Evolutionary analyses of Georgian Rabies Lyssavirus strains and their genetically most closely related conspecifics from elsewhere provide evidence suggesting risks for dog-to-cattle rabies transmission as well as this disease coinfections and transborder transmission

Leila Tabatadze
The Scientific-Research Center of Agriculture
Ekaterine Gabashvili
The Scientific-Research Center of Agriculture
Saba Kobakhidze
The Scientific-Research Center of Agriculture
George Lomidze
The Scientific-Research Institute of Agriculture
Jimsher Loladze
The Scientific-Research Center of Agriculture
Levan Tsitskishvili
The Scientific-Research Center of Agriculture
Mamuka Kotetishvili (m.kotetishvili@hygienege.onmicrosoft.com)
The G. Natadze Scientific-Research Institute of Sanitary, Hygiene and Medical Ecology
https://orcid.org/0000-0002-3882-0878

Research Article

Keywords:

Posted Date: March 31st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1434624/v1

Abstract

Genetic relationships between Rabies Virus (RABV) isolates recovered from dog, jackal, and cattle in Georgia, and their genetically most closely related conspecifics from elsewhere, were determined, using the DNA analysis of nucleoprotein (N) gene loci. Multiple isolates from dogs and cattle were found to share the same sequence types (STs), being suggestive of the risks of dog-to-cattle rabies transmission in Georgia. Exhibiting population selective sweeps and expansion, as well as genetic recombination, the evolutionary analyses of Georgian RABV strains and their conspecifics from Russia, Turkey, Iran, and elsewhere, provided further evidence for rabies coinfections and its transborder transmission.

Full Text

Rabies is one of the most dreadful zoonotic diseases, exhibiting almost 100% fatality rates in infected mammals if left untreated [1]. The etiological agent of rabies, *Rabies Lyssavirus*, otherwise called as Rabies Virus (RABV), is a nonsegmented negative-strand RNA virus, with the broadest host range among lyssaviruses, including humans, various domestic animals, livestock, and wildlife. In Georgia, rabies has been officially documented since 1930, where nearly one third of animal rabies cases were found to occur in livestock, including 83% of cases associated with cattle [1]. This pilot study provides some initial but important information on the population structure and mechanisms of evolutionary divergence of RABV strains recovered in Georgia, and on those of their genetically most closely related conspecifics from elsewhere.

We analyzed the DNA sequences of the nucleoprotein-encoding (N) gene loci (the 1350-bp region) of a total of 72 RABV isolates, recovered (during 2015-2016) from dogs, jackals, and cattle in Georgia. The information on both the above DNA sequences and the sources of these isolates (Table S1) was retrieved from the GenBank database of the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/). The above sequences were submitted to the NCBI GenBank database by the Pox and Rabies branch, Centers for Disease Control and Prevention, Atlanta, GA, the United States of America. The Maximum Likelihood (ML) method, implemented in the Molecular Evolutionary Genetics Analysis software MEGA X (v. 10.1.8) [2], was applied to determine STs of the N gene loci of these RABV isolates, and to construct a phylogenetic tree. In the phylogenetic analysis, we also included their genetically most closely related conspecifics from elsewhere. The DNA sequences of these conspecifics were obtained using the Basic Local Alignment Search Tool (BLAST) with the megablast algorithm searching for the genetically closest homologs of the N gene loci of the determined RABV STs in the blastn database.

The Tajima's Neutrality Test [3] was applied to determine a number of segregating sites and an average number of nucleotide differences across the targeted N gene loci of the RABV isolates. Besides, the FUBAR (Fast, Unconstrained Bayesian AppRoximation) model [4] was applied to infer nonsynoymous (d_N) versus synonymous (d_S) substitution rates, as well as the d_N/d_S ratios (), on a per-site basis for the coding alignment of the above genetic loci. The FEL (Fixed Effects Likelihood) model [5] was applied to

infer d_N versus d_S substitution rates on a per-site basis for the same coding DNA sequence alignment across the RABV phylogenetic clusters. Using the Tamura-Nei model (+G) [6], we also estimated a shape parameter for the discrete Gamma Distribution in these N gene loci. The Kimura 2-parameter model [7] was applied to determine a transition/transversion (*ti*/*tv*) bias (*R*) across the targeted nucleotide sequences by computing a tree topology with the maximum Log likelihood 3605.928. A maximum composite likelihood estimate of the pattern of nucleotide substitution was also determined as described previously [8].

