

Epigenetics and Environmental Effects on Gene Expression

Bhavya Gupta

Abstract

This paper explores how environmental stress can influence epigenetic regulation and change the way genes are expressed without altering the underlying DNA sequence. It focuses on mechanisms such as DNA methylation, histone modifications, chromatin structure, and non-coding RNAs, with particular attention to how these processes help plants respond to stresses such as drought, heat, cold, and salinity. The paper uses a literature-based approach, comparing findings from human studies, controlled plant and animal experiments, and epigenome-wide research rather than generating new experimental data. Different research methods, including bisulfite sequencing, ChIP-seq, and RNA sequencing, were also examined to understand how epigenetic changes are detected and analysed.

The literature shows that environmental exposure can produce measurable changes in the epigenome and influence stress-response genes, although these effects are often context- and tissue-specific. Comparisons between cold and heat stress also show that plants do not rely on one general epigenetic response. Some epigenetic changes may persist and contribute to stress memory, while studies of sperm RNA suggest that chemically modified RNAs can also carry environmental information between generations. However, distinguishing causation from correlation remains difficult, particularly because of reverse causation, tissue-access limitations, and other confounding factors. Overall, epigenetics appears to add an important regulatory layer to heredity and gene expression, but many questions about causal mechanisms, stability, and inheritance remain open.

I. Introduction

Background of the topic

Epigenetics refers to changes in gene expression that occur without changing the underlying DNA sequence. This makes it different from a genetic mutation, where the sequence of DNA itself is altered. Instead, epigenetic mechanisms determine how accessible certain genes are and whether they are expressed or kept silent. The four major mechanisms involved are DNA methylation, histone modification, chromatin remodeling, and non-coding RNA. DNA methylation, particularly around CpG-rich promoter regions, can suppress transcription by making the region less accessible. Histone modifications also affect accessibility; acetylation generally loosens the interaction between DNA and histones, allowing transcription machinery to reach the DNA more easily.

Chromatin remodeling controls gene expression by repositioning nucleosomes and changing how exposed a region of DNA is. Non-coding RNAs add another layer of regulation by interacting with regulatory proteins and, in some cases, directing them towards specific regions of the genome to influence gene activity. These mechanisms can work together to determine whether a gene is active or silent, while the actual DNA sequence remains unchanged. Epigenetics therefore provides cells with a way to regulate the use of their genetic information without having to alter the information itself.

Existing research

One of the landmark examples of environmental epigenetics comes from the Dutch Hunger Winter of 1944–45, when severe famine exposed pregnant women and their developing babies to prolonged nutritional deprivation. Decades later, researchers examined people who had been exposed to the famine before birth and compared them with their unexposed same-sex siblings. Those who had experienced famine during development showed lower DNA methylation at the IGF2 gene than their siblings. Since IGF2 plays an important role in growth and development, this finding suggested that an environmental condition experienced before birth could leave a measurable epigenetic mark many decades later.

The importance of the Dutch Hunger Winter studies goes beyond the IGF2 finding itself. People exposed to famine before birth were also found to have higher risks of health problems later in life, including cardiovascular and metabolic conditions. Together, these observations provided some of the earliest human evidence that the environment during a critical period of development can influence the epigenome and potentially shape health much later in life. This is why the Dutch famine studies became foundational to environmental epigenetics: they connected a clearly documented environmental exposure, a persistent change in DNA methylation, and later-life health outcomes without requiring a change in the underlying DNA sequence.

Research Gap

Although environmental exposures have been linked to epigenetic changes, it is still unclear whether these changes can truly be inherited across generations in humans. One major problem is separating genuine epigenetic inheritance from genetic, environmental, and cultural inheritance. For an epigenetic mark to be considered truly transgenerational, it would need to survive in germ cells and affect later generations even when they are no longer directly exposed to the original environmental factor. This is difficult to establish in humans because controlled experiments across several generations are not possible, and shared family environments can create similar patterns without the epigenetic mark actually being passed through the germline.

