PHYLOGENETIC ANALYSIS OF NOROVIRUSES BASED ON RNA-DEPENDENT RNA POLYMERASE GII.P16 GENE SEQUENCES

S.V. Oparina[1,2*], *N.V. Epifanova*[1], *N.A. Novikova*[1]

[1] Academician IN Blokhina Nizhny Novgorod Scientific Research Institute of Epidemiology and Microbiology, Federal Service for Surveillance on Customers Rights Protection and Human Wellbeing, 71 Mallaya Yamskay St., Nizhny Novgorod, 603950, Russia;
[2] National Research Lobachevsky State University of Nizhny Novgorod, 23 Prospekt Gagarina, Nizhny Novgorod, 603950, Russia.

* Corresponding author: svetlanochka.o@mail.ru

Abstract. Noroviruses with the GII.P16 polymerase gene have long been considered rare. However, since 2015, there have been reports from different geographical regions about the spread of new recombinant strains of norovirus in which GII.P16 polymerase is associated with a capsid protein of various genotypes. In the autumn of 2016, a sharp increase in the frequency of detection of noroviruses was observed in Nizhny Novgorod, which coincided with the appearance of new recombinants GII.4_Sydney and GII.2 – with GII.P16 polymerase. Based on the sequences of the RNA-dependent RNA polymerase gene present in the GenBank database, a phylogenetic analysis of noroviruses with GII.P16 specificity was performed by constructing Bayesian phylogenetic trees. Analysis of amino acid sequences showed that representatives of the novel variant of the GII.P16 polymerase gene had five substitutions compared to earlier strains. These substitutions are located near sites responsible for the enzyme activity of polymerase and could affect the transmissivity of the virus. Acquisition of a novel variant of the GII.P16 polymerase gene by noroviruses with different capsid protein genotypes probably provides certain advantages for recombinants and creates prerequisites for their wide distribution.

Keywords: norovirus, polymerase, GII.P16, recombinant genotypes, phylogenetic analysis.

List of Abbreviations

RdRp – RNA-dependent RNA polymerase
VP – Viral protein
ORF – open reading frame
NS – Nonstructural protein
CDC – Centers for Disease Control and Prevention

Introduction

Noroviruses belong to the family *Caliciviridae*, genus *Norovirus*. Caliciviruses infect a wide range of vertebrate species, including humans, causing acute gastroenteritis. According to the new classification, noroviruses are divided into ten genogroups (GI-GX). Based on the analysis of the amino acid sequence of the main capsid protein VP1, 48 genotypes are isolated. Based on nucleotide diversity in the RdRp region, noroviruses are divided into 60 P-types (Chhabra *et al*., 2019; Chhabra *et al*., 2020).

The human norovirus genome is a linear single-stranded RNA of positive polarity, approximately 7.5-7.7 kb in length, is organized into three open reading frames (ORFs) (Bull *et al*., 2010). ORF1 encodes a large polyprotein, precursor of six non-structural proteins (NS1/2 – NS7), including RNA-dependent RNA polymerase (Lee *et al*., 2019). ORF2 encodes the major structural protein of the VP1 capsid, while ORF3 encodes the minor structural protein of the capsid VP2, which is located inside the viral particle and is involved in the assembly of the capsid and encapsulation of the genome (Vongpunsawad *et al*., 2013).

At the junction of reading frames encoding non-structural and structural proteins, there is a «hot spot» in which recombinations often occur in the norovirus genome; therefore, a dual nomenclature was adopted, taking into account the genotypes of noroviruses in two reading frames. For typing, the ORF1 region encoding the RNA polymerase and the ORF2 regions encoding the N/S domain and the P domain of the capsid protein VP1 are most often used (Kroneman *et al*., 2013).

RNA-dependent RNA polymerase plays a critical role in the replication of the norovirus

genome. The region encoding the polymerase changes quite quickly, while there is a difference in the rate of accumulation of mutations for different genotypes (Ozaki *et al*., 2018). Noroviruses containing the GII.P16 polymerase gene have long been considered rare. However, since 2015, there have been reports from different geographical regions about the spread of new recombinant strains of norovirus in which GII.P16 polymerase is associated with a capsid protein of various genotypes (Bidalot *et al*., 2017; Ruis *et al*., 2017; Han *et al*., 2018; Hata *et al*., 2018, Zhirakovskaya *et al*., 2020).

In the autumn of 2016, an increase in the frequency of detection of noroviruses was observed in Nizhny Novgorod, which coincided with the appearance of new recombinants GII.4_Sydney and GII.2 – with GII.P16 polymerase.

The purpose of this work was to analyze recombinant genotypes of noroviruses carrying the gene of RNA-dependent RNA polymerase with the GII.P16 specificity, presented in the GenBank database.

Materials and Methods

The basis for the selection of sequences for the study was taken from the previously formed database of nucleotide sequences NoroGen, containing complete nucleotide sequences of the genome of noroviruses, as well as partial genome sequences of noroviruses identified in the territory of the Russian Federation, extracted from the GenBank database by the keywords «Norovirus complete genome» as of 2 June 2021 (Epifanova & Epifanov, 2021). The NoroGen database for this study was supplemented with full-genome sequences deposited in GenBank in the period from 3 June 2021 to 28 May 2022. Next, sequences of noroviruses with GII.P16 polymerase were filtered, uploaded in a format convenient for subsequent phylogenetic analysis, and partial sequences of GII.P16 noroviruses identified outside the Russian Federation that are not included in NoroGen were added. In total, the analysis included 400 sequences of noroviruses circulating in the territories of different countries in the period 1975–2021, including 241 full-genome and 159 partial sequences.

