



# Recombination drives the emergence of orf virus diversity: evidence from the first complete genome sequence of an Indian orf virus isolate and comparative genomic analysis

Basanta Pravas Sahu<sup>1</sup> · Prativa Majee<sup>1</sup> · Ravi Raj Singh<sup>1</sup> · Niranjana Sahoo<sup>2</sup> · Debasis Nayak<sup>3</sup>

Received: 27 November 2021 / Accepted: 7 March 2022 / Published online: 11 May 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

## Abstract

Contagious pustular dermatitis is a disease that primarily infects small ruminants and possesses zoonotic potential. It is caused by orf virus (ORFV), a member of the genus *Parapoxvirus*. In this study, we evaluated an ORFV outbreak in goats in Madhya Pradesh, a state in central India, during 2017. The transboundary potential of this virus was evaluated by constructing phylogenetic trees. The complete genome sequence of an ORFV isolate named Ind/MP/17 was found to be 139,807 bp in length with 63.7% GC content and 132 open reading frames (ORFs) flanked by 3,910-bp inverted terminal repeats (ITRs). An investigation into evolutionary parameters such as selection pressure ( $\theta = dN/dS$ ) and nucleotide diversity ( $\pi$ ) demonstrated that ORFV has undergone purifying selection. A total of 40 recombination events were identified, 21 of which were evident in the Ind/MP/17 genome, indicating its ability to generate new variants.

Contagious ecthyma, otherwise known as contagious pustular dermatitis, is a neglected zoonotic disease caused by orf virus (ORFV), a member of the genus *Parapoxvirus*. It primarily affects sheep, goats, and wild animals and can be transmitted to humans through direct or indirect contact [1, 2]. The virus is extremely contagious and can exist within the animal's wool and faeces for several years [2]. The infection presents as intensified skin wounds with exacerbated blisters throughout the buccal cavity, leading to weight loss and anorexia. The morbidity rate often reaches 100%, resulting in emaciation in adults and young animals, which negatively affects the herd economy [2]. Animal handlers are prone to zoonotic infection, often manifested by the

formation of pustules that can cause severe pain on the hands and can spread to other regions, such as the genitals and face [3]. The ORFV genome consists of double-stranded DNA (dsDNA) and contains almost 130 distinct genes. Genes in the central region are relatively conserved and are involved in mature virion formation and virus replication. In contrast, genes in the terminal regions are more variable and are often associated with virulence and immune modulation [4]. Despite its global distribution, only 14 complete genome sequences are available thus far. The absence of a complete genome sequence of an Indian isolate makes it difficult to perform genetic analysis and thus hinders further functional studies. Therefore, we performed molecular detection of ORFV isolates prevalent within the black Bengal goat breed of central India, followed by a complete genome sequence analysis using a next-generation sequencing (NGS) platform, and we utilized comparative genomics approaches to determine phylogenetic relationships and perform evolutionary and recombination analysis.

The study area included the Dhar district of Madhya Pradesh, a state in central India (75.30E, 22.59 N). During 2017, 10 clinical samples were collected from infected goats aged between one and 11 months that showed typical skin lesions on their lips. The scab samples were stored at  $-80^{\circ}\text{C}$  until further analysis. Total genomic DNA was isolated from the skin tissue according to the protocol described by Sarker et al., using a DNeasy Blood and Tissue Purification

Handling Editor: William G. Dundon.

✉ Debasis Nayak  
debasis@iiserb.ac.in  
https://nayakdn.wixsite.com/virology

- <sup>1</sup> Department of Biological Sciences and Biomedical Engineering, Indian Institute of Technology Indore, Indore, MP 453331, India
- <sup>2</sup> Centre for Wildlife Health, College of Veterinary Sciences and Animal Husbandry, Bhubaneswar 751030, India
- <sup>3</sup> Department of Biological Science, Indian Institute of Science Education and Research (IISER) Bhopal, Bhopal Bypass Road, Bhauri, Madhya Pradesh 462066, India

