

# SCIENTIFIC REPORTS

OPEN

## Molecular and microscopic characterization of a novel *Eastern grey kangaroopox virus* genome directly from a clinical sample

Subir Sarker<sup>1</sup>, Hayley K. Roberts<sup>1</sup>, Naomie Tidd<sup>2</sup>, Shayne Ault<sup>3,4</sup>, Georgia Ladmore<sup>4</sup>, Andrew Peters<sup>3</sup>, Jade K. Forwood<sup>5</sup>, Karla Helbig<sup>1</sup> & Shane R. Raidal<sup>2,3</sup>

Poxviruses are large DNA viruses with varying zoonotic potential, and are recognised in a broad range of wildlife. Although poxviruses have been detected in kangaroos, their genetic relationships to poxviruses in other animals and humans is not well understood. Here, we present a novel genome sequence of a marsupial poxvirus, the *Eastern grey kangaroopox virus* (EKPV-NSW), isolated from a wild eastern grey kangaroo. In the present study, histopathologically confirmed epidermal pox lesions were used to recover the full-length viral genome and perform electron microscopic analysis, with both immature virions and intracellular mature virions detected. Subsequent analysis of the EKPV-NSW genome demonstrated the highest degree of sequence similarity with EKPV-SC strain (91.51%), followed by WKPV-WA (87.93%), and MOCV1 (44.05%). The novel EKPV-NSW complete genome encompasses most of the chordopoxviruses protein coding genes (138) that are required for genome replication and expression, with only three essential protein coding genes being absent. The novel EKPV-NSW is missing 28 predicted genes compared to the recently isolated EKPV-SC, and carries 21 additional unique genes, encoding unknown proteins. Phylogenetic and recombination analyses showed EKPV-NSW to be the distinct available candidate genome of chordopoxviruses.

The family *Poxviridae* includes a wide range of pathogenic enveloped, double-stranded DNA (dsDNA) viruses with varying zoonotic potential that have been recognised in a broad range of wildlife taxa including hundreds of bird species, reptiles, marine mammals, macropods, marsupials, monotremes, ungulates, equids, and primates<sup>1–7</sup>. The subfamily *Entomopoxvirinae* delineates those infecting insects whereas the subfamily *Chordopoxvirinae* includes all genera infecting vertebrates. The International Committee on Taxonomy of Viruses (ICTV; <http://ictvonline.org/virusTaxonomy.asp>) recognises ten genera and one genus unassigned in the *Chordopoxvirinae* with the greatest diversity among the *Avipoxvirus* and *Orthopoxvirus* genera. The natural host range is a factor for establishing the taxonomic status of new poxvirus species. In some species the host range is narrow including only one known host. An example of this is the now extinct *Variola virus*, the cause of smallpox in humans. In other cases the natural host range may be very broad with regular cross-species infection, such as is the case for cowpox<sup>1,8</sup>.

As has recently been demonstrated for other DNA viruses<sup>9,10</sup>, the discovery of individual novel poxvirus species in new hosts can provide more accurate phylogenetic trees and improved taxonomic classification for the family *Poxviridae*<sup>11</sup>. At least 41 poxvirus species have had their whole genomes characterized and compared to most other pathogenic poxviruses, all of which have a large linear genome (128–360 kb) with a central core of approximately 80 relatively conserved genes that are recognized as the minimum essential poxvirus genome<sup>12</sup>. All poxviruses also contain a set of more diverse and presumably, host specific genes which are located towards the flanking or terminal regions of each genome<sup>13</sup>.

<sup>1</sup>Department of Physiology, Anatomy and Microbiology, School of Life Sciences, La Trobe University, Melbourne, VIC 3086, Australia. <sup>2</sup>Veterinary Diagnostic Laboratory, Charles Sturt University, Wagga Wagga, NSW 2678, Australia. <sup>3</sup>School of Animal and Veterinary Sciences, Charles Sturt University, Wagga Wagga, NSW 2678, Australia. <sup>4</sup>Lake Road Veterinary Clinic, 327 Lake Albert Road, Koorlingal, NSW 2650, Australia. <sup>5</sup>School of Biomedical Sciences, Charles Sturt University, Wagga Wagga, NSW 2678, Australia. Correspondence and requests for materials should be addressed to S.S. (email: [s.sarker@latrobe.edu.au](mailto:s.sarker@latrobe.edu.au))

*Eastern grey kangaroopox virus* (EKPV) has been nominated as a member of the subfamily *Chordopoxvirinae*, based on histopathological lesion, and electron microscopy<sup>14,15</sup>, as has the recently described novel *Eastern grey kangaroopox virus*<sup>16</sup>, but the ICTV has not yet approved it as a species. In Australia, kangaroo pox has been recognised in a wide range of native macropod species<sup>14,15,17</sup> including the eastern (*Macropus giganteus*) and western grey kangaroo (*Macropus fuliginosus*) species, red kangaroo (*Macropus rufus*), common wallaroo (*Macropus robustus*), tammar wallaby (*Macropus eugenii*), agile wallaby (*Macropus agilis*), swamp wallaby (*Wallabia bicolor*), Tasmanian pademelon (*Thylogale billardierii*) and quokka (*Setonix brachyurus*)<sup>18</sup>. So there is a strong likelihood that EKPV represents at least one naturally occurring pathogenic poxvirus species of kangaroos in Australia. It is also possible that macropod poxviruses may be species specific, as Speare (1988) cited an example of a captive mixed macropod colony, where only newly introduced eastern grey kangaroos were affected, while other species did not develop lesions. Nevertheless, the epidemiology of poxvirus infections in macropods, or marsupials more broadly, is not well characterised, with the common brushtail possum (*Trichosurus vulpecular*) and common ring-tail possum (*Pseudocheirus peregrinus*) being the only other marsupials to have reports of pox<sup>4</sup>. Other Australian indigenous animals, have also been reported to contract pox, with descriptions of the virus in the monotreme species of the short-beaked echidna (*Tachyglossus aculeatus*)<sup>14,17–20</sup>, and an anecdotal report of poxvirus infection in a southern bent-wing bat (*Miniopterus schreibersii bassanii*).

It is plausible that a poxvirus with wide host-specificity may be responsible for disease across a diverse range of marsupials, or there may be multiple poxvirus species present within Australia's highly diverse native mammals. This paper describes the genomic structure of a novel EKPV-NSW virus from a wild eastern grey kangaroo with pox.

## Results

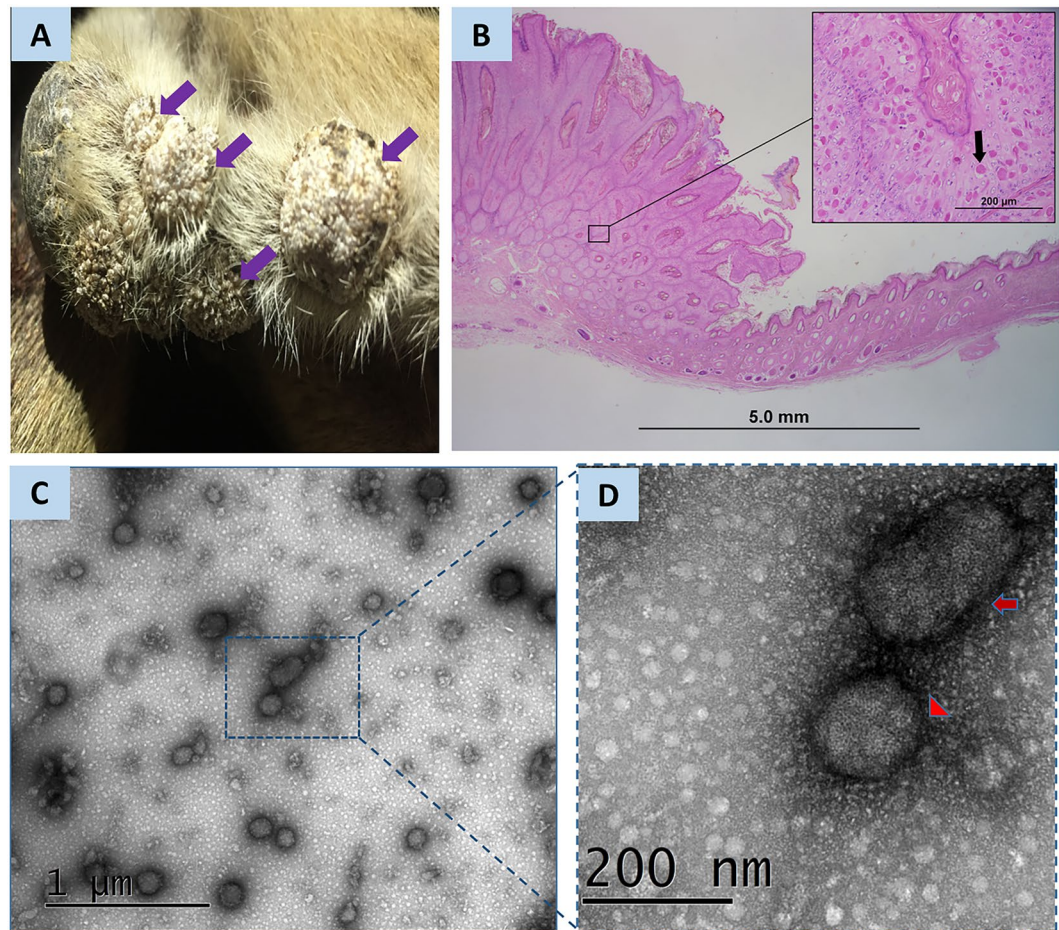
**Evidence of poxvirus infection in an Eastern grey kangaroo.** Affected nodules involved haired-skin of the limbs, tail, head and neck (Fig. 1A). Lesions were well demarcated and papillomatous with focal acanthosis, orthokeratotic hyperkeratosis and massive hyperplasia of the *stratum spinosum* with keratinocytes in this zone containing marked vacuolation and intracytoplasmic eosinophilic inclusions 20–30 µm in diameter which often displaced nuclei (Fig. 1B). At the margins of the lesions the thickened epidermis merged with normal epidermis. The centre of the largest lesions had prominent rete pegs and a pseudolobulated epidermal proliferation with closely associated elongated folds of epidermis containing centres of lamellated keratin. Beneath the affected areas there was a mild perivascular to diffuse infiltration of mixed mostly mononuclear inflammatory cells in the superficial dermis with lymphocytes and plasma cells dominating the reaction. There was mild superficial oedema present as well. The poxvirus infection was further confirmed by performing a PCR targeting 630 bp sequence of the conserved genomic domain encoding the RNA polymerase subunit gene<sup>21</sup>.

