Genomic and biological characterization of exon 4 KRAS mutations in human cancer

Manickam Janakiraman[1], Efsevia Vakiani[1,2], Zhaoshi Zeng[3], Christine A. Pratilas[4], Barry S. Taylor[5], Dhananjay Chitale[1,2], Ensar Halilovic[6], Manda Wilson[5], Kety Huberman[1], Julio Cezar Ricarte Filho[1], Yogindra Persaud[1], Douglas A. Levine[3], James A. Fagin[1,7], Suresh C. Jhanwar[2], John M. Mariadason[8], Alex Lash[5], Marc Ladanyi[1,2], Leonard B. Saltz[7], Adriana Heguy[1], Philip B. Paty[3], David B. Solit[1,7*]

[1]Human Oncology and Pathogenesis Program and the Departments of [2]Pathology, [3]Surgery, [4]Pediatrics, [5]Computational Biology, [6]Molecular Pharmacology and Chemistry, and [7]Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY and the [8]Ludwig Institute for Cancer Research, Melbourne, Australia,

Running Title: Exon 4 KRAS mutations in human cancers.

SUPPLEMENTAL METHODS

Cell culture and Immuno blotting. Cancer cell lines were obtained as specified in Supplemental Table 2. LS1034, HCC2998, CCCL-18, and CCCL-23 were grown in MEM with 10mM HEPES, non-essential amino acids (NEAA); C32, C80, C84 and C125PM in Iscove's modified DMEM; and SW1222 and SW1417 were grown using αMEM. All cell lines were supplemented with 2mM glutamine, 50units/ml each of penicillin and streptomycin, and 10% heat inactivated fetal bovine serum (Gemini Bioproducts, Carlsbad, CA) and incubated at 37°C in 5% CO_2. For proliferation assays, cells were plated in 96 well plates, at a density of 2500-5000 cells per well. After 24 hours, cells were treated with PD0325901 or gefitinib at a range of concentrations prepared by serial dilution. PD0325901 and gefitinib for *in vitro* studies was dissolved in DMSO to yield 1 and 10mM stock solutions, respectively, and stored at -20°C. For proliferation studies, the cells were exposed to Alamar Blue (AccuMed International, OH) three to five days following drug treatment, and plates were read using a fluorescence spectrophotometer. The dose required to inhibit growth by 50% (IC_{50}) was calculated using the SoftMaxPro ver.5 software. For focus formation assays, $5x10^5$ cells growing in log phase were either mock transfected with transfection reagent alone or with non-targeting siRNA pool, or KRAS siRNA pool (Dharmacon, CA). Cells were incubated at 37ºC for 2 weeks. Colonies were then stained with crystal violet (Sigma-Aldrich, St. Louis, MO) for 1 hour and plates were scanned using flat-bed scanner with foci counted manually and graphed.

Immunblots were performed as previously reported (19). p42/44 MAPK, phospho-p42/44 MAPK, Akt, phospho-Akt (ser473), RB, cleaved PARP, and cleaved caspase-3 antibodies were

obtained from Cell Signaling Technology. Anti-Cyclin D1 and p27 antibodies were obtained from Santa Cruz Biotechnology. To measure apoptotic cells, cells were seeded in 10cm dishes at a density of 1 x 10^6 cells/dish and the following day were treated with the indicated concentration of drug or vehicle (DMSO) for the indicated times. Both adherent and floating cells were harvested and stained with ethidium bromide using the method of Nüsse [48].

Additional Sequenom method details. The Sequenom MassARRAY system is based on matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS). In these assays, the mutant and germline alleles for a given point mutation produce single-allele base extension reaction products of different masses that are then resolved by MALDI-TOF MS. Both the amplification and extension primers were designed using Sequenom Assay Designer v3.1 software (Sequenom, San Diego, CA). The amplification primers were designed with a 10mer tag sequence to increase their mass so that they fall outside the range of detection of the MALDI-TOF mass spectrometer. Results were generated using the SpectroTYPER v3.4 software (Sequenom, San Diego, CA). All the positive cases were confirmed by visually reviewing the spectra. For the PCR amplification, a total of 15 ng of genomic DNA (in 1 μl) was amplified in a 5 μl reaction mixture containing 0.1 μl (0.5 U) HotStarTaq enzyme (Qiagen, Valencia, CA), 0.625 μl of 10x HotStar buffer, 0.325 μl of 25 mM (total) MgCl2, 0.25 μl of 10mM (each) deoxynucleotide triphosphate, 1 μl of 100 nM of each forward and reverse primers and 1.7 μl of water. The PCR step was initiated with a 95°C soak for 15 min, followed by 45 cycles, consisting of 95°C for 20 sec, 56°C for 30 sec, 72°C for 60 sec, and a final extension of 3 min at 72°C. After PCR, the remaining unincorporated dNTPs were dephosphorylated by adding 2 μl of the SAP cocktail, containing 1.33 μl of water, 0.17 μl of reaction buffer (Sequenom, San

Diego, CA) and 0.5 µl of SAP (Sequenom, San Diego, CA). The 384-well plate was then sealed and placed in a thermal cycler with the following conditions: 37°C for 40 min, 85°C for 5 min and then held at 4°C indefinitely. After the SAP treatment, a 2 µl cocktail, consisting of 0.755 µl water; 0.2 µl iPlex 10x buffer (Sequenom, San Diego, CA), 0.2 µl iPlex terminator mix (Sequenom, San Diego, CA); 0.804 µl of 7 µM/ 14 µM (depending on the low vs. high mass primers) extension primer mixture and 0.041 µl iPlex enzyme (Sequenom, San Diego, CA) was added. After the iPlex cocktail addition, the plate was again sealed and placed in a thermal cycler with the following program: 94°C for 2 min followed by 40 cycles of 94°C for 5 sec, [5 cycles (52°C for 5 sec, 80°C for 5 sec) and 72°C for 5 sec]. The reaction mixture was then desalted by adding 16 µl of water and 6 mg cationic resin mixture, SpectroCLEAN (Sequenom, San Diego, CA). The plate was then sealed and placed in a rotating shaker for 20 min to desalt the iPlex solution. Completed genotyping reactions were spotted in nanoliter volumes onto a matrix arrayed silicon chip with 384 elements (Sequenom SpectroCHIP) using the MassARRAY Nanodispenser. SpectroCHIPs were analyzed using the Bruker Autoflex MALDI-TOF mass spectrometer and the spectra were processed using the SpectroTYPER v3.4 software (Sequenom, San Diego, CA). Mutations included in the sequenom assay panel are listed in Table S5.

Multiplexed PCR and MS-based genetic fingerprinting assay and analysis method useful for sample identification, matching and tracking.

1) Selection of SNP assay loci. We downloaded all HapMap [49] non- redundant individual genotype data (organized by chromosome and population) from the HapMap FTP site [50]. After parsing through all loci on all chromosomes and in all HapMap populations, we filtering out any polymorphic loci of length greater than one or containing an 'N' character, flagged bi-allelic

SNP haplotypes appearing in AB:AA:BB ratios of greater than 0.45:0.225:0.225 (i.e., close to Hardy-Weinberg equilibrium for bi-allelic loci, where alleles are found in roughly equal proportions). We then counted the number of populations per locus passing this filter, and sorted the results. The loci with the best counts in each chromosome were retained. Allowing for ties, this resulted in 89 HapMap equally bi-allelic loci in a number of populations, which were also scattered across the genome and from which we could design a multiplexed SNP assay (see Table S6).

2) Multiplexed assay primer design. Forty-two out of the selected 89 SNPs, covering all chromosomes (most chromosomes are covered by 2 SNPs, see Table S6), that were amenable to assay design for mass spectrometry-based genotyping, were selected to design 4 multiplexed assays using the Sequenom Assay Design V.3.1 software (Sequenom, San Diego, CA). Two of the multiplexed wells also contained a sex determination assay based on the amelogenin gene which has two versions, one on the X chromosome and one in the Y chromosome and has areas of divergent sequence between the two versions of the gene [51]. Table S7 lists the selected SNP, specific primers and multiplexing scheme.

3) Sequenom assay. See above additional Sequenom method details

4) Sample matching and clustering tool. We created a tool that reads in the assay results and does a pair-wise comparison between all samples to look for samples that might be related. For each sample pair we assume no relation and ask, what is the probability of the two samples having at least *m* matching loci by chance? Since there are only two possible outcomes for two

samples at a locus: match and not matched, and the loci are independent, we used the binomial distribution to calculate the probability of at least *m* matching loci of the loci examined:

$$P = \sum_{k=m}^{N} \binom{N}{k} p^k (1-p)^{N-k}$$

, where *m* is the number of loci that match between the samples, *N* is the total number of loci for which we had results in both samples, and p^k is the probability of a match at a locus. p^k, the probability of two unrelated samples matching at a locus is given by: $(GG \cap GG) \cup (GT \cap GT) \cup (TT \cap TT) = (0.25 \bullet 0.25) + (0.5 \bullet 0.5) + (0.25 \bullet 0.25) = 0.0625 + 0.25 + 0.0625 = 0.375$. If a sample was run more than once and the tool finds more than one sample name and assay identifier pair with the same genotype, the sample name and assay pair is only included once. If the duplicates conflict, all results for that sample and assay pair are excluded. We used the Bonferroni method to correct for multiple comparisons. Any sample pair with a corrected p-value of < 0.05, we consider unlikely to be unrelated. It is also possible that two similar samples have been missed. Sample pairs deemed *likely* to be related are examined manually to determine if a labeling error occurred.

A simple clustering tool reads the results and groups samples that are likely to be related. In this clustering procedure, samples are clustered together in a cluster if they are related (at a corrected p-value of 0.05) to at least one other sample in the cluster. Samples are assigned to one cluster exclusively, and singlet clusters are allowed. Both macros can be downloaded from the web site[52]

5) **Validation of the MS-based genetic fingerprinting assay and analysis method:** In order to cross-validate our assay and compare with other methodologies, we tested DNA from 16 cell

lines that had been previously genotyped by SNP microarray analysis [53]. Seven of the SNPs we used in our assay were represented in the Affymetrix 250K array and only 109 out of the 112 overlapping calls generated on 16 unrelated samples were concordant (p-value = 9.90×10^{-17}, one-tailed Fisher exact test, see Table S8). Additionally, when a panel of colorectal cell lines was analyzed using this assay we found a list of colorectal cancer cell lines that share identical genetic background (see Table S9).

SUPPLEMENTAL REFERENCES

48. Nusse M, Beisker W, Hoffmann C, Tarnok A. Flow cytometric analysis of G1- and G2/M-phase subpopulations in mammalian cell nuclei using side scatter and DNA content measurements. Cytometry 1990; 11: 813-21.
49. The International HapMap Project. Nature 2003; 426: 789-96.
50. http://ftp.hapmap.org/genotypes/2008-11_phaseII+III/forward/non-redundant/
51. Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro LJ. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. Am J Hum Genet 1992; 50: 303-16.
52. http://cbio.mskcc.org/dna_fingerprinting/
53. Schweppe RE, Klopper JP, Korch C, et al. Deoxyribonucleic acid profiling analysis of 40 human thyroid cancer cell lines reveals cross-contamination resulting in cell line redundancy and misidentification. J Clin Endocrinol Metab 2008; 93: 4331-41.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1. Representative Sanger traces of tumor and germline DNA confirming that A146T and K117N mutations were somatic. Both normal and tumor DNA were sequenced to confirm the somatic nature of the exon 4 mutations. Representative Sanger sequencing traces from normal and tumor DNA are shown for codon A146 (**A)** and K117N (**B)**. **C.** Cell lines for which matched tumor tissue was available were shown to harbor identical RAS mutations. Representative Sanger sequencing traces (left) and mass spectrometry peaks (right) are shown for the cell line CCCL-23 (above) and tumor (below).

