




## ORIGINAL RESEARCH

# The draft genome of *Actinia tenebrosa* reveals insights into toxin evolution

Joachim M. Surm<sup>1,2</sup>  | Zachary K. Stewart<sup>3,4</sup> | Alexie Papanicolaou<sup>5</sup>  | Ana Pavasovic<sup>1</sup> | Peter J. Prentis<sup>3,4</sup> 

<sup>1</sup>Faculty of Health, School of Biomedical Sciences, Queensland University of Technology, Kelvin Grove, Qld, Australia

<sup>2</sup>Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, Qld, Australia

<sup>3</sup>Science and Engineering Faculty, School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, Qld, Australia

<sup>4</sup>Institute for Future Environments, Queensland University of Technology, Brisbane, Qld, Australia

<sup>5</sup>Hawkesbury Institute for the Environment, Sydney, NSW, Australia

**Correspondence**

Joachim M. Surm, Department of Ecology, Evolution and Behavior, Faculty of Science, Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel.  
Emails: joachim.surm@hdr.qut.edu.au; joachim.surm@mail.huji.ac.il

**Abstract**

Sea anemones have a wide array of toxic compounds (peptide toxins found in their venom) which have potential uses as therapeutics. To date, the majority of studies characterizing toxins in sea anemones have been restricted to species from the superfamily, Actinioidea. No highly complete draft genomes are currently available for this superfamily, however, highlighting our limited understanding of the genes encoding toxins in this important group. Here we have sequenced, assembled, and annotated a draft genome for *Actinia tenebrosa*. The genome is estimated to be approximately 255 megabases, with 31,556 protein-coding genes. Quality metrics revealed that this draft genome matches the quality and completeness of other model cnidarian genomes, including *Nematostella*, *Hydra*, and *Acropora*. Phylogenomic analyses revealed strong conservation of the Cnidaria and Hexacorallia core-gene set. However, we found that lineage-specific gene families have undergone significant expansion events compared with shared gene families. Enrichment analysis performed for both gene ontologies, and protein domains revealed that genes encoding toxins contribute to a significant proportion of the lineage-specific genes and gene families. The results make clear that the draft genome of *A. tenebrosa* will provide insight into the evolution of toxins and lineage-specific genes, and provide an important resource for the discovery of novel biological compounds.

**KEYWORDS**

Cnidaria, concerted evolution, sea anemone, venom

## 1 | INTRODUCTION

Cnidarian venom consists of a diverse array of peptides that have distinct biochemical and pharmacological properties (Frazão, Vasconcelos, & Antunes, 2012; Jouiaei, Sunagar, et al., 2015). These toxins are used for a variety of different roles, consistent with nematocyst morphology and function (Beckmann & Özbek, 2012; Fautin, 2009; Fautin & Mariscal, 1991; Kass-Simon & Scappaticci, 2002; Özbek, 2010). Multiple toxin types have been pharmacologically

characterized in cnidarians, including neurotoxins, pore-forming toxins, and enzymatic toxins (Casewell, Wüster, Vonk, Harrison, & Fry, 2013; Daly, 2016; Jouiaei, Sunagar, et al., 2015; Jouiaei, Yanagihara, et al., 2015; Prentis, Pavasovic, & Norton, 2018). Consistent with other venomous lineages, cnidarian venoms are a rich source of novel biological compounds, often being encoded by genes that lack homology to sequences other than cnidarians (Moran, Praher, et al., 2012; Sebé-Pedrós et al., 2018; Sunagar et al., 2018; Surm et al., 2019).

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

Recent studies have revealed a high frequency of cnidarian-specific genes is enriched within the cnidocyte (Sebé-Pedrós et al., 2018; Sunagar et al., 2018). Many of these cnidarian-specific genes expressed in the cnidocytes encode for toxin peptides (Columbus-Shenkar et al., 2018; Sebé-Pedrós et al., 2018). This highlights that cnidarians possess both morphological and biochemical novelties and that the evolution of these innovations may be related. This is consistent with evidence that acrorhagin 1 and acrorhagin 2 are novel toxin-coding genes which are localized to the acrorhagi, a morphological structure used for envenomation that is unique to sea anemones from Actinioidea (Honma et al., 2005; Macrander, Brugler, & Daly, 2015).

Indeed, understanding the evolution of venom and its delivery in cnidarians can provide insights into the innovation of morphological and biochemical novelties. While the majority of cnidarian toxin research has focussed on sea anemones from the Actinioidea superfamily (Prentis et al., 2018), no highly complete sequenced genomes for members of this superfamily currently exist (Urbarova et al., 2018). This lack of genomic resources limits our collective ability to understand the phylogenetic and molecular evolutionary histories of toxin-encoding genes within this superfamily. Such a resource would provide an excellent model to investigate the evolution of novel morphological and cellular structures, and their relationship with novel genes.

*Actinia tenebrosa* is a sea anemone from the superfamily Actinioidea. This species is similar in morphology to the northern hemisphere species, *Actinia equina* (Farquhar, 1898; Sherman, Peucker, & Ayre, 2007; Watts, Allcock, Lynch, & Thorpe, 2000), both of which have been used as model organisms for the investigation of sea anemone toxins (Honma et al., 2005; Maček & Lebez, 1988; Minagawa, Sugiyama, Ishida, Nagashima, & Shiomi, 2008; Moran et al., 2008; Norton, Maček, Reid, & Simpson, 1992; O'Hara, Caldwell, & Bythell, 2018; Prentis et al., 2018; Surm et al., 2019; Watts et al., 2000). Here, we have assembled and annotated the first draft genome for *A. tenebrosa*. We performed phylogenomic analyses and provide insights into the evolution of lineage-specific genes in cnidarians, specifically revealing that these novel genes undergo increased rates of expansions compared with gene families that have a wider taxonomic distribution. Moreover, genetic innovations restricted to Actinioidea are found to be enriched for functions related to venom and its delivery. The suite of toxin and toxin-like (TTL) genes identified in *A. tenebrosa* reveal an abundance of gene families evolving through lineage-specific duplications and, in some cases, concerted evolution. This study shows that gene duplication and divergent selective pressures have shaped the genetic variation in genes encoding toxins in actinarians.

## 2 | METHODS

### 2.1 | Genome assembly of *Actinia tenebrosa*

#### 2.1.1 | Sample preparation, sequencing, and assembly

Samples of *A. tenebrosa* were collected from the intertidal zone at Coolum, (QLD, Australia). Tissue from a single individual was used

to extract high-quality gDNA using the E.Z.N.A. Mollusc DNA Kit (Omega Bio-Tek; Stefanik, Wolenski, Friedman, Gilmore, & Finnerty, 2013). Extracted gDNA was used to construct four paired-end (PE) libraries sequenced on Illumina 2500 HiSeq platform using multiple insert sizes (170, 500, 2,000, 5,000 bp) with a read length of 100 bp (NCBI BioProject PRJNA505921). Sequencing resulted in over 150 million PE reads per library, with over 96% being high-quality ( $Q > 30$ ,  $[N(\text{ambiguous bases}) < 1\%]$ ). Contiguous sequences were generated and scaffolded using a manual operation of ALLPATH-LG (Butler et al., 2008) with a focus on removing redundant sequences.

The presence of the complete mitochondrial genome of *A. tenebrosa* in the draft genome was investigated. Assembled contigs were queried using BLASTN against a database which consisted of the complete mitochondrial genome of *A. equina*. Contigs receiving a significant hit ( $e$  value  $1e^{-05}$ ) were imported into Geneious 9.1.6 and aligned using a global alignment with free end gap and 100% identity. This resolved a single sequence, of 20,691 bp, and was aligned to the complete mitochondrial genome of *A. equina* using eight iterations of MUSCLE. Gene order and annotation of the mitochondrial genome of *A. tenebrosa* were performed as per Wilding and Weedall (2019).

#### 2.1.2 | Annotation

##### Repeat library generation

Homology and ab initio-based methods were used to identify repeat regions and low-complexity DNA sequences. Miniature Inverted-repeat Terminal Elements (MITEs) were predicted with MITE-HUNTER v.11-2011 (Han & Wessler, 2010) and detectMITE v.20170425 (Ye, Ji, & Liang, 2016). MITE predictions were clustered using CD-HIT v.4.6.4 (Fu, Niu, Zhu, Wu, & Li, 2012). Parameters = "cd-hit-est -c 0.8 -s 0.8 -aL 0.99 -n 5" (same parameters used by detectMITE). Prediction of long terminal repeat retrotransposons (LTR-RTs) was performed using LTRharvest (GT 1.5.10; Ellinghaus, Kurtz, & Willhoeft, 2008) and LTR\_FINDER v.1.06 (Xu & Wang, 2007), and these results were combined using LTR\_retriever commit 9b1d08d (Ou & Jiang, 2018) to identify canonical and noncanonical (i.e., non-TGCA motif) LTR-RTs. MITE and LTR-RT libraries were concatenated, and the genome sequence was masked using RepeatMasker open-4.0.7 (Smit, Hubley, & Green, 2013) with settings "-e ncbi -nolow -no\_is -norna." De novo repeat prediction was performed using RepeatModeler open-1.0.10 (Smit & Hubley, 2008) with the masked genome as input.

All repeat models were curated to remove models putatively part of protein-coding genes. Any models confidently annotated by LTR\_retriever or RepeatModeler (i.e., not classified as "Unknown") were removed from consideration as they are not likely to be part of protein-coding genes. Open reading frames from the remaining repeat models were extracted and examined using HMMER 3.1b2 (Eddy, 2011) to identify models that only contained domains associated with transposable elements. For this purpose, we collated a list of transposon-associated domains which primarily consisted of domains identified by Piriyaopongsa, Rutledge, Patel, Borodovsky, and

Jordan (2007) with additional Pfam (Finn et al., 2014) and NCBI CDD (Marchler-Bauer et al., 2015) domains included on the basis of manual inspection of domain prediction results for putative transposable elements. Repeat models that contained a TE-associated domain prediction were removed from consideration and assumed to be true-positives. A custom database of known genes was created to enable BLAST comparison of remaining repeat models and subsequent removal of false predictions from protein-coding genes. The database includes the UniProtKB/Swiss-Prot proteins as well as the gene models of *Nematostella vectensis* (v.2.0; Putnam et al., 2007; Schwaiger et al., 2014), *Exaiptasia pallida* (v.1.1; Baumgarten et al., 2015), *Acropora digitifera* (v.0.9; Shinzato et al., 2011), and *Hydra vulgaris* (Chapman et al., 2010). This database had probable transposons removed via the same process detailed above using HMMER 3.1b2 and domain organization. Any remaining repeat models were removed from the initial custom repeat library (CRL) if they had significant BLASTX hits ( $e$  value  $< 1e^{-02}$ ) when queried against the gene model database. The final curated CRL was used to soft-mask the *A. tenebrosa* genome using RepeatMasker (-e ncbi -s -nolow -no\_is -norna -xsmall) for later gene prediction. Scripts were produced to automate this process and are available from [https://github.com/zkstewart/Genome\\_analysis\\_scripts/tree/master/repeat\\_pipeline\\_scripts](https://github.com/zkstewart/Genome_analysis_scripts/tree/master/repeat_pipeline_scripts).

