DOI: https://doi.org/10.17816/brmma354241

Review Article

EVOLUTIONARY MECHANISMS OF VIRUS VARIABILITY

A.V. Moskalev[1], B.Yu. Gumilevskiy[1], V.Ya. Apchel[1, 2], V.N. Tsygan[1]

[1] Kirov Military Medical Academy, Saint Petersburg, Russia

[2] A.I. Herzen Russian State Pedagogical University of the Ministry of Education and Science of the Russian Federation, Saint Petersburg, Russia

ABSTRACT. The evolutionary changes of viruses are primarily associated with the replication processes of viruses containing deoxyribonucleic and ribonucleic acids, which differ significantly. The genomes of most viruses containing ribonucleic acid are replicated with much less accuracy compared to the genomes of viruses containing deoxyribonucleic acid. Comparing the number of mutations in an infected cell reflects an inverse relationship between genome size and the frequency of mutations carried out by these two categories of viruses. Viruses with double-stranded deoxyribonucleic acid genomes have a low mutation rate compared to single-stranded genomes. The genome of viruses is not a stable unique structure, but rather an average, variable number of different amino acid sequences. It is in the virus population that a high mutation rate is maintained, and low variability is not beneficial for the preservation of viruses in nature. Some animal species may be intermediate hosts when new epidemic viruses appear. The introduction of non-viral nucleic acid into the viral genome can also contribute to the evolutionary changes of the virus, lead to the formation of defective genomes or to the emergence of hypervirulent strains. Viral genomes encode numerous molecules that modulate a wide range of protective immune mechanisms. The variability of viruses is also facilitated by the simultaneous integration of several proviral genomes into one cell, which activates the processes of recombination and genetic shift. An important evolutionary point may be the conversion of ribonucleic acid ribose into deoxyribose of deoxyribonucleic acid, which increases the stability of nucleic acids by more than 100 times. Horizontal gene transfer between viruses that infect different hosts is a central feature of the evolution of viruses containing ribonucleic acid. Eukaryote viruses with single-stranded deoxyribonucleic acid probably evolved from bacterial plasmids after they acquired capsid protein genes from the (+) ribonucleic acid chain of viruses. In addition to megaviruses and adenoviruses, polintons are likely precursors to bidnaviruses and virophages.

Keywords: viruses; evolution of viruses; gene; proviral genomes; hypervirulent strains; immune system; mutations; nucleic acids; single-stranded deoxyribonucleic acid.

To cite this article:
Moskalev AV, Gumilevskiy BYu, Apchel VYa, Tsygan VN. Evolutionary mechanisms of virus variability. *Bulletin of the Russian Military Medical Academy.* 2023;25(2):301–316. DOI: https://doi.org/10.17816/brmma354241

Received: 29.03.2023 **Accepted:** 16.05.2023 **Published:** 15.06.2023

УДК 576.8.097.31

DOI: https://doi.org/10.17816/brmma354241

Обзорная статья

ЭВОЛЮЦИОННЫЕ МЕХАНИЗМЫ ИЗМЕНЧИВОСТИ ВИРУСОВ

А.В. Москалев[1], Б.Ю. Гумилевский[1], В.Я. Апчел[1, 2], В.Н. Цыган[1]

[1] Военно-медицинская академия имени С.М. Кирова, Санкт-Петербург, Россия

[2] Российский государственный педагогический университет имени А.И. Герцена, Санкт-Петербург, Россия

Резюме. Эволюционные изменения вирусов в первую очередь связаны с процессами репликации вирусов, содержащих дезоксирибонуклеиновую и рибонуклеиновую кислоты, которые значительно отличаются. Геномы большинства вирусов, содержащих рибонуклеиновую кислоту, реплицируются с гораздо меньшей точностью по сравнению с геномами вирусов, содержащих дезоксирибонуклеиновую кислоту. Сравнение количества мутаций в инфицированной клетке отражает обратную зависимость между размером генома и частотой мутаций, осуществляемых этими двумя видами вирусов. Для вирусов с двухцепочечными геномами дезоксирибонуклеиновой кислоты характерна низкая частота мутаций по сравнению с одноцепочечными геномами. Геном вирусов представляет собой не стабильную уникальную структуру, а скорее всего усредненное, вариабельное количество различных аминокислотных последовательностей. Именно в популяции вирусов поддерживается высокая частота мутаций, а низкая изменчивость является невыгодной для сохранения вирусов в природе. Некоторые виды животных могут быть промежуточными хозяевами при появлении новых эпидемических вирусов. Введение невирусной нуклеиновой кислоты в вирусный геном может также способствовать эволюционным изменениям вируса, приводить к образованию дефектных геномов или появлению гипервирулентных штаммов. Вирусные геномы кодируют многочисленные молекулы, модулирующие широкий спектр защитных иммунных механизмов. Вариабельности вирусов способствует и одновременная интеграция нескольких провирусных геномов в одну клетку, что активизирует процессы рекомбинации и генетического сдвига. Важным эволюционным моментом может быть превращение рибозы рибонуклеиновой кислоты в дезоксирибозу дезоксирибонуклеиновой кислоты, что увеличивает стабильность нуклеиновых кислот более чем в 100 раз. Горизонтальный перенос генов между вирусами, которые заражают различных хозяев, является центральной особенностью эволюции вирусов, содержащих рибонуклеиновую кислоту. Вирусы эукариотов с одноцепочечной дезоксирибонуклеиновой кислотой, вероятно, эволюционировали из бактериальных плазмид после приобретения ими генов капсидных белков из (+) цепи вирусов, содержащих рибонуклеиновую кислоту. Кроме мегавирусов и аденовирусов полинтоны являются вероятными предшественниками биднавирусов и вирофагов.

Ключевые слова: вирусы; эволюция вирусов; ген; провирусные геномы; гипервирулентные штаммы; иммунная система; мутации; нуклеиновые кислоты; одноцепочечная дезоксирибонуклеиновая кислота.

Как цитировать:
Москалев А.В., Гумилевский Б.Ю., Апчел В.Я., Цыган В.Н. Эволюционные механизмы изменчивости вирусов // Вестник Российской военно-медицинской академии. 2023. Т. 25, № 2. С. 301–316. DOI: https://doi.org/10.17816/brmma354241

Рукопись получена: 29.03.2023 **Рукопись одобрена:** 16.05.2023 **Опубликована:** 15.06.2023

INTRODUCTION

Researchers have been concerned for centuries with the question of what evolutionary mechanisms allow viruses to escape the control of the immune system (IS). Why is immunological memory not sufficiently effective to eliminate the virus from the macroorganism? Thanks to improvements in rapid sequencing techniques, we now have information about the variability of viruses and the peculiarities of immunopathogenesis of viral infections. At present, researchers can detect traces of viruses not only in nature but also in the deoxyribonucleic acid (DNA) of living organisms. Evolutionary changes concern not only viruses but also their hosts; as a result, viruses that successfully overcome established immunological defense mechanisms are selected. In fact, the evolution of viruses can be viewed as a kind of "arms race", in which viral genes are selected from their host cells. Although viruses circulate in a vast genetically variable host population and in constantly changing conditions, owing to their genetic diversity caused by mutations, recombinations of viral genes are preserved with the most favorable properties of their genomes. In viral infections, hundreds or thousands of viral particles with different characteristics are formed even after a single cycle of reproduction in a single cell. This diversity in viruses ensures their survival under the most unfavorable conditions. This results in a significantly different virus genotype even after a single selective event. Positive and negative selection of pre-existing virus genotypes can occur at any stage of the life cycle of viruses. This is influenced by the need for viral spread in the host organism, population density, immune homeostasis, and immunological memory. Two general strategies for virus survival are distinguished. The first is the r-propagation strategy characterized by short reproductive cycles and high reproductive output. The second is the K-propagation strategy with a low output of viral particles during reproduction with the development of latent variants of the course of infections. Under these conditions, the viral genome is in a stable environment, and the virus can exist as long as its host lives. In some cases of infections caused by the herpes simplex virus and the human immunodeficiency virus (HIV), both the R-strategy of reproduction (productive infection) and the K-strategy characterizing the latent course of infection with low reproduction occur.

Genomes of viruses containing ribonucleic acid (RNA) often replicate with errors close to the threshold. By contrast, the replication of DNA viruses proceeds with higher accuracy, which is well below the error threshold. Moreover, mutations accumulate with each replication cycle of viruses, and with their high frequency, viruses of the same species and type in a population can have different antigenic properties. The peculiarities of virus evolution should be considered in the light of two hypotheses of virus origin: (1) viruses arose before cells and may have participated in their organization, and (2) viruses arose after the cells as their genetic elements and acquired the ability for autonomous replication. Information obtained about the amino acid sequences of viruses in host cell genomes provides important information about the consequences of the interaction between the virus and host IC cells over evolutionary time. In the last decade, advanced computational methods and an expanded database of viral genomes made possible by metagenomic studies have enabled outlining of the major general trends in viral evolution. All these considerations are the reasons why we are interested in considering the evolutionary features of changes in the antigenic structures of viruses.

This study aimed to summarize new literature data on the mechanisms of virus variability and peculiarities of their escape from IS control.

MATERIALS AND METHODS

The current scientific literature has focused on the evaluation of the evolutionary mechanisms of the variability of viruses that support their vital functions.

