Acute Myeloid Leukemia with Del(9q) is Characterized by Frequent Mutations of *NPM1*, *DNMT3A, WT1* and Low Expression of *TLE4*

Tobias Herold,[1,2,3,4,*] Klaus H. Metzeler,[1,2,3,4] Sebastian Vosberg,[1,2,3,4] Luise Hartmann,[1,2,3,4] Vindi Jurinovic,[5] Sabrina Opatz,[1,2,3,4] Nikola P. Konstandin, [1] Stephanie Schneider,[1] Evelyn Zellmeier,[1] Bianka Ksienzyk,[1] Alexander Graf,[7] Stefan Krebs,[7] Helmut Blum,[7] Maria Cristina Sauerland,[8] Thomas Büchner,[9] Wolfgang E. Berdel,[9] Bernhard J. Wörmann,[10] Ulrich Mansmann,[3,4,5] Wolfgang Hiddemann,[1,2,3,4] Stefan K. Bohlander,[11] Karsten Spiekermann[1,2,3,4] and Philipp A. Greif[1,2,3,4]

[1]Department of Internal Medicine 3, University Hospital Grosshadern, Ludwig-Maximilians-Universität (LMU) München, München, Germany
[2]Clinical Cooperative Group Leukemia, Helmholtz Center Munich for Environmental Health, München, Germany
[3]German Cancer Consortium (DKTK), Heidelberg, Germany
[4]German Cancer Research Center (DKFZ), Heidelberg, Germany
[5]Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität (LMU) München, München, Germany
[6]Center for Human Genetics, Philipps University, Marburg, Germany
[7]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität (LMU) München, München, Germany
[8]Institute of Biostatistics and Clinical Research, and [9]Department of Medicine A-Hematology, Oncology and Pneumology, University of Münster, Münster, Germany
[10] Department of Hematology, Oncology and Tumor Immunology, Charité University Medicine, Campus Virchow, Berlin, Germany
[11]Department of Molecular Medicine and Pathology, The University of Auckland, Auckland, New Zealand

* Correspondence to: Tobias Herold, MD, Marchioninistr. 15, 81377 München, Phone: +49 89 4400 75833, FAX: +49 89 4400-72221, tobias.herold@med.uni-muenchen.de

Running title: ***AML del(9q) mutation profile***

Supported by: The German Cancer Aid (grant 109031) to P.A.G. and S.K.B., the Wilhelm-Sander-Stiftung (grant 2014.162.1) to P.A.G. and H.B., start-up funding from the Ludwig-Maximilians-Universität to T.H. and K.H.M. (FöFoLe 798/774 and 783) and the Wilhelm-Sander-Stiftung (grant 2013.086.1) to T.H., K.S. and U.M.. S.K.B. is supported by Leukaemia & Blood Cancer New Zealand and the family of Marijanna Kumerich.

Deletions of the long arm of chromosome 9 [del(9q)] are a rare but recurring aberration in acute myeloid leukemia (AML). Del(9q) can be found as the sole abnormality or in combination with other cytogenetic aberrations such as t(8;21) and t(15;17). *TLE1* and *TLE4* were identified to be critical genes contained in the 9q region. We performed whole exome sequencing of 5 patients with del(9q) as the sole abnormality followed by targeted amplicon sequencing of 137 genes of 26 patients with del(9q) as sole or combined with other aberrations. We detected frequent mutations in *NPM1* (10/26 38%), *DNMT3A* (8/26; 31%) and *WT1* (8/26; 31%) but only few *FLT3*-ITDs (2/26; 8%). All mutations affecting *NPM1* and *DNMT3A* were exclusively identified in patients with del(9q) as the sole abnormality and were significantly more frequent compared to 111 patients classified as intermediate-II according to the European LeukemiaNet (10/14, 71% vs. 22/111, 20%; $P<0.001$, 8/14, 57% vs. 26/111, 23%; $P=0.02$). Furthermore, we identified *DNMT3B* to be rarely but recurrently targeted by truncating mutations in AML. Gene expression analysis of 13 patients with del(9q) and 454 patients with normal karyotype or various cytogenetic aberrations showed significant down regulation of *TLE4* in patients with del(9q) ($P=0.02$). Interestingly, downregulation of *TLE4* was not limited to AML with del(9q), potentially representing a common mechanism in AML pathogenesis. Our comprehensive genetic analysis of the del(9q) subgroup reveals a unique mutational profile with the frequency of *DNMT3A* mutations in the del(9q) only subset being the highest reported so far in AML, indicating oncogenic cooperativity.

INTRODUCTION

Acute myeloid leukemia (AML) is an extremely aggressive neoplasm of the bone marrow. Recent studies revealed distinct molecular subgroups of AML harboring defined combinations of somatic mutations. We and others found striking associations of gene mutations within specific AML subgroups (Wechsler et al., 2002; Dicker et al., 2007; Greif et al., 2012; Herold et al., 2014). For example, patients with biallelic *CEBPA* mutations exhibit a high frequency of mutations in *GATA2* (Greif et al., 2012). More recently, we could demonstrate that patients with a gain of chromosome 13 as the sole abnormality are characterized by the co-occurrence of *RUNX1* and *SRSF2* mutations and a uniform gene expression profile (Herold et al., 2014). These subgroup analyses provide hope to identify mechanisms of oncogenic cooperativity that could serve as the basis for tailored therapies.

Deletions affecting the long arm of chromosome 9 [del(9q)] are rare in AML (Langabeer et al., 1998; Grimwade et al., 2010). The overall incidence was reported to be 2%, and del(9q) is considered as marker of intermediate risk according to the MRC classification (Dohner et al., 2010). An association between del(9q) and mutations in *CEBPA* was reported (Frohling et al., 2005). Del(9q) is significantly associated with t(8;21) and t(15;17), but can also be observed as a sole abnormality or in association with other cytogenetic aberrations (Langabeer et al., 1998; Dohner et al., 2010). Since the *RUNX1*/*RUNX1T1* fusion, which results from the t(8;21), by itself is not sufficient to cause leukemia, the commonly deleted region (CDR) of del(9q) was extensively studied to identify cooperating factors (Sweetser et al., 2005). Two genes closely related to the CDR (*TLE1* and *TLE4*) were identified to contribute to leukemogenesis due to haploinsufficiency (Dayyani et al., 2008). These genes were found to be expressed at lower levels also in cases of AML not harboring del(9q) (Dayyani et al.,

2008). Since del(9q) also occurs as the sole cytogenetic abnormality, we were interested in the mutational profile of this subgroup. We performed whole exome sequencing in 5 patients with del(9q) as the sole abnormality and subsequently screened 26 additional patients with del(9q) by targeted amplicon sequencing for mutations in 137 genes identified by exome sequencing in the del(9q) cases, recurrently mutated genes in AML and genes in the CDR of del(9q). In addition, we used a large gene expression data set and publically available copy number alteration, methylation and gene expression data to study the expression and regulation of *TLE1* and *TLE4*.

MATERIALS AND METHODS

Patients

We studied patients who were enrolled in the multicenter AMLCG-1999 trial of the German AML Cooperative Group (NCT00266136) (for details, see supplementary Fig. S1) (Buchner et al., 2006). The gene mutation and expression analysis included three additional patients that were registered in this trial but were not randomized or died before receiving first treatment. All other patients received intensive induction chemotherapy as described elsewhere (Buchner et al., 2006). The AMLCG clinical trial was approved by the local institutional review boards of all participating centers and informed consent was obtained from all patients in accordance with the Declaration of Helsinki.

Sequencing

To perform exome sequencing, genomic DNA of available paired diagnostic and remission samples from 5 patients with AML and del(9q) as the sole cytogenetic abnormality was extracted from archived bone marrow samples. Target enrichment,

sequencing and variant calling were performed as described previously (Greif et al., 2012; Opatz et al., 2013; Herold et al., 2014). The quality metrics are summarized in Supplementary Table S1.

A panel of 137 genes, which included selected genes identified by exome sequencing of the 5 del(9q) patients, genes encompassed in the minimal commonly deleted region of del(9q) and genes recurrently mutated in AML were studied by Haloplex targeted amplicon sequencing (Agilent, Santa Clara, CA) in all AMLCG-1999 AML del(9q) patients with available material (n=26). A detailed gene list is shown in the Supplementary Table S2. The libraries were sequenced on a MiSeq instrument, resulting in a mean target coverage of ~200x (for quality metrics, see Supplementary Table S3). Gene panel sequencing data were analyzed as described previously (Herold et al., 2014).

Sanger sequencing of PCR-amplified genomic DNA was performed for additional validation of selected mutations. Primer sequences and PCR conditions (for *GPR98*, *CELSR2*, *PPM1E*, *ASXL2*, *CEBPZ*, *WDR52* and *DNMT3B*) are given in the supplement (Table S4). PCR products were purified using NucleoFast® 96 PCR Clean-up Kit (Macherey Nagel, Düren, Germany) and both strands were sequenced on an ABI 3500xL Genetic Analyzer using the BigDye® Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems; Foster City, CA). Sequences were aligned and compared to the reference sequences (NCBI accession numbers in the Supplementary) using the Sequencher software (Gene Codes Corporation, Ann Arbor, MI).

Gene Expression Analysis

To characterize further the del(9q) subgroup, we compared gene expression profiles of 13 patients with del(9q) (n=6 solely, n=7 with t(8;21) or other additional

aberrations) to 454 AML patients with normal karyotype (CN-AML) or various cytogenetic abnormalities (except for deletion of 9q). The gene expression data set was published previously and is publicly available through the Gene Expression Omnibus (GEO) website (GSE37642) (Herold et al., 2014). Details of sample preparation, hybridization and image acquisition were described previously (Li et al., 2013). The Linear Models for Microarray Data (Limma) package was used to compute differentially expressed probe sets. Differential regional gene expression on chromosome 9 was analyzed using MACAT (MicroArray Chromosome Analysis Tool) as described previously (Toedling et al., 2005; Herold et al., 2011).

Analysis of Copy Number Alterations and Methylation

Copy number alteration (CNA) and methylation data sets of AML patients were downloaded from the TCGA website (http://cancergenome.nih.gov/) (TCGA 2013). The CNA analysis focused on the CDR region of del(9q) as reported previously and surrounding sequences including *TLE1* and *TLE4* (Sweetser et al., 2005). We used the segment mean value that represents the extent of CNA for a particular genomic segment. As threshold, a cut-off of 0.2 for amplifications and −0.2 for deletions was used as published previously (Laddha et al., 2014). For DNA-methylation analysis with focus on *TLE4*, beta values were compared using Spearman's test, and adjustment for multiple hypothesis testing was performed using the method described by Benjamini and Hochberg (Benjamini and Hochberg 1995).

Statistical Analyses

All statistical analyses were performed using the R 3.2.0 software (R Foundation for Statistical Computing, Vienna, Austria) and routines from the biostatistics software repository Bioconductor. Two-sided Fisher's exact test was used to compare categorical variables while the Wilcoxon-Mann-Whitney-Test was applied for

continuous variables. Complete remission (CR), relapse-free survival (RFS) and overall survival (OS) were defined as described previously (Cheson et al., 2003; Buchner et al., 2006). Patients alive without an event were censored at the time of their last follow-up. The prognostic impact of del(9q) was evaluated according to the Kaplan-Meier method and the log-rank test. A *P*-value of ≤0.05 was considered significant.

RESULTS

Patient Characteristics

Among 3,214 patients of the multicenter AMLCG-1999 study cohort with cytogenetic data available we identified 61 with del(9q) (incidence 1.9%). Detailed patient characteristics are shown in the Supplementary Table S5. In 23 cases del(9q) was observed as the sole abnormality, whereas 15 cases were associated with t(8;21) and 23 with other cytogenetic abnormalities, including 12 cases with a complex karyotype. According to the European Leukemia Net classification, 15 patients were classified as favorable, 33 as intermediate-II and 13 as adverse risk. Interestingly, relapse free (RFS) and overall survival (OS) showed no significant differences between these subgroups (Fig. 1A and B), although there was a trend for better relapse free survival of patients with a t(8;21). When we compared the subgroup with del(9q) as the sole abnormality to a control group of 362 patients treated in the same trial and classified as intermediate II without alterations of del(9q), only bone marrow blasts at day 16 showed a significant differences between the two groups (Table 1). There was no significant difference regarding RFS and OS (Fig. 1C and D).

High Frequency of Mutations in *NPM1, DNMT3A* and *WT1* in AML Del(9q)

We were able to identify 26 del(9q) patients with material available for genomic analysis. In five patients with del(9q) as the sole abnormality we performed whole

exome sequencing of paired diagnostic and remission samples (selected based on availability). Targeted amplicon sequencing of 137 genes, including genes found in the exome sequencing of the five patients, was performed in all 26 patients with del(9q) and available material. Details of all detected non-synonymous variants are shown in Figure 2 and supplementary Table S6. Detailed cytogenetic results are shown in Table S7. In total, we found 107 mutations in 48 genes. At least one mutation was found in each patient. On average, we detected 7 (4-14) non-synonymous somatic mutations per AML exome. There was a median of 3 (1-14) variants per patient in the whole cohort of del(9q) patients.

Genes found in at least three patients altered in the 26 patients with del(9q) were *NPM1* (n=10), *DNMT3A* (n=8), *FLT3* (n=7), *WT1* (n=8), *ASXL1*, *CEBPA*, *CELSR2*, *IDH1*, *NRAS*, and *PTPN11* (each n=3). We detected a high frequency of mutations in *WT1* distributed among del(9q) subgroups (n=8/26; 31%). Furthermore, we identified mutations in *NPM1* (10/26; 38%), *DNMT3A* (8/26; 31%), *PTPN11* (3/26; 12%), *GATA2*, *KALRN, NCAN, SLC28A1* and *TUBG1* (each 2/26; 8%) that were exclusively present in the subgroup of patients with del(9q) as the single cytogenetic alteration. The association of *NPM1* and *DNMT3A* mutations with this subgroup was significant among all patients with del(9q) (P<0.001 and P =0.002, respectively). In addition, we found five *CEBPA* mutations in three AML patients with del(9q) (3/26; 12%). In total, we identified eight alterations of *FLT3* in seven patients (*FLT3*-ITD n=2, *FLT3*-TKD n=2, and *FLT3*-other n=4 [p.N676 n=2, p.Y842C n=1 and p.T329N n=1]). *FLT3* p.N676 and p.Y842C mutations had previously been shown to have leukemogenic potential (Kindler et al., 2005; Opatz et al., 2013). Remarkably, only 2/26 (8%) of patients had a FLT3-ITD.

