# Mother-child transmission of epigenetic information by tunable polymorphic imprinting

Short title: Environmental influences on polymorphic imprinting

Brittany L. Carpenter[1, †], Wanding Zhou[1, †], Zachary Madaj[1], Ashley K. DeWitt[1], Jason P. Ross[2], Kirsten Grønbæk[3], Gangning Liang[4], Susan J. Clark[5], Peter L. Molloy[2], and Peter A. Jones[1*].

**Affiliations:**

[1]Van Andel Research Institute, Grand Rapids, MI, USA.

[2]CSIRO Health and Biosecurity, North Ryde, NSW, Australia.

[3]Department of Hematology, Rigshospitalet, Faculty of Health Sciences, University of Copenhagen, Copenhagen, 2100, Denmark.

[4]Department of Urology, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, USA.

[5]Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, New South Wales 2010, Australia.

* Corresponding author: Peter.Jones@vai.org

†These authors contributed equally to the work.

**Abstract**

Genomic imprinting mediated by DNA methylation restricts gene expression to a single allele determined by parental origin and is not generally considered to be under genetic or environmental influence. Here, we focused on a differentially methylated region (DMR) of about 1.9 kb that includes a 101-bp noncoding RNA gene (nc886/VTRNA2-1), which is maternally imprinted in approximately 75% of humans. This is unlike other imprinted genes, which demonstrate monoallelic methylation in 100% of individuals. The DMR includes a CTCF binding site on the centromeric side defining the DMR boundary and is flanked by a CTCF binding site on the telomeric side. The centromeric CTCF binding site contains an A/C polymorphism (rs2346018); the C allele is associated with less imprinting. The frequency of imprinting of the nc886 DMR in infants was linked to at least two non-genetic factors, maternal age at delivery and season of conception. In a separate cohort, nc886 imprinting was associated with lower BMI in children at 5 years of age. Thus, we propose that the imprinting status of the nc886 DMR is "tunable" in that it is associated with maternal haplotype and prenatal environment. This provides a potential mechanism for transmitting information, with phenotypic consequences, from mother to child.

**Significance Statement**

*First,* our work provides critical biological interpretation of intermediate DNA methylation readouts at the nc886 DMR. nc886 was identified in multiple large-scale EWAS studies that failed to recognize that this region acts as a contiguous DMR imposed by genomic imprinting, highlighting the need to re-examine several 450k datasets. *Second*, strict control of genomic imprinting was thought to be required for organismal viability. Reports of polymorphic imprinting are limited to specific tissue types such as placenta and brain. In blood and somatic tissues, we show nc886 imprinting is mosaic in the population and influenced by maternal environment.

**Main Text**

Epigenome-wide association studies (EWAS) have used the Infinium HumanMethylation 450 (HM450) array to identify CpG sites that are differentially methylated in the DNA of individuals with differing phenotypes or diseases. Whereas most studies are cross-sectional, some have sought to identify risk-associated epigenetic marks that hold predictive value (1, 2). It has been known for many years that genetic polymorphisms can alter the frequency of DNA methylation, that is, they can alter allele-specific DNA methylation (ASM) (3-5). DNA methylation is also associated with environmental perturbuations (6-10): for example, maternal and paternal exposures such as nutrition and smoking are linked to the methylation states of CpG sites in children (11-15).

Genomic imprinting is a type of ASM that is usually associated with extended regions of parental-specific differentially methylated DNA (DMRs) and is established in the gametes. Most of the hundred or so imprinted genes are monoallelically methylated at imprinting control regions, where the gene is expressed from the unmethylated allele in nearly all somatic tissues at a given developmental stage (16). The exceptions include a small subset of genes that are imprinted in a tissue- or isoform-specific manner, in brain and placenta, across most individuals (16, 17). Importantly, unlike non-imprinted ASMs, imprinting is not generally altered by genetic or nongenetic factors, but is instead fixed in the population (18). Tight control of nearly all imprinted genes across the human population is consistent with the known importance of imprinting in regulating embryonic development and neuronal function (19, 20). Disruptions in imprinting can lead to human disorders, including Beckwith-Wiedemann, Silver-Russel, Prader-Willi, and Angelman syndromes; all of which have severe phenotypes impacting human growth and development (21, 22).

Work from our lab uncovered a unique pattern of DNA methylation at the promoter of a non-coding gene (nc886), where approximately 75% of individuals demonstrated ASM and 25% were completely unmethylated (23). Romanelli and colleagues expanded on this finding to determine that this region of DNA was polymorphically imprinted in a small sample set (24). Prior to the conclusion that nc886 was polymorphically imprinted, interindividual variation of imprinting status had only been reported for placental-specific imprinted genes (25, 26).

Genomic imprinting is considered an all-or-none process, in which DMRs on either the maternal or paternal autosomes cause monoallelic expression of the corresponding gene (27, 28). Imprinting is commonly, but not always, established by silencing of a single allele through DNA methylation (29). Maternally methylated imprinted genes are often associated with decreased size of the developing fetus (30). Whereas, paternally expressed imprinted genes are often associated with increased fetal weight (30). Importantly, maternal exposures, during pregnancy, shift the percentage of DNA methylation or level of expression at imprinted genes, primarily in the placenta (24, 31, 32). The best studied example of genetic and environmental impacts on imprinted genes is that of the H19/IGF2 locus, where quantitatively small but additive effects on levels of DNA methylation are seen (33). However, the possibility that a parental cue from the environment or parental genetics might shift the likelihood that genomic imprinting is established in a human child has yet to be investigated.

Here, we studied the frequency of imprinting of a RNA polymerase III–transcribed, noncoding RNA of 101 bp called nc886 (also referred to as VTRNA2-1), which can be silenced by DNA methylation (23, 34). The complete nc886 sequence is found in higher-order primates, although portions of it are present in other mammals (*SI Appendix*, Fig. S1), which may indicate its importance in primate evolution. ASM at the nc886 locus has previously been studied and

described as a human “metastable epiallale”, which is an allele that shows epigenetic heterogeneity in a population and is sensitive to environmental conditions (6, 35). However, an appreciation of the imprinting biology of this locus has not been emphasized in prior studies. In this work, we show that the region surrounding nc886 acts as an imprinted DMR, as opposed to a intermediately methylated region. This locus represents an instance of non-placental polymorphic imprinting in humans (24) and we show that variation of imprinting in the population is associated with prenatal environment.

**nc886 is part of a 1.9-kb polymorphically imprinted region**

Genomic imprinting occurs as parental-specific DMRs that tightly regulate gene dosage (28). Non-imprinted ASMs, on the other hand, depend on genomic context rather than parental origin, and local polymorphisms to affect CpG methylation states, which are generally confined to shorter regions of DNA (3, 18, 36). We and others have reported that nc886 expression is tightly regulated by DNA methylation and that the nc886 locus shows polymorphic imprinting in the human population when measured in peripheral blood cells: roughly 75% of individuals worldwide are monoallelically methylated and about 25% of individuals are biallelically unmethylated at nc886 (23, 24, 34, 37).

We first used HM450 data to examine DNA methylation at known imprinted DMRs (Fig. 1) (38). We calculated the variance of beta values in 26 imprinted DMRs that had a minimum of five HM450 probes using data from peripheral blood of 2,664 Europeans and Indian Asians from the London Life Sciences Prospective Population (LOLIPOP) study (39). We confirm that all imprinted DMRs investigated, except for probes adjacent to nc886, have limited inter-individual

variance in this cohort. This variance is much higher for nc886 than other included imprinted DMRs (Fig. 1A).

We chose to survey DNA methylation in more detail for known imprinted loci containing a maternally methylated DMR (PEG3), a paternally-methylated DMR (H19), and to expand our window of investigation of nc886, a polymorphically imprinted locus. We analyzed data from two independent studies of sperm DNA (GSE47627 and GSE64096) and also peripheral blood from the LOLIPOP study (39-41). The H19 DMR is fully methylated and the PEG3 DMR generally lacks DNA methylation in 66 sperm samples (Fig. 1B and 1E respectively). We found that the nc886 locus and its flanking regions were unmethylated in all sperm samples, similar to PEG3 a maternally-methylated DMR (Fig. 1H, dark blue), with partial methylation at a CTCF site on the telomeric side. As expected of imprinted loci, we saw approximately 50% methylation in all peripheral blood samples for both H19 and PEG3 (indicated by green; Figs. 1C and 1F). For nc886 in peripheral blood, the majority of samples showed about 50% methylation (Fig. 1I) and the remainder showed methylation close to 0% (blue). Based on these HM450 data we defined the nc886 region as a maternally derived DMR in the human population with boundaries marked by two CTCF sites (spaced ~ 3 kb apart; Figs. 1H and 1I). Consistent with previous work from Monk and colleagues (24), whole genome bisulfite sequencing data of various normal tissues from TCGA and BLUEPRINT (*SI Appendix*, Fig. S2) better resolves the boundaries of the nc886 DMR, which is 1979 bp in size and includes the nc886 locus. nc886 is overlapped by a CpG island and the DMR is bounded on the centromeric side by repetitive sequences just outside a variably methylated CTCF site (Fig 1H; 1I) (24, 42). The nc886 DMR is flanked on the telomeric side by a CTCF site that is unmethylated in peripheral blood (Fig. 1I).

We noted that three Type-II probes (cg04515200, cg13581155, and cg11978884) within the nc886 DMR on the HM450 platform displayed a consistent bias towards hypomethylation relative to the rest of the probes in the DMR (Fig. 1H). We believe this is inherent to the Infinium probe design, because we could still distinguish a separation in the methylation beta values for individuals having monoallelic methylation at these probes and based on WGBS data, the three CpGs do not appear biased from an averaged methylation level of 0.5 (Fig. 1I and *SI Appendix,* S2). The high correlation of beta values between these three probes and probes across the nc886 DMR also support the idea that these three CpGs bear consistent methylation with the entire DMR and that this DMR is acting as a contiguous unit (*SI Appendix,* Fig. S3).

