SUPPLEMENTARY TABLES

Supplementary Table 1. Targeted Genes

ABL1	CTNNB1	IDH2	MYCL1	RET
ABL2	EGFR	IGF1R	MYCN	RICTOR
AKT1	EPHA3	IKBKE	NF1	RUNX1
AKT2	EPHA5	IKZF1	NF2	RUNX1T1
AKT3	EPHB6	JAK2	NKX2-1	SMAD2
ALK	ERBB2	JAK3	NOTCH1	SMAD3
APC	ERBB3	KDR	NOTCH2	SMAD4
ATM	ERBB4	KEAP1	NOTCH3	SMARCA4
AURKA	FAM123B	KIAA0774	NOTCH4	SMARCB1
BCL2	FBXW7	KIAA1303	NPM1	SMO
BRAF	FGFR1	KIT	NRAS	SOCS1
BRCA1	FGFR2	KRAS	NTRK1	SRC
BRCA2	FGFR3	MAP2K1	NTRK2	STK11
CCND1	FGFR4	MAP2K2	NTRK3	SUFU
CCNE1	FHIT	MAP2K4	PAX5	TCF4
CDC73	FKBP9	MCL1	PDGFRA	TERT
CDH1	FLT1	MDM2	PDGFRB	TET2
CDK4	FLT3	MDM4	PDPK1	TGFBR2
CDK6	FLT4	MEN1	PIK3CA	TNFAIP3
CDK8	FRAP1	MET	PIK3R1	TOP1
CDKN1A	GATA1	MITF	PTCH1	TP53
CDKN2A	GNAQ	MLH1	PTEN	TSC1
CEBPA	GNAS	MLL	PTK2B	TSC2
CHEK1	GUCY1A2	MPL	PTPN11	TSHR
CHEK2	HNF1A	MRE11A	PTPRD	VHL
CREBBP	HRAS	MSH2	RAF1	WT1
CRKL	HSP90AA1	MSH6	RB1	ZNF668
CSF1R	IDH1	MYC	REL	

Exons from 137 genes were targeted for hybrid selection and massively parallel sequencing.

Supplementary Table 2. Pharmacogenomic Loci

GENE	LOCUS	RELEVANT DRUGS
ABCB1	chr7:86976581	Idarubicin, AraC, Paclitaxel
ABCB1	chr7:86998554	Idarubicin, AraC, Taxanes, Platinums
ABCC2	chr10:101610761	Docetaxel
ABCC4	chr13:94613416	6MP
ABCG2	chr4:89252551	Methotrexate
ABCG2	chr4:89271347	Gefitinib
ABCG2	chr4:89274403	Methotrexate
C1orf144	chr1:16578662	Daunorubicin
CYP1B1	chr2:38151707	Daunorubicin, Paclitaxel
CYP2C19	chr10:96509051	Tamoxifen
CYP2C19	chr10:96511647	Tamoxifen
CYP2C8	chr10:96786964	Paclitaxel
CYP2C8	chr10:96788739	Paclitaxel
CYP2C8	chr10:96808096	Paclitaxel
CYP2C8	chr10:96808109	Paclitaxel
CYP2C8	chr10:96817020	Paclitaxel
CYP2D6	chr22:40853554	Tamoxifen
CYP2D6	chr22:40853749	Tamoxifen
CYP2D6	chr22:40853887	Tamoxifen
CYP2D6	chr22:40854122	Tamoxifen
CYP2D6	chr22:40854188	Tamoxifen
CYP2D6	chr22:40854891	Tamoxifen
CYP2D6	chr22:40855030	Tamoxifen
CYP2D6	chr22:40855078	Tamoxifen
CYP2D6	chr22:40855716	Tamoxifen
CYP2D6	chr22:40856638	Tamoxifen
CYP3A4	chr7:99196395	Multiple
CYP3A4	chr7:99196460	Multiple
CYP3A4	chr7:99197606	Multiple
CYP3A4	chr7:99204017	Multiple
CYP3A4	chr7:99204029	Multiple
CYP3A4	chr7:99205328	Multiple
CYP3A4	chr7:99205363	Multiple
CYP3A4	chr7:99219597	Multiple
CYP3A4	chr7:99220032	Multiple
CYP3A5	chr7:99088330	Multiple
CYP3A5	chr7:99100771	Multiple
CYP3A5	chr7:99108475	Multiple
DPYD	chr1:97688202	5-FU
DPYD	chr1:97753983	5-FU
DPYD	chr1:97937679	5-FU, Capecitabine
DPYD	chr1:98121473	5-FU
ERCC2	chr19:50546759	5-FU
ESR1	chr6:152205074	Tamoxifen
ESR2	chr14:63769569	Tamoxifen

