Genomic landscape and clonal evolution of acute myeloid leukemia with t(8;21): an international study on 331 patients

Friederike Christen¹*, Kaja Hoyer¹*, Kenichi Yoshida²*, HsIn-An Hou³*, Nils Waldhueter¹, Michael Heuser⁴, Robert K. Hills⁵, Willy Chan¹, Raphael Hablesreiter¹, Olga Blau¹, Yotaro Ochi², Piroska Klement⁴, Wen-Chien Chou³, Igor-Wolfgang Blau¹, Jih-Luh Tang³, Tomasz Zemojtel⁶, Yuichi Shiraishi⁷, Yusuke Shiozawa², Felicitas Thöl⁴, Arnold Ganser⁴, Bob Löwenberg⁸, David C. Linch⁹, Lars Bullinger¹,¹⁰, Peter J.M. Valk⁸, Hwei-Fang Tien³, Rosemary E. Gale⁹, Seishi Ogawa²§, and Frederik Damm¹,¹⁰§

1. Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Hematology, Oncology, and Tumor Immunology, Berlin, Germany
2. Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan
3. Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan
4. Department of Hematology, Hemostasis, Oncology and Stem Cell Transplantation, Hannover Medical School, Hannover, Germany
5. Centre for Trials Research, Cardiff University, Cardiff, United Kingdom
6. Berlin Institute of Health (BIH) Core Genomics Facility, Charité, University Medical Center, Berlin, Germany
7. Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan
8. Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands
9. Department of Haematology, University College London Cancer Institute, London, United Kingdom
10. German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Heidelberg, Germany

* These authors contributed equally to this work as first authors.
§ These authors share senior authorship.

Running title: Landscape and clonal evolution of AML with t(8;21)

Keywords: AML, core binding factor, clonal heterogeneity, ASXL2, ZBTB7A
Christen et al. Landscape and clonal evolution of AML with t(8;21)

Abstract word count: 246
Main text word count: 3866
Tables: 1
Figures: 6

Corresponding author:
Frederik Damm, Department of Hematology, Oncology, and Tumor Immunology, Charité, University Medical Center, Augustenburger Platz 1, 13353 Berlin, Germany. Tel: +49-30-450-553615, Fax: +49-30-450-559975, Email: frederik.damm@charite.de

Key points

- RAS/RTK mutations occur in 63% of patients with t(8;21) AML and confer poor prognosis

- One third of patients with t(8;21) AML who relapse are genetically distinct from diagnostic findings
Abstract

Acute myeloid leukemia (AML) with t(8;21)(q22;q22) is characterized by considerable clinical and biological heterogeneity leading to relapse in up to 40%. We sequenced coding regions or hotspot areas of 66 recurrently mutated genes in a cohort of 331 t(8;21) patients. At least one mutation in addition to t(8;21) was identified in 95% with a mean of 2.2 driver mutations per patient. Recurrent mutations occurred in genes related to RAS/RTK signaling (63.4%), epigenetic regulators (45%), cohesin complex (13.6%), MYC signaling (10.3%), and the spliceosome (7.9%). Our study identified mutations in previously unappreciated genes GIGY2F, DHX15, and G2E3. Based on high mutant levels, pairwise precedence, and stability at relapse, epigenetic regulator mutations were likely to occur before signaling mutations. In 34% of RAS/RTK$^{\text{mutated}}$ patients, we identified multiple mutations in the same pathway. Deep sequencing (~42,000x) of 126 mutations in 62 CR samples from 56 patients identified 16 persisting mutations in 12 patients, of whom five lacked RUNX1-RUNX1T1 in qPCR analysis. KIT$^{\text{high}}$ mutations defined by a mutant level ≥25% were associated with inferior relapse-free survival (HR=1.96; 95%-CI 1.22-3.15; $P=.005$). Together with age and white blood cell counts, JAK2, FLT3-ITD$^{\text{high}}$, and KIT$^{\text{high}}$ mutations were identified as significant prognostic factors for overall survival in multivariate analysis. Whole-exome sequencing was performed on 19 paired diagnosis, remission, and relapse trios. Exome-wide analysis showed an average of 16 mutations with signs of substantial clonal evolution. Based on the resemblance of diagnosis and relapse pairs, genetically stable (n=13) and unstable (n=6) subgroups could be identified.
Introduction

Core binding factor (CBF) acute myeloid leukemia (AML), the most common cytogenetic subtype of AML, is defined by the presence of t(8;21)(q22;q22) or inv(16)(p13q22)/t(16;16)(p13;q22). AML with t(8;21)(q22;q22) is recognized by the World Health Organization as a unique entity within the category of “AML with recurrent genetic abnormalities”\(^1\). Compared with other cytogenetic subsets of AML, patients with t(8;21) are considered a favorable risk group according to their high remission and survival rates\(^2\).

