Journal of General Virology (2005), **86**, 2969–2977 DOI 10.1099/vir.0.81331-0

Complete coding sequences of the rabbitpox virus genome

G. Li,[1]† N. Chen,[1]†‡ R. L. Roper,[1]§ Z. Feng,[2] A. Hunter,[1] M. Danila,[1] E. J. Lefkowitz,[3] R. M. L. Buller[2] and C. Upton[1]

[1]Department of Biochemistry and Microbiology, University of Victoria, Ring Road, Petch Bldg, Rm 150, Victoria, BC, Canada V8W 3P6

[2]Department of Molecular Microbiology and Immunology, St Louis University School of Medicine, St Louis, MO 63104, USA

[3]Department of Microbiology, University of Alabama (Birmingham), Birmingham, AL 35294-2170, USA

Correspondence
C. Upton
cupton@uvic.ca

Received 11 July 2005
Accepted 12 August 2005

Rabbitpox virus (RPXV) is highly virulent for rabbits and it has long been suspected to be a close relative of vaccinia virus. To explore these questions, the complete coding region of the rabbitpox virus genome was sequenced to permit comparison with sequenced strains of vaccinia virus and other orthopoxviruses. The genome of RPXV strain Utrecht (RPXV-UTR) is 197 731 nucleotides long, excluding the terminal hairpin structures at each end of the genome. The RPXV-UTR genome has 66·5 % A+T content, 184 putative functional genes and 12 fragmented ORF regions that are intact in other orthopoxviruses. The sequence of the RPXV-UTR genome reveals that two RPXV-UTR genes have orthologues in variola virus (VARV; the causative agent of smallpox), but not in vaccinia virus (VACV) strains. These genes are a zinc RING finger protein gene (RPXV-UTR-008) and an ankyrin repeat family protein gene (RPXV-UTR-180). A third gene, encoding a chemokine-binding protein (RPXV-UTR-001/184), is complete in VARV but functional only in some VACV strains. Examination of the evolutionary relationship between RPXV and other orthopoxviruses was carried out using the central 143 kb DNA sequence conserved among all completely sequenced orthopoxviruses and also the protein sequences of 49 gene products present in all completely sequenced chordopoxviruses. The results of these analyses both confirm that RPXV-UTR is most closely related to VACV and suggest that RPXV has not evolved directly from any of the sequenced VACV strains, since RPXV contains a 719 bp region not previously identified in any VACV.

INTRODUCTION

Poxviruses are large double-stranded DNA viruses; they have genomes ranging from 130 to 380 kb and replicate in the cytoplasm of cells of their hosts, which include both vertebrates and insects (Esposito & Fenner, 2001; Moss, 2001). Two of the most important members of the *Poxviridae* are variola virus (VARV), the causative agent of smallpox, and the closely related vaccinia virus (VACV) that was successfully used as the live vaccine for smallpox. VARV and VACV belong to the genus *Orthopoxvirus* of the *Chordopoxvirinae*. While VARV no longer circulates naturally in the human population, it remains a serious bioterrorism concern as does monkeypox virus (MPXV), which occasionally produces serious zoonotic infections of humans (Ligon, 2004; Reynolds *et al.*, 2004). The origins of VACV are complicated and confused because of the many sources of vaccine material and methods of propagation and distribution during the centuries of its use (Bazin, 2000). A comparison of the recently sequenced horsepox virus genome with other orthopoxviruses (OPVs) found the highest degree of similarity with genomes of several VACV strains, leading to the suggestion that VACV is possibly a natural horse pathogen (D. L. Rock, personal communication). It is also interesting that VACV strains closely related to those employed as vaccines in the smallpox eradication programme have maintained natural infection cycles in

†These authors contributed equally to this paper.

‡Present address: Genelux Corporation, 3030 Bunker Hill Street, Suite 310, San Diego, CA 92109, USA.

§Present address: Department of Microbiology & Immunology, Brody School of Medicine, East Carolina University, Greenville, NC 27834, USA.

The GenBank/EMBL/DDBJ accession number of the RPXV-UTR genome sequence is AY484669.

A full list of RPXV-UTR annotated ORFs is available as a supplementary table in JGV Online.

cattle in Brazil (Damaso *et al.*, 2000) and water buffalo in India (Kolhapure *et al.*, 1997).

Rabbitpox virus (RPXV) was first isolated and reported in 1932 from one of a series of epidemics of a highly lethal, airborne infection of laboratory rabbits at the Rockefeller Institute in New York (Greene, 1933). At this time, the clinical, pathological and virological features of the disease were thoroughly studied by Greene (1934a, b). In 1941 in Utrecht, the Netherlands, another outbreak occurred. An isolate from this outbreak was designated RPXV strain Utrecht (RPXV-UTR) (Fenner, 1994; Esposito & Fenner, 2001). VACV infections of the skin of rabbits by intradermal scarification (dermal VACV) or the respiratory tract with aerosol infectious doses as high as $1{\cdot}3\times10^4$ p.f.u. failed to produce a fatal disease; however, similar aerosol infection with RPXV-UTR produces almost uniform fatalities with a dose of 15 p.f.u. per rabbit (Westwood *et al.*, 1966). Virological studies determined that dermal VACV and RPXV-UTR replicated to similar titres in the rabbit lung, but only RPXV-UTR showed consistent and significant titres in internal organs such as spleen, liver, kidney, gonads and brain, which suggests that dermal VACV is unable to disseminate from the lung (Westwood *et al.*, 1966). Since RPXV is spread via the respiratory tract, it may prove to be a good non-primate model for the study of smallpox. For many years it was suspected that RPXV may have been derived from VACV, as it arose at Rockefeller University in laboratories where work on VACV in rabbits was ongoing, and its genome is similar to VACV by restriction endonuclease analysis (Wittek *et al.*, 1977); however, the molecular basis of RPXV's enhanced virulence for the rabbit compared with VACV was unknown.

In this study we report the DNA sequence of the entire protein-coding region of the genome of RPXV-UTR and the identification of three RPXV-UTR genes that are conserved in VARV but not in sequenced VACV strains and thus may contribute to the enhanced virulence of RPXV and VARV over VACV. In addition, phylogenetic analysis of these genomes shows that VACV and RPXV are so closely related that they may be considered strains of the same virus.

