Anim. Reprod, v.7, n.3, p.103-116, Jul./Sept. 2010

Epigenetics and reproduction and the developmental origins of health and disease

K.M. Aagaard-Tillery[1,2,3], M.A. Suter[1], A. Harris[2], A. Abramovici[1], J. Cantu[1]

[1]Department of Obstetrics and Gynecology, Div. of Maternal-Fetal Medicine, Baylor College of Medicine, Houston, TX, USA.
[2]Bioinformatics Research Laboratory, Baylor College of Medicine, Houston, TX, USA.

Abstract

Over the past century, studies of development and reproductive biology have transcended our understandings of what constitutes heritability and the acquisition of phenotypic traits from one generation to the next. While our early research defined "particulate genetic inheritance" as a primary mechanism for the heritability of traits, more recent work in past decades in lower eukaryotes and early mammalian species have included "epigenetic" (or "upon the genome") modifications to the genomic backbone as a primary mechanism in the complex series of molecular interactions which ultimately enabling coordinate regulation of development. In recent years such investigations have evolved to focus on the role of *epigenetic modifications* to DNA and core histones in higher mammalian developmental processes. What are *epigenetic modifications*? While almost all cells of an individual bear near identical genomic constitutions, phenotype is ultimately determined by the gene expression profile. Gene expression is maintained by two major mechanisms: (1) transcription factors and post-transcriptional modifiers, and (2) epigenetic modifications, in particular DNA and core-histone modifications, that can be inherited during mitosis from one cell generation to another. This epigenetic code is essential in directing the tremendous number of gene expression changes that must occur for a cell to leave its pluripotent state and become fully differentiated to then function in adaptive homeostasis processes of the organism. It may therefore be stated that one's epigenetic signatures are the net outcome of genotype, developmental lineage, and environmental exposures. These epigenetic signatures are stable and/or heritable patterns of gene activity and expression that do not result from changes in the genomic sequence. Covalent modifications to histones (i.e., histone H3 acetylation and methylation) and DNA methylation (meCpG) are examples of such epigenetic events which collectively act as a "memory" to maintain gene expression profiles after cell division. Research is rapidly demonstrating the importance of the epigenetic code to normal human development as well as the burden of disease that occurs when the epigenetic code or machinery malfunctions. In essence, epigenetic modulation results in functional adaptations of the genomic response to the environment and is believed to play a fundamental role in early developmental plasticity. This review provides and overview on studies related to reproduction and epigenetic inheritance, many of which have arisen from the Developmental Origins of Adult Health and Disease fields.

Keywords: chromatin, developmental origins of disease, DNA methylation, histone modification, nutrition, pregnancy, tobacco.

Introduction

Genomic and epigenomic inheritance

Epigenomics overview

As schematically summarized in Fig. 1, while genomic DNA is the template of our heredity, it is the coordination and regulation of its expression that results in the wide complexity and diversity seen among organisms. In recent "postgenomic era" years, higher order architectural features (i.e., beyond primary sequence information) of the genome have been detected. Converging lines of evidence suggest that regulation of gene expression occurs by single nucleotide variance in gene regulatory regions (e.g., single nucleotide polymorphisms, or SNPs), epigenetic regulation (e.g., chromatin modifications, DNA methylation and RNA interference) and, most recently, large scale genomic structural variation (i.e., copy number imbalance through large insertions or deletions, as well as balanced chromosomal rearrangements; Eicherl *et al*., 2007; Estivill and Armengol, 2007; Pennisi, 2007). Collectively, structural rearrangements acting in concert with allelic polymorphisms are thought to contribute to the "*genomic*" landscape from which disease phenotypes arise. Alongside our SNP data as previously reported (Aagaard-Tillery *et al*., 2010), the high mutability of CNVs creates signature differences in mouse strains and may play a significant role in disease associated mutations (Egan *et al*., 2007).

It is becoming increasingly evident that epigenomics plays an equally important and possibly more prevalent role in the development of common and complex diseases as well as phenotypic traits (Aagaard-Tillery *et al*., 2008a; Bocock and Aagaard-Tillery, 2009; Segars and Aagaard-Tillery, 2009; Suter and Aagaard-Tillery, 2009). The concept term "epigenetics" was first coined by Conrad Waddington many decades ago and referred to "the interactions between genes and their products which bring phenotype into being" (Waddington, 1968). By way of overview, every normal and diseased cell has an epigenetic signature that is based on its genotype, developmental lineage, and environmental exposures in its history. These epigenetic signatures are stable and/or heritable patterns of gene activity and expression that do not result from changes in the genomic sequence (Holliday, 1987; Riggs *et al*., 1996).

[1]Corresponding author: aagaardt@bcm.tmc.edu

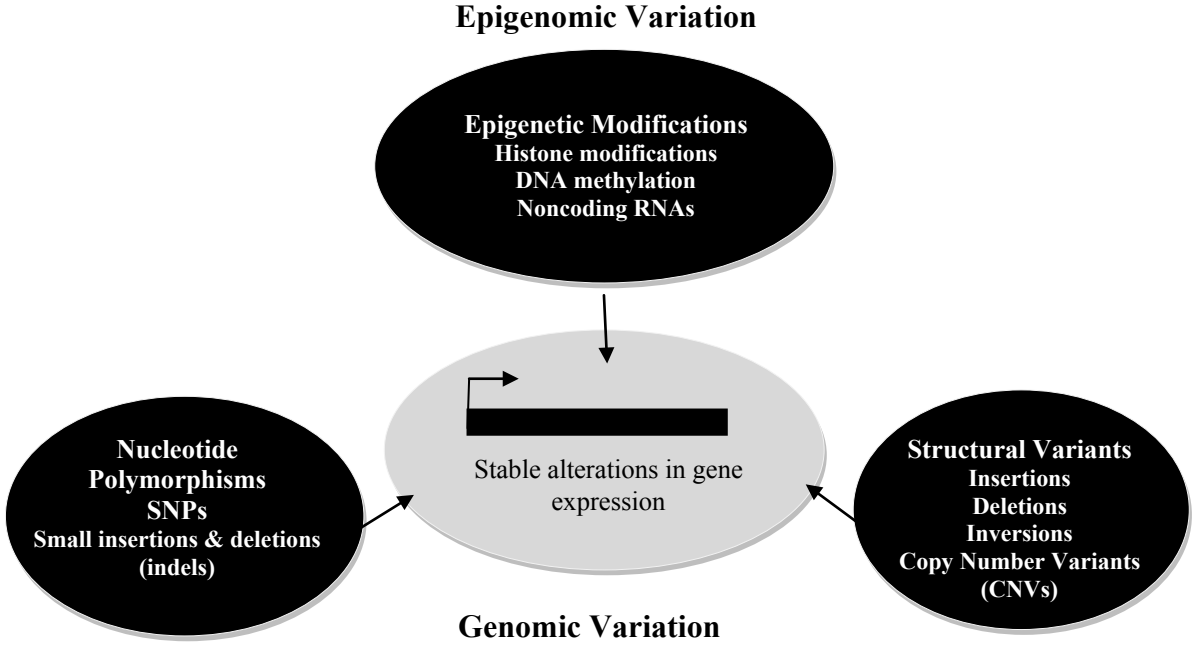

Figure 1. Genomic and epigenomic regulators of gene expression. Structural genomic variation comprises the nucleotide backbone of an individual. Epigenomic variation is comprised of epigenetic modifications, including both covalent histone modifications, DNA methylation, and non-coding RNAs. This review focuses on epigenetic modifications layered upon the genomic backbone as mediators of fetal growth and the developmental origins of health and disease.

Cells contain an epigenetic code and machinery to direct the huge number of normal gene expression changes that must occur for a cell to leave its pluripotent state and become a differentiated cell that comprises a tissue, participates in organogenesis, and then functions in adaptive homeostasis throughout the aging process of the organism. The epigenetic mechanisms identified to date include DNA methylation of CpG dinucleotides, RNA-associated gene silencing, chromatin remodeling, and post-translational, ATP-dependent histone modifications. In sum of this work, the epigenome is clearly a key factor in determining when and where genes are expressed and how they respond to the environment (Fig. 2). Despite significant effort and progress to date, the field of epigenomics is still in its infancy and an analogous effort to map the human epigenome will provide countless resources to scientists in the form of biomarkers of development, disease, environmental exposure, response to therapy, and a basic understanding of normal development and disease processes.