Supplemental Figure 2. Kaplan-Meier plots of disease-specific survival of 186 patients with stages 1-3 colorectal cancer as a function of KRAS/NRAS mutational status. **A.** Patients were divided into three cohorts: (1) those with KRAS G12/G13 mutations (shown in blue), (2) those with non-G12/G13 mutations in KRAS or in NRAS (shown in red) and (3) those wild-type for KRAS (shown in green). Patients with exon 3 and 4 KRAS or NRAS mutations had a statistically significant improvement in disease-specific survival (p-value = 0.006). A trend towards improved disease-specific survival was observed in a pairwise comparison of the KRAS wild-type cohort versus the cohort of patients with G12/G13 KRAS mutations but this did not reach statistical significance (p-value = 0.07). A statistically significant improvement in disease-specific survival was observed in a pairwise comparison of the G12/G13 KRAS mutant cohort versus the cohort of patients with non-G12/G13 KRAS mutations or NRAS mutations (p-value = 0.02). All differences in disease-specific survival among any two classes of KRAS mutation were tested by the log-rank test statistic. **B.** Patients were divided into two cohorts: (A) KRAS G12/G13 mutant (shown in blue) or (B) KRAS G12/G13 wild-type (shown in green).

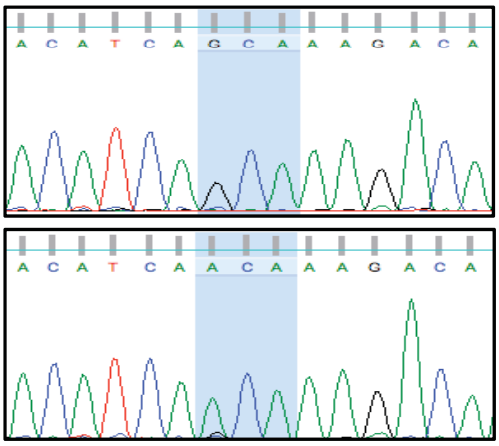
In a pairwise comparison of KRAS G12/G13 mutant versus KRAS G12/G13 wild-type, a statistically significant improvement in disease-specific survival was observed in the KRAS G12/G13 wild-type cohort (p-value = 0.02, log rank).

Supplemental Figure 3. Frequency and concordance of KRAS, NRAS and BRAF mutations in the colorectal cancer cell lines. A. Genomic DNA was extracted from 70 colorectal cancer cell lines and was screened for mutations in *KRAS*, *NRAS*, *HRAS*, *BRAF*, *MEK1* and *PIK3CA* using the MALDI-TOF based genotyping platform. Exon 2 KRAS mutations (G12 and G13) are shown in dark blue, exon 3 KRAS mutations (Q61) in light blue, exon 4 KRAS mutations (K117 and A146) in red, and BRAF mutations in green. **B.** Exon 4 KRAS mutations (all either K117N or A146T/V) were non-overlapping in distribution with mutations in KRAS exons 2 and 3 and BRAF. No colorectal cancer cell lines with NRAS mutation were identified. **C.** Graphical representation of the hotspot mutations in KRAS as mapped to the *KRAS4b* gene. Introns are represented as thin lines whereas exons (not to scale) are shown as boxes with the blue shading representing the coding regions and the green shading representing the untranslated regions (UTR). Red dots below the exons represent sites for hotspot mutations found in the human tumors and cell lines.

Figure S1

A

KRAS p.A146T (c.436 **G**CA > **A**CA)

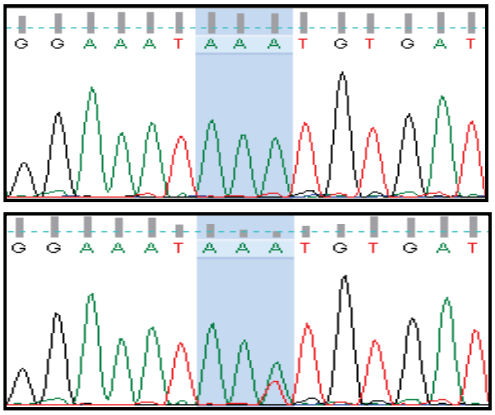


Normal

Tumor

B

KRAS p.K117N (c.351 AA**A** > AA**T**)

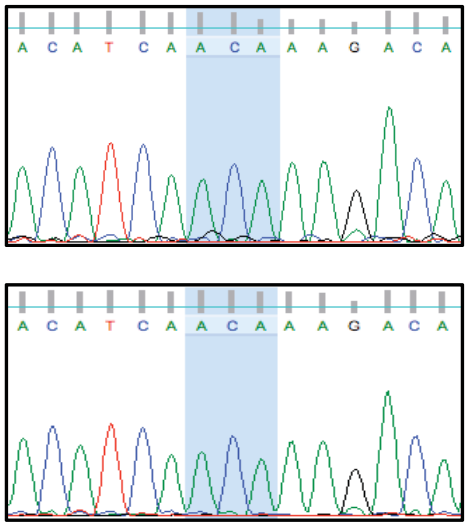


Normal

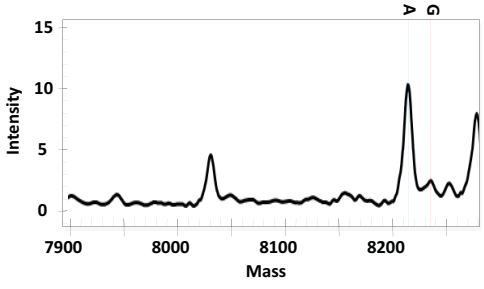
Tumor

C

KRAS p.A146T (c.436 **G**CA > **A**CA)



CCCL-23 Cell line



CCCL-23 Tumor

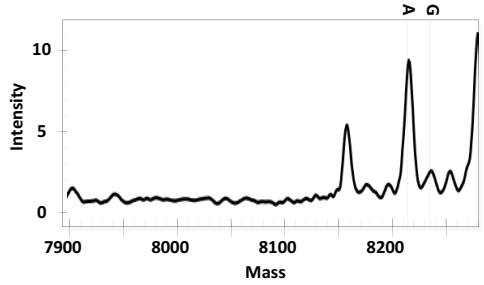
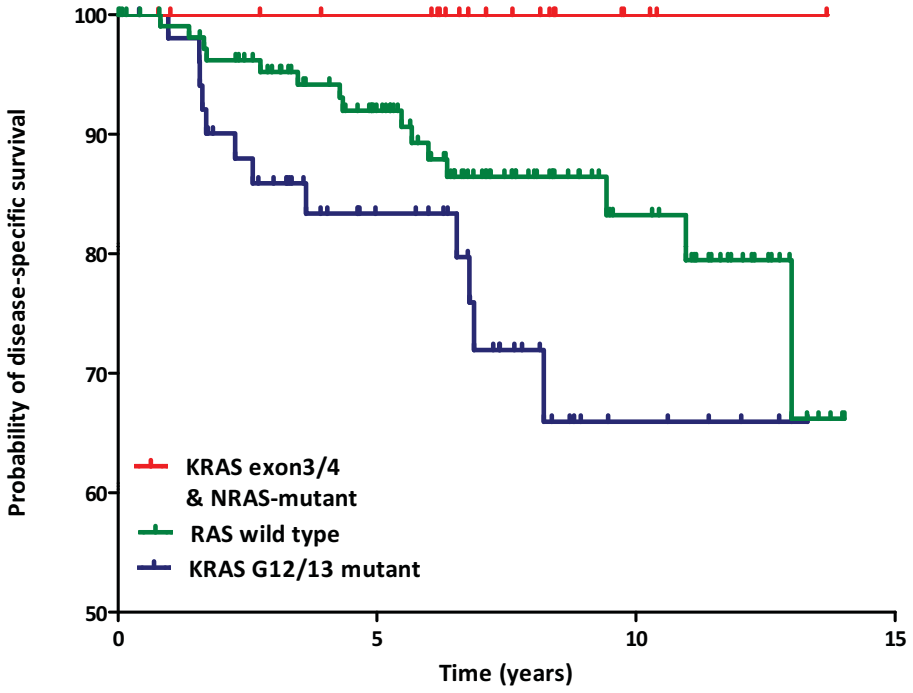


Figure S2

A


Probability of disease-specific survival
100
90
80
70
60
50
0
5
10
15
Time (years)
KRAS exon3/4
& NRAS-mutant
RAS wild type
KRAS G12/13 mutant

B

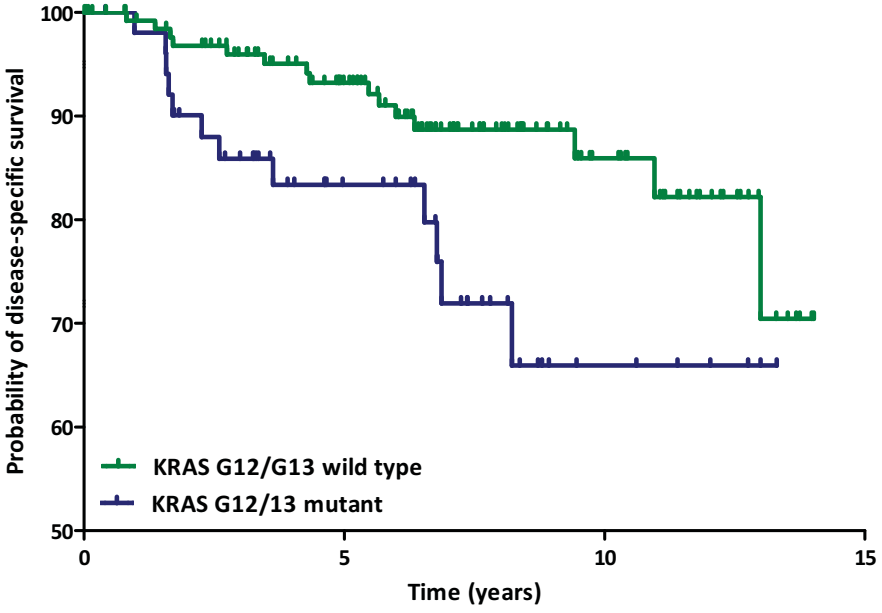

Probability of disease-specific survival
100
90
80
70
60
50
0
5
10
15
Time (years)
KRAS G12/G13 wild type
KRAS G12/13 mutant