Finally, knowledge of the protein-coding regions of the RPXV genome will facilitate the use of this virus in testing the efficacy of OPV antivirals and vaccines. Currently most antivirals and vaccines are tested in mouse-based OPV challenge models, and additional small animal (non-primate) models will be necessary for US Food and Drug Administration approval of smallpox prophylactics and therapeutics as described in the Animal Rule (FDA, 2002). Since RPXV-UTR encodes several virulence factors that are also present in VARV but absent from VACV, RPXV-UTR infections of rabbits may prove to be such a model.

METHODS

Cells and viruses. RPXV-UTR, originally obtained from the American Type Culture Collection (ATCC, catalogue number VR-157), was generously provided by Dr Dick Moyer of the University of Florida. PK 15 cells (CCL-33) were obtained from the ATCC and propagated in Dulbecco's modified Eagle's medium (DMEM; Bio-Whittaker) containing 10 % FETALCLONE II (HyClone Laboratories). Virus infectivity was estimated as described previously (Chen *et al.*, 2000).

Purification of RPXV genomic DNA, genome sequencing and assembly. Standard procedures were used and have been described previously (Chen *et al.*, 2003). With the exception of the terminal hairpin loops, the entire RPXV genome was divided into 19 overlapping fragments of an average size of 11 kb. Each fragment was amplified from genomic DNA and sequenced with a bank of sequencing primers. The final DNA consensus sequence represented on average a 3·7-fold redundancy, with each nucleotide being covered by at least one high-quality sequence read in each direction.

In order to improve reliability in the DNA sequence the following precautions were taken: (i) DNA templates pooled from multiple PCRs (typically 8–18) were used in sequencing reactions; (ii) a relatively large amount of template was used in the original PCRs; (iii) uncloned PCR products were used as sequencing templates; the Expand Long Template PCR System was chosen because it employs a high-fidelity DNA polymerase. We have sequenced two VACV isolates, separated by only a few passages in tissue culture, with this method and found no differences between the genomes (data not shown).

Annotation. An open reading frame (ORF) was defined as a continuous stretch of DNA that translated into a polypeptide initiated by a methionine residue and extending for at least 60 amino acids prior to a termination codon (TGA, TAA or TAG). ARTEMIS (Mural, 2000), the Poxvirus Orthologous Clusters database (POCs) (Ehlers *et al.*, 2002; Upton *et al.*, 2003), BLASTP (Altschul *et al.*, 1997) and GeneStar software (Windows) (Burland, 2000) were used to detect and annotate ORFs. In addition, for some ORFs, BLASTN, TBLASTN and BLASTP searches were carried out at the NCBI website (Altschul *et al.*, 1990). EMBOSS Showorf software (Rice *et al.*, 2000) and nucleotide-amino acid alignment program (NAP) (Huang & Zhang, 1996) were used on a Linux platform to compare the nucleotide sequence of the fragmented ORF regions of RPXV against the corresponding longest protein sequence encoded by orthologous chordopoxvirus (ChPV) genes. Based on the results of Showorf and NAP, we constructed a physical map of the genome that includes the fragmented ORF regions I–XII. Viral genome organizer (VGO) software (Upton *et al.*, 2000) was used to analyse the position and arrangement of genes.

Phylogeny analysis. Phylogenetic analysis was carried out using a genomic nucleotide sequence alignment of the conserved central region of all OPV genomic sequences currently available in GenBank (see Sequence availability section below). The alignment extended from base 20352 to base 163510 of the RPXV-UTR genome. This alignment starts with ORF 13 and extends just past ORF 159 of RPXV-UTR and corresponds to the VACV-Copenhagen (VACV-COP) genes C7L to A51R (which is fragmented in RPXV-UTR). Sequence alignments were generated using a combination of the programs MAVID and Multi-LAGAN (Brudno *et al.*, 2003; Bray & Pachter, 2004). The final computational alignment was then hand-edited extensively to optimize the alignment.

Phylogenetic inferences were generated using both maximum-parsimony and Bayesian inference methods. Maximum-parsimony trees were constructed using PAUP* version 4.0b10 (Swofford, 2003), while MrBayes version 3.1 (Ronquist & Huelsenbeck, 2003) was used for Bayesian inference methods. Maximum-parsimony trees were constructed using the branch-and-bound search method and employed 1000 replicates for bootstrap resampling analysis. Bayesian inference using Markov chain Monte Carlo methods used a general time

reversible (GTR) model of nucleotide substitution and allowed for gamma-distributed variation across sites with a proportion of invariable sites. Tree analysis was performed using 100 000 generations with a sampling frequency of 100. Standard deviation of split frequencies converged to 0 following 45 000 generations, resulting in one final tree with a probability of 99 %.

Sequence availability. The RPXV-UTR genome sequence has been deposited in GenBank under accession number AY484669 and at http://www.poxvirus.org. The genomes used for comparison are: VACV-COP, M35027 (Goebel *et al.*, 1990); VACV-TT, AF095689 (Q. Jin and others, unpublished); VACV-MVA, U94848 (Antoine *et al.*, 1998); VACV-WR, AY243312.1 (J. J. Esposito and others, unpublished); VARV-Bangladesh-1975 (VARV-BSH), L22579 (Massung *et al.*, 1994); VARV-India-1967 (VARV-IND), X69198 (S. N. Shchelkunov and others, unpublished); VARV-Garcia-1966 (VARV-GAR), Y16780 (S. N. Shchelkunov and others, unpublished); MPXV-Zaire-96-I-16 (MPXV-ZAI), AF380138 (Shchelkunov *et al.*, 2001); MPXV-WRAIR, AY603973 (Chen *et al.*, 2005); ectromelia virus-Moscow (ECTV-MOS), AF012825 (Chen *et al.*, 2003); cowpox virus-Brighton Red (CPXV-BR), AF482758 (D. J. Pickup, unpublished); CPXV-GRI-90, X94355 (S. N. Shchelkunov and others, unpublished); camelpox virus (CMLV)-M-96, AF438165 (Afonso *et al.*, 2002); and CMLV-CMS, AY009089 (Gubser & Smith, 2002).

