

Title	Comparative analyses of glycerotoxin expression unveil a novel structural organization of the bloodworm venom system
Authors	Richter, S; Helm, C; Meunier, FA; Hering, L; Campbell, LI; Drukewitz, SH; Undheim, EAB; Jenner, RA; Schiavo, G; Bleidorn, C
Description	© The Author(s). 2017 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License ( <a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a> ), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver ( <a href="http://creativecommons.org/publicdomain/zero/1.0/">http://creativecommons.org/publicdomain/zero/1.0/</a> ) applies to the data made available in this article, unless otherwise stated. The attached file is the published version of the article.
Date Submitted	2017-05-04

RESEARCH ARTICLE

Open Access



# Comparative analyses of glycerotoxin expression unveil a novel structural organization of the bloodworm venom system

Sandy Richter<sup>1,2,3\*</sup>, Conrad Helm<sup>4</sup>, Frederic A. Meunier<sup>5</sup>, Lars Hering<sup>6</sup>, Lahcen I. Campbell<sup>3,7</sup>, Stephan H. Drukewitz<sup>1</sup>, Eivind A. B. Undheim<sup>8</sup>, Ronald A. Jenner<sup>3</sup>, Giampietro Schiavo<sup>9</sup> and Christoph Bleidorn<sup>1,2,10\*</sup>

## Abstract

**Background:** We present the first molecular characterization of glycerotoxin (GLTx), a potent neurotoxin found in the venom of the bloodworm *Glycera tridactyla* (Glyceridae, Annelida). Within the animal kingdom, GLTx shows a unique mode of action as it can specifically up-regulate the activity of Ca<sub>v</sub>2.2 channels (N-type) in a reversible manner. The lack of sequence information has so far hampered a detailed understanding of its mode of action.

**Results:** Our analyses reveal three ~3.8 kb GLTx full-length transcripts, show that GLTx represents a multigene family, and suggest it functions as a dimer. An integrative approach using transcriptomics, quantitative real-time PCR, *in situ* hybridization, and immunocytochemistry shows that GLTx is highly expressed exclusively in four pharyngeal lobes, a previously unrecognized part of the venom apparatus.

**Conclusions:** Our results overturn a century old textbook view on the glycerid venom system, suggesting that it is anatomically and functionally much more complex than previously thought. The herein presented GLTx sequence information constitutes an important step towards the establishment of GLTx as a versatile tool to understand the mechanism of synaptic function, as well as the mode of action of this novel neurotoxin.

**Keywords:** Annelida, Glyceridae, Calcium channel, Neurotoxin, GLTx, Venom system

## Background

Complex venoms and associated venom systems have independently evolved in a broad phylogenetic range of animals where they play a diversity of biological roles including defense, competition and predation [1–3]. Despite their similar usage, highly diverse venom systems with remarkable variation evolved not only between different venomous clades but also within single clades [1]. Potent neurotoxins, which act as modulators on a variety of ion channels, are a conspicuous component of many venom cocktails with proven pharmacological and therapeutic potential [4]. One specific venom peptide,  $\omega$ -conotoxin GVIA, from marine snails of the genus

*Conus*, is a widely used pharmacological tool in neuroscience because of its ability to block Ca<sub>v</sub>2.2 (N-type) calcium channels [5, 6]. Another blocker of this channel,  $\omega$ -conotoxin MVIIA, isolated from *Conus magus*, was the first FDA-approved drug for intractable chronic pain [7]. Calcium channels regulate the permeability of cell membranes to calcium ions (Ca<sup>2+</sup>) in a voltage-controlled manner, and they function as key transducers for the intracellular flow of calcium [4, 8]. Ca<sup>2+</sup> is in turn of critical importance for the regulation of many biological processes in eukaryotic cells. In neurons, Ca<sup>2+</sup> is instrumental for the transmission of nerve signals and synaptic activity [9] as its entry leads to fusion of synaptic vesicles and subsequent neurotransmitter release into the synaptic cleft [10]. Calcium channel blockers have been isolated from various animal venoms such as snakes, spiders and mollusks [2]. In contrast, there are

\* Correspondence: sr.sandyrichter@gmail.com; christoph.bleidorn@gmail.com

<sup>1</sup>Institute of Biology - Molecular Evolution and Systematics of Animals, University of Leipzig, Talstraße 33, D-04103 Leipzig, Germany  
Full list of author information is available at the end of the article

very few accounts of calcium channel activators within animal venoms. A voltage-gated calcium channel ( $Ca_v$ ) agonist has been described from centipede venom ( $\omega$ -SLPTX-Ssm1a) in the genus *Scolopendra*, although its mode of action and target specificity remain unknown [11, 12]. Agonists of  $Ca_v$  sub-type 2.2 (N-type) are particularly rare [2], with glycerotoxin, which was purified from a marine annelid of the genus *Glycera*, being the best example [13].

Glycerotoxin-producing polychaetes belong to the taxon Glyceridae Grube, 1850 (Annelida), and are commonly known as bloodworms. These venomous annelids show a broad worldwide distribution and are easily recognizable by an eversible pharynx that possesses four cross-arranged teeth consisting of a hook-shaped jaw and an aileron (supporting structure), each of which is associated with a putative venom gland [14–21]. The putative venom glands are each surrounded by massive layers of musculature [22] whose contraction plays an essential role in the process of envenomation [17]. These glands are further connected by special ducts to the teeth that exhibit a series of pores through which the venom is delivered [14, 17, 23]. Furthermore, biomineralization with the copper-based biomineral atacamite [ $Cu_2(OH)_3Cl$ ] enhances the stiffness and hardness of the jaws and makes them remarkably resistant to abrasion. Beyond its mechanical function, it is speculated that copper may play a role in the activation of venom during injection [24].