The t(8;21)(q22;q22) fuses the *RUNX1* gene located on chromosome 21 to the *RUNX1T1* gene located on chromosome 8 leading to the *RUNX1-RUNX1T1* fusion gene (also known as *AML1-ETO* fusion)\(^3\). The functional consequences of this fusion have been well studied over the last decade\(^4,5\). Unequivocally, *RUNX1-RUNX1T1* inhibits wildtype *RUNX1* function in a dominant negative manner by competing for heterodimerization and DNA binding at Runx-binding sites\(^5\). Although recognized as a leukemia-initiating event, the fusion is not sufficient to induce leukemia in murine or human hematopoietic cells\(^6-8\), suggesting additional genetic aberrations are required for overt full-blown leukemia. Recently, massively parallel sequencing studies have unraveled a long list of somatic gene mutations underlying the leukemogenesis of t(8;21) AML. In addition to known mutations in the RAS (*K/NRAS*) and in tyrosine kinase signaling pathways (RAS/RTK; *CBL, FLT3, JAK2, KIT, PTPN11*), chromatin modifiers/epigenetic regulators of transcription (*ASXL1/2, BCOR/L1, TET2*), MYC signaling (*MGA, MYC, CCND1/2*), and components of the cohesin complex are recurrently altered\(^9-19\). Some of these genetic aberrations, such as *KIT* mutations and *FLT3-ITD*, have been reported to confer poor prognosis in some but not all studies\(^20\). It has been suggested that this poor prognosis is restricted to patients with a high allele burden of these mutant genes\(^21\). Previous studies have been limited by
relatively small sample numbers, a focus on limited gene sets, or had pooled both pediatric and adult patients. We therefore embarked on a comprehensive molecular study on 331 adult AML patients with t(8;21) using a 66-gene targeted sequencing panel approach to more definitively review these issues which are potentially relevant to the choice of therapy.

Current knowledge about genetic aberrations and patterns of clonal evolution at relapse of t(8;21) AML is limited. Two studies investigated paired diagnosis/relapse specimen (5 and 7 cases, respectively)\(^9,17\). Whole-exome sequencing (WES) demonstrated dynamic patterns of clonal evolution during disease progression, including loss, retention or gain of mutations. However, no relapse-specific genetic aberrations have been identified to date. To provide further insight into disease progression we have therefore carried out WES of 19 paired diagnosis, complete remission (CR), and relapse trios tracking the different genetic clones during disease evolution.
Methods

Patients

Bone marrow or peripheral blood samples from 331 adolescent/adult patients (age, 15-84 years) with t(8;21) AML were collected from collaborating institutions in Germany (n=65), the Netherlands (n=50), Taiwan (n=80), and the United Kingdom (n=136). Over 75% of patients were enrolled into multicenter treatment trials of the German AML study group (AML SHG 0199, -0295, AMLSG 07-04), HOVON/SAKK protocols -04, -29, -42, -42A, and -102, and UK MRC AML10, AML12 and AML15 trials. Details of treatment protocols have been previously reported\textsuperscript{22-33}. All but five patients were treated with an intensive cytarabine/anthracycline induction backbone and subsequent cytarabine based consolidation therapy. Cumulative dosages of high-dose cytarabine varied between the different treatment strategies: AMLSG (18-54 g/m\textsuperscript{2}), HOVON/SAKK (13-25 g/m\textsuperscript{2}), MRC (1-45 g/m\textsuperscript{2}), and Taiwan (48-64 g/m\textsuperscript{2}). 27 patients received allogeneic hematopoietic stem cell transplantation (HSCT) in first remission. Written consent was obtained in accordance with the Declaration of Helsinki and with ethical approval obtained from the local ethics committees.

Mutation analysis by targeted sequencing

305 patients were screened for 66 genes recurrently mutated in AML using a customized version of the TruSight Myeloid Sequencing panel (Supplemental Table S1) according to the manufacturer’s instruction. Libraries were paired-end sequenced with a mean sequencing depth of ~4000x on a NextSeq and of ~1000x on a MiSeq sequencer (Supplemental Figure S1). Variants were detected with a variant allele frequency (VAF) cutoff of 5%. Of all detected variants, 38% were validated by amplicon-based targeted deep sequencing (TDS; n=243), digital droplet PCR
(ddPCR; n=27), or Sanger sequencing (n=9) as previously described\textsuperscript{34-38}. Variant calling and criteria used for candidate selection are detailed in supplemental methods.

**Other cytogenetic and molecular analyses**

The t(8;21) rearrangement was determined by karyotyping (as well as additional cytogenetic abnormalities) and/or fluorescence in situ hybridization and/or evidence of \textit{RUNX1-RUNX1T1} fusion transcripts, as previously described\textsuperscript{39,40}. The screening for \textit{FLT3}-internal tandem duplication (ITD) was performed by Genescan-based fragment analysis enabling quantification of the allelic ratio (AR)\textsuperscript{41-43}.

**Whole-exome sequencing and clonal evolution**

WES with an average reading depth of 138x was performed for 26 paired diagnosis/CR samples (Supplemental Table S2). Of these, 19 patients were also sequenced at first relapse (Rel1), and in four also a second relapse (Rel2) was available. WES was performed using SureSelect Human All Exon V5 and V6 (Agilent Technologies, Santa Clara, CA, USA) enrichment followed by sequencing on the HiSeq 2500 platform (Illumina, San Diego, CA, USA)\textsuperscript{44-46}. For patients suffering from two relapses, clonal evolution was evaluated by amplicon-based TDS of all single nucleotide variations (SNVs) detected at diagnosis, CR1, Rel1, CR2, and Rel2 (n=650 SNVs).

