Transposable Elements: Targets for Early Nutritional Effects on Epigenetic Gene Regulation

Robert A. Waterland and Randy L. Jirtle*

Department of Radiation Oncology, Duke University Medical Center, Durham, North Carolina 27710

Received 7 January 2003/Returned for modification 14 April 2003/Accepted 8 May 2003

Early nutrition affects adult metabolism in humans and other mammals, potentially via persistent alterations in DNA methylation. With viable yellow agouti (A⁺) mice, which harbor a transposable element in the agouti gene, we tested the hypothesis that the metastable methylation status of specific transposable element insertion sites renders them epigenetically labile to early methyl donor nutrition. Our results show that dietary methyl supplementation of a/a dams with extra folic acid, vitamin B₁₂, choline, and betaine alter the phenotype of their A⁺/a offspring via increased CpG methylation at the A⁺ locus and that the epigenetic metastability which confers this lability is due to the A⁺ transposable element. These findings suggest that dietary supplementation, long presumed to be purely beneficial, may have unintended deleterious influences on the establishment of epigenic gene regulation in humans.

Human epidemiologic and animal model data indicate that susceptibility to adult-onset chronic disease is influenced by persistent adaptations to prenatal and early postnatal nutrition (1, 2, 13, 14); however, the specific biological mechanisms underlying such adaptations remain largely unknown. Cytosine methylation within CpG dinucleotides of DNA acts in concert with other chromatin modifications to heritably maintain specific genomic regions in a transcriptionally silent state (4). Genomic patterns of CpG methylation are reprogrammed in the early embryo and maintained thereafter (2). Because diet-derived methyl donors and cofactors are necessary for the synthesis of S-adenosylmethionine, required for CpG methylation (23), early nutrition may therefore influence adult phenotype via DNA methylation (26).

Accordingly, it is important to identify genomic regions that are likely targets for early nutritional influences on CpG methylation. Most regions of the adult mammalian genome exhibit little interindividual variability in tissue-specific CpG methylation levels. Conversely, CpG methylation is determined probabilistically at specific transposable element insertion sites in the mouse genome, causing cellular epigenetic mosaicism and individual phenotypic variability (19). Transposable elements (including retrotransposons and DNA transposons) are parasitic elements which are scattered throughout and constitute over 35% of the human genome (32). Most transposable elements in the mammalian genome are normally silenced by CpG methylation (32). The epigenetic state of a subset of transposable elements, however, is metastable and can affect regions encompassing neighboring genes. (19). We hypothesized that the epigenetic metastability of such regions renders them susceptible to nutritional influences during early development.

We tested this hypothesis in viable yellow agouti (A⁺) mice. The murine agouti gene encodes a paracrine signaling molecule that signals follicular melanocytes to switch from producing black eumelanin to yellow phaeomelanin. Transcription is initiated from a hair cycle-specific promoter in exon 2 of the agouti (A) allele (Fig. 1A). Transient agouti expression in hair follicles during a specific stage of hair growth results in a sub-apical yellow band on each hair, causing the brown (agouti) coat color of wild-type mice (8). The nonagouti (a) allele was caused by a loss-of-function mutation in A (5); a/a homozygotes are therefore black. The A⁺ allele (Fig. 1A) resulted from the insertion of an intracisternal A particle (IAP) retrotransposon into the 5’ end of the A allele (8). Ectopic agouti transcription is initiated from a cryptic promoter in the proximal end of the A⁺ IAP. CpG methylation in this region varies dramatically among individual A⁺ mice and is correlated inversely with ectopic agouti expression. This epigenetic variability causes a wide variation in individual coat color (Fig. 2A), adiposity, glucose tolerance, and tumor susceptibility among isogenic A⁺/a littermates (15).