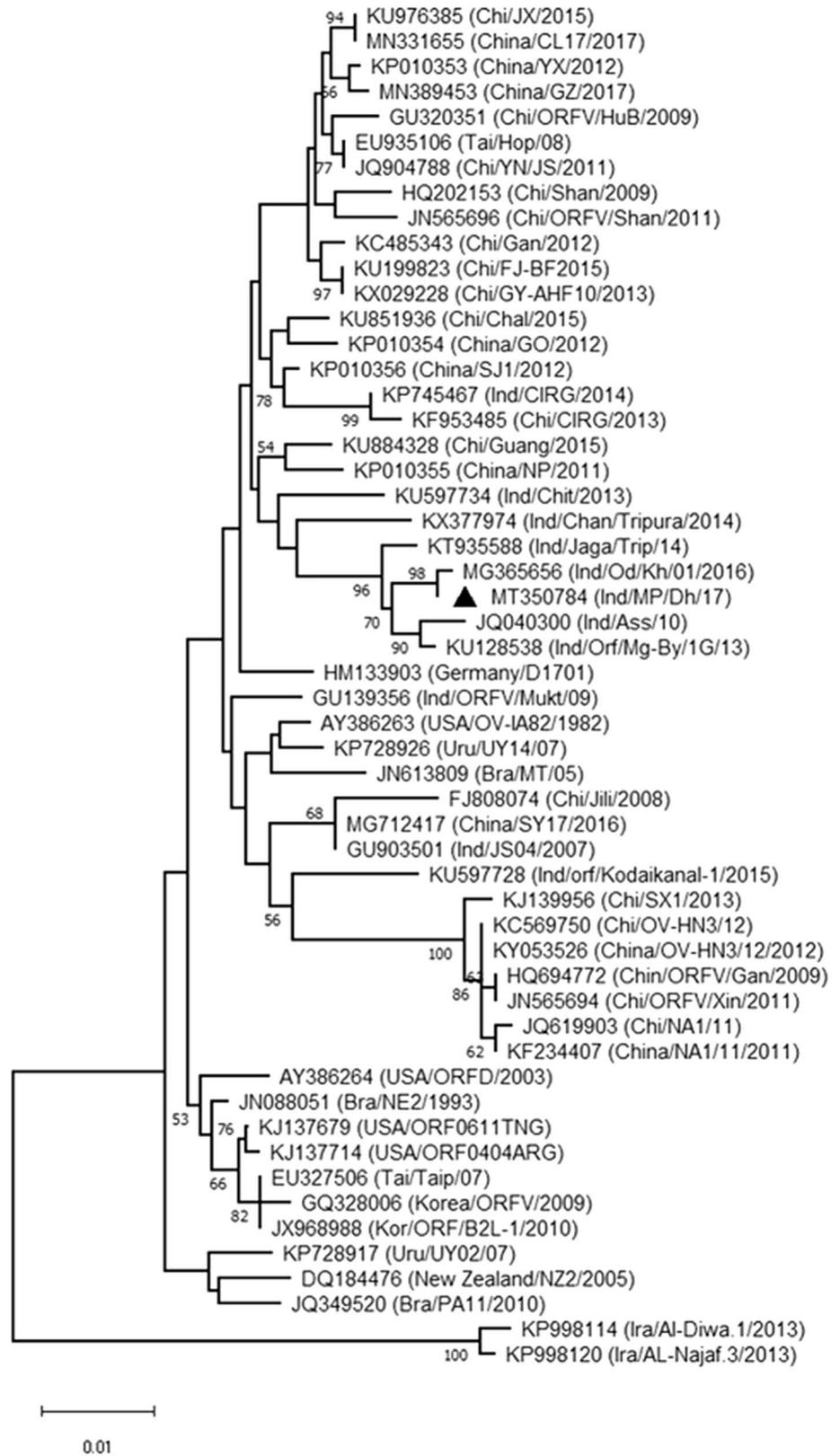
Kit (QIAGEN, Germany) [5]. The presence of the virus was confirmed by PCR, utilizing four sets of primers targeting ORFV011, ORFV020, ORFV059, and ORFV108, commonly known as B2L, E3L, F1L, and A32L, respectively (Supplementary Table S1) [6]. The PCR-amplified products were purified using a MiniElute Gel Extraction Kit (QIAGEN, Germany) and sent to a sequencing service for Sanger sequencing. Subsequently, gene-specific phylogenetic trees were constructed to infer the genetic relatedness and transboundary potential of the present isolate compared to other isolates reported in India and other regions of the world. Phylogenetic trees were constructed by the maximum-likelihood (ML) method, using the general time-reversible (GTR) substitution model and 1,000 bootstrap replicates in MEGA 6.0 (Fig. 1, Supplementary Figs. S1-S3). Nucleotide BLAST and phylogenetic analysis based on these four genes confirmed that the isolate from this study was 99%-100% identical to other isolates with sequences in the GenBank database. At the global level, the highest similarity was observed with the isolates China/NP/2011, China/FJ-YT2015, China/HuB13/2013, and USA/ORFD/2003. At the country level, the present isolate shared the highest nucleotide sequence similarity with an ORFV isolate from Odisha, a state in eastern India. The global comparison was consistent with a previous study showing its closeness to Chinese ORFV isolates [6], indicating the transboundary potential of this virus to spread to neighbouring states or countries. We speculate that trade and animal movement are the likely cause of the observed epidemiological linkage between isolates from the two distinct geographical areas.