**Statistical analysis**

To determine the sequential order of acquired mutations we applied Bradley-Terry maximum likelihood model fitting to cases with $\geq$2 mutations (n=212). The gene \textit{MGA} was used as reference. We used Pearson Goodness-of-Fit test to test whether
mutations in RAS/RTK signaling genes showed statistical evidence for clonal heterogeneity. We assumed that mutations occurred in different clones when $P<.05$.
The definition of CR, overall survival (OS), and relapse-free survival (RFS) followed recommended criteria$^{47}$. Primary analysis was performed on OS. Sensitivity analyses were performed on CR and RFS. Pairwise comparisons of variables for exploratory purposes were performed using Mann-Whitney or chi-squared tests. Kaplan-Meier analysis was performed to construct survival curves and log-rank test was applied to evaluate differences between subgroups. Logistic regression models including treatment strategies of contributing study groups trial as a covariable were used to study associations between individual variables and CR rate. Cox models stratified for treatment strategies of contributing study groups used for time to-event variables (OS and RFS), and $P$-values were calculated using the Wald test. Multivariate logistic regression models were constructed for factors associated with achievement of CR, and multivariate Cox proportional hazards models were used to study factors associated with survival endpoints. Cytogenetic aberrations and gene mutations were included in multivariate models if detected in $\geq10$ patients with an unadjusted univariate $P\leq.1$. As mutation burden has been shown to carry important prognostic value in t(8;21) AML$^{21}$, we investigated the impact of mutant levels of FLT3-ITD and KIT mutations. Using maximally selected log-rank statistics, allelic ratios $\geq0.35$ and mutant levels $\geq25\%$ were defined as $\text{FLT3-ITD}^\text{high}$ or $\text{KIT}^\text{high}$, respectively. To provide quantitative information on the relevance of results, 95% confidence intervals (CIs) of hazard ratios (HR) were computed. Two-sided $P$-values $\leq.05$ were considered significant in the primary analysis, and as indicators for a trend in all additional analyses. All analyses were carried out using the software package of SPSS Version 23.0 (IBM, Armonk, NY).
Results

Mutation and variant allele spectrum in t(8;21) AML

Sequencing of 331 diagnostic t(8;21) AML samples revealed a total of 729 mutations (Supplemental Tables S3 and S4). Mutations were found in 49 genes, of which 42 were recurrently mutated (Figure 1A). In 316 patients (95%), we found at least one mutation with a mean of $2.2 \pm 1.4$ mutations per patient (Supplemental Figure S2). 210 patients (63.4%) harbored mutations in RAS/RTK signaling pathways with KIT (27%), NRAS (14.8%), FLT3 (16.9%; 10% of all patients harbored a FLT3-ITD), KRAS (5.7%), and CBL (5%) representing the most common RAS/RTK aberrations. Mutations in the RTK signaling regulator GIGYF2 were found in 5 patients (1.5%) and occurred in a hotspot affecting arginine codons at positions 791 and 792\textsuperscript{48}. Approximately half of all patients (149/331=45%) harbored a mutation in genes involved in epigenetic regulation (chromatin remodeling and DNA methylation): ASXL2 (15.7%), ASXL1 (12.4%), TET2 (7.9%), EZH2 (5.7%), and KDM6A (4.2%). Genes encoding for components of the cohesin complex or the splicing machinery were identified in 13.6% and 7.9% of patients and rarely co-occurred (Supplemental Figure S3). Hotspot mutations in DHX15, an RNA helicase implicated in pre-mRNA splicing, mainly affected codon R222 and occurred in 6% of patients. MYC signaling was altered in 10.3%, mainly due to alterations in CCND2 (4.3%) and MGA (3.6%). Additionally, a high rate of loss of function mutations in ZBTB7A (13%), one of the transcription factors that regulates hematopoietic differentiation, was found\textsuperscript{16}.

In our cohort, we observed a median VAF of 0.28. We detected the highest median VAF in genes involved in DNA methylation, including DNMT3A (0.43), TET2 (0.40), and IDH2 (0.39), suggesting that these mutations are present in the majority of the cells and therefore play a role in early stages of leukemogenesis (Figure 1B). We observed lower median VAFs for mutations of transcription factors (0.19), RAS/RTK
signaling (0.22) and cohesin complex (0.20), especially for those in FLT3 (0.11), KIT (0.21) and the RAS GTPases NRAS (0.16) and KRAS (0.20), indicating that alterations in these genes are later events. Further, we modelled the sequential order of mutation acquisition in pairwise precedences\(^49\). Applying this approach in 212 patients with \(\geq 2\) mutations, we confirmed genes involved in RAS/RTK pathways and chromatin remodeling to be later events, whereas TET2, ZBTB7A, and CCND2 aberrations were acquired earlier (Supplemental Figure S4). We noted that 34% of patients with a RAS/RTK mutation harbored additional alterations of the same pathway. A total of 71 patients harbored \(\geq 2\) signaling mutations, with 35 patients having mutations in different RAS/RTK genes, 22 cases having different variants in the same gene, or both (n=14) (Figure 2A and Supplemental Figure S5). Out of 89 KIT\(^\text{mutated}\) patients, 21 (23.6%) showed additional KIT mutation(s) resulting in the highest double mutation rate in one gene. We found 113 KIT mutations in 89 patients, of which 86 mutations were located in exon 17, mainly SNVs at D816 or N822. In exon 8 we found 17 mutations, mainly in-frame insertions/deletions at positions 416 to 419. (Figure 1C). Most of the patients (95%) with multiple KIT mutations harbored at least one exon 17 mutation. To test for clonal heterogeneity, we used Pearson Goodness-of-Fit tests\(^49\). While the sum of VAFs were below 50% and thus too low for prediction tests in 54 cases, statistical evidence for clonal heterogeneity was found in 11 patients, indicating respective mutations occurred in different clones (Figure 2B, C and Supplemental Figure S5). In 6 patients no evidence for clonal heterogeneity could be shown (\(P > .05\)).

