

ARTICLE

Open Access

Evolutionary evidence for multi-host transmission of cetacean morbillivirus

Wendy K. Jo¹, Jochen Kruppa², Andre Habierski¹, Marco van de Bildt³, Sandro Mazzariol⁴, Giovanni Di Guardo⁵, Ursula Siebert⁶, Thijs Kuiken³, Klaus Jung², Albert Osterhaus¹ and Martin Ludlow¹

Abstract

Cetacean morbillivirus (CeMV) has emerged as the pathogen that poses the greatest risk of triggering epizootics in cetacean populations worldwide, and has a high propensity for interspecies transmission, including sporadic infection of seals. In this study, we investigated the evolutionary history of CeMV by deep sequencing wild-type viruses from tissue samples representing cetacean species with different spatiotemporal origins. Bayesian phylogeographic analysis generated an estimated evolutionary rate of 2.34×10^{-4} nucleotide substitutions/site/year and showed that CeMV evolutionary dynamics are neither host-restricted nor location-restricted. Moreover, the dolphin morbillivirus strain of CeMV has undergone purifying selection without evidence of species-specific mutations. Cell-to-cell fusion and growth kinetics assays demonstrated that CeMV can use both dolphin and seal CD150 as a cellular receptor. Thus, it appears that CeMV can readily spread among multiple cetacean populations and may pose an additional spillover risk to seals.

Introduction

Cetacean morbillivirus (CeMV) is known to infect a wide range of cetacean species from both Odontoceti (toothed whales) and Mysticeti (baleen whales) subgroups¹, including threatened and endangered species, such as the fin whale (*Balaenoptera physalus*) and sperm whale (*Physeter macrocephalus*). This emerging morbillivirus causes variable degrees of morbidity and mortality and thus may have a major impact on the demography and conservation of cetacean populations. Five distinct strains of CeMV have now been documented^{2–6} since the identification of porpoise morbillivirus (PMV) in stranded harbor porpoises (*Phocoena phocoena*) in Ireland and the Netherlands in 1988 and 1990, respectively^{2,7}. Dolphin morbillivirus (DMV) is the best characterized strain, having caused several epizootics in addition to

isolated cases that were retrospectively identified using molecular and pathological diagnostic protocols. The first documented CeMV outbreak killed thousands of striped dolphins (*Stenella coeruleoalba*) in the Mediterranean Sea from 1990 to 1992³. However, CeMV was later identified as the etiological agent responsible for an earlier epizootic in bottlenose dolphins (*Tursiops truncatus*) along the northwestern Atlantic coast in 1987–1988⁸. The trans-oceanic occurrence of CeMV infections continued with later outbreaks in the Gulf of Mexico in 1993–1994⁹, the Mediterranean Sea in 2006–2008¹⁰, the coast of South Australia in 2013¹¹, along the eastern USA seaboard in 2013–2015¹², and in southeastern Brazil in 2017¹³, causing mass die-offs of striped dolphins, bottlenose dolphins, Indo-Pacific bottlenose dolphins (*T. aduncus*), short-beaked common dolphins (*Delphinus delphis*), long-finned pilot whales (*Globicephala melas*), and Guiana dolphins (*Sotalia guianensis*). Additionally, DMV infections have been diagnosed in cases of fin whale and sperm whale strandings, possibly playing a role in their demise^{14–16}. Pilot whale morbillivirus (PWMV) has been sporadically reported to have infected pilot whales

Correspondence: Martin Ludlow (martin.ludlow@tiho-hannover.de)

¹Research Center for Emerging Infections and Zoonoses, University of Veterinary Medicine Hannover, 30559 Hannover, Germany

²Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover, 30559 Hannover, Germany

Full list of author information is available at the end of the article.

© The Author(s) 2018



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

(*G. melas* and *G. macrorhynchus*) in the northeastern and northwestern Atlantic^{4,17}. In recent years, two novel CeMV strains have been identified, namely beaked whale morbillivirus (BWMV) in the central Pacific⁵ and a more divergent strain, which is called Guiana dolphin morbillivirus (GDMV) in this study for simplification purposes, that was detected in Guiana and Indo-Pacific bottlenose dolphins found stranded along the Brazilian⁶ and West Australian coasts¹⁸, respectively.

CeMV is an enveloped negative-sense single-stranded RNA virus in the genus Morbillivirus within the family *Paramyxoviridae* along with measles virus (MV, the prototypic species), the recently eradicated rinderpest virus (RPV), peste des petits ruminants virus (PPRV), canine distemper virus (CDV), phocine distemper virus (PDV), and feline morbillivirus (FeMV). A characteristic of animal morbilliviruses (i.e., RPV, PPRV, CDV, and CeMV) is a propensity for interspecies transmission. For instance, RPV infected multiple wild and domestic species of ungulates, whereas CDV can infect multiple carnivores and spread from carnivores to non-human primates, rodents, and artiodactyls¹⁹. In addition to multiple dolphin and whale species¹, CeMV has also crossed the species barrier to infect marine mammals belonging to a different order. In 1997, DMV was associated with an epizootic that halved the already endangered population of Mediterranean monk seals (*Monachus monachus*) in Mauritania²⁰, although it was later proposed that algal intoxication directly caused these deaths²¹. Consequently, the direct cause of this epizootic remains a matter of debate. Additional documented cases of CeMV infection of seal species include a monk seal in Greece²² and a captive harbor seal (*Phoca vitulina*) in Italy²³ that were reported to be PMV-infected and DMV-infected, respectively. However, the host range of CeMV and potential barriers to cross-species infections remain to be further determined.

In common with other morbilliviruses, CeMV displays a tropism for lymphoid and epithelial cells, and induces typical lesions including bronchointerstitial pneumonia and lymphocytic depletion in lymphoid organs^{3,6,14,16,17,24}. In addition, CeMV is neurotropic as non-suppurative encephalitis is commonly observed in infected cetaceans. Cell entry and spread of morbilliviruses is mediated by the viral fusion (F) and hemagglutinin (H) glycoproteins. Signaling lymphocyte activation molecule (SLAM/CD150) and poliovirus receptor-like 4 (PVRL-4/Nectin-4) have been previously identified as universal morbillivirus cellular receptors^{25,26}. SLAM has been found in immune cells and appears to be the primary receptor for virus entry and spread to other tissues, whereas Nectin-4 is the primary receptor of epithelial cells and is important for virus exit and transmission. The route of transmission is usually through aerosol droplets as well as direct or

indirect contact. In the case of cetaceans, virus dissemination in the population is speculated to occur via direct contact or aerogenic means¹. It is assumed that morbillivirus excretion and transmission occur predominantly from virus-infected epithelial cells at mucosal surfaces. High levels of CeMV-positive epithelial cells have previously been detected in the respiratory and urogenital tracts of infected dolphins and whales^{9,11,17}. Vertical transmission has also been discussed as another possible route of virus infection¹⁴.

Although a large number of case studies investigating CeMV epidemiology via serology and partial sequences have been reported¹, no studies on CeMV evolution in relation to trans-oceanic viral spread based on whole genome sequences are available. In the present study, we have deep-sequenced European CeMV strains from a collection of historical and recent CeMV-infected tissue samples and virus isolates to investigate the evolutionary history of CeMV and to identify possible molecular signatures of host adaptation arising from interspecies virus transmission.

Results

Genome analyses of wild-type and laboratory-adapted CeMVs

A collection of historical and recent CeMV-infected tissue samples from Europe was investigated along with cell culture-passaged CeMV strains (Table 1). Samples were first screened by real-time PCR (RT-PCR) to assess the viral load with Ct values ranging from 12 to 32. Complete CeMV genome sequences from tissue samples and virus isolates were generated by a combination of next generation sequencing (NGS) (Fig. 1a) and rapid amplification of cDNA ends (RACE; Fig. 1b and Supplementary Fig. S1). In addition, gap sequencing was performed for CeMV strains DMV-Bph, DMV-NL/LA/11.1, and PMV-53 due to low coverage in a number of regions. Host reference alignments of evaluated samples permitted confirmation of the identity of the different host species (dolphin or whale). An analysis of genome termini showed the presence of two nt changes at positions 5 (C → U) and 12 (G → A) of the leader sequence in all CeMV sequences compared to the lab-adapted 1990 DMV strain (GenBank Accession No. AJ608288) that was used as a reference genome in this study.

The genetic variation within individual CeMV lineages was low, with a pairwise identity distance of 98.3–99.9% between DMV strains and 99–7–99.8% between PMV strains. In contrast, an 86.3–86.8% pairwise identity distance was observed between DMV and PMV strains. A Bayesian estimation of the substitution rate of CeMV was 2.34×10^{-4} nucleotide substitutions/site/year (subs/site/year) with 95% of the highest posterior density intervals at $1.86\text{--}2.83 \times 10^{-4}$ subs/site/year. The substitution

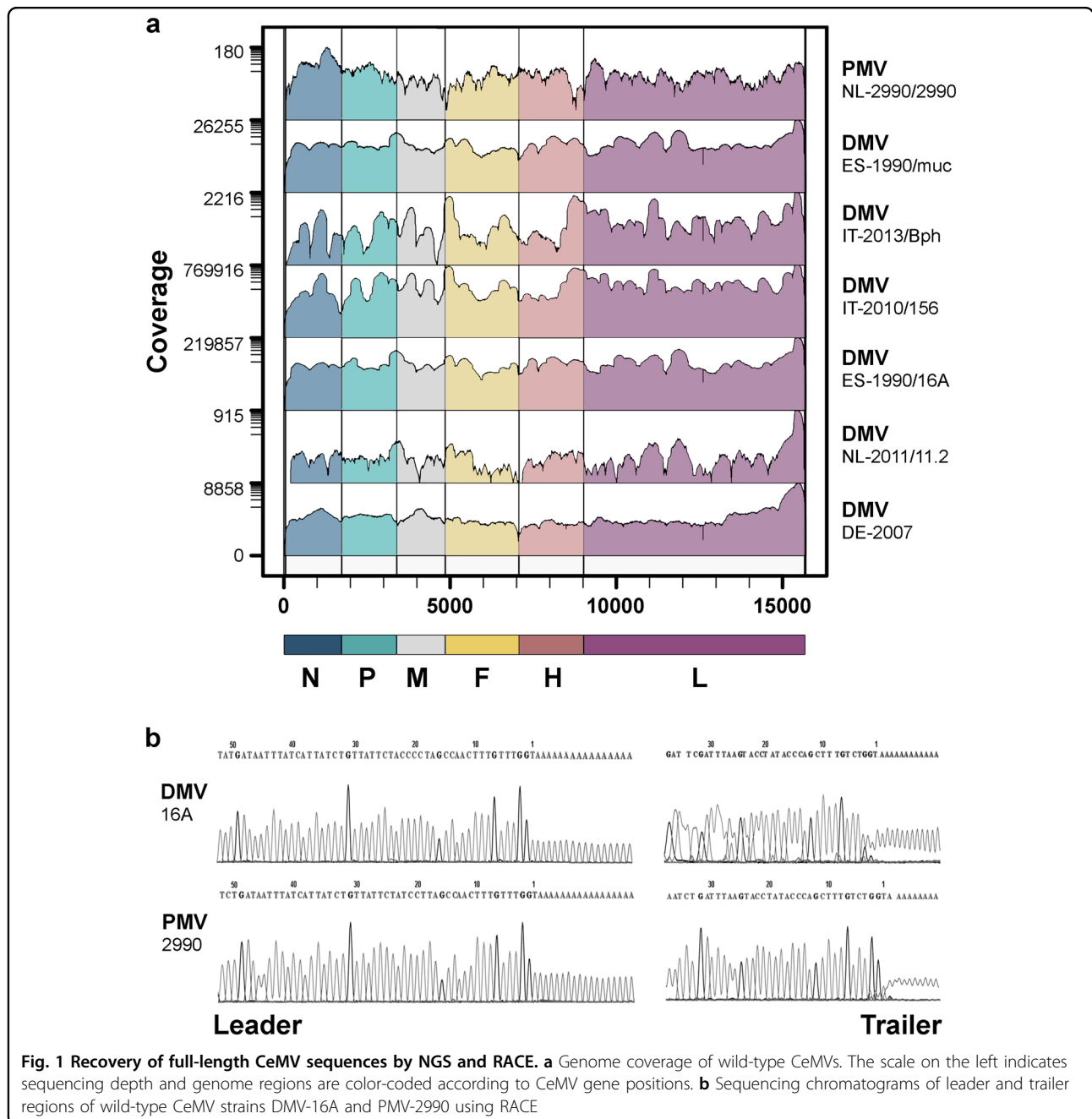
Table 1 Cetacean morbillivirus detection in cetacean samples from the North Sea and the Mediterranean Sea

CeMV strain	Variant	Stranding year	Location	Host	Sample material	RT-PCR (Ct)	NGS coverage	GenBank accession no.	References
PMV	Ulster/88	1988	Northern Ireland, U.K.	Harbor porpoise	Vero cells	NI	5149X	MH430942	2
DMV	16A	1990	Spain	Striped dolphin	Lung	12	6023X	MH430934	28
DMV	16A-cc	1990	Spain	Striped dolphin	Vero cells	13	27X	MH430932	This study
DMV	16A-cc-vdol	1990	Spain	Striped dolphin	Vero-dolSLAM cells	13	1188X	MH430933	This study
DMV	muc	1990	Spain	Striped dolphin	Lung	13	1248X	MH430935	28
DMV	muc-cc	1990	Spain	Striped dolphin	Vero cells	13	20X	MH430936	This study
PMV	53	1990	Netherlands	Harbor porpoise	Vero cells	NI	9120X	MH430943	7
PMV	2990	1990	Netherlands	Harbor porpoise	Brain	32	26X	MH430945	7
PMV	2990-cc	1990	Netherlands	Harbor porpoise	Vero cells	NI	11,933X	MH430944	This study
DMV	DE/2007	2007	Germany	White-beaked dolphin	Brain	19	283X	MH430940	48
DMV	156	2010	Italy	Striped dolphin	Lung	19	50,169X	MH430937	This study
DMV	DMV/LA/NL/11.2	2011	Netherlands	White-beaked dolphin	Lung	23	21X	MH430941	24
DMV	DMV_Bph	2013	Italy	Fin whale	Brain	32	144X	MH430938	14, 49
DMV	DK/16 ^a	2016	Denmark	Fin whale	Lung	27	NI	MH430939	16

Ct value ≤ 40 = positive; Ct value > 40 = negative
 RT-PCR reverse transcriptase PCR, Ct value threshold cycle, NGS next-generation sequencing, NI not investigated
^aGenome obtained by Sanger sequencing