RESULTS AND DISCUSSION

The R-strategy of multiplication leads to the production of up to 10 000 viral particles by a single cell infected with poliovirus in 8 h. This rate of viral multiplication over 3–4 cycles of multiplication can ensure the production of viral particles enough to infect every cell in the body. However, this does not happen for several reasons: with IP control, viruses can multiply only in certain tissues and cell types. Nevertheless, *the R-strategy* of reproduction is a characteristic of many infections, and high rates of viral particle production can persist for years. For example, in the late stages of HIV infection, it takes only 1.6 days for thousands of viral particles to leave the cell and reach other T-lymphocytes, followed by lysis of these cells. Mutations that accompany genome replication accumulate, although these errors are not yet easy to quantify, and they depend on the peculiarities of the experiment [1].

Replications of DNA and RNA viruses differ considerably. The genomes of most RNA viruses replicate with much less precision than those of DNA viruses. The average error rate of RNA virus genomes is approximately one missense integration per 104–105 polymerized nucleotides, which

corresponds to an average of one mutation in each replicated genome. However, not all viral RNA genomes have the same mutation rate. In viruses with larger genomes, RNA replications proceed with higher precision. This is also true for coronavirus; however, the inactivation of the correcting exonuclease leads to a 15–20-fold increase in the frequency of viral mutations, both *in vitro* and in mouse model [2].

The estimated replication error rate of viral DNA is approximately 10^{-6} to 10^{-8}, which is closer to the replication rate of host cell genomes than for most RNA viral genomes. A reason for this difference is that most DNA-dependent polymerases can excise and replace incorrectly incorporated nucleotides, whereas many RNA polymerases have no such error correction mechanisms. A comparison of the number of mutations in an infected cell reflects an inverse relationship between the genome size and frequency of mutations carried out in both RNA and DNA viruses. However, these values can also be higher because the frequency of polymerase errors caused by RNA editing or spontaneous damage of viral nucleic acids (DNs) by oxygen radicals or ionizing radiation is not taken into account. DNA viruses with double-stranded (ds) genomes are characterized by a low mutation rate because complementary strands ensure the repair of altered sites. The replication of small single-stranded (ss) genomes of DNA viruses (*Parvoviridae and Circoviridae*) is accompanied by more pronounced mutational changes than DNA genomes of larger viruses [3].

The study of the RNA population of bacteriophage Qβ has led to a unique conclusion. Although the Qβ bacteriophage population is in dynamic equilibrium, viral mutants arise at a high rate. Therefore, the Qβ phage genome is not a stable unique structure but an averaged, variable number of different amino acid sequences. This conclusion was subsequently confirmed for many virus populations. Accordingly, viral populations exist as a dynamic distribution of non-identical but related replicons, called quasispecies. Such a particular population of viral quasispecies must contain a very large number of viral particles because such equilibrium cannot be achieved in small populations. Under these conditions, extreme gene and phenotype fluctuations are possible [4].

The RNA virus population is characterized by the average, consensus sequences of its genome; however, nearly every genome is different. Rare mutational changes in the genome can persist in all genomes of the viral progeny. Other mutations tend to also persist in the genome. Therefore, the product of selection after replication is a new, sufficiently diverse population of genomes in which selected and closely related mutations are preserved. This concept of quasispecies shows that the viral population is not just a set of diverse mutants but a group of interactive variants characterizing a particular population. Population diversity is crucial to the survival of viruses, and viral populations, not individual mutants, are the objects of selection, limiting viral diversity. For example, some spontaneous mutants of HIV type 1 that are resistant to the reverse-transcriptase inhibitor lamivudine show a 3.2-fold reduction in the error rate. Another example is the poliovirus, whose replication is accompanied by errors and production of a variable population. Thus, ribavirin-resistant mutant poliovirus strains are characterized by approximately sixfold reduced variability; however, such mutants are low pathogenic for animals compared with parental viruses. This reduction in diversity has resulted in the weakening and loss of neurotropism by polioviruses. Thus, in a genetically diverse population, viral mutants complement each other, i.e., the population develops, not the individual. A high mutation rate is maintained in the population of viruses, and low variability is disadvantageous for the conservation of viruses in nature. Contrary to quasispecies dynamics, viruses with low mutational capacity turn out to be more adapted to survive and reproduce in a certain environment in which viruses with high mutational variability are found. Such a population is characterized by the so-called quasispecies effect [5, 6].

Despite the high mutation rate, not all viral genome sequences change. For example, the cis-regulatory elements of RNA viruses slightly change during reproduction. These sequences include signals necessary for genome replication, matrix RNA (mRNA) synthesis, and genome packaging. They are often binding sites for one or more viral or cellular proteins. The mutational genome encoding the corresponding viral binding protein may not replicate at all. Changes in both interacting components, i.e., the binding protein and the mutation genome sequences, are necessary to restore function. Their close functional relationship markedly reduces evolutionary changes. Many biological parameters influence virus reproduction and survival. Naturally, among others, population dynamics and seasonal variations in temperature, humidity, etc., are the most important [7].

Takata et al. [8] added a high dose of ribavirin to the medium of cells infected with poliovirus, whereas the part of the cells remained intact. The extracted viral RNA from the resulting progeny was injected into new host cells by transfection. A 50% loss of infectivity index (LI50) was assessed. At LI50, two mutations per genome, 50% of viral genomes mutate lethally. This is nearly two times higher than naturally occurring mutations. The deterioration of genomic viability at normal mutation rates in the population (approximately 10%) does not occur. At 1.5 mutations per

genome, no significant loss of function occurs. Two mutations per genome reduce the infectivity of wild-type virus RNA to 30%, seven mutations per genome make the virus RNA avirulent, and the poliovirus is on the verge of viability.

Thus, genetic diversity in the viral population by recombination or reassortment facilitates the construction of genomes with minimal mutational changes, which compensates for the presence of harmful mutations. This phenomenon is quite rare; however, the offspring of this rare virus will eventually prevail in the population.

Diversity in viral populations is provided by two mechanisms, namely, genetic drift and genetic shift. Diversity arising from genome replication errors and immune selection of single-site mutants (drift) differs from the recombination between genomes or recombination of genome segments (shift). Drift occurs every time a genome replicates, whereas shear is relatively rare. Episodic influenza pandemics support this conclusion. For example, only six cases of hemagglutinin (H) genetic shift in influenza viruses have been identified since 1889. Nevertheless, the combination of drift and shift, together with the presence of intermediate virus species, contributes significantly to annual influenza infection outbreaks (Fig. 1) [9].

Major influenza pandemics (Fig. 1) are characterized by viral reassortants. The reassortants carried H and neuraminidase (N) genes that had not been in circulation in humans for some time; thus, there was little or no immunity. With the emergence of the new subtype, an influenza pandemic characterized by a new H, or a new combination of H and N had occurred. Segments of the viral genome are illustrated in three colors, each representing a particular viral genotype. Segments and gene products of pandemic strains are labeled in each human silhouette. The numbers next to the arrows show the number of segments of the viral genome replaced in each episode.

New strains of influenza virus type A can appear after the recombination of human and avian influenza viruses in pigs. Such pandemic strains are the result of an exchange of segments of the genomes of human, swine, and bird influenza viruses. Consequently, human and avian influenza viruses replicate well in the pig organism regardless of the H and N composition. This occurs because the epithelium of the mucous membrane of the upper respiratory tract of pigs contains the same receptors as those of birds and humans. Such a recombination of H and N segments of influenza viruses is infrequent; however, the high population density creates conditions for the emergence of new pandemic strains. For example, virologists have confirmed that strains of the pandemic influenza virus type A from 1957 and 1968 that originated in China circulate in wild bird populations. Studies in Italy in the late 1980s confirmed the reassortment processes of avian and human influenza viruses in pigs. Sequencing and phylogenetic analysis techniques confirmed that pigs could be an intermediate host for new pandemic influenza viruses. Indeed, such recombination led to the emergence of the 2009 “swine influenza” type A–H1N1 pandemic [10].

The H1N1 influenza virus probably caused 105,700–400 000 deaths worldwide in the first year alone, with 46 000–179 000 people probably dying from cardiovascular complications. The pandemic officially ended in 2010, and H1N1 is now considered a seasonal flu virus. This genetic exchange is an important source of variation, as confirmed by reoviruses and orthopoxviruses. The introduction of nonviral NK into the viral genome can also contribute to evolutionary changes in the virus, leading to the formation of defective genomes or emergence of hypervirulent strains. As a result of recombination processes, cytopathic viruses can appear in nonpathogenic bovine diarrhea viruses, and oncogenic viruses can appear after infection with non-oncogenic strains. This is a consequence of viral genes acquiring the genetic material of the host cell, which is characteristic of transforming retroviruses such as the Rous sarcoma virus [11, 12].