Imprinting is essentially dichotomous: DMRs are either monoallelically methylated or not. As such, methylation data cannot be considered a continuous variable in which the methylation values of all the assayed DNA molecules are averaged. We reexamined the LOLIPOP data set, taking this biology into account. For the H19 and PEG3 DMR, all data points were near a beta value of 0.5, as expected (Figs. 1D and 1G). However, for nc886 the dichotomous nature of the data became clear, in that the nc886 DMR showed beta values primarily at either 50% or near 0% with the exception of the three CpGs previously mentioned (Fig. 1J). Thus, by using the data from the large cohort (39), we have confirmed the previous observations that the nc886 DMR is polymorphically imprinted in humans (23, 24).

**An A/C SNP is associated with local DNA methylation density**

It has been unclear whether the polymorphic nature of imprinting of the nc886 DMR is governed by genetic or nongenetic factors, given the lack of evidence in the literature that genomic imprinting can be associated with the local haplotype (5). Similar to Van Baak et al.

(35), we reanalyzed DNA methylation data from twins, generated by Grundberg and colleagues (43). We confirmed their findings that monozygotic twins, but not dizygotic twins, are concordant for DNA methylation across the nc886 DMR (*SI Appendix,* Fig. S4), strongly suggesting that the nc886 DMR could be influenced by genetic factors. The nc886 locus is located within a haplotype block ranging in size from approximately 5 kb to 35 kb, depending on the population (*SI Appendix,* Fig. S5). The 2-kb spanning the nc886 DMR has been examined for potentially causal SNPs without success or with the conclusion that SNPs did not impact DNA methylation (6, 23, 24, 37, 42). Given the central role of CTCF in imprinting, we focused on an A/C SNP (rs2346018; A allele frequency range: 31-45% and C allele frequency range: 54-68% from the 1000 Genomes Project) located in the variably methylated centromeric CTCF binding site (36, 44, 45). Using the R package motifbreakR we found that having a C SNP "breaks" the CTCF binding motif of the centromeric CTCF site, and could, therefore, impact the ability of CTCF to bind its target (46). To determine if SNP status was associated with local DNA methylation density, we analyzed DNA methylation at the centromeric CTCF site and across the nc886 locus in peripheral blood DNA of 31 cancer-free individuals using bisulfite conversion followed by clonal sequencing (primer locations shown in Fig. 2A). Eight examples are shown in Fig. 2B, with the data from the remaining 23 individuals shown in *SI Appendix,* Fig. S6. Clearly, the A and C alleles can be densely methylated in both homozygotes and heterozygotes, but there was also sporadic methylation in both predominantly methylated or unmethylated DNA strands (Fig. 2B).

We chose to test the hypothesis that genetic background could "tune" the likelihood that an individual displays imprinting at the nc886 DMR, thus switching from imprinted to not-imprinted. Examining data from both heterozygous and homozygous individuals, we determined

that the A allele has higher average methylation density than the C allele (Fig. 2C; $P < e–9$). Furthermore, the odds ratio that a clone of the centromeric CTCF site is fully methylated in an individual with the A polymorphism (versus C) was 2.86. These data demonstrate that DNA methylation of the centromeric CTCF site, 900 bp from nc886, correlates with a common SNP, with an A allele more likely to be methylated than a C allele. Given the role of CTCF in imprinting, the presence of the A allele might directly alter the likelihood of establishing DNA methylation at the nc886 DMR through reduced sequence binding affinity of CTCF or other DNA-binding proteins (e.g. DNMTs, ZFP57, Kaiso, etc.) (47, 48). Alternatively, the A allele might indirectly impact imprinting through increased DNA methylation density and therefore, binding affinity of CTCF (49) (Fig. 5).

**Linked DNA methylation of the centromeric CTCF binding site and nc886**

Previous work has shown that DNA methylation at the centromeric CTCF site serves as the boundary of ASM for the nc886 region (24, 42). Because the entire 1.9 kb region between the CTCF sites acts as a DMR (Figs. 1, and *SI Appendix,* Figs. S2 and S3), we hypothesized that DNA methylation of the centromeric CTCF site, or of any region within the DMR, might predict DNA methylation at nc886. While other studies have suggested that the entire region can be imprinted [23,32], they have not shown a lack of DNA methylation of this region in sperm, as we have (Fig 1H) and it is technically challenging to show that DNA methylation is linked across a single DNA strand.

Figure 2B shows clonal bisulfite sequencing analysis of genomic DNA from individuals who have either monoallelic methylation or no methylation, as indicated by SNP status in the centromeric CTCF site. We classified 50% methylation, representative of monoallelic

methylation at nc886 as being "1" and 0% methylation, or biallelic lack of methylation as being "0" and modeled the data against the percentage of DNA methylation at the centromeric CTCF site for all 31 individuals (Fig. 2D). Lack of DNA methylation at the centromeric CTCF site was associated with a very low probability of having DNA methylation at nc886. With increasing density of DNA methylation at the centromeric CTCF site, the probability of nc886 being classified as "imprinted" increased; thus, there was a strong positive correlation between DNA methylation at the centromeric CTCF site and at nc886 (Fig. 2D; $P < 0.0001$; $R^2$ 0.77). DNA methylation of the centromeric CTCF site, therefore, explains much of the variation in DNA methylation at nc886. While we were not able to sequence individual DNA strands greater than 1 kb in length, these data and the high correlation between beta values across the nc886 DMR (*SI Appendix,* Fig. S3) suggest that DNA methylation of this region is indeed present on individual alleles (Fig. 2).

**DNA methylation of the centromeric CTCF binding site is maternally derived**

As summarized in *SI Appendix,* Table S2, previous studies analyzing SNPs have independently concluded that DNA methylation across the nc886 region is not dependent on genetic context but is maternally derived (23, 24, 37, 42). Additionally, when analyzing HM450 data from sperm we find that the nc886 DMR lacks DNA methylation, indicating that methylation of this region is likely maternally derived (Fig. 1H). To confirm that DNA methylation of the centromeric CTCF site could be included in this region of maternally derived methylation, we used parent-offspring trios informative for the A/C SNP (rs2346018) and performed bisulfite conversion and clonal sequencing. As shown in Fig. 3, DNA methylation of the centromeric CTCF site was maternally derived. Thus, the existing literature and the lack of

DNA methylation across the nc886 DMR in sperm (Fig. 1H), along with this result supports that imprinting of the nc886 DMR is not paternally derived.

**Maternal age and nutrition are associated with the nc886 DMR**

Recent EWAS studies have identified nc886 as a "metastable epiallele" that is altered in the population and is dependent on maternal nutritional status (6, 35, 50). However, these studies did not consider the data in terms of the percentage of each cohort demonstrating imprinting, which might account for the biology. Therefore, we reanalyzed the results from three independent studies in terms of percentage of each subpopulation that demonstrates imprinting, in order to determine if the likelihood that an individual is called "imprinted" shifts in each experimental group (*SI Appendix,* Fig. S7). To address global shifts in beta values introduced by different methods used to preprocess the HM450 data or by residual experimental batch effects, we applied unsupervised hierarchical clustering, using the DNA methylation beta values from the entire region to determine whether an individual had imprinting at the nc886 DMR and whether this call is conclusive.

We used the valuable data from Silver and colleagues (6), who examined DNA methylation in 114 children in Gambia conceived in either the rainy season (low calorie, nutritionally rich) or dry season (high calorie, nutritionally poor). Hierarchical clustering identified three groups of individuals, two of which showed the expected dichotomy of being imprinted or not imprinted. The third cluster, consisting of five children (4.4%) who had evidence of DNA methylation in only about half of the probes in the nc886 DMR, was classified as inconclusive (*SI Appendix,* Figs. S7A and S8B). When comparing imprinted individuals and non-imprinted individuals we found that season of conception was significantly ($P = 0.0417$)

associated with the frequency of imprinting of the nc886 DMR in infants (Fig. 4). In this data set, we performed likelihood ratio tests (LRT) to adjust for available confounding factors including Aflatoxin exposure, a common foodborne toxin in Africa, and gender. We found that neither variable was a significant predictor of imprinting at nc886 and after adjusting for these factors, season of conception significance slightly improves ($P = 0.03902$). Thus, season of conception, and by association maternal nutrition, may contribute to "tuning" the likelihood that an individual has imprinting at the nc886 DMR. Furthermore, recent analysis of a Scottish birth cohort identified differential DNA methylation of nc886 in response to folate supplementation during pregnancy (50).

The same hierarchical clustering analysis was performed on data from Markunas et al.(51) for DNA methylation of the nc886 DMR in children born to 855 mothers of various ages (*SI Appendix,* Figs. S7C and S7D). We found that children born to mothers younger than 20 years old had significantly less imprinting than those born to older mothers (Fig. 4; $P = 0.0167$). After adjustment for gender, maternal age is still significantly associated with imprinting of the nc886 DMR (LRT $P = 0.0281$). Our data suggest that teenage pregnancy is associated with a decreased likelihood that a child will have imprinting at the nc886 DMR.

We replotted data from van Dijk et al.(52) where HM450 analysis was performed on blood spots obtained from children at birth. This study identified lack of imprinting at nc886 measured at birth as being strongly associated with increased body mass index (BMI) in children assessed at five years of age. Hierarchical clustering of the nc886 DMR was used to define imprinted, not imprinted, and inconclusive groups, and BMI was analyzed as a continuous variable (*SI Appendix,* Figs. S7E and S7F). We confirmed that imprinting of the nc886 DMR was associated with lower BMI in children when analyzed as discrete quartiles (Fig. 4) and as a

continuous variable (Chi-squared $P = 1.75\times10^{-3}$ and logistic regression $P = 1.94\times10^{-3}$, respectively). We conclude that imprinting of the nc886 DMR is associated with a lower childhood BMI. Collectively, these three data sets demonstrate that maternal environment is linked to the establishment of imprinting across the nc886 DMR in children and that this epigenetic mark potentially impacts human phenotypes later in life.

Lastly, we calculated the percentage of each population, across all samples, that demonstrated imprinting (Fig. 4). When examining all samples from the BMI data set, we found that 76% of individuals were imprinted and 24% were not imprinted at the nc886 DMR, similar to the maternal age study and consistent with our previous findings where 75% of individuals were imprinted and 25% were not (23). In the smallest data set, which examined maternal nutrition in Gambia, we found 82% of the population with imprinting. Whether this is due to environmental differences or discrepancies in the genetic makeup of different populations needs to be investigated.

