Supplementary Data: MinimuMM-seq: Genome sequencing of circulating tumor cells for minimally invasive molecular characterization of multiple myeloma pathology

Authors

Ankit K. Dutta[1,2,3,8], Jean-Baptiste Alberge[1,2,3,8], Elizabeth D. Lightbody[1,2,3], Cody J. Boehner[1,2,3], Andrew Dunford[3], Romanos Sklavenitis-Pistofidis[1,2,3], Tarek H. Mouhieddine[1,2,3], Annie N. Cowan[1,2], Nang Kham Su[1,2,3], Erica M. Horowitz[1,2], Hadley Barr[1,2], Laura Hevenor[1,2], Jenna B. Beckwith[1,2], Jacqueline Perry[1,2], Amanda Cao[1,2], Ziao Lin[3], Frank K. Kuhr[4], Richard G. Del Mastro[4], Omar Nadeem[1,2], Patricia T. Greipp[5], Chip Stewart[3], Daniel Auclair[6], Gad Getz[3,7,*], Irene M. Ghobrial[1,2,3*]

Affiliations

[1]Center for Prevention of Progression of Blood Cancers, Dana-Farber Cancer Institute, Boston, MA, 02215, USA

[2]Department of Medical Oncology, Harvard Medical School, Boston, MA, 02215, USA

[3]Cancer Program, Broad Institute of MIT and Harvard, Cambridge, MA, 02142, USA

[4]Menarini Silicon Biosystems, Huntingdon Valley, PA, 19006, USA

[5]Department of Laboratory Medicine and Pathology, Mayo Clinic Comprehensive Cancer Center, Rochester, MN, 55905, USA

[6]Multiple Myeloma Research Foundation, Norwalk, CT, 06851, USA

[7]Cancer Center and Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, 02114, USA

Author List Footnotes

[8] These authors contributed equally.

*Corresponding authors: irene_ghobrial@dfci.harvard.edu (lead contact) and gadgetz@broadinstitute.org

Corresponding Authors:

- Dr. Irene M. Ghobrial, Dana-Farber Cancer Institute, 450 Brookline Ave., Boston, MA 02215
617.632.4198 tel / 617.582.8608 fax / irene_ghobrial@dfci.harvard.edu
- Dr. Gad Getz, Broad Institute, 75 Ames St., Cambridge, MA 02142
617.714.7471 tel / gadgetz@broadinstitute.org

Table of Contents

Supplementary Materials and Methods

Fish analysis

FISH analysis was performed on BM aspirate cells with fluorescence in situ pretreatment, hybridization and fluorescence microscopy in accordance with laboratory specimen-specific protocols. Fifty plasma cell (PC) nuclei were analyzed per probe set, as available when at least 15 PC were available. Otherwise, analysis was considered insufficient. FISH analysis was performed by two qualified clinical cytogenetic technologists and interpreted by a board-certified (American Board of Medical Genetics and Genomics) clinical cytogeneticist. BM aspirate samples were subjected to one of two MM FISH panels. A limited panel included probes designed to detect high risk multiple myeloma abnormalities; loss of chromosome 17p (TP53 [17p13.1] / D17Z1 [CEN17], Abbott Molecular, Des Plaines, IL), gain or amplifications of chromosome 1q (TP73 [1p36.3] / 1q22 [1q22], laboratory-developed test) and IGH gene rearrangements (break-apart probe (BAP), laboratory-developed test, and a IGH::CCND1 dual color dual fusion (DF) probe set for t(11;14) (Abbott Molecular). If the IGH break-apart probe was abnormal (separation of 5' and 3' IGH probe sets; i.e. 1R1G1F, 1R1F or 1G1F) without evidence of IGH::CCND1 fusion, double fusion (DF) probe sets to identify classic partners t(4;14) (IGH::NSD2 or FGFR3, Abbott Molecular), t(14;16) (IGH::MAF, Abbott Molecular), t(14;20) (IGH::MAFB, laboratory-developed test) were subsequently performed. In addition to the probes comprised on the limited panel, the extended panel also included probes to detect loss of chromosome 13/13q (RB1 [13q14] / LAMP1 [13q34] / D4Z1 [CEN4], Abbott Molecular) and MYC rearrangements (BAP, Abbott Molecular). Reflex analyses for the extended panel included those from the initial panel in addition to a DF probe to identify t(6;14) (IGH::CCND3, laboratory-developed test) in the setting of an IGH rearrangement. If ploidy status could not be determined by flow cytometry, investigation for gains of chromosomes 9, 15 (D9Z1 [CEN9] / D15Z4 [CEN15], Abbott Molecular), 3 and 7 (D3Z1 [CEN3] / D7Z1 [CEN7], Abbott Molecular) was also sought. In addition, a smaller panel including FISH probes to detect classic multiple myeloma progression markers; gain or amplification of chromosome

1q, loss of TP53 at chromosome 17p13 and MYC rearrangements. The level of detection required to identify abnormalities was as follows: a minimum of 3 cells displaying fusion signals in the setting of DF probes, a minimum 5 cells with disrupted or separated signals in the setting of BAP probes and a minimum 5 cells with tetraploidy for tetraploid clones or 10 supporting cells for enumeration probes.

Cytoplasmic immunoglobulin in situ hybridization (cIg-FISH)

Pre-analysis to assess adequacy of PC content of samples prior to cIg-FISH was performed using flow cytometry. Samples with more than 0.1% PC (identified with anti-CD19-PerCP 5.5 (clone SJ25C1, BD Biosciences), anti-CD38-APC (clone REA671, Miltenyi Biotec), anti-CD138-BV421 (clone MI15, BD Biosciences), anti-CD45-BB515 (clone HI30, BD Biosciences), anti-cytoplasmic kappa and lambda) were deemed satisfactory for cIg-FISH analysis. After hybridization of slides and post-hybridization wash steps, slides were washed with PBS and left to air dry. PCs were stained with fluorescein isothiocyanate (FIT-C)-conjugated antibodies directed against the kappa and lambda light chains. Only light-chain positive cells were targeted for scoring during FISH analysis. Samples processed before 2020 underwent cIg-FISH-based PC enrichment

Fluorescence activated cell sorting (FACS-FISH)

BM cells (approximately 20 x 10^6) were lysed in ACK lysis buffer for 5 minutes, followed by PBS wash x2 (lyse-wash procedure). The cell pellet was re-suspended in 3% BSA/PBS. Next, 10x10^6 cells were incubated for 15 minutes with the following antibodies: anti-CD19-PerCP 5.5 (clone SJ25C1, BD Biosciences), anti-CD38-APC (clone REA671, Miltenyi Biotec), anti-CD45-BB515 (clone HI30, BD Biosciences), anti-CD56-PE-Cy7 (clone NCAM16.2, BD Biosciences), anti-CD138-BV421 (clone MI15, BD Biosciences), and anti-CD319-PE (clone REA150, Miltenyi Biotec). The specimen was centrifuged and re-suspended in 1.5 mL of PBS. Sorting was performed on BD FACSMelody cell sorter (BD

Biosciences, San Jose, CA). Sorting streams were defined for each case separately, using gates to include CD138-positive, CD319-positive, CD38-bright, CD56-positive and/or CD45-negative plasma cells, and separate them from normal plasma cells. A purity of at least 95% was achieved and verified by Kaluza software (Beckman Coulter Life Sciences, Indianapolis, IN). A minimum of 1000 sorted PCs collected in methanol/acetic acid was required to carry out FISH analysis. The sorted specimen was then processed for FISH analysis. Samples processed between 2020 and 2022 underwent PC enrichment via FACS (FACS-FISH).

DNA library construction for preliminary Ultra Low-Pass Whole-Genome Sequencing

Minipools of CTCs that were sorted by DEPArray underwent whole genome amplification using the Ampli1 kit (Menarini Silicon Biosystems), followed by PCR-free library preparation with unique dual indices (KAPA HyperPrep kit and KAPA Unique Dual Indexed Adapter kit), library quantification and ultra low pass whole genome sequencing (ULP-WGS) on RapidRun flowcell of HiSeq2500 (Illumina). ULP-WGS was used for molecular assessment to detect hyperdiploidy and copy changes as genomic biomarker events of MM disease, with ichorCNA (Adalsteinsson et al., 2017) analyses performed to determine copy number variant (CNV) events and infer tumor fraction.

