

Concept Paper

Not peer-reviewed version

---

# Past Epigenetic Information Archive with Blockchain Data Architecture in a Biological Cellular Substructure

---

[Jung In Choi](#) \*

Posted Date: 25 December 2023

doi: [10.20944/preprints202312.1424.v2](https://doi.org/10.20944/preprints202312.1424.v2)

Keywords: epigenetic information; epigenetic modification; blockchain; cellular network; data archive; evolution; data immutability; distributed ledger; p2p communication; cellular substructure



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Concept Paper

# Past Epigenetic Information Archive with Blockchain Data Architecture in a Biological Cellular Substructure

Jung In Choi

AICT, Seoul National University, jichoi@snu.ac.kr

**Abstract:** Past epigenetic information refers to the epigenetic modifications that occurred in the past that have since been updated. These modifications may no longer be present and not actively influencing cellular functions. When past epigenetic modifications disappear, the precise records of those modifications typically vanish. If this is indeed true, a critical question arises: Can an organism sustain continuous survival and evolution without the information about past epigenetic modifications up to the present? It is reasonable to assume that among the information about past modifications, there could be crucial data essential for survival and evolution. If all these experiences were to vanish, rendering them unknowable in the present, it would undoubtedly pose significant challenges to survival and evolution. Therefore, the argument that such epigenetic in-formation has been securely stored somewhere within the cell is not merely speculative but likely a factual reality. In this study, a new term, "Epigenetic Archive", is introduced in biology. This refers to a data repository that encompasses all experience data on past epigenetic modifications. The argument presented here underscores the importance of the Epigenetic Archive and the need for its secure preservation. According to the argument, the information that cells experience should be continuously and securely stored, and any distortion of this information is said to be inconsistent with the principles of evolution and survival. This information is seen as evidence of a cell's successful existence up to the present, and it is emphasized that this successful record must be preserved. From this perspective, the Epigenetic Archive should be designed to safely store past biological experiences and link this information securely to the present data. This study proposes that the data architecture in the Epigenetic Archive is indeed akin to the principles of blockchain technology, particularly in terms of data immutability and a distributed ledger that securely preserves information across generations, ensuring that all cells possess consistent epi-genetic information across a biological organism. The applications and implications of such an analogy of blockchain to epigenetics could be an original contribution to future research on biology and genetics.

**Keywords:** epigenetic information; epigenetic modifications; epigenetic archive; blockchain; data immutability; distributed ledger; p2p communication; cellular substructure

---

## I. Introduction

The information stored within a cell is composed of genetic information and epigenetic information. Genetic information is in the form of sequential DNA sequences, while epigenetic information is contextual information that modulates the expression of genes.

Genetic information is the blueprint for the cell's structure and function, and it is responsible for the inheritance of traits from one generation to the next. This information is encoded in the DNA sequence, which is made up of four nucleotides (adenine, thymine, guanine, and cytosine) arranged in a specific order. The sequence of these nucleotides determines the sequence of amino acids in proteins, which in turn, determine the structure and function of the cell [1-2].

Epigenetic information, on the other hand, refers to modifications that occur in the DNA or the proteins of that package DNA (histones). It does not change the underlying DNA sequence but can influence gene expression. These modifications can be heritable and can be passed down from cell to cell during cell division, and from one generation to the next. Epigenetic modifications can be influenced by environmental factors, such as diet, stress, and exposure to toxins [3-6].

It is reasonable to think of epigenetic information as a contextual layer that is added onto the sequential layer of genetic information, and which continuously changes over time to modulate cellular function. Epigenetic modifications can affect how genes are expressed and can be influenced by a variety of factors, such as environmental cues and developmental signals, leading to a complex interplay between genetic and epigenetic information.

Epigenetic memory refers to the ability of cells to maintain past epigenetic information over time, even after cell division or changes in environmental conditions. This memory is thought to be important for past cellular experiences and can be passed down through generations of cells [7–10].

When epigenetic modifications accumulate in cells, they combine with genetic information in the nucleus to regulate complex cellular activities. Epigenetic modifications, such as changes in DNA methylation or histone modification, can alter the accessibility of DNA and influence gene expression patterns, ultimately affecting cellular behavior.