**Clinico-biological associations and prognostic impact of gene mutations in t(8;21) AML**
The study population included 331 t(8;21) AML patients with a median age of 41.7 years (Table 1). Most patients were cytomorphologically classified as AML with maturation. While 94% of patients presented with de novo AML, secondary or therapy-related AML were each diagnosed in 3% of patients. JAK2 mutations were significantly less often found in de novo AML (4.8% vs. 26.7%, \( P < .001 \); Supplemental Table S5). Median white blood cell (WBC) count was 17.9 \( \times 10^9 \)/L and highest in cases with a FLT3, KIT, or NRAS mutation. An extramedullary manifestation was present at diagnosis in 13 of 124 patients (10.5%) and correlated with the presence of DHX15 and \( KIT^{high} \) mutations, especially when a KIT mutation located in exon 17 (25.9% vs. 6.2%, \( P = .008 \); Supplemental Table S6). For 316 patients, cytogenetic data from karyotyping were available. Loss of a sex chromosome (LOS) was the most frequent additional aberration (loss of Y: \( n = 108 \); loss of X: \( n = 43 \)). Taking molecular and cytogenetic data together, 99% showed a genetic aberration in addition to the \( RUNX1-RUNX1T1 \) fusion (Figure 3 and Supplemental Table S7).

Clinical follow-up data were available for 322/331 patients. Median follow-up time for patients who remained alive was 6.74 years (range, 0.25 to 21.1 years). Overall, 296 (92.2%) patients reached CR after induction therapy, 17 patients died during induction (5.3%) and 8 experienced primary induction failure (2.5%).

Multivariate logistic regression analysis revealed a reduced CR rate in patients with a cytogenetic aberration of chromosome 7 (-7/del7q; \( P = .026 \)), \( FLT3\)-ITD\( ^{high} \) (\( P = .05 \)), or a JAK2 mutation (\( P = .04 \)). Patients with a higher platelet count obtained more often a CR (\( P = .04 \); Supplemental Table S8). We next evaluated the prognostic impact of clinico-biological variables on RFS. Six variables were selected for multivariate analysis stratified for treatment strategies of contributing study groups. Restricted to male patients, loss of Y was the only independent factor for longer RFS. Together with high WBC counts and \( FLT3\)-ITD\( ^{high} \), \( KIT\) mutated patients with a high mutant level...
≥25% were at higher risk for relapse (HR=1.96; 95%-CI 1.22-3.15; \( P=0.005 \); Supplemental Table S9). With respect to OS, nine variables were included in multivariate analysis. \( \text{KIT}^{\text{high}} \) and \( \text{JAK2} \) mutations were the strongest poor prognostic factors in our cohort (Figure 4). Additionally, age, high WBC counts, trisomy 8, and \( \text{FLT3-ITD}^{\text{high}} \) were also independently associated with inferior OS. In male patients, loss of Y associated with longer survival (HR=0.49; 95%-CI 0.3-0.8; \( P=0.004 \)). Censoring our analysis for patients receiving allogeneic HSCT in CR1 resulted in similar results for RFS and OS. No difference was observed for any survival endpoint with respect to number of signaling mutations or clonal heterogeneity status (Supplemental Figure S6). Collectively, we observed that mutations predicted to occur later in leukemogenesis appear to act as predictors for clinical outcome.

**Profiling of somatic mutations in AML with t(8;21) at diagnosis, CR, and relapse**

WES was performed on paired diagnosis, CR and relapse trios of 19 t(8;21) AML patients. A total of 425 SNVs and small Indels with translational consequence were detected in 385 genes (Supplemental Table S10), of which 19 genes were mutated in more than one patient (Supplemental Figure S7). Only 8 SNVs were located in areas affected by one of the 30 CNAs we detected (Supplemental Table S11). The total number of somatic mutations did not change at relapse, with a mean of 16.6 mutations per patient at diagnosis and 15.5 mutations at relapse. In total 131 mutations (31%) present at diagnosis were lost at relapse and 110 mutations (26%) were selected/acquired during disease progression, while 184 mutations (43%) were found in both diagnosis and relapse samples of patients (Figure 5A). Of note, the t(8;21) fusion was detected in all patients at diagnosis and relapse suggesting common ancestral clones. All but two AML trio samples shared at least one additional genomic alteration at diagnosis and relapse. An increase of C>A and A>T
transversions was observed in relapse-specific as compared with diagnostic mutations ($P=0.06$; Supplemental Figure S8).

