

Metastable Epialleles, Imprinting, and the Fetal Origins of Adult Diseases

DANA C. DOLINOY, RADHIKA DAS, JENNIFER R. WEIDMAN, AND RANDY L. JIRTLE

Department of Radiation Oncology [D.C.D., R.D., J.R.W., R.L.J.], University Program in Genetics and Genomics [D.C.D., R.D., R.L.J.], Duke University Medical Center, Durham, North Carolina 27710; Integrated Toxicology Program [D.C.D., R.L.J.], Duke University, Durham, North Carolina 27708

ABSTRACT: Epigenetics is the study of the heritable changes in gene expression that occur without a change in the DNA sequence itself. These heritable epigenetic changes include chromatin folding and attachment to the nuclear matrix, packaging of DNA around nucleosomes, histone modifications, and DNA methylation. The epigenome is particularly susceptible to dysregulation during gestation, neonatal development, puberty, and old age. Nevertheless, it is most vulnerable to environmental factors during embryogenesis because the DNA synthetic rate is high, and the elaborate DNA methylation patterning and chromatin structure required for normal tissue development is established during early development. Metastable epialleles are alleles that are variably expressed in genetically identical individuals due to epigenetic modifications established during early development and are thought to be particularly vulnerable to environmental influences. The viable yellow agouti (A^{vy}) allele, whose expression is correlated to DNA methylation, is a murine metastable epiallele, which has been used as an epigenetic biosensor for environmental factors affecting the fetal epigenome. In this review, we introduce epigenetic gene regulation, describe important epigenetic phenomenon in mammals, summarize literature linking the early environment to developmental plasticity of the fetal epigenome, and promote the necessity to identify epigenetically labile genes in the mouse and human genomes. (*Pediatr Res* 61: 30R–37R, 2007)

Historically, DNA has been considered the sole unit of biologic inheritance. Therefore, research was designed to investigate how individuals with different genotypes respond to various environmental factors and how these responses change over time. Recently, however, the revelation that epigenetic marks are influenced by environmental factors (1,2), and may also be inherited transgenerationally (3,4) has promoted the investigation of how epigenetic variability affects phenotype. If the genome is thought of as being similar to the hardware in a computer, the epigenome is the software that directs the computer's operation. Thus, identifying epigenetic targets and defining how they are dysregulated in human disease by environmental exposures will allow for the development of innovative novel diagnostic, treatment, and

prevention strategies that target the “epigenomic software” rather than the “genomic hardware.”

The term “epigenetics” was first defined in the 1940s by developmental biologist Conrad Waddington as “the interactions of genes with their environment, which bring the phenotype into being (5).” Subsequently, in 1975, Holliday and Pugh proposed covalent chemical DNA modifications, including methylation of cytosine-guanine (CpG) dinucleotides, as the molecular mechanism to explain Conrad's hypothesis (6). Several years later, the revelations that X-inactivation in mammals and genomic imprinting are regulated by epigenetic mechanisms highlighted the heritable nature of epigenetic gene regulation mechanisms (7,8). The genomics revolution inspired the investigation of global rather than local gene analyses, and the term “epigenomics” was coined as the study of the “effects of chromatin structure including the higher order of chromatin folding and attachment to the nuclear matrix, packaging of DNA around nucleosomes, covalent modifications of histone tails (acetylation, methylation, phosphorylation, ubiquitination), and DNA methylation (9).” Additionally, evidence demonstrating the resistance of certain gene loci to methylation reprogramming during embryogenesis revealed that epigenetic modifications can be inherited not only mitotically but also transgenerationally (3,4,10).

EPIGENETIC MECHANISMS

The two most extensively studied epigenetic mechanisms in mammals are methylation of cytosine at the carbon-5 position in CpG dinucleotides and chromatin packaging of DNA *via* histone variants and posttranslational histone modifications as well as subsequent nonhistone protein recruitment to specific regions of DNA. Both chromatin condensation and DNA methylation are generally associated with gene silencing. They are not necessarily independent events, but may act together to alter gene transcription. The influence of regulatory small RNA/micro RNA on gene transcription is another field of epigenetic gene regulation that is now emerging.

DNA methylation. Cytosine methylation of CpG dinucleotides is a postsynthetic DNA modification in which a methyl

Received November 3, 2006; accepted December 7, 2006.

Correspondence: Randy L. Jirtle, Ph.D., Box 3433, Duke University Medical Center, Durham, NC 27710, e-mail: jirtle@radonc.duke.edu

This work was supported by DOE grant DE-FG02-05ER64101 and National Institutes of Health grants ES13053, ES08823, ES015165, and T32-ES07031.

DOI: 10.1203/pdr.0b013e31804575f7

Abbreviations: **A**, agouti; A^{vy} , viable yellow agouti; **Axin^{fl}**, axin fused; **CpG**, cytosine-guanine dinucleotide; **IAP**, intracisternal A particle; **LOI**, loss of imprinting; **RISC**, RNA-induced silencing complex; **RITS**, RNA-induced initiation of transcriptional silencing; **SAM**, S-adenosylmethionine

group is enzymatically transferred from *s*-adenosylmethionine (SAM) to the carbon-5 position of the cytosine ring. This reaction is catalyzed by DNA methyl transferases Dnmt1, Dnmt3a, Dnmt3b, and Dnmt2 as well as accessory proteins like Dnmt3L (11). The resulting 5-methylcytosine (5mC) plays an important role in many aspects of cellular function. The methyl group of the 5mC extends into the major groove of DNA, inhibiting transcription by interfering with transcription factor binding proteins. In addition, DNA methyl transferases and methylated DNA interact with higher order chromatin proteins, such as the repressive PcG protein, EZH2, to affect histone modifications and further compact chromatin (12).