The current state of epigenetic modifications can be thought of as an additional information layer stored on top of the genetic information in the form of context-specific information. This information layer affects gene expression and regulates various cellular activities. While current epigenetic information can be stored and inherited, it is not known if past epigenetic information remains in a cell or how it can be retrieved in the form of memory in a cell.

Further research is needed to understand the mechanisms involved in the storage and recall of epigenetic memory with past epigenetic information [11–13].

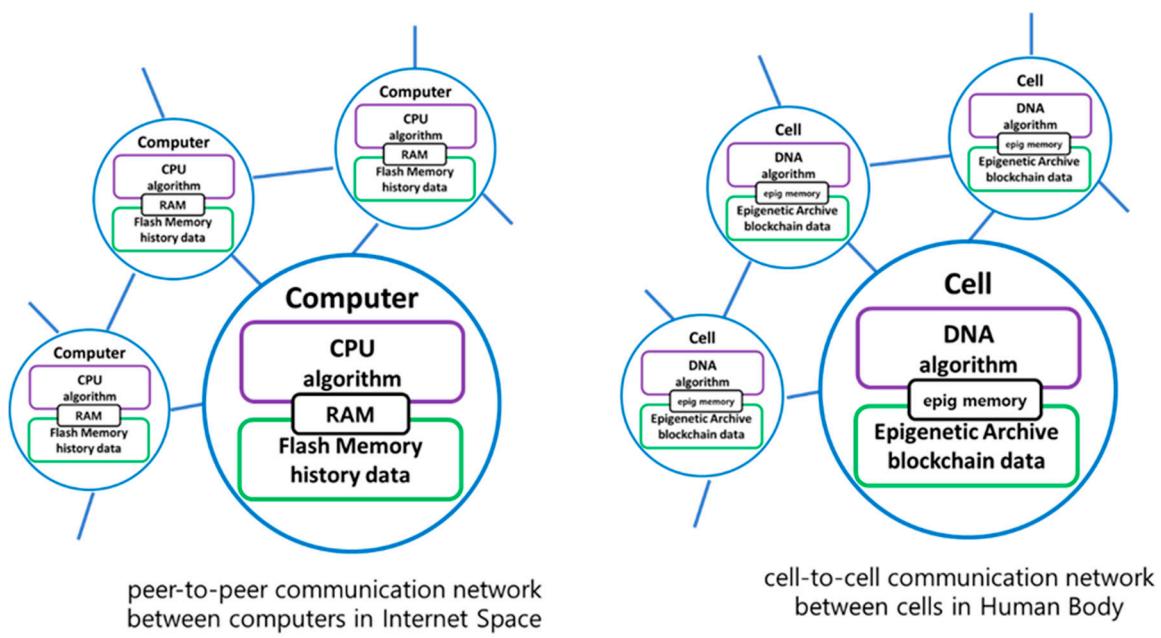
The concept of the Epigenetic Archive is distinct from the epigenetic memory. While epigenetic memory refers to the biological material that remains in the nucleus of the cell, the Epigenetic Archive refers to the biological information encoded somewhere in the past. The Epigenetic Archive specifically denotes a data repository where information about all past epigenetic modifications, even those that may have changed or disappeared by now, is recorded [14–17].

This concept bears similarities to the characteristics of blockchain, emphasizing data immutability and a distributed ledger to maintain a continuous record of epigenetic modifications over time. With reference to another term, 'cellular blockchain', this study seeks to illustrate the biological mechanism of how the 'Epigenetic Archive' is formed and preserved in a cell population.

## II. Cellular Blockchain Concept

Blockchain is a decentralized database technology used to securely record and manage data across a peer-to-peer network of computers. It ensures transparency, security, and immutability of recorded information [18–23].

The cellular blockchain concept is inspired by the analogy of digital computers and biological living cells. Figure 1 shows the analogy of the peer-to-peer network among computers in the internet space and the cell-to-cell network among cells in the human body. It is assumed that the algorithm in the CPU of the digital computer corresponds to the genetic algorithm in the DNA of the cell, and history data in the memory of the computer corresponds to the epigenetic data in the cellular blockchain of the cell.



