Supplementary Information

Global DNA Hypomethylation Coupled to Repressive Chromatin Domain Formation and Gene Silencing in Breast Cancer

Gary Hon[1], R. David Hawkins[1], Otavia L. Caballero[3], Christine Lo[4], Ryan Lister[6], Mattia Pelizzola[6], Armand Valsesia[7], Zhen Ye[1], Samantha Kuan[1], Lee E. Edsall[1], Anamaria Aranha Camargo[5], Brian J Stevenson[7], Joseph R. Ecker[6], Vineet Bafna[4], Robert L. Strausberg[3,8], Andrew J. Simpson[3,8], & Bing Ren[1,2]

[1]Ludwig Institute for Cancer Research,
[2]Department of Cellular and Molecular Medicine, Moores Cancer Center, and Institute of Genomic Medicine, University of California San Diego, La Jolla, California 92093, USA
[3]Ludwig Collaborative Laboratory for Cancer Biology and Therapy, Department of Neurosurgery, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
[4]Department of Computer Science, University of California-San Diego, San Diego, CA, USA
[5]Ludwig Institute for Cancer Research, Sao Paulo, Brazil
[6]Genomic Analysis Laboratory, Howard Hughes Medical Institute, and The Salk Institute for Biological Studies, La Jolla, California 92037, USA
[7]Swiss Institute of Bioinformatics, Ludwig Institute for Cancer Research, 1015 Lausanne, University of Lausanne, 1011 Lausanne, Switzerland
[8]Ludwig Institute for Cancer Research Ltd, New York, New York, United States of America

Correspondence to:
Bing Ren[1,2]
biren@ucsd.edu
Phone: 858 822 5766
Fax: 858 534 7750

Running Title: DNA Hypomethylation Linked to Chromatin Silencing
Keywords: epigenetics, cancer, chromatin, hypomethylation

Table of Contents

1. Supplementary Figures and Legends

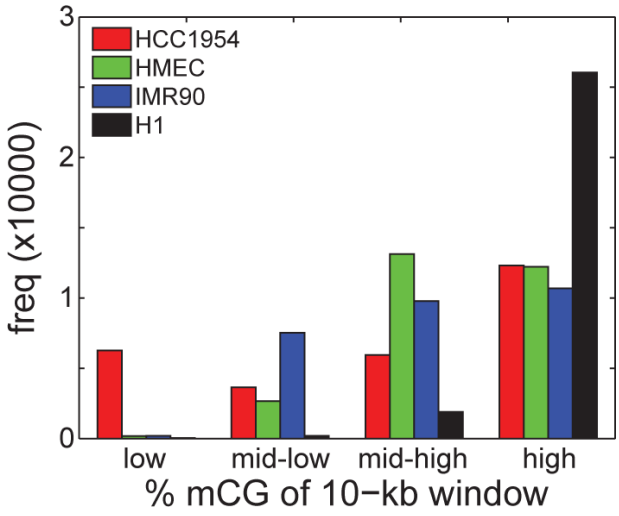

Figure S1 – Distribution of low, mid-low, mid-high, and high mCG windows.
Frequency of low [0% to 25%), mid-low [25% to 50%), mid-high [50% to 75%), and high [75% to 100%] mCG in 10kb bins spanning the human genome for various cells.

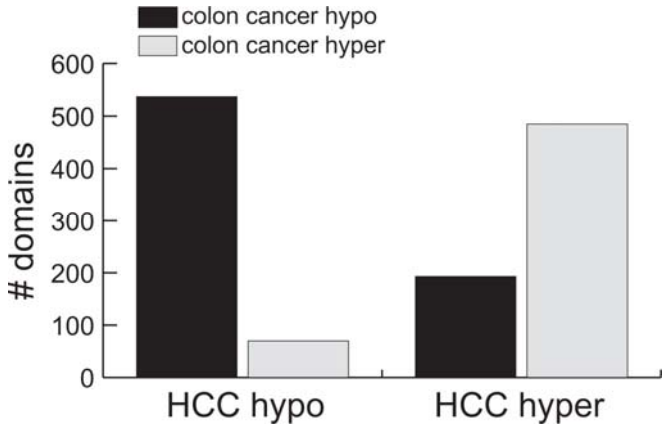

Figure S2 – Hypomethylation commonalities in colon cancer and HCC1954.
A set of 730 and 555 domains found hypomethylated and hypermethylated, respectively, in colon cancer (Irizarry et al. 2009) were examined for methylation change in HCC1954. Hypo: hypomethylated; hyper: hypermethylated.

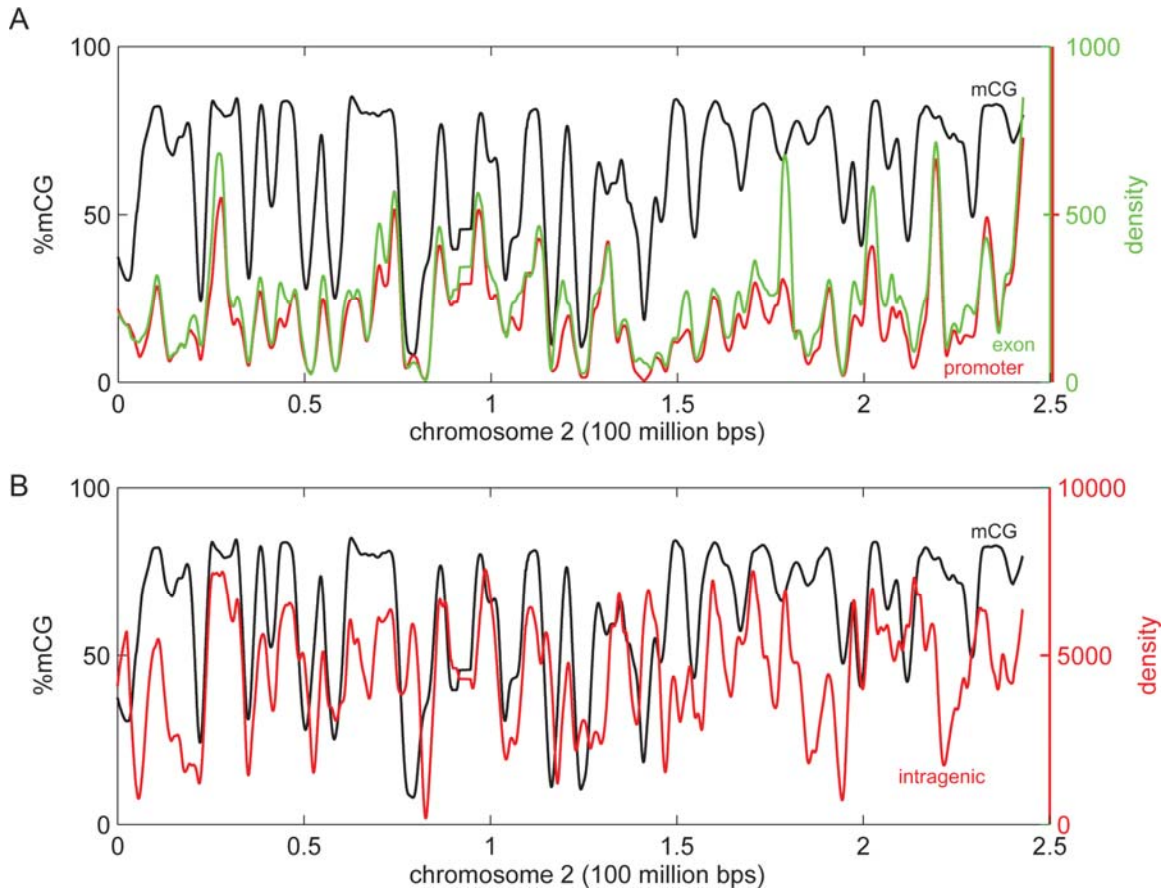

Figure S3 – Correlation of mCG and genic features.

(A) Distribution of %mCG (black) and promoter (red) / exon (green) density on chromosome 2 for HCC1954. Density is expressed as the number of base pairs of the gene feature in 10-kb windows spanning the chromosome. Promoters are defined to be +/- 1000 from the TSS.