Computer analysis of nucleotide sequences to determine the norovirus genotype was performed using the web service for automatic norovirus genotyping Norovirus Genotyping Tool Version 2.0 (Kroneman *et al*., 2011, http://www.rivm.nl/mpf/norovirus/typingtool), for identification closely related strains of human noroviruses using the BLAST software package (Altschul *et al*., 1990). Nucleotide sequences alignment and genetic analysis were performed using MEGA X software, version 10.0.5 (Kumar *et al*., 2018).

Phylogenetic analysis using molecular clocks was carried out by constructing Bayesian phylogenetic trees based on the Markov chain Monte Carlo algorithm and the SRD06 substitution model implemented in BEAUti 1.10 programs.4. and BEAST 1.10.4. and visualized in FigTree v1.4.4 (Suchard *et al*., 2018).

The amino acid substitutions on the RdRp of GII.P16 strains were mapped on the structural model of the GII.4 virus (PDB number 4QPX (Zamyatkin *et al*., 2014)) using UCSF Chimera v 1.16 (Pettersen *et al*., 2004).

The protocol of the study was approved by the Local Ethics Committee of the Academician I.N. Blokhina Nizhny Novgorod Scientific Research Institute of Epidemiology and Microbiology of the Federal Service for Surveillance on Customers Rights Protection and Human Wellbeing the Russian Federation.

Results

Phylogenetic analysis

The full-genome nucleotide sequences of GII.P16 noroviruses selected for the study were labeled according to the year and country of detection of the isolate, as well as genotype by the capsid protein gene VP1. After alignment, regions with a length of 1533 nucleotides corresponding to the complete RNA-dependent RNA polymerase gene were used for further analysis. When constructing the phylogenetic tree (Fig. 1), these sequences were divided into three large groups – lines A, B and C, whose common ancestor existed around the end of the 60s of the last century.

The common ancestor of the noroviruses forming line A existed around 1977. Line A

contains sequences of non-recombinant noroviruses GII.16[P16], including isolate AY772730 (Hu/NLV/GII/Neustrelitz260/2000/DE), identified in 2000 in Germany and currently considered a reference strain for the type specificity of GII.16 both by the capsid gene and by the gene polymerases. Single noroviruses GII.16[P16], identified in 2005-2008 in Russia, Thailand, and Paraguay, are also located here. The remaining sequences of this line belong to recombinants, including single ones, with the type specificity of the capsid protein GII.5 (Russia, 2010), GII.17 (Japan, 2002, Great Britain, 2014) and more numerous with capsid GII.2 circulating in Japan from 2009 to 2016. and forming a clearly differentiated subline.

The common ancestor of noroviruses, whose sequences form the B line, existed around 2003. Line B contains one non-recombinant norovirus GII.16[P16], identified in Russia (Novosibirsk) in 2012, as well as recombinants with capsid GII.2 (Canada, USA, Japan, Brazil, 2011–2014), GII.3 (Russia, Brazil, USA, Peru, 2011–2013) and GII.13 (Thailand, USA, Japan, Russia, 2012–2016). At the same time, clustering of polymerase sequences on a subline is observed according to the genotypes of capsid protein associated with it.

Line C began to form around 2013 and immediately split into two large sublines. One of them, which we called «subline GII.2», includes sequences associated with the GII.2 capsid gene. These recombinant noroviruses, in all probability, did not evolve from the previously circulating recombinants GII.2[P16] belonging to the A and B lines, but arose anew as a result of a recombination event between the parent strains that have not yet been established. Noroviruses of this subline, detected since 2016, are widely circulating in the world and have been found in China, Japan, Thailand, Australia, the USA, Canada, Russia, and the UK.

The second subline of line C, which we conditionally called «subline GII.4», is heterogeneous and includes previously unknown recombinants GII.4[P16], which have been widely distributed in the world since 2016 due to the successful combination of a novel variant of polymerase P16 with a variant of the capsid protein GII.4_Sydney, which appeared in 2012. This subline includes sequences of noroviruses circulating in China, Japan, Thailand, USA, Argentina, Russia in 2016–2020. The same subline in the tree constructed on the basis of the complete sequence of the polymerase gene includes single recombinants associated with the capsid gene GII.1 (China, Russia, Canada, 2017–2019), GII.3 (Russia, 2017), GII.12 (USA, 2017–2018), GII.13 (Russia, 2017).

The divergence of nucleotide sequences between lines A, B and C ranges from 7% to 13%, within line A – 0.9–10%, B – 0.2–10%, C – 0.5–3.1%.

Analysis of the complete gene of any of the viral proteins is the most informative for studying the evolution of this gene. However, there are many partial sequences of certain viral genes in the GenBank database. To assess possible changes in the topology of the phylogenetic tree, a Bayesian tree was constructed based on the same sequences as in Figure 1, but containing only the 3′-end of ORF1 with a length of 264 nucleotides, which is usually used to determine the P-type of norovirus (Fig. 2).

It can be observed that clustering of polymerase lines persists. There were no significant changes in line A, line B in this tree does not represent a monophyletic group, but the compact arrangement of the sequences included in it has been preserved. In line C, there was a transfer of strains with the genotype of the capsid protein GII.1, GII.3, GII.12 and GII.13 from subline GII.4 to subline GII.2. This means that for this region of the polymerase gene, they turned out to be closer to noroviruses having a capsid protein of genotype GII.2. Sublinia GII.4 now contains only representatives of the GII.4 genotype VP1. Thus, the analysis of a short region of the polymerase gene allows for adequate differentiation of sequences into large clusters, but has limitations for assessing the relationships of some single sequences.