**Figure 1.** Analogy of decentralized communication networks in Internet Space and in Human Body.

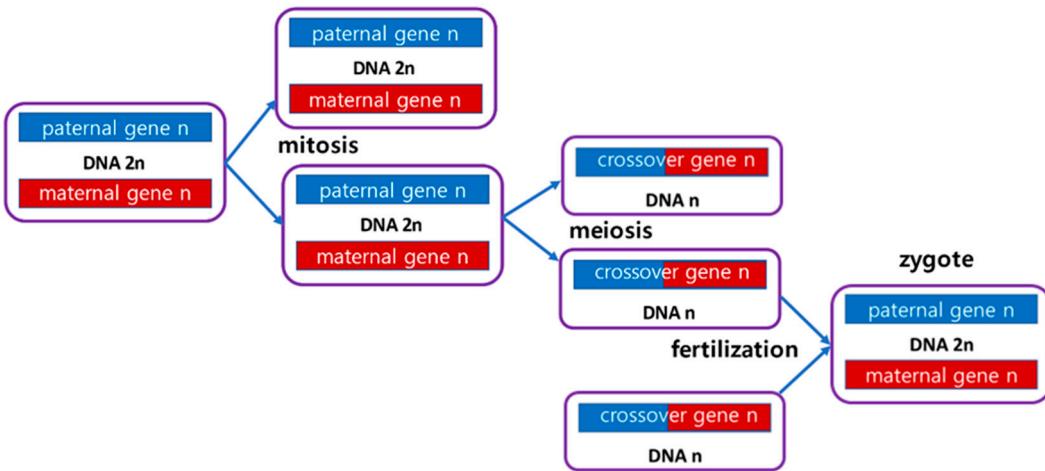
The current epigenetic modifications of DNA can be likened to the role of random access memory (RAM) in a computer, as both serve to store immediately needed information for the cell or process. Being volatile, both can undergo rapid changes. On the other hand, the Epigenetic Archive, responsible for storing past epi-genetic data, can be compared to flash memory in a computer. Both act as non-volatile storage, capable of preserving data for extended periods. Therefore, in internet space, data created through a peer-to-peer network are distributed and stored in a block-chain manner. Similarly, in the human body, epigenetic data created through a cell-to-cell network among cells are distributed and stored in the Epigenetic Archive using a blockchain manner.

The Epigenetic Archive is a blockchain-like database of the cell that connects epi-genetic information from the original cells to the present by making it a long chain of data blocks. The current epigenetic information is implemented through the accumulated epigenetic modification and is attached to the nuclear DNA with genetic information in the form of methylation. This is implemented as a biochemical material, affects cellular activity, and can be passed on to the next generation. At the same time, during modification, data are stored in the Epigenetic Archive as a blockchain, that is, as a Current Block, using biological data logic. Afterwards, when a modification is updated, a new data block is created, and the old data block becomes the Previous Block. As blocks are added in this way, a long cellular blockchain is created. This blockchain becomes an Epigenetic Archive. Here, only the data in the Current Block remain in the DNA in the form of modification and affects cellular activities such as gene expression.

However, the Epigenetic Archive must include all data related to the modification of other cells in the cell population. In other words, the blockchain collects all modifications, including those of the surrounding cells, that occurred within the same time period, creates one block, and shares it with a common ledger. Therefore, all cells have the same blockchain, but they attach their own to DNA, causing an epigenetic effect.

