Deep targeted sequencing of *TP53* in chronic lymphocytic leukemia: clinical impact at diagnosis and at time of treatment

Christian Brieghel,[1] Savvas Kinalis,[2] Christina W. Yde,[2] Ane Y. Schmidt,[2] Lars Jønson,[2] Michael A. Andersen,[1] Caspar da Cunha-Bang,[1] Lone B. Pedersen,[1] Christian H. Geisler,[1] Finn C. Nielsen[2] and Carsten U. Niemann[1]

[1]Department of Hematology, Rigshospitalet and [2]Center for Genomic Medicine, Rigshospitalet, Copenhagen, Denmark

Supplementary information

Deep targeted sequencing of *TP53* in chronic lymphocytic leukemia:

clinical impact at diagnosis and at time of treatment.

Christian Brieghel[1], Savvas Kinalis[2], Christina W. Yde[2], Ane Y. Schmidt[2], Lars Jønson[2], Michael A. Andersen[1], Caspar da Cunha-Bang[1], Lone B. Pedersen[1], Christian Geisler[1], Finn C. Nielsen[2], and Carsten U. Niemann[1].

1. Department of Hematology, Rigshospitalet, Copenhagen, Denmark.
2. Center for Genomic Medicine, Rigshospitalet, Copenhagen, Denmark.

Content

Supplemental methods

Supplemental results

Supplemental references

Supplemental table legends

Supplemental figure legends

Supplemental methods

Deep-targeted sequencing

Peripheral blood mononuclear cells (PBMC) were isolated from fresh EDTA blood with Lymfoprep (Axis-Shield, Oslo, Norway). Genomic DNA was extracted using QIAamp DNA Blood Midi kit (Qiagen, Hilden, Germany) and quantified on a Qubit® 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The SU-DHL4 cell line was cultured in 90% RPMI 1640 media, 19% FCS, and 1% streptomycin at 37°C. Primers were designed to flank *TP53* exons 2-10 (RefSeq NM_00546.5) including splicing sites with 2 bp intronic overlap. We used one primer-pair for exons 2, 3, 7, 8, and 10, two primer-pairs for exon 5 and 9, while exon 4 was covered using 4 primer-pairs. Patient samples were amplified using Phusion® Hot Start II High-Fidelity DNA polymerase (Thermo Fisher Scientific) with 30 cycles of PCR. Library preparation was performed on a SciClone G3 Liquid Handling Workstation (Perkin Elmer, Waltham, MA, USA) according to the manufacturer's protocol: KAPA DNA Library Preparation, Rev 1 (Roche Nimblegen, Madison, WI, USA). In brief, amplicons were purified using SeqCap Pure Capture Bead Kit containing AMPure® XP Beads (Beckman, Brea, CA, USA). Following end repair and poly(A)-tailing, adaptor ligation was performed with SeqCap Adapter Kit A and B (Roche NimbleGen) or NEXTflex™ DNA Barcodes 96 (Bioo Scientific, Austin, TX, USA). Following seven cycles of PCR and an additional AMPure step, libraries were pooled either 24 or 96 samples per lane and sequenced as paired-end on a HiSeq2500. Using HiSeq® SBS Kit v4 (2x125 base PE; Illumina, San Diego, CA, US), 75% of reads in the target region were paired.

Determination of sensitivity of the assay

To establish the sensitivity of the *TP53* mutation assay, serial dilution of Sanger validated *TP53* mutations (*TP53*muts) were sequenced. Pilot study samples were handled as described above using only SeqCap Adapters (Roche NimbleGen) and sequenced as paired-end on a MiSeq using MiSeq Reagent Kit v2 (Illumina, 2x250 base PE). With mutations detected as low as 0.023% variant allele frequency (VAF), the limit of detection (LOD) was established at 0.2% VAF while a dilution step (dilution factor 5) was included to increase the robustness of the assay (Figure S2A). By subtraction of the VAF from variants called in pure cell line DNA, the correlation between dilution grade and detected VAF was improved ($r^2 > 0.9957$ vs $r^2 > 0.9967$, squared Pearson's correlation coefficient). Thus, this approach was chosen for future analyses, replacing any negative adjusted VAF with zero. The range of the dilution step and the ability to detect insertions and deletions (indels) was confirmed based on similar serial dilution of patient samples with known *TP53* indels (Figure S2B). Finally, the performance of the assay was evaluated assessing eight del(17p) samples with unknown *TP53* mutational status. As expected, seven of eight samples harbored *TP53* mutations (Table S1).

Bioinformatic workflow

Based on the pilot studies, a workflow for detection of low burden variants was developed in CLC Biomedical Genomics Workbench 3.0 (CLC BGW, Qiagen). Paired reads were trimmed 20 bp 5` to

remove primers and 1 bp 3` to remove bad quality reads further allowing for a quality-score of 0.01. Reads were mapped to the hg19 (ensembl GRCh37.73) reference genome and realigned using 2 multipasses. Directional frequency filter was set to zero to ensure variant calls in unpaired reads (25% of target region). The combined call of both a low frequency (VAF ≥0.001%) and basic variant detection tool (VAF ≥2%) were used to ensure detection of both high and low burden mutations. To describe the distribution of inevitable low frequency errors (error distribution), all variants with a minimum of 1 variant read and VAF ≥0.001% were called.

Dilution match algorithm

In analyses downstream of CLC BGW, the adjusted dilution ratios (aDRs) and the dilution grades (DGs: reference allele frequency of p.Arg273Cys) were plotted and expected to match a line with a slope of one (Figure S3A). However, a cluster with aDRs around 1 (i.e. a distance to a line with slope of 1 above 0.5) was also observed and consequently excluded. We next calculated the 99.9% reference range (mean ± 3.27 SD) of remaining distances to a line with a slope of 1 and trimmed the reference range to an absolute distance ≤0.16. Variants within this distance were considered to have a dilution match (Figure S3B).

$$\left| \left(RAF(p.R273C)_{diluted} - \frac{VAF_{diluted} - DHL4\,VAF_{undiluted}}{VAF_{undiluted} - DHL4\,VAF_{undiluted}} \right) \right| \leq 0.16$$

Dilution match algorithm

Stereotypic error model

We further modeled the distribution of stereotypic errors[1] to identify outliers considered as true mutations.[2] As modeling of variant reads included inevitable differences in read depth (coverage), we modeled VAF in two steps as summarized in Figures S4-S6. First, each unique genomic position including nucleotide change (e.g. g.17:7579472G>C) were grouped and modeled only if observed in 20 samples or more. Secondly, we grouped each unique nucleotide change (e.g. A>T) and modeled only if observed in 20 samples or more. Extreme outliers were excluded from modeling by truncating VAFs that were further than 2 and 4 standard deviations (SDs) from the mean of the log transformed VAF distributions. Subsequently, truncated VAFs were modeled to fit gamma distributions that allowed us to identify outliers using Bonferroni adjusted right-sided *P*-values of 10^{-2}, 10^{-4}, and 10^{-6}. Unfit for modeling, variants with a rare nucleotide change observed in less than 20 samples were called with a LOD of 2% VAF. Due to the design of the dilution match algorithm, c.817C>T (p.Arg273Cys) could not be modeled by this approach. Thus, this mutation was called with a LOD of 2% VAF.

Reporting variant findings

Variants identified by the dilution match algorithm (DMA) and stereotypic error model (SEM) with a minimum of 10 variant reads and a VAF ≥0.2% were called and referenced in the IARC *TP53* database (http://p53.iarc.fr) excluding validated polymorphisms, synonymous mutations and mutations predicted to encode functional p53 (Table S5).

Comparison of results from the DMA and SEM was performed with 2x2 contingency tables using combinations of annotated *P*-values and SDs for LODs 0.2%, 0.3%, and 1% VAF to evaluate the performance of both methods (Table S3). The three LODs represented our threshold, the lower threshold used in previous publications,[2-4] and the threshold of minor *TP53* mutations,[3] respectively.

All analyses downstream of CLC BGW were performed with R version 3.4.1 using packages tidyverse, fitdistrplus, robustHD, epiR, Publish, survival and survminer.[5] Source code is available upon request.

Validation by droplet digital PCR and capture based NGS

Droplet digital PCR (ddPCR) was used to validate the first 30 low burden *TP53* mutations. Genomic DNA was fragmented to approximately 150 bp using S220 Focused-ultrasonicator (Covaris, Woburn, MA, USA). Custom made allele specific Prime Assay™ probes (Bio-Rad, Hercules, CA, USA) were added to triplicates of fragmented DNA from either *TP53* aberrated patients or healthy donors. Water acted as non-template controls. Following droplet generation and 39 cycles of PCR using ddPCR Supermix for Probes (no dUTP, Bio-Rad), fluorescent FAM and HEX signals were read and analyzed on QX200™ Droplet Digital™ PCR System using QuantaSoft™ 1.7 (Bio-Rad).

Samples from newly diagnosed patients referred between 2008 and 2014 were sequenced using a custom made SeqCap EZ Choice gene panel (Roche Nimblegen) containing *TP53* exons 2-10. DNA extracted from peripheral blood mononuclear cells (PBMC) was fragmented (Covaris) and purified using AMPure XP Beads (Beckman). Following end repair and poly(A)-tailing, fragments were adaptor ligated using NEXTflex™ DNA Barcodes 96 (Bioo Scientific) using KAPA Library Preparation Kit (Roche NimbleGen) and amplified with 7 cycles of PCR. Pooled libraries were hybridized twice using SeqCap EZ Kit (Roche NimbleGen) to capture target regions and sequenced as pair-end on a HiSeq2500 using HiSeq® SBS Kit v4 (2x125 base PE; Illumina). All samples were sequenced in a single HiSeq run to obtain 80% of output and subsequently analyzed in CLC BGW allowing for a VAF >1%. CLC BGW workflow is available upon request.