The existence of a venom apparatus strongly suggested the presence and utilization of venom in Glyceridae. Further insights were gained through proteomic studies on purified fractions of the venom gland cocktail of *Glycera tridactyla* (formerly *G. convoluta*). Michel, Keil [25] recovered low and high molecular weight toxins, and further demonstrated a paralytic function of the latter. Subsequent studies revealed the ability of a high-molecular weight component to reversibly increase spontaneous neurotransmitter release [26, 27]. This neurotoxic activity of *G. tridactyla* venom correlated with the presence of a 300–320 kDa glycoprotein [13, 22, 28], known as glycerotoxin (GLTx). This neurotoxin was found to target  $Ca_v2.2$  channels (N-type), which are expressed in the presynaptic plasma membrane, and causes an increased  $Ca^{2+}$  influx at resting potential [13]. By its ability to specifically up-regulate the activity of  $Ca_v2.2$  channels, GLTx follows a so far unique mode of action. As a result, this neurotoxin dramatically increases neurotransmitter release in a variety of preparations [13, 29], which makes it a versatile tool to analyze the physiological mechanisms affecting neuroexocytosis. Furthermore, GLTx is able to up-regulate the process of presynaptic vesicle recycling at the frog neuromuscular junction and therefore has been pivotal to secure a detailed understanding of the

mechanism of bulk endocytosis at nerve terminals [30]. However, whereas the functional properties of glycerotoxin are well-described, missing sequence data and poor anatomical characterization of the venom apparatus have hampered understanding of the mode of envenomation and actual mode of action on  $Ca^{2+}$  channels.

In our study, we performed the first molecular characterization of glycerotoxin in the bloodworm species *G. tridactyla*. We followed a multidisciplinary approach combining transcriptomic and proteomic analyses, qPCR experiments, *in situ* hybridization, antibody staining and bioinformatic analyses. Our studies elucidated three full-length transcripts of GLTx, and provided new insights into its mechanisms of action and evolution. Moreover, our integrated approach led to the discovery of previously unrecognized pharyngeal toxin-producing structures, demonstrating that the glycerid venom apparatus is anatomically and functionally more complex than previously thought. This unexpected complexity provides a step change in our understanding of the structure of the venom system in glycerid annelids, breaking a century old consensus in the field. Furthermore, the first cloning and sequence analysis of GLTx paved the way for more in-depth analyses of this unique neurotoxin capable of stimulating neuronal communication.

## Results

### Characterization of glycerotoxin (GLTx)

Edman sequencing of purified glycerotoxin (GLTx) yielded a series of short amino acid sequences varying in length between 7–18 amino acids (Additional file 1). Using them as reference sequences in BLAST-searches against transcriptome data led to the identification of GLTx in *G. tridactyla*. In the transcriptome library of the pharyngeal lobes (SR23, single specimen), we were able to recover three full-length transcripts of GLTx. In contrast, the putative venom glands (SR21, single specimen and SR22, pooled multiple specimen) exhibited only a few fragments of the GLTx full-length transcript, whilst in the body tissues (SR25, single specimen and SR26, pooled multiple specimen) we were not able to identify any GLTx transcripts.

Cloning of the nearly full-length GLTx gene (length ~3600 bp) amplified from cDNA (using SR23, contig 5772 as reference) and subsequent primer walking revealed the nucleotide sequence of glycerotoxin (Additional file 2). The cloned sequences are concordant in length to the recovered GLTx full-length transcript from the transcriptome assembly (SR23, contig 5772). The highly similar clones 5A and 7 each harbor a 12 bp insertion which is missing in clone 6A. Translations of obtained GLTx sequences contain an unexpected stop codon at the 5'-end of clone 5A

(Additional file 3). Two further GLTx full-length transcripts (SR23, contig 4317 and contig 4318) are similar to each other but remarkably different from contig 5772. They presumably represent a different GLTx paralog (Additional file 3).

Without a signal peptide, the translated GLTx full-length transcript (SR23, contig 5772) has a length of 1257 amino acids, and the two full-length transcripts (SR23, contig 4317 and contig 4318) have a length of 1256 amino acids. According to a size estimation performed in CLC Main Workbench, the molecular weight of the proteins deduced from the three GLTx full-length transcripts is around 140 kDa each. SDS-PAGE of *G. tri-dactyla* venom followed by in-gel digestion and identification by tandem mass spectrometry concordantly revealed a molecular weight of the complete GLTx polypeptide chain of around 150 kDa (Additional file 4: Figure S1 and Additional file 5). The 320 kDa band which correlates with the activity of GLTx could not be recovered.

In all analyzed GLTx transcripts, protein domain searches yielded a calcium-binding EGF domain, two WSC domains (cell wall integrity and stress response component), and a CCP domain (complement control protein, also known as short consensus repeats SCRs or sushi repeats), all of which are found at the N-terminal end of the protein (Fig. 1a). The computationally determined full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) each have a signal peptide (Fig. 1a). However, no known protein domains could be identified within the C-terminal end of GLTx. Our data reveals that GLTx is a unique neurotoxin with 80% of its sequence displaying a completely unknown domain organization. BLAST-searches of GLTx clones and GLTx full-length transcripts in NCBI GenBank yielded hits to an uncharacterized protein in *Branchiostoma*, and a collectin-12 sequence in *Exaiptasia pallida* (Cnidaria, KXJ16027.1). Moreover, sequence matches were found in expressed sequence tags (ESTs) of the annelids *Myzostoma cirriferum* (FN428144.1) and *Pomatoceros lamarckii* (GR311097.1). A maximum likelihood phylogenetic analysis on 52 clones shows the presence of at least three paralogs, named GLTx paralog 1, paralog 2, and paralog 3 (Fig. 1b and Additional file 6).