Furthermore, we compared the incidences of mutations in *NPM1*, *DNMT3A* and *FLT3*-ITD in the group of patients with del(9q) as the single alteration with 111 patients classified as intermediate-II without del(9q) (Metzeler et al., 2016). *NPM1* and *DNMT3A* mutations were significantly more frequent in patients with del(9q) as the sole abnormality (*NPM1*: 10/14, 71% vs. 22/111, 20%; *DNMT3A*: $P<0.001$; 8/14, 57% vs. 26/111, 23%; $P=0.02$), whereas *FLT3*-ITD was rare (1/14, 7% vs. 41/111, 37%; $P=0.03$). The whole group of del(9q) patients showed frequent mutations in *WT1* (8/26, 31% vs. 12/111, 11%; $P=0.03$) but significantly fewer *FLT3*-ITDs (2/26, 8% vs. 41/111, 37%; $P=0.004$).

Finally, we tested if *TLE4* expression was associated with *DNMT3A* mutations in the data set GSE37642. In total, 381 patients (n=108 with *DNMT3A* mutation) could be analyzed (Metzeler et al., 2016). Interestingly, there was no significant association between *DNMT3A* mutations and *TLE4* expression when we tested the whole group ($P=0.48$) and the 10% of samples with the lowest *TLE4* expression ($P=0.84$).

DNMT3B is Rarely but Recurrently Mutated in AML

Whole exome sequencing identified one AML del(9q) patient with a truncating somatic mutation in *DNMT3B* (p.I622fs, variant allele frequency [VAF] 27%). In this patient there was no evidence for genetic lesions of *DNMT3A*, suggesting that alteration of *DNMT3B* may represent an alternative route to perturbed DNA methylation in AML. In the TCGA-cohort, a single missense mutation of *DNMT3B* (p.R538C) as well as a single out of frame fusion (*MDM4*/*DNMT3B*) have been reported (TCGA 2013). These findings prompted us to look for *DNMT3B* mutations in gene panel sequencing data from ongoing projects in our research group. Indeed, we found another truncating *DNMT3B* (p.R571X, VAF 42%) mutation in a patient from a cohort of 56 AML t(8;21) patients. This patient was not found to carry an additional

del(9q) alteration. Both truncating *DNMT3B* mutations (p.I622fs and p.R571X) identified in our AML patients were confirmed by Sanger sequencing of both DNA strands (Fig. S2).

_WT1_ Mutational Status and Survival

We used univariate Cox regression analysis to identify prognostically relevant biomarkers. We included only the most common alterations (*NPM1*, *DNMT3A* and *WT1)* into the analysis. Only *WT1* showed a significant prognostic relevance. Within the cohort of 26 AML patients with del(9q) analyzed for *WT1* mutational status, 23 patients were treated on the AMLCG-1999 trial and included in further analysis. Six patients had mutations in *WT1* and 17 patients had wild type status. Within the group of patients with del(9q), there was no difference in RFS depending on the *WT1* mutational status; however, a mutation in *WT1* was associated with a more favorable OS (P=0.05; Fig. S3). As reported before (Gaidzik et al., 2009; Hou et al., 2010; Krauth et al., 2015), *WT1* mutations were significantly associated with younger age (P<0.001). Five of six patients in the *WT1* mutated group reached a complete remission (CR), in contrast to only 5 of 17 patients with *WT1* wild type (P=0.05). The main reason for this difference was the higher early death rate in the *WT1* wild type group (53%), whereas no patient in the group with *WT1* mutations died in the first 60 days after diagnosis. In a multivariate model including *WT1* and age, only age was left as significant variable (data not shown). Therefore, the effect on survival by *WT1* reflects most probably the favorable impact of younger age and could furthermore be influenced by the very small patient cohort and seems not to be specifically associated with del(9q).

Expression of _TLE1_ and _TLE4_ in AML Subgroups

Expression of *TLE1* (location: 9q21.32) was not significantly different between patients with del(9q) as the sole abnormality (n=6; P=0.43) or in combination with other

aberrations (n=13; *P*=0.71) in comparison to a large group of AML samples (n=454) with different cytogenetic abnormalities or CN-AML (GSE37642, Fig. 3A and B). In contrast, *TLE4* (location: 9q21.31) showed a trend towards or significantly lower expression in del(9q) patients in this data set [del(9q) solely: *P*=0.06; any del(9q): *P*=0.02] (Fig. 3C and D). Interestingly, several AML samples without del(9q) also showed equally low *TLE4* expression (Fig. 3 C and D). When analyzing *TLE4* in the publically available TCGA microarray data set of 200 AML patients we also observed a weak trend (*P*=0.21) towards lower *TLE4* expression in the cohort with del(9q) and any additional cytogenetic alteration (Fig. 3F). Again, several samples without del(9q) showed low expression of *TLE4* (Fig. 3E).

Reduced *TLE4* Expression is not Related to Copy Number Alterations in the Absence of Del(9q)

To identify potential copy number alterations leading to reduced *TLE4* expression we used a publically available TCGA data set of 200 AML patients. We filtered for patients with deletions in the CDR of del(9q) and the surrounding regions including *TLE1* and *TLE4* (genomic location: GRCh37: chr9:79216292-87648505). Only two out of 200 patients showed a deletion of *TLE4* (and large parts of chromosome 9) (Fig. S4). Therefore, altered gene expression in samples that do not have a del(9q) does not seem to be related to copy number alteration.

DNA Methylation Status is not Associated with Reduced *TLE4* Expression in the Absence of Del(9q)

We used the TCGA data set to identify differences in CpG-island methylation of the *TLE4* promoter associated with *TLE4* expression. Sixteen methylation sites covering the gene and promoter region were evaluated. After adjustment for multiple hypothesis testing no significant hyper- or hypomethylation associated with the level of

TLE4 expression could be detected. The same result was observed when only samples with low *TLE4* expression (lowest 10%) were considered.

Potential Regulators of *TLE4* Expression in the Absence of Del(9q)

To assess if reduced expression of *TLE4* could be linked to the expression of specific transcription factors, we assessed the predicted transcription factor binding sites in the *TLE4* promotor region using the UCSC genome browser (Encode Txn Factor ChIP V1). We were able to identify several transcription factors whose expression was significantly associated with *TLE4* expression in our own gene expression set and the TCGA data set. A list of significantly associated transcription factors in both data sets is given in Table 2. Among others, the expression of *TCF7L2* and *SP1* were significantly associated with *TLE4* expression (Fig. 4) (Tickenbrock et al., 2005; Liu et al., 2010). In a subgroup analysis incorporating only patients with del(9q) four of the identified transcription factors (*WRNIP1*, *CHD2*, *NR3C1* and *E2F6*) were significant with $P<0.05$. However, potentially due to the small sample size, no factor was significant after adjustment for multiple testing.

Analysis of Differential and Regional Gene Expression of AML with Del(9q)

No probe set was significantly deregulated ($P\leq0.05$ after adjustment for multiple testing) in AML with del(9q) patients (n=13) when compared to AML patients with various other cytogenetic abnormalities (n=454). When we applied a non-adjusted *P*-value of ≤0.001 only 26 probe sets were significantly deregulated (16 down-regulated). Interestingly, 12/16 down regulated probe sets were located on chromosome 9, whereas no up-regulated probe set was located on this chromosome. No specific gene expression profile of AML with del(9q) was detected as shown in Figure S5.

The analysis of regional gene expression on chromosome 9 using the MACAT tool showed reduced expression of several genes on 9q (Fig. S6). Significantly deregulated genes of this region are shown in the Supplementary Table S8.

DISCUSSION

Our study describes, to our knowledge for the first time, a characteristic mutational pattern in AML with del(9q) as the sole cytogenetic abnormality. In this subgroup, we found mutations in *NPM1* and *DNMT3A* at very high frequencies. *NPM1* mutations are common in AML and reach a frequency of 60% in younger patients (Schneider et al., 2012). The incidence of *DNMT3A* in AML is 18% in unselected patients younger than 60 years of age (Thol et al., 2011) and is significantly higher (up to 35%) in sAML or patients with CN-AML (Fried et al., 2012; Marcucci et al., 2012). Our finding of a *DNMT3A* mutation frequency of 57% in AML with del(9q) as the sole abnormality is therefore quite remarkable and differs significantly from other patients classified as ELN intermediate-II. Furthermore, it is important to mention that *DNMT3A* and *NPM1* mutations were only found in the del(9q) sole subgroup and not in del(9q) patients with other cytogenetic changes. This association of *NPM*1 and *DNMT3A* mutation with deletion on the long arm of chromosome 9 suggests a synergistic effect in leukemogenesis. The co-occurence of *DNMT3A* and *NPM1* mutations was recently described and is frequently associated with alterations of *FLT3* (TCGA 2013). However, our study showed that *FLT3*-ITD is a rare event in patients with del(9q).

In patients with del(9q) and other cytogenetic alteration we found a high incidence of *WT1* mutations (31%). In the TCGA cohort, only 12 *WT1* mutations were found in 200 samples (incidence 6%) (TCGA 2013). This is in line with previous studies that showed

a frequency of *WT1* mutations of 6.8% to 12.6% in unselected AML patient cohorts (Gaidzik et al., 2009; Ho et al., 2010; Hou et al., 2010).

The prevalence of *CEBPA* mutations in cases with del(9q) and a noncomplex aberrant karyotype was 41% in previous reports (Frohling et al., 2005). However, in our analysis the incidence of *CEBPA* mutations in patients with non-complex karyotype and del(9q) was only 13%. All patients with a mutation in *CEBPA* (n=3) also carried a mutation in *WT1*. An association of *WT1* and *CEBPA* mutations was described previously (Gaidzik et al., 2009). To exclude the possibility that targeted amplicon sequencing might have missed *CEBPA* mutations, all patients were reanalyzed for *CEBPA* mutations using our standard diagnostic test (Benthaus et al., 2008). No additional mutations were detected (data not shown). A difference in the composition of data sets, with a high rate of del(9q) as the single abnormality in our cohort might account for this difference. Moreover, we found *ASXL2* frame shift mutations in two patients, one with del(9q) solely and the other one with both del(9q) and t(8;21). Recurring mutations in *ASXL2* have been described in AML with t(8;21) (Micol et al., 2014).

The truncating *DNMT3B* mutations in two of our AML patients, together with another two patients with genetic lesions of *DNMT3B* in the TCGA cohort (TCGA 2013), suggest that alteration of DNMT3B may disturb DNA methylation and that *DNMT3B* mutations are drivers of AML development. Expression of truncated DNMT3B protein was reported in several cancer cell lines of multiple tissue origins (Ostler et al., 2007; Shah et al., 2010; Ostler et al., 2012). Murine models show that *Dnmt3b* acts as a tumor suppressor during the development of leukemia and lymphoma (Vasanthakumar et al., 2013; Schulze et al., 2016). All in all, these findings support that rare but recurring mutations may impair the function of DNMT3B in AML.

It is challenging to compare clone size either based on cytogenetics or on variant allele frequencies (VAF) from targeted amplicon sequencing. Whereas sequencing is performed on amplified DNA reflecting the ratio of malignant and normal cells in a given sample, chromosomal analysis represents cells that grow in culture with only a limited number of cells routinely analyzed. However, we could recently demonstrate, that despite these considerations both techniques strongly correlate (Vosberg et al., 2016). In our cohort, several patients with *DNMT3A*, *NPM1* or *WT1* mutations had dominant del(9q) clones cytogenetically (Table S9 and Fig. S7). This argues for del(9q) being an early event in leukemogenesis.

The survival rate of patients with del(9q) as the sole abnormality showed no differences to other patients classified as intermediate risk II. The unexpected unfavorable and favorable outcome rates in the del(9q) subgroup with t(8;21) and for the del(9q) subgroup with complex cytogenetics, respectively, could be explained by a relatively high relapse and poor survival rate in the favorable group and the small patient cohort for a three-way comparison.

Our gene expression analysis could verify the reduced expression of *TLE4* in patients with del(9q) (Dayyani et al., 2008). However, a significant change in *TLE1* expression, which is located very close to *TLE4* on 9q, was not observed. For *TLE4,* some patients without cytogenetic alterations on chromosome arm 9q showed equally low expression. *TLE4* was shown to act as a tumor suppressor gene and to cooperate with t(8;21) in leukemogenesis (Dayyani et al., 2008). To identify an effect of low *TLE4* expression beyond the del(9q) subgroup we used the publicly available data set of the TCGA consortium. We were able to demonstrate that reduced expression of *TLE4* was not associated with CNA affecting 9q. In addition, no significant difference in hyper- or hypomethylation of the CpG-island located in the *TLE4* promoter region was detected

that could explain the reduced expression of *TLE4* in the absence of a deletion of 9q. In subsequent analyses, the expression of several transcription factors (e.g. *TCF7L2* and *SP1)* could be linked to *TLE4* expression. Both *TCF7L2* and *SP1* are known to be involved in pathways critical for AML development (Tickenbrock et al., 2005; Liu et al., 2010). These results led us to the hypothesis that the contributing effect of *TLE4* on leukemogenesis is not restricted to patients with a t(8;21) and del(9q) but is more widely found in AML (Dayyani et al., 2008).

Interestingly, we did not observe any mutation in genes located in the del(9q) region, but were able to demonstrate a significant impact of haplo-insufficiency of del(9q) (Fig. S6 and Table S8). Therefore, we hypothesize that in patients with del(9q) a gene dosage effect may contribute to leukemogenesis, whereas in patients without del(9q) and reduced TLE4 expression several transcription factors might play a similar role.

In summary, our study reveals a unique mutational profile in AML with del(9q). In particular, AML with del(9q) as the sole cytogenetic abnormality is a rather homogeneous entity as far as additional mutations are concerned. The striking associations of del(9q) with a distinct group of potentially cooperating mutations suggest several alternative but defined oncogenic pathways to leukemogenesis in del(9q) AML.

ACKNOWLEDGMENT

The authors thank all participants of the AMLCG trial and recruiting centers.

REFERENCES

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

Benthaus T, Schneider F, Mellert G, Zellmeier E, Schneider S, Kakadia PM, Hiddemann W, Bohlander SK, Feuring-Buske M, Braess J, Spiekermann K, Dufour A. 2008. Rapid and sensitive screening for CEBPA mutations in acute myeloid leukaemia. Br J Haematol 143:230-239.

Buchner T, Berdel WE, Schoch C, Haferlach T, Serve HL, Kienast J, Schnittger S, Kern W, Tchinda J, Reichle A, Lengfelder E, Staib P, Ludwig WD, Aul C, Eimermacher H, Balleisen L, Sauerland MC, Heinecke A, Wormann B, Hiddemann W. 2006. Double induction containing either two courses or one course of high-dose cytarabine plus mitoxantrone and postremission therapy by either autologous stem-cell transplantation or by prolonged maintenance for acute myeloid leukemia. J Clin Oncol 24:2480-2489.

Cheson BD, Bennett JM, Kopecky KJ, Buchner T, Willman CL, Estey EH, Schiffer CA, Doehner H, Tallman MS, Lister TA, Lo-Coco F, Willemze R, Biondi A, Hiddemann W, Larson RA, Lowenberg B, Sanz MA, Head DR, Ohno R, Bloomfield CD, International Working Group for Diagnosis SoRCTO, Reporting Standards for Therapeutic Trials in Acute Myeloid L. 2003. Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia. J Clin Oncol 21:4642-4649.