Evaluation of CLL cell fraction by flow cytometry

To evaluate the fraction of CLL cells in purified PBMCs, we reanalyzed the diagnostic fcs files. Samples were originally run on a FACSCalibur or FACSCanto II (BD, Franklin Lakes, NJ, USA). Reanalysis was performed using FlowJo® vX.0.7 (Flowjo, Ashland, OR, USA). The fraction of CLL cells (CD5+, CD19+) was calculated from mononuclear cells gated by CD45 and SSC.

Supplemental results

SEM identified a c.848G>A (p.Arg283His) in 45 patients (range: 0.20-0.43% VAF, Figure S7A). However, none of these were positive by DMA. This specific mutation was considered low frequent noise and excluded from downstream analyses comparing DMA and SEM. Four variants were only detected by DMA including three c.848_849delGC (p.Arg283fs*22) positioned at a problematic position with a wide distribution of background noise (Figure S7 A-B) and one c.488A>G (p.Tyr163Cys) that was validated by ddPCR but excluded based on SEM modeled on only 24 events (Figure S7C). Twenty-two variants were only detected by SEM. This included one c.338T>G (p.Phe113Cys) identified at 0.6% VAF, while the remaining 21 variants fell below 0.3% VAF (Table S4); the LOD applied in previous studies.[2, 4] All 70 true negative variants with VAF ≥1% excluded by both the SEM and the DMA consisted of a splicing site variant c.74+2T>G. This variant showed a multimodal distribution considered unfit for modeling and is thought to be a mapping error rather than a result of PCR (Figure S7D).

Due to redundant mutations, validation of the initial 20 *TP53*muts by ddPCR included 30 low burden *TP53*mut (VAF ≤10%) with a median VAF of 0.91% (IQR: 0.41-2.95%). All tested samples were validated (Table S6). In addition, three high burden *TP53*muts (VAF >10%) and two minor *TP53*muts (VAF <1%) called with a LOD of 0.1% VAF were also validated. Among *TP53*muts above 1% VAF tested by capture based targeted NGS, 24 (92%) were validated: 2 mutations could not be validated (VAF of 1.2% and 2.9%, Table S4).

Supplemental references

1.	Newman AM, Lovejoy AF, Klass DM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. Nat Biotechnol. 2016;34(5):547-555.

2.	Rossi D, Khiabanian H, Spina V, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. Blood. 2014;123(14):2139-2147.
3.	Malcikova J, Stano-Kozubik K, Tichy B, et al. Detailed analysis of therapy-driven clonal evolution of TP53 mutations in chronic lymphocytic leukemia. Leukemia. 2015;29(877-885.
4.	Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. Blood. 2016;127(17):2122-2130.
5.	Team RC. R: A Language and Environment for Statistical Computing. 2017.

Supplemental table legends

Table S1. Analysis of 8 del(17p) test samples with unknown *TP53* status. Nine *TP53* mutations were identified in the 7 of the 8 samples using the dilution match algorithm. Dilution grade (DG) and dilution ratio (DR) were highly similar with short distances from a slope of one (DFSO = DG - DR).

Table S2. The primers used for PCR amplification of *TP53* exons 2-10.

Table S3. Contingency tables show results for the entire study cohort (n=308) of the dilution match algorithm (DMA) and stereotypic error modeling (SEM) at three different limits of detection (LOD of VAF ≥0.2%, VAF ≥0.3%, and VAF ≥1%) and 6 different SEM settings (*P*-values and SD). Standard deviations (SD) were used for truncating extreme outliers before modeling and Bonferroni adjusted *P*-values were used to identify modeled outliers.

Table S4. *TP53* mutations identified by dilution match algorithm (DMA) and stereotypic error model (SEM). Only variants identified by both DMA and SEM were considered true mutations. *Indicates newly diagnosed patients.

Table S5. *TP53* mutations excluded from analysis. Mutations predicted to encode functional p53 and silent mutations were excluded.

Table S6. Validation by droplet digital PCR. Probes for initial *TP53* findings validated all 33 tested mutations above the limit of detection (LOD) of the assay (0.2% VAF). Two mutations below the LOD were also validated, but not included in the final report. Note that 14 mutations in 7 non-diagnostic patients, who were neither sampled 0-200 days before treatment were also tested and validated.

Table S7. Patient characteristics in newly diagnosed patients stratified according to *TP53* (A) and IGHV status (B). *TP53* wild-type (TP53wt), *TP53* mutations without del(17p) (TP53mut), del(17p) regardless of *TP53* mutational status (Del(17p)), beta-2-microglobulin (B2M), fluorescent in situ hybridization (FISH).

Table S8. Patient characteristics at time of treatment stratified according to *TP53* (A) and IGHV status (B). *TP53* wild-type (TP53wt), *TP53* mutations without del(17p) (TP53mut), del(17p) regardless of *TP53* mutational status (Del(17p)), beta-2-microglobulin (B2M), fluorescent in situ hybridization (FISH).

Case	g description	Undiluted VAF (%)	Diluted VAF (%)	aDR	DG	DFSO
Pilot1	17:7577536 T>C	11.40	2.38	0.209	0.216	-0.007
Pilot2	17:7577536 T>C	11.52	2.32	0.201	0.250	-0.049
Pilot3	17:7579446 T>G	0.31	0.07	0.226	0.161	0.065
Pilot3	17:7579440^7579441 ->G	0.29	0.05	0.189	0.161	0.028
Pilot4	17:7578239 C>-	77.38	8.79	0.114	0.134	-0.020
Pilot5	17:7578203 C>G	11.01	1.15	0.104	0.113	-0.009
Pilot6	17:7578403 C>T	75.47	7.34	0.097	0.117	-0.020
Pilot7	17:7578205 C>A	39.71	8.77	0.221	0.252	-0.031
Pilot7	17:7578268 A>C	13.98	3.16	0.226	0.252	-0.026

Table S1. Analysis of 8 del(17p) test samples with unknown *TP53* status. Nine *TP53* mutations were identified in the 7 of the 8 samples using the dilution match algorithm. Dilution grade (DG) and the adjusted dilution ratio (aDR) showed high correlation as seen by short distances from a slope of one (DFSO = DG - aDR).

TP53	Strand	Oligo
Exon 2	Forward	AGTGTCTCATGCTGGATCCC
Exon 2	Reverse	GTGGGCCTGCCCTTCCAATG
Exon 3	Forward	AGAGACCTGTGGGAAGCGAA
Exon 3	Reverse	CCCAGCCCAACCCTTGTCCT
Exon 4a	Forward	CAGGGCAGCTACGGTTTCCG
Exon 4a	Reverse	CTGACAGGAAGCCAAAGGGT
Exon 4b	Forward	CCCAGAATGCCAGAGGCTGC
Exon 4b	Reverse	AGTCACAGACTTGGCTGTCC
Exon 4c	Forward	CCAGAATGCCAGAGGCTGCT
Exon 4c	Reverse	GGGCAACTGACCGTGCAAGT
Exon 4d	Forward	CTGAGGACCTGGTCATCTGA
Exon 4d	Reverse	TGTAGGAGCTGCTGGTGCAG
Exon 5a	Forward	CCCTGTGCAGCTGTGGGTTG
Exon 5a	Reverse	CTGGGCAACCAGCCCTGTCG
Exon 5b	Forward	TTGTGCCCTGACTTTCAACT
Exon 5b	Reverse	GCTGTGACTGCTTGTAGATG
Exon 6	Forward	CAGGCCTCTGATTCCTCACT
Exon 6	Reverse	CTCCTCCCAGAGACCCCAGT
Exon 7	Forward	ACTGGCCTCATCTTGGGCCT
Exon 7	Reverse	GCCAGTGTGCAGGGTGGCAA
Exon 8a	Forward	GAACAGCTTTGAGGTGCGTG
Exon 8a	Reverse	CGCTTCTTGTCCTGCTTGCT
Exon 8b	Forward	CAGGTAGGACCTGATTTCCT
Exon 8b	Reverse	CCTGGGGGCAGCTCGTGGTG
Exon 9a	Forward	GCAGTTATGCCTCAGATTCA
Exon 9a	Reverse	AACTTTCCACTTGATAAGAG
Exon 9b	Forward	GCAGTTATGCCTCAGATTCA
Exon 9b	Reverse	GATAAGAGGTCCCAAGACTT
Exon 10	Forward	TGAACCATCTTTTAACTCAG
Exon 10	Reverse	AGGAAGGGGCTGAGGTCACT

Table S2. The primers used for PCR amplification of *TP53* exons 2-10.

LOD

Contingency tables organized by P (10^{-2}, 10^{-4}, 10^{-6}), SD (2, 4), and LOD (VAF ≥0.2%, VAF ≥0.3%, VAF ≥1%).

