Alma Mater Studiorum – Università di Bologna

Dipartimento di Medicina Specialistica, Diagnostica e Sperimentale

Dottorato di Ricerca in Oncologia, Ematologia e Patologia

XXX Ciclo

Settore Scientifico Disciplinare: MED/15

Settore Concorsuale:06/D3

Molecular characterization of acute myeloid leukemia by Next Generation Sequencing: identification of novel biomarkers and targets of personalized therapies

Presentata da: **Antonella Padella**

Coordinatore Prof. Pier-Luigi Lollini

Supervisore: Prof. Giovanni Martinelli

Esame finale anno 2018

Abstract

Acute myeloid leukemia (AML) is a hematopoietic neoplasm that affects myeloid progenitor cells and it is one of the malignancies best studied by next generation sequencing (NGS), showing a highly heterogeneous genetic background. The aim of the study was to characterize the molecular landscape of 2 subgroups of AML patients carrying either chromosomal number alterations (i.e. aneuploidy) or rare fusion genes. We performed whole exome sequencing and we integrated the mutational data with transcriptomic and copy number analysis. We identified the cell cycle, the protein degradation, response to reactive oxygen species, energy metabolism and biosynthetic process as the pathways mostly targeted by alterations in aneuploid AML. Moreover, we identified a 3-gene expression signature including *RAD50, PLK1* and *CDC20* that characterize this subgroup.

Taking advantage of RNA sequencing we aimed at the discovery of novel and rare gene fusions. We detected 9 rare chimeric transcripts, of which partner genes were transcription factors (*ZEB2*, *BCL11B* and *MAFK*) or tumor suppressors (*SAV1* and *PUF60*) rarely translocated across cancer types. Moreover, we detected cryptic events hiding the loss of *NF1* and *WT1*, two recurrently altered genes in AML. Finally, we explored the oncogenic potential of the *ZEB2-BCL11B* fusion, which revealed no transforming ability in vitro. However, further studies may elucidate its role in AML.

Taken together, our results highlight the need for a deep molecular characterization of AML heterogeneity and identified potential biomarkers and targets for personalized therapies. Further studies will elucidate the role of these markers as drivers of leukemogenesis, prognostic factors and predictors of therapeutic response.

Table of Contents

Introduction

Acute myeloid leukemia

Acute myeloid leukemia (AML) originate from the defective regulation of the differentiation and self-renewing programs of multipotent hematopoietic stem cells, resulting in the bone marrow (BM) expansion of myeloid precursors, with limited or abnormal differentiation capacity. The etiological classification of AML includes

- De novo or primary AML, which arises in patients not exposed to risk factors;
- AML "secondary to leukemogenic agent exposure";
- AML "secondary to myelodysplastic syndromes".

AML is the most common type of leukemia among adult patients and the incidence is around 4 cases per 100,000 of population, giving a predicted Europe-wide incidence of around 30,000 cases per year, with a poor 5-year survival of less than 30%. The incidence increases sharply with age (mean age at diagnosis 67 years), and with population aging, it is likely to rise in the future. A further increase is expected from the rising incidence of therapy-related myeloid neoplasms (i.e. myelodysplastic syndromes or AML occurring in cancer survivors after successful treatment of a primary tumor).

The French-American-British (FAB) classification divides AML into eight subtypes, based on cell morphology defined by cytologic and cytochemical analyses (Table 1).

FAB	Definition	Cytogenetics
M0	AML, minimally differentiated	
M1	AML, without maturation	
M2	AML, with granulocytic maturation	t(8;21)(q22;q22)t(6;9)
M3	Acute Promyelocitic Leukemia (APL)	t(15;17)
M4	Acute myelomonocitic leukemia	inv(16)(p13q22),del(16q)
M4eos	Myelomonocitic with BM eosinophilia	inv(16),t(16;16)
M5	Acute monoblastic leukemia (M5a) or monocitic leukemia (M5b)	del(11q),t(9;11),t(11;19)
M6	Acute erythroleukemia (M6a) or rare erythroid leukemia (M6b)	
M7	Acute megakaryoblatic leukemia	t(1;22)

Table 1. FAB classification of AML.

In 2016, the World Health Organization provided an updated classification system incorporating morphology, cytogenetics, molecular genetics and immunological markers[1].

In the last few years, the development of Next Generation Sequencing (NGS) technologies for high-resolution analysis of cancer genome has dramatically improved our understanding of

AML pathogenesis, showing that a number of genetic hits participate to the malignant transformation of hematopoietic stem-progenitor cells. It has been demonstrated that the landscape of somatic alterations is the results of a relative small number of "driver mutations" which typically occur with "passenger mutations", thus contributing to the mutational spectrum of genetic variation of leukemic cells.

By analyzing the genome of 200 AML patients, 9 functional categories of significantly mutated genes and their distinct patterns of cooperativity and mutual exclusivity, have been defined (transcription factor gene fusions, *NPM1*, tumor suppressor genes, two groups of epigenetic modifier genes, signaling genes, myeloid transcription factor genes, cohesin- complex genes and spliceosome-complex genes; Figure 1)[2,3]. An average of 5 recurrently mutated genes and 1.5 gene-fusion event per case were identified. Most patients were characterized by clonal heterogeneity at the time of diagnosis, with the presence of both a founding clone and at least one subclone.

Clonal evolution studies on AML demonstrated that genes involved in the epigenetic regulation such as *DNMT3A*, *ASXL1*, *IDH2*, and *TET2* were present in the pre-leukemic clone and persisted during remission, leading to relapse[4]. In addition, two independent studies showed that clonal hematopoiesis occur in healthy individual with somatic mutations involving the same genes (*DNMT3A*, *TET2*, and *ASXL1*), increase as people age and were associated with an increased risk of hematologic cancer and in all-cause mortlity[5,6].

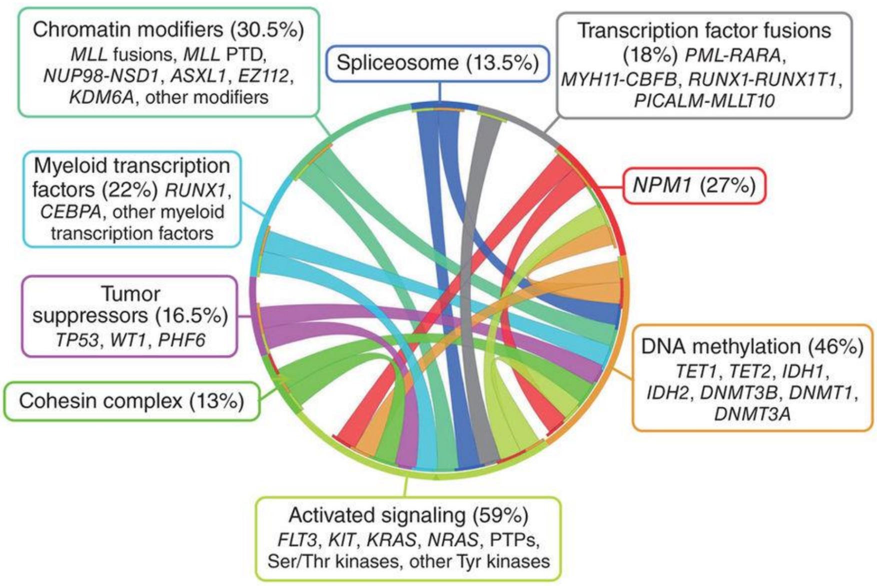


Figure 1. Circos plot showing the 9 functional categories and their pattern of co-occurrence and mutual exclusivity. Ribbons connecting different categories reflect the co-occurency of alterations in genes involved in that pathways. PTPs, protein tyrosine phosphatases; *MLL* PTD, *MLL* partial tandem duplication[3].

Recently, a study has provided a more detailed genomic classification of AML and the relative correlation with prognosis, identifying 11 subgroups of patients with different patterns of genomic lesions. However, two additional groups were either without a molecular driver (4%) or with a driver not falling into a class defining-lesion (11%, Figure 2)[7]. Papaemmanuil and colleagues shed light into the different roads that lead to AML and how the specific path from normal hematopoietic cell to leukemia has important biologic and clinical implications. The clinical consequence of this molecular-based classification led to the conclusion that *NPM1* and *CEBPA*^{biallelic} mutations represent the category which confers favorable prognosis, *TP53*-complex karyotype subgroup has adverse outcome; patients with mutations in the chromatin-spliceosome category are usually older, with lower leucocytes and blast counts, lower responsivity to induction chemotherapy, high probability to relapse and a short overall survival. The latter group of patients would be considered at intermediate risk but their diseases behave as adverse outcome leukemias. Then, *IDH*^{R172} mutation has a really a low frequency in patients but responsible for an outcome similar to *NPM1*-mutated AML. Taken together the deep molecular characterization of AML carried out in the last years helped to re-classify the disease and re-define the diagnosis procedures and the management of patients (Figure 3), as outlined by the 2017 European Leukemia Net (ELN) recommendations[1].

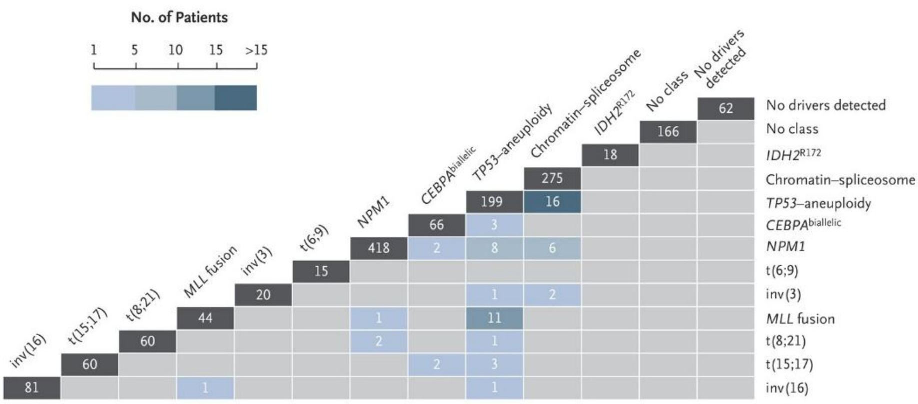


Figure 2. AML molecular subgroups. Patients distribution and intersection across molecular types identified by Papaemmanuil et al[7]. The numbers on the first row of each column represent patients belonging only to the respective class. The numbers along the columns represent patients meeting criteria for more than one subgroup.

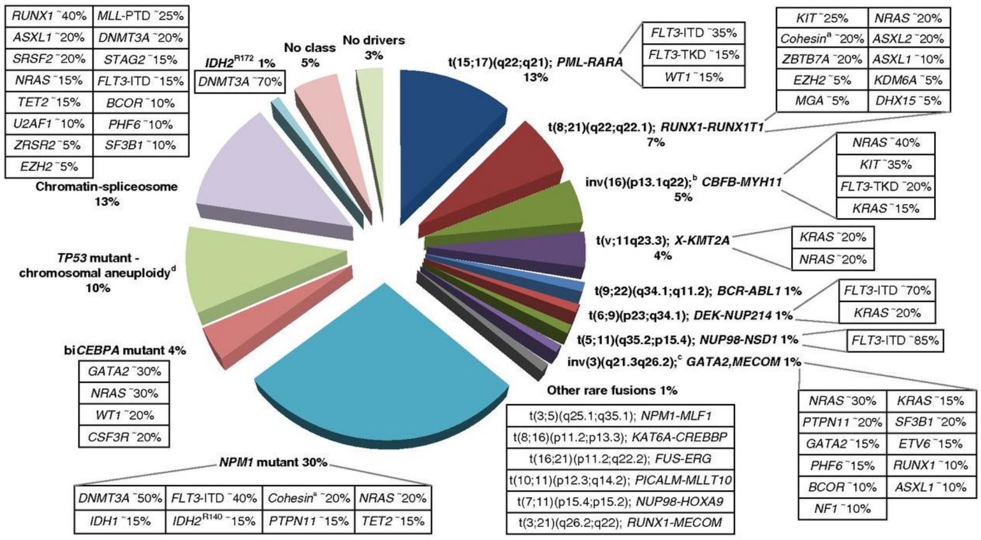


Figure 3. Molecular classes of AML according to 2017 ELN recommendations[1]. For each molecular subgroup, the frequency across AML, and co-occurring mutations with their relative frequency are shown in the boxes.

Even if the genetic stratification of AML patients has been updated, it still relies on few molecular markers and, most importantly, how this deep molecular characterization of patients may be clinically actionable is still unknown[1].

Few innovative therapeutic concepts were effectively translated into the clinical practice and despite some first line therapy success, the prognosis remains poor for a considerable number of cases, which either do not respond to therapy or become incurable when relapsing due to clonal evolution and to the failure of current therapeutic strategies to eradicate the leukemia stem cells. Indeed, chemotherapies have reached their plateau in cure rates and survival in hematology. Optimal treatment is inpatient-based, highly-toxic and very expensive, involving multiple courses of combination chemotherapy and stem cell transplantation. This therapeutic approach cures less than 50% of patients under 60 years of age. The outcome of older patients is even poorer, in particular for those who are considered unfit for intensive chemotherapy. Long-term survival is a dismal 10-20%.

A key problem is how to address the right therapy to any individual patient. Approaches with novel drugs (i.e. hypomethylating agents, monoclonal antibodies, molecular target drugs) are failing in drastically augmenting cure rates and overall survival in general hemato-oncology population. The most recent approaches for personalized therapies are aimed at tailoring clinical trials to patients' specific genomic background and response rates are lower than expected. The employment of a single drug to tailor a given mutation (supposed to be a driver mutation) might not be the correct approach. The advent of both powerful methods for patient characterization (such as genomics, proteomics, metabolomics and drug response assays), and computational tools for analyzing large sets of data has dramatically improved the knowledge of hematological malignancies and has boosted the development of large-scale databases. However, the potentiality of omics data remains largely unexploited and the lack of a multidimensional analysis is reflected in the insufficient characterization of hematological diseases and poor stratification of patients.

Under review (Cancer Research)

Aneuploidy: cancer strength or vulnerability?

Giorgia Simonetti[1], Samantha Bruno[1], Antonella Padella[1], Elena Tenti[1], Giovanni Martinelli[1]

[1] Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy

Running title. Aneuploidy and cancer.

Keywords. Tumor promotion and progression, cell cycle checkpoints, genotype-phenptype correlations, aneuploidy, anti-cancer therapies.

Financial support. Supported by the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n° 306242-NGS-PTL. The study was supported by Associazione Italiana per la Ricerca sul Cancro (AIRC) Investigator grant, n. 19226 to Giovanni Martinelli, Investigator Grant n. 15762.

Corresponding author. Giorgia Simonetti, via Massarenti, 9 - 40138 Bologna, Italy. Phone: 0039 0512143791; fax: 0039 0512144037; e-mail: giorgia.simonetti3@unibo.it.

Abstract

Aneuploidy causes a proliferative disadvantage, mitotic and proteotoxic stress in non-malignant cells and has been associated with defects in the spindle assembly checkpoint (SAC). Aneuploidy is also the hallmark of cancer and evidence from mouse models suggests a complex relationship between chromosome number alterations, SAC genes and tumor susceptibility. We here discuss the oncogenic and tumor suppressor functions of aneuploidy, which is affected by the genomic and environmental background, and on its therapeutic potential. The genome-destabilizer effect induced by the aneuploid condition, driving an increased adaptive capacity, coupled with the stem-cell like quiescent state and the immune escape potential is the strength of aneuploid cancer. However, chromosome instability, mitotic defects and aneuploidy-tolerating mechanisms can be suitable targets for *ad hoc* therapeutic strategies taking into account synthetic lethal combinations.

Aneuploidy: a normal and abnormal condition

Normal human diploid cells contain 23 pairs of chromosome (44 autosomes and 2 sex chromosomes). In some circumstances, the chromosome number is altered, a condition known as aneuploidy. Aneuploidy is physiological during cellular development in a tissue-specific way[8], likely due to its contribution to cellular diversity, that provides a selective advantage in response to injuries. Hepatocytes can develop as polyploid cells and then undergo reductive division leading to massive chromosome loss to near-diploid cells. Aneuploid mitotic cells and post-mitotic neurons have been also detected in normal murine and human brain. Moreover, aneuploidy associates with aging and age-related disorders in different tissues. The frequency of chromosome segregation errors in meiosis I increases in oocytes with increasing maternal age[9] and the aneuploid condition may favour neurodegeration during aging[8]. Down syndrome individuals frequently develop Alzheimer's disease by the age of 40 and normal patients affected by this neurodegenerative disorder showed an increased number of cells carrying trisomy of chromosomes 21 or 17, which locate many susceptibility genes.

The rate of constitutive aneuploidy matters in terms of beneficial and detrimental effects. Low levels of aneuploidy can be tolerated or even provide an advantage under specific conditions in non-malignant tissues[10], while increased rates of aneuploidy can become pathogenic, as observed in neurodegenerative diseases and in cancer[8]. This phenomenon is of particular interest since 1914, when Theodor Boveri proposed that an unfitting chromosome number can promote cancer. Over the past 100 years, a number of studies investigated the cellular and molecular events that cause aneuploidy and studied its potential involvement in cancer development. We here speculate on the complex relationship between aneuploidy and cancer, including its oncogenic and tumor suppressor properties and its therapeutic potentials.

Origin and molecular players of aneuploidy

Eukaryotic cells have developed sophisticated control systems to ensure a correct cellular division, called cell cycle checkpoints, which consists of a family of proteins regulating progression through the different phases. The three checkpoints are: (i) restriction point, acting at the end of G1 phase to promote DNA replication and entry into S phase, when external and internal cellular conditions are favorable; (ii) G2/M checkpoint, also known as the DNA damage checkpoint, which ensures the fidelity of the DNA replication process before the cell starts to divide; (iii) metaphase checkpoint before transition to anaphase, allowing mitosis to be completed once all chromosome are properly aligned at the equator and correctly connected with the spindle (Figure 4). The metaphase checkpoint, also known as mitotic checkpoint, is relevant to the faithful segregation of chromosomes, thus ensuring a correct chromosome number in daughter cells.

However, some cells are capable to escape the mitotic control and to survive and divide despite the presence of mitotic damage, allowing the development of viable aneuploid cells. Several mechanisms were shown to be involved in aneuploidy[11]: (i) errors in centrosome duplication, leading to the generation of more than two centrosomes that result in multiple spindle poles and multipolar division; (ii) cohesion defects, likely due to persistence or premature loss of chromatid cohesion during anaphase; (iii) merotelic attachment, causing chromatid missegregation or exclusion from both daughter cells when one kinetochore is attached to microtubules at both poles; (iv) alterations of mitotic checkpoint signaling; (v) failure of cytokinesis. Physiologically, defects in the segregation process induce a “stop” signal during anaphase. A weakened mitotic checkpoint might allow cells to enter anaphase in the presence of unattached or misaligned chromosome and, both copies of one chromosome might be deposited into a single daughter cell. Therefore, failure of the mitotic checkpoint machinery has been an obvious candidate mechanism involved in the generation of chromosome instability (CIN) during mitosis. However, its molecular players are rarely targeted by mutations in human cancers.

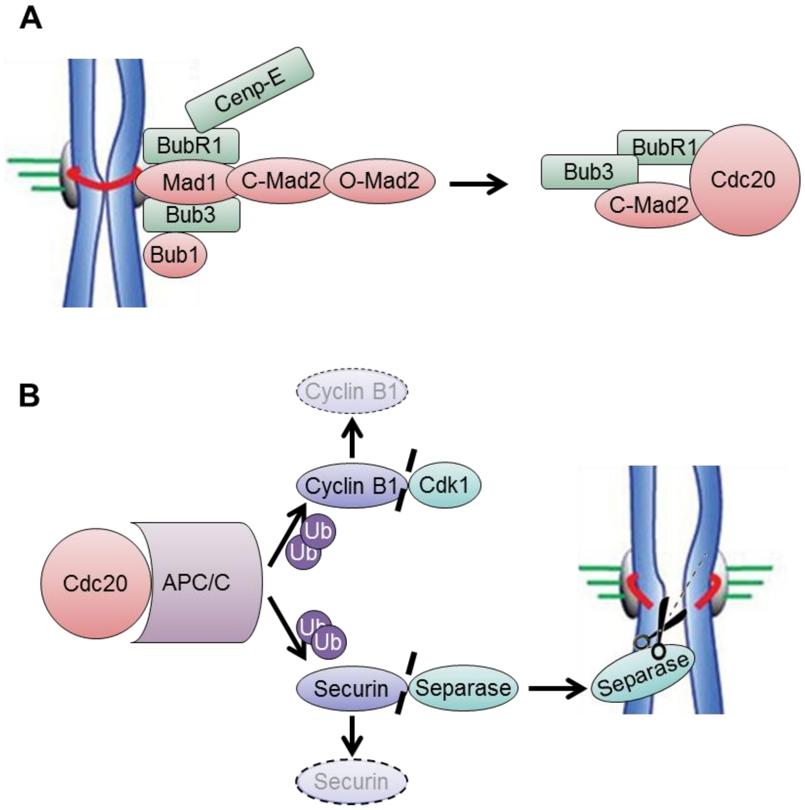


Figure 4. The metaphase checkpoint machinery. (**A**) The correct and timely regulated assembly of Mad2, Bub3 and Cenp-E at the unattached kinetochores leads to the generation of a diffusible Mad2 STOP-signal, depending on the conversion of Mad2 from an open (O-Mad2) to a closed conformation (C-Mad2). C-Mad2 sequesters Cdc20, causing its inactivation, that prevents the anaphase promoting complex/cyclosome (APC/C) from degrading cyclin B1 and securin. Under these conditions, the separation of sister chromatids cannot occur. (**B**) When the last kinetochore pair is attached to microtubules at opposite spindle poles, the inhibitory signal of C-Mad2 is extinguished and Cdc20 is released. Therefore, Cdc20 binds and activate APC/C, that in turn polyubiquitinates cyclin B1 and securin, Cdk1 and separase partners, respectively Cyclin B1 is degraded through the proteasome, leading to a rapid decline of Cdk1 activity. Securin is released from separase, thus activating the degradation of the cohesion complex at and near sister chromatid kinetochores. These events are needed for a correct metaphase-to-anaphase transition and faithful chromosome segregation.

On the contrary, the checkpoint machinery is frequently hyperactivated in chromosomally unstable malignant cells, resulting in mitotic delay, abnormal stabilization of cyclin B1 and securin, and increased incidence of merotelic attachments and lagging chromosomes, generating CIN both *in vitro* and *in vivo*[11].

Notably, eukaryotic cells start cell cycle immediately after they exit from the quiescent G_0 phase, thus implying that the entire cellular architecture needs to be prepared to sustain mitosis and accomplish a correct cell division. This observation clearly suggests that aneuploidy can be caused by dysfunction of cellular component not directly involved in mitosis, but essential in the other step of cell cycle, either through mutations, copy number alterations, epigenetic modifications or deregulated expression.

Tumor-protecting and tumor-promoting effects of aneuploidy

One of the most active research field is currently focusing on the complex relationship between aneuploidy and cancer. Evidence suggests that aneuploidy can exert an anti-tumorigenic or a pro-tumorigenic effect (Figure 5).

The first studies conducted on yeast strains, murine and human cells showed that aneuploidy impairs the proliferative capacity of non-malignant cells and the phenotype is independent of the identity of the individual chromosomes, while being potentially proportional to its size[12–15]. This was linked to an imbalance in the cellular protein composition, which may saturate key chaperones, prohibiting them from folding client proteins required for cell viability, thus eliciting a proteotoxic stress response and altering the redox anabolic homeostasis, with increased reactive oxygen species (ROS)[16]. To bypass the unfitness barrier exerted by the abnormal chromosome number, aneuploid cells undergo a metabolic reprograming, characterized by heightened glucose and/or glutamine consumption[13,15]. This phenotype is sustained by aneuploidy-related transcriptomic and proteomic signatures, including genes involved in cell cycle, ribosome biogenesis, energy production and response to stress[12]. Several lines of evidence argue for a negative effect of aneuploidy on the fitness of non-malignant cells (Figure 2A). First, a recent single cell sequencing study revealed that aneuploidy is an extremely rare event in normal conditions, even in brain and liver tissues, accounting for less than 5% of all cells[17]. Second, trisomic murine embryonic fibroblasts (MEFs) show contact inhibition properties, proliferation arrest in low-serum medium, lack of clonogenic capacity and senescence features after 7-10 passages in culture[18]. Third, individuals affected by Down syndrome display a reduced incidence of solid tumors, including breast, lung, and prostate cancers[19]. Forth, trisomic cells can revert to the euploid state by losing extra chromosomes both *in vitro* and *in vivo*, in order to acquire a growth advantage[18].

These observations suggest that single-chromosome aneuploidy is not sufficient *per se* to induce malignant transformation. In addition, despite the observation of spontaneous increase of micronuclei in *BubR1*$^{+/-}$ MEFs, mice did not show any increase in spontaneous tumor formation, which was observed once they were challenged with chemical carcinogens (Table 2). Accordingly, some mouse models of CIN are at a significantly decreased risk of developing tumors (Table 2), even under various oncogenic backgrounds.

Despite the detrimental effect of chromosome number alterations on cellular fitness, aneuploidy is a hallmark of cancer, a disease of cells undergoing uncontrolled proliferation. According to the Felix Mitelman Database[20], about 90% of solid tumors and 50% of hematological neoplasms are aneuploid. How can this be reconciled with the previously reported findings?

Aneuploidy drives an adaptive response by inducing genome[21,22] and chromosome[18,23] instability. It has been recently proposed that oncogenic alterations targeting *TP53*, *RB*[24], *KRAS*[25] and aneuploidy itself[23] cause replicative stress, which perturbs SAC genes, as *MAD2*, resulting in CIN[26], even in the absence of mutations in genes involved in chromosome segregation or mitotic checkpoints. The progressive acquisition of mutations and/or copy number variants increases the cell tolerability towards the negative consequences of the altered chromosome number (Figure 5B). In particular, the genome destabilizing effect of aneuploidy confers an evolutionary flexibility that may contribute to the aggressive growth of advanced malignancies with complex karyotypes and some genetically engineered models of CIN develop tumors at accelerated rate (Table 2), particularly when combined with inactivation of the *TP53* tumor suppressor gene[27]. Moreover, the adaptive response forced by aneuploidy includes (i) heightened anchorage-independent growth and migration capacity, as shown in a colorectal model[28]; (ii) redistribution of cellular resources, leading to the reduction of ribosome synthesis in favor of telomerase components and other cellular proteins, as demonstrated in yeasts[29]; (iii) decreased neoantigen load in most tumors, possibly mediated by limited neoantigen generation and presentation through the MHC complex, which is relevant to tumor recognition by the immune system[30]. This in turn results in decreased immune cell infiltration and makes aneuploid neoplasms less responsive to immunomodulating agents (Figure 5B).

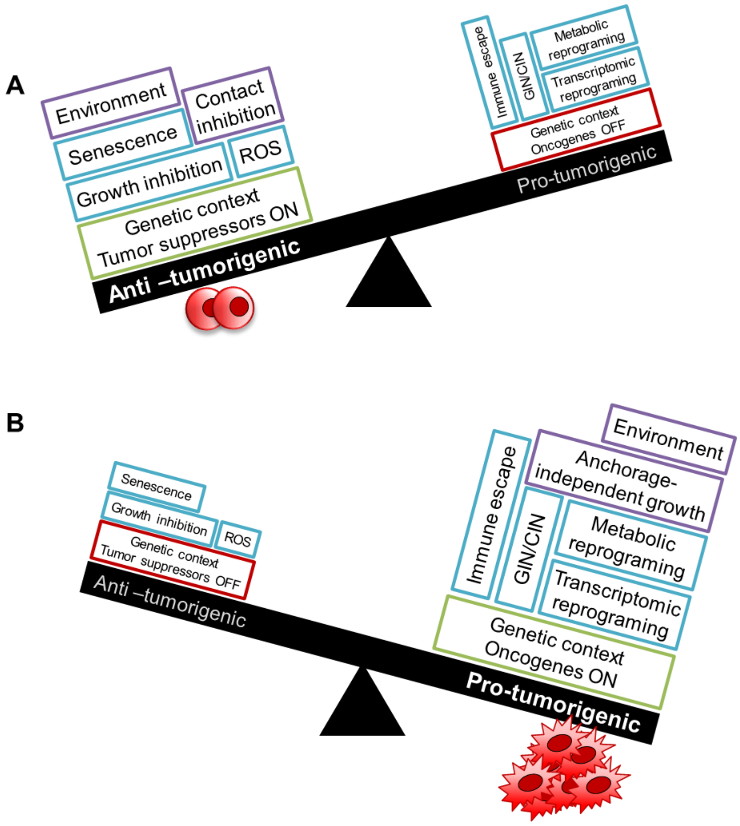


Figure 5. The complex relationship between aneuploidy and cancer (ROS = Reactive oxygen species; CIN: Chromosomal instability; GIN: Genomic Instability).

Aneuploidy and cancer: the cell type, genomic background and environmental conditions matter.

Although Down syndrome patients have a ten-fold reduction of solid tumors-related mortality compared with the general population, they are prone to leukemia development[19], and gain of chromosome 21 is a common event in sporadic leukemia. This apparent paradox argues for a tissue and chromosome-specific oncogenic effect of aneuploidy. Accordingly, malignant transformation does not occur randomly in the majority of transgenic and knock-out mouse models of aneuploidy (Table 2). *Cenp-e*$^{+/-}$ mice have reduce incidence of developing spontaneous liver tumors, while they are more prone to hematological and lung cancers. In parallel, *MAD2* overexpression specifically increases the susceptibility to hepatoma and hepatocellular carcinoma, lung adenomas, fibrosarcomas and lymphomas.

Although the spectrum and the degree of aneuploidy change across tumors, many human cancers share recurrent aneuploidies[31]. According to computational modeling, chromosome number alterations do not occur by chance: a selective pressure forces the acquisition of specific oncogenes and loss of tumor suppressors[32]. This is coupled to common phenotypic consequences induced by the unbalanced protein load[12–15], and participate to cellular adaptation to specific biological and environmental conditions. Indeed, in normal conditions aneuploidy is well tolerated under "population flush" effects, when rapid cell expansion is needed (e.g. during embryogenesis) and in non-regenerating tissues, as brain and liver, in which the potentially dangerous consequences of aneuploidy are prevented by the non-proliferative cellular features. On the contrary, aneuploidy is physiologically selected against in tissues that undergo self-renewal, including the hematopoietic compartment, the skin, and the intestines[10]. However, aneuploidy improves the survival rate under stress conditions, as hypoxia and chemotherapy pressure, both in cancer models[28] and budding yeasts[33]. This scenario recapitulates tumor biology, since both cancer and leukemia stem cells mainly localize in hypoxic niches. If we consider aneuploid cells as a pre-malignant state, their genomic plasticity confers the ability to evolve to a more malignant phenotype, in order to tolerate adverse environmental conditions. The DNA replication stress, that fuels defective chromosome condensation and segregation in human aneuploid pluripotent stem cells[26] may also propagate genome instability in cancer stem cells. Karyotypic heterogeneity may in turn result in phenotypic variation allowing specific aneuploid cells to be more "fit" under stress conditions.

Therapeutic potential of aneuploidy in cancer patients

Aneuploidy has been correlated with a transcriptomic signature of CIN in malignant cells[34]. Overexpression of the signature predicts inferior outcome across several cancer types. However, cases with extreme CIN score, according to their gene expression profiling[34], displayed better prognosis compared with the ones having intermediate score in breast, ovarian, gastric, and non-small cell lung cancer[35]. Moreover, hyperdiploidy (>50 chromosomes) generally predicts good prognosis in pediatric and adult acute lymphoblastic leukemia, while hypodiploidy (<44 chromosomes) is associated with poor outcome and a progressively reduced survival along with chromosome number decrease[36].

This evidence suggests a dual relationship between aneuploid malignancies and anti-tumor therapeutic strategies: aneuploidy can be a cancer strength or an Achilles' heel. Aneuploidy promotes cancer immune escape and correlates with bad prognosis in response to immune checkpoint blockade agents, according to two different clinical trials in metastatic melanoma patients[30]. However, tumors characterized by high rate of pre-existing CIN (e.g. the one s with extreme CIN score), which seem to have reduced fitness, may be induced to mitotic catastrophe by drugs acting at the chromosome segregation level, in particular by enhancing the chromosome missegregation rate.

Mouse model	Phenotype	Reference
Cenp-e$^{+/-}$	• Near-diploid aneuploidy and chromosomal instability. • Prone to develop splenic lymphomas (10%), lung adenomas (3-fold increase). • 50% decrease in liver tumor incidence.	Weaver B.A.A. et al. *Cancer Cell.* 2007
MAD2-tg	• Aneuploid and tetraploid cells, chromosomal breaks and fragments, end-to-end fusions (dicentric and acentric chromosomes), chromatid breaks and gaps. • 50% of mice were dead by 75 weeks. • Prone to develop hepatoma and hepatocellular carcinoma, lung adenomas, fibrosarcomas and lymphomas.	Sotillo R. et al. *Cancer Cell.* 2007
Mad2$^{+/-}$	• Defective mitotic checkpoint and chromosome missegregation. • High rate of papillary lung adenocarcinomas in aged mice.	Michel L.S. et al. *Nature.* 2001
Mad1$^{+/-}$	• Aneuploidy. • Prone to develop lung adenocarcinoma, hepatocellular carcinoma, rhabdomyosarcoma, osteosarcoma, hemangiosarcoma, and uterine sarcoma (2-fold increase) by 18 months of age.	Iwanaga Y. et al. *Cancer Research.* 2007
Bub1-tg	• Chromosome missegregation due to misalignment and near diploid aneuploidy. • Prone to develop d lymphomas, lipomas, sarcomas, liver and skin tumors (≈67%). • Premature onset of Eμ-*Myc*-mediated lymphoma.	Ricke M.R. et al. *The Journal of Cell Biology*. 2011
BubR1$^{+/-}$	• Defective in spindle checkpoint activation, reduced securin and CDC20 expression, increased level of micronuclei. • No effects on the frequency or rate of spontaneous tumors. • High incidence and premature onset of colon adenocarcinoma when primed with azoxymethane. • Develop lung and liver tumors when primed with azoxymethane.	Dai W., Wang Q. et al. *Cancer Research*. 2004
BubR1-tg	• Genomic integrity is preserved through correction of mitotic checkpoint impairment and microtubule-kinetochore attachment defects. • Resistance to Ras-mediated tumorigenesis.	Baker D.J., Dawlaty M.M. et al. *Nature Cell Biology*. 2013
Bub1$^{-/H}$ *Bub1*$^{H/H}$ *Bub1*$^{+/-}$	• Weakened mitotic checkpoint and aneuploidy, with a milder phenotype and a higher Bub1 expression in *Bub1*$^{+/-}$ mice. • *Bub1*$^{-/H}$: prone to develop sarcomas, lymphomas, and lung tumor. • Bub1$^{H/H}$: prone to develop sarcomas and highly susceptible to hepatocellular carcinomas. • *Bub1*$^{+/-}$: decreased tumor incidence, especially in the liver and the lung.	Jeganathan K. *Journal of Cell Biology*. 2007; Baker D.J. et al. *Nature Genetics*. 2004
Bub3$^{+/-}$	• Aneuploidy, premature sister-chromatid separation and chromatid breaks. • No effects on the frequency or rate of spontaneous tumors.	Kalitsis P. et al. *Genes Chromosomes Cancer*. 2005
TetO-Hec1-tg	• Hyperactive mitotic checkpoint and increased Mad2 expression. • Prone to develop lung adenomas (12.8%), liver tumors (25.5%), hemangiosarcoma.	Diaz-Rodriguez E., Sotillo, R. et al. *PNAS.* 2008

Cdh1$^{+/-}$	• Numeric and structural chromosomal aberrations. • Increased susceptibility to spontaneous tumours.	Garcia-Higuera I. et al. *Nature Cell Biology*. 2008
Plk4$^{+/-}$	• Increased centrosomal amplification, multipolar spindle formation and aneuploidy. • Prone to develop liver and lung cancers (15-fold higher) with aging.	Ko, M. A. et al. *Nature Genetics.* 2005
CDC20$^{+/AAA}$	• Functional loss of spindle assembly checkpoint, premature anaphase and aneuploidy. • Prone to develop tumors (50% by 24 months of age), especially hepatomas and lymphomas.	Li, M., Fang, X. et al. *The Journal Cell Biology*. 2009
MMTV-AuroraA-tg	• Genetic instability associated with activation of the AKT pathway, centrosome amplification, chromosome tetraploidization, premature sister chromatid segregation. • Prone to develop mammary tumors between 9 and 20 months of age (40%).	Wang, X. et al. *Oncogene*. 2006
Rae1$^{+/-}$ *Bub3*$^{+/-}$	• Defective mitotic checkpoint and chromosome missegregation. • Prone to develop carcinogen-induced lung tumors.	Babu, J. R. et al. *The Journal Cell Biology*. 2003

Table 2. Mouse models with hyperactive or defective mitotic checkpoint showing evidence of increased or reduced predisposition to tumor development.

This includes compounds that disrupt microtubule dynamics, either by inducing over-polymerization (stabilizing drugs, e.g. taxane), or by reducing polymerization (destabilizing drugs, e.g. vinblastine), drugs that target the machineries regulating kinetochore-microtubule attachment, correction of misattachments (e.g. Aurora B, required to destabilize incorrect anchoring kinetochore-microtubules) or silencing of the SAC. These hypotheses have been tested at preclinical level and need to be verified in clinical settings. The severe bone marrow toxicity of mitotic drugs should be taken into account while designing *ad hoc* combination therapies, which may develop on a chemotherapy backbone, in order to succeed in tumor debulking and disease eradication, while reducing side effects. Recently, a clinical trial comparing paclitaxel response with CIN level in breast cancer patients has opened the recruitment phase (NCT03096418, clinical trials.gov). In parallel it has been demonstrated that cancer cell lines with defects in chromatid cohesion were resistant to paclitaxel, but highly responsive to inhibition of the SAC in case of intact microtubule pulling forces[37].

This approach is built on the concept of synthetic lethality, which refers to the simultaneous perturbation of two genes resulting in cell or organism death. Certain drugs can cause lethality in malignant cells carrying structural or functional alterations in specific genes or pathways. These “lethal” combinations should be exploited to target aneuploidy-supporting cellular functions. Indeed, besides their neutropenic effects, mitotic drugs are not expected to be effective against tumors displaying a negative correlation between CIN and survival. The strength of these aneuploid tumors reside in their increased tolerability towards stress conditions, their stem cell-like quiescent state and genomic complexity, which likely favor resistance to chemotherapy and progression to a very aggressive phenotype. Given aneuploid cell-dependance on chaperone pathways and heightened protein turnover required to fight the unbalanced protein load, they may collapse in response to proteasome and protease inhibition, due to exaggerated proteotoxic stress. The forced cellular metabolism needed to sustain aneuploidy (e.g. glycolysis or glutamine dependence) and the weak ribosome biogenesis may be additional valuable therapeutic targets to be exploited under the aneuploid condition.

Conclusions and outlook

An improved understanding of the molecular mechanisms underlying aneuploidy and of its consequences on cell biology has revealed a complex relationship between chromosome gain/loss and cancer. Aneuploidy can increase malignant cell strength while causing their vulnerability to specific conditions or therapeutic interventions. The tissue type, the genetic background and the microenvironment play a pivotal role in the match. However, the genetic determinants of the pro-tumorigenic or anti-tumorigenic effects of aneuploidy and their interplay with the biology of the cell of origin remain unclear. Therefore, the identification of the genomic patterns that synergize with aneuploid phenotypic profiles in promoting tumor development and the association between chromosome missegregation frequencies and adaptive levels of CIN might be a prerequisite to any therapeutic decision. These approaches will link genetic variability, drug-resistant growth and acquisition of stem cell characteristics, while defining lineage-specific vulnerabilities for aneuploid tumors. Such knowledge, complemented by the availability of rationally designed targeted agents, which have returned promising results, will serve as a map for personalized synthetic lethal therapeutic strategies.

Fusion genes in AML

Fusion genes resulting from chromosomal translocations are an important class of cancer-associated alterations due to their role in the pathogenesis of the disease, as a molecular biomarker for disease monitor and as attractive targets of therapy. Gene fusions commonly exert their oncogenic role by forming a novel chimeric transcript with an oncogenic functionality (e.g. leading a constitutive activation of a tyrosine kinase), deregulating one of the involved genes (e.g. by juxtapositioning a strong promoter or an enhancer region to an oncogene), or inducing a loss of function (e.g. by truncating a tumor suppressor gene). As an iconic example, the *BCR-ABL* fusion gene was firstly described in chronic myeloid leukemia (CML): the oncogenic function of the chimera derives from the constitutively active tyrosine kinase activity of *ABL*, which leads to the phosphorylation of several cellular substrates and activation of a number of signal pathways involved in control of cell proliferation and differentiation, adhesion and cell survival[38]. Nowadays, *BCR-ABL* is the target of tyrosine kinase inhibitor (TKI) treatments and it is a powerful diagnostic molecular biomarker to monitor molecular response to target therapies (also in an minimal residual disease setting)[39].

According to the Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer, the number of annotated gene fusions account for 10993 entities in 67625 cases[20]. Several gene fusion have been described in hematological malignancies and, in particular, more than 30% of AML patients are characterized by the presence of a recurrent fusion gene[1]. The frequency of the 4 most recurrent fusions is between 1% and 13% of patients, they are associated to chromosomal translocations detected by fluorescence *in situ* hybridization (FISH, Table 3)[40] and they are currently used as prognostic and diagnostic markers. The chromosomal translocation t(15;17), which lead to the expression of the *PML-RARα* chimera, characterize patients with AML M3. In physiological conditions, RARα interacts with RXR and binds DNA, and, upon the binding of ATRA, a conformational change lead to the dissociation of the complex[40]. The fusion protein is insensitive to physiological concentrations of ATRA and it complexes with RXR forming an oligomeric complex essential to exert its functions as an oncogene[41,42]. *PML-RARα* acts as a transcriptional repressor, interfering with gene expression programs, which in turn control differentiation, apoptosis and self-renewal. Patients with this subtype of AML have favorable prognosis and treatments with all-trans-retinoic acid (ATRA) and/or As_2O_3 is able to overcome the transforming potential of PML-RARα[40].

The transcript *RUNX1-RUNXT1* and *CBFβ-MYH11* are associated to the t(8;21) and inv(16), respectively, which are classified also as favourable prognosis, but no target agents have been developed yet. *RUNXT1* encodes for a transcriptional repressor, while *RUNX1* is a transcription factor part of the core binding factor (CBF) transcriptional complex that regulates important target genes in hematopoiesis. *RUNX1* fusion proteins disrupt the normal myeloid gene expression and it seems it acts in a dominant negative fashion to inhibit the normal transcriptional activity of *RUNX1/CBFβ*[43]. In addition to t(8;21), more than 50 chromosome translocations and different types of alterations including somatic and germline point mutations, affect *RUNX1*[44,45]. *CBFβ* participates in the formation of the of CBF complex and, consequently, it interacts with *RUNX1*. The fusion protein *CBFβ-MYH11* predispose cells to leukemic transformation by interfering in a dominant-negative manner with CBF, thereby impairing hematopoietic differentiation. Both *RUNX1-RUNXT1* and *CBFβ-MYH11* are not sufficient to induce the leukemic phenotype in vivo[46] and additional alterations such as *FLT3-ITD, KRAS* and *KIT* mutations are needed[47–50].

The t(6;9), inv(3)/t(3;3), t(v;11q23.3) and t(9;22) abnormalities results in the expression of *DEK-NUP214*, *GATA2*/*MECOM* fusions, *KMT2A*-fusions and *BCR-ABL*, respectively, which correlates with adverse prognosis[1]. The fusion *DEK-NUP214* leads to increased protein synthesis[51], promotes proliferation via mTOR signalling[52] and induces leukemia in mice[53].On the other hand, *MECOM* and *GATA2* are two transcription factors involved in the development and proliferation of hematopoietic cells whose rearrangements has been linked to leukemogenesis by the overexpression[54] and the displacement of an enhancer sequence[55], respectively. *KMT2A* encodes for a histone methyltransferase and it has been shown to be required for the development and maintenance of hematopoiesis[56]. The translocation usually involves the N-terminal of the gene and the most frequent fusion in AML is *KMT2A-AF9* (or *MLL-AF9*)[57], where the fusion protein induces the aberrant expression of a self-renewal–associated gene-expression program[58]. Moreover, *KMT2A*-fusions are generally associated with other acute leukemias, where it fuses to 94 different partner genes, resulting in a KMT2A-fusion protein that acts as a potent oncogene[57]. Except for patients with the *MLL3-KMT2A* rearrangements which falls into the intermediate risk class, all other *KMT2A*-fusions have adverse prognosis[1]. The *BCR-ABL* fusion gene encodes for a constitutively active tyrosine kinase and it is most commonly associated with chronic myelogenous leukemia, a subset of precursor B cell acute lymphoblastic leukemia (B-ALL) and acute biphenotypic leukemia. In addition, 1 % of de-novo AML are associated with the Philadelphia Chromosome[1].

Given that in most cases of AML expressing fusion genes, one of the translocated partner gene is represented by a transcriptional factor, the mechanism of leukemogenesis is associated to an aberrant transcriptional regulation and, consequently, to a change in the expression profile[59].

All fusion proteins converge in the interference of the process of myeloid differentiation, suggesting that it is a common molecular mechanism that must be disrupted in order to acquire a transformed phenotype. However, like *RUNX1-RUNXT1*, the expression of a fusion protein may be not sufficient to induce the leukemic phenotype and co-operating alterations are needed to the onset of AML[43,47,48,60].

Translocation	Prognosis	Oncofusion protein	Occurrence in AML
t(8;21)(q22;q22)	Favorable	*RUNX1-RUNXT1*	7%
t(15;17)(q24;q21)	Favorable	*PML-RARα*	13%
inv(16)(p13;q22) or t(16;16)(p13;q22)	Favorable	*CBFβ-MYH11*	5%
t(v;11q23)	Adverse	*KMT2A*-fusions	4%
t(9;22)(q34;q11)	Adverse	*BCR-ABL1*	1%
t(6;9)(p23;q34)	Adverse	*DEK-NUP214*	<1%
inv(3)(q21;q26) or t(3;3)(q21;q26)	Adverse	*RPN1-MECOM (GATA2)*	<1%

Table 3. Recurrent balanced translocation and their relative fusion gene and frequency in AML.

The study of the Cancer Genome Atlas Research Network (TCGA) carried out on 179 AML patients enabled the detection of 118 fusions by RNAseq, with an average of 1.5 fusion per patient[2]. Of these, 74 were in-frame events and included many recurrent and previously described events such as *PML-RARα*, *MYH11-CBFβ*, *RUNX1-RUNXT1*, *KMT2A*-fusions. An independent study on the same dataset observed a statistically significant reduction of the frequencies of significant gene mutation in AML patients with recurrent in-frame fusion transcripts compared with those without recurrent in-frame fusion transcripts, suggesting and/or supporting the driver role of these fusions in the pathogenesis of AML[61]. Fifteen novel in-frame chimera were identified and none of these were recurrent among the characterized cohort of patients. However, some genes involved were mutated or translocated across analyzed samples. Forty-two chimera were out of frame, suggesting a potential loss of function for one or both genes involved. Notably, most of the novel fusion event were not detected by means of routine cytogenetic analysis[2].

The rapid increase of sequencing studies has led to the creation of several databases that collects fusion genes. One of the earliest effort is the Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer[20], arose before the advent of deep sequencing. It is a heavily curated database of fusions supplemented with clinical association information, like karyotype abnormalities associated with a particular tumor type or patient prognosis.

On the other hand, the TCGA Fusion Gene Data Portal database was recently developed and it is based on integrated analysis of paired-end RNA sequencing and DNA copy number data from the TCGA dataset, providing a bona-fide fusion list across many tumor types[61]. Several other databases collecting gene fusions exist[62] and, according to our knowledge, we may have discovered only a fraction of chimeras. Furthermore, current databases reflect the fact that i) druggable fusion are not so frequent but relevant across cancer types; ii) we know little about certain classes of fusion, such as fusions involving genes encoding long non-coding RNAs; iii) exploring certain rare and poorly understood fusions which are perhaps not directly related to cancer is likely to synergistically improve our understanding of cancer-related fusions.
Therefore, the identification of fusion events, even if private or in a small subgroup of poorly characterized patents, it is of clinical significance. In a perspective of precision medicine approach, a comprehensive knowledge of the landscape of alterations, although rare, must be carried out.

Aims

The overall aim of the study was the molecular characterization of AML for a better stratification of patients through identification of novel biomarkers. To this purpose, the study took advantage of advanced next generation sequencing (NGS) technologies, both at DNA and RNA level. Specific aims were the followings:

1. **Dissecting the molecular mechanisms of aneuploidy in AML.** Aneuploidy, the presence of an abnormal number of chromosome, characterize ≈90% of solid tumors and more than 20% of AML cases. However, aneuploidy *per se* seems to act as a barrier to malignant transformation. The study aimed to elucidate the molecular mechanisms associated with aneuploidy in AML patients, by analysing the genomic and transcriptomic landscape of aneuploid and euploid cases by whole exome sequencing, single nucleotide polymorphism array and gene expression profiling.

2. **Identification of novel fusion genes in AML patients.** Chromosomal rearrangements and fusion genes have a crucial diagnostic, prognostic and therapeutic role in cancer. Most AML cases are associated with non-random chromosomal translocations, which result in the expression of a fusion gene. Therefore, the second aim of the study was the identification of novel and rare fusion transcripts by performing RNA sequencing on AML patients carrying rare or poorly described chromosomal translocation(s).

3. **Analysis of the leukemogenic, prognostic and therapeutic potentials of AML genomic lesions**. The third aim of the study was to perform functional studies (i) to assess the oncogenic potential of identified gene fusion(s); (ii) to discover novel insights into the mechanisms involved in leukemogenesis and explore their potential as therapeutic targets; (iii) to develop genetic models that accurately define novel leukemia subtypes based on the genomic profile of individual patients.

The identification of inter-individual differences that may play a role in leukemogenesis or affect response to the diverse therapeutic interventions, promises to be crucial for the development of strategies to personalize treatments and tailor therapies to different subgroups of AML patients or to each patient in the era of precision medicine.