Dayyani F, Wang J, Yeh JR, Ahn EY, Tobey E, Zhang DE, Bernstein ID, Peterson RT, Sweetser DA. 2008. Loss of TLE1 and TLE4 from the del(9q) commonly deleted region in AML cooperates with AML1-ETO to affect myeloid cell proliferation and survival. Blood 111:4338-4347.

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

Dohner H, Estey EH, Amadori S, Appelbaum FR, Buchner T, Burnett AK, Dombret H, Fenaux P, Grimwade D, Larson RA, Lo-Coco F, Naoe T, Niederwieser D, Ossenkoppele GJ, Sanz MA, Sierra J, Tallman MS, Lowenberg B, Bloomfield CD. 2010. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood 115:453-474.

Fried I, Bodner C, Pichler MM, Lind K, Beham-Schmid C, Quehenberger F, Sperr WR, Linkesch W, Sill H, Wolfler A. 2012. Frequency, onset and clinical impact of somatic DNMT3A mutations in therapy-related and secondary acute myeloid leukemia. Haematologica 97:246-250.

Frohling S, Schlenk RF, Krauter J, Thiede C, Ehninger G, Haase D, Harder L, Kreitmeier S, Scholl C, Caligiuri MA, Bloomfield CD, Dohner H, Dohner K. 2005. Acute myeloid leukemia with deletion 9q within a noncomplex karyotype is associated with CEBPA loss-of-function mutations. Genes Chromosomes Cancer 42:427-432.

Gaidzik VI, Schlenk RF, Moschny S, Becker A, Bullinger L, Corbacioglu A, Krauter J, Schlegelberger B, Ganser A, Dohner H, Dohner K, German-Austrian AMLSG. 2009. Prognostic impact of WT1 mutations in cytogenetically normal acute myeloid leukemia: a study of the German-Austrian AML Study Group. Blood 113:4505-4511.

Greif PA, Dufour A, Konstandin NP, Ksienzyk B, Zellmeier E, Tizazu B, Sturm J, Benthaus T, Herold T, Yaghmaie M, Dorge P, Hopfner KP, Hauser A, Graf A, Krebs S, Blum H, Kakadia PM, Schneider S, Hoster E, Schneider F, Stanulla M, Braess J, Sauerland MC, Berdel WE, Buchner T, Woermann BJ, Hiddemann W, Spiekermann K, Bohlander SK. 2012. GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. Blood 120:395-403.

Grimwade D, Hills RK, Moorman AV, Walker H, Chatters S, Goldstone AH, Wheatley K, Harrison CJ, Burnett AK. 2010. 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.

Herold T, Jurinovic V, Mulaw M, Seiler T, Dufour A, Schneider S, Kakadia PM, Feuring-Buske M, Braess J, Spiekermann K, Mansmann U, Hiddemann W, Buske C, Bohlander SK. 2011. Expression analysis of genes located in the minimally deleted regions of 13q14 and 11q22-23 in chronic lymphocytic leukemia-unexpected expression pattern of the RHO GTPase activator ARHGAP20. Genes Chromosomes Cancer 50:546-558.

Herold T, Metzeler KH, Vosberg S, Hartmann L, Rollig C, Stolzel F, Schneider S, Hubmann M, Zellmeier E, Ksienzyk B, Jurinovic V, Pasalic Z, Kakadia PM, Dufour A, Graf A, Krebs S, Blum H, Sauerland MC, Buchner T, Berdel WE, Woermann BJ, Bornhauser M, Ehninger G, Mansmann U, Hiddemann W, Bohlander SK, Spiekermann K, Greif PA. 2014. Isolated trisomy 13 defines a homogeneous AML subgroup with high frequency of mutations in spliceosome genes and poor prognosis. Blood 124:1304-1311.

Ho PA, Zeng R, Alonzo TA, Gerbing RB, Miller KL, Pollard JA, Stirewalt DL, Heerema NA, Raimondi SC, Hirsch B, Franklin JL, Lange B, Meshinchi S. 2010. Prevalence and prognostic implications of WT1 mutations in pediatric acute myeloid leukemia (AML): a report from the Children's Oncology Group. Blood 116:702-710.

Hou HA, Huang TC, Lin LI, Liu CY, Chen CY, Chou WC, Tang JL, Tseng MH, Huang CF, Chiang YC, Lee FY, Liu MC, Yao M, Huang SY, Ko BS, Hsu SC,

Wu SJ, Tsay W, Chen YC, Tien HF. 2010. WT1 mutation in 470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system. Blood 115:5222-5231.

Kindler T, Breitenbuecher F, Kasper S, Estey E, Giles F, Feldman E, Ehninger G, Schiller G, Klimek V, Nimer SD, Gratwohl A, Choudhary CR, Mueller-Tidow C, Serve H, Gschaidmeier H, Cohen PS, Huber C, Fischer T. 2005. Identification of a novel activating mutation (Y842C) within the activation loop of FLT3 in patients with acute myeloid leukemia (AML). Blood 105:335-340.

Krauth MT, Alpermann T, Bacher U, Eder C, Dicker F, Ulke M, Kuznia S, Nadarajah N, Kern W, Haferlach C, Haferlach T, Schnittger S. 2015. WT1 mutations are secondary events in AML, show varying frequencies and impact on prognosis between genetic subgroups. Leukemia 29:660-667.

Laddha SV, Ganesan S, Chan CS, White E. 2014. Mutational landscape of the essential autophagy gene BECN1 in human cancers. Mol Cancer Res 12:485-490.

Langabeer SE, Grimwade D, Walker H, Rogers JR, Burnett AK, Goldstone AH, Linch DC. 1998. A study to determine whether trisomy 8, deleted 9q and trisomy 22 are markers of cryptic rearrangements of PML/RARalpha, AML1/ETO and CBFB/MYH11 respectively in acute myeloid leukaemia. MRC Adult Leukaemia Working Party. Medical Research Council. Br J Haematol 101:338-340.

Li Z, Herold T, He C, Valk PJ, Chen P, Jurinovic V, Mansmann U, Radmacher MD, Maharry KS, Sun M, Yang X, Huang H, Jiang X, Sauerland MC, Buchner T, Hiddemann W, Elkahloun A, Neilly MB, Zhang Y, Larson RA, Le Beau MM, Caligiuri MA, Dohner K, Bullinger L, Liu PP, Delwel R, Marcucci G, Lowenberg B, Bloomfield CD, Rowley JD, Bohlander SK, Chen J. 2013. Identification of a 24-gene prognostic signature that improves the European LeukemiaNet risk classification of acute myeloid leukemia: an international collaborative study. J Clin Oncol 31:1172-1181.

Liu S, Wu LC, Pang J, Santhanam R, Schwind S, Wu YZ, Hickey CJ, Yu J, Becker H, Maharry K, Radmacher MD, Li C, Whitman SP, Mishra A, Stauffer N, Eiring AM, Briesewitz R, Baiocchi RA, Chan KK, Paschka P, Caligiuri MA, Byrd JC, Croce CM, Bloomfield CD, Perrotti D, Garzon R, Marcucci G. 2010. Sp1/NFkappaB/HDAC/miR-29b regulatory network in KIT-driven myeloid leukemia. Cancer Cell 17:333-347.

Marcucci G, Metzeler KH, Schwind S, Becker H, Maharry K, Mrozek K, Radmacher MD, Kohlschmidt J, Nicolet D, Whitman SP, Wu YZ, Powell BL, Carter TH, Kolitz JE, Wetzler M, Carroll AJ, Baer MR, Moore JO, Caligiuri MA, Larson RA, Bloomfield CD. 2012. Age-related prognostic impact of different types of DNMT3A mutations in adults with primary cytogenetically normal acute myeloid leukemia. J Clin Oncol 30:742-750.

Metzeler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Gorlich D, Schneider S, Konstandin NP, Dufour A, Braundl K, Ksienzyk B, Zellmeier E, Hartmann L, Greif PA, Fiegl M, Subklewe M, Bohlander SK, Krug U, Faldum A, Berdel WE, Wormann B, Buchner T, Hiddemann W, Braess J, Spiekermann K, Group AS. 2016. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood 128:686-698.

Micol JB, Duployez N, Boissel N, Petit A, Geffroy S, Nibourel O, Lacombe C, Lapillonne H, Etancelin P, Figeac M, Renneville A, Castaigne S, Leverger G, Ifrah N, Dombret H, Preudhomme C, Abdel-Wahab O, Jourdan E. 2014. Frequent ASXL2 mutations in acute myeloid leukemia patients with t(8;21)/RUNX1-RUNX1T1 chromosomal translocations. Blood 124:1445-1449.

Opatz S, Polzer H, Herold T, Konstandin NP, Ksienzyk B, Zellmeier E, Vosberg S, Graf A, Krebs S, Blum H, Hopfner KP, Kakadia PM, Schneider S, Dufour A, Braess J, Sauerland MC, Berdel WE, Buchner T, Woermann BJ, Hiddemann W, Spiekermann K, Bohlander SK, Greif PA. 2013. Exome sequencing identifies recurring FLT3 N676K mutations in core-binding factor leukemia. Blood 122:1761-1769.

Ostler KR, Davis EM, Payne SL, Gosalia BB, Exposito-Cespedes J, Le Beau MM, Godley LA. 2007. Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins. Oncogene 26:5553-5563.

Ostler KR, Yang Q, Looney TJ, Zhang L, Vasanthakumar A, Tian Y, Kocherginsky M, Raimondi SL, DeMaio JG, Salwen HR, Gu S, Chlenski A, Naranjo A, Gill A, Peddinti R, Lahn BT, Cohn SL, Godley LA. 2012. Truncated DNMT3B isoform DNMT3B7 suppresses growth, induces differentiation, and alters DNA methylation in human neuroblastoma. Cancer Res 72:4714-4723.

Schneider F, Hoster E, Schneider S, Dufour A, Benthaus T, Kakadia PM, Bohlander SK, Braess J, Heinecke A, Sauerland MC, Berdel WE, Buechner T, Woermann BJ, Feuring-Buske M, Buske C, Creutzig U, Thiede C, Zwaan MC, van den Heuvel-Eibrink MM, Reinhardt D, Hiddemann W, Spiekermann K. 2012. Age-dependent frequencies of NPM1 mutations and FLT3-ITD in patients with normal karyotype AML (NK-AML). Ann Hematol 91:9-18.

Schulze I, Rohde C, Scheller-Wendorff M, Baumer N, Krause A, Herbst F, Riemke P, Hebestreit K, Tschanter P, Lin Q, Linhart H, Godley LA, Glimm H, Dugas M, Wagner W, Berdel WE, Rosenbauer F, Muller-Tidow C. 2016. Increased DNA methylation of Dnmt3b targets impairs leukemogenesis. Blood 127:1575-1586.

Shah MY, Vasanthakumar A, Barnes NY, Figueroa ME, Kamp A, Hendrick C, Ostler KR, Davis EM, Lin S, Anastasi J, Le Beau MM, Moskowitz IP, Melnick A, Pytel P, Godley LA. 2010. DNMT3B7, a truncated DNMT3B isoform expressed in human tumors, disrupts embryonic development and accelerates lymphomagenesis. Cancer Res 70:5840-5850.

Sweetser DA, Peniket AJ, Haaland C, Blomberg AA, Zhang Y, Zaidi ST, Dayyani F, Zhao Z, Heerema NA, Boultwood J, Dewald GW, Paietta E, Slovak ML, Willman CL, Wainscoat JS, Bernstein ID, Daly SB. 2005. Delineation of the

minimal commonly deleted segment and identification of candidate tumor-suppressor genes in del(9q) acute myeloid leukemia. Genes Chromosomes Cancer 44:279-291.

TCGA. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368:2059-2074.

Thol F, Damm F, Ludeking A, Winschel C, Wagner K, Morgan M, Yun H, Gohring G, Schlegelberger B, Hoelzer D, Lubbert M, Kanz L, Fiedler W, Kirchner H, Heil G, Krauter J, Ganser A, Heuser M. 2011. Incidence and prognostic influence of DNMT3A mutations in acute myeloid leukemia. J Clin Oncol 29:2889-2896.

Tickenbrock L, Schwable J, Wiedehage M, Steffen B, Sargin B, Choudhary C, Brandts C, Berdel WE, Muller-Tidow C, Serve H. 2005. Flt3 tandem duplication mutations cooperate with Wnt signaling in leukemic signal transduction. Blood 105:3699-3706.

Toedling J, Schmeier S, Heinig M, Georgi B, Roepcke S. 2005. MACAT--microarray chromosome analysis tool. Bioinformatics 21:2112-2113.

Vasanthakumar A, Lepore JB, Zegarek MH, Kocherginsky M, Singh M, Davis EM, Link PA, Anastasi J, Le Beau MM, Karpf AR, Godley LA. 2013. Dnmt3b is a haploinsufficient tumor suppressor gene in Myc-induced lymphomagenesis. Blood 121:2059-2063.

Vosberg S, Herold T, Hartmann L, Neumann M, Opatz S, Metzeler KH, Schneider S, Graf A, Krebs S, Blum H, Baldus CD, Hiddemann W, Spiekermann K, Bohlander SK, Mansmann U, Greif PA. 2016. Close correlation of copy number aberrations detected by next-generation sequencing with results from routine cytogenetics in acute myeloid leukemia. Genes Chromosomes Cancer 55:553-567.

Wechsler J, Greene M, McDevitt MA, Anastasi J, Karp JE, Le Beau MM, Crispino JD. 2002. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. Nat Genet 32:148-152.

Figure legends

Figure 1. Relapse free and overall survival of AML patients with del(9q) in different cytogenetic subgroups. Kaplan–Meier estimates of AML patients with del(9q) as the sole abnormality, in combination with t(8;21) or with various additional cytogenetic alterations (A,B). AML patients with del(9q) as the sole abnormality compared to other AML patients from the intermediate II group (C,D).

Figure 2. Mutation profile of AML with del(9q). Distribution of mutated genes in 26 AML patients with del(9q) as the sole abnormality, in combination with t(8;21) or with various additional cytogenetic alterations. Each column represents one patient. Dark grey boxes indicate mutations, whereas light grey boxes indicate wild-type status. Gene names are indicated to the left, mutation frequencies are indicated to the right and patient identifiers are shown at the bottom. Patients show a high frequency of mutations in *NPM1*, *WT1* and *DNMT3A*. Arrows indicate the five patients that were exome-sequenced. Stars indicate patients with a complex karyotype. Only mutations with a frequency of ≥ 8 % and *DNMT3B* are shown.

Figure 3. Expression of *TLE1* and *TLE4* in different cytogenetic AML subgroups. Boxplot of *TLE1* (A,B) and *TLE4* (C,D,E,F) expression in various cytogenetic subgroups (A,C,E) and the whole group (B,D,F). A-D show the analysis of the data set GSE37642 and E-F of the TCGA data set. Abbreviations: CN-AML, AML with normal karyotype.

Figure 4. Expression of *TCF7L2* and SP1 in relation to *TLE4.* Scatterplot of *TLE4* combined with *TCF7L2* and *SP1* expression in the GSE37642 and TCGA data set.