Genomic sequence alignment and processing

Sequencing reads were aligned to the hg19 reference genome with the bwa mem v0.7.7 algorithm (Li, 2013) and the -M option. Duplicates were marked with the MarkDuplicates function from picard tools v1.475. BAM files were then processed for indel realignment with the RealignerTargetCreator (parameters -dcov 250 -nt 1 -L 9) and IndelRealigner functions and for base calling quality recalibration with the BaseRecalibrator function of GATK 3.4 and with the Broad institute's b37 bundle reference dbSNP 138, known indels, and variantEvalGoldStandard. Samples were checked for absence of contamination and sample mismatch with the CrossCheckLaneFingerprints function from picard v1.475 and with the ContEst tool (Cibulskis et al., 2011).

Panel of normals

Paired germline sequencing data were used as a panel of normals (PoN) to normalize and control for artifacts and variability of unknown source in copy number profiling and false-positive mutations. The copy number PoN was generated with the ReCapSeg algorithm (Lichtenstein et al., 2016). The token file for point mutations was generated with the CGA_Token_PoN_Maker v0.1 Firecloud task (available here https://portal.firecloud.org/?return=terra#methods/getzlab/CGA_Token_PoN_Maker_v0.1_Jan2019/2 with a free account) and run on Terra.

Copy number analysis

Copy number profiling was performed with the AllelicCapSeg algorithm (Landau et al., 2013). AllelicCapSeg uses CNV and SNP haplotyping to infer allelic local copy number. The GATK CNV task was used to calculate segment copy number ratio between tumor and normal. CNV values within [-0.1, 0.1] in the log2 ratio space were considered normal. Results were normalized with the PoN built from matched-germline samples of this cohort processed and sequenced with the same protocol. The heterozygotes sites were obtained from MuTect1 call_stats results, and used as an input of the AllelicCapSeg algorithm. Copy number estimate and minor allele frequencies were reported per tumor-normal pair and used in mutation calling step to estimate cancer cell fraction (CCF) of mutations with ABSOLUTE (Carter et al., 2012).

Structural variant detection

Three algorithms were used to detect and filter structural variants (SVs) genome-wide similar to Morton, Karyadi, Stewart and colleagues (Morton et al., 2021). Briefly, Manta (Chen et al., 2016), dRanger, and SVaBa (Wala et al., 2018) algorithms are executed in parallel on paired tumor-normal BAM files. dRanger was run with the following parameters: tminmapq = 5, minpairs = 2, windowsize = 2000, nminwindow = 2000, minsomratio = 50, nminspanfrac = 0.5, minscoreforbp = 0.01. Manta and SVaBa were run with the default parameters and the following filters were applied to each output: minscoreforbp=0.1,

min_span=200, max_pon=1, max_norm=2, min_tum_SR=0 (1 for Manta), min_tum_RP=1 (0 for Manta), min_tum=4. Specific capture of reads supporting immunoglobulin (IG) translocations restricted to regions of interest was assessed with parameters min_tum=2 and minimum number of split-read and read-pairs both set to 0. Regions of interest were pre-defined as genomic intervals encompassing any IGH translocation from the CoMMpass study release IA15 and found in 1% of participants. Next, BreakPointer (Drier et al., 2013) was used to aggregate results from the SV detection algorithms and to score structural variants after local assembly with the Smith-Waterman algorithm. SVs were all manually reviewed in IGV.

Mutation calling

Mutations and short indels were detected with MuTect1 and Strelka2 respectively. MuTect1 (Cibulskis et al., 2013) was used in matched tumor-normal pairs genome-wide. To estimate the power to detect somatic variants given purity, ploidy, and cancer cell fraction, the MuTect formula was used. Additionally, Strelka2 (Kim et al., 2018) was used to characterize short insertions and deletions (indels). The DeTiN algorithm (Taylor-Weiner et al., 2018) was used to estimate tumor-in-normal contamination and to rescue somatic mutations originally discarded by MuTect1 and Strelka2. Filtering of mutations was done with in-house code and included detection and filtering of oxoG artefacts (Costello et al., 2013) detection of mutation in the panel of normals. Additionally, mutations were inspected with the BLAT algorithm and for each sequencing read supporting a somatic mutation, the alternative alignments suggested by BLAT are examined. Mutations that are only supported by reads which are ambiguously mapped are removed. Finally, SNVs and indels were annotated with GATK's Funcotator v1.6 and used as an input of the ABSOLUTE algorithm (Carter et al., 2012).

Absolute copy number, mutations, and phylogeny trees

ABSOLUTE (Carter et al., 2012) was used to estimate purity, ploidy, and subclonal composition of mutations and copy number abnormalities. ABSOLUTE solutions were all reviewed manually and chosen based on optimal fit of subclonal SNV multiplicity and fraction of alternate reads and when available, BM

samples were cross-validated with the fraction of cells bearing arm-level abnormalities according to matched FISH reports. For participants with matched BMPCs and CTCs available, and for the participant with serial sampling over time, the union of mutations between both compartments (or between both timepoints) was additionally force-called with the forcecaller task (available with a free Terra account at the following location https://portal.firecloud.org/?return=terra#methods/danielr/forcecall_snps_and_indels/7). Cancer cell fraction from the union of mutations was then calculated again with ABSOLUTE. Phylogeny trees between both compartments (or between timepoints) were reconstructed with the PhylogicNDT algorithm (Leshchiner et al., 2019). PhylogicNDT was run with the following parameters: minimum cancer cell fraction: 20%, minimum coverage: 10, number of iterations: 1,000.

Mutational signature analysis

Mutational processes were weighted with the ARD-NMF decomposition provided in the SignatureAnalyzer method (Kasar et al., 2015; Kim et al., 2016; Taylor-Weiner et al., 2019). SignatureAnalyzer was run with the following parameters: reference=pcawg_COMPOSITE, objective=poisson, n=100. The PCAWG composite reference dataset is used to annotate single-base substitutions in their pentanucleotide neighborhood (SBS with 1536 context possibilities), indels, and double-based substitutions, and assigned to most likely reference signature based on cosine similarity metrics. Stability of NMF decomposition was assessed in this study with aggregation of signatures weights across all ARD-NMF runs. For matched BMPCs and CTCs sequencing data, bootstrapping (N=1000) was used to estimate mean and 95% confidence intervals of the mutations cosine similarity between assays.

BCR alignment

BCR sequences were reconstructed with the mixcr (Bolotin et al., 2017; Bolotin et al., 2015) set of algorithms in “shotgun analyze” mode with default parameters for DNA BCR sequence reconstruction and with the "--only-productive" flag. Input regions included sequence mapping to reference immunoglobulin

heavy chain loci hg19 coordinates. BCR hits were then compared between BM and PB and allele frequencies in both compartments were systematically reported.

References to supplementary materials and methods

Adalsteinsson, V. A., Ha, G., Freeman, S. S., Choudhury, A. D., Stover, D. G., Parsons, H. A., Gydush, G., Reed, S. C., Rotem, D., Rhoades, J., Loginov, D., Livitz, D., Rosebrock, D., Leshchiner, I., Kim, J., Stewart, C., Rosenberg, M., Francis, J. M., Zhang, C. Z., . . . Meyerson, M. (2017). Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat Commun*, *8*(1), 1324. https://doi.org/10.1038/s41467-017-00965-y

Bolotin, D. A., Poslavsky, S., Davydov, A. N., Frenkel, F. E., Fanchi, L., Zolotareva, O. I., Hemmers, S., Putintseva, E. V., Obraztsova, A. S., Shugay, M., Ataullakhanov, R. I., Rudensky, A. Y., Schumacher, T. N., & Chudakov, D. M. (2017). Antigen receptor repertoire profiling from RNA-seq data. *Nat Biotechnol*, *35*(10), 908-911. https://doi.org/10.1038/nbt.3979

Bolotin, D. A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I. Z., Putintseva, E. V., & Chudakov, D. M. (2015). MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*, *12*(5), 380-381. https://doi.org/10.1038/nmeth.3364

Carter, S. L., Cibulskis, K., Helman, E., McKenna, A., Shen, H., Zack, T., Laird, P. W., Onofrio, R. C., Winckler, W., Weir, B. A., Beroukhim, R., Pellman, D., Levine, D. A., Lander, E. S., Meyerson, M., & Getz, G. (2012). Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol*, *30*(5), 413-421. https://doi.org/10.1038/nbt.2203

Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A. J., Kruglyak, S., & Saunders, C. T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*, *32*(8), 1220-1222. https://doi.org/10.1093/bioinformatics/btv710

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S., & Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*, *31*(3), 213-219. https://doi.org/10.1038/nbt.2514

