Use of Next Generation Sequencing to study two cowpox virus outbreaks

Markus H. Antwerpen[1], Enrico Georgi[1], Alexandra Nikolic[1], Gudrun Zoeller[1], Peter Wohlsein[2], Wolfgang Baumgärtner[3], Christophe Peyrefitte[4], Remi Charrel[4] and Hermann Meyer[1]

[1] Bundeswehr Institute of Microbiology, Munich, Germany
[2] Small Animal Clinic, University of Veterinary Medicine, Hanover, Germany
[3] Department of Pathology, University of Veterinary Medicine, Hanover, Germany
[4] Unité des Virus Emergents, Aix Marseille Université, Marseille, France

ABSTRACT

Background. Between 2008 and 2011 about 40 cases of human cowpox were reported from Germany and France. Infections had been acquired via close contact to infected, young pet rats. An identical and unique sequence of the hemagglutinin gene was found in various cowpox virus (CPXV) isolates pointing to a common source of infection. In a second CPXV outbreak in cats in a small animal clinic in Germany in 2015, four out of five hospitalized cats showed identical hemagglutinin sequences and thus, a hospital-acquired transmission had been assumed. Next-Generation Sequencing was performed in order to re-investigate the outbreaks, as epidemiological data could not confirm all cases.

Methods. Homogenates of lesion material from rats, cats and humans were cultivated in cell culture. The genomes of four virus isolates, nine CPXVs from our strain collections and from DNA of three paraffin-embedded lesion materials were determined by Next Generation Sequencing (NGS). For phylogenetic analyses a MAFFT-alignment was generated. A distance matrix based on concatenated SNPs was calculated and plotted as dendrogram using Unweighted Pair Group Method with Arithmetic mean (UPGMA) for visualization.

Results. Aligning of about 200.000 nucleotides of 8 virus isolates associated with the pet rat outbreak revealed complete identity of six genomes, the remainder two genomes differed in as little as 3 SNPs. When comparing this dataset with four already published CPXV genomes also associated with the pet rat outbreak, again a maximum difference of 3 SNPs was found. The outbreak which lasted from 2008 till 2011 was indeed caused by a single strain which has maintained an extremely high level of clonality over 4 years. Aligning genomic sequences from four cases of feline cowpox revealed 3 identical sequences and one sequence which differed in 65 nucleotides. Although identical hemagglutinin sequences had been obtained from four hospitalized cats, genomic sequencing proved that a hospital-acquired transmission had occurred in only three cats.

Conclusion. Analyzing the rather short sequence of the hemagglutinin gene is not sufficient to conduct molecular trace back analyses. Instead, whole genome sequencing is the method of choice which can even be applied to paraffin-embedded specimens.

Corresponding author
Markus H. Antwerpen,
markusantwerpen@bundeswehr.org

INTRODUCTION

The genus *Orthopoxvirus*, family *Poxviridae*, subfamily *Chordopoxvirinae*, presently contains four species which are pathogenic to humans: variola virus (VARV), the causative agent of smallpox, monkeypox virus (MPXV), vaccinia virus (VACV), and cowpox virus (CPXV). Mousepox (ECTV) and camelpox virus (CMLV) can affect mice and camels, respectively. Including tatera poxvirus (TATV), these species are collectively referred to as "Old World" orthopoxviruses (OPVs). In addition, three species of North American origin, i.e., raccoon, vole and skunk poxvirus, are currently known (*ICTV, 2018*).

While naturally occurring VARV infections are indeed no longer of concern—except in a bioterrorist context—human infections with MPXV and CPXV are increasingly reported. Whereas human monkeypox are commonly epidemiological attributed to the tropical rainforest of Africa, CPXV is endemic in Western Eurasia. Although initially seen in cows, it is now generally accepted that the species CPXV is misnamed (*Reynolds et al., 2018*). Rodents, particularly vole species, are regarded to be the true reservoir hosts from which spillover infections to accidental hosts are regularly observed. Accidental hosts are cows, domestic cats, horses and also exotic animals in zoos, such as large felidae and elephants, which regularly develop severe disease. Humans in direct contact with infected accidental hosts, especially cats and pet rats, can develop local lesions, in rare cases a generalization is possible (*Fassbender et al., 2016*).

Historically, OPVs have been differentiated from other poxviruses by demonstration of a hemagglutinin (HA)-positive phenotype able to agglutinate chicken erythrocytes. Within the genus *Orthopoxvirus* single species had been differentiated based on biological characteristics, including mouse virulence, size and color of pocks on the chorioallantoic membrane as well as formation of A-type inclusion bodies. A pathognomonic feature of isolates assigned to the species CPXV is the induction of hemorrhagic pocks on the chorioallantoic membrane and the formation of intracytoplasmatic A-type inclusion bodies (*Fenner, 1990*). However, conducting these assays is laborious and time consuming. Therefore, attempts had been made to correlate the phenotype with the genotype of OPV species either by generation of Restriction Fragment Length Polymorphism data or by sequencing single genes, such as the hemagglutinin (HA) gene (*Ropp et al., 1995*). Both genome-based methods successfully confirmed the current species concept originally based on phenotypic characteristics. Sequencing of the HA gene had become the method of choice to assign a given OPV isolate to a species of the genus *Orthopoxvirus* and numerous HA sequences from various cases are available in GenBank (*Ninove et al., 2009*; *Hemmer et al., 2010*; *Cardeti et al., 2011*; *Dabrowski et al., 2013*; *Ducournau et al., 2013*; *Fassbender et al., 2016*; *Mauldin et al., 2017*; *Franke et al., 2017*). As compared to VARV and MPXV, the HA sequences of CPXVs are much more diverse. This genetic variation can be explained by the existence of local variants which had evolved over time in a rodent reservoir (*Reynolds et al., 2018*). Recently, results of phylogenetic analyses of 58 full-length CPXV genomes confirmed the polyphyletic character of the species CPXV which consists of at least four different clades and several variants (*Mauldin et al., 2017*; *Franke et al., 2017*). The classification of CPXV strains into clades roughly followed their geographic origin.

A rather large outbreak occurred in Germany and France during 2008 and 2011 with about 40 cases of human cowpox virus infection (*Campe et al., 2009*; *Ninove et al., 2009*; *Ducournau et al., 2013*). All humans had acquired their infection via young pet rats. Most of the pet rats had died rather quickly due to pulmonary infection caused by CPXV. Characterization of 14 viruses/clinical materials by sequencing the respective HA genes proved a 100% identity (*Campe et al., 2009*; *Ninove et al., 2009*; *Dabrowski et al., 2013*; *Ducournau et al., 2013*) indicating that a single CPXV strain circulating in rat breeder facilities could had been the common source. Recently an atypical manifestation of CPXV infection in 5 cats in the larger area of Hannover was described by *Jungwirth et al. (2018)*. Identical HA gene sequences were identified in four cats, and in combination with epidemiological data a hospital-acquired transmission had been assumed (*Jungwirth et al., 2018*).