RESULTS AND DISCUSSION

Overview of the structure of the genome of RPXV-UTR

The genome of RPXV-UTR was assembled into a contiguous sequence of 197 731 nucleotides. This sequence includes all of the predicted protein coding regions, but does not include the hairpin structures located at each end of the genome. The leftmost, confirmed nucleotide of the end corresponding to the *Hin*dIII C fragment of VACV was designated nucleotide 1, as is customary. Like other OPVs, the RPXV-UTR genome is A + T rich (66·5 %). This RPXV-UTR sequence can be subdivided into a central region of 177 687 bp flanked by two inverted terminal repeats (ITRs) of 10 022 bp each (Fig. 1). Each ITR contains two sequence blocks made up of direct repeats (DR1 and DR2 in the left ITR and DR3 and DR4 in the right ITR). DR1/4 are 612 bp in length, including 8·5 copies of a 70 bp tandem direct repeat element, and repeats DR2/3 are 363 bp in length, consisting of 6·5 copies of a 54 bp tandem direct repeat

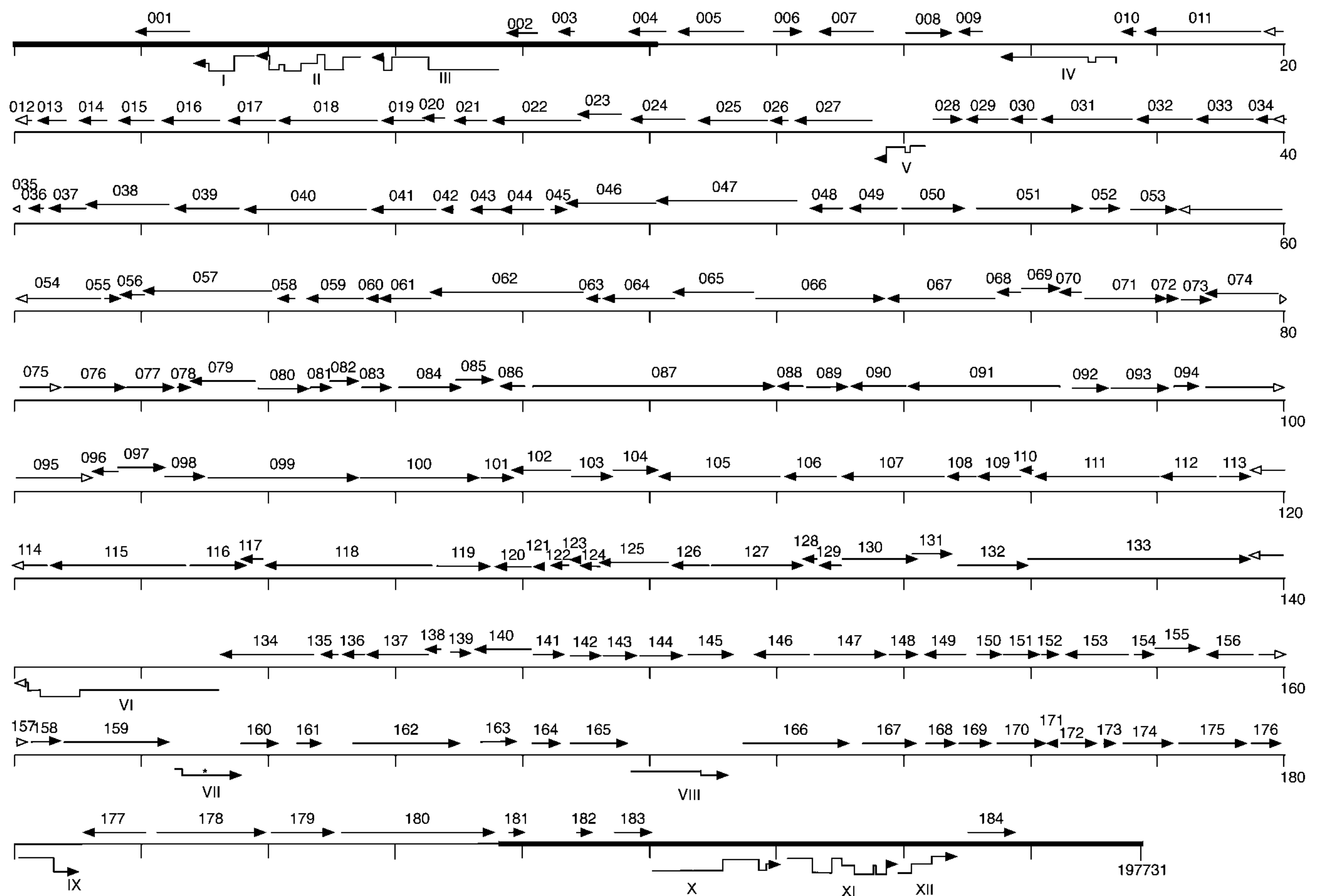

Fig. 1. RPXV-UTR genome: predicted genes are numbered and shown as straight arrows; regions containing fragments of genes present in other OPVs are shown with staggered arrows to represent frame changes and have been given roman numerals. Open arrowheads indicate that an ORF is split over two lines of the diagram. The scale is shown in kb; thickened lines represent the ITRs of the genome: *, Stop codon.

element; these are essentially identical to the repeat sequence elements found in VACV-COP (data not shown). Dot-plot comparisons of the complete RPXV-UTR genome with other OPV genomes indicate that it is made up of a subset of the sequences within the CPXV-BR genome (data not shown).

All ORFs that could encode proteins of at least 60 amino acids were initially annotated as genes. Subsequently, we included two small ORFs, RPXV-UTR-010 and RPXV-UTR-123, due to the existence of highly conserved orthologues in other poxvirus genomes (ORFs CPXV-BR-026 and VACV-COP-A14.5L); we also reclassified a number of predicted genes as 'fragmented ORFs' because they correspond to small regions of larger ORFs identified in other orthopoxviruses and thus may not be functional. Thus, in total, we assigned 184 putative genes predicted to encode proteins ranging in size from 53 to 1286 aa (see Supplementary Table S1 available in JGV Online) and 12 regions that contain fragmented ORFs in RPXV-UTR compared with other OPVs (Table 1). Each ITR contains three complete genes, ORF 001/184, ORF 002/183 and ORF 003/182, as well as three fragmented ORF regions, I/XII, II/XI and III/X (Fig. 1). ORF 004 spans the left ITR junction and is an orthologue of CPXV-BR-218; this ORF is fragmented in VACV-COP (C13L and C14L). Only the C-terminal half of RPXV-UTR-004 is present in the right ITR as RPXV-UTR-181.

