

Molecular epidemiology and phylodynamics of goose hemorrhagic polyomavirus

Running title: Phylodynamics of goose hemorrhagic polyomavirus

Eszter Kaszab¹, Szilvia Marton¹, Ádám Dán², Attila Farsang³, Ádám Bálint⁴, Krisztián Bányai¹, Enikő Fehér¹

¹ Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungária krt. 21, Budapest, Hungary, H-1143

² University of Veterinary Medicine, István utca 2, Budapest, Hungary, H-1078

³ Ceva-Phylaxia, Szállás utca 5, Budapest, Hungary, H-1107

⁴ Veterinary Diagnostic Directorate, National Food Chain Safety Office, Tábornok utca 2, Budapest, Hungary, H-1143

Correspondence:

Enikő Fehér and Krisztián Bányai, Institute for Veterinary Medical Research, Centre for Agricultural Research, Hungária krt. 21, Budapest, Hungary, H-1143

E-mail: feher.eniko@agrar.mta.hu, bkrota@hotmail.com

Summary

Goose hemorrhagic polyomavirus (GHPV, or *Anser anser polyomavirus 1*) is a small dsDNA virus that may cause fatal disease in young geese. In this study, GHPV positivity was examined in goose and duck samples collected in Hungary between 2005 and 2019. In this period, 384 of the investigated 1111 specimens were diagnosed as GHPV positive by PCR assay. Twenty-two GHPV genomes were sequenced and subjected to phylogenetic and evolutionary analysis. Based on the sequence data, the mean evolutionary rates were estimated $6.57 \times 10^{-6} - 5.82 \times 10^{-5}$ s/s/y for both GHPV complete genomes and individual genes, with negative selection acting on each gene. When GHPV VP1 sequences originating from wild birds were also included in the analyses, the nt and aa mutations inflated the substitution rate to 1.54×10^{-4} s/s/y that may imply adaptation of the virus to novel host species. Our data suggested the co-circulation of various GHPV strains in Hungarian goose farms; the source of these may be persistently infected domesticated or migratory wild birds. Vaccination of domestic waterfowls and detection, as well as characterization of GHPV in other avian species may help to elaborate new strategies for more effective disease control and prevention.

Keywords: goose polyomavirus, phylodynamics, wild birds

Introduction

Hemorrhagic nephritis and enteritis, a high morbidity and mortality disease of geese, was first described in 1969 in Hungary (Bernáth & Szalai, 1970). A few years later it occurred in France (Schettler, 1977; Vuillaume et al., 1982), Germany (Schettler, 1980), Poland (Gawel et al., 2014; Kozdrun et al., 2012) and in Belgium (Garmyn et al., 2017). The causative dsDNA virus, the goose hemorrhagic polyomavirus (GHPV), or recently named *Anser anser*

polyomavirus 1, belongs to the *Gammapolyomavirus* genus of the *Polyomaviridae* family (Guerin et al., 2000). The genome of GHPV is 5252-5256 bp long and comprises the genes of the VP1-VP3 structure proteins, the large and small tumor antigens (LTA and STA), and the predicted ORF-X with unknown function (Fehér et al., 2014; Johne & Müller, 2003). GHPV is known to cause disease in geese but may also replicate in other birds species without causing clinical symptoms (Corrand et al., 2011; Pingret et al., 2008). The infection affects the internal organs via tropism for endothelial cells, resulting in vascular dysfunctions, edema, ascites, hemorrhages, enteritis and neurological signs before death of young geese (Bernáth & Szalai, 1970; Dobos-Kovács et al., 2005; Guerin et al., 2000; Lacroux et al., 2004). GHPV may also persist in waterfowl farms without any signs of disease (Chunhe et al., 2018; Garmyn et al., 2017; Kozdrun et al., 2012; Leon et al., 2013; Palya et al., 2004; Wan et al., 2018).