Dietary methyl supplementation of a/a dams shifts the coat color distribution of their A⁺/a offspring (30). Because A⁺/a coat color correlates with A⁺ methylation status (15), it has been inferred that supplementation alters phenotype via A⁺ methylation (7). Nevertheless, no study has yet compared A⁺ methylation among the offspring of supplemented and unsupplemented dams (25). Therefore, to test our hypothesis that transposable elements are targets for early nutritional effects on epigenic gene regulation, we had to determine if CpG methylation plays a role in diet-induced phenotypic alterations in A⁺/a mice.

Agouti pseudoexon 1A (PS1A) was formed when a 4.1-kb genomic region containing exon 1A underwent duplication and inversion (Fig. 1A) (6). In mice that carry the light-bellied agouti (A⁺) allele, exon 1A is oriented properly with respect to the agouti gene and drives agouti expression (and yellow pigmentation) throughout the hair growth cycle in ventral follicles. The orientation of the duplication is reversed in mice carrying the A allele (Fig. 1A). In these animals, exon 1A points away from agouti, causing a loss of ventral follicle-specific agouti expression (6). Early genetic analyses concluded...
that the $A^v$ IAP is located within agouti exon 1A (8); however, this conclusion has not been reevaluated since the subsequent characterization of the inverted repeat in the region (6).

In this study, after determining the actual location of the $A^v$ IAP, we showed that dietary methyl donor supplementation of a/a dams alters $A^v/a$ offspring phenotype by increasing CpG methylation at the agouti locus. Furthermore, we demonstrated that the epigenetic metastability which confers this lability is due to the $A^v$ IAP.

MATERIALS AND METHODS

Animals and diets. $A^v$ mice were obtained from the colony at the Oak Ridge National Laboratory (29). The $A^v$ mutation arose spontaneously in the C3H/HeJ strain. Mice carrying the mutation were backcrossed with C57BL/6J mice for one to three generations before being propagated by sibling mating. These animals therefore include 6.25% to 25% of the C3H/HeJ genome and 75% to 93.75% of the C57BL/6J genome (31). This congenic colony has been propagated by sibling mating and forced heterozygosity for the $A^v$ allele for over 200 generations, resulting in an essentially invariant genetic background.

Virgin $a/a$ females, 8 weeks of age, were assigned randomly to NIH-31 diet or NIH-31 supplemented with the methyl donors and cofactors folic acid, vitamin B$_{12}$, choline chloride, and anhydrous betaine (Harlan Teklad) (30). All ingredients were provided by Harlan Teklad except for the anhydrous betaine (Finnsgur Bioprodutcs). The diets were provided for 2 weeks before the females were mated with $A^v/a$ males and throughout pregnancy and lactation. Upon weaning to a stock maintenance diet at age 21 days, the $A^v/a$ offspring were weighed, tail tipped, photographed, and rated for coat color phenotype (Fig. 2A). Animals in this study were maintained in accordance with all relevant federal guidelines, and the study protocol was approved by the Duke University Animal Care and Use Committee.
Phenotypic classification. The coat color phenotype of \( A^{\text{av}}/a \) mice was assessed at 21 and 100 days of age. A single observer classified coat color by visual estimation of the proportion of brown fur: yellow (≤5% brown), slightly mottled (≥5% but less than half), mottled (about half), heavily mottled (greater than half but ≤95%), and pseudoagouti (>95%). The term pseudoagouti is used to describe \( A^{\text{av}}/a \) animals in which ectopic \textit{agouti} expression was silenced (or nearly silenced) by CpG methylation, recapitulating the brown agouti phenotype of an \( A^{\text{av}}/a \) mouse.

Long-range PCR. The Expand Long-Template PCR system (Roche) was used per the manufacturer’s instructions. Primers were designed to amplify either the PS1A or exon 1A region (6) from genomic DNA (22): exon 1A forward, 1A-F (TCAGATTCTGGAGTGCAGACATGGATCC) and reverse, 1A-R1 (TTCCAG
GENOMIC DNA (22) from two Avy/Avy IAP, we used forward primer IAPF3 (ATTTTTAGGAAAAGAGAGTAAGA (TGAATCAGAAAGGATTTAGTAAAATGGCTC). PS1A-R2 (GAAGAAAACCATGAATCAGAAAGGATTTAG), and PS1A-R3 (TGAGTCAAGGAAGGGTGTTAGTATGAC). PS1A-R1 (AAAGCATTTTTGAAGAAAACCATGAATC), GATTCATCAATAATCGCT), 1A-F, and reverse, PS1A-R1 (AAAGCATTTTTGAAGAAAACCATGAATC), GATTCATCAATAATCGCT), 1A-R2 (AGGTACTGAAATTAACACGGCG)

