Teacher Background

Epigenetics and Inheritance

Note: The Teacher Background Section is meant to provide information for the teacher about the topic and is tied very closely to the PowerPoint slide show. For greater understanding, the teacher may want to play the slide show as he/she reads the background section. For the students, the slide show can be used in its entirety or can be edited as necessary for a given class.

What is Epigenetics?

Epigenetics is generally defined "as relating to or arising from nongenetic influences on gene expression". (1) In Greek, the prefix *epi* means *upon, above, in addition to*, or *near*, so epigenetics means a way of changing the expression of genes without changing the DNA sequence. (2) The word epigenetics is derived from the word epigenesis which "can be traced back to Aristotle, who proposed that humans developed from the interplay of nature, or 'preformation', and nurture, which he called 'epigenesis'". (3) The term epigenetics was coined in the early 1940s by Conrad Waddington to explain "the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being." Today, based on several more recent studies, epigenetics is defined as "the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence."(4) The DNA genome has not changed in epigenetic inheritance but what has changed is whether or not the gene is expressed and whether that change in phenotype can be passed on the next generation(s).

Epigenetics is an emerging basic field of genetics at the epicenter of modern medicine which explains certain phenotypes that cannot be traced directly to the genome of the individual. As a mechanism, it explains how genetically identical monozygotic twins (resulting from the fertilization of a single oocyte that separates into two identical developing embryos shortly after fertilization) may develop differently. They have the same DNA base pair sequence in each of the cells and are always the same sex. However, it has been observed that the phenotype in monozygotic twins can be different. For example, one twin can develop cancer while the other never does; one can be bisexual while the other is heterosexual; one may seem to physically age more than the other. With identical DNA in each of their cells, how can they have such different outcomes? Is it possible for some of these traits to be passed on to future generations?

How are Genes Packaged?

The DNA in genes is packaged into chromosomes in order to fit inside the nucleus. The shortest human chromosome contains 46,900,000 base pairs of DNA, 1400 nm in diameter, and 2 μm long. Since each base pair is approximately 3.4 Å or 0.34 nm, the DNA has to be organized in some way in order to form the chromosome. The first level of organization is a bead-like structure called a nucleosome. Nucleosomes are made of two copies each of H2A, H2B, H3, and H4 histone proteins around which 146 base pairs of DNA are wrapped twice and then connected to the second group of histones by linker DNA containing 8 - 114 base pairs, depending on the species. Another 146 base pairs will wrap twice around this second group of histones and link to the third group of histones via linker DNA, eventually forming a long strand of chromatin which is called "beads on a string" and which is 11 nm in diameter. (5)

The second level of organization is the coiling of the "beads on a string" into a helical structure producing a 30 nm in diameter chromatin fiber that is found in both interphase chromatin and mitotic chromatin. The stability of the 30 nm fiber seems to be due to the presence of another histone, H1, since experiments that strip H1 from the chromatin maintain the nucleosome but not the 30-nm structure. The final organization of the DNA is in loops, scaffolds, and domains found in loosely folded interphase chromosomes and more densely folded mitotic chromatids and are around 700 nm in diameter. The interphase chromosome has areas of single-copy genetically-active DNA (euchromatin) where proteins are actively being transcribed and translated, while the mitotic chromatin contains tightly coiled and genetically inactive DNA (heterochromatin) when the cell is undergoing mitotic cell division and no protein synthesis is taking place. (5)

Chromosomes are depicted in the mitotic metaphase condensed state largely so the centromere location, banding patterns, and relative size of the chromosome can be seen. The centromere is the highly condensed region within the chromosome that is responsible for the accurate segregation of the replicated chromatin material during mitosis and meiosis. The area above the centromere is referred to as the p arm of the chromosome and the area below is called the q arm. For example, Cri-du-Chat Syndrome is characterized by a cat-like cry, mental retardation, and distinctive facial features in the offspring and is due to a deleted upper portion of the 5^{th} chromosome. The syndrome is denoted as 5 p- (5 p minus). (5)

How Are Genes Expressed?

When genes are expressed, the DNA is transcribed into mRNA in the nucleus and then eventually translated into protein at the ribosome via the efforts of tRNA which delivers the amino acids of the protein in sequence, according to the triplet codon of the mRNA. As the sequence of amino acids is being arranged, post-translational changes to the protein occur, such as the addition of functional groups like phosphate, acetyl, methyl, and others, cleavage of subunits, or degradation of entire

proteins. The modified proteins then perform their assigned task (e.g., structural protein, enzyme, or immunoglobulin, just to name a few). Many of these modifications regulate aspects of normal development, such as differentiation and stem cell maintenance.

What Are Epigenetic Modifications of DNA?

Methylation modifications of DNA and expressed proteins have been known for years. In 1969, Griffith and Mayler suggested these modifications may modulate gene expression. (4) Epigenetic modifications influence condensation of the chromatin and affect the accessibility of the DNA to be transcribed into mRNA and translated into proteins. Known epigenetic mechanisms include:

1) DNA methylation ($-CH_3$) of the cytosine at a CpG island in promoters and other areas of the genome causes genes to be switched off as the DNA condenses. Genes can be switched on by removing the methyl group from the same site which causes the DNA to uncondense or unravel.

2) Histone modifications of the N-terminal histone tail domains, including methylation (CH_3), acetylation ($-COCH_3$), phosphorylation ($-PO_4^{2-}$), ubiquitination (-ubiquitin polypeptide), or sumyolation (-small ubiquitin-related modifier), can cause the gene to be turned on or off depending on the modification.

3) RNA interference (RNAi), including microRNAs (miRNA), small interfering (siRNA), and PIWI-interacting (piRNA), can alter the transcribed mRNA and thus prevent it from being translated into protein. (4, 6, 7)

What Is DNA Methylation?

Methylation of the DNA at a cytosine is the best understood of the epigenomic modifications and takes place at the CpG sites where cytosine is next to guanine in the DNA sequence. The “p” in CpG refers to the phosphodiester bond between the cytosine and the guanine, showing they are next to each other on the same side of a DNA strand, in both single- and double-stranded DNA (writing C-G would denote a base pairing across from each other in the double-stranded DNA). Methylation of the DNA causes the DNA to condense and genes within the condensed DNA cannot be expressed since they have no access to the various enzymes involved in transcription of DNA. The methyl groups can be removed from the DNA which causes the DNA to unravel and genes to be expressed. (8)

Within the genome, 70 – 80% of the CpG dinucleotides are methylated. However, there are areas that contain an unusually high number of CpG dinucleotides, known as CpG islands, which are typically free of methylation. (9) The CpG island concept was defined by Gardiner-Garden in 1987 as being a 200-base pair section of DNA with a C+ G content of 50% and an observed CpG/expected CpG in excess of 0.6. More recently, Daiya Takai and Peter A. Jones studied chromosomes 21 and 22 from the human genome sequences and defined CpG island as regions of DNA of greater than 500 base pairs with a C+ G equal to or greater than 55% and observed CpG/expected CpG of 0.65. By this definition, these islands were

more likely to be associated with 5' regions of genes (this definition excludes most *Alu*-repetive elements). CpG islands are now understood to be associated with the 5' end of all of the housekeeping genes and many tissue-specific genes, since methylation of cytosine can only occur when a cytosine lies directly on the five prime side of a guanine. 40% of mammalian genes contain CpG islands in their promoters and exon regions, which are usually not methylated. Otherwise, there are few CpG dinucleotides in other regions of the mammalian genome and they are mostly methylated. One of the main roles of the any methylation in the CpG islands is to silence the gene during the process of differentiation, X-chromosome inactivation, imprinting, and silencing intragenomic parasites. (10, 11)

There are five known enzymes which methylate DNA. All are DNA methyltransferase (DNMT) enzymes found in mammalian cells and include DNMT 1, 2, 3A, 3B, and 3L.

1) Human DNMT1 is called DNA (cytosine-5)-methyltransferase 1 and is the most abundant and catalytically active in most cell types. The gene that codes for this enzyme is located on chromosome 19 at position 13.2 on the p arm and consists of an N-terminal ($-NH_2$) domain which plays a regulatory role and a C-terminal (-COOH) domain which has a catalytic effect on methylation. Its main function is recognizing hemimethylated cytosines (DNA in which one strand is methylated and the other strand is not) when the DNA is being copied during cell division and adding a methyl group to the newly made side of the DNA. It also regulates reactions involving proteins and lipids and controls the processing of chemicals that relay signals in the nervous system (neurotransmitters).