(B) Distribution of %mCG (black) and intragenic density (red) on chromosome 2 for HCC1954. Density is expressed as the number of base pairs of gene bodies in 10-kb windows spanning the chromosome.

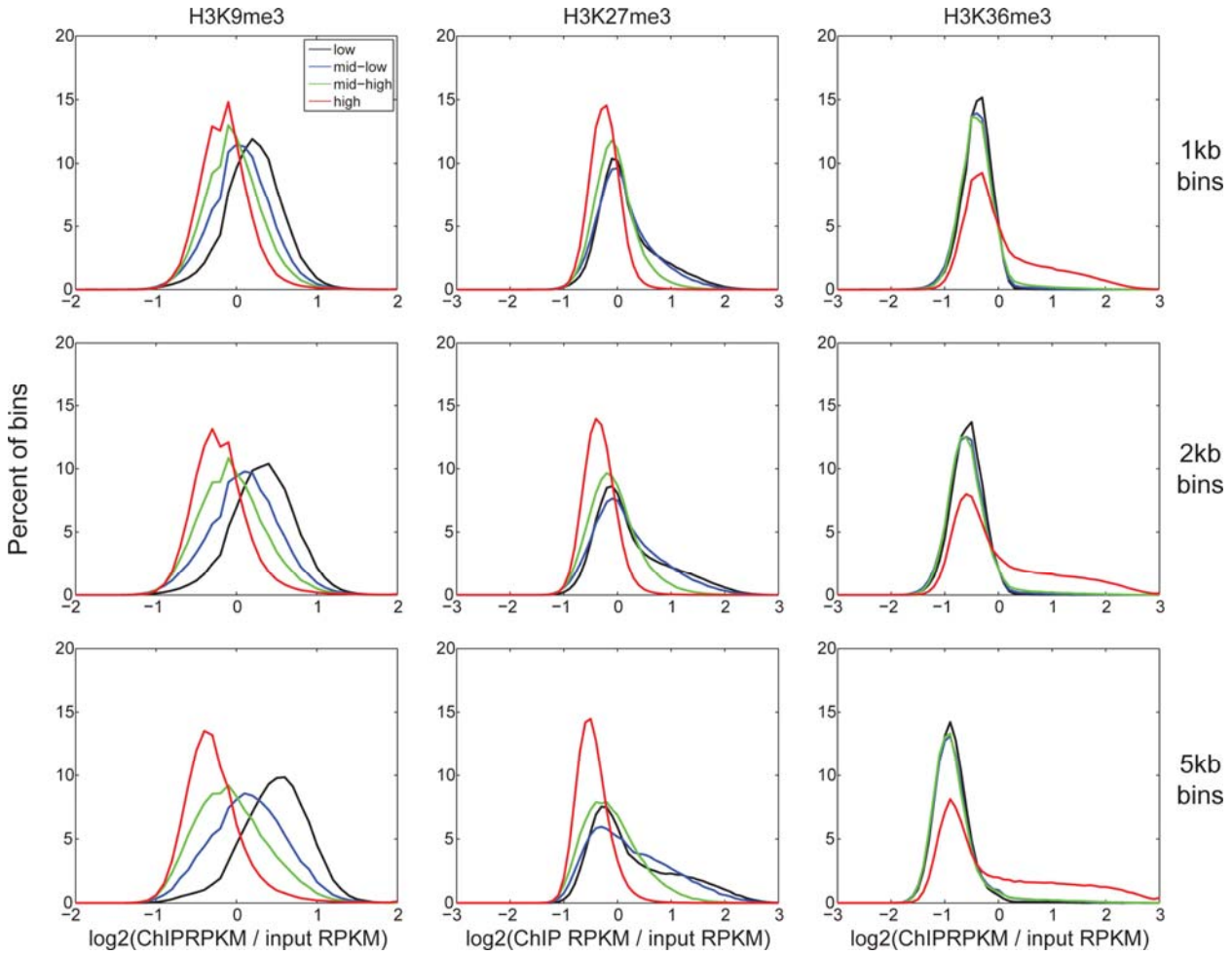

Figure S4 –H3K9me3/H3K27me3 opposes DNA methylation exclusion, various bin sizes
As in Figure 2D-E, but using bin sizes of 1-kb (top), 2-kb (middle), and 5-kb (bottom).

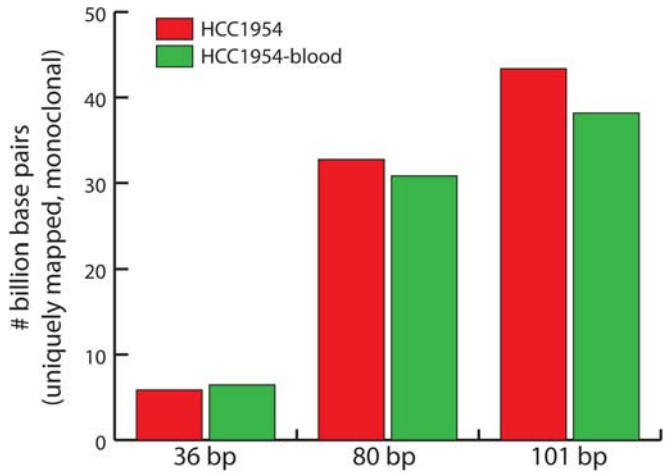

Figure S5 – Length distribution of genome sequencing.
Each end of paired-end genomic libraries were sequenced either 36, 80, or 101 base pairs.

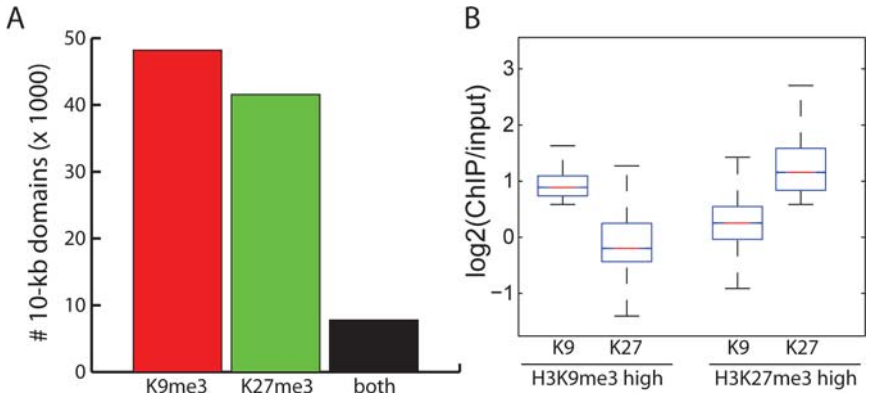

Figure S6 – H3K9me3 and H3K27me3 are exclusive of each other.

(A) The number of 10kb windows having (ChIP RPKM / input RPKM) ≥ 0.5 for H3K9me3 and H3K27me3, as well as where both marks are enriched.

(B) Enrichment of H3K9me3 and H3K27me3 at H3K9me3-enriched 10kb windows and H3K27me3-enriched 10kb windows.

