JOURNAL OF GENERAL VIROLOGY

RESEARCH ARTICLE

Virtanen *et al.*, *Journal of General Virology* 2019;100:227–236

DOI 10.1099/jgv.0.001187

MICROBIOLOGY SOCIETY

Co-circulation of highly diverse Aleutian mink disease virus strains in Finland

Jenni Virtanen,[1,*] Teemu Smura,[2] Kirsi Aaltonen,[1] Anna-Maria Moisander-Jylhä,[1] Anna Knuuttila,[1]† Olli Vapalahti[1,2] and Tarja Sironen[1,2]

Abstract

Aleutian mink disease virus (AMDV) is the causative agent of Aleutian disease (AD), which affects mink of all genotypes and also infects other mustelids such as ferrets, martens and badgers. Previous studies have investigated diversity in Finnish AMDV strains, but these studies have been restricted to small parts of the virus genome, and mostly from newly infected farms and free-ranging mustelids. Here, we investigated the diversity and evolution of Finnish AMDV strains by sequencing the complete coding sequences of 31 strains from mink originating from farms differing in their virus history, as well as from free-ranging mink. The data set was supplemented with partial genomes obtained from 26 strains. The sequences demonstrate that the Finnish AMDV strains have considerable diversity, and that the virus has been introduced to Finland in multiple events. Frequent recombination events were observed, as well as variation in the evolutionary rate in different parts of the genome and between different branches of the phylogenetic tree. Mink in the wild carry viruses with high intra-host diversity and are occasionally even co-infected by two different strains, suggesting that free-ranging mink tolerate chronic infections for extended periods of time. These findings highlight the need for further sampling to understand the mechanisms playing a role in the evolution and pathogenesis of AMDV.

INTRODUCTION

Aleutian mink disease virus (AMDV), a member of the family *Parvoviridae* and genus *Amdoparvovirus* [1–3], belongs to the species *Carnivore amdoparvovirus 1* [4]. It has an icosahedral, non-enveloped virion and a 4.7 kb ssDNA genome [2, 5, 6], which encodes two structural proteins (VP1 and VP2) and three non-structural proteins (NS1, NS2 and NS3) [1, 7–9]. AMDV causes Aleutian disease (AD), one of the most important infectious diseases of farmed mink, causing significant welfare problems for the animals as well as financial losses to the farmers. AMDV induces high antibody titres, plasmacytosis and immune complex disease, the clinical signs of which range from subclinical to severe and fatal [10]. Despite the efforts to identify, diagnose and eradicate it, AD has spread to all mink-producing countries [11]. AMDV is resistant to many standard physical and chemical treatments, making it difficult to destroy the virus and prevent its spread [12, 13].

AD was first described in 1956 in the USA [14], and soon thereafter in Sweden and Denmark [11], being detected first in the highly susceptible Aleutian-type mink but later reported in other mink genotypes as well [11]. The virus likely existed in wild mink long before the first Aleutian-type mink were born in 1941, but it remained undetected until the disease-susceptible Aleutian mink was bred [15]. Currently, AD diagnosis is based on pathological findings, on serological methods such as counterimmunoelectrophoresis and ELISA, and on PCR [16–18]. In Finland, most farmed mink are annually screened with automatized ELISA, with the seroprevalence of mink tested from 1980 to 2014 fluctuating between 3 and 60 % [19].

In addition to mink, AMDV infects a wide range of other species like ferrets, polecats, martens, otters, raccoons, striped skunks, bobcats and common genets [19–22]. Other amdoparvoviruses have been found from striped skunks, raccoons and foxes [3, 23–26].

Received 13 June 2018; Accepted 9 November 2018; Published 10 December 2018

Author affiliations: [1]Department of Veterinary Biosciences, Faculty of Veterinary Medicine, University of Helsinki, Agnes Sjöbergin katu 2, 00790, Helsinki, Finland; [2]Department of Virology, Faculty of Medicine, University of Helsinki, Haartmaninkatu 3, 00290, Helsinki, Finland.

***Correspondence:** Jenni Virtanen, jenni.me.virtanen@helsinki.fi

Keywords: AMDV; NGS; phylogeny; molecular epidemiology.

Abbreviations: AD, Aleutian disease; AMDV, Aleutian mink disease virus; dNTP, deoxyribonucleotide triphosphate; GTR, general time-reversible model; HDP, highest posterior density; HKY, Hasegawa—Kishino—Yano model; MRCA, most recent common ancestor; NGS, next generation sequencing.

†Present address: Anna Knuuttila, Fimmic Oy, Helsinki, Finland.

GeneBank accession numbers for the sequences are MG821234–MG821259 and MG821261-MG821309.

Three supplementary figures and three supplementary tables are available with the online version of the article.

Studies on the molecular epidemiology of AD in Finland have mainly focused on newly infected farms or free-ranging animals, and only partial sequences are known [19, 27]. Here we sequenced complete or partial sequences of Finnish AMDV strains from newly infected farms and farms that had struggled with the disease longer. We also recovered additional complete sequences from free-ranging mink. These sequences were compared with each other and global strains [28–33] to understand diversity, phylogeny and transmission of the virus in Finland.

RESULTS

AMDV sequenced with next generation and Sanger sequencing

We obtained 52 samples from Finnish fur farms and 45 free-ranging mustelid samples from Finland and Estonia [19], all of which were known to be AMDV-positive. From these 97 samples, we attempted to sequence the whole AMDV coding sequence in one part or in two fragments using a previously published approach [29] with modified primers. The entire coding sequence was acquired from 31 samples, including nine free-ranging mink and 22 farmed mink, and a sequence missing the first 1500 bp was acquired from three additional samples (F4, F9 and F11). Only smaller AMDV fragments were obtained from mink F31 and marten W106, and no sequence was obtained from the rest of the samples with this method.

Smaller fragments of the genome were sequenced from the samples of farmed mink if whole-genome sequencing was unsuccessful. This was done by using two PCR reactions targeting either nt 578–951 or nt 1662–2302 (nt sites are according to strain AMDV-G (GenBank accession no. M20036.1) throughout the manuscript). With this approach, we obtained additional sequences from 17 samples. Four additional partial sequences (same two regions) from Finnish farm samples (F45–F48) were also included in further analysis. Altogether, we obtained AMDV sequences from 57 strains (Table 1).