PCR experiments using genomic DNA revealed at least four introns (Additional file 4: Figure S2 and Additional file 7). We focused mainly on the intron-exon-structure at the 3'-end of the gene, especially regions adjacent to the exon 3 analyzed in the context of qPCR studies and *in situ* hybridization experiments. The intron-exon-structure at the 5'-end (exon 1) remains unknown.

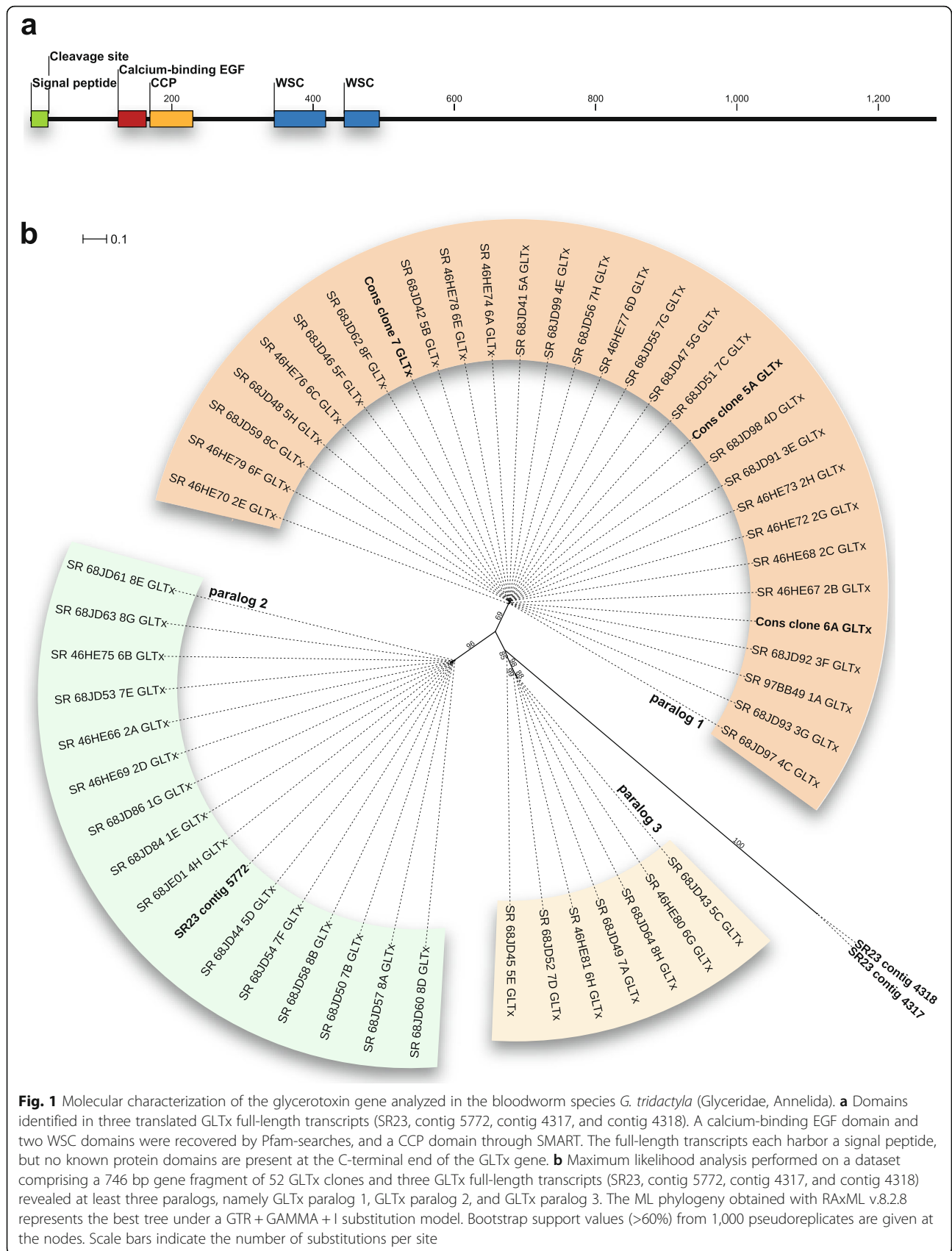
### GLTx localization within the glycerid venom apparatus

