1

How Do Genes Work?

In this chapter, we will discuss fundamental concepts central to molecular biology and to understanding epigenetics. We will briefly discuss DNA, genes and proteins, mutation, genotype and phenotype and the relationships between each, and finally provide an overview as to how genes are controlled, outlining gene regulation and repression. We will outline the concept of molecular homeostasis and the interaction between genome and environment, briefly differentiating between epigenetics and genetics.

1.1 DNA, Genes and Proteins

DNA is comprised of four nucleotide bases: adenine, guanine, cytosine and thymine. DNA is a stable double helix with guanine always pairing with cytosine and adenine always pairing with thymine, or uracil in the case of RNA. The human genome is diploid, so it is comprised of 2 sets of 23 chromosomes, 1 inherited from each parent. Every nucleated cell contains the same 46 chromosomes with the same genetic information – this is our genome. Approximately 10% of the human genome is estimated to be coding, that is, specifically encodes for proteins; the remainder is noncoding DNA including repetitive sequences such as microsatellites, minisatellites, transposable elements (SINES, LINES), satellite DNA and triplet repeats.

Epigenetics and Health: A Practical Guide, First Edition. Michelle McCulley.

Type of region	Description
Protein-coding genes	Segments of DNA that encode proteins.
Noncoding RNA genes	Segments of DNA that encode RNA molecules that do not translate into proteins. Examples include microRNAs and long noncoding RNAs.
Regulatory regions	DNA sequences that control gene expression, such as promoters, enhancers, silencers and insulators. These regions can be located upstream or downstream of the gene, or even within the gene itself.
Repetitive DNA	Segments of DNA that are repeated many times throughout the genome, such as transposable elements, satellite DNA and minisatellites.
Intergenic regions	DNA sequences located between genes that do not have any known function. These regions make up the majority of the human genome.
Introns	Segments of DNA located within genes that do not encode proteins. Introns are transcribed into RNA but are removed during the process of RNA splicing, which generates the final mRNA transcript that is used to produce proteins.

The human genome project suggests that there are around 20 000 genes, with each human chromosome on average containing 1300 genes. Genes are transcribed and translated to produce their encoded protein. This two-stage process takes the message embedded in double-stranded DNA and transcribes the coding sequence as single-stranded mRNA, which is then translated by the ribosome using tRNA and rRNA to produce the protein. Genes range in size from 1 kb, as in the case of insulin, to 2.5 Mb for larger genes, such as dystrophin. Almost all genes contain introns and exons. Exons are expressed coding regions and introns are non-expressed intervening sequences. Introns are transcribed into primary RNA and then spliced out of mature RNA in the cytoplasm. The average number of exons for a human gene is 9, and therefore introns are 8. However, there is considerable variation, e.g. 79 for dystrophin and 3 for beta-globin. The average size of an exon is 145 bp.

In addition to introns and exons, genes also have an adjacent upstream (5′) regulatory promoter region as well as other regulatory sequences such as enhancers, silencers and sometimes a locus control region. The promoter region contains specific conserved sequences such as TATA box, CG box and CAAT box, which provide binding sites for transcription factors. The first and last exons also contain untranslated regions (UTRs) known as the 5′UTR and 3′UTR. The 5′UTR signals the start of transcription and contains ATG, the initiator codon that initiates the site of the start of translation. The 3′UTR contains a termination codon, which

marks the end of translation, plus nucleotides that encode a sequence of adenosine residues known as the poly (A) tail; the addition of a poly (A) tail is an essential step in the process of transcription that enables the pre-mRNA to exit the nucleus and move into the cytoplasm for translation.

Gene regulatory sequence	Function
Promoters	DNA sequences located upstream of the transcription start site that recruit the transcriptional machinery to initiate gene expression
Enhancers	DNA sequences that can be located far away from the gene they regulate and can enhance gene expression by increasing transcription rates and/or making expression more specific to certain cells or conditions
Silencers	DNA sequences that can be located near or far from the gene they regulate and can reduce or turn off gene expression
Insulators	DNA sequences that act as boundaries between different gene regulatory regions, preventing their influence on each other
Scaffold/matrix attachment regions	DNA sequences anchor the chromatin to the nuclear matrix, thereby organising and stabilising the structure of chromatin and regulating gene expression
CpG islands	Regions of DNA that have a high density of CpG dinucleotides are often associated with gene promoters and are involved in the regulation of gene expression through DNA methylation
miRNA target sites	Specific RNA sequences within the 3′ UTR of messenger RNAs that are recognised by microRNAs and lead to translational repression or mRNA degradation
Cis-regulatory modules	Clusters of enhancers, silencers and promoter sequences that work together to regulate the expression of a specific gene or set of genes in response to different signalling pathways

