

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to [508 standards](#) due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehp508@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

Supplemental Material

Identification of Smoking-Associated Differentially Methylated Regions Using Reduced Representation Bisulfite Sequencing and Cell type-Specific Enhancer Activation and Gene Expression

Ma Wan, Brian D. Bennett, Gary S. Pittman, Michelle R. Campbell, Lindsay M. Reynolds, Devin K. Porter, Christopher L. Crowl, Xuting Wang, Dan Su, Neal A. Englert, Isabel J. Thompson, Yongmei Liu, and Douglas A. Bell

Table of Contents

Figure S1. (A–J) Differential methylation tracks (Δ Methylation) and SM-DMRs displayed across sample groups and across CRU-RRBS-Cell types with CD14+ histone modifications and transcription factor clusters (TFs) for (A) *HOXA5* region, (B) *AHRR*, (C) *AHRR-C5orf55-EXOC-AS*, (D) *ALPPL2*, (E) *PPP1R15A (GADD34)*, (F) *F2RL3, F2RL3-CPAMD8*, (G) *SLC24A3*, (H) *RDX*, (I) *SASH1*, and (J) *LRP5*. Boxed areas in each panel show the region of interest. UCSC genome browser views display RRBS results and additional Roadmap Project histone modifications and ENCODE project ChIPseq data.

Figure S2. The correspondence between RRBS and pyrosequencing measurements of DNA methylation at SM-DMRs for overlapping CRU subjects. (A–G) Plots show RRBS and pyrosequencing determined measurements of DNA methylation at SM-DMRs. Individual regions (A–F) and overall (G) with their corresponding Pearson correlations are denoted in different color and plotted separately as indicated. *LRP5* is not shown.

Figure S3. Pyrosequencing measurements on *HOXA5* and *RDX* SM-DMRs. (A, B) Upper panels show genome browser tracks for (top) Roadmap whole genome bisulfite sequencing (WGBS) in monocytes shows location of all CpGs in region, reference gene, differentially methylated CpGs by RRBS (red downward bar = loss of methylation, blue upward bar, gain of methylation), and RRBS DMR (black bar).

Figure S4. The correlation between AHRR SM-DMR, mRNA, eRNA, and cotinine levels among a subset of eight smokers and eight nonsmokers from the CRU-RRBS, CRU-mRNA and CRU-eRNA subgroups. (A) The correlation between AHRR mRNA, and eRNA expression ($R^2=0.98$, $p < 0.0001$). (B) eRNA was significantly correlated with serum cotinine levels ($R^2=0.65$, $p=0.0007$). (C) The correlation between average AHRR SM-DMR methylation from RRBS and mRNA expression ($R^2=0.45$, $p=0.0042$). Squares represent nonsmokers, while circles indicate smokers.

Table S1. Pyrosequencing and BSAS assay primers.

Table S2. mRNA Real-time Polymerase Chain Reaction Assays.

Table S3. Enhancer RNA RT-PCR Primers.