Mouse knockouts of the *DNMT1* gene are lethal at 8.5 days in the embryo. In humans, a mutation in the *DNMT1* gene is named hereditary sensory and autonomic neuropathy type IE (HSAN IE) and is characterized by a gradual loss of intellectual function (dementia), deafness, and sensory problems in the feet. The mutation occurs in exon 20 of the gene which reduces or eliminates the function of the translated protein. A mutation in exon 21 of the gene is named autosomal dominant cerebellar ataxia, deafness, and narcolepsy and is distinct from the mutation in exon 20. There are several normal polymorphisms in the *DNMT1* gene which have been associated with an increased risk of cancer, especially of the breast and stomach. In addition, overexpression of the *DNMT1* gene has been identified in certain brain cancers called gliomas. (12, 13)

2) Human DNMT2, called DNA (cytosine-5)-methyltransferase 2, has low enzymatic activity, and mouse knockouts of the *DNMT2* gene show no change in phenotype. It was found in 2006 to be an RNA methytransferase which methylates cytosine 38 in the anitcodon loop of tRNA, rather than methylating DNA. It is significantly different from other *DNMT* genes in that it has only the C-terminal catalytic domain and is missing the N-terminal regulatory domain. The gene which codes for this enzyme is located on chromosome 10 on the p arm. (9, 12, 14)

3) The human *DNMT3* gene family is thought to be involved primarily in *de novo* methylation and is highly expressed at the stage of embryonic development when waves of *de novo* methylation are occurring in the genome (embryo implantation). They are also important for imprinting and X-chromosome inactivation. *DNMT3A,* called DNA (cytosine-5)-methyltransferase 3 alpha, and *DNMT3B,* called DNA (cytosine-5)-methyltransferase 3 beta, are the main two genes in this family, although it has been shown that methylation levels are significantly higher when *DNMT3A* and *DNMT3L*, called DNA (cytosine-5)-methyltransferase 3-like, are both expressed than when *DNMT3A* is expressed alone. However, the same effect is not seen with DNMT3B so it does not appear that DNMT3L is necessary for DNMT3B to fully function. Alone, DNMT3L is inactive due to the lack of critical catalytic motifs, although mouse knockouts of the *DNMT3L* gene result in maternal DNA methylation imprint failure and male sterility. (9, 12, 14) The *DNMT3L* gene is located on chromosome 21q, in the distal part of the Down's syndrome critical region. The *DNMT3A* gene is located on chromosome 2 at position 23.3 on the p arm while the *DNMT3B* gene is located on chromosome 20 at the position 11.2 on the q arm. (15, 16, 17)

Mouse knockouts of the *DNMT3A* gene are born live but die before reaching four weeks of age. They exhibit subtle DNA methylation defects in maternally imprinted regions. Mouse knockouts of the *DNMT3B* gene are lethal by day 14.5 in the embryo and show marked demethylation of pericentromeric (around the centromere of a chromosome) satellite repeats. (12) *DNMT3A* gene somatic mutations are found in 20% of individuals with acute myeloid leukemia (AML) although a direct link between mutant *DNMT3A* gene, epigenetic changes and pathogenesis still remains to be established. (18) Mutations in the *DNMT3B* gene cause a rare autosomal recessive disease called Immunodeficiency, Centric Region Instability, Facial Anomalies Syndrome (ICF). Individuals with this syndrome have low serum levels of IgG, IgM, IgE, and/or IgA and thus have recurrent infections, hypomethylation and DNA rearrangement of the juxtacentromeric heterochromatin in chromosomes 1 and 16 and sometimes 9, and facial abnormalities including a broad flat nasal bridge, very widely spaced eyes, and epicanthal folds. The patient's prognosis varies depending on which opportunistic infections and pulmonary infections the person contracts. Recently, bone marrow transplantation has been successful in two children. (19, 20)

What Are Histones?

Histone modifications are another mechanism of epigenetics. Since the DNA is a long strand of dinucleotides and would not fit into the cell nucleus as is, corralling the DNA is accomplished by winding it around proteins called histones. Eight histone proteins – 2 each of H2A, H2B, H3, and H4 cluster- first cluster in an H2A-H2B dimer and H3-H4 dimer. Then each dimer combines with a like dimer to form an H2A-H2B tetramer and H3-H4 tetramer. Finally, the two tetramers combine to form the histone octomer around which 146 - 147 base pairs of DNA wrap 1.67 times in a left-handed super-helical turn (21) to form the nucleosome.

DNA is acidic and negatively charged and histones are basic and positively charged so they combine by the attraction of opposite charges. The helix-dipoles from the histone alpha helices cause a net positive

charge to accumulate at the point of interaction with the negatively charged phosphate groups on the DNA. In addition, hydrogen bonds form between the DNA backbone -P=O and the H-N- amide group on the main chain of histone proteins. Non-polar interactions occur between the histone, such as at a proline amino acid, and a deoxyribose sugar on the DNA . Salt bridges form between basic side chains, like lysine and arginine with their $-NH_3^+$ group on the amino acid side chain, and the acidic PO_4^- on the DNA. Lastly, there are non-specific groove insertions of the H3 and H2B N-terminal tails into two minor grooves each on the DNA molecule. (22) The H1 linker histone keeps the DNA in the nucleosome from unwrapping so the DNA can wrap around the next cluster of histones to form another nucleosome. (23)

In order for DNA to be repaired and for transcription to occur, the nucleosome must unravel to expose the DNA. The nucleosome structure is regulated by numerous modifications found on the N-terminus histone "tails" which can cause condensation or uncondensation (unraveling) of the DNA, depending on the type of modification. Genes on the DNA can be expressed only when in the uncondensed (unraveled) form.

What are Histone Modifications

The N-terminus of the histone tails (amino, or NH_2, end beginning with amino acid #1) are highly basic and extend beyond the DNA into the cytoplasm, providing a site for post-translational histone modification –methylation[1], acetylation[2], phosphorylation[3], ubiquitylation[4], and sumyolation[5] – mostly

[1]Adding the methyl ($-CH_3$) group to a histone amino acid does not change its charge. Just as methylation of DNA suppresses gene expression and demethylation induces gene expression, methylation (me) of lysine (K) in the H3 histone at the 9th amino acid position on the histone tail (H3K9me) and the 27th amino acid position (H3K27me) causes suppression of gene expression. However, some trimethylations (me3), such as on lysine (K) at the 4th amino acid position on H3 histone (H3K4me3), induce gene expression. (25) The histone methylations are catalyzed by histone methyltransferase (HMT). HMT is also known as euchromatic histone-lysine N-methyltransferase 1 *(*EHMT1). The *EHMT1* gene which codes for this enzyme is found on chromosome 9 at position 34.3 on the q arm. (26) Demethylation of lysine in the histone tail proteins by the enzyme histone demethylase (HDM) can cause either activation or repression of gene expression based on which lysine residue is demethylated. There are many HDM enzymes that have been found that demethylate the histone tails. (25)

[2]Histone acetylation by the enzyme histone acetyltransferase (HAT) activates gene expression while deacetylation by the enzyme histone deacetylase (HDAC) suppresses gene expression. The activation is accomplished by the elimination of lysine's positive charge (R group $-(CH_2)_4-NH_3^+$) by adding the acetyl group ($-COCH_3$), which is polar with a small negative cloud around the oxygen atom. This action has the potential to weaken the interactions between the histones and DNA allowing unfolding of the DNA and the expression of the gene. (24)

[3]Histone phosphorylation ($-PO_4^{-2}$) has divergent roles. Sometimes it is involved in chromatin condensation associated with mitosis and meiosis as when phosphorylation (ph) of serine (S) on histone at the 10th amino acid position (H3S10ph) and the 28th amino acid position (H3S28ph) occur. Sometimes phosphorylation is involved in chromatin relaxation by influencing demethylation and acetylation of adjacent amino acid residues on the histone. (25)

[4]Ubiquitin is a small 76-amino acid polypeptide that can be covalently bonded to histone H2A and H2B. It was discovered nearly 30 years ago as a covalent modifier of H2A and now is known to be one of the cell's most broadly used protein modifications. When added to a histone tail in ubiquitylation, it can either suppress or enhance gene expression depending on which amino acid in H2A or H2B it is attached.(27)

[5]Sumyolation is a reversible post-translational modification of histone proteins by small ubiquitin-related modifiers (SUMOs). They are polypeptides which are just a little longer than ubiquitin polypeptides (around 100 amino acids) and attach to histone lysines which are in the ψ K X E (ψ is a hydrophobic amino acid, K is lysine, X is an undetermined amino acid, and E is glutamic acid). Sumyolation helps with gene silencing through recruitment of histone deacetylase and heterochromatin protein 1. (28, 29)

on the lysine, serine, and arginine amino acids of the tail. (24)(Post-translational means that the modifications are added after the histone proteins are transcribed and translated.) These modifications influence the chromatin structure by unfolding to form euchromatin (where the genes can be transcribed), or by tightly coiling to form heterochromatin (where the genes cannot be expressed), depending on the modification.