A unique phenomenon in the life activity of viruses is their basic properties being barriers to pronounced genetic changes. Thus, only retrovirus genomes can be transformed into DNA and vice versa. This occurs because the mechanisms

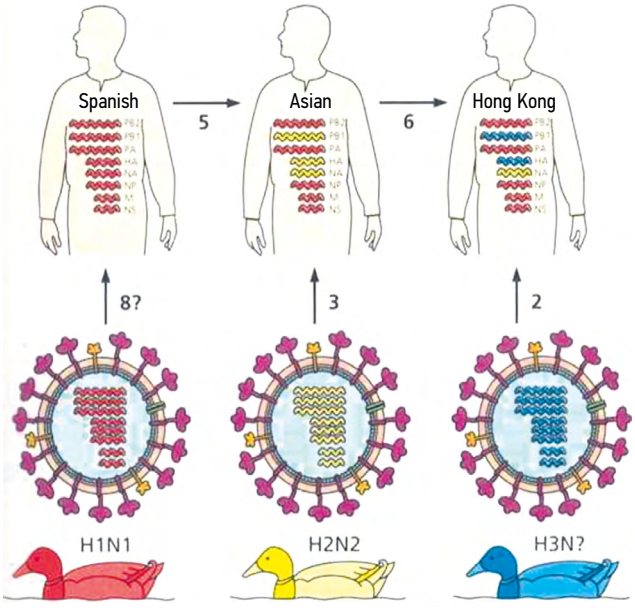

Fig. 1. Appearance and transmission of distinct serotypes of influenza A virus in human pandemics in the 20th century (on J. Flint, V. Racaniello, G. Rall et al. Principles of Virology. Fifth edition. Vol. II. 2020)

Рис. 1. Появление и передача различных серотипов вируса гриппа А при пандемиях в XX в. (по Дж. Флинту, В. Раканьелло, Г. Ралла и др. Принципы вирусологии. 5-е изд. Т. II. 2020)

that ensure genome replication and expression are selected in the course of evolution, and viruses most often die under extreme influences. In addition, each stage of virus reproduction requires its close interaction with the host cell. Any change in the viral component, which is not supported by compensatory changes in the cellular mechanism, can lead to impaired reproduction and the spread of the virus in the host. Evolutionary changes are also limited by the nature of the capsid, which is necessary for genome transmission. The capsid determines the size of the NCs that are packaged in them. Finally, viral genomes encode numerous molecules that modulate a wide range of immune mechanisms [13].

Evolutionary changes are also enhanced because viruses can occupy wider niches, infecting several intermediate hosts. When a virus is eliminated from one organism, its population can survive in another host, which is more typical of RNA viruses. A high virulence also makes it easier for viruses to infect a wide range of organisms. For example, the 2013–2016 outbreaks of Ebola viruses in West Africa are extremely contagious. Viral variability is also facilitated by the simultaneous integration of several proviral genomes into a single cell, which contributes to recombination and genetic shift. Electron microscopy of lambda and related bacteriophages revealed that different pairs of viral DNA sequences formed heteroduplexes when homologous ds sites were linked to corresponding nonhomologous ss sites. The genomes of this group of phages were mosaic, containing blocks of genes shuffled by recombination during evolution. Such nonhomologous recombination of the genetic materials of bacteria and bacteriophages is a central feature in their evolution. This allowed us to establish that bacteriophage genomes have hereditary relationships with *Eukarya* and *Archaea* viruses [14].

An important evolutionary point may be the change of the ribose RNA to deoxyribose DNA, which increases the stability of NK by more than 100-fold. This is important because DNA can store much more information. In addition, the presence of a complementary strand for ss-gap repair is an additional advantage for dsDNA genomes (dsDNA). The transition from RNA to DNA genomes was made possible by the evolution of genes encoding reverse transcriptase and enzymes required for deoxynucleotide synthesis (e.g., ribonucleotide reductase and thymidylate synthetase) [14].

Sequences of millions of viral genomes are currently known, and this is constantly increasing. In addition, it is now possible to take a sample from an ecosystem and determine the nature and diversity of viral genomes. Metagenomic analysis has revealed an amazing diversity among known families of viruses. More surprisingly, the vast majority of the viral sequences identified represent previously unknown viral genomes. Although the origin of viruses remains largely unclear, advances in metagenomic analysis and bioinformatics have provided data on the ancestors of modern eukaryotic virus families. Studies of proteins necessary for genome replication (e.g., RNA-dependent RNA polymerases) have made it possible to trace the lineage of viruses that affect eukaryotic organisms. Such phylogenomic reconstructions have provided new insights into the likely sequences that led to the emergence of various eukaryotic virus families [15].

Although RNA viruses are rare in bacteria and archaea, they are frequently detected in eukaryotes. Physical barriers associated with eukaryotic cell compartmentalization partially explain this striking difference. The genomes of many eukaryotic RNA viruses replicate in the cytoplasm in viral compartments or on cell membranes. Viruses containing the RNA(+) chain have a simple multiplication strategy because their genomes can be converted into proteins. These viruses have become the main pool for studying the evolution of RNA viruses in eukaryotes. Phylogenetic comparisons of polymerase gene sequences have shown that bacterial levivirus is a likely precursor of eukaryotic (+) strand RNA replicons, such as mitoviruses and narnaviruses, which reproduce in the mitochondria. Migration into the cytoplasm of one such replicon and its acquisition of a picornavirus capsid is a plausible explanation for the origin of plant ermyaviruses. Phylogeny, i.e., the relationships between different groups of organisms and their evolutionary development, up to the precursor viruses containing the RNA (+) chain of other eukaryotic RNA viruses, has been studied. These precursors give rise to four branches. The first branch includes the widespread picornaviruses, and the second branch includes flaviviruses and alphaviruses. Comparison of the sequences of their RNA-dependent RNA polymerases also suggests that viruses with ds RNA genomes evolved from two separate branches of (+) strand RNA viruses [16, 17].

A group of dsRNA viruses arose from the same branch as the picornaviruses, and another group of viruses can be traced directly back to the ancestor of the primary (+) RNA chain. In this scheme, viruses with (–) chain RNA genomes, such as bunyaviruses, filoviruses, and orthomyxoviruses, are among the last viruses to emerge, through the loss of the (+) chain dsRNA viruses in this second group. Genome analysis of the phylogenetic tree reflects the increasing accumulation of additional genes encoding chelicases and extensive replacement of functional modules. Horizontal gene transfer between viruses that infect different hosts is a central feature of RNA viral evolution and is consistent with the proposed phylogenetic scheme [18].

After infection, retroviral reverse transcriptase and integrase transform the RNA genomes into a DNA copy,

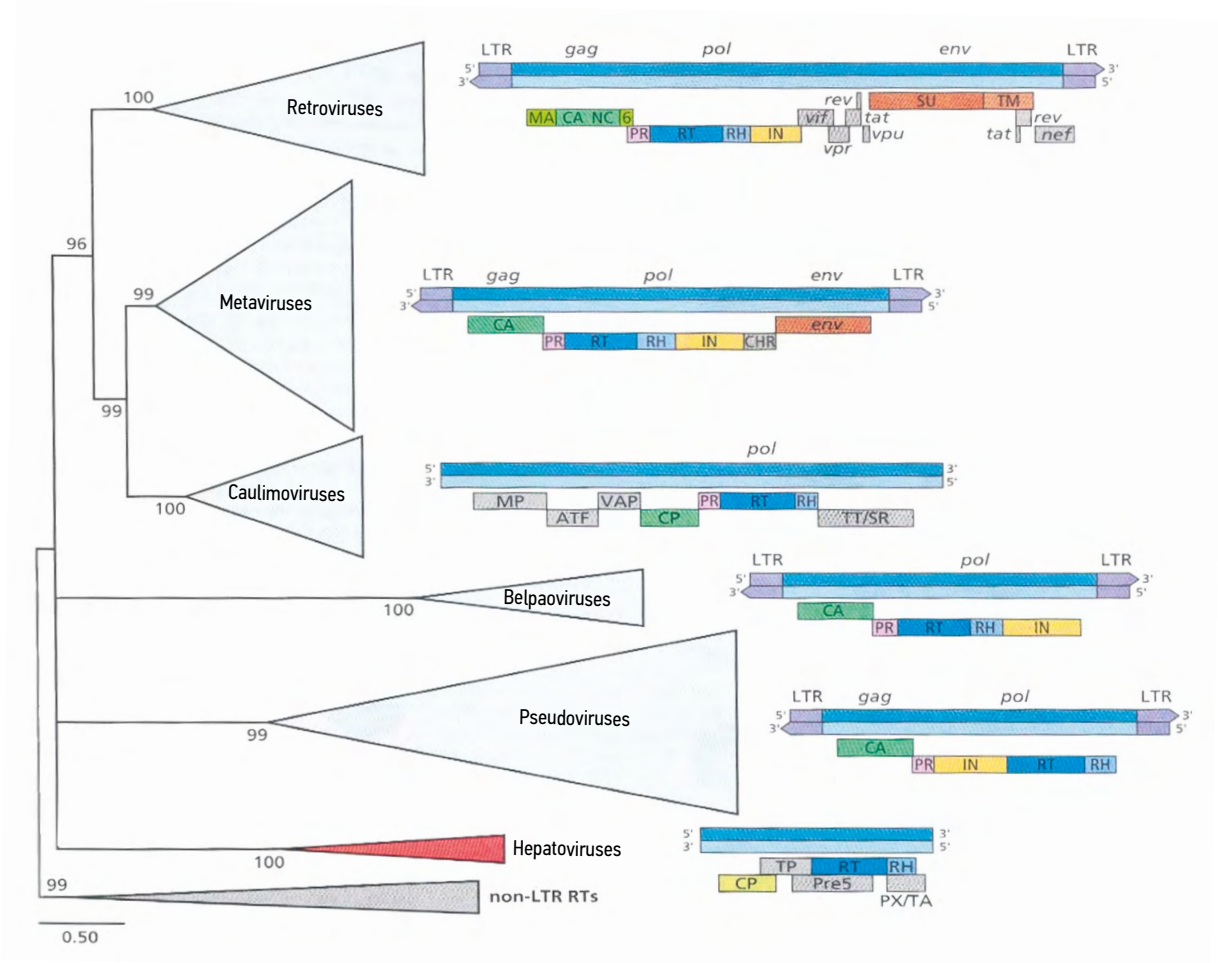

Fig. 2. Phylogeny of reverse transcriptases in retroviruses and pararetroviruses: env — envelope genes; gag — group-specific antigen; IN — integrase; LTR — long terminal repeat; MA — matrix protein; MP — motion protein; NC — nucleocapsid; nef tat, rev, vif, vpr, vpu — human immunodeficiency virus type 1 genes that express regulatory proteins via mRNA; P — polymerase; pol — polymerase genes; PR — protease; PreS — pre-surface protein (shell); PX/TA — activator of protein X/transcription; RH — RNase H; RT — reverse transcriptase; SU — surface glycoprotein; TM — transmembrane glycoprotein; TP — RNase H; TT/SR — translational transactivator/suppressor of RNA interference; VAP — virion-associated protein (adapted from M. Krupovic et al. 2018. J. Virol 92:e00515-18)