The PS1A and exon 1A regions were amplified from genomic DNA (22) present in both exon 1A and PSA1 that allowed the a and A’ alleles to be cleaved selectively by Alol and SacI, respectively. We digested 0.5 μg of genomic DNA (22) overnight with 5 U of either Alol or SacI. Bisulfite modification was performed by a protocol adapted from the recently optimized protocol of Granaou et al. (9). Specifically, each digested genomic DNA sample was precipitated with ethanol, washed, and denatured in 50 μl of 0.3 M NaOH (20 min at 37°C). Deamination was initiated by addition of 450 μl of a solution of saturated sodium bisulfite (Sigma) and 10 mM hydroquinone (Sigma), pH 5.0. Each sample was overlaid with mineral oil and incubated for 4 h at 55°C in the dark. The samples were desalted with the Wizard DNA clean-up system (Promega) and resolubilized in 50 μl 1 mM Tris-Cl, pH 8.0. DNA samples were desulfonated by addition of 5.5 μl of 3 M NaOH and incubation at 37°C for 20 min, then ethanol precipitated, washed in 75% ethanol, and suspended in 10 μl of 1 mM Tris-Cl, pH 8.0. Because of the possible differential stability of methylated and unmethylated DNA under long-term storage conditions, all studies were conducted with freshly isolated genomic DNA.

Bisulfite-modified DNA (4 μl) was PCR amplified in 50-μl reactions with Platinum Taq DNA polymerase (Invitrogen) per the manufacturer’s instructions (40 cycles). PCR bands were agarose gel purified (GenElute Minus EtBr Spin Columns; Sigma) and sequenced manually (Thermo Sequenase radiolabeled terminator cycle sequencing kit; USB Corporation) according to the manufacturer’s instructions. Sequencing products were resolved by polyacrylamide gel electrophoresis, with blank lanes between the C and T lanes to avoid signal overlap. Percent methylation at each CpG site was quantitated by phosphor imaging (percent methylation = 100 × ([C volume]/[C volume + T volume])). Sequencing confirmed that endonuclease digestion was complete. The seven CpG sites studied in the PS1A region (and corresponding exon 1A region) are located at positions 910, 916, 934, 943, 956, 972, and 991 of accession number AF540972. These sites were chosen due to their relative proximity to the regions of dissimilarity between PS1A and exon 1A, enabling region-specific PCR amplification following bisulfite conversion. To analyze CpG methylation within the proximal long terminal repeat of the IAP, we PCR amplified the IAP-PS1A junction with the A’ IAP sequence (8). The nine CpG sites studied are located 138, 130, 124, 101, 80, 40, 27, 24, and 12 nucleotides upstream of the IAP-PS1A junction (8).

Primers used for bisulfite sequencing studies. For the exon 1A region, we used forward primer 1ABF3 (ATTTTGTGTTAATAAGGTATGATAG), reverse primer 1ABR3 (TAAACTAAATCAATAACACCCACC), and sequencing primer 1ABF4 (ATTTTGATGAAATTTCCTTG), for the PS1A region, we used forward primer 1ABF3, reverse primer 1ABR2 (CCATAATACTAAAAAAAAAAATAACCCACC), and sequencing primer 1ABF4. For the A’ IAP, we used forward primer IAPF3 (ATTTTGTGAAATAAAGGATAGTAAGTAAG), reverse primer IAPBR4 (TAAACTAAATCAATAACCCACC), and sequencing primer IAPFS (ATATTATTTTGTAGTGATGTATTTG).

