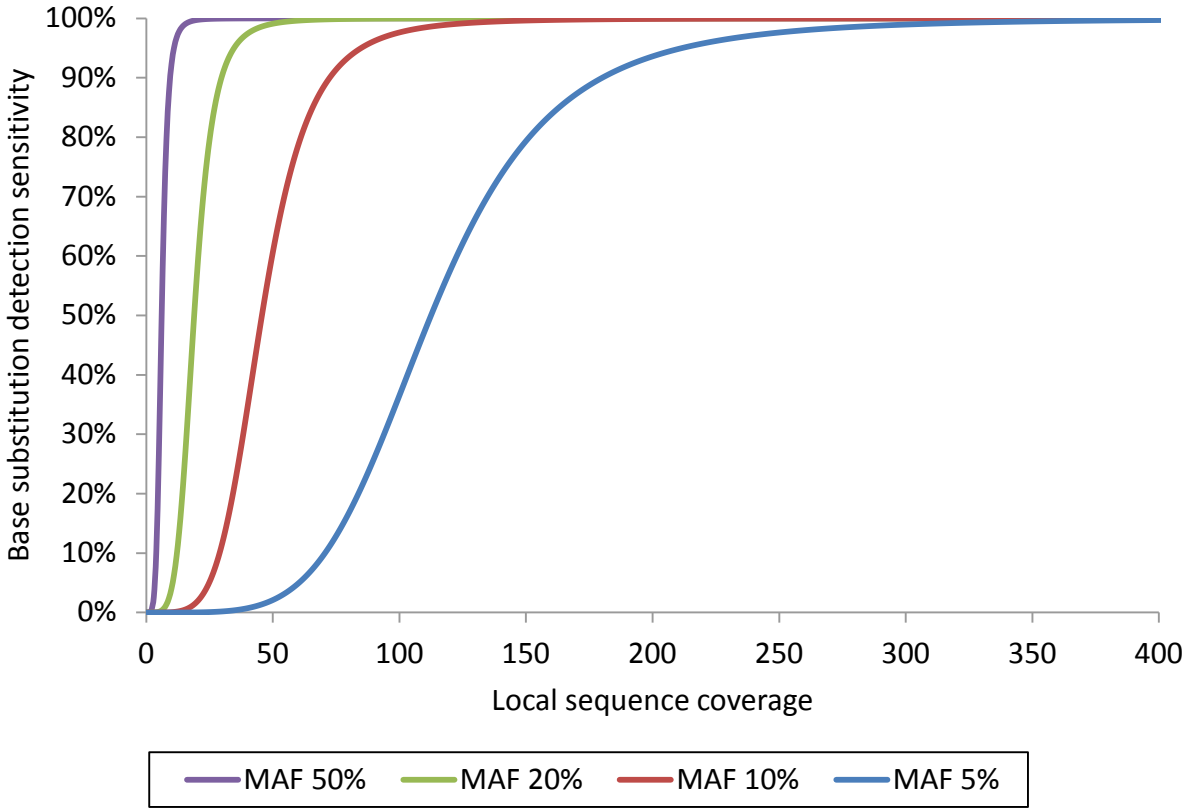
SUPPLEMENTARY INFORMATION

Frampton et al.

Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing

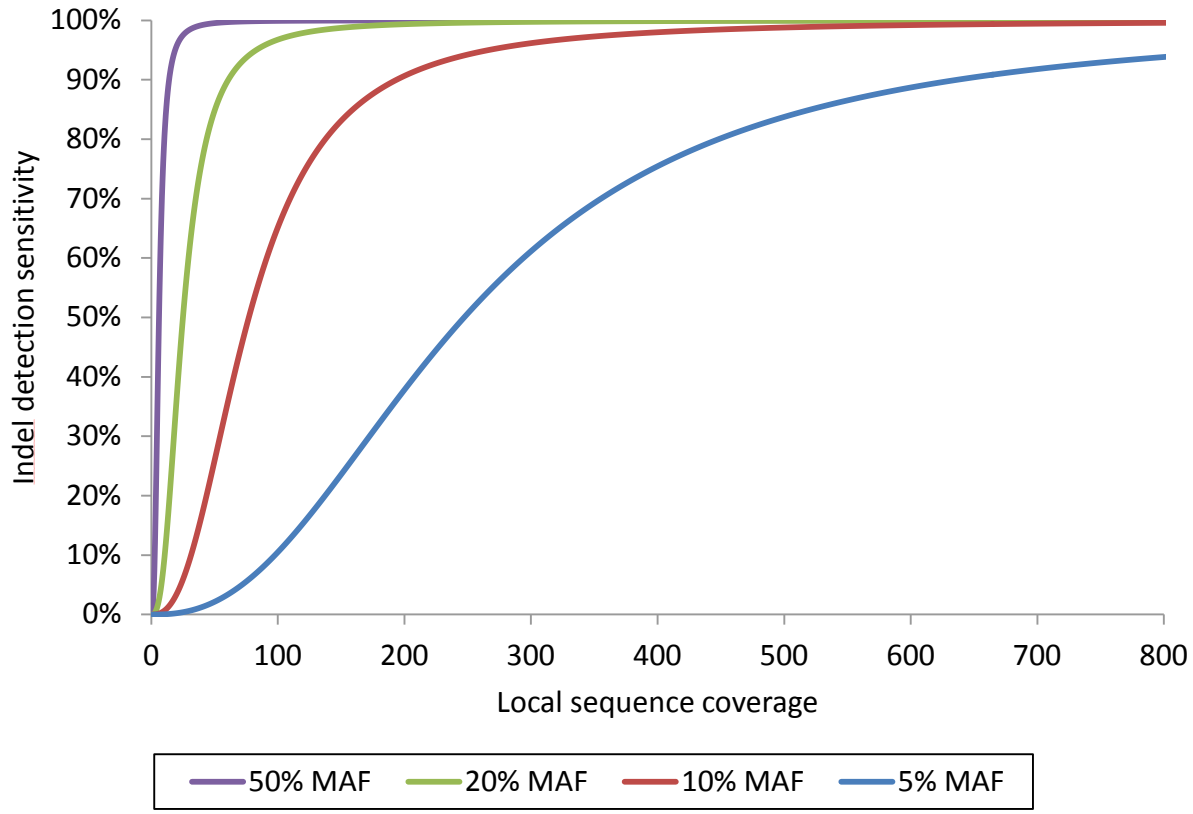
Nature Biotechnology: doi:10.1038/nbt.2696

Supplementary Figure 1A
Base substitution detection sensitivity as a function of coverage and mutant allele frequency



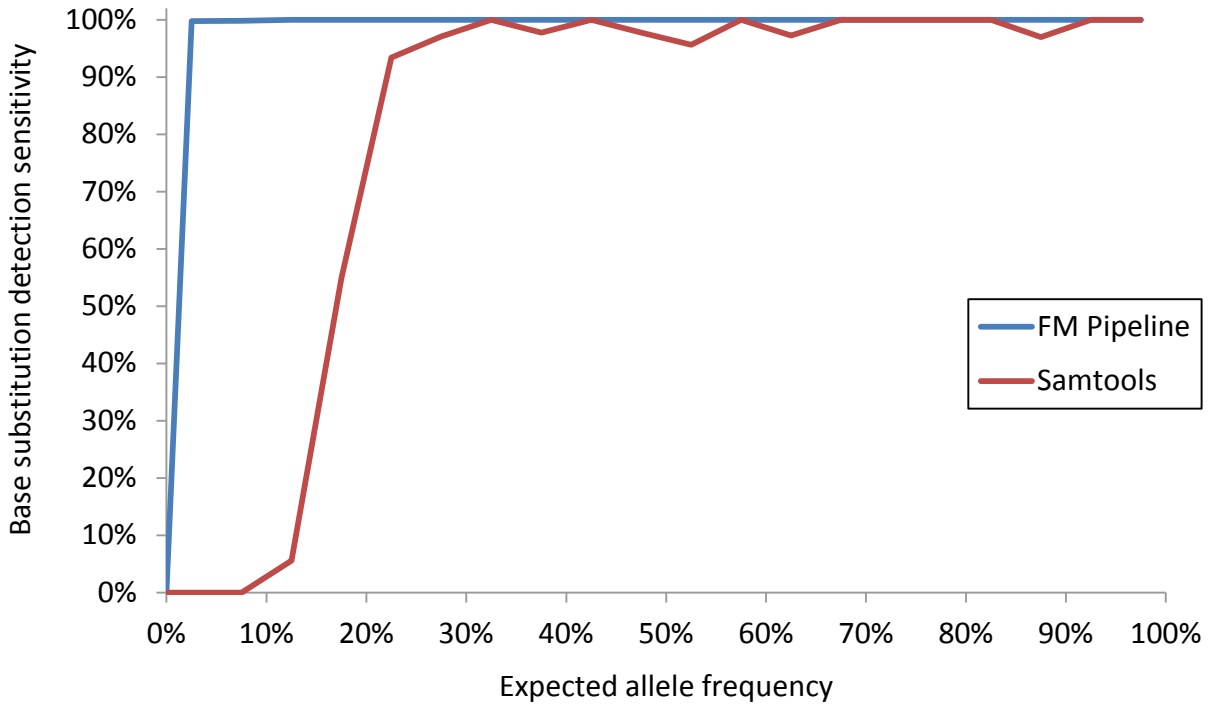
Supplementary Figure 1A caption:
A smoothed representation base substitution detection sensitivity as a function of local sequence coverage was generated based on test set calls in all down-sample datasets. Expected detection sensitivity was plotted at mutant allele frequencies 50%, 20%, 10%, and 5%.

