Journal of Venom Research ISSN: 2044-0324 J Venom Res, 2020, Vol 10, 7-12

RESEARCH REPORT

Horizontal transfer of a natterin-like toxin encoding gene within the holobiont of the reef building coral *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia) and across multiple animal linages

Ranko Gacesa[1], Julia Yun-hsuan Hung[2], David G Bourne[2,3] and Paul F Long[1,4,*]

[1]School of Cancer & Pharmaceutical Sciences, Faculty of Life Sciences & Medicine, King's College London, 150 Stamford Street, London SE1 9NH, United Kingdom

[2]College of Science and Engineering, James Cook University, Townsville, Queensland 4811, Australia

[3]Australian Institute of Marine Science, Townsville, Queensland 4810, Australia

[4]Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, Av. Prof Lineu Prestes, 580, B16, 05508-000 São Paulo, SP, Brazil

***Correspondence to:** Paul Long, Email: paul.long@kcl.ac.uk, Tel/Fax: 00 44 (0)20 7848 4842

Received: 18 March 2020 | **Revised:** 25 March 2020 | **Accepted:** 01 April 2020 | **Published:** 17 April 2020

ABSTRACT

Phylogenetic evidence is provided for horizontal transfer of a natterin-like toxin encoding gene from fungi into the genome of the coral *Acropora digitifera*. Sequencing analysis of the coral tissues supported that a fungal taxon predicted to be the most likely gene donor was represented in the coral microbiome. Further bioinformatics data suggested widespread recruitment of the natterin-like gene into venomous terrestrial invertebrates, and repositioning of this gene to non-toxic functions in non-venomous teleost fish.

KEYWORDS: Horizontal gene transfer, Cnidaria, venom

INTRODUCTION

A widely held contention amongst contemporary scientists is that venom innovation is driven by convergent evolution of non-toxin encoding genes that undergo gene duplication and mutation giving rise to toxic functions (Fry et al, 2003). Yet, the genetic origin of these non-toxic progenitors that subsequently become recruited and then evolve venom purposes remains obscure. Non-sexual transmission of genetic material between distantly related species is termed horizontal gene transfer (HGT, also known as lateral gene transfer [LGT]) and is recognised as a pervasive source of genetic novelty in prokaryotes (Koonin et al, 2001) and is widespread between animal associated microorganisms, probably due to increased cell-to-cell contact as a result of close physical proximity (Jeong et al, 2019; Murphy et al, 2019). In contrast, convincing examples of HGT in eukaryotic genomes are rare (Ku and Martin, 2016; Martin, 2017), although genome sequencing is increasingly highlighting prospective HGT events in both animals and plants (Dunning-Hotopp et al, 2007; Leger et al, 2018; Rossoni et al, 2019).

To date, only three examples of possible HGT involvement in venom evolution have been reported and include the toxin sphingomyelinase D (SMase D), which is found almost exclusively in *Loxosceles* and *Sicarius* spider venoms (Binford et al, 2009). Comparison of SMase D sequences have shown a shared C-terminal structural motif suggesting HGT from bacteria to spiders has resulted as a single divergent event in the evolution of this toxin family (Cordes and Binford, 2006). A second example of possible HGT leading to toxin diversification is that of the aerolysin family of pore-forming toxins. Clustering based on pairwise similarity and phylogenetic analysis support HGT from bacteria to

the sea anemone *Nematostella vectensis*, with at least five other recurrent transfer events between distantly related species belonging to different kingdoms of life (Moran et al, 2012). A third example of a likely HGT event is that of the GH19 chitinase (Glycoside hydrolase family 19) gene from a unicellular microsporidia yeast donor into parasitoid wasps. The GH19 chitinase gene was shown to be highly expressed in wasp venom glands and has not been found in any other sequenced animal genome or in any other fungal taxon, strongly reinforcing evidence that HGT is an innovation leading to genetic novelty in venom toxin recruitment in parasitoid wasps (Martinson et al, 2016).