### III. Epigenetic Archive Inheritance

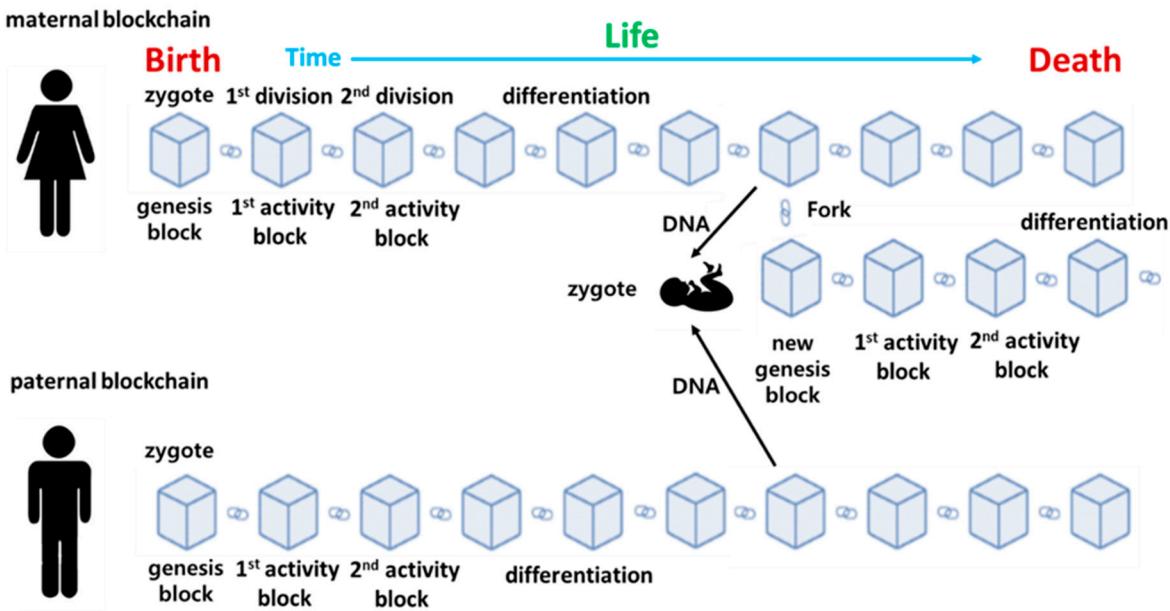
Human life begins with a fertilized egg, or zygote, and ends with death. As shown in Figure 2, the zygote undergoes division, differentiation, and growth through mitosis, and finally becomes an adult composed of 40 trillion cells. Here, the sperm and egg, produced through meiosis of the germ cells, combine to form a new fertilized egg, and life continues.



**Figure 2.** Human cell's genetic DNA process with mitosis and meiosis, and fertilization for zygote.

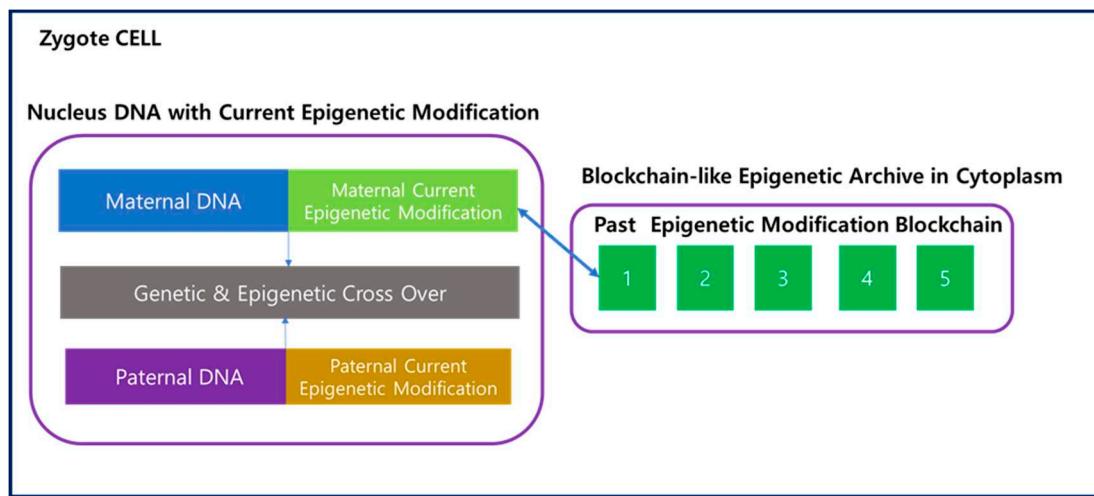
In Figure 3, assume that human life from birth to death is recorded as a cellular blockchain in an Epigenetic Archive. The starting point is the genesis block, which records the cellular activity of a single cell fertilized egg. It divides into two identical cells through the first mitosis. The description of this process is linked to the genesis block as the 1st activity block. As cell activity continues through further cell division and differentiation, corresponding data blocks are sequentially created and connected, and the blockchain ends at the moment when life activity ends.