P = 10^{-2}, SD = 2

A — VAF ≥0.2%

	DMA +	DMA −	Total
SEM +	94	22	116
SEM −	4	386	390
Total	98	408	506

B — VAF ≥0.3%

	DMA +	DMA −	Total
SEM +	84	1	85
SEM −	3	127	130
Total	87	128	215

C — VAF ≥1%

	DMA +	DMA −	Total
SEM +	52	0	52
SEM −	0	70	70
Total	52	70	122

P = 10^{-2}, SD = 4

D — VAF ≥0.2%

	DMA +	DMA −	Total
SEM +	94	22	116
SEM −	4	386	390
Total	98	408	506

E — VAF ≥0.3%

	DMA +	DMA −	Total
SEM +	84	1	85
SEM −	3	127	130
Total	87	128	215

F — VAF ≥1%

	DMA +	DMA −	Total
SEM +	52	0	52
SEM −	0	70	70
Total	52	70	122

P = 10^{-4}, SD = 2

G — VAF ≥0.2%

	DMA +	DMA −	Total
SEM +	94	22	116
SEM −	4	386	390
Total	98	408	506

H — VAF ≥0.3%

	DMA +	DMA −	Total
SEM +	84	1	85
SEM −	3	127	130
Total	87	128	215

I — VAF ≥1%

	DMA +	DMA −	Total
SEM +	52	0	52
SEM −	0	70	70
Total	52	70	122

P = 10^{-4}, SD = 4

J — VAF ≥0.2%

	DMA +	DMA −	Total
SEM +	93	22	115
SEM −	5	386	391
Total	98	408	506

K — VAF ≥0.3%

	DMA +	DMA −	Total
SEM +	83	1	84
SEM −	4	127	131
Total	87	128	215

L — VAF ≥1%

	DMA +	DMA −	Total
SEM +	52	0	52
SEM −	0	70	70
Total	52	70	122

P = 10^{-6}, SD = 2

M — VAF ≥0.2%

	DMA +	DMA −	Total
SEM +	93	22	115
SEM −	5	386	391
Total	98	408	506

N — VAF ≥0.3%

	DMA +	DMA −	Total
SEM +	83	1	84
SEM −	4	127	131
Total	87	128	215

O — VAF ≥1%

	DMA +	DMA −	Total
SEM +	52	0	52
SEM −	0	70	70
Total	52	70	122

P = 10^{-6}, SD = 4

P — VAF ≥0.2%

	DMA +	DMA −	Total
SEM +	91	22	113
SEM −	7	386	393
Total	98	408	506

Q — VAF ≥0.3%

	DMA +	DMA −	Total
SEM +	82	1	83
SEM −	5	127	132
Total	87	128	215

R — VAF ≥1%

	DMA +	DMA −	Total
SEM +	52	0	52
SEM −	0	70	70
Total	52	70	122

Table S3. Contincency tables

Sample ID	c description	p description	VAF (%)	Exon	Effect	Tool	Validated
CLL6*	c.848_849delGC	p.Arg283fs	0.36	Exon 8	frame shift	DMA	Not tested
CLL14	c.613T>A	p.Tyr205Asn	0.22	Exon 6	missense	SEM	Not tested
CLL15*	c.672+1G>A	NA	97.45	Exon 6	splice	SEM & DMA	Not tested
CLL18	c.667C>A	p.Pro223Thr	0.21	Exon 6	missense	SEM	Not tested
CLL20*	c.667C>A	p.Pro223Thr	0.29	Exon 6	missense	SEM	Not tested
CLL20*	c.565G>T	p.Ala189Ser	0.27	Exon 6	missense	SEM	Not tested
CLL21*	c.832C>T	p.Pro278Ser	0.44	Exon 8	missense	SEM & DMA	ddPCR
CLL21*	c.667C>A	p.Pro223Thr	0.29	Exon 6	missense	SEM	Not tested
CLL21*	c.1086C>A	p.Ser362Arg	0.26	Exon 10	missense	SEM	Not tested
CLL21*	c.565G>T	p.Ala189Ser	0.25	Exon 6	missense	SEM	Not tested
CLL22	c.413C>T	p.Ala138Val	5.34	Exon 5	missense	SEM & DMA	Not tested
CLL22	c.395A>G	p.Lys132Arg	4.77	Exon 5	missense	SEM & DMA	ddPCR
CLL22	c.767C>G	p.Thr256Arg	2.81	Exon 7	missense	SEM & DMA	Not tested
CLL22	c.451C>T	p.Pro151Ser	1.63	Exon 5	missense	SEM & DMA	Not tested
CLL22	c.742C>T	p.Arg248Trp	1.47	Exon 7	missense	SEM & DMA	ddPCR
CLL22	c.673-2A>T	NA	1.05	Exon 7	splice	SEM & DMA	Not tested
CLL22	c.673-2A>G	NA	0.99	Exon 7	splice	SEM & DMA	Not tested
CLL22	c.659A>G	p.Tyr220Cys	0.94	Exon 6	missense	SEM & DMA	ddPCR
CLL22	c.701A>G	p.Tyr234Cys	0.91	Exon 7	missense	SEM & DMA	ddPCR
CLL22	c.673-2A>C	NA	0.87	Exon 7	splice	SEM & DMA	Not tested
CLL22	c.577C>G	p.His193Asp	0.71	Exon 6	missense	SEM & DMA	Not tested
CLL22	c.736A>G	p.Met246Val	0.49	Exon 7	missense	SEM & DMA	Not tested
CLL22	c.542G>A	p.Arg181His	0.33	Exon 5	missense	SEM & DMA	Not tested
CLL22	c.783-2A>T	NA	0.29	Exon 8	splice	SEM	Not tested
CLL22	c.667C>A	p.Pro223Thr	0.29	Exon 6	missense	SEM	Not tested
CLL22	c.1086C>A	p.Ser362Arg	0.28	Exon 10	missense	SEM	Not tested
CLL22	c.565G>T	p.Ala189Ser	0.28	Exon 6	missense	SEM	Not tested
CLL22	c.560-2A>T	NA	0.23	Exon 6	splice	SEM & DMA	Not tested
CLL22	c.404G>T	p.Cys135Phe	0.22	Exon 5	missense	SEM	Not tested
CLL22	c.577C>A	p.His193Asn	0.20	Exon 6	missense	SEM	Not tested
CLL22	c.395A>C	p.Lys132Thr	0.20	Exon 5	missense	SEM	Not tested
CLL33*	c.809T>C	p.Phe270Ser	2.55	Exon 8	missense	SEM & DMA	ddPCR
CLL45*	c.814G>A	p.Val272Met	0.45	Exon 8	missense	SEM & DMA	ddPCR
CLL56*	c.82G>T	p.Glu28*	0.24	Exon 3	nonsense	SEM	Not tested
CLL65	c.596G>T	p.Gly199Val	2.95	Exon 6	missense	SEM & DMA	ddPCR
CLL65	c.700T>A	p.Tyr234Asn	0.41	Exon 7	missense	SEM & DMA	ddPCR
CLL74	c.814G>A	p.Val272Met	30.15	Exon 8	missense	SEM & DMA	ddPCR
CLL74	c.239delC	p.Pro80fs	10.79	Exon 4	frame shift	SEM & DMA	Not tested
CLL74	c.488A>G	p.Tyr163Cys	3.50	Exon 5	missense	SEM & DMA	ddPCR
CLL74	c.332T>C	p.Leu111Pro	3.14	Exon 4	missense	SEM & DMA	Not tested
CLL74	c.743G>A	p.Arg248Gln	3.05	Exon 7	missense	SEM & DMA	Not tested
CLL74	c.838A>G	p.Arg280Gly	2.24	Exon 8	missense	SEM & DMA	Not tested
CLL74	c.722C>G	p.Ser241Cys	2.12	Exon 7	missense	SEM & DMA	Not tested
CLL74	c.524G>A	p.Arg175His	1.98	Exon 5	missense	SEM & DMA	Not tested
CLL74	c.920-2A>G	NA	1.90	Exon 9	splice	SEM & DMA	Not tested
CLL74	c.581T>G	p.Leu194Arg	1.75	Exon 6	missense	SEM & DMA	Not tested
CLL74	c.993+2T>C	NA	1.01	Exon 9	splice	SEM & DMA	Not tested
CLL74	c.818G>A	p.Arg273His	0.78	Exon 8	missense	SEM & DMA	Not tested
CLL74	c.338T>G	p.Phe113Cys	0.64	Exon 4	missense	SEM	Not tested
CLL74	c.701A>G	p.Tyr234Cys	0.52	Exon 7	missense	SEM & DMA	ddPCR
CLL74	c.659A>G	p.Tyr220Cys	0.48	Exon 6	missense	SEM & DMA	Not tested
CLL74	c.830G>A	p.Cys277Tyr	0.47	Exon 8	missense	SEM & DMA	Not tested
CLL74	c.272G>A	p.Trp91*	0.45	Exon 4	nonsense	SEM & DMA	Not tested
CLL74	c.413C>T	p.Ala138Val	0.38	Exon 5	missense	SEM & DMA	Not tested
CLL74	c.536A>G	p.His179Arg	0.27	Exon 5	missense	SEM & DMA	Not tested
CLL74	c.395A>G	p.Lys132Arg	0.27	Exon 5	missense	SEM & DMA	ddPCR

(continued)

