

Supplemental Information

A high proportion of germline variants in pediatric Chronic Myeloid Leukemia

Krumbholz et al.

1. Supplemental Methods
2. Supplemental Figures
 - Supplemental Figure 1
 - Supplemental Figure 2
 - Supplemental Figure 3
 - Supplemental Figure 4
3. Supplemental Tables
 - Supplemental Table 1
 - Supplemental Table 2
 - Supplemental Table 3
 - Supplemental Table 4
 - Supplemental Table 5
4. References

1. Supplemental Methods

Patient samples and DNA isolation

To minimize the possible influence of different ethnic backgrounds for our analysis, we included only pediatric patients from the German pediatric CMLpaed II trial who were classified as individuals of European origin by SNP genotyping. Patients from foreign countries were excluded from this study, resulting in 118 pediatric patients with CML. Sufficient material and informed consent for genetic analyses were available in 62 individuals (one case was excluded due to insufficient data quality). The investigated cohort is representative of the whole CMLpaed II study cohort. Patient characteristics are listed in Supplemental Table 1.

DNA was isolated from frozen blood and/or bone marrow samples at diagnosis or in remission (*BCR-ABL1* <5%) using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) or AllPrep DNA/RNA Kit (Qiagen) according to the manufacturer's instructions. For samples at diagnosis 100 µl – 200 µl blood or bone marrow was isolated in dependency of the initial leukocyte count. For remission samples 400 µl of material was used for DNA isolation.

When available, saliva samples or fingernails were utilized as reference biomaterials in addition to remission samples. DNAs were isolated with the blackPREP Swab DNA Kit (Analytik Jena GmbH, Jena, Germany) for saliva samples and the innuPREP Forensic Kit (Analytik Jena GmbH, Jena, Germany) for fingernails according to the manufacturer's instructions.

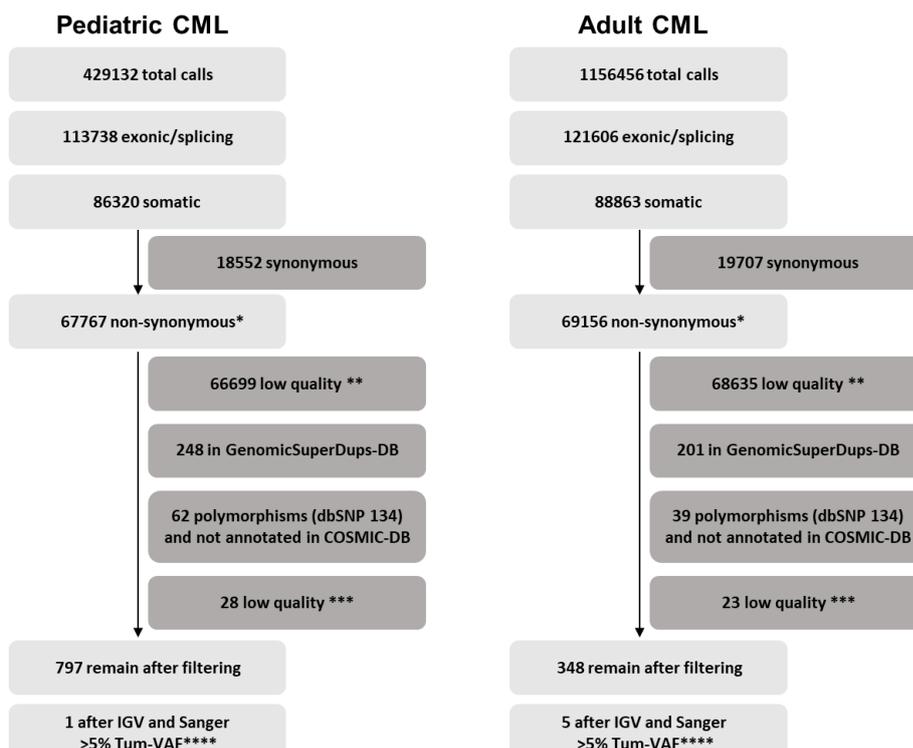
Whole exome sequencing

The libraries were sequenced on HiSeq 2000 with 100 bp paired-end reads using a 200-cycle TruSeq SBS v3 kit as described previously (1).

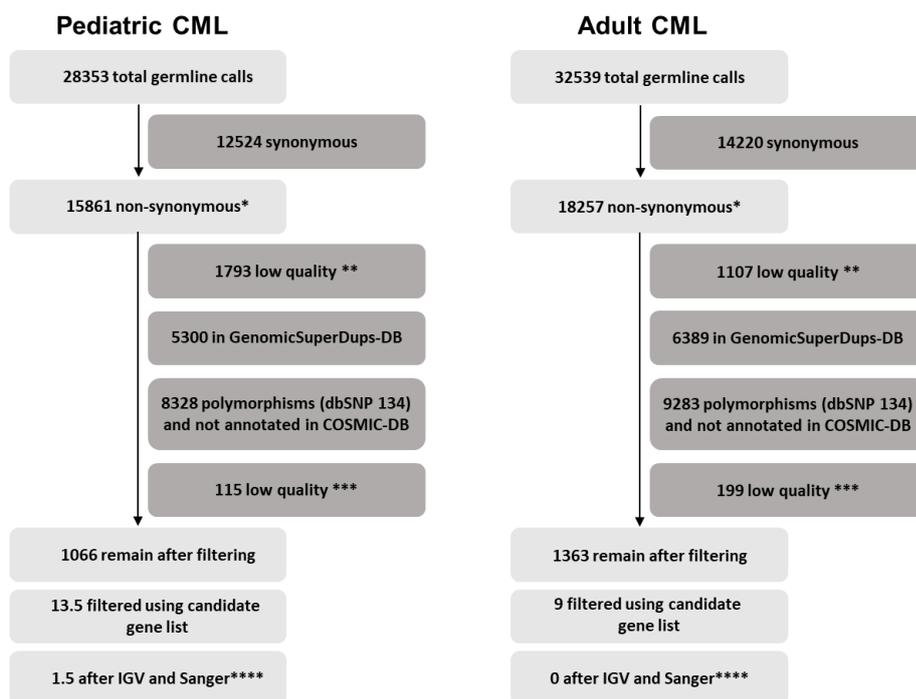
Filtering was performed on the output of VarScan2 v 2.3.9 after annotation using Annovar release 20150322. Total calls, shown in the flow chart for somatic calls top box, were classified as exonic, splicing, or exonic; splicing, based on Annovar annotation, and further as germline and somatic based on presence in tumor and remission samples. Calls were separated in categories somatic and germline and were filtered using the same criteria, as shown in the two flow charts.