The method of split decomposition [9], implemented in SplitsTree (v.4.14.4) [10], was used to reconstruct possible genetic recombination events of the N gene loci between the RABV strains. For the SplitsTree-generated splits graphs, the bootstrap values being ≥ 95 (from 1000 replicates) for each node of a parallelogram(s), and the fit values being ≥ 95 for each splits network, were considered to be statistically significant. SimPlot (v3.5.1) [43] [11] was applied to detect recombination hotspots across the above N gene loci.

The DNA sequence analysis discriminated the local 72 RABV isolates into 41 STs (Table S1). As shown, while certain STs of the N gene loci (STs: 9, 10, 12, 13, 15-20, 23, 25-28, 30-32, 35, 36, 38, 39, and 41) were uniquely associated with the RABV isolates from dogs, several STs (STs: 1, 7, 11, 29, and 37) were shared by the isolates recovered from both dogs and cattle. Two RABV isolates (MW055109.1 and MT079888.1), recovered from the jackals, were found to share the ST22 with the isolate (MT079902.1) from a dog. STs 4, 5, 6, 8, 33, and 34 appeared to be associated with the RABV isolates recovered from cattle in the neighbor regions across the Western Georgia (Imereti and Guria) or the north-western part of the country (Samegrelo-Zemo Svaneti), while STs 14, 21, 24, and 40 were uniquely associated with the South-East region (Kvemo Kartli) or the eastern part (Kakheti) of Georgia.

The BLAST-identified genetically most closely related conspecifics of the local RABV isolates were associated mainly with the strains from the neighbor countries, such as Russia, Azerbaijan and Turkey followed by Iran, Iraq, Hungary, and Tajikistan (Table S2), sharing the DNA-DNA identity in a range of 96.74-99.93%. The ancestral ST and the most recent ancestral ST were found to be associated respectively with the RABV strain (KY860610.1) isolated from a dog in Turkey, and the strain (MK598340.1) isolated from a red fox in Hungary, when the above conspecifics of the local RABV isolates were analyzed alone using ML. However, the ancestral ST versus the most recent ancestral ST appeared to be associated respectively with the strain (JQ944706.1) recovered from a dog in Russia, and two strains (MT079899.1 and MT079940.1 sharing the same ST37) from a dog and cattle in Georgia, when the entire RABV population, including the local isolates, were analyzed using the same algorithm. The phylogenetic analysis resulted in one large genetic clade with multiple clusters and sub-clusters incorporating a great majority of the STs, and one relatively distinct phylogenetic clade including only four STs 39, 41, 40 and 38 (Fig. 1). As shown, the former also included the RABV strains recovered in Turkey and Russia, while the latter encompassed the strains recovered in Azerbaijan, Hungary, and Tajikistan. For a better phylogenetic resolution, in this analysis, we also included multiple RABV isolates from Iran, because, based on the BLAST analysis, certain strains from this country were genetically

closely related to the ones from Turkey, while others to a single strain from Iraq, constituting collectively another distinct phylogenetic clade (Fig. 1). The above phylogenetic clades may represent the RABV evolutionary lineages, although, a significantly larger strain collection from all these countries, including Georgia, need to be analyzed to support the above statement.

