Materials and Methods

Patient samples

The patients provided informed consent to collect blood and marrow for research in accordance with the Declaration of Helsinki and the Belmont Report. The University of Utah Institutional Review Board approved the protocol (IRB #45880). Peripheral blood (PB) and BM samples were obtained both at initial diagnosis and on imatinib treatment as detailed in Supplementary Table 1. Mononuclear cells (MNCs) were isolated from PB or BM by density gradient separation. $CD34^+$, $CD14^+$, and $CD3^+$ cells were enriched by AutoMACS magnetic beads (Miltenyi, Auburn, CA). Sorted cell fractions were confirmed to be >90% pure by fluorescence-activated cell sorting (FACS).

Whole exome sequencing

DNA exome enrichment was performed using the SureSelect Human All Exon + UTR (v5) Kit (Agilent Santa Clara, CA). Enriched exomes were sequenced on an Illumina HiSeq 2500 sequencer, as described[6].

Sequenom MassARRAY

Further validation of the detected mutations in serial samples was performed using the MassARRAY platform (Sequenom, San Diego, CA). Briefly, multiplexed PCR was initially performed using Sequenom's iPlexPro chemistry. The PCR contained: 0.8 µL H_2O (molecular biology grade), 0.5 µL PCR buffer, 0.4 µL 25 mM $MgCl_2$, 0.1 µL 25 mM dNTPs, 0.2 µL PCR enzyme (5 U/µL), 1 µL multiplexed primer mix (Supplementary Table 4), and 2 µL DNA 10 ng/µL in Tris-EDTA. PCR was performed with a Mastercycler ProS (Eppendorf), using the following conditions: denaturation at 95°C for 2 minutes followed by 45 cycles of 95°C for 30 seconds, 56°C for 30 seconds, and 72°C for 1 minute, followed by 72°C for 5 minutes. The PCR products were treated with a cocktail of 1.53 µL H_2O, 0.17 µL 10× SAP Buffer and 0.3 µL shrimp alkaline phosphatase (1.7 U/µL) (SAP) in a thermal cycler at 37°C for 40 minutes followed by 85°C for 5 minutes. The extend PCR contained 7 µL SAP treated PCR products, 0.62 µL H_2O, 0.2 µL 10x iPLEX buffer, 0.2 µL 10x Termination Mix, 0.041 µL Thermo Sequenase (5 U/µL) and 0.94 µL of the multiplexed extend primer mix. The extend PCR was performed in a thermal cycler,

using the following conditions: denaturation at 95°C for 30 seconds followed by 40 cycles of 95°C for 5 seconds, 52°C for 5 seconds, 80°C for 5 seconds (the 52 and 80°C steps were repeated 5 times), followed by 72°C for 3 minutes. A total of 41 µL of molecular grade water and ion exchange resin was added to each sample. The plate was rotated for approximately 30 minutes and centrifuged at 4000 rpm for 5 minutes. Samples were spotted on the SpectroCHIP II G96 using the MassARRAY Nanodispenser. Results were visualized on the MassARRAY analyzer 4 system, using the autorun settings. Data was analyzed using Sequenom Typer version 4.0.

Sanger sequencing

Four mutations were further validated by amplification of the target region followed by Sanger sequencing. Primers were designed using Primer3Web version 4.0.0 (Supplementary Table 5).

Colony genotyping

$CD34^+$ cells from samples at diagnosis, day 67 and day 78 after starting imatinib were suspended in MethoCult™ H4230 (StemCell Technologies) in the presence of Stem Span CC100 (StemCell Technologies, Vancouver, BC), 10ng/ml GM-CSF (Miltenyi) and 10ng/ml G-CSF (Amgen, USA). After 2 weeks, at least 100 individual colonies were picked from each sample. DNA and RNA were extracted using the AllPrep DNA/RNA Micro kit (Qiagen). RNA was converted to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA). The cDNA from each colony was used for detection of the *BCR-ABL*1 transcripts by real-time quantitative PCR (RT-qPCR), using glucuronidase b (*GUSB)* as the housekeeping gene[7,8].

Quantitation of mutant alleles by pyrosequencing

Primers for amplification and sequencing of the mutant alleles were designed using QIAGEN PyroMark Assay Design software (Supplementary Table 6). The amplified fragments were subjected to pyrosequencing in accordance with manufacturer's instructions in PyroMark Q24 (Pyrosequencing Qiagen, Germany). The percentages of mutant and non-mutated alleles were determined by the Allele Quantitation Algorithm using PyroMark Q24 Software version 2.0.7.

Bioinformatics

Fastq files generated by the Illumina HiSeq 2500 sequencer were aligned to the human genome (build 37) with Novoalign (version 2.08.01); potential PCR duplicates were eliminated following standard GATK protocols (version 3.3-2). Samtools (version 1.0) was used to generate mpileup files, and VarScan 2 (version 2.3.7) was used to call variants. Three population databases (1000 genomes, NHLBI 6500, and UCSC genome browser) were used to cull SNPs found to occur at >1 in 500 alleles in any of these databases. Annovar was used to annotate mutations. Non-exonic mutations (except for slice-site mutations) were filtered out[9].