Co-infections

To study the presence of possible co-infections, we randomly selected six farm strains (F1, F2, F19, F46, F32 and F43) and all nine free-ranging strains for analysis. We amplified one of the regions (nt 578–951), cloned the PCR-product, and obtained sequences from seven to ten clones of each sample. Based on the acquired sequences and a neighbour-joining tree built from those (Fig. S1, available in the online version of this article), none of the farmed mink seemed to be infected by multiple strains whereas one of the free-ranging mink (W249) was infected by two distinct virus strains. Hence, W249 was excluded from further analysis.

General description of the sequences

Non-structural proteins of all the sequenced strains had the same length: NS1 of 641 aa and NS2 of 114 aa. The length of the VP1 was either 689 or 690 aa, and that of VP2 646 or 647 aa, with some of the strains having one aa deletion in the hypervariable region (aa 232–243 of VP2). The pairwise nucleotide distances of the entire coding sequences (nt 206–4349) were almost equally high between the Finnish strains (0–8.6 %) as they were between all the sequenced strains from this study and GenBank (0–9.3 %). There was no significant difference in distances within Finnish farm strains (0–8.2 %) or Finnish free-ranging strains (1.0–7.9 %). Distances were higher within the NS coding region (0–13.2 %, nt 206–2213) than the VP coding region (0–6.2 %, nt 2206–4351) using all sequenced complete coding regions. Distances within farms ranged from 0 to 0.6 %, apart from farm 19, which had a mean distance of 6.7 % based on nt 578–951.

Variation within samples

To study the frequencies of single nucleotide polymorphisms within samples, the quality filtered next generation sequencing (NGS) reads were assembled against the consensus sequence of a given sample, and the frequency of single nucleotide variants in each sequence position was calculated. AMDV strains sequenced from free-ranging mink had a significantly higher intra-strain variant frequency than strains sequenced from farmed mink (Fig. S2). The mean variant frequency in complete coding sequences was 2.11×10^{-3} for free-ranging strains and 5.52×10^{-4} for farm strains (Mann–Whitney *U*-test, $P<0.01$). The respective frequencies for free-ranging and farmed strains were 2.19×10^{-3} and 8.55×10^{-4} ($P<0.01$) when the NS coding region of the genome was used (nt 206–2211) and 2.07×10^{-3} and 3.10×10^{-4} ($P<0.01$) for the structural part of the genome (nt 2206–4349). The intra-strain variant frequencies of the farm strains were also higher in the NS coding region compared to the structural region (related-samples Wilcoxon signed-rank test, $P<0.01$). The difference between these two genome regions was not statistically significant in the virus strains sequenced from the free-ranging mink.

Complete coding sequences analysed for recombination

We searched for possible recombination events in all 30 complete coding sequences with Simplot and seven methods implemented in the package RDP V4.95. RDP detected potential recombination in 20 sequences. Events that were detected by at least four programs, together with their *P*-values, are presented in Table S1. Phylogenetic trees were then constructed from nt 1–700, 1165–1864 and 2446–3443 where RDP did not detect any clear evidence of recombination. Four sequences representing the Danish outbreak [30] and all the other AMDV strains for which the complete coding sequences and sampling years are available in GenBank were also included. To visualize the recombination events, the sequences were classified into three groups according to highly supported deep nodes in a phylogenetic tree constructed from nt 1–700 (Fig. 1a). Based on this region, all sequences clustered into three groups (referred to as A–C) and further into several sub-clusters. The phylogenetic trees constructed based on different genomic regions were clearly incongruent. While three major clusters

Table 1. Sample information

Sample ID	Farm no.	AMDV status*	Parts sequenced†	Country/region	GenBank accession no.
F1	1	AMDV free	98–4447	Finland/Ostrobothnia	MG821234
F2	2	AMDV free	98–4447	Finland/Ostrobothnia	MG821235
F3	2	AMDV free	98–4447	Finland/Ostrobothnia	MG821236
F4	3	AMDV free	1469–4447	Finland/Northern Ostrobothnia	MG821265
F5	4	AMDV free	94–4439	Finland/Ostrobothnia	MG821237
F6	5	AMDV free	98–4447	Finland/Central Ostrobothnia	MG821238
F7	6	AMDV free	127–4429	Finland/Ostrobothnia	MG821239
F8	7	AMDV free	98–4446	Finland	MG821240
F9	8	AMDV free	1469–4447	Finland/Central Ostrobothnia	MG821266
F10	8	AMDV free	589–940, 1677–2288	Finland/Central Ostrobothnia	MG821270, MG82129
F11	9	AMDV free	1469–4439	Finland	MG821267
F12	10	AMDV free	98–4447	Finland/Southern Ostrobothnia	MG821241
F13	11	AMDV free	143–4439	Finland	MG821242
F14	12	AMDV free	98–4447	Finland/Ostrobothnia	MG821243
F15	12	AMDV free	98–4447	Finland/Ostrobothnia	MG821244
F16	13	AMDV free	98–4447	Finland/Central Ostrobothnia	MG821245
F17	20	AMDV free	578–951	Finland/Ostrobothnia	MG821271
F18	21	AMDV free	578–951, 1676–2296	Finland/Northern Ostrobothnia	MG821272, MG821292
F19	22	AMDV free	589–942, 1682–2290	Finland/Southern Ostrobothnia	MG821273, MG821293
F20	23	AMDV free	578–942, 1676–2294	Finland/Ostrobothnia	MG821274, MG821294
F22	24	AMDV free	589–941, 1676–2294	Finland/Southern Ostrobothnia	MG821275, MG821295
F23	26	AMDV free	589–940, 1682–2291	Finland/Central Ostrobothnia	MG821276, MG821296
F24	14	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821246
F25	14	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821247
F26	14	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821248
F27	14	AMDV positive	98–4446	Finland/Northern Ostrobothnia	MG821249
F28	15	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821250
F29	15	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821251
F30	15	AMDV positive	151–4429	Finland/Northern Ostrobothnia	MG821252
F31	15	AMDV positive	Fractions between 146–3111‡	Finland/Northern Ostrobothnia	MG821268
F32	15	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821253
F33	15	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821254
F34	15	AMDV positive	98–4447	Finland/Northern Ostrobothnia	MG821255
F35	16	AMDV positive	578–951, 1675–2294	Finland/Northern Ostrobothnia	MG821277, MG821297
F36	17	AMDV free	588–942, 1676–2288	Finland/Central Ostrobothnia	MG821278, MG821298
F37	17	AMDV free	588–940, 1679–2290	Finland/Central Ostrobothnia	MG821279, MG821299
F38	17	AMDV positive	578–951, 1717–2294	Finland/Central Ostrobothnia	MG821280, MG821300
F39	17	AMDV positive	578–951	Finland/Central Ostrobothnia	MG821281
F40	17	AMDV positive	578–951, 1675–2295	Finland/Central Ostrobothnia	MG821282, MG821301
F41	17	AMDV positive	578–951, 1675–2295	Finland/Central Ostrobothnia	MG821283, MG821302
F42	17	AMDV positive	578–951, 1682–2295	Finland/Central Ostrobothnia	MG821284, MG821303
F43	17	AMDV positive	578–951, 1675–2295	Finland/Central Ostrobothnia	MG821285, MG821304
F44	18	AMDV positive	578–951, 1711–2258	Finland/Northern Ostrobothnia	MG821286, MG821305
F45	19	AMDV positive	627–942, 1691–2250	Finland/Northern Ostrobothnia	MG821287, MG821306
F46	19	AMDV positive	627–942, 1676–2250	Finland/Northern Ostrobothnia	MG821288, MG821307
F47	19	AMDV positive	638–942, 1676–2250	Finland/Northern Ostrobothnia	MG821289, MG821308
F48	19	AMDV positive	627–942, 1676–2250	Finland/Northern Ostrobothnia	MG821290, MG821309
W47	Free-ranging mink		98–4447	Finland/South Karelia	MG821256
W51	Free-ranging mink		98–4446	Finland/South Karelia	MG821257
W181	Free-ranging mink		98–4447	Finland/Tavastia Proper	MG821258
W235	Free-ranging mink		98–4447	Finland/South Karelia	MG821259
W249	Free-ranging mink		150–4447‡§	Finland/Southern Savonia	