With the introduction of benchtop sequencers, enabling even small routine diagnostic laboratories to perform whole genome sequencing, sequencing of only a part of the microorganisms' genome is not a riddle anymore. Determination of drug resistances or virulence genes can be performed easily based on the same sequence dataset, as well as identification and *in-silico* typing and clonal assignment of the investigated strains.

In this study Next Generation Sequencing (NGS) was used to re-analyze the pet rat and the feline cowpox outbreak. We determined the genomes of further strains and lesion materials and conclude that the pet rat-associated outbreak which lasted from 2008 till 2011 had been indeed caused by a unique CPXV strain. In addition, we determined genomic sequences of the 2015 Hannover outbreak in cats and conclude that hospital acquired cowpox infections had occurred in three out of four hospitalized cats.

METHODS

Viruses

Four clinical specimens suspected to contain CPXV were received in Germany in 2009. Two specimens were collected from a 13-year old boy and his pet rat; the boy presented with local skin lesions and his rat died with severe pulmonary symptoms. One specimen was lesion material from a boa which had been bitten by a feeder rat, which had been bought in a pet store in Marl, Germany. The fourth specimen originated from a rat, which had been bought in the same pet store in Marl and had died showing severe respiratory symptoms. Homogenates of the lesions were prepared in cell culture medium using a bead-beater and aliquots were inoculated onto African Green Monkey Kidney cell line MA 104 as described recently (*Fassbender et al., 2016*). Four viral strains, each corresponding to a specific specimen, have been isolated and designated as strains "Boy Biederstein", "Rat Biederstein", "Rat Marl" and "Boa Marl", respectively (Table 1). In the feline CPXV outbreak in Hanover 2015 (listed in Table 2), a total of five CPXV strains could be isolated ("75/01", "Moritz 2015/3a", "Sammy 2015/4a", "Leo 2015/5" and "S2216/04") (*Jungwirth et al., 2018*). Homogenatization of lesions and inoculation of cell-cultures were performed as described above. CPXV identification had been achieved through sequencing the PCR product corresponding to the entire ORF of the HA gene (*Campe et al., 2009*).

Table 1 List of cases associated with the pet rat outbreak 2008 till 2011 (HA-sequence group C030; Table S1) (See last pages).

Case	Strain designation	Host	Country	Town	Year	Reference	Accession #	Full-length genome determined
#1 GER 2008	CPXV HumKre08/1	Human	Germany	Krefeld	2008	Dabrowski	KC813512	Dabrowski et al.
	CPXV RatKre08/2	Rat					KC813505	
#2 GER 2009	CPXV Rat Koelle/Shiva	Rat	Germany	Muenchen	2009	Campe	FJ654467	this study
	CPXV Rat breeder	Rat		Hengersberg			FJ654467	n.a.
	CPXV RatGer09/1	Rat		Germering		Dabrowski	KC813503	Dabrowski et al.
	CPXV Rat Biederstein	Rat		Muenchen		unpublished	n.a.	this study
	CPXV Boy Biederstein	Human						this study
#3 GER 2009	CPXV HumAac09/1	Human	Germany	Aachen	2009	Dabrowski	KC813508	Dabrowski
	CPXV RatAac09/1	Rat					KC813501	
#4 GER 2009	CPXV Rat Marl	Rat	Germany	Marl	2009	unpublished	LN864565	Hoffmann et al.
	CPXV Boa	Boa					n.a.	this study
#5 FRA 2009	CPXV Compiegne 2009	Human	France	Compiegne	2009	Ninove	FJ754355	this study
	CPXV DG	Human					FJ754357	n.a.
	CPXV LP	Human					FJ754356	n.a.
#6 FRA 2009	CPXV R09	Rat	France		2009	unpublished	FJ790310	n.a.
#7 FRA 2011	CPXV Cepad 332	Human	France	Epinal	2011	Ducournau	KC592396	this study
	CPXV Cepad 333	Human					KC592397	this study
	CPXV Cepad 336	Human					KC592399	n.a.

Table 2 List of known strains of the hospital-associated CPXV outbreak in Germany in 2015 with their accession numbers including reference isolates showing the same HA-sequence (HA-sequence group C029; Table S1) and used for further analyses in this study.

Case	Strain designation	Host	Country	Town	Year	Reference	Accession#	Full-length genome determined
feline cowpox	Moritz 2015/3a	Cat	Germany	Aerzen	2015	Jungwirth	MF948141	this study
	Sammy 2015/4a	Cat	Germany	Bad Münder	2015	Jungwirth	MF948143	this study
	Sammy 33-37 (paraffin)					Jungwirth		this study
	Leo 2015/5	Cat	Germany	Extertal	2015	Jungwirth	MF948139	this study
	Leo 26-29 (paraffin)					Jungwirth		this study
	Cookie S963/15 (paraffin)	Cat	Germany	Benthe	2015	Jungwirth	MF948138	this study
	CPXV HumPad07/1	Human	Germany	Paderborn	2007	Dabrowski	KC813496	Dabrowski
	CPXV S2216/04	Cat	Germany	Hannover	2005	Jungwirth	MF948142	this study
	CPXV 75/01	Human	Germany	Detmold	2001	unpublished	AY902293	this study
	CPXV K779	Cat	England	Bristol	2000	Mauldin	KY549146	Mauldin
	CPXV-922-99	Cat	Germany		1999	unpublished	AF377885	n.a.
	CPXV-EP-4 lidil	Elephant	Germany	Hameln	1980	Carroll	HQ420895	Carroll

Whole genome sequencing

Genomic DNA was prepared from virus-infected cell cultures. A total of eight CPXV strains associated with the pet rat outbreak were analyzed. Four strains were part of our strain collections ("Compiegne 2009", "Rat Koelle", "Cepad 332", and "Cepad 333") and four

strains had been isolated (strains "Rat Marl", "Boa Marl", "Rat Biederstein" and "Boy Biederstein"; Table 1). With regard to the feline cowpox outbreak (*Jungwirth et al., 2018*) a total of five CPXV strains ("75/01", "Moritz 2015/3a", "Sammy 2015/4a", "Leo 2015/5" and "S2216/04") of our collections were investigated (Table 2).

Infected cell cultures were freeze-thawed three times and genomic DNA was extracted by using the automated DNA extraction system MagNA Pure Compact (Roche Diagnostics, Penzberg, Germany). Genomic DNA was also extracted directly from paraffin-embedded skin tissues of cats Sammy, Leo and Cookie using the FFPE-DNAeasy kit (Qiagen, Hilden, Germany).