**Discussion**

We conclude that the nc886 DMR, when imprinted, is maternally methylated, given our evidence that this region behaves as a contiguous DMR that is not methylated in sperm (Fig 1H). The fact that the entire 1.9-kb nc886 DMR is subject to tunable polymorphic imprinting has been largely overlooked, emphasizing the need to scrutinize data obtained from the HM450 platform for possible dichotomization. We find that some of the polymorphic nature of the imprinting can be explained by local genetic makeup, as measured using an A/C SNP in the centromeric CTCF site. The mechanism by which this polymorphism alters the likelihood of imprinting is not completely clear but is possibly associated with chromatin conformation variation triggered by

CTCF binding, given reports that CTCF binding sites are involved in genomic imprinting (53-55). This possibility is supported by the observation that mutations in CTCF binding sites in the XIST promoter alter CTCF binding efficiency and choice of X chromosome inactivation (48). Additionally, the presence of a genetic influence is supported by the fact that monozygotic twins are found to be more concordant in DNA methylation at the nc886 DMR than dizygotic twins (35). However, Van Baak et al. concluded that DNA methylation at nc886 could not be explained by genetics alone and classified it as a region of "epigenetic supersimilarity" (35). While our data provide high resolution analysis of local DNA methylation in 31 individuals, a very recent paper by Zink et al. (37) found no genetic polymorphisms associated with imprinting of nc886 in an Icelandic population, suggesting a need for further genotyping and concurrent DNA methylation analysis in individuals from different populations and environmental contexts.

We find that the frequency of imprinting of the nc886 DMR in children is also associated with the mothers' age and season of conception indicating parental environment as potentially another mechanism for tuning the likelihood that the nc886 DMR will be imprinted (6, 51). Furthermore, the frequency of imprinting of the nc886 DMR at birth is maintained and is directly associated with the BMI of children at the age of five (52). Collectively, these results suggest that both genetic and environmental factors may affect the establishment of imprinting of a DMR, which is closely associated with human physiology (Fig. 5). Unfortunately, there are no known SNPs in nc886 which would allow us to examine tissues for allele specific expression. However, we have shown that DNA methylation can silence transcription of nc886, making it very likely that differential methylation of this region is functionally important (23, 34).

A further understanding of causality from all of these observations will require unraveling potential biological roles of nc886 or the nc886 DMR. nc886 has been variously

described as a regulator of the double-stranded RNA dependent protein kinase R (PKR) (23, 56-58) and dicer (59). It was also suggested as a possible tumor suppressor (60) or oncogene (61). Alternatively, some other genetic or epigenetic factors within or around the 1.9- kb DMR sequence may be drivers of phenotypic effects associated with nc886. Thus far, EWAS studies have primarily focused on variations in DNA methylation levels without considering the essential role of imprinting in human development as we have done here.

**FIGURES AND LEGENDS**

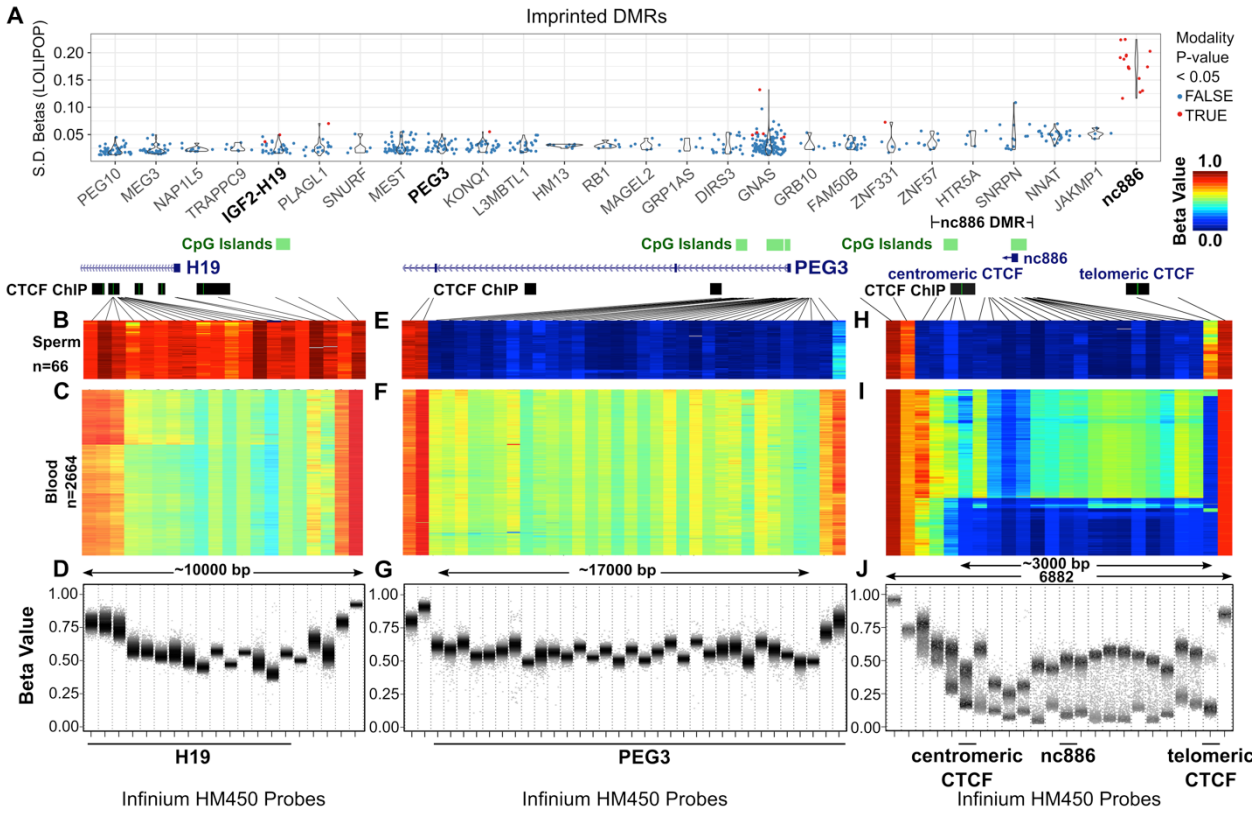


**Fig. 1. The nc886 DMR displays maternal polymorphic imprinting.** A) Populational variance for beta values in known imprinting control regions in peripheral blood. Each dot corresponds to one CpG interrogated by a HM450 probe. Populational standard deviations (Y-axis) are plotted for CpGs located in known imprinting DMRs and nc886 (X-axis). While all imprinting DMRs were examined, only imprinting control regions with more than five probes are included in the display and sorted by the mean standard deviation in the population. Imprinting control regions were retrieved from an earlier study (38). Probes with multi-modal distribution of beta values in the population are colored red. DNA methylation data was retrieved from the LOLIPOP study (39). B,E,H) DNA methylation data from the Infinium HM450 BeadChip platform in 66 sperm samples (GSE47627, GSE64096) for a paternally methylated DMR for the H19 gene (B), a maternally methylated DMR for the PEG3 gene (E) and the nc886 DMR (H). Black lines point to the genomic location of each probe relative to the respective DMR. C,F,I) Heat map of HM450 beta values from peripheral blood for H19 (C), PEG3 (F) and nc886 (I); data obtained from Wahl et al. 2017. D,G,J) Beta values of each sample are plotted for Infinium HM450 probes for H19 (D) PEG3 (G), and nc886 (J). HM450 probe IDs can be found in *SI Appendix*, Table 1. Only probes with sequence mapped optimally are included.

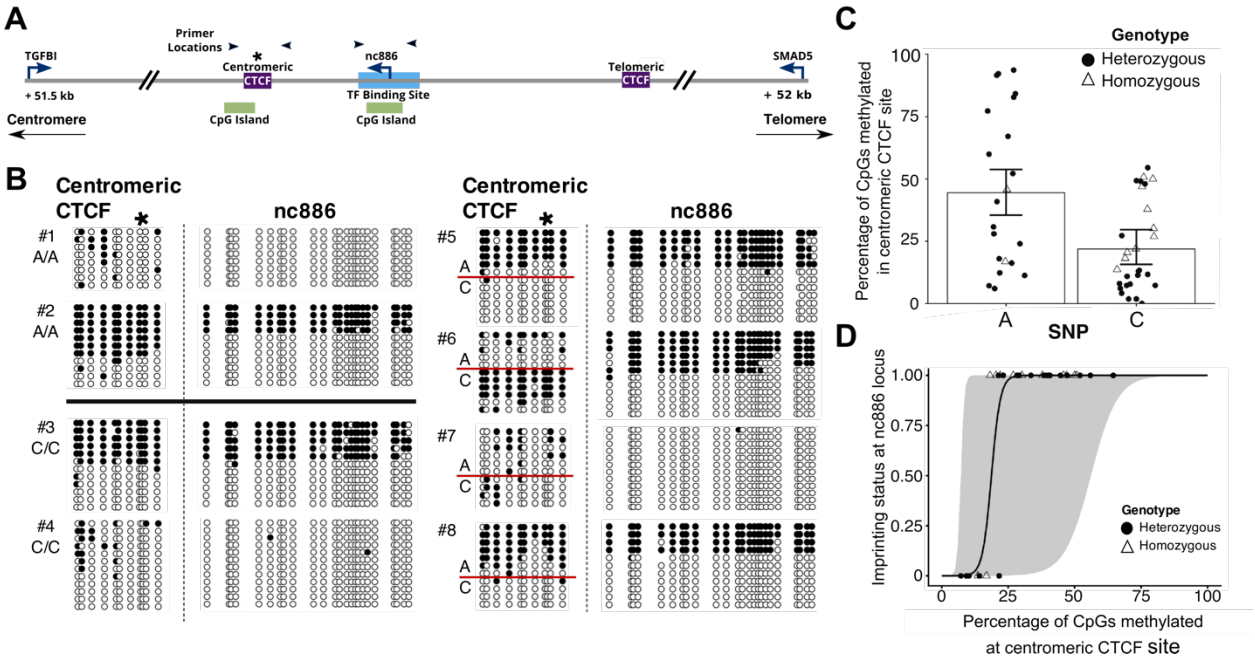


**Fig. 2. An A polymorphism (rs2346018) is associated with higher local DNA methylation density at the centromeric CTCF site, which has a strong positive correlation to DNA methylation at nc886. A)** Diagram of the nc886 DMR. Black arrows indicate primer locations for bisulfite sequencing in Figs. 2, 3, and *SI Appendix,* S6. **B)** Bisulfite conversion and clonal sequencing of genomic DNA from white blood cells of eight individuals at the centromeric CTCF site and nc886. The asterisk (*) indicates SNP location. **C)** The percentage of CpGs methylated for each allele at the centromeric CTCF site in homozygotes and heterozygotes ($n$ = 31; $P$ < 1e–9). **D)** DNA methylation of the centromeric CTCF plotted against DNA methylation of nc886 as a discrete variable, with 0 being biallelically unmethylated and 1 being monoallelically methylated ($n$ = 31; $P$ = 4.7e–7). Gray ribbon represents a 95% confidence interval.