By virtue of the functional importance of venom for prey capture and in defence against predation or deterrence of competitors, acquisition of the venom trait has been a key innovation that underpinned the explosive radiation of many animal species, including inhabitants of marine environments (Casewell et al, 2013). To begin our investigations into toxin diversification in Cnidaria (Weston et al, 2012; Starcevic and Long, 2013), we took as our model marine venomous animal the reef building coral *A. digitifera* Dana 1846 (Cnidaria: Anthozoa: Scleractinia). A total of 55 predicted toxins encoded by the genome of *A. digitifera* could be annotated by homology to known animal toxins, and these could be grouped into 13 clusters with two or more sequences and 19 singlets. To our knowledge, *A. digitifera* has the greatest percentage of toxin encoding gene duplications reported in the genome of any venomous animal to date (Dunlap et al, 2013; Gacesa et al, 2015). Natterins are a family of 5 toxins isolated from the venom of the medically significant *Thalassophryne nattereri* fish, found in the Amazon River and tributaries of north and north-eastern Brazil. Natterins induce oedema, intensify nociceptive pain and cleave human kininogen releasing kallidin (Lys-bradykinin), thus provoking vasodilation (Magalhães et al, 2005; Lopes-Ferreira et al, 2014). A single copy of a gene sequence with homology to natterin-4 was located during the annotation for putative toxin genes encoded by the genome of *A. digitifera* (Gacesa et al, 2015). The genetic origin of this singlet gene was investigated using phylogenetic and experimental approaches to resolve whether the gene had been acquired following an episode of HGT.

MATERIALS AND METHODS

Phylogeny

Putative homologs of natterin-4 were identified by blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins) searches against the NCBI NR protein database, using the BLOSUM 62 substitution matrix. Sequences corresponding to major animal and fungal phyla, with a blastp e-value of less than 1e-5, were used for phylogenetic analysis. Multiple sequence alignments were constructed using ClustalW and MUSCLE algorithms implemented in MEGA X (Kumar et al, 2016), using default settings. A further structure-based alignment was also built using the Expresso procedure from the open server of T-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:expresso). Structural alignment was performed using PDB structures 1W3GA (hemolytic lectin from the mushroom *Laetiporus sulphureus* complexed with two N-acetyllactosamine molecules) and 5MH1A (crystal structure of an IclR homolog from *Microbacterium* sp. strain HM58-2), identified by a blastp search against the PDB implemented in the T-Coffee expresso software. MEGA X software was used to construct maximum likelihood (ML) phylogenetic trees for all three multiple alignments, using the Le_Gascuel_2008 model (Le and Gascuel, 2008) with gamma-distributed rate of evolution (5 discrete categories). The model of evolution was selected based on AICc and BIC criteria scores determined by model selection framework implemented in MEGA X. Bootstrapping test with 500 bootstrap replicates was used to determine the confidence of phylogenetic trees.

To assess the robustness of the phylogenetic analysis, each of the multiple alignments was also used to reconstruct phylogenetic trees using Bayesian analysis implemented in MrBayes software (version 3.2.7, http://mrbayes.sourceforge.net). Bayesian analysis was performed with a mixed amino acid model (prset aamodelpr=mixed), a gamma distributed rate of evolution with 4 discrete categories (lset rates=gamma), and prior probabilities set to default values. A Markov chain Monte Carlo simulation was executed for 20,000,000 generations using 4 runs and 5 chains per run. The simulation convergence was confirmed by reviewing standard deviation of split frequencies and potential scale reduction values. The trees from MrBayes 3.2.7 were displayed by Figtree (http://tree.bio.ed.ac.uk/software/figtree).

In addition to phylogenetic analysis, homology between natterin-like sequences was assessed using the CLuster ANalysis of Sequences (CLANS) tool to construct a Fruchterman-Reingold graph-layout from all-against-all blastp searches (Frickey and Lupas, 2004). The graph was constructed from pair-wise blastp results with e-value below 1.0e-10, and the network-based clustering implemented in CLANS was performed on the resulting graph. Sequences used for phylogenetic reconstruction, multiple alignments, results of model selection, detailed results of phylogenetic reconstruction, and MrBayes batch file are included as Supplementary Figures S1-S4.