## 2.2 | Gene model prediction and annotation

Following the masking of repeat regions, gene models were predicted using ab initio methods guided by transcriptional expression. These reads included the Red and Brown ecotypes obtained from NCBI (Bioproject PRJNA313244; van der Burg, Prentis, Surm, & Pavasovic, 2016). Raw reads were quality trimmed using Trimmomatic (Bolger, Lohse, & Usadel, 2014) with parameters used by the Trinity de novo assembler (Haas et al., 2013; MacManes, 2014). Trimmed sequences were aligned against the genome using STAR 2.5 (Dobin et al., 2013) using the two-pass procedure for the de novo identification of transcription splice sites. The SAM file produced by STAR was converted to BAM format and sorted using samtools v.1.5 (Li et al., 2009). Gene models were predicted by BRAKER1 v1.11 (Hoff, Lange, Lomsadze, Borodovsky, & Stanke, 2016) using the soft-masked genome assembly and the STAR alignment file as inputs. The completeness of the protein-coding genes was then assessed using BUSCO (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015; Waterhouse et al., 2018).

Gene models were annotated by querying models against the Uniclust90 database (Mirdita et al., 2017) using MMseqs2 with an  $e$  value  $< 1e^{-05}$  (Steinegger & Söding, 2017). Gene Ontology (GO) terms associated with the representative UniProtKB sequence for each Uniclust90 hit were attributed to the *A. tenebrosa* gene model using the idmapping\_selected.tab file provided by UniProtKB. Protein domain predictions were performed by HMMER 3.1b2 using a custom domain database, which included NCBI's CDD in addition to CATH (S35 v.4.1.0; Dawson et al., 2017) and SUPERFAMILY (1.75; Gough, Karplus, Hughey, & Chothia, 2001), and tabulated using scripts available from [https://github.com/zkstewart/Genome\\_analysis\\_scripts/tree/master/annotation\\_table](https://github.com/zkstewart/Genome_analysis_scripts/tree/master/annotation_table).

## 2.3 | Gene family evolution

Using translated gene models from *Nematostella vectensis*, *Exaiptasia pallida*, *Acropora digitifera*, *Amplexidiscus fenestrafer*, *Discosoma* sp., and *Hydra vulgaris*, an "all-against-all" BLASTP analysis ( $e$  value  $< 10e^{-5}$ ) was performed. ORTHOMCL version 2.0.9 (Li, Stoeckert, & Roos, 2003) was used, with default parameters, to assign proteins into orthologous gene groups. Phylogenetic analyses were performed using single-copy orthologs (SCO) for each species. A total of 1,314 SCO were identified and aligned using clustal-omega (Sievers et al., 2011). The alignments were concatenated, and the best evolutionary mode protein model (JTT+F+I+G4) was determined. Finally, a maximum-likelihood tree with 1,000 ultrafast bootstrap replicates was generated using IQ-TREE (Nguyen, Schmidt, von Haeseler, & Minh, 2015).

Following the generation of a cnidarian species tree, the gain and loss of gene families across Cnidaria were inferred using the DOLLOP program from the PHYLIP package version 3.696 (Felsenstein, 1989; <http://evolution.genetics.washington.edu/phylip.html>). The species tree and a presence/absence matrix of gene families were imported into the DOLLOP program. The most parsimonious evolutionary scenario for the gain and loss of gene families was estimated using Dollo's parsimony law, which assumes genes arise once on the evolutionary tree and can be lost independently in different evolutionary lineages (Farris, 1977). The predicted proteomes from cnidarian species with sequenced genomes were used to investigate the evolution of protein domains. Protein domains were predicted using HMMER 3.1b2 against the Pfam database ( $e$  value  $< 1e^{-05}$ ), the best hit was retained, and overlapping domains were removed. A Fisher exact test was performed to determine Pfam enrichment with  $p$ -value of .05. Finally, we investigated the proportion of shared and unique gene families in actiniarian species. A BLASTP analysis ( $e$  value  $< 1e^{-05}$ ) was performed with OrthoVenn (Wang, Coleman-Derr, Chen, & Gu, 2015) using gene models from *A. tenebrosa*, *N. vectensis*, and *E. pallida* to determine the number of shared and unique gene families in each species.

The presence of toxin and toxin-like (TTL) genes was investigated in *A. tenebrosa*. The TTL genes were identified as per Surm et al. (2019). Briefly, BLASTP was performed against the manually curated Swiss-Prot database ( $e$  value  $< 1e^{-05}$ ). Significant queries with top BLAST annotations from sequences in the Tox-Prot database (Jungo & Bairoch, 2005) were considered candidate proteins. Candidate proteins were further interrogated and required to contain a signal peptide identified using SignalP (Petersen, Brunak, Heijne, & Nielsen, 2011).

The phylogenetic and evolutionary histories of multiple toxin gene families were investigated. Candidate sea anemone sodium channel inhibitory toxin (NaTx), sea anemone type 1 potassium channel toxin (KTx), sea anemone type 3 (BDS-LIKE) KTx, and membrane attack complex/perforin (MACPF) sequences were used for phylogenetic analysis to determine their distribution among cnidarian taxa and aligned to functionally characterized sequences (Jouiaei, Sunagar, et al., 2015; Sunagar & Moran, 2015). The fluorescent

**TABLE 1** Comparative genome metrics across Cnidaria

Annotation metrics	ADIG	AFEN	ATEN	DSPP	EPAL	NVEC	HVUG
Genome size (Mbp)	420	350	255	428	260	329/450	1,300
Assembly size (Mbp)	419	370	238	444	258	356	852
Total contig size (Mbp)	365	305	206	364	213	297	785
Total contig size (% of assembly)	87	82.43	86.56	81.98	82.5	83.4	92.2
Contig N50 (kbp)	10.9	20	8.4	18.7	14.9	19.8	9.7
Scaffold N50 (kbp)	191	510	159	769	440	472	92.5
Percent repetitive DNA	13	30.7	19.57	37.8	26	26	57
BUSCO (%)	74.7	83.7	89.6	86.3	87.3	91.6	77

Abbreviations: ADIG, *Acropora digitifera*; AFEN, *Amplexidiscus fenestrafer*; ATEN, *Actinia tenebrosa*; DSPP, *Discosoma* sp.; EPAL, *Exaiptasia pallida*; HVUG, *Hydra vulgaris*; NVEC, *Nematostella vectensis*.

protein (FP) gene family was also investigated to explore the evolution of nontoxin gene families. Sequences were identified in cnidarian genomes by the presence of GFP PFAM domain (PF01353) and aligned to sequences used in previous studies (Alieva et al., 2008; Ikmi & Gibson, 2010).

Protein alignments of candidate gene families were imported into IQ-TREE (v1.4.2; Nguyen et al., 2015) to determine a best fit of protein model evolution. Phylogenetic trees were generated from the alignments using 1,000 ultrafast bootstrap iterations and visualized using Figtree (v1.4.3, <http://tree.bio.ed.ac.uk/software/figtree/>). Selection analyses were performed on these gene families using previously published methods (Jouiaei, Sunagar, et al., 2015; Sunagar & Moran, 2015).

### 3 | RESULTS

#### 3.1 | Genome assembly

Using a whole-genome shotgun strategy, we sequenced and assembled the genome of *A. tenebrosa*. A total of 1.2 billion paired-end reads, with a length of 100 bp, were sequenced across four different insert size libraries (170, 500, 2,000, and 5,000 bp; Table S1). Raw reads were used to assemble the *A. tenebrosa* genome using ALLPATHS-LG. The genome size of *A. tenebrosa* is estimated to be ~255 Mbp (Table 1). The draft genome assembled is of similar quality to other cnidarian genomes (Table 1). Although the assembly resulted in the scaffold and contig N50 lower than other cnidarian genomes, the predicted genome completeness using metazoan Augustus gene models is among the highest (89.6%) for cnidarian genomes, with only *N. vectensis* having a more complete assembly (91.6%). The assembly contains ~19% repetitive DNA sequences, which is similar to reported values for other cnidarians (Tables 1 and S2).

#### 3.2 | Functional annotation of predicted gene models

The ab initio gene model prediction identified 31,556 protein-coding genes in *A. tenebrosa*. All gene models were validated, receiving significant BLAST hits against multiple *A. tenebrosa* transcriptomes (van der Burg et al., 2016). Our ab initio gene model prediction was highly complete compared with other cnidarian genomes, increasing the previous BUSCO score to 94.6% (Table 1). Only *E. pallida* gene models were more complete (94.7%). Of the 31,556 protein-coding genes, 19,022 and 25,478 returned a significant BLAST hit ( $e$  value  $1e^{-05}$ ) against the Swiss-prot and TREMBL database, respectively. This highlights that ~80% of the predicted proteome shares sequence similarity to known protein sequences, with ~20% having no similarity to other proteins. In contrast, only 6.56% of *E. pallida* predicted proteome returned no hits to known proteins at this stringency. However, other cnidarian genomes returned similar levels of novelty, with *Discosoma* sp. having 16.17% of proteins returning no hits. The annotation of protein domains revealed 19,056 (~60%) gene models contained identifiable Pfam domains. This is less than other sea anemone genomes, with 78.64% and 68.35% of *E. pallida* and *N. vectensis* gene models having a protein domain, respectively. Additionally, both corallimorpharians genomes reported less than 60% of gene models to encode proteins with known protein domains. Taken together, these results highlight that the draft genome of *A. tenebrosa* is mostly complete, yet a significant proportion of its genes are unique (Table 2).

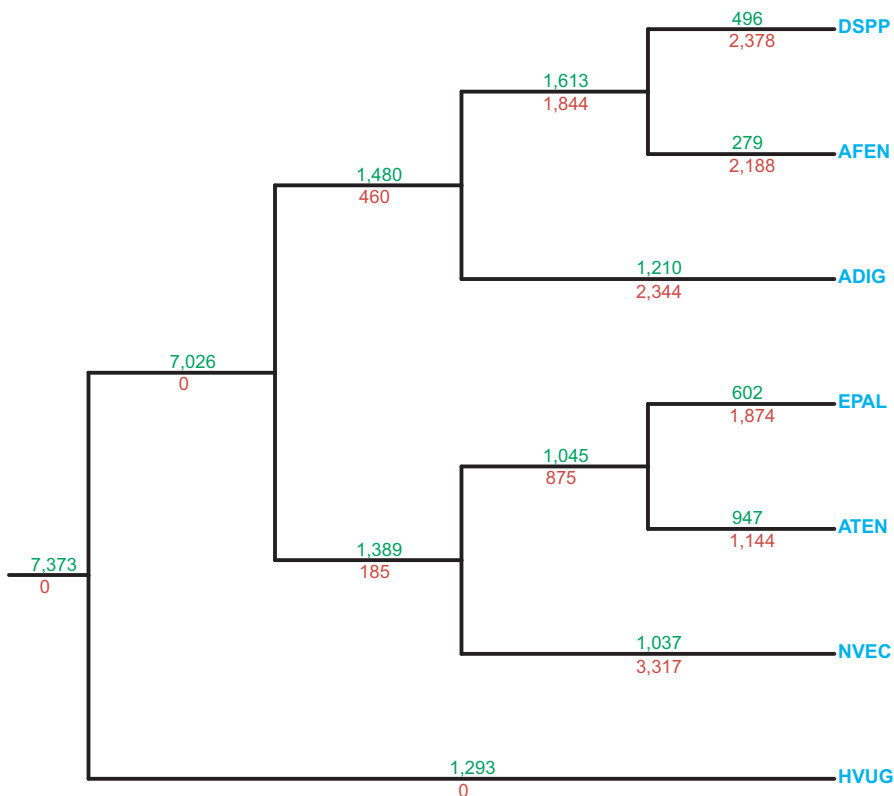
Our assembly also resolved the complete mitochondrial genome for *A. tenebrosa* (GenBank accession MK291977), shown to be 20,691 bp long (Figure S1). The mitochondrion of *A. tenebrosa* was aligned to the recently completed *A. equina* mitochondrion (Wilding & Weedall, 2019), revealing identical gene order and protein-coding