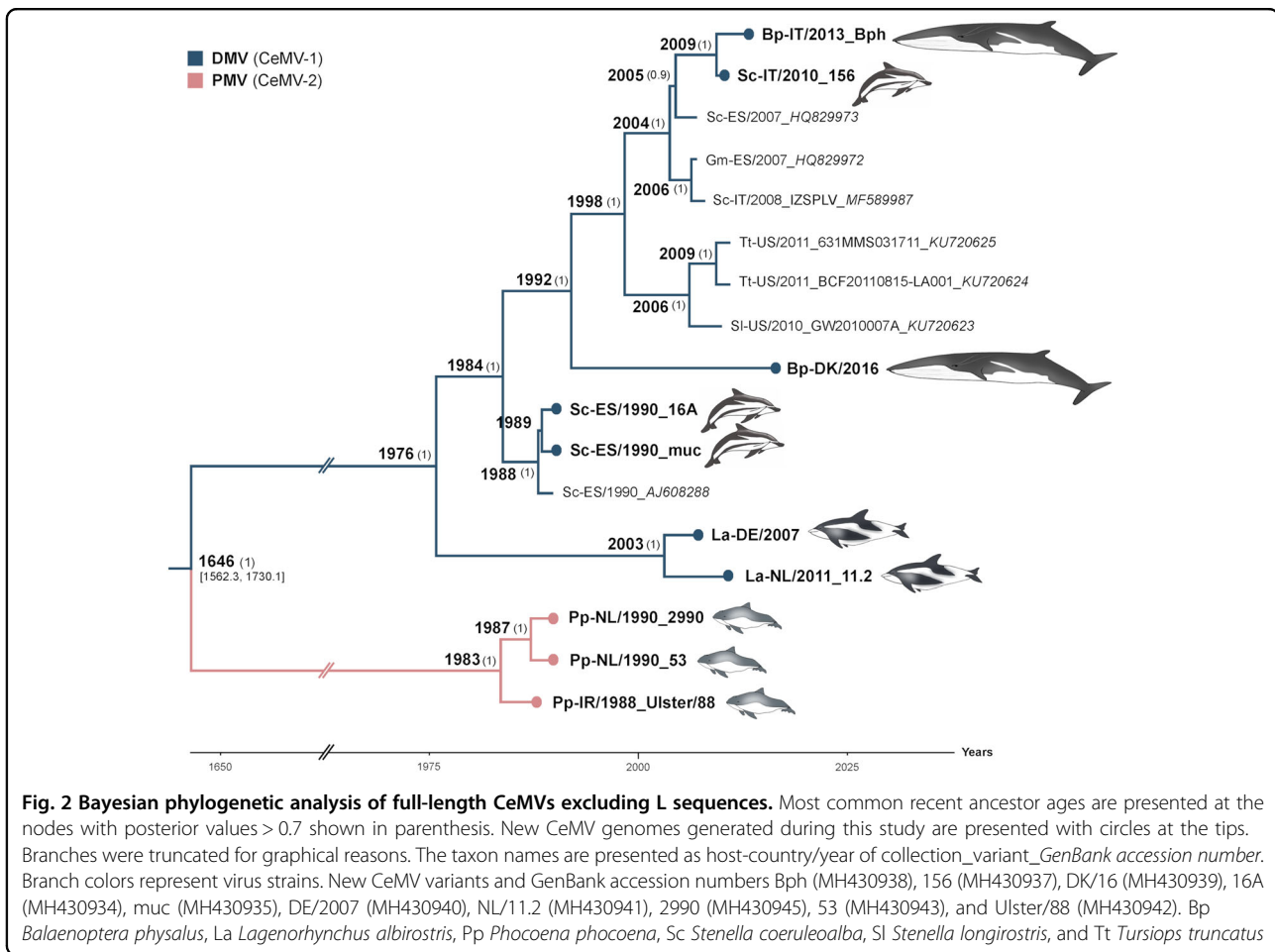
rates of individual coding sequences for N, P, M, F, H, and L resulted in a similar range (Supplementary Fig. S2). However, the non-coding regions were estimated to have a higher mutation rate of 5.58×10^{-4} subs/site/year, as these regions are usually less often subjected to selective pressures in comparison to coding regions. We excluded L gene sequences from subsequent Bayesian analyses as these were absent from the partial genome sequences of two published DMVs from the 2006–2008 outbreak²⁷. Moreover, the L gene sequence of the reference variant was excluded from the analyses due to the presence of many unique mutations, which may be artefactual, in comparison to all other DMV strains, including two new L gene sequences derived from dolphin tissues obtained from the same epizootic in 1990 (GenBank Accession Nos. MH430934–35). Non-synonymous substitutions in this variant accounted for 41% of the total amino acid (aa) changes in L (Supplementary Fig. S3). The tree topology was not changed when these sequences were taken out for analyses of complete DMV genomes (Supplementary Fig. S4). Phylogenetic analyses based on Bayesian inference suggested that DMV and PMV shared a most recent common ancestor (MRCA) approximately 400 years ago (Fig. 2). Moreover, the DMV strain responsible for the mass die-off in the Mediterranean Sea in 1990–1992 had a more basal position relative to other DMV variants, which emerged in subsequent years in the Mediterranean Sea (isolate_GenBank Accession Nos. HQ829972, HQ829973, IZSPLV_MF589987, 156_MH430937, Bph_MH430938), North Sea (isolate_GenBank Accession No. DK/16_MH430939), and the Gulf of Mexico (GenBank accession nos. KU720623–25). A full-length DMV sequence recovered from a fin whale stranded on the Danish coast in 2016 was also located within this DMV clade. However, additional North Sea variants from white-beaked dolphins (GenBank accession nos. DE/2007_MH430940 and NL/11.2_MH430941) were exceptions, as they shared a common ancestor with this Mediterranean DMV variant in 1976 (Fig. 2).

Phylogenetic analysis of individual genes displayed different tree topologies (Supplementary Fig. S5), of which the F and H gene phylogenies best resembled the tree derived from full-length CeMV sequences. Not surprisingly, this analysis is indicative of differing selective pressures in shaping the evolution of CeMV, in which the F and H genes apparently have key roles. Nevertheless, a distinction between different strains can still be made using more conserved viral genes. Therefore, we extended our analysis of global CeMV distribution patterns by examining a much larger data set of available partial CeMV P gene sequences. Phylogeographic analysis of 38 partial P gene sequences, which were previously deposited in GenBank or derived from this study, demonstrated that CeMV strains are neither host-specific nor location-



specific (Fig. 3a), as also suggested by the full-length sequence analyses. The multi-host transmission of CeMV among cetacean species is best represented by DMV and BWMV, which have been identified in multiple odontocete and mysticete species^{3-5,14,15}. The trans-oceanic spread of DMV is also readily apparent in closely related strains detected in the Gulf of Mexico, the Mediterranean Sea (Figs. 2 and 3, sequences 8–10 and 1–5), the east and west coasts of the North Atlantic Ocean and the North Sea (Fig. 3, sequences 11–14). Different virus dispersal routes through the Atlantic Ocean and from the

Atlantic to the Indian Oceans are proposed (Fig. 3b). Furthermore, each CeMV strain formed a distinctive clade grouping the sequences of DMV (CeMV-1), PMV (CeMV-2), BWMV (CeMV-3), PWMV (CeMV-4), and GDMV (CeMV-5; Fig. 3a). This last strain was the most basal and divergent, sharing an MRCA with other CeMV strains a few hundred years ago. The MRCA shared by DMV and PMV was found to have different dates upon comparison of either complete or partial P gene sequences (Figs. 2 and 3a). While full-length genome sequence analysis would be expected to be more accurate, only two



strains (DMV and PMV) could be included. In contrast, analysis of partial P gene sequences included sequences from all five known CeMV strains. However, given the differences in the evolutionary rate of the P gene compared with full-genome sequences, an analysis based on partial sequence may be less representative.

Identification of molecular signatures of CeMV adaptation associated with in vitro replication

Given that historical morbillivirus sequences, including CeMV, were generated using virus isolates, we have also examined the effects of virus isolation and passaging on the resulting consensus genome sequences. Complete genome sequences from DMV-16A and DMV-muc, which were first isolated and passaged in Vero cells 25 years ago²⁸, were compared to consensus sequences of the same viruses generated directly from original stored tissue samples. In addition, we generated Vero cells expressing SLAM derived from a Pacific white-sided dolphin (*Lagenorhynchus obliquidens*; Vero-dolSLAM) to investigate if expression of a natural CeMV cellular receptor better maintains the authentic consensus virus sequence. These cells were used to re-isolate DMV-16A

from a stored tissue sample collected during the CeMV epizootic in 1990. A single large syncytium was observed 2–3 days following infection, and the virus was subsequently passaged four times before NGS was used to generate a complete consensus viral sequence.

Sequence analysis of wild-type viruses present in tissue samples and viruses passaged in Vero cells and/or Vero-dolSLAM cells revealed minimal genetic variation (Table 2). The isolate DMV-16A grown in Vero-dolSLAM cells showed less variation from the wild-type consensus sequence than the same variant grown in Vero cells. Mutations S459P in P and I873T in L emerged in both isolates following in vitro passage. However, due to the presence of these substitutions in other wild-type DMV sequences (Fig. 4), we used the raw NGS data to calculate the frequencies of these variants in the virus population. Whereas the Vero cell isolates homogeneously expressed 459P in the P gene and 873T in the L gene, these positions were heterogeneous in the wild-type tissue-derived DMV-16A sequence (P-459: S = 53%, P = 46%; L-873: I = 52%, T = 47%), which was less heterogeneous in the DMV-16A isolated from Vero-dolSLAM (P-459: S = 1%, P = 99%; L-873: I = 4%, T = 96%) and showing different degrees

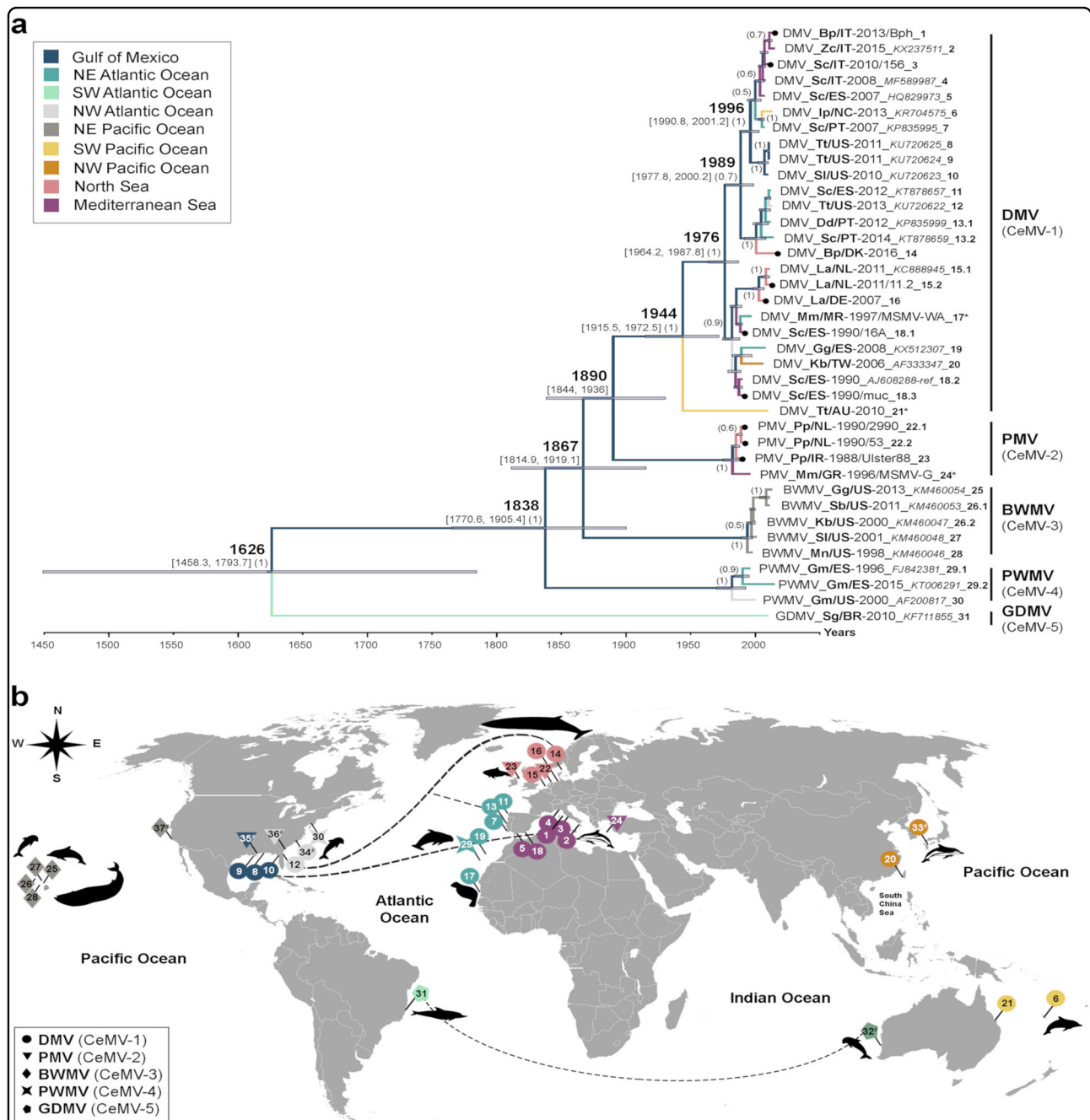
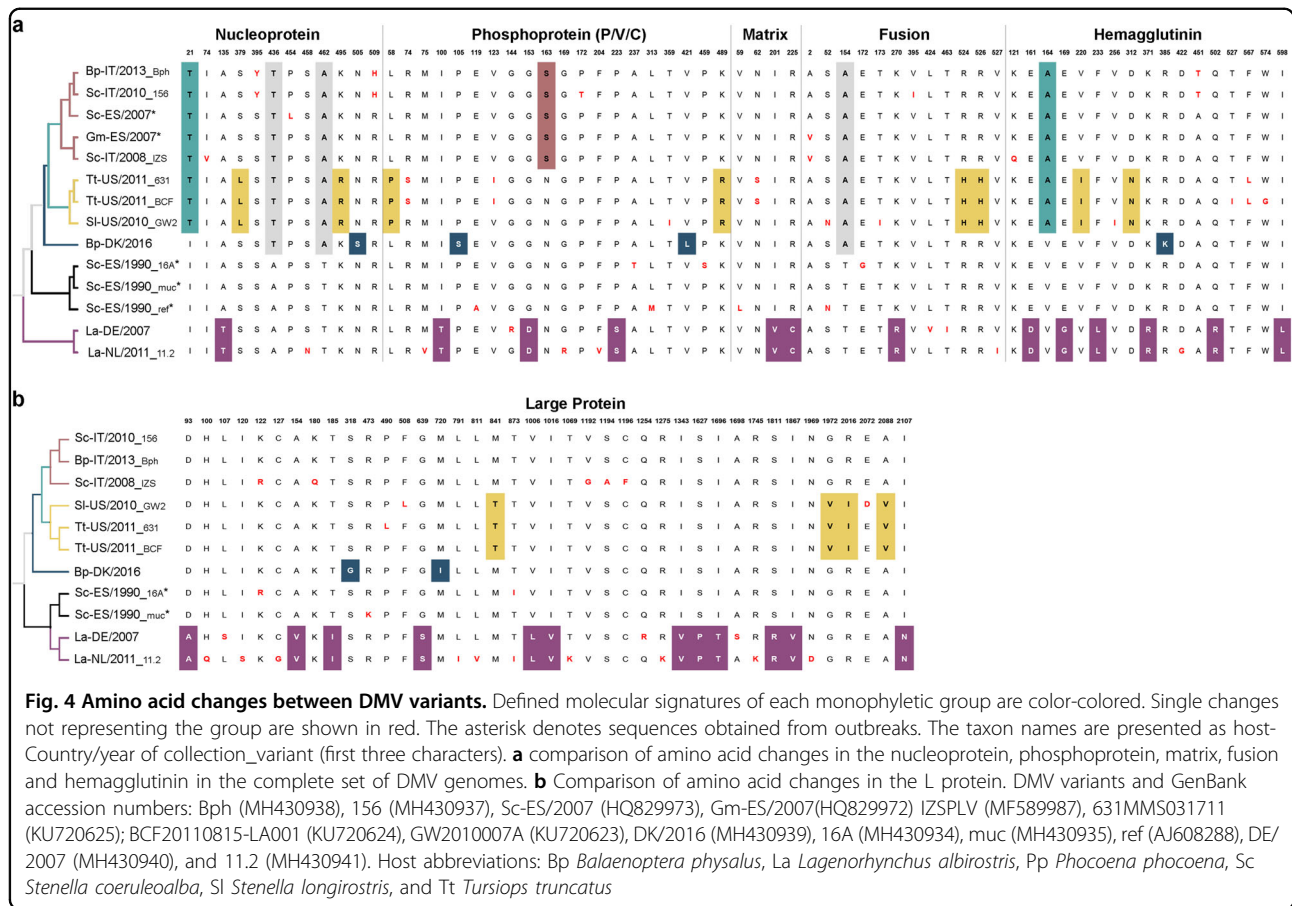


