

To Bi or not to Bi: Acute Erythroid Leukemias and hematopoietic lineage choice

Cristina Di Genua,¹ Claus Nerlov¹

¹MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, OX3 9DS, UK

Abstract

Acute erythroid leukemia (AEL) is an acute leukemia characterized by erythroid lineage transformation. The WHO 2008 classification recognized two subtypes of AEL: bi-lineage erythroleukemia (erythroid/myeloid leukemia) and pure erythroid leukemia. In the updated 2016 WHO classification the erythroleukemia subtype was removed with around half of cases re-classified as myelodysplastic syndrome (MDS) and half as acute myeloid leukemia (AML). Diagnosis and classification are currently based on morphology using standard blast cutoffs, without integration of underlying genomic and other molecular features. Key outstanding questions are therefore whether AEL can be accurately diagnosed solely based on morphology or whether genetic or other molecular criteria should be included in its classification, and whether considering AEL as an entity distinct from AML and MDS is clinically relevant. We will here discuss recent work on the molecular basis of AEL, including the identification of mutations causative of AEL, and of transcriptional and epigenetic features that can be used to distinguish AEL from MDS and non-erythroid AML, and the prognostic value of these molecular features.

Morphological classification of acute erythroid leukemia

Acute erythroid leukemia (AEL) is a subtype of acute myeloid leukemia (AML), representing < 5% of all AML cases. It typically occurs in older patients, and is characterized by erythroblastic proliferation and differentiation arrest [1, 2]. Diagnosis and classification of AEL has been debated as it shares characteristics with both AML and myelodysplastic syndrome (MDS). The WHO 2008 classification defined two AEL subtypes: bi-lineage erythroleukemia (erythroid/myeloid leukemia – the most common form, accounting for 97% of AEL cases [3]) was defined by $\geq 50\%$ erythroid precursors in the bone marrow (BM), and $\geq 20\%$ myeloblasts in the non-erythroid BM component. Pure erythroid leukemia was defined by $\geq 80\%$ erythroid precursors in the BM [4]. AEL was classified as AML-M6 in the FAB classification [5]. Both AEL subtypes were considered as AML and therefore intensive chemotherapy was used as the curative treatment [6]. However, the non-erythroid compartment, where myeloid blasts are scored, generally constitutes only a minor fraction of the BM. Therefore, bi-lineage AEL generally presents with <20% myeloblasts in total BM, making the significance of the myeloid blast count difficult to interpret, with considerable implications for treatment selection [3]. Moreover, in the updated 2016 WHO criteria blast percentage calculation was changed from non-erythroid cells to all nucleated BM cells. This change meant many of the cases previously classified as AEL with a total myeloid blast count of <20% were re-classified as MDS, usually MDS with excess blasts. Cases with $\geq 50\%$ erythroid precursors and $\geq 20\%$ total myeloblasts were classified as AML with myelodysplasia-related changes (AML-MRC). Pure erythroid leukemia remained as an AML subtype in the WHO 2016 classification as previously defined by $\geq 80\%$ erythroid precursors in the BM [7]. However, for the purposes of this review we will define AEL according to the WHO 2008 classification, as this is the AEL definition most commonly used in the current literature.

A key challenge when using morphology in diagnosis is that pathologists follow different approaches when counting cells [8]. In addition, small variations in blast percentages may reclassify cases that are on the border between MDS and AML [9]. However, morphology will still play an important role in diagnosing AEL, as it allows detection of structural abnormalities, such as poikilocytosis and the presence of normoblasts in blood smears [1], that may be valuable for disease classification and prognosis. The development of standardized algorithms that detect and quantify cellular abnormalities and use of machine learning to integrate cell morphology with molecular modalities [10], is also likely to be important.

In terms of prognosis, retrospective studies showed that overall survival of AEL patients was comparable to that for MDS or AML-MRC, and survival related to cytogenetic risk group rather than to blast count or morphological dysplasia [11]. Therefore, for patients that can be classified as AEL, molecular rather than morphological features may offer the best prognostic value. Consistent with this, re-classification of erythroleukemia samples into MDS and AML did not show significant differences between mutation pattern or survival [12].