Cibulskis, K., McKenna, A., Fennell, T., Banks, E., DePristo, M., & Getz, G. (2011). ContEst: estimating cross-contamination of human samples in next-generation sequencing data. *Bioinformatics*, *27*(18), 2601-2602. https://doi.org/10.1093/bioinformatics/btr446

Costello, M., Pugh, T. J., Fennell, T. J., Stewart, C., Lichtenstein, L., Meldrim, J. C., Fostel, J. L., Friedrich, D. C., Perrin, D., Dionne, D., Kim, S., Gabriel, S. B., Lander, E. S., Fisher, S., & Getz, G. (2013). Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Research*, *41*(6), e67-e67. https://doi.org/10.1093/nar/gks1443

Drier, Y., Lawrence, M. S., Carter, S. L., Stewart, C., Gabriel, S. B., Lander, E. S., Meyerson, M., Beroukhim, R., & Getz, G. (2013). Somatic rearrangements across cancer reveal classes of samples with distinct patterns of DNA breakage and rearrangement-induced hypermutability. *Genome Res*, *23*(2), 228-235. https://doi.org/10.1101/gr.141382.112

Kasar, S., Kim, J., Improgo, R., Tiao, G., Polak, P., Haradhvala, N., Lawrence, M. S., Kiezun, A., Fernandes, S. M., Bahl, S., Sougnez, C., Gabriel, S., Lander, E. S., Kim, H. T., Getz, G., & Brown, J. R. (2015). Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun*, *6*, 8866. https://doi.org/10.1038/ncomms9866

Kim, J., Mouw, K. W., Polak, P., Braunstein, L. Z., Kamburov, A., Kwiatkowski, D. J., Rosenberg, J. E., Van Allen, E. M., D'Andrea, A., & Getz, G. (2016). Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. *Nat Genet*, *48*(6), 600-606. https://doi.org/10.1038/ng.3557

Kim, S., Scheffler, K., Halpern, A. L., Bekritsky, M. A., Noh, E., Kallberg, M., Chen, X., Kim, Y., Beyter, D., Krusche, P., & Saunders, C. T. (2018). Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods*, *15*(8), 591-594. https://doi.org/10.1038/s41592-018-0051-x

Landau, D. A., Carter, S. L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M. S., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., Wan, Y., Zhang, W., Shukla, S. A., Vartanov, A., Fernandes, S. M., Saksena, G., Cibulskis, K., Tesar, B., Gabriel, S., . . . Wu, C. J. (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*, *152*(4), 714-726. https://doi.org/10.1016/j.cell.2013.01.019

Leshchiner, I., Livitz, D., Gainor, J. F., Rosebrock, D., Spiro, O., Martinez, A., Mroz, E., Lin, J. J., Stewart, C., Kim, J., Elagina, L., Bozic, I., Mino-Kenudson, M., Rooney, M., Ou, S.-H. I., Wu, C. J., Rocco, J. W., Engelman, J. A., Shaw, A. T., & Getz, G. (2019). Comprehensive analysis of tumour initiation, spatial and temporal progression under multiple lines of treatment. *bioRxiv*, 508127. https://doi.org/10.1101/508127

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Lichtenstein, L., Woolf, B., MacBeth, A., Birsoy, O., & Lennon, N. (2016). ReCapSeg: validation of somatic copy number alterations for CLIA whole exome sequencing. In: AACR.

Morton, L. M., Karyadi, D. M., Stewart, C., Bogdanova, T. I., Dawson, E. T., Steinberg, M. K., Dai, J., Hartley, S. W., Schonfeld, S. J., Sampson, J. N., Maruvka, Y. E., Kapoor, V., Ramsden, D. A., Carvajal-Garcia, J., Perou, C. M., Parker, J. S., Krznaric, M., Yeager, M., Boland, J. F., . . . Chanock, S. J. (2021). Radiation-related genomic profile of papillary thyroid carcinoma after the Chernobyl accident. *Science*, *372*(6543). https://doi.org/10.1126/science.abg2538

Taylor-Weiner, A., Aguet, F., Haradhvala, N. J., Gosai, S., Anand, S., Kim, J., Ardlie, K., Van Allen, E. M., & Getz, G. (2019). Scaling computational genomics to millions of individuals with GPUs. *Genome Biol*, *20*(1), 228. https://doi.org/10.1186/s13059-019-1836-7

Taylor-Weiner, A., Stewart, C., Giordano, T., Miller, M., Rosenberg, M., Macbeth, A., Lennon, N., Rheinbay, E., Landau, D. A., Wu, C. J., & Getz, G. (2018). DeTiN: overcoming tumor-in-normal contamination. *Nat Methods*, *15*(7), 531-534. https://doi.org/10.1038/s41592-018-0036-9

Wala, J. A., Bandopadhayay, P., Greenwald, N. F., O'Rourke, R., Sharpe, T., Stewart, C., Schumacher, S., Li, Y., Weischenfeldt, J., Yao, X., Nusbaum, C., Campbell, P., Getz, G., Meyerson, M., Zhang, C. Z., Imielinski, M., & Beroukhim, R. (2018). SvABA: genome-wide detection of structural variants and indels by local assembly. *Genome Res*, *28*(4), 581-591. https://doi.org/10.1101/gr.221028.117

Supplementary Figures

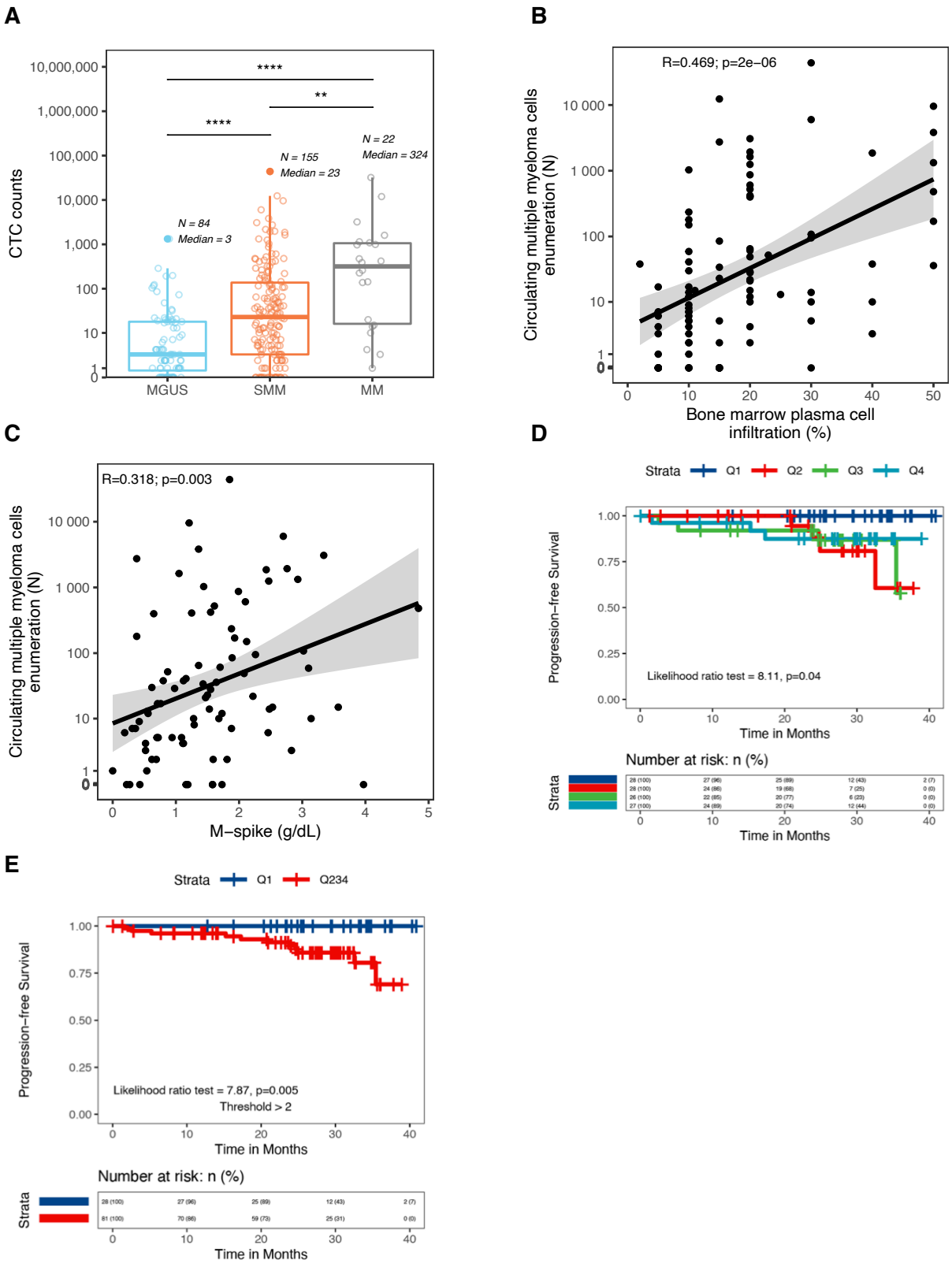

Figure S1

(legend on next page)

Figure S1

Correlation between clinical measures of disease pathology, survival, and circulating tumor cells enumeration. A. Boxplot of CTCs enumerated between disease precursor stages MGUS and SMM, and participants with overt disease (MM). **B-C**. Comparison and correlation of enumeration results from circulating tumor cells (CTCs) of multiple myeloma precursor patients with (B) plasma cells involvement in the bone marrow given as BMPC percentage (N=92 with successful count), and (C) M-spike protein concentration (N=85 with successful quantification). **D-E**. Kaplan-Meier curves depicting progression-free survival for smoldering multiple myeloma patients (N=109) based on CTC count quartiles with overall p-value (D) and quartiles Q1 vs Q2, Q3, and Q4 (E).

