PNAS
www.pnas.org

Supplementary Information for

Mother-child transmission of epigenetic information by tunable polymorphic imprinting

Brittany L. Carpenter, Wanding Zhou, Zachary Madaj, Ashley K. DeWitt, Jason P. Ross, Kirsten Grønbæk, Gangning Liang, Susan J. Clark, Peter L. Molloy, and Peter A. Jones.

Peter A. Jones
Email: Peter.Jones@vai.org

This PDF file includes:

Figs. S1 to S7
Tables S1 and S2

www.pnas.org/cgi/doi/10.1073/pnas.1815005115

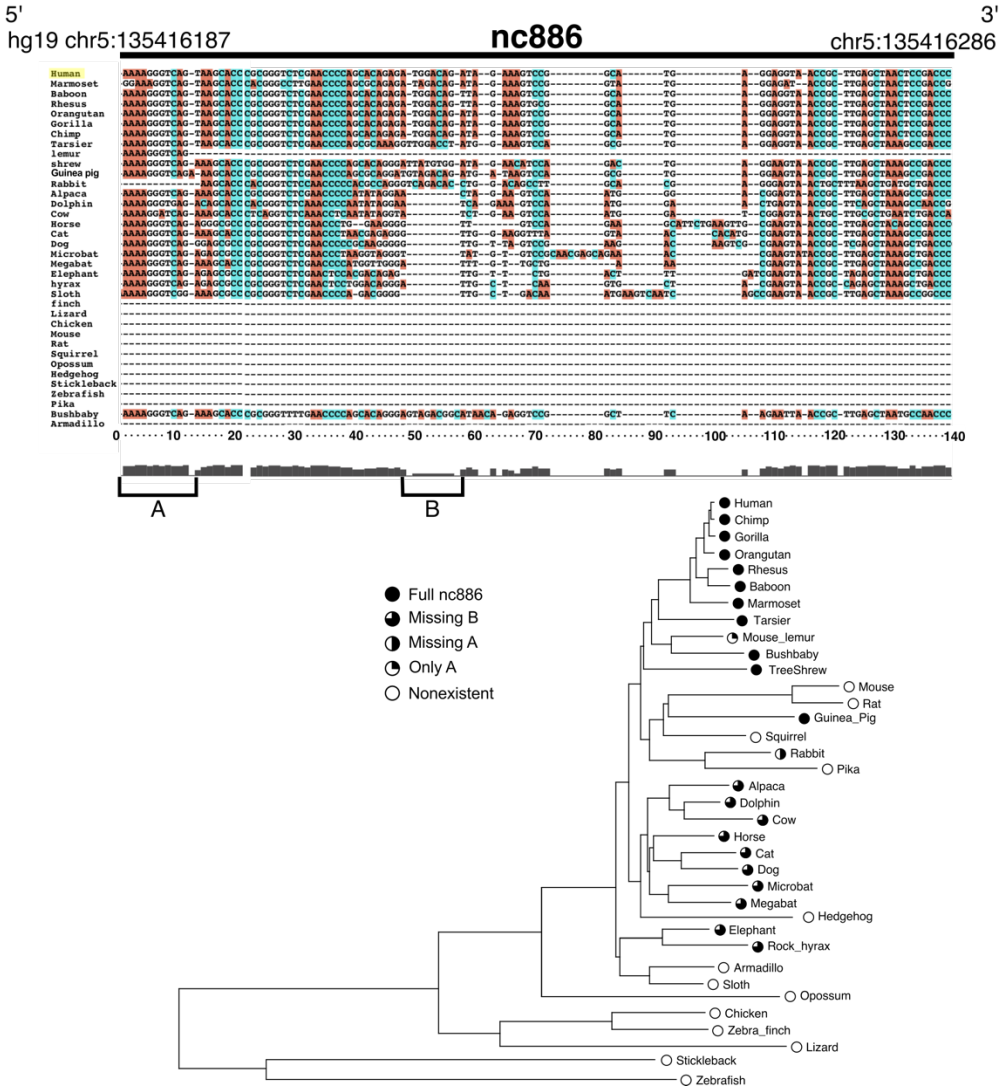

Fig. S1. Species alignment for nc886. Sequence alignment of the nc886 locus and a pruned tree from the UCSC 46-vertebrate species tree.

