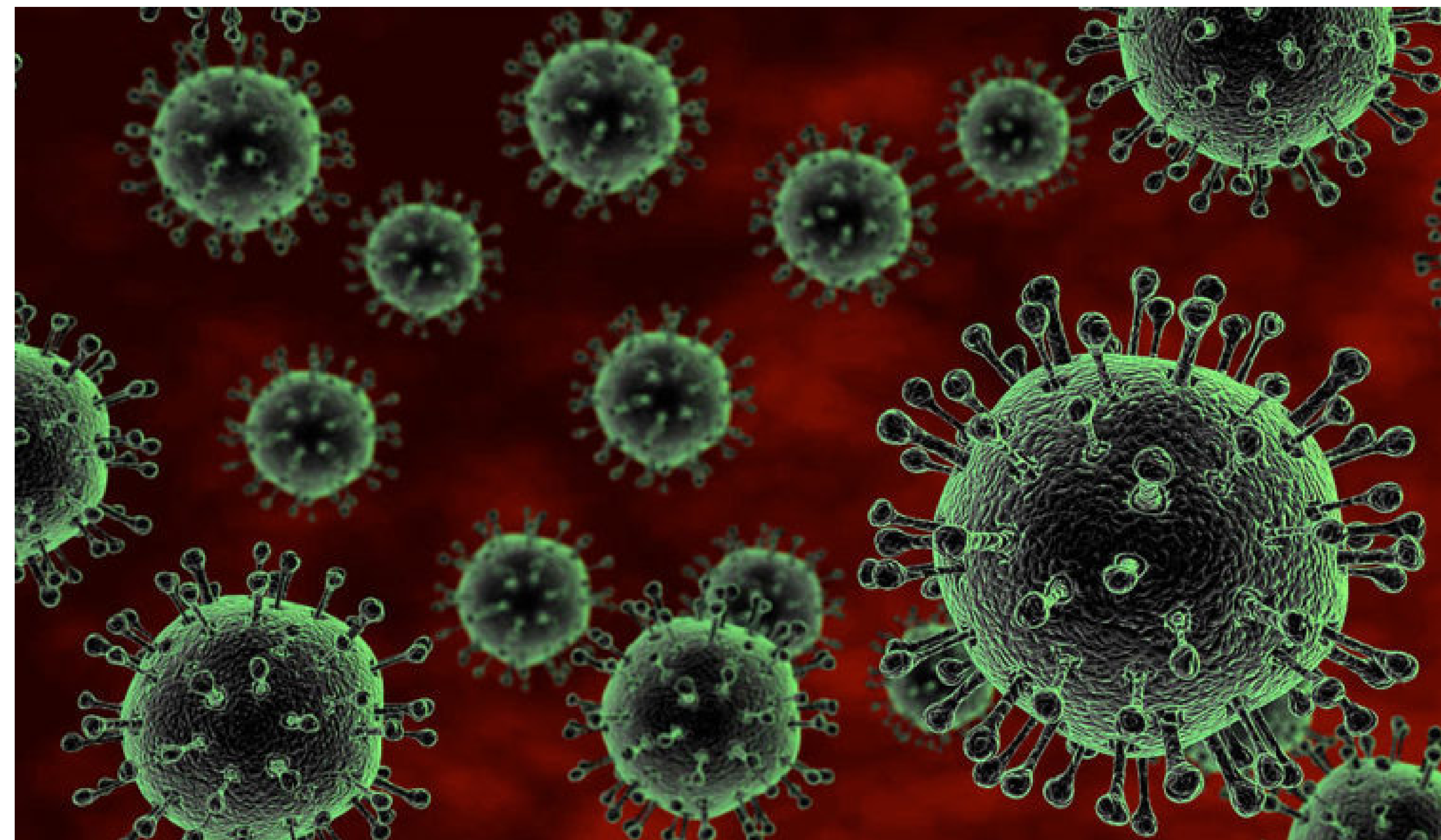
Modern approaches in the field of molecular genetics of viruses in the study of representatives of the Coronaviridae family — Site of Mikhail Supotnitsky

https://www-supotnitskiy-ru.translate.goog/stat/stat154.htm?_x_tr_sl=ru&_x_tr_tl=en&_x_tr_hl=en&_x_tr_pto=sc

---



Received September 1, 2021 Accepted for publication September 23, 2021

The existing suspicions about the artificial origin of the COVID-19 pandemic and about the use of reverse genetics technology to create the SARS-CoV-2 virus require an understanding of its capabilities in constructing new viruses. *The purpose of this work* is to show how the use of reverse genetics allows us to construct previously non-existent

coronaviruses, technologies and major achievements in their creation. For the preparation of this article, information was used that is in the public domain and easily verified from the sources cited. The name of the technology - "reverse genetics" came from the fact that when obtaining RNA viruses capable of reproducing, they do not go from DNA to RNA, as is usually done in a cell during protein synthesis, but vice versa, from the RNA of the virus to its complementary DNA (cDNA), and from it with the help of T7 phage RNA polymerase “back” to the infectious RNA. Since the resulting coronavirus genome plus RNA mimics the cellular messenger RNA (mRNA), it is immediately recognized by the cell's translation machine and triggers the formation of its own infectious viral particles. Two systems of reverse genetics have been developed, involving the production of infectious plus RNA - under conditions *in vitro* and *in vivo* . The problem of obtaining a full-length cDNA of the giant genome of coronaviruses is solved by its fragmentation and subsequent cross-linking of the fragments using standard molecular biology approaches. The article provides examples of how this technology makes it possible to obtain synthetic coronaviruses that are indistinguishable from those isolated from nature, change the range of their hosts, increase virulence and resistance to specific antibodies, and influence the pathogenesis of the disease. It also shows the prospects for using recombinant viruses in cell screening assays and *in vivo* infection models to identify prophylactic and therapeutic approaches to the treatment of viral infections.

Key words: atypical pneumonia; coronavirus; interspecies transfer; reverse genetics; pandemic; COVID-19; SARS-CoV-2 .

Bibliographic description: *Supotnitsky M.V. The use of reverse genetics for the construction of previously non-existent coronaviruses: technologies and achievements // Bulletin of the RCB Protection Troops.*

2021. V. 5. No. 3. S. 217–235. https://doi.org/10.35825/2587-5728-2021-5-3-217-235

Coronaviruses have been a routine of genetic engineering for at least 25 years. The creation of their recombinant derivatives has become commonplace in the practice of two collaborating / competing scientific schools: the University of North Carolina at Chapel Hill (UNC, UNC-Chapel Hill; United States) - head Ralph Steven Barik (Ralph Steven Baric, born in 1954) [1, 2]; and the Wuhan Institute of Virology (WIV; China) [1] – head of the school Zhengli-Li Shi, born in 1964 [3]; But these are not the only scientific schools experimenting with the genomes of coronaviruses. In the pre-pandemic period, research on changing the genome of coronaviruses was carried out openly, with noble, publicly announced goals [2]: to determine which protein is responsible for the ability of coronaviruses to move from one animal species to another; whether they can get from animals to people; and whether they can be spread by airborne droplets between people; how to create a life-saving vaccine in case "if" ..., etc. SARS-CoV-2 was manipulated by reverse genetics even before the announcement of the COVID-19 pandemic [4] [3]. After the announcement of a pandemic, work on obtaining synthetic variants of SARS-CoV-2 became explosive [4–8]. In 2021, cell and vector systems were created to obtain the virus in amounts beyond the needs of diagnostic studies [8]. Obtaining variants of bat coronaviruses capable of causing infection in other animal species and in human cells [9], as well as variants of SARS-CoV-2 that did not previously exist in nature [5–8], indicates the transition of genetic studies of coronaviruses to the level of *synthetic biology* [10, 11]. In the context of the protracted COVID-19 pandemic and the remaining unknown natural reservoir of SARS-CoV-2, such studies and their methodology should be closely monitored, and its results should be publicly available.

The purpose of this work is to show how the use of reverse genetics allows us to construct previously non-existent coronaviruses, technologies and major achievements in their creation.

All information used to prepare this article is in the public domain and can be verified using the links to sources. Her search was carried out using the text medical database PubMed; search capabilities of Google Scholar, Scientific Electronic Library eLIBRARY.RU and scientific specialized publications.

The first coronavirus chimeras. One of the hallmarks of viruses of the *Coronaviridae* family is their high host specificity. The main task of research on obtaining coronavirus chimeras in the late 1990s. was the identification of the molecular basis of the interaction of coronaviruses with their corresponding host cell receptors. Its solution would make it easier to understand the pathogenesis of coronavirus infections, but this was hindered by ignorance of the details of the virion assembly process. The opportunity to begin studying the role of individual proteins in the morphogenesis of coronavirus was provided by a new technology - the assembly of coronavirus-like particles (coronavirus **-** like particles, VLP **)** from proteins M (membrane matrix protein), E (membrane protein) and spike glycoprotein (S-protein) co-expressed in cell culture without the participation of the virus nucleocapsid (protein N). VLPs were released from cells and formed a homogeneous population morphologically indistinguishable from normal virions [12].

Proteins M and E could independently form VLPs. The role of the S protein in virion assembly and budding remained unclear. It was itself transported to the plasma membrane and retained in the Golgi complex due to its association with the M protein. S-multimers somehow specifically fit into the voids of the M (or M and E) monomer arrays, but they did not make a special contribution to their overall

stability. VLP [13]. It became obvious that the S protein, although not required for virus assembly, is required for some other important function, for example, for cell infection [14].

In order to prove the role of the S-protein in the specific recognition of receptors on the surface of target cells and thus show its participation in the initiation of the infectious process, L. Kuo et al. [14] [4] constructed *a mutant* mouse hepatitis virus (MHV) in which the spike glycoprotein (S) ectodomain was replaced by a highly divergent feline infectious peritonitis virus (FIPV) S protein ectodomain. MHV and FIPV belong to two different groups of coronaviruses, and each is highly specific for its respective host species. The S proteins of MHV and FIPV share only 26% amino acid identity, with the greatest difference occurring in the amino-terminal half of each molecule. They recognize different receptors: MHV, members of the mouse biliary glycoprotein family; FIPV, feline aminopeptidase N (fAPN).

The resulting viable chimeric virus, designated fMHV, acquired the ability to *cross the species barrier* —i.e. to infect cat cells, and at the same time it lost the ability to infect mouse cells in tissue culture. This reciprocal species-specific switch has convinced researchers that the range of coronavirus host cells is determined primarily by interactions between the S protein and the viral receptor on the host cell. The S-protein of the coronavirus is the main and possibly the only factor in its species specificity. At the same time, the researchers were aware of the limitations of the site-specific mutagenesis method they used to study gene expression and function due to the extremely large genome size of coronaviruses – 28–32 kb. [fourteen].

Development of a reverse genetic system for constructing coronaviruses. Obtaining a full-length cDNA [5] of a giant (for viruses!) genome, and even in combination with areas of instability capable of forming a complete infectious virus RNA in permissive cell

lines, was for molecular biologists of the 1990s. was a difficult task. At the beginning of the 2000s, the group of Ralph Barik solved it by following the path of deconstructing the coronavirus RNA genome into cDNA fragments obtained using RT-PCR [6] or chemical synthesis. These fragments were then sequentially ligated according to the sequence of the RNA chain of the virus, maximizing the stability of the genome. If necessary, nucleotide substitutions were made in individual fragments. Full-length cDNA of the virus *in vitro* was used as a template for RNA transcription using T7 phage RNA polymerase . The resulting plus RNA of the coronavirus genome mimics cellular mRNA. Therefore, when introduced into a permissive cell, it is immediately recognized by its translation machine and triggers the formation of its own infectious viral particles. That is, the creators of this method of virus synthesis went in the "opposite direction" - not from DNA to RNA, as is usually done in a cell during protein synthesis, but vice versa, from the RNA of the virus to its cDNA, from it "back" - to infectious RNA. Nucleotide substitutions and deletions introduced into cDNA fragments before their cross-linking into the full-length cDNA of the virus, after transcription of RNA from it and the formation of viral particles, if one does not know about their artificial origin, will be considered in epidemic chains as mutations of the natural strain of the virus, for example, Wuhan -Hu-1. This is the essence of the work of the reverse genetic system (English reverse genetic system,

Currently, two systems of reverse genetics are used to obtain synthetic coronaviruses, which involve the production of infectious plusRNA under *in vitro* conditions [4–6, 16]; and two systems that allow the production of infectious virus plusRNA under *in vivo conditions* [7, 8].