References

Additional File- Excel Document

Figure S1

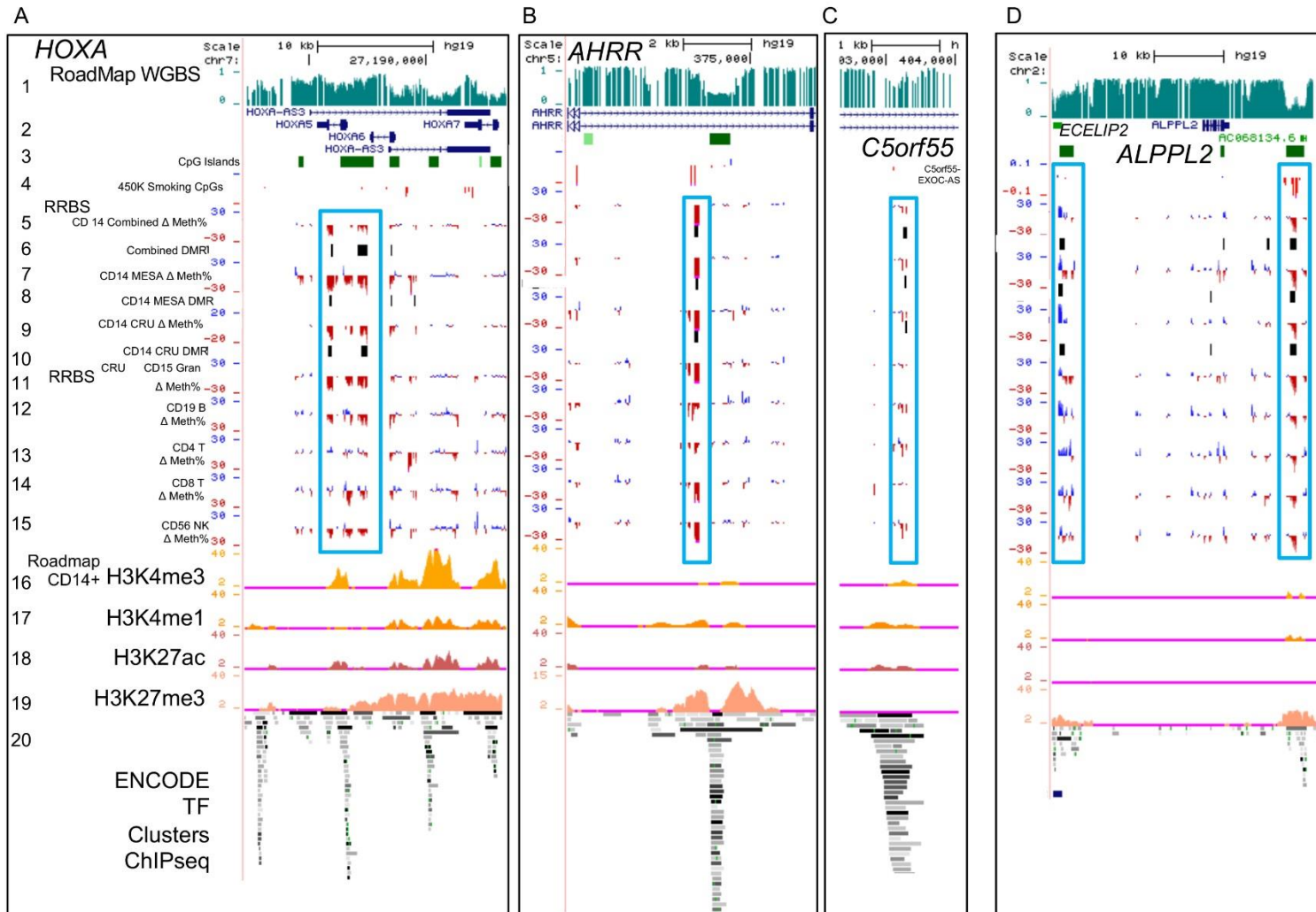
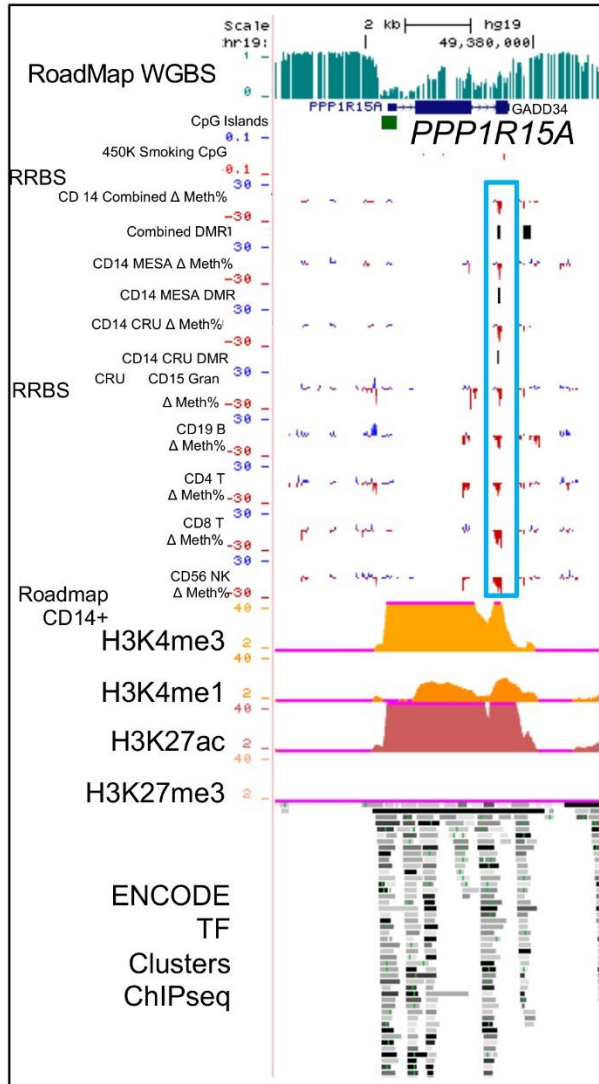
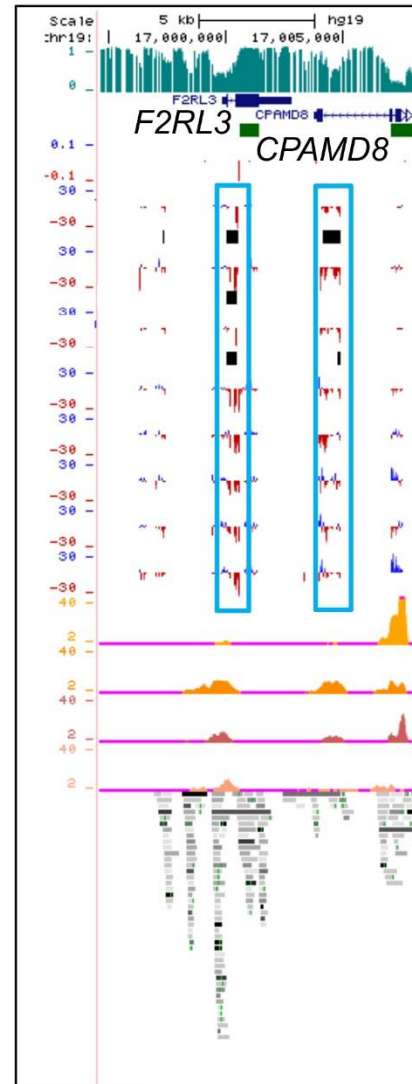