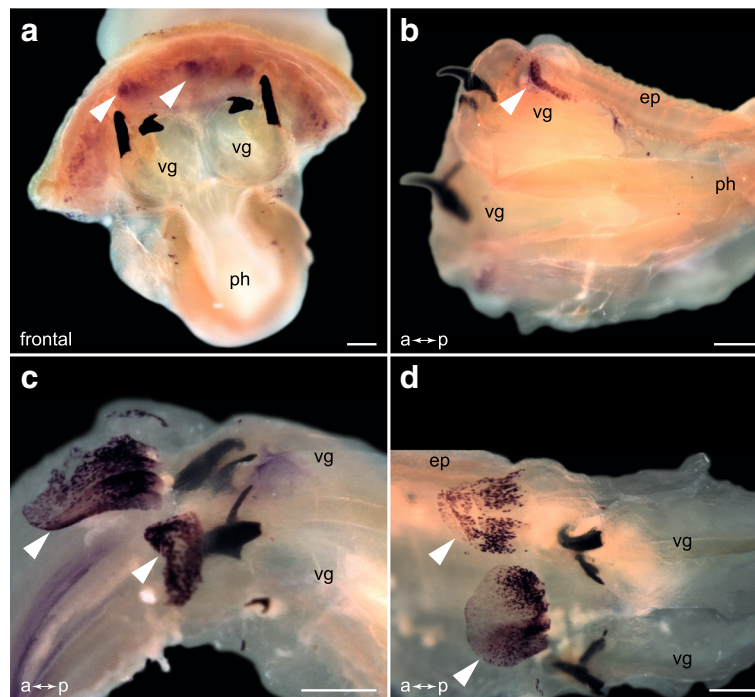
We next aimed to identify the site of toxin expression. We performed *in situ* hybridization experiments with digoxigenin-labeled RNA probes on everted glycerid pharynges, which were dissected into two halves. Prominent GLTx expression is detected in lobate structures located near the base of the teeth, attached to the wall of the proboscis (Fig. 2a,b). Expression signal is restricted to these clearly defined pharyngeal lobes (Fig. 2c,d), and to tissue located at the base of the teeth (Fig. 4b). No other distinct expression was visible in the putative venom glands or elsewhere in the pharynx (Fig. 2). The basal parts of the lobes show strong GLTx expression, whereas the distal parts exhibit a fainter signal (Fig. 2c,d).

To identify pharynx components presumably responsible for GLTx storage, we performed further anti-GLTx antibody staining on everted glycerid pharynx samples also equally dissected into two halves. Distinct anti-GLTx-immunoreactivity (GLTx-IR) is visible in the clearly defined pharyngeal lobes (Fig. 3a), and is most prominent at the base and appears fainter in distal parts of the lobe (Fig. 3b–d). Moreover, fluorescence *in situ* hybridization (FISH) coupled with antibody staining against GLTx reveals that neurotoxin expression and storage is restricted to the same set of cells within the lobe (Fig. 4), indicating that these cells express and secrete GLTx. The double-staining approach performed on an entire inverted pharynx revealed altogether four pharyngeal lobes, each one associated with a single tooth. Additional GLTx-IR staining outside the pharyngeal lobes was not detectable when analyzing pharynx samples that were not embedded (Fig. 3a and Fig. 4c). Notably, analyses on paraffin embedded cross sections of the putative venom glands revealed a prominent GLTx-IR signal inside the lumen (Additional file 4: Figure S3). These results agree with previous work in which GLTx was identified through activity tests in venom fractions extracted from putative venom glands [13, 22, 28]. In what extent the putative venom glands are involved in the storage of GLTx as well as whether there are ducts connecting the pharyngeal lobes and putative venom glands remain to be further investigated. However, antibody staining against GLTx unveils a network of duct-like structures (which we herein refer to as ducts) leading from the lobes to the associated tooth (Fig. 3c,d). It therefore seems that both pharyngeal lobes and putative venom glands are directly connected to the teeth. GLTx-IR also showed that GLTx is released through a series of pores on the teeth (Fig. 5 and Additional file 8).

To analyze whether the lobes of the glycerid pharynx could be part of the nervous system and innervated by prominent nerves and muscle fibers, we carried out







**Fig. 2** Expression of GLTx in bisected pharynges of adult *G. tridactyla*. Stereomicrographs. Arrows indicate *in situ* hybridization signal. **a–b** Frontal and lateral view on a glycerid pharynx cut into two halves. GLTx expression is restricted to lobate structures that are attached to the wall of the proboscis near the base of the teeth, and is absent from any other pharyngeal tissues. **c–d** GLTx expression occurs in clearly defined pharyngeal lobes. ep, epithelium; ph, pharynx; vg, putative venom gland. Scale bars: 100  $\mu$ m (**a–d**)

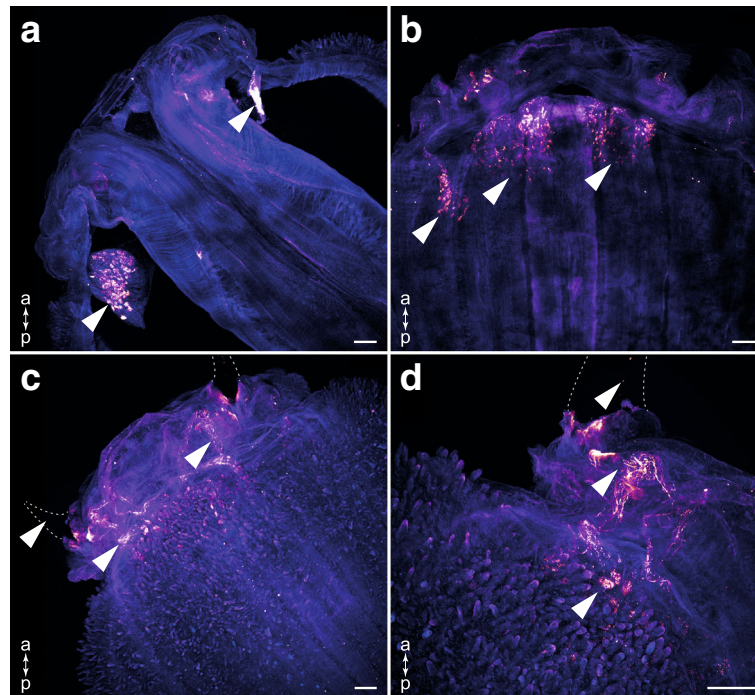
antibody staining using the widespread neurotransmitter serotonin (5-HT) and labeled the surrounding musculature (F-actin fibers) with phalloidin. Whereas the 5-HT staining revealed a dense meshwork of nerves and somata within the entire pharynx region (Additional file 4: Figure S4), the lobes themselves exhibit only faint 5-HT-IR. Nevertheless, the anti-serotonin staining shows that the lobes are neural innervated, but that prominent somata clusters or neuropils are absent. Phalloidin staining shows dense muscle bundles within the entire pharynx, whereas F-actin labelling is almost lacking within the lobe (Additional file 4: Figure S4).