Throughout the RPXV genome, as in other poxviruses, these putative genes and ORF fragments of known genes are located on both strands of the genome with very short sequences between them. The majority of genes present within 50 kb of each terminus are transcribed towards that end of the genome, as is the case with most other poxvirus genomes (Fig. 1). Among the 184 predicted genes, 183 genes have annotated orthologues in other OPVs. RPXV-UTR-171, which could encode a 77 aa protein, is the only gene not previously annotated in any fully sequenced poxvirus genome. Despite the significant length of this ORF, it is probably not a functional gene, since there is no classical promoter sequence in the 100 nucleotides upstream of the ORF and the same region in CPXV-BR overlaps the significantly larger CPXV-BR-203 gene that is transcribed in the opposite orientation; CPXV-BR-203 is a member of the myxoma virus M-T4 virulence factor family (Barry *et al.*, 1997; Hnatiuk *et al.*, 1999).

Phylogenetic analyses

Phylogenetic relationships among all available OPV genomes were inferred using two different computational methods to ensure that the final evolutionary tree was not dependent on the method used. Inferences based on maximum-parsimony analysis and Bayesian inference each produced a single tree that showed identical topologies and no significant differences in branch lengths (Fig. 2). While confidence values based on bootstrap analysis of the maximum-parsimony tree showed 100 % confidence for the majority of branch points, confidence was lower for a few of the lineages. In contrast, when assessing tree reliability using Bayesian inference, the final tree showed a probability of 99 % with the posterior probability of all bipartitions (branch points) equal to 1·0, providing a great deal of confidence in the final topology.

Table 1. Regions in the RPXV-UTR genome corresponding to fragments of annotated ORFs in other OPVs

RPXV-UTR				Longest OPV orthologue		Annotation
Region	Start	Stop	Significant ORF fragments (bp)	Gene name (bp)	Function/motif	OPVs with intact gene
I/XII*	2870/193907	3835/194859	192, 369	CPXV-BR-005 (1065)	TNF-α receptor-like	I: CPXV, CMLV, MPXV, ECTV. XII: CPXV, CMLV, MPXV, ECTV, VARV
II/XI*	3899/192162	5570/193833	234, 330, 342, 450	CPXV-BR-006 (1857)	Ankyrin repeats	II: CPXV, CMLV, MPXV, ECTV. XI: CPXV, CMLV, MPXV, ECTV, VARV
III/X*	5687/190073	7605/192045	492, 1158	CPXV-BR-008 (2016)	Ankyrin repeats	III: CPXV, CMLV. X: CPXV
IV	15338	17304	216, 234, 1230	CPXV-BR-025 (2004)	Host range	CPXV, MPXV
V	33549	34339	255, 366	ECTV-MOS-024 (831)	Putative monoglyceride lipase	ECTV, MPXV, CPXV
VI	139433	143226	684, 702, 2178	CPXV-BR-158 (3852)	ATI protein	CPXV, ECTV
VII	162502	163510	255, 363, 381	VACV-COP-A51R (1002)	Unknown	All other sequenced OPVs
VIII	169747	171243	375, 660	CPXV-BR-197 (1515)	Unknown	CMLV, CPXV, ECTV, MPXV
IX	180049	181031	405, 624	ECTV-MOS-163 (984)	IL1β receptor-like	ECTV, CPXV, VACV, MPXV

*Regions present in the ITRs of RPXV-UTR.

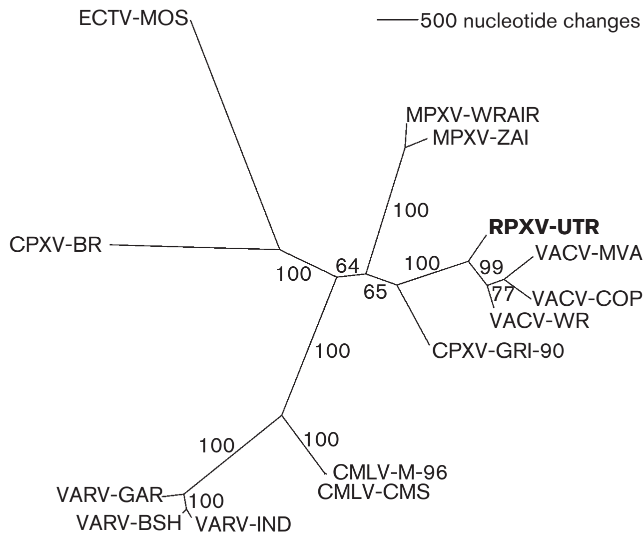

Fig. 2. Phylogeny of OPVs. Predictions are based upon the multiple nucleic acid sequence alignment of the central conserved genomic region of each representative OPV species, strain or isolate. The tree represents the phylogenetic inference generated from a branch-and-bound search using maximum-parsimony as the optimality criterion. Branch lengths are proportional to the number of nucleotide changes. Bootstrap resampling confidence values on 1000 branch-and-bound replicates are displayed for each branch point as a percentage of the total replicates.

This tree places RPXV-UTR together with all other VACV strains on one branch separate from the other OPVs. A similar phylogenetic relationship was obtained for a merged tree created from separate multiple alignments using the predicted protein sequences of the 49 genes that are conserved amongst all completely sequenced ChPVs (data not shown). This tree topology is the same as that previously derived using similar methods for analysis of OPV evolution using the same core region nucleotide sequence multiple alignment (Chen *et al.*, 2005). OPV evolutionary analysis which compared tree topologies using amino acid and nucleotide sequence alignments for different OPV gene sets also displayed the same overall tree topology, but showed slightly different branching arrangements for ECTV and CPXV-BR (Gubser *et al.*, 2004). In fact, this previous work demonstrated that the branching arrangement for ECTV and CPXV-BR varied depending on whether genes located near the left or right ends of these viruses were used in the analysis, suggesting that recombination may have occurred in the evolutionary history of these virus strains, resulting in different tree topologies for different regions of their genomes. The present analysis is consistent with the previously reported branching order for VACV isolates as well as CPXV-GRI-90.

Our analyses support the idea that RPXV-UTR is very closely related to VACV. How RPXV-UTR compares to the RPXV strain that arose at Rockefeller University is unknown, since we have been unable to locate any stocks of the Rockefeller Institute strain of RPXV. The clouded history of VACV together with the fact that it was passaged as a crude mixture in animals at the time RPXVs were isolated leaves the connection between these viruses unclear. However, one hypothesis is that RPXV was present as a natural variant in VACV crude stocks and was selected *in vivo* from this complex mixture after infection of rabbits.