Histone modification and chromatin remodeling molecules as markers of chromatin status

Nucleosomes

All eukaryotes maintain their genome as a nucleoprotein complex, which consists of DNA wrapped around four histone proteins. The basic repeating unit of chromatin is the nucleosome, and the central core of the nucleosome consists of two copies each of four histone proteins. As schematically represented in Fig. 2, two copies of H3 and H4 join to form a histone tetramer, and two histone H2A/ H2B dimers form the histone octamer (Luger *et al*., 1997). Around the octamer, 147 base pairs of DNA wraps approximately 1.7 times in a left-handed superhelix to form the nucleosome. High resolution crystallography of nucleosome structures have revealed that each of the four histones contain a globular histone-fold domain, which is involved in dimer-tetramer interfaces within the nucleosome. Each histone also contains an N-terminal domain, called the histone "tail", which extrudes from the nucleosome surface. The histone tails do not contribute to the structure of the individual nucleosomes but rather maintain dimensional confirmation and allow for nucleosome stacking (Clayton *et al*., 2006; Luger, 2006).

Histone modifications as epigenetic determinants of chromatin structure

While nucleosomes represent the initial step in the formation of higher-order chromatin structures, histones maintain an epigenetic code that has a role in the regulation of gene expression, DNA replication, recombination, and repair (Clayton *et al*., 2006;

Nightingale *et al*., 2006; Bernstein *et al*., 2007). Each of these processes has been demonstrated in eukaryotic models to employ posttranslational modifications of chromatin structure in their regulation (Bernstein *et al*., 2005, 2006, 2007; Shi., 2007; Shi and Whestine, 2007). Histones have a protruding charged 15-38 amino acid N-terminus ("histone tail") that influences nucleosome assembly into higher order chromatin structure. In its condensed state, chromatin folds so the nucleosomes are stacked, a structural configuration not readily accessible to gene activation. However, covalent modifications (i.e., acetylation, methylation, phosphorylation, poly-ADP ribosylation and ubiquitination) of the H3 and H4 tails alter the interaction between histones and DNA to effect nucleosome interactions and higher order chromatin folding. These posttranslational covalent modifications regulate the contact between the octamer core and DNA, and determine DNA accessibility to transcription factor complexes (Fig. 2). The ability to store information appears to reside in the amino-terminal tails of the four core histones which are exposed on the nucleosome surface and are subject to enzyme-catalyzed posttranslational modifications of select amino acids, including lysine acetylation, lysine and arginine methylation, serine or threonine phosphorylation, lysine ubiquitination, lysine sumoylation, or glutamine ADP ribosylation (Bernstein *et al*., 2005, 2006, 2007; Shi and Whetstine, 2007). The globular portion of histones also contains amino acids that are subject to modification. In sum, more than forty years of work support the paradigm that acetylated histones are markers of transcriptionally active genes (Fig. 2). Ergo, it is generally accepted that methylation of distinct lysine and arginine residues along with hypoacetylation of lysine residues are markers of transcriptionally silent genes.

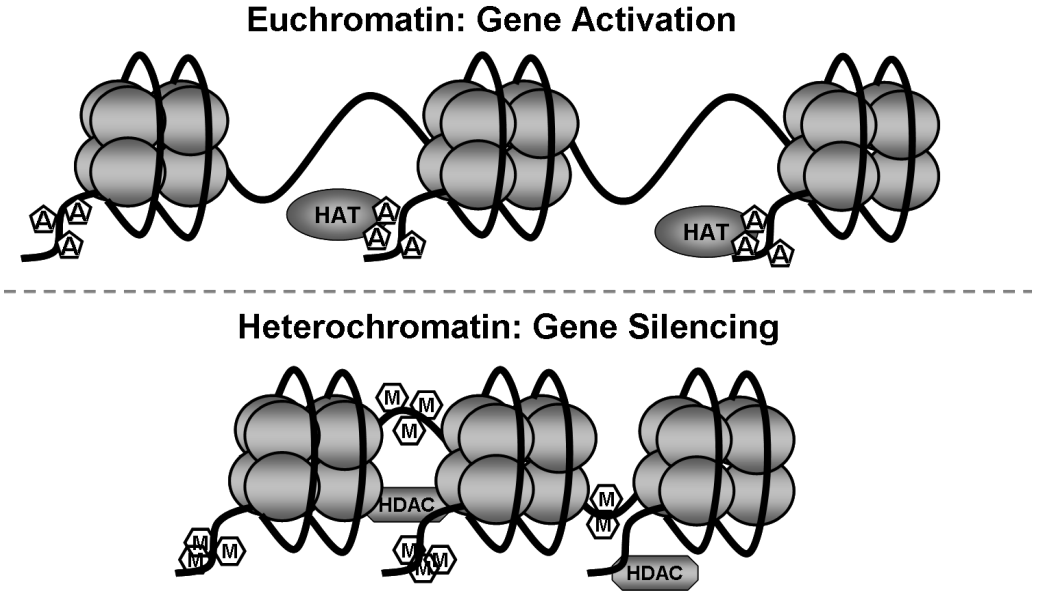

Figure 2. Characteristics of euchromatin and heterochromatin. Euchromatin is considered to be a more open chromatin state, generally characterized by hyperacetylation of the histone tails by HATs and a lack of methylated DNA. Heterochromatin, which is a compact chromatin state that inhibits transcription, is generally characterized by site specific methylation of the histone tails and regions of DNA methylation.

Histone methylation

Histone methylation occurs on select lysine (K) and arginine (R) residues. More than 50 SET domain-containing proteins are known to methylate lysine residues of histones (Shi, 2007). Lysines can be mono-, di-, or tri-methylated and patterns of histone H3 methylation are the most studied to date. Methylation of H3K4, H3K9, H3K27, H3K36, H3K79, and H4K20 has been linked to chromatin and transcriptional regulation as well as to the DNA damage response (Shi, 2007; Shi and Whetstine, 2007). Monomethylation of H3K27, H3K9, H4K20, H3K79, or H2BK5 at promoters, insulators, enhancers, and transcribed regions of human genes is linked to gene activation, whereas dimethylation of H3K9 or trimethylation of H3K27, H3K9, or H3K79 is linked to gene repression (Barski *et al*., 2007). However, and of interest to our proposal herein, a subsequently published paradigm-shifting paper suggested that, on a genome-wide scale, trimethylation of H3K4, acetylation of H3K9, 14, and RNA polymerase II occupy the promoters of most protein-coding genes in human ES cells as well as in differentiated tissues (Guenther *et al*., 2007). Only a subset of these protein-coding genes yield full-length transcripts and those that do also contain nucleosomes

with trimethylated H3K36 modifications, a hallmark of elongation. The remainder of the protein-coding genes experience transcription initiation but elongation does not occur. Although histone methylation has long been thought to be a permanent mark, the first histone demethylase, LSD1, was only recently identified in 2004 (Barski *et al*., 2007). Early data suggested that histone demethylases may have important roles throughout development including stem cell maintenance and differentiation, X-chromosome inactivation, imprinting, cell cycle regulation, cell differentiation, organogenesis, and DNA repair (Wolfe and Matzke, 1999; Barski *et al*., 2007; Beisel *et al*., 2007). There is scant evidence for histone methylation in initiating transcription per se (Barski *et al*., 2007).

Histone acetylation

In contrast to histone methylation, histone acetylation is recognized as one of the more prominent epigenetic marks leading to the directed initiation of gene expression and there are well characterized histone acetylases (HATs) and deacetylases (HDACs; Bernstein *et al*., 2005, 2006, 2007; Brickner *et al*., 2007). As we have reviewed above, acetylation of the N-terminal (ε) amino groups of specific lysine residues on the H3 tail is thought to induce an open chromatin confirmation that allows the transcriptional machinery to access promoters, enhancers, and insulators (Bernstein *et al*., 2005, 2006, 2007). Acetylation of lysines in histones has long been known to neutralize the positive charge between the negatively charged DNA nucleotide backbone, thereby promoting "chromatin relaxation". Thus acetylation of histones, particularly H3, targets and directs the epigenetic transcriptional machinery (Fig. 2).

As such, histone acetylation may be considered an important epigenetic "indexing system" demarcating transcriptional active and inactive chromatin domains in the eukaryotic genome. While a hierarchical paradigm has yet to be established in primates, collective evidence from other eukaryotes suggests that while di/trimethylation of H3-K4 increases transcription, di/trimethylation of H3-K9 and trimethylation of H3-K27 repress transcription. This likely occurs by alternately facilitating or preventing binding of the epigenetic machinery via unique recognition sequences (chromodomains) with high-specificity for methylation modifications on lysine residues. Similarly, alternate recognition sequences (bromodomains) of effecter molecules associated with transcriptional activation have high affinity for H3 acetylation motifs at H3-K9 and H3-K14. Together, these data suggest that histone modifications are extremely dynamic and highly regulated, but much additional work is necessary to demystify the histone code particularly with respect to development.