#### Expression levels of GLTx

Transcriptome libraries were used to ascertain the GLTx expression level in putative venom glands (SR21 + SR22), pharyngeal lobes (SR23 + SR24) and body (SR25 + SR26). The highest relative number of mapped reads (212–457 matched reads per million reads) was present in the lobes. Substantially lower numbers (0.47–0.78 matched reads per million reads) were identified in the putative venom glands (Table 1). A comparison of the normalized number of mapped reads (normalized in reference to the total number of filtered reads, Table 1 and Additional file 9) between putative venom glands and pharyngeal lobes indicates that the GLTx expression level, based on the

full-length transcript (SR23, contig 5772), is around 960 times higher in the lobes. In contrast, the expression of GLTx transcripts is only around 300 times higher in the lobes than in the putative venom glands for the other two GLTx full-length transcripts (SR23, contig 4317 and contig 4318). Since the body samples yielded no GLTx transcripts (maximal 5 reads) in transcriptome libraries and GLTx clones (Additional file 9), this tissue was used as calibrator sample in quantitative real-time PCR (qPCR) experiments.

Analyzing 10 specimens as biological samples with paralog-unspecific GLTx primers spanning an intron-exon-border (GLTx-3' and GLTx-5') revealed a higher GLTx expression in both the lobes and the putative venom glands, compared to the body samples, with the highest expression occurring in the lobes (Fig. 6a). A more detailed pattern was found in qPCR experiments using paralog-specific GLTx primers. We recovered GLTx paralog 1 as the most highly expressed paralog, followed by GLTx paralog 2 and paralog 3 in putative venom glands and lobes in comparison to the body samples (Fig. 6a). GLTx expression levels in lobes and putative venom glands are significantly different from the body, as well as from each other (Fig. 6 and Additional file 10b,c). Consistent with our RNAseq results (SR23, contig 5772 and GLTx clones, Table 1), a direct comparison between



**Fig. 3** Confocal maximum projections of everted *G. tridactyla* pharynges cut into two halves. Anti-GLTx staining (glow-mode) and phalloidin-rhodamine counterstaining (blue). Arrows indicate GLTx-IR staining. **a** Distinct GLTx-IR staining occurs in clearly defined pharyngeal lobes. Note there is no additional staining inside other pharyngeal tissues. **b** Anti-GLTx staining revealed a radial color pattern that is most prominent at the base and appears faint in apical parts of the lobes. **c-d** A network of duct-like structures (which we refer to as ducts) connects and transports the GLTx from the pharyngeal lobes to the teeth. Through a series of pores the venom is delivered. Scale bars: 100  $\mu$ m (**a-d**)

lobes and putative venom glands (using the latter as calibrator sample) shows that the relative expression level of GLTx is around 500–1300 times higher in the lobes (Fig. 6b). Furthermore, GLTx expression within the pharyngeal lobes and putative venom glands seems to be coordinated, as the most highly expressed paralog in the lobes was also the most highly expressed paralog in the putative venom glands (Additional file 4: Figure S5 and Additional file 10d,e).