Virulence genes

RPXV-UTR was compared with available sequences of strains of VACV and other OPVs to identify ORFs that may contribute to the enhanced virulence of RPXV-UTR over other VACV strains for rabbits. The VACV-MVA sequence was not used in this analysis as it contains a large number of deletions and grows poorly in mammalian cells (Antoine *et al.*, 1998; Blanchard *et al.*, 1998; Drexler *et al.*, 1998). Furthermore, we have avoided relying heavily on the genome sequence of VACV-TT because our group, and others, has identified a number of errors in this sequence (Upton *et al.*, 2003). This analysis indicates that RPXV-UTR contains three genes that are not present in VACV-COP or VACV-WR; these genes encode a zinc RING finger protein (ORF 008), an ankyrin repeat-containing protein (ORF 180) and a chemokine-binding protein (ORF 001/184).

RPXV-UTR-008 is predicted to encode an orthologue of the ECTV strain Moscow 012 gene product (ECTV-MOS-012; also known as p28), a key ECTV virulence protein (Senkevich *et al.*, 1994). Orthologues are also present in VARV, CPXV, MPXV, CMLV and a number of other ChPVs, but are deleted in VACV-COP and fragmented in VACV-TT (Upton *et al.*, 2003) and VACV-WR. The most prominent feature of the predicted ORF 008 protein is a C-terminal zinc-binding motif, known as the RING finger motif (PROSITE database PDOC00449; http://www.expasy.ch); similar motifs are present in a wide variety of proteins with diverse functions (Lovering *et al.*, 1993; Boddy *et al.*, 1997; Lyngso *et al.*, 2000; Kaiser *et al.*, 2003). This poxvirus protein localizes to the virus factory in the cytoplasm of infected cells (Upton *et al.*, 1994), reduces apoptosis in infected cells (Brick *et al.*, 1998) and has recently been found to have a ubiquitin ligase activity (Huang *et al.*, 2004; Nerenberg *et al.*, 2005). The predicted protein product of RPXV-UTR-008 is 96·4 % identical to ECTV p28, which is not required for virus multiplication in cell culture, but is an important determinant of ECTV pathogenicity. Disruption of this gene in ECTV increases the LD_{50} by several orders of magnitude (Senkevich *et al.*, 1994) and functional p28 protein was found necessary for ECTV replication in some primary murine macrophages (Senkevich *et al.*, 1995). By analogy, if a functional p28 protein is necessary for efficient virus replication in alveolar macrophages and subsequent spread to, and replication in, the draining hilar lymph node following a respiratory infection, this might explain the reduced dermal VACV titres compared with RPXV-UTR in the hilar lymph node and internal organs (Westwood *et al.*, 1966).

The protein product of RPXV-UTR-180, which is predicted to be a 791 aa orthologue of VARV-BSH-B18R, contains three ankyrin repeat motifs (PROSITE database PDOC50088; http://www.expasy.ch). This is an interesting protein because evidence is accumulating that ankyrin repeat motifs mediate protein–protein interaction events, such as those between integral membrane and cytoskeletal proteins (Lambert *et al.*, 1990; Lux *et al.*, 1990). Moreover, poxvirus proteins containing ankyrin repeat motifs are thought to influence virus host range and pathogenesis (Shchelkunov *et al.*, 1993, 1998). Orthologues of RPXV-UTR-180 are present in many of the OPVs but are fragmented in VACV (Fig. 3). A notable exception among the virulent OPVs is ECTV, which lacks an orthologue of RPXV-UTR-180 due to a series of small deletions that shift the reading frame. However, both ECTV and CPXV contain a paralogue of this gene at the left end of the genome that has approximately 46 % amino acid identity; thus it is possible that this ankyrin repeat motif-containing protein may still be important for virulence in RPXV and that ECTV-MOS-005 functions to complement the loss of the RPXV-UTR-180 orthologue in ECTV. The C-terminal region of RPXV-UTR-180 contains some similarity to the F-box motif (PROSITE database PDOC50181; http://www.expasy.ch). This protein motif is believed to play a general role in protein–protein interactions and functions in the association of the Skp1–cullin–F-box protein ligase complexes; it is associated with ubiquitination and degradation of several proteins (Bai *et al.*, 1996). The RPXV-UTR-180 gene is also interesting because it contains 719 nucleotides (from 186989 to 187708; Fig. 3) that are not present anywhere in the genomes of any of the four sequenced VACV strains; this encodes the C terminus of the RPXV-UTR-180 protein. Although an orthologous region is present in other sequenced OPVs, they possess significant minor differences that exclude the possibility that this is a contaminating laboratory sequence arising from the processes of DNA sequencing. This finding has implications for the evolutionary relationship of the OPVs, suggesting that RPXV is not a direct descendant of the known VACV strains without recombination with a second unknown OPV but could have been present in the uncloned population of VACV in use at that time.

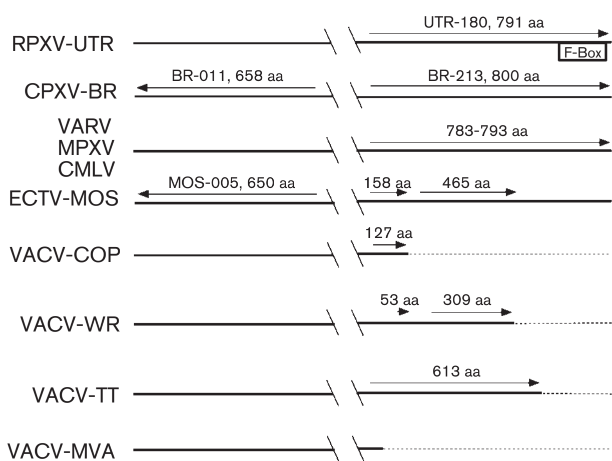

Fig. 3. Organization of RPXV-UTR-180 orthologues, paralogues and gene fragments at the left and right ends of OPV genomes. The omitted central portion of the genomes is indicated by the two slashes. Dotted lines indicate sequences deleted in VACV genomes. Lengths of predicted proteins for potential genes and ORFs are indicated.

