

Recombination Drives Emergence of Orf Virus Diversity: Evidence From the First Complete Genome of Indian Orf Virus and Comparative Genomic Analysis

Debasis Nayak (✉ debasis@iiserb.ac.in)

Indian Institute of Science Education and Research Bhopal <https://orcid.org/0000-0002-4762-4553>

Basanta Sahu

Indian Institute of Technology Indore

Prativa Majee

Indian Institute of Technology Indore

Ravi Singh

Indian Institute of Technology Indore

Niranjan Sahoo

Odisha University of Agriculture and Technology

Research Article

Keywords: ORFV, Phylogenetics, Evolution, NGS, Recombination

Posted Date: December 8th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1119090/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Contagious pustular dermatitis is a disease that primarily infects small ruminants and has the zoonotic potential evoked by a Parapoxvirus, Orf virus (ORFV). This study evaluated an ORFV outbreak in goats that arose in Madhya Pradesh, a state of central India, during 2017 by constructing phylogenetic trees and unveiling its transboundary potential. Thereafter, the complete genome of an ORFV strain named Ind/MP has revealed the presence of 139,807bp nucleotide sequences, GC content 63.7%, 132 open reading frames (ORFs) circumscribed by inverted terminal repeats (ITRs) of 3,910bp. Evolutionary parameters such as selection pressure ($\theta=dN/dS$), nucleotide diversity (π), etc., demonstrate the ORFV exhibit purifying selection. A total of forty recombination events were observed, out of which Ind/MP strains were engaged in twenty-one recombination events indicating this strain can recombine for the generation of new variants.

Main Text

Orf or contagious ecthyma is otherwise known as contagious pustular dermatitis, is a neglected zoonotic disease caused by the Orf virus (ORFV), a member of the genus *parapoxvirus*. It primarily affects the sheep and goats, occasionally other ruminants and wild animals, as well as animal handlers, veterinarians, and para veterinarians [1, 2]. The virus is extremely contagious and can exist within the animals' wool and feces for several years [2]. The infection is exhibited with the presence of intensified skin wounds with exacerbated blisters throughout the buccal cavity, often leading to weight loss and anorexia. The morbidity rate often reaches up to 100%, resulting in emaciation in adults, and kids, thereby negatively affecting the herding economy [2]. Animal handlers are prone to the zoonotic potential and are often manifested in the form of unbearable pustules on hands that can expand to other body parts such as genitals and the face of affected persons [3]. The ORFV genome consists of double-stranded DNA (dsDNA), accommodating almost 130 distinctive genes. Genes in the central region are relatively more conserved and involved in mature virion formation and virus replication. In contrast, genes in the terminal regions are more variable and are often attributed to virulence and immune modulation [4]. Despite its global distribution, only fourteen complete genome information is available so far. The absence of a complete genome sequence of the Indian isolates makes it difficult to comprehend genetic analysis and thus hinders further functional studies. Therefore we performed molecular detection of ORFV isolates prevalent in central India infecting the black Bengal goat breed followed by complete genome sequence analysis for the first time through next-generation sequencing (NGS) platform and utilized comparative genomics approaches to decipher phylogenetic relationship, evolutionary and recombination analysis.

The study area is in the Dhar district of Madhya Pradesh, a central Indian state (75.30E, 22.59 N). Samples (n=10) were collected from naturally occurring infected goats aged between one to eleven months, showing typical Orf skin lesions on their lips during 2017. Scab samples collected were subsequently stored at $-80\text{ }^{\circ}\text{C}$ for virus isolation and further analysis. Total genomic DNA was isolated from the skin tissue according to the protocol described by Sarker et al. 2017 using the DNeasy Blood and tissue purification Kits (QIAGEN, Germany) [5]. The viral presence was confirmed by PCR utilizing

