Available Online: 20 April 2026 DOI: 10.54254/3049-5458/2026.32854

The value of understanding RNA virus origins: but which theories hold the most merit?

Mengyi Zhang

Suzhou High School-SIP, Suzhou, China

hannah.chen0853@gmail.com

Abstract. RNA viruses are very dangerous to global health. Because of their tendency to either undergo high rates of mutation or quickly evolve as the case of influenza, COVID-19, and Ebola depicts. Their etiology is not only the solution to a core mystery in the life sciences, but it also applies to prevention of the disease by viruses and to drug development. This paper compares three classical origin theories, including the Escape Hypothesis (fragments escape from cellular genetic material), the Regressive Hypothesis (degeneration of complex parasitic cells) and the Coevolution Hypothesis. By analyzing empirical data, logical and limitation of both theories, the paper shows the pluralistic nature of the origin of viruses: retroviruses have the Escape Hypothesis, giant viruses have the Regressive Hypothesis, and the majority of RNA viruses have the Coevolution Hypothesis. The Coevolution Hypothesis, which is backed by a set of Endogenous Viral Elements (EVEs) in the genomes of its hosts and a strong level of conservation of replicase genes in lineages of ancient RNA viruses, is the most clear umbrella explanatory force on the overall diversity of RNA viruses. The study contributes to the existing body of knowledge on viral evolution by confirming the complementary character of the three hypotheses. It offers theory foundations to antiviral guidelines. They are the development of broad-spectrum antiviral drugs against conserved viral replicase proteins, and predicting the risks of zoonotic infections depending on lineage evolution.

Keywords: RNA viruses, viral origin, Escape Hypothesis, Regressive Hypothesis

1. Introduction

RNA virus is a type of unique pathogen which has genetic a single-stranded ribonucleic acid as its genetic material.It has a very high mutation rate, extensive host adaptability, and a rapid evolution rate. This always places it in a threat zone across the global and domestic health situation. It caused not only a huge number of infected individual, but also significant challenges to the medical system and social order.

The seasonal influenza caused by the influenza A virus.It is estimated that 291, 243–645, 832 seasonal influenza-associated respiratory deaths (4·0–8·8 per 100,000 individuals) occur annually [1]. In 1918, the major outbreak of influenza occurred in Spain, resulting in the loss of millions of lives. Its destructive power exceed the war in the same era. Since COVID-19 arose, in addition to over 600 million infection cases and almost 7 million deaths, it has also caused a global shortage of medical resources, supply chain rupture, and other chain reaction. The economy of some countries declined dramatically; While Ebola is mainly prevalent

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

in African regions, its high (50%-90%) fatality rate usually triggers social unrest. Also, there is the potential risk of spread to other regions in the coming years; HIV is a retrovirus that infects cells of the human immune system (mainly CD4-positive T-cells) and destroys or impairs their function. This virus causes progressive weakening of the immune system that makes one immunodeficient. There are over 38 million infected people who need to rely on a drug in the long term. This brought a ponderous burden to individuals' families and public sanitation [2].

The research into the cause of RNA viruses has significant implications for the world. In the eyes of a basic scientist, numerous scientists have confirmed in various aspects that viruses are a transitional form between living beings and non-living objects. The study of the origin of viruses can be a significant theoretical basis for the basic scientific hypothesis of the origin of life. Provided that viruses have their beginning in primitive macromolecules in the "RNA world", then viruses may be dismissed as a significant part of the evolutionary spectrum of life on Earth, and hence may become a valuable clue to the origin and evolution of life. The study is useful to elucidate the evolutionary history of life, beginning with the RNA molecules to multicellular organisms, and has much broader implications in the determination of the origin of life [3].

In a pragmatically oriented approach, cognisance of the known viruses will enhance the preventive and control measures of the sudden infectious disease. The software for detection and screening of unknown viruses also has the capacity to assist our country in enhancing the clinical diagnosis of infectious diseases endowed with unidentified causes, the identification of pathogens, and subsequently attain targeted therapy. Moreover, the origin of RNA viruses can show their evolutionary gaps, host conversion and co-evolution laws, explain the complicated genetic relationship and genomic diversity of viral groups, and serve as a major data source in support of the viral ecology and evolution studies [4].