Transmission electron microscopic analysis of affected tissues clearly identified the two stages of enveloped, brick-shaped virus particles (Fig. 1C,D) indicative of active poxvirus infection in the eastern grey kangaroo. According to Harrion *et al.*<sup>22</sup>, two different stages of virus particles such as immature virion (IV) (Fig. 1D, red arrow head) and intracellular mature virion (IMV) (Fig. 1D, red arrow) were detected in the affected tissue collected from the eastern grey kangaroo. Morphologically, the immature virions were rectangular or ovoid shaped with rounded corners, whereas the intracellular mature virions were brick-shaped with the surface filaments have a spiral or criss-cross arrangement. The virions had a length of approximately 200 to 250 nm and a width of 120 to 150 nm.

**Genome structure and analysis of EKPV-NSW.** The EKPV-NSW genome is a linear double-stranded DNA molecule of 172,979 bp, with a relatively high GC (57.7%) content (Fig. 2) compared to other chordopoxviruses. The inverted terminal repeats (ITR) of EKPV-NSW encompass 2,919 bp each, which is similar to other poxviruses, and coordinates 1–2,919 sense orientation and 170,061–172,979 antisense orientation, however, the length of these ITRs were almost double in size compared to the recently isolated EKPV-SC strain. Interestingly, each of the inverted repeats, in turn, constitutes arrays of direct repeats. There were eight tandem repeats detected within each inverted terminal repeat region, and consisted of a 99, 66, two 48, two 35 and two 33 bp repeat units with 78%, 91%, 83%, 97% and 88% identity match, respectively. These direct repeat arrays are much smaller than those detected in other chordopoxviruses, however, the possibility that they extend beyond the sequenced portion of the genome cannot be ruled out.

The EKPV-NSW genome contained 169 predicted protein-coding genes and were numbered from left to right (Fig. 2, and Supplementary Table S1). Comparative analysis of the protein sequences encoded by the predicted ORFs to the sequences in the nonredundant protein sequence database at the National Center for Biotechnology Information (NIH, Bethesda, MD) using BLASTP identified homologs with significant protein sequence similarity (E value ≤ e-4) for 141 genes (Supplementary Table S1). Interestingly, EKPV-NSW contained 28 predicted protein coding genes that, at this time, appear to be unique to this virus. Among these unique protein-coding genes, 7 contained predicted transmembrane helices and/or a signal peptide, whereas 21 genes remained completely uncharacterized (Supplementary Table 1), and demonstrated no significant homology with known proteins in the NIH database.

Among the predicted protein-coding gene products of EKPV-NSW, 138 gene products had detected homologs to other chordopoxviruses (Supplementary Table S1). Among these conserved chordopoxvirus gene products, the highest number of protein-coding genes (132 genes) demonstrated homologs to the recently isolated EKPV-SC strain, with the remaining six gene products (ORF028, ORF050, ORF051, ORF079, ORF108 and ORF131) being homologous to ORFs from other poxviruses, including avipoxviruses, parapoxviruses and orthopoxvirus (Supplementary Table S1). All conserved genes of EKPV-NSW showed the highest sequence similarity to orthologs of the recently isolated EKPV strain SC, along with other chordopoxviruses, and these observations imply that the conserved EKPV-NSW genes shares a common evolutionary history, at least within the poxviruses



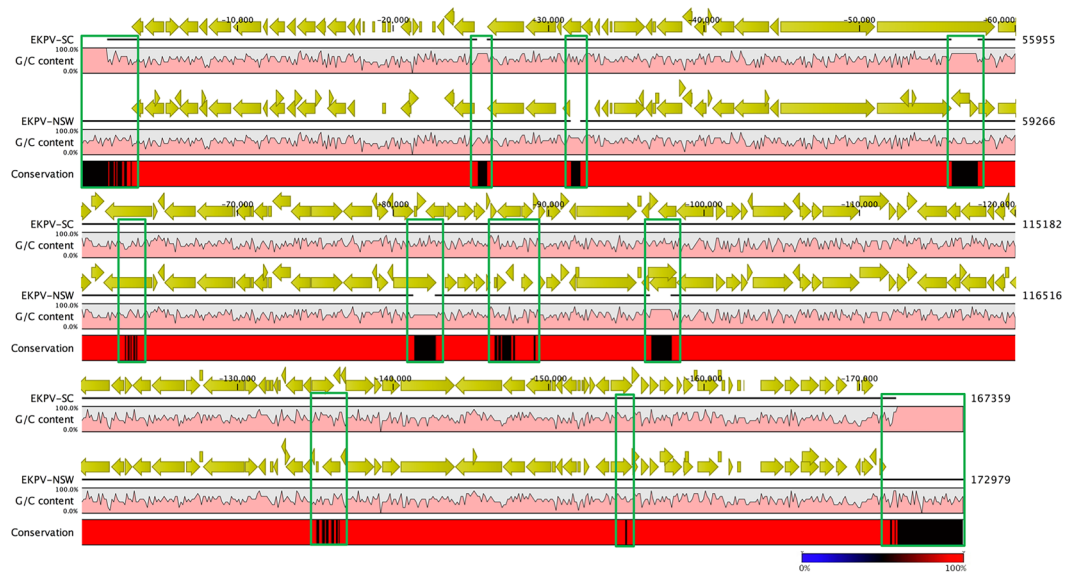
**Figure 1.** Pathological and transmission electron microscopic analysis of cutaneous papilloma infected with a EKPV-NSW. **(A)** Non-ulcerated, alopecic, raised, verrucous nodules (6–24 mm in diameter) with a sessile base (arrow). **(B)** Histological changes are characterized by massive hyperplasia of the *stratum spinosum* with keratinocytes in this zone containing marked vacuolation and intracytoplasmic eosinophilic inclusions 20–30 µm in diameter which often displaced nuclei (arrow). Stages of virus maturation are shown as immature virion (IV) (Fig. 1D, red arrow head), and intracellular mature virion (IMV) (Fig. 1D, red arrow).

(Supplementary Table S1). Overall, the central portion of the EKPV-NSW genome showed a nearly perfect conservation of gene synteny with other chordopoxviruses, with the terminal regions being highly divergent containing most of the predicted unique genes (Fig. 2 and Supplementary Table S1); which is also consistent with other chordopoxviruses<sup>23,24</sup>.

Considering the conserved genes among the tetrapod-infecting chordopoxviruses<sup>12</sup>, only three ancestral poxvirus genes; thioredoxin-like protein (A2.5L), endoplasmic reticulum-localized MP (E8R), and actin tail microtubule (F12L) appeared to have been lost in EKPV-NSW. In comparison to EKPV-SC, there were several occurrences of gene translocations, including the position of the genes (EKPV-NSW ORF072–078 and ORF125–130), which were observed to be translocated in EKPV-NSW than EKPV-SC (highlighted as grey in Supplementary Table S1). Relative to the EKPV-SC strain, the EKPV-NSW strain genome contained several inserted genes that had homologs to the FeP2 major envelope protein (ORF028) and the B22R family (ORF051); MOCV1 hypothetical protein (ORF050) and putative structural protein (ORF079); ORFV hypothetical protein (ORF108) and a FWPV putative myristylated membrane protein (ORF131) (Supplementary Table S1). Moreover, EKPV-NSW genome was also missing 28 predicted ORFs relative to the EKPV-SC strain. Among the missing genes, only six of the EKPV-SC genes, including palmitylated EEV membrane glycoprotein, virus entry/fusion complex component, virion phosphoprotein early morphogenesis, internal virion protein, IMV heparin binding surface protein and DNA helicase transcript release factor (EKPV-SC ORF30, 54, 69, 75, 86 and 124, respectively) appeared to be conserved among chordopoxviruses. All of the other predicted missing genes encode hypothetical protein located in the terminal regions of the genome.

Interestingly, three of the predicted genes (ORF048, ORF052, ORF163) had homologs to other cellular organisms, including eukaryotes and bacteria (Supplementary Table S1). ORF048 was found to have unexpected weak similarity (34% aa identity) to a bacterium (*Microcystis aeruginosa*) hypothetical protein (Supplementary Table S1), whereas the ORF052 showed very low similarity with the hypothetical protein of pectate lyase of *Paenibacillus* sp. (39% aa identity). Furthermore, the EKPV-NSW ORF163, translated to a predicted 367