Pre-mRNA contains the entire transcribed sequences of exons and introns; the introns must be removed to produce mRNA that comprises the precise code for translation into a functional protein product. The removal of introns occurs through splicing, which occurs in a spliceosome, itself composed of hundreds of proteins and 5 RNAs. Once a transcript has been spliced and its 5′ and 3′ ends modified, a piece of mature functional mRNA has been produced, suitable for translation into a protein. There are specific nucleotide recognition sequences that aid in splicing. These sequences are present: towards the end of an exon, at the

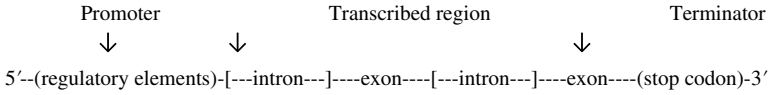

Figure 1.1 The gene is shown in a linear arrangement, with the promoter at the beginning, the transcribed region in the middle and the terminator at the end. The transcribed region is shown with a series of exons (which code for protein) separated by introns (which do not code for protein). The regulatory elements, which can be located upstream or downstream of the promoter, are also shown at the beginning of the gene. At the end of the transcribed region, there is a stop codon that signals the end of the protein-coding sequence.

beginning of an intron, at the end of an intron, at the beginning of the next exon and at a region within the intron but close to the 3′ end, which provides a binding site for intron removal. Such sequences are recognised by small nuclear ribonucleoproteins (snRNPs), and they cut the RNA at the intron-exon borders and connect the exons together. Alternative splicing is a mechanism that allows more information to be packed into a single gene. That is, from a single gene, through splicing exons together in different combinations, multiple RNAs and functional proteins can be produced; this allows cells to produce related but distinct proteins from a single gene. For example, one type of protein may be produced in one tissue, whereas another form may be produced in another tissue (Figure 1.1).

The estimated 20 000 protein-coding genes comprising the human genome are spread between the lengths of the human chromosomes. Each gene has a promoter where transcription of mRNA by RNA polymerase II is initiated by transcription factors. There are also remote elements called enhancers that will modulate the activity of the promoter. Different cell types activate gene expression in a different manner by making use of the genome's extensive system of regulatory *cis*-acting elements. The mechanisms underlying this process involve physical changes to the chromatin that either promote or inhibit gene expression. This is fundamental to understanding epigenetics, as it is how gene expression is suppressed or enhanced.

1.2 Gene Control, Homeostasis and Epigenetics

In complex multicellular organisms, differential gene expression is fundamental during embryonic development and in the maintenance of the adult state. It is key to understand that different cells make different proteins, which means different genes are switched on in different tissues even though all cells carry the same comprehensive set of genetic instructions. Therefore, there must be a way in which the body controls which gene is switched on, to what extent and when. Unused genetic information is not discorded – just not switched; for this to happen, there are specific mechanisms that can activate specific portions of the

genome and repress the expression of other genes. The activation and repression of genetic loci can be seen as a form of molecular homeostasis, a delicate balancing act for a healthy organism, given expression of the wrong gene at the wrong time in the wrong cell type or in the wrong amount can lead to a harmful phenotype even when the gene itself is normal, such as in cancer or cell death.

Housekeeping genes need to be expressed in all types of nucleated cells because they encode a vital product needed to fulfil an important cellular function, for example, protein synthesis, energy production. Many other genes, however, show a much more restricted pattern of expression that is tissue specific. Spatial restriction of gene expression can occur at many distinct levels: multiple organ/tissue, specific tissue/cell lineage/cell type, individual cells and intracellular distribution. Similarly, gene expression is also temporally restricted at various levels: cell cycle stage, developmental stage, differentiation stage and inducible expression.