Figure S3

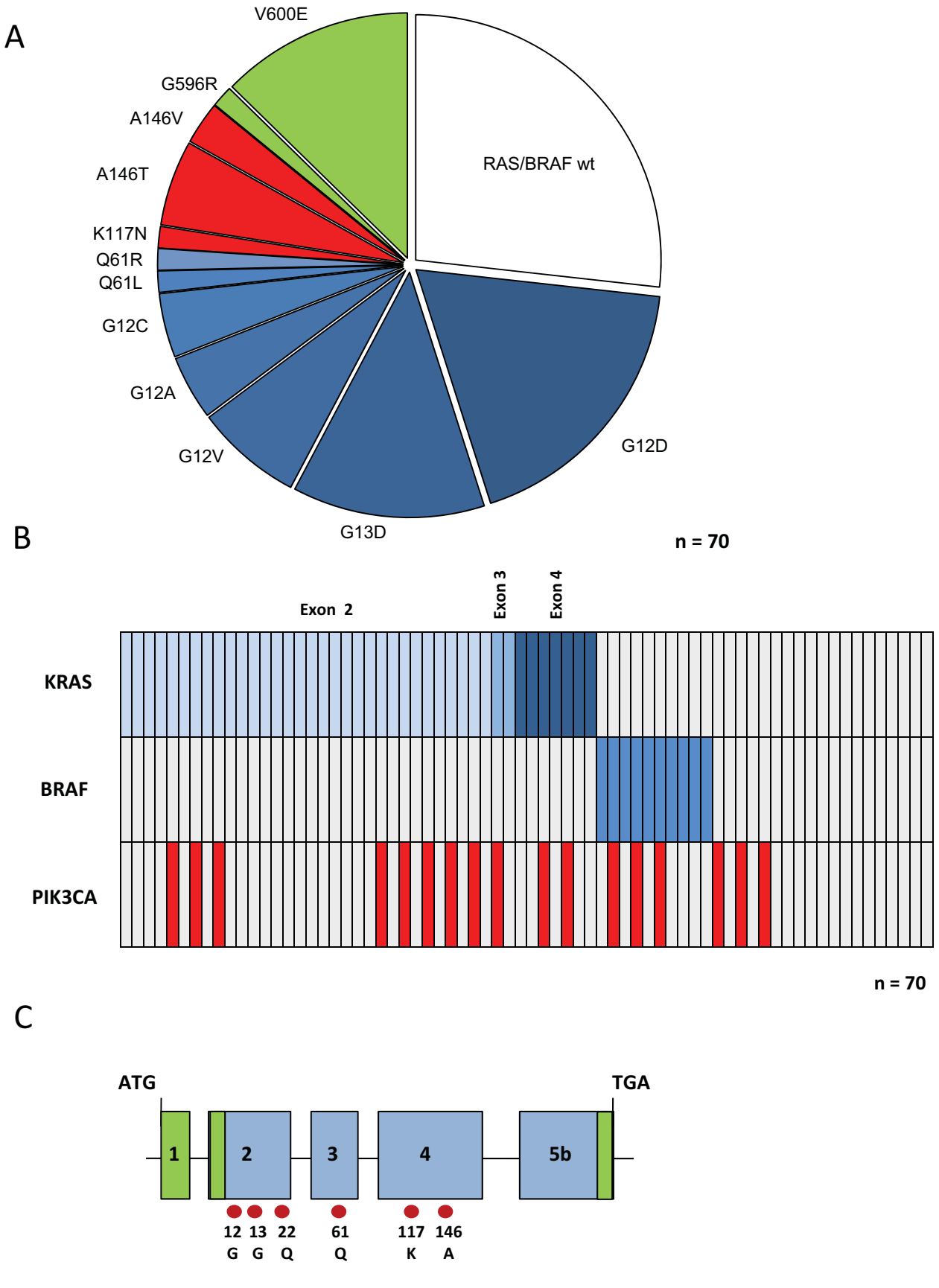

A
V600E
G596R
A146V
A146T
K117N
Q61R
Q61L
G12C
G12A
G12V
G13D
G12D
RAS/BRAF wt
B
n = 70
Exon 2
Exon 3
Exon 4
KRAS
BRAF
PIK3CA
n = 70
C
ATG
TGA
1
2
3
4
5b
12 13 22
G G Q
61
Q
117 146
K A

Table S1

Table S1. KRAS K117N & A146T/V mutations in patients for which annotated clinical data was available

Tissue Site	KRAS	PIK3CA	TP53	Sex	Age	Primary Location	Stage	Primary Size	FUP Status	FUP Interval (yrs)	MSI Status
polyp	A146T	wt	wt	M	67	Rectum					MSS
polyp	A146T	wt	wt	M	72	Ascending					MSS
polyp	A146T	wt	wt	M	77	Ascending					MSS
primary colon	A146T	wt	R213*	F	68	Rectum	I	3.5	NED	10.3	MSS
primary colon	A146T	wt	wt	F	61	Rectum	I	5.0	NED	6.6	MSS
primary colon	A146T	M1043I	wt	M	85	Cecum	I	3.5	DOC	13.7	MSI
primary colon	A146T	wt	wt	M	69	Cecum	I	5.0	NED	6.2	MSS
primary colon	K117N	H1047R	wt	M	75	Cecum	I	6	NED	1.0	MSI
primary colon	A146T	wt	R282W	F	73	Sigmoid	II	5.0	NED	6.3	MSS
primary colon	A146T	wt	wt	M	67	Rectum	II	4.5	NED	7.1	MSS
primary colon	A146T	wt	R273H	M	70	Cecum	II	6.0	NED	0.8	MSS
primary colon	A146V	wt	H214R	F	63	Transverse	II	4.2	NED	6.2	MSS
primary colon	A146T	wt	wt	F	41	Transverse	III	4.5	NED	10.4	MSS
primary colon	K117N	wt	wt	M	57	Rectum	III	4.0	NED	7.6	MSS
primary colon	A146T	wt	wt	F	70	Transverse	IV	5.0	DOD	1.8	MSS
primary colon	A146T	wt	wt	F	67	Sigmoid	IV	3.3	DOD	2.6	MSS
primary colon	K117N	wt	wt	F	75	Sigmoid	IV	3.0	DOD	0.8	MSS
liver met	A146T	wt	wt	F	50	Sigmoid	IV	4.5	NED	13.0	N/A
liver met	A146T	wt	S20L/R337C	F	59	Sigmoid	IV	N/A	NED	12.5	N/A
liver met	A146T	wt	R342*	M	59	Sigmoid	N/A	N/A	NED	19.5	N/A
liver met	A146T	wt	R306*	F	75	Rectum	IV	5.5	NED	1.7	N/A

Table S2

Table S2. List of KRAS, BRAF and PIK3CA mutations in 70 colorectal cancer cell lines.

Cell Lines	KRAS	BRAF	PIK3CA	Source
GEO	G12A	wt	wt	LICR
C84	G12A	wt	wt	CRUK
SW1116	G12A	wt	wt	ATCC
RW7213	G12C	wt	wt	LICR
SW837	G12C	wt	wt	LICR
Gp5d	G12D	wt	H1047L	LICR
LS174T	G12D	wt	H1047R	LICR
LS180	G12D	wt	H1047R	LICR
ALA	G12D	wt	wt	LICR
CCCL-2	G12D	wt	wt	MSKCC
CCCL-11	G12D	wt	wt	MSKCC
CCCL-13	G12D	wt	wt	MSKCC
Isreco1	G12D	wt	wt	LICR
Isreco2	G12D	wt	wt	LICR
Isreco3	G12D	wt	wt	LICR
LS513	G12D	wt	wt	LICR
SNU-C2B	G12D	wt	wt	LICR
TC71	G12D	wt	wt	LICR
CCCL-3	G12V	wt	wt	MSKCC
SKCO-1	G12V	wt	wt	LICR
SW403	G12V	wt	wt	LICR
SW480	G12V	wt	wt	LICR
SW620	G12V	wt	wt	LICR
T84	G13D	wt	E542K	LICR
DLD-1	G13D	wt	E545K	LICR
HCT-15	G13D	wt	E545K	LICR
HCT-8	G13D	wt	E545K	LICR
HCT116	G13D	wt	H1047R	LICR
CCCL-19	G13D	wt	wt	MSKCC
CCCL-28	G13D	wt	wt	MSKCC
LoVo	G13D	wt	wt	LICR
NCI-H747	G13D	wt	wt	ATCC
SW948	Q61L	wt	E542K	LICR
RW2982	Q61R	wt	wt	LICR
C125PM	K117N	wt	wt	CRUK
HCC2998	A146T	wt	wt	LICR
LS1034	A146T	wt	wt	ATCC
CCCL-18	A146T	wt	E542K	MSKCC
CCCL-23	A146T	wt	E545K	MSKCC
C80	A146V	wt	wt	CRUK
SW1222	A146V	wt	wt	ATCC
NCI-H508	wt	G596R	wt	ATCC
RKO	wt	V600E	H1047R	LICR
Vaco-5	wt	V600E	H1047R	LICR
HT29	wt	V600E	P449T*	LICR
Co-115	wt	V600E	wt	LICR
COLO-201	wt	V600E	wt	LICR
COLO-205	wt	V600E	wt	LICR
LS411N	wt	V600E	wt	ATCC
SW1417	wt	V600E	wt	LICR
WiDR	wt	V600E	wt	LICR
CCCL-10	wt	wt	E545K	MSKCC
CCCL-6	wt	wt	H1047R	MSKCC
CoCM-1	wt	wt	R1023Q	MSKCC
C2BBe1	wt	wt	wt	ATCC
Caco-2	wt	wt	wt	ATCC
C32	wt	wt	wt	CRUK
CCCL-14	wt	wt	wt	MSKCC
CCCL-20	wt	wt	wt	MSKCC
CCCL-21	wt	wt	wt	MSKCC
CCCL-26	wt	wt	wt	MSKCC
CCCL-30	wt	wt	wt	MSKCC
COLO-320	wt	wt	wt	LICR
DIFI	wt	wt	wt	LICR
HCA7	wt	wt	wt	LICR
HTB135	wt	wt	wt	ATCC
KM12	wt	wt	wt	NCI
LIM1215	wt	wt	wt	LICR
SW48	wt	wt	wt	LICR
V9P	wt	wt	wt	LICR

*From COSMIC (http://www.sanger.ac.uk/perl/genetics/CGP/core_line_viewer?action=sample&id=905939); MSKCC-Memorial Sloan-Kettering Cancer Center, New York; LICR-Ludwig Institute for Cancer Research, Melbourne; CRUK-Cancer Research, UK; ATCC-American Type Culture Collection, Manassas; NCI-National Cancer Institute, Bethesda.

Table S3

Table S3. Focal regions of copy number gain and loss identified by RAE Analysis of 128 colorectal tumors.