The distribution of CpG sequences in mammalian genomes is nonrandom (13). CpG dinucleotides are greatly underrepresented in the mammalian genome due to evolutionary spontaneous deamination of 5-methylcytosine to thymine. The majority of unmethylated sites occur in CpG islands, defined as discreet regions containing more than 50% CpG content. Normally, they are located within or near gene promoters or first exons of housekeeping genes. In contrast, the promoter and regulatory regions of transposable elements are methylated, thereby inhibiting the parasitic transposable and repetitive elements from replicating. Finally, the *cis*-acting imprinting regulatory elements involved in parental-specific silencing of one allele of imprinted genes are differentially methylated (13,14). Thus, the monoallelic, parent-of-origin-dependent expression of imprinted genes is epigenetically regulated, and therefore susceptible to expression deregulation by environmental factors that alter the epigenome.

Chromatin packaging. Epigenetic manipulation of cellular phenotype is also driven by alteration of chromatin structure *via* covalent histone modifications and incorporation of histone variants into the nucleosome (15). Chromatin is a nucleoprotein complex that packages linear genomic DNA *via* an array of nucleosomes. Each nucleosome consists of 147 base pairs of DNA coiled around an octamer of histone proteins. Each octamer contains two copies each of the four core histones, H2A, H2B, H3, and H4. Chromatin may be further modified by association with linker histones, histone variants, and nonhistone proteins as well as a myriad of posttranslational modifications of histone proteins, including histone acetylation, methylation, ubiquitinylation, phosphorylation, and ADP-ribosylation (16,17).

Histone acetylation is usually associated with transcriptional activation because the affinity of histone proteins for DNA is reduced and chromatin packaging is relaxed. Histone methylation results in various transcriptional consequences depending on histone number and the lysine residue modified. For example, histone H3 methylation at lysine 9 is associated with heterochromatin and gene silencing, while histone H3 methylation at lysine 4 or 27 is associated with transcriptional activation. Each lysine residue may be methylated in the form of mono-, di-, or tri-methylation, adding enormous complexity to the histone code. Furthermore, histone modifications interact with DNA methylation patterns to recruit multi-subunit chromatin-protein complexes, such as the repressive polycomb group proteins or the activating SWI-SNF proteins, adding yet another layer of complexity to epigenetic gene regulation.

Small RNA. Complementary RNA has long been known to induce gene-silencing, but the exact mechanisms are still being elucidated. Short antisense RNA transcripts are produced within the nucleus by the action of the enzyme Dicer, which cleaves double-stranded RNA precursors into 21–26 nucleotide long RNA species (18,19). These then associate with silencing-effector complexes, such as RISC (which directs cleavage of cognate mRNA or causes translational repression) and RITS (which mediates heterochromatin formation at target loci), and abrogate gene-expression. Thus, regulation mediated by small RNA occurs both at the post-transcriptional and transcriptional level.

The latter transcriptional regulation is referred to as “epigenetic silencing,” and is mediated either by covalent modifications of chromatin (such as H3 methylation at Lys9) or by DNA methylation. Heterochromatin formation, particularly at centromeric regions, is initiated by siRNA produced *via* transposon-derived repeat regions, in association with the RITS complex. Spreading and maintenance of this repressed state involves a self-enforcing feedback loop of siRNA production and recruitment of heterochromatic proteins such as Swi6 (in fission yeast) or HP1 (in *Drosophila* and vertebrates). Recently, the ribonuclease Eri1 has been implicated in the down-regulation of this heterochromatin assembly, by preventing excess siRNA accumulation.

EPIGENETIC PHENOMENA AND THE FETAL ENVIRONMENT

Environmental exposure to nutritional, chemical, and physical factors can alter gene expression, and affect adult phenotype by not only mutating promoter regions of genes, but also by modifying CpG methylation or other epigenetic modifications at critical epigenetically labile genomic regions (2). Three potential epigenetic susceptibility targets for such environmentally induced effects are transposable elements, the promoter regions of housekeeping genes, and *cis*-acting regulatory elements of imprinted genes. These genomic targets contain CpG islands that are normally methylated, unmethylated, or differentially methylated, respectively. Of these epigenetically labile targets, the transposable elements and imprinted genes are particularly important.

Transposable elements and metastable epialleles. Most regions of the adult mammalian genome exhibit little interindividual variability in tissue-specific CpG methylation levels. In contrast, CpG methylation is determined stochastically at specific transposable element insertion sites, causing cellular epigenetic mosaicism and individual phenotypic variability (2).

Transposons are parasitic, repetitive mobile elements that are dispersed throughout mammalian genomes. They are remnants of ancestral infections which became fixed in the germline DNA, and subsequently increased in copy number. The sequencing of the human genome provided the rather humbling finding that transposons comprise roughly 45% of our genome (20). Most transposable elements are silenced by CpG methylation (21); however, the epigenetic state of a subset of transposable elements is metastable (22). They can therefore potentially affect the expression of neighboring genes.

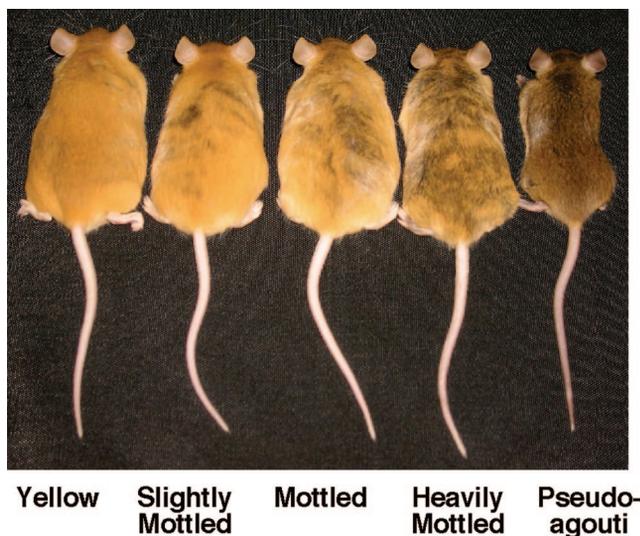


Figure 1. Genetically identical 3-mo-old A^{vy} mice representing the five coat color phenotypes. Yellow mice are hypomethylated at the transposable element upstream of the *Agouti* gene allowing maximal ectopic expression, whereas hypermethylation of this site silences ectopic *Agouti* expression in the pseudoagouti animals. Mice that are predominately yellow are also clearly more obese than brown mice. Reprinted from Dolinoy DC *et al.* 2006 *Environ Health Perspect* 114:567–572, with permission.