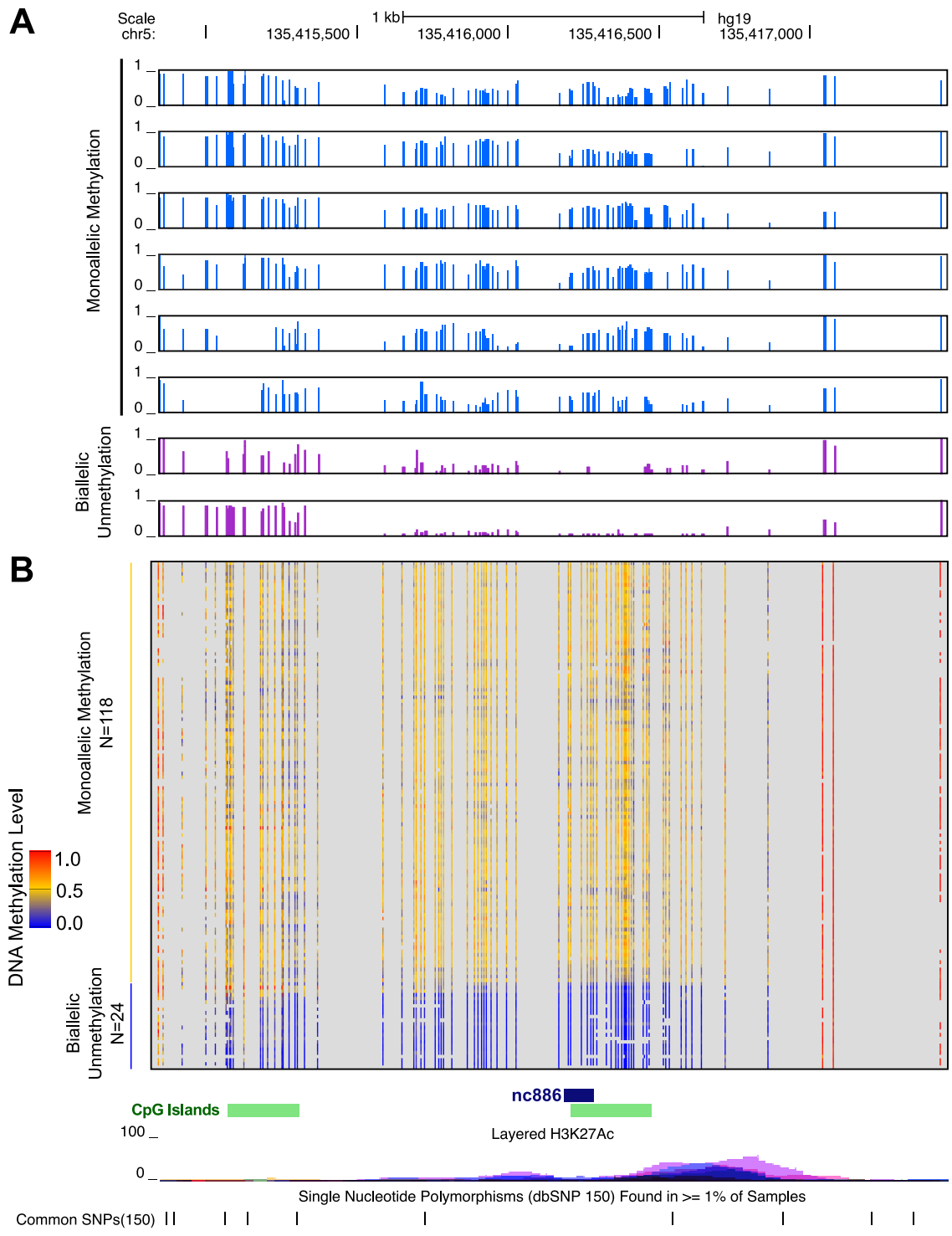

Fig. S2. Whole genome bisulfite sequencing for the nc886 DMR. The nc886 DMR was defined as 1979 bps using WGBS data for individuals from TCGA (A) and BLUEPRINT (B) data sets.