In order to extend the geographical representation of noroviruses and to cover more widely the time interval of their circulation, an expanded version of the tree was constructed based on the nucleotide sequences of a region

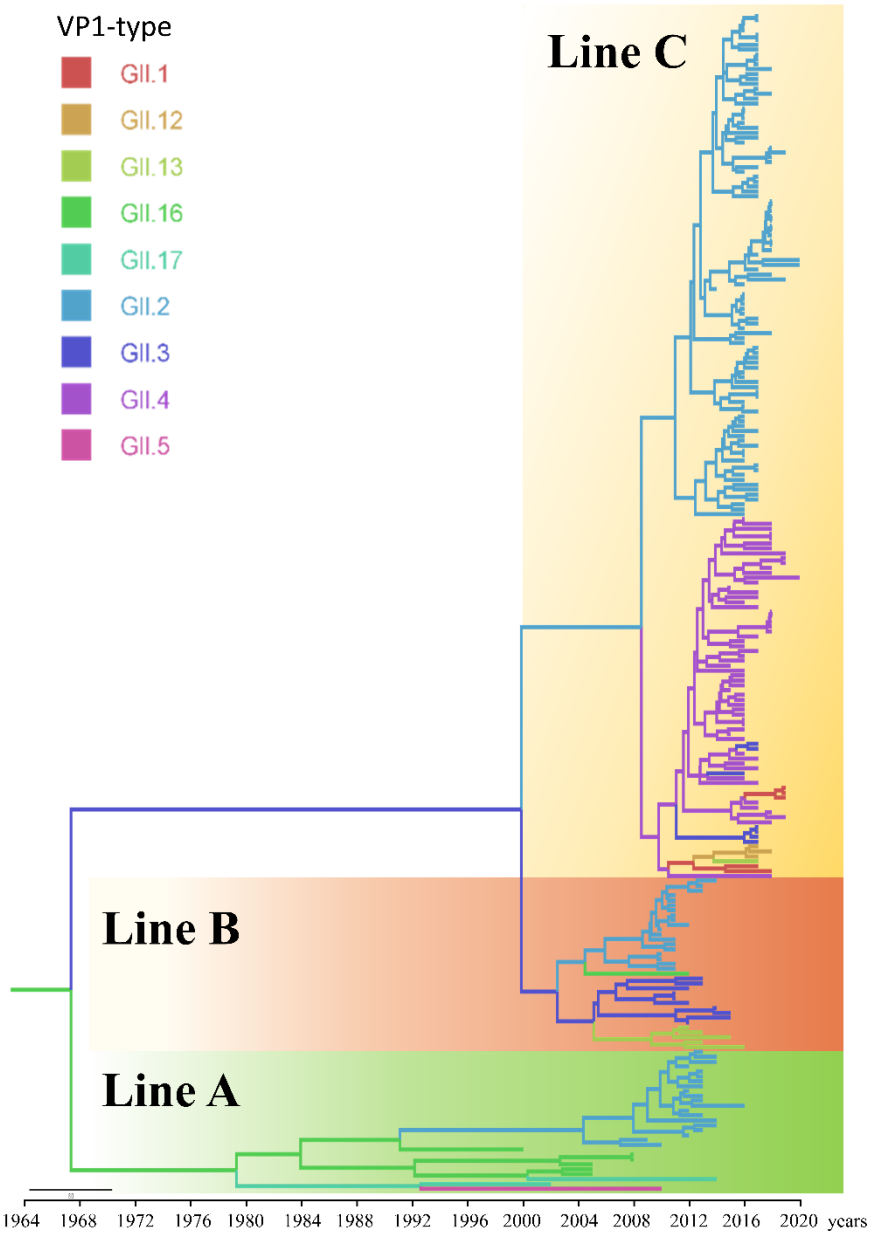

Fig. 1. Phylogenetic analysis based on the complete polymerase gene (1533 nucleotides) of noroviruses (n = 241). Phylogenetic tree was constructed by the BEAST software. Branches were colored by VP1-type. Phylogenetic clusters based on lines: line A is represented by green, line B is red and line C is orange

of the polymerase gene with a length of 264 nucleotides, including an additional 159 sequences, among which 22 sequences from Nizhny Novgorod. The total number of analyzed sequences was 400 (Fig. 3).

In the expanded tree, clustering on the polymerase line is traced according to a short region, while maintaining the general topology of the tree and increasing clusters by adding new isolates, including variants of the GII.10 genotype.

Additional sequences of GII.16 noroviruses joined Line A, including isolate AB684676 (Hu/GII/22-2/Tokyo/1975/JPN) identified in Japan in 1975, as well as isolates from France (1999) and China (2008). Among others, it is possible to note the GII.5 isolate from the USA (2005) and subline GII.2, which now contains,