Supplementary Figure 1B
Indel detection sensitivity as a function of coverage and mutant allele frequency



Supplementary Figure 1B caption:
A smoothed representation indel detection sensitivity as a function of local sequence coverage was generated based on test set calls in all down-sample datasets. Expected detection sensitivity was plotted at mutant allele frequencies 50%, 20%, 10%, and 5%.

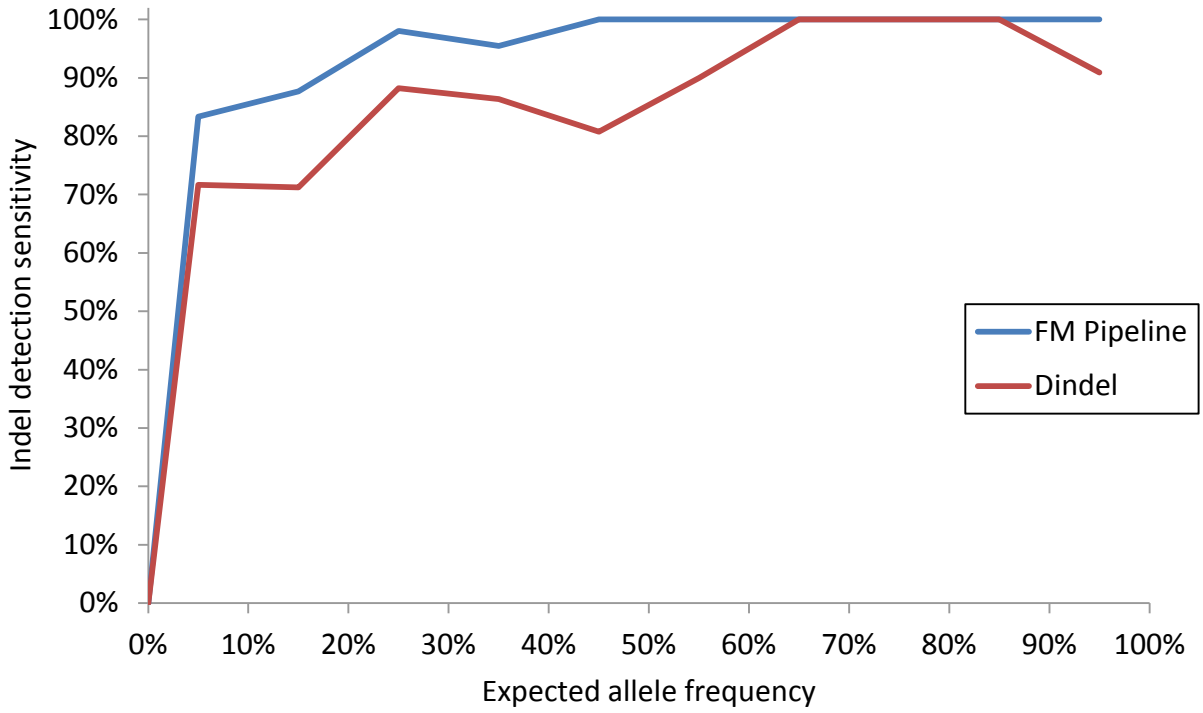
Supplementary Figure 2A
Base substitution detection sensitivity for the Samtools algorithm compared to the FM pipeline



Supplementary Figure 2A caption:
Detection sensitivity for base substitutions was examined for the two normal cell-line pools (pool XX and pool YY) using the FM pipeline and Samtools 0.1.16. For Samtools, the variant calling procedure described at the Samtools homepage (http://samtools.sourceforge.net/mpileup.shtml) was used, with the commands:

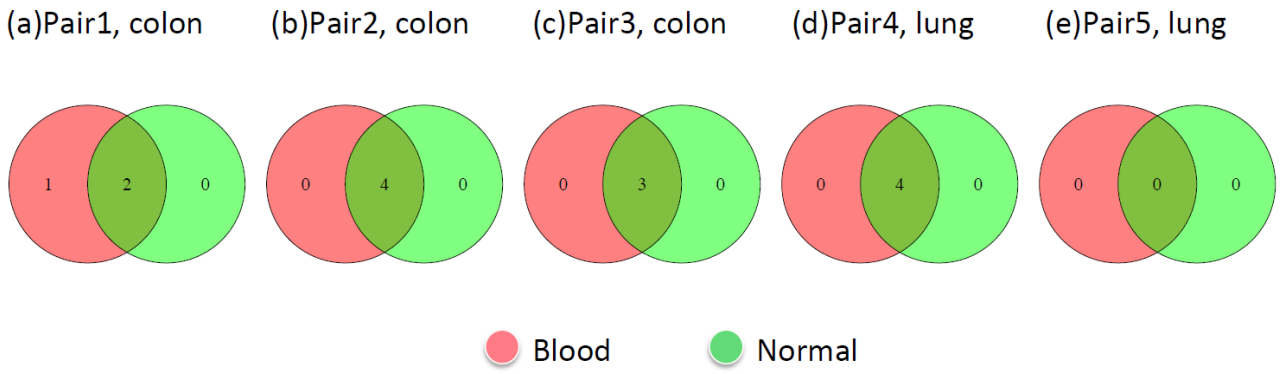
- samtools mpileup -uBD -Q 30 -q 20 -l <exon regions> -f hg19.fa <sampleX.bam> | bcftools view -bv -> <outputFile.bcf>
- bcftools view <outputFile.bcf> | vcfutils.pl varFilter > outputFile.flt.vcf

Supplementary Figure 2B
Indel detection sensitivity for the Dindel algorithm compared to the FM pipeline



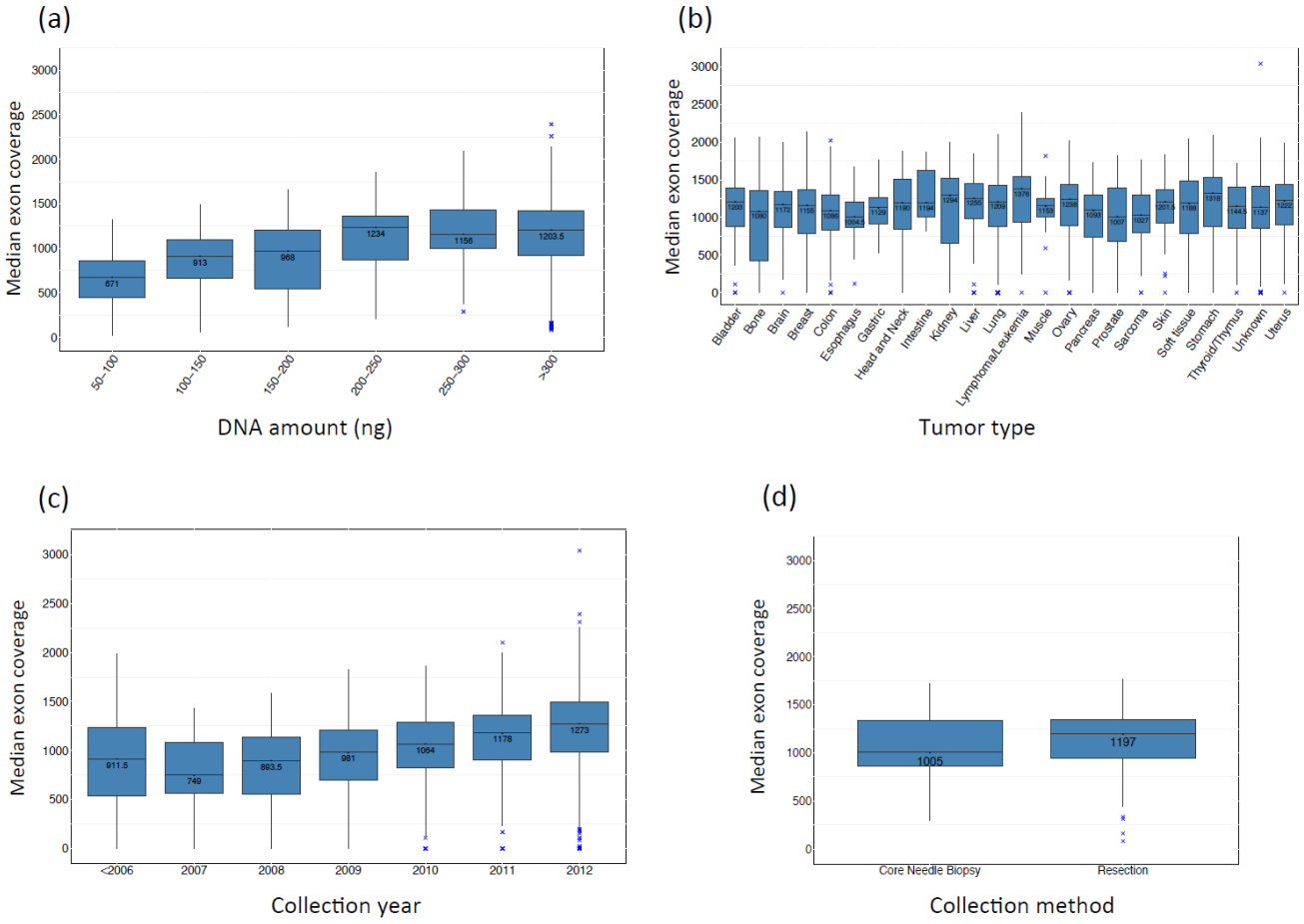
Supplementary Figure 2B caption:
Detection sensitivity for indel calls were generated for pure tumor cell lines and pools using the FM pipeline and Dindel 1.01. For Dindel, the four-step procedure for calling indels in pooled samples outlined in the Dindel 1.01 manual (ftp://ftp.sanger.ac.uk/pub4/resources/software/dindel/manual-1.01.pdf) was used with default settings. The "outside coverage range" variant filter was not applied.