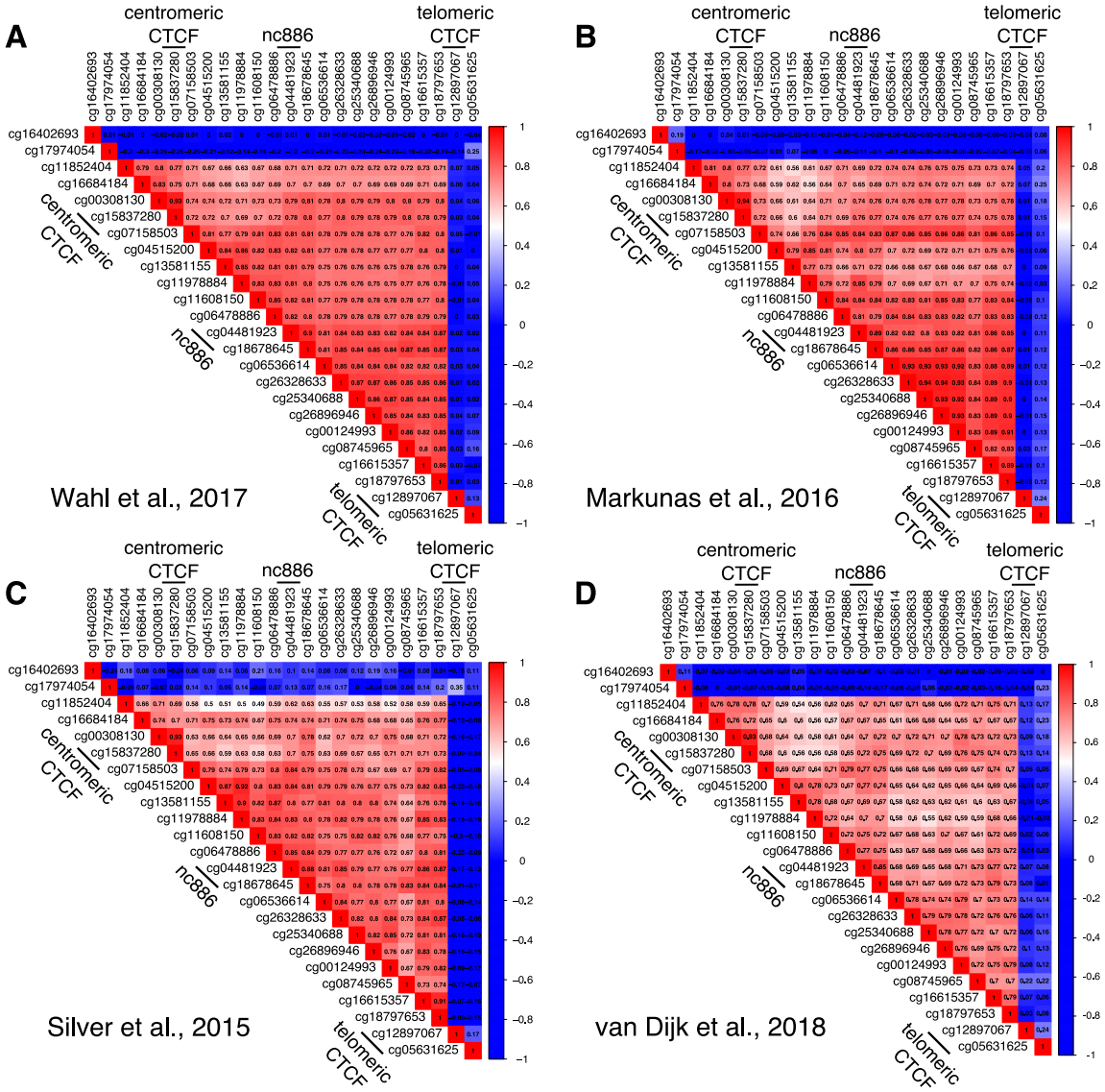

Fig. S3. Correlation of beta values across the nc886 DMR. Correlograms were generated based on beta values across the nc886 DMR and flanking invariable CpG sites. The data are from four independent data sets that were analyzed in this study.