four sets of primers targeting ORFV011, ORFV020, ORFV059, and ORFV108, commonly known as B2L, E3L, F1L, and A32L (Supplementary Table 1). The PCR amplified DNA was purified using the MiniElute gel extraction kit (QIAGEN, Germany) and sent for Sanger sequencing. Subsequently, gene-specific phylogenetic trees were constructed to infer the genetic relationship and transboundary potential of the circulating strain by comparing within the country and with other global isolates using a general-time-reversible (GTR) substitution model for the maximum likelihood (ML) phylogeny with 1,000 bootstrap values using MEGA 6.0 (Figure 1A-2D). Nucleotide BLAST and phylogenetic analysis based on these four genes confirmed that the present isolate has 99%-100%. At the global level, the maximum similarity was observed with China/NP/2011, China/FJ-YT2015, China/HuB13/2013, and USA/ORFD/2003 isolates. The sequence and phylogenetic analysis revealed 99-100% nucleotide similarities with the ORFV isolate of Odisha, a state of eastern India. The global comparison was consistent with the previous study showing its closeness with the Chinese ORFV strains [6]. This indicates the transboundary potential of the present virus isolates to spread to the neighboring states or countries. We speculate that trade exchange and transport can be a potential route to establish the epidemiological linkage between the strains isolated from the two distinct geographical areas.

The concatenated phylogenetic tree analysis indicated that the present isolate Ind/MP/2017 (Ind/MP) is closely related to the China/GO/2012 (Chi/GO) (KP010354) isolate; thus, it was further considered as the reference genome for mapping and mutations analysis (Supplementary Figure 1). Meantime, virus isolation was attempted bypassing clinical samples in African green monkey kidney (Vero) cells and primary lamb testicle cell line [7]. However, the virus could not be recovered until the sixth blind passage. So, we moved ahead with the NGS experiment by isolating viral DNA directly from the clinical samples through NextSeq 500 NGS platform [8, 9]. The total length of the assembled genome exhibited 139,807bp in length, and the assigned NCBI accession number is MT332357. Like other PPV genomes, the genome possessed a high (63.7%) G+C content. The Inverted Terminal Repeats (ITRs), which spanned throughout ORFV001 and ORFV134 having a total length of 3,910bp. Each ITR is composed of a terminal *Bam*HI site. Telomere resolution motifs are composed of TAAAT, followed by a spacer sequence, ACCCGACC, and six T residues, which form the terminal hairpin loop (Figure 2). Using NCBI's ORF Finder tool and NCBI's BLAST, we obtained 132 ORFs for a distinct set of genes. In comparison to the reference genome Chi/GO by the help of utilizing the BioEdit and ExPaSy tools, our analysis showed nearly 488 unique mutations in the current isolate (Supplementary Table 2). These observed mutations led to both synonymous and non-synonymous amino acid substitutions. The highest number of synonymous and non-synonymous amino acid substitutions was recorded in RNA helicase NPH-II, RNA-polymerase subunit RPO147, virion core protein P4a precursor, and EEV maturation protein, Poly(A)-polymerase catalytic subunit PAPL, NF-kappa pathway inhibitor, DNA-binding protein, Ankyrin/F-box protein, respectively. We observed that the highest number of non-synonymous substitutions within immune regulatory genes such as NF-kappa pathway inhibitor, and Ankyrin/F-box protein, etc. These mutations might be responsible for maintaining the heterogeneity and mimicking the virulence of this pathogen. By taking into account all the fourteen available complete genome sequences of ORFV and utilizing DnaSP, the nucleotide diversity (π) and haplotype diversity (Hd) were observed to be 0.02815 and 1.000, respectively. Selection pressure analysis

($\theta = dN/dS$), with a value of 0.02911 revealed that ORFV resides under 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 θ value was obtained from recently studied avipoxvirus and ORFV partial genes, which ranged 0.065-0.200 [10, 11]. This confirmed dN and dS impel selection pressure to alter the rate of evolution. Identification of perfect mono, di, tri, tetra, penta, hexa as well as compound microsatellites was made by IMEx software [12] with the parameter: type of repeat: perfect; repeat size: all; minimum repeat number: 6, 3, 3, 3, 3, 3 for mono, di, tri, tetra, penta and hexanucleotide repeats, respectively, and the distance between two SSRs (dMAX) was ten nucleotides. Our study revealed 1,108 and 94 numbers of SSRs and cSSR scattered throughout the ORFV genome. The ORFV genome is rich mostly with dinucleotide repeats (76.5%), followed by trinucleotide (18.14%), and mononucleotide repeats (5.14%). The hexanucleotide microsatellites most scarcely presented and constituted only 0.18% of the ORFV genome (Supplementary Figure 2). The distribution of classified repeats suggests that dinucleotide GC/CG is more prevalent in most of the ORFV genomes, similar to other DNA viruses like Human papilloma [13]. Di-nucleotide repeat could form Z confirmation or other alternative secondary DNA to facilitate the recombination activity [14]. By using a single mononucleotide repeat, Houg et. al., could follow the transmission dynamics of a human adenovirus during an epidemic [15]. Therefore, these microsatellites could potentially be used as a powerful tool for epidemiological and evolutionary studies for ORFV.