Results – I

Aneuploidy in AML

Under submission

Aneuploid AML exhibits a signature of genomic alterations in the cell cycle and protein degradation machinery

Giorgia Simonetti,[1*] Antonella Padella,[1*] Italo Farìa do Valle,[2,3] Maria Chiara Fontana,[1] Eugenio Fonzi,[1] Samantha Bruno,[1], Carmen Baldazzi,[1] Viviana Guadagnuolo,[1] Marco Manfrini,[1] Anna Ferrari,[1] Stefania Paolini,[1] Cristina Papayannidis,[1] Giovanni Marconi,[1] Eugenia Franchini,[1] Elisa Zuffa,[1] Maria Antonella Laginestra,[1] Federica Zanotti,[1] Annalisa Astolfi,[4] Ilaria Iacobucci,[1#] Simona Bernardi,[5] Marco Sazzini,[6] Elisa Ficarra,[7] Jesus Maria Hernandez,[8] Peter Vandenberghe,[9] Jan Cools,[9] Lars Bullinger,[10] Emanuela Ottaviani,[1] Nicoletta Testoni,[1] Michele Cavo,[1] Torsten Haferlach,[11] Gastone Castellani,[2] Daniel Remondini,[2] and Giovanni Martinelli,[1] on behalf of the NGS-PTL consortium.

[1]Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy; [2]Department of Physics and Astronomy, University of Bologna, Bologna, Italy; [3]CAPES Foundation, Ministry of Education of Brazil – Brasília (DF), Brazil; [4]Giorgio Prodi" Cancer Research Center, University of Bologna, Bologna, Italy; [5]Unit of Blood Diseases and Stem Cell Transplantation, University of Brescia, Brescia, Italy; [6]Department of Biological Geological and Environmental Sciences, University of Bologna, Bologna, Italy; [7]Politecnico di Torino, Torino, Italy; [8]Fundación de Investigación del Cáncer de la Universidad de Salamanca, Salamanca, Spain; [9]KU Leuven, Leuven, Belgium; [10]University Hospital of Ulm, Ulm, Germany; [11]MLL Munich Leukemia Laboratory, Munich, Germany.

* These authors contributed equally to this work.
#Current: Department of Pathology, ST Jude Children's Research Hospital, Memphis (TN), USA

Running title: Genomic dissection of aneuploid AML

Corresponding authors:
Giorgia Simonetti and Giovanni Martinelli Via Massarenti,9
phone: 0039 0512143791 fax: 0039 0512144037
e-mail: giorgia.simonetti3@unibo.it, giovanni.martinelli2@unibo.it

Abstract

Aneuploidy occurs in more than 20% of acute myeloid leukemia (AML) cases and correlates with adverse prognosis. To understand the molecular bases of aneuploid (A-) AML, we studied the mutational and transcriptional profile in 42 A-AML and 35 euploid (E-) AML. A-AML was characterized by genomic instability based on exonic variants, with an average of 26 somatic mutations per sample compared with 15 lesions in E-AML. Integration of exome, copy number and gene expression data revealed alterations in genes involved in DNA repair (e.g. *SLX4IP*, *RINT1*, *HINT1*, *ATR*) and cell cycle phases (e.g. *MCM2/4/5/7/8/10*, *UBE2C*, *USP37*, *CK2/3/4*, *BUB1B*, *NUSAP*1, *E2F*) in A-AML, which associated with a 3-gene signature defined by *PLK1* and *CDC20* upregulation and *RAD50* downregulation and with silencing of the p53-transcriptional program either at structural or functional level. Moreover, A-AML was enriched for alterations in the protein ubiquitination and degradation pathway, response to reactive oxygen species, energy metabolism and biosynthetic process, which may help facing the unbalanced protein load. E-AML was associated with *BCOR/BCORL1* mutations and overexpression of HOX-family genes. Aneuploidy causes a proliferative disadvantage, mitotic and proteotoxic stress in non-malignant cells. Our findings indicate that aneuploidy-related and leukemia-specific alterations cooperate to tolerate an abnormal chromosome number in AML and point to the mitotic and protein degradation machineries as potential therapeutic targets for synthetic lethal strategies in A-AML.

Introduction

Aneuploidy originates from defects in chromosome segregation and can be due to deregulated centrosome duplication [1,2], alterations in sister chromatin cohesion [3], weakened or hyperactive mitotic checkpoint [4], failure of chromosome detachment from microtubules [5] and/or telomere dysfunction [6].

This chromosomal imbalance is detrimental for fitness and development in yeasts [7,8], Drosophila [9], maize[10], rice[11] and mice[12]. At cellular level, trisomic yeast strains[8], human fibroblast[13] and mouse embryonic fibroblasts[14] undergo a massive transcriptional reprogramming[15,16] and display impaired proliferation, mitotic and proteotoxic stress and metabolic alterations[17]. A decreased proliferative capacity was also observed in aneuplois hematopoietic stem cells[18]. As such, aneuploidy *per se* seems to act as a barrier to malignant transformation. However, ≈90% of solid tumors carry an abnormal chromosome number, and also a fraction of hematological malignancies display chromosome gains or losses [19]. In acute myeloid leukemia (AML), more than 20% of cases carry a whole chromosome trisomy or monosomy either alone or in combination with other cyotogenetic abnormalities [20], with monosomy 7 and trisomy 8 being the most common numerical alterations [21]. Chromosome gains and losses detected at diagnosis are generally preserved at disease progression and relapse, supporting a role as disease initiating events [22]. Moreover, monosomies (e.g. 5 and 7 losses) and the monosomal karyotype predict dismal outcome [21]. Isolated trisomies (e.g. trisomy 13 [23]) have in some but not all series been associated with adverse prognosis [20,24].

A number of genes and pathways participate to prevent the propagation of aneuploidy. Beside *TP53*, a guardian of ploidy [25], other genes have been proposed as involved in mitotic checkpoint and homologous recombination [25,26]. However, they are rarely mutated and their deregulated expression in mice can result either in increased or decreased cancer incidence [27].

Recent studies have analyzed the molecular profile of aneuploid (A-) AML subsets. Isolated trisomy 13 was associated with a high frequency of *RUNX1*, *ASXL1*, *BCOR* and spliceosome-complex gene mutations, along with upregulated FOXO1 and FLT3 and downregulated SPRY expression [23,28]. AML with trisomy 8 showed high frequency *ASXL1* and *RUNX1* mutations [29] and deregulated expression of cell adhesion and apoptosis-regulating genes [30], while *CUX1* was identified as a haploinsufficient tumor suppressor gene in −7/del(7q) [31,32] cases.

Taken together, this evidence suggests that leukemia-specific mechanisms may cooperate to overcome the unfitness barrier associated with aneuploidy.

To elucidate the molecular mechanisms associated with A-AML, we analyzed the genomic and transcriptomic landscape of aneuploid and euploid (E-) leukemia cases by whole exome sequencing (WES), single nucleotide polymorphism (SNP)-array and gene expression profiling (GEP). We here show that A-AML is characterized by high genomic instability and by a cell cycle-related pattern of somatic mutations, copy number and transcriptomic alterations, along with the downregulation of the p53 transcriptional program and with genomic lesions affecting genes belonging to the protein ubiquitination and degradation machinery.

Materials and Methods

Patients

Primary samples from AML patients (≥18 years) were obtained after informed consent as approved by the Institutional Ethical Committee (protocol number 253/2013/O/Tess and 112/2014/U/Tess of Policlinico Sant'Orsola-Malpighi, Bologna, and internal MLL board and SOP EN ISO 15189 of Munich Leukemia Laboratory, Munich) in accordance with the Declaration of Helsinki. Statistical significance was determined using the Mann-Whitney test for continuous variables and the Fisher's exact test or chi-square test for categorical variables.

Sample preparation

Leukocytes were enriched by separation on Ficoll density gradient and lysed in RLT buffer. Buccal swab samples, used as normal matching, were collected with the Oragene Discover kit (DNA Genotek). Genomic DNA, RNA and proteins were extracted by column purification (AllPrep DNA/RNA/Protein Mini Kit and QIAcube, Qiagen) and from saliva by paramagnetic particles (Maxwell® 16 LEV DNA Blood Purification Kit and Maxwell® MDx Instrument), according to the manufacturer's recommendations.

Chromosome Banding Analysis

Chromosome banding analysis was performed on bone marrow cells after short-term cultures (24 and/or 48 hours) as previously reported [33]. Karyotypes were examined after G-banding and described according to International System for Human Cytogenomic Nomenclature (ISCN 2016) [34]. A complex karyotype was defined when three or more chromosomal abnormalities occurred in the same clone. Aneuploidy is defined as the gain or loss of one or more whole chromosomes. According to ISCN criteria, chromosomal gains or structural abnormalities and loss had to be detected in at least 2 and 3 metaphases, respectively, to be acknowledged as clonal.

WES and identification of somatic mutations

Paired-end DNA libraries were prepared from matched tumor and germline DNA from 77 cases using TruSeq Exome Enrichment Kit or Nextera Rapid Capture Expanded kits (Illumina Inc.) according to the manufacturer's protocol. Libraries were sequenced using Illumina HiSeq 1000 (Personal Genomics, Verona, Italy) or HiScan SQ ("Giorgio Prodi" Cancer Research Center, University of Bologna, Italy) and 100-bp paired-end sequences were generated. Identification of tumor-specific variants, DNA Sanger sequencing, targeted resequencing and mutational signature analysis are described in the Supplemental Data.

Copy number alteration (CNA) analysis

Genome-wide CN analysis was carried out on 70 AML samples included in the WES cohort, using Human Cytoscan HD or SNP 6.0 arrays (Affymetrix). Microarray data are available at the following link: https://ngs-ptl.unibo.it:5006 (access through WebDAV protocol; username SimonettiPadella17_1; password: Revisor_002). Data analysis is described in the Supplemental Data.

Enrichment analysis

Enrichment analyses were conducted in R v3.3.2 [35] and Bioconductor v3.4 (BiocInstaller 1.24.0) using the following packages: "org.Hs.eg.db" v3.4.0,[36] "clusterProfiler" v3.2.11 [37], "GO.db" v3.4.0 [38]. CNA events were grouped as follows: heterozygous + homozygous amplifications (gain+duplication) and heterozygous + homozygous deletions (loss+deletion). Multiple events of the same type in the same gene were considered as one. Fisher's exact test was used to test events at single gene level and p-values were corrected for multiple testing with Benjamini-Hochberg method (FDR < 0.05) [39].

Genes affected by CNAs were annotated according to Gene Ontology Biological Processes (GO-BP) [40]. An over-representation test (based on hypergeometric distribution [37]) on each pathway was performed at patient level. Then, the adjusted *p*-values values (FDR < 0.05) obtained for a certain pathway across all patients were used as predictor variable in a logistic regression model fitted against the case (A-AML)/control (E-AML) classification as dependent variable (0=case, 1=ctrl). P-values from all the logistic regression tests were adjusted for multiple testing (FDR < 0.05; 99,9999% CI). Over-representation analysis of cytogenic bands was performed using WEB-based GEne SeT AnaLysis Toolkit [41].

Gene expression profiling (GEP)

Labeled cDNA was prepared and hybridized to GeneChip Human Transcriptome Array 2.0 (Affymetrix) according to manufacturer's recommendations. Raw data were processed by Expression Console software with Robust Multi-Array (RMA) normalization. Supervised data analysis was carried out with Transcriptome Analysis Console v3.0 software (Affymetrix). Downstream analyses are described in the Supplemental Data. Microarray data can be accessed as reported above.

Western blot analysis

Protein extracts were separated by SDS-PAGE and transferred onto nitrocellulose membranes. The following antibodies were used: rabbit anti-RAD50, rabbit anti-CDC20 (D6C2Q), rabbit anti-PLK1 (28G4; all Cell Signaling Technologies), mouse anti–β-actin (AC-74; Sigma-Aldrich); HRP-conjugated anti–rabbit IgG and anti–mouse IgG (GE Healthcare). ECL Prime (GE Healthcare) and SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific) reagents were used for detection. Quantitative analysis was performed using the ImageJ software (1.45s; National Institutes of Health).

Results

Genomic complexity in A-AML

To investigate the genetic lesions associated with the aneuploid phenotype in AML, we performed WES of 42 A-AML and 35 E-AML cases. Patient characteristics are reported in Table 1 and Table S1. The A-AML cohort included six cases with isolated trisomy; one case displaying trisomy plus another alteration; two *EVI1*-related, three *KMT2A*-rearranged and eight core binding factor (CBF) AML carrying additional abnormalities (including whole chromosome gain/loss) and 22 cases with complex (CK) or monosomal (MK) karyotype. E-AML cases were normal karyotype AML or carried structural chromosomal abnormalities in the absence of clonal numerical alterations. A-AML patients were older (median age: 62 vs. 56 years of E-AML, p=0.02) and most of them had adverse prognosis, according to the 2017 European Leukemia Net recommendations [42].

A-AML displayed a significantly higher mutation load than E-AML, with an average number of somatic mutations of 26 (15 in E-AML, p<0.001, Figure 1A and Table S2). The TCGA cohort showed an average number of 16 and 12 somatic mutations in A-AML and E-AML, respectively (p=0.027). More than 50% of A-AML displayed ≥20 mutations per case, compared to 17% of E-AML (p=0.002, Figure 1B). *In silico* analysis indicated that an average number of 11 and 7 amino acid substitutions in A-AML and E-AML, respectively (p=0.008), had an impact on protein function, thus contributing to the higher complexity and heterogeneity of A-AML mutational background. The increased number of mutations in A-AML was confirmed in the TCGA dataset (16 and 12 mutations in A-AML and E-AML respectively, p=0.027), where no differences were detected in terms of patients' age between the two cohorts. To understand whether the number of somatic mutations was dependent on patients' age, we performed linear regression analysis. We observed an age-dependent increase in the mutation load specifically in A-AML, with no correlation in E-AML, both in our cohort (Figure S1A) and in the TCGA dataset (Figure S1B). Moreover, no significant difference in the number of somatic mutations was observed between CK-A-AML and non-CK-A-AML cases both in our cohort and in the TCGA one. Taken together, the data suggest a higher complexity and heterogeneity of A-AML mutational background.

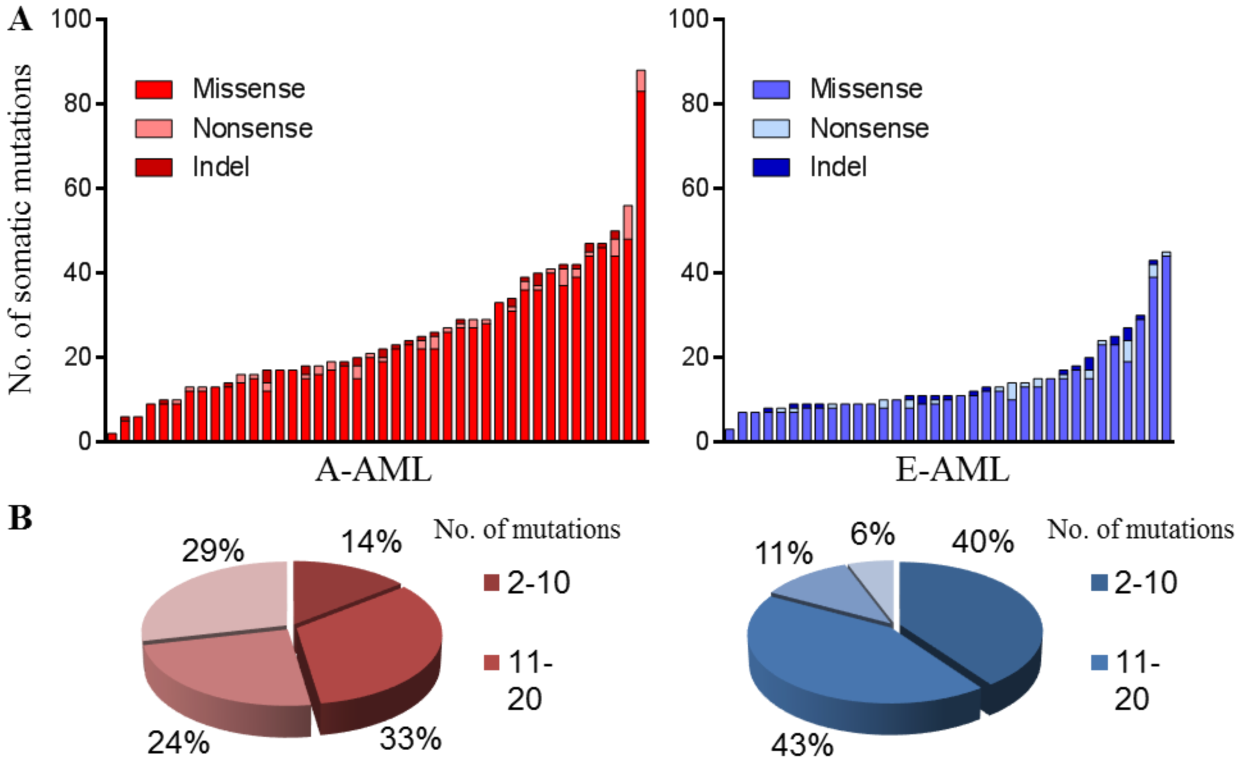


Figure 1. Genetic instability in A-AML. (A) Number and type of non-silent mutations detected by WES in each A-AML and E-AML case. (B) Frequency of A-AML and E-AML cases classified according to the number of mutations

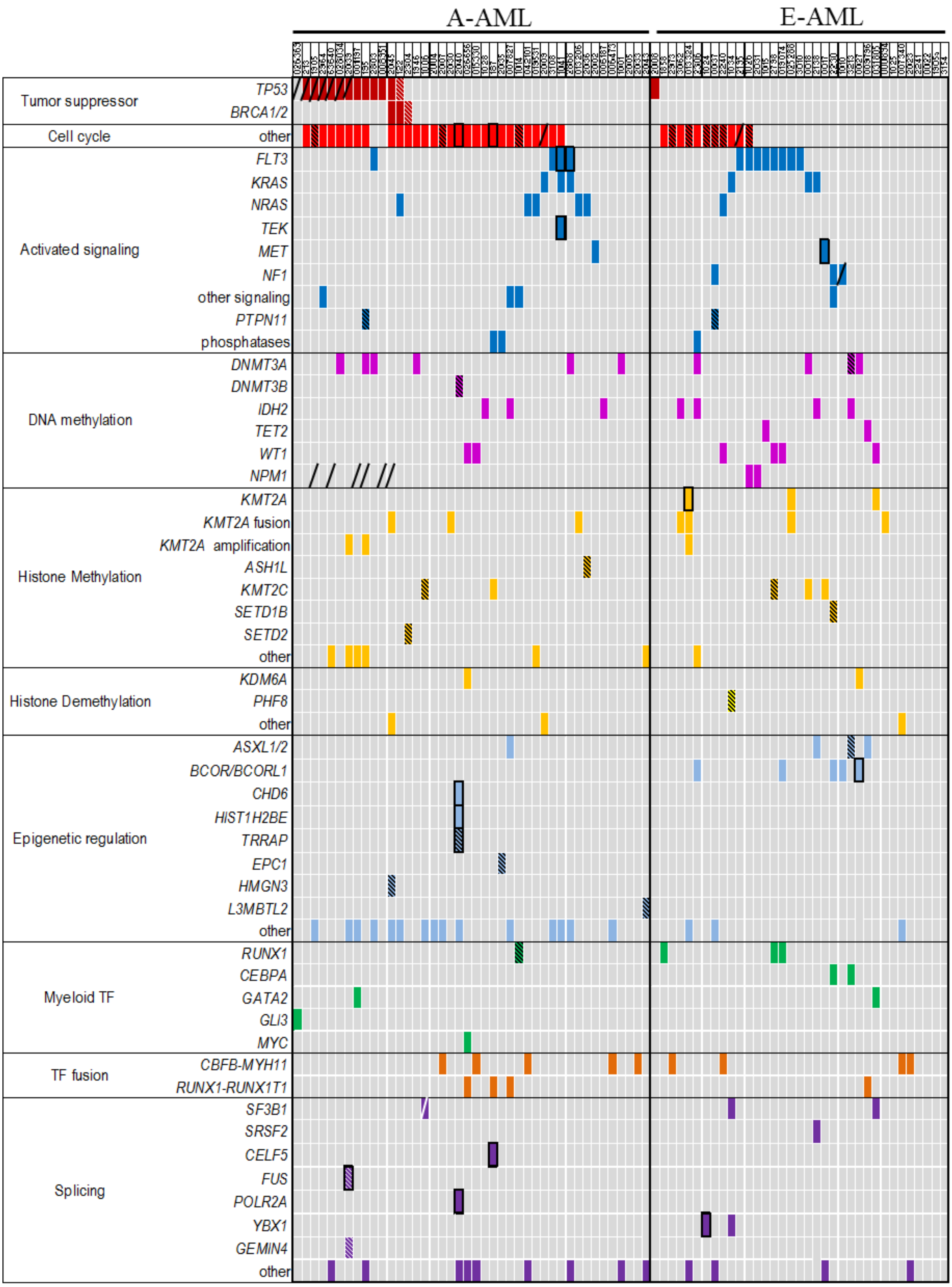


Figure 2. Pattern of genomic lesions in AML. Pattern of genomic lesions in A-AML and E-AML. Each row denotes one or more specific gene(s) or group of genes (other). Columns represent (left to right): functional categories, mutated genes/group of genes or other genomic alterations, single patients. Colours indicates functional categories; bars indicate CN loss; squares indicate CN gains. Striped cells indicates LOH.

Patterns of gene mutations in A-AML and E-AML

Recurrent genomic lesions, identified in both cohorts, included *DNMT3A*, *IDH2*, *KRAS* and *FLT3* alterations (Figure 2). The *TP53* gene was differentially mutated in A-AML (28.6% *vs.* 2.9% of E-AML, *p*=0.004), while E-AML was significantly enriched for mutations of the transcription factor *BCOR/BCORL1* (14.3%, absent in A-AML, *p*=0.02, Figure S2). Additional somatic mutations exclusively detected in E-AML (though significance was not reached) targeted *NF1*, *KMT2A* (8.6% of E-AML), *TET2*, *NPM1*, *CEBPA* (5.7% of E-AML), *EZH2*, *ASXL1* (2.9% of E-AML, Figure S2). These data were partially confirmed by the TCGA cohort: i) TP53 mutations were exclusively present in the aneuploidy cases; ii) *NPM1* and *BCOR* mutations were exclusively detected in the euploid cohort, although the statistical significance was not reached for *BCOR*.

To better investigate the role of the mutations in aneuploidy, we compared the percentage of A-AML and E-AML cases carrying at least one lesion in each functional category (Figure 3A). Tumor suppressor genes (*TP53*, *BRCA1*, *BRCA2*) were among those preferentially mutated in A-AML (31.0% *vs.* 2.9% of E-AML, *p*=0.002), along with genes involved in trafficking of proteins between cellular compartments (50% *vs.* 22.9% of E-AML, *p*=0.02, Figure S3 and Table S3), ubiquitination (45.2% *vs.* 20% of E-AML, *p*=0.029, Figure S4 and Table S3), cell adhesion (42.9% *vs.* 11.4% of E-AML, *p*=0.003, Figure S5 and Table S3) and cell cycle (69.0% *vs.* 31.4% of E-AML, *p*=0.001, Figure 2 and Figure 3A). Of note, A-AML was enriched for mutations with a predicted functional impact on ubiquitination (*p*=0.05) and cell cycle processes (*p*=0.04).

Cell-cycle-related mutations, which characterized both CK-A-AML and non-CK-A-AML (68.2% and 70.0% of cases, respectively), and were also enriched in the TCGA A-AML cohort (p=0.04), were mostly private, since alterations in the same gene were not recurrent among patients. In A-AML they targeted cell cycle regulators and direct players involved in many cell cycle phases (Figure 3B), with DNA replication and S phase, G2/M transition, spindle and centrosome dynamics and chromosome segregation being the most frequently mutated (24%, 13%, 9%, 9% and 9%, respectively, Figure 3C). Moreover, mutations targeting the same cell cycle phase co-occurred very rarely. Few cell cycle-related mutations were also detected in E-AML, which mainly affected centrosome and microtubule dynamics (38% and 13%, respectively) and G2/M transition (13%, Figure 3C).

These results suggest that deregulated cell cycle functionality may be involved in the development and propagation of the aneuploid status in AML and changes in the protein balance regulation may favor adaptation of the aneuploid leukemic cells.

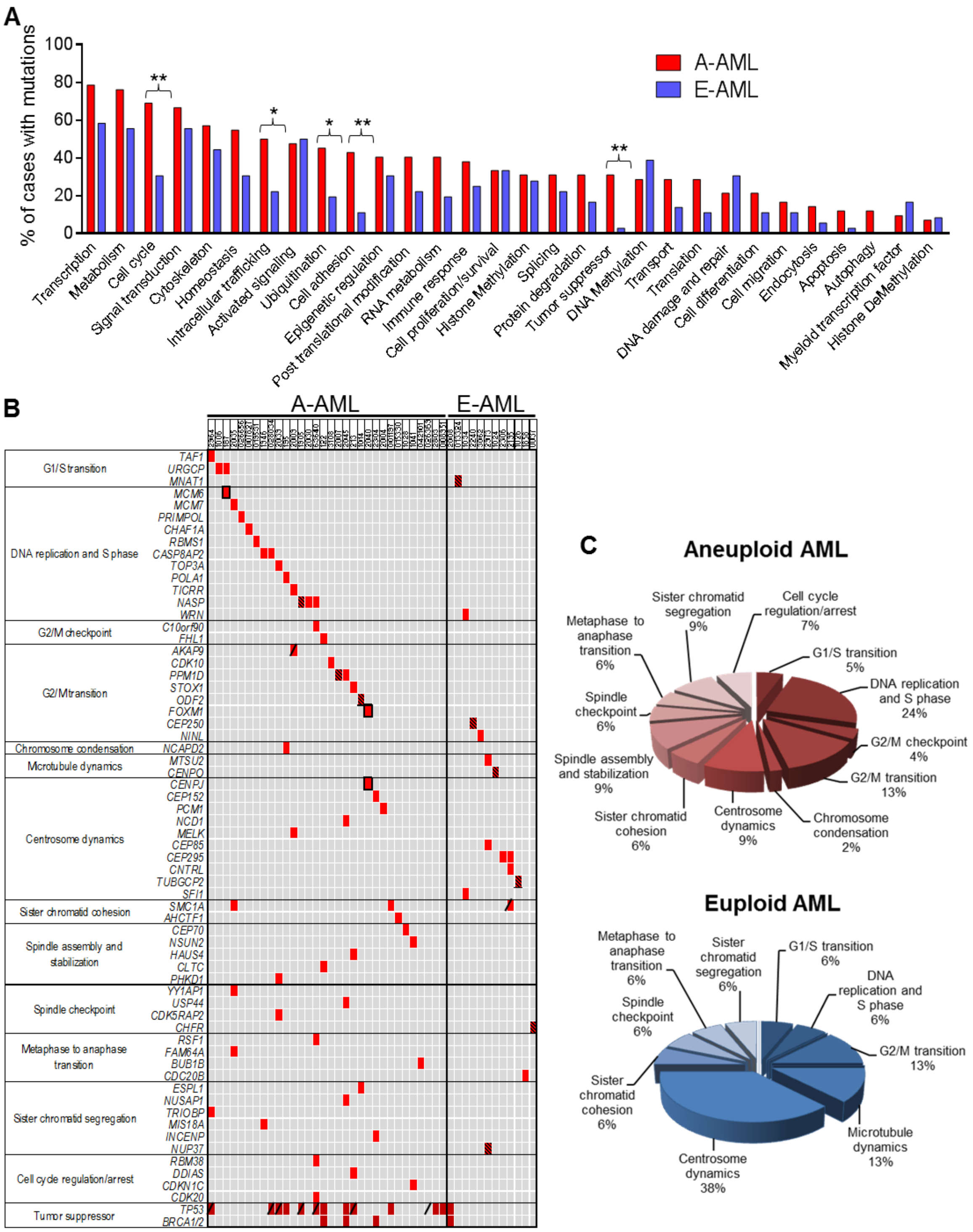


Figure 3. Spectrum of somatic mutation categories distinguishing A-AML and E-AML. (A) Frequency of cases carrying mutations according to functional categories. Statistical significance was determined by Fisher's exact test (*, p <0.05; **, p <0.01; ***, p <0.001). (B) Distribution of mutations targeting cell cycle-related genes. Each row denotes one gene; columns represent (left to right): cell cycle phases, mutated genes, single patients. Colors indicates functional categories; bars indicate CN loss; squares indicate CN gains. (C) Frequency of mutations according to cell cycle phases in A-AML and E-AML.

Mutational signatures in A-AML and E-AML

To gain insights in the mutational processes active in A-AML and E-AML, we analyzed somatic base substitutions in the two cohorts. Both A-AML and E-AML showed a preponderance of C>T transitions (37.7% and 39.7% of overall SNVs in A-AML and E-AML, respectively), as previously reported [43,44], followed by C>A transversions (26.0% and 25.1% in A-AML and E-AML, respectively). However, when considering non-synonymous SNVs (excluding SNPs), the frequency of C>T transitions was reduced in A-AML (24.9% and 34.3% in A-AML and E-AML, respectively, $p<0.001$) and C>A transversions were the most common substitutions, with a different frequency ($p<0.001$) between A-AML (57.1%) and E-AML (45.1%, Figure 4A). The higher C>A incidence was the major determinant of the increase in transversion frequency in A-AML (66.3% of non-synonymous SNVs *vs.* 56.5% in E-AML, $p<0.001$, Figure 4B).

Overall, mutational signature analysis revealed that two signatures contributed to the mutational diversity of our WES cohort (Figure 4C). Signature #1, which characterized 61.9% A-AML and 71.4% E-AML, was dominated by C>T transitions at NpCpG trinucleotides, a mutational process linked to spontaneous hydrolytic deamination of 5-methylcytosines [43] and correlated with age in many cancer types [45]. Signature #2 was characterized by C>A transversions mainly at GpCpN sites and was enriched in 38.1% A-AML and 28.6% E-AML (Figure 4D).

Signature #2 was associated with increased patients' age, especially in the A-AML cohort (median age: 70.5 and 60.5 years in A-AML and E-AML, respectively, *vs.* 62 and 55 years in A-AML and E-AML patients with signature #1 enrichment, $p=0.01$) and increased disease-related mutation load (median number of nonsynonymous SNVs: 34.5 and 21 in A-AML and E-AML, respectively, *vs.* 15.5 and 9 in A-AML and E-AML patients with signature #1 enrichment, $p=0.005$), with A-AML patients enriched for signature #2 being the oldest and carrying the highest number of mutations.

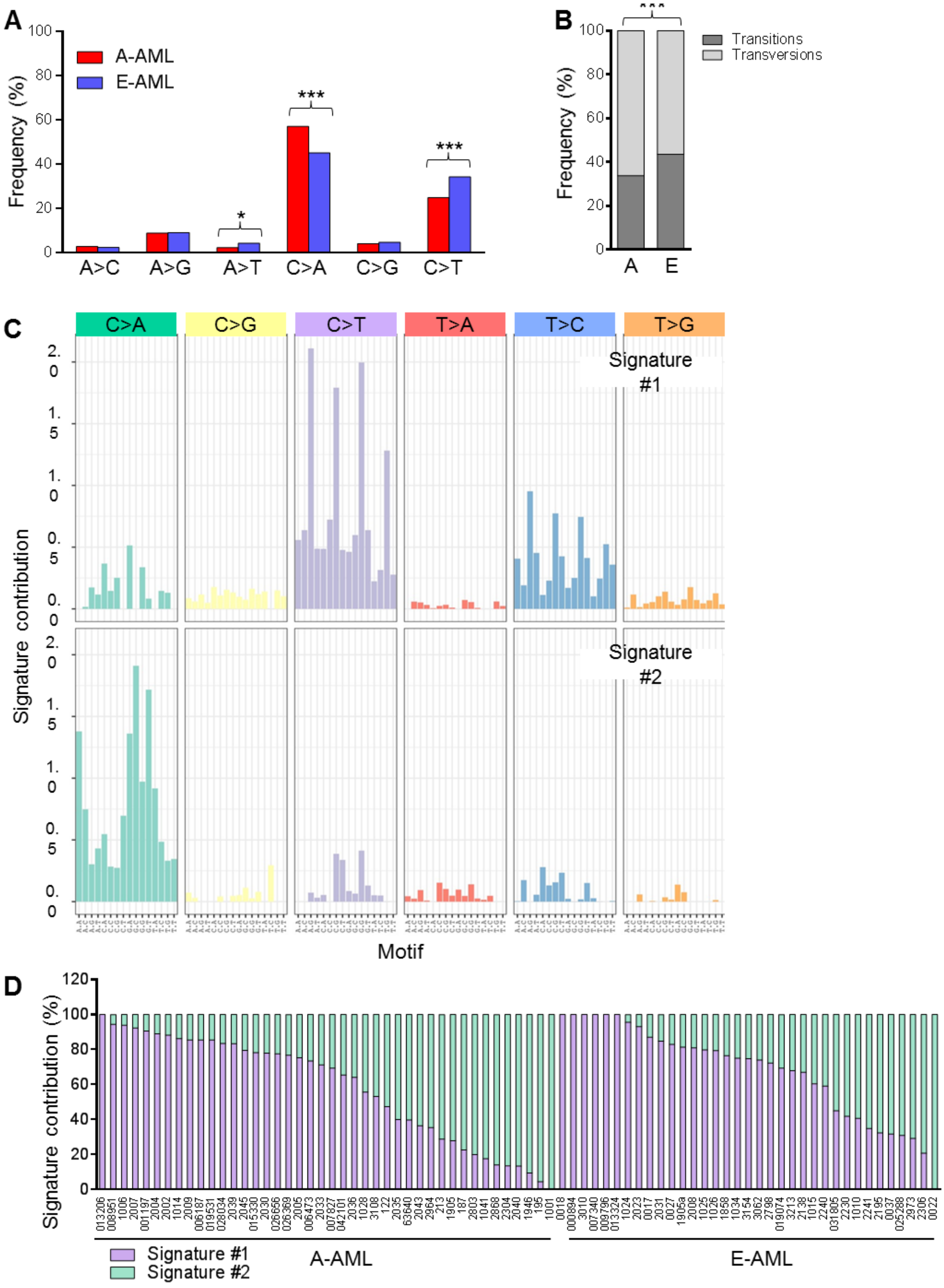


Figure 4. Mutational signatures in A-AML and E-AML. (A) Nucleotide targeting of non-silent mutations detected by WES (excluding SNPs) in A-AML and E-AML. Statistical significance was analyzed by Fisher's exact test (*, p <0.05; ***, p <0.001). (B) Percentage of transitions and transversions among non-silent mutations detected by WES (excluding SNPs) in A-AML and E-AML. Fisher's exact test was used for statistical significance

(*p* <0.001). (C) Mutational signatures according to the 96 substitution classification defined by the substitution class and sequence context immediately 5′ and 3′ to the mutated base. Mutation types are reported on the horizontal axes using different colors; the percentage of each specific mutation type is represented by vertical axes. Synonymous and nonsynonymous SNVs were considered in the analysis. (D) Contribution of the identified signatures to the mutational processes in A-ML and E-AML. Rows indicate cases displaying *DNMT3A* or *NF1* mutations.

CNAs in A-AML and E-AML

Data from SNP profiling were available for 38/42 A-AML and 32/35 E-AML patients. Considering whole chromosome and focal CNAs, A-AML was significantly associated with CN gains affecting cell cycle, nucleotide biosynthesis, glucose, carbohydrate and amino acid metabolism, bioenergetics pathways, protein assembly and degradation, response to reactive oxygen species (ROS), stem cell-related pathways, kinase signaling (Table S4). No differences occurred between CK-A-AML and non-CK-A-AML.

To prioritize CNAs with a putative role in the aneuploid phenotype, we considered gene gains and losses with a significantly different frequency (according to Fisher's exact test) among the two cohorts and excluded the events simply caused by whole chromosome trisomy and monosomy. The remaining genes defined chromosome cytobands preferentially affected by CNAs in A-AML ($p \leq 0.05$, Table S5). These included minimal common regions in chromosomes frequently found as monosomic or trisomic in AML: gain of 8p11-p12, 8p21, 8q11-q13, 8q21-q24, 13q33, 21q21-q22 and loss of 5q14-q15, 5q21-q23, 5q31-q33, 7q21-22, 7q31-34. A-AML was associated with additional CNAs in regions rarely targeted by whole chromosome gain and loss: gain at 6p12, 6p21-p25, 6q13-q16, 9p13, 12p13, 20q11, 22q12 and loss at 11p13, 12p12-p13, 17p13 where genes involved in the aneuploid phenotype likely localize. In particular, loss of *TP53*, mapping at 17p13, was detected in seven A-AML cases (p=0.01) and the remaining allele was mutated in six of them. Moreover, six A-AML cases carried *NPM1* loss at 5q35 (p=0.03), while *NPM1* mutations were only detected in E-AML (Figure 2).

Among the CNAs discriminating A-AML and E-AML, we identified 40 genes gained and 23 genes in lost regions with a known role in AML pathogenesis, defining hotspots of CN gain localized at 6p22 and CN loss at 5q31 and 12p13 (p<0.05, Table S6). Genes located in the hotspot regions were involved in cell cycle regulation and DNA replication (gain of *E2F2*, loss of *KIF20A*, *CDKN1B*, *PURA*), double-strand break repair (loss of *RAD50*), chromatin organization (gain of *DEK*, loss of *KDM3B*), regulation of leukemia stem-cell phenotype and differentiation (gain of *SOX4*, loss of *TIFAB*, *CTNNA1*). For the selected genes, we also computed the frequency of CN events co-occurring in the two cohorts. A-AML and E-AML shared the co-occurrence of CN loss at 5q with CN gain of the *MYB* oncogene or the tyrosine protein kinase *JAK2* (Figure 5 A-B), although the frequency of these events was higher in A-AML. Moreover, in A-AML loss at 5q was frequently combined with *TP53* loss, as previously reported [46], with loss of *CDKN1B* and the hematopoietic gene *ETV6* (chromosome 12), and with CN gain of the regulators of hematopoiesis *RUNX1* and *ERG* (chromosome 21), which frequently co-occurred, as 5q loss, with gain of *MYC* or of 8q (Figure 5B).

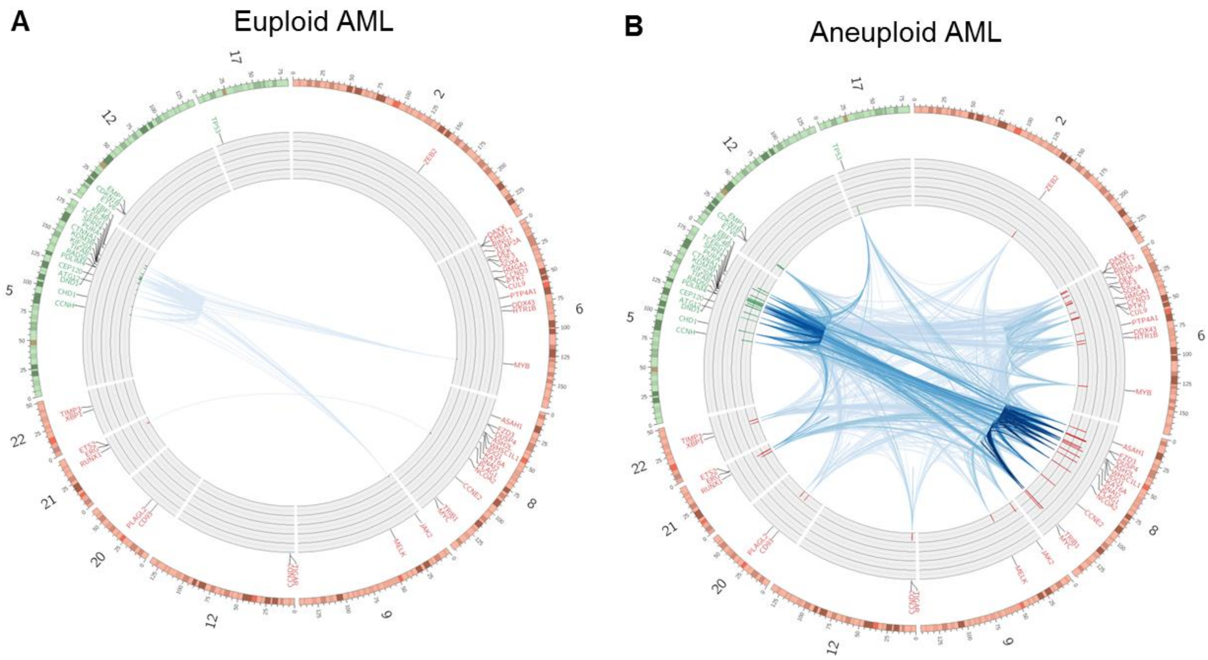


Figure 5. Frequency and co-occurrence of CNAs in leukemia-related genes in E-AML and A-AML. The Circos plots depict AML-related genes that were associated to the aneuploid phenotype for CN gains/duplications (in red) and loss/deletions (in green) in euploid (A) and aneuploid (B) AML. The barplots on the periphery represent the percentage of patients with CN events in each gene (on a 0-100% scale). Links connect CNAs co-occurring in the same patient; the intensity of a link's color reflects the absolute frequency of patients harboring that co-occurrence (min=1; max=17). Mutually exclusive alterations may exist in areas that are not connected.

Networks of genomic events characterizing A-AML and E-AML

We asked whether, overall, the genomic events including mutations and CNAs presented with differential frequency across GO-BP pathways in A-AML and E-AML. For both cohorts, we built networks in which the nodes and links represented the pathways and number of patients with enrichment of the two pathways, respectively. By considering links with weight ≥2, we identified in A-AML and E-AML 165 and 48 nodes and 4768 and 281 edges, respectively. Genomic alterations in A-AML targeted genes derived from many pathways (Figure S6A), while they occurred in a more restricted way in E-AML patients (Figure S6B). High impact genomic alterations may disrupt many pathways at once, and those most concomitantly affected are highlighted by high degree values in the networks. Regulation of hematopoiesis and myeloid cell differentiation were ranked among the top disease-related pathways, according to their degree and betweenness centrality both in A-AML and E-AML (Table S7). Moreover, A-AML was characterized by alterations affecting DNA replication-dependent nucleosome organization and assembly, and leucocyte differentiation, while regulation of Smoothened (SMO) signaling pathway and cell-matrix adhesion distinguished the euploid network.

Deregulated expression of leukemia-specific and aneuploidy-related genes in A-AML

To identify transcriptional properties contributing to the aneuploid phenotype in AML, we performed GEP of 22 aneuploid (characterized by different types of whole chromosome gain and loss) and 27 euploid cases (normal karyotype, Table S1). Principal component analysis showed no differences in the A-AML cohort according to karyotypic complexity (data not shown). Supervised analysis identified differential expression of 204 coding genes (56 up- and 148 downregulated) between A-AML and E-AML. We detected increased transcript level of *CDKN2C*, *MCM2* and *PLK1* and decreased expression of *HINT1* and *HOXB5*, which were also identified in a previous A-AML microarray dataset [47], along with overexpression of genes associated with chromosomal instability in solid tumors (*MCM2*, *CDC20* and *UBE2C*)[48].

A panel of genes was related to AML pathogenesis (Figure 6A). These include transcription factors as the *KMT2A* partner *MLLT10*, HOX family members (*HOXA3/5/6/7/8/10*, *HOXB3/5*, *MEIS1*, *NKX2-3*) and the regulator of DNA hydroxymethylation *WT1*, which showed lower expression in A-AML. *RUNX3* and the WNT-related gene *FRAT2* were upregulated in A-AML. Additional signaling molecules showed reduced expression in A-AML, including the inositol 1,4,5-trisphosphate receptor *ITPR2*, the leukemia stem-cell marker *CD47*, the

CALM/AF10-related gene *COMMD3* and the RAS pathway genes *BRAF*, *PIK3CB*, *SOS1*, *PIK3C3*, suggesting that distinct molecular mechanisms drive A-AML and E-AML.

A particularly relevant finding with regard to the aneuploid phenotype was the enrichment of upregulated genes with known functions in protein modification, ubiquitination, metabolic processes and telomere maintenance (Figure 5B and Table S8), coupled to the downregulation of genes involved in macromolecule biosynthesis and nucleic acid metabolic process (Figure 6B and Table S9). Such profile is indicative of A-AML cells attempt to face the unfavorable aneuploid condition by managing the unbalanced protein load and by controlling the proliferation rate. Indeed, A-AML cases had a significantly lower white blood cell (WBC) count compared with E-AML both in our cohort (median value: 7.1×10^9/L in A-AML *vs.* 15.6×10^9/L in E-AML, p=0.038, Table S1) and in the TCGA dataset (median value: 10.0×10^9/L in A-AML *vs.* 29.5×10^9/L in E-AML, p=0.02).

Transcriptomic signatures of A-AML

A significant fraction of the differentially expressed genes was involved in cell cycle and DNA repair (Figure 6A). These included the DNA damage sensors *ATR* and *RAD50* and its interacting protein RINT1 and the positive regulators of the p53-mediated program *DMTF1* and *HINT1*, which were downregulated in A-AML. Moreover, A-AML showed deregulated expression of ubiquitin-related genes involved in cell cycle progression (Figure 7A and Table S8): reduced levels of the ubiquitin-activating enzyme *UBA3* and upregulation of *CCNF*, a subunit of the SCFs complex, of the ubiquitin ligase *UHRF*1, the ubiquitin-conjugating enzyme *UBE2C* and of *CDC20*, which regulate APC/C activity during metaphase to anaphase transition. We sought to identify a suitable transcriptomic signature of A-AML, with therapeutic potentials. By combining computational analysis and biological significance, we defined a 3-gene signature composed of overexpressed *PLK1* and *CDC20* and downregulated *RAD50*, which discriminated 73% of patients between the A-AML and E-AML cohorts. *RAD50* downregulation (2-fold, p=0.041), *PLK1* and *CDC20* upregulation (2.5-fold and 3.7-fold, p=0.024 and p=0.004, respectively) were confirmed at protein level (Figure 7C-D), indicating that a multi-step process, involving different cell cycle phases is finely tuned in A-AML.

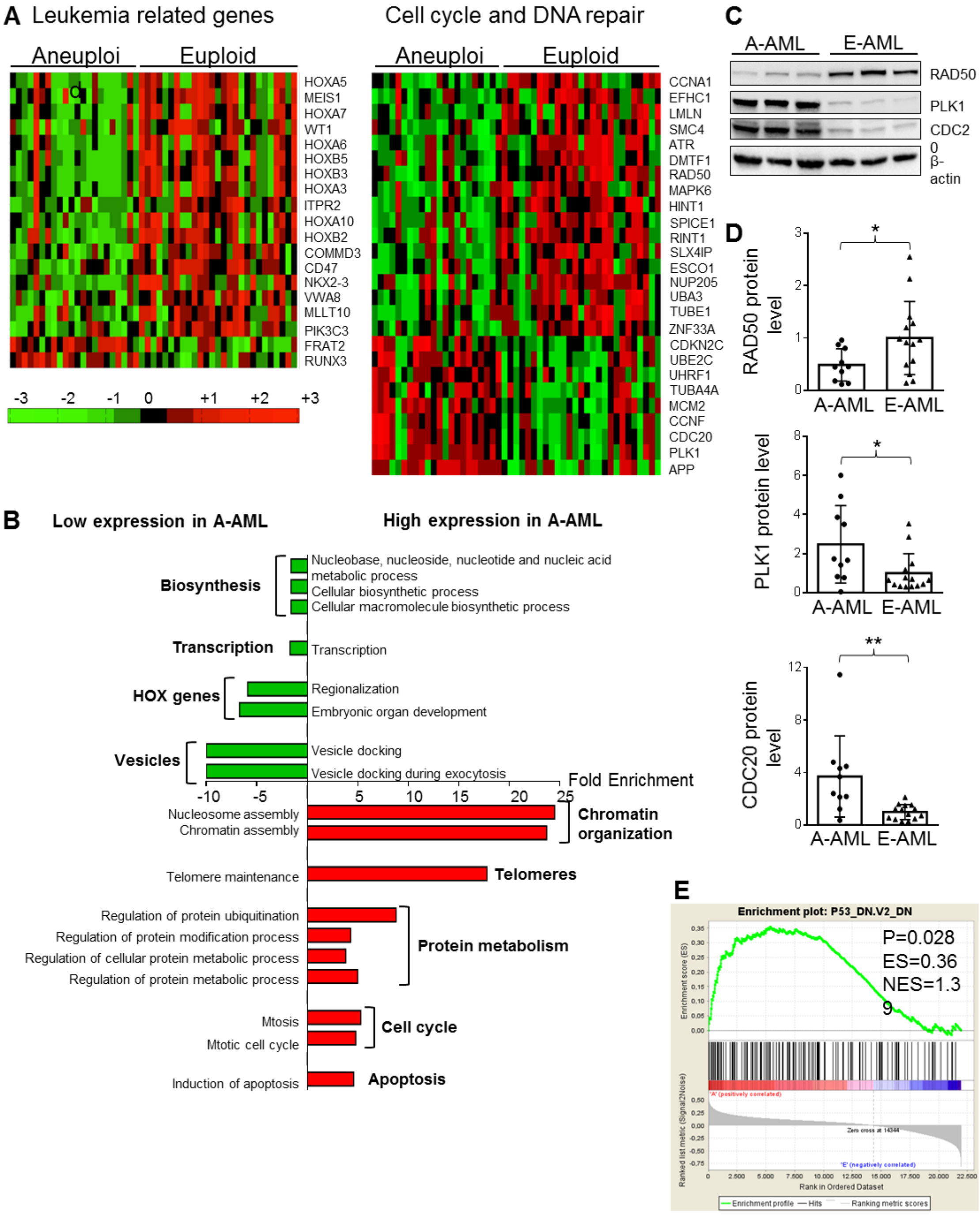


Figure 6. GEP analysis of A-AML and E-AML. (A) Gene expression differences in leukemia-related and cell cycle- and DNA repair-related genes between A-AML and E-AML were determined by supervised analysis (n=22, A-AML, n=27, E-AML). Visualized data are standardized through z-score transform; color changes within a row indicate expression levels relative to the mean and rescaled on transcript standard deviation (red: upregulated, green: downregulated). (B) Biological processes significantly enriched among differentially expressed genes in A-AML *vs.* E-AML according to David analysis (p<0.05). (C-D) Downregulation of *RAD50* and upregulation of *PLK1* and *CDC20* in A-AML determined at protein level. (C) Western blot of three representative cases of each cohort. Statistical significance was determined by Student's t test (*, p <0.05; **, p <0.01; ***, p <0.001). (D) Densitometry after normalization for the mean value across all E-AML cases, with β-actin serving as control; (E) Signature of p53-downregulation in A-AML identified by GSEA.