Molecular profiling of AEL

AEL is often under-represented in genomic profiling studies of AML and MDS. However, recently several studies have performed mutational and transcriptomic analysis to characterize and molecularly stratify AEL patients. The most frequent mutations identified across studies include *TP53* (31%), *NPM1* (14%) and *DNMT3A* (11%) (Table 1). Ping *et al.*, observed high prevalence of *CEBPA* and *GATA2* mutations in AEL, and that *GATA2* mutations were more frequent in AEL patients when compared to non-AEL AML, and showed a statistically significant association with *CEBPA* mutations [13]. Other genomic profiling studies observed a lower frequency of *CEBPA* and *GATA2* mutations. The difference in frequency of specific mutations may be due to differences in the patient cohorts analyzed. The cohort in the Ping *et al.*, study was younger with a median age of 50 compared to median age of 61 in the other studies [6, 12-15]. The ethnic background of the patient cohorts could also lead to a difference in mutation frequency. Previous studies have shown a higher frequency of *CEBPA* mutations in Eastern Asian compared to Caucasian cohorts even when excluding an age bias, which may explain the higher frequency of *CEBPA* mutation detected in the Ping *et al.*, study [16, 17]. Further studies would be needed to determine if *GATA2* mutation frequency differs between different populations.

Several studies created subgroups of AEL patients according to mutational association and exclusivity, with *TP53* mutant patients representing the largest subgroup. Other subgroups include *NPM1* mutant, *KMT2A* (also known as *MLL*) mutant/rearranged (11%), *DDX41* mutant (3%), *NUP98* rearranged (20%, specific to pediatric AEL), mutated in epigenetic regulator genes (33%), mutated in transcription factor genes (6%) and mutated in splicing/chromatin modifying genes (27%) [6, 12, 15]. Among the genomic subgroups *TP53* mutations were associated with poor prognosis, whereas *NPM1* mutations had a superior outcome with a 5-year survival of 88% [6].

Interestingly, Iacobucci *et al.*, identified age-related differences in the AEL mutational spectrum. Pediatric AEL patients more frequently had mutations in transcriptional regulators such as *WT1*. Patients <20 years old had frequent *RAS* mutations, whereas mutations associated with clonal hematopoiesis of indeterminant potential (CHIP; *DNMT3A* and *TET2*) and *NPM1* mutations were more frequent in 20-59–

year old patients, and *TP53* mutations were more frequent in patients >60 years old. They also identified a statistically significant difference between age matched AEL, MDS and non-erythroid AML. For example, childhood AEL patients had higher frequency of *FLT3* and *WT1* mutations and lower frequency of *GATA2* and *ASXL1* mutations compared to childhood MDS. When compared to non-erythroid childhood AML, childhood AEL had a higher frequency of mutations in RAS pathway and epigenetic regulators. Adult AEL was found to be an intermediate of adult MDS and non-erythroid AML in terms of genetic profile: Adult AEL had lower frequency of *FLT3* and *NPM1* mutations compared to non-erythroid adult AML but higher frequency than adult MDS, while MDS-associated genes such as *SF3B1* and *ASXL1* were less frequent in adult AEL, but more common than in non-erythroid adult AML [6].

Transcriptome analysis of AEL samples also identified distinct gene expression subgroups that were correlated with genomic lesions. For example, 55% of patients with *TP53* mutations fell within a gene expression subgroup characterized by over-expression of *LTF*, *DLK1*, and *MECOM*. Interestingly, all but one sample could be categorized using gene expression, whereas 45/130 (35%) of samples lacked a recurrent genetic alteration that allowed categorization by mutation, indicating that gene expression analysis has the potential to improve AEL diagnosis and patient stratification [6]. Furthermore, a study by Fagnan *et al.* showed that the transcriptional program of AEL patients differed from MDS and other AML subtypes, independent of molecular subgroups. Using erythroid vs. myeloid transcriptome-based trajectories they showed 78% of AEL cases mapped closer to the erythroid axis and 22% mapped closer to the myeloid trajectory [15]. Subgrouping by gene expression also had marked variation in patient outcome [6].

Precise classification of AEL is important in order to create a framework for diagnosis, risk stratification, and rational for testing targeted therapies. Such classification will likely require a multi-pronged approach that incorporates morphological evaluation, immunophenotyping by flow cytometry, and molecular testing at the genomic and transcriptomic level [11, 12, 18]. In particular, the mutation pattern in AEL is highly complex and heterogeneous with not all patients fitting into a discrete mutational subgroup, and gene expression could provide a more accurate method to subgroup patients. The molecular profiling of AEL samples is still limited and further studies, including single cell transcriptomics, proteomics and mutational analysis of serial patient samples, and of purified cell populations from the hematopoietic hierarchy of patient samples will deepen our understanding on this complex subtype.

Modeling AEL

Molecular characterization of AEL provides prognostic value. However, the complexity of the mutational spectrum in AEL makes it difficult to ascertain which mutations are causative of AEL, in particular as the mutations observed in AEL are also found in MDS and AML, raising the question of how erythroid-lineage transformation occurs. In addition, AEL most commonly involves both erythroid and myeloid lineage transformation, and the identity of the disease-propagating cell(s) remains unclear. Identifying the cells that sustain AEL, and well as the mutations that lead to their transformation, will be critical for understanding the molecular and cellular etiology of AEL.