Metastable epialleles are alleles that are variably expressed in genetically identical individuals due to epigenetic modifications that were established during early development (22). The term “metastable” refers to the labile nature of the epigenetic state of these alleles, while “epiallele” defines their potential to maintain epigenetic marks transgenerationally. In addition, gestational exposure to nutritional agents and other environmental factors has been demonstrated to alter epigenetic marks at metastable epialleles (1,23,24)

One model for such epigenetically based phenotypic variability is the viable yellow agouti (A^{vy}) mouse (Fig. 1). The murine *Agouti* gene encodes a paracrine signaling molecule that promotes follicular melanocytes to produce yellow pheomelanin pigment rather than black eumelanin pigment. Transcription is normally initiated from a hair-cycle specific promoter in exon 2 of the *Agouti* (A) allele (Fig. 2A). Transcription of the A allele normally occurs only in the skin where transient expression in hair follicles during a specific stage of hair growth results in a subapical yellow band on each black hair, causing the brown (agouti) coat color of wild-type mice (25).

The A^{vy} allele resulted from the insertion of an IAP murine retrotransposon upstream of the transcription start site of the *Agouti* gene (Fig. 2A) (1,25). A cryptic promoter in the proximal end of the A^{vy} IAP promotes constitutive ectopic *Agouti* transcription, leading to yellow fur, obesity, and tumorigenesis (3,26). CpG methylation in the A^{vy} IAP correlates inversely with ectopic *Agouti* expression. The degree of methylation varies dramatically among individual isogenic A^{vy}/a mice (Fig. 2B), causing a wide variation in coat color ranging from yellow (unmethylated) to pseudoagouti (methylated) (Fig. 1) (3). Increased body weight is also positively correlated to ectopic *agouti* expression, as seen in the wk 15 isogenic A^{vy}/a litter mates shown in Figure 1.

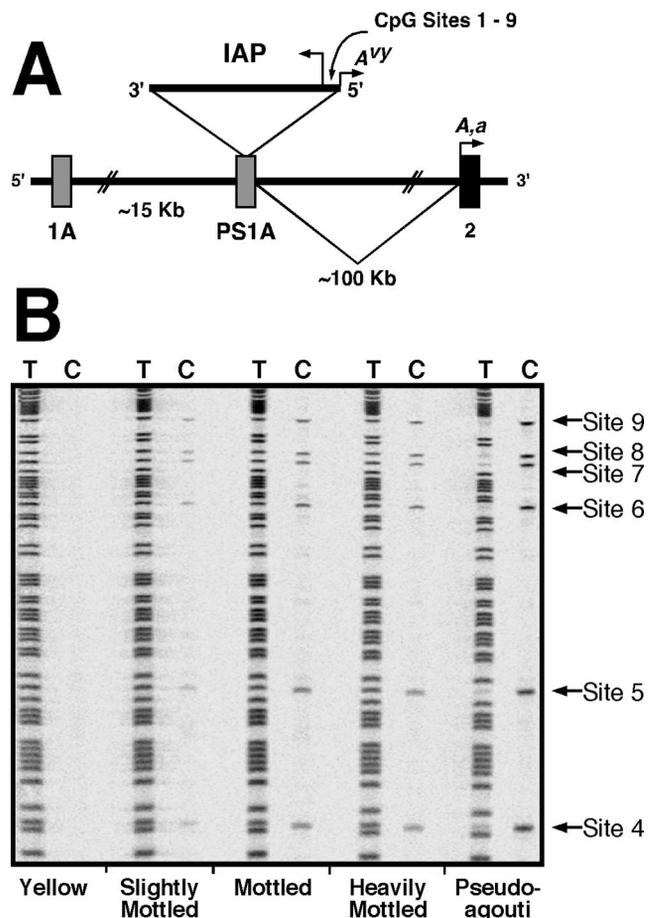


Figure 2. The A^{vy} metastable epiallele. (A) A contraoriented IAP insertion within pseudoxon 1A (PS1A) of the murine *Agouti* gene contains a cryptic promoter (short arrowhead labeled A^{vy}) that drives ectopic *Agouti* expression. Transcription of A and a alleles initiates from a hair-cycle specific promoter in exon 2 (short arrowhead labeled A , a). (B) Pseudoagouti animals exhibit the highest degree of CpG methylation at sites 4 to 9, as revealed by bisulfite sequencing analysis. Reprinted from Dolinoy DC *et al.* 2006 *Environ Health Perspect* 114:567–572, with permission.

Recently, the A^{vy} model has been used as an epigenetic biosensor for determining if maternal dietary exposures affect the fetal epigenome. Cooney *et al.* (27) and Waterland *et al.* (1) demonstrated that maternal dietary methyl supplementation with extra folic acid, vitamin B₁₂, choline, and betaine shifts the coat color distribution of the offspring toward the pseudoagouti phenotype. Waterland *et al.* (1) further demonstrated that the shift in coat color distribution was caused by increased methylation at each of seven A^{vy} pseudoxon 1A (PS1A) CpG sites. Methylation profiles at the seven CpG sites were highly correlated in tissues from ectodermal (brain and tail), endodermal (liver), and mesodermal (kidney) lineages, indicating that methylation profiles at the A^{vy} locus are established before embryonic stem cell differentiation. In addition, methylation in day 21 tissues was correlated to methylation in day 100 tissues demonstrating that methylation of this metastable epiallele is efficiently maintained over time.