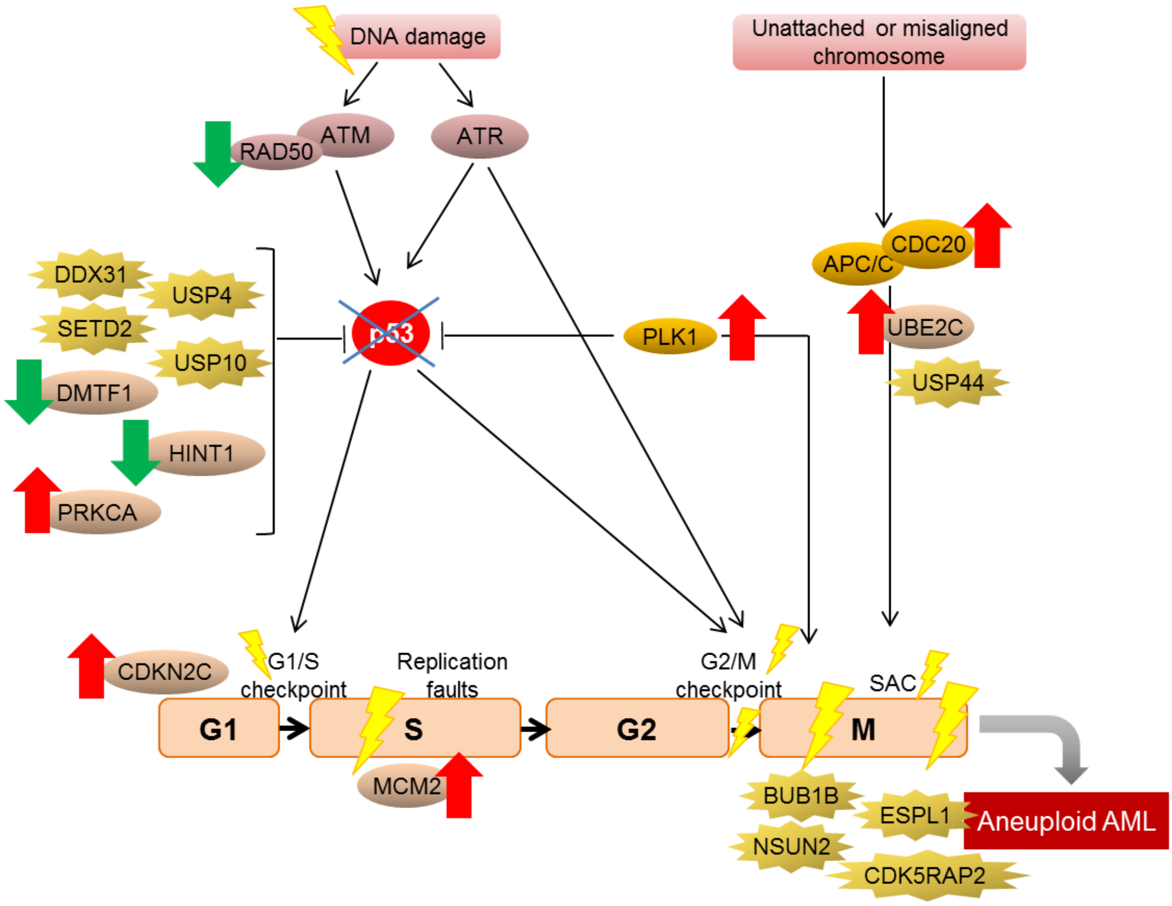
GSEA identified a significant association of A-AML with a gene expression signature of p53-deficiency (NES=1.39, *p*=0.03, Figure 7E). This finding was particularly relevant, since the A-AML cases analyzed by GEP included 27% of MK-AML/CK-AML, and we expected an overall rate of *TP53* abnormalities of ≈16% in our A-AML cohort [46]. To verify this hypothesis, we screened the mutational hotspots of the*TP53* gene by Sanger sequencing. Four out of 22 patients (18%) carried *TP53* genomic alterations (2 mutations and 2 chromosome 17 monosomies, Table S10). The signature enrichment remained significant by excluding these cases from the A-AML cohort (Figure S7), thus indicating that a general mechanism of down-modulation of the p53-related transcriptional program cooperates with structural abnormalities to silence the p53 pathway.

Discussion

The study of aneuploidy is of clinical and biological relevance in AML, since more than 20% of cases display numerical chromosome aberrations [20]. However, few studies have so far focused on the entire coding genome of a limited number of aneuploid cases [23,31,32,49,50].

To shed light into the molecular processes associated with A-AML, we integrated WES, CNA and GEP analysis of a large A-AML cohort and compared it with E-AML cases.

Aneuploidy associates with genomic instability in AML, as previously observed in solid tumors [51,52]. We found an average number of coding mutations of 26 and 15 per sample in A-AML and E-AML, respectively, by integrating two variant calling tools, a recently suggested strategy to improve cancer genome analysis [53]. Besides patients' age, which is related to the total number of mutations in AML [54], the stress caused by the aneuploid condition may provide a selective pressure towards accumulation of further genetic lesions leading to phenotypic changes which enable cells to tolerate chromosome imbalances[63,66]. The reduced number of circulating WBC in A-AML compared with E-AML, reflects the proliferative disadvantage of aneuploid leukemic cells and points to a more quiescent and stem-cell-like state [18]. This may contribute to drug resistance in poor prognosis aneuploid patients [55]. The overall mutational spectrum of A-AML and E-AML was dominated by C>T base substitutions and by transitions, as previously reported [43,44] and by signature #1, which is dominant in AML and other solid tumors [43]. However, compared with euploid cases, A-AML showed, among non-synonymous SNVs (excluding SNPs), a higher prevalence of C>A transversions, which participate to defective DNA mismatch repair or reactive oxygen species-related signatures [56]. This evidence



indicates that C>A transversions are not only related to chemotherapy and disease relapse [44,57], as our A-AML and E-AML cohorts did not

Figure 7. Mechanisms potentially inducing and supporting aneuploidy in AML. Model incorporating the genomic and transcriptomic results.

differ in terms of prevalence of secondary, therapy-related and relapsed cases (Table 1). Moreover, the association between *DNMT3A* mutations and signature #2 in A-AML suggests that an altered DNA methylation landscape may reduce the rate of spontaneous deamination of 5-methyl-cytosine [58] and favor aneuploidy, as a consequence of DNA hypomethylation at the centromere [59]. Larger patient cohorts and analysis of the effect of the single mutations are needed to confirm this observation.

The integration of the genomic and transcriptomic patterns characterizing A-AML points to cellular functions with a potential causal role in aneuploid leukemia, including deregulation of cell cycle-related processes occurring inside or outside of mitosis [60,61] (Figure 7). Moreover, our GEP data indicate that aneuploidy shapes the transcriptional profile of leukemic cells, by affecting not only the expression of genes located on trisomic or monosomic chromosomes [62], but also of a set of genes, which is independent of the identity of the individual chromosomes, as observed in different aneuploid models [8,14-16].The observed lesions may promote genomic instability, hamper cell cycle checkpoints and force its progression [63]. Evidence is available in the literature for some genes, including *BUB1B* [64], *NSUN2* [65], *ESPL1* [3], *CDK5RAP2* [66], *NDC1* [67], *USP44* [68], which were mutated in A-AML and *NPM1*[69], targeted by CN loss. However, dysregulation of most mitotic checkpoint genes does not induce spontaneous tumorigenesis and their cooperation to the A-AML phenotype remains to be confirmed. Among the mutated genes, the tumor suppressor *TP53* has been associated with A-AML [70] and CK-AML [46]. We show here that the p53 transcriptional program is generally silenced in A-AML either through structural or functional inactivation, which can be mediated by a number of events [71], including mutations of the p53 regulators *SETD2*, *DDX31*, *USP10*, *USP4*, by decreased expression of *DMTF1* and *HINT1* or increased levels of *PRKCA*. Reduced expression of *RAD50*, suggestive of an impaired DNA damage response and checkpoint arrest, and upregulation of *PLK1* may hamper p53 activation in A-AML, while overexpression of *CDC20*, could help bypassing the spindle-assembly checkpoint (Figure 7).

The protein ubiquitination and degradation pathway are deregulated or targeted by genomic abnormalities in A-AML, along with genes involved in response to ROS, as previously reported in aneuploid models [15,72]. These alterations may help face the accumulation of ROS causing oxidative DNA damage [73] and the unbalanced gene dosage induced by aneuploidy, that leads to heightened energy metabolism [8], supported by CN gains of genes involved in glucose uptake and catabolism and in the biosynthetic processes.

Our findings unravelling the molecular basis of A-AML may be relevant to the design of *ad hoc* therapies. Microtubule depolymerizing drugs and PLK1 inhibitors, which showed synthetic lethal interaction [74], targeting of centrosome clustering and chromosomal instability through kinetochores (e.g. Aurora kinase inhibitors) and chemical inhibition of the APC/C, either alone or when combined with topoisomerase poisons [75] or defective sister chromatid cohesion [76] represent candidate strategies. Additional approaches may take advantage of aneuploid cells dependency on chaperone pathways, protein turnover and forced metabolism, thus suggesting proteasome, protease or glycolysis [77] inhibition as potential synthetic lethal strategies under the aneuploid condition.

Acknowledgments

We thank all the members of the "Next Generation Sequencing platform for targeted Personalized Therapy of Leukemia" - NGS-PTL consortium: Paolo Garagnani (University of Bologna, Bologna, Italy) for the effort in the NGS-PTL project, Massimo Delledonne (Personal Genomics, Verona, Italy), Alberto Ferrarini and Marianna Garonzi (University of Verona, Verona, Italy), Nadine Vincent and Julien Schira (Fasteris, Plan-les-Ouates, Switzerland) for sequencing service and advice, Marco Capoferri and Mirko Fortuna (Sinaptica IT, Rome, Italy) for database curation, Raffaele Calogero (University of Torino, Torino, Italy), Clelia Tiziana Storlazzi (University of Bari, Bari, Italy), Jiri Mayer and Sarka Pospisilova (Masaryk University, Brno, Czech Republic) for discussion. We also thank Stefano A. Pileri (University of Bologna, Bologna, Italy) and Guido Biasco ("Giorgio Prodi" Cancer Research Center, University of Bologna, Bologna, Italy) for access to MiSeq and HiScan instrument, respectively.

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n° 306242-NGS-PTL. The study was supported by Associazione Italiana per la Ricerca sul Cancro (AIRC) Investigator grant, n. 19226 to Giovanni Martinelli, Investigator Grant n. 15762 and AIRC 5x1000, n. 10007 to Stefano A. Pileri; Programma di ricerca Regione-Università 2010-2012 (L. Bolondi). Lars Bullinger was supported in part by the Deutsche Forschungsgemeinschaft (Heisenberg-Professur BU 1339/8-1).

References

1. Ganem NJ, Godinho SA, Pellman D. A mechanism linking extra centrosomes to chromosomal instability. *Nature*. 2009;460(7252):278-282.

2. Hassold T, Abruzzo M, Adkins K, et al. Human aneuploidy: incidence, origin, and etiology. *Environmental and molecular mutagenesis*. 1996;28(3):167-175.

3. Zhang N, Ge G, Meyer R, et al. Overexpression of Separase induces aneuploidy and mammary tumorigenesis. *Proc Natl Acad Sci U S A*. 2008;105(35):13033-13038.

4. Sotillo R, Schvartzman JM, Socci ND, Benezra R. Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. *Nature*. 2010;464(7287):436-440.

5. Bakhoum SF, Thompson SL, Manning AL, Compton DA. Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nat Cell Biol*. 2009;11(1):27-35.

6. Pampalona J, Soler D, Genesca A, Tusell L. Whole chromosome loss is promoted by telomere dysfunction in primary cells. *Genes Chromosomes Cancer*. 2010;49(4):368-378.

7. Niwa O, Tange Y, Kurabayashi A. Growth arrest and chromosome instability in aneuploid yeast. *Yeast*. 2006;23(13):937-950.

8. Torres EM, Sokolsky T, Tucker CM, et al. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. *Science*. 2007;317(5840):916-924.

9. Lindsley DL, Sandler L, Baker BS, et al. Segmental aneuploidy and the genetic gross structure of the Drosophila genome. *Genetics*. 1972;71(1):157-184.

10. McClintock B. A Cytological and Genetical Study of Triploid Maize. *Genetics*. 1929;14(2):180-222.

11. Singh K, Multani DS, Khush GS. Secondary trisomics and telotrisomics of rice: origin, characterization, and use in determining the orientation of chromosome map. *Genetics*. 1996;143(1):517-529.

12. Gropp A, Winking H, Herbst EW, Claussen CP. Murine trisomy: developmental profiles of the embryo, and isolation of trisomic cellular systems. *J Exp Zool*. 1983;228(2):253-269.

13. Segal DJ, McCoy EE. Studies on Down's syndrome in tissue culture. I. Growth rates and protein contents of fibroblast cultures. *J Cell Physiol*. 1974;83(1):85-90.

14. Williams BR, Prabhu VR, Hunter KE, et al. Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. *Science*. 2008;322(5902):703-709.

15. Sheltzer JM, Torres EM, Dunham MJ, Amon A. Transcriptional consequences of aneuploidy. *Proc Natl Acad Sci U S A*. 2012;109(31):12644-12649.

16. Upender MB, Habermann JK, McShane LM, et al. Chromosome transfer induced aneuploidy results in complex dysregulation of the cellular transcriptome in immortalized and cancer cells. *Cancer research*. 2004;64(19):6941-6949.

17. Dephoure N, Hwang S, O'Sullivan C, et al. Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. *Elife*. 2014;3:e03023.

18. Pfau SJ, Silberman RE, Knouse KA, Amon A. Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues in vivo. *Genes Dev*. 2016;30(12):1395-1408.

19. Mitelman F, Johansson B, Mertens F. Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer; 2017.

20. Grimwade D, Hills RK, Moorman AV, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood*. 2010;116(3):354-365.

21. Breems DA, Van Putten WL, De Greef GE, et al. Monosomal karyotype in acute myeloid leukemia: a better indicator of poor prognosis than a complex karyotype. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2008;26(29):4791-4797.

22. Parkin B, Ouillette P, Li Y, et al. Clonal evolution and devolution after chemotherapy in adult acute myelogenous leukemia. *Blood*. 2013;121(2):369-377.

23. Herold T, Metzeler KH, Vosberg S, et al. Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. *Blood*. 2014;124(8):1304-1311.

24. Farag SS, Archer KJ, Mrozek K, et al. Isolated trisomy of chromosomes 8, 11, 13 and 21 is an adverse prognostic factor in adults with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B 8461. *International journal of oncology*. 2002;21(5):1041-1051.

25. Thompson SL, Compton DA. Proliferation of aneuploid human cells is limited by a p53-dependent mechanism. *J Cell Biol*. 2010;188(3):369-381.

26. Yuen KW, Warren CD, Chen O, Kwok T, Hieter P, Spencer FA. Systematic genome instability screens in yeast and their potential relevance to cancer. *Proc Natl Acad Sci U S A*. 2007;104(10):3925-3930.

27. Weaver BA, Silk AD, Montagna C, Verdier-Pinard P, Cleveland DW. Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell*. 2007;11(1):25-36.

28. Dicker F, Haferlach C, Kern W, Haferlach T, Schnittger S. Trisomy 13 is strongly associated with AML1/RUNX1 mutations and increased FLT3 expression in acute myeloid leukemia. *Blood*. 2007;110(4):1308-1316.

29. Alpermann T, Haferlach C, Eder C, et al. AML with gain of chromosome 8 as the sole chromosomal abnormality (+8sole) is associated with a specific molecular mutation pattern including ASXL1 mutations in 46.8% of the patients. *Leuk Res*. 2015;39(3):265-272.

30. Virtaneva K, Wright FA, Tanner SM, et al. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proc Natl Acad Sci U S A*. 2001;98(3):1124-1129.

31. Jerez A, Sugimoto Y, Makishima H, et al. Loss of heterozygosity in 7q myeloid disorders: clinical associations and genomic pathogenesis. *Blood*. 2012;119(25):6109-6117.

32. McNerney ME, Brown CD, Wang X, et al. CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia. *Blood*. 2013;121(6):975-983.

33. Zuffa E, Franchini E, Papayannidis C, et al. Revealing very small FLT3 ITD mutated clones by ultra-deep sequencing analysis has important clinical implications in AML patients. *Oncotarget*. 2015;6(31):31284-31294.

34. McGowan-Jordan J, Simons A, Schmid M. ISCN(2016): An International System for Human Cytogenomic Nomenclature. *Karger Publishers*. 2016.

35. Team RC. R: A language and environment for statistical computing. 2016.

36. Carlson M. org.Hs.eg.db: Genome wide annotation for Human. *R package version 341*. 2016.

37. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-287.

38. Carlson M. GO.db: A set of annotation maps describing the entire Gene Ontology. *R package version 340*. 2016.

39. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57(1):289-300.

40. Gene Ontology C. Gene Ontology Consortium: going forward. *Nucleic Acids Res*. 2015;43(Database issue):D1049-1056.

41. Zhang B, Kirov S, Snoddy J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res*. 2005;33(Web Server issue):W741-748.

42. Dohner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults:

2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447.

43. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.

44. Garg M, Nagata Y, Kanojia D, et al. Profiling of somatic mutations in acute myeloid leukemia with FLT3-ITD at diagnosis and relapse. *Blood*. 2015;126(22):2491-2501.

45. Alexandrov LB, Jones PH, Wedge DC, et al. Clock-like mutational processes in human somatic cells. *Nat Genet*. 2015;47(12):1402-1407.

46. Rucker FG, Schlenk RF, Bullinger L, et al. TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. *Blood*. 2012;119(9):2114-2121.

47. Neben K, Tews B, Wrobel G, et al. Gene expression patterns in acute myeloid leukemia correlate with centrosome aberrations and numerical chromosome changes. *Oncogene*. 2004;23(13):2379-2384.

48. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet*. 2006;38(9):1043-1048.

49. Larsson N, Lilljebjorn H, Lassen C, Johansson B, Fioretos T. Myeloid malignancies with acquired trisomy 21 as the sole cytogenetic change are clinically highly variable and display a heterogeneous pattern of copy number alterations and mutations. *Eur J Haematol*. 2012;88(2):136-143.

50. McNerney ME, Brown CD, Peterson AL, et al. The spectrum of somatic mutations in high-risk acute myeloid leukaemia with -7/del(7q). *Br J Haematol*. 2014;166(4):550-556.

51. Sheltzer JM, Blank HM, Pfau SJ, et al. Aneuploidy drives genomic instability in yeast. *Science*. 2011;333(6045):1026-1030.

52. Solomon DA, Kim T, Diaz-Martinez LA, et al. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science*. 2011;333(6045):1039-1043.

53. Bodini M, Ronchini C, Giaco L, et al. The hidden genomic landscape of acute myeloid leukemia: subclonal structure revealed by undetected mutations. *Blood*. 2015;125(4):600-605.

54. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012;150(2):264-278.

55. Kusumbe AP, Bapat SA. Cancer stem cells and aneuploid populations within developing tumors are the major determinants of tumor dormancy. *Cancer Res*. 2009;69(24):9245-9253.

56. Nagarajan N, Bertrand D, Hillmer AM, et al. Whole-genome reconstruction and

mutational signatures in gastric cancer. *Genome Biol*. 2012;13(12):R115.

57. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506-510.

58. Abdel-Wahab O, Levine RL. Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia. *Blood*. 2013;121(18):3563-3572.

59. Costa G, Barra V, Lentini L, Cilluffo D, Di Leonardo A. DNA demethylation caused by 5-Aza-2'-deoxycytidine induces mitotic alterations and aneuploidy. *Oncotarget*. 2016;7(4):3726-3739.

60. Ricke RM, van Deursen JM. Aneuploidy in health, disease, and aging. *J Cell Biol*. 2013;201(1):11-21.

61. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*. 2017;355(6322).

62. Schoch C, Kern W, Kohlmann A, Hiddemann W, Schnittger S, Haferlach T. Acute myeloid leukemia with a complex aberrant karyotype is a distinct biological entity characterized by genomic imbalances and a specific gene expression profile. *Genes Chromosomes Cancer*. 2005;43(3):227-238.

63. Janssen A, van der Burg M, Szuhai K, Kops GJ, Medema RH. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. *Science*. 2011;333(6051):1895-1898.

64. Hanks S, Coleman K, Reid S, et al. Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B. *Nat Genet*. 2004;36(11):1159-1161.

65. Hussain S, Benavente SB, Nascimento E, et al. The nucleolar RNA methyltransferase Misu (NSun2) is required for mitotic spindle stability. *J Cell Biol*. 2009;186(1):27-40.

66. Lizarraga SB, Margossian SP, Harris MH, et al. Cdk5rap2 regulates centrosome function and chromosome segregation in neuronal progenitors. *Development*. 2010;137(11):1907-1917.

67. Chial HJ, Giddings TH, Jr., Siewert EA, Hoyt MA, Winey M. Altered dosage of the Saccharomyces cerevisiae spindle pole body duplication gene, NDC1, leads to aneuploidy and polyploidy. *Proc Natl Acad Sci U S A*. 1999;96(18):10200-10205.

68. Zhang Y, Foreman O, Wigle DA, et al. USP44 regulates centrosome positioning to prevent aneuploidy and suppress tumorigenesis. *J Clin Invest*. 2012;122(12):4362-4374.

69. Grisendi S, Bernardi R, Rossi M, et al. Role of nucleophosmin in embryonic development and tumorigenesis. *Nature*. 2005;437(7055):147-153.

70. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis

in Acute Myeloid Leukemia. *N Engl J Med*. 2016;374(23):2209-2221.

71. Prokocimer M, Molchadsky A, Rotter V. Dysfunctional diversity of p53 proteins in adult acute myeloid leukemia: projections on diagnostic workup and therapy. *Blood*. 2017;130(6):699-712.

72. Torres EM, Dephoure N, Panneerselvam A, et al. Identification of aneuploidy-tolerating mutations. *Cell*. 2010;143(1):71-83.

73. Li M, Fang X, Baker DJ, et al. The ATM-p53 pathway suppresses aneuploidy-induced tumorigenesis. *Proc Natl Acad Sci U S A*. 2010;107(32):14188-14193.

74. Hugle M, Belz K, Fulda S. Identification of synthetic lethality of PLK1 inhibition and microtubule-destabilizing drugs. *Cell Death Differ*. 2015;22(12):1946-1956.

75. Eguren M, Alvarez-Fernandez M, Garcia F, et al. A synthetic lethal interaction between APC/C and topoisomerase poisons uncovered by proteomic screens. *Cell Rep*. 2014;6(4):670-683.

76. de Lange J, Faramarz A, Oostra AB, et al. Defective sister chromatid cohesion is synthetically lethal with impaired APC/C function. *Nat Commun*. 2015;6:8399.

77. Liu LL, Long ZJ, Wang LX, et al. Inhibition of mTOR pathway sensitizes acute myeloid leukemia cells to aurora inhibitors by suppression of glycolytic metabolism. *Mol Cancer Res*. 2013;11(11):1326-1336.

Supplementary Figures

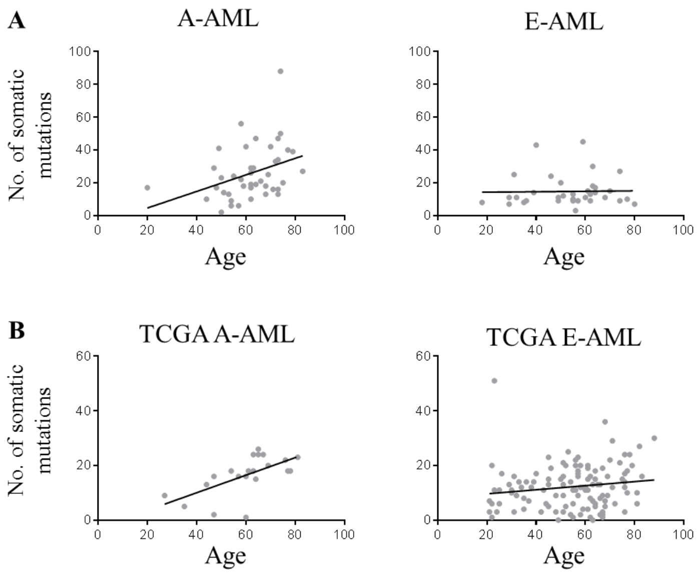


Figure S1. Age-dependent increase of the mutation load in A-AML patients. (A) Linear regression analysis showing the mutation load dependency on age only in the A-AML cohort. (B) Linear regression analysis on the TCGA AML cohort confirming the age-dependent increase in the mutation load specifically in A-AML.

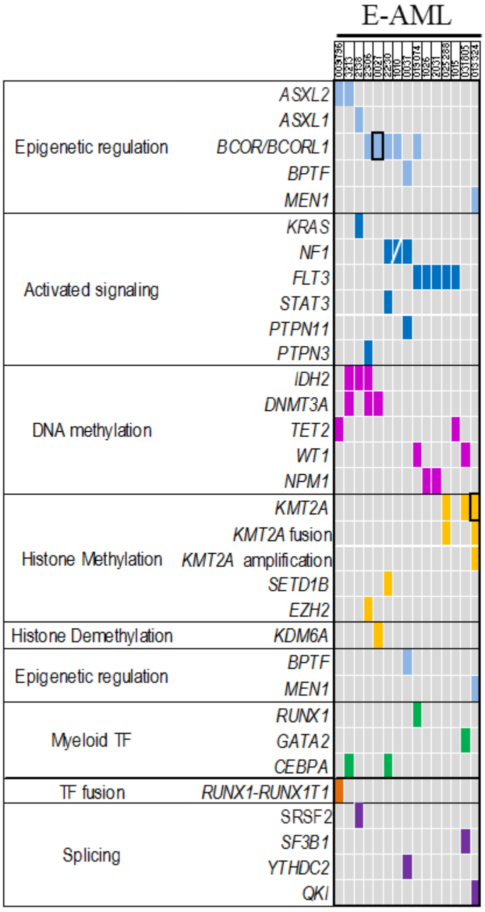


Figure S2. Somatic mutations preferentially detected in E-AML. Pattern of genomic lesions in E-AML cases carrying mutations not detected in the A-AML cohort (BCOR/BCORL1, NF1, TET2, NPM1, KMT2A, CEBPA). Each row denotes one or more specific gene(s); columns represent (left to right): functional categories, mutated genes/group of genes or other genomic alterations, single patients. Colours indicates functional categories; bars indicate CN loss; squares indicate CN gains.

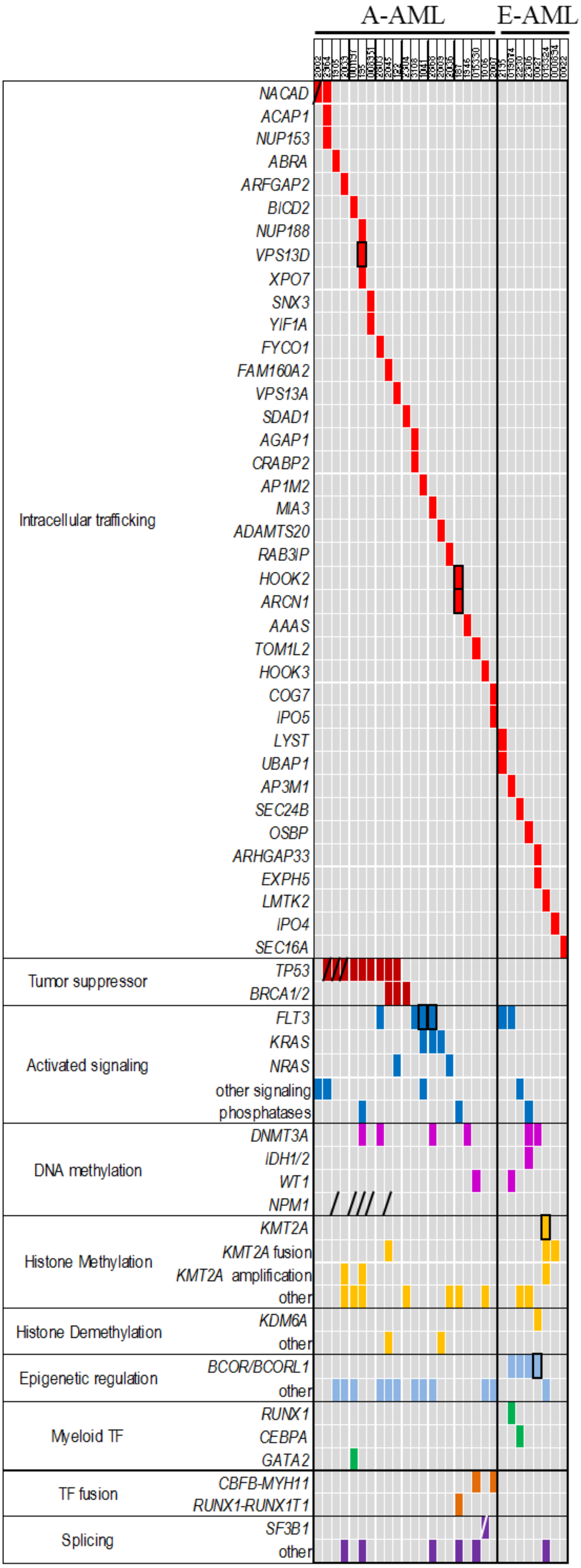


Figure S3. Pattern of genomic lesions in A-AML and E-AML cases carrying mutations in genes involved in intracellular trafficking. Each row denotes one or more specific gene(s); columns represent (left to right): functional categories, mutated genes/group of genes or other genomic alterations, single patients. Colours indicates functional categories; bars indicate CN loss; squares indicate CN gains.

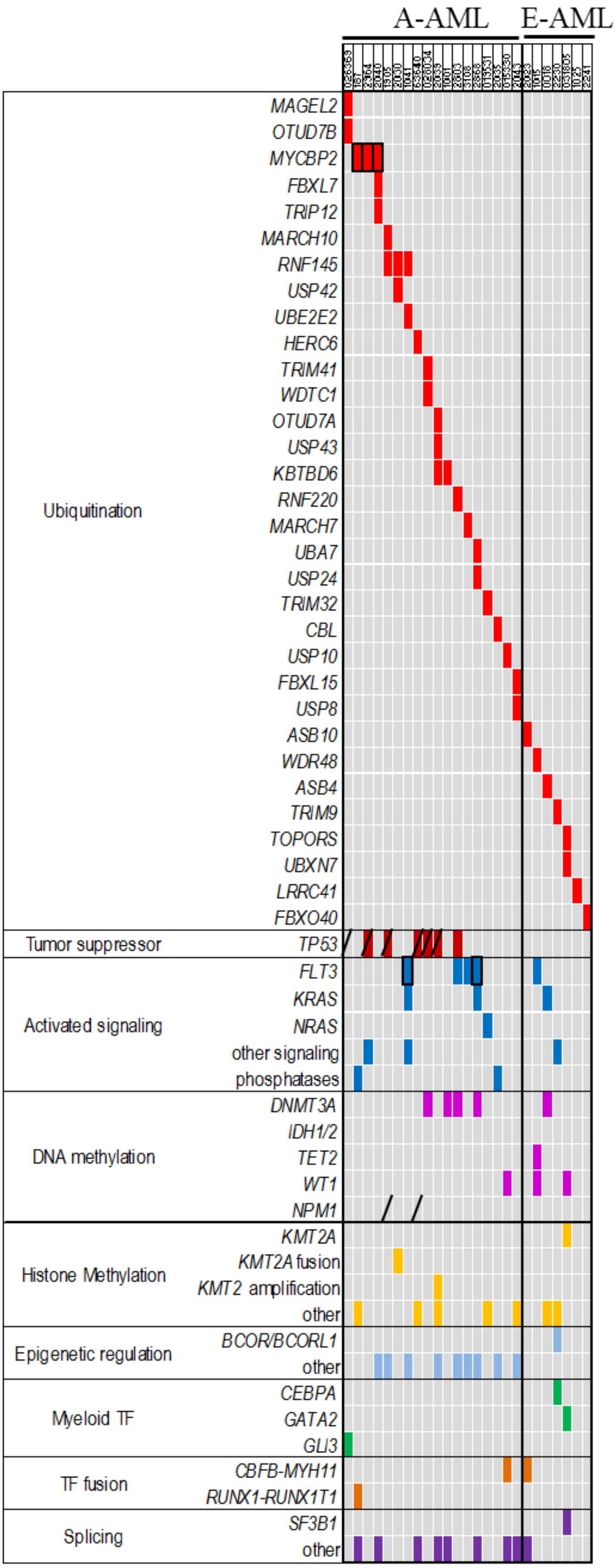


Figure S4. Pattern of genomic lesions in A-AML and E-AML cases carrying mutations in genes involved in protein ubiquitination. Each row denotes one or more specific gene(s); columns represent (left to right): functional categories, mutated genes/group of genes or other genomic alterations, single patients. Colours indicates functional categories; bars indicate CN loss; squares indicate CN gains.

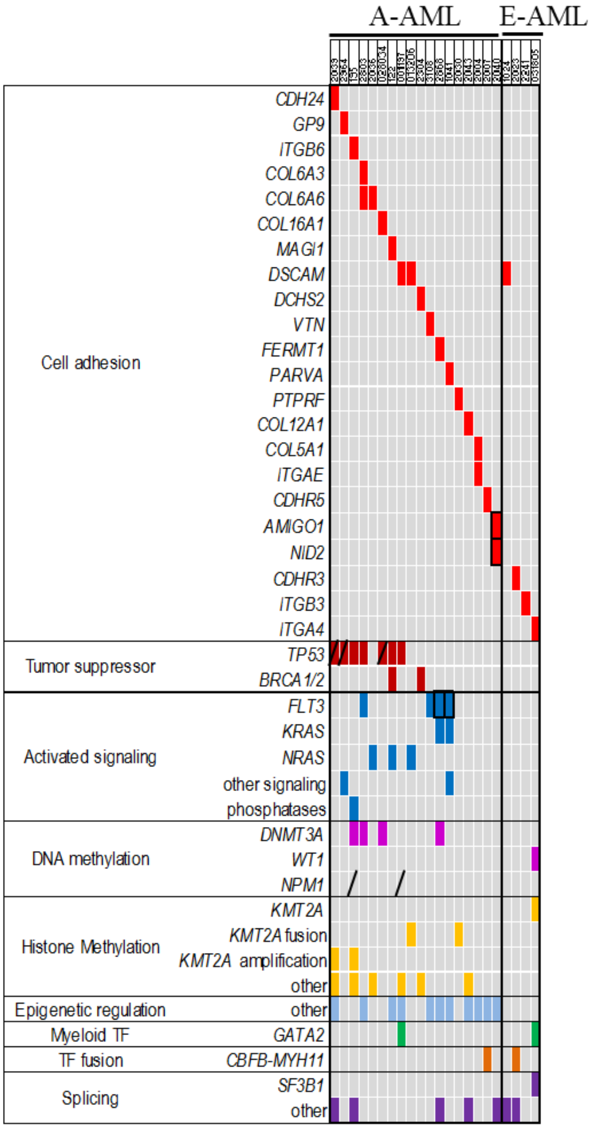


Figure S5. Pattern of genomic lesions in A-AML and E-AML cases carrying mutations in genes involved in cell adhesion. Each row denotes one or more specific gene(s); columns represent (left to right): functional categories, mutated genes/group of genes or other genomic alterations, single patients. Colors indicates functional categories; bars indicate CN loss; squares indicate CN gains.

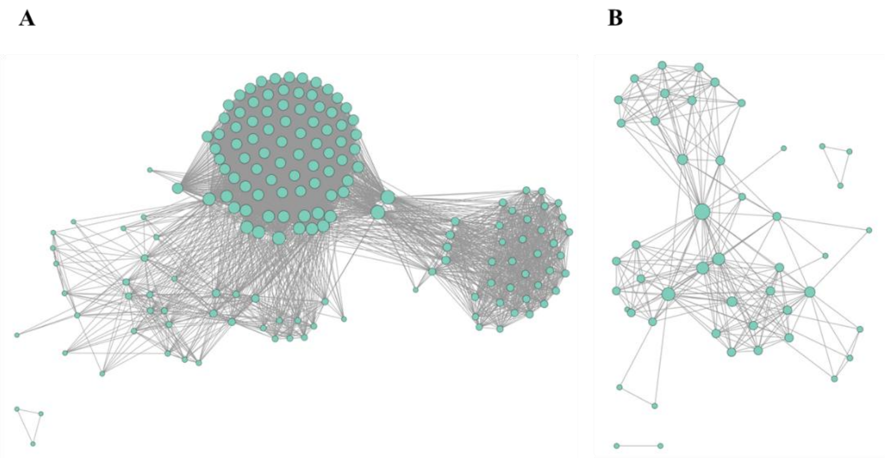


Figure S5. Overview of aneuploid (A) and euploid (B) networks. Each node is a GO-BP pathway and the links connect pathways enriched in least two patients.

Supplementary tables

Table S1. Aneuploid and Euploid AML patients' characteristics.

Patient ID	WES	SNP	GEP	Karyotype	De novo/sec /t-AML	FAB type	Diagnosis (D)/ Relapse (R)	Sex	Age	Sample type (BM/PB)	WBC count ($x10^9/L$)	Genetic group
Aneuploid AML												
001197	x	x		45-48,XX,idic(5)(q11),dic(6;22)(p11;p11),+r(6),del(8)(p11),+der(8)t(6;8)(?;p11)x1-2,der(8;12)dic(8;12)(p11;p11)t(12;15)(q24;q?)[15/15]	de novo	M2	D	F	51	BM	31.0	Adverse
006187	x	x		50,XY,+4,+14,+21,+22[5/8]/46,XY[3/8]	de novo	M1	D	M	59	BM	0.9	Adverse
008951	x	x		48,XX,der(3)t(3;5)(p11;p11),del(3)(q21q26),del(5)(q11q35),+der(5)del(5)(q14q34),+8,i(9)(p10)[7/19]/48,XX,der(3)t(3;5)(p11;p11),del(3)(q21q26),del(5)(q11q35),+der(5)del(5)(q14q34),der(7)t(5;7)(?;p11),+8,i(9)(p10),der(12)t(7;12)(p11;p11)[3/19]/46,XX[9/19]	sec (MDS)	M6	D	F	73	BM	1.6	Adverse
019531	x			45,XY,t(2;3)(p21;q26),-7,der(12)ins(12;7)(p13;q22q34)[19/23]/46,XY[4/23]	de novo	M0	D	M	53	BM	na	Adverse
026369	x	x		46,XX,-5,der(7)t(5;7)(?;q11),+der(8)del(8)(p11p23),+11,der(17)t(5;17)(?;p13),-18,-20,+der(22)t(5;22)(?;q13)[3/23]/47,XX,-5,+6,der(7)t(5;7)(?;q11),+del(8)(p11p23),+11,der(17)t(5;17)(?;p13),-18,-20,+der(22)t(5;22)(?;q13)[4/23]/46,XX[16/23]	de novo	M2	D	F	73	BM	1.0	Adverse
026656	x	x		45,X,-Y,t(2;21;8)(p25;q22;q22)[16/20]/46,XY[4/20]	de novo	M1	D	M	20	BM	6.4	Favorable
028034	x	x		46,X,-X,del(5)(q13q35),+6,-7,der(11)t(7;11)(q11;p11),del(8)(q11),der(16;17)(p10;q10),+22[9/11]/47,XX,del(5)(q13q35),+6,+8,der(16;17)(p10;q10)[2/11]	de novo	M2	D	F	58	BM	3.7	Adverse
015330	x			47,XY,t(3;6;3)(p26;p24;p21),+8,inv(16)(p13q22)[18/20]/46,XY[2/20]	de novo	M1	D	M	75	BM	na	Favorable
007827	x	x		45,X,-Y,t(8;21)(q22;q22)[9/20]/46,XY[11/20]	de novo	M2	D	M	62	BM	7.1	Favorable
013206	x	x		52,XY,+6,ins(6;11)(q27;q13q23),+8,+9,+13,+19,+21[21/21]	de novo	M1	D	M	44	BM	na	Adverse
006473	x	x		46,XY,inv(16)(p13q22)[14/21]/48,XY,inv(16)(p13q22),+8,+13[1/21]/93,XXYY,inv(16)(p13q22)x2,-2,+8,+8[3/21]/46,XY[3/21]	de novo	M4eo	D	M	62	BM	na	Favorable

042101	x			48,XX,+8,+22.ish der(16)ins(16;16)(q22;p13p13)[19/22]/46,XX[3/22]	de novo	M4eo	D	F	64	BM	na	Favorable
2803	x	x		47,XY,del(5)(q11q33),der(8)t(5;8)(q33;q24),+der(8)t(5;8)(q33;q24)[18/20]/46,XY[2/20]	sec (MDS)	M4	D	M	73	PBL	56.9	Adverse
2030	x	x	x	47,XX,+8,t(9;11)(p22;q23)[16/20]/48,XX,+8,+8,t(9;11)(p22;q23)[1/20]/50,XX,+8,+8,t(9;11)(p22;q23),+13,+19[3/20]	de novo	M5	D	F	62	BM	23.1	Adverse
37			x	45,X,-Y,t(8;21)(q22;q22)[11/20]/46,X,-Y,t(8;21),+mar(3/20)/46,XY[6/20]	de novo	M2	D	M	39	BM	na	Favorable
3108	x	x		46,XX[19/21]/47,XX,+4[2/21]	de novo	M5	D	F	69	BM	129.0	Intermediate
2004	x			48-50,XX,+2,-5,+6,der(8),del(9)(q22),der(11)t(8;11;17)(q12;q13;p13),+13,+der(13)del(13)(q14q32),+1~2 mar[18/19]/46,XX[1/19]	de novo	M2	D	F	55	BM	0.8	Adverse
63640	x	x		41-45,XX,t(1;2;?)(q21;q12;?),-3,t(5;11)(q13;p13),-7,add(8)(q24),-16,-17,-22,3~4mar[13/21]; 46,XX[8/21]	de novo	na	D	F	70	BM	1.0	Adverse
2002	x	x	x	45,XY,t(3;3)(q21;q26),-7[20/20]	t-AML	M0	D	M	68	PBL	5.2	Adverse
2007	x	x		47,XY,+8,inv(16)(p13q22)[9/20]/46,XY,t(9;17)(q34;q21),inv(16)(p13q22)[8/20]/46,XY,add(8)(q24),inv(16)(p13q22)[3/20]	t-AML	M2-M4	D	M	68	BM	14.0	Favorable
1014	x	x		47,XX,+4[3/20]/46,XX[17/20]	de novo	M4	D	F	48	PBL	4.7	Intermediate
69			x	47,XX,+X,i(17)(q10)[14/20]/46,XX[6/20]	de novo	M2	D	F	72	BM	35.1	Intermediate
2035	x	x		46,XY,t(3;12)(p22;q24),+4,-15,+mar[19/20]/46,XY[1/20]	de novo	na	R	M	62	BM	na	Adverse
2868	x	x		47,XX,del(5)(q13q33),+8[20/20]	de novo	na	D	F	72	BM	8.9	Adverse
187	x	x		74,XXX,t(8;21)(q22;q22)X2,-7,+8,,+9,+13,+16,+17,+19[9/10]/46,XX[1/10]	de novo	na	D	F	74	BM	1.7	Favorable
70			x	45,X,-Y,t(8;21)(q22;q22)[19/20]/46,XY[1/20]	de novo	M4	D	M	31	BM	5.1	Favorable
2964	x	x		44,XX,+der(3)t(3;20)(p12;p11),del(5)(q13q33),-7,-13,t(13;20)(q12;p11),-17,der(21)t(17;21)(q11;q22),+mar[14/20]/45,XX,t(1;16)(q12;q11),del(5)(q13q33),del(6)(q21q25),-7,add(22)(q13)[6/20]	sec (MDS)	sec (MDS)	D	F	60	BM	1.5	Adverse
2009	x	x		45,XX,t(3;14;16)(q21;q22;q22),add(7)(q34),-7,del(14)(q23q32),add(16)(q22),-21,+mar [16/20]/46,XY,t(3;14;16)(q21;q22;q22),add(7)(q34),der(7),del(14)(q23q32),add(16)(q22),-21,+mar[4/20]	de novo	M2	D	M	50	PBL	14.5	Adverse
5			x	45,XY,t(1;3;13)(p34;q26;q14),-7[18/20]/45,XY,t(1;3;13)(p34;q26;q14),-7, der(21)t(7;21)(q10;p10)[1/20]/46,XX[1/20]	de novo	na	D	F	68	BM	8.6	Adverse

54			x	47,XY,del(7)(q32q36),t(16;16)(p13;q22),+22[15/20]/48,XY,del(7)(q32q36),t(16;16)(p13;q22),+21,+22[2/20]/46,XY[3/20]	de novo	M4	D	M	61	BM	7.4	Favorable
21			x	48,XX,+4,+8[30/30]	de novo	M4	D	F	70	BM	na	Intermediate
13			x	44,XX,t(3;17)(p21;p13),del(5)(q13q33),-12,del(13)(q14),der(14)t(12;14)(p11p11),-18,add(21)(q13)[24/25]/46,XX[1/25]	na	M4	D	F	69	BM	na	Adverse
1006	x	x	x	47,XX,+4[13/20]/46,XX[7/20]	t-AML	M1	D	F	62	BM	13.4	Intermediate
1			x	47,XX,+21[16/20]/46,XX[4/20]	de novo	M0-M1	D	F	54	BM	38.9	Intermediate
122	x	x		51,XX,+X,t(1;3)(p36;p21),del(5)(q13q33),-7,+8,+9,add(10)(p15),+13,+20,+22[20/20]	de novo	na	D	F	83	BM	1.8	Adverse
24			x	43,XY,-7,hsr(11)(q13q23),-13,-17,del(20)(q11q13),-21,-der(22)add(22)(p13),+mar,1~3dmin[19/23]/44,XY,-7,hsr(11)(q13q23),-13,-17,del(20)(q11q13),-21,-der(22)add(22)(p13),+2mar,1~3dmin[4/23]	t-AML	M5	D	M	69	BM	238	Adverse
2304	x	x		52,XY,inv(3)(q21q26),+8,+10,+13,+15,+21,+22[18/20]/45,XY,inv(3)(q21q26),-7[2/20]	de novo	M1	D	M	47	BM	1.0	Adverse
2033	x	x		47,XX,del(11)(p11p15);t(15;17)(q24;q25)*,inv(16)(p13q22),+8[20/20]	de novo	M1	D	F	57	BM	48.3	Favorable
1001	x	x		47,XY,+13[20/20]	de novo	na	D	M	54	BM	66.0	Intermediate
71			x	44-47,XX,t(4;17)(p15;q21),del(5)(q13q33),-7,-18,der(X),1~3mar[9/20]/46,XX[11/20]	de novo	M2	D	M	67	BM	3.6	Adverse
1946	x	x		46,XY,+8,add(11)(p15),-13,+mar[7/19]/46,XY,+8,add(11)(p15)x2,-13,+mar[5/19]/46,XY,+8,add(11)(p15)[5/19]/46,XY,+8,add(11)(p15)x2[2/19]	de novo	M5	D	F	49	BM	3.0	Adverse
2045	x	x		44,XX,-3,del(4)(q21q31),-5,del(7)(q22q36),der(8)t(3;8)(q25;p21),del(10)(q22q24),inv(11)(q13q23),der(12),add(13)(q34),del(15)(q11q24),del(16)(q22),add(17)(p13),-18,+r[15/20]/46,XX[5/20]	de novo	M0	D	F	64	BM	3.1	Adverse
56			x	45,XX,t(3;21)(q26;q22),der(5)(q?),-7,del(12)(p11p13)[15/18]/45,XX,t(3;21)(q26;q22),der(5)(q?),-7,del(11)(p13p15),del(12)(p11p13)[3/18]	t-AML	na	D	F	62	BM	77	Adverse
2043	x	x		46,XY,del(12)(p11p13)(7/20)/47,XY,del(12)(p11p13)+13[2/20]/48,XY,del(12)(p11p13),+13,+14[3/20]/49,XY,del(12)(p11;p13),+der(12)(p11;p13),+13,+14[3/20]/46,XY[5/20]	de novo	M1	R	M	63	BM	na	Adverse

195	x	x		44,XY,t(Y;1;5)(p11;p32;q33),-7,del(12)(p12;p13) (5/19)/44,XY,t(Y;1;5)(p11;p32;q33),-7,add(11)(q23),del(12)(p12;p13),-18 [3/19]/44,XY,t(Y;1;5)(p11;p32;q33),der(2), -7, del(12)(p12;p13),t(12?;19)(q13?;p13),-18,del(X)(p21)[9/19]/46,XY[2/19]	sec (MDS)	sec (MDS)	D	M	58	BM	20.3	Adverse
1905	x	x		42,XY,-4,del(5)(q13q33),-7,-12,der(16),add(17)(p13),der(19)t(4;19)(q31;p13),-20,-21,-22,+mar[9/21]/42,XY,-4,-5,del(7)(q11q36),der(16),add(17)(p13),der(19)t(4;19)(q31;p13),-20,-21,+mar[2/21]/69-72,XXY,id[10/21]	de novo	na	D	M	79	BM	3.0	Adverse
2005	x	x		45,XX,-7[9/17]/46,XX[8/17]	de novo	M1	D	F	50	BM	1.6	Adverse
58			x	47,XY,+8[11/20]/46,X,-Y,+8[9/20]	de novo	M5	D	M	42	BM	66.9	Intermediate
2040	x	x	x	92-104,XXXX,+5,+8,+8,+9,+13,+13,+13,+20,+20,+21,+22,+22[20/20]	t-AML	M5	D	F	74	BM	2.8	Adverse
23			x	48,XY,+1,+13[16/27]/46,XY[11/27]	de novo	na	D	M	82	BM	na	Intermediate
1028	x	x		47,XX,+21[6/8]/46,XX[2/8]	de novo	M1	D	F	54	BM	1.2	Intermediate
2039	x	x		45 XY del(5)(q13;q33),dup(11)(q13;q25),t(12;16)(p13;p13),-13,-17,+r[20/20]	de novo	na	D	M	73	BM	25.0	Adverse
12			x	46,XX[11/20]/47,XX,+der(13)i(13)(q10)[8/20]/47,XX,+13[1/20]	t-AML	na	D	F	76	BM	na	Intermediate
68			x	48,XX,+14,inv(16)(p13q22),+21[18/20]/46,XX,inv(16)(p13q22)[2/20]	t-AML	M4	D	F	57	BM	10.5	Favorable
213	x	x		42-48,XX,del(5)(q13q33),i(21)(q10),+der(21)i(21)(q10)x2[18/20]/42-48,XX,del(5q),del(11)(p13p15),-17,i(21q),+der(21)i(21)(q10)x2[2/20]	de novo	na	D	F	71	BM	2.5	Adverse
2036	x	x	x	49,XY,+3r[17/20]/46,XY[3/20]	de novo	M4	D	M	66	BM	115.0	Adverse
25			x	43,XY,del(2)(q?),+der(3)del(3)(q?),-5,-7,i(8)(q10),-13,-14,der(16),-17,add(22)(p13),+r[16/20]/43,XY,del(2)(q?),+der(3)del(3)(q?),-5,-7,der(8)t(8q?;11q?),-11,der(16),-17[3/20]/46,XY[1/20]	de novo	M0	D	M	62	BM	1.5	Adverse
55			x	47,XX,+8[20/20]	de novo	M4	D	F	71	BM	90	Intermediate
1041	x	x		47,XY,+8[14/20]/46,XY[6/20]	de novo	M5	D	M	77	BM	12.6	Intermediate
Euploid AML												
025288	x	x		46,XX,t(6;11)(q27;q23)[20/21]/46,XX[1/21]	de novo	M4	D	F	49	BM	76.0	Adverse
007340	x			46,XX,del(7)(q34),inv(16)(p13q22)[13/20]/46,XX[7/20]	de novo	M4eo	D	F	35	BM	na	Favorable