In other words, the connected blockchain records of life activities are stored equally in all cells. Here, through reproduction, this process can be passed on to the offspring. At this time, the new fertilized egg receives half and half of the DNA from the mother and father and starts life activities again from the newly fertilized egg. From this point on, the child cell division, differentiation, and growth process repeats the process of the parent cell. In other words, the genesis block for the fertilized egg of the parent cell starts through a fork in the parent cell blockchain, and it is connected to the parent cell before the fork and continues to extend the blockchain. Therefore, if you trace it back, the blockchain of the child cell is connected to the genesis block of the parent cell, which is again connected to the blockchain of the grandparent cell, which is ultimately connected to the genesis block of the first human ancestor.



**Figure 3.** Human Life with Cellular Blockchain in Epigenetic Archive.

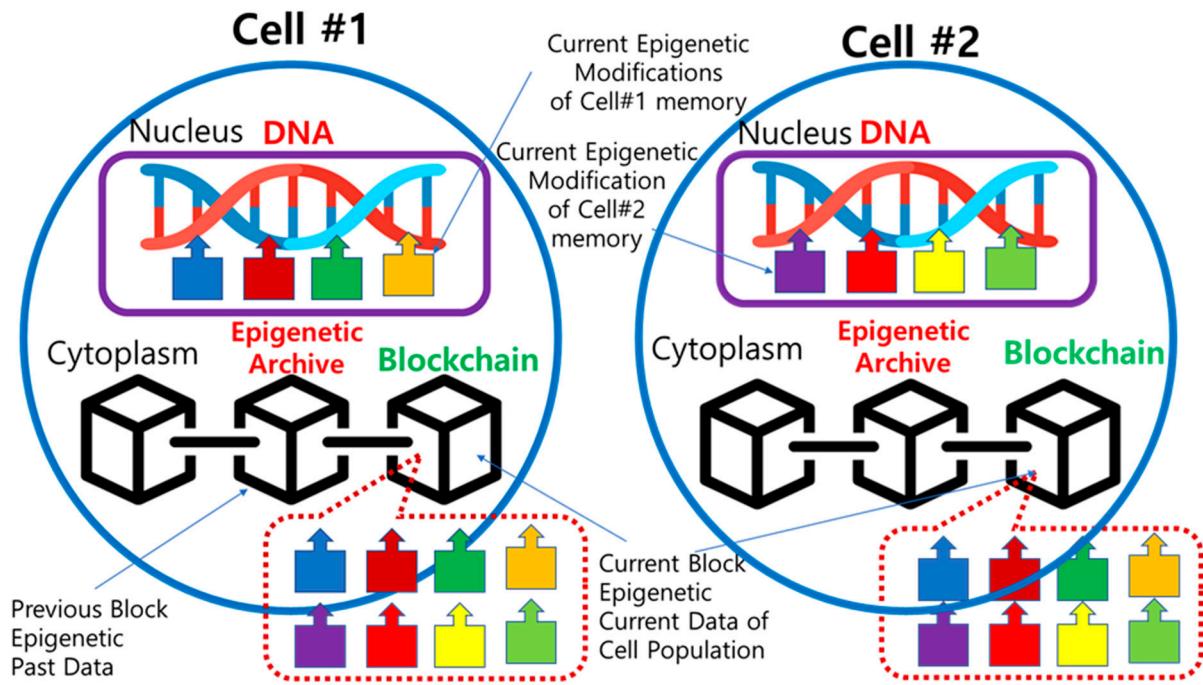
As shown in Figure 4, when the egg and sperm merge to form a zygote, not only nucleus genetic DNA but active epigenetic modifications undergo crossover, with data from both the maternal and paternal lineages. On the other hand, information about past epigenetic modifications is stored in the Epigenetic Archive within the cytoplasm in a blockchain-like fashion. The merging of these two blockchains in the zygote is practically challenging, and it is reasonable to consider a fork scenario, where one of the blockchains diverges to start a new one. Consequently, the transmission of information to the next generation through the fork occurs via the maternal lineage, while the paternal lineage loses its blockchain. Therefore, in the zygote, the maternal epigenetic blockchain is replicated, and new blocks are added, extending the chain. This process takes place in the maternal uterus, and the existing maternal blockchain continues to be connected even after the fork.