Sample ID	c description	p description	VAF (%)	Exon	Effect	Tool	Validated
CLL74	c.673-2A>T	NA	0.21	Exon 7	splice	SEM	Not tested
CLL77*	c.844C>G	p.Arg282Gly	0.26	Exon 8	missense	SEM & DMA	Not tested
CLL94*	c.745A>G	p.Arg249Gly	11.06	Exon 7	missense	SEM & DMA	Capture
CLL98*	c.472C>G	p.Arg158Gly	6.32	Exon 5	missense	SEM & DMA	ddPCR & Capture
CLL104*	c.742C>T	p.Arg248Trp	0.29	Exon 7	missense	SEM & DMA	ddPCR
CLL107*	c.388C>T	p.Leu130Phe	0.21	Exon 5	missense	SEM & DMA	Not tested
CLL109*	c.902delC	p.Pro301fs	0.24	Exon 8	frame shift	SEM	Not tested
CLL111*	c.823T>C	p.Cys275Arg	0.58	Exon 8	missense	SEM & DMA	Not tested
CLL112*	c.524G>A	p.Arg175His	48.92	Exon 5	missense	SEM & DMA	Capture
CLL116*	c.716A>G	p.Asn239Ser	11.86	Exon 7	missense	SEM & DMA	Capture
CLL120*	c.736A>G	p.Met246Val	18.24	Exon 7	missense	SEM & DMA	Capture
CLL120*	c.527G>T	p.Cys176Phe	0.20	Exon 5	missense	SEM & DMA	Not tested
CLL135*	c.610delG	p.Glu204fs	77.27	Exon 6	frame shift	SEM & DMA	Capture
CLL137	c.530C>T	p.Pro177Leu	2.25	Exon 5	missense	SEM & DMA	Not tested
CLL139*	c.726C>G	p.Cys242Trp	0.66	Exon 7	missense	SEM & DMA	Not tested
CLL144*	c.734G>A	p.Gly245Asp	79.37	Exon 7	missense	SEM & DMA	Capture
CLL144*	c.536A>T	p.His179Leu	1.22	Exon 5	missense	SEM & DMA	Not validated (Capture)
CLL145*	c.848_849delGC	p.Arg283fs	0.45	Exon 8	frame shift	DMA	Not tested
CLL146*	c.848_849delGC	p.Arg283fs	0.34	Exon 8	frame shift	DMA	Not tested
CLL150*	c.626_627delGA	p.Arg209fs	40.10	Exon 6	frame shift	SEM & DMA	ddPCR & Capture
CLL161	c.376-1G>A	NA	4.50	Exon 5	splice	SEM & DMA	Not tested
CLL165*	c.309C>G	p.Tyr103*	1.94	Exon 4	nonsense	SEM & DMA	Capture
CLL177*	c.830G>A	p.Cys277Tyr	2.19	Exon 8	missense	SEM & DMA	Capture
CLL179*	c.524G>A	p.Arg175His	0.45	Exon 5	missense	SEM & DMA	Not tested
CLL187*	c.318C>G	p.Ser106Arg	2.87	Exon 4	missense	SEM & DMA	Capture
CLL191*	c.438G>A	p.Trp146*	26.41	Exon 5	nonsense	SEM & DMA	Capture
CLL193	c.439_440insG	p.Val147fs	41.24	Exon 5	frame shift	SEM & DMA	Not tested
CLL193	c.733G>A	p.Gly245Ser	29.85	Exon 7	missense	SEM & DMA	Not tested
CLL193	c.481G>A	p.Ala161Thr	0.29	Exon 5	missense	SEM	Not tested
CLL193	c.488A>G	p.Tyr163Cys	0.23	Exon 5	missense	DMA	Not tested
CLL193	c.128T>A	p.Leu43*	0.21	Exon 4	nonsense	SEM & DMA	Not tested
CLL223*	c.644G>T	p.Ser215Ile	40.75	Exon 6	missense	SEM & DMA	Capture
CLL223*	c.581T>G	p.Leu194Arg	13.62	Exon 6	missense	SEM & DMA	Capture
CLL233*	c.772G>A	p.Glu258Lys	0.49	Exon 7	missense	SEM & DMA	Not tested
CLL250*	c.572_573insCC	p.Gln192fs	0.48	Exon 6	frame shift	SEM & DMA	Not tested
CLL251*	c.869G>A	p.Arg290His	50.49	Exon 8	missense	SEM & DMA	Capture
CLL258*	c.848G>C	p.Arg283Pro	0.27	Exon 8	missense	SEM & DMA	Not tested
CLL261*	c.455_456insT	p.Pro153fs	0.36	Exon 5	frame shift	SEM & DMA	Not tested
CLL269*	c.440T>G	p.Val147Gly	33.23	Exon 5	missense	SEM & DMA	Capture
CLL275*	c.524G>A	p.Arg175His	3.02	Exon 5	missense	SEM & DMA	Not tested
CLL288*	c.581T>G	p.Leu194Arg	59.85	Exon 6	missense	SEM & DMA	Capture
CLL294*	c.577C>T	p.His193Tyr	27.33	Exon 6	missense	SEM & DMA	Capture
CLL299*	c.388C>T	p.Leu130Phe	0.94	Exon 5	missense	SEM & DMA	Not tested
CLL302*	c.314G>A	p.Gly105Asp	0.31	Exon 4	missense	SEM & DMA	Not tested
CLL310*	c.482C>A	p.Ala161Asp	20.83	Exon 5	missense	SEM & DMA	Capture
CLL310*	c.626_627delGA	p.Arg209fs	13.80	Exon 6	frame shift	SEM & DMA	ddPCR & Capture
CLL325*	c.737T>G	p.Met246Arg	8.84	Exon 7	missense	SEM & DMA	Capture
CLL327*	c.830G>T	p.Cys277Phe	3.92	Exon 8	missense	SEM & DMA	Capture
CLL328	c.469G>T	p.Val157Phe	79.62	Exon 5	missense	SEM & DMA	Not tested
CLL328	c.764delT	p.Ile255fs	0.43	Exon 7	frame shift	SEM & DMA	Not tested
CLL328	c.529C>A	p.Pro177Thr	0.42	Exon 5	missense	SEM & DMA	Not tested
CLL328	c.645T>G	p.Ser215Arg	0.37	Exon 6	missense	SEM & DMA	Not tested
CLL328	c.993+1G>A	NA	0.30	Exon 9	splice	SEM & DMA	Not tested
CLL342*	c.626_627delGA	p.Arg209fs	1.60	Exon 6	frame shift	SEM & DMA	ddPCR & Capture
CLL342*	c.722C>T	p.Ser241Phe	0.42	Exon 7	missense	SEM & DMA	Not tested
CLL346*	c.742C>G	p.Arg248Gly	2.94	Exon 7	missense	SEM & DMA	Not validated (Capture)

(continued)

Sample ID	c description	p description	VAF (%)	Exon	Effect	Tool	Validated
CLL346*	c.743G>A	p.Arg248Gln	0.93	Exon 7	missense	SEM & DMA	Capture
CLL346*	c.641A>G	p.His214Arg	0.41	Exon 6	missense	SEM & DMA	Not tested
CLL348*	c.430C>T	p.Gln144*	0.29	Exon 5	nonsense	SEM	Not tested
CLL348*	c.482C>T	p.Ala161Val	0.27	Exon 5	missense	SEM	Not tested
CLL350*	c.523C>T	p.Arg175Cys	14.10	Exon 5	missense	SEM & DMA	Capture
CLL350*	c.919+1G>T	NA	0.24	Exon 8	splice	SEM & DMA	Not tested
CLL358*	c.824G>A	p.Cys275Tyr	0.41	Exon 8	missense	SEM & DMA	Not tested
CLL358*	c.80C>A	p.Pro27His	0.22	Exon 3	missense	SEM	Not tested

Table S4. *TP53* mutations identified in full study cohort (n=308) using the dilution match algorithm (DMA) and stereotypic error model (SEM). Only variants identified by both DMA and SEM were considered true mutations. *Indicates newly diagnosed patients.

Sample ID	c description	p description	VAF (%)	Exon	Exclusion criterion
CLL47	c.704A>G	p.Asn235Ser	49.18	Exon 7	Encoding functional p53
CLL61	c.672G>A	p.Glu224Glu	0.69	Exon 6	Silent mutation
CLL93	c.114A>C	p.Gln38His	0.21	Exon 4	Encoding functional p53
CLL122	c.826G>A	p.Ala276Thr	2.54	Exon 8	Encoding functional p53
CLL140	c.1060C>A	p.Gln354Lys	52.51	Exon 10	Encoding functional p53
CLL214	c.704A>G	p.Asn235Ser	48.96	Exon 7	Encoding functional p53
CLL239	c.410T>C	p.Leu137Pro	1.04	Exon 5	Encoding functional p53
CLL247	c.837G>A	p.Gly279Gly	1.43	Exon 8	Silent mutation
CLL250	c.570T>C	p.Pro190Pro	0.52	Exon 6	Silent mutation
CLL267	c.1015G>A	p.Glu339Lys	49.56	Exon 10	Encoding functional p53
CLL329	c.612G>A	p.Glu204Glu	50.05	Exon 6	Silent mutation

Table S5. *TP53* mutations excluded from analysis. Mutations predited to encode functional p53 and silent mutations were excluded.

Sample ID	Assay	VAF by ddPCR (%)	VAF by NGS (%)	H20 Control (%)	Cohort
CLL21	P278S	0.33	0.44	0.00	Newly diagnosed
CLL33	F270S	2.34	2.55	0.00	Newly diagnosed
CLL45	V272M	0.23	0.45	0.00	Newly diagnosed
CLL98	R158G	6.63	6.32	0.00	Newly diagnosed
CLL104	R248W	0.46	0.29	0.00	Newly diagnosed
CLL150	R209Kfs*6	40.94	40.10	0.00	Newly diagnosed
CLL148	R209Kfs*6	0.22	0.15	0.00	Newly diagnosed
CLL310	R209Kfs*6	14.38	13.80	0.00	Newly diagnosed
CLL315	R209Kfs*6	0.08	0.13	0.00	Newly diagnosed
CLL342	R209Kfs*6	1.67	1.60	0.00	Newly diagnosed
CLL22	Y220C	0.90	0.94	0.00	At time of treatment
CLL22	K132R	6.22	4.77	0.00	At time of treatment
CLL22	R248W	1.36	1.47	0.00	At time of treatment
CLL22	Y234C	1.01	0.91	0.00	At time of treatment
CLL65	G199V	3.26	2.95	0.00	At time of treatment
CLL65	Y234N	0.37	0.41	0.00	At time of treatment
CLL74	Y163C	3.28	3.50	0.00	At time of treatment
CLL74	V272M	32.32	30.15	0.06	At time of treatment
CLL74	Y234C	0.48	0.52	0.00	At time of treatment
CLL74	K132R	0.40	0.27	0.00	At time of treatment
CLL193	Y163C	0.14	0.23	0.00	At time of treatment
CLL4	Q331*	5.41	4.90	0.00	Non-diagnostic, not treated
CLL4	R209Kfs*6	2.07	1.82	0.00	Non-diagnostic, not treated
CLL10	P316Sfs*21	0.23	0.20	0.00	Non-diagnostic, not treated
CLL26	C135F	4.48	4.20	0.00	Non-diagnostic, not treated
CLL40	Y163C	0.45	0.53	0.00	Non-diagnostic, not treated
CLL40	K132R	0.39	0.35	0.00	Non-diagnostic, not treated
CLL40	R248W	0.32	0.29	0.00	Non-diagnostic, not treated
CLL40	Y234C	1.46	1.20	0.00	Non-diagnostic, not treated
CLL67	I195T	0.35	0.36	0.00	Non-diagnostic, not treated
CLL134	R248W	1.36	0.90	0.00	Non-diagnostic, not treated
CLL134	I195T	0.67	0.45	0.00	Non-diagnostic, not treated
CLL134	Y234C	0.43	0.34	0.00	Non-diagnostic, not treated
CLL186	R248W	0.41	0.44	0.00	Non-diagnostic, not treated
CLL186	Y234C	2.54	2.61	0.00	Non-diagnostic, not treated