**TABLE 2** Functional annotation of gene models from seven cnidarian species

Annotation metrics	ADIG	AFEN	ATEN	DSPP	EPAL	HVUG	NVEC
BUSCO (%)	80.5	72.8	94.6	68.6	94.7	91.5	93.8
Protein-coding genes	33,878	21,372	31,556	23,199	26,087	21,990	24,780
SP annotation	24,094	12,959	19,022	13,562	20,515	15,923	18,974
SP annotation (%)	71.12	60.64	60.28	58.46	78.64	72.41	76.57
No SP annotation (%)	28.88	39.36	39.72	41.54	21.36	27.59	23.43
TREMBL annotation	30,116	18,106	25,478	19,447	24,376	19,992	208,698
TREMBL annotation (%)	88.90	84.72	80.74	83.83	93.44	90.91	84.21 <sup>a</sup>
No TREMBL annotation (%)	11.10	15.28	19.26	16.17	6.56	9.09	15.78 <sup>a</sup>
Pfam	24,000	12,686	19,056	13,283	20,514	15,665	16,938
Pfam annotated (%)	70.84	59.36	60.39	57.26	78.64	71.24	68.35
No Pfam annotated (%)	29.16	40.64	39.61	42.74	21.36	28.76	31.65
Total Pfam found	52,242	27,154	42,834	27,355	45,944	28,984	30,605
Pfam per gene	1.54	1.27	1.36	1.18	1.76	1.32	1.24

Abbreviations: ADIG, *Acropora digitifera*; AFEN, *Amplexidiscus fenestrafer*; ATEN, *Actinia tenebrosa*; DSPP, *Discosoma* sp.; EPAL, *Exaiptasia pallida*; HVUG, *Hydra vulgaris*; NVEC, *Nematostella vectensis*.

<sup>a</sup>As the predicted proteome of *N. vectensis* is incorporated into the TREMBL protein database, a subset of TREMBL's database with *N. vectensis* predicted proteins removed was used instead.



**FIGURE 1** Comparative analysis of gene families within Cnidaria. Maximum-likelihood protein tree generated to determine cnidarian phylogeny, all bootstrap support 100%. TTL gene family gains (green) and losses (red) are represented above and below branches, respectively. ADIG, *Acropora digitifera*; AFEN, *Amplexidiscus fenestrafer*; ATEN, *Actinia tenebrosa*; DSPP, *Discosoma* sp.; EPAL, *Exaiptasia pallida*; HVUG, *Hydra vulgaris*; NVEC, *Nematostella vectensis*

sequence similarity. Nucleotide differences in the mitochondrion of *A. tenebrosa* and *A. equina* included a thymine insertion in the intergenic region between genes *ND6* and *CYTB* in *A. tenebrosa*, a transversion SNP was identified in the large RNA subunit, and a transition SNP was identified in the intergenic region between *COIII* and *COI* genes.

### 3.3 | Gene family evolution

Manual curation and a phylogenomic characterization of seven Cnidarian species drove our investigation of Cnidarian gene turnover. Using 1,314 genes, we built a representative cnidarian species tree from all seven genomes (Figure 1). This species tree confirmed

**TABLE 3** Expansion of shared and lineage-specific gene families in cnidarians

	ADIG	AFEN	ATEN	DSPP	EPAL	HVUG	NVEC
Total genes	33,878	21,372	31,556	23,199	26,087	21,990	24,780
Singletons	4,053	5,261	6,705	5,752	2,590	2,800	5,492
Singletons (%)	11.96	24.62	21.25	24.79	9.93	12.73	22.16
Total gene families	14,285	13,279	15,576	13,306	14,501	8,666	13,323
Total genes in gene families	29,825	16,111	24,851	17,447	23,497	19,190	19,288
Expansion	2.09	1.21	1.6	1.31	1.62	2.21	1.45
Lineage-specific gene families	1,210	279	947	496	602	1,293	1,037
Lineage-specific gene families (%)	8.47	2.1	6.08	3.73	4.15	14.92	7.78
Lineage-specific genes	4,238	659	3,963	1,232	1,830	6,451	3,447
Expansion	3.5	2.36	4.18	2.48	3.04	4.99	3.32
Shared gene families	13,075	13,000	14,629	12,810	13,899	7,373	12,286
Shared gene families (%)	91.53	97.90	93.92	96.27	95.85	85.08	92.22
Shared genes	25,587	15,452	20,888	16,215	21,667	12,739	15,841
Expansion	1.96	1.19	1.43	1.27	1.56	1.73	1.29

Abbreviations: ADIG, *Acropora digitifera*; AFEN, *Amplexidiscus fenestrafer*; ATEN, *Actinia tenebrosa*; DSPP, *Discosoma* sp.; EPAL, *Exaiptasia pallida*; HVUG, *Hydra vulgaris*; NVEC, *Nematostella vectensis*.

the phylogenetic position of *A. tenebrosa* with previously published species trees (Daly et al., 2017; Rodríguez et al., 2014; Wang et al., 2017). We found 7,373 gene families were shared among all cnidarian taxa investigated. An additional 7,026 gene families were gained in Anthozoa following their divergence from Medusozoa (*H. vulgaris*). In the actiniarian lineage (which includes *A. tenebrosa*, *E. pallida*, and *N. vectensis*), 1,389 and 185 gene families were gained and lost, respectively. Examination of the genome of *A. tenebrosa* found that 947 gene families (3,963 genes) were gained in this species following divergence from other sea anemone taxa investigated. In all cnidarians, lineage-specific gene families have undergone a greater expansion compared with gene families shared among cnidarians (Table 3). This is most apparent in *A. tenebrosa* and *H. vulgaris*, with lineage-specific gene families having a mean copy number of 4.18 and 4.99 genes, respectively. Additional novelty is observed with 6,705 (21.26%) singletons (lineage-specific genes not in gene families) found in the gene models of *A. tenebrosa*. These results suggest significant gene family conservation across cnidarians, particularly in Anthozoa, but with lineage-specific genes contributing to a significant proportion of the genome.

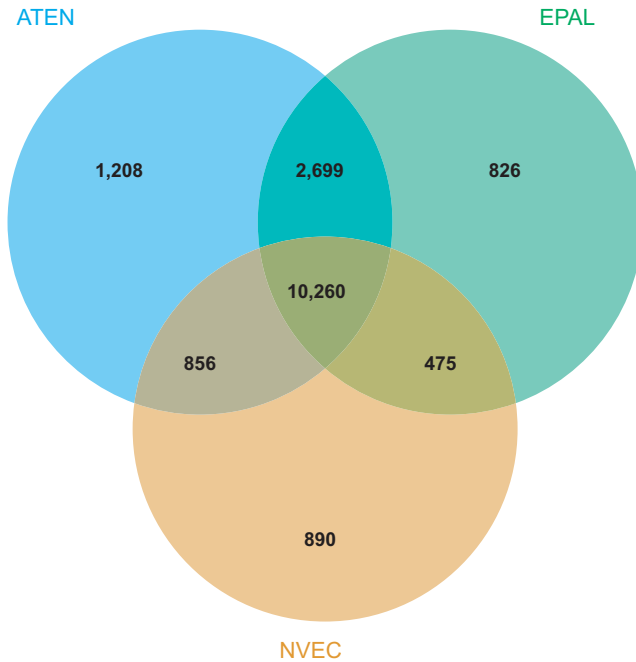
A closer examination of gene families within Actiniaria revealed 10,260 orthologs shared across the three actiniarian genomes investigated (Figure 2). These 10,260 actiniarian orthologs, however, do not exhibit any GO term enrichment. Five GO terms, including nematocyst (GO: 0042151; Table S3), were over-represented in the predicted protein sequences from the 1,208 genes unique to *A. tenebrosa*. This highlights that a significant proportion of genes unique to *A. tenebrosa* have roles related to envenomation. Although all actiniarians are venomous, we observe, therefore, the first expansion of lineage-specific genes is related to venom delivery.

To better understand the evolution of protein domains across cnidarian genomes, we also investigated Pfam domain enrichment.

Using a Fisher exact test, 25 Pfam domains were significantly enriched in *A. tenebrosa*, in comparison with other cnidarian genomes (Figure 3). Enrichment of ShK and Defensin\_4 domains underpinned much of the expansion of toxin-related genes in *A. tenebrosa*. Both ShK and Defensin\_4 domains are associated with potassium channel-blocking toxins in sea anemones, specifically sea anemone type 1 potassium channel toxin (KTx) and type 3 (BDS-LIKE) KTx, respectively (Castañeda et al., 1995).

With evidence supporting that genetic innovations in the genome of *A. tenebrosa* are related to venom, we further investigated its total and toxin-like gene (TTL) complement. Overall, we identified 113 TTL gene families in *A. tenebrosa* (Table S4). Manual curation of TTL genes revealed that sea anemone type 3 (BDS-LIKE) KTx family is the most highly expanded TTL gene family (15 copies, 11 of which are full-length sequences). A phylogeny of sea anemone type 3 (BDS-LIKE) KTx was generated from these full-length sequences (Figure 4), as well as functionally characterized sequences from other sea anemones (Jouiaei, Sunagar, et al., 2015; Sunagar & Moran, 2015). The 11 *A. tenebrosa* sequences clustered into four distinct clades, one of which includes only *A. tenebrosa* paralogs (Clade A). This suggests a process of concerted-like evolution. Investigation into the genomic localization of the 11 *A. tenebrosa* sequences revealed no presence of tandem duplication, a common mechanism observed during concerted evolution. Furthermore, the sequence identity among *A. tenebrosa* sea anemone type 3 (BDS-LIKE) KTx paralogs is highly divergent at 34.2% and 43.5% at the nucleotide and protein level, respectively.

