

SUPPLEMENTARY INFORMATION

- 1. Supplementary Materials and Methods.**
- 2. Supplementary Tables: Table 1-4**
- 3. Supplementary Figures: Figure 1-9**

Supplementary Materials and Methods

1. Statistical Analyses

cMethDNA data analyses were performed using GraphPad Prism version 5.0 (GraphPad Software, San Diego, CA), SAS software (v 9.2, SAS Institute Inc., Cary, NC), or with R version 2.15.2 (2012-10-26). All statistical tests were two-sided and considered statistically significant at $p < 0.05$ unless otherwise stated. Distributions of cMethDNA data between independent groups were described using box plots and difference was tested using nonparametric Mann-Whitney test. The Wilcoxon signed-rank test was performed when comparing two related samples (e.g., measurements on the same subjects). Where age-dependent difference in methylation were examined, based on donor age serum samples were assembled into four quartiles, then mean CMI were tested using one-way ANOVA and Kruskal Wallis, which showed no age related methylation. Reproducibility of the duplicate measurements was evaluated with the coefficient of variation (CV) and inter-user reproducibility was evaluated using intraclass correlation coefficient (ICC). Receiver Operating Characteristic (ROC) analyses were used to define a laboratory threshold that maximizes the sum of sensitivity and specificity in the training set and the classifier was subsequently validated in an independent test set. The performance of the 10-gene panel was characterized through estimating the area under the ROC curve (AUC), sensitivity, specificity, classification accuracy and likelihood ratio along with the 95% confidence intervals. For determination of marker profile concordance in the Rapid Autopsy set, methylation levels were binarized as 'positive' or 'negative', and concordance between two samples from the same patient was measured as the number of markers (out of 10) on which the two samples agreed. To obtain the distribution of concordances that would be expected by chance alone, we first determined the expected methylation status (methylated or unmethylated) of each marker in the entire population depending on whether the frequency of methylation was greater than or less than 50%, finally calculating the concordances between sample and expected methylation status as above. Performance of the 10-gene panel in arrays of a variety of tumor types was tested using TCGA Illumina Infinium HumanMethylation27 array data downloaded from the UC Santa Cruz Cancer Genomics Browser <https://genome-cancer.ucsc.edu/>.

2. DNA Methylation array

Details of processing of the tissue DNA methylome array have been previously published (11) and data deposited (GEO accession number GSE 31979). Briefly, frozen tissue samples were sectioned, washed free of embedding media with ethanol washes, harvested into microcentrifuge tubes for lysis in TNES/PK (10 mM Tris, 150 mM NaCl, 2 mM EDTA, 0.5% SDS and 30 μ g proteinase K) prior to conversion with sodium bisulfite. Sodium bisulfite conversion was performed using EZ DNA Methylation kit (Zymo Research, Orange, CA, USA) as directed by the manufacturer. Eluted DNA was processed with the Illumina Infinium Human Methylation27 BeadChip (Illumina, Inc., San Diego, CA; WG-311-1202) in the JHU DNA Microarray Core. Methylation levels were defined according to the manufacturer, as β -values (similar to % methylation) ranging from 0-1 (low to high, respectively). DiffScores were computed after Benjamini and Hochberg correction for false discovery ($p = 0.05$). The DiffScore is a transformation of the p-value that provides directionality to the p-value based on the difference between the average signals in the reference group vs. the comparison group. The formula is: $\text{DiffScore} = 10 * \text{sgn}(\mu_{\text{cond}} - \mu_{\text{ref}}) * \log_{10} p$; For a p-value of 0.05, $\text{DiffScore} = \pm 13$;

For a p-value of 0.01, DiffScore = ± 22 ; For a p-value of 0.001, DiffScore = ± 33 ; $p = 10^{(\text{DiffScore} * \text{sgn}(\mu_{\text{cond}} - \mu_{\text{ref}}) / 10)}$.

For the serum DNA methylation array (GEO accession #GSE52621), whole blood was collected in an SST Vacutainer tube (BD Diagnostics, Franklin Lakes, NJ; #367988), allowed to clot for 30 min, and separated according to manufacturer's directions within 2 hr. Serum (6 ml) was digested overnight at 48°C by inverting continuously with 12 ml buffer containing 1% SDS, 100 mM Tris pH 8.5, 2.5 mM EDTA, and 120 μ l 2% w/v proteinase K (Roche Applied Sciences, Indianapolis, IN). Thereafter, three times daily 120 μ l 2% proteinase K was added for a total of 72 hours of digestion. The sample was centrifuged 10 min and supernatant was extracted in an equal volume of warm (50 °C) Tris buffer-saturated phenol/chloroform, pH 8.0. Phase Lock Heavy Gel (5 Prime, Gaithersburg, MD) was used to isolate the aqueous layer. DNA was precipitated with ethanol in the presence of glycogen and NaCl, washed in 70% ethanol, and resuspended in 1 mM Tris, pH 8. DNA was converted with sodium bisulfite and processed for array as described for tissue.

3. The cMethDNA assay

Overview: For each gene, cMethDNA requires: 1) A standard (STDgene) to operate as a gene-specific reference DNA. This has 5' and 3' sequences (~20 bp each "external" sequences) homologous to the TARGETgene which flank a short internal non-human DNA sequence (i.e. 140-300 bp of lambda phage DNA). This cassette is packaged into a plasmid (e.g. pCR2.1; Life Technologies); 2) A forward and reverse "external" primer pair used for multiplex PCR capable of hybridizing to the external 5' and 3' sequences of the TARGETgene/STDgene; 3) A forward and reverse primer pair used for real-time PCR capable of hybridizing to sequences located internally relative to external sequences, and which are specific to the methylated TARGETgene; 4) A pair of forward and reverse internal primers used for real-time PCR which are capable of specifically hybridizing to the reference STDgene; 4) Probes for TARGETgene and STDgene internal sequences, labeled in distinguishable colors (e.g., 6FAM/TAMRA or VIC/TAMRA; used for two-color real-time PCR). All assay primers and probes are listed in Supplementary Table 3.

cMethDNA assay primers/probes were designed to overlap or lie within 100 bases of the differentially methylated loci identified by methylome array. The cMethDNA methylation-specific target gene primers (two) and probe (one) jointly contain about 9-11 CpG dinucleotides (ranging from 7-12) depending on the desired melting temperatures (T_m) of the primers/probe (calculated as C or G = 4, and A or T = 2 T_m units) and the density of CG dinucleotides in the region; independent C residues (about 8; ranging from 6-10) are also present to ensure selective hybridization only to sodium bisulfite-converted DNA.

Standards: For individual genes the standard was designed so that STD and endogenous gene amplicons resulting from Step 1 and Step 2 PCR reactions would be the same size. Forward and reverse cloning primers encompassed 5' and 3' external primer sequences (~20 bp) fused to lambda phage sequences (predicted after sodium bisulfite conversion; ~20 bp; Supplementary Table 3). Each STDgene was designated to have a unique phage sequence to eliminate cross-reactivity between standards. PCR-mediated cloning was performed and sequences were verified by restriction digestion, as well as DNA sequencing. Plasmid copy number was determined by OD₂₆₀ (Nanodrop, Thermo Scientific, Wilmington, DE), considering the molecular mass of the recombinant plasmid using online OligoCalc, software (Northwestern

University, <http://www.basic.northwestern.edu/biotoools/OligoCalc.html>). Prior to beginning the study, stocks of STDgene plasmids were prepared, a cocktail of all 10 individual STDgene plasmids was mixed together, then aliquoted and stored frozen in the presence of salmon sperm carrier DNA (50 µg/ml) at -80°C.

Processing of Serum DNA for cMethDNA assay

Whole blood was collected in an SST Vacutainer tube (BD Diagnostics, Franklin Lakes, NJ; #367988), allowed to clot for 30 min, and separated by centrifugation (2000 x g for 20 min at 4°C) according to manufacturer's directions within 2 hr. Samples were stored at -80°C and thawed on ice before using. Some samples were transported frozen on dry ice to a central location before thawing for the cMethDNA assay. Serum was purified using the MinElute Virus Spin Kit (Qiagen catalog #57704), essentially according to the manufacturer's protocol although modified for cMethDNA to accommodate 300 µl of serum. For each new kit, stocks were prepared of protease (1.4 ml AVE to protease vial, mix, aliquot and store at -20 °C), and of carrier RNA (310 µl AVE in 310 µg carrier RNA, dissolve, freeze at -20 °C in aliquots @ 1 µg/ml; use 5.6 µg/sample). All steps were performed at room temperature, including centrifugation; Buffer AVE was equilibrated to room temperature for the elution step; an incubator oven was pre-warmed to 56 °C; a working dilution of carrier RNA was prepared (per sample = 300 µl Buffer AL + 5.6 µl carrier RNA in Buffer AVE). The protocol given below was followed:

1. Pipet 37 µl Qiagen Protease (prepared in AVE and stored in -20 °C freezer) into a 2 ml microcentrifuge tube.
2. Add 300 µl serum, tap to mix.
3. Add 300 µl Buffer AL containing 5.6 µg of carrier RNA (5.6 µl). Mix.
4. Immediately before using, thaw and dilute the frozen stock STDgene cocktail: Mix 2 µl of stock 10³/µl + 198 µl dilution buffer (1X MSP buffer, 50 µg/ml tRNA, 50 µg/ml salmon sperm DNA). Then add 5 µl of the diluted STDgene cocktail to the serum dilution in step 3; for our study 12 unique STDgenes, 50 copies total of each STDgene were present in the final serum mixture. Mix for 15 s by pulse vortexing.
5. Incubate at 56 °C. for 15 min in a pre-warmed rack.
6. Briefly centrifuge.
7. Add 375µl absolute ethanol, close the cap, mix by pulse-vortexing for 15 s. Incubate the lysate with the ethanol for 5 min at room temperature (not to exceed 25 °C).
8. Briefly centrifuge the 2 ml tube.
9. Transfer all the lysate onto the QIAamp MinElute column without wetting the rim. Close the cap and centrifuge at 8000 rpm 1 min. Transfer the column to a new collection tube. Discard the filtrate.
10. Wash the column with 500 µl Buffer AW1 without wetting the rim. Centrifuge 8000 rpm 1 min. Transfer the column to a new collection tube.
11. Wash the column with 500 µl Buffer AW2 without wetting the rim. Centrifuge 8000 rpm 1 min. Transfer the column to a new collection tube.
12. Wash the column with 500 µl Absolute ethanol without wetting the rim. Centrifuge 8000 rpm 1 min. Transfer the column to a new collection tube.
13. Centrifuge the empty column full speed (13,000 rpm) 3 min. Transfer the column to a new collection tube. Make sure no liquid is visible in or attached to the column.

14. Add 22 μ l Buffer AVE directly to the center of the membrane. Incubate at room temperature 5 min. Centrifuge at full speed 1 min. Repeat with another 22 μ l Buffer AVE and centrifuge full speed 1 min. Pool.
15. Transfer the contents to a 500 μ l microcentrifuge tube for sodium bisulfite treatment. Keep on ice.