Fig. 3 Phylogeography of CeMV. **a** Bayesian phylogenetic analysis of partial P genes (400 bp). Most common recent ancestor ages are presented at the nodes in cursive with 95% highest posterior density interval values in brackets and as gray horizontal bars. Posterior values > 0.5 are shown in parenthesis. CeMV genomes in this study are presented with black circles at the tips. Each branch is color-coded according to the ocean/sea in which the cetaceans were found. The taxon names are presented as virus_host/Country-year of collection/variant_GenBank accession number_ID for **b**. Sequences 17* and 24* were extracted from a published paper²². Sequence 21* was kindly provided by Dr. Stone and Jianning Wang. Host abbreviations: Bp *Balaenoptera physalus*, Dd *Delphinus delphis*, Gg *Grampus griseus*, Gm *Globicephala melas*, Ip *Indopacetus pacificus*, Kb *Kogia breviceps*, La *Lagenorhynchus albirostris*, Mm *Monachus monachus*, Mn *Megaptera novaeangliae*, Pp *Phocoena phocoena*, Sb *Steno bredanensis*, Sc *Stenella coeruleoalba*, Sg *Sotalia guianensis*, Sl *Stenella longirostris*, Tt *Tursiops truncatus*, and Zc *Ziphius cavirostris*. **b** A world map of CeMV migration using sequences from **a**. Locations of viruses in the map are relative and were obtained from the publications in which the sequences were published. The hashtag (#) indicates no sequence availability (No. 32: CeMV-5 from AU/2010 and No. 33: CeMV-1_JP/1999). Proposed virus migratory routes (dashed lines) were based on sequence clades from the phylogenetic tree in **a**

Table 2 Sequence variation between wild-type and isolated CeMV strains

Strain-variant	Gene	Position (ORF)	Codon	Wild-type	Vero	V. doISLAM	Mutation
DMV-16A	Leader	26	—	a	a/t	a	g.26A>T
	P	693	231	c	g	c	dS
	P	1375	459	Ser ¹	Pro	Pro	S459P
	M	310	104	Thr	Ala	Thr	T104A
	M	502	168	Asn	His	Asn	N168H
	F	1339	447	Arg	Gly	Arg	R447G
	L	2618	873	Ile ²	Thr	Thr	I873T
DMV-muc	Leader	11	—	g	a	—	g.11G>A
	M	523	175	Arg	Gly	—	R175G
	M	681	227	Lys	Asn	—	K227N
PMV-2990	M	265	89	Glu	Lys	—	E89K
	H	224	75	Asn	Asn	—	S75N
	L	406	126	Trp	Trp	—	G126W
	L	4125	1375	t	c	—	dS

dS Synonymous substitution



of heterogeneity in all other wild-type sequences for S/P at 459 (DMV-muc: S = 41%, P = 48%; DMV-156: S = 18%, T = 81%; DMV-Bph: S = 3% P = 97%; DMV-DE/2007: S = 1%, P = 98%; and DMV/LA/NL/11.2: S = 40%, P = 60%) in the P gene and isoleucine and threonine at 873 (DMV-muc: I = 49%, T = 51%; DMV-156: I = 29%, T = 71%; DMV-Bph: I = 27%, T = 73%; DMV-DE/2007: I = 2%, T = 95%; and DMV/LA/NL/11.2: I = 65%, T = 35%) in the L gene.

Upon comparison of whole genome sequences of *in vitro* isolates of DMV-16A, DMV-muc, and PMV-2990 with analogous sequences generated from original tissue samples, we did not detect common mutations that had arisen as a result of cell culture adaptation. Instead, we identified changes in the individual variants that have been previously detected in MV studies^{29–31}. These changes include the nucleotide (nt) substitution g.26A>T²⁹ and aa substitutions in the M protein, namely E89K (PMV-2990)³¹ and R175G (DMV-muc)³⁰. An analysis of the genome sequence of DMV-16A, which had been passaged in Vero cells, showed the presence of T at nt 26, although A was still present as a minor variant. Upon the inclusion of other morbilliviruses in an alignment of genome termini, wild-types had an A at nt 26, whereas vaccine strains had a T at this position (Supplementary Fig. S1).

The capacity of CeMV to cross-species barriers

Alignment of full-length CeMV sequences showed no apparent aa substitutions suggestive of adaptation to dolphin or whale species. Instead, patterns of non-synonymous substitutions (*dS*) in CeMV-1 characterizing each cluster were identified (Fig. 4). The aa sites 74, 95, and 454 in N; aa site 172 in P; aa site 245 in F; aa sites 121, 256, 422, 451, and 567 in H; and aa sites 122, 136, 399, 873, 1006, 1194, 1196, 1396, and 1698 in L were found to be statistically significant ($p < 0.05$) episodic-positive selection sites using MEME (mixed effects model of evolution)³². However, no pervasive positive selection site was identified using an FEL (fixed effects likelihood)³³. Instead, a greater number of negative selection sites were detected (Supplementary Table S1). Furthermore, the potential variability in SLAM and PVRL-4 binding sites based on MV-H interaction with marmoset SLAM (91% similarity with human SLAM)³⁴ and human PVRL-4³⁵ were also investigated (Supplementary Table S2). Based on the alignment, cetacean SLAM has approximately 32 and 7% aa differences from human SLAM and PVRL-4, respectively. We identified a number of SLAM interaction sites in the H protein that were conserved among all morbilliviruses, including CeMV, namely D501, D503, D526, S528, R529, P550, Y520, and Y539 (aa position is based on CeMV). For PVRL-4, the sites Y520 (also in SLAM), Y537, Y539, Y454, L460,

G461, and D501 (also in SLAM) were found to be conserved.

Previously documented cross-species infections of CeMV in monk and harbor seals^{20,22,23} prompted us to extend our analysis to examine the capacity of CeMV to use phocine SLAM as a cellular receptor. This was evaluated using a quantitative cell-to-cell fusion assay and multistep growth analyses of DMV in Vero cells, newly generated Vero-dolSLAM cells and Vero cells expressing phocine SLAM (Vero-phocaSLAM) (Fig. 5a). DMV F and H glycoproteins induced significantly higher levels of cell-to-cell fusion than the glycoproteins from two CDV strains (of dog and seal origin) in the evaluated cell lines (Fig. 5b), clearly indicating that DMV could use both dolphin and phocine SLAM as cellular receptors. The stringency of the assay was validated by the additional observation that DMV glycoproteins were unable to use human SLAM as a cellular receptor (data not shown). Moreover, when the growth curves were compared (Fig. 5c), DMV growth yielded high titers in both Vero cells expressing SLAM but not in Vero cells. However, DMV showed better replication in Vero-dolSLAM cells compared to Vero-phocaSLAM. Collectively, these results indicate that DMV can readily use the SLAM receptor of seals.

Discussion

Over the past thirty years, CeMV has emerged as a major pathogen in cetacean species worldwide^{3,9,11,13}. During this period, the known host range of CeMV has continually expanded to encompass multiple toothed^{3,15,17} and baleen^{5,14,16} whale species, including endangered populations such as the fin whale¹⁴. Moreover, instances of spillover into harbor and endangered monk seals have also been documented^{20,23}. However, the role of viral molecular adaptation in shaping the evolutionary pathways of CeMV interspecies transmission has remained largely unexplored. In this study, we have used NGS and RACE to determine the first complete CeMV genome sequences from original cetacean tissue samples, including the first complete genome sequence of PMV from a harbor porpoise. We found that CeMV requires minimal genetic variation to maintain multi-host transmission cycles between populations of odontocete and mysticete cetaceans. Further analyses also demonstrated that CeMV can effectively use both dolphin and phocine SLAM as cellular receptors.

Analysis of CeMV sequences enabled us to estimate an overall substitution rate of 2.34×10^{-4} subs/site/year, which is lower than what has been proposed for both MV ($6.5\text{--}8.7 \times 10^{-4}$ subs/site/year)³⁶ and CDV ($4.8\text{--}7.41 \times 10^{-4}$ subs/site/year)^{37,38} based on their H gene sequences. In future, additional full-length sequences of CeMV strains, especially BMWV, PWMV, and GDMV, will help

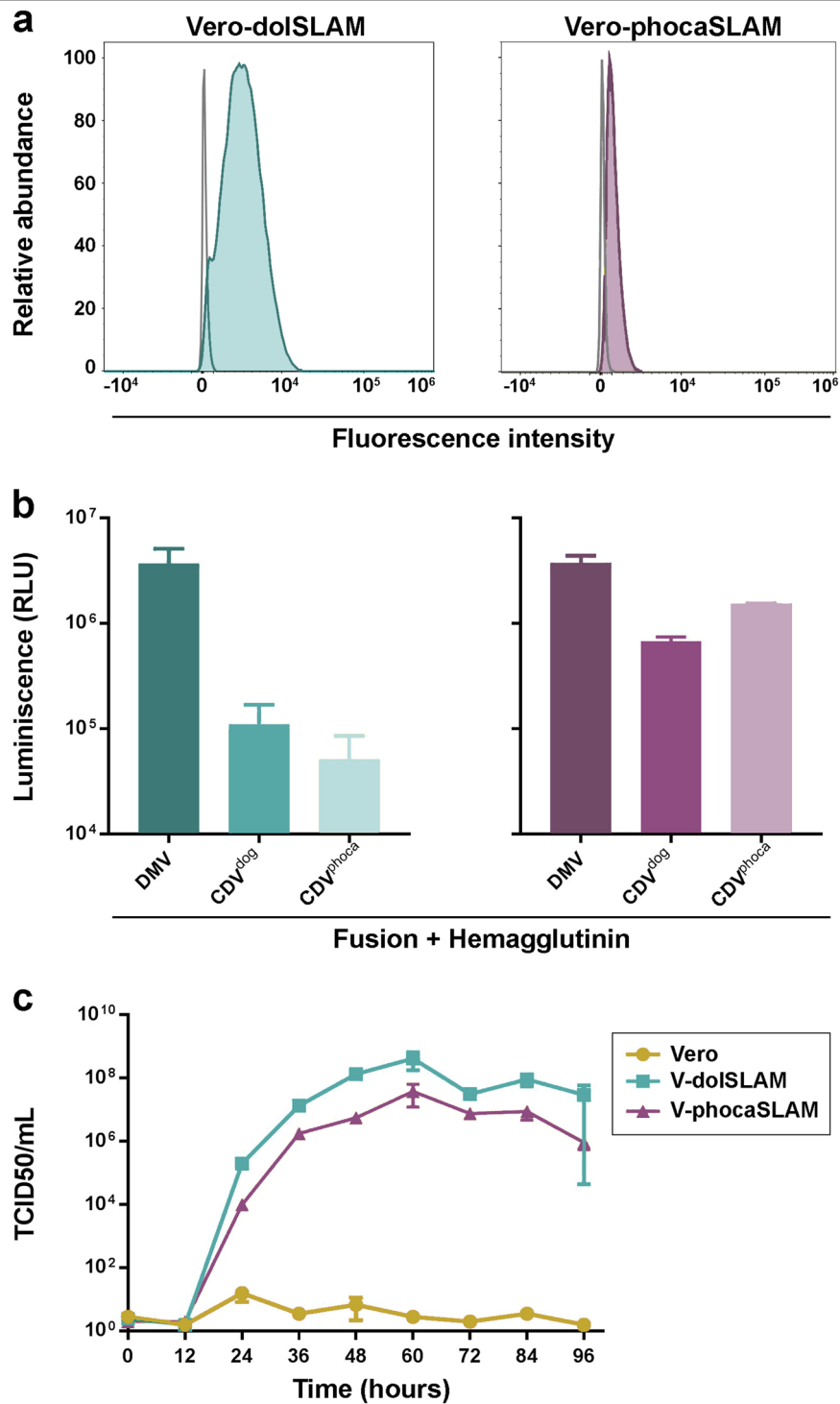


Fig. 5 CeMV use of the seal SLAM receptor. **a** Characterization of SLAM expression on Vero cells. Vero-dolSLAM cells (teal colored), Vero-phocaSLAM cells (purple colored) and untransfected Vero cells (gray line) were stained with a HA-tag antibody and analyzed by flow cytometry. **b** Cell-to-cell fusion activity of DMV, CDV-dog/2016, and CDV-phoca/1988 glycoproteins in Vero-dolSLAM cells (left) and Vero-phocaSLAM cells (right) as measured by β -galactosidase activity. Relative luminiscence unit (RLU) values on the Y-axis were calculated by subtracting the RLU values obtained from Vero-dolSLAM or Vero-phocaSLAM cells with values generated from Vero cells (control). Experiments were performed in quadruplicates ($n = 3$). **c** Multistep growth curve analyses of DMV-16A inoculated at an moi of 0.001 in Vero cells (control), Vero-dolSLAM and Vero-phocaSLAM cells. Virus titers were determined in triplicates

to further calibrate the substitution rate generated in this study. The Bayesian reconstruction of CeMV evolutionary history has indicated that the DMV strain involved in the Mediterranean outbreak in 1990–1992 was more basal to the variants that have emerged within this clade in subsequent years (as also confirmed by maximum likelihood analysis), including the source of the later DMV epizootic in the Mediterranean Sea in 2006–2008. Of particular interest were the DMV variant recovered from a fin whale stranded along the North Sea coast in 2016 and the recently published sequences of CeMV strains from dolphins in the Gulf of Mexico³⁹, as these sequences were also derived from a common ancestor. This finding suggests that undefined cetacean species may mediate the trans-oceanic spread of CeMV, driving the dispersal and evolution of this virus, as previously proposed. Pilot whales⁴⁰ and melon-headed whales (*Peponocephala electra*)⁴¹ have been suggested as viral vectors due to their extensive patterns of migration and their interaction with certain dolphin populations. However, the identity of the hosts acting as vectors that were responsible for triggering CeMV epizootics is a question that requires additional research. Moreover, it remains to be determined if porpoises and/or other cetacean species are the true hosts of PMV, given that this strain has only been previously detected in European waters and the northwest Atlantic Ocean without re-emergence since the last documented case in 1996²². In other morbilliviruses such as MV, only 8/24 known genotypes have been detected since 2009⁴², suggesting that many genotypes no longer circulate. Therefore, it is reasonable to speculate that more CeMV strains are likely to be discovered, as shown by the recent detection of the novel strains BWMV and GDMV, and some classical strains, such as PMV, may go unnoticed or disappear.

With respect to viral ecology, anthropogenic factors may have impacted the global ecology and evolution of CeMV. High levels of commercial whaling from the eighteenth to twentieth centuries decimated populations of targeted cetacean species including fin, minke (*Balaenoptera acutorostrata*), humpback (*Megaptera novaeangliae*), and sperm whales. This may well have created bottlenecks for CeMV evolution by reducing cetacean populations to levels that were unable to sustain CeMV endemicity. However, high levels of persistent environmental pollutants, including organochlorine pesticides and polychlorinated biphenyls, especially in the northern Atlantic Ocean, have also been speculated to lead to an increased susceptibility of aquatic mammals to morbillivirus infection⁴³.