DNA methylation

DNA methylation

Most methylation in mammals occurs at the number five carbon of the cytosine pyrimidine ring. Genomic methylation patterns are propagated during cell division by DNA methyltransferases. These enzymes are categorized into maintenance (DNA Methyltransferase 1, Dnmt1) and de novo (Dnmt3a and Dnmt3b) methyltransferases, but all three enzymes can act in both modes (Bestor *et al*., 1988; Bestor, 1992; Chen *et al*., 2003). One of the most important sites of gene regulation by DNA methylation are CpG-enriched regions associated with promoters called "CpG islands" (Takai and Jones, 2002). DNA methylation acts with other enzymes to covalently modify histones to cause gene silencing and to maintain a repressive chromatin state. It is thought that DNA methylation acts as a memory to maintain gene expression profiles after cell division, thus defining the state of cellular differentiation. It has long been recognized that DNA methylation is inversely related to both the expression of developmentally regulated genes and the potency of cells.

DNA methylation as a common endpoint to covalent modifications in chromatin structure

As detailed above, multiple lines of evidence from lower eukaryotes and mammalian cell lines have shown that gene expression in general, and transgene expression in particular, is regulated in part by DNA methylation, posttranslational histone modifications, and timing of replication. *In vivo*, DNA methylation is associated with diverse biologic processes such as X chromosome inactivation, genomic imprinting, silencing of mobile genetic elements, and the developmental regulation of tissue-specific genes. However, it remains unknown whether DNA methylation *per se* is required for the establishment or maintenance of the silent chromatin state, or whether (re)activation of the DNA methylation machinery occurs in response to maintained histone modifications with DNA replication events.

Two independent investigators have recently addressed this issue with mammalian transgenes and transgenic promoters (Feng *et al*., 2006; Yan and Boyd, 2006). Feng *et al.* (2006) concluded that DNA methylation within the transgene is unique among epigenetic marks in that it serves to confer epigenetic memory to prevent gene reactivation. Conversely, Yan and Boyd concluded that DNA methylation is indeed sufficient, rather because it results from histone modifications localized to discrete sites in the mammalian epigenome. Thus, two independent hypotheses arise: either (1) DNA methylation, but not necessarily the histone code, is self-replicating and is able to confer intrinsic epigenetic memory, or (2) the histone code confers alterations in DNA methylation, and therefore it is not necessary *per se* since it may be

generated in response to appropriate (conserved) histone code alterations.

Epigenetic processes in development

DNA methylation and early embryogenesis

The genome of mammals is progressively demethylated during the preimplantation period, which is believed to be important for reprogramming the genome to a pluripotent state in the preimplantation embryo (Mayer *et al*., 2000; Oswald *et al*., 2000). After implantation, DNA is progressively remethylated and cell-type-specific DNA methylation patterns are established as cells differentiate (Ehrlic *et al*., 1992). Imprinted genes represent a small subset of methylated genes in which allele-specific expression correlates with allele-specific methylation, with some genes showing methylation on the paternal allele and others on the maternal allele (McGrath, 1984; Surani *et al*., 1984). The methylation of these imprinted genes does not change during development, and correct imprinting is believed to be essential for normal development and differentiation.

Two additional characteristics of epigenetic signatures make them of interest in this review. First, it is understood that the mammalian genome is demethylated during preimplantation and remethylated at differentially methylated regions (DMRs) during differentiation. Animal models have previously demonstrated that significant *in utero* alterations related to fetal growth and metabolism are associated with epigenetic perturbations to allow for developmental plasticity. Second, it is increasingly accepted that DNA methylation patterns are both dependent on environmental exposures early in development and influenced by inborn factors, including the DNA sequence (Kerkel *et al*., 2008; Foley *et al*., 2009). However, it remains largely unknown whether these DNA methylation events in adults are stable or are in fact dynamic and subject to ongoing influence by virtue of phenotype or exogenous exposures. Importantly, CpG methylation is a biochemically stable modification of a particularly stable macromolecule (DNA), making this modification extremely attractive as a potential biomarker. As we will discuss further below, we have focused on this observation in some of our work on *in utero* tobacco exposure as the environmental influence, which may interact with inborn and pre-existing fetal genetic factors to ultimately render susceptibility to fetal growth restriction.

Epigenetic regulation and the developmental origins of disease

Modification and plasticity of DNA methylation

Although DNA methylation is considered a relatively stable epigenetic mark, an example of potential plasticity of DNA methylation and that environmental changes can affect methylation status are demonstrated by experiments with the agouti viable yellow (A^{vy}) strain of mice. Expression of the dominant agouti viable yellow (A^{vy}) allele causes a yellow coat, obesity, diabetes, and cancer predisposition. Among isogenic A^{vy}/a mice, individual variation in methylation at the A^{vy} promoter leads to extensive phenotypic variation. Pseudoagouti A^{vy}/a females (brown coat, A^{vy} silenced by hypermethylation) give birth to a higher percentage of pseudoagouti offspring than do yellow A^{vy}/a mothers due to heritable epigenetic silencing of the A^{vy} promoter (Waterland *et al*., 2006). There is also a shift to the pseudoagouti phenotype in offspring born to mothers provided with a diet high in methyl donors and cofactors (Waterland *et al*., 2006) demonstrating that maternal environment can affect establishment of DNA methylation at specific loci (Waterland *et al*., 2006).

Histone modifications and the developmental origins of disease

According to the fetal or developmental origins of adult disease hypothesis, perturbations in the gestational milieu influence the development of diseases later in life through the static reprogramming of gene expression via alterations in chromatin infrastructure (Fig. 3). Others have previously shown that uteroplacental insufficiency induced through bilateral uterine artery ligation of the pregnant rat dam results in asymmetrical IUGR and, similar to the human, causes demonstrable postnatal disease; these alterations are associated with modifications of the fetal epigenome (changes in postnatal gene expression (Lane *et al*., 2001; Fu *et al.*, 2004, 2006; MacLennan *et al.*, 2004; Ke *et al.,* 2005, 2006).

Emerging answers to questions in epigenetics research and development

Epigenetic signatures in relationship to the genome and fetal development

While recent work has shown that DNA methylation, histone modifications, and other epigenetic changes play crucial roles in many biological processes such as gene expression, chromatin accessibility, DNA replication, and imprinting, a paucity of data exists on their role in modulating adult metabolic disease states and obesity. There are significant technical challenges which have inhibited significant advances in this arena.

First, tissue and developmental lineage-specific reference data for epigenomic demarcations in humans has been lacking. Second, characterization of profiling histone modifications requires significant amounts of fresh cells or tissue. Third, since by definition epigenetic signatures are tissue-specific, studies aimed

at examining their relationship to fetal development is problematic at an attainment and ascertainment level. For example, undoubtedly appropriate fetal growth velocity involves coordinated epigenetic modifications to chromatin and DNA in liver, skeletal muscle, adipose tissue and the hypothalamus. However, assaying such tissues is not possible. Fourth, from existing GWAS data (such as the Wellcome Trust) it is possible - if not probable - that many DNA polymorphisms that affect human complex disease are in fact regulatory SNPs or copy number variants (CNVs), but generally not missense changes in the coding regions of genes. These sites are therefore likely to be "hot spots" for detection of DNA methylation changes and are potentially detectable by approaches such as ours of allele-specific methylation (Pastinen and Hudson, 2004; Bertram, 2005; McMinn *et al.*, 2006; Pastinen *et al.,* 2006; Kerkel *et al.*, 2008; Rollins, 2008; Suzuki, 2008).

With this in mind, we will now focus on two arenas where we have had recent success in addressing such issues: a non human primate model of maternal high fat diet exposure, and a human model of genomic and epigenomic interplay in response to maternal smoking.