Gene expression requires two events to occur: first, the recruitment and activation of chromatin remodelling enzymes that alter the structure of the nucleosome and make the promoter sites on the DNA accessible. The second is the recruitment of coactivators to help assemble the factors needed for transcription; this includes transcription factors, RNA polymerase II, etc. Most eukaryotic genes are regulated at the transcriptional level either via the complex interplay between transcription factors, gene promoters and enhancers and/or through epigenetic mechanisms including chromatin remodelling and DNA methylation. Posttranscriptional gene regulation is through the regulation of splicing and mRNA processing, regulation of mRNA transport, degradation of mRNA, translational regulation or by modifying the translated protein to alter its activity.

Chromatin remodelling is essential for many processes including transcription, replication, DNA repair and recombination. Many of these changes are mediated by chemical changes to the N-terminal tails of core histones in a nucleosome. The histone tails are exposed on the surface of the nucleosome and are amenable to modification through acetylation, phosphorylation and methylation, allowing access to the underlying DNA. Acetylation adds an acetyl group to specific lysine residues; acetylated histones are found in regions of open chromatin where transcription occurs, which allows transcriptional enzymes to have access to the DNA. Acetylation is mediated by histone acetyltransferases (HATS) and deacetylation by histone deacetylases (HDACs) – both important to gene expression.

1.3 DNA Methylation and Regulation of Gene Expression

The DNA of most eukaryotes is modified after replication through the addition of methyl groups to bases and sugars. Base methylation most often involves enzyme-mediated addition of methyl groups to cytosine; in most eukaryotes, approximately 5% of cytosine residues are methylated; however, the extent of methylation

can be tissue specific and can vary from less than 2% to over 7%. In eukaryotes, low amounts of methylation are associated with elevated levels of gene expression and elevated levels of methylation are associated with low levels of gene expression. In mammalian females, the inactivated X chromosome has a higher level of methylation than the active X chromosome. Methylation patterns are tissue specific and, once established, heritable for all cells of that tissue. The precise mechanism as to how methylation affects gene regulation is still uncertain. It is possible that there are proteins that bind to the methyl group attached to the cytosine; these proteins could recruit corepressors or HDACs to remodel chromatin, changing it from open to closed.

1.4 Post-transcriptional Regulation of Gene Expression

As outlined above, primary mRNA resulting from transcription is then further modified prior to translation – noncoding introns are removed, exons are spliced together and mRNA is modified through 5′ cap and poly A tail, prior to export to the cytoplasm for translation. Whilst many opportunities exist for further regulation during these steps, the major two are alternative splicing and regulation of the stability of the mRNA itself. Alternative splicing can generate multiple forms of a protein, so the expression of one gene can produce a family of related proteins.

Post-translational regulation	Mechanism	Examples
Protein phosphorylation	Addition of phosphate group to serine, threonine or tyrosine residues	Activation of Cyclin-dependent kinase 1 during cell cycle progression
Protein methylation	Addition of methyl group to arginine or lysine residues	Histone methylation by EZH2 in Polycomb repressive complex
Protein acetylation	Addition of acetyl group to lysine residues	Histone acetylation by histone acetyltransferases (HATs) in chromatin remodelling
Protein ubiquitination	Addition of ubiquitin to lysine residues	Ubiquitination of tumour suppressor p53 leading to its degradation
Protein sumoylation	Addition of small ubiquitin-like modifier (SUMO) protein to lysine residues	SUMOylation of transcription factor Sp3 leading to repression of its activity

Post-translational regulation	Mechanism	Examples
Protein neddylation	Addition of NEDD8 to lysine residues	Neddylation of Cullin proteins leading to activation of Cullin-RING E3 ubiquitin ligases
Protein glycosylation	Addition of carbohydrate to serine, threonine or asparagine residues	Glycosylation of E-cadherin promoting its adhesive function in cell–cell junctions
Protein lipidation	Addition of lipid group to cysteine residue	Palmitoylation of Hedgehog protein for proper signalling in development

1.5 Promoters and Enhancers

Transcription in eukaryotes is controlled through the interactions of promoters and enhancers; there are other localised sequence-based elements such as the CCAAT box, which also have an important regulatory effect and are found closer to the gene itself within the promoter region. Enhancers control the rate of transcription and can be located before, after or within the gene expressed. Transcription factors are proteins that bind to DNA-recognition sequences within promoters and enhancers and then activate transcription through protein–protein interaction. A gene can have different methylation patterns in different tissues and will be correlated with different regulatory patterns.