The sea anemone sodium channel inhibitory toxin family (NaTx) has also previously been shown to evolve via concerted evolution in multiple different sea anemone species (Moran et al., 2008). Here we generated a phylogeny for the NaTx gene family (Figure 5), using sequences from a previously published alignment



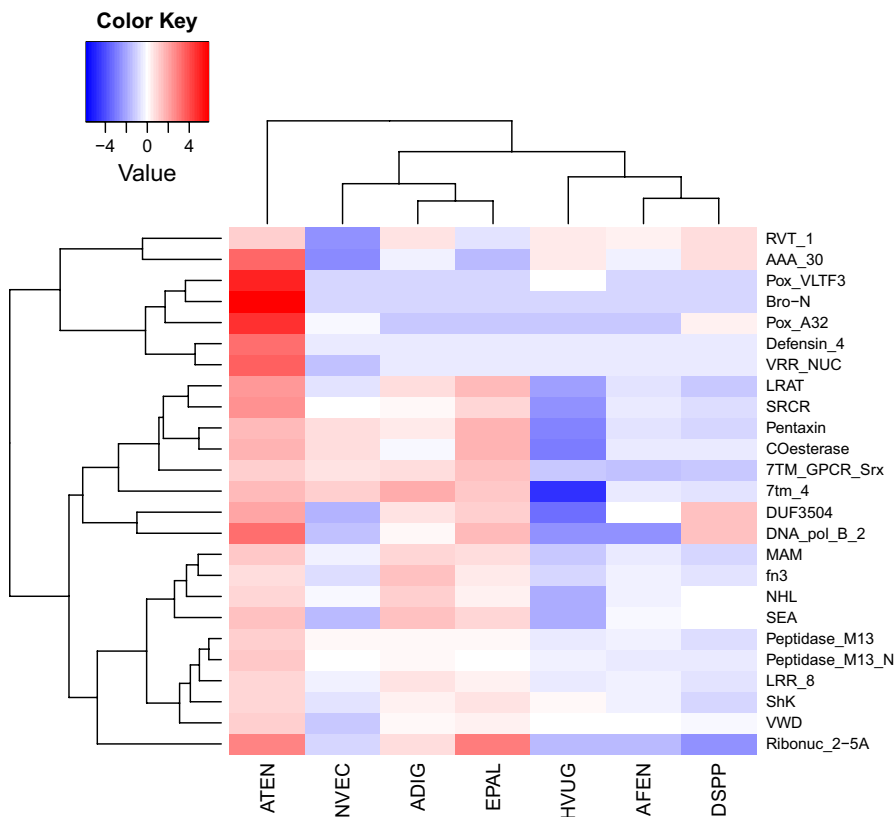
**FIGURE 2** Comparative analysis of gene families among Actinarians. Venn diagram highlighting orthologous genes between Actinarian genomes. EPAL, *Exaiptasia pallida*; NVEC, *Nematostella vectensis*; and ATEN, *Actinia tenebrosa*

(Jouiaei, Sunagar, et al., 2015; Sunagar & Moran, 2015), as well as newly identified sequences from *N. vectensis* (Nv4, Nv5, Nv6, Nv7, and Nv8; Sachkova et al., 2019). The phylogeny of NaTx gene

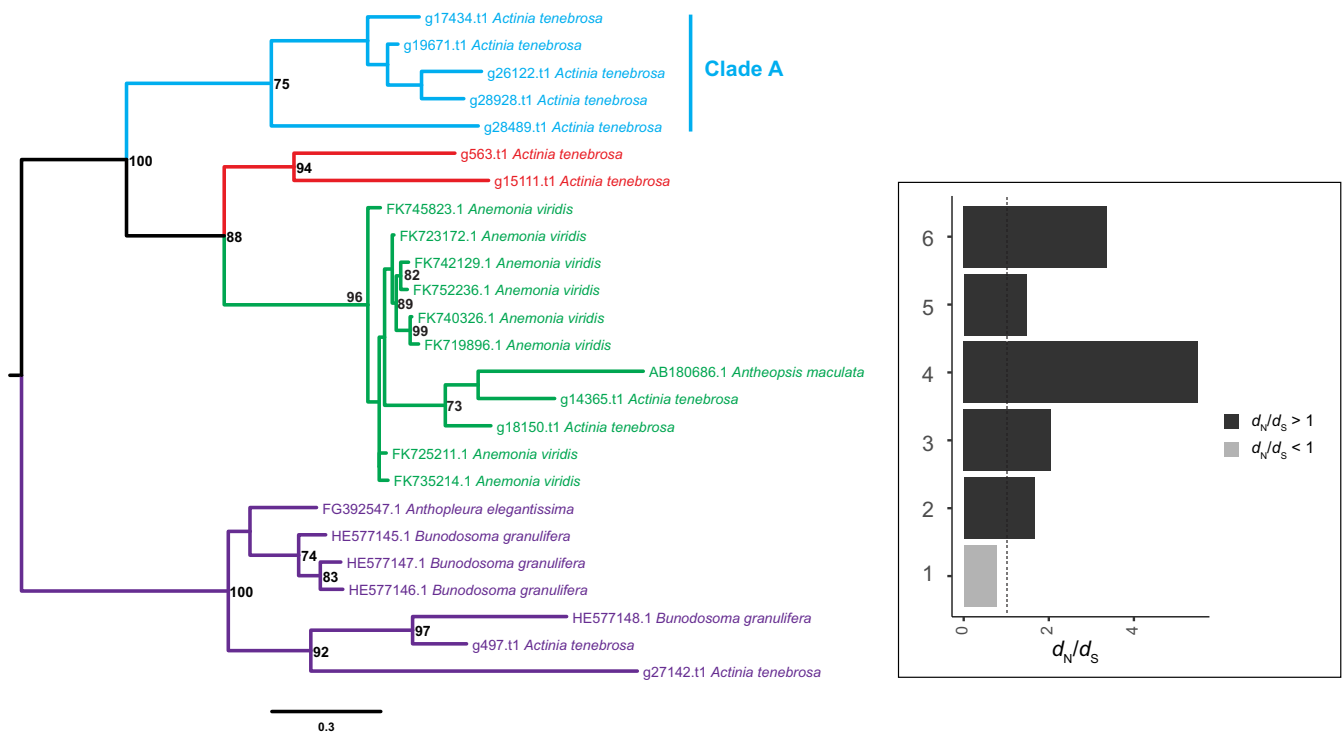
family confirmed evidence of concerted evolution in multiple species, with paralogs clustering strongly together in *N. vectensis*, *A. viridis*, and *A. equina*. Three copies of NaTx were identified in *A. tenebrosa*, with two copies clustering together and another paralog clustering with *A. equina* sequences. The three *A. tenebrosa* paralogs share 61.1% and 52.9% sequence similarity at the nucleotide and protein level, respectively. Additionally, the genome of *A. tenebrosa* revealed no evidence of tandem duplication for the three NaTx paralogs.

Investigating the phylogenetic histories of cnidarian MACPF gene family also revealed evidence of concerted-like evolution (Figure 6). This included two paralogs of *A. tenebrosa* MACPF sequences clustering together. Genomic localization further revealed these sequences evolved through tandem duplication (Figure S2). Evidence of concerted evolution was also revealed with *A. tenebrosa* MACPF sequences being highly homogenous, sharing 94.3% and 92.8% similarity at the nucleotide and protein level, respectively. In fact, clustering of paralogs of MACPF was observed in the majority of the anthozoan genomes investigated, including all sea anemones. Multiple tandem duplications were observed in *E. pallida*; however, this was not consistent in all sea anemones with *N. vectensis* paralogs not adjacent to each other in the genome. We also found evidence of concerted-like evolution in the sea anemone type 1 KTx family (Figure 7). While we did not observe this for *A. tenebrosa* paralogs, this process was observed for *A. viridis* paralogs.

While concerted-like evolution appears to be a consistent pattern of TTL genes families in sea anemones, similar pattern is also observed broadly in cnidarians fluorescent protein (FP) family (Figure



**FIGURE 3** Protein domain enrichment across Cnidaria. Heat map of Pfam domains enriched in *Actinia tenebrosa*. Abundance of Pfam domains in cnidarians log<sub>2</sub> and median centered. ADIG, *Acropora digitifera*; AFEN, *Amplexidiscus fenestrafer*; ATEN, *Actinia tenebrosa*; DSPP, *Discosoma* sp.; EPAL, *Exaiptasia pallida*; HVUG, *Hydra vulgaris*; NVEC, *Nematostella vectensis*



**FIGURE 4** Maximum-likelihood tree with midpoint root depicting relationships among sea anemone type 3 (BDS-LIKE) KTx coding sequences. Bootstrap values after 1,000 iterations are shown next to nodes, values under 70% not reported. The GenBank accession numbers for the protein-coding gene used in this phylogenetic analysis are described in Sunagar and Moran (2015). A corresponding bar plot is provided which shows the computed dN/dS value for orthologs and paralogs. 1 = Actiniaria orthologs, 2 = *Anemonia viridis* paralogs, 3 = *Actinia tenebrosa* paralogs, 4 = *Bunodosoma granulifera* paralogs, 5 = *Actinia tenebrosa* Clade A paralogs (Clade A), and 6 = *Actinia tenebrosa* paralogs diverged

S3), a nontoxin gene family. Combining published data (Alieva et al., 2008; Ikmi & Gibson, 2010), with the genomic datasets from this study, we observed a consistent pattern of paralogs clustering together. This is observed for *A. tenebrosa* FP paralogs that cluster together in a clade consisting of other sea anemone chromoprotein sequences. While these sequences from *A. tenebrosa* appear to be evolving via concerted evolution, this appears to not be reliant on tandem duplication. While the *A. tenebrosa* paralogs sequence have escaped tandem duplication, they maintain a high level of sequence identity of 95.6% and 93.6% at the nucleotide and protein level, respectively. A similar pattern is also observed in other Hexacorallia taxa including *N. vectensis* and *A. fenestrafer*.

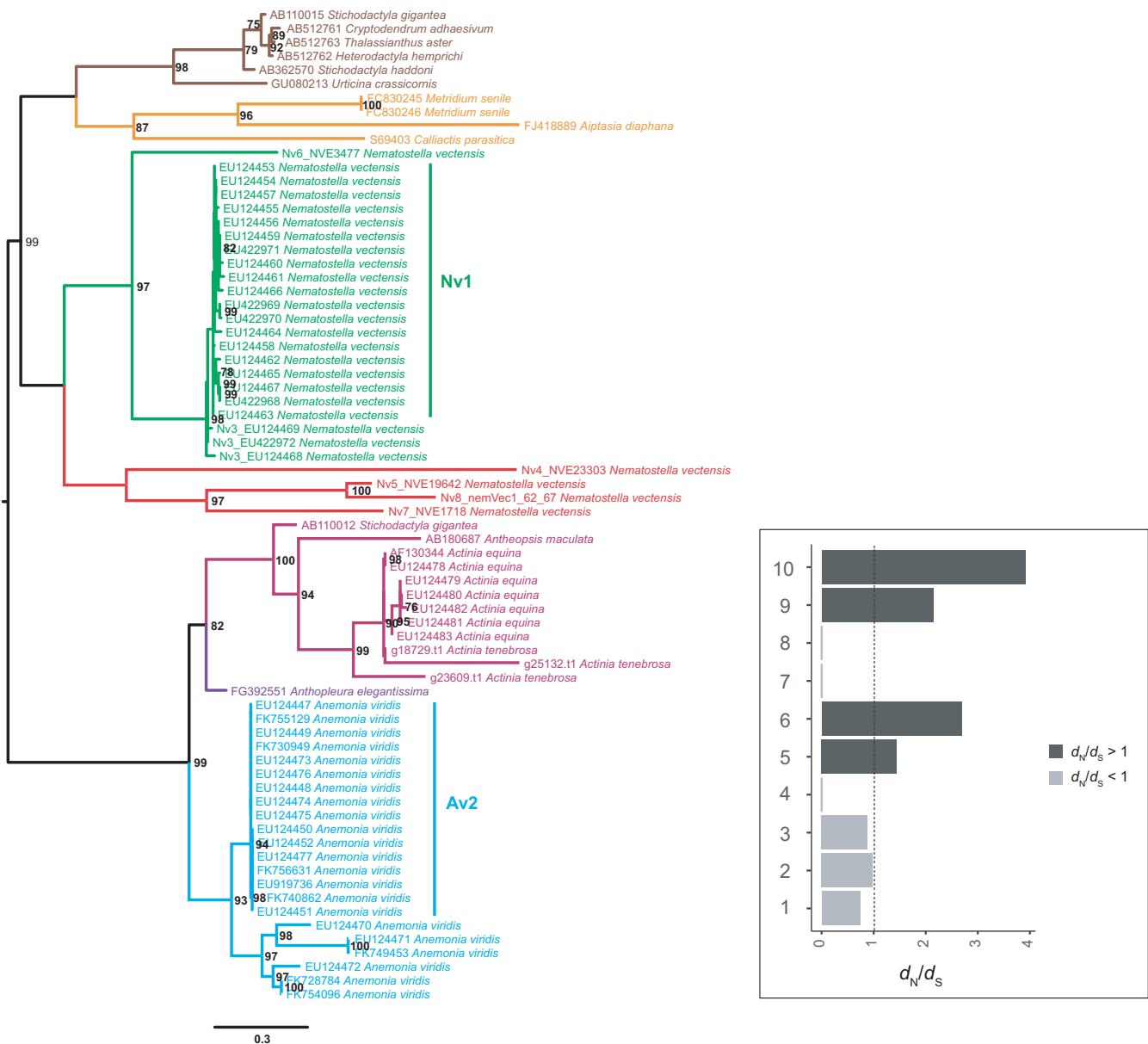
### 3.4 | Selection patterns on toxin gene families

