

# Mutation accumulation in cancer genes relates to nonoptimal outcome in chronic myeloid leukemia

Shady Adnan Awad,<sup>1-4,\*</sup> Matti Kankainen,<sup>1,2,4-6,\*</sup> Teija Ojala,<sup>7</sup> Perttu Koskenvesa,<sup>1</sup> Samuli Eldfors,<sup>8</sup> Bishwa Ghimire,<sup>8</sup> Ashwini Kumar,<sup>8</sup> Soili Kytölä,<sup>6</sup> Mahmoud M. Kamel,<sup>3</sup> Caroline A. Heckman,<sup>5,8</sup> Kimmo Porkka,<sup>1,4,5</sup> and Satu Mustjoki<sup>1,2,4,5</sup>

<sup>1</sup>Hematology Research Unit Helsinki, Department of Hematology, Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland; <sup>2</sup>Department of Clinical Chemistry and Hematology, University of Helsinki, Helsinki, Finland; <sup>3</sup>Clinical and Chemical Pathology Department, National Cancer Institute, Cairo University, Giza, Egypt; <sup>4</sup>Translational Immunology Research Program, University of Helsinki, Helsinki, Finland; <sup>5</sup>iCAN Digital Precision Cancer Medicine and <sup>6</sup>Department of Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland; and <sup>7</sup>Department of Pharmacology, Faculty of Medicine, and <sup>8</sup>Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

## Key Points

- High mutational load and variants in cancer genes predicts nonoptimal treatment outcome and are new independent prognostic markers of CML.
- Dysregulation of DNA repair and the JAK-STAT signaling pathways relates to progression.

Chronic myeloid leukemia (CML) is a myeloproliferative neoplasm accounting for ~15% of all leukemia. Progress of the disease from an indolent chronic phase to the more aggressive accelerated phase or blast phase (BP) occurs in a minority of cases and is associated with an accumulation of somatic mutations. We performed genetic profiling of 85 samples and transcriptome profiling of 12 samples from 59 CML patients. We identified recurrent somatic mutations in *ABL1* (37%), *ASXL1* (26%), *RUNX1* (16%), and *BCOR* (16%) in the BP and observed that mutation signatures in the BP resembled those of acute myeloid leukemia (AML). We found that mutation load differed between the indolent and aggressive phases and that nonoptimal responders had more nonsilent mutations than did optimal responders at the time of diagnosis, as well as in follow-up. Using RNA sequencing, we identified other than *BCR-ABL1* cancer-associated hybrid genes in 6 of the 7 BP samples. Uncovered expression alterations were in turn associated with mechanisms and pathways that could be targeted in CML management and by which somatic alterations may emerge in CML. Last, we showed the value of genetic data in CML management in a personalized medicine setting.

## Introduction

Chronic myeloid leukemia (CML) is a clonal hematopoietic stem cell disorder characterized by a reciprocal translocation between the long arms of chromosomes 9 and 22 that generates a hybrid gene between breakpoint cluster region (*BCR*) and Abelson murine leukemia viral oncogene homolog 1 (*ABL1*).<sup>1</sup> The disease occurs with an incidence rate of ~0.9 to 1.1 cases per 100 000 population per year and is more common in adulthood.<sup>2,3</sup> Its frontline treatment is tyrosine kinase inhibitor (TKI) therapy. This treatment targets the *BCR-ABL1* fusion protein and has enabled almost normal life expectancy for CML patients.<sup>3-5</sup> Regardless of the emergence of second- and third-generation TKIs with more rapid and deeper response, up to 2% to 3% of patients are unresponsive to TKI treatment and progress to the accelerated phase (AP) and/or blastic phase (BP), with a dismal survival of 7 to 11 months.<sup>6,7</sup>

Studies have linked unfavorable outcomes in CML with increased genetic instability.<sup>8-10</sup> Genetic examinations of AP and/or BP samples have identified recurrent point mutations within the *ABL1* kinase domain and in *RUNX1*,<sup>11-13</sup> *IKZF1*,<sup>11,14</sup> *TP53*,<sup>15</sup> and *ASXL1*.<sup>15,16</sup> More recently, specific mutations, especially in epigenetic modifiers, in chronic phase (CP) samples have been linked with an unfavorable

Submitted 9 September 2019; accepted 18 December 2019; published online 11 February 2020. DOI 10.1182/bloodadvances.2019000943.

\*S.A.A. and M.K. contributed equally to this study.

Whole-exome sequencing, targeted, and RNA sequencing data, and deidentified individual participant data are available from the corresponding author upon request.

Regulations pertaining to the authors' ethical permit prohibit deposition of these data in public repositories.

The full-text version of this article contains a data supplement.

© 2020 by The American Society of Hematology

TKI response.<sup>17-20</sup> Mutations in cancer-associated genes at the time of initial diagnosis and their acquisition during treatment are also related to an unfavorable outcome.<sup>19</sup> Nevertheless, knowledge about the prognostic value of genomic data in CML management is still in its infancy.

In this study, we used a combination of whole-exome sequencing (WES) and targeted deep sequencing to characterize CML mutations at diagnosis and in the follow-up/progression phases and explored activation of cellular pathways in CML by using RNA sequencing. We identified recurrent somatic mutations in *ABL1*, *ASXL1*, *RUNX1*, and *BCOR*, especially in BP, and discovered similarities in mutation signatures of AP/BP and acute myeloid leukemia (AML) samples. We found that patients with unfavorable treatment outcomes exhibited more nonsilent mutations of potential relevance to CML at the time of diagnosis and acquired such mutations during the treatment. Using RNA sequencing, we identified new hybrid genes. Expression alterations were identified in known cancer pathways and in DNA repair machinery components, illustrating the value of RNA sequencing in finding treatment targets and pathomechanisms. Here, we highlight the value of integrated precision medicine strategy and describe the use of patient-specific molecular data in CML management.

## Materials and methods

### Patients

A total of 59 patients and 5 unrelated age-matched controls were included in the genomic screening experiment. A total of 12 patients and 4 unrelated controls were included in the RNA-sequencing experiment. All patients were recruited from Helsinki University Hospital (HUU), Helsinki, Finland, and the National Cancer Institute, Cairo University, Giza, Egypt. CML diagnosis and progression were defined according to World Health Organization criteria for myeloid neoplasms.<sup>21</sup> All control samples were collected by HUU. All subjects had given written informed consent in accordance with the Declaration of Helsinki.

### Primary cell isolation

Bone marrow mononuclear cells and peripheral blood mononuclear cells (PMNCs) were isolated by Ficoll-Hypaque density gradient centrifugation. CD34<sup>+</sup> cells and CD3<sup>+</sup> cells were purified with human CD34<sup>+</sup> and CD3<sup>+</sup> cell isolation kits (Miltenyi Biotec, Bergisch Gladbach, Germany). The purity of isolated cells was evaluated by flow cytometry with CD34-fluorescein isothiocyanate and CD3-allophycocyanin staining (BD Biosciences, San Jose, CA). The final pools contained >90% purified cells.

### WES and targeted sequencing

Genomic DNA was extracted from AP/BP samples with a high blast count (median ~45%), control samples without CD34<sup>+</sup> enrichment, and from CP samples with a low blast count (median ~3%) after CD34<sup>+</sup> enrichment. Exome capture was performed with the SureSelect XT Clinical Research Exome kit (Agilent, Santa Clara, CA). Targeted sequencing capture was performed with the SeqCap EZ Comprehensive Cancer Design (Roche NimbleGen, Madison, WI), comprising 578 cancer-associated genes. Sequencing libraries were prepared and subjected to 2 × 100-bp paired-end sequencing on HiSeq instruments (Illumina, San Diego, CA),

according to the manufacturer's instructions. WES data from 11 AML cases have been described earlier.<sup>22</sup>

### RNA sequencing

Total RNA was extracted from 2 AP/BP, all 5 CP, and all 4 control samples after CD34<sup>+</sup> enrichment and, in the case of 5 AP/BP samples, without CD34<sup>+</sup> enrichment. RNA-sequencing libraries were produced by using a ribo-depletion-based approach and sequenced on Illumina instruments with paired-end 2 × 100-bp reads, and possible confounding factors were adjusted in statistical testing, as described in the supplemental Methods.

### Drug-sensitivity and -resistance testing

The oncology compound library consisted of 125 approved anticancer drugs and 127 investigational and preclinical compounds (supplemental Data set 12). The experiment was performed as previously described<sup>22</sup> and as indicated in the supplemental Methods.

### Analysis of genomic and transcriptomic data

Details of bioinformatics analysis are available in the supplemental Methods.

### Validation of hybrid genes

Hybrid genes were validated with single-step RT-PCR, fluorescence in situ hybridization (FISH), and/or karyotyping, as described in the supplemental Methods.

### Statistical analysis

Two-tailed Student *t* test, Mann-Whitney *U* test, Fisher's exact test,  $\chi^2$  test, Spearman correlation test, Pearson correlation test, and Fisher's exact test with simulated *P* value on 1e+07 replicates were computed using GraphPad Prism 7 software or R 3.5.0. The statistical significance of the difference in microbial counts per million (CPM) mapped reads was examined using the 2-tailed Student *t* test with unequal variance in Microsoft Excel.

## Results

### Patient characteristics

Samples from a total of 59 CML patients with a median age of 50 years (range, 24-78) were included in the genomic profiling (Table 1; supplemental Data set 1). These were from 16 patients with samples from AP/BP, 40 patients with samples from CP, and 3 patients with samples from both. The CP patients had median Sokal and Hasford scores of 0.89 (0.51-3.45) and 891 (100-2384), respectively. Twenty CP patients achieved major molecular remission (MMR) by 12 months (classified as optimal responders) and 13 after 12 months (classified as suboptimal responders); 7 patients failed to achieve MMR at any time and/or progressed to AP/BP (classified as poor responders). The median time to MMR was 12 (range, 3-70) months. Five unrelated age-matched subjects served as controls. In addition to genomic screens, samples from 7 AP/BP cases, 5 CP cases, and 4 unrelated healthy controls were analyzed by RNA sequencing. Nine of the CP cases underwent genomic screening, as well. For the remaining 3 CML cases, no patient demographic information was available.