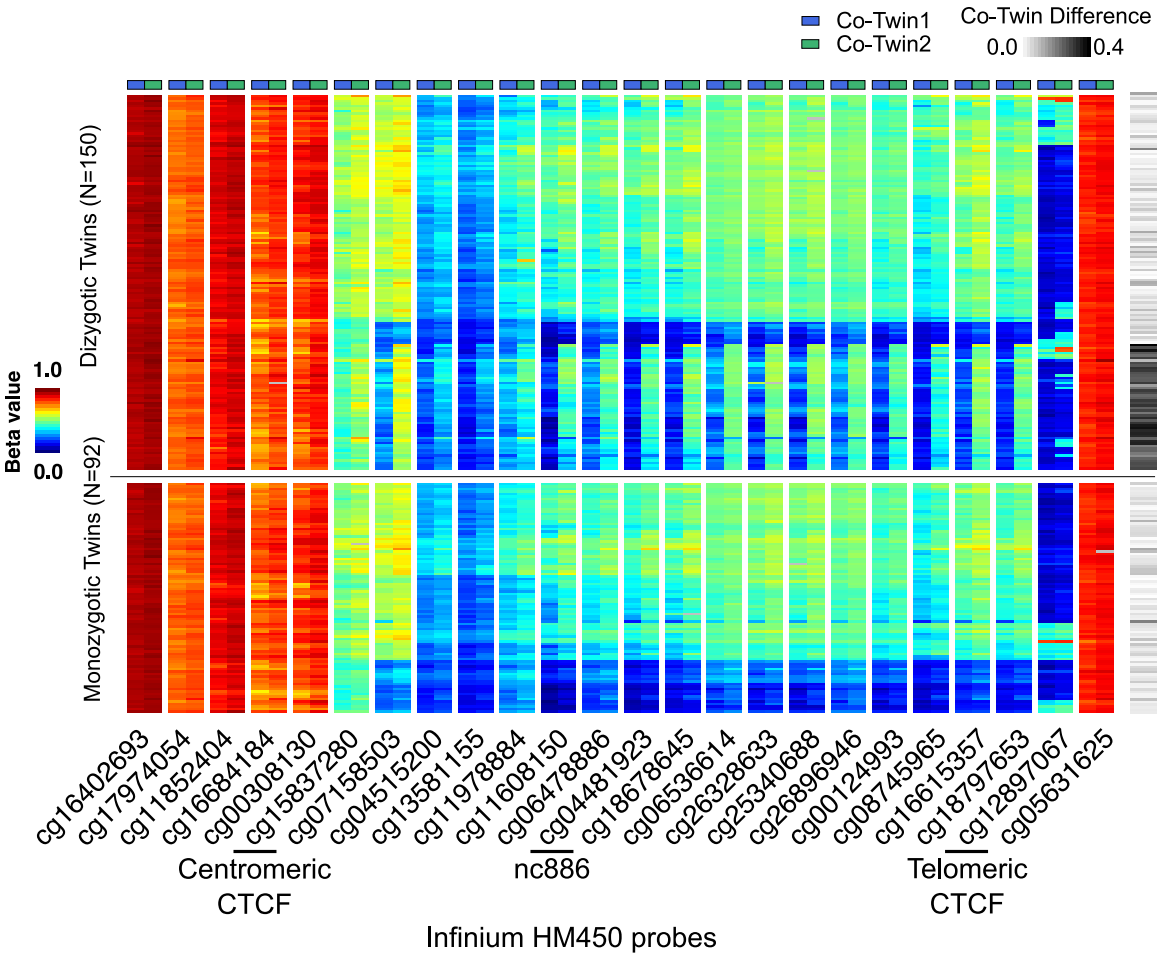

Fig. S4. Monozygotic, but not dizygotic, twins are concordant for DNA methylation across the nc886 DMR. HM450 data was analyzed from adipose tissue of monozygotic (bottom panel) and dizygotic (top panel) twins. Heat map displays beta values for co-Twin1 and co-Twin2 at each CpG probe across the DMR. Difference in beta values between co-twins for cg04481923 (nc886) are represented on the right.