Рис. 2. Филогения обратных транскриптаз ретровирусов и параретровирусов: env — гены оболочки; gag — группоспецифический антиген; IN — интеграза; LTR — длинный терминальный повтор; MA — матричный белок; MP — белок движения; NC — нуклеокапсид; nef tat, rev, vif, vpr, vpu — гены вируса иммунодефицита человека 1-го типа, которые экспрессируют регуляторные белки через мРНК; P — полимераза; pol — гены полимеразы; PR — протеаза; PreS — предповерхностный белок (оболочечный); PX/TA — активатор белка X/транскрипции; RH — RNase H; RT — обратная транскриптаза; SU — поверхностный гликопротеин; TM — трансмембранный гликопротеин; TP — РНКаза H; TT/SR — трансляционный трансактиватор/супрессор РНК-интерференции; VAP — вирион-ассоциированный белок (адаптировано из M. Krupovic et al. 2018. J. Virol 92:e00515-18)

which is inserted into the host DNA to form provirus and retroelements. Retroelements and protoviruses acquire new cellular gene sequences encoding ribonuclease H (RNase), integrase, and structural genes and become infectious viral particles. Phylogenetic comparisons show that the reverse transcriptases of group II self-spliced introns in these unicellular organisms and the long interspersed nuclear elements (LINE) of eukaryotic cells are the evolutionary ancestors of retroviral reverse transcriptase. Comparison analysis of retroviral genome sequences shows extensive shuffling with the genomes of other viruses, with host cell genomes and their transposable genetic elements. The genomes of eukaryotic pararetroviruses, vertebrate *Hepadnaviridae,* and plant *Caulimoviridae* are also replicated by reverse transcription. However, these viruses insert DNA into the capsid, which does not need integration into the host genome for the virus to replicate. Hepadnaviruses also lack several of the characteristic features of retroviruses and other pararetroviruses. Phylogenetic comparison of all known viruses with reverse transcription places them outside the order *Ortervirales* (Fig. 2) [19].

However, in the origin of DNA viruses, the following features stand out. Thus, sequence comparison shows that

eukaryotic viruses containing ssDNA probably evolved from bacterial plasmids after they acquired the genes for capsid proteins from the (+) chain of RNA viruses. How this occurred is unknown. Possibly, some ssDNAs, such as parvoviruses, may have arisen before the formation of major eukaryotic kingdoms. Other virus lineages appear to have a later history. Genome comparison shows that dsDNA viruses, such as polyomaviruses and papillomaviruses, most likely arose from osDNA viruses through the encapsidation of a DNA replication intermediate product. Clearly, the extensive gene shuffling between functional modules from different groups of plasmids and viruses makes it difficult to trace the ancestors of these viruses. Nevertheless, genomic studies have supported a general scheme of origin for the main eukaryotic osDNA and dsDNA virus lineages, which include the acquisition of genes for various replication-related enzymes and structural proteins. The genome sizes of eukaryotic dsDNA viruses vary considerably. Nuclear cytoplasmic dsDNA viruses are the largest and most common group. This group includes vertebrate viruses, such as poxviruses, aspharviruses, pridoviruses, invertebrate ascoviruses, and viruses that infect amoebas and other protists [20, 21].

Gene set analysis shows that large viruses probably evolved from smaller viruses, obtaining genes from various sources, including host cells by gene duplication. These same studies have suggested that megaviruses and vertebrate herpesviruses belong to two separate evolutionary branches that arose from two unrelated families of viruses that infect bacteria. The origin of megaviruses can be traced to members of the bacteriophage family *Tectiviridae.* In genomic reconstructions of the megavirus branch, a Tectivirus-like ancestor enters a primitive eukaryotic host cell in the bacterial endosymbiont. This virus subsequently forms large DNA transposons (15 000–20 000 base pairs; kilobases, [kb]) called polyntons. These are large and complexly arranged DNA transposons, which were discovered in the mid-2000s. One polynton encodes up to 10 different proteins, which are common in many unicellular and multicellular eukaryotes. Polyntons encode a protein-primed DNA polymerase and a retrovirus-like integrase, from which their name is derived. Most of them also encode an adenovirus-like cysteine protease and icosahedral viral capsid proteins [22–24].

Polyntons are located in the nucleus; thus, the appearance of cytoplasmic plasmids and viruses from these precursors is likely related to transcriptional mechanisms. In addition to megaviruses and adenoviruses, polyntons are probable precursors of bidnaviruses and virophages. The evolutionary origin of some large dsDNA viruses, such as baculoviruses and other insect viruses, remains unknown, despite reports that many of them represent a distant offshoot of megaviruses. A probable ancestor *of Herpesvirales* is a caudate bacteriophage with an icosahedral head belonging to the order *Caudovirales*, and it is the most common virus on Earth (Fig. 2) [25].

Figure 2 shows that the phylogenetic tree is constructed based on the analysis of all reverse-transcribing viruses and long terminal repeat (LTR) retrotransposons that form virus-like particles recognized by the International Committee on Virus Taxonomy. The scheme includes sequences of nonviral, bacterial group II introns and eukaryotic long embedded sequences (non-LTR LINEs; gray triangle). The genomic DNA organizations of the encoded proteins are adjacent to the corresponding branches. Families in the order *Ortervirales* are indicated by blue triangles, with sizes proportional to the number of members. *Belpaoviridae* include retrotransposons Bel and Poa LTR, *Metaviridae* include the yeast Ty3, and *Pseudoviridae* include the yeast Tyl. *Hepadnaviridae* pararetroviruses are indicated by a red triangle, and a representative circular hepatitis B virus genome is shown as a 6.6-kDa linear protein.

The history of the discovery of giant viruses is interesting. For example, during an outbreak of pneumonia in Bradford, England, in 1992, the largest virus was isolated at the time. Researchers isolated pathogenic amoebas from cooling towers of hospitals, which turned out to be gram-positive bacteria and did not infect people. Accordingly, they had nothing to do with the disease. Electron microscopy of *Acanthamoeba polyphaga* infected with this agent detected 400-nm icosahedral viral particles in the cytoplasm. Mature particles are surrounded by an abundance of fibers whose bases form the outer protein capsid. The virus was named mimivirus because it mimicked a microbe. It is now called *Acanthamoeba polyphaga mimivirus.* Ten years later, two other amoeba pathogens were discovered: one was isolated from marine sediments off the coast of central Chile, and the other from a freshwater pond near Melbourne. These new giant viral particles bore no morphological or genomic resemblance to any previously known virus families. They were suggested to be the first members of the new genus *Pandoravirus*. A fourth giant virus was discovered in 2014 in a sample from permafrost in Siberia, and it was more than 30 000 years old. Amazingly, this virus could still infect cultivated amoebas. This ancient virus was somewhat similar to pandoraviruses; however, the replication cycle and genomic features were similar to those of icosahedral DNA viruses. The genomes of these viruses contained a large proportion (21.2%) of multiple, regularly interspersed copies of two nucleotide bases (n.o.)

in tandem arrays of the conserved palindrome. These giant *Pithovirus sibericum* viruses of amoebae and protists are now widely distributed [26, 27].

Despite the unknown original evolution of viruses, analysis of NK sequences has revealed many evolutionary relationships between modern viruses, their precursors, and their hosts. Nevertheless, many additional ecological niches have yet to be explored to clarify or adjust evolutionary branches. In addition, future searches for the sources of life on Earth will likely stimulate new or refined hypotheses about the origin of viruses and their place in the biosphere.

Coevolution with the host is characteristic of small eukaryote DNA viruses, parvoviruses, polyomaviruses, and papillomaviruses. Evidence for coevolution comes from the discovery of a close association of viral DNA sequences with a particular host group. This relationship was particularly striking in the comparison of the distributions of human papillomaviruses of types 16 and 18. Variants of each type were associated with a specific racial and geographic distribution in the human population. Another example is human polyomavirus JC, which causes severe brain oligodendrocyte damage, often with a fatal outcome. Five genotypes have been identified in this virus, which have been identified in the USA, Africa, some European countries, and Asia. Sequence analysis of the NC subtypes of JC viruses shows that they not only evolved with humans but also in specific human subgroups [28].

The question of how can the evolution of viruses be related to specific human populations arises The unusual pathogenesis of human papillomaviruses partially answers this question. Infection of basal adult skin keratinocytes leads to viral multiplication that coordinated with cell differentiation, and the assembly of viral particles occurs only when the cells undergo terminal differentiation in the superficial skin layers. In this case, reactivation and persistence of the virus in pregnancy contribute to a very high probability of vertical transmission. This mode of transmission is the predominant mechanism for papillomaviruses [29].