Sodium Bisulfite-Mediated Conversion of Serum DNA

1. Add 7.5 μ l M-dilution buffer (Zymo Research) to the DNA sample (represents all the DNA recovered from the purification step; final volume is \sim 50 μ l). Incubate at 42 $^{\circ}$ C for 15-20 minutes. Briefly centrifuge.
2. Prepare the CT Conversion Reagent (for each sample: water = 71.4 μ l, M-dilution buffer = 17.6 μ l, CT granular conversion reagent (Zymo Research) = 54 mg). Alternatively, use a ready-made vial per suitable for 11 serum DNA samples: add 750 μ l ddH₂O + 185 μ l M-dilution buffer to the 1.7 ml brown vial containing 567 mg of CT reagent. This is enough for 11 serum DNA samples. Rotate the solution in the dark at room temp for 10 min to dissolve. Use the reagent within 10 min.
3. Mix together 97.5 μ l of CT Conversion Reagent with the sample thoroughly. Cap the tube, vortex briefly (or invert), centrifuge briefly. The final volume is \sim 150 μ l. Incubate the DNA solution in a PCR machine with a hot lid cycling 16 cycles of (95 $^{\circ}$ C 30 sec, 50 $^{\circ}$ C 1 hr). An oil overlay is not needed if the hot lid is used. Hold at 4 $^{\circ}$ C. Briefly centrifuge.
4. To perform DNA cleanup place 600 μ l of M-Binding Buffer (Zymo Research) into a Zymo Spin Column (IC) in a collection tube. Add the sample. Gently pipet up and down 10 times to mix the DNA with the binding buffer. Centrifuge at 13000 rpm for 30 seconds and change to a new collection tube. Discard the collection tube containing the flow through.
5. Add 100 μ l of M-Wash Buffer (Zymo Research) to the column and centrifuge at 13000 rpm for 30 seconds.
6. Add 200 μ l M-Desulfonation Buffer (Zymo Research) to the column and let the mixture sit at room temperature for 20 minutes. Centrifuge at 13000 rpm for 30 seconds. During this time, warm an aliquot of molecular grade water to 70 $^{\circ}$ C.
7. Wash the column with 200 μ l of M-Wash Buffer (Zymo Research), invert the tube and centrifuge at 13000 rpm for 30 seconds. Discard the flow through. Change collection tubes. Repeat once, centrifuging a full minute.
8. Transfer the column to a new collection tube. Add 15 μ l of the pre-heated water right to the top of the resin in the column. Allow the column to sit for 5-10 minutes at room temperature. Centrifuge the column at 13000 rpm for 1 minute to recover the DNA. Chill the sample on ice and then perform multiplex PCR using the entire sample.

Multiplex PCR (PCR Assay #1)

Overview: In cMethDNA assays, the aim of the multiplex reaction is to co-amplify up to 12 genes using one pair of external primers per gene. The multiplex step is independent of the methylation status of the genes because primers hybridize to sequences flanking but not covering CpGs of interest; methylated and standard amplicons are quantitated later by real-time PCR. When CpGs are unavoidable in the external primers due to CG content of the region, degenerate primers with equimolar ratio of mismatched nucleotides are used, and mismatches are limited to

no more than two CpGs. The relatively low annealing temperature (56°C), extensive annealing time (45 sec), high primer concentration (2 ng/μl) and high polymerase concentration (10U in 50 μl) of the PCR reaction are intended to avoid biased amplification of unmethylated or methylated DNA. In previous studies, QM-MSP and qMSP (using ACTB as reference) were found to yield equivalent results (Fackler et al., 2004).

Reaction conditions: The multiplex reaction was performed in a 50 μl reaction volume, in buffer containing 16.6 mM NH₄SO₄, 67 mM Tris pH 8.8, 6.7 mM MgCl₂, 1.25 mM dNTP (each; Denville Scientific, Metuchen, NJ), 10 mM β-mercaptoethanol, 0.1% DMSO, 2 ng/ul per primer and 10 U/50μl Platinum Taq polymerase (Invitrogen). Thermocycler settings were 95 degrees 5 min, followed by 36 cycles of 95 °C for 30 sec, 56 °C for 45 sec, and 72 °C for 45 sec, then a final extension of 7 min at 72 °C. After completion of PCR, the reaction mix was diluted 1:5 by adding 200 μl 1X Multiplex Dilution Buffer [containing 1X MSP buffer, 50 μg/ml Salmon Sperm DNA (Invitrogen), 50 μg/ml tRNA (Roche Applied Science, Indianapolis, IN)] before performing the quantitative real-time PCR assay. The reaction mix was stored frozen at -20 °C for archival purposes.

Quantitative Real-Time Methylation Specific PCR (PCR Assay #2):

DNA amplicons synthesized in the Multiplex Reaction (PCR #1) were quantified using two-color real-time PCR (absolute quantitation method) after diluting the reaction mix. TARGETgene and STDgene amplicons were assayed independently but in the same well using sets of primers internal to the multiplex primers. Each gene and standard amplicon type was tested with a unique sets of primers (forward and reverse) and hydrolysis probes specific for gene target or reference standard.

1. **Master mix:** The real-time PCR reaction was performed in 20 μl containing 16.6 mM NH₄SO₄, 67 mM Tris pH 8.8, 6.7 mM MgCl₂, 10 mM β-mercaptoethanol, 0.1% DMSO, 300 nM ROX (Invitrogen/Life Technologies, Grand Island, MY), 200 μM dNTP (Denville Scientific), 5 μg/ml tRNA (Roche Applied Science, Indianapolis, IN), 1.0 U RAMP Taq polymerase (Denville Scientific; #CB4080-9), 700 nM each forward and reverse primer (Life Technologies/Foster City, CA), and 200 nM for each hydrolysis probe (labeled with 6FAM/TAMRA or VIC/TAMRA; Applied Biosystems/Life Technologies).
2. **Samples and multiplex controls:** For each sample or multiplex reaction control, 4 μl was tested at a final 1:500 or 1:50,000 dilution (diluted in 1X MSP dilution buffer: 16.6 mM NH₄SO₄, 67 mM Tris pH 8.8, 6.7 mM MgCl₂, 10 mM β-mercaptoethanol, 0.1% DMSO, 50 μg/ml salmon sperm, 50 μg/ml tRNA).
3. **Standard Curve (for absolute quantitation):** To make the stocks of the curve DNA, TARGETgene and STDgene templates were amplified independently for each gene using the multiplex PCR reaction protocol, 10 ng template DNA (either universally methylated DNA made by SssI treatment of cell line DNA, or plasmid DNA containing standard), using 4 ng forward and reverse primers. To prepare a standard curve for each gene, aliquots of stock were prepared containing equal copy number of TARGETgene and STDgene (as determined by overlapping amplification plots in real-time PCR). For each assay the curve DNA was serially diluted 1:10 for 7 logs of dynamic range into salmon sperm DNA (50 μg/ml), then discarded after use. Four microliters was assayed in

duplicate wells by real-time. These stocks were stored at -80° C and are stable indefinitely.

4. **Quantification:** Absolute quantitation of multiplex amplicons was performed using ABI 7500 software according to instructions from manufacturer, with several modifications. The automatic baseline setting was used, and the threshold was set manually at the point where overlap of the curves for TARGETgene or STDgene was evident (e.g. usually ΔR_n between 0.02 and 0.04), then standard curve serial dilutions were assigned an arbitrary copy number values beginning with 200,000,000 copies/well for the most concentrated points. Sample values were extrapolated from the curve for target and reference DNAs and only reference sample values falling within the range of the standard curve were accepted. The Methylation Index (MI) for each gene was calculated as: $MI = \frac{\# \text{ copies } M}{\# \text{ copies } (M+STD)} (100)$ and the cumulative methylation index (CMI) was determined as equal to the sum of MI for all genes.

For curves where the average ΔC_t was ≤ 1.0 cycle (considering all dilutions) but not perfectly overlapping, the copy number was calculated in this way: 1) Based on the average ΔC_t for all points of the curve (7 logs) the fold-difference was calculated (where 1 cycle = 2-fold). 2) The value of the most concentrated sample was divided by this difference. Example: If the ΔC_t STDgene -TARGETgene = 1.5 cycles, then $2^{1.5}=2.828$, and 200,000,000 copies/2.828 fold = 70,700,000 copies. Assign STDgene = 70,700,000 and TARGETgene = 200,000,000 copies per well. Assays were repeated after the standards remade if the curves exceeded 1.0 cycle overlap.

5. **Plate set up:** An Applied Biosystems (ABI) MicroAmp Fast Optical 96 Well Plate (#4346906) was used with the ABI Fast 7500 Real-Time PCR System. Usually the standard curve samples were placed in Row A wells 1-12, and in Row B wells 1-2. Row B well 3 contained NTC (no template control) from the multiplex reaction, well 4 contained NTC from the real-time reaction mix, well 5 contained STD cocktail (no target DNA) from the multiplex, and well 6 contained universally methylated genomic DNA (no STD reference DNA). The remaining wells were occupied by samples.

Assay Performance Criteria (example in Supplementary Figure 9):

1. Standard curve: $R^2 = \geq 0.99$, efficiency = $90 \pm 10\%$ and slope = $-3.33 \pm 10\%$. TARGETgene and STDgene C_t overlap ≤ 1.0 cycles, averaged over the whole curve.
2. Sample C_t values may not exceed the upper range of the TARGETgene or STDgene standard curves or be lower than the lowest point of the STDgene curve.
4. Control water (after both multiplex and real-time PCR), $C_t \geq 38$.
5. Control fully methylated TARGETgene DNA (universally methylated DNA template, after multiplex and real-time PCR), MI = 100.
6. Control STDgene DNA (STDgene plasmid, after multiplex and real-time PCR), MI = 0.
7. For patient samples, duplicate serum aliquots are assayed and averaged.
8. Implausible extreme values (e.g. due to contamination) are discarded/repeated.

Supplementary Table 1. Patient sample sets used in this study

Sample sets			
A. Assay development and analytical validation			
Sample sets	Blood		
	Serum Cancer	Serum Normal	
Training cMethDNA set-J0425, J0214, TBCRC005 (normal)	24	28	
Test cMethDNA set-TBCRC005 (cancer and normal)	33	27	
B. Monitoring response to therapy: Serial serum sample collection points			
Sample sets	Baseline + first follow-up	Baseline + multiple follow-up	
Stage 4: Progression after prior therapy- J0425, J0214	29	13	
C. Concordance between sera and tissue methylation pattern			
Sample sets	Blood	Tissue	
	Serum Cancer	Tumor Primary	Tumor Metast.
Stage 4: Newly diagnosed- TBCRC 013	18	18	12
Stage 4: Progressed after prior therapy- J0214, J0425	10	10	3
Stage 4: At autopsy- J0611	10	10	61
Patient characteristics described in Supplementary Table 2			
Studies at J: Johns Hopkins; TBCRC: Translational Breast Cancer Consortium			

Supplementary Table 2. 10-gene marker panel and its performance

A

Selected methylated CpG loci markers - 10-gene panel					Fackler Array*	TCGA Array**
Array ID	Symbol	Distance to TSS	CpG Island	Product	Ratio Cancer to Normal	Ratio Cancer to Normal
cg13801416	AKR1B1	31	TRUE	Aldo-keto reductase family 1, member B1	7.4	14.0
cg00557354	ARHGEF7	117	TRUE	Rho guanine nucleotide exchange factor 7 isoform a	2.9	2.7
cg21513553	COL6A2	595	TRUE	Alpha 2 type VI collagen isoform 2C2 precursor	6.0	9.5
cg16557944	GPX7	153	TRUE	Glutathione peroxidase 7	4.9	14.7
cg25438963	HIST1H3C	24	TRUE	H3 histone family, member C	2.2	3.0
cg06760035	HOXB4	498	TRUE	Homeo box B4	4.5	4.1
cg09952204	RASGRF2	106	TRUE	Ras protein-specific guanine nucleotide-releasing factor 2	8.4	12.7
cg00777121	RASSF1	176	TRUE	Ras association domain family 1 isoform A	2.3	2.8
cg14696396	TM6SF1	66	TRUE	Transmembrane 6 superfamily member 1	4.7	5.1
cg06856528	TMEFF2	239	TRUE	Tomoregulin-2	5.2	4.3

Tissue methylation arrays of *Fackler et al. 2011 (GEO Accession number GSE31979) and **The Cancer Genome Atlas project (<http://tcga-data.nci.nih.gov/>). Ratios are based on the average beta methylation value within each group of arrayed tissues, cancer versus normal breast, performed on Illumina Infinium HumanMethylation 27 arrays.