**Table 1. Demographics and clinical characteristics of patients subjected to genomic profiling**

	CP							
	CML	AP/BP	CP-All	Progressive	Poor	All	Response Suboptimal	Optimal
Patients, n	59	16	43*	3	4	33	13	20
Age, median (range), y	50 (24-78)	47.89 (24-74)	51.18 (27-78)	41 (35-46)	45 (27-54)	51 (29-78)	50 (29-67)	59 (29-78)
Sex, male:female, n (%)	32:23 (58:42)	10:6 (63:37)	21:15 (60:40)	1:2 (33:67)	3:1 (75:25)	17:14 (52:48)	7:6 (54:46)	10:8 (56:44)
Disease stage at diagnosis, CP/AP/BP, n (%)	51/3/5 (86:5:9)	8/3/5 (50/19/31)	43/0/0 (100/0/0)	3/0/0 (100/0/0)	4/0/0 (100/0/0)	33/0/0 (100/0/0)	13/0/0 (100/0/0)	20/0/0 (100/0/0)
Hb at diagnosis, median (range), g/L	118 (78-157)	119 (78-157)	117.5 (78-156)	130 (85-144)	93 (93-95)	120.5 (78-156)	105 (78-143)	122 (91-156)
WBC count at diagnosis, median (range), × 10 <sup>9</sup> /L	59 (3.78-365)	55.5 (3.78-317)	67 (14-365)	21.6 (16.9-174.1)	319 (238-345)	46 (14-365)	75 (28-327)	44 (14-365)
Platelet count at diagnosis, median (range), × 10 <sup>9</sup> /L	457 (19-2674)	210.5 (19-1048)	480 (120-2674)	860 (156-2570)	565 (393-584)	480 (120-2674)	325 (120-620)	655 (127-2674)
PB blast at diagnosis, n (%)	1 (0-46)	10 (4-46)	0.5 (0-6)	0 (0-6)	2 (2-3)	0 (0-6)	0 (0-6)	0 (0-4)
BM blast at diagnosis, n (%)	3 (0-90)	45 (6-90)	3 (0-8)	2 (2-3)	3 (2-4)	3 (0-8)	2 (0-6)	3 (0-8)
Spleen size at diagnosis, median (range), cm	13.5 (0-30)	15 (0-30)	13 (0-30)	12 (0-15)	25 (12-30)	12 (0-23)	13.75 (0-23)	9 (0-21)
Sokal score, low/intermediate/high/NA, median (range)	12/10/11/26	0/0/0/16	0.89 (0.51-3.45) 12/10/11/7	1.18 (0.56-7.4) 1/1/1/0	1.02 (0.73-1.89) 1/1/1/0	0.89 (0.51-3.45) 10/8/9/6	0.82 (0.52-1.76) 5/3/2/3	0.9 (0.51-3.45) 5/5/7/3
Hasford score, low/intermediate/high/NA, median (range)	11/13/8/27	0/0/0/16	891 (100-2384) 11/13/8/8	720 (328-1299) 1/2/0/0	908 (199-2176) 1/1/1/0	873 (100-2384) 9/11/5/8	866 (124-1306) 4/5/0/4	1106 (100-2384) 5/6/5/4
Additional cytogenetic abnormalities at diagnosis, n/N (%)	13/66 (23.2)	10/16 (62.5)	3/40 (7.5)	0/3 (0)	1/4 (25)	2/33 (6)	0/13 (0)	2/20 (10)
Frontline treatment, imatinib/2G TKI/Chemotherapy/NA, n	34/21/3/1	10/3/3/0	21/18/0/1	3/0/0/0	2/2/0/0	18/15/0/0	8/5/0/0	10/10/0/0
Clinical responses poor/suboptimal/optimal/NA, n	4/13/20/22	0/0/0/16	4/13/20/3	3/0/0/0	4/0/0/0	0/13/20/0	0/13/0/0	0/0/20/0
Interval between diagnosis and follow-up samples, median (range), mo	—	—	6 (3-30)	—	3 (3)	6 (3-30)	6 (3-30)	4.5 (3-19)
Response at follow-up sampling time, minCyR/PCyR/CCyR/MMR, n	—	—	3/2/15/5	—	2/0/0/0	1/2/15/5	1/1/9/0	0/1/6/5
Time to achieve MMR, median (range) (at 12 mo/after 12 mo/never), mo	—	—	12 (3-70) 18/14/4/4	—	0/0/4/0	12 (3-70) 18/14/0/0	30 (15-70) 0/14/0/0	6 (3-12) 18/0/0/0

CML, all patients in our cohort (n = 59); AP/BP, the blast phase subset (n = 16); CP/CP-All, chronic phase subset (n = 43). Fractionation of CP patients into subsets based on their treatment responses: progressive, patients progressed to AP/BP; poor, patients who failed to achieve MMR at any time; responsive, patients who achieved MMR either by 12 months (optimal) or after 12 months (suboptimal).

2G, second generation; BM, bone marrow; CCyR, complete cytogenetic response; Hb, hemoglobin; MMR, major molecular response; minCyR, minimal cytogenetic response; NA, no available data; PB, peripheral blood; PCyR, partial cytogenetic response; WBC, white blood cell.

\*Clinical data were missing for 3 CP patients.

## Somatic mutations in CML samples

We analyzed 125 samples from 59 CML patients and 4 unrelated healthy individuals (supplemental Figure 1) by using WES (mean, 89×; supplemental Figures 2 and 3) or targeted sequencing (mean, 187×; supplemental Figures 2 and 3). These included 40 diagnostic CP samples taken at the time of the initial presentation, with a low median blast count (~3%), 19 diagnostic AP/BP samples with a high median blast count (~45%) from 16 patients with only AP/BP samples and from 3 patients with CP and AP/BP samples. CP samples were enriched for malignant cells to make them more comparable with AP/BP samples and to detect variants in clones of low abundance. In addition, 1 or more follow-up samples from 28 cases (median, 6 months; range, 3-30), sorted T-cell or remission PMNC samples from 24 cases, and PMNC samples from 5 controls were analyzed. Matched skin samples were available for all WES samples and served as controls.

We identified 501 somatic mutations in 456 genes among the genomic samples (supplemental Data set 2). These included 418 single-nucleotide variants (SNVs; Figure 1B) and 83 frameshifting (Figure 1C) variants. The mean SNV burden was 0.16 and 0.43 per million base pairs (mbp) among WES and targeted sequencing samples, respectively. The SNV burden was higher in AP/BP samples ( $0.22 \pm 0.13$  for WES and  $0.95 \pm 0.72$  for targeted) than in CP samples ( $0.12 \pm 0.05$  and  $0.24 \pm 0.47$ ; Figure 2A-B). Similarly, CP cases that did not achieve optimal response ( $0.14 \pm 0.07$  and  $0.63 \pm 0.79$ ) had a higher SNV burden when compared with cases achieving MMR ( $0.10 \pm 0.03$  and  $0.19 \pm 0.43$ ; supplemental Figure 4A-B). Notably, the SNV burden did not correlate with other clinical criteria (Figure 2C-F; supplemental Figure 5). By conducting a mutation signature analysis of variants pooled over subgroups, we observed a marked amount of DNA mismatch and double-strand repair signature type variants in CML samples (Figure 2G; supplemental Figure 6). The age-related signature 1 and double-strand repair signature 3 accounted for most of the variants in AP/BP. In general, mutation signatures of AP/BP were more similar to those found in AML ( $r = 0.90$ ) than in CP ( $r = 0.24$ ; Figure 2H; supplemental Data set 3). Within the CP subsets, patients responding poorly had proportionally more variants assigned to signatures 1, 7, and 9 than did the optimal or suboptimal responders, who showed dominance of mismatch repair signatures 6 and 15.

## Viral pathogen screening

Reactivation of viruses is a known complication of TKI treatment, and viral infections have been associated with TKI use.<sup>23-25</sup> Therefore, we performed mapping of sequence reads to viral genomes. This analysis revealed an average viral load of 0.36 CPM across our samples. Elevated amounts of viral reads (CPM  $\geq 1.5$ ) were seen in 2 AP/BP samples from 1 individual (Figure 1D; supplemental Figures 7 and 8). However, no significant difference ( $P > .50$ ) was seen in viral loads between follow-up and diagnostic CP samples (supplemental Data set 4). Statistically significant differences were also not found between tumor control and follow-up samples ( $P > .20$ ) or between tumor control and diagnostic samples ( $P > .40$ ). Thus, in the present study, we did not find genomic evidence of enrichment of viral genome material in the samples of CML patients treated with TKI therapy.

