JOURNAL OF GENERAL VIROLOGY

RESEARCH ARTICLE

Ehmann *et al.*, *Journal of General Virology* 2021;102:001357
DOI 10.1099/jgv.0.001357

MICROBIOLOGY SOCIETY

OPEN ACCESS

Molecular and genomic characterization of a novel equine molluscum contagiosum-like virus

Rosina Ehmann[1,*], K. Brandes[2], M. Antwerpen[1], M. Walter[1], K. v. Schlippenbach[3], E. Stegmaier[4], S. Essbauer[1], J. Bugert[1], J. P. Teifke[5] and H. Meyer[1]

Abstract

Cases of pox-like lesions in horses and donkeys have been associated with poxviruses belonging to different genera of the family *Poxviridae*. These include the orthopoxviruses vaccinia virus (VACV), horsepoxvirus (HPXV) and cowpoxvirus (CPXV), as well as a potentially novel parapoxvirus and molluscum contagiosum virus (MOCV). However, with the exception of VACV, HPXV and CPXV, the genomic characterization of the causative agents remains largely elusive with only single short genome fragments available. Here we present the first full-length genome sequence of an equine molluscum contagiosum-like virus (EMCLV) directly determined from skin biopsies of a horse with generalized papular dermatitis. Histopathological analysis of the lesions revealed severe epidermal hyperplasia with numerous eosinophilic inclusion bodies within keratinocytes. Virions were detected in the lesions in embedded tissue by transmission electron microscopy. The genome sequence determined by next- and third-generation sequencing comprises 166 843 nt with inverted terminal repeats (ITRs) of 3473 nt. Overall, 20 of the predicted 159 ORFs have no equivalents in other poxviruses. Intriguingly, two of these ORFs were identified to encode homologues of mammalian proteins involved in immune signalling pathways, namely *secreted and transmembrane protein 1* (SECTM1) and *insulin growth factor-like family receptor 1* (IGFLR1), that were not described in any virus family so far. Phylogenetic analysis with all relevant representatives of the *Poxviridae* suggests that EMCLV should be nominated as a new species within the genus *Molluscipoxvirus*.

INTRODUCTION

The *Poxviridae* constitute a diverse family of complex viruses with a linear, double-stranded DNA genome of 130–350 kb. Members of the subfamily *Chordopoxvirinae* infect a large variety of vertebrates whereas insects are the hosts of *Entomopoxvirinae*. While some poxviruses like variola virus (VARV) and ectromelia virus (ECTV) are extremely species-specific, most poxviruses exhibit a broader host range as illustrated by cowpoxvirus (CPXV) [1].

Horses seem to be susceptible hosts for a surprising number of different poxvirus species that cause papular or pustular cutaneous lesions. However, due to the highly similar clinical features, characterization of the respective causative agents demands molecular analyses, which have been only rarely achieved. The spectrum of the etiological viruses and their genome sequences is therefore poorly understood. 'Classical' horsepox is associated with different disease types including localized forms of the face ('buccal type') or pasterns ('grease') and a generalized form called 'viral papular dermatitis' [2]. Horsepox was frequently described in Europe in the nineteenth and twentieth century but became progressively rare to potentially extinct in modern times [3]. Its causative agent horsepoxvirus (HPXV), a vaccinia virus-like orthopoxvirus was recently sequenced from historical specimens of Mongolian horses [4]. Outbreaks of vaccinia virus (VACV) infection in horses have been historically documented after contact with vaccinated persons. In a recent description, VACV infection in horses was detected in Brazil, and linked to potential vaccine escapees that established natural reservoirs [5–7]. Severe cowpoxvirus (CPXV) infections in horses were reported in at least three fatal cases including two foals and might be linked to immunocompromised individuals [8, 9]. Another poorly characterized orthopoxvirus is associated

Received 17 September 2019; Accepted 06 November 2019; Published 10 January 2020
Author affiliations: [1]Bundeswehr Institute of Microbiology, Munich, Germany; [2]Animal Pathology Augsburg, Augsburg, Germany; [3]Veterinarian Practice, Zusamaltheim, Germany; [4]Makoa Farm, Uswaa, Tanzania; [5]Federal Research Institute for Animal Health, Greifswald - Insel Riems, Germany.
***Correspondence:** Rosina Ehmann, rosinaehmann@bundeswehr.org
Keywords: equine molluscum; novel poxvirus, Molluscipoxvirus; SECTM1; IGFLR1.
Abbreviations: EMCLV, equine molluscum contagiosum-like virus; IGFLR1, insulin growth factor-like family receptor 1; ITR, inverted terminal repeats; MOCV, molluscum contagiosum virus; SECTM1, secreted andtransmembraneprotein 1.

with Uasin Gishu, a disease reported to cause pox-like disease in nonindigenous horse breeds in several African countries. In contrast to viral papular dermatitis, Uasin Gishu has a protracted disease progression and is considerably less contagious [10, 11]. A novel parapoxvirus species has been linked to a fatal infection in a horse in Finland and to skin lesions of two American patients with contact to horses [12, 13]. However, only sequence information of short diagnostic PCR fragments is available for these viruses. Lastly, a condition similar to molluscum contagiosum in humans with numerous small, indolent papules has been sporadically reported in horses and donkeys in Europe, Africa and Northern America [14–16]. Based on macroscopic and histologic presentations, electron microscopic studies and *in situ* hybridization assays the causative agent has been shown to be related to molluscum contagiosum virus (MOCV) [17]. Information on the genome sequence of the virus causing equine molluscum contagiosum is so far limited to a 630 bp fragment of the conserved core [15]. The phylogenetic relationship of this virus to MOCV and the respective host ranges therefore remain yet uncertain.

Here we report molluscum contagiosum-like lesions in a horse with detailed characterization of the causative virus. Successful determination of the first full-length genome sequence offered new insights into virus evolution of the GC-rich poxviruses and identified two novel viral homologues of mammalian genes that extend our known portfolio of how viruses interact with the host immune system.

RESULTS

Clinical and histological presentation

In 2016 a 10-year-old mare used for horseback riding tours in the region of Kilimanjaro, Tanzania, presented with numerous nodular skin lesions affecting extensive parts of the body but clustering most densely on the trunk and neck. The alterations appeared as round, verrucous, indolent masses of 5–15 mm in diameter surrounded by tufted hair (Fig. 1a). Several horses on the same farm were reported to show similar symptoms over the last years. Horses affected by equine molluscum on

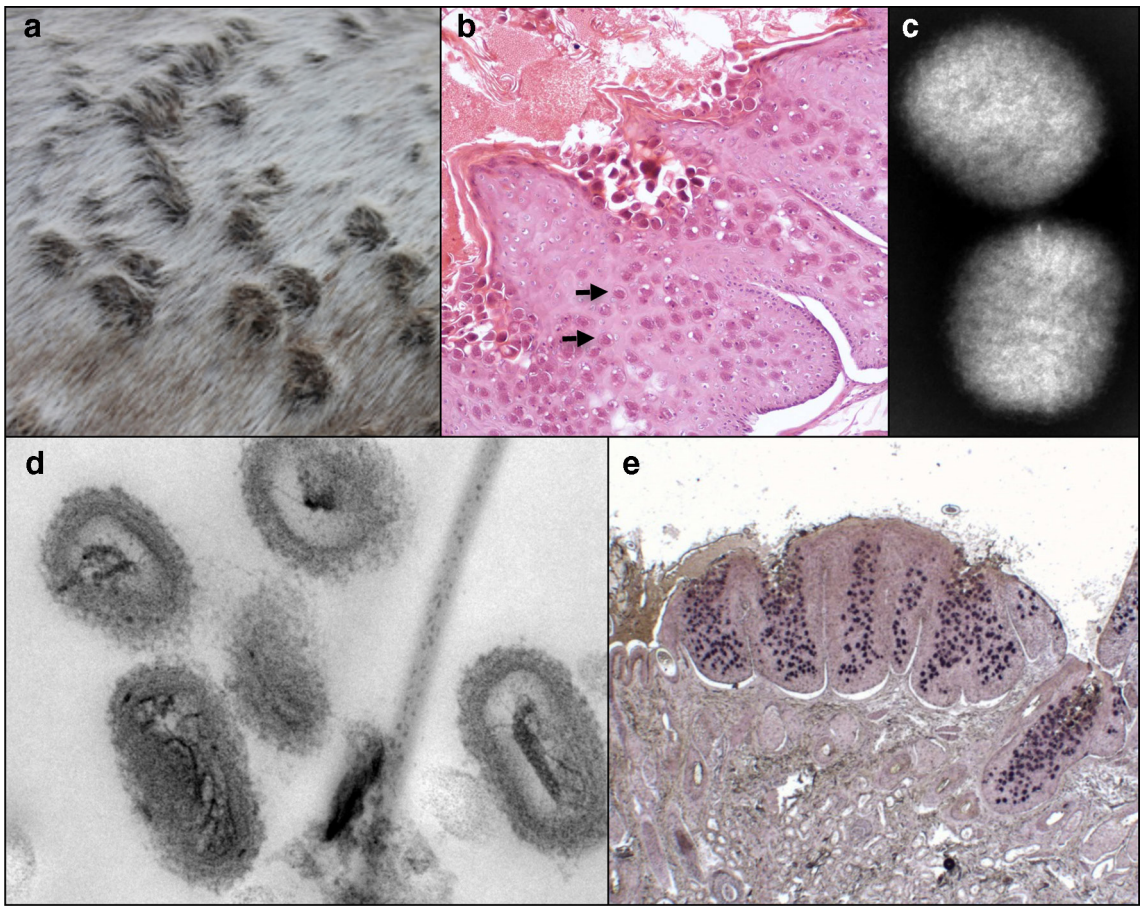

Fig. 1. (a–e) Clinical presentation and microscopic analysis of equine molluscum contagiosum lesions. (a) Raised, papulonodular skin lesions on the neck of the horse (5–15 mm in diameter). (b) Histological changes of the skin including epidermal hyperplasia, ballooning degeneration and numerous eosinophilic inclusion bodies (arrows) in the cytoplasm of keratinocytes, HE, magnification x200. Electron microscopic study of the particles with (c) negative stain and (d) ultra-thin sections. (e) *In situ* hybridization using a digoxigenin-labelled MOCV-specific DNA probe. A specific signal could be detected within the cytoplasmic inclusion bodies, magnification x100.

this farm presented lesions throughout their entire lives but the number of lesions varied in correlation with the immune status of the individual animals. Horses with poor body condition or other stressors like pregnancy or vaccination developed more nodules. A history of immunosuppression and *Corynebacterium pseudotuberculosis* infection preceded blooming of the skin lesions in the 10-year-old mare described here, which showed the first lesions at the age of 5 years. Biopsies of several papules were obtained for analyses. Histologically, the nodules consisted of a sharply delimited and severe epidermal hyperplasia. The enlarged stratum spinosum showed numerous large, eosinophilic, intracytoplasmic inclusion bodies within the keratinocytes (Fig. 1b). This histologic pattern is characteristic of molluscum contagiosum [18].

Identification of the etiological agent

Nucleic acid extracted from the skin biopsy was tested for the presence of poxvirus DNA in a diagnostic pan-pox PCR with two profiles detecting low-GC and high-GC content poxviruses, respectively [19]. The high-GC PCR produced a specific amplicon of 630 bp. Sequencing of the amplicon (corresponding to nt 79878 to 80507 in GenBank acc. no. MN339351) resulted in a unique sequence that shared 99% nucleotide identity with a MOCV-like poxvirus derived from two donkeys (GenBank acc. no. JQ269324.1). Nucleotide identity to the MOCV reference sequence (GenBank acc. no. NC_001731.1) was considerably lower with only 90.5%.

Electron microscopy was conducted with both negative staining and ultra-thin sections. Negatively stained samples revealed particles of 250–280×150 nm with a raspberry-like surface structure or inner structure resembling that of MOCV or MOCV-like particles (Fig. 1c). In the sections, virions were detected in the epidermal layers with dimensions of 200–280×150 nm. The inner core consisted of a dumb-bell-shaped electron-dense structure, which was surrounded by a less electron-dense laminated capsule typical of poxviruses (Fig. 1d). *In situ* hybridization using a 300 bp digoxigenin-labelled MOCV-specific probe resulted in pronounced hybridization signals within the intracytoplasmic inclusions bodies of the keratinocytes (Fig. 1e).