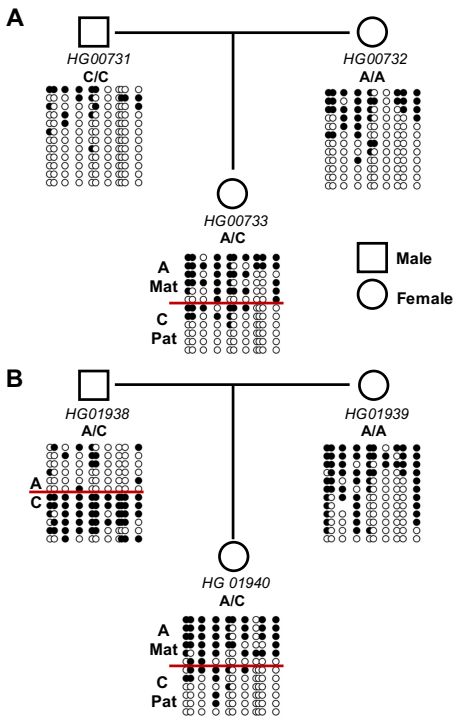


**Fig. 3. Maternal imprinting of the nc886 region extends to the centromeric CTCF site.** Locus-specific bisulfite sequencing was performed on genomic DNA at the centromeric CTCF site. Genomic DNA was isolated from benign lymphoblastoid cell lines derived from parent-offspring trios (mother, father, child) of disease-free individuals, where SNP status can be determined from the sequence. Two representative parent-offspring trios are shown. Clones are sorted based first on SNP status and then on DNA methylation.

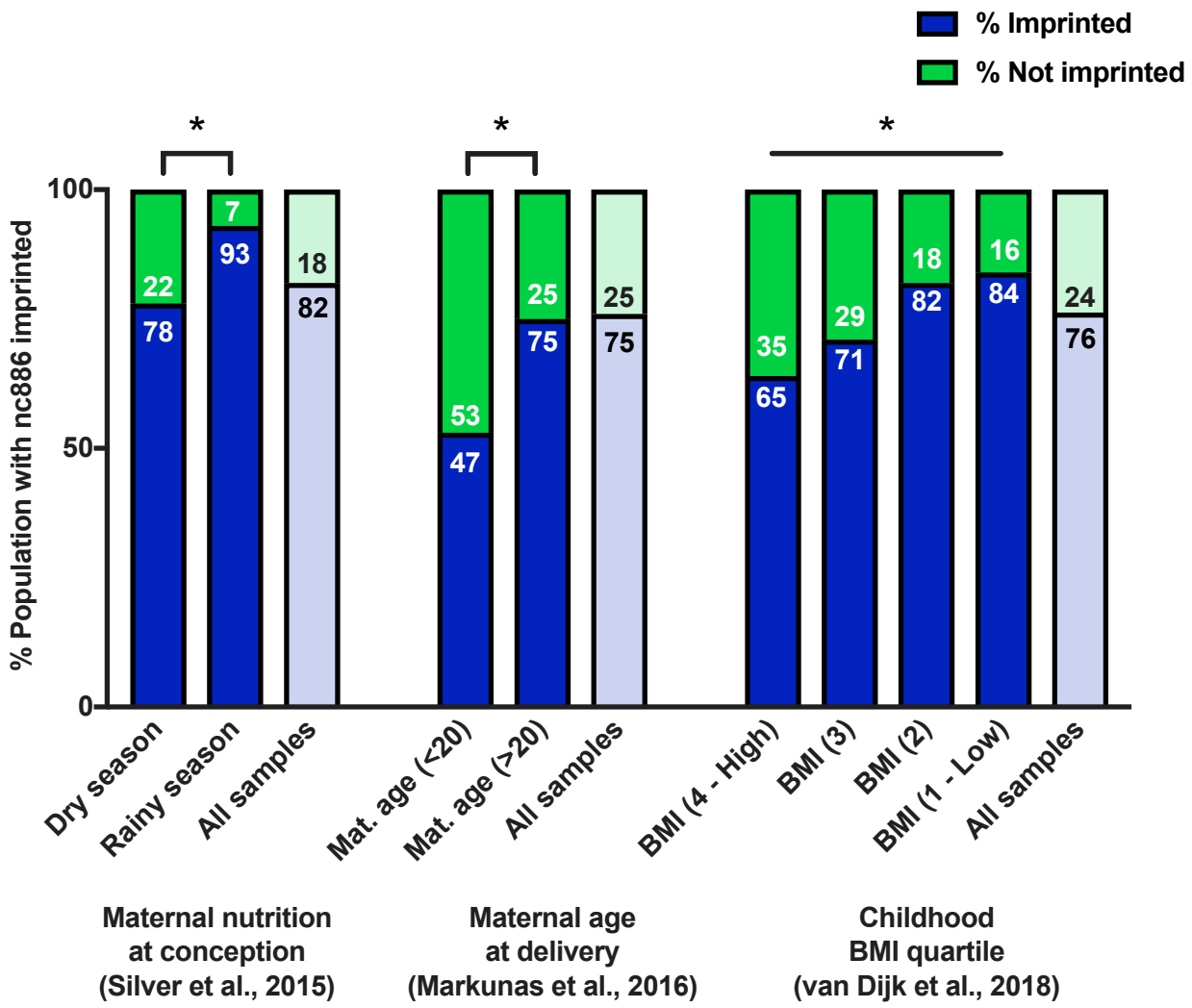


**Fig. 4. Maternal environment *in utero* impacts DNA methylation of nc886 in infants.** HM450 data for the nc886 DMR was analyzed from three independent studies of peripheral blood of infants at birth. Each group is separated by the percentage of the population that is imprinted vs. not imprinted based on hierarchical clustering. Individuals with inconclusive evidence of imprinting have been removed from this analysis.

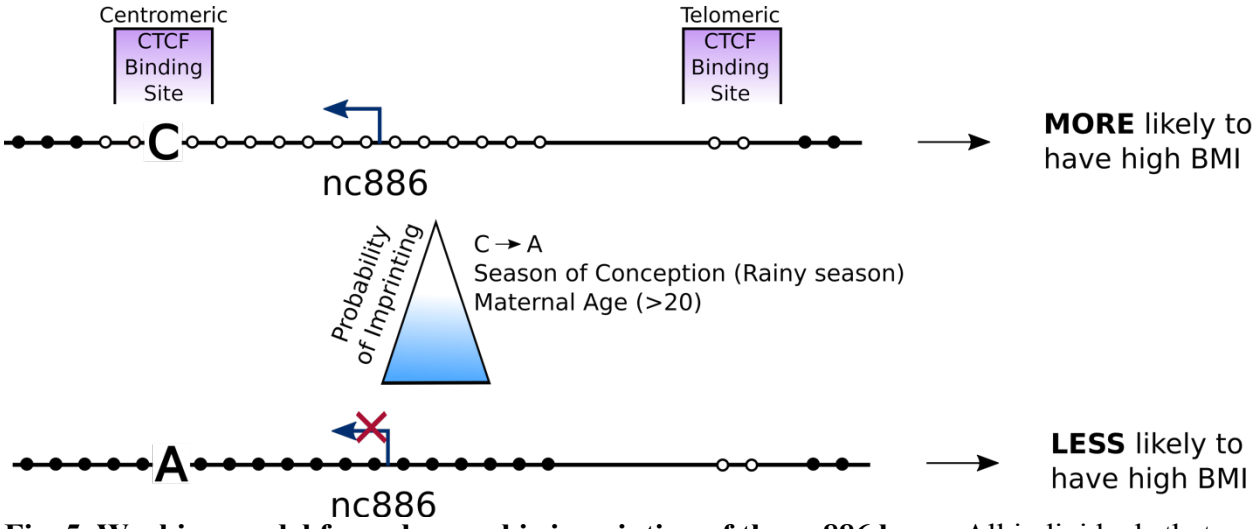


**Fig. 5. Working model for polymorphic imprinting of the nc886 locus.** All individuals that we examined, have the single paternally derived unmethylated allele at the nc886 DMR. In 75% of the human population the nc886 DMR is monoallelically methylated on the maternally derived allele (bottom), with the remaining 25% having no methylation on either allele (top). Here, we demonstrate that the presence of an A or C polymorphism in the centromeric-CTCF binding site is associated with local DNA methylation density in the nc886 DMR. We hypothesize that the presence of the A allele or increased DNA methylation density reduces CTCF binding, allowing DNA methylation to spread into the DMR. We reanalyzed data from Silver *et al.*, and Markunas *et al.*, and found that maternal age and season of conception contribute to the likelihood that a child will have imprinting at the nc886 DMR (6, 51). Therefore, we propose that a combination of maternal factors shifts the balance between imprinted and not-imprinted in children, with downstream phenotypic consequences such as BMI.

## METHODS

DNA methylation analysis of Illumina HumanMethylation450 data

Normalized beta values were downloaded from GEO, with accession GSE47627 and GES64096 for sperm and GSE55763 for genomic DNA from peripheral blood of 2,664 individuals assayed on the Infinium HM450 platform (LOLIPOP data). Probes falling in the region chr5:135413937-135419936 (Genome Build GRCh37) were extracted and their signals were visualized using the R package, SeSAMe (10). Genomic features were extracted from the UCSC genome browser. Probes were annotated with potential cross-hybridization issues and SNP-bias were excluded according to an earlier study (10, 62). This analysis was also performed on matched normal tissues from the TCGA database, as shown in *SI Appendix,* Fig. S2.