In three WES-investigated patients we found persisting mutations in morphologically defined CR samples (VAF range: 5 to 14%). While flow-cytometry based minimal residual disease (MRD) monitoring was negative in one patient, the $RUNX1$-$RUNX1T1$ fusion was detectable at low mRNA levels in all cases (Supplemental Table S12). This prompted us to investigate mutation persistence using amplicon-based ultra-deep sequencing for 126 mutations in 62 CR samples from 56 patients (mean coverage: 41430 reads/amplicon). A total of 16 mutations were identified in 12 samples (VAF range: 0.1% to 12%; Supplemental Table S13), while the $RUNX1$-$RUNX1T1$ fusion was detectable by qPCR during CR in 26 patients. Of note, five patients showed mutational residual disease while the fusion was not detectable (Supplemental Figure S9).

To identify types of clonal evolution, we defined each mutation as stable, if present at diagnosis and relapse, or as lost or gained, if it was detected only at diagnosis or relapse, respectively. Except for AML59, in all patients parts of the individual diagnostic mutation spectrum were lost and replaced with at least one new mutation at relapse. These data suggest that some clones were successfully eradicated during induction chemotherapy while others escaped and/or were selected and subsequently expanded at relapse by acquiring one or more new mutations. When calculating the ratio of the number of stable to gained or lost mutations, we found two evolutionary groups. Patients with more than 40% stable mutations defined group A and cases with over 60% of gained/lost variants were pooled in group B (Figure 5B and Supplemental Figures S10-11). Group B patients ($n=6$) had a dynamic tumor development with relapse samples showing little or no resemblance to the cancer found at diagnosis. In the course of their disease, one major clone replaced another.
For patients from group A, diagnosis and relapse showed a more similar genetic composition (Figure 5C). While clonal evolution seemed to be patient-specific, we noted some gene-related patterns. Mutations in epigenetic regulators and genes involved in cell cycle control were stable or got lost, i.e. they were always found at diagnosis and never acquired just at relapse, further supporting their important role in leukemic initiation. In contrast, mutations in transcription factors, RAS/RTK signaling pathway genes, members of the cohesin complex and splicing machinery were equally often found to be stable, gained or lost (Supplemental Figure S12). When comparing baseline characteristics of relapsing t(8;21) AML patients according to their evolutionary patterns “A” or “B”, no major differences were observed (Supplemental Table S14). Group B patients had a significantly longer OS compared to patients with a more analogous mutation profile (group A) at diagnosis and relapse ($P=0.006$; Supplemental Figure S13). These data warrant further investigations in larger paired diagnosis/relapse cohorts for threshold fine-tuning as also an “intermediate” subgroup might be extracted from our study.

Next, we investigated dynamic evolution of multiclonal RAS/RTK mutations to address biological relevance of clonal heterogeneity at relapse. Already $KIT$ itself, the most frequently mutated gene of the RAS/RTK signaling pathways, showed a very diverse clonal evolution with gained or lost mutations in $7/10$ $KIT^{mutated}$ patients (Supplemental Figure S14). In all five cases with multiple RAS/RTK mutations, the mutations where present in different clones at diagnosis and showed opposing clonal dynamics over time (Supplemental Figure S15).

To obtain more detailed insights into the clonal architecture of t(8;21) AML and its dynamic evolution, we investigated four patients that suffered from a Rel2. To increase accuracy of VAF quantification, variants were re-sequenced with a mean coverage of 39400 reads/amplicon per sample (Supplemental Table S15). All
patients harbored a dominant mutation cluster representing the founding clone as well as at least one subclone at diagnosis. Neither the founding clone nor the major diagnosis subclone, but a second subclone drove Rel1 and was only detectable with low VAFs at diagnosis. In group B patients (AML47 and AML57) (Figure 6A, B and Supplemental Figure S16), this second clone outcompeted the other clones after induction chemotherapy, while the original clones were eradicated. In contrast, the other two patients (AML05 and AML07) kept either the AML-initiating clone and/or the first major subclone as a considerable part of the tumor composition at Rel1 (Figure 6C, D and Supplemental Figure S17). Therefore, it seems likely that these primary clones harbored therapy resistance properties themselves. The high similarity between Rel1 and Rel2 indicates that both relapses were driven by the same clone. However, in Rel2 new subclones appeared or the first major subclone came back up suggesting either ongoing DNA damage due to prolonged effects from cytotoxic therapy and/or clonal selection of very minor subclones that existed below the limit of detection.
Discussion

In this study, we embarked on a comprehensive molecular investigation unravelling molecular genetics underlying diagnosis and relapse of t(8;21) AML in the so far largest cohort comprising 331 patients. In all but four patients, we detected additional genetic aberration(s) cooperating with the RUNX1-RUNX1T1 fusion to overt full-blown leukemia. In line with previous reports, most common additional aberrations were mutations in the RAS/RTK pathways in 63%, alterations of epigenetic regulators in 45% and LOS in 48%. Noteworthy, and in contrast to other AML subtypes such as cytogenetically normal AML, we failed to identify genetic patterns of co-occurrence or exclusivity suggesting no major mutation-based clusters/subgroups subdividing t(8;21) AML (Supplemental Figure S18).

Approximately one fifth of t(8;21) patients had persisting mutations during CR. Surprisingly, some of these patients tested negative for the RUNX1-RUNX1T1 fusion. While recent reports provided evidence that persisting molecular MRD is associated with relapse and poor prognosis in AML, it will be of major interest to couple mutation and fusion quantification to improve patient monitoring in this entity in the future.

