

Comparative Genomics of Amphibian-like Ranaviruses, Nucleocytoplasmic Large DNA Viruses of Poikilotherms



Stephen J. Price^{1,2}

¹Genetics, Evolution and Environment department, UCL Genetics Institute, London, UK. ²Institute of Zoology, Zoological Society of London (ZSL), London, UK.

Supplementary Issue: Evolutionary Genomics

ABSTRACT: Recent research on genome evolution of large DNA viruses has highlighted a number of incredibly dynamic processes that can facilitate rapid adaptation. The genomes of amphibian-like ranaviruses – double-stranded DNA viruses infecting amphibians, reptiles, and fish (family *Iridoviridae*) – were examined to assess variation in genome content and evolutionary processes. The viruses studied were closely related, but their genome content varied considerably, with 29 genes identified that were not present in all of the major clades. Twenty-one genes had evidence of recombination, while a virus isolated from a captive reptile appeared to be a mosaic of two divergent parents. Positive selection was also found to be acting on more than a quarter of *Ranavirus* genes and was found most frequently in the Spanish common midwife toad virus, which has had a severe impact on amphibian host communities. Efforts to resolve the root of this group by inclusion of an outgroup were inconclusive, but a set of core genes were identified, which recovered a well-supported *species* tree.

KEYWORDS: comparative genomics, pathogen, aquaculture, phylogenetics

SUPPLEMENT: Evolutionary Genomics

CITATION: Price. Comparative Genomics of Amphibian-like Ranaviruses, Nucleocytoplasmic Large DNA Viruses of Poikilotherms. *Evolutionary Bioinformatics* 2015:11(S2) 71–82 doi: 10.4137/EBO.S33490.

TYPE: Original Research

RECEIVED: July 27, 2016. **RESUBMITTED:** September 12, 2016. **ACCEPTED FOR PUBLICATION:** September 15, 2016.

ACADEMIC EDITOR: Liuyang Wang, Associate Editor

PEER REVIEW: Four peer reviewers contributed to the peer review report. Reviewers' reports totalled 1046 words, excluding any confidential comments to the academic editor.

FUNDING: This work was funded by NERC grants NE/M000338/1, NE/M000591/1, and NE/M00080X/1 and ERC grant 260801-BIG-IDEA. The author confirms that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Author discloses no potential conflicts of interest.

CORRESPONDENCE: s.price@ucl.ac.uk

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY 4.0 License.

Paper subject to independent expert blind peer review. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Published by Libertas Academica. Learn more about this journal.

Introduction

Large DNA viruses have traditionally been viewed as somewhat evolutionarily inactive, exhibiting low substitution rates and cospeciation alongside their hosts.¹ This may have sprung from an overemphasis on the effect of the molecular mechanisms of replication (eg, the specific enzymes utilized and the resultant error rates, which ultimately determine mutation rates) and a lack of consideration of other biological factors that determine substitution rates such as generation time, life cycle, and selection.² Indeed, although phylogenetic concordance between hosts and viruses has been observed frequently, in some cases it has been shown to stem from preferential host switching rather than cospeciation.³ Several lines of evidence from recent studies show that large DNA virus genomes can be highly dynamic. Genome evolution via lateral gene transfer occurs frequently with examples of the acquisition of genetic material from eukaryotes, bacteria, and bacteriophages.^{4–8} Double-stranded DNA viruses exhibit considerable variation in substitution rates.² Viruses such as the variola virus (a poxvirus causing smallpox) appear to have substitution rates on par with RNA viruses, and there is tentative evidence for high rates in other virus groups too.^{9,10} Recombination between viruses is also now known to be extremely widespread in some large DNA viruses.^{11,12}

Nucleocytoplasmic large DNA viruses form a monophyletic group of seven virus families (*Asfarviridae*, *Ascoviridae*, *Iridoviridae*, *Marseilleviridae*, *Mimiviridae*, *Phycodnaviridae*, and *Poxviridae*).¹³ Nucleocytoplasmic large DNA viruses share a common ancestor and a core set of genes, and their origins are thought to be concomitant with eukaryogenesis.^{4,14} They are differentiated from other large DNA viruses by their capacity to establish replication *factories* in the cytoplasm of host cells.¹⁵

The family *Iridoviridae* consists of five genera that infect vertebrates and invertebrates.¹⁶ They have received increased attention recently due to the impact on amphibian populations by members of the genus *Ranavirus*.^{17,18} Ranaviruses have a very broad host range – infecting amphibians, reptiles, and fish – and are thought to have undertaken a number of host jumps.¹⁹ They can have severe impacts on individual host species or entire amphibian communities.^{20,21} The amphibian-like ranaviruses (ALRVs)¹⁹ are considered emerging pathogens, and humans are thought to have contributed to their geographic distribution through the movement of hosts, which prompted their inclusion on the World Organization for Animal Health's list of notifiable pathogens.^{22,23} Humans may also have affected *Ranavirus* evolution; virus strains isolated from aquaculture



have increased virulence compared to wild relatives,^{24,25} while the cohousing of diverse amphibian species through trade may present enhanced opportunities for recombination among divergent viruses.^{26–28}

ALRVs with sequenced whole genomes are geographically widespread and have been isolated over a 50-year period. Genomes range from 100 to 130 kb in size and contain approximately 100 genes with low functional annotation rates.^{18,29} The ALRVs form a distant sister clade to grouper iridovirus-like ranaviruses (GIV-like; pathogens of marine fish) and are frequently divided into three groups: *Ambystoma tigrinum virus*-like (ATV-like), common midwife toad virus-like (CMTV-like), and *Frog virus 3*-like (FV3-like).³⁰ This division of the ALRVs is suggestive of underlying systematic relationships but has not yet been explicitly defined. There has been a gradual increase in the number of published *Ranavirus* whole genomes, but genomes have generally been sequenced one at a time applying a variety of approaches. There have been few comparative genomic analyses within this group with even fewer attempts to examine the group's evolutionary dynamics. However, a study examining *Ranavirus* genes for evidence of positive selection associated with host-switching found recombination in a high proportion of studied genes as well as positive selection in a number of newly acquired genes that might be associated with host switching events.³¹ Another recent study focused on closely related, monophyletic isolates of ATV in North America and found little variation in genome content but did find evidence of positive selection and recombination among the regional isolates.³²

Methods

Data. Sixteen published ALRV genomes were downloaded from NCBI nucleotide database (full GenBank [gbk] format and the nucleotide sequence only format [FASTA]) (Table 1). These viruses were isolated in the USA, Europe, China, and Australia over a 50-year period from 1965 to 2013 (Table 1). The same files were also downloaded for the GIV-like ranaviruses for use as outgroups and in some comparative analyses (Table 1).

Comparative genomics. Genome contiguity was assessed by aligning complete virus genomes with NUCmer v3.1, which is a fast aligner for multiple, closely related genomes, using a minimum cluster length of 100 nucleotides. The output was visualized with Mummerplot.³³ To identify orthologous genes, ALRV genomes were first aligned using the Mugsy whole genome aligner (version 1, release 2.3) with default settings³⁴ and putative orthologs were clustered using Mugsy-Annotator 0.5³⁵ (coverage cutoff = 0.8, identity = 0.8, and query coverage = 0.8) with GenBank files containing original annotations for each virus supplied as input. Mugsy-Annotator is a command line program, which uses a whole genome multiple alignment and the genomes' associated features to identify annotation anomalies such as fragmentation, inconsistent starts, and missing features. Orthologous

genes are grouped in clusters, and the algorithm requires loci to encode functional proteins (ie, no premature stop codons) for cluster membership. The sequence indices generated by Mugsy-Annotator were then used to extract ortholog sequences from whole genomes using the EMBOSS program Extractseq.³⁶ Sequences were then aligned with reference to translated amino acid sequences using TranslatorX,³⁷ which was run locally from the command line (with MAFFT as the aligner, amino acid sequences translated from the nucleotide sequences, and default settings used for other parameters).