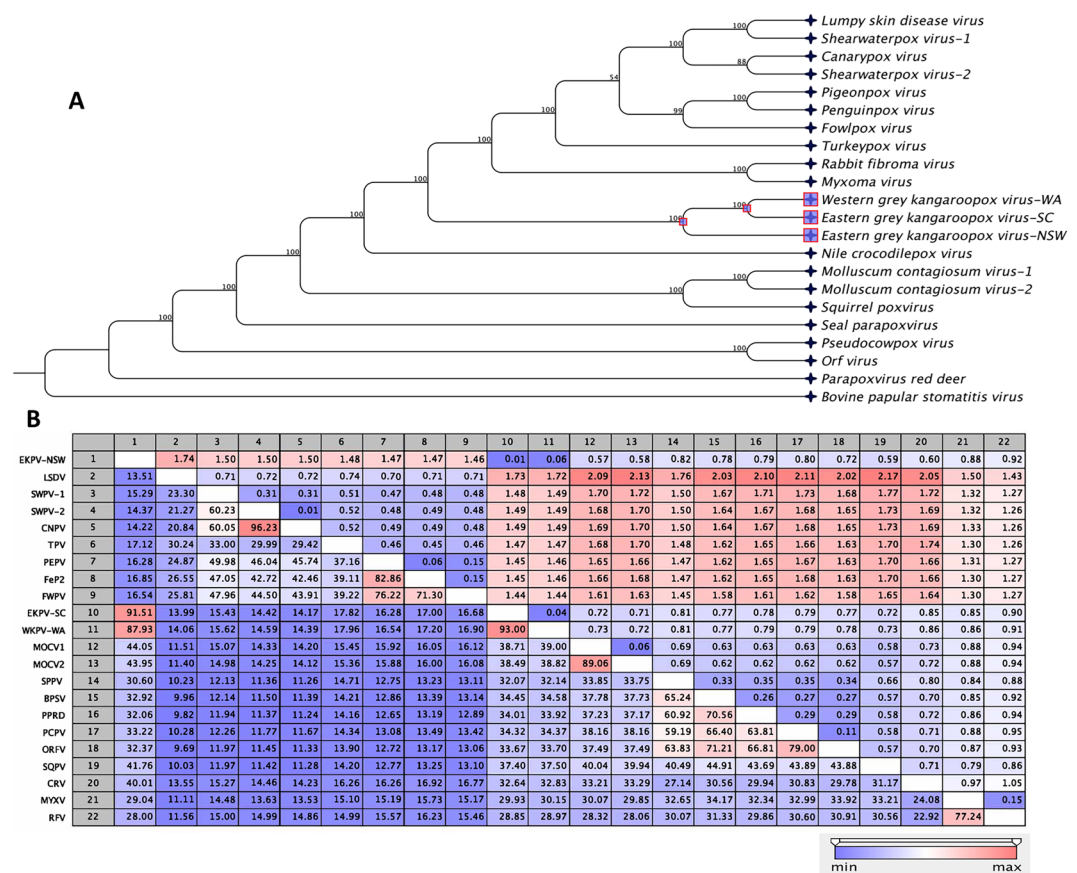


**Figure 2.** Comparative genome architectures between EKPV-NSW and EKPV-SC strain. Sequence alignment of *Eastern grey kangaropox virus* strain NSW (EKPV-NSW, GenBank accession number MF661791) to the reference *Eastern grey kangaropox virus* strain SC (EKPV-SC, MF467281) genome. The alignment was performed using the global alignment program contained in CLC Genomic Workbench (tool for Classical sequence analysis). The upper row represents the comparative ORFs map between EKPV-NSW and EKPV-SC. Protein coding ORFs, with banana colour depicting the direction of transcription. The middle row of each representative genome highlights the G + C% with each line plot represents the %G + C content of the indicated viral genome. The bottom graph represents the sequence conservation between the aligned EKPV-NSW and EKPV-SC sequences at a given coordinate at each position in the alignment. The gradient of the colour reflects the conservation of that particular position is in the alignment. If one position is 100% conserved the colour will be shown in red, and lower level of conservations are specified as gradient for colour. 50% conservation will be coloured as black, and it signifies that the consensus sequence has nucleotides support from one genome, and thus lower than 50% conservation in consensus sequence will not be observed in this Figure. Green boxes are used to highlights the most variable region of the genomes.

amino acids, and showed a significant similarity (55% aa identity) to a eukaryotic (*Elephantulus edwardii*) 3- $\beta$ -hydroxysteroid dehydrogenase (Supplementary Table S1).

**Evolutionary relationships of EKPV-NSW.** We used multiple-nucleotide alignments among the selected complete poxviruses genome sequences to construct a ML phylogenetic tree and calculate the distant matrix. The genome of EKPV-NSW strain shared the highest nucleotide sequence similarity to the EKPV-SC strain (91.51%), followed by WKPV-WA (87.93%), *Mollusum contagiosum virus* (MOCV1) (44.05%), *Squirrel poxvirus* (41.76%) and *Nile crocodilepox virus* (40.01%) (Fig. 3B). In the resulting ML tree, EKPV-NSW stain was placed in the same clade with the recently isolated kangaropox viruses (EKPV-SC and WKPV-WA) indicating that the kangaropox viruses (KPVs) were descendent from a common ancestor (Fig. 3A). Excepting the recently isolated KPVs, a greater distance between the EKPV-NSW and other selected poxvirus genome sequences were exhibited by highlighting the very low level of sequence identity to other poxviruses (14.27 to 44.05%) (Fig. 3A,B).

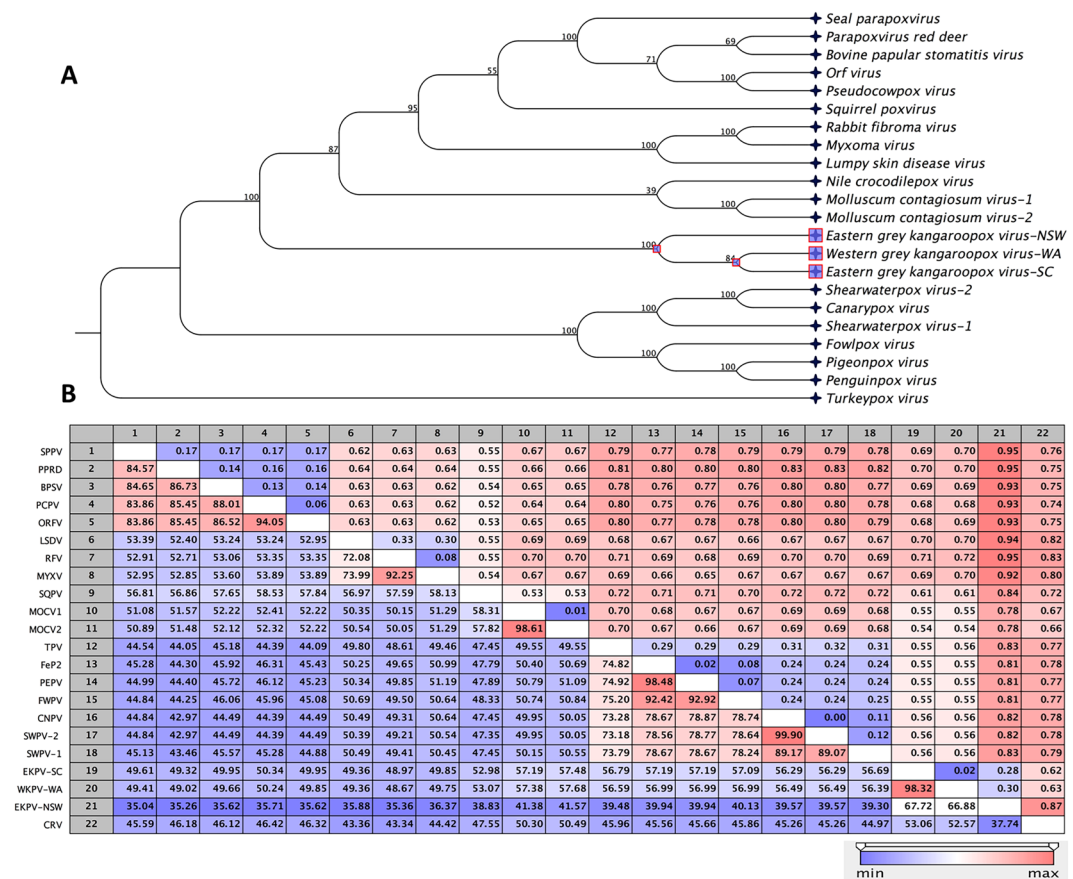
To gain further understanding into the evolution of the EKPV-NSW, we generated phylogenetic trees and a distance matrix calculated with the complete coding sequences of the DNA polymerase, DNA Topoisomerase and viral B22R-like gene, as has been performed previously<sup>24,25</sup>. The phylogenetic tree analysis together with the pairwise amino acid comparisons using the DNA polymerase gene demonstrated that KPVs placed between a clade consisted of MOCV and CRV, and the avipoxviruses (APVs) (Fig. 4A). However, EKPV-NSW generated a sister clade along with the recently isolated EKPV-SC and WKPV-WA<sup>16</sup>. The DNA polymerase gene orthologs of EKPV-NSW showed amino acid identities ranging from 35 to 67.72%, to other chordopoxviruses (Fig. 4B). This low-level average amino acid identity makes a much greater range of variation in the genome level of identity. However, a much closer and well-resolved evolutionary relationship was revealed using the DNA Topoisomerase coding sequences (Fig. 5A,B). Interestingly, there was a very strong clade support (100%) between the EKPV-NSW and the other two KPVs (EKPV-SC and WKPV-WA) (Fig. 5A), and the protein sequence of DNA Topoisomerase of EKPV-NSW showed the highest (100% aa identity) similarity with EKP-SC (Fig. 5B). By also building a phylogenetic tree with the viral B22R-like gene coding sequences, we discovered that several other viruses (e.g. *Turkeypox virus*, *Nile crocodilepox virus*, *Lumpy skin disease virus*) were distantly related within the EKPV-NSW clade (Supplementary Figure S1A and B).