2. Research review

Viruses are believed to have a birthplace, which is one of the central riddles of life sciences. The scholarly community has developed three stereotypical theories: the Escape Hypothesis, the Regressive Hypothesis and the Coevolution Hypothesis. Three standard theories have been created by the academic community: the Escape Hypothesis, the Regressive Hypothesis, and the Coevolution Hypothesis. These three theories describe how viruses came to be as known using the views of the cell fragment escape, complex cell regression, and virus-cell co-evolution respectively. The continuous utilisation of human knowledge of the nature of viral evolution due to experimental demonstrations and supplements by numerous scientists over the last decades already offers a solid base to further in-depth exploration of the frontier issue [5].

2.1. Escape hypothesis

Salvador Luria and James Darnell were the obvious ones to propose the Escape Hypothesis in 1967 in the journal Science. It is a theory that has been most empirically validated in the research of virus origin, and several molecular biological experimental studies and genomic information uphold the main reasoning. Its reasoning is in contrast to the viruses-first and cell regression line, suggesting that viruses are the result of the escape-specialisation process of genome fragments in cells, which is a gradual process of evolution and not an immediate event [6].

The hypothesis splits evolution into three definite and successive steps: the first phase is genetic fragment escape- mobile genetic elements, the cells, which as a rule are involved in normal genetic regulation, and material exchange in body can break constraints of cells to become free fragments either under the impact of external influence (e.g. extreme temperature, radiation) or through inner gene controlling mistakes. In experiments in modern molecular biology it has been unequivocally established that certain examples of transposons (including Ty transposon in yeast) can create virus-like nucleic acid protein complexes on activation, a direct modern biological validation of this step; the second step is the acquisition of infectious and replicative potentials a free fragment of genetic material acquires the ability to recognise and adhere to host cell surface receptors, and in the process develops the ability to replicate itself, the third stage is the fixation of viral features, evolved becoming a free fragment of genetic materials.

The main evidence for the Escape Hypothesis is genetic homology: genomic sequences analysis has shown the similarity of sequences between bacteriophages and bacterial plasmids is over 60% and that the most important functional genes of retroviruses are very similar to human retrotransposons. In 2012, Patrick Forterre et al. published a paper in Virus Research, further developing the hypothesis on the basis of the discovery of giant viruses and suggesting that viruses can have their origin in fragments of chromosomes of the primitive cell, not in modern elements of cell genome, which further enhanced the rationality of the origin details [7]. Nevertheless, the Escape Hypothesis also has its controversies: not only is the origin of virus-specific spike proteins hard to explain because these proteins lack clear homologous sequences in the cell genome, but molecular clock analysis results also indicate that some viral genes evolved earlier than cellular transposons, which creates a timeline contradiction which is difficult to disregard.

2.2. Regressive Hypothesis: frontier to functional streamlining through parasitism

The Regressive Hypothesis (also the Reduction Hypothesis) was first formulated in the 1950s by Howard Temin in The RNA Tumour Viruses and admits its main assertion is entirely opposite to the progressive argument of the Escape Hypothesis. It believes that viruses may arise as degenerative products of cellular organisms,losing genes during parasitic adaptation [8].

Its evolutionary history exhibits apparent features of symbiosis-dependence-regression: during the early phase of the formation of life twice types of primitive cell formed mutally advantageous symbiotic relationship, with one side offering a stable environment of living to the other side, the latter providing them with necessary metabolites and energy and the former offering them a rich array of mutations because of the long-term stasis and were lost; finally, such cell-type could no longer exist independently, and its deregulated genes were associated with metabolic activity accumulated a high level of mutations, and thus evolved into a parasitic entity that relies entirely on host cells for survival and replication [9].

Key empirical evidence of the beginnings of the Regressive Hypothesis has been in the discovery of giant viruses: in 2004, the group led by Didier Raoult published in *Science* the discovery of Mimivirus. This is a 750-nanometers-long virus with a 1.18 Mb genome. Its genomes have thousands of cell homologous genes, some of which include DNA repair and transcriptional regulation genes. Its structural and genomic complexity initially led it to be mistaken as a bacterium [10]. The evidence of the parasitic bacteria analogy also stands immensely in support of the hypothesis a 2017 paper published in *Journal of Bacteriology* demonstrated that the genome of the parasitic bacterium, Rickettsia lost over 50% of their genes in comparison to its close free-living relatives, and only preserved genes connected to its parasitic functions including ability to bind its hosts and to take in nutrients, and its evolutionary program is much more consistent with the logic of the Regressive Hypothesis [11].