Initial modeling of AEL involved virus-induced erythroleukemia produced by two retroviruses: the replication-defective spleen focus-forming virus (SFFV) and the replication-competent Friend murine leukemia virus (F-MuLV). SFFV initially induces uncontrolled proliferation of erythroid cells resulting in

erythroid hyperplasia, enlargement of the spleen and liver and Epo-independent differentiation, resulting in polycythemia. This effect is caused by the SFFV gp55 envelope glycoprotein, which interacts with the tyrosine kinase sf-Stk (short form of Stk, encoded by *Mst1r*) and Epo receptor, resulting in activation of the Ras/Raf/MAPK and PI-3K pathways [19, 20]. This is followed by erythroblast transformation due to SFFV integration into the host DNA, most commonly at the *Sfpi1* locus, leading to ectopic expression of the encoded myeloid transcription factor PU.1 [21]. In contrast, the most common integration site for F-MuLV was the *Fli1* gene, leading to Fli1 overexpression [22]. FLI1 and PU.1 both cross-antagonize erythroid transcription factors such as KLF-1, or GATA-1 and GATA-2 respectively, and thereby block erythroid differentiation program [23-25]. While these models do not directly reflect the genetic events observed in human AEL, they do highlight key molecular mechanisms that contribute to erythroid lineage transformation.

More recently, AEL modeling has been performed based on the genetic and transcriptional aberrations observed in human AEL patients [6, 15, 26, 27]. Iacobucci *et al.*, identified rare AEL cases (3/159 or 1.9% of those analyzed) with mutation in *NTRK1* which encodes the TrkA tyrosine kinase, two of which also contained *TP53* mutations. When virally expressed in murine hematopoietic stem and progenitor cells (HSPCs), mutant *NTRK1* and *TP53* co-operated to induce an aggressive leukemia, which was pure erythroid (expressing GATA-1, Ter119 and Glycophorin A, but not MPO) [6]. Another pure erythroid leukemia model was generated through overexpression of *Cdx4* in murine HSPCs. In human leukemia *CDX4* is highly expressed in AEL compared to non-erythroid AML, and *Cdx4*-overexpressing HSPCs induced a transplantable leukemia that infiltrated the spleen and liver. The vast majority of infiltrated cells were erythroid precursors, and were Ter119 positive, but a minority of cells were MPO positive, indicative of myeloid lineage involvement [26]. Finally, Fagnan *et al.*, identified that *TP53* mutant AEL samples was associated with aberrant expression of *ERG*, and showed that *TP53* R248Q mutation and *ERG* over-expression co-operated to induce a fatal erythroleukemia featuring erythroid and to a lesser extent myeloid cells in the BM (Table 2) [15].

While these models clearly implicate specific genomic and transcriptomic alteration observed in AEL tumors in erythroid leukemogenesis, viral overexpression does have the potential to induce phenotypes not replicated by somatic or germline mutation. Accurate genetic modeling also has the advantage that it allows the pre-leukemic effects of mutations to be characterized, and performs an unbiased interrogation of the hematopoietic hierarchy for the target cell(s) of transformation. Recently, Di Genua *et al.*, used knockin mutagenesis to model AEL induced by combined bi-allelic C/EBP α and GATA-2 zinc finger 1 (ZnF1) mutations, a combination of mutations that showed significant association across myeloid leukemias, and are observed in both pure myeloid (predominantly the M1 and M2 morphological subtypes) and erythroid (M6) leukemias [13, 28].

Bi-allelic C/EBP α mutation involves the combination of N-terminal and C-terminal mutations. N-terminal C/EBP α mutation involves the selective loss of the longer 42kDa translational isoform of the C/EBP α protein (p42), while retaining expression of the shorter 30kDa isoform (p30). In contrast, C-terminal mutations involve mutation of the C/EBP α basic region-leucine zipper domain, leading to loss of DNA binding of both p42 and p30 [29]. Both N- and C-terminal mutations disable C/EBP α -mediated cell cycle arrest [30], and in combination drive pre-leukemic hematopoietic stem cell (HSC) expansion. In addition,

combined N- and C-terminal terminal mutations alter the lineage programming of pre-leukemic HSCs, increasing erythroid and suppressing myeloid lineage gene expression (Figure 1) [31].

The majority of *GATA2* mutations that co-occur with bi-allelic *CEBPA* are found within ZnF1 of *GATA2*, distinct from the germline *GATA2* mutations within ZnF2 that are found in autosomal dominant inherited hematological malignancies such as Emberger syndrome [32-34], MonoMac syndrome [35, 36] and dendritic cell, monocyte, B and natural killer, lymphoid (DCML) deficiency syndrome [37, 38]. While *GATA2* ZnF2 mutations are associated with pre-disposition to MDS and AML, little is known about the leukemogenic effects of *GATA2* ZnF1 mutations.