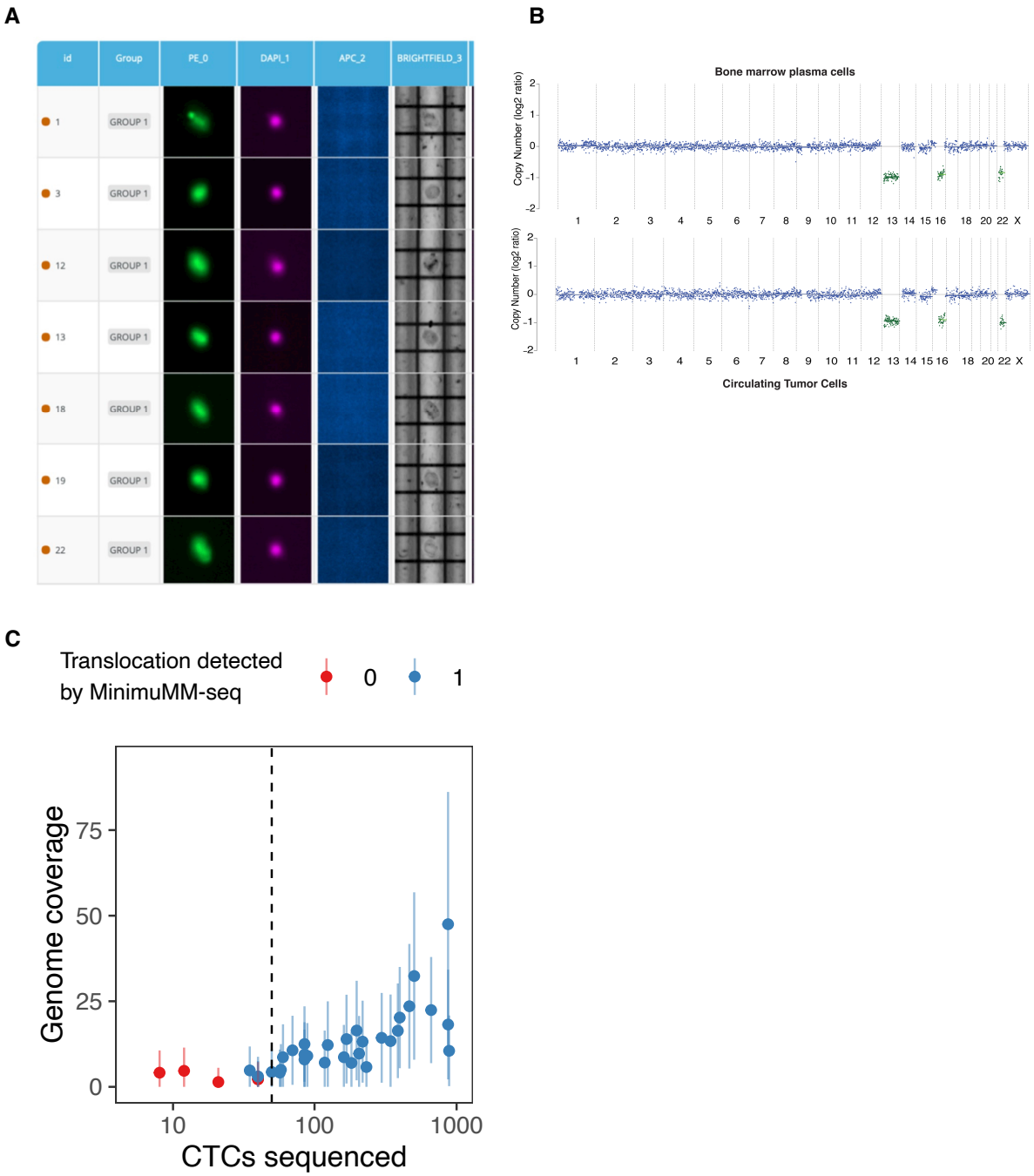

Figure S2

(legend on next page)

Figure S2: Isolation of pure circulating tumor cells from peripheral blood of precursor disease patients. A. Enriched CTCs are intact cells with immunophenotype of PCs (138+38+45-). **B.** Copy number profiling of MM cells by ultra-low pass whole-genome sequencing in the bone marrow (top) and in matched circulating tumor cells (bottom) illustrates tumor origin and genomic abnormalities. In this example SMM patient concordance of deletion 13, 16 and 22 is observed in both samples. **C.** Detection of IGH translocations per number of cells sequenced and sequence genome coverage. Dashed line at 50 CTCs. Genome coverage is indicated as mean (X) with 95% central interval.