Locus -specific bisulfite sequencing analysis of the centromeric CTCF site and nc886

Genomic DNA isolated from white blood cells of disease-free individuals was obtained from the Van Andel Research Institute's Pathology and Biorepository Core. For trio analysis, the following DNA samples were obtained from the NHGRI Sample Repository for Human Genetic Research at the Coriell Institute for Medical Research: HG00731, HG00732, HG00733, HG01938, HG01939, and HG01940. Bisulfite conversion and clean-up of 2 μg of genomic DNA was performed using the EZ DNA Methylation Kit (Zymo Research) according to the manufacturer's protocol. Locus-specific PCR was performed using primers specific for bisulfite-converted DNA and PCR products cloned using the pGEM-T Easy vector and NEB 5-alpha-competent *E. coli* (New England Biolabs Inc). Colonies were screened for positive inserts by PCR and sequencing performed using the M13 promoter.

Statistical Analysis

**Determining whether SNP status is associated with DNA methylation in the centromeric CTCF region**

Beta binomial mixed-effects regression, weighted by the number of CpG sites in a region (10 for SNP A, 11 for C) was used to model data obtained from bisulfite sequencing and PCR via the R package (v3.4.3, https://www.r-project.org/) glmmTMB) (63). A random intercept for each individual was used to account for the paired heterozygous and homozygous measures; a second random intercept was included to account for DNA methylation of the nc886 region as determined by the presence of at least 10% of clones being completely methylated. Beta-binomial was chosen because it is well-suited for modeling the binary methylated/unmethylated status of each CpG in the region, but the probability of a given site being methylated increases with each additional site methylated in the region. A likelihood ratio test confirmed weighted beta-binomial was a better fit for these data than a weighted binomial, and both Akaike information criterion (AIC) and Bayesian information criterion (BIC) showed this regression was better than negative binomial regression.

**Determining the correlation between DNA methylation at the centromeric CTCF and nc886 regions**

The percentage of DNA methylation was calculated for each region analyzed by bisulfite conversion and sequencing and the data were analyzed via logistic regression. Bisulfite sequencing was used to estimate DNA methylation in both the centromeric CTCF and nc886 regions. Logistic regression was used to determine whether the probability of nc886 being methylated changed as the percentage of methylation of the CTCF region increased. The reported p-value was calculated using a likelihood ratio test. The mean percent methylation of the centromeric CTCF site was estimated using a mixed-effects negative binomial regression via the R package glmmTMB (63). Data were plotted as the mean change in probability per percent increase of methylation in the centromeric CTCF, with a 95% confidence band.

Population analysis of DNA methylation at nc886 from published data sets

We downloaded from GEO Infinium HM450 data from genomic DNA of infants in studies of season of conception in Gambia (6) and maternal age at delivery (51) with accession numbers GSE59592 and GSE82273 respectively. For the childhood BMI study (52), the GEO Infinium HM450 data (accession GSE103657) was not optimal due to technical effects introduced by batch correction. Instead, the raw array data was renormalized consistent with the methods of the published study and BMI measurement data were accessed through the authors of van Dijk et al (2018).

Based on bisulfite sequencing analysis (*SI Appendix,* Fig. S6) which provides single base resolution of individual DNA strands for DNA methylation and previous studies, we concluded that there are only two possibilities for DNA methylation in this region, near either 0% DNA methylation (biallelically unmethylated) or near 50% DNA methylation (monoallelically methylated) (6, 23, 24, 64). However, when data were plotted it was clear that not all individuals exhibited this expected dichotomy of 50% methylated or not methylated (*SI Appendix,* Fig. S7). Upon further investigation, we found the individuals who did not agree with this dichotomy comprised a third cluster based on hierarchical clustering via Manhattan distances. Therefore, we identified three clusters of individuals which we called imprinted, not imprinted, or inconclusive based on partial methylation readouts throughout the region (*SI Appendix,* Fig. S7). To remain consistent with previously observed data we restricted our analyses to focus on the patients not in this third, inconsistent cluster. Results without this restriction are reported in the supplement.

Treating imprinting as the response variable, we applied logistic regression using the glm function in R. The season of conception, maternal age (dichotomized by teenage pregnancy), and childhood BMI Z-scores were used as the regressors of the three data sets, respectively. P-values for the coefficients were obtained using a Wald test.

**REFERENCES**


1. Flanagan JM (2015) Epigenome-wide association studies (EWAS): past, present, and future. *Methods Mol Biol* 1238:51-63.
2. Lappalainen T & Greally JM (2017) Associating cellular epigenetic models with human phenotypes. *Nat Rev Genet* 18(7):441-451.
3. Chandler LA, Ghazi H, Jones PA, Boukamp P, & Fusenig NE (1987) Allele-specific methylation of the human c-Ha-ras-1 gene. *Cell* 50(5):711-717.
4. Hitchins MP*, et al.* (2011) Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5'UTR. *Cancer Cell* 20(2):200-213.
5. Kerkel K*, et al.* (2008) Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat Genet* 40(7):904-908.
6. Silver MJ*, et al.* (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol* 16:118.
7. Dolinoy DC, Das R, Weidman JR, & Jirtle RL (2007) Metastable epialleles, imprinting, and the fetal origins of adult diseases. *Pediatr Res* 61(5 Pt 2):30R-37R.
8. Jirtle RL & Skinner MK (2007) Environmental epigenomics and disease susceptibility. *Nat Rev Genet* 8(4):253-262.
9. Do C*, et al.* (2016) Mechanisms and Disease Associations of Haplotype-Dependent Allele-Specific DNA Methylation. *Am J Hum Genet* 98(5):934-955.
10. Zhou W, Triche TJ, Jr., Laird PW, & Shen H (2018) SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res*.
11. Soubry A, Hoyo C, Jirtle RL, & Murphy SK (2014) A paternal environmental legacy: evidence for epigenetic inheritance through the male germ line. *Bioessays* 36(4):359-371.
12. Lee HS (2015) Impact of Maternal Diet on the Epigenome during In Utero Life and the Developmental Programming of Diseases in Childhood and Adulthood. *Nutrients* 7(11):9492-9507.
13. Chango A & Pogribny IP (2015) Considering maternal dietary modulators for epigenetic regulation and programming of the fetal epigenome. *Nutrients* 7(4):2748-2770.
14. Joubert BR*, et al.* (2016) DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet* 98(4):680-696.
15. Joubert BR*, et al.* (2016) Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat Commun* 7:10577.
16. Babak T*, et al.* (2015) Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nat Genet* 47(5):544-549.
17. Prickett AR & Oakey RJ (2012) A survey of tissue-specific genomic imprinting in mammals. *Mol Genet Genomics* 287(8):621-630.
18. Tycko B (2010) Allele-specific DNA methylation: beyond imprinting. *Hum Mol Genet* 19(R2):R210-220.
19. Bartolomei MS (2009) Genomic imprinting: employing and avoiding epigenetic processes. *Genes Dev* 23(18):2124-2133.
20. Hitchins MP & Moore GE (2002) Genomic imprinting in fetal growth and development. *Expert Rev Mol Med* 4(11):1-19.
21. Butler MG (2009) Genomic imprinting disorders in humans: a mini-review. *J Assist Reprod Genet* 26(9-10):477-486.

22. Eggermann T*, et al.* (2015) Imprinting disorders: a group of congenital disorders with overlapping patterns of molecular changes affecting imprinted loci. *Clin Epigenetics* 7:123.
23. Treppendahl MB*, et al.* (2012) Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML. *Blood* 119(1):206-216.
24. Romanelli V*, et al.* (2014) Variable maternal methylation overlapping the nc886/vtRNA2-1 locus is locked between hypermethylated repeats and is frequently altered in cancer. *Epigenetics* 9(5):783-790.
25. Hanna CW*, et al.* (2016) Pervasive polymorphic imprinted methylation in the human placenta. *Genome Res* 26(6):756-767.
26. Jinno Y*, et al.* (1994) Mosaic and polymorphic imprinting of the WT1 gene in humans. *Nat Genet* 6(3):305-309.
27. Bartolomei MS & Ferguson-Smith AC (2011) Mammalian genomic imprinting. *Cold Spring Harb Perspect Biol* 3(7).
28. Reik W & Walter J (2001) Genomic imprinting: parental influence on the genome. *Nat Rev Genet* 2(1):21-32.
29. Inoue A, Jiang L, Lu F, Suzuki T, & Zhang Y (2017) Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* 547(7664):419-424.
30. Haig D (1997) Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proc Biol Sci* 264(1388):1657-1662.
31. Susiarjo M, Sasson I, Mesaros C, & Bartolomei MS (2013) Bisphenol a exposure disrupts genomic imprinting in the mouse. *PLoS Genet* 9(4):e1003401.
32. Kappil M, Lambertini L, & Chen J (2015) Environmental Influences on Genomic Imprinting. *Curr Environ Health Rep* 2(2):155-162.
33. Tobi EW*, et al.* (2012) Prenatal famine and genetic variation are independently and additively associated with DNA methylation at regulatory loci within IGF2/H19. *PLoS One* 7(5):e37933.
34. Helbo AS, Lay FD, Jones PA, Liang G, & Gronbaek K (2017) Nucleosome positioning and NDR structure at RNA polymerase III promoters. *Sci Rep* 7:41947.
35. Van Baak TE*, et al.* (2018) Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol* 19(1):2.
36. Do C*, et al.* (2017) Genetic-epigenetic interactions in cis: a major focus in the post-GWAS era. *Genome Biol* 18(1):120.
37. Zink F*, et al.* (2018) Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. *Nat Genet*.
38. Court F*, et al.* (2014) Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res* 24(4):554-569.
39. Wahl S*, et al.* (2017) Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 541(7635):81-86.
40. Krausz C*, et al.* (2012) Novel insights into DNA methylation features in spermatozoa: stability and peculiarities. *PLoS One* 7(10):e44479.
41. Jenkins TG*, et al.* (2015) Intra-sample heterogeneity of sperm DNA methylation. *Mol Hum Reprod* 21(4):313-319.
42. Paliwal A*, et al.* (2013) Comparative anatomy of chromosomal domains with imprinted and non-imprinted allele-specific DNA methylation. *PLoS Genet* 9(8):e1003622.