The analysis of molecular evolutionary patterns of the N gene loci of RABV isolates revealed that the *ti* substitution rates were notably higher than the *tv* substitution rates, especially the T (U)→C and C→T (U) rates (Table S3). The maximum composite likelihood estimates of the pattern of nucleotide substitutions resulted in the following nucleotide frequencies: 29.64% (A), 26.28% (T/U), 20.56% (C), and 23.52% (G). The *ti*/*tv* rate ratios respectively were k_1 = 9.007 (purines) and k_2 = 17.013 (pyrimidines), with the overall *ti*/*tv* bias being R = 6.214. The FEL-inferred d_N/d_S rates were in a range of 0.0035-0.05. The FUBAR could reveal pervasive negative/purifying selection at 147 sites across the N gene loci examined. No evidence for pervasive positive/diversifying selection in these gene loci could be found in the FUBAR analysis (Fig. S1). The Tajima's neutrality test of the nucleotide variations of the N gene loci resulted in a negative value: D = - 1.237845 (Table 1).

SplitsTree generated the splits graphs with the strongest fit (100), exhibiting multiple parallelograms, some nodes of which were supported by the strong bootstrap values (94.9-99.9) (Fig. 2), when examining the above N gene loci. As shown, the genetic recombination events of the N gene loci, displayed in a capacity of these parallelograms, involved not only certain STs of the RABV isolates from Georgia (ST40, ST21, ST 15, ST17) (Fig. 2a), but also multiple strains from elsewhere, mainly from Iran (Fig. 2b-d). SimPlot could determine multiple genetic recombination hotspots across the N gene loci of some of these RABV strains (Fig. S2a-d). A great majority of the recombining RABV isolates were from dogs.

Based on the results of the above analysis having identified the specific STs shared by multiple RABV isolates from two different isolation sources (dogs and cattle), it is suggested that rabid dogs could contribute to the transmission of rabies in cattle in Georgia. Besides, as shown, while the distribution of RABV strains reflect some regional patterns (e.g., in Kvemo Kartli and Kakheti), certain strains from dogs seem to have spread across some regions of this country (such as Imereti, Guria, and Samegrelo-Zemo Svaneti). We assume that at least stray dogs could be the transmission vehicles of these strains across the above regions.

It must be noted that the ST22, shared by certain RABV strains recovered from the dog and jackals, belongs to one of the RABV cosmopolitan phylogenetic clades described previously [12]. While the immediate sources of rabies acquisition by the above animals remain unknown, jackals, as being generally a part of the dominant wildlife reservoirs for emerging rabies in domestic and agricultural animals [13], can be thought to contribute to this disease transborder transmission involving Georgia. As shown, the genetically most closely related conspecifics of the local RABV isolates appeared to be associated with Russia, Azerbaijan, Turkey, Tajikistan, Iran, Hungary or Iraq. The ML analysis results of the ancestral STs can be suggestive of a history of transborder transmission of certain RABV strains from

Russia, Turkey, and Hungary, including Georgia that seems to have contributed to the recent events of the above process exhibiting this disease global transmission [14, 15].

The *ti/ tv* substitution rates, determined for the targeted N gene loci of the RABV strains in our analysis, are consistent with the results from the previous studies demonstrating a significantly higher frequency of C→U changes compared to that of their reverse (U→C) and other *ti*s changes at site in many mammalian RNA viruses with <5% of heterogeneity [16]. Here, we show that the U→C and C→U substitutions were one of the major drivers of DNA sequence diversification of the N gene loci of the RABV strains examined. Besides, the *ti/ tv* bias estimate, obtained from our analysis, was also very close to the *ti/ tv* value (5.682) determined previously for the N gene loci of the RABV strains recovered elsewhere [17], thus suggesting some common evolutionary trends of this viral agent at least across certain regions globally.

The previous studies show that the evolution of the RABV N gene has been strongly subjected to purifying selection pressures [18], and that it has been highly constrained especially at nonsynonymous sites, with no clear evidence for positive selection [19]; however, it is noteworthy that some codons of this gene have been affected by nonsynonymous substitutions et relatively higher rates, which may signify the existence of localized selection pressures [17]. In our analysis, the FEL-inferred d_N/d_S rates on the per-site basis revealed no clear tendency towards amino acid changes, being further supported by the FUBAR inferences exhibiting pervasive negative/purifying selection at multiple sites across the targeted N gene loci. In contrast, no evidence for pervasive positive/diversifying selection could be found in the above gene loci. The Negative Tajima's *D*negative value, obtained from the analysis of the RABV strains in this study, could be suggestive of negative selection and recent selective sweeps [20], as well as of a recent population expansion [21].