B

Gene Biomarker Methylation Values		
Symbol	Ratio Serum: Cancer vs. Normal	DiffScore* (Adjusted FDR)
AKR1B1	3.71	70
ARHGEF7	3.80	94
COL6A2	2.55	26
GPX7	2.34	34
HIST1H3C	1.97	49
HOXB4	4.50	117
RASGRF2	4.00	58
RASSF1	1.39	8
TM6SF1	2.58	57
TMEFF2	3.05	40

* FDR adjusted p-value 0.05 = DiffScore of ± 13 ; p 0.01 = ± 20 , p 0.001 = ± 30 ; 902 CpG loci

Supplementary Table 3A. cMethDNA and QM-MSP primer and probe sequences

Type*	Primer/Probe Name	5' To 3' Sequence	Array Target ID	TSS	M Primer Location	M Primer Distance to TSS	Size
External	AKR1B1_Ext_F	GYGTAATTAATTAGAAGGTTTTTT	cg13801416	Chr. 7; -134143888	134143923	minus 35 bp	216 bp
External	AKR1B1_Ext_R	AACACCTACCTTCCAAATAC					
M Target	AKR1B1_FM	GCGCGTTAATCGTAGGCGTTT					
M Target	AKR1B1_RM	CCCAATACGATACGACCTTAAC					
M Target	AKR1B1_M_Probe	CGTACCTTTAAATAACCCGTAAAATCGA					
U Target	AKR1B1_FUM	TGGTGTGTTAATTGTAGGTGTTTT					
U Target	AKR1B1_RUM	CCCAATACAAATACAACTTAAC					
Standard	AKR1B1_U_Probe	ACATACCTTTAAATAACCCATAAAATCAAC					
Standard	STD_AKR1B1_F	TTTGTGTGTTTTGTGGAAGTAAG					
Standard	STD_AKR1B1_R	ATTTCATCAATACTTTCAAATAACACA					
Standard	STD_AKR1B1_Probe	AAATACATTATCCTACCACCTAACCAATACA					
Standard	STD_AKR1B1 Sequence	GYGTAATTAATTAGAAGGTTTTTTATTGGTGTAGGTAATTTGTATTGTTATTGTGATTGATGATTTTTTTGATTGTAGATAGTGGTGTGTTTGTGATGTTTTGTGGAAGTAAGTGATTGTTAGTGGTAGGATAATGTATTGATGTGTTATTTGAAAGTATTGATGAATGGTGTGGTGATTTATGATGGTATTGTTGGAAGGTAGGTGTT					
External	ARHGEF7_F_Ext	YGTTTYGAGGTGAAGGYGYG					
External	ARHGEF7_R_Ext	CTCCAACAACCTACAAAAAAC					
M Target	ARHGEF7_FM	GTTTTTCGGGTCGTAGCGATG					
M Target	ARHGEF7_RM	CAAAAAACCTCCGAATCCGAA					
M Target	ARHGEF7_M_Probe	AAACCACGTAACGATTTACTCGACGA					
U Target	ARHGEF7_FUM2	GTTGTAGTGATGAATTTTGTGAG					
U Target	ARHGEF7_RUM2	CAAAAAACCTCCAAATCCAAAT					
Standard	ARHGEF7_U_Probe	AATAAACCAATACAAATTTACTCAACAAA					
Standard	STD_ARHGEF7_F	GGAAATGTGATTTTGTGTTTTATGTT					
Standard	STD_ARHGEF7_R	ACCCAACACTATTCTTAATCAC					
Standard	STD_ARHGEF7_Probe	ACCTACACATCACTAACAAACATATACAA					
Standard	STD_ARHGEF7 Sequence	YGTTTYGAGGTGAAGGYGYGTGTTGATGTTAGTGGGTTGGGAAAATGTGATTTTGTGTTTATGTTTATGTTTATGTTTATGGTATTTGTAGATATTTTGTGATATGTTTGTAGTGATGTTAGGTTATGGTGATTAAGAAATAGGTGTTGGGTATTAGTGTGGTTTTGGTTTTGTAGGTTGTTGGAG					
External	COL6A2_Ext_F	AGGTTTAGGAGAAGTTGTAGA	cg21513553	Chr 21; 47518033	47518733	700 bp (intronic)	154 bp
External	COL6A2_Ext_R	TACCAACAATAAAAACCCAAAC					
M Target	COL6A2_FM	ATTTCGGGTTGATAGCGATTTCGT					
M Target	COL6A2_RM	CGATTCCACCAACGCCCCG					
M Target	COL6A2_M_Probe	ATCCCAAAAACGAATATAAACGACCCG					
U Target	COL6A2_FUM	GATTTGGGTTGATAGTATTGTA					
U Target	COL6A2_RUM	CAATTCACCAACACCCCAAC					
Standard	COL6A2_U_Probe	ATCCCAAAAACAAATATAAACCAACCCAAAC					
Standard	STD_COL6A2_F	TTGGTAAGGTGGTGATGGTGAA					
Standard	STD_COL6A2_R	TCTTCTACCATCAACAAACATCC					
Standard	STD_COL6A2_Probe	AACTTTCACCATACAAATAATCACTTCC					
Standard	STD_COL6A2 Sequence	AGGTTTAGGAGAAGTTGTAGATGTTTTGTGGTGGGTTAGTAGTATTGGTAAGGTGGTGATGGTGAAGGAAGTGATTTTGTATGGTGAAGTTATTAATGTGGGATGTTTTGTATGGTAGAAGATTGTAGTTTGGGTTTTATTGTTGGTA					

*cMethDNA primers/probes = External, M Target and Standard types; QM-MSP primers/probes = External, M Target and U Target types

Supplementary Table 3B. cMethDNA and QM-MSP primer and probe sequences

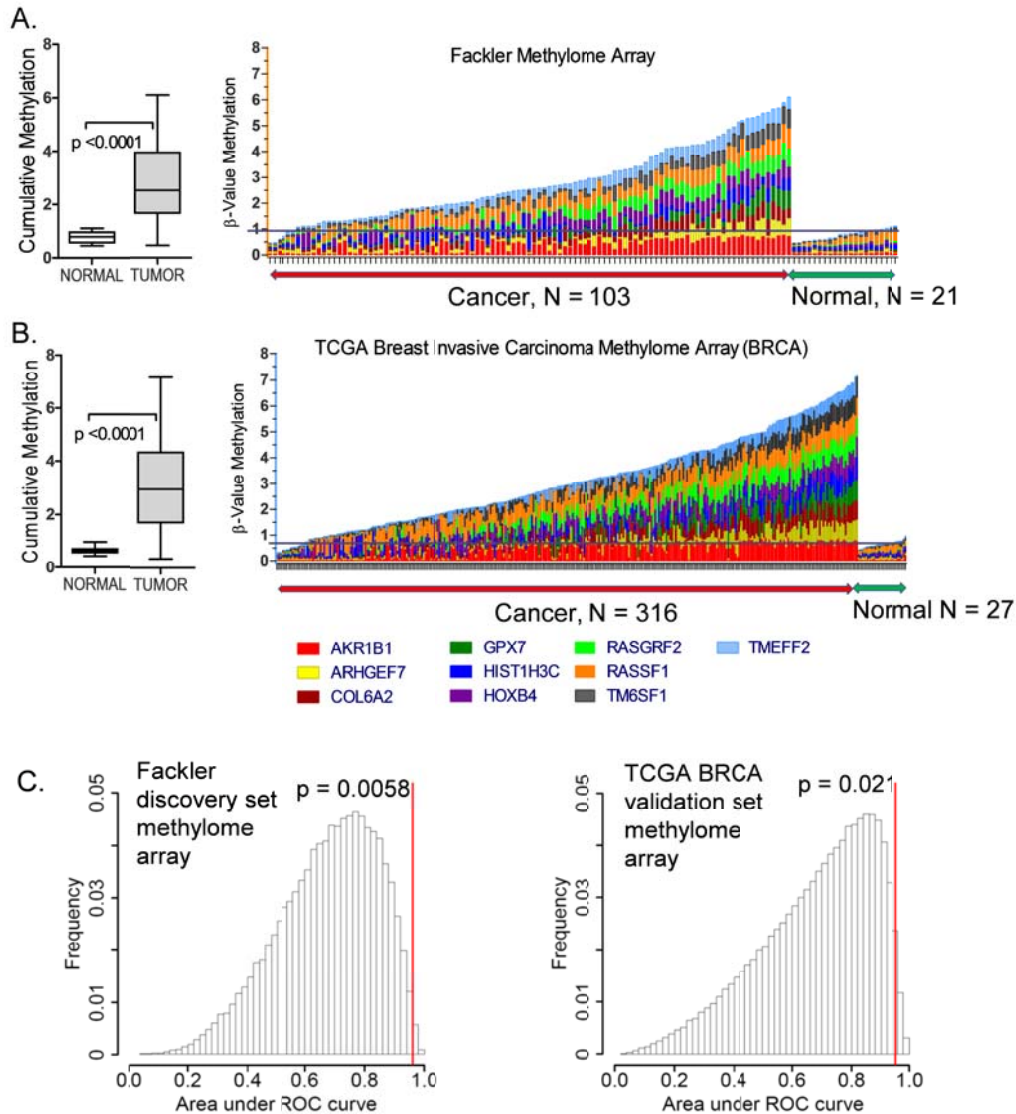
Type	Primer/Probe Name	5' To 3' Sequence	Array Target ID	TSS	M Primer Location	M Primer Distance to TSS	Size
External	GPX7_Ext_F	GGTGAATTGAGGTTTAGAG	cg16557944	Chr 1; 53068043	53068094	51 bp	299 bp
External	GPX7_Ext_R	ATACTTCTCCAACRACACCA					
Target	GPX7_FM	ACGGTGGTAGCGGCGTGGTT					93 bp
Target	GPX7_RM	ACCCCGAATATTAACCGCCTTA					
Target	GPX7_M_Probe	TACTACGCGCAAACCGCAACCCAC					
Target	GPX7_FUM	TGATGGTGTAGTGGTGTGG					95 bp
Target	GPX7_RUM	ACCCCAAAATTAACCAACCTTAA					
Standard	GPX7_U_Probe	CTACTACACACAAACCACAACCCAC					
Standard	STD_GPX7_F	TGATAGTATTAGAAAGGGATTGTAG					298, 94 bp
Standard	STD_GPX7_R	CAAACTCAACCACATCCAAACTC					
	STD_GPX7_Probe	AAATCACCTACCAATTCAACCATACCA					
Standard	STD_GPX7 Sequence	GGTGAATTGAGGTTTAGAGGATGGAGGATGATGTAATGTTGATGATAGTATTAGAAAGGGATTGTAGGAGGAGTTGGTATGGT TGAAATTGGTAGGTGATTTGGTTGTTGATTTGGATTTGGATGTGGTTAGATTGGATGAGTAGATGGTTAGAGTTAGGTTGTTATTTTT GGTATGGAAAGTGTAGTGAAAAAAATAGTGGTAGTTGTTGAATAGTTGTTGAGTTGATAGGTTGGTTGTATAGAAAAGTGGGA TTTTTGTGGGTAGTATAAAGTGGTGTGGTGGAGAAGTAT					
External	HIST1H3C_Ext_F2	GTGTGTGTTTTTATGTAAATGG	cg25438963	Chr 6; 26045639	26045655	16 bp	139 bp
External	HIST1H3C_Ext_R2	ATAAAATTTCTTACRCCACC					
Target	HIST1H3C_FM2	AATAGTTCGTAAGTTTATCGGCG					92 bp
Target	HIST1H3C_RM2	CTTACGCCACCGATAACCGA					
Target	HIST1H3C_M_Probe	TACTTACGCGAAACTTTACGCCGA					
Target	HIST1H3C_FUM2	GTAATAGTTGTAAAGTTTATTGGTG					95 bp
Target	HIST1H3C_RUM2	TTTCTTACACCACCAATAACCAA					
Standard	HIST1H3C_U_Probe	AACTACTTACACAAAACTTTACCACCAA					
Standard	STD_HIST1H3C_F	GATTTAGAGTTGGATGTGTGGAT					140, 92 bp
Standard	STD_HIST1H3C_R	ACCACCATACTAATAATCAAATCTA					
	STD_HIST1H3C_Probe	AAATATCACTCATCACCAAAATAATCCAA					
Standard	STD_HIST1H3C Sequence	GTGTGTGTTTTTATGTAAATGGTGGATTTAGAGTTGGATGTGTGGATGGAGTTTGGATTTTGGTGGTATGATATTTTGGTA TTGTTAGATTGATTATTAGTATGGTGGTTA GGTGGTGTGAAGAAATTTAT					
External	HOXB4_Ext_F	TTAGAGGYGAGAGAGTAGTT	cg06760035	Chr 17; 46655743	46655478	265 bp	224 bp
External	HOXB4_Ext_R	AAACTACTACTAACRCCCTC					
Target	HOXB4_FM	CGGGATTTTGGGTTTTTCGTCG					94 bp
Target	HOXB4_RM	CGACGAATAACGACGCAAAAAC					
Target	HOXB4_M_Probe	AACCGAACGATAACGAAAACGACGAA					
Target	HOXB4_FUM3	GTGGTGTGATTGTGTAGTGTTA					96 bp
Target	HOXB4_RUM2	CAAACCAAAACAATAACAAAACAAC					
Standard	HOXB4_U2_Probe	CAAAATCCCAACAAACCATATAACT					
Standard	STD_HOXB4_F	GTTAGTTTGTAGTGTATTGAGTAT					228, 94 bp
Standard	STD_HOXB4_R	CATCTTCCACAATAAACTTCCAATT					
	STD_HOXB4_Probe	TAACTCCACTATTCTACCTACCATTT					
Standard	STD_HOXB4 Sequence	TTAGAGGYGAGAGAGTAGTTAATAATGGTGTGATTGTTATGTTTTTGAATTGTTAGTTTTGTAGTGTATTGAGTATTTTGTTTGAT GAAATGGTAGGTAGAAATAGGTGGAGTTAGATAGTAATTGGAAGTTTATGTTGGAAGATGTTATTAGAATTGGTGTTTTTGGTGG TGATGTTTTTGGTATAATTTATTGTAGAAGAGAGGTTGTTAGTAGTATT					