Figure S1

E



F



G

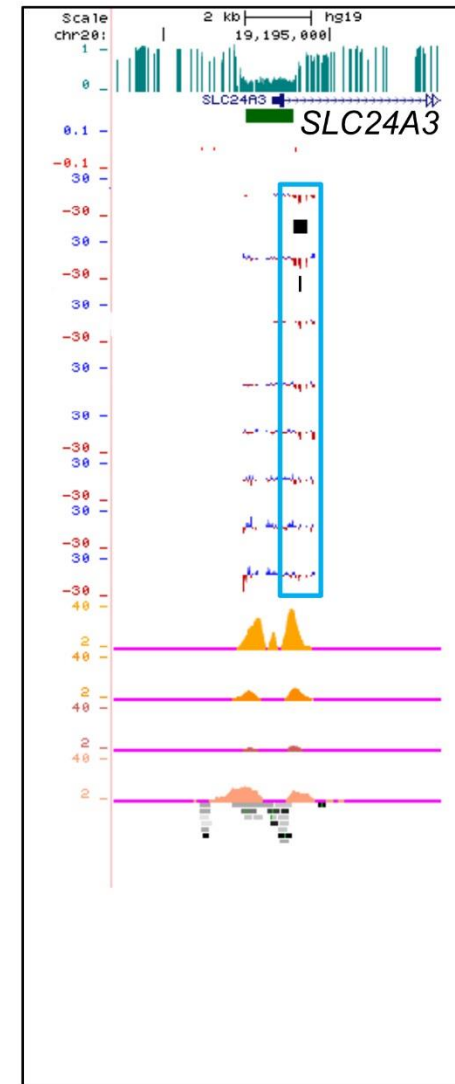
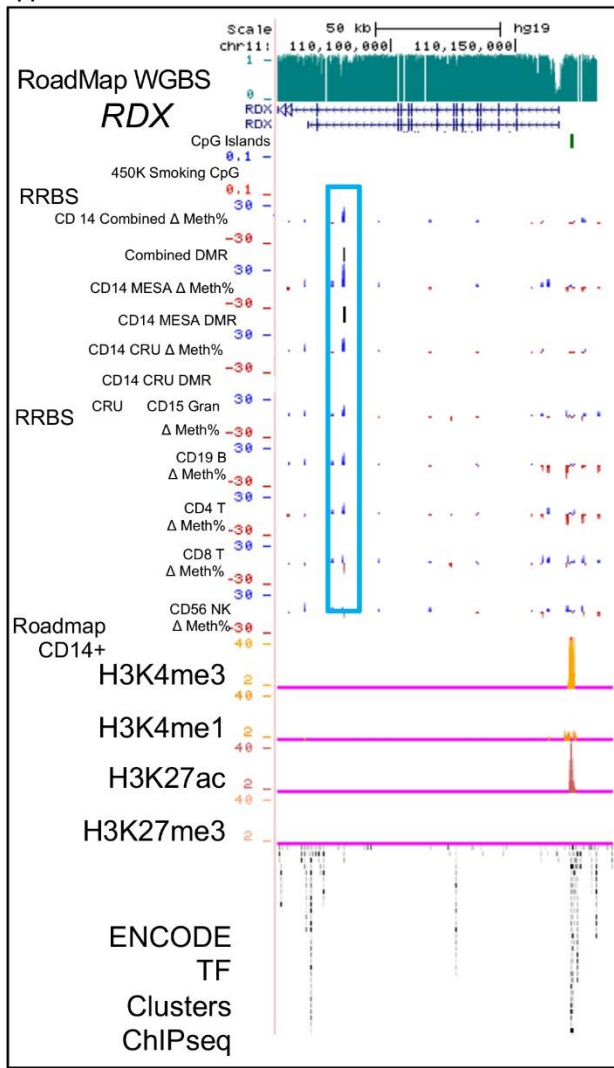
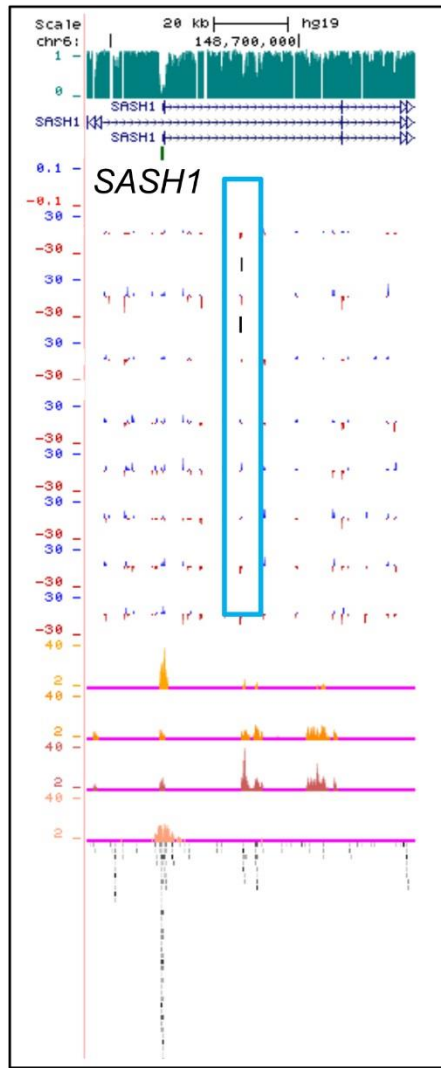