DNA quantification was performed using the fluorometric system Qubit 2.0 (Life Technologies, Darmstadt, Germany) and the Qubit® dsDNA high sensitivity assay kit (Thermo Fisher Scientific, Waltham, MA, USA). Nextera® XT DNA Library Preparation kit (Illumina) with an input DNA amount of 3–5 ng was used for library preparation. Finally, whole genome sequencing was performed on a MiSeq instrument (Illumina) with corresponding MiSeq Reagent Kit v3 (600 cycle) chemistry.

Bioinformatics

Sequences encoding the entire open reading frame (ORF) of the hemagglutinin gene of "Old World" OPVs were downloaded from NCBI nucleotide archive. The respective strains ($n = 462$) and their accession numbers are listed (Table S1). For sequence clustering a Unweighted Pair Group Method with Arithmetic mean (UPGMA)-analysis was performed using software Bionumerics and visualized as Minimum-Spanning-tree (MST).

Sequence analysis

The sequenced paired-end reads were mapped against chromosomes of *Catus felis* (RefSeq NC_018723–NC_018741) or *Chlorocebus sabaeus* (RefSeq NC_023642–NC_023672) in order to discard sequence reads of eukaryotic background DNA. Remaining reads were assembled de-novo with SPAdes v3.9 (*Bankevich et al., 2012*). Hence, resulting contigs were aligned against CPXV reference strain Brighton Red (RefSeq NC_003663.2) using blastn (*Camacho et al., 2009*) to determine their order. Afterwards, an in-house developed Python script was used to concatenate sequences into one single scaffold by resolving overlaps or creating gaps filled with Ns of a length estimated from the mapping. BWA (*Li, 2013*) was subsequently used to map the raw reads to the chromosomal contigs. Finally, Pilon (*Walker et al., 2014*) was applied to undo mis-assemblies, polish the sequence by correcting SNPs and small InDels, as well as to close gaps in order to obtain one continuous chromosome based on almost the complete whole genome sequence. The Genome Annotation Transfer Utility (GATU (*Tcherepanov, Ehlers & Upton, 2006*)) was used for annotation of genomes. The criteria for annotation included 90% nt similarity with CPXV strain Brighton Red. After the transfer of annotation, assigned open reading frames of all genomes were visually inspected and corrected manually, if needed. All generated genomes were deposited into NCBI GenBank as a BioProject PRJNA369073.

For phylogenetic analyses, a MAFFT-alignment (*Katoh & Standley, 2013*) of all sequenced and downloaded reference virus genomes (Tables 1 and 2) was generated

using the FFT-NS-2 algorithm and inspected and manually corrected using MEGA 6.0, if needed. Using an in-house python script, only those positions present in all datasets were used for further calculations. Furthermore, for determining of SNP-calling, only those SNPS were used, of which within a surrounding area of 10 nucleotides no further SNPs were observed in the used dataset, and no gap in the alignment for at least one sequence was observed. 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 1,000 replicates.

RESULTS

Hemagglutinin-based phylogenetic relationships of Orthopoxviruses

We reconstructed the phylogeny of Old World OPVs by using sequences covering the entire ORF of the HA gene since the number of HA sequences by far exceeded those of other genes. As of June 2018, a total of 462 sequences were retrieved from Genbank and aligned. Details about the species, the year and location of isolation, the host and the length of the ORFs are given in Table S1. A total of 152 unique sequences were identified and identical sequences were assigned common group numbers. Bayesian inference showed a consensus Minimum-Spanning-Tree (Fig. 1). Circle sizes are relative to the number of identical sequences and branch lengths are relative to the number of nucleotide changes. In general, all sequences of strains and isolates belonging to the species MPXV ($n = 112$ sequences), VARV ($n = 87$), CMLV ($n = 39$) and ECTV ($n = 11$) cluster closely together with little variation. The four species are clearly separated from each other, thus reflecting accurately the current taxonomy of the genus *Orthopoxvirus*.

The species VACV is monophyletic as well, however, variation is much higher with as many as 34 different variants out of 52 reported sequences. This can be best explained because VACV vaccine strains have been repeatedly passaged in various cell cultures in different laboratories. Since there was no selective pressure to survive in a reservoir host, *in vitro* replication has favored evolution into pools of highly diverse quasi-species (*Qin et al., 2015*). The genetic variation of VACVs described here mirrors results of recent full-genome analysis (*Damaso, 2018*). The authors conclude that VACV was once a circulating virus in animals in Europe in the 19th century that later disappeared in this reservoir.

Comparative analyses of CPXV sequences display a much higher genetic diversity: 144 strains can be allocated to 79 unique sequence groups. In contrast to the *orthopoxvirus* species mentioned above, phylogenetic analyses clearly demonstrate that CPXV cannot be regarded as a monophyletic species: there are several different clades as already reported by *Dabrowski et al. (2013)*, *Mauldin et al. (2017)* and *Franke et al. (2017)*. However, some strains might even be erroneously attributed to VACV while indeed belonging to CPXV. The genetic diversity of CPXVs is also reflected by different lengths of the HA ORF, ranging from 816 bp to 963 bp. A recent phylogenetic analysis of 83 full-length OPV genomes (*Franke et al., 2017*) confirmed the polyphyletic character of the species CPXV and the existence of at least four different clades had been assumed.
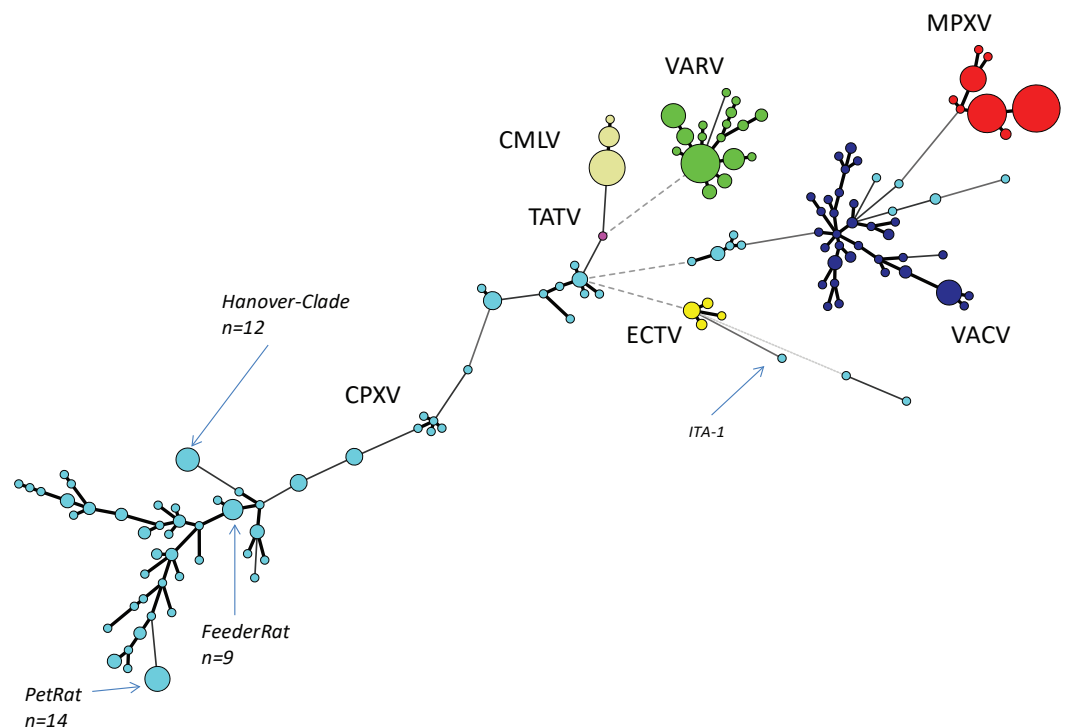
Figure 1 **Minimum-Spanning-Tree of 462 hemagglutinin gene sequences of "Old-World"-orthopoxviruses retrieved from NCBI (March 2018).** Each circle corresponds to a unique sequence with circle sizes relative to the number of identical sequences. Individual species of the genus Orthopoxvirus are labelled by different colors. CPXV, Cowpox virus, light-blue; MPXV, Monkeypox virus, red; VARV, Variola virus, green; VACV, Vaccinia virus, dark-blue; TATV, Tatera poxvirus, pink; ECTV, Ectromelia virus, yellow; CMLV, Camelpox virus, beige.