GHPV constitute a genetically distinct gammapolyomavirus with a maximum of 69% nt identity compared to the genome of other avian polyomaviruses. Similarly to other polyomaviruses its genome seems to be highly conserved with genes typical for the genus. Although the GHPV associated syndrome has been known for fifty years, limited number of sequences and molecular epidemiological data are available. Here we processed the GHPV diagnostic results generated in 2005-2019, Hungary. Representative GHPV genomes were characterized and subjected to evolutionary analysis to reveal the mutation rate and the selection constraints affecting the viral genome.

Materials and Methods

Sample processing

Carcasses with suspected GHPV infection were sent for differential diagnosis to the Veterinary Diagnostic Directorate, National Food Chain Safety Office, Budapest, Hungary. Altogether, 1111 domestic goose and 20 duck (*Anas platyrhynchos domestica* and mallard) origin internal organ or cloacal swab samples were tested for GHPV by PCR between 2005 and 2019, and 384 goose and 3 duck (1 domestic and 2 mallard) samples proved to be positive.

Next generation sequencing

In the present study, 19 goose and 2 duck origin archived GHPV positive samples were processed for next generation sequencing of the viral genome. The genome sequence of the virus strain isolated during the first recorded GHPV-associated disease outbreak in 1969 was also sequenced (Bernáth et al., 2006; Dobos-Kovács et al., 2005). Total DNA from tissue samples (kidney and cloacal bursa) was extracted with DNeasy Blood & Tissue Kit (Qiagen). Overlapping PCR amplicons, covering the whole GHPV genome, were amplified in PCR mixtures containing 1 µl of the extracted nucleic acid, 200 nM of primers, 200 µM of dNTP mix, 1x Phusion Green buffer and 0.3 U of Phusion DNA Polymerase (Thermo Fisher Scientific) (Fehér et al., 2014). Oligonucleotide primers are available from the authors upon request. The cycling protocol included a denaturation step at 98 °C for 30 sec, followed by 40 cycles with the steps of 98 °C for 10 sec, 57 °C for 30 sec and 72 °C for 2.5 min, and a final extension step at 72 °C for 10 min. The PCR products were purified from agarose gel with the Geneaid Gel/PCR DNA Fragments Extraction Kit, subsequently pooled for each sample. Library preparation and sequencing was performed using the Ion Torrent PGM system as described elsewhere (Fehér et al., 2017; Homonnay et al., 2014). Consensus sequences were generated with CLC Genomics Workbench v7 with mapping to reference sequences

(www.qiagenbioinformatics.com/products/clc-genomics-workbench/). Genomic sequences were deposited in GenBank under the following accession numbers: MN311452-MN311473.

Evolutionary analysis

Alignments were prepared with the MUSCLE algorithm implemented in the Mega6 software (Tamura et al., 2013). Potential recombination events were detected with the RDP4 software (Martin et al., 2015). Evolutionary rate estimation and phylogenetic reconstruction was performed using the Bayesian Markov chain Monte Carlo method of the BEAST package 1.10.4 applying the best fit substitution models determined with three different software (TOPALi v2.5; Mega6; JModelTest) (Milne et al., 2008; Posada 2008; Suchard et al., 2018; Tamura et al., 2013). The data were tested with uncorrelated log normal relaxed clock, 50 million iterations and constant size coalescent tree prior. The effective sample size was confirmed with Tracer v1.6 (Rambaut et al., 2018). Substitution rates are given as substitutions per nucleotide site per year (s/s/y). Estimate uncertainty was indicated by 95% highest posterior density (HPD) values. Maximum clade credibility tree was generated and visualized with FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree>). Investigation of the selection constraint and calculation of the non-synonymous to synonymous (dN/dS) mutations was carried out with the fixed effect likelihood (FEL) and the fast unconstrained bayesian approximation for inferring selection (FUBAR) methods of the Datamonkey server (Weaver et al., 2018).

Results and discussion

The length of the 22 newly determined GHPV genomes varied between 5252-5254 bp that was the consequence of indel mutations occurring exclusively in the noncoding regions. The genome structure and the length of the open reading frames were identical to the nine

reference GHPV genomes available in GenBank, regardless of the host (duck or geese), or region (Europe or Asia) (Chunhe et al., 2018; Fehér et al. 2014). The 31 complete GHPV genomes shared very high (99.7-99.9%) nt identities with each other.