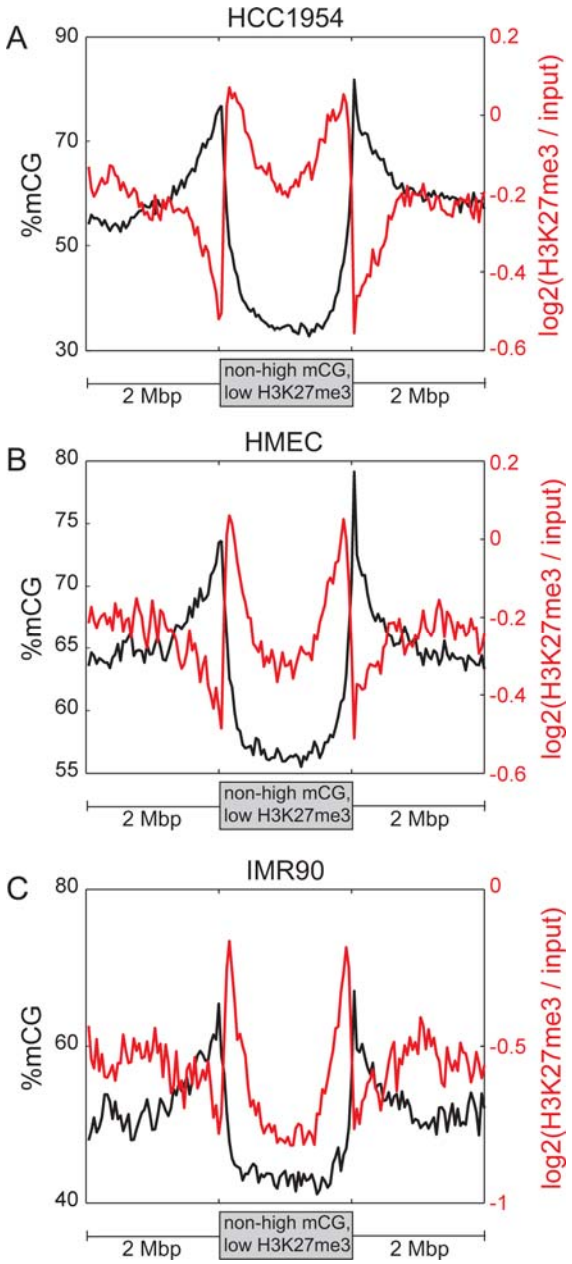

Figure S7 – Flanking H3K27me3 enrichment around PMDs.
Enrichment of H3K27me3 and %mCG around domains of non-high %mCG coinciding with low H3K27me3 in (A) HCC1954, (B) HMEC, and (C) IMR90.

2. Supplementary Methods

2.1. Cell culture

HCC1954 cells were grown in RPMI 1640 media made fresh from powder (Gibco 31800-089) according to manufacturer instructions. Final media was supplemented with 10% FBS, 1X non-essential amino acids (Invitrogen 11140050), and 1X L-glutamine (Invitrogen 45000-676-1). Cells were grown at 37°C/5% CO2, and split with 0.1% Trypsin/0.02% EDTA. Cells submitted to epigenetic analysis were between passage 34 and 42.

Cryo-preserved HMECs at passage 6-7 were purchased from Lonza (CC-2551) and grown according to manufacturer instructions at 37°C/5% CO_2. The cells were split 2 times before harvesting.

Depending on the number of population doublings, HMECs grown in culture enter distinct transition states on the way to becoming immortal cancer cells(Romanov et al. 2001). The earliest stage during primary culture, called pre-stasis, is characterized by high levels of p16/CDKN2A expression prior to entering an initial growth plateau called stasis or selection. RNA-seq expression data indicates that harvested HMECs highly express p16/CDKN2A (RPKM = 16.5968), and methylC-seq data indicates the promoter of this gene is not DNA methylated, suggesting that these cells have not reached stasis(Novak et al. 2008). Furthermore, harvested HMECs express KLK6 (RPKM = 28.02643), COX7A1 (RPKM = 0.43879), EPCAM (RPKM = 23.4664), KRT19 (RPKM = 22.9986), and PRDM1 (RPKM = 2.08479), all of which are hallmarks of early passage pre-stasis HMECs that are not expressed in either late passage pre-stasis HMECs or post-stasis HMECs. Finally, harvested HMECs express TP53 (RPKM = 18.83497), indicating that they have not entered crisis(Garbe et al. 2009).

2.2. Genome sequencing library preparation

2.2.1. Genomic DNA sonication

2.2.1.1. Extracted DNA was sonicated using the Diagenode Biorupter on high for 30 cycles (30 sec on, 1 min off)

2.2.2. End repair

2.2.2.1. Prepare fresh ER mix: 10 µl T4 DNA ligase buffer with 10 mM ATP, 4 µl 10 mM dNTP mix, 5 µl 3U/µl T4 DNA polymerase, 1 µl Klenow 5U/µl DNA polymerase, 5 µl T4 PNK, 75 µl of sonicated DNA and water.

2.2.2.2. Incubate 30min at 20C.

2.2.2.3. Purify cDNA using MinElute purification according to manufacturer instructions. Elute twice in 16 µl EB.

2.2.3. A-tailing

2.2.3.1. Prepare A-tailing reaction mix per reaction: 5 µl 10X Klenow buffer, 10 µl 1 mM dATP, 3 µl 5U/µl Klenow 3'-5' exo minus.

2.2.3.2. Mix 16 µl to DNA sample.

2.2.3.3. Incubate 30 min at 37°C.

2.2.3.4. Purify cDNA using MinElute purification according to manufacturer instructions. Elute once in 10 µl EB.

2.2.4. Adapter ligation

2.2.4.1. To each reaction, add 25 µl 2X DNA ligase buffer, 1 µl 1:10 sequencing adapters, 5 µl DNA ligase (Quick Ligation Kit (NEB Cat.M2200L)), 9 µl water.
2.2.4.2. Mix tube gently, spin briefly, incubate at room temperature for 15 min at 20C.
2.2.4.3. Perform MinElute PCR cleanup. Elute twice in 15 µl EB.

2.2.5. Size selection

2.2.5.1. Run samples in 8% polyacrylamide gel.
2.2.5.2. Excise 280-320bp main fragment.
2.2.5.3. Put gel pieces in 0.5mL tube (with bottom of tube punctured 3 times with needle). Put 0.5mL tube into a 2mL tube, and centrifuge at 14k for 2 min.
2.2.5.4. Add 2X volume EB Buffer. Shake for overnight at 4C.
2.2.5.5. The next day, quick spin the tubes and place in 50C for 15 min. Spin 2 min at 14K rpm.
2.2.5.6. Transfer the supernatant to two Nanoseq columns (Nanosep MF Filter tube(VWR Cat.29300-642). Spin 2 min @ 14k.
2.2.5.7. Collect supernatant into new 1.5ml eppendorf tube.
2.2.5.8. Add EB Buffer to bring to a total volume of 300µl.
2.2.5.9. Add 1/10 volume(50µl) 3M NaOAC(ph 5.2) and vortex to mix.
2.2.5.10. Add 2µl glycogen(20mg/ml) and 500 ml cold 100% ETOH(2.5volumn), vortex.
2.2.5.11. Freeze at -80°C for 30 min,spin 14K for 15min at 4C. Remove supernatant.
2.2.5.12. Wash with 1mL 70% EtOH, vortex, spin in cold room for 5min @14K rpm, remove all traces of EtOH, and air dry 5 minutes.
2.2.5.13. Resuspend precipitated DNA in 30µl EB.

2.2.6. Library amplification

2.2.6.1. Perform PCR using sequencing primers: 98C 30 sec, 10 x (98C 10 sec, 65C 30 sec, 72C 30 sec), 72C 5 min, 4C hold.
2.2.6.2. Perform Minelute PCR purification according to manufacturer instructions. Elute twice in 15 µl EB.

2.2.7. Size selection

2.2.7.1. Run samples in 8% polyacrylamide gel.
2.2.7.2. Excise 280-320bp main fragment.
2.2.7.3. Put gel pieces in 0.5mL tube (with bottom of tube punctured 3 times with needle). Put 0.5mL tube into a 2mL tube, and centrifuge at 14k for 2 min.
2.2.7.4. Add 2X volume EB Buffer. Shake for overnight at 4C.
2.2.7.5. The next day, quick spin the tubes and place in 50C for 15 min. Spin 2 min at 14K rpm.
2.2.7.6. Transfer the supernatant to two Nanoseq columns (Nanosep MF Filter tube(VWR Cat.29300-642). Spin 2 min @ 14k.
2.2.7.7. Collect supernatant into new 1.5ml eppendorf tube.
2.2.7.8. Add EB Buffer to bring to a total volume of 300µl.
2.2.7.9. Add 1/10 volume(50µl) 3M NaOAC(ph 5.2) and vortex to mix.
2.2.7.10. Add 2µl glycogen(20mg/ml) and 500 ml cold 100% ETOH(2.5volumn), vortex.
2.2.7.11. Freeze at -80°C for 30 min,spin 14K for 15min at 4C. Remove supernatant.
2.2.7.12. Wash with 1mL 70% EtOH, vortex, spin in cold room for 5min @14K rpm, remove all traces of EtOH, and air dry 5 minutes.