Isolation of the causative virus was attempted by inoculation and two blind passages in both MA104, a standard cell line commonly used for the isolation of poxviruses, and UcP-R, bovine oesophageal cells (primary cells successfully used for the propagation of difficult to grow parapoxviruses). No cytopathic effect or increase in viral nucleic acid could be detected over time. Co-infection trials with vaccinia virus MVA as a self-limiting helper virus were also unsuccessful in both cell types. MOCV is known to induce abortive infections in standard cell-culture systems. The infectivity of the virions detected in the skin sample was therefore also assessed using a reporter assay established for MOCV [20]. This assay uses reporter plasmids expressing firefly luciferase under the control of the strong early/late poxviral consensus promoter to detect the activity of the viral transcription machinery in infected susceptible cells. No luciferase signal could be detected after inoculation of HEK-293 cells with sample material.

Determination of the full-length genome sequence

Sequencing of the viral genome was conducted by both Illumina MiSeq and MinION technology. 812 629 nanopore reads were sequenced with 823 360 204 bases in total. After filtering, 269 129 of these sequences with an average length of 1814 bases were used for assembly. Flye assembler [21] created a single contig related with molluscum contagiosum virus with an untrimmed length of 181 471 bases and a mean coverage of 295× as well as nine contigs related to *Equus caballus,* which were discarded. The Illumina sequencing run produced 612 011 paired-end reads related to MOCV with an average coverage of 160×. Illumina short-read sequencing data was then used for polishing of the long-read MinION molluscum-related contig with pilon tool [22]. The finally assembled and polished molluscum-related contig contained signs of a concatemer with four concatemer resolution motifs and four ITR regions. The sequence was therefore manually trimmed at the genome ends.

The final full-length genome sequence contained 166 843 nt with inverted terminal repeats (ITRs) of 3473 nt at both flanks. The tandem repeat cassettes (nt 212–1078 and 165 766–166 632) in the ITRs and two coding regions (nt 39 753–40 001 and 140 568–140 897) contained repetitive sequence motifs that could only be resolved to their true length with the long-read capacity of the MinION platform. The conserved concatemer junction resolution signals close to the genome ends locate to nt 67–83 (5′-ATTTATAGGCAGAAAAA-3′) and nt 166 761–166 777 5′-TTTTTCTGCCTATAAAT-3′).

Phylogeny

Phylogenetic analyses to determine the relationship of the causative virus of equine molluscum contagiosum with other poxviruses was based on an alignment of the newly determined genome sequence with 38 representative poxvirus species. For the phylogeny the alignment was shortened to concatenated sequence blocks conserved in all 39 virus genomes (Fig. 2a). Within this tree the virus causing the molluscum contagiosum-like lesions in the horse from Tanzania clustered with MOCV as the closest phylogenetic neighbour. However, when compared in a separate phylogenetic analysis with all available 15 MOCV full-length genome sequences, it appears highly divergent on a separate branch from both MOCV subtypes (Fig. 2b). These findings suggest that the virus derived from the equine molluscum lesions represents a novel species in the genus *Molluscipoxvirus* and was therefore tentatively designated equine molluscum contagiosum-like virus (EMCLV).

General genome features of EMCLV

The genome sequence of the EMCLV strain Tanzania 2016 contains 166 843 nt. It is markedly shorter than that of MOCV with an average length of approximately 190 000 nt. Also, the G+C content is slightly higher in EMCLV (66.8 %) than in

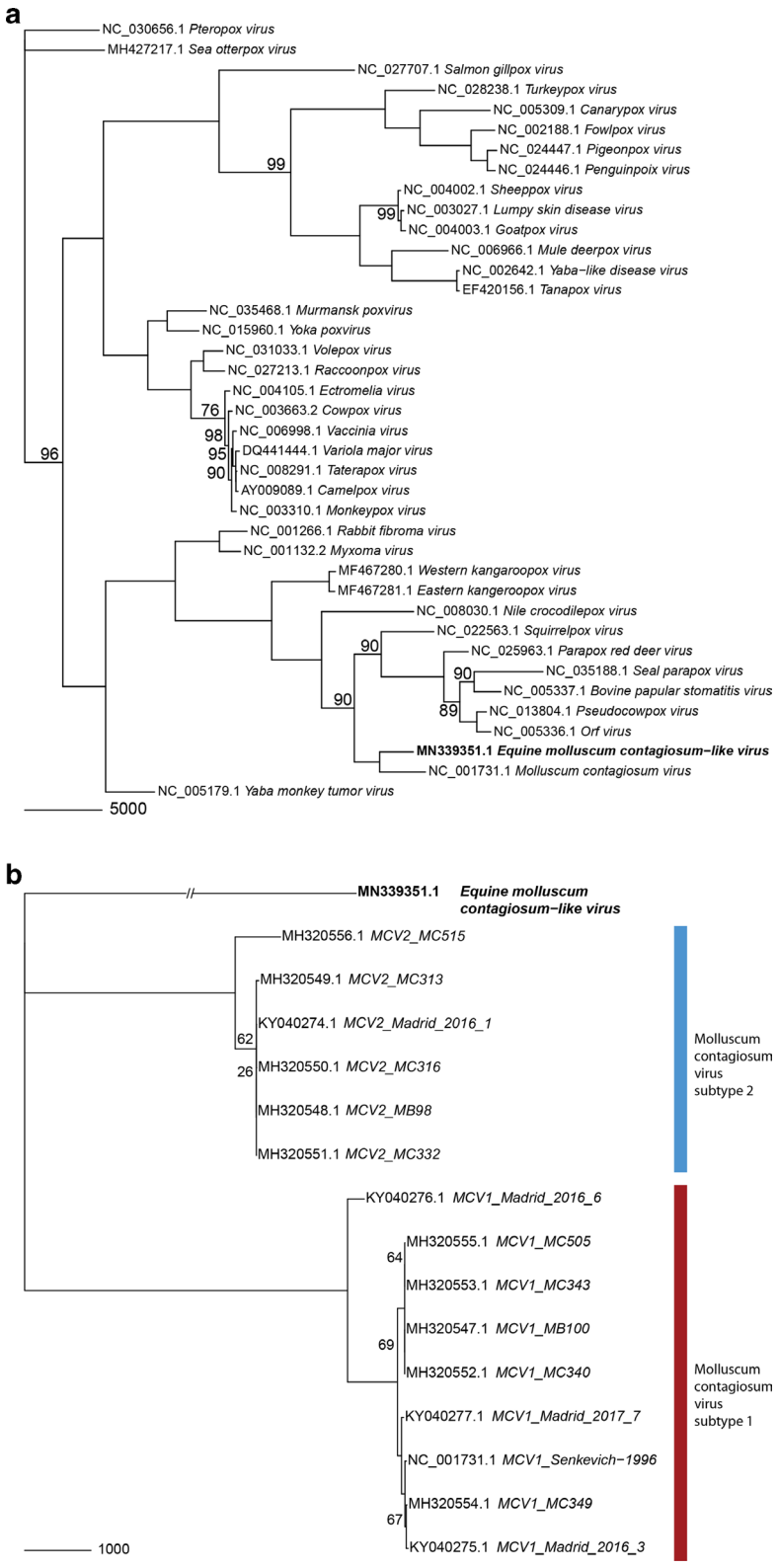

Fig. 2. Phylogenetic analysis of EMCLV. When compared to 38 representative poxviruses across all genera, the genome sequence retrieved from the equine molluscum contagiosum lesions shares a common branch with MOCV (Fig. 2a). A phylogenetic tree constructed with all full-length genome sequences available for MOCV clearly indicates that EMCLV is highly divergent from both subtype 1 and subtype 2 of MOCV (Fig. 2b). The scale bar indicates the number of nucleotide substitutions.

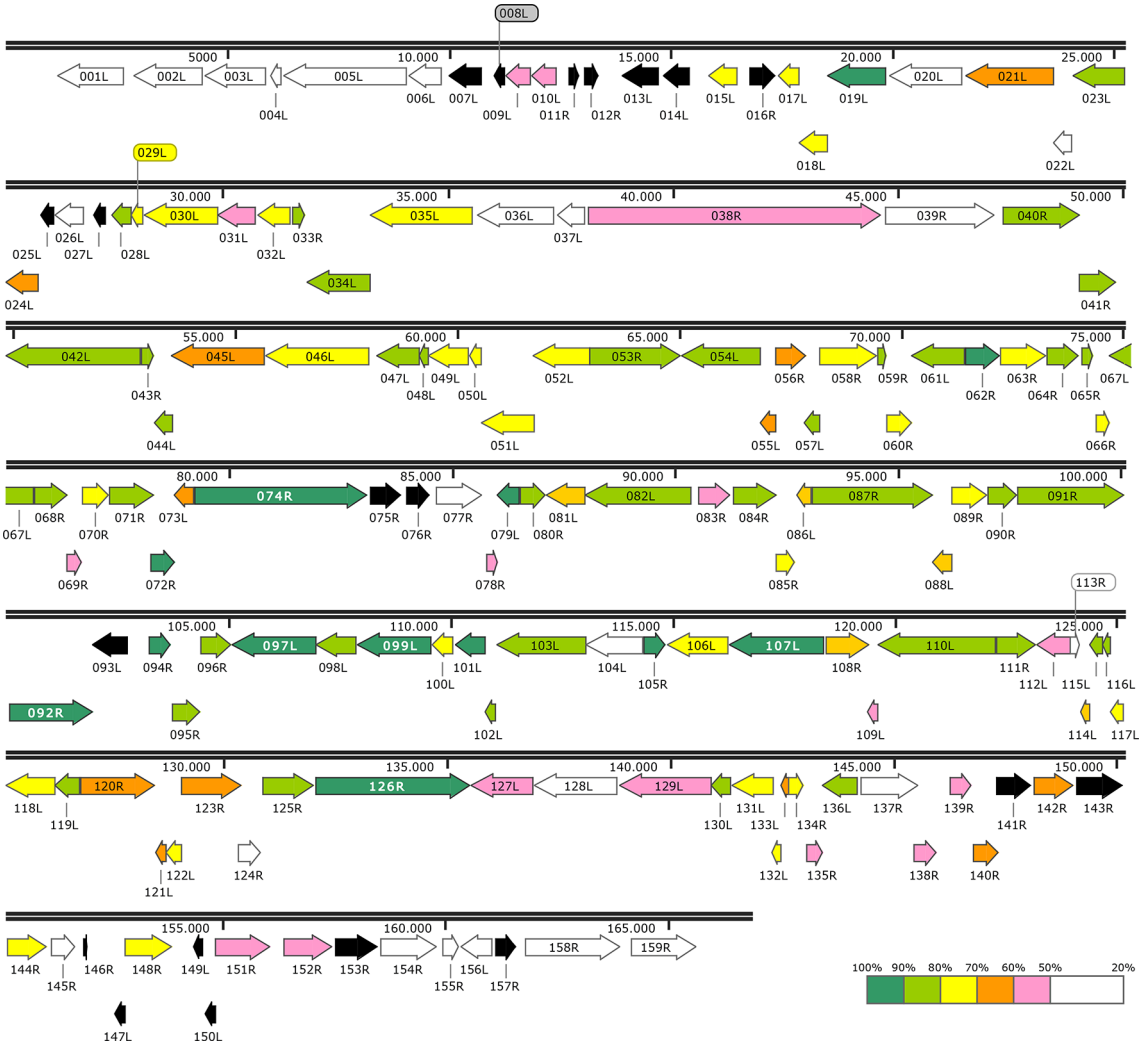

Fig. 3. Genome map of EMCLV. Different colours were used to visualize amino acid sequence similarity to MOCV (reference sequence GenBank acc. no. NC_001731). ORFs depicted in black are unique to EMCLV.

MOCV (63.7 % on average). The tandem repeat structure in the inverted terminal repeats differs considerably between EMCLV and MOCV. EMCLV shows only one motif (5′-GGAT TTGGCCATACTGAGTGCGCGCGCAAAG-3′) with minor nucleotide variations in a cassette of 29 repetitions whereas MOCV features a complex pattern of different motifs [23].

Annotation of EMCLV and comparison with MOCV

A conservative approach was used for the genome annotation of EMCLV with a minimum ORF length of 40 codons. Shorter ORFs were only annotated if unambiguous poxviral promoter sequences were present. In total, 159 ORFs named EMCLV1L to EMCLV159R (and EMCLV1 to EMCLV159 for the respective predicted proteins) were identified as the most probable predicted gene set for EMCLV (Fig. 3 and Table 1). Only 139 of these ORFS were orthologues of other poxvirus genes with MOCV as the closest hit in nearly all cases [EMCLV109L is most closely related to Western kangaroopox virus (WKPV)].