Fungal community profiling

Specimens of *A. digitifera* were collected from the northeast of Orpheus Island (18°34’19.5»S, 146°29’31.4»E) in the central Great Barrier Reef. The coral nubbins were air-blasted (80psi) with sterilised artificial seawater to remove coral tissue and associated microbes from the coral skeleton. Genomic DNA was extracted with a DNeasy PowerBiofilm Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. The fungal ITS region was amplified with forward ITS1F 5’-CTTGGTCATTTAGAGGAAGTAA-3’ and reverse ITS4 5’TCCTCCGCTTATTATTGATATGC-3’ primer pairs (White et al, 1990; Gardes et al, 1993). PCR was carried out in 20 µl reactions that included 6.8 µl DNA template, 10 µl AmpliTaq Gold 360 Mastermix (Applied Biosystems, CA, USA) and 1.6 µl of each 10 µM fungal ITS primer. PCR conditions were: initial denaturation at 95°C for 10min; 35 cycles of 30sec denaturation at 95°C, 30sec annealing at 55°C and 60sec amplification at 72°C; followed by a 7min final extension and hold at 4°C. Amplification products were separated using agarose gel electrophoresis and PCR products of the expected size of 850bp were sequenced

(Ramaciotti Centre for Genomics, University of New South Wales, Sydney, Australia) using the BASE protocol (Bissett et al, 2016) on an Illumina MiSeq, with 300bp paired-end sequencing targeting the fungal ITS region (ITS1F-ITS4). Bioinformatics processing of the Illumina datasets was performed using the QIIME 2 pipeline (https://qiime2.org, Caporaso et al, 2010). The quality of the reads was visually assessed using FastQC, the reads were then filtered and trimmed to 270 bp using DADA 2 single-end reads (Callahan et al, 2016). Good quality sequences were then filtered with the UNITE database (UNITE Community 2017) with exclude-seqs and filter-feature command lines to remove *Symbiodinium* and host coral contaminating ITS sequences. A taxonomy classification was also executed using the UNITE database developer file of the QIIME 2 package. Fungal sequences were also aligned using ClustalW with other fungal ITS sequences extracted from the NCBI GenBank. Phylogenetic analysis was carried out with both Neighbor-Joining (NJ) and ML algorithms using MEGA 7.

RESULTS AND DISCUSSION

The gene sequence with homology to natterin-4 had the accession number adi_1.18989 in the open access annotated database of the *A. digitifera* genome ('ZoophyteBase', Dunlap et al, 2013), and is also deposited in NCBI protein database (accession XP_015754265) and the UniProt database (UniParc accession UPI00077AABEE). Extra caution is advised when assigning possible HGT events in Cnidaria because of the potential for contamination of the sequencing template by a significant number of microorganisms harboured by these animals (Starcevic et al, 2008). The natterin-4 encoding gene was unlikely to be of symbiont origin because it was located at position 1,014,831 to 1,017,287 on an assembly scaffold with the NCBI sequence accession number NW_015441109.1. This scaffold is large (1,039,139 bp) and the predicted taxonomy of genes that flanked 100 kb upstream and 21852 bp downstream of the natterin homologue could be assigned using a BLASTn search of the NR database with a 0.1 e-value cutoff. Annotation of these flanking regions was ascribed based on the highest scoring hits from the NR database and, the resultant 503 hits corresponded to stony coral sequences only, with no bacterial or fungal hits detected. This implied that it was unlikely the natterin-4 encoding homologue was a chimera that had arisen as a sequence assembly artefact, or that the gene was sequenced owing to fungal or bacterial contamination of the original genome sequencing template.

Accurate detection and classification of HGT events is a major computational and conceptual challenge (Ravenhall et al, 2015). Herein we used a phylogeny-based approach for gene-species tree reconstruction to better detect HGT events across diverse prokaryotic and eukaryotic species that co-exist in close symbiotic relationships. When the *A. digitifera* natterin homologue was used as a query to search the NCBI NR database by BLASTp, the highest-scoring hits were surprisingly similar to sequences of fungal and bacterial origin, rather than to animals. This was unexpected because natterin was first described and, the toxic action thereafter, extensively studied in teleost fish (Xue et al, 2012). Homologues of natterin-4 were uncovered in seemingly very divergent animal lineages including coral *Orbicella faveolata*, butterflies, flatworms as well as several other teleost fish. The amino acid sequence of the natterin-4 homologue from *A. digitifera* could be roughly divided into N-terminal and C-terminal portions. The C-terminal portion of *A. digitifera* and *O. faveolata* natterins aligned closely to the corresponding portion of the *natterin-4* gene sequence of *T. nattereri*, which is also believed to be the region of the protein most closely associated with toxic biological actions (Magalhães et al, 2005). *Natterin*-like sequences from non-coral invertebrates (insects and flatworms), however, showed high divergence in alignment to C-terminal region of fish and coral natterins. A Bayesian method (MrBayes) for tree construction used the Metropolis–Hastings Markov Chain Monte Carlo (MCMC) simulation (Ronquist et al, 2012) with ClustalW2, MUSCLE and T-Coffee multiple sequence alignments as inputs. All simulations gave good convergence, with average standard deviation of split frequencies <0.01 and a convergence value (Potential Scale Reduction Factor) approaching 1.0 (Gelman and Rubin, 1992). Analysis of the phylogenetic tree (Figure 1) positioned the natterin-like sequences into five distinct clades corresponding to fungi, bacteria, corals, fish and non-coral invertebrates.