Obtaining infectious plusRNA coronavirus *in vitro* . The technology was pioneered by Ralph Barrick's group. Using the viruses

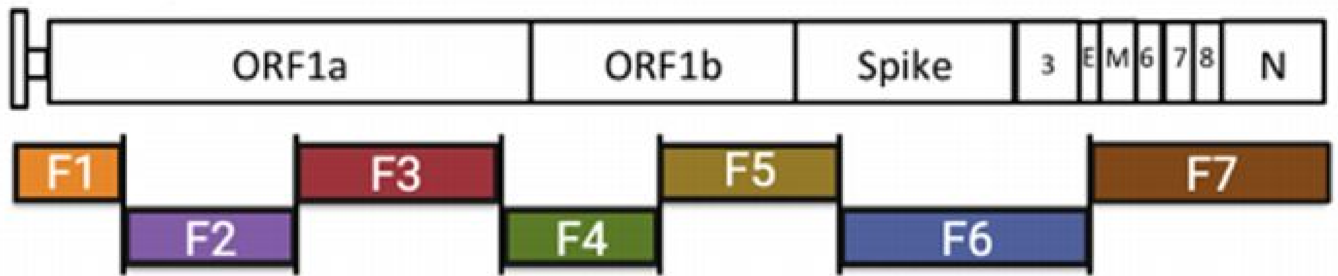
of transmissible gastroenteritis (ransmissible gastroenteritis virus, TGEV) and murine hepatitis (strain MHV-A59), then the most studied among coronaviruses, they assembled full-length cDNA of both viruses. On plasmids, cDNA subclones covering the entire virus genome were obtained. The cDNA subclones were then stitched together under *in vitro conditions.* and received an intact cDNA construct, fully consistent with the original plus-strand RNA coronavirus. There were no signs of full-length cDNA assembly. Transcripts obtained from the full-length cDNA of the coronavirus were introduced into permissive cell lines (Vero E6) by electroporation, and the assembly of full-fledged, infectious viral particles began in them [1, 2] [7] . The general scheme for the assembly of TGEV cDNA clones into a full-length cDNA of an infectious RNA virus is shown in Figure 1.

Figure 1 - *Reverse genetic system for constructing coronaviruses. The general assembly scheme of six cDNA clones (TGEV A–TGEVF) into full-length TGEV cDNA. The plaque-derived virus replicated efficiently and exhibited similar plaque morphology in permissive cell lines. Numerical designations correspond to the position of nucleotides in the virus genome. Abbreviations: Bgl1 and BxtX1, class IIS and IIG restriction endonucleases, which formed "sticky ends" along which "seamless" joining of cDNA clones took place; PL, papain-like protease; GFL, growth factor-like domain; Pol, polymerase motif; MIB, metal-binding motif; Hel, helicase motif; VD, variable domain; CD, conserved domain; ↑ – intergenic starts. Details in B. Yount et al. [one].*

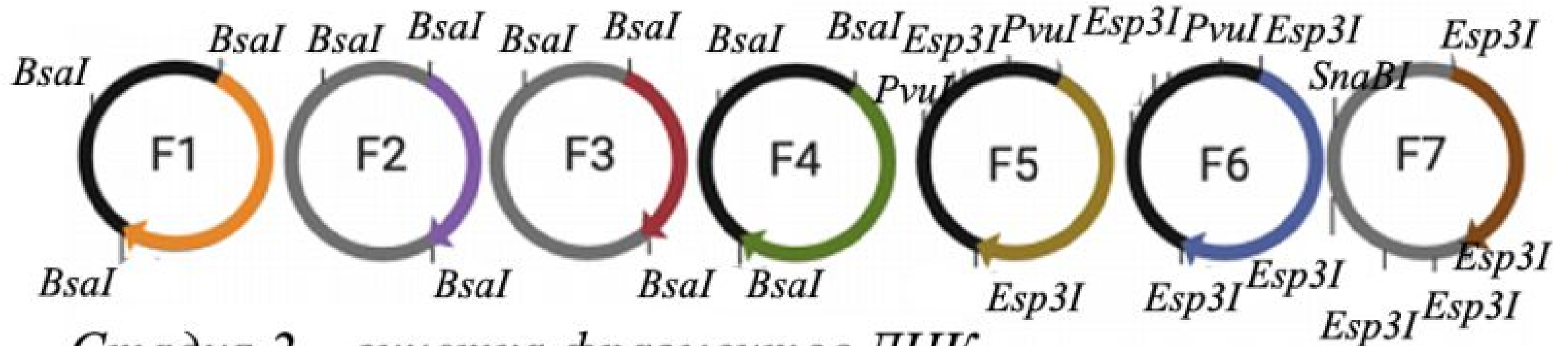
The complete genome of SARS-CoV Urbani was assembled in the same way as six contiguous subclones. Through the unique sites of the restriction endonuclease BglI, they were connected (ligated) into a full-length cDNA of the virus and used *in vitro* as a template for RNA transcription using T7 phage RNA polymerase . The resulting RNA transcripts were introduced into permissive cells by electroporation, where they were used by the cellular translation machine to form infectious viral particles [1].

The same technology, adapted by H. Xie et al. [5, 6] [8] to obtain synthetic derivatives of SARS-CoV-2, is shown in Figure 2. It allows: 1) to generate mutant and reporter SARS-CoV-2 and other viruses by manipulating a plasmid containing a cDNA fragment with the required mutation (or mutations), reducing the risk of off-target mutations or deletions inadvertently included in the recombinant virus; 2) simultaneously manipulate multiple mutations from different cDNA fragments, since more than one mutation from different cDNA fragments can be constructed in parallel to create combinatorial mutant viruses. Such flexibility is important in characterizing the combinatorial effect of multiple viral mutations on host immune response or disease progression; 3) quickly insert mutations into the genome of an artificial virus,

Геном SARS-CoV-2



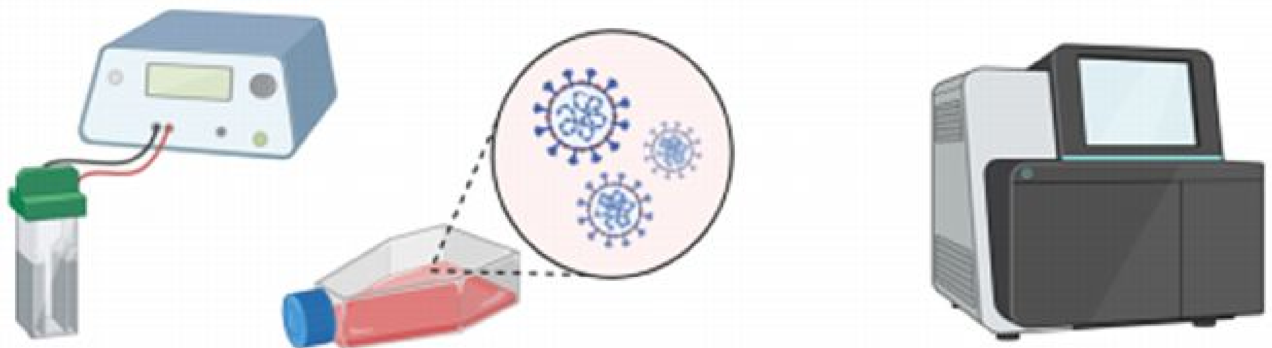
Стадия 1 – приготовление плазмид с фрагментами F1–F7



Стадия 2 – очистка фрагментов ДНК

Стадия 3 – in vitro сборка фрагментов ДНК в полноразмерную кДНК SARS-CoV-2

Стадия 4 – in vitro транскрипция полноразмерной РНК вируса



Стадия 5 – электропорация в пермисивные клетки

Стадия 6 – полногеномное секвенирование вируса

Figure 2 – *The main six stages of obtaining synthetic SARS-CoV-2 and its derivatives when obtaining virus plusRNA under in vitro. The stages are divided into 108 stages. Steps 1-4 are performed in a shared laboratory.*

Steps 5-6 procedures related to SARS-CoV-2 manipulation should be performed in a biosafety level 3 (BSL-3) laboratory.

Stage 1 – preparation of seven plasmids containing fragments F1–F7 of SARS-CoV-2. Unwanted mutations in plasmids prior to assembly of full-length SARS-CoV-2 DNA are excluded by restriction analysis and Sanger sequencing.

Stage 2 - preparation of high-quality DNA fragments for subsequent experiments by digestion (hydrolysis) of plasmids with restriction enzymes.

Stage 3 - assembly of seven DNA fragments into full-length SARS-CoV-2 DNA under in vitro using T4 DNA ligase. The full-length ligation product is immediately purified by phenol-chloroform extraction and precipitation with isopropanol.

Stage 4 - transcription under in conditions vitro full-length RNA and N-gene RNA.

Stage 5 - electroporation into permissive cells (Vero E6 or BHK-21 and VeroE6) full-length virus RNA and isolation of the recombinant SARS-CoV-2 virus from cell culture.

Stage 6 - Sanger whole genome sequencing of the virus to check the entire genome sequence of the resulting virus. According to H. Xie et al. [6].

TTN Thao et al. [4] proposed a platform for the assembly of large genomes of coronaviruses, an alternative to that of Ralph Barik [9] . Their platform uses the yeast *Saccharomyces cerevisiae to create*

synthetic RNA viruses. The rationale for using the cloning system in yeast is the ability of yeast to recombine overlapping DNA fragments *in vivo* , which has led to the development of a technique called transformation-associated recombination cloning (TAR). Subgenomic fragments of viruses have been created TTN Thao et al. [4] using viral isolates, cloned viral DNA, clinical samples, or synthetic DNA, and then these fragments were reassembled into one genome in *S. cerevisiae* using TAR technology, which allows the resulting recombinant to be preserved as an artificial chromosome in yeast **[10]** [17]. The infectious RNA of the replication-capable virus was then obtained using T7 RNA polymerase. Initially, this technology was tested on the MHV A59 (Figure 3).

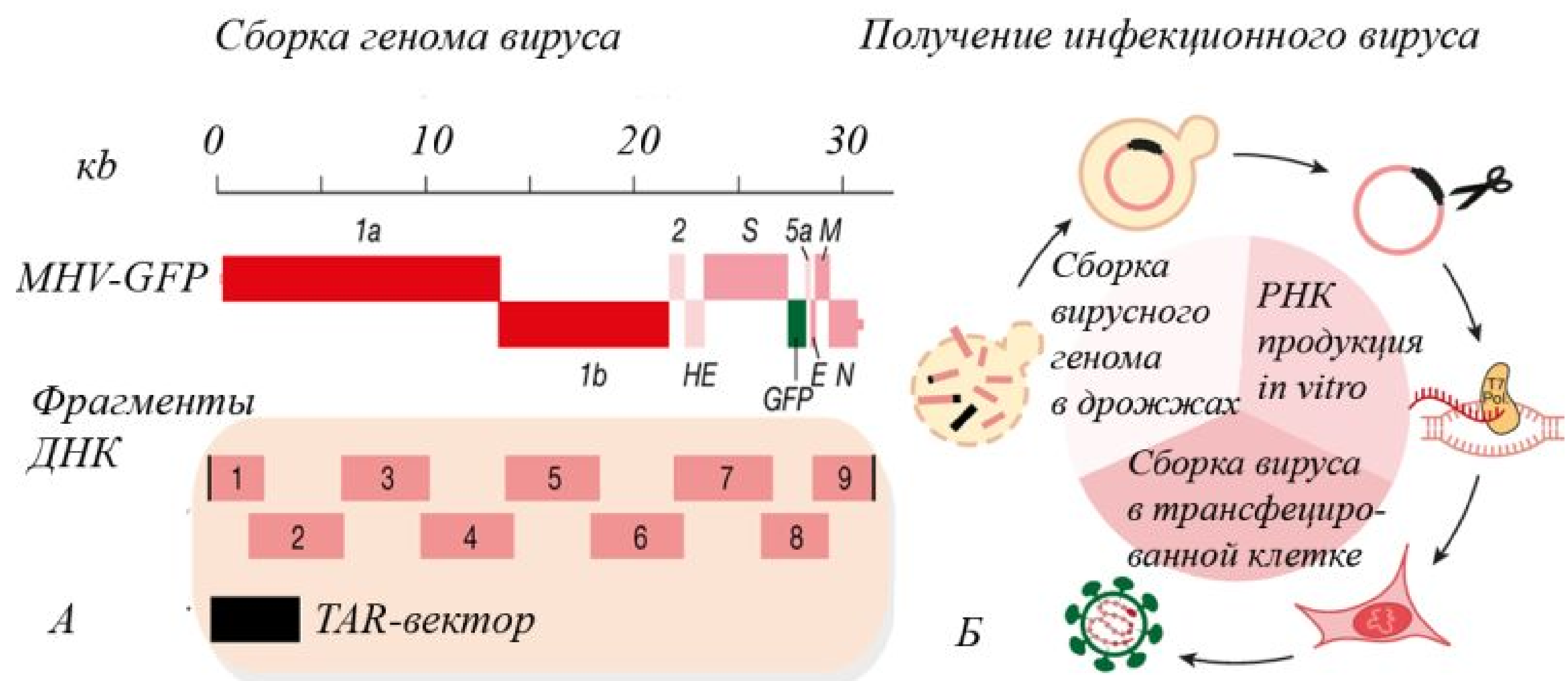


Figure 3. ***Scheme for creating a synthetic infectious coronavirus using TAR technology. A. Viral RNA was obtained from mouse 17Cl-1 cells infected with MHV-GFP and used for amplification by RT-PCR overlapping DNA fragments spanning the MHV-GFP genome from 2024 to 29672 nucleotides. DNA fragments containing 5' and 3' ends were cloned on vaccinia virus and***

amplified by PCR. B. All DNA fragments were simultaneously transformed into S. cerevisiae (strain VL6-48N). The transformed DNA fragments were assembled by homologous recombination in yeast to generate a YAC that contains the full length viral cDNA sequence. In vitro production of infectious capped viral RNA begins with YAC isolation followed by plasmid linearization to provide a DNA template for transcription of the viral RNA by T7 RNA polymerase. Infectious virus production is initiated by electroporation of BHK-MHV-N cells, after which the production and amplification of the virus is carried out by culturing the virus in permissive cell lines. For a detailed description of the technology, see TTN Thao et al. [four].