43. Grundberg E*, et al.* (2013) Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am J Hum Genet* 93(5):876-890.
44. Maurano MT, Wang H, Kutyavin T, & Stamatoyannopoulos JA (2012) Widespread site-dependent buffering of human regulatory polymorphism. *PLoS Genet* 8(3):e1002599.
45. Lewis A & Reik W (2006) How imprinting centres work. *Cytogenet Genome Res* 113(1-4):81-89.
46. Coetzee SG, Coetzee GA, & Hazelett DJ (2015) motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. *Bioinformatics* 31(23):3847-3849.
47. Prickett AR*, et al.* (2013) Genome-wide and parental allele-specific analysis of CTCF and cohesin DNA binding in mouse brain reveals a tissue-specific binding pattern and an association with imprinted differentially methylated regions. *Genome Res* 23(10):1624-1635.
48. Pugacheva EM*, et al.* (2005) Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. *Hum Mol Genet* 14(7):953-965.
49. Wang H*, et al.* (2012) Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res* 22(9):1680-1688.
50. Richmond RC*, et al.* (2018) The long-term impact of folic acid in pregnancy on offspring DNA methylation: follow-up of the Aberdeen Folic Acid Supplementation Trial (AFAST). *Int J Epidemiol*.
51. Markunas CA*, et al.* (2016) Maternal Age at Delivery Is Associated with an Epigenetic Signature in Both Newborns and Adults. *PLoS One* 11(7):e0156361.
52. van Dijk SJ*, et al.* (2017) DNA methylation in blood from neonatal screening cards and the association with BMI and insulin sensitivity in early childhood. *Int J Obes (Lond)*.
53. Szabo PE, Tang SH, Silva FJ, Tsark WM, & Mann JR (2004) Role of CTCF binding sites in the Igf2/H19 imprinting control region. *Mol Cell Biol* 24(11):4791-4800.
54. Lin S, Ferguson-Smith AC, Schultz RM, & Bartolomei MS (2011) Nonallelic transcriptional roles of CTCF and cohesins at imprinted loci. *Mol Cell Biol* 31(15):3094-3104.
55. Bell AC & Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 405(6785):482-485.
56. Lee K*, et al.* (2011) Precursor miR-886, a novel noncoding RNA repressed in cancer, associates with PKR and modulates its activity. *RNA* 17(6):1076-1089.
57. Jeon SH*, et al.* (2012) Characterization of the direct physical interaction of nc886, a cellular non-coding RNA, and PKR. *FEBS Lett* 586(19):3477-3484.
58. Calderon BM & Conn GL (2017) Human noncoding RNA 886 (nc886) adopts two structurally distinct conformers that are functionally opposing regulators of PKR. *RNA* 23(4):557-566.
59. Ahn JH*, et al.* (2018) nc886 is induced by TGF-beta and suppresses the microRNA pathway in ovarian cancer. *Nat Commun* 9(1):1166.
60. Fort RS*, et al.* (2018) Nc886 is epigenetically repressed in prostate cancer and acts as a tumor suppressor through the inhibition of cell growth. *BMC Cancer* 18(1):127.
61. Lee EK*, et al.* (2016) nc886, a non-coding RNA and suppressor of PKR, exerts an oncogenic function in thyroid cancer. *Oncotarget* 7(46):75000-75012.

62. Zhou W, Laird PW, & Shen H (2017) Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res* 45(4):e22.
63. Magnusson A*, et al.* (2017) glmmTMB: Generalized Linear Mixed Models using Template Model Builder. R package version 0.1.3. https://github.com/glmmTMB.
64. Waterland RA & Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 23(15):5293-5300.
65. Zhou W*, et al.* (2018) DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nat Genet* 50(4):591-602.
66. Kulis M*, et al.* (2015) Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet* 47(7):746-756.
67. Adams D*, et al.* (2012) BLUEPRINT to decode the epigenetic signature written in blood. *Nat Biotechnol* 30(3):224-226.

**Acknowledgements:** We would like to thank the VARI Pathology and Biorepository Core for providing samples for bisulfite sequencing analysis. The VARI Bioinformatics and Biostatistics Core was instrumental in providing statistical support. We would like to thank Bev Muhlhausler for curating and allowing access to BMI data from the DOMInO study. Grant support for this study was provided by R35CA209859 from the National Cancer Institute to PAJ.

**Author Contributions:** BLC was responsible for experimental design and execution and for coordinating efforts between authors. WZ was primarily responsible for bioinformatics analysis of public datasets. ZM was responsible for statistics and corresponding text. AKD performed extensive bisulfite sequencing analysis. JR was responsible for analysis of BMI data. BLC and PAJ were responsible for writing the manuscript, with intellectual input and editing from all other authors.

**Conflict of Interest Statement:** PAJ is a paid consultant for Zymo Research. In this study, Zymo reagents were used to analyze locus specific DNA methylation patterns.

**Author Information:**
Correspondence and requests for materials should be addressed to Peter A. Jones, Peter.Jones@vai.org

## Supplemental Information

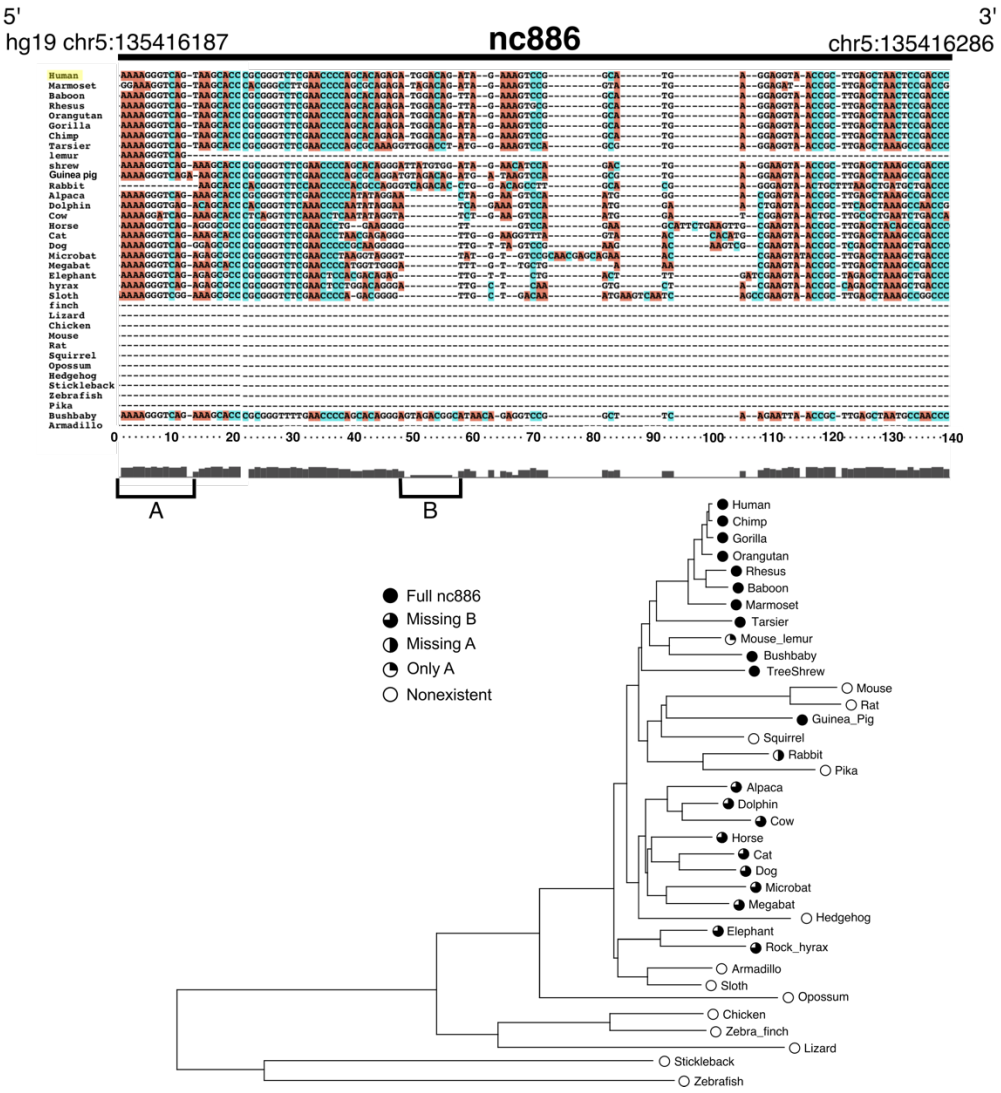


**Fig. S1. Species alignment for nc886.** Sequence alignment of the nc886 locus and a pruned tree from the UCSC 46-vertebrate species tree.