*cMethDNA primers/probes = External, M Target and Standard types; QM-MSP primers/probes = External, M Target and U Target types

Supplementary Table 4: Comparison and reproducibility of the cMethDFNA assay using three DNA purification methods

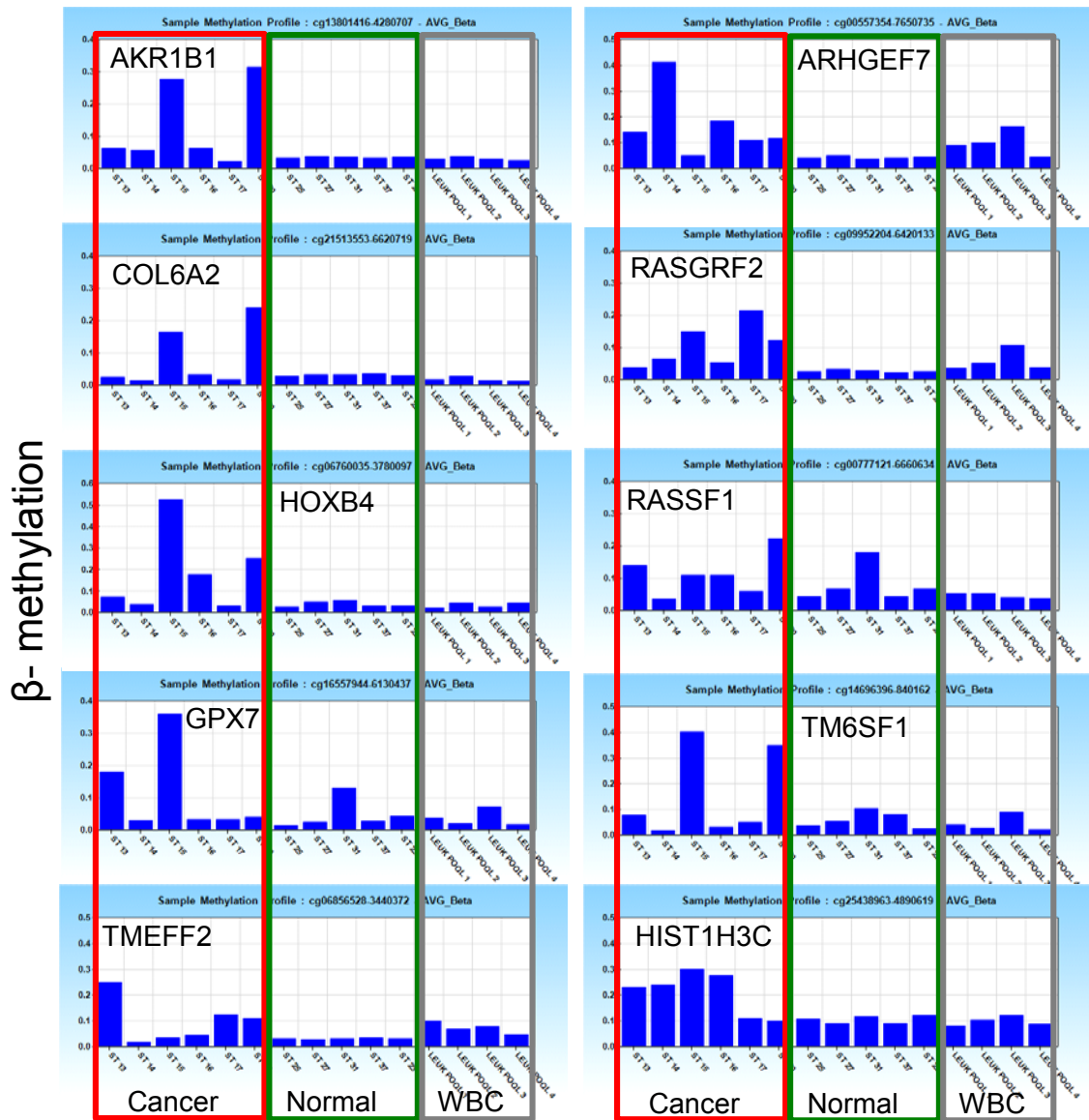
	QIAamp MinElute Virus Kit				
Methylated Copies	0	50	200	800	3200
Replicates	5	6	6	6	6
CV (%)	--	29.1	12.5	4.6	2.5
	QIAamp UltraSens Virus Kit				
Methylated Copies	0	50	200	800	3200
Replicates	1	6	6	6	6
CV (%)	--	43.9	9.7	6	3.4
	Zymo Research Quick gDNA				
Methylated Copies	0	50	200	800	3200
Replicates	5	6	6	6	6
CV (%)	--	38	21.5	11.3	7.3

Supplementary Figure 1



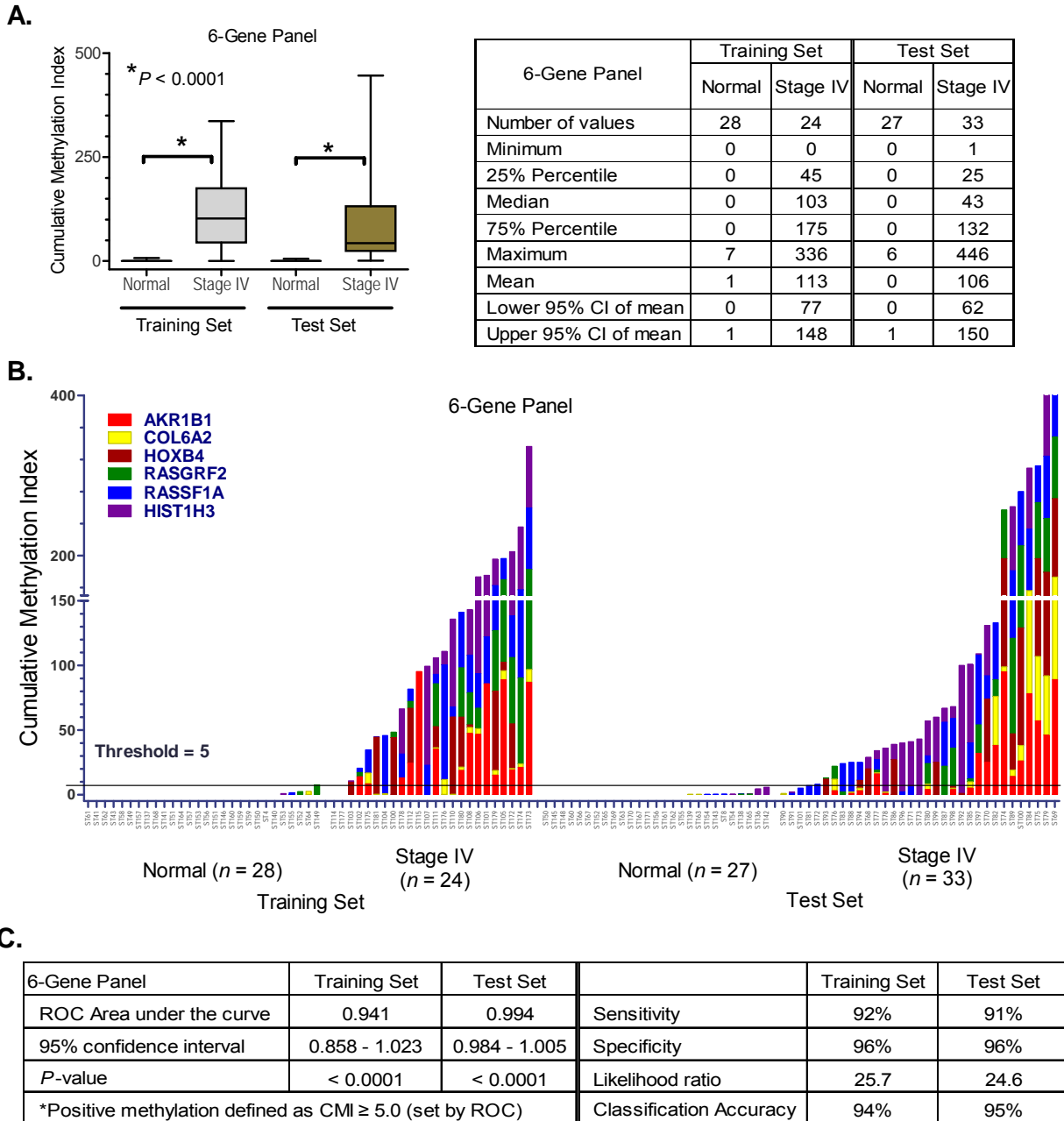
Supplementary Figure 1: Infinium Human Methylation27 tissue DNA methylation array profile of the 10-gene marker panel. Box-whiskers plots indicate average cumulative methylation among the 10 genes in cancer tissue compared to normal breast (box-whiskers plot, $p < 0.0001$; Mann Whitney). At right, histogram plots indicate level and frequency of methylation in individual samples (Histogram bar height indicates cumulative β -methylation of the 10 biomarkers, colors indicate individual genes). Cancer (red arrows) and normal (green arrows) samples are indicated below X-axis. Horizontal line indicates 95th percentile of normal sample methylation. A) Differential analysis was performed on the Fackler Methylome Array (GSE 31979; N = 103 breast cancers, N = 15 normal breast, N = 6 microdissected normal breast adjacent to tumor) to select samples frequently and highly methylated in cancer but not normal tissues. B) Data from TCGA (BRCA, N = 316 breast cancer, N = 27 normal breast tissues) support and independently validate the Fackler methylome array data. C) Computational analyses were performed on 100,000 randomly created 10-probe panels drawn from the Fackler methylome array study (left histogram) or from the TCGA BRCA methylome array study (right histogram), among 8376 probes with SD > 0.1 and probe detection p-value < 0.00001 in the Fackler array. Receiver operator characteristic (ROC) AUC values are plotted on the X-axis, AUC frequency of random gene panels is plotted on the Y-axis. The 10-gene test panel (red line) outperformed 97.9% of the random panels in the TCGA array ($p = 0.021$, one-sided) and 99.4% in the Fackler array ($p = 0.0058$, one-sided). In TCGA data, information on 911/8376 probes was suppressed because probes were located near single nucleotide polymorphisms.