We retrieved fourteen available ORFV complete genome sequences for the GenBank database along with five sequences of *Parapoxvirus* (PPV) and Orthopoxvirus, consisting of two *Pseudocowpox virus* (PCPV), one *Bovine papular stomatitis virus* (BPSV), and two *Monkeypox virus* (MPV), respectively to create a phylogenetic tree. It revealed that six ORFV strains originating in goats and eight strains belonging to sheep formed two separate clades except for Ger/D1701 with 61-100% bootstrap support. The present ORFV strain showed a close relationship with Chi/GO and USA/ORFD isolates. Our analysis also showed that all ORFVs were more closely related to PPVs (PCPVs and BPSV) than to Orthopoxvirus (MPVs) (Figure 3). However, our analysis, in comparison to the previous study, showed an increase in the heterogeneity and inability to maintain the perfection of the host-specific clade. This kind of ambiguity was also observed during phylodynamic analysis of the parapoxvirus genus in Mexico (2007–2011), where Ger/D1701, with several other isolates, exhibited a separate clade rather than host-specific clade [11]. To understand the source of genetic variation among all the ORFV complete genomes, we looked for evidence of recombination using the RDP, GENECONV, Bootscan, MaxChi, Chimaera, Siscan, PhylPro, LARD, and 3Seq methods contained in the RDP4 program [16]. We observed a total of 40 potential recombination events with significant *P*-values detected across the ORFV genomes (Supplementary Table 3). Viruses undergo genetic recombination to form new variants in the population by deleting many of their non-essential genes or by acquiring new host genes [17]. Viral genome sequencing elucidates that recombination plays a vital role in understanding human and animal pathogens' evolution, including *Vaccinia* and *Variola viruses* [18, 19]. However, in this study, we identified forty potential recombination events where Ind/MP actively participated in more than 50% of events by forming recombinant as well as

major and minor parents. Thus, the Ind/MP strain has the potential to evolve via recombination and can act as a major or minor parent to form new variants.

In conclusion, we report the complete genome of circulating ORFV isolate from central India. Subsequently, by in-depth analysis through a comparative genomic approach, we propose recombination events that may be responsible for ORFV evolution and generation of new strain types. We hope that the current genomic information would be greatly useful for further understanding of ORFV biology, epidemiology, and research carried in front of diagnosis and vaccine development.

Declarations

Acknowledgments

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

Author Contributions Statement

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.

Live Vertebrate Experiment Statement

Field veterinarians collected clinical samples as a routine practice from symptomatic goats. The guideline (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 was followed during the sample collection procedure.

Conflict of Interest Statement

The authors declare no conflict of interest.

Availability of data and material

The partial gene information ORFV011, ORFV020, ORFV059, and ORFV108 were submitted to GenBank under accession no. MT350784, MT350785, MT350786 and MT350787, respectively. The comprehensive complete genome was submitted with the accession number MT332357 to the GenBank.

References

1. Guo, J., Rasmussen, J., Wünschmann, A., and de La Concha-Bermejillo, A.J.V.m. (2004). Genetic characterization of orf viruses isolated from various ruminant species of a zoo. *99*, 81-92.
2. Haig, D.M., and Mercer, A.A. (1998). Ovine diseases. *Orf. Veterinary research* *29*, 311-326.

