
[32] Pieter Van Vlierberghe, et al., ETV6 mutations in early immature human T cell leukemias, J. Exp. Med. 208.13 (2011) 2571–2579.
[33] Steven Goossens, et al., ZEB2 drives immature T-cell lymphoblastic leukaemia development via enhanced tumour-initiating potential and IL-7 receptor signalling, Nat. Commun. 6 (2015) 5794.
[34] Triona Ni Chonghaile, et al., Maturation stage of T-cell acute lymphoblastic leukemia determines BCL-2 versus BCL-XL dependence and sensitivity to ABT-199, Cancer Discov. (2014) CD-14.
[35] Krzysztof Mrózek, et al., Molecular cytogenetic characterization of the KG-1 and KG-1a acute myeloid leukemia cell lines by use of spectral karyotyping and fluorescence in situ hybridization, Genes Chromosomes Cancer 38.3 (2003) 249–252.