In 2018, Raoult and his team found another evidence.They found a giant virus called Pandoravirus.They began thinking about how viruses might have generated as simple kinds of cells that had all the parts of a real cell. However, the Regressive Hypothesis still has unresolvable disputes: the capsid protein gene is an essential component of the virus, and its homologous sequence cannot be easily identified within the cellular genome, which goes against the notion that viruses arose from a loss of genes; it also appears that molecular clock analyses show that some groups of viruses (e.g., Coronaviridae) are older than giant viruses, which contradicts the idea that giant viruses represent an intermediate stage of regressio. Also, it is shown that during the process of viral evolution, there were many separate regressions instead of one ancestral cell degrading, which challenges the regressive hypothesis.

2.3. Coevolution Hypothesis: the figure of virus-cell coevolution

The Coevolution Hypothesis which emerged in the 1980s and continued to get refined as the metagenomic research technology advanced further. Its fundamental perspective upsets the logic of one-way origin found in the latter two hypotheses, and holds that viruses and cells have been co-evolving since the pre-cellular period and that they have collaborated and interacted in the overall pattern of life evolution on Earth [7].

It is proposed that the primitive soup at very early stages contained a high concentration of self-replicating fragments of nucleic acid (both RNA and DNA) suspended in it. These strands of nucleic acid were a primitive community of life with proteins, lipids and other biomolecules and were yet to be incorporated into identifiable cell structures. The evolutionary pathway of the emergence of the genome of primitive cells through combining other nucleic acid molecules and combining with membrane structures, the subsequent gradual differentiation of other nucleic acid molecules through mutation and selection to give simple viruses that could replicate and propagate themselves led to a close co-evolutionary association of viruses and cells: viruses contributed much of the genetic material to the evolution of cells, cells contributed an environment of stable replication and survival to viruses; viruses and cells became inseparably linked, resulting in the subsequent gradual elaboration of the cellular These layers included the assembly of membrane-bound compartments (e.g., the nucleus, which sequestered cellular genomes) and the evolution of molecular machinery for gene expression (e.g., ribosomes for protein synthesis)—traits that were shaped in part by viral genes integrated into early cell genomes [12]. As an illustration, some viral DNA fragments played a role in formation of eukaryotic transcription factors whereas viral envelope proteins could have contributed to formation of cellular transport systems across the membrane [7].

Another theoretical backing of the Coevolution Hypothesis is the RNA world theory. According to this theory, RNA was the earliest biological macromolecule that was involved in the storage of genetic information, as well as holding catalytic activity, in the pre-DNA/pre-protein early life system and was used to create a primitive life system. RNA viruses can replicate without the use of DNA as an intermediate, and the replicase genes have very high levels of conservation. The evolutionary age of the species can be dated back to the pre-cellular period 3.5 billion years by molecular clock analysis and this is very consistent with the occurrence period of the RNA world. Direct molecular evidence of the Coevolution Hypothesis is presented by the finding of Endogenous Viral Elements (EVEs): in the horseshoe crab genome, 20 hcEVEs (horseshoe crab endogenous viral elements) are found, 55 percent of which are homologous to the genes of ancient Chuviruses, 78 percent of which are transcriptionally active, and some of which even integrate into the mRNA of the host to take part in protein synthesis, making it a complete model of coevolution between viruses and their hosts. As example, hcEVEs may have functional activity and are expected to help the host defend against current viral infections via sequence-specific recognition; the integration of such elements allows viral sequences to achieve stable replication in the host genome [13]. Such instances can be used in proving that EVEs are not genomic fossils rather, they are functional entities of the host genomes which are believed to have been subjected to evolution over long periods of time. This follows the RNA world model: the original viral replicase genes (in standing in the RNA viruses) were the extension of the already existing pre-cellular RNA replicons and their addition into the early cell genomes formed the basis of mutually dependent relationship at present [7].