FCGR3A	chr1:159781166	Cetuximab
FGFR4	chr5:176452849	Multiple
GSTP1	chr11:67109265	Multiple
GSTP1	chr11:67110155	Multiple
ITPA	chr20:3141842	6MP
LRP2	chr2:169719231	Cisplatin
MAN1B1	chr9:139102689	Daunorubicin
MTHFR	chr1:11777044	Methotrexate
MTHFR	chr1:11777063	Methotrexate
MTHFR	chr1:11778965	Methotrexate
NQO1	chr16:68302646	Cisplatin, Doxorubicin, Anthracyclines
NRP2	chr2:206360545	Daunorubicin
SLC19A1	chr21:45782222	Methotrexate
SLC22A2	chr6:160590272	Cisplatin
SLCO1B3	chr12:20936961	Docetaxel
SOD2	chr6:160033862	Cyclophosphamide
SULT1A1	chr16:28524986	Tamoxifen
SULT1A1	chr16:28525015	Tamoxifen
SULT1A1	chr16:28528073	Tamoxifen
SULT1A1	chr16:28528301	Tamoxifen
TMPT	chr6:18247207	Purines
TPMT	chr6:18238897	6MP
TPMT	chr6:18238991	6MP
TPMT	chr6:18251934	6MP
TYMS	chr18:647646	5-FU
TYMS	chr18:663451	5-FU
UGT1A1	chr2:234255266	Irinotecan
UGT1A1	chr2:234255709	Irinotecan
UGT1A1	chr2:234330398	Irinotecan
UGT1A1	chr2:234330521	Irinotecan
UGT1A1	chr2:234333620	Irinotecan
UGT1A1	chr2:234333883	Irinotecan
UGT1A1	chr2:234334358	Irinotecan
UMPS	chr3:125939432	5-FU

Abbreviations: 6MP = 6-mercaptopurine, 5FU = 5-fluorouracil, AraC = cytarabine

79 loci in 34 genes which might predict sensitivity or resistance to cancer therapies were also targeted for sequencing. Drugs related to polymorphisms at these loci are listed.

Supplementary Table 3. Known genomic alterations in 10 cancer cell lines.

Cell Line	Primary Disease	Somatic Alterations in COSMIC
HL-60	Acute Myelogenous Leukemia	CDKN2A, NRAS, TP53
HT-29	Colon Cancer	APC, BRAF, PIK3CA, SMAD4, TP53
MCF7	Breast Cancer	CDKN2A, PIK3CA
MDA-MB-231	Breast Cancer	BRAF, CDKN2A, KRAS, NF2, TP53
NCI-H1395	Lung Adenocarcinoma	BRAF, STK11
NCI-H358	Lung Bronchoalveolar Carcinoma	KRAS
NCI-H69	Lung Small Cell Carcinoma	PIK3CA, RB1, TP53
ZR-75-30	Breast Cancer	CDH1
PC-3	Prostate Cancer	PTEN, TP53
SK-MEL-2	Melanoma	BRAF, EGFR, TP53

DNA from 10 cancer cell lines were sequenced. Known genomic alterations in these cell lines were obtained from the COSMIC database.