009796	x	x		46,XX,ins(8;21)(q22;q22q22),del(9)(q12q34)[12/15]/46,XX[3/15]	de novo	M2	D	F	29	BM	6.1	Favorable
019074	x	x		46,XX,der(7)t(3;7)(q26;q11.2)[20/20]	t-AML	M2	D	F	50	BM	20.4	Intermediate
000894	x	x		46,XY,t(11;19)(q23;p13)[20/20]	t-AML	na	D	M	29	BM	na	Intermediate
013324	x	x		46,XY,der(19)t(17;19)(q21;p13).ish der(10)ins(10;11)(p12;q23q23)[15/20]/46,XY[5/20]	de novo	M1	D	M	49	BM	115.0	Adverse
031805	x	x		46,XY,t(3;3)(q21;q26)[12/12]	de novo	na	D	M	31	BM	1.1	Adverse
2973	x	x		46,XX,t(16;16)(p13;q22)[20/20]	t-AML	M4	D	F	46	BM	95.0	Favorable
6			x	46,XY[20/20]	de novo	M4	D	M	57	BM	2.9	Intermediate
18			x	46,XX[20/20]	na	na	D	F	42	BM	14.5	Intermediate
1024	x	x		46,XY,dup(1)(p22p36)[20/20]	sec (MDS)	sec (MDS)	D	M	74	BM	1.5	Intermediate
1026	x	x		46,XX,del(9)(q12q34)[20/20]	sec (MDS)	sec (MDS)	D	F	63	BM	21.8	Intermediate
48			x	46,XX[20/20]	de novo	M1	D	F	60	BM	3.2	Intermediate
2306	x			46XX,del(7)(q22q32)[15/20]/46,XX[5/20]	de novo	M1	D	F	40	BM	1.1	Intermediate
0027	x	x		46,XX[13/15]/46,XX,del(5)(q31q33)[2/15]	de novo	M0-M1	D	F	70	BM	3.8	Adverse
47			x	46,XX[20/20]	de novo	M1-M2	D	F	66	BM	35.9	Intermediate
66			x	46,XY[20/20]	de novo	na	D	M	70	BM	234	Intermediate
2195	x	x	x	46,XX[10/10]	de novo	M2	D	F	63	BM	44.3	Intermediate
14			x	46,XX[20/20]	de novo	M0	D	F	51	BM	3.8	Intermediate
2241	x	x		46,XY[20/20]	de novo	M5	D	M	62	BM	50.0	Intermediate
64			x	46,XX[20/20]	sec (MDS)	sec (MDS)	D	F	66	BM	na	Intermediate
41			x	46,XX[20/20]	de novo	na	D	F	60	BM	68.5	Intermediate
40			x	46,XX[20/20]	na	na	D	F	76	BM	na	Intermediate

50			x	46,XX[20/20]	de novo	M1	D	F	34	BM	102	Interme diate
59			x	46,XY[20/20]	de novo	M1	D	M	64	BM	1.9	Interme diate
0017	x	x		46, XY[20/20]	de novo	M0-M1	D	M	57	BM	3.1	Interme diate
49			x	46,XX[20/20]	de novo	na	D	F	72	BM	26.1	Interme diate
85			x	46,XX[20/20]	de novo	M2-M4	D	F	67	BM	46.7	Interme diate
2798	x	x		46,XY[20/20]	de novo	na	R	M	60	BM	37.6	Interme diate
65			x	46,XY[28/28]	de novo	M1	D	M	64	BM	65.1	Interme diate
2008	x	x		46,XY,der(1)r(1p),t(5;13)(q22;q32),t(9;20)(q13;q11),der(19)t(1;19)(q21;q 13),HSR[17/20]/46,XY[3/20]	sec (MDS)	sec (MDS)	D	M	55	BM	2.6	Adverse
45			x	46,XY[20/20]	de novo	na	D	M	42	BM	163.9	Interme diate
83			x	46,XY[20/20]	de novo	na	D	M	73	BM	2	Interme diate
2031	x	x		46,XX,del(13)(q14q22)[10/20]/46,XX,del(8)(p21),del(13)(q14q22)[3/20]/ 46,XX[7/20]	t-AML	M1	D	F	64	BM	13.0	Interme diate
1858	x	x		46,XX[20/20]	de novo	M1	D	F	36	BM	5.6	Interme diate
0018	x	x		46,XX[20/20]	sec (MDS)	M2	D	F	51	PBL	61.6	Interme diate
3154	x	x		46,XX[20/20]	de novo	M5	D	F	62	BM	10.7	Interme diate
3010	x	x		46,XX[20/20]	de novo	M1	D	F	56	BM	84.3	Interme diate
15			x	46,XX[20/20]	de novo	M2	D	F	47	BM	89.4	Interme diate
46			x	46,XY[20/20]	de novo	M5	D	M	45	BM	88	Interme diate
16			x	46,XY[30/30]	de novo	M4	D	M	67	BM	9.7	Interme diate
0037	x	x	x	46,XX[20/20]	de novo	M2	D	F	59	BM	1.6	Interme diate
0022	x	x	x	46,XY[20/20]	de novo	M2	D	M	77	BM	6.9	Interme diate
2240	x	x		46,XY,inv(16)(p13q22)[19/20]/46,XY[1/20]	de novo	M4	D	M	55	BM	10.0	Favorabl e

39			x	46,XX[20/20]	de novo	M5	D	F	50	BM	77.7	Intermediate
2230	x	x		46,XY[20/20]	sec (MDS)	na	D	M	74	BM	1.8	Favorable
3213	x	x		46,XX[20/20]	de novo	M4	D	F	32	BM	15.6	Favorable
51			x	46,XX[20/20]	de novo	M1	D	F	38	BM	37.2	Intermediate
20			x	46,XY[20/20]	sec (MDS)	na	D	M	71	BM	2.3	Intermediate
3062	x	x		46XY,t(6;11)(q27;q23)[19/20]/46XY[1/20]	de novo	M2	D	M	18	BM	127.0	Adverse
1015	x	x		46,XX,t(2;14)(q21;q32),t(11;12)(p15;q22)[17/20]/46,XX[3/20]	de novo	M0-M1	D	F	39	BM	50.0	Intermediate
1010	x	x		46,XY,t(6;17)(p21;q11)[20/20]	sec (MDS)	M2	D	M	64	BM	30.5	Intermediate
26			x	46,XX[20/20]	de novo	M0	D	F	67	BM	108.6	Intermediate
44			x	46,XY[20/20]	de novo	M4	D	M	66	BM	18.9	Intermediate
2138	x	x		46,XX[20/20]	sec (MDS)	M2	D	F	68	BM	5.2	Intermediate
1025	x	x		46,XX[14/28]/46,XX,der(9)t(1;9)(q11;q34)[14/28]	de novo	M4	D	F	55	BM	13.7	Intermediate
1905a	x	x		46,XY[20/20]	sec (MDS)	sec (MDS)	D	M	80	BM	3.9	Intermediate
2023	x			46,XX,del(7)(q22;q32),inv(16)(p13q22)[15/20]/47,XX,del(7)(q22;q32),inv(16)(p13;q22),+9[1/20]/46,XX[4/20]	t-AML	na	D	F	62	BM	1.2	Favorable
1034	x	x		46,XX,inv(3)(q21q26)[8/20]/46,XX[12/20]	de novo	M7	D	F	61	BM	12.7	Adverse

*FAB=French-American-British; idic=isodicentric chromosome; i=isochromosome; dic=dicentric chromosome; dmin=double minute; der=derivative chromosome; ish=in situ hybridization ; r=ring chromosome; mar=marker chromosome; * not involving PML-RARA.*

Table S2. Number and type of mutations detected by WES.

	No. mutations	Missense	Frameshift deletion	Frameshift insertion	In-frame deletion	In-frame insertion	Stop-gain	Stop-loss
Aneuploid AML								
2005	2	2						
1001	6	5	1					
2033	6	6						
1028	9	9						
013206	10	9		1				
006473	10	9					1	
026369	13	12					1	
019531	13	12					1	
001197	14	13	1					
2002	13	13						
213	16	15					1	
008951	16	14					2	
2030	17	17						
1014	17	17						
006187	18	15			1	1	1	
2007	18	16					2	
1006	19	17					2	
026656	17	12		2	1		2	
042101	19	18			1			
015330	20	15	1	1			3	
2036	21	20					1	
028034	22	19			2		1	
2009	23	22	1					
2004	24	23			1			
007827	26	22			1		3	
122	27	26					1	
2043	29	28					1	
2035	29	27					2	
2304	29	27	1				1	
3108	25	22			1		2	
2868	33	33						
2803	34	31	1	1			1	
1905	39	36	1				2	
1041	40	36			1	2	1	
63640	42	37				1	4	
1946	41	40					1	
2964	42	39		1			2	
2039	47	46		1				
2045	47	44	1		1		1	
187	50	44	1		1		4	
195	56	48					8	
2040	88	83					5	

Euploid AML								
3010	3	3						
1905a	7	7						
000894	7	7						
3062	8	7			1			
007340	8	7					1	
1024	9	8					1	
2241	9	7	1				1	
2008	9	8	1					
1858	9	9						
025288	9	8			1			
3154	9	9						
0017	9	9						
0022	10	8					2	
2240	10	10						
2798	11	8		1			2	
3213	11	9	1				1	
013324	11	11						
009796	11	9		2				
2138	11	10	1					
0018	12	11			1			
2023	13	12		1				
1015	14	13					1	
1010	14	10					4	
1025	13	12					1	
0027	15	13					2	
1034	15	15						
2031	17	15		1			1	
1026	18	17		1				
019074	20	15		3			2	
2973	24	23						1
031805	25	23		2				
2230	27	19	3				5	
2195	30	29				1		
2306	43	39	1				3	
0037	45	44					1	

Table S3. List of mutations in the functional categories associated with the aneuploid phenotype.

Functional category	Genes mutated in A-AML	Genes mutated in E-AML
Transcription	AR; ARNTL; BDP1; CDX4; EVX2; FOXP3; GABPB2; HOXA1; HOXC10; HOXC9; ICE1; JAZF1; LIMD1; MAZ; MED11; MED20; MED23; MGA; MLLT10; MSC; MYBL1; NCOA6; NFATC2IP; NFATC4; NFE2L1; NFKB2; NPAS1; NPAS2; NR2C2; NR4A1; NRIP2; PARP14; POLR2B; POLR2H; POLR3A; POLR3B; POLR3E; RAI1; RORB; SALL3; SATB2; SERTAD3; SOX13; SOX30; SPEN; SPI1; TAF4B; TCF20; TLE2; TSHZ3; VGLL3; ZBTB38; ZEB2; ZHX1; ZHX2; ZKSCAN5; ZMYM2; ZNF107; ZNF205; ZNF227; ZNF28; ZNF282; ZNF383; ZNF407; ZNF423; ZNF57; ZNF626; ZNF653; ZNF696; ZNF845; ZNF91; ZSCAN26; ZXDC	AFF1; ARNT2; BTAF1; CAMTA1; DNTTIP2; HIC1; HNF1B; HOXD13; MED10; MGA; NOTCH3; PATZ1; PITX1; PLAG1; POLR1A; RAI1; TEAD2; TRERF1; UBTF; ZFP82; ZKSCAN7; ZNF148; ZNF229; ZNF333; ZNF407; ZNF460; ZNF483; ZNF491; ZNF558; ZNF572; ZNF595,ZNF718; ZNF621; ZNF628; ZNF891
Metabolism	AASDHPPT; ACOT11; ACSL3; ADCY8; AGXT2; AKR7A3; ALOXE3; APOB; CES2; CHST9; COL4A3BP; CPOX; CPT2; CTH; CTNNA1; CUBN; CYP3A43; DAOA; ENPP3; ENTPD4; EXT2; FAR2; FLAD1; FOXRED2; GALNT8; GCNT3; GLA; GPD1L; GPR119; HK3; IDO2; IMPDH2; LDHD; LSR; MACROD2; MORC2; NAAA; NADSYN1; NAGS; NCF2; NNT; NT5E; NUDT18; P4HB; PANK2; PASK; PCK1; PDE12; PENK; PHKB; PIBF1; PIGQ; PIP5K1C; PLA2G4E; POMGNT2; PPAPDC2; PTGS1; RNF219; RRP8; SDSL; SOD2; SSTR4; UGT3A1; VNN3; VSTM4	AASDHPPT; AASS; ACOT11; ACOX1; ACPT; ACSL3; ADCK3; ADCY8; AGXT2; AK9; AKR7A3; ALG9; ALOXE3; APOB; ATP8A2; CES2; CHDH; CHST5; CHST9; CIDEA; COL4A3BP; CPOX; CPT2; CTH; CTNNA1; CUBN; CYP1A1; CYP1B1; CYP3A43; DAOA; ENPP3; ENTPD4; EXT2; FAR2; FLAD1; FOXRED2; GALNT8; GCNT3; GLA; GLP1R; GPD1L; GPR119; GYG2; H6PD; HK3; HMGCR; IDO2; IMPDH2; LDHD; LSR; MACROD2; MAN2C1; MBOAT1; MORC2; MTCH2; NAAA; NADSYN1; NAGS; NAT2; NCF2; NDST1; NDST4; NDUFS1; NNT; NPC1L1; NT5E; NUDT18; OSBPL7; P4HB; PANK2; PASK; PC; PCK1; PDE12; PENK; PGAP2; PGM2; PHKB; PIBF1; PIGQ; PIGT; PIGZ; PIP5K1C; PKM; PLA2G4E; POMGNT2; PPAPDC2; PTGS1; PYGM; RNF219; RRP8; SDSL; SMPD3; SOD2; SSTR4; TYSND1; UGT3A1; VNN3; VSTM4; XDH
Cell cycle	AHCTF1; AKAP9; BUB1B; C10orf90; CASP8AP2; CDK10; CDK20; CDK5RAP2; CDKN1C; CENPJ; CEP152; CEP70; CHAF1A; CLTC; DDIAS; ESPL1; FAM64A; FHL1; FOXM1; HAUS4; INCENP; MCM6; MCM7; MELK; MIS18A; NASP; NCAPD2; NDC1; NSUN2; NUSAP1; ODF2; PCM1; PKHD1; POLA1; PPM1D; PRIMPOL; RBM38; RBMS1; RSF1; SMC1A; STOX1; TAF1; TICRR; TOP3A; TRIOBP; URGCP; USP44; YY1AP1	CDC20B; CENPO; CEP250; CEP295; CEP85; CHFR; CNTRL; MNAT1; MTUS2; NINL; NUP37; SFI1; SMC1A; TUBGCP2; WRN
Signal transduction	ADCYAP1R1; AGER; AKAP12; AKAP6; ALPK2; AMER3; APLP2; ARHGAP31; ARHGEF11; ARHGEF37; ARHGEF40; ARHGEF6; CALCRL; CARD10; CSK; DACT2; DDR1; DENND5A; DENND6A; DLK1; EDN2; EPHA10; EPHA5; ERRFI1; ESR1; ESRRB; EVI2A; FRS3; FZD2; GAREML; GPC6; GPR153; GPRC5C; IL22RA1; INSRR; ITPR3; MAGI3; MAST3; MCF2L2; MERTK; MFNG; MPP6; NLRC3; NRK; OSMR; PAK6; PIK3R1; PLA2R1; PLCD1; PLD1; PLEKHG4; PLEKHG5; PRKCE; PRKRIR; PTK7; PYGO1; RASGEF1A; RASGRP3; RGS22; RHOBTB1; RHOQ; RPTOR; SBF1; SGK1; SIPA1; SIPA1L2; SMEK2; SOCS4; ST5; SULF1; TBC1D9; TIFAB; TNFRSF1A; TRAF2; TRAF4; TRIP6; TSC1; VRK3; WWC2	APLP2; ARHGEF12; ARHGEF33; CALCOCO1; CALCRL; CGN; CSNK2A1; DAPK2; DIRAS2; DKK1; EGFR; EPHA3; EPHA7; GNB1; GPR34; GRAP2; GRK6; IBTK; INPPL1; INSR; IQSEC1; ITPR2; KCTD16; LPAR5; MAST1; NLRC5; PDE5A; PLA2R1; PPP2R1A; PYGO1; RALGDS; RAP1B; RASAL2; RASD2; RGS3; RIPK2; RIT1; SIK1; STK11IP; WDR24; ZFYVE1
Cytoskeleton	ACTBL2; ACTR3B; ALMS1; ARHGAP26; BAIAP2L2; CAPZA2; CDC42EP1; COBL; DIAPH2; DST; ELMO1; ELMO2; EPB41L4B; FLNA; KANK3; KIF1B; KLC4; LAMA2; LAMA4; LAMB1; LMNB2; MICALL2; MTSS1; MYLK;	CTTN; EPB41L1; GAS2L2; KIF1A; KIF21A; KLC2; LAMC3; LMNA; MARK2; MICAL3; MYH14; MYH4; MYO1H; OBSL1; RICTOR; RLTPR; SDC3; SHROOM1; SNTB2; SORBS3; UBXN11

	MYO1A; PPP1R18; QRICH1; RAC2; SPTA1; SPTB; SVIL; SYNE2; SYNPO2L; TBCB; TNS3; TUBA8; VCL; VILL	
Homeostasis	ANO8; ASPH; ATP13A2; BEST2; CACNA1C; CACNA2D2; CACNB4; CFTR; CHRNB2; DPP6; HEPH; HTR3B; KCNA5; KCNA6; KCNC1; KCNH4; KCNJ3; KCNK5; KCNQ1; KCNT1; KCNV1; KCNV2; LRRC8A; MCOLN2; ORAI1; PDZD2; PIEZO2; SARAF; SLC12A8; SLC34A1; SLC38A10; SLC38A6; SLC4A10; TPCN2	ANO3; CACNA1H; CACNA2D3; CLCA2; KCNG4; KCNK13; PIEZO2; SLC34A1; SLC4A4; SLC9A7; SLC9A9
Intracellular trafficking	AAAS; ABRA; ACAP1; ADAMTS20; AGAP1; AP1M2; ARCN1; ARFGAP2; BICD2; COG7; CRABP2; FAM160A2; FYCO1; HOOK2; HOOK3; IPO5; MIA3; NACAD; NUP153; NUP188; RAB3IP; SDAD1; SNX3; TOM1L2; VPS13A; VPS13D; XPO7; YIF1A	AP3M1; ARHGAP33; EXPH5; IPO4; LMTK2; LYST; OSBP; SEC16A; SEC24B; UBAP1
Activated signaling	DUSP10; FLT3; KIT; KRAS; MAP3K4; MAPK14; MET; NRAS; PPP2R2A; PRKCA; PTPN11; PTPN5; PTPRM; RET; STAT5A; TEK	FLT3; KRAS; MET; NF1; NRAS; PHLPP1; PTPN11; PTPN3; STAT3
Ubiquitination	CBL; FBXL15; FBXL7; HERC6; KBTBD6; MAGEL2; MARCH10; MARCH7; MYCBP2; OTUD7A; OTUD7B; RNF145; RNF220; TRIM32; TRIM41; TRIP12; UBA7; UBE2E2; USP10; USP24; USP42; USP43; USP8; WDTC1	ASB4; ASB10; FBXO40; LRRC41; TOPORS; TRIM9; UBXN7; WDR48
Cell adhesion	AMIGO1; CDH24; CDHR5; COL12A1; COL16A1; COL5A1; COL6A3; COL6A6; DCHS2; DSCAM; DSCAM; FERMT1; GP9; ITGAE; ITGB6; MAGI1; NID2; PARVA; PTPRF; VTN	CDHR3; DSCAM; ITGA4; ITGB3
Epigenetic regulation	ARID1A; ARID5A; ASXL2; BAZ1B; C11orf30; C17orf49; CCDC101; CECR2; CHD1; CHD4; CHD6; EPC1; HIST1H2BA; HIST1H2BE; HIST1H2BF; HMGN3; KAT6B; L3MBTL2; MSL2; MTA1; PAF1; PHC1; RTF1; TRRAP	ASXL1; ASXL2; BCOR; BCORL1; BPTF; KANSL1; MEN1
Post-translational modification	ADAMTSL1; ART1; ART5; B3GALT5; B3GALTL; CPN2; FNTB; FUT2; GXYLT1; HHAT; HYOU1; IBA57; METAP1; MIPEP; NAA25; PADI3; PPIG; RPN1; SDF2; SIAE; TMX1; TPST1; UGGT2	COLGALT1; DAD1; MTMR8; NOTUM; PFDN5; TGM1; ZDHHC4; ZDHHC8
RNA metabolism	ADARB2; AGO1; ATXN1; DDX1; DIS3L; DNAJB11; ELAC1; FTSJ3; IMP3; INTS1; LCMT2; MEX3B; MPHOSPH10; NOP2; PUS1; PXDNL; QARS; RC3H1; RNASEK; RNH1; SMG7; WIBG; YBX3; ZC3H13; ZCCHC11; ZCCHC6	AGO1; BICD1; CNOT1; DCP2; HNRNPL; MATR3; RNMTL1; RPUSD2; SAMD4A
Immune response	BPI; CD1D; CD84; CD96; CNR2; DMBT1; F8; F9; FCRL3; IL10RA; IL16; IL2RB; IRF5; MAVS; NLRP14; NLRX1; PRB3; SERPINA3; SIGLEC5; TLR10	CLEC4M; DEFB134; DMBT1; EPX; HAVCR1; HHLA2; IFI44L; LRRC32; MARCO; MMRN1; MNDA; SERPINA5; XPNPEP2
Cell proliferation/survival	ADNP2; BAG3; CSF3R; DOCK1; ENDOG; FGF1; FGF6; FGFR1; LGALS1; LIX1L; LTBP3; MACC1; MTUS1; PACRG; RPS6; TACSTD2; TGFBR2; ZFP36L1; ZNF217	ADNP2; AHI1; C1orf56; CASP5; DOCK4; MCC; PDGFA; PEG3; PLAGL1; REL; SETBP1; VWCE
Histone methylation	ASH1L; KMT2C; KMT2D; MLL amplification; MLL fusion; PRDM14; PRMT6; SETD2; SETD4; SETD5; SMYD1	EZH2; KMT2C; MLL; MLL amplification; MLL fusion; SETD1B
Splicing	CELF5; CLASRP; DDX50; EIF4A3; FUS; GEMIN4; LUC7L2; POLR2A; PRPF40B; PTBP2; SF3B1; SNRNP200; SRRM2; ZNF326	DDX41; PABPC1; QKI; SF3B1; SRSF2; YBX1; YTHDC2
Protein degradation	ADAMTS12; ADAMTS15; ADAMTS5; CTBS; DDI1; DDI2; HGFAC; KLK12; PSMB6; PSMC2; RCE1; WFDC13; ZFAND2B	A2ML1; ADAMTS16; ADAMTS17; CPA5; KLK8; PSME4

Tumor suppressor	BRCA1; BRCA2; TP53	TP53
DNA Methylation	DNMT3A; DNMT3B; IDH2; WT1	DNMT3A; IDH2; NPM1; TET2; WT1
Transport	ABCA13; ABCA2; ABCC1; ATP6V1G1; MFI2; SLC28A3; SLC2A9; SLC38A5; SLC45A3; SLC5A3; SLC6A14; SLCO2B1; STRA6	ABCA12; ABCA13; ABCC1; SLCO1C1; STEAP3
Translation	DDX31; EEF1D; EIF2S3; EIF3D; EPRS; GTPBP2; METTL17; MRPL14; MTRF1; RPL19; RPL22L1; RPL6; SARS	EIF5; HARS; MARS2; MRPS2; NARS
DNA damage and repair	ANKLE1; APTX; CHEK2; ERCC6L2; FANCE; HERC2; MLH3; NEIL3; RIF1; SPIDR; XAB2; ZRANB3	ATM; CDK12; CHEK2; HIPK2; LIG3; MSH5; NEIL3; PARP1; REV3L; RNF169; RTEL1; SLX4; SPRTN
Cell differentiation	ANXA13; DMXL1; EDRF1; GLIPR1; MFHAS1; NBEAL2; NFE2; WNT8A; ZBTB7A; ZNF3	AHR; DLL3; MFHAS1; MYADM
Cell migration	FAT1; LIMS2; LRRC16A; PLXNB1; PLXND1; PPIA; S1PR1	FAT1; FLT1; PEAK1; SPATA13
Endocytosis	ANKRD13B; CD163; CHODL; EEA1; RILP; STAB2	MRC2
Apoptosis	PIDD1; RTN3; TM7SF3; UNC5D; ZNF420	UNC5D
Autophagy	ATG2B; WDFY3; SOGA1; WDFY3; ATG2B	
Myeloid transcription factor	GATA2; GLI3; MYC; RUNX1	CEBPA; GATA2; RUNX1
Histone DeMethylation	KDM6A; KDM6B; KDM7A	KDM5C; KDM6A; PHF8

Table S4. Summary of GO-BP pathways enriched in A-AML according to CNAs.

Pathway ID	Pathway name	Logistic regression coeffcient *	*p* value (adj) †	% A-AML cases 3	%E-AML cases ‡	Genes and frequency in A-AML §	Genes and frequency in E-AML §
CELL CYCLE							
GO:1901989	positive regulation of cell cycle phase transition	-5.015	0.012	7.89	0	ZNF16 (7.89)	
GO:1901992	positive regulation of mitotic cell cycle phase transition	-4.931	0.012	7.89	0	TMOD3 (7.89)	
GO:1901991	negative regulation of mitotic cell cycle phase transition	-3.744	0.013	21.05	3.12	RGCC (18.42); TPRA1 (5.26); ZFP36L1 (10.53)	RGCC (3.12)
GO:0000082	G1/S transition of mitotic cell cycle	-4.291	0.012	76.32	65.62	ACVR1 (10.53); ACVR1B (7.89); AKT1 (10.53); CCND1 (10.53); CCNE1 (7.89); CCNE2 (39.47); CDC6 (5.26); CDC7 (5.26); CDK2 (5.26); CDK3 (10.53); CDK4 (7.89); CDK6 (7.89); CDK7 (15.79); CDKN1A (13.16); CDKN1B (5.26); CDKN2A (18.42); CDKN2C (7.89); CDKN2D (7.89); CRLF3 (5.26); CUL3 (7.89); EIF4EBP1 (5.26); HINFP (10.53); INHBA (5.26); IQGAP3 (5.26); ITGB1 (7.89); LATS1 (18.42); LATS2 (2.63); MARK4 (7.89); MAX (10.53); MCM10 (7.89); MCM2 (5.26); MCM4 (50); MCM5 (18.42); MCM6 (7.89); MCM7 (2.63); MCM8 (10.53); MYC (52.63); ORC1 (7.89); ORC2 (7.89); ORC3 (18.42); ORC4 (10.53); PLK2 (7.89); PLK3 (7.89); POLE (7.89); POLE3 (13.16); PPP6C (10.53); PRIM1 (2.63); RANBP2 (5.26); RANBP3 (7.89); RANBP3L (2.63); RB1 (57.89); RBBP8 (5.26); RPA1 (5.26); RPA2 (10.53); SPDYA (5.26); TFDP3 (7.89); USP37 (5.26); WEE1 (7.89)	CDK3 (3.12); CDK6 (6.25); CDK7 (3.12); CDKN2A (3.12); CDKN2C (6.25); LATS1 (3.12); LATS2 (3.12); MARK4 (3.12); ORC1 (3.12); PHF8 (6.25); PLK3 (3.12); RB1 (59.38)
GO:1902806	regulation of cell cycle G1/S phase transition	-4.041	0.012	55.26	0	C8orf4 (47.37); FBXW7 (13.16); TAF1 (7.89)	
GO:2000045	regulation of G1/S transition of mitotic cell cycle	-4.083	0.011	26.32	3.12	CCND1 (10.53); E2F1 (10.53); ECD (5.26); ID2 (5.26); INO80 (7.89); KIF14 (5.26); PSME1 (10.53); PSME2 (10.53); SENP2 (5.26); TCF3 (7.89)	SENP2 (3.12)
GO:1902807	negative regulation of cell cycle G1/S phase transition	-3.781	0.012	10.53	0	CDKN2D (7.89); MEN1 (7.89)	
GO:2000134	negative regulation of G1/S transition of mitotic cell cycle	-3.763	0.012	65.79	65.62	BCL2 (5.26); BRD7 (7.89); CDC73 (5.26); CDK2AP2 (7.89); CDKN1A (13.16); CTDSP1 (5.26); DACT1 (10.53); DCUN1D3 (5.26); E2F7 (7.89); EZH2 (2.63); FBXO7 (21.05); FHL1 (10.53); GPNMB (5.26); MYO16 (5.26); PRMT2 (2.63); PTEN (10.53); RB1 (57.89); SMARCA4 (7.89); ZNF655 (2.63)	FBXO7 (3.12); PTEN (9.38); RB1 (59.38)
GO:0000083	regulation of transcription involved in G1/S transition of mitotic cell cycle	-8.165	0.011	68.42	59.38	BACH1 (23.68); BRD4 (7.89); CCNA1 (18.42); CCNE1 (7.89); CDC6 (5.26); E2F1 (10.53); HINFP (10.53); NPAT (7.89); ORC1 (7.89); PCNA (10.53); RB1 (57.89); RRM2 (5.26); TFDP1 (23.68)	CCNA1 (3.12); ORC1 (3.12); RB1 (59.38); TFDP1 (3.12)
GO:0007062	sister chromatid cohesion	-5.659	0.011	65.79	28.12	AURKB (7.89); BUB1B (7.89); BUB3 (7.89); CDC20 (10.53); CDCA8 (7.89); CENPA (7.89); CENPC (15.79); CENPH (15.79); CENPI (7.89); CENPK (10.53); CENPM	CDC20 (3.12); CDCA8 (6.25); CENPH (3.12); CENPK (3.12); CLIP1 (3.12); KIF2B (3.12); KIF2C (3.12);

						(13.16); CENPO (7.89); CENPP (13.16); CENPU (13.16); CKAP5 (7.89); CLIP1 (7.89); DSN1 (7.89); ERCC6L (7.89); ESCO1 (5.26); FBXW7 (13.16); HDAC8 (7.89); KIF22 (5.26); KIF2B (5.26); KIF2C (7.89); KNTC1 (2.63); MAD2L1 (13.16); MCMBP (7.89); MIS12 (5.26); NDEL1 (7.89); NSL1 (5.26); NUF2 (5.26); NUP133 (7.89); NUP37 (7.89); NUP43 (18.42); NUP85 (10.53); NUP98 (7.89); PDS5B (2.63); PHB2 (13.16); PLK1 (5.26); PPP1CC (7.89); RAD21 (47.37); RAD21L1 (2.63); RANBP2 (5.26); REC8 (10.53); RPS27 (5.26); SEC13 (5.26); SMC1A (7.89); SMC5 (2.63); SPC24 (5.26); SPDL1 (7.89); STAG1 (5.26); STAG2 (13.16); WAPL (5.26); XPO1 (5.26); ZW10 (10.53); ZWILCH (10.53)	NUP43 (3.12); NUP85 (3.12); SMC1A (3.12); STAG2 (12.5)
GO:0007064	mitotic sister chromatid cohesion	-4.666	0.013	26.32	12.5	CDC20 (10.53); CHTF8 (7.89); DSCC1 (7.89); GSG2 (5.26); MAU2 (7.89); NIPBL (15.79); PDS5B (2.63); SMC1A (7.89); WAPL (5.26)	CDC20 (3.12); NIPBL (6.25); SMC1A (3.12)
PROTEIN POST-TRANSLATIONAL MODIFICATION							
GO:0043254	regulation of protein complex assembly	-6.775	0.009	47.37	40.62	BBS10 (2.63); DAB2IP (2.63); GFAP (5.26); HJURP (5.26); HRG (5.26); HSP90AA1 (10.53); HSPA8 (10.53); INSM1 (13.16); IRGM (5.26); LATS1 (18.42); NCLN (7.89); PTPN11 (23.68)	HJURP (3.12); HRG (3.12); LATS1 (3.12); PTPN11 (34.38)
GO:0031334	positive regulation of protein complex assembly	-7.783	0.009	34.21	9.38	AJUBA (10.53); FAF1 (7.89); FNIP1 (5.26); ICE1 (10.53); IFNG (10.53); SLF1 (7.89); SLF2 (5.26); SUMO1 (7.89); TAL1 (7.89); TNF (10.53); VCP (18.42); WARS (10.53)	FAF1 (9.38); ICE1 (3.12); TAL1 (3.12); VCP (3.12)
GO:0006513	protein monoubiquitination	-9.532	0.020	57.89	28.12	PEX12 (5.26); RYBP (7.89); LEO1 (7.89); KLHL12 (5.26); UBE2E1 (5.26); UBB (7.89); CUL4B (7.89); CUL3 (7.89); SKP1 (5.26); DTX3L (5.26); WDR48 (5.26); PCGF1 (7.89); UHRF1 (7.89); CTR9 (7.89); WAC (5.26); BIRC2 (10.53); KDM2B (7.89); CBL (10.53); DDB2 (7.89); UBE2W (44.74); BCOR (13.16); RING1 (15.79); HUWE1 (13.16); RBX1 (13.16); UBE2T (5.26); DTL (5.26); CDC73 (5.26)	CUL4B (6.25); BCOR (28.12); HUWE1 (15.62)
STEM CELL REGULATING PATHWAYS							
GO:0007224	smoothened signaling pathway	-3.497	0.013	60.53	31.25	ARL13B (5.26); ARL3 (5.26); BBS7 (13.16); BMP4 (10.53); CENPJ (21.05); DHH (7.89); DISP1 (5.26); DYRK2 (7.89); DZIP1 (18.42); EVC2 (13.16); FKBP8 (7.89); GLI1 (7.89); GLI2 (5.26); GLI3 (15.79); HHIP (13.16); HIPK1 (7.89); HSPB11 (7.89); IFT27 (18.42); IFT46 (13.16); IFT52 (2.63); IFT57 (7.89); KIAA0586 (10.53); KIF3A (5.26); NDST1 (5.26); PDX1 (23.68); PKD2L1 (5.26); PTCH1 (13.16); SEPT2 (5.26); SHH (2.63); SMO (2.63); STIL (7.89); TCTN2 (2.63); TGFBR2 (7.89); TMEM17 (5.26); TROVE2 (5.26); WDR19 (13.16)	ARL13B (3.12); CENPJ (3.12); DISP1 (3.12); DZIP1 (3.12); GLI3 (15.62); HSPB11 (3.12); IFT57 (3.12); PDX1 (3.12); PTCH1 (3.12); STIL (3.12); TGFBR2 (3.12)
GO:0008589	regulation of smoothened signaling pathway	-4.021	0.012	36.84	3.12	CREBBP (7.89); FGF10 (2.63); FGFR2 (10.53); GAS1 (13.16); GLI1 (7.89); GLI2 (5.26); INTU (13.16); OTX2 (10.53); PTCH1 (13.16); RPGRIP1L (7.89); TCTN1 (7.89); TULP3 (13.16)	PTCH1 (3.12)

GO:0045879	negative regulation of smoothened signaling pathway	-4.610	0.013	81.58	84.38	CD3E (13.16); GLI3 (15.79); GPC3 (50); GPR37L1 (5.26); HHIP (13.16); KCTD11 (5.26); KIF7 (10.53); PTCH1 (13.16); PTCH2 (7.89); RB1 (57.89); RUNX2 (31.58); SERPINE2 (7.89); SUFU (5.26); TULP3 (13.16); ULK3 (7.89)	GLI3 (15.62); GPC3 (65.62); PTCH1 (3.12); PTCH2 (3.12); RB1 (59.38); RUNX2 (12.5)
METABOLISM							
GO:0009116	nucleoside metabolic process	-5.870	0.010	31.58	46.88	ATIC (5.26); FPGS (10.53); HPRT1 (28.95); NME5 (5.26); NT5C1A (7.89); PRPS1 (10.53); UMPS (5.26)	HPRT1 (40.62); NT5C1A (6.25); PRPS1 (3.12)
GO:0009163	nucleoside biosynthetic process	-6.447	0.012	52.63	46.88	TXNDC9 (5.26); PDCL2 (15.79); UMPS (5.26); NT5E (18.42); TK2 (5.26); UCK1 (2.63); NME5 (5.26); PDCL (2.63); PUDP (18.42); CMPK1 (7.89); ADAL (7.89); GMPS (5.26); ADA (7.89); CAD (7.89); DCK (13.16); NME1 (5.26); NME2 (5.26); HPRT1 (28.95); CTPS1 (7.89); IMPDH2 (5.26)	PUDP (15.62); CMPK1 (3.12); HPRT1 (40.62); CTPS1 (6.25)
GO:0042278	purine nucleoside metabolic process	-6.337	0.010	7.89	0	MACROD1 (7.89)	
GO:0042451	purine nucleoside biosynthetic process	-7.442	0.012	42.11	40.62	NT5E (18.42); NME5 (5.26); ADAL (7.89); GMPS (5.26); ADA (7.89); NME1 (5.26); NME2 (5.26); HPRT1 (28.95); IMPDH2 (5.26)	HPRT1 (40.62)
GO:0009123	nucleoside monophosphate metabolic process	-4.625	0.011	81.58	62.5	NDUFAB1 (5.26); NDUFB11 (7.89); SDHAF2 (7.89); COX7A2L (7.89); SDHD (10.53); ENTPD8 (10.53); GMPR2 (10.53); SLC25A23 (7.89); TPI1 (13.16); NUDT9 (13.16); ATP5G3 (10.53); RFK (10.53); PPIF (7.89); UQCC2 (15.79); CARD11 (2.63); COX15 (5.26); PFKFB1 (7.89); HKDC1 (7.89); COX8C (10.53); VCP (18.42); NMNAT1 (10.53); AMPD1 (5.26); ATP5B (2.63); CLPX (10.53); CCNB1 (15.79); UMPS (5.26); HSPA8 (10.53); P2RX7 (2.63); MSH2 (7.89); NT5E (18.42); NDUFAF7 (5.26); CDK1 (7.89); HTR2A (18.42); BAD (7.89); MYH3 (7.89); SURF1 (10.53); ATP6V1B2 (5.26); NDUFA8 (2.63); NDUFB9 (50); UQCC3 (7.89); DLG3 (21.05); SDHA (10.53); UQCRB (47.37); TK2 (5.26); PGAM1 (7.89); UCK1 (2.63); ATP5J (21.05); DUT (7.89); ALDOC (5.26); PGM1 (7.89); DNM1L (7.89); PGAM4 (7.89); SDHC (5.26); PARK7 (10.53); ATIC (5.26); ATP6V1A (5.26); GPI (5.26); GPD1 (7.89); COX6B1 (7.89); RHOA (5.26); GADD45GIP1 (5.26); NDUFA11 (7.89); MECP2 (2.63); ATP1A2 (5.26); NDUFA6 (13.16); EIF6 (10.53); MPP3 (5.26); FIGNL1 (2.63); COX6C (44.74); DLD (2.63); SLC25A25 (2.63); AMPD2 (5.26); NDUFV2 (5.26); OLA1 (10.53); AK2 (7.89); ALDOB (13.16); ENO2 (13.16); HK1 (7.89); ACTN3 (7.89); PRKAA1 (2.63); NUDT4 (7.89); COX7B (10.53); TEFM (5.26); NDUFV3 (2.63); PFAS (7.89); CMPK1 (7.89); DLG4 (5.26); NT5C1A (7.89); ATP5F1 (5.26); ATP5E (7.89); OGT (7.89); GMPS (5.26); NT5C2 (7.89); LEXM (7.89); ECD (5.26); NDUFB1 (10.53); ATP5EP2 (2.63); ARNT (5.26); NDUFS6 (13.16); NCOR1 (7.89); ADA (7.89); COQ9 (5.26); OGDHL (5.26); NDUFS5 (7.89); HIF1A (10.53); BPGM (2.63); ADSL (15.79); TIGAR (15.79); HSPA1B (15.79); UQCRHL (7.89); AK4 (7.89);	VCP (3.12); NMNAT1 (3.12); CCNB1 (3.12); HTR2A (6.25); DLG3 (31.25); SDHA (3.12); PGM1 (3.12); PARK7 (3.12); MECP2 (9.38); AK2 (3.12); COX7B (3.12); CMPK1 (3.12); NT5C1A (6.25); LEXM (3.12); ATP5EP2 (3.12); NDUFS6 (3.12); NDUFS5 (6.25); ADSL (3.12); AK4 (3.12); NDUFB5 (3.12); NDUFA3 (3.12); STOML2 (3.12); HPRT1 (40.62); ATP5H (3.12); ATP7A (12.5); ATP5L (3.12); ENPP1 (3.12); PRPS1 (3.12)

						NDUFAF1 (7.89); COX8A (7.89); DNPH1 (18.42); PAICS (15.79); NDUFB7 (7.89); COA6 (5.26); AK1 (2.63); PGK1 (7.89); CAD (7.89); DCK (13.16); PGK2 (18.42); NDUFB5 (5.26); NDUFA2 (5.26); ATP5A1 (5.26); NDUFA7 (7.89); NDUFB4 (5.26); NDUFA3 (13.16); NUDT5 (7.89); COX6A1 (7.89); STOML2 (18.42); SLC25A13 (2.63); STAT3 (5.26); GART (21.05); NDUFC1 (13.16); NDUFB3 (7.89); HPRT1 (28.95); DCTD (15.79); NDUFA12 (5.26); NUDT3 (15.79); COX4I2 (2.63); DGUOK (7.89); PFKL (2.63); ATP5H (10.53); IMPDH2 (5.26); MLXIPL (2.63); SIRT6 (7.89); CYC1 (7.89); NDUFA13 (7.89); COX6A2 (7.89); OGDH (2.63); ATP5G2 (2.63); NUDT10 (7.89); COX5A (7.89); DDIT4 (5.26); FLCN (7.89); ATP7A (15.79); ATP5L (13.16); TJP2 (2.63); ATP6V0A2 (7.89); CBFA2T3 (7.89); ENPP1 (18.42); ATP6V1B1 (7.89); IGF1 (2.63); ENO4 (7.89); PRPS1 (10.53); SCRIB (42.11); MYOG (5.26); ATP5J2 (2.63); UQCRQ (5.26)	
GO:0009124	nucleoside monophosphate biosynthetic process	-6.730	0.015	10.53	0	ENTPD8 (10.53)	
GO:0009132	nucleoside diphosphate metabolic process	-5.954	0.011	71.05	40.62	ENTPD8 (10.53); TPI1 (13.16); NUDT9 (13.16); CARD11 (2.63); PFKFB1 (7.89); HKDC1 (7.89); P2RX7 (2.63); HTR2A (18.42); BAD (7.89); DLG3 (21.05); PGAM1 (7.89); ALDOC (5.26); PGM1 (7.89); PGAM4 (7.89); GPI (5.26); GPD1 (7.89); NME5 (5.26); EIF6 (10.53); MPP3 (5.26); AK2 (7.89); ALDOB (13.16); ENO2 (13.16); HK1 (7.89); ACTN3 (7.89); PRKAA1 (2.63); CMPK1 (7.89); DLG4 (5.26); NUDT18 (5.26); OGT (7.89); AK8 (10.53); ECD (5.26); ARNT (5.26); NCOR1 (7.89); OGDHL (5.26); HIF1A (10.53); BPGM (2.63); TIGAR (15.79); AK4 (7.89); AK1 (2.63); PGK1 (7.89); PGK2 (18.42); NME1 (5.26); NME2 (5.26); NUDT5 (7.89); ENTPD4 (5.26); STAT3 (5.26); PFKL (2.63); MLXIPL (2.63); SIRT6 (7.89); OGDH (2.63); ENTPD3 (5.26); ENTPD2 (2.63); DDIT4 (5.26); TJP2 (2.63); CBFA2T3 (7.89); IGF1 (2.63); ENO4 (7.89); SCRIB (42.11); CMPK2 (5.26); MYOG (5.26)	HTR2A (6.25); DLG3 (31.25); PGM1 (3.12); AK2 (3.12); CMPK1 (3.12); AK4 (3.12)
GO:0046390	ribose phosphate biosynthetic process	-6.351	0.012	84.21	59.38	GNAI3 (5.26); GUCY1A3 (15.79); ATP5G3 (10.53); LPAR1 (2.63); CRH (47.37); DRD3 (5.26); RFK (10.53); PANK3 (7.89); EDNRA (13.16); PANK2 (2.63); VCP (18.42); AMPD1 (5.26); ATP5B (2.63); WFS1 (10.53); MC3R (7.89); HRH3 (5.26); AKAP9 (2.63); GUCA1A (18.42); PTGIR (7.89); UMPS (5.26); NF1 (21.05); NPPB (10.53); HTR1B (18.42); SURF1 (10.53); DCAKD (5.26); RXFP2 (31.58); GPR65 (15.79); UQCC3 (7.89); GNAI2 (5.26); APLP1 (10.53); AVP (2.63); UCK1 (2.63); ATP5J (21.05); TSHR (10.53); GUCY2F (7.89); ADCY7 (7.89); CCR2 (5.26); GHRH (7.89); VIP (18.42); PTH (7.89); DRD1 (5.26); CALCA (7.89); ATIC (5.26); NPPA (10.53); OPRL1 (2.63); PDZD3 (10.53); PPCS (7.89); PANK1 (7.89); MRAP (2.63); GUCY2C (5.26); RCVRN (2.63); NME5 (5.26); RLN2 (18.42); AMPD2 (5.26); AK2 (7.89); COASY (5.26); GUCA1B (18.42); GALR3 (15.79); APOE (7.89); ADCY10	VCP (3.12); AKAP9 (6.25); PTGIR (3.12); NF1 (34.38); NPPB (3.12); RXFP2 (12.5); GPR65 (9.38); VIP (3.12); NPPA (3.12); PPCS (6.25); RLN2 (3.12); AK2 (3.12); APOE (3.12); CMPK1 (3.12); EDNRB (3.12); TAAR1 (3.12); GUCA2A (6.25); GIPR (3.12); NPR3 (3.12); ATP5EP2 (3.12); GUCA2B (6.25); ADSL (3.12); STOML2 (3.12); SLC26A1 (18.75); HPCA (3.12); GALR2 (3.12); HPRT1 (40.62); ATP5H (3.12); CTPS1 (6.25); ATP5L (3.12); AKAP12 (3.12); ADGRG6 (3.12); PRPS1 (3.12)

						(5.26); PFAS (7.89); CMPK1 (7.89); ADCY1 (2.63); ACAT1 (7.89); ATP5F1 (5.26); EDNRB (21.05); RAF1 (5.26); ATP5E (7.89); TAAR1 (18.42); GUCA2A (7.89); NOS1 (7.89); GMPS (5.26); GIPR (5.26); NPR3 (15.79); ATP5EP2 (2.63); MC4R (5.26); AKAP5 (10.53); NPY2R (13.16); ADRB2 (5.26); TMIGD3 (5.26); ADORA3 (5.26); ADNP (10.53); GUCA2B (7.89); ADCY3 (7.89); ADSL (15.79); RUNDC3A (5.26); PAICS (15.79); CAD (7.89); NME1 (5.26); ATP5A1 (5.26); GRM2 (5.26); NME2 (5.26); STOML2 (18.42); SLC25A13 (2.63); GABBR1 (2.63); PAPSS2 (5.26); SLC26A1 (13.16); HPCA (7.89); STAT3 (5.26); GART (21.05); GALR2 (7.89); HPRT1 (28.95); ADORA2B (7.89); LTB4R2 (10.53); ATP5H (10.53); CTPS1 (7.89); IMPDH2 (5.26); ADRB1 (7.89); CYC1 (7.89); GPER1 (2.63); GUCY1B2 (2.63); ATP5G2 (2.63); ADCY4 (10.53); FLCN (7.89); ATP5L (13.16); PAPSS1 (13.16); NPR2 (15.79); ADGRD1 (7.89); ATP6V0A2 (7.89); P2RY11 (7.89); GPR161 (5.26); AKAP12 (21.05); PTK2B (39.47); GUCY2D (7.89); ADGRG6 (18.42); PRPS1 (10.53); RAMP2 (5.26); ATP5J2 (2.63); ADRB3 (5.26)	
GO:0009119	ribonucleoside metabolic process	-6.410	0.009	73.68	50	TXNDC9 (5.26); GMPR2 (10.53); GNAI3 (5.26); AHCY (13.16); CARD11 (2.63); PDCL2 (15.79); UMPS (5.26); NT5E (18.42); DLG3 (21.05); UCK1 (2.63); RHOA (5.26); NME5 (5.26); PDCL (2.63); MPP3 (5.26); CMPK1 (7.89); DLG4 (5.26); ADAL (7.89); NUDT18 (5.26); NT5C1A (7.89); APOBEC3C (13.16); GMPS (5.26); NT5C2 (7.89); NFS1 (2.63); MOCOS (5.26); ADA (7.89); RAB23 (13.16); ENPP4 (18.42); AK4 (7.89); CAD (7.89); NME1 (5.26); NME2 (5.26); AHCYL1 (5.26); PEMT (7.89); APOBEC2 (15.79); ENTPD4 (5.26); MOCS3 (10.53); HPRT1 (28.95); DGUOK (7.89); CTPS1 (7.89); IMPDH2 (5.26); MOCS1 (15.79); TJP2 (2.63); GPHN (10.53); SCRIB (42.11)	DLG3 (31.25); CMPK1 (3.12); NT5C1A (6.25); AK4 (3.12); HPRT1 (40.62); CTPS1 (6.25)
GO:0042454	ribonucleoside catabolic process	-9.301	0.014	47.37	40.62	AHCY (13.16); ADAL (7.89); NUDT18 (5.26); APOBEC3C (13.16); ADA (7.89); ENPP4 (18.42); APOBEC2 (15.79); ENTPD4 (5.26); HPRT1 (28.95)	HPRT1 (40.62)
GO:0046128	purine ribonucleoside metabolic process	-6.301	0.010	68.42	50	GMPR2 (10.53); GNAI3 (5.26); AHCY (13.16); CARD11 (2.63); NT5E (18.42); DLG3 (21.05); RHOA (5.26); NME5 (5.26); MPP3 (5.26); DLG4 (5.26); ADAL (7.89); NUDT18 (5.26); NT5C1A (7.89); GMPS (5.26); NT5C2 (7.89); NFS1 (2.63); MOCOS (5.26); ADA (7.89); RAB23 (13.16); ENPP4 (18.42); AK4 (7.89); NME1 (5.26); NME2 (5.26); AHCYL1 (5.26); PEMT (7.89); MOCS3 (10.53); HPRT1 (28.95); DGUOK (7.89); IMPDH2 (5.26); MOCS1 (15.79); TJP2 (2.63); GPHN (10.53); SCRIB (42.11)	DLG3 (31.25); NT5C1A (6.25); AK4 (3.12); HPRT1 (40.62)
GO:0046129	purine ribonucleoside biosynthetic process	-7.442	0.012	42.11	40.62	NT5E (18.42); NME5 (5.26); ADAL (7.89); GMPS (5.26); ADA (7.89); NME1 (5.26); NME2 (5.26); HPRT1 (28.95); IMPDH2 (5.26)	HPRT1 (40.62)
GO:0009156	ribonucleoside monophosphate biosynthetic process	-6.532	0.013	26.32	3.12	GART (21.05); PFAS (7.89); PRPS1 (10.53)	PRPS1 (3.12)