Histone modifications can positively or negatively affect other histone modifications by 'crosstalk'. For instance, ubiquitylation (ub) of lysine (K) at the 123rd amino acid position on H2B (H2BK123ub) can promote methylation of lysine (K) at 4th amino acid position on H3 (H3K4me). Likewise, acetylation of lysine (K) at the 14th amino acid position on H3 (H3K14ac) can block methylation of lysine (K) at the 9th amino acid position on H3 (H3K9). In addition, phosphorylated amino acid residues on histones are heavily involved in 'crosstalk' among various other amino acid residues on histone tails, thereby affecting their modifications and behavior.

Most of the research on histone tail modification has been conducted on the N-terminus of histones H2A, H2B, H3, and H4. Much less is known about the C-terminus (carboxyl end) of histones. The C-terminus of the H2A protrudes from the nucleosome and is located where the DNA leaves the nucleosome. This unstructured histone tail is very small – only 15 amino acid residues beyond the golubular domain – and is divided into two parts. Amino acids (aa) 115-122 (8 aa long) pass between the strands of DNA which are wrapped around the nucleosome, and amino acids 123-129 (7 aa long) protrude from the nucleosome structure. The H2A tail can interact with the H1 linker histone so is important in chromatin structure and it has been shown to have a key role in nucleosome stability and mobility *in vivo* and *in vitro*. The C-terminus tail of the H2B histone amino acid residues arginine (R) at the 119th amino acid position (R119) and threonine (T) at the 122nd amino acid position (T122) play a direct role in controlling the ubiquitylation of proline (P) at the 1st amino acid position on the N-terminus of H2B (H2BP1ub) and methylation of lysine (K) at the 4th amino acid position on H3 (H3K4me) methylation by cross talk.The C-terminus of H3 and H4 do not protrude from the nucleosome and thus do not have modifications. (96, 97)

<u>Are There Interactions Between DNA Methylations and Histone Modifications?</u>

DNA methylations and histone modifications can interact with each other to change the expression of the gene.The transcriptionally inactive heterochromatin has increased affinity for methylated DNA-binding proteins. These proteins further recruit histone deacetylases (HDAC) and attract DNA and histone methyltransferases (DNMT and HMT), which, in turn, further reinforce histone modification patterns conducive to silencing. Methylated promoters are associated with unique repressive histone markers, such as trimethylation of histone 3 (H3) at lysine (K) at the 9th amino acid position (H3K9me3), and at lysine (K) on the 27thamino acid position (H3K27me3).

The transcriptionally active euchromatin has unmethylated promoters associated with gene transcription. They have increased affinity for histone acetylases (HAT), DNA and histone demethylases (DME and HDM), and histone markers associated with active chromatin, such as acetylated (ac)H3K9ac and trimethlated H3K4me3. (30) This methylation prevents the finding of DNMT3L-DNMT3A methyltransferase complex and thus prevents methylation of DNA which would silent the gene. (31)

In transcriptionally active euchromatin, methylated DNA wrapped around histones with acetylated histone tails can direct the deacetylation after the DNA is methylated by DNMT1 thus causing the DNA to coil into heterochromatin. Once the acetyl groups are removed by HDAC, then HMT adds methyl groups and heterochromatin protein 1 (HP1) to the histone tails to maintain the heterochromatin state. HP1 proteins are dominant suppressors binding to H3K9me, causing transcriptional repression. (32, 33)

What Is Post-Transcriptional Gene Silencing by RNA Interference (RNAi)?

RNA interference (RNAi) is another mechanism of epigenetics. RNAi is called post-transcriptional gene silencing (PTGS). RNAi is a conserved biological response to double-stranded RNA that mediates resistance to both endogenous parasitic and exogenous pathogenic nucleic acids, and regulates the expression of protein-coding genes. This is accomplished by causing the destruction of specific mRNA molecules in endogenous parasitic and in exogenous pathogenic nucleic acids.(34) There are several types of RNAi, including microRNAs (miRNAs), small interfering RNAs (siRNAs), and PIWI-interacting RNAs (piRNAs).(35)

In order for a gene to be expressed into a protein, DNA segments of base pairs in genes are transcribed into precursor-messenger RNA (pre-mRNA) and then altered into the final mRNA which is able to leave the nucleus. The mRNA attaches to a ribosome made of ribosomal RNA (rRNA) and protein and is translated into a protein via the ordering of the amino acid sequences by the complementary pairing of nitrogenous bases of the mRNA and the transfer RNA (tRNA) within the ribosome.

There are regions of the DNA which code for functional RNAs which are not translated into protein; these RNAs are called non-coding RNA (ncRNA). Examples of ncRNA include: tRNA; rRNA; small nuclear RNA (snRNA) which functions to process pre-mRNA into mRNA; small nucleolar RNA (snoRNA) which modifies rRNA, tRNA, and snRNA by methylation and by converting uridine to pseudouridine (ψ); transfer-messenger RNA (tmRNA) found in bacteria which combine tRNA and mRNA; and others, including miRNA, siRNA, and piRNA. All play a role in gene expression but are not translated into proteins themselves.

MicroRNAs (miRNAs) are short single-stranded RNA molecules about 18-25 nucleotides long which regulate gene expression by cleaving the mRNA before it can be translated into protein. Most of the mi-RNA work is done in the cytoplasm and stops mRNA translation into protein by cleavage, removal of the polyA tail (deadenylation), disruption of the cap-tail interactions, and/or degradation of mRNAs by exonucleases.[6] (35, 36, 37)

Small interfering RNAs (siRNAs) are double stranded RNA and have been found in plants and fungi; roles for siRNAs have been found in animals but they are less well understood. They can be introduced in the lab as well and play a role in cleaving mRNA, thus silencing the gene from being translated into protein.[7] (37)

PIWI-interacting RNAs (piRNAs) are about 25-30 nucleotides in length. They play a role in silencing sequences of DNA called transposable elements. Transposable elements are segments of DNA that can move from place to place in the genome – Barbara McClintock, who was first to describe transposable genes, called them "jumping genes". As they move around the genome, some are harmless while others can cause mutations when they reinsert in the middle of another gene. PIWI-interacting RNAs stop these sequences from moving around by cleaving the transposon transcript and then combining with Argonaute 3 proteins called PIWI proteins to cleave them further. (37)

How Are Epigenetics Modifications Involved in the Zygote and Embryo?

In the one-celled fertilized egg (zygote) stage, all of the methylation and other epigenetic information are removed. At this point, the cell is a stem cell and could potentially form any type of differentiated tissue. As the zygote begins to divide, each cell has identical DNA. The cells divide repeatedly as the developing embryo moves down the fallopian tube and enters the uterus in 3 – 5 days after fertilization. In the uterus, the cells continue to divide, becoming a hollow ball of cells called a blastocyst. Between 5 and 8 days after fertilization, the blastocyst attaches to the lining of the uterus, usually near the top. This process, called implantation, is completed by day 9 and 10. (38)

Before implantation, most CpGs in the embryonic genome are unmethylated, but some regions are packaged with nucleosomes containing methylated lysine (K) at the 4th amino acid position (H3K4me), perhaps as a result of RNA polymerase binding. At the time of implantation, the methyltransferases *DNMT3a* and *DNMT3b* are expressed. DNA methylation is facilitated by *DNMT3L* which binds to