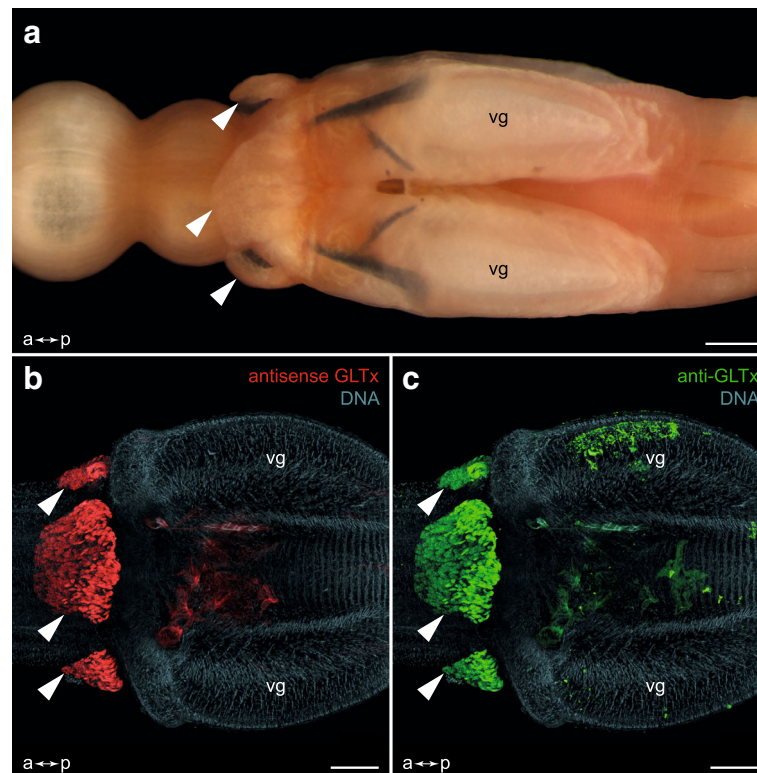
## Discussion

### GLTx is an unusual neurotoxin with novel functional organization

In this study we revealed the full coding sequence of glycerotoxin for the first time. Three computationally recovered GLTx full-length transcripts each harbor a signal peptide at the 5'-end (Fig. 1a), which represents a typical feature for secreted toxins [1]. The identification of several paralogs (Fig. 1b) further indicates the existence of a GLTx multigene family, which is another common feature of toxin genes [1]. GLTx constitutes an uncharacterized protein family with 80% of the GLTx gene displaying an unknown domain organization. However, sequence similarity observed in an uncharacterized protein in *Branchiostoma*, a collectin-12 sequence in Cnidaria, and two annelid ESTs outside Glyceridae suggest that the GLTx family may have evolved from an

existing gene family. Indeed, such a scenario has been proposed to be a general model for the evolution of venom toxins [1, 31], where gene duplication creates novel paralogs that can evolve new functions via sub- or neo-functionalization. However, further phylogenetic analyses and broader sampling of GLTx within Glyceridae is necessary to confirm this hypothesis, and to elucidate the details of the evolutionary origin of GLTx. Moreover, the GLTx gene shows an intron-exon-structure (Additional file 4: Figure S2), which makes it extremely unlikely that this neurotoxin is encoded and expressed by symbiotic bacteria. Our mass spectrometry results further suggest that the molecular weight of the complete GLTx polypeptide chain (around 150 kDa, Additional file 4: Figure S1) is around half the weight previously reported in the literature (300–320 kDa) [13, 22, 28]. Crucially, only the large molecular weight form is active [13], suggesting that GLTx may function as a dimer. Interestingly, the GLTx polypeptide chain has a similar size as the monomer of  $\alpha$ -latrotoxin (LTX), the vertebrate-specific pore-forming neurotoxin isolated from the black widow spider venom (genus *Latrodectus*). LTX is also an excitatory neurotoxin that increases neurotransmitter release [32]. Unlike GLTx, the effects of LTX are irreversible and cause the complete depletion of pre-existent synaptic vesicles [32–34]. Another excitatory



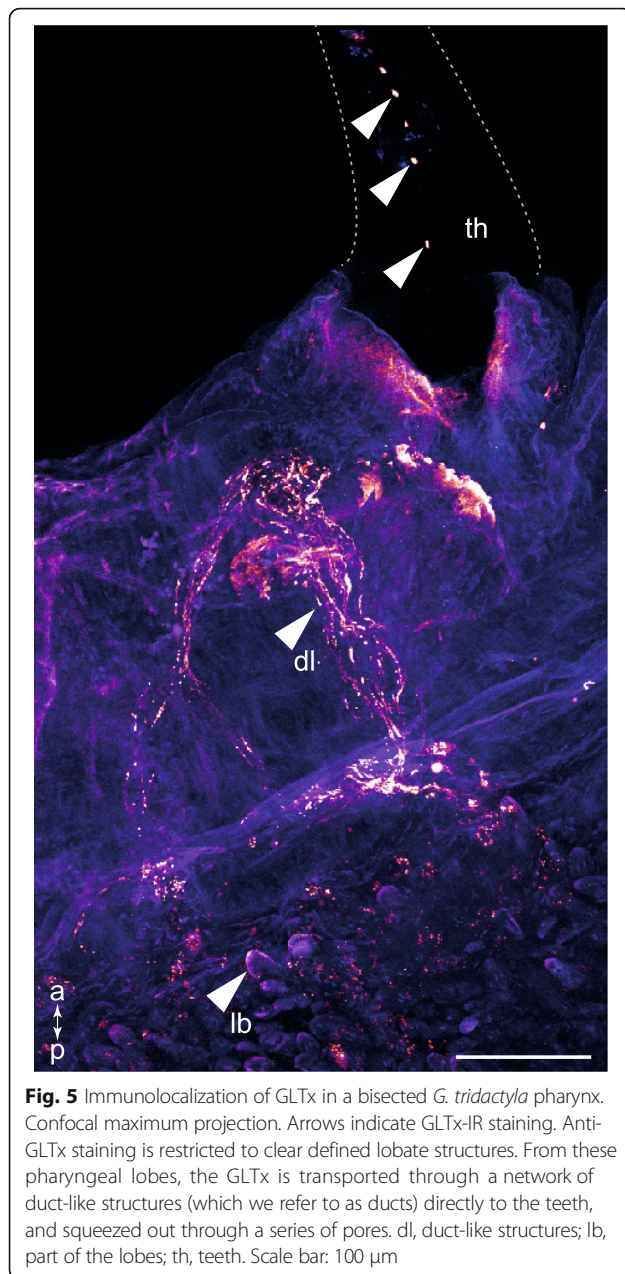


**Fig. 4** Confocal maximum projections of an inverted pharynx of *G. tridactyla* analyzed as total. Fluorescence *in situ* hybridization (FISH) coupled with antibody staining against GLTx, and TO-PRO®-3 iodide counterstaining. **a** Overview of an inverted glycerid pharynx comprising four cross arranged putative venom glands each connected to a tooth, and four corresponding pharyngeal lobes (three of them marked by an arrow). **b** Fluorescence *in situ* hybridization (FISH) revealed a clear GLTx signal in the lobes (marked by arrows) and tissue at the base of the teeth. Note that there is no distinct staining visible in the putative venom glands. **c** Distinct GLTx-IR staining (marked by arrows) is solely restricted to the pharyngeal lobes. Note that there is no staining signal inside the putative venom glands or additional pharynx tissues. vg, putative venom gland. Scale bars: 100  $\mu$ m (**a-c**)

neurotoxin is trachynilysin isolated from the venom of the stonefish *Synanceia trachynis*. This neurotoxin may as well form pores by insertion into the cell membrane, and leads to selective depletion of synaptic vesicles in an irreversible manner [35–37].