Synonymous calls were filtered out (all non-synonymous, stop-gain, stop-loss, splicing, frameshift, and non-frameshift mutations were kept for further filtering, all calls are labeled as solely “non-synonymous” in the flow chart and marked with one asterisk to underline further call categories kept for further filtering). This step was followed by removing calls, which had very low coverage (tumor_reads2 < 4, marked with two asterisks, were in areas of segmental duplications (had an entry for GenomicSuperDups-database (DB)), or polymorphisms (had an entry for SNP-DB 134 but not for cancer database COSMIC-DB 70). Afterwards, another round of filtering of remaining mutations was done on a low number of reads (0 of tumor_reads2_plus and 0 of tumor_reads2_minus, removing calls with a strong disbalance of reads, marked with three asterisks). Remaining somatic calls showing tumor VAF higher as 5% were checked in later versions of the polymorphisms database dbSNP 150 and database of mutations in cancer, COSMIC 89, and if not present in healthy populations at frequency higher as 1%, were chosen for visual inspection in IGV. The remaining germline calls were screened using a candidate gene list of 87 genes. Both, identified somatic and germline calls, were taken in validation with Sanger sequencing.

Filtering statistics of somatic calls

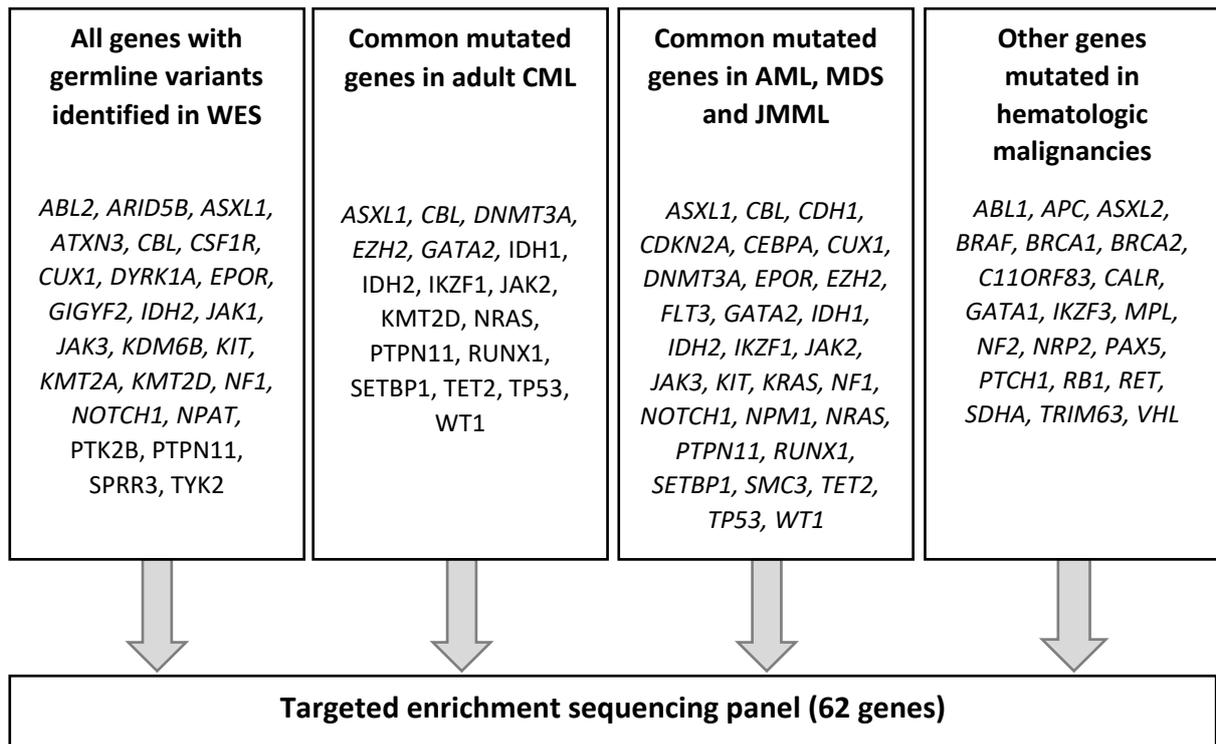


Filtering statistics of germline calls



Targeted enrichment sequencing panel

A custom amplicon panel, HaloPlex HS (Agilent Technologies, CA, USA), was designed for the targeted approach. A complete list of the 62 target genes is provided in Supplementary Table 2. In accordance with the whole exome sequencing data, all genes for which germline variants have been identified in pediatric and adult CML patients were included in the panel. In addition, genes commonly mutated in CML and other hematologic malignancies were included.



Exome-sequencing and amplicon sequencing data analysis

For the mutation analysis, a previously described pipeline was implemented (1). The reads were aligned using a BWA-MEM 0.7.10 aligner, SAMtools 0.1.19 was implemented for processing, optical duplicate removal was performed with Picard 2.18.26, and GATK 3.4-46 was applied for local realignment. BEDtools 2.24.0 was implemented for coverage calculation. Variant calling was performed with VarScan 2 version (v.) 2.3.9, and variants were annotated using ANNOVAR v. 20150322. Several databases were used for the annotation of the results: RefSeq for the prediction of coding, non-coding, and splicing-affecting mutations; genomicSuperDups for the detection of artifacts in repetitive regions; and esp6500, 1000G, and dbSNP138 for the retrieval of genetic polymorphisms; and COSMIC70 for the annotation of variants detected in cancers. Only variants with an effect on coding sequence (exonic and splicing) supported by at least six reads and at tumor allele frequency >5% were retained. Variants annotated in polymorphism databases and present in a healthy population at >1% frequency in database SNP134 were filtered out. In addition, germline variants were screened

for genes recurrently mutated in cancers and for genes of DNA damage repair pathways. All mutations were reviewed in IGV Genome Browser, checked in databases of later releases, namely, SNP150 and COSMIC 89, to exclude sequence variants present in healthy populations at >1% and retain pathologic variants, and were validated by Sanger sequencing.