2.2.7.13. Resuspend precipitated DNA in 10µl EB.

2.3. Strand-specific mRNA-seq library preparation

2.3.1. Homogenization

2.3.1.1. Homogenize cell pellet of ~2 x 10^6 cells in 1ml of TRIZOL reagent using a 20 gauge needle. The sample volume should not exceed 10% of the volume of TRIZOL Reagent used for the homogenization.

2.3.2. Phase separation

2.3.2.1. Add 0.2 ml of chloroform per 1ml of TRIZOL Reagent.

2.3.2.2. Cap sample tubes securely. Vortex samples vigorously by hand for 15 seconds and incubate them at room temperature for 2 to 3 minutes.

2.3.2.3. Centrifuge the samples at no more tha12,000g for 15 minutes at 4C.

2.3.2.4. Following centrifugation, the mixture separates into lower red, phenol-chloroform phase, an inter phase, and a colorless upper aqueous phase. RNA remains exclusively in the aqueous phase. Transfer upper aqueous phase carefully without disturbing the interphase into fresh tube. Measure the volume of the aqueous phase.

2.3.3. RNA precipitation

2.3.3.1. Precipitate the RNA from the aqueous phase by adding 0.5 ml of isopropyl alcohol per 1 ml of TRIZOL Reagent used for the initial homogenization.

2.3.3.2. Incubate samples at room temperature for 10 minutes.

2.3.3.3. Centrifuge at not more than 12,000g for 10 minutes at 4C.

2.3.3.4. The RNA precipitate, often invisible before centrifugation, forms a gel-like pellet on the side and bottom of the tube.

2.3.4. RNA wash

2.3.4.1. Remove the supernatant completely.

2.3.4.2. Wash the RNA pellet once with 75% ethanol, adding at least 1 ml of 75% ethanol per 1 ml of TRIZOL Reagent used for the initial homogenization.

2.3.4.3. Mix the samples by vortexing and centrifuge at no more than 7,500g for 5 minutes at 4C.

2.3.4.4. Repeat above washing procedure once. Remove all leftover ethanol.

2.3.5. Re-dissolving RNA

2.3.5.1. Air-dry or vacuum dry RNA pellet for 5-10 minutes. (Do not dry the RNA pellet by centrifuge under vacuum. It is important not to let the RNA pellet dry completely as this will greatly decrease its solubility. Partially dissolved RNA samples have an A260/A280 ratio < 1.6.)

2.3.5.2. Dissolve RNA in 50µl DEPC-treated water by passing solution a few times through a pipette tip.

2.3.5.3. Incubate 10 min 55C.

2.3.5.4. Store at -80°C.

2.3.6. polyA+ RNA isolation

2.3.6.1. Dilute 10µg of RNA in a total of 100µL of DEPC water.

2.3.6.2. Add 100µl of Binding Buffer (20 mM Tris-HCl, pH 7.5, 1.0 M LiCl, 2 mM EDTA) to the RNA sample.

2.3.6.3. Resuspend Dynabeads Oligo (dT)25 thoroughly in the vial to obtain a uniform brown suspension. Transfer 200µl (1 mg) of beads to a tube for each sample.
2.3.6.4. Place tubes on a magnet (Dynal MPC™) for 1-2min. The Dynabeads Oligo (dT)25 will migrate to the side of the tube nearest the magnet.
2.3.6.5. Remove the supernatant with a pipette while the tube remains on the magnet.
2.3.6.6. Remove the tube from the magnet and add 100µl Binding Buffer to resuspend the beads. Wash beads by using the magnet to pull the beads to the side of the tube, 3-4 times.
2.3.6.7. Again place the tube on the magnet for 1-2 min.
2.3.6.8. Remove the supernatant while the tube remains on the magnet.
2.3.6.9. Resuspend the beads in 100µl Binding Buffer.
2.3.6.10. Heat RNA samples to 65°C for 2 min to disrupt secondary structures. Immediately place on ice for 5 min.
2.3.6.11. Add the 200µl of total RNA to the 100µl washed beads.
2.3.6.12. Mix thoroughly and anneal by rotating continuously on a mixer for 5 min at room temperature.
2.3.6.13. Place the tube on the magnet for 1-2 min and carefully remove all the supernatant.
2.3.6.14. Remove the tube from the magnet and add 200µl Washing Buffer B (10 mM Tris-HCl, pH 7.5, 0.15 M LiCl, 1 mM EDTA). Mix by using the magnet to pull the beads to the side of the tube. Carefully remove all the supernatant.
2.3.6.15. Repeat the washing step as described in step 6.
2.3.6.16. After the last wash, place all samples on ice and proceed with elution one-by-one.
2.3.6.17. To elute mRNA from the beads, remove the washing solution, quick spin, place on magnet and remove the remaining of the washing buffer.
2.3.6.18. Take the magnet out and add the desired amount (12 µl) of cold 10 mM Tris-HCl directly on the beads.
2.3.6.19. Heat to 78°C for 2 min and place the tube immediately on the magnet. Quickly transfer the eluted mRNA to a new RNase-free tube. Keep on ice until you are done with all samples.

2.3.7. First strand synthesis from polyA$^+$ RNA

2.3.7.1. Prepare fresh 2.1X First Strand synthesis mix per reaction using the SuperScript® III First-Strand Synthesis System (Invitrogen 18080-051): 2 µl of 10X reverse transcription buffer, 1 µl of 10 mM dNTPs, 1 µl 100 mM $MgCL_2$, 2 µl 100 mM DTT, 1.6 µl random primers, 1 µl oligo dT primers, 1 µl Actinomycin D.
2.3.7.2. Mix with polyA+ RNA and place in PCR machine with the following program: 98C 1 min, 70C 5 min, down to 15C by 0.1C/sec.
2.3.7.3. Add 1 µl RNAseOUT (Invitrogran 10777-019) and 1 µl Superscript III enzymes.
2.3.7.4. In PCR machine, perform: 0.1 C/sec to 25C, 25C 10 min, 0.1C/sec to 42C, 42C 45 min, 0.1C/sec to 50C, 50C 15 min, 75C 15 min, 15C hold.

2.3.8. Second strand synthesis

2.3.8.1. Autoclave 5ml aliquots of Sephadex G50 beads (medium) in 1mM TrisHCl, pH 7.0.
2.3.8.2. With a clean scalpel, cut a hole into the tops of 1.7 ml tubes and place an unfiltered 200 µl pipet tip. Insert 4x4 mm^2 piece of GF/C into tip as a filter.
2.3.8.3. Mix well the G50 suspension and add 100 µl into filtered tip. Shake solution down.

2.3.8.4. Add another 150 µl of G50 suspension.
2.3.8.5. Spin down the empty column for 0.5 min at 2000 rpm.
2.3.8.6. Insert the column into a new collection tube (with opening on the cap).
2.3.8.7. Load the RNA sample and spin down 1 min at 2000 rpm
2.3.8.8. Measure volume of the eluate and transfer it to a 1.5 ml tube.
2.3.8.9. Put tubes on ice while you prepare SS-mix.
2.3.8.10. Prepare fresh 3.3x SS-mix (Second Strand synthesis mix) per reaction: 1 µl reverse transcription buffer 10X, 15 µl second strand synthesis buffer, 0.5 µl $MgCl_2$ 100 mM, 1 µl 100 mM DTT, 2 µl 10 mM dUNTPs, 0.5 µl E. coli DNA ligase, 2 µl DNA polymerase I, 0.5 µl RNase H.
2.3.8.11. Add to RNA sample, incubate for 2hrs at 16C.
2.3.8.12. Purify cDNA using QIA-quick-QG purification according to manufacturer instructions. Elute twice in 26 µl EB.