GO:0009167	purine ribonucleoside monophosphate metabolic process	-4.665	0.011	81.58	62.5	NDUFAB1 (5.26); NDUFB11 (7.89); SDHAF2 (7.89); COX7A2L (7.89); SDHD (10.53); GMPR2 (10.53); SLC25A23 (7.89); TPI1 (13.16); NUDT9 (13.16); ATP5G3 (10.53); PPIF (7.89); UQCC2 (15.79); CARD11 (2.63); COX15 (5.26); PFKFB1 (7.89); HKDC1 (7.89); COX8C (10.53); VCP (18.42); NMNAT1 (10.53); AMPD1 (5.26); ATP5B (2.63); CLPX (10.53); CCNB1 (15.79); HSPA8 (10.53); P2RX7 (2.63); MSH2 (7.89); NT5E (18.42); NDUFAF7 (5.26); CDK1 (7.89); HTR2A (18.42); BAD (7.89); MYH3 (7.89); SURF1 (10.53); ATP6V1B2 (5.26); NDUFA8 (2.63); NDUFB9 (50); UQCC3 (7.89); DLG3 (21.05); SDHA (10.53); UQCRB (47.37); PGAM1 (7.89); ATP5J (21.05); ALDOC (5.26); PGM1 (7.89); DNM1L (7.89); PGAM4 (7.89); SDHC (5.26); PARK7 (10.53); ATIC (5.26); ATP6V1A (5.26); GPI (5.26); GPD1 (7.89); COX6B1 (7.89); RHOA (5.26); GADD45GIP1 (5.26); NDUFA11 (7.89); MECP2 (2.63); ATP1A2 (5.26); NDUFA6 (13.16); EIF6 (10.53); MPP3 (5.26); FIGNL1 (2.63); COX6C (44.74); DLD (2.63); SLC25A25 (2.63); AMPD2 (5.26); NDUFV2 (5.26); OLA1 (10.53); AK2 (7.89); ALDOB (13.16); ENO2 (13.16); HK1 (7.89); ACTN3 (7.89); PRKAA1 (2.63); NUDT4 (7.89); COX7B (10.53); TEFM (5.26); NDUFV3 (2.63); PFAS (7.89); DLG4 (5.26); ATP5F1 (5.26); ATP5E (7.89); OGT (7.89); GMPS (5.26); NT5C2 (7.89); LEXM (7.89); ECD (5.26); NDUFB1 (10.53); ATP5EP2 (2.63); ARNT (5.26); NDUFS6 (13.16); NCOR1 (7.89); ADA (7.89); COQ9 (5.26); OGDHL (5.26); NDUFS5 (7.89); HIF1A (10.53); BPGM (2.63); ADSL (15.79); TIGAR (15.79); HSPA1B (15.79); UQCRHL (7.89); AK4 (7.89); NDUFAF1 (7.89); COX8A (7.89); PAICS (15.79); NDUFB7 (7.89); COA6 (5.26); AK1 (2.63); PGK1 (7.89); PGK2 (18.42); NDUFB5 (5.26); NDUFA2 (5.26); ATP5A1 (5.26); NDUFA7 (7.89); NDUFB4 (5.26); NDUFA3 (13.16); NUDT5 (7.89); COX6A1 (7.89); STOML2 (18.42); SLC25A13 (2.63); STAT3 (5.26); GART (21.05); NDUFC1 (13.16); NDUFB3 (7.89); HPRT1 (28.95); NDUFA12 (5.26); NUDT3 (15.79); COX4I2 (2.63); PFKL (2.63); ATP5H (10.53); IMPDH2 (5.26); MLXIPL (2.63); SIRT6 (7.89); CYC1 (7.89); NDUFA13 (7.89); COX6A2 (7.89); OGDH (2.63); ATP5G2 (2.63); NUDT10 (7.89); COX5A (7.89); DDIT4 (5.26); FLCN (7.89); ATP7A (15.79); ATP5L (13.16); TJP2 (2.63); ATP6V0A2 (7.89); CBFA2T3 (7.89); ENPP1 (18.42); ATP6V1B1 (7.89); IGF1 (2.63); ENO4 (7.89); PRPS1 (10.53); SCRIB (42.11); MYOG (5.26); ATP5J2 (2.63); UQCRQ (5.26)	VCP (3.12); NMNAT1 (3.12); CCNB1 (3.12); HTR2A (6.25); DLG3 (31.25); SDHA (3.12); PGM1 (3.12); PARK7 (3.12); MECP2 (9.38); AK2 (3.12); COX7B (3.12); LEXM (3.12); ATP5EP2 (3.12); NDUFS6 (3.12); NDUFS5 (6.25); ADSL (3.12); AK4 (3.12); NDUFB5 (3.12); NDUFA3 (3.12); STOML2 (3.12); HPRT1 (40.62); ATP5H (3.12); ATP7A (12.5); ATP5L (3.12); ENPP1 (3.12); PRPS1 (3.12)
GO:0009168	purine ribonucleoside monophosphate biosynthetic process	-6.476	0.013	31.58	3.12	ADA (7.89); ADSL (15.79); AMPD1 (5.26); AMPD2 (5.26); ATIC (5.26); GART (21.05); GMPS (5.26); IMPDH2 (5.26); PAICS (15.79); PFAS (7.89)	ADSL (3.12)
GO:0009185	ribonucleoside diphosphate metabolic process	-5.749	0.011	71.05	40.62	TPI1 (13.16); NUDT9 (13.16); CARD11 (2.63); PFKFB1 (7.89); HKDC1 (7.89); P2RX7 (2.63); HTR2A (18.42); BAD (7.89); DLG3 (21.05); PGAM1 (7.89); ALDOC (5.26);	HTR2A (6.25); DLG3 (31.25); PGM1 (3.12); AK2 (3.12)

						PGM1 (7.89); PGAM4 (7.89); GPI (5.26); GPD1 (7.89); EIF6 (10.53); MPP3 (5.26); AK2 (7.89); ALDOB (13.16); ENO2 (13.16); HK1 (7.89); ACTN3 (7.89); PRKAA1 (2.63); DLG4 (5.26); NUDT18 (5.26); OGT (7.89); ECD (5.26); ARNT (5.26); NCOR1 (7.89); OGDHL (5.26); HIF1A (10.53); BPGM (2.63); TIGAR (15.79); PGK1 (7.89); PGK2 (18.42); NUDT5 (7.89); ENTPD4 (5.26); STAT3 (5.26); PFKL (2.63); MLXIPL (2.63); SIRT6 (7.89); OGDH (2.63); ENTPD2 (2.63); DDIT4 (5.26); TJP2 (2.63); CBFA2T3 (7.89); IGF1 (2.63); ENO4 (7.89); SCRIB (42.11); MYOG (5.26)	
GO:0006140	regulation of nucleotide metabolic process	-4.503	0.013	78.95	50	SLC25A23 (7.89); GNAI3 (5.26); GUCY1A3 (15.79); LPAR1 (2.63); CRH (47.37); DRD3 (5.26); PPIF (7.89); EDNRA (13.16); UQCC2 (15.79); CXCL9 (13.16); VCP (18.42); WFS1 (10.53); MC3R (7.89); CCNB1 (15.79); HRH3 (5.26); AKAP9 (2.63); GUCA1A (18.42); PTGIR (7.89); P2RX7 (2.63); NF1 (21.05); CDK1 (7.89); HTR2A (18.42); HTR1B (18.42); PDE5A (13.16); RXFP2 (31.58); GPR65 (15.79); GNAI2 (5.26); ACMSD (5.26); APLP1 (10.53); AVP (2.63); PGAM1 (7.89); TSHR (10.53); SSTR4 (13.16); ADCY7 (7.89); CCR2 (5.26); DNM1L (7.89); PGAM4 (7.89); GHRH (7.89); VIP (18.42); PTH (7.89); DRD1 (5.26); CALCA (7.89); PARK7 (10.53); GPD1 (7.89); OPRL1 (2.63); PDZD3 (10.53); RHOA (5.26); GADD45GIP1 (5.26); TBL1XR1 (5.26); MRAP (2.63); RCVRN (2.63); EIF6 (10.53); RLN2 (18.42); ACTN3 (7.89); MAPK7 (5.26); PRKAA1 (2.63); CXCL10 (13.16); GUCA1B (18.42); GALR3 (15.79); APOE (7.89); ADCY1 (2.63); EDNRB (21.05); RAF1 (5.26); OGT (7.89); EGLN1 (7.89); GUCA2A (7.89); NOS1 (7.89); GIPR (5.26); LEXM (7.89); CHGA (7.89); NPR3 (15.79); ECD (5.26); MC4R (5.26); ARNT (5.26); AKAP5 (10.53); NPY2R (13.16); NCOR1 (7.89); ADRB2 (5.26); TMIGD3 (5.26); ADORA3 (5.26); ADNP (10.53); HIF1A (10.53); GUCA2B (7.89); BPGM (2.63); ADCY3 (7.89); TIGAR (15.79); RUNDC3A (5.26); GRM2 (5.26); NME2 (5.26); GABBR1 (2.63); HPCA (7.89); STAT3 (5.26); GALR2 (7.89); ADORA2B (7.89); LTB4R2 (10.53); MLXIPL (2.63); ADRB1 (7.89); SIRT6 (7.89); GPER1 (2.63); ADCY4 (10.53); DDIT4 (5.26); FLCN (7.89); ATP7A (15.79); ADGRD1 (7.89); P2RY11 (7.89); CBFA2T3 (7.89); GPR161 (5.26); AKAP12 (21.05); FZD2 (5.26); PTK2B (39.47); IGF1 (2.63); ADGRG6 (18.42); RAMP2 (5.26); MYOG (5.26); ADRB3 (5.26)	VCP (3.12); CCNB1 (3.12); AKAP9 (6.25); PTGIR (3.12); NF1 (34.38); HTR2A (6.25); RXFP2 (12.5); GPR65 (9.38); VIP (3.12); PARK7 (3.12); TBL1XR1 (3.12); RLN2 (3.12); APOE (3.12); EDNRB (3.12); GUCA2A (6.25); GIPR (3.12); LEXM (3.12); NPR3 (3.12); GUCA2B (6.25); HPCA (3.12); GALR2 (3.12); ATP7A (12.5); AKAP12 (3.12); ADGRG6 (3.12)
GO:0006164	purine nucleotide biosynthetic process	-6.167	0.012	52,63	43,75	ADSL (15.79); ATIC (5.26); GART (21.05); GMPR2 (10.53); GMPS (5.26); HPRT1 (28.95); IMPDH2 (5.26); MTHFD1 (10.53); OAS1 (2.63); PAICS (15.79); PFAS (7.89); PRPS1 (10.53)	ADSL (3.12); HPRT1 (40.62); PRPS1 (3.12)
GO:1900542	regulation of purine nucleotide metabolic process	-4.326	0.013	78.95	50	SLC25A23 (7.89); GNAI3 (5.26); GUCY1A3 (15.79); LPAR1 (2.63); CRH (47.37); DRD3 (5.26); PPIF (7.89); EDNRA (13.16); UQCC2 (15.79); CXCL9 (13.16); VCP (18.42); WFS1 (10.53); MC3R (7.89); CCNB1 (15.79);	VCP (3.12); CCNB1 (3.12); AKAP9 (6.25); PTGIR (3.12); NF1 (34.38); HTR2A (6.25); RXFP2 (12.5); GPR65 (9.38); VIP (3.12); PARK7 (3.12); TBL1XR1 (3.12); RLN2 (3.12); APOE (3.12); EDNRB (3.12);

						HRH3 (5.26); AKAP9 (2.63); GUCA1A (18.42); PTGIR (7.89); P2RX7 (2.63); NF1 (21.05); CDK1 (7.89); HTR2A (18.42); HTR1B (18.42); PDE5A (13.16); RXFP2 (31.58); GPR65 (15.79); GNAI2 (5.26); APLP1 (10.53); AVP (2.63); PGAM1 (7.89); TSHR (10.53); SSTR4 (13.16); ADCY7 (7.89); CCR2 (5.26); DNM1L (7.89); GHRH (7.89); VIP (18.42); PTH (7.89); DRD1 (5.26); CALCA (7.89); PARK7 (10.53); GPD1 (7.89); OPRL1 (2.63); PDZD3 (10.53); RHOA (5.26); GADD45GIP1 (5.26); TBL1XR1 (5.26); MRAP (2.63); RCVRN (2.63); EIF6 (10.53); RLN2 (18.42); ACTN3 (7.89); MAPK7 (5.26); PRKAA1 (2.63); CXCL10 (13.16); GUCA1B (18.42); GALR3 (15.79); APOE (7.89); ADCY1 (2.63); EDNRB (21.05); RAF1 (5.26); OGT (7.89); EGLN1 (7.89); GUCA2A (7.89); NOS1 (7.89); GIPR (5.26); LEXM (7.89); CHGA (7.89); NPR3 (15.79); ECD (5.26); MC4R (5.26); ARNT (5.26); AKAP5 (10.53); NPY2R (13.16); NCOR1 (7.89); ADRB2 (5.26); TMIGD3 (5.26); ADORA3 (5.26); ADNP (10.53); HIF1A (10.53); GUCA2B (7.89); ADCY3 (7.89); TIGAR (15.79); RUNDC3A (5.26); GRM2 (5.26); NME2 (5.26); GABBR1 (2.63); HPCA (7.89); STAT3 (5.26); GALR2 (7.89); ADORA2B (7.89); LTB4R2 (10.53); MLXIPL (2.63); ADRB1 (7.89); SIRT6 (7.89); GPER1 (2.63); ADCY4 (10.53); DDIT4 (5.26); FLCN (7.89); ATP7A (15.79); ADGRD1 (7.89); P2RY11 (7.89); CBFA2T3 (7.89); GPR161 (5.26); AKAP12 (21.05); FZD2 (5.26); PTK2B (39.47); IGF1 (2.63); ADGRG6 (18.42); RAMP2 (5.26); MYOG (5.26); ADRB3 (5.26)	GUCA2A (6.25); GIPR (3.12); LEXM (3.12); NPR3 (3.12); GUCA2B (6.25); HPCA (3.12); GALR2 (3.12); ATP7A (12.5); AKAP12 (3.12); ADGRG6 (3.12)
GO:0006195	purine nucleotide catabolic process	-6.267	0.013	28,95	9,38	DNPH1 (18.42); GDA (10.53); GPX1 (5.26); ITPA (2.63); NT5C (10.53); NT5C1A (7.89); NT5C2 (7.89); NT5E (18.42); NUDT15 (2.63)	NT5C (3.12); NT5C1A (6.25)
GO:0009259	ribonucleotide metabolic process	-5.708	0.011	18,42	3,12	ATIC (5.26); RNASEH2B (18.42)	RNASEH2B (3.12)
GO:0009260	ribonucleotide biosynthetic process	-6.587	0.012	84.21	59.38	GNAI3 (5.26); GUCY1A3 (15.79); ATP5G3 (10.53); LPAR1 (2.63); CRH (47.37); DRD3 (5.26); RFK (10.53); PANK3 (7.89); EDNRA (13.16); PANK2 (2.63); VCP (18.42); AMPD1 (5.26); ATP5B (2.63); WFS1 (10.53); MC3R (7.89); HRH3 (5.26); AKAP9 (2.63); GUCA1A (18.42); PTGIR (7.89); UMPS (5.26); NF1 (21.05); NPPB (10.53); HTR1B (18.42); SURF1 (10.53); DCAKD (5.26); RXFP2 (31.58); GPR65 (15.79); UQCC3 (7.89); GNAI2 (5.26); APLP1 (10.53); AVP (2.63); UCK1 (2.63); ATP5J (21.05); TSHR (10.53); GUCY2F (7.89); ADCY7 (7.89); CCR2 (5.26); GHRH (7.89); VIP (18.42); PTH (7.89); DRD1 (5.26); CALCA (7.89); ATIC (5.26); NPPA (10.53); OPRL1 (2.63); PDZD3 (10.53); PPCS (7.89); PANK1 (7.89); MRAP (2.63); GUCY2C (5.26); RCVRN (2.63); NME5 (5.26); RLN2 (18.42); AMPD2 (5.26); AK2 (7.89); COASY (5.26); GUCA1B (18.42); GALR3 (15.79); APOE (7.89); ADCY10 (5.26); PFAS (7.89); CMPK1 (7.89); ADCY1 (2.63); ACAT1 (7.89); ATP5F1 (5.26); EDNRB (21.05); RAF1 (5.26); ATP5E (7.89); TAAR1 (18.42); GUCA2A (7.89); NOS1	VCP (3.12); AKAP9 (6.25); PTGIR (3.12); NF1 (34.38); NPPB (3.12); RXFP2 (12.5); GPR65 (9.38); VIP (3.12); NPPA (3.12); PPCS (6.25); RLN2 (3.12); AK2 (3.12); APOE (3.12); CMPK1 (3.12); EDNRB (3.12); TAAR1 (3.12); GUCA2A (6.25); GIPR (3.12); NPR3 (3.12); ATP5EP2 (3.12); GUCA2B (6.25); ADSL (3.12); STOML2 (3.12); SLC26A1 (18.75); HPCA (3.12); GALR2 (3.12); HPRT1 (40.62); ATP5H (3.12); CTPS1 (6.25); ATP5L (3.12); AKAP12 (3.12); ADGRG6 (3.12); PRPS1 (3.12)

						(7.89); GMPS (5.26); GIPR (5.26); NPR3 (15.79); ATP5EP2 (2.63); MC4R (5.26); AKAP5 (10.53); NPY2R (13.16); ADRB2 (5.26); TMIGD3 (5.26); ADORA3 (5.26); ADNP (10.53); GUCA2B (7.89); ADCY3 (7.89); ADSL (15.79); RUNDC3A (5.26); PAICS (15.79); CAD (7.89); NME1 (5.26); ATP5A1 (5.26); GRM2 (5.26); NME2 (5.26); STOML2 (18.42); SLC25A13 (2.63); GABBR1 (2.63); PAPSS2 (5.26); SLC26A1 (13.16); HPCA (7.89); STAT3 (5.26); GART (21.05); GALR2 (7.89); HPRT1 (28.95); ADORA2B (7.89); LTB4R2 (10.53); ATP5H (10.53); CTPS1 (7.89); IMPDH2 (5.26); ADRB1 (7.89); CYC1 (7.89); GPER1 (2.63); GUCY1B2 (2.63); ATP5G2 (2.63); ADCY4 (10.53); FLCN (7.89); ATP5L (13.16); PAPSS1 (13.16); NPR2 (15.79); ADGRD1 (7.89); ATP6V0A2 (7.89); P2RY11 (7.89); GPR161 (5.26); AKAP12 (21.05); PTK2B (39.47); GUCY2D (7.89); ADGRG6 (18.42); PRPS1 (10.53); RAMP2 (5.26); ATP5J2 (2.63); ADRB3 (5.26)	
GO:0009261	ribonucleotide catabolic process	-6.581	0.013	65.79	40.62	NUDT9 (13.16); PDE7A (47.37); ITPA (2.63); NT5E (18.42); PDE5A (13.16); MAPK7 (5.26); PDE4C (7.89); NUDT4 (7.89); NUDT18 (5.26); EGLN1 (7.89); ENTPD4 (5.26); HPRT1 (28.95); NUDT3 (15.79); PDE1B (2.63); NUDT10 (7.89)	HPRT1 (40.62)
GO:0009150	purine ribonucleotide metabolic process	-5.418	0.011	84.21	62.5	NDUFAB1 (5.26); NDUFB11 (7.89); SDHAF2 (7.89); COX7A2L (7.89); SDHD (10.53); GMPR2 (10.53); SLC25A23 (7.89); GNAI3 (5.26); GUCY1A3 (15.79); TPI1 (13.16); NUDT9 (13.16); PDE7A (47.37); ATP5G3 (10.53); LPAR1 (2.63); CRH (47.37); DRD3 (5.26); PANK3 (7.89); PPIF (7.89); EDNRA (13.16); UQCC2 (15.79); CARD11 (2.63); ITPA (2.63); COX15 (5.26); CXCL9 (13.16); PANK2 (2.63); PFKFB1 (7.89); HKDC1 (7.89); COX8C (10.53); VCP (18.42); NMNAT1 (10.53); AMPD1 (5.26); ATP5B (2.63); CLPX (10.53); WFS1 (10.53); MC3R (7.89); CCNB1 (15.79); HRH3 (5.26); AKAP9 (2.63); GUCA1A (18.42); PTGIR (7.89); HSPA8 (10.53); P2RX7 (2.63); MSH2 (7.89); NF1 (21.05); NT5E (18.42); NDUFAF7 (5.26); CDK1 (7.89); NPPB (10.53); HTR2A (18.42); HTR1B (18.42); BAD (7.89); PDE5A (13.16); MYH3 (7.89); SURF1 (10.53); DCAKD (5.26); ATP6V1B2 (5.26); NDUFA8 (2.63); NDUFB9 (50); RXFP2 (31.58); GPR65 (15.79); UQCC3 (7.89); GNAI2 (5.26); DLG3 (21.05); APLP1 (10.53); AVP (2.63); SDHA (10.53); UQCRB (47.37); PGAM1 (7.89); ATP5J (21.05); ALDOC (5.26); TSHR (10.53); SSTR4 (13.16); GUCY2F (7.89); ADCY7 (7.89); CCR2 (5.26); PGM1 (7.89); DNM1L (7.89); PGAM4 (7.89); GHRH (7.89); VIP (18.42); PTH (7.89); DRD1 (5.26); CALCA (7.89); SDHC (5.26); PARK7 (10.53); ATIC (5.26); ATP6V1A (5.26); GPI (5.26); GPD1 (7.89); NPPA (10.53); COX6B1 (7.89); OPRL1 (2.63); PDZD3 (10.53); PPCS (7.89); RHOA (5.26); GADD45GIP1 (5.26); TBL1XR1 (5.26); NDUFA11 (7.89); PANK1 (7.89); MECP2 (2.63); MRAP (2.63); GUCY2C (5.26); ATP1A2 (5.26); RCVRN	VCP (3.12); NMNAT1 (3.12); CCNB1 (3.12); AKAP9 (6.25); PTGIR (3.12); NF1 (34.38); NPPB (3.12); HTR2A (6.25); RXFP2 (12.5); GPR65 (9.38); DLG3 (31.25); SDHA (3.12); PGM1 (3.12); VIP (3.12); PARK7 (3.12); NPPA (3.12); PPCS (6.25); TBL1XR1 (3.12); MECP2 (9.38); RLN2 (3.12); AK2 (3.12); COX7B (3.12); APOE (3.12); EDNRB (3.12); TAAR1 (3.12); GUCA2A (6.25); GIPR (3.12); LEXM (3.12); NPR3 (3.12); ATP5EP2 (3.12); NDUFS6 (3.12); NDUFS5 (6.25); GUCA2B (6.25); ADSL (3.12); AK4 (3.12); NDUFB5 (3.12); NDUFA3 (3.12); STOML2 (3.12); SLC26A1 (18.75); HPCA (3.12); GALR2 (3.12); HPRT1 (40.62); ATP5H (3.12); MCCC2 (3.12); ATP7A (12.5); ATP5L (3.12); ENPP1 (3.12); AKAP12 (3.12); ADGRG6 (3.12); PRPS1 (3.12)

						(2.63); NME5 (5.26); NDUFA6 (13.16); EIF6 (10.53); MPP3 (5.26); FIGNL1 (2.63); COX6C (44.74); DLD (2.63); SLC25A25 (2.63); RLN2 (18.42); AMPD2 (5.26); NDUFV2 (5.26); OLA1 (10.53); AK2 (7.89); ALDOB (13.16); ENO2 (13.16); HK1 (7.89); ACTN3 (7.89); MAPK7 (5.26); PDE4C (7.89); PRKAA1 (2.63); NUDT4 (7.89); CXCL10 (13.16); COX7B (10.53); COASY (5.26); TEFM (5.26); GUCA1B (18.42); GALR3 (15.79); APOE (7.89); NDUFV3 (2.63); ADCY10 (5.26); PFAS (7.89); DLG4 (5.26); ADCY1 (2.63); NUDT18 (5.26); ACAT1 (7.89); ATP5F1 (5.26); EDNRB (21.05); RAF1 (5.26); HMGCR (7.89); ATP5E (7.89); OGT (7.89); EGLN1 (7.89); TAAR1 (18.42); GUCA2A (7.89); NOS1 (7.89); GMPS (5.26); NT5C2 (7.89); GIPR (5.26); LEXM (7.89); NFS1 (2.63); CHGA (7.89); NPR3 (15.79); ECD (5.26); NDUFB1 (10.53); ATP5EP2 (2.63); MOCOS (5.26); MC4R (5.26); ARNT (5.26); AKAP5 (10.53); BPNT1 (7.89); NPY2R (13.16); NDUFS6 (13.16); NCOR1 (7.89); ADRB2 (5.26); TMIGD3 (5.26); ADORA3 (5.26); COQ9 (5.26); OGDHL (5.26); RAB23 (13.16); NDUFS5 (7.89); ADNP (10.53); HIF1A (10.53); GUCA2B (7.89); BPGM (2.63); SULT1E1 (13.16); ADCY3 (7.89); ADSL (15.79); TIGAR (15.79); HSPA1B (15.79); UQCRHL (7.89); AK4 (7.89); NDUFAF1 (7.89); RUNDC3A (5.26); COX8A (7.89); PAICS (15.79); NDUFB7 (7.89); COA6 (5.26); SULT1B1 (13.16); AK1 (2.63); PGK1 (7.89); PGK2 (18.42); NDUFB5 (5.26); NDUFA2 (5.26); NME1 (5.26); ATP5A1 (5.26); GRM2 (5.26); NDUFA7 (7.89); NDUFB4 (5.26); NDUFA3 (13.16); NME2 (5.26); NUDT5 (7.89); COX6A1 (7.89); STOML2 (18.42); SLC25A13 (2.63); GABBR1 (2.63); MOCS3 (10.53); PAPSS2 (5.26); SLC26A1 (13.16); HPCA (7.89); STAT3 (5.26); GART (21.05); NDUFC1 (13.16); NDUFB3 (7.89); GALR2 (7.89); HPRT1 (28.95); NDUFA12 (5.26); NUDT3 (15.79); COX4I2 (2.63); PDE1B (2.63); ADORA2B (7.89); PFKL (2.63); LTB4R2 (10.53); ATP5H (10.53); IMPDH2 (5.26); MLXIPL (2.63); MCCC2 (13.16); MOCS1 (15.79); ADRB1 (7.89); SIRT6 (7.89); CYC1 (7.89); NDUFA13 (7.89); COX6A2 (7.89); OGDH (2.63); GPER1 (2.63); GUCY1B2 (2.63); ATP5G2 (2.63); NUDT10 (7.89); ADCY4 (10.53); COX5A (7.89); DDIT4 (5.26); FLCN (7.89); ATP7A (15.79); IMPAD1 (50); ATP5L (13.16); PAPSS1 (13.16); NPR2 (15.79); TJP2 (2.63); ADGRD1 (7.89); ATP6V0A2 (7.89); P2RY11 (7.89); CBFA2T3 (7.89); ENPP1 (18.42); GPR161 (5.26); SULT6B1 (5.26); AKAP12 (21.05); FZD2 (5.26); ATP6V1B1 (7.89); TPST2 (18.42); PTK2B (39.47); GUCY2D (7.89); IGF1 (2.63); GPHN (10.53); ENO4 (7.89); ADGRG6 (18.42); PRPS1 (10.53); SCRIB (42.11); RAMP2 (5.26); MYOG (5.26); ATP5J2 (2.63); ADRB3 (5.26); UQCRQ (5.26)	
GO:0009152	purine ribonucleotide biosynthetic process	-6.177	0.012	15,79	3,12	ADSL (15.79)	ADSL (3.12)

GO:0009154	purine ribonucleotide catabolic process	-6.444	0.013	65.79	40.62	NUDT9 (13.16); PDE7A (47.37); ITPA (2.63); NT5E (18.42); PDE5A (13.16); MAPK7 (5.26); PDE4C (7.89); NUDT4 (7.89); NUDT18 (5.26); EGLN1 (7.89); HPRT1 (28.95); NUDT3 (15.79); PDE1B (2.63); NUDT10 (7.89)	HPRT1 (40.62)
GO:0008643	carbohydrate transport	-5.810	0.009	57,89	21,88	AKT1 (10.53); AKT2 (7.89); PEA15 (5.26); SLC2A1 (10.53); SLC2A10 (7.89); SLC2A11 (28.95); SLC2A12 (18.42); SLC2A2 (5.26); SLC2A4 (5.26); SLC2A5 (10.53); SLC2A6 (10.53); SLC2A8 (10.53); SLC35A4 (5.26); SLC35A5 (5.26); SLC35C1 (7.89); SLC35D1 (7.89); SLC35D2 (13.16); SLC35D3 (15.79); SLC37A1 (2.63); SLC37A2 (10.53); SLC37A4 (10.53); SLC45A1 (10.53); SLC5A11 (5.26); TMEM241 (5.26)	SLC2A1 (6.25); SLC2A11 (12.5); SLC2A12 (3.12); SLC2A2 (3.12); SLC2A5 (3.12); SLC35D3 (3.12); SLC45A1 (3.12)
GO:1901136	carbohydrate derivative catabolic process	-4.572	0.011	84.21	78.12	NUDT9 (13.16); AHCY (13.16); GPD1L (5.26); PDE7A (47.37); NT5C (10.53); GBA2 (15.79); HYAL2 (5.26); OVGP1 (5.26); NEIL2 (44.74); CD44 (7.89); HYAL1 (5.26); SDC2 (7.89); ITPA (2.63); UNG (7.89); ACAN (7.89); SMUG1 (7.89); STAB2 (7.89); NT5E (18.42); CST3 (10.53); CHP1 (7.89); PDE5A (13.16); LYVE1 (7.89); OMD (13.16); SDC4 (7.89); GLA (10.53); CEMIP (10.53); NEU1 (15.79); TMEM2 (2.63); CSPG4 (10.53); SGSH (10.53); GPC3 (50); STT3B (5.26); DUT (7.89); LUM (7.89); GPD1 (7.89); NTHL1 (5.26); GNPDA1 (5.26); NUDT15 (2.63); CTBS (5.26); FBXO6 (10.53); MAPK7 (5.26); PDE4C (7.89); NUDT4 (7.89); ABHD10 (5.26); HYAL4 (2.63); ADAL (7.89); NUDT18 (5.26); NT5C1A (7.89); FUCA1 (10.53); CHIT1 (5.26); EDEM2 (10.53); KERA (7.89); EGLN1 (7.89); FMOD (5.26); FGF2 (13.16); DCTPP1 (7.89); APOBEC3C (13.16); DCN (2.63); ADA (7.89); PNLIPRP2 (7.89); GNS (2.63); PGLYRP3 (5.26); GALNS (7.89); GM2A (5.26); HGSNAT (5.26); ENPP4 (18.42); FBXO2 (10.53); BCAN (5.26); OGN (13.16); HYAL3 (5.26); GBA3 (15.79); HPSE (13.16); APOBEC2 (15.79); ENTPD4 (5.26); PGLYRP1 (7.89); HPRT1 (28.95); NEIL1 (10.53); NUDT3 (15.79); GPC4 (52.63); HEXB (10.53); PDE1B (2.63); NCAN (7.89); PRKCD (5.26); NUDT10 (7.89); NAGA (13.16); SAMHD1 (7.89); PGM2 (13.16); NEU2 (7.89); CSPG5 (5.26); NAGLU (5.26); DPYS (44.74)	NT5C (3.12); GLA (3.12); CEMIP (3.12); SGSH (3.12); GPC3 (65.62); FBXO6 (3.12); NT5C1A (6.25); FBXO2 (3.12); GBA3 (3.12); PGLYRP1 (3.12); HPRT1 (40.62); GPC4 (65.62)
GO:0015758	glucose transport	-7.026	0.009	57,89	12,5	AKT1 (10.53); BRAF (2.63); EDNRA (13.16); FABP5 (39.47); G6PC3 (5.26); HK1 (7.89); PLA2G1B (7.89); PPARD (13.16); PRKAG3 (5.26); SLC2A1 (10.53); SLC2A10 (7.89); SLC2A12 (18.42); SLC2A2 (5.26); SLC2A4 (5.26); SLC2A5 (10.53); SLC2A8 (10.53); SLC37A4 (10.53); SLC45A3 (5.26); STXBP3 (5.26)	SLC2A1 (6.25); SLC2A12 (3.12); SLC2A2 (3.12); SLC2A5 (3.12)
GO:0010827	regulation of glucose transport	-6.931	0.009	44,74	12,5	AAAS (7.89); FFAR4 (5.26); NDC1 (7.89); NUP133 (7.89); NUP155 (2.63); NUP188 (10.53); NUP210 (5.26); NUP214 (10.53); NUP37 (7.89); NUP43 (18.42); NUP58 (21.05); NUP85 (10.53); NUP98 (7.89); NUPL2 (5.26); RAE1 (5.26); RANBP2 (5.26); TPR (5.26); TRIB3 (10.53)	NDC1 (3.12); NUP43 (3.12); NUP58 (3.12); NUP85 (3.12)
GO:0010828	positive regulation of glucose transport	-6.953	0.009	13,16	0	CLIP3 (7.89); NR4A3 (13.16)	

GO:0046323	glucose import	-6.516	0.009	23,68	9,38	DRD1 (5.26); HNF1A (7.89); SLC2A10 (7.89); SLC2A12 (18.42); SLC2A4 (5.26); SLC2A6 (10.53); SLC2A8 (10.53); SORT1 (5.26); TSC1 (10.53)	SLC2A12 (3.12); TSC1 (6.25)
GO:0046324	regulation of glucose import	-6.451	0.009	18,42	6,25	APPL1 (5.26); ASPSCR1 (10.53); RTN2 (5.26); SLC25A27 (15.79)	ASPSCR1 (3.12); RTN2 (3.12)
GO:0046326	positive regulation of glucose import	-7.037	0.009	65,79	65,62	ADIPOQ (5.26); AKT1 (10.53); AKT2 (7.89); C1QTNF2 (5.26); CLTCL1 (18.42); GPC3 (50); IGF1 (2.63); ITLN1 (5.26); NFE2L2 (10.53); PIK3R1 (13.16); PTH (7.89); RAP1A (5.26); TERT (13.16)	ADIPOQ (3.12); GPC3 (65.62); TERT (3.12)
GO:0005977	glycogen metabolic process	-5.997	0.013	36,84	9,38	AKT1 (10.53); AKT2 (7.89); EPM2A (18.42); GNMT (15.79); IL6ST (7.89); PHKA1 (7.89); PPP1CA (7.89); PPP1CB (5.26); PPP1CC (7.89); PPP1R1A (2.63); PPP1R2P3 (5.26); PPP1R3B (5.26); PPP1R3C (7.89); PPP1R3D (7.89); PPP1R3E (10.53); PYGB (2.63); SLC37A4 (10.53); STBD1 (13.16); STK40 (7.89)	EPM2A (3.12); STK40 (6.25)
GO:1901607	alpha-amino acid biosynthetic process	-8.741	0.012	63.16	46.88	AHCY (13.16); MTHFD1 (10.53); CBS (2.63); SHMT2 (2.63); SLC1A3 (2.63); AGXT2 (2.63); SERINC3 (2.63); SEPHS2 (7.89); CPS1 (5.26); PLOD2 (5.26); GOT1L1 (44.74); NAGS (5.26); PARK7 (10.53); NOXRED1 (10.53); MRI1 (7.89); CTH (5.26); GLUD2 (10.53); CAD (7.89); GOT2 (5.26); OTC (34.21); ENOPH1 (13.16); GOT1 (7.89); ASNSD1 (10.53); GLS2 (2.63); AASS (2.63); AASDHPPT (10.53); ALDH4A1 (7.89)	PARK7 (3.12); OTC (46.88)
GO:0006525	arginine metabolic process	-9.549	0.013	18,42	3,12	ARG1 (18.42); ART4 (5.26); DDAH2 (15.79)	ARG1 (3.12)
GO:0009084	glutamine family amino acid biosynthetic process	-7.768	0.013	42.11	46.88	SLC1A3 (2.63); CPS1 (5.26); NAGS (5.26); NOXRED1 (10.53); GLUD2 (10.53); CAD (7.89); OTC (34.21); GLS2 (2.63); ALDH4A1 (7.89)	OTC (46.88)
BIOENERGETICS							
GO:0030819	positive regulation of cAMP biosynthetic process	-6.418	0.013	73.68	21.88	ADCY7 (7.89); ADORA2B (7.89); ADRB1 (7.89); AKAP12 (21.05); AKAP5 (10.53); AVP (2.63); CALCA (7.89); CRH (47.37); DRD1 (5.26); GHRH (7.89); GIPR (5.26); GNAS (10.53); GPER1 (2.63); GPR161 (5.26); GPR65 (15.79); MC3R (7.89); MC4R (5.26); MRAP (2.63); NME2 (5.26); PTGIR (7.89); PTH (7.89); RAMP2 (5.26); RLN2 (18.42); RXFP2 (31.58); TSHR (10.53)	AKAP12 (3.12); GIPR (3.12); GPR65 (9.38); PTGIR (3.12); RLN2 (3.12); RXFP2 (12.5)
GO:0046395	carboxylic acid catabolic process	-9.206	0.015	2.63	0	PON1 (2.63); PON3 (2.63)	
PROTEIN KINASE SIGNALING							

GO:0006469	negative regulation of protein kinase activity	-4.823	0.011	86.84	75	AKT1 (10.53); ASPN (13.16); CAMK2N2 (5.26); CAV3 (5.26); CDKN2A (18.42); CEP85 (10.53); CHAD (5.26); CHP1 (7.89); DBNDD2 (7.89); DCN (2.63); DEPTOR (7.89); EPHA1 (2.63); FABP4 (7.89); FGFR1OP (15.79); FLRT1 (7.89); FLRT2 (10.53); FLRT3 (2.63); GADD45A (7.89); GADD45B (7.89); GMFB (10.53); GNAQ (13.16); IL6 (5.26); INPP5K (5.26); LRP6 (5.26); LRRC15 (5.26); LRRC3 (2.63); LRRC3C (5.26); LRRC4 (2.63); LRRC4B (10.53); LRRTM1 (7.89); MLLT1 (7.89); NCK1 (15.79); NF1 (21.05); NF2 (52.63); NYX (7.89); PARK7 (10.53); PKIA (47.37); PKIG (10.53); PODNL1 (7.89); PPM1E (10.53); PPP1R1A (2.63); PPP1R1B (5.26); PREX1 (2.63); PRKAR1A (10.53); PSEN1 (10.53); PTPRC (5.26); QARS (5.26); RB1 (57.89); RGN (7.89); RTN4RL2 (7.89); SOCS1 (5.26); SOCS2 (2.63); SOCS3 (10.53); TARBP2 (7.89); TESC (2.63); TRIB1 (55.26); TRIB2 (7.89); TRIB3 (10.53); TRIM27 (15.79); TSC2 (5.26); UBASH3B (10.53); WARS (10.53); WWTR1 (5.26)	CAMK2N2 (3.12); CDKN2A (3.12); CHAD (3.12); FGFR1OP (3.12); FLRT2 (3.12); LRRC15 (3.12); NCK1 (9.38); NF1 (34.38); NF2 (56.25); PARK7 (3.12); PPM1E (6.25); PRKAR1A (3.12); RB1 (59.38); SOCS3 (3.12)
GO:0033673	negative regulation of kinase activity	-4.870	0.011	18.42	0	AJUBA (10.53); CDKN1B (5.26); CSK (7.89); MSTN (10.53); MYCNOS (7.89); NPRL2 (5.26)	
GO:0071901	negative regulation of protein serine/threonine kinase activity	-6.300	0.013	47.37	3.12	ABL1 (10.53); CDK5RAP3 (5.26); CDKN1B (5.26); CDKN2D (7.89); DAB2IP (2.63); FAM212A (5.26); FAM212B (5.26); HEXIM2 (5.26); LRP6 (5.26); PKIA (47.37); PKIG (10.53); PPP1R1B (5.26); PYCARD (7.89)	CDK5RAP3 (3.12)
GO:0000187	activation of MAPK activity	-6.589	0.012	76.32	43.75	ADORA2B (7.89); ADRA2B (5.26); ALK (5.26); APP (26.32); ARRB1 (10.53); AVPI1 (7.89); BMP2 (2.63); C1QTNF2 (5.26); C5 (13.16); C5AR1 (10.53); CD74 (5.26); CDK1 (7.89); CHRNA7 (15.79); CSPG4 (10.53); CXCR4 (7.89); DUSP6 (7.89); EGF (13.16); ERP29 (7.89); FCER1A (5.26); FGF10 (2.63); FGF2 (13.16); GHR (2.63); GHRL (5.26); HGF (2.63); IGF1 (2.63); IL1B (5.26); IQGAP3 (5.26); KARS (7.89); KIT (18.42); LPAR1 (2.63); MAP2K1 (10.53); MAP2K2 (7.89); MAP2K6 (10.53); MAP3K2 (5.26); MAPK1 (21.05); MAPKAPK3 (5.26); MAPKAPK5 (2.63); MOS (7.89); NOD1 (5.26); NOD2 (7.89); NRG1 (50); NTRK3 (13.16); P2RX7 (2.63); PAK3 (18.42); PDE6H (5.26); PEA15 (5.26); PLA2G1B (7.89); PRKAA1 (2.63); PROK1 (5.26); PROK2 (7.89); PTPN11 (23.68); RIPK2 (42.11); S1PR2 (7.89); SAA1 (7.89); SHC1 (5.26); SOD1 (2.63); SYK (13.16); TAB1 (15.79); TAB2 (18.42); TAB3 (7.89); TDGF1 (5.26); TGFB3 (10.53); TLR4 (2.63); TNF (10.53); UBA52 (7.89); UBB (7.89); UBC (2.63); UBE2N (7.89)	C5AR1 (3.12); CHRNA7 (25); MAP2K6 (3.12); MAPK10 (6.25); PAK3 (21.88); PTPN11 (34.38); TAB2 (3.12)

GO:0043406	positive regulation of MAP kinase activity	-4.448	0.013	71.05	37.5	AJUBA (10.53); CD40 (2.63); CSK (7.89); DIRAS2 (13.16); EDN3 (7.89); EGF (13.16); EGFR (5.26); ELANE (5.26); ERBB2 (5.26); EZH2 (2.63); FGF2 (13.16); FGFR1 (44.74); FLT1 (23.68); FLT3 (31.58); HRAS (5.26); HTR2A (18.42); KIT (18.42); KITLG (2.63); KRAS (13.16); MST1R (5.26); NEK10 (5.26); PDCD10 (5.26); PDE5A (13.16); PDGFA (2.63); PDGFB (13.16); PDGFRB (5.26); PIK3CG (2.63); PIK3R5 (7.89); PSEN1 (10.53); S100A12 (5.26); SRC (2.63); TAB1 (15.79); TNF (10.53); TNFSF11 (2.63)	EGFR (15.62); FLT1 (3.12); FLT3 (25); HTR2A (6.25); TENM1 (12.5)
GO:0051056	regulation of small GTPase mediated signal transduction	-5.159	0.011	65.79	43.75	AMOT (7.89); ARAP1 (10.53); ARAP3 (5.26); ARHGAP1 (7.89); ARHGAP12 (7.89); ARHGAP19 (7.89); ARHGAP21 (7.89); ARHGAP22 (5.26); ARHGAP25 (7.89); ARHGAP26 (5.26); ARHGAP29 (5.26); ARHGAP30 (5.26); ARHGAP33 (7.89); ARHGAP40 (7.89); ARHGAP44 (7.89); ARHGAP6 (10.53); ARHGAP9 (2.63); ARHGEF12 (10.53); ARHGEF19 (7.89); ARHGEF2 (5.26); ARHGEF26 (5.26); ARHGEF3 (5.26); ARHGEF6 (18.42); ARHGEF7 (5.26); ARHGEF9 (10.53); BCR (28.95); CHN1 (10.53); DEPDC1B (7.89); ECT2 (5.26); FGD1 (7.89); GARNL3 (2.63); GMIP (7.89); GNA13 (13.16); INPP5B (7.89); PIK3R2 (7.89); PLEKHG2 (7.89); PREX1 (2.63); RAC1 (5.26); RALGAPA1 (10.53); RALGAPA2 (15.79); RALGAPB (2.63); RHOA (5.26); RHOBTB1 (7.89); RHOBTB2 (5.26); RHOF (7.89); RHOG (7.89); RHOH (13.16); RHOV (7.89); SIPA1 (7.89); SIPA1L3 (10.53); SRGAP3 (5.26); TAGAP (15.79); TIAM1 (2.63); TRIP10 (7.89); TSC2 (5.26); VAV2 (2.63)	ARHGAP26 (3.12); ARHGAP6 (6.25); ARHGEF6 (18.75); ARHGEF9 (18.75); BCR (3.12); DEPDC1B (3.12); ECT2 (3.12); GNA13 (3.12); INPP5B (6.25); MCF2 (6.25); OPHN1 (3.12); TAGAP (3.12)
GO:0007265	Ras protein signal transduction	-3.614	0.013	71.05	71.88	BRAP (7.89); CCNA2 (13.16); CDK2 (5.26); CDKN1A (13.16); CDKN2A (18.42); CNKSR1 (10.53); DNMT1 (7.89); DOK1 (7.89); DOK2 (39.47); DOK3 (5.26); FGF2 (13.16); G3BP1 (5.26); G3BP2 (13.16); HRAS (5.26); IGF1 (2.63); IQGAP3 (5.26); JUN (7.89); KRAS (13.16); MAPKAPK3 (5.26); MAPKAPK5 (2.63); NF1 (21.05); NRAS (5.26); PARK7 (10.53); PLD1 (5.26); PLK2 (7.89); RALA (5.26); RALGDS (10.53); RAPGEF6 (5.26); RASSF1 (5.26); RB1 (57.89); RFXANK (7.89); RGL2 (15.79); RIT1 (5.26); SHC1 (5.26); SHTN1 (7.89); SYNGAP1 (15.79); TP53 (7.89); ZNF304 (13.16)	CDKN2A (3.12); JUN (3.12); NF1 (34.38); PARK7 (3.12); PLD1 (3.12); RB1 (59.38); ZNF304 (3.12)
GO:0046578	regulation of Ras protein signal transduction	-4.619	0.013	13.16	0	FOXM1 (13.16); SQSTM1 (5.26)	
GO:0046580	negative regulation of Ras protein signal transduction	-8.542	0.015	55.26	37.5	DAB2IP (2.63); MFN2 (10.53); NF1 (21.05); PPP2CB (42.11); RABGEF1 (2.63); RASA2 (5.26); RASAL1 (2.63); RASAL3 (7.89); SPRY2 (5.26); SYNGAP1 (15.79); TNK1 (5.26); TRIM67 (7.89)	MFN2 (3.12); NF1 (34.38)

GO:1904893	negative regulation of STAT cascade	-5.210	0.013	65.79	62.5	SOCS3 (10.53); SOCS2 (2.63); LRRC4 (2.63); SOCS4 (10.53); LRRC15 (5.26); RTN4RL2 (7.89); LRRTM1 (7.89); LRRC3C (5.26); NF2 (52.63); LRRC4B (10.53); PIBF1 (21.05); PODNL1 (7.89); LRRC3 (2.63); FLRT1 (7.89); FLRT3 (2.63); SOCS1 (5.26); NYX (7.89); DCN (2.63); PTPN2 (5.26); FLRT2 (10.53); ADIPOR1 (5.26); BCL3 (7.89); SOCS5 (7.89); PPP2CA (5.26); ASPN (13.16); PPP2R1A (10.53); VHL (5.26); CHAD (5.26)	SOCS3 (3.12); LRRC15 (3.12); NF2 (56.25); PIBF1 (3.12); FLRT2 (3.12); BCL3 (3.12); PPP2R1A (3.12); CHAD (3.12)
GO:0046426	negative regulation of JAK-STAT cascade	-5.210	0.013	60.53	62.5	ADIPOR1 (5.26); ASPN (13.16); BCL3 (7.89); CHAD (5.26); DCN (2.63); FLRT1 (7.89); FLRT2 (10.53); FLRT3 (2.63); HMGA2 (7.89); LRRC15 (5.26); LRRC3 (2.63); LRRC3C (5.26); LRRC4 (2.63); LRRC4B (10.53); LRRTM1 (7.89); NF2 (52.63); NYX (7.89); PODNL1 (7.89); RTN4RL2 (7.89); SOCS1 (5.26); SOCS2 (2.63); SOCS3 (10.53); SOCS4 (10.53); SOCS5 (7.89); VHL (5.26)	BCL3 (3.12); CHAD (3.12); FLRT2 (3.12); LRRC15 (3.12); NF2 (56.25); SOCS3 (3.12)
RESPONSE TO REACTIVE OXYGEN SPECIES							
GO:0072593	reactive oxygen species metabolic process	-5.626	0.013	60.53	18.75	ALOX12 (7.89); AOX1 (7.89); BCL2 (5.26); CTGF (18.42); CYBA (7.89); CYR61 (5.26); DDIT4 (5.26); EPHX2 (50); GLS2 (2.63); IL19 (5.26); NDUFA13 (7.89); P2RX7 (2.63); PDGFB (13.16); PDK4 (2.63); PLA2R1 (10.53); PMAIP1 (5.26); PREX1 (2.63); RFK (10.53); SOD1 (2.63)	CTGF (3.12); EPHX2 (15.62)
GO:1901031	regulation of response to reactive oxygen species	-8.661	0.028	55.26	37.5	BMP7 (7.89); RGN (7.89); HSPH1 (2.63); TNF (10.53); PARK7 (10.53); MET (18.42); SESN3 (7.89); GPR37 (2.63); HGF (2.63); SZT2 (10.53); PAWR (7.89); ENDOG (10.53); PSAP (5.26); FOXO3 (31.58); FBLN5 (10.53); NFE2L2 (10.53); GCH1 (10.53); STK26 (7.89); GPR37L1 (5.26)	PARK7 (3.12); MET (28.12); SZT2 (3.12); FOXO3 (15.62)

[*] The coefficients reflect ORs<1; which indicates association of altered pathway state to A-AML, since the aneuploid state was set to 0 and euploid state was set to 1 in the model.
[†] Cut-off was set at 0.05.
[‡] Percentages of patients who had at least one gene altered in a certain pathway.
[§] Genes belonging to each pathway, with percentage of A-AML and E-AML patients having a CNV in that gene.