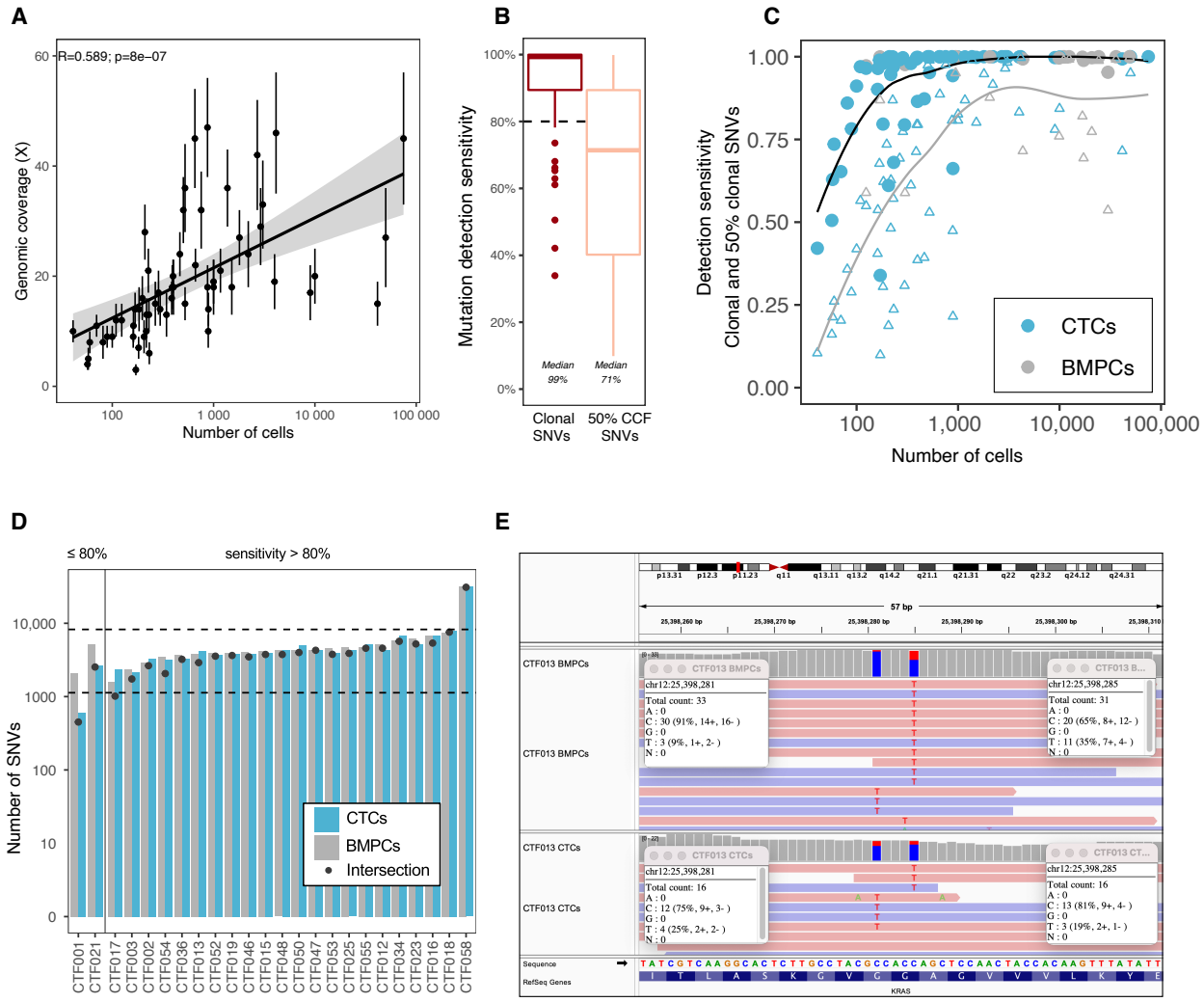

Figure S3

CTCs sequencing enables cohort-level genomic characterization of tumor in MM precursor stages A. Effective sequence coverage calculated post-alignment with 95% intervals by number of cells captured. **B.** Estimated power to detect single-nucleotide variants (SNVs) given actual sequence coverage and tumor purity achieved in this study in the context of a clonal mutation (left, cancer cell fraction CCF 100%) and of a subclonal mutation (right, CCF=50% shown). **C.** Power of SNV detection by number of cells sequenced given purity, ploidy, and genomic coverage in scenarii of a 100% cancer cell fraction (CCF), in plain circles, and of a 50% CCF in triangles. **D.** For patients underpowered (left, N=2), and powered at 80% to detect SNVs (right, N=15/17), bar plot representation of total number of mutations detected in each compartment (BMPCs and CTCs) by our assay with CCF estimate >33%, and intersection between both compartments (black dots). Dashed lines represent reference range obtained from the Oben et al. study. **E.** Double-hit hotspot mutation at KRAS p.G12 and p.G13 in matched bone marrow and peripheral blood from participant CTF013.

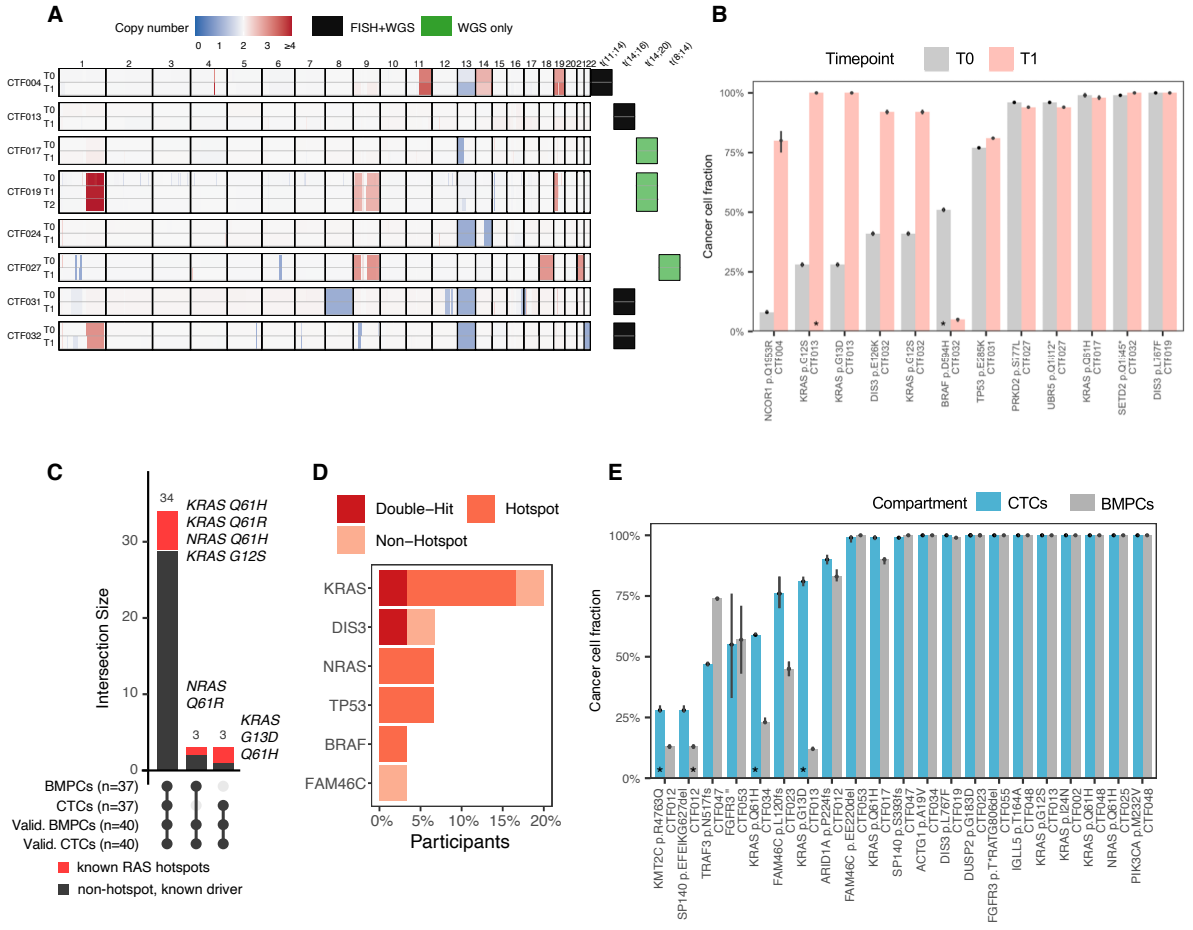

Figure S4

Longitudinal and tissue-matched genomic characterization of driver mutations. A. Genome wide copy number abnormalities heatmap (left), and translocation discovery (right), with comparison to clinical FISH reports. Each row is split into a top panel dedicated to CTCs profiling at initial screening date (T0), and subsequent follow-up screenings below (T1, T2). **B.** Cancer cell fraction of non-silent mutations in myeloma driver genes at T0 (grey) and T1 (pink). Asterisks represent mutations discovered only at the highlighted timepoint, but CCF is given in validation (force-calling) mode. **C.** Categorical classification of mutations in known commonly mutated genes detected in this cohort. Top BMPC and CTC rows represent de novo detection of mutation, while bottom validated (Valid.) BMPC and CTC rows represent cross-compartment validation of mutation by force-calling. Known hotspots (G12, G13, Q61 from RAS mutants) are shown in red. **D.** Fraction of participants with 6 common myeloma driver genes with one or more mutations detected in a known hotspot or outside of a known hotspot. **E.** Cancer cell fraction of non-silent mutations in known commonly mutated genes of myeloma between matched CTCs (blue) and BMPCs (grey). Similar to panel B, asterisks represent mutations discovered only at the highlighted timepoint, but CCF is given in validation (force-calling) mode. Quotation marks represent splice site variants.