The amplicon sequencing pipeline was identical, except that the reads were trimmed with cutadapt 1.18 to remove sequencing adaptors before alignment, and as the HaloPlex HS (Agilent Technologies) workflow contained molecular barcodes, both optical duplicates and sequence duplicates were removed using Picard 2.18.26, which reduced the number of artifacts. All mutations detected were validated by Sanger sequencing in tumor and remission samples for classification into germline and somatic categories.

Analyses of “Alexandrov” signatures

Known mutational processes were investigated with the sigProfilerExtractor.py script (v1.1.4, <https://github.com/AlexandrovLab/SigProfilerExtractor> (2) and mutation signatures v.3.2 deposited to COSMIC (3)) with 5000 replicates and corresponding reference genome GRCh37. Only somatic cells, both non-coding and coding, synonymous and non-synonymous, sequence variants, and mutations non-annotated in the respective polymorphism databases generated in the exome-sequencing pipeline were used as inputs for the analysis. To remove technology-related artifacts, the input lists were prefiltered in the following way: variants annotated in the genomicSuperDups database and variants annotated by ANNOVAR as “loss of heterozygosity,” as well as mutations in *MUC* genes (*MUC0-9*) were removed from the analysis. Only variants supported by at least 10 reads and not <5 forward reads and <5 reverse reads were taken into the analysis.

Classification and assessment of germline variants

Hematological Predictor of Pathogenicity (HePPy) is an ensemble in silico predictor combining 10 in silico predictor scores and 4 phylogenetic conservation scores trained on 371 well-

defined hematopathological somatic missense variants. Prediction scores were classified as previously described (HePPy (4): >0.75, probably pathogenic; 0.5-0.75, possibly pathogenic; <0.5 probably benign. AlphaMissense (5): >0.564, likely pathogenic; 0.340-0.564, ambiguous; ≤0.339, likely benign. VIPUR (6): >0.5, deleterious; <0.5, neutral.

The effect of stop and frameshift mutations on the protein domain architecture was assessed by a SMART search against the SMART and PFAM domain databases (7, 8). The PyMOL software was used to generate protein models illustrating the amino acid substitution.

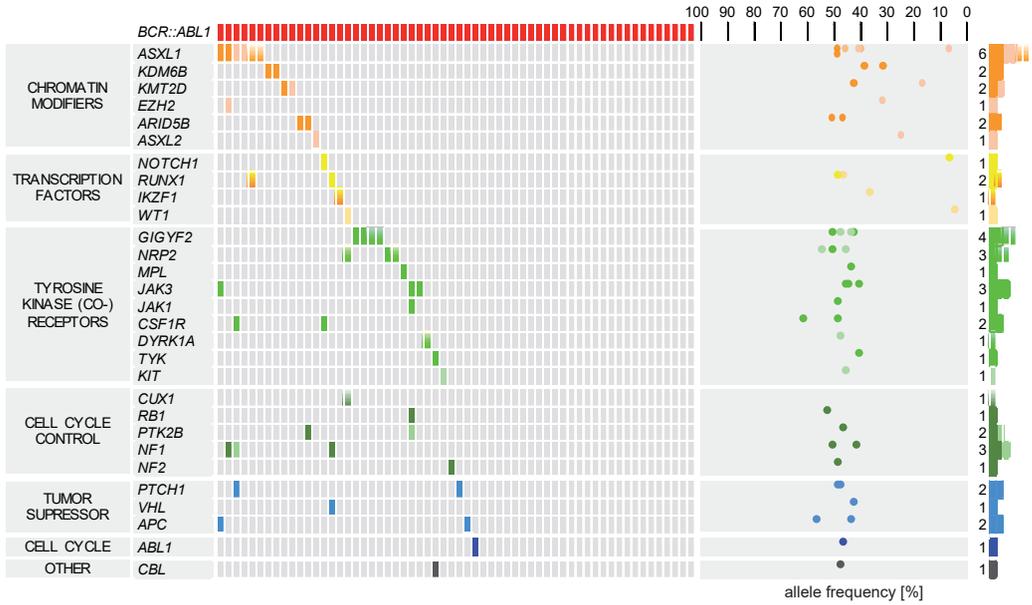
Statistical analyses

Statistical analyses were conducted using GraphPad prism software (v9.5.1) (GraphPad Software. San Diego, California USA, www.graphpad.com (2023)). To compare numbers of individuals affected with germline and/or somatic variants Fisher's exact test was used.