Table S5. **Chromosome regions enriched for CNAs in the aneuploid cohort.**

Cytoband	*p* value (adj)	Genes
gain(6p12.1)	0.003	ICK;FBXO9;HCRTR2;KLHL31;BMP5;GCM1
gain(6p12.2)	0.003	EFHC1;TMEM14A;GSTA3;GSTA4;TRAM2
gain(6p12.3)	<0.001	CRISP3;C6orf141;CRISP1;OPN5;ADGRF4;ADGRF5;TDRD6;ADGRF2;DEFB110;DEFB112;DEFB113;DEFB114;ADGRF1;TNFRSF21;GLYATL3;DEFB133;CYP39A1;PGK2;RHAG;CRISP2;TFAP2D;SLC25A27
gain(6p21.1)	<0.001	TOMM6;DNPH1;CNPY3;FRS3;SLC22A7;APOBEC2;CAPN11;TAF8;PTCRA;SPATS1;C6orf223;RSPH9;TSPO2;UNC5CL;ENPP4;CUL9;GLTSCR1L;ZNF318;USP49;GNMT;RPL7L1;GUCA1A;GUCA1B;PRICKLE4;HSP90AB1;CRIP3;MDFI;C6orf226;NFKBIE;PEX6;TREM1;POLH;GTPBP2;MRPS18A;MRPS10;PPP2R5D;TMEM63B;TRERF1;AARS2;PTK7;ENPP5;PRPH2;C6orf132;MRPL14;DLK2;TBCC;BYSL;TFEB;TTBK1;CCND3;ABCC10;TJAP1;NCR2;MED20;MAD2L1BP
gain(6p21.2)	<0.001	DAAM2;MOCS1;SAYSD1;CPNE5;TDRG1;KCNK16;KCNK5;KCNK17
gain(6p21.31)	<0.001	RPS10-NUDT3;NUDT3;IP6K3;C6orf1;LEMD2;SCUBE3;HMGA1;MLN;GGNBP1;DEF6;MAPK13;BAK1;TCP11;UQCC2
gain(6p21.32)	<0.001	PFDN6;C6orf10;COL11A2;DAXX;ZBTB9;HLA-DMA;HLA-DMB;HLA-DOA;HLA-DOB;HLA-DPA1;HLA-DPB1;HLA-DPB2;HLA-DQA1;HLA-DQA2;HLA-DQB1;HLA-DQB2;HLA-DRA;HLA-DRB1;HLA-DRB5;HLA-DRB6;KIFC1;HCG23;HCG25;CUTA;PHF1;PSMB8;PSMB9;RGL2;RING1;RPS18;RXRB;VPS52;TAP1;TAP2;SLC39A7;HSD17B8;B3GALT4;SYNGAP1;WDR46;ZBTB22
gain(6p21.33)	<0.001	EHMT2;CLIC1;CSNK2B;DDAH2;LY6G6F;HLA-B;HLA-C;HSPA1B;HSPA1L;SAPCD1;MSH5;NEU1;C6orf48;POU5F1;APOM;LSM2;C6orf47;LY6G5B;LY6G6D;VARS;PRRC2A;BAG6;GPANK1;ABHD16A;SLC44A4;VWA7;C6orf25;LY6G6C;LY6G5C
gain(6p22.1)	<0.001	TRIM10;ZBED9;MAS1L;TRIM40;TRIM39-RPP21;ZSCAN23;HLA-E;HLA-F;HLA-L;ZKSCAN4;HCG17;TRIM27;ZSCAN31;ZNF192P1;TRIM26;ZKSCAN8;ZSCAN9;ZKSCAN3;PGBD1;TRIM15;ZSCAN12
gain(6p22.2)	<0.001	SLC17A4;SLC17A2;TRIM38;SCGN;SLC17A3;HIST1H2AA;HIST1H2BA;HIST1H1C;HIST1H1D;HIST1H2AE;HIST1H2BB;HIST1H1A;SLC17A1;HIST1H2AB;HIST1H2BG;HIST1H2BH;HIST1H2BI;HIST1H3A;HIST1H3C;HIST1H3E;HIST1H3G;HIST1H3B;HIST1H4A;HIST1H4F;HIST1H4H;HIST1H4B;HIST1H4G;HIST1H3F
gain(6p22.3)	<0.001	C6orf229;NRSN1;MBOAT1;E2F3;RBM24;GMPR;MYLIP;KAAG1;JARID2;STMND1;ACOT13;PRL;MRS2;SOX4;DEK;KIAA0319
gain(6p23)	<0.001	RANBP9;RNF182;SIRT5;NOL7;MCUR1;CD83
gain(6p24.2)	<0.001	TMEM170B;SMIM13;SYCP2L;ERVFRD-1;ELOVL2;GCM2
gain(6p25.2)	<0.001	ECI2;FAM217A;FAM50B;TUBB2B;C6orf201;SERPINB6;SERPINB9;WRNIP1;SLC22A23;PRPF4B
gain(6q13)	<0.001	KHDC1L;MB21D1;SDHAF4;KHDC3L;DPPA5;OOEP;DDX43;LMBRD1;OGFRL1;KHDC1
gain(6q14.1)	<0.001	RWDD2A;COX7A2;IRAK1BP1;DOPEY1;SENP6;HTR1B;IMPG1;PGM3;FAM46A;ELOVL4;TPBG;TTK;SH3BGRL2;UBE3D
gain(6q14.3)	0.004	SYNCRIP;CEP162;NT5E;TBX18
gain(6q15)	<0.001	PNRC1;PM20D2;SRSF12;CFAP206;ORC3;GABRR2;GJB7;UBE2J1;RARS2;SMIM8;LYRM2;RRAGD;SPACA1;RNGTT
gain(6q16.1)	<0.001	KLHL32;UFL1;MMS22L;NDUFAF4;GPR63;FHL5

gain(8p11.21)	<0.001	AP3M2;CHRNB3;SMIM19;GPAT4;NKX6-3;IDO2;DKK4;ANK1;IKBKB;IDO1;GOLGA7;PLAT;POLB;THAP1;C8orf4;SLC20A2;VDAC3;KAT6A;RNF170;GINS4;HOOK3;CHRNA6
gain(8p11.22)	0.004	HTRA4;ADAM2;PLEKHA2;TM2D2
gain(8p11.23)	<0.001	ERLIN2;PROSC;GOT1L1;LETM2;FGFR1;DDHD2;ADGRA2;LSM1;WHSC1L1;BRF2;STAR;RAB11FIP1;PLPP5;ASH2L;BAG4
gain(8p12)	<0.001	DUSP4;PURG;NRG1;SARAF;PPP2CB;WRN;DUSP26;RNF122;UBXN8;TTI2;MAK16;FUT10
gain(8p21.1)	0.009	EXTL3;NUGGC;ELP3;INTS9;FZD3
gain(8p21.2)	<0.001	PNMA2;CHRNA2;NKX2-6;EPHX2;PTK2B;TRIM35;GNRH1;NEFM;NKX3-1;ADAM7
gain(8p21.3)	<0.001	SORBS3;NPM2;R3HCC1;DMTN;C8orf58;CSGALNACT1;CCAR2;PDLIM2;TNFRSF10B;FGF17;DOK2;CHMP7
gain(8q11.21)	<0.001	CEBPD;MCM4;C8orf22;SNAI2;EFCAB1
gain(8q11.23)	0.002	OPRK1;ATP6V1H;RP1;SOX17;TCEA1
gain(8q12.1)	<0.001	SBF1P1;CYP7A1;SDR16C5;SDR16C6P;PENK;PLAG1;IMPAD1;RAB2A;CHCHD7;FAM110B;TOX
gain(8q12.3)	0.015	BHLHE22;UG0898H09;TTPA;GGH
gain(8q13.1)	<0.001	TCF24;C8orf44-SGK3;COPS5;ADHFE1;CRH;MCMDC2;RRS1;C8orf46;PPP1R42;MYBL1;PDE7A;C8orf44;CSPP1;VCPIP1;TRIM55;DNAJC5B
gain(8q13.3)	<0.001	NCOA2;TRAM1;XKR9;LACTB2;MSC
gain(8q21.11)	<0.001	C8orf89;RDH10;LY96;GDAP1;UBE2W;RPL7;ELOC
gain(8q21.13)	<0.001	STMN2;FABP5;HEY1;ZC2HC1A;PKIA;PEX2;ZBTB10;ZFHX4
gain(8q21.3)	<0.001	C8orf88;WWP1;TMEM64;NBN;RMDN1;NECAB1;OSGIN2;RUNX1T1;RIPK2;CPNE3
gain(8q22.1)	<0.001	FSBP;FAM92A;NDUFAF6;C8orf37;RAD54B;KIAA1429;GEM;DPY19L4;RBM12B;GDF6;MTERF3;PDP1;ESRP1;INTS8;UQCRB;PLEKHF2;TSPYL5;TMEM67;CCNE2;TP53INP1;PTDSS1
gain(8q22.2)	<0.001	OSR2;COX6C;ERICH5;RNF19A;FBXO43;KCNS2;POLR2K;RPL30;SPAG1
gain(8q22.3)	<0.001	CTHRC1;DPYS;DCAF13;RRM2B;UBR5;AZIN1;KLF10;SLC25A32;DCSTAMP;FZD6
gain(8q23.1)	<0.001	ABRA;RSPO2;EIF3E;ENY2;TRHR;NUDCD1;PKHD1L1;EMC2
gain(8q24.11)	0.002	EXT1;RAD21;EIF3H;MED30
gain(8q24.13)	<0.001	ZHX1-C8orf76;TRIB1;RNF139;ZHX1;ZNF572;ZHX2;NSMCE2;ATAD2;HAS2;NDUFB9;WDYHV1;SQLE;DERL1;TATDN1;C8orf76;FAM83A;TBC1D31;WASHC5
gain(8q24.21)	<0.001	CASC11;CCAT2;PRNCR1;PCAT2;CCDC26;FAM84B;MYC;POU5F1B;GSDMC;TMEM75;CASC8
gain(8q24.3)	<0.001	CCDC166;MINCR;TOP1MT;CYP11B2;PUF60;ARC;SCRIB;GLI4;ZNF707;ZFP41;COMMD5;ZNF517;WDR97;LY6E;HGH1;ADGRB1;ZNF250;RPL8;SCX;ZNF7;ZNF34;NAPRT;RECQL4;ZNF623
gain(9p13.3)	<0.001	GNE;CREB3;CLTA;ATP8B5P;SPAG8;FAM221B;NPR2;TMEM8B;GBA2;HRCT1;MSMP;TLN1;TPM2;FAM166B;CA9;RECK;HINT2;ARHGEF39;CCIN;CD72;RGP1;RUSC2
gain(12p13.32)	0.005	GALNT8;KCNA5;KCNA6;TIGAR;CCND2
gain(13q33.1)	0.032	ERCC5;CCDC168;KDELC1;TEX30
gain(20q11.21)	0.022	ABALON;TPX2;FOXS1;POFUT1;PLAGL2;MYLK2;TM9SF4

gain(21q21.3)	<0.001	RWDD2B;USP16;CCT8;ADAMTS5;GABPA;LTN1;APP;ATP5J;MRPL39;BACH1;JAM2
gain(21q22.11)	<0.001	OLIG2;OLIG1;TCP10L;GART;DONSON;IFNAR2;SMIM11A;DNAJC28;C21orf62;C21orf59;SON;SYNJ1;PAXBP1;CRYZL1
gain(21q22.12)	0.007	MORC3;SETD4;CLIC6;CHAF1B;RUNX1
gain(21q22.13)	<0.001	DYRK1A;DSCR10;KCNJ6;KCNJ15;PIGP;SIM2
gain(21q22.2)	<0.001	LCA5L;ERG;ETS2;HMGN1;PCP4;BRWD1;WRB;PSMG1
gain(21q22.3)	<0.001	TCONS_00029157;FRGCA;CRYAA;CSTB;RRP1B;AATBC;MX1;FAM3B;RIPK4;AGPAT3;PWP2;TMPRSS3;SUMO3;TFF1;TFF2;TFF3;TMPRSS2;SSR4P1;U2AF1;UBE2G2;C21orf33;FAM207A;RRP1
gain(22q12.1)	<0.001	CHEK2;CRYBB1;PITPNB;TFIP11;C22orf31;SRRD;MIAT;ASPHD2;XBP1;KREMEN1;ZNRF3;TPST2;HPS4
gain(22q12.2)	<0.001	PIK3IP1;RNF215;EWSR1;MORC2;SEC14L2;PATZ1;PISD;PRR14L;SEC14L3;INPP5J;SEC14L4;LIMK2;DRG1;PLA2G3;MTFP1;CCDC157;SEC14L6;RNF185;GAL3ST1;DEPDC5;SFI1
gain(22q12.3)	<0.001	HMGXB4;TOM1;IFT27;C1QTNF6;FBXO7;HMOX1;MCM5;MYH9;TIMP3;YWHAH
loss(5q14.3)	<0.001	LUCAT1;CETN3;LYSMD3;TMEM161B;ARRDC3;RASA1;ADGRV1;CCNH
loss(5q15)	<0.001	POU5F2;RHOBTB3;ELL2;KIAA0825;RGMB;LNPEP;PCSK1;SPATA9;SLF1
loss(5q21.1)	<0.001	CHD1;PPIP5K2;GIN1;ST8SIA4;C5orf30
loss(5q22.3)	<0.001	TMED7-TICAM2;CDO1;TICAM2;TMED7;FEM1C;ATG12
loss(5q23.1)	<0.001	HNCAT21;LVRN;HSD17B4;FAM170A;LOX;FTMT
loss(5q23.2)	<0.001	MARCH3;ZNF474;MGC32805;CEP120;TEX43;C5orf63;ALDH7A1;PHAX;ZNF608;GRAMD3;SNX2;SNCAIP
loss(5q23.3)	<0.001	KIAA1024L;HINT1;ISOC1;SLC12A2;LYRM7
loss(5q31.1)	<0.001	RAD50;TH2LCRR;WSPAR;DCANP1;CSF2;SEPT8;ACSL6;AFF4;HSPA4;IL3;IL5;IL13;IRF1;NEUROG1;TIFAB;CDKL3;C5orf15;TRPC7;SLC22A4;SLC22A5;MEIKIN;UBE2B;VDAC1;PDLIM4;P4HA2;CDKN2AIPNL;CXCL14;FNIP1
loss(5q31.2)	<0.001	KIF20A;BRD8;CTNNA1;ETF1;LRRTM2;GFRA3;PKD2L2;PROB1;MZB1;PAIP2;FAM13B;FAM53C;KDM3B;SIL1;WNT8A;NME5;CDC23;MYOT;MATR3;CDC25C;SLC23A1
loss(5q31.3)	<0.001	GNPDA1;ARHGAP26;RELL2;IK;DND1;ANKHD1-EIF4EBP3;NDUFA2;PCDH1;PCDHB18P;WDR55;ANKHD1;TMCO6;PCDHAC2;VTRNA1-3;VTRNA1-2;VTRNA1-1;KCTD16;HMHB1;PURA;ARAP3;NDFIP1;YIPF5;SPRY4;SLC4A9;EIF4EBP3;FCHSD1;CD14;RNF14;KIAA0141
loss(5q32)	<0.001	TCERG1;SCGB3A2;GPR151;CSNK1A1;SH3RF2;PLAC8L1;ABLIM3;ARSI;C5orf46;PDGFRB;POU4F3;GRXCR2;SPINK1;JAKMIP2
loss(5q33.1)	<0.001	G3BP1;CTB-113P19.1;CTB-12O2.1;SLC36A2;SLC36A1;CCDC69;GPX3;ANXA6;ATOX1;SPARC;CD74
loss(5q33.2)	<0.001	FAXDC2;FAM114A2;LARP1;GEMIN5;KIF4B;MRPL22;MFAP3;GALNT10;CNOT8
loss(5q33.3)	<0.001	FAM71B;EBF1;HAVCR1;ITK;HAVCR2;TIMD4;MED7
loss(7q21.12)	<0.001	DBF4;KIAA1324L;SLC25A40;SRI;STEAP4;DMTF1
loss(7q21.2)	<0.001	AKAP9;CDK6;CYP51A1;SAMD9L;FAM133B;LRRD1;SAMD9;VPS50;GATAD1;MTERF1;RBM48;KRIT1
loss(7q21.3)	<0.001	SLC25A13;BET1;DLX5;DLX6;PEG10;GNG11;PDK4;ASB4;SDHAF3;CASD1;SHFM1;TFPI2;SGCE
loss(7q22.1)	<0.001	ATP5J2-PTCD1;CYP3A7-CYP3A51P;STAG3L5P-PVRIG2P-PILRB;CYP3A4;CYP3A5;ZSCAN25;FAM200A;TMEM130;GATS;NPTX2;SMURF1;ZNF655;PVRIG;TRRAP;TRIM4

loss(7q22.3)	0.008	HBP1;GPR22;PIK3CG;CBLL1
loss(7q31.1)	<0.001	HRAT17;THAP5;LSMEM1;IFRD1;DNAJB9;PNPLA8;GPR85;TMEM168
loss(7q31.2)	<0.001	ASZ1;TES;ST7-OT4;MET;WNT2;CAPZA2;CAV1;CAV2;ST7-OT3
loss(7q31.31)	0.014	LVCAT5;TSPAN12;LSM8
loss(7q32.1)	<0.001	SMKR1;FLNC;SND1;TPI1P2;FSCN3;TSPAN33;FAM71F2;ARF5;LEP;PAX4;STRIP2;LRRC4;SMO;GCC1;FAM71F1
loss(7q32.2)	<0.001	CPA2;KLF14;SSMEM1;KLHDC10;CPA4;TMEM209
loss(7q33)	<0.001	SLC13A4;WDR91;FAM180A;C7orf73;CREB3L2;C7orf49;SLC35B4;TRIM24
loss(7q34)	<0.001	C7orf55-LUC7L2;PRSS37;CLEC2L;FMC1;TRY2P;UBN2;ATP6V0A4;LUC7L2;KIAA1147;KIAA1549;MGAM2;FAM131B
loss(11p13)	0.023	FJX1;PAMR1;TRIM44;SLC1A2;WT1
loss(12p12.3)	<0.001	STRAP;HIST4H4;ERP27;C12orf60;ARHGDIB;ART4;MGP;MGST1;SMCO3;PDE6H;WBP11;H2AFJ;LMO3;RERG
loss(12p13.1)	<0.001	CDKN1B;CREBL2;EMP1;GPR19;RPL13AP20;DDX47;FAM234B;PLBD1;GSG1;GPRC5A
loss(12p13.2)	<0.001	SMIM10L1;KLRC4-KLRK1;BORCS5;TMEM52B;ETV6;GABARAPL1;CLEC9A;KLRD1;CLEC12B;OLR1;LOH12CR2;TAS2R9;TAS2R8;TAS2R7;CLEC7A;YBX3
loss(17p13.1)	<0.001	RPL29P2;FBXO39;EFNB3;SLC13A5;GUCY2D;C17orf100;TNFSF12-TNFSF13;RNASEK;MED31;XAF1;TP53;TEKT1;NAA38;TXNDC17;KIAA0753
loss(17p13.2)	<0.001	ZFP3;NLRP1;AIPL1;ALOX15;SMTNL2;SCIMP;FAM64A;NCBP3;DHX33;C1QBP;ZNF232;RPAIN;ZNF594;USP6;RABEP1

Table S6. Chromosome regions enriched for CNAs in AML-related genes in the aneuploid cohort.

Cytoband	A-AML (% of cases)	E-AML (% of cases)	Candidate genes	*p value*
gain (6p22.3)	15.8	0	*E2F3, SOX4, DEK*	*<0.05*
loss (5q31.1)	36.8	6.3	*RAD50, TIFAB, PDLIM4*	*<0.05*
loss (5q31.2)	36.8	6.3	*KIF20A, CTNNA1, KDM3B*	*<0.01*
loss (5q31.3)	36.8	6.3	*DND1, PURA, SPRY4*	*<0.05*
loss (12p13.1)	18.4	0	*CDKN1B, EMP1*	*<0.05*

Table S7. Top 5 disease-related pathways ranked according to their degree and betweenness centrality in the A-AML and E-AML networks.

Pathway ID	Pathway name	Degree	Betweenness centrality
Aneuploid AML			
GO:0034723	DNA replication-dependent nucleosome organization	124	0.136
GO:0006335	DNA replication-dependent nucleosome assembly	123	0.136
GO:1903706	regulation of hemopoiesis	113	0.050
GO:0030099	myeloid cell differentiation	107	0.027
GO:0002521	leucocyte differentiation	107	0.027
Euploid AML			
GO:1903706	regulation of hemopoiesis	34	0.029
GO:0045879	negative regulation of smoothened signaling pathway	26	0.141
GO:0008589	regulation of smoothened signaling pathway	23	0.031
GO:0001953	negative regulation of cell-matrix adhesion	18	0.077
GO:0030099	myeloid cell differentiation	17	0.036

Table S8. Pathway enrichment analysis for genes upregulated in A-AML.

Source	Term	Fold Enrichment	*p* value	Genes
GO-BP	GO:0006334~nucleosome assembly	24.5	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0031497~chromatin assembly	23.7	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0065004~protein-DNA complex assembly	22.6	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0034728~nucleosome organization	22.1	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0006323~DNA packaging	17.6	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0006333~chromatin assembly or disassembly	16.2	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0016043~cellular component organization	2.5	<0.001	HIST1H2AB, LDLR, HIST1H2AE, ITSN1, APP, HIST1H2BM, HIST1H2BI, TGFBI, RUNX3, EHD4, PRKCA, H1F0, HIST1H3J, HIST1H2BB, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, CCNF, NID1, CDC20, MCM2, UBE2C, PLK1, HIST1H3A, HIST1H3B, TUBA4A, SETD7, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0034622~cellular macromolecular complex assembly	7.4	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, TUBA4A, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0006325~chromatin organization	6.2	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, SETD7, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I

GO-BP	GO:0051276~chromosome organization	4.9	0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, SETD7, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
REACTOME	REACT_7970:Telomere Maintenance	17.8	<0.001	HIST1H2AB, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0034621~cellular macromolecular complex subunit organization	6.6	<0.001	HIST1H2AB, H1F0, HIST1H2BB, HIST1H3J, HIST1H2BC, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2AE, MCM2, HIST1H2BM, HIST1H2BI, HIST1H3A, HIST1H3B, TUBA4A, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I
GO-BP	GO:0032268~regulation of cellular protein metabolic process	4.3	0.005	PRKCA, APP, PLK1, MKNK2, PAX5, CDC20, UBE2C
GO-BP	GO:0031400~negative regulation of protein modification process	9.9	0.007	PRKCA, PAX5, CDC20, UBE2C
GO-BP	GO:0051246~regulation of protein metabolic process	3.8	0.009	PRKCA, APP, PLK1, MKNK2, PAX5, CDC20, UBE2C
GO-BP	GO:0031399~regulation of protein modification process	5.0	0.016	PRKCA, PLK1, PAX5, CDC20, UBE2C
GO-BP	GO:0051437~positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	13.0	0.022	PLK1, CDC20, UBE2C
GO-BP	GO:0032269~negative regulation of cellular protein metabolic process	6.5	0.022	PRKCA, PAX5, CDC20, UBE2C
GO-BP	GO:0051443~positive regulation of ubiquitin-protein ligase activity	12.6	0.023	PLK1, CDC20, UBE2C
GO-BP	GO:0051439~regulation of ubiquitin-protein ligase activity during mitotic cell cycle	12.4	0.023	PLK1, CDC20, UBE2C
GO-BP	GO:0051248~negative regulation of protein metabolic process	6.3	0.024	PRKCA, PAX5, CDC20, UBE2C
GO-BP	GO:0031401~positive regulation of protein modification process	6.3	0.024	PRKCA, PLK1, CDC20, UBE2C
GO-BP	GO:0051351~positive regulation of ligase activity	12.1	0.025	PLK1, CDC20, UBE2C
GO-BP	GO:0051438~regulation of ubiquitin-protein ligase activity	11.3	0.028	PLK1, CDC20, UBE2C
GO-BP	GO:0051340~regulation of ligase activity	10.9	0.030	PLK1, CDC20, UBE2C
GO-BP	GO:0031398~positive regulation of protein ubiquitination	10.5	0.032	PLK1, CDC20, UBE2C
GO-BP	GO:0031396~regulation of protein ubiquitination	8.8	0.044	PLK1, CDC20, UBE2C
GO-BP	GO:0051247~positive regulation of protein metabolic process	4.8	0.047	PRKCA, PLK1, CDC20, UBE2C
GO-BP	GO:0032270~positive regulation of cellular protein metabolic process	5.0	0.042	PRKCA, PLK1, CDC20, UBE2C

GO-BP	GO:0031399~regulation of protein modification process	5.0	0.016	PRKCA, PLK1, PAX5, CDC20, UBE2C
GO-BP	GO:0000278~mitotic cell cycle	4.8	0.007	APP, CDKN2C, PLK1, CCNF, CDC20, UBE2C
GO-BP	GO:0022403~cell cycle phase	4.3	0.012	APP, CDKN2C, PLK1, CCNF, CDC20, UBE2C
GO-BP	GO:0007049~cell cycle	3.0	0.014	APP, UHRF1, CDKN2C, PLK1, CCNF, CDC20, MCM2, UBE2C
KEGG	hsa04110:Cell cycle	6.3	0.022	CDKN2C, PLK1, CDC20, MCM2
GO-BP	GO:0000280~nuclear division	5.3	0.037	PLK1, CCNF, CDC20, UBE2C
GO-BP	GO:0007067~mitosis	5.3	0.037	PLK1, CCNF, CDC20, UBE2C
GO-BP	GO:0000087~M phase of mitotic cell cycle	5.3	0.038	PLK1, CCNF, CDC20, UBE2C
GO-BP	GO:0022402~cell cycle process	3.1	0.039	APP, CDKN2C, PLK1, CCNF, CDC20, UBE2C
GO-BP	GO:0048285~organelle fission	5.1	0.041	PLK1, CCNF, CDC20, UBE2C
REACTOME	REACT_152:Cell Cycle, Mitotic	3.3	0.048	PLK1, TUBA4A, CDC20, MCM2, UBE2C
GO-BP	GO:0030198~extracellular matrix organization	8.5	0.047	APP, TGFBI, NID1
GO-BP	GO:0010324~membrane invagination	5.3	0.037	APP, LDLR, ITSN1, EHD4
GO-BP	GO:0006897~endocytosis	5.3	0.037	APP, LDLR, ITSN1, EHD4
GO-BP	GO:0006917~induction of apoptosis	4.6	0.022	PRKCA, APP, CDKN2C, ITSN1, RUNX3
GO-BP	GO:0012502~induction of programmed cell death	4.6	0.022	PRKCA, APP, CDKN2C, ITSN1, RUNX3

Table S9. Pathway enrichment analysis for genes downregulated in A-AML.

Category	Term	Fold Enrichment	*p* value	Genes
GO-BP	GO:0048562~embryonic organ morphogenesis	8.7	<0.001	HOXB3, HOXB4, HOXB2, HOXA3, HOXA5, HOXA6, HOXB5, HOXA7, ZEB1
GO-BP	GO:0048568~embryonic organ development	6.7	<0.001	HOXB3, HOXB4, HOXB2, HOXA3, HOXA5, HOXA6, HOXB5, HOXA7, ZEB1
GO-BP	GO:0003002~regionalization	5.9	<0.001	HOXB3, HOXB4, HOXB2, HOXA3, HOXA5, HOXA6, HOXB5, HOXA7, HOXA10
GO-BP	GO:0007389~pattern specification process	4.8	<0.001	HOXB3, HOXB4, HOXB2, HOXA3, HOXA5, HOXA6, HOXB5, HOXA7, HOXA10, ZEB1
GO-BP	GO:0048598~embryonic morphogenesis	4.2	0.001	HOXB3, HOXB4, HOXB2, HOXA3, HOXA5, HOXA6, HOXB5, HOXA7, HOXA10, ZEB1
GO-BP	GO:0009887~organ morphogenesis	2.7	0.004	HOXB3, HOXB4, HOXB2, HOXA3, BRAF, HOXA5, HOXA6, HOXB5, HOXA7, ANGPT1, ZEB1, NKX2-3
GO-BP	GO:0009790~embryonic development	2.5	0.012	HOXB3, HOXB4, HOXB2, HOXA3, HOXA5, PIK3CB, HOXA6, HOXB5, HOXA7, HOXA10, ZEB1
GO-BP	GO:0006139~nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1.6	0.001	SUPT3H, OCLN, ZKSCAN1, NFIX, ZEB1, MEIS1, NAA38, WT1, POT1, ZFC3H1, WDR36, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, LARS, HOXA10, TWISTNB, LUC7L3, HIP1, ATP8B4, NKX2-3, ESCO1, ZNF33A, GMDS, MAT2A, ZNF25, ATR, RAD50, NME7, HOXB3, ZNF138, MED31, HOXB4, HOXB2, RPAIN, DMTF1, ATP2C1, HOXB5, NFE2L3, ZNF33B
GO-BP	GO:0019219~regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1.5	0.021	SUPT3H, NFIX, ZKSCAN1, ZEB1, MEIS1, WT1, POT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, NKX2-3, HIP1, ZNF33A, RFX7, ZNF25, ATR, RAD50, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3
GO-BP	GO:0009119~ribonucleoside metabolic process	8.4	0.049	OCLN, MAT2A, NME7
GO-BP	GO:0044249~cellular biosynthetic process	1.6	0.001	SUPT3H, ZKSCAN1, NFIX, ZEB1, ALG8, MEIS1, WT1, POT1, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, LARS, HOXA10, OGT, TWISTNB, HIP1, ATP8B4, NKX2-3, ZNF33A, TBXAS1, GMDS, MAT2A, ZNF25, ATR, RAD50, NME7, HOXB3, ZNF138, MED31, HOXB4, HOXB2, RPAIN, ST8SIA6, DMTF1, ATP2C1, HOXB5, NFE2L3, ZNF33B, MRPL45
GO-BP	GO:0034645~cellular macromolecule biosynthetic process	1.6	0.002	SUPT3H, ZKSCAN1, NFIX, ZEB1, ALG8, MEIS1, WT1, POT1, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, LARS, HOXA10, OGT, TWISTNB, HIP1, NKX2-3, ZNF33A, ZNF25, ATR, RAD50, MED31, HOXB3, ZNF138, HOXB4, HOXB2, RPAIN, ST8SIA6, DMTF1, HOXB5, MRPL45, ZNF33B, NFE2L3

GO-BP	GO:0009059~macromolecule biosynthetic process	1.6	0.002	SUPT3H, ZKSCAN1, NFIX, ZEB1, ALG8, MEIS1, WT1, POT1, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, LARS, HOXA10, OGT, TWISTNB, HIP1, NKX2-3, ZNF33A, ZNF25, ATR, RAD50, MED31, HOXB3, ZNF138, HOXB4, HOXB2, RPAIN, ST8SIA6, DMTF1, HOXB5, MRPL45, ZNF33B, NFE2L3
GO-BP	GO:0009058~biosynthetic process	1.5	0.002	SUPT3H, ZKSCAN1, NFIX, ZEB1, ALG8, MEIS1, WT1, POT1, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, LARS, HOXA10, OGT, TWISTNB, HIP1, ATP8B4, NKX2-3, ZNF33A, TBXAS1, GMDS, MAT2A, ZNF25, ATR, RAD50, NME7, HOXB3, ZNF138, MED31, HOXB4, HOXB2, RPAIN, ST8SIA6, DMTF1, ATP2C1, HOXB5, NFE2L3, ZNF33B, MRPL45
GO-BP	GO:0043170~macromolecule metabolic process	1.3	0.005	DNAJC10, ZKSCAN1, ZEB1, ALG8, WDR36, RRN3, P4HA1, PIK3C3, CAT, OGT, TWISTNB, LUC7L3, ESCO1, ZNF33A, BRAF, TRPM7, PIK3CB, RPS6KC1, RAD50, TBCK, ZNF138, PROK2, RPAIN, MAPK6, NFE2L3, ZNF33B, MRPL45, PMPCB, SUPT3H, ERMP1, CLU, NFIX, MEIS1, NAA38, WT1, POT1, ZFC3H1, ZNF600, HOXA3, HOXA5, UFM1, HOXA6, HOXA7, LARS, HOXA10, LMLN, HIP1, NKX2-3, ZNF25, ATR, MSRB3, HOXB3, MED31, HOXB4, HOXB2, ST8SIA6, DMTF1, HOXB5, UBA3
GO-BP	GO:0044260~cellular macromolecule metabolic process	1.4	0.005	DNAJC10, ZKSCAN1, ZEB1, ALG8, WDR36, RRN3, P4HA1, PIK3C3, OGT, CAT, TWISTNB, LUC7L3, ESCO1, ZNF33A, BRAF, TRPM7, PIK3CB, RPS6KC1, RAD50, TBCK, ZNF138, PROK2, RPAIN, MAPK6, NFE2L3, ZNF33B, MRPL45, SUPT3H, NFIX, MEIS1, NAA38, WT1, POT1, ZFC3H1, HOXA3, ZNF600, HOXA5, UFM1, HOXA6, HOXA7, HOXA10, LARS, HIP1, NKX2-3, ZNF25, ATR, MSRB3, HOXB3, MED31, HOXB4, HOXB2, ST8SIA6, DMTF1, HOXB5, UBA3
GO-BP	GO:0044238~primary metabolic process	1.3	0.006	OCLN, DNAJC10, ZKSCAN1, ZEB1, ALG8, WDR36, RRN3, P4HA1, PIK3C3, CAT, OGT, TWISTNB, LUC7L3, ATP8B4, ESCO1, ZNF33A, BRAF, TRPM7, PIK3CB, RPS6KC1, RAD50, NME7, TBCK, ZNF138, PROK2, RPAIN, MAPK6, ATP2C1, NFE2L3, ZNF33B, MRPL45, PMPCB, SUPT3H, ERMP1, CLU, NFIX, MEIS1, NAA38, POT1, WT1, ZFC3H1, ZNF600, HOXA3, HOXA5, UFM1, HOXA6, HOXA7, LARS, HOXA10, HSD17B4, LMLN, NKX2-3, HIP1, TBXAS1, GMDS, PLEK, MAT2A, ZNF25, ATR, MSRB3, HOXB3, MED31, HOXB4, HOXB2, ST8SIA6, DMTF1, HOXB5, UBA3
GO-BP	GO:0044237~cellular metabolic process	1.2	0.015	OCLN, DNAJC10, ZKSCAN1, ZEB1, ALG8, WDR36, RRN3, P4HA1, PIK3C3, CAT, OGT, TWISTNB, LUC7L3, ATP8B4, ESCO1, ZNF33A, BRAF, TRPM7, PIK3CB, RPS6KC1, RAD50, NME7, TBCK, ZNF138, PROK2, RPAIN, MAPK6, ATP2C1, NFE2L3, ZNF33B, MRPL45, SUPT3H, NFIX, MEIS1, NAA38, POT1, WT1, ZFC3H1, ZNF600, HOXA3, HOXA5, UFM1, HOXA6, HOXA7, LARS, HOXA10, HSD17B4, NKX2-3, HIP1, TBXAS1, GMDS, PLEK, MAT2A, ZNF25, ATR, MSRB3, HOXB3, MED31, HOXB4, HOXB2, ST8SIA6, DMTF1, HOXB5, UBA3

GO-BP	GO:0010556~regulation of macromolecule biosynthetic process	1.4	0.037	SUPT3H, NFIX, ZKSCAN1, ZEB1, MEIS1, WT1, POT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, NKX2-3, HIP1, ZNF33A, RFX7, ZNF25, ATR, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3
GO-BP	GO:0031326~regulation of cellular biosynthetic process	1.4	0.037	SUPT3H, ZKSCAN1, NFIX, ZEB1, MEIS1, WT1, POT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, NKX2-3, HIP1, ZNF33A, PLEK, RFX7, ZNF25, ATR, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3
GO-BP	GO:0009889~regulation of biosynthetic process	1.4	0.040	SUPT3H, ZKSCAN1, NFIX, ZEB1, MEIS1, WT1, POT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, NKX2-3, HIP1, ZNF33A, PLEK, RFX7, ZNF25, ATR, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3
GO-BP	GO:0031323~regulation of cellular metabolic process	1.3	0.045	SUPT3H, ZKSCAN1, NFIX, ZEB1, MEIS1, WT1, POT1, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, PDGFD, HIP1, NKX2-3, ZNF33A, PLEK, PIK3CB, RFX7, ZNF25, ATR, RAD50, MED31, HOXB3, ZNF138, PROK2, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3
GO-BP	GO:0050789~regulation of biological process	1.2	0.048	HINT1, DNAJC10, ZKSCAN1, ZEB1, CD47, AGAP6, RRN3, PIK3C3, ANGPT1, CAT, PDGFD, OGT, RAB27B, AGAP4, ZNF33A, BRAF, PIK3CB, G3BP1, RINT1, RPS6KC1, IPO8, RAD50, TBCK, ZNF138, PROK2, MAPK6, ATP2C1, KRIT1, NFE2L3, ZNF33B, SUPT3H, RALGPS2, CLU, NFIX, ABHD2, MEIS1, FAM13B, POT1, WT1, ZNF600, HOXA3, RASGRP3, HOXA5, HOXA6, SOS1, HOXA7, HOXA10, EXOC4, RHOBTB1, SUCNR1, NKX2-3, HIP1, PLEK, RFX7, ZNF25, ATR, TAX1BP1, ITPR2, HOXB3, MED31, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, RAP1B
GO-BP	GO:0006807~nitrogen compound metabolic process	1.5	0.003	SUPT3H, OCLN, ZKSCAN1, NFIX, ZEB1, MEIS1, NAA38, WT1, POT1, ZFC3H1, WDR36, ZNF600, HOXA3, RRN3, HOXA5, P4HA1, HOXA6, HOXA7, LARS, HOXA10, TWISTNB, LUC7L3, HIP1, ATP8B4, NKX2-3, ESCO1, ZNF33A, GMDS, MAT2A, ZNF25, ATR, RAD50, NME7, HOXB3, ZNF138, MED31, HOXB4, HOXB2, RPAIN, DMTF1, ATP2C1, HOXB5, NFE2L3, ZNF33B
GO-BP	GO:0034641~cellular nitrogen compound metabolic process	1.5	0.003	SUPT3H, OCLN, ZKSCAN1, NFIX, ZEB1, MEIS1, NAA38, WT1, POT1, ZFC3H1, WDR36, ZNF600, HOXA3, RRN3, HOXA5, HOXA6, HOXA7, LARS, HOXA10, TWISTNB, LUC7L3, HIP1, ATP8B4, NKX2-3, ESCO1, ZNF33A, GMDS, MAT2A, ZNF25, ATR, RAD50, NME7, HOXB3, ZNF138, MED31, HOXB4, HOXB2, RPAIN, DMTF1, ATP2C1, HOXB5, NFE2L3, ZNF33B
GO-BP	GO:0051171~regulation of nitrogen compound metabolic process	1.4	0.023	SUPT3H, NFIX, ZKSCAN1, ZEB1, MEIS1, WT1, POT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, NKX2-3, HIP1, ZNF33A, RFX7, ZNF25, ATR, RAD50, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3

GO-BP	GO:0046486~glycerolipid metabolic process	4.0	0.037	PLEK, PIK3CB, PIK3C3, CAT, NKX2-3
GO-BP	GO:0006355~regulation of transcription, DNA-dependent	1.7	0.016	SUPT3H, ZNF33A, RFX7, ZNF25, ZKSCAN1, NFIX, ZEB1, MEIS1, WT1, HOXB3, ZNF138, HOXB4, HOXB2, HOXA3, HOXA5, DMTF1, HOXA6, HOXB5, HOXA7, HOXA10, ZNF33B, NFE2L3, NKX2-3
GO-BP	GO:0006350~transcription	1.6	0.009	SUPT3H, NFIX, ZKSCAN1, ZEB1, MEIS1, WT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, TWISTNB, NKX2-3, HIP1, ZNF33A, ZNF25, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, ZNF33B, NFE2L3
GO-BP	GO:0045449~regulation of transcription	1.4	0.037	SUPT3H, NFIX, ZKSCAN1, ZEB1, MEIS1, WT1, HOXA3, ZNF600, RRN3, HOXA5, HOXA6, HOXA7, HOXA10, CAT, NKX2-3, HIP1, ZNF33A, RFX7, ZNF25, MED31, HOXB3, ZNF138, HOXB4, HOXB2, DMTF1, HOXB5, UBA3, ZNF33B, NFE2L3
GO-BP	GO:0051252~regulation of RNA metabolic process	1.6	0.020	SUPT3H, ZNF33A, RFX7, ZNF25, ZKSCAN1, NFIX, ZEB1, MEIS1, WT1, HOXB3, ZNF138, HOXB4, HOXB2, HOXA3, HOXA5, DMTF1, HOXA6, HOXB5, HOXA7, HOXA10, ZNF33B, NFE2L3, NKX2-3
GO-BP	GO:0007264~small GTPase mediated signal transduction	3.8	0.002	RALGPS2, RASGRP3, BRAF, SOS1, KRIT1, G3BP1, RHOBTB1, RAP1B, RAB27B
PANTHER	P00047:PDGF signaling pathway	4.4	0.003	MAPK6, BRAF, PIK3CB, SOS1, PIK3C3, RPS6KC1, ITPR2
KEGG	hsa04722:Neurotrophin signaling pathway	5.7	0.030	BRAF, PIK3CB, SOS1, RAP1B
PANTHER	P00034:Integrin signalling pathway	3.1	0.032	MAPK6, BRAF, PIK3CB, SOS1, PIK3C3, RAP1B
PANTHER	P04393:Ras Pathway	5.1	0.038	BRAF, PIK3CB, SOS1, PIK3C3
PANTHER	P00056:VEGF signaling pathway	5.9	0.026	MAPK6, BRAF, PIK3CB, PIK3C3
GO-BP	GO:0015031~protein transport	2.5	0.002	RPGR, PLEK, SNX14, RINT1, IPO8, WDR19, COG5, RPAIN, COG6, PEX1, NUP205, EXOC4, EXOC6, RAB27B, MRPL45
GO-BP	GO:0045184~establishment of protein localization	2.5	0.002	RPGR, PLEK, SNX14, RINT1, IPO8, WDR19, COG5, RPAIN, COG6, PEX1, NUP205, EXOC4, EXOC6, RAB27B, MRPL45
GO-BP	GO:0008104~protein localization	2.2	0.008	RPGR, PLEK, SNX14, RINT1, IPO8, WDR19, COG5, RPAIN, COG6, PEX1, NUP205, EXOC4, EXOC6, RAB27B, MRPL45
GO-BP	GO:0006886~intracellular protein transport	2.7	0.026	WDR19, RPGR, RPAIN, PEX1, NUP205, EXOC4, IPO8, MRPL45
GO-BP	GO:0034613~cellular protein localization	2.5	0.040	WDR19, RPGR, RPAIN, PEX1, NUP205, EXOC4, IPO8, MRPL45
GO-BP	GO:0006904~vesicle docking during exocytosis	16.0	0.015	PLEK, EXOC4, EXOC6
GO-BP	GO:0048278~vesicle docking	14.8	0.017	PLEK, EXOC4, EXOC6
GO-BP	GO:0022406~membrane docking	12.4	0.024	PLEK, EXOC4, EXOC6
KEGG	hsa05221:Acute myeloid leukemia	12.1	0.004	BRAF, PIK3CB, SOS1, CCNA1
KEGG	hsa04510:Focal adhesion	4.4	0.023	BRAF, PIK3CB, SOS1, RAP1B, PDGFD

Table S10. ***TP53*** **mutational status in the aneuploid and euploid GEP cohorts.**

Patient ID	Aneuploid/Euploid	*TP53*
1	Aneuploid	wt
5	Aneuploid	wt
12	Aneuploid	wt
13	Aneuploid	c.742C>T, p.R248W
21	Aneuploid	wt
23	Aneuploid	wt
24	Aneuploid	-17
25	Aneuploid	-17
37	Aneuploid	wt
54	Aneuploid	wt
55	Aneuploid	wt
56	Aneuploid	wt
58	Aneuploid	wt
68	Aneuploid	wt
69	Aneuploid	wt
70	Aneuploid	wt
71	Aneuploid	c. C577>T, p.H193Y
1006	Aneuploid	wt
2002	Aneuploid	wt
2030	Aneuploid	wt
2036	Aneuploid	wt
2040	Aneuploid	wt
6	Euploid	wt
14	Euploid	wt
15	Euploid	wt
16	Euploid	wt
18	Euploid	wt
20	Euploid	wt
0022	Euploid	wt
26	Euploid	wt
0037	Euploid	wt
39	Euploid	wt
40	Euploid	wt
41	Euploid	wt
44	Euploid	wt
45	Euploid	wt
46	Euploid	wt
47	Euploid	wt
48	Euploid	wt
49	Euploid	wt
50	Euploid	wt
51	Euploid	wt

59	Euploid	wt
64	Euploid	wt
65	Euploid	wt
66	Euploid	wt
83	Euploid	wt
85	Euploid	wt
2195	Euploid	wt

Table S11. List of WES mutations analyzed by Sanger Sequencing or Next Generation Sequencing.