Locus	Region (Mb)	Effective frequency†	Q-value‡	Number of genes§	Genetic element of interest¥
Amplifications					
7p22.3-p21.3*	0.140-9.112	50 (3.1)	<4.86E-08	72,2	*CARD11,MMD2,FBXL18,PMS2,RPA3,ICA1,CHST12,PDGFA*
7q32.3-q34*	131.599-141.403	41.4 (2.3)	<4.86E-08	62,1	*TRIM24,HIPK2,DGKI,BRAF,CREB3L2*
7q36.1-q36.3*	152.011-158.811	39.8 (7)	<4.86E-08	20,2	hsa-mir-153-2,hsa-mir-595,*SHH,UBE3C,PAXIP1*
8q24.12 *	120.518-122.057	50 (10.2)	<4.86E-08	8,0	near *MYC; ENPP2,COL14A1,SNTB1*
9p22.1	19.286-19.331	21.1 (1.6)	0.0902	1,0	*DENND4C*
12p13.33*	0.544-0.551	21.1 (0.8)	0.0972	1,0	-
12p13.33 *	0.939-1.472	21.9 (0.8)	0.0642	1,0	*ERC1*
12p13.33-p13.31*	1.926-8.201	25.8 (2.3)	0.0042	88,2	*CACNA1C,NTF3,ING4,ZNF384,PTPN6,GDF3,CCND2,ZNF384*
13q12.13-q12.2*	26.194-27.744	56.2 (15.6)	<4.86E-08	15,0	*FLT3,CDX2,USP12; near CDK8*
13q14.11-q21.1*	40.932-56.650	57 (11.7)	<4.86E-08	72,2	hsa-mir-16-1,hsa-mir-15a,*HTR2A,RB1,LCP1*
16p11.2	27.812-28.187	21.1 (0.8)	0.0933	2,0	*XPO6*
16p11.2	28.774-29.090	21.1 (0.8)	0.0926	7,0	*CD19,SH2B*
20p13-p12.3*	0.590-6.362	39.1 (12.5)	<4.86E-08	73,1	-
20q11.21-q12*	29.592-39.033	68 (25.8)	<4.86E-08	110,2	*MYLK2,C20orf160,HCK,KIF3B,DNMT3B,DHX35,MAFB,NDRG3*
Deletions					
3p14.2	60.448-60.462	18 (6.2)	0.1308	1,0	*FHIT*
4q22.1*	92.062-92.096	27.3 (0.8)	1.83E-05	0,0	-
4q35.2*	189.387-189.417	28.1 (2.3)	6.13E-06	0,0	-
5q14.1	80.703-81.063	19.5 (0.8)	0.1172	2,0	*SSBP2*
5q14.2	81.609-82.016	19.5 (0.8)	0.1171	2,0	-
5q14.3	83.148-91.578	21.9 (0.8)	0.0203	13,1	*hsa-mir-9-2,EDIL3,MEF2C,COX7C*
5q15	93.191-93.309	19.5 (0.8)	0.1152	1,0	-
5q21.1-q22.3	97.997-114.112	26.6 (5.5)	0.0123	31,0	*APC,FBXL17,EFNA5*
5q23.1	115.292-116.096	19.5 (0.8)	0.1188	3,0	*SEMA6A*
5q23.1	120.408-120.477	18.8 (1.6)	0.1430	0,0	-
6q26	162.762-162.790	19.5 (1.6)	0.1472	1,0	*PARK2*
8p11.21	41.074-41.131	18.8 (0)	0.1364	0,0	*SFRP1*
10q23.2	89.459-89.573	25 (0.8)	0.0043	2,0	*PAPSS2* (near *PTEN*)
10q25.1	107.961-111.569	21.1 (0.8)	0.0609	1,0	*SORCS1*
10q25.2	114.214-114.844	20.3 (0.8)	0.1014	2,0	*TCF7L2*
16p13.2	6.429-6.883	34.4 (12.5)	4.10E-06	1,0	*A2BP1*
17p13.1-p11.2*	11.090-16.778	48.4 (1.6)	<7.34E-08	32,1	*hsa-mir-744,MAP2K4*
18p11.32-p11.31*	0.004-3.647	48.4 (1.6)	<7.34E-08	21,0	-
18p11.21*	11.126-12.656	49.2 (1.6)	<7.34E-08	9,0	*IMPA2,CIDEA*
18q21.1-q21.2*	45.118-50.105	60.2 (7)	<7.34E-08	21,0	*SMAD4,DCC*

† Single-copy gain, amplification in parenthesis; otherwise heterozygous and homozygous deletion

‡ False-discovery corrected p-value

§ Number of genes (RefSeq) and microRNAs (hg18)

* See supplementary table 4 (arm-length events)

¥ Genetic elements of interest selected from COSMIC, Cancer Gene Cesnsus, or other evidence

Notes:

Regional boundaries of the 10q23.2 deletion (85.796-107.064) span 200 additional genes and 7 microRNAs: hsa-mir-346, hsa-mir-107, hsa-mir-607, hsa-mir-608, hsa-mir-146b, hsa-mir-936, hsa-mir-609, *WNT8B, MINPP1, TNKS2, PLCE1, BMPR1A, FAS, TLX1, NFKB2, SUFU, GHITM, PCDH21, GRID1, OPN4, SNCG, C10orf116, GLUD1, ATAD1, C10orf59, STAMBPL1, CH25H, LIPA, IFIT1, PANK1, CPEB3, KIF11, HHEX, CYP26A1, CEP55, RBP4, LGI1, CYP2C18, CYP2C19, CYP2C9, CYP2C8, SORBS1, ALDH18A1, ENTPD1, DNTT, PIK3AP1, SLIT1, ARHGAP19, PGAM1, MMS19, SFRP5, CRTAC1, HPSE2, CNNM1, GOT1, ENTPD7, COX15, CUTC, CHUK, PKD2L1, SCD, NDUFB8, HIF1AN, BTRC, POLL, FGF8, PPRC1, PSD, FBXL15, ACTR1A, CYP17A1, CNNM2, NT5C2, INA, TAF5, PDCD11, NEURL, SH3PXD2A, COL17A1, GSTO2, SORCS3, C10orf99, LRRC22, LRRC21, RGR, KIAA1128, WAPAL, LDB3, MMRN2, KIAA1975, FAM35A, FAM22A, FAM22D, LIPL1, LIPF, LIPL2, LIPN, ANKRD22, ACTA2, IFIT2, IFIT3, IFIT1L, IFIT5, SLC16A12, MPHOSPH1, HTR7, RPP30, ANKRD1, PCGF5, HECTD2, PPP1R3C, FGFBP3, BTAF1, MARCH5, IDE, EXOC6, CYP26C1, FER1L3, GPR120, PDE6C, C10orf4, TMEM20, NOC3L, TBC1D12, HELLS, C10orf129, PDLIM1, TCTN3, C10orf130, CCNJ, ZNF518, BLNK, TMEM10, TLL2, SMBP, LCOR, C10orf12, FRAT1, FRAT2, KIAA0690, EXOSC1, ZDHHC16, UBTD1, ANKRD2, C10orf65, C10orf62, C10orf83, PI4KII, AVPI1, ZFYVE27, C10orf132, C10orf28, LOXL4, C10orf33, HPS1, NKX2-3, SLC25A28, ABCC2, DNMBP, CPN1, ERLIN1, CWF19L1, BLOC1S2, SEC31L2, PAX2, C10orf6, SEMA4G, MRPL43, PEO1, LZTS2, PDZD7, SFXN3, KAZALD1, TD1, LBX1, RP11-529I10.4, FBXW4, NPM3, MGEA5, KCNIP2, C10orf76, HPS6, LDB1, NOLC1, ELOVL3, PITX3, GBF1, CUEDC2, C10orf95, C10orf77, TRIM8, ARL3, SFXN2, C10orf26, C10orf32, AS3MT, PCGF6, USMG5, FAM26B, FAM26C, OBFC1, SLK, C10orf78, C10orf79, GSTO1, KIAA1754, C10orf80*

Table S4

Table S4. Arm length events identified by array CGH in 128 colorectal tumors.

Alteration	Arm(s)*	Q-value	Frequency(%)§
Gain	7p	<4.86E-08	47.7
Gain	7q	<4.86E-08	37.5
Gain	8q	<4.86E-08	46.1
Gain	12p	<4.86E-08	21.9
Gain	13q	<4.86E-08	54.7
Gain	20p	<4.86E-08	35.9
Gain	20q	<4.86E-08	65.6
Loss	1p	<7.34E-08	23.4
Loss	4p	3.38E-07	28.1
Loss	4q	<7.34E-08	25
Loss	8p	<7.34E-08	36.7
Loss	14q	<7.34E-08	21.9
Loss	15q	<7.34E-08	29.7
Loss	17p	<7.34E-08	44.5
Loss	18p	<7.34E-08	47.7
Loss	18q	<7.34E-08	57.8
Loss	21q	<7.34E-08	19.5
Loss	22q	2.44E-05	20.3

* Contiguous region of significant alteration spanning >50% of the chromosome arm by genomic coverage

§ Frequencies calculated from tumors whose median copy number alteration status for the arm was either single-copy gain or heterozygous loss

Table S5

Table S5. List of mutations screened using the colorectal Sequenom assay panel.

KRAS	NRAS	BRAF	PIK3CA	MEK1
G12V	G12V	G466E	K111N	K57N
G12D	G12D	G466V	C420R	
G12A	G12A	G466A	P539R	
G12S	G13D	G466R	E542K	
G12R	G13A	G469E	E542G	
G12C	G13V	G469A	E542V	
G13D	G13R	G469V	E542Q	
G13A	A18T	G469R	E545K	
G13V	Q61K	V600E	E545G	
G13S	Q61E	V600K	E545A	
G13R	Q61P	V600R	E545D	
G13C	Q61R	V600M	Q546K	
L19F	Q61L	K601E	Q546E	
Q22K	Q61H	L618S	Q546P	
Q22R			Q546R	
A59T			H1047R	
A59E			H1047L	
Q61K			H1047Y	
Q61E				
Q61P				
Q61R				
Q61L				
Q61H				
A146T				
A146V				
A146P				