Table 1. cont.

Sample ID	Farm no.	AMDV status*	Parts sequenced†	Country/region	GenBank accession no.
W327	Free-ranging mink		98–4447	Finland/Central Finland	MG821261
W454	Free-ranging mink		98–4447	Finland/South Karelia	MG821262
W456	Free-ranging mink		98–4447	Finland/Uusimaa	MG821263
W458	Free-ranging mink		98–4447	Finland/Uusimaa	MG821264
W106	Free-ranging marten		Fractions between 110–1343‡	Estonia	MG821269

*AMDV status of the farm at the time samples were taken. Samples from farms with AMDV-free status were first AMDV-positive samples from those farms.

†Parts sequenced compared to AMDV-G.

‡Small coverage values (between 10 and 100).

§Sequence excluded from the analysis because of co-infection.

were formed in both trees constructed based on NS protein-coding sequences (i.e. 1–700 and 1165–1864), distinct strains/subclusters changed their positions both within and between the clusters. In addition, the clustering pattern in the VP protein-coding region was incongruent to that observed in the NS protein-coding regions. For example, strains W181 and W454 were placed differently in the second tree compared to the other two trees, and strain F13 changed position between the NS and VP protein-coding regions (Fig. 1a–c).

The changes in the clustering pattern of distinct recombinant strains (suggested by RDP) into major clades A–C (pre-determined based on nt 1–700) were further assessed using a grouping scan (Fig. 1d) method implemented in the SSE package. These analyses revealed major recombination breakpoints in the NS coding region (approx. nt 1000) and in the VP coding region (approx. nt 2000) and suggested that many of the virus strains had two or more recombination breakpoints at varying positions in the genome.

Despite the apparently common incongruence in the deep nodes of the phylogenetic tree, all the subclusters except 2, 4 and 10 remained congruent throughout the coding sequence (Fig. 1a–c). This suggests that some of the recombination events detected have occurred between the ancestors of the strains in the dataset or that the parental strains were not represented.

Evolutionary relationships analysed with phylogenetic trees

To assess the rate and timeline of AMDV evolution and the relationships between the virus strains derived from different

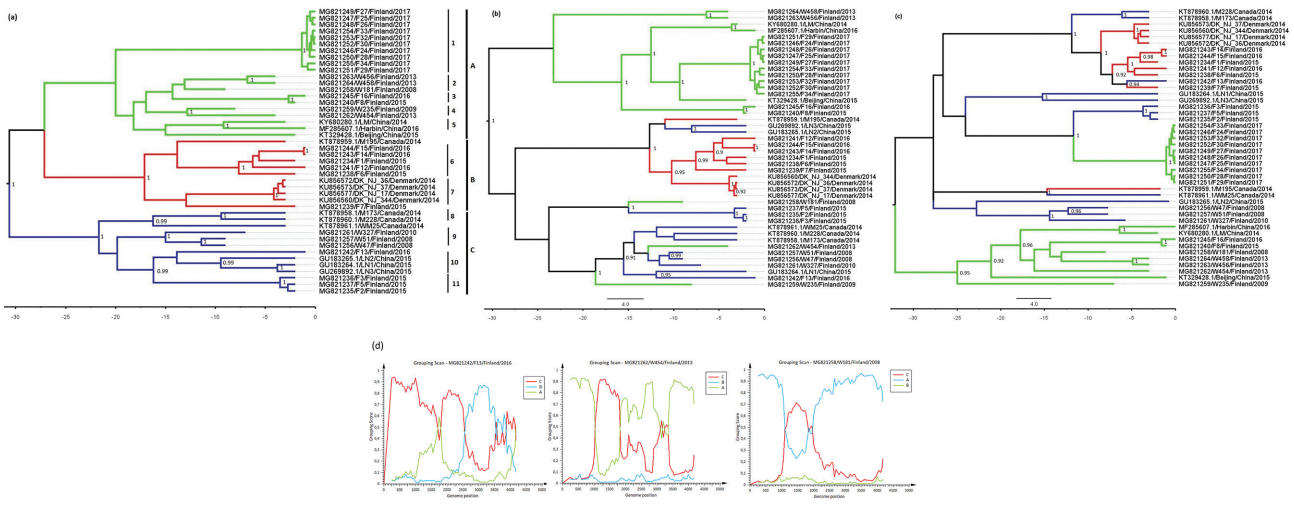

Fig. 1. Recombination and phylogenetic comparison of different regions of AMDV genome. Trees were constructed from the regions 1–700 (a), 1165–1864 (b) and 2446–3443 (c) using Beast v1.8.0. Best fitting evolutionary models were defined using MEGA. Timescales are shown under trees, posterior probabilities above 0.9 next to the nodes, and branches are colour coded based on clusters in the tree A to visualize recombination. The grouping scan (500 nt window and 50 nt steps) from F13, W181 and W454 (d) using groups determined in (a) was performed with SSE. Finnish strains starting with F are from farmed mink and strains starting with W are from free-ranging mink.

farms and free-ranging animals, trees were constructed from three partial regions: nt 578–951, nt 1662–2302 and nt 2417–3413, including all GenBank sequences with known sampling years from each region (Figs 2 and S3). The tree from nt 2417–3413 includes a molecular clock but attempts to build a molecular clock from nt 578–951 and nt 1662–2302 were unsuccessful. Regions nt 578–951 and nt 1662–2302 were chosen based on the large amount of sequences available, and nt 2417–3413 because of the absence of recombination. The entire coding sequences were not used because of the observed recombination events.