Our data show that signaling mutations are rather late events and impact patient outcome: JAK2 and FLT3-ITD high associated with reduced CR rates, KIT high with shorter RFS, and JAK2, FLT3-ITD high and KIT high with shorter OS. A comparable evolutionary mutation pattern has been recently described in transformed MDS patients with sAML-associated mutations (Type I mutations) almost always having lower VAFs than mutations in age-related clonal hematopoiesis and other mutations (Type II mutations). In contrast to their high frequency and early nature in age-related clonal hematopoiesis, ASXL1 mutations are rather late events in t(8;21) AML, which holds also true for its homologue ASXL2. We confirm that ASXL1
mutations do not harbor prognostic impact in t(8;21) and - in line with the current ELN risk classification\(^2\) - should not be used as an adverse marker if co-occurring with t(8;21). RUNX1 mutations, the other mutation marker of the ELN classification, did also not influence patient outcome in this cohort, however due to small patient numbers (n=8) an even larger patient cohorts is necessary for a final verdict.

Observations pointing to the importance of subclonal tumoral heterogeneity have been highlighted in various hematologic malignancies\(^{45,58-61}\). Recently, Itzykson et al. reported that clonal interference in RAS/RTK pathways, a phenomenon defined by the co-existence of clones sharing a common ancestor and harboring independent lesions targeting the same pathway, occurs in up to 36% of t(8;21) AML and is associated with shorter event-free survival\(^{62}\). In that study, clonal interference was defined by the presence of \(\geq 2\) signaling mutations assuming mutations occur in different clones. We observed multiple RAS/RTK mutations in 22% of patients without an impact on CR rate, RFS, or OS. Using prediction tools, evidence for clonal heterogeneity could be identified in some of these cases. However, clonal heterogeneity and interference are not the same as evolutionary independent genetic lesions are required for the latter. Only single-cell experiments will be able to address clonal independency of multiple RAS/RTK mutations. Nevertheless, we observed many expanding RAS/RTK mutations at relapse pointing to the importance of clonal heterogeneity for our understanding of relapse mechanisms. With respect to relapse-specific mutations, few previously unknown genes were identified. In 11% of patients, we discovered a frameshift mutation affecting the E3 ubiquitin ligase G2E3 that has previously been implicated as a regulator of DNA damage response and cell death\(^{63}\). Whether G2E3 and/or other novel mutations such as GIGYF2 or DHX15 might serve for target-specific treatment approaches should be addressed in future studies. While KIT mutations have been associated with unfavorable outcome in several but not all
studies\textsuperscript{20,21,64}, the prognostic importance of \textit{FLT3}-ITD is not very well established in \textit{t(8;21)} AML\textsuperscript{20}. In our large cohort, high mutant levels of both genes were associated with poor prognosis. Thus therapeutic targeting with midostaurin, a multitargeted kinase inhibitor approved for \textit{FLT3}\textsuperscript{mutated} AML\textsuperscript{65}, the multikinase inhibitor dasatinib, which showed encouraging results in a phase Ib/Ila trial when added to intensive chemotherapy in CBF AML patients\textsuperscript{66}, or other RTK inhibitors are promising treatment approaches to improve patient outcome.

In summary, we provide a comprehensive overview on the mutational landscape, the importance of driver genes, and models of genetic relapse. We hope our data will serve as a basis for guided and risk-adapted treatment strategies.
### Tables

**Table 1:** Pretreatment characteristics of 331 t(8;21) AML patients.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>t(8;21) AML (n=331)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>41.7</td>
</tr>
<tr>
<td>range</td>
<td>15-84</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>male - no. (%)</td>
<td>188 (57%)</td>
</tr>
<tr>
<td>female - no. (%)</td>
<td>143 (43%)</td>
</tr>
<tr>
<td>Diagnostic timeframe</td>
<td></td>
</tr>
<tr>
<td>1990-1999 - no. (%)</td>
<td>75 (23%)</td>
</tr>
<tr>
<td>2000-2004 - no. (%)</td>
<td>76 (23%)</td>
</tr>
<tr>
<td>2005-2010 - no. (%)</td>
<td>93 (28%)</td>
</tr>
<tr>
<td>2010-2016 - no. (%)</td>
<td>87 (26%)</td>
</tr>
<tr>
<td>Morphology</td>
<td></td>
</tr>
<tr>
<td>AML with minimal differentiation - no. (%)</td>
<td>9 (3%)</td>
</tr>
<tr>
<td>AML without maturation - no. (%)</td>
<td>28 (9%)</td>
</tr>
<tr>
<td>AML with maturation - no. (%)</td>
<td>260 (84%)</td>
</tr>
<tr>
<td>Acute myelomonocytic leukemia - no. (%)</td>
<td>11 (3%)</td>
</tr>
<tr>
<td>Acute monoblastic/monocytic leukemia - no. (%)</td>
<td>3 (1%)</td>
</tr>
<tr>
<td>missing data – no.</td>
<td>20</td>
</tr>
<tr>
<td>Type of AML</td>
<td></td>
</tr>
<tr>
<td>de novo - no. (%)</td>
<td>305 (94%)</td>
</tr>
<tr>
<td>secondary - no. (%)</td>
<td>10 (3%)</td>
</tr>
<tr>
<td>therapy-related – no. (%)</td>
<td>9 (3%)</td>
</tr>
<tr>
<td>missing data – no.</td>
<td>7</td>
</tr>
<tr>
<td>Additional cytogenetic aberration</td>
<td></td>
</tr>
<tr>
<td>Loss of X chromosome - no.</td>
<td>43</td>
</tr>
<tr>
<td>Loss of Y chromosome - no.</td>
<td>108</td>
</tr>
<tr>
<td>Trisomy 8- no.</td>
<td>10</td>
</tr>
<tr>
<td>-7/del7q - no.</td>
<td>11</td>
</tr>
<tr>
<td>del9q - no.</td>
<td>37</td>
</tr>
<tr>
<td>others - no.</td>
<td>64</td>
</tr>
<tr>
<td>Bone marrow blasts</td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>56.8</td>
</tr>
<tr>
<td>missing data – no.</td>
<td>50</td>
</tr>
<tr>
<td>WBC count</td>
<td></td>
</tr>
<tr>
<td>median - (x10⁹/l)</td>
<td>17.9</td>
</tr>
<tr>
<td>range - (x10⁹/l)</td>
<td>0.9-153</td>
</tr>
<tr>
<td>missing data – no.</td>
<td>10</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td></td>
</tr>
<tr>
<td>median – g/L</td>
<td>8.1</td>
</tr>
<tr>
<td>range – g/L</td>
<td>2.5-15.4</td>
</tr>
<tr>
<td>missing data – no.</td>
<td>60</td>
</tr>
<tr>
<td>Platelet count</td>
<td></td>
</tr>
<tr>
<td>median - (x10⁹/l)</td>
<td>51.5</td>
</tr>
<tr>
<td>range - (x10⁹/l)</td>
<td>3-647</td>
</tr>
<tr>
<td>missing data – no.</td>
<td>40</td>
</tr>
</tbody>
</table>