In 2006, Dolinoy *et al.* (23) demonstrated that maternal dietary genistein supplementation of mice during gestation, at levels comparable to humans consuming high soy diets, shifted the coat color of viable yellow *Agouti* (A^{vy}/a) offspring

toward pseudoagouti by increasing methylation of six CpG sites in a retrotransposon upstream of the *Agouti* gene. Moreover, the genistein-induced hypermethylation protected *A^{vy}/a* animals from obesity in adulthood. Thus, epigenetic alterations at metastable epialleles, such as transposable elements, can link early nutrition to adult chronic disease susceptibility, not only in mice, but also potentially in humans.

A second identified murine metastable epiallele is the *Axin^{Fu}* allele. The wild-type *Axin* gene encodes the axin protein, which inhibits *Wnt* signaling and is therefore involved in mammalian embryonic axis formation (28). The *Axin* gene is ubiquitously expressed during embryonic development and adulthood (28). The *Axin^{Fu}* allele contains a spontaneous IAP insert within intron 6 of the gene. This results in expression of a truncated biologically active 3' transcript of *Axin* that originates within the transposable element, and causes axial duplications during development (28,29). *Axin^{Fu}* mice have kinked tails of varying severity. Moreover, the extent of the kink in the tail is inversely related to the degree of IAP methylation at the *Axin^{Fu}* locus (4). Mice with severe tail kinks are less methylated at the *Axin^{Fu}* IAP, and mice with silent, phenotypically normal tails are more methylated within the IAP. This metastable epiallele is therefore similar to *A^{vy}*, but the location of the IAP within the gene is different.

Recently, Waterland *et al.* (24) repeated the maternal methyl donor study using the *Axin^{Fu}* mouse model to determine whether the ability of maternal dietary methyl supplementation with extra folic acid, vitamin B₁₂, choline, and betaine to induce epigenetic changes in offspring is gene locus dependent. The percentage of *Axin^{Fu}/+* offspring nonpenetrant for a kinky tail increased in the supplemented offspring when compared with controls. Like the *A^{vy}* model, methylation profiles at number of CpG sites near the *Axin^{Fu}* locus

showed increased methylation in supplemented offspring, indicating that nutritional supplementation affects offspring epigenotype at more than one locus. In contrast to the *A^{vy}* locus, methylation profiles were not similar in brain, liver, or kidney tissues, indicating that in the *Axin^{Fu}* model, the window of epigenetic vulnerability may occur later in development.

Imprinted genes. The vast majority of autosomal genes are expressed from both parentally contributed alleles; however, the expression of an increasing number of growth regulatory genes are controlled by an unusual epigenetic phenomenon referred to as genomic imprinting (13,30). Genomic imprinting is a non-Mendelian inherited epigenetic form of gene regulation that does not involve alterations in the DNA sequence but rather changes in DNA methylation and histone alterations, which are heritable during cell division (Fig. 3). Moreover, expression of the single functional allele of an imprinted gene is parent-of-origin dependent.

The health consequences of genomic imprinting are potentially disastrous. It eliminates the protection that diploidy normally affords against deleterious effects of recessive mutations. Furthermore, the epigenetic regulatory system that maintains the functionally haploid state of these genes can be deregulated by nongenotoxic environmental agents. Thus, imprinted genes significantly increase susceptibility to developmental and behavioral disorders, and diseases like asthma, cancer, diabetes, and obesity.

The most widely debated theory of why imprinting evolved, "the conflict hypothesis," predicts that imprinting arose because of a genetic tug-of-war between the parents to control the amount of nutrients extracted from the mother by her offspring (31,32). We demonstrated that imprinting evolved approximately 180 million years ago following the divergence of Prototherian (*i.e.* monotremes) from Therian (*i.e.* marsupi-

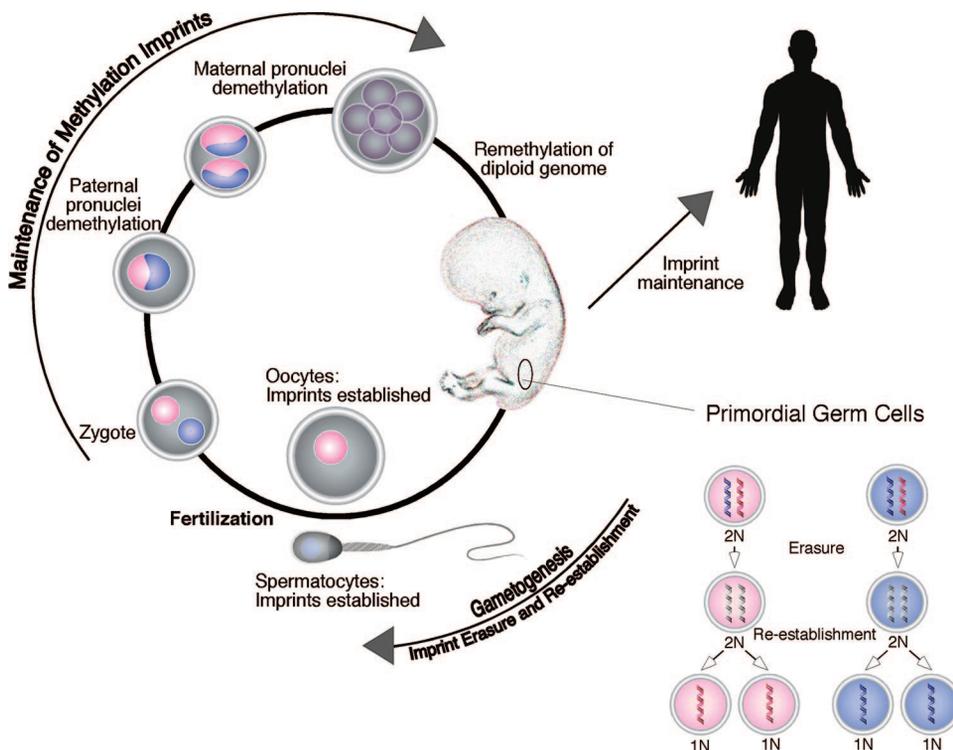


Figure 3. Imprinting throughout development. Methylation marks associated with imprinted genes are established at two distinct phases of development. During gametogenesis, the imprint marks present on the maternal (*pink, M*) and paternal (*blue, P*) chromosomes are erased (*gray*), followed by establishment of primary imprints which then reflect only the sex of the individual in which they reside. Just after fertilization, a global demethylation event occurs in the zygote, first in the paternal pronucleus (*blue, P*), followed by the maternal pronucleus (*pink, M*). Imprint marks that were established in the gametes must resist this demethylation process. Remethylation of the diploid genome occurs postimplantation, and includes setting of the secondary imprints, which are maintained throughout the lifespan of the individual. Reprinted from Murphy SK and Jirtle RL 2003 *Bioessays* 25:577–588, © 2003 Wiley Periodicals, Inc., with permission.