Statistical analysis. Because the supplements were provided to the dams, they were the appropriate units of analysis. Therefore, all analyses used within-litter averages for 9 litters (30/A’ offspring) and 10 litters (39/A’ offspring) for the unsupplemented and supplemented groups, respectively. Group comparisons of litter size and day 21 weights were performed by t test. Percent methylation data were not distributed normally and were therefore transformed dichotomously (<20% = 0, ≥20% = 1) before analysis. Relationships among supplementation, A’ methylation, and coat color were analyzed by mediational regression analysis (3) with SAS software.

Nucleotide sequence accession numbers. The following sequence data were submitted to GenBank: A’’ PS1A (accession number AF540972), A’’ PS1A (accession number AF540973), A’’ exon 1A (accession number AF540974), and A’’ exon 1A (accession number AF540975).

RESULTS AND DISCUSSION

To determine if the IAP insertion causes the epigenetic metastability in the A’’ region, we needed first to determine the IAP location in the A’’ allele. Exploiting sequence dissimilarities between exon 1A and PSA1 (6) (Fig. 1A), long-range PCR was used to amplify the sequence bracketing the consensus IAP insertion site of both regions. This demonstrated clearly that the 4.5-kb IAP insert is contained within PSA1 and not in exon 1A, as previously reported (8) (Fig. 1B).

To distinguish between the A’’ and a alleles and thus enable A’’-specific quantitation of PSA1 methylation in A’’/a mice, we sequenced the PSA1 region downstream from the consensus IAP insertion site in a/a and A’’/A’’ homozygotes. We identified and exploited a single-nucleotide polymorphism within an Alol consensus sequence to cleave the PSA1 region of the a allele while leaving the A’’ allele intact. Hence, by digesting genomic DNA with Alol and employing reverse primers specific to PSA1, bisulfite sequencing (9) was used to quantify site-specific CpG methylation of the A’’ PSA1 in A’’/a mice. Importantly, because each A’’/a cell contains only one copy of the A’’ allele, our assay quantitates the percentage of cells in which each A’’ CpG site examined is methylated.

Dietary supplementation of a/a dams throughout the reproductive cycle did not affect litter size or offspring body weight at age 21 days (data not shown). Supplementation did shift the coat color distribution of A’’/a offspring toward the brown (pseudoagouti) phenotype (Fig. 2B). To determine if this phenotypic change was caused by increased A’’ CpG methylation, we quantitated PSA1 CpG methylation at seven CpG sites (~600 bp downstream from the IAP insertion site) in tail tip DNA from all A’’/a offspring born to nine unsupplemented and 10 supplemented dams. We chose to measure CpG methylation in this region for two reasons: (i) amplification of bisulfite-treated DNA was more reliable in this region than in the IAP long terminal repeat, and (ii) the correlations between average percent methylation and coat color phenotype are comparable in the two regions (IAP long terminal repeat r² = 0.82, downstream of PSA1 r² = 0.85), indicating that measurements of CpG methylation 600 bp downstream of the IAP insertion site are representative of average methylation levels throughout the PSA1 region encompassing the A’’ transcription start site.

Percent methylation in A’’ PSA1 was distributed bimodally in unsupplemented A’’/a offspring (Fig. 3A), suggesting a probabilistic epigenetic switch that tends to assume one of two methylation states (19). Maternal supplementation caused a general increase in methylation at each site (Fig. 3A). We examined the relationships among supplementation, CpG methylation, and coat color by mediational regression analysis (3). The highly significant effect of supplementation on coat coloration vanished when A’’ methylation was included in the model (Fig. 3B). This provides the first experimental evidence that A’’ CpG methylation mediates the effect of supplementation on A’’/a coat color.