3. Duchateau, N.C., Aerts, O., and Lambert, J. (2013). Autoinoculation with Orf virus (ecthyma contagiosum). *International Journal of Dermatology* 53, e60-e62.
4. Cottone, R., Büttner, M., Bauer, B., Henkel, M., Hettich, E., and Rziha, H.-J. (1998). Analysis of genomic rearrangement and subsequent gene deletion of the attenuated Orf virus strain D1701. *Virus Research* 56, 53-67.
5. Sarker, S., Roberts, H.K., Tidd, N., Ault, S., Ladmore, G., Peters, A., Forwood, J.K., Helbig, K., and Raidal, S.R.J.S.r. (2017). Molecular and microscopic characterization of a novel Eastern grey kangaropox virus genome directly from a clinical sample. 7, 1-13.
6. Kumar, N., Wadhwa, A., Chaubey, K.K., Singh, S.V., Gupta, S., Sharma, S., Sharma, D.K., Singh, M.K., and Mishra, A.J.V.G. (2014). Isolation and phylogenetic analysis of an orf virus from sheep in Makhdoom, India. 48, 312-319.
7. Mazur, C., and Machado, R.D. (1989). Detection of contagious pustular dermatitis virus of goats in a severe outbreak. *Veterinary Record* 125, 419.
8. Bennett, M., Tu, S.-L., Upton, C., McArtor, C., Gillett, A., Laird, T., and O'Dea, M.J.V.r. (2017). Complete genomic characterisation of two novel poxviruses (WKPV and EKPV) from western and eastern grey kangaroos. 242, 106-121.
9. Günther, T., Haas, L., Alawi, M., Wohlsein, P., Marks, J., Grundhoff, A., Becher, P., and Fischer, N.J.S.r. (2017). Recovery of the first full-length genome sequence of a parapoxvirus directly from a clinical sample. 7, 1-8.
10. Le Loc'h, G., Bertagnoli, S., and Ducatez, M.F. (2015). Time scale evolution of avipoxviruses. *Infection, Genetics and Evolution* 35, 75-81.
11. Velazquez-Salinas, L., Ramirez-Medina, E., Bracht, A.J., Hole, K., Brito, B.P., Gladue, D.P., Carrillo, C.J.I., *Infection, Genetics and Evolution* (2018). Phylodynamics of parapoxvirus genus in Mexico (2007–2011). 65, 12-14.
12. Mudunuri, S.B., and Nagarajaram, H.A. (2007). IMEx: Imperfect Microsatellite Extractor. *Bioinformatics* 23, 1181-1187.
13. Singh, A.K., Alam, C.M., Sharfuddin, C., and Ali, S. (2014). Frequency and distribution of simple and compound microsatellites in forty-eight Human papillomavirus (HPV) genomes. *Infection, Genetics and Evolution* 24, 92-98.
14. Treco, D., and Arnheim, N. (1986). The evolutionarily conserved repetitive sequence d(TG.AC)_n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. *Molecular and cellular biology* 6, 3934-3947.
15. Houn, H.S., Lott, L., Gong, H., Kuschner, R.A., Lynch, J.A., and 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.
16. Martin, D.P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B.J.V.e. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. 1.

17. Fleming, S.B., Lyttle, D.J., Sullivan, J.T., Mercer, A.A., and Robinson, A.J.J.J.o.G.V. (1995). Genomic analysis of a transposition-deletion variant of orf virus reveals a 3.3 kbp region of non-essential DNA. *76*, 2969-2978.
18. Smithson, C., Purdy, A., Verster, A.J., and Upton, C.J.P.o. (2014). Prediction of steps in the evolution of variola virus host range. *9*.
19. Odom, M.R., Hendrickson, R.C., and Lefkowitz, E.J.J.V.r. (2009). Poxvirus protein evolution: family wide assessment of possible horizontal gene transfer events. *144*, 233-249.

Figures

Figure 1

Phylogenetic trees constructed based on the four ORFV genes, namely ORFV011, ORFV020, ORFV059, and ORFV108. The phylogenetic relationship was constructed by the GTR model of the maximum-likelihood method using MEGA 6.0 software. Numbers at the branching points indicate the bootstrap support calculated for 1,000 replicates (A) ORFV011, (B) ORFV020, (C) ORFV059, and (D) ORFV108 based on partial nucleotide sequences. The black triangle represents the accession number of the present study and was mentioned as MT350784 (B2L), MT350785 (E3L) MT350786 (F1L), and MT350787 (A32L),

Figure 2

Analysis of ITR-ends of ORFV. (A) The left end (5') sequence alignment of 14 complete ORFV genomes consists of a terminal BamHI site (green box) along with telomere resolution motifs (red box) (ATTTTTT-N(8)-TAAAT). (B) The right end (3') sequence alignment of 14 complete ORFV genomes consists of a terminal BamHI site (green box) along with telomere resolution motifs (red box) (ATTTTTTN(8)-TAAAT).