Supplementary Table 4. Summary of sequencing results for cell lines

CELL LINE	TUMOR TYPE	PF READS	PERCENT OF TOTAL PF READS IN POOL	PERCENT SELECTED BASES	MEAN TARGET COVERAGE	PERCENT OF TARGET BASES WITH AT LEAST 30X COVERAGE
HAPMAP	N/A	14,316,958	8%	55%	469	96%
HT-29 (50% input)	Colon	7,809,802	5%	60%	295	95%
MCF7	Breast	14,643,284	9%	60%	540	96%
MDA-MB-231	Breast	15,609,408	9%	59%	568	96%
ZR-75-30	Breast	17,709,492	10%	57%	593	96%
HL-60	Leukemia	15,110,292	9%	58%	543	96%
NCI-H69	Lung	12,108,024	7%	61%	451	96%
NCI-358	Lung	12,119,682	7%	61%	441	96%
PC-3	Prostate	17,359,026	10%	55%	584	96%
SK-MEL-28	Melanoma	13,823,238	8%	65%	552	97%
NCI-1395	Lung	14,095,256	8%	55%	510	96%
HT-29	Colon	14,432,968	9%	62%	543	96%

Pools of genomic DNA from cancer cell lines were subject to exon capture and sequenced in a single 100-bp paired-end Illumina HiSeq2000 lane. Purity filtered (PF) sequence reads for each sample are shown; the percent of total PF reads in the pool demonstrate the relative representation of each sample within the pool. Percent selected bases indicate the percent of bases that mapped within 250 bp of a target exon, including both on-bait and near-bait bases. Mean target coverage represents the average number of unique reads in which each base was sequenced.

Supplementary Table 5. Summary of genomic alterations in cell lines

CELL LINE	MEAN TARGET COVERAGE	SNVs	MISSENSE SNVs	NONSENSE SNVs	INDELS	COPY NUMBER GAINS	COPY NUMBER LOSSES
HT-29 (50% input)	295	12	7	2	2	0	2
MCF7	540	9	5	1	0	2	1
MDA-MB-231	568	10	6	2	1	0	3
ZR-75-30	593	22	9	1	0	2	0
HL-60	543	7	4	1	0	1	1
NCI-H69	451	10	6	2	1	1	0
NCI-358	441	9	7	0	0	0	2
PC-3	584	2	0	0	1	0	1
SK-MEL-28	552	13	9	0	0	0	0
NCI-1395	510	5	5	0	1	2	0
HT-29	543	25	14	5	2	0	0

Summary of single nucleotide variants, indels, and copy number alterations for the 12 pooled cell lines. Copy number alterations are calculated as the log2 ratio of coverage as compared to a normal diploid sample. Copy number gains are defined as a log2 ratio greater than 1.58 (3-fold); copy number losses are defined as log2 ratio less than -1.58 (3-fold).

Supplementary Table 6. Single-nucleotide variants and indels in breast cancer cell line MDA-MB-231

GENE	CODING CHANGE	PROTEIN CHANGE	TYPE OF SNV	NUMBER OF READS	ALLELE FREQUENCY
BRAF	**c.1391G>T**	**p.G464V**	**Missense**	**701**	**56%**
EPHA3	c.2515G>T	p.E839*	Nonsense	428	32%
GNAS	c.18C>T	p.N6N	Synonymous	354	20%
KRAS	**c.38G>A**	**p.G13D**	**Missense**	**70**	**66%**
MYCL1	c.676G>A	p.E226K	Missense	996	35%
NF1	c1398_1399insC	p.T467Hfs*3	Frameshift	227	100%
NF2	**c.691G>T**	**p.E231***	**Nonsense**	**350**	**100%**
NOTCH3	c.1102A>T	p.T368S	Missense	951	66%
PDGFRA	c.515A>T	p.Y172F	Missense	996	35%
PIK3CA	c.363C>T	p.I121I	Synonymous	123	39%
TP53	**c.839G>A**	**p.R280K**	**Missense**	**769**	**100%**

The single nucleotide variants and indels detected in one of the 12 pooled samples, breast cancer cell line MD-MBA-231. Total number of reads and the frequency of the variant (out of the total reads) are shown. Mutations listed in bold have previously been reported in the COSMIC database for this cell line.

Supplementary Table 7. Correlation of copy number alterations detected by exon capture and microarrays in cell lines

Cell Line	Correlation Coefficient
HT-29 (50% input)	0.90
MCF7	0.96
MDA-MB-231	0.94
ZR-75-30	0.94
HL-60	0.96
NCI-H69	0.94
NCI-358	0.98
PC-3	0.89
SK-MEL-28	0.97
NCI-1395	0.98
HT-29	0.98

Correlation coefficients between gene-level copy number alterations as detected by exon capture and copy number data previously obtained using a high-density SNParray (Affymetrix SNP 6.0 platform).