The divergence of DMV sequences derived from North Sea white-beaked dolphins from the Mediterranean 1990 DMV sequences supports the endemic circulation of this variant in the white-beaked dolphin population, as

suggested previously²⁴, even before the first documented epizootic in 1990–1992 in the Mediterranean Sea. This implies that the MRCA resided in an unknown cetacean species from the North Atlantic Ocean in the mid-1970s. Furthermore, the numerous aa substitutions identified as unique to this clade might be indicative of virus adaptation to this unknown host species. The population size of white-beaked dolphins in the northeastern Atlantic and North and Barents Seas⁴⁴ might be large enough to sustain endemic transmission, supporting previous findings in which the pathogenicity of this DMV variant was reportedly reduced in this species²⁴.

The full-length genomes of DMV and PMV had a divergence of approximately 14%, which is higher than the previously observed variations between CDV strains (5–9%). However, under the current International Committee on Taxonomy of Viruses classification system, DMV and PMV are still considered as CeMV strains, as only CeMV is officially recognized as a species. Moreover, the diversity between recognized morbilliviruses is much higher, ranging between 25 and 54% at the nucleotide level. The genetic evidence presented in this study supports and extends previous observations that the transmission cycle of CeMV, in common with other animal morbilliviruses such as CDV, involves spread between multiple host species. However, this has resulted in a rather arbitrary development of CeMV strain classification. The ability of CeMV strains (e.g., DMV and BWMV) to infect multiple species from both toothed and baleen whale populations creates confusion as the original given strain names are related to the species of first identification. Therefore, we propose to rename the currently known CeMV strains DMV, PMV, PWMV, BWMV, and GDMV to CeMV-1, CeMV-2, CeMV-3, CeMV-4, and CeMV-5, respectively.

In an attempt to investigate the genetic variations driving adaptive evolution of CeMV, we highlighted the molecular footprints characterizing each DMV subgroup. We were unable to identify species-specific variations because no shared mutation was found in a comparison between viruses retrieved from two infected fin whales (DMV-DK/2016 and DMV-Bph). In contrast, the identified aa substitutions were either unique or pertained to their own monophyletic group, regardless of their host species. For instance, full CeMV genome sequences from a fin whale (DMV-Bph) and a striped dolphin (DMV-156), both from the same clade, were highly similar, with 99.9% shared pairwise sequence identity. However, we cannot discard the possibility that DMV-Bph might represent a direct spillover case from an infected dolphin. Additionally, the two fin whale cases differ with respect to age and pathology^{14,16}. Further analyses showed evidence that a number of aa sites were under episodic-positive selection, indicating instances of selective pressures that affected

only a small subset of DMV sequences within a branch. However, no single pervasive positive selection site was detected in contrast to the large number of sites under purifying selection (negative selection). Purifying selection is a major driver of evolution, removing mutations that would otherwise be deleterious. This process is particularly notable in RNA viruses, which have a high mutation rate due to error-prone virus replication. This selection might explain the low genetic diversity found among DMV variants in this study, further implying that DMV can switch between hosts belonging to both cetacean subgroups (Mysticeti and Odontoceti) and sustain transmission without adaptive evolution. Interestingly, strong purifying selection was also observed in an analysis of rabies virus genomes⁴⁵. This virus predominantly infects members of the order *Carnivora* with spillover into other mammalian species, including humans. The potential host range of CeMV may also not be restricted to only one order, given previous incidences of viral spread to monk and harbor seals^{20,22,23}. Therefore, we also assessed the ability of CeMV to use phocine SLAM as a cellular receptor. Surprisingly, CeMV efficiently used phocine SLAM for virus entry and cell-cell fusion, suggesting that in addition to PDV and CDV, CeMV may also pose a risk to seal populations.

Research on many viruses infecting marine mammals, including CeMV, is usually limited by the challenges inherent to studying viral replication and transmission in both the marine environment and appropriate in vitro model systems. To facilitate the study of wild-type CeMV strains, we generated Vero cells that constitutively expressed a dolphin SLAM receptor, thus enabling us to compare authentic wild-type sequences with those of historical and recent virus isolates. We found that the sequence of the DMV-16A isolate cultured in newly generated Vero-dolSLAM cells was comparable to that of the wild-type, except at two sites, S459P in P and I873T in L. However, we also found variable allele frequencies at these two sites in all wild-type DMV sequences. Furthermore, we compared historical laboratory-adapted DMV and PMV strains grown in Vero cells to their wild-type counterpart. A number of substitutions were found, some of which have been previously identified in MV studies. For instance, the substitution g.26A>T has been associated with MV vaccine strains²⁹ and has been shown to be important for virus replication⁴⁶. The aa substitutions in the M protein, namely E89K (PMV-2990) and R175G (DMV-muc), have also been shown to be important in MV cell culture adaptation^{30,31}. In particular, experiments with the E89K substitution in MV have demonstrated that this mutation can facilitate virus infection in Vero cells and cotton rats^{31,47}. Interestingly, some morbillivirus vaccine strains also contain these changes at positions 5 and 12 of the leader sequence,

which we identified in this study (Supplementary Fig. S1). The significance of these substitutions with respect to delineating wild-type versus laboratory virus isolates should be investigated further.

In summary, it appears that no or minimal adaptive changes are required for CeMV transmission among cetacean species, reinforcing the threat posed by this virus to endangered cetacean populations. However, the capacity of CeMV to jump into non-cetacean host species besides pinnipeds remains unknown. A more extensive survey of the interaction of CeMV with heterologous SLAM and Nectin-4 receptors may identify additional species that may be susceptible to CeMV infection. Additional full-genome CeMV sequences, and when possible, more extensive serosurveys among cetacean species, will further illuminate the evolutionary history and trajectories of this virus. This approach will necessitate new cross-disciplinary collaborations to acquire samples from regions where gaps in surveillance currently exist such as the Arctic, Indian, southern Atlantic, and western Pacific Oceans, enabling a more comprehensive global overview of this emerging and re-emerging morbillivirus, which has a major impact on the natural ecology and conservation status of dolphin and whale populations.

Materials and methods

Tissue samples and viruses

Tissues from CeMV-infected cetaceans analyzed in this study were previously investigated for the presence of morbilliviruses. The samples PMV-2990/brain⁷, DMV-16A/lung²⁸, DMV-muc/lung²⁸, DMV-DE/2007/brain⁴⁸, DMV-156/lung, DMV-LA/NL/11.2/lung²⁴, DMV-Bph/brain^{14,49}, and DMV-DK/2016/lung¹⁶ were collected from animals stranded in European waters (North Sea and Mediterranean Sea). Other samples included PMV-Ulster/88² and PMV-53⁷_passage 4 (P4), from which no original material could be acquired. We also included virus strains DMV-16A_P7, DMV-muc_P5, and PMV-2990_P4, which were previously isolated in Vero cells (green monkey kidney). In addition, DMV-16A was re-isolated from dolphin lung tissue in Vero-dolSLAM cells. Additional tissue samples from a CDV-infected dog from Germany (CDV-dog/2016) and a CDV-infected Baikal seal (*Pusa sibirica*) from Russia (CDV-phoca/1988) were used as the templates for generating additional plasmids expressing morbillivirus glycoproteins.

Generation of expression plasmids

A plasmid expressing dolphin SLAM (CD150) from a Pacific white-sided dolphin was constructed by subcloning a synthesized open reading frame (ORF) from a dolphin SLAM (Thermo Fischer Scientifics, Waltham, MA, USA) into the mammalian expression vector pCXN2 (Addgene, Cambridge, MA, USA). The design of the

synthetic fragment was analogous to previously published constructs⁵⁰ and thus contained the immunoglobulin Igk leader sequence, followed by the influenza virus HA epitope sequence with a linker sequence to the dolphin SLAM sequence (GenBank Accession No. AB428366) in which the signal sequence was removed (aa 1–18). The signal peptide cleavage site was predicted with signal P 4.1 software⁵¹. A similar design strategy was used for the construction of a plasmid expressing harbor seal SLAM. In brief, harbor seal SLAM was amplified from harbor seal spleen tissue using previously published primers⁵² and was then Sanger-sequenced (Eurofins Genomics, Munich, Germany). Primers containing sequences flanking the adjacent regions (with a linker sequence on the 5' side and pCXN2 vector sequences on the 3' side) were designed for the amplification of harbor seal SLAM sequence (GenBank Accession No. MH430950), in which the signal sequence was removed (aa 1–18). A second fragment containing the immunoglobulin Igk leader sequence, influenza virus HA epitope sequence and linker sequence, flanked by the pCXN2 vector sequence on the 5' side and seal SLAM sequence on the 3' side, was amplified. All fragments were assembled using the NEBuilder HiFi DNA assembly master mix (NEB) into pCXN2, which had been predigested with *EcoRI* to generate pCXN2-phoca-SLAM. The F and H glycoproteins from DMV-16A, CDV-dog/2016, and CDV-seal/1988 glycoproteins (GenBank Accession Nos. MH430946–49) were cloned into a pCG expression plasmid (kindly provided by Dr. Jürgen Schneider-Schaulies, Institute of Virology and Immunobiology, Würzburg).

Cell lines

Vero and 293T cells were cultured in advanced MEM and DMEM media respectively, Media were supplemented with 10% FBS, 1% penicillin/streptomycin, and 1% GlutaMax. Vero cells were transfected either with pCXN2-dolSLAM or pCXN2-phoca-SLAM using standard Lipofectamine 3000 protocols (Thermo Fischer Scientific). At 1 day post-transfection, the cells were cultured in growth media supplemented with 1 mg/ml geneticin G418 (Thermo Fischer Scientific) as a selective antibiotic. Vero-dolSLAM cells and Vero-phocaSLAM cells were sorted into single clones using a MoFlo XDP cell sorter (Beckman Coulter) to ensure an equivalent level of SLAM expression within the cell population. Successfully transfected cells were sorted by staining with anti-HA tag antibody [HA.C5] (1:200; ab18181, Abcam) as the primary antibody and chicken anti-mouse IgG (H+L) Alexa Fluor 488 (1:500; A-21200, Thermo Fischer Scientific) as the secondary antibody. Cell clones were expanded and assessed for SLAM expression using an Attune NxT flow cytometer (Thermo Fischer Scientific).

Virus infection

CeMV growth kinetics were assessed in Vero, Vero-dolSLAM and Vero-phocaSLAM cells. In brief, cells seeded into six-well plates were inoculated with DMV-16A (virus isolated in Vero-dolSLAM) at a multiplicity of infection (MOI) of 0.001 before 1 ml of OptiMEM media was added to the media and incubated at 37 °C. The combined supernatant and cell-associated virus from triplicate wells of each cell line were harvested every 12 h starting from time point 0 until 96 h (4 days). The virus titer was determined by TCID₅₀ assay in Vero-dolSLAM cells. The wells were scored as positive when syncytia formation was observed.

Cell-to-cell fusion assay

Morbillivirus cell-to-cell spread was assessed in Vero cells expressing either dolphin or harbor seal SLAM using a quantitative fusion assay as previously described⁵³, with modifications. Plasmids expressing β -galactosidase alpha and omega fragments were kindly provided by Dr. Imke Steffen (RIZ, Hannover). In brief, 293T cells were transfected using calcium phosphate co-precipitation with 3 μ g of β -galactosidase omega-fragment expression plasmid, 1.5 μ g of pCG-CeMV-F and 1.5 μ g of pCG-CeMV-H per well in a 6-well tray. Media were replaced after 8 h. Vero, Vero-dolSLAM, and Vero-phocaSLAM cells were transfected with 200 ng of β -galactosidase alpha-fragment expression plasmid per well in a 96-well tray using Lipofectamine 3000 reagents (Thermo Fischer Scientific). After 48 h, the media in the wells containing 293T cells was removed, and the cells were scraped and overlaid onto β -galactosidase alpha-fragment-transfected cells. The plates were centrifuged for 5 min at 500 \times g and incubated at 37 °C for 6 h. Cell-to-cell fusion was evaluated using a Galacto-Star™ kit (Thermo Fischer Scientific). Enzyme activity was measured using a TECAN Infinite 200 plate reader.

Sample preparation

Virus-infected tissue samples were prepared following a viral enrichment and random amplification protocol as previously described⁵⁴, with modifications. In brief, 22–221 mg of freshly frozen tissue samples from each animal and 500 μ l of the virus isolates were lysed following three freeze/thaw/homogenization cycles to ensure cell disruption. Homogenates were centrifuged and filtered (0.45 μ m). RNA was extracted using TRIzol (Thermo Fischer Scientific) and reverse-transcribed with Superscript IV (Thermo Fischer Scientific) using non-ribosomal hexamers flanked by an adaptor (5' gccggagctctgcagatc3'). Second-strand cDNA synthesis was performed with Klenow fragment DNA polymerase (New England Biolab [NEB], Ipswich, MA, USA). Random amplification of samples was performed using Phusion

polymerase (NEB) and the same primer sequence as above. To determine the amount of CeMV present in the samples before performing NGS, DMV and PMV RT-PCRs were performed as previously described⁵⁵.

Generation of full-length genomes by next-generation sequencing (NGS)

Samples were prepared according to the Nextera XT DNA Library Prep Kit protocol (Illumina, San Diego, CA, USA) using Qubit fluorometric DNA quantitation and a 2100 Bioanalyzer instrument (Agilent, Santa Clara, CA, USA) to assess quality control of the library. Paired-end sequencing of the library was conducted on an Illumina MiSeq system with the MiSeq Reagent Kit v3 (600 cycles; Illumina). Quality trimming and reference assembly were performed using CLC Genomics Workbench 10 software. A full-length DMV sequence (GenBank Accession No. AJ608288) was used as a reference. To confirm the host identity of each sample, reads were also aligned to bottlenose dolphin, bowhead whale, and green monkey sequences, which were retrieved from Ensembl and the bowhead whale genome resource⁵⁶. Quality control of single reads was performed using Trimmomatic V.0.36⁵⁷. Afterwards, high quality reads were mapped to the reference genomes using Bowtie2 V.2.3.4⁵⁸.

Completion of sequencing gaps and genome termini

Primers (Supplementary Table S2) were designed to cover regions of genomes for which we were unable to recover complete sequences by NGS (DMV_Bp, LA/NL/11.1, DMV-DE/2007, and PMV-53). Gap regions were amplified by PCR using Phusion polymerase and Sanger-sequenced (Eurofins Genomics). In addition, due to the low quality of reads from formalin-fixed paraffin-embedded brain tissue of sample DMV-DK/16, primer sets were designed to amplify the full-length genome (Supplementary Table S3). A RACE protocol was performed to determine the sequence of both 3' and 5' termini of CeMV in all positive tissues and virus isolates. Briefly, to determine the 3' end or leader sequence, RNA was first polyadenylated on the 3' end by a poly(A) polymerase (NEB) followed by cDNA generation using a poly(T) adaptor as the primer (5'-gactcgagtcgacatcg(T)₁₇-3'). A PCR was next performed using the same poly(T) adaptor as forward primer and a reverse primer designed according to consensus sequences at the 3' end of CeMV genomes obtained in this study (Supplementary Table S4). To determine the 5' end or trailer sequence, RNA was first converted to cDNA using a primer designed according to consensus sequences at the 5' end of CeMV genomes obtained in this study (Supplementary Table S4). A poly(A) tail was then added to the cDNA at the 3' end by a terminal transferase followed by a PCR using a CeMV primer and a poly(T) adaptor as the reverse primer.

Sequence analyses

Full-length genomes of morbilliviruses (virus-isolate_GenBank accession no.) DMV-GW2010007A_KU720623, DMV-Bph_MH430938, DMV-156_MH430937, DMV-DK/16_MH430939, DMV-16A_MH430934, DMV-muc_MH430935, DMV-DE/2007_MH430940, DMV-NL/11.2_MH430941, PMV-2990_MH430945, PMV-53_MH430943, PMV-Ulster/88_MH430942, DMV_AJ608288 (used as reference in this study), DMV-631IMMS031711_KU720625, DMV-BCF20110815-LA001_KU720624, MV-Zagreb_AF266290, MV-Mvi/Arizona.USA/11.08/2_JN635406.1, RPV-RBOK_Z30697, RPV-Kabete-O_X98291, PPRV-Nigeria/75/1_X74443, PPRV-CH/HNZM/2014_KM089832, CDV-Onderstepoort_AF305419, CDV-SNP/1994/spottel_hyena_1_KU578255, PDV-Wadden_Sea.NLD/1988_KC802221, and partial phosphoprotein sequences of CeMV strains were retrieved from the NCBI database. Alignments were performed with MAFFTv7.