A phylogenetic tree based on concatenated sequences indicated that the present isolate, Ind/MP/17, is closely related to the isolate China/GO/2012 (KP010354) (Supplementary Fig. S4). Thus, China/GO/2012 was used as a reference genome for mapping and mutation analysis. Viral DNA was isolated from a clinical sample, and sequenced using a NextSeq 500 NGS platform [7, 8]. The resulting genome sequence, with a total length of 139,807 bp, was submitted to the GenBank database under accession number MT332357. Like those of other PPV isolates, the genome of Ind/MP/17 has a high G+C content (63.7%). The inverted terminal repeats (ITRs) are 3,910 bp in length and cover the region of ORFV001 and ORFV134. Each ITR was found to contain a terminal *Bam*HI site. Telomere resolution motifs were composed of TAAAT, followed by a spacer sequence, ACCCGACC, and (T)<sub>6</sub> residues, which formed the terminal hairpin loop (Fig. 2). Using the NCBI ORF Finder tool and NCBI BLAST, we identified 132 ORFs for a distinct set of genes. By comparing our sequence with that of the reference genome using the BioEdit and ExPaSy tools, we found 488 unique mutations in the current isolate (Supplementary Table S2). The largest numbers of synonymous amino acid substitutions were observed in RNA

helicase NPH-II, RNA polymerase subunit RPO147, and the virion core protein P4a precursor. We observed the largest number of nonsynonymous substitutions within immune regulatory genes, including the NF-kappa pathway inhibitor and ankyrin/F-box protein. These mutations might be responsible for maintaining the heterogeneity and virulence of this pathogen. Using DnaSP to compare all 14 available complete genome sequences of ORFV, the nucleotide diversity ( $\pi$ ) and haplotype diversity (Hd) were estimated to be 0.02815 and 1.000, respectively. Selection pressure analysis ( $\theta = dN/dS$ ; 0.02911) revealed that ORFV had undergone purifying selection. Tajima's D test of neutrality resulted in a significant negative value (-0.14928), suggesting that this virus might be undergoing a period of evolutionary expansion. A similar pattern of  $\theta$  values was obtained using partial gene sequences of recently studied avipoxviruses and ORFV, with values ranging from 0.065 to 0.200 [9, 10]. These results were consistent with the observed dN and dS values and suggested that selection pressure had altered the rate of evolution. Perfect mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeat microsatellites, as well as compound microsatellites, were identified using IMEx software [11] with the following parameters: type of repeat, perfect; repeat size, all; minimum repeat number, 6, 3, 3, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta- and hexanucleotide repeats, respectively. For **??**compound simple sequence repeat**??** (cSSR) identification, the distance between two **??**simple sequence repeats**??** (SSRs) (dMAX) was kept at 10 nucleotides. Our study revealed 1,108 SSRs and 94 cSSRs scattered throughout the ORFV genome. The ORFV genome is rich in dinucleotide repeats (76.5%), followed by trinucleotide (18.14%) and mononucleotide repeats (5.14%). Hexanucleotide microsatellites occurred infrequently and constituted only 0.18% of the total microsatellites (Supplementary Fig. S5). The distribution of classified repeats suggested that dinucleotide GC/CG repeats are prevalent within the ORFV genome, which is consistent with what has been reported for other DNA viruses, such as human papillomavirus [12]. Dinucleotide repeats can adopt a Z confirmation or other alternative secondary DNA structures to facilitate recombination activity [13]. By monitoring a single mononucleotide repeat, Houg et al. revealed the transmission dynamics of a human adenovirus during an epidemic [14]. Therefore, these microsatellites could be used as a tool for epidemiological and evolutionary studies of ORFV.