The three main subfamilies, namely, *Alphahesvirinae, Betaherpesvirinae, and Gammaherpesvirinae*, of viruses in the family *Herpesviridae are now* easily distinguished by their genome sequence features. However, the original taxonomic division of these families was based on common, often quite, arbitrary biological characteristics. Studies have revealed an evolutionary scale of changes in the genome of herpesviruses comparable to that of the hosts. Therefore, the earliest herpesvirus entered the body of the progenitor of modern humans. Subsequent viruses evolved by coevolution with their hosts. This conclusion is confirmed by the sequencing of the genomes of all members of the three main subfamilies. Their genomes contained a basic block of genes, often organized into similar clusters in the genome. The three main groups of herpesviruses are believed to have originated between 180 and 220 million years ago. That is, these three subfamilies must have existed long before mammals, which appeared on Earth approximately 60–80 million years ago. The herpesviruses of fish, oysters, and amphibians have nearly identical architecture but little or no homology to the genome sequences of the herpesviruses of the major subfamilies. They most likely represent a very early branch of this ancient family [30].

In contrast to the large genomes of DNA viruses, RNA virus genomes are small and contain few genes that share characteristics with the host cell genome and that can be used in correlating virus and host evolutionary processes. However, by comparing the nucleotide sequences of many (+) and (–) chains of viral RNA genomes, blocks of genes encoding proteins with similar functions can be identified. Thus, common coding strategies are deduced, also suggesting a common pedigree. A feature of the sequences of the genomes of many (–) strand RNA viruses is a limited number of genes encoding 4–13 proteins. Most of these proteins can be classified into three functional classes, i.e., core proteins that interact with the RNA genome, envelope glycoproteins that are required for adhesion and penetration of viral particles, and polymerases that are required for replication and mRNA synthesis (Fig. 3) [31].

Figure 3 shows that the gene maps *of Rhabdoviridae, Paramyxoviridae, Bunyaviridae, Arenaviridae*, and *Orthomyxoviridae* illustrate the similarity of gene products. The individual gene segments *of Orthomyxoviridae are* arranged according to functional similarity with the other two groups of segmented viruses. Within a given genome, the genes must be scaled. For segmented genomes, genes with blue outlines are those that encode multiple proteins from different open-reading frames. Genes with red outlines are ambisens (arrow).

Studies of the ecological and biological characteristics of influenza viruses have shown that the same virus population can infect many different host species. In each new host, influenza viruses undergo selection processes necessary for their reproduction and spread. Thus, the gene pool of this (–) chain RNA virus is enormous, with powerful changes in genetic information. Large-scale genome sequencing has provided information on the state of the viral gene pool of influenza viruses during its transmission from humans to animals, from animals to humans, and from humans to humans. Hemann et al. [32] sequenced genomes from over 200 influenza virus isolates and studied nearly three million

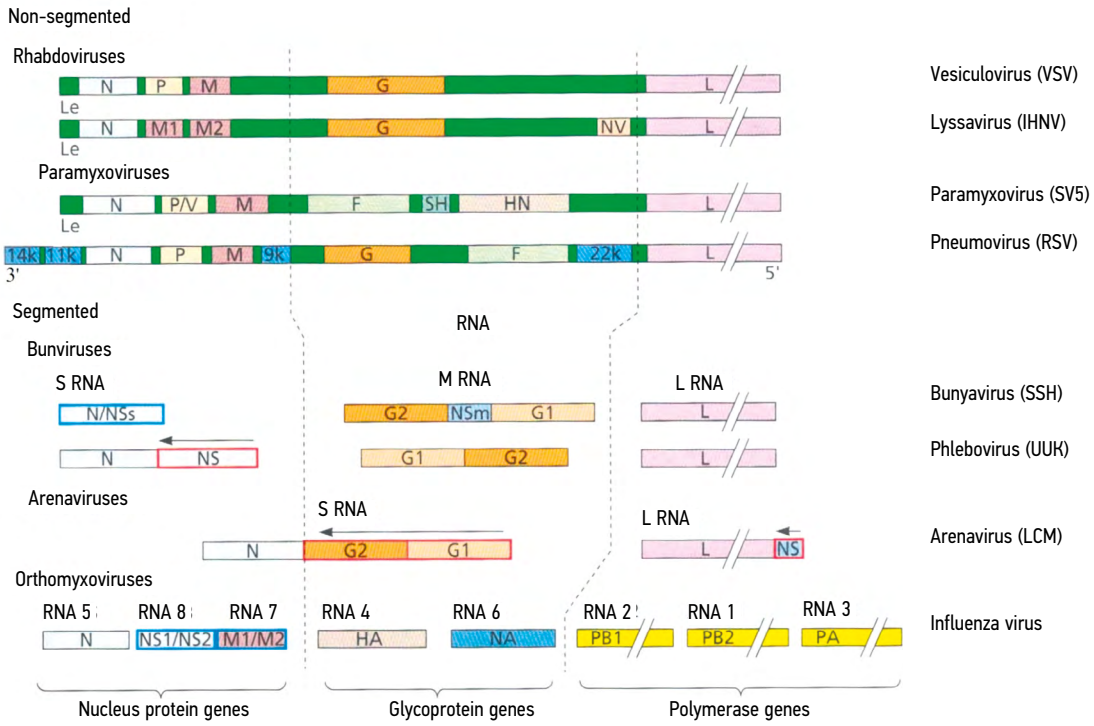

Fig. 3. Genetic maps of selected (–) strand RNA viral genomes (on J. Flint, V. Racaniello, G. Rall et al. Principles of Virology. Fifth edition. Vol. II. 2020)
Рис. 3. Генетические карты выделенных (–) нить РНК-вирусов (по Дж. Флинту, В. Раканьелло, Г. Ралла и др. Принципы вирусологии. 5-е изд. Т. II. 2020)

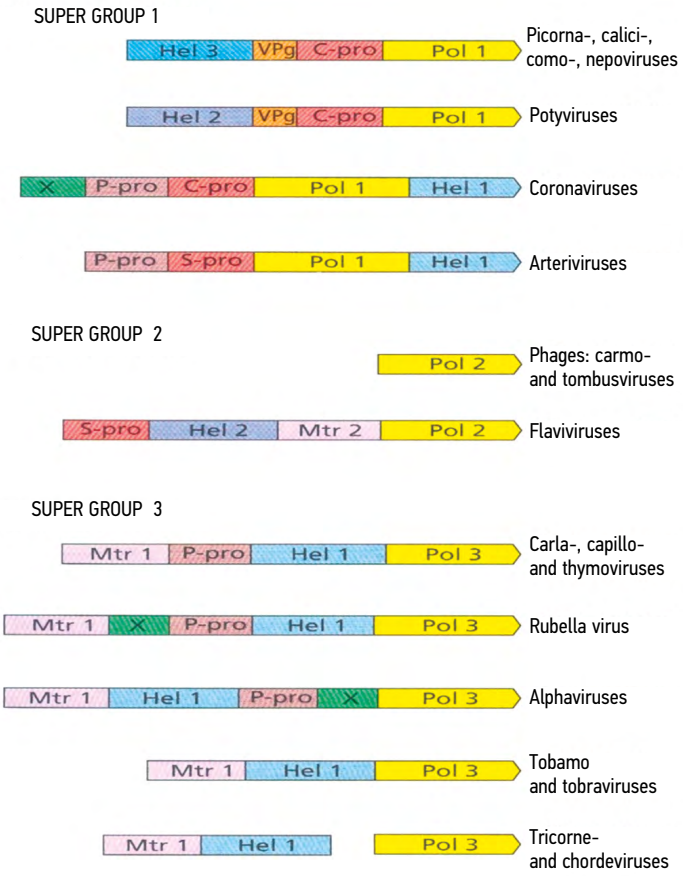

Fig. 4. RNA virus genomes and evolution (on J. Flint, V. Racaniello, G. Rall et al. Principles of Virology. Fifth edition. Vol. II. 2020)
Рис. 4. Эволюция геномов РНК-вирусов (по Дж. Флинту, В. Раканьелло, Г. Ралла и др. Принципы вирусологии. 5-е изд. Т. II. 2020)

nucleotide sequences. The most important conclusion was that this population of influenza viruses contains several lineages at any given time. In addition, alternate minor lineages exchange information with the dominant lineage. As conditions change, the number of immune mutants can either increase or decrease. The number of mutants tropic to host cell receptors also changes. Important clues to understanding the epidemiological features of influenza viruses were obtained by sequencing and analyzing the genomes of over 1 300 isolates from various locations. The viral genome changed as a result of frequent recombination of genes with the appearance of subtypes with other antigens and different dynamics. However, all of them obeyed the classical viral model of distribution of influenza viruses.

The largest group of viruses form the (+) chain RNA viruses (except for retroviruses). Just like (–) chain RNA viruses, they encode 3–12 proteins. Most proteins can be also subdivided into three functional groups, although the organization of their genomes is not necessarily identical. Genomes of (+) chain RNA viruses contain several genes that provide replicative functions. Over time, the genes have been mixed and matched in selected combinations. Helicase, genome-associated proteins, chymotrypsin-like proteases, polymerases, papain-like protease I, and methyltransferases provided replicative functions. Differences in the polymerase gene distinguished three supergroups. Each of these supergroups contained viruses infecting various animals and plants. These viruses could infect hosts in all branches of evolutionary periods (Fig. 4) [5].