As mentioned previously, there are two types of regulatory sequences that control gene transcription: promoters and enhancers. Promoters are the recognition point for RNA polymerase binding; they are located immediately next to a gene, typically several hundred nucleotides long and need to be able to allow binding of RNA polymerase II to transcribe primary mRNA. The promoter region has several key elements that aid in its recognition by the polymerase enzyme; the promoter itself – the TATA box (made of 8-bp consensus sequence only AT base pairs), typically flanked either side by GC-rich regions. Many promoters also contain CAAT box and GC box- both bind transcription factors and function like enhancers- mutations in either can affect the rate of transcription. RNA polymerase II requires transcription factors to aid in the start of transcription; they are assembled at the promoter in a specific order and provide a platform that RNA polymerase II can recognise and bind to. The organisation of the upstream region of a gene including the promoter region is variable with respect to the nature, number and arrangement of controlling elements, in some cases including sites for enhancer binding and tissue-specific enhancer binding.

Enhancers can be on either side of a gene, so upstream (5′) or downstream (3′), they can also be at a reasonable distance from the gene or even within the gene itself. If they are adjacent to the gene, they are termed *cis*-regulators, as opposed to *trans* regulators such as binding proteins, which can regulate a gene on any chromosome. Typically, enhancers will interact with multiple regulatory proteins and transcription factors to increase the rate of transcription efficiency or promoter activation. Within enhancer sites, it is common to find binding sites for both positive and negative gene regulators. The main difference between enhancers and promoters is that enhancers do not have a fixed position (i.e., can be upstream, downstream or within the gene they regulate); the orientation of the enhancer can be inverted without significant effect on its action, and if an unrelated gene is placed near an enhancer, the transcription of that gene becomes enhanced. Enhancers are responsible for time- and tissue-specific expression. They exert their effect when the transcription factor binds to the enhancer, altering the chromatin configuration, and second, through binding/looping DNA, they bring distant enhancers and their promoters into direct contact to form complexes with transcription factors and polymerases. The new resulting configuration increases the overall rate of RNA synthesis.

1.6 Mutation Genotype, Phenotype, Epigenotype

Now we understand the structure and function of a gene, which is to carry instructions to produce a specific protein, we can begin to understand what the consequences are of mutations to DNA and disruptions to normal control over gene expression, that is, the switching on and off of a gene. In the simplest terms, the DNA spanning a particular area of a chromosome will contain the information needed to produce a particular protein in response to a signal to do so. If there are disruptions or changes to the DNA sequence of a particular gene, then this can have an impact on how or even whether the protein is produced. Given the complexity of transcription and translation, including the splicing of introns, exons and mechanisms that control gene expression, there are multiple locations where this can have a detrimental effect in terms of influencing the production of a protein. The resulting physical/physiological effect or the effect on a person's health is termed the phenotype; the change to the DNA is termed the genotype. There are also changes to the gene control mechanisms external to the DNA; these are termed epigenotypes.

What then is the significance of mutation? Are all mutations equal? Fundamental to molecular pathology is trying to understand why a particular genetic/epigenetic change or genotype should result in a specific phenotype or clinical condition. When trying to understand mutation, we can take a tiered approach, with

the first stage being to ascertain whether the mutation results in loss of function or gain of function. With a loss of function, the protein product will have reduced or no functional capability. With a gain of function mutation, the protein product will do something atypical or unusual. Deletions of a whole gene, nonsense mutations and frameshifts are almost certain to destroy gene function. Mutations that change conserved sequence flanking introns affect splicing and will typically knock out the function of the gene; splicing can be changed by many other sequence changes too. Missense mutation is more likely to be pathogenic if it affects a part of the protein that is known to be functionally important, for example, a DNA-binding domain. Changing an amino acid is more likely to affect function if that changed amino acid happens to be conserved in related genes. Amino acid substitutions are more likely to affect function if they are non-conservative.

We will first look at the main classes of DNA mutations, which are deletions, insertions, frameshifts, dynamic mutations and single base substitutions (including missense mutations, nonsense mutations and splice site mutations). If we look at these from the perspective of how you stop a gene from working, then it can help to determine what meaningful interpretations of such data are. Genetic variants can also have an impact on epigenetic mark; single SNPs can have an impact on more than one proximal CpG sites. Correlation between SNPs and methylation is referred to as mQTLs – methylation quantitative trait loci. The added complexity is that SNPs can affect both gene expression and methylation independently, as outlined in Table 1.1.