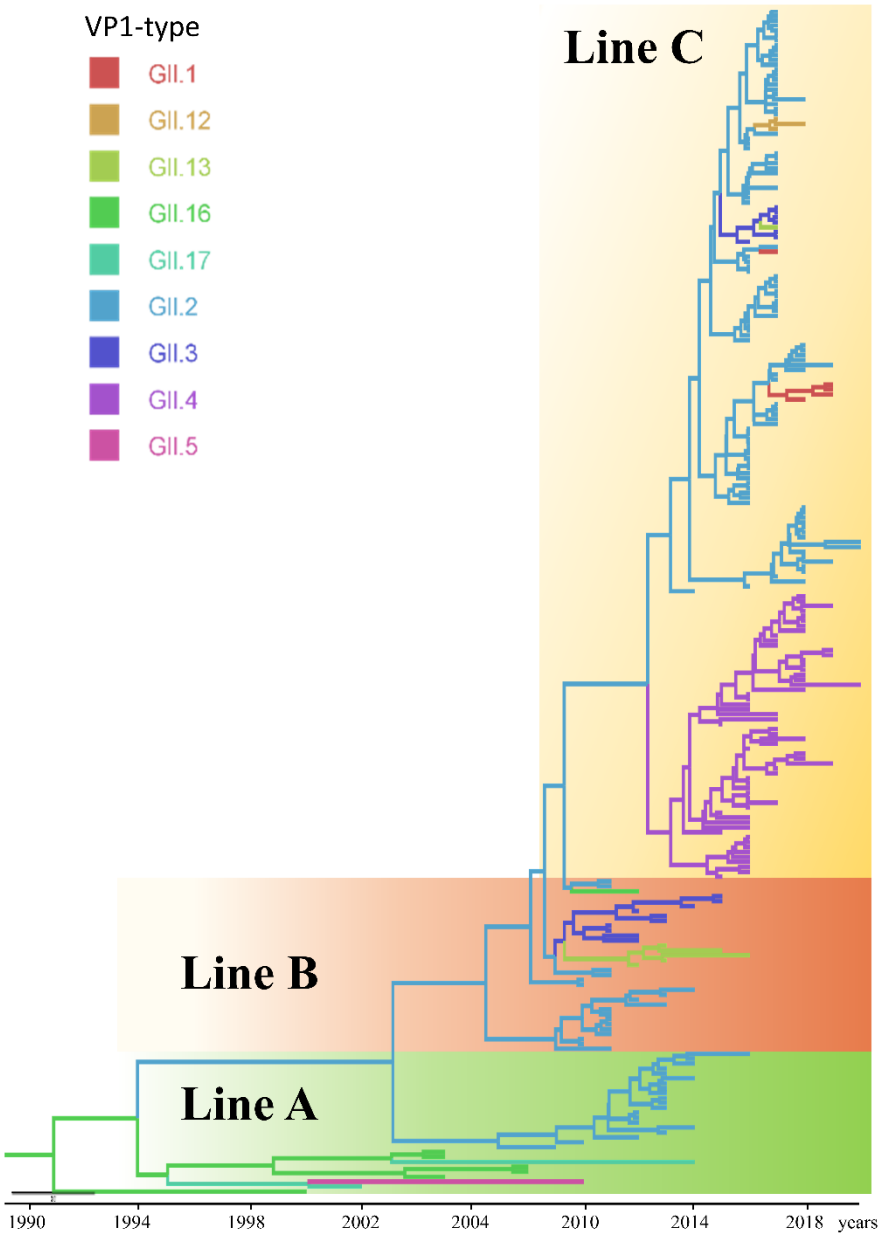

Fig. 2. Phylogenetic analysis based on the polymerase gene region (264 nucleotides) of noroviruses (n = 241). Phylogenetic tree was constructed by the BEAST software. The set of sequences is similar to the one shown in Fig.1. Branches were colored by VP1-type. Phylogenetic clusters based on lines: line A is represented by green, line B is red and line C is orange

in addition to isolates from Japan, also strains from Korea (2010, 2012), China (2010–2011) and Taiwan (2012). A subline containing variants of the GII.10 genotype, identified in India (2003 and 2013) and Korea (2012), is also presented here.

Line B contains one norovirus GII.17, identified in South Africa in 2011, as well as additional recombinants with capsid GII.3 (Russia, Brazil, USA, Peru, Bangladesh, Spain, Italy, Germany, 2011–2015) and GII.13 (Thailand, Taiwan, Germany, USA, Japan, China, Spain, Italy, Russia, 2010–2016), there is one Nizhny Novgorod strain GII.13 (2016), GII.16 (China 2013).

Subline GII.2 in line C is heterogeneous and includes new recombinants associated with the capsid gene: GII.13 (Spain 2020), GII.2, de-