Supplementary Figure 3
Comparison of variant calls in paired blood and FFPE normal samples



Supplementary Figure 3 caption:
Five pairs of blood and FFPE normal samples were used to compare pipeline variant calls. Total number of calls (post-filtering) is shown with overlap: calls in blood in red; calls in normal FFPE in green. FFPE tissue from pairs 1, 2, and 3 are colon; pairs 4 and 5 are lung.

Supplementary Figure 4
Impact of pre-analytic parameters on assay performance



Supplementary Figure 4 caption:
Assay performance, as represented by median unique exon sequencing coverage, is shown for sub-groups of clinical patient samples. The upper and lower box edges denote the first and third quartiles, medians are represented by solid black lines in the boxes, and median numbers were showed below the lines. Whiskers extend to the largest and smallest data within 1.5 times of the interquartile ranges. Outliers, if any, are marked individually as blue crosses out of box.
(a) Median exon coverage stratified by assay DNA input amount (ng).
(b) Median exon coverage stratified by tumor type.
(c) Median exon coverage stratified by sample collection year.
(d) Median exon coverage stratified by two primary sample collection methods, where annotated.

Supplementary Table 1A
Genes targeted in hybridization capture (current test)

ABL1	BRIP1	CHEK2	FANCA	GNAQ	MAP2K1	NFKBIA	PPP2R1A	SMARCD1
AKT1	BTG1	CHUK	FANCC	GNAS	MAP2K2	NKX2-1	PRDM1	SMO
AKT2	BTK	CIC	FANCD2	GPR124	MAP2K4	NOTCH1	PRKAR1A	SOCS1
AKT3	C17orf39	CRBN	FANCE	GSK3B	MAP3K1	NOTCH2	PRKDC	SOX10
ALK	CARD11	CREBBP	FANCF	HGF	MAP3K13	NOTCH3	PRSS8	SOX2
ALOX12B	CASP8	CRKL	FANCG	HLA-A	MCL1	NOTCH4	PTCH1	SPEN
APC	CBFB	CRLF2	FANCI	HRAS	MDM2	NPM1	PTEN	SRC
APCDD1	CBL	CSF1R	FANCL	IDH1	MDM4	NRAS	PTPN11	STAG2
AR	CCND1	CTCF	FANCM	IDH2	MED12	NSD1	RAD50	STAT4
ARAF	CCND2	CTNNB1	FAT3	IGF1	MEF2B	NTRK1	RAD51	STK11
ARFRP1	CCND3	CUL4A	FBXW7	IGF1R	MEN1	NTRK2	RAD51B	SUFU
ARID1A	CCNE1	CUL4B	FGF10	IGF2	MET	NTRK3	RAD51C	SYK
ARID2	CD79A	CYP17A1	FGF12	IKBKE	MITF	NUP93	RAD51D	TBX3
ASXL1	CD79B	DAXX	FGF14	IKZF1	MLH1	PAK3	RAD52	TET2
ATM	CDC73	DDR2	FGF19	IL7R	MLL	PAK7	RAD54L	TGFBR2
ATR	CDH1	DIS3	FGF23	INHBA	MLL2	PALB2	RAF1	TIPARP
ATRX	CDK12	DNMT3A	FGF3	IRF4	MPL	PARP1	RARA	TNFAIP3
AURKA	CDK4	DOT1L	FGF4	IRS2	MRE11A	PARP2	RB1	TNFRSF14
AURKB	CDK6	EGFR	FGF6	JAK1	MSH2	PARP3	REL	TOP1
AXL	CDK8	EMSY	FGF7	JAK2	MSH6	PARP4	RET	TP53
BACH1	GRIN2A	EP300	FGFR1	JAK3	MTOR	PAX5	RICTOR	TRRAP
BAP1	SPOP	EPHA3	FGFR2	JUN	MUTYH	PDGFRA	RNF43	TSC1
BARD1	SETD2	EPHA5	FGFR3	KDM5A	MYC	PDGFRB	RPA1	TSC2
BCL2	PBRM1	EPHB1	FGFR4	KDM5C	MYCL1	PDK1	RPTOR	TSHR
BCL2L2	FOXL2	ERBB2	FLT1	KDM6A	MYCN	PIK3C2G	RUNX1	VHL
BCL6	CTNNA1	ERBB3	FLT3	KDR	MYD88	PIK3C3	RUNX1T1	WISP3
BCOR	CDKN1B	ERBB4	FLT4	KEAP1	MYST3	PIK3CA	SF3B1	WT1
BCORL1	CDKN2A	ERG	GATA1	KIT	NBN	PIK3CG	SH2B3	XPO1
BLM	CDKN2B	ESR1	GATA2	KLHL6	NCOR1	PIK3R1	SMAD2	XRCC3
BRAF	CDKN2C	EZH2	GATA3	KRAS	NF1	PIK3R2	SMAD4	ZNF217
BRCA1	CEBPA	FAM123B (WTX)	GNA11	LMO1	NF2	PMS2	SMARCA4	ZNF703
BRCA2	CHEK1	FAM46C	GNA13	LRP1B	NFE2L2	PNRC1	SMARCB1	

Genes targeted for rearrangement detection

ALK	BRAF	ETV1	ETV5	EWSR1	MYC	PDGFRA	RARA	ROS1
BCL2	EGFR	ETV4	ETV6	MLL	NTRK1	RAF1	RET	TMPRSS2
BCR								

Supplementary Table 1B
Genes targeted in hybridization capture (earlier test)

ABL1	BRAF	CEBPA	ERG	HSP90AA1	MAP2K1	NF2	PTEN	SUFU
ABL2	BRCA1	CHEK1	ESR1	IDH1	MAP2K2	NKX2-1	PTPN11	TBX22
AKT1	BRCA2	CHEK2	EZH2	IDH2	MAP2K4	NOTCH1	PTPRD	TET2
AKT2	CARD11	CRKL	FANCA	IGF1R	MCL1	NPM1	RAF1	TGFBR2
AKT3	CBL	CRLF2	FBXW7	IGF2R	MDM2	NRAS	RARA	TNFAIP3
ALK	CCND1	CTNNB1	FGFR1	IKBKE	MDM4	NTRK1	RB1	TNKS
APC	CCND2	DDR2	FGFR2	IKZF1	MEN1	NTRK2	RET	TNKS2
AR	CCND3	DNMT3A	FGFR3	INHBA	MET	NTRK3	RICTOR	TOP1
ARAF	CCNE1	DOT1L	FGFR4	INSR	MITF	PAK3	RPTOR	TP53
ARFRP1	CD79A	EGFR	FLT1	IRS2	MLH1	PAX5	RUNX1	TSC1
ARID1A	CD79B	EPHA3	FLT3	JAK1	MLL	PDGFRA	SMAD2	TSC2
ATM	CDH1	EPHA5	FLT4	JAK2	MPL	PDGFRB	SMAD3	USP9X
ATR	CDH2	EPHA6	FOXP4	JAK3	MRE11A	PHLPP2	SMAD4	VHL
AURKA	CDH20	EPHA7	GATA1	JUN	MSH2	PIK3CA	SMARCA4	WT1
AURKB	CDH5	EPHB1	GNA11	KDM6A	MSH6	PIK3CG	SMARCB1	
BAP1	CDK4	EPHB4	GNAQ	KDR	MTOR	PIK3R1	SMO	
BCL2	CDK6	EPHB6	GNAS	KIT	MUTYH	PKHD1	SOX10	
BCL2A1	CDK8	ERBB2	GPR124	KRAS	MYC	PLCG1	SOX2	
BCL2L1	CDKN2A	ERBB3	GUCY1A2	LRP1B	MYCL1	PRKDC	SRC	
BCL2L2	CDKN2B	ERBB4	HOXA3	LRP6	MYCN	PTCH1	STAT3	
BCL6	CDKN2C	ERCC2	HRAS	LTK	NF1	PTCH2	STK11	