**Figure 4.** Epigenetic Archive Inheritance with Maternal Blockchain Fork .

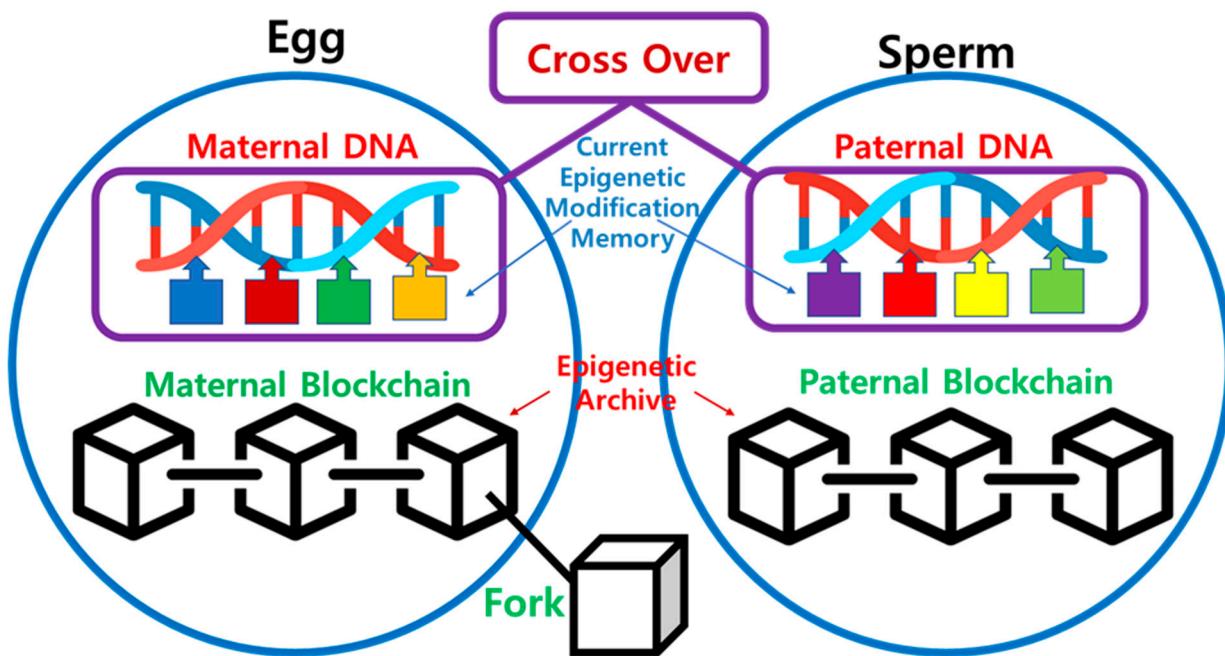
#### IV. Results

Two illustrative demonstrations are presented for the new term “Epigenetic Archive” based on the cellular blockchain concept.



**Figure 5.** Blockchain based Epigenetic Archive in a Cell Population.

In Figure 5, there is a cell population consisting of Cell #1 and Cell #2. In each cell, four different epigenetic modifications occurred in the nucleus during the current time period. They are attached to the DNA in the form of biological memory and affect cell activity and heredity. At the same time, all modifications are copied as epigenetic in-formation into the Epigenetic Archive in the cytoplasm. Throughout the whole population, the data of the eight epigenetic modifications are collected in the common ledger and stored as the Current Block of the blockchain in the Epigenetic Archive of each cell. When a new block is created, the Current Block moves to the Previous Blocks, continuing to extend the blockchain. Therefore, both Cells #1 and #2 have a common blockchain, while each cell has its own current epigenetic modifications attached to the DNA.



**Figure 6.** Epigenetic Archive Inheritance Model.

In Figure 6, during the formation of the zygote, the merging of active epigenetic modifications occurs alongside the crossover of the nucleus DNA, blending the maternal and paternal. On the other hand, information about past epigenetic modifications is stored in the Epigenetic Archive within the cytoplasm in a blockchain-like manner. It involves a fork scenario, with the maternal lineage serving as the continuous epigenetic chain. The maternal epigenetic blockchain is preserved and expanded, while the paternal lineage loses its blockchain. This scenario mirrors the mitochondrial inheritance pattern.