Supplemental Figures

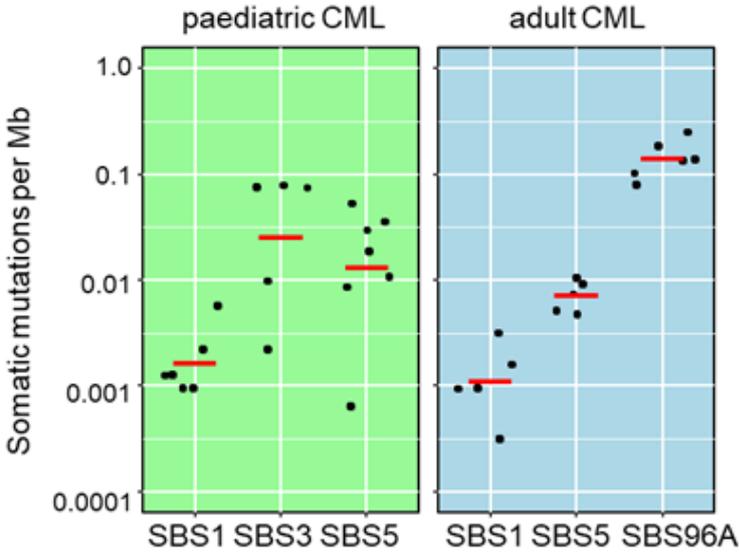
Supplemental Figure 1

A



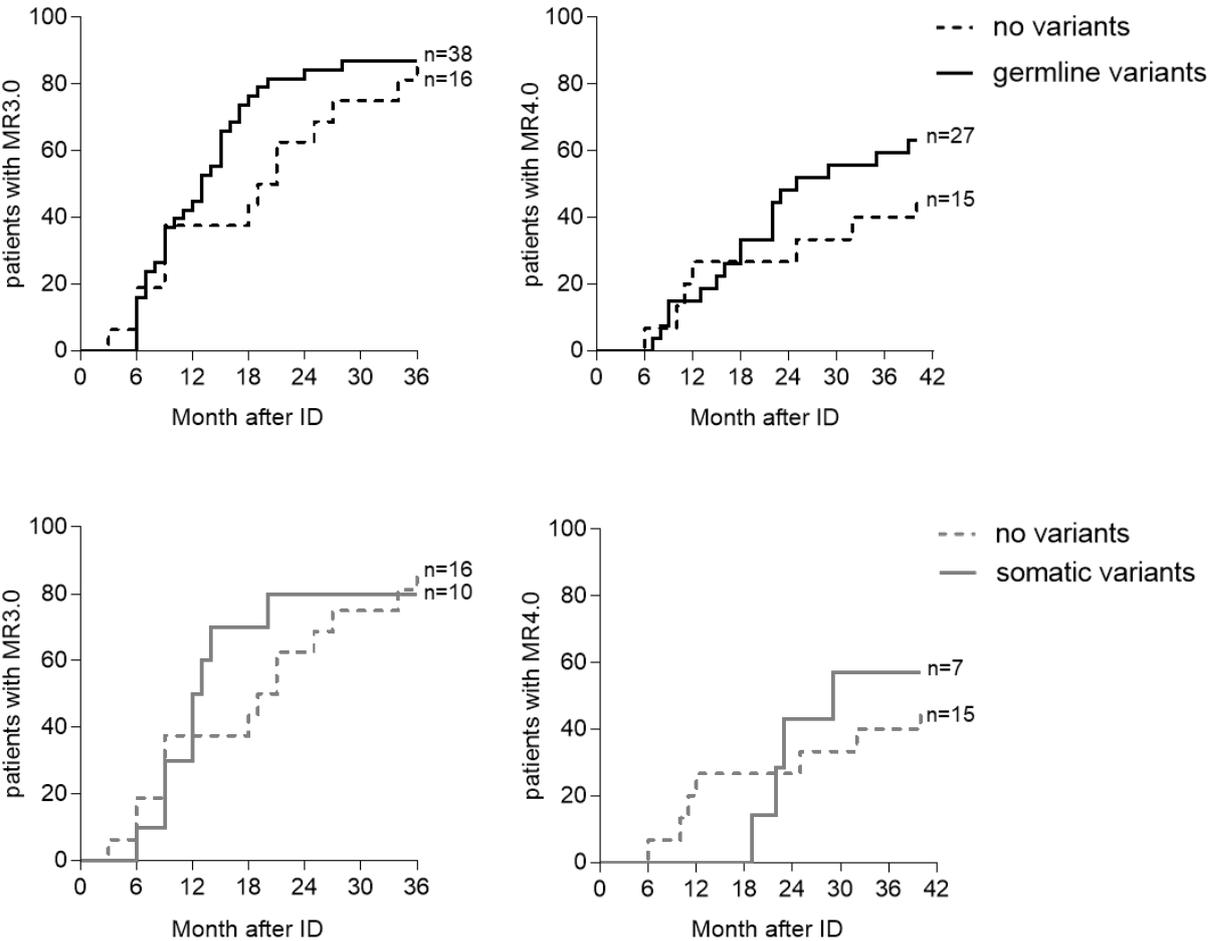
Supplemental Figure 1: Germline (dark colors) and somatic variants (light colors) in 61 adult patients with CML are displayed, with variant allele frequency (VAF). Variants with unknown mutation status are presented using gradient colors. Each column represents one patient, and genes are functionally grouped.

Supplemental Figure 2



Supplemental Figure 2: Results of somatic variant signature analysis for pediatric and adult patients with CML. The charts indicate the frequency of mutations per megabase linked to the specific mutation signature. Identified signatures include SBS1 (caused by spontaneous or enzymatic deamination of 5-methylcytosine to thymine), SBS3 (due to defective homologous recombination-based DNA damage repair), SBS5 (occurring more often in various cancer types) and SBS96A (with no identifiable pattern).

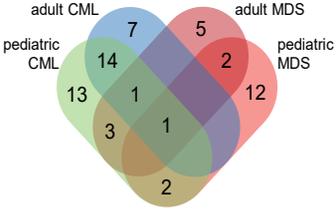
Supplemental Figure 3



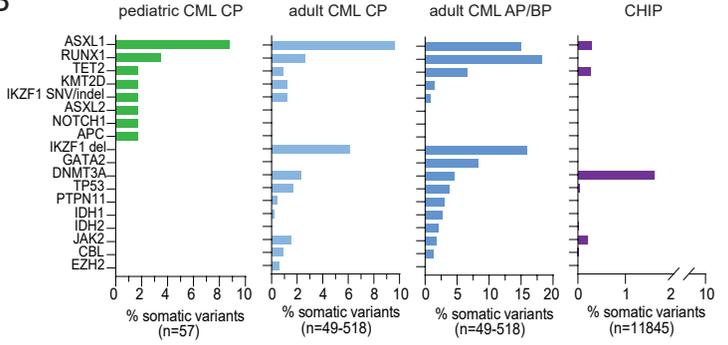
Supplemental Figure 3: Cumulative incidence of major molecular response (MR3.0; $BCR::ABL1 \leq 0.1\%$) or deep molecular response (MR4.0; $BCR::ABL1 \leq 0.01\%$) according to the $BCR::ABL1$ transcript level. Comparison of patients without detectable variants versus patients with somatic variants (upper graphs) or germline variants (lower graphs).

Supplemental Figure 4

A



B



Supplemental Figure 4: The variant spectrum of pediatric CML in comparison to other myeloid neoplasms in childhood and adulthood. (A) A Venn diagram displaying germline variants in pediatric CML, adult CML, pediatric MDS, and adult MDS highlights that childhood and adult CML share the most similarities. (B) The distribution pattern of somatic variants in pediatric and adult patients with CML differs from those present in CHIP variations.