The amount of variation in amino acid and nucleotide alignments was summarized with AMAS,³⁸ a python command line utility, which allows the computation of basic statistics from alignments. Alignments were checked for the presence or absence of sequences from each of the 16 isolates to generate a matrix (Supplementary Table 1; 16 viruses and 130 orthologs), which was used to summarize the number of genomes containing the ortholog as well as a summary of its occurrence across the virus phylogeny. In addition, the occurrence of each ortholog in the genomes of the GIV-like ranaviruses was assessed by a BLASTP search (Blast 2.2.29+, *e*-value cutoff at 1×10^{-10} , four threads, and otherwise default BLASTP settings) against the published coding regions of GIV-like viruses. Orthologs found in all 16 ALRVs as well as both GIV-like viruses were classed as *Ranavirus-core*. Orthologs were then processed through a published pipeline to find functional annotations.³⁹

After removing stop codons, orthologs were assessed for recombination with GARD,^{40,41} which is implemented in HyPhy and uses a genetic algorithm to find recombination breakpoints in multiple sequence alignments. GARD was run in parallel in an MPI environment with five threads using the HKY85 substitution model. Alignments were split at breakpoints identified by GARD for further analysis. Prunier⁴² identifies likely occurrences of lateral gene transfer by analyzing the lack of congruence between the topologies of a reference *species* tree and a given gene tree. All orthologous genes were screened for such conflict. A phylogeny generated from an ALRV-core gene set after removing genes with evidence of recombination (tree construction method described below) was supplied as the *species* tree. Prunier called RAXML to generate gene trees from nucleotide alignments for each orthologous gene using the GTR model of nucleotide substitution and 100 bootstraps using the rapid bootstrapping algorithm.

Each parent alignment and any split alignments (generated with reference to the GARD output using a custom script) were converted to interleaved phylip format using Prank v.140603.⁴³ These alignments were used to construct phylogenetic trees by maximum likelihood using RAXML⁴⁴, applying the GTR model of nucleotide substitution with rate variation modeled by a gamma distribution with four categories. Ten maximum-likelihood trees were generated on distinct starting trees in RAXML's rapid hill-climbing mode,

**Table 1.** Isolation history and source of data for *Ranavirus* genomes used in comparative genomics analyses.

VIRUS ID	ABBREVIATION	NCBI ACCESSION NO.	TYPE	ORIGIN	HOST	YEAR COLLECTED [†]	LENGTH (BASE PAIRS)	PUBLICATION
Frog virus 3	FV3	AY548484	ALRV	USA	<i>Rana pipiens</i> (Northern leopard frog)	1965 (approx.)	105903	Tan et al., 2004 ⁶³
Frog virus 3 isolate SSME	FV3_SSME	KJ175144	ALRV	USA	Uncertain	1998	105070	Morrison et al., 2014 ⁶⁴
Rana grylio iridovirus	RGV	JQ654586	ALRV	China	<i>Rana grylio</i> (pig frog)	1995	105791	Lei et al., 2012 ⁵⁹
Soft-shelled turtle iridovirus	STIV	EU627010	ALRV	China	<i>Trionyx sinensis</i> (soft-shelled turtle)	1998 (approx.)	105890	Huang et al., 2009 ⁶⁵
Tiger frog virus	TFV	AF389451	ALRV	China	<i>Rana tigrina rugulosa</i> (tiger frog)	1998 (approx.)	105057	He et al., 2002 ⁶³
German gecko ranavirus isolate 2000/99	GGRV	KP266742	ALRV	Germany	<i>Uroplatus fimbriatus</i> (giant leaf-tailed gecko)	1999	103681	Stöhr et al., 2015 ³⁰
Common midwife toad ranavirus isolate Mesotriton alpestris/2008/E	CMTV_SP	JQ231222	ALRV	Spain	<i>Mesotriton alpestris</i> (alpine newt)	2008	106878	Mavian et al., 2012 ⁶⁶
Common midwife toad ranavirus isolate Pelophylax kl. esculentus/2013/NL	CMTV_NL	KP056312	ALRV	Netherlands	<i>Pelophylax kl. esculentus</i> (edible frog)	2013	107772	van Buerden et al., 2014 ⁶⁷
Testudo hermanni ranavirus isolate CH8/96	THR896	KP266741	ALRV	Switzerland	<i>Testudo hermanni</i> (Hermann's tortoise)	1996	105811	Stöhr et al., 2015 ³⁰
Andrias davidianus ranavirus isolate 2010SX	ADRV2010	KF033124	ALRV	China	<i>Andrias davidianus</i> (Chinese giant salamander)	2010	106719	Wang et al., 2014 ⁶⁸
Andrias davidianus ranavirus isolate 1201	ADRV1201	KC865735	ALRV	China	<i>Andrias davidianus</i> (Chinese giant salamander)	2012	106734	Chen et al., 2013 ⁶⁹
Chinese giant salamander iridovirus isolate CGSIV-HN1104	CGSIV	KF512820	ALRV	China	<i>Andrias davidianus</i> (Chinese giant salamander)	2013	105375	Li et al., 2014 ⁷⁰
Tortoise ranavirus isolate 1 (882/96)	ToRV1	KP266743	ALRV	Germany	<i>Testudo kleinmanni</i> (Kleinmann's tortoise)	1996	103876	Stöhr et al., 2015 ³⁰
Ambystoma tigrinum virus	ATV	NC_005832	ALRV	USA	<i>Ambystoma tigrinum</i> (Tiger salamander)	1996	106332	Jancovich et al., 2003 ⁷¹
Epizootic haematopoietic necrosis virus	EHNV	FJ433873	ALRV	Australia	<i>Perca fluviatilis</i> (redfin perch)	1986 (approx.)	127011	Jancovich et al., 2010 ¹⁹
European sheatfish virus	ESV	JQ724856	ALRV	Uncertain	Uncertain	Uncertain	127732	Mavian et al., 2012 ⁷²
Grouper iridovirus	GIV	AY666015	GIV	Uncertain	<i>Epinephelus awoara</i> (yellow grouper)	Uncertain	139793	Tsai et al., 2005 ⁷³
Singapore grouper iridovirus	SGIV	NC_006549	GIV	Singapore	<i>Epinephelus chlorostigma</i> (brown-spotted grouper)	1998 (approx.)	140131	Song et al., 2004 ⁷⁴

Notes: Two virus types are included; most analyses focused on amphibian-like ranaviruses (ALRV). Grouper Iridovirus-like (GIV) viruses were included as an outgroup in some phylogenetic analyses. [†]For some isolates, the year of collection is provided in the GenBank record or accompanying publication. For others, the year of collection is estimated from other information in the accompanying publication, and this is indicated in the table by appending (approximately) to the year.



and 100 bootstrap replicates were calculated and annotated on the best maximum-likelihood tree.

The adaptive version of Branch-Site REL⁴⁵ (implemented in HyPhy) was used to detect episodic selection. Branch-Site REL uses random effects models to allow variation in substitution rates between branches across a phylogeny in order to overcome shortcomings in previous branch-site models of selective pressure that can lead to erroneous inferences.⁴⁵ Each alignment was analyzed for positive selection using the random effects branch-site model, applying the universal genetic code and allowing alpha to vary from branch to branch and omega to vary among branch-site combinations. The RAxML trees were provided as input, and all branches were tested before applying the default Holm–Bonferroni correction for multiple comparisons. Positive selection was also analyzed with BUSTED, which uses the Branch-Site REL framework to test for evidence of episodic selection by estimating the proportion of selected codons across all branches of the phylogenetic tree instead of the branch-specific method of Branch-Site REL.⁴⁶ Again, the universal genetic code was used.