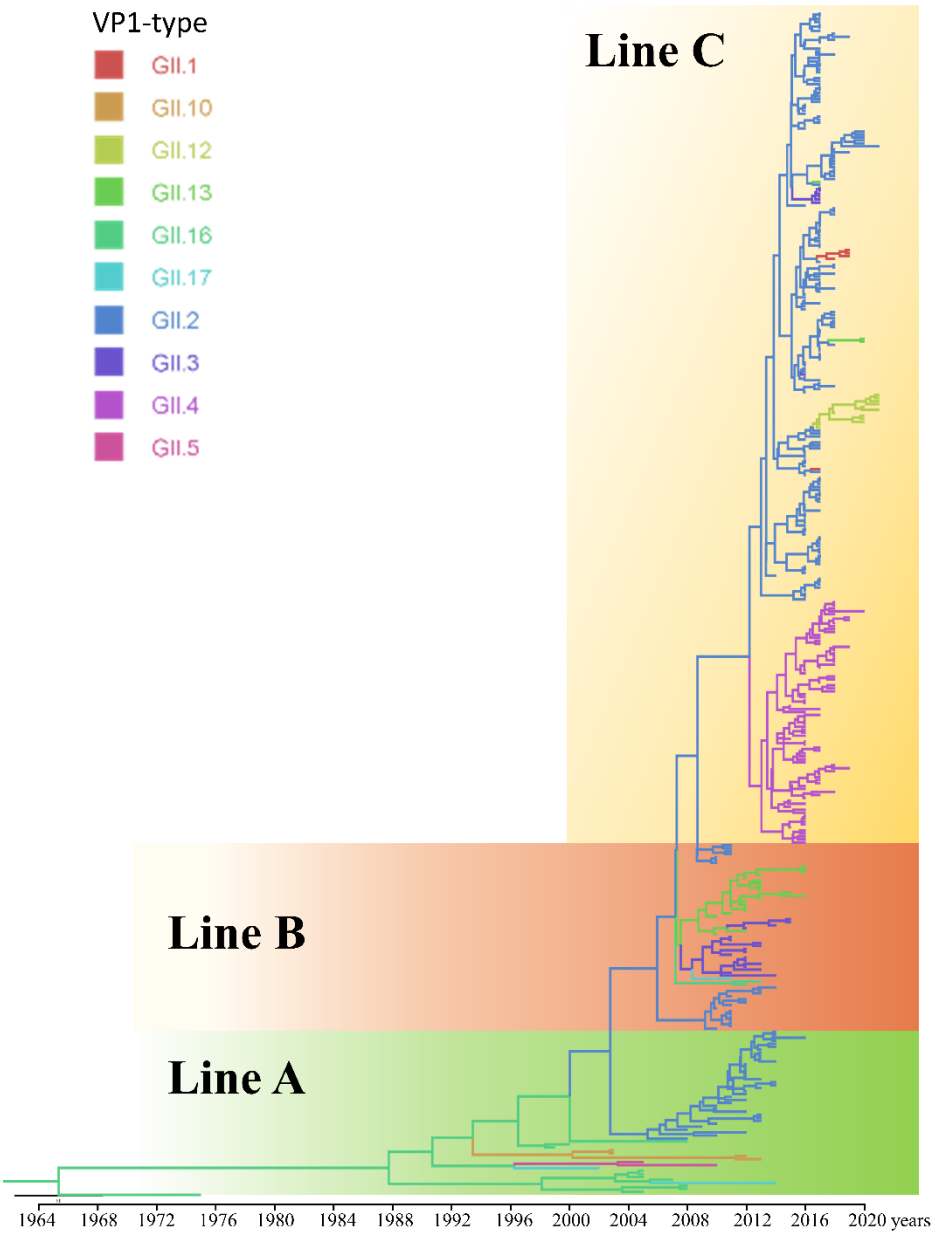

Fig. 3. Expanded phylogenetic analysis based on the polymerase gene (264 nucleotides) of noroviruses (n = 400). Phylogenetic tree was constructed by the BEAST software and including 241 full-genome and 159 partial sequences. Branches were colored by VP1-type. Phylogenetic clusters based on lines: line A is represented by green, line B is red and line C is orange

tected from 2014 to 2021 and found in China, Japan, Thailand, Australia, USA, Argentina, Canada, Russia, Great Britain, Belarus (Nizhny Novgorod isolates are located together with strains from Russia, USA and China). In subline GII.12, Nizhny Novgorod strains formed a separate group, which included the Belarusian isolate (2021), isolates from the USA (2018) and Brazil (2020) are also represented here.

Subline GII.4 in line C includes sequences of norovirus circulating in China, Korea, Japan, Thailand, Taiwan, Great Britain, Canada, USA, Argentina, Brazil, Paraguay, France, Russia in 2015–2020. Isolates of GII.4 from Nizhny Novgorod (2016–2017) are located together with other Russian strains (isolates from Novosibirsk 2016).

Thus, the addition of short sequences to the tree made it possible to increase the analysis in-

terval for the period from 1975 to 2021, extend the geography of the study, in particular due to isolates from a number of European countries, and show that representatives of the A and B lines have not been found since 2016, and all currently circulating noroviruses with P16 polymerase belong to line C.

Analysis of nucleotide substitutions

The strain GII.P16 – AY772730 (Hu/NLV/GII/Neustrelitz260/2000/DE from Germany) was used as a reference for comparison.

Fig. 4 shows the nucleotide substitutions detected along the polymerase lines and are informative (they were found in more than 10% of the sequences). Substitutions in positions C519G, T877A, G994A, A1069C, A1078G are nonsynonymous, since they lead to substitutions in the amino acid sequence of polymerase, and also differentiate line C from lines A and B.

For line A, differences from the reference strain are observed in such positions as C144T, T603A, G831A. For line B – C105T, C483T, A792G. For line C – A54T, T225C, C492A.

Within line A, subline GII.16 has characteristic substitutions in positions A227G, C645T, A864G; subline GII.2 in positions C72A, A234G, C495T; subline GII.17 – A159G, T612A; subline GII.5 – T240C, C594T, T1186G.

The line B in the sublines also have distinctive substitutions: GII.2 – G30A, G204A, C1158T; GII.3 – C273T, C838T, C1365T; GII.13 – T957C, A1380G.

In line C, substitutions in sublines are presented as follows: for GII.1 – A313T, G591A, T634A; for GII.2 – T418C, C1140T, A1161T; for GII.3 – G117A, A459G, A969G; for GII.12 – G360T, C367T, T1065C.

In the short region of 264 nucleotides indicated in Figure 5, it is possible to observe the preservation of differentiation on the line and subline, for example, subline GII.4 has a replacement in position A1530G, which is characteristic only for it in line C, while GII.2, GII.3 – common in position T1302C, GII.12 has a characteristic replacement in position A1308G, and GII.1 – in A1404G. Line C and B have a common replacement in position G1505A, and line A has characteristic replacements in position G1284A, C1416T. Therefore, the analysis of a short region of the gene (264 nucleotides) of the polymerase is correct, but has limitations when differentiating on sublines.