Finnish virus strains can be found from almost all of the major branches in the phylogenetic trees, forming 3–8 different clusters instead of a single country-specific cluster (Figs 2a and S3). Typically, sequences from one farm clustered together, with the exception of farm 19 (F45–F48), in which highly divergent strains were present at the time of sampling.

The smallest estimated time of most recent common ancestor (tMRCA) for the Finnish strains and strains from other countries in the phylogenetic tree based on nt 2417–3413 was 16 years (95 % HDP=9–28 years) (Fig. 2b, node E). The smallest estimated separation time of Finnish free-ranging strains and farm strains was 18 years (95 % HDP=12–30 years) (Fig. 2b, node H).

The mean evolutionary rates calculated from the trees of Fig. 1, the rates were 3.3867×10^{-3} (95 % highest posterior density, HPD, interval [1.8161×10^{-3}, 5.0036×10^{-3}]), 4.478×10^{-3} (95 % HPD interval [2.4571×10^{-3}, 6.7893×10^{-3}]) and 1.8778×10^{-3} (95 % HPD interval [9.4357×10^{-4}, 2.9117×10^{-3}]) nt substitutions/site/year for the regions 1–700, 1165–1864 and 2446–3443. The mean rate from the tree containing all available sequences of region 2417–3413 (Fig. 2b) was 9.0529×10^{-4} (95 % HPD interval [4.3423×10^{-4}, 1.3848×10^{-3}]). The evolutionary rate varied between some branches, e.g. the branch separating farms with continuously circulating virus (samples F24–F34) from others had a higher evolutionary rate in the capsid protein-encoding region (Fig. 2b, node A).

DISCUSSION

In this study, we sequenced the entire genome from 30 samples and shorter regions from several other samples from

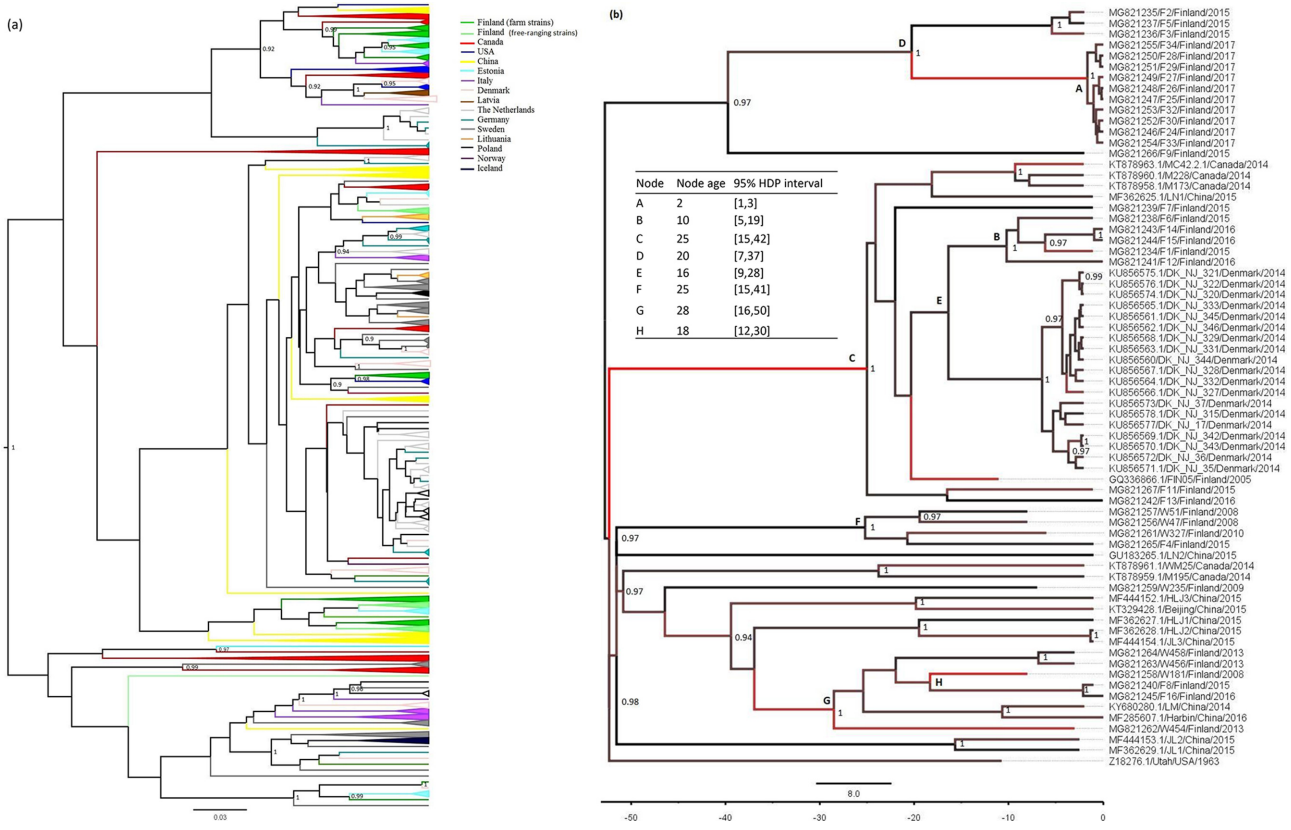

Fig. 2. Evolutionary rate and molecular clock analysis using Bayesian phylogeny. Phylogenetic tree from nt 578–951 (a) without molecular clock and from nt 2417–3413 (b) with molecular clock. Posterior probabilities above 0.9 are shown next to the nodes. Sequences from different countries in (a) are colour coded. Node ages and ranges of selected nodes of (b) are marked next to the tree, timescale is shown under it, and evolutionary rate from slower to faster is indicated with black-red gradient. Finnish strains starting with F are from farmed mink and strains starting with W are from free-ranging mustelids.

both farmed and free-ranging mink. We demonstrated that AMDV strains are highly diverse in Finland and globally (Figs 2a and S3), which has also been observed previously [34], and suggest that there has been extensive virus exchange between countries. These results support the hypothesis that AMDV has, in recent decades, been introduced into Finnish mink multiple times [27]. The first mink were transported to Scandinavia from Canada and the USA by 1920, and mink farming in Finland began in the 1930s, but the exact origin of Finnish mink is unclear [35, 36]. The first observations of free-ranging mink in Finland occurred already in 1932 [37]. Currently, around 300 mink farms operate in Finland and over 95 % of them are located in Ostrobothnia, the western part of the country [35].