Modeling of bi-allelic C/EBP α mutant [31, 39] and C/EBP α + FLT3-ITD mutant AML [40] by knockin mutagenesis of the mutations seen in human AML resulted in the generation of pure myeloid acute leukemias, but not AEL. However, combining bi-allelic C/EBP α knockin mutations with the *GATA-2* ZnF1 G320D mutation, a point mutation observed in multiple studies of AML and AEL [13, 28], ca. 40% of leukemias displayed bi-lineage neutrophil-erythroid phenotype (Table 2) [27]. Importantly, the bi-lineage leukemia phenotype was transplantable by transformed leukemia-initiating cells (LIC) with a neutrophil-monocyte progenitor immunophenotype (NMPs), which were bi-potent and generated both myeloid and erythroid blasts at the single cell level. Analysis of human AEL samples identified, in addition to myeloid and erythroid blasts, a Lin[−]CD34⁺CD38⁺CD71[−]CD235a[−]CD117⁺ myeloid progenitor population with a leukemic stem cell signature, making it a candidate human AEL LIC [27]. Therefore, the erythroid and myeloid components of bi-lineage AEL generated by mutant C/EBP α and *GATA-2* is sustained by a single bi-potent leukemia-propagating cell population, and the cellular structure of the murine AEL model is analogous to human AEL.

The neutrophil and erythroid lineages develop through distinct progenitor pathways [41], and therefore no normal oligo-potent progenitor exists that generates both erythroid and neutrophil output. As mentioned above, combined N- and C-terminal C/EBP α mutations lead to ectopic expression of erythroid genes in HSCs [31]. This ectopic transcriptional lineage programming was sustained in pre-leukemic bi-allelic C/EBP α mutant phenotypic NMPs, including the increased expression of erythroid-lineage transcription factors (TFs; *GATA-1*, *GATA-2*, *FOG-1*, *KLF1*), but was insufficient to promote erythroid lineage transformation [27]. However, in the additional presence of ZnF1 mutant *GATA-2* the chromatin access of erythroid TFs was increased and access of myeloid-lineage TFs (C/EBP, PU.1) decreased in pre-leukemic NMPs, conferring ectopic erythroid lineage potential on the C/EBP α + *GATA-2* mutant NMPs (Figure 1).

Bi-allelic *CEBPA* and *GATA2* ZnF1 mutations co-occur in both human AEL and non-erythroid AML, and both leukemia type was observed in C/EBP α + *GATA-2* mutant murine leukemias. Comparing the AEL- and AML leukemic NMPs did not identify major transcriptional differences. However, the altered chromatin accessibility state observed in pre-leukemic C/EBP α + *GATA-2* mutant NMPs was conserved in the AEL leukemic NMP. Therefore, epigenetic, rather than transcriptional, profiling was key to distinguishing AEL from pure myeloid AML in this scenario.

Other combinatorial knockin mouse models will be required to determine if the alterations in gene expression and chromatin accessibility in the NMP that develop in C/EBP α + GATA-2 mutant model are shared with other models of AEL.

Finally, cell lines such as F-36P, HEL and OCI-M2, derived from AEL patients, express markers of multiple lineages [3]. These cell lines have been shown to be *TP53*-mutated and have abnormal copy number alterations in *EPOR* and *ERG*, which were also detected in a *TP53*-mutated AEL patient sample [42]. This substantiates the importance of *TP53* mutations and alterations in *EPOR*, *ERG*, *PU.1* and *FLI1* in erythroid transformation, and indicates that such cell lines can act as preclinical models of specific AEL genetic subtypes.

Future perspectives

Accurately defining AEL as a separate and discrete disease entity remains a challenge since the driver mutations found in AEL tumors overlap with those found in MDS and pure myeloid AML, and morphology does not unequivocally distinguish AEL from MDS. However, the recent findings discussed above provide evidence that both epigenetic and transcriptional features can be used to classify AEL.

In particular, epigenetic changes may be an important step in the early stage of AEL development. In the presence of bi-allelic C/EBP α mutations, the GATA-2 ZnF1 mutation acts as a 'non-canonical' chromatin regulator by altering the accessibility of lineage-specific TF motifs controlling the phenotype of the resulting leukemia. In addition, mutations in the epigenetic modifier *DNMT3A* are one of the most common mutations identified in AEL patients, and loss of *Dnmt3a* in mice has been shown to induce a spectrum of hematopoietic malignancies including AML, MDS, T- and B-cell acute lymphocytic leukemia. Notably, loss of *DNMT3A* promoted lineage-specific methylation alterations that were distinct in lymphoid and myeloid diseases [43]. Together, these observations support the notion that mutations in chromatin modifiers can control the lineage identity of the resulting leukemia, and that the epigenetic signatures generated can be used to distinguish AEL from other disease entities.