2.3.9. Ultrasonic digestion with Diagenode Bioruptor

2.3.9.1. Put samples in Biorupter TPX 1.5 ml tubes. Add EB buffer to bring total volume to 100 µl (recommended volume for sonication in 1.5 ml tubes).
2.3.9.2. Sonicate on high power for 10 min (30 s on, 30 s off) x 6 times.

2.3.10. End repair

2.3.10.1. Prepare fresh ER mix: 12 µl T4 DNA ligase buffer with 10 mM ATP, 3 µl 10 mM dNTP mix, 2 µl 3U/µl T4 DNA polymerase, 0.4 µl Klenow 5U/µl DNA polymerase, 2 µl T4 PNK.
2.3.10.2. Mix 19.4 µl of ER mix with 100 µl sheared cDNA.
2.3.10.3. Incubate 30min at 20C.
2.3.10.4. Purify cDNA using QIA-quick-QG purification according to manufacturer instructions. Elute twice in 18 µl EB.

2.3.11. A-tailing

2.3.11.1. Prepare A-tailing reaction mix per reaction: 5 µl 10X Klenow buffer, 10 µl 1 mM dATP, 0.5 µl water, 0.5 µl 5U/µl Klenow 3'-5' exo minus.
2.3.11.2. Mix 16 µl to cDNA.
2.3.11.3. Incubate 30 min at 37°C.
2.3.11.4. Purify cDNA using MinElute-QG purification according to manufacturer instructions. Elute once in 10 µl EB.

2.3.12. Adaptor ligation

2.3.12.1. Prepare the Adaptor mix (13pmol/µl) per reaction: 1.3 µl 100 pmol/µl sequencing adapter 1, 1.3 µl 100 pmol/µl sequencing adapter 2, 7.4 µl DEPC water.
2.3.12.2. Add 2 µl to cDNA, mix carefully and let stand for 5 min on ice.
2.3.12.3. Add 15 µl of DNA ligase buffer 2x (NEB Quick Ligation Kit M2200L).
2.3.12.4. Add 3 µl of DNA ligase per sample.
2.3.12.5. Incubate 30 min at 20C.
2.3.12.6. Purify cDNA using QIAquick-QG purification according to manufacturer instructions. Elute once in 12 µl EB.

2.3.13. Size selection

2.3.13.1. Run samples in 2% high-resolution agarose gel.

2.3.13.2. Excise 150-200bp main fragment and perform QIAquick-QG gel purification according to manufacturer instructions.

2.3.13.3. Elute twice in 26 µl EB.

2.3.14. UDGase treatment

2.3.14.1. Add following reagents to the non-amplified library and mix carefully: 5.7 µl 10X UDGase buffer, 1 µl 1U/µl UDGase (AmpErase UNG from AB SybrGreen kit).

2.3.14.2. Incubate 30min at 37C.

2.3.15. Real-Time PCR check of the non-amplified sequencing library

2.3.15.1. Perform RT PCR on each sample to determine the number of cycles necessary to reach the middle of the logarithmic phase.

2.3.15.2. Go to: http://seq.molbiol.ru/lib_ampl.html to calculate the number of PCR cycles to run per sample.

2.3.16. Library amplification and size selection

2.3.16.1. Perform PCR using sequencing primers, using the number of cycles calculated above for each sample: 98C 30 sec, cycles x (98C 15 sec, 65C 30 sec, 72C 30 sec), 4C hold.

2.3.16.2. Run samples in 2% high-resolution agarose gel.

2.3.16.3. Excise 150-200bp main fragment and perform QIAquick-QG gel purification according to manufacturer instructions.

2.3.16.4. Elute twice in 16 µl EB.

2.4. methylC-seq library preparation

2.4.1. Extract DNA using Qiagen DNeasy Kit (69504)

2.4.1.1. Thaw frozen pellet containing a maximum of 5 million cells until pellet can be dislodged by gently flicking the tube.

2.4.1.2. Resuspend pellet in 200 µl PBS.

2.4.1.3. Add 20 µl proteinase K.

2.4.1.4. Add 4 µl RNase A (100 mg/ml). Mix by vortexing, and incubate for 2 min at room temperature.

2.4.1.5. Add 200 µl Buffer AL (w/o EtOH). Mix thoroughly, vortex, and incubate 56 C for 10 min.

2.4.1.6. Add 200 µl 100% EtOH, mix thoroughly, vortex.

2.4.1.7. Pipet mixture into DNeasy Mini spin column placed in 2 ml collection tube. Centrifuge 8000 rpm for 1 min. Discard flow-through and collection tube.

2.4.1.8. Place column in new 2 ml collection tube, add 500 µl Buffer AW1, centrifuge 8000 rpm for 1 min. Discard flow-through and collection tube.

2.4.1.9. Place column in new 2 ml collection tube, add 500 µl Buffer AW2, centrifuge 14000 rpm for 3 min to dry. Discard flow-through and collection tube.

2.4.1.10. Place column in clean tube. Perform twice: add 100 µl Buffer AE to membrane. Incubate 1 min. Centrifuge 8000 rpm for 1 min.

2.4.2. Fragmentation of gDNA

2.4.2.1. Starting with 5.5 µg gDNA, spike in lambda DNA at 0.5%. Add EB until volume is 300 µl for sonication.

2.4.2.2. Sonicate on high power for 10 min (30 s on, 30 s off) x 5 times.

2.4.2.3. Perform MinElute cleanup according to manufacturer instructions. Final elution volume is 36 µl.

2.4.3. Perform End Repair using the End-It Kit (Epicentre):

2.4.3.1. To each reaction, add 5 µl 10X End-It Buffer, 5 µl 10 mM dNTP mix, 5 µl 10 mM ATP, 1 µl End-It Enzyme Mix.
2.4.3.2. Mix tube gently, spin briefly, incubate at room temperature for 45 min.
2.4.3.3. Perform MinElute PCR cleanup, using 2 columns per sample. Elute each with 16 µl EB.

2.4.4. A-tailing

2.4.4.1. To each reaction, add 5 µl Klenow buffer, 10 µl 1 mM dATP mix, 3 µl Klenow (3' to 5' exo minus).
2.4.4.2. Mix tube gently, spin briefly, put in thermal block for 30 min at 37C.
2.4.4.3. Perform MinElute PCR cleanup, using 2 columns per sample. Elute sequentially with 20 µl EB, using first eluate in second column elution.
2.4.4.4. SpeedVac to concentrate sample to 4 µl.

2.4.5. Adapter ligation with methylated adapters

2.4.5.1. To each reaction, add 10 µl 2X DNA ligase buffer, 5 µl 15 uM methylated adapters, 1 µl T4 DNA ligase.
2.4.5.2. Mix tube gently, spin briefly, incubate at room temperature for 15 min.
2.4.5.3. Perform MinElute PCR cleanup, using 2 columns per sample.
2.4.5.4. Elute sequentially with 20 µl EB, using first eluate in second column elution.

2.4.6. Size selection

2.4.6.1. Run samples in 2% high-resolution agarose gel.
2.4.6.2. Excise 150-300bp main fragment and perform QIAquick-QG gel purification according to manufacturer instructions.
2.4.6.3. Elute twice in 16 µl EB.

2.4.7. Bisulfite Conversion of gDNA

2.4.7.1. Convert 450 ng of adapter-ligated DNA as per MethylCode Kit instructions (Invitrogen MECOV-50).
2.4.7.2. Elute in 32 µl elution buffer.

2.4.8. Low Amplification (5 cycle PCR) of Converted gDNA

2.4.8.1. To each reaction, add 5 µl 10X Pfu Cx buffer (Agilent 600410), sequencing primers, 0.75 µl 20 mM dNTPs, 28.25 µl water, 1 µl Pfu Turbo Cx Polymerase.
2.4.8.2. Perform PCR: 95C 2 min, 98C 30 sec, 5 x (98C 15 sec, 60C 30 sec, 72C 4 min), 72C 10 min, 4C hold.

2.4.9. Size selection

2.4.9.1. Run samples in 2% high-resolution agarose gel.
2.4.9.2. Excise 150-300bp main fragment and perform QIAquick-QG gel purification according to manufacturer instructions.
2.4.9.3. Elute twice in 10 µl EB.

