Infection, Genetics and Evolution 41 (2016) 135–141

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Phylogenetic analysis of canine distemper virus in South America clade 1 reveals unique molecular signatures of the local epidemic

Cristine D.B. Fischer [a,b], Tiago Gräf [a,*], Nilo Ikuta [a,b,c], Fernanda K.M. Lehmann [a], Daniel T. Passos [e], Aline Makiejczuk [a,b], Marcos A.T. Silveira Jr [a,b], André S.K. Fonseca [c], Cláudio W. Canal [d], Vagner R. Lunge [a,b,c]

[a] Laboratório de Diagnóstico Molecular, Programa de Pós-Graduação em Biologia Celular e Molecular Aplicada à Saúde, Universidade Luterana do Brasil, Av. Farroupilha, 8001, 92425-900 Canoas, Rio Grande do Sul, Brazil
[b] Hospital Veterinário, Curso de Medicina Veterinária, Universidade Luterana do Brasil, Av. Farroupilha, 8001, 92425-900 Canoas, Rio Grande do Sul, Brazil
[c] Simbios Biotecnologia, Rua Caí, 541, 94940-030 Cachoeirinha, Rio Grande do Sul, Brazil
[d] Laboratório de Virologia, Faculdade de Veterinária, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9090, 91540-000 Porto Alegre, Rio Grande do Sul, Brazil
[e] Department of Biochemistry, Western University, London, ON N6A 5C1, Canada

ARTICLE INFO

Article history:
Received 13 November 2015
Received in revised form 27 March 2016
Accepted 29 March 2016
Available online 6 April 2016

Keywords:
Canine distemper virus
Phylogeny
Molecular epidemiology
South America
Amino acid signature

ABSTRACT

Canine distemper virus (CDV) is a highly contagious pathogen for domestic dogs and several wild carnivore species. In Brazil, natural infection of CDV in dogs is very high due to the large non-vaccinated dog population, a scenario that calls for new studies on the molecular epidemiology. This study investigates the phylodynamics and amino-acid signatures of CDV epidemic in South America by analyzing a large dataset compiled from publicly available sequences and also by collecting new samples from Brazil. A population of 175 dogs with canine distemper (CD) signs was sampled, from which 89 were positive for CDV, generating 42 new CDV sequences. Phylogenetic analysis of the new and publicly available sequences revealed that Brazilian sequences mainly clustered in South America 1 (SA1) clade, which has its origin estimated to the late 1980's. The reconstruction of the demographic history in SA1 clade showed an epidemic expanding until the recent years, doubling in size every nine years. SA1 clade epidemic distinguished from the world CDV epidemic by the emergence of the R580Q strain, a very rare and potentially detrimental substitution in the viral genome. The R580Q substitution was estimated to have happened in one single evolutionary step in the epidemic history in SA1 clade, emerging shortly after introduction to the continent. Moreover, a high prevalence (11.9%) of the Y549H mutation was observed among the domestic dogs sampled here. This finding was associated ($p < 0.05$) with outcome-death and higher frequency in mixed-breed dogs, the later being an indicator of a continuous exchange of CDV strains circulating among wild carnivores and domestic dogs. The results reported here highlight the diversity of the worldwide CDV epidemic and reveal local features that can be valuable for combating the disease.

1. Introduction

Canine distemper (CD) is an important viral infectious disease affecting many wild carnivores and domestic dogs. The manifestations of CD vary from mild to severe signs and frequently can lead to the death of the host. Among dogs, CD mortality rate is only smaller than that of rabies (Deem et al., 2000). The etiological agent, canine distemper virus (CDV), belongs to the genus *Morbillivirus*, family Paramyxoviridae, and contains genetic information in a negative sense RNA strand that encodes six main proteins, the hemagglutinin (H) and fusion (F) glycoproteins being responsible for virion cell attachment and fusion process, respectively (Sawatsky and von Messling, 2010). H protein also plays an important role in the host-specific immunity and due to external pressures on the immune system, displays the highest genetic variability in the CDV genome (Iwatsuki et al., 2000).