Table S6. Droplet digital PCR (ddPCR) probes for initial *TP53* findings validated all 33 tested mutations above the threshold of 0.2% VAF. Two mutations below the VAF threshold of 0.2% were also validated, but not included in the final report. Note that 14 mutations in 7 non-diagnostic patients not sampled before 200 days of treatment were also tested and validated.

A) Patient characteristics according to TP53 mutational status in newly diagnosed patients

Variable		TP53wt (N=249)	VAF<1%, N (%)	VAF 1-10%, N (%)	VAF>10%, N (%)
			TP53 mutated (N=41)		
Age	<65 years	111 (44.6)	2 (13.3)	2 (20.0)	5 (31.2)
	>65 years	138 (55.4)	13 (86.7)	8 (80.0)	11 (68.8)
Binet	A	203 (85.7)	12 (92.3)	8 (88.9)	9 (60.0)
	B/C	34 (14.3)	1 (7.7)	1 (11.1)	6 (40.0)
	Unknown	12	2	1	1
B2M	<4.0 mg/L	184 (87.6)	12 (100.0)	7 (77.8)	10 (66.7)
	>4.0 mg/L	26 (12.4)	0 (0.0)	2 (22.2)	5 (33.3)
	Unknown	39	3	1	1
IGHV	Mutated	170 (69.1)	12 (80.0)	4 (44.4)	9 (56.2)
	Unmutated	76 (30.9)	3 (20.0)	5 (55.6)	7 (43.8)
	Inconclusive	3	0	1	0
FISH	No del(17p)	248 (99.6)	15 (100.0)	10 (100.0)	10 (62.5)
	Del(17p)	1 (0.4)	0 (0.0)	0 (0.0)	6 (37.5)

B) Patient characteristics according to IGHV mutational status in newly diagnosed patients

Variable		Mutated, N (%)	Unmutated, N (%)	Inconclusive, N (%)
		IGHV mutational status (N=290)		
Age	>65 years	74 (38.5)	43 (47.8)	3 (37.5)
	<65 years	118 (61.5)	47 (52.2)	5 (62.5)
Binet	A	166 (91.2)	59 (70.2)	7 (87.5)
	B/C	16 (8.8)	25 (29.8)	1 (12.5)
	Unknown	10	6	0
B2M	<4.0 mg/L	148 (89.7)	58 (78.4)	7 (100.0)
	>4.0 mg/L	17 (10.3)	16 (21.6)	0 (0.0)
	Unknown	27	16	1
TP53 status	TP53wt	166 (86.5)	75 (83.3)	7 (87.5)
	TP53mut <1%	12 (6.2)	3 (3.3)	0 (0.0)
	TP53mut 1-10%	4 (2.1)	5 (5.6)	1 (12.5)
	TP53mut >10%	6 (3.1)	4 (4.4)	0 (0.0)
	Del(17p)	4 (2.1)	3 (3.3)	0 (0.0)

Table S7. Patient characteristics in newly diagnosed patients stratified according to *TP53* (A) and IGHV mutational status (B). *TP53* wild-type (TP53wt), *TP53* mutation without del(17p) (TP53mut), del(17p) regardless of TP53 mutational status (Del(17p)), beta-2-microglobulin (B2M), fluorescent in situ hybridization (FISH).

A) Patient characteristics according to TP53 mutational status at time of treatment

Variable		TP53wt (N=44)	TP53 mutated (N=17)		
			VAF<1%, N (%)	VAF 1-10%, N (%)	VAF>10%, N (%)
Age	<65 years	27 (61.4)	1 (25.0)	3 (50.0)	4 (57.1)
	>65 years	17 (38.6)	3 (75.0)	3 (50.0)	3 (42.9)
Binet	A	13 (29.5)	3 (75.0)	2 (33.3)	2 (28.6)
	B/C	31 (70.5)	1 (25.0)	4 (66.7)	5 (71.4)
B2M	<4.0 mg/L	21 (77.8)	4 (100.0)	1 (50.0)	1 (25.0)
	>4.0 mg/L	6 (22.2)	0 (0.0)	1 (50.0)	3 (75.0)
	Unknown	17	0	4	3
IGHV	Mutated	15 (36.6)	1 (25.0)	1 (20.0)	0 (0.0)
	Unmutated	26 (63.4)	3 (75.0)	4 (80.0)	7 (100.0)
	Inconclusive	3	0	1	0
FISH	No del(17p)	43 (97.7)	4 (100.0)	5 (83.3)	3 (42.9)
	Del(17p)	1 (2.3)	0 (0.0)	1 (16.7)	4 (57.1)

B) Patient characteristics according to IGHV mutational status at time of treatment

Variable		IGHV mutational status (N = 61)		
		Mutated, N (%)	Unmutated, N (%)	Inconclusive, N (%)
Age	>65 years	8 (44.4)	27 (64.3)	4 (66.7)
	<65 years	10 (55.6)	15 (35.7)	2 (33.3)
Binet	A	6 (37.5)	13 (37.1)	3 (60.0)
	B/C	10 (62.5)	22 (62.9)	2 (40.0)
	Unknown	2	7	1
B2M	<4.0 mg/L	9 (75.0)	16 (69.6)	2 (100.0)
	>4.0 mg/L	3 (25.0)	7 (30.4)	0 (0.0)
	Unknown	6	19	4
TP53 status	TP53wt	16 (88.9)	25 (59.5)	4 (66.7)
	TP53mut <1%	1 (5.6)	3 (7.1)	0 (0.0)
	TP53mut 1-10%	1 (5.6)	5 (11.9)	1 (16.7)
	TP53mut >10%	0 (0.0)	3 (7.1)	0 (0.0)
	Del(17p)	0 (0.0)	6 (14.3)	1 (16.7)

Table S8. Patient characteristics at time of treatment stratified according to *TP53* (A) and IGHV mutational status (B). *TP53* wild-type (TP53wt), *TP53* mutation without del(17p) (TP53mut), del(17p) regardless of *TP53* mutational status (Del(17p)), beta-2-microglobulin (B2M), fluorescent in situ hybridization (FISH).

Supplemental figure legends

Figure S1. Consort diagram of patients included in the study. Newly diagnosed patients and patients sampled at time of treatment were collected within 200 days of a diagnostic flow cytometry and up to 200 days before treatment, respectively.

Figure S2. Establishing the limit of detection and range of dilution grade. (A) By serial 10-fold dilutions of four Sanger validated *TP53* mutations, all four mutations could be detected down to an adjusted variant allele frequency (VAF) of 0.016% (unadjusted 0.023% VAF). Variants were not detected at the lowest dilution level in two samples expected to fall below a VAF of 0.01%. A previously unknown low burden mutation was also identified (CLL617 g.17:7578394 T>A). (B) To test the range of the dilution step and indel detection capability, four Sanger validated *TP53* indels were initially diluted to low burden levels (1:10) and subsequently diluted 1:3 and 1:10.

Figure S3. Determining a dilution match. (A) Scatter plot of dilution grades (DGs) and adjusted dilution ratios (aDRs) for variants identified in both undiluted and diluted samples from the same patient. The DGs and aDRs were expected to match a line with a slope of one (solid line), however a cluster of aDRs with distances from a slope of one (DFSO) above 0.5 (thin line) was also identified. Notice the distribution of DGs in tilted box plot (top). (B) Histogram of DFSO for all variants identified in both undiluted and diluted samples from the same patient. The 99.9% reference range (striped lines) for variants with a DFSO less than 0.5 (solid line) determined a dilution match in which variants were considered true mutations. The threshold was trimmed to an absolute DFSO of 0.16 (dotted lines) based on the 99.9% reference range.

Figure S4. Modeling stereotypic errors. (A) The distribution of single nucleotide variants (SNV) showed enrichment of errors for C:G>A:T as well as C:G deletions. (B) For each position specific nucleotide change (e.g. g.17:7578406C>T) called 20 times or more, the distribution of variant allele frequencies (VAF) could be modeled to fit gamma distributions (red curve). (C) SNVs with less than 20 position specific occurrences showed a more even distribution of A, T, C, and G with very few deletions left to model. (D) Regardless of the genomic position, low occurring position specific variants (<20) were pooled according to their unique nucleotide change (e.g. A>G) and VAFs were modeled using best fitted gamma distribution. Variants above the overall (red line) and modeled limits of detection (striped red line) were called as true mutations.