To determine if the nutritional effect on A’’ PSA1 methylation in tail DNA extends to other tissues, average percent methylation of A’’ PSA1 was also measured in liver, kidney, and brain samples from animals representing the five coat color phenotypes. PSA1 methylation in the tail correlated
highly with that in the other tissues (Fig. 4A). The tissues studied were derived from the three germ layers of the early embryo: endoderm (liver), mesoderm (kidney), and ectoderm (brain). These data thus indicate that $Avy$ methylation is determined in the early embryo and maintained with high fidelity throughout development. This is consistent with previous studies showing high agouti expression in all tissues of yellow but not pseudoagouti $Avy/a$ mice (31). Hence, the nutritional effect
on \( A^{\nu} \) methylation likely occurs during early embryonic development and affects all tissues.

Coat color phenotype and \( A^{\nu} \) methylation also persisted to adulthood. Independent classification of coat color phenotype at age 100 days agreed with the day 21 classification in 48 of 50 \( A^{\nu}/a \) mice (data not shown). Similarly, mean \( A^{\nu} \) PS1A methylation in day 21 tail DNA predicted that in day 100 liver DNA (Fig. 4B). These data further support the idea that \( A^{\nu} \) PS1A methylation levels in tail DNA reflect those in other tissues and demonstrate that average \( A^{\nu} \) methylation is maintained quantitatively from weaning to adulthood. Hence, transient exposure of \( A^{\nu}/a \) mice to methyl supplementation in utero causes a shift in epigenotype that persists to influence the adult phenotype.

Alleles such as \( A^{\nu} \), whose epigenetic marks are determined probabilistically but subsequently maintained stably, have been termed metastable epialleles (19). Metastable epialleles are commonly associated with transposable elements (19), which can interfere with the expression of neighboring genes (27). Nonetheless, it remains unclear if the epigenetic metastability of the \( A^{\nu} \) allele is due to the IAP insert or if some unique characteristic of the PS1A sequence contributes to its probabilistic variability in Cpg methylation. Because the exon 1A region is nearly identical to the PS1A region (Fig. 1A), it provides an ideal cis negative control region by which to examine the IAP’s influence on epigenetic metastability. Similarly, the PS1A region on the \( a \) allele of \( A^{\nu}/a \) heterozygotes provides a negative control region in trans.

We therefore measured site-specific methylation of seven homologous CpG sites within exon 1A and PS1A of the \( a \) and \( A^{\nu} \) alleles in \( A^{\nu}/a \) animals of divergent phenotypes (Fig. 5). The extreme interindividual variability of \( A^{\nu} \) PS1A (\( P < 0.0001 \) by analysis of variance; Fig. 5B) was not observed in consensus sites within PS1A and exon 1A on the \( a \) allele (Fig. 5C and D). Instead, methylation in those regions was tightly regulated in a site-specific fashion and did not differ significantly among individuals. Notably, the interindividual coefficient of variation in percent methylation at each site was only slightly greater than that obtained by serial digestion and bisulfite sequencing of replicate DNA samples from a single individual (data not shown). Percent methylation at individual CpG sites was quantitatively similar between the two regions on \( a \). Methylation of \( A^{\nu} \) exon 1A, which lies approximately 15 kb upstream of the IAP, was also not significantly different among individual mice (Fig. 5A), but was hypermethylated relative to consensus sites on the \( a \) allele (\( P < 0.0001 \)). Clearly, epigenetic instability that renders \( A^{\nu} \) nutritionally labile is associated with the IAP insertion.

Conversely, proximity to PS1A apparently destabilizes IAP methylation. We performed bisulfite sequencing of the IAP/PS1A junction to quantify methylation at nine CpG sites within the cryptic promoter of the IAP (data not shown). Each animal’s average percent methylation in the PS1A region was predicted by that in the neighboring IAP regardless of maternal diet (\( r^2 = 0.92, n = 22 \) animals). Hence, whereas most IAPs in the mouse genome are heavily methylated (24), methylation at the \( A^{\nu} \) IAP correlates with that in the neighboring PS1A region and varies dramatically among individuals. Therefore, epigenetic metastability at the \( A^{\nu} \) locus occurs via a mutual interaction between a transposable element and its specific genomic region.