Understanding the mutational and transcriptional landscape of AEL, combined with an accurate model of AEL can potentially guide treatment. 45% of AEL cases contained mutations in at least one signaling pathway that could be targeted by a tyrosine kinase/JAK2/RAS inhibitor. This was demonstrated by murine leukemias co-mutated with *NTRK1* and *TP53* were highly sensitive to the TRK inhibitor Larotrectinib *in vivo* [6]. *TP53* mutations in AML and AEL have been shown to have a high response rate to hypomethylating agents [44, 45]. *DNMT3A* mutations have also been linked to an increased response to hypomethylating agents [46], therefore *TP53* and *DNMT3A* mutant AELs could benefit from treatment with hypomethylating agents. The C/EBP α + GATA-2 mutant model fully recapitulates bi-lineage AEL, providing a pre-clinical model that could be used for drug screening, and in addition identified a specific leukemia-initiating progenitor cell population capable of propagating bi-lineage AEL. Further functional analysis of the C/EBP α + GATA-2 mutant AEL model could lead to identification and therapeutic targeting of human AEL-propagating cells.

In summary, analysis of the genomic and transcriptomic landscape of human AEL shows that AEL is genetically an intermediate between MDS and AML. However, the majority of AEL cases can be sub-grouped by gene expression, potentially providing a discrete classification not achievable using

morphology or mutation analysis alone. Integrated analysis of genomic and gene expression information should therefore be considered, in addition to morphological criteria, when diagnosing and classifying AEL. The clinical value of an independent AEL classification will have to await prospective studies where tailored treatments, based on the specific molecular characteristics of AEL, are tested against the current standard-of-care.

Acknowledgements

We thank Dr Paresh Vyas for helpful comments on the manuscript. CdG was supported by an MRC Studentship, and the work described was supported by a Bloodwise Project Grant to CN and and MRC Unit Program Grant to CN.

1. Zuo, Z., et al., *Acute erythroid leukemia*. Arch Pathol Lab Med, 2010. **134**(9): p. 1261-70.
2. Liu, W., et al., *Pure erythroid leukemia: a reassessment of the entity using the 2008 World Health Organization classification*. Mod Pathol, 2011. **24**(3): p. 375-83.
3. Boddu, P., et al., *Erythroleukemia-historical perspectives and recent advances in diagnosis and management*. Blood Rev, 2018. **32**(2): p. 96-105.
4. Vardiman, J.W., et al., *The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes*. Blood, 2009. **114**(5): p. 937-51.
5. Bennett, J.M., et al., *Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group*. Br J Haematol, 1976. **33**(4): p. 451-8.
6. Iacobucci, I., et al., *Genomic subtyping and therapeutic targeting of acute erythroleukemia*. Nat Genet, 2019. **51**(4): p. 694-704.
7. Arber, D.A., et al., *The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia*. Blood, 2016. **127**(20): p. 2391-405.
8. Kasyan, A., et al., *Acute erythroid leukemia as defined in the World Health Organization classification is a rare and pathogenetically heterogeneous disease*. Mod Pathol, 2010. **23**(8): p. 1113-26.
9. Wang, S.A. and R.P. Hasserjian, *Acute Erythroleukemias, Acute Megakaryoblastic Leukemias, and Reactive Mimics: A Guide to a Number of Perplexing Entities*. Am J Clin Pathol, 2015. **144**(1): p. 44-60.
10. Maria, I.J., T. Devi, and D. Ravi, *Machine Learning Algorithms For Diagnosis Of Leukemia*. International Journal of Scientific & Technology Research, 2020. **9**(1): p. 267-270.
11. Hasserjian, R.P., et al., *Acute erythroid leukemia: a reassessment using criteria refined in the 2008 WHO classification*. Blood, 2010. **115**(10): p. 1985-92.
12. Cervera, N., et al., *Revisiting gene mutations and prognosis of ex-M6a-acute erythroid leukemia with regard to the new WHO classification*. Blood Cancer J, 2017. **7**(8): p. e594.
13. Ping, N., et al., *Exome sequencing identifies highly recurrent somatic GATA2 and CEBPA mutations in acute erythroid leukemia*. Leukemia, 2017. **31**(1): p. 195-202.
14. Grossmann, V., et al., *Acute erythroid leukemia (AEL) can be separated into distinct prognostic subsets based on cytogenetic and molecular genetic characteristics*. Leukemia, 2013. **27**(9): p. 1940-3.