TABLE 1. Patient Characteristics

Variable	**Del(9q) solely[a]**	**ELN Intermediate-II Control Group[a]**	***P*-value**
No. of patients	23	362	
Median age, years (range)	60 (22-76)	61 (18-82)	0.65
Male sex, no. (%)	9 (39)	207 (57)	0.13
White cell count, G/l, median (range)	18.6 (1.3-150)	11.4 (0.6-341)	0.57
Hemoglobin, g/dl, median (range)	8.2. (4.2-13.6)	9.1 (3.8-16.9)	0.09
Platelet count, G/l ,median (range)	37 (5-260)	54 (1-1760)	0.39
LDH (U/l), median(range)	525 (96-1367)	397 (115-11140)	0.83
Bone marrow blasts, %, median (range)	80 (19-100)	80 (11-100)	0.4
Bone marrow blasts at day 16, %, median (range)	2 (0-10)	5 (0-100)	**0.004**
Performance Status (ECOG) ≥ 2 (%)	8 (38)	112 (34)	0.81
de novo AML, no. (%)	20 (87)	268 (74)	0.22
Allogeneic transplantation in first CR, no. (%)	1 (4)	29 (8)	1
Complete remission, no. (%)	15 (65)	195 (54)	0.39
Relapse, no. (%)	11 (73)	154 (79)	0.53
Deceased, no. (%)	15 (65)	289 (80)	0.11

[a]All patients were enrolled in the AMLCG-99 trial and received intensive induction treatment. All patients are classified as ELN Intermediate-II; Del(9q) solely: AML patients with isolated del(9q).

TABLE 2. Transcription Factors Significantly Associated with TLE4 Expression in the AMLCG and TCGA Data Sets

	AMLCG		TCGA	
Gene	**rho**	**adjusted *P*-value**	**rho**	**adjusted *P*-value**
WRNIP1	-0.325	<0.001	-0.167	<0.001
BRCA1	-0.283	<0.001	-0.175	<0.001
TCF3	-0.266	<0.001	-0.308	<0.001
TEAD4	-0.220	<0.001	-0.309	<0.001
TBP	-0.193	<0.001	-0.267	<0.001
ZNF143	-0.188	<0.001	0.233	<0.001
PML	-0.171	<0.001	-0.245	<0.001
MEF2C	-0.151	<0.001	0.265	<0.001
CREB1	-0.150	<0.001	0.358	<0.001
E2F6	-0.109	<0.001	-0.187	<0.001
NFYA	-0.107	<0.001	0.245	<0.001
MAZ	0.115	<0.001	-0.202	<0.001
RAD21	0.171	<0.001	0.270	<0.001
NR3C1	0.176	<0.001	0.311	<0.001
EZH2	0.180	<0.001	-0.195	<0.001
EGR1	0.187	<0.001	0.219	<0.001
SP1	0.190	<0.001	0.465	<0.001
MEF2A	0.197	<0.001	0.440	<0.001
IRF1	0.209	<0.001	0.283	<0.001
MAX	0.244	<0.001	0.198	<0.001
STAT3	0.264	<0.001	0.197	<0.001
FOS	0.290	<0.001	0.402	<0.001
CHD2	0.294	<0.001	0.282	<0.001
BACH1	0.320	<0.001	0.505	<0.001
CHD1	0.322	<0.001	0.251	<0.001
TCF7L2	0.361	<0.001	0.396	<0.001
ELF1	0.362	<0.001	0.384	<0.001

Figure 1: Relapse free and overall survival of AML patients with del(9q) in different cytogenetic subgroups

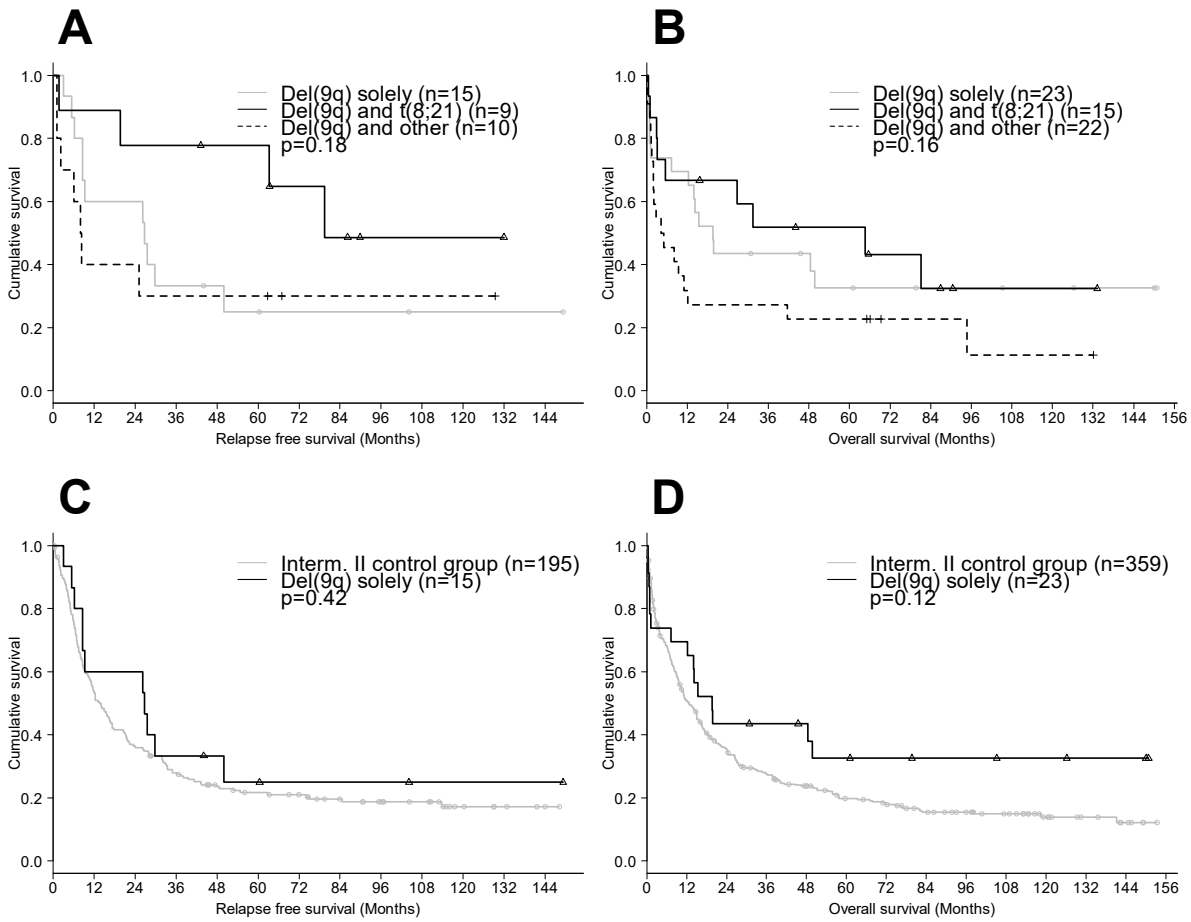


Figure 2: Frequency distribution of recurrently mutated genes in AML with del(9q)

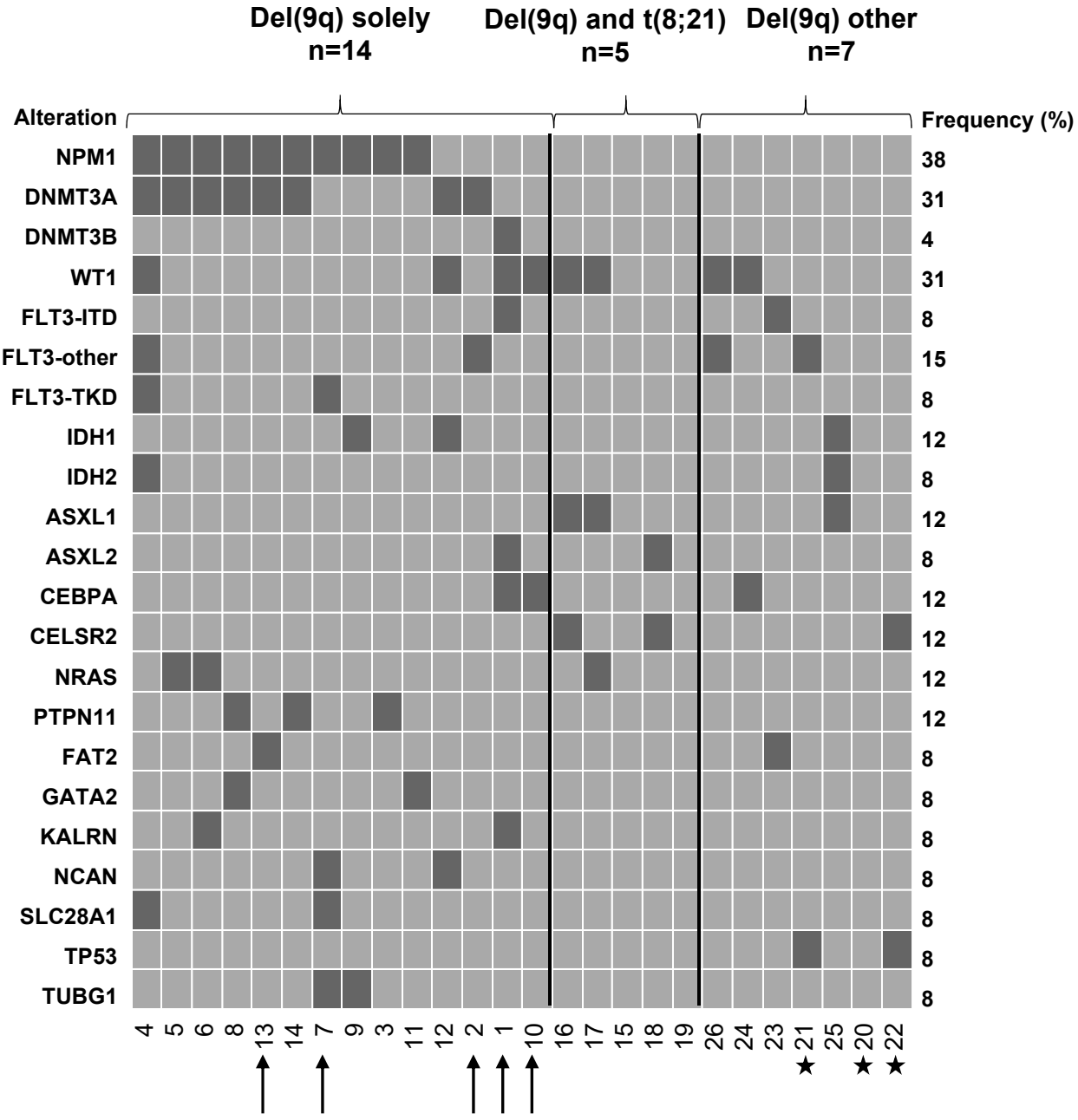


Figure 3: Expression of *TLE1* and *TLE4* in different cytogenetic AML subgroups

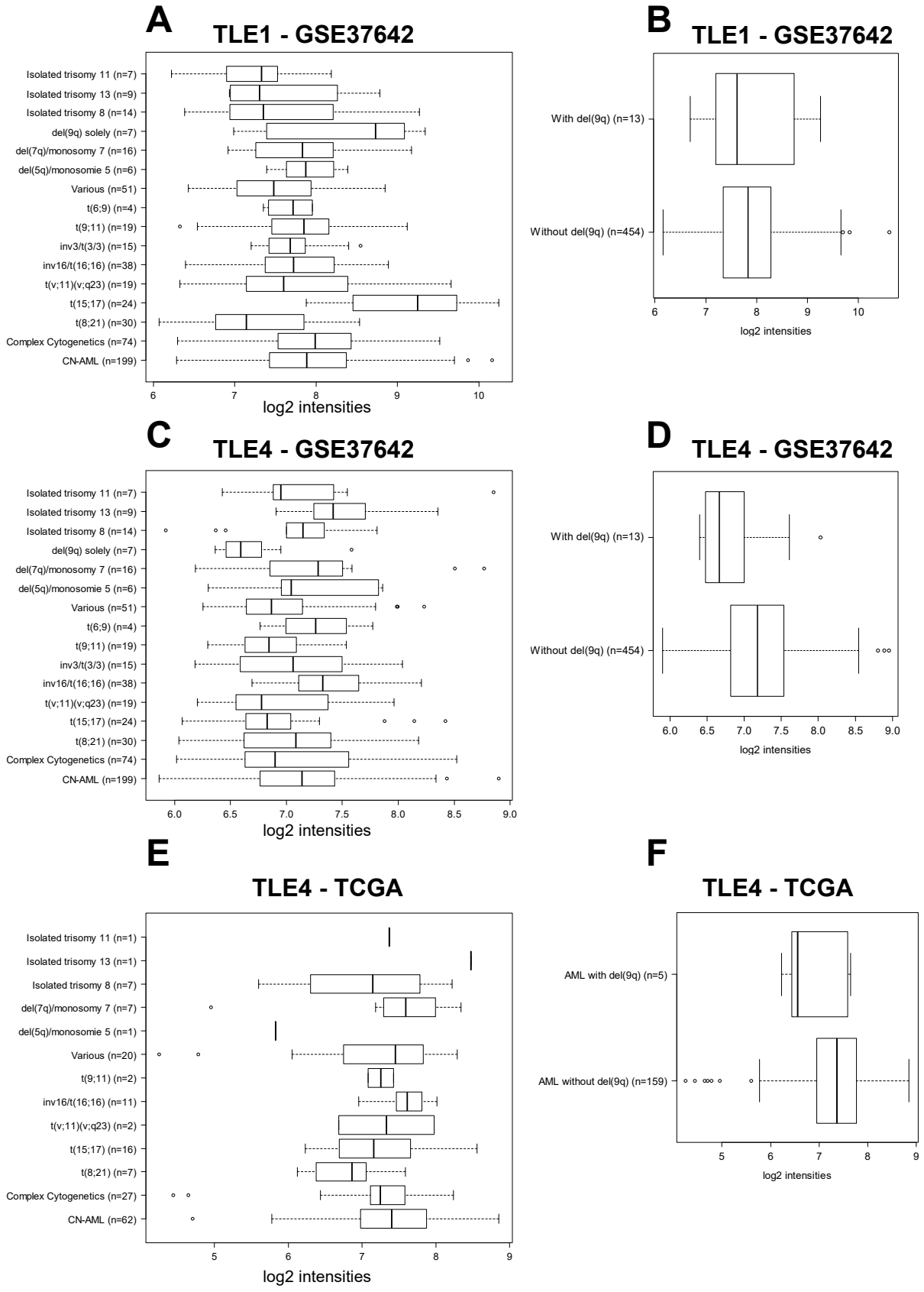
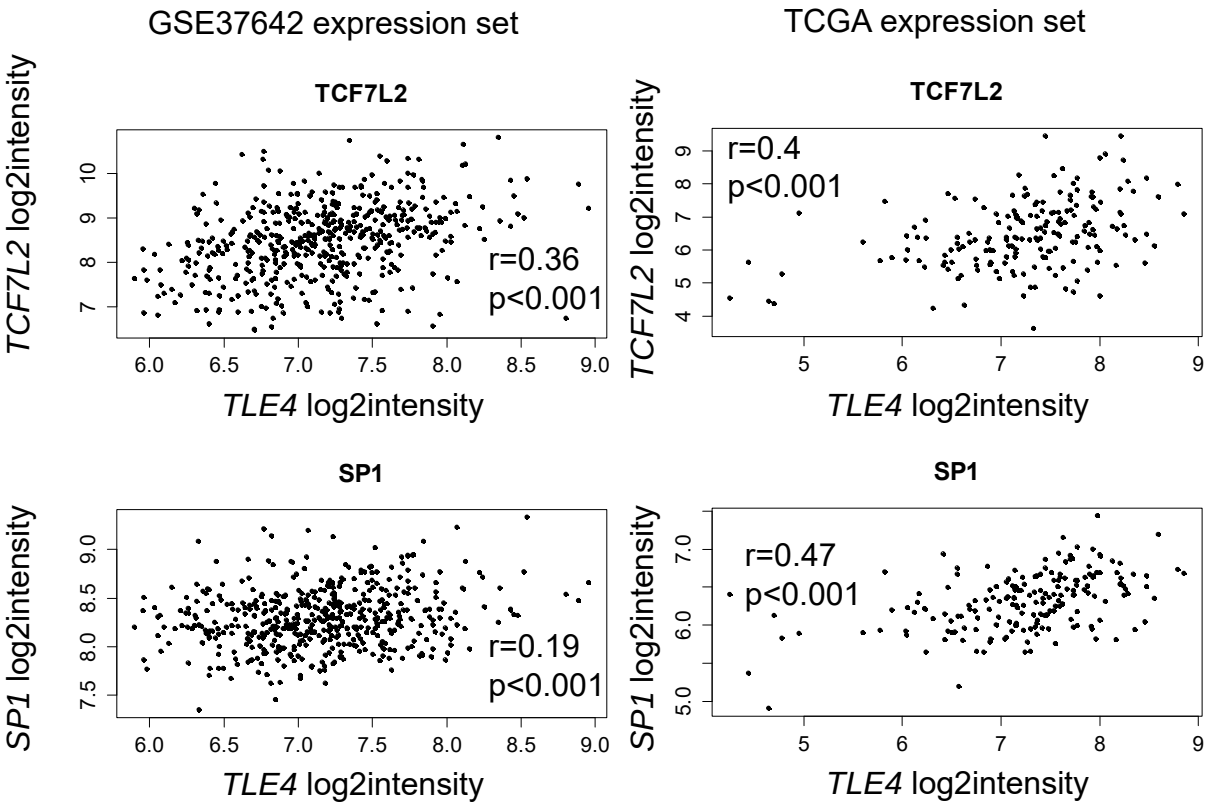