Analysis of amino acid replacement

When analyzing the derived amino acid sequences, representatives of a novel variant of the GII.P16 polymerase gene, which has been globally distributed in the world since 2015, revealed five substitutions (D173E, S293T, V332I, K357Q, T360A) compared to earlier strains (Fig. 5). Moreover, the last four substitutions are the most significant, since they are near sites responsible for the enzyme activity of polymerase and could affect the transmissivity of the virus.

Lysine at position 291 (K291) is important for polymerase, since it was shown that this variant of the residue allows to increase the replication rate (Bull *et al*., 2010), therefore, polymerase GII.P16 underwent additional mutations during its evolution (5 more substitutions), which allowed it to become widespread.

Discussion

The acquisition of alternative RNA polymerases by recombination is an important mechanism for the evolution of norovirus (Kim *et al*., 2014). Recombination of genomes of various norovirus strains occurs mainly at the junction of open reading frames encoding non-structural (ORF1) and structural proteins (ORF2 and ORF3).

Analysis of the nucleotide sequences of the norovirus GII.P16 polymerase gene available in the GenBank database showed that this gene was present in recombinants with different capsid protein specificity detected in the world since the 70s of the last century. At the moment, it is known that GII.P16 polymerase is associated with 11 different capsids VP1 (GII.1, GII.2, GII.3, GII.4, GII.5, GII.10, GII.12, GII.13, GII.16, GII.17, and variant GII.21 is only in the CDC collection (Barclay *et al*., 2018)), although recombination events with are

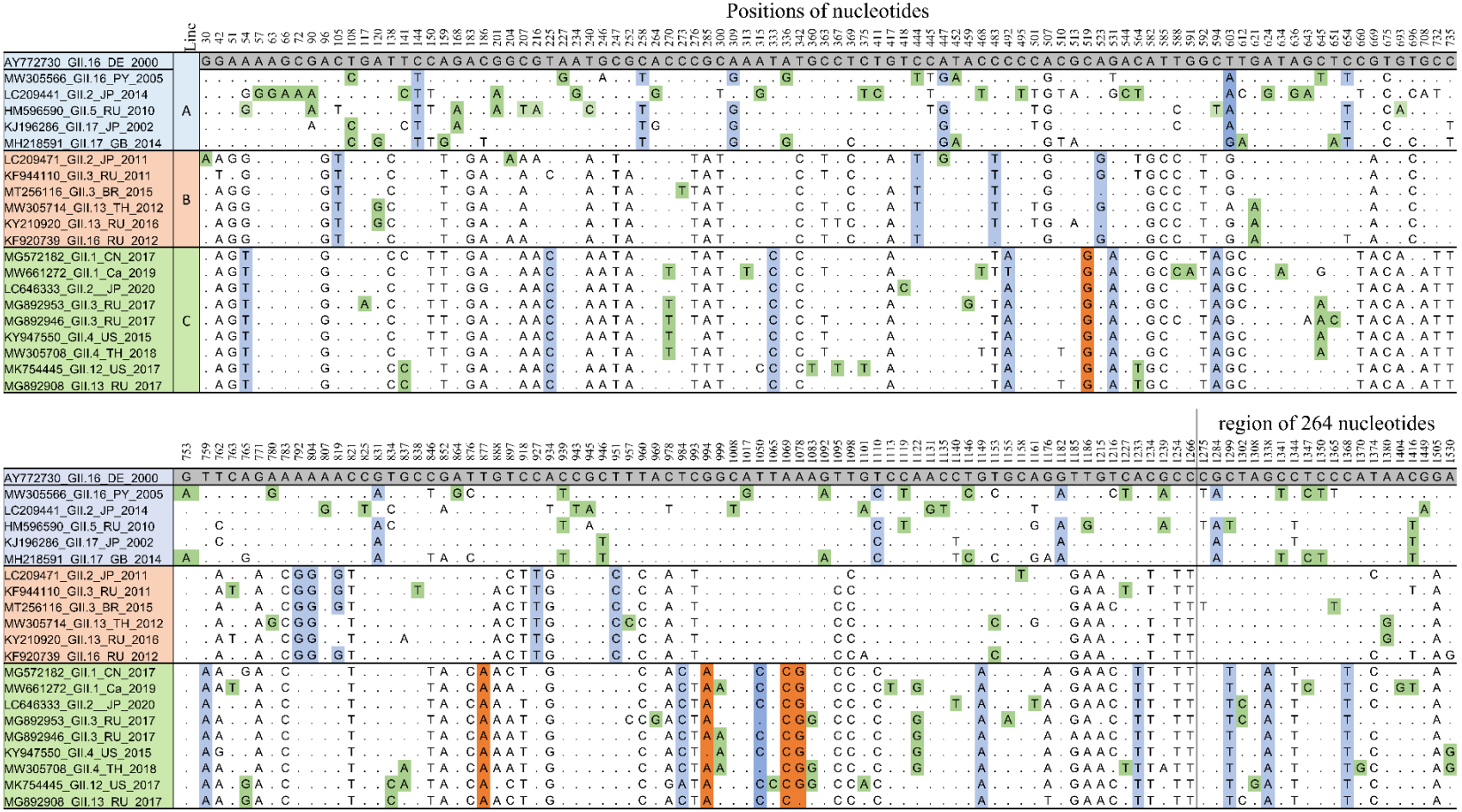

Fig. 4. Comparative analysis of the nucleotide sequences of GII.P16 gene relative to the reference strain (GenBank number AY772730 (Hu/NLV/GII/Neustrelitz260/2000/DE from Germany))

substitutions leading to changes in amino acids in the region of the active center of polymerase

nucleotide substitutions characteristic of lines

nucleotide substitutions characteristic of sublines

Substitutions leading to changes in amino acids in the region of the active center of polymerase are indicated by orange, nucleotide substitutions characteristic of lines indicated by blue, nucleotide substitutions characteristic of sublines indicated by green