Table S6

Table S6. SNPs included in the HapMap Sequenom assay

Chromosome	dbSNP id	Population Count	Populations and Genotype Counts
chrX	rs28539253	3	US_African-30-trios[CT:45|CC:26|TT:28] , US_Gujarati[CT:45|CC:30|TT:24] , Maasai_Kenyan-60-trios[CT:51|CC:25|TT:23]
chrX	rs6645103	3	CEPH-60-trios[CT:45|CC:31|TT:22] , US_Chinese[CT:51|CC:25|TT:22] , US_Gujarati[CT:47|CC:25|TT:26]
chr13	rs9573824	6	US_African-30-trios[AG:50|AA:25|GG:23] , CEPH-60-trios[AG:50|AA:26|GG:23] , Han_Chinese[AG:50|AA:24|GG:25] , US_Chinese[AG:49|AA:24|GG:26] , Japanese[AG:47|AA:29|GG:23] , Italian[AG:51|AA:23|GG:24]
chr13	rs630706	6	US_African-30-trios[CT:47|CC:26|TT:25] , CEPH-60-trios[CT:46|CC:28|TT:25] , Han_Chinese[CT:52|CC:24|TT:23] , US_Chinese[CT:48|CC:24|TT:27] , Luhya_Kenyan[CT:46|CC:30|TT:22] , Maasai_Kenyan-60-trios[CT:50|CC:23|TT:26]
chr13	rs1927014	6	US_African-30-trios[CT:49|CC:25|TT:25] , CEPH-60-trios[CT:46|CC:25|TT:28] , Han_Chinese[CT:51|CC:24|TT:24] , Luhya_Kenyan[CT:46|CC:22|TT:30] , Maasai_Kenyan-60-trios[CT:50|CC:26|TT:23] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr13	rs7326426	6	Han_Chinese[CT:49|CC:24|TT:25] , US_Chinese[CT:52|CC:22|TT:24] , US_Gujarati[CT:54|CC:22|TT:22] , Japanese[CT:51|CC:24|TT:24] , Luhya_Kenyan[CT:45|CC:27|TT:27] , Yoruba-60-trios[CT:52|CC:23|TT:24]
chr12	rs2433651	7	US_African-30-trios[AG:50|AA:25|GG:23] , US_Chinese[AG:51|AA:22|GG:25] , US_Gujarati[AG:48|AA:22|GG:28] , Japanese[AG:48|AA:26|GG:24] , Luhya_Kenyan[AG:45|AA:27|GG:26] , Italian[AG:48|AA:25|GG:25] , Yoruba-60-trios[AG:47|AA:22|GG:30]
chr12	rs7297243	7	US_African-30-trios[AG:49|AA:23|GG:26] , CEPH-60-trios[AG:47|AA:27|GG:25] , Han_Chinese[AG:47|AA:24|GG:28] , US_Chinese[AG:47|AA:25|GG:27] , Japanese[AG:50|AA:23|GG:26] , Luhya_Kenyan[AG:46|AA:28|GG:24] , Italian[AG:45|AA:31|GG:23]
chr11	rs10899035	7	US_African-30-trios[CT:47|CC:26|TT:25] , CEPH-60-trios[CT:51|CC:24|TT:24] , Han_Chinese[CT:46|CC:25|TT:28] , US_Chinese[CT:46|CC:28|TT:24] , US_Gujarati[CT:49|CC:26|TT:24] , Luhya_Kenyan[CT:51|CC:22|TT:25] , Yoruba-60-trios[CT:48|CC:28|TT:23]
chr10	rs1624525	7	US_African-30-trios[GT:54|GG:22|TT:22] , Han_Chinese[GT:45|GG:24|TT:30] , Japanese[GT:46|GG:23|TT:30] , Luhya_Kenyan[GT:46|GG:28|TT:24] , US_Mexican-30-trios[GT:46|GG:22|TT:30] , Maasai_Kenyan-60-trios[GT:53|GG:22|TT:23] , Yoruba-60-trios[GT:45|GG:22|TT:31]
chr17	rs4617924	6	US_African-30-trios[AC:47|AA:28|CC:23] , CEPH-60-trios[AC:49|AA:23|CC:27] , Luhya_Kenyan[AC:48|AA:22|CC:28] , US_Mexican-30-trios[AC:46|AA:30|CC:22] , Maasai_Kenyan-60-trios[AC:46|AA:29|CC:23] , Yoruba-60-trios[AC:48|AA:26|CC:24]
chr17	rs2619118	6	CEPH-60-trios[CT:52|CC:23|TT:23] , Han_Chinese[CT:47|CC:25|TT:26] , US_Chinese[CT:48|CC:25|TT:25] , US_Gujarati[CT:46|CC:26|TT:26] , Japanese[CT:50|CC:23|TT:25] , Italian[CT:46|CC:25|TT:27]
chr16	rs9928312	6	US_African-30-trios[AG:47|AA:26|GG:25] , Han_Chinese[AG:46|AA:24|GG:29] , US_Chinese[AG:51|AA:25|GG:22] , US_Gujarati[AG:45|AA:26|GG:27] , Luhya_Kenyan[AG:49|AA:27|GG:22] , Yoruba-60-trios[AG:50|AA:26|GG:22]
chr16	rs31042	6	CEPH-60-trios[AG:50|AA:25|GG:23] , Han_Chinese[AG:45|AA:24|GG:30] , Japanese[AG:46|AA:23|GG:30] , Luhya_Kenyan[AG:46|AA:25|GG:27] , Italian[AG:50|AA:23|GG:26] , Yoruba-60-trios[AG:48|AA:23|GG:27]
chr15	rs12442455	6	US_African-30-trios[CT:49|CC:28|TT:22] , CEPH-60-trios[CT:45|CC:24|TT:30] , Japanese[CT:48|CC:28|TT:23] , US_Mexican-30-trios[CT:45|CC:23|TT:30] , Italian[CT:46|CC:23|TT:29] , Yoruba-60-trios[CT:52|CC:22|TT:25]
chr14	rs10140137	6	US_African-30-trios[AG:46|AA:29|GG:23] , US_Gujarati[AG:48|AA:24|GG:27] , Japanese[AG:50|AA:24|GG:25] , Luhya_Kenyan[AG:51|AA:24|GG:24] , Maasai_Kenyan-60-trios[AG:49|AA:23|GG:27] , Yoruba-60-trios[AG:51|AA:23|GG:25]
chr14	rs4905376	6	US_African-30-trios[CT:53|CC:22|TT:23] , CEPH-60-trios[CT:50|CC:26|TT:22] , US_Gujarati[CT:45|CC:25|TT:28] , Luhya_Kenyan[CT:45|CC:26|TT:27] , Maasai_Kenyan-60-trios[CT:45|CC:26|TT:27] , Yoruba-60-trios[CT:45|CC:25|TT:29]
chr14	rs1026504	6	US_African-30-trios[AC:53|AA:23|CC:22] , CEPH-60-trios[AC:48|AA:29|CC:22] , US_Chinese[AC:50|AA:24|CC:25] , Luhya_Kenyan[AC:53|AA:24|CC:22] , Maasai_Kenyan-60-trios[AC:51|AA:23|CC:24] , Italian[AC:50|AA:25|CC:23]
chr19	rs7250431	6	US_African-30-trios[GT:45|GG:30|TT:23] , CEPH-60-trios[GT:48|GG:24|TT:27] , Han_Chinese[GT:46|GG:25|TT:28] , US_Chinese[GT:47|GG:27|TT:25] , Maasai_Kenyan-60-trios[GT:47|GG:29|TT:23] , Yoruba-60-trios[GT:50|GG:22|TT:26]
chr19	rs273628	6	US_African-30-trios[AG:47|AA:23|GG:28] , US_Gujarati[AG:48|AA:28|GG:22] , Japanese[AG:51|AA:23|GG:25] , Luhya_Kenyan[AG:51|AA:24|GG:24] , Maasai_Kenyan-60-trios[AG:46|AA:26|GG:26] , Italian[AG:45|AA:25|GG:28]
chr18	rs2287434	6	US_African-30-trios[CT:47|CC:26|TT:25] , CEPH-60-trios[CT:50|CC:24|TT:25] , Han_Chinese[CT:45|CC:23|TT:31] , US_Chinese[CT:47|CC:22|TT:30] , Italian[CT:46|CC:29|TT:23] , Yoruba-60-trios[CT:50|CC:23|TT:25]
chr18	rs8096468	6	CEPH-60-trios[CT:49|CC:27|TT:22] , Han_Chinese[CT:51|CC:23|TT:24] , US_Chinese[CT:47|CC:28|TT:24] , Japanese[CT:47|CC:28|TT:24] , US_Mexican-30-trios[CT:45|CC:22|TT:32] , Italian[CT:49|CC:24|TT:25]
chr22	rs165583	7	US_African-30-trios[AC:50|AA:25|CC:23] , CEPH-60-trios[AC:53|AA:23|CC:23] , Han_Chinese[AC:48|AA:25|CC:25] , Japanese[AC:48|AA:24|CC:26] , US_Mexican-30-trios[AC:50|AA:25|CC:23] , Italian[AC:46|AA:24|CC:28] , Yoruba-60-trios[AC:47|AA:26|CC:25]
chr20	rs11698886	7	CEPH-60-trios[AG:50|AA:26|GG:23] , Han_Chinese[AG:46|AA:25|GG:28] , US_Chinese[AG:52|AA:22|GG:24] , US_Gujarati[AG:54|AA:22|GG:22] , Japanese[AG:46|AA:23|GG:30] , Luhya_Kenyan[AG:49|AA:26|GG:24] , US_Mexican-30-trios[AG:50|AA:25|GG:23]
chr21	rs9984703	5	CEPH-60-trios[AG:46|AA:29|GG:24] , US_Chinese[AG:47|AA:28|GG:24] , US_Gujarati[AG:48|AA:28|GG:22] , US_Mexican-30-trios[AG:47|AA:28|GG:23] , Italian[AG:53|AA:23|GG:23]
chr21	rs233621	5	CEPH-60-trios[CT:45|CC:29|TT:25] , Han_Chinese[CT:47|CC:28|TT:24] , US_Gujarati[CT:45|CC:27|TT:26] , Japanese[CT:45|CC:29|TT:25] , Italian[CT:50|CC:25|TT:23]
chr21	rs2823809	5	CEPH-60-trios[AG:46|AA:30|GG:22] , Luhya_Kenyan[AG:47|AA:28|GG:24] , Maasai_Kenyan-60-trios[AG:50|AA:24|GG:25] , Italian[AG:45|AA:29|GG:24] , Yoruba-60-trios[AG:50|AA:24|GG:24]
chr21	rs233622	5	CEPH-60-trios[AG:46|AA:27|GG:25] , Han_Chinese[AG:47|AA:28|GG:24] , US_Gujarati[AG:46|AA:28|GG:25] , Japanese[AG:45|AA:29|GG:25] , Italian[AG:50|AA:24|GG:24]
chr21	rs373037	5	CEPH-60-trios[CT:45|CC:22|TT:32] , Han_Chinese[CT:53|CC:23|TT:23] , Japanese[CT:45|CC:26|TT:28] , Luhya_Kenyan[CT:50|CC:26|TT:22] , Maasai_Kenyan-60-trios[CT:48|CC:23|TT:27]
chr21	rs9647139	5	CEPH-60-trios[AG:46|AA:26|GG:27] , US_Gujarati[AG:54|AA:22|GG:22] , Japanese[AG:47|AA:26|GG:25] , Luhya_Kenyan[AG:54|AA:22|GG:22] , Italian[AG:50|AA:25|GG:23]