[6]The miRNAs are transcribed by RNA polymerase II forming the large primary-microRNA (pri-miRNA) transcripts. An RNase III endonuclease called Drosha and its co-factor Pasha process the pri-miRNA into a double stranded structure of 60 – 70 nucleotides with a short hairpin loop structure on one end; this is called precursor-miRNA (pre-miRNA). The resulting double stranded structure has a 2 nucleotide overhang at the 3'end. The pre-miRNA binds to Exportin 5 to move through the nuclear pore where it is transferred to a Dicer enzyme grasped by the 3' overhang end with the hairpin loop sticking out of the Dicer molecule. The Dicer molecule cleaves the pre-miRNA into a double stranded miRNA molecule of 18-25 nucleotides in length. One strand of the miRNA is chosen and transferred to an RNA-induced silencing complex (RISC) which includes one of the RNA strands from the double stranded miRNA, an Argonaute 2 enzyme which will cleave the mRNA, and other proteins. The chosen miRNA strand guides the RISC complex to the mRNA and binds with it. If the miRNA binds perfectly with the mRNA strand, the RISC complex will cleave the mRNA at a site generally found in the coding sequence (or open reading frame)of the mRNA target. If the miRNA binds imperfectly, it only needs to match a few of the RNA bases; this is called a seed. When there is less than perfect pairing, the miRNA can remain bound with RISC to the mRNA and repress the translation of the mRNA in the 3' untranslated regions (3'UTRs).(35, 36, 37)

[7]These long double stranded siRNA molecules also are cut by the Dicer enzyme into short double stranded siRNAs which are about 21-25 nucleotides in length. One strand is chosen to combine with the RISC complex whose main component is Argonaute 2. The siRNA guides the complex to the mRNA where most siRNAs find a perfect match and then Argonaute 2 cleaves the mRNA. This silences the mRNA since it cannot be translated into protein. Endonucleases degrade the mRNA into its components. (37)

Before implantation, most CpGs in the embryonic genome are unmethylated, but some regions are packaged with nucleosomes containing methylated lysine (K) at the 4th amino acid position (H3K4me), perhaps as a result of RNA polymerase binding. At the time of implantation, the methyltransferases *DNMT3a* and *DNMT3b* are expressed. DNA methylation is facilitated by *DNMT3L* which binds to chromatin by recognizing the lysine (K4) amino acid residue on histone H3. If this histone moiety is methylated, however, the complex cannot bind and the underlying DNA region is thus protected from *de novo* methylation. This may be one of the mechanisms used to generate a bimodal methylation pattern characterized by methylation over most of the genome, but not at CpG islands. (39)

In order for the embryo to develop ectoderm, mesoderm, and endoderm (precursor layers from which all differentiated cells will arise), some genes will have to be switched off. This is done by DNA methyltransferase 3 (*DNMT3*) enzymes which will begin to add methyl groups to genes at *de novo* (new) CpG sites which shut the gene off. (12) As epigenetic modifications continue in the developing embryo, ectoderm (outer layer) will differentiate into skin cells of the epidermis, neurons of the nerve cells, and pigment cells. The mesoderm (middle layer) will differentiate into cardiac muscle, skeletal muscle cells, red blood cells, tubule cells of the kidney, and smooth muscle. The endoderm (inner layer) will differentiate into alveolar cells of the lungs, thyroid cells, and pancreatic cells.

For instance as the circulatory system begins to form complete with circulating red blood cells (hemoglobin) during fetal development, the fetal hemoglobin gene remains switched on while the adult hemoglobin gene is switched off. Fetal hemoglobin has an oxygen-binding affinity that is higher than that of normal adult hemoglobin since the fetus is relying on obtaining oxygen from the lower levels of oxygen in the mother's circulating blood. Closer to birth, the fetal hemoglobin gene is switched off by methylation which causes the DNA to condense, and the normal adult hemoglobin gene is switched on by demethylation which causes the DNA to uncondensed and be expressed to accommodate moving from the womb to breathing air in the outside environment. (7)

Little is known about how DNA methylation is targeted to specific regions but it has been suggested that it involves interactions between DNA methyltransferases and chromatin-associated proteins. (40) In addition to DNA methylation, histone modifications, including methylation, take place, and non-coding RNA, such as PIWI-interacting RNAs, can be transcribed from the DNA and bind to PIWI and Argonaute molecules to mediate genome regulation by cleaving the transposons that potentially could move from place to place in the DNA. (41)

<u>What Is Imprinting And How Is It Related To Epigenetics?</u>

Genomic imprinting is an epigenetic process that causes monoallelic gene expression by silencing or activating one of the alleles of an autosomal gene, depending on the parent of origin. The resulting allelic asymmetry distinguishes imprinting from other forms of epigenetic regulation. (42) Imprinted genes are genes marked in a parental-specific way and are reset during egg and sperm formation. The

epigenetic tags on imprinted genes usually stay put for life. Soon after the egg and sperm unite to form the zygote, most of the epigenetic tags that activate and silence genes are stripped from the DNA. However, in mammals, imprinted genes begin the process of development with epigenetic tags in place. For most genes, humans inherit two working copies – one from mom and one from dad. But with imprinted genes, humans inherit only one working copy. Depending on the gene, either the copy from mom or the copy from dad is epigenetically silenced. Regardless of whether they came from mom or dad, certain genes are always silenced in the egg while others are always silenced in the sperm. (43)

Imprinting of autosomal chromosomes is thought to be a result of the Genetic Conflict Theory which predicts imprinted genes will be growth enhancing when paternally inherited and growth suppressing when maternally inherited. The theory is supported by the fact that genomic imprinting usually affects the growth in the womb and behavior after birth. (44) For example, humans inherit a copy of the *Igf2* gene from each parent and usually both copies are expressed. But certain gene copies are switched on depending on whether they are of male or female origin. The *Igf2* gene is one such imprinted gene from the male. The male makes IGF-II protein and more IGF-II means a bigger placenta that drives more nutrients to the fetus and produces a bigger baby which ensures that the father's genes will make it into the next generation. In contrast, it is in the female's best interest to combat this effect through her own imprinted genes that control the size of her baby to ensure that she can reproduce more than once and as many of her genes as possible survive into the next generation. (45)

Are There Examples of Imprinting In Humans?

Imprinted genes were discovered when children with Prader-Willi Syndrome (PWS) and Angelman Syndrome (AS) both had a deletion in chromosome 15 but each syndrome had very different symptoms. PWS and AS are genetic disorders that occur in approximately one out of every 15,000 births. PWS and AS affect males and females with equal frequency and affect all races and ethnicities. PWS and AS are two very different disorders, but they are both linked to the same imprinted region of chromosome 15. In 1983, by studying the inherited heteromorphisms of chromosome 15 in individuals with PWS, the deleted 15q11.2-q13 region was shown to be of paternal origin.

By 1987, studies of individuals with AS showed that they also had a deletion in the same region of 15q. In 1988, the deletions in PWS and AS were shown to be indistinguishable by molecular analysis. In 1989, it was found that the deleted region in AS was always derived from the mother, and in 1990, the parent-of-origin effect on the phenotype became known as genomic imprinting. In PWS and AS, some of the genes in the imprinted 15q region are silenced in the egg, and at least one gene is silenced in the sperm. Therefore, someone who inherits a defect on chromosome 15q is missing different active genes, depending on whether the chromosome came from mom or dad. (46)

Prader-Willi Syndrome was first described by Swiss doctors Andrea Prader, Alexis Labhart, and Heinrich Will in 1956, based on common clinical characteristics of nine children they had examined. These

included small hands and feet, insatiable hunger and extreme obesity, small stature, very low lean body mass, low muscle tone at birth, poor motor skills, undescended testicles and underdeveloped sex organs, sucking complications, and intellectual disability. People with PWS have distinctive facial features such as a narrow forehead, almond-shaped eyes, and a triangular mouth and often have behaviorial problems such as temper outbursts, stubbornness, compulsive behavior, and sleep abnormalities. (47) Children with Prader-Willi Syndrome can be helped by monitoring their diet and daily exercise to keep their weight gain within healthy ranges and by monitoring growth hormones to build muscles.