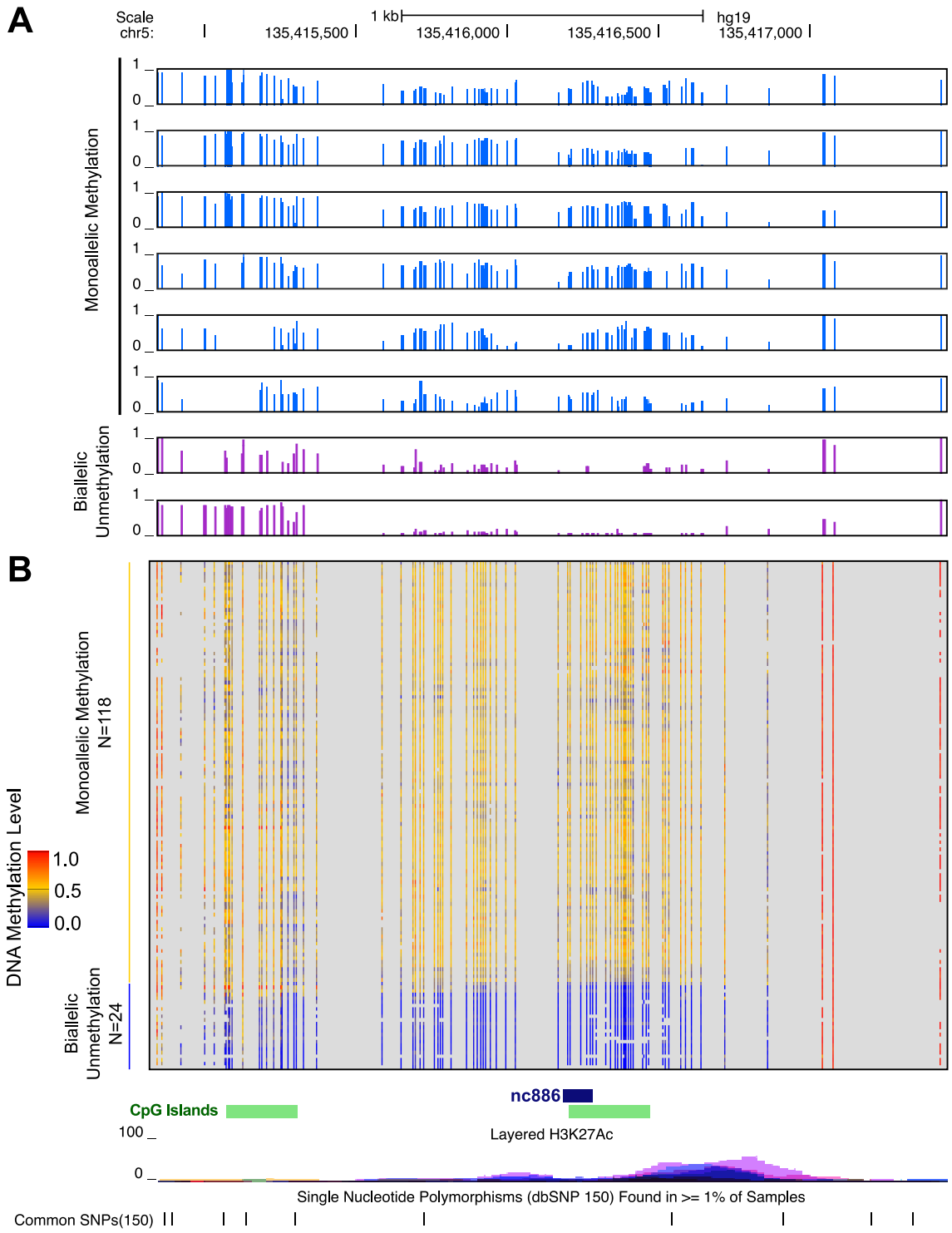


**Fig. S2. Whole genome bisulfite sequencing for the nc886 DMR.** The nc886 DMR was defined as 1979 bps using WGBS data for individuals from TCGA (65) (A) and BLUEPRINT (66, 67) (B) data sets.

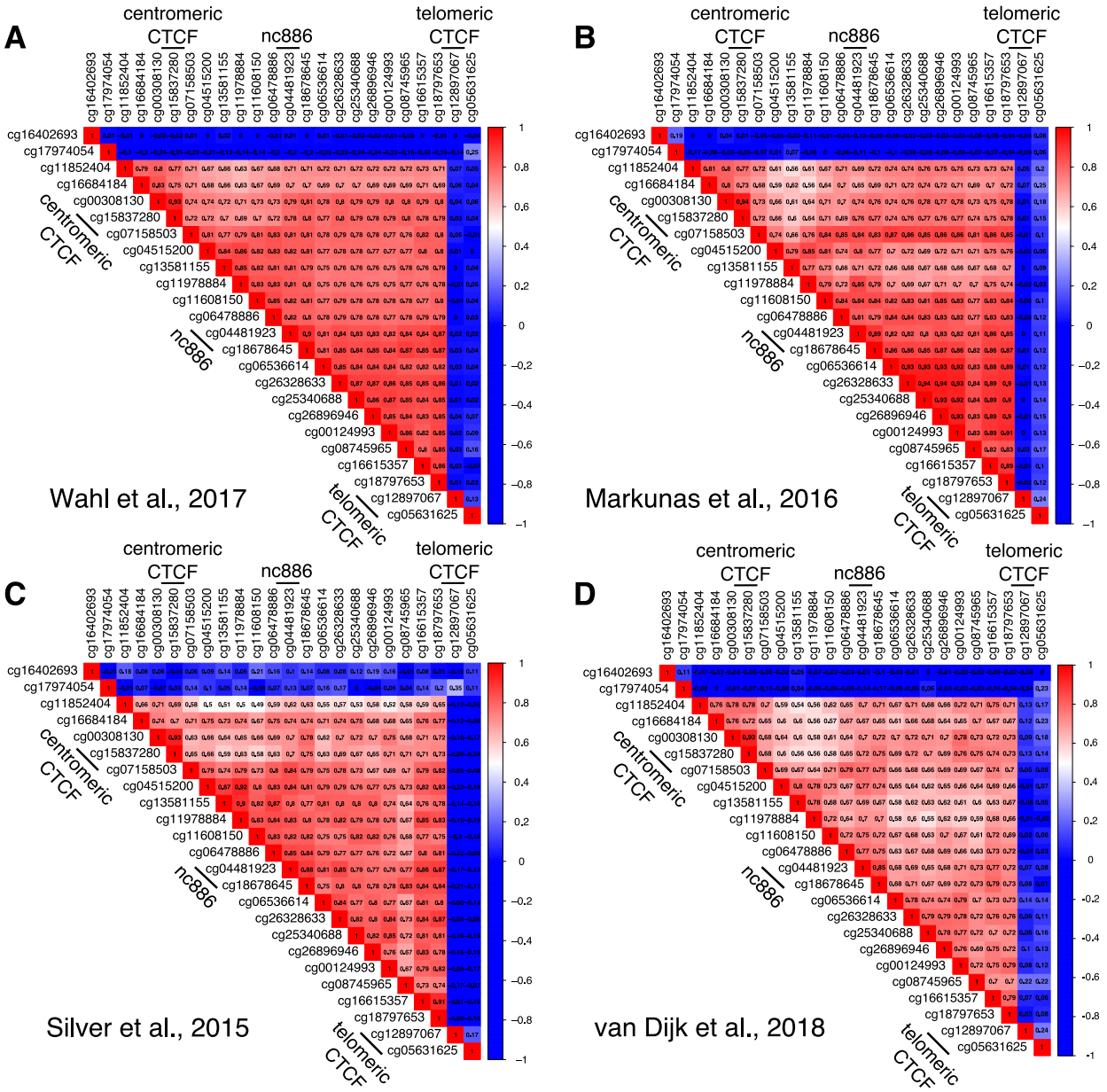


**Fig. S3. Correlation of beta values across the nc886 DMR.** Correlograms were generated based on beta values across the nc886 DMR and flanking invariable CpG sites. The data are from four independent data sets that were analyzed in this study.

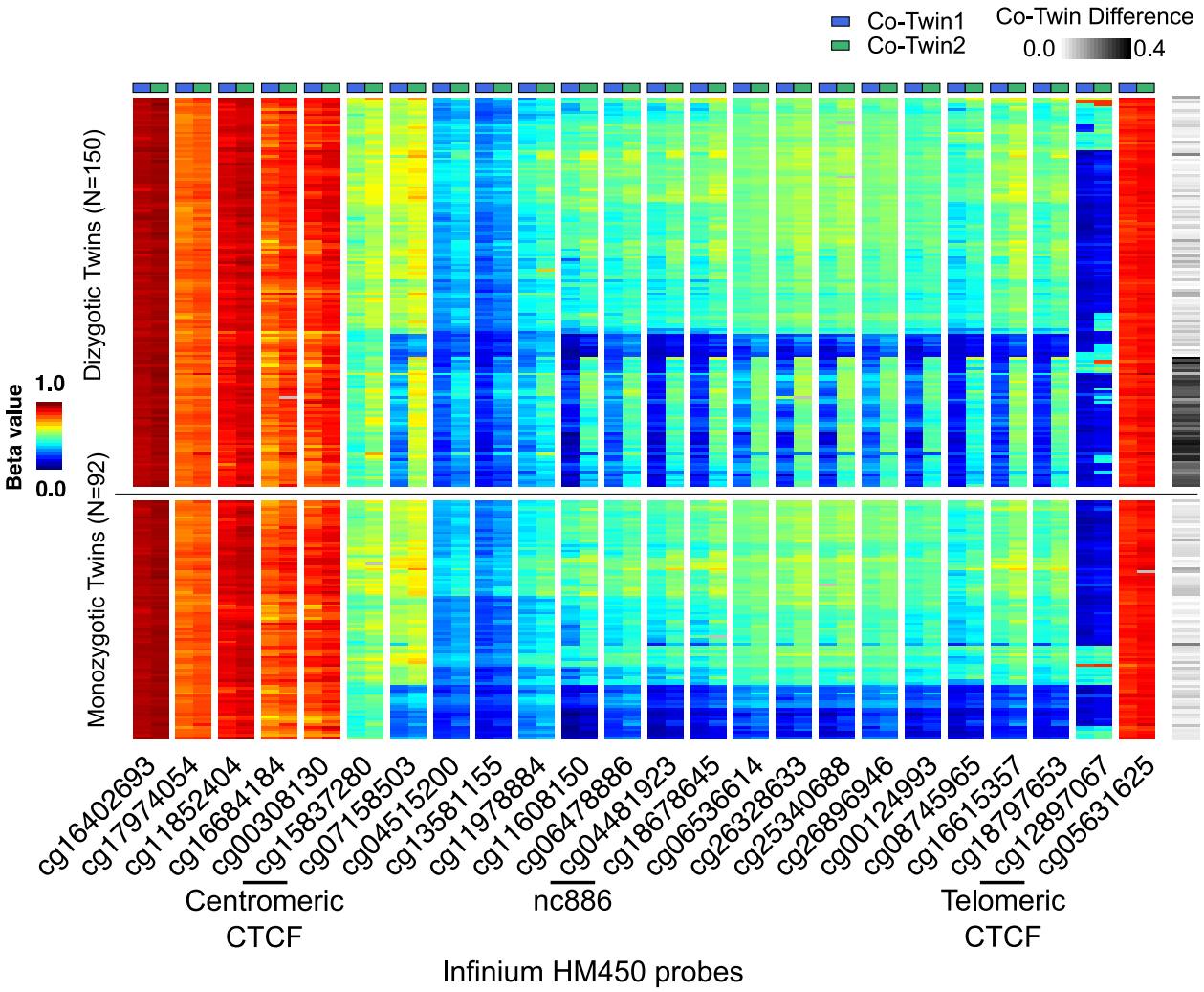


**Fig. S4. Monozygotic, but not dizygotic, twins are concordant for DNA methylation across the nc886 DMR.** HM450 data was analyzed from adipose tissue of monozygotic (bottom panel) and dizygotic (top panel) twins. Heat map displays beta values for co-Twin1 and co-Twin2 at each CpG probe across the DMR.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