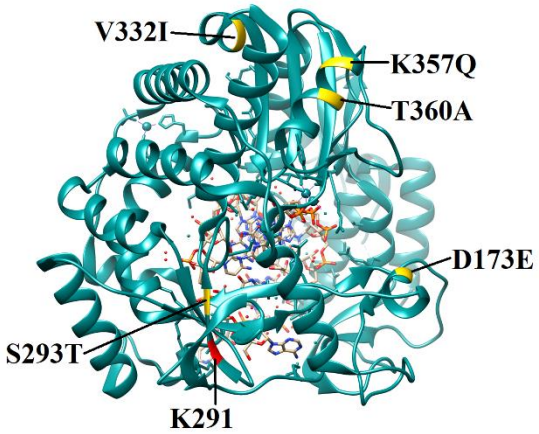

Fig. 5. Structural model of amino acid substitutions in GII.P16 norovirus based on polymerase GII.4 (PDB number 4QPX). The structural model was visualized using Chimera v.1.16. Amino acid substitutions are indicated in yellow. Residue K291, which affects the replication rate (Bull *et al.*, 2010), is indicated in red. D – Aspartic acid, S – Serine, V – Valine, K – Lysine, T – Threonine, E – Glutamic acid, I – Isoleucine, Q – Glutamine, A – Alanine

most often recorded GII.2, GII.4 and GII.13 (Kendra *et al*., 2022).

According to the data obtained by Barclay *et al*. (2018), the phylogeny of GII.P16 polymerases showed the presence of three different lines of GII.P16 polymerase: «existing A», «existing B» and «new». Our phylogenetic analysis confirmed the existence of three lines, which we named A, B and C.

Strains of the A and B lines descended from a common ancestor, and early representatives of the A line were found in the period from 1975 to 2002 among viruses with capsid GII.16 and GII.17. and continued to circulate until 2014. The first strains of line B appeared in 2010 in association with capsid GII.2 and then in the form of recombinants with different types of capsid protein (GII.2, GII.3 GII.13, GII.16 and GII.17) were detected up to 2016. Noroviruses of line C have a common ancestor with line B, have been detected since 2014 to the present in association with capsids of genotypes GII.1, GII.2, GII.3, GII.4, GII.12 and GII.13 with the predominance of GII.4[P16] and GII.2[P16] recombinants both in Nizhny Novgorod and around the world.

The data obtained on the presence of amino acid substitutions in representatives of a novel variant of the GII.P16 polymerase gene are consistent with the hypothesis of Ruis *et al.* (2017) that the worldwide distribution of strains with a novel variant of polymerase is associated with changes that occurred in the active center of polymerase, which could increase the transmissivity of norovirus (Zhirakovskaya *et al*., 2020). In addition to the frequency of mutations, the replication rate is considered another important factor determining the fitness of the virus, and the K291 residue contained in the motif located in the «finger» domain of the enzyme contributes to this (Bull *et al*., 2010).

Thus, the acquisition of a new variant of the GII.P16 polymerase gene by noroviruses with different genotypes of the capsid protein probably provided certain advantages to recombinants and created prerequisites for their widespread distribution.

Further study of various mutations occurring in polymerase will contribute to a better understanding of emerging features affecting the transmissivity of noroviruses (Tohma *et al*., 2017). It is planned to continue this study and expand it by including the complete polymerase GII.P16 genes identified in Nizhny Novgorod.

Acknowledgements

The reported study was funded by the Federal Service for Surveillance on Customers Rights Protection and Human Wellbeing the Russian Federation №121092100042-4.

We express our gratitude to V.Y. Epifanov for developing the database architecture for storing sequences, creating the NoroGen database, cleaning and preparing data for uploading, developing tools for editing the database, developing tools for analyzing sequences and uploading in the required formats.

References

ALTSCHUL S.F., GISH W., MILLER W., MYERS E.W. & LIPMAN D.J. (1990): Basic local alignment search tool. *Journal of molecular biology* **215**(3), 403–410.

BARCLAY L., CANNON J.L., WIKSWO M.E., PHILLIPS A.R., BROWNE H., MONTMAYEUR A.N., Roman L. TATUSOV R.L., BURKE R.M., Hall A.J. & VINJE J. (2019): Emerging Novel GII.P16 Noroviruses Associated with Multiple Capsid Genotypes. *Viruses* **11**(6):535. P. 17.

BIDALOT M., THERYL., KAPLON J., DE ROUGEMONT A. & AMBERT-BALAY K. (2017): Emergence of new recombinant noroviruses GII.p16-GII.4 and GII.p16-GII.2, France, winter 2016 to 2017. *Euro surveillance* **22**(15), 1–5.

BULL R.A., EDEN J-S., RAWLINSON W.D. & WHITE P.A. (2010): Rapid Evolution of Pandemic Noroviruses of the GII.4 Lineage. *PLoS Pathogens* **6**(3), 1-10.