We retrieved the 14 available complete ORFV genome sequences from the GenBank database, along with those of two pseudocowpox virus (PCPV) isolates and one bovine papular stomatitis virus (BPSV) isolate, belonging to the genus *Parapoxvirus* (PPV), and two monkeypox virus (MPV) isolates, belonging to the genus *Orthopoxvirus*, to construct a phylogenetic tree. Our analysis revealed that all of the ORFV isolates were more closely related to the

**Fig. 1** Phylogenetic analysis based on ORFV011 sequences of ORFV. The phylogenetic tree was constructed by the maximum-likelihood method with the GTR model, using MEGA 6.0 software. The numbers at the branching points indicate the bootstrap support calculated for 1,000 replicates. The black triangle indicates the isolate from the present study.





host-specific clades with 100% bootstrap support [15]. Later, Velazquez-Salinas et al. observed another distinct clade of ORFVs isolated from both sheep and goats in Mexico from 2007 to 2011 [10]. Interestingly, this clade included an isolate from sheep (Ger/D1701) for which a complete genome sequence was available [10], allowing us to verify the status of our isolate within the clade by comparing all available complete genome sequences of ORFV. The present isolate shows a close phylogenetic relationship to members of the goat-specific clade and to the isolates Chi/GO/2012 and USA/ORFD/1982, with 63–100% bootstrap support. However, the isolate Ger/D1701 did not appear in the host-specific clade and was associated with other clades of goat isolates, supporting the results of a study carried out by Velazquez-Salinas et al. [10] (Fig. 3). To investigate the source of genetic variation among all of the complete ORFV genomes from diverse geographical areas, we looked for evidence of recombination using the RDP, GENECONV, Bootscan, MaxChi, Chimaera, Siscan, PhylPro, LARD, and 3Seq methods in the RDP4 program [16]. Forty potential recombination events with significant *P*-values were detected across the complete ORFV genome (Supplementary Table S3). Viruses undergo genetic recombination to form new variants in the population by deleting many of their nonessential genes or acquiring new host genes [17]. The genome sequencing results have shown that recombination plays a vital role in the evolution of viruses, including vaccinia virus and variola virus [18, 19]. The 40 potential recombination events identified in Ind/MP/17 could have actively contributed more than 50% of the genome. Thus, the Ind/MP/17 isolate has the potential to evolve via recombination and act as a major or minor parent to form new variants.

In conclusion, we report here the complete genome sequence of an ORFV isolate circulating in central India. Through in-depth comparative genomic analysis, we identified recombination events that were possibly involved in ORFV evolution and the generation of novel variants. This information will be useful for further understanding ORFV biology, epidemiology, and vaccine development.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00705-022-05443-5>.

**Acknowledgements** BPS is thankful to the University Grant Commission (UGC, Govt. India) for providing a Ph.D. fellowship.

**Author contributions** BPS processed the samples, did the molecular analysis, and wrote the manuscript. PM, RS, and NS constructed the phylogenetic tree and figures. DN designed the study, executed it, and wrote the manuscript.

**Availability of data and materials** Partial gene sequences of ORFV011, ORFV020, ORFV059, and ORFV108 were submitted to the GenBank database under accession nos. MT350784, MT350785, MT350786, and MT350787, respectively. The complete genome sequence of isolate

Ind/MP/17 was submitted to the GenBank database under the accession number MT332357.

## Declarations

**Conflict of interest** The authors declare no conflict of interest.

**Live vertebrate experiment statement** Field veterinarians collected clinical samples as a routine practice from symptomatic goats. The guidelines (clause 3.2.2) set by the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA), Ministry of Environment and Forest, Govt. of India, were followed during the sample collection procedure.