Figure 4 shows that the genomes of (+) strand RNA viruses contain several genes of replicative functions that are mixed and matched in selected combinations over time. These functions include helicase (Hel), genome-associated protein (VPg), chymotrypsin-like protease (C- or S-pro), polymerase (Pol), papain-like protease I (P-pro), methyltransferase (Mtr), and region of unknown function (X). Differences in the polymerase gene define the three supergroups. In the figure, the genes are not scaled, and the structural proteins are omitted.

Host genes associated with antiviral defense were also selected. Individuals encoding ineffective alleles died from viral infections and were eliminated from the population. As a result, viruses with compensatory mutations emerged. Constitutively expressed host cell genes encode antiviral proteins with cell-autonomous functions. These proteins contribute to internal cellular defense by inhibiting the multiplication of infecting viruses at various stages. These include the APOBEC3 family of cytidine deaminases, which induce mutations in viral DNA, the tripartite motif proteins that interact with the capsids of infecting retroviruses, and the membrane protein tetterin. Points of interaction between host antiviral proteins and viral proteins can be identified by studying orthologous genes in related host species whose codons exhibit a higher proportion of nucleotide substitutions (dN substitution). Unchanged codons are called synonymous codons (dS). A dN/dS ratio of > 1 at a particular site indicates a high probability of positive selection. A study of numerous cellular proteins interacting with viruses showed that > 30% of codon changes in human and mammalian genes occur [33].

The dimeric transferrin receptor protein (TFR1) located on the cell surface controls various functions of living cells. This protein is also a receptor for various viruses. The opposing functions of this protein, namely, avoidance of infection and maintenance of iron uptake function, were apparently balanced during the evolution of the *TFR1* gene. Analysis of this gene in several evolutionarily related rodent species showed that although most of the amino acids in the encoded proteins were conserved, several residues were quite variable, with dN/dS of > 1, indicating a high probability of positive selection. Structures of the human ectodomain TFR1 are included in the binding sites of arenavirus and mouse breast tumor virus but separated from the transferrin-binding site. An experiment revealed natural substitutions of these residues to block virus entry while preserving TFR1 iron uptake in both rodents and humans [34].

Modern paleovirusology has made it possible to establish traces of virus sequences that are relics of ancient viral infections in many animal species. Endogenous retroviral sequences are detected in 6%–14% of vertebrate genomes. In humans, these endogenous sequences account for up to 8% of genomic DNA. Endogenous retroviral sequences are a consequence of the integration of proviral DNA into the genomes of host germline cells, which are passed on to subsequent generations. Consequently, a comparison of orthologous endogenous retroviral sequences of modern vertebrate species and knowledge of species evolution may allow an assessment of when common viral sequences were inserted into the ancestral host germline. For example, most endogenous human proviruses were between 10 and 50 million years old and were derived from retroviruses circulating on Earth long before *Homo sapiens* appeared. The *Spumaretrovirinae* subfamily is approximately 460–550 million years old, i.e., they belong to the Paleozoic era. In 2012, phylogenetic analysis of the coelacanth endogenous foamy-like virus (CoeEFV) element in the genome *of Latimeria chalumnae* was presented, which suggested the ancient oceanic origin of foamy viruses. Apparently, these viruses accompanied their vertebrate hosts

in the evolutionary transition from water to land more than 400 million years ago. Using CoeEFV reverse transcriptase *as a* probe, additional foamy virus-like sequences were identified in the genomes of amphibians, fish, salamanders, frogs, rayfish, spiny finfish, and sharks in 2017. Sequence differences in LTRs suggested that the mutation rate corresponds to the rate of mutational changes in the host genome. As a result, retroviruses appeared together with their vertebrate hosts in the ocean approximately 460–550 million years ago, at the beginning of the Paleozoic era, if not earlier [28, 31].

Most endogenous retroviruses are defective. However, some retain functional genes, and if their transcription is not suppressed, they can reappear as infectious agents. For example, the high incidence of spontaneous leukogenesis in AKR mice was associated with the formation of replication-competent leukemic viruses resulting from recombination between the genomes of three different endogenous mouse retroviruses. Such endogenous sequences can also serve as genetic reservoirs for recombination with exogenously infecting viruses. Endogenous retroviral sequences can have pronounced effects on the evolution and function of their host genome. For example, recombination between sequences integrated at different loci can explain several large-scale deletions, duplications, and other chromosomal rearrangements that occurred during genome evolution in primates. In humans, such recombination has contributed to the extensive duplication of gene blocks that make up the major locus of major histocompatibility complex class I. The diversity that has arisen in such duplications and the heterozygosity of the locus give a strong selective advantage against circulating pathogens in humans [35].

Transcription processes lead to the fact that host cell genes repurpose endogenous proviruses to perform their new functions. Thus, resistance to retroviral infection in some mouse strains is ensured by the expression of endogenous sequences linked to the *Fv-1* retroviral capsid gene. Another *Ty3/gypsy* capsid gene, the ancient transmissible retrotransposon LTR, is the progenitor of the mammalian gene *Arc*, which is expressed in the brain and required for information storage [17].

Unfortunately, the age of other RNA viruses remains under-investigated. However, sequences associated with other RNA viruses have been found in host genomes. These sequences were found to be associated with flaviviruses and picornaviruses and were included in plant and insect genomes. Sequences of the genomes of known RNA viruses, when compared with the library of vertebrate genomes, were found to be between 30 and 40 million years old. Among them were viruses causing severe hemorrhagic fevers [16].

The study of evolutionary changes in DNA viruses has also revealed ancient associations of circoviruses, parvoviruses with vertebrate genomes. Some insertions are more than 50 million years old; however, others occurred more recently. These viruses have tiny genomes that encode only two reading frames: rep and cap. Host enzymes recruit proteins encoded by the *rep* gene (Rep 78/68) to the hairpin region of the viral genomes, where their synthesis is initiated. Sequences of parvoviruses and adeno-associated virus DNA were inserted into the host genome, which are recognized by the viral protein Rep. In the absence of a helper virus, integrated parvovirus genomes can go into a latent state, which is activated when the host cell is subsequently infected by the helper virus. Therefore, germline insertions likely lead to the formation of endogenous sequences at different times during their host evolution and are the result of random copying of circovirus and parvovirus DNA in the loci of their host genomes that resemble the hairpin regions of viral replication [13].

Another mechanism explains the presence of endogenous hepadnavirus genome sequences in some bird species. Modern hepadnavirus genomes can reside as minichromosomes in the nuclei of host cells and can sometimes be inserted into the host DNA through nonhomologous recombination at random sites. These viruses infect birds and are found to have endogenous sequences associated with the viral dsDNA genome. Phylogenetic analysis of endogenous avian viral sequences shows that birds were ancestral hosts of viruses in the *Hepadnaviridae* family, and this family of viruses is more than 82 million years old. Since endogenous avian hepadnavirus sequences are not found in mammalian genomes, mammalian hepadnaviruses were suggested to appear much later, after an interspecies exchange between birds and mammals [36].

Unfortunately, it is still unclear why viral NCs integrated into the eukaryotic host DNA millions of years ago are not very different from modern NCs. Genome replication error rates of currently circulating RNA viruses and some small octDNA viruses are quite high (10^{-2} to 10^{-5} mutations per site/replication round) compared with host DNA (10^{-7}–10^{-9} nucleotide substitutions). When comparing viral sequences, measuring the rate of change over several years in a particular host, these rates of some viral genomes were consistent with the original estimates. Therefore, many of these viral lines originated relatively recently in evolutionary time, on the order of hundreds or thousands of years ago. However, ancient endogenous sequences associated with some octDNA and RNA viruses other than retroviruses were

found to have circulated in various hosts many millions of years ago. For example, two independent phylogenetic analyses of currently circulating filoviruses (including Ebola and Marburg viruses) have shown that this family was 10 000–150 000 years old, whereas paleovirus analyses were over 40 million years old. The discrepancy between such estimates is even greater for circoviruses, ranging from < 500 years to over 40 million years since the identification and analysis of related endogenous viral sequences. Thanks to paleovirus analysis, viral genomes have been established to change much more slowly [27].

CONCLUSION

The relationship between viruses and their hosts is constantly changing. Therefore, combined efforts of evolutionary biologists, ecologists, and virologists are needed to better understand the dynamics of evolutionary viral change. Unfortunately, despite great advances in the study of evolutionary mechanisms, the interaction between virus populations and their hosts remains largely unknown. Clearly, the rapid reproduction of viruses with the formation of numerous offspring, adaptation to changes in host populations, and ability for enormous genetic diversity ensure their survival. However, the state of the host, including its survival, depends on the state of the evolutionary mechanisms of innate and adaptive immunity, which can recognize and then block the reproduction or destroy the invading viruses. Modern hosts represent the offspring of survivors of ancient infections. The review presented herein considered the evolution of viruses and hosts mainly in the context of currently circulating populations to which access is available. Therefore, except for retroviruses, no information has been revealed about other viral families to estimate how old they may be or how they may have changed over evolutionary periods. Expanding evidence of changes in virus DNA over evolutionary time is coupled with changes in the gene pool of surviving hosts. Unfortunately, until now, many "intimate" mechanisms of virus evolution and variability have not been fully studied; thus, we cannot give an unambiguous answer or single out any standard patterns of viral genome changes. Although not many mechanisms and schemes are known, they are influenced by numerous environmental factors and peculiarities of host cell biology, making the results of variability unpredictable. In the host cell, the genetic programs of the cell and the virus collide, which also explains the nature of viruses. Thus, viruses are submicroscopic supramolecular noncellular life forms that can reproduce only in the cell. Parasitologists consider viruses to be intracellular genetic parasites, biochemists consider them to be protein–nucleic complexes, and molecular geneticists consider them mobile genetic elements. Based on what is currently known, we can identify the following promising directions: the development of antiviral drugs, using viruses as expression vectors, and the creation of potentially the most effective live-weakened vaccines. In general, viral infections are accompanied by serious consequences ranging from host IS activation and adaptation to the death of entire viral populations. However, given the constantly changing viral populations, radical modifications of Earth's ecosystems, and changes in host IS, including humans, it is very difficult to understand which viruses will prevail.