A unique feature of GLTx is its ability to reversibly up-regulate the activity of presynaptic  $\text{Ca}_v2.2$  channels [13]. Although a detailed structural functional characterization of the glycoprotein and its mode of action are not yet available, the GLTx sequence data presented in this manuscript (Additional files 2 and 3) enable us to speculate about putative mechanisms. The calcium-binding EGF domain of GLTx (Fig. 1a) may allow the binding of the neurotoxin to extracellular recognition sites on the presynaptic plasma membrane as EGF-like domains are supposed to favor protein-protein interactions [38]. The requirement of  $\text{Ca}^{2+}$  for GLTx activity is in agreement with the observation that GLTx functions in a calcium-dependent manner [26]. The GLTx transcript further encodes a CCP domain (complement control protein, Fig. 1a), also known as sushi domain or short consensus repeat (SCR). CCP modules are involved in specific

protein–protein and protein–carbohydrate interactions and are found within regulators of complement activation (RCA) [39, 40]. One of the homologous protein classes that belong to the RCA family is factor H [41], which is a glycoprotein regulator entirely composed of CCP domains [42]. Through its ability to recognize polyanionic markers on host cell surfaces, factor H can interact with host cell membranes or self-surfaces [42]. In venoms, CCP containing proteins have been reported from the venom cocktail of parasitic wasps of the genus *Leptopilina* [43, 44], but functional assays are to our knowledge not yet available. The GLTx transcripts further encode two WSC domains (cell wall integrity and stress response component, Fig. 1a), which represent putative carbohydrate binding domains. The human plasma membrane protein polycystin 1, which also contains a WSC domain, is suggested to function as a mechanosensor regulating proliferation, adhesion and differentiation [45, 46]. In yeast, this domain is found in regulators of cell wall integrity and the stress response pathway [47, 48]. A fungal  $\beta$ -1,3-exoglucanase that hydrolyzes laminarin to glucose monomers also contains tandem WSC domains, like GLTx [49]. The sequence



information presented here constitutes an important step towards understanding the GLTx mode of action and illuminates its value as a neurobiological tool.

#### The glycerid venom apparatus is a complex system

Our study substantially changes our current understanding of the glycerid venom system. For over a century the glycerid venom apparatus was assumed to be constructed of four putative venom glands that are each connected through a duct to a tooth [14–20]. However, our *in situ* and fluorescence *in situ* hybridization experiments revealed a clear GLTx expression signal restricted to four pharyngeal lobes and an area at the base of the

teeth (Fig. 2 and Fig. 4b). Ehlers [15] (using the name “Lappen”) and Gravier [50] (using the name “membrane quadrilobée”) thought these lobate structures were part of the nervous system, a view that is not supported by our immunohistochemical studies (Additional file 4: Figure S4). Oppenheimer [51] already doubted that the lobes are exclusively part of the nervous system as she already distinguished different cell types within the lobe. Raphaël [52] denied that the lobes were part of the nervous system and rather proposed that the lobes function in the fixation and excretion of hemoglobin. Later studies of Michel [20, 53] (using the name “languettes”) again accepted that the basal parts of the pharyngeal lobes are part of the nervous system, and suggested further the presence of glandular cells thought to secrete a proteolytic enzyme having a digestive function. However, the four pharyngeal lobes have never been recognized to be part of the glycerid venom system.

Whereas the pharyngeal lobes show an obvious staining signal, we were not able to detect any GLTx *in situ* signal inside the putative venom glands (Fig. 2). Moreover, comparative transcriptomics and qPCR experiments carried out on three tissue types of *G. tridactyla*, namely putative venom glands, pharyngeal lobes and posterior body wall, show that the relative GLTx expression is 500–1,300 times higher in the lobes than in the putative venom glands (Fig. 6b and Table 1). Furthermore, antibody staining against GLTx highlighted a network of ducts that connects lobes and teeth (Fig. 3c,d and Fig. 5 and Additional file 8). Taken together, these results strongly support the conclusion that the lobes are the main site for neurotoxin expression, and that the neurotoxin might be transferred from the pharyngeal lobes directly to the teeth where it is injected into prey through a series of pores [14, 17, 23]. Our milking of glycerid venom confirms that there is a direct link between putative venom glands and teeth, a connection that may be independent of the link between pharyngeal lobes and teeth. A network of canals between putative venom glands and pharyngeal lobes could not be identified even though both tissue types show a correlated GLTx expression profile (Additional file 4: Figure S5). This is especially interesting as GLTx antibody staining on paraffin embedded pharynx clearly recovered the neurotoxin inside the lumen of the putative venom glands (Additional file 4: Figure S3). Yet, the GLTx protein detected in the lumen of the putative venom glands may have been produced in the pharyngeal lobes, as there was no GLTx *in situ* signal detectable inside the putative venom glands (Fig. 2). Furthermore, our investigation cannot exclude the possibility that the low GLTx expression signal inside the putative venom glands (qPCR studies and transcriptome analyses, Fig. 6 and Table 1) comes from the ducts connecting the lobes and