As typically seen for related poxviruses, sequence similarity on the amino acid level is higher in the conserved core of the genome whereas similarity levels decrease significantly towards the more variable flanking regions of the genome (colour coding illustrates the amino acid similarity levels in Fig. 3). The gene sets of EMCLV and MOCV with the respective homologues and gene functions are listed in Table 1. ORFs conserved in EMCLV and MOCV show a complete synteny with only a single exception presented by MC054L and EMCLV145R, which code for a viral homologue of the IL18-binding protein (IL18-BP). EMCLV lacks 24 of the 163 genes annotated in MOCV, namely MC004L, MC010R, MC011L, MC012L, MC014R, MC024L, MC051L, MC052R, MC053L, MC055R, MC063L, MC064R, MC066L, MC096L, MC132R, MC146R, MC147R, MC148R, MC150R, MC151L, MC155R, MC156R, MC160L and MC161R. Interestingly, several of the missing genes in EMCLV listed above are members of gene families with duplication or triplication in

Table 1. Comparative list of the gene content of EMCLV and MOCV. Amino acid similarity of conserved ORFs was determined by EMBOSS Needle pairwise alignment with MOCV reference sequence NC001731

EMCLV gene	MCV orthologue	Amino acid similarity (%)	ORF position in EMCLV	ORF size in EMCLV (bp)	Gene function
EMCLV001L (=EMCLV160R)	**MC001L; MC163R**	32,7 (MC163R); 13,8 (MC001L)	1180…2655	1476	protein with unknown function, contains a semaphorin domain
EMCLV002L	**MC002L**	44,8	2887…4443	1557	protein with unknown function, similarity to human SLAM
EMCLV003L	**MC003L**	40,7	4489…5847	1359	protein with unknown function, related to VACV F5L
–	**MC004L**	–	–	–	protein with unknown function
EMCLV004L	**MC005L**	25	5965…6213	249	inhibitor of NF-κB activation
EMCLV005L	**MC006L**	38,4	6263…9040	2778	protein with unknown function
EMCLV006L	**MC007L**	46,1	9111…9833	723	inactivator of retinoblastoma protein (pRb)
EMCLV007L	–	–	9992…10732	741	viral homologue of equine *secreted and transmembrane protein 1* (SECTM1)
EMCLV008L	–	–	11026…11241	216	hypothetical protein
EMCLV009L	**MC008L**	56,4	11277…11831	555	hypothetical protein
EMCLV010L	**MC009L**	57,3	11879…12412	534	hypothetical protein
–	**MC010R**	–	–	–	protein with unknown function
–	**MC011L**	–	–	–	protein with unknown function
–	**MC012L**	–	–	–	protein with unknown function
EMCLV011R	–	–	12715…12939	225	hypothetical protein
EMCLV012R	–	–	13073…13384	312	hypothetical protein
EMCLV013L	–	–	13922…14728	807	hypothetical protein
EMCLV014L	–	–	14838…15434	597	hypothetical protein
EMCLV015L	**MC013L**	70,3	15852…16499	648	protein with unknown function, similarity to human DnaJ molecular chaperone
–	**MC014R**	–	–	–	protein with unknown function
EMCLV016R	–	–	16785…17369	585	hypothetical protein
EMCLV017L	**MC015L**	72,4	17439…17891	453	hypothetical protein
EMCLV018L	**MC016L**	77	17916…18551	636	S-S bond formation pathway protein
EMCLV019L	**MC017L**	92,3	18535…19863	1329	putative serine/threonine protein kinase
EMCLV020L	**MC018L**	48	19943…21574	1632	RhoA signalling inhibitor, virus release protein
EMCLV021L	**MC019L**	65,3	21667…23643	1977	EEV maturation protein
EMCLV022L	**MC020L**	34,5	23633…24043	411	protein with unknown function
EMCLV023L	**MC021L**	79,8	24071…25219	1149	major envelope protein
EMCLV024L	**MC022L**	61,5	25212…25928	717	hypothetical protein
EMCLV025L	–	–	25983…26267	285	hypothetical protein
EMCLV026L	**MC023L**	42,1	26305…26919	615	protein with unknown function
–	**MC024L**	–	–	–	protein with unknown function
EMCLV027L	–	–	27149…27409	261	hypothetical protein
EMCLV028L	**MC025L**	89,2	27551…27997	447	protein with unknown function
EMCLV029L	**MC026L**	71,4	27999…28250	252	APC11-like protein with RING finger domain

Continued

Table 1. Continued

EMCLV gene	MCV orthologue	Amino acid similarity (%)	ORF position in EMCLV	ORF size in EMCLV (bp)	Gene function
EMCLV030L	**MC027L**	73,8	28283…29896	1614	hypothetical protein
EMCLV031L	**MC028L**	51,9	29898…30734	837	hypothetical protein with a Src homology 2 (SH2) domain found in Carboxyl-Terminal Src Kinase (Csk)
EMCLV032L	**MC029L**	74,4	30794…31510	717	serine recombinase
EMCLV033R	**MC030R**	81,9	31585…31866	282	virion core DNA-binding phosphoprotein
EMCLV034L	**MC031L**	88,7	31882…33297	1416	catalytic subunit of poly(A) polymerase
EMCLV035L	**MC032L**	71,1	33294…35543	2250	IEV morphogenesis protein
EMCLV036L	**MC033L**	41,5	35659…37368	1710	hypothetical protein, contains a class I major histocompatibility complex (MHC) alpha chain immunoglobulin domain
EMCLV037L	**MC034L**	40,7	37452…38063	612	DNA-dependent RNA polymerase 30 kDa subunit (RPO30)
EMCLV038R	**MC035R**	54,5	38140…44655	6516	integral membrane protein, homologue of VARV B22R
EMCLV039R	**MC036R**	34,3	44719…47166	2448	hypothetical protein, contains a BEN domain
EMCLV040R	**MC037R**	89,4	47346…49049	1704	IMV virion morphogenesis protein
EMCLV041R	**MC038R**	82,2	49036…49848	813	membrane phosphoprotein
EMCLV042L	**MC039L**	81,1	49853…52864	3012	DNA-dependent DNA polymerase
EMCLV043R	**MC040R**	86,1	52893…53189	297	sulfhydryl oxidase (FAD-linked)
EMCLV044L	**MC041L**	82,2	53186…53587	402	virion core protein
EMCLV045L	**MC042L**	66,1	53574…55679	2106	IEV morphogenesis protein
EMCLV046L	**MC043L**	73,5	55709…58024	2316	IEV morphogenesis protein
EMCLV047L	**MC044L**	89,7	58206…59141	936	DNA-binding core protein
EMCLV048L	**MC045L**	80,6	59150…59365	216	IMV membrane protein
EMCLV049L	**MC046L**	74,6	59369…60253	885	ssDNA-binding phosphoprotein
EMCLV050L	**MC047L**	78	60303…60542	240	IMV membrane protein
EMCLV051L	**MC048L**	73,5	60564…61730	1167	telomere-binding protein, 36 kDa major membrane protein
EMCLV052L	**MC049L**	75,3	61727…62995	1269	virion core cysteine protease
EMCLV053R	**MC050R**	82,7	63004…65049	2046	ATP-dependent RNA helicase (virion NTPase II)
–	**MC051L**	–	–	–	protein with unknown function, similarity to IL18-BP
–	**MC052R**	–	–	–	hypothetical protein
–	**MC053L**	–	–	–	protein with unknown function, similarity to IL18-BP
EMCLV145R	**MC054L**	20,6	151189…151731	543	viral homologue of human (MC054L)/equine (EMCLV145R) IL18-BP
–	**MC055R**	–	–	–	hypothetical protein
EMCLV054L	**MC056L**	86,8	65051…66844	1794	metalloendopeptidase
EMCLV055L	**MC057L**	66,4	66841…67173	333	entry/fusion complex component
EMCLV056R	**MC058R**	68,1	67167…67874	708	viral late transcription elongation factor (VLTF)
EMCLV057L	**MC059L**	82,5	67808…68179	372	glutaredoxin-like protein
EMCLV058R	**MC060R**	72,4	68183…69484	1302	FEN1-like nuclease
EMCLV059R	**MC061R**	88,9	69491…69682	192	DNA-dependent RNA polymerase 7 kDa subunit (RPO7)

Continued

Table 1. Continued

EMCLV gene	MCV orthologue	Amino acid similarity (%)	ORF position in EMCLV	ORF size in EMCLV (bp)	Gene function
EMCLV060R	**MC062R**	70,9	69683…70273	591	Nlpc/p60 superfamily protein
–	**MC063L**	–	–	–	hypothetical protein, similarity to non-histone chromatin phosphoprotein
–	**MC064R**	–	–	–	hypothetical protein, prenylated and membrane-associated
EMCLV061L	**MC065L**	79,7	70226…71416	1191	virion phosphoprotein, early morphogenesis
–	**MC066L**	–	–	–	glutathione peroxidase, inhibitor of apoptosis
EMCLV062R	**MC067R**	94,6	71446…72228	783	viral late transcription trans-activator (VLTF-1)
EMCLV063R	**MC068R**	73,5	72254…73288	1035	myristylprotein of the poxvirus entry/fusion-complex
EMCLV064R	**MC069R**	89,7	73289…74020	732	myristoylated virion core protein
EMCLV065R	**MC070R**	79,6	74069…74341	273	crescent membrane and immature virion formation protein
EMCLV066R	**MC071R**	70,8	74386…74703	318	hypothetical protein
EMCLV067L	**MC072L**	86,8	74693…75601	909	internal virion protein
EMCLV068R	**MC073R**	80,1	75627…76382	756	DNA-binding virion core protein
EMCLV069R	**MC074R**	51,4	76375…76722	348	entry/fusion IMV membrane protein
EMCLV070R	**MC075R**	73,9	76712…77320	609	IMV membrane protein, virion morphogenesis
EMCLV071R	**MC076R**	86,3	77332…78345	1014	cap-specific mRNA (nucleoside-2′-O-)-methyltransferase
EMCLV072R	**MC077R**	91,4	78245…78808	564	DNA-dependent RNA polymerase 22 kDa subunit (RPO22)
EMCLV073L	**MC078L**	67,6	78759…79205	447	myristylprotein of the poxvirus entry/fusion-complex
EMCLV074R	**MC079R**	94,4	79246…83121	3876	DNA-dependent RNA polymerase 147 kDa subunit (RPO147)
EMCLV075R	–	–	83167…83883	717	hypothetical protein
EMCLV076R	–	–	83991…84503	513	hypothetical protein
EMCLV077R	**MC080R**	32,3	84648…85694	1047	MHC-I homologue, inhibits cell surface MHC-I protein presentation
EMCLV078R	**MC081R**	58,6	85788…86039	252	protein with unknown function
EMCLV079L	**MC082L**	91,7	86008…86517	510	tyrosine/serin protein phosphatase, IFN-gamma inhibitor
EMCLV080R	**MC083R**	85,3	86532…87107	576	IMV membrane protein; subunit of the poxvirus multiprotein entry-fusion complex
EMCLV081L	**MC084L**	64,5	87126…87998	873	IMV heparin-binding surface protein
EMCLV082L	**MC085L**	89,8	87999…90377	2379	RNA polymerase-associated transcription specificity factor (RAP94)
EMCLV083R	**MC086R**	54,9	90529…91257	729	viral late transcription factor VLTF4
EMCLV084R	**MC087R**	83,3	91334…92299	966	DNA topoisomerase type I
EMCLV085R	**MC088R**	74,3	92284…92715	432	crescent membrane and immature virion formation protein
EMCLV086L	**MC089L**	64,9	92746…93066	321	hypothetical protein
EMCLV087R	**MC090R**	81,2	93095…95821	2727	mRNA capping enzyme large subunit
EMCLV088L	**MC091L**	66,1	95783…96226	444	virion core protein
EMCLV089R	**MC092R**	74,7	96219…97025	807	virion core protein
EMCLV090R	**MC093R**	82,3	97031…97693	663	uracil-DNA glycosylase