Supplementary Table 8. Summary of genomic alterations in FFPE tumor samples

SAMPLE	TUMOR TYPE	TUMOR PURITY	MEAN TARGET COVERAGE	SNVs	MISSENSE SNVs	NONSENSE SNVs	INDELS	COPY NUMBER GAINS	COPY NUMBER LOSSES
FFPE 1	Colon	60%	457	12	7	0	0	0	0
FFPE 2	Colon	10%	353	8	5	0	1	0	0
FFPE 3	Colon	20%	498	33	20	1	5	0	0
FFPE 4	Colon	60%	300	12	6	3	0	0	0
FFPE 5	Breast	80%	472	9	5	0	0	2	2
FFPE 6	Breast	70%	532	15	7	0	2	0	0
FFPE 7	Colon	50%	250	30	15	1	1	0	0
FFPE 8	Breast	80%	537	5	4	0	1	0	0
FFPE 9	Colon	60%	116	4	3	1	2	0	0
FFPE 10	Colon	50%	410	27	13	3	2	0	0

Summary of single nucleotide variants, indels, and copy number alterations for the pooled FFPE samples. Copy number alterations are calculated as the log2 ratio of coverage as compared to a normal diploid sample. Copy number gains are defined as a log2 ratio greater than 1.58 (3-fold); copy number losses are defined as log2 ratio less than -1.58 (3-fold).

Supplementary Table 9. Pharmacogenomic polymorphisms in UGT1A1 and ERCC2

Samples	Cancer Type	UGT1A1-G3156A		ERCC2-K751QC	
		Reads	Genotype	Reads	Genotype
FFPE 1	Colon	478	G/A	463	T
FFPE 2	Colon	318	G/A	480	T/G
FFPE 3	Colon	636	G/A	442	T
FFPE 4	Colon	394	G	244	T
FFPE 5	Breast	321	G/A	266	T
FFPE 6	Breast	589	G	696	T
FFPE 7	Colon	303	G/A	272	T
FFPE 8	Breast	580	G	759	T
FFPE 9	Colon	121	G	99	T/G
FFPE 10	Colon	411	G	471	T

Allele incidence for 2 exemplary pharmacogenomics loci across all 10 FFPE samples.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Approach to tumor genomic profiling. Genomic DNA is extracted from tumor samples and used to generate sequencing libraries. A DNA barcode is appended to each library by PCR. Following quantification of libraries, equimolar pools consisting of 12 barcoded tumor DNAs and normal diploid control DNAs are made. The DNA pools are subjected to solution-phase hybrid capture with the biotinylated RNA baits corresponding to the coding sequence of the 137 “druggable” or potentially “actionable” genes known to undergo somatic genomic alterations in cancer. Massively parallel sequencing is performed on the captured sequences. The sequencing data were deconvoluted to match all high-quality reads with the corresponding tumor samples and call single-nucleotide variations, small insertions/deletions, and copy number alterations.

Supplementary Figure 2. Exon capture bait performance. (A) The percent of targets in a normal diploid sample (HapMap) with at least the specified sequence coverage. As shown, 97% of targets in the HapMap sample were covered at 30x or higher. The 80th percentile target coverage was 191x. (B) The number of targets in a HapMap sample with specified sequence coverage. (C) Coverage for each target in a normal diploid sample (HapMap) was plotted as a function of GC content. **(**D) Correlation of target coverage for two independently generated libraries from the HT-29 cell line that were captured and sequenced within a single pool. Coverage for each genomic target was highly correlated with a correlation coefficient of 0.86. (E) Correlation of target coverage for two separate captures and sequencing runs for a single HapMap library, with a correlation coefficient of 0.96. (F) Correlation of target coverage for two separate captures and sequencing runs for a single HT-29 library, with a correlation coefficient of 0.97.

Supplementary Figure 3. Copy number by quantitative PCR in sample FFPE 9. Copy number correlation between exon capture and QPCR in sample FFPE 9. Quantitative PCR of *FGFR1, CCND1*, and *NOTCH1* using 3 independent sets of primers was performed and average values for each gene were compared to exon capture copy number.