3. Supplemental Tables

Supplemental Table 1

Overview of the study cohort (N=61) in comparison to all pediatric CML patients diagnosed in Germany and enrolled in the CMLpaed II cohort (N=118). None of the characteristics in the study cohort differed significantly from the characteristics of the overall CMLpaed II cohort.

Patient characteristics				
	CMLpaed II trial (Germany) N = 118		Study cohort N = 61	
	Number of patients	Percentage (%)	Number of patients	Percentage (%)
Sex				
Male	72	61.0	38	62.3
Female	46	39.0	23	37.7
Stage at diagnosis				
Chronic Phase	109	92.4	57	93.4
Accelerated Phase	4	3.4	2	3.3
Blastic Phase	5	4.2	2	3.3
Age at diagnosis (median, range)	13.5y (1.3-18)		11.9y (3.4-17.8)	
Transcript type				
b2a2	50	42.7	25	41.0
b3a2	60	51.3	34	55.7
b2a2 and b3a2	7	6.0	2	3.3
Treatment response*				
Optimal	60	53.6	29	48.3
Non optimal	52	46.4	31	51.7
Cytogenetics				
<i>BCR::ABL1</i> only			47	77.0
Variant translocation			3	4.9
ACA			7	11.5
unknown			4	6.6

* Patients with irregular TKI intake were excluded

ACA additional chromosomal aberrations

Supplemental Table 2

Gene list for targeted amplicon sequencing. All coding exons were designed (the corresponding genes are marked as “hotspots,” and the design was limited to exons recurrently associated with cancers).

<i>ABL1</i>	<i>CSF1R</i>	<i>JAK3</i>	<i>PTK2B</i>
<i>ABL2</i>	<i>CUX1</i>	<i>KDM6B</i>	<i>PTPN11</i>
<i>APC</i>	<i>DNMT3A</i>	<i>KIT</i>	<i>RB1</i>
<i>ARID5B</i>	<i>DYRK1A</i>	<i>KMT2A</i>	<i>RET</i>
<i>ASXL1</i>	<i>EPOR</i>	<i>KMT2D</i>	<i>RUNX1</i>
<i>ASXL2</i>	<i>EZH2</i>	<i>KRAS</i>	<i>SDHA</i>
<i>ATXN3</i>	<i>FLT3</i>	<i>MPL</i>	<i>SETBP1</i>
<i>BRAF</i>	<i>GATA1</i>	<i>NF1</i>	<i>SPRR3</i>
<i>BRCA1</i>	<i>GATA2</i>	<i>NF2</i>	<i>TET2</i>
<i>BRCA2</i>	<i>GIGYF2</i>	<i>NOTCH1</i>	<i>TP53</i>
<i>C11ORF83</i>	<i>IDH1</i>	<i>NPAT</i>	<i>TRIM63</i>
<i>CALR</i>	<i>IDH2</i>	<i>NPM1</i>	<i>TYK2</i>
<i>CBL</i>	<i>IKZF1</i>	<i>NRAS</i>	<i>VHL</i>
<i>CDH1</i>	<i>IKZF3</i>	<i>NRP2</i>	<i>WT1</i>
<i>CDKN2A</i>	<i>JAK1</i>	<i>PAX5</i>	
<i>CEBPA</i>	<i>JAK2</i>	<i>PTCH1</i>	