Continued

Table 1. Continued

EMCLV gene	MCV orthologue	Amino acid similarity (%)	ORF position in EMCLV	ORF size in EMCLV (bp)	Gene function
EMCLV091R	**MC094R**	84,7	97717…100080	2364	NTPase, DNA primase
EMCLV092R	**MC095R**	97,5	100077…101984	1908	viral early transcription factor (VETF) 70 kDa small subunit
EMCLV093L	–	–	101944…102726	783	hypothetical protein
–	**MC096L**	–	–	–	hypothetical protein
EMCLV094R	**MC097R**	90,7	103227…103712	486	DNA-dependent RNA polymerase 18 kD subunit (RPO18)
EMCLV095R	**MC098R**	85,4	103746…104384	639	NTP phosphohydrolase of the MutT family, mRNA decapping enzyme
EMCLV096R	**MC099R**	86,7	104381…105082	702	NTP phosphohydrolase of the MutT family, mRNA decapping enzyme
EMCLV097L	**MC100L**	89,6	105096…106994	1899	virion ATPase I
EMCLV098L	**MC101L**	87,1	106995…107882	888	mRNA capping enzyme (small subunit) and transcription initiation factor
EMCLV099L	**MC102L**	91,8	107918…109561	1644	rifampicin resistance protein
EMCLV100L	**MC103L**	75,1	109597…110052	456	viral late transcription factor (VLTF-2)
EMCLV101L	**MC104L**	96,9	110124…110801	678	viral late transcription factor (VLTF-3)
EMCLV102L	**MC105L**	87,1	110798…111010	213	S-S bond formation pathway protein
EMCLV103L	**MC106L**	88,2	111041…113053	2013	P4b major core protein precursor
EMCLV104L	**MC107L**	47,3	113072…114337	1266	putative core protein
EMCLV105R	**MC108R**	92,8	114376…114876	501	DNA-dependent RNA polymerase 19 kD subunit (RPO19)
EMCLV106L	**MC109L**	70,9	114890…116257	1368	virion morphogenesis, core protein
EMCLV107L	**MC110L**	90,7	116281…118407	2127	viral early transcription factor large subunit (VETF-L)
EMCLV108R	**MC111R**	61,6	118470…119471	1002	viral intermediate transcription factor VITF-3 32 kDa small subunit
EMCLV109L	**MC112L**	52,3	119400…119627	228	viral membrane associated, early morphogenesis protein
EMCLV110L	**MC113L**	85,4	119628…122282	2655	P4a major core protein precursor
EMCLV111R	**MC114R**	81,1	122297…123214	918	viral membrane formation protein
EMCLV112L	**MC115L**	53,9	123215…123961	747	virion core and cleavage-processing protein
EMCLV113R	**MC116R**	49,3	123976…124203	228	putative virion membrane protein
EMCLV114L	**MC117L**	69,4	124187…124399	213	IMV membrane protein
EMCLV115L	**MC118L**	83,2	124400…124687	288	phosphorylated IMV membrane protein
EMCLV116L	**MC119L**	81,1	124704…124865	162	IMV membrane protein, non-essential
EMCLV117L	**MC120L**	71,9	124869…125159	291	core protein
EMCLV118L	**MC121L**	77	125143…126243	1101	myristylprotein of the poxvirus entry/fusion-complex
EMCLV119L	**MC122L**	82,1	126251…126790	540	IMV virion membrane protein
EMCLV120R	**MC123R**	61,5	126805…128493	1689	DNA helicase
EMCLV121L	**MC124L**	63,9	128483…128734	252	zinc finger-like protein
EMCLV122L	**MC125L**	78,1	128735…129079	345	IMV membrane protein; subunit of the poxvirus multiprotein entry-fusion complex
EMCLV123R	**MC126R**	67,1	129078…130433	1356	DNA polymerase processivity factor

Continued

Table 1. Continued

EMCLV gene	MCV orthologue	Amino acid similarity (%)	ORF position in EMCLV	ORF size in EMCLV (bp)	Gene function
EMCLV124R	**MC127R**	43,1	130354…130875	522	Holliday junction resolvase
EMCLV125R	**MC128R**	84,9	130900…132051	1152	intermediate transcription factor VITF-3 45 kDa large subunit
EMCLV126R	**MC129R**	95,5	132084…135554	3471	DNA-dependent RNA polymerase 132 kDa subunit (RPO132)
EMCLV127L	**MC130L**	55,9	135565…136932	1368	A-type inclusion protein
EMCLV128L	**MC131L**	47,5	136983…138830	1848	A-type inclusion protein
–	**MC132R**	–	–	–	inhibitor of NF-κB activation
EMCLV129L	**MC133L**	51,1	138873…140942	2070	A-type inclusion protein
EMCLV130L	**MC134L**	84,4	140943…141368	426	IMV membrane protein
EMCLV131L	**MC135L**	78,2	141410…142318	909	DNA-dependent RNA polymerase 35 kDa subunit (RPO35)
EMCLV132L	**MC136L**	74,6	142287…142490	204	IMV membrane protein for virion morphogenesis
EMCLV133L	**MC137L**	60,7	142491…142661	171	protein with unknown function
EMCLV134R	**MC138R**	73,5	142671…143024	354	hypothetical protein
EMCLV135R	**MC139R**	59,3	143064…143435	372	hypothetical protein
EMCLV136L	**MC140L**	88,2	143432…144199	768	ATPase/DNA packaging protein
EMCLV137R	**MC141R**	42,4	144284…145588	1305	protein with unknown function
EMCLV138R	**MC142R**	50,5	145468…145998	531	EEV membrane phosphoglycoprotein
EMCLV139R	**MC143R**	52,9	146276…146782	507	IEV and EEV membrane glycoprotein
EMCLV140R	**MC144R**	64,4	146816…147388	573	MHC class II antigen presentation inhibitor
EMCLV141R	–	–	147334…148107	774	hypothetical protein
EMCLV142R	**MC145R**	65	148171…149061	891	concanavalin-like precursor protein
EMCLV143R	–	–	149126…150157	1032	hypothetical protein
–	**MC146R**	–	–	–	protein with unknown function
–	**MC147R**	–	–	–	protein with unknown function
–	**MC148R**	–	–	–	CC-chemokine homologue
EMCLV144R	**MC149R**	70,2	150203…151102	900	putative extracellular enveloped virion protein
EMCLV145R	**MC054L**	20,6	151190…151732	543	viral homologue of human (MC054L)/equine (EMCLV146R) IL18-binding protein
EMCLV146R	–	–	151920…152006	87	hypothetical protein
EMCLV147L	–	–	152600…152857	258	hypothetical protein
–	**MC150R**	–	–	–	protein with unknown function
–	**MC151L**	–	–	–	protein with unknown function
EMCLV148R	**MC152R**	74,6	152856…153917	1062	viral homologue of equine 3-beta-hydroxysteroid dehydrogenase
EMCLV149L	–	–	154391…154591	201	hypothetical protein
EMCLV150L	–	–	154651…154860	210	hypothetical protein
EMCLV151R	**MC153R**	57	154859…156115	1257	hypothetical protein
EMCLV152R	**MC154R**	50,7	156399…157490	1092	protein with unknown function
EMCLV153R	–	–	157559…158512	954	hypothetical protein

Continued

Table 1. Continued

EMCLV gene	MCV orthologue	Amino acid similarity (%)	ORF position in EMCLV	ORF size in EMCLV (bp)	Gene function
–	**MC155R**	–	–	–	putative prenylated, membrane-associated protein
–	**MC156R**	–	–	–	protein with unknown function
EMCLV154R	**MC157R**	40,1	158574…159836	1263	protein with unknown function, contains a CD48-like immunoglobulin domain
EMCLV155R	**MC158R**	29,9	159956…160321	366	protein with unknown function
EMCLV156L	**MC159L**	47,5	160368…161054	687	viral FLIP, inhibitor of apotosis, IRF-3, NF-κB and NEMO polyubiquitination
EMCLV157R	–	–	161144…161605	462	viral homologue of the extracellular ligand-binding domain of equine IGFLR1
–	**MC160L**	–	–	–	inhibitor of NF-κB activation
–	**MC161R**	–	–	–	protein with unknown function, contains an immunoglobulin domain, similarity to human SLAM
EMCLV158R	**MC162R**	33,6	161804…163957	2154	protein with unknown function, contains an immunoglobulin domain, similarity to human SLAM
EMCLV159R	**MC163R**	32,7	164189…165664	1476	putative inhibitor of apoptosis, contains a semaphorin domain

MOCV but only a single copy in EMCLV. An example is the IL18-BP gene family with MC051L, MC053L and MC054L corresponding to EMCLV145R. EMCLV is predicted to code for 20 proteins that are unique to this virus (ORFs illustrated in black in Fig. 2). While 18 of these translated ORFs did not produce any hit when searched for in public sequence, structure or domain databases, two – EMCLV007 and EMCLV158 – showed unambiguous similarity to the mammalian proteins SECTM1 and IGFLR1, respectively. For better understanding, detailed comparison of the gene set of EMCLV with MOCV was subdivided into three sections: genes encoding proteins needed for transcription and replication (I), modulation of cell biology (II) and modulation of the immune system (III).

(I) Comparison of the genes associated with the transcription and replication machinery

The genes coding for the polymerases and associated cofactors as well as IMV, EEV and core proteins are entirely conserved among EMCLV and MOCV. However, the homology levels are surprisingly low when compared with other poxvirus genera. An artificial amino acid sequence produced by the concatenation of seven gene products of the highly conserved core region that is frequently used for the comparison of different poxvirus species (J6R, H4L, D1R, D5R, A7L, A10L and A24R in VACV Copenhagen [24] corresponding to MC079R, MC085L, MC090R, MC094R, MC110L, MC113L and MC129R in MOCV and EMCLV074R, EMCLV082L, EMCLV087R, EMCLV091R, EMCLV107L, EMCLV110L and EMCLV126R in EMCLV) revealed only 81.4% amino acid sequence identity between EMCLV and MOCV. Instead, VARV major (strain GBR_harv, GenBank acc. no. DQ_441444) and VACV (strain Western Reserve GenBank acc. no. NC_006998.1), two members of different species of the genus *Orthopoxvirus*, share 98.4% amino acid identity in the same alignment.

Another interesting finding was that three ORFs from the EMCLV conserved core differ considerably in length from those of MOCV and might raise the impression of truncation on first sight. For example, the RPO30 subunit of the RNA polymerase (EMCLV037) consists of only 203 aa whereas in MOCV it contains 440 aa on average. Closer inspection with other RPO30 orthologues reveals that most poxviruses encode an RPO30 protein of only 200–260 aa while the MOCV RPO30 is extensively longer at the C-terminus. Interestingly, EMCLV shares the short protein length of RPO30 with Eastern kangaroopox virus (EKPV), WKPV and several avipoxviruses (APVs). A comparable situation applies to EMCLV109L, which codes for a viral membrane protein. Again, EMCLV shares a short protein version with WKPV (75 aa in both cases) whereas the orthologue in MOCV is substantially longer spanning 113–128 aa. Also, the DNA helicase (EMCLV120) is significantly shorter than its MOCV orthologue. However, in this case most poxviruses including the kangaroopoxviruses (KPVs) and APVs encode proteins of an average size of 460–500 aa. A predicted DNA helicase of 562 aa in EMCLV is larger than average poxviral helicases and MOCV again encodes the largest protein with 694–800 aa. Recombination events among MOCV and KPVs were already identified independently by Bennett *et al.* [25] and Sarker *et al.* [26]. A common ancestor for MOCV, KPVs and APVs was proposed based on these findings and the close clustering in phylogenetic trees. Our observations support a common evolutionary history of EMCLV, MOCV, KPVs and possibly also APVs.

(II) Cell biology modulators in EMCLV and MOCV

Of the MOCV genes implicated in the modulation of cellular pathways, two members of the SLAM gene family (MC002L, MC161R and MC162R) are conserved in EMCLV (EMCLV001L and EMCLV158R). An inactivator of the retinoblastoma protein (pRb), which interacts with cell proliferation control can be identified in both MOCV (MC007L) and EMCLV (EMCLV006L) [27]. Also present (EMCLV29L) is a homologue of MC026L, an APC11-like protein with a RING motif domain, which is conserved in several G+C rich poxviruses including Nile and saltwater crocodilepox virus, parapoxviruses, squirrelpox virus and avipoxviruses [28] and interferes with cell-cycle control [29]. EMCLV156L is an orthologue of MC159L, which was found to act as a viral FLIP protein to inhibit apoptosis [30, 31]. MC163R also acts as inhibitor of apoptosis with EMCLV159R as an orthologue [32].

In contrast, the MC066L selenoprotein, which is able to inhibit UV induced cell death, has no equivalent in EMCLV [33].

(III) Immune modulators in EMCLV and MOCV

Several MOCV genes are known to be implicated in immune modulation in the host and are present in EMCLV. MC005L (EMCLV004L) inhibits NF-κB activation [34]. The IL18-BP gene family is represented by MC051L, MC053L, MC054L and EMCLV145R. MC080R (EMCLV077R) is a MHC class I homologue and interferes with antigen presentation on the cell surface [35]. However, no orthologues were detected in EMCLV for MC132R and MC160L, further inhibitors of NF-κB activation [36, 37], and neither for MC148R, a CC-chemokine homologue [38].