Figure S1

H



I



J

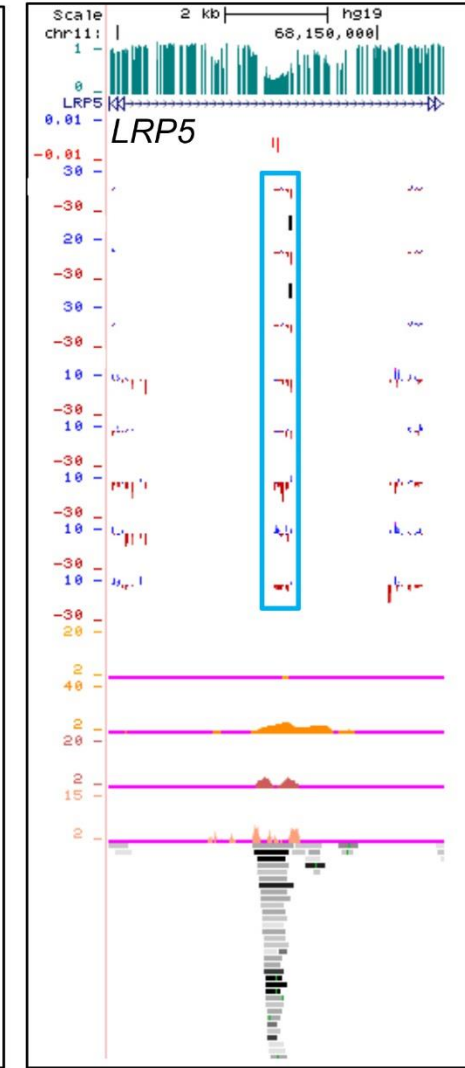


Figure S1 K

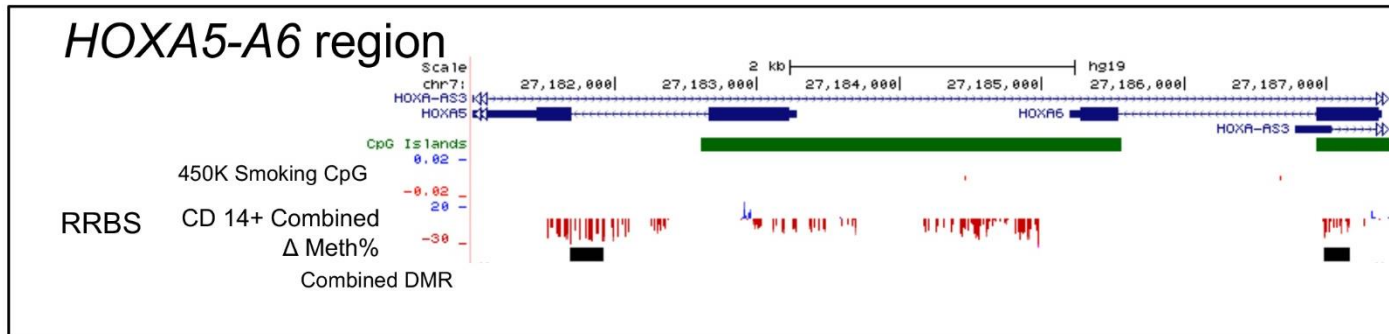


Figure S1. (A–J) Differential methylation tracks (Δ Methylation) and SM-DMRs displayed across sample groups and across CRU-RRBS-Cell types with CD14+ histone modifications and transcription factor clusters (TFs) for (A) *HOXA5* region, (B) *AHRR*, (C) *AHRR-C5orf55-EXOC-AS*, (D) *ALPPL2*, (E) *PPP1R15A (GADD34)*, (F) *F2RL3, F2RL3-CPAMD8*, (G) *SLC24A3*, (H) *RDX*, (I) *SASH1*, and (J) *LRP5*. Boxed areas in each panel show the region of interest. UCSC genome browser views display RRBS results and additional Roadmap Project histone modifications and ENCODE project ChIPseq data. Each panel includes the following browser tracks (from top to bottom):