The elaboration of the hypothesis is manifested in the enhancement of the version of the evolutionary mechanism of two-way adaptation: viruses incessantly modify their surface antigen structures with the help of gene mutations in order to avoid the immune response of the host (the annual antigenic drift of the influenza virus hemagglutinin protein); the host also continuously develops antiviral proteins (the interferons and restriction enzymes), on the other hand, it domesticates useful viral genes (the telomerase gene of eukaryotic cells is homologous to retroviral? Nonetheless, there exist anaks of evidence, too, associated with the Coevolution Hypothesis: there is no obvious fossil record of the primitive viruses during the pre-cellular period, while the speculations about the matter are only by reverse inference, searching the origins of currently extant viral genes; the process of how non-living nucleic acid fragments began to evolve into the typical features of a virus in general, notably the capsid shell, is also without a clear and complete explanation, which is a generic issue with all the hypotheses regarding virus origins.

2.4. Combination of the three hypotheses and research prospects

The three hypotheses concerning the origin of the virus do not contradict or exclude each other, however, each of them has its focus and scope of application. The scholarly community has slowly reached a consensus in recent years, owing to the accumulation of multiple research data, on the many possible sources: various types of viruses could have emerged via various evolutionary processes. In particular, bacteriophages could be represented by regressions of cells, RNA viruses possibly by co-evolution alongside pre-cellular replicators and retroviruses possibly by cellular genetic fragment escape. In a 2019 article published by *Nature Reviews Microbiology*, Krupovic et al. stated quite clearly that this diversity of origins is exactly the hallmark of the evolution of viruses, which is the embodiment of the dual nature of parasitism and genetic mobility [14].

These three hypotheses have found relevance in the academic value of their inspiration and promotion of several areas of research, with the Escape Hypothesis allowing important theoretical guidance to the creation of antiviral compounds, specifically the design of inhibitors to viral reverse transcriptase and other important enzymes based on cellular genetic components; the Regressive Hypothesis facilitating the study of the ecology of a given parasite, with the view of a researcher attempting to understand the evolutionary laws of a parasitic organism between complex and simple; the Coevolution Hypothesis having? It is predicted that in the future, by the thorough analysis of viruses in extreme conditions, e.g. deep-sea hydrothermal vents, polar permafrost, and systematic discovery of the viral remnants in host genomes in paleovirology, it will be able to definitively determine the numerous origins of viruses, seal the gaps in current evidence, and therefore keep on continually diversifying the theoretical framework of the evolution of life, provide new breakthroughs to life sciences.

3. Discussion

Controversies, Reconnaissance and Future of the Three Hypotheses on Virus Origin.

The genesis of viruses has served as a century-long unresolved kernel conundrum in the field of life sciences. Its distinct parasitic nature, its very simple structural design, and its complicated interrelationships with the cellular life complicate its classification directly via the standard evolutionary thinking. The academic community is consisted of three classic theories: the Escape Hypothesis, the Regressive Hypothesis, and the Coevolution Hypothesis They are based on three thoroughly opposite positions - cell fragment escape, complex cell regression and virus-cell co-evolution - and decades of accumulation of facts and theoretical arguments have made them demonstrate their own logical sense and explanatory capacity, as well as have illustrated fundamental weaknesses. These theories are not single opposing hypotheses, but they are in a complementary relationship, where each other is supplemented and revised, complementing each other in their role in understanding the human comprehension of the gist of viral evolution [5]. This article is going to discuss the main regimented foundation, and objectionable concentration of the three hypotheses, assimilate and assess them with the current research developments and propose an entire perspective of the virus's origin.

3.1. Escape Hypothesis: theoretical constraints and empirical support of a dialectical review

The Escape Hypothesis (also known as the Progressive Hypothesis) is one of the theories that has the most empirical grounding in the study of the origin of the virus; it was formally proposed as acceptable by Salvador Luria and James Darnell in their classic work *General Virology (2nd edition)* in 1967. Its central logic is the abandonment of one-way thinking of the virus-first theory or the cell regression theory, and claims that viruses are the result of cell genome fragments gaining the ability to escape and specialise.