Phylogenetics. To address the evolutionary relationships among ALRVs, a phylogenetic tree including GIV as an outgroup was constructed. The results of *Ranavirus* genome content comparisons (described above) were used to select a set of orthologs conserved among ALRVs and GIV-like ranaviruses. Orthologs with evidence of recombination were removed from this core set. Nucleotide sequences were aligned with reference to amino acid sequences with TranslatorX.³⁷ Nucleotide and amino acid alignments were each concatenated separately using phyutility,⁴⁷ and all columns containing gaps were removed using trimAl (*nogaps* option).⁴⁸ The final multiple sequence alignments were used to construct *species* trees in Mr. Bayes v3.2.2⁴⁹ and RAxML,⁴⁴ using each ortholog as a partition and applying GTR + gamma (four rate categories) across all partitions. Markov chain Monte Carlo (MCMC) chains (two runs and four chains each) were run for 250,000 generations in Mr. Bayes. Twenty starting trees were generated in RAxML and support values were generated from 100 bootstrap replicates.

In order to obtain the maximum amount of data possible for the reconstruction of phylogenetic relationships among ALRVs, nucleotide sequences of the ALRV-core orthologs (only those present in all sixteen ALRVs) without evidence of recombination were used in phylogeny reconstruction. Alignments were concatenated and processed following the same pipeline described above (using 500,000 generations for MCMC chains). This tree was used as the *species* tree in Prunier analyses.

Results

Genome content. A total of 130 orthologous clusters were found in ALRV genomes. The attempted functional annotation of these orthologs conducted here resulted in only 40 annotations (Supplementary Table 2). The majority of genes were shorter than 1000 nucleotide bases in length (Supplementary Fig. 1). The 26 genes identified previously²⁹

as core genes that are conserved across the whole family (*Iridoviridae*) were among the orthologous clusters and were each found in all sixteen ALRV genomes. An additional 35 orthologs (a total of 61) were found in all viruses in the genus (*Ranavirus*-core genes; ALRVs and GIV-like viruses). A further four orthologs (total = 65) were present in all 16 ALRVs, but a total of 101 orthologs occurred in at least one member of all three major clades of the ALRVs.

Twenty-nine orthologs were therefore considered clade-specific within the ALRV. Each major clade had four clade-specific orthologs. There were also 12 orthologs in common between the FV3-like and CMTV-like clades that were not found in ATV-likes, and there was one ortholog that was lost in CMTV-likes (ie, present in ATV- and FV3-likes) compared with four orthologs lost in FV3-likes. The rate of functional annotation was very low among these accessory genes: the only orthologs with some available indication of function were a putative integrase-like protein, which was absent in the ATV-likes, a putative capsid maturation protease that occurred only in ATV-likes (and only in the fish pathogens of this group – European sheatfish virus [ESV] and epizootic haematopoietic necrosis virus [EHNV]), and a novel US-22 family protein in the CMTV-likes.

In addition, 38 singletons – defined as novel open reading frames present in only one genome – were found in ALRV genomes. Thirty of these singletons occurred in the two fish pathogens of the ATV-like clade (25 in ESV and five in EHNV), which have larger genomes than the other ALRVs (Table 1). Singletons were also found in ADRV2010 (2), CGSIV (2), CMTV_NL (1), STIV (1), and TFV (2).

Removing genes with signs of recombination resulted in a set of 51 orthologous genes that were found in all 16 ALRVs (ALRV-core; Supplementary Table 2). The final concatenated alignment contained 16 taxa and 44,940 nucleotide bases. Reconstruction of phylogenetic relationships of ALRVs using this 51-gene set resulted in a well-supported phylogeny (Fig. 1). When rooted on ATV-like viruses, the tree returned ATV-like, CMTV-like, and FV3-like viruses as three monophyletic clades with tortoise ranavirus 1 (ToRV1) intermediate to the ATV-likes and CMTV-like/FV3-like clade.

Systematics and genome arrangement. After removing genes with evidence of recombination, a set of 39 orthologous genes present in the GIV-like viruses and all 16 ALRVs (Supplementary Table 2) were retained for use in phylogeny reconstruction. The concatenated nucleotide alignment contained 17 taxa and 39,498 nucleotide bases. The concatenated amino acid alignment contained 13,166 residues. In spite of the large amount of data, phylogenetic evidence for the root of the ALRVs was inconclusive – there was support for ATV on its own or as part of a clade of ATV-like viruses as the root with the topology sensitive to the method followed for tree construction (Supplementary Fig. 2). It was therefore not possible to determine whether ATV-like viruses should be considered monophyletic or paraphyletic.

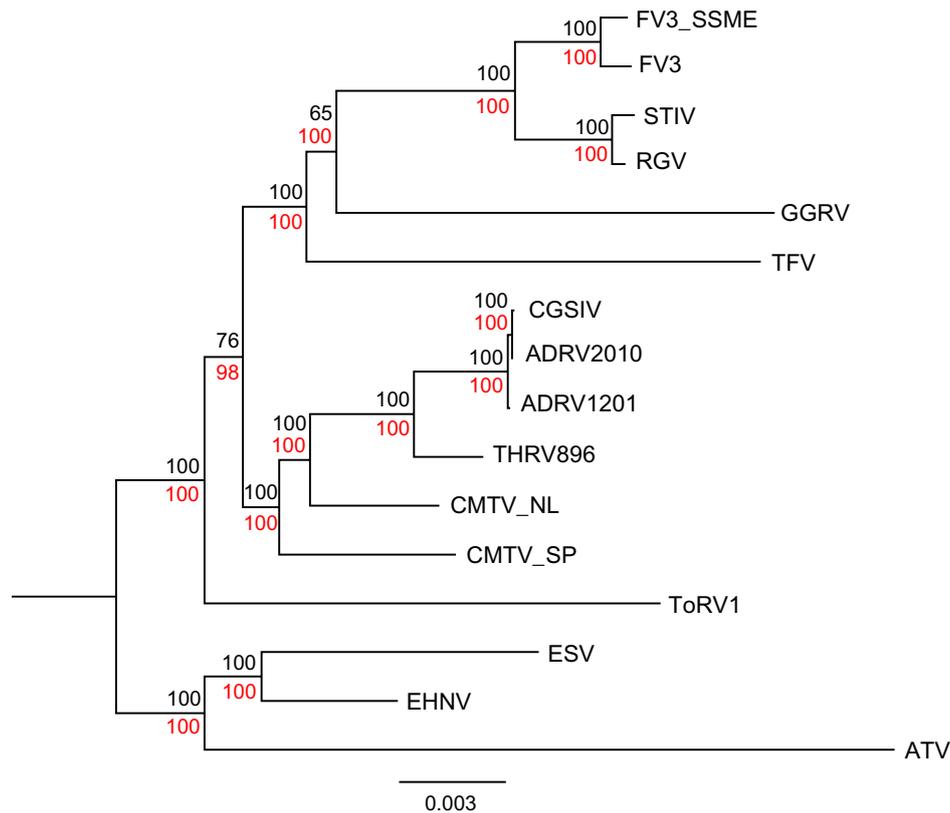


Figure 1. Phylogenetic relationships of amphibian-like ranaviruses (ALRVs) based on 51 ALRV-core genes having removed genes with evidence of recombination. Maximum likelihood and Bayesian trees were constructed from a concatenated nucleotide alignment 44,940 nucleotide bases in length, which was partitioned by loci.

Notes: Support values were generated from 100 maximum likelihood bootstrap replicates (black) and Bayesian posterior probabilities (red; as percentage) and are annotated on the Bayesian tree. Scale is in nucleotide substitutions per site.