To evaluate the possibility of applying the synthetic genomics platform to other coronaviruses, the researchers recreated MERS-CoV from eight overlapping DNA fragments. As a result, recombinant rMERS-CoV and rMERS-CoV-GFP were obtained. With this experiment, they showed that the synthetic genomics platform is suitable for modifying the coronavirus genome. Synthetic rMERS-CoV-GFP is unique in that it contained an insert of the green fluorescent protein (GFP) gene, consisting of 238 amino acids (MM 27 kD). Clones of synthetic viruses were passaged on permissive cells 15–17 times; subsequent sequencing showed that the obtained genomes were stably preserved during passages.

Further, TTN Thao et al. [4] cloned several other coronaviruses: HCoV-229E2, HCoV-HKU1 (GenBank: NC_006577), MERS-CoV-Riyadh-1734-2015 (GenBank: MN481979); and other families of viruses such as the ZIKA virus of the Flaviviridae family (GenBank: KX377337) and

human respiratory syncytial virus of the Pneumoviridae family (hRSV). Cloning of these viral genomes using TAR technology was successful in all cases, regardless of the source of the virus, the nucleic acid template, or the number of DNA fragments. Cloning of hRSV-B was carried out *without any prior information about the virus genome*, directly from a clinical specimen (nasopharyngeal aspirate) of four overlapping DNA fragments (GenBank: MT107528). Taken together, these results demonstrate that the TAR synthetic genomics platform provides technological advances for the rapid generation of molecular clones of various RNA viruses using viral isolates, cloned DNA, synthetic DNA, or clinical specimens as starting material.

Obtaining infectious plusRNA directly in permissive cells. Systems of this type differ from those described above in that they lack an error-prone cDNA transcription step under *in vitro* conditions . The first such reverse genetics system was created in 2020 by Ye Ch. et al. [7] [11] based *on the bacterial artificial chromosome* (BAC) [12] . It was used to generate an infectious recombinant SARS-CoV-2 (rSARS-CoV-2), which exhibits *in vivo* properties similar to those of a natural virus isolate.

The researchers chemically synthesized five fragments of the genome of the SARS-CoV-2 strain isolated from an oropharyngeal swab taken from a patient with respiratory symptoms in Snohomish County, Washington (USA). They were assembled in the pBeloBAC plasmid [13] using standard molecular biology approaches (Figure 4).

To facilitate the assembly of the viral genome, genetic marks were included in it to distinguish the rSARS-CoV-2 clone from the natural isolate - two silent mutations. One in the S-protein gene (21895 nucleotides), the other in the M-protein gene (26843 nucleotides); removing the BstBI and MluI restriction sites, respectively (Figure 4B).

To restore rSARS-CoV-2, Vero E6 cells were transfected with BAC SARS-CoV-2 and BAC (control), and in comparison with the control, the presence of a cytopathic effect in the cells was monitored, which manifested itself 72 h after transfection. The production of the infectious virus (passage 0) by the transfected cells was 3.4×105 PFU/mL [14] (Figure 4D). Recovery of rSARS-CoV-2 was confirmed by the detection of viral antigen in Vero E6 cells infected with tissue culture supernatants previously harvested from Vero E6 cells transfected with SARS-CoV-2 BAC.

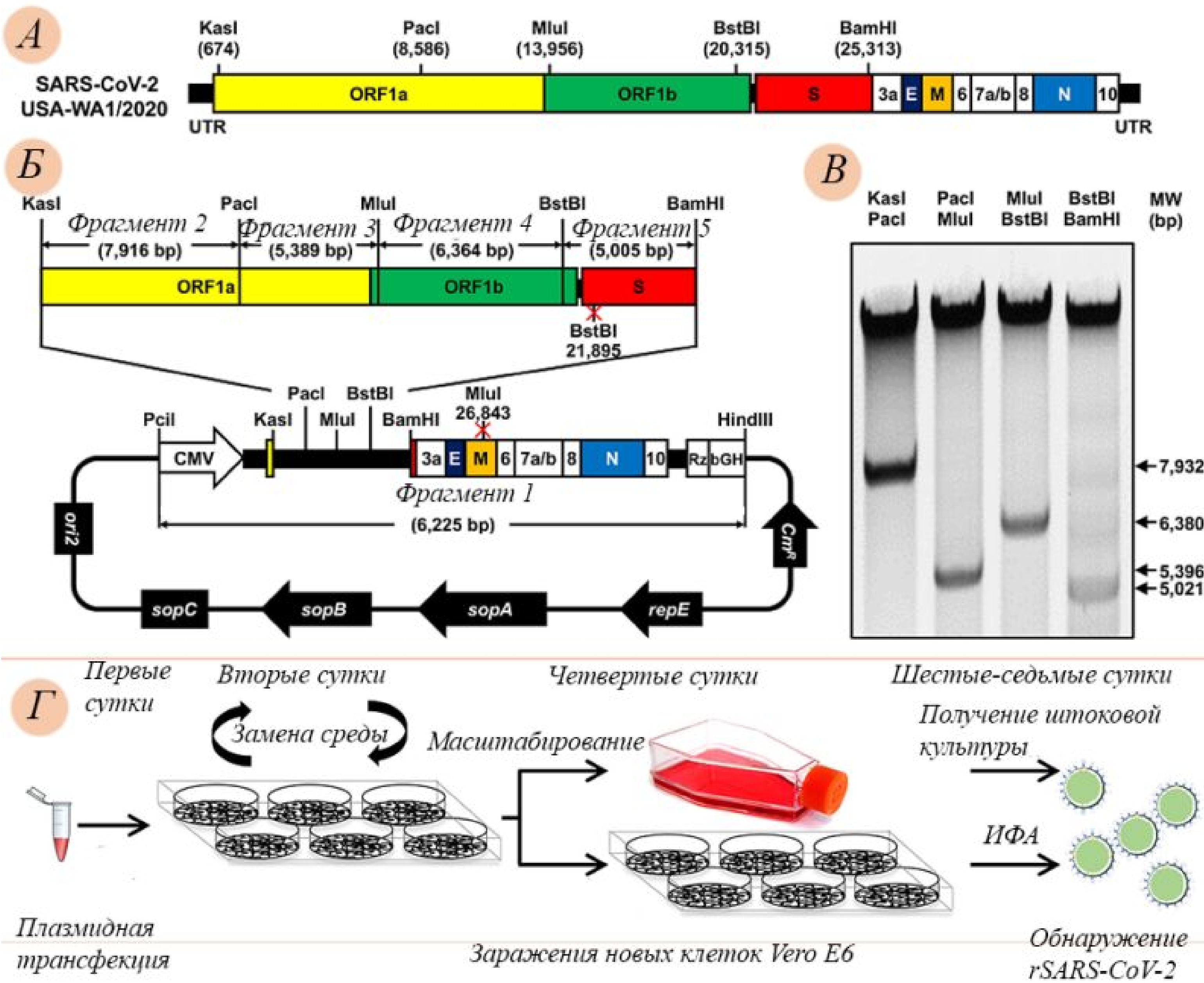


Figure 4 – ***Assembly of a synthetic SARS-CoV-2 genome in BAC and production of a synthetic virus directly in permissive cells. A. Schematic representation of the SARS-CoV-2 genome. These restriction sites were used to clone the entire viral genome (29,903 nucleotides) of SARS-CoV-2 (strain USA-WA1/2020) into the pBeloBAC plasmid. Open reading frames of structural proteins 1a, 1b, spike (S), shell (E), template (M) and nucleocapsid (N) and additional proteins (3a, 6, 7a, 7b, 8 and 10) are shown. UTR. Length not to scale. B and C. Assembly of the viral genome. B. A full-length infectious cDNA clone was assembled by sequentially cloning chemically synthesized fragments 1 to 5 covering the entire viral genome into pBeloBAC plasmid using the indicated***

restriction sites under the control of the cytomegalovirus (CMV) promoter; the clone was flanked at the 3' end with a hepatitis delta virus (HDV) ribozyme (Rz) and bovine growth hormone (bGH) termination and polyadenylation sequences. The length of each of the chemically synthesized viral fragments is indicated. Ori2 indicates the origin of BAC replication. sopA, sopB, and sopC are plasmid elements that ensure that each bacterial cell receives a copy of the BAC. CmR indicates the chloramphenicol resistance gene. B. After assembly, the DNA of the BAC clone carrying the entire viral genome was digested with the indicated restriction enzymes (top) and the DNA products were analyzed on a 0.5% agarose gel. D. Schematic representation of the rSARS-CoV-2 rescue approach. Vero cells The length of each of the chemically synthesized viral fragments is indicated. Ori2 indicates the origin of BAC replication. sopA, sopB, and sopC are plasmid elements that ensure that each bacterial cell receives a copy of the BAC. CmR indicates the chloramphenicol resistance gene. B. After assembly, the DNA of the BAC clone carrying the entire viral genome was digested with the indicated restriction enzymes (top) and the DNA products were analyzed on a 0.5% agarose gel. D. Schematic representation of the rSARS-CoV-2 rescue approach. Vero cells The length of each of the chemically synthesized viral fragments is indicated. Ori2 indicates the origin of BAC replication. sopA, sopB, and sopC are plasmid elements that ensure that each bacterial cell receives a copy of the BAC. CmR indicates the chloramphenicol resistance gene. B. After assembly, the DNA of the BAC clone carrying the entire viral genome was digested with the indicated restriction enzymes (top) and the DNA products were analyzed on a 0.5% agarose gel. D. Schematic representation of the rSARS-CoV-2 rescue approach. Vero cells digested with the indicated restriction enzymes (top) and the DNA products analyzed on a 0.5% agarose gel. D. Schematic

representation of the rSARS-CoV-2 rescue approach. Vero cells digested with the indicated restriction enzymes (top) and the DNA products analyzed on a 0.5% agarose gel. D. Schematic representation of the rSARS-CoV-2 rescue approach. Vero cells E6 BAC were transfected with SARS-CoV-2. After 24 hours, the transfection medium was changed to the post-infection medium. On day 4, cells were added to T75 flasks and tissue culture supernatant was used to infect new Vero cells. E6. 48 hours after infection in Vero cells E6 using immunofluorescence detected rSARS-CoV-2. As a control experiment, Vero cells E6 were transfected with BAC. By Ye Ch. et al. [7].

A few months later, a similar system was developed by S. Rihn et al. [8] [15] . From the Ye Ch system described above. et al. [7] it differs in that a *plasmid vector* , rather than BAC, is used to introduce the DNA encoding the plusRNA of the virus into permissive cells . Direct visualization and quantification of replicating virus in cells was provided by cassettes containing markers. The researchers called it *the reverse genetics (RG) plasmid system* . They created an infectious cDNA clone (icDNA) SARS-CoV-2 based on the Wuhan-Hu-1 virus using the low-copy plasmid pCC1-4K (which does not contain the F factor responsible for plasmid conjugation) .

The SARS-CoV-2 icDNA genome is assembled from 5 synthetic DNA fragments, each flanked by unique SanDI (1524), PacI (8586), MluI (13956), Bsu36I (18176), and BamHI (25313) restriction sites. The human cytomegalovirus (CMV) promoter was inserted at the 5' end of the genome; double ribozyme of hepatitis virus (HDV) [16] and a simian virus 40 (SV40) terminator sequence were added after the poly-

A tail of the 3' end of the viral genome. These elements ensure efficient transcription and homogeneous processing of the 3' end during infectious virus rescue. Direct visualization and quantification of the replicating virus in cells was provided by cassettes containing sequences encoding fluorescent (mCherry and ZsGreen) and bioluminescent (nanoluciferase or NLuc) protein markers [17] inserted into the backbone of the plasmid. To avoid deletion of viral sequences, the markers were cloned inside the reading frame (in-frame) closer to the C-terminus of the ORF7a protein using the FMDV 2A “ribosomal skip” (2A) linker [18]. In this way, they succeeded in releasing the emerging reporter protein from the ORF7a protein (Figure 5).