chr21	rs2837501	5	CEPH-60-trios[CT:45|CC:24|TT:29] , US_Chinese[CT:47|CC:24|TT:28] , US_Gujarati[CT:48|CC:22|TT:28] , US_Mexican-30-trios[CT:46|CC:25|TT:28] , Italian[CT:53|CC:23|TT:23]
chr21	rs4819214	5	US_African-30-trios[AG:47|AA:25|GG:26] , Luhya_Kenyan[AG:45|AA:24|GG:30] , US_Mexican-30-trios[AG:47|AA:22|GG:29] , Maasai_Kenyan-60-trios[AG:50|AA:26|GG:22] , Yoruba-60-trios[AG:48|AA:26|GG:24]
chr21	rs8130107	5	US_African-30-trios[CT:49|CC:28|TT:22] , US_Chinese[CT:49|CC:27|TT:23] , US_Gujarati[CT:48|CC:27|TT:24] , Japanese[CT:47|CC:23|TT:29] , Yoruba-60-trios[CT:46|CC:27|TT:26]
chr21	rs2009879	5	Han_Chinese[CT:45|CC:28|TT:26] , US_Chinese[CT:47|CC:22|TT:30] , Japanese[CT:50|CC:25|TT:24] , US_Mexican-30-trios[CT:49|CC:26|TT:23] , Yoruba-60-trios[CT:48|CC:22|TT:28]
chr7	rs362813	6	US_African-30-trios[CT:47|CC:25|TT:27] , Han_Chinese[CT:46|CC:24|TT:29] , US_Chinese[CT:50|CC:27|TT:22] , US_Gujarati[CT:48|CC:26|TT:25] , US_Mexican-30-trios[CT:47|CC:22|TT:29] , Italian[CT:48|CC:23|TT:28]
chr7	rs701286	6	US_African-30-trios[AG:53|AA:22|GG:23] , Han_Chinese[AG:48|AA:23|GG:28] , US_Chinese[AG:46|AA:23|GG:30] , Luhya_Kenyan[AG:48|AA:26|GG:25] , US_Mexican-30-trios[AG:45|AA:25|GG:29] , Yoruba-60-trios[AG:50|AA:25|GG:24]
chr7	rs2237628	6	US_African-30-trios[CT:47|CC:25|TT:26] , Han_Chinese[CT:48|CC:24|TT:26] , US_Chinese[CT:50|CC:27|TT:22] , US_Gujarati[CT:48|CC:26|TT:25] , US_Mexican-30-trios[CT:47|CC:22|TT:29] , Italian[CT:48|CC:23|TT:28]
chr7	rs12216600	6	CEPH-60-trios[CT:46|CC:25|TT:27] , US_Chinese[CT:47|CC:25|TT:27] , US_Mexican-30-trios[CT:45|CC:25|TT:29] , Maasai_Kenyan-60-trios[CT:47|CC:25|TT:26] , Italian[CT:51|CC:24|TT:23] , Yoruba-60-trios[CT:48|CC:25|TT:25]
chr7	rs6962246	6	US_African-30-trios[CT:50|CC:26|TT:22] , CEPH-60-trios[CT:45|CC:30|TT:23] , US_Chinese[CT:48|CC:27|TT:24] , US_Gujarati[CT:45|CC:26|TT:27] , Japanese[CT:48|CC:28|TT:23] , Maasai_Kenyan-60-trios[CT:46|CC:30|TT:23]
chr6	rs2517478	6	US_African-30-trios[GT:49|GG:25|TT:25] , Han_Chinese[GT:48|GG:25|TT:25] , Japanese[GT:48|GG:25|TT:25] , Luhya_Kenyan[GT:50|GG:26|TT:22] , Italian[GT:48|GG:24|TT:27] , Yoruba-60-trios[GT:48|GG:23|TT:28]
chr6	rs2523836	6	US_African-30-trios[CT:50|CC:23|TT:25] , Han_Chinese[CT:48|CC:25|TT:25] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:50|CC:26|TT:22] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:48|CC:23|TT:28]
chr6	rs2523843	6	US_African-30-trios[AG:50|AA:22|GG:26] , Han_Chinese[AG:48|AA:25|GG:25] , Japanese[AG:48|AA:25|GG:25] , Luhya_Kenyan[AG:51|AA:25|GG:22] , Italian[AG:48|AA:24|GG:27] , Yoruba-60-trios[AG:48|AA:23|GG:28]
chr6	rs2905759	6	US_African-30-trios[CT:49|CC:25|TT:25] , Han_Chinese[CT:48|CC:25|TT:25] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:50|CC:26|TT:22] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:48|CC:23|TT:28]
chr6	rs2523884	6	US_African-30-trios[CT:48|CC:25|TT:25] , Han_Chinese[CT:50|CC:25|TT:24] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:49|CC:27|TT:22] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr6	rs2517489	6	US_African-30-trios[AG:49|AA:25|GG:25] , Han_Chinese[AG:50|AA:24|GG:25] , Japanese[AG:48|AA:25|GG:25] , Luhya_Kenyan[AG:49|AA:22|GG:27] , Italian[AG:48|AA:27|GG:23] , Yoruba-60-trios[AG:46|AA:30|GG:23]
chr6	rs2256436	6	US_African-30-trios[AG:50|AA:23|GG:25] , Han_Chinese[AG:50|AA:25|GG:24] , Japanese[AG:48|AA:25|GG:25] , Luhya_Kenyan[AG:49|AA:27|GG:22] , Italian[AG:48|AA:24|GG:27] , Yoruba-60-trios[AG:46|AA:23|GG:30]
chr6	rs2517477	6	US_African-30-trios[AG:49|AA:25|GG:25] , Han_Chinese[AG:48|AA:25|GG:25] , Japanese[AG:48|AA:25|GG:25] , Luhya_Kenyan[AG:50|AA:22|GG:26] , Italian[AG:48|AA:27|GG:24] , Yoruba-60-trios[AG:48|AA:28|GG:23]
chr6	rs2523883	6	US_African-30-trios[CT:48|CC:25|TT:25] , Han_Chinese[CT:49|CC:25|TT:24] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:49|CC:27|TT:22] , Italian[CT:47|CC:25|TT:27] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr6	rs2517502	6	US_African-30-trios[CT:50|CC:23|TT:25] , Han_Chinese[CT:50|CC:25|TT:24] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:49|CC:27|TT:22] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr6	rs2523881	6	US_African-30-trios[GT:49|GG:25|TT:25] , Han_Chinese[GT:50|GG:24|TT:25] , Japanese[GT:48|GG:25|TT:25] , Luhya_Kenyan[GT:49|GG:22|TT:27] , Italian[GT:48|GG:27|TT:24] , Yoruba-60-trios[GT:46|GG:30|TT:23]
chr6	rs2517500	6	US_African-30-trios[CT:50|CC:23|TT:25] , Han_Chinese[CT:50|CC:25|TT:24] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:49|CC:27|TT:23] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr6	rs154989	6	US_African-30-trios[AG:45|AA:23|GG:30] , CEPH-60-trios[AG:46|AA:26|GG:26] , Han_Chinese[AG:47|AA:29|GG:23] , US_Gujarati[AG:50|AA:22|GG:26] , Japanese[AG:47|AA:25|GG:26] , Italian[AG:49|AA:24|GG:25]
chr6	rs2535311	6	US_African-30-trios[AG:50|AA:23|GG:25] , Han_Chinese[AG:48|AA:25|GG:25] , Japanese[AG:48|AA:25|GG:25] , Luhya_Kenyan[AG:50|AA:26|GG:22] , Italian[AG:48|AA:24|GG:27] , Yoruba-60-trios[AG:48|AA:23|GG:28]
chr6	rs2535310	6	US_African-30-trios[AC:47|AA:22|CC:29] , CEPH-60-trios[AC:51|AA:25|CC:22] , Han_Chinese[AC:48|AA:25|CC:25] , Japanese[AC:48|AA:25|CC:25] , Luhya_Kenyan[AC:53|AA:22|CC:24] , Yoruba-60-trios[AC:47|AA:23|CC:29]
chr6	rs2106067	6	US_African-30-trios[CT:50|CC:25|TT:24] , Han_Chinese[CT:48|CC:25|TT:25] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:50|CC:22|TT:26] , Italian[CT:48|CC:27|TT:24] , Yoruba-60-trios[CT:48|CC:28|TT:23]
chr6	rs2523891	6	US_African-30-trios[CT:49|CC:25|TT:25] , Han_Chinese[CT:50|CC:25|TT:24] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:49|CC:27|TT:22] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr6	rs2535315	6	US_African-30-trios[CT:51|CC:22|TT:25] , Han_Chinese[CT:48|CC:25|TT:25] , Japanese[CT:48|CC:25|TT:25] , Luhya_Kenyan[CT:50|CC:26|TT:22] , Italian[CT:48|CC:24|TT:27] , Yoruba-60-trios[CT:48|CC:23|TT:28]
chr6	rs154988	6	US_African-30-trios[AG:45|AA:30|GG:23] , CEPH-60-trios[AG:46|AA:26|GG:26] , Han_Chinese[AG:47|AA:23|GG:29] , US_Gujarati[AG:50|AA:26|GG:22] , Japanese[AG:47|AA:26|GG:25] , Italian[AG:49|AA:25|GG:24]
chr5	rs4958301	7	US_African-30-trios[GT:50|GG:22|TT:26] , CEPH-60-trios[GT:51|GG:22|TT:25] , US_Gujarati[GT:51|GG:22|TT:25] , US_Mexican-30-trios[GT:45|GG:29|TT:25] , Maasai_Kenyan-60-trios[GT:49|GG:27|TT:22] , Italian[GT:49|GG:24|TT:25] , Yoruba-60-trios[GT:48|GG:23|TT:27]
chr4	rs1521696	6	US_African-30-trios[AT:45|AA:25|TT:29] , CEPH-60-trios[AT:50|AA:26|TT:22] , Han_Chinese[AT:51|AA:24|TT:24] , US_Chinese[AT:47|AA:22|TT:30] , US_Gujarati[AT:45|AA:27|TT:26] , Japanese[AT:46|AA:30|TT:23]