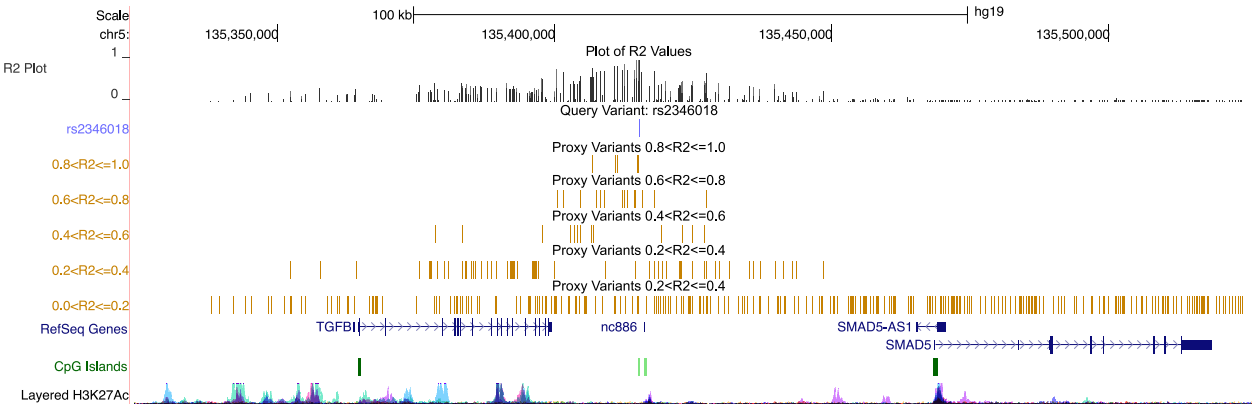


**Fig. S5. The local haplotype for nc886.** LDlink was used to calculate R2 values between variant rs2346018 and surrounding variants. R2 values are displayed for all populations from the 1000 genomes project. Local haplotypes range in size from 5 kb – 35 kb from rs2346018 depending on the population.

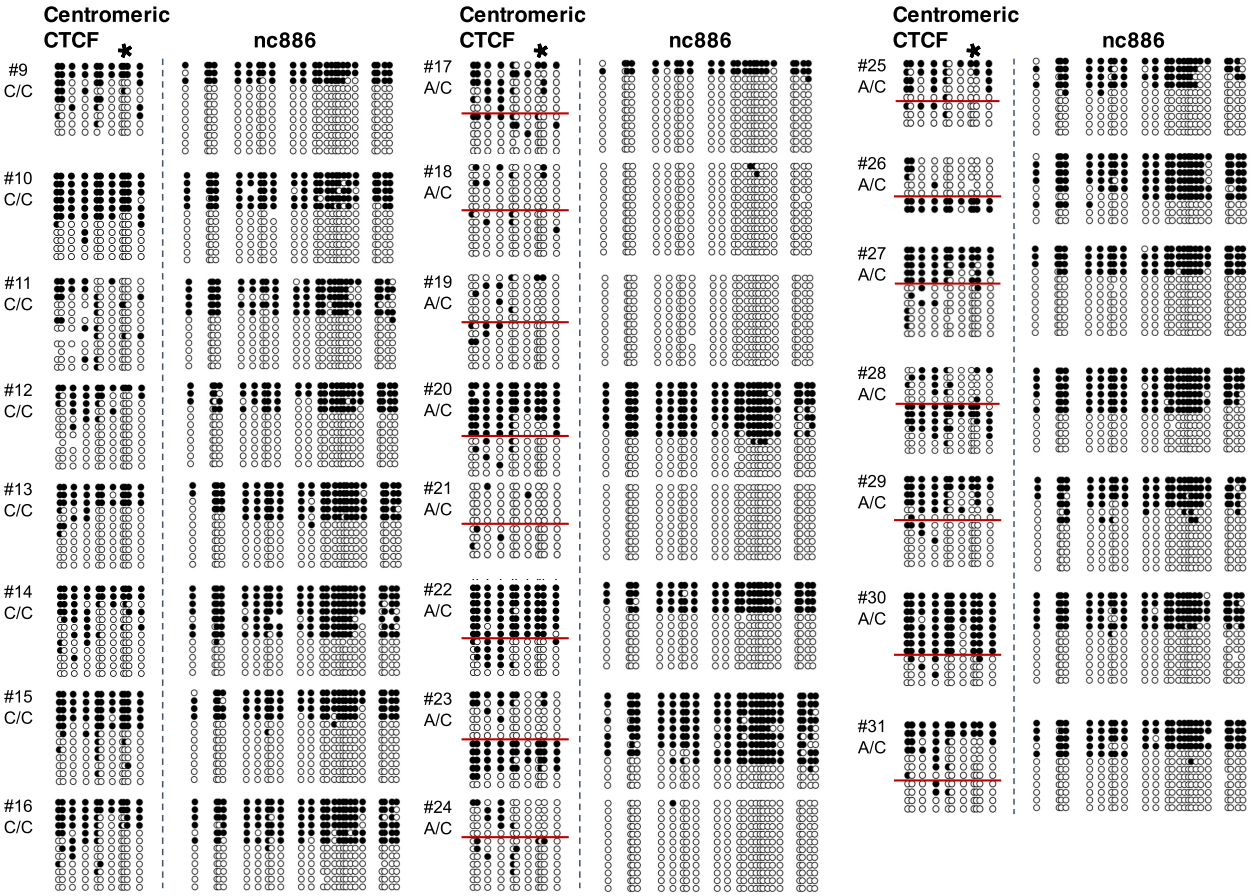


**Fig. S6. Bisulfite sequencing of centromeric CTCF and nc886.** Locus-specific bisulfite sequencing analysis for individuals not represented in Figs. 2. Clones for centromeric CTCF are not physically connected on an individual DNA strand. Clones for heterozygous individuals are sorted first by A/C and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

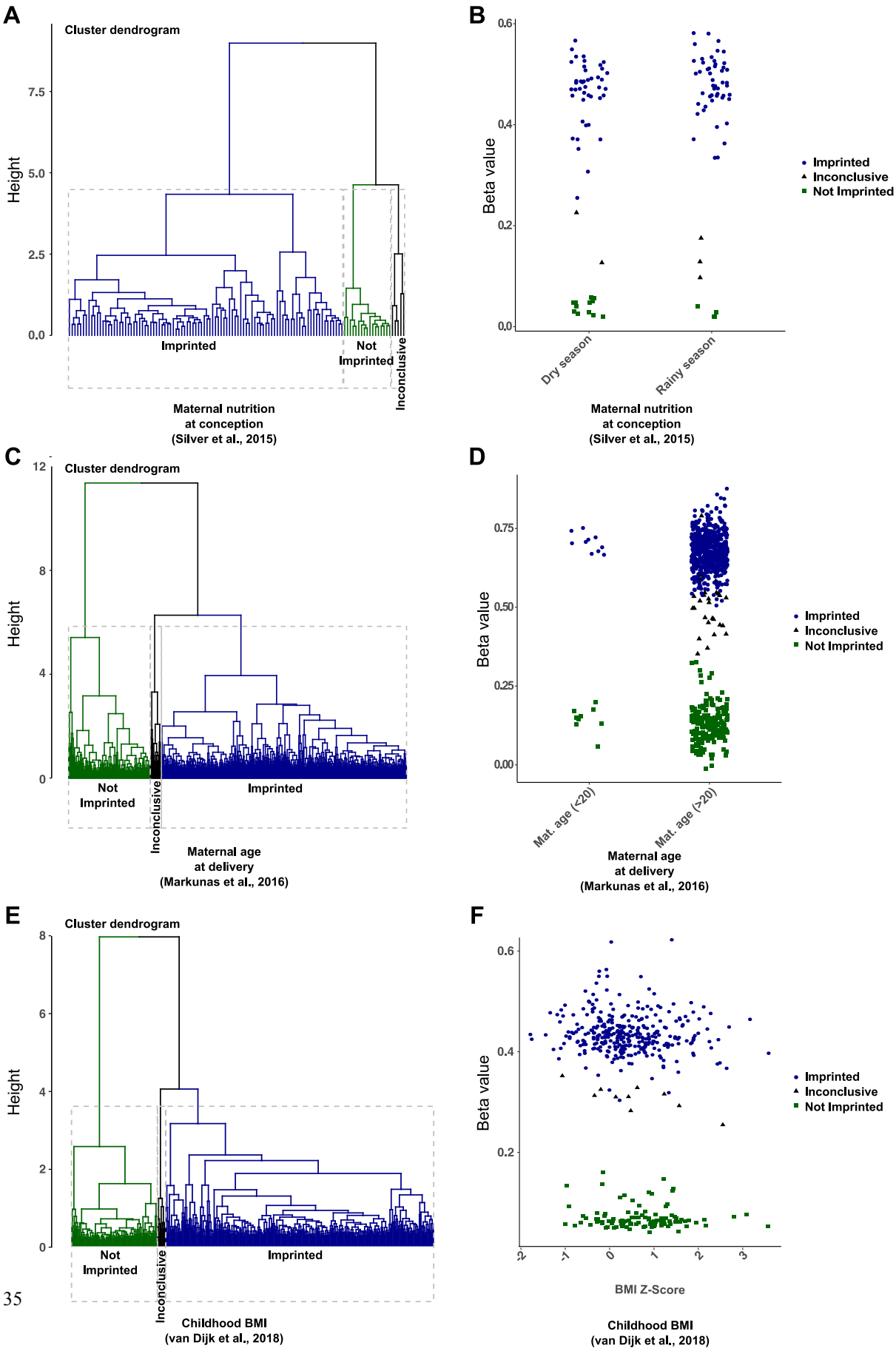
A
Cluster dendrogram
Height
7.5
5.0
2.5
0.0
Imprinted
Not Imprinted
Inconclusive
Maternal nutrition at conception (Silver et al., 2015)
B
Beta value
0.6
0.4
0.2
0.0
Dry season
Rainy season
Maternal nutrition at conception (Silver et al., 2015)
Imprinted
Inconclusive
Not Imprinted
C
Cluster dendrogram
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<br>rs2346019<br>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.