Examination of the phylogenetic tree placed the natterin-4 homologue of *A. digitifera* in closer evolutionary proximity to microorganisms rather than clustering with natterin-4 sequences of other animals (Figure 1). This positioning of corals close to microorganisms was observed in both Bayesian and maximum likelihood based phylogenetic reconstruction (Figure 1), and was consistent between trees reconstructed using ClustalW and MUSCLE multiple alignment algorithms (Figure 1 and Figures S1-S4), but not with the tree reconstructed from the T-Coffee Expresso alignment, which grouped coral natterin-like sequences with invertebrate (Figures S5 and S6). This discrepancy between structure based alignment (Expresso) and more traditional sequence-based alignments (ClustalW and MUSCLE) was possibly due to lack of appropriate protein structures for natterin-like sequences. To identify potential bias in multiple alignment, we reviewed alignments of input sequences and structural temples used by Expresso (PDB structures 1W3GA and 5H1A), and found that alignment of templates to input sequences covers less than 30% of input sequences. In addition, templates were found to be highly dissimilar (pairwise BLASTp e-value 1.3 and 17% coverate). Thus, we discarded Expresso alignments as biased due to the lack of appropriate templates, and instead used alignment-free clustering approaches to test the validity of phylogenetic trees based on MUSCLE and ClustalW alignments. The CLANS method identified two clusters in natterin-like sequences: coral and fungi sequences formed a cluster, as did fish sequences, while invertebrates and bacteria did not make clear clusters. The coral-fungi cluster was found to be well connected to fishes, while both clusters showed loose association with invertebrates (Figure 2). Given that corals are metazoans, clustering of corals with fungi would seem counterintuitive to the established tree of life and, thus, provides endorsement for a possible HGT event from microorganisms to *A. digitifera*.

Fungal targeted ITS barcode sequences were acquired from tissues removed from *A. digitifera* specimens also collected