An examination of ALRV genome arrangement, using whole genome alignments, did not resolve the ALRV root either. Genome contiguity is very consistent among ALRVs, but there have been two major rearrangements during ALRV evolution (Fig. 2). These rearrangements support the division of the ALRVs into three groups but do not resolve whether there is a systematic basis for all three (ie, groupings which correspond to monophyletic clades). CMTV-like viruses are distinguished from ATV-like viruses by an approximately 30 kb inversion. This inversion bisects an annotated open reading frame at one end in ATV. A steroid oxidoreductase-like protein occurs as a full-length gene in CMTV-like viruses but occurs as a truncated version in ATV (ortholog identifier = C367o). However, this gene is present at full length in EHN and in the same orientation as CMTV-like but absent all together in ESV. Therefore, two alternative hypotheses remain concerning the root and ancestral genome arrangement of ALRVs (Fig. 2).

A second major inversion of approximately 88 kb – containing the older 30 kb rearrangement described earlier – occurred on the branch to the ancestor of all FV3-like viruses (Fig. 2). Again, this inversion bisects an open reading frame at one end when the original annotation of the CMTV_SP virus is used as a reference. There were a

number of additional smaller rearrangements, the largest an approximately 13 kb long inversion in CGSIV relative to the very closely related ADRV isolates.

Recombination and selection. Twenty-one orthologous genes were found with evidence of recombination ($P < 0.05$) using GARD (Table 2). Genes affected by recombination include neurofilament triplet H1-like protein genes, a myristoylated membrane protein gene, and a D5 family NTPase/ATPase, but two-thirds lacked functional annotation (Supplementary Table 2). Prunier revealed 36 orthologous genes where there was significant conflict between the gene tree and the *species* tree provided. Phylogenies constructed from individual ortholog alignments were used to assess the phylogenetic position of ToRV1 across its genome. Moving along the genome, small, spatially clustered groups of genes supported equivalent phylogenetic positions for ToRV1 – either FV3-like or a sister group to the FV3-like/CMTV-like clade – but the supported position alternated (Fig. 3).

Twenty-eight genes were found with evidence of episodic selection using the branch-site method (Table 3). Again, functional annotation was lacking, but eight were annotated and three of the 29 were members of the *Iridoviridae*-core group. Most of the selected genes were present in all clades, but this may merely reflect lower power to detect these effects with

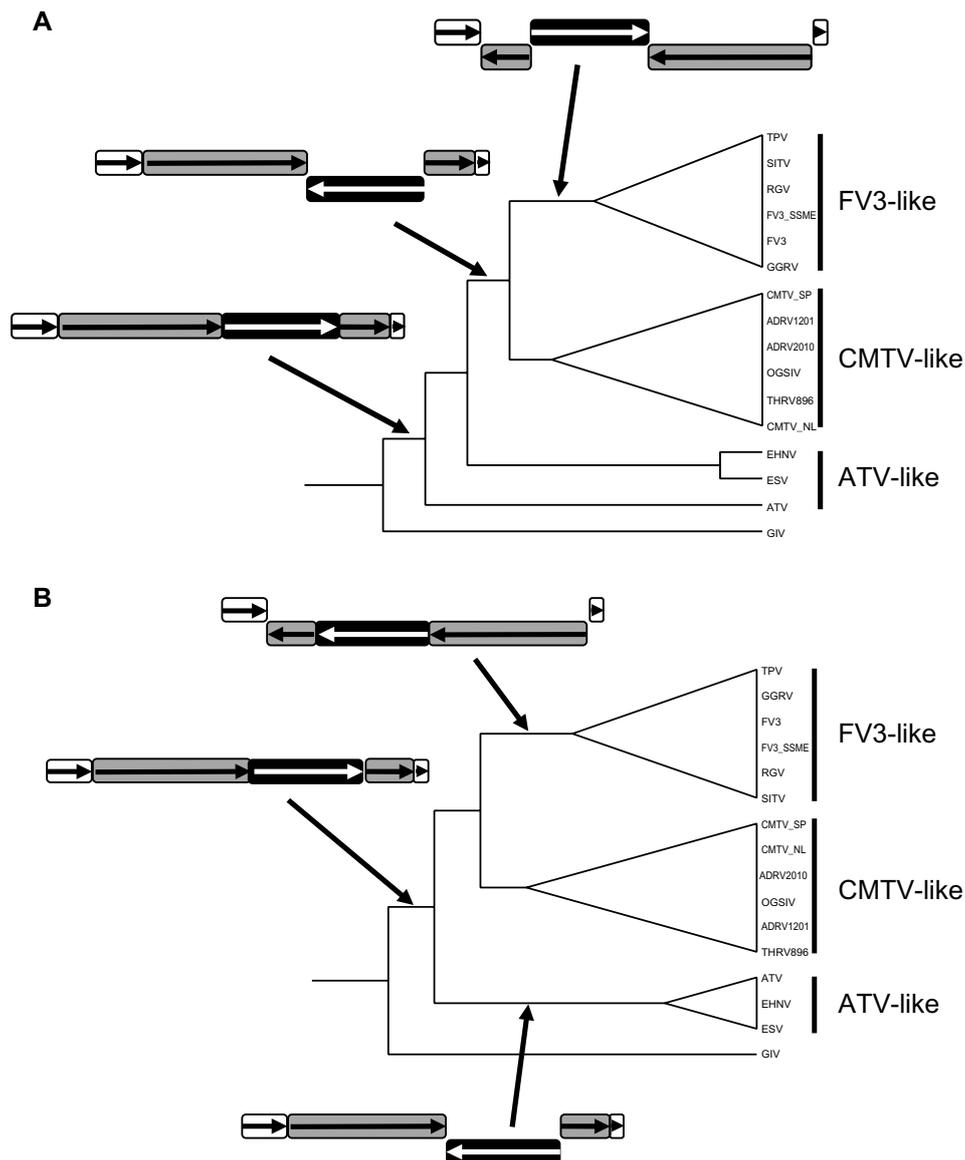


Figure 2. Alternative hypotheses for the evolutionary history of the ALRVs accounting for phylogenetic signal and genome arrangement. Genome segments involved in major rearrangements are color filled: black for the 30 kb rearrangement and the combined black and gray segments for the 88 kb rearrangement (see text). Orientation of genome segments is indicated by arrows and position relative to horizontal axis. **(A)** Ancestral ATV-like arrangement and paraphyly of ATV-likes. **(B)** Ancestral CMTV-like arrangement and monophyly of ATV-likes.

fewer taxa. Two genes only found in the FV3/CMTV clades (functional annotation lacking for both) were also shown to have undergone positive selection.

Branch-Site REL identifies the branches in a phylogeny where selection has occurred. Among the 28 genes with evidence of selection, the terminal branches leading to ATV and ESV (five genes), TPV (six genes), and CMTV_SP (seven genes) featured most frequently. Other terminal branches where positive selection was detected were EHN (three genes); ToRV1, STIV, and GGRV (all two); CGSIV and RGV (one). In terms of selection on ancestral branches, the ATV-likes and the viruses isolated from Chinese giant salamanders were found within the affected clades most frequently. Annotated orthologs that have experienced selection included a nuclear

calmodulin-binding protein gene, neurofilament triplet H1-like protein genes, myristoylated membrane protein genes, and an ATPase-dependent protease.

The alignment-wide approach using BUSTED identified many of the same genes as Branch-Site REL (Table 3): 24 of the 28 genes identified by Branch-Site REL were also significant for positive selection using BUSTED, while there were six additional genes identified by BUSTED as candidates for having undergone episodic positive selection (including DNA polymerase).

Discussion

The traditional view of large DNA viruses was that they were slow evolving and their genomes were lacking in

Table 2. Orthologous genes with evidence of recombination.