2.5. ChIP-seq library preparation

2.5.1. Cross-linking and isolation of chromatin with adherent cells

2.5.1.1. Grow cells until near confluent.
2.5.1.2. Remove media.
2.5.1.3. Add 8 ml of fresh media to each plate.
2.5.1.4. Add 800 µl cross-linking buffer (0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM Hepes pH 8, 11% fresh formaldehyde, water) and incubate 15 min at room temperature.
2.5.1.5. Add 0.055 volume of culture (484 µl) of 2.5 M glycine to quench cross-linking. Mix well by shaking gently but several times (swirl once after initial swirling of plates), and incubate 5 min at room temperature.
2.5.1.6. Remove media and dispose properly.
2.5.1.7. Add 10-12 ml PBS, remove by aspiration. Repeat.
2.5.1.8. Add 5 ml PBS. Scrape cells. Rinse plate with a extra PBS, collect cells.
2.5.1.9. Centrifuge 3K rpm for 15 min. Remove supernatant.

2.5.2. Chromatin isolation

2.5.2.1. Add 5 mL buffer L1 (50 mM Hepes KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100, water, 1 protease inhibitor tablet (1:50) (Roche Complete 1697498)) to cells, let sit 5-10 minutes. Resuspend cells in 25 mL (30 mL total) of buffer L1. Rock at 4°C for 10 minutes to completely resuspend.
2.5.2.2. Spin at 3000 rpm in 10 minutes at 4°C. Remove supernatant.
2.5.2.3. Resuspend each tube of cells in 24mL of buffer L2 (200 mM NaCl, 1 mM EDTA pH 8, 0.5 mM EGTA, 10 mM Tris, pH 8, water, 1 protease inhibitor tablet).
2.5.2.4. Rock at room temperature for 10 minutes.
2.5.2.5. Pellet nuclei in a tabletop centrifuge by spinning at 3K rpm, 10 minutes at 4°C. Remove supernatant.
2.5.2.6. Resuspend nuclei in 5mL buffer L3 (1 mM EDTA pH 8, 0.5 mM EGTA pH 8, 10 mM Tris pH 8, water, 1 protease inhibitor tablet). Transfer to a 15mL conical tube.
2.5.2.7. Place the 15mL conical tube on ice. Sonicate the cell suspension with a microtip attached to Branson 450 sonifier, setting at output ~15-20, constant power. Sonicate 30 times for 30 seconds, allowing the suspension to cool 1 minute on ice between pulses. Use sandpaper to smooth tip every 10 cycles. Let the tip cool down for 2 minutes every 5 cycles.
2.5.2.8. Spin out debris at 4K rpm, 15 minutes.
2.5.2.9. Adjust glycerol concentration to 10% with 80% stock.
2.5.2.10. Adjust the final DNA concentrations to 500 µg/mL by adding L3 buffer.
2.5.2.11. Chromatin samples stored at -80°C.

2.5.3. Chromatin immunoprecipitation – Day 1

2.5.3.1. Make 5mg/mL BSA (Sigma BSA powder, cat. A-3350) (0.125g BSA in 25mL 1X PBS) immediately before use, put on ice.
2.5.3.2. Add 50µl Dynal beads (Invitrogen 100-07D) to 1.5ml eppendorf tube. Place on magnetic strip. Remove supernatant with a pipette/vacuum. Resuspend beads in 1mL PBS containing 5mg/ml BSA.

2.5.3.3. Wash 3 times with PBS/BSA. Resuspend beads in 1000µl PBS/BSA minus volume of antibody to be added.
2.5.3.4. With tube against the magnet, add 5 µg antibody for a total volume of 1000µl. Incubate for **at least 2 hrs** on a rotating platform at 4ºC.
2.5.3.5. Immediately before use: make 10% sodium deoxycholate (DOC) (0.1g DOC in 1mL dH2O) immediately before use; make 5mg/mL BSA (0.125g BSA in 25mL PBS) immediately before use. make 50X Complete (1 tablet in 1.0mL water).
2.5.3.6. Place bead/antibody solution on magnets, remove supernatant with a pipette/aspirate, and resuspend in 1mL PBS/BSA (on ice).
2.5.3.7. Wash 3 times with PBS/BSA.
2.5.3.8. Resuspend beads in 100µl 1X PBS/BSA.
2.5.3.9. Set up IP reactions with chromatin extract: 1% Triton-X, 0.10% DOC, 1X Complete, bead/PBS/BSA mixture, 500 µg chromatin, and TE.
2.5.3.10. Incubate overnight on a rotating platform.

2.5.4. Chromatin immunoprecipitation – Day 2

2.5.4.1. Make 10% DOC immediately before use.
2.5.4.2. Make RIPA buffer immediately before use. Add the stock solutions in the order listed: water, 50 mM Hepes pH 8, 1% NP-40, 0.7% DOC, 0.5 M LiCl, 1 Complete tablet.
2.5.4.3. Wash beads with RIPA buffer with tubes on magnet. Keep tubes on ice. Wash 5 times with 1mL RIPA buffer. Remove buffer by aspiration.
2.5.4.4. Wash once with 1mL TE.
2.5.4.5. After removing the TE by aspiration, quickly spin the tubes or place on magnet and remove remaining liquid with a pipet.
2.5.4.6. To elute beads, add elution buffer (10 mM Tris pH 8, 1 mM EDTA, 1% SDS, water)
2.5.4.7. Add 150µl of elution buffer, vortex briefly to resuspend the beads and incubate at 65ºC for 10 minutes.
2.5.4.8. Spin for 30 seconds at maximum speed or place on magnet and transfer liquid to a new tube.
2.5.4.9. Reverse crosslink at 65ºC overnight.
2.5.4.10. Also reverse crosslink 30µl input chromatin in 120µl elution buffer.

2.5.5. Chromatin immunoprecipitation – Day 3

2.5.5.1. Let samples sit until they reach room temperature.
2.5.5.2. Add 250µl TE to each sample
2.5.5.3. Add 8µl of 10mg/ml RNAse A (final conc. = 0.2mg/ml). Incubate at 37C for 1hr.
2.5.5.4. Add 8µl of 20mg/ml Proteinase K (final conc. = 0.4mg/ml). Incubate for 1hr at 55C.
2.5.5.5. Prepare Phase Lock tube (5Prime 2302800) per IP, spin at 14K rpm for 1min.
2.5.5.6. Add 400µl Phenol:Chloroform: Isoamyl 25:24:1 alcohol to each tube.
2.5.5.7. Add sample to Phase Lock tube and shake vigorously, do not vortex.
2.5.5.8. Spin 4 min at 14K rpm. **Note**: if aqueous phase is cloudy, extract again.
2.5.5.9. Transfer aqueous layer to new tube.
2.5.5.10. Add 16µl of 5M NaCl (final conc. = 200mM), plus 1.5µl of 20mg/ml (30µg of glycogen). Vortex.
2.5.5.11. Add 920µl cold 100% EtOH, vortex briefly.
2.5.5.12. Incubate at -80ºC for 30min. Spin at 14K rpm for 15 minutes at 4C.
2.5.5.13. Wash pellet with 1mL cold 70% EtOH, vortex, spin 5min. at 4C at 14K rpm.

2.5.5.14. Resuspend thoroughly in 34 µl of 10mM Tris.

2.5.6. Perform End Repair using the End-It Kit (Epicentre):

2.5.6.1. To each reaction, add 5 µl 10X End-It Buffer, 5 µl 2.5 mM dNTP mix, 5 µl 10 mM ATP, 1 µl End-It Enzyme Mix.
2.5.6.2. Mix tube gently, spin briefly, incubate at room temperature for 45 min.
2.5.6.3. Perform MinElute PCR cleanup according to manufacturer instructions. Elute twice with 16 µl EB.

2.5.7. A-tailing

2.5.7.1. To each reaction, add 5 µl NEB Buffer 2, 10 µl 1 mM dATP mix, 3 µl Klenow (3' to 5' exo minus).
2.5.7.2. Mix tube gently, spin briefly, put in thermal block for 30 min at 37C.
2.5.7.3. Perform MinElute PCR cleanup. Elute twice in 10 µl EB.
2.5.7.4. SpeedVac to concentrate sample to 4 µl.