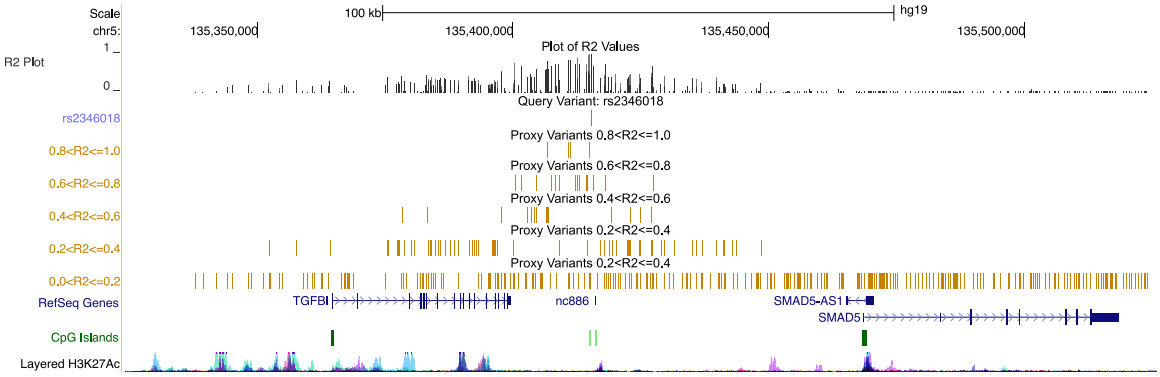

Fig. S5. The local haplotype for nc886. LDlink was used to calculate R2 values between variant rs2346018 and surrounding variants. R2 values are displayed for all populations from the 1000 genomes project. Local haplotypes range in size from 5 kb – 35 kb from rs2346018 depending on the population.

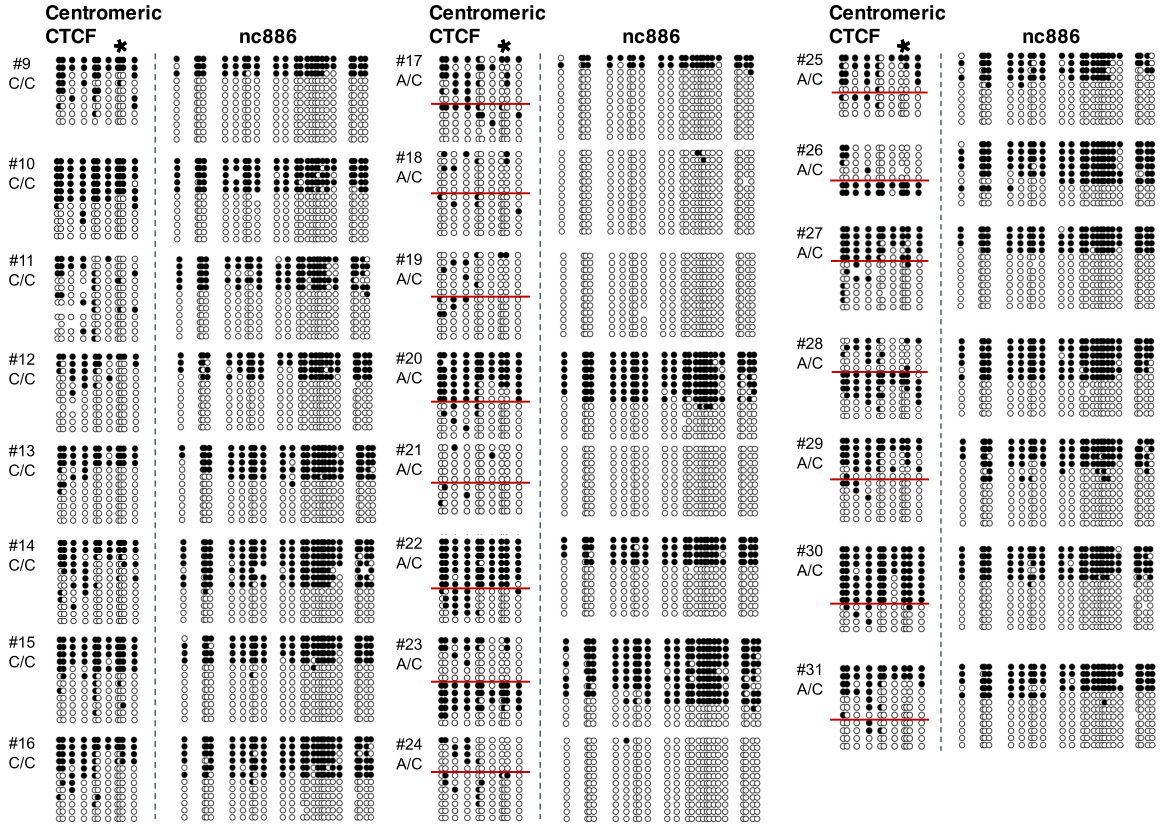

Fig. S6. Bisulfite sequencing of centromeric CTCF and nc886. Locus-specific bisulfite sequencing analysis for individuals not represented in Figs. 2. Clones for centromeric CTCF are not physically connected on an individual DNA strand. Clones for heterozygous individuals are sorted first by A/C and then by decreasing amount methylation at the centromeric CTCF. Clones for homozygous individuals and all individuals at nc886 are sorted by decreasing amount of methylation.