Abbreviations: AML, acute myeloid leukemia; WBC, white blood cell.
Figure legends

**Figure 1: Overview of mutations identified by targeted sequencing. A)** Histogram showing the frequency of mutated patients per gene for all genes with detected variants. Bars were colored according to the functional category of the gene (Supplemental Table S16). Mutation frequencies were shown over the bar in percent. *FLT3* variants include *FLT3*-ITD and *FLT3*-TKD. **B)** Boxplot showing the median, 25%-quantile, and 75%-quantile of the variant allele frequency (VAF) for all genes that are mutated in more than four patients. The dashed line indicated 50% VAF that was expected for heterozygous mutations. VAFs were corrected for patient specific chromosomal aberrations and sex bias. Genes were color coded according to their assigned category. *FLT3*-ITDs were not included in this analysis. **C)** Pie chart showing the distribution of 113 *KIT* mutations in 89 patients for most frequent codons and exons, with mutation frequencies shown in percentage.

**Figure 2: Mutational landscape and clonal heterogeneity of RAS/RTK signaling genes. A)** Mutational landscape showing patients with mutations in RAS/RTK pathway. The mutation frequency of each gene was shown on the left. **B)** Variant allele frequencies of RAS/RTK pathway mutations with evidence for clonal heterogeneity for three illustrative patients. **C)** Variant allele frequencies of RAS/RTK pathway mutations without evidence for clonal heterogeneity. Clonality prediction was determined using Pearson Goodness-of-Fit test.

**Figure 3: Mutational landscape of 327 mutated t(8;21) AML patients.** The landscape showed all genetic and cytogenetic aberrations for each patient. Mutations
were color coded by mutation type. The histogram on the left showed the frequency distribution of all aberrations. ND=missing cytogenetic data.

**Figure 4: Multivariate analysis with clinical and biological variables.** Gene mutations were included in the Cox regression analysis if they were detected in at least 10 patients and had a univariate $P$-value for OS, not adjusted for multiple comparisons, of $\leq 0.10$. The model was stratified for center treatment strategy. KIT-mutated patients with a VAF $\geq 25\%$ and FLT3-ITD positive patients with an allelic ratio $\geq 0.35$ were defined as $KIT^{\text{high}}$ or $FLT3^{\text{ITD}^\text{high}}$.

**Figure 5: Distribution of somatic mutations in paired diagnosis relapse samples.** **A)** Histogram showing the total number of SNVs per patient. For each patient the number of mutations in diagnosis and relapse were depicted next to each other as either shared (grey), unique to diagnosis (orange) or unique to relapse (blue). **B)** Ratio of stable to unstable (gained or lost) mutations. Patients with $\geq 60\%$ gained/lost mutations were considered group B. **C)** Variant allele frequency (VAF) plots of two exemplary patients from each group (A: AML27, B: AML38) showing the different evolutionary patterns observed from diagnosis to relapse. Known AML driver genes were highlighted. VAFs were corrected for frequent chromosomal aberrations and sex bias. The VAF plots of the remaining 17 patients can be found in the supplement (Supplemental Figures S10 and S11)