Bayesian analysis was performed to determine the most recent common ancestor (MRCA) and estimate the rate of nucleotide substitution per site, per year using BEAST v2.4.6 software⁵⁹. The sequences were partitioned into coding regions of N, P, M, F, H, and L and non-coding sequences including the leader, trailer and the gene start and gene end of each gene. The test was run for 50 million generations with sampling at every 1000 steps and the following priors: a coalescent-constant population, strict clock model rate, and Hasegawa–Kishino–Yano as the substitution model. Tracer was used to verify quality of the analysis. The initial 10% of trees were discarded as burn-ins using TreeAnnotator, and the best tree was visualized with FigTree. Partial P gene (300 bp) was analyzed using the MASCOT package⁶⁰ and the same priors as described before. Individual complete genes were analyzed based on maximum likelihood in MEGA7.0 with 1000 bootstraps. Nucleotide substitution models were selected as a best-fit model according to Bayesian information criterion. Transient and pervasive selection sites were detected by MEME and FEL available from the Datamonkey web server (www.datamonkey.org).

Data availability

The sequences generated in this study from full-length CeMV genomes and the *Phoca vitulina* SLAM receptor have been deposited under GenBank accession numbers (MH430932-50). Plasmid sequences, cloning primers, and other data are available upon request.

Acknowledgements

We wish to thank Brett Stone (QML Vetnostics, Queensland Australia) and Jianning Wang (CSIRO, Geelong, Australia) for providing a partial P gene sequence of CeMV strain 11-00358. We acknowledge Mareike Schubert and Ilka Baumann for their technical assistance as well as Cinzia Centelleghie and Mattia Panin for their technical support with the shipment of Mediterranean cetacean samples. We also thank the Genomics Lab at the Institute of Animal Breeding and Genetics in Hannover, Germany, for operating the MiSeq system. This study was supported by a grant from the Niedersachsen-Research Network on Neuroinfectiology (N-RENNT) from the Ministry of Science and

Culture of Lower Saxony, Germany and by the R2N Project funded by the Federal State of Lower Saxony, Germany.

Author details

¹Research Center for Emerging Infections and Zoonoses, University of Veterinary Medicine Hannover, 30559 Hannover, Germany. ²Institute for Animal Breeding and Genetics, University of Veterinary Medicine Hannover, 30559 Hannover, Germany. ³Department of Viroscience, Erasmus Medical Center, 3015 CE Rotterdam, The Netherlands. ⁴Department of Comparative Biomedicine and Food Science, University of Padova, 35020 Legnaro, Padua, Italy. ⁵Faculty of Veterinary Medicine, University of Teramo, 64100 Teramo, Italy. ⁶Institute for Terrestrial and Aquatic Wildlife Research, University of Veterinary Medicine Hannover, 25761 Bűsum, Germany

Authors' contributions

A.O. and M.L. designed the study. W.K.J., J.K., A.H., M.v.d.B. and K.J. performed the laboratory work and data interpretation. T.K., U.S., S.M. and G.D.G. collected field samples as part of regional stranding networks. W.K.J. and M.L. wrote the manuscript. All authors approved the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41426-018-0207-x>).

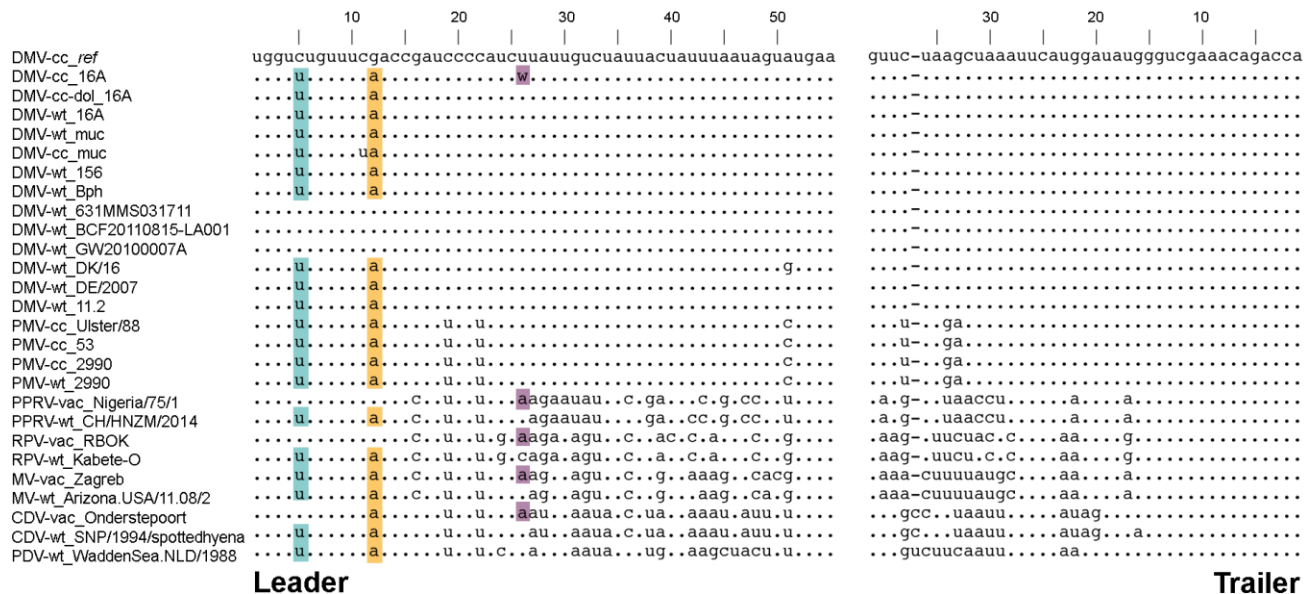
Received: 13 July 2018 Revised: 5 November 2018 Accepted: 11 November 2018

Published online: 05 December 2018

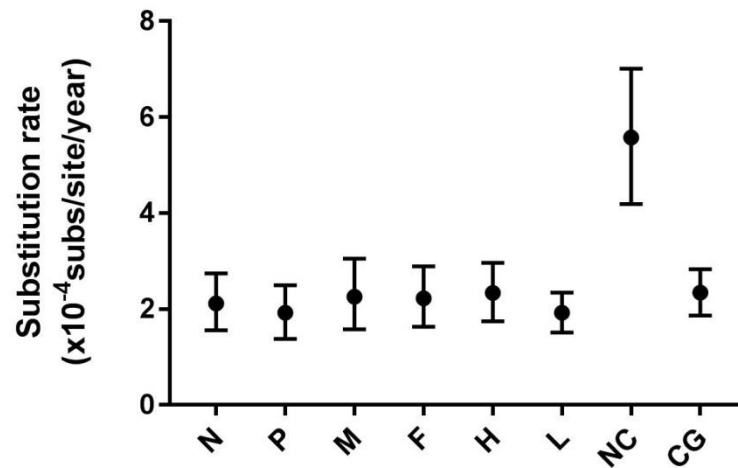
References

1. Van Bresse, M.-F. et al. Cetacean morbillivirus: current knowledge and future directions. *Viruses* **6**, 5145–5181 (2014).
2. Kennedy, S. et al. Viral distemper now found in porpoises. *Nature* **336**, 21–21 (1988).
3. Domingo, M. et al. Morbillivirus in dolphins. *Nature* **348**, 21–21 (1990).
4. Taubenberger, J. K. et al. Molecular genetic evidence of a novel morbillivirus in a long-finned pilot whale (*Globicephalus melas*). *Emerg. Infect. Dis.* **6**, 42–45 (2000).
5. Jacob, J., West, K., Levine, G., Sanchez, S. & Jensen, B. Initial characterization of novel beaked whale morbillivirus in Hawaiian cetaceans. *Dis. Aquat. Organ.* **117**, 215–227 (2016).
6. Groch, K. R. et al. Novel cetacean morbillivirus in Guiana dolphin, Brazil. *Emerg. Infect. Dis.* **20**, 511–513 (2014).
7. Visser, I. K. G. et al. Characterization of morbilliviruses isolated from dolphins and porpoises in Europe. *J. Gen. Virol.* **74**, 631–641 (1993).
8. Taubenberger, J. et al. Two morbilliviruses implicated in bottlenose dolphin epizootics. *Emerg. Infect. Dis.* **2**, 213–216 (1996).
9. Lipscomb, T. P. et al. Morbilliviral epizootic in bottlenose dolphins of the Gulf of Mexico. *J. Vet. Diagn. Investig.* **8**, 283–290 (1996).
10. Raga, J. A. et al. Dolphin morbillivirus epizootic resurgence, Mediterranean Sea. *Emerg. Infect. Dis.* **14**, 471–473 (2008).
11. Kemper, C. M. et al. Morbillivirus-associated unusual mortality event in South Australian bottlenose dolphins is largest reported for the Southern Hemisphere. *R. Soc. Open Sci.* **3**, 160838 (2016).
12. Morris, S. E. et al. Partially observed epidemics in wildlife hosts: modelling an outbreak of dolphin morbillivirus in the northwestern Atlantic, June 2013–2014. *J. R. Soc. Interface* **12**, 20150676 (2015).
13. Groch, K. R. et al. Guiana dolphin unusual mortality event and link to cetacean morbillivirus, Brazil. *Emerg. Infect. Dis.* **24**, 1349–1354 (2018).
14. Mazzariol, S. et al. Mediterranean fin whales (*Balaenoptera physalus*) threatened by dolphin morbillivirus. *Emerg. Infect. Dis.* **22**, 302–305 (2016).
15. Mazzariol, S. et al. Dolphin morbillivirus associated with a mass stranding of sperm whales, Italy. *Emerg. Infect. Dis.* **23**, 144–146 (2017).
16. Jo, W. K. et al. Dolphin morbillivirus in a fin whale (*Balaenoptera physalus*) in Denmark, 2016. *J. Wildl. Dis.* **53**, 921–924 (2017).
17. Sierra, E. et al. Morbillivirus and pilot whale deaths, Canary Islands, Spain, 2015. *Emerg. Infect. Dis.* **22**, 740–742 (2016).
18. Stephens, N. et al. Cetacean morbillivirus in coastal Indo-Pacific bottlenose dolphins, Western Australia. *Emerg. Infect. Dis.* **20**, 672–676 (2014).
19. Martinez-Gutierrez, M. & Ruiz-Saenz, J. Diversity of susceptible hosts in canine distemper virus infection: a systematic review and data synthesis. *BMC Vet. Res.* **12**, 78 (2016).
20. Osterhaus, A. et al. Morbillivirus in monk seal mass mortality. *Nature* **388**, 838–839 (1997).
21. Hernández, M. et al. Did algal toxins cause monk seal mortality? *Nature* **393**, 28–29 (1998).
22. van de Bildt, M. W. et al. Identification of morbilliviruses of probable cetacean origin in carcasses of Mediterranean monk seals (*Monachus monachus*). *Vet. Rec.* **146**, 691–694 (2000).
23. Mazzariol, S. et al. Dolphin morbillivirus infection in a captive harbor seal (*Phoca vitulina*). *J. Clin. Microbiol.* **51**, 708–711 (2013).
24. van Elk, C. E. et al. Is dolphin morbillivirus virulent for white-beaked dolphins (*Lagenorhynchus albirostris*)? *Vet. Pathol.* **51**, 1174–1182 (2014).
25. Yanagi, Y., Tatsu, H., Ono, N. & Tanaka, K. SLAM (CDw150) is a cellular receptor for measles virus. *Nature* **406**, 893–897 (2000).
26. Noyce, R. S. et al. Tumor cell marker PVRL4 (Nectin 4) is an epithelial cell receptor for measles virus. *PLoS Pathog.* **7**, e1002240 (2011).
27. Bellière, E. N., Esperón, F. & Sánchez-Vizcaíno, J. M. Genetic comparison among dolphin morbillivirus in the 1990–1992 and 2006–2008 Mediterranean outbreaks. *Infect. Genet. Evol.* **11**, 1913–1920 (2011).
28. Osterhaus, A. D. et al. Morbillivirus threat to Mediterranean monk seals? *Vet. Rec.* **130**, 141–142 (1992).
29. Parks, C. L. et al. Analysis of the noncoding regions of measles virus strains in the edmonston vaccine lineage. *J. Virol.* **75**, 921–933 (2001).
30. Takeuchi, K. et al. Recombinant wild-type and edmonston strain measles viruses bearing heterologous H proteins: role of H protein in cell fusion and host cell specificity. *J. Virol.* **76**, 4891–4900 (2002).
31. Tahara, M., Takeda, M. & Yanagi, Y. Contributions of matrix and large protein genes of the measles virus edmonston strain to growth in cultured cells as revealed by recombinant viruses. *J. Virol.* **79**, 15218–15225 (2005).
32. Murrell, B. et al. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764 (2012).
33. Kosakovsky Pond, S. L. & Frost, S. D. W. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208–1222 (2005).
34. Hashiguchi, T. et al. Structure of the measles virus hemagglutinin bound to its cellular receptor SLAM. *Nat. Struct. Mol. Biol.* **18**, 135–141 (2011).
35. Zhang, X. et al. Structure of measles virus hemagglutinin bound to its epithelial receptor nectin-4. *Nat. Struct. Mol. Biol.* **20**, 67–72 (2013).
36. Pomeroy, L. W., Bjørnstad, O. N. & Holmes, E. C. The evolutionary and epidemiological dynamics of the paramyxoviridae. *J. Mol. Evol.* **66**, 98–106 (2008).
37. Panzera, Y., Sarute, N., Iraola, G., Hernández, M. & Pérez, R. Molecular phylogeography of canine distemper virus: geographic origin and global spreading. *Mol. Phylogenet. Evol.* **92**, 147–154 (2015).
38. Ke, G.-M. et al. Phylogenetic analysis of the canine distemper virus hemagglutinin gene. *BMC Vet. Res.* **11**, 164 (2015).
39. Fauquier, D. et al. Evaluation of morbillivirus exposure in cetaceans from the northern Gulf of Mexico 2010–2014. *Endanger. Species Res.* **33**, 211–220 (2017).
40. Duignan, P. J. et al. Morbillivirus infection in two species of pilot whales (*Globicephala* sp.) from the western Atlantic. *Mar. Mammal Sci.* **11**, 150–162 (1995).
41. Stone, B. M., Blyde, D. J., Saliki, J. T. & Morton, J. M. Morbillivirus infection in live stranded, injured, trapped, and captive cetaceans in southeastern Queensland and northern New South Wales, Australia. *J. Wildl. Dis.* **48**, 47–55 (2012).
42. Moss, W. J. Measles. *Lancet* **390**, 2490–2502 (2017).
43. Aguilar, A. & Borrell, A. Abnormally high polychlorinated biphenyl levels in striped dolphins (*Stenella coeruleoalba*) affected by the 1990–1992 Mediterranean epizootic. *Sci. Total Environ.* **154**, 237–247 (1994).
44. Hammond, P. S. et al. *Lagenorhynchus albirostris*. The IUCN Red List of Threatened Species 2012 e.T11142A17875454. <https://doi.org/10.2305/IUCN.UK.2012.RLTS.T11142A17875454.en>. Accessed 20 March 2018 (2012)