als and eutherians) mammals (13,33). Thus, genomic imprinting arose in mammals with the evolution of the placenta and advent of viviparity. Although *Igf2* (*IGF 2*) is imprinted in all Therian mammals investigated including humans, imprinting at the *Igf2r* locus was lost approximately 75 million years ago in an ancestral mammal that ultimately gave rise to primates (13,34). *NNAT* and *MEG3* are imprinted genes found only in eutherian mammals (35). Although *DLK1* and *p57* (*KIP2*) are present in marsupials, they are only imprinted in eutherians (36,37). These findings demonstrate that the expression of imprinted genes is species, tissue, and developmental stage dependent, and indicate that imprinting may play an important role in mammalian speciation.

Abnormal expression of imprinted genes during development results in a number of severe pediatric developmental disorders such as Prader-Willi syndrome (PWS), Angelman syndrome (AS), and Beckwith-Wiedemann syndrome (BWS) (13). In all three of these imprinting disorders, epigenetic alterations have an important contributory or causative role. Furthermore, *IGF2* loss of imprinting (LOI) during *in vitro* fertilization (IVF) is associated with a significant increase in the incidence of these developmental disorders (38).

Imprinted gene dysregulation can also occur in somatic cells, either by epigenetic or genetic mutations, causing cancer (39,40). Furthermore, only a single event is required to completely inactivate an imprinted tumor suppressor because imprinted genes are functionally haploid. Imprinted genes are therefore at a much greater risk of somatic cell inactivation by mutation, loss of heterozygosity (LOH) and epigenetic alterations in gene expression because one allele is already inactive because of imprinting. The imprinted, silenced allele has been equated to the "first hit," as proposed by Knudson in his two-step model for carcinogenesis.

Imprinted genes that encode for growth factors can also be inappropriately overexpressed in somatic cells through LOI (39,40). Cui and colleagues (41) demonstrated a surprising finding that lymphocytes in approximately 10% of the human population have LOI at the *IGF2* locus. Moreover, *IGF2* biallelic expression in peripheral lymphocytes is strongly correlated with *IGF2* LOI in normal colonic mucosa, and a personal history of colorectal cancer. The mechanism for this potentially systemic epigenetic alteration is presently unknown. Nevertheless, *IGF2* LOI is inherited or acquired early in life in a subset of people (41,42), and they appear to have an enhanced risk of developing cancer.

Early postnatal nutritional status also influences murine epigenetic gene regulation of the imprinted *Igf2*, a gene implicated in the etiology of a number of human cancers, including breast cancer (43,44). In 2006, Waterland *et al.* (45) demonstrated that postweaning diet affects the developmental relaxation of *Igf2* imprinting. A methyl-donor deficient diet administered for 60 d postweaning caused a significant *Igf2* LOI (expression from both rather than one parental allele) relative to that observed in mice receiving a natural control diet. Thus, the effects of nutrition on the epigenome are not limited only to the fetal stage of development, but may also influence early infant and childhood development.

EPIGENETIC REPROGRAMMING IN MAMMALS

Epigenetic marks, including CpG methylation, are generally stable in somatic cells; however, during at least two developmental time periods, the epigenome undergoes extensive reprogramming. These critical windows of development include gametogenesis as well as early preimplantation embryos (46). At gametogenesis genome-wide demethylation happens during the development of the primordial germ cells. In the mouse this process occurs from E11.5 to E12.5 (47). In the murine male germ cells, *de novo* methylation occurs around E16–18.5, whereas, in female germ cells *de novo* methylation occurs after birth in mature oocytes. This demethylation and remethylation cycle is thought to erase previous paternal imprints and reestablish sex-specific imprints. It may also be important for clearance of acquired epigenetic marks *via* genetic factors, environmental exposures, or disease state.

At fertilization both parental genomes undergo further epigenetic modifications. Initially, the paternal genome, which exists as a single copy and is densely packaged, exchanges protamines for histones and undergoes active demethylation before DNA replication (48). Secondly, the maternal genome, which exists in two copies and is arrested at metaphase II, completes meiosis and undergoes passive demethylation after several cleavages (47). This wave of epigenetic demethylation is thought to restore totipotency of the fertilized egg; however, some sequences are protected from epigenetic reprogramming at fertilization, including those regulating monoallelic expression of imprinted genes, repeat sequences such as IAP, and heterochromatin near chromosome centromeres (10,46,49). *De novo* methylation of both parental genomes occurs around implantation, with the embryonic lineages, such as the inner cell mass, showing hypermethylation in comparison to extra-embryonic lineages, such as the trophectoderm.