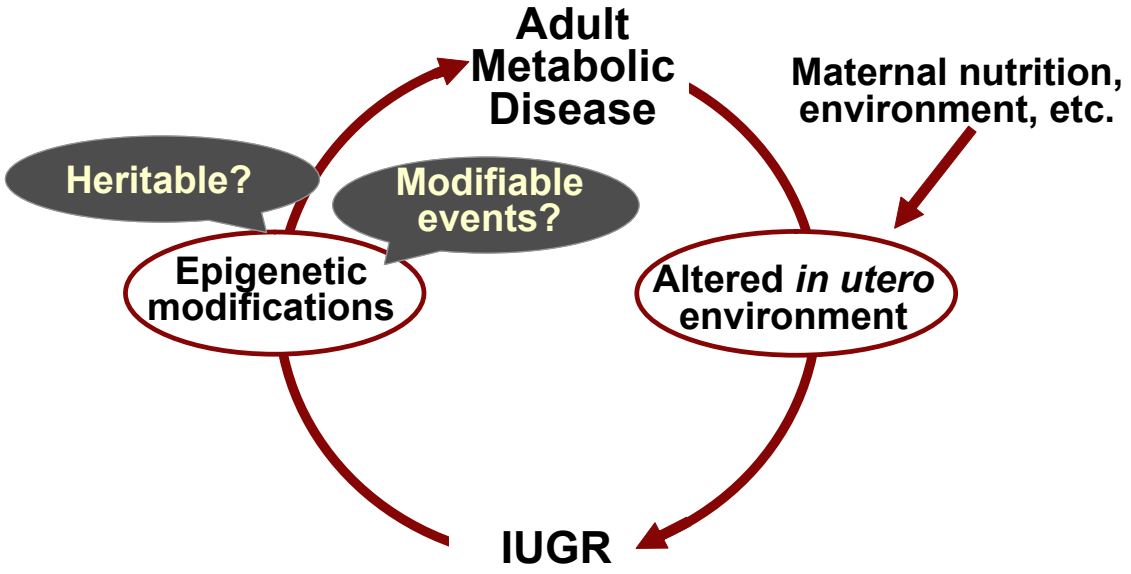

Figure 3. Schematic representation illustrating the interactions between the in utero environment, fetal development and adult phenotype. The fetal *in utero* environment is affected by maternal environmental conditions, especially maternal nutrition. This can result in intrauterine growth restriction (IUGR), which in turn causes the modification of the developing fetal epigenome through mechanisms such as DNA methylation, histone methylation/acetylation, and RNA interference (not reviewed here). These epigenetic modifications can lead to the propensity to develop adult metabolic disease, which again affects the *in utero* environment in a potentially self perpetuating cycle. It is still unknown to what extent these epigenetic changes are heritable (if so, how many generations?) and if they are modifiable.

Non human primate model of maternal high fat diet exposure

Collective data from models of nutritional constraint and uteroplacental insufficiency suggest that the gestational milieu influences the postnatal phenotype to render a susceptibility to childhood obesity through metabolic reprogramming (Fig. 3; Aagaard-Tillery *et al*., 2008a; Bocock and Aagaard-Tillery, 2009; Segars and Aagaard-Tillery, 2009; Suter and Aagaard-Tillery, 2009). Whether this is due to maternal diet (high fat, caloric dense), maternal phenotype (obesity), or a combination of the two has remained poorly understood. While multiple pathways converge to epigenetically modify the genome, an exploding volume of literature and our recent data (Aagaard-Tillery *et al*., 2008a, 2010; Cox *et al*., 2009; Suter *et al*., 2010) suggests that histone modifications and DNA methylation are competent sensors of the metabolic environment and are likely mechanistic explanations.

Along these lines, in a primate model we have shown that maternal diet can alter the fetal phenotype (Aagaard-Tillery *et al*., 2008a; Cox *et al*., 2009; McCurdy *et al*., 2009) and we demonstrate that such alterations modify the epigenome to result in predictable alterations in the maternal and fetal transcriptome and metabolome. However, comprehensive epigenome-wide interrogations with layered functional gene expression data have not yet been performed as there are significant technical challenges which have inhibited significant advances in this arena. First, tissue and developmental lineage-specific reference data for epigenomic demarcations in primate models has been lacking. Second, characterization of profiling histone modifications requires significant amounts of fresh cells

or tissue. Third, since by definition epigenetic signatures are tissue-specific, studies aimed at examining their relationship to later in life disease is problematic at an attainment and ascertainment level. For example, undoubtedly metabolic syndromes involve coordinated epigenetic modifications to chromatin and DNA in liver, skeletal muscle, adipose tissue and the hypothalamus. However, assaying such tissues across the lifetime of an individual is not possible. It has yet to be determined whether epigenetic signatures in placenta can be employed at a genome-wide level to predict disease. Our primate model is uniquely poised and developed to address these issues (Aagaard-Tillery *et al*., 2008a; Cox *et al*., 2009; McCurdy *et al*., 2009).

Prior to our work in non-human primates, a number of crucial pieces of evidence integral to our understanding of Barker's hypothesis as it relates to human disease were missing. First, modifications of the primate fetal epigenome in response to perturbations in the gestational milieu were assumed but unproven. This is of teleological importance, as the physiologic stress "felt" by a fetus may differ in animals designed to carry a singleton gestation rather than litter; the fetal epigenome would be anticipated to modify accordingly. Second, it was unknown whether nutritional overabundance of epidemic proportions (e.g., maternal obesity) would profoundly modify fetal chromatin structure similar to that observed in models of nutritional constraint.

We have demonstrated that a maternal high fat diet results in lysine site-specific acetyl modifications of fetal hepatic chromatin structure (Aagaard-Tillery *et al.*, 2008a; Cox *et al.*, 2009) alterations in expression and function of related key components of the epigenetic machinery accompany these modifications. Taken together, our observations imply significance for several reasons. First, these data are a novel demonstration that the fetal primate epigenome is modified in response to a clinically relevant and evolutionarily recent *in utero* nutritional stress: a caloric dense, high fat maternal diet leading to obesity. Second, at a molecular level, we observe that the fetal chromatin structure is not globally altered but rather modified via site-specific alterations in H3 acetylation (Fig. 3).

We find it of interest that we did not observe global alterations in histone marks (Aagaard-Tillery *et al.*, 2008a). We hypothesize that this may be due to the importance of maintaining site specificity in the fetal primate epigenome. It is possible that fetal hepatic chromatin resides predominantly as inert heterochromatin, and is reliant upon stress-induced deimination or demethylation for recruitment of histone acetylases to ultimately render chromatin accessible for transcription (Guenther *et al.*, 2007). Following gene activation, resumption of a repressed methylated state would employ HDAC-dependent recruitment of the methyltransferases to provide the methylation "marks" integral to the control of epigenetic events. Our observation that an ordered change in the ratio of modified acetyl variants occurs under conditions associated with a marked decrease in HDAC1 expression and function alongside relatively minor alterations in site-specific methylation is certainly consistent with this notion. Moreover, organisms are classically assumed to adapt to any given fetal stress overtime with a patterned, dichotomous site-specific histone modification (i.e., AcH3K14 hyperacetylation and H3K27me2me3) which is integral to regulation of gene-specific transcription. However, in the face of a *de novo* evolutionary event (nutritional abundance), site-specific acetylation modifications of the histone code likely enable regulation of genes integral to glucose and lipid homeostasis. In this fashion, modification of the epigenome is poised to occur via heritable site-specific covalent modifications rather than a potentially disastrous global modification of fetal chromatin structure.

While our non human primate studies provided strong evidence for an influence of the maternal diet on the fetal epigenome, with a focus on histone modifications, it remained unclear whether there was a role for DNA methylation changes (outside of imprinting models). In these studies, we chose to examine the concomitancy of genomic susceptibility with epigenetic modifications in response to in utero tobacoo exposure. In such a fashion, we were attempting to define a potential mechanism for understanding the role of environmental influence on the genomic backbone in mediating disease.

Maternal tobacco use, fetal genomic variation, and developmental programming of fetal growth

Rationale for genomic studies

Maternal tobacco use has been identified in multiple population-based studies as the strongest modifiable risk factor for intrauterine growth restriction (IUGR) and multiple other adverse pregnancy outcomes (Peacock *et al*., 1998; Cnattingius *et al*., 1999; Secker-Walker and Vacek, 2003; Hammoud *et al*., 2005; Aagaard-Tillery *et al*., 2008b). However, while many fetuses are exposed to tobacco smoke, not all experience adverse outcomes. This discrepancy cannot be accounted for by dose effect alone and smoking-related weight reduction has historically been considered to be largely independent of other maternal and fetal risk factors which influence birth weight (Aagaard-Tillery *et al*., 2008b). Thus, current efforts aimed at understanding the potential genetic and metabolic basis of this variable susceptibility to tobacco smoke exposure are of importance in perinatal medicine.

Tobacco metabolism and known polymorphisms

Mechanisms leading to growth restriction following *in utero* tobacco exposure are poorly understood, but have generally often been attributed to

chronic fetal hypoxia. Nicotine, a principal alkaloid of tobacco smoke, has been shown to mediate constriction of the intrauterine vessels and result in increased apoptosis of placental syncytiotrophoblasts. Nicotine, cotinine, and potentially harmful DNA adducts (metabolic products of polycyclic aromatic hydrocarbons; PAH) are known to cross or collect in the placenta of smokers (Daube *et al*., 1997; Gladen *et al*., 2000). Thus while chronic hypoxia may be a mediator of growth restriction in response to in utero tobacco exposure, it is also plausible that the discrepant variation in fetal susceptibility to smoking-related growth restriction results from fetal and/or maternal metabolic gene polymorphisms.