Genes targeted for rearrangement detection

ALK	BRAF	ETV1	ETV5	EWSR1	RAF1	RARA	RET	TMPRSS2
BCR	EGFR	ETV4	ETV6	MLL				

Supplementary Table 1 notes:
Genes are identified by HUGO nomenclature. The total gene count for the current test (Supplementary Table 1A) is 295 genes and the total gene count for the earlier test (Supplementary Table 1B) is 189 genes.

Supplementary Table 2A
Summary of cell-lines used for assessment of base substitution detection

Cell line	Pool	Number of known heterozygous SNP sites	Number of known homozygous SNP sites
NA07014	pool XX	210	102
NA10840	pool XX	195	117
NA18595	pool XX	248	127
NA18957	pool XX	236	123
NA18488	pool XX	183	123
NA18511	pool XX	259	120
NA18867	pool XX	276	128
NA18924	pool XX	160	156
NA19108	pool XX	243	121
NA19114	pool XX	235	122
NA12386	pool YY	189	110
NA12767	pool YY	187	124
NA18618	pool YY	200	137
NA19059	pool YY	161	149
NA19147	pool YY	272	119
NA19179	pool YY	257	112
NA19190	pool YY	249	131
NA19214	pool YY	229	124
NA19235	pool YY	238	138
NA19247	pool YY	223	156

Supplementary Table 2B

Summary of sequencing metrics for cell-lines and pools used for assessment of base substitution detection

Normal HapMap cell line	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
NA07014	19,124,723	92%	7,364,455	360	100%	0.34%
NA10840	16,332,031	92%	6,509,830	319	100%	0.34%
NA18595	17,422,928	92%	6,849,036	335	100%	0.35%
NA18957	23,172,522	91%	8,979,557	438	100%	0.27%
NA18488	23,980,245	91%	8,908,221	438	100%	0.36%
NA18511	22,972,561	92%	8,569,119	416	100%	0.35%
NA18867	18,482,783	92%	7,252,368	357	100%	0.35%
NA18924	17,304,309	91%	6,698,243	331	100%	0.35%
NA19108	22,123,874	92%	8,497,021	412	100%	0.35%
NA19114	18,786,684	92%	7,388,077	354	100%	0.27%
NA12386	16,937,149	92%	6,669,594	325	100%	0.26%
NA12767	17,654,744	92%	6,935,027	343	100%	0.26%
NA18618	23,619,618	91%	9,224,735	457	100%	0.27%
NA19059	21,789,629	92%	8,203,279	407	100%	0.27%
NA19147	16,340,118	91%	6,320,632	315	100%	0.27%
NA19179	25,175,272	91%	9,471,701	466	100%	0.27%
NA19190	23,514,670	92%	9,123,334	443	100%	0.28%
NA19214	19,183,067	91%	7,434,387	365	100%	0.28%
NA19235	24,336,984	91%	9,519,772	468	100%	0.28%
NA19247	23,490,106	91%	8,827,832	420	100%	0.29%

Pool ID	Number of cell lines in pool	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
pool XX	10	51,303,128	92%	15,700,438	753	100%	0.32%
pool YY	10	35,501,862	92%	12,145,181	597	100%	0.32%

Supplementary Table 2 notes:

Pure cell lines used to establish the base substitution test set were sequenced at reduced coverage, as alterations in pure cell lines are expected to be present at high alternate allele frequencies.

Supplementary Table 3
Distribution of expected mutant allele frequencies in base substitution test set

Expected mutant allele frequency	Number of sites in pool XX	Number of sites in pool YY
< 5%	206	201
5 - 10%	314	300
10 - 15%	130	103
15 - 20%	75	78
20 - 25%	64	42
25 - 30%	41	27
30 - 35%	29	31
35 - 40%	21	24
40 - 45%	24	16
45 - 50%	20	25
50 - 55%	11	12
55 - 60%	9	11
60 - 65%	12	24
65 - 70%	12	14
70 - 75%	9	17
75 - 80%	16	9
80 - 85%	11	16
85 - 90%	21	12
90 - 95%	17	13
95 - 100%	15	25
Total	1,057	1,000

Supplementary Table 4A
Summary of base substitution detection performance (sensitivity)

Down-sampled coverage	MAF < 5% n = 407			MAF 5 - 10% n = 614			MAF > 10% n = 1,036		
	False negatives	Mean	Confidence interval	False negatives	Mean	Confidence interval	False negatives	Mean	Confidence interval
738, 580	6	98.5%	96.8 – 99.5%	2	99.7%	98.8 – 100%	0	100%	99.6 – 100%
500	16	96.1%	93.7 – 97.7%	4	99.3%	98.3 – 99.8%	0	100%	99.6 – 100%
450	20	95.1%	92.5 – 97.0%	4	99.3%	98.3 – 99.8%	0	100%	99.6 – 100%
400	24	94.1%	91.4 – 96.2%	3	99.5%	98.6 – 99.9%	0	100%	99.6 – 100%
350	43	89.4%	86.0 – 92.2%	5	99.2%	98.1 – 99.7%	0	100%	99.6 – 100%
300	67	83.5%	80.0 – 87.0%	7	98.9%	97.7 – 99.5%	0	100%	99.6 – 100%
250	104	74.4%	69.9 – 78.6%	13	97.9%	96.4 – 98.9%	1	99.9%	99.5 – 100%
200	149	63.4%	58.5 – 68.1%	36	94.1%	92.0 – 95.9%	1	99.9%	99.5 – 100%
150	227	44.2%	39.3 – 49.2%	105	82.3%	79.7 – 85.8%	5	99.5%	98.9 – 99.8%

Supplementary Table 4B
Summary of base substitution detection performance (specificity)

Down-sampled coverage	True positives	False positives		Positive predictive value	
		MAF ≥ 5%	MAF < 5%	Mean	Confidence interval
738, 580	2,577	0	2	99.9%	99.7 – 100%
500	2,564	0	0	100%	99.9 – 100%
450	2,557	0	0	100%	99.9 – 100%
400	2,550	0	0	100%	99.9 – 100%
350	2,521	0	0	100%	99.9 – 100%
300	2,488	0	0	100%	99.9 – 100%
250	2,425	0	0	100%	99.8 – 100%
200	2,341	0	0	100%	99.8 – 100%
150	2,148	0	0	100%	99.8 – 100%

Supplementary Table 4C
Summary of base substitution detection performance (full sensitivity results)

Available in online materials

Supplementary Table 4 notes:
Confidence intervals were calculated as the exact 95% binomial confidence interval. False positives are base substitution calls in pooled samples that are absent from constituent cell-line sequence data. To represent typical reportable alterations, call settings for recurrent sites of mutation were applied. Initial coverage listed (738x/580x) is full read depth obtained in pools, subsequently down-sampled in-silico.