In this study, we further explored the evolutionary histories of TTL gene families to provide insights into the selective pressures acting on them (Table S5). Here we report evidence of purifying selection acting on all TTL gene families. Given the evidence of concerted-like evolution acting on many of the gene families investigated, we tested the selective pressures of paralogs where possible. In *A. tenebrosa*, paralogs from the sea anemone type 3 (BDS-LIKE) KTx (Figure 4; dN/dS = 2.0515) gene family revealed evidence of diversifying selection. Similarly, in *A. viridis*, sea anemone type 3

(BDS-LIKE) KTx paralogs also appear to be evolving under diversifying selection (Figure 4; dN/dS = 1.6665). We further explored the selective pressures acting on NaTx paralogs. Paralogs from *A. tenebrosa* and *A. viridis* appear to be evolving under diversifying selection, with a dN/dS ratio of 1.4438 and 2.6865, respectively. In *A. equina*, however, we cannot confirm diversifying selection acting on paralogs. Differences in selective pressures among a subset of NaTx orthologs were also observed, with orthologs from *Actinia* genus (dN/dS = 0.9825) and appear to be evolving under a relaxed rate of purifying selection compared to among actiniarian orthologs (dN/dS = 0.7397). *Nematostella vectensis* NaTx paralogs had a dN/dS ratio (0.881) consistent with the action of purifying selection. Additionally, all sea anemone type 1 KTx paralogs are inferred to be evolving under purifying selection, with the exception of *A. viridis* (Figure 7; dN/dS = 3.0389).

Divergent evolutionary histories were also observed among paralogs of gene families that appear to be evolving through a process of concerted-like evolution. Specifically, in NaTx and sea anemone type 3 (BDS-LIKE) KTx gene families, some paralogs are evolving through a process of concerted evolution and others are escaping this process. This is observed for *N. vectensis* and *A. viridis* NaTx paralogs, and *A. tenebrosa* sea anemone type 3 (BDS-LIKE) KTx paralogs. For *N. vectensis*, no evidence of positive selection could be inferred among the highly homogenous Nv1 sequences (Figure 5); however,





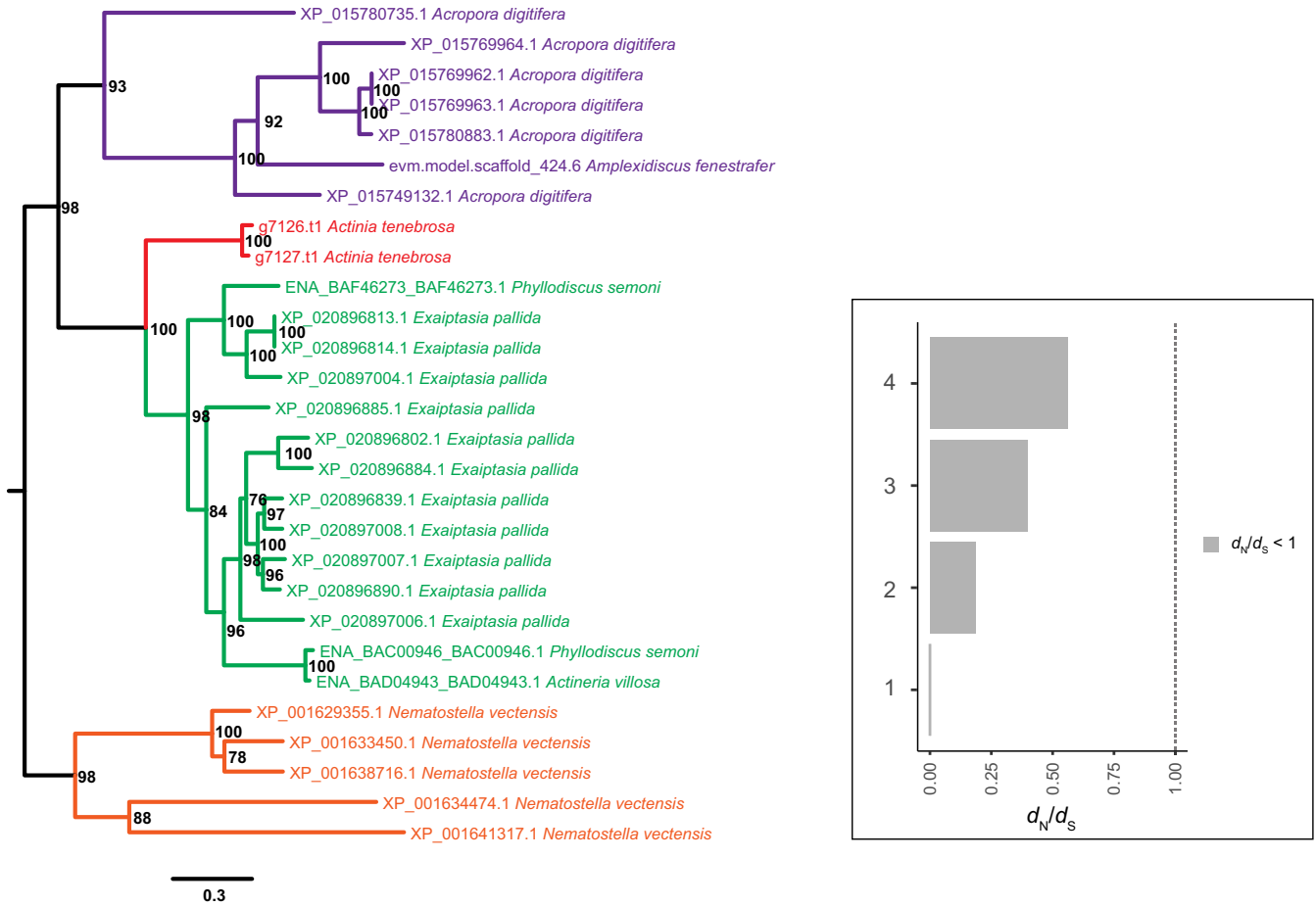
**FIGURE 5** Maximum-likelihood tree with midpoint root depicting relationships among NaTx coding sequences. Bootstrap values after 1,000 iterations are shown next to nodes, values under 70% not reported. The GenBank accession numbers for the protein-coding gene used in this phylogenetic analysis are described in Sunagar and Moran (2015) and Sachkova et al. (2019). A corresponding bar plot is provided which shows the computed  $d_N/d_S$  value for orthologs and paralogs. 1 = Actiniaria orthologs, 2 = *Actinia* orthologs, 3 = *Nematostella vectensis* paralogs, 4 = *Actinia equina* paralogs, 5 = *Actinia tenebrosa* paralogs, 6 = *Anemonia viridis* paralogs, 7 = Nv1 paralogs, 8 = Av2 paralogs, 9 = *Nematostella vectensis* paralogs diverged, and 10 = *Anemonia viridis* paralogs diverged

paralogs that have escaped this homogenization are inferred to be evolving under diversifying selection (Nv3-8;  $d_N/d_S = 2.1451$ ). This is also observed for *A. viridis* NaTx paralogs. While we could not infer the selective pressures acting on the highly homogenous Av2 sequences, those that have diverged are undergoing diversifying selection ( $d_N/d_S = 3.9116$ ). In *A. tenebrosa*, some sea anemone type 3 (BDS-LIKE) KTx sequences paralogs also show strong clustering (Figure 4 Clade A), while other paralogs cluster with sequences from other sea anemones. While both sets of paralogs are evolving under diversifying selection (Clade A  $d_N/d_S = 1.4761$ , diverging paralogs in *A. tenebrosa*  $d_N/d_S = 3.3732$ ), those that have diverged

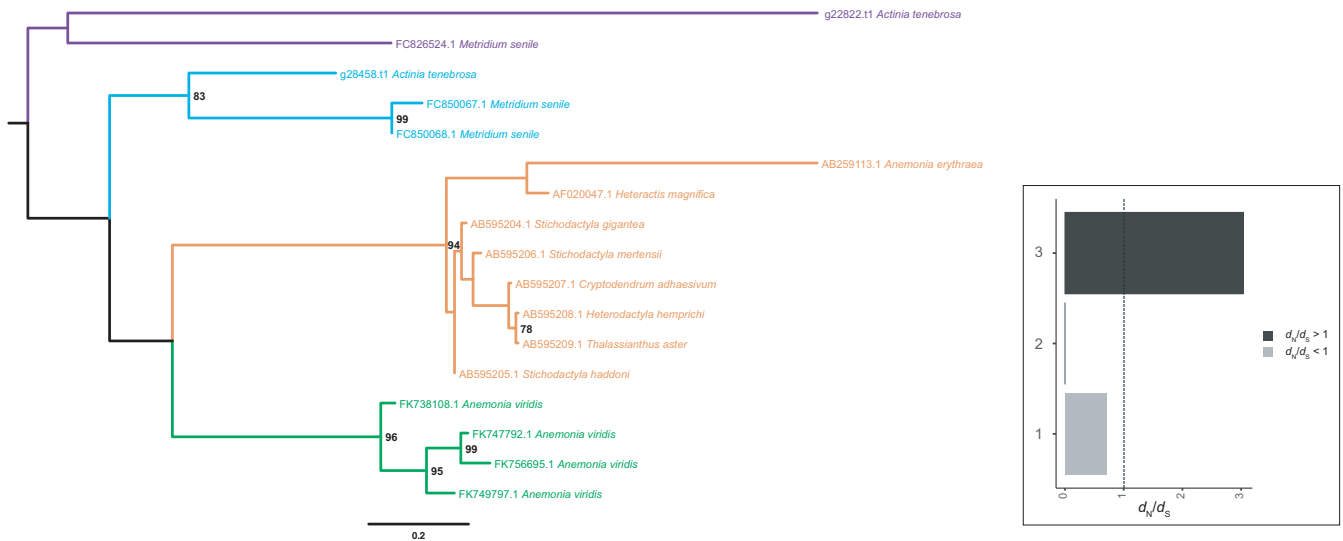
show pronounced signatures of diversifying selection. Additionally, we also have evidence of concerted-like evolution for the FP and MACPF gene families; however, we did not observe any paralogs escaping this process (Figures 6 and S3).