REFERENCES

1. Katze MG, Korth MJ, Law GL, et al. *Viral pathogenesis: From basics to systems biology*. San Diego: Academic Press, 2016. 422 p.

2. Ahmad L, Mostowy S, Sancho-Shimizu S. Autophagy-virus interplay: From cell biology to human disease. *Front Cell Dev Biol*. 2018;19:155. DOI: 10.3389/fcell.2018.00155

3. Luoa LY, Hahnb WC. Oncogenic signaling adaptor proteins. *J Genet Genomics*. 2015;42(10):521–529. DOI: 10.1016/j.jgg.2015.09.001

4. Griffin DE. The immune response in measles: Virus control, clearance and protective immunity. *Viruses*. 2016;10(8):282–291. DOI: 10.3390/v8100282

5. Gong B-L, Mao R-Q, Xiao Y, et al. Improvement of enzyme activity and soluble expression of an alkaline protease isolated from oil-polluted mud flat metagenome by random mutagenesis. *Enzyme Microb Technol*. 2017;106:97–105. DOI: 10.1016/j.enzmictec.2017.06.015

6. Domingo E, Perales C. Quasispecies and virus. *Eur Biophys J*. 2018;4(47):443–457. DOI:10.1007/s00249-018-1282-6

7. Guo Y-J, Pan W-W, Liu S-B, et al. ERK/MAPK signaling pathway and tumorigenesis. *Exp Ther Med*. 2020;19(3):1997–2007. DOI: 10.3892/etm.2020.8454

8. Takata MA, Gonçalves-Carneiro D, Zang TM, et al. CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature*. 2017;550(7674):124–127. DOI: 10.1038/nature24039

9. Thapa RJ, Ingram JP, Ragan KB, et al. DAI senses influenza a virus genomic RNA and activates RIPK3-Dependent cell death. *Cell Host Microbe*. 2016;20(5):674–681. DOI: 10.1016/j.chom.2016.09.014

10. Hemann EA, Green R, Turnbull JB, et al. Interferon-λ modulates dendritic cells to facilitate T cell immunity ion with influenza A virus. *Nat Immunol*. 2019;20:1035–1045. DOI: 10.1038/s41590-019-0408-z

11. Stecca B, Rovida E. Impact of ERK5 on the hallmarks of cancer. *Int J Mol Sci*. 2019;20(6):1426. DOI: 10.3390/ijms20061426

12. Yang L, Shi P, Zhao G, et al. Targeting cancer stem cell pathways for cancer therapy. *Signal Transduct Target Ther*. 2020;5(8):8. DOI: 10.1038/s41392-020-0110-5

13. Burrell C, Howard C, Murphy F. *Fenner and White's medical virology. 5th edition*. San Diego: Academic Press, 2016. 454 p.

14. Nash A, Dalziel R, Fitzgerald J. *Mims' pathogenesis of infectious disease. 6th edition*. San Diego: Academic Press; 2015. 348 p.

15. Maillard PV, van der Veen AG, Poirier EZ, et al. Slicing and dicing viruses: antiviral RNA interference in mammals. *EMBO J*. 2019;38(8):e100941. DOI: 10.15252/embj.2018100941

16. Hayward A. Origin of the retroviruses: when, where, and how? *Curr Opin Virol*. 2017;25:23–27. DOI:10.1016/j.coviro.2017.06.006

17. Krupovic M, Koonin EV. Multiple origins of viral capsid proteins from cellular ancestors. *PNAS USA*. 2017;114(12):E2401–E2410. DOI: 10.1073/pnas.1621061114

18. Lee S, Liu H, Wilen CB, et al. A secreted viral nonstructural protein deters intestinal norovirus pathogenesis. *Cell Host Microbe*. 2019;25(6):179–187. DOI: 10.1016/j.chom.2019.04.005845–857

19. Horie M. The biological significance of bornavirus-derived genes in mammals. *Curr Opin Virol*. 2017;25:1–6. DOI: 10.1016/j.coviro.2017.06.004

20. Hadjidj R, Badis A, Mechri S, et al. Purification, biochemical, and molecular characterization of novel protease from Bacillus licheniformis strain K7A. *Int J Biol Macromol*. 2018;114:1033–1048. DOI: 10.1016/j.ijbiomac.2018.03.167

21. Jeong YJ, Baek SC, Kim H. Cloning and characterization of a novel intracellular serine protease (IspK) from *Bacillus megaterium* with a potential additive for detergents. *Int J Biol Macromol*. 2018;108:808–816. DOI: 10.1016/j.ijbiomac.2017.10.173

22. Ashraf NM, Krishnagopal A, Hussain A, et al. Engineering of serine protease for improved thermostability and catalytic activity using rational design. *Int J Biol Macromol*. 2019;126:229–237. DOI: 10.1016/j.ijbiomac.2018.12.218

23. Ashraf NM, Krishnagopal A, Hussain A, et al. Engineering of serine protease for improved thermo stability and catalytic activity using rational design. *Int J Biol Macromol*. 2019;126:229–237. DOI: 10.1016/j.ijbiomac.2018

24. Ho SYW, Lanfear R, Bromham L, et al. Time-dependent rates of molecular evolution. *Mol Ecol*. 2011;20(15):3087–3101. DOI: 10.1111/j.1365-294X.2011.05178.x

25. Katzourakis A, Gifford RJ. Endogenous viral elements in animal genomes. *PLoS Genet*. 2010;11(6):e1001191. DOI: 10.1371/journal.pgen.1001191

26. Aiewsakun P, Katzourakis A. Endogenous viruses: Connecting recent and ancient viral evolution. *Virology*. 2015;479-480:26–37. DOI: 10.1016/j.virol.2015.02.011

27. Parrish NF, Tomonaga K. Endogenized viral sequences in mammals. *Curr Opin Microbiol*. 2016;31:176–183. DOI: 10.1016/j.mib.2016.03.002

28. Frank JA, Feschotte C. Co-option of endogenous viral sequences for host cell function. *Curr Opin Virol*. 2017;25:81–89. DOI: 10.1016/j.coviro.2017.07.021

29. Garcia-Sastre A. Ten strategies of interferon evasion by viruses. *Cell Host Microbe*. 2017;22(2):176–184. DOI: 10.1016/j.chom.2017.07.012

30. Diner BA, Lum KK, Javitt A, et al. Interactions of the antiviral factor interferon gamma-inducible protein 16. NIFI16 mediate immune signaling and herpes simplex virus-1 immunosuppression. *Mol Cell Proteomics*. 2015;14(9):2341–2356. DOI: 10.1074/mcp.M114.047068

31. Stoye JP. Studies of endogenous retroviruses reveal a continuing evolutionary saga. *Nat Rev Microbiol*. 2012;6(10):395–406. DOI: 10.1038/nrmicro2783

32. Hemann EA, Green R, Turnbull JB, et al. Interferon-λ modulates dendritic cells to facilitate T cell immunity ion with influenza A virus. *Nat Immunol*. 2019;20:1035–1045. DOI: 10.1038/s41590-019-0408-z

33. Enard D, Cai L, Gwennap C, Petrov DA. Viruses are a dominant driver of protein adaptation in mammals. *Elife*. 2016;5:e12469. DOI: 10.7554/eLife.12469

34. Xu X, Zhang M, Xu F, Jiang S. Wnt signaling in breast cancer: biological mechanisms, challenges and opportunities. *Mol Cancer*. 2020;19(165):165. DOI: 10.1186/s12943-020-01276-5

35. Wang B, Li X, Liu L, Wang M. β-Catenin: oncogenic role and therapeutic target in cervical cancer. *Biol Res*. 2020;53:33. DOI: 10.1186/s40659-020-00301-7

36. Ma Z, Damania B. The cGAS-STING defense pathway and its counteraction by viruses. *Cell Host Microbe*. 2016;19(2):150–158. DOI: 10.1016/j.chom.2016.01.010

СПИСОК ЛИТЕРАТУРЫ

1. Katze M.G., Korth M.J., Law G.L., et al. Viral pathogenesis: From basics to systems biology. San Diego: Academic Press, 2016. 422 p.