There are also major gaps in understanding the mechanism itself. During germ-cell development and early embryonic development, epigenetic information is extensively reprogrammed, meaning many existing marks are erased before the next generation develops. Identifying what, if anything, escapes this reprogramming is particularly difficult in female germ cells because they are scarce and difficult to study, while germ-cell samples can also be contaminated by surrounding somatic cells. Another possibility is that long-lived RNA molecules, especially those present in sperm, may carry epigenetic information across generations, but how these RNAs could preserve and transmit environmental information is still not well understood. Demonstrating that a specific germ-cell factor actually causes a phenotype in the next generation—and experimentally removing or adding that factor—is also extremely difficult in humans. These gaps are why current evidence suggests that transgenerational epigenetic inheritance is possible, but has not yet been conclusively demonstrated in humans.

Research Question / Objective

This paper aims to examine how environmental stresses such as drought, heat, cold, and salinity influence the epigenetic regulation of stress-response genes in plants. Rather than looking at epigenetics as a whole, the focus is on how plants use changes in DNA methylation, histone modifications, and small non-coding RNAs to alter gene activity when exposed to stressful conditions. These mechanisms can change chromatin accessibility and the expression of genes involved in processes such as stress signalling, growth, and tolerance, without changing the underlying DNA sequence.

The paper will therefore investigate **how environmental stress triggers these epigenetic changes and how they contribute to plant stress responses and memory**. Particular attention will be given to DNA methylation, stress-associated histone marks such as H3K4me3, and small RNAs involved in pathways such as RNA-directed DNA methylation. For example, heat stress can be associated with changes in histone modifications at stress-responsive regions, while drought can alter DNA methylation and small-RNA-mediated regulation. By focusing on these mechanisms and their effects on stress-response genes, the paper aims to understand how plants translate environmental signals into changes in gene expression and adaptation.

II. Methodology

Participants / Data Source

Environmental epigenetics draws on several types of biological data, with human studies often using epigenome-wide association studies (EWAS). An EWAS compares epigenetic patterns across individuals with a particular exposure or phenotype, usually by measuring DNA methylation at thousands of genomic sites. Human studies can use case-control groups, longitudinal birth cohorts, family or twin cohorts, or even compare the same individuals before and after an exposure. Blood is one of the most commonly studied tissues, although the choice of tissue is important because epigenetic patterns can differ between cell types. EWAS can identify associations between an environmental exposure and changes in DNA methylation, but these associations do not by themselves prove that the epigenetic change caused the observed trait.

Human cohorts are valuable because they capture environmental exposures and biological responses in real-world populations, including factors that cannot easily be recreated in a laboratory. However, they also contain genetic differences, differences in cell-type proportions, and many environmental and lifestyle factors that can complicate interpretation. Model organisms and plants provide the opposite advantage: researchers can control factors such as temperature, drought, nutrition, and exposure duration while collecting specific tissues at defined stages. Plant experiments, including those using species such as *Arabidopsis* and crops, therefore make it easier to directly test how a particular stress changes DNA methylation, histone modifications, or other molecular regulators. Together, human cohorts and controlled experimental systems provide complementary evidence, with human studies showing whether patterns occur in real populations and model systems helping investigate the mechanisms behind them.

Design / Approach

EWAS can be carried out using several study designs, each with different strengths for interpreting environmental and disease-related epigenetic changes. A cross-sectional case-control design compares individuals who have a particular disease or exposure with those who do not, but it is often considered a suboptimal choice

because samples are usually collected at only one point in time. This makes it difficult to determine whether an observed epigenetic difference contributed to the disease or instead developed as a consequence of it. Monozygotic twin studies provide another approach, particularly when identical twins differ in a disease or trait. Because the twins share almost all of their DNA, differences in their epigenetic patterns can help researchers investigate the contribution of environmental or other non-genetic factors.

Longitudinal designs are generally more informative because the same individuals can be studied at different time points. For example, samples can be collected before and after disease onset or before and after an environmental exposure, allowing researchers to observe whether an epigenetic change occurs before or after the outcome. The authors therefore highlight the value of comparing cells or individuals before and after exposure, since this can provide stronger evidence about the direction of an association than a single cross-sectional comparison. Even then, longitudinal studies do not automatically prove causation, but they reduce some of the uncertainty that comes from measuring exposure, epigenetic state, and disease at only one time point.