Both GHPV genome sequences and individual viral genes (VP1-VP3, LTA, STA, ORF-X) were subjected to evolutionary rate estimation, respectively. Two sequence groups were set for the VP1 gene, one extracted from complete AaPyV1 genome sequences (n=31), and another that also contained individual VP1 sequences from the GenBank (n=61), including those originating from unusual host species, i.e. white stork (*Ciconia Ciconia*), rock dove (*Columba livia*), European herring gull (*Larus argentatus*), common buzzard (*Buteo buteo*), grey heron (*Ardea cinerea*) and Eurasian jackdaw (*Corvus monedula*). The lowest VP1 nt identity values were measured for the sequences of white stork that showed $\geq 97.9\%$ identity with sequences from other hosts. Recombination events were not detected among GHPV sequences. The average evolutionary rates were $1.25 \times 10^{-5} - 5.82 \times 10^{-5}$ s/s/y for the complete genomes, and for the VP1, VP2, LTA, STA and ORF-X genes extracted from the complete genomes, and slightly lower (6.57×10^{-6} s/s/y) for the VP3 (Table 1). A higher evolutionary rate of 1.54×10^{-4} s/s/y (95% HPD 5.77×10^{-5} to 2.53×10^{-4}) was calculated for VP1 sequences when all 61 sequences were included in the analysis.

The dN/dS ratio was < 1 for all the protein coding genes, ranging between 0.0916-0.963, with the highest value for the ORF-X. Both the FEL and FUBAR methods predicted sites under negative selection in the VP1, LTA, STA and ORF-X. In case of the VP1 the dN/dS ratio was 0.442 and 0.111 for the 61 and the 31 sequences, respectively. The higher evolutionary and dN/dS ratios may be consequences of aa substitutions in the GHPV VP1 sequences of unusual host species (e.g. white stork and pigeon) that might be the result of the accumulation of mutations during possible adaptation processes (Table 1). However, the short sampling interval might also contribute to the elevated substitution rate of viruses from the unusual host

as well (Duchêne et al., 2014; Firth et al., 2010). It is of note that a broad range of 95% HPD value was detected for the substitution rate of the single genes suggesting that these data sets may be not eligible for precise evolutionary analyses because of the few nt substitution sites and short sequences (Firth et al., 2010). Positive selection (calculated with FUBAR) was predicted to act only on one site of the LTA with G100V substitution found in three Hungarian strains collected in 2015.

To evaluate the evolutionary mechanisms of GHPV on a spatiotemporal scale we used sampling dates and geographic origin of GHPV positive cases. The Hungarian GHPV positive goose samples have been collected at 71 sites in Hungary, and most of the samples originated from the middle area between the rivers Duna and Tisza, and from South Hungary. These are traditionally outstanding regions of the domestic waterfowl industry. Thus, samples used for complete genome sequencing were selected mainly from this area (Figure 1).

Fluctuation of positive case numbers was observed with peaks in 2008, 2013 and 2017 (Figure 1). No data were available about the number of goose/duck farms and sampling at different country sites and years. GHPV infection of ducks was described in three cases and all these samples were collected in Bócsa (in 2012, 2015 and 2016).

The time-scaled phylogenetic tree generated from all 61 available VP1 sequences showed segregation of the strains into two clades, an event that dated back to 1944 (95% HPD 1889-1969) (Figure 2). The Hungarian strain 1969 represented a unique clade at basal position of the phylogenetic tree. The other clade diverged into two subclades in 1976 (95% HPD 1951-1999). The Hungarian strains collected in 2007-2008 grouped in the subclade I with a Hungarian strain collected in Felpéc, 2013. Three reference sequences belonged to this subclade, including one goose origin strain from France (2000) and two duck origin strains from China (2016). Regarding the Hungarian GHPV strains collected in 2013, the sequence from Felpéc was the only which separated in different subclade and was solely sampled in the

northwestern area of Hungary. The subclade II comprised goose and duck origin Hungarian strains collected in 2011-2015, as well as European and Chinese reference strains (2001-2015). The novel Hungarian GHPV sequences separated with the date 2007 (95% HPD 2002-2009) in a distinct lineage in this subclade with a number of Polish sequences gained from goose, duck and wild birds. Some previously sequenced Hungarian GHPV strains (collected in 2000-2001) constituted another lineage.