45. Troupin, C. et al. Large-scale phylogenomic analysis reveals the complex evolutionary history of rabies virus in multiple carnivore hosts. *PLoS Pathog.* **12**, e1006041 (2016).
46. Liu, X., Bankamp, B., Xu, W., Bellini, W. J. & Rota, P. A. The genomic termini of wild-type and vaccine strains of measles virus. *Virus Res.* **122**, 78–84 (2006).
47. Dong, J. et al. The E89K mutation in the matrix protein of the measles virus affects in vitro cell death and virus replication efficiency in human PBMC. *Open Virol. J.* **6**, 68–72 (2012).
48. Wohlsein, P., Puff, C., Kreutzer, M., Siebert, U. & Baumgärtner, W. Distemper in a dolphin. *Emerg. Infect. Dis.* **13**, 1959–1961 (2007).
49. Beffagna, G., Centellegho, C., Franzo, G., Di Guardo, G. & Mazzariol, S. Genomic and structural investigation on dolphin morbillivirus (DMV) in Mediterranean fin whales (*Balaenoptera physalus*). *Sci. Rep.* **7**, 41554 (2017).
50. Tatsuo, H., Ono, N. & Yanagi, Y. Morbilliviruses use signaling lymphocyte activation molecules (CD150) as cellular receptors. *J. Virol.* **75**, 5842–5850 (2001).
51. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).
52. Ohishi, K. et al. Host–virus specificity of morbilliviruses predicted by structural modeling of the marine mammal SLAM, a receptor. *Comp. Immunol. Microbiol. Infect. Dis.* **33**, 227–241 (2010).
53. Steffen, I. et al. Characterization of the Bas-Congo virus glycoprotein and its function in pseudotyped viruses. *J. Virol.* **87**, 9558–9568 (2013).
54. Piewbang, C. et al. Canine bocavirus type 2 infection associated with intestinal lesions. *Vet. Pathol.* **55**, 434–441 (2018).
55. Grant, R. J., Banyard, A. C., Barrett, T., Saliki, J. T. & Romero, C. H. Real-time RT-PCR assays for the rapid and differential detection of dolphin and porpoise morbilliviruses. *J. Virol. Methods* **156**, 117–123 (2009).
56. Keane, M. et al. Insights into the evolution of longevity from the bowhead whale genome. *Cell Rep.* **10**, 112–122 (2015).
57. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
58. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
59. Drummond, A. & Bouckaert, R. *Bayesian Evolutionary Analysis with BEAST* (Cambridge University Press, Cambridge, 2015).
60. Müller, N. F., Rasmussen, D. A. & Stadler, T. The structured coalescent and its approximations. *Mol. Biol. Evol.* **34**, 2970–2981 (2017).



Supplementary Fig. 1 Genome termini alignment of CeMV sequences and other morbilliviruses. Abbreviations: cc, cell-culture; wt, wild-type; vac, vaccine. GenBank accession numbers: DMV-cc_ref (AJ608288); DMV-cc_16A (MH430932); DMV-cc_dol_16A (MH430933), DMV-wt_16A (MH430934); DMV-wt_muc (MH430935); DMV-cc_muc (MH430936); DMV-wt_156 (MH430937); DMV-wt_Bph (MH430938); DMV-wt_631IMMS031711 (KU720625); DMV-wt_BCF20110815-LA001 (KU720624); DMV-wt_GW2010007A (KU720623), DMV-wt_DK/16 (MH430939); DMV-wt_DE/2007 (MH430940); DMV-wt_11.2 (MH430941); PMV-cc_Ulster/88 (MH430942); PMV-cc_53 (MH430943); PMV-cc_2990 (MH430944); PMV-wt_2990 (MH430945); PPRV-vac_Nigeria/75/1 (X74443); PPRV-wt_CH/HNZM/2014 (KM089832); RPV-vac_RBOK (Z30697), RPV-wt_Kabete-O (X98291); MV-vac_Zagreb (AF266290); MV-Mvi/Arizona.USA/11.08/2 (JN635406.1); CDV-vac_Onderstepoort (AF305419); CDV-wt_SNP/1994/spottedhyena_1 (KU578255); PDV-wt_Wadden_Sea.NLD/1988 (KC802221).



Supplementary Fig. 2 Bayesian estimates of CeMV substitution rates of coding sequences of N, P, M, F, H, and L; non-coding sequences (NC); and complete genome (CG). Values are in $\times 10^{-4}$ nucleotide substitutions/site/year (highest posterior density interval): N, 2.12 (1.55, 2.74); P, 1.92 (1.38, 2.5); M, 2.56 (1.58, 3.04); F, 2.22 (1.63, 2.89); H, 2.33 (1.74, 2.96); L, 1.93 (1.51, 2.34); NC, 5.58 (4.19, 7.01); CG, 2.34 (1.86, 2.83). Sequences used for estimation (GenBank accession numbers): DMV-Bph (MH430938), DMV-156 (MH430937), DMV-DK/2016 (MH430939), DMV-16A (MH430934), DMV-muc (MH430935), DMV-ref (AJ608288), DMV-DE/2007 (MH430940), DMV-11.2 (MH430941), PMV-2990 (MH430945), PMV-53 (MH430943), PMV-Ulster/88 (MH430942), DMV-Sc/ES/2007 (HQ829973), DMV-Gm/ES/2007 (HQ829972), DMV-631IMMS031711 (KU720625); DMV-BCF20110815-LA001 (KU720624); DMV-GW2010007A (KU720623), and DMV-IZSPLV (MF589987). For L and CG, the sequences AJ608288, HQ829973, HQ829972 were excluded from analyses.

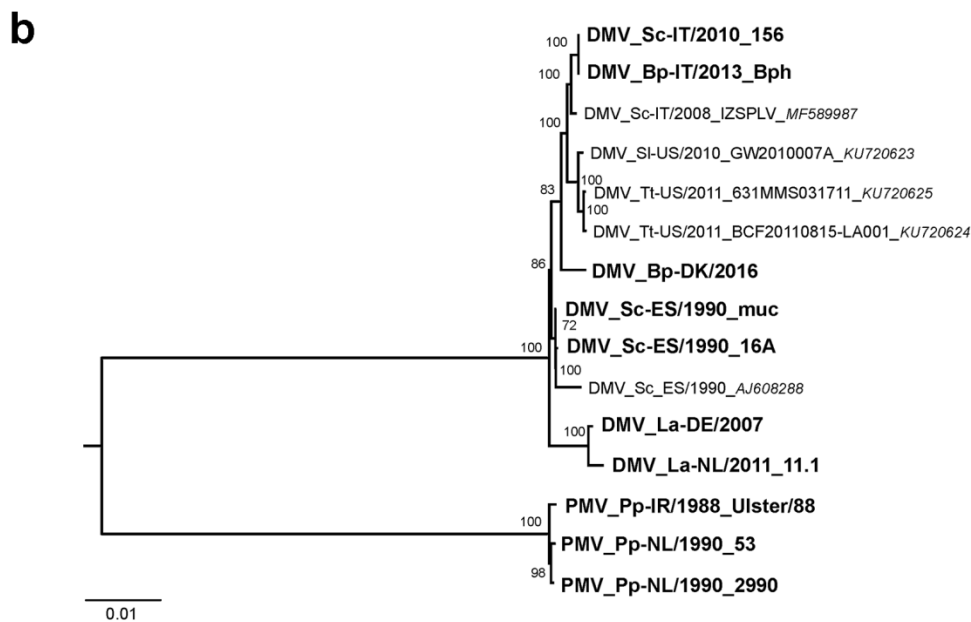
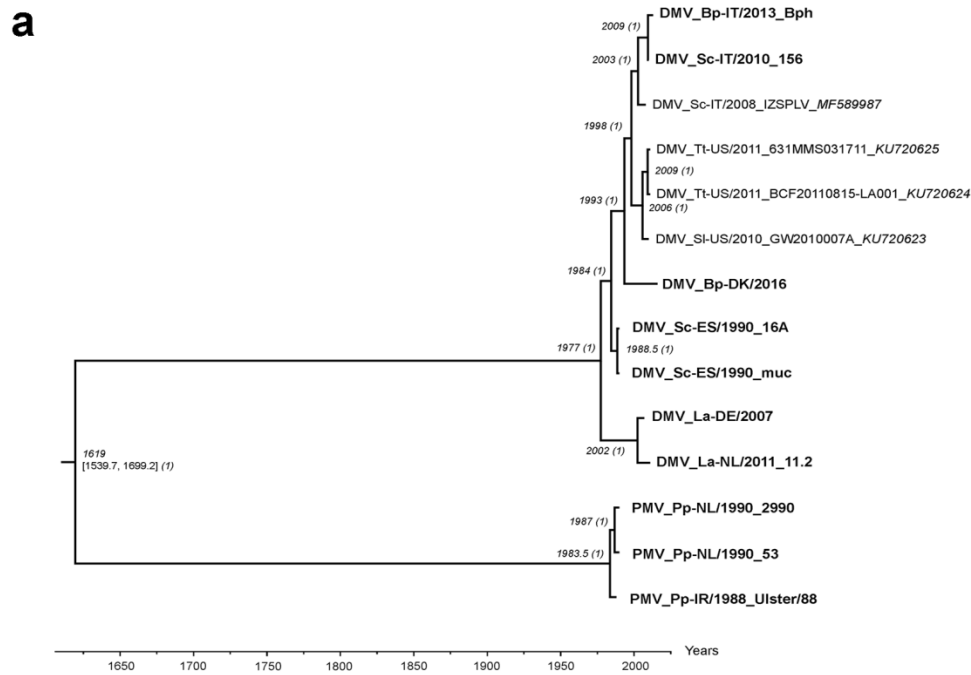
Sequence	34	93	100	107	120	122	127	138	154	180	185	189	318	473	490	508	552
Sc-ES/1990_AJ608288	V	D	H	L	I	G	C	S	A	K	T	H	S	R	P	F	Q
Sc-ES/1990 16A	L	D	H	L	I	R	C	G	A	K	T	R	S	R	P	F	K
Sc-ES/1990 muc	L	D	H	L	I	K	C	G	A	K	T	R	S	K	P	F	K
Bp-DK/2016	L	D	H	L	I	K	C	G	A	K	T	R	G	R	P	F	K
Sc-IT/2008 IZSPLV_MF589987	L	D	H	L	I	R	C	G	A	Q	T	R	S	R	P	F	K
Sc-IT/2010 156	L	D	H	L	I	K	C	G	A	K	T	R	S	R	P	F	K
Bp-IT/2013 Bph	L	D	H	L	I	K	C	G	A	K	T	R	S	R	P	F	K
Sl-US/2010 GW2010007A_KU720623	L	D	H	L	I	K	C	G	A	K	T	R	S	R	P	L	K
Tt-US/2011 631MMS031711_KU720625	L	D	H	L	I	K	C	G	A	K	T	R	S	R	L	F	K
Tt-US/2011 BCF20110815-LA001_KU720624	L	D	H	L	I	K	C	G	A	K	T	R	S	R	P	F	K
La-DE/2007	L	A	H	S	I	K	C	G	V	K	I	R	S	R	P	F	K
La-NL/2011 11.2	L	A	Q	L	S	K	G	G	V	K	I	R	S	R	P	F	K

Sequence	639	720	791	811	841	873	1006	1016	1036	1037	1039	1041	1056	1069	1070	1075	1111
Sc-ES/1990_AJ608288	G	M	L	L	M	T	V	I	A	D	H	V	R	A	V	L	S
Sc-ES/1990 16A	G	M	L	L	M	I	V	I	D	E	Q	E	I	T	G	I	F
Sc-ES/1990 muc	G	M	L	L	M	T	V	I	D	E	Q	E	I	T	G	I	F
Bp-DK/2016	G	I	L	L	M	T	V	I	D	E	Q	E	I	T	G	I	F
Sc-IT/2008 IZSPLV_MF589987	G	M	L	L	M	T	V	I	D	E	Q	E	I	T	G	I	F
Sc-IT/2010 156	G	M	L	L	M	T	V	I	D	E	Q	E	I	T	G	I	F
Bp-IT/2013 Bph	G	M	L	L	M	T	V	I	D	E	Q	E	I	T	G	I	F
Sl-US/2010 GW2010007A_KU720623	G	M	L	L	T	T	V	I	D	E	Q	E	I	T	G	I	F
Tt-US/2011 631MMS031711_KU720625	G	M	L	L	T	T	V	I	D	E	Q	E	I	T	G	I	F
Tt-US/2011 BCF20110815-LA001_KU720624	G	M	L	L	T	T	V	I	D	E	Q	E	I	T	G	I	F
La-DE/2007	S	M	L	L	M	T	L	V	D	E	Q	E	I	T	G	I	F
La-NL/2011 11.2	S	M	I	V	M	I	L	V	D	E	Q	E	I	K	G	I	F

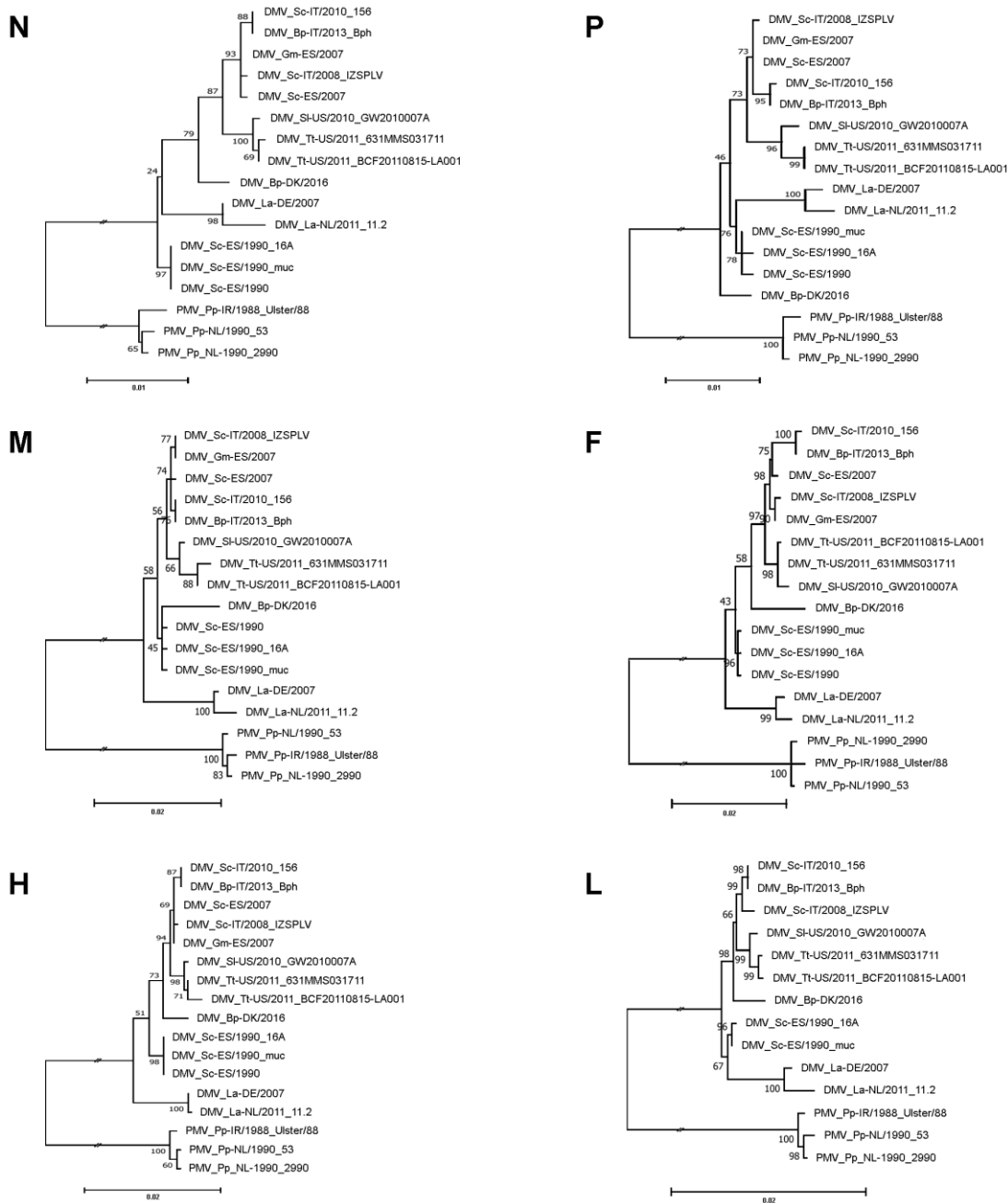
Sequence	1120	1129	1161	1170	1192	1194	1196	1228	1237	1253	1254	1275	1343	1388	1611	1627	1696
Sc-ES/1990_AJ608288	V	R	V	I	G	A	C	A	S	D	Q	R	I	S	Y	S	I
Sc-ES/1990 16A	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Sc-ES/1990 muc	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Bp-DK/2016	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Sc-IT/2008 IZSPLV_MF589987	G	S	D	M	G	A	F	V	A	E	Q	R	I	R	H	S	I
Sc-IT/2010 156	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Bp-IT/2013 Bph	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Sl-US/2010 GW2010007A_KU720623	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Tt-US/2011 631MMS031711_KU720625	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
Tt-US/2011 BCF20110815-LA001_KU720624	G	S	D	M	V	S	C	V	A	E	Q	R	I	R	H	S	I
La-DE/2007	G	S	D	M	V	S	C	V	A	E	R	R	V	R	H	P	T
La-NL/2011 11.2	G	S	D	M	V	S	C	V	A	E	Q	K	V	R	H	P	T