Supplemental Table 3

WES coverage statistics

Patient_ID	# reads on target	Average coverage	No coverage	1x Coverage	10x Coverage	15x Coverage	50x Coverage	120x Coverage
upn_136D	37.63%	179.6	0.22%	99.78%	97.98%	96.82%	86.63%	68.33%
upn_136-CR	38.43%	163	0.25%	99.75%	97.74%	96.42%	84.80%	63.94%
upn_15D	36.09%	150.6	1.75%	98.25%	88.41%	84.53%	65.75%	46.75%
upn_15-CR	36.40%	181.3	0.34%	99.66%	97.41%	96.06%	85.39%	68.05%
upn_315D	37.86%	94.6	1.49%	98.51%	90.31%	86.46%	61.78%	34.25%
upn_315-CR	23.32%	71	1.07%	98.93%	95.08%	92.85%	62.65%	18.93%
upn_327D	36.29%	59.2	2.46%	97.54%	84.32%	77.99%	43.09%	17.57%
upn_327-CR	37.49%	164.1	0.26%	99.74%	97.44%	95.95%	83.80%	63.72%
upn_331D	36.46%	214.4	0.25%	99.75%	98.43%	97.70%	90.46%	74.88%
upn_331-CR	34.76%	206.7	0.32%	99.68%	98.19%	97.43%	90.24%	74.88%
upn_338D	37.01%	148.4	0.36%	99.64%	96.45%	94.52%	80.26%	58.84%
upn_338-CR	36.90%	209.8	0.16%	99.84%	98.58%	97.77%	89.90%	73.99%
upn_1D	48.57%	48.2	8.25%	91.75%	74.14%	66.99%	36.64%	14.23%
upn_1-CR	48.09%	54.3	6.00%	94.00%	83.46%	78.28%	44.02%	14.27%
upn_248D	46.64%	39.4	7.87%	92.13%	74.96%	67.39%	29.86%	7.34%
upn_248-CR	45.39%	81.7	5.07%	94.93%	84.92%	80.85%	56.84%	30.13%
upn_255D	44.45%	55.4	5.81%	94.19%	83.87%	78.62%	43.43%	14.73%
upn_255-CR	45.95%	41.2	6.30%	93.70%	79.12%	71.75%	30.02%	7.51%
upn_295D	46.29%	81.3	5.39%	94.61%	84.58%	80.28%	55.47%	29.59%
upn_295-CR	47.65%	42.5	7.22%	92.78%	75.89%	68.96%	32.72%	8.77%
upn_308D	46.86%	67.9	5.89%	94.11%	82.69%	78.13%	49.04%	21.59%
upn_308-CR	44.40%	46.4	6.30%	93.70%	80.11%	73.61%	35.64%	10.14%
upn_336D	43.41%	39.4	6.11%	93.89%	78.70%	71.13%	28.26%	6.77%
upn_336-CR	43.29%	55.7	5.65%	94.35%	82.27%	77.01%	43.63%	15.13%
CM1D	44.78%	74.9	4.48%	95.52%	91.26%	89.31%	64.44%	14.59%
CM1-CR	43.67%	78.5	5.15%	94.85%	90.82%	88.94%	66.13%	17.59%
CM3D	31.43%	45.9	3.56%	96.44%	90.37%	85.37%	34.15%	3.25%
CM3-CR	28.70%	46.8	3.53%	96.47%	90.55%	85.77%	35.29%	3.36%
CM4D	41.09%	83	3.63%	96.37%	92.82%	91.12%	69.36%	16.92%
CM4-CR	28.37%	47.4	3.83%	96.17%	90.59%	86.55%	37.23%	3.17%
CM5D	34.14%	46.1	3.76%	96.24%	89.86%	84.77%	34.22%	3.46%
CM5-CR	42.82%	70.5	4.41%	95.59%	91.04%	88.70%	59.92%	12.81%
CM6D	27.66%	40.7	3.73%	96.27%	89.45%	83.77%	28.42%	1.92%
CM6-CR	39.70%	67	3.26%	96.74%	92.56%	89.95%	55.54%	9.92%
CM7D	28.07%	143.9	0.25%	99.76%	97.63%	96.16%	82.93%	59.12%
CM7-CR	28.25%	102.3	0.36%	99.64%	96.35%	93.97%	71.91%	39.77%
CM8D	28.84%	96.1	0.41%	99.59%	96.00%	93.41%	69.84%	36.95%
CM8-CR	28.51%	88.5	0.47%	99.53%	94.86%	91.69%	65.49%	32.81%
CM9D	27.24%	86.3	0.47%	99.53%	94.90%	91.71%	64.91%	31.50%
CM9-CR	28.17%	89.5	0.44%	99.56%	95.40%	92.39%	65.79%	32.79%
CM10D	26.56%	130.2	0.19%	99.81%	97.52%	95.81%	79.61%	52.84%
CM10-CR	28.10%	121	0.19%	99.81%	97.41%	95.59%	78.14%	48.99%
CM11D	29.79%	69.3	0.40%	99.60%	94.06%	89.98%	55.23%	21.13%
CM11-CR	28.51%	97.6	0.28%	99.72%	96.01%	93.34%	69.90%	37.42%
CM12D	29.66%	97	0.28%	99.72%	95.88%	93.13%	69.60%	37.19%
CM12-CR	29.63%	113.4	0.24%	99.76%	96.66%	94.46%	75.59%	45.92%

Supplemental Table 4

Amplicon sequencing coverage statistics

SampleID	Average coverage	No coverage	1x Coverage	10x Coverage	15x Coverage	50x Coverage	120x Coverage	200x Coverage	500x Coverage	1000x Coverage
126D	354.195	2.22%	97.78%	95.85%	94.82%	88.16%	79.10%	75.42%	59.92%	23.73%
128D	404.466	2.13%	97.87%	95.19%	94.02%	87.60%	80.52%	77.70%	67.44%	32.91%
195D	481.647	2.17%	97.83%	96.15%	95.67%	92.22%	86.92%	85.16%	76.96%	42.74%
210D	483.697	1.84%	98.16%	96.65%	96.02%	93.08%	88.09%	85.71%	77.56%	42.78%
283D	529.968	2.24%	97.76%	95.48%	94.80%	91.56%	86.75%	84.98%	77.65%	47.38%
296D	236.337	3.28%	96.72%	94.63%	93.56%	85.97%	74.45%	69.64%	51.43%	8.78%
303D	439.062	2.07%	97.93%	96.30%	95.57%	91.72%	85.91%	83.56%	74.15%	37.01%
304Rz	186.817	3.42%	96.58%	93.75%	92.51%	82.76%	67.97%	61.69%	39.81%	3.44%
305D	380.409	1.95%	98.05%	96.36%	95.76%	92.05%	85.41%	82.47%	71.72%	29.19%
309D	446.392	1.94%	98.06%	96.47%	95.73%	92.39%	86.57%	84.48%	75.37%	37.46%
310D	225.064	3.64%	96.36%	92.97%	91.89%	83.87%	71.53%	66.12%	48.12%	8.19%
312D	296.796	3.59%	96.41%	91.82%	89.95%	79.31%	68.52%	64.67%	50.87%	18.84%
313D	268.497	2.56%	97.44%	93.70%	91.91%	83.59%	73.47%	69.58%	54.71%	15.13%
314D	258.466	2.48%	97.52%	94.18%	92.97%	86.30%	76.90%	72.70%	56.40%	11.44%
316D	106.318	3.47%	96.53%	92.98%	90.93%	73.36%	47.25%	37.51%	12.33%	0.04%
317D	408.279	1.88%	98.12%	96.61%	96.02%	92.75%	86.73%	84.29%	74.64%	34.09%
323D	583.652	1.81%	98.19%	96.65%	96.33%	94.06%	90.12%	88.55%	81.69%	53.47%
326D	134.422	3.52%	96.48%	88.70%	85.67%	71.27%	54.18%	47.57%	25.37%	0.66%
328D	643.544	1.75%	98.25%	96.68%	96.27%	93.35%	89.07%	87.30%	81.06%	55.14%
329D	137.086	2.27%	97.73%	94.65%	93.23%	80.97%	61.41%	53.79%	23.62%	0.07%
334D	303.921	2.17%	97.83%	95.86%	95.08%	89.47%	81.31%	77.95%	64.15%	17.49%
335D	255.174	3.65%	96.35%	92.84%	91.93%	85.55%	74.90%	70.70%	53.68%	12.20%
337D	492.071	1.81%	98.19%	96.65%	96.05%	92.55%	86.78%	84.49%	75.77%	42.94%
340D	490.227	2.07%	97.93%	96.17%	95.46%	91.37%	85.73%	83.83%	75.44%	42.87%
344D	437.81	2.97%	97.03%	93.95%	93.13%	87.89%	81.73%	79.76%	70.76%	37.38%
345D	150.932	3.34%	96.66%	93.83%	92.28%	79.14%	60.26%	52.90%	28.82%	1.18%
347D	557.854	1.77%	98.23%	96.56%	96.07%	92.95%	87.97%	86.16%	78.92%	48.22%
348D	454.619	1.74%	98.26%	96.74%	96.20%	92.57%	86.83%	84.26%	75.05%	37.97%
349D	823.09	1.73%	98.27%	96.73%	96.15%	93.32%	89.08%	87.50%	82.07%	61.62%
350D	462.404	2.81%	97.19%	95.72%	95.20%	91.71%	86.30%	83.81%	73.78%	39.17%
352D	293.848	3.20%	96.80%	94.81%	93.93%	87.79%	78.01%	74.05%	59.02%	17.83%
353D	404.269	1.92%	98.08%	96.33%	95.76%	92.09%	85.74%	83.40%	73.30%	32.64%
354R	448.909	1.84%	98.16%	96.71%	96.25%	93.23%	87.93%	85.92%	77.27%	39.06%
356D	386.306	1.87%	98.13%	96.68%	96.01%	92.53%	86.12%	83.54%	73.02%	30.02%
357D	420.013	2.27%	97.73%	95.70%	94.93%	90.86%	84.70%	82.59%	73.34%	34.77%
359D	466.97	1.78%	98.22%	96.20%	95.47%	91.72%	86.18%	84.08%	75.34%	40.09%
360D	34.7776	4.33%	95.67%	84.00%	75.79%	24.96%	2.11%	0.78%	0.00%	0%
362D	600.925	1.76%	98.24%	96.79%	96.22%	93.98%	89.99%	88.29%	81.47%	54.05%
363D	161.085	3.37%	96.63%	94.05%	92.70%	81.59%	64.05%	56.75%	32.40%	1.35%
365D	375.815	2.07%	97.93%	96.21%	95.63%	91.51%	84.96%	82.20%	71.09%	28.73%
367R	423.864	1.88%	98.12%	96.50%	95.89%	91.86%	85.66%	83.34%	73.42%	34.94%
370D	353.491	2.99%	97.01%	94.03%	92.93%	86.75%	79.80%	77.10%	65.48%	26.94%
373D	365.743	2.09%	97.91%	95.91%	95.18%	90.45%	83.55%	81.11%	69.28%	26.99%
374D	475.234	1.96%	98.04%	96.34%	95.65%	92.22%	86.95%	84.82%	75.70%	40.91%
375D	223.34	5.22%	94.78%	84.94%	81.40%	66.33%	52.69%	48.49%	35.91%	12.95%
378D	319.55	2.14%	97.86%	96.28%	95.46%	90.63%	82.85%	79.74%	66.54%	19.92%
379D	416.409	1.75%	98.25%	96.59%	96.02%	92.80%	86.71%	84.42%	74.57%	34.43%
55D	395.711	3.55%	96.45%	94.02%	93.05%	87.41%	79.99%	77.08%	66.27%	32.14%
60D	198.316	3.51%	96.49%	92.55%	91.28%	82.84%	69.60%	64.25%	43.46%	4.59%