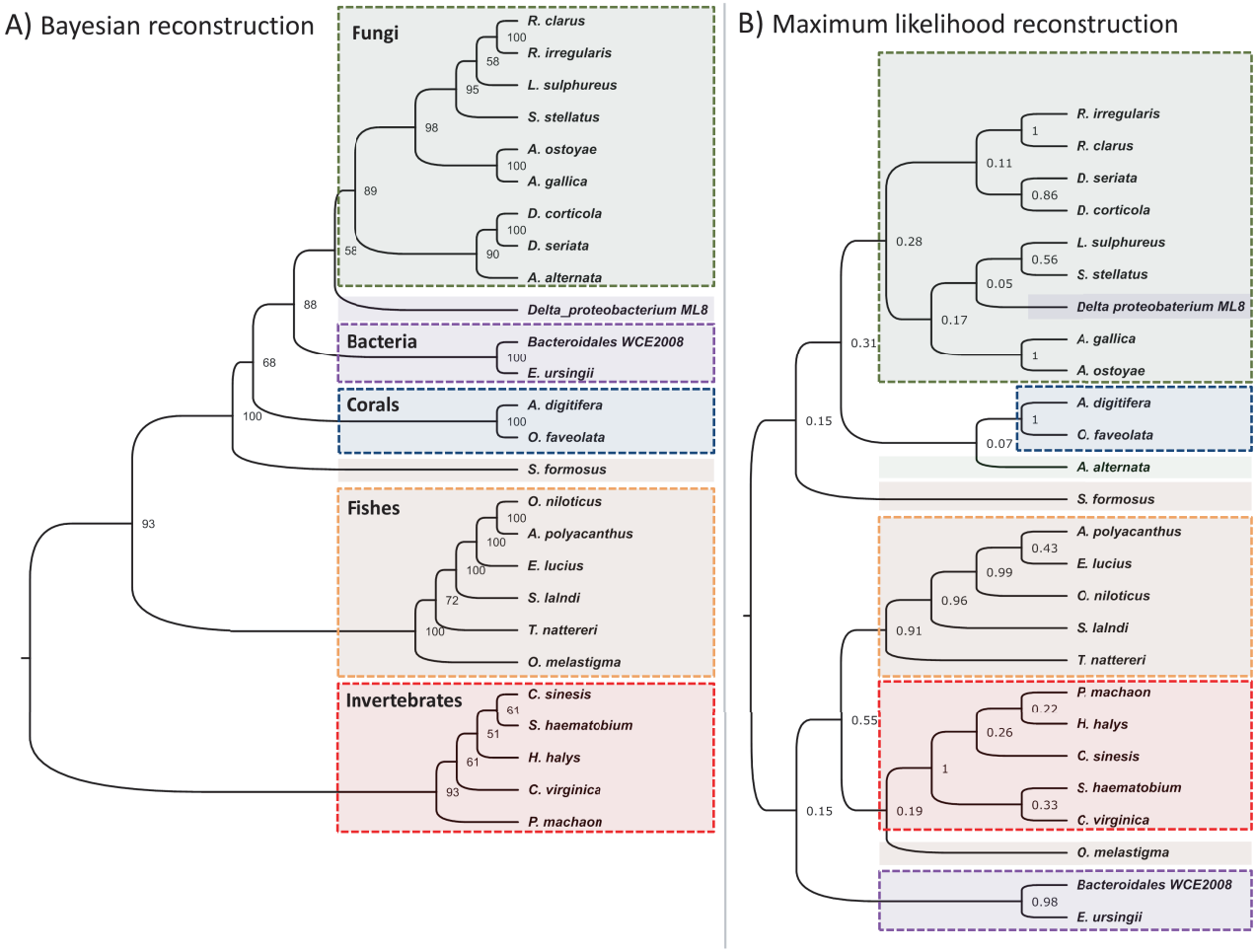

Figure 1. Phylogeny of natterin-4 homologues from different kingdoms of life. **A** shows phylogenetic tree reconstructed using MrBayes tool for Bayesian phylogeny, while **B** shows maximum likelihood (ML) phylogenetic tree generated using MEGA X. Fungal sequences are highlighted in green; Bacteria in purple; Cnidaria in blue; Fishes in orange; and invertebrate animals in red.

on the Great Barrier Reef. The sequences clustered with corresponding ascomycete and basidiomycete ITS sequences by both NJ and ML approaches. An abundant fungal_sv_5 sequence clustered using a naïve Bayes classifier with a high matching homology confidence score of 0.98 to a species of the genus *Phoma*. A genome sequence for *Phoma* is not presently available to mine for natterin-like homologues and, no natterin-like gene sequences from *Phoma* species have been deposited in any publically available database. Nonetheless, species of the genera *Phoma* and *Epicoccum* are closely related based on phylogenetic analysis (Arenal et al, 2008; Fávaro et al, 2011), and so provides anecdotal evidence that a fungal donor of the natterin-4 gene is possible from the fungal microbiota of *A. digitifera.*

Cnidaria are arguably the most basal of the extant Metazoa to be venomous for which fossil records exist dating to at least the Neoproterozoic times, ~650 million years ago, long before the Cambrian radiation and the evolutionary split with bilateria (Van Iten et al, 2014; Starcevic and Long, 2013). Despite the early divergence of cnidarians, the toxin diversity of cnidarian venoms is just as complex as the toxin complement of major groups of venomous bilaterians including marine snakes, perhaps reflecting their long evolutionary history and diverse lifestyles (Jaimes-Becerra et al, 2019). Phylogenetic analysis in Figure 1 indicated that the *A. digitifera* natterin-4 homologue formed a cluster with microorganisms rather than animals, which implied that a fungal to coral HGT event potentially occurred prior to cnidaria – bilateria split. The *A. digitifera natterin-4* homologue grouped closer to fish sequences rather than to the cluster composed of arthropod sequences, which again is contrary to passage down standard evolutionary lines following vertical transmission, but consistent with both coral and ancestral fish acquiring the gene possibly via a second HGT event. Interestingly, the arthropod sequences could be attributed to known venomous arthropod taxa, however, with the exception of *T. nattereri* this was not the case for the remaining representatives of the fish cluster. We suggest that following a HGT evident into fish, the natterin-4 homologue was passed across fish taxa by divergent evolution but in doing so, toxic function was lost and has now been repurposed into a non-toxic alternative physiological function in modern day fish. Evidence in support of a non-toxic function for natterin-like proteins found in other non-toxic animals comes from a very recent study of the Pacific oyster *Crossostrea gigas* (Unno et al, 2016) which described a novel type of Mannose-specific Lectin designated CGL1, named natterin-3 in the Uniprot database (entry: K1QRB6). This protein shared 26 % amino acid sequence identity (43 %