## 4 | DISCUSSION

Here we present a draft genome assembly and annotation of *A. tenebrosa*. This complete draft assembly is the first from any species of the superfamily Actinioidea. Overall, the assembly was of



**FIGURE 6** Maximum-likelihood tree with midpoint root depicting relationships among MACPF coding sequences. Bootstrap values after 1,000 iterations are shown next to nodes, values under 70% not reported. The European Nucleotide accession numbers for the protein-coding genes are reported. A corresponding bar plot is provided which shows the computed  $d_N/d_S$  value for orthologs and paralogs. 1 = Anthozoan orthologs, 2 = *Nematostella vectensis* paralogs, 3 = *Exaiptasia pallida* paralogs, and 4 = *Acropora digitifera* paralogs



**FIGURE 7** Maximum-likelihood tree with midpoint root depicting relationships among sea anemone type 1 KTx coding sequences. Bootstrap values after 1,000 iterations are shown next to nodes, values under 70% not reported. The GenBank accession numbers for the protein-coding gene used in this phylogenetic analysis are described in Sunagar and Moran (2015). A corresponding bar plot is provided which shows the computed  $d_N/d_S$  value for orthologs and paralogs. 1 = Actiniaria orthologs, 2 = *Metridium senile* paralogs, and 3 = *Anemonia viridis* paralogs

similar quality and completeness to currently published anthozoan genomes (Baumgarten et al., 2015; Chapman et al., 2010; Putnam et al., 2007; Shinzato et al., 2011; Wang et al., 2017), verifying its suitability for comparative genomic studies. Insights into the evolution of gene families across Cnidaria revealed significant conservation among anthozoan species, with the many gene families gained in either the last common ancestor of Cnidaria or Anthozoa. Notably, all anthozoans used in this study are from Hexacorallia, highlighting a high conservation of gene families shared among this subclass. This is consistent with previous studies that have suggested that this shared gene set plays an important role in the evolution of traits essential to Hexacorallia taxa, including symbiosis with dinoflagellates, stress response, and delivery of venom (Baumgarten et al., 2015; Rachamim et al., 2015; Wang et al., 2017).

The *A. tenebrosa* genome is the most gene dense among cnidarians, with only *E. pallida* having a smaller genome and only *A. digitifera* having more protein-coding genes. However, flow cytometry revealed that the genome size of *A. equina* is larger (~520 Mb) than that predicted here ~255 Mb (Adachi, Miyake, Kuramochi, Mizusawa, & Okumura, 2017). One hypothesis for the discrepancy observed between estimated genome sizes may be associated with repeat regions that have not been fully captured in our assembly. The *A. tenebrosa* genome also contained a higher proportion of lineage-specific genes compared with other cnidarian genomes. Previous studies have identified this pattern in species from the superfamily Actinioidea, particularly those genes that encode for peptide toxins (Prentis et al., 2018; Surm et al., 2019). It is shown that there is relatively little overlap of toxin genes among cnidarian species and that a high proportion are restricted to specific lineages (Rachamim et al., 2015; Surm et al., 2019). In addition, many lineage-specific toxins from *A. tenebrosa* have expression restricted to acrorhagi, a novel structure used for envenomation. These data support the hypothesis that novel genes are expressed in novel morphological structures. Evidence in support of this hypothesis in other cnidarian species is equivocal. For example, although *Nematostella*-specific genes comprise a significant proportion of genes expressed in the nematostome, a novel structure only found in this genus, many of these genes were also expressed in tissues common to all sea anemone species (Babonis, Martindale, & Ryan, 2016).

The origin of new genes is considered to be an important source of evolutionary novelty, by providing the substrate upon which natural selection can act. New genes may be formed through multiple processes, ranging from gene duplication through exon shuffling to de novo gene formation (Kaessmann, 2010; McLysaght & Hurst, 2016; Tautz & Domazet-Lošo, 2011). Genes created through these processes produce copies of a gene that are identical to the ancestral sequence or generate genes with novel sequences that are restricted to specific lineages (Capra, Pollard, & Singh, 2010). Here, we have revealed that lineage-specific gene families undergo increased rates of gene duplication compared with gene families shared among cnidarian orders. This suggests that following the formation of new genes in cnidarian taxa, repeated duplication events occur. However, this also suggests that few new genes arise through de

novo gene evolution in cnidarians, as genes generated through this mechanism have been reported to undergo limited gene duplications (Schlötterer, 2015).

Sea anemones, and in particular species from the superfamily Actinioidea, are an important group used to understand the evolution of toxins. Our results support this, with gene families encoding peptide toxins enriched in *A. tenebrosa* relative to sea anemones from other superfamilies. For example, genes involved in venom production (peptide toxins) or delivery (cnidocyte) are associated with the nematocyst GO term, which are significantly over-represented in the gene families restricted to *A. tenebrosa*. This GO term was not enriched for toxin genes restricted to *N. vectensis* or *E. pallida*. This result, however, may be a consequence of ascertainment bias as the majority of toxins characterized in actinarians to date have been identified in the superfamily, Actinioidea (206 of the 236 cnidarian toxins; Prentis et al., 2018). Furthermore, difference in genome assemblies and annotations methods may also contribute to differences observed in gene family evolution among cnidarian genomes.

We propose that the major contributor to the evolution of new genes in cnidarians is through a process of gene duplication. Significant expansions of neurotoxins are observed in *A. tenebrosa*. This was evident from the increased copy number of Pfam domains (ShK and Defensin\_4) which are associated with neurotoxins that modulate potassium ion channels. The Defensin\_4 domain is associated with the sea anemone type 3 (BDS-LIKE) potassium channel toxin family, and both the gene family and protein domain are restricted to Actinioidea (Diochot, Schweitz, Béress, & Lazdunski, 1998). This is supported by evidence of the sea anemone type 3 (BDS-LIKE) potassium channel toxin family identified to be gained in the 947 Actinioidean-specific gene families (Figure 1). Furthermore, sea anemone type 3 (BDS-LIKE) KTx appears to be the most highly duplicated toxin-encoding gene in *A. tenebrosa*.

In this study, we explored the selective pressures acting on orthologs and paralogs in TTL gene families to investigate the adaptive evolution of lineage-specific duplications. We revealed repeated evidence of paralogs evolving at an accelerated rate compared with orthologs. Our findings identified that TTL paralogs often cluster together, suggesting recent duplications undergo accelerated rates of nonsynonymous substitutions, whereas nucleotide variation in ancient duplications is driven by selective forces that limit deleterious mutations (purifying selection). This pattern is supported by the work of Sunagar and Moran (2015) who observed this pattern of divergent selective pressures among ancient and young venomous lineages. The authors suggest that the evidence of diversifying selection acting on younger venomous lineages is driven by recent duplications allowing for the adaptations to an ecological niche. In ancient venomous lineages, however, TTL genes encoding toxins that resulted from ancient duplications events are dominated by purifying selection to limit deleterious mutations. This suggests that toxins in ancient venomous lineages have become specialized to their ecological requirements. Our study supports these findings, and we further suggest that lineage-specific duplications may drive the adaptive evolution of toxins in ancient venomous lineages required to meet their ecological

and life history requirements. This pattern was not conserved for all TTL gene families, however, with MACPF paralogs and *N. vectensis* NaTx paralogs evolving under purifying selection. This may be due to members of both gene families instead evolving via a process consistent with the action of concerted evolution.

Diverse evolutionary trajectories exist following gene duplication including pseudogenization, neofunctionalization, and subfunctionalization. An additional trajectory includes conservation which can be driven through a process of concerted evolution. Concerted evolution is the homogenization of paralogs that results in sequence similarity greater within species compared to between species (Liao, 1999; Nei & Rooney, 2005). This homogenization is typically attributed to gene conversion or unequal-crossing over (Brown, Wensink, & Jordan, 1972; Eickbush & Eickbush, 2007; Szostak & Wu, 1980). Here we observe concerted-like evolution in multiple TTL gene families including sea anemone types 1 and 3 (BDS-LIKE) KTx, NaTx, MACPF, and the nontoxin gene family FP. Whether the concerted-like evolution observed is through lineage-specific duplications or concerted evolution remains elusive.

Concerted evolution of a sea anemone toxin gene family has previously been reported in multiple species (Moran, Genikhovich, et al., 2012; Moran et al., 2008). Nv1, a member of NaTx TTL gene family, is the major adult venom component in *N. vectensis*. Nv1 has evolved via concerted evolution. This is supported by the evidence of Nv1 copies being encoded by a cluster of at least 12 highly conserved sequences (Moran, Genikhovich, et al., 2012; Moran et al., 2008). This is further supported in the NaTx phylogeny we generated in this study, with Nv1 sequences clustering strongly together (Figure 5). From our selection analyses, we could not infer that these highly homogenous Nv1 sequences are evolving under diversifying selection. Divergently, the *N. vectensis* paralogs that escaped this homogenization are inferred to be evolving under diversifying selection, which consists of Nv3-8. Recent experimental evidence supports the adaptive evolution of these paralog escaping the process of concerted evolution (Sachkova et al., 2019). This is evident with Nv4 and Nv5 paralogs expression being mostly restricted to early life stages compared with Nv1, suggesting neofunctionalization or subfunctionalization. The Nv4 and Nv5 paralogs also exhibit divergent activity being highly toxic to fish, compared with Nv1 which has greater activity against arthropods (Sachkova et al., 2019). Indeed, a similar pattern is also observed in *A. viridis* NaTx paralogs with the Av2 copies being highly similar and other copies escaping this homogenization. These escaped paralogs are also inferred to be evolving under diversifying selection. While evidence supports that both Nv1 and Av2 are evolving through a process of gene conversion or unequal-crossing over consistent with concerted evolution (Moran et al., 2008), the escaped paralogs, however, may have resulted from lineage-specific duplications (Sachkova et al., 2019).

Overall, our phylogenetic analyses provide repeated evidence of paralogs clustering closer together than orthologs for multiple TTL gene families in cnidarians. Whether this occurs through a process of concerted evolution (gene conversion or unequal-crossing over) or lineage-specific duplications is unclear, especially given that a

combination of both processes may be occurring in parallel. We propose that concerted evolution is an important process in the evolution of ancient actiniarian venom, occurring in gene families recruited into the venom of at least last common ancestor. Subsequently, lineage-specific duplications allow paralogs to escape the homogenizing process associated with concerted evolution, with selection driving these new duplicates to undergo neofunctionalization or subfunctionalization.

In venomous animals, biochemical and morphological innovations result in phenotypic adaptations, such as toxin peptides and an envenomation system. Although the cnidarian envenomation system is largely conserved across this phylum, our analysis revealed duplication events in gene families enriched in *A. tenebrosa* include many nematocyte-related proteins such as toxin peptides. We propose that the genome sequence of *A. tenebrosa* will aid future research to improve our understanding of Actinioidean innovations involved in venom production and its delivery.

## ACKNOWLEDGMENTS

The authors would like to thank the Evolutionary and Physiological Genomics Lab (ePGL), in particular Chloé A. van der Burg, for their continual help. The authors would like to thank Yehu Moran for his support. The authors would also like to thank QUT Marine group for their help and advice caring for the animals. The authors would like to acknowledge QUT Molecular Genetics Research Facility for the use of their facilities and the Hawkesbury Institute for the Environment for computational resources. The data reported in this paper were generated at the Central Analytical Research Facility operated by the Institute for Future Environments. Computational resources and services used in this work were provided by the High Performance Computing and Research Support Group, Queensland University of Technology, Brisbane, Australia.

## CONFLICT OF INTEREST

We declare there are no conflicts of interest.