15. Fagnan, A., et al., *Human erythroleukemia genetics and transcriptomes identify master transcription factors as functional disease drivers*. Blood, 2020. **136**(6): p. 698-714.
16. Su, L., et al., *CEBPA mutations in patients with de novo acute myeloid leukemia: data analysis in a Chinese population*. Onco Targets Ther, 2016. **9**: p. 3399-403.
17. Wei, H., et al., *Distinct genetic alteration profiles of acute myeloid leukemia between Caucasian and Eastern Asian population*. J Hematol Oncol, 2018. **11**(1): p. 18.
18. Qiu, S., et al., *An analysis of 97 previously diagnosed de novo adult acute erythroid leukemia patients following the 2016 revision to World Health Organization classification*. BMC Cancer, 2017. **17**(1): p. 534.
19. Finkelstein, L.D., et al., *Sf-Stk kinase activity and the Grb2 binding site are required for Epo-independent growth of primary erythroblasts infected with Friend virus*. Oncogene, 2002. **21**(22): p. 3562-70.
20. Nishigaki, K., et al., *Erythroid cells rendered erythropoietin independent by infection with Friend spleen focus-forming virus show constitutive activation of phosphatidylinositol 3-kinase and Akt kinase: involvement of insulin receptor substrate-related adapter proteins*. J Virol, 2000. **74**(7): p. 3037-45.
21. Moreau-Gachelin, F., A. Tavitian, and P. Tambourin, *Spi-1 is a putative oncogene in virally induced murine erythroleukaemias*. Nature, 1988. **331**(6153): p. 277-80.
22. Ben-David, Y., E.B. Giddens, and A. Bernstein, *Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus*. Proc Natl Acad Sci U S A, 1990. **87**(4): p. 1332-6.
23. Nerlov, C. and T. Graf, *PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors*. Genes Dev, 1998. **12**(15): p. 2403-12.
24. Schuetze, S., et al., *Role of the PU.1 transcription factor in controlling differentiation of Friend erythroleukemia cells*. Mol Cell Biol, 1992. **12**(7): p. 2967-75.
25. Starck, J., et al., *Spi-1/PU.1 is a positive regulator of the Fli-1 gene involved in inhibition of erythroid differentiation in friend erythroleukemic cell lines*. Mol Cell Biol, 1999. **19**(1): p. 121-35.
26. Thoene, S., et al., *The ParaHox gene Cdx4 induces acute erythroid leukemia in mice*. Blood Adv, 2019. **3**(22): p. 3729-3739.
27. Di Genua, C., et al., *C/EBPalpha and GATA-2 Mutations Induce Bilineage Acute Erythroid Leukemia through Transformation of a Neomorphic Neutrophil-Erythroid Progenitor*. Cancer Cell, 2020. **37**(5): p. 690-704 e8.
28. Fasan, A., et al., *GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis*. Leukemia, 2013. **27**(2): p. 482-5.
29. Nerlov, C., *C/EBPalpha mutations in acute myeloid leukaemias*. Nat Rev Cancer, 2004. **4**(5): p. 394-400.
30. Lopez, R.G., et al., *C/EBPalpha and beta couple interfollicular keratinocyte proliferation arrest to commitment and terminal differentiation*. Nat Cell Biol, 2009. **11**(10): p. 1181-90.
31. Bereshchenko, O., et al., *Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML*. Cancer Cell, 2009. **16**(5): p. 390-400.

32. Ostergaard, P., et al., *Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome)*. Nat Genet, 2011. **43**(10): p. 929-31.
33. Mansour, S., et al., *Emberger syndrome-primary lymphedema with myelodysplasia: report of seven new cases*. Am J Med Genet A, 2010. **152A**(9): p. 2287-96.
34. Kazenwadel, J., et al., *Loss-of-function germline GATA2 mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature*. Blood, 2012. **119**(5): p. 1283-91.
35. Hsu, A.P., et al., *Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome*. Blood, 2011. **118**(10): p. 2653-5.
36. Vinh, D.C., et al., *Autosomal dominant and sporadic monocytopenia with susceptibility to mycobacteria, fungi, papillomaviruses, and myelodysplasia*. Blood, 2010. **115**(8): p. 1519-29.
37. Bigley, V., et al., *The human syndrome of dendritic cell, monocyte, B and NK lymphoid deficiency*. J Exp Med, 2011. **208**(2): p. 227-34.
38. Dickinson, R.E., et al., *Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency*. Blood, 2011. **118**(10): p. 2656-8.
39. Kirstetter, P., et al., *Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells*. Cancer Cell, 2008. **13**(4): p. 299-310.
40. Reckzeh, K., et al., *Molecular and cellular effects of oncogene cooperation in a genetically accurate AML mouse model*. Leukemia, 2012. **26**(7): p. 1527-36.
41. Drissen, R., et al., *Identification of two distinct pathways of human myelopoiesis*. Sci Immunol, 2019. **4**(35).
42. Adelaide, J., et al., *Gains of EPOR and ERG genes in adult erythroleukaemia*. Br J Haematol, 2020. **189**(4): p. e174-e177.
43. Mayle, A., et al., *Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation*. Blood, 2015. **125**(4): p. 629-38.
44. Welch, J.S., et al., *TP53 and Decitabine in Acute Myeloid Leukemia and Myelodysplastic Syndromes*. N Engl J Med, 2016. **375**(21): p. 2023-2036.
45. Almeida, A.M., et al., *Clinical Outcomes of 217 Patients with Acute Erythroleukemia According to Treatment Type and Line: A Retrospective Multinational Study*. Int J Mol Sci, 2017. **18**(4).
46. Coombs, C.C., et al., *Mutational correlates of response to hypomethylating agent therapy in acute myeloid leukemia*. Haematologica, 2016. **101**(11): p. e457-e460.