The transmission of RABV strains in non-flying mammals was shown to occur at a low rate at a global level, resulting in the phylogenetic structure that reflects a significant population subdivision rather than gene flow [22]. However, some studies provided initial evidence for genetic recombination events involving the polymerase gene [23] and the N gene [24] in some natural populations of RABV. Here, the results obtained from our SplitsTree and SimPlot-analyses, provide additional evidence for intra-species recombination, being in a strong agreement with the previous notion [24] suggesting that the N gene could be affected by homologous genetic recombination across the RABV natural populations. The above findings pinpoint collectively to the presence of rabies coinfections in certain rabid animals, such as e.g., rabid dogs. However, more in-depth studies are needed to determine whether HGT contributes to virulence and/or pathogenicity of this viral pathogen, as well as to its adaptation to various host environments.

Statements And Declarations

Acknowledgement

We thank Stylianos Koulouris and Yves Van der Stede for their helpful discussions and editing of the manuscript.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that they have no competing interests. † Leila Tabatadze and Ekaterine Gabashvili contributed equally to the research described in this paper.

Ethics Approval

N/A

Author Contributions

Leila Tabatadze and Ekaterine Gabashvili mined the DNA sequence data in the NCBI GenBank database, and performed the ST and genetic recombination analyses. Saba Kobakhidze and George Lomidze contributed to the phylogenetic and recombination analyses. Jimsher Loladze and Levan Tsitskishvili provided their expertise for the study, and contributed to a drafting of the manuscript. Mamuka Kotetishvili developed the conception and design for, and supervised the study, as well as drafted and formulated the manuscript.

References

1. Mersini K (August, 2020) Cost-effectiveness Evaluation of the National Rabies Control Program in Georgia and Recommendations for Improvement. the FAO Division/Office FEGEO, pp 1–36
2. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096
3. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595. https://doi.org/10.1093/genetics/123.3.585
4. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol 30:1196–1205. https://doi.org/10.1093/molbev/mst030
5. Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL (2010) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26:2455–2457. https://doi.org/10.1093/bioinformatics/btq429

6. Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023
7. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120. https://doi.org/10.1007/BF01731581
8. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101:11030–11035. https://doi.org/10.1073/pnas.0404206101
9. Bandelt H-J, Dress AWM (1992) Split decomposition: A new and useful approach to phylogenetic analysis of distance data. Mol Phylogenet Evol 1:242–252. https://doi.org/10.1016/1055-7903(92)90021-8
10. Huson DH, Bryant D (2006) Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol 23:254–267. https://doi.org/10.1093/molbev/msj030
11. Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC (1999) Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 73:152–160. https://doi.org/10.1128/JVI.73.1.152-160.1999
12. Troupin C, Dacheux L, Tanguy M, Sabeta C, Blanc H, Bouchier C, Vignuzzi M, Duchene S, Holmes EC, Bourhy H (2016) Large-Scale Phylogenomic Analysis Reveals the Complex Evolutionary History of Rabies Virus in Multiple Carnivore Hosts. PLoS Pathog 12:e1006041. https://doi.org/10.1371/journal.ppat.1006041
13. Noman Z, Anika TT, Haque ZF, Rahman AKMA, Ward MP, Martínez-López B (2021) Risk factors for rabid animal bites: a study in domestic ruminants in Mymensingh district, Bangladesh. Epidemiol Infect 149:e76. https://doi.org/10.1017/S095026882100056X
14. Johnson N, Freuling C, Vos A, Un H, Valtchovski R, Turcitu M, Dumistrescu F, Vuta V, Velic R, Sandrac V, Aylan O, Müller T, Fooks AR (2008) Epidemiology of rabies in Southeast Europe. Dev Biol (Basel) 131:189–198
15. Turcitu MA, Barboi G, Vuta V, Mihai I, Boncea D, Dumitrescu F, Codreanu MD, Johnson N, Fooks AR, Müller T, Freuling CM (2010) Molecular epidemiology of rabies virus in Romania provides evidence for a high degree of heterogeneity and virus diversity. Virus Res 150:28–33. https://doi.org/10.1016/j.virusres.2010.02.008
16. Simmonds P, Ansari MA (2021) Extensive C->U transition biases in the genomes of a wide range of mammalian RNA viruses; potential associations with transcriptional mutations, damage- or host-mediated editing of viral RNA. PLoS Pathog 17:e1009596. https://doi.org/10.1371/journal.ppat.1009596
17. Bourhy H, Kissi B, Audry L, Smreczak M, Sadkowska-Todys M, Kulonen K, Tordo N, Zmudzinski JF, Holmes EC (1999) Ecology and evolution of rabies virus in Europe. J Gen Virol 80(Pt 10):2545–2557.