As mentioned above, PWS is caused by a lack of active genetic material in the 15q11.2-q13 region. Normally, individuals inherit one copy of chromosome 15 from their mother and one from their father. The genes in the PWS region are normally only active on the chromosome that came from the person's father. In PWS, the genetic defect causing inactivity of the paternal chromosome 15 imprint can occur in one of three ways: by deletion (70% of cases), by uniparental disomy (UPD) which means there are 2 maternal copies (25% of cases), or by translocation or imprinting mutation. The PWS characteristic features are likely the result of loss of function of genes that provide instructions for making molecules called small nucleolar RNAs (snoRNAs) which help to modify other types of RNA, such as rRNA, tRNA, and snRNA (which processes pre-mRNA into mRNA). Studies suggest that loss of a particular group of snoRNA genes called the SNORD116 cluster may play a major role in causing the symptoms of PWS. (47, 48)

Angelman Syndrome was named after Harry Angelman who first reported "puppet children" as he called them. Symptoms included low weight and small cranium, loss of full control of bodily movements, recurrent seizures, low muscle tone, absence of speech, balance disorders, neurological problems, and unusual faces characterized by a large jaw and open-mouthed expression showing the tongue. (46, 49) Other symptoms include delayed development, intellectual disability, and children with AS typically have an easily provoked excitable demeanor with frequent smiling, laughter, and hand-flapping movements. Hyperactivity, short attention span, and a fascination with water as well as needing less than usual sleep are typical of AS children. (50)

As discussed above, AS is caused by a lack of active genetic material in the 15q11.2-p13 region. Normally, individuals inherit one copy of chromosome 15 from their mother and one from their father. The genes in the AS region are normally only active on the chromosome that came from a person's mother. In AS, the genetic defect causing inactivity of maternal chromosome 15 imprint can occur in one of three ways: by deletion (70% of cases), mutation in the maternal chromosome (11% of cases), or uniparental disomy (UPD) – 2 paternal copies in a small number of cases. (48, 50) Many of the characteristics of AS are the result of the loss of function of the gene *UBE3A*. Individuals inherit a copy of the gene from each parent and both copies are actively turned on in most tissue of the body. However, in the brain, only the maternal copy of the gene is active. If that maternal gene is faulty, the person will have no active copies of the gene in the brain thus causing the AS symptoms. (50)

Another epigenetic imprinting disease is Beckwith-Wiedemann Syndrome (BWS). It is caused by imprinting and methylation errors in several genes on the chromosomal band 11p15, including insulin-like growth factor-II *(IGF2*) at 11p15.5 and *H19* and *Lit1,* which encode untranslated RNAs. Symptoms of this syndrome include overgrowth of organs such as macroglossia (enlarged tongue), macrosomia (excessive infant birth weight >8 pounds, 13 ounces, ear pits or earlobe creases, midline abdominal-wall defects, and increased incidence of cancer in infancy and young child. Researchers have shown that many cases of BWS are entirely epigenetic in origin since there are reports of monozygotic twins who are discordant for this syndrome. (42, 51)

What Is X-Inactivation and How Is It Related To Epigenetics?

Even though females have two X chromosomes (XX) and males have only 1 (XY), the amount of protein transcribed and translated from those genes on the X chromosome is the same in females as males. One of the X chromosomes in each cell of the female is inactivated by epigenetic mechanisms to form heterochromatin. While imprinting can affect multiple contiguous genes within extended chromosomal regions, X-inactivation affects the entire chromosome. Imprinted chromosomal domains are measured in megabases whereas X-inactivation affects the majority of the genes across much of an entire chromosome. (42)

The inactivated X chromosomes are called Barr bodies and are visible using a light microscope. X-inactivation takes place during the late blastocyst stage, beginning on day 2 of development and ending around day 10 as the blastocyst embeds in the uterine wall. A normal XX female has 1 Barr body, an XXX female has 2, an XXXX female has 3. A normal XY male would have 0 Barr bodies, a Klinefelter's Syndrome male XXY would have 1, an XXXY male would have 2, and an XXXXY male would have 3. Turner Syndrome XO females have 0 Barr bodies and an active X chromosome. However, Turner's individuals differ from normal females in phenotype since X-inactivation occurs during days 2-10 in embryonic development and 2 copies of some X-linked genes are necessary before X-inactivation occurs in early development and later for normal ovarian function. The X chromosome(s) that is inactivated is determined randomly and all future generations of each cell will have identical cells. Therefore, females are mosaics with respect to the X chromosome. Patches of body cells will have the maternally inherited X chromosome inactivated and other patches will have the paternally inherited one inactivated.[8] (52)

[8]The *Xist* (X-inactive-specific transcript) gene is found on the long arm (q) of the X chromosome itself. The *Xist* gene produces a long non-coding RNA (>200 nucleotides). This long non-coding *Xist* is a 17 kilobase siRNA which is expressed only in cells containing at least 2 X chromosomes and isn't normally expressed in male cells. When expressed on the inactive X chromosome (Xi), *Xist* stays in the nucleus and begins to coat the entire length of the X chromosome which is becoming inactive. As it does , it recruits various silencing protein complexes to label the future inactive X chromosome. It also appears that the siRNA might direct DNA methylation and histone hypoacetylation, methylation, and ubiquitination on the Xi chromosome, further compacting it. Of the 233 genes on the X chromosome, 43 escape inactivation, mostly on the Xi p arm. Some genes on the Xi q arm also remain active to maintain ovarian function until menopause. (53, 54, 55)

Are Identical Twins Really Identical?

Genetically identical monozygotic twins result from the fertilization of a single egg that separates into two developing embryos shortly after fertilization. They have the same DNA base pair sequence in each of the cells and are always the same sex. However, the phenotype can be different due to epigenetic changes. For instance, since X-inactivation Barr bodies form randomly in each cell, two twin females might have a different mosaic pattern of Barr bodies in their tissues.

Epigenetic mechanisms, such as DNA methylation, histone modifications, and RNAi, can be affected by development *in utero* and/or childhood, environmental chemicals, drugs/pharmaceuticals, aging, and diet. This can explain why twins can be different phenotypically while having the same DNA sequence. Some of the epigenetic changes can activate genes and some can repress genes and can lead to health issues, such as cancer, autoimmune disease, mental disorders, and diabetes.(56)

Some of these changes in epigenetics can be revealed through a process called methyl mapping. A sodium bisulfite treatment of the study DNA converts unmethylated cytosine residues to uracil. Methylated cytosines resist conversion and are identified by microarrays or sequencing. (57) In a study of 160 identical twins, 35% of the group showed methylation pattern differences between very young twins and middle-aged twins. It was found that twins who had spent the most time apart and had more divergent medical histories exhibited the greatest epigenetic differences. (58)

Can Epigenetic Changes Be Inherited?

Imprinted genes are not the only genes that bypass epigenetic reprogramming in the early embryo as has been found in inherited epigenetic traits discovered more recently.(43) Epigenetic expression has been found to be affected by many environmental factors. Some studies have shown that the environment can affect not only the phenotype of the current generation but also generations to follow through the inheritance of these epigenetic factors. One factor is diet, both *in utero* and in childhood and adulthood. Especially important are methionine and folic acid-rich food, which affect methylation of DNA. A genetic connection between diet of one generation and the health of the next generation has been demonstrated in a number of epidemiological studies.

A study was done in the 1980's by Dr. Lars Olav Bygren of the Karolinska Institute in Stockholm using 19th century feast and famine data collected in an isolated Överkalix region in Norbotten, Sweden. It was found that pre-pubertal boys who enjoyed a rare abundance of food over a single winter produced children and grandchildren who lived an average of 6 years less than that their peers. (59) The boys were age 9 – 12 and were going through a slow growth period when cells are formed which give rise to

sperm. Since the building blocks for methylation come from food, the diet would allow greater methylation during the feast period and this methylated epigenome would be copied along with the DNA in the formation of sperm.

The agouti gene in mice was studied by Dr. Randy Jirtle and Dr. Rob Waterland at Duke University in 2003. They found that mice with an unmethylated agouti gene were ravenous, yellow, and obese and were prone to cancer and diabetes. When folic acid was added to the diet of pregnant agouti mice, it altered the gene expression by methylating the agouti gene in the offspring and produced brown, normal weight offspring. (60)

A November, 2008, study done by L. H. Lumey, et al, and published in the *Proceedings of the National Academy of Science* showed an incidence of inherited epigenetic expression during World War II. From November 1944 until April 1945 when they were liberated by the Allies, the Dutch had little to eat except tulip bulbs; thousands starved to death. Children in the womb at that time were born with hypomethylation of the imprinted insulin-like growth factor2 (*IFG2*) gene unlike their siblings who were unexposed to the famine *in utero*. This hypomethalation was caused by the mother's diet which lacked methionine and folic-acid rich foods and led to children who had impaired glucose tolerance, hypercholesterolemia, raised blood pressure, and higher rates of obesity in adulthood. (61)