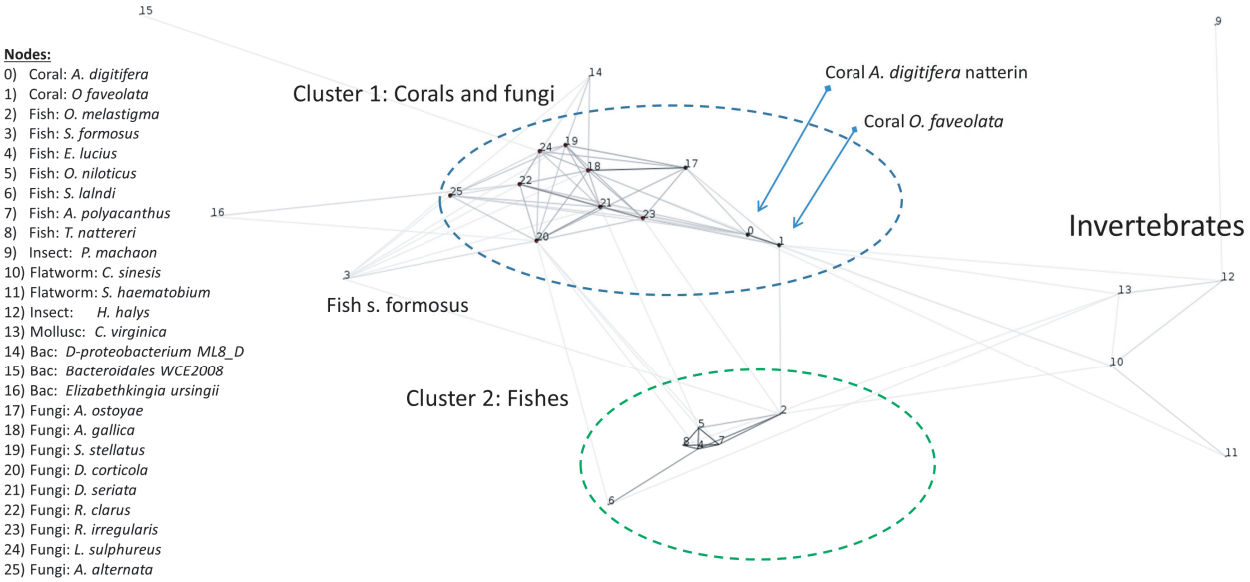

Figure 2. Blast-based clustering of natterin-4 sequences. Figure shows CLANS-generated network of all-against-all BLAST searches of natterin-like sequences from animals, fungi and bacteria. Network based clustering identified two main clusters in data: Cluster 1 (encircled in blue) is composed of coral sequence (nodes 0,1) and fungal sequences (nodes 17-25), while Cluster 2 (encircled in green) comprises fish sequences. Invertebrate sequences (nodes 9-13) were found to be connected, while bacterial sequences (nodes 14-16) do not connect to other bacteria. Connections indicate pair-wise blast results with e-value below 1.0e-10.

sequence similarity) when we used a pair-wise alignment to compare with the toxic C-terminal region of natterin-4 of *T. nattereri*. The CGL1 amino acid sequence has been shown to share between 30-39 % identity with the other proteins from non-venomous animals, for example, DM9 protein from *Drosophila melanogaster* and with 170 other *Diptera* proteins of unknown function (Ponting et al, 2001; Unno et al, 2016).