96D	604.47	1.85%	98.15%	96.85%	96.46%	93.92%	90.07%	88.38%	81.83%	54.75%
Ad01	388.223	2.30%	97.70%	95.39%	94.51%	90.55%	84.19%	81.63%	71.28%	30.66%
Ad02	498.359	2.34%	97.66%	95.24%	94.53%	90.29%	84.81%	82.89%	74.91%	43.69%
Ad03	551.847	2.20%	97.80%	95.80%	95.02%	91.95%	87.18%	85.32%	78.20%	49.26%
Ad04	548.086	2.22%	97.78%	95.92%	95.11%	91.61%	86.83%	84.93%	78.03%	49.19%
Ad05	406.894	3.11%	96.89%	93.94%	92.94%	87.48%	81.04%	78.76%	69.37%	34.09%
Ad06	466.148	2.83%	97.17%	94.49%	93.64%	89.00%	83.33%	81.23%	73.23%	40.44%
Ad07	284.077	2.16%	97.84%	96.14%	95.55%	90.40%	81.52%	78.15%	61.84%	14.28%
Ad08	603.94	1.48%	98.52%	97.24%	96.66%	93.77%	89.45%	87.86%	80.77%	53.33%
Ad09	553.605	1.56%	98.44%	97.07%	96.47%	93.12%	87.99%	85.93%	78.22%	48.66%
Ad10	492.81	1.61%	98.39%	96.82%	96.30%	93.37%	87.88%	85.95%	77.96%	42.86%
Ad11	598.005	1.58%	98.42%	97.23%	96.74%	94.33%	90.27%	88.57%	81.54%	53.13%
Ad12	415.079	1.87%	98.13%	96.30%	95.49%	91.46%	85.32%	83.03%	72.83%	34.07%
Ad13	473.708	1.82%	98.18%	96.63%	96.16%	92.59%	86.91%	84.81%	75.75%	40.71%
Ad14	186.798	2.13%	97.87%	95.44%	94.37%	84.79%	70.17%	64.04%	40.83%	2.23%
Ad15	222.067	2.20%	97.80%	95.09%	93.77%	83.57%	68.52%	63.06%	43.67%	9.73%
Ad16	476.59	1.90%	98.10%	96.09%	95.03%	89.90%	83.49%	80.92%	71.22%	40.36%
Ad17	540.362	1.97%	98.03%	96.50%	95.90%	92.26%	86.86%	85.25%	77.16%	46.59%
Ad18	632.068	1.73%	98.27%	96.84%	96.32%	93.57%	89.64%	87.98%	81.61%	55.91%
Ad19	503.41	1.67%	98.33%	96.74%	96.09%	92.69%	87.70%	85.82%	77.91%	44.94%
Ad20	331.892	1.81%	98.19%	96.28%	95.50%	90.01%	82.46%	79.35%	66.06%	22.02%
Ad21	386.624	1.80%	98.20%	96.28%	95.67%	91.64%	84.50%	81.94%	71.72%	29.81%
Ad22	405.045	1.85%	98.15%	96.43%	95.72%	91.87%	85.30%	82.80%	72.77%	32.81%
Ad23	452.955	1.87%	98.13%	96.71%	96.18%	93.20%	87.66%	85.49%	76.77%	38.97%
Ad24	491.536	1.80%	98.20%	96.81%	96.30%	92.87%	87.50%	85.45%	77.43%	42.61%
Ad25	513.2	1.68%	98.32%	96.83%	96.36%	93.53%	88.54%	86.63%	78.54%	45.67%
Ad26	727.01	1.60%	98.40%	97.16%	96.69%	94.85%	91.74%	90.42%	85.19%	62.79%
Ad27	406.641	1.87%	98.13%	96.54%	95.98%	92.55%	86.54%	83.99%	73.80%	32.76%
Ad28	462.574	1.90%	98.10%	96.38%	95.64%	91.19%	85.06%	82.83%	73.81%	40.11%
Ad29	505.413	1.84%	98.16%	96.51%	95.88%	92.24%	86.52%	84.55%	76.10%	44.36%
Ad30	485.564	2.71%	97.29%	96.13%	95.72%	92.81%	87.55%	85.24%	75.66%	41.33%
Ad31	302.065	2.68%	97.32%	95.64%	94.86%	88.57%	79.13%	75.09%	60.02%	18.51%
Ad32	560.688	2.61%	97.39%	96.25%	95.92%	93.30%	88.44%	86.54%	78.12%	48.31%
Ad33	609.315	2.70%	97.30%	96.25%	95.92%	93.50%	89.09%	87.52%	80.20%	52.29%
Ad34	490.799	2.67%	97.33%	96.11%	95.76%	93.15%	87.65%	85.65%	76.26%	42.64%
Ad35	573.062	2.70%	97.30%	96.31%	95.96%	93.35%	88.11%	86.25%	78.39%	49.04%
Ad36	477.385	2.73%	97.27%	96.01%	95.49%	91.98%	85.78%	83.33%	73.46%	40.26%
Ad37	394.747	2.71%	97.29%	95.84%	95.33%	91.18%	83.77%	80.66%	68.62%	30.93%
Ad38	413.644	2.72%	97.28%	96.02%	95.60%	91.96%	85.14%	82.26%	71.01%	33.09%
Ad39	303.884	14.69%	85.32%	74.13%	71.81%	65.60%	59.68%	57.78%	50.12%	24.47%
Ad40	605.98	1.60%	98.40%	97.06%	96.64%	94.41%	89.96%	88.15%	81.26%	52.78%
Ad41	590.762	1.62%	98.38%	96.93%	96.42%	94.09%	89.81%	87.99%	80.87%	52.29%
Ad42	655.026	1.63%	98.37%	97.13%	96.77%	94.92%	91.89%	90.16%	84.31%	58.96%
Ad43	649.905	1.65%	98.35%	97.05%	96.66%	94.68%	91.14%	89.59%	83.29%	57.36%
Ad44	702.827	1.84%	98.16%	96.89%	96.57%	94.46%	90.66%	89.39%	83.44%	60.23%
Ad45	448.276	1.83%	98.17%	96.74%	96.25%	93.52%	88.13%	86.20%	77.74%	39.45%
Ad46	546.412	1.96%	98.04%	96.62%	96.12%	93.60%	88.97%	86.94%	80.17%	49.30%
Ad47	527.556	1.86%	98.14%	96.84%	96.32%	93.82%	89.26%	87.40%	80.17%	47.91%
Ad48	555.22	1.87%	98.13%	96.76%	96.28%	93.89%	89.64%	88.05%	81.23%	51.05%
Ad49	465.01	2.05%	97.95%	96.14%	95.43%	91.58%	85.86%	83.86%	75.57%	41.18%