Supplementary Figure 2



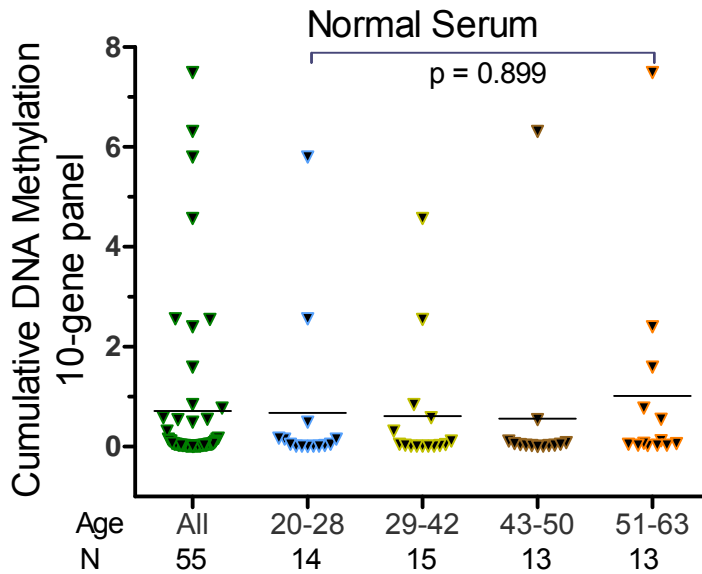
Supplementary Figure 2: Illumina Infinium Human Methylation27 serum DNA array profile of the 10-gene marker panel. Cancer-specific biomarkers were selected with very low background methylation in normal sera (Normal) or leukocyte pools (WBC), compared to cancer sera (Cancer), and with tumor-specific hypermethylation in a previous breast tissue array (Supplementary Figure 1). Histogram plots show relative methylation (Y-axis) of individual cancer (red box) and normal (green box) samples (X-axis). Data was deposited in the Gene Expression Omnibus (GEO) repository (accession #GSE52621).

Supplementary Figure 3



Supplementary Figure 3: cMethDNA assay using a 6-gene marker panel. Six most robust markers in the 10-gene panel were identified by Mann-Whitney test, and then evaluated as a minimal marker set. A) Box plot shows that cancer sera display significantly higher average cumulative methylation for the panel markers than normal controls ($p < 0.0001$, Mann-Whitney). B) Within individual samples, cumulative cMethDNA assay values (CMI; Y-axis; indicated by bar height) are shown, each colored segment representing the methylation index for an individual gene, indicated by color. The Training set consisted of patients (N = 24) from trial J0425/J0214 and normal control women (N = 28). ROC analysis was performed on data collected from this set of individuals to define a normal laboratory methylation threshold (CMI \geq 5 units; Y-axis). The Test set consisted of patients (N = 33), and normal control women (N = 27), from trial TCRC 005. C) Statistical analysis of Training and Test sets for the 6-gene panel. Performance was nearly identical to the 10-gene panel.

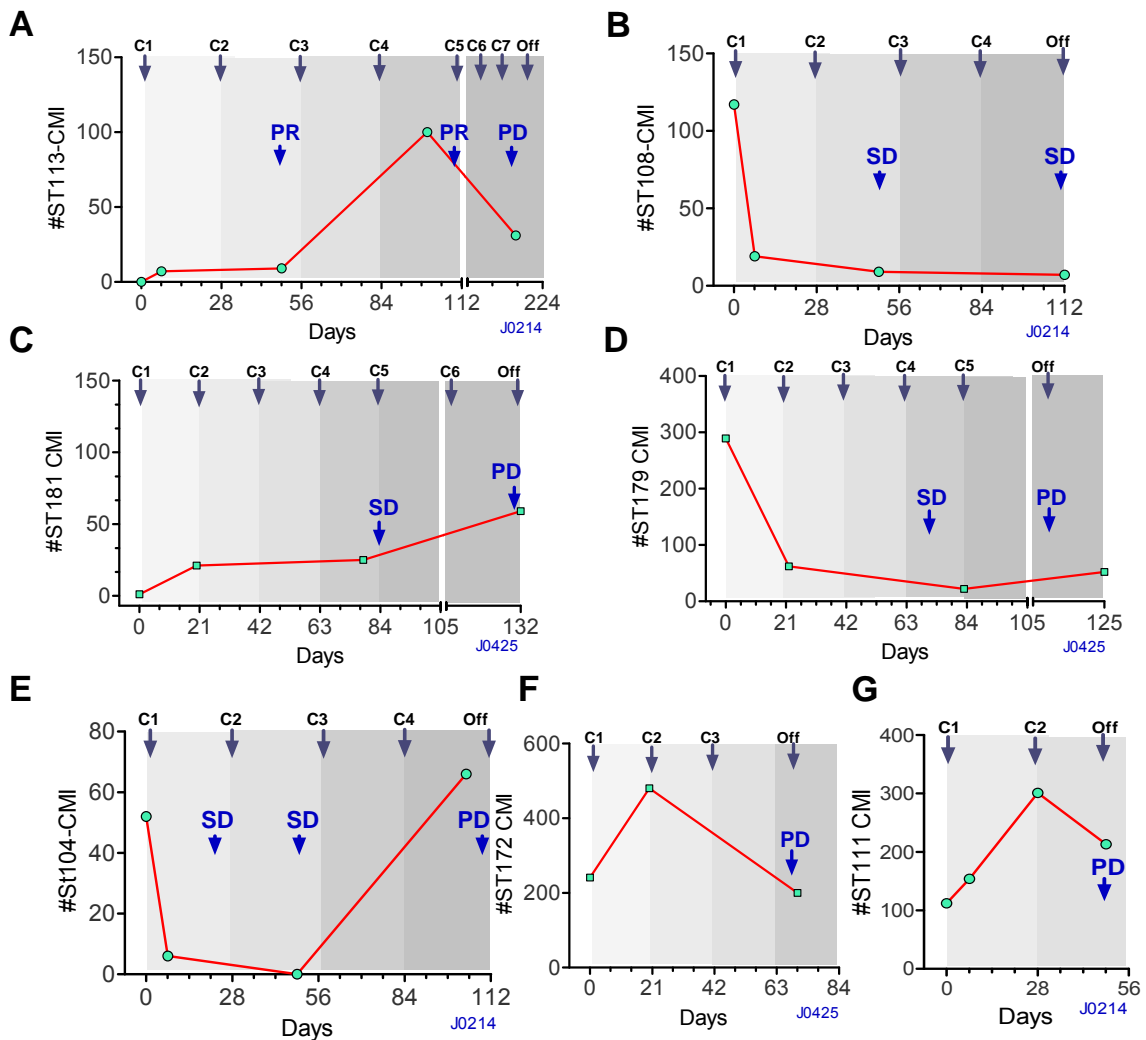
Supplementary Figure 4



	All	20-28	29-42	43-50	51-63
Donor Age					
Number of women	55	14	15	13	13
Cumulative Methylation:					
Minimum	0.00	0.00	0.01	0.00	0.02
25% Percentile	0.02	0.01	0.01	0.02	0.03
Median	0.05	0.05	0.04	0.04	0.07
75% Percentile	0.49	0.25	0.58	0.10	1.18
Maximum	7.49	5.80	4.57	6.31	7.49
Mean	0.71	0.68	0.61	0.56	1.02
Std. Deviation	1.65	1.62	1.28	1.73	2.08
Std. Error	0.22	0.43	0.33	0.48	0.58
Lower 95% CI of mean	0.27	-0.26	-0.10	-0.49	-0.24
Upper 95% CI of mean	1.16	1.61	1.32	1.61	2.27

Supplementary Figure 4: Scatter plot and data analysis of cumulative methylation of the 10-gene panel in normal serum DNA. Age ranged from 20-63 years, (median = 42.0, mean = 40.6 years). Quartile ranges are indicated. No significant difference among quartiles was observed ($p = 0.899$, one-way ANOVA; $p = 0.420$, Kruskal-Wallis), nor between youngest and oldest quartiles ($p = 0.643$, unpaired t test with Welch's correction; $p = 0.224$, Mann-Whitney test).

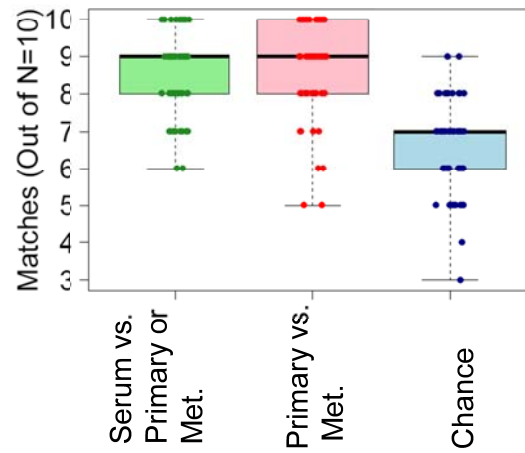
Supplementary Figure 5



Supplementary Figure 5. Monitoring of treatment response by the cMethDNA assay.

Representative plots of CMI of patient (ST#) sera assessed by cMethDNA at different time points during the course of treatment. Patients (a-g) were given either 28 day cycles of docetaxel or 21 day cycles of capecitabine, as indicated by shading. C: cycles of treatment; PD, SD, PR: Imaging- and RECIST criteria-assessed progressive disease, stable disease or partial response, respectively.