1. Roadmap Epigenome Whole Genome Bisulfite Sequencing CD14+ cells (nonsmoker)
2. Refseq genes
3. CpG islands
4. 450K Smoking-associated CpGs (from Joehanes et al. 2016)
- 5-10: CD14+ monocytes RRBS methylation difference relative to nonsmokers (Δ Meth%, red, negative change = loss of methylation, blue, positive change = gain of methylation) and SM-DMR location (RRBS-Combined, MESA-RRBS and CRU-RRBS analyses)
- 11-15: Differential methylation tracks for five CRU-RRBS-Celltypes (CD15+ Δ Meth%, CD19+ Δ Meth%, CD4+ Δ Meth%, CD8+ Δ Meth%, CD56+ Δ Meth%)
- 16-19: Roadmap Histone modifications, CD14+ monocytes (nonsmoker)
20. ENCODE transcription factor ChIPseq clusters (compressed view).

(K) *HOXA* region showing nominally significant differentially methylated CpGs at previously unknown smoking-associated locations in or near *HOXA5-A6*. 450K Smoking-associated CpGs are CpG locations identified as smoking associated by Joehanes et al (2016.)

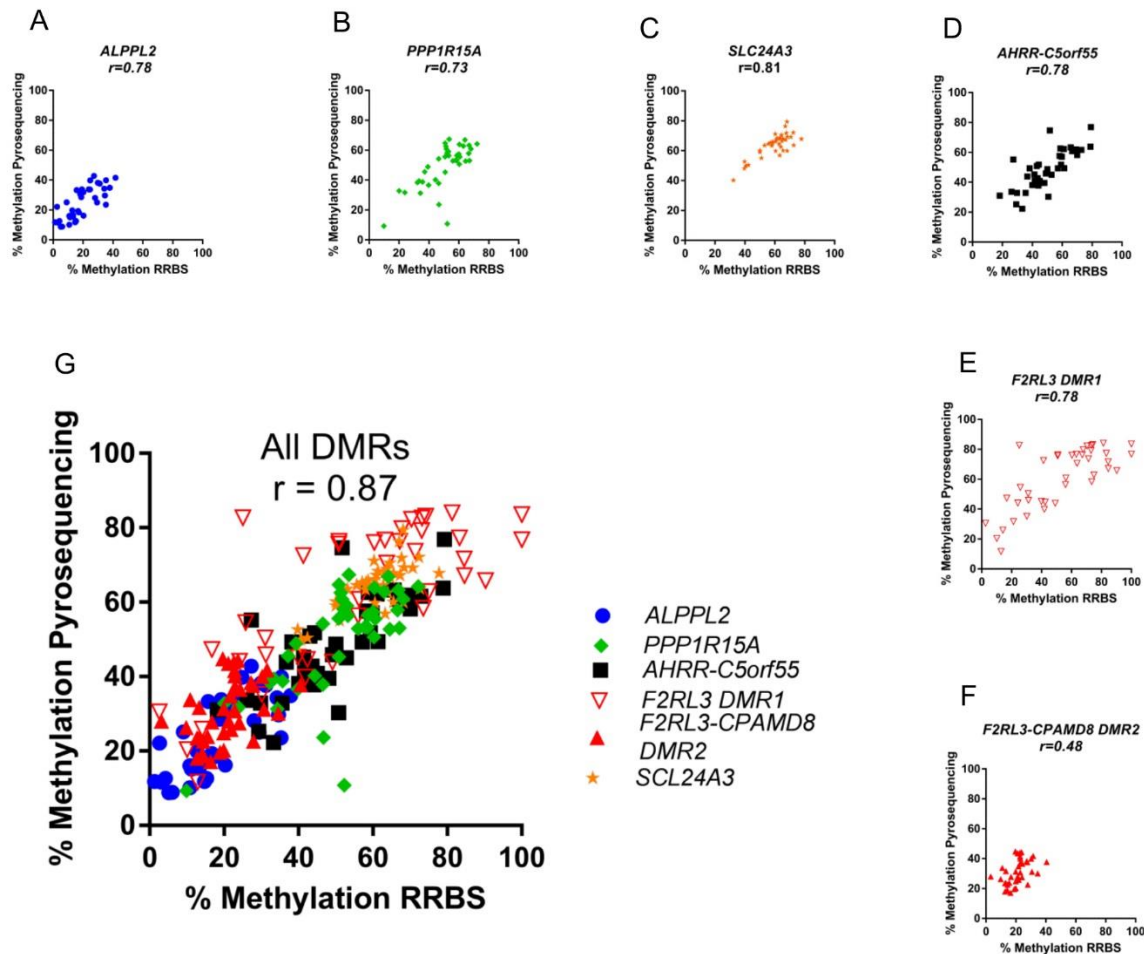


Figure S2. The correspondence between RRBS and pyrosequencing measurements of DNA methylation at SM-DMRs for overlapping CRU subjects. (A-G) Plots show RRBS and pyrosequencing determined measurements of DNA methylation at SM-DMRs. Individual regions (A-F) and overall (G) with their corresponding Pearson correlations are denoted in different color and plotted separately as indicated. LRP5 is not shown.