Figure 4: Expression of *TCF7L2* and *SP1* in relation to *TLE4*



Supplementary Information

Acute Myeloid Leukemia with Del(9q) is Characterized by Frequent Mutations of *NPM1*, *DNMT3A, WT1* and Low Expression of *TLE4*

Tobias Herold,[1,2,3,4,*] Klaus H. Metzeler,[1,2,3,4] Sebastian Vosberg,[1,2,3,4] Luise Hartmann,[1,2,3,4] Vindi Jurinovic,[5] Sabrina Opatz,[1,2,3,4] Nikola P. Konstandin, [1] Stephanie Schneider,[1] Evelyn Zellmeier,[1] Bianka Ksienzyk,[1] Alexander Graf,[7] Stefan Krebs,[7] Helmut Blum,[7] Maria Cristina Sauerland,[8] Thomas Büchner,[9] Wolfgang E. Berdel,[9] Bernhard J. Wörmann,[10] Ulrich Mansmann,[3,4,5] Wolfgang Hiddemann,[1,2,3,4] Stefan K. Bohlander,[11] Karsten Spiekermann[1,2,3,4] and Philipp A. Greif[1,2,3,4]

[1]Department of Internal Medicine 3, University Hospital Grosshadern, Ludwig-Maximilians-Universität (LMU) München, München, Germany
[2]Clinical Cooperative Group Leukemia, Helmholtz Center Munich for Environmental Health, München, Germany
[3]German Cancer Consortium (DKTK), Heidelberg, Germany
[4]German Cancer Research Center (DKFZ), Heidelberg, Germany
[5]Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität (LMU) München, München, Germany
[6]Center for Human Genetics, Philipps University, Marburg, Germany
[7]Laboratory for Functional Genome Analysis (LAFUGA), Gene Center, Ludwig-Maximilians-Universität (LMU) München, München, Germany
[8]Institute of Biostatistics and Clinical Research, and [9]Department of Medicine A-Hematology, Oncology and Pneumology, University of Münster, Münster, Germany
[10] Department of Hematology, Oncology and Tumor Immunology, Charité University Medicine, Campus Virchow, Berlin, Germany
[11]Department of Molecular Medicine and Pathology, The University of Auckland, Auckland, New Zealand

* Correspondence to: Tobias Herold, MD, Marchioninistr. 15, 81377 München, Phone: +49 89 4400 75833, FAX: +49 89 4400-72221, tobias.herold@med.uni-muenchen.de

Supplementary Figure legends:

Figure S1. Flow Chart.

Figure S2. Results of Sanger sequencing of *DNMT3B.*

Figure S3. Relapse free and overall survival of AML patients with del(9q) with or without mutations in *WT1.* Kaplan–Meier estimates of AML patients with del(9q) as sole abnormality with or without a mutation of *WT1*.

Figure S4. Copy number variations of *TLE1*, *TLE4* and the CDR of del(9q). Visualization of chromosome 9, focusing on the genomic region of *TLE1*, *TLE4* and the CDR of del(9q) using the copy number data provided by TCGA (n=200). Only samples with significant deletions in the analyzed region are shown (n=27). Deletions are marked in red. Chromosomal locations of genes are shown by blue dashed lines. No significant clustering of deletions in *TLE4* was observed.

Figure S5. Differential gene expression of AML with del(9q). Heatmap showing the 26 probe sets significantly deregulated using a non-adjusted *P*-value of ≤0.001 comparing del(9q) with all kind of other cytogenetic alteration. As shown in the grey bar, samples with del(9q) did not cluster together (light gray) in comparison to samples without this alteration (dark grey).

Figure S6. Regional gene expression on chromosome 9 in AML with del(9q). Expression of probe sets located on chromosome 9 displayed by MACAT in patients with del(9q) (n=13) compared with del(9q) negative AML cases (n=454). The region of del(9q) shows significantly lower gene expression in case of a deletion (yellow dots).

Figure S7. Percentages of metaphases with del(9q) and variant allele frequencies of *DNMT3A*, *NPM1* or *WT1* mutations. Figure showing the association of the percentages of metaphases with del(9q) and variant allele frequencies of *DNMT3A*, *NPM1* or *WT1* mutations. Abbreviations: VAF, variant allele frequencies

Supplementary Table S1. Quality metrics summary of exon sequencing.

ID	1	1 R	2	2 R	7	7 R	10	10 R	13	13 R
# of total reads:	45,755,476	42,870,854	46,549,028	29,985,508	43,564,000	43,625,108	41,361,080	46,226,186	49,045,644	42,285,960
% of reads mapped to genome:	94.2971	99.2009	99.1673	99.2536	99.146	99.1921	99.2017	99.2064	99.525	99.4237
% of reads mapped with MQ >= Q20:	92.59	97.35	97.36	97.55	97.44	97.51	97.52	97.45	98.26	97.98
# of sequenced bases mapped to genome:	3,394,887,086	3,348,127,966	3,633,721,261	2,341,234,462	3,401,812,383	3,406,503,321	3,227,790,798	3,610,354,347	3,841,577,579	3,305,786,373
% of bases mapped in HQ reads:	98.21	98.15	98.2	98.31	98.31	98.33	98.32	98.25	98.75	98.56
% of >=Q20 bases mapped in HQ reads:	97.21	97.17	97.23	97.31	97.35	97.32	97.31	97.27	97.75	97.55
% of target sequenced:	95.69	95.64	95.72	95.16	95.58	95.59	95.53	95.75	95.82	95.78
% of target sequenced, min 10x coverage:	75.05	75.25	77.43	66.18	75.59	75.07	76.11	77.63	82.02	79.66
% of target sequenced, min 20x coverage:	52.21	52.75	56.2	38.94	53.73	52.9	53.74	56.51	64.67	59.13
% of target sequenced, min 30x coverage:	35.31	35.94	39.47	22.39	37.29	36.42	36.43	39.7	48.99	41.77
mean target coverage:	28.7	29.08	31.44	20.77	29.96	29.47	29.14	31.39	38.49	32.51
% bases on target:	29.54	30.35	30.24	30.84	30.73	30.18	31.46	30.4	34.91	34.3

Abbreviation: R=Remission

Supplementary Table S2. Genes included in targeted amplicon sequencing.

Nr.	Entrez Gene Id	Gene Symbol	GeneDescription
1	5243	ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1
2	9429	ABCG2	ATP-binding cassette, sub-family G (WHITE), member 2
3	91703	ACY3	aspartoacylase (aminocyclase) 3
4	375790	AGRN	agrin
5	374860	ANKRD30B	ankyrin repeat domain 30B
6	307	ANXA4	annexin A4
7	171023	ASXL1	additional sex combs like 1 (Drosophila)
8	55252	ASXL2	additional sex combs like 2 (Drosophila)
9	54880	BCOR	BCL6 corepressor
10	63035	BCORL1	BCL6 corepressor-like 1
11	673	BRAF	v-raf murine sarcoma viral oncogene homolog B1
12	9024	BRSK2	BR serine/threonine kinase 2
13	84267	C9orf64	
14	867	CBL	Cas-Br-M (murine) ecotropic retroviral transforming sequence
15	978	CDA	cytidine deaminase
16	51654	CDK5RAP1	CDK5 regulatory subunit associated protein 1
17	1029	CDKN2A	cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)
18	1050	CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha
19	10153	CEBPZ	CCAAT/enhancer binding protein (C/EBP), zeta
20	1952	CELSR2	cadherin, EGF LAG seven-pass G-type receptor 2 (flamingo homolog, Drosophila)
21	29121	CLEC2D	C-type lectin domain family 2, member D
22	27151	CPAMD8	C3 and PZP-like, alpha-2-macroglobulin domain containing 8
23	1387	CREBBP	CREB binding protein
24	1436	CSF1R	colony stimulating factor 1 receptor
25	1441	CSF3R	colony stimulating factor 3 receptor (granulocyte)
26	1620	DBC1	deleted in bladder cancer 1
27	9937	DCLRE1A	DNA cross-link repair 1A
28	1788	DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha
29	1789	DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta
30	1837	DTNA	dystrobrevin, alpha
31	136227	EMID2	EMI domain containing 2
32	2120	ETV6	ets variant 6
33	149371	EXOC8	exocyst complex component 8
34	2146	EZH2	enhancer of zeste homolog 2 (Drosophila)
35	339479	FAM5C	family with sequence similarity 5, member C
36	2196	FAT2	FAT tumor suppressor homolog 2 (Drosophila)
37	80204	FBXO11	F-box protein 11
38	55294	FBXW7	F-box and WD repeat domain containing 7
39	2322	FLT3	fms-related tyrosine kinase 3
40	2332	FMR1	fragile X mental retardation 1
41	257019	FRMD3	FERM domain containing 3
42	10690	FUT9	fucosyltransferase 9 (alpha (1,3) fucosyltransferase)
43	2623	GATA1	GATA binding protein 1 (globin transcription factor 1)

44	2624	GATA2	GATA binding protein 2
45	2625	GATA3	GATA binding protein 3
46	80318	GKAP1	G kinase anchoring protein 1
47	84059	GPR98	G protein-coupled receptor 98
48	81502	HM13	histocompatibility (minor) 13
49	3190	HNRNPK	heterogeneous nuclear ribonucleoprotein K
50	3265	HRAS	v-Ha-ras Harvey rat sarcoma viral oncogene homolog
51	3417	IDH1	isocitrate dehydrogenase 1 (NADP+), soluble
52	3418	IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial
53	3575	IL7R	interleukin 7 receptor
54	83729	INHBE	inhibin, beta E
55	9922	IQSEC1	IQ motif and Sec7 domain 1
56	3716	JAK1	Janus kinase 1
57	3717	JAK2	Janus kinase 2
58	3718	JAK3	Janus kinase 3
59	8997	KALRN	kalirin, RhoGEF kinase
60	83473	KATNAL2	katanin p60 subunit A-like 2
61	55582	KIF27	kinesin family member 27
62	3815	KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog
63	84861	KLHL22	kelch-like 22 (Drosophila)
64	3845	KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
65	3955	LFNG	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase
66	11024	LILRA1	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 1
67	9947	MAGEC1	melanoma antigen family C, 1
68	55700	MAP7D1	MAP7 domain containing 1
69	58508	MLL3	myeloid/lymphoid or mixed-lineage leukemia 3
70	55904	MLL5	myeloid/lymphoid or mixed-lineage leukemia 5 (trithorax homolog, Drosophila)
71	4312	MMP1	matrix metallopeptidase 1 (interstitial collagenase)
72	29074	MRPL18	mitochondrial ribosomal protein L18
73	55149	MTPAP	mitochondrial poly(A) polymerase
74	4583	MUC2	mucin 2, oligomeric mucus/gel-forming
75	4588	MUC6	mucin 6, oligomeric mucus/gel-forming
76	79784	MYH14	myosin, heavy chain 14, non-muscle
77	1463	NCAN	neurocan
78	26155	NOC2L	nucleolar complex associated 2 homolog (S. cerevisiae)
79	4851	NOTCH1	notch 1
80	4869	NPM1	nucleophosmin (nucleolar phosphoprotein B23, numatrin)
81	4893	NRAS	neuroblastoma RAS viral (v-ras) oncogene homolog
82	4915	NTRK2	neurotrophic tyrosine kinase, receptor, type 2
83	4926	NUMA1	nuclear mitotic apparatus protein 1
84	119765	OR4B1	olfactory receptor, family 4, subfamily B, member 1
85	54726	OTUD4	OTU domain containing 4
86	5042	PABPC3	poly(A) binding protein, cytoplasmic 3
87	84295	PHF6	PHD finger protein 6
88	22843	PPM1E	protein phosphatase, Mg2+/Mn2+ dependent, 1E
89	65121	PRAMEF1	PRAME family member 1
90	158471	PRUNE2	prune homolog 2 (Drosophila)