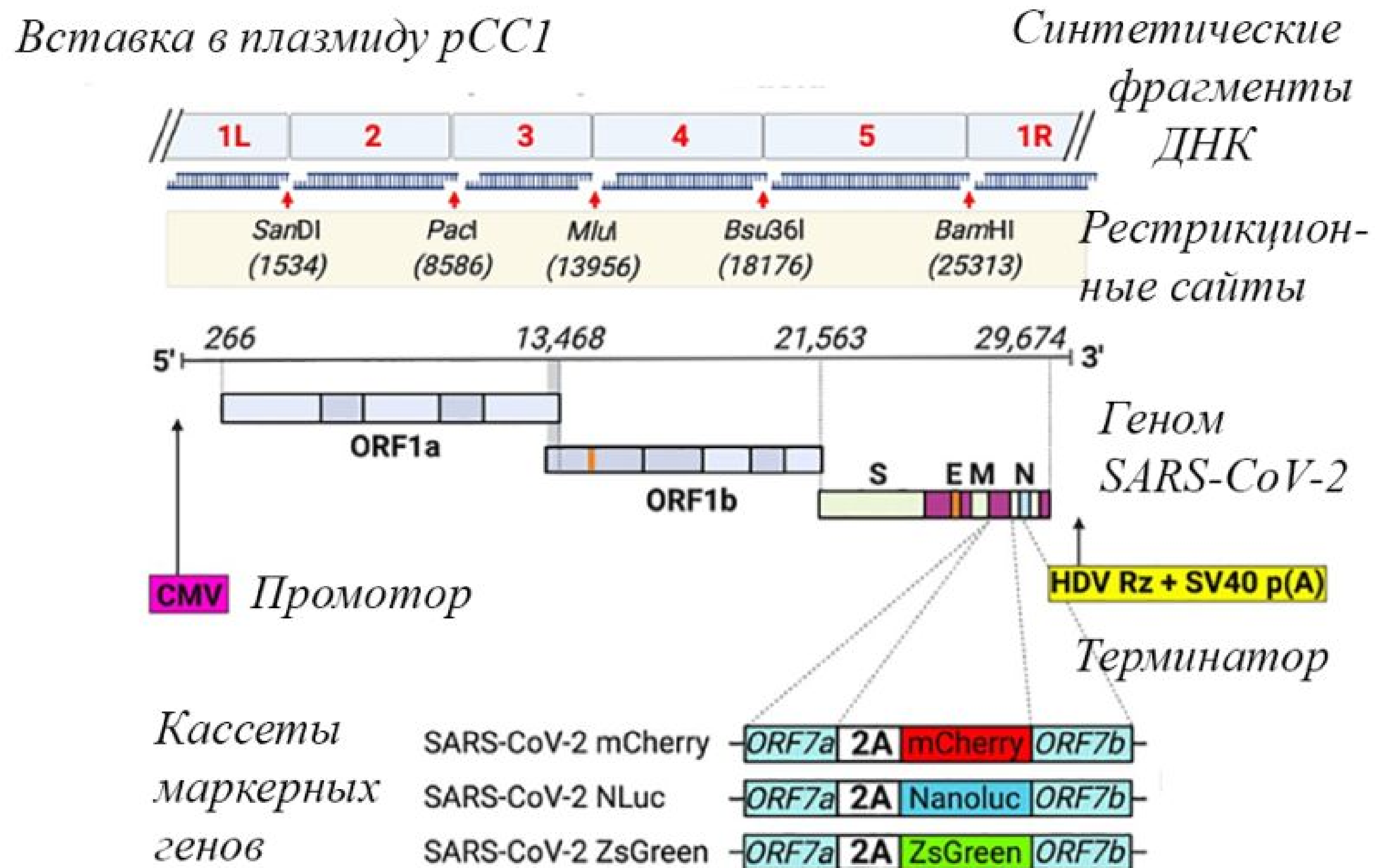


Figure 5 – *icDNA clone construction scheme pCC1-4K-SARS-CoV-2-*

Wuhan-Hu-1. Synthetic DNA fragments 1, 2 and 3 based on the SARS-CoV-2-Wuhan-Hu1 sequence were cloned into pCC1 plasmids (derived from plasmid pCC1BAC), and fragments 4 and 5 were cloned into high copy plasmid pUC57Kan by Gene Synthesis Company (Genscript) . The fragments were designed to contain the SanD1, PacI, Mlul, Bsu36I and BamHI specific restriction cloning sites for cloning purposes. The sequences encoding mCherry, ZsGreen, and NLuc markers were cloned in frame to the C-terminus of the ORF7a protein via an FMDV 2A linker. A significant portion of the regulatory sequences of plasmid pCC1 was removed to obtain plasmid pCC1-4K (which lacks intact factor F). The pCC1-4K system is maintained at one copy per cell, which ensures the stability of the SARS-CoV-2 cDNA. To increase the functionality of this system, cassettes of various reporter molecules (for example, mCherry, ZsGreen, and Nanoluciferase (NLuc)) were inserted into the icDNA genome. The viral genome itself was flanked by a eukaryotic promoter and terminator. According to S. Rihn et al. [8].

Developed by S. Rihn et al. [8], the reverse genetics plasmid system makes it possible to obtain replication-competent SARS-CoV-2 directly from cell culture supernatants. After transfection of the plasmid into BHK-21 cells, they were placed in six-well plates, and after three days of growth, the supernatant with the virus was transferred to T25 flasks containing Vero E6 cells. In addition, this SARS-CoV-2 construct is easily amenable to genetic manipulation (to study virus variants) and the insertion of reporters such as fluorescent or bioluminescent proteins, which can be used in various studies both *in vitro* and *in vivo and* provide direct detection and quantification of the kinetics of virus replication in the cell.

Study of the biology of chimeric coronaviruses. Researchers needed to understand the extent of the change in the S-protein and in which direction to change them in order to expand the specificity of the coronavirus in the organs and tissues of the traditional host, and expand the range of its new hosts. Such experiments were carried out with MHV. An attempt was made to expand the host range and specificity of the virus by switching it from a highly specific receptor (individual mouse biliary glycoproteins) to a non-specific one, heparin sulfate (a linear polysaccharide found in all animal tissues). What came of this is not clear from the article. But times have changed. The deadly SARS epidemic has shifted researchers' interests from the S-protein variations of the model MHV and FIPV to the S-protein that provides the specificity of more dangerous coronaviruses. The key problem was the mechanism by which they can acquire the ability to change the "owner", i.e.[19] .

KE Follis et al. [19] of The University of Montana, (Missoula, USA) drew attention to the absence of a cleavage site by furin and furin-like cellular proteases in the SARS-CoV S-protein, which led to its incomplete cleavage upon interaction with the receptor, and, consequently, to the failure to use the entire pathogenic potential of the virus. This finding contradicted the then existing ideas about the special role of proteolytic maturation in the structure and function of other class I glycoproteins included in the envelopes of viruses. Cleavage of the S protein by cellular proteases is required to provide the fusion potential of the envelope glycoproteins of retroviruses, orthomyxoviruses, paramyxoviruses, filoviruses, arenaviruses, and many coronaviruses (MHV, avian coronaviruses, CoV OC43), since they are initially synthesized as inactive precursors. They need proteolytic cleavage for maturation and full functional activity. After the subsequent activation of the mature class I envelope glycoprotein due to binding to the receptor and/or low endosomal pH, these complexes

undergo a deep structural reorganization with the formation of highly stable structures, hairpins, which facilitate efficient fusion of the viral and cellular membranes [20].

Using the MAXHOM algorithm used to align the putative S1–S2 junction region of the S-glycoprotein coronaviruses, KE Follis et al. [19] found remnants of a furin site in SARS-CoV, CoV 229E, and NL63 glycoproteins that disappeared as a result of spontaneous deletions. A once-existing protease-responsive site was signaled by a single arginine (R) at position 667 of the SARS-CoV S-glycoprotein. Introduction of the synthetic furin recognition sequence SLLR into R667, i.e. into the putative S1–S2 junction region, enabled efficient cleavage of S-glycoprotein to form discrete S1 and S2 subunits, and markedly increased the ability of the spike complex to mediate cell fusion, i.e. form syncretism.

The absence of a furin site in SARS-CoV, and without it leading to a fatal outcome in almost every tenth who fell ill with SARS, aroused the interest of Japanese researchers. R. Watanabe et al. [21] from the National Institute of Infectious Diseases (Tokyo, Japan). They introduced a furin-like cleavage sequence into the S protein at amino acids 798 to 801 and found that the S protein was now cleaved when expressed on the cell surface and induced cell fusion without trypsin treatment. In addition, they found that a pseudotyped virus carrying a cleaved S protein infects cells in the presence of a lysosomotropic agent as well as a protease inhibitor, both of which normally block SARS-CoV from entering the cell through endosomes. The results of R. Watanabe et al. [21] showed that the insertion of a furin site into the S protein of SARS-CoV allows the virus to enter the cell directly from its surface.

The discovery of SARS-like CoVs (SL-CoVs) that are identical in genomic organization to SARS-CoV but differ in binding to ACE2 [20]

has raised the question of how insurmountable the interspecies barrier between humans and bats is for such viruses. In other words, can SARS-like CoV animals, as a result of random recombination with the genomes of other coronaviruses, acquire the ability to cause an infectious process in humans.

The first attempt to answer this question was made by the Shi Zhengli group. To do this, based on the human immunodeficiency virus (HIV), they constructed a pseudoviral system with cell lines expressing human, civet, or horseshoe bat ACE2 molecules. Pseudoviruses included the full-length S protein of SL-CoV and SARS-CoV, and a series of S chimeras that included inserts of different SARS-CoV S protein sequences into the backbone of the SL-CoV S protein. They showed that the S protein of SL-CoV cannot use ACE2 of different species to enter cells. The SARS-CoV S protein also cannot bind the ACE2 molecule of the horseshoe bat *Rhinolophus pearsonii* . However, when the receptor binding domain (RBD) of the SL-CoV S protein was replaced with the RBD of the SARS-CoV S protein, the fusion S protein acquired the ability to use human ACE2 to enter the cell (albeit with different efficiency for different constructs), which meant the structural and functional similarity of the SL-CoV S-protein with the SARS-CoV S-protein. These results suggest that while the SL-CoVs found in bats at the time of the study are unlikely to infect humans, it remains to be seen if they are capable of using other surface molecules of certain human cell types. It is also possible that these viruses can become infectious to humans if they undergo an N-terminal sequence variation, for example by recombination with other CoVs, which in turn

Ralph Barrick's group, unlike Shi Zhengli's, did not use pseudotyped viruses in their virus-receptor interaction studies. They considered this system safe, but too artificial. The results obtained on its basis are difficult to extrapolate to a real infectious process, since in principle it

cannot provide the correct structural expression of the S-protein on virions. Therefore, using synthetic biology techniques and a reverse genetics system (discussed above for the construction of TGEV and MHV), they obtained a series of isogenic strains corresponding to strains found in palm civets and raccoon dogs, as well as SARS-CoV isolates, covering early, middle and late phases of the SARS epidemic. Synthesized by them *in vitro* recombinant viruses replicated efficiently in cell culture and showed different sensitivity to neutralization by antibodies. Human, but not zoonotic, variants of viruses replicate efficiently in *human respiratory* epithelial cultures., confirming earlier hypotheses that zoonotic isolates are less pathogenic to humans but can develop into highly pathogenic strains. All artificial viruses reproduced efficiently in permissive cell lines. Severe lung injury, manifested by diffuse alveolar damage, hyaline membrane formation, alveolitis, and death, has been reported in 12-month-old mice infected intranasally with palm civet strain HC/SZ/61/03 or SARS-CoV variant GZ02 isolated during the early phase of the epidemic. . Related lines of SARS-CoV strains of the middle and late stages of the epidemic or raccoon dogs did not cause lung damage [22].