The H gene diversity is useful to investigate the molecular epidemiology of this virus. CDV clades have been identified worldwide, most of them following a geographical pattern of distribution. Currently, 14 clusters have been reported: America 1 and 2; Arctic-like, Asia 1–4; Europe 1/South America 1 (EU1/SA1); European wildlife; Rockborn-like; South America 2 (SA2) and 3 (SA3); and Africa 1 and 2 (Iwatsuki et al., 1997; Mochizuki et al., 1999; Calderon et al., 2007; Panzera et al., 2012; Sarute et al., 2014; Espinal et al., 2014; Zhao et al., 2010; Bi et al., 2015; Ke et al., 2015). Among them, the America 1 clade, mainly composed of sequences from the US, was showed to be the origin of the current worldwide CDV epidemic and to play a central role in the initial dissemination of the virus to other regions (Panzera et al., 2015).

Abbreviations: CD, canine distemper; CDV, canine distemper virus; H, hemagglutinin; SA1, South America 1 (clade); EU1, Europe 1 (clade); MLE, marginal likelihood estimation; tMRCA, time of most recent common ancestor; *pp*, posterior probability.

* Corresponding author at: Universidade Luterana do Brasil, Av. Farroupilha, 8001, 92425-900 Canoas, Rio Grande do Sul, Brazil.
E-mail address: akograf@gmail.com (T. Gräf).

http://dx.doi.org/10.1016/j.meegid.2016.03.029
1567-1348/

Polymorphisms found in the H gene have also been proposed to affect host switches, severity of signs and/or viral fitness. Molecular evolutionary analyses identified substitutions at positions 530 and 549 of the signaling lymphocytic activation molecule (SLAM) receptor-binding site of the H gene as being determinant to the host cell tropism (McCarthy et al., 2007). While domestic dogs are infected with 549Y strains, wild canids and non-canids are more frequently, though not exclusively, infected with 549Y and 549H strains, respectively. However a statistical association between viral strain and host species has not yet been demonstrated (Sekulin et al., 2011; Nikolin et al., 2012). By its turn, a changing number of substitutions in position 530 can be observed regardless of the host species (McCarthy et al., 2007; Sekulin et al., 2011; Nikolin et al., 2012). More recently, the H protein substitution R580Q was reported to strongly impact CDV fitness in in vitro experiments (Sattler et al., 2014). Moreover, a series of four compensatory substitutions (N71S, V159I, M195V and M500L) was described to restore the functional efficiency of the virus. In vivo, R580Q is rarely observed and when it does occur is always followed by one or more compensatory substitutions.

Although immunization efforts with attenuated strains have been largely used to prevent CD, outbreaks have been frequently reported in the last decade throughout the world (Benetka et al., 2011; Demeter et al., 2007; Martella et al., 2006; Pardo et al., 2005; Decaro et al., 2004). In Brazil, epidemiological/prevalence studies suggested that CDV is endemic in urban canine populations, showing up to 58% of sero-prevalence (Headley et al., 2012). Previous studies reported that the Brazilian epidemic is almost exclusively composed of strains clustering in EU1/SA1 clade (Budaszewski et al., 2014; Negrão et al., 2013), however a better characterization of the molecular epidemic in Brazil is needed.

In the current study we investigated the phylodynamic history of CDV in Brazil and in EU1/SA1 clade. By applying Bayesian phylogenetics methods in new and readily available sequences we describe the time of epidemic onset, epidemic growth and the evolutionary process of the emergence of amino-acid signatures observed in SA1 clade.

2. Materials and methods

2.1. Animals

This study was approved by the Committee for Use of Animals in Research (CEP-2009-023A) of the Lutheran University of Brazil (Universidade Luterana do Brasil, ULBRA). Dogs with CD clinical signs (anorexia, conjunctivitis, respiratory, gastroenterical and neurological signs) were examined and treated in the normal routine of a Veterinary University Hospital (HV-ULBRA, Canoas, Rio Grande do Sul, Brazil) from March 2010 to June 2011. General data for the animals (age, breed, gender, vaccination history) and clinical signs were recorded during the supportive care or by periodic monitoring.

2.2. CDV detection and sequencing of the H gene

Whole blood, urine, conjunctive and rectal swabs were collected from animals for both immunochromatography and RT-nested-PCR (RT-nqPCR) assays. CDV antigen was detected in the conjunctive swabs by the immunochromatography assay (Anigen CDV Ag test kit, BioNote Inc., Seoul, Korea) (An et al., 2008). CDV RNA was detected in whole blood, urine, rectal and conjunctive swabs by RT-nqPCR as previously described (Fischer et al., 2013).