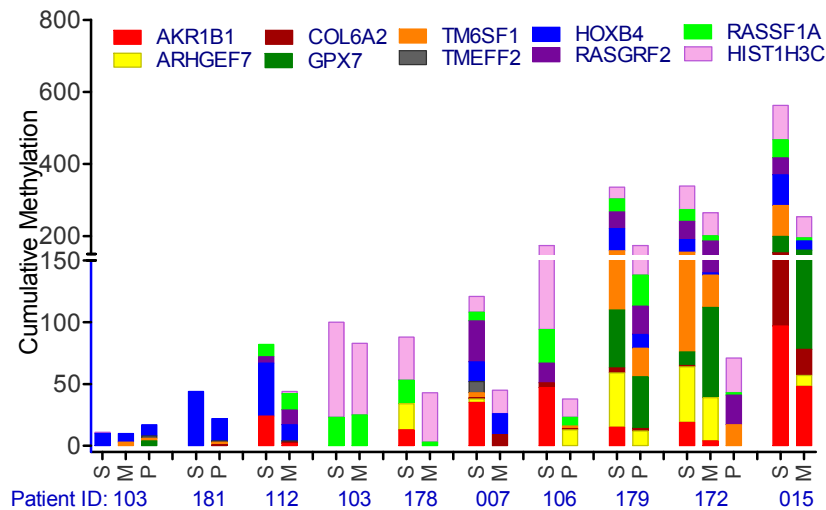
Supplementary Figure 6



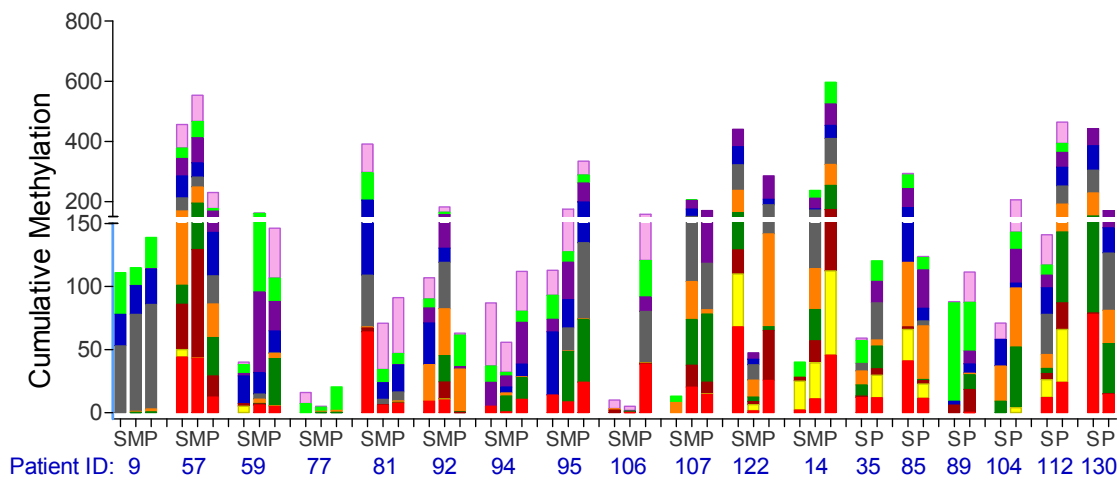
Supplementary Figure 6: Methylation Profiles in serum, primary and metastatic tumor DNA. Cumulative methylation in DNA from serum and metastatic tissue collected at autopsy were compared to each other. A match is indicated where gene was hypermethylated in two or more samples of the same individual (Figure 5). By chance alone, median match = 7. Results indicate samples within individuals were more alike than between individuals.

Supplementary Figure 7

A.

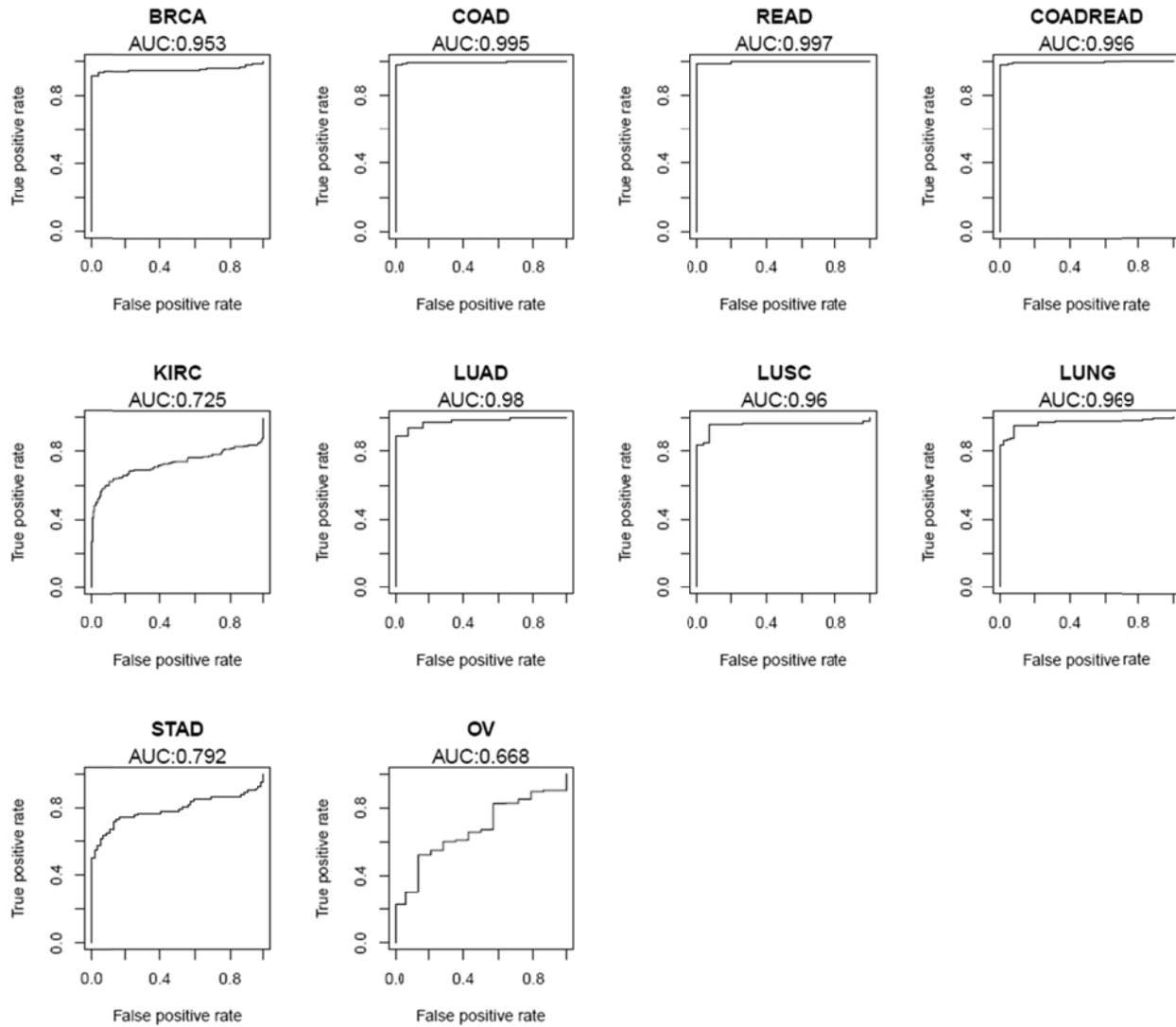


B.



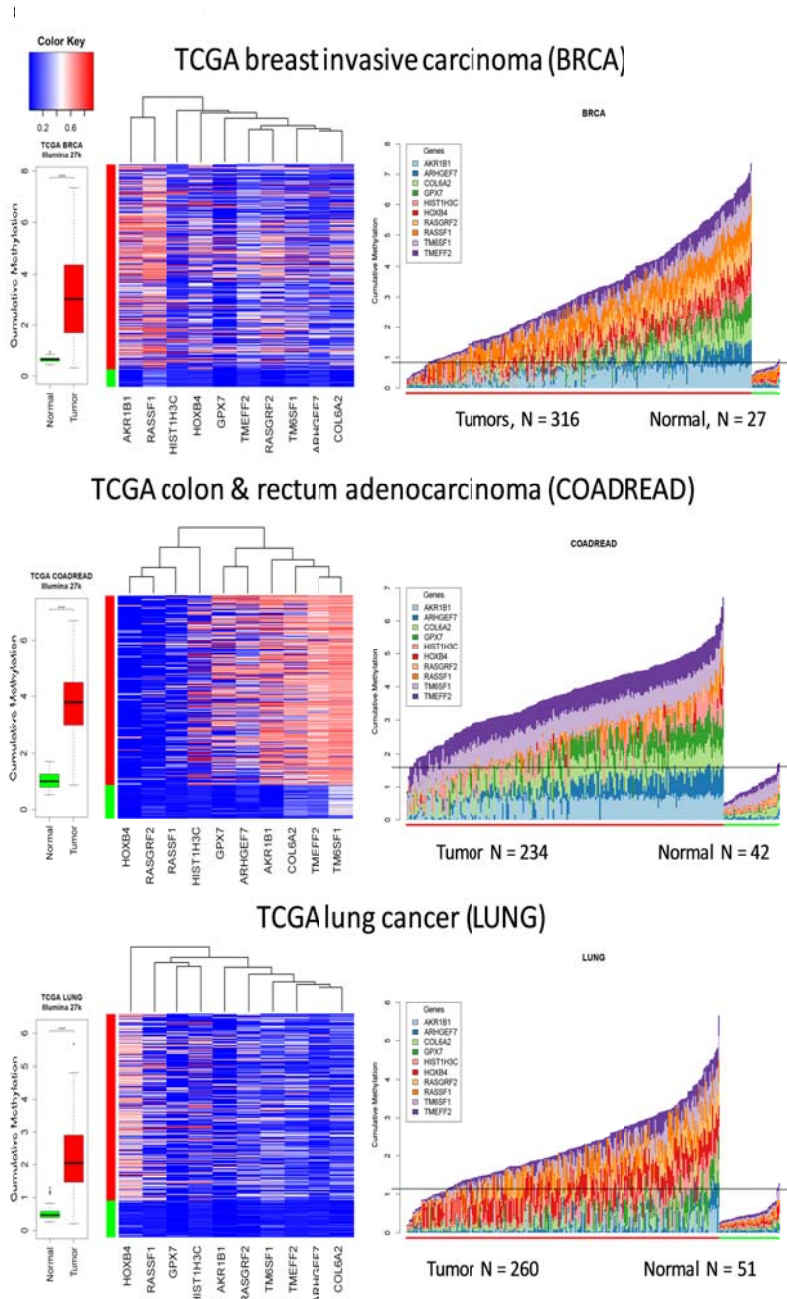
Supplementary Figure 7. Methylation profiles of serum, primary tumor and/or metastasis in patients with Stage 4 breast cancer. Serum (S), tissue biopsy of the metastasis (M) and surgically resected primary (P) (formalin fixed paraffin embedded tumor sections) from the same patient were tested by cMethDNA (serum) and QM-MSP (tissues) assays. A) Samples from a subset of patients with metastatic breast cancer in the Training set ($n = 10$) whose serum was collected at the time of diagnosis of recurrent metastatic breast cancer. B) Serum, biopsy of the distant metastasis, and primary tumor samples from patients with newly diagnosed Stage 4 breast cancer (TBCRC013; $n = 25$) were analyzed by cMethDNA and QM-MSP. Source of sample and Patient ID are indicated on the X-axis and the cumulative methylation index (CMI) is shown on the Y-axis, each colored segment representing one gene.