1.7 Conclusion

Virtually any disease is the result of the combined action of genes and the environment, but the relative role of the genetic component may be large or small. For many of the common diseases affecting humans, such as diabetes, heart disease and cancer, there is not a single gene responsible. We are complex organisms, and as such, our pathophysiology is often a combined effect of what we inherit and our individual environment or circumstances.

There are many contributory factors to human variation both genetic, environmental and the interaction between the two. In terms of genetic differences, this is the existence of alternative forms of a gene (alleles) in the population. These variations can occur at greater or lesser frequency in a population, can be well established from a distant ancestor or can be new recently occurring variants. The impact of any change in genetic code depends on the resulting impact that change has on the production and activity of the encoded protein. Therefore, it is commonplace for similar phenotypes to arise as a consequence of mutation and variation at different loci.

Table 1.1 How do you stop a gene from working?

How to disrupt gene function	Example
Delete entire gene or part of the gene	Normal gene sequence: 5′--ATGCGTTAACCGGTACCG--3′ Met-Arg-Asn-Arg-Thr Mutated gene sequence: 5′--ATGCGTTCACCGGTACCG--3′ Met-Arg-His-Arg-Thr In this example, a single nucleotide mutation has changed the third codon from 'AAC' to 'CAC', which results in the substitution of the amino acid asparagine (Asn) with histidine (His). This change can potentially affect the function of the protein product of this gene, depending on its role and the location of the mutated residue within the protein structure. If the mutation occurs in a critical part of the protein, it may disrupt its activity or stability, leading to a loss of function. Alternatively, if the mutation occurs in a non-critical region, the protein may still be functional, but its activity may be altered or reduced. In some cases, mutations can also create new or altered functions, which can have both positive and negative consequences.
Insert a sequence into the gene	
Disrupt the gene structure • By translocation • By inversion	
Prevent the promoter working • By mutation • By methylation	
Destabilise mRNA	
Polyadenylation site mutation	
Nonsense-mediated RNA decay	
Prevent correct splicing • Inactivate donor splice site • Inactivate acceptor splice site • Alter an exonic splicing enhancer • Activate a cryptic splice site	
Introduce a frameshift in translation	
Convert a codon into a stop codon	
Replace an essential amino acid	
Prevent post-transcriptional processing	
Prevent correct cellular localisation of product	

This table summarises the many ways in which gene function can be disrupted.

Genes and their encoded protein products do not act in isolation; most genes will have multiple roles and functions, often depending on the cell type, tissue and developmental stage of the organism. There are many complex mechanisms that regulate when and how a gene is transcribed and translated into a protein and how the protein product is then modified and contributes to a particular biochemical pathway. Understanding human health and disease is therefore incredibly complex and far more than a simple genotype/phenotype correlation, with a need to further our understanding of the importance of gene–gene and gene–environmental interactions in disease.

In addition to mutation, epigenetic modifications and miRNAs can enhance or suppress gene expression. In the following chapter, we will explore further DNA

methylation, the post-translational modifications of histones, chromatin remodelling and non-coding RNAs and their roles in controlling gene expression. As we have already discussed, these epigenetic mechanisms are intricately connected to developmental plasticity and cellular differentiation and being a means to relay environmental influences to the cell nucleus, bridging the gap between lifestyle and the genome.

Task

Questions to ask yourself about your research

- *Is epigenomics relevant to my research?*
- *How is your research interest an epigenetic/epigenomic issue? Or how is it not?*

Task 1. Make a mind map of known contributing factors to your disease/health phenotype of interest. Think about the condition, the associated phenotypes, the organism you are interested in studying, the time course of the disease and known contributory environmental and genetic factors. Think about whether there are differences in terms of gender, age, ethnicity and occupation. Think about how environmental and lifestyle factors contribute to the aetiology; what is known about the epigenetics of your condition?

Look at the obesity foresight map to see an incredibly detailed map.
Link to foresight map: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/296290/obesity-map-full-hi-res.pdf

Further Reading

Tompkins JD. Discovering DNA Methylation, the history and future of the writing on DNA. *J Hist Biol*. 2022 Dec;55(4):865–87. doi: 10.1007/s10739-022-09691-8. Epub 2022 Oct 14. PMID: 36239862; PMCID: PMC9941238.