By 2008, the same group had synthesized a SARS-like CoV with a size of 29.7 kb. (Bat-SCoV), what they then believed to be the probable precursor of the epidemic SARS-CoV. At the start of their study, four Bat-SCoVs (HKU3-1, HKU3-2, HKU3-3, and RP3) had been identified, but none had been isolated in culture. The infectivity of these viruses was hypothesized, since their genomic RNA sequences were obtained by RT-PCR sequencing of samples of the genetic material of the viruses from faeces or rectal swabs of bats. Ralph Barik's group used *consensus design to synthesize the sequence of a previously non-existent virus*. Based on four Bat-SCoV sequences taken from the GenBank database (accession number FJ211859), the researchers designed the

coronavirus consensus sequence and “split” it into cDNA fragments with junction points exactly matched with the existing SARS-CoV reverse genetics system. The virus was replaced by the binding domain of the Bat-SCoV Spike receptor (RBD) to SARS-CoV RBD (Bat-SRBD). Defined and functional SARS-CoV 5'UTRs and transcriptional regulatory sequences were used because the Bat-SCoV 5'UTR was found to be incomplete. The synthesized genomic cDNA fragments were inserted into plasmid vectors and assembled into full length cDNA. It was transcribed *in vitro* with the production of coronavirus RNA, designated Bat-SRBD. It acquired the ability to infect cultured cells and mice [23].

Apparently, the time has come to unite the efforts of both groups. In their joint work, which, according to the authors, has the goal of “predicting and preparing for future viruses”, a previously non-existing virus, SHC014-MA15, was created using *reverse genetics* , *capable of replicating in the respiratory tract of humans and animals [9].* To do this, the nucleotide sequence of the S1 subunit, which circulates among bats of the SHC014 coronavirus, the closest “relative” of SARS-CoV, and which did not manifest itself as a pathogen for humans due to a difference in 14 amino acid residues in the region of the spike that binds to human ACE2, was point changes have been made [21]. The researchers changed the nucleotide sequence of the gene encoding the S1 subunit of SHC014 to that of SARS-CoV. The rest of the genes, i.e. those that determine the formation of the transcription complex of viral replication and the assembly of its particles in the cell did not undergo changes [22] . The new chimeric SARS-like coronavirus has been designated SHC014-MA15. Like SARS for infecting human, civet, and bat lung epithelial cells, it can use ACE2 as a target receptor and its orthologues, and replicate in them to high titers comparable to those of the natural strain SARS-CoV Urbani, remaining on other genes old SHC014. Experiments under *in vivo*

conditions demonstrated the replication of a chimeric virus in the lungs of model mice with a pronounced pathological process (Figure 6) [23] .

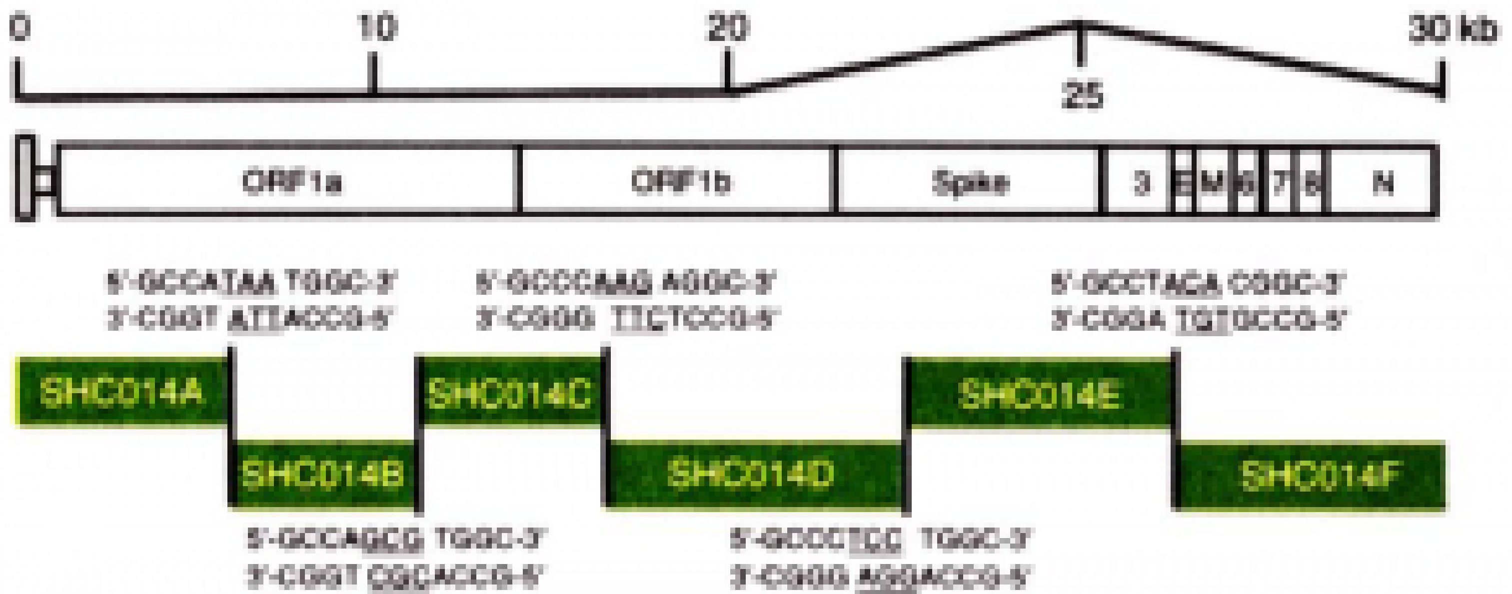


Figure 6 - Chimeric virus SHC014-MA15 capable of replication in the respiratory tract of humans and animals that have the ACE2 receptor or its orthologues. The genomic and amino acid sequences of the spike subunit S1 domains representative of CoV were downloaded from Genbank or the Pathosystems Resource Integration Center (PATRIC). The viral genome was synthesized as six contiguous cDNA segments (designated as SHC014A, SHC014B, SHC014C, SHC014D, SHC014E, and SHC014F) flanked by

unique BglI restriction sites, which provided targeted assembly of the full-length cDNA. The new chimera proved to be more virulent in human cells than the original virus, SHC014 [9].

But the similarity of both viruses ended with the common ability of the S1 subunits to recognize the human ACE2 receptor [24] , and the SHC014 virus's own pathogenic potential, which was not suspected before these experiments, was revealed, for example, the ability to overcome artificially created immunity. The researchers tried *in vitro* to determine the neutralizing efficacy against SHC014-MA15 of monoclonal antibodies (mAbs) of a wide protective spectrum (109.8, SHC014-MA15, 230.15 and 227.14) [25] , which showed a good neutralizing effect against SARS-CoV. The effect of these antibodies on the replication of SHC014-MA15 was insignificant, while the replication of SARS-CoV Urbani was suppressed by them at relatively low concentrations. Only the use of a high concentration (10 μg/mL) of mAb109.8 made it possible to achieve 50% neutralization of SHC014-MA15 under in vitro conditions [9] .

To evaluate the effectiveness of existing vaccines against HC014-MA15, the investigators vaccinated mice (Balb/cAnNHsD) with a candidate double inactivated whole SARS-CoV (DIV) vaccine. Previous work has shown that the DIV vaccine can protect mice from infection with a homologous virus [25].

However, vaccination against SARS-CoV not only did not protect animals from infection with HC014-MA15, but also aggravated the

course of the infectious process caused by it. Serum obtained from mice vaccinated with DIV failed to neutralize SHC014-MA15 *in vitro* [9].

An attempt to make an attenuated live vaccine based on HC014-MA15 was also unsuccessful. Its use conferred little cross-protection against SARS-CoV infection, confirming the presence of conserved common epitopes, but there were safety concerns with the vaccination. At doses that provided some protective effect against SARS-CoV infection, the vaccine itself caused a pathological process in experimental animals [9]. In general, experiments to obtain from a coronavirus that infects only bats, its genetically modified version, which poses a danger to humans, have yielded a lot of epidemiology. They made it possible to understand the mechanism of the interspecies "jump" of coronaviruses from bats to humans that exists in nature. I had to get rid of the illusion that the spread of new types of coronaviruses in human populations,[26] .

The search for approaches to designing synthetic CoVs and studying their biological properties has increasingly led researchers to viral chimeras that are not controlled by vaccination. This phenomenon was encountered by researchers at the Icahn School of Medicine at Mount Sinai (New York, NY), who created a chimeric WIV1-MA15 virus, in which the S protein SARS-CoV has been replaced by a similar one from WIV1-CoV. It turned out that vaccination with two different SARS-CoV-based inactivated vaccines failed to protect immunized mice against WIV1-MA15 [27].

Biological properties of artificial SARS-CoV-2. For coronavirus reverse genetics platforms, obtaining SARS-CoV-2 is nothing more than a special case. Judging by the statements of the authors of published works, reverse genetic systems that provide rapid synthesis of SARS-CoV-2 infectious wild-type strains, its mutant and reporter strains, are

being developed exclusively for studying the pathogenesis of viral infection, transmission mechanisms, developing therapeutic methods and creating vaccines [5, 6]. Using such a clone, it is possible to evaluate the effect of genetic changes in the virus by removing certain sequences from SARS-CoV-2 and studying their effect on virus replication, S-protein processing, its immunogenic and toxic properties. Recombinant SARS-CoV-2 (rSARS-CoV-2) are routinely compared to their natural isolates across a range of markers used to characterize the virus. Clones of synthetic SARS-CoV-2 retained the stability at the level of the original strains after 15–17 passages in tissue cultures with the preservation of genomes [4]. To further characterize the genetic identity of rSARS-CoV-2 to the parent strain, sequencing and restriction analysis are used. Growth kinetics, peak titers, cytopathic effect are usually similar to the original strain, the virus yield by transfected cells could be from $3.4\text{-}10^5$ PFU/ml [7] to 2.9×10^6 PFU/ml virus [5]. Recovery of rSARS-CoV-2 was also confirmed by the detection of viral antigen in Vero E6 cells infected with tissue culture supernatants previously harvested from Vero E6 cells transfected with rSARS-CoV-2 [7].

For comparative experiments to study the properties of genetically modified and natural variants of rSARS-CoV-2 *in vivo*, *golden Syrian hamsters* (*Mesocricetus auratus*) are currently used [28]. Ye Ch. et al. [7] to confirm that rSARS-CoV-2 exhibits the same ability to replicate, virulence and pathogenicity as the natural SARS-CoV-2 isolate, infected both strains of golden Syrian hamsters intranasally at a dose of $2*10^4$ POE. On the 2nd and 4th day after infection, they removed the upper and lower respiratory tract from infected animals, as well as from the control group of animals, and assessed the general pathological changes (lungs) and the degree of viral infection (upper respiratory tract and lungs) - Figure 7 .