Next Generation Sequencing of cowpoxviruses

For each isolate or sample described in this study, one single sequence of at least 200.000 nucleotides with an average sequencing depth of 74 (minimum of 45-fold) could be determined. Each of the sequences spans the entire highly conserved core genome. The outermost terminal inverted repeats (*Massung, Knight & Esposito, 1995*) could not be reliable determined using a short-read-sequencing technology in combination with the described conservative approach, and were therefore discarded.

In order to determine the reliability of the sequencing method used here on the Illumina platform we re-sequenced two CPXV strains for comparison with their already published genome sequences in order to estimate inter-study reliability. CPXV EP4 Hannover (Table 1) from our in-house collection was compared to Germany_1980_EP4 (HQ420895) which originates from the same isolate, but was sequenced by a different institution using Illumina Sequencing technology as well (*Carroll et al., 2011*). The second CPXV strain was Ratpox09 (Table 1, Case #4), whose sequence (LN864565) had already been determined on the Illumina platform by *Hoffmann et al. (2015)* and in this study, too. Not a single nucleotide difference between the sequences could be observed.

In a second approach we wanted to determine whether the process of virus isolation in cell culture would result in nucleotide differences as compared to a sequence directly generated from DNA of lesion material. To this end we sequenced both, two virus isolates obtained from the cats Leo and Sammy (Table 1) and DNA directly extracted from paraffin-embedded lesion material of the same cats. Comparison of the genome of CPXV_2015_5 and DV400026 PF (Leo) as well as CPXV_2015_4a and CPXV_DV400033 PF (Sammy) revealed identical sequences in both cases.

In order to compare the genomes described in this study with already published CPXV genomes, we selected prototypic members of three genomic CPXV-clades described by *Franke et al. (2017)*. Strain HumGra07 (KC813510) belongs to the VARV-like CPXV clade marked in red, strain Finland2000Hum (HQ420893) belongs to the VACV-like CPXV clade marked in green, all other strains EP4, K779, K780, MKY2002, CatBer07, HumLan08 and RatHei09 (HQ420895, AY902281, KY549147, HQ420898, KC813502, KC813492 and KC813504, resp.) belong to the genetic diverse CPXV-1 lineage marked in blue (Fig. 2). Whereas HumLan08 and RatHei09 are epidemiologically linked to each other and show identical genome sequences, two genomes from isolates from the city Leipzig in East Germany (2015-1 and 2015-6) (*Franke et al., 2017*; *Eder et al., 2017*) originate from two epidemiological unlinked human cases. This was done in order to achieve an idea of the dimension of genetic distances of unrelated samples with the same geospatial and temporal, as well as host attribution. These latter two sequences differ in 111 nucleotide positions from each other.

All strains belonging to HA-sequence type C030 (Table S1) are clustered to the VARV-like strains and strains of HA-sequence type C029 (Table S1) show strong similarity to the EP4-strain (HQ420895) within the CPXV-1 lineage.