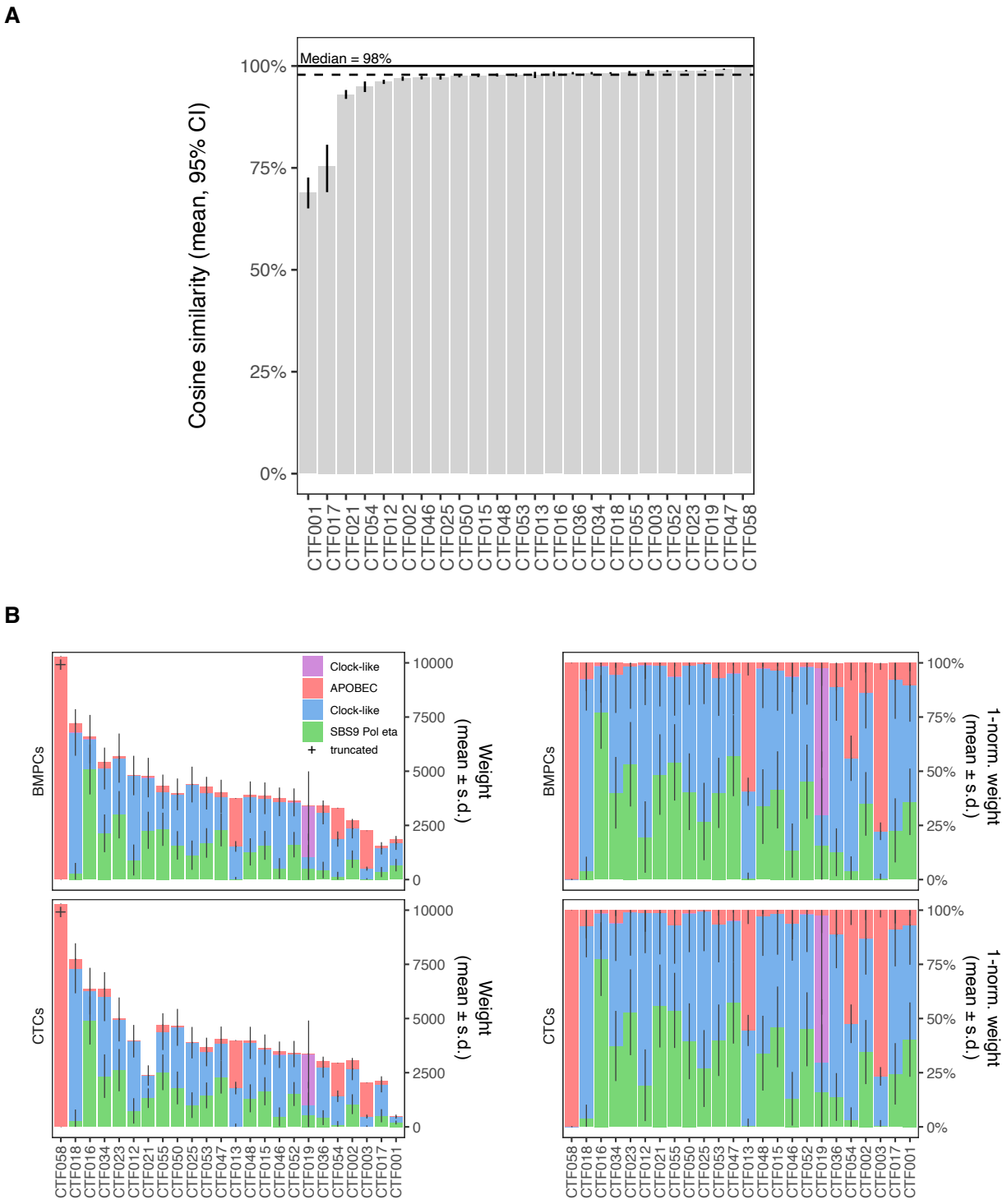

Figure S5

(legend on next page)

Figure S5. Comparison of mutational processes between BMPCs and CTCs assigned to most likely PCAWG composite reference signature. A. Cosine similarity between BMPCs and CMMCs for each participant with matched samples using raw mutational data and bootstrapping, Mean and 95% confidence intervals are shown. **B.** Bar graph representing signature weight (left) and normalized weight (right) in BMPCs and CTCs, per each matched sample. Mean is aggregated from all NMF runs. Plus (+) sign symbolizes truncation for CTF058 (APOBEC Weight, CTCs:30,828 ± 71, BMPCs: 31,455 ± 91).

Supplementary Tables

Table S1

Patient ID	CTF ID	Age	Gender	Stage	IgH and Light Chain (LC)	Patient Sample
1	CTF001	41	F	SMM	IgG Kappa	Matched
2	CTF003	51	F	SMM	IgG Kappa	Matched
3	CTF002	50	F	SMM	IgG Lambda	Matched
4	CTF004	72	M	SMM	IgG Lambda	PB
5	CTF005	56	F	SMM	IgA Lambda	PB
6	CTF010	51	M	SMM	IgG Kappa	PB
7	CTF011	75	M	SMM	Biclonal IgG Kappa/IgG Lambda	PB
8	CTF012	65	F	SMM	IgG Kappa	Matched
9	CTF013	61	F	SMM	IgG Lambda	Matched
10	CTF015	78	F	SMM	IgG Kappa	Matched
11	CTF016	76	F	SMM	Lambda LC	Matched
12	CTF017	50	F	SMM	IgG Lambda	Matched
13	CTF018	78	F	SMM	IgG Kappa	Matched
14	CTF019	61	M	MM	IgG Kappa	Matched
15	CTF021	63	M	MM	IgG Kappa	Matched
16	CTF023	84	M	SMM	IgG Kappa	Matched
17	CTF022	72	M	MM	IgG Kappa	PB
18	CTF024	73	F	SMM	IgA Lambda	PB
19	CTF025	77	F	SMM	IgG Kappa	Matched
20	CTF026	70	M	MGUS	IgA Kappa	PB
21	CTF027	65	F	SMM	Lambda LC	PB
22	CTF028	57	M	MGUS	IgG Kappa	PB
23	CTF029	81	F	SMM	Biclonal IgA/IgG	PB
24	CTF030	66	F	SMM	IgA Kappa	PB
25	CTF031	64	M	SMM	IgG Kappa	PB
26	CTF032	71	F	SMM	IgG Lambda	PB
27	CTF033	64	F	SMM	IgG Kappa	PB
28	CTF034	67	M	MM	IgG Lambda	Matched
29	CTF035	77	M	SMM	IgG Kappa	PB
30	CTF036	54	M	SMM	IgG Lambda	Matched
31	CTF038	62	M	SMM	IgG Kappa	PB
32	CTF039	64	M	SMM	IgG Lambda	PB
33	CTF040	62	M	SMM	IgG Lambda	PB
34	CTF041	58	F	SMM	IgG Kappa	PB
35	CTF042	68	F	SMM	IgG Kappa	PB
36	CTF043	75	M	SMM	Kappa LC	PB
37	CTF044	67	F	MGUS	IgG Lambda	PB
38	CTF045	37	F	SMM	IgG Kappa	PB
39	CTF046	48	M	SMM	IgG Kappa	Matched
40	CTF047	56	M	MM	IgA Lambda LC	Matched
41	CTF048	68	F	MGUS	IgG Lambda	Matched
42	CTF049	46	F	MM	IgG Lambda	PB
43	CTF050	74	M	SMM	Lambda LC	Matched
44	CTF051	51	F	SMM	IgG Kappa	PB
45	CTF052	76	F	SMM	IgG Kappa	Matched
46	CTF053	71	M	MM	IgA Kappa	Matched
47	CTF054	53	M	MGUS	IgG Kappa	Matched
48	CTF055	60	M	MM	IgG Lambda	Matched
49	CTF056	47	F	SMM	IgA Kappa	PB
50	CTF057	68	M	MM	IgG Kappa	PB
51	CTF058	54	F	MM	IgG Kappa	Matched

Clinical characteristics and sampling of participants in this study. PB: Peripheral Blood.

Table S2

Characteristics	**Mean (±SD)**	**Range**
Insert size mode	314 bp (±37bp)	197 to 365
Coverage	19.1 (±11.1)	3 to 47
% Genome coverage at …		
5X	92.4% (±13.7%)	10% to 98.6%
10X	76.2% (±27.4%)	1% to 98.3%
15X	55.4% (±34.8%)	0% to 97.6%
30X	18.4% (±28.3%)	0% to 89.9%
Heterozygous SNP Q-score	14.1 (±3.47)	3 to 19
% Bases excluded due to …		
Low Mapping Quality	7.3% (±1.3%)	5% to 11.6%
Duplicate Reads	44.7% (±23.9%)	0% to 80.1%
Unpaired Reads	0.04% (±0.04%)	0% to 0.16%
Low Base Quality	1.1% (±0.8%)	0% to 3.4%
Overlapping Reads	5.0% (±4.5%)	1% to 17.9%
Above Coverage Cap	0.8% (±0.4%)	0% to 1.7%
Total % Bases excluded	58.9% (±19.4%)	10% to 88.9%

WGS coverage and library metrics. Coverage and SNP sensitivity are calculated as effective post-exclusion of low-mapping reads, duplicates, overlapping, and low-quality reads**.**