2.5.8. Adapter ligation

2.5.8.1. To each reaction, add 5 µl 2X DNA ligase buffer, 0.5 µl 1:10 sequencing adapters, 0.5 µl DNA ligase (Quick Ligation Kit (NEB Cat.M2200L)).
2.5.8.2. Mix tube gently, spin briefly, incubate at room temperature for 15 min.
2.5.8.3. Perform MinElute PCR cleanup. Elute twice in 20 µl EB.

2.5.9. Size selection

2.5.9.1. Run samples in 8% polyacrylamide gel.
2.5.9.2. Excise 200-400bp main fragment.
2.5.9.3. Put gel pieces in 0.5mL tube (with bottom of tube punctured 3 times with needle). Put 0.5mL tube into a 2mL tube, and centrifuge at 14k for 2 min.
2.5.9.4. Add 2X volume EB Buffer. Shake for overnight at 4C.
2.5.9.5. The next day, quick spin the tubes and place in 50C for 15 min. Spin 2 min at 14K rpm.
2.5.9.6. Transfer the supernatant to two Nanoseq columns (Nanosep MF Filter tube(VWR Cat.29300-642). Spin 2 min @ 14k.
2.5.9.7. Collect supernatant into new 1.5ml eppendorf tube.
2.5.9.8. Add EB Buffer to bring to a total volume of 500µl.
2.5.9.9. Add 1/10 volume(50µl) 3M NaOAC(ph 5.2) and vortex to mix.
2.5.9.10. Add 4µl glycogen(20mg/ml) and 1ml cold 100% ETOH(2.5volumn), vortex.
2.5.9.11. Freeze at -80°C for 30 min,spin 14K for 15min at 4C. Remove supernatant.
2.5.9.12. Wash with 1mL 70% EtOH, vortex, spin in cold room for 5min @14K rpm, remove all traces of EtOH, and air dry 5 minutes.
2.5.9.13. Resuspend precipitated DNA in 20µl EB.

2.5.10. Library amplification

2.5.10.1. Perform PCR using sequencing primers: 98C 30 sec, 18 x (98C 10 sec, 65C 30 sec, 72C 30 sec), 72C 5 min, 4C hold.
2.5.10.2. Perform Minelute PCR purification according to manufacturer instructions. Elute twice in 20 µl EB.

2.5.11. Size selection

2.5.11.1. Run samples in 8% polyacrylamide gel.

2.5.11.2. Excise 200-400bp main fragment.

2.5.11.3. Put gel pieces in 0.5mL tube (with bottom of tube punctured 3 times with needle). Put 0.5mL tube into a 2mL tube, and centrifuge at 14k for 2 min.

2.5.11.4. Add 2X volume EB Buffer. Shake for overnight at 4C.

2.5.11.5. The next day, quick spin the tubes and place in 50C for 15 min. Spin 2 min at 14K rpm.

2.5.11.6. Transfer the supernatant to two Nanoseq columns (Nanosep MF Filter tube(VWR Cat.29300-642). Spin 2 min @ 14k.

2.5.11.7. Collect supernatant into new 1.5ml eppendorf tube.

2.5.11.8. Add EB Buffer to bring to a total volume of 500µl.

2.5.11.9. Add 1/10 volume(50µl) 3M NaOAC(ph 5.2) and vortex to mix.

2.5.11.10. Add 4µl glycogen(20mg/ml) and 1ml cold 100% ETOH(2.5volumn), vortex.

2.5.11.11. Freeze at -80°C for 30 min,spin 14K for 15min at 4C. Remove supernatant.

2.5.11.12. Wash with 1mL 70% EtOH, vortex, spin in cold room for 5min @14K rpm, remove all traces of EtOH, and air dry 5 minutes.

2.5.11.13. Resuspend precipitated DNA in 20µl EB.

2.6. ChIP-methylC-seq library preparation

ChIP-methylC-seq was performed by combining the ChIP protocol with methylC-seq library construction. Specifically, ChIP was performed as in steps 2.5.1 through 2.5.5. Following spike-in of sonicated lambda DNA at a final concentration of 0.5% (assuming ChIP yield of 10 ng), library construction was performed using the methylC-seq protocol (steps 2.4.3 through 2.4.9), with the following changes: 1) methylated adapters were diluted 1:3, and 2) library amplification used 20 PCR cycles with 2 minutes of extension at 72C.

2.7. Sequencing

We performed paired-end sequencing of the HCC1954-blood and HCC1954 genomes with read sizes of 36-bp, 75-bp, or 101-bp on each end (Figure S2).

Single-end methylC-seq libraries were sequenced at 100-bp, while single-end RNA-seq and ChIP-seq libraries were sequenced at 36-bp. With the exception of H3K4me3 in HCC1954 and genome sequencing of HCC1954/HCC1954-blood, each experiment is represented by at least two biological replicates.

2.8. Read mapping

We mapped paired-end genome sequencing reads using the Bowtie program(Langmead et al. 2009) in 2 passes: the first attempting to map in paired-end mode, and the second mapping unmapped reads in single-end mode. We kept only those reads that mapped uniquely to the human genome (NCBI build hg18) with at most 3 mismatches. Specifically, the parameters used in paired-end mode were: "--solexa1.3-quals -I 0 -X 400 -y -v 3 -m 1 --best –strata". Similarly, the single-end parameters were: “"--solexa1.3-quals -v 3 -m 1 --best –strata”.

We mapped methylC-seq reads using the pipeline developed in Lister et al(Lister et al. 2009). Specifically: reads were trimmed until encountering a base with PHRED score <= 2 and bases near the 3' end were trimmed if the adapter sequence was included. For the purpose of mapping, all cytosines were computationally converted to thymines, to be reverted back to their original state after mapping. The Bowtie program(Langmead et al. 2009) was then used to align these computationally converted reads to computationally converted copies of hg18. Non-uniquely mapping reads were discarded.

We mapped ChIP-seq reads using the Bowtie program, keeping only those reads that mapped uniquely to hg18 with at most 3 mismatches. Specifically, the parameters used were: "--solexa1.3-quals -v 3 -m 1 --best --strata -S ". We used the TopHat program(Trapnell et al. 2009) to map RNA-seq reads, using default parameters.

With the exception of RNA-seq experiments, PCR duplicates for each library were removed with the Picard program(Picard 2011). All biological replicates were combined and compressed/indexed by the SAMtools suite into BAM format(Li et al. 2009).

Read coverage of uniquely mapping, monoclonal reads for genome sequencing, methylC-seq, and ChIP-seq can be found in Table S1 and S2. The number of RNA-seq reads mapped by TopHat is summarized in Table S3.

2.9. Quality control of methylC-seq

For methylC-seq, attaining nearly complete bisulfite conversion is essential to accurately call methylated cytosines. As in Lister et al(Lister et al. 2009), we spiked in the unmethylated lambda genome at 0.5% of the total DNA to measure the conversion rate. For HCC1954 samples, we achieved conversion rates between 99.28% and 99.54% for all 3 replicates, with a weighted average of 99.45%. For HMEC, the conversion rates ranged from 99.55% to 99.56% for an average of 99.56%.

For ChIP-methylC-seq experiments, bisulfite conversion rates for spiked-in unmethylated lambda controls were between 98.0% and 98.5%.

2.10. Integrative analysis with breast tumor panel expression

Microarray gene expression data from a panel of breast tumors and normal breast samples generated by Perou and colleagues(Hennessy et al. 2009; Herschkowitz et al. 2008; Herschkowitz et al. 2007; Hoadley et al. 2007; Hu et al. 2009; Mullins et al. 2007; Oh et al. 2006; Parker et al. 2009; Perreard et al. 2006; Prat et al. 2010; Weigelt et al. 2005) were downloaded from the Gene Expression Omnibus. For simplicity in merging experiments, we focused only on platform GPL1390, corresponding to "Agilent Human 1A Oligo UNC custom Microarrays"

spanning 22576 genes. After combining the samples from all experiments, redundant samples were removed and filtered for those samples having patient survival information. Using gene identifiers from the platform, we mapped 17268 genes on the microarray to genes from our RNA-seq experiments, spanning 13262 unique gene identifiers. Based on sample annotations, we split the samples 3 ways: normal-like, HER2+ cancer, and HER2- cancer.