Pet rat-associated outbreak
Comparison of sequences generated in this study
In this study we successfully isolated 4 CPXVs from specimens of a boy, his pet rat as well as from a boa and a rat. By including additional four strains of our strain collections, which were attributed to this outbreak based on their HA-sequence a total of eight pet rat-associated CPXVs genomic sequences were determined. Following alignment a total of 197.964 nucleotide positions could be used for further investigations. Comparison of the eight sequences determined here revealed that six sequences were identical: the strains "Cepad 332" and "Cepad 333" (Case #7) and "Compiegne" (Case #5) had been isolated in France, strains "Boa Marl" and "Rat Marl" had been isolated in the same town in West-Germany (Case #4) and strain "Rat Koelle" originates from South Germany (Case #2). Compared to these six sequences, two sequences—obtained from a sick rat with a proven transmission to a boy (Case #3, South-Germany)—differed in as little as either two or three SNPs, respectively. In gene CPXV089 (according to the nomenclature of the CPXV reference strain Brighton) a nucleotide exchange was leading to an amino-acid change (A → G) and in gene CPXV139 a non-synonymous mutation (R → W) was observed. In addition, strain "Boy Biederstein" has one more nucleotide exchange downstream the gene CPXV035 in an intergenic region. As the young boy was indeed infected by his rat ("Rat
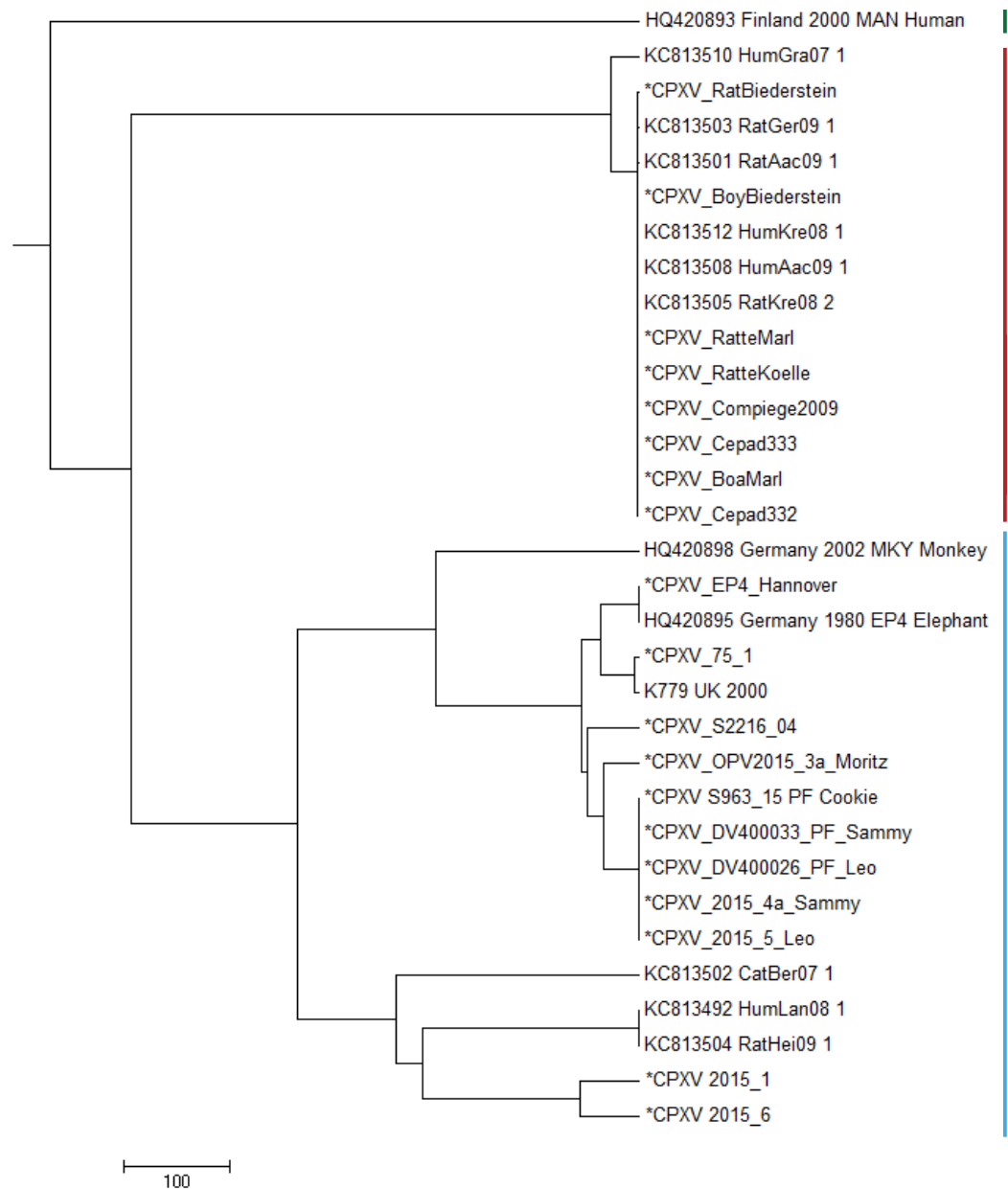
Figure 2 Maximum Parsimony-Dendrogram of extracted and concatenated SNP-positions. Branch lengths show nucleotide differences between the investigated strains. All nodes show bootstrap values of 100%. Sequences marked with an asterisk (*) had been determined in this study. Colored bars indicate three different CPXV-clades according to *Franke et al. (2017)*.

Biederstein"), this third and unique mutation of the isolate "Boy Biederstein" indicates a mutually new introduced point mutation during the period of infection.

Comparison with sequences derived from GenBank

In a next step, we compared our dataset with five sequences already deposited in GenBank from other pet rat-associated CPXVs (*Dabrowski et al., 2013*). A maximum difference
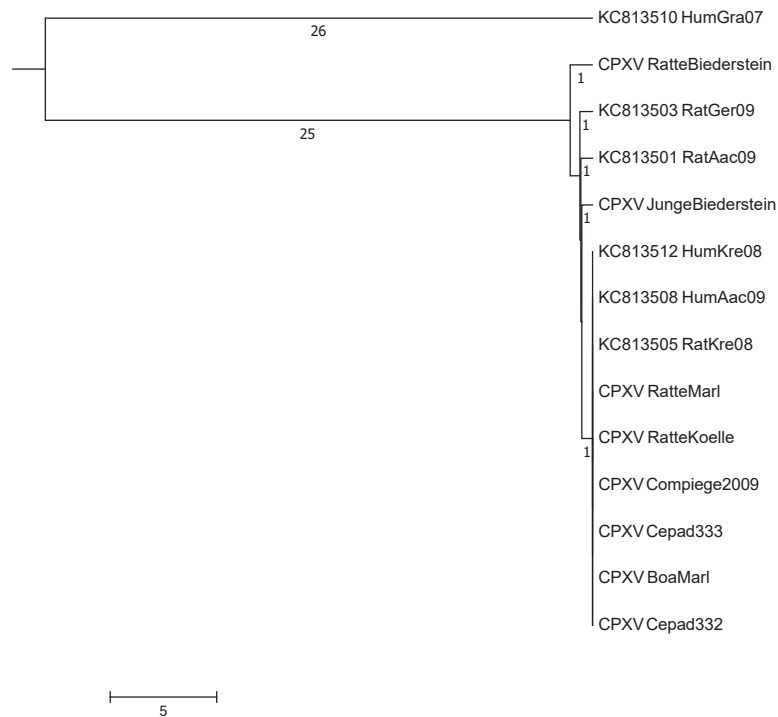
Figure 3 **Maximum-Parsimony Dendrogram of extracted and concatenated SNP-positions of pet rat-associated cowpox virus strains.** Branch lengths show nucleotide differences between the investigated strains. All nodes show bootstrap values of 100%.

of up to eight nucleotides was initially observed between the isolates "HumAac09" and "Rat Biederstein". These sequences had been generated (*Dabrowski et al., 2013*) by using the Roche 454 technology which is known to produce sequencing errors especially in homopolymer regions (*Quail et al., 2012*). Visual inspection and comparison of the alignment confirmed that most differences could be attributed to homopolymer regions. As no raw reads are publically available, a re-calculation with actually available error-correction algorithms could not be performed. By ignoring those positions, strains HumKre08 and RatKre08 from a rat and its owner (Case #1) were identical to the six sequences determined here as well as with strain HumAac09 (Case #3) from a different rat owner. The associated rat isolate "RatAac09" however, differed in two nucleotides as compared to the human isolate. The genetic difference of strain "RatGer09" (Case # 2) was determined to be of two nucleotides. By comparing this pet rat associated-dataset with sequences already published at NCBI, the closest next neighbor was determined to be strain HumGra07 (KC813510), which originates from a case of human cowpox in Graz, Austria, 2007, and differed in a total of 51 SNPs (Fig. 3).