Of the over 4000 substances in tobacco smoke, PAH compounds together with nitrosamines comprise likely carcinogenic species in tobacco smoke. The majority of chemical carcinogens are metabolized in a sequential series of two-phase enzymatic metabolic reactions. Phase I enzymes such as cytochrome P450 metabolically activate PAH compounds into oxidized derivatives, resulting in reactive oxygen intermediates capable of covalently binding DNA to form adducts (Fig. 4). In turn, these reactive electrophilic intermediates can be detoxified by phase II enzymes, such as the glutathione S-transferase (GST) family, via conjugation with endogenous species to form hydrophilic glutathione conjugates which are then readily excretable. Thus the coordinated expression of these enzymes and their relative balance may determine the extent of cellular DNA damage and related development of adverse outcomes. CYP1A1 is a phase I metabolic enzyme which encodes the aryl hydrocarbon hydroxylase (AHH) enzymes responsible for the activation of the PAH compounds to their potentially harmful reactive intermediates. Following high-affinity binding of PAH compounds to their intracellular aryl hydrocarbon (AH) ligands, the complex is translocated to the nucleus where it dissociates then heterodimerizes to form a DNA binding complex (AH:ARNT) to modulate chromatin disruption and regulate induction of CYP1A1 expression (The CYP1A1 $Ile_{462}Val$ (AA > AG/GG) allele carriers exhibit higher levels of CYP1A1 enzymatic activity and inducibility, and smokers who carry this variant have increased cellular PAH-DNA adducts. We have shown that maternal smoking induces placental expression of CYP1A1 (Suter *et al*., 2010).

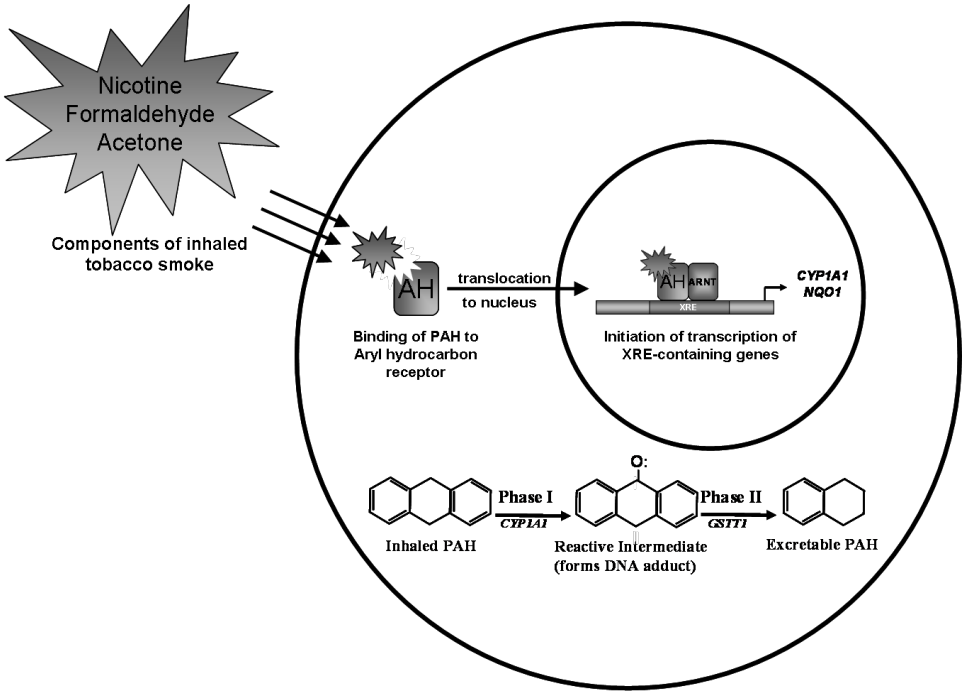

Figure 4. Cellular processing of PAH compounds. The PAH found in tobacco smoke are metabolized in a sequential series of two-phase enzymatic metabolic reactions. Upon entering the cell these compounds are recognized by ligands which bind the PAH, translocate to the nucleus, and initiate transcription of Phase I and Phase II enzymes which contain an XRE in their promoter. Phase I enzymes such as CYP1A1 activate PAH compounds into reactive oxygen intermediates capable of covalently binding DNA to form adducts. Phase II enzymes, such as members of the glutathione S-transferase (GST) family, metabolize the reactive species into hydrophilic glutathione conjugates which are then readily excretable.

Polymorphisms leading to enzymatic inactivity in the phase II metabolic enzyme glutathinone S-transferase theta 1 (GSTT1), e.g., GSTT1(del), are prevalent and have been extensively studied in the context of individual susceptibility to tobacco-mediated carcinogenesis, albeit with variable attributable risk associations. Theoretically, any combinatorial association of increase Phase I activity (e.g., increased expression of CYP1A1 via functional polymorphisms) in combination with decreased Phase II activity (e.g., decreased GSTT1 expression) may yield increased susceptibility to tobacco-related adverse outcomes. Association of these metabolic gene polymorphisms with smoking behavior in healthy controls have also been attempted in large population-based sample banks, with no associations observed for polymorphisms in the genes encoding phase I nor phase II enzymes (Mooney *et al*., 1997). However, variable expression of alternate cytochrome P450 enzymes (e.g., CYP2A6) have been shown to modify daily cigarette consumption (Oscarson *et al*., 1999; Hong *et al*., 2003). CYP2A6 is a highly polymorphic allele, and functions as the rate-limiting enzyme in the metabolism of nicotine to cotinine. Individuals with diminished activity of CYP2A6 activity at the CYP2A6*2 allele (CYP2A6 $Lys_{160}His$ T > A) inherit the slowest metabolism of nicotine and have been associated with lower cigarette consumption, shorter duration of smoking, and increased ability to quit smoking.

For these reasons, we hypothesized that maternal and/or fetal metabolic gene polymorphisms would alter the tobacco-related risk of adverse pregnancy outcomes (Fig. 4). Utilizing prospectively acquired biologic samples from a multi-institutional study, we assessed whether the functional polymorphisms associated with increased formation of carcinogenic adducts (phase I CYP1A1 $Ile_{462}Val$ polymorphism), inability to excrete reactive intermediates along the phase II metabolic pathways (GSTT1(del)), or altered metabolism of nicotine to cotinine (CYP2A6 $Lys_{160}His$) might account for the varying susceptibility to tobacco-mediated adverse pregnancy outcomes.

To summarize our now published work, the association of smoking with different outcomes in the full cohort of 1004 women were as anticipated with respect to smoking and as we have previously reported in other data sets (Aagaard-Tillery *et al*., 2010). Overall there were 99 patients with SGA < 10^{th} percentile and 46 < 5^{th} percentile, 150 patients who delivered preterm, 63 who suffered a pregnancy loss, 150 with preterm delivery, and 285 with at least one of the events making up the composite adverse outcome. Since there were only eight cases of abruption, further analysis of this as a separate outcome was not undertaken. Smoking was significantly associated with an increased risk of SGA < 10^{th} percentile (14.4% SGA among smokers and 8.2% among non-smokers, Odds Ratio [OR] 1.9, 95% Confidence Interval [CI] 1.2 – 2.9; P = 0.004). Smoking was not significantly associated with an increase in SGA < 5^{th} percentile (OR 1.7, 95% CI 0.9 – 3.2, P = 0.07), although the point estimate of the odds ratio was very similar; this was likely due to the smaller numbers. Birth weight (mean for smokers 3115 ± 657 g versus 3243 ± 631 g, P = 0.009) and birth weight ratio (mean for smokers = 0.97 ± 0.14 compared with 1.01 ± 0.15 for non-smokers, P = 0.0004) were significantly lower among the smokers. Smoking was marginally associated with an increased risk of the composite adverse outcome (OR 1.3, 95% CI 1.0 – 1.7). P = 0.05. However, smoking was not associated with pregnancy loss, preterm birth or gestational age at delivery.

The availability of DNA and the distribution of genotypes were as anticipated. Of the 1004 maternal samples, 213 (21%) were discarded for an inability to determine the GSTT1 allelic deletion variant or for poor quality DNA. For CYP1A1 and CYP2A6, only 45 (4.5%) and 55 (5.5%) samples could not be genotyped respectively. Of the corresponding 1004 conceptuses, DNA samples were available for 772 (76.9%). Of these, GSTT1 deletion could not be determined in 197 samples (25.5%), and CYP1A1 and CYP2A6, in 157 (20.3%) and 205 (26.6%) samples respectively. The distribution of maternal and fetal genotypes was consistent with that reported for the general population and did not demonstrate bias with respect to maternal tobacco use. Since the detected prevalence of the CYP1A1 and CYP2A6 recessive alleles were uncommon in the study population and prior studies have not supported a recessive model, the AG/GG (CYP1A1) and AT/AA (CYP2A6) genotypes were combined in the data analyses.