The free-ranging and farm strains were mixed in the phylogenetic trees, suggesting that the virus is transmitted between farms and the wild. However, it is not clear how much of this occurred decades ago and how much is still happening today, as we found no evidence of recent virus transmission between farms and the wild, and the smallest suggested separation time between the virus lineages detected from farm and free-ranging animals was 18 years. The strains from Finnish free-ranging animals do not form their own group, in contrast to the AMDV strains in wild mink in Poland [38]. Yet many of the virus lineages detected from free-ranging mink appear to evolve independently from farm strains, and it is possible that the virus has spread to Finnish wildlife as soon as mink farming started in Finland. However, we cannot exclude the possibility that there is also current virus transport between farms and the wild based on the limited sample size. It should also be noted that, with the current, limited dataset, the reported tMRCAs should be treated as rough estimates, especially regarding the deep nodes of the phylogenetic trees, where the confidence intervals of the estimated tMRCAs span several decades.

The AMDV strains from free-ranging mink appear to have significantly higher within-sample variation than farmed strains have. One explanation for this observation is that free-ranging mink may have carried the virus for a longer time than farmed animals, which mostly live for less than a year, and hence the virus has less time to evolve. Another explanation for the difference in variation could be that free-ranging animals are more frequently infected with more than one virus strain. The mink is a territorial animal and might not be exposed to several virus strains directly, but since AMDV remains infectious for a long time in the environment, indirect exposures are likely to occur frequently. Although evidence of co-infection was observed only in one free-ranged individual (W249) (Fig. S1), this does not rule out the possibility of co-infection with a low copy number virus strain. Furthermore, co-infections have been reported earlier in parvoviruses [28, 39], and co-infections by closely related strains are difficult to detect.

AMDV strains seem to spread efficiently between and within farms. Their efficient spread may lead to farms having multiple strains simultaneously, which, if several virus strains infect one animal, may lead to recombination. Recombination events increase the genetic diversity of circulating virus strains, and potentially, may lead to the appearance of more virulent strains. Previous studies have shown frequent recombination among parvoviruses [40–42]. For AMDV, the Canadian virus strains have reportedly undergone recombination during intense farming practices [28], but previously, no recombination has been detected in Danish [30] or Finnish strains [27]. However, these studies focused on one particular epidemic [30] or used a small, partial genome region [27], hampering the detection of recombination event. Only one of the Finnish farms in our study displayed signs of having more than one virus strain. However, the sample size from the farms was relatively small, eight AMDV-positive samples per farm at most, which is too few to exclude the possibility that other farms may also have been infected with more than one virus strain. Clear evidence of recombination was detected both in farm and in free-ranging samples, and several recombination events were supported both by phylogenetic trees and by grouping scan analysis (e.g. W181, W454 and F13, Fig. 1). Most evident recombination breakpoints were similar to those reported previously [28] and some of the possible recombination events appear to have happened a long time ago. For example, one event appeared in all the sequences in group 6, suggesting that it had occurred years ago (Table S1, event six), possibly even before the ancestral virus strain was introduced in Finland. This high level of variation and number of possible recombination events in the distant past were clearly reflected in the phylogenetic analysis: the phylogenetic trees showed highly variable topology depending on the genomic region. The clustering pattern of closely related sequences remained similar in all trees, while the relationships between more distant groups varied. Variation of this kind and the several recombination breakpoints in different regions of the genome complicate the phylogenetic analysis of AMDV strains.

Our analysis suggested a somewhat higher substitution rate for AMDV as compared to those observed in canine parvovirus, feline panleukopenia virus and several other ssDNA viruses [43–46]. Hypothetically, this might be due to the intensive farming practices, assuming that in crowded conditions virus transmission between animals occurs more frequently than in less-crowded wild and companion animal populations, leading to more virus replication cycles per year. Further, in several cases, the evolutionary rate was clearly higher in certain branches of the tree (Fig. 2b). Interestingly, within sample variation was higher in the NS region of farm but not free-ranging strains. Among the farmed animals, this may be either due to more stringent negative selection posed on the VP region than the NS region or increased positive selection in the NS region. However, further studies are needed to understand this observation.

Attempts to obtain complete genome sequences from free-ranging animals other than mink, and some samples from AMDV-positive farms, were unsuccessful. This is probably due to the high diversity of AMDV strains, and thereby mismatches with our primers. Amplification of these strains would require further optimization of the primers able to amplify more divergent ADMV strains, or the development of alternative, PCR-free, approaches. This new information on the diversity of Finnish strains will enable us to improve diagnostic tests, locate sources of infection, and thus hopefully have better success in eradication efforts. These results will also help us to define the effect farming and man-made populations have on disease ecology and virus–host dynamics.

METHODS

Samples and selection of AMDV-positive individuals

This study included samples from Finnish fur farms differing in their AMDV status and free-ranging mustelids from Finland and from Estonia. Fin Furlab provided 52 samples from 24 fur farms whose AMDV seroprevalence varied between 0 and 86 %. DNA had already been isolated from spleen tissue, and all samples were AMDV DNA-positive based on diagnostic PCR described by Knuuttila *et al.* [19]. The samples were collected in 2015–2017. Because all the farm samples were originally collected for diagnostic purposes, no ethical permission for animal experiments was mandatory. The AMDV DNA-positive free-ranging mustelid samples originated from Knuuttila *et al.* [19] and included 33 samples from mink, five from martens, two from polecats and four from badgers originating from southern and eastern Finland, and from Estonia.