## Relevant variants in diagnostic samples

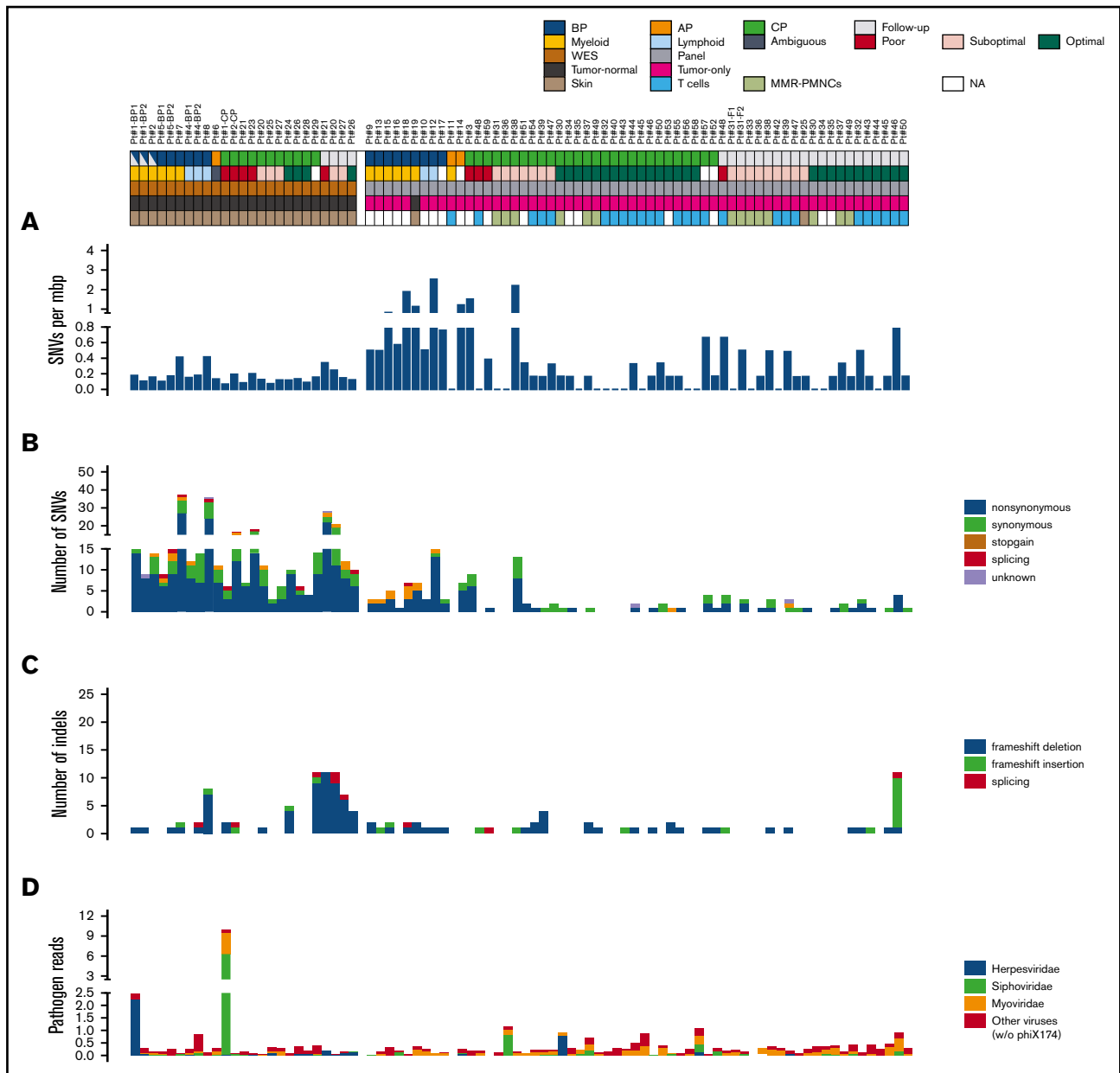
We identified 57 potentially relevant variants across the 59 diagnostic samples through the selection of nonsilent variants (ie, nonsynonymous SNVs and frameshifting insertions and deletions) that were linked with cancer or occurred in genes mutated in 2 or more patients (Figure 3; all 279 nonsilent variants are listed in supplemental Data set 5). The majority occurred in AP/BP samples (frequency, 37). At least 1 variant was identified in 85% of AP/BP samples (Figure 4), and they had a mean of 1.9 potentially relevant nonsilent variants. Among the most recurrently mutated genes in AP/BP samples ( $n = 19$ ) were *ABL1* ( $n = 7$ ), *ASXL1* ( $n = 5$ ), *RUNX1* ( $n = 3$ ), and *BCOR* ( $n = 3$ ). Two patients had mutations in the *JAK3* (p.R870Q and p.A966T) gene and 2 in the *BRD3* (p.P24fs) gene. We also detected variants in DNA repair components, such as *MSH6* (p.E226fs) and *PAPD7* (p.K528fs). Regarding CP, a lower number of mutations was found (frequency, 20). CP cases carried a mean of 0.47 variants, and only 35% of cases had potentially relevant nonsilent mutations. In optimal responders (20%) mutations were rare, whereas in suboptimal (31%) and poor (71%) responders they were more common (Figure 4). The few genes recurrently mutated in CP cases ( $n = 40$ ) were *ASXL1* ( $n = 6$ ), *PRKDC* ( $n = 2$ ), and *KIAA1549* ( $n = 2$ ). We also detected more mutations in epigenetic modifier genes in poor ( $n = 3$  of 7) and suboptimal ( $n = 3$  of 13) responders than in optimal CP responders ( $n = 3$  of 20; supplemental Figure 9), albeit without statistical significance ( $P > .20$ ). Interestingly, these mutations were also enriched in AP/BP samples ( $n = 10$  of 19;  $P \leq .05$ ). In addition, mutations were seen in the protein tyrosine phosphatase receptor (PTPR) genes *PTPRD* and *PTPRJ*. Additional mutations were also discovered in isolated cases in *PTPRF* and *PTPRS* (supplemental Data set 5).

## Relevant variants in longitudinal samples

We analyzed longitudinal samples taken at a median of 6 months off the diagnosis from 5 patients with poor (median interval of 6 months), 11 with suboptimal (median interval of 6 months), and 12 with optimal (median interval of 4.5 months) responses. This analysis gained insights into the clonal dynamics and identified the loss of 5, acquisition of 15, and presence of 10 truncal nonsilent variants of potential relevance to CML between follow-up and diagnostic samples (Figure 5; supplemental Figure 10). Corresponding numbers for all nonsilent mutations were 57, 130, and 20, respectively (supplemental Data set 6). The presence of truncal variants (ie, shared by all clones) or acquisition of new potentially relevant variants was nominal among poor responders, as exemplified by truncal mutations in *RUNX1* (p.R162K) and *MECOM* (p.A200G) and acquisition of mutations in *KDM6A* (p.R1163P) and *RERE* (p.S1434F). In contrast, truncal or new potentially relevant mutations were identified only in 4 suboptimal and 3 optimal responders (Figure 5; supplemental Figure 10). Further evaluation of variant allele frequency data revealed the presence of preleukemic mutations (variants with a stable frequency at diagnosis and follow-up and coupled with *BCR1-ABL1* decline) and clonal evolution in Philadelphia-negative clone (variants acquired despite *BCR-ABL1* clearance). Such variants resided mainly in epigenetic modifier genes (*TET2* and *DNMT3A*) and were associated with mixed responses.

## Transcriptional analysis of CML samples

Seven AP/BP, 5 CP, and 4 healthy control subjects had RNA sequencing performed and were assessed for hybrid transcripts and gene deregulation events (Figure 6). These analyses defined

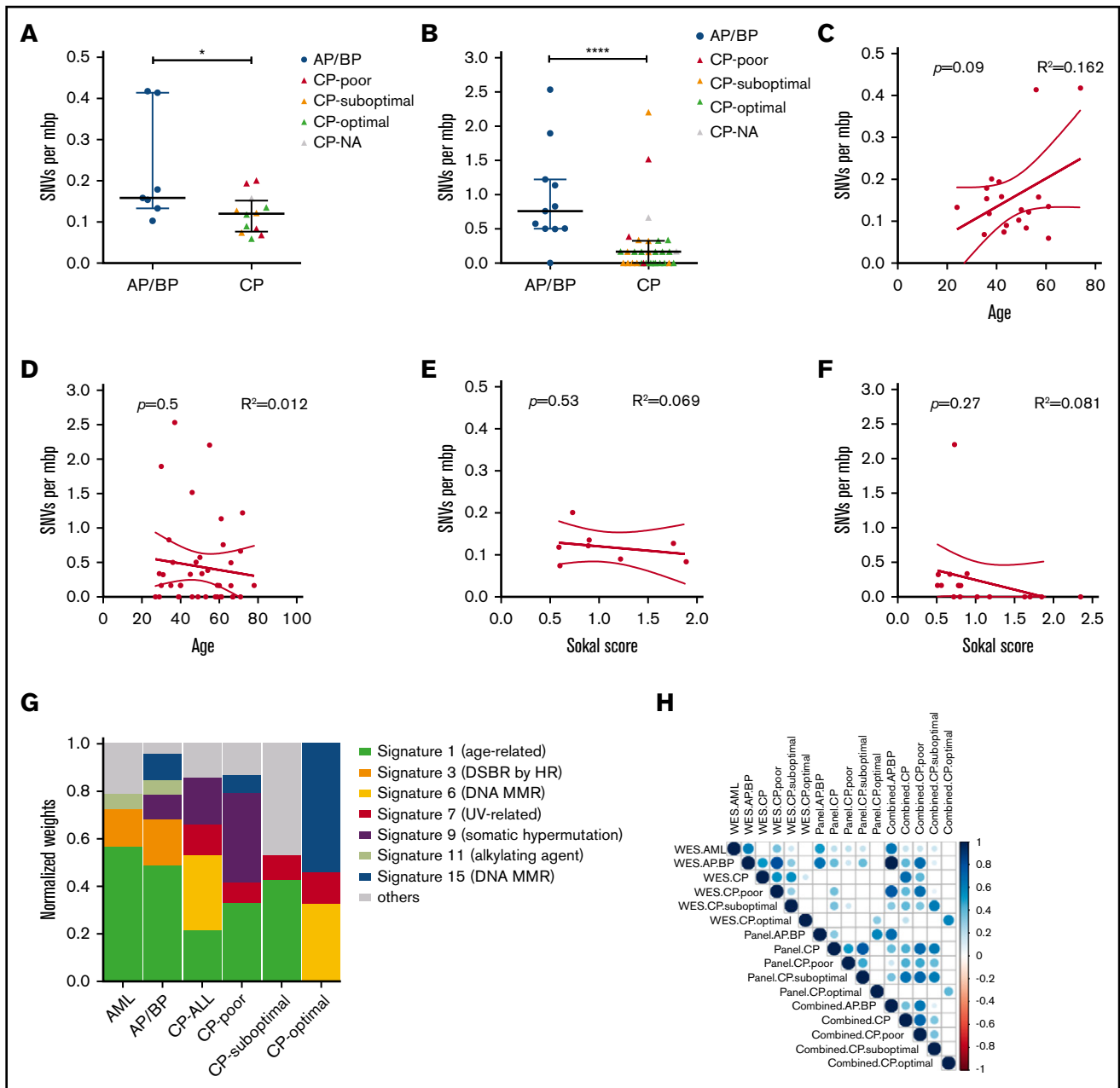


**Figure 1. Mutational landscape.** Explanatory tracks below sample names indicate the sampling point (diagnostic AP, BP, CP, or follow-up sample), treatment response for CP cases (poor, suboptimal, or optimal), expansion compartment (myeloid, lymphoid, or ambiguous) for AP/BP cases, sequencing strategy (WES or panel sequencing), variant calling strategy (tumor normal or tumor only), and control sample type (skin, T cells, or PMNC). The following tracks show mutation load calculated as number of SNVs per mbp (A), the number of SNVs identified in each sample by the consequence of the mutation (B), the number of small frameshift insertions and deletions identified in each sample by the consequence of the frameshift insertion and deletion (C), and the number of viral reads in each sample expressed as CPMs (D). Diagnostic samples from patients #22, #33, #41, #42 and follow-up sample from patient #42 were excluded from the analysis.

P210 *BCR-ABL1* hybrid in 7 AP/BP and in 3 CP samples. In addition, we found 8 other hybrids in AP/BP samples, but none in CP samples. These included 3 known leukemia-associated hybrid genes, *CBFB-MYH11*, *RUNX1-DYRK1A*, and *TMEM236-MRC1* (detected in 2 cases), and 5 putative hybrid genes. All known hybrid genes were validated in relevant samples by at least 1 method (Figure 6A; supplemental Data set 7).