Table 1: Mutations identified across studies

Gene	Number of patients with the mutation	Total number of patients analyzed	Percentage of patients with the mutation	Reference
<i>TP53</i>	121	391	31	[6, 12-15]
<i>NCOR2</i>	6	31	19	[15]
<i>NPM1</i>	55	391	14	[6, 12-15]
<i>ETV6</i>	26	210	12	[6, 12]
<i>TET2</i>	36	299	12	[6, 12, 13, 15]
<i>DNMT3A</i>	44	391	11	[6, 12-15]
<i>WT1</i>	27	278	10	[6, 12, 15]
<i>CD36</i>	3	31	10	[15]
<i>NOTCH2</i>	3	31	10	[15]
<i>SEC31A</i>	3	31	10	[15]
<i>RUNX1</i>	37	391	9	[6, 12-15]
<i>NF1</i>	15	159	9	[6]
<i>PTPN11</i>	19	210	9	[6, 12]
<i>KMT2A</i>	27	300	9	[6, 12, 14]
<i>ASXL1</i>	31	387	8	[6, 12-15]
<i>SMC3</i>	4	51	8	[12]
<i>U2AF1</i>	17	217	8	[6, 13]
<i>RB1</i>	12	159	8	[6]
<i>CEBPA</i>	26	388	7	[6, 12-15]

<i>STAG2</i>	16	241	7	[6, 12, 15]
<i>ABL1</i>	2	31	6	[15]
<i>CAD</i>	2	31	6	[15]
<i>CLASP2</i>	2	31	6	[15]
<i>EP300</i>	2	31	6	[15]
<i>MLH1</i>	2	31	6	[15]
<i>MUC20</i>	2	31	6	[15]
<i>MYH10</i>	2	31	6	[15]
<i>POLE</i>	2	31	6	[15]
<i>IDH2</i>	21	333	6	[6, 13-15]
<i>GNAS</i>	3	51	6	[12]
<i>FLT3</i>	19	333	6	[6, 13-15]
<i>SRSF2</i>	5	89	6	[13, 15]
<i>KDM5A</i>	8	159	5	[6]
<i>PDE4DIP</i>	8	159	5	[6]
<i>GATA2</i>	13	268	5	[12, 13]
<i>NRAS</i>	17	359	5	[6, 12-14]

Mutations identified across studies that are present in $\geq 5\%$ of AEL patients.

Table 2: Models of AEL

Gene	Model type	Leukemia phenotype	Reference
<i>Sfpi1</i> overexpression	SFFV retrovirus	Pure erythroid	[21]
<i>Fli1</i> overexpression	F-MuLV retrovirus	Pure erythroid	[22]
NTRK1 mutation and TP53 mutation	Viral overexpression	Pure erythroid	[6]
<i>Cdx4</i> overexpression	Viral overexpression	Predominately erythroid with minority myeloid cells	[26]
TP53 mutation and <i>ERG</i> overexpression	Viral overexpression	Predominately erythroid with minority myeloid cells	[15]
Bi-allelic C/EBP α mutations	Knockin	Pure myeloid	[31]
Bi-allelic C/EBP α mutations and FLT3-ITD mutation	Knockin	Pure myeloid	[40]
Bi-allelic C/EBP α mutations and GATA-2 mutation	Knockin	Pure myeloid and bi-lineage leukemia	[27]