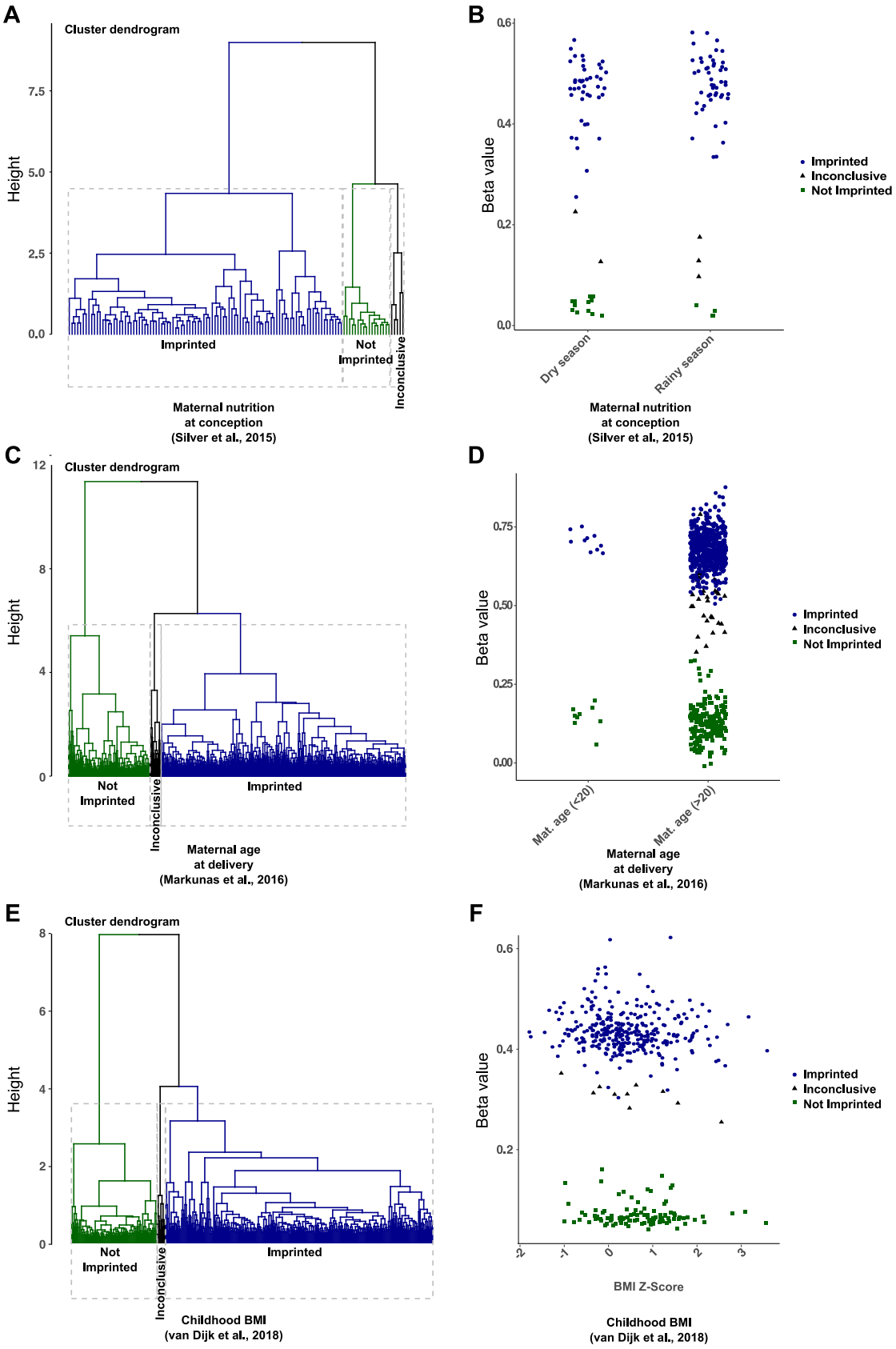
A
Cluster dendrogram
Height
7.5
5.0
2.5
0.0
Imprinted
Not Imprinted
Inconclusive
Maternal nutrition at conception (Silver et al., 2015)
B
Beta value
0.6
0.4
0.2
0.0
Dry season
Rainy season
Maternal nutrition at conception (Silver et al., 2015)
Imprinted
Inconclusive
Not Imprinted
C
Cluster dendrogram
Height
12
8
4
0
Not Imprinted
Inconclusive
Imprinted
Maternal age at delivery (Markunas et al., 2016)
D
Beta value
0.75
0.50
0.25
0.00
Mat. age (<20)
Mat. age (>20)
Maternal age at delivery (Markunas et al., 2016)
Imprinted
Inconclusive
Not Imprinted
E
Cluster dendrogram
Height
8
6
4
2
0
Not Imprinted
Inconclusive
Imprinted
Childhood BMI (van Dijk et al., 2018)
F
Beta value
0.6
0.4
0.2
-2
-1
0
1
2
3
BMI Z-Score
Childhood BMI (van Dijk et al., 2018)
Imprinted
Inconclusive
Not Imprinted