91	5728	PTEN	phosphatase and tensin homolog
92	5781	PTPN11	protein tyrosine phosphatase, non-receptor type 11
93	5885	RAD21	RAD21 homolog (S. pombe)
94	158158	RASEF	RAS and EF-hand domain containing
95	80010	RMI1	RMI1, RecQ mediated genome instability 1, homolog (S. cerevisiae)
96	9810	RNF40	ring finger protein 40
97	23168	RTF1	Rtf1, Paf1/RNA polymerase II complex component, homolog (S. cerevisiae)
98	861	RUNX1	runt-related transcription factor 1
99	6262	RYR2	ryanodine receptor 2 (cardiac)
100	6294	SAFB	scaffold attachment factor B
101	26040	SETBP1	SET binding protein 1
102	7536	SF1	splicing factor 1
103	10291	SF3A1	splicing factor 3a, subunit 1, 120kDa
104	23451	SF3B1	splicing factor 3b, subunit 1, 155kDa
105	10326	SIRPB1	signal-regulatory protein beta 1
106	9154	SLC28A1	solute carrier family 28 (sodium-coupled nucleoside transporter), member 1
107	64078	SLC28A3	solute carrier family 28 (sodium-coupled nucleoside transporter), member 3
108	2030	SLC29A1	solute carrier family 29 (nucleoside transporters), member 1
109	6543	SLC8A2	solute carrier family 8 (sodium/calcium exchanger), member 2
110	8243	SMC1A	structural maintenance of chromosomes 1A
111	9126	SMC3	structural maintenance of chromosomes 3
112	55627	SMPD4	sphingomyelin phosphodiesterase 4, neutral membrane (neutral sphingomyelinase-3)
113	8470	SORBS2	sorbin and SH3 domain containing 2
114	6427	SRSF2	serine/arginine-rich splicing factor 2
115	10735	STAG2	stromal antigen 2
116	94121	SYTL4	synaptotagmin-like 4
117	50840	TAS2R14	taste receptor, type 2, member 14
118	7062	TCHH	trichohyalin
119	54790	TET2	tet methylcytosine dioxygenase 2
120	7088	TLE1	transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila)
121	7089	TLE2	transducin-like enhancer of split 2 (E(sp1) homolog, Drosophila)
122	7090	TLE3	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)
123	7091	TLE4	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)
124	7157	TP53	tumor protein p53
125	7283	TUBG1	tubulin, gamma 1
126	7307	U2AF1	U2 small nuclear RNA auxiliary factor 1
127	11338	U2AF2	U2 small nuclear RNA auxiliary factor 2
128	29979	UBQLN1	ubiquilin 1
129	55779	WDR52	WD repeat domain 52
130	7490	WT1	Wilms tumor 1
131	7508	XPC	xeroderma pigmentosum, complementation group C
132	79038	ZFYVE21	zinc finger, FYVE domain containing 21
133	10520	ZNF211	zinc finger protein 211
134	10782	ZNF274	zinc finger protein 274
135	90333	ZNF468	zinc finger protein 468
136	100287226	ZNF729	zinc finger protein 729
137	8233	ZRSR2	zinc finger (CCCH type), RNA-binding motif and serine/arginine rich 2

Supplementary Table S3. Quality metrics summary of targeted amplicon sequencing.

	Median
Region Size [kbp]	486
Percent Target bases 30x [%]	92
Paired reads past filter [k]	1064001
Individual reads past filter [k]	2128002
Raw read length [bp]	251
Raw sequence before quality trimming [Mb]	534129
Reads surviving quality trim [% of raw reads]	89453
Reads post trimming [k]	1476356
Mean length of trimmed reads [bp]	158518
Reads aligned to hg19 with MAPQ>20 [k]	1434737
Confidently aligned [% of trimmed reads]	97.17
Confidently aligned sequence [Mb]	227607
Confidently aligned HQ sequence (Q>30) [Mb]	221438
Substitution rate [%]	0.174
Indel rate [%]	0.009
Reads aligned in pairs [% of paired reads]	99.99

Supplementary Table S4. Primer sequences for Sanger sequencing and NCBI accession numbers.

Primer name	Sequence	NCBI Reference sequences
ASXL2_fwd_1	GAATGCTTTCGCAACATGG	gDNA : NC_000002.12 range: 25878443-25733753 mRNA: NM_018263.4
ASXL2_rev_1	AAGATTCTTGGCTTTCACTTTTG	
ASXL2_fwd_2	AGGATGAGGATCTCTTGGAGC	
ASXL2_rev_2	TCAAATATATCCAACTGCTTCTGC	
CEBPZ_fwd_1	CTGCAAAGCTTTCTCCTTCTG	gDNA: NC_000002.12 range: 37231597-37201629 mRNA: NM_005760.2
CEBPZ_rev_1	TCTGGCTGACATGAGTATTAATTTG	
CEBPZ_fwd_2	AAACCCACATGGAATAGCAG	
CEBPZ_rev_2	CAAATTTAGGAGAGTTCATTTCCC	
CELSR2_fwd_1	GCTGGAGAACAGTGATTAAGGC	gDNA: NC_000001.11 range: 109250019-109275756 mRNA: NM_001408.2
CELSR2_rev_1	AGCAGTGTTGCCCTCAGTG	
CELSR2_fwd_2	TGACACCAATGACCATGACC	
CELSR2_rev_2	CCACCATCCTGTGCTCG	
CELSR2_fwd_3	TTACTGAAGGTGGGTGGAGG	
CELSR2_rev_3	CAGTCCTGCTTCCCAAGAAC	
GPR98_fwd_1	TTTCCAGCGATTAAAAGCTAATG	gDNA: NC_000005.10 range: 90558800-91164216 mRNA: NM_032119.3
GPR98_rev_1	AATGGCTTTCATCCACACAAG	
GPR98_fwd_2	AACTTGTCACCCTTCATGGC	
GPR98_rev_2	TGTCATGGGTCCATGCTAC	
GPR98_fwd_3	TGCTGTAATTAACATCAGTTTTACTTG	
GPR98_rev_3	CCACCAGTTATCTGAATTGGC	
GPR98_fwd_4	TCAAGTATTCAATGCAAGTTAAATGTC	
GPR98_rev_4	CCAAATCAGCAGGGAGAGG	
GPR98_fwd_5	CAACTTGACATTGGGAAAGC	
GPR98_rev_5	TCTTTGGTTCTTGTACAGCACAG	
GPR98_fwd_6	TGAAATGGATGAAAATTGATTACTG	
GPR98_rev_6	GCACACACCAGCACAGAGAG	
PPM1E_fwd_1	TTGGTTCATTTGAATCATGCTATAC	gDNA: NC_000017.11 range: 58755869-58985179 mRNA: NM_014906.4
PPM1E_rev_1	AAATAGCCCTTTAACCACACAG	
PPM1E_fwd_2	GAACCCGAGTCCGAGCC	
PPM1E_rev_2	CAGTTTGAGGGGCCGAG	
WDR52_fwd_1	TTTCTGTTGCCATACAAAATCTG	gDNA: NC_000003.12 range: 113445190-113286930 mRNA: NM_001164496.1
WDR52_rev_1	TCACCTGGATAATTCCCTTTC	
WDR52_fwd_2	TCAACTACTTCAACCTCACATCTTC	
WDR52_rev_2	ATACTGTGCATGGTGTTGGC	
DNMT3B_fwd_1	GGAGTGGAGGAAATGAGCTG	gDNA: NC_000020.11 range: 32762385-32809356 mRNA: NM_006892
DNMT3B _rev_1	TTTCAGGGTGTGGAGGACTG	
DNMT3B _fwd_2	CAGAGCTTGAGTCTTTGCCC	
DNMT3B _rev_2	CCTGGCTACCCTGTTGTGAC	

Supplementary Table S5. Additional patient characteristics.

Variable	Del(9q) solely	Del(9q) and t(8;21)	Del(9q) with other aberrations
No. of patients	23	15	23
Median age, years (range)	60 (22-76)	48 (32-81)	62 (28-74)
Male sex, no. (%)	9 (39)	8 (53)	15 (65)
White-cell count, G/l, median (range)	18.6 (1.3-150)	6.2 (1-22)	4.5 (0.6-327)
Hemoglobin, g/dl, median (range)	8.2. (4.2-13.6)	7.1 (3.5-11.8)	9.2 (5.8-12.8)
Platelet count, G/l ,median (range)	37 (5-260)	29 (3-344)	60 (1-335)
LDH (U/l), median(range)	525 (96-1367)	344 (200-2034)	348 (211-1880)
Bone marrow blasts, %, median (range)	80 (19-100)	70 (30-90)	50 (20-95)
Bone marrow blasts at day 16, %, median (range)	2 (0-10)	4 (0-90)	18 (0-90)
Performance Status (ECOG) ≥ 2 (%)	8 (38)	7 (47)	6 (30)
de novo AML (%)	20 (87)	11 (73)	16 (70)
Allogeneic transplantation, no. (%)	6 (26)	3 (20)	8 (35)
Complete remission, no. (%)	15 (65)	9 (60)	10 (43)
Relapse, no. (%)	11 (73)	4 (44)	7 (70)
Deceased, no. (%)	15 (65)	9 (60)	18 (78)

Supplementary Table S6. Non-synonymous variants identified by exome and targeted amplicon sequencing.

ID	Cyto-genetics	Chr	Position	Gene	cDNA change	Transcript ID	AA change	VAF (%)	Exome	Sanger
1	Del9q solely	2	25973057	ASXL2	p.N456fs	NM_018263	p.N456fs	27	yes	yes
1	Del9q solely	20	31979989	CDK5RAP1	p.S168F	NM_016408	p.S168F	51	yes	
1	Del9q solely	19	33792404	CEBPA	p.R306P	NM_004364	p.R306P	35	yes	
1	Del9q solely	19	33793261	CEBPA	p.Q20fs	NM_004364	p.Q20fs	34		
1	Del9q solely	20	31388064	DNMT3B	p.I622fs	NM_006892	p.I622fs	27	yes	yes, confirmed somatic
1	Del9q solely	2	48061846	FBXO11	p.Y185fs	NM_025133	p.Y185fs	53	yes	
1	Del9q solely	13	28608254	FLT3-ITD	p.L601delinsPLQVTGSSDNEYFYVDFREYEYDL	NM_004119	p.L601delinsPLQVTGSSDNEYFYVDFREYEYDL	10		
1	Del9q solely	3	124413262	KALRN	p.R2497W	NM_001024660	p.R2497W	37	yes	
1	Del9q solely	16	30785389	RNF40	p.N987fs	NM_014771	p.N987fs	46	yes	
1	Del9q solely	11	32417908	WT1	p.A382fs	NM_024426	p.A382fs	18		
2	Del9q solely	2	25457242	DNMT3A	p.R882H	NM_175629	p.R882H	29		
2	Del9q solely	18	32392010	DTNA	p.T179M	NM_032979	p.T179M	28	yes	
2	Del9q solely	13	28592620	FLT3-other	p.Y842C	NM_004119	p.Y842C	11		
2	Del9q solely	5	90106881	GPR98	p.Q5268H	NM_032119	p.Q5268H	16	yes	yes, confirmed somatic
3	Del9q solely	9	21971057	CDKN2A	p.R115L	NM_058195	p.R115L	56		
3	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	55		
3	Del9q solely	12	112888163	PTPN11	p.G60V	NM_080601	p.G60V	41		
4	Del9q solely	2	25463182	DNMT3A	p.R771X	NM_175629	p.R771X	44		
4	Del9q solely	13	28592642	FLT3-TKD	p.N676K	NM_004119	p.D835H	15		
4	Del9q solely	13	28592642	FLT3-TKD	p.D835H	NM_004119	p.D835Y	9		
4	Del9q solely	13	28602340	FLT3-other	p.D835Y	NM_004119	p.N676K	16		
4	Del9q solely	15	90631934	IDH2	p.R140Q	NM_002168	p.R140Q	43		
4	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	30		
4	Del9q solely	15	85448881	SLC28A1	p.R239W	NM_004213	p.R239W	40		
4	Del9q solely	11	32417912	WT1	p.R380fs	NM_024426	p.R380fs	5		
5	Del9q solely	2	25457242	DNMT3A	p.R882H	NM_175629	p.R882H	46		
5	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	39		
5	Del9q solely	1	115256529	NRAS	p.Q61L	NM_002524	p.Q61L	45		
6	Del9q solely	2	25463298	DNMT3A	p.F732S	NM_175629	p.F732S	18		
6	Del9q solely	3	124153238	KALRN	p.G970S	NM_003947	p.G970S	46		
6	Del9q solely	5	170837545	NPM1	p.W288fs		p.W288fs	27		

6	Del9q solely	1	115258744	NRAS	p.G13D	NM_002524	p.G13D	32		
7	Del9q solely	13	28592642	FLT3-TKD	p.D835Y	NM_004119	p.D835Y	32	yes	
7	Del9q solely	X	147019042	FMR1	p.P350S	NM_002024	p.P350S	37	yes	
7	Del9q solely	6	96651896	FUT9	p.V289M	NM_006581	p.V289M	30	yes	
7	Del9q solely	12	57850266	INHBE	p.A230P	NM_031479	p.A230P	39	yes	
7	Del9q solely	22	20800837	KLHL22	p.D478H	NM_032775	p.D478H	36	yes	
7	Del9q solely	19	19360658	NCAN	p.R1302C	NM_004386	p.R1302C	39	yes	
7	Del9q solely	1	881898	NOC2L	p.V563M	NM_015658	p.V563M	26	yes	
7	Del9q solely	5	170837545	NPM1	p.W288fs		p.W288fs	33		
7	Del9q solely	17	57047057	PPM1E	p.D314A	NM_014906	p.D314A	36	yes	yes
7	Del9q solely	19	5653399	SAFB	p.D498fs	NM_002967	p.D498fs	33	yes	
7	Del9q solely	15	85488071	SLC28A1	p.Y616C	NM_004213	p.Y616C	36	yes	
7	Del9q solely	17	40764512	TUBG1	p.R156Q	NM_001070	p.R156Q	27	yes	
7	Del9q solely	3	14200277	XPC	p.E369G	NM_004628	p.E369G	22	yes	
7	Del9q solely	14	104196143	ZFYVE21	p.H180Q	NM_001198953	p.H180Q	36	yes	
8	Del9q solely	2	25466814	DNMT3A	p.K630_R631 delinsSPSX	NM_175629	p.K630_R631delinsSPSX	40		
8	Del9q solely	3	128200744	GATA2	p.T354M	NM_001145661	p.T354M	43		
8	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	55		
8	Del9q solely	12	112888166	PTPN11	p.D61V	NM_080601	p.D61V	38		
9	Del9q solely	2	37438172	CEBPZ	p.D952N	NM_005760	p.D952N	55		yes
9	Del9q solely	2	209113113	IDH1	p.R132C	NM_005896	p.R132C	30		
9	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	45		
9	Del9q solely	17	40767055	TUBG1	p.Q451P	NM_001070	p.Q451P	28		
10	Del9q solely	19	33792381	CEBPA	p.V314delinsKV	NM_004364	p.V314delinsKV	45	yes	
10	Del9q solely	20	30155907	HM13	p.P354L	NM_178581	p.P354L	46	yes	
10	Del9q solely	19	17942058	JAK3	p.P986R	NM_000215	p.P986R	42	yes	
10	Del9q solely	11	48238420	OR4B1	p.A20V	NM_001005470	p.A20V	40	yes	
10	Del9q solely	11	32413566	WT1	p.R462W	NM_024426	p.R462W	93		
11	Del9q solely	3	128200754	GATA2	p.N351D	NM_001145661	p.N351D	57		
11	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	38		
12	Del9q solely	2	25463241	DNMT3A	p.F751fs	NM_175629	p.F751fs	44		
12	Del9q solely	2	25463568	DNMT3A	p.I705T	NM_175629	p.I705T	46		
12	Del9q solely	2	209113113	IDH1	p.R132C	NM_005896	p.R132C	30		
12	Del9q solely	19	19356128	NCAN	p.E1167Q	NM_004386	p.E1167Q	51		
12	Del9q solely	11	32414242	WT1	p.Q437X	NM_024426	p.Q437X	72		
13	Del9q solely	12	9840618	CLEC2D	p.S98X	NM_013269	p.S98X	54	yes	
13	Del9q solely	19	17091360	CPAMD8	p.R558L	NM_015692	p.R558L	50	yes	