Supplementary Table 1. Clinical and Laboratory Parameters

	Diagnosis	Day 67	Day 78	Day 92	Day 124
Disease status	CML-AP	CML-AP	CML-AP	CMML-1	CMML-1
Treatment	None	Imatinib	Imatinib	Imatinib, hydrea	Imatinib, 5-azacytidine
WBC count (10^3/µL)	270	19	15	73	21
Segmented neutrophils (%)	33	66	56	63	72
Bands (%)	19	1	0	10	0
Metamyelocytes (%)	5	0	1	5	0
Myelocytes (%)	42	0	0	2	0
Promyelocytes (%)	1	0	0	0	0
Myeloblasts (%)	0	0	0	0	0
Basophils (%)	0	1	2	0	0
Eosinophils (%)	0	0	0	0	0
Monocytes (%)	0	23	26	19	16
Lymphocytes (%)	0	9	15	1	12
Hemoglobin (g/dL)	9	8.8	8.7	9	9
Platelets (10^3/µL)	55	87	82	80	90
Cytogenetics (Ph^+ met)	100%	NA	NA	0%	NA
BM blast count (%)	1.1	NA	NA	2.7	NA
PB blast count (%)	0	0	0	0	0
***BCR-ABL1* % (IS)**	10.1%	1.8	NA	0.12	NA

NA – not available

Supplementary Table 2. Somatic SNVs Detected by WES on Day 92

Gene	Chromosome	Position	Ref	Variant	Amino Acid Exchange	COSMIC ID
EZH2	7	148506219	T	C	I669M	1448968
KRAS	12	25398285	C	G	G12R	516
MSLN	16	816896	C	A	P462H	NR
NTRK3	15	88669547	C	T	V443I	NR

NR – not reported

Supplementary Table 3. Sequenom MassARRAY Analysis

							Mononuclear cells									
					CD3$^+$ Cells		Diagnosis		Day 67		Day 78		Day 92		Day 124	
Gene	Ch	Position	W	V	WT	Var	WT	Var	WT	Var	WT	Var	WT	Var	WT	Var
ASXL1	20	20031022937	C	Del	0.52	0.47	0.5	0.5	0.5	0.5	0.49	0.51	0.5	0.5	0.44	0.56
EZH2	7	7148506219	T	C	1	0	1	0	0.81	0.19	0.82	0.18	0.76	0.24	0.75	0.25
KRAS	12	12025398285	C	G	1	0	1	0	0.76	0.24	0.82	0.18	0.63	0.37	0.57	0.43
MSLN	16	16000816896	C	A	1	0	1	0	0.82	0.19	0.9	0.1	0.77	0.23	0.74	0.26
NTRK3	15	15088669547	C	T	1	0	1	0	0.67	0.33	0.8	0.2	0.6	0.4	0.6	0.4
TET2	4	4106156316	T	Del	0.74	0.26	0.73	0.27	0.73	0.27	0.73	0.27	0.68	0.32	0.72	0.28
TET2	4	4106196599	A	Del	0.5	0.5	0.49	0.51	0.53	0.47	0.48	0.52	0.48	0.52	0.48	0.52

Ch - chromosome; W - wild type nucleotide; V - variant nucleotide; WT - wild type allele; Var - variant allele.