An expression analysis between the AP/BP or CP group and the control group revealed 1237 differentially expressed genes

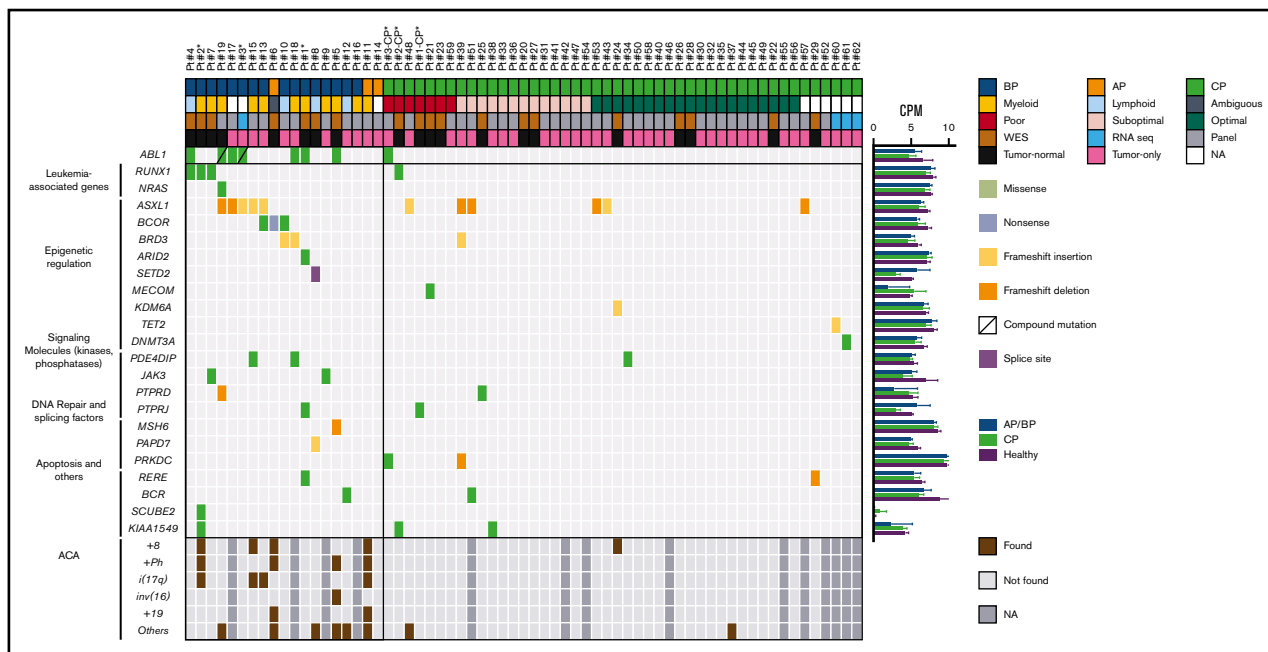
(Figure 6B; supplemental Data set 8). Genes called as differentially expressed ( $Q \leq .05$ ) in both disease phases included known CML markers, such as *DPP4* (*CD26*) and *RXFP1*, as well as other cancer-related genes, such as *CD69*, *ST18*, *IL2RA*, *MLLT4*, and *MUC4*. Genes implicated in important cancer signaling (JAK/STAT and TNF-signaling pathways), immune checkpoints, and inflammation pathways were also found to be altered. Moreover, analyses revealed expression changes that were disease-phase specific (Figure 6B-C; supplemental Figure 11). Deregulation of *MRC1*,



**Figure 2. Mutation loads and cancer signatures.** Numbers of SNVs per mbp in individual samples in WES (A) and targeted sequencing (B). Samples were categorized into AP/BP (WES,  $n = 7$ ; panel,  $n = 12$ ) and CP (WES,  $n = 12$ ; panel,  $n = 31$ ) cases. CP cases with optimal (WES,  $n = 4$ ; panel,  $n = 16$ ), suboptimal (WES,  $n = 3$ ; panel,  $n = 10$ ), and poor (WES,  $n = 4$ ; panel,  $n = 3$ ) responses are marked with green, orange, and red, respectively. CP samples with missing response data (WES,  $n = 1$ ; panel,  $n = 2$ ) are marked with gray. Scatterplots compare mutation load per mbp and age in WES (C) and targeted sequencing (D) data sets and mutation load per mbp and Sokal score in WES (E) and targeted sequencing (F). Correlations between variables were assessed using Pearson's correlation coefficient. (G) Normalized weights of trinucleotide signatures identified in AML ( $n = 11$ ), AP/BP ( $n = 19$ ), and CP ( $n = 39$ ) cases. CP samples were also subset to optimal ( $n = 19$ ), suboptimal ( $n = 10$ ), and poor ( $n = 7$ ) responders. Variants were pooled across WES and panel sequencing cases. Weights of most frequent signatures in each cancer type are shown across cancers as separate signatures. Summed weight of other signatures is represented by the category "other." (H) Correlation matrix showing Spearman correlation between normalized weights of signatures of AML, AP/BP, CP, and CP subsets calculated using WES, panel, or combined WES and panel data. Blue dot color indicates positive correlations and red negative correlations as expressed in  $r$  values. The dot size represents  $-\log_{10} P$  values. Nonsignificant correlations ( $P > .05$ ) were removed. \* $P < .05$ ; \*\*\*\* $P < .0001$ . Diagnostic samples from patients #22, #33, #41, #42 and follow-up sample from patient #42 were excluded from the analysis.

*IL21R*, and *HGF* genes was identified only in AP/BP vs control samples and deregulation of *GLI2*, *GAS2*, *IRF8*, and *PIEZO2* in CP vs control samples. Activation of the KRAS pathway was specific for

AP/BP and activation of the CXCR4 pathway for CP. Pathway enrichment analysis also revealed expression alterations between CP and AB/BP in the IL-2/STAT5 and IL-6/JAK/STAT3 pathways



**Figure 3. Landscape of relevant nonsilent variants at diagnosis.** Potentially relevant nonsilent (ie, nonsynonymous SNVs and frameshifting insertions and deletions that were linked with cancer or occurred in genes that mutated in 2 or more patients) variants. Explanatory tracks from top to bottom show the sampling point (AP, BP, CP or follow-up sample), treatment response for CP cases (poor, suboptimal, or optimal), and expansion compartment (myeloid, lymphoid, or ambiguous) for AP/BP cases, sequencing strategy (WES, panel sequencing, or RNA sequencing), and variant calling strategy (tumor normal or tumor only). The color of the variant box indicates the type of mutation. The average expression of the genes in 7 AP/BP, 5 CP, and 4 healthy subjects are shown on the right, expressed as CPMs. Bar lengths indicate means and error bars represent range. At the bottom, additional chromosomal abnormalities (ACA) found using karyotyping techniques are shown. \*Patients with samples from both CP and AP/BP.

(Figure 6C; supplemental Data set 9). Pairwise comparison between CML phases, however, uncovered only a handful of genes between CP and AB/BP ( $n = 21$ ). The few genes that reached statistical significance in all comparisons included *IL2RA* (*CD25*) and chimerin-1 (*CHN1*; supplemental Data set 8).

### A precision medicine approach to CML management

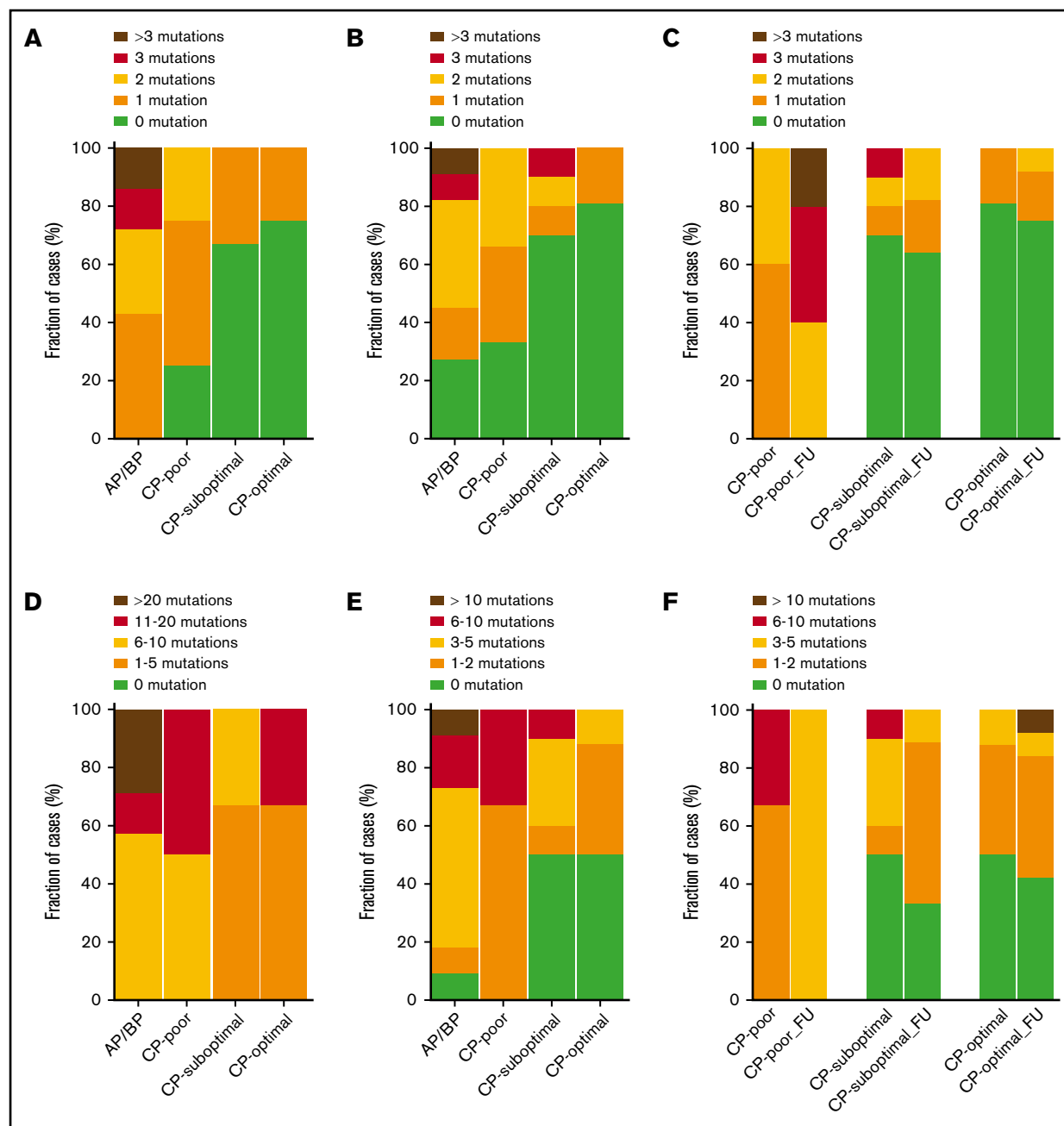
An integrative personalized approach was applied to a 35-year-old man with lymphoid CML failing to respond to TKI treatment (Patient #4; Figure 7A). The patient achieved partial remission (blast count, 5%), but eventually relapsed (blast count, 85%). WES and RNA-sequencing analysis of initial samples revealed somatic mutations in *ABL1* (p.T315I) and *RUNX1* (p.R177Q; Figure 7B) and deregulation of several *RUNX1* targeting pathways, such as coagulation, complement, and hematopoietic stem cell pathways (supplemental Figure 12A; supplemental Data sets 10 and 11). Using drug-sensitivity profiling, ponatinib was found to be the effective *BCR-ABL1* TKI. Other effective drugs included glucocorticoids and mTOR, VEGFR, and MEK inhibitors (supplemental Figure 12B-C; supplemental Data set 12). Building on these results, the VEGFR-selective TKI inhibitor axitinib was administered<sup>26</sup> at first, and it induced clearance of blast cells. The treatment was changed to ponatinib followed by another course of axitinib. At relapse, drug-sensitivity profiling was repeated. WES revealed disappearance of the original clone and emergence of a clone with a mutation in *EZH2* (p.A687V). We also observed upregulation of genes implicated in IL-8- and CXCR1-mediated signaling and in fatty acid metabolism. In agreement, drug-sensitivity profiling showed increased sensitivity to PF-3845, a fatty acid amide hydrolase inhibitor (supplemental Figure 12B-C). Because of the