Whole-genome sequencing

The complete coding regions were first amplified by PCR; the amplification products were then subjected to Illumina MiSeq sequencing. Modified primers from Hagberg *et al.* [29] in PCR allowed amplification of the entire coding sequence. Primers were modified to match all the complete AMDV sequences published in GenBank by summer 2016, and the AMDV genome was amplified in one fragment by the primers AMDVF1 and AMDVR3, or in two fragments by the primers AMDVF1 and AMDVR2 as well as AMDVF2 and AMDVR3. All the primer sequences of this study are shown in Table S2. The product sizes were 4390, 3240 and 3022 bp. PCR reactions contained 0.5 µl of PrimeSTAR GXL DNA polymerase (TaKaRa), 5 µl of 5x PrimeSTAR GXL buffer, 2 µl of 2.5 mM deoxyribonucleotide triphosphate (dNTP), 0.75 µl of 10 µM forward primer, 0.75 µl of 10 µM reverse primer, 11 µl of water and 5 µl of template DNA. The cycling conditions were: 35 cycles (PCR in one fragment) or 30 cycles (PCR in two fragments) of denaturation at 98 °C for 10 s, annealing at 57 °C for 15 s, and extension at 68 °C for 4.5 min. PCR products were analysed in a 0.8 % agarose gel stained with GelRed (Biotium). These products were purified according to manufacturer's instructions with the GeneJET PCR purification kit (Thermo Scientific) with products from the same sample combined. NGS libraries were prepared with the Nextera XT DNA Sample Preparation and the Nextera XT Index Kit for 24 Indexes (Illumina), and sequencing was performed with MiSeq (Illumina) using the MiSeq Reagent Kit v2-150.

Filling the gaps in the genome

Gaps present in the sequence after NGS were filled with Sanger sequencing. These regions were amplified with PCR using 12.5 µl of 2x Phusion Flash PCR Master Mix (Thermo Scientific), 2.5 µl of 10 µM primers and 2.5 µl of water. Primers AMDV754 and AMDV1005 amplified the region 754–1027, AMDV1256 and AMDV1526 the region 1256–1549 and AMDV2439 and AMDV2633 the region 2439–2656 of AMDV-G. The cycling conditions were as follows: initial denaturation at 98 °C for 10 s, followed by 30 cycles of denaturation at 98 °C for 1 s, annealing at 56 °C for 5 s and extension at 72 °C for 7 s, and final extension at 72 °C for 1 min. PCR products were run on a 1.5 % agarose gel, purified as described earlier, and sequenced with Sanger sequencing using PCR primers. If the gap-filling PCR was unsuccessful, Sanger sequencing was performed with the original complete genome PCR product as a template.

Other sequencing methods

If complete genome sequencing from samples from farms with a long-term infection was unsuccessful, two smaller parts were sequenced using two pan-AMDV PCRs. PCR amplifying the region 1662–2302 was described by Knuuttila *et al.* [19]. The PCR reaction amplifying the region 578–951 contained 12.5 µl of Fermentas SYBR MasterMix (Thermo Scientific), 6 µl of H_2O, 0.75 µl 10 of µM primers D_AMDV F 7 h PN1 and D_AMDV R 7 h PN2 [47] and 5 µl of template DNA. The real-time PCR was performed with Stratagene Mx3005P (Agilent Technologies) and the cycling conditions were as follows: 10 min at 95 °C, followed by 40 cycles of 30 s at 95 °C, 1 min at 55 °C and 1 min at 72 °C, and melting curve analysis with 1 min at 95 °C, 30 s at 55 °C and 30 s at 95 °C. PCR products were sequenced with Sanger sequencing as described earlier.

Co-infections

To study co-infections, nt 578–951 from selected samples was amplified with a PCR reaction that contained 5 µl of 10x Taq Buffer with $(NH_4)_2SO_4$, 1 µl of 10 mM dNTP, 1 µl of 10 µM D_AMDV F 7 h PN1, 1 µl of 10 µM D_AMDV F 7 h PN1, 4 µl 25 mM $MgCl_2$, 1.25 U of Taq DNA polymerase (Thermo Scientific), 5 µl of template DNA, and water added to 50 µl. The cycling conditions were as follows: initial denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 95 °C for 30 s, annealing at 47 °C for 30 s and extension at 72 °C for 25 s, and final extension at 72 °C for 7 min. PCR products were purified as earlier, cloned with PGEM-T Vector system I (Promega), plasmids from ten clones from each sample were isolated with GeneJET Plasmid Miniprep kit (Thermo Scientific) and sequenced with Sanger sequencing.

Data analysis

The data were mapped against the reference sequence (AMDV-G) with UGENE [48] using Bowtie2 [49] and checked using BWA-MEM 0.7.17 [50]. Full genomes were aligned in MEGA6 [51] using ClustalW. *P*-distances with pairwise deletion and the best evolutionary models were defined with MEGA6 [51]. Similarity scans were performed with Simplot [52], grouping scans (500 nt window and 50 nt steps) with SSE [53], and Bayesian phylogeny with Beast 1.8.2 [54], Tracer v1.6 [55] and FigTree v1.4.2 [56]. The combinations of three clock models (strict clock, lognormal relaxed clock and exponential clock) and three demographic models (constant size, exponential growth and Bayesian skyline model) were assessed using path sampling (PS) and stepping-stone sampling (SS) methods and calculating Bayes factors between the different model combinations [57, 58]. The results are given in Table S3. Lognormal relaxed clock and the Bayesian skyline model showed the highest marginal likelihoods with BF>20 when compared to other model comparisons. For Bayesian trees with a molecular clock from regions 1–700, 1165–1864 and 2446–3443 (Fig. 1a–c) and 2417–3413 (Fig. 2c), analysis was performed using general time-reversible model (GTR)+G, GTR+G, Hasegawa−Kishino−Yano model (HKY)+G+I and HKY+G+I as evolutionary models, a lognormal relaxed clock as a lock model, Bayesian skyline as tree prior, and 50 000 000 (Fig. 1a-c) and 100 000 000 (Fig. 2c) as the chain length. For Bayesian trees without a molecular clock from the regions 578–951 and nt 1662–2302 (Fig. 2a–b), analysis was performed using GTR+G+I and GTR+G as evolutionary models and 20000000 as the chain length. Recombination was analysed with the programs RDP [59], GENECONV [60], BootScan [61], MaxChi [62], Chimaera [63], SiScan [64] and 3Seq [65] of RDP v.4.92 package [66] with the highest acceptable *P*-value of 0.05. In order to estimate the intra-host variation of the virus populations, the primer sequences and sequences with quality score <30 were removed using Trimmomatics [67]. The NGS data was assembled using the BWA-MEM algorithm [50] and the potential PCR duplicates were removed using SAMTools version 1.8 [68, 69]. The single nucleotide variants were called using LoFreq version 2 [70]. Statistical tests were performed with IBM SPSS statistics 24.