ID	CTF ID	Stage	IgH and Light Chain	BM Biopsy PCs%	BMPCs	CTCs	Translocation	Hyperdiploid	1q dup	1p del	13q del	16q del	17p del	*Notes*
1	CTF001	SMM	IgG Kappa	20%	170	41	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Insufficient number of plasma cells
2	CTF003	SMM	IgG Kappa	15%	300	100	t(14;16), 90%	n/a	n/a	n/a	75%	n/a	n/a	
3	CTF002	SMM	IgG Lambda	20%	124	170	t(14;16), 70%	n/a	n/a	n/a	n/a	n/a	n/a	
8	CTF012	SMM	IgG Kappa	20%	8300	210	n/a	Trisomy 9, 11, 15	40%	n/a	95%	n/a	n/a	
9	CTF013	SMM	IgG Lambda	10%	30608	398	t(14;16), 35%	n/a	n/a	n/a	n/a	n/a	n/a	
10	CTF015	SMM	IgG Kappa	30%	8300	4135	t(4;14), 95%	n/a	n/a	n/a	n/a	n/a	n/a	
11	CTF016	SMM	Lambda LC	20%	8300	398	t(11;14), 80%	n/a	n/a	n/a	n/a	n/a	n/a	
12	CTF017	SMM	IgG Lambda	15%	4500	886	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Insufficient number of plasma cells
13	CTF018	SMM	IgG Kappa	40%	11474	2911	n/a	n/a	95%	n/a	n/a	n/a	n/a	
14	CTF019	MM	IgG Kappa	50%	947	1169	n/a	Trisomy 9	65%	n/a	n/a	n/a	n/a	
15	CTF021	MM	IgG Kappa	80%	12423	182	14q IGH sep, 60%	Trisomy 3, 9, 11, 15	n/a	n/a	n/a	n/a	n/a	
16	CTF023	SMM	IgG Kappa	20%	23344	343	n/a	Trisomy 3,7,9,11,15(tetra)	n/a	n/a	n/a	n/a	n/a	
19	CTF025	SMM	IgG Kappa	30%	36368	198	n/a	n/a	70%	n/a	n/a	n/a	n/a	
28	CTF034	MM	IgG Lambda	80%	16880	654	n/a	Trisomy 3, 9 , 15	n/a	n/a	n/a	n/a	n/a	
30	CTF036	SMM	IgG Lambda	20%	49395	286	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Insufficient number of plasma cells
39	CTF046	SMM	IgG Kappa	30%	2051	523	n/a	n/a	n/a	n/a	n/a	n/a	n/a	Insufficient number of plasma cells
40	CTF047	MM	IgA Lambda LC	80%	16700	32071	t(14;16), 100%	n/a	n/a	n/a	n/a	n/a	n/a	
41	CTF048	MGUS	IgG Lambda	10%	9906	266	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
43	CTF050	SMM	Lambda LC	20%	4343	227	n/a	n/a	72%	n/a	n/a	n/a	n/a	
45	CTF052	SMM	IgG Kappa	50%	8300	168	t(11;14), 78%	n/a	n/a	n/a	88%	n/a	n/a	
46	CTF053	MM	IgA Kappa	50%	8300	124	t(4;14), 100%	n/a	n/a	n/a	n/a	n/a	n/a	
47	CTF054	MGUS	IgG Kappa	5%	17243	70	t(14;20), 72%	n/a	n/a	n/a	n/a	n/a	n/a	
48	CTF055	MM	IgG Lambda	70%	20783	871	t(4;14), 77%	Trisomy 3, 9 , 15	76%	n/a	n/a	n/a	n/a	
51	CTF058	MM	IgG Kappa	80%	8300	3061	t(14;16), 92%	n/a	86%	n/a	n/a	n/a	n/a	

Table S3

Clinical BM FISH results and cells recovered for cohort with matched samples**.**

Participant	V_segment	D_segment	J_segment	CDR3 Aminoacid sequence	Clone Allele Fraction BMPCs	Clone Allele Fraction CTCs
CTF001	IGHV3-11*00	IGHD3-22*00	IGHJ6*00	CTRGHYYDSSGYSLIKGYNYYYYLDVW	100%	80%
CTF001	IGHV4-39*00	IGHD3-10*00	IGHJ5*00	CARQGVLGFGVSNWFDPW	0%	20%
CTF002	IGHV3-21*00	IGHD5-12*00	IGHJ3*00	CARDLSKLAEAFDIW	71%	80%
CTF002	IGHV3-13*00	IGHD3-22*00	IGHJ4*00	CATRTMIVV_LLPTPPDYW	29%	20%
CTF003	IGHV1-24*00	IGHD6-13*00	IGHJ6*00	CATEISPAIPPLGYGLGVW	42%	57%
CTF003	IGHV3-16*00	IGHD3-22*00	IGHJ4*00	CVRKRVVLL**_SGYYWGFDYW	33%	43%
CTF003	IGHV3-35*00	IGHD3-22*00	IGHJ4*00	CVRKRVVLL**_SGYYWGFDYW	25%	0%
CTF012	IGHV1-46*00	IGHD4-17*00	IGHJ6*00	CARMEASYAPTHNFYGLDVW	94%	100%
CTF012	IGHV4-28*00	IGHD4-17*00	IGHJ6*00	CARMEASYAPTHNFYGLDVW	6%	0%
CTF013	IGHV4-59*00	IGHD3-16*00	IGHJ4*00	CARDQGGPFDHW	87%	100%
CTF013	IGHV3-15*00	IGHD3-10*00	IGHJ4*00	CTLDYFGSGSNYNKYW	7%	0%
CTF013	IGHV4-59*00	IGHD3-16*00	IGHJ5*00	CARDQGGPFDHW	7%	0%
CTF015	IGHV1-3*00	IGHD3-16*00	IGHJ2*00	CATLPDDYGVDYGYWYFDLW	95%	94%
CTF015	IGHV1-67*00	IGHD3-16*00	IGHJ2*00	CATLPDDYGVDYGYWYFDLW	5%	0%
CTF015	IGHV1-46*00	IGHD3-16*00	IGHJ2*00	CATLPDDYGVDYGYWYFDLW	0%	6%
CTF017	IGHV3-13*00	IGHD2-2*00	IGHJ5*00	CARGL***QL_MRCNWFDPW	33%	100%
CTF017	IGHV1-18*00	IGHD3-3*00	IGHJ5*00	CARGVRITIFGVAGGWTSPEDGEKDGLDPW	11%	0%
CTF017	IGHV4-39*00	IGHD3-22*00	IGHJ4*00	CARHRLATYYYESSGYYFDYW	11%	0%
CTF017	IGHV4-34*00	IGHD6-19*00	IGHJ5*00	CAGWGRTLPFLSNWFDPW	11%	0%
CTF017	IGHV1-18*00	IGHD5-24*00	IGHJ4*00	CARDWDMATIRGGGDYW	11%	0%
CTF017	IGHV3-15*00	IGHD2-8*00	IGHJ4*00	CTTQIYCTNGVCADYW	11%	0%
CTF017	IGHV4-39*00	IGHD6-19*00	IGHJ4*00	CVLPLVGTVYVGYW	11%	0%
CTF018	IGHV4-31*00	IGHD1-7*00	IGHJ5*00	CARDWNYGTNSFWFDPW	100%	95%
CTF018	IGHV6-1*00	IGHD1-7*00	IGHJ5*00	CARDWNYGTNSFWFDPW	0%	5%
CTF019	IGHV2-70*00	IGHD3-10*00	IGHJ4*00	CARIRDYYASGAHDFW	100%	100%
CTF021	IGHV3-21*00	IGHD3-3*00	IGHJ6*00	CARFGGRDFWSGYSGYYHYGMDVW	100%	100%
CTF023	IGHV2-5*00	IGHD3-10*00	IGHJ4*00	CVHRQGRTLLRGAMSPYFDFW	100%	0%
CTF025	IGHV3-43*00	IGHD4-23*00	IGHJ1*00	CVKGDYGRNPGHFEYW	100%	100%
CTF034	IGHV3-33*00	IGHD4-17*00	IGHJ4*00	CARDCDS**H_GVTTL*FDYW	67%	72%
CTF034	IGHV3-48*00	IGHD3-9*00	IGHJ3*00	CARILSDFGDHDRRDAFDVW	33%	22%
CTF034	IGHV3-21*00	IGHD3-9*00	IGHJ3*00	CARILSDFGDHDRRDAFDVW	0%	6%
CTF036	IGHV1-3*00	IGHD6-19*00	IGHJ4*00	CASEIVGWAFDYW	90%	83%
CTF036	IGHV1-67*00	IGHD6-19*00	IGHJ4*00	CASEIVGWAFDYW	10%	0%
CTF036	IGHV1-46*00	IGHD6-19*00	IGHJ4*00	CASEIVGWAFDYW	0%	17%
CTF046	IGHV1-24*00	IGHD3-16*00	IGHJ4*00	CATSALGQVDNW	100%	67%
CTF046	IGHV3-53*00	IGHD6-13*00	IGHJ5*00	CARSYSSSLRGDWFDPW	0%	33%
CTF047	IGHV3-49*00	IGHD5-12*00	IGHJ4*00	CSRDGLGIVATGDVDSGGLDRW	100%	100%
CTF048	IGHV3-9*00	IGHD3-22*00	IGHJ6*00	CVKDLSNGYYSLDANHYFGMDVW	100%	100%
CTF052	IGHV1-69*00	IGHD2-21*00	IGHJ6*00	CVRGEDEVTAIDYYYFGMDVW	100%	0%
CTF053	IGHV1-18*00	IGHD4-17*00	IGHJ4*00	CAREGDDYDDYNYLDYW	100%	100%
CTF054	IGHV5-51*00	IGHD3-10*00	IGHJ4*00	CARRFGGSYFDYW	100%	100%
CTF055	IGHV3-30*00	IGHD6-13*00	IGHJ4*00	CAKGIYSSSFTRARDSW	100%	100%
CTF058	IGHV4-31*00	IGHD6-13*00	IGHJ5*00	CARGLSWEVPAAAWFDPW	100%	100%