RPXV-UTR-001 and RPXV-UTR-184 are ITR genes that encode proteins that belong to the chemokine-binding protein family. Although the promoter and gene are very similar in all OPVs, in several VACVs a small deletion close to the 5′ end of the gene generates a frame-shift mutation that results in the synthesis of a severely truncated protein; this 7·5 kDa gene product gives its name to the well-characterized p7·5 early promoter in VACV-WR (Wittek *et al.*, 1980a, b; Wittek & Moss, 1980). It has been shown that RPXV secretes this chemokine-binding protein and that VACV-WR and several other strains do not (Martinez-Pomares *et al.*, 1995; Alcami *et al.*, 1998b). Unexpectedly, deletion of the RPXV chemokine-binding protein gene did not lead to attenuation in mice or rabbits, but rather to slight enhancement in the onset of illness; it was suggested that this protein might enhance spread of the virus by lessening symptoms early during sickness (Martinez-Pomares *et al.*, 1995). Similarly, insertion of the VACV-Lister gene for this chemokine-binding protein into VACV-WR reduced virulence in mice (Reading *et al.*, 2003). This phenotype is not without precedent; deletion of the IL1-binding protein gene resulted in greater disease symptoms due to increased levels of IL1 (Alcami & Smith, 1992). Thus, it is reasonable to still consider the chemokine-binding protein as a virulence factor.

It is also interesting that RPXV-UTR-134, the orthologue of the OPV structural protein (P4c) that directs intracellular mature virus particles into A-type inclusions, is only full-length in the WR strain of VACV (McKelvey *et al.*, 2002). The general conservation of this gene in a number of OPVs that do not produce a full-length A-type inclusion protein suggests both that the P4c proteins and the partial A-type inclusion proteins may be providing a selectable advantage to these viruses.

Fragmented regions

To be consistent with our previous work (Chen *et al.*, 2000), in annotating the RPXV genome, we intentionally avoided including a number of small ORFs that are clearly fragments of larger genes present in other OPVs. However, we describe 12 regions that clearly correspond to disrupted ORFs and have named them fragmented regions I to XII (Table 1; Fig. 1). Mutations in these regions were mapped by comparing the RPXV DNA sequence to the protein sequence of the longest OPV orthologues using the NAP program. It is very unlikely that these 12 regions generate functional gene products, since the ORFs contained within them are all significantly truncated with respect to their orthologues.

However, further work is required to confirm this loss of function.

Most of the fragmented regions in RPXV-UTR correspond to regions that are also fragmented in other VACV genomes. However, fragment region VII corresponds to a fragmented orthologue of VACV-COP-A51R; the function of this gene product is unknown but it is present in all other OPVs. The mutation that disrupts the orthologue of VACV-COP-A51R was very clear in the sequence trace data (data not shown); since DNA sequencing was performed in two directions using multiple pooled PCR products as the DNA sequencing template, we are confident that this and other mutations that disrupt coding sequences are not from PCR mutations or sequencing errors.

The fragmented haemagglutinin (HA) gene was annotated as RPXV-UTR-163 because it represents approximately the C-terminal two-thirds of the gene; however, this is additionally annotated as 'fragmented' in the VOCs database (http://www.biovirus.org). The RPXV HA gene is frameshifted because of an additional adenine residue after a run of six adenines; the gene was resequenced and analysed using two different sequencing machines and software packages to remove any systematic error. Fragmentation of the HA gene was also confirmed by visualizing the plaque phenotype of RPXV-UTR (data not shown), in which the cells fuse at the edges of the plaques (Ichihashi & Dales, 1971). Thus, although the gene encoding HA is conserved in all other OPVs, the sequencing and experimental data indicate that this isolate of RPXV is indeed HA negative.

In conclusion, we have identified three genes that are present in RPXV-UTR but absent from other VACVs. Each of these gene products has features that associate them with poxvirus virulence; some are better characterized than others. For example, knock-out experiments have clearly shown a role for the RING finger protein in ECTV infections of mice (Senkevich *et al.*, 1994, 1995) and orthologues of the RPXV-UTR-001 protein have been identified as binding host chemokines (Alcami *et al.*, 1998a; Lalani *et al.*, 1998, 1999), but, since deletion of the chemokine gene did not attenuate RPXV in mice or rabbits (Martinez-Pomares *et al.*, 1995), it is more likely that one or both of the other genes are responsible for the enhanced RPXV virulence over VACV in rabbits. Although bioinformatics analysis of the RPXV genome cannot substitute for a thorough biochemical characterization of the contribution that each of these three genes makes to virus virulence, it may be prudent to ensure that all three of these genes are absent from any VACV strains that are engineered for human vaccines or therapeutics. It should also be noted that there are a large number of minor differences between the genomes that could affect virulence. For example, even single nucleotide changes in poxvirus promoters may significantly alter transcription levels and single amino acid changes in proteins can result in relatively major changes in the protein–protein interactions required for a viral protein to bind a specific host cytokine.

ACKNOWLEDGEMENTS

The authors would like to thank Angelika Ehlers, Ryan Brodie, Vasily Tcherepanov and Ross Gibbs for their work in software and database development and Melissa Da Silva for critical review of this manuscript. This work was supported by an NIAID/DARPA grant U01AI48653-02 to R. M. L. B. and C. U., an NIAID/DARPA grant U01AI48706 to E. J. L., an NIH/NIAID contract HHSN266200400036C to E. J. L. and C. U. and by an NSERC Discovery grant to C. U.

REFERENCES

Afonso, C. L., Tulman, E. R., Lu, Z., Zsak, L., Sandybaev, N. T., Kerembekova, U. Z., Zaitsev, V. L., Kutish, G. F. & Rock, D. L. (2002). The genome of camelpox virus. *Virology* **295**, 1–9.

Alcami, A. & Smith, G. L. (1992). A soluble receptor for interleukin-1 beta encoded by vaccinia virus: a novel mechanism of virus modulation of the host response to infection. *Cell* **71**, 153–167.

Alcami, A., Symons, J. A., Khanna, A. & Smith, G. L. (1998a). Poxviruses: capturing cytokines and chemokines. *Semin Virol* **8**, 419–427.

Alcami, A., Symons, J. A., Collins, P. D., Williams, T. J. & Smith, G. L. (1998b). Blockade of chemokine activity by a soluble chemokine binding protein from vaccinia virus. *J Immunol* **160**, 624–633.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.

Antoine, G., Scheiflinger, F., Dorner, F. & Falkner, F. G. (1998). The complete genomic sequence of the modified vaccinia Ankara strain: comparison with other orthopoxviruses. *Virology* **244**, 365–396.

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebl, M., Harper, J. W. & Elledge, S. J. (1996). SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* **86**, 263–274.