Figure S5. Limit of detection for stereotypic error model. Multiple boxplot exemplifying the stereotypic error modeling (SEM). (A) Displaying only the first 100 bp 5' of *TP53* exon 4, the limits of detection (LODs) for position specific nucleotide change allowed variant calls at various thresholds (dotted line) according to the hg19 position and nucleotide change. Variants below Bonferroni corrected LODs (horizontal bars) and overall LOD (i.e. VAF <0.2%) were excluded (open circle), while variants above both modeled and overall LODs were called as true mutations (solid circle). Outliers above the modeled LODs but below the overall LOD indicated (open triangle). (B) For unique nucleotide changes (only SNVs are shown), box plots demonstrate that all modeled LODs (horizontal bar) are below the overall LOD (red line).

Figure S6. Schematic workflow of bioinformatic pipeline. The combined variant call format (VCF) files from a low frequency detection tool and a basic detection tool in CLC Genomics Workbench (Qiagen) was analyzed in R using both a dilution match algorithm (DMA) and a

stereotypic error model (SEM). For DMA, the number of total variants is indicated (N), while the number of unique variants (N) and number of unique variants modeled (n) is indicated for SEM. Variants were further cross referenced in the IARC *TP53* database. Only true positive mutations from DMA and SEM were included in the final call.

Figure S7. Stereotypic error distribution of problematic variants. (A) g.17:7577090C>T (p.Arg283His) showed a distribution with a wide right tail. Stereotypic error modeling (SEM) identified 45 false positive mutations for this variant between 0.2-0.65% VAF. (B) At the same position, g.17:7577089_7577090GC>- (p.Arg283fs*22) was called in three patients using the dilution match algorithm (DMA), but excluded using SEM. (C) With only 24 observations, SEM excluded a ddPCR validated g.17:7578442T>C (p.Tyr163Cys) with VAF of 0.23% (red arrow) that was called by the dilution match algorithm (DMA). (D) g.17:77579837A>C (c.74+2T>G) splicing site mutation shows a multimodal distribution unfit for modeling. This variant was called by CLC in 116 undiluted samples with a VAF up to 1.7% but was excluded by both DMA and SEM. Best fitted gamma distribution (red curves), modeled position specific limit of detection (LOD; dotted red line), and overall LOD of 0.2% VAF (red line).

Figure S8. Subanalysis of overall and treatment-free survival in newly diagnosed patients. Kaplan Meier curves comparing (A-B) overall (OS) and (C-D) treatment-free survival (TFS) based on *TP53* status. Patients with *TP53*mut 1-10% VAF and *TP53*ab >10% VAF shown separately (A, C). Patients with *TP53*mut 1-10% VAF and *TP53*mut >10% VAF shown separately (B, D). Significant difference in TFS between *TP53*ab >10% VAF and *TP53* wild-type (*TP53*wt) was demonstrated. *P*-values indicated in tables.

Figure S1.

Figure S2.

A Serial 10–fold dilution of TP53 mutations
$R^2 = 0.9967$

SampleID/TP53 mutation
- DIL1 g.17:7577538 C>T
- DIL2 g.17:7578190 T>C
- DIL3 g.17:7578414 A>–
- DIL3 g.17:7578417..7578424 CTCCGTCA>–
- DIL4 g.17:7577547 C>T
- DIL4 g.17:7578394 T>A

Variant allele frequency (%)

Dilution grade (%)

B Dilution of known low burden TP53 indels
$R^2 = 0.9985$

SampleID/TP53 mutation
- BC1 g.17:7576868^7576869 –>TC
- BC147 g.17:7576900^7576901 –>A
- BC77 g.17:7574003 G>–
- BC96 g.17:7576928 T>–

Variant allele frequency (%)

Dilution grade (%)

Figure S3.

A

Distribution of dilution grades and adjusted dilution ratios for variants in both undiluted and diluted sample from the same patient

B Distance from slope of one

Determing the threshold for correct dilution using the distance from a line with slope of one

Figure S4.

A Distribution of SNV modeled for position specific nucleotide change

B Stereotypic error distribution and fitted gamma distribution
Position specific nucleotide change for g.7578518C>T (hg19)
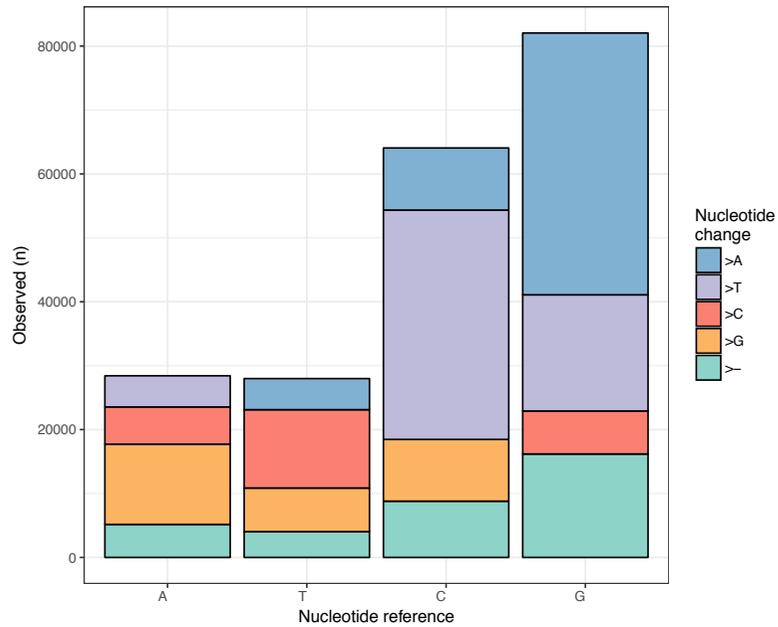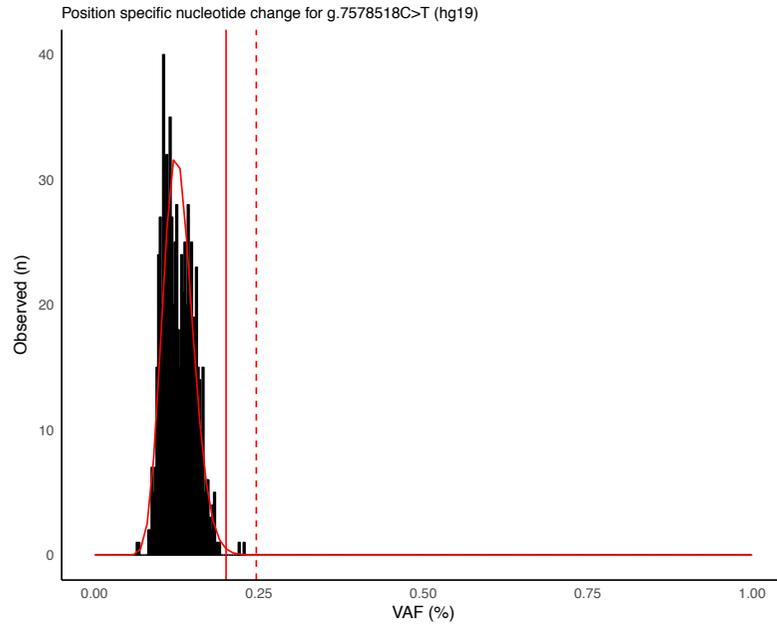
C Distribution of SNV modeled for specific nucleotide change

D Stereotypic error distribution and fitted gamma distribution
Unique nucleotide change for G>A
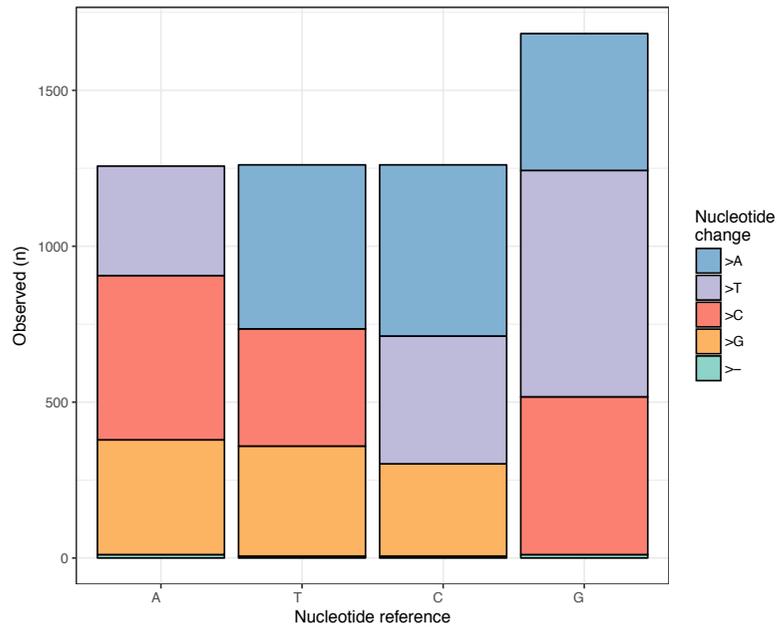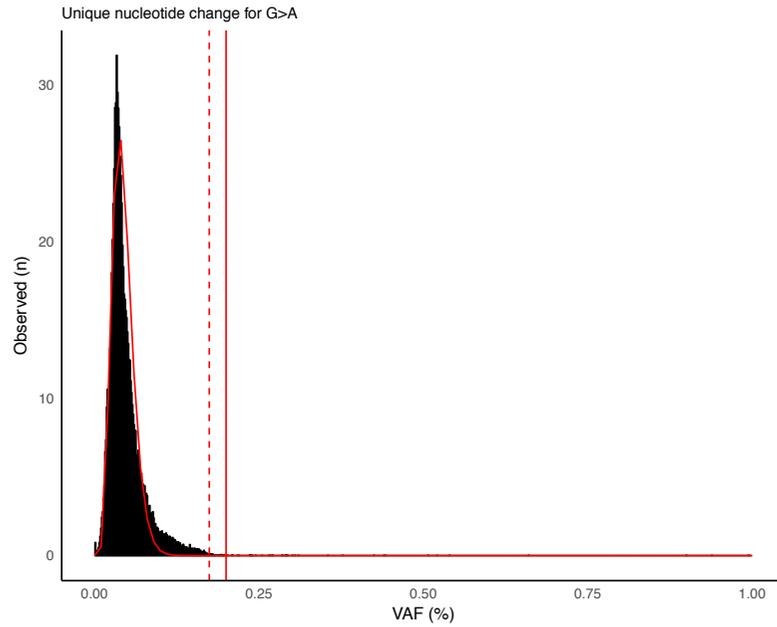
Figure S5.