## AUTHOR CONTRIBUTIONS

All authors conceived and designed the project. JMS, PJP, and AnP collected organism samples. JMS performed DNA extraction. AIP assembled genome, and ZKS annotated genome. JMS performed comparative genomics and phylogenetic analysis. JMS led the draft of the manuscript with contributions from all authors. All authors read and approved the final version.

## DATA AVAILABILITY STATEMENT

A description and overview of the project are available under the BioProject accession number PRJNA505921. A description of the complete mitochondrion is available through GenBank accession number MK291977.

## ORCID

Joachim M. Surm  <https://orcid.org/0000-0001-6970-029X>

Alexie Papanicolaou  <https://orcid.org/0000-0002-3635-6848>

Peter J. Prentis  <https://orcid.org/0000-0001-6587-8875>

## REFERENCES

- Adachi, K., Miyake, H., Kuramochi, T., Mizusawa, K., & Okumura, S. (2017). Genome size distribution in phylum Cnidaria. *Fisheries Science*, 83(1), 107–112. <https://doi.org/10.1007/s12562-016-1050-4>
- Alieva, N. O., Konzen, K. A., Field, S. F., Meleshkevitch, E. A., Hunt, M. E., Beltran-Ramirez, V., ... Matz, M. V. (2008). Diversity and evolution of coral fluorescent proteins. *PLoS ONE*, 3(7), e2680. <https://doi.org/10.1371/journal.pone.0002680>
- Babonis, L. S., Martindale, M. Q., & Ryan, J. F. (2016). Do novel genes drive morphological novelty? An investigation of the nematosomes in the sea anemone *Nematostella vectensis*. *BMC Evolutionary Biology*, 16(1), 114. <https://doi.org/10.1186/s12862-016-0683-3>
- Baumgarten, S., Simakov, O., Esherrick, L. Y., Liew, Y. J., Lehnert, E. M., Michell, C. T., ... Voolstra, C. R. (2015). The genome of *Aiptasia*, a sea anemone model for coral symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, 112(38), 11893–11898. <https://doi.org/10.1073/pnas.1513318112>
- Beckmann, A., & Özbek, S. (2012). The nematocyst: A molecular map of the cnidarian stinging organelle. *International Journal of Developmental Biology*, 56(6–7–8), 577–582. <https://doi.org/10.1387/ijdb.113472ab>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Brown, D. D., Wensink, P. C., & Jordan, E. (1972). A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: The evolution of tandem genes. *Journal of Molecular Biology*, 63(1), 57–73. [https://doi.org/10.1016/0022-2836\(72\)90521-9](https://doi.org/10.1016/0022-2836(72)90521-9)
- Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I. A., Belmonte, M. K., Lander, E. S., ... Jaffe, D. B. (2008). ALLPATHS: *De novo* assembly of whole-genome shotgun microreads. *Genome Research*, 18(5), 810–820. <https://doi.org/10.1101/gr.7337908>
- Capra, J. A., Pollard, K. S., & Singh, M. (2010). Novel genes exhibit distinct patterns of function acquisition and network integration. *Genome Biology*, 11(12), R127. <https://doi.org/10.1186/gb-2010-11-12-r127>
- Casewell, N. R., Wüster, W., Vonk, F. J., Harrison, R. A., & Fry, B. G. (2013). Complex cocktails: The evolutionary novelty of venoms. *Trends in Ecology & Evolution*, 28(4), 219–229. <https://doi.org/10.1016/j.tree.2012.10.020>
- Castañeda, O., Sotolongo, V., Amor, A. M., Stöcklin, R., Anderson, A. J., Harvey, A. L., ... Karlsson, E. (1995). Characterization of a potassium channel toxin from the Caribbean sea anemone *Stichodactyla helianthus*. *Toxicon*, 33(5), 603–613. [https://doi.org/10.1016/0041-0101\(95\)00013-C](https://doi.org/10.1016/0041-0101(95)00013-C)
- Chapman, J. A., Kirkness, E. F., Simakov, O., Hampson, S. E., Mitros, T., Weinmaier, T., ... Steele, R. E. (2010). The dynamic genome of *Hydra*. *Nature*, 464(7288), 592–596. <https://doi.org/10.1038/nature08830>
- Columbus-Shenkar, Y. Y., Sachkova, M. Y., Macrander, J., Fridrich, A., Modepalli, V., Reitzel, A. M., ... Moran, Y. (2018). Dynamics of venom composition across a complex life cycle. *eLife*, 7, e35014. <https://doi.org/10.7554/eLife.35014>
- Daly, M. (2016). Functional and genetic diversity of toxins in sea anemones. In P. Gopalakrishnakone & A. Malhotra (Eds.), *Evolution of venomous animals and their toxins* (pp. 1–18). Dordrecht, The Netherlands: Springer.
- Daly, M., Crowley, L. M., Larson, P., Rodríguez, E., Saucier, E. H., & Fautin, D. G. (2017). *Anthopleura* and the phylogeny of Actinioidea (Cnidaria: Anthozoa: Actiniaria). *Organisms Diversity & Evolution*, 17(3), 545–564. <https://doi.org/10.1007/s13127-017-0326-6>
- Dawson, N. L., Lewis, T. E., Das, S., Lees, J. G., Lee, D., Ashford, P., ... Sillitoe, I. (2017). CATH: An expanded resource to predict protein function through structure and sequence. *Nucleic Acids Research*, 45(D1), D289–D295. <https://doi.org/10.1093/nar/gkw1098>
- Diochot, S., Schweitz, H., Béress, L., & Lazdunski, M. (1998). Sea anemone peptides with a specific blocking activity against the fast inactivating potassium channel Kv3.4. *Journal of Biological Chemistry*, 273(12), 6744–6749. <https://doi.org/10.1074/jbc.273.12.6744>
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., ... Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15–21. <https://doi.org/10.1093/bioinformatics/bts635>
- Eddy, S. R. (2011). Accelerated profile HMM searches. *PLOS Computational Biology*, 7(10), e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>
- Eickbush, T. H., & Eickbush, D. G. (2007). Finely orchestrated movements: Evolution of the ribosomal RNA genes. *Genetics*, 175(2), 477–485. <https://doi.org/10.1534/genetics.107.071399>
- Ellinghaus, D., Kurtz, S., & Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinformatics*, 9(1), 18. <https://doi.org/10.1186/1471-2105-9-18>
- Farquhar, H. (1898). Preliminary account of some New-Zealand Actiniaria. *Journal of the Linnean Society of London, Zoology*, 56(171), 527–536. <https://doi.org/10.1111/j.1096-3642.1898.tb00409.x>
- Farris, J. S. (1977). Phylogenetic analysis under Dollo's law. *Systematic Biology*, 26(1), 77–88. <https://doi.org/10.1093/sysbio/26.1.77>
- Fautin, D. G. (2009). Structural diversity, systematics, and evolution of cnidae. *Toxicon*, 54(8), 1054–1064. <https://doi.org/10.1016/j.toxicol.2009.02.024>
- Fautin, D. G., & Mariscal, R. N. (1991). Cnidaria: Anthozoa. In F. Harrison & J. Westfall (Eds.), *Microscopic anatomy of invertebrates: Protozoa* (Vol. 2, pp. 267–358). New York, NY: Wiley-Liss.
- Felsenstein, J. (1989). PHYLIP - Phylogeny inference package (version 3.2). *Cladistics*, 5(2), 163–166. <https://doi.org/10.1111/j.1096-0031.1989.tb00562.x>
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., ... Punta, M. (2014). Pfam: The protein families database. *Nucleic Acids Research*, 42(D1), D222–D230. <https://doi.org/10.1093/nar/gkt1223>
- Frazão, B., Vasconcelos, V., & Antunes, A. (2012). Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: An overview. *Marine Drugs*, 10(8), 1812–1851. <https://doi.org/10.3390/md10081812>
- Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Gough, J., Karplus, K., Hughey, R., & Chothia, C. (2001). Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. *Journal of Molecular Biology*, 313(4), 903–919. <https://doi.org/10.1006/jmbi.2001.5080>
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., ... Regev, A. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, 8(8), 1494–1512. <https://doi.org/10.1038/nprot.2013.084>
- Han, Y., & Wessler, S. R. (2010). MITE-Hunter: A program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Research*, 38(22), e199. <https://doi.org/10.1093/nar/gkq862>
- Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., & Stanke, M. (2016). BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, 32(5), 767–769. <https://doi.org/10.1093/bioinformatics/btv661>