Barry, M., Hnatiuk, S., Mossman, K., Lee, S. F., Boshkov, L. & McFadden, G. (1997). The myxoma virus M-T4 gene encodes a novel RDEL-containing protein that is retained within the endoplasmic reticulum and is important for the productive infection of lymphocytes. *Virology* **239**, 360–377.

Bazin, H. (2000). *The Eradication of Smallpox*. London: Academic Press.

Blanchard, T. J., Alcami, A., Andrea, P. & Smith, G. L. (1998). Modified vaccinia virus Ankara undergoes limited replication in human cells and lacks several immunomodulatory proteins: implications for use as a human vaccine. *J Gen Virol* **79**, 1159–1167.

Boddy, M. N., Duprez, E., Borden, K. L. & Freemont, P. S. (1997). Surface residue mutations of the PML RING finger domain alter the formation of nuclear matrix-associated PML bodies. *J Cell Sci* **110**, 2197–2205.

Bray, N. & Pachter, L. (2004). MAVID: constrained ancestral alignment of multiple sequences. *Genome Res* **14**, 693–699.

Brick, D. J., Burke, R. D., Schiff, L. & Upton, C. (1998). Shope fibroma virus RING finger protein N1R binds DNA and inhibits apoptosis. *Virology* **249**, 42–51.

Brudno, M., Do, C. B., Cooper, G. M., Kim, M. F., Davydov, E., Green, E. D., Sidow, A. & Batzoglou, S. (2003). LAGAN and Multi-LAGAN:

efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res* **13**, 721–731.

Burland, T. G. (2000). DNASTAR's Lasergene sequence analysis software. *Methods Mol Biol* **132**, 71–91.

Chen, N., Buller, R. M., Wall, E. M. & Upton, C. (2000). Analysis of host response modifier ORFs of ectromelia virus, the causative agent of mousepox. *Virus Res* **66**, 155–173.

Chen, N., Danila, M. I., Feng, Z., Buller, R. M., Wang, C., Han, X., Lefkowitz, E. J. & Upton, C. (2003). The genomic sequence of ectromelia virus, the causative agent of mousepox. *Virology* **317**, 165–186.

Chen, N., Li, G., Liszewski, M. K. & 13 other authors (2005). Virulence differences between monkeypox virus isolates from West Africa and the Congo basin. *Virology* **340**, 46–63.

Damaso, C. R., Esposito, J. J., Condit, R. C. & Moussatche, N. (2000). An emergent poxvirus from humans and cattle in Rio de Janeiro State: Cantagalo virus may derive from Brazilian smallpox vaccine. *Virology* **277**, 439–449.

Drexler, I., Heller, K., Wahren, B., Erfle, V. & Sutter, G. (1998). Highly attenuated modified vaccinia virus Ankara replicates in baby hamster kidney cells, a potential host for virus propagation, but not in various human transformed and primary cells. *J Gen Virol* **79**, 347–352.

Ehlers, A., Osborne, J., Slack, S., Roper, R. L. & Upton, C. (2002). Poxvirus Orthologous Clusters (POCs). *Bioinformatics* **18**, 1544–1545.

Esposito, J. & Fenner, F. (2001). Poxviruses. In *Fields Virology*, 4th edn, pp. 2885–2921. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

FDA (2002). New drug and biological drug products; evidence needed to demonstrate effectiveness of new drugs when human efficacy studies are not ethical or feasible. 21 CFR parts 314 and 601. In *Federal Register*, vol. 67, no. 105, pp. 37988–37998. Washington, DC: US Food and Drug Administration.

Fenner, F. (1994). Rabbitpox virus. In *Virus Infections of Rodents and Lagomorphs*. Edited by A. D. M. E. Osterhaus. Amsterdam: Elsevier Science.

Goebel, S. J., Johnson, G. P., Perkus, M. E., Davis, S. W., Winslow, J. P. & Paoletti, E. (1990). The complete DNA sequence of vaccinia virus. *Virology* **179**, 247–266.

Greene, H. S. N. (1933). A pandemic of rabbit-pox. *Proc Soc Exp Biol Med* **30**, 892–894.

Greene, H. S. N. (1934a). Rabbit pox. I. Clinical manifestations and cause of disease. *J Exp Med* **60**, 427–440.

Greene, H. S. N. (1934b). Rabbit pox. II. Pathology of the epidemic disease. *J Exp Med* **60**, 441–455.

Gubser, C. & Smith, G. L. (2002). The sequence of camelpox virus shows it is most closely related to variola virus, the cause of smallpox. *J Gen Virol* **83**, 855–872.

Gubser, C., Hue, S., Kellam, P. & Smith, G. L. (2004). Poxvirus genomes: a phylogenetic analysis. *J Gen Virol* **85**, 105–117.

Hnatiuk, S., Barry, M., Zeng, W., Liu, L., Lucas, A., Percy, D. & McFadden, G. (1999). Role of the C-terminal RDEL motif of the myxoma virus M-T4 protein in terms of apoptosis regulation and viral pathogenesis. *Virology* **263**, 290–306.

Huang, X. & Zhang, J. (1996). Methods for comparing a DNA sequence with a protein sequence. *Comput Appl Biosci* **12**, 497–506.

Huang, J., Huang, Q., Zhou, X. & 11 other authors (2004). The poxvirus p28 virulence factor is an E3 ubiquitin ligase. *J Biol Chem* **279**, 54110–54116.

Ichihashi, Y. & Dales, S. (1971). Biogenesis of poxviruses: interrelationship between hemagglutinin production and polykaryocytosis. *Virology* **46**, 533–543.

Kaiser, F. J., Moroy, T., Chang, G. T., Horsthemke, B. & Ludecke, H. J. (2003). The RING finger protein RNF4, a co-regulator of transcription, interacts with the TRPS1 transcription factor. *J Biol Chem* **278**, 38780–38785.

Kolhapure, R. M., Deolankar, R. P., Tupe, C. D. & 8 other authors (1997). Investigation of buffalopox outbreaks in Maharashtra State during 1992–1996. *Indian J Med Res* **106**, 441–446.

Lalani, A. S., Ness, T. L., Singh, R., Harrison, J. K., Seet, B. T., Kelvin, D. J., McFadden, G. & Moyer, R. W. (1998). Functional comparisons among members of the poxvirus T1/35kDa family of soluble CC-chemokine inhibitor glycoproteins. *Virology* **250**, 173–184.