In 2013, Canadian and American researchers found that children born to obese women who had had the biliopancreatic diversion with duodenal switch (BPD-DS) gastric bypass surgery were much slimmer than their siblings who were conceived when the mother was obese. When the researchers looked at gene methylation and expression levels in the cardiometabolic pathways of these children, they found agreement showing improvements in insulin resistance and in overrepresentation of genes with glucose-metabolic and inflammation-related functions. Because the BPD-DS operation is malabsorptive (rearrangement of the digestive system limits the amount of calories and nutrients the body can absorb) while the purely gastric bariatric banding operations are restrictive (shrinkage of the stomach reduces the amount of food it can hold), the BPD-DS method of gastric bypass surgery was found to help these obese patients and their offspring much more than the purely gastric bariatric banding did. In the purely gastric bariatric banding surgery (or if the obese patient has no gastric banding surgery at all), Dr. John Kral of New York's SUNY Downstate Medical Center who co-authored the study said that fetuses are essentially marinated while in the womb and they're differently marinated depending on the mother's weight and health. Overweight mothers have higher levels of sugar and fat in the bloodstream which in turn makes it to the womb. The results of this methylation study was published in July, 2013, in the *Proceedings of the National Academy of Sciences*.[9] (62, 63, 64, 65, 66, 67)

[9] While epigenetic inheritance has not yet been studied in regards to physical activity, a 2013 study was published in *PLOS Genetics* in which 31 sedentary but healthy men from Malmö, Sweden (mean age=37.4 years;mean MBI=27.8 kg/m2) were asked to attend a one hour spinning class and 2 hours of aerobics each week for 6 months. They were asked not to change their diet and daily activity during the study. It was found that there was a general global increase in adipose tissue DNA methylation in response to the 6 months of exercise. In particular, the exercise intervention was associated with a decrease in waist circumference and waist-hip ratio. The study found there were increased levels of DNA methylation associated with waist-hip ratio. (77)

Other factors affecting the epigenome and inheritance are alcohol consumption, smoking habits, and exposure to other toxic materials, such as bisphenyl A (BPA), stress, and abuse can lead to harmful epigenetic changes.

The relationship between genetics and alcohol consumption leading to alcoholism has been established. Now environmental factors that affect epigenetics and thus gene expression in alcoholics are increasingly being understood. Fetal alcohol spectrum disorders (FASD) are caused by *in utero* exposure to alcohol and have serious consequences for the developing fetus. The severity of the disorders is related to dose, timing, and duration of exposure. Ethanol-exposed fetuses have significantly reduced levels of DNA methylase activity which, in turn, causes gene expression in genes that normally would have been shut off by condensation of the nucleosome. The same has been found in histone methylation and is largely due to folic acid deficiency in the diet of the chronic alcoholic mother. Memory circuits can be affected in the fetus by changes in histone acetylation. MicroRNA has been implicated in possibly causing brain damage in the fetus exposed to prenatal alcohol exposure. FASD affects learning and memory, understanding and following directions, controlling emotions,communicating, and performing daily life activities, such as dressing and feeding. Fetal alcohol syndrome is the most serious type of FASD and includes facial abnormalities, growth problems, and nervous system abnormalities. FASD last a lifetime and require ongoing care throughout the life of the affected person. (95)

A study published in the *European Journal of Human Gene*tics in 2006 showed that men who reported beginning smoking before age 11 had children who did not live as long as their peers and were subject to heart disease. Dr. Marcus Pembrey, a British geneticist, collaborated with Dr. Lars Olav Bygren (author of the Överkalix, Norbotten Swedish study). They used data from the Avon Longitudinal Study of Parents and Children (ALSPAC) and found that of 14,024 fathers, 166 of them said they had started smoking before age 11. Environmentally imprinted factors mark the Y chromosome when the sperm first form. Sons of these 166 early smokers had significantly higher body mass indexes than other boys by age 9. That means that sons of men who smoke in prepuberty will be at higher risk for obesity and other health problems well into adulthood, even if they, themselves, don't smoke. These boys will have a shorter lifespan, just like the overeaters in Överkalix. (59)

A study published in 2012 in *Human Molecular Genetic*s found strong evidence that tobacco use can chemically modify and affect the activity of genes known to increase the risk of developing cancer. Scientists have identified specific epigenetic patterns on the genes of people who smoke. Compared with people who had never smoked, the smokers had fewer methylations on 20 different regions of their DNA. Further study showed the modifications on several sites located in 4 genes that have been weakly linked to cancer. Hypomethalation activates genes and earlier had been shown to cause cancer. This was the first study to establish a close link between epigenetic modificatons on a cancer gene and the risk of developing the disease. (68)

Another study published in 2012 in *Environmental Health Perspective* identified a set of genes including *AHRR* (aryl hydrocarbon receptor repressor) and *CYP1A1* (cytochrome P4501A1) with methylation changes present at birth in children (2nd generation) whose mothers (1st generation) smoked during pregnancy. Aryl hydrocarbon receptor (AHR) is a protein receptor for environmental contaminants which acts as a mediator of chemical toxicity. *CYP1A1* and *AHRR* are genes which are regulated through the AHR-mediated pathway and which are known to be sensitive target genes for toxin exposure. (69, 70) These exposed children are more likely to have problems like low birth-weight, asthma, and possibly obesity, cancer, and high blood pressure. The children will be part of a follow-up study as they get older since it is thought that smoking can affect the reproductive cells of the offspring and perhaps affect the 3rd generation. (71) To prove epigenetics in a woman who smoked around her fetal daughter thus potentially affecting 3 generations (if the fetal daughter's eggs are considered as the 3rd generation), one would need to look at the results of the 4th generation's health. Smoking has been shown to cause altered global methylation. Once one gives up smoking, these altered methylations start to disappear but they never quite match the unmarked DNA of a non-smoker. (72)

In a study by Michael Skinner et al in 2006, it was shown that exposure to vinclozolin, a fungicide sprayed on fruits and vegetables, predisposes male rats to develop a variety of adult-onset diseases earlier than normal, effects which are still detectable over four subsequent generations without being diminished. These alterations in brain and behavior occur in young animals at the time of embryonic sex determination and promote in the F1 generation a spermatogenic cell defect and subfertility in the male. (73, 74)

According to the National Institutes of Health (NIH), bisphenol A (BPA) and phthalates are epigenetically toxic. The epigenetic effect of BPA was clearly demonstrated in viable yellow mice by decreasing CpG methylation upstream of the Agouti gene and by preventing the hypomethylating effect of BPA by maternal dietary supplementation with a methyl donor, such as folic acid. Human placental cell lines exposed to BPA show altered microRNA expression levels, (specifically, miRNA-146a was strongly induced by BPA treatment). (75)

BPA can be found in reusable plastic water and baby bottles, in polycarbonate plastic containers and tableware (labeled #7 with PC underneath the number),and in the lining of canned foods and canned drinks. By choosing glass, porcelain, or stainless steel containers and using plastic containers that are labeled BPA-free, BPA can be avoided. Additionally, avoid microwaving food in plastic food containers containing BPA since BPA can be imparted to food. The Food and Drug Administration (FDA) no longer allows the use of BPA-based materials in baby bottles, sippy cups, and infant formula packaging. (76)

In Quebec in 2010, Michael Meaney et al studied the effect of abuse of children and epigenetic changes. He studied brain samples of 36 human children – 12 who were suicide victims who had been abused as children, 12 who were suicide victims who had suffered no known abuse, and 12 who died suddenly in accidents. Only those who had been abused showed a methylated glucocorticoid receptor (*GR*) gene

that would have led to increased levels of cortisol in the blood. The emotional signs of chronic stress in humans are anxiety, depression, feeling loss of control, nervousness, and difficulty sleeping. [10, 11] (80)

In 2011, a study in Germany by T. Elbert et al used the hypothalamic-pituitary-adrenal axis research, such as the glucocorticoid receptor (*GR*) gene, from Quebec to analyze the methylation status of *GR* in 25 mothers who were exposed to intimate partner violence (IPV) and in their children age 10-19 years after birth. Methylation of the mother's *GR* gene was not affected by IPV. However, they did find methylation of the GR gene of adolescent children by way of epigenetic modifications *in utero* caused by prenatal exposure to cortisol due to IPV experienced by the mother while pregnant. The conclusion of the study was that this may be a plausible mechanism for explaining how prenatal stress may cause adult psychosocial issues. (81)

How Is Epigenetics Related to Cancer?