Supplementary Table 5A
Summary of tumor cell lines and known indel alterations used for assessment of indel detection

Tumor cell line	Gene	Indel type	Indel length (bp)	Measured mutant allele frequency	Genomic coordinates (hg19)	Reference sequence	Alternate sequence
A2058	PTEN	Deletion	35	35%	chr10:89711905-89711940	GTGTATTATTATAGCTACCTG TTAAAGAATCATCTG	G
A253	CDKN2A	Deletion	13	100%	chr9:21974766-21974779	CCGCGGCCGTGGCC	C
	TP53	Deletion	1	100%	chr17:7578390-7578391	CT	C
A549	SMARCA4	Deletion	23	100%	chr19:11121116-11121139	TGCAGTCCTACTATGCCGTG GCCC	T
AGS	CDH1	Insertion	1	100%	chr16:68855925-68855925	C	CC
AsPC-1	CDKN2A	Deletion	2	97%	chr9:21971126-21971128	GAG	G
	TP53	Deletion	1	95%	chr17:7578529-7578530	AA	A
CCRF-CEM	NOTCH1	Insertion	36	47%	chr9:139399385-139399385	G	GTGGGGAAGGCGCGGCAG GAAGTGGAAGGAGCTGTTG
DU-145	STK11	Deletion	5	100%	chr19:1220438-1220443	CAAGCC	C
HCC1187	NFKBIA	Insertion	1	100%	chr14:35872476-35872476	G	GG
	TP53	Deletion	3	100%	chr17:7579362-7579365	AACC	A
HCC1395	-	-	-	-	-	-	-
HCC1599	BRCA2	Deletion	10	99%	chr13:32913041-32913051	AAAGAACCTAC	A
HCC1937	BRCA1	Insertion	1	96%	chr17:41209082-41209082	G	GG
HCT-15	APC	Deletion	1	100%	chr5:112175538-112175539	GC	G
	BRCA2	Deletion	1	44%	chr13:32913842-32913843	AA	A
	BRCA2	Deletion	2	50%	chr13:32912090-32912092	TGT	T
	MSH6	Deletion	1	49%	chr2:48025989-48025990	CC	C
J-RT3-T3-5	PTEN	Insertion	7	44%	chr10:89717675-89717675	C	CCCCATGG
	PTEN	Insertion	39	45%	chr10:89717712-89717712	C	CCTGAAGTTCATGTACTTTG AGTTCCCTCAGCCCTGGGTT
	TP53	Deletion	1	43%	chr17:7573943-7573944	TC	T
LoVo	SMO	Insertion	6	40%	chr7:128829061-128829061	G	GCTGCTG
	APC	Deletion	1	63%	chr5:112175580-112175581	CC	C
	NF1	Deletion	1	not found [a]	-	-	-
MV-4-11	FLT3	Insertion	30	64%	chr13:28608284-28608284	T	TGATCATATTCATATTCTCTG AAATCAACGT
NCI-H128	ARID1A	Deletion	1	38%	chr1:27092812-27092813	GG	G
	CUL4B	Deletion	3	97%	chrX:119694438-119694441	TCCT	T
	RB1	Deletion	1	95%	chr13:48951089-48951090	AA	A
NCI-H146	TP53	Deletion	19	100%	chr17:7576874-7576893	ATCCAGTGGTTTCTTCTTTG	A
NCI-H1693	SMARCA4	Deletion	17	100%	chr19:11138497-11138514	CTTGATAGAATTCTTCCC	C
NCI-H2009	EPHB1	Deletion	1	not found [a]	-	-	-
NCI-H2122	CDH1	Deletion	18	95%	chr16:68842431-68842449	AGAAAATGAAAAAGGCCCA	A
	PARP1	Deletion	1	66%	chr1:226551693-226551694	CC	C
	STK11	Deletion	1	90%	chr19:1221318-1221319	CC	C
NCI-H2228	RB1	Deletion	1	93%	chr13:48934154-48934155	GG	G
NCI-H727	TP53	Insertion	9	85%	chr17:7578443-7578443	A	ACTGCTTGTA
RL95-2	BRCA2	Insertion	1	43%	chr13:32912776-32912776	T	TT
	BRCA2	Insertion	1	44%	chr13:32913957-32913957	A	AA
	PTEN	Insertion	1	37%	chr10:89720817-89720817	A	AA
	NF2	Deletion	1	47%	chr22:30050711-30050712	AA	A
	PTEN	Deletion	1	43%	chr10:89720816-89720817	AA	A
	TP53	Deletion	3	41%	chr17:7578200-7578203	CCAC	C
SCC-9	TP53	Deletion	32	100%	chr17:7577084-7577116	TCTGTGCGCCGGTCTCTCCCA GGACAGGCACAA	T
SW1417	TP53	Deletion	14	98%	chr17:7577557-7577571	AGGAACTGTTACACA	A
SW48	BRCA2	Deletion	1	41% b	chr13:32913842-32913843	AA	A
	FBXW7	Deletion	1	43%	chr4:153244160-153244161	CC	C
THP-1	TP53	Deletion	26	97%	chr17:7578386-7578412	AGCGCTCATGGTGGGGGCA GCGCCTCA	A
TUR	PTEN	Insertion	4	93%	chr10:89692905-89692905	G	GCCCG

Supplementary Table 5A notes:

[a] No evidence for presence of event in sequence data, excluded from sensitivity analysis.

[b] Indel not in COSMIC database for SW48, but identical to indel in HCT-15.

Supplementary Table 5B
Sequencing metrics for tumor cell lines used for assessment of indel detection

Tumor cell line	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
A2058	22,634,933	91%	7,666,148	373	100%	0.37%
A253	18,157,377	92%	6,564,360	304	98%	0.40%
A549	17,816,231	91%	6,186,531	300	99%	0.38%
AGS	24,118,544	92%	8,016,723	371	100%	0.40%
AsPC-1	23,431,383	92%	8,119,725	402	100%	0.39%
CCRF-CEM	19,272,769	92%	6,486,064	312	100%	0.40%
DU-145	20,787,366	91%	7,927,454	386	99%	0.37%
HCC1187	13,034,172	91%	4,731,947	224	98%	0.38%
HCC1395	17,785,269	91%	6,844,908	331	98%	0.36%
HCC1599	25,317,227	92%	9,089,866	436	99%	0.38%
HCC1937	16,879,720	92%	6,325,353	309	97%	0.36%
HCT-15	20,597,740	92%	7,395,353	354	100%	0.39%
J-RT3-T3-5	18,481,482	92%	6,318,735	302	99%	0.37%
LoVo	22,913,590	91%	7,301,191	338	100%	0.37%
MV-4-11	20,115,065	92%	6,494,045	315	100%	0.37%
NCI-H128	20,654,831	91%	7,209,021	356	100%	0.40%
NCI-H146	18,413,662	92%	6,235,752	295	99%	0.37%
NCI-H1693	17,973,760	92%	6,749,231	333	97%	0.37%
NCI-H2009	19,010,570	92%	6,273,869	293	98%	0.37%
NCI-H2122	19,957,036	91%	7,207,395	344	99%	0.35%
NCI-H2228	24,226,669	92%	8,766,637	416	100%	0.38%
NCI-H727	19,853,855	92%	6,798,088	328	99%	0.37%
RL95-2	16,165,504	92%	5,828,319	279	99%	0.39%
SCC-9	19,900,014	91%	6,492,864	305	99%	0.38%
SW1417	18,515,389	92%	6,456,540	309	99%	0.36%
SW48	17,235,671	92%	6,360,536	298	100%	0.38%
THP-1	24,543,517	91%	8,748,070	426	99%	0.36%
TUR	20,843,773	92%	6,938,859	339	100%	0.37%

Supplementary Table 5B notes:

Pure cell lines used to establish indel test set were sequenced at reduced average coverage in comparison with pooled samples. As many tumor cell lines carry homozygous deletions, all exons are not expected to be covered >100x.

Supplementary Table 6A
Summary of tumor cell line pools used for assessment of indel detection

Pools of two cell lines (14)

Pool	Components	
pool A	NCI-H1693	TUR
pool B	NCI-H146	THP-1
pool C	A549	HCT-15
pool D	A253	NCI-H2122
pool E	J.RT3-T3.5	NCI-H2228
pool F	AsPC-1	DU-145
pool G	NCI-H128	NCI-H727
pool H	NCI-H2009	SW1417
pool I	CCRF-CEM	HCC1395
pool J	AGS	HCC1187
pool K	HCC1937	SCC-9
pool L	HCC1599	MV-4-11
pool M	A2058	LoVo
pool N	RL95-2	SW48

Pools of three cell lines (8)

pool O	HCC1187	J.RT3-T3.5	MV-4-11
pool P	NCI-H128	NCI-H2122	NCI-H727
pool Q	AGS	NCI-H2009	RL95-2
pool R	HCC1937	SW48	TUR
pool S	LoVo	SCC-9	THP-1
pool T	CCRF-CEM	HCC1599	HCT-15
pool U	A253	AsPC-1	NCI-H146
pool V	A2058	A549	DU-145

Pools of four cell lines (9)

pool W	HCC1395	NCI-H1693	NCI-H2228	SW1417
pool X	HCC1187	MV-4-11	NCI-H1693	RL95-2
pool Y	A2058	A549	NCI-H2228	TUR
pool Z	HCC1395	SCC-9	SW1417	THP-1
pool AA	A253	AGS	CCRF-CEM	J.RT3-T3.5
pool BB	DU-145	HCC1937	LoVo	NCI-H128
pool CC	HCC1599	HCT-15	NCI-H146	SW48
pool DD	AsPC-1	NCI-H2009	NCI-H2122	NCI-H727
pool EE	AGS	DU-145	MV-4-11	SCC-9