Supplementary Figure 8A



Supplementary Figure 8A: Performance of the 10-gene panel in arrays of a variety of tumor types. TCGA Illumina Infinium HumanMethylation27 array data was downloaded from the UC Santa Cruz Cancer Genomics Browser <https://genome-cancer.ucsc.edu/>. A) ROC analyses were performed to identify the extent to which the 10-gene markers distinguish tumor from normal. AUC is indicated.

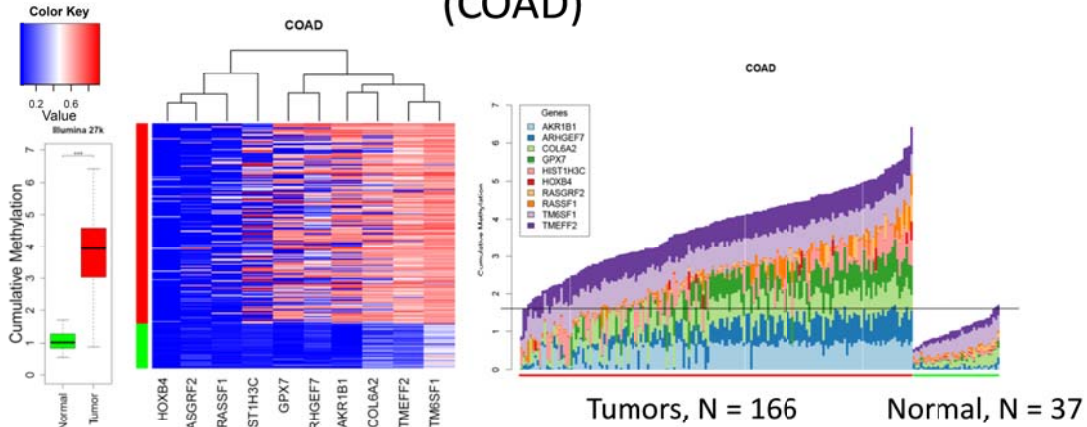
Supplementary Figure 8B



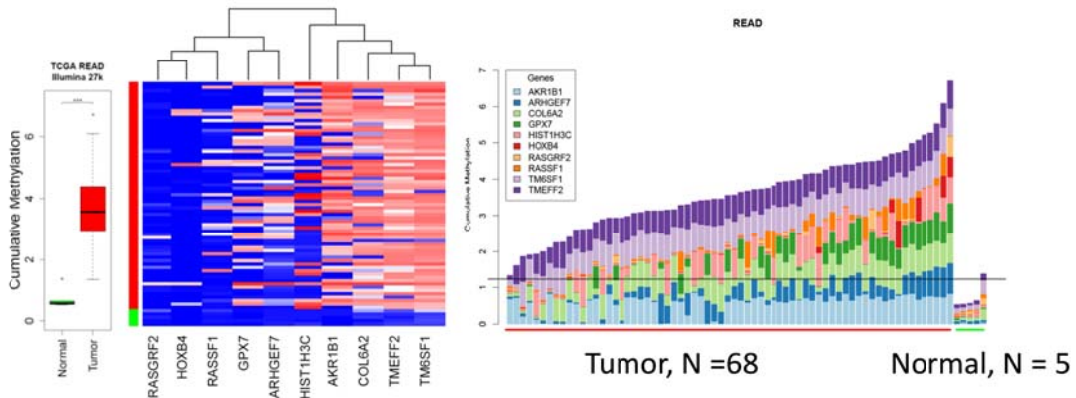
Supplementary Figure 8B: Performance of the 10-gene panel in arrays of a variety of tumor types. TCGA Illumina Infinium HumanMethylation27 array data was downloaded from the UC Santa Cruz Cancer Genomics Browser. Significant differences between median β -methylation in samples of tumor (red) and normal (green) tissue are indicated by box plot p-values: * = <0.05 , ** = <0.01 , *** = <0.001 . Two-dimensional hierarchical analyses were performed based on Euclidian distances, samples on the Y-axis (green = normal, red = tumor) and biomarkers on the X-axis; array β -methylation levels are indicated by the Color Key, (blue low, red high). Histogram bar plots are shown for the 10-gene panel, samples on the X-axis, and cumulative β -methylation within sample on the Y-axis; each gene is indicated by a different color, shown by the legend. Horizontal line represents the 95th percentile of cumulative methylation in normal samples.

Supplementary Figure 8C

TCGA colon adenocarcinoma (COAD)

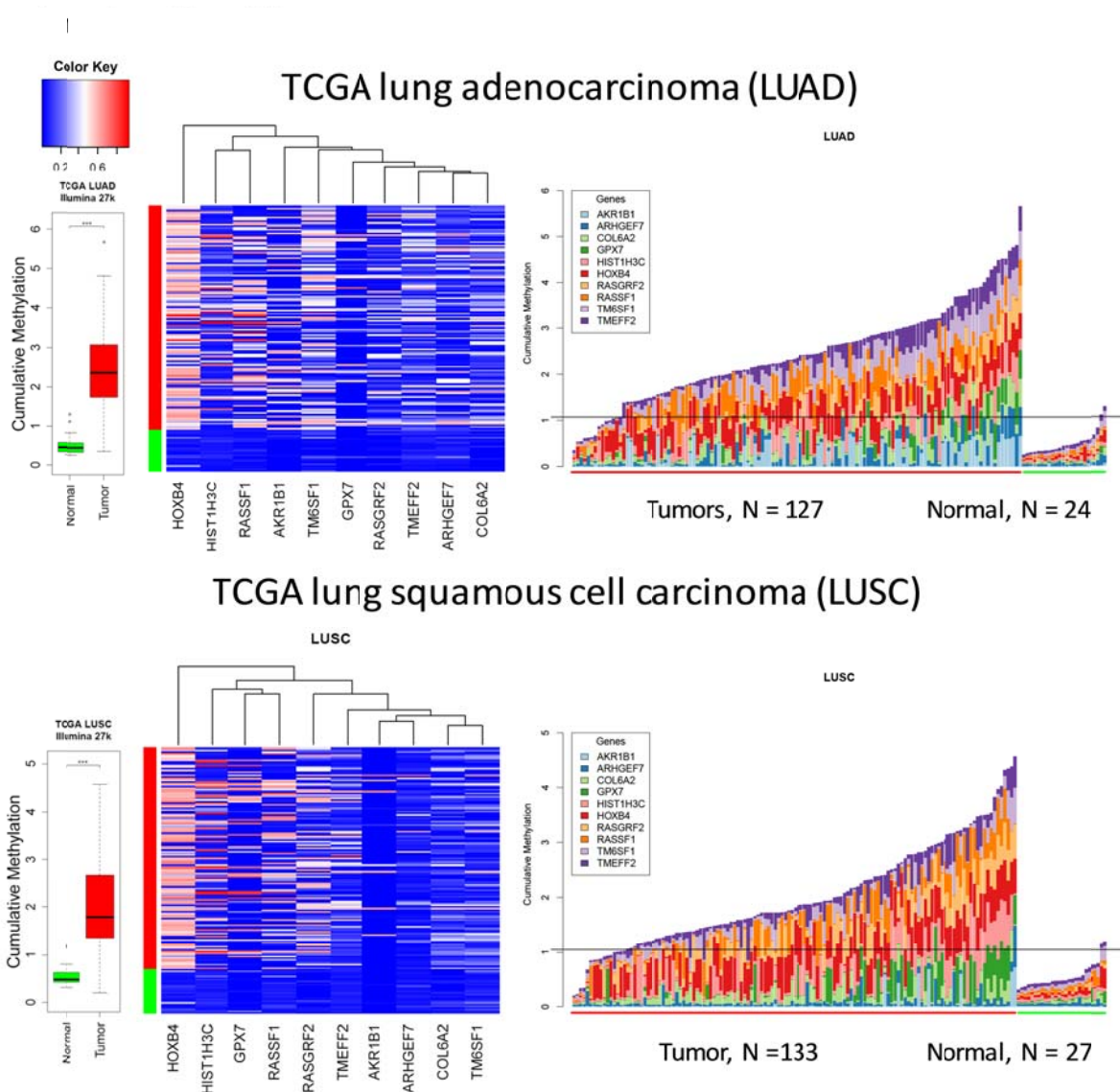


TCGA rectum adenocarcinoma (READ)



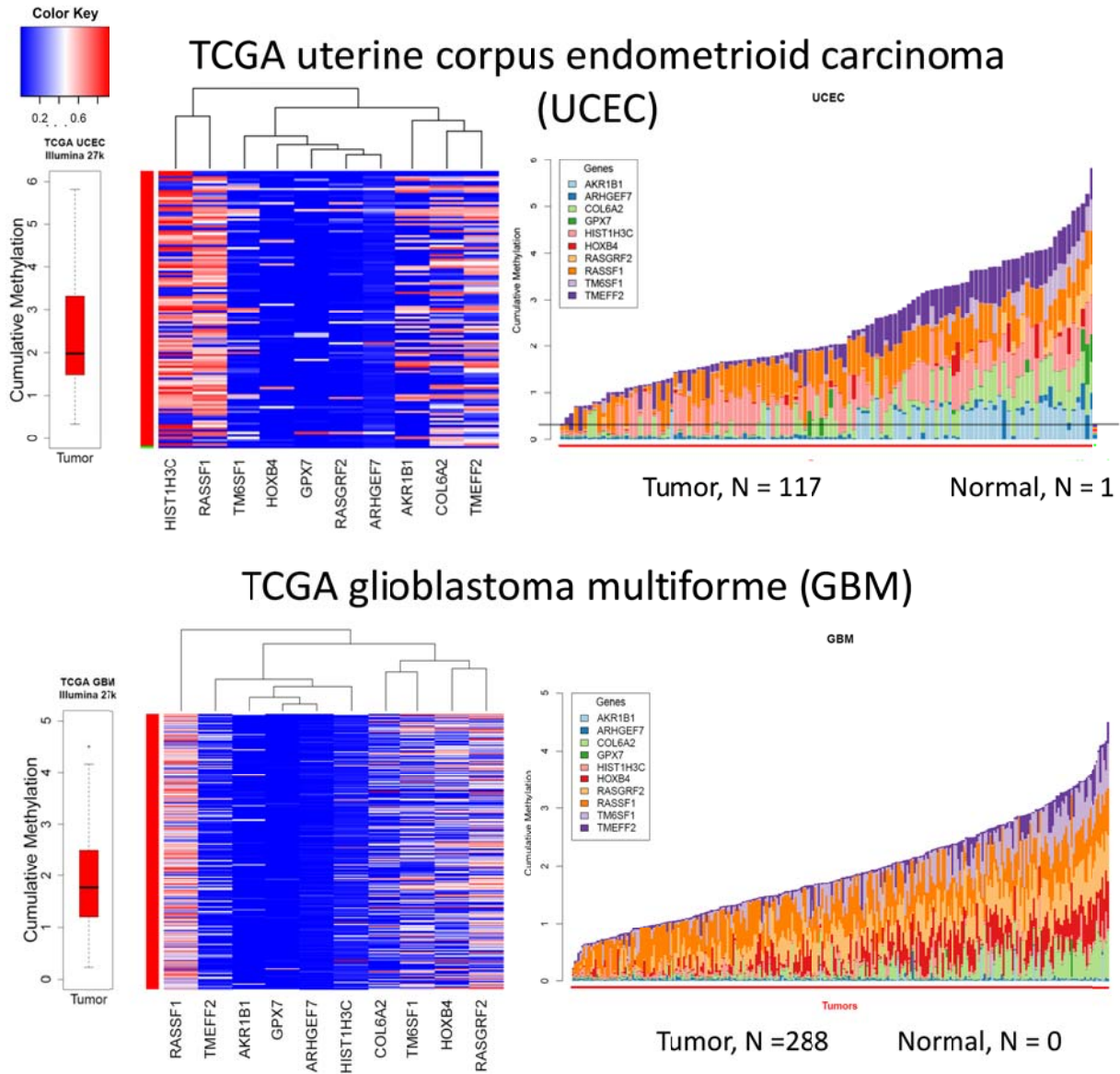
Supplementary Figure 8C: Performance of the 10-gene panel in arrays of a variety of tumor types. TCGA Illumina Infinium HumanMethylation27 array data was downloaded from the UC Santa Cruz Cancer Genomics Browser. Significant differences between median β -methylation in samples of tumor (red) and normal (green) tissue are indicated by box plot p-values: * = <0.05 , ** = <0.01 , *** = <0.001 . Two-dimensional hierarchical analyses were performed based on Euclidian distances, samples on the Y-axis (green = normal, red = tumor) and biomarkers on the X-axis; array β -methylation levels are indicated by the Color Key, (blue low, red high). Histogram bar plots are shown for the 10-gene panel, samples on the X-axis, and cumulative β -methylation within sample on the Y-axis; each gene is indicated by a different color, shown by the legend. Horizontal line represents the 95th percentile of cumulative methylation in normal samples.