CONCLUSION

The classic 'birth and death' model of venom toxin diversification through gene duplication, followed by evolution of a toxin function by one of the copies has been modified to include the possibility that gene evolution is a dynamic process, and can include the possibility of reverse gene recruitment to non-toxic functions. We offer data that shows potential phylogenetic evidence for HGT of a natterin-4 homologue from fungi to the coral *A. digitifera*. The role of HGT of toxin encoding genes in the evolution of animal venoms maybe contrary to conventional thinking on the evolution of toxin diversification, but perhaps warrants more profound experimental consideration, particularly for toxin evolution in early diverging metazoans such as exemplified by the Cnidaria (Jaimes-Becerra et al, 2019). The recent plethora of publicly available cnidarian genome sequences would greatly facilitate such investigations (Kim et al, 2019; Ohdera et al, 2019; Surm et al, 2019).

ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council UK grant number G82144A.

COMPETING INTERESTS

None declared.

REFERENCES

Arenal F, Platas G, Monte E and Peláez, F. 2008. ITS sequencing support for *Epicoccum nigrum* and *Phoma epicoccina* being the same biological species. Mycol Res, 104, 301-303.

Binford GJ, Bodner MR, Cordes MHJ, et al. 2009. Molecular evolution, functional variation, and proposed nomenclature of the gene family that includes sphingomyelinase D in sicariid spider venoms. Mol Biol Evol, 26, 547-566.

Bissett A, Fitzgerald A, Meintjes T, et al. 2016. Introducing BASE: the biomes of Australian soil environments soil microbial diversity database. GigaScience, 5, 21.

Caporaso J G, Kuczynski J, Stombaugh J, et al. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods, 7, 335-336.

Callahan BJ, McMurdie, PJ, Rosen MJ, Han AW, Jahnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods, 13, 581-583.

Casewell NR, Wüster W, Vonk FJ, Harrison RA and Fry BG. 2013. Complex cocktails: The evolutionary novelty of venoms. Trends Ecol Evol, 28, 219-229.

Cordes MHJ and Binford GJ. 2006. Lateral gene transfer of a dermonecrotic toxin between spiders and bacteria. Bioinformatics, 22, 264-268.

Dunlap WC, Starcevic A, Baranasic D, Diminic J, Zucko J, Gacesa R, et al. 2013. KEGG orthology-based annotation of the predicted proteome of *Acropora digitifera*: ZoophyteBase - an open access and searchable database of a coral genome. BMC Genomics, 14, 509.

Dunning-Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, Torres MCM, et al. 2007. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science, 317, 1753-1756.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39, 783-791.

Fávaro LC, de Melo FL, Aguilar-Vildoso CI and Araújo, W. L. 2011. Polyphasic analysis of intraspecific diversity in *Epicoccum nigrum* warrants reclassification into separate species. PLoS ONE, 6, e14828.

Frickey T and Lupas AN. 2004. CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics, 20:3702-3704.

Fry BG, Wüster W, Kini RM, et al. 2003. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. J Mol Evol, 57, 110-29.

Gacesa R, Chung R, Dunn SR, et al. 2015. Gene duplications are extensive and contribute significantly to the toxic proteome of nematocysts isolated from *Acropora digitifera* (Cnidaria: Anthozoa: Scleractinia). BMC Genomics, 16, 774.

Gelman A and Rubin DB. 1992. Inference from iterative simulation using multiple sequences. Stat Sci, 7, 457-511.

Jaimes-Becerra A, Gacesa R, Doonan LB, et al. 2019. "Beyond Primary Sequence"-Proteomic Data Reveal Complex Toxins in Cnidarian Venoms. Integr Comp Biol, 59, 777-785.

Jeong H, Arif B, Caetano-Anollés G, Kim KM and Nasir A. 2019. Horizontal gene transfer in human-associated microorganisms inferred by phylogenetic reconstruction and reconciliation. Sci Rep, 9, 5953.

Kim HM, Weber JA, Lee N, et al. 2019.The genome of the giant Nomura's jellyfish sheds light on the early evolution of active predation. BMC Biol, 17, 28.

Koonin EV, Makarova KS and Aravind L. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. Annu Rev Microbiol, 55, 709-742.

Le SQ and Gascuel O. 2008. An improved general amino acid replacement matrix. Mol Biol Evol, 25, 1307-1320.

Leger MM, Eme L, Stairs CW and Rodgers AJ. 2018. Demystifying eukaryote lateral gene transfer (Response to Martin 2017 DOI: 10.1002/bies.201700115). BioEssays, 40, 1700242.