Pools of five cell lines (8)

pool FF	A549	AGS	DU-145	HCT-15	TUR
pool GG	A2058	NCI-H2122	NCI-H727	RL95-2	THP-1
pool HH	A253	HCC1395	LoVo	NCI-H146	SCC-9
pool II	CCRF-CEM	HCC1599	NCI-H128	NCI-H2228	SW48
pool JJ	HCC1599	HCC1937	MV-4-11	NCI-H1693	NCI-H2009
pool KK	A2058	HCC1395	NCI-H1693	NCI-H2009	NCI-H2228
pool LL	AGS	AsPC-1	DU-145	MV-4-11	SCC-9
pool MM	AsPC-1	HCC1187	HCC1937	J.RT3-T3.5	SW1417

Pools of six cell lines (1)

pool NN	A253	HCC1187	J.RT3-T3.5	LoVo	NCI-H2122
	SW1417				

Pools of ten cell lines (1)

pool OO	A549	CCRF-CEM	HCT-15	NCI-H128	NCI-H146
	NCI-H727	RL95-2	SW48	THP-1	TUR

Supplementary Table 6B
Sequencing metrics for cell line pools used for assessment of indel detection

Pool ID	Number of cell lines in pool	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
pool A	2	56,997,382	92%	17,929,700	871	100%	0.34%
pool B	2	48,694,571	92%	15,462,657	739	100%	0.34%
pool C	2	55,572,417	92%	16,738,366	787	100%	0.35%
pool D	2	47,141,202	92%	15,500,432	755	99%	0.36%
pool E	2	38,432,222	92%	12,847,382	613	100%	0.36%
pool F	2	59,996,187	91%	17,765,328	840	100%	0.37%
pool G	2	46,596,923	91%	14,822,939	736	100%	0.34%
pool H	2	59,444,095	92%	17,702,973	825	100%	0.33%
pool I	2	52,122,793	92%	17,313,786	834	100%	0.31%
pool J	2	40,776,658	92%	13,403,473	647	100%	0.32%
pool K	2	35,979,025	92%	12,228,130	590	100%	0.31%
pool L	2	28,147,995	92%	9,950,582	488	100%	0.43%
pool M	2	42,839,903	92%	13,657,350	649	100%	0.30%
pool N	2	32,407,654	92%	11,040,621	531	100%	0.30%
pool O	3	48,140,867	92%	14,805,569	703	100%	0.31%
pool P	3	36,466,712	92%	12,288,023	605	100%	0.44%
pool Q	3	45,502,492	92%	14,650,015	697	100%	0.43%
pool R	3	39,033,530	92%	13,046,950	634	100%	0.34%
pool S	3	41,299,953	92%	13,204,933	645	100%	0.35%
pool T	3	37,251,264	92%	11,857,641	579	100%	0.35%
pool U	3	36,059,025	92%	11,274,897	533	100%	0.35%
pool V	3	35,912,841	91%	11,312,939	554	100%	0.35%
pool W	4	48,919,377	92%	15,104,316	719	100%	0.34%
pool X	4	40,021,398	92%	13,454,664	667	100%	0.30%
pool Y	4	45,584,864	92%	14,473,673	709	100%	0.30%
pool Z	4	41,980,486	92%	14,108,281	677	100%	0.31%
pool AA	4	52,663,991	92%	16,326,344	776	100%	0.34%
pool BB	4	53,572,468	92%	17,063,546	817	100%	0.34%
pool CC	4	46,024,523	92%	15,010,423	721	100%	0.34%
pool DD	4	66,455,797	92%	19,117,598	898	100%	0.36%
pool EE	4	44,255,734	92%	13,611,473	646	100%	0.44%
pool FF	5	45,781,116	92%	14,602,374	693	100%	0.37%
pool GG	5	49,662,505	92%	16,125,949	778	100%	0.33%
pool HH	5	31,463,295	92%	10,597,876	510	100%	0.33%
pool II	5	53,921,196	92%	17,056,368	823	100%	0.34%
pool JJ	5	49,921,714	92%	15,905,111	761	100%	0.31%
pool KK	5	45,745,404	92%	14,735,480	697	100%	0.43%
pool LL	5	45,313,285	92%	13,006,280	632	100%	0.31%
pool MM	5	52,922,680	92%	16,733,324	797	100%	0.44%
pool NN	6	44,477,558	92%	14,835,358	711	100%	0.30%
pool OO	10	55,203,747	92%	16,756,696	814	100%	0.31%

Supplementary Table 7A
Summary of indel detection performance (sensitivity)

Down-sampled coverage	MAF < 10% n = 60			MAF 10 - 20% n = 73			MAF > 20% n = 94		
	False negatives	Mean	Confidence interval	False negatives	Mean	Confidence interval	False negatives	Mean	Confidence interval
667	7	88.3%	77.4 – 95.2%	2	97.3%	90.5 – 99.7%	2	97.9%	92.5 – 99.7%
500	11	81.7%	69.6 – 90.5%	2	97.3%	90.5 – 99.7%	2	97.9%	92.5 – 99.7%
430	12	80.0%	67.7 – 89.2%	2	97.3%	90.5 – 99.7%	2	97.9%	92.5 – 99.7%
355	16	73.3%	60.3 – 83.9%	3	95.9%	88.5 – 99.1%	2	97.9%	92.5 – 99.7%
276	21	65.0%	51.6 – 76.7%	6	91.8%	82.3 – 96.9%	2	97.9%	92.5 – 99.7%
190	21	65.0%	51.6 – 76.7%	10	86.3%	76.2 – 93.2%	2	97.9%	92.5 – 99.7%

Supplementary Table 7B
Summary of indel detection performance (specificity)

Down-sampled coverage	True positives	False positives		Positive predictive value	
		MAF ≥ 20%	MAF < 20%	Mean	Confidence interval
667	875	0	3	99.7%	99.0 – 99.9%
500	873	0	0	100%	99.6 – 100%
430	871	0	3	99.7%	99.0 – 99.9%
355	873	0	1	99.9%	99.4 – 100%
276	844	0	1	99.9%	99.3 – 100%
190	790	0	5	99.4%	98.5 – 99.8%

Supplementary Table 7C
Summary of indel detection performance (full sensitivity results)

Available in online materials

Supplementary Table 7 notes:
Indel calls in pooled samples were considered to agree with indel calls in constituent cell lines, if they had the same base composition (number of each nucleotide added or deleted) and position (+/- 25bp). Confidence intervals were calculated as the exact 95% binomial confidence interval. False positives are indel calls in pooled samples that are absent from constituent cell-line sequence data. Coverage listed is the average of the median exon coverage of 41 pools. Down sampling was performed in-silico.

Supplementary Table 8

Known indels alterations in pure cell lines and pools not detected by Dindel algorithm

Known indel	Tumor cell line / Pool ID	Indel length	Expected allele frequency
A2058_PTEN_D35	A2058	35	53.0%
A2058_PTEN_D35	pool M	35	26.4%
A2058_PTEN_D35	pool V	35	23.9%
A2058_PTEN_D35	pool Y	35	16.3%
A2058_PTEN_D35	pool GG	35	10.5%
A2058_PTEN_D35	pool KK	35	10.1%
A549_SMARCA4_D23	pool V	23	22.2%
A549_SMARCA4_D23	pool OO	23	9.3%
CCRFCEM_NOTCH1_I36	CCRF-CEM	36	48.0%
CCRFCEM_NOTCH1_I36	pool I	36	19.1%
CCRFCEM_NOTCH1_I36	pool T	36	16.6%
CCRFCEM_NOTCH1_I36	pool AA	36	7.8%
CCRFCEM_NOTCH1_I36	pool II	36	6.5%
CCRFCEM_NOTCH1_I36	pool OO	36	3.0%
JRT3T35_PTEN_I39	J-RT3-T3-5	39	44.0%
JRT3T35_PTEN_I39	pool E	39	21.6%
JRT3T35_PTEN_I39	pool O	39	15.4%
JRT3T35_PTEN_I39	pool AA	39	13.9%
JRT3T35_PTEN_I39	pool MM	39	13.9%
JRT3T35_PTEN_I39	pool NN	39	6.0%
JRT3T35_PTEN_I7	J-RT3-T3-5	7	44.0%
JRT3T35_PTEN_I7	pool E	7	21.6%
JRT3T35_PTEN_I7	pool O	7	15.4%
JRT3T35_PTEN_I7	pool AA	7	13.9%
JRT3T35_PTEN_I7	pool MM	7	13.9%
JRT3T35_PTEN_I7	pool NN	7	6.0%
MV411_FLT3_I30	MV-4-11	30	39.0%
MV411_FLT3_I30	pool L	30	19.0%
MV411_FLT3_I30	pool EE	30	8.1%
MV411_FLT3_I30	pool O	30	7.6%
MV411_FLT3_I30	pool LL	30	7.1%
MV411_FLT3_I30	pool JJ	30	6.0%
MV411_FLT3_I30	pool X	30	6.0%
NCIH727_TP53_I9	pool P	9	13.5%
RL952_PTEN_D1	pool Q	1	17.8%
RL952_PTEN_D1	pool GG	1	6.7%
RL952_PTEN_D1	pool OO	1	4.3%
RL952_PTEN_I1	RL95-2	1	37.0%
RL952_PTEN_I1	pool N	1	16.2%
RL952_PTEN_I1	pool Q	1	14.3%
RL952_PTEN_I1	pool X	1	7.3%
RL952_PTEN_I1	pool GG	1	5.4%
SCC9_TP53_D32	SCC-9	32	99.0%
SCC9_TP53_D32	pool K	32	30.4%
SCC9_TP53_D32	pool S	32	20.2%
SCC9_TP53_D32	pool EE	32	15.3%
SCC9_TP53_D32	pool LL	32	15.3%
SCC9_TP53_D32	pool Z	32	10.1%
SCC9_TP53_D32	pool HH	32	7.8%
THP1_TP53_D26	THP-1	26	99.0%
THP1_TP53_D26	pool B	26	49.9%
THP1_TP53_D26	pool S	26	34.7%
THP1_TP53_D26	pool Z	26	17.2%
THP1_TP53_D26	pool GG	26	13.6%
THP1_TP53_D26	pool OO	26	8.4%