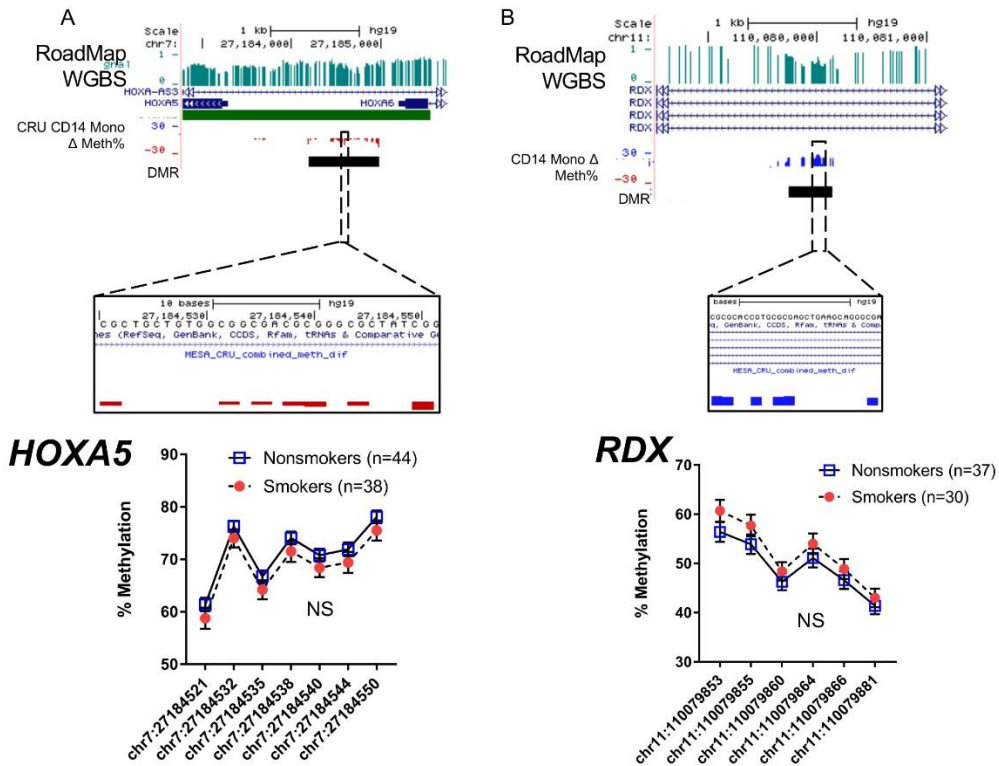


Figure S3. Pyrosequencing measurements on HOXA5 and RDX SM-DMRs. (A, B) Upper panels show genome browser tracks for (top) Roadmap whole genome bisulfite sequencing (WGBS) in monocytes shows location of all CpGs in region, reference gene, differentially methylated CpGs by RRBS (red downward bar = loss of methylation, blue upward bar, gain of methylation), and RRBS DMR (black bar).

Insert shows the RRBS CpGs that were tested. Lower panels show the differential methylated CpGs from HOXA5 and RDX DMRs examined by pyrosequencing in the CRU-Pyro group. At each CpG the difference between smokers and nonsmokers was similar to the RRBS result but this difference was not significant. Subject numbers for successful pyrosequencing assays are indicated for nonsmokers and smokers. Blue open squares represent nonsmokers, red circles indicate smokers.