Careful analysis of the EMCLV genes interacting with the host immune response led to the identification of three predicted proteins that merit further attention. One is the IL18-BP with interesting implications on MOCV evolution. The other two proteins are homologues of mammalian proteins implicated in immune-signalling pathways – SECTM1 and IGFLR1 – that to the best of our knowledge have not been described in any virus so far.

IL18-binding protein

Genes coding for IL18-BP were found in both EMCLV (EMCLV145R) and MOCV (MOCV054L). In contrast to all other conserved genes, EMCLV146R and MOCV054L do not follow the syntenic gene order within the two virus species and also differ in transcription orientation (L/R). MC054L localizes to the left boundaries of the conserved core region in the MOCV genome whereas EMCLV145R is situated in the right variable region close to the genome end. It is also noteworthy that EMCLV145R has no other orthologue in the EMCLV genome while MOCV codes for three related proteins MC051L, MC053L and MC054L that form a gene family (although only MC054L was shown to have IL18-binding activity *in vitro* [39]. The most compelling difference of MC054 and EMCLV145, however, lies in the number of protein domains. Poxviral IL18-BPs contain an N-terminal signal peptide and the actual IL18-binding domain. MC054 is unique among poxviral IL18-BPs because of a cleavable C-terminal tail [40]. Alignment of the protein sequences of MC054 and EMCLV145 reveals that the characteristic C-terminal tail structure is not present in EMCLV146 (Fig. 4).

Secreted and transmembrane protein 1

EMCLV007L was found to encode a homologue of equine SECTM1 (eqSECTM1) isoform X2 (NCBI RefSeq. XP_005597056.1; blastp search E-value $7e^{-50}$). SECTM1 is a transmembrane glycoprotein with a soluble secreted and a transmembrane form. It contains a signal peptide, a larger extracellular domain with an Ig-fold, a single conserved transmembrane domain and a smaller intracellular domain [41]. Crystal structures of the protein are not yet available but Bianchetti and colleagues [42] proposed a structure prediction for the Ig-fold in the extracellular domain of SECTM1. They also identified two highly conserved cysteines in the extracellular domain that are predicted to lead to an atypical disulfide bridge and a highly conserved amino acid motif of $G_{112}Y_{115}W_{117}L_{119}G_{121}Q_{123}$. Though SECTM1 structures vary across species, all the listed characteristics are conserved in eqSECTM1. The EMCLV007 homologue covers the entire eqSECTM1 protein and contains the two cysteines in the extracellular domain, the $G_{112}Y_{115}W_{117}L_{119}G_{121}Q_{123}$ motif and the transmembrane domain (Fig. 5). However, apart from the conserved regions and motifs, the overall amino acid identity of the two proteins is low (36.7 %). Several models for the prediction of protein structure and transmembrane domains lead to contradictory results with regard to number and location of potential additional transmembrane helices and the resulting protein topology for EMCLV007. TMHMM server predicts two transmembrane domains for EMCLV007 (aa 56–78 and 171–192) while the Phobius algorithm identifies aa 172–193 and 214–236 in EMCLV007 as transmembrane proteins above the prediction threshold. The 'DAS' transmembrane prediction server returns three transmembrane domains at aa positions 57–78, 173–193 and 213–234. All of these predictions would lead to conformational changes of EMCLV007 with disruption of either the extracellular or cytoplasmic domain. *In vitro* experiments will be needed to determine the actual protein topology of EMCLV007 for further implications of its function.

Insulin growth factor-like family receptor 1

BLASTP search for the translation of EMCLV157R revealed a high similarity to equine IGFLR1 (eqIGFLR1; NCBI RefSeq. XP_014693751.1; BLASTP search E-value $2e^{-47}$). IGFLR1 is a member of the tumor necrosis factor receptor family and shows a similar general architecture with a signal peptide, an extracellular domain that mediates ligand binding, a transmembrane domain anchoring the receptor to the cell membrane and an intracellular domain for signal transduction with similarities to a cytoplasmic death domain (Fig. 6) [43]. The extracellular domain contains ten highly conserved cysteines for five predicted disulfide bonds. Nine of these cysteines are conserved in equine IGFLR1 [43]. However,

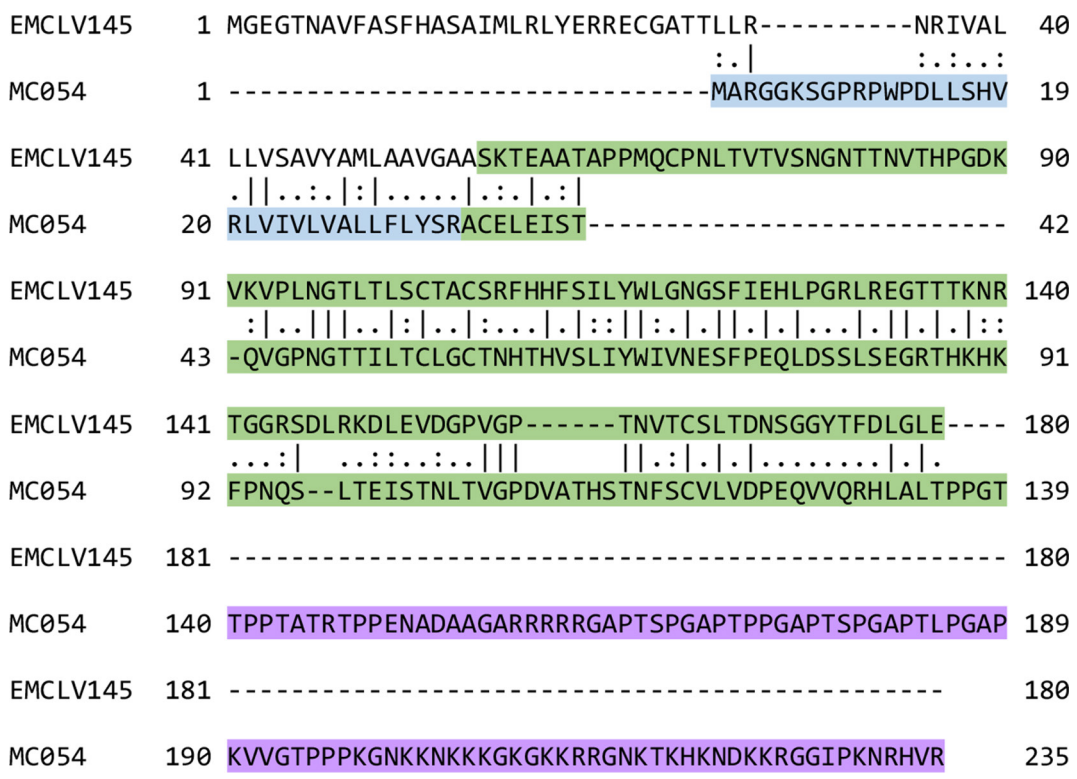

Fig. 4. Protein alignment of EMCLV145 and MC054, viral homologues of the IL18-BP. Signal peptides are marked in blue; the IL18-binding domain is highlighted in green. Purple indicates the C-terminal tail of the IL18-BP of MOCV.

in contrast to EMCLV007, the viral homologue eqIGFLR1 covers only a small fraction of the protein including only the signal peptide and the extracellular ligand-binding domain (Fig. 6). The 100% amino acid identity stretch from residue 19 to 93 in EMCLV157 to eqIGFLR1 is an indicator of a recent viral acquisition of this protein from the equine host genome.

DISCUSSION

Horses and other equines are susceptible to a surprisingly large number of poxviruses that cause exanthemic lesions of the skin but only a few of the respective causative agents are well understood and characterized. Beside the universal interest of basic research there are many more reasons why any skin disease of poxviral etiology in equines merits further evaluation. A major point of interest was caused by the controversial debate of the role of HPXV in the evolution of VACV and its historical role in the development of Jenner's smallpox vaccine(s). Many questions are left unanswered as HPXV has become very rare or even extinct in modern times [44] and only a single full-length genome derived from equine samples was determined so far [4]. Some smallpox vaccines show unambiguous genetic traits of HPXV like the Dryvax and IOC strain [45, 46]. The Mulford 1902 smallpox vaccine strain is most closely related to HPXV but carries a deletion pattern found only in current VACV strains [47]. If HPXV still exists in certain biological niches close monitoring of skin affections in horses and donkeys may help to detect further extant strains, which would contribute substantially to the understanding of the enthralling evolutionary history of VACV and other orthopoxviruses.

The skin lesions in this report were caused by a virus related to molluscum contagiosum virus, which is no less interesting than HPXV. After the eradication of variola virus in the human population MOCV was considered the only poxvirus with humans as its sole host species [18, 48, 49]. However, this silently ignores that molluscum contagiosum-like skin lesions were sporadically reported in several animal species including horses [16, 50, 51], kangaroos [52, 53], chimpanzees [54] and sea lions [55]. Without modern diagnostic techniques older case reports could only yield evidence of a similarity to MOCV based on gross pathology, histology and electron microscopy making it impossible to delineate any conclusions on the true identity of the respective viruses and their host range and zoonotic potential. Thompson and colleagues demonstrated in 1998 [17] a close relationship of equine and human molluscum contagiosum virus by *in situ* hybridization assays, however, the question whether it was the same virus or another species could not be answered. A few years ago, Fox and colleagues [15] provided the first genetic evidence that molluscum contagiosum in horses might be caused by a virus related to but different from MOCV based on the low nucleotide identity in a gene of the conserved core. Unfortunately, the sequence data was limited to a 630 bp fragment of the RPO147 RNA polymerase subunit gene. Metagenomics

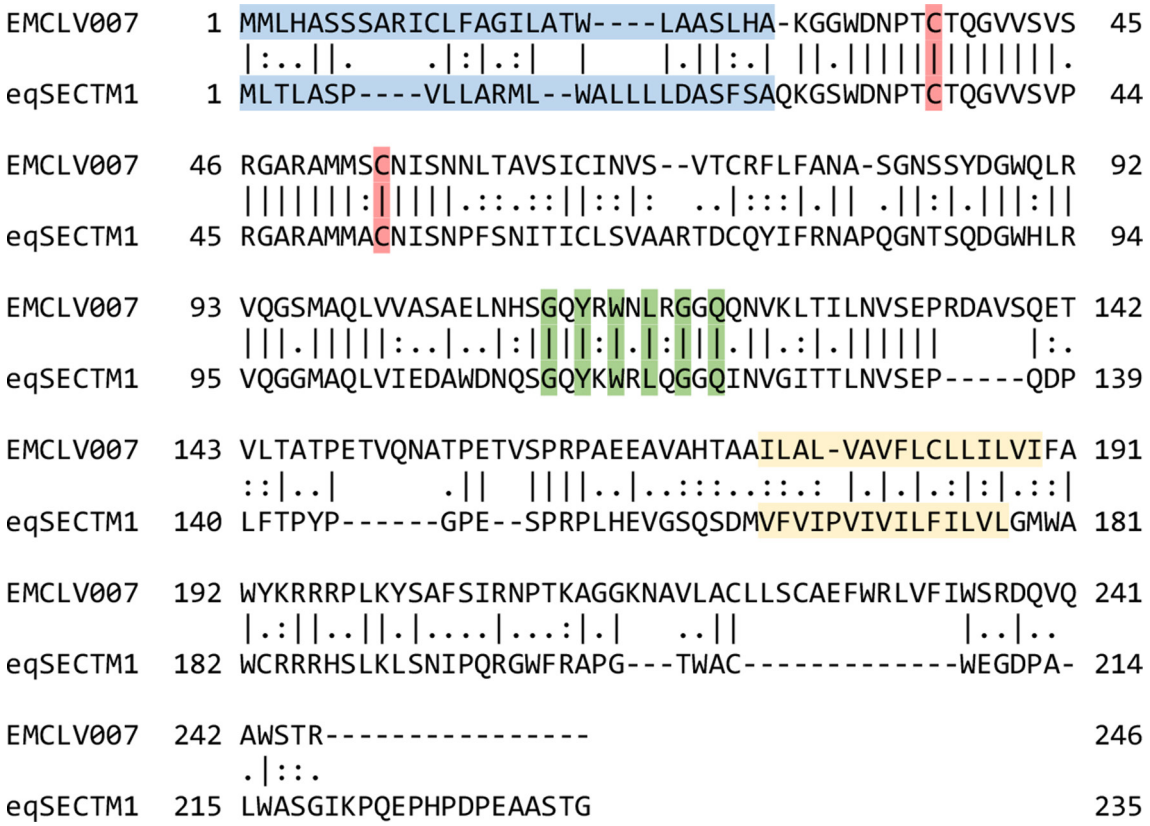

Fig. 5. Protein alignment of EMCLV007 and equine SECTM1 (RefSeq accession XP_005597056.1). Signal peptides are marked in blue. The two conserved cysteines in the extracellular domain are highlighted in red and the conserved G112Y115W117L119G121Q123 motif in green. Yellow illustrates the transmembrane domain.