Funding information
This study was supported by the Helve Foundation, Orion-Farmos Research Foundation, and Finnish Veterinary Foundation grants.

Acknowledgements
We thank Marja Isomursu and Minna Nylund (Evira), Petri Timonen (Natural Resources Institute Finland) and Tiit Maran (Tallinn Zoological Gardens/Estonian University of Life Sciences) for taking part in the sample collection. We also thank Johanna Martikainen for excellent technical assistance.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Shahrabadi MS, Cho HJ, Marusyk RG. Characterization of the protein and nucleic acid of Aleutian disease virus. *J Virol* 1977;23:353–362.
2. Bloom ME, Race RE, Wolfinbarger JB. Characterization of Aleutian disease virus as a parvovirus . *J Virol* 1980;35:836–843.
3. Canuti M, Whitney HG, Lang AS. Amdoparvoviruses in small mammals: expanding our understanding of parvovirus diversity, distribution, and pathology. *Front Microbiol* 2015;6:1119.
4. Cotmore SF, Agbandje-McKenna M, Chiorini JA, Mukha DV, Pintel DJ *et al.* The family Parvoviridae. *Arch Virol* 2014;159:1239–1247.
5. Bloom ME, Alexandersen S, Perryman S, Lechner D, Wolfinbarger JB. Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus (ADV): sequence comparisons between a nonpathogenic and a pathogenic strain of ADV. *J Virol* 1988;62:2903–2915.
6. Bloom ME, Alexandersen S, Garon CF, Mori S, Wei W *et al.* Nucleotide sequence of the 5'-terminal palindrome of Aleutian mink disease parvovirus and construction of an infectious molecular clone. *J Virol* 1990;64:3551–3556.
7. Qiu J, Cheng F, Burger LR, Pintel D. The transcription profile of Aleutian mink disease virus in CRFK cells is generated by alternative processing of pre-mRNAs produced from a single promoter. *J Virol* 2006;80:654–662.
8. Qiu J, Cheng F, Pintel D. The abundant R2 mRNA generated by aleutian mink disease parvovirus is tricistronic, encoding NS2, VP1, and VP2. *J Virol* 2007;81:6993–7000.
9. Huang Q, Luo Y, Cheng F, Best SM, Bloom ME *et al.* Molecular characterization of the small nonstructural proteins of parvovirus Aleutian mink disease virus (AMDV) during infection. *Virology* 2014;452-453:23–31.
10. Bloom ME, Kanno H, Mori S, Wolfinbarger JB. Aleutian mink disease: puzzles and paradigms *Infect. Agents Dis* 1994;3:279–301.
11. Aasted B. Aleutian disease of mink. Virology and immunology *Acta Pathol. Microbiol Immunol Scand Suppl* 1985;287:1–47.
12. Hussain I, Price GW, Farid AH. Inactivation of Aleutian mink disease virus through high temperature exposure in vitro and under field-based composting conditions. *Vet Microbiol* 2014;173:50–58.
13. Cho HJ. Purification and structure of aleutian disease virus. *Front Biol* 1976;44:159–174.
14. Hartsough GR, Gorham JR. Aleutian disease in mink. *Nat Fur News* 1956:10–11.
15. Gorham J, Henson J, Crawford T, Padgett G. The epizootiology of Aleutian disease. 1976:135–158.
16. Knuuttila A, Aronen P, Saarinen A, Vapalahti O. Development and evaluation of an enzyme-linked immunosorbent assay based on recombinant VP2 capsids for the detection of antibodies to Aleutian mink disease virus. *Clin Vaccine Immunol* 2009;16:1360–1365.
17. Knuuttila A, Aronen P, Eerola M, Gardner IA, Virtala AM *et al.* Validation of an automated ELISA system for detection of antibodies to Aleutian mink disease virus using blood samples collected in filter paper strips. *Virol J* 2014;11:141-422X-11-141.
18. Cho HJ, Ingram DG. Antigen and Antibody in Aleutian Disease in Mink. I. Precipitation Reaction by Agar-Gel Electrophoresis. *The Journal of Immunology* 1972;108:555–557.
19. Knuuttila A, Aaltonen K, Virtala AM, Henttonen H, Isomursu M *et al.* Aleutian mink disease virus in free-ranging mustelids in Finland - a cross-sectional epidemiological and phylogenetic study. *J Gen Virol* 2015;96:1423–1435.
20. Farid AH. Aleutian mink disease virus in furbearing mammals in Nova Scotia, Canada. *Acta Vet Scand* 2013;55:10-0147–10-0155.
21. Fournier-Chambrillon C, Aasted B, Perrot A, Pontier D, Sauvage F *et al.* Antibodies to Aleutian mink disease parvovirus

in free-ranging European mink (Mustela lutreola) and other small carnivores from southwestern France. *J Wildl Dis* 2004; 40:394–402.