Supplementary Figure 8D



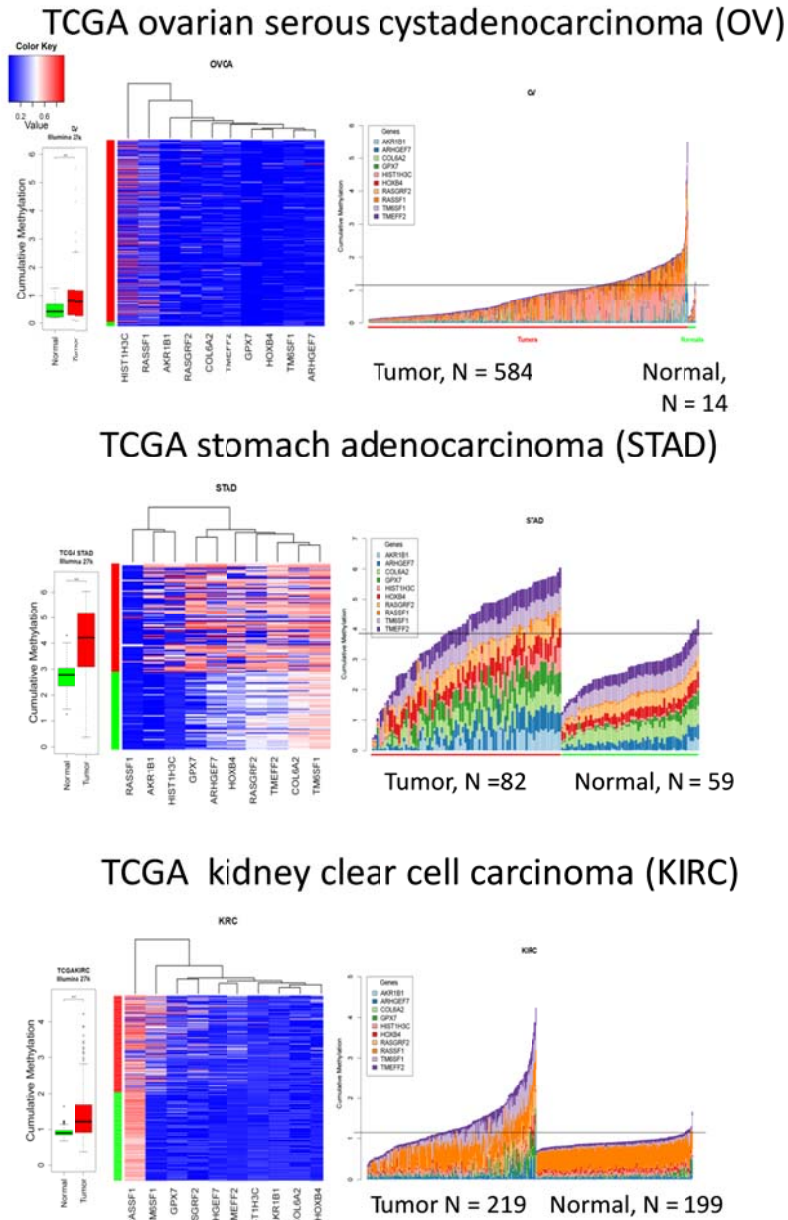
Supplementary Figure 8D: Performance of the 10-gene panel in arrays of a variety of tumor types. TCGA Illumina Infinium HumanMethylation27 array data was downloaded from the UC Santa Cruz Cancer Genomics Browser. Significant differences between median β -methylation in samples of tumor (red) and normal (green) tissue are indicated by box plot p-values: * = <0.05 , ** = <0.01 , *** = <0.001 . Two-dimensional hierarchical analyses were performed based on Euclidian distances, samples on the Y-axis (green = normal, red = tumor) and biomarkers on the X-axis; array β -methylation levels are indicated by the Color Key, (blue low, red high). Histogram bar plots are shown for the 10-gene panel, samples on the X-axis, and cumulative β -methylation within sample on the Y-axis; each gene is indicated by a different color, shown by the legend. Horizontal line represents the 95th percentile of cumulative methylation in normal samples.

Supplementary Figure 8E



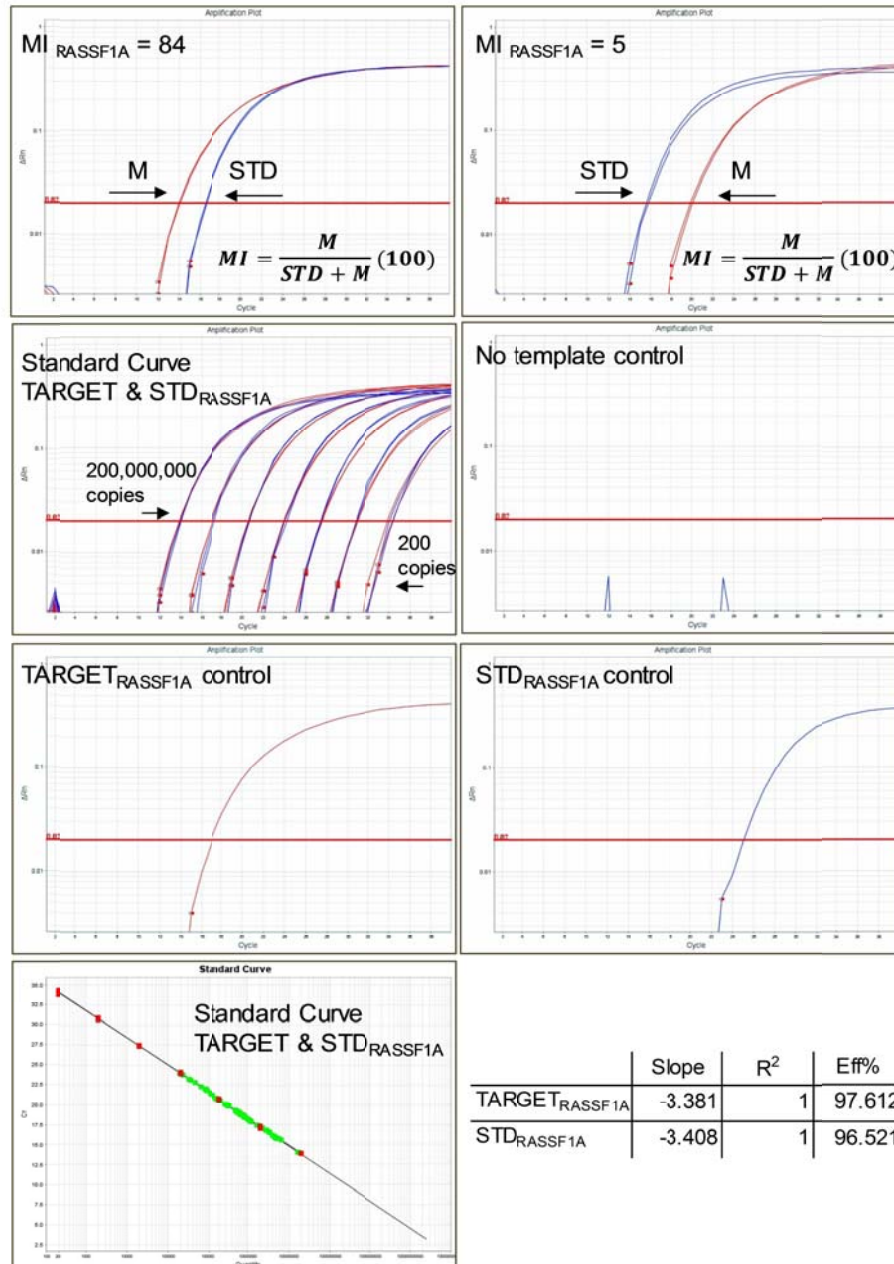
Supplementary Figure 8E: Performance of the 10-gene panel in arrays of a variety of tumor types. TCGA Illumina Infini HumanMethylation27 array data was downloaded from the UC Santa Cruz Cancer Genomics Browser. Significant differences between median β -methylation in samples of tumor (red) and normal (green) tissue are indicated by box plot p-values: * = <0.05 , ** = <0.01 , *** = <0.001 . Two-dimensional hierarchical analyses were performed based on Euclidian distances, samples on the Y-axis (green = normal, red = tumor) and biomarkers on the X-axis; array β -methylation levels are indicated by the Color Key, (blue low, red high). Histogram bar plots are shown for the 10-gene panel, samples on the X-axis, and cumulative β -methylation within sample on the Y-axis; each gene is indicated by a different color, shown by the legend. Horizontal line represents the 95th percentile of cumulative methylation in normal samples.

Supplementary Figure 8F



Supplementary Figure 8F: Performance of the 10-gene panel in arrays of a variety of tumor types. TCGA Illumina Infinium HumanMethylation27 array data was downloaded from the UC Santa Cruz Cancer Genomics Browser. Significant differences between median β -methylation in samples of tumor (red) and normal (green) tissue are indicated by box plot p-values: * = <0.05 , ** = <0.01 , *** = <0.001 . Two-dimensional hierarchical analyses were performed based on Euclidian distances, samples on the Y-axis (green = normal, red = tumor) and biomarkers on the X-axis; array β -methylation levels are indicated by the Color Key, (blue low, red high). Histogram bar plots are shown for the 10-gene panel, samples on the X-axis, and cumulative β -methylation within sample on the Y-axis; each gene is indicated by a different color, shown by the legend. Horizontal line represents the 95th percentile of cumulative methylation in normal samples.

Supplementary Figure 9



Supplementary Figure 9: Representative examples of amplification plots of cMethDNA real-time PCR. Real-time PCR was achieved using the absolute quantitation method with the ABI 7500 Real Time PCR System. Copy number of TARGET_{RASSF1A} (red curve) and STD_{RASSF1A} (blue curve) is extrapolated from a set of two standard curves, shown as serial 1:10 dilutions of TARGET/STD overlapping curves. No template control (multiplexed water only), TARGET_{RASSF1A} control (universally methylated bisulfite converted template), and STD_{RASSF1A} control (eg.STD_{RASSF1A} plasmid) were all derived from the multiplex PCR reaction in Step #1 and tested by real-time PCR in Step #2. Statistical evaluation of the standard curves for linearity and efficiency is shown.