## V. Conclusion

This study offers a novel perspective on the question of “life”, defining it as the harmonious interplay between immutable genetic algorithms and continuously updating epigenetic historical records within the activities of cells. Unlike the static genetic algorithms, the intriguing focus here is on the less understood Epigenetic Archive. Drawing inspiration from the similarity between the data characteristics of the Epigenetic Archive and the blockchain concept, this study endeavors to elucidate the underlying mechanisms.

The conclusion drawn from this study is that the newly introduced term “Epigenetic Archive” serves as a record of epigenetic modifications, starting with the initial cell and preserving data securely up to the present moment, encompassing not only individual cells but the entire cell population. The data characteristics of the Epigenetic Archive, akin to a vast, spatiotemporally updated database, bear a striking resemblance to the principles of blockchain technology, an insight crucially derived from the study.

Furthermore, the study postulates an argument that the Epigenetic Archive, un-like genetic algorithms, is conveniently preserved within the cell’s cytoplasm, potentially influencing the heredity process. By presenting this innovative conceptual model, this research opens the door to further groundbreaking studies, suggesting a promising avenue for future research in biology and genetics.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not Applicable

**Informed Consent Statement:** Not Applicable

**Data Availability Statement:** Not Applicable

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C. Initial sequencing and analysis of the human genome. *Nature* 2001, 409, 860–921.
2. Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Kalush, F.; et al. The sequence of the human genome. *Science* 2001, 291, 1304–1351.
3. Andrew, P.F.; Jirtle, R.L. Epigenetics: A means to explore the interface between genetics and the environment. *Ann. N. Y. Acad. Sci.* 2003, 983, 154–165.
4. Howard, C.; Bergman, Y. Linking DNA methylation and histone modification: Patterns and paradigms. *Nat. Rev. Genet.* 2009, 10, 295–304.
5. Roadmap Epigenomics Consortium; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015, 518, 317–330.
6. Robin, H. Epigenetics: An overview. *Dev. Genet.* 2002, 31, 173–177.
7. D’Urso, A.; Bricknerrends, J.H. Mechanisms of epigenetic memory. *Trends Genet.* 2014, 30, 230–236.
8. Allis, C.D.; Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 2016, 17, 487–500.
9. Zentner, G.E.; Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* 2013, 20, 259–266.

10. Jablonka, E.; Lamb, M.J. The inheritance of acquired epigenetic variations. *Int. J. Epidemiol.* 2012, **41**, 10–13.
11. Yidong, J.; Jirtle, R. Environmental epigenomics in human health and disease. *Environ. Mol. Mutagen.* 2008, **49**, 4–8.
12. Aaron, D.G.; et al. Epigenetic memory of transcriptional gene silencing and reprogramming in plants. *Nature* 2007, **447**, 88–91.
13. Mayumi Iwasaki; Jerzy Paszkowski; Epigenetic memory in plants. Published online 2014 Aug 7. doi: 10.15252/embj.201488883
14. Alexander, P.F.; Waxman, D.J. Establishment of epigenetic memory in the hematopoietic lineage during development. *Epigenetics Chromatin* 2015, **8**, 1–12.
15. Goldberg, A.D.; Banaszynski, L.A.; Noh, K.M.; Lewis, P.W.; Elsaesser, S.J.; Stadler, S.; Dewell, S.; Law, M.; Guo, X.; Li, X.; et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 2010, **140**, 678–691.
16. Hiroshi, K. Histone modifications for human epigenome analysis. *J. Hum. Genet.* 2013, **58**, 439–445.
17. Bob, W. Epigenetics: The science of change. *Environ. Health Perspect.* 2006, **114**, A160–A167.
18. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System; 2008.
19. Antonopoulos, A.M. Mastering Bitcoin: Unlocking Digital Cryptocurrencies; O'Reilly Media: Sonoma County, CA, USA, 2014.
20. Drescher, D. Blockchain Basics: A Non-Technical Introduction in 25 Steps; Publisher Apress: <https://www.apress.com/gp/2017>.
21. Antonopoulos, A.M. The Internet of Money; 2016.
22. Tapscott, D.; Tapscott, A. Blockchain Revolution: How the Technology Behind Bitcoin and Other Cryptocurrencies is Changing the World; Portfolio: Greenwich, CT, USA, 2016.
23. Narayanan, A.; Bonneau, J.; Felten, E.; Miller, A. Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction; Princeton University Press: Princeton, NJ, USA, 2016.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.