Univariate analyses of the relationship between smoking and the dichotomous adverse outcomes by maternal and fetal genotype were completed. Within each genotype, the odds ratios and 95% confidence intervals for the association of smoking with the adverse outcome is given. The P-values showed that despite different magnitudes and even different directions of odds ratios between the two genotypes of each polymorphism, there was no statistically significant interaction between any maternal or fetal genotype and smoking. This indicates that maternal and fetal genotype did not significantly modify any association between smoking and adverse outcome. Odds ratios for each less common genotype with the adverse outcome were then calculated, adjusting for smoking status in addition to confounders. When the effect of the simultaneous presence of maternal and fetal gene polymorphisms at each of the three alleles were examined for association with SGA 10^{th}%, fetal GSTT(del) was unique in its observed association with growth restriction in combinatorial allelic models. For example, maternal GSTT1(del) with a fetal GSTT1 non-deletion did not show an increased risk of smoking-associated growth restriction (OR 1.97, 95% CI 0.40 – 9.71), while fetal

GSTT1(del) with maternal GSTT1 non-deletion (OR 6.38, 95% CI 1.30 – 31.41), fetal GSTT1(del) with CYP1A1 AA (OR 5.16, 95% CI 1.42 – 18.7), and fetal GSTT1(del) with CYP2A6 TT (OR 5.38, 95% CI 1.49 – 19.46; Aagaard-Tillery *et al*., 2010).

Complete univariate and multivariate analyses of birthweight and birthweight ratio by smoking status and maternal/fetal GSTT1 genotype was performed. There was no interaction between genotype and smoking except for fetal GSTT1 (P = 0.02). The mean difference (Δ = 0.03) in the birthweight ratio between smokers and non-smokers among GSTT1 non-deletion fetuses was non-significant (P > 0.5), whereas the mean birthweight ratio difference (Δ = 0.09) between smokers and non-smokers among GSTT1(del) fetuses was significant (P = 0.0002). After adjusting for race, maternal BMI, and multiparity versus primiparity, fetal GSTT1(del) persisted as an effect modifier for the relationship between smoking and birthweight ratio (corresponding to an additional reduction (Δ = 0.07) between tobacco exposed and non-exposed fetuses for the GSTT deletion group reduction, P = 006). Although in univariable analysis, there was no significant interaction term between GSTT1 fetal genotype and smoking (P = 0.14), after adjusting for race, maternal BMI, and multiparity versus primiparity, the interaction term was significant (P = 0.002). This corresponded to an additional 256 reduction between tobacco exposed and non-exposed fetuses for the GSTT deletion group (Aagaard-Tillery *et al*., 2010).

Summary of fetal tested polymorphism studies

Our secondary analysis recognizes that there are fetal genomic underpinnings which render susceptibility to tobacco-mediated fetal growth restriction. The implications of our preliminary genomic findings are two-fold. First, our data illustrate that a fetal metabolic gene (GSTT1) which is integral in the excretion of reactive intermediates of aromatic hydrocarbons modifies fetal growth specifically in response to in utero tobacco exposure. These findings imply that tobacco metabolites may reach the fetus and thus modify fetal growth if not excreted. Second, our proposed studies aimed at illuminating the complex interplay of genomic-epigenomic-environmental interactions may help dissect multifactorial etiologies and identify at-risk populations for the common adverse pregnancy outcomes.

In utero tobacco exposure epigenetically modifies placental CYP1A1 expression

Rationale for epigenetic studies

As noted in the above section, we have recently and significantly expanded the scope of other authors' original analysis with paired maternal and fetal samples from a large, prospective study conducted through the NICHD Maternal-Fetal Medicine Units Network (Aagaard-Tillery *et al*., 2010). Specifically, we performed blinded genotyping for known functional allelic variants of CYP1A1 (Ile_{462}ValAA > AG/GG), GSTT1(del), and CYP2A6 (Lys_{160}HisT > A) in smokers and their offspring alongside 1:1 matched controls. In our analysis, deletion of fetal GSTT1 was singularly observed to significantly reduce the fetal birth weight ratio among smokers (P = 0.02, for interaction). However, our study failed to fully account for susceptibility to fetal growth restriction per se. Thus, it remained a formal possibility that non-allelic encoded dysregulation in the expression of these integral genes (or their metabolic pathways) may play a significant role in modifying fetal growth in response to maternal tobacco use.

In considering potential candidates, two lines of evidence led us to focus on regulation of expression of the Phase I CYP1A1 gene. First, aromatic hydrocarbon emissions are derived from both combustion of fossil fuels (coal, diesel, and gasoline) and environmental tobacco smoke. Multiple population-based analyses have demonstrated that the risk of fetal death, premature birth, and low birth weight is significantly higher for those with high prenatal ambient PAH exposure from all sources. Extension of these studies to include direct exposure measures (i.e., quantitation of PAH level by personal air monitoring) reveal significant interactions between maternal CYP1A1 haplotype and exposure to hydrocarbons on the detected level of PAH-DNA adducts present in cord blood (Wang *et al*., 2008). Second, in both human and animal models, environmental tobacco smoke induces aryl hydrocarbon hydroxylase activity and placental expression of CYP1A1 (Huel *et al*., 1989).

Emerging evidence has shown that in addition to genomic base pair differences, gene expression can be silenced by non-allelic mechanisms including epigenetic influences such as covalent modifications of histones and DNA methylation (as discussed earlier). Along these lines, other authors have previously observed that CYP1A1 is inducible in its placental expression among smokers and that well-characterized xenobiotic response elements (XREs) in the proximal promoter are differentially methylated at CpG islands in lung tissue of smokers (Anttila *et al.*, 2003). Given these published observations of others alongside our prior observations we hypothesized that non-allelic modulation of CYP1A1 expression may contribute to risk of adverse pregnancy outcomes. Since hypermethylation in key gene regulatory sequences at CpG islands is generally associated with gene silencing, we sought to compare placental gene expression of multiple CYP family members among gravidae. In order to specifically characterize the methylation status of the CYP1A1 proximal promoter, bisulfite modification and sequencing of the entirety of the 1 kb promoter

(containing 59 CpG sites) was performed. We thereafter correlated site-specific methylation with placental CYP1A1 expression.

Region I of the CYP1A1 proximal promoter in placenta is hypomethylated in smokers compared with non-smokers. CYP1A1 expression is regulated by an aryl hydrocarbon receptor (AhR; Fig. 4). The AhR binds the PAH from tobacco smoke in the cytoplasm, translocates to the nucleus to form a heterodimer with ARNT, and binds tightly and specifically to the XRE located within the proximal promoter of CYP1A1. The methylation sites of the CYP1A1 proximal promoter region have been previously characterized in lung tissue from smokers and non-smokers (Anttila *et al.*, 2003). We amplified four fragments of the CYP1A1 promoter region: Region I (-1411 to -1295), Region II (-1295 to -1006), Region III (-583 to -395) and Region IV (-395 to -228). In total, this proximal promoter contains a total of 59 CpG sites which have the potential to undergo differential methylation. In order to fully characterize the methylation status of each of these sites in response to maternal tobacco use, we isolated placental genomic DNA from a total of 15 smokers and 19 non-smokers. The CYP1A1 promoter region was amplified and cloned after sodium bisulfite treatment to determine CpG methylation across four primed regions (Regions I-IV), including the XRE transcriptional binding element in Region I. Each PCR reaction was subcloned and transformed into E. coli. For each subject we obtained a minimum of four reads per region, confirming with our sequencing data complete conversion in the bisulfite reaction (data not shown). The total percent methylation for each region was calculated for both the smokers and non-smokers, and differences were compared by the two-tailed Student t-test. The first core primed region (Region I) was the only primed region to contain an XRE transcriptional binding element and was unique in significant rate of methylation in smokers compared with non-smokers (55.6 vs 45.9% meCpG, P = 0.027). In support of previously published data utilizing primary lung tissue partial or no methylation was observed in placentas from smokers versus nonsmokers in primed regions II-IV. Interestingly Region I contains an XRE which is known to regulate transcription of CYP1A1.