chr4	rs720	6	US_African-30-trios[CT:45|CC:32|TT:22] , CEPH-60-trios[CT:53|CC:23|TT:23] , US_Chinese[CT:45|CC:24|TT:30] , US_Gujarati[CT:46|CC:26|TT:26] , Maasai_Kenyan-60-trios[CT:48|CC:28|TT:22] , Italian[CT:48|CC:24|TT:27]
chr4	rs2160043	6	US_African-30-trios[CT:52|CC:24|TT:22] , CEPH-60-trios[CT:49|CC:24|TT:25] , Han_Chinese[CT:47|CC:23|TT:29] , US_Chinese[CT:45|CC:31|TT:22] , Luhya_Kenyan[CT:45|CC:24|TT:30] , Maasai_Kenyan-60-trios[CT:47|CC:28|TT:23]
chr3	rs6800661	6	US_African-30-trios[AC:45|AA:32|CC:22] , Han_Chinese[AC:46|AA:23|CC:30] , US_Chinese[AC:51|AA:24|CC:24] , US_Gujarati[AC:45|AA:24|CC:30] , Japanese[AC:48|AA:26|CC:24] , Maasai_Kenyan-60-trios[AC:49|AA:23|CC:26]
chr3	rs744876	6	US_African-30-trios[AG:46|AA:28|GG:25] , US_Chinese[AG:52|AA:22|GG:24] , Luhya_Kenyan[AG:50|AA:25|GG:24] , Maasai_Kenyan-60-trios[AG:48|AA:23|GG:28] , Italian[AG:45|AA:29|GG:24] , Yoruba-60-trios[AG:49|AA:25|GG:25]
chr3	rs1450344	6	US_African-30-trios[AG:50|AA:22|GG:26] , Han_Chinese[AG:47|AA:25|GG:26] , US_Chinese[AG:45|AA:24|GG:30] , US_Gujarati[AG:51|AA:22|GG:25] , US_Mexican-30-trios[AG:46|AA:25|GG:28] , Yoruba-60-trios[AG:45|AA:30|GG:23]
chr3	rs4263316	6	US_African-30-trios[CT:45|CC:32|TT:22] , CEPH-60-trios[CT:48|CC:25|TT:26] , US_Gujarati[CT:50|CC:24|TT:25] , Luhya_Kenyan[CT:49|CC:26|TT:24] , US_Mexican-30-trios[CT:47|CC:26|TT:25] , Yoruba-60-trios[CT:47|CC:26|TT:25]
chr3	rs949632	6	US_African-30-trios[AG:45|AA:22|GG:32] , Han_Chinese[AG:46|AA:30|GG:23] , US_Chinese[AG:51|AA:24|GG:24] , US_Gujarati[AG:45|AA:30|GG:24] , Japanese[AG:48|AA:24|GG:26] , Maasai_Kenyan-60-trios[AG:49|AA:26|GG:23]
chr3	rs4073109	6	Han_Chinese[CT:51|CC:23|TT:24] , US_Chinese[CT:47|CC:22|TT:30] , US_Gujarati[CT:45|CC:26|TT:28] , Luhya_Kenyan[CT:45|CC:28|TT:25] , US_Mexican-30-trios[CT:49|CC:28|TT:22] , Yoruba-60-trios[CT:50|CC:26|TT:22]
chr3	rs7609911	6	Han_Chinese[AC:48|AA:25|CC:25] , US_Chinese[AC:51|AA:22|CC:25] , Luhya_Kenyan[AC:45|AA:27|CC:26] , US_Mexican-30-trios[AC:53|AA:23|CC:22] , Maasai_Kenyan-60-trios[AC:49|AA:27|CC:23] , Italian[AC:45|AA:29|CC:24]
chr3	rs11710939	6	Han_Chinese[AG:47|AA:26|GG:25] , US_Chinese[AG:50|AA:25|GG:24] , Japanese[AG:48|AA:28|GG:23] , Luhya_Kenyan[AG:49|AA:25|GG:25] , Maasai_Kenyan-60-trios[AG:45|AA:25|GG:28] , Italian[AG:45|AA:24|GG:29]
chr3	rs6437577	6	US_African-30-trios[AC:45|AA:32|CC:22] , CEPH-60-trios[AC:48|AA:25|CC:26] , US_Gujarati[AC:50|AA:24|CC:25] , Luhya_Kenyan[AC:49|AA:26|CC:24] , US_Mexican-30-trios[AC:47|AA:26|CC:25] , Yoruba-60-trios[AC:47|AA:26|CC:25]
chr3	rs16863691	6	Han_Chinese[GT:53|GG:23|TT:23] , Japanese[GT:50|GG:26|TT:23] , Luhya_Kenyan[GT:48|GG:22|TT:28] , US_Mexican-30-trios[GT:46|GG:30|TT:22] , Maasai_Kenyan-60-trios[GT:50|GG:23|TT:26] , Yoruba-60-trios[GT:50|GG:22|TT:26]
chr3	rs7616178	6	US_African-30-trios[CT:47|CC:28|TT:23] , CEPH-60-trios[CT:48|CC:23|TT:28] , Han_Chinese[CT:46|CC:27|TT:26] , US_Gujarati[CT:46|CC:24|TT:29] , Luhya_Kenyan[CT:47|CC:28|TT:24] , US_Mexican-30-trios[CT:47|CC:28|TT:23]
chr3	rs12634576	6	US_African-30-trios[AG:45|AA:32|GG:22] , Han_Chinese[AG:46|AA:23|GG:30] , US_Chinese[AG:51|AA:24|GG:24] , US_Gujarati[AG:46|AA:25|GG:27] , Japanese[AG:48|AA:26|GG:24] , Maasai_Kenyan-60-trios[AG:49|AA:23|GG:26]
chr3	rs12634593	6	US_African-30-trios[GT:45|GG:22|TT:32] , Han_Chinese[GT:46|GG:30|TT:23] , US_Chinese[GT:51|GG:24|TT:24] , US_Gujarati[GT:45|GG:27|TT:26] , Japanese[GT:48|GG:24|TT:26] , Maasai_Kenyan-60-trios[GT:49|GG:26|TT:23]
chr3	rs6551382	6	US_African-30-trios[AG:46|AA:29|GG:23] , Han_Chinese[AG:46|AA:23|GG:30] , US_Chinese[AG:51|AA:24|GG:24] , US_Gujarati[AG:45|AA:24|GG:30] , Japanese[AG:48|AA:26|GG:24] , Maasai_Kenyan-60-trios[AG:49|AA:23|GG:26]
chr3	rs2141603	6	US_African-30-trios[AG:45|AA:22|GG:32] , CEPH-60-trios[AG:48|AA:26|GG:25] , US_Gujarati[AG:50|AA:25|GG:24] , Luhya_Kenyan[AG:48|AA:24|GG:26] , US_Mexican-30-trios[AG:47|AA:25|GG:26] , Yoruba-60-trios[AG:47|AA:25|GG:26]
chr3	rs6764404	6	US_African-30-trios[CT:45|CC:22|TT:32] , Han_Chinese[CT:46|CC:30|TT:23] , US_Chinese[CT:51|CC:24|TT:24] , US_Gujarati[CT:45|CC:30|TT:24] , Japanese[CT:48|CC:24|TT:26] , Maasai_Kenyan-60-trios[CT:49|CC:26|TT:23]
chr2	rs2419407	7	US_African-30-trios[CT:49|CC:22|TT:28] , Han_Chinese[CT:50|CC:26|TT:23] , Japanese[CT:51|CC:23|TT:25] , Luhya_Kenyan[CT:47|CC:26|TT:25] , Maasai_Kenyan-60-trios[CT:47|CC:26|TT:25] , Italian[CT:46|CC:29|TT:23] , Yoruba-60-trios[CT:49|CC:23|TT:27]
chr1	rs12136079	7	US_African-30-trios[AG:46|AA:26|GG:26] , Japanese[AG:47|AA:24|GG:28] , Luhya_Kenyan[AG:45|AA:28|GG:25] , US_Mexican-30-trios[AG:46|AA:28|GG:25] , Maasai_Kenyan-60-trios[AG:50|AA:24|GG:25] , Italian[AG:49|AA:25|GG:24] , Yoruba-60-trios[AG:47|AA:29|GG:22]
chr9	rs1335143	6	US_African-30-trios[AG:45|AA:23|GG:30] , CEPH-60-trios[AG:48|AA:24|GG:26] , US_Gujarati[AG:48|AA:24|GG:27] , Luhya_Kenyan[AG:50|AA:26|GG:22] , US_Mexican-30-trios[AG:45|AA:25|GG:29] , Yoruba-60-trios[AG:50|AA:22|GG:26]
chr9	rs11103233	6	US_African-30-trios[CT:49|CC:23|TT:26] , CEPH-60-trios[CT:51|CC:25|TT:22] , US_Chinese[CT:45|CC:22|TT:31] , US_Gujarati[CT:48|CC:28|TT:22] , Maasai_Kenyan-60-trios[CT:47|CC:23|TT:28] , Italian[CT:50|CC:23|TT:25]
chr8	rs10102929	6	US_African-30-trios[CT:47|CC:22|TT:29] , CEPH-60-trios[CT:45|CC:27|TT:27] , Han_Chinese[CT:50|CC:24|TT:25] , Luhya_Kenyan[CT:48|CC:27|TT:24] , US_Mexican-30-trios[CT:49|CC:25|TT:25] , Maasai_Kenyan-60-trios[CT:50|CC:23|TT:25]
chr8	rs12541274	6	US_African-30-trios[CT:45|CC:26|TT:28] , CEPH-60-trios[CT:46|CC:22|TT:30] , US_Chinese[CT:48|CC:24|TT:27] , Japanese[CT:47|CC:27|TT:25] , Luhya_Kenyan[CT:46|CC:30|TT:22] , Yoruba-60-trios[CT:46|CC:23|TT:30]
chr8	rs7828662	6	US_African-30-trios[CT:47|CC:25|TT:26] , US_Gujarati[CT:46|CC:30|TT:22] , Luhya_Kenyan[CT:50|CC:22|TT:26] , US_Mexican-30-trios[CT:50|CC:23|TT:25] , Maasai_Kenyan-60-trios[CT:46|CC:23|TT:29] , Yoruba-60-trios[CT:49|CC:22|TT:27]
chr8	rs4875552	6	CEPH-60-trios[CT:48|CC:25|TT:25] , Han_Chinese[CT:45|CC:23|TT:31] , Japanese[CT:52|CC:24|TT:23] , US_Mexican-30-trios[CT:45|CC:30|TT:23] , Maasai_Kenyan-60-trios[CT:48|CC:25|TT:26] , Italian[CT:46|CC:27|TT:25]
chr8	rs11136962	6	Han_Chinese[CT:48|CC:26|TT:24] , US_Chinese[CT:50|CC:25|TT:24] , US_Gujarati[CT:46|CC:30|TT:22] , Japanese[CT:47|CC:24|TT:28] , Luhya_Kenyan[CT:45|CC:24|TT:30] , Yoruba-60-trios[CT:47|CC:30|TT:22]
chr8	rs7842255	6	US_African-30-trios[CT:53|CC:22|TT:23] , CEPH-60-trios[CT:51|CC:24|TT:24] , Japanese[CT:46|CC:28|TT:25] , US_Mexican-30-trios[CT:50|CC:22|TT:26] , Maasai_Kenyan-60-trios[CT:53|CC:23|TT:23] , Yoruba-60-trios[CT:46|CC:31|TT:22]
chr8	rs7824027	6	US_African-30-trios[AG:53|AA:23|GG:22] , CEPH-60-trios[AG:51|AA:24|GG:24] , Japanese[AG:46|AA:25|GG:28] , US_Mexican-30-trios[AG:50|AA:26|GG:22] , Maasai_Kenyan-60-trios[AG:53|AA:23|GG:23] , Yoruba-60-trios[AG:46|AA:22|GG:31]