**Figure 3.** Phylogenetic tree and pairwise comparison among selected complete genome sequences of poxviruses. **(A)** The ML tree was constructed from a multiple-nucleotide alignment from 22 complete genome sequences of poxviruses. The numbers on the left show bootstrap values as percentages, and EKPV-NSW was highlighted with red box. **(B)** Pairwise comparison among selected complete genome sequences of poxviruses. Upper comparison gradient indicated the distance between two sequences, and lower comparison gradient indicated percentage identity between two sequences. The abbreviations and GenBank accession details for poxviruses were used: MOCV1 (*Molluscum contagiosum virus* subtype 1, MCU60315); MOCV2 (*Molluscum contagiosum virus* subtype 2, KY040274); SQPV (*Squirrel poxvirus*, HE601899); CRV (*Nile crocodilepox virus*, DQ356948); SPPV (*Seal parapoxvirus*, KY382358); BPSV (*Bovine papular stomatitis virus*, KM875470); PPRD (*Parapoxvirus red deer*, KM502564); ORFV (*Orf virus*, DQ184476); PCPV (*Pseudocowpox virus*, GQ329670); MYXV (*Myxoma virus*, KP723391); RFV (*Rabbit fibroma virus*, AF170722); LSDV (*Lumpy skin disease virus*, NC\_003027); TPV (*Turkeypox virus*, KP728110); FeP2 (*Pigeonpox virus*, KJ801920); PEPV (*Penguinpox virus*, KJ859677); FWPV (*Fowlpox virus*, AF198100); SWPV-1 (*Shearwaterpox virus-1*, KX857216); SWPV-2 (*Shearwaterpox virus-2*, KX857215); CNPV (*Canarypox virus*, AY318871); EKPV-NSW (*Eastern Grey kangaroopox virus* strain NSW, MF661791); EKPV-SC (*Eastern Grey kangaroopox virus* strain Sunchine Coast, MF467281); WKPV-WA (*Western Grey kangaroopox virus* strain Western Australia, MF467280).

### Evidence of extensive recombination between EKPV-NSW and other Chordopoxviruses.

Recombination plays a critical role in the emergence of a new virus by avoiding genetic decline and creating novel traits<sup>26</sup>. Consequently, we further investigated the recombination pattern among selected poxvirus used in this study with the EKPV-NSW genome sequence using the RDP4 program<sup>27</sup>. Interestingly, we noted that there were 48 recombination events with significant *P*-values detected across the EKPV-NSW genome (Table 1). Importantly, the strongest support for recombination was overlapped within DNA polymerase protein gene (events 15, 20, 43 and 48), where the EKPV-NSW carried a form of this gene that was essentially descended from a common ancestral genome of MOCV1, MOCV2, ORFV, BPST EKPV-SC, WKPV-WA, TPV, SWPV-1 and -2, PEPV, FeP2, CNPV, FWPV and LSDV. Another potential example of recombination's affecting EKPV-NSW corresponded to the regions that contained the membrane protein gene (events 11, 16 and 31) and the RNA polymerase gene (events 12, 29, 34 and 36), which are likely to be a mixture of these gene within the member of the genera orthopoxvirus, avipoxviruses and parapoxviruses. A sequence fragment of approximately 1431 bp belonging to the polymerase catalytic subunit VP55, IEV morphogenesis (event 2) was the best supported (*p*-value: 1.94E-304) recombination event, and derived from EKPV-SC as a potential major parent with unknown minor parent. Remarkably, a large number of apparent recombination events (1, 5, 7–9, 11–12, 22–28, 34, 36–37, 40, 41, 44, 45–46) were involved mostly with MOCV-1 and MOCV-2. Recombination event 01 (ATPase NPH1;



**Figure 4.** Phylogenetic tree and pairwise comparison of the DNA polymerase genes. **(A)** The Neighbor Joining (NJ) tree was constructed from the protein sequences of selected poxviruses. The numbers on the left show bootstrap values as percentages, and EKPV-NSW was highlighted with red box. **(B)** Pairwise comparison of protein sequences of DNA polymerase genes of poxviruses. Upper comparison gradient indicated the distance between two protein sequences, and lower comparison gradient indicated percentage identity between two protein sequences. The abbreviations for poxviruses were used: MOCV1, *Molluscum contagiosum virus* subtype 1; MOCV2, *Molluscum contagiosum virus* subtype 2; SQPV, *Squirrel poxvirus*; CRV, *Nile crocodilepox virus*; SPPV, *Seal parapoxvirus*; BPSV, *Bovine papular stomatitis virus*; PPRD, *Parapoxvirus red deer*; ORFV, *Orf virus*; PCPV, *Pseudocowpox virus*; MYXV, *Myxoma virus*; RFV, *Rabbit fibroma virus*; LSDV, *Lumpy skin disease virus*; TPV, *Turkeypox virus*; FeP2, *Pigeonpox virus*; PEPV, *Penguinpox virus*; FWPV, *Fowlpox virus*; SWPV-1, *Shearwaterpox virus-1*; SWPV-2, *Shearwaterpox virus-2*; CNPV, *Canarypox virus*; EKPV-NSW, *Eastern Grey kangaroopox virus* strain NSW; EKPV-SC, *Eastern Grey kangaroopox virus* strain Sunchine Coast; WKPV-WA, *Western Grey kangaroopox virus* strain Western Australia.

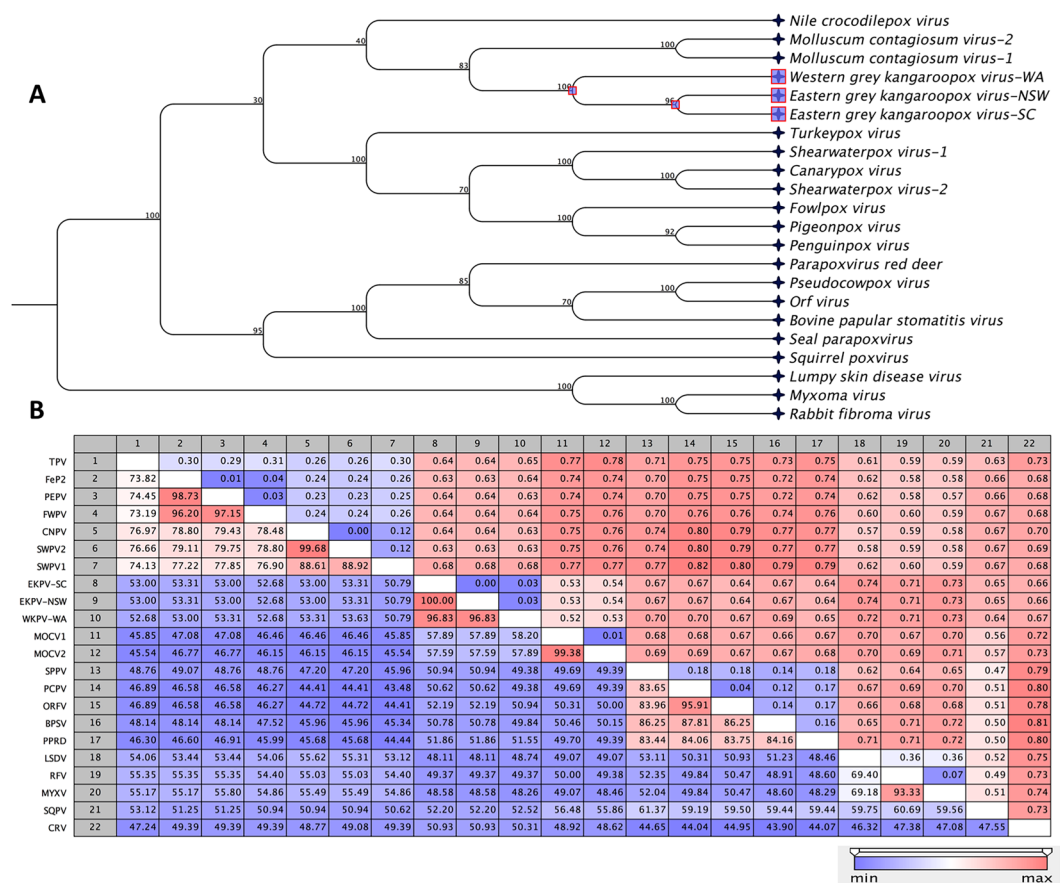
$p$ -value:  $7.53E-297$ ) for example, where the isolate EKPV-NSW was identified as an apparent recombinant of isolates MOCV-1 and MOCV-2 as minor parents, and EKPV-SC and WKPV-WA as major parents. It should be noted that there were also a few apparent recombination events overlapping within the intergenic region and or hypothetical proteins, which might suggest the recombination events with another as yet unidentified poxvirus genome.

## Discussion

This paper fully characterises a novel poxvirus genome from a marsupial. Pox lesions have been recognised in kangaroos for at least 3 decades with cases reported in a female eastern grey kangaroo held in captivity on a property in Nerang, Queensland, along with molluscum contagiosum in a red kangaroo in the 1970s. With the exception of very recently isolated KPVs<sup>16</sup>, poxvirus-like particles were confirmed to be present in the lesions electronmicroscopically<sup>14,15</sup>, however, no taxonomic classification has been granted yet by the ICTV and until now, unresolved relationships with other members of the family *Poxviridae*. In this study, we confirmed the infection of grey kangaroo with a novel species of poxvirus, and determined the complete genome sequence with predicted full-length coding regions including ITRs, and nominate this as the type genome for *Eastern grey kangaroopox virus* strain NSW (EKPV-NSW). We also further confirmed the evolutionary relationship of EKPV with other members of the *Chordopoxvirinae* subfamily.

The genome of EKPV-NSW is sufficiently diverse from recently isolated KPVs<sup>16</sup>, as well as all other known poxvirus genomes sequenced to date, and we propose for it to be classified within a new poxvirus genus





**Figure 5.** Phylogenetic tree and pairwise comparison of the DNA topoisomerase I genes. **(A)** The Neighbor Joining (NJ) tree was constructed from the protein sequences of selected poxviruses. The numbers on the left show bootstrap values as percentages, and EKPV-NSW was highlighted with red box. **(B)** Pairwise comparison of protein sequences of DNA topoisomerase I genes of poxviruses. Upper comparison gradient indicated the distance between two protein sequences, and lower comparison gradient indicated percentage identity between two protein sequences. The abbreviations for poxviruses were used: MOCV1, *Molluscum contagiosum virus* subtype 1; MOCV2, *Molluscum contagiosum virus* subtype 2; SQPV, *Squirrel poxvirus*; CRV, *Nile crocodilepox virus*; SPPV, *Seal parapoxvirus*; BPSV, *Bovine papular stomatitis virus*; PPRD, *Parapoxvirus red deer*; ORFV, *Orf virus*; Pseudocowpox virus; MYXV, *Myxoma virus*; RFV, *Rabbit fibroma virus*; LSDV, *Lumpy skin disease virus*; TPV, *Turkeypox virus*; FeP2, *Pigeonpox virus*; PEPV, *Penguinpox virus*; FWPV, *Fowlpox virus*; SWPV-1, *Shearwaterpox virus-1*; SWPV-2, *Shearwaterpox virus-2*; CNPV, *Canarypox virus*; EKPV-NSW, *Eastern grey kangaroopox virus-strain NSW*; EKPV-SC, *Eastern grey kangaroopox virus-strain SC*; WKPV-SC, *Western grey kangaroopox virus-strain WA*.