Supplemental Table 5

Result of the trio analysis of an additional 12 patients to investigate the presence of germline variants in the parents of affected individuals.

Patient #	Gene	Mother	Father
136	<i>GIGYF2</i> *	Yes	Yes
	<i>CSFR1</i>	Yes	No
248	<i>KIT</i>	Yes	n.a.
	<i>NPAT</i>	Yes	n.a.
308	<i>GIGYF2</i>	No	Yes
	<i>PTPN11</i>	No	Yes
314	<i>GIGYF2</i>	No	Yes
316	<i>NF2</i>	No	Yes
	<i>PTK2B</i>	No	Yes
327	<i>NOTCH1</i>	No	Yes
	<i>IDH1</i> *	Yes	No
329	<i>ASXL2</i>	Yes	No
331	<i>SPRR3</i>	Yes	Yes
	<i>ABL2</i>	Yes	No
338	<i>SPRR3</i>	Yes	Yes
348	<i>KMT2A</i>	Yes	No
352	<i>GIGYF2</i>	Yes	No
	<i>SDHA</i>	No	Yes
354	<i>TET2</i>	Yes	No

* SNPs with allele frequencies of >1%

n.a., not available

References

1. Cacciardi S, Dolnik A, Kapp-Schwoerer S, Rucker FG, Lux S, Blatte TJ, et al. Clonal evolution patterns in acute myeloid leukemia with NPM1 mutation. *Nat Commun.* 2019;10(1):2031.
2. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. *Nature.* 2020;578(7793):94-101.
3. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* 2019;47(D1):D941-D7.
4. Hutter S, Baer C, Walter W, Kern W, Haferlach C, Haferlach T. A Novel Machine Learning Based in silico Pathogenicity Predictor for Missense Variants in a Hematological Setting. *Blood.* 2019;134(Supplement_1):2090-.
5. Cheng J, Novati G, Pan J, Bycroft C, Zengulyte A, Applebaum T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science.* 2023;381(6664):eadg7492.
6. Baugh EH, Simmons-Edler R, Muller CL, Alford RF, Volfovsky N, Lash AE, et al. Robust classification of protein variation using structural modelling and large-scale data integration. *Nucleic Acids Res.* 2016;44(6):2501-13.
7. Letunic I, Khedkar S, Bork P. SMART: recent updates, new developments and status in 2020. *Nucleic Acids Res.* 2021;49(D1):D458-D60.
8. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* 2021;49(D1):D412-D9.