ORTHOLOG ID	GENE ID	UNIPROT REF.	DELTA AICc	MAX. NO. OF BREAKPOINTS	MEAN SPLITS	SH TEST P-VALUE
C3	NA	Q6GZT5	6.45	3	0.600	0.002
C5	Putative tyrosine protein kinase	Q6GZU9	713.53	3	0.545	0.001
C6	Putative D5 family NTPase/ATPase	Q6GZV3	101.9	4	0.208	0.003
C9	Neurofilament triplet H1-like protein	Q6GZU4	1375.63	10	0.038	0.021
C10	Putative ATPase-dependent protease	Q6GZP6	68.54	4	0.278	0.020
C11	NA	Q6GZN8	115.2	3	0.667	0.001
C14	NA	Q6GZS6	723.69	5	0.070	0.001
C17	Putative phosphotransferase	Q6GZR8	47.76	3	0.444	0.008
C18	NA	Q6GZN6	25.49	3	0.375	0.001
C18o	NA	C3RWM1	3.76	3	0.250	0.017
C20	NA	Q6GZQ1	139.9	3	0.286	0.000
C21	NA	Q6GZX2	18.48	3	0.400	0.002
C31	NA	Q6GZV2	3.78	3	0.000	0.005
C35	Putative myristoylated membrane protein	Q6GZX3	253.36	3	0.375	0.014
C68	Neurofilament triplet H1-like protein	Q6GZT0	123.24	4	0.056	0.041
C69	NA	Q6GZX0	14.84	3	0.833	0.034
C71o	NA	Q6GZV6	41.96	3	0.600	0.003
C200	NA	Q6GZR4	22.24	3	0.000	0.000
C116120	NA	Q6GZN9	39.55	3	0.400	0.000
C78117145	NA	Q6GZU0	431.92	4	0.244	0.000
C6210013531o	NA	D3TTS0	64.47	3	0.400	0.000

Notes: A summary of output from GARD analysis to detect recombination breakpoints, including the number of breakpoints, and the *P*-value from the Shimodaira and Hasegawa test.

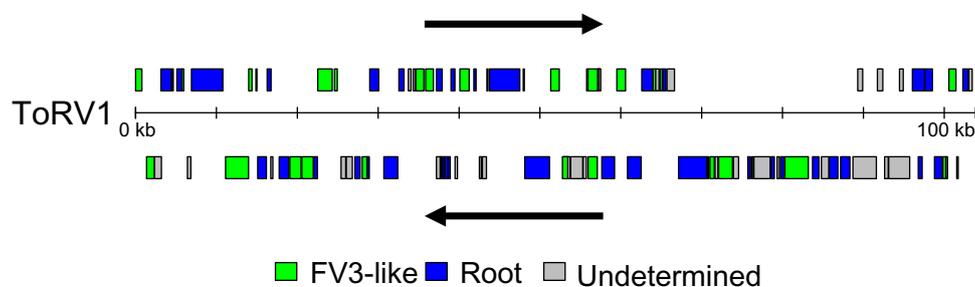


Figure 3. ToRV1 as a mosaic virus. Colored blocks represent putative genes (the position above or below the line along with the respective arrows indicates orientation). Colors represent ToRV1's position in the gene tree (within the FV3-like clade, the sister taxon to the FV3-like/CMTV-like clade [root], or unresolved). Tick marks are positioned at 10 kbp intervals.

dynamic changes. This study suggests how dynamic genome change in ranaviruses is linked to important aspects of their biology and adds to work on other virus families that challenge this traditional view.^{9,11} The ALRVs of the family *Iridoviridae* are important, emerging pathogens of amphibians, fish, and reptiles. Their genomes have been subjected to widespread recombination and positive selection and have a considerable accessory genome. Although the viruses examined here are closely related, they exhibit a number of processes leading to dynamic change; their genome content varies systematically and individually, there is frequent

recombination, and many genes have been the target of positive selection.

ALRV genomes were examined for the presence or absence of functional coding sequences to establish core gene sets as well as their accessory genomes. Some of the variation in genome content maps to the virus phylogeny with groups of genes that are lineage specific. ALRVs are highly variable in their host range and virulence,^{20,21,50,51} and variation in genome content is considered an important factor in explaining observed phenotypic differences between virus types in such key traits.⁵² Unfortunately, this study found

Table 3. Amphibian-like ranavirus orthologous genes under positive selection.

ORTHOLOG	GENE ID	UNIPROT	SPLIT	BRANCH-SITE REL TAXON	P-VALUE		BUSTED	
					SELECTED?	P-VALUE	SELECTED?	P-VALUE
C4	DNA polymerase-like protein	Q6GZR5	1	None	no	0.009	yes	0.009
C9	Neurofilament triplet H1-like protein	Q6GZU4	1	(THR896, CMTV_NL)	yes	0.000	yes	0.000
				TFV	no	0.001	no	0.001
				(ATV, (EHN, ESV))		0.001		0.001
				(RGV, (FV3, (ADRV1201, (ADRV2010, (STIV, (CGSIV, FV3_SSME))))))		0.043		0.043
				None	no	0.000	yes	0.000
				(ToRV1, ((ESV, (FV3_SSME, FV3)), GGRV))	yes	0.017	yes	0.000
				CMTV_SP		0.024		0.024
				(ESV, (ATV, (ToRV1, GGRV)))	yes	0.002	yes	0.000
				(ATV, (ToRV1, GGRV)))		0.045		0.045
				None	no	0.000	yes	0.000
				EHN	yes	0.000	yes	0.000
				None	no	0.000	yes	0.000
C10	Putative ATPase-dependent protease	Q6GZP6	1	ESV	yes	0.033	yes	0.005
				CGSIV	yes	0.043	no	0.105
				ATV	yes	0.000	yes	0.000
				None	no	0.012	yes	0.012
C11	NA	Q6GZN8	2	None	no	0.001	yes	0.000
				(FV3_SSME, FV3)	yes	0.007	yes	0.007
C14	NA	Q6GZS6	1	None	no	0.000	yes	0.000
				CMTV_SP	yes	0.043	yes	0.043
				(CMTV_NL, (THR896, ((STIV, (FV3_SSME, FV3)), RGV))),		0.043	yes	0.043
				(CMTV_SP, ((CGSIV, (ADRV2010, ADRV1201)), GGRV)))		0.004	yes	0.004
				TFV	yes	0.045	yes	0.045
				ESV	yes	0.000	yes	0.000
C15	Myristoylated membrane protein	Q6GZS3	1	(ATV, (EHN, ESV)), (CMTV_SP, ((ToRV1, (THR896, CMTV_NL)), (ADRV1201, (CGSIV, ADRV2010)), (FV3_SSME, FV3)), (RGV, STIV))))		0.031		0.031
				ATV	yes	0.000	yes	0.000
C18	NA	Q6GZN6	2	((ADRV1201, ADRV2010), (THR896, CMTV_NL))	yes	0.000	yes	0.000
				TFV	yes	0.003	yes	0.003
C20	NA	Q6GZQ1	1	ATV	yes	0.007	no	0.389
				RGV	yes	0.000	yes	0.000
				(ATV, (GGRV, (CMTV_SP, ((ADRV1201, CGSIV), ADRV2010), (EHN, (THR896, CMTV_NL))))))		0.005		0.005
				EHN		0.024		0.024
				GGRV		0.033		0.033
				None	no	0.038	yes	0.038
C31	NA	Q6GZV2	1	None	no	0.038	yes	0.038