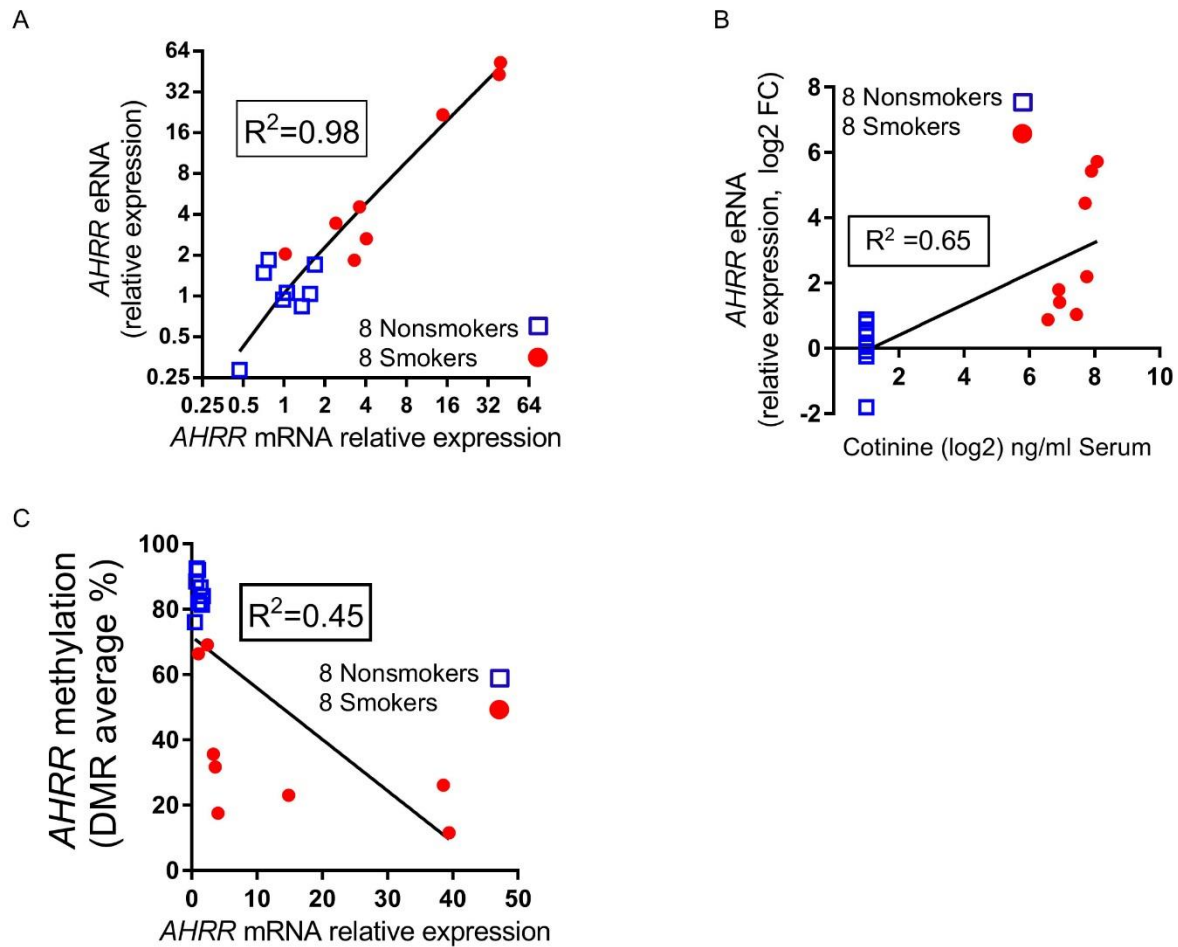


Figure S4. The correlation between AHRR SM-DMR, mRNA, eRNA, and cotinine levels among a subset of eight smokers and eight nonsmokers from the CRU-RRBS, CRU-mRNA and CRU-eRNA subgroups. (A) The correlation between AHRR mRNA, and eRNA expression ($R^2=0.98$, $p < 0.0001$). (B) eRNA was significantly correlated with serum cotinine levels ($R^2=0.65$, $p=0.0007$). (C) The correlation between average AHRR SM-DMR methylation from RRBS and mRNA expression ($R^2=0.45$, $p=0.0042$). Squares represent nonsmokers, while circles indicate smokers.

Supplemental Table S1. Pyrosequencing and BSAS assay primers

GeneID	Forward Primer	Reverse Primer	Sequencing Primer	CpGs
<i>ALPPL2</i>	TTGGTGGAGTTGTTTTTGAGAGAAGT	CATCCCATAAACCCAAATAACCC	ACCTCCATACCATAAAACCC	chr2:233284491, 233284494, 233284501, 233284504, 233284510, 233284517
<i>AHRR-C5orf55-EXOC-AS</i>	TGGGAGGGGGATAAAGAAGG	CCCCACTCCCCTCCTCAA	AGGTAAATTAGGAGGTAGAAA	chr5:403298, 403307, 403312
<i>F2RL3 DMR1</i>	TTGTAGAGGTAGGGGTTGATTGAG	AAACACCCAACCCAAAAACAATA	GATAGTGATATTTTGGAGTTT	chr19:17000465, 17000476
<i>F2RL3 DMR2</i>	AGTTGTTTATAGGTTGGATGTTTTGTTAAT	CCCACAACCTAACATTCTATTC	GTTTGTTTAGAAAGAGGTT	chr19:17004845, 17004874
<i>HOXA5</i>	TTTTGGTAAGTTTGAAGTTAGGATTT	CCCAACAAAACCCAATCTCT	GGGTTTTTAGTTGGGTTTTGAA	chr7:27184521, 27184532, 27184535, 27184538, 27184540, 27184544, 27184550
<i>LRP5</i>	GGTTTAGGTAGTTTGTGTTTGGTATA	BIO-CCTAAAATTAACCTTTAACATATCTCC	ATTTGTTTTGTTTTTTTATTGAAG	chr11:68148333, 68148355, 68148365
<i>PPP1R15A</i>	GTTAGGGTTTTGGGAGTAGT	AATTCCTAAAACCTACCCAAACTCTAACCC	TTTTGGGAGTAGTTTGT	chr19:49378964, 49378970, 49378976, 49378981, 49378985, 49378988
<i>RDX</i>	AGTGGGTATAGTTAGTGGA	BIO-AACTCCCTAACCTCTTAAACTT	GGGTATAGTTAGTGAT	chr11:110079853, 110079855, 110079860, 110079864, 110079866, 110079881
<i>SLC24A3</i>	TGAAGGATTTAAGGGGATTGTTTAATATA	CCCATAACTAAAAACCAACACTTTTTCTC	TTTTTTAATTTTTTTTAGAATAGG	chr20:19194143
<i>AHRR-BSAS</i>	TTTTTGGTTGTAGTGAATTTG	AAAAACCAACCTATCCCCTAC	NA	chr5:373193, 373199, 373203, 373248, 373250, 373315, 373353, 373355, 373378, 373398, 373423, 373476, 373490, 373494