studies in straw-coloured fruit-bat colonies in Africa revealed sequence information corresponding to 23 poxvirus genes, most closely related to MOCV, without any link to clinical symptoms [56]. These bat-derived sequence fragments showed a higher deviation from MOCV than EMCLV, however, a full-length genome could not be obtained. In 2017 within a very short time frame two independent groups characterized the genome sequences of EKPV and WKPV, respectively [25, 26]. These were the first full-length genome sequences available from molluscum contagiosum-like skin affections in animals. While sharing some of the characteristics of MOCV like the high G+C content, a probable common ancestor and certain gene sets that were likely derived from recombination events with MOCV, EKPV and WKPV were found to be different from all other poxvirus genera characterized so far and proposed to be members of a novel genus tentatively named *Thylacopoxvirus* [25]. In contrast, the virus found to cause the lesions in the horse from Tanzania is a clear member of the genus *Molluscipoxvirus* highlighted by the stringent synteny and similar gene complement with MOCV. EMCLV is therefore the first molluscum contagiosum-like virus found in animals. The fact that all host-derived genes [EMCLV007L (SECTM1), EMCLV077R (MHC-I homologue), EMCLV145R (IL18-binding protein), EMCLV147R (3-beta-hydroxysteroid dehydrogenase) and EMCLV157R (IGFLR1)] show close relatedness to the respective genes in horses and donkeys emphasizes the adaptation to equine hosts. Observations from caretakers and farmers that equine molluscum contagiosum may affect several horses in the same stables or herds but transmission to staff was never reported even under poor hygienic conditions support the hypothesis that EMCLV is highly adapted to equids and zoonotic potential is either very low or absent (personal communication, E. Stegmaier, K. v. Schlippenbach).

The benign nature of molluscum contagiosum with its limited impact on livestock in the case of EMCLV and immunocompetent persons with regard to MOCV is one of the reasons why the disease and their causative viruses are rather neglected study objects. However, in immunocompromised individuals MOCV may develop severe and extensive lesions that can be difficult to manage. A broader understanding of molluscipoxviruses is therefore highly desirable. The first genome sequence of a MOCV subtype 1 was published by Senkevich *et al.* [49] and further genome sequences were added only recently with four sequences by López-Bueno *et al.* [57] including the first MOCV subtype 2 full-length genome, and 11 more full-length genomes by Zorec *et al.* [48] and Huang *et al.* [58]. With this substantially broader genetic information several interesting findings could be made when

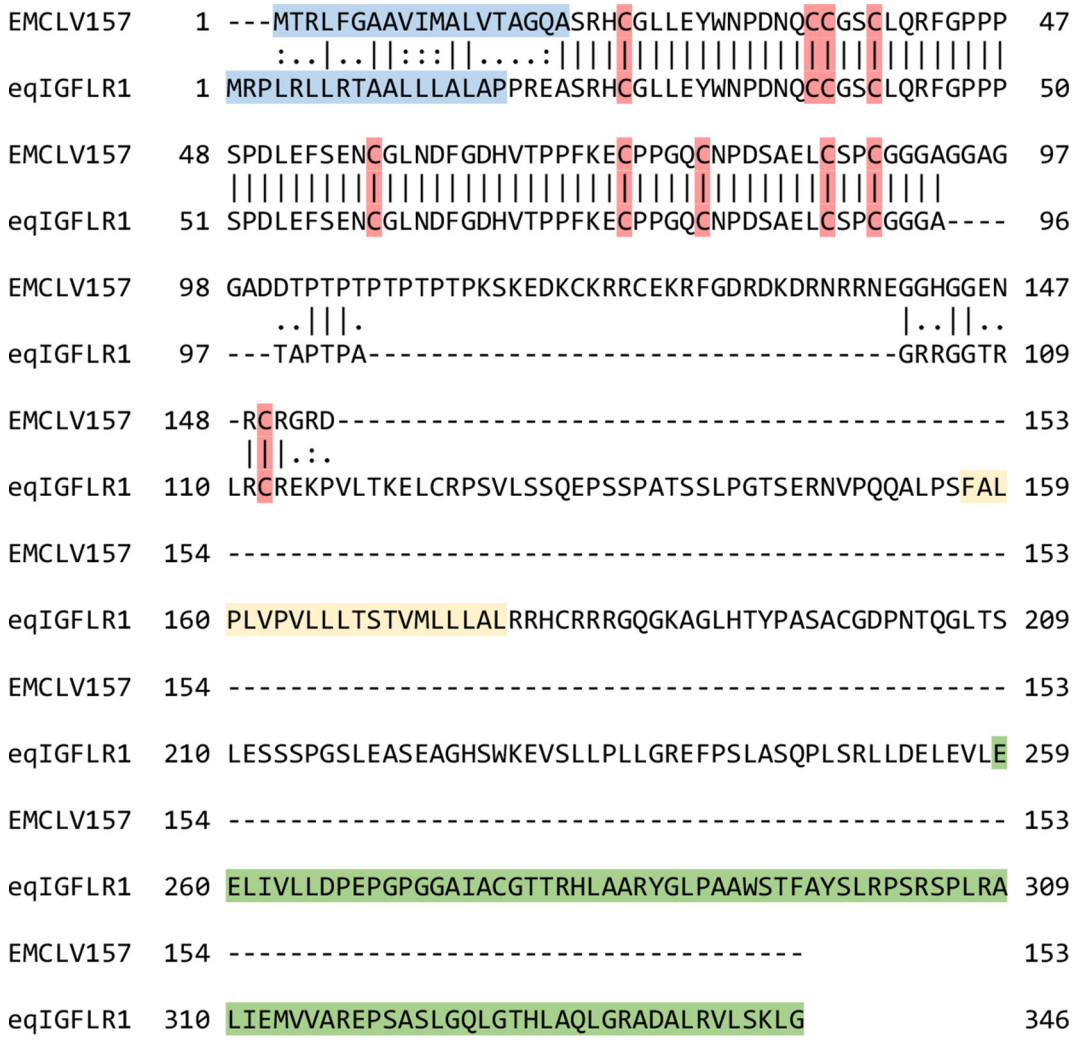

Fig. 6. Protein alignment of EMCLV157 and equine IGFLR1 (RefSeq accession XP_014693751.1). Signal peptides are marked in blue. Conserved cysteines in the extracellular domain are highlighted in red. Yellow illustrates the transmembrane domain while green marks the putative death domain in the intracellular C-terminal part of IGFLR1.

MOCV and the newly described EMCLV were compared. Firstly, the tandem repeat patterns of EMCLV in the inverted terminal repeats deviate considerably from those in MOCV, which are far more complex and diverse. Secondly, though showing a syntenic genome and similar gene set as MOCV, the overall amino acid identity levels of MOCV and EMCLV are surprisingly low even when analysing only conserved genome regions. An alignment of seven concatenated protein sequences from the conserved core frequently used for phylogenetic comparisons [24] demonstrated that the clinically similar MOCV and EMCLV show a considerably lower amino acid identity than the clinically and biologically diverse orthopoxviruses VARV and VACV. When the first MOCV subtype 2 full-length genome sequence was published the authors found the two subtypes sufficiently divergent to consider them two separate species [57]. However, they also argued that in light of the similar clinical and biological features there is no sense in classification of two different species for the MOCV subtypes 1 and 2. Based on the genetic findings and following this line of discussion, EMCLV can be confidently proposed as a novel species in the genus *Molluscipoxvirus*.

A third point of interest addresses the different strategies poxviruses evolved to cope and interact with the host immune system. MOCV differs from many other poxviruses in its highly potent immunosuppression, which permits the virus to easily persist for weeks or even months rather than causing acute illness that ends in either death of the host or rapid viral clearance [31]. Equine molluscum clinically shows an equal ability of persistence and EMCLV apparently developed both similar and new strategies to achieve successful host immune evasion when compared to MOCV. An intriguing example for a similar strategy is the IL18-BP of EMCLV. Viral homologues of mammalian IL18-BPs are encoded by many members of the *Poxviridae* across different species and genera [59–61]. IL18 is a pro-inflammatory cytokine with the ability

to induce INF-gamma, a critical factor of antiviral response [62]. In the mammalian host IL18-BP regulates the levels of soluble (and therefore active) IL18 whereas poxviruses use homologues of these proteins to modulate the host immune response. While on first sight it may not be surprising that both EMCLV and MOCV encode an IL18-BP homologue it certainly is striking that the respective ORFs are the only genes that are not syntenic among the two virus species in the entire genomes. It is also noteworthy that EMCLV145R has no other homologue in the EMCLV genome while MOCV codes for three related proteins MC051L, MC053L and MC054L that form a gene family although only MC054L was shown to have IL18-binding activity *in vitro* [63]. The most compelling difference of MC054 and EMCLV145, however, lies in the respective protein structures. MC054 was found to be special among poxviral IL18-binding proteins because it does not only consist of the conserved signal peptide and the actual IL18-binding domain but contains also an additional C-terminal tail that is not present in any other viral homologue [40]. Alignment of the protein sequences of MC054 and EMCLV145 reveals that this characteristic tail structure is not present in the equine virus (Fig. 4). All these findings strongly indicate that the common ancestor of EMCLV and MOCV did not contain a gene for IL18-BP and that both viruses acquired it by independent gene transfer events from their respective hosts. This emphasizes the obvious importance of IL18-BPs for members of the genus *Molluscipoxvirus* specifically and poxviruses in general to deal with the host immune defense.

Examples for complementary immune evasion strategies of EMCLV can be found among the 20 predicted genes unique to this virus. EMCLV007L and EMCLV157R code for viral homologues of mammalian proteins involved in immune signalling that to the best of our knowledge have not been described in any other virus so far.

The first one is SECTM1, a poorly studied type I transmembrane glycoprotein that was found to exist in two forms – a soluble secreted and a transmembrane form – that gave the protein its name. The transmembrane form of the protein seems to be retained in the Golgi apparatus [41] but the soluble form was shown to be a ligand of CD7, a co-receptor found on T cells, natural killer (NK) cells and some B cells [64]. Interaction of hSECTM1 with CD7 plays a role in T-cell and NK-cell activation and homeostasis as a co-mitogenic signal after initial priming by MHC-mediated TCR/CD3 signalling and is strongly augmented in the presence of CD28 [65]. SECTM1 expression is regulated by IFN in the mode of an early response gene [66] and interaction of CD7, CD28 and SECTM1 in human cells and tissues leads to the production of IL-2, IFN-γ and the upregulation of cell-surface molecules like CD25, CD54 and CD69 in T cells and NK cells [65–67]. However, SECTM1 seems to have major species-specific differences. While humans harbour only one gene for SECTM1, many species like mice, rats and cattle encode two proteins – SECTM1a and SECTM1b – with different functional profiles [42]. Murine SECTM1a and SECTM1b were shown to have a different receptor-binding behaviour promoting T-cell activation for mSECTM1a and inhibitory effects on T cells for mSECTM1b [68]. A single SECTM1 gene was annotated in the equine genome so far coding for a SECTM1 protein that exists in two highly similar isoforms. It is highly likely that EMCLV007 is interfering with T- and NK-cell activation, however, understanding of the exact mechanism needs future experimental data about the interaction of eqSECTM1 with CD7 and CD28 in equine immune cells and information on the amino acid residues involved in SECTM1 for interaction with CD7.

The role of the viral IGFLR1 homologue (EMCLV157) is more obvious from sequence data alone. Mammalian IGFLR1 is a member of the tumor necrosis factor receptor family and binds IGF-like protein 1 (IGFL1). In mice IGFLR1 is most abundantly expressed on T cells. In both human and mouse skin models IGFLR1 was induced under inflammatory conditions suggesting that it might act as a regulatory element in T cell response in the skin [43]. The EMCLV homologue of IGFLR1 spans only the N-terminal part of the protein with the signal peptide and the ligand-binding extracellular domain where amino acid identity reaches 100%. This is a strong implication that the viral protein represents a soluble decoy receptor molecule that acts as a cytokine scavenger to block activity of the ligand – in this case IGFL1. Examples of such cytokine scavengers or decoy receptors are numerous among poxviruses including VACV protein B8, which binds IFN-gamma [69], the cytokine response modifier family CrmA-CrmE, which represent TNF decoy receptors [70] and VACV B15, which binds IL-1beta [71].