Sequences originating from the same area and date grouped on diverse branches of the phylogenetic tree implying co-circulation of variable strains in the sampling area, which may be the consequence of reintroduction of strains from recovered animals or from the environment. Furthermore, export of live animals or migratory wild birds acting as reservoirs may also transmit GHPVs introducing distinct sequence variants to distant goose farms.

Polyomaviruses are described as slowly evolving viruses with conserved genomic sequences that helps co-divergence with limited number of host species (Torres et al., 2018). However, elevated substitution rates have been described for small dsDNA viruses with the increasing number of sequence data, the development of analysis tools, and with the examination of viral evolution independently from the host (Domingo-Calap et al., 2018; Firth et al., 2010). Recent studies determined evolutionary rates in the order of $10^{-8} - 10^{-3}$ for dsDNA viruses, including polyomaviruses, emphasizing that these are calculated values for a given data set (Aiewsakun & Katzourakis, 2016; Torres et al., 2018). In our study, a complete genome based evolutionary rate in the order of 10^{-5} s/s/y was estimated for the GHPV with purifying selection acting on its genes. The variable capsid protein sequences detected in distinct hosts may indicate that these polyomaviruses are not strictly species specific, and incorporation of these sequences into the VP1 analysis resulted in a higher substitution rate. It is conceivable, that similarly to duck, the GHPV does not cause clinical and pathological signs in the newly recognized hosts. Migratory birds may have a reservoir role promoting the development of a

prosperous transmission network of the virus. Further investigations may reveal the significance of GHPV infection in different bird species. Follow-up of the viral outbreaks at goose farms and genome analysis may help to answer questions about the epizootiology of GHPV infections.

Acknowledgements

The work was supported by the National Research, Development and Innovation Office (NKFI-OTKA, PD115519 – EF, and K120201 - KB). EF and SM was supported by the Bolyai János Fellowship program awarded by the Hungarian Academy of Sciences.

Ethics Statement

The authors confirm that no ethical approval was required as this work was carried out with collected diagnostic samples. No animal experimentation was conducted in the present study.

Conflict of Interest Statement

The authors declare no conflict of interest.

Data Availability Statement

The sequence data supporting this study are available in the GenBank with the accession numbers MN311452-MN311473.

References

Aiewsakun, P., & Katzourakis, A. (2016). Time-Dependent Rate Phenomenon in Viruses. *Journal of Virology*, 90(16):7184-95. <https://doi.org/10.1128/JVI.00593-16>

Bernáth, S., Farsang, A., Kovács, A., Nagy, E., & Dobos-Kovács, M. (2006). Pathology of goose haemorrhagic polyomavirus infection in goose embryos. *Avian Pathology*, 35(1):49-52. <https://doi.org/10.1080/03079450500465759>

Bernáth, S., & Szalai, F. (1970). Investigation for clearing the etiology of the disease appeared among goslings in 1969 (in Hungarian). *Magyar Állatorvosok Lapja*, 25:531–536.

Chunhe, W., Cuiteng, C., Longfei, C., Rongchang, L., Guanghua, F., Shaohua, S., ... Yu, H. (2018). Genomic analysis of Sheldrake origin goose hemorrhagic polyomavirus, China. *Journal of Veterinary Science*, 19(6):782-787. <http://dx.doi.org/10.4142/jvs.2018.19.6.782>

Corrand, L., Gelfi, J., Albaric, O., Etievant, M., Pingret, J.L., & Guerin, J.L. (2011). Pathological and epidemiological significance of goose haemorrhagic polyomavirus infection in ducks. *Avian Pathology*, 40(4):355-60. <https://doi.org/10.1080/03079457.2011.582481>