H gene amplification was carried out using reverse transcription (RT) followed by two PCR reactions (nested-PCR). RT and first strand PCR (RT-PCR) assays were run in a volume of 25 μL using 2 μL of RNA template and 23 μL of reaction mixture (MMLV reverse transcriptase, Taq DNA polymerase, dNTPs, appropriate buffer) containing the primers H2F (5′-AATATGCTAACCGCTATCTC–3′) and H3R (5′-TCAAGGTTTTGA ACGGTTAC–3′) (An et al., 2008). Amplification was performed in a Veriti 96 thermo cycler (Applied Biosystems Inc., Norwalk, CT, USA) with one cycle at 37 °C for 30 min and 35 cycles at 94 °C for 30 s, 55 °C for 30 c and 72 °C for 1 min. Nested PCR was carried out in a volume of 25 μL using 1 μL of DNA template and 24 μL of reaction mixture containing primers CDVF10 (5′-TATCATGACRGYARTGGTTC–3′) and CDVR10 (5′-AATYYTCRAYACTGGWTGTG–3′) (Hashimoto et al., 2001). Thermal cycling was performed in the same equipment with one cycle at 94 °C for 3 min, 30 cycles at 94 °C for 20 s, 55 °C for 40 s and 72 °C for 1 min. PCR products were visualized by electrophoresis on polyacrylamide gels stained with silver nitrate.

Sequencing reactions were carried out in a volume of 25 μL using a sequencing PCR mixture (Big Dye Terminator Cycle Sequencing, Applied Biosystems, Norwalk, USA) and the same primers of the nested PCR. Chromatograms were assembled and visually assessed using SeqMan software, LaserGene package (DNAStar, Madison, WI, USA). Sequence data was deposited in Genbank under the accession numbers KC879042–KC879083.

2.3. Sequence dataset compilation

In order to better characterize the CDV clades circulating in Brazil and the temporal history of the epidemic, all available H gene sequences with sampling year, country of origin and host species data were downloaded from GenBank. The dataset was aligned with MAFFT (Katoh and Toh, 2008) and trimmed aiming to keep the highest number of sequences from Brazil. Sequences that were too short or identical to another sequences were removed from the dataset. Recombinant sequences were detected using RPD v4.0 software under default settings and removed from the dataset afterwards (Martin et al., 2010). An initial phylogenetic tree was reconstructed in RAxML (Stamatakis, 2014) and assessed for its temporal signal by using Path-O-gen (http://tree.bio.ed.ac.uk/software/pathogen/). Outlier sequences in a regression of root-to-tip divergence versus sampling time were removed from the downstream analyses.

2.4. Phylodynamics analyses

Time-scaled phylogenetic tree reconstruction was performed using BEAST/BEAGLE software (Drummond et al., 2012; Ayres et al., 2012) through the Cipres Science Gateway (https://www.phylo.org). This software allows for the combination of different clock, substitution and demographic models, demanding an appropriate model test approach. In the current study a marginal likelihood estimation (MLE) (Baele et al., 2012 and Baele et al., 2013) was applied to compare alternative models in a Bayesian framework. Trees were reconstructed using SRD06 substitution model (Shapiro et al., 2006) and the uncorrelated lognormal (ucld) relaxed molecular clock (Drummond et al., 2006), which outperformed alternative models. Also, the non-parametric Bayesian skyline coalescent model was applied (Strimmer and Pybus, 2001).

To investigate the demographic history of CDV in Brazil, phylogenetic analyses of sequences from the Brazilian clade were performed with the same models as described above and testing for the best demographic model by applying MLE. Informative priors extracted from the analysis of the complete dataset were used on the time of most recent common ancestor (tMRCA) of the Brazilian clade tree root. Epidemic doubling time (λ) was calculated by applying growth rate (r) to the formula $\lambda = \ln(2)/r$. Monte Carlo Markov Chains (MCMC) were run sufficiently long to ensure stationary and adequate effective sample size (ESS) for the main parameters. Tracer software (available at: http://beast.bio.ed.ac.uk/Tracer) was used to diagnose MCMC, adjust initial burn-in and to perform skyline demographic reconstruction.