https://doi.org/10.1099/0022-1317-80-10-2545

18. Kuzmina NA, Kuzmin IV, Ellison JA, Taylor ST, Bergman DL, Dew B, Rupprecht CE (2013) A reassessment of the evolutionary timescale of bat rabies viruses based upon glycoprotein gene sequences. Virus Genes 47:305–310. https://doi.org/10.1007/s11262-013-0952-9
19. Holmes EC, Woelk CH, Kassis R, Bourhy H (2002) Genetic constraints and the adaptive evolution of rabies virus in nature. Virology 292:247–257. https://doi.org/10.1006/viro.2001.1271
20. Zhai W, Nielsen R, Slatkin M (2009) An investigation of the statistical power of neutrality tests based on comparative and population genetic data. Mol Biol Evol 26:273–283. https://doi.org/10.1093/molbev/msn231
21. Sharma M, Fomda BA, Mazta S, Sehgal R, Singh BB, Malla N (2013) Genetic diversity and population genetic structure analysis of *Echinococcus granulosus sensu stricto* complex based on mitochondrial DNA signature. PLoS ONE 8:e82904. https://doi.org/10.1371/journal.pone.0082904
22. Bourhy H, Reynes J-M, Dunham EJ, Dacheux L, Larrous F, Huong VTQ, Xu G, Yan J, Miranda MEG, Holmes EC (2008) The origin and phylogeography of dog rabies virus. J Gen Virol 89:2673–2681. https://doi.org/10.1099/vir.0.2008/003913-0
23. Liu W, Liu Y, Liu J, Zhai J, Xie Y (2011) Evidence for inter- and intra-clade recombinations in rabies virus. Infect Genet Evol 11:1906–1912. https://doi.org/10.1016/j.meegid.2011.08.031
24. He C-Q, Meng S-L, Yan H-Y, Ding N-Z, He H-B, Yan J-X, Xu G-L (2012) Isolation and identification of a novel rabies virus lineage in China with natural recombinant nucleoprotein gene. PLoS ONE 7:e49992. https://doi.org/10.1371/journal.pone.0049992

Table

Table 1. The results obtained from the evolutionary analysis of the RABV N gene loci using Tajima's Neutrality Test

m	*S*	*n*	p_s	Θ	π	*D*
61	186	1350	0.137778	0.029441	0.019096	-1.237845

m = a number of sequences examined; *n* = a total number of sites, *S* = a number of segregating sites, p_s = S/n, $\Theta = p_s/a_1$, π = nucleotide diversity; *D* - the Tajima test statistic.