Dobos-Kovács, M., Horváth, E., Farsang, A., Nagy, E., Kovács, A., Szalai, F., & Bernáth, S. (2005). Haemorrhagic nephritis and enteritis of geese: pathomorphological investigations and proposed pathogenesis. *Acta Veterinaria Hungarica*, 53(2):213-23. <https://doi.org/10.1556/AVet.53.2005.2.6>

Domingo-Calap, P., Schubert, B., Joly, M., Solis, M., Untrau, M., Carapito, R., ... Bahram, S. (2018). An unusually high substitution rate in transplant-associated BK polyomavirus in vivo

is further concentrated in HLA-C-bound viral peptides. *PLoS Pathogens*, 14(10):e1007368.

<https://doi.org/10.1371/journal.ppat.1007368>

Duchêne, S., Holmes, E.C., & Ho, S.Y. (2014). Analyses of evolutionary dynamics in viruses are hindered by a time-dependent bias in rate estimates. *Proceedings of The Royal Society B Biological Sciences*, 281:20140732. <https://doi.org/10.1098/rspb.2014.0732>

Fehér, E., Lengyel, G., Dán, A., Farkas, S.L., & Bányai, K. (2014). Whole genome sequence of a goose haemorrhagic polyomavirus detected in Hungary. *Acta Microbiologica et Immunologica Hungarica*, 61:221-227. <https://doi.org/10.1556/AMicr.61.2014.2.11>

Fehér, E., Marton, S., Tóth, Á.G., Ursu, K., Wernike, K., Beer, M., ... Bányai, K. (2017). Sequence analysis of Schmallerberg virus genomes detected in Hungary. *Acta Microbiologica et Immunologica Hungarica*, 64(4):373-384.

<https://doi.org/10.1556/030.64.2017.038>

Firth, C., Kitchen, A., Shapiro, B., Suchard, M.A., Holmes, E.C., & Rambaut, A. (2010). Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Molecular Biology and Evolution*, 27(9):2038-51. <https://doi.org/10.1093/molbev/msq088>

Garmyn, A., Verlinden, M., Bosseler, L., Adriaensen, C., & Martel, A. (2017). Persistent Goose Hemorrhagic Polyomavirus Infection on a Belgian Goose Farm. *Avian Diseases*, 61(4):536-538. <https://doi.org/10.1637/11604-020317-Case.1>

Gaweł, A., Woźniakowski, G., Samorek-Salamonowicz, E., Kozdruń, W., Bobrek, K., Bobusia, K., & Nowak, M. (2014). Hemorrhagic nephritis and enteritis in a goose flock in Poland - Disease course analysis and characterisation of etiologic agent. *Avian Diseases*, 58:518–522. <https://doi.org/10.1637/10845-041014-Reg.1>

Guerin, J.L., Gelfi, J., Dubois, L., Vuillaume, A., Boucraut-Baralon, C., & Pingret, J.L. (2000). A novel polyomavirus (goose hemorrhagic polyomavirus) is the agent of hemorrhagic nephritis enteritis of geese. *Journal of Virology*, 74(10):4523-9. <https://doi.org/10.1128/JVI.74.10.4523-4529.2000>

Homonnay, Z.G., Kovács, E.W., Bányai, K., Albert, M., Fehér, E., Mató, T., ... Palya, V. (2014). Tembusu-like flavivirus (Perak virus) as the cause of neurological disease outbreaks in young Pekin ducks. *Avian Pathology*, 43(6):552-60. <https://doi.org/10.1080/03079457.2014.973832>

Johne, R., & Müller, H. (2003). The genome of goose hemorrhagic polyomavirus, a new member of the proposed subgenus Avipolyomavirus. *Virology*, 308:291–302. [https://doi.org/10.1016/S0042-6822\(02\)00103-4](https://doi.org/10.1016/S0042-6822(02)00103-4)

Kozdruń, W., Woźniakowski, G., Samorek-Salamonowicz, E., & Czekaj, H. (2012). Viral infections in goose flocks in Poland. *Polish Journal of Veterinary Sciences*, 15(3):525-30. <https://doi.org/10.2478/v10181-012-0080-9>