A Thresholds for position specific nucleotide change
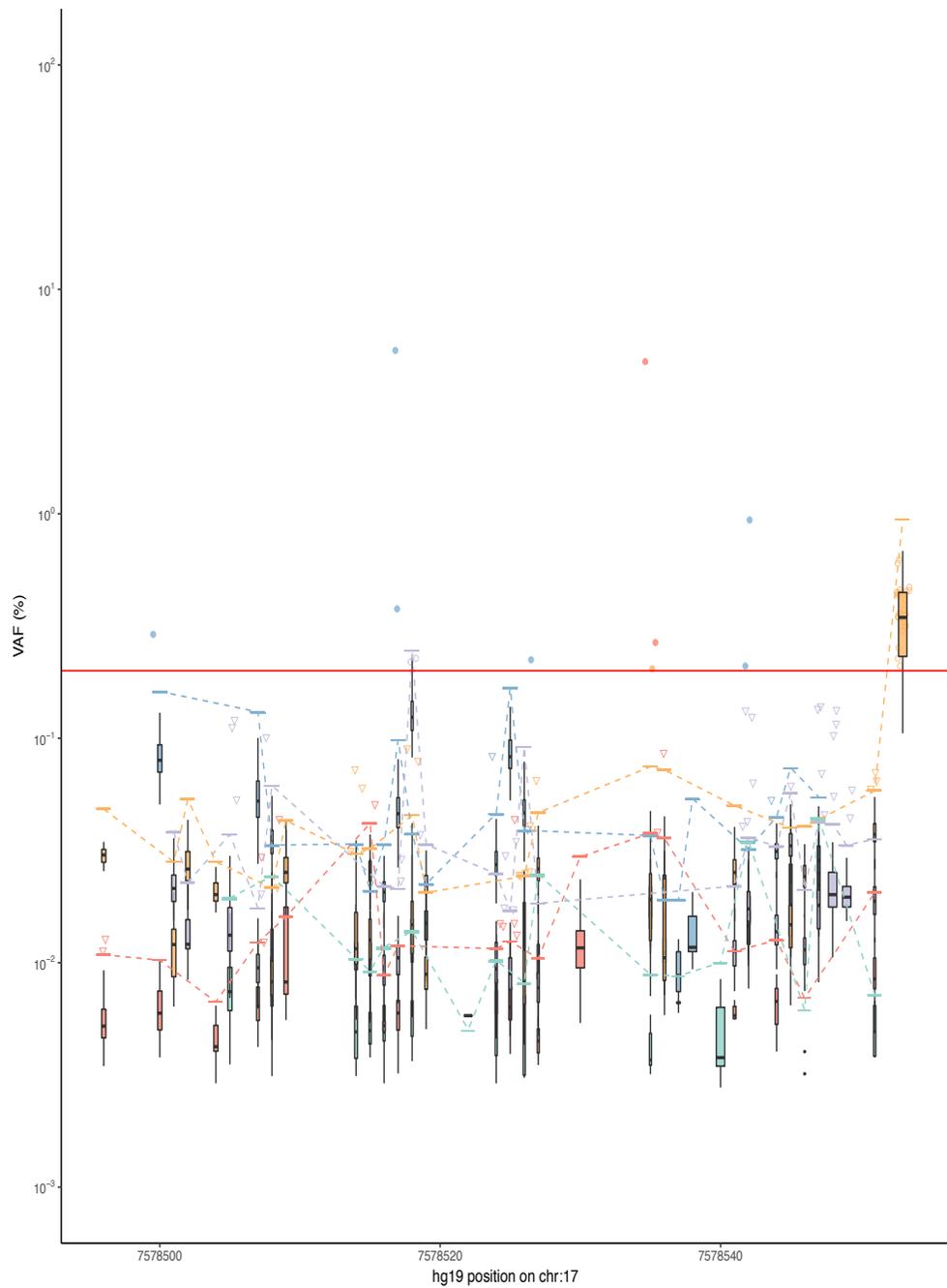
B Thresholds for nucleotide specific change
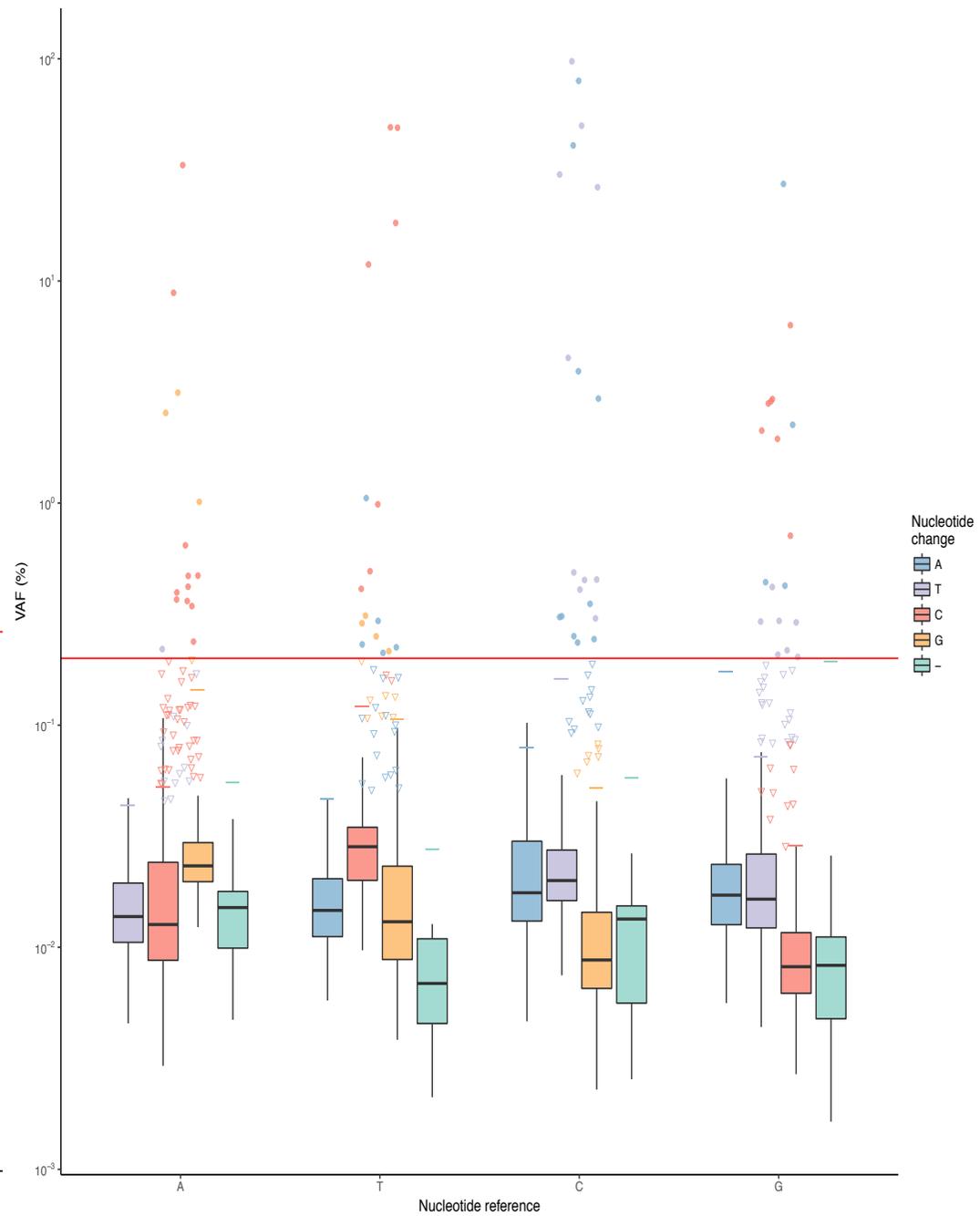
Figure S6.
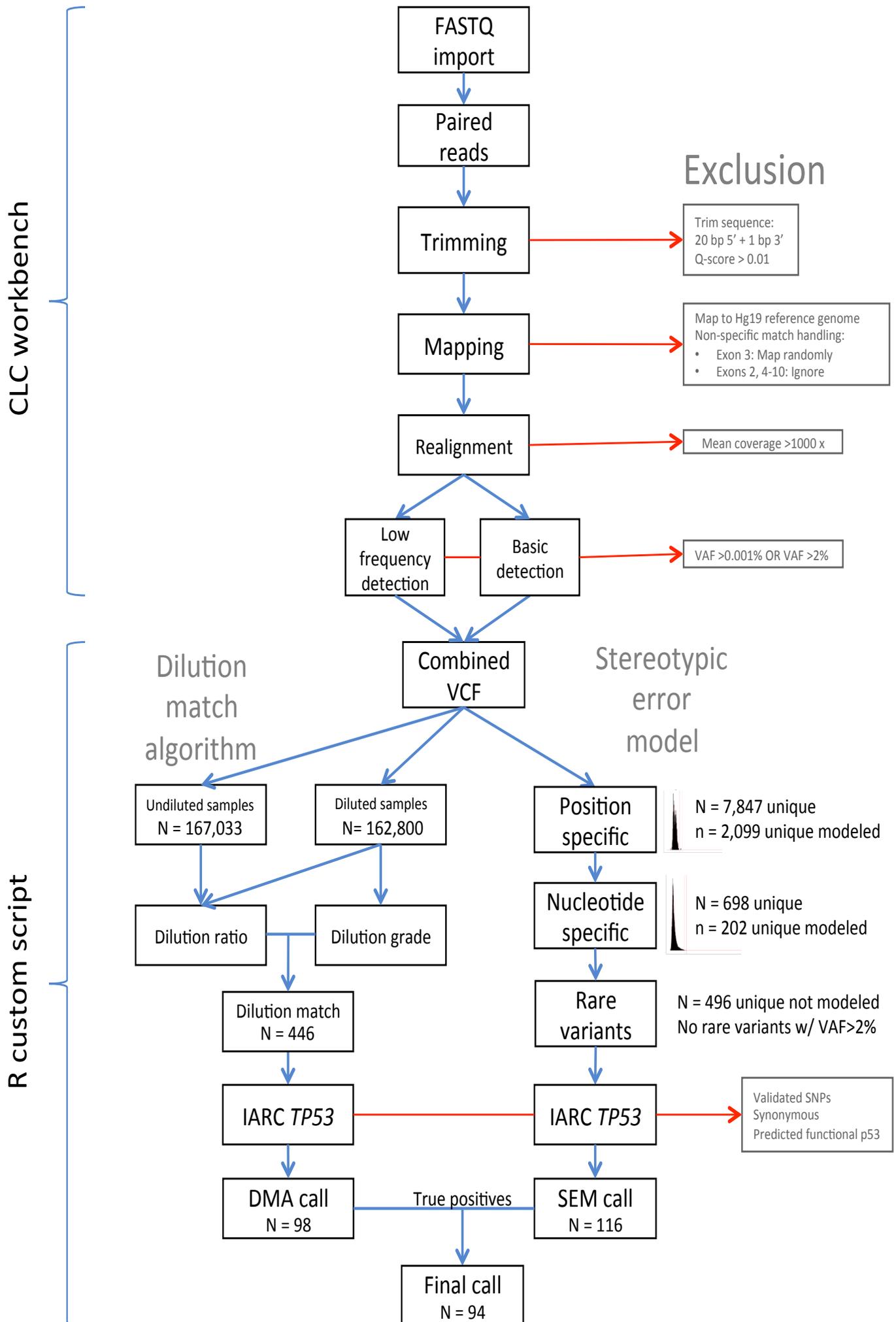
Figure S7.