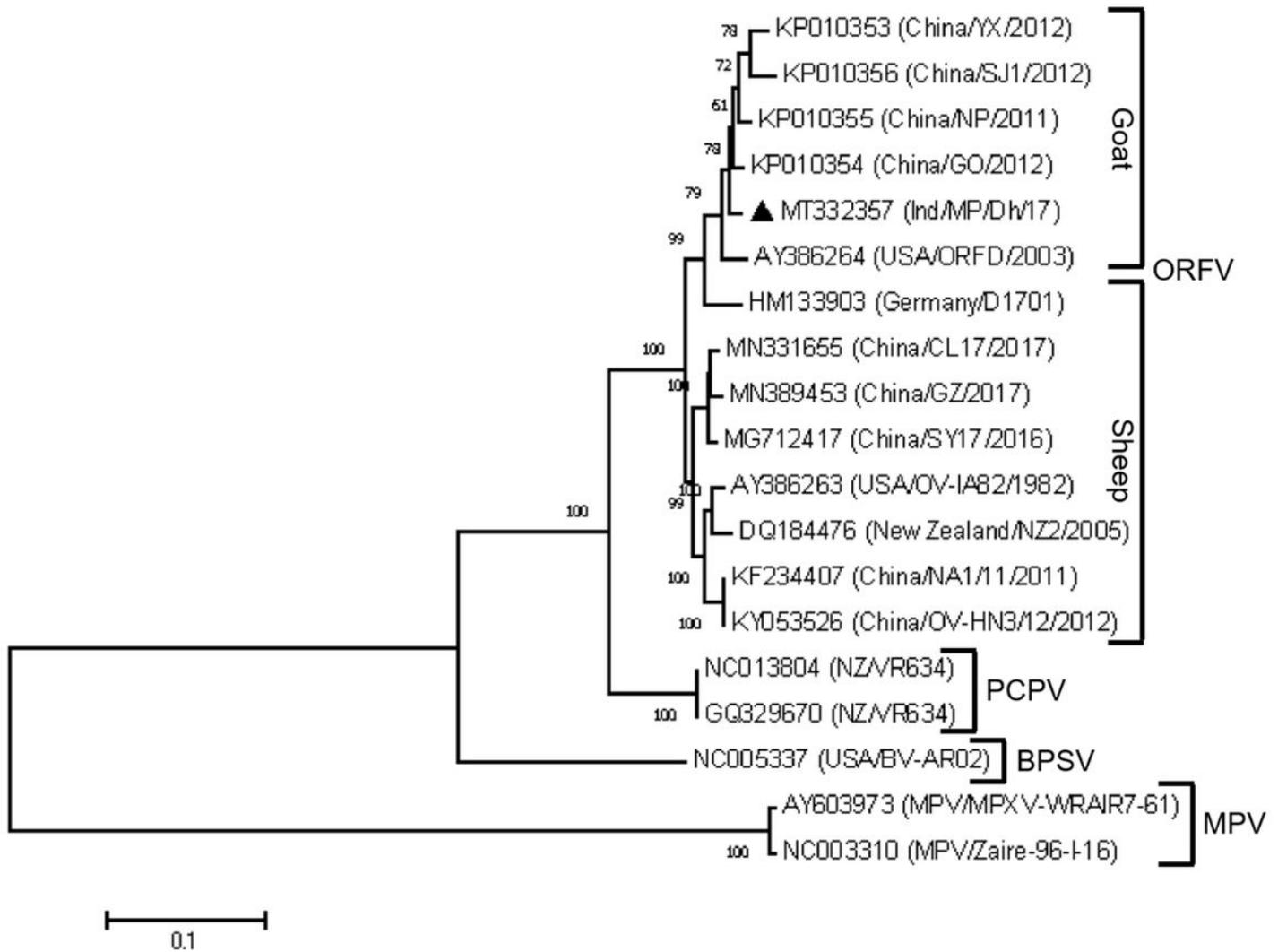


Figure 3

Comparison of PPVs by constructing a Phylogenetic tree. Nineteen complete genome sequences, including the terminal repetitions, were aligned to construct a phylogenetic tree with bootstrap value 1000 using the GTR model of maximum likelihood method. The black triangle represents the ORFV isolates of the present investigation showing its relationship with fourteen global strains.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ElectronicSupplementaryfileORFVcompletegenome.docx](#)