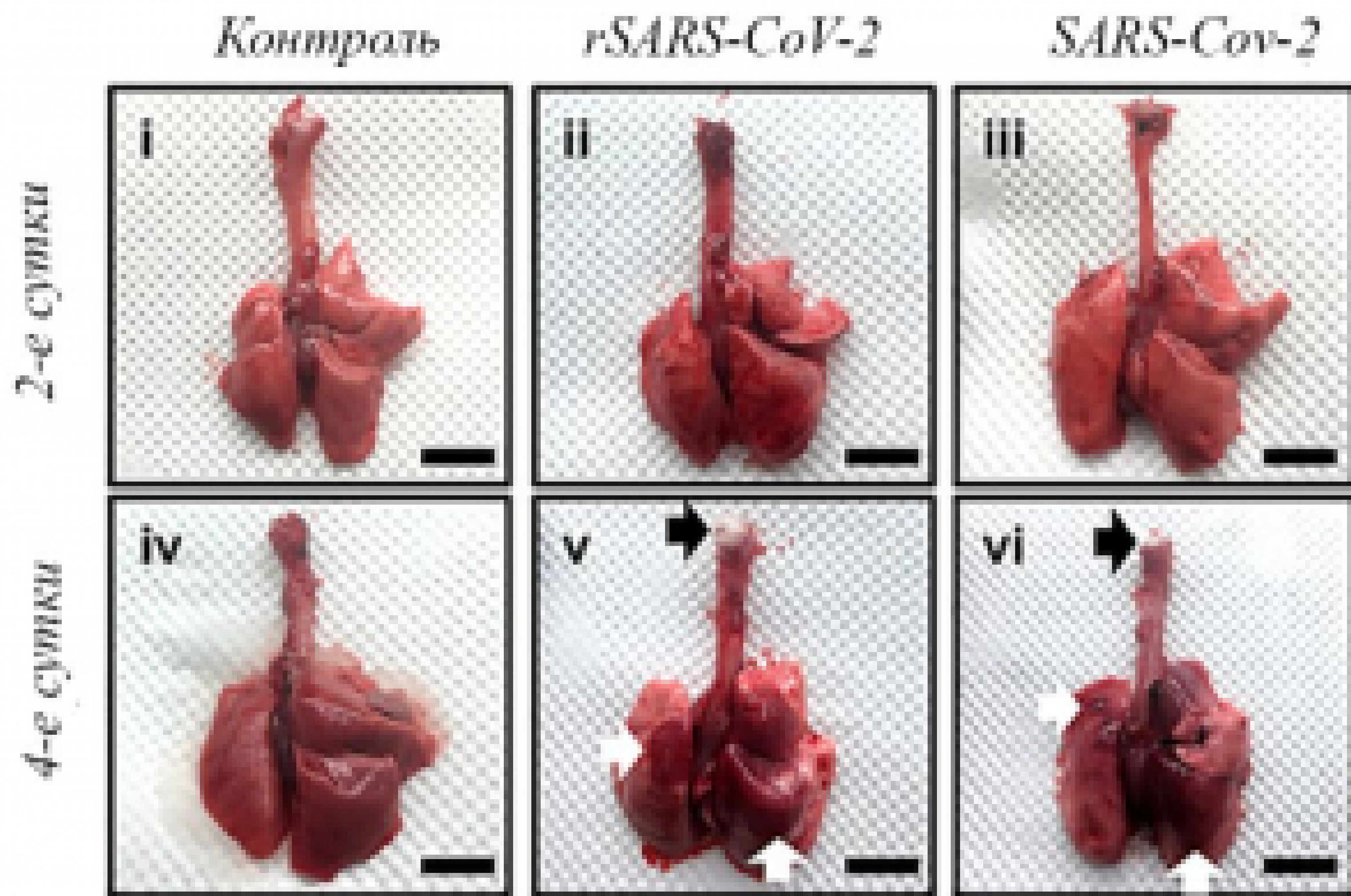


Figure 7 – *Comparison of the lethality of recombinant and natural SARS-CoV-2. In animals infected with rSARS-CoV-2 (ii) and SARS-CoV-2 (iii), moderate multifocal hyperemia and induration were observed in the lungs on the 2nd day after infection. Gross pathological lesions in the lungs were expressed on the 4th day after infection, with severe multifocal or locally extensive congestion and consolidation (white arrows) on 40–50% of*

lung surfaces (v and vi). These lesions were widespread, involving both the right (cranial, medial, and caudal lobes) and left lobe of the lungs. In particular, the presence of foamy exudate (black arrows) in the trachea of hamsters infected with rSARS-CoV-2 or SARS-CoV-2 on day 4 post-infection indicates ongoing bronchopneumonia.

Received Ye Ch. et al. [7], the results of experiments on intrazonal infection of hamsters showed the impossibility of distinguishing lesions caused by natural isolates of SARS-CoV-2 and its synthetic copies by clinical and pathomorphological features.

Creation of traceable strains of the SARS-CoV-2 virus. Such strains are created by incorporating the auxiliary protein ORF7a of the gene cassette virus with the marker gene into the frame. The gene cassette provides the marker gene with efficient expression without detrimental effect on the replicating virus. The greatest advantage among researchers is fluorescent and bioluminescent markers. The molecular weight of currently used marker proteins is in the range of 26.7–19.0 kDa. As a rule, SARS-CoV-2 traceable strains constructed by reverse genetics do not differ in their cultural properties from natural isolates [4–8, 29]. Using nano-luciferase as a marker of virus replication A. Pickard et al. [29] identified 35 drugs that inhibit the replication of SARS-CoV-2 in cells *Vero and human hepatocytes (amodiaquine, atovaquone, bedaquiline, ebastine, LY2835219,* manidipine , panobinostat, vitamin D3 , etc.). Thus, this direction in the design of synthetic SARS-CoV-2 is promising for the accelerated selection of

drugs that have a therapeutic effect in COVID-19. The very method of screening samples with the desired properties by the brightness of the marker is easy to automate.

Experiments with synthetic coronaviruses of farm animals. In recent years, coronaviruses from farm animals have become involved in the construction of coronavirus chimeras. The results of individual experiments made it possible not only to show the possibility of increasing the virulence of synthetic coronaviruses, but also to detect changes in the pathogenesis of the disease caused by them. For example, the introduction of a furin site into the S-protein of the chicken bronchitis virus (IBV) [27] , which causes damage to the respiratory organs, reproductive organs and nephrosonephritis syndrome in young animals, switched its tropism from the cells of the respiratory tract and genitourinary system to cells of the central nervous system . How this is done is shown in Figure 8.

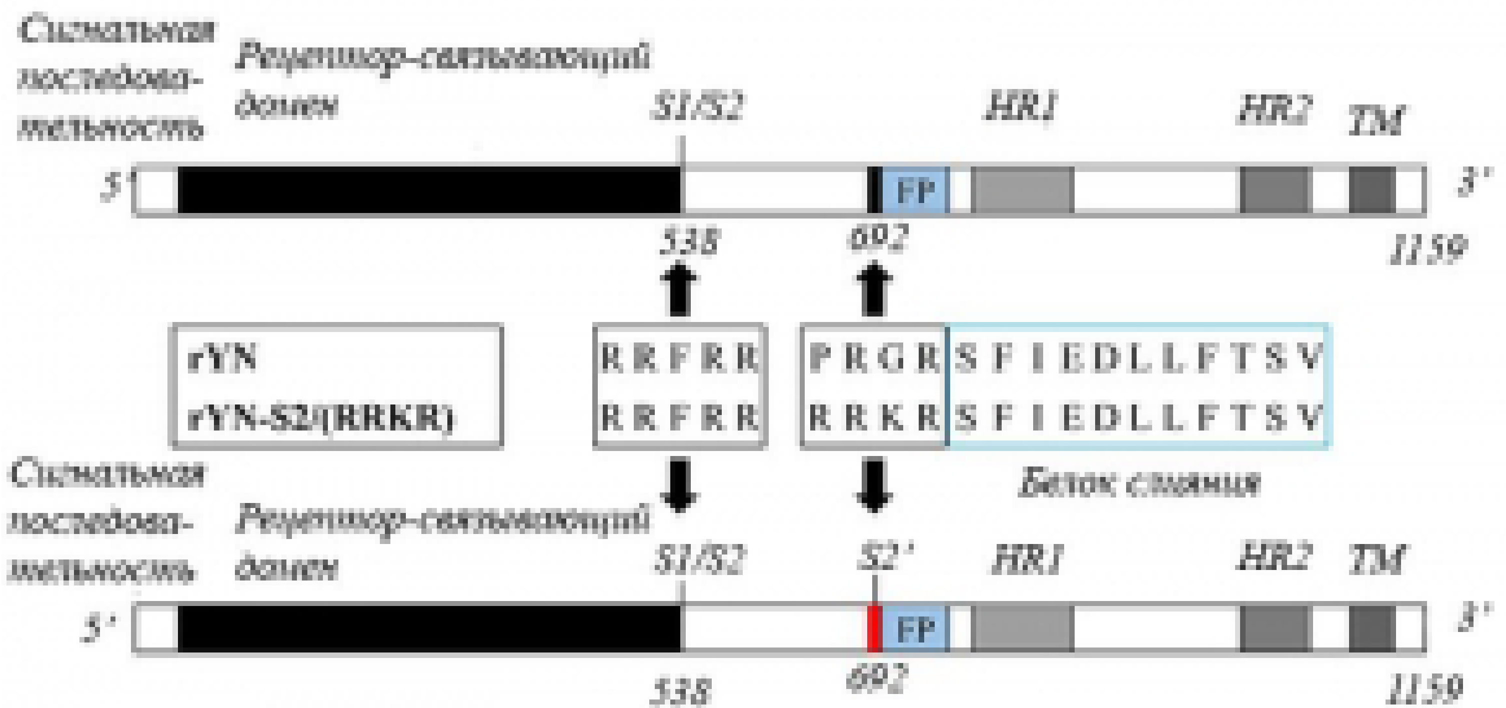


Figure 8 – *Production of synthetic chicken bronchitis virus rYN-S2/RRKR. A schematic representation of the S protein is shown. S protein is a fusion protein that mediates attachment to the host receptor. It is normally cleaved by the host cell's furin-like protease into two distinct polypeptides (subunits): S1 (left) and S2 (right). S1 is the receptor-binding domain of the S protein. The S2 subunit initiates the entry of the coronavirus into the cell. It includes a fusion protein (FP), a central helix (CH), a binding domain (CD), and a heptad repeat domain (HR1/2). Mediates integration between the viral membrane and the*

host cell membrane. After the S-protein binds to the receptor, it undergoes further conformational changes, allowing the cell proteases to sequentially cleave it at two sites: first at the S1/S2 interface (i.e. at the S1/S2 site - shown by arrows), which leads to the cleavage of S1 from S2 and its penetration into the blood. In the S2' site of the molecular clone of strain IBV-YN rYN-S2/RRKR, the PTKR (proline–threonine–lysine–arginine) sequence located before FP was replaced by a sequence for furin cleavage, which also consists of four amino acids, RRKR (arginine–arginine– lysine-arginine). The IBV-rYN strain is an infectious molecular clone of the IBV-YN strain (GenBank accession number: JF893452), parental to rYN-S2/RRKR, used as a control. FP, S2 fusion protein; HR1/2, heptad repeats; TM, S2 transmembrane domain [32]. also consisting of four amino acids - RRKR (arginine-arginine-lysine-arginine). The IBV-rYN strain is an infectious molecular clone of the IBV-YN strain (GenBank accession number: JF893452), parental to rYN-S2/RRKR, used as a control. FP, S2 fusion protein; HR1/2, heptad repeats; TM, S2 transmembrane

domain [32]. also consisting of four amino acids - RRKR (arginine-arginine-lysine-arginine). The IBV-rYN strain is an infectious molecular clone of the IBV-YN strain (GenBank accession number: JF893452), parental to rYN-S2/RRKR, used as a control. FP, S2 fusion protein; HR1/2, heptad repeats; TM, S2 transmembrane domain [32].

The synthetic rYN-S2/RRKR virus proved to be more fatal to 10-day-old chicken eggs compared to its parent strain rYN. Inoculation of rYN-S2/RRKR resulted in the death of all embryonic eggs within 36 hours, while the rYN strain required more than 96 hours for the same result. In addition, the 50% embryo infectious dose (EID_{50}) rYN-S2/RRKR in eggs was approximately ten times less than that of rYN. When evaluating the pathogenicity of rYN-S2/RRKR, it was found that lethal outcomes during rYN inoculation in chickens reached 10%, the clinical picture was typical of chicken bronchitis; while chickens in the rYN-S2/RRKR inoculated group developed neurological signs such as tremors and paralysis. They have never been reported before. The lethality of infected chickens was 90%. Chickens that died after infection with rYN showed obvious lesions of the respiratory and urinary systems, including mucus and petechial hemorrhage in the larynx, significant deposits of urate in the larynx, swelling of the ureter and kidneys. Chickens infected with the rYN-S2/RRKR strain had mucus and petechial hemorrhages in the larynx and had no obvious lesions in the kidneys. No brain lesions were found in the brain samples in the rYN groups and the negative control. CNS lesions were observed in the rYN-S2/RRKR group: significant microglial

hyperplasia, microglial nodule formation, and perivascular inflammatory infiltrates [32][28] .

This example shows the consequences of changing just two amino acids at the furin protease cleavage site. As a result, there was such a significant increase in the virulence of IBV, a change in the pathogenesis and symptoms of chicken bronchitis that this disease could be mistaken for a previously unknown one, if one does not know the genetic history of the virus strain that caused it. In addition to bronchitis in chickens, the following causes great damage to animal husbandry: bovine coronavirus (BCoV) - causes respiratory infection and diarrhea in cattle; transmissible gastroenteritis virus (TGEV) and porcine epidemic diarrhea virus (PEDV) – cause diarrhea in pigs; hemagglutinating encephalomyelitis virus (PHEV) - causes vomiting and malnutrition in pigs.

Exhaustion of the possibilities of experiments with the S -protein of coronaviruses. Considering the achieved level of reverse genetics in the construction of coronaviruses, one cannot fail to notice that experiments with coronavirus spikes by the beginning of the current decade are close to the limit in their development, they began to be repeated [29]. And there are serious reasons for this. Each type of viral architecture has its own structural limitations and the limit of variation in the shape, size, or configuration of a viral particle that can be implemented using a specific set of structural proteins and their modifications. When this tolerance is exceeded, the result of constructing a viral particle becomes uncertain, the process of its assembly begins to make mistakes, and here the virus is not up to the "interspecies jump" with any spikes. In addition, all "chimeras" were created according to a well-known pattern, i.e. according to the nucleotide sequences of real viruses and their individual genes, even if they belong to different species. They were simply swapped and the amino acid sequences of the spike were adjusted for closer contact

with the target cell receptor (replacement of non-polar amino acids by polar ones),[30] . At the same time, the role of mutations in non-structural and accessory proteins in the pathogenesis and outcome of COVID-19 is less studied, but even the limited data available suggest that it is significant (Table 1).