13	Del9q solely	2	25457242	DNMT3A	p.R882H	NM_175629	p.R882H	50	yes	
13	Del9q solely	5	150934102	FAT2	p.S1256C	NM_001447	p.S1256C	12	yes	
13	Del9q solely	12	25398284	KRAS	p.G12D	NM_033360	p.G12D	46	yes	
13	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	29		
13	Del9q solely	15	41750007	RTF1	p.M199fs	NM_015138	p.M199fs	44	yes	
14	Del9q solely	2	25457243	DNMT3A	p.R882C	NM_175629	p.R882C	42		
14	Del9q solely	5	170837543	NPM1	p.L287fs	NM_002520	p.L287fs	38		
14	Del9q solely	12	112888211	PTPN11	p.E76G	NM_080601	p.E76G	39		
15	Del9q t(8;21)	7	87180129	ABCB1	p.A342V	NM_000927	p.A342V	29		
16	Del9q t(8;21)	20	31022441	ASXL1	p.G642fs	NM_015338	p.G642fs	16		
16	Del9q t(8;21)	1	109793756	CELSR2	p.R352H	NM_001408	p.R352H	45		yes
16	Del9q t(8;21)	11	32417910	WT1	p.S381X	NM_024426	p.S381X	28		
17	Del9q t(8;21)	20	31022441	ASXL1	p.G642fs	NM_015338	p.G642fs	24		
17	Del9q t(8;21)	1	115256529	NRAS	p.Q61R	NM_002524	p.Q61R	36		
17	Del9q t(8;21)	11	32421544	WT1	p.C350R	NM_024426	p.C350R	42		
18	Del9q t(8;21)	2	25972652	ASXL2	p.R591fs	NM_018263	p.R591fs	28		yes
18	Del9q t(8;21)	1	109815537	CELSR2	p.S2742delinsSEEEE	NM_001408	p.S2742delinsSEEEE	36		yes
19	Del9q t(8;21)	3	113138899	WDR52	p.R179X	NM_018338	p.R179X	40		yes
20	Del9q other	21	44524456	U2AF1	p.S34F	NM_006758	p.S34F	11		
21	Del9q other	13	28602341	FLT3-other	p.N676T	NM_004119	p.N676T	15		
21	Del9q other	4	106156729	TET2	p.R544X	NM_017628	p.R544X	24		
21	Del9q other	17	7577539	TP53	p.R209G	NM_001276761	p.R209G	82		
22	Del9q other	1	109806955	CELSR2	p.G1753S	NM_001408	p.G1753S	57		yes
22	Del9q other	17	7577539	TP53	p.R209W	NM_001276761	p.R209W	30		
23	Del9q other	5	150889735	FAT2	p.G3969D	NM_001447	p.G3969D	48		
23	Del9q other	13	28608246	FLT3-ITD	p.E604delinsREYEYDLKWE	NM_004119	p.E604delinsREYEYDLKWE	9		
23	Del9q other	13	28608268	FLT3-ITD	p.E596delinsDFYVDFRE	NM_004119	p.E596delinsDFYVDFRE	10		
23	Del9q other	13	28608268	FLT3-ITD	p.E596delinsEMVQVTGSSDNEYFYVDFRE	NM_004119	p.E596delinsEMVQVTGSSDNEYFYVDFRE	24		
23	Del9q other	13	28608276	FLT3-ITD	p.F594delinsDGTGDRLLRX	NM_004119	p.F594delinsDGTGDRLLRX	2		
24	Del9q other	19	33792350	CEBPA	p.L324Q	NM_004364	p.L324Q	59		
24	Del9q other	19	33792365	CEBPA	p.S319delinsKQRNVETQQKVLELTS	NM_004364	p.S319delinsKQRNVETQQKVLELTS	25		
24	Del9q other	11	32456465	WT1	p.E143X	NM_024424	p.E143X	46		
25	Del9q other	20	31022403	ASXL1	p.630_637del	NM_015338	p.630_637del	31		
25	Del9q other	2	209113113	IDH1	p.R132C	NM_005896	p.R132C	32		

25	Del9q other	15	90630490	IDH2	p.Y274C	NM_002168	p.Y274C	48		
25	Del9q other	17	74732959	SRSF2	p.P95R	NM_001195427	p.P95R	29		
26	Del9q other	13	28623571	FLT3-other	p.T329N	NM_004119	p.T329N	45		
26	Del9q other	11	32417910	WT1	c.1142_1142delinsAGGT	NM_024426	c.1142_1142delinsAGGT	9		
26	Del9q other	11	32417942	WT1	p.R370fs	NM_024426	p.R370fs	35		

Supplementary Table S7. Detailed cytogenetic results of molecular characterized patients.

ID	International System for Human Cytogenetic Nomenclature (ISCN)
1	46,XX,del(9)(q22)[8]/46,XX [18]
2	46,XX,del(9)(q22)[2]/46,XX[23]
3	46,XX,del(9)(q22q34)[16]/46,XX[4]
4	46,XY,del(9)(q22)[6]/46,XY[19]
5	46,XX,del(9)(q22q34)[19]/46,XX[1]
6	46,XY,del(9)(q22)[20]
7	46,XX,del(9)(q22q34)[5]/46,XX[20]
8	46,XY,del(9)(q11q32)[19]/46,XY[1]
9	46,XX,del(9)(q21)[20]
10	46,XX,del(9)(q21q34)[6]/46,XX[16]
11	46,XX,del(9)(q21q34)[15]
12	46,XX,del(9)(q22)[4]/46,XX[13]
13	46,XY,del(9)(q22q34)[19]/46,XY[1]
14	46,XY,del(9)(q22q32)[22]/46,XY[3]
15	46,XX,der(8)t(8;21)(q22;q22),del(9)(q22),der(21)t(8;21)(q22;q22) ins(21;9)(q22;q22q34) [10]/46,XX [11]
16	46,XY,t(8;21)(q22;q22),del(9)(q22)[20]
17	46,XY,t(8;21)(q22;q22),del(9)(q22)[18]/46,XY[2]
18	46,XY,t(8;21)(q22;q22),del(9)(q22)[20]
19	46,XY,t(8;21)(q22;q22)[1]/46,idem,del(9)(q22)[12]/45,idem,-Y,del(9)(q22)[4]
20	48~51,XY,+8,del(9)(q22),+1~4mar[cp17]
21	46,XY,del(5)(q13q33)[2]/63,XX,+1,del(5)(q13q33),+del(5)(q13q33),+8,+9, +del(9)(q21),+10,+11,+i(11)(q10),+12,+13,+14,+16,+19,+20,+21,+22,+22[9]/46,XX[1]
22	46~47,XX,der(2)t(2;8)(q35;?),der(4)t(4;9)(q12;?),der(5)t(5;()(q31;?),del(7)(q22), der(8)t(8;20)(p21;?)t(2;8)(q35;q12),der(9)t(9;21)(?;?),+del(9)(q11),t(19;20)(q11;q13), der(20)t(5;20)(?;q11)[cp15]
23	47,XX,+6,del(9)(q21)[20]
24	46,XY,del(9)(q13q32)[13]/47,XY,+21[5]/46,XY[2]
25	47,XY,+10[7]/46,XY,del(9)(q22)[3]/46,XY[11]
26	46,XX,i(7)(p10),del(9)(q13)[17]/46,XX[5]

Supplementary Table S8. Genes within significant regions as identified by MACAT.

ProbeSet ID	Cytoband	Gene Symbol	Gene Description	Score	p-Value
220925_at	9q22.31	FAM120A	family with sequence similarity 120A	-1.6	0.017
219362_at	9q22.31	FAM120A	family with sequence similarity 120A	-1.78	0.002
217771_at	9q22	NFIL3	nuclear factor, interleukin 3 regulated	-0.42	0.561
209273_s_at	9q21.3-q22	GAS1	growth arrest-specific 1	-1.46	0.04
209274_s_at	9q21.3-q22	GAS1	growth arrest-specific 1	-1.35	0.068
221425_s_at	9q22.31	AUH	AU RNA binding protein/enoyl-CoA hydratase	-1.89	0.008
242776_at	9q22.31	PHF2	PHD finger protein 2	-2.15	0.001
238800_s_at	9q22	SYK	spleen tyrosine kinase	-1.73	0.002
236243_at	9q22	SYK	spleen tyrosine kinase	-1.27	0.041
236155_at	9q21.33	ISCA1	iron-sulfur cluster assembly 1	-1.76	0.009
220933_s_at	9q21.33	ISCA1	iron-sulfur cluster assembly 1	-1.54	0.024
204456_s_at	9q22.31	FAM120A	family with sequence similarity 120A	-0.01	0.972
204457_s_at	9q22.31	PHF2	PHD finger protein 2	-0.86	0.21
228187_at	9q21.33	GOLM1	golgi membrane protein 1	0.05	0.913
238463_at	9q21.33	NAA35	N(alpha)-acetyltransferase 35, NatC auxiliary subunit	-0.82	0.048
229761_at	9q22.2	DIRAS2	DIRAS family, GTP-binding RAS-like 2	-0.43	0.324
219619_at	9q12	BARX1	BARX homeobox 1	-0.29	0.33
240122_at	9q21.33	NAA35	N(alpha)-acetyltransferase 35, NatC auxiliary subunit	0.07	0.843
244023_at	9q21	ZCCHC6	zinc finger, CCHC domain containing 6	-0.83	0.151
226068_at	9q21.33	ISCA1	iron-sulfur cluster assembly 1	-1.66	0.014
209269_s_at	9q22.31	FAM120AOS	family with sequence similarity 120A opposite strand	-1.07	0.092
207540_s_at	9q22	SYK	spleen tyrosine kinase	-2.52	0.001
205052_at	9q22.32	MIRLET7D	microRNA let-7d	-1.69	0.017
203574_at	9q21.3-q22	GAS1	growth arrest-specific 1	-1.04	0.145
225395_s_at	9q22.32	PTPDC1	protein tyrosine phosphatase domain containing 1	-2.18	0
230308_at	9q21	LOC440173	uncharacterized LOC440173	-1.01	0.023
239391_at	9q22.31	FAM120AOS	family with sequence similarity 120A opposite strand	-1.32	0.005
210516_at	9q21	ZCCHC6	zinc finger, CCHC domain containing 6	-0.64	0.066
200774_at	9q21	ZCCHC6	zinc finger, CCHC domain containing 6	-1.97	0.004
200767_s_at	9q21	LOC100506834	uncharacterized LOC100506834	-1.75	0
207138_at	9q22.32	ZCCHC6	zinc finger, CCHC domain containing 6	-0.37	0.401
212726_at	9q22.31	PTPDC1	protein tyrosine phosphatase domain containing 1	-1.45	0.01
219845_at	9q22.2	FAM120AOS	family with sequence similarity 120A opposite strand	-0.57	0.201
229517_at	9q21	DIRAS2	DIRAS family, GTP-binding RAS-like 2	-0.97	0.04

238841_at	9q22	ZCCHC6	zinc finger, CCHC domain containing 6	-1.38	0.005
227793_at	9q22.31	SYK	spleen tyrosine kinase	-1.04	0.059

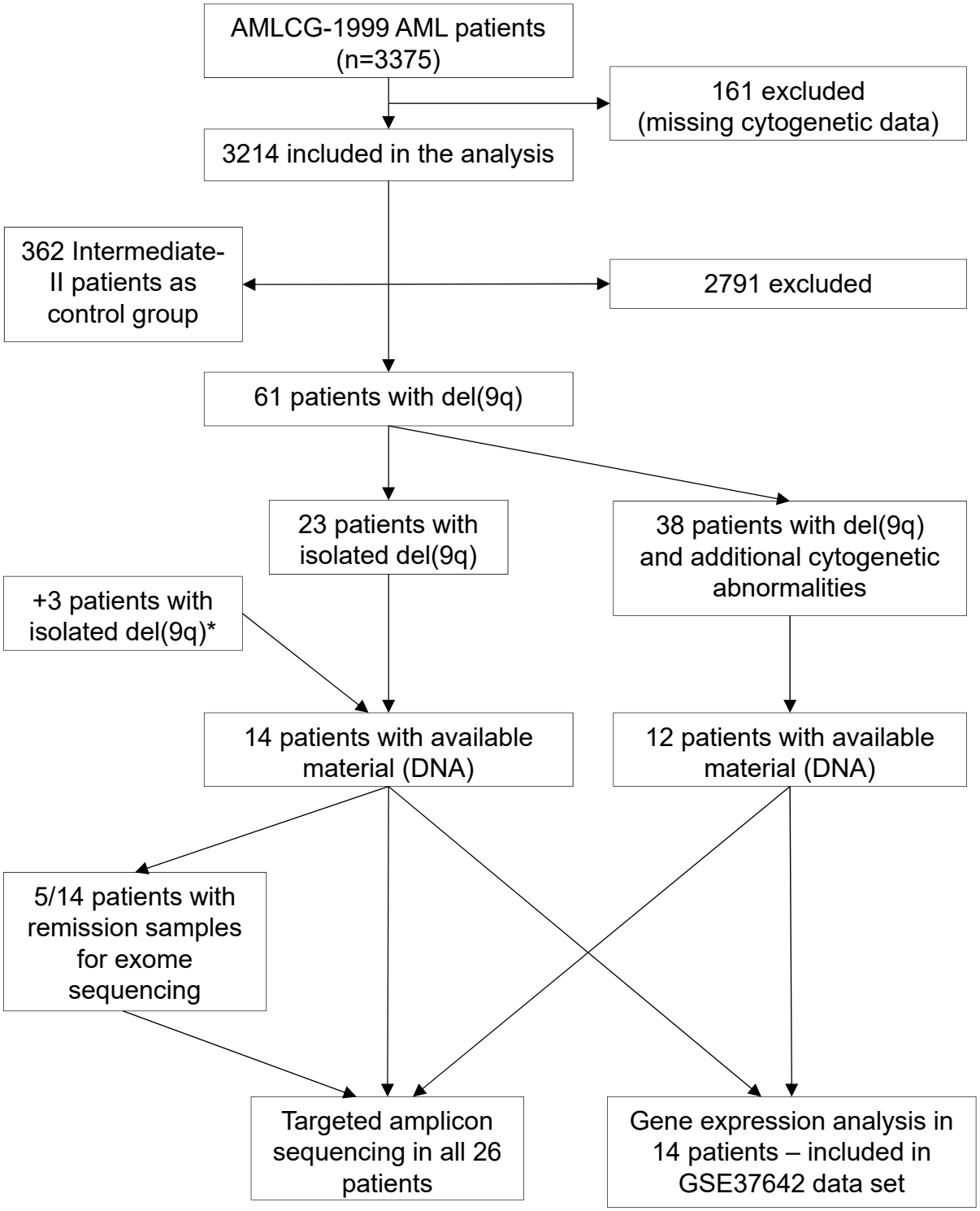
Supplementary Table S9. Percentages of metaphases with del(9q) and variant allele frequencies of *DNMT3A*, *NPM1* or *WT1* mutations.