*Marsupialipoxvirus*. The observed sequence similarity of EKPV-NSW with other poxviruses is expected, given the extensive genetic diversity, and phylogeographic separation of KPVs from other *Poxviridae*<sup>28</sup>. Genome analysis of selected complete poxvirus sequences established the position of this EKPV-NSW as closest to the EKPV-SC (91.51% sequence identity), followed by WKPV-WA (87.93%), MOCV1 (44.05% sequence identity), SQPV (41.76% sequence identity) and CRV (40.01% sequence identity) (Fig. 3A,B). Further phylogenetic and distance matrix analysis using DNA polymerase, DNA topoisomerase and viral B22R-like genes did not show any obvious trend associated in the organisation of the genera except KPVs in phylogenetic trees. When coding sequences of DNA polymerase gene was compared, there was a significant genetic distance (0.28–0.30) estimated between EKPV-NSW and other KPVs (EKPV-SC and WKPV-WA). However, comparing the genetic distances of viral B22R-like genes, there was no obvious trend between EKPV-NSW and other poxviruses. A well-supported phylogenetic tree was generated using the DNA topoisomerase protein coding sequences. This demonstrated that EKPV-NSW was closely related to EKPV-SC and WKPV-WA. These findings infer that EKPV-NSW may be more closely related at a conserved gene level than across the complete genome, and highlights the value of complete genome characterization compared to single gene phylogenies.

Like many other members of the *Chordopoxvirinae*, and given the distribution of lesions along the limbs and extremities of affected kangaroos, transmission of EKPV is likely via arthropod vectors or close contact, although we cannot attribute the actual causality of this infection in the eastern grey kangaroo. Previous reports of kangaroo pox have described single lesions as the typical clinical presentation<sup>14,18</sup> whereas this report and others<sup>29</sup> describe multiple skin lesions akin to swinepox and smallpox diseases that resolve spontaneously over months.

Event number	BP* start in EKPV-NSW	BP* end in EKPV-NSW	Recombinant gene in EKPV-NSW	Recombinant sequence(s)	Minor parental sequence(s)	Major parental sequence(s)	Detection methods	P-value
1	110227	111051	ATPase NPH1	EKPV-NSW	MOCV-1, MOCV-2	EKPV-SC, WKPV-WA	RGBMCST	7.53E-297
2	40036	41467	polymerase catalytic subunit VP55, IEV morphogenesis	EKPV-NSW	Unknown	EKPV-SC	GMCS	1.94E-304
3	87937	88163	polymerase small subunit VP39	EKPV-NSW	PCPV, SPPV, BPSV, PPRD, ORFV	EKPV-SC, WKPV-WA	RGBMCT	5.22E-80
4	88163	88697	polymerase small subunit VP39, RNA polymerase	EKPV-NSW	Unknown	EKPV-SC	RGBMCS	7.44E-60
5	148077	148987	RNA polymerase, IMV protein, A31R conserved protein	PCPV	EKPV-NSW, EKPV-SC, WKPV-WA	MOCV-1, MOCV-2	RGBMCS	3.95E-31
6	3493	3956	hypothetical protein	EKPV-SC	WKPV-WA	EKPV-NSW	RB	6.03E-15
7	55947	56368	hypothetical protein	MOCV-1, MOCV-2	EKPV-NSW, EKPV-SC, WKPV-WA	PPRD, SPPV, BPSV	RGBST	2.61E-25
8	114663	148077	VLTF-2	CRV	MOCV-1, MOCV-2	EKPV-NSW, EKPV-SC, WKPV-WA	MCT	3.94E-17
9	131907	133223	putative myristylated membrane protein, IMV membrane protein entry/fusion complex component	CRV	EKPV-NSW, EKPV-SC, WKPV-WA	MOCV-1, MOCV-2	BMCS	3.29E-32
10	60028	60569	ER-localized membrane protein virion core protein	LSDV	EKPV-NSW, EKPV-SC, WKPV-WA	FWPV, SWPV-1, PEPV, FeP2	GMT	2.76E-18
11	65353	66460	membrane protein	MOCV-1, MOCV-2	EKPV-NSW, EKPV-SC, WKPV-WA	BPSV	RBM	2.96E-13
12	137003	137580	RNA polymerase	CRV	EKPV-NSW, WKPV-WA	MOCV-1, MOCV-2	BMCS	1.26E-14
12	87904	87936	polymerase small subunit VP39	EKPV-NSW	Unknown	EKPV-SC	GB	5.19E-18
14	125418	128688	P4a precursor, viral membrane formation, virion core and cleavage processing protein, IMV membrane protein virion maturation	MYXV, RFV	Unknown	EKPV-NSW, WKPV-WA	RMCST	8.10E-27
15	60882	61305	DNA polymerase	TPV, SWPV-1	EKPV-NSW, EKPV-SC	LSDV	ST	9.92E-12
16	64771	64898	membrane protein	FWPV, PEPV, FeP2	EKPV-NSW	LSDV, EKPV-SC	GT	1.20E-11
17	57782	57970	hypothetical protein	EKPV-NSW, EKPV-SC, WKPV-WA	FWPV, PEPV, FeP2	SQPV	RBMCT	1.45E-10
18	67136	67344	hypothetical protein	BPSV, SPPV, PPRD, PCPV, ORFV	Unknown	EKPV-NSW, EKPV-SC, WKPV-WA	RBMCT	1.16E-11
19	32702	25273	RhoA signaling inhibitor virus release protein, EEV maturation protein, major envelope antigen, hypothetical protein	FWPV, PEPV, FeP2	EKPV-NSW, EKPV-SC, WKPV-WA	ORFV	RGMCL	3.37E-21
20	62057	62533	DNA polymerase	PEPV, SWPV-1, SWPV-2, CNPV, TPV, FeP2, FWPV	WKPV-WA	LSDV, EKPV-NSW, EKPV-SC	RST	1.31E-08
21	67162	67344	hypothetical protein	EKPV-NSW, EKPV-SC, WKPV-WA	SWPV-2	CRV	RB	1.36E-08
22	149100	149425	ATPase/DNA packaging protein	CRV	EKPV-NSW, EKPV-SC	MOCV-1, MOCV-2	RBMCT	1.63E-08
23	146182	147838	P4c precursor, IMV MP/virus entry, RNA polymerase	MOCV-1, MOCV-2	EKPV-NSW, EKPV-SC, WKPV-WA	LSDV	MCS	1.02E-40
24	18695	19706	hypothetical protein	MOCV-1, MOCV-2	SPPV	EKPV-NSW, EKPV-SC, WKPV-WA	RBMCT	3.75E-10
25	108137	108316	morphogenesis VETF-s	BPSV, SPPV, PPRD	MOCV-2	EKPV-NSW, EKPV-SC, WKPV-WA	RGBMCS	2.87E-22
26	67625	67842	hypothetical protein	EKPV-NSW, EKPV-SC, WKPV-WA, MOCV-1, MOCV-2	TPV, SWPV-1, SWPV-2, CNPV, FWPV	CRV	RT	2.81E-06
27	23134	23473	S-S bond formation pathway protein substrate	CRV	MOCV-1, MOCV-2	EKPV-NSW	BS	8.95E-09
28	42072	43581	IEV morphogenesis	CRV	MOCV-1, MOCV-2	EKPV-NSW	MC	5.28E-07
29	136915	137595	RNA polymerase	PPRD, BPSV, PCPV, ORFV	Unknown	EKPV-NSW, EKPV-SC, WKPV-WA	RBCST	2.91E-07
30	78247	78586	VLTF	EKPV-NSW, EKPV-SC, WKPV-WA	Unknown	PPRD, BPSV	RGMCT	1.00E-04
31	65769	65979	membrane protein	EKPV-NSW, EKPV-SC, WKPV-WA	PEPV, SWPV-1, SWPV-2, CNPV, FeP2, FWPV	SPPV, PPRD, PCPV, ORFV	RT	6.28E-05
32	48582	48707	surface glycoprotein	SQPV	EKPV-NSW	PPRD, ORFV	MCT	1.79E-04