Feline cowpox outbreak
Comparison of sequences generated in this study
Plaque-like skin lesions had been observed in five cats which had been hospitalized in a period of 5 weeks in a veterinary clinic. CPXV infections were diagnosed and based on
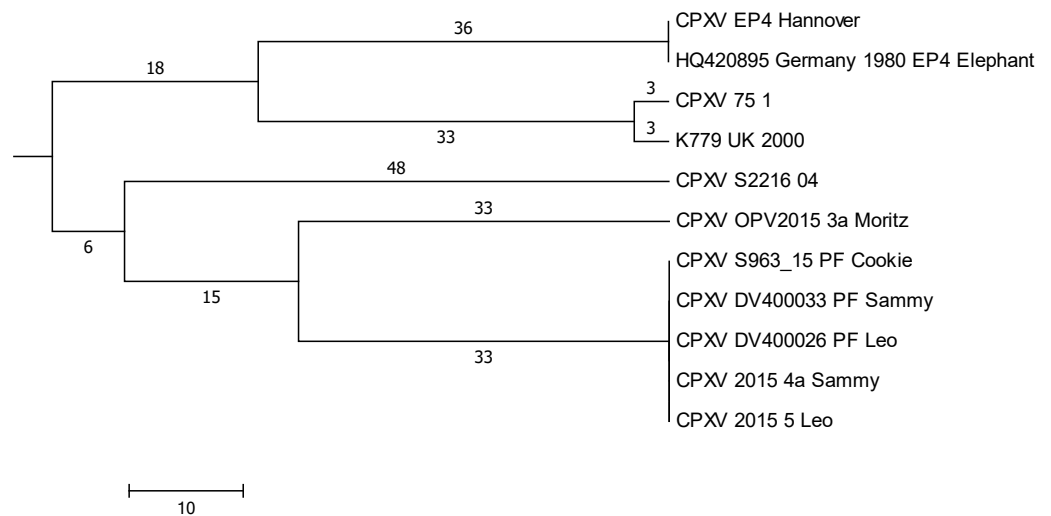
Figure 4 Maximum-Parsimony Dendrogram of extracted and concatenated SNP-positions of feline cowpox-associated strains. Branch lengths show nucleotide differences between the investigated strains. All nodes show bootstrap values of 100%.

identical HA-sequences in four out of five cats (Leo, Sammy, Cookie, and Moritz), it was assumed that infections could have been acquired in the clinic due to contact (*Jungwirth et al., 2018*). DNA extracted from three virus isolates and from three paraffin-embedded tissues (labelled "PF" in Fig. 4) was used to generate whole genomic sequences. In depth-analysis of the assembly of genomes derived from the paraffin-embedded samples showed a higher initial miscalling rate of nucleotides within reads, similar to observations made before (*Arreaza et al., 2016*). By using a sequence depth >120-fold, the consensus bases were all called correctly as compared to the sequence derived from isolates. Comparison of all genomes demonstrated that five genomes were identical. The sixth sequence derived from cat Moritz (CPXV_OPV2015_3a) differed in 65 positions (Fig. 4) with SNPs located throughout the whole genome.

Comparison with sequences derived from GenBank

In order to find a possible relationship to previously reported cases of cowpox, sequences generated here were compared to genomes with identical HA sequences. A total of 201.073 nucleotide positions could be used for evaluation. Strain "CPXV S2216/04" which was isolated from a cat in 2004 in the same region, differed in 96 SNPs and strains CPXV "K779 UK2000" and "CPXV 75/1" differed in 108 nucleotides.

DISCUSSION

Conducting PCR screening assays targeting the HA gene in combination with sequencing of the HA ORF allow (i) unambiguous identification of orthopoxviruses and (ii) assigning a given isolate to an OPV species. Sequencing can be performed even in laboratories, which do not have access to a routine NGS-pipeline but rely on Sanger sequencing. Here, we aligned 447 HA sequences of orthopoxviruses and could show that the HA genetic signatures of

VARV, MPXV, CMLV and ECTV are indeed species-specific. This can be used as a useful tool for rapid typing of new isolates and assigning them to a given OPV species without conducting time-consuming and laborious phenotypic characterizations. Typing of and differentiating the species CPXV and VACV is also possible with the restriction that some CPXV strains might be mistyped as VACVs (Fig. 1 and Table S1, CPXV groups C001-3, C053 and C063). According to whole genome analysis, these strains form a separate clade referred to as "VACV-like" CPXV (Franke et al., 2017) which has a common ancestor with the species VACV. According to the species definition of the International Committee on Taxonomy of Viruses, the taxonomic classification of CPXVs has to be revisited to avoid confusion and to recognize five monophyletic clades of cowpox viruses as lowest taxa, which is supported by current data (Mauldin et al., 2017).

The discriminatory power of the HA gene sequence is excellent despite it is a short sequence of about 940 nucleotides: HA analysis allows to unambiguously identify VARV, MPXV, ECTV, CMLV, VACV and CPXV. In case of VARV and MPXV further differentiation is possible: variola major strains can be differentiated from variola minor strains and Western African MPXVs can be differentiated from Central African MPXVs.

In contrast, the heterogeneity of CPXVs is much higher with 79 unique sequences out of 144 strains (Fig. 1 and Fig. S1). The existence of unique isolates can be best explained by evolution over time in a local rodent reservoir and strains therefore cluster according to their geographic origin. This is a prerequisite for conducting phylogeographic attributions and trace back analyses to understand cowpox outbreaks. Researchers have taken advantage of the sequence diversity to investigate cowpox outbreaks and conducted trace-back analysis. For example, in a rather small outbreak involving a rat, an elephant and the animal caretaker three identical HA sequences were obtained. Since this sequence had not been reported before, a rat-elephant-human transmission was assumed (Hemmer et al., 2010). In another outbreak with generalized infections in banded mongooses (*Mungos mungo*), jaguarundis (*Herpailurus yagouaroundi*) and a time-delayed case of human cowpox, an identical and unique HA sequence indicates a common source of infection, i.e., infected food rats (Kurth et al., 2009). If HA sequences are distinct, then it can be concluded that they are not epidemiologically linked, if HA sequences are identical, they could or could not be epidemiologically linked. This scenario merits to be further investigated using a supplementary method (Kurth et al., 2009; Hemmer et al., 2010). To this end we have investigated two HA sequence groups (C029 and C030; Table S1) which are composed of 11 and 14 identical sequences, respectively, by determining the sequence of the entire genome which has the highest discriminatory power.

Sequence group C030 consists of various specimens derived from seven pet rat-associated cases (Table 1). We could show here, that all 14 genomes had maintained an extremely high level of clonality. The same high level of clonality (maximum of 3 SNPs) was observed in 4 genomes obtained from a smallpox outbreak, which spread form in Iran in 1968 to Syria and Yugoslavia in 1972. A highly efficient proofreading mechanism of the viral DNA-polymerase and/or a low selective pressure in the respective hosts might be the reason. The origin of the pet rat outbreak could not be identified with certainty: close to the German-Czech border a die-off, especially of young rats was observed in 2009 at a rat

breeder facility and the outbreak in Munich (case #2) could be traced back to this facility (*Campe et al., 2009*). The authors (*Ducournau et al., 2013*) claim that the young rats in France (case #7) in 2011 had been imported from the Czech Republic. Further attempts involving Czech authorities to unveil the origin of feeder or pet rats and their trading routes failed.