- Lacroux, C., Andreoletti, O., Payre, B., Pingret, J.L., Dissais, A., & Guerin, J.J. (2004). Pathology of spontaneous and experimental infections by goose haemorrhagic polyoma virus. *Avian Pathology*, 33:351–358. <https://doi.org/10.1080/0307945042000220525>
- Leon, O., Corrand, L., Bich, T.N., Le Minor, O., Lemaire, M., Guérin, J.L., 2013. Goose Hemorrhagic polyomavirus detection in geese using real-time PCR assay. *Avian Dis.* 57(4):797-9.
- Martin, D.P., Murrell, B., Golden, M., Khoosal, A., & Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, 1:vev003. <https://doi.org/10.1093/ve/vev003>
- Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D.F., & Wright, F. (2008). TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, 25(1):126-127. <https://doi.org/10.1093/bioinformatics/btn575>
- Palya, V., Ivanics, E., Glávits, R., Dán, A., Mató, T., & Zarka, P. (2004). Epizootic occurrence of haemorrhagic nephritis enteritis virus infection of geese. *Avian Pathology*, 33:244–250. <https://doi.org/10.1080/0307945042000195740>
- Pingret, J.L., Boucraut-Baralon, C., & Guérin, J.L. (2008). Goose haemorrhagic polyomavirus infection in ducks. *Veterinary Record*, 162(5):164. <http://dx.doi.org/10.1136/vr.162.5.164-a>

Posada, D. (2008). jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, 25(7):1253-6. <https://doi.org/10.1093/molbev/msn083>

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., & Suchard, M.A. (2018). Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, 67:901-904. <https://doi.org/10.1093/sysbio/syy032>

Schettler C.H. (1980). Clinical picture and pathology of haemorrhagic nephritis and enteritis in geese. *Tierärztliche Praxis Kleintiere*, 8:313–320. doi

Schettler, C.H. (1977). Détection en France de la néphrite hémorragique et entérite de l'oie (NHEO). [Detection of Hemorrhagic Nephritis Enteritis of Geese (HNEG) in France]. *Recueil de Médecine Vétérinaire*, 153:353–355.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, 4:vey016. <https://doi.org/10.1093/ve/vey016>

Tamura, K., Stecher, G., Peterson, D., Filipowski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution*, 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>

Torres, C., Barrios, M.E., Cammarata, R.V., Victoria, M., Fernandez-Cassi, X., Bofill-Mas, S., ... Mbayed, V.A. (2018). Phylodynamics of Merkel-cell polyomavirus and human

polyomavirus 6: A long-term history with humans. *Molecular Phylogenetics and Evolution*, 126:210-220. <https://doi.org/10.1016/j.ympev.2018.04.025>

Vuillaume, A., Tournut, J., & Banon, H. (1982). A propos de la maladie des oisons d'apparition tardive ou néphrite hémorragique-entérite de l'oie (N.H.E.O.) [About the late disease onset in goslings or goose hemorrhagic enteritis nephritis (HNEG)]. *Revue de Médecine Vétérinaire*, 133:341–346.

Wan, C., Cheng, L., Fu, G., Chen, C., Liu, R., Shi, S., ...Huang, Y. (2018). Rapid detection of goose hemorrhagic polyomavirus using TaqMan quantitative real-time PCR. *Molecular and Cellular Probes*, 39:61-64. <https://doi.org/10.1016/j.mcp.2018.04.003>

Weaver, S., Shank, S.D., Spielman, S.J., Li, M., Muse, S.V., & Pond, S.L.K. (2018). Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Molecular Biology and Evolution*, 35(3):773–777. <https://doi.org/10.1093/molbev/msx335>