- Honma, T., Minagawa, S., Nagai, H., Ishida, M., Nagashima, Y., & Shiomi, K. (2005). Novel peptide toxins from acrorhagi, aggressive organs of the sea anemone *Actinia equina*. *Toxicon*, 46(7), 768–774. <https://doi.org/10.1016/j.toxicon.2005.08.003>
- Ikmi, A., & Gibson, M. C. (2010). Identification and *in vivo* characterization of NvFP-7R, a developmentally regulated red fluorescent protein of *Nematostella vectensis*. *PLoS ONE*, 5(7), e11807. <https://doi.org/10.1371/journal.pone.0011807>
- Jouiaei, M., Sunagar, K., Federman Gross, A., Scheib, H., Alewood, P. F., Moran, Y., & Fry, B. G. (2015). Evolution of an ancient venom: Recognition of a novel family of cnidarian toxins and the common evolutionary origin of sodium and potassium neurotoxins in sea anemone. *Molecular Biology and Evolution*, 32(6), 1598–1610. <https://doi.org/10.1093/molbev/msv050>
- Jouiaei, M., Yanagihara, A. A., Madio, B., Nevalainen, T. J., Alewood, P. F., & Fry, B. G. (2015). Ancient venom systems: A review on cnidaria toxins. *Toxins*, 7(6), 2251–2271. <https://doi.org/10.3390/toxins7062251>
- Jungo, F., & Bairoch, A. (2005). Tox-Prot, the toxin protein annotation program of the Swiss-Prot protein knowledgebase. *Toxicon*, 45(3), 293–301. <https://doi.org/10.1016/j.toxicon.2004.10.018>
- Kaessmann, H. (2010). Origins, evolution, and phenotypic impact of new genes. *Genome Research*, 20(10), 1313–1326. <https://doi.org/10.1101/gr.101386.109>
- Kass-Simon, G., & Scappaticci, A. A. (2002). The behavioral and developmental physiology of nematocysts. *Canadian Journal of Zoology*, 80(10), 1772–1794. <https://doi.org/10.1139/z02-135>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, L., Stoeckert, C. J., & Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, 13(9), 2178–2189. <https://doi.org/10.1101/gr.1224503>
- Liao, D. (1999). Concerted evolution: Molecular mechanism and biological implications. *American Journal of Human Genetics*, 64(1), 24–30. <https://doi.org/10.1086/302221>
- Maček, P., & Lebez, D. (1988). Isolation and characterization of three lethal and hemolytic toxins from the sea anemone *Actinia equina* L. *Toxicon*, 26(5), 441–451. [https://doi.org/10.1016/0041-0101\(88\)90183-3](https://doi.org/10.1016/0041-0101(88)90183-3)
- MacManes, M. D. (2014). On the optimal trimming of high-throughput mRNA sequence data. *Frontiers in Genetics*, 5, 13. <https://doi.org/10.3389/fgene.2014.00013>
- Macrander, J., Brugler, M. R., & Daly, M. (2015). A RNA-seq approach to identify putative toxins from acrorhagi in aggressive and non-aggressive *Anthopleura elegantissima* polyps. *BMC Genomics*, 16, 221. <https://doi.org/10.1186/s12864-015-1417-4>
- Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., ... Bryant, S. H. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Research*, 43(D1), D222–D226. <https://doi.org/10.1093/nar/gku1221>
- McLysaght, A., & Hurst, L. D. (2016). Open questions in the study of *de novo* genes: What, how and why. *Nature Reviews Genetics*, 17(9), 567–578. <https://doi.org/10.1038/nrg.2016.78>
- Minagawa, S., Sugiyama, M., Ishida, M., Nagashima, Y., & Shiomi, K. (2008). Kunitz-type protease inhibitors from acrorhagi of three species of sea anemones. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 150(2), 240–245. <https://doi.org/10.1016/j.cbpb.2008.03.010>
- Mirdita, M., von den Driesch, L., Galiez, C., Martin, M. J., Söding, J., & Steinegger, M. (2017). Uniclust databases of clustered and deeply annotated protein sequences and alignments. *Nucleic Acids Research*, 45(D1), D170–D176. <https://doi.org/10.1093/nar/gkw1081>
- Moran, Y., Genikhovich, G., Gordon, D., Wienkoop, S., Zenkert, C., Özbek, S., ... Gurevitz, M. (2012). Neurotoxin localization to ectodermal gland cells uncovers an alternative mechanism of venom delivery in sea anemones. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732), 1351–1358. <https://doi.org/10.1098/rspb.2011.1731>
- Moran, Y., Praher, D., Schlesinger, A., Ayalon, A., Tal, Y., & Technau, U. (2012). Analysis of soluble protein contents from the nematocysts of a model sea anemone sheds light on venom evolution. *Marine Biotechnology*, 15(3), 329–339. <https://doi.org/10.1007/s10126-012-9491-y>
- Moran, Y., Weinberger, H., Sullivan, J. C., Reitzel, A. M., Finnerty, J. R., & Gurevitz, M. (2008). Concerted evolution of sea anemone neurotoxin genes is revealed through analysis of the *Nematostella vectensis* genome. *Molecular Biology and Evolution*, 25(4), 737–747. <https://doi.org/10.1093/molbev/msn021>
- Nei, M., & Rooney, A. P. (2005). Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics*, 39, 121–152. <https://doi.org/10.1146/annurev.genet.39.073003.112240>
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1), 268–274. <https://doi.org/10.1093/molbev/msu300>
- Norton, R. S., Maček, P., Reid, G. E., & Simpson, R. J. (1992). Relationship between the cytotoxins tenebrosin-C from *Actinia tenebrosa* and equinatoxin II from *Actinia equina*. *Toxicon*, 30(1), 13–23. [https://doi.org/10.1016/0041-0101\(92\)90497-5](https://doi.org/10.1016/0041-0101(92)90497-5)
- O'Hara, E. P., Caldwell, G. S., & Bythell, J. (2018). Equistatin and equinatoxin gene expression is influenced by environmental temperature in the sea anemone *Actinia equina*. *Toxicon*, 153, 12–16. <https://doi.org/10.1016/j.toxicon.2018.08.004>
- Ou, S., & Jiang, N. (2018). LTR\_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiology*, 176(2), 1410–1422. <https://doi.org/10.1104/pp.17.01310>
- Özbek, S. (2010). The cnidarian nematocyst: A miniature extracellular matrix within a secretory vesicle. *Protoplasma*, 248(4), 635–640. <https://doi.org/10.1007/s00709-010-0219-4>
- Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10), 785–786. <https://doi.org/10.1038/nmeth.1701>
- Piriyaopongsa, J., Rutledge, M. T., Patel, S., Borodovsky, M., & Jordan, I. K. (2007). Evaluating the protein coding potential of exonized transposable element sequences. *Biology Direct*, 2, 31. <https://doi.org/10.1186/1745-6150-2-31>
- Prentis, P. J., Pavasovic, A., & Norton, R. S. (2018). Sea anemones: Quiet achievers in the field of peptide toxins. *Toxins*, 10(1), 36. <https://doi.org/10.3390/toxins10010036>
- Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., ... Rokhsar, D. S. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science*, 317(5834), 86–94. <https://doi.org/10.1126/science.1139158>
- Rachamim, T., Morgenstern, D., Aharonovich, D., Brekman, V., Lotan, T., & Sher, D. (2015). The dynamically evolving nematocyst content of an anthozoan, a scyphozoan, and a hydrozoan. *Molecular Biology and Evolution*, 32(3), 740–753. <https://doi.org/10.1093/molbev/msu335>
- Rodríguez, E., Barbeitos, M. S., Brugler, M. R., Crowley, L. M., Grajales, A., Gusmão, L., ... Daly, M. (2014). Hidden among sea anemones: The first comprehensive phylogenetic reconstruction of the order Actiniaria (Cnidaria, Anthozoa, Hexacorallia) reveals a novel group of hexacorals. *PLoS ONE*, 9(5), e96998. <https://doi.org/10.1371/journal.pone.0096998>
- Sachkova, M. Y., Singer, S. A., Macrander, J., Reitzel, A. M., Peigneur, S., Tytgat, J., & Moran, Y. (2019). The birth and death of toxins with distinct functions: A case study in the sea anemone *Nematostella*. *Molecular Biology and Evolution*, 36(9). <https://doi.org/10.1093/molbev/msz132>

- Schlötterer, C. (2015). Genes from scratch – the evolutionary fate of *de novo* genes. *Trends in Genetics*, 31(4), 215–219. <https://doi.org/10.1016/j.tig.2015.02.007>
- Schwaiger, M., Schonauer, A., Rendeiro, A. F., Pribitzer, C., Schauer, A., Gilles, A. F., ... Technau, U. (2014). Evolutionary conservation of the eumetazoan gene regulatory landscape. *Genome Research*, 24(4), 639–650. <https://doi.org/10.1101/gr.162529.113>
- Sebé-Pedrós, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhé, M.-P., Renno, J., ... Marlow, H. (2018). Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-seq. *Cell*, 173(6), 1520.e20–1534.e20. <https://doi.org/10.1016/j.cell.2018.05.019>
- Sherman, C. D. H., Peucker, A. J., & Ayre, D. J. (2007). Do reproductive tactics vary with habitat heterogeneity in the intertidal sea anemone *Actinia tenebrosa*? *Journal of Experimental Marine Biology and Ecology*, 340(2), 259–267. <https://doi.org/10.1016/j.jembe.2006.09.016>
- Shinzato, C., Shoguchi, E., Kawashima, T., Hamada, M., Hisata, K., Tanaka, M., ... Satoh, N. (2011). Using the *Acropora digitifera* genome to understand coral responses to environmental change. *Nature*, 476(7360), 320–323. <https://doi.org/10.1038/nature10249>
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539. <https://doi.org/10.1038/msb.2011.75>
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Smit, A. F. A., & Hubley, R. (2008). *RepeatModeler Open-1.0*. 2008–2015. Retrieved from <http://www.repeatmasker.org>
- Smit, A. F. A., Hubley, R., & Green, P. (2013). *RepeatMasker Open-4.0*. 2013–2015. Retrieved from <http://www.repeatmasker.org>
- Stefanik, D. J., Wolenski, F. S., Friedman, L. E., Gilmore, T. D., & Finnerty, J. R. (2013). Isolation of DNA, RNA and protein from the starlet sea anemone *Nematostella vectensis*. *Nature Protocols*, 8(5), 892–899. <https://doi.org/10.1038/nprot.2012.151>
- Steinegger, M., & Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology*, 35, 1026–1028. <https://doi.org/10.1038/nbt.3988>
- Sunagar, K., Columbus-Shenkar, Y. Y., Fridrich, A., Gutkovich, N., Aharoni, R., & Moran, Y. (2018). Cell type-specific expression profiling unravels the development and evolution of stinging cells in sea anemone. *BMC Biology*, 16(1), 108. <https://doi.org/10.1186/s12915-018-0578-4>
- Sunagar, K., & Moran, Y. (2015). The rise and fall of an evolutionary innovation: Contrasting strategies of venom evolution in ancient and young animals. *PLoS Genetics*, 11(10), e1005596. <https://doi.org/10.1371/journal.pgen.1005596>
- Surm, J. M., Smith, H. L., Madio, B., Undheim, E. A. B., King, G. F., Hamilton, B. R., ... Prentis, P. J. (2019). A process of convergent amplification and tissue-specific expression dominates the evolution of toxin and toxin-like genes in sea anemones. *Molecular Ecology*, 28(9), 2272–2289. <https://doi.org/10.1111/mec.15084>
- Szostak, J. W., & Wu, R. (1980). Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. *Nature*, 284(5755), 426. <https://doi.org/10.1038/284426a0>
- Tautz, D., & Domazet-Lošo, T. (2011). The evolutionary origin of orphan genes. *Nature Reviews Genetics*, 12(10), 692–702. <https://doi.org/10.1038/nrg3053>
- Urbarova, I., Patel, H., Forêt, S., Karlsen, B. O., Jørgensen, T. E., Hall-Spencer, J. M., & Johansen, S. D. (2018). Elucidating the small regulatory RNA repertoire of the Sea Anemone *Anemone viridis* based on whole genome and small RNA sequencing. *Genome Biology and Evolution*, 10(2), 410–426. <https://doi.org/10.1093/gbe/evy003>
- van der Burg, C. A., Prentis, P. J., Surm, J. M., & Pavasovic, A. (2016). Insights into the innate immunome of actinarians using a comparative genomic approach. *BMC Genomics*, 17, 850. <https://doi.org/10.1186/s12864-016-3204-2>
- Wang, X., Liew, Y. J., Li, Y., Zoccola, D., Tambutte, S., & Aranda, M. (2017). Draft genomes of the corallimorpharians *Amplexidiscus fenestrafer* and *Discosoma* sp. *Molecular Ecology Resources*, 17(6), e187–e195. <https://doi.org/10.1111/1755-0998.12680>
- Wang, Y., Coleman-Derr, D., Chen, G., & Gu, Y. Q. (2015). OrthoVenn: A web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Research*, 43(W1), W78–W84. <https://doi.org/10.1093/nar/gkv487>
- Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., ... Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution*, 35(3), 543–548. <https://doi.org/10.1093/molbev/msx319>
- Watts, P. C., Allcock, A. L., Lynch, S. M., & Thorpe, J. P. (2000). An analysis of the nematocysts of the beadlet anemone *Actinia equina* and the green sea anemone *Actinia prasina*. *Journal of the Marine Biological Association of the United Kingdom*, 80(4), 719–724.
- Wilding, C. S., & Weedall, G. D. (2019). Morphotypes of the common beadlet anemone *Actinia equina* (L.) are genetically distinct. *Journal of Experimental Marine Biology and Ecology*, 510, 81–85. <https://doi.org/10.1016/j.jembe.2018.10.001>
- Xu, Z., & Wang, H. (2007). LTR\_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Research*, 35, W265–W268. <https://doi.org/10.1093/nar/gkm286>
- Ye, C., Ji, G., & Liang, C. (2016). *detectMITE*: A novel approach to detect miniature inverted repeat transposable elements in genomes. *Scientific Reports*, 6, 19688. <https://doi.org/10.1038/srep19688>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Surm JM, Stewart ZK, Papanicolaou A, Pavasovic A, Prentis PJ. The draft genome of *Actinia tenebrosa* reveals insights into toxin evolution. *Ecol Evol*. 2019;9:11314–11328. <https://doi.org/10.1002/ece3.5633>