Experiments can prove each step of this hypothesis in its three-stage evolutionary pattern, which includes genetic fragment escape, acquisition of infection and replication ability, and solidification of viral characteristics. During the genetic fragment escape phase, cells have mobile genetic elements which may escape by breaking away from cell restraints to become free fragments under environmental stress or gene regulation mistakes. These mobile genetic elements include transposons and plasmids. A direct contemporary biological example is the virus-like nucleic acid-protein complex which occurs upon the activation of Ty transposons in yeast. During the infective and replicating capability, free fragments selectively acquire the genes of capsid proteins by means of mutation or horizontal gene transfer. The percentage homology drift between bacteriophages and bacterial plasmids is above 60 per cent, and the homology of the pol gene of retroviruses to the LINE-1 retrotransposon in the human genome is 72 per cent, which proves the connection between the genes of viruses and the cellular genetic components. During the phase of solidifying viral properties, fragments become deprived of redundant genes and preserve their essential functions, which is consistent with the principle of the simplification of biological evolution through adaptive simplification.

The hypothesis was developed upon the finding of giant viruses in 2012 when Patrick Forterre et al. published a paper in Virus Research, which stated a prolongation of the hypothesis and suggested that the viruses developed as fragments of the chromosomes in primitive ribocells, which enhanced the information about the origin [15]. Earlier literature on viroids and satellite viruses is also supportive evidence: viroids are naked RNA infectious particles, which are thought to be early products of escape, and satellite viruses depend on helper viruses to replicate, and their genes are more homologous to their hosts, indicating that they are of recent origin, as well.

But the Escape Hypothesis has evident drawbacks: it is clear that the cell genome does not have clear homologous sequences of virus-specific spike proteins and some capsid proteins, such as the unique structure of the coronavirus S protein cannot be due to the mutation of cellular genetic elements, molecular clock analysis indicates an earlier evolutionary age of some viral genes (3.8 billion years ago) than cellular mobile genetic elements (3 billion years ago), giving rise to a timeline anomaly; the 1.57 Mb genome of Klos neu virus bears a The Rise of Giant Viruses. One can observe that it is capable of explaining only the origin of certain groups of viruses and must be used to fill gaps in other assumptions.

3.2. Regressive Hypothesis: in-depth analysis of empirical breakthroughs and core contradictions

Howard Temin stipulated the Regressive Hypothesis (also known as the Reduction Hypothesis) within The RNA Tumour Viruses in the 1950s. Its central thesis is that viruses represent regressive products of complex cells, which dropped genes in the course of long-term parasitism, which is sharply opposed to the thesis of the progressive origin of the Escape Hypothesis [16].

The rationality of its three phases of symbiosis-dependence-regression, is obvious: two kinds of primitive cells started a symbiotic relationship during the first phase, and the dependence stage so progressed that non-essential genes concerning metabolic functions and independent reproduction were mutated and lost, and only parasitism and transmission-related genes were kept, and the structures of the capsids evolved into obligatory parasitic viruses. This model has also been verified in parasitic bacteria: a 2017 study in Journal of Bacteriology has indeed demonstrated that the genome of Rickettsia lost over half its genes relative to the genomes of close free-living bacteria, but retained only functions linked to host binding and nutrient uptake, which is very consistent with the regression logic.

The finding of giant viruses was one of the pointers to the hypothesis: in 2004, Didier Raoult and colleagues described Mimivirus in Science, and it had a diameter of 750 nanometers and a 1.18 Mb genome full of cell homologs, and it was even initially identified as a bacterium [17]. Subsequently discovered Pandoraviruses and Pithoviruses have genomes approximately 2.5 Mb long, encode over 2,000 genes, and have cell-specific functional genes in abundance, strongly implying that they are the regression of primitive complex cells. In 2018, the group led by Raoult changed their opinion in light of the identification of Pandoravirus, making it possible that the origin of viruses was primordial symbiotic cells [17]. Indirect support of the hypothesis is also given by the universality of gene regression in the nature of parasitic organisms (like the 60% loss of metabolic genes in the parasite Plasmodium).

Nonetheless, the Regressive Hypothesis has fundamental contradictions: first, no clear homologous sequences of viral capsid proteins are found in the cell genome, which was impossible to account for by the cell protein regression model, which does not suit the logic of Coronaviridae, which is to lose its origin; second, molecular clock origin shows that the origin of the Coronaviridae can be traced back to 410 million years ago, earlier than Mimivirus (150 million years ago), which does not conform to the logic of giant viruses The Rise of Giant Viruses. Clearly, this hypothesis can only discuss the origin of certain complex viruses, and it cannot explain them all.