Gene/genome	Evol. rate model	Evolutionary rate	FEL (p=0.1/p=0.05)	FUBAR (0,9<)	dN/dS
GHPV VP1 (n=61)	HKY+G	1.54x10 ⁻⁴ (5.77x10 ⁻⁵ – 2.53x10 ⁻⁴)	5/3 neg	4 neg	0.442
GHPV VP1 (n=31)	HKY+G+I	4.82 x10 ⁻⁵ (1.03 x10 ⁻⁸ – 1.24 x10 ⁻⁴)	2/0 neg	2 neg	0.111
GHPV VP2	HKY	1.25x10 ⁻⁵ (6.76 x10 ⁻¹² – 5.46x10 ⁻⁵)	-	-	0.194
GHPV VP3	HKY	6.57x10 ⁻⁶ (5.27x10 ⁻¹³ – 3.38x10 ⁻⁵)	1/0 neg	-	0.644
GHPV LTA	HKY+G	4.87x10 ⁻⁵ (4.39 x10 ⁻⁶ – 1.00 x10 ⁻⁴)	7/4 neg	13 neg 1 pos	0.0916
GHPV STA	HKY+G	5.82x10 ⁻⁵ (3,05x10 ⁻¹⁰ – 1.37x10 ⁻⁴)	2/2 neg	2 neg	0.111
GHPV 1 ORF-X	HKY	2.98x10 ⁻⁵ (1.73 x10 ⁻⁹ – 8.53x10 ⁻⁵)	1/0 neg	1 neg	0.963

complete genome	HKY+G+I	3.027×10^{-5} ($1.09 \times 10^{-5} - 5.33 \times 10^{-5}$)	-	-	-
-----------------	---------	--	---	---	---

TABLE 1 Evolutionary rate estimation and dN/dS ratios for GHPV genomes and individual genes. Substitution rates are expressed as substitution per nucleotide site per year (s/s/y). The table contains the number of sites under selection constraint (negative, neg; positive; pos)