A Stereotypic error distribution and fitted gamma distribution

Position specific nucleotide change for hg19 g.17:7577090C>T/p.Arg283His
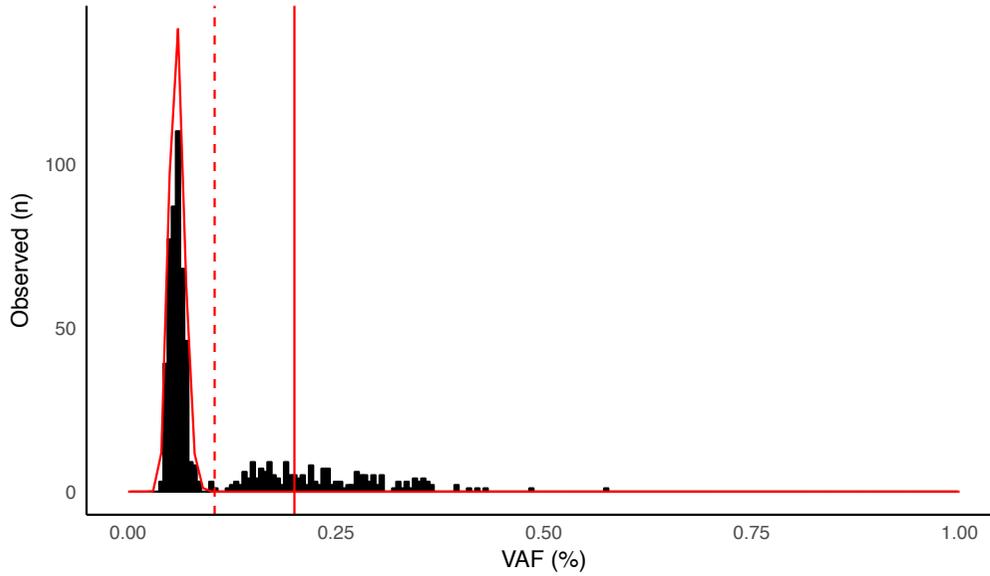
B Stereotypic error distribution and fitted gamma distribution

Position specific nucleotide change for hg19 g.17:7577089_7577090delGC/p.Arg283fs*22
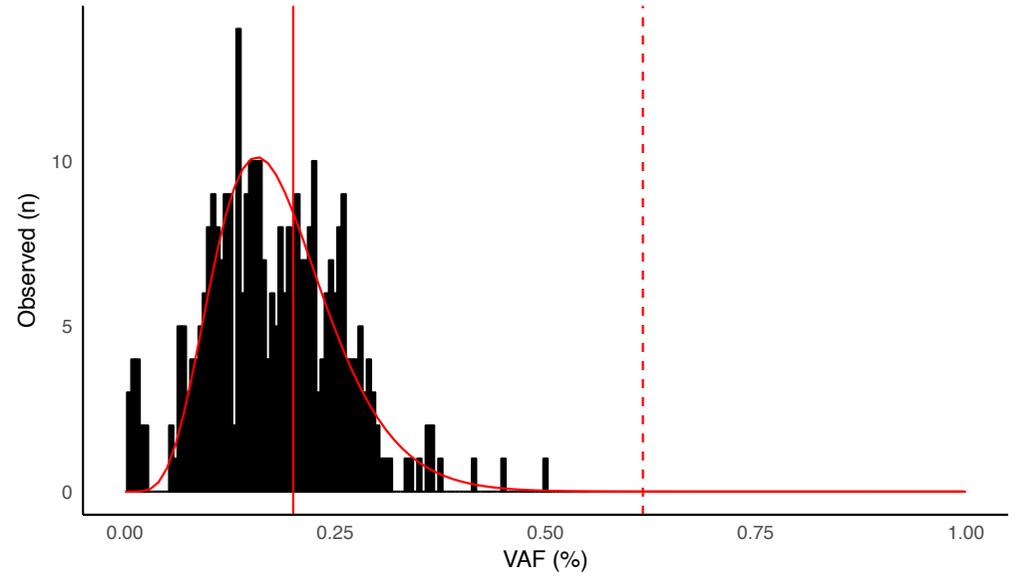
C Stereotypic error distribution and fitted gamma distribution

Position specific nucleotide change for hg19 g.17:7578442T>C/p.Tyr163Cys
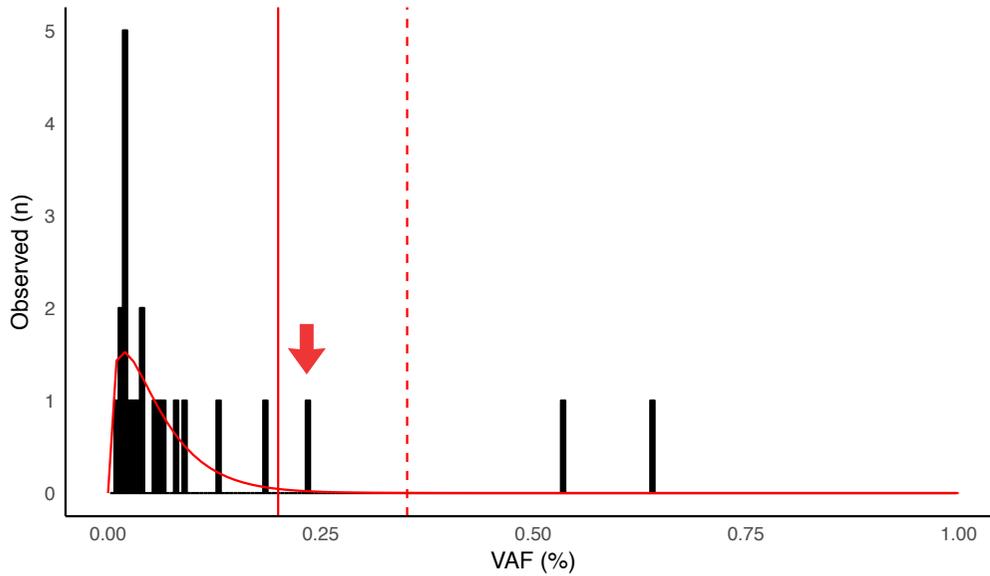
D Stereotypic error distribution and fitted gamma distribution

Position specific nucleotide change for hg19 g.17:77579837A>C/c.74+2T>G

Figure S8.

OS

TFS

TP53wt   TP53mut <1%   TP53ab 1–100%

A

248	223	149	90	43
15	15	11	5	4
27	21	13	5	1

No. patients

B

248	176	104	62	29
15	11	9	5	4
27	13	8	2	1

No. patients

TP53wt   TP53mut <1%   TP53mut 1–100% w/o del(17p)   Del(17p)

C

248	223	149	90	43
15	15	11	5	4
20	17	12	5	1
7	4	1	0	0

No. patients

D

248	176	104	62	29
15	11	9	5	4
20	11	8	2	1
7	2	0	0	0

No. patients