C35	Putative myristoylated membrane protein	Q6GZK3	2	((CGSIV, ADRV2010), THRV896), CMTV_NL) ESV ToRV1	0.000 0.009 0.032	yes	yes	0.000	0.000
C38	NA	Q6GZR6	1	(CMTV_SP, ((THRV896, ((CGSIV, ADRV2010), ADRV1201)), CMTV_NL	0.024	yes	yes	0.024	0.008
C43	NA	Q6GZT6	1	EHN (THRV896, (ADRV2010, ADRV1201)) (ESV, EHN)	0.000 0.000 0.047	yes	yes	0.000	0.000
C47	NA	ADA0A0VJ51	2	STV	0.000	yes	no	0.000	0.436
C68	Neurofilament triplet H1-like protein	Q6GZT0	3	TFV CMTV_SP (EHN, (ESV, ATV))	0.002 0.025 0.030	yes	yes	0.002	0.000
C69	NA	Q6GZX0	2	None		no	yes		0.000
C90	NA	Q6GZU8	1	(ToRV1, (ATV, (EHN, ESV)))	0.007	yes	yes	0.007	0.044
C101	Putative myeloid cell leukemia protein	Q6GZM8	1	ESV	0.000	yes	yes	0.000	0.002
C106	NA	Q2WEW6	1	(ATV, (ESV, EHN))	0.024	yes	no	0.024	0.731
C118	NA	A0A0A0VCE2	1	CMTV_SP	0.000	yes	yes	0.000	0.007
C127	NA	Q6GZP8	1	GRV	0.005	yes	yes	0.005	0.010
C129	NA	A0A0D3R3F0	1	ATV	0.013	yes	yes	0.013	0.005
C141	NA	Q6GZN3	1	TFV	0.010	yes	yes	0.010	0.001
C176	Putative nuclear calmodulin binding protein	Q6GZS2	1	CMTV_SP	0.026	yes	yes	0.026	0.009
C200	NA	Q6GZR4	2	None		no	yes		0.001
C204	NA	Q6GZU6	1	CMTV_SP	0.022	yes	no	0.022	1.000
C25160	NA	Q6GZT4	3	TFV ESV	0.000 0.001	yes	yes	0.000	0.000
C116120	NA	Q6GZN9	2	((ATV, EHN), ((ADRV1201, ADRV2010), CGSIV)) (ATV, EHN)	0.000 0.016	yes	yes	0.000	0.000
C53175183	eIF-2 alpha-like protein	P68101	1	(CMTV_SP, RGV)	0.000	yes	yes	0.000	0.000
C78117145	NA	Q6GZU0	2	None		no	yes		0.000
C20	NA	Q67475	3	None		no	yes		0.000
C330	NA	Q6GZQ6	1	ToRV1	0.033	yes	yes	0.033	0.025
C568170	NA	Q6GZU1	1	STV (ADRV2010, ADRV1201), CGSIV)	0.000 0.003	yes	yes	0.000	0.003
C62100135310	NA	D3TTS0	3	(ADRV2010, ADRV1201), CGSIV)	0.014	yes	yes	0.014	0.033
C710	NA	Q6GZV6	1	((FV3_SSME, FV3), (RGV, STIV)) (EHN, ESV) ATV CMTV_SP	0.000 0.001 0.001 0.012	yes	no	0.000	0.415
						yes	yes		0.000

Notes: Summary of output from Branch-Site REL and BUSTED methods, which were used to identify episodic, positive selection along specific branches or across all branches in the phylogeny, respectively. Taxon represents either the tip of a selected terminal branch or a clade (presented in Newick format) under a selected ancestral branch. Full virus names and the abbreviations used here are listed together in Table 1.



a similar, low rate of functional information for *Ranavirus* genes to previous studies; less than one-third of orthologs returned functional information from homology searches.^{19,53} This lack of functional annotation precludes predictions about how such lineage-specific genes might impact host-pathogen interactions, but gene knockout approaches will continue to be useful methods to investigate these roles.⁵⁴

Recombination was also found to be widespread in *Ranavirus* genomes and appeared to involve viruses separated by large geographic distances (based on existing patchy knowledge of *Ranavirus* geographic distribution). This result confirms and extends previous findings: Abrams et al.³¹ found evidence of widespread recombination when they examined a subset of the viruses studied here in positive selection analyses, while Epstein and Storfer³² also showed that regional isolates of ATV have been recombining.

In addition to assessing orthologs independently for evidence of recombination, the phylogenetic relationships of ToRV1 were compared across its genome. Phylogenetic analysis of each gene along the genome revealed two alternative positions for ToRV1, either at the base of the FV3-like clade or as a sister taxon to the FV3-like/CMTV-like clade. Genes returning each position were spatially clustered, with short sections of the genome supporting one position or the other, suggesting that ToRV1 is a mosaic virus. Of course, alternative explanations could be a sequencing library made up of two divergent viruses arising via lab contamination or coinfection.⁵⁵ These alternative hypotheses could be assessed - though not distinguished - with access to the original short reads. This virus was isolated from a captive, exotic reptile originating from the pet trade and as such can be considered at higher risk of exposure to diverse, geographically distant viruses resulting in a genuine mosaic virus or the type of mixed infection that might lead to an error in genome assembly.

There have been a number of attempts to use phylogenetics to address *Ranavirus* phylogeography and host jumps,^{19,30,56} but the apparent mosaic nature of ToRV1 shows that this can be very challenging. Recombination within genes can result in inaccurate inferences when those genes are used as candidates for phylogeny reconstruction.⁵⁷ *Ranavirus* genes with evidence of recombination include neurofilament triplet H1-like protein genes, which have been used previously for phylogenetics.⁵⁸

The 26 genes conserved across the family *Iridoviridae* have been frequently used in *Ranavirus* phylogenetics as a more robust alternative to candidate gene approaches, but six of these were shown here to be recombinant among ALRVs (Table 2). Stöhr et al.³⁰ classed ToRV1 as CMTV-like based on its genome arrangement - which matches viruses of that clade - but considered it an *intermediate* virus, given its placement in an FV3-like clade in a phylogenetic tree constructed from amino acid sequences of 17 genes. However, the phylogeny constructed here from 51 ALRV-core genes after removing recombinant loci provides strong support for the placement of ToRV1 as the outgroup to a CMTV-like/FV3-like clade. This

remains consistent with *Ranavirus* genome arrangements: the most recent common ancestor of ToRV1 and all CMTV-like and FV3-like ranaviruses shared the genome arrangement of CMTV-likes, with a major genome rearrangement occurring in an ancestor of FV3-likes following divergence of CMTV-likes and FV3-likes.

The finding that positive selection on *Ranavirus* genes is common (acting on ~20% of the genome) agrees with other recent studies.^{31,32} Selection on eIFalpha - a gene thought to be involved in immune evasion - has been shown previously in all studied strains of Ambystomatid salamander viruses and was found under selection again here in the Spanish CMTV (CMTV_SP; a virulent virus shown to have had a community-level impact²⁰) and RGV (a virus isolated from an aquaculture facility⁵⁹). Genes sharing homology with myristoylated membrane proteins were also found to have undergone positive selection on multiple branches. Myristoylation can play a key role in the envelopment of large DNA viruses and anchoring of virus particles to cell membranes facilitating entry to host cells.^{60,61}

In total, seven genes (more than any other virus studied) were found under selection along the terminal branch, leading to the Spanish isolate of CMTV, which is notable for the observations of its severe impacts on entire amphibian communities.²⁰ A putative nuclear calmodulin-binding protein was among these. Calmodulin is a calcium receptor that plays a critical role in a variety of cell responses to virus infection.⁶² The potential capacity of this protein to modify conserved host pathways makes it an interesting target in efforts to understand the host range of this virus isolate.

As discussed, this study follows a number of others in attempting to use phylogenetics to reconstruct the evolutionary history of ALRVs. Jancovich et al.¹⁹ proposed two alternate patterns of host switching consistent with their analysis of phylogenetic data, and the root of the ALRV was key to distinguish between them. Unfortunately, the analysis presented here failed to identify the root with any confidence since the topology and support was sensitive to the data supplied and the method applied. There was some support for ATV alone as the outgroup (which would mean that ATV-like viruses are paraphyletic) but monophyly of ATV-likes (ATV, EHNV and ESV) was also supported. The monophyly of ATV-likes with ATV at the root of this clade is significant as it suggests that the ancestor of ALRVs was an amphibian pathogen and there was a jump back into fish.