ID	Percentages of del(9q) metaphases	DNMT3A VAF (%)	NPM1 VAF (%)	WT1 VAF (%)
1	31	0	0	18
2	8	29	0	0
3	80	0	58	0
4	24	44	30	5
5	95	46	39	0
6	100	18	0	0
7	20	0	0	0
8	95	40	0	0
9	100	0	45	0
10	27	0	0	93
11	100	0	38	0
12	24	46; 44*	0	72
13	95	50	29	0
14	88	42	0	0
15	48	0	0	0
16	100	0	0	28
17	90	0	0	41
18	100	0	0	0
19	94	0	0	0
20	74	0	0	0
21	75	0	0	0
22	100	0	0	0
23	100	0	0	0
24	65	0	0	46
25	14	0	0	0
26	77	0	0	34; 7; 8; 7; 9*

* More than one clone

Abbreviations: VAF, variant allele frequencies

Figure S1: Flow Chart

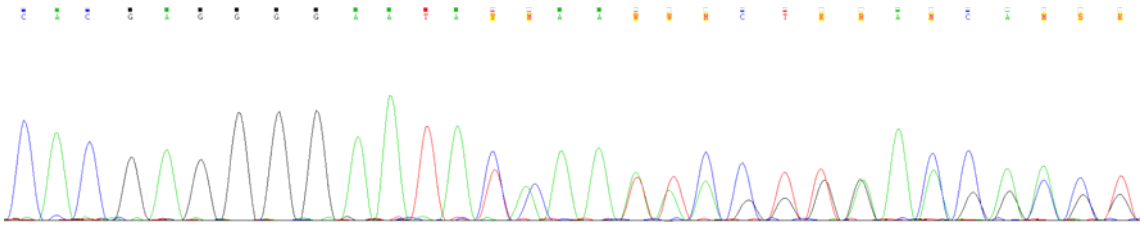


*The biological analysis included three additional patients that were just registered in the AMLCG-1999 trial and were not randomized or died before receiving first treatment.
All patients included in the clinical analysis received intensive induction chemotherapy.

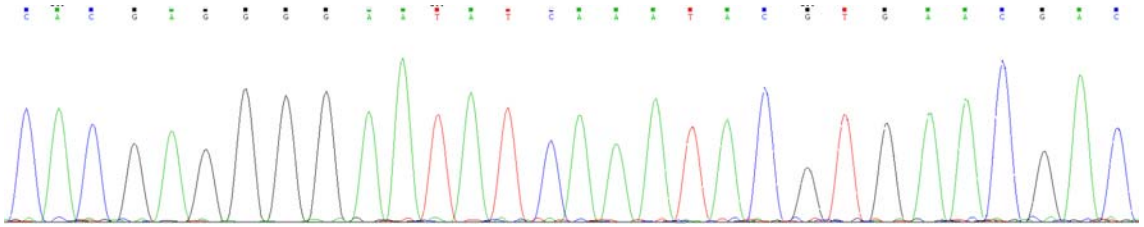
Figure S2: Sanger sequencing of *DNMT3B*

Patient ID 1 (del 9q soley): Diagnosis

DNMT3B (NM_006892) c.1865delT; p.I622fs



Patient ID 1 (del 9q soley): Remission



Patient with AML t(8;21): Diagnosis

DNMT3B (NM_006892) c.1711C>T; p.R571X: validated

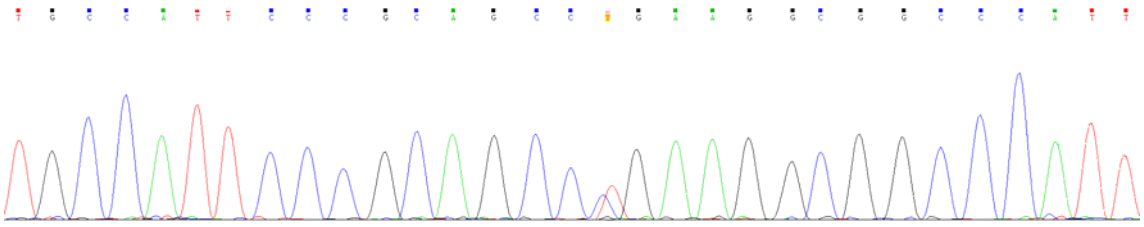


Figure S3: Relapse free and overall survival of AML patients with del(9q) with or without mutations in *WT1*

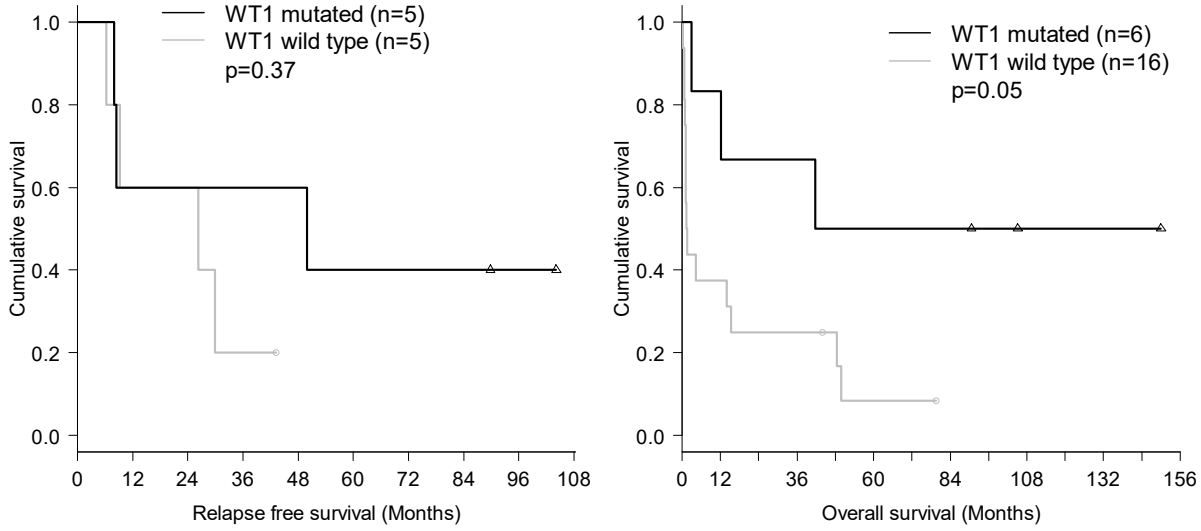


Figure S4: Copy number alterations of *TLE1*, *TLE4* and the CDR of del(9q)

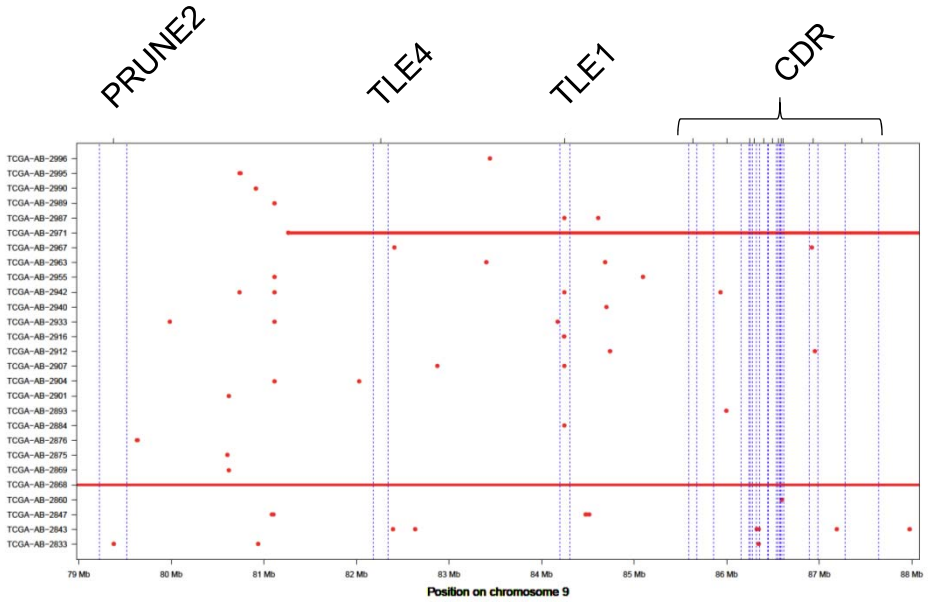


Figure S5: Differential gene expression of AML with del(9q)

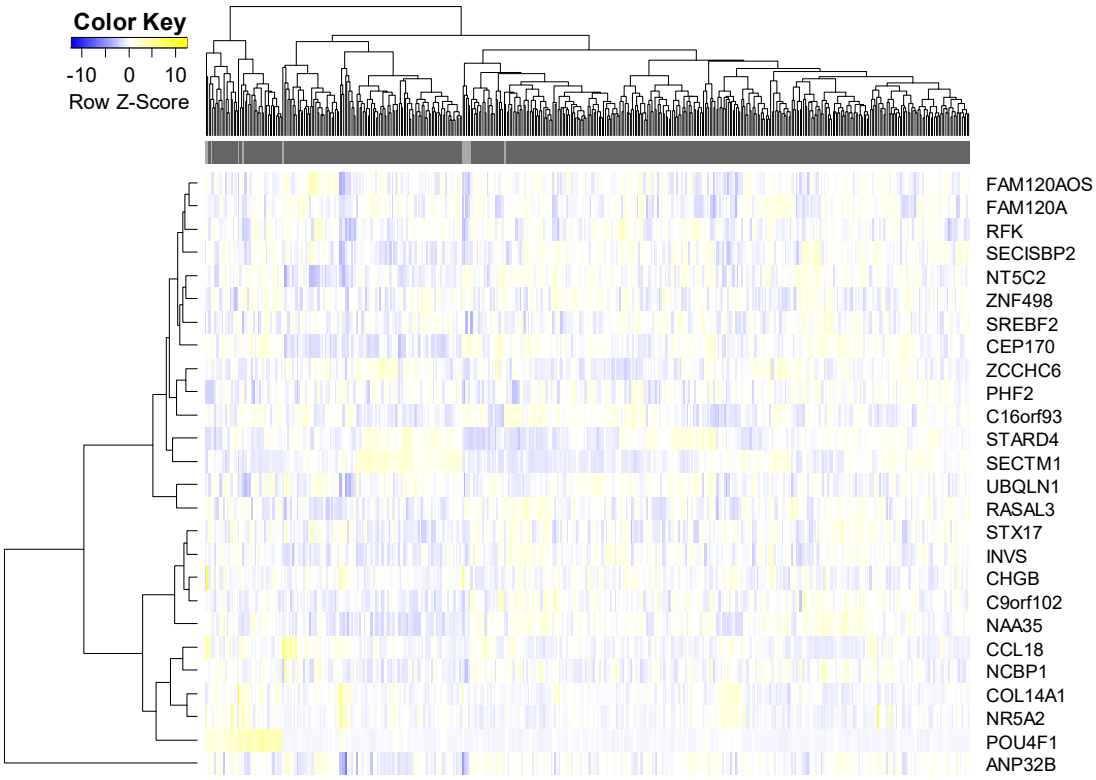


Figure S6: Regional gene expression on chromosome 9 in AML with del(9q)

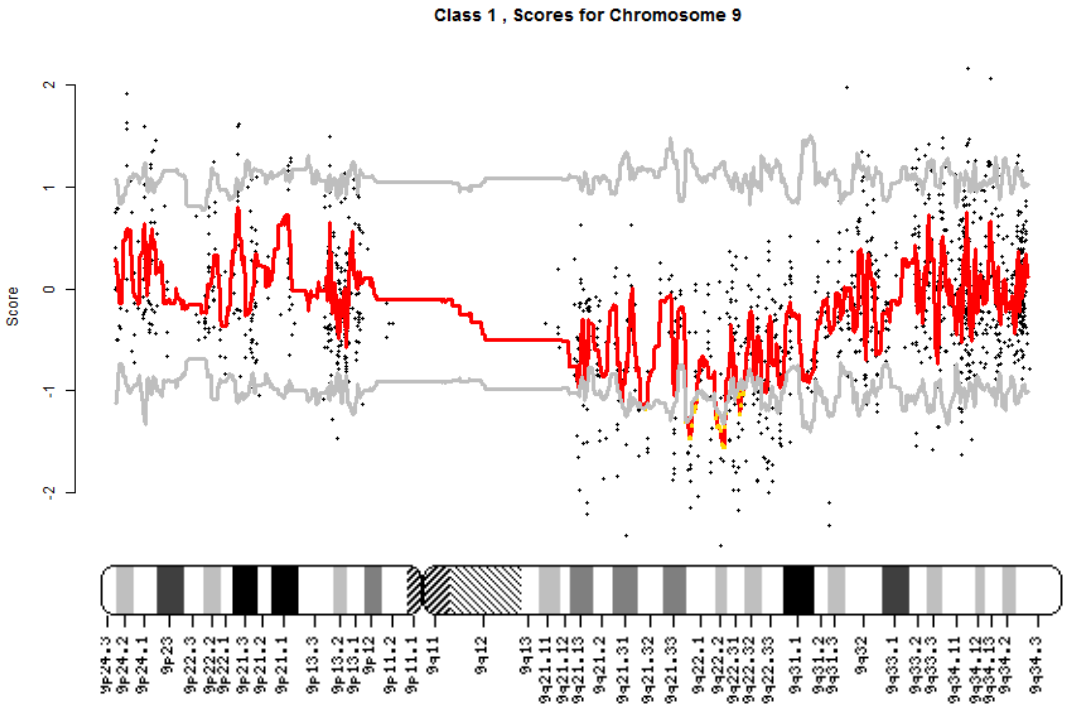


Figure S7: Percentages of metaphases with del(9q) and variant allele frequencies of *DNMT3A*, *NPM1* or *WT1* mutations