Data Collection

Different laboratory techniques are used to measure different types of epigenetic marks. For DNA methylation, bisulfite sequencing is one of the main approaches. In this method, DNA is treated with sodium bisulfite, which converts unmethylated cytosines into uracil, while methylated cytosines remain largely unchanged. After sequencing, the pattern of converted and unconverted cytosines can therefore be used to determine which sites were methylated. Whole-genome bisulfite sequencing (WGBS) can examine methylation across the genome at single-base resolution, while reduced representation bisulfite sequencing (RRBS) focuses on a selected, CpG-rich portion of the genome.

For histone modifications, chromatin immunoprecipitation (ChIP) uses antibodies that specifically recognise a particular histone mark, such as H3K27me3. DNA regions associated with that modified histone are pulled down along with the chromatin, after which the DNA is purified and sequenced. In ChIP-seq, mapping these sequences back to the genome shows where the particular histone modification is concentrated and allows researchers to identify genomic regions associated with active or repressed gene regulation. For non-coding RNAs, RNA sequencing can identify and quantify RNA molecules produced in a sample, while small-RNA sequencing is particularly useful for profiling shorter regulatory RNAs such as microRNAs and other small RNAs. Together, these methods allow researchers to examine DNA methylation, histone modifications, and RNA-based regulation as different layers of the epigenetic response.

Data Analysis

After methylation data are collected, statistical tests are used to determine whether the observed differences between groups, such as stressed and control plants, are likely to represent genuine differential methylation rather than random variation. Common approaches include the Wilcoxon rank-sum test, which is a non-parametric alternative to the t-test; the t-test, which compares mean methylation levels between two groups; and permutation tests, which repeatedly rearrange the group labels to estimate how likely an observed difference is under the null hypothesis. More complex approaches include the empirical Bayes method, which uses a linear model with moderated variance estimates, and bump hunting, which considers patterns across nearby CpG sites to identify genomic regions showing coordinated methylation differences. Because thousands of loci may be tested simultaneously, p-values are commonly adjusted to control the false discovery rate (FDR).

Methylation data can be represented using either β -values or M-values, which describe the same underlying methylation measurements in different ways. β -values range from 0 to 1 and are easier to interpret because they roughly represent the proportion of methylated DNA at a CpG site. M-values are a log-ratio transformation of methylated versus unmethylated signal and can be more suitable for statistical modelling because their scale is less restricted. Thus, β -values are generally easier to understand biologically, while M-values are often preferred for statistical analysis. The choice of statistical method and methylation scale can affect power and false-discovery control, so the analysis needs to match the study design and dataset.

III. Results

Primary Findings

Research on environmental epigenomics shows that exposures such as diet, metabolic stress, hypoxia, toxins, and disease can influence molecular regulators of gene expression, including DNA methylation, histone modifications, and chromatin states. These changes can alter which genes are more or less active in response to an environmental stimulus. The response is not always a direct effect on the epigenome, however. Environmental signals can first activate cell-signalling pathways and transcription factors, which then target particular regions of the genome and produce changes in chromatin or gene expression. Genetic variation can also influence how an individual responds. For instance, meQTLs are genetic variants associated with differences in DNA methylation, while response eQTLs can appear only after cells experience a particular environmental challenge.

Another major finding is that changes in cell-type composition can look like genuine epigenetic changes when analysing a whole tissue. If an environmental exposure or disease increases the proportion of one cell type while decreasing another, the overall DNA methylation or chromatin profile of the tissue will also change, even if the epigenetic state within each individual cell type has not changed. NAFLD provides a useful example because disease progression can change the proportions of inflammatory cells, hepatocytes, and other liver cell subtypes. Therefore, an EWAS signal may reflect *which cells are present* rather than reprogramming within those cells. This is why cell composition, genetic variants, and signalling pathways need to be considered when interpreting environmental epigenetic findings rather than automatically treating every methylation difference as a direct epigenetic effect.

Group / Category Comparisons

DNA methylation, histone modification, and non-coding RNAs do not act as completely separate mechanisms. Instead, they can interact to coordinate gene expression. Long non-coding RNAs (lncRNAs) are particularly important because they can act as molecular scaffolds, bringing regulatory proteins and epigenetic enzymes to particular genomic regions. Some lncRNAs interact directly with DNA methyltransferases (DNMTs) and help influence where DNA methylation is established or maintained. In this way, an RNA molecule can help connect transcriptional regulation with DNA methylation rather than acting only as a regulator of RNA itself.