Figures

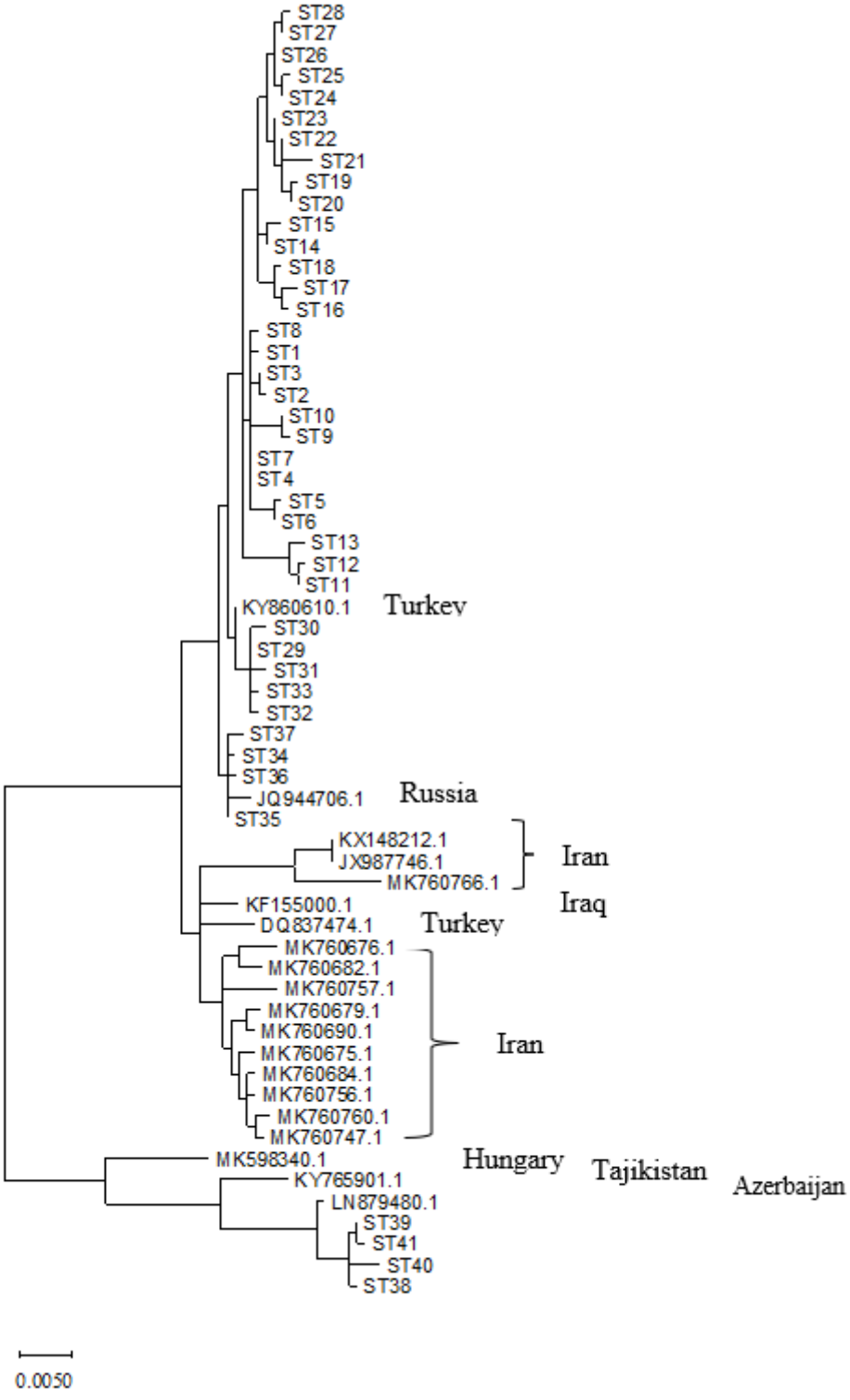

Figure 1

Genetic Relationships between the STs of the RABV strains isolated in Georgia and their genetically most closely related conspecifics recovered globally

Figure 2

The SplitsTree-generated splits graphs exhibiting independent genetic recombination events of the N gene loci between different RABV strains and STs

The parallelogram(s) elucidating: (a) HGT events identified exclusively between some RABV STs from Georgia (Fit: 100); (b) HGT events identified between RABV isolates from Iran and Iraq (Fit: 100); (c) HGT events involving multiple RABV strains isolated in Georgia, Iran and Iraq (Fit: 100); (d) HGT events involving RABV strains from Georgia, Iran, and Iraq (Fit: 100).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.docx