lack of effective drugs for the found clone, the patient was treated with allogeneic hematopoietic stem cell transplantation (allo-HSCT).

The same approach was also applied in a 49-year old male patient who presented with myeloid BP CML (Patient #5; Figure 7C). WES of the initial sample revealed variants in *IL1R1* (p.G420V) and *DOT1L* (p.R347H). The patient achieved a partial response to dasatinib treatment, but eventually relapsed after 4 months. At relapse, WES revealed acquisition of somatic mutations in *ABL1* (p.T315I), *MSH6* (p.E226fs), and *SETD1B* (p.K1717X) genes (Figure 7D). RNA sequencing revealed acquisition of a hybrid gene *CBFB-MYH11* confirmed by FISH and *inv(16)* in karyotyping. In addition, eosinophilic markers, including *EPX* and *PRG2* genes and genes from *CBFB-MYH11*-related pathways, were upregulated. Drug-sensitivity testing revealed that ponatinib was the only active *BCR-ABL1* TKI at relapse, in addition to the VEGFR and mTOR inhibitors (supplemental Figure 13). Axitinib was administered for 3 months, together with short CVAD cycles and bosutinib to inhibit the wild-type *BCR-ABL1*. Treatment induced rapid clearance of blasts, and the patient proceeded to allo-HSCT.

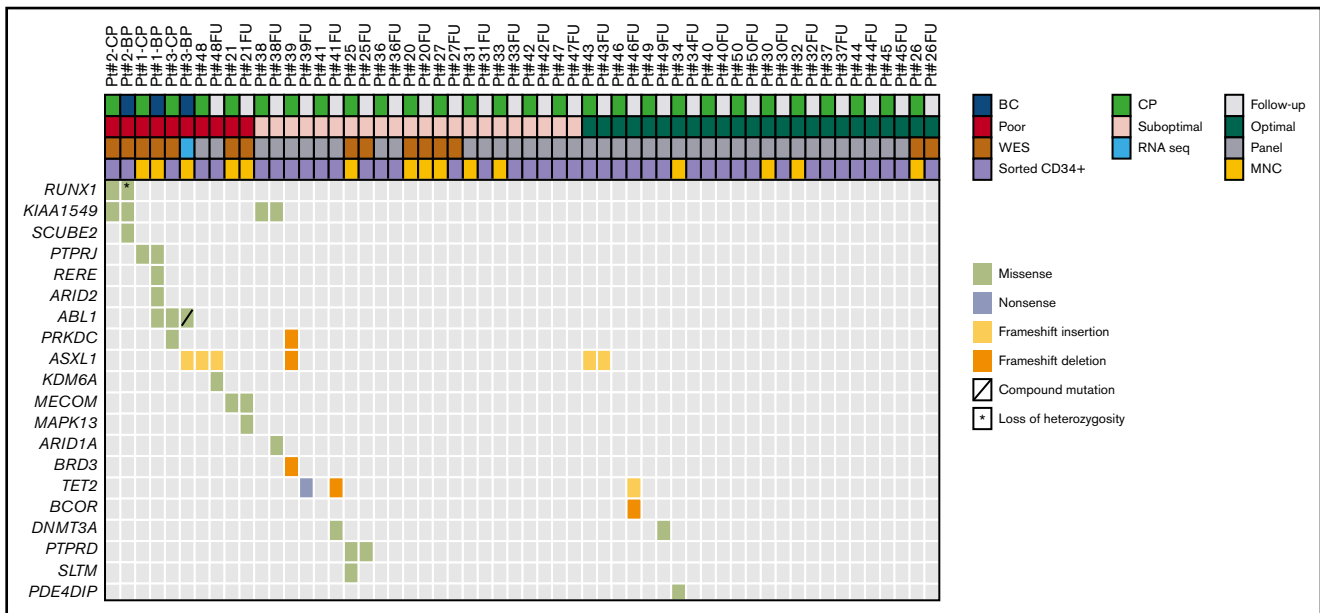
### Discussion

In the present study, we characterized genetic factors underlying CML. We compared samples from indolent CP and aggressive AP/BP stages and from patients with various responses to TKI. Our results showed that AP/BP samples shared genetic characteristics similar to those of AML samples and that cancer-relevant nonsilent variants at diagnosis were associated with treatment outcome. We also found that AP/BP and CP patients with poor response had



**Figure 4. Mutation fractions of nonsilent mutations.** (A) Stacked columns show the fraction of cases with a given number of potentially relevant nonsilent variants (ie, nonsynonymous SNVs and frameshifting insertions and deletions that were linked with cancer or occurred in genes mutated in 2 or more patients) among AP/BP cases and CP cases with poor, suboptimal, or optimal responses sequenced by WES ( $n = 18$ ). Differences in distribution of variants within all diagnostic groups ( $P = .51$ ) and diagnostic CP subsets ( $P = .81$ ) were assessed with Fisher's exact test. (B) Fraction of cases with a given number of potentially relevant nonsilent variants among cases sequenced by panel sequencing ( $n = 42$ ).  $P$  values by Fisher's exact test: all diagnostic groups ( $P = .04$ ) and diagnostic CP subsets ( $P \leq .15$ ). (C) Fraction of cases with potentially relevant nonsilent variants at diagnosis and follow-up (FU) sequenced using either WES or panel sequencing. Only cases with more than 1 sample were included ( $n = 40$ ).  $P$  values by Fisher's exact test: diagnostic CP subset ( $P = .003$ ) and follow-up CP subsets ( $P = .02$ ). (D) Fraction of cases with a given number of nonsilent variants among cases sequenced by WES ( $n = 17$ ).  $P$  values by Fisher's exact test: all diagnostic groups ( $P = .09$ ) and diagnostic CP subsets ( $P = .26$ ). (E) Fraction of cases with a given number of nonsilent variants among cases sequenced by panel sequencing ( $n = 40$ ).  $P$  values by Fisher's exact test: all diagnostic groups ( $P = .01$ ) and diagnostic CP subsets ( $P = .08$ ). (F) Fraction of cases with nonsilent variants at diagnosis and follow-up, sequenced using either WES or panel sequencing. Only cases with more than 1 sample were included ( $n = 28$ ).  $P$  values by Fisher's exact test: diagnostic CP subset ( $P = .08$ ) and follow-up CP subsets ( $P = .44$ ).





**Figure 5. Landscape of relevant nonsilent variants in serial samples.** Potentially relevant nonsilent variants (ie, nonsynonymous SNVs and frameshifting insertions and deletions that were linked with cancer or occurred in genes mutated in 2 or more patients). Explanatory tracks from top to bottom show the sampling point (BP, CP, or follow-up [FU] sample), treatment response for CP cases (poor, suboptimal, or optimal), sequencing strategy (WES, RNA sequencing, or panel sequencing), and sorting status of the sample (sorted CD34<sup>+</sup> cells or unsorted MNCs). The color of the variant box indicates the type of mutation.

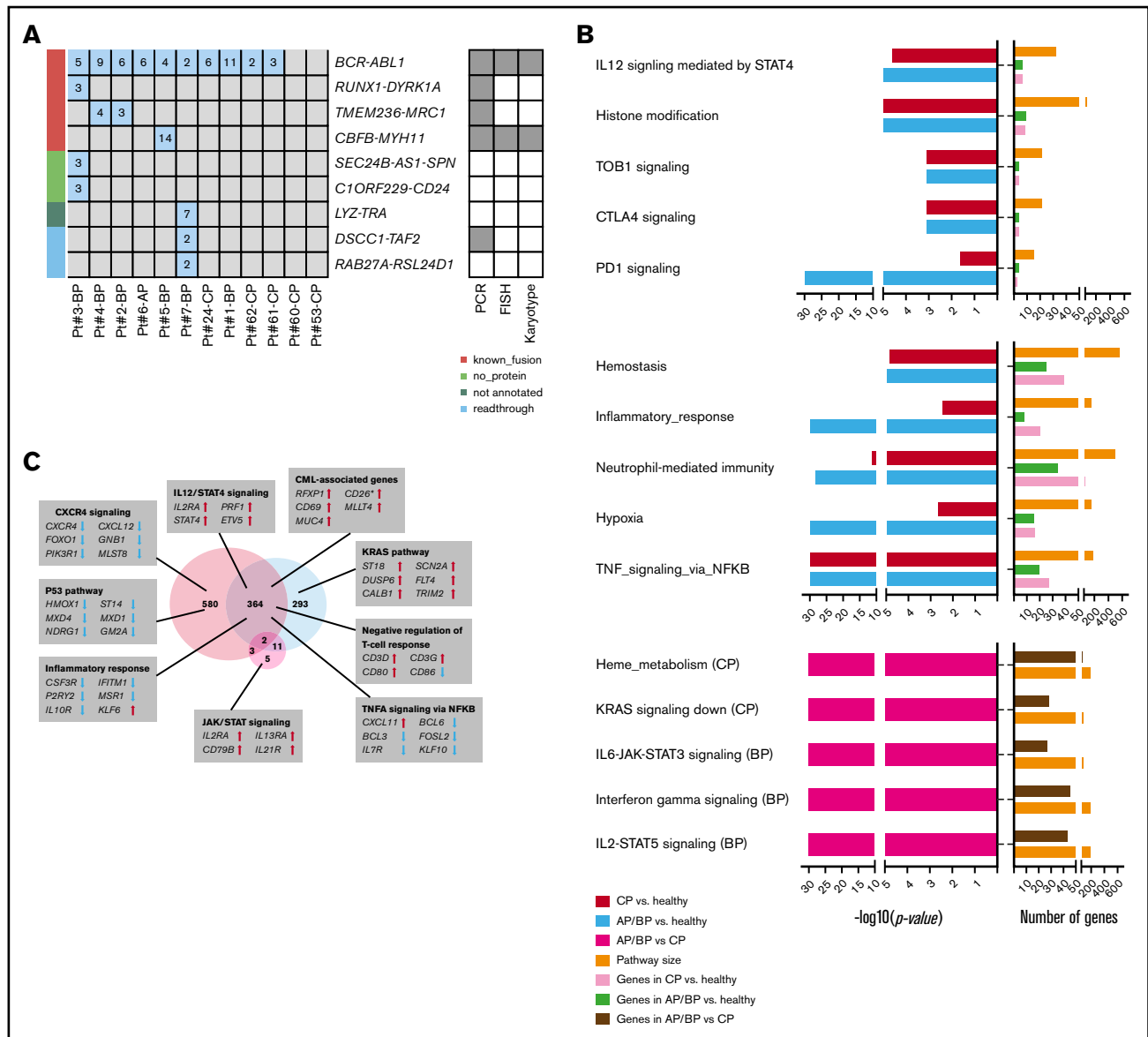
more truncal and newly acquired variants of relevance to CML than the other patients. Using RNA sequencing, we identified hybrid genes and detected expression perturbations in cancer markers and pathways.