To assess the genes that were differentially expressed between each pair of categories, we used the Wilcoxon rank sum test ($p \leq 0.01$) in conjunction with a cutoff of at least 50% difference in median RNA expression.

2.11. Assessing allele-specific epigenetic modifications

To assess allele-specific ChIP-seq enrichment at a given haplotype block, we first counted the number of ChIP reads landing in each of the two haplotypes, here labeled as hap1 and hap2. To ensure we only counted ChIP reads that unambiguously mapped to one haplotype and not the other, we only counted those reads where the number of the number of SNPs belonging to one haplotype was at least 2 times that for the other haplotype. As the background, we also counted the number of occurrences of hap1 and hap2 from genome sequencing. Finally, to assess allele-specificity of ChIP-seq, we used Fisher's Exact Test on these four values (#ChIP reads in hap1, #ChIP reads in hap2, #genome sequencing reads in hap1, #genome sequencing reads in hap2), with a cutoff of $p = 0.05$.

Similarly, to assess allele-specific DNA methylation at a given haplotype block, we counted the number of methylated and unmethylated cytosines in CG context (given the reference genome sequence hg18 and with a Phred score ≥ 20) from sequenced methylC-seq reads landing into each haplotype block. Again, to ensure we only counted ChIP reads that unambiguously mapped to one haplotype and not the other, we only counted those reads where the number of the number of SNPs belonging to one haplotype was at least 2 times that for the other haplotype. Since we used bisulfite converted reads in this analysis, C/T SNPs cannot be unambiguously called as belonging to one haplotype or another, and so these SNPs are discarded from this analysis. In contrast, bisulfite SNPs spanning C/X SNPs, where X is either A or G, can either be X, C (methylated), or T (unmethylated). In this case, a bisulfite read containing X at the SNP can be unambiguously assigned to one haplotype, whereas a read containing C or T at the SNP can be assigned to the other haplotype.

For each haplotype, we summed the number of methylated and unmethylated cytosines in CG context on both strands. Finally, to assess allele-specificity of methylC-seq, we used Fisher's Exact Test on these four values (#methylated CGs on hap1, #unmethylated CGs on hap1, #methylated CGs on hap2, # unmethylated CGs on hap2), with a cutoff of $p = 0.05$.

3. Supplementary Tables

Table S1: Coverage of uniquely mapping monoclonal reads (relative to hg18)

	HCC1954				HMEC			HCC1954-blood	
	rep1	rep2	rep3	combined	rep1	rep2	combined	rep1	combined
methylC-Seq	9.45	10.24	7.33	27.02	9.87	10.21	20.08		
genome-Seq	27.64			27.64				25.45	25.45

Table S2: Number of uniquely mapping monoclonal reads (relative to hg18, in millions)

		HCC1954			HMEC		
		rep1	rep2	combined	rep1	rep2	combined
ChIP-Seq	H3K4me1	11.26	6.85	18.11	16.84	15.19	32.03
	H3K4me3	10.47		10.47	14.79	11.22	26.01
	H3K9me3	76.48	5.33	81.81			
	H3K27ac	11.27	7.80	19.07	11.87	8.12	19.99
	H3K27me3	87.87	3.41	91.28	7.57	8.29	15.86
	H3K36me3	92.25	14.51	106.76	8.42	10.24	18.65
	input	11.39	13.18	24.57	4.70	12.33	17.03

Table S3: Number of TopHat mapped reads (relative to hg18, in millions)

	HCC1954			HMEC		
	rep1	rep2	combined	rep1	rep2	combined
RNA-Seq	21.96	22.98	44.78	23.26	25.31	48.23

4. References

Garbe, J.C., S. Bhattacharya, B. Merchant, E. Bassett, K. Swisshelm, H.S. Feiler, A.J. Wyrobek, and M.R. Stampfer. 2009. Molecular distinctions between stasis and telomere attrition senescence barriers shown by long-term culture of normal human mammary epithelial cells. *Cancer Res* **69:** 7557-7568.

Hennessy, B.T., A.M. Gonzalez-Angulo, K. Stemke-Hale, M.Z. Gilcrease, S. Krishnamurthy, J.S. Lee, J. Fridlyand, A. Sahin, R. Agarwal, C. Joy et al. 2009. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res* **69:** 4116-4124.

Herschkowitz, J.I., X. He, C. Fan, and C.M. Perou. 2008. The functional loss of the retinoblastoma tumour suppressor is a common event in basal-like and luminal B breast carcinomas. *Breast Cancer Res* **10:** R75.

Herschkowitz, J.I., K. Simin, V.J. Weigman, I. Mikaelian, J. Usary, Z. Hu, K.E. Rasmussen, L.P. Jones, S. Assefnia, S. Chandrasekharan et al. 2007. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8:** R76.

Hoadley, K.A., V.J. Weigman, C. Fan, L.R. Sawyer, X. He, M.A. Troester, C.I. Sartor, T. Rieger-House, P.S. Bernard, L.A. Carey et al. 2007. EGFR associated expression profiles vary with breast tumor subtype. *BMC Genomics* **8:** 258.

Hu, Z., C. Fan, C. Livasy, X. He, D.S. Oh, M.G. Ewend, L.A. Carey, S. Subramanian, R. West, F. Ikpatt et al. 2009. A compact VEGF signature associated with distant metastases and poor outcomes. *BMC Med* **7:** 9.

Langmead, B., C. Trapnell, M. Pop, and S.L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078-2079.

Lister, R., M. Pelizzola, R.H. Dowen, R.D. Hawkins, G. Hon, J. Tonti-Filippini, J.R. Nery, L. Lee, Z. Ye, Q.M. Ngo et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462:** 315-322.

Mullins, M., L. Perreard, J.F. Quackenbush, N. Gauthier, S. Bayer, M. Ellis, J. Parker, C.M. Perou, A. Szabo, and P.S. Bernard. 2007. Agreement in breast cancer classification between microarray and quantitative reverse transcription PCR from fresh-frozen and formalin-fixed, paraffin-embedded tissues. *Clin Chem* **53:** 1273-1279.

Novak, P., T. Jensen, M.M. Oshiro, G.S. Watts, C.J. Kim, and B.W. Futscher. 2008. Agglomerative epigenetic aberrations are a common event in human breast cancer. *Cancer Res* **68:** 8616-8625.

Oh, D.S., M.A. Troester, J. Usary, Z. Hu, X. He, C. Fan, J. Wu, L.A. Carey, and C.M. Perou. 2006. Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol* **24:** 1656-1664.

Parker, J.S., M. Mullins, M.C. Cheang, S. Leung, D. Voduc, T. Vickery, S. Davies, C. Fauron, X. He, Z. Hu et al. 2009. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* **27:** 1160-1167.

Perreard, L., C. Fan, J.F. Quackenbush, M. Mullins, N.P. Gauthier, E. Nelson, M. Mone, H. Hansen, S.S. Buys, K. Rasmussen et al. 2006. Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. *Breast Cancer Res* **8:** R23.

Picard. 2011. http://picard.sourceforge.net.

Prat, A., J.S. Parker, O. Karginova, C. Fan, C. Livasy, J.I. Herschkowitz, X. He, and C.M. Perou. 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12:** R68.

Romanov, S.R., B.K. Kozakiewicz, C.R. Holst, M.R. Stampfer, L.M. Haupt, and T.D. Tlsty. 2001. Normal human mammary epithelial cells spontaneously escape senescence and acquire genomic changes. *Nature* **409:** 633-637.

Trapnell, C., L. Pachter, and S.L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25:** 1105-1111.

Weigelt, B., Z. Hu, X. He, C. Livasy, L.A. Carey, M.G. Ewend, A.M. Glas, C.M. Perou, and L.J. Van't Veer. 2005. Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res* **65:** 9155-9158.