Importantly, some genomic loci fully or partially escape epigenetic reprogramming during *in utero* development. In the late 1970s breeding studies involving both *A^{vy}* and *Axin^{Fu}* revealed inheritance of coat color (50) or tail kink phenotype (51). For example, pseudoagouti *A^{vy}* mothers but not the fathers produced more pseudoagouti offspring. In contrast, penetrant *Axin^{Fu}* mothers and fathers produced more offspring with tail kinks. These observations led to the hypothesis that epigenetic modifications are not only mitotically heritable, but also transgenerationally heritable through inefficient erasure of epigenetic marks during gametogenesis. Using embryo transfer experiments in inbred mouse strains, Whitelaw and colleagues (3,4,52) determined that the observed inheritance patterns were not associated with unlinked modifier genes or differences in intrauterine environments, but rather with epigenetic inheritance resulting from incomplete epigenetic reprogramming. Recently Whitelaw's group assessed DNA methylation in *A^{vy}* gametes, zygotes, and blastocysts to determine whether methylation was the transgenerationally inherited epigenetic mark (53). Interestingly, DNA methylation was absent from the blastocyst, indicating that an epigenetic mechanism other than DNA methylation is the inherited mark.

In addition, epigenetic reprogramming may be affected by environmental factors, even in the absence of continued exposure. Interestingly, the influence of environmental factors on epigenetic gene regulation may also persist transgenerationally despite the lack of continued exposure in second, third, and fourth generations (54). Anway *et al.* (54) observed epigenetic inheritance of altered DNA methylation patterns in the male germ line following transient exposure to the anti-androgenic fungicide vinclozolin and the estrogenic pesticide methoxychlor. The fact that epigenetic marks are inherited transgenerationally, and that they are influenced by the environment supports the concept that genome-wide epigenetically labile genes need to be identified in both the mouse and human.

GENOME-WIDE IDENTIFICATION OF METASTABLE EPIALLELES

The developmental origins hypothesis states that environmental factors, such as nutrition acting in early life, program the risks for adverse health outcomes in adult life (55). Evidence shows that transposons adjacent to genes like the *Agouti* in A^{vy} mice (1,2), and the regulatory elements of imprinted genes are the epigenomic targets linking environmental exposures during gestation to the development of diseases in adulthood (2,56). The epigenome is particularly vulnerable to environmental perturbations during embryogenesis. Thus, to understand the etiology of chronic adult diseases, the epigenomic targets that are particularly vulnerable to dysregulation during early development must be identified.

Transposable elements and human disease. Although the human genome is riddled with transposable elements (20), it is presently unknown if any of them are regulating gene expression in a manner comparable to that seen at the *Agouti* locus in A^{vy} mice, and, if so, whether they affect human disease susceptibility. The A^{vy} mouse model dramatically demonstrates the importance of determining if epigenetically regulated genes with a methylation and expression profile similar to that of the *Agouti* gene in this mouse (*i.e.* “*Agouti* Expression Fingerprint”) are also present in the human genome.

The “*Agouti* Expression Fingerprint” is defined by a large variability in gene expression between individuals concomi-

tant with a low variability in gene expression between tissues in the three germ layers within an individual. The ratio of these two gene expression variances will therefore be large. Consequently, genome-wide expression chips should be useful in detecting genes with an “*Agouti* Expression Fingerprint” in both inbred mice and monozygotic twins where genetic variation is eliminated. This experimental approach will only detect genes whose expression pattern is epigenetically established before embryonic stem cell differentiation. Fortunately, this is the stage of development where the epigenome is most vulnerable to environmental perturbations.

Imprinted genes and human disease. The majority of autosomal genes are biallelically expressed. Nevertheless, most of the human diseases that currently are known or suggested to result from epigenetic abnormalities involve imprinted genes (13,44,57). Although epigenetic dysregulation of imprinted genes might just be easier to detect and correlate with disease than for nonimprinted genes, it is more likely that the additional complexity of epigenetic regulation for imprinted genes makes them more susceptible to epigenetic dysregulation compared with nonimprinted genes. Thus, it is critical to identify those genes that are imprinted in the human genome.

To date, most efforts to identify imprinted genes have been experimental, focusing on small regions of a chromosome. A robust method for genome-wide identification of imprinted genes involves the use of machine-learning algorithms trained to identify genomic motifs predictive of imprinted genes. Luedi *et al.* (58) recently developed such a bioinformatic approach for interrogating the entire mouse genome to identify genes with high probability of being imprinted. Data were collected on repeated elements, CpG islands, transcription factor binding sites, and other DNA characteristics within the upstream and downstream regions, introns, and exons of all annotated genes in the mouse genome. The most important genomic features for predicting the imprint status of a gene include the intronic presence of endogenous retrovirus (ERV) elements, and their insertion orientation relative to that of the gene (Fig. 4).