Table 1 - The role of mutations in non-structural and accessory proteins of SARS-CoV-2 in the pathogenesis and outcome of COVID-19*

protein type	Mutation	Protein	Impact on the course of the disease
non-structural	L37F	NSP6	Mild illness
	F308Y	NSP4	Same
	A97V, P323L	NSP12	Severe illness
	S1197R, T1198K	NSP3	Same
	L71F	NSP7	Fatal outcome
Auxiliary	L84S	ORF8	Mild illness
	G196V	ORF3a	Same
	I33T	ORF6	Severe illness
	Q57H, G251V	ORF3a	Same
	S253P	ORF3a	Fatal outcome

* Based on the work of A. Nagy et al. [33].

Apparently, the next stage of reverse genetics in the construction of coronaviruses will be the tuning of *non-structural and auxiliary proteins* involved in the penetration of the virus into cell endosomes, suppression of the interferon activity of the infected cell, enhancement of virus replication in specific cells, penetration into other body media, etc. In the future, there will be a transition to completely synthetic coronaviruses, when "templates" and "consensus design" for constructing chimeras are no longer required, and the virus will be designed by a computer from scratch for specific tasks, of course, in order to obtain new vaccines and drugs.

***

The ability to create viruses using reverse genetics approaches is a powerful tool for answering important questions in the biology of viral infections. It allows understanding the mechanisms of viral infection, identifying viral and host factors, and the interactions that control entry, replication, assembly, and budding of viruses in cells. In addition, reverse genetics facilitates the design of recombinant viruses expressing reporter genes for use in cell screening assays or in vivo models of infection *to quickly* and easily identify prophylactic and therapeutic approaches to treat viral infections, as well as to create attenuated forms of viruses for them. use as attenuated vaccines.

Author Contribution

Idea and concept of an article, search and analysis of literature, writing an article, digital image processing.

Conflict of Interest Information

The author declares that the study was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Review Details

The article was openly reviewed by two reviewers, experts in the field. Reviews are in the editorial office of the journal.

List of sources

1. Young B., Curtis K., Baric R. Strategy for Systematic Assembly of Large RNA and DNA Genomes: Transmissible Gastroenteritis Virus Model // J. Virol. 2000. V. 74 (22). P. 10600–10611. https://doi.org/10.1128/JVI.74.22.10600-10611.2000

2. Young B., Denison M. Weiss S., Baric R. Systematic Assembly of a Full-Length Infectious cDNA of Mouse Hepatitis Virus Strain A59 // J. Virol. 2002. V. 76(21). P. 11065–11078. https://doi.org/10.1128/JVI.76.21.11065-11078.2002

3. Ren W., QuX., Li W., Shi1 Z. et al. Difference in Receptor Usage between Severe Acute Respiratory Syndrome (SARS) Coronavirus and SARS-Like Coronavirus of Bat Origin // J. Virol. 2008. V. 82(4). P. 1899–1907. https://doi.org/10.1128/JVI.01085-07

4. Thao TTN, Labroussaa F., Thiel V. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform // Nature. 2020. V. 582. P. 561–565. https://doi.org/10.1038/s41586-020-2294-9

5. Xie X., Muruato A., Lokugamage KG et al. An Infectious cDNA Clone of SARS-CoV-2 // Cell Host Microbe. 2020 May 13. V. 27(5). P. 841–848.e3. https://doi.org/10.1016/j.chom.2020.04.004

6. Xie X., Lokugamage KG, Zhang X. et al. Engineering SARS-CoV-2 using a reverse genetic system // Nat Protoc. March 2021; v. 16(3). P. 1761–1784. doi:10.1038/s41596-021-00491-8

7. Ye Ch., Chiem K., Park JG et al. Rescue of SARS-CoV-2 from a Single Bacterial Artificial Chromosome // mBio. 2020 Sep-Oct; 11(5): e02168-20. https://doi.org/10.1128/mBio.02168-20

8. Rihn S., Merits A., Bakshi S. et al. A plasmid DNA-launched SARS-CoV-2 reverse genetic system and coronavirus toolkit for COVID-19 research // PLoS Biol. Feb. 2021 v. 19(2). e3001091. doi:10.1371/journal.pbio.3001091

9. Menachery VD, Yount BL, Debbink K., Baric R. et al. SARS-like cluster of circulating bat coronavirus pose threat for human emergence // Nat. Med. 2015. V. 21, No. 12. P. 1508–1513. https://doi.org/10.1038/nm.3985/ (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4797993/)

10 Lakota Jan. Today`s Biothreats - Where the Past Predictions Meet the Future // Journal of NBC Protection Corps. 2020. V. 4. No 4. P. 421–440. https://doi.org/10.35825/2587-5728-2020-4-4-421-430

11 Lakota Jan. Synthetic Biology – Friend or Foe? What Kind of Threats Should We Expect? // Journal of NBC Protection Corps. 2021. V. 5. No. 2. P. 103–122. https://doi.org/10.35825/2587-5728-2021-5-2-103-122

12. Godeke GJ., de Haan CAM, Rossen JWA et al. Assembly of spikes into coronavirus particles is mediated by the carboxy-terminal domain of the spike protein // J. Virol. 1999. V. 74. P. 1566–1571. https://doi.org/10.1128/jvi.74.3.1566-1571.2000

13. Opstelten DJ.E., Raamsman MJB, Wolfs K. et al. Envelope glycoprotein interactions in coronavirus assembly // J. Cell Biol. 1995. V. 131. P. 339–349. https://doi.org/10.1083/jcb.131.2.339

14. Kuo L., Godeke GJ., Raamsman MJ, Masters P., Rottier PGM Retargeting of Coronavirus by Substitution of the Spike Glycoprotein Ectodomain: Crossing the Host Cell Species Barrier // J. Virol. 2000. V. 74(3). P. 1393–1406. https://doi.org/10.1128/jvi.74.3.1393-1406.2000

15. Perez DR, Cockrell AS, Beall A. et al. Efficient Reverse Genetic Systems for Rapid Genetic Manipulation of Emergent and Preemergent Infectious Coronaviruses // Reverse Genetics of RNA Viruses. 2017; 1602:59–81. doi: 10.1007/978-1-4939-6964-7_5

16. Young B., Curtis KM, Fritz EA et al. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus // Proc Natl Acad Sci USA. 2003 . v. 100(22). P. 12995–13000. https://doi.org/10.1073/pnas.1735582100

17. Kouprina N., Larionov V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast Saccharomyces cerevisiae // Nat Protoc. 2008. V. 3(3). P. 371-377. https://doi.org/10.1038/nprot.2008.5

18. de Haan C., Li Z., Lintelo E. et al. Murine Coronavirus with an Extended Host Range Uses Heparan Sulfate as an Entry Receptor // J. of Virol. 2005. V. 79(22). P. 14451–14456. https://doi.org/10.1128/JVI.79.22.14451-14456.2005

19. Follis KE, York J., Nunberg JH Furin cleavage of the SARS coronavirus spike glycoprotein enhances cell–cell fusion but does not affect virion entry // Virology. 2006. V. 350(2). P. 358–369. https://doi.org/10.1016/j.virol.2006.02.003

20. Eckert DM, Kim PS Mechanisms of Viral Membrane Fusion and Its Inhibition // Annual Review of Biochemistry. 2001. V. 70. P. 777–810. https://doi.org/10.1146/annurev.biochem.70.1.777

21. Watanabe R., Matsuyama S., Shirato K. et al. Entry from the Cell Surface of Severe Acute Respiratory Syndrome Coronavirus with Cleaved S Protein as Revealed by Pseudotype Virus Bearing Cleaved S Protein // J. Virol. 2008. V. 82(23). P. 11985–11991. https://doi.org/10.1128/JVI.01412-08

22. Rockx B., Sheahan T., Donaldson E., Baric R. et al. Synthetic reconstruction of zoonotic and early human severe acute respiratory syndrome coronavirus isolates that produce fatal disease in aged mice // J. Virol. 2007. V. 81. P. 7410–7423. https://doi.org/10.1128/JVI.00505-07

23. Becker MM, Graham RL, Donaldson F., Baric R. et al. Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice // Proc. Natl. Acad. sci. USA. 2008. V. 105(50). P. 19944–19949. https://doi.org/10.1073/pnas.0808116105

24. Sheahan T. Rockx B., Donaldson E. et al. Mechanisms of zoonotic severe acute respiratory syndrome coronavirus host range expansion in human airway epithelium // Journal of virology. 2008. V. 82. P. 2274–2285. https://doi.org/10.1128/JVI.02041-07

25. Bolles M., Deming D., Long K. et al. A double-inactivated severe acute respiratory syndrome coronavirus vaccine provides incomplete protection in mice and induces increased eosinophilic proinflammatory pulmonary response upon challenge // J. Virol. 2011. V. 85. P. 12201–12215. https://doi.org/10.1128/JVI.06048-11

26. Li F. Structure, Function, and Evolution of Coronavirus Spike Proteins // Annu Rev Virol. 2016. V. 29; P. 237–261. https://doi.org/10.1146/annurev-virology-110615-042301

27. Menachery V., Yount B., Sims A. et al. SARS-like WIV1-CoV poised for human emergence // PNAS. 2016. V. 113 (11). P. 3048–3053. https://doi.org/10.1073/pnas.1517719113

28 Sia SF, Yan LM, Chin AWH et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters // Nature. 2020. V. 583. P. 834–838. https://doi.org/10.1038/s41586-020-2342-5.

29. Pickard A., Calverley B., Chang J. et al. Discovery of re-purposed drugs that slow SARS-CoV-2 replication in human cells // PLoS Pathog. 2021 Sep 9; V.17(9). e1009840. doi:10.1371/journal.ppat.1009840

30. Medical aspects of chemical and biological warfare / Edit. FR Sidell, ET Tafuqi, DR Franz. Washington, 1997 .

31. Jackson PJ, Ramsay AJ, Christensen CD et al. Expression of mouse interleukin-4 by a recombinant ectromelia virus suppresses cytolytic lymphocyte responses and overcomes genetic resistance to mousepox // J. Virol. 2001. V. 75, No 3. P. 1205–1210. https://doi.org/:10.1128/JVI.75.3.1205-1210.2001

32. Cheng J., Zhao Y., Xu G. et al. The S2 Subunit of QX-type Infectious Bronchitis Coronavirus Spike Protein Is an Essential Determinant of Neurotropism // Viruses. 2019. V. 11(10). 972. https://doi.org/10.3390/v11100972

33. Nagy A., Pongor S., Gyorffy B. Different mutations in SARS-CoV-2 associate with severe and mild outcome // Int. J. Antimicrob. agents. 2021. V. 57. 106272. https://doi.org/10.1016/j.ijantimicag.2020.106272

34. Deigin Y., McCairn K., Sousa A. et al. Have artificial lighting and noise pollution caused zoonosis and the COVID-19 pandemic? A review // Environ. Chem. Lett. 2021. Jul 31: 1–10. doi: 10.1007/s10311-021-01291-y

35. Markson S. What really happened in Wuhan. Harper Collins Publishers Australia Pty Limited. ISBN 978 1 4607 6092 5. 2021.

about the author

Federal State Budgetary Institution "27 Scientific Center" of the Ministry of Defense of the Russian Federation, 105005, Russian Federation, Moscow, Brigadirsky pereulok, 13.

Supotnitsky Mikhail Vasilievich. Chief Specialist, Cand. biol. Sciences, art. scientific collaborator

Contact person: Supotnitsky Mikhail Vasilyevich; 27nc_1@mil.ru

Modern approaches to molecular genetics of viruses in the study of the members of the family Coronaviridae

MV Supotnitskiy

Federal State Budgetary Establishment "27 Scientific Centre" of the Ministry of Defense of the Russian Federation, Brigadirskii Lane 13, Moscow 105005, Russian Federation

Received 01 September 2021. Accepted 23 September 2021.

The existence of certain suspicions about the artificial origin of the COVID-19 pandemic and about the possible use of reverse genetic technology to create the SARS-CoV-2 virus require an understanding of its capabilities in the design of new viruses. The *aim of this work* is to

show how the use of reverse genetics allows the design of previously non-existent coronaviruses, technologies and the main achievements in their creation. Only the information in the public domain was used for the preparation of this article. The technology is called "reverse genetics" because when obtaining RNA viruses capable of replication, the process is going not from DNA to RNA, as usual, but on the contrary, from the RNA of the virus to its complementary DNA (cDNA), and from it with the help of T7 RNA polymerase – “back” to the infectious RNA. Since the resulting plus-RNA of the coronavirus genome mimics cellular messenger RNA (mRNA), it is immediately recognized by the cell's translation machine and triggers the formation of its own infectious viral particles. Two systems of reverse genetics have been developed, *in vitro* and in vivo. The problem of obtaining a full-length cDNA of the giant genome of coronaviruses is solved by fragmentation and subsequent stitching of fragments using standard molecular biology approaches. The article provides the examples of how this technology makes it possible to obtain synthetic coronaviruses that are indistinguishable from those isolated from nature, to change the range of their hosts, to enhance virulence and resistance to specific antibodies, and to influence the pathogenesis of the disease . The article also shows the prospects for the use of recombinant viruses in cellular screening analyses and infection models in vivo for the identification of preventive and therapeutic approaches to the virus disease treatment.