Continued



Event number	BP* start in EKPV-NSW	BP* end in EKPV-NSW	Recombinant gene in EKPV-NSW	Recombinant sequence(s)	Minor parental sequence(s)	Major parental sequence(s)	Detection methods	P-value
33	68041	68135	hypothetical protein	EKPV-NSW, EKPV-SC, WKPV-WA	PEPV, SWPV-1, SWPV-2, CNPV, FeP2, FWPV, CNPV	CRV	RCT	8.82E-03
34	91766	91963	RNA polymerase	SQPV	EKPV-NSW, EKPV-SC, WKPV-WA	MOCV-1, MOCV-2	RGCT	1.07E-04
35	68689	68889	hypothetical protein	EKPV-NSW, EKPV-SC, WKPV-WA	FWPV	SPPV	RBST	7.65E-19
36	91818	91968	RNA polymerase	MOCV-1, MOCV-2	Unknown	EKPV-NSW, EKPV-SC, WKPV-WA	MS	2.01E-05
37	149766	150084	ATPase/DNA packaging protein	MOCV-1, MOCV-2	CRV	EKPV-NSW, EKPV-SC, WKPV-WA	RS	1.08E-07
38	141264	142194	A type inclusion protein	PPRD, SPPV, BPSV, PCPV, ORFV	EKPV-NSW, EKPV-SC, WKPV-WA	SWPV-1, SWPV-2	RGMC	3.50E-03
38	131065	131179	IMV membrane protein	EKPV-NSW, EKPV-SC, WKPV-WA	TPV	SQPV	RT	5.62E-04
40	102834	103852	virion core, uracil-DNA glycosylase DNA polymerase processivity factor	EKPV-NSW, EKPV-SC, WKPV-WA	MOCV-1, MOCV-2	Unknown	RMC	9.64E-12
41	50108	51197	surface glycoprotein	MOCV-1, MOCV-2	SPPV	EKPV-NSW	RMCS	2.85E-43
42	114234	114479	trimeric virion coat protein	SPPV	EKPV-NSW, PPRD	Unknown	MCS	1.35E-159
43	60232	60409	DNA polymerase	EKPV-NSW, EKPV-SC, WKPV-WA	FWPV	ORFV	RB	1.65E-04
44	95069	96086	RAP94	ORFV, PPRD, PCPV	MOCV-2	EKPV-NSW, EKPV-SC, WKPV-WA	BCT	4.88E-03
45	4303	5126	Hypothetical protein	EKPV-NSW, EKPV-SC	Unknown	MOCV-2	ST	7.20E-04
46	47997	49575	surface glycoprotein	MOCV-1, MOCV-2	EKPV-NSW	PPRD	BMCT	1.16E-20
47	39048	39159	nonfunctional serine recombinase	EKPV-NSW	SWPV-1, CNPV	PPRD	RC	2.04E-04
48	61169	61342	DNA polymerase	PPRD, SPPV, BPSV, PCPV, ORFV	EKPV-NSW	SQPV	RB	2.01E-04

**Table 1.** Predicted recombination events between EKPV-NSW and other Chordopoxviruses. Details recombination events detected between EKPV-NSW and other Chordopoxviruses. The 48 recombination events were detected from entire poxviruses genomes analysed in this study using RDP4 program. Detection method coding R, G, B, M, C, S, P, L and T represents methods RDP, GENECONV, Bootscan, MaxChi, Chimaera, SiScan, PhylPro, LARD and 3Seq, respectively. The P value for the detection method in bold is shown. The abbreviations for poxviruses were used: MOCV, *Molluscum contagiosum virus*; SQPV, *Squirrel poxvirus*; CRV, *Nile crocodilepox virus*; SPPV, *Seal parapoxvirus*; BPSV, *Bovine papular stomatitis virus*; PPRD, *Parapoxvirus red deer*; ORFV, *Orf virus*; PCPV, *Pseudocowpox virus*; MYXV, *Myxoma virus*; RFV, *Rabbit fibroma virus*; LSDV, *Lumpy skin disease virus*; TPV, *Turkeypox virus*; FeP2, *Pigeonpox virus*; PEPV, *Penguinpox virus*; FWPV, *Fowlpox virus*; SWPV-1, *Shearwaterpox virus-1*; SWPV-2, *Shearwaterpox virus-2*; CNPV, *Canarypox virus*; EKPV-NSW, *Eastern grey kangaroopox virus-strain NSW*; EKPV-SC, *Eastern grey kangaroopox virus-strain SC*; WKPV-SC, *Western grey kangaroopox virus-strain WA*.

The eastern grey kangaroo is the most common and the second largest marsupial and is widespread throughout eastern Australia, preferring dry mesic habitat. Depending on climatic and seasonal variations the population fluctuates between 11–30 million animals<sup>30</sup>. Australian macropods encroaching on urban and wild-forest areas, are currently the major reservoir of human pathogens such as *Ross river virus*, macropods are required for both the maintenance of the virus in nature and its transmission to humans<sup>31</sup>. Therefore, a key concern with EKPV is whether or not it has zoonotic potential.

Recent studies on viral genome sequencing suggests that recombination has facilitated the evolution of human and animal pathogens, including *Variola virus* and *Vaccinia virus*<sup>26,32,33</sup>. Remarkably, the EKPV-NSW complete genome also provides evidence of extensive recombination among chordopoxviruses, which may indicate past host switching and recombination in the evolution of poxviruses more generally. A large number of conserved genes in EKPV-NSW appear to be evolved via homologous or non-homologous site-specific recombination with other poxviruses. For example, the genes encoding the DNA polymerase gene, membrane protein gene, RNA polymerase, polymerase catalytic subunit VP55, IEV morphogenesis, and ATPase NPH1 were predicted to be recombinants between EKPV-NSW and other poxviruses including species of the genera *Molluscipoxvirus*, *Orthopoxvirus*, *Parapoxviruses* and *Avipoxviruses*. Such evidence further indicates that the EKPV-NSW might be a mosaic of genetic materials obtained via multiple recombination events.

The EKPV-NSW genome has a relatively high GC (57.7%) content compared to recently isolated KPVs (54%), as well as other chordopoxviruses which range from 25–67%<sup>34</sup>, although, the biological significance for this variation is not yet known. High GC ratio poxviruses include those belonging to the *molluscipoxvirus*, *Crocodylidpoxvirus* and *Parapoxvirus*<sup>35</sup>, which might explain further why EKPV-NSW showed the second highest similarity with the *Molluscum contagiosum virus*, *Nile crocodilepox virus* and *Squirrel poxvirus*. Studies by Hatcher

*et al.*<sup>35</sup> also postulated that high GC ratio viruses may be subject to different evolutionary mechanisms, including a lack of gene reduction as compared to viruses with higher AT content.

The disparity between the gene complements of EKPV-NSW and EKPV-SC are relative consistent in identity levels, with 21 genes remaining completely uncharacterized (Supplementary Table 1) and a further 28 EKPV-SC genes being missing. Most of the genes involved in genome replication and expression, viral membrane biosynthesis, as well as core and capsid structure and morphogenesis are well conserved in EKPV-NSW. However, extreme variation in identity levels were observed with other proteins, with the EKPV-NSW DNA polymerase (ORF055) having a high sequence identity of 67% to the homologous protein in EKPV-SC, whereas the putative structural protein precursor (ORF079) shares no more than 44% identity with other poxvirus proteins.

Recent studies have shown that some genes encoded by the members of the subfamily *Chordopoxvirinae* exhibit greater similarity to eukaryotic and bacterial genes, highlighting their potential roles in the evolution of poxvirus in various host species<sup>33</sup>. Three of the predicted genes of EKPV-NSW had high scores against bacteria (ORF048, ORF052) and eukaryotic proteins (ORF163), and essentially insignificant BALSTP scores against viral proteins. For instance, the predicted EKPV-NSW ORF048 gene which returned an unexpected weak similarity to a bacterium (*Microcystis aeruginosa*) hypothetical protein, whereas the ORF052 showed very low similarity with the hypothetical protein of pectate lyase of *Paenibacillus* sp (Supplementary Table S1). Furthermore, the EKPV-NSW ORF163 is most likely of eukaryotic origin considering the significant BALSTP scores against a eukaryotic (*Elephantulus edwardii*) 3-beta-hydroxysteroid dehydrogenase, although this catalytic enzyme recently been identified in several poxviruses such as skunkpox, volepox, and raccoonpox<sup>36</sup>. This may suggest that the members of the family poxvirus may carry these eukaryotic and or bacterial genes simply as poxviruses are more effective in capturing or maintaining host genes. Although few of these genes have homologs to those found in other poxviruses, indicating a potential for horizontal gene transfer events, and this unique complement likely underlines EKPV mechanisms of virulence and host adaption.

## Conclusion

We have characterized a novel full-length *Eastern grey kangaropox virus*-NSW strain from a marsupial, which is significantly divergent, but most similar to a recently isolated *Eastern grey kangaropox virus*-SC strain. Except recently isolated KPVs, EKPV-NSW is not closely related to any other Chordopoxvirinae genome so far isolated from other natural host species. In addition, it encodes 21 predicted protein-coding genes which have remained completely uncharacterized. This novel EKPV-NSW complete genome sequence is missing 28 genes, and it has been observed that there are several occurrences of gene translocations compared to EKPV-SC. Further experimental study of EKPV pathogenesis, and an understanding of the functions of uncharacterized proteins will provide a unique approach to better assess the risk associated with the zoonotic potential.

## Methods

**Source of samples.** Samples were obtained from a sub-adult male eastern grey kangaroo in good body condition that was died accidentally by a motor vehicle in Wagga Wagga, NSW (ID: 17–1662; year of sampling: 2017; GPS latitude: −35.126°S; 147.378°E). Animal sampling was carried out by the attending veterinarian for the investigation of the novel clinical presentation, and sent to the Veterinary Diagnostic Laboratory, Charles Sturt University for further analysis. Animal sampling was obtained in accordance with approved guidelines set by the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes (1997) and approved by the Charles Sturt University Animal Ethics Committee (Research Authority permit 09/046). The kangaroo had multifocal skin lesions along both hind limbs extending from the dorsal aspects of the feet to the stifle. On each leg there were 10–14 affected areas of skin measuring 6–24 mm in diameter as non-ulcerated, alopecic, raised, verrucous nodules with a sessile base (Fig. 1A). There were also numerous circular to ovoid non-ulcerated, partially haired demarcations where healing seemed to be underway. Nodules were collected into formalin for histopathological examination. Other nodules were also collected and stored at −20 °C.

**DNA extraction.** Total genomic DNA was isolated from the affected skin tissue according to the protocol described by Sarker *et al.*<sup>37</sup> using a ReliaPrep gDNA Tissue Miniprep System (Promega, USA). A total of approximately 25 mg of skin tissue were aseptically dissected and chopped into small pieces and transferred into a 1.5 mL microcentrifuge tube (Eppendorf). Virion enrichment was performed by centrifugation for 2 minutes at 800 × g to remove tissue debris, and the supernatants were subsequently filtered through 5 µm centrifuge filters (Millipore)<sup>38</sup>. The filtrates were nuclease treated to remove unprotected nucleic acids using 20 µL of RNase A (Promega, USA), and incubated in microcentrifuge tube at 56 °C for 10 min. Viral nucleic acids were subsequently extracted using ReliaPrep binding columns (Promega, USA).