This study also highlighted some interesting results surrounding viruses from aquaculture. There was evidence of frequent positive selection in the tiger frog virus (TFV), which was isolated from an aquaculture facility in China.⁵³ These findings may therefore support the hypothesis that such facilities provide suitable environments for the rapid evolution of increased virulence.⁵⁰ *Ranavirus* isolates from aquaculture facilities in North America have already been shown to have increased virulence compared to related viruses from the wild, and the potential for spillback of viruses from industry to the wild should be considered a significant conservation threat.⁵⁰



Acknowledgments

The author thanks Richard Nichols for helpful discussions and also thanks Florent Lasalle for advice and sharing custom scripts to run HYPHY programs in batch. Finally, the author thanks four anonymous reviewers who provided helpful feedback, which improved the article.

Author Contributions

Conceived and designed the experiments: SJP. Analyzed the data: SJP. Wrote the article: SJP. The author reviewed and approved the final article.

Supplementary Material

Supplementary Figure 1. Distribution of ALRV orthologous gene lengths.

Supplementary Figure 2. Phylogenetic relationships among ranaviruses yielded inconsistent evidence for the root of ALRVs with overall placement of the root and support sensitive to the method followed. Alignments of 39 concatenated genes were used in tree construction with (a) a Bayesian method and nucleotide sequences (nt); (b) maximum likelihood method and nucleotide sequences; (c) Bayesian method and amino acid sequences (aa); and (d) maximum likelihood method and amino acid sequences. All trees were rooted on grouper iridovirus. The position of ATV is highlighted in red. Branch support values represent posterior probabilities from Mr. Bayes (Bayesian) or bootstrap values generated with RAxML (maximum likelihood; 100 replicates).

Supplementary Table 1. Orthologous gene presence-absence matrix (orthologs in rows and virus isolates in columns).

Supplementary Table 2. Summary of base composition, variation, recombination, positive selection, and presence/absence across *Iridoviridae* phylogeny for all ALRV orthologous genes.

REFERENCES

1. Holmes EC. Evolutionary history and phylogeography of human viruses. *Annu Rev Microbiol.* 2008;62(1):307–28.
2. Duffy S, Shackelton LA, Holmes EC. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet.* 2008;9(4):267–76.
3. Ramsden C, Holmes EC, Charleston MA. Hantavirus evolution in relation to its rodent and insectivore hosts: no evidence for codivergence. *Mol Biol Evol.* 2009;26(1):143–53.
4. Yutin N, Koonin EV. Hidden evolutionary complexity of nucleocytoplasmic large DNA viruses of eukaryotes. *Virology.* 2012;9:161.
5. Shackelton LA, Holmes EC. The evolution of large DNA viruses: combining genomic information of viruses and their hosts. *Trends Microbiol.* 2004;12(10):458–65.
6. Moreira D, Brochier-Armanet C. Giant viruses, giant chimeras: the multiple evolutionary histories of mimivirus genes. *BMC Evol Biol.* 2008;8:12.
7. Gilbert C, Peccoud J, Chateigner A, Moumen B, Cordaux R, Herniou EA. Continuous influx of genetic material from host to virus populations. *PLoS Genet.* 2016;12(2):e1005838.
8. Gilbert C, Chateigner A, Ernenwein L, et al. Population genomics supports baculoviruses as vectors of horizontal transfer of insect transposons. *Nat Commun.* 2014;5:3348.
9. Firth C, Kitchen A, Shapiro B, Suchard MA, Holmes EC, Rambaut A. Using time-structured data to estimate evolutionary rates of double-stranded DNA viruses. *Mol Biol Evol.* 2010;27(9):2038–51.
10. Chateigner A, Bézier A, Labrousse C, Jiolle D, Barbe V, Herniou EA. Ultra deep sequencing of a baculovirus population reveals widespread genomic variations. *Viruses.* 2015;7(7):3625–46.
11. Sijmons S, Thys K, Ngwese MM, et al. High-throughput analysis of human cytomegalovirus genome diversity highlights the widespread occurrence of gene-disrupting mutations and pervasive recombination. *J Virol.* 2015;89(15):7673–95.
12. Lassalle F, Depledge DP, Reeves MB, et al. Islands of linkage in an ocean of pervasive recombination reveals two-speed evolution of human cytomegalovirus genomes. *Virus Evol.* 2016;2(1):vew017.
13. Fischer MG. Giant viruses come of age. *Curr Opin Microbiol.* 2016;31:50–7.
14. Koonin EV, Yutin N. Origin and evolution of eukaryotic large nucleocytoplasmic DNA viruses. *Intervirology.* 2010;53(5):284–92.
15. Colson P, Lamballerie X, Yutin N, et al. “Megavirales”, a proposed new order for eukaryotic nucleocytoplasmic large DNA viruses. *Arch Virol.* 2013;158(12):2517–21.
16. Williams T, Barbosa-Solomieu V, Chinchar VG. A decade of advances in iridovirus research. *Adv Virus Res.* 2005;65:173–248.
17. Gray MJ, Miller DL, Hoverman JT. Ecology and pathology of amphibian ranaviruses. *Dis Aquat Organ.* 2009;87(3):243–66.
18. Chinchar VG, Hyatt A, Miyazaki T, Williams T. Family iridoviridae: poor viral relations no longer. *Curr Top Microbiol Immunol.* 2009;328:123–70.
19. Jancovich JK, Bremont M, Touchman JW, Jacobs BL. Evidence for multiple recent host species shifts among the ranaviruses (family iridoviridae). *J Virol.* 2010;84(6):2636–47.
20. Price SJ, Garner TWJ, Nichols RA, et al. Collapse of amphibian communities due to an introduced ranavirus. *Curr Biol.* 2014;24(21):2586–91.
21. Teacher AGF, Cunningham AA, Garner TWJ. Assessing the long-term impact of ranavirus infection in wild common frog populations. *Anim Conserv.* 2010;13(5):514–22.
22. Schloegel LM, Daszak P, Cunningham AA, Speare R, Hill B. Two amphibian diseases, chytridiomycosis and ranaviral disease, are now globally notifiable to the World Organization for Animal Health (OIE): an assessment. *Dis Aquat Organ.* 2010;92(2–3):101–8.
23. Schloegel LM, Picco AM, Kilpatrick AM, Davies AJ, Hyatt AD, Daszak P. Magnitude of the US trade in amphibians and presence of *Batrachochytrium dendrobatidis* and ranavirus infection in imported North American bullfrogs (*Rana catesbeiana*). *Biol Conserv.* 2009;142(7):1420–6.
24. Hoverman JT, Gray MJ, Haislip NA, Miller DL. Phylogeny, life history, and ecology contribute to differences in amphibian susceptibility to ranaviruses. *Ecohealth.* 2011;8(3):301–19.
25. Storfer A, Alfaro ME, Ridenhour BJ, et al. Phylogenetic concordance analysis shows an emerging pathogen is novel and endemic. *Ecol Lett.* 2007;10(11):1075–83.
26. Robertson DL, Hahn BH, Sharp PM. Recombination in AIDS viruses. *J Mol Evol.* 1995;40(3):249–59.
27. Pisoni G, Bertoni G, Puricelli M, Maccali M, Moroni P. Demonstration of coinfection with and recombination by caprine arthritis-encephalitis virus and maedi-visna virus in naturally infected goats. *J Virol.* 2007;81(10):4948–55.
28. Jehle JA, Fritsch E, Huber J, Backhaus H. Intra-specific and inter-specific recombination of torricid-specific granuloviruses during co-infection in insect larvae. *Arch Virol.* 2003;148(7):1317–33.
29. Eaton HE, Metcalf J, Penny E, Tcherpanov V, Upton C, Brunetti CR. Comparative genomic analysis of the family iridoviridae: re-annotating and defining the core set of iridovirus genes. *Virology.* 2007;4:11.
30. Stöhr AC, López-Bueno A, Blahak S, et al. Phylogeny and differentiation of reptilian and amphibian ranaviruses detected in Europe. *PLoS One.* 2015;10(2):e0118633.
31. Abrams AJ, Cannatella DC, Hillis DM, Sawyer SL. Recent host-shifts in ranaviruses: signatures of positive selection in the viral genome. *J Gen Virol.* 2013;94(pt 9):2082–93.
32. Epstein B, Storfer A. Comparative genomics of an emerging amphibian virus. *G3 (Bethesda).* 2015;6(1):15–27.
33. Kurtz S, Phillippy A, Delcher AL, et al. Versatile and open software for comparing large genomes. *Genome Biol.* 2004;5(2):R12.
34. Angiuoli SV, Salzberg SL. Mugsy: fast multiple alignment of closely related whole genomes. *Bioinformatics.* 2011;27(3):334–42.
35. Angiuoli SV, Hotopp JCD, Salzberg SL, Tettelin H. Improving pan-genome annotation using whole genome multiple alignment. *BMC Bioinformatics.* 2011;12(1):272.
36. Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet.* 2000;16(6):276–7.
37. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res.* 2010;38(suppl 2):W7–13.
38. Borowiec ML. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ.* 2016;4:e1660.
39. De Wit P, Pespeni MH, Ladner JT, et al. The simple fool’s guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Mol Ecol Resour.* 2012;12(6):1058–67.
40. Pond SLK, Posada D, Gravenor MB, Woelck CH, Frost SDW. GARD: a genetic algorithm for recombination detection. *Bioinformatics.* 2006;22(24):3096–8.



41. Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SDW. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol.* 2006;23(10):1891–901.
42. Abby SS, Tannier E, Gouy M, Daubin V. Detecting lateral gene transfers by statistical reconciliation of phylogenetic forests. *BMC Bioinformatics.* 2010;11:324.
43. Loytynoja A, Goldman N. An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A.* 2005;102(30):10557–62.
44. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–3.
45. Kosakovsky Pond SL, Murrell B, Fourment M, Frost SDW, Delpont W, Scheffler K. A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol.* 2011;28(11):3033–43.
46. Murrell B, Weaver S, Smith MD, et al. Gene-wide identification of episodic selection. *Mol Biol Evol.* 2015;32(5):1365–71.
47. Smith SA, Dunn CW. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics.* 2008;24(5):715–6.
48. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972–3.
49. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* 2001;17(8):754–5.
50. Hoverman JT, Gray MJ, Miller DL. Anuran susceptibilities to ranaviruses: role of species identity, exposure route, and a novel virus isolate. *Dis Aquat Organ.* 2010;89(2):97–107.
51. Duffus ALJ, Nichols RA, Garner TWJ. Experimental evidence in support of single host maintenance of a multihost pathogen. *Ecosphere.* 2014;5(11):142.
52. Chinchar VG, Yu KH, Jancovich JK. The molecular biology of frog virus 3 and other iridoviruses infecting cold-blooded vertebrates. *Viruses.* 2011;3(10):1959–85.
53. He JG, Lu L, Deng M, et al. Sequence analysis of the complete genome of an iridovirus isolated from the tiger frog. *Virology.* 2002;292(2):185–97.
54. Robert J, Jancovich JK. Recombinant ranaviruses for studying evolution of host-pathogen interactions in ectothermic vertebrates. *Viruses.* 2016;8(7):E187.
55. Hatwell JN, Sharp PM. Evolution of human polyomavirus JC. *J Gen Virol.* 2000;81(pt 5):1191–200.
56. Hyatt AD, Gould AR, Zupanovic Z, et al. Comparative studies of piscine and amphibian iridoviruses. *Arch Virol.* 2000;145(2):301–31.
57. Schierup MH, Hein J. Consequences of recombination on traditional phylogenetic analysis. *Genetics.* 2000;156(2):879–91.
58. Holopainen R, Ohlemeyer S, Schütze H, Bergmann SM, Tapiovaara H. Ranavirus phylogeny and differentiation based on major capsid protein, DNA polymerase and neurofilament triplet H1-like protein genes. *Dis Aquat Organ.* 2009;85(2):81–91.
59. Lei X-Y, Ou T, Zhu R-L, Zhang Q-Y. Sequencing and analysis of the complete genome of rana grylio virus (RGV). *Arch Virol.* 2012;157(8):1559–64.
60. Maurer-Stroh S, Eisenhaber F. Myristoylation of viral and bacterial proteins. *Trends Microbiol.* 2004;12(4):178–85.
61. Strecker T, Maisa A, Daffis S, Eichler R, Lenz O, Garten W. The role of myristoylation in the membrane association of the lassa virus matrix protein Z. *Virol J.* 2006;3:93.
62. Taylor JE, Chow JYH, Jeffries CM, et al. Calmodulin binds a highly extended HIV-1 MA protein that refolds upon its release. *Biophys J.* 2012;103(3):541–9.
63. Tan WGH, Barkman TJ, Chinchar VG, Essani K. Comparative genomic analyses of frog virus 3, type species of the genus ranavirus (family iridoviridae). *Virology.* 2004;323(1):70–84.
64. Morrison EA, Garner S, Echaubard P, Lesbarrères D, Kyle CJ, Brunetti CR. Complete genome analysis of a frog virus 3 (FV3) isolate and sequence comparison with isolates of differing levels of virulence. *Virol J.* 2014;11:46.
65. Huang Y, Huang X, Liu H, et al. Complete sequence determination of a novel reptile iridovirus isolated from soft-shelled turtle and evolutionary analysis of iridoviridae. *BMC Genomics.* 2009;10(1):224.
66. Mavian C, Lopez-Bueno A, Balseiro A, Casais R, Alcami A, Alejo A. The genome sequence of the emerging common midwife toad virus identifies an evolutionary intermediate within ranaviruses. *J Virol.* 2012;86(20):11413.
67. van Beurden SJ, Hughes J, Saucedo B, et al. Complete genome sequence of a common midwife toad virus-like ranavirus associated with mass mortalities in wild amphibians in the Netherlands. *Genome Announc.* 2014;2(6):e1293–14.
68. Wang N, Zhang M, Zhang L, et al. Complete genome sequence of a ranavirus isolated from Chinese Giant Salamander (*Andrias davidianus*). *Genome Announc.* 2014;2(1):e1032–13.
69. Chen Z, Gui J, Gao X, Pei C, Hong Y, Zhang Q. Genome architecture changes and major gene variations of *Andrias davidianus* ranavirus (ADRV). *Vet Res.* 2013;44:101.
70. Li W, Zhang X, Weng S, Zhao G, He J, Dong C. Virion-associated viral proteins of a Chinese giant salamander (*Andrias davidianus*) iridovirus (genus ranavirus) and functional study of the major capsid protein (MCP). *Vet Microbiol.* 2014;172(1–2):129–39.
71. Jancovich JK, Mao JH, Chinchar VG, et al. Genomic sequence of a ranavirus (family iridoviridae) associated with salamander mortalities in North America. *Virology.* 2003;316(1):90–103.
72. Mavian C, López-Bueno A, Fernández Somalo MP, Alcami A, Alejo A. Complete genome sequence of the European sheatfish virus. *J Virol.* 2012;86(11):6365–6.
73. Tsai CT, Ting JW, Wu MH, Wu MF, Guo IC, Chang CY. Complete genome sequence of the grouper iridovirus and comparison of genomic organization with those of other iridoviruses. *J Virol.* 2005;79(4):2010–23.
74. Song WJ, Qin QW, Qiu J, Huang CH, Wang F, Hew CL. Functional genomics analysis of Singapore grouper iridovirus: complete sequence determination and proteomic analysis. *J Virol.* 2004;78(22):12576–90.