Somatic mutations at diagnosis are associated with treatment outcome in CML.<sup>17-19,27</sup> In our study, we found 65% of CP and 95% of AB/BP cases to harbor at least 1 somatic mutation. The fraction of CP patients carrying a mutation was higher than that reported previously in studies using a panel of 92 genes<sup>19</sup> (37%) or WES<sup>13</sup> (50%). Regarding nonsilent variants, 35% of CP cases and 85% of AP/BP cases had at least 1 mutation of relevance to CML. Interestingly, the amount of these variants was associated with the outcome, and they were enriched in patients with a poor response in comparison with cases with optimal or suboptimal outcome. SNV burden information also supported the link between the number of variants and treatment outcome. Compared with solid tumors and lymphomas,<sup>28,29</sup> the SNV burden in our cohort rate was low. However, AP/BP samples had significantly higher SNV burden than did CP samples. Within the CP subset, patients with a poor response showed a tendency toward higher SNV burden than the others. Notably, SNV burden was independent of other prognostic parameters.

Mutational signature profiles were constructed by pooling SNVs across sample sets. These analyses revealed the presence of variants from cancer aging-related signature (signature 1) as well as DNA double-strand break repair (signature 3), and mismatch repair (signatures 6 and 15) signatures in diagnostic AP/BP and CP samples in an agreement with the known role of *BCR-ABL1* to induce DNA damage and inhibiting DNA repair.<sup>30,31</sup> In general, mutation signatures in AP/BP and AML were similar. This may be related to variants in DNA repair and spliceosomes genes, commonly mutated in AML<sup>32-34</sup> and in our AP/BP and CP patients with poor or suboptimal response. Within the CP subset, signatures

7 and 9 were enriched in poor and suboptimal responders. Of these, signature 9 has been related to acute lymphoblastic leukemia (ALL)<sup>35</sup> and AP/BP patients with *RUNX1* mutations.<sup>36</sup> It is also related to mutations caused by polymerase  $\eta$ . The UV-related signature 7 has been described in pediatric B-ALL.<sup>37</sup> CP patients with optimal response showed dominance of mismatch-repair signatures.

Our analyses identified several genes recurrently mutated in CML. In addition to *ABL1* mutations, recurrently mutated genes in AP/BP included *RUNX1*, *ASXL1*, and *BCOR* in corroboration with previous results.<sup>11-13,38</sup> A novel gene recurrently mutated in AP/BP was *BRD3*, which belongs to bromodomain and extraterminal family with a role in AML.<sup>39</sup> Interestingly, the particular variant (p.P24fs) is supported by several observations in a cancer knowledge base (Catalog of Somatic Mutations in Cancer #COSM21288).<sup>40</sup> Other cancer drivers with mutations, although in solitary cases, in AP/BP cases were *RERE*<sup>41</sup> and *SETD2*.<sup>42</sup> As in previous reports,<sup>13,19</sup> mutations in leukemia driver genes were scarce in CP. Those patients that carried such mutations typically responded unfavorably. For example, a patient with *RUNX1* p.R177K at diagnosis progressed into BP at 7 months.<sup>43</sup> We also detected variants in epigenetic modifiers. Notably, *ASXL1* variants were found in 5 AP/BP and 6 CP cases. The high-mutation frequency of *ASXL1* has been described previously.<sup>17,18,44-46</sup> Mutations in other epigenetic modifiers were also found and associated with unfavorable outcome, as recently reported.<sup>19,47,48</sup> AP/BP and CP cases with unfavorable response also harbored mutations in *PTPR* genes, which are tumor suppressors<sup>49</sup> and negative regulators of JAK/STAT signaling.<sup>50</sup> They have been reported to be downregulated in CML patients.<sup>51</sup> We speculate that loss of *PTPR* could result in activation of JAK/STAT signaling and be associated with TKI resistance.

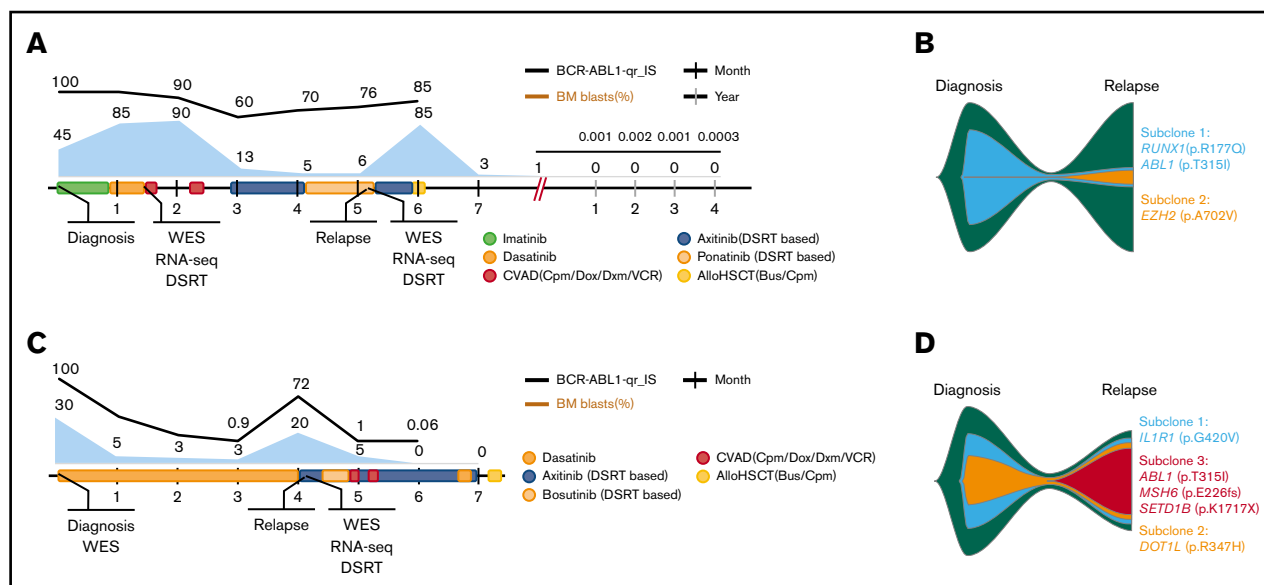


**Figure 6. Expression landscape.** (A) Fusion genes found in 7 AP, BP, or CP samples. Fusions were classified into 4 categories: known fusions, fusions including genes without a known protein product, unannotated fusions, and readthroughs. Blue (left) indicates detection of the fusion, and numbers indicate the number of unique reads supporting the fusion. The methods employed to validate the fusions are shown (right). (B) Depiction of the main molecular pathways with altered expression between healthy and CP cases, healthy and AP/BP cases, and CP and AP/BP cases. Upregulated pathways (top) and downregulated pathways (middle) in comparison with healthy cases. Deregulated pathways (bottom) in AP/BP vs CP cases. Pathways with  $P = .00$  were set to the minimum  $P$  of all pathways divided by 2 in the visualization. (C) Venn diagram showing the number of differentially expressed genes between AP/BP and healthy (light blue), CP and healthy (light red), and AP/BP and CP. Differentially expressed genes relevant for CML pathogenesis are highlighted in the figure. Upregulated genes are highlighted with arrows pointing upward and downregulated genes with arrows pointing downward. \*Genes significantly differentially expressed in AP/BP in comparison with healthy, but borderline significance ( $Q > .05$  and  $Q < .10$ ) in CP, in comparison with healthy cases.

We also explored mutational dynamics of CP under TKI treatment. In general, our findings supported association between clone dynamics and TKI response.<sup>19</sup> They showed that CP patients with unfavorable outcome had more truncal variants and/or acquired more variants during the treatment than optimal responders. Our analysis also indicated that variants were seldom lost during follow-up.

A hybrid gene analysis highlighted that hybrids other than *BCR-ABL1* were frequent in AP/BP, but not in CP. This was in support with a finding suggesting a role for hybrid genes in CML progression.<sup>13</sup>

Among the hybrid genes that we identified were known leukemia-associated hybrids such as *CBFB-MYH11*, which is associated with the AML subtype M4Eo, and indicates aggressive outcome in AP/BP,<sup>52</sup> and *RUNX1-DYRK1A*, which has been linked with ALL pathogenesis.<sup>53</sup> Other potentially oncogenic hybrids included *DSSC1-TAF2* identified in breast cancer and *RSL24D1-RAB27A* detected in lung adenocarcinoma.<sup>54</sup> The *TMEM236-MRC1* hybrid was found in 2 cases, being the only recurrent hybrid other than *BCR-ABL1*. Interestingly, both of its partners (*TMEM236* and *MRC1*;  $Q \leq .05$ ) were among the top altered genes in the expression analysis.