Supplemental Table S2. mRNA Real-time Polymerase Chain Reaction Assays

Gene	Assay ID	
<i>ACTB</i>	HS99999903_m1	
<i>ALPPL2</i>	HS00741068_g1	
<i>AHRR</i>	HS01005075_m1	
<i>AHRR-C5orf55-EXOC-AS</i>	HS00744706_s1	
<i>F2RL3-CPAMD8</i>	HS00610855_m1	
<i>F2RL3</i>	HS01006385_g1	
<i>HOXA5</i>	HS00430330_m1	
<i>LRP5</i>	HS00182031_m1	
<i>PPP1R15A</i>	HS00169585_m1	
<i>RDX</i>	NT	
<i>SASH1</i>	HS00323932_m1	
<i>SLC24A3</i>	HS00915003_m1	

Supplemental Table 3. Enhancer RNA RT-PCR Primers

SM-DMR enhancer regions	Locus	Primers-Forward	Primers-Reverse
AHRR - Region 2	chr5:371604-371689	TGTCCTAGACCCACCTGTCC	GCCTGGAAGGACAGTCAGAG
AHRR - Region 3	chr5:371947-372028	TTCCTCTCATCTGGCTGCTT	AGGTCCCTATTGGCAGGACT
AHRR - Region 4	chr5:372157-372267	CTCAGTGCCTTCCCTCCTC	ACAGTCCCCTCTGTGCAGTC
AHRR - Region 5	chr5:372483-372586	CATCTCGCCTTCTTTTCAGC	GGCTGAAGTTTCTGCTCTG
AHRR - Region 6	chr5:372694-372825	CTACCAGGGAATTGGCTTCA	ACTCGGTCTGACCCACAGTC
AHRR - Region 7	chr5:373202-373306	TGGGGACTGTTACCTCTG	GCAATGAGAGGCTGGAAGG
AHRR - Region 8	chr5:373630-374020	GACTGGGGCTGGAGAGGT	GCCTTTTCCCCGAGAAGC
AHRR - Region 9	chr5:374334-374463	GGTCTGCACTTGGGAGAAAG	GAAACCCAGCATCCTAGC
AHRR - Region 10	chr5:374798-375014	CTGAATCAATGGCTCCCAGT	TGCTCTGAGCTCTCCTCACC
AHRR - Region 11	chr5:376293-376403	AGGAGGGGACAGAGAGCAGT	CTCAGATCAGCCAGGAGAGG
AHRR - Region 12	chr5:375746-375863	CCCTCAGTCTCTCAGAAGC	GACTGACCAAGGGCTGAGTT
AHRR - Region 14	chr5:369226-369349	CCATGTGTCAAGAGCAGCAC	CCTCAGGTGTCCCAGAAAAC
AHRR - Control Region 1 (Locus A)	chr5:349510-349652	GTGAAACCCCGTCTCCACTA	GTGATCTTGGCTCACTGCAA
AHRR - Control Region 21 (Locus B)	chr6:2613040-2613144	GCAGCCATAAAAAGGGTGAG	GCGGTGTTTGGTTTTCTGTT
ALPPL2-6eRNA	chr2:232421401-232421482	TCACCAGGGAGCTTCTTTA	GAGACAACCAGAAGCCCAGA
C5orf55-EXOC-AS-1eRNA	chr5:402729-402817	GCATTTACCCAAAGGAGCTG	CTGTGCATTCACTGGTGGAC
F2RL3-1eRNA	chr19:16889457-16889569	TGTCATGTCCAGCTGTTTCC	GGAGAGAGAGGGGACCTCAG
F2RL3-CPAMD8-1eRNA	chr19:16893971-16894052	CGGCTGGATGTTTTGCTAAT	CATTCCCAGAAAGCACTCGT
HOXA5-1eRNA	chr7:27143756-27143844	CGAATCACGTGCTTTTGTTG	CTCCACCCAACTCCCCTATT
LRP5-2eRNA	chr11:68380478-68380570	CCGACAGCTAGAAGGAAGGA	TGCCACTCTCACACAAAAGC
SASH1-1eRNA	chr6:148363394-148363485	GGCCTCCCAACACTTCTTTT	AGGTTGCAGTGAGCGAAGAT

References

Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. 2016. Epigenetic signatures of cigarette smoking. *Circ Cardiovasc Genet* 9:436-447.