CHHABRA P., DE GRAAF M., PARRA G.I. CHAN M.C-W., GREEN K., MARTELLA V., WANG Q., WHITE P.A., KATAYAMA K., VENNEMA H., KOOPMANS M.P. & VINJE J. (2019): Updated classification of norovirus genogroups and genotypes. *The Journal of general virology* **100**(10), 1393–1406.

CHHABRA P., DE GRAAF M., PARRA G.I. CHAN M.C-W., GREEN K., MARTELLA V., WANG Q., WHITE P.A., KATAYAMA K., VENNEMA H., KOOPMANS M.P. & VINJE J. (2020): Corrigendum: Updated classification of norovirus genogroups and genotypes. *The Journal of general virology* **101**(8). P. 893.

EPIFANOVA N.V. & EPIFANOV V.Y. (2021): Nucleotide sequences of the norovirus genome (NoroGen), version 2.0. Database registration certificate 2021622566, https://www.fips.ru/registers-doc-view/fips_servlet

HAN J., WU X., CHEN L., FU Y., XU D., ZHANG P. & JI L. (2018): Emergence of norovirus GII.P16-GII.2 strains in patients with acute gastroenteritis in Huzhou, China, 2016–2017. BMC *Infectious diseases* **18**(1), 1–9.

HATA M., NALAMURA N., KOBAYASHI S., ONOUCHI A., SAITO T., HIROSE E., ADACHI H., SAITO N., ITO M., YASUI Y., MATSUMOTO M. & MINAGAWA H. (2018): Emergence of new recombinant noroviruses GII.P16-GII.2 and GII.P16-GII.4 in Aichi, Japan, during the 2016/17 season. *Japanese journal of infectious diseases* **71**(4), 319–322.

KENDRA J.A., TOHMA K. & PARRA G.I. (2022): Global and regional circulation trends of norovirus genotypes and recombinants, 1995–2019: A comprehensive review of sequences from public databases. *Reviews in medical virology* **32**(1), 1–10.

KIM J.E., LEE S.G., CHO H.G. HAN S-H., KANG L-H., LEE Y-M., PARK C-J & PAIK S. (2014): Genetic Analysis of the Capsid Region of Norovirus GII.4 Variants Isolated in South Korea. *Journal of microbiology* **52**, 427–434.

KRONEMAN A., VENNEMA H. & DEFORCHE K. (2011): An automated genotyping tool for enteroviruses and noroviruses. *Journal of clinical virology* **51**(2), 121–125.

KRONEMAN A., VEGA E., VENNEMA H., VINJE J., WHITE P.A., HANSMAN G., GREEN K., MARTELLA V., KATAYAMA K. & KOOPMANS M.P. (2013): Proposal for a unified norovirus Nomenclature and genotyping. *Archives of virology* **158**, 2059–2068.

KUMAR S., STECHER G. & LI M. (2018): MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular biology and evolution* **35**, 1547-1549.

LEE S., LIU H., WILEN C.B., SYCHEY Z.E., DESAI C., HYKES B.L., ORCHARD R.C., McCUNE B.T., KIM K-W., NICE T.J., HANDLEY S.A., BALDRIDGE M.T., Gaya K AMARASINGHE G.K. & VIRGIN H.W. (2019): A secreted viral nonstructural protein determines intestinal norovirus pathogenesis. *Cell Host Microbe* **25**, 845–857.

OZAKI K., MATSUSHIMA Y., NAGASAWA K., MOTOYA T., RYOA., KURODA M., KATAYAMA K. & KIMURA H. (2018): Molecular evolutionary analyses of the RNA-dependent RNA polymerase region in norovirus genogroup II. *Frontiers in microbiology* **9**, 1–11.

PETTERSEN E.F., GODDARD T.D., HUANG C.C., COUCH G.S., GREENBLATT D.M., MENG E.C. & FERRIN T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**(13), 1605–1612.

RUIS C., ROY S., BROWN J.R., ALLEN D.J., GOLDSTEIN R.A. & BREUER J. (2017): The emerging GII.P16-GII.4 Sydney 2012 norovirus lineage is circulating worldwide, arose by late-2014 and contains polymerase changes that may increase virus transmission. *PLoS One* **12**(6), 1–9.

SUCHARD M.A., LLEMEY P., BBAELE G., AYYRES D.A., DRUMMOND A.J. & RAMBAUT A. (2018): Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus evolution* **4**(1). 1–5.

TOHMA K., LEPORE C.J., FORD-SILTZ L.A. & PARRA G.I. (2017): Phylogenetic Analyses Suggest that Factors Other Than the Capsid Protein Play a Role in the Epidemic Potential of GII.2 Norovirus. *Clinical Science and Epidemiology* **2**(3), 1–13.

VONGPUNSAWAD S., VENKATARAM-PRASAD B.V. & ESTES M.K. (2013) Norwalk virus minor capsid protein VP2 associates within the VP1 shell domain. *Journal of virology* **87**, 4818–4825.

ZAMYATKIN D., RAO C., HOFFARTH E., JURCA G., RHO H., PARRA F., GROCHULSKI P., & NG, K.K. (2014). Structure of a backtracked state reveals conformational changes similar to the state following nucleotide incorporation in human norovirus polymerase. *Acta crystallographica. Section D, Biological crystallography* **70**(12), 3099–3109

ZHIRAKOVSKAYA E.V., TIKUNOV A.Y., SOKOLOV S.N., KRAVCHUK B.I., KRASNOVA E.I. and TIKUNOVA N.V. (2020): Characterization of the complete genome sequence of the recombinant norovirus GII.P16/GII.4_Sydney_2012 revealed in Russia. *Vavilov Journal of Genetics and Breeding* **24**(1), 69–79.