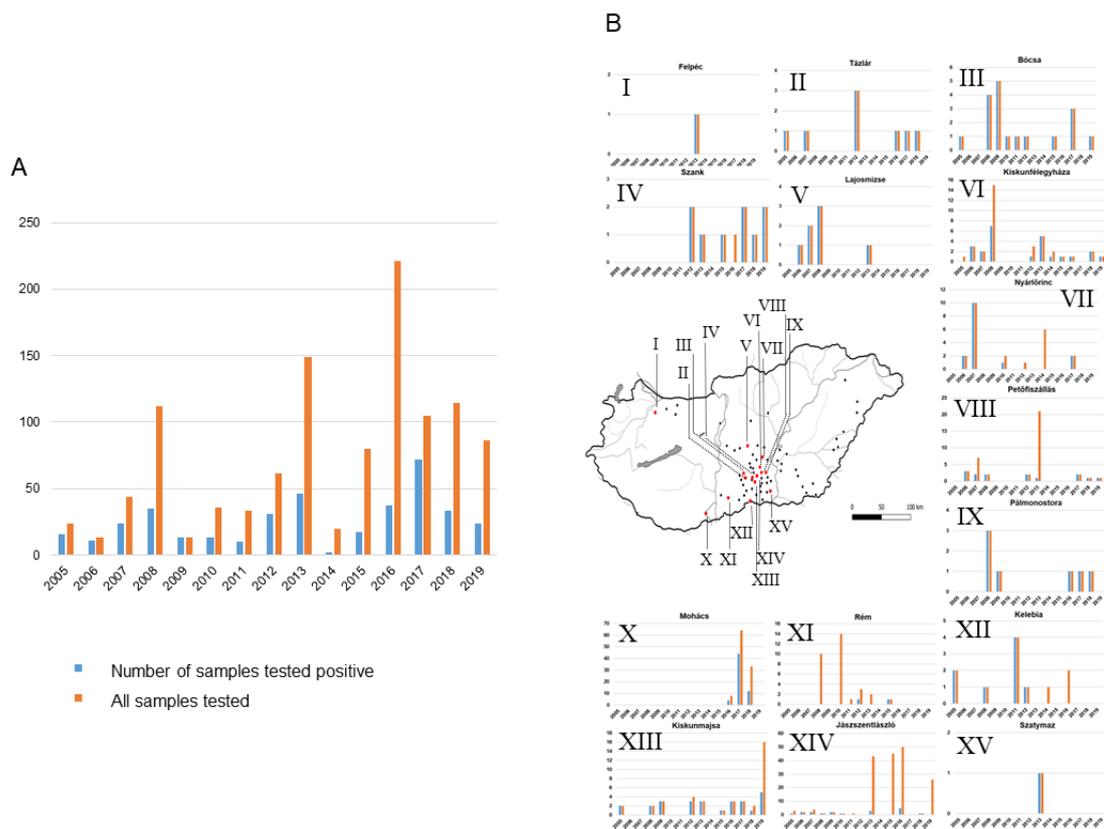


FIGURE 1 Spatiotemporal distribution of GHPV in Hungary. (a) Number of samples tested positive by PCR (blue) and all samples tested (orange). (b) Location of 71 sampling sites where GHPV was detected in the last 15 years. Diagrams represent GHPV positivity at some highlighted Hungarian sites. X axis, sampling year; Y axis, number of samples tested positive by PCR (blue) and all samples tested (orange)

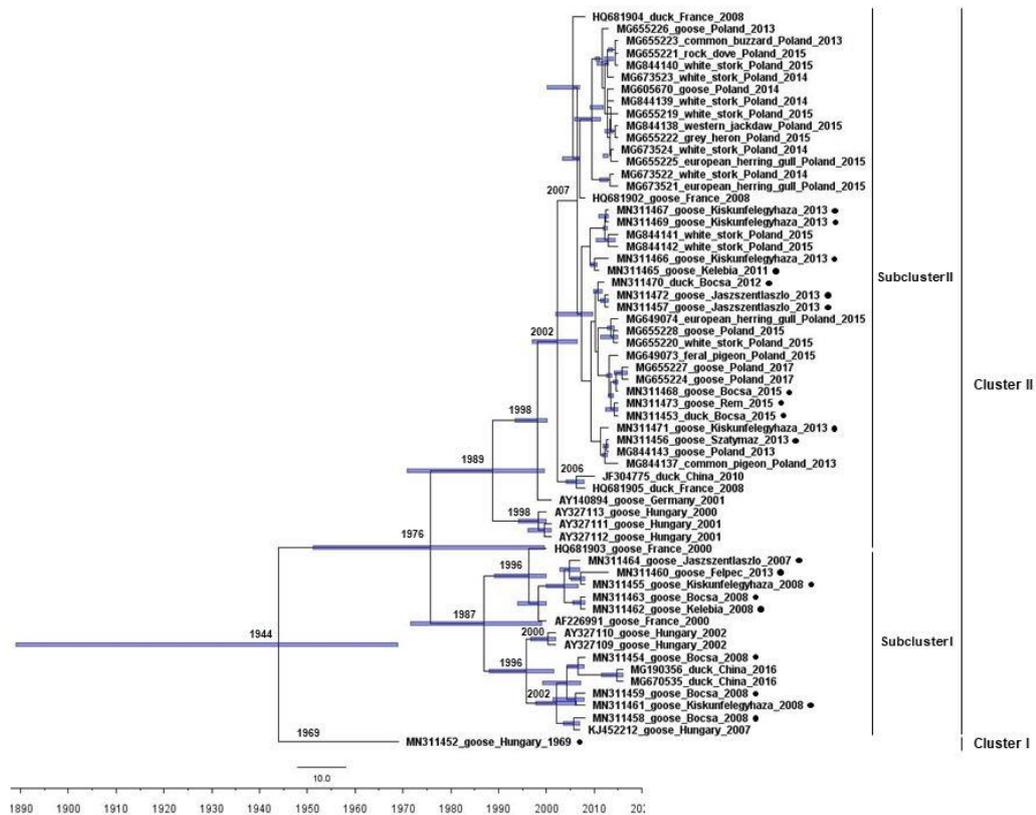


FIGURE 2 Time-calibrated maximum clade credibility tree of GHPV VP1 sequences. Blue bars indicate 95% highest posterior density intervals for the age of the nodes. The novel Hungarian sequences are highlighted by black dots