Lopes-Ferreira M, Grund LZ and Lima C. 2014. *Thalassophryne nattereri* fish venom: from the envenoming to the understanding of the immune system. J Venom Anim Toxins Incl Trop Dis, 20, 35.

Magalhães GS, Lopes-Ferreira M, Junqueira-de-Azevedo IL, et al. 2005. Natterins, a new class of proteins with kininogenase activity characterized from *Thalassophryne nattereri* fish venom. Biochimie, 87, 687-699.

Martin WF. 2017. Too Much Eukaryote LGT. BioEssays, 37, 1700115.

Martinson EO, Martinson VG, Edwards R, Mrinalini WJH and Werren JH. 2016. Laterally transferred gene recruited as a venom in parasitoid wasps. Mol Biol Evol, 33, 1042-1052.

Moran Y, Fredman D, Szczesny P, Grynberg M and Technau U. 2012. Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. Mol Biol Evol, 29, 2223-2230.

Murphy CL, Youssef NH, Hanafy RA, et al. 2019. Horizontal Gene Transfer as an Indispensable Driver for Evolution of Neocallimastigomycota into a Distinct Gut-Dwelling Fungal Lineage. Appl Environ Microbiol, 85, e00988-19.

Nei M, Gu X and Sitnikova T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA, 94, 7799-7806.

Ohdera A, Ames CL, Dikow RB, et al. 2019. Box, stalked, and upside-down? Draft genomes from diverse jellyfish (Cnidaria, Acraspeda) lineages: *Alatina alata* (Cubozoa), *Calvadosia cruxmelitensis* (Staurozoa), and *Cassiopea xamachana* (Scyphozoa). Gigascience, 8, giz069.

Ponting CP, Mott R, Bork P and Copley RR. 2001. Novel protein domains and repeats in *Drosophila melanogaster*: insights into structure, function, and evolution. Genome Res, 11, 1996-2008.

Ravenhall M, Škunca N, Lassalle F and Dessimoz C. 2015. Inferring horizontal gene transfer. PLoS Comp Biol, 11, e1004095.

Ronquist F, Teslenko M, van der Mark P, et al. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol, 61, 539-542.

Rossoni AW, Price DC, Seger M, et al. 2019. The genomes of polyextremophilic cyanidiales contain 1% horizontally transferred genes with diverse adaptive functions. eLife, 8, e45017.

Saitou N and Nei M. 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol, 4, 406-425.

Starcevic A, Akthar S, Dunlap WC, et al. 2008. Enzymes of the shikimic acid pathway encoded in the genome of a basal metazoan, *Nematostella vectensis*, have microbial origins. Proc Natl Acad Sci, USA 105, 2533-2537.

Starcevic A and Long PF. 2013. Diversification of animal venom peptides-were jellyfish amongst the first combinatorial chemists? ChemBioChem, 14, 1407-1409.

Surm JM, Stewart ZK, Papanicolaou A, Pavasovic A, Prentis PJ. 2019.The draft genome of *Actinia tenebrosa* reveals insights into toxin evolution. Ecol Evol, 9, 11314-11328.

Unno H, Matsuyama K, Tsuji Y, et al. 2016. Identification, Characterization, and X-ray crystallographic analysis of a novel type of mannose-specific lectin CGL1 from the Pacific Oyster *Crassostrea gigas*. Sci Rep, 6, 29135.

Van Iten H, Marques AC, Leme JDM, Pacheco MLF and Simões MG. 2014. Origin and early diversification of the phylum Cnidaria Verrill: Major developments in the analysis of the taxon's Proterozoic-Cambrian history. Palaeontology, 57, 677-690.

Weston AJ, Dunlap WC, Shick JM, et al. 2012. A Profile of an Endosymbiont-enriched Fraction of the Coral *Stylophora pistillata* Reveals Proteins Relevant to Microbial-Host Interactions. Mol Cell Proteomics, 11, M111.015487.

Whelan S and Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol, 18, 691–699.

Xue Z, Liu X, Pang Y, et al. 2012. Characterization, phylogenetic analysis and cDNA cloning of natterin-like gene from the blood of lamprey, *Lampetra japonica*. Immunol Letts, 148, 1-10.

Zuckerkandl E and Pauling L. 1965. Evolutionary divergence and convergence in proteins. In: Bryson V and Vogel HJ (Eds) *Evolving Genes and Proteins*, Academic Press, New York, pp 97-166.