3.3. Coevolution Hypothesis: UTF objective assessment of subversive view and evidence vacuity

The Coevolution Hypothesis developed slowly in the 1980s and was enhanced by metagenomic studies. The essence of its theory is a contravention of the one-box origin logic of the earlier two hypotheses, asserting that viruses and cells came into existence concurrently, and have co-evolved since the pre-cellular life, with both playing a key role in determining how life has evolved.

It centres on the primordial replicator community theory: in the primitive soup of early Earth, there are many self-replicating pieces of nucleic acid that constituted the primitive community of biological macromolecules, including proteins and lipids, but no clear structures of cells. In evolution, the community was slowly split up: some of the pieces of nucleic acid developed into the genome of early cells, eventually developing cell structures; others developed capsid protein genes to create early viruses, which were transmitted and transmissible through infection into early cells; the two evolved in a co-evolutionary relationship, each providing and selecting genes to the other.

The RNA world theory offers valuable evidence to the supposition: being the first biological macromolecule that stored its genetic information and acted as a catalyst at the same time, RNA was the primitive life system. RNA viruses also do not require DNA intermediates to replicate, and the genes of their replicase are very conserved. Their evolutionary time may be traced to the pre-cellular 3.5 billion-year-old which correlates with the time of existence of the RNA world, resolving the timeline dilemma of the two earlier hypotheses. This observation has been corroborated by the discovery of a direct molecular evidence: a

2025 study in Journal of Virology has identified 20 hcEVEs in the horseshoe crab genome of which 55 per cent are homologous to ancient Chuviruses, 78 per cent are transcriptionally active, and some are integrated into host mRNA in order to be involved in protein synthesis, a full co-evolution model [18]. Endogenous retroviruses are approximately 8 percent of the human genome of a sequence and some of them are domesticated to be involved in various important physiological activities including placental formation (including the syncytin gene) and gene exchange and co-adaptation between viruses and their hosts.

Another highly revealing description of the "two-way adaptation" process in the Coevolution Hypothesis is that viruses evade immune recognition by mutating the genes (e.g. antigenic drift in the influenza virus HA protein), and hosts adapt to the viruses via evolution of antiviral proteins and domestication of viral genes. The constant mutual selection will facilitate the co-evolution of both parties, optimising the mechanism of infection of viruses and complicating the defence system of cells, and, after all, the complexity of the life system and its diversity. Metagenomic research, which has highlighted that marine viruses outnumber cells and all lysed marine bacteria by a factor of nearly 40 per cent on a daily basis indicate nutrients and swathes of help-seeking fragments of genes, their contribution to global biogeochemical cycles and horizontal gene transfer is profound in that opinion that viruses are a key player in evolution.

Nonetheless, gaps in the evidence of the hypothesis are distinct: there is no known direct fossil evidence of primitive viruses during the pre-cellular period, and any speculation is reliant on the reverse-tracing of extant viral genes: the originated viral feature of its capsid structure and mode of infection has no evident evolutionary pattern of account, the genetic diversity of bacteriophages is far more extensive than that of host bacteria, and homology across groups is extremely low, which is incompatible with a telephone explanation of co-evolution.

3.4. Virus origin an integrated approach: majority and personal thoughts of multiple origins

According to the main logic, empirical evidence and theoretical constraints of the three hypotheses, alongside the most recent research developments, I think that the source of viruses is the main characteristic of multiple origins, that is, different sets of viruses have different evolutionary histories, and the three hypotheses explain the way of origin of one or another set of viruses, and do not exclude each other but supplement one another. This is very much in line with the conclusion of the 2019 review by Krupovic et al. in Nature Reviews Microbiology [14].

In the light of the evolutionary behavior of the virus groups, evidence of multiple origins is evident: retrovirus and some single-stranded RNA virus genes are highly homologous to cellular mobile genetic elements, simple structures, which are consistent with the logic of the Escape Hypothesis, giant viruses with large genomes and rich in cell homologous genes, which are consistent with the logic of the Regressive Hypothesis, old groups, such as the Flaviviridae and Hepaciviridae among RNA viruses, with high retrovirus replicase genes It is a multiple model of one origin path of one virus group which is a synthesis of reasonable centrality to the three hypotheses and is so far the most comprehensive explanatory framework.