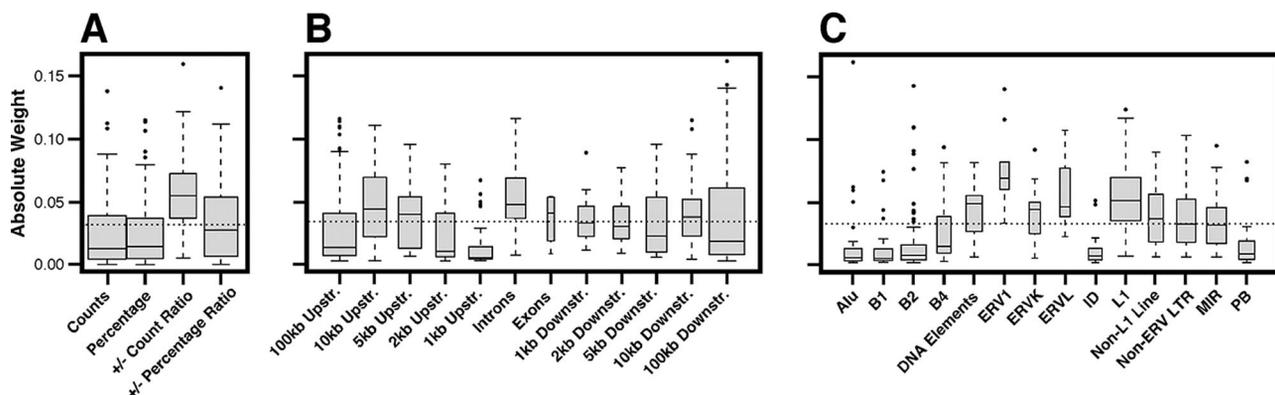


Figure 4. Box plots of the absolute weights of the features used in the imprinted vs nonimprinted classifier. (A) On average, the ratios of counts of repetitive elements carried the greatest absolute weight ($p = 3 \times 10^{-12}$). (B) Data on repetitive elements within the introns were the most important ($p = 4 \times 10^{-4}$), followed by the 10-kb upstream region ($p = 6 \times 10^{-3}$), while the 1-kb upstream window was of least importance ($p = 5 \times 10^{-10}$). (C) Among the repetitive elements, ERV1 ($p = 2 \times 10^{-5}$) and ERVL ($p = 5 \times 10^{-3}$) were of greatest average importance, followed by LINE L1 elements ($p = 4 \times 10^{-15}$). Redrawn from Luedi PP *et al.* 2005 Genome Res 15:875–884 ©2005 by Cold Spring Harbor Laboratory Press, with permission.

Of 23,788 annotated autosomal mouse genes, the imprinted gene prediction algorithm identified 600 (2.5%) genes to be potentially imprinted, 64% of which are predicted to exhibit maternal expression. The real power of this bioinformatics approach for predicting imprinted genes, however, lies in its ability to readily interrogate the genomes of any eutherian species for which complete genomic sequence is available. The probability that a gene is actually imprinted would be expected to increase if it is predicted to be imprinted in more than one species, particularly if the species are distantly related. Furthermore, phylogenetic comparison of the orthologous domains for those genes predicted to be imprinted in distantly related species should allow for the identification of evolutionarily conserved *cis*-acting elements potentially involved in imprinting regulation. These putative imprinting control elements would be expected to be the targets for epigenetic modification by environmental exposures during early development (56).

CONCLUSION

There is now compelling evidence that the risk of many chronic adult diseases and disorders result from exposure to environmental factors early in development (55,59). Moreover, it appears that a link between what we were exposed to *in utero* and disease formation in adulthood involves epigenetic modifications such as DNA methylation of transposable elements and *cis*-acting imprinting regulatory elements. Therefore, the development of genome-wide approaches, including expression arrays and novel bioinformatic tools (58), to determine those epigenetically labile targets in the human genome that are involved in the etiology of chronic diseases is critical. The application of the *A^{vy}* mouse model as an epigenetic biosensor for environmental effects on the fetal epigenome has identified several important nutritional factors affecting epigenetic gene regulation. The determination of other key chemical and physical environmental factors affecting the fetal epigenome will allow for better health intervention and preventive strategies for protecting individuals from environmentally induced changes in the fetal epigenome.