Table S7

Table S7. Primers for the SNPs in HapMap sequenom assay

WELL	SNP_ID	1st-PCR Primer	2nd-PCR Primer	Extension Primer
W1	rs12541274	ACGTTGGATGAGAGCAGAGCCCATCGAGA	ACGTTGGATGACCTTTACAAGGTCCTGCAC	CTCACTCCTTCTTGGCT
W1	rs10102929	ACGTTGGATGAGGAAGGTGGTCATAGGTTC	ACGTTGGATGAAGCCCGTAAGTCTATGAGG	AGTCTATGAGGGATCCA
W1	AMELXY.a	ACGTTGGATGCACTGGGATGTGGTGATGAG	ACGTTGGATGGGCTGCACCACCAAATCATC	CCACCAAATCATCCCCGTG
W1	rs6645103	ACGTTGGATGCCACCACTGCTCACATAATC	ACGTTGGATGACTTTTCTGACAGTGCAGCC	GCCTGCAGATAACATCTAA
W1	rs10899035	ACGTTGGATGCTCTCTTCCATGCCTTTTAC	ACGTTGGATGTAAAATGGAGTAGTACAGGC	GGAGTAGTACAGGCATAAAA
W1	rs7297243	ACGTTGGATGGGCTAACATTTTAACTCAGG	ACGTTGGATGCTTCTTAGACTGACAAAATGC	TGACAAAATGCTTGTATCTTT
W1	rs4958301	ACGTTGGATGGCTATTTGATAGACCAGGTG	ACGTTGGATGCTTCTGACCACTTCTCAAAC	CACTACACTCTAATCCAAGCTA
W1	rs744876	ACGTTGGATGAGAAGCATGCACATTCAGGG	ACGTTGGATGAGTGCCTTGTATGGGTGTG	GGTGGCTAGTCCTCAAATATGG
W1	rs2433651	ACGTTGGATGGTTGCCTGAGATACGTTTCG	ACGTTGGATGATGCAGGACCTCTTCCTATG	CAGGACCTCTTCCTATGTCATGC
W1	rs2535310	ACGTTGGATGTCATGTGCTTGGATCCAAGG	ACGTTGGATGGCTGCTTGTAACAGAGACAC	ACACAAATGCCACTGGTTTAAATT
W1	rs11698886	ACGTTGGATGGCGCACAGGAGGTTCCAG	ACGTTGGATGGGCATGCCACACAGGCAG	GGCATGCCACACAGGCAGAGAGAA
W1	rs1624525	ACGTTGGATGTTTTAATTAGGAGAAGCAG	ACGTTGGATGAATGAGGATTCCAACAGGTC	GCTGCACAAATAATCTCTCAACCACA
W2	rs120434	ACGTTGGATGAGGGAGCCAGTTGGCAACAT	ACGTTGGATGGCCTTTCAATGCCTTGTTCC	TGTGCTGGAATGCTGAT
W2	rs233621	ACGTTGGATGTGGAAAGCTGGAGGCCTAAG	ACGTTGGATGTTTCTTTTGGTTCTCAGGGC	CTCAGTCTAGATCCACAA
W2	rs701286	ACGTTGGATGGGAAAGCTGTCAGGAATTG	ACGTTGGATGGGAAAGTATAGCACAAAAGC	GCAAAGACAAATGGGAGA
W2	rs10140137	ACGTTGGATGCGGCAAATCCTGTTTCTATG	ACGTTGGATGGTAACAAATAATGTTTTAGTC	TTAGTCTTAAATGTACCCCA
W2	rs2051068	ACGTTGGATGCTGGAAACCATTTCAGAATGC	ACGTTGGATGCTGAAGGTAAAGATTCAAG	TCAAGAGCGCACAGGTATAG
W2	rs1365740	ACGTTGGATGGTACTTACCTCCTGAGGTAG	ACGTTGGATGCAGCACTGGGATGACTTACA	TACACCTTTTTCCTCTTCATT
W2	rs1434199	ACGTTGGATGCCTCTGCCCCCTAATTTCTT	ACGTTGGATGTAGGGAGCTGAATCAATAC	AGCTGAATCAATACATATCACT
W2	rs2419407	ACGTTGGATGAACCTGACAAGTCAGTCTGC	ACGTTGGATGTCTCCAGTCTTTCAGAGAG	AGTCTTTCAGAGAGTTGATCAT
W2	rs273628	ACGTTGGATGGGGAATGTTTTGTCAGTGTC	ACGTTGGATGCTCATGGATAAATAACTGGG	GTGAAATTGCTGGATCATAGGA
W2	rs6800661	ACGTTGGATGTCTGTAAGTGTTTTAATAG	ACGTTGGATGCTAATCTTTCTCCAAATGC	AATCTTTCTCCAAATGCAAGGTTT
W2	rs31042	ACGTTGGATGATTACCTGGGCATTGTTTCC	ACGTTGGATGGTTCAGAAGTATTTCAGGAC	GACACATTTACCAGTGAGTTAGGC
W2	rs1026504	ACGTTGGATGCTCCATTTTGACCTTGTACC	ACGTTGGATGCCATTTACTTCGTGCAGCAT	TGCAGCATGATTAGAAGTAAAGGA
W3	rs8096468	ACGTTGGATGTCACTCATCCACATGGCAGG	ACGTTGGATGAGCACAGCGCACCACACCA	CCTGGGCTGAGCAACCA
W3	rs2619118	ACGTTGGATGTCTGTCTCTCCAGCTACACT	ACGTTGGATGACCTGGGAGACGCGAAAGGA	GCGAAAGGAGGAAATGC
W3	rs2287434	ACGTTGGATGGGTTAAAGGGACTCCATCAG	ACGTTGGATGGGTACTAAGACAATTTTGGC	ACAATTTTGGCAAGACTG
W3	rs362813	ACGTTGGATGGTCTTCCCAATGTCTGTCTC	ACGTTGGATGCAGCCATCACTACCATCATC	TCAAAAACACTATTCTGCC
W3	rs265005	ACGTTGGATGAGCCTCTCTGCTAAGCTACA	ACGTTGGATGTTGTTTGTCTGCTCAACTGG	aTGCTCAACTGGACTATAAA
W3	rs9573824	ACGTTGGATGTCTAAGGCATCCTGGCAGTA	ACGTTGGATGAAAGCATTATAGTTGGGAGG	AAAACAAACAGGGTAGCAAG
W3	rs11103233	ACGTTGGATGATAAACCATTGCCATCGAAG	ACGTTGGATGAGGAACTCACGGGAGGACAA	TGGGCTTTTGCTCTTACTACTC
W3	rs1521696	ACGTTGGATGTTCATTGTATTCACTGTTC	ACGTTGGATGGTCTAGGAATCTTCCAAATG	TAGATTTGGACTTTAAGCAATC
W3	rs2347790	ACGTTGGATGATGCTCCACCCACACAGTTC	ACGTTGGATGATAACCTATGCAACAGCTGG	gTGGAAGATGTGCGTATGCCTTA
W3	rs7250431	ACGTTGGATGGAGAAAGAAGGATGCTGCTC	ACGTTGGATGACTTCTCCTGGCTCCTATGC	GGGCTGGTCCACCCTCCTGTCTCA
W3	rs12136079	ACGTTGGATGTGGAAAACGGAAGAAAGTG	ACGTTGGATGTTCTGGAACTCCTGCTATTC	TATTCAAATAAATACTGGCTGTTA
W3	rs1335143	ACGTTGGATGATACTTATGTCCAAGGCAA	ACGTTGGATGCTTCTTATCCATAATAATCCC	TGCAAAATACATTTATTGCATCTTTA
W4	rs154989	ACGTTGGATGATTGCAACCATCGAGAGGAC	ACGTTGGATGTTCCCTCTTGTATCTCTGGC	TCGCATAGCATCAGGCT
W4	rs165583	ACGTTGGATGCCCTTTCCCTGAACAGAAGC	ACGTTGGATGCTCCCGTGGTTTTATAGCAG	GGCGAAGGAAGGAGACC
W4	rs630706	ACGTTGGATGGTATCTAGCATTATCGAGAC	ACGTTGGATGCACAAACTTGATGTCTCCCC	ATGTCTCCCCATTTACTTA
W4	rs12442455	ACGTTGGATGATATGTCACGCATAGCCCAG	ACGTTGGATGAACCTGGCTCATGGATGACA	AGTGGATTATCCTTGGAAT
W4	rs2160043	ACGTTGGATGAGTATTGACAAGGATGTGG	ACGTTGGATGGTTTTCCAAAGTAGCTACAG	CAATTTATATTCCTACCAGCA
W4	rs4617924	ACGTTGGATGCTCGTGGTACCTGAAAATTC	ACGTTGGATGGCAGTTATGTGTACATTAG	ATTTAGACTGAGTTTTGTCTT
W4	AMELXY.b	ACGTTGGATGACCACCAAATCATCCCCGTG	ACGTTGGATGAGCTGGCACCACTGGGATGT	ACCACTGGGATGTGGTGATGAG
W4	rs9928312	ACGTTGGATGTAGTTCAGATCCCAACTGGC	ACGTTGGATGAAAAGCGGTGGCCAGGAGTG	GGCACAGGATGGCCCCAAAGCAT

Table S8

Table S8. Cross-validation of MS based genetic fingerprinting assay

Sample	Call Type	rs2287434	rs4617924	rs265005	rs11698886	rs1521696	rs120434	rs6800661
ARO	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	NoCall	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	GC	CC
HT29	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	Het	NoCall	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	GC	CC
8305c	SNPchip Haplotype	Hom2	Het	Hom2	Hom1	Het	Hom2	Hom1
	MS-Genetic fingerprinting	CT	CA	TT	GG	AT	GG	CA
BHP10_3	SNPchip Haplotype	Het	Hom1	Hom2	Het	Hom2	Hom2	Het
	MS-Genetic fingerprinting	CT	AA	TT	GA	TT	GG	CA
BHP18_21	SNPchip Haplotype	Het	Hom1	Hom2	Het	Hom2	Hom2	Het
	MS-Genetic fingerprinting	CT	AA	TT	GA	TT	GG	CA
BHP2_7	SNPchip Haplotype	Het	Hom1	Hom2	Het	Hom2	Hom2	NoCall
	MS-Genetic fingerprinting	CT	AA	TT	GA	TT	GG	CA
BHP7_13	SNPchip Haplotype	Het	Hom1	Hom2	Het	Hom2	Hom2	Het
	MS-Genetic fingerprinting	CT	AA	TT	GA	TT	GG	CA
DRO	SNPchip Haplotype	NoCall	Hom1	Hom2	Hom1	Hom1	NoCall	Het
	MS-Genetic fingerprinting	CT	AA	TT	GG	AA	GG	CA
KAK-1	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	NoCall	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	NoCall	CC
KAT-10	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	Hom2	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	NoCall	CC
KAT-18	SNPchip Haplotype	Hom2	Het	Hom2	NoCall	NoCall	Hom2	Hom1
	MS-Genetic fingerprinting	CC	CA	TT	GG	AT	GG	CC
KAT-4	SNPchip Haplotype	NoCall	Hom1	Hom2	Hom1	NoCall	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	NoCall	CC
KAT-5	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	Hom2	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	NoCall	CC
KAT-50	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	Hom2	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	NoCall	CC
KAT-7	SNPchip Haplotype	Hom1	Hom1	Hom2	Hom1	Hom2	Het	Hom1
	MS-Genetic fingerprinting	TT	AA	TT	GG	TT	NoCall	CC
NPA	SNPchip Haplotype	Het	Hom1	Het	Hom2	Hom1	Het	Het
	MS-Genetic fingerprinting	NoCall	AA	CT	AA	AA	GC	CA

Hom1=homozygous for allele 1; Hom2=homozygous for allele 2; Het=heterozygous

Table S9

Table S9. HapMap for cell line identity typing.

Cell Lines	Match	P Value
*SW620	SW480	3.97E-11
§LS174T	LS180	7.73E-06
#DLD-1	HCT-8	1.08E-12
#HCT-15	DLD-1	2.06E-05
αCOLO-201	COLO-205	1.18E-12
βIS1	IS2	4.83E-08
ªHT-29	WiDr	9.81E-06
¶C2BBe1	Caco-2	5.74E-08
RKO-ATCC	RKO	6.24E-14

*SW480 developed from primary and SW620 established at 6 months follow-up from lymph node metastasis of the same patient (Leibovitz et al., & Gagos et al., PMID: 1000501 & 7763008)

§ LS174T is a trypsinized variant of LS180 cell line (Tom et al., PMID: 1262041)

Using DNA fingerprinting, these four cell lines found to have an identical genetic background (Vermeulen et al., & Chen et al., PMID:7621404 & 9809040)

α Tumor cell lines (COLO 201 and COLO 205) have been established from ascites fluid obtained from a male patient with adenocarcinoma of the colon (Semple et al., PMID:565251)

β Cell lines developed from a single patient (Cajot et al., PMID:9205061)

ª Chromosomal and isozymic evidence indicating that WiDr is a derivative of HT-29 (chen at al., PMID:3472642)

¶ Clone derived from Caco-2 (Peterson et al., PMID: 1506435)