22. **Murakami M, Matsuba C, Une Y, Nomura Y, Fujitani H.** Nucleotide sequence and polymerase chain reaction/restriction fragment length polymorphism analyses of Aleutian disease virus in ferrets in Japan. *J Vet Diagn Invest* 2001;13:337–340.
23. **Li L, Pesavento PA, Woods L, Clifford DL, Luff J *et al.*** Novel amdovirus in gray foxes. *Emerg Infect Dis* 2011;17:1876–1878.
24. **Shao XQ, Wen YJ, Hx B, Zhang XT, Yue ZG *et al.*** Novel amdoparvovirus infecting farmed raccoon dogs and arctic foxes. *Emerg Infect Dis* 2014;20:2085–2088.
25. **Bodewes R, Ruiz-Gonzalez A, Schapendonk CM, van den Brand JM, Osterhaus AD *et al.*** Viral metagenomic analysis of feces of wild small carnivores. *Virol J* 2014;11:89-422X-11-89.
26. **Canuti M, Doyle HE, P Britton A, Lang AS.** Full genetic characterization and epidemiology of a novel amdoparvovirus in striped skunk (Mephitis mephitis). *Emerg Microbes Infect* 2017;6:e30.
27. **Knuuttila A, Uzcategui N, Kankkonen J, Vapalahti O, Kinnunen P.** Molecular epidemiology of Aleutian mink disease virus in Finland. *Vet Microbiol* 2009;133:229–238.
28. **Canuti M, O'Leary KE, Hunter BD, Spearman G, Ojkic D *et al.*** Driving forces behind the evolution of the Aleutian mink disease parvovirus in the context of intensive farming. *Virus Evol* 2016;2: vew004.
29. **Hagberg EE, Krarup A, Fahnøe U, Larsen LE, Dam-Tuxen R *et al.*** A fast and robust method for whole genome sequencing of the Aleutian Mink Disease Virus (AMDV) genome. *J Virol Methods* 2016;234:43–51.
30. **Hagberg EE, Pedersen AG, Larsen LE, Krarup A.** Evolutionary analysis of whole-genome sequences confirms inter-farm transmission of Aleutian mink disease virus. *J Gen Virol* 2017;98:1360–1371.
31. **Li Y, Huang J, Jia Y, Du Y, Jiang P *et al.*** Genetic characterization of Aleutian mink disease viruses isolated in China. *Virus Genes* 2012;45:24–30.
32. **van Dawen S, Kaaden OR, Roth S.** Propagation of Aleutian disease parvovirus in cell line CCC clone 81. *Arch Virol* 1983;77:39–50.
33. **Xi J, Wang J, Yu Y, Zhang X, Mao Y *et al.*** Genetic characterization of the complete genome of an Aleutian mink disease virus isolated in north China. *Virus Genes* 2016;52:463–473.
34. **Ryt-Hansen P, Hagberg EE, Chriél M, Struve T, Pedersen AG *et al.*** Global phylogenetic analysis of contemporary aleutian mink disease viruses (AMDVs). *Virol J* 2017;14:231-017-0898-y.
35. **Suomen Turkiseläinten Kasvattajain Liitto (STKL).** Suomen Turkiseläinten Kasvattajain Liitto (STKL). https://profur.fi/.
36. **Nes N, Einarsson E, Lohi O, Jørgensen G.** Beautiful fur animals – and their colour genetics1988.
37. **Nummi P.** Suomeen istutetut riistaeläimet. *Helsingin yliopisto, Maatalous- ja Metsäeläintieteen Laitos* 1988:13–15.
38. **Jakubczak A, Kowalczyk M, Kostro K, Jezewska-Witkowska G.** Comparative molecular analysis of strains of the Aleutian Disease Virus isolated from farmed and wild mink. *Ann Agric Environ Med* 2017;24:366–371.
39. **Hoelzer K, Shackelton LA, Holmes EC, Parrish CR.** Within-host genetic diversity of endemic and emerging parvoviruses of dogs and cats. *J Virol* 2008;82:11096–11105.
40. **Shackelton LA, Hoelzer K, Parrish CR, Holmes EC.** Comparative analysis reveals frequent recombination in the parvoviruses. *J Gen Virol* 2007;88:3294–3301.
41. **Ohshima T, Mochizuki M.** Evidence for recombination between feline panleukopenia virus and canine parvovirus type 2. *J Vet Med Sci* 2009;71:403–408.
42. **Wang J, Cheng S, Yi L, Cheng Y, Yang S *et al.*** Evidence for natural recombination between mink enteritis virus and canine parvovirus. *Virol J* 2012;9:252–259.
43. **Duffy S, Shackelton LA, Holmes EC.** Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet* 2008;9:267–276.
44. **Hoelzer K, Shackelton LA, Parrish CR, Holmes EC.** Phylogenetic analysis reveals the emergence, evolution and dispersal of carnivore parvoviruses. *J Gen Virol* 2008;89:2280–2289.
45. **Hoelzer K, Parrish CR.** The emergence of parvoviruses of carnivores. *Vet Res* 2010;41:39.
46. **Shackelton LA, Parrish CR, Truyen U, Holmes EC.** High rate of viral evolution associated with the emergence of carnivore parvovirus*Proc. Natl Acad Sci USA* 2005;102:379–384.
47. **Jensen TH, Christensen LS, Chriél M, Uttenthal A, Hammer AS.** Implementation and validation of a sensitive PCR detection method in the eradication campaign against Aleutian mink disease virus. *J Virol Methods* 2011;171:81–85.
48. **Okonechnikov K, Golosova O, Fursov M.** Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 2012;28:1166–1167.
49. **Langmead B, Salzberg SL.** Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359.
50. **Li H.** Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM2013 arXiv:1303.3997.
51. **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S.** MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 2013;30:2725–2729.
52. **Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS *et al.*** Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J Virol* 1999;73:152–160.
53. **Simmonds P.** SSE: a nucleotide and amino acid sequence analysis platform. *BMC Res Notes* 2012;5:50-0500-5-50.
54. **Drummond AJ, Suchard MA, Xie D, Rambaut A.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 2012;29: 1969–1973.
55. **Rambaut A, Suchard M, Xie D, Drummond A.** Tracer v1.6. 2014;6 http://tree.bio.ed.ac.uk/software/tracer/.
56. **Rambaut A.** FigTree. http://tree.bio.ed.ac.uk/software/figtree/.
57. **Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA *et al.*** Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty *Mol. Biol Evol* 2012;29:2157–2167.
58. **Baele G, Wl L, Drummond AJ, Suchard MA, Lemey P.** Accurate model selection of relaxed molecular clocks in bayesian phylogenetics. *Mol Biol Evol* 2013;30:239–243.
59. **Martin D, Rybicki E.** RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 2000;16:562–563.
60. **Padidam M, Sawyer S, Fauquet CM.** Possible emergence of new geminiviruses by frequent recombination. *Virology* 1999;265:218–225.
61. **Salminen MO, Carr JK, Burke DS, McCutchan FE.** Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res Hum Retroviruses* 1995;11:1423–1425.
62. **Smith JM.** Analyzing the mosaic structure of genes. *J Mol Evol* 1992;34:126–129.
63. **Posada D, Crandall KA.** Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proc Natl Acad Sci USA* 2001;98:13757–13762.
64. **Gibbs MJ, Armstrong JS, Gibbs AJ.** Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 2000;16:573–582.
65. **Boni MF, Posada D, Feldman MW.** An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 2007;176:1035–1047.
66. **Martin DP, Murrell B, Golden M, Khoosal A, Muhire B.** RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol* 2015;1:vev003.

67. **Bolger AM, Lohse M, Usadel B.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
68. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J *et al.*** The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
69. **Li H.** A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987–2993.
70. **Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH *et al.*** LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res* 2012;40:11189–11201.