Supplementary Table 4. Sequences of Sequenom Primers

Gene	Ch	Position	Forward PCR Primer Sequence	Reverse PCR Primer Sequence	Extend PCR Primer Sequence
FAT3	11	11092534968	ACGTTGGATGTGACAATGCACCAGTCTTCG	ACGTTGGATGTAACGTCGGATGTGTCTCTG	CGCGCAGGAAGTGTACC
TET2	4	4106156316	ACGTTGGATGGATTCCTTTTCTGCCACTAC	ACGTTGGATGCTGAGGAACCTGTGGAAGAG	CCATCACAATTGCTTCTT
MKL1	22	22040815309	ACGTTGGATGACAGGGCTGATTTGGTCTTG	ACGTTGGATGAGCTGAAGTTGCGATCACTG	GAGCTGATTGAGCGCCTTC
FAT1	4	4187525083	ACGTTGGATGCGGATAGATGCTCTCCTCAA	ACGTTGGATGGGCAGATAATGGAAAGCCTC	GCTCTCCTCAATTACCCTAA
EZH2	7	7148506219	ACGTTGGATGAAAACAGCTCTTCGCCAGTC	ACGTTGGATGGCAGTTATGATGGTTAACGG	GGCTCTCTTGGCAAAAATACC
TET2	4	4106196599	ACGTTGGATGGAAACCTATCAGTGGACAAC	ACGTTGGATGACCTATACAGATCCATCGGC	GGGAGAATAGGAACCCAGATA
NTRK3	15	15088669547	ACGTTGGATGTGGACCGTCGACCATATTTG	ACGTTGGATGGGTATCCATAGCAGTTGGAC	GTGTCCTGTTGGTGGTTCTCTTC
ZNF521	18	18022902100	ACGTTGGATGACAGCTTCGTCTTCCAACTC	ACGTTGGATGAAGACTGAAGATGGAGAGGC	TGAAGATGGAGAGGCACTAGATT
FAT1	4	4187539216	ACGTTGGATGCCACACTTTGTGACTGATCC	ACGTTGGATGGTAATTCAGATCAGGGCATC	CTGACTCAGGAACCAAC
GATA5	20	20061050495	ACGTTGGATGACAGGTAGGACAGCATCGAG	ACGTTGGATGTACGCCGACTCGGGCTCCTT	ACAAACATCGGAGAGCCG
KRAS	12	12025398285	ACGTTGGATGGCTGTATCGTCAAGGCACTC	ACGTTGGATGAGGCCTGCTGAAAATGACTG	ACTTGTGGTAGTTGGAGCT
UBE2S	19	19055912995	ACGTTGGATGGTCGGTGGAGGAAGCTTCA	ACGTTGGATGTCTGCTCACAGAGATCCACG	TTCGGCCCTGCCGCTGGGCC
ERCC4	16	16014015921	ACGTTGGATGCTGCCCTGTATTAAATAGCC	ACGTTGGATGTCATTTGTTACACGGCGAGG	AATCAGCTGAAGATAGAAGGA
METTL13	1	1171752976	ACGTTGGATGTGTATGATGTGGGCTATCGG	ACGTTGGATGGGGTGGCATTACATTCCTTC	TTACATTCCTTCATTTGCTTGA
MYC	8	8128753155	ACGTTGGATGTCTGAAGAGGACTTGTTGCG	ACGTTGGATGCAAGAGTTCCGTAGCTGTTC	AGAGGACTTGTTGCGGAAACGAC
ALK	2	2029448423	ACGTTGGATGAGAGAGGATCAGCGAGAGTG	ACGTTGGATGTGGTGTGGTTGTCAATACCC	TGGGACCTGTCTTCCAGTGTCACCC
MSLN	16	16000816896	ACGTTGGATGTCACTGTCCACCCACCGTGT	ACGTTGGATGTTGGGATAGAGGACGTCCAG	ACGTGTCCAGGTCCTGG
BDNF	11	11027681197	ACGTTGGATGTCTAGCTAAGAAAGCTCAAC	ACGTTGGATGAGTAGGATAAACTCAGAGCG	TGTGTGTGTGTGTGTGTG
TCF3	19	19001627417	ACGTTGGATGTCTCCCGAAGGAGGCATAG	ACGTTGGATGTCCCACTAACCTTCCTTCTC	TCCCCTTGCAGGCAAGAGC
ASXL1	20	20031022937	ACGTTGGATGGAACTGAATGTGAGTCTGGC	ACGTTGGATGGAGACAGAATGGGACCATTG	AATGGGACCATTGTCTGCAG
MLL3	7	7151945334	ACGTTGGATGAGAGTCCATCAATCCAGTAG	ACGTTGGATGATGTCCTGAGGAACAGTTGG	CAATCCAGTAGAAAGTTCAGAAT
ABL2	1	1179086420	ACGTTGGATGCTGAACATCGTAAGAGAAGC	ACGTTGGATGGGGTAAGGTTGAAAGGGAC	TCTGTGTCATTTTATTTTTTTTTT

Supplementary Table 5: Sequences of Primers for Sanger Sequencing

Primer ID	Sequence
EZH2-F	CATGGCAAAGTGACCCATCA
EZH2-R	ACTCCCTTTTCAGTCCTGTGA
KRAS-F	GGCCTGCTGAAAATGACTGA
KRAS-R	TGTATCAAAGAATGGTCCTGCAC
MSLN-F	GACACCCTAGACACCCTGAC
MSLN-R	CCTTGGGATAGAGGACGTCC
NTRK3-F	TCTCAGAGAGCAATGGGAGA
NTRK3-R	TGCAGTTCTGAGAAGGCTACA

Supplementary Table 6: Sequences of Primers for Pyrosequencing

Primer ID	Sequence
EZH2-F	AAACAGCTCTTCGCCAGTCT
EZH2-R (Biotinylated)	CAGTGTGTCTCTTTGCAGTTATGA
EZH2-S	CTCTTGGCAAAAATACC
KRAS-F	TATTCGTCCACAAAATGATTCTGA
KRAS-R (Biotinylated)	TATAAGGCCTGCTGAAAATGACT
KRAS-S	CTCTTGCCTACGCCA
MSLN-F	AACTCTGCCCGGCAAGGTG
MSLN-R (Biotinylated)	CCCAGGAAGGACTGGATCTTCA
MSLN-S	CCAGGGCGGTCAGGC
NTRK3-F (Biotinylated)	AAATTTGGACCGTCGACCATA
NTRK3-R	CTTGCTGCTTTTGCCTGTGT
NTRK3-S	CCTGTTGGTGGTTCTC