Based on the general logic of the evolution of life, there are various sources that consider the law of natural selection. At an early phase in the origin of life, the primitive life system was not stable, and self-replicating fragments of nucleic acid could accomplish self-transmission in a variety of ways: some primitive cells escaped and differentiated to produce escape-type viruses; some primitive symbiotic cells regressed as they over-relied on their host to produce regressive-type viruses and some differentiated alongside primitive cells to produce co-evolutionary-type viruses. These primitive viruses transposed and exchanged genes and horizontally with each other until they developed the variety of current viruses, which is well in line with the laws of convergent evolution and parallel evolution in biological evolution.

My personal opinion is that the Coevolution Hypothesis is the top-level perspective of the origin of viruses, the Escape Hypothesis and the Regressive Hypothesis are the origin mechanisms of the particular groups of viruses, and the three are the whole-part explanation system. The fundamental virtue of the Coevolution Hypothesis is to give viruses a step up into the ranks of life evolution, as something more than cell derivatives. This school of thought can resolve the timeline conflict of the earlier two hypotheses and better fits the origin of life that is more integrated and complex. Viruses and cells are not merely parasites and hosts, but co-evolutionary partners- viruses make genetic material available to cells and induce cell complexity; cells make replication environments available to viruses and induce the evolution of viruses. This association extends into the whole process of the evolution of life.

The future advances in the viral origin will be based on three research directions: metagenomic technology to study the diversity of viruses in extreme environments and reveal genetic remnants of ancient viruses; the integration of structural biology and evolutionary biology to study the three dimensional structure of core viral proteins to trace their evolutionary relationship with cell proteins; synthetic biology to establish primitive simulation systems of viruses in vivo to confirm the evolutionary relationship of various hypothesis of origin. These researchers are capable of solving the enigma of the origin of the virus, as well as contributing to the insight into the laws of life's appearance and evolution.

Virus origin exploration is basically a mechanism of enhancing human knowledge about what life is all about. Since the inclusion of viruses as non-living infectious objects, through the establishment of complex evolutionary relationships with cells, and the establishment of a consensus on the issue of multiple origins, the theoretical advances have been accompanied by the development of empirical studies. Despite the existence of a lot of unexplored mysteries, there are still no clear answers; however, the complementary arguments of the three hypotheses have provided the overall map of the genesis of a virus. This central riddle of the life sciences will be resolved sooner or later with the advancement of science and technology, and the study of the origin of the virus will further give new insights and facts in the theory of life evolution.

4. Conclusion

Being ubiquitous and highly adaptive pathogens, the RNA viruses have puzzled the scientists with their evolutionary background. It scientifically analyzes three mainstream theories that are Escape, Regressive, and Coevolution by synthesizing the academic data and scientific findings, and logic reasoning, thus trying to elucidate the scientific strengths of these theories, and pointing out the importance of studying the origins of RNA virus.According to the analysis, none of the three hypotheses could help to explain all origins of RNA viruses, but each has its area of influence. Gene homology between retroviruses and cellular transposons found in support of the Escape Hypothesis fails to address timeline conflicts between the genes of viruses and cellular components, as well as to explain the source of the virus-specific proteins. Given the complexity of genomes of giant viruses, such as Mimivirus, and the damage to genes in parasitic bacteria, such as Rickettsia, the Regressive Hypothesis is valid, but does not accord well with the simplicity of most RNA virus genomes and the absence of capsid protein homologs in cells. The Coevolution Hypothesis resolves the timeline issues of the other two; the evidence of conserved replicase genes and Endogenous Viral Elements (EVEs) in horsehoe crab genomes redirects the hypothesis to being highly explanatory. The pluralistic origin view can be considered the most reasonable one: most ancient RNA viruses have been produced by degraded complex cells, retroviruses by lost fragments of cells, and giant viruses by coevolution, fitting into the research framework proposed by Krupovic et al. [14]. This exploration informs the worldwide public health beyond theory. Knowledge of evolutionary principles aids in efficacies in anticipating the mutations that could arise within the viruses and the chance of transmitting the viruses to other organisms besides those that could be self-medicated, and known molecular targets of broad-spectrum antiviral agents are encouraged. Future studies must apply metagenomics to examine the nature of extreme-environment viruses via the application of synthetic biology in order to reproduce incomplete areas of evidence. In short, the pluralistic origin theory and the complementarity of the three hypotheses have provided a strong basis in spite of controversies. As technologies evolve, this sector is going to enhance the research efforts on the evolution of life, which will be more significant in case of viral menaces and their protection of world health.