Lalani, A. S., Masters, J., Graham, K., Liu, L., Lucas, A. & McFadden, G. (1999). Role of the myxoma virus soluble CC-chemokine inhibitor glycoprotein, M- T1, during myxoma virus pathogenesis. *Virology* **256**, 233–245.

Lambert, S., Yu, H., Prchal, J. T., Lawler, J., Ruff, P., Speicher, D., Cheung, M. C., Kan, Y. W. & Palek, J. (1990). cDNA sequence for human erythrocyte ankyrin. *Proc Natl Acad Sci U S A* **87**, 1730–1734.

Ligon, B. L. (2004). Monkeypox: a review of the history and emergence in the Western hemisphere. *Semin Pediatr Infect Dis* **15**, 280–287.

Lovering, R., Hanson, I. M., Borden, K. L. & 7 other authors (1993). Identification and preliminary characterization of a protein motif related to the zinc finger. *Proc Natl Acad Sci U S A* **90**, 2112–2116.

Lux, S. E., John, K. M. & Bennett, V. (1990). Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**, 36–42.

Lyngso, C., Bouteiller, G., Damgaard, C. K., Ryom, D., Sanchez-Munoz, S., Norby, P. L., Bonven, B. J. & Jorgensen, P. (2000). Interaction between the transcription factor SPBP and the positive cofactor RNF4. An interplay between protein binding zinc fingers. *J Biol Chem* **275**, 26144–26149.

Martinez-Pomares, L., Thompson, J. P. & Moyer, R. W. (1995). Mapping and investigation of the role in pathogenesis of the major unique secreted 35-kDa protein of rabbitpox virus. *Virology* **206**, 591–600.

Massung, R. F., Liu, L. I., Qi, J., Knight, J. C., Yuran, T. E., Kerlavage, A. R., Parsons, J. M., Venter, J. C. & Esposito, J. J. (1994). Analysis of the complete genome of smallpox variola major virus strain Bangladesh-1975. *Virology* **201**, 215–240.

McKelvey, T. A., Andrews, S. C., Miller, S. E., Ray, C. A. & Pickup, D. J. (2002). Identification of the orthopoxvirus p4c gene, which encodes a structural protein that directs intracellular mature virus particles into A-type inclusions. *J Virol* **76**, 11216–11225.

Moss, B. (2001). Poxviruses. In *Fields Virology*, 4th edn, pp. 2849–2884. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

Mural, R. J. (2000). ARTEMIS: a tool for displaying and annotating DNA sequence. *Brief Bioinform* **1**, 199–200.

Nerenberg, B. T., Taylor, J., Bartee, E., Gouveia, K., Barry, M. & Fruh, K. (2005). The poxviral RING protein p28 is a ubiquitin ligase that targets ubiquitin to viral replication factories. *J Virol* **79**, 597–601.

Reading, P. C., Symons, J. A. & Smith, G. L. (2003). A soluble chemokine-binding protein from vaccinia virus reduces virus virulence and the inflammatory response to infection. *J Immunol* **170**, 1435–1442.

Reynolds, M. G., Cono, J., Curns, A., Holman, R. C., Likos, A., Regnery, R., Treadwell, T. & Damon, I. (2004). Human monkeypox. *Lancet Infect Dis* **4**, 604–605.

Rice, P., Longden, I. & Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276–277.

Ronquist, F. & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

Senkevich, T. G., Koonin, E. V. & Buller, R. M. L. (1994). A poxvirus protein with a RING zinc finger motif is of crucial importance for virulence. *Virology* **198**, 118–128.

Senkevich, T. G., Wolffe, E. J. & Buller, R. M. L. (1995). Ectromelia virus RING finger protein is localized in virus factories and is required for virus replication in macrophages. *J Virol* **69**, 4103–4111.

Shchelkunov, S. N., Blinov, V. M. & Sandakhchiev, L. S. (1993). Ankyrin-like proteins of variola and vaccinia viruses. *FEBS Lett* **319**, 163–165.

Shchelkunov, S. N., Safronov, P. F., Totmenin, A. V., Petrov, N. A., Ryazankina, O. I., Gutorov, V. V. & Kotwal, G. J. (1998). The genomic sequence analysis of the left and right species-specific terminal region of a cowpox virus strain reveals unique sequences and a cluster of intact ORFs for immunomodulatory and host range proteins. *Virology* **243**, 432–460.

Shchelkunov, S. N., Totmenin, A. V., Babkin, I. V. & 11 other authors (2001). Human monkeypox and smallpox viruses: genomic comparison. *FEBS Lett* **509**, 66–70.

Swofford, D. L. (2003). PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4. Sunderland, MA: Sinauer Associates.

Upton, C., Schiff, L., Rice, S. A., Dowdeswell, T., Yang, X. & McFadden, G. (1994). A poxvirus protein with a RING finger motif binds zinc and localizes in virus factories. *J Virol* **68**, 4186–4195.

Upton, C., Hogg, D., Perrin, D., Boone, M. & Harris, N. L. (2000). Viral genome organizer: a system for analyzing complete viral genomes. *Virus Res* **70**, 55–64.

Upton, C., Slack, S., Hunter, A. L., Ehlers, A. & Roper, R. L. (2003). Poxvirus orthologous clusters: toward defining the minimum essential poxvirus genome. *J Virol* **77**, 7590–7600.

Westwood, J. C., Boulter, E. A., Bowen, E. T. & Maber, H. B. (1966). Experimental respiratory infection with poxviruses. I. Clinical virological and epidemiological studies. *Br J Exp Pathol* **47**, 453–465.

Wittek, R. & Moss, B. (1980). Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. *Cell* **21**, 277–284.

Wittek, R., Menna, A., Schumperli, D., Stoffel, S., Muller, H. K. & Wyler, R. (1977). *Hin*dIII and *Sst*I restriction sites mapped on rabbit poxvirus and vaccinia virus DNA. *J Virol* **23**, 669–678.

Wittek, R., Cooper, J. A., Barbosa, E. & Moss, B. (1980a). Expression of the vaccinia virus genome: analysis and mapping of mRNAs encoded within the inverted terminal repetition. *Cell* **21**, 487–493.

Wittek, R., Barbosa, E., Cooper, J. A., Garon, C. F., Chan, H. & Moss, B. (1980b). Inverted terminal repetition in vaccinia virus DNA encodes early mRNAs. *Nature* **285**, 21–25.