Sequence	1698	1745	1748	1811	1867	1969	1972	1977	2000	2001	2005	2016	2072	2088	2107	2147	2155
Sc-ES/1990_AJ608288	A	R	N	R	I	N	G	L	Y	L	I	R	E	A	I	R	P
Sc-ES/1990 16A	A	R	S	S	I	N	G	I	C	I	V	R	E	A	I	K	Q
Sc-ES/1990 muc	A	R	S	S	I	N	G	I	C	I	V	R	E	A	I	K	Q
Bp-DK/2016	A	R	S	S	I	N	G	I	C	I	V	R	E	A	I	K	Q
Sc-IT/2008 IZSPLV_MF589987	A	R	S	S	I	N	G	I	C	I	V	R	E	A	I	K	Q
Sc-IT/2010 156	A	R	S	S	I	N	G	I	C	I	V	R	E	A	I	K	Q
Bp-IT/2013 Bph	A	R	S	S	I	N	G	I	C	I	V	R	E	A	I	K	Q
Sl-US/2010 GW2010007A_KU720623	A	R	S	S	I	N	V	I	C	I	V	I	D	V	I	K	Q
Tt-US/2011 631MMS031711_KU720625	A	R	S	S	I	N	V	I	C	I	V	I	E	V	I	K	Q
Tt-US/2011 BCF20110815-LA001_KU720624	A	R	S	S	I	N	V	I	C	I	V	I	E	V	I	K	Q
La-DE/2007	S	R	S	R	V	N	G	I	C	I	V	R	E	A	N	K	Q
La-NL/2011 11.2	A	K	S	R	V	D	G	I	C	I	V	R	E	A	N	K	Q

Supplementary Fig. 3 Amino acid changes in DMV Large protein. 28/68 changes (green) were accounted for by Sc-ES/1990 (GenBank accession no. AJ608288). Taxon names are presented as host_country/year of collection|variant_GenBank Accession No. Abbreviations: Bp, *Balaenoptera physalus*; Sc, *Stenella coeruleoalba*; Tt, *Tursiops truncatus*; Sl, *Stenella longirostris*; La, *Lagenorhynchus albirostris*; DMV, dolphin morbillivirus; ES, Spain; IT, Italy; US, United States of America; DK, Denmark; DE, Germany; NL, the Netherlands.



Supplementary Fig. 4 Phylogenetic analyses of full-length CeMVs. **a** Bayesian reconstruction of full-length CeMVs. Ages are presented at nodes, posterior values in parenthesis. **b** Maximum-likelihood reconstruction of full-length CeMVs. Bootstrap values are presented at nodes. Taxon names are presented as virus_host_country/year of collection_variant. GenBank accession number in parentheses: Bph (MH430938), 156 (MH430937), DK/2016 (MH430939), 16A (MH430934), muc (MH430935), DE/2007 (MH430940), 11.2 (MH430941), 2990 (MH430945), 53 (MH430943), Ulster/88 (MH430942). Abbreviations: Bp, *Balaenoptera physalus*; Sc, *Stenella coeruleoalba*; Tt, *Tursiops truncatus*; Sl, *Stenella longirostris*; La, *Lagenorhynchus albirostris*; Pp, *Phocoena phocoena*; DMV, dolphin morbillivirus; PMV, porpoise morbillivirus; ES, Spain; IT, Italy; US, United States of America; DK, Denmark; DE, Germany; NL, the Netherlands; IR, Northern Ireland, U.K.



Supplementary Fig. 5 Maximum likelihood reconstruction of individual genes. Bootstrap values at nodes. Taxon names are presented as virus_host_country/year of collection_variant. GenBank accession number in parentheses: Bph (MH430938), 156 (MH430937), IZSPLV (MF589987), 631MMS031711 (KU720625); BCF20110815-LA001 (KU720624), GW2010007A (KU720623), DK/2016 (MH430939), 16A (MH430934), muc (MH430935), DE/2007 (MH430940), 11.2 (MH430941), 2990 (MH430945), 53 (MH430943), Ulster/88 (MH430942). Abbreviations: Bp, *Balaenoptera physalus*; Sc, *Stenella coeruleoalba*; Tt, *Tursiops truncatus*; SI, *Stenella longirostris*; La, *Lagenorhynchus albirostris*; Pp, *Phocoena phocoena*; DMV, dolphin morbillivirus; PMV, porpoise morbillivirus; ES, Spain; IT, Italy; US, United States of America; DK, Denmark; DE, Germany; NL, the Netherlands; IR, Northern Ireland, U.K.

Supplementary Table 1 Negative selection sites based on FEL

Protein	Site	Partition	alpha	beta	alpha=beta	LRT	p-value	Total branch length
Nucleoprotein	161	1	21,474	0.000	3,089	5,980	0.014	3,116
	103	1	18,114	0.000	3,008	5,675	0.017	2,628
	397	1	11,559	0.000	2,194	5,294	0.021	1,677
	248	1	7,437	0.000	1,728	4,993	0.025	1,079
	470	1	8,020	0.000	1,876	5,046	0.025	1,164
	76	1	17,641	0.000	1,045	4,931	0.026	2,560
	439	1	8,775	0.000	1,912	4,911	0.027	1,273
	319	1	9,182	0.000	1,805	4,840	0.028	1,332
	514	1	44,868	1,483	10,066	4,634	0.031	6,982
	85	1	14,984	0.000	0.819	4,587	0.032	2,174
	34	1	8,120	0.000	1,955	4,514	0.034	1,178
	520	1	16,961	0.000	0.968	4,467	0.035	2,461
	151	1	17,663	0.000	1,032	4,083	0.043	2,563
	38	1	11,171	0.000	0.837	4,023	0.045	1,621
	487	1	20,358	0.000	1,817	4,012	0.045	2,954
	56	1	13,549	0.000	1,031	3,969	0.046	1,966
	63	1	14,905	0.000	1,207	3,969	0.046	2,163
	412	1	6,034	0.000	1,012	3,917	0.048	0.875
	90	1	6,974	0.000	2,397	3,861	0.049	1,012
	283	1	6,034	0.000	1,068	3,877	0.049	0.875
Phosphoprotein	480	1	18,973	0.000	2,621	6,617	0.010	1,618
	103	1	49,500	0.000	1,914	5,441	0.020	4,221
	337	1	20,081	0.000	3,959	5,442	0.020	1,713
	347	1	36,637	0.000	1,937	5,244	0.022	3,124
	190	1	16,486	0.000	5,598	4,844	0.028	1,406
	216	1	14,115	0.000	3,378	4,787	0.029	1,204
	498	1	23,567	0.000	2,954	4,777	0.029	2,010
	396	1	10,971	0.000	3,168	4,416	0.036	0.936
	386	1	11,729	0.000	2,157	4,043	0.044	1,000
	338	1	11,626	0.000	2,254	3,830	0.050	0.992
Matrix	140	1	25156	0.000	5109	7315	0.007	4740
	242	1	17547	0.000	2443	6328	0.012	3306
	169	1	9921	0.000	1952	5210	0.022	1869
	58	1	10901	0.000	1081	4924	0.026	2054
	83	1	8547	0.000	0.986	4979	0.026	1610
	179	1	18199	0.000	0.864	4760	0.029	3429
	279	1	10305	0.000	0.948	4769	0.029	1942
	158	1	7439	0.000	2548	4692	0.030	1402

259	1	13350	0.000	0.778	4672	0.031	2515	
262	1	13350	0.000	0.795	4630	0.031	2515	
290	1	13561	0.000	0.776	4646	0.031	2555	
253	1	9087	0.000	2129	4575	0.032	1712	
181	1	12855	0.000	0.753	4299	0.038	2422	
280	1	6766	0.000	1838	4321	0.038	1275	
5	1	14152	0.000	0.828	4200	0.040	2667	
211	1	11916	0.000	0.754	4226	0.040	2245	
216	1	10209	0.000	0.604	4160	0.041	1924	
237	1	12472	0.000	0.736	4158	0.041	2350	
302	1	7036	0.000	1066	4168	0.041	1326	
141	1	11841	0.000	0.738	4151	0.042	2231	
117	1	6176	0.000	0.742	4044	0.044	1164	
131	1	8988	0.000	0.594	4012	0.045	1693	
205	1	6176	0.000	0.731	4019	0.045	1164	
272	1	8938	0.000	0.593	4006	0.045	1684	
84	1	5896	0.000	0.736	3974	0.046	1111	
87	1	8988	0.000	0.597	3971	0.046	1693	
294	1	11841	0.000	0.823	3997	0.046	2231	
69	1	6027	0.000	0.922	3897	0.048	1136	
27	1	7986	0.000	0.593	3826	0.050	1505	
Fusion	437	1	17483	0.000	1528	8418	0.004	2451
	284	1	18822	0.000	0.883	6146	0.013	2639
	96	1	33271	0.000	1735	5979	0.014	4664
	454	1	36047	0.000	1785	6007	0.014	5054
	160	1	14105	0.000	2608	5897	0.015	1977
	333	1	16036	0.000	2239	5861	0.015	2248
	398	1	15120	0.000	2443	5476	0.019	2120
	138	1	7753	0.000	1136	4948	0.026	1087
	190	1	9270	0.000	1714	4989	0.026	1300
	196	1	9056	0.000	2206	4931	0.026	1270
	448	1	13512	0.000	1263	4906	0.027	1894
	226	1	8679	0.000	1355	4779	0.029	1217
	342	1	7677	0.000	1820	4796	0.029	1076
	400	1	13858	0.000	2070	4651	0.031	1943
	412	1	15893	0.000	1161	4627	0.031	2228
	489	1	7753	0.000	1356	4675	0.031	1087
	517	1	13858	0.000	2070	4651	0.031	1943
	442	1	10453	0.000	2121	4544	0.033	1465
	44	1	13009	0.000	0.691	4266	0.039	1824
	515	1	9853	0.000	3270	4254	0.039	1381

	100	1	12463	0.000	0.994	4159	0.041	1747
	184	1	10387	0.000	1065	4107	0.043	1456
	334	1	9052	0.000	1249	4007	0.045	1269
	514	1	9460	0.000	0.888	4010	0.045	1326
	328	1	15346	0.000	0.907	3971	0.046	2151
	40	1	8880	0.000	0.904	3945	0.047	1245
	394	1	6118	0.000	1921	3898	0.048	0.858
Hemagglutinin	77	1	34493	0.000	0.905	6481	0.011	5606
	97	1	16121	0.000	2608	5812	0.016	2620
	7	1	16730	0.000	2181	5687	0.017	2719
	57	1	11406	0.000	1596	5664	0.017	1854
	575	1	24476	0.000	0.761	5682	0.017	3978
	571	1	12695	0.000	2101	5615	0.018	2063
	282	1	28277	0.000	1011	5429	0.020	4596
	446	1	24476	0.000	0.954	5409	0.020	3978
	43	1	10589	0.000	1214	5227	0.022	1721
	63	1	7578	0.000	1443	5279	0.022	1232
	281	1	21260	0.000	0.953	5275	0.022	3455
	18	1	7531	0.000	1368	5175	0.023	1224
	324	1	23772	0.000	0.717	5135	0.023	3864
	377	1	18423	0.000	0.765	4912	0.027	2994
	529	1	6742	0.000	1647	4811	0.028	1096
	55	1	11263	0.000	2177	4688	0.030	1830
	11	1	18423	0.000	0.899	4654	0.031	2994
	459	1	7578	0.000	1962	4613	0.032	1232
	340	1	7531	0.000	1738	4567	0.033	1224
	349	1	7531	0.000	1738	4567	0.033	1224
	399	1	20008	0.000	0.746	4507	0.034	3252
	508	1	7578	0.000	1785	4519	0.034	1232
	504	1	8034	0.000	1448	4378	0.036	1306
	100	1	12532	0.000	0.697	4127	0.042	2037
	286	1	14890	0.000	1064	4005	0.045	2420
	535	1	21708	0.000	0.837	4020	0.045	3528
Large Protein	1418	1	45.309	0	5.836	11.187	0.001	6.998
	778	1	31.28	0	4.482	9.806	0.002	4.831
	586	1	34.653	0	6.049	9.104	0.003	5.352
	384	1	21.853	0	4.574	8.082	0.004	3.375
	411	1	37.341	0	5.485	8.257	0.004	5.767
	439	1	49.136	0	1.175	7.96	0.005	7.589
	527	1	33.01	0	2.311	7.652	0.006	5.098
	2183	1	42.315	0	1.087	7.697	0.006	6.536

1150	1	19.223	0	3.121	7.325	0.007	2.969
1280	1	80.844	0	0.992	7.236	0.007	12.486
1297	1	18.469	0	4.325	7.312	0.007	2.853
64	1	28.325	0	2.697	6.95	0.008	4.375
99	1	29.628	0	1.455	7.069	0.008	4.576
1772	1	16.775	0	3.557	6.93	0.008	2.591
58	1	15.031	0	3.232	6.823	0.009	2.322
225	1	22.491	0	2.388	6.56	0.01	3.474
1227	1	22.491	0	2.388	6.56	0.01	3.474
1723	1	22.001	0	2.425	6.455	0.011	3.398
2089	1	31.247	0	0.883	6.32	0.012	4.826
310	1	22.351	0	1.606	6.121	0.013	3.452
575	1	17.615	0	2.555	6.179	0.013	2.721
952	1	34.226	0	1.06	6.119	0.013	5.286
954	1	12.479	0	1.52	6.123	0.013	1.927
1811	1	24.175	0	0.728	6.17	0.013	3.734
429	1	17.402	0	2.321	6.051	0.014	2.688
604	1	30.403	0	0.911	6.054	0.014	4.696
1377	1	29.485	0	0.912	6.054	0.014	4.554
2127	1	10.786	0	1.507	5.997	0.014	1.666
2164	1	16.211	0	2.323	5.997	0.014	2.504
420	1	23.208	0	0.983	5.968	0.015	3.585
562	1	29.485	0	1.054	5.912	0.015	4.554
1519	1	24.175	0	0.806	5.972	0.015	3.734
1861	1	24.175	0	0.875	5.894	0.015	3.734
203	1	23.467	0	2.194	5.854	0.016	3.625
1590	1	29.485	0	1.06	5.808	0.016	4.554
293	1	19.522	0	5.83	5.721	0.017	3.015
1632	1	25.676	0	0.726	5.684	0.017	3.966
80	1	10.758	0	1.578	5.578	0.018	1.662
512	1	24.175	0	1.045	5.629	0.018	3.734
1139	1	9.702	0	1.623	5.602	0.018	1.499
406	1	15.714	0	0.938	5.524	0.019	2.427
1158	1	23.601	0	0.805	5.5	0.019	3.645
1193	1	11.096	0	1.531	5.468	0.019	1.714
1483	1	24.175	0	1.052	5.525	0.019	3.734
226	1	22.248	0	0.972	5.436	0.02	3.436
536	1	19.182	0	0.964	5.233	0.022	2.963
1219	1	19.182	0	0.964	5.233	0.022	2.963
1249	1	24.175	0	1.345	5.187	0.023	3.734
1459	1	16.895	0	1.718	5.189	0.023	2.609