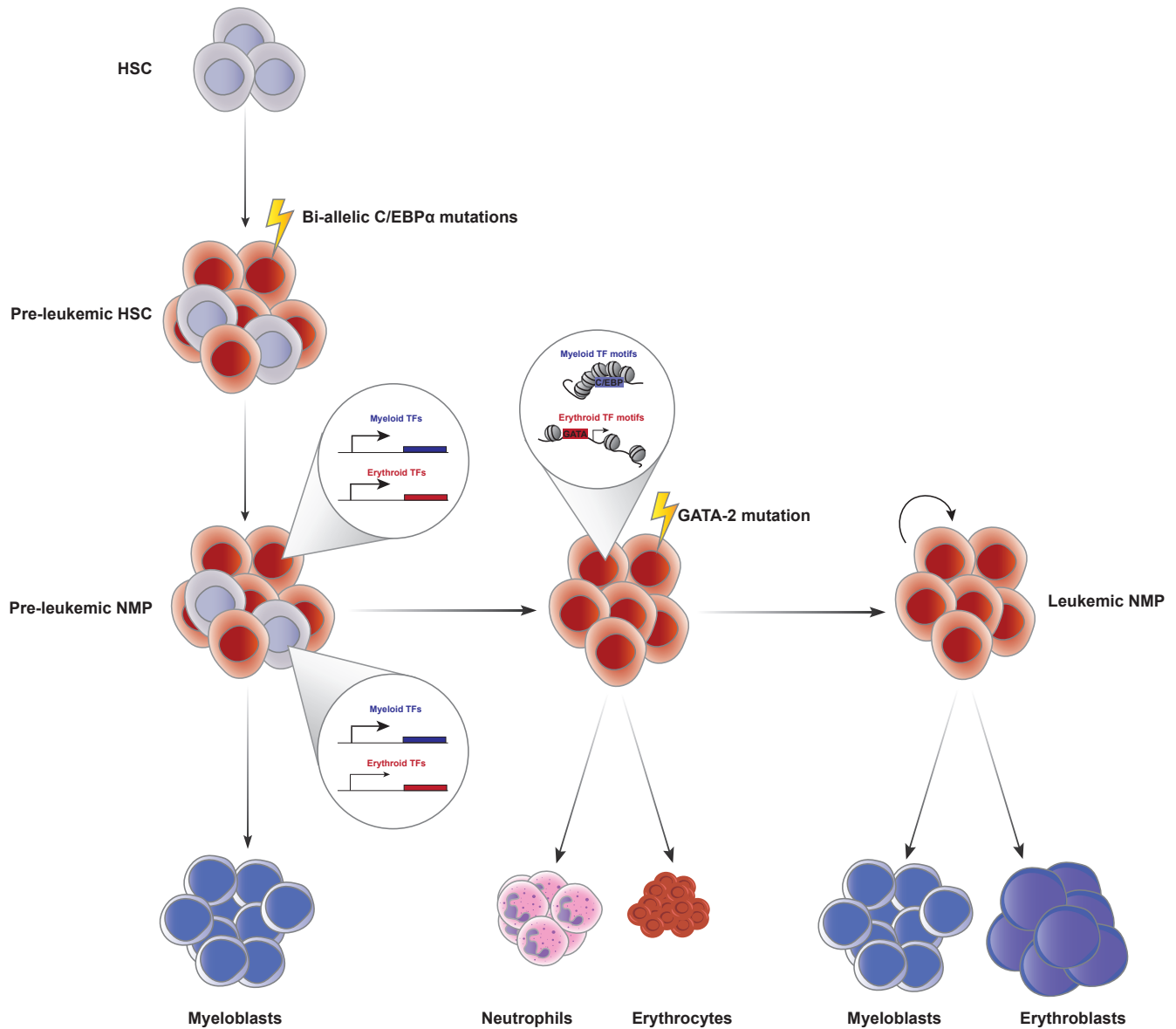


Figure 1. Schematic of the AEL development. Pre-leukemic HSCs carrying the bi-allelic C/EBP α mutations (red cell) have a competitive advantage over non-mutated HSCs (blue cell) and increases the number of committed progenitors. The bi-allelic C/EBP α mutations then lead to an increased expression of erythroid transcription factors genes in the pre-leukemic NMPs. The GATA-2 mutation adds an additional regulatory layer by increasing the chromatin accessibility of erythroid transcription factor motifs and decrease in the accessibility of myeloid transcription factor motifs within the pre-leukemic NMP. This creates a neomorphic neutrophil-erythroid progenitor that can produce immature myeloid and erythroid cells. This chromatin state is preserved upon transformation that creates a bi-potent LIC that sustains the bi-lineage leukemia, producing both the myeloblasts and erythroblasts. The erythroid-permissive chromatin state caused by the GATA-2 mutation is essential for AEL development as only bi-allelic C/EBP α + GATA2 mutant mice and not bi-allelic C/EBP α mutant mice develop an AEL.