A survey of host range genes in poxvirus genomes

Kirsten A. Bratke [a], Aoife McLysaght [a], Stefan Rothenburg [b],*

[a] *Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland*
[b] *Laboratory for Host-Specific Virology, Division of Biology, Kansas State University, KS 66506, USA*

ARTICLE INFO

ABSTRACT

Poxviruses are widespread pathogens, which display extremely different host ranges. Whereas some poxviruses, including variola virus, display narrow host ranges, others such as cowpox viruses naturally infect a wide range of mammals. The molecular basis for differences in host range are poorly understood but apparently depend on the successful manipulation of the host antiviral response. Some poxvirus genes have been shown to confer host tropism in experimental settings and are thus called host range factors. Identified host range genes include vaccinia virus K1L, K3L, E3L, B5R, C7L and SPI-1, cowpox virus CP77/CHOhr, ectromelia virus p28 and 022, and myxoma virus T2, T4, T5, 11L, 13L, 062R and 063R. These genes encode for ankyrin repeat-containing proteins, tumor necrosis factor receptor II homologs, apoptosis inhibitor T4-related proteins, Bcl-2-related proteins, pyrin domain-containing proteins, cellular serine protease inhibitors (serpins), short complement-like repeats containing proteins, KilA-N/RING domain-containing proteins, as well as inhibitors of the double-stranded RNA-activated protein kinase PKR. We conducted a systematic survey for the presence of known host range genes and closely related family members in poxvirus genomes, classified them into subgroups based on their phylogenetic relationship and correlated their presence with the poxvirus phylogeny. Common themes in the evolution of poxvirus host range genes are lineage-specific duplications and multiple independent inactivation events. Our analyses yield new insights into the evolution of poxvirus host range genes. Implications of our findings for poxvirus host range and virulence are discussed.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Poxviruses are widespread pathogens, which as a family infect a wide range of animals and have been responsible for mass mortalities of animals and humans (Fenner, 2000). They include variola virus (VARV; abbreviations are shown in Table 1), the causative agent of smallpox, which is believed to have caused more fatalities throughout human history than all other infectious diseases combined (McFadden, 2005), and vaccinia virus, which has been successfully used as a vaccine in the eradication of smallpox, one of the biggest achievements in the history of medicine (Fenner, 2000). Despite this success, poxviruses continue to be constant threats for animals, including humans. In fact more than 10 poxviruses are known to infect humans, including monkeypox virus (MPXV), an emerging zoonotic disease, causing high fatality rates (Essbauer et al., 2010; Lewis-Jones, 2004). Strikingly, poxviruses display extremely different host ranges. Whereas some poxviruses such as VARV and camelpox virus (CMLV) have very narrow host ranges, others such as MPXV and cowpox virus (CPXV) naturally infect a broad spectrum of vertebrates (Fenner, 2000; McFadden, 2005).

Poxviruses are DNA viruses that exclusively replicate in the cytoplasm of infected cells (Lefkowitz et al., 2006). They are grouped into two subfamilies: *Entomopoxvirinae*, which infect insects, and *Chordopoxvirinae*, which infect vertebrates. Currently, nine different genera of *Chordopoxvirinae* are recognized: avipoxviruses, molluscipoxviruses, parapoxviruses, orthopoxviruses, yatapoxviruses, leporipoxviruses, capripoxviruses, cervidpoxviruses and suipoxviruses. In the past, classification and grouping into different poxvirus genera and species has relied on pathological findings and host range as well as structural and antigenic properties (Fenner, 2000; Lefkowitz et al., 2006). Recent advances in the genomic characterization of poxviruses lay the groundwork for more precise phylogenetic characterization (Bratke and McLysaght, 2008; Lefkowitz et al., 2006; McLysaght et al., 2003). The complete genomes of about 25–29 different poxvirus species, depending on species definition, have been determined to date. For some species multiple strains were sequenced (http://www.poxvirus.org, Lefkowitz et al., 2005). Poxvirus genomes possess large genomes, ranging from 130 to 365 kb and often contain more than 200 genes (Lefkowitz et al., 2006). A core of 33 gene families are found as single copy genes in all sequenced genomes, and have been used to reconstruct the poxvirus phylogeny (Bratke and McLysaght, 2008). Approximately 90 genes are commonly found in *Chordopoxvirinae*, with the remaining genes being unique

Table 1
Names and abbreviations of poxviruses used in this study.

Name	Abbreviation	Genus
Amsacta moorei entomopoxvirus	AMV	Entomopoxvirus B
Melanoplus sanguinipes entomopoxvirus	MSV	Unassigned
Canarypox virus	CNPV	Avipoxvirus
Fowlpox virus	FPV	Avipoxvirus
Crocodilepox virus	CRV	Unclassified
Molluscum contagiosum virus subtype 1	MOCV	Molluscipoxvirus
Bovine papular stomatitis virus	BPSV	Parapoxvirus
Orf virus	ORFV	Parapoxvirus
Camelpox virus	CMLV	Orthopoxvirus
Cowpox virus – Brighton Red	CPXV-BR	Orthopoxvirus
Cowpox virus – GRI-90	CPXV-GRI	Orthopoxvirus
Ectromelia virus	ECTV	Orthopoxvirus
Horsepox virus	HSPV	Orthopoxvirus
Monkeypox virus	MPXV	Orthopoxvirus
Rabbitpox virus	RPXV	Orthopoxvirus
Taterapox virus	TATV	Orthopoxvirus
Vaccinia virus	VACV	Orthopoxvirus
Variola virus	VARV	Orthopoxvirus
Yaba-like disease virus	YLDV	Yatapoxvirus
Yaba monkey tumor virus	YMTV	Yatapoxvirus
Tanapox virus	TPV	Yatapoxvirus
Myxoma virus	MYXV	Leporipoxvirus
Rabbit fibroma virus	RFV	Leporipoxvirus
Deerpox virus W-848-83	DPV-W83	Cervidpoxvirus
Deerpox virus W-1170-84	DPV-W84	Cervidpoxvirus
Goatpox virus	GTPV	Capripoxvirus
Lumpy skin disease virus	LSDV	Capripoxvirus
Sheeppox virus	SPPV	Capripoxvirus
Swinepox virus	SWPV	Suipoxvirus

to some lineages or representing lineage-specific gene duplications (Lefkowitz et al., 2006).