Supplementary Table 9A
Summary of normal and tumor cell lines used for assessment of copy number alteration detection

Tumor cell line	Ploidy	Normal cell line	Previously known amplifications (copy number)	Previously known deletions (copy number)	Novel amplifications (copy number)	Novel deletions (copy number)
HCC1143	3-4	HCC1143 BL	CCND1 (12), MYC (8)	-	AKT1 (11), FGF3 (12), FGF4 (12), FGF19 (12), MDM2 (9), SRC (7)	PBRM1 (0)
HCC2218	2	HCC2218 BL	ERBB2 (9)	-	-	CDH1 (0)
NCI-H1395	2-3	NCI-BL1395	MCL1 (10)	-	MYC (10)	-
NCI-H1770	2-3	NCI-BL1770	MYCN (13)	-	-	-
NCI-H2009	2	NCI-BL2009	MET (4) [a]	RB1 (0)	FGF10 (8), FGF14 (6), IRS2 (6), RICTOR (8)	-
NCI-H2122	2-3	NCI-BL2122	MYC (10)	CDKN2A (0), CDKN2B (0)	-	LRP1B (0)
NCI-H2126	3-4	NCI-BL2126	-	CDKN2A (0), CDKN2B (0)	MCL1 (6), RICTOR (7)	LRP1B (0)

Supplementary Table 9A notes:
[a] MET was not found to be significantly amplified in NCI-H2009 (4 copies) and was excluded from the analysis.

Supplementary Table 9B

Sequencing metrics for normal and tumor cell lines used for assessment of copy number alteration detection

Tumor cell line	Normal cell line	Percent tumor/normal	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
HCC1143	HCC1143 BL	100% / 0%	29,836,251	92%	11,329,859	515	99%	0.28%
HCC1143	HCC1143 BL	80% / 20%	35,582,339	92%	13,707,084	596	100%	0.29%
HCC1143	HCC1143 BL	50% / 50%	30,391,462	92%	12,082,932	561	100%	0.32%
HCC1143	HCC1143 BL	40% / 60%	29,607,219	92%	11,370,365	532	100%	0.26%
HCC1143	HCC1143 BL	30% / 70%	43,592,209	92%	16,551,377	775	100%	0.29%
HCC1143	HCC1143 BL	20% / 80%	37,359,326	92%	13,525,907	637	100%	0.27%
HCC2218	HCC2218 BL	100% / 0%	22,605,297	92%	7,125,071	266	97%	0.29%
HCC2218	HCC2218 BL	80% / 20%	36,937,520	92%	13,076,437	592	100%	0.29%
HCC2218	HCC2218 BL	50% / 50%	33,793,687	92%	11,673,760	537	100%	0.36%
HCC2218	HCC2218 BL	40% / 60%	36,611,814	92%	7,514,667	280	99%	0.31%
HCC2218	HCC2218 BL	30% / 70%	43,637,833	92%	13,828,282	593	100%	0.28%
HCC2218	HCC2218 BL	20% / 80%	36,495,222	92%	12,532,424	593	100%	0.37%
NCI-H1395	NCI-BL1395	100% / 0%	29,065,742	92%	10,700,333	502	100%	0.37%
NCI-H1395	NCI-BL1395	80% / 20%	36,038,287	92%	11,361,481	525	100%	0.31%
NCI-H1395	NCI-BL1395	50% / 50%	33,366,113	92%	12,138,176	573	100%	0.32%
NCI-H1395	NCI-BL1395	40% / 60%	25,539,233	92%	7,818,250	329	100%	0.27%
NCI-H1395	NCI-BL1395	30% / 70%	40,361,718	92%	7,346,586	288	99%	0.30%
NCI-H1395	NCI-BL1395	20% / 80%	36,374,467	92%	12,401,341	600	100%	0.27%
NCI-H1770	NCI-BL1770	100% / 0%	46,569,279	92%	16,853,355	758	100%	0.27%
NCI-H1770	NCI-BL1770	80% / 20%	37,030,829	92%	10,396,701	416	100%	0.36%
NCI-H1770	NCI-BL1770	50% / 50%	31,830,465	92%	11,862,333	569	100%	0.29%
NCI-H1770	NCI-BL1770	40% / 60%	33,966,162	92%	12,887,759	615	100%	0.26%
NCI-H1770	NCI-BL1770	30% / 70%	30,690,667	92%	11,192,910	551	100%	0.28%
NCI-H1770	NCI-BL1770	20% / 80%	45,235,990	92%	13,730,324	555	100%	0.36%
NCI-H2009	NCI-BL2009	100% / 0%	35,758,125	92%	13,784,120	649	100%	0.28%
NCI-H2009	NCI-BL2009	80% / 20%	32,871,962	92%	12,124,715	571	100%	0.27%
NCI-H2009	NCI-BL2009	50% / 50%	36,068,520	92%	12,877,298	602	100%	0.29%
NCI-H2009	NCI-BL2009	40% / 60%	21,327,300	92%	8,402,288	406	99%	0.31%
NCI-H2009	NCI-BL2009	30% / 70%	33,856,799	92%	12,479,365	618	100%	0.27%
NCI-H2009	NCI-BL2009	20% / 80%	41,030,097	92%	14,059,098	677	100%	0.29%
NCI-H2122	NCI-BL2122	100% / 0%	50,870,155	92%	19,240,735	882	99%	0.33%
NCI-H2122	NCI-BL2122	80% / 20%	34,549,033	92%	13,083,989	555	99%	0.26%
NCI-H2122	NCI-BL2122	50% / 50%	30,676,188	91%	11,712,539	576	99%	0.25%
NCI-H2122	NCI-BL2122	40% / 60%	33,701,500	92%	12,869,190	615	100%	0.29%
NCI-H2122	NCI-BL2122	30% / 70%	37,369,863	92%	13,740,965	604	99%	0.27%
NCI-H2122	NCI-BL2122	20% / 80%	29,152,070	92%	10,413,365	454	100%	0.26%
NCI-H2126	NCI-BL2126	100% / 0%	27,692,780	92%	9,882,180	412	99%	0.29%
NCI-H2126	NCI-BL2126	80% / 20%	39,255,911	92%	14,565,890	690	100%	0.29%
NCI-H2126	NCI-BL2126	50% / 50%	27,808,638	92%	10,810,327	522	100%	0.27%
NCI-H2126	NCI-BL2126	40% / 60%	36,326,389	92%	13,132,951	652	100%	0.36%
NCI-H2126	NCI-BL2126	30% / 70%	38,957,179	92%	14,683,179	706	100%	0.28%
NCI-H2126	NCI-BL2126	20% / 80%	40,503,947	92%	15,530,310	771	100%	0.28%

Supplementary Table 9B notes:

Many tumor cell lines have homozygous chromosomal deletions, so some exons are not expected to be covered >100x in these samples.

Supplementary Table 10A
Summary of copy number alteration detection performance (sensitivity)

CNA type	Tumor fraction = 20%				Tumor fraction ≥ 30%			
	Total CNAs	False negatives	Mean	Confidence interval	Total CNAs	False negatives	Mean	Confidence interval
Amplification CN ≥ 8	14	1	92.9%	66.1 – 99.8%	56	0	100%	93.6 – 100%
Amplification CN ≥ 6	19	3	84.2%	60.4 – 96.6%	76	8	89.5%	80.3 – 95.3%
Deletion	9	1	88.9%	51.8 – 99.7%	36	1	97.2%	85.5 – 99.9%

Supplementary Table 10B
Summary of copy number alteration detection performance (specificity)

CNA type	All positives	False positives	Positive predictive value	
			Mean	Confidence interval
Amplification	84	0	100%	95.7 – 100%
Deletion	43	0	100%	91.8 – 100%

Supplementary Table 10A-B notes:
CNA amplification calls were considered false positives if the copy level of that gene in the pure tumor cell line was estimated to be <5 copies. Confidence intervals were calculated as the exact 95% binomial confidence interval.