The identification of two new viral genes for host immune evasion is a rare event and underlines the benefits of discovering novel virus species. During the last decade the growing technical capabilities of sequencing technology has led to an exponential increase in genomic information and facilitated the discovery of a multitude of novel species of all kingdoms of life. Simultaneously, criticism has arisen about the trend of sheer collection and publication of genomics data while neglecting the experimental work to characterize these new species and their actual meaning and importance in the respective ecosystems [72]. However, equine molluscum contagiosum-like virus is a good example that it is still worth being vigilant for novel virus species anywhere as even the hundredth member may still provide unexpected characteristics and features that triggers important future experimental work. In the era of growing application of monoclonal antibodies to treat immune-mediated diseases and cancer the plethora of viral immune modulators represent a treasure box of potential useful drugs that are worth exploring in depth.

METHODS

Sample collection

From a horse with molluscum contagiosum-like skin lesions two nodules were surgically removed. One was fixed in 4% formalin for histopathologic and electron microscopic analyses while the other was stored in NaCl for viral isolation and DNA extraction.

Histopathology

The fixed material was processed for histopathology according to standard laboratory protocols. Out of the paraffin block 4 µm sections were cut and stained with hematoxylin-eosin (HE).

In situ hybridization

For *in situ* hybridization a PCR-generated, digoxigenin-labelled 300 bp probe of MOCV DNA was used. Briefly, 3 µm sections of the skin biopsy were dewaxed before proteolytic digest by proteinase K (20 µg ml^{-1} for 30 min at 37 °C). After refixation in absolute ethanol, sections were air-dried. Hybridization was carried out at 40 °C overnight with approximately 100 ng probe. After repeated washing, colour reaction was performed during 30 min, and tissues were sealed with aqueous mounting medium [73].

Electron microscopy

Freeze-thawed fragments of the skin biopsy (2×2 mm^2) were incubated for 2 h at 4 °C in 5% glutaraldehyde buffered with 0.1 M cacodylate buffer to pH 7.4, washed three times with 0.1 M cacodylate buffer with 0.2 M sucrose and post-fixed for 2 h in 1% OsO4 in 0.1 M cacodylate buffer with 0.2 M sucrose. After washing samples were dehydrated in graded ethanol and embedded in Agar 100 Resin (Agar Scientific, Stansted, UK). The blocks were sectioned on an Ultracut microtome, sections were stained for 20 min with uranyl acetate and lead citrate [74] and examined with a transmission electron microscope LIBRA 120 (Carl Zeiss, Oberkochen, Germany). In parallel negative staining was performed on a 10 µl aliquot of the homogenated skin, which was stained for 5 s with the same volume of 2% phosphotungstic acid on formvar-carbon coated copper grids, dried, and examined.

Cell-culture isolation

Papular lesion material was homogenized with the Fast-PrepTM System (Q-biogene, Heidelberg, Germany) using lysing matrix A 2 ml tubes (MP Biomedicals). After a brief centrifugation step (1 min at 1000 ***g***), the supernatant of homogenized lesions was used for further analyses. For virus isolation, 0.4 ml of the supernatant was inoculated onto monolayers of MA104 (African green monkey kidney cell line, ATCC CRL-2378.1) and UcP-R (fetal bovine oesophageal primary cells, kind gift of Mathias Büttner, Oberschleißheim). After 1 h at 37 °C, the inoculum was removed, and the cells were washed five times with 5 ml of minimum essential medium with penicillin/streptomycin (MEM-AB) before addition of 5 ml of MEM-AB containing 1% fetal calf serum. Cells were monitored daily for cytopathic effects and two blind passages were conducted with infected cells after two freeze-thaw cycles as inoculum for the next passage.

Reporter assays for viral replication

Experimental conditions with homogenized original sample material of EMCLV, MVA as an orthopoxvirus control (m.o.i. 0.01) and mock were set up in triplicate using human HEK293 cells (ATCC CRL1573) and plasmid p240 (pRB21-pE/L-FF luciferase) following the procedure described in [20] Firefly luciferase signals were read in a VIKTOR multireader instrument (Perkin Elmer).

DNA extraction and diagnostic PCR

Total DNA was directly isolated from 0.1 ml supernatant of homogenized skin nodules by addition of 0.1 ml of lysis-buffer containing proteinase K (1 mg ml^{-1}) and Tween 20 (0.5%). After incubation at 56 °C for 3 h and enzyme inactivation at 95 °C for 20 min, DNA was extracted with standard phenol/chlorophorm/isoamylalcohol procedure. A pan-poxvirus PCR [19] was used for detection of poxviral DNA.

Library preparation, high-throughput sequencing and assembly

The library for the Illumina sequencing was prepared using the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (New England BioLabs) without previous shearing of the DNA and according to the protocol for use with inputs ≤100 ng (no size selection). The sequencing was performed on an Illumina MiSeq device using the MiSeq Reagent Kit v2 (500-cycles). Quality metrics were obtained using FastQC [75].

The library for the nanopore sequencing was prepared using the Ligation Sequencing Kit 1D SQK-LSK108 (Oxford Nanopore Technologies) and sequenced on a MinION device for 24 h. Basecalling was done separately using Guppy v3.0.2 with the high accuracy basecalling model. Quality metrics were obtained using NanoStat [76].

Raw sequence read data was submitted to NCBI under Bioproject ID PRJNA579182 together with sample information under BioSample ID SAMN13105282. Before assembly, filtering was applied to the nanopore reads to consider only those reads with a minimum length of 5000 bases or a minimum base quality of 10 at a minimum length of 700 bases. Preselected quality reads were then assembled using flye [21]. A similarity search of the resulting contigs against the NCBI nucleotide database (nr) was performed using BLASTN [77, 78] to separate virus-related contigs from background-related contigs (equine, bacterial, etc.). The virus-related contigs were then polished with the Illumina reads in three iterations using pilon [22]. The final genome sequence was deposited in GenBank with accession number MN339351.

Phylogenetic analysis

For phylogenetic analyses, a MAFFT-alignment of the EMCLV genome with 38 representative poxvirus full-length genome sequences was generated using the FFT-NS-2 algorithm [79]. GBlocks was used to extract conserved regions using standard parameters [80] leading to a gap-free alignment of 34214 nt for the 39 different chordopoxviruses (Fig. 2a) and 151 694 nt for the molluscum alignment (Fig. 2b). A Maximum Parsimony distance matrix using the accelerated transformation (ACCTRAN) was calculated and plotted as dendrogram using R (packages Phangorn and Phytools). Bootstrap values were calculated with 1000 replicates.

Genome annotation

The Genome Annotation Transfer Utility tool (GATU) from the Viral Bioinformatics Resource Centre [81] was used to screen the EMCLV genome for common ORFs with the MOCV reference strain (GenBank acc. no. NC_001731). Search parameters included a minimum ORF size of 40 codons with maximum overlap of 25%. ORFs were annotated as conserved MOCV genes from the reference genome with amino acid similarities as low as 25% if database searches gave a clear hit for MOCV or characteristic amino acid residues like cysteines were conserved. After the initial round ORFs were manually checked for correct start and stop codons. Subsequently, intergenic regions were scanned for further ORFs. Annotation as a potential gene in the second search round was based on localization, ORF size, overlap with other genes, presence of typical poxvirus promoter and termination motifs, as well as similarity to other genes found in public databases. All ORF translations were analysed in BLASTP [77]. Hypothetical proteins were also checked for conserved domains or secondary structures with CDD [82] and HHPred [83, 84] as well as common protein motifs to assist functional protein identification by InterPro, ScanProsite and Uniprot [85–87]. Protein alignments were made with EMBOSS Needle pairwise alignment [88]. Transmembrane domain and protein topology predictions were made with TMHMM server v2.0 [89], Phobius [90] and DAS [91]. The EMCLV genome map was created with SnapGene Viewer software (GSL Biotech LLC).

Funding information
This work received no specific grant from any funding agency.

Acknowledgements
We thank Claudia Kahlhofer, Aileen Lorber, Gudrun Zöller, Josua Zinner and Willi Eggeling for excellent technical assistance.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
Skin samples from the horse were obtained for diagnostic purposes and to enable adequate disease prevention strategies in the respective herd. Therefore no ethical approval was required for the use of these samples. Retrieval of the skin biopsy was performed according to the Code of good veterinarian practice and respecting international and national guidelines for the care and use of animals.

References

1. Haller SL, Peng C, McFadden G, Rothenburg S. Poxviruses and the evolution of host range and virulence. *Infection, Genetics and Evolution* 2014;21:15–40.
2. Damaso CR. Revisiting Jenner's mysteries, the role of the Beaugency lymph in the evolutionary path of ancient smallpox vaccines. *Lancet Infect Dis* 2018;18:e55–e63.
3. Esparza J, Schrick L, Damaso CR, Nitsche A. Equination (inoculation of horsepox): an early alternative to vaccination (inoculation of cowpox) and the potential role of horsepox virus in the origin of the smallpox vaccine. *Vaccine* 2017;35:7222–7230.
4. Tulman ER, Delhon G, Afonso CL, Lu Z, Zsak L *et al*. Genome of horsepox virus. *J Virol* 2006;80:9244–9258.
5. Borges IA, Reynolds MG, McCollum AM, Figueiredo PO, Ambrosio LLD *et al*. Serological Evidence of *Orthopoxvirus* Circulation Among Equids, Southeast Brazil. *Front Microbiol* 2018;9:402.
6. Brum MCS, Anjos BLD, Nogueira CEW, Amaral LA, Weiblen R *et al*. An outbreak of orthopoxvirus-associated disease in horses in southern Brazil. *J Vet Diagn Invest* 2010;22:143–147.
7. Campos RK, Brum MCS, Nogueira CEW, Drumond BP, Alves PA *et al*. Assessing the variability of Brazilian vaccinia virus isolates from a horse exanthematic lesion: coinfection with distinct viruses. *Arch Virol* 2011;156:275–283.
8. Franke A, Kershaw O, Jenckel M, König L, Beer M *et al*. Fatal cowpox virus infection in an aborted foal. *Vector Borne Zoonotic Dis* 2016;16:431–433.
9. Pfeffer M, Burck G, Meyer H. Kuhpockenviren in Deutschland: Eine analyse von 5 Fällen AUS dem Jahr 1998. *Berl Munch Tierarztl Wochenschr* 1999;112:334–338.
10. Kaminjolo JJ, Muhammed SI, Johnson LW, Berger J. Uasin Gishu skin disease of horses in Kenya 1975.
11. Kaminjolo JS, Nyaga PN, Gicho JN. Isolation, cultivation and characterization of a poxvirus from some horses in Kenya. *Zentralblatt Veterinarmedizin Reihe B* 1974;21:592–601.
12. Airas N, Hautaniemi M, Syrjä P, Knuuttila A, Putkuri N *et al*. Infection with possible novel parapoxvirus in horse, Finland, 2013. *Emerg Infect Dis* 2016;22:1242–1245.
13. Osadebe LU, Manthiram K, McCollum AM, Li Y, Emerson GL *et al*. Novel poxvirus infection in 2 patients from the United States. *Clin Infect Dis* 2015;60:195–202.
14. Cooley AJ, Reinhard MK, Gross TL, Fadok VA, Levy M. Molluscum contagiosum in a horse with granulomatous enteritis. *J Comp Pathol* 1987;97:29–34.
15. Fox R, Thiemann A, Everest D, Steinbach F, Dastjerdi A *et al*. Molluscum contagiosum in two donkeys. *Vet Rec* 2012;170:649.2–64649.
16. Rahaley RS, Mueller RE. Molluscum contagiosum in a horse. *Vet Pathol* 1983;20:247–250.
17. Thompson CH, Yager JA, Van Rensburg IB. Close relationship between equine and human molluscum contagiosum virus demonstrated by in situ hybridisation. *Res Vet Sci* 1998;64:157–161.
18. Chen X, Anstey AV, Bugert JJ. Molluscum contagiosum virus infection. *Lancet Infect Dis* 2013;13:877–888.
19. Li Y, Meyer H, Zhao H, Damon IK. Gc content-based pan-pox universal PCR assays for poxvirus detection. *J Clin Microbiol* 2010;48:268–.
20. Sherwani S, Blythe N, Farleigh L, Bugert JJ. New method for the assessment of molluscum contagiosum virus infectivity. *Methods Mol Biol* 2012;890:135–146.
21. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019;37:540–546.
22. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A *et al*. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;9:e112963.
23. Senkevich TG, Koonin EV, Bugert JJ, Darai G, Moss B. The genome of molluscum contagiosum virus: analysis and comparison with other poxviruses. *Virology* 1997;233:19–42.
24. Emerson GL, Nordhausen R, Garner MM, Huckabee JR, Johnson S *et al*. Novel poxvirus in big brown bats, northwestern United States. *Emerg Infect Dis* 2013;19:1002–1004.
25. Bennett M, Tu S-L, Upton C, McArtor C, Gillett A *et al*. Complete genomic characterisation of two novel poxviruses (WKPV and EKPV) from Western and eastern grey kangaroos. *Virus Res* 2017;242:106–121.
26. Sarker S, Roberts HK, Tidd N, Ault S, Ladmore G *et al*. Molecular and microscopic characterization of a novel eastern grey kangaroopox virus genome directly from a clinical sample. *Sci Rep* 2017;7:16472.
27. Mohr S, Grandemange S, Massimi P, Darai G, Banks L *et al*. Targeting the retinoblastoma protein by MC007L, gene product of the molluscum contagiosum virus: detection of a novel virus-cell interaction by a member of the poxviruses. *J Virol* 2008;82:10625–10633.