This interaction can also extend to histone modifications and chromatin structure. For example, DNA methylation is closely associated with histone deacetylation, and complexes containing DNMTs can interact with histone-modifying enzymes such as HDAC1. Similarly, the Polycomb Repressive Complex 2 (PRC2) component EZH2 can interact with DNMTs and contribute to their recruitment to particular genomic regions, linking histone-based repression with DNA methylation. These examples show why the major epigenetic mechanisms are better understood as an interconnected system: an environmental signal or regulatory RNA can influence one layer, which can then recruit or alter another layer to produce a coordinated change in gene expression.

Secondary Comparisons

Cold and heat stress can both produce epigenetic changes in plants, but the mechanisms and genes involved are not identical. Under cold stress, one of the clearest examples is the regulation of the *FLC* (FLOWERING LOCUS C) gene during vernalization. Cold exposure promotes changes in chromatin at *FLC*, particularly involving the repressive histone mark H3K27me₃, which helps maintain *FLC* silencing and contributes to cold memory. Cold stress can also alter H3K4me₃ and histone acetylation at cold-responsive genes such as *COR15A* and *ATGOLS3*, affecting their accessibility and expression. DNA methylation and histone deacetylation also contribute to the regulation of other cold-response pathways.

Heat stress, in contrast, is strongly associated with activation of heat-shock pathways and transcriptional memory. The HSFA2 heat-shock transcription factor is an important regulator, while histone modifications such as increased H3K4me₃ have been observed at heat-memory genes including *APX2* and *HSP22*. These changes help maintain a transcriptionally responsive state after the initial heat exposure. Thus, although both stresses can involve DNA methylation and histone modifications, cold stress is strongly linked with *FLC* repression and vernalization pathways, whereas heat stress involves HSFA2 and heat-shock-protein pathways. This suggests that plant epigenetic responses are stress-specific rather than one generic epigenetic response to environmental stress.

Key / Unexpected Findings

One of the more unexpected findings in research on intergenerational inheritance is that sperm RNA may carry information about a father's environmental experience, and that this information can depend on chemical modifications of the RNA itself. In studies involving males fed a high-fat diet, sperm contained altered populations of tRNA-derived fragments, including changes in RNA modifications such as 5-methylcytosine. When these modified RNA fragments were introduced into early embryos, they were able to reproduce aspects of the metabolic phenotype associated with the paternal diet, including impaired glucose regulation in offspring.

What makes this particularly interesting is that injecting a synthetic, chemically unmodified version of the same RNA sequence did not produce the same effect. This suggests that the sequence of the RNA alone may not contain all of the information needed for transmission; its chemical modifications can also influence its biological activity. This changes the way inheritance can be viewed: information passed from one generation to the next may involve not only DNA sequence, but also molecular states carried by gametes. However, these findings should not be taken as proof that environmental experiences are routinely inherited in humans, since much of the mechanistic evidence comes from controlled animal experiments.

IV. Discussion and Conclusions

Interpretation

The findings discussed throughout this paper suggest that epigenetic mechanisms do not create a completely separate hereditary code alongside DNA. Instead, they add regulatory layers that determine when, where, and to what extent the information already encoded in DNA is used. DNA methylation, histone modifications, chromatin structure, and non-coding RNAs can make particular regions more or less accessible and can influence transcription, while still leaving the underlying DNA sequence unchanged. In this sense, they act more like a regulatory filter on the classical DNA → RNA → protein pathway rather than replacing it. Chromatin topology can determine which regulatory regions come into contact with genes, while non-coding RNAs can recruit chromatin-modifying complexes or influence RNA stability and translation.

This also suggests that the traditional idea of a gene as simply a stretch of DNA that produces a particular protein is incomplete. Gene function depends not only on its sequence but also on the regulatory environment surrounding it, including chromatin state, transcriptional regulators, and non-coding RNAs. However, this does not mean that every regulatory state should be treated as an independent hereditary information system. The DNA sequence remains the primary template for the canonical flow of genetic information, while epigenetic mechanisms influence how that information is interpreted and expressed. Epigenetics therefore extends the classical understanding of heredity by showing that inheritance and gene function involve not only the sequence itself, but also the molecular context in which that sequence operates.