**Figure 7. Clinical course of 2 index patients subjected to precision-medicine-based treatment management.** (A) Clinical course of the first patient highlighting the collection of samples for molecular diagnostics, blast levels, and administration of cancer drugs. The timeline provides a monthly view of the clinical events during the disease course and a yearly view of follow-up observations after allo-HSCT. (B) Depiction of the clonal evolution of the tumor of the first patient during transformation from diagnosis (blast level, 45%), through partial remission (blast level, 5%), to relapse (blast level, 85%). Analyses highlighted the major loss of the subclone, with mutations in *ABL1* and *RUNX1* at the second sampling. (C) Clinical course of the second patient, highlighting collection of samples for molecular diagnostics, blast levels, and administration of cancer drugs. The timeline provides a monthly view of clinical events during the disease course. (D) Depiction of the clonal evolution of the tumor of the second patient during transformation from diagnosis (blast level, 30%), through partial remission (blast level, 5%), to relapse (blast level, 20%). Analyses highlighted the acquisition of a subclone with mutations in *ABL1*, *MSH6*, and *SETD1B* genes at the second sampling.

Transcriptional changes taking place at CML were identified by means of differential gene expression analysis, accounting for confounding factors. These analyses revealed deregulation of genes reported to be altered in previous CML investigations, such as *RXFP1*,<sup>43</sup> *PIEZO2*,<sup>55</sup> and *CD26*,<sup>56</sup> and in leukemia studies, such as *CD69*,<sup>57</sup> *ST18*,<sup>58</sup> and *MUC4*.<sup>59</sup> Expression analysis also suggested that DNA damage could be related to the down-regulation of DNA-repair machinery genes and that dysregulation events in genes, including *DNMT1*,<sup>60</sup> *SEPP1*,<sup>61</sup> *NEIL1*,<sup>62</sup> and *WT1*,<sup>63,64</sup> may have contributed to the high occurrence of DNA damage-associated variants in our cohort. In addition, several immune checkpoint genes were found to be deregulated, suggesting an explanation for how CML cells evade immune-cell clearance. Our data also highlighted progression-associated events, such as expression alterations in *HGF*<sup>43</sup> and *CD26*<sup>56</sup> between AP/BP and CP. The pathway-enrichment analysis between AP/BP and CP in turn highlighted the IL-2-induced JAK-STAT signaling pathway along with the KRAS- and IL-6-induced JAK-STAT signaling pathways, providing support for the role of IL-2-induced JAK-STAT signaling in CML progression<sup>56</sup> and indicating that the JAK-STAT signaling pathways could be targets of CML treatment. However, it must be emphasized that the small sample size and variations in the cell fractions studied (supplemental Figure 12) were limitations and that a larger study is needed to clarify the findings.

An integrative approach of WES, RNA sequencing, and ex vivo drug-sensitivity profiling was applied to 2 independent BP cases at diagnosis and relapse. In the first case, this approach identified *RUNX1* (p.R177Q) and *ABL1* (p.T315I) mutations and pinpointed ponatinib and drugs previously related to *RUNX1* mutations and

*RUNX1-RUNX1T1* hybrids in AML, including glucocorticoid,<sup>65,66</sup> mTOR,<sup>65</sup> and VEGFR<sup>67</sup> inhibitors. Evidence of translational capability was provided by the success of axitinib in inducing partial remission. At relapse, replacement of the original clone with an *EZH2* mutation clone resulted in downregulation of *EZH2/PRC2* complex target genes<sup>68,69</sup> and loss of activity of the administered drugs. The patient was successfully treated at relapse with allo-HSCT. In the second case, the integrative approach was applied to a relapsed patient with *ABL1* (p.T315I), *MSH6* (p.E226fs), and *SETD1B* (p.K1717X) mutations. The patient underwent drug-sensitivity and -resistance testing-based axitinib treatment and successfully translated into remission.

In conclusion, our study provides insights into CML pathogenesis. We showed that AP/BC and CP samples differed in their mutation load and variant signatures and that the frequency of variants was linked with unfavorable treatment outcomes. Transcriptome profiling identified new hybrid genes and revealed that a notable portion of transcriptome changes are shared between the AP/BP and CP phases. From a clinical prospective, our results suggest that clone architecture and mutations in cancer genes at diagnosis are genetic biomarkers that should be considered when tailoring precision medicine strategies for CML management and stratifying patients.

## Acknowledgments

The authors thank the personnel of the FIMM Technology Center and Hematology Research Unit, Helsinki, for assistance. CSC (IT Center for Science Ltd) is acknowledged for assistance and computing resources.

This work was supported by the Academy of Finland (grants 292605 and 287224); the Finnish Funding Agency for Innovation (Dnro 6113/31/2016), the Finnish Special Governmental Subsidy for Health Sciences, Research, and Training; the Signe and Ane Gyllenberg Foundation, the Finnish Cultural Foundation, the Helsinki Institute of Life Science, the Cancer Society of Finland, the Relander Foundation, the Nordic Cancer Union, an Incyte Nordic Hematology grant, and the Finnish Cancer Institute.

## Authorship

Contribution: S.A.A. designed the study, performed experiments, and analyzed DNA, RNA-sequencing, and drug-sensitivity data; M.K. designed the study and performed DNA and RNA-sequencing data analyses; T.O. performed pathogen screening analyses; B.G. and S.E. helped in performing the data analyses; P.K., A.K., S.K., M.M.K., C.A.H., and K.P. contributed and prepared biological samples and clinical data; S.M. conceived of and designed the study and directed

and supervised the research; S.A.A., M.K., T.O., and S.M. wrote the manuscript; and all authors read and approved the final manuscript.

Conflict-of-interest disclosure: S.M. has received honoraria and research funding from Novartis, Pfizer, and Bristol-Myers Squibb (all 3 unrelated to this study). S.A.A. has received research funding from Incyte. The remaining authors declare no competing financial interests.

ORCID profiles: S.A.A., 0000-0001-6205-7799; M.K., 0000-0002-4714-9481; T.O., 0000-0001-8186-1707; P.K., 0000-0002-7121-321X; S.E., 0000-0001-5056-4750; B.G., 0000-0001-5076-8366; K.P., 0000-0003-4112-5902; S.M., 0000-0002-0816-8241.

Correspondence: Satu Mustjoki, Hematology Research Unit Helsinki, University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center, Haartmaninkatu 8, PO Box 700, FIN-00290 Helsinki, Finland; e-mail: satu.mustjoki@helsinki.fi.

## References

1. Deininger MWN, Goldman JM, Melo JV. The molecular biology of chronic myeloid leukemia. *Blood*. 2000;96(10):3343-3356.
2. Hoffmann VS, Baccarani M, Hasford J, et al. The EUTOS population-based registry: incidence and clinical characteristics of 2904 CML patients in 20 European Countries. *Leukemia*. 2015;29(6):1336-1343.
3. Apperley JF. Chronic myeloid leukaemia. *Lancet*. 2015;385(9976):1447-1459.
4. Bower H, Björkholm M, Dickman PW, Höglund M, Lambert PC, Andersson TM. Life Expectancy of Patients With Chronic Myeloid Leukemia Approaches the Life Expectancy of the General Population. *J Clin Oncol*. 2016;34(24):2851-2857.
5. Hochhaus A, Saussele S, Rosti G, et al. Chronic myeloid leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2017;28(suppl 4):iv41-iv51.
6. Hehlmann R. How I treat CML blast crisis. *Blood*. 2012;120(4):737-747.
7. Jain P, Kantarjian H, Ghorab A, et al. Prognostic factors and survival outcomes in patients with chronic myeloid leukemia in blast phase in the tyrosine kinase inhibitor era: Cohort study of 477 patients. *Cancer*. 2017;123(22):4391-4402.
8. Burke BA, Carroll M. BCR-ABL: a multi-faceted promoter of DNA mutation in chronic myelogenous leukemia. *Leukemia*. 2010;24(6):1105-1112.
9. Nieborowska-Skorska M, Kopinski PK, Ray R, et al. Rac2-MRC-clll-generated ROS cause genomic instability in chronic myeloid leukemia stem cells and primitive progenitors. *Blood*. 2012;119(18):4253-4263.
10. Bolton-Gillespie E, Schemionek M, Klein H-U, et al. Genomic instability may originate from imatinib-refractory chronic myeloid leukemia stem cells. *Blood*. 2013;121(20):4175-4183.
11. Grossmann V, Kohlmann A, Zenger M, et al. A deep-sequencing study of chronic myeloid leukemia patients in blast crisis (BC-CML) detects mutations in 76.9% of cases. *Leukemia*. 2011;25(3):557-560.
12. Zhao L-J, Wang Y-Y, Li G, et al. Functional features of RUNX1 mutants in acute transformation of chronic myeloid leukemia and their contribution to inducing murine full-blown leukemia. *Blood*. 2012;119(12):2873-2882.
13. Branford S, Wang P, Yeung DT, et al. Integrative genomic analysis reveals cancer-associated mutations at diagnosis of CML in patients with high-risk disease. *Blood*. 2018;132(9):948-961.
14. Krumbholz M, Tauer J, te Kronnie G, Ekici A, Suttorp M, Metzler M. Copy Number Variations and IKZF1 Mutations In Pediatric CML [abstract]. *Blood*. 2013;122(21). Abstract 1473.
15. Menezes J, Salgado RN, Acquadro F, et al. ASXL1, TP53 and IKZF3 mutations are present in the chronic phase and blast crisis of chronic myeloid leukemia. *Blood Cancer J*. 2013;3(11):e157.
16. Makishima H, Jankowska AM, McDevitt MA, et al. CBL, CBLB, TET2, ASXL1, and IDH1/2 mutations and additional chromosomal aberrations constitute molecular events in chronic myelogenous leukemia. *Blood*. 2011;117(21):e198-e206.
17. Schmidt M, Rinke J, Schäfer V, et al. Molecular-defined clonal evolution in patients with chronic myeloid leukemia independent of the BCR-ABL status. *Leukemia*. 2014;28(12):2292-2299.
18. Togasaki E, Takeda J, Yoshida K, et al. Frequent somatic mutations in epigenetic regulators in newly diagnosed chronic myeloid leukemia. *Blood Cancer J*. 2017;7(4):e559.
19. Kim T, Tyndel MS, Kim HJ, et al. Spectrum of somatic mutation dynamics in chronic myeloid leukemia following tyrosine kinase inhibitor therapy. *Blood*. 2017;129(1):38-47.