Fig. S7. Hierarchical clustering of beta values in the nc886 DMR. A) Cluster dendrogram for data from GSE59592 based on beta values in the nc886 DMR. B) Individuals are grouped by dry or rainy season of conception, and beta values for cg04481923 are plotted. The imprinted, not imprinted, and inconclusive groups are based on the cluster dendrogram from panel A. C) Clustering analysis was performed as for panel A, using data from GSE82273. D) Individuals are grouped by age (< 20 years and > 20 years), and beta values for cg04481923 are plotted. The imprinted, not imprinted, and inconclusive groups are based on the cluster dendrogram in panel C. E) Clustering analysis was performed as for panel A, using data from GSE103657. F) Individuals are plotted by BMI measurement from the DOMInO study and beta values for cg04481923 are plotted. The imprinted, not imprinted, and inconclusive groups are based on the cluster dendrogram in panel E. When we performed the analyses including the "inconclusive" group we find that season of conception is sensitive to this inclusion ($P = 0.096$); beta regression estimated the odds ratio for rainy season being methylated at the nc886 probe versus the dry season as 1.395 ($P = 0.0322$). Maternal age and BMI were not sensitive to inclusion of the "inconclusive" group ($P = 0.0289$ and $P = 2.22 \times 10^{-3}$ respectively).

H19	PEG3	nc886
cg16675558	cg22220806	cg16402696
cg03996735	cg13946792	cg17974054
cg18104242	cg15473155	cg11852404
cg27300742	cg19041006	cg16684184
cg25281616	cg17663463	cg00308130
cg01895612	cg19771589	cg15837280
cg23476401	cg02793099	cg07158503
cg00237904	cg01656470	cg04515200
cg06765785	cg27519373	cg13581155
cg25821896	cg07310951	cg11978884
cg18454954	cg02478023	cg11608150
cg25579157	cg22354595	cg06478886
cg02886509	cg19335327	cg04481923
cg02657360	cg14849423	cg18678645
cg16574793	cg15777825	cg06536614
cg09452478	cg10204755	cg26328633
cg20891060	cg12205903	cg25340688
cg20049005	cg26349266	cg26896946
cg19462210	cg02162069	cg00124993
cg00220736	cg13960339	cg08745965
	cg13369939	cg16615357
	cg20628335	cg18797653
	cg22294267	cg12897067
	cg19098268	cg05631625
	cg24844423	
	cg15678121	
	cg25458871	
	cg18668753	
	cg18706888	
	cg06652523	
	cg01054891	
	cg13374648	
	cg26917367	
	cg22927979	

Table S1. Infinium HM450 probe IDs used to analyze DNA methylation across a paternally methylated DMR (H19), a maternally methylated DMR (PEG3) and a polymorphically imprinted DMR (nc886). Probe IDs correspond to heat maps in Figure 1.

Study	Index SNP	Tissue Investigated	# of Informative Individuals	Conclusion
Paliwal et al., 2013	rs2346019	Placenta	10	Allele specific DNA methylation of nc886 is maternally derived
Treppendahl et al., 2012	rs9327740	Peripheral Blood	1	Allele specific DNA methylation of nc886 is maternally derived
Romanelli et al., 2014	rs2346018 rs2346019 rs9327740	Placenta & Cord Blood	6	nc886 is a maternally methylated DMR

Table S2. Published studies concluding that DNA methylation of the nc886 DMR is maternally derived.