Comparison with Previous Studies

The classical central dogma, associated with Francis Crick, describes the directional flow of sequence information from DNA to RNA and from RNA to protein. DNA therefore acts as the main hereditary substrate, while RNA serves primarily as an intermediate for expressing that information. Modern epigenetics does not necessarily overturn this framework. Instead, it shows that the flow of genetic information is controlled by additional molecular layers that determine which parts of the DNA are accessible and how strongly they are expressed.

DNA methylation can silence or alter the activity of particular genes, while histone modifications and chromatin structure can make DNA more or less accessible to the transcription machinery. Non-coding RNAs add another regulatory layer by interacting with DNA, RNA, or protein complexes and influencing transcription and chromatin states. These mechanisms therefore reshape the use of genetic information rather than replacing the DNA-to-RNA-to-protein pathway. At the same time, findings involving cytoplasmic inheritance and unusual DNA structures suggest that heredity may involve more molecular components than the classical model originally emphasised. The modern view is consequently less about abandoning the central dogma and more about recognising that the expression and inheritance of biological information occur within a much more complex molecular context.

Practical Implications

One practical application of plant epigenetics is the use of stress memory in crop breeding. When a parent plant experiences drought, heat, or salt stress, some of the resulting epigenetic changes can persist and influence how its offspring respond to a similar stress, even though their DNA sequence has not been changed. Repeated or multi-generational stress exposure can therefore be investigated as a way of identifying plants that maintain stronger stress responses in later generations. This could be particularly useful for crops that are difficult to improve through conventional breeding alone, while clonal crops may also retain some epigenetic states through repeated cell division. However, because epigenetic marks can be unstable or disappear across generations, stress priming would need to be tested carefully rather than assumed to produce permanently tolerant offspring.

Another emerging approach is epigenome editing using CRISPR-based systems. Instead of using CRISPR to alter the DNA sequence itself, catalytically inactive Cas proteins such as dCas9 can be directed to a specific genomic region and linked to an epigenetic modifier. This could allow researchers to increase or decrease DNA methylation or alter histone-associated regulation at a stress-response gene without changing its underlying sequence. In principle, such targeted epigenome editing could help breeders fine-tune genes involved in drought, heat, or salt tolerance and test whether particular epigenetic states improve stress responses. This makes epigenetic breeding a potential complement to conventional genetic breeding, although the stability and inheritance of these engineered epigenetic states remain important challenges.

V. Limitations & Future Research

One of the biggest challenges in environmental epigenetics is establishing causation. Reverse causation occurs when the outcome or disease causes the observed epigenetic change, rather than the epigenetic change contributing to the outcome. This is particularly difficult in cross-sectional studies because methylation and the

outcome are often measured at the same time, making their temporal order unclear. Other factors, such as smoking, treatment, genetics, or environmental exposures, can also confound the relationship. Replication across different populations and longitudinal studies can help, but an association alone still cannot establish a causal pathway.

A second problem is tissue accessibility. Researchers cannot directly measure DNA methylation in the brain of a living person, so studies of brain-related conditions often use accessible tissues such as blood or saliva. However, methylation patterns can be tissue-specific, meaning a change detected in blood may not accurately represent what is happening in the brain. Post-mortem brain tissue provides more direct evidence but can only be studied after disease has already developed, creating its own limitations. Future research can strengthen causal evidence by combining larger longitudinal cohorts, better tissue-specific datasets, experimental or quasi-experimental approaches, and Mendelian randomization using genetic variants associated with methylation (mQTLs). The authors also emphasise the need for replication across diverse populations, since many existing mQTL resources are based mainly on people of European ancestry. Using several complementary approaches together should make it easier to distinguish genuine causal mechanisms from biomarkers and correlations.

References

- [1]. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3257095/>
- [2]. <https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2026.1804465/full>
- [3]. <https://www.nature.com/articles/s41467-018-05445-5>
- [4]. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2026.1738299/full>
- [5]. <https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1010567>
- [6]. <https://www.mdpi.com/2076-3417/15/18/9940>
- [7]. <https://link.springer.com/article/10.1186/s12859-015-0641-x>
- [8]. <https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2022.1067406/full>
- [9]. <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.1075279/full>
- [10]. <https://www.mdpi.com/1422-0067/24/6/5889>
- [11]. <https://www.mdpi.com/2075-1729/16/1/79>
- [12]. <https://www.mdpi.com/2223-7747/10/6/1226>
- [13]. <https://www.mdpi.com/2073-4425/10/3/193>