Expression of CYP1A1 is inversely correlated with methylation status of Region I

Methylation of CpG sites in DNA is generally considered to correlate with a decrease in transcription. However, the direct evidence for true correlations in complex mammalian systems is limited. Given our differential methylation surrounding the XRE element in Region I, we therefore sought to better correlate placental CYP1A1 expression with the level of site-specific methylation of Region I. To do so, we plotted the relative expression level of CYP1A1 against the percent methylation for each region of the CYP1A1 promoter in both smokers and non-smokers and interrogated the relationship with bivariate correlations (Pearson's correlation for variance) employing a two-tailed test for significance. We found that the percent methylation of Region I inversely correlates with expression level (r = -0.737, P = 0.007); this correlation held true regardless of maternal smoking behavior. In a linear regression model controlling for the potential covariates of fetal gender and maternal comorbidities, percentage CpG methylation in Region I independently predicted CYP1A1 expression (data not shown). Moreover, there was no correlation between methylation status of Regions II-IV and placental CYP1A1 expression (data not shown).

Summary and implications of perinatal tobacco exposure studies

First, these data provide evidence that fetal homozygous deletion of the singular phase II PAH gene integral to excretion of DNA adduct forming reactive intermediates (GSTT1) significantly and specifically modifies fetal growth patterns in response to maternal smoking. These findings persisted in multiple allelic interaction models to suggest an interaction between the fetal metabolic gene GSTT1, maternal smoking, and modification of birth weight. Of note, 18-22% of the U.S. population carries a homozygous deletion of GSTT1. As discussed, phase I gene-products, such as CYP1A1, are integral in metabolic activation of PAH compounds into oxidized derivatives, resulting in reactive oxygen intermediates capable of covalently binding DNA to form adducts; as a balance to such intermediary forming reactions, conjugation with endogenous species to form hydrophilic glutathione conjugates which are then readily excreted occurs.

Second, we have built on these observations and demonstrated that increased placental CYP1A1 expression was specifically and significantly associated with hypomethylation of the CYP1A1 promoter region in smokers compared with non-smokers. Region I, which contains an XRE element that is involved in regulation of CYP1A1 expression, was the only region which demonstrated significant differential methylation within the proximal promoter. This association held true within an individual, as there was a significant correlation between CYP1A1 expression and Region I hypomethylation. With this preliminary data we present the first evidence that maternal smoking alters gene-specific DNA methylation in human placenta. Because of the limitations of biologic material we did not assay for CYP1A1 levels nor characterize its promoter directly in fetal blood.

Next steps

To address ongoing unanswered questions and extend our initial work in our non human primate model

as well as maternal smoking studies, we are currently utilizing of state-of-the-art high throughput technologies developed in our laboratory over the past two years. Taking a genome-wide approach, we have successfully developed working protocols and bioinformatic methodologies necessary for primate work, as well as advanced statistical analysis and programming for profiling epigenetic markers using both CpG microarrays and cutting edge single molecule sequencing (SMS) technologies.

Summary

In summary, data are beginning to emerge suggesting that epigenetic modifications of both DNA and chromatin structure, such as DNA methylation and histone modifications, may play a role in the postnatal morbidity plaguing the infant with poor growth and nutrition. This mechanism is thought to have arisen as a means to allow the genome a certain level of inheritable plasticity in order to adapt to fluctuating environmental conditions that random mutations in DNA sequence alone would not be able to adjust for. Modern day conditions often present our bodies to rapidly changing environments (e.g., shifting nutrient sources in the Developing World) that overwhelm our bodies' ability to cope, often resulting in self perpetuating morbidities, such as metabolic syndrome, that can last multiple generations. This is an important arena of research as much remains unknown. As we learn more, it will be possible to intervene and effect these epigenetic modifications, resulting in significant changes to phenotype and, ultimately, disease.

Acknowledgments

This work was supported by the NIH Director New Innovator Pioneer Award DP2120OD001500-01, NIH NIDDK R01DK080558-02 and the NICHD-Women's Reproductive Health Research (WRHR) Career Development Program #K12HD050128. We would like to thank the members of the Aagaard lab for helpful discussions regarding the manuscript including Dr. Min Hu, Ms. Lori Showalter, and Ms. Cynthia Shope.

References

Aagaard-Tillery KM, Grove K, Bishop J, Ke X, Fu Q, McKnight R, Lane RH. 2008a. Developmental origins of disease and determinants of chromatin structure: maternal diet modifies the primate fetal epigenome. *J Mol Endocrinol*, 41:91-102.

Aagaard-Tillery KM, Porter TF, Lane RH, Varner MW, LaCoursiere DY. 2008b. In utero tobacco exposure is associated with modified effects of maternal factors on fetal growth. *Am J Obstet Gynecol*, 198:66.e1-6.

Aagaard-Tillery K, Spong CY, Thom E, Sibai B, Wendel G Jr, Wenstrom K, Samuels P, Simhan H, Sorokin Y, Miodovnik M, Meis P, O'Sullivan MJ, Conway D, Wapner RJ. 2010. Pharmacogenomics of maternal tobacco use: metabolic gene polymorphisms modify risk of adverse pregnancy outcomes. *Obstet Gynecol*, 115:568-577.

Anttila S, Hakkola J, Tuominen P, Elovaara E, Husgafvel-Pursianen K, Karjalainen A, Hirvonen A, Nurminen T. 2003. Methylation of cytochrome P4501A1 promoter in the lung is associated with tobacco smoking. *Cancer Res*, 63:8623-8628.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell*, 129:823-837.

Beisel C, Buness A, Roustan-Espinosa IM, Koch B, Schmitt S, Haas SA, Hild M, Katsuyama T, Paro R. 2007. Comparing active and repressed expression states of genes controlled by the Polycomb/Trithorax group proteins. *Proc Natl Acad Sci USA*, 104:16615-16620.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ 3rd, Gingeras TR, Schreiber SL, Lander ES. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*, 120:169-181.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125:315-326.

Bernstein BE, Meissner A, Lander ES. 2007. The mammalian epigenome. *Cell*, 128:669-681.

Bertram L, Hiltunen M, Parkinson M, Ingelsson M, Lange C, Ramasamy K, Mullin K, Menon R, Sampson AJ, Hsiao MY, Elliott KJ, Velicelebi G, Moscarillo T, Hyman BT, Wagner SL, Becker KD, Blacker D, Tanzi RE. 2005. Family-based association between Alzheimer's disease and variants in UBQLN1. *N Engl J Med,* 352:884-894.

Bestor T, Laudano A, Mattaliano R, Ingram V. 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol,* 203:971-983.

Bestor TH. 1992. Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *Embo J*, 11:2611-2617.

Bocock PN, Aagaard-Tillery KM. 2009. Animal models of epigenetic inheritance. *Semin Reprod Med*, 27:369-379.

Brickner DG, Cajigas I, Fondufe-Mittendorf Y, Ahmed S, Lee PC, **Widom J, Brickner JH**. 2007. H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous

transcriptional state. *PLoS Biol*, 5:e81.

Chen T, Ueda Y, Dodge JE, Wang Z, Li E. 2003. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol*, 23:5594-5605.

Clayton AL, Hazzalin CA, Mahadevan LC. 2006. Enhanced histone acetylation and transcription: a dynamic perspective. *Mol Cell*, 23:289-296.

Cnattingius S, Granath F, Petersson G, Harlow BL. 1999. The influence of gestational age and smoking habits on the risk of subsequent preterm deliveries. *N Engl J Med*, 341:943-948.

Cox J, Williams S, Grove K, Lane RH, Aagaard-Tillery KM. 2009. A maternal high fat diet is accompanied by alterations in the fetal primate metabolome. *Am J Obstet Gynecol*, 201:281.e1-9.

Daube H, Scherer G, Riedel K, Ruppert T, Tricker AR, Rosenbaum P, Adlkofer F. 1997. DNA adducts in human placenta in relation to tobacco smoke exposure and plasma antioxidant status. *J Cancer Res Clin Oncol*, 123:141-151.

Egan CM, Sridhar S, Wigler M, Hall IM. 2007. Recurrent DNA copy number variation in the laboratory mouse. *Nat Genet*, 39:1384-1389.

Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C. 1992. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res*, 10:2709-2721.

Eichler EE, Nickerson DA, Altshuler D, Bowcock AM, Brooks LD, Carter NP, Church DM, Felsenfeld A, Guyer M, Lee C, Lupski JR, Mullikin JC, Pritchard JK, Sebat J, Sherry ST, Smith D, Valle D, Waterston RH. 2007. Completing the map of human genetic variation. *Nature*, 447:161-165.

Estivill X, Armengol L. 2007. Copy number variants and common disorders: filling the gaps and exploring complexity in genome-wide association studies. *PLoS Genet*, 3:1787-1799.