Sequence group C029 consists of a total of 11 identical HA sequences (Table 2), four sequences from an outbreak of feline cowpox and seven sequences from single cases of cowpox which had occurred in the larger area of Hannover and in the UK in previous years (*Schaudien et al., 2007*; *Carroll et al., 2011*; *Dabrowski et al., 2013*). The latter cases are not linked to the feline cowpox outbreak. Genome sequencing clearly proved that a hospital-acquired transmission had occurred in three out of four cats. In combination with epidemiological data (*Jungwirth et al., 2018*), cat Leo could be identified as the index case with few nodular skin lesions at initial presentation. At this time an infection with CPXV was not considered and Leo was hospitalized. During this time Sammy with a history of a car accident and Cookie, which was treated for renal insufficiency, were hospitalized as well. A fourth cat, Moritz (with an identical HA sequence) was also hospitalized. All cats could have been potentially exposed. However, the genome of Moritz differed in 61 nucleotides compared to identical genomes of Leo, Sammy and Cookie. We conclude that Moritz has not become infected in the clinic.

CONCLUSION

Here, we demonstrated that traceback analyses based on HA sequences only are not sufficient to detect the true chains of transmission in an outbreak. Whole genome sequencing is the method of choice since it provides the highest discriminatory power and in combination with epidemiological data outbreak investigations can be performed.

For orthopoxviruses the specific viral load within lesions is fortunately very high, so that even with smaller sequencing instruments, sufficient sequencing depth can be achieved using clinical samples. Even sequencing the entire genome of FFPE treated lesion material can be performed.

With use of Next-Generation-Sequencing, knowledge of genetic epidemiology of orthopoxviruses has increased during the last two to three years (*Hoffmann et al., 2015*; *Mauldin et al., 2017*; *Franke et al., 2017*). Creating and updating reliable databases of orthopoxvirus genomes would even allow a better understanding of genetic diversity and clonality. In combination with new sequencing technologies, this knowledge might help to be prepared for future orthopoxvirus outbreaks of MPXV or CPXV. This has already been successfully applied during the recent Ebola outbreak in West Africa: Sequences were generated on site (*Arias et al., 2016*; *Quick et al., 2016*), linked to clinical cases and enabled outbreak reconnaissance and containment of the outbreak.

Funding

This publication was supported by the European Virus Archive goes Global (EVAg) project that has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 653316. This study was also supported in part by the European Union's Horizon 2020 research and innovation program under grant agreement No 643476 (COMPARE). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
European Union's Horizon 2020 research and innovation program: 653316, 643476.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Markus H. Antwerpen and Hermann Meyer conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.
- Enrico Georgi analyzed the data, authored or reviewed drafts of the paper, approved the final draft.
- Alexandra Nikolic performed the experiments, analyzed the data, approved the final draft.
- Gudrun Zoeller conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, approved the final draft.
- Peter Wohlsein, Wolfgang Baumgärtner, Christophe Peyrefitte and Remi Charrel performed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

   All generated genomes are available in NCBI GenBank as BioProject PRJNA369073.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.6561#supplemental-information.

REFERENCES

Arias A, Watson SJ, Asogun D, Tobin EA, Lu J, Phan MVT, Jah U, Wadoum REG, Meredith L, Thorne L, Caddy S, Tarawalie A, Langat P, Dudas G, Faria NR,

Dellicour S, Kamara A, Kargbo B, Kamara BO, Gevao S, Cooper D, Newport M, Horby P, Dunning J, Sahr F, Brooks T, Simpson AJH, Groppelli E, Liu G, Mulakken N, Rhodes K, Akpablie J, Yoti Z, Lamunu M, Vitto E, Otim P, Owilli C, Boateng I, Okoror L, Omomoh E, Oyakhilome J, Omiunu R, Yemisis I, Adomeh D, Ehikhiametalor S, Akhilomen P, Aire C, Kurth A, Cook N, Baumann J, Gabriel M, Wölfel R, Caro AD, Carroll MW, Günther S, Redd J, Naidoo D, Pybus OG, Rambaut A, Kellam P, Goodfellow I, Cotten M. 2016. Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases. *Virus Evolution* **2(1)**:Article vev016 DOI 10.1093/ve/vew016.

Arreaza G, Qiu P, Pang L, Albright A, Hong LZ, Marton MJ, Levitan D. 2016. Pre-analytical considerations for successful Next-Generation Sequencing (NGS): challenges and opportunities for Formalin-Fixed and Paraffin-Embedded Tumor Tissue (FFPE) samples. *International Journal of Molecular Sciences* **17(9)**:Article 1579 DOI 10.3390/ijms17091579.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov SA, Lesin MV, Nikolenko IS, Pham S, Prjibelski DA, Pyshkin VA, Sirotkin VA, Vyahhi N, Tesler G, Alekseyev AM, Pevzner AP. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* **19**:455–477 DOI 10.1089/cmb.2012.0021.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden LT. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**:421 DOI 10.1186/1471-2105-10-421.

Campe H, Zimmermann P, Glos K, Bayer M, Bergemann H, Dreweck C, Graf P, Weber KB, Meyer H, Büttner M, Busch U, Sing A. 2009. Cowpox virus transmission from pet rats to humans, Germany. *Emerging Infectious Diseases* **15**:777–780 DOI 10.3201/eid1505.090159.

Cardeti G, Brozzi A, Eleni C, Polici N, D'Alterio G, Carletti F, Scicluna MT, Castilletti C, Capobianchi MR, Di Caro A, Autorino GL, Amaddeo D. 2011. Cowpox virus in Llama, Italy. *Emerging Infectious Diseases* **17**:1513–1515 DOI 10.3201/eid1708.101912.

Carroll DS, Emerson GL, Li Y, Sammons S, Olson V, Frace M, Nakazawa Y, Czerny CP, Tryland M, Kolodziejek J, Nowotny N, Olsen-Rasmussen M, Khristova M, Govil D, Karem K, Damon IK, Meyer H. 2011. Chasing Jenner's Vaccine: revisiting cowpox virus classification. *PLOS ONE* **6(8)**:e23086 DOI 10.1371/journal.pone.0023086.

Dabrowski WP, Radonić A, Kurth A, Nitsche A. 2013. Genome-wide comparison of cowpox viruses reveals a new clade related to Variola virus. *PLOS ONE* **8(12)**:e79953 DOI 10.1371/journal.pone.0079953.

Damaso CR. 2018. Revisiting Jenner's mysteries, the role of the Beaugency lymph in the evolutionary path of ancient smallpox vaccines. *The Lancet Infectious Diseases* **18**:e55–e63 DOI 10.1016/S1473-3099(17)30445-0.

Ducournau C, Ferrier-Rembert A, Ferraris O, Joffre A, Favier A-L, Flusin O, Van Cauteren D, Kecir K, Auburtin B, Védy S, Bessaud M, Peyrefitte CN. 2013.