1594	1	18.589	0	0.909	5.154	0.023	2.871
394	1	7.726	0	1.014	5.073	0.024	1.193
739	1	11.367	0	2.37	5.069	0.024	1.756
835	1	7.726	0	1.014	5.073	0.024	1.193
1790	1	7.726	0	1.014	5.073	0.024	1.193
2055	1	7.726	0	1.014	5.073	0.024	1.193
2145	1	10.273	0	1.874	5.032	0.025	1.587
405	1	7.369	0	1.02	4.972	0.026	1.138
460	1	12.608	0	3.511	4.923	0.026	1.947
1001	1	7.726	0	1.088	4.931	0.026	1.193
1151	1	7.369	0	1.02	4.972	0.026	1.138
1183	1	6.087	0	1.113	4.938	0.026	0.94
46	1	13.961	0	1.225	4.915	0.027	2.156
934	1	7.726	0	1.103	4.8	0.028	1.193
1259	1	7.369	0	1.078	4.854	0.028	1.138
1648	1	7.369	0	1.078	4.854	0.028	1.138
2037	1	5.779	0	1.1	4.855	0.028	0.893
376	1	23.767	0	1.567	4.689	0.03	3.671
680	1	6.087	0	1.201	4.727	0.03	0.94
714	1	7.369	0	1.095	4.723	0.03	1.138
1097	1	7.369	0	1.095	4.723	0.03	1.138
1550	1	7.369	0	1.095	4.723	0.03	1.138
1082	1	16.584	0	0.995	4.664	0.031	2.561
1213	1	20.627	0	1.002	4.673	0.031	3.186
2050	1	5.792	0	1.188	4.646	0.031	0.895
655	1	7.726	0	1.289	4.611	0.032	1.193
1721	1	6.087	0	1.205	4.596	0.032	0.94
1927	1	22.248	0	1.56	4.611	0.032	3.436
1953	1	7.726	0	1.289	4.611	0.032	1.193
248	1	17.357	0	0.727	4.548	0.033	2.681
1484	1	7.369	0	1.275	4.536	0.033	1.138
359	1	5.792	0	1.193	4.514	0.034	0.895
492	1	5.792	0	1.193	4.514	0.034	0.895
1201	1	11.514	0	1.747	4.505	0.034	1.778
1349	1	11.785	0	0.688	4.499	0.034	1.82
1755	1	11.785	0	0.688	4.499	0.034	1.82
102	1	16.208	0	0.995	4.428	0.035	2.503
205	1	13.794	0	1.028	4.465	0.035	2.131
244	1	17.15	0	0.695	4.463	0.035	2.649
391	1	13.794	0	1.028	4.465	0.035	2.131
1478	1	7.369	0	1.39	4.421	0.035	1.138

1629	1	6.33	0	1.763	4.452	0.035	0.978
2168	1	13.794	0	1.028	4.465	0.035	2.131
1218	1	15.509	0	0.771	4.416	0.036	2.395
587	1	14.836	0	0.693	4.334	0.037	2.291
1476	1	8.292	0	1.327	4.358	0.037	1.281
2156	1	5.594	0	1.708	4.335	0.037	0.864
759	1	17.143	0	0.709	4.325	0.038	2.648
1119	1	13.158	0	0.926	4.321	0.038	2.032
1589	1	20.569	0	0.812	4.306	0.038	3.177
538	1	14.453	0	0.766	4.268	0.039	2.232
626	1	6.469	0	1.971	4.245	0.039	0.999
170	1	17.249	0	0.788	4.2	0.04	2.664
181	1	7.369	0	1.454	4.179	0.041	1.138
351	1	7.369	0	1.454	4.179	0.041	1.138
1072	1	5.988	0	1.127	4.158	0.041	0.925
1173	1	5.988	0	1.127	4.158	0.041	0.925
1178	1	7.369	0	1.454	4.179	0.041	1.138
797	1	15.022	0	0.788	4.116	0.042	2.32
1934	1	17.411	0	0.808	4.125	0.042	2.689
98	1	19.696	0	0.858	4.11	0.043	3.042
511	1	18.341	0	1.027	4.081	0.043	2.833
1375	1	16.735	0	0.807	4.089	0.043	2.585
145	1	17.033	0	0.85	4.057	0.044	2.631
1819	1	18.415	0	0.855	4.04	0.044	2.844
1935	1	13.37	0	0.731	4.001	0.045	2.065
2147	1	32.089	1.177	4.033	3.964	0.046	5.381
1911	1	6.994	0	0.688	3.947	0.047	1.08
1920	1	13.526	0	0.641	3.952	0.047	2.089
118	1	13.37	0	0.802	3.898	0.048	2.065
1399	1	16.722	0	1.018	3.896	0.048	2.583
471	1	6.994	0	0.717	3.868	0.049	1.08
1323	1	6.994	0	0.717	3.868	0.049	1.08
1622	1	6.994	0	0.717	3.868	0.049	1.08
2133	1	13.37	0	0.842	3.83	0.05	2.065

Supplementary Table 2 Wildtype morbilliviruses Hemagglutinin-SLAM interactions sites

Receptor	H	MV	RPV	PPRV	CeMV		CDV	PDV	
					DMV	PMV			
SLAM*	Site I	D 505	D 505	D 505	D 501	D 501	D 501	D 501	
		D 507	D 507	D 507	D 503	D 503	D 503	D 503	
	Site II	D 530	D 530	D 530	D 526	D 526	D 526	D 526	
		T 531	I 531	V 531	T 527	T 527	V 527	V 527	
		S 532	S 532	S 532	S 528	S 528	S 528	S 528	
		R 533	R 533	R 533	R 529	R 529	R 529	R 529	
		F 552	Y 552	F 552	F 548	L 548	T 548	T 548	
		P 554	P 554	P 554	P 550	P 550	P 550	P 550	
	Site III	P 191	P 191	R 191	S 191	S 191	A 191	A 191	
		T 192	T 192	T 192	V 192	V 192	T 192	T 192	
		T 193	T 193	V 193	I 193	I 193	T 193	T 193	
		I 194	T 194	T 194	T 194	T 194	S 194	S 194	
		R 195	K 195	R 195	R 195	R 195	V 195	M 195	
	Site IV	Y 524	Y 524	Y 524	Y 520	Y 520	Y 520	Y 520	
		Y 541	H 541	Y 541	Y 537	Y 537	Y 537	Y 537	
		Y 543	Y 543	Y 543	Y 539	Y 539	Y 539	Y 539	
		F 552	Y 552	F 552	F 548	L 548	T 548	T 548	
		F 483	L 483	L 483	L 479	L 479	L 479	L 479	
		P 545	T 545	T 545	T 541	T 541	P 541	P 541	
	PVRL4 [#]	Site I	Y 524	Y 524	Y 524	Y 520	Y 520	Y 520	Y 520
			L 526	S 526	T 526	T 522	T 522	I 522	V 522
			Y 541	H 541	Y 541	Y 537	Y 537	Y 537	Y 537
			Y 543	Y 543	Y 543	Y 539	Y 539	Y 539	Y 539
			P 458	P 458	P 458	P 454	P 454	P 454	P 454
			M 459	L 459	Y 459	L 455	L 455	K 455	K 455
			L 462	S 462	S 462	S 458	S 458	T 458	T 458
			A 463	A 463	F 463	V 459	V 459	V 459	I 459
			L 464	L 464	L 464	L 460	L 460	L 460	L 460
G 465			G 465	G 465	G 461	G 461	G 461	G 461	
L 482		I 482	I 482	T 478	T 478	V 478	I 478		
F 483		L 483	L 483	L 479	L 479	L 479	L 479		
Site II		T 498	T 498	I 498	V 494	V 494	I 494	I 494	
		Y 499	Y 499	E 499	E 495	E 495	Q 495	Q 495	
		D 505	D 505	D 505	D 501	D 501	D 501	D 501	
		L 500	L 500	L 500	L 496	L 496	T 496	T 496	
	Site III	K 387	R 387	K 387	K 383	K 383	Q 383	K 383	
		G 388	D 388	T 388	S 384	G 384	R 384	R 384	
K 389		R 389	R 389	R/K 385[§]	R 385	K 385	R 385		
Q 391		P 391	P 391	P 387	P 387	Y 387	Y 387		
Y 499		Y 499	E 499	E 495	E 495	Q 495	Q 495		
L 500		L 500	L 500	L 496	L 496	T 496	T 496		

Conserved regions are marked in bold, sites in which DMV and PMV differ are in green

*Sites are based on MV-H interaction with marmoset SLAM (Hashiguchi et al 2011)

[#]Sites are based on MV-H interaction with human nectin-4 (Zhang et al 2013)

[§]In site 385, only variant DMV-DK/16 (North Sea fin whale) has a K instead of R

Supplementary Table 3 Primers sets covering gap regions of DMV and PMV

Strain	Genome	Gap region (bp)	Forward primer (5' – 3')	Reverse primer (5' – 3')
DMV	DMV/LA/NL/11.2, DE/2007, DMV_Bp	1 – 975	accaRacaaagYtggtca	agaattcatgcaaccctaagg
	DMV_Bp	1290 – 1805	atcagcaggcaagggtgag	agatttgagacactcaagtccc
	DMV/LA/NL/11.2, DE/2007	1929 – 2905	atcaagcgagagtcacacc	agttctacgtctgcagtaggac
	DMV/LA/NL/11.2	3856 – 4652	atgtgtcaatcttatccac	aRcttMtatgtgcatgc
	DMV/LA/NL/11.2	5671 – 6257	agataactgcaggagttgcc	ttagatagccattggtcgc
	DMV/LA/NL/11.2	6888 – 7898	aagcctgatctaacaggtactacc	ttgtcatctgcagcacKg
	DMV/LA/NL/11.2, DE/2007	8420 – 9456	agtagcaatcaagactggctg	actgcYtatccctagccttg
	DMV/LA/NL/11.2	9720 – 10201	atggacttatgtactgcatgtag	ttcatcatagtctcgtagttgataac
	DMV/LA/NL/11.2	12189 – 12997	atggatagacacatMatcaKacctag	tggtgtatacctcgctactcgg
	DMV/LA/NL/11.2	14479 – 14917	gtcagtcgagcttagatctgg	aatactagataacactcagtgagat
PMV	53	1459 – 1946	atgccataccaagaggtacaag	tgacgactctcgcttgagtc

Supplementary Table 4. Primer sets used to amplify and sequence full-length genome of DMV-DK/16

Gap region (bp)	Forward primer (5' – 3')	Reverse primer (5' – 3')	Type of primer
176 – 1325	ttcaggaggagccataagagggattaagcatg	agctctgctgcattagcctg	PCR
1220 – 2530	atcagcaggcaaggtag	cggctcctgcccagataa	PCR
2442 – 4486	aattccaacctcagtgcc	ggatgaactagctatgagcct	PCR
4046 – 4959	atacatgagcatgatgagagacc	ctcttgctatgctgagatct	PCR
4939 – 5885	agatctcgcataaggcaaagag	aatttctgcttagcatctcgc	PCR
5791 – 7015	actgtcttagctgtcaaggggt	ctcaggtatgaccaataatccagta	PCR
6888 – 8538	gcctgatctaacaggtactaccaa	cgactgcagttgtaagggtat	PCR
8331 – 9893	gaggactcaggtctagagcttgat	ggctctaataatgccacgatt	PCR
9767 – 11177	ctgcaatgactgtggacccta	ggacagtgaggatcgctta	PCR
10991 – 12578	aggcttacgagactgtcagtg	cataatgaacaagactcatgacg	PCR
12483 – 14023	cgcaagaggaaggtcgat	tcggtcatcattccaatcc	PCR
13882 – 15702	ctgcatgactgtactgtaatgc	accagacaaagctgggtatag	PCR
662	agatactgcagctgactccgagacg	-	Sequencing
955	ccttagggtgcatgaattctc	-	Sequencing
1750	agtccaaggaattgacctccacatc	-	Sequencing
2174	atgctgacctgctcgtggtcc	-	Sequencing
2953	actgcagacgtagaactcaatcccg	-	Sequencing
3183	gagtgtgcttcgctccatgattaagtc	-	Sequencing
3528	cctcaagttagattatagatccaggct	-	Sequencing
5400	agtccagctataaagtgatgaccagg	-	Sequencing
6132	gcatacccgactctatcagaggtcaa	-	Sequencing
6548	agttgaggtagacggaatcacaatcc	-	Sequencing
7379	aagatcatcgagacgaagtgg	-	Sequencing
7802	gacctgatgatccattgagttccaa	-	Sequencing
8222	actcaagcatttgcaagagtagaccacc	-	Sequencing
8802	ggcaactctggtgtcatcactactgcat	-	Sequencing
9196	gggtcgagaatggttactccaatca	-	Sequencing
9559	ctgagatgcggtccatcatcaa	-	Sequencing
10179	atcaactacgagactatgatgaagggtc	-	Sequencing
10597	atggctcctatttgaagaccctg	-	Sequencing
11456	aatcctaaggcagcgattgcacg	-	Sequencing
11822	gattaagatggcactactacgtgctc	-	Sequencing
12053	atgtgtccagagtatcactcggc	-	Sequencing
12945	acatcgtaatccgagtagcgagg	-	Sequencing
13248	ccaatcattgagaaggatgcgta	-	Sequencing
13726	atctcacggtctgtaacatgatctaccac	-	Sequencing
14318	gaactcatctgctgtacaaggca	-	Sequencing
14595	ccagaggtgacatgggtaggtaatgtaga	-	Sequencing
14986	taagaacatcaccgggattttagcac	-	Sequencing
15189	gcgtctggatcagagggctctgtcaa	-	Sequencing

Supplementary Table 5 RACE primer for DMV and PMV

Strain	Primer	Region (bp)	Primer (5' – 3')
DMV	Leader-outer	291	RV-ggtcaccagcaagtctactaa
	Leader-inner	142	RV-ttgaacagagctagactccg
	Trailer-outer	15622	FW-tcctgtcagagtgatcag
	Trailer-inner	15558	FW-aagttgatagggtacgggtgc
PMV	Leader-outer	172	RV-attgctcctcctgaacctg
	Leader-inner	103	RV-agaaggtcgcctatgattg
	Trailer-outer	15601	FW-tagaacggatccctactcct
	Trailer-inner	15556	FW-ttaagctgatagggtatggc
-	Adaptor ^a	-	gactcgagtcgacatcg
-	T17 adaptor ^a	-	gactcgagtcgacatcg(T)17

FW = forward; RV = reverse, ^areference