Supplementary Table 10C
Summary of copy number alteration detection performance (full sensitivity results)

Tumor cell line	Gene	Copy number	Percent tumor cell line in pool				
			75%	50%	40%	30%	20%
HCC1143	CCND1	12	X	X	X	X	X
HCC1143	MYC	8	X	X	X	X	X
HCC1143	AKT1	11	X	X	X	X	X
HCC1143	FGF3	12	X	X	X	X	X
HCC1143	FGF4	12	X	X	X	X	X
HCC1143	FGF19	12	X	X	X	X	X
HCC1143	MDM2	9	X	X	X	X	O
HCC1143	SRC	7	X	O	O	O	X
HCC1143	PBRM1	0	X	X	O	X	O
HCC2218	ERBB2	9	X	X	X	X	X
HCC2218	CDH1	0	X	X	X	X	X
NCI-H1395	MCL1	10	X	X	X	X	X
NCI-H1395	MYC	10	X	X	X	X	X
NCI-H1770	MYCN	13	X	X	X	X	X
NCI-H2009	RB1	0	X	X	X	X	X
NCI-H2009	FGF10	8	X	X	X	X	X
NCI-H2009	FGF14	6	O	O	O	X	X
NCI-H2009	IRS2	6	X	O	O	X	X
NCI-H2009	RICTOR	8	X	X	X	X	X
NCI-H2122	MYC	10	X	X	X	X	X
NCI-H2122	CDKN2A	0	X	X	X	X	X
NCI-H2122	CDKN2B	0	X	X	X	X	X
NCI-H2122	LRP1B	0	X	X	X	X	X
NCI-H2126	CDKN2A	0	X	X	X	X	X
NCI-H2126	CDKN2B	0	X	X	X	X	X
NCI-H2126	MCL1	6	X	X	X	X	O
NCI-H2126	RICTOR	7	X	X	X	X	O
NCI-H2126	LRP1B	0	X	X	X	X	X

Supplementary Table 10C notes:
X = Detected, O = Not detected

Supplementary Table 11A
Summary of NGS/Sequenom discordances

NGS + / Sequenom - discordances	
Probable discordance reason	**# calls**
Sensitivity / Tissue heterogeneity (weak evidence by Sequenom)	5
Unresolved	2
Total	7

NGS - / Sequenom + discordances	
Probable discordance reason	**# calls**
Sequenom miscall	1
NGS miscall	1
Sensitivity / Tissue heterogeneity (weak evidence by NGS)	1
Unresolved	1
Total	4

Supplementary Table 11B
Summary of sequencing metrics for short variant (Sequenom) concordance samples (n=118)

Percentile	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
5	28,074,541	89%	4,756,532	326	90%	0.25%
25	33,877,061	92%	8,298,227	765	93%	0.30%
50 (median)	40,722,157	95%	10,178,557	921	95%	0.35%
75	47,360,898	97%	15,441,199	1,227	97%	0.41%
95	53,594,991	97%	18,497,822	1,465	100%	0.47%

Supplementary Table 12A
Summary of concordance between NGS and FISH/IHC

Study	Tissue	Gene tested	Method	Number of samples [b]	FISH/IHC positive samples	NGS positive samples	Discordant calls	NGS Accuracy vs. FISH/IHC [i]
A	Prostate	AR - Amp.	FISH	25	5	5	0	100%
A	Prostate	PTEN - Del.	FISH	22	6	7	1 [f]	95%
B	Breast	HER2 - Amp.	FISH	42	6	6 [e]	0	100%
C	Breast	HER2 - Amp.	FISH	30 [c]	14	13	1 [g]	97%
D	Head and Neck	CCND1 - Amp.	IHC [a]	34	8	9	1 [h]	97%
D	Head and Neck	PTEN - Del.	IHC [a]	32 [d]	3	3	0	100%

Supplementary Table 12A notes:
[a] Strong expression for amplification or loss of staining for homozygous deletion
[b] Number of samples with copy number calls by both NGS and FISH/IHC (purity cut-off >20% for NGS), all NGS calls made blinded to confirmation method data
[c] Excludes one case classified as intermediate (5 copies) by NGS
[d] Excludes two cases with PTEN mutation by NGS and loss of staining
[e] Excludes one case with amplification of HER2 exons 17-27 only
[f] FISH analysis repeated and NGS data re-examined – discordance confirmed
[g] No evidence of gain by NGS
[h] Intermediate FISH expression
[i] 1 – (number of discordant calls / number of samples)

Supplementary Table 12B
Summary of sequencing metrics for CNA (FISH/IHC) concordance samples (n=131)

Percentile	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
5	27,052,257	93%	2,395,330	164	84%	0.26%
25	40,105,861	96%	7,512,241	571	100%	0.41%
50 (median)	45,942,100	97%	12,781,379	969	100%	0.49%
75	49,895,221	97%	15,992,746	1,188	100%	0.55%
95	58,774,526	98%	19,927,762	1,458	100%	0.75%

Supplementary Table 13
Summary of specimens and alterations in reproducibility (precision) assessment

			TEST REPLICATE				
Short variants (base substitutions and indels)			**1A**	**1B**	**1C**	**2**	**3**
Gene	**Variant type**						
ATRX	base sub.	splice	X	X	X	X	X
CDKN2A	base sub.	nonsense	X	X	X	X	X
CDKN2A	base sub.	nonsense	X	X	X	X	X
FGFR1	base sub.	missense	X	X	X	X	X
MLL2	indel	frameshift	X	X	X	X	X
NOTCH1	indel	frameshift	O	O	O	O	X
NPM1	indel	frameshift	X	X	X	X	X
PIK3CA	base sub.	missense	X	X	X	X	X
TP53	base sub.	splice	X	X	X	X	X
TP53	base sub.	missense	X	X	X	X	X
TP53	indel	nonframeshift	X	X	X	X	X
TP53	base sub.	missense	X	X	X	X	X
TP53	base sub.	nonsense	X	X	X	X	X
TP53	indel	frameshift	X	X	X	X	X

Copy number alterations			**1A**	**1B**	**1C**	**2**	**3**
Gene	**CNA type**	**Copy level**					
CCND1	amplification	16	X	X	X	X	X
CCND1	amplification	8	X	X	X	X	X
CCNE1	amplification	16	X	X	X	X	X
EGFR	amplification	12	X	X	X	X	X
ERBB2	amplification	16	X	X	X	X	X
ERBB2	amplification	16	X	X	X	X	X
FGF19	amplification	16	X	X	X	X	X
FGF19	amplification	8	X	X	X	X	X
FGF3	amplification	16	X	X	X	X	X
FGF3	amplification	8	X	X	X	X	X
FGF4	amplification	16	X	X	X	X	X
FGF4	amplification	8	X	X	X	X	X
FGFR1	amplification	16	X	X	X	X	X
IGF1R	amplification	8	O	X	O	O	O
MCL1	amplification	8	O	X	X	O	X
RAF1	amplification	16	X	X	X	X	X
ZNF703	amplification	16	X	X	X	X	X
CDKN2A	deletion	0	X	X	X	X	X
CDKN2B	deletion	0	X	X	X	X	X
RB1	deletion	0	X	O	X	X	X
TP53	deletion	0	X	O	X	X	X

Supplementary Table 13 notes:
X = Detected, O = Not detected. CNA calls judged equivocal on blinded model review were excluded, including all amplifications of copy number 6-7, as well as two amplifications at copy number 8 in replicate 1C. Overall concordance was calculated as the average of pair-wise replicate concordances. Pair-wise replicate concordance was calculated as (2 x # of concordant alterations detected) / (# of alterations detected in replicate 1 + # of alterations detected in replicate 2).

Supplementary Table 14
Summary of sequencing metrics for clinical tumor samples (n=2,221)

Percentile	Total read pairs	Fraction uniquely aligned to hg19	On-target unique read pairs	Median exon coverage	Fraction exons covered >= 100x	Sequence error rate
5	26,315,828	96%	5,234,501	358	99%	0.24%
25	39,920,154	97%	11,567,899	885	100%	0.29%
50 (median)	46,064,967	97%	15,169,372	1,189	100%	0.34%
75	53,252,037	97%	18,075,165	1,409	100%	0.41%
95	65,478,936	98%	22,085,190	1,728	100%	0.56%

Supplementary Table 15
Alteration details for Figure 6B

Available in online materials