Start	End	Reference base(s)	Altered base(s)	Gene	Detection Method	Validation Method	Tumour status	Germline status	Sanger Sequencing Forward Primer (5'-3')	Sanger Sequencing Reverse Primer (5'-3')
35033566	35033566	G	T	AGXT2	MuTect	SS	mut	na	TGCGTTCTTAGAAATCAGAGGTG	CAGAGCCTTGCAGTTTACTTGAT
105407657	105407657	A	G	AHNAK2	VarScan	SS	mut	wt	TGTCTTCCTCTGAAATCGAAGGA	GATTCAAAGTGAGGACCAGTGAG
46246569	46246569	G	T	ARID2	MuTect	SS	wt	na	TCCCGACTCAGGATCAAAAGTAT	ATGAGACATGGAAAACAGTGCAT
31023076	31023079	TGAT	-	ASXL1	VarScan	SS	mut	wt	AGAGGACCTGCCTTCTCTGAGAAA	TTCGATGGGATGGGTATCCAATGC
17314703	17314703	G	A	ATP13A2	MuTect	SS	mut	wt	GAGCATGGCCAGTATTGAGTG	ACGTCATCTATTCTGGGACCTG
25022705	25022705	C	T	CENPO	MuTect	SS	mut	na	AATTGTTAGCTCCCCTGGTTTTA	CCCAAATCACGTTTGTTCATTAT
57743938	57743938	T	G	CLTC	Both	SS	mut	na	ATGCGCCTCAAGTATGTGTTTTA	GGTTAAGAAGATGGGTGTGAACC
165542528	165542528	G	-	COBLL1	VarScan	SS	mut	na	TGCTCTTGTCCTTTGTGTTGATT	GGCATTAAAAGCCACAACACAAA
135487552	135487552	G	A	DDX31	Both	SS	mut	na	TAATTGTCCTGCTCATGTTTCGT	AGTGCTTGTTTATTGGGAGAAGG
6350850	6350850	G	A	FAM64A	MuTect	SS	mut	na	TGGTTTTGTCTGTGCTAACCTTT	GGATTGTGCCCAGGTTAAAAGG
15937239	15937239	G	T	FBXL7	MuTect	SS	mut	wt	AGAACTGCACCAAACTCAAATCC	TTTGCTTTGTTCAGGTTTGTGTG
28592642	28592642	C	A	FLT3	VarScan	SS	mut	na	CCGCCAGGAACGTGCTTG	GCAGACGGGCATTGCCCC
75874786	75874786	C	G	GLIPR1	MuTect	SS	mut	wt	CTCCAATTATTCACACACAGCAA	CAGAAAACAGGAAGTGTCCAAAG
51749144	51749144	G	A	GRM2	MuTect	SS	mut	wt	GGTTCCATGTTAGGGTGAATGTT	GAATTCGTCCAATCGGTACTCAT
1960012	1960012	-	ACT	HIC1	VarScan	SS	mut	na	CTCCTGCTCCTTCTCCTGGTC	CATGTCATGGTCCAGGTTGAG
176316714	176316714	T	C	HK3	Both	SS	mut	na	ATGTATCTCCTTCAAAGCCAGGG	GGTATGGTCGAAGGTGGTCAG
37524829	37524829	C	A	IL2RB	MuTect	SS	wt	na	GGGAGTGCGGGGCTATAATC	TTCCTCTGAGTAGGGGTCGT
98662167	98662167	G	A	IPO5	MuTect	SS	mut	na	TTCTGGTCTTTGTGTTTTGCCTC	ACCGATAGCTCCTTCTAAAGACA
41620056	41620056	G	T	L3MBTL2	MuTect	SS	mut	wt	AACACCACACCTTCCCTGTC	ACTCTTCAGCCCTCGAAACC
29820029	29820029	G	A	MAZ	MuTect	SS	mut	na	TGTCACTCCCATTTCCTACAGAT	GAACTGGCTTTCTTGACTACTCC
99966347	99966347	C	T	METAP1	MuTect	SS	mut	na	GAGCTTCTGTTGGGCAATAACTA	CCAAGAGATTCCCAGATCATCCT
82335748	82335748	G	T	MEX3B	MuTect	SS	mut	wt	CGGTATCTTCTTCCTGCTCTTCT	CCGCTGCCTTTAAGAAAAGAT
158819011	158819011	C	T	MNDA	VarScan	SS	mut	na	ACTAACGAGCTTTCATAGGGGAT	GCCTCGTGAATGTCATAAAAGCA
18258091	18258091	C	T	NAT2	VarScan	SS	mut	wt	ATTGACGGCAGGAATTACATTGT	ACTTCTTCAACCTCTTCCTCAGT
29562747	29562747	G	A	NF1	Both	SS	mut	na	AATGGGATTGTTTGCACTAACCT	CTAACATGTTGCCAATCAGAGGA
25457049	25457054	CTCCCA	-	NINL	VarScan	SS	mut	na	CTGTGGAGTGGATGGGGATATT	CATGTCATCCTTCTCTCTCTCCA
57080473	57080473	G	A	NLRC5	MuTect	SS	mut	wt	TCCAAGTCTGGGAGTCCAAT	ACCCACGCCTCTTTCTTTTC

120474836	120474836	A	G	PDE5A	Both	SS	mut	wt	AACTGCACAGAGGGAACTCA	TGCAAATAAGGCAAAAACTCTAGCA
81242148	81242149	TT	-	PKD1L2	VarScan	SS	mut	na	GTGGGGTCTGGAATATGGTATCT	CATGCCATAAAATCAGAGGGACA
42281236	42281236	A	C	PKDCC	VarScan	SS	mut	na	ATAGAAAATCACTGGCCTCCTCT	TAGGCATTATAGAGGTTCCGCTT
42284989	42284989	C	A	PLA2G4E	Both	SS	mut	na	GTTAGGGTTCTCAATGGCCTG	TCTAACCTAATCCCCTGTGTGC
79785404	79785404	G	C	POLR3A	MuTect	SS	mut	na	GCTTCCTTCCATCTCCTCAATTC	CACCACTCACAGTTCCTAAGTTC
106857363	106857363	T	C	POLR3B	Both	SS	mut	na	ACTCTATGGTAGGCATGAAATGA	GACAGCATAGAGGAGCAAGTCTA
58740498	58740498	C	G	PPM1D	MuTect	SS	mut	wt	TGAATGCATACCCCGTTTTT	TCTTTCGCTGTGAGGTTGTG
54153157	54153157	C	T	PSME4	VarScan	SS	mut	wt	CTTGACTTCTGTATTTGGCCCTT	TCCTTTTATAACTTCAGGAGCACC
43111336	43111336	G	T	PTK7	MuTect	SS	wt	na	CCTCATGTTCTACTGCAAGAAGC	CCAGGTTGCTCAGAAGACGAG
37628876	37628876	A	G	RAC2	Both	SS	mut	na	ACGGGTAGGAAAAGGATTAAGAGA	CTTAAGGGGAGAGGTAGGGTTTC
37640155	37640155	C	G	RAC2	MuTect	SS	mut	wt	AGCTTGAGTAAGTTCCCCTTCC	TGCATCCACAGAGTAAAGACTGA
38967300	38967300	A	C	RICTOR	MuTect	SS	mut	wt	AATTAATAGGAATGGGCCAAAAA	CCTGGCATGAAAGAATCTGTTAG
1551180	1551193	GCTGCT TCCGGA CA	-	RILP	VarScan	SS	mut	na	GAGAGGGACAGTACAAAGGGTT	GTCATCAGGCTCAGCAGAATG
122265671	122265671	A	T	SETD1B	Both	SS	mut	na	CAGTGACCTGCTCAAGTTCAAC	CACTTGGTGGCGTCGATGATG
134491555	134491555	G	A	SGK1	Both	SS	mut	na	GCTTGAAGTGGGTGATTATGGAA	GCTCCACCAAAAGGCTAACTAAA
124517319	124517319	G	A	SIAE	Both	SS	mut	na	GGAGGAGAGTAATGTGTGGTCAT	GACCAGCACATTATGAGGACAAA
96964395	96964395	G	A	SNRNP20 0	Both	SS	mut	na	TGCCTGGTTACTTTTATAGCTCG	AAGACCCCACCATATACTCACTC
101245695	101245695	G	-	SPAG1	VarScan	SS	mut	wt	GGTCTGTCTCTTCCACTTGATTG	TAACCTCAATCCCATCCCAAGAT
40474366	40474366	G	-	STAT3	VarScan	SS	mut	na	CATGATCTTTCCTTCCCATGTCC	GCTGTATCCCCTCTTTAGACTCA
7578190	7578190	T	C	TP53	Both	SS	mut	na	CACTTGTGCCCTGACTTTCA	TTGCACATCTCATGGGGTTA
7578535	7578535	T	C	TP53	MuTect	SS	mut	na	CACTTGTGCCCTGACTTTCA	TTGCACATCTCATGGGGTTA
7577574	7577574	T	C	TP53	VarScan	SS	mut	na	CACTTGTGCCCTGACTTTCA	TTGCACATCTCATGGGGTTA
68834972	68834972	A	C	TPCN2	MuTect	SS	mut	wt	GAGATCCTGAGTTTGGTCCTGTC	TGTCACCCTTTCTTCTCCACTTA
98524941	98524941	G	A	TRRAP	MuTect	SS	mut	wt	GAAAAGACCATCCCCAATGTTAT	GAAAATCAATCAGCCAAACTCAG
42288436	42288436	A	T	UBTF	MuTect	SS	mut	wt	GAGATAGGGCACCATGCAGT	CTCAGACAGGTCGTTCCACA
6303900	6303900	G	A	WFS1	MuTect	SS	mut	na	ATCAACATGCTCCCGTTCTTCAT	AGGATGGTGCTGAACTCGATG
32417913	32417913	-	TT	WT1	VarScan	SS	mut	wt	AGGGAGTAGTTAGACTTTGGGAC	TATCTCTTATTGCAGCCTGGGTA
75245170	75245170	T	A	YLPM1	MuTect	SS	mut	mut	ATTTGGGGGAGGAACTGAAA	TTACCGGCTCTGGTGTATCC
37618563	37618563	A	G	ZNF420	Both	SS	mut	na	TCAGACGAGCCTCACACCTA	ACTTTTTGATGTCGGGTAAGTTGT
39923059	39923059	G	A	BCOR	MuTect	NGS	mut	na		

39911649	39911649	G	A	BCOR	MuTect	NGS	mut	na
39911499	39911499	T	C	BCOR	MuTect	NGS	mut	na
119148966	119148966	T	C	CBL	MuTect	NGS	mut	na
33792321	33792321	C	T	CEBPA	MuTect	NGS	mut	na
25462068	25462068	A	G	DNMT3A	VarScan	NGS	mut	na
25470551	25470551	C	A	DNMT3A	Both	NGS	mut	na
25467482	25467482	C	T	DNMT3A	Both	NGS	mut	na
25457242	25457242	C	T	DNMT3A	VarScan	NGS	mut	na
25457242	25457242	C	T	DNMT3A	VarScan	NGS	mut	na
25463271	25463271	G	C	DNMT3A	MuTect	NGS	mut	na
25467198	25467198	G	T	DNMT3A	MuTect	NGS	mut	na
28608327	28608327	G	C	FLT3	MuTect	NGS	mut	na
90631934	90631934	C	T	IDH2	MuTect	NGS	mut	na
44945176	44945176	T	G	KDM6A	MuTect	NGS	mut	na
25398284	25398284	C	T	KRAS	MuTect	NGS	mut	na
25398284	25398284	C	T	KRAS	MuTect	NGS	wt	na
25398281	25398281	C	T	KRAS	MuTect	NGS	wt	na
25398281	25398281	C	T	KRAS	MuTect	NGS	mut	na
25398284	25398284	C	T	KRAS	MuTect	NGS	mut	na
25380285	25380285	G	A	KRAS	MuTect	NGS	mut	na
170837543	170837543	-	TCTG	NPM1	VarScan	NGS	mut	na
170837543	170837543	-	TCTG	NPM1	VarScan	NGS	mut	na
115256529	115256529	T	C	NRAS	Both	NGS	mut	na
115258748	115258748	C	T	NRAS	MuTect	NGS	mut	na
198273279	198273279	C	A	SF3B1	MuTect	NGS	wt	na
198266834	198266834	T	C	SF3B1	MuTect	NGS	mut	na
198267371	198267371	G	C	SF3B1	MuTect	NGS	wt	na
106196267	106196267	C	T	TET2	VarScan	NGS	mut	na
7577120	7577120	C	T	TP53	Both	NGS	mut	na
7577538	7577538	C	T	TP53	Both	NGS	mut	na
7577505	7577505	T	A	TP53	VarScan	NGS	mut	na

7577082	7577082	C	T	TP53	Both	NGS	mut	na
7578388	7578388	C	T	TP53	VarScan	NGS	mut	na
7577557	7577557	G	-	TP53	VarScan	NGS	mut	na
7577136	7577136	C	-	TP53	VarScan	NGS	mut	na
7577570	7577570	C	T	TP53	MuTect	NGS	mut	na
7578265	7578265	A	G	TP53	MuTect	NGS	mut	na

Results – II

Gene fusions discovery

Under submission

RNA Sequencing Reveals Novel and Rare Fusion Transcripts in Acute Myeloid Leukemia

Antonella Padella[1*], Giorgia Simonetti[1*], Giulia Paciello[2], Anna Ferrari[1], Elisa Zago[3], Carmen Baldazzi[1], Viviana Guadagnuolo[1], Cristina Papayannidis[1], Valentina Robustelli[1], Enrica Imbrogno[1], Andrea Ghelli Luserna di Rorà[1], Nicoletta Testoni[1], Gerardo Musuraca[4], Simona Soverini[1], Massimo Delledonne[3], Ilaria Iacobucci[1#], Clelia Tiziana Storlazzi[5], Elisa Ficarra[2] and Giovanni Martinelli[1]

[1] Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy;[2] Dept. of Control and Computer Engineering DAUIN, Politecnico di Torino, Turin, Italy; [3] Personal Genomics, Verona, Italy; [4] Istituto Scientifico Romagnolo Per Lo Studio E La Cura Dei Tumori, Meldola (FC), Meldola, Italy; [5] Department of Biology, University of Bari, Bari, Italy.

*G.S. and A.P. contributed equally to this study.

#Current: Department of Pathology, ST Jude Children's Research Hospital, Memphis (TN), USA

Abstract

Chromosomal rearrangements and fusion genes have a crucial diagnostic, prognostic and therapeutic role in acute myeloid leukemia (AML) and about 30% of patients are characterized by the presence of a fusion transcript. We characterize by RNA sequencing the transcriptome of 8 AML patients with a rare or poorly described chromosomal translocation(s) detected by cytogenetic analysis, with the aim of identifying novel and rare fusion transcripts. We found 9 fusions, of which only one was previously described. Partner genes were transcription factors or oncosuppressors rarely altered across cancer types. Moreover, we detected cryptic events hiding the loss of *NF1* and *WT1*, two recurrently mutated genes in AML. The novel bioinformatic tool FuGePrior was exploited to prioritize detected fusions, which reduced the lists to fusions to those that were highly reliable from a structural point of view and with high probability of being drivers of the oncogenic process. Taken together, our results suggest that fusion genes, even if rare, contribute to the heterogeneity of AML and may be crucial for the development of the disease and the response to therapy.

Introduction

Acute myeloid leukemia (AML) is a neoplasm characterized by the accumulation of undifferentiated myeloid progenitors, which disrupt normal hematopoiesis. In the European Leukemia Net (ELN) 2017 classification of AML, fusion genes represent a major criterion of diagnosis and prognostic risk stratification[1]. The chromosomal translocation t(15;17), leading to the expression of *PML-RARα* chimera, characterizes patients with AML M3, which generally have favorable prognosis. AML expressing the transcript *RUNX1-RUNXT1* and *CBFβ-MYH11*, which associate with the t(8;21) and inv(16), respectively, are also classified as favourable prognosis. The t(6;9), inv(3)/t(3;3), t(v;11q23.3) and t(9;22) abnormalities results in the expression of *DEK-NUP214*, *GATA2*/*MECOM* fusions, *KMT2A*-fusions and *BCR-ABL*, respectively, which correlates with adverse prognosis[1].

Recently, a new classification model based on genomic aberrations has been proposed, which allowed the identification of 11 subtypes of molecularly-distinct AML[7]. Among the 1540 screened patients, 18.2% cases (280 patients, 6 out of 11 subgroups) were characterized by the presence of a known fusion genes as the main driver event[7].

Moreover, fusion genes resulting from chromosomal aberrations are common features of cancers and they represent an extremely attractive therapeutic targets. *PML-RARA* and *BCR-ABL* are two of the best examples were a fusion gene drives leukemogenesis and a target therapies are able to revert the leukemogenic phenotype.

Recent advances in Next Generation Sequencing (NGS) technologies, allowed the identification of novel fusion events in acute leukemias, which remained cryptic in the routine cytogenetic analysis. Togni and collegues disocvered the *NUP98-PHF23* fusion gene in paediatric cytogenetically normal AML carrying a cryptic chromosomal translocation between chromosomes 11 and 17[63]. In acute lymphoblastic leukemia, the *EPOR* rearrangement had been detected and it leads to the truncation of the cytoplasmatic tail of the receptor and its overexpression. Human leukemic cells with the *EPOR* rearrangement were sensitive to inhibition of the JAK-STAT signalling pathway, suggesting a therapeutic option for patients carrying this aberration[64].

Hence, the identification of fusion events, even if private or shared between a small subgroup of poorly characterized patents, may be of clinical significance. To identify fusion transcripts with a potential leukemogenic/pathogenetic role, we performed RNA sequencing (RNAseq) on 8 AML patients characterized by the presence of a rare or poorly described chromosomal translocation(s). We combined different approaches including cytogenetic analysis, RNAseq, state of the art bioinformatics pipelines and literature search to guide the identification of novel and rare fusion events. We validate the presence of 9 fusion genes involving either transcription factors, oncosuppressor, or fusions associated to a loss event involving a gene known to be altered in AML, demonstrating that the landscape of alterations in AML is not limited to known genes and fusion genes, although rare, may play an important role in the disease development.

Methods

Patients and samples

Primary adult AML samples were obtained after informed consent as approved by the Institutional Ethical Committee (protocol number 253/2013/O/Tess and 112/2014/U/Tess) of Policlinico Sant'Orsola-Malpighi (Bologna, Italy) in accordance with the Declaration of Helsinki.

Leukocytes were enriched by separation on Ficoll density gradient and lysed in RLT buffer. Genomic DNA and RNA were extracted by column purification (AllPrep DNA/RNA/Protein Mini Kit and QIAcube, Qiagen).

Chromosome Banding Analysis

Chromosome banding analysis was performed on bone marrow cells after short-term cultures (24 and/or 48 hours) as previously reported[65]. Karyotypes were examined after G banding technique and described according to International System for Human Cytogenomic Nomenclature (ISCN 2016)[66].

RNA Sequencing and fusion detection

Libraries for RNA sequencing were prepared with the TruSeq stranded mRNA kit (Illumina) following manufacturer's instructions. RNAseq libraries were subjected to 2×75 bp paired-end sequencing and run on a HiSeq 2500 or 1000 instrument (Illumina), and following manufacturer's instructions. An average of 50 million reads were obtained for each sample. We applied FuGePrior pipeline to the gene fusion lists output of ChimeraScan[67] and deFuse[68] tools ran with default parameter configuration. According to FuGePrior workflow, fusions with the following characteristics were removed to be further deepened by wet lab experiments: i) not supported by split reads (i.e., reads harboring the fusion breakpoint), ii) involving at least an unannotated partner gene, iii) shared by at least a healthy sample, iv) characterized by a not reliable structure and/or v) having at least a driver score probability (DS) lower than 0.7, which was a measure of the probability of the fusion of being an oncogenic event, according to several fusion properties according to Pegasus[69] and Oncofuse[70].

Validation of fusions

Selected fusions were validated by RT-PCR and Sanger sequencing. In particular, cDNA synthesis and PCR amplification were performed using standard protocols that come with M-MLV Reverse Transcriptase and Random Hexamers (Invitrogen) and Fast Start Taq DNA Polymerase (Roche). PCR primers were designed to amplify 200–400 bp fragments containing the fusion boundary detected by RNAseq using Primer3 (http://primer3.ut.ee/, Table 1). PCR products were purified with the QIAquick PCR purification kit (Qiagen) or, in cases where multiple PCR products were detected, conventional agarose gel electrophoresis and extraction of specific bands using the QIAquick Gel Extraction kit (Qiagen) was performed. PCR products were sequenced by Sanger Sequencing using an ABI PRISM 3730 automated DNA sequencer (Applied Biosystems) and the Big Dye Terminator DNA sequencing kit (Applied Biosystems). Fusion detection was performed using NCBI Blast alignment and BLAT software tool (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) to reference genome GRCh37/hg19.

Primer	Forward/Reverse	Sequence 5'-3'
ZEB2-BCL11B F	Forward	TCTTATCAATGAAGCAGCCGATC
ZEB2-BCL11B R	Reverse	AAGGGGAAGTTCATTTGACACTG
BCL11Bex2	Reverse	CTGCAATGTTCTCCTGCTTGG
BCL11Bex3	Reverse	CTGACAACTGACACTGGCATCCA
BCL11Bex4	Reverse	GCTGTCGCCCAGGAAATTCA
BCL11B-ZEB2 F	Forward	CGATGCCAGAATAGATGCCG
BCL11B-ZEB2 R	Reverse	ACTCATGGTTGGGCACACTA
WT1-CNOT2 F	Forward	CTTTTCACCTGTATGAGTCCTGG
WT1-CNOT2 R	Reverse	CGTCAACCCTGCTGTAATATCTC
CNOT2-WT1 F	Forward	GGGAGAACTTTCGCTGACAA
CNOT2-WT1 R	Reverse	CTCACCCTCCAACCTTTCCT
CPD-PXT1 F	Forward	GGCTCAGTGGTAGCAAGCTA
CPD-PXT1 R	Reverse	AAAGAGGGAGACGGAGAAGG
SAV1-GYPB F	Forward	AATGGAGACTCTGGTTCCCG
SAV1-GYPB R	Reverse	AGGTTGAAGTGTGCATTGCC
OAZ1-MAFK F	Forward	ATGGTGAAATCCTCCCTGCA
OAZ1-MAFK R	Reverse	CCAGCTTCTCCACCTCCTG
PUF60-TYW1 F	Forward	CTCCTATCCCGGTCACCATC
PUF60-TYW1 R	Reverse	TGTGTGCTATCAGGAGGCAA
UTP6-CRLF6 F	Forward	GGAACGCATAGAAGATCGGC
UTP6-CRLF6 R	Reverse	GCATGGCGATTTCACCTTCT
HINFP-RSRC2 F	Forward	TGTTTGTGTGTGTGTGTGTGTTT
HINFP-RSRC2 R	Reverse	GCGGGCGGATCACTTGAGGT
VPS13C-NCRNA00188 F	Forward	GAGCCAAGATCGCACCACTG
VPS13C-NCRNA00188 R	Reverse	TAAGCAATCCTCCCACCTTGTAG
FEZ1-TAOK3 F	Forward	AATCACCCAAAGTCAGGAATTCG
FEZ1-TAOK3 R	Reverse	GATCTCGGCTCACTACAACCTC
NUMA1-SLC35E3 F	Forward	CTTAAGAATAGACAAGACCGGGC
NUMA1-SLC35E3 R	Reverse	GTTCAAGTGATTCTCCTGCCTC
HIC2-PI4KA R	Reverse	TTTGGGTTGACTTGCTTCCG
PRKRIP1-PMS2 F	Forward	AGCATACGGGGATGTGATGT
PRKRIP1-PMS2 R	Reverse	TACCGCAGTGCTTAGACTCA
TPT1-EEF1A1 F	Forward	TTGATTTGTTCTGCAGCCCC
TPT1-EEF1A1 R	Reverse	CCATCCAGGCCAAATAAGCG

Table 1. Primers used for validation by RT-PCR and Sanger sequencing.

Results

RNAseq cohort selection

We screened our biobank of AML biological samples collected between 2010 and 2015. We identified 40 patients (less than 1% of our biobank) carrying a rare chromosomal translocation as the solely alteration (sample 20) or in association with other chromosomal abnormalities. Of these, biological material of 8 patients was available for sequencing (Table 2).

ID	Karyotype	Blasts	FAB	Phase	Validated fusion(s)
59810	46,XX,t(2;14)(q21;q32),t(11;12)(p15;q22)[17]/46,XX[3]	80%	M0	Diagnosis	4
20	46,XY,t(6;17)(p21;q11)[20]	90%	M2	Diagnosis	2
21	46,XY,t(3;12)(p22;q24),+4,-15,+mar [19]/46,XY[1]	80%	NOS	Relapse	1
32	45,XY,der(12)t(12;18)(p13;q12),-18 [12]/45,XY,t(4;16)(q31;q22),der(12)t(12;18)(p13;q12),-18[4]/45,XY,der(6)t(6;12;18)(p21;p13,q12),-18 [3]/46,XY [1]	80%	NOS	Relapse	0
84*	47,XX,+8,del(11)(p11p15),t(15;17)(q24q25),inv(16)(p13q22)[20]	80%	M1	Diagnosis	1*
68187	46,XX,add(8)(p23),der(16)t(1;16)(q11;q11) [18]/46,XX[2]	70%	II AML	Diagnosis	1
63569	46,XY,add(10)(p15)[9]/46,XY,add(10)(p15),t(1;8)(p36;q13)[2]/46,XY[20]	70%	M2	Relaspe	0
125	44~47,XX,t(4;17)(p15;q21),del(5)(q13q33),-7,-18,der(X),+1~3mar[cp9]/46,XX[11]	50%	M2	Diagnosis	1

Table 2. Characteristic of the characterized patients and number of validated fusion for each patient. * positive control.

Identification and validation of chimera

Paired-end RNAseq data were analyzed with deFuse and Chimerascan in order to detect fusion transcripts. To select biologically relevant fusions, the list of putative chimeric transcripts obtained by the aforementioned tools were prioritized using the novel FuGePrior tool.

We identified an average of 416 fusions per patient (range 133-895): 130 of them were in frame and 284 were out-of-frame events (range 47-27 and 72-631, respectively). Firstly, we focused on fusion genes originating from chromosomal translocations detected by the cytogenetic analysis. Then, to detect cryptic fusions we excluded non-specific fusions involving genes showing a large diversity among partner genes (such as *HBB*, *HBA*, *HBD*, *MPO*, *DLG2*, ecc.), conjoined genes and fusions recurring in our cohort. We also evaluated the base composition of the fusion sequences, i.e. splitting reads with sequence homology were excluded. Finally, a manual curation approach was used to filter out fusions without biologically relevant functions. We classified the candidate fusion transcripts for validation into three tiers based on level of evidence:

- Tier 1: fusion genes whose partners mapped in the genomic region translocated, according to cytogenetic analysis.
- Tier 2: chimeras with the driver score obtained with Pegasus and/or Oncofuse ≥0.7.
- •Tier 3: fusions deriving from biologically and functional plausible translocation mechanisms.

The recurrent gene fusion *CBFβ-MYH11* was identified in sample 84 [47,XX,del(11)(p11p15);t(15;17)(q24q25),inv(16)(p13q22)], thus confirming the reliability of our bioinformatic analysis. Nineteen fusions were selected for experimental validation by RT-PCR and Sanger sequencing. Of these, 9 were validated (47% of selected fusions, Table 3). No chimeras were detected and/or confirmed in samples 32 and 63569.

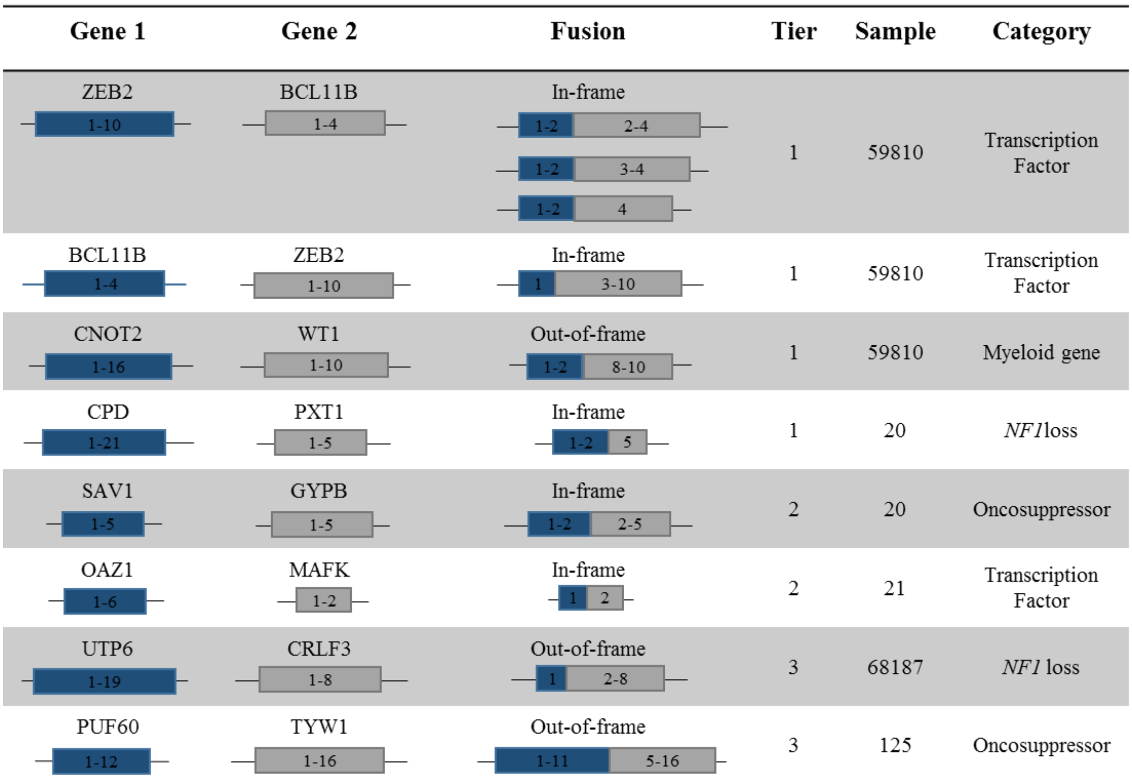


Figure 1. Schematic representation of validated fusion genes. Exon numbers are reported in the boxes. Reading frames, tier, and potential functional category/altered pathway of the putative fusion protein is specified.

Sample	Program	Chr1	Chr2	Gene 1	Gene2	Breakpoint1	Breakpoint2	Reading Frame	Validation	Pegasus	Oncofuse	Tier
68187	defuse	7	7	PRKRIP1	PMS2	102016769	6045662	FrameShift	NO	0.0084	0.004	3
63569	defuse	22	22	HIC2	PI4KA	21797146	21088841	FrameShift	NO	0.0084	0.02	3
125	defuse	13	6	TPT1	EEF1A1	45914213	74227825	FrameShift	NO	0.7772	0.0004	2
59810	defuse	11	12	WT1	CNOT2	32414299	70688076	FrameShift	NO	0.0077	0.009	1
59810	defuse	11	12	HINFP	RSRC2	118995969	122999034	FrameShift	NO	0.0065	0.01	1
59810	defuse	11	12	FEZ1	TAOK3	125322956	118636293	FrameShift	NO	0.0079	0.003	1
59810	defuse	11	12	NUMA1	SLC35E3	71777209	69158894	-	NO	0.0053	0.02	1
84	defuse	15	17	VPS13C	NCRNA00188	62146254	16345042	-	NO	*	*	1
21	chimerascan	12	3	CPSF6	PPP4R2	69656342	73114548	InFrame	NO	0.6992	0.21	3
68187	defuse	17	17	UTP6	CRLF3	30228555	29131126	FrameShift	YES	0.0084	0.006	3
125	chimerascan	8	7	PUF60	TYW1	144899080	66648113	FrameShift	YES	0.0076	0.02	3
59810	both tools	14	2	BCL11B	ZEB2	99737497	145187592	InFrame	YES	0.7365	0.99	1
59810	both tools	2	14	ZEB2	BCL11B	145274845	99724176	InFrame	YES	0.7365	0.99	1
59810	defuse	2	14	ZEB2	BCL11B	145274845	99697894	InFrame	YES	0.7365	0.99	1
59810	chimerascan	12	11	CNOT2	WT1	70672053	32414300	FrameShift	YES	0.0084	0.02	1
20	chimerascan	17	6	CPD	PXT1	28712254	36359651	InFrame	YES	0.0084	0.08	1
20	chimerascan	14	4	SAV1	GYPB	51131897	144922436	InFrame	YES	0.8774	0.02	2
21	chimerascan	19	7	OAZ1	MAFK	2269743	1578785	InFrame	YES	0.0053*	0.97	2
84	both tools	16	16	CBFβ	MYH11	67116211	15814908	InFrame	YES (+)	0.7968	0.91	1

Table 3. List of fusions selected for validation by RT-PCR and Sanger Sequencing (Chr: chromosome).

***ZEB2-BCL11B* fusion**

Sample 59810 carried the fusion transcript *ZEB2-BCL11B* (tier 1, Figure 1 and 2), which is an in-frame fusion and a rare event in AML associated with t(2;14)(q21-22;q32)[71]. The breakpoint of the fusion mapped in exon 2 of *ZEB2* (NM_014795) and exon 2 of *BCL11B* (NM_00128223; Figure S1A). Of note, we identified 3 splicing isoforms (Figure S1A-C), two of which were not reported before[71]. Type 1 isoform is the full-length chimera and it retains all exons of both genes involved in the translocation. Type 2 isoform was also detected by the bioinformatic pipeline and it was characterized by the junction of exon 2 of *ZEB2* and exon 3 of *BCL11B*. In type 3 isoform, exon 2 and 3 of *BCL11B* were removed, resulting in an mRNA composed by exon 2 of *ZEB2* and exon 4 of *BCL11B.* The reciprocal fusion transcript was also detected and validated and it was formed by exon 1 of *BCL11B* and exon 3 to 10 of *ZEB2* (Figure S1D). *BCL11B* is a Kruppel family zinc finger family gene and is a transcriptional co-repressor. It associates with NURD nucleosome remodelling and histone deacetylase complex. It has a pivotal role in differentiation and survival of T-cell[72,73]. *ZEB2* is a fundamental transcriptional factor for hematopoiesis: it controls adult hematopoietic differentiation[74] and its knockdown in AML cell lines resulted in the releases of the granulocytic differentiation block and proliferation arrest[75]. Moreover, according to the Mitelman Database of Aberrations of Chromosomes and Gene Fusion in Cancer[20], we found that 12 hematological patients were characterized by the translocation t(2;14)(q21-22;q32), suggesting that, even though it is a rare event, the resulting fusion gene may have a pathogenic role. The *ZEB2-BCL11B* fusion has been described in AML and its oncogenic role may be linked to the overexpression of *BCL11B*[71,76], while the reciprocal fusion has never been described and its role in leukemogenesis is unknown. The expression of the chimera was confirmed at protein level by immunohistochemistry using an anti-BCL11B antibody (data not shown).

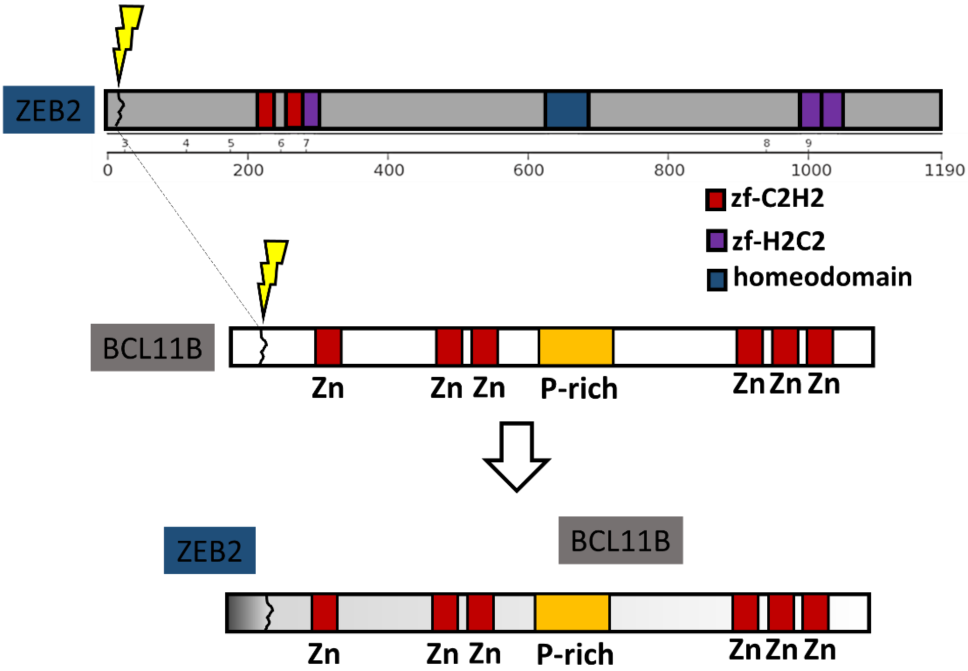


Figure 2. Putative ZEB-BCL11B fusion protein.

OAZ1-MAFK fusion

In sample 21 we identified a novel fusion event between chromosomes 19 and 7, involving the genes *OAZ1* and *MAFK* (tier 2, Figure 1). *OAZ1* is an Ornithine decarboxylase (*ODC*) antizyme protein that negatively regulates *ODC* activity. ODC controls polyamine homeostasis and OAZ1 suppress polyamine production by inhibiting the functional assembly of ODC homodimer[77]. A high concentration of polyamines leads to the translation of *OAZ1* full length, which binds to *ODC* and it inhibits its activity. *MAFK* is a transcriptional factor, whose role of activator or repressor of transcription depends on the interacting proteins[78]. The breakpoint of the fusion mapped in exon 1 of *OAZ1* (NM_004152), which encodes for a polyamine sensing region and a proteasome interaction domain. The breakpoint at 3' mapped in exon 2 of *MAFK* (NM_002360), which, together with exon 3, encodes for the bZIP domain (Figure S2B). The prediction of the putative fusion protein is outlined in Figure 3[77,78]: The transcript was formed at 5' by the sensing regions of polyamine that normally controls the transcription of OAZ1, while the 3' encodes for the bZIP domain of MAFK.

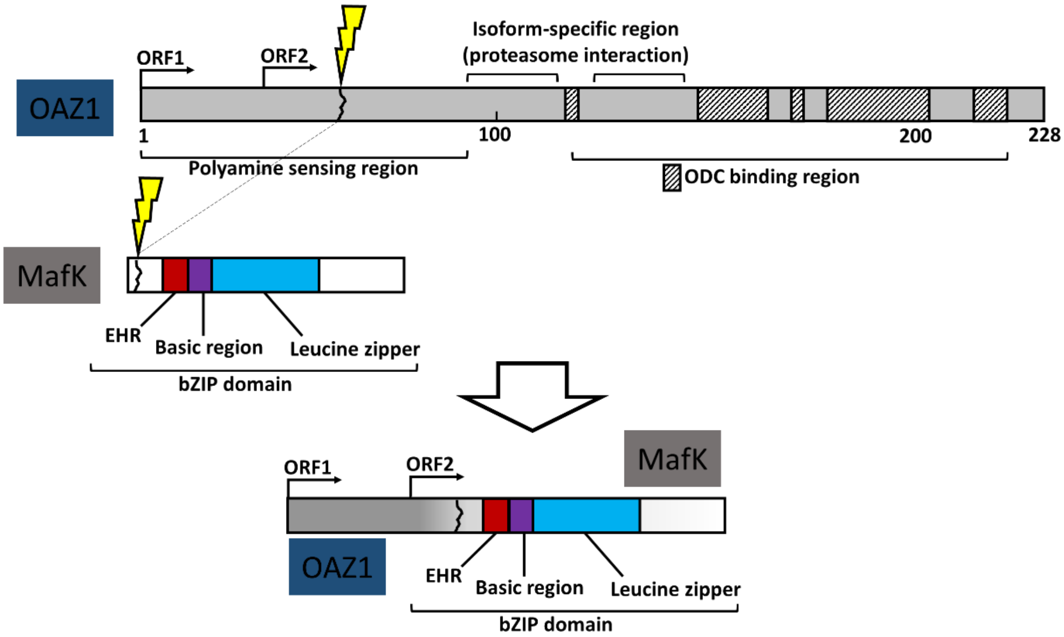


Figure 3. Putative OAZ1-MAFK fusion protein

SAV1-GYPB **fusion**

We confirmed in sample 20 the in-frame transcript *SAV1-GYPB*, which remained cryptic at cytogenetic analysis and the driver score predicted by Pegasus (DS=0.87) enabled us to detected the chimera as a potential driver of leukemogenesis (tier 2, Figure 1). The breakpoint mapped in chromosome 14p22, exon 2 of *SAV1* (NM_021818) and chromosome 4q31, exon 2 of *GYPB* (NM_002100, Figure S2B). *SAV1* encodes for a protein characterized by two WW domanis, a SARAH domain and a coiled-coil region (N- to C-terminal). It is a tumor suppressor involved in the Hippo pathway, where it negatively regulates proliferation and apoptosis by interacting with the kinases STK3/MST2 and STK4/MST1 via the coiled-coil domain, and promoting the exit from the cell cycle and differentiation[79]. Studies on high-grade clear cell renal cell carcinoma revealed its oncosuppressive role: *SAV1* is downregulated due to a copy number loss in the 14q22 region[80]. In the chromosomal rearrangement described in sample 20, *SAV1* lost the stabilization and interaction domains including the WW domain and the colied-coil domain[81,82], while *GYBP*, a sialoglycoproteins of the human erythrocyte membrane, lost the N-terminal domains and retained the dimeric transmembrane domain in the fusion (Figure 4).

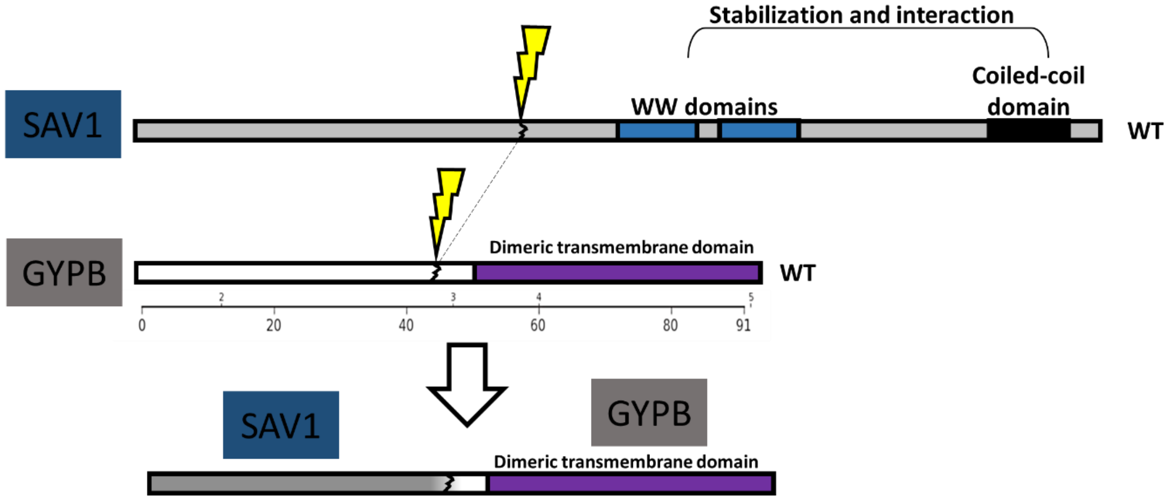


Figure 4. Putative SAV1-GYPB fusion protein

CPD-PXT1 fusion

A new in-frame fusion gene were identified in sample 20: *CPD-PXT1* (tier 1, Figure 1), which appeared as the reciprocal fusion product of t(6;17)(p21;q11) translocation. The breakpoint of the fusion mapped in exon 2 of *CPD* (NM_001304) and exon 5 of *PXT1* (NM_152990, Figure S3B). *CPD* encodes for a metallocarboxypeptidase and it maps in chromosome 17q11, while the role of *PXT1* (6p21) is unknown. The translocation hid a copy-neutral loss of heterozigosity involving *NF1* (data not shown), which has been reported as deleted in 5% of AML[83]. Moreover, the patient was characterized by a mutation in *NF1* (NM_001042492:exon29:c.C3916T:p.R1306X) detected by Whole Exome Sequencing (data not shown), suggesting the lack of negative regulation of the RAS pathway.

Out-of-frame fusions

The sample 59810 showed the *CNOT2-WT1* chimera, which is a novel out-of-frame fusion (tier 1, Figure 1) related to t(11;12)(p15;q22) translocation, identified by cytogenetic analysis. The breakpoint mapped in exon 2 of *CNOT2* (forward strand, NM_014515) and exon 8 of *WT1* (reverse strand, NM_024424, Figure S3A). We also detected a splicing variant mapping in exon 3 of a non-coding transcript variant of *CNOT2* (NR_037615, data not shown). The partner genes mapped in opposite strands, therefore *CNOT2* sequence conserved its orientations, while *WT1* sequence had inverted orientation. The *CNOT2-WT1* transcript has never been annotated in cancer. Moreover, the translocation was associated to a deletion at 5' of *WT1* and 3' of *CNOT2*, which remained cryptic at cytogenetic analysis, leading to a homozygous loss of *WT1*. *CNOT2* encodes for a subunit of the multi-component CCR4-NOT complex, which is involved in transcriptional regulation, mRNA degradation, miRNA-mediated repression and translational repression during translational initiation[84–86]. *WT1* is a transcription factor and it is recurrently altered in haematological malignancies, including AML[7,87].

In sample 68187 we validated the out-of-frame fusion *UTP6-CRLF3* (tier 3, figure 1). The two breakpoints mapped in chromosome 17, specifically in exon 1 of *UTP6* (NM_018428) and exon 2 of *CRLF3* (NM_015986, Figure S3C). *UTP6* is involved in nucleolar processing of pre-18S ribosomal RNA, while *CRLF3* is a cytokine receptor-like factor that may negatively regulate cell cycle progression at the G0/G1 phase[88]. The breakpoint of the chimera mapped on the reverse strand of chr17: 29131126 and chr17: 30228555, suggesting a loss event leading to the haploinsufficiency of *NF1*, which maps in the forward strand of chromosome 17: 29421945-29709134 (GRCh37).

We confirmed the presence of the frameshifted *PUF60-TYW1* in sample 125 (tier 3, Figure 1). The chimera involved the exons 1-11 of *PUF60* (chromosome 8, NM_001271098) and the exons 5-16 of *TWY1* (chromosome 7, NM_018264, Figure S3D). *PUF60* participates in the splicing machinery[89] while the role of *TYW1* may be a component of the wybutosine biosynthesis pathway. It has been demonstrated that *PUF60* haploinsufficiency was involved in *TP53*-dependent progression of a T-cell acute lymphoblastic leukaemia (T-ALL) in a mouse model[89,90]. A database search revealed that *TYW1* is frequently involved in fusion formation with a variety of partners in different tumor types[61], suggesting that it maps in an unstable genomic region.

Discussion

Fusion genes are frequently present in cancer and they are often caused by chromosomal rearrangements such as translocations, inversions and deletions, which may involve the same or different chromosome. With the advent of NGS, driver alterations in AML have been extensively described, however, 4% of AML patients have no known driver alterations[7]. The TCGA study revealed that an average of 1.5 fusions characterizes each AML patient and it depicted an heterogeneous landscape of chimeras[87], where very few fusions and partner genes were recurrently rearranged and altered, respectively. Here we described novel and rare fusion events with a predicted pathogenic role in AML, that allow a better characterization of AML patients in a precision medicine perspective and may provide insights for the design of novel targeted therapies. We selected a cohort of AML patients characterized by the presence of a rare or never described chromosomal translocation according to cytogenetic analysis, in order to detect the putative fusion gene associated to the translocation. Moreover, to identify cryptic events, we took advantage of FuGePrior. State of the art bioinformatics pipelines for gene fusion discovery from RNAseq data are characterized by a two-fold drawback. First, these pipelines provide in output a large number of putative chimeric transcripts, generally plagued with considerable amounts of false positive predictions[91]. This is mainly due to systematic errors including read-through artefacts, reverse transcriptase template switching events or mapping biases. Second, gene fusion identification tools provide no information regarding the oncogenic relevance of the output fusions. These features make the systematic experimental validation of gene fusion lists obtained from in silico pipelines unfeasible. To select a reduced number of biologically relevant fusions, we prioritized the list of putative chimeric transcripts with the novel FuGePrior tool. FuGePrior combines results from state of the art bioinformatic tools for chimeric transcripts identification and prioritization, several filtering and processing steps designed on up-to-date literature on gene fusions and analysis of the potential functionality of the fusion according to its structure.

The shortlist of fusions to be validated were classified into tier 1 (translocated partner genes mapping in the genomic region identified by cytogenetic analysis), tier 2 (DS≥0.7), tier 3 (fusions deriving from biologically and functional plausible translocation mechanisms). In a cohort of 8 patients, we validated 5 fusion genes associated with the cytogenetic translocations (tier 1) and 4 fusions which remained cryptic at cytogenetic analysis (2 fusions were classified as tier 2 and 2 were tier 3). The tier-1 fusions were the two isoform of *ZEB2-BCL11B* and its reciprocal *BCL11B-ZEB2* associated with the translocation t(2;14), *CNOT2-WT1* which derived from the translocation t(11;12) and *CPD-PXT1* related to the t(6;17) aberration (Table 2, Figure 1). The cryptic fusions (tier 2 and 3) included *UTP6-CRLF3*, *PUF60-TYW1*, *SAV1-GYPB* and *OAZ1-MAFK* (Table 2, Figure 1). The nature of these cryptic fusions may be ascribable to low-frequency genomic alterations (i.e. structural alterations present in few clones) or to a post-translational events where a chimeric transcript may generate from a trans-splicing event leading to the expression of a chimeric transcript[92,93].The fusions *ZEB2-BCL11B*, *BCL11B-ZEB2* and *OAZ1-MAFK* involved genes which encode for transcription factors and the putative mechanism of action of the fusion proteins may be linked to alterations of the cellular transcriptional program by the novel chimera. In particular, the fusion protein *ZEB2-BCL11B* is formed by 24 residues of *ZEB2* and 803 out of 823 residues of *BCL11B.* We hypothesize a mechanism of action linked to the overexpression of *BCL11B*[71,76,94] driven by the hematopoietic transcription factor *ZEB2* . However, for the reciprocal product *BCL11B-ZEB2*, studies may elucidate its role in the leukemogenesis.

By integrating array and mutational data, we linked the presence of fusions like *UTP6-CRLF3* and *CPD-PXT1* to the loss of *NF1*, which is a negative regulator of the RAS pathway. These alterations suggest that RAS signalling may be frequently deregulated at different levels. Therefore, in a precision medicine perspective, these "hidden" alterations requiring a combination approach to be revealed, must be taken into account. The consequences of *CNOT2-WT1* and *PUF60-TYW1*, which are out of frame events, may be linked to the loss of function of partner genes, which in turn, may lead to a potential mechanism of haploinsufficiency. Indeed, alterations of *WT1* including point mutations and small indels was described associated to AML and it was detected in approximately 5% of cases[7,87], while the haploinsufficiency of *PUF60* has been associated with the progression of T-ALL in a mouse model[90]. Functional studies are needed to elucidate its role in AML. *CNOT2* and *TYW1* are frequently translocated with different partners in other cancer types[61], suggesting that their genomic location may be prone to chromosomal rearrangements[95,96]. The fusion gene *SAV1-GYPB* may be of interest due to the role of the oncosuppressor *SAV1* in controlling cell cycle and apoptosis in the Hippo pathway[81]. It is a scaffold protein that interacts with the kinases Mst1 and Mst2 via the coiled-coil domain and form a complex which may be in an active (phosphorylated) or inactive (not phosphorylated) state. Moreover, Sav1 not bound to Mst is less stable than Sav1 bound to Mst. When activated by upstream stimuli, the kinases Mst may phosphorylate Sav1, which induce a conformational change in Sav facilitating recruitment of substrates to the complex and downstream effects[82]. The coiled coil and the stabilization domains of Sav1 is disrupted in the fusion, which may lead to a decreased stability deriving from the interaction and phosphorylation by Mst[82] and the loss of its function in the Hippo pathway.

Our understanding of the genomic landscape of AML has been dramatically improved, leading to a progressive shift from driver discovery to diagnostic application, in which the identification of alterations in each patient is fundamental to tailor a personalized therapeutic intervention. However, targeted sequencing approaches are limited to the detection of alterations in few genes and we have shown in this study that the landscape of AML genomic lesions is not restricted to known genes, and fusion genes, although rare, may play an important role in the disease development and progression. Functional studies will elucidate the potentiality of rare fusions as driver events of leukemogenesis and as therapeutic targets.