The role of epigenetic modifications in cancer etiology and progression is well established. Global hypomethylation in cancer was reported in the early 1980s at about the same time the first genetic mutation in an oncogene was discovered. Hypomethylation of the DNA causes genome instability and aberrant transcription initiations, including activation of transposons which may lead to gene disruption. Since then, hypermethylation of the DNA has been reported in some cancers. Hypermethylation affects gene transcription by turning off the gene. For example, hypermethylation of CpG islands in promotor regions of genes, which are usually unmethylated and available to initiate gene expression, would cause genes to turn off. (82)

Histone modification patterns are also altered in human tumors, especially in the promotor regions of tumor suppressor genes which normally are turned on. In particular, levels of trimethylation of histone H4 lysine (K) at amino acid position 20 (H4K20me3) and monoacetylation of histone H4 lysine (K) at amino acid position 16 (H4K16ac) are severely disturbed in cancer cells both globally and at particular loci. The deactylation and disturbance of the trimethylation turn off the tumor suppressor genes. In other parts of the chromosome, gene repeats, (which are normally turned off by methylation) are unmethylated and then acetylated, turning on those genes. (83)

[10]Stress happens to everyone and cortisol (a glucocortoid) is released from the adrenal glands during stressful circumstances and is meant to be part of the fight-or-flight short term response. After the 'danger' is gone, cortisol levels lower and the body goes back to its normal state. However, in people who experience chronic high levels of stress, elevated levels of cortisol continue to circulate in the body over long periods of time. High levels of cortisol in the blood leads to anxiety, nervousness, depression, difficulty sleeping, and feeling loss of control. These are key factors in the development of alcoholism because alcohol consumption can temporarily reduce the anxiety. (78)

[11]Early emotional environments can lead to long lasting changes in the brain. As reported by Michael Meaney et al in *Nature Neuroscience* in 2004, rat pups who were nurtured by their mothers who licked and groomed them were more assertive and confident (less anxious) than those pups that weren't nurtured. The nurtured pups had brains that had an unmethylated (therefore expressed) stress regulation gene called glucocorticoid receptor (GR) in the hippocampus of the brain which lowered cortisol levels in the blood..Pups who were not nurtured by their mothers had a methylated *GR* gene blocking the glucocorticoid receptor from being expressed which caused increased levels of cortisol in the blood, and this change lasted into adulthood. On-going high levels of cortisol lead to anxiety and nervousness in rats. In neglected pups, the stress-altered gene returned to the normal state of activation when the pups were transferred to more attentive mothers. (79)

MicroRNAs with growth-inhibitory functions, such as miRNA-124a and miRNA-34b/c, undergo epigenetic inactivation because the sequences surrounding their respective transcription start sites become hypermethylated. Because they are inactivated, the intended messenger RNA target is translated into protein when it should have been blocked. Instead of normal rates of cellular growth, proliferation, differentiation, and cell death, abnormal rates can lead to formation of tumors. The reduction or deletion of a miRNA can occur at any stage of miRNA biogenesis and ultimately leads to the inappropriate expression of the miRNA target. If the reduction or deletion of a miRNA that functions as a tumor suppressor occurs, a tumor can form. The overall outcome might involve increased cell proliferation, invasiveness or angiogenesis, and/or decreased levels of apoptosis, ultimately leading to tumor formation. (84, 85)

Scientists are examining changes in the cancer epigenome to provide understanding of tumor development and progression, including the acquisition of drug resistance during treatment. These changes provide diagnostic, prognostic, and predictive biomarkers for improving clinical management of patients. (86)

A number of drugs have been developed that target epigenetic methylation alterations in pathological states. The nucleoside analogs 5-azacytidine and hydrophilic decitabine (5-aza-2'-deoxycytidine) are DNMT (DNA methyltransferase) inhibitors which prevent methylation, thus keeping the gene turned on. They are approved by the US Food and Drug Administration (FDA) to treat high-risk myelodysplastic syndrome (MDS). MDS is a group of diseases that cause immature blood cells to accumulate in the bone marrow. This can lead to several types of leukemia and is caused by chemotherapy with alkalating agents, radiation therapy, and exposure to industrial solvents. The drug hydralazine is currently being investigated in clinical trials as a demethylating agent against solid tumors. No drugs targeting histone methylation are FDA approved or in clinical trials. (87)

Histone acetylation is another epigenetic therapy that has been studied. Several HDAC (histone deacetylase) inhibitors are FDA approved, including hydrophobic vorinostat (suberoylanilide hydroxamic acid, SAHA) and the depsipeptide romidepsin. Vorinostat and romidepsin are used to treat cutaneous T-cell lymphoma, a rare cancer of the immune system which attacks the skin. Histone acetylation allows for gene expression so the histone deacetylase inhibitor would allow gene expression to continue. Preclinical trials suggest antitumor activity of the oligoamine analog SL11144, which inhibits LSD1 (part of a deacetylase complex), and the *S*-adenosylhomocysteine hydrolase inhibitor, DZNep, which depletes cellular levels of PRC2 (polycomb repressive complex 2 which has histone methyltransferase activity) components.[12, 13, 14] (88)

Are There Other Diseases That Are Thought To Be Related To Epigenetics?

While epigenetics and cancer have been well established, looking at the epigenetic relationship in other diseases remains sketchy. One disease, rheumatoid arthritis, is a disease that usually affects joints symmetrically and symptoms include painfully swollen and inflamed synovial membranes in the joints. (89, 90) It has long been thought to be an autoimmune disease. However, a study published in *Nature Biotechnology* in January, 2013 found that 4 CpG sites showed an association between genotype and variance of methylation between those with rheumatoid arthritis and the controls who did not have rheumatoid arthritis. (91) With this data, it is suggested that other autoimmune diseases could have an epigenetic relationship.

Are There Further Epigenomic Studies To Be Done?

The Human Genome Project was completed in April, 2003, (92) and by then, plans to do a Human Epigenome Project to study tissue-specific methylation and epigenetic heterogeneity among individuals were well underway. The Human Epigenome Project (HEP) is a public/private collaboration run by the members of the Human Epigenome Consortium. The HEP consortium is made up of The Wellcome Trust Sanger Institute in the United Kingdom, Epigenomics AD in Germany and the United States, and The Centre National de Génotypage in France.

The Human Epigenome Project aims to identify, catalogue and interpret genome-wide DNA methylation patterns of all human genes in all major tissues. Differentially methylated cytosines give rise to distinct

[12]During drug development, many molecules are developed and tested on *in vitro* tumor tissue to see which ones change the epigenomic expression of the gene. A commonly used cell-based assay for testing both DNMT and HDAC inhibitors is the quantification of the re-expression of known epigenetically-silenced genes by reverse transcription polymerase chain reaction (RT-PCR) and western blot analysis. However, this approach is not high-throughput and may produce gene specific results. Other assays that have been used include estimation of global DNA methylation using capillary electrophoresis, DNA digestion with methylation-sensitive restriction enzymes, analysis of specific DNA methylation using bisulfite sequencing, and methylation specific-PCR. However, these assay systems are time-consuming, cumbersome, and subject to misinterpretation.

[13]The most recent method for screening epigenetic drugs in cells is the novel cell-based assay system EPISSAY, reported in BMC (BioMed Central) Cancer in 2013. The EPISSAY bioassay system provides a rapid system to compare the efficiencies of existing and newly formulated drugs that reactivate gene expression and it was developed to compare the activity of different epigenetic drugs. This assay system is based on mammalian MCF10A cells expressing a fusion protein between red fluorescent protein (RFP) and bacterial nitroreductase (TMnfsB) driven by CMV promotor. Epigenetic silencing has been shown to drive genes driven by CMV promotor in both stable transfected cells and transgenic pigs. Silenced CMV promotor driven genes were shown to be reactivated after treatment with epigenetic drugs, such as decitabine. (88)

[14]Once the molecules are experimentally developed and tested on cell-assays *in vitro*, then they must be tested *in vivo* before they can be approved by the Food and Drug Administration (FDA). First the drugs are tested on non-human animals, like mice or rats, in preclinical trials and then tested on cancer patients in clinical trials. If the drug works on a small number of people in phase I without deleterious side effects, then it will continue in phase II and III trials. If the drug continues to work and side effects are minimal, then it will be approved by the FDA for use with the general public.

patterns specific for tissue type and disease state. Such methylation variable positions (MVPs) are common epigenetic markers. A subset of the DNA methylation patterns within a cell are characteristic tothat cell type. Cell type-specific and tissue-specific DNA methylation have organ-to-organ variations in the clusters of methylated CpGs within the same individual. Despite overall consistency in tissue-specific DNA methylation patterns, variations in these patterns exist among different individuals.