REFERENCES

- Waterland RA, Jirtle RL 2003 Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 23:5293–5300
- Waterland RA, Jirtle RL 2004 Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. *Nutrition* 20:63–68
- Morgan HD, Sutherland HG, Martin DI, Whitelaw E 1999 Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 23:314–318
- Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV, Whitelaw E 2003 Transgenerational inheritance of epigenetic states at the murine AxinFu allele occurs after maternal and paternal transmission. *Proc Natl Acad Sci U S A* 100:2538–2543
- Waddington CH 1940 *Organisers and Genes*. Cambridge University Press, Cambridge, UK
- Holliday R, Pugh J 1975 DNA modification mechanisms and gene activity during development. *Science* 187:226–232
- Willard HF, Brown CJ, Carrel L, Hendrich B, Miller AP 1993 Epigenetic and chromosomal control of gene expression: molecular and genetic analysis of X chromosome inactivation. *Cold Spring Harb Symp Quant Biol* 58:315–322
- Monk M 1988 Genomic imprinting. *Genes Dev* 2:921–925
- Murrell A, Rakyan VK, Beck S 2005 From genome to epigenome. *Hum Mol Genet* 14:R3–R10
- Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W 2003 Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* 35:88–93
- Goll MG, Bestor TH 2005 Eukaryotic methyltransferases. *Annu Rev Biochem* 74:481–514
- Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden J, Bollen M, Esteller M, Di Croce L, de Launoit Y, Fuks F 2006 The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439:871–874
- Murphy SK, Jirtle RL 2003 Imprinting evolution and the price of silence. *Bioessays* 25:577–588
- Gardiner-Garden M, Frommer M 1987 CpG islands in vertebrate genomes. *J Mol Biol* 196:261–282
- Saha A, Wittmeyer J, Cairns BR 2006 Chromatin remodelling: the industrial revolution of DNA around histones. *Nat Rev Mol Cell Biol* 7:437–447
- Caiafa P, Zampieri M 2005 DNA methylation and chromatin structure: the puzzling CpG islands. *J Cell Biochem* 94:257–265
- Cheung P, Lau P 2005 Epigenetic regulation by histone methylation and histone variants. *Mol Endocrinol* 19:563–573
- Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D 2004 RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* 303:672–676
- Matzke MA, Birchler JA 2005 RNAi-mediated pathways in the nucleus. *Nat Rev Genet* 6:24–35
- Bestor TH 2003 Cytosine methylation mediates sexual conflict. *Trends Genet* 19:185–190
- Yoder JA, Walsh CP, Bestor TH 1997 Cytosine methylation and the ecology of intragenomic parasite. *Trends Genet* 13:335–340
- Rakyan VK, Blewitt ME, Druker R, Preis JJ, Whitelaw E 2002 Metastable epialleles in mammals. *Trends Genet* 18:348–351
- Dolinoy DC, Wiedman J, Waterland R, Jirtle RL 2006 Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. *Environ Health Perspect* 114:567–572
- Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG 2006 Maternal methyl supplements increase offspring DNA methylation at AxinFused. *Genesis* 44:401–406
- Duhl DM, Vrieling H, Miller KA, Wolff GL, Barsh GS 1994 Neomorphic agouti mutations in obese yellow mice. *Nat Genet* 8:59–65
- Miltenberger RJ, Mynatt RL, Wilkinson JE, Woychik RP 1997 The role of the agouti gene in the Yellow Obese syndrome. *J Nutr* 127:1902S–1907S.
- Cooney CA, Dave AA, Wolff GL 2002 Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J Nutr* 132:2393S–2400S.
- Zeng L, Fagotto F, Zhang T, Hsu W, Vasicek TJ, Perry WL, Lee JJ, Tilghman SM, Gumbiner BM, Costantini F 1997 The mouse fused locus encodes axin, an inhibitor of the wnt signaling pathway that regulates embryonic axis formation. *Cell* 90:181–192
- Flood WD, Ruvinsky A 2001 Alternative splicing and expressivity of the Axin(Fu) allele in mice. *Heredity* 87:146–152
- Reik W, Walter J 2001 Genomic imprinting: parental influence on the genome. *Nat Rev Genet* 2:21–32
- Wilkins JF, Haig D 2003 What good is genomic imprinting: the function of parent-specific gene expression. *Nat Rev Genet* 4:359–368
- Haig D, Graham C 1991 Genomic imprinting and the strange case of the insulin-like growth factor II receptor. *Cell* 64:1045–1046
- Killian JK, Byrd JC, Jirtle JV, Munday BL, Stoskopf MK, MacDonald RG, Jirtle RL 2000 M6P/IGF2R imprinting evolution in mammals. *Mol Cell* 5:707–716
- Killian JK, Nolan CM, Wylie AA, Li T, Vu TH, Hoffman AR, Jirtle RL 2001 Divergent evolution in M6P/IGF2R imprinting from the Jurassic to the Quaternary. *Hum Mol Genet* 10:1721–1728
- Evans HK, Weidman JR, Cowley DO, Jirtle RL 2005 Comparative Phylogenetic Analysis of Bicap/Nnat Reveals Eutherian-Specific Imprinted Gene. *Mol Biol Evol* 22:1740–1748
- Suzuki S, Renfree M, Pask A, Shaw G, Kobayashi S, Kohda T, Kaneko-Ishino T, Ishino F 2005 Genomic imprinting of IGF2, p57(KIP2) and PEG1/MEST in a marsupial, the tammar wallaby. *Mech Dev* 122:213–222
- Weidman JR, Maloney KA, Jirtle RL 2006 Comparative phylogenetic analysis reveals multiple non-imprinted isoforms of opossum Dlk1. *Mamm Genome* 17:157–167
- Niemitz EL, Feinberg AP 2004 Epigenetics and assisted reproductive technology: a call for investigation. *Am J Hum Genet* 74:599–609
- Feinberg AP, Tycko B 2004 The history of cancer epigenetics. *Nat Rev Cancer* 4:143–153
- Feinberg AP 2004 The epigenetics of cancer etiology. *Semin Cancer Biol* 14:427–432
- Cruz-Correa M, Cui H, Giardiello FM, Powe NR, Hylind L, Robinson A, Hutcheon DF, Kafonek DR, Brandenburg S, Wu Y, He X, Feinberg AP 2004 Loss of imprinting of insulin growth factor II gene: a potential heritable biomarker for colon neoplasia predisposition. *Gastroenterology* 126:964–970
- Jirtle RL 2004 IGF2 loss of imprinting: a potential heritable risk factor for colorectal cancer. *Gastroenterology* 126:1190–1193
- Feinberg AP, Tycko B 2004 The history of cancer epigenetics. *Nat Rev Cancer* 4:143–153
- Falls JG, Pulford DJ, Wylie AA, Jirtle RL 1999 Genomic imprinting: implications for human disease. *Am J Pathol* 154:635–647

45. Waterland RA, Lin J-R, Smith CA, Jirtle RL 2006 Post-weaning diet affects genomic imprinting at the insulin-like growth factor 2 (Igf2) locus. *Hum Mol Genet* 15:705–716
46. Reik W, Dean W, Walter J 2001 Epigenetic reprogramming in mammalian development. *Science* 293:1089–1093
47. Morgan HD, Santos F, Green K, Dean W, Reik W 2005 Epigenetic reprogramming in mammals. *Hum Mol Genet* 14:R47–R58
48. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T 2000 Demethylation of the zygotic paternal genome. *Nature* 403:501–502
49. Santos F, Hendrich B, Reik W, Dean W 2002 Dynamic reprogramming of DNA methylation in the early mouse embryo. *Dev Biol* 241:172–182
50. Wolff GL 1978 Influence of maternal phenotype on metabolic differentiation of agouti locus mutants in the mouse. *Genetics* 88:529–539
51. Belyaev DK, Ruvinsky AO, Borodin PM 1981 Inheritance of alternative states of the fused gene in mice. *J Hered* 72:107–112
52. Rakyan VK, Preis J, Morgan HD, Whitelaw E 2001 The marks, mechanisms and memory of epigenetic states in mammals. *Biochem J* 356:1–10
53. Blewitt ME, Vickaryous NK, Paldi A, Koseki H, Whitelaw E 2006 Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. *PLoS Genet* 2:399–405
54. Anway MD, Cupp AS, Uzumcu M, Skinner MK 2005 Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308:1466–1469
55. Barker DJ 2004 The developmental origins of adult disease. *J Am Coll Nutr* 23:588S–595S
56. Allred CD, Allred KF, Ju YH, Virant SM, Helferich WG 2001 Soy diets containing varying amounts of genistein stimulate growth of estrogen-dependent (MCF-7) tumors in a dose-dependent manner. *Cancer Res* 61:5045–5050
57. Jiang YH, Bressler J, Beaudet AL 2004 Epigenetics and human disease. *Annu Rev Genomics Hum Genet* 5:479–510
58. Luedi PP, Hartemink AJ, Jirtle RL 2005 Genome-wide prediction of imprinted murine genes. *Genome Res* 15:875–884
59. McMillen IC, Robinson JS 2005 Developmental origins of the metabolic syndrome: prediction, plasticity, and programming. *Physiol Rev* 85:571–633