**Figure 6: Fish plots showing the clonal evolution from primary tumor to first and second relapse for all double relapse patients.** **A)** Model of clonal progression during course of disease for patient AML47 and **B)** AML57. For founding clone (gray), primary specific clone (yellow), and relapse-enriched clone (blue),
exemplary clone defining genes were shown. The remaining genes and their affiliation to one of the clones as well as a more detailed evolutionary tree can be found in the supplement (Supplemental Figures S16 and S17). For diagnosis, first and second relapse, the proportion of each sub clone was shown in percentage of the whole cell content. C) AML05 and D) AML07 from the group with an analogous mutation profile at diagnosis and relapse (group A) were depicted in the same manner. The RUNX1-RNX1T1 fusion was detected in all time points.
Acknowledgements

This study was supported by grants #DA1787/1-1 from the Deutsche Forschungsgemeinschaft, a Mechthild-Harf research fellowship from the DKMS Giving Life Foundation, grant #2017_EKES.33 from the Else Kröner-Fresenius-Stiftung, grant 17/4 from the Gutermuth Stiftung, a Lady Tata Memorial Trust awarded to F.D., and a DKTK research grant awarded to F.D. and L.B.; W.C. received a fellowship from the Deutsche José Carreras Leukämie-Stiftung. M.H. was supported by DFG grants HE 5240/5-1, HE 5240/6-1. This work was partially sponsored by grants MOST 106-2314-B-002-226-MY3 and 106-2314-B-002-231-MY3 from the Ministry of Science and Technology (Taiwan) and MOHW MOHW107-TDU-B-211-114009 from the Ministry of Health and Welfare (Taiwan). D.C.L. and R.E.G. were supported by Bloodwise. This work was partially sponsored by Project for Cancer Research and Therapeutics Evolution (P-CREATE) from Japan Agency for Medical Research and Development (16cm0106501h0001) to S.O.

Authorship contributions


Conflict of interest: The authors indicated no potential conflicts of interest.
References


Figure 1

(a) Fraction of mutated patients

(b) Variant Allele Frequency

(c) KIT mutations (n=113)
Figure 2

A

% Mutant

% Mutant

n=210

Tyrosine kinase signaling

B

083

0.0

0.2

0.4

0.6

VAF

GIGYF2 (0.35)

KIT (0.34)

NRAS (0.11)

p < 0.001

p < 0.001

0.0

0.2

0.4

0.6

VAF

CBL (0.45)

FLT3 (0.07)

CBL (0.06)

KRAS (0.06)

p < 0.001

p < 0.001

0.0

0.2

0.4

0.6

VAF

CSF3R (0.36)

KIT (0.24)

p = 0.001

0.0

0.2

0.4

0.6

VAF

AML55

KIT (0.09)

NRAS (0.38)

KRAS (0.28)

p = 0.21

0.0

0.2

0.4

0.6

VAF

098

281

295

CSF3R (0.54)

NRAS (0.44)

KIT (0.27)

JAK2 (0.23)

p = 0.32

p = 0.62

0.0

0.2

0.4

0.6

VAF

p = 0.62

p = 0.21

0.0

0.2

0.4

0.6

VAF

RAS/RTK mutation

RAS/RTK mutation

RAS/RTK mutation

C

RAS/RTK mutation

RAS/RTK mutation

RAS/RTK mutation

B

RAS/RTK mutation

RAS/RTK mutation

RAS/RTK mutation

A

RAS/RTK mutation

RAS/RTK mutation

RAS/RTK mutation
<table>
<thead>
<tr>
<th>Parameter</th>
<th>HR</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age: above vs. below median</td>
<td>1.79</td>
<td>[1.22–2.63]</td>
<td>0.003</td>
</tr>
<tr>
<td>WBC: above vs. below median</td>
<td>1.61</td>
<td>[1.10–2.36]</td>
<td>0.015</td>
</tr>
<tr>
<td>Platelets: above vs. below median</td>
<td>0.82</td>
<td>[0.56–1.22]</td>
<td>0.330</td>
</tr>
<tr>
<td>Loss of chr. Y: yes vs. no</td>
<td>0.74</td>
<td>[0.49–1.11]</td>
<td>0.146</td>
</tr>
<tr>
<td>only male cohort</td>
<td>0.49</td>
<td>[0.30–0.80]</td>
<td>0.004</td>
</tr>
<tr>
<td>Trisomy 8: yes vs. no</td>
<td>3.41</td>
<td>[1.51–7.69]</td>
<td>0.003</td>
</tr>
<tr>
<td>KIT&lt;sup&gt;high&lt;/sup&gt;: yes vs. no</td>
<td>2.27</td>
<td>[1.44–3.59]</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>FLT3–ITD&lt;sup&gt;high&lt;/sup&gt;: yes vs. no</td>
<td>3.47</td>
<td>[1.51–7.97]</td>
<td>0.003</td>
</tr>
<tr>
<td>RAD21: mutated vs. wildtype</td>
<td>0.39</td>
<td>[0.12–1.25]</td>
<td>0.113</td>
</tr>
<tr>
<td>JAK2: mutated vs. wildtype</td>
<td>3.26</td>
<td>[1.64–6.48]</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Figure 4
Figure 5

(A) Bar chart showing the number of somatic mutations for different patients grouped into Group A and Group B.

(B) Bar chart for Group B patients, showing stable mutations in green and gain/loss mutations in grey.

(C) Graph for AML38 and AML27, displaying mutation VAF over time with different colors for different types of mutations.
Figure 6
Genomic landscape and clonal evolution of acute myeloid leukemia with t(8;21): an international study on 331 patients


Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.