2. Ahmad L., Mostowy S., Sancho-Shimizu S. Autophagy-virus interplay: From cell biology to human disease // Front Cell Dev Biol. 2018. Vol. 19. ID 155. DOI: 10.3389/fcell.2018.00155

3. Luoa L.Y., Hahnb W.C. Oncogenic signaling adaptor proteins // J Genet Genomics. 2015. Vol. 42, No. 10. P. 521–529. DOI: 10.1016/j.jgg.2015.09.001

4. Griffin D.E. The immune response in measles: Virus control, clearance and protective immunity // Viruses. 2016. Vol. 10, No. 8. P. 282–291. DOI: 10.3390/v8100282

5. Gong B.-L., Mao R.-Q., Xiao Y., et al. Improvement of enzyme activity and soluble expression of an alkaline protease isolated from oil-polluted mud flat metagenome by random mutagenesis // Enzyme Microb Technol. 2017. Vol. 106. P. 97–105. DOI: 10.1016/j.enzmictec.2017.06.015

6. Domingo E., Perales C. Quasispecies and virus // Eur Biophys J. 2018. Vol. 4, No. 47. P. 443–457. DOI:10.1007/s00249-018-1282-6

7. Guo Y.-J., Pan W.-W., Liu S.-B., et al. ERK/MAPK signaling pathway and tumorigenesis // Exp Ther Med. 2020. Vol. 19, No. 3. P. 1997–2007. DOI: 10.3892/etm.2020.8454

8. Takata M.A., Gonçalves-Carneiro D., Zang T.M., et al. CG dinucleotide suppression enables antiviral defence targeting non-self RNA // Nature. 2017. Vol. 550, No. 7674. P. 124–127. DOI: 10.1038/nature24039

9. Thapa R.J., Ingram J.P., Ragan K.B., et al. DAI senses influenza a virus genomic RNA and activates RIPK3-Dependent cell death // Cell Host Microbe. 2016. Vol. 20, No. 5. P. 674–681. DOI: 10.1016/j.chom.2016.09.014

10. Hemann E.A., Green R., Turnbull J.B., et al. Interferon-λ modulates dendritic cells to facilitate T cell immunity ion with influenza A virus // Nat Immunol. 2019. Vol. 20. P. 1035–1045. DOI: 10.1038/s41590-019-0408-z

11. Stecca B., Rovida E. Impact of ERK5 on the hallmarks of cancer // Int J Mol Sci. 2019. Vol. 20, No. 6. ID 1426. DOI: 10.3390/ijms20061426

12. Yang L., Shi P., Zhao G., et al. Targeting cancer stem cell pathways for cancer therapy // Signal Transduct Target Ther. 2020. Vol. 5, No. 8. ID 8. DOI: 10.1038/s41392-020-0110-5

13. Burrell C., Howard C., Murphy F. Fenner and White's medical virology. 5th edition. San Diego: Academic Press, 2016. 454 p.

14. Nash A., Dalziel R., Fitzgerald J. Mims' pathogenesis of infectious disease. 6th edition. San Diego: Academic Press, 2015. 348 p.

15. Maillard P.V., van der Veen A.G., Poirier E.Z., et al. Slicing and dicing viruses: antiviral RNA interference in mammals // EMBO J. 2019. Vol. 38, No. 8. ID e100941. DOI: 10.15252/embj.2018100941

16. Hayward A. Origin of the retroviruses: when, where, and how? // Curr Opin Virol. 2017. Vol. 25. P. 23–27. DOI:10.1016/j.coviro.2017.06.006

17. Krupovic M., Koonin E.V. Multiple origins of viral capsid proteins from cellular ancestors // PNAS USA. 2017. Vol. 114, No. 12. P. E2401–E2410. DOI: 10.1073/pnas.1621061114

18. Lee S., Liu H., Wilen C.B., et al. A secreted viral nonstructural protein deters intestinal norovirus pathogenesis // Cell Host Microbe. 2019. Vol. 25, No. 6. P. 179–187. DOI: 10.1016/j.chom.2019.04.005845–857

19. Horie M. The biological significance of bornavirus-derived genes in mammals // Curr Opin Virol. 2017. Vol. 25. P. 1–6. DOI: 10.1016/j.coviro.2017.06.004

20. Hadjidj R., Badis A., Mechri S., et al. Purification, biochemical, and molecular characterization of novel protease from Bacillus licheniformis strain K7A // Int J Biol Macromol. 2018. Vol. 114. P. 1033–1048. DOI: 10.1016/j.ijbiomac.2018.03.167

21. Jeong Y.J., Baek S.C., Kim H. Cloning and characterization of a novel intracellular serine protease (IspK) from *Bacillus megaterium* with a potential additive for detergents // Int J Biol Macromol. 2018. Vol. 108. P. 808–816. DOI: 10.1016/j.ijbiomac.2017.10.173

22. Ashraf N.M., Krishnagopal A., Hussain A., et al. Engineering of serine protease for improved thermostability and catalytic activity using rational design // Int J Biol Macromol. 2019. Vol. 126. P. 229–237. DOI: 10.1016/j.ijbiomac.2018.12.218

23. Ashraf N.M., Krishnagopal A., Hussain A., et al. Engineering of serine protease for improved thermo stability and catalytic activity using rational design // Int J Biol Macromol. 2019. Vol. 126. P. 229–237. DOI: 10.1016/j.ijbiomac.2018

24. Ho S.Y.W., Lanfear R., Bromham L., et al. Time-dependent rates of molecular evolution // Mol Ecol. 2011. Vol. 20, No. 15. P. 3087–3101. DOI: 10.1111/j.1365-294X.2011.05178.x

25. Katzourakis A., Gifford R.J. Endogenous viral elements in animal genomes // PLoS Genet. 2010. Vol. 11, No. 6. ID e1001191. DOI: 10.1371/journal.pgen.1001191

26. Aiewsakun P., Katzourakis A. Endogenous viruses: Connecting recent and ancient viral evolution // Virology. 2015. Vol. 479-480. P. 26–37. DOI: 10.1016/j.virol.2015.02.011

27. Parrish N.F., Tomonaga K. Endogenized viral sequences in mammals // Curr Opin Microbiol. 2016. Vol. 31. P. 176–183. DOI: 10.1016/j.mib.2016.03.002

28. Frank J.A., Feschotte C. Co-option of endogenous viral sequences for host cell function // Curr Opin Virol. 2017. Vol. 25. P. 81–89. DOI: 10.1016/j.coviro.2017.07.021

29. Garcia-Sastre A. Ten strategies of interferon evasion by viruses // Cell Host Microbe. 2017. Vol. 22, No. 2. P. 176–184. DOI: 10.1016/j.chom.2017.07.012

30. Diner B.A., Lum K.K., Javitt A., et al. Interactions of the antiviral factor interferon gamma-inducible protein 16. NIFI16 mediate immune signaling and herpes simplex virus-1 immunosuppression // Mol Cell Proteomics. 2015. Vol. 14, No. 9. P. 2341–2356. DOI: 10.1074/mcp.M114.047068

31. Stoye J.P. Studies of endogenous retroviruses reveal a continuing evolutionary saga // Nat Rev Microbiol. 2012. Vol. 6, No. 10. P. 395–406. DOI: 10.1038/nrmicro2783

32. Hemann E.A., Green R., Turnbull J.B., et al. Interferon-λ modulates dendritic cells to facilitate T cell immunity ion with influenza A virus // Nat Immunol. 2019. Vol. 20. P. 1035–1045. DOI: 10.1038/s41590-019-0408-z

33. Enard D., Cai L., Gwennap C., Petrov D.A. Viruses are a dominant driver of protein adaptation in mammals // Elife. 2016. Vol. 5. ID e12469. DOI: 10.7554/eLife.12469

34. Xu X., Zhang M., Xu F., Jiang S. Wnt signaling in breast cancer: biological mechanisms, challenges and opportunities // Mol Cancer. 2020. Vol. 19, No. 165. ID 165. DOI: 10.1186/s12943-020-01276-5

35. Wang B., Li X., Liu L., Wang M. β-Catenin: oncogenic role and therapeutic target in cervical cancer // Biol Res. 2020. Vol. 53. ID 33. DOI: 10.1186/s40659-020-00301-7

36. Ma Z., Damania B. The cGAS-STING defense pathway and its counteraction by viruses // Cell Host Microbe. 2016. Vol. 19, No. 2. P. 150–158. DOI: 10.1016/j.chom.2016.01.010

AUTHORS INFO

***Alexander V. Moskalev,** MD, Dr. Sci. (Med.), professor;
ORCID: https://orcid.org/0000-0002-3403-3850;
eLibrary SPIN: 8227-2647; e-mail: alexmav195223@yandex.ru

Boris Yu. Gumilevsky, MD, Dr. Sci. (Med.), professor;
Scopus Author ID: 6602391269; Reseacher ID: J-1841-2017;
eLibrary SPIN: 3428-7704

Vasiliy Ya. Apchel, MD, Dr. Sci. (Med.), professor;
ORCID: https://orcid.org/0000-0001-7658-4856;
Scopus Author ID: 6507529350; Researcher ID: E-8190-2019;
Scholar ID: g9EKlssAAAAJ&hl; eLibrary SPIN: 4978-0785

Vasiliy N. Tsygan, MD, Dr. Sci. (Med.), professor;
ORCID: https://orcid.org/0000-0003-1199-0911;
eLibrary SPIN: 7215-6206

ОБ АВТОРАХ

***Александр Витальевич Москалев,** д-р мед. наук, профессор; ORCID: https://orcid.org/0000-0002-3403-3850;
eLibrary SPIN: 8227-2647; e-mail: alexmav195223@yandex.ru

Борис Юрьевич Гумилевский, д-р мед. наук, профессор;
Scopus Author ID: 6602391269; Reseacher ID: J-1841-2017;
eLibrary SPIN: 3428-7704

Василий Яковлевич Апчел, д-р мед. наук, профессор;
ORCID: https://orcid.org/0000-0001-7658-4856;
Scopus Author ID: 6507529350; Researcher ID: E-8190-2019;
Scholar ID: g9EKlssAAAAJ&hl; eLibrary SPIN: 4978-0785

Василий Николаевич Цыган, д-р мед. наук, профессор;
ORCID: https://orcid.org/0000-0003-1199-0911;
eLibrary SPIN: 7215-6206

* Автор, ответственный за переписку / Corresponding author