## References

- Guo J, Rasmussen J, Wünschmann A, de La Concha-Bermejillo A (2004) Genetic characterization of orf viruses isolated from various ruminant species of a zoo. *Vet Microbiol* 99:81–92
- Haig DM, Mercer AA (1998) Ovine diseases. *Orf. Vet Res* 29:311–326
- Duchateau NC, Aerts O, Lambert J (2013) Autoinoculation with Orf virus (ecthyma contagiosum). *Int J Dermatol* 53:60–62
- Cottone R, Büttner M, Bauer B, Henkel M, Hettich E, Rziha H-J (1998) Analysis of genomic rearrangement and subsequent gene deletion of the attenuated Orf virus strain D1701. *Virus Res* 56:53–67
- Sarker S, Roberts HK, Tidd N, Ault S, Ladmore G, Peters A, Forwood JK, Helbig K, Raidal SRJSR (2017) Molecular and microscopic characterization of a novel Eastern grey Kangaroo virus genome directly from a clinical sample. *Sci Rep* 7:1–13
- Venkatesan G, De A, Arya S, Kumar A, Muthuchelvan D, Deb-nath BC et al (2018) Molecular evidence and phylogenetic analysis of orf virus isolates from outbreaks in Tripura state of North-East India. *Virus Dis* 29:216–220
- Bennett M, Tu S-L, Upton C, McArtor C, Gillett A, Laird T, O’Dea MJVR (2017) Complete genomic characterisation of two novel poxviruses (WKPV and EKPV) from western and eastern grey kangaroos. *Virus Res* 242:106–121
- Günther T, Haas L, Alawi M, Wohlsein P, Marks J, Grundhoff A, Becher P, Fischer NJSR (2017) Recovery of the first full-length genome sequence of a parapoxvirus directly from a clinical sample. *Sci Rep* 7:1–8
- Le Loc’h G, Bertagnoli S, Ducatez MF (2015) Time scale evolution of avipoxviruses. *Infect Genet Evol* 35:75–81
- Velazquez-Salinas L, Ramirez-Medina E, Bracht AJ, Hole K, Brito BP, Gladue DP, Carrillo CJI, Genetics, and Evolution (2018) Phylodynamics of parapoxvirus genus in Mexico (2007–2011). *Infect Genet Evol* 65:12–14
- Mudunuri SB, Nagarajaram HA (2007) IMEx: imperfect microsatellite extractor. *Bioinformatics* 23:1181–1187
- Singh AK, Alam CM, Sharfuddin C, Ali S (2014) Frequency and distribution of simple and compound microsatellites in forty-eight Human papillomavirus (HPV) genomes. *Infect Genet Evol* 24:92–98
- Treco D, Arnheim N (1986) The evolutionarily conserved repetitive sequence d(TG.AC)<sub>n</sub> promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. *Mol Cell Biol* 6:3934–3947
- Houng HS, Lott L, Gong H, Kuschner RA, Lynch JA, Metzgar D (2009) Adenovirus microsatellite reveals dynamics of

- transmission during a recent epidemic of human adenovirus serotype 14 infection. *J Clin Microbiol* 47:2243–2248
15. Chi X, Zeng X, Li W, Hao W, Li M, Huang X, Huang Y, Rock DL, Luo S, Wang S (2015) Genome analysis of orf virus isolates from goats in the Fujian Province of southern China. *Front Microbiol* 6:1135
  16. Martin DP, Murrell B, Golden M, Khoosal A, Muhire BJVE (2015) RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1:1
  17. Fleming SB, Lyttle DJ, Sullivan JT, Mercer AA, Robinson AJJ-JOGV (1995) Genomic analysis of a transposition-deletion variant of orf virus reveals a 3.3 kbp region of non-essential DNA. *J Gen Virol* 76:2969–2978
  18. Smithson C, Purdy A, Verster AJ, Upton CJPO (2014) Prediction of steps in the evolution of variola virus host range. *PLoS One* 9:e91520
  19. Odom MR, Hendrickson RC, Lefkowitz EJJVR (2009) Poxvirus protein evolution: family wide assessment of possible horizontal gene transfer events. *Virus Res* 144:233–249

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.