These results indicate that epigenetic metastability caused by juxtaposition of transposable elements and genomic promoter region DNA renders a subset of mammalian genes epigenetically labile to the effects of nutrition and other environmental influences during early development. Our findings have important implications for humans because transposable elements constitute over 35% of the human genome (32) and are found within about 4% of human genes (16). Furthermore, many human genes are transcribed from a cryptic promoter within the L1 retrotransposon (17), analogous to ectopic \textit{agouti} transcription originating in the \( A^{\nu} \) IAP. It has been proposed that transposable elements in the mammalian genome cause considerable phenotypic variability, making each individual mammal a “compound epigenetic mosaic” (27). Our results provide compelling evidence that the specific composition of each individual’s “epigenetic mosaic” is influenced by early nutrition.
Our findings are also important in the context of epigenetic inheritance at the $A^v$ locus. When $A^v/a$ animals inherit the $A^v$ allele maternally, agouti expression and coat color phenotype are correlated with maternal phenotype in that yellow dams produce fewer pseudoagouti offspring than do pseudoagouti dams (28, 30). This phenotypic inheritance was originally attributed to a maternal effect on metabolic differentiation (28). A recent study (15), however, suggests that this parental effect...

FIG. 5. Percentage of cells methylated at each of seven CpG sites in $A^v$ exon 1A (A), $A^v$ PS1A (B), a exon 1A (C), and a PS1A (D). Each graph shows data from the same two slightly mottled and two pseudoagouti $A^v/a$ animals (four total). In the $A^v$ PS1A region (B), CpG methylation correlates with coat color; the pseudoagouti animals (circles) were heavily methylated, and the slightly mottled animals (triangles) were hypomethylated. In the other three regions (A, C, and D), methylation was independent of coat color (each point shows the mean ± standard error of the mean). On the $a$ allele, methylation was tightly regulated and site dependent and did not differ significantly between the exon 1A and PS1A regions (C and D). Note that CpG site 5 is missing in the exon 1A region of the $a$ allele in C due to a single-nucleotide polymorphism. The $A^v$ exon 1A region (A) was hypermethylated relative to consensus sites on the $a$ allele ($P < 0.0001$). Given the 99% sequence identity of these four regions, the epigenetic metastability of $A^v$ PS1A is clearly associated with the neighboring IAP.
is caused by incomplete erasure of epigenetic marks at the \( A\)' locus in the female germ line. Our findings show that early nutrition can influence the establishment of epigenetic marks at the \( A\)' locus in the early embryo, thereby affecting all tissues, including, presumably, the germ line. Hence, incomplete erasure of nutritionally induced epigenetic alterations at \( A\)' provides a plausible mechanism by which adaptive evolution (10) may occur in mammals.

The moderate nature of the nutritional treatment used in these studies further underscores their relevance to humans. Whereas severe methyl donor deficiency has been demonstrated to induce gene-specific DNA hypomethylation in rodents (12), we show here that merely supplementing a mother’s nutritionally adequate diet with extra folic acid, vitamin B\(_{12}\), choline, and betaine can permanently affect the offspring’s DNA methylation at epigenetically susceptible loci. This finding supports the conjecture that population-based supplementation with folic acid, intended to reduce the incidence of neural tube defects, may have unintended influences on the establishment of epigenetic gene-regulatory mechanisms during human embryonic development (21).

It is increasingly evident that epigenetic gene regulatory mechanisms play important roles in the etiology of human diseases (11, 18). Our findings demonstrate that mammalian metastable epialleles associated with transposable elements enable early environmental influences, including nutrition, to persistently affect these important regulatory mechanisms. Hence, epigenetic alterations at metastable epialleles are a likely mechanistic link between early nutrition and adult chronic disease susceptibility.

ACKNOWLEDGMENTS

We thank George Wolff for providing \( A\)' animals, Michael Babyak for statistical advice, and Kay Nolan and Susan Murphy for suggestions on the manuscript. This work was supported by a Dannon Institute fellowship (R.A.W.) and NIH grants CA25951 and ES08823.

REFERENCES