The HEP study was conducted between 1999 and 2006 and they have identified MVPs in approximately 150 loci in the Major Histocompatibility Complex in chromosome 6 , which is associated with many diseases. In 2003, HEP released about 135,000 CpG methylation sites in a variety of individuals. (93, 94)

The National Institutes of Health (NIH) Roadmap Epigenomics Program began in 2008 as a consortium to investigate the epigenomic origins of health and susceptibility to disease. In January, 2015, the group reported having mapped the epigenome of more than 100 tissues and cell types and their findings were published in the journal *Nature*, February 19, 2015. (98, 99)

Bibliography

1. oxforddictionaries.com/us/definition/american_english/epigenetic
2. dictionary.reference.com/browse/epi-
3. www.nature.com/ni/journal/v11/n7/full/ni0710-565.html
4. www.ncbi.nlm.nih.gov/pmc/articles/PMC2791696/
5. www.ndsu.edu/pubweb/~mcclean/plsc431/eukarychrom/eukaryo3.htm
6. **grants.nih.gov/grants/guide/rfa-files/RFA-RM-07-013.html**
7. Biochemistry, Voet, Donald & Voet, Judith G., 2011, John Wiley & Sons
8. www.cs.hunter.cuny.edu/~saad/courses/compbio/lectures/lecture9.pdf
9. www.helicase.pbworks.com/w/page/17605706/Sabrie-Breland
10. www.ncbi.nlm.nih.gov/pubmed/11891299
11. www.ncbi.nlm.nih.gov/pmc/articles/PMC1292996/
12. www.georgiahealth.edu/cancer/people/robertson/research.html
13. www.ghr.nlm.nih.gov/gene/DNMT1
14. www.sciencedirect.com/science/article/pii/S096921260800049X
15. www.ncbi.nlm.nih.gov/gene/1788
16. www.ghr.nlm.nih.gov/gene/DNMT3B
17. hmg.oxfordjournals.org/content/9/1
18. www.nature.com/ng/journal/v43/n4/full/ng0411-289.html
19. www.ncbi.nlm.nih.gov/gene/29947
20. www.ncbi.nlm.nih.gov/pmc/articles/PMCI1459120/

21. *http://www.ncbi.nlm.nih.gov/pubmed/9305837*
22. *http://www.ncbi.nlm.nih.gov/pubmed/19138067*
23. www.ncbi.nlm.nih.gov/books/NBK45788/?report=printable
24. www.ncbi.nlm.nih.gov/pmc/articles/PMC3193420/
25. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3469451/
26. ghr.nlm.nih.gov/gene/EHMT1
27. depts.washington.edu/phcol/faculty/gardner.php
28. www.pnas.org/content/100/23/13225
29. www.genetik.uni-koeln.de/groups/Dohmen/ResearchTopics.html
30. http://www.nature.com/nrg/journal/v10/n5/full/nrg2540.html
31. **www.nature.com/nature/journal/v461/n7262/fig_tab/461359a_F1.html**
32. **www.ncbi.nlm.nih.gov/pmc/articles/PMC1779566**
33. http://www.nature.com/nrg/journal/v3/n9/execsumm/nrg887.html
34. www/ncbi.nlm.nih.gov/projects/genome/probe/doc/TechRnai.shtml
35. www.nature.com/nrc/journal/v6/n4/fig_tab/nrc1840_F1.html
36. www.nature.com/nrd/journal/v6/n6/fig_tab/nrd2310_F1.html
37. www.nature.com/nrg/multimedia/mai/index.html
38. www.merckmanuals.com/home/womens_health_issues_pregnancy/stages_
39. http://www.nature.com/nrg/journal/v10/n5/execsumm/nrg2540.html
40. www.rcsb.org/pdb/101/motm.do?momID=139
41. http://www.sciencedirect.com/science/article/pii/S0955067411000226
42. www.ncbi.nlm.nih.gov/pmc/articles/PMC5174789/
43. learn.genetics.utah.edu/content/epigenetics/inheritance/
44. voices.yahoo.com/genomic-imprinting-1599959.html?cat=58
45. www.nature.com/news/2002/020627/full/news020624-6.html
46. herkules.oulu.fi/isbn9514270274/html/c273.html
47. ghr.nlm.nih.gov/condition/prader-willi-syndrome
48. www.fpwr.ca/about-prader-willi-syndrome/
49. ananceleste.hubpages.com/hub/Puppet-Children-Angelman-Syndrome
50. ghr.nlm.nih.gov/condition/angelman-syndrome
51. www.ncbi.nlm.nih.gov/pmc/articles/PMC384940/
52. faculty.clintoncc.suny.edu/faculty/michael.gregory/files/Bio%20100
53. faculty.virginia.edu/mammgenetics/805-2ndXo3.html
54. www.nature.com/scitable/topicpage/X-Chromosome-X-Inactivation-323
55. www.nature.com/scitable/topicpage/Genetic-Imprinting-and-X-Inactivation-1066
56. www.biotek.com/resources/articles/automation-generic-fluorescent-methyltransferase.html
57. http://cen.acs.org/articles/87/i37/Mapping-Epigenome.html
58. www.scientificamerican.com/article.cfm?id=identical-twins-exhibit-d

59. www.time.com/time/magazine/article/0,917,1952313-2000.html
60. *http://www.nature.com/news/2003/030728/full/news030728-12.html*
61. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2579375/
62. http://www.pnas.org/content/110/28/11439.full
63. http://www.bariatric-surgery-source.com/types-of-bariatric-surgery.html
64. www.nutritionremarks.com/2013/05/10/relationship-between-obese-moms-and-kids-epigenetics-home-environment-or-both
65. www.sciencedaily.com/releases/2013/06/130611111708.htm
66. http://www.cdc.gov/healthyyouth/obesity/facts.htm
67. http://www.nhlbi.nih.gov/health/health-topics/topics/obe/risks
68. news.sciencemag.org/2012/12/smoking-smothers-your-genes
69. www.ncbi.nlm.nih.gov/pmc/articles/PMC2712099
70. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2701375/
71. ehp.neihs.nih.gov/september-podcast
72. www3.imperial.ac.uk/newsandeventsgrp/imperialcollege/newssummary/news_6-11-2012
73. www.pnas.org/cgi/doi/10.1073/pnas.1118512109
74. http://press.endocrine.org/doi/pdf/10.1210/en.2005-1058
75. www.ncbi.nlm.nih.gov/pmc/articles/PMC3431850/
76. www.fda.gov/NewsEvents/PublicHealthFocus/ucm064437.htm
77. www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1003572
78. pubs.niaaa.nih.gov/publications/arcr344/495-505.htm
79. http://www.nature.com/nrn/journal/v11/n9/full/nrn2897.html
80. www.ncbi.nlm.nih.gov/pmc/articles/PMC2944040
81. www.nature.com/tp/journal/v1/n7/full/tp201121a.html
82. www.nature.com/nm/journal/v17/n3/fig_tab/nm.2305_F1.html
83. http://www.nature.com/nrg/journal/v8/n4/execsumm/nrg2005.html
84. www.the-scientist.com/?articles.view/articleNo/30739/title/Epigenetic-Changes-in-Cancer.html
85. **www.nature.com/nrc/journal/v6/n4/fig_tab/nrc1840_F2.html**
86. www1.imperial.ac.uk/resources/B80D3D73-9DAA-4AB6-8559-24BFC332AC79/epigenetics_large.jpg
87. www.nature.com/nbt/journal/v28/n10/fig_tab/nbt.1678_ft.html
88. www.ncbi.nlm.nih.gov/pmc/articles/PMC3637807
89. www.healthcentral.com/rheumatoid-arthritis/causes-000048_1-145.html
90. www.medicinenet.com/rheumatoid_arthritis/article.htm
91. www.nature.com/nbt/journal/v31/n2/full/nbt.2487.html
92. http://www.genome.gov/10001772
93. http://www.epigenome.org/index.php?page=project
94. http://en.wikipedia.org/wiki/Human_Epigenome_Project

95. http://pubs.niaaa.nih.gov/publications/arcr351/37-46.htm
96. http://www.ncbi.nlm.nih.gov/pubmed/21170357
97. http://mcb.asm.org/content/30/13/3216.long
98. http://www.nih.gov/news/health/feb2015/niehs-18.htm
99. http://www.nature.com/nature/journal/v518/n7539/full/nature14248.html