28. Sarker S, Isberg SR, Milic NL, Lock P, Helbig KJ. Molecular characterization of the first saltwater crocodilepox virus genome sequences from the world's largest living member of the Crocodylia. *Sci Rep* 2018;8:5623.
29. Mo M, Fleming SB, Mercer AA. Cell cycle deregulation by a poxvirus partial mimic of anaphase-promoting complex subunit 11. *Proc Natl Acad Sci USA* 2009;106:19527–19532.
30. Schmotz C, Uğurlu H, Vilen S, Shrestha S, Fagerlund R *et al*. MC159 of molluscum contagiosum virus suppresses autophagy by recruiting cellular SH3BP4 via an SH3 domain-mediated interaction. *J Virol* 2019;93.
31. Shisler JL. Immune evasion strategies of molluscum contagiosum virus. *Adv Virus Res* 2015;92:201–252.
32. Coutu J, Ryerson MR, Bugert J, Brian Nichols D. The molluscum contagiosum virus protein MC163 localizes to the mitochondria and dampens mitochondrial mediated apoptotic responses. *Virology* 2017;505:91–101.
33. Shisler JL, Senkevich TG, Berry MJ, Moss B. Ultraviolet-induced cell death blocked by a selenoprotein from a human dermatotropic poxvirus. *Science* 1998;279:102–105.
34. Brady G, Haas DA, Farrell PJ, Pichlmair A, Bowie AG. Molluscum contagiosum virus protein MC005 inhibits NF-κB activation by targeting NEMO-Regulated IκB kinase activation. *J Virol* 2017;91.
35. Harvey IB, Wang X, Fremont DH. Molluscum contagiosum virus MC80 sabotages MHC-I antigen presentation by targeting tapasin for ER-associated degradation. *PLoS Pathog* 2019;15:e1007711.
36. Brady G, Haas DA, Farrell PJ, Pichlmair A, Bowie AG. Poxvirus protein MC132 from molluscum contagiosum virus inhibits NF-κB activation by targeting p65 for degradation. *J Virol* 2015;89:8406–8415.
37. Nichols DB, Shisler JL. Poxvirus MC160 protein utilizes multiple mechanisms to inhibit NF-kappaB activation mediated via components of the tumor necrosis factor receptor 1 signal transduction pathway. *J Virol* 2009;83:3162–3174.
38. Damon I, Murphy PM, Moss B. Broad spectrum chemokine antagonistic activity of a human poxvirus chemokine homolog. *Proc Natl Acad Sci USA* 1998;95:6403–6407.
39. Xiang Y, Moss B. Determination of the functional epitopes of human interleukin-18-binding protein by site-directed mutagenesis. *J Biol Chem* 2001;276:17380–17386.
40. Xiang Y, Moss B. Molluscum contagiosum virus interleukin-18 (IL-18) binding protein is secreted as a full-length form that binds cell surface glycosaminoglycans through the C-terminal tail and a furin-cleaved form with only the IL-18 binding domain. *J Virol* 2003;77:2623–2630.
41. Slentz-Kesler KA, Hale LP, Kaufman RE. Identification and characterization of K12 (SECTM1), a novel human gene that encodes a Golgi-associated protein with transmembrane and secreted isoforms. *Genomics* 1998;47:327–340.
42. Bianchetti L, Tarabay Y, Lecompte O, Stote R, Poch O *et al*. Tex19 and Sectm1 concordant molecular phylogenies support co-evolution of both eutherian-specific genes. *BMC Evol Biol* 2015;15:222.
43. Lobito AA, Ramani SR, Tom I, Bazan JF, Luis E *et al*. Murine insulin growth factor-like (IGFL) and human IGFL1 proteins are induced in inflammatory skin conditions and bind to a novel tumor necrosis factor receptor family member, IGFLR1. *J Biol Chem* 2011;286:18969–18981.
44. Esparza J. Has horsepox become extinct? *Vet Rec* 2013;173:272–273.
45. Medaglia MLG, Moussatché N, Nitsche A, Dabrowski PW, Li Y *et al*. Genomic analysis, phenotype, and virulence of the historical Brazilian smallpox vaccine strain IOC: implications for the origins and evolutionary relationships of vaccinia virus. *J Virol* 2015;89:11909–11925.
46. Qin L, Favis N, Famulski J, Evans DH. Evolution of and evolutionary relationships between extant vaccinia virus strains. *J Virol* 2015;89:1809–1824.
47. Schrick L, Tausch SH, Dabrowski PW, Damaso CR, Esparza J *et al*. An early American smallpox vaccine based on Horsepox. *N Engl J Med Overseas Ed* 2017;377:1491–1492.
48. Zorec T, Kutnjak D, Hošnjak L, Kušar B, Trčko K *et al*. New insights into the evolutionary and genomic landscape of molluscum contagiosum virus (MCV) based on nine MCV1 and six MCV2 complete genome sequences. *Viruses* 2018;10:586.
49. Senkevich TG, Bugert JJ, Sisler JR, Koonin EV, Darai G *et al*. Genome sequence of a human tumorigenic poxvirus: prediction of specific host response-evasion genes. *Science* 1996;273:813–816.
50. Moens Y, Kombe AH. Molluscum contagiosum in a horse. *Equine Vet J* 1988;20:143–145.
51. Van Rensburg IB, Collett MG, Ronen N, Gerdes T. Molluscum contagiosum in a horse. *J S Afr Vet Assoc* 1991;62:72–74.
52. Bagnall BG, Wilson GR. Molluscum contagiosum in a red kangaroo. *Australas J Dermatol* 1974;15:115–120.
53. McKenzie RA, Fay FR, Prior HC. Poxvirus infection of the skin of an eastern grey kangaroo. *Aust Vet J* 1979;55:188–190.
54. Douglas JD, Tanner KN, Prine JR, Van Riper DC, Derwelis SK. Molluscum contagiosum in chimpanzees. *J Am Vet Med Assoc* 1967;151:901–904.
55. Wilson TM, Poglayen-Neuwall I. Pox in South American sea lions (Otaria byronia). *Can J Comp Med* 1971;35:174–177.
56. Baker KS, Leggett RM, Bexfield NH, Alston M, Daly G *et al*. Metagenomic study of the viruses of African straw-coloured fruit bats: detection of a chiropteran poxvirus and isolation of a novel adenovirus. *Virology* 2013;441:95–106.
57. López-Bueno A, Parras-Moltó M, López-Barrantes O, Belda S, Alejo A. Recombination events and variability among full-length genomes of co-circulating molluscum contagiosum virus subtypes 1 and 2. *J Gen Virol* 2017;98:1073–1079.
58. Huang B, Jennison A, Whiley D, McMahon J, Hewitson G *et al*. Illumina sequencing of clinical samples for virus detection in a public health laboratory. *Sci Rep* 2019;9:1–8.
59. Esteban DJ, Nuara AA, Buller RML. Interleukin-18 and glycosaminoglycan binding by a protein encoded by variola virus. *J Gen Virol* 2004;85:1291–1299.
60. Nazarian SH, Rahman MM, Werden SJ, Villeneuve D, Meng X *et al*. Yaba monkey tumor virus encodes a functional inhibitor of interleukin-18. *J Virol* 2008;82:522–528.
61. Smith VP, Bryant NA, Alcamí A. Ectromelia, vaccinia and cowpox viruses encode secreted interleukin-18-binding proteins. *J Gen Virol* 2000;81:1223–1230.
62. Dinarello CA. Il-18: a TH1-inducing, proinflammatory cytokine and new member of the IL-1 family. *J Allergy Clin Immunol* 1999;103:11–24.
63. Xiang Y, Moss B. Correspondence of the functional epitopes of poxvirus and human interleukin-18-binding proteins. *J Virol* 2001;75:9947–9954.
64. Lyman SD, Escobar S, Rousseau AM, Armstrong A, Fanslow WC. Identification of CD7 as a cognate of the human K12 (SECTM1) protein. *J Biol Chem* 2000;275:3431–3437.
65. Wang T, Huang C, Lopez-Coral A, Slentz-Kesler KA, Xiao M *et al*. K12/SECTM1, an interferon-γ regulated molecule, synergizes with CD28 to costimulate human T cell proliferation. *J Leukoc Biol* 2012;91:449–459.
66. Huyton T, Göttmann W, Bade-Döding C, Paine A, Blasczyk R. The T/NK cell co-stimulatory molecule SECTM1 is an IFN "early response gene" that is negatively regulated by LPS in Human monocytic cells. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2011;1810:1294–1301.
67. Lam GK, Liao H-X, Xue Y, Alam SM, Scearce RM *et al*. Expression of the CD7 ligand K-12 in human thymic epithelial cells: regulation by IFN-gamma. *J Clin Immunol* 2005;25:41–49.
68. Howie D, Garcia Rueda H, Brown MH, Waldmann H. Correction: secreted and transmembrane 1A is a novel co-stimulatory ligand. *PLoS One* 2013;8.

69. Smith VP, Alcami A. Inhibition of interferons by ectromelia virus. *J Virol* 2002;76:1124–1134.
70. Pontejo SM, Alejo A, Alcami A. Poxvirus-Encoded TNF decoy receptors inhibit the biological activity of transmembrane TNF. *J Gen Virol* 2015;96:3118–3123.
71. Alcamí A, Smith GL. A soluble receptor for interleukin-1 beta encoded by vaccinia virus: a novel mechanism of virus modulation of the host response to infection. *Cell* 1992;71:153–167.
72. Canuti M, van der Hoek L. Virus discovery: are we scientists or genome collectors? *Trends Microbiol* 2014;22:229–231.
73. Teifke JP, Kidney BA, Löhr CV, Yager JA. Detection of papillomavirus-DNA in mesenchymal tumour cells and not in the hyperplastic epithelium of feline sarcoids. *Vet Dermatol* 2003;14:47–56.
74. Reynolds ES. The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J Cell Biol* 1963;17:208–212.
75. Andrews S, Krueger F, Segonds-Pichon A, Biggins L, Krueger C. 2012. FastQC as found at. http://www.bioinformatics.babraham.ac.uk/projects/fastqc
76. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 2018;34:2666–2669.
77. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
78. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J *et al*. BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421.
79. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002;30:3059–3066.
80. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S *et al*. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008;36:W465–W469.
81. Tcherepanov V, Ehlers A, Upton C. Genome annotation transfer utility (GATU): rapid annotation of viral genomes using a closely related reference genome. *BMC Genomics* 2006;7:150.
82. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ *et al*. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 2017;45:D200–D203.
83. Söding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 2005;33:W244–W248.
84. Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J *et al*. A completely Reimplemented Mpi bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* 2018;430:2237–2243.
85. de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS *et al*. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res* 2006;34:W362–W365.
86. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P *et al*. Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res* 2018.
87. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 2017;45:D158–D169.
88. Madeira F, Park YM, Lee J, Buso N, Gur T *et al*. The EMBL-EBI search and sequence analysis tools Apis in 2019. *Nucleic Acids Res* 2019;47:W636–W641.
89. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;305:567–580.
90. Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* 2004;338:1027–1036.
91. Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A. Prediction of transmembrane alpha-helices in procariotic membrane proteins: the dense alignment surface method; Prot. *Eng* 1997;10:673–676.