Key words: *atypical pneumonia; with oronavirus; interspecies transmission; reserve genetics; pandemic; COVID-19; SARS-CoV-2*

For citation : *Supotnitskiy MV Modern approaches to molecular genetics of viruses in the study of the members of the family Coronaviridae // Journal of NBC Protection Corps. 2021. V. 5. No. 3. P. 217–235. https://doi.org/10.35825/2587-5728-2021-5-3-217-235*

conflict of interest statement

The author declare that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Peer review information

The article has been peer reviewed by two experts in the respective field. Peer reviews are available from the Editorial Board.

References . See p. 233–234.

Authors

Federal State Budgetary Establishment "27 Scientific Center" of the Ministry of Defense of the Russian Federation. Brigadirskii Lane 13, Moscow 105005, Russian Federation.

Mikhail Vasilyevich Supotnitskiy. senior researcher. chief specialist. Candidate of Biological Sciences.

Contact person: Mikhail Vasilievich Supotnitskiy; 27nc_1@mil.ru

[1] The institute was founded in 1956 as the Wuhan Laboratory of Microbiology of the Chinese Academy of Sciences. In January 2015, a French contractor from Lyon built the first laboratory in China with a BSL-4 protection level at the institute. The contract value is USD 44 million. Employees were trained at the BSL-4 laboratory in Galveston, Texas (The Galveston National Laboratory in Galveston, Texas, United States). The new laboratory building has 3000 m2 of BSL-4 space, as well as 20 BSL-2 laboratories and two BSL-3 laboratories. BSL-4 facilities were accredited by the China National Accreditation Service for Conformity Assessment (CNAS) in January 2017, and the BSL-4 lab became operational in January 2018. See https://en.wikipedia.org/wiki/Wuhan_Institute_of_Virology (accessed 25.05.2021).

[2] Details of the manipulations with the genome of coronaviruses preceding the COVID-19 pandemic can be found in the scientific "essay" by Russian scientist Yuri Deigin. In the same work, Yuri Deigin proposes approaches to determining signs of interference in the virus genome. See Deigin Y. Lab-Made? SARS-CoV-2 Genealogy through the Lens of Gain-of-Function Research. https://yurideigin.medium.com/lab-made-cov2-genealogy-through-the-lens-of-gain-of-function-research-f96dd7413748 (Accessed: 06/16/2021).

[3] The name of the new disease and the virus that caused it was determined by WHO on February 11, 2020 as follows: the *causative agent* of the disease is severe acute respiratory syndrome

coronavirus 2 (SARSCoV-2). *The disease* is coronavirus disease 2019 (Coronavirus Disease 2019, COVID-19).

[4] Work performed at David Axelrod Institute, Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany (New York 12201) and Institute of Virology, Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, and Institute of Biomembranes, Utrecht University (3584 CL Utrecht, The Netherlands).

[5] cDNA (complementary DNA, cDNA) is DNA synthesized on a mature mRNA template in a reaction catalyzed by reverse transcriptase.

[6] Reverse transcription polymerase chain reaction (RT-PCR) is a method for amplifying a specific fragment of ribonucleic acid (RNA). A single-stranded RNA molecule is converted in reverse transcription reactions (RT, English RT, reverse transcription) into complementary DNA (cDNA) and then the already single-stranded DNA molecule is amplified using traditional PCR.

[7] Work performed at the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; Department of Pediatrics and Microbiology and Immunology, Elizabeth B. Lamb Center for Pediatric Research, Vanderbilt University Medical Center, Nashville, Tennessee; Department of Microbiology, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; Department of Microbiology and Immunology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

[8] The work was done in the Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston TX, USA. It is also unique in that on 40 pages, using SARS-CoV-2 as an

example, a detailed description of 108 stages of obtaining synthetic coronaviruses is given, indicating possible errors and critical points in their implementation. For example, the section "TROUBLESHOOTING" (search and elimination of errors) is worth looking at just out of curiosity [6]. In the same institution, Chinese scientists from the Institute of Virology in Wuhan were trained (see https://en.wikipedia.org/wiki/Wuhan_Institute_of_Virology ; accessed 05/25/2021).

[9] The work was done at the 1Institute of Virology and Immunology (IVI), Switzerland; Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, Switzerland; Graduate School for Biomedical Science, University of Bern, Switzerland; Institute of Veterinary Bacteriology, Vetsuisse Faculty, University of Bern, Switzerland; Institute for Infectious Diseases, University of Bern, Switzerland; Department for Molecular and Medical Virology, Ruhr-Universität Bochum, Germany; Institute of Virology, Charite-Universitäts corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Germany; German Center for Infection Research, associated partner Charite, Berlin, Germany; Institute of Medical Parasitology, Tropical and Transmissible Diseases. E.I. Marcinovsky, Moscow,

[10] Yeast artificial chromosomes (YAC) are the chromosomes of the yeast *Saccharomyces cerevisiae*, designed for gene cloning. The main components of the YAC are the autonomously replicating sequence, *S. cerevisiae* centromere and telomeres . Genes for selectable markers such as antibiotic resistance or imaging markers are used to select transformed yeast cells. DNA fragments ranging in size from 100 to 1000 kb can be inserted into such a chromosome.

[11] Work performed at Texas Biomedical Research Institute, San Antonio, Texas, USA; Department of Veterinary Microbiology,

University of Ilorin, Ilorin, Nigeria; Department of Molecular and Cell Biology, Centro Nacional de Biotecnologia (CNB-CSIC), Madrid, Spain; Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, California, USA.

[12] A bacterial artificial chromosome (BAC) is a DNA construct based on a functional fertility plasmid (or F-plasmid) used to transform and clone bacteria, usually *E. coli* . F-plasmids contain separation genes that promote even distribution of plasmids after bacterial cell division. The usual size of a bacterial artificial chromosome insert is 150–350 kb.

[13] pBeloBAC is a single-copy plasmid vector that replicates in Escherichia coli. Designed to create YOU. For a detailed description see https://www.snapgene.com/resources/plasmid-files/?set=basic_cloning_vectors&plasmid=pBeloBAC11 (Accessed 07/20/2021)

[14] A plaque forming unit (PFU) is a plaque on a continuous cell culture lawn caused by the lytic action of a single infectious viral unit.

[fifteen] The work was done at the MRC-University of Glasgow Center for Virus Research (CVR), Glasgow, United Kingdom; Institute of Technology, University of Tartu, Tartu, Estonia; MRC Protein Phosphorylation and Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dundee, United Kingdom; Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, United Kingdom; Emerging Viruses, Inflammation and Therapeutics Group, Menzies Health Institute Queensland, Griffith University, Gold Coast, Queensland, Australia; Division of Biological Sciences, College of Arts and Sciences, University of the Philippines Visayas, Miagao, Iloilo, Philippines; West of Scotland Specialist Virology Centre, Glasgow, United Kingdom; Indian Immunologicals Ltd (IIL), Rakshapuram, Gachibowli Post, Hyderabad Telangana, India;

Drug Discovery Unit (DDU), Wellcome Center for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, United Kingdom; School of Medical Sciences, Griffith University, Gold Coast, Queensland, Australia.

[16] Hepatitis delta virus (HDV) ribozyme is a non-coding RNA found in the HDV genome that is required for its replication. HDV is the only known human virus that uses the activity of a ribozyme to infect its host. The ribozyme is active *in vivo* in the absence of any protein factors.

[17] mCherry is a protein belonging to the group of monomeric red fluorescent proteins mFruits. Its natural precursor protein DsRed (RFP) was isolated from corals of the Discosoma genus of the sea anemone order. Unlike DsRed, mCherry is a monomeric protein and its fluorescence is more stable than that of its natural precursor. The mCherry gene contains 711 base pairs. The protein includes 236 amino acids, MM 26.7 kDa.

ZsGreen is an exceptionally bright green fluorescent protein derived from the rift coral Zoanthus sp. (MM 26.1 kDa).

NLuc is a mutant form of luciferase (171 amino acids, MM 19 kDa), which has better characteristics than the wild-type protein.

[18] 2A-peptides (eng. *2A self-cleaving peptides*) are a group of peptides with a length of 18 to 22 amino acids. Separate two proteins, self-excision from the polypeptide chain occurs after translation. It is assumed that when 2A peptides are excised, not a real proteolytic cleavage occurs, but a “breakthrough” of the ribosome, as a result of which the peptide bond between the glycine and proline residues is simply not formed.

[19] Yuri Deygin drew attention to the fact that the name Shi Zhengli appeared for the first time in this article. Apparently in 2005 she trained under the guidance of Peter Rottier in Utrecht. (Faculty of Veterinary Medicine and Institute of Biomembranes, Utrecht University, The Netherlands). This institute was listed as her affiliation.

[20] ACE2 (angiotensin-converting enzyme 2, ACE2) is an entry receptor. Allows SARS-like viruses to bind to the cell and initiate their entry into the cell. Found on the surface of many human cells, including cells of the nasal cavity (olfactory epithelium), lungs (ciliated cells of the bronchial epithelium and pneumocytes), kidneys (cells of the proximal tubules of the kidneys), bladder (urothelial cells), pancreas (islets of Langerhans), intestines (epithelial cells of the ileum and esophagus), brain (neurons and glial cells), heart (myocardium), blood vessels (endothelium).

[21] See T. Sheahan et al. [24].

[22] The work was carried out under conditions of biosafety level BSL3. It is used to work with microorganisms that can cause potentially fatal diseases if the aerosol is inhaled. For more information on biosafety levels see https://en.wikipedia.org/wiki/Biosafety_level#Biosafety_level_3 (Accessed 03/04/2020).

[23] The work at the end of 2015 caused a debate about the safety of the described experiments. Simon Wain-Hobson, a virologist at the Pasteur Institute in Paris, said the researchers have created a new virus that "grows remarkably well" in human cells. “If the virus breaks out, no one will be able to predict the trajectory.” See https://www.nature.com/news/engineered-bat-virus-stirs-debate-over-

risk-research-1.18787 (Accessed 03/01/2020). It is even less reliable to "predict the trajectory" if such experiments are not carried out.

[24] For more details on the structure, function, and evolution of the S protein, see F. Li [26].

[25] By reducing the number of plaques on Vero E6 cells after co-incubation of mAb and virus.

[26] The composition of research organizations is also interesting: *Departments of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC (USA)*; Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC (USA); National Center for Toxicological Research, Food and Drug Administration, Jefferson, AR (USA); Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill, Chapel Hill, NC (USA); Marsico Lung Institute/Cystic Fibrosis Center, University of North Carolina at Chapel Hill, Chapel Hill, NC (USA); Institute for Research in Biomedicine, Bellinzona, Switzerland Institute of Microbiology, ETH Zurich, Zurich (Sweden); Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute (USA); Department of Medicine, Harvard Medical School, Boston Massachusetts (USA); *Key Laboratory of Special Pathogens and Biosafety, Wuhan Institute of Virology, Chinese Academy of Sciences (Yuhan, PRC)* . The Vero E6 cells used to construct the chimeric SHC014-CoV virus were obtained from USAMRIID (US Army Medical Research Institute of Infectious Diseases, Fort Detrick, MD, USA). At the end of the list of authors - *Zhengli - Li Shi* and *Ralph S Baric* - the founders of the Chinese and American schools for creating coronavirus chimeras.

[27] Refers to gamma-coronaviruses (γ-CoV).

[28] The work was done at the Key Laboratory of Animal Epidemiology of the Ministry of Agriculture, College of Veterinary Medicine, China Agricultural University, Beijing 100193, China.

[29] See [9] and [27].

[30] The possibility of artificial origin of SARS-CoV-2 was discussed in detail by Y. Deigin et al. [34]. As an instrument of political pressure on China, the version of the artificial origin of the virus is described in detail in the book by S. Markson [35].

Mikhail Vasilievich Supotnitsky

Russian microbiologist, colonel of the medical service of the reserve, inventor, author of books and articles on the history of plague epidemics and other especially dangerous infections, the history of the development and use of chemical and biological weapons. Deputy editor-in-chief of the scientific and practical journal "Bulletin of the

RCB Protection Troops" of the Ministry of Defense of the Russian Federation.