SUPPLEMENTARY METHODS

Blocking Oligonucleotides:

Oligonucleotide	Sequence (5'-3')
IndexBlock_1	CAA GCA GAA GAC GGC ATA CGA GAT ATC ACG GTG ACT GGA GTT C
IndexBlock_2	CAA GCA GAA GAC GGC ATA CGA GAT CGA TGT GTG ACT GGA GTT C
IndexBlock_3	CAA GCA GAA GAC GGC ATA CGA GAT TTA GGC GTG ACT GGA GTT C
IndexBlock_4	CAA GCA GAA GAC GGC ATA CGA GAT TGA CCA GTG ACT GGA GTT C
IndexBlock_5	CAA GCA GAA GAC GGC ATA CGA GAT ACA GTG GTG ACT GGA GTT C
IndexBlock_6	CAA GCA GAA GAC GGC ATA CGA GAT GCC AAT GTG ACT GGA GTT C
IndexBlock_7	CAA GCA GAA GAC GGC ATA CGA GAT CAG ATC GTG ACT GGA GTT C
IndexBlock_8	CAA GCA GAA GAC GGC ATA CGA GAT ACT TGA GTG ACT GGA GTT C
IndexBlock_9	CAA GCA GAA GAC GGC ATA CGA GAT GAT CAG GTG ACT GGA GTT C
IndexBlock_10	CAA GCA GAA GAC GGC ATA CGA GAT TAG CTT GTG ACT GGA GTT C
IndexBlock_11	CAA GCA GAA GAC GGC ATA CGA GAT GGC TAC GTG ACT GGA GTT C
IndexBlock_12	CAA GCA GAA GAC GGC ATA CGA GAT CTT GTA GTG ACT GGA GTT C

Primers for QPCR:

Oligonucleotide	Sequence (5'-3')
FGFR1_A_FWD	TGA GCT GTC AAG GAC AGT GG
FGFR1_A_REV	GAC AGA TGT GCC TTC TGC AA
FGFR1_B_FWD	GTG GTG TTG GCA GAG GCT AT
FGFR1_B_REV	TGC AAG GAC AGA AGC ATC AC
FGFR1_C_FWD	GAG CCT GAA GTG GGT GAG AG
FGFR1_C_REV	CTC CGT GTT GCT GTT TCT GA
CCND1_A_FWD	TGA AGA ATC CCT GGA TGG AG
CCND1_A_REV	GCC TGG GGT GAG ATA CAA GA
CCND1_B_FWD	CCC TTC TCT CCC GCT AGA AC
CCND1_B_REV	ACC CCT TCC TCC TTC AGA AA
CCND1_C_FWD	TGA ACT ACC TGG ACC GCT TC
CCND1_C_REV	GGG GAT GGT CTC CTT CAT CT
NOTCH1_A_FWD	TCT GGG GTC CTC TTT TTC CT
NOTCH1_A_REV	ACA GAG CCG AAT CCA GCT TA
NOTCH1_B_FWD	AGG CCG TGC AGA GTA AGT GT
NOTCH1_B_REV	GGT AGC AAC TGG CAC AAA CA
NOTCH1_C_FWD	GAC TGC AGC GAG AAC ATT GA
NOTCH1_C_REV	GGG ACA CTC GCA GTA GAA GG
LINE_CTRL_FWD	AAA GCC GCT CAA CTA CAT GG
LINE_CTRL_REV	TGC TTT GAA TGC GTC CCA GAG

SUPPLEMENTARY APPENDIX

“Supplementary Appendix.xlsx”

TAB 1: SNVs – cell lines
TAB 2: Indels – cell lines
TAB 3: Target Coverage – cell lines
TAB 4: Gene Coverage – cell lines
TAB 5: Gene Copy Number (log2 ratio) – cell lines
TAB 6: SNP Array Copy Number – cell lines
TAB 7: SNVs – FFPE samples
TAB 8: Indels – FFPE samples
TAB 9: Target Coverage – FFPE samples
TAB 10: Gene Coverage – FFPE samples
TAB 11: Gene Copy Number (log2 ratio) – FFPE samples
TAB 12: Bait sequences