**Library construction and MiSeq sequencing.** The protocol for DNA-seq library preparation was adapted using the Illumina Nextera XT DNA Library Prep V3 Kit. The library was generated using one ng of total genomic DNA (gDNA). Tagmentation of gDNA was performed in a mixture contained 10 µL Tagment DNA buffer (Illumina<sup>®</sup> Inc., San Diego, CA, USA) with one ng of gDNA, and the reaction mixture was run in a thermal cycler at 55 °C for 5 min. When the temperature reached 10 °C, 5 µL of Neutralize Tagment buffer (Illumina<sup>®</sup> Inc., San Diego, CA, USA) was immediately added and the reaction incubated for 5 min at room temperature to inactivate the transposome.

Amplification of the tagmented DNA was performed using a limited-cycle PCR program. PCR amplification of selectively tagmented DNA fragments (50 µL) contained 25 µL of tagmented DNA, 5 µL of index 1 (i7), 5 µL of index 2 (i5) and 15 µL of Nextera PCR Master Mix (Illumina<sup>®</sup> Inc., San Diego, CA, USA). PCR amplification

was performed with the following temperature cycling profile: 72 °C initial denaturation for 3 min; 95 °C denaturation for 30 s; 12 cycles of 95 °C for 10 s, 55 °C for 30 s, and 72 °C for 30 s; and 72 °C final extension step for 5 min. The amplified library was cleaned to remove PCR-generated adaptor-dimers using an AMPure XP beads (Invitrogen™, USA) according to the protocol described in Illumina Nextera XT DNA Library Prep V3 Kit with final elution in 50 µL of resuspension buffer.

The quality and quantity of the prepared library was assessed using an Agilent Tape Station (Agilent Technologies) by Genomic Platform, La Trobe University, and it was confirmed that the average size of the insert was 625 bp. The prepared library was normalized and pooled in equimolar quantity. The quality and quantity of the final pooled library was assessed as described above prior to sequencing by the facility. Cluster generation and sequencing of the pooled DNA-library was sequenced as paired-end on Illumina® MiSeq chemistry according to the manufacturer's instructions.

**Bioinformatics.** The sequencing with the Illumina Miseq chemistry gave a total of 1,194,060 pairs of 301-bp reads. The sequencing data was analysed according to an established pipeline<sup>11</sup> in CLC Genomics workbench 9.5.4 under the La Trobe University Genomics Platform. Briefly, the preliminary quality evaluation for each raw read was generated using quality control (QC) report. The raw data were pre-processed to remove ambiguous base calls, and bases or entire reads of poor quality using default parameters. The datasets were trimmed to pass the quality control based on PHRED score 0.01 (equivalent to 20) or per base sequence quality score. Trimmed sequence reads were mapped against an Australian kangaroo (*Macropus eugenii*, GenBank accession number GL044074T)<sup>39</sup> to remove possible remaining host DNA contamination, and a total of 4.96% reads were excluded from further analysis that had mapped to the reference genome. Using the *de novo* assembler contained in CLC Genomic workbench, the remainder of the reads (95.04%) were assembled, and a total of 20,411 contigs were generated. The contigs that had high confidence (average coverage > 100) were selected for downstream analysis. A BLASTN search on the resulted contigs confirmed the closest match with human *Molluscan contagious virus* sequences. The selected contigs were then mapped against *Molluscan contagious virus* genome sequence as a reference genome in Geneious with medium sensitivity and 1000-times iteration (version 10.2.3, Biomatters, New Zealand). Mapped consensus sequence was checked carefully, and manually edited for gaps tweaked, ordered and oriented where necessary. The draft genome thus generated from *de novo* contigs were then used as a reference genome to assemble against the trimmed reads (95.04% reads) to validate further, which matched perfectly and produced a 172,979 bp (> 172 Kbs) consensus genome for the *Eastern grey kangaroopox virus* obtained from an Eastern Grey Kangaroo.

**Genome annotations.** The EKPV-NSW genome was annotated using GATU<sup>40</sup> to capture all the potential open reading frames (ORFs) that have more than 50 amino acids with minimal overlapping (overlaps cannot exceed 25% of one of the genes). Intergenic regions were further checked for the presence of ORFs using the CLC Genomics Workbench (CLC) analysis tool (version 9.5.4). ORFs longer than 50 codons were annotated to be predicted protein-coding genes, and all the ORFs were subsequently extracted into a FASTA file, and similarity searches including nucleotide (BLASTN) and protein (BLASTP) were performed on annotated ORFs as potential genes if they shared significant sequence similarity to known viral or cellular genes (BLAST E value ≤ e-4), or contained a putative conserved domain as predicted by BLASTP<sup>41</sup>. The final EKPV-NSW annotations were further examined with other poxvirus ortholog alignments to determine the correct methionine start site, correct stop codons, signs of truncation, and validity of overlaps. Transmembrane helices and single peptides were predicted using Geneious (version 10.2.3, Biomatters, New Zealand)<sup>42</sup>. The tandem direct repeats were detected using the Tandem Repeats Finders<sup>43</sup>.

**Genome sequence analysis and phylogenetic trees.** Nucleotide sequences of selected poxviruses that showed at least 15% similarity with newly sequenced EKPV-NSW genome were downloaded from GenBank: MOCV1 (*Molluscum contagiosum virus* subtype 1, MCU60315); MOCV2 (*Molluscum contagiosum virus* subtype 2, KY040274); SQPV (*Squirrel poxvirus*, HE601899); CRV (*Nile crocodilepox virus*, DQ356948); SPPV (*Seal parapoxvirus*, KY382358); BPSV (*Bovine papular stomatitis virus*, KM875470); PPRD (*Parapoxvirus red deer*, KM502564); ORFV (*Orf virus*, DQ184476); PCPV (*Pseudocowpox virus*, GQ329670); MYXV (*Myxoma virus*, KP723391); RFV (*Rabbit fibroma virus*, AF170722); LSDV (*Lumpy skin disease virus*, NC\_003027); TPV (*Turkeypox virus*, KP728110); FeP2 (*Pigeonpox virus*, KJ801920); PEPV (*Penguinpox virus*, KJ859677); FWPV (*Fowlpox virus*, AF198100); SWPV-1 (*Shearwaterpox virus-1*, KX857216); SWPV-2 (*Shearwaterpox virus-2*, KX857215); CNPV (*Canarypox virus*, AY318871); EKPV-NSW (*Eastern grey kangaroopox virus* strain NSW, MF661791); EKPV-SC (*Eastern grey kangaroopox virus* strain Suncine Coast, MF467281); WKPV-WA (*Western grey kangaroopox virus* strain Western Australia, MF467280).

For phylogenetic analysis, selected poxvirus complete genome sequences were aligned in Geneious (version 10.2.3) using MAFFT (version 7.309, algorithm; E-INS-I, scoring matrix; 100PAM/k = 2)<sup>44</sup>, and then manually edited in Geneious. To select the best construction method, a model testing was performed in CLC Genomics Workbench (version 9.5.4) using default parameters, and it favoured a general-time-reversible (GTR) substitution model for the maximum likelihood (ML) phylogeny. CLC Genomics Workbench was used to create a ML tree based on the Neighbor Joining (NJ) method and tested by bootstrapping with 1000 replicates.

Amino acid sequences of single protein (DNA polymerase, DNA topoisomerase I and variola B22R-like protein) were extracted from the complete genome sequences, and aligned using MAFFT multiple protein sequences alignment tool (version 7.309, algorithm; E-INS-I, scoring matrix; BLOSUM62) in Geneious (version 10.2.3). Tree topology with 1000 bootstrap re-samplings for each protein was inferred by NJ<sup>45</sup> using tools available in Geneious. Jukes-Cantor parameters was used for the measurement of protein and nucleotide distance.



To understand the source of genetic variation among the selected poxviruses and EKPV-NSW, we looked further for evidence of recombination using the RDP, GENECONV, Bootscan, MaxChi, Chimaera, Siscan, PhylPro, LARD and 3Seq methods contained in the RDP4 program<sup>27</sup>. Events that were detected by at least two of the afore-said methods with significant p-values were considered plausible recombinant events.

**Transmission electron microscopy.** Tissue sample collected from pox lesions were suspended in 1:10 in phosphate buffered saline (PBS) and chopped with a stainless-steel scalpel, followed by grinding with sterile round glass body homogeniser. Suspensions were clarified by centrifugation at 14,000 g for 5 min, followed by filtration of the supernatant through a 0.45 µm filter. The filtrate was then adsorbed onto a 400-mesh copper EM grid coated with a thin film of carbon for 5 minutes. Excess solution was removed using 3MM filter paper (Whatman) and the grid rinsed briefly with distilled water before negative staining with three 10 sec applications of 2% [w/v] uranyl acetate (Electron Microscopy Sciences, PA, USA) followed by removal of excess stain on filter paper after each application. Grids were air-dried for 20 minutes before examination under a JEOL JEM-2100 transmission electron microscope (TEM) operated at an accelerating voltage of 200 kV. High resolution digital images were recorded on a Gatan Orius SC200D 1 wide angle camera with Gatan Microscopy Suite and Digital Micrograph (Version 2.32.888.0) imaging software. Viruses were measured using ImageJ software.

**Data availability.** The complete sequence of the EKPV-NSW genome was deposited in GenBank under the accession number MF661791.