20. Branford S, Kim DDH, Apperley JF, et al; International CML Foundation Genomics Alliance. Laying the foundation for genomically-based risk assessment in chronic myeloid leukemia. *Leukemia*. 2019;33(8):1835-1850.
21. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.
22. Pemovska T, Kontro M, Yadav B, et al. Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov*. 2013;3(12):1416-1429.
23. Ikeda K, Shiga Y, Takahashi A, et al. Fatal hepatitis B virus reactivation in a chronic myeloid leukemia patient during imatinib mesylate treatment. *Leuk LymPhoma*. 2006;47(1):155-157.
24. Prestes DP, Arbona E, Nevett-Fernandez A, et al. Dasatinib Use and Risk of Cytomegalovirus Reactivation After Allogeneic Hematopoietic-Cell Transplantation. *Clin Infect Dis*. 2017;65(3):510-513.
25. Lai G-M, Yan S-L, Chang C-S, Tsai C-Y. Hepatitis B reactivation in chronic myeloid leukemia patients receiving tyrosine kinase inhibitor. *World J Gastroenterol*. 2013;19(8):1318-1321.
26. Pemovska T, Johnson E, Kontro M, et al. Axitinib effectively inhibits BCR-ABL1(T315I) with a distinct binding conformation. *Nature*. 2015;519(7541):102-105.
27. Magistrini V, Sharma N, Piazza R, et al. Integrated Analysis of Whole-Exome Sequencing and Micrnas Expression in Blast Crisis Transformation of Chronic Myeloid Leukemia [abstract]. *Blood*. 2012;120(21). Abstract 3727.
28. Alexandrov LB, Nik-Zainal S, Wedge DC, et al; ICGC PedBrain. Signatures of mutational processes in human cancer [published correction appears in *Nature* 2013;502:258]. *Nature*. 2013;500(7463):415-421.
29. Korfi K, Mandal A, Furney SJ, Wiseman D, Somerville TC, Marais R. A personalised medicine approach for ponatinib-resistant chronic myeloid leukaemia. *Ann Oncol*. 2015;26(6):1180-1187.
30. Stoklosa T, Poplawski T, Koptyra M, et al. BCR/ABL inhibits mismatch repair to protect from apoptosis and induce point mutations. *Cancer Res*. 2008;68(8):2576-2580.
31. Tobin LA, Robert C, Rapoport AP, et al. Targeting abnormal DNA double-strand break repair in tyrosine kinase inhibitor-resistant chronic myeloid leukemias. *Oncogene*. 2013;32(14):1784-1793.
32. Wimmer K, Etzler J. Constitutional mismatch repair-deficiency syndrome: have we so far seen only the tip of an iceberg? *Hum Genet*. 2008;124(2):105-122.
33. Makishima H, Visconte V, Sakaguchi H, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood*. 2012;119(14):3203-3210.
34. Wang H, Wen J, Chang CC, Zhou X. Discovering transcription and splicing networks in myelodysplastic syndromes. *PLoS One*. 2013;8(11):e79118.
35. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet*. 2014;46(2):116-125.
36. Awad S, Kankainen M, Dufva O, et al. RUNX1 Mutations Identify an Entity of Blast Phase Chronic Myeloid Leukemia (BP-CML) Patients with Distinct Phenotype, Transcriptional Profile and Drug Vulnerabilities [abstract]. *Blood*. 2018;132(suppl 1). Abstract 4257.
37. Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature*. 2018;555(7696):371-376.
38. Damm F, Chesnais V, Nagata Y, et al. BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. *Blood*. 2013;122(18):3169-3177.
39. Carretta M, Brouwers-Vos AZ, Bosman M, et al. BRD3/4 inhibition and FLT3-ligand deprivation target pathways that are essential for the survival of human MLL-AF9+ leukemic cells. *PLoS One*. 2017;12(12):e0189102.
40. Wang K, Yuen ST, Xu J, et al. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat Genet*. 2014;46(6):573-582.
41. Waerner T, Gardellin P, Pfizenmaier K, Weith A, Kraut N. Human RERE is localized to nuclear promyelocytic leukemia oncogenic domains and enhances apoptosis. *Cell Growth Differ*. 2001;12(4):201-210.
42. Mancini M, Santis SD, Monaldi C, et al. SETD2 Loss of Function Is a Recurrent Event in Advanced-Phase Chronic Myeloid Leukemia [abstract]. *Blood*. 2017;130(suppl 1). Abstract 43.
43. Giustacchini A, Thongjuea S, Barkas N, et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat Med*. 2017;23(6):692-702.
44. Boulwood J, Perry J, Zaman R, et al. High-density single nucleotide polymorphism array analysis and ASXL1 gene mutation screening in chronic myeloid leukemia during disease progression. *Leukemia*. 2010;24(6):1139-1145.
45. Ernst T, Busch M, Rinke J, et al. Frequent ASXL1 mutations in children and young adults with chronic myeloid leukemia. *Leukemia*. 2018;32(9):2046-2049.
46. Roche-Lestienne C, Marceau A, Labis E, et al; Fi-LMC group. Mutation analysis of TET2, IDH1, IDH2 and ASXL1 in chronic myeloid leukemia. *Leukemia*. 2011;25(10):1661-1664.
47. Baer CR, Kern W, Haferlach C, Haferlach T. Mutation Screening at Diagnosis Reveals a High Frequency of ASXL1 Mutations in CML Patients Who Fail to Achieve Molecular Remission Criteria after One Year of TKI Treatment [abstract]. *Blood*. 2017;130(suppl 1). Abstract 1585.

48. Nteliopoulos G, Bazeos A, Claudiani S, et al. Somatic variants in epigenetic modifiers can predict failure of response to imatinib but not to second-generation tyrosine kinase inhibitors. *Haematologica*. 2019;104(12):2400-2409.
49. Du Y, Grandis JR. Receptor-type protein tyrosine phosphatases in cancer. *Chin J Cancer*. 2015;34(2):61-69.
50. Lui WY, Peyser ND, Ng PK-S, et al. Frequent mutation of receptor protein tyrosine phosphatases provides a mechanism for STAT3 hyperactivation in head and neck cancer. *Proc Natl Acad Sci USA*. 2014;111(3):1114-1119.
51. Della Peruta M, Martinelli G, Moratti E, et al. Protein tyrosine phosphatase receptor type gamma is a functional tumor suppressor gene specifically downregulated in chronic myeloid leukemia. *Cancer Res*. 2010;70(21):8896-8906.
52. Salem A, Loghavi S, Tang G, et al. Myeloid neoplasms with concurrent BCR-ABL1 and CBFβ rearrangements: A series of 10 cases of a clinically aggressive neoplasm. *Am J Hematol*. 2017;92(6):520-528.
53. Gu Z, Churchman M, Roberts K, et al. Genomic analyses identify recurrent *MEF2D* fusions in acute lymphoblastic leukaemia. *Nat Commun*. 2016;7:13331.
54. Yoshihara K, Wang Q, Torres-Garcia W, et al. The landscape and therapeutic relevance of cancer-associated transcript fusions. *Oncogene*. 2015;34(37):4845-4854.
55. Avilés-Vázquez S, Chávez-González A, Hidalgo-Miranda A, et al. Global gene expression profiles of hematopoietic stem and progenitor cells from patients with chronic myeloid leukemia: the effect of in vitro culture with or without imatinib. *Cancer Med*. 2017;6(12):2942-2956.
56. Warfvinge R, Geironsen L, Sommarin MNE, et al. Single-cell molecular analysis defines therapy response and immunophenotype of stem cell subpopulations in CML. *Blood*. 2017;129(17):2384-2394.
57. Huang S-Y, Liu Y-H, Chen Y-J, Yeh Y-Y, Huang H-M. CD69 partially inhibits apoptosis and erythroid differentiation via CD24, and their knockdown increase imatinib sensitivity in BCR-ABL-positive cells. *J Cell Physiol*. 2018;233(9):7467-7479.
58. Steinbach D, Schramm A, Eggert A, et al. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. *Clin Cancer Res*. 2006;12(8):2434-2441.
59. Sadras T, Heatley SL, Kok CH, et al. Differential expression of MUC4, GPR110 and IL2RA defines two groups of CRLF2-rearranged acute lymphoblastic leukemia patients with distinct secondary lesions. *Cancer Lett*. 2017;408:92-101.
60. Ha K, Lee GE, Pali SS, et al. Rapid and transient recruitment of DNMT1 to DNA double-strand breaks is mediated by its interaction with multiple components of the DNA damage response machinery. *Hum Mol Genet*. 2011;20(1):126-140.
61. Barrett CW, Reddy VK, Short SP, et al. Selenoprotein P influences colitis-induced tumorigenesis by mediating stemness and oxidative damage. *J Clin Invest*. 2015;125(7):2646-2660.
62. Sengupta S, Yang C, Hegde ML, et al. Acetylation of oxidized base repair-initiating NEIL1 DNA glycosylase required for chromatin-bound repair complex formation in the human genome increases cellular resistance to oxidative stress. *DNA Repair (Amst)*. 2018;66-67:1-10.
63. Bordin F, Piovan E, Masiero E, et al. *WT1* loss attenuates the TP53-induced DNA damage response in T-cell acute lymphoblastic leukemia. *Haematologica*. 2018;103(2):266-277.
64. Oji Y, Tatsumi N, Kobayashi J, et al. Wilms' tumor gene *WT1* promotes homologous recombination-mediated DNA damage repair. *Mol Carcinog*. 2015;54(12):1758-1771.
65. Fuka G, Kantner H-P, Grausenburger R, et al. Silencing of *ETV6/RUNX1* abrogates PI3K/AKT/mTOR signaling and impairs reconstitution of leukemia in xenografts. *Leukemia*. 2012;26(5):927-933.
66. Kilbey A, Terry A, Wotton S, et al. Runx1 Orchestrates Sphingolipid Metabolism and Glucocorticoid Resistance in Lymphomagenesis. *J Cell Biochem*. 2017;118(6):1432-1441.
67. Imai N, Shikami M, Miwa H, et al. t(8;21) acute myeloid leukaemia cells are dependent on vascular endothelial growth factor (VEGF)/VEGF receptor type2 pathway and phosphorylation of Akt. *Br J Haematol*. 2006;135(5):673-682.
68. Du J, Li L, Ou Z, et al. *FOXC1*, a target of polycomb, inhibits metastasis of breast cancer cells [published correction appears in *Breast Cancer Res Treat*. 2016;156:609-610]. *Breast Cancer Res Treat*. 2012;131(1):65-73.
69. Ezhkova E, Pasolli HA, Parker JS, et al. Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell*. 2009;136(6):1122-1135.