This paper explores the history and future of the discovery of DNA methylation. The author provides a comprehensive overview of the key scientists and research breakthroughs that have contributed to our understanding of DNA methylation, including the discovery of 5-methylcytosine in the 1950s and the development of techniques for mapping DNA methylation patterns in the genome. The article also discusses the potential applications of DNA methylation research in fields such as epigenetics, cancer research and personalised medicine. The author concludes that continued research

in this field has the potential to transform our understanding of the role of DNA methylation in human health and disease and to lead to the development of new diagnostic and therapeutic tools.

Choudhuri S. From Waddington's epigenetic landscape to small noncoding RNA: some important milestones in the history of epigenetics research. *Toxicol Mech Methods*. 2011 May;21(4):252–74. doi: 10.3109/15376516.2011.559695. PMID: 21495865.

This paper provides an historical overview of the field of epigenetics, tracing its origins from Conrad Waddington's concept of the 'epigenetic landscape' in the 1940s to the discovery of small noncoding RNA in the 1990s. The author discusses key milestones in the development of the field, including the discovery of DNA methylation in the 1970s and the identification of histone modifications and chromatin remodelling complexes in the 1980s. The paper also highlights some of the key challenges facing the field of epigenetics, including the need for more comprehensive and systematic analyses of epigenetic marks and their functional consequences.

Van Soom A, Peelman L, Holt WV, Fazeli A. An introduction to epigenetics as the link between genotype and environment: a personal view. *Reprod Domest Anim*. 2014 Sep;49(Suppl 3):2–10. doi: 10.1111/rda.12341. PMID: 25220743.

In this article, the authors introduce the field of epigenetics and its role as a link between genotype and environment. The article first defines epigenetics and its various mechanisms, including DNA methylation, histone modification and non-coding RNAs. The authors then discuss how epigenetic modifications can be influenced by environmental factors, such as nutrition, stress and toxins, and how these modifications can have downstream effects on gene expression and disease susceptibility. The article concludes by highlighting the importance of understanding epigenetic regulation in the context of reproductive biology and animal breeding, as well as in the field of human health and disease.

Rivera RM, Bennett LB. Epigenetics in humans: an overview. *Curr Opin Endocrinol Diabetes Obes*. 2010 Dec;17(6):493–9. doi: 10.1097/MED.0b013e3283404f4b. PMID: 20962634.

This paper provides an overview of epigenetics in humans. They discuss the different types of epigenetic marks, such as DNA methylation, histone modifications and non-coding RNA regulation, and their impact on gene expression. The authors also describe how epigenetics is involved in various processes, including development, ageing and disease. They discuss the role of environmental factors, such as diet and stress, in shaping the epigenome and influencing health outcomes. Lastly, the authors touch on the potential clinical applications of epigenetics, such as epigenetic biomarkers for disease diagnosis and treatment.

Peixoto P, Cartron PF, Serandour AA, Hervouet E. From 1957 to nowadays: a brief history of epigenetics. *Int J Mol Sci*. 2020 Oct 14;21(20):7571. doi: 10.3390/ijms21207571. PMID: 33066397; PMCID: PMC7588895.

This paper provides a brief history of the field of epigenetics from 1957 to the present day. The authors discuss the key discoveries and milestones in the field, including the identification of histone modifications, DNA methylation and the role of non-coding RNAs in epigenetic regulation. The article also highlights the importance of epigenetics in various biological processes, including development, ageing and disease. The authors conclude by discussing the current state of epigenetics research and the potential for epigenetic therapies for disease treatment and prevention.

Bošković A, Rando OJ. Transgenerational epigenetic inheritance. *Annu Rev Genet*. 2018 Nov 23;52:21–41. doi: 10.1146/annurev-genet-120417-031404. Epub 2018 Aug 30. PMID: 30160987.

This article focuses on transgenerational epigenetic inheritance, which refers to the transfer of epigenetic information across multiple generations in the absence of changes to the underlying DNA sequence. The authors describe various mechanisms of transgenerational epigenetic inheritance, including DNA methylation, histone modification and small RNA-mediated processes, and highlight the role of these mechanisms in various model organisms. They also discuss the potential impact of environmental exposures on transgenerational epigenetic inheritance and the implications for human health and disease. The authors suggest that further research in this area is necessary to fully understand the mechanisms underlying transgenerational epigenetic inheritance and its potential implications.