Table S4

(legend on next page)

Table S4: Comparison of BCR sequence with VDJ and CDR3 exact match between BMPCs and CTCs obtained by mixcr algorithm. Rows are colored by detection in both CTCs and BMPCs (green). Red rows depict major clone found in BMPCs but not reconstructed in CTC. Allele fraction is given as a percentage of all BCR detected. No IGH BCR was reconstructed for matched patient CTF016 in either CTCs or BMPCs.

ID	CTF ID	Stage	IgH and Light Chain	BM Biopsy PCs%	CMMCs	Translocation	Hyperdiploid	1q dup	1p del	13q del	16q del	17p del	*Notes*
4	CTF004	SMM	IgG Lambda	10%	60	t(11;14), 80%	n/a	n/a	n/a	n/a	n/a	n/a	
5	CTF005	SMM	IgA Lambda	10%	57	t(14;16)	n/a	n/a	n/a	n/a	n/a	n/a	
6	CTF010	SMM	IgG Kappa	20%	402	14q IGH sep, 50%	n/a	n/a	n/a	85%	n/a	n/a	Trisomy 6p (75%)
7	CTF011	SMM	Biclonal IgG Kappa/IgG Lambda	15%	297	n/a	n/a	✓	n/a	85%	n/a	n/a	
19	CTF022	SMM	IgG Kappa	20%	206	n/a	Trisomy 5, 9, 15	n/a	n/a	✓	n/a	n/a	
20	CTF024	SMM	IgA Lambda	13%	217	n/a	n/a	n/a	n/a	✓	n/a	n/a	
22	CTF026	MGUS	IgA Kappa	10%	81	n/a	n/a	n/a	n/a	n/a	n/a	n/a	No BM FISH ever performed
23	CTF027	SMM	Lambda LC	10%	4979	n/a	n/a	✓	n/a	n/a	n/a	n/a	
24	CTF028	MGUS	IgG Kappa	10%	162	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
25	CTF029	SMM	Biclonal IgA/IgG	20%	58	t(11;14), 100%	n/a	n/a	n/a	n/a	n/a	n/a	
26	CTF030	SMM	Lambda high	20%	503	t(4;14)	Trisomy 3	✓	n/a	✓	n/a	n/a	
27	CTF031	SMM	IgG Kappa	30%	41500	t(14;16), 85%	n/a	n/a	n/a	85%	n/a	70%	8q del, 95%
28	CTF032	SMM	IgG Lambda	50%	9372	t(14;16), 70%	n/a	75%	n/a	90%	n/a	n/a	
29	CTF033	SMM	IgG Kappa	20%	755	14q IGH sep, 40%	Trisomy 3, 7, 9, 11, 15	n/a	n/a	n/a	n/a	n/a	
31	CTF035	SMM	IgG Kappa	70%	184	n/a	Trisomy 9, 11, 15	90%	n/a	n/a	n/a	n/a	
33	CTF038	SMM	IgG Kappa	20%	523	n/a	n/a	76%	n/a	n/a	n/a	n/a	
34	CTF039	SMM	IgG Lambda	20%	232	t(11;14)	n/a	n/a	✓	n/a	n/a	n/a	
35	CTF040	SMM	IgG Lambda	20%	171	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
36	CTF041	SMM	IgG Kappa	10%	386	t(11;14)	✓	n/a	n/a	n/a	n/a	n/a	
37	CTF042	SMM	IgG Kappa	20%	230	14q IGH sep	n/a	n/a	n/a	✓	n/a	n/a	
38	CTF043	SMM	Kappa LC	30%	218	t(11;14), 100%	n/a	n/a	n/a	100%	n/a	n/a	
39	CTF044	MGUS	IgG Lambda	n/a	872	n/a	n/a	n/a	n/a	n/a	n/a	n/a	
40	CTF045	SMM	IgG Kappa	20%	1371	t(11;14), 63%	n/a	n/a	n/a	n/a	n/a	n/a	
42	CTF049	MM	IgG Lambda	80%	1513	t(14;20), 94%	n/a	n/a	n/a	n/a	n/a	n/a	
44	CTF051	SMM	IgG Kappa	10%	161	t(14;20), 94%	n/a	n/a	n/a	✓	n/a	n/a	
49	CTF056	SMM	IgA Kappa	40%	89	t(14;16)	Trisomy 9	✓	n/a	✓	n/a	n/a	
50	CTF057	MM	IgG Kappa	70%	109	n/a	n/a	n/a	n/a	n/a	n/a	n/a	

Table S5

Clinical BM FISH results of peripheral blood only cohort and CTCs recovered.

Single Nucleotide Variants and short indels	Total	Median per participant (range)	Median clonal per participant (range)	Median subclonal per participant (range)
Total	257423	3676 (314 to 30929)	2413 (281 to 8920)	1248 (31 to 22009)
By variant class				
3'UTR	1638	20 (2 to 262)	13 (2 to 80)	7 (0 to 182)
5'Flank	10090	130 (11 to 1472)	84 (10 to 396)	43 (1 to 1076)
5'UTR	650	7 (1 to 143)	4 (0 to 31)	3 (0 to 112)
Could not be determined	4	1 (0 to 1)	1 (0 to 1)	0 (0 to 1)
De novo start in frame	6	1 (0 to 1)	1 (0 to 1)	0 (0 to 1)
De novo start out of frame	8	1 (0 to 1)	0 (0 to 1)	0 (0 to 1)
Frameshift deletion	52	2 (0 to 5)	1 (0 to 2)	1 (0 to 3)
Frameshift insertion	12	1 (0 to 1)	1 (0 to 1)	0 (0 to 1)
Intergenic region	130641	1946 (179 to 14476)	1285 (159 to 4338)	650 (19 to 10138)
Inframe deletion	18	1 (0 to 2)	0 (0 to 2)	1 (0 to 1)
Inframe insertion	3	1 (0 to 1)	1 (0 to 1)	0 (0 to 1)
Intron	88794	1202 (96 to 11712)	777 (87 to 3253)	417 (9 to 8459)
Missense mutation	1530	19 (2 to 242)	10 (1 to 55)	8 (0 to 187)
Nonsense mutation	95	1 (0 to 24)	1 (0 to 5)	1 (0 to 19)
Nonstop mutation	1	1 (0 to 1)	1 (1 to 1)	0 (0 to 0)
RNA	23136	342 (24 to 2482)	218 (21 to 735)	116 (2 to 1747)
Silent	622	7 (0 to 100)	5 (0 to 23)	2 (0 to 77)
Splice site	117	2 (0 to 11)	1 (0 to 6)	1 (0 to 7)
Start codon SNP	5	1 (0 to 1)	0 (0 to 1)	1 (0 to 1)
Translation start site	1	1 (0 to 1)	0 (0 to 0)	1 (1 to 1)

Table S6

Enumeration of single nucleotide variants and short insertions and deletions discovered from WGS of CTCs.

Structural Variants	Total	Median (Range)
Total	545 (100.0%)	4 (1 - 72)
By variant class		
Deletion	161 (29.5%)	2 (0 to 23)
Inter-chromosomal translocation	91 (16.7%)	2 (0 to 15)
Inversion	84 (15.4%)	2 (0 to 22)
Long range structural variant	161 (29.5%)	2 (0 to 34)
Tandem duplication	48 (8.8%)	1 (0 to 5)

Table S7

Enumeration of structural variants reconstructed from WGS of CTCs discovered by tumor-normal matched analysis with the Structural Variant detection workflow (Supplementary Materials and Methods).