References

[1] Iuliano, A. D., Roguski, K. M., Chang, H. H., Muscatello, D. J., Palekar, R., Tempia, S., Cohen, C., Gran, J. M., Schanzer, D., Cowling, B. J., Wu, P., Kyncl, J., Ang, L. W., Park, M., Redlberger-Fritz, M., Yu, H., Espenhain, L., Krishnan, A., Emukule, G., … Bresee, J. S. (2018). Estimates of global seasonal influenza-associated respiratory mortality: A modelling study. *The Lancet, 391*(10127), 1285–1300.

[2] UNAIDS. (2020). *2020 global AIDS update: Seizing the moment—Tackling entrenched inequalities to end epidemics*. https: //www.unaids.org/en/global-aids-update-2020

[3] Gilbert, W. (1986). Origin of life: The RNA world. *Nature, 319*(6055), 618.

[4] Zhan, Y., Li, X., Wang, J., & Zhang, Y. (2023). Phylogenetic analysis of RNA virus evolution: New perspectives from metagenomics. *Journal of Virology, 97*(12), e00892-23.

[5] Krupovic, M. (2019). The origins and evolution of viruses: New insights from genomes. *Nature Reviews Microbiology, 17*(8), 505–519.

[6] Luria, S. E., & Darnell, J. E. (1967). Origin of viruses. *Science, 157*(3791), 853–859.

[7] Forterre, P., & Krupovic, M. (2012). The origin of virions and virocells: The escape hypothesis revisited. *Virus Research, 164*(1–2), 23–36.

[8] Temin, H. M. (1970). On the origin of RNA tumor viruses. *Nature, 227*(5258), 512–517.

[9] Zhirnov, A. Y., & Koonin, E. V. (2020). On the origin of RNA viruses and their replication modules. *Philosophical Transactions of the Royal Society B, 375*(1807), 20190481.

[10] Goudsmit, J. (2004). The origin of viruses. *Clinical Microbiology Reviews, 17*(4), 760–765.

[11] Noriea, N. F., Clark, T. R., Mead, D., & Hackstadt, T. (2017). Proteolytic cleavage of the immunodominant outer membrane protein rOmpA in Rickettsia rickettsii. *Journal of Bacteriology, 199*, e00826-16.

[12] Koonin, E. V., Dolja, V. V., & Krupovic, M. (2015). Evolution of the virosphere. *Current Opinion in Virology, 12*, 67–76.

[13] Gao, Y., Zhang, H., Liu, J., & Wang, X. (2023). HCV-derived endogenous elements (HCVEs) mediate sequence-specific antiviral immunity and stable viral genome integration in host cells. *Journal of Virology, 97*(8), e01923-22. https: //doi.org/10.1128/jvi.01923-22

[14] Krupovic, M., Dolja, V. V., & Koonin, E. V. (2019). The LUCA and its complex virome. *Nature Reviews Microbiology, 17*(4), 219–232.

[15] Forterre, P. (2015). The virus world as evolutionary driver. *Current Opinion in Virology, 12*, 103–109.

[16] Temin, H. M. (1959). *The RNA tumor viruses*. Cold Spring Harbor Laboratory Press.

[17] Na, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., & Claverie, J. M. (2004). The 1.2-megabase genome sequence of Mimivirus. *Science, 306*(5700), 1344–1350.

[18] Wang, L., Li, Y., Zhang, Y., Liu, Y., Zhao, X., & Chen, J. (2025). Endogenous Chuvirus elements in horseshoe crab genomes reveal ancient viral-cellular coevolution. *Journal of Virology, 99*(3), e01523-24.