Supplementary figures

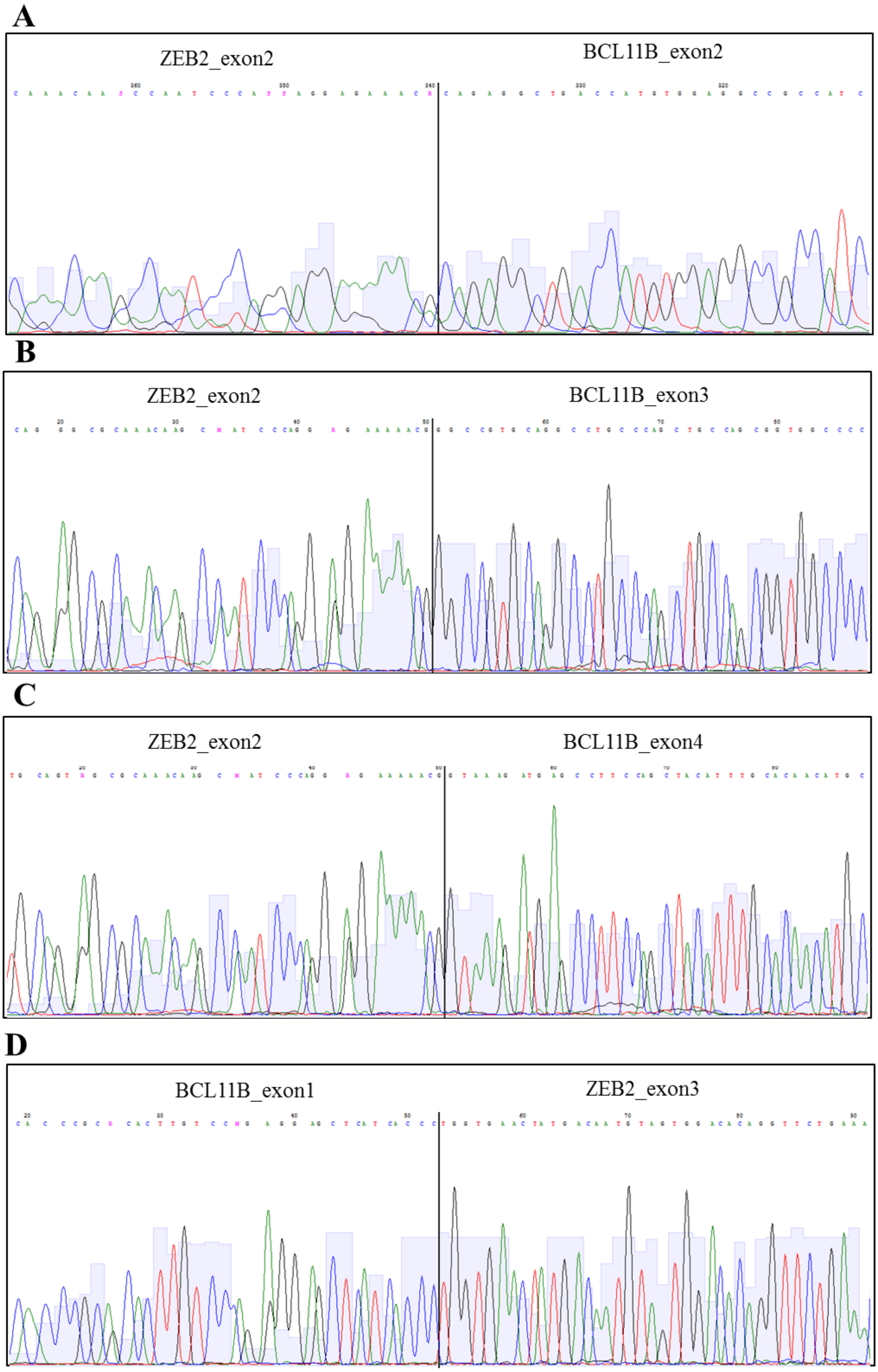


Figure S1. Electropheogram of breakpoint of *ZEB2-BCL11B* and related fusions. (A) *ZEB2-BCL11B* full-length (tier 1). (B) *ZEB2-BCL11B* isoform 2. (C) *ZEB2-BCL11B* isoform 3 (D) *BCL11B-ZEB2* (tier 1).

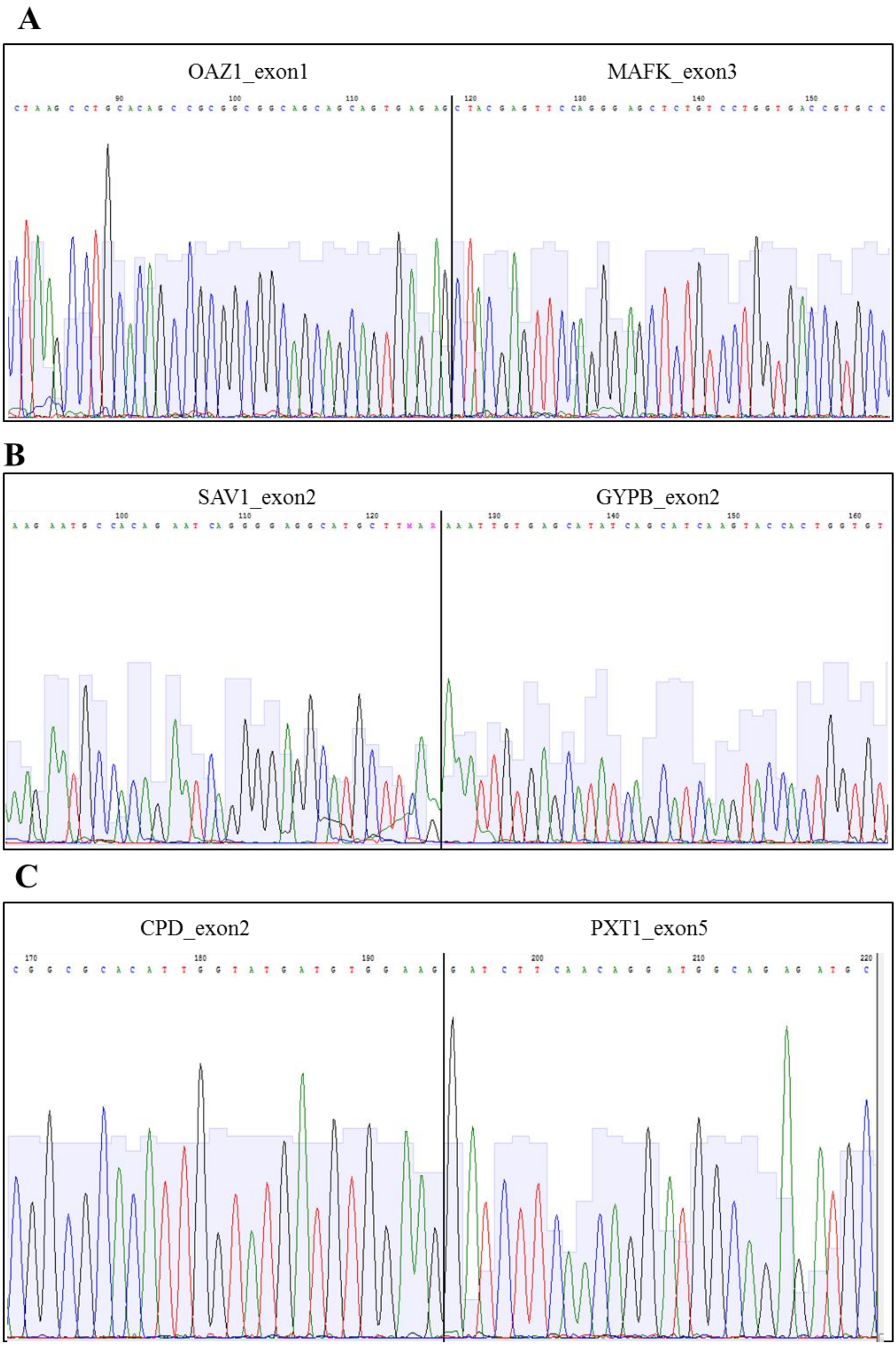


Figure S2. Electropheogram of breakpoint of in-frame fusion. (A) OAZ1-MAFK (tier 2) (B) SAV1-GYPB (tier 2). (C) CPD-PXT1 (tier 1).

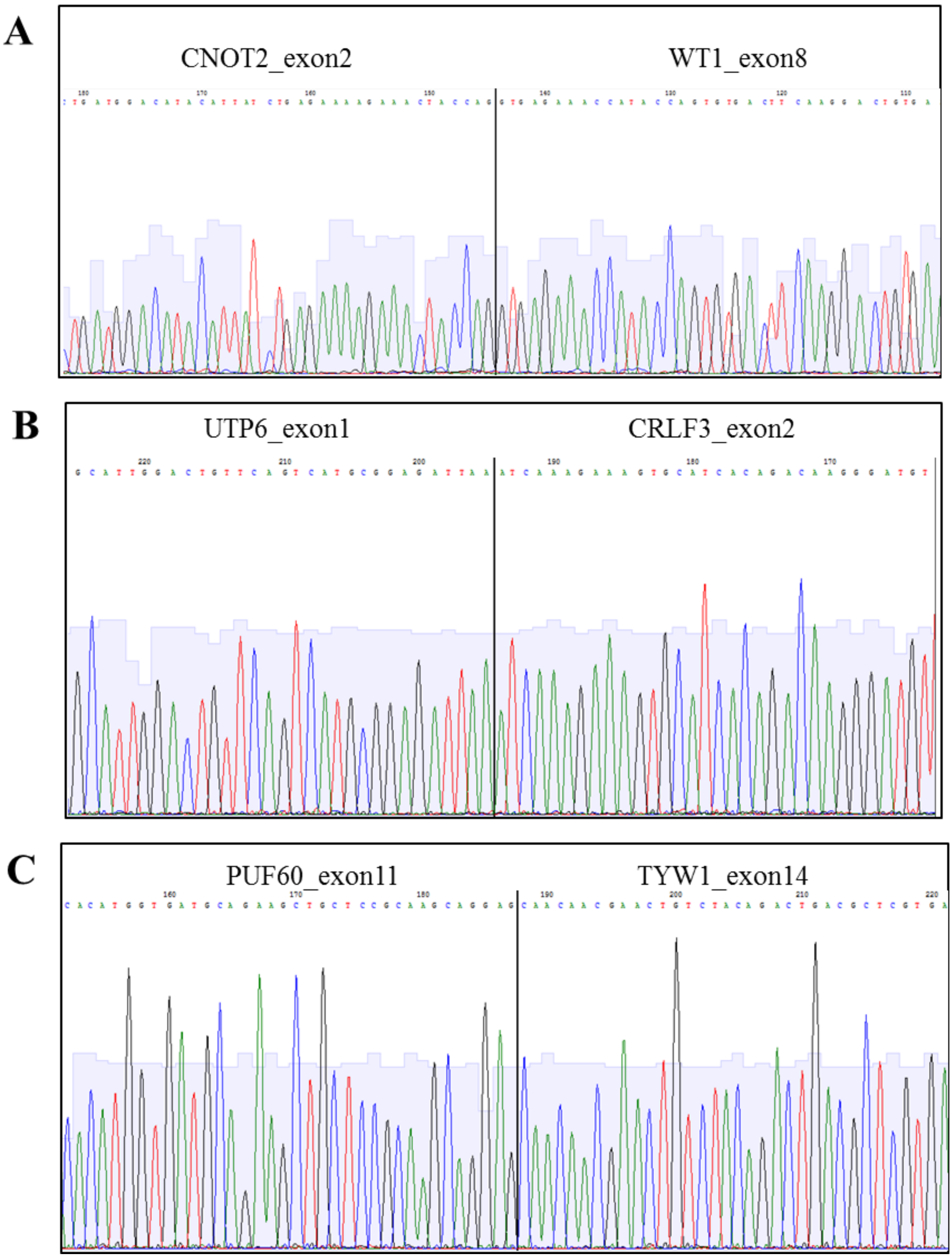


Figure S3. Electropheogram of breakpoint of out-of-frame fusions. (A) CNOT2-WT1 (tier 1). (B) UTP6-CRLF3 (tier 3). (C) PUF60-TYW1 (tier3).

Results – III

Functional studies on *ZEB2-BCL11B* fusion

Introduction

Among 8 patients characterized by RNAseq, we detected the rare fusion transcript *ZEB2-BCL11B* (sample 59810) and its reciprocal *BCL11B-ZEB2*. The *ZEB2-BCL11B* chimeric transcript fuses exon 2 of *ZEB2* (NM_014795) and exon 2 of *BCL11B* (NM_001282238) and we also identified 2 splicing isoforms and the reciprocal fusion transcript, which was formed by exon 1 of *BCL11B* and exon 3 to 10 of *ZEB2*. Co-operating mutations in *TET2* (NP_001120680.1:p.Gln1534Ter) and *FLT3* (NP_004110.2:p.Asp839Gly) were detected in the same patient. *BCL11B* is a Kruppel family zinc finger family gene located at 14q32 and is a transcriptional co-repressor complexes which associates with NURD nucleosome remodeling and histone deacetylase complex[94]. It has a pivotal role in differentiation and survival of hematopoietic cells. In particular, the encoded protein is a tumor suppressor in the T-cell lineage, thus being involved in T-cell differentiation and proliferation[72,73]. The 14q32 region and *BCL11B* are subject to translocation and cytogenetic alteration in T cell malignancies including T-cell acute lymphoblastic leukemia (T-ALL) and acute mixed lineage leukemia[94,97–99]. However, *BCL11B* was also identified as a putative oncogene in AML[76] and Torkildsen and colleagues detected the fusion gene Z*EB2-BCL11B* associate with the translocation t(2;14)(q22;q32)[71]. *ZEB2* is a transcription factor involved in the epithelial to mesenchymal transition and, therefore, its expression is linked to metastasis formation in various cancer types[100]. Two recent studies have uncovered the role of *ZEB2* in hematopoiesis and AML. In particular, a mouse model with a conditional deletion of *ZEB2* showed impaired differentiation of myeloid progenitors, B-cell, myeloid precursors as well as terminally differentiated cell[74]. On the other hand, using AML and non-AML cell lines, H. Li and colleagues identified *ZEB2* as an AML-specific dependency, the role of which has also been validated in vivo through a murine model of AML[75]. Both studies reported the overexpression of *ZEB2* in hematopoietic stem cell and AML, and a granulocyte differentiation shift upon *ZEB2* deletion or inhibition in both normal and malignant hematopoiesis, suggesting a crucial role of *ZEB2* in stemness maintenance[101].

To assess the oncogenic potential of the fusion *ZEB2-BCL11B* and its reciprocal transcript, together with characterization of the mechanism(s) of leukemogenesis promoted in AML, we performed functional studies on the full length fusion transcript in collaboration with Prof. Brian Huntly at the Department of Heamatology, University of Cambridge.

Methods

Retroviral transduction assays

The TY1-tagged full length transcripts *ZEB2-BCL11B* was subcloned into a retroviral vector using EcoRI restriction sites. After E.Coli bacteria was transformed, the plasmid's sequence was checked by Sanger sequencing to screen for spontaneous mutations. Murine stem cell virus–based (MSCV-based) retroviral constructs carrying the tagged *ZEB2-BCL11B* sequence upstream of an internal ribosomal entry site–green fluorescent protein (IRES-GFP) cassette were generated using 293T packaging cell line. Vectors containing the fusion gene (ZEB2-BCL11B), the *MLL-AF9* fusion (positive control) or the empty vector (EV) was then used to transduce mouse c-Kit+ bone marrow (BM) cells. Mouse whole BM were positively selected using the CD117 (c-Kit) MicroBeads and the LS MACS column according manufacturers protocol (Miltenyi Biotec). Retroviral transduction were performed as previously described[102].

Serial replating assays

To assess the effect of the transcript on the clonogenic ability, colony forming unit assay (CFU-A) was performed by seeding 1000 CD117+ transduced cells in Methocult M3434 methylcellulose medium (StemCell Technologies) supplemented with 100 ng/ml stem cell factor (SCF), IL3 10 ng/ml and IL-6 20 ng/ml (PEPROTECH). Cells were plated in duplicate and after 7-12 days colonies were scored, pooled and identical numbers of cells were re-plated under the same conditions.

Flow cytometry analysis

Sinlge cell suspension of transduced cells were prepared as described[102]. Dead cells were excluded by gating on 7AAD (Miltenyi Biotec)-negative cells. Flow cytometry analysis were performed on an LSR Fortessa cell analyser (BD Biosciences) and data were analysed with FlowJo software v 10 (Tree Star).

Immunoblotting

Whole-cell lysates were prepared from 10×10^6 cells in 6× Laemmli buffer. Lysates were run on SDS–PAGE gels and transferred to PVDF membranes (Millipore) using standard protocols. Membranes were probed with the anti-GAPDH (Abcam), anti-TY1 (ThermoFisher) and anti-BCL11B (Abcam) primary antibodies at 1:10000, 1:2000 and 1:10000 dilutions, respectively. Membranes were probed with secondary antibodies conjugated to IRDye 680RD or IRDye 800 CW (LI-COR Biosciences Ltd) at 1:10000 dilution and proteins were detected using the

Odyssey Infrared Imaging System (LI-COR Biosciences Ltd). Restore Western Blot Stripping Buffer (Thermo Scientific) were used to remove primary and secondary antibodies from PVDF membrane to reprobe with the anti-BCL11B antibody.

Quantitative real-time PCR (qPCR)

Total RNA was isolated using an RNeasy Mini Kit (Qiagen). cDNA was prepared from 1 µg RNA using the SuperScript III First-Strand Synthesis System (Invitrogen) qRT-PCR was performed on diluted cDNA (1:10 in water), using Brilliant III Ultra-Fast QPCR Master Mix (Agilent) and gene-specific primers (Sigma-Aldrich) on an Mx3000p qPCR system (Agilent) and standard cycling set-up. Primer sequences were retrieved from Primer3 (http://primer3.ut.ee) and are listed below (F, forward; R, reverse; all 5′-3′):

ZEB2 F1, 5'-TGCCATCTGATCCGCTCTTA-3'

BCL11B R1, 5'-TTCCAGTCCTTCATCCTCTTCC-3'

Gapdh_F, 5′-TGACGTGCCGCCTGGAGAAA-3′

Gapdh_R, 5′-AGTGTAGCCCAAGATGCCCTTCAG-3′.

Gene expression levels were determined by the $2^{-\Delta\Delta CT}$ method following normalization to *Gapdh*.

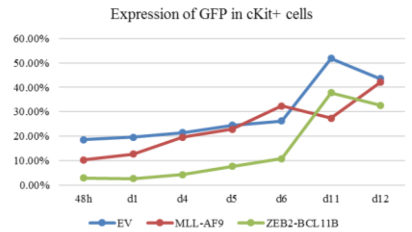
Results

Generation of c-Kit+ cells expressing the tagged *ZEB2-BCL11B* fusion

To check the successful transduction of c-Kit+ cell and to test the level of expression of the fusion gene, GFP+ cells were monitored by flow cytometry for 14 days (as a surrogate of fusion gene expression) and the levels of mRNA expression were measured by quantitative PCR (qPCR) using primers spanning the breakpoint of the fusion gene. Flow cytometry analyses showed that the expression of the GFP in transduced cells was increasing over time and it reached nearly 40% in cells transduced with the MSCV-ZEB2-BCL11B-IRES-GFP after 14 days post-transduction (Figure 1A). The qPCR confirmed that the chimera *ZEB2-BCL11B* were specifically expressed in cells transuded with the relative vector, however its expression were relatively low (Figure 1B).

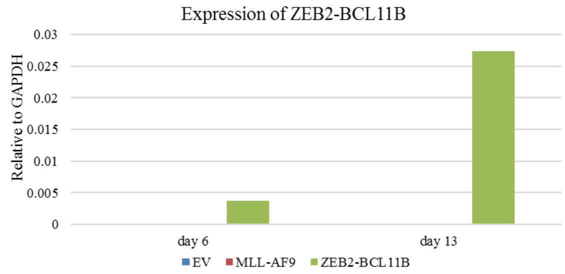
A

GFP+ cells							
Vector	48h	Day 1	Day 4	Day 5	Day 6	Day 11	Day 12
EV	18.60%	19.60%	21.40%	24.50%	26.20%	51.80%	43.50%
MLL-AF9	10.20%	12.60%	19.50%	22.80%	32.40%	27.40%	42.00%
ZEB2-BCL11B	2.88%	2.54%	4.23%	7.58%	10.80%	37.80%	32.60%



B

qPCR gene expression		
Vector	Day 6	Day 13
EV	0.0001207	2.731E-05
MLL-AF9	6.563E-05	3.21E-05
ZEB2-BCL11B	0.0036779	0.0274061



C

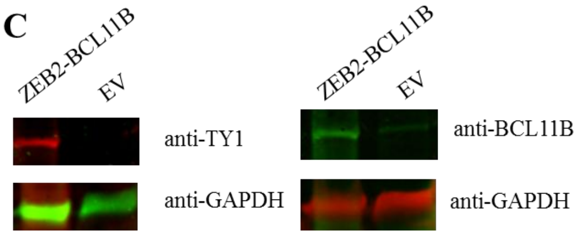


Figure 1. **Expression of ZEB2-BCL11B in transduced cells**. (A) Flow cytometry of transduced c-Kit+ cells over time. GFP expression increased over time and 43% cells were GFP+ at day 12 post transducition. (B) qPCR performed on cells harvested at day 13. The fusion is specifically expressed in cells transduced with the relative vector, however its expression is relatively low when compared to *GAPDH*. (C) Western blot on transfected 293T cell lysates showing the presence of the tagged ZEB2-BCL11B protein, which is specifically expressed in cells transfected with the MSCV-ZEB2-BCL11B-IRES-GFP construct. The same membrane was stripped and reprobed with the anti-BCL11B antibody.

Therefore, to check the presence of the protein product, Western Blot was performed. The anti-TY1 antibody were used in 293T cells lysate transfected with either the EV or MSCV-ZEB2-BCL11B-IRES-GFP. Since the fusion had the C-terminal of BCL11B intact and the molecular weight of the fusion and the wilde-type BCL11B were the same, the anti-BCL11B antibody was also used to detect the putative fusion protein (Figure 1C). In conclusion, 293T and c-Kit+ cells were successfully transduced, thus confirming the functionality of the vector and the proper translation of the relative mRNA.

c-Kit+ cells showed no self-renewal in vitro upon ZEB2-BCL11B expression

To assess the ability of the fusion gene to alter the self-renewal of murine hematopoietic cell progenitors, serial replating assays in methylcellulose were performed. No evidence of transformation were detected after two re-plating: cells transduced with the empty vector or the *ZEB2-BCL11B* chimera did not proliferate, while *MLL-AF9* transduced cells clearly showed the transformed phenotype.

Discussion

Chromosomal rearrangements and fusion gene products have a crucial diagnostic, prognostic and potentially therapeutic role in AML. Among 8 patients characterized by RNAseq, we detected the rare fusion transcript *ZEB2-BCL11B* (sample 59810) and its reciprocal *BCL11B-ZEB2,* together with point mutations in *TET2* and *FLT3*. In particular, the fusion protein *ZEB2-BCL11B,* which was previously described in AML[71] is formed by 24 residues of *ZEB2* and 803 out of 823 residues of *BCL11B*. To assess the oncogenic potential of the fusion, we transformed murine c-Kit+ cell with retroviral vector expressing the chimera and we performed serial replating assays, to test the ability of the transcripts to alter self-renewal. No evidence of transforming ability was detected in vitro, even though we confirmed transduced cells were expressing the fusion protein. The role of *BCL11B* as an oncogene in AML has been hypothesized, however functional studies on a myeloid cell line such as 32D cells resulted in a decrease proliferation[76]. Moreover, recent studies have shed light into the role of *ZEB2* in normal and malignant haematopoiesis[74,75], suggesting it may also play a role in the transformation. These studies suggested that a gain of function of *ZEB2* may have role, however, some data from the mouse phenotype and the nature of the mutations which are predicted to be inactivating also suggest that a loss of function may be crucial also. To this purpose, it would be of interest investigating the functional roles of the reciprocal *BCL11B-ZEB2* chimera. Similarly, *AML-ETO* (*RUNX1-RUNXT1*), which characterize 7% of AML[1], is not able on its own to induce leukemia in experimental in vitro and in vivo models, but requires additional mutations in other genes for the induction of haematological disease[48,60]. In order to identify co-occurring mutations in patients carrying the *ZEB2-BCL11B* fusions, we performed a mutational screen on 5 additional patients with a diagnosis of acute leukemia and characterized by the presence of t(2;14)(q-v;q32), in collaboration with the MLL Munich Leukemia Laboratory. Notably, 4 patients were positive for the *FLT3-ITD* internal tandem duplication, suggesting *FLT3* as a cooperating partner in leukemia induction (data not shown). Taken together our data demonstrated that the fusion *ZEB2-BCL11B* alone is not able to induce transformation in a murine model in vitro and alterations such as *FLT3-ITD* may be required for the leukemia development. The co-expression of these alterations in vitro and in vivo in murine or human hematopoietic progenitor cells, together with the investigation on the role of the *BCL11B-ZEB2* transcript, may elucidate the mechanism(s) of leukemogenesis promoted and investigate its potentiality as a target of therapies for this small subgroup of AML.

Discussion

The aim the thesis was to perform systematic deep whole exome/transcriptome studies on AML patients in order to identify novel biomarkers to improve outcomes for therapeutic interventions and to develop strategies to personalize treatments to different stratified groups of leukemia patients. Our results shed light into two different molecular aspects of AML, aneuploidy and fusion genes, highlighting inter-individual differences that may play a role in the leukemia development and may be markers of differential success of therapeutic interventions.

Aneuploidy, whole-chromosomes gains or losses, can exert an anti-tumorigenic or a pro-tumorigenic effect and, despite the detrimental effect of chromosome number alterations on cellular fitness, it is a hallmark of cancer. Aneuploidy in AML is generally associated with adverse prognosis, with monosomies (e.g. chromosome 5 and 7 losses) and monosomal karyotypes having poorer outcome[103] and isolated trisomies (e.g. trisomy 13) usually associating with adverse prognosis[104,105]. Hence, the clinical need of a deep molecular characterization that elucidates the mechanisms that allow aneuploid cells to survive and proliferate, in order find targets that may be exploited as target of therapy. We selected a cohort of 42 aneuploid and 35 euploid AML patients to be characterized. We detected a lower white blood cell count in our A-AML cohort compared with the E-AML (p=0.038), suggesting that other molecular and cellular mechanisms may be at play to tolerate the aberrant number of chromosomes and overcome the reduced fitness of aneuploid cells. Therefore, to detect and investigate alterations that contribute in the anuploid phenotype, we integrated and compared WES, CNA and GEP data from the two cohorts. Aneuploidy associates with genomic instability in AML, reflected also by the average number of coding mutations of 26 and 15 per sample in A-AML and E-AML, respectively. We detected alterations in genes involved in DNA repair and cell cycle phases, which were preferentially associated to the A-AML cohort. In particular, gene expression profiling revealed a 3-gene signature defined by *PLK1* and *CDC20* upregulation and *RAD50* downregulation, which characterize A-AML. Notably, a functional and structural silencing of the p53-transcriptional program was detected in A-AML patients. This cohort was also enriched in alterations in the protein ubiquitination and degradation pathway, response to reactive oxygen species, energy metabolism and biosynthetic process, processes that may help facing the unbalanced protein load deriving from the aneuploidy phenotype[12,22]. Taken together our results indicates that AML patients characterized by an aneuploidy phenotype have alterations in several cellular processes including the cell cycle control, metabolism and protein degradation, that may be target of *ad hoc* therapeutic strategies. Synthetic lethal approaches with microtubule depolymerizing agents and PLK1 inhibitors, APC/C inhibitors and topoisomerase[106] or defective sister chromatid cohesion inhibition[37] and the targeting of altered DNA repair pathways in combination with mutagenic chemotherapies[107,108] are some of candidate strategies that may be evaluated.

We characterized by RNA sequencing the transcriptome of 8 AML patients with rare or poorly described chromosomal translocations, with the aim of identifying novel and rare fusion transcripts. We validated 5 novel and rare fusion genes associated with the translocations (two isoform of *ZEB2-BCL11B* and its reciprocal *BCL11B-ZEB2*, *CNOT2-WT1* and *CPD-PXT1*), and 4 fusions which remained cryptic at cytogenetic analysis, including 2 fusions that were classified as driver events by either Pegasus or Oncofuse (*SAV1-GYPB* and *OAZ1-MAFK*) and additional 2 derived from a manual curation (*UTP6-CRLF3*, *PUF60-TYW1*;). The genesis of these cryptic fusions may be due to rare structural aberrations or post-transcriptional events. Translocated partner genes were transcription factors (*ZEB2*, *BCL11B* and *MAFK*) or tumor suppressors (*SAV1* and *PUF60*) rarely altered across cancer types. Moreover, we detected cryptic events hiding the loss of *NF1* and *WT1*, two recurrently altered genes in AML[2,7,83]. We exploited the potentiality of the novel tool FuGePrior, which combined results from state of the art bioinformatics tools for chimeric transcript identification and prioritization. The tool has been implemented with several filtering and processing steps in order to consider up-to-date literature on gene fusions and the analysis on the functional reliability of gene fusion structure. We demonstrated that the landscape of alterations in AML is not limited to known alterations, and fusion genes, though rare, may play an important role in the disease development and progression.

Among fusion transcripts identified by RNAseq, we detected the rare fusion transcript *ZEB2-BCL11B*[71] and its reciprocal *BCL11B-ZEB2*, which are in-frame fusions associated with t(2;14)(q21;q32). The translocation and the *ZEB2-BCL11B* chimera have been previously described in haematological malignancies[20,71] and, aiming to investigate its leukemogenic functionand explore the potentiality as therapeutic target, we tested the ability of the *ZEB2-BCL11B* transcript to alter self-renewal of murine hematopoietic precursor (c-Kit+ cells). Colony-forming unit assays returned no evidence of transforming ability *in vitro in* cells that were efficiently expressing the transcript and the fusion protein. The key roles of *ZEB2* and *BCL11B* in haematological malignancies and, in particular, in the myeloid lineage have been described[94,101]. Our results suggested that, although the fusion *ZEB2-BCL11B* alone is not able to induce transformation in murine cells in vitro, it may cooperate with co-occurring alterations (such as *FLT3-ITD*). We hypothesized a mechanism in which the expression of the chimera leads to a novel transcriptional program. *BLC11B* have been described to be involved in the differentiation of T cells[97], therefore its aberrant expression in myeloid precursors may also lead to the expression of T-cell markers in leukemic cells. Moreover, the contemporary loss of function of *ZEB2*, which blocks myeloid differentiation[74] and the co-occurrence of proliferative stimuli given by *FLT3* alterations, drive uncontrolled proliferation of undifferentiated cells. Future *in vitro* and *in vivo* studies on genetically engineered murine and human hematopoietic progenitor cells may elucidate the leukemogenic mechanisms promoted or enhanced by the fusion and test its potentiality as a target of therapies for this small subgroup of AML.

In conclusion we analysed the landscape of alterations that characterize two classes of AML patients. We found that aneuploidy-related alterations cooperate with known AML mutations, allowing cells to tolerate chromosome number imbalances and point to the DNA damage, mitotic and protein degradation machineries as potential therapeutic targets for synthetic lethal strategies. Moreover, fusion genes, even if rare, contribute to the heterogeneity of AML and may play a role in leukemia development and progression. Functional studies will further elucidate their potentials as pathogenic alterations and target of therapies. Taken together, the study highlighted the relevance of a multi-layer approach and the need of dissecting the molecular landscape of AML patients, in order to identify biomarkers within each molecular subgroup. This will allow a better stratification of patients and may enable the prediction of the clinical response to a specific targeted therapy.

Bibliography

1. Döhner, H. *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* **129,** 424–447 (2017).
2. Cancer Genome Atlas Research Network *et al.* Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368,** 2059–74 (2013).
3. Chen, S.-J., Shen, Y. & Chen, Z. A panoramic view of acute myeloid leukemia. *Nat. Genet.* **45,** 586–587 (2013).
4. Döhner, H., Weisdorf, D. J. & Bloomfield, C. D. Acute Myeloid Leukemia. *N. Engl. J. Med.* **373,** 1136–1152 (2015).
5. Genovese, G. *et al.* Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N. Engl. J. Med.* **371,** 2477–2487 (2014).
6. Jaiswal, S. *et al.* Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N. Engl. J. Med.* **371,** 2488–2498 (2014).
7. Papaemmanuil, E. *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N. Engl. J. Med.* **374,** 2209–2221 (2016).
8. Siegel, J. J. & Amon, A. New Insights into the Troubles of Aneuploidy. *Annu. Rev. Cell Dev. Biol.* **28,** 189–214 (2012).
9. Jones, K. T. Meiosis in oocytes: predisposition to aneuploidy and its increased incidence with age. *Hum. Reprod. Update* **14,** 143–158 (2008).
10. Pfau, S. J., Silberman, R. E., Knouse, K. A. & Amon, A. Aneuploidy impairs hematopoietic stem cell fitness and is selected against in regenerating tissues in vivo. *Genes Dev.* **30,** 1395–1408 (2016).
11. Schvartzman, J.-M., Sotillo, R. & Benezra, R. Mitotic chromosomal instability and cancer: mouse modelling of the human disease. *Nat. Rev. Cancer* **10,** 102–15 (2010).
12. Sheltzer, J. M., Torres, E. M., Dunham, M. J. & Amon, A. Transcriptional consequences of aneuploidy. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 12644–9 (2012).
13. Torres, E. M. *et al.* Effects of Aneuploidy on Cellular Physiology and Cell Division in Haploid Yeast. *Science (80-.).* **317,** 916–924 (2007).
14. Upender, M. B. *et al.* Chromosome Transfer Induced Aneuploidy Results in Complex Dysregulation of the Cellular Transcriptome in Immortalized and Cancer Cells. *Cancer Res.* **64,** 6941–6949 (2004).
15. Williams, B. R. *et al.* Aneuploidy Affects Proliferation and Spontaneous Immortalization in Mammalian Cells. *Science (80-.).* **322,** 703–709 (2008).

16. Gorrini, C., Harris, I. S. & Mak, T. W. Modulation of oxidative stress as an anticancer strategy. *Nat. Rev. Drug Discov.* **12,** 931–947 (2013).
17. Knouse, K. A., Wu, J., Whittaker, C. A. & Amon, A. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc. Natl. Acad. Sci.* **111,** 13409–13414 (2014).
18. Sheltzer, J. M. *et al.* Single-chromosome Gains Commonly Function as Tumor Suppressors. *Cancer Cell* **31,** 240–255 (2017).
19. Yang, Q., Rasmussen, S. A. & Friedman, J. M. Mortality associated with Down's syndrome in the USA from 1983 to 1997: a population-based study. *Lancet (London, England)* **359,** 1019–25 (2002).
20. Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2017). Mitelman F, Johansson B and Mertens F (Eds.), http://cgap.nci.nih.gov/Chromosomes/Mitelman".
21. Sheltzer, J. M. *et al.* Aneuploidy Drives Genomic Instability in Yeast. *Science (80-.).* **333,** 1026–1030 (2011).
22. Torres, E. M. *et al.* Identification of Aneuploidy-Tolerating Mutations. *Cell* **143,** 71–83 (2010).
23. Passerini, V. *et al.* The presence of extra chromosomes leads to genomic instability. *Nat. Commun.* **7,** 10754 (2016).
24. Schvartzman, J.-M., Duijf, P. H. G., Sotillo, R., Coker, C. & Benezra, R. Mad2 Is a Critical Mediator of the Chromosome Instability Observed upon Rb and p53 Pathway Inhibition. *Cancer Cell* **19,** 701–714 (2011).
25. Sotillo, R., Schvartzman, J.-M., Socci, N. D. & Benezra, R. Mad2-induced chromosome instability leads to lung tumour relapse after oncogene withdrawal. *Nature* **464,** 436–440 (2010).
26. Lamm, N. *et al.* Genomic Instability in Human Pluripotent Stem Cells Arises from Replicative Stress and Chromosome Condensation Defects. *Cell Stem Cell* **18,** 253–261 (2016).
27. Li, M. *et al.* The ATM-p53 pathway suppresses aneuploidy-induced tumorigenesis. *Proc. Natl. Acad. Sci.* **107,** 14188–14193 (2010).
28. Rutledge, S. D. *et al.* Selective advantage of trisomic human cells cultured in non-standard conditions. *Sci. Rep.* **6,** 22828 (2016).
29. Millet, C., Ausiannikava, D., Le Bihan, T., Granneman, S. & Makovets, S. Cell populations can use aneuploidy to survive telomerase insufficiency. *Nat. Commun.* **6,**

8664 (2015).

30. Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science (80-.).* **355,** eaaf8399 (2017).

31. Zack, T. I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45,** (2013).

32. Davoli, T. *et al.* Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155,** 948–62 (2013).

33. Pavelka, N. *et al.* Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468,** 321–5 (2010).

34. Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N. & Szallasi, Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet* **38,** 1043–1048 (2006).

35. Birkbak, N. J. *et al.* Paradoxical Relationship between Chromosomal Instability and Survival Outcome in Cancer. *Cancer Res.* **71,** 3447–3452 (2011).

36. Hunger, S. P. & Mullighan, C. G. Redefining ALL classification: toward detecting high-risk ALL and implementing precision medicine. *Blood* **125,** 3977–87 (2015).

37. de Lange, J. *et al.* Defective sister chromatid cohesion is synthetically lethal with impaired APC/C function. *Nat. Commun.* **6,** 8399 (2015).

38. Salesse, S. & Verfaillie, C. M. BCR/ABL: from molecular mechanisms of leukemia induction to treatment of chronic myelogenous leukemia. *Oncogene* **21,** 8547–8559 (2002).

39. Soverini, S., De Benedittis, C., Mancini, M. & Martinelli, G. Best Practices in Chronic Myeloid Leukemia Monitoring and Management. *Oncologist* **21,** 626–33 (2016).

40. Martens, J. H. A. & Stunnenberg, H. G. The molecular signature of oncofusion proteins in acute myeloid leukemia. *FEBS Lett.* **584,** 2662–2669 (2010).

41. Zeisig, B. B. *et al.* Recruitment of RXR by Homotetrameric RARα Fusion Proteins Is Essential for Transformation. *Cancer Cell* **12,** 36–51 (2007).

42. Zhu, J. *et al.* RXR Is an Essential Component of the Oncogenic PML/RARA Complex In Vivo. *Cancer Cell* **12,** 23–35 (2007).

43. Higuchi, M. *et al.* Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell* **1,** 63–74 (2002).

44. De Braekeleer, E. *et al.* *RUNX1* translocations and fusion genes in malignant

hemopathies. *Futur. Oncol.* **7,** 77–91 (2011).
45. Sood, R., Kamikubo, Y. & Liu, P. Role of RUNX1 in hematological malignancies Running Title: RUNX1 in leukemia. **129,** 2070–2083 (2017).
46. Liu, P. P. *et al.* The fusion gene Cbfb-*MYH11* blocks myeloid differentiation and predisposes mice to acute myelomonocytic leukaemia. *Nat. Genet.* **23,** 144–146 (1999).
47. Zhao, S. *et al.* KRAS (G12D) Cooperates with AML1/ETO to Initiate a Mouse Model Mimicking Human Acute Myeloid Leukemia. *Cell. Physiol. Biochem.* **33,** 78–87 (2014).
48. Schessl, C. *et al.* The AML1-ETO fusion gene and the FLT3 length mutation collaborate in inducing acute leukemia in mice. *J. Clin. Invest.* **115,** 2159–68 (2005).
49. Castilla, L. H. *et al.* Identification of genes that synergize with Cbfb-MYH11 in the pathogenesis of acute myeloid leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 4924–9 (2004).
50. Bacher, U., Haferlach, T., Schoch, C., Kern, W. & Schnittger, S. Implications of NRAS mutations in AML: a study of 2502 patients. *Blood* **107,** 3847–3853 (2006).
51. Ageberg, M., Drott, K., Olofsson, T., Gullberg, U. & Lindmark, A. Identification of a novel and myeloid specific role of the leukemia-associated fusion protein DEK-NUP214 leading to increased protein synthesis. *Genes, Chromosom. Cancer* **47,** 276–287 (2008).
52. Sandén, C., Ageberg, M., Petersson, J., Lennartsson, A. & Gullberg, U. Forced expression of the DEK-NUP214 fusion protein promotes proliferation dependent on upregulation of mTOR. *BMC Cancer* **13,** 440 (2013).
53. Oancea, C., Rüster, B., Henschler, R., Puccetti, E. & Ruthardt, M. The t(6;9) associated DEK/CAN fusion protein targets a population of long-term repopulating hematopoietic stem cells for leukemogenic transformation. *Leukemia* **24,** 1910–1919 (2010).
54. Lugthart, S. *et al.* High EVI1 levels predict adverse outcome in acute myeloid leukemia: prevalence of EVI1 overexpression and chromosome 3q26 abnormalities underestimated. *Blood* **111,** 4329–4337 (2008).
55. Yamazaki, H. *et al.* A Remote GATA2 Hematopoietic Enhancer Drives Leukemogenesis in inv(3)(q21;q26) by Activating EVI1 Expression. (2014). doi:10.1016/j.ccr.2014.02.008
56. Jude, C. D. *et al.* Unique and independent roles for MLL in adult hematopoietic stem

cells and progenitors. *Cell Stem Cell* **1,** 324–37 (2007).
57. Meyer, C. *et al.* The MLL recombinome of acute leukemias in 2017. *Leukemia* (2017). doi:10.1038/leu.2017.213
58. Somervaille, T. C. P. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10,** 257–268 (2006).
59. Bullinger, L. & Valk, P. J. M. Gene expression profiling in acute myeloid leukemia. *J. Clin. Oncol.* **23,** 6296–305 (2005).
60. Fenske, T. S. *et al.* Stem cell expression of the AML1/ETO fusion protein induces a myeloproliferative disorder in mice. *Proc. Natl. Acad. Sci.* **101,** 15184–15189 (2004).
61. Yoshihara, K. *et al.* The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene* **34,** 4845–4854 (2015).
62. Latysheva, N. S. & Babu, M. M. Discovering and understanding oncogenic gene fusions through data intensive computational approaches. *Nucleic Acids Res.* **44,** 4487–4503 (2016).
63. Togni, M. *et al.* Identification of the NUP98-PHF23 fusion gene in pediatric cytogenetically normal acute myeloid leukemia by whole-transcriptome sequencing. *J. Hematol. Oncol.* **8,** 69 (2015).
64. Iacobucci, I. *et al.* Truncating Erythropoietin Receptor Rearrangements in Acute Lymphoblastic Leukemia. *Cancer Cell* **29,** 186–200 (2016).
65. Zuffa, E. *et al.* Revealing very small FLT3 ITD mutated clones by ultra-deep sequencing analysis has important clinical implications in AML patients. *Oncotarget* **6,** 31284–31294 (2015).
66. International Standing Committee on Human Cytogenomic Nomenclature, McGowan-Jordan, J., Simons, A. & Schmid, M. (Michael). *ISCN : an international system for human cytogenomic nomenclature (2016).*
67. Iyer, M. K., Chinnaiyan, A. M. & Maher, C. A. ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics* **27,** 2903–2904 (2011).
68. McPherson, A. *et al.* deFuse: An Algorithm for Gene Fusion Discovery in Tumor RNA-Seq Data. *PLoS Comput. Biol.* **7,** e1001138 (2011).
69. Abate, F. *et al.* Pegasus: a comprehensive annotation and prediction tool for detection of driver gene fusions in cancer. *BMC Syst. Biol.* **8,** 97 (2014).
70. Shugay, M., De Mend??bil, I. O., Vizmanos, J. L. & Novo, F. J. Oncofuse: A computational framework for the prediction of the oncogenic potential of gene fusions.

Bioinformatics **29,** 2539–2546 (2013).

71. Torkildsen, S. *et al.* Novel ZEB2-BCL11B Fusion Gene Identified by RNA-Sequencing in Acute Myeloid Leukemia with t(2;14)(q22;q32). *PLoS One* **10,** e0132736 (2015).
72. Wakabayashi, Y. *et al.* Bcl11b is required for differentiation and survival of αβ T lymphocytes. *Nat. Immunol.* **4,** 533–539 (2003).
73. Li, L., Leid, M. & Rothenberg, E. V. An Early T Cell Lineage Commitment Checkpoint Dependent on the Transcription Factor Bcl11b. *Science (80-.).* **329,** 89–93 (2010).
74. Li, J. *et al.* The EMT transcription factor Zeb2 controls adult murine hematopoietic differentiation by regulating cytokine signaling. *Blood* **129,** (2017).
75. Li, H. *et al.* The EMT regulator ZEB2 is a novel dependency of human and murine acute myeloid leukemia. *Blood* **129,** (2017).
76. Abbas, S. *et al.* Integrated genome-wide genotyping and gene expression profiling reveals BCL11B as a putative oncogene in acute myeloid leukemia with 14q32 aberrations. *Haematologica* **99,** 848–57 (2014).
77. Wu, H.-Y. *et al.* Structural basis of antizyme-mediated regulation of polyamine homeostasis. *Proc. Natl. Acad. Sci.* **112,** 11229–11234 (2015).
78. Katsuoka, F. & Yamamoto, M. Small Maf proteins (MafF, MafG, MafK): History, structure and function. *Gene* **586,** 197–205 (2016).
79. Luo, X. *et al.* The human WW45 protein enhances MST1-mediated apoptosis in vivo. *Int. J. Mol. Med.* **23,** 357–62 (2009).
80. Matsuura, K. *et al.* Downregulation of SAV1 plays a role in pathogenesis of high-grade clear cell renal cell carcinoma. *BMC Cancer* **11,** 523 (2011).
81. Mardin, B. R. *et al.* Components of the Hippo pathway cooperate with Nek2 kinase to regulate centrosome disjunction. *Nat. Cell Biol.* **12,** 1166–1176 (2010).
82. Callus, B. A., Verhagen, A. M. & Vaux, D. L. Association of mammalian sterile twenty kinases, Mst1 and Mst2, with hSalvador via C-terminal coiled-coil domains, leads to its stabilization and phosphorylation. *FEBS J.* **273,** 4264–4276 (2006).
83. Haferlach, C. *et al.* Deletion of the tumor-suppressor gene NF1 occurs in 5% of myeloid malignancies and is accompanied by a mutation in the remaining allele in half of the cases. *Leukemia* **26,** 834–839 (2012).
84. Jayne, S., Zwartjes, C. G. M., Van Schaik, F. M. A. & Timmers, H. T. M. Involvement of the SMRT/NCoR–HDAC3 complex in transcriptional repression by the CNOT2

subunit of the human Ccr4–Not complex. *Biochem. J.* **398,** 461–467 (2006).

85. Zwartjes, C. G. M., Jayne, S., van den Berg, D. L. C. & Timmers, H. T. M. Repression of Promoter Activity by CNOT2, a Subunit of the Transcription Regulatory Ccr4-Not Complex. *J. Biol. Chem.* **279,** 10848–10854 (2004).

86. Ito, K. *et al.* CNOT2 depletion disrupts and inhibits the CCR4-NOT deadenylase complex and induces apoptotic cell death. *Genes to Cells* **16,** 368–379 (2011).

87. The Cancer Genome Atlas Research Network. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N. Engl. J. Med.* **368,** 2059–2074 (2013).

88. Yang, F. *et al.* Cloning and characterization of a novel intracellular protein p48.2 that negatively regulates cell cycle progression. *Int. J. Biochem. Cell Biol.* **41,** 2240–50 (2009).

89. Hastings, M. L., Allemand, E., Duelli, D. M., Myers, M. P. & Krainer, A. R. Control of Pre-mRNA Splicing by the General Splicing Factors PUF60 and U2AF65. *PLoS One* **2,** e538 (2007).

90. Matsushita, K. *et al.* Haploinsufficiency of the <i>c-myc</i> transcriptional repressor <i>FIR</i>, as a dominant negative-alternative splicing model, promoted p53-dependent T-cell acute lymphoblastic leukemia progression by activating Notch1. *Oncotarget* **6,** 5102–5117 (2015).

91. Kumar, S., Vo, A. D., Qin, F. & Li, H. Comparative assessment of methods for the fusion transcripts detection from RNA-Seq data. *Sci. Rep.* **6,** 1–10 (2016).

92. Gingeras, T. R. Implications of chimaeric non-co-linear transcripts. *Nature* **461,** 206–211 (2009).

93. Rickman, D. S. *et al.* SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. *Cancer Res.* **69,** 2734–8 (2009).

94. Huang, X., Du, X. & Li, Y. The role of BCL11B in hematological malignancy. *Exp. Hematol. Oncol.* **1,** 22 (2012).

95. Lin, C., Yang, L. & Rosenfeld, M. G. *Molecular Logic Underlying Chromosomal Translocations, Random or Non-Random? Advances in Cancer Research* **113,** (Elsevier Inc., 2012).

96. Shugay, M., Ortiz de Mendíbil, I., Vizmanos, J. L. & Novo, F. J. Genomic Hallmarks of Genes Involved in Chromosomal Translocations in Hematological Cancer. *PLoS Comput. Biol.* **8,** e1002797 (2012).

97. Ha, V. L. *et al.* The T-ALL related gene BCL11B regulates the initial stages of human

T-cell differentiation. *Leukemia* (2017). doi:10.1038/leu.2017.70

98. Hayashi, Y. *et al.* 14q32 translocations are associated with mixed-lineage expression in childhood acute leukemia. *Blood* **76,** 150–6 (1990).

99. Georgy, M., Yonescu, R., Griffin, C. A. & Batista, D. A. S. Acute mixed lineage leukemia and a t(6;14)(q25;q32) in two adults. *Cancer Genet. Cytogenet.* **185,** 28–31 (2008).

100. Hill, L., Browne, G. & Tulchinsky, E. ZEB/miR-200 feedback loop: At the crossroads of signal transduction in cancer. *Int. J. Cancer* **132,** 745–754 (2013).

101. Meyer, S. E. From EMT to HSC to AML: ZEB2 is a cell fate switch. *Blood* **129,** 400–401 (2017).

102. Giotopoulos, G. *et al.* A novel mouse model identifies cooperating mutations and therapeutic targets critical for chronic myeloid leukemia progression. *J. Exp. Med.* **212,** 1551–1569 (2015).

103. Breems, D. A. *et al.* Monosomal Karyotype in Acute Myeloid Leukemia: A Better Indicator of Poor Prognosis Than a Complex Karyotype. *J. Clin. Oncol.* **26,** 4791–4797 (2008).

104. Grimwade, D. *et al.* Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood* **116,** 354–365 (2010).

105. Farag, S. S. *et al.* Isolated trisomy of chromosomes 8, 11, 13 and 21 is an adverse prognostic factor in adults with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B 8461. *Int. J. Oncol.* **21,** 1041–51 (2002).

106. Eguren, M. *et al.* A synthetic lethal interaction between APC/C and topoisomerase poisons uncovered by proteomic screens. *Cell Rep.* **6,** 670–83 (2014).

107. Al-Ahmadie, H. *et al.* Synthetic lethality in ATM-deficient RAD50-mutant tumors underlies outlier response to cancer therapy. *Cancer Discov.* **4,** 1014–21 (2014).

108. Lord, C. J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science (80-.).* **355,** 1152–1158 (2017).