Feng Y-Q, Desprat R, Fu H, Olivier E, Lin CM, Lobell A, Gowda SN, Aladjem MI, Bouhassira E. 2006. DNA methylation supports intrinsic epigenetic memory in mammalian cells. *PLoS Genet*, 2:e65.

Foley DL, Craig JM, Morley R, Olsson CJ, Dwyer T, Smith K, Saffery R. 2009. Prospects for epigenetic epidemiology. *Am J Epidemiol*, 169:389-400.

Fu Q, McKnight RA, Yu X, Wang L, Callaway CW, Lane RH. 2004. Uteroplacental insufficiency induces site-specific changes in histone H3 covalent modifications and affects DNA-histone H3 positioning in day 0 IUGR rat liver. *Physiol Genomics*, 20:108-116.

Fu Q, McKnight RA, Yu X, Callaway CW, Lane RH. 2006. Growth retardation alters the epigenetic characteristics of hepatic dual specificity phosphatase 5. *FASEB J*, 20:2127-2129.

Gladen BC, Zadorozhnaja TD, Chislovska N, Hryhorczuk DO, Kennicutt MC 2nd, Little RE. 2000. Polycyclic aromatic hydrocarbons in placenta. *Hum Exp Toxicol*, 19:597-603.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, 130:77-88,

Hammoud AO, Bujold E, Sorokin Y, Schild C, Krapp M, Baumann P. 2005. Smoking in pregnancy revisited: findings from a large population-based study. *Am J Obstet Gynecol*, 192:1856-1862.

Holliday R. 1987. The inheritance of epigenetic defects. *Science*, 238:163-170.

Hong Y-C, Lee K-H, Son B-K, Ha EH, Moon HS, Ha M. 2003. Effects of the GSTM1 and GSTT1 polymorphisms on the relationship between maternal exposure to environmental tobacco smoke and neonatal birth weight. *J Occup Environ Med*, 45:492-498.

Huel G, Godin J, Moreau T, Girard F, Sahuquillo J, Hellier G, Blot P. 1989. Aryl hydrocarbon hydroxylase activity in human placenta of passive smokers. *Environ Res*, 50:173-183.

Ke X, McKnight RA, Wang ZM, Yu X, Wang L, Callaway CW, Albertine KH, Lane RH. 2005. Nonresponsiveness of cerebral p53-MDM2 functional circuit in newborn rat pups rendered IUGR via uteroplacental insufficiency. *Am J Physiol Regul Integr Comp Physiol*, 288:R1038-1045.

Ke X, Lei Q, James SJ, Kelleher SL, Melnyk S, Jernigan S, Yu X, Wang L, Callaway CW, Gill G, Chan GM, Albertine KH, McKnight RA, Lane RH. 2006. Uteroplacental insufficiency affects epigenetic determinants of chromatin structure in brains of neonatal and juvenile IUGR rats. *Physiol Genomics*, 25:16-28.

Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, Li K, Murty VV, Schupf N, Vilain E, Morris M, Haghighi F, Tycko B. 2008. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat Genet*, 40:904-908.

Lane RH, Kelley DE, Gruetzmacher EM, Devaskar SU. 2001. Uteroplacental insufficiency alters hepatic fatty acid-metabolizing enzymes in juvenile and adult rats. *Am J Physiol Regul Integr Comp Physiol*, 280:R183-R190.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, 389:251-260.

Luger K. 2006. Dynamic nucleosomes. *Chromosome Res*, 14:5-16.

MacLennan NK, James SJ, Melnyk S, Piroozi A, Jernigan S, Hsu JL, Janke SM, Pham TD, Lane RH. 2004. Uteroplacental insufficiency alters DNA methylation, one-carbon metabolism, and histone acetylation in IUGR rats. *Physiol Genomics*, 18:43-50.

Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. 2000. Demethylation of the zygotic paternal genome. *Nature*, 403:501-502.

McCurdy CE, Bishop JM, Williams SM, Grayson BE, Smith MS, Friedman JE, Grove KL. 2009. Maternal high-fat diet triggers lipotoxicity in the fetal livers of nonhuman primates. *J Clin Invest*, 119:323-335.

McGrath J, Solder D. 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell,* 37:179-183.

McMinn J, Wei M, Sadovsky Y, Thaker HM, Tycko B. 2006. Imprinting of PEG1/MEST isoform 2 in human placenta. *Placenta*, 27:119-126.

Mooney LA, Bell DA, Santella RM, Van Bennekum AM, Ottman R, Paik M, Blaner WS, Lucier GW, Covey L, Young TL, Cooper TB, Glassman AH, Perera FP. 1997. Contribution of genetic and nutritional factors to DNA damage in heavy smokers. *Carcinogenesis*, 18:503-509.

Nightingale KP, O'Neill LP, Turner BM. 2006. Histone modifications: signaling receptors and potential elements of a heritable epigenetic code. *Curr Opin Genet Dev*, 16:125-136.

Oscarson M, McLellan RAN, Gullsten H, Agúndez JA, Benítez J, **Rautio A, Raunio H, Pelkonen O, Ingelman-Sundberg M**. 1999. Identification and characterization of novel polymorphisms in the CYP2A locus: implications for nicotine metabolism. *FEBS Lett*, 460:321-327.

Oswald J, Engemann S, Lane N, Mayer W, Olek A, **Fundele R, Dean W, Reik W, Walter J**. 2000. Active demethylation of the paternal genome in the mouse zygote. *Curr Biol*, 10:475-478.

Pastinen T, Hudson TJ. 2004. Cis-acting regulatory variation in the human genome. *Science*, 306:647-650.

Pastinen T, Ge B, Hudson TJ. 2006. Influence of human genome polymorphism on gene expression. *Hum Mol Genet*, 15(spec no 1):R9-R16.

Peacock JL, Cook DG, Carey IM, Jarvis MJ, Bryant AE, Anderson HR, Bland JM. 1998. Maternal cotinine level during pregnancy and birth weight for gestational age. *Int J Epidemiol*, 27:647-656.

Pennisi E. 2007. Breakthrough of the year. Human genetic variation. *Science*, 318:1842-1843.

Riggs A, Martienssen R, Russo V. 1996. Introduction. *In*: Russo VEA, Martienssen RA, Riggs AD (Ed.). *Epigenetic Mechanisms of Gene Regulation*. Plainview, NY: Cold Spring Harbor Lab. Press. pp. 1-4.

Rollins RA, Haghighi F, Edwards JR, Das R, Zhang MQ, Ju J, Bestor TH. 2006. Large-scale structure of genomic methylation patterns. *Genome Res*, 16:157-163.

Secker-Walker RH, Vacek PM. 2003. Relationship between cigarette smoking during pregnancy, gestational age, maternal weight gain, and infant birth weight. *Addict Behav*, 28:55-66.

Segars JH, Aagaard-Tillery KM. 2009. Epigenetics in Reproduction. *Semin Reprod Med*, 27:349-350.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Casero RA, Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119:941-953.

Shi Y. 2007. Histone lysine demethylases: emerging roles in development, physiology and disease. *Nat Rev Genet,* 8:829-833.

Shi Y, Whetstine JR. 2007. Dynamic regulation of histone lysine methylation by demethylases. *Mol Cell,* 25:1-14.

Surani MA; Barton SC ; Norris ML. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature*, 308:548-550.

Suter MA, Aagaard-Tillery KM. 2009. Environmental influences on epigenetic profiles. *Semin Reprod Med,* 27:380-390.

Suter MA, Abramovici A, Hu M, Showalter L, Shope C, **Varner M, Aagaard-Tillery K**. 2010. In utero tobacco exposure epigenetically modifies placental CYP1A1 gene expression. *Metabolism.* (doi: 10.1016/j.metabd.2010.01.013).

Takai D, Jones PA. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci* USA, 99:3740-3745.

Waddington CH. 1968. Towards a theoretical biology. *Nature,* 218:525-527.

Wang S, Chanock S, Tang D, Li Z, Jedrychowski W, Perera FP. 2008. Assessment of interactions between PAH exposure and genetic polymorphisms on PAH-DNA adducts in African American, Dominican, and Caucasian mothers and newborns. *Cancer Epidemiol Biomarkers Prev,* 17:405-413.

Waterland RA. 2006. Assessing the effects of high methionine intake on DNA methylation. *J Nutr,* 136:1706S-1710S.

Waterland RA, Lin JR, 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.

Wolffe AP, Matzke MA. 1999. Epigenetics: regulation through repression. *Science* 286:481-486.

Yan C, Boyd DD. 2006. Histone H3 acetylation and H3K4 methylation define distinct chromatin regions permissive for transgene expression. *Mol Cell Biol*, 26:6357-6371.