Poxviruses are excellent models for studying genome evolution. They display a relatively low rate of accumulation of point mutations (Li et al., 2007), whereas gene duplications, losses, gain by horizontal gene transfer (HGT) and recombination between different species occur frequently (Bratke and McLysaght, 2008; Hughes and Friedman, 2005; McLysaght et al., 2003). These events are important for adapting to their hosts and for subverting the host antiviral response. The terminal regions of poxvirus genomes are especially prone to recombination, important for the assimilation of new genes, and contain identical inverted terminal repeats (ITR), which range in size between 0.1 and 13 kb (Lefkowitz et al., 2006).

Poxviruses contain many genes that are non-essential for viral replication in cell culture but important for modulating and circumventing the host-response and thus influence the course of poxvirus infection and pathology. Such genes are designated as virulence genes (McFadden, 2005). A portion of these genes has been identified that are important for viral replication in only a subset of tissue culture cells, which were derived from different tissues or animal species and are commonly referred to as host range genes or factors and are thought to be responsible for poxvirus-specific differences in tropism and host range (McFadden, 2005; Werden et al., 2008). Unlike many other vertebrate viruses, poxviruses do not rely on specific receptors to enter cells, but harness molecules that are ubiquitously present in many different cell-types in many animal species. Whether a virus can replicate after entry into a cell, depends on the successful manipulation of the cellular antiviral response (McFadden, 2005). Approximately 12 different host range genes or gene families have been identified to-date (reviewed in Werden et al., 2008).

Although some progress has been made in recent years in understanding poxvirus host range, the molecular basis for it is only poorly understood. We conducted systematic analyses on the presence or absence of known poxvirus host range genes. We found evidence for lineage-specific gene inactivation, deletions, duplications and recombination events. Furthermore, we analyzed the evolution and phylogeny of an extended family of host range genes, which often included known virulence factors that have not yet been identified as host range genes. The data presented here fill an important gap in our understanding and knowledge of poxvirus host range, have important implications for disease control and identify threats that might emerge through the assimilation of novel host range factors.

2. Materials and methods

2.1. Construction of poxvirus phylogenetic trees

We used a concatenated sequence alignment containing protein sequences of 28 genes (Supplementary Table 1) that are present in single copy (sc) in 104 completely sequenced poxviruses that were used in this study and performed phylogenetic analyses. ClustalW (Thompson et al., 1994) was used to infer a neighbor-joining (NJ) phylogenetic tree based on this alignment. While it provides the backbone for the topology presented in Fig. 1, the large number of genomes and small number of families used to determine their relationship leave some branches where the bootstrap values are low or where the inferred topology is in conflict with the accepted poxvirus phylogeny. We have used concatenated alignments of larger numbers of protein families present in sc in subsets of the 104 genomes to resolve these branches. The tree shown in Fig. 1 is therefore a composite tree. Different colored bootstraps are derived from different subset trees as follows.

The inferred monophyly of molluscum contagiosum virus (MOCV) and crocodilepox virus (CRV) is an artifact according to the author of the CRV genome paper (Afonso et al., 2006). When using 36 sc families present in CRV, MOCV, BPSV, FPV-I and AMV to draw a neighbor-joining tree, the topology observed is that in Fig. 1, represented by orange bootstrap values.

The relationship of genera in clade II poxviruses was clarified using a maximum-likelihood (ML) tree based on 95 families from all clade II genomes plus orthopoxvirus outgroup, and this corresponds to the green bootstraps in Fig. 1.

Black bootstraps are derived from a NJ tree of orthopoxvirus representatives plus outgroup based on 83 shared sc gene families.

In agreement with previous analyses containing fewer MPXV genomes (Likos et al., 2005), MPXVs form two clades, which we term clade I (includes MXPV strains Z96, CON, Z79) and clade II (includes strains SL, LIB, WRA, COP, USA39 and USA44) strains. As there was weak bootstrap support for a branch within clade II MPXV, the within-genus topology was resolved with a NJ tree of MPXV plus outgroup based on 128 sc families, and the bootstrap values are noted in gray in Fig. 1.

The 49 VARV isolates that are fully sequenced and were analyzed by (Esposito et al., 2006) are difficult to separate in the poxvirus tree due to their high sequence similarity. A ML tree based on 114 sc families present in all VARV genomes, CMLV, Taterapox virus (TATV) and the ECTV outgroup agrees to a large extent with the topology favored by (Esposito et al., 2006). A NJ tree of the same genomes is presented by purple bootstraps in Fig. 1.

The poorly-supported topology of 11 VACV isolates and closely related genomes horsepox (HSPV) and rabbitpox (RPXV) is resolved by a NJ tree of 100 families present in sc in these genomes and the closest outgroup CPXV-GRI-90. This tree is the source of the red bootstrap values.