## References

- McFadden, G. Poxvirus tropism. *Nat Rev Microbiol* **3**, 201–213, <https://doi.org/10.1038/nrmicro1099> (2005).
- Baker, K. S. & Murcia, P. R. Poxviruses in Bats ... so What? *Viruses* **6**, 1564–1577, <https://doi.org/10.3390/v6041564> (2014).
- Di Giulio, D. B. & Eckburg, P. B. Human monkeypox: an emerging zoonosis. *The Lancet. Infectious diseases* **4**, 15–25 (2004).
- Vogelneust, L., Stewart, S. & Sangster, C. Poxvirus infection outbreak in common ringtails (*Pseudocheirus peregrinus*). *Aust Vet J* **90**, 143–145, <https://doi.org/10.1111/j.1751-0813.2011.00880.x> (2012).
- Afonso, C. L. *et al.* Genome of Crocodilepox Virus. *J Virol* **80**, 4978–4991, <https://doi.org/10.1128/JVI.80.10.4978-4991.2006> (2006).
- Fox, R. *et al.* Molluscum contagiosum in two donkeys. *Veterinary Record* **170**, 649–649, <https://doi.org/10.1136/vr.100721> (2012).
- Thompson, C. H., Yager, J. A. & Van Rensburg, I. B. Close relationship between equine and human molluscum contagiosum virus demonstrated by *in situ* hybridisation. *Res Vet Sci* **64**, 157–161 (1998).
- Bratke, K. A., McLysaght, A. & Rothenburg, S. A survey of host range genes in poxvirus genomes. *Infect Genet Evol* **14**, 406–425, <https://doi.org/10.1016/j.meegid.2012.12.002> (2013).
- Das, S. *et al.* Evolution of circoviruses in lorikeets lags behind its hosts. *Molecular Phylogenetics and Evolution* **100**, 281–291, <https://doi.org/10.1016/j.ympev.2016.04.024> (2016).
- Buck, C. B. *et al.* The Ancient Evolutionary History of Polyomaviruses. *PLoS Pathog* **12**, e1005574, <https://doi.org/10.1371/journal.ppat.1005574> (2016).
- Sarker, S. *et al.* Genomic characterization of two novel pathogenic avipoxviruses isolated from pacific shearwaters (*Ardenna* spp.). *BMC genomics* **18**, 298, <https://doi.org/10.1186/s12864-017-3680-z> (2017).
- Upton, C., Slack, S., Hunter, A. L., Ehlers, A. & Roper, R. L. Poxvirus Orthologous Clusters: toward Defining the Minimum Essential Poxvirus Genome. *J Virol* **77**, 7590–7600, <https://doi.org/10.1128/JVI.77.13.7590-7600.2003> (2003).
- Offerman, K., Carulei, O., van der Walt, A. P., Douglass, N. & Williamson, A.-L. The complete genome sequences of poxviruses isolated from a penguin and a pigeon in South Africa and comparison to other sequenced avipoxviruses. *BMC genomics* **15**, 1–17, <https://doi.org/10.1186/1471-2164-15-463> (2014).
- McKenzie, R. A., Fay, F. R. & Prior, H. C. Poxvirus infection of the skin of an eastern grey kangaroo. *Aust Vet J* **55**, 188–190 (1979).
- Bagnall, B. G. & Wilson, G. R. Molluscum contagiosum in a red kangaroo. *The Australasian journal of dermatology* **15**, 115–120 (1974).
- Bennett, M. *et al.* Complete genomic characterisation of two novel poxviruses (WKPV and EKPV) from western and eastern grey kangaroos. *Virus Res* **242**, 106–121, <https://doi.org/10.1016/j.virusres.2017.09.016> (2017).
- Ladds, P. 9–24 (CSIRO Publishing, Collingwood, Victoria, 2009).
- Papadimitriou, J. M. & Ashman, R. B. A Poxvirus in a Marsupial Papilloma. *Journal of General Virology* **16**, 87–89, <https://doi.org/10.1099/0022-1317-16-1-87> (1972).
- Rothwell, T. L., Keep, J. M., Xu, F. N. & Middleton, D. J. Poxvirus in marsupial skin lesions. *Aust Vet J* **61**, 409–410 (1984).
- Arundel, J., Beveridge, I. & Presidente, P. Parasites and Pathological Findings in Enclosed and Free-Ranging Populations of *Macropus rufus* (Demarest) (Marsupialia) at Menindee, New South Wales. *Wildlife Research* **6**, 361–379, <https://doi.org/10.1071/WR9790361> (1979).
- Li, Y., Meyer, H., Zhao, H. & Damon, I. K. GC Content-Based Pan-Pox Universal PCR Assays for Poxvirus Detection. *Journal of Clinical Microbiology* **48**, 268–276, <https://doi.org/10.1128/JCM.01697-09> (2010).
- Harrison, S. C. *et al.* Discovery of antivirals against smallpox. *Proc Natl Acad Sci USA* **101**, 11178–11192, <https://doi.org/10.1073/pnas.0403600101> (2004).
- Senkevich, T. G., Koonin, E. V., Bugert, J. J., Darai, G. & Moss, B. The genome of molluscum contagiosum virus: analysis and comparison with other poxviruses. *Virology* **233**, 19–42, <https://doi.org/10.1006/viro.1997.8607> (1997).
- Giessing, M. C. *et al.* Salmon Gill Poxvirus, the Deepest Representative of the Chordopoxvirinae. *J Virol* **89**, 9348–9367, <https://doi.org/10.1128/jvi.01174-15> (2015).
- Günther, T. *et al.* Recovery of the first full-length genome sequence of a parapoxvirus directly from a clinical sample. *Scientific Reports* **7**, 3734, <https://doi.org/10.1038/s41598-017-03997-y> (2017).
- Qin, L. & Evans, D. H. Genome Scale Patterns of Recombination between Coinfecting Vaccinia Viruses. *J Virol* **88**, 5277–5286, <https://doi.org/10.1128/jvi.00022-14> (2014).
- Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**, vev003, <https://doi.org/10.1093/ve/vev003> (2015).
- Zenger, K. R., Eldridge, M. D. B. & Cooper, D. W. Intraspecific variation, sex-biased dispersal and phylogeography of the eastern grey kangaroo (*Macropus giganteus*). *Heredity* **91**, 153–162 (2003).
- Vogelneust, L. & Portas, T. 133–166 (CSIRO Publishing, Collingwood, Victoria, 2008).
- Coghlan, B. A., Goldizen, A. W., Thomson, V. A. & Seddon, J. M. Phylogeography of Eastern Grey Kangaroos, *Macropus giganteus*, Suggests a Mesic Refugium in Eastern Australia. *PLoS one* **10**, e0128160, <https://doi.org/10.1371/journal.pone.0128160> (2015).
- Clafin, S. B. & Webb, C. E. Ross River Virus: Many Vectors and Unusual Hosts Make for an Unpredictable Pathogen. *PLOS Pathogens* **11**, e1005070, <https://doi.org/10.1371/journal.ppat.1005070> (2015).

32. Smithson, C., Purdy, A., Verster, A. J. & Upton, C. Prediction of Steps in the Evolution of Variola Virus Host Range. *PloS one* **9**, e91520, <https://doi.org/10.1371/journal.pone.0091520> (2014).
33. Odom, M. R., Hendrickson, R. C. & Lefkowitz, E. J. Poxvirus protein evolution: Family-wide assessment of possible horizontal gene transfer events. *Virus Res* **144**, 233–249, <https://doi.org/10.1016/j.virusres.2009.05.006> (2009).
34. Darby, A. C. *et al.* Novel host-related virulence factors are encoded by squirrelpox virus, the main causative agent of epidemic disease in red squirrels in the UK. *PloS one* **9**, e96439, <https://doi.org/10.1371/journal.pone.0096439> (2014).
35. Hatcher, E. L., Wang, C. & Lefkowitz, E. J. Genome Variability and Gene Content in Chordopoxviruses: Dependence on Microsatellites. *Viruses* **7**, 2126–2146, <https://doi.org/10.3390/v7042126> (2015).
36. Smithson, C. *et al.* The genomes of three North American orthopoxviruses. *Virus genes* **53**, 21–34, <https://doi.org/10.1007/s11262-016-1388-9> (2017).
37. Sarker, S., Das, S., Helbig, K., Peters, A. & Raidal, S. R. Genome sequence of an Australian strain of *Canid alphaherpesvirus 1*. *Aust Vet J*, (In press) (2017).
38. Jensen, R. H. *et al.* Target-dependent enrichment of virions determines the reduction of high-throughput sequencing in virus discovery. *PloS one* **10**, e0122636, <https://doi.org/10.1371/journal.pone.0122636> (2015).
39. Renfree, M. B. *et al.* Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development. *Genome Biology* **12**, R81–R81, <https://doi.org/10.1186/gb-2011-12-8-r81> (2011).
40. Tcherepanov, V., Ehlers, A. & Upton, C. Genome Annotation Transfer Utility (GATU): rapid annotation of viral genomes using a closely related reference genome. *BMC genomics* **7**, 1–10, <https://doi.org/10.1186/1471-2164-7-150> (2006).
41. Benson, D. A. *et al.* GenBank. *Nucleic Acids Research* **41**, D36–42, <https://doi.org/10.1093/nar/gks1195> (2013).
42. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567–580, <https://doi.org/10.1006/jmbi.2000.4315> (2001).
43. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27**, 573–580 (1999).
44. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**, 772–780, <https://doi.org/10.1093/molbev/mst010> (2013).
45. Saitou, N. & Nei, M. The Neighbor-Joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425 (1987).

## Acknowledgements

The authors are grateful to Vaheesan Rajabal and Steve Petrovski for technical assistance with Illumina library preparation and MiSeq sequencing. The authors also like to acknowledge Peter Lock to assist with TEM, and the Latrobe Institute for Molecular Science (LIMS) Bioimaging Facility for the use of Transmission Electron Microscope.

## Author Contributions

S.S. and S.R.R. conceived the experiment; S.A. and G.L. collected the sample; N.T., A.P. and S.R.R. performed the histological analysis; S.S. performed the library preparation, sequencing and TEM experiments; H.K. performed PCR; S.S. analysed the sequencing data; S.S. and S.R.R. wrote the manuscript; all authors participated in reviewing and editing the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-16775-7>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017