Concomitant human infections with 2 cowpox virus strains in related cases, France, 2011. *Emerging Infectious Diseases* **19**:1996–1999 DOI 10.3201/eid1912.130256.

Eder I, Vollmar P, Pfeffer M, Naether P, Rodloff AC, Meyer H. 2017. Two distinct clinical courses of human cowpox, Germany, 2015. *Viruses* **9(12)**:Article 375 DOI 10.3390/v9120375.

Fassbender P, Zange S, Ibrahim S, Zoeller G, Herbstreit F, Meyer H. 2016. Generalized cowpox virus infection in a patient with HIV, Germany, 2012. *Emerging infectious diseases* **22**:553–555 DOI 10.3201/eid2203.151158.

Fenner F. 1990. Poxviruses. In: *Fields virology*. New York: Raven Press, 2113–2133.

Franke A, Pfaff F, Jenckel M, Hoffmann B, Höper D, Antwerpen M, Meyer H, Beer M, Hoffmann D. 2017. Classification of cowpox viruses into several distinct clades and identification of a novel lineage. *Viruses* **9(6)**:Article 142 DOI 10.3390/v9060142.

Hemmer CJ, Littmann M, Löbermann M, Meyer H, Petschaelis A, Reisinger EC. 2010. Human cowpox virus infection acquired from a circus elephant in Germany. *International Journal of Infectious Diseases* **14(3)**:e338–e340 DOI 10.1016/j.ijid.2010.03.005.

Hoffmann D, Franke A, Jenckel M, Tamošiunaite A, Schluckebier J, Granzow H, Hoffmann B, Fischer S, Ulrich RG, Höper D, Goller K, Osterrieder N, Beer M. 2015. Out of the Reservoir: phenotypic and genotypic characterization of a novel cowpox virus isolated from a common vole. *Journal of Virology* **89**:10959–10969 DOI 10.1128/JVI.01195-15.

ICTV. 2018. International Committee on Taxonomy of Viruses (ICTV). *Available at https://talk.ictvonline.org/taxonomy/* (accessed on 17 April 2018).

Jungwirth N, Puff C, Köster K, Mischke R, Meyer H, Stark A, Thoma B, Zöller G, Seehusen F, Hewicker-Trautwein M, Beineke A, Baumgärtner W, Wohlsein P. 2018. Atypical cowpox virus infection in a series of cats. *Journal of Comparative Pathology* **158**:71–76 DOI 10.1016/j.jcpa.2017.12.003.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780 DOI 10.1093/molbev/mst010.

Kurth A, Straube M, Kuczka A, Dunsche AJ, Meyer H, Nitsche A. 2009. Cowpox virus outbreak in banded mongooses (mungos mungo) and jaguarundis (herpailurus yagouaroundi) with a time-delayed infection to humans. *PLOS ONE* **4**:e6883 DOI 10.1371/journal.pone.0006883.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv preprint. arXiv:13033997.

Massung RF, Knight JC, Esposito JJ. 1995. Topography of variola smallpox virus inverted terminal repeats. *Virology* **211**:350–355 DOI 10.1006/viro.1995.1416.

Mauldin MR, Antwerpen M, Emerson GL, Li Y, Zoeller G, Carroll DS, Meyer H. 2017. Cowpox virus: what's in a name? *Viruses* **9(12)**:Article 101 DOI 10.3390/v9050101.

Ninove L, Domart Y, Vervel C, Voinot C, Salez N, Raoult D, Meyer H, Capek I, Zandotti C, Charrel NR. 2009. Cowpox virus transmission from pet rats to humans, France. *Emerging infectious diseases* **15**:781–784 DOI 10.3201/eid1505.090235.

Qin L, Favis N, Famulski J, Evans DH. 2015. Evolution of and evolutionary relationships between extant vaccinia virus strains. *Journal of Virology* **89**(**1809**):1809–1824 DOI 10.1128/JVI.02797-14.

Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13**:341 DOI 10.1186/1471-2164-13-341.

Quick J, Loman JN, Duraffour S, Simpson TJ, Severi E, Cowley L, Bore AJ, Koundouno R, Dudas G, Mikhail A, Ouédraogo N, Afrough B, Bah A, Baum JJH, Becker-Ziaja B, Boettcher PJ, Cabeza-Cabrerizo M, Camino-Sánchez Á, Carter LL, Doerrbecker J, Enkirch T, García-Dorival I, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou EL, Koropogui M, Kosgey A, Kuisma E, Logue HC, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallasch E, Patrono VL, Portmann J, Repits GJ, Rickett YN, Sachse A, Singethan K, Vitoriano I, Yemanaberhan LR, Zekeng GE, Racine T, Bello A, Sall AA, Faye O, Faye O, Magassouba N, Williams VC, Amburgey V, Winona L, Davis E, Gerlach J, Washington F, Monteil V, Jourdain M, Bererd M, Camara A, Somlare H, Camara A, Gerard M, Bado G, Baillet B, Delaune D, Nebie YK, Diarra A, Savane Y, Pallawo BR, Gutierrez JG, Milhano N, Roger I, Williams JC, Yattara F, Lewandowski K, Taylor J, Rachwal P, Turner JD, Pollakis G, Hiscox AJ, Matthews AD, O'Shea KM, Johnston MA, Wilson D, Hutley E, Smit E, Di Caro A, Wölfel R, Stoecker K, Fleischmann E, Gabriel M, Weller AS, Koivogui L, Diallo B, Keïta S, Rambaut A, Formenty P, Günther S, Carroll WM. 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature* **530**:228–232 DOI 10.1038/nature16996.

Reynolds MG, Guagliardo SAJ, Nakazawa YJ, Doty JB, Mauldin MR. 2018. Understanding orthopoxvirus host range and evolution: from the enigmatic to the usual suspects. *Current Opinion in Virology* **28**:108–115 DOI 10.1016/j.coviro.2017.11.012.

Ropp SL, Jin Q, Knight JC, Massung RF, Esposito JJ. 1995. PCR strategy for identification and differentiation of small pox and other orthopoxviruses. *Journal of Clinical Microbiology* **33**:2069–2076.

Schaudien D, Meyer H, Grunwald D, Janssen H, Wohlsein P. 2007. Concurrent infection of a cat with cowpox virus and feline parvovirus. *Journal of Comparative Pathology* **137**:151–154 DOI 10.1016/j.jcpa.2007.05.007.

Tcherepanov V, Ehlers A, Upton C. 2006. Genome Annotation Transfer Utility (GATU): rapid annotation of viral genomes using a closely related reference genome. *BMC Genomics* **7**:150 DOI 10.1186/1471-2164-7-150.

Walker JB, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo AC, Zeng Q, Wortman J, Young KS, Earl MA. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS ONE* **9**(**11**):e112963 DOI 10.1371/journal.pone.0112963.