**Table 1** Comparison of GLTx expression in different tissues (pharyngeal lobes versus pvg, putative venom glands) of *G. tridactyla* based on RNAseq data. Filtered Illumina reads were mapped against three GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) and three full-length GLTx clones (Cons\_clone\_5A, 6A, and 7). Note that the Fold Change for each contig equals the relative number of mapped reads from the pharyngeal lobes library divided by the relative number of mapped reads from the library of the putative venom glands (see also Additional file 9)

Contig	Length [bp]	Lobes Mapped Reads <sup>a</sup>	PVG Mapped Reads <sup>a</sup>	Lobes vs. PVG Fold Change, normalized
Cons_clone_5A	3605	454.87	0.78	579.72
Cons_clone_6A	3598	453.01	0.78	577.35
Cons_clone_7	3607	444.99	0.78	567.12
SR23_contig_5772	3846	456.91	0.47	963.83
SR23_contig_4317	3841	212.89	0.70	302.87
SR23_contig_4318	3841	212.32	0.70	302.05

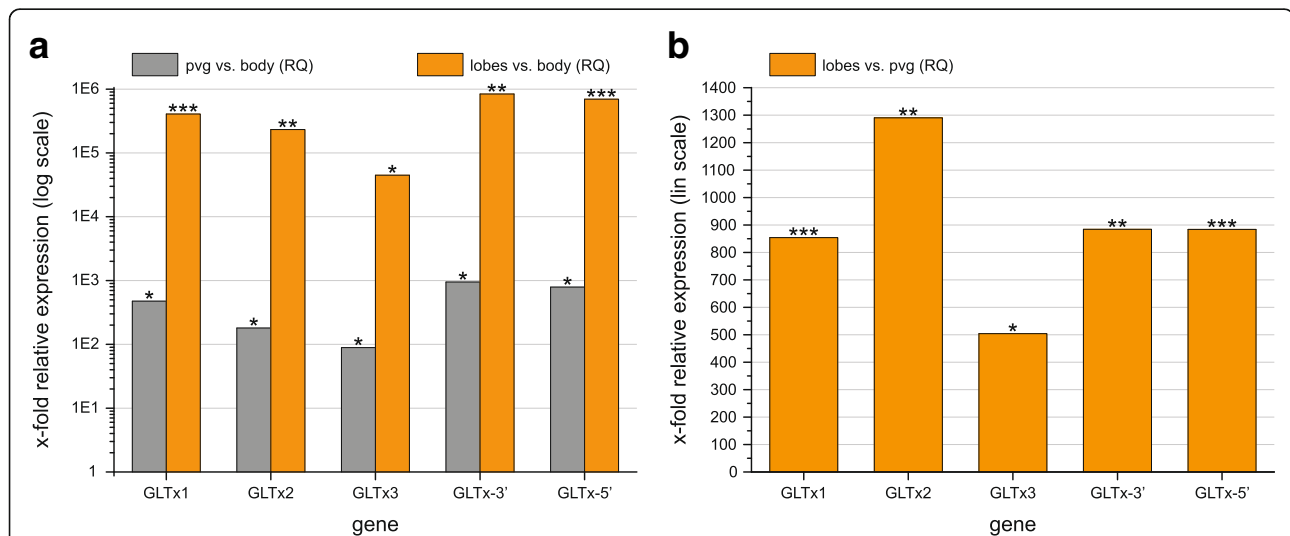
<sup>a</sup>number of mapped reads per million reads of the respective RNAseq library

teeth. However, a recent transcriptome study on venom gland tissue of *G. tridactyla* revealed a complex mixture of putative toxin transcripts [54], which suggests a glandular function of the putative glands alongside its function as a venom storage site. Whether or not the pharyngeal lobes and putative venom glands may be involved in the differential expression of different venom toxins needs further investigation. Our results clearly show that the functional morphology of the glycerid venom system is more complex than hitherto thought.

Compartmentalization of toxin production has been reported from different venomous taxa [55–57]. Within the large protostome clade Lophotrochozoa, complex venom systems comprising several structural subunits

have also been described. The cephalopod venom apparatus comprises two pairs of histologically different venom glands, named posterior and anterior venom glands [58]. Whereas analyses on the posterior venom glands in cephalopods revealed toxins that convergently evolved in other venomous animals [59–61], the role of the anterior glands remains poorly investigated. These glands are considered as mucus secreting organs [58] but their contribution to the cephalopod venom cocktail remains unclear, even though it was recently shown that they may also express some toxin transcripts [62].

In Glyceridae, we show distinct GLTx expression patterns in pharyngeal lobes and putative venom glands (Fig. 6 and Table 1). In this respect, the glycerid venom



**Fig. 6** Quantitative real-time PCR (qPCR) expression levels (shown as Fold Change, RQ) between biological groups (putative venom glands (pvg), pharyngeal lobes, and posterior body wall) of five analyzed GLTx transcripts (GLTx paralogs 1–3, and adjacent gene regions GLTx-3' and GLTx-5'; for details see Material and methods section “Quantitative real-time PCR”) in *G. tridactyla* ( $n = 10$ ; \*\*\* $p \leq 0.001$ ; \*\* $p \leq 0.01$ ; \* $p \leq 0.05$ ). **a** Relative GLTx expression (logarithmic scale) in putative venom glands (grey) and pharyngeal lobes (orange) in comparison to the GLTx expression signal exhibited by the body tissue (RQ = 1). Relative GLTx expression in the pharyngeal lobes and putative venom glands is significantly different from the expression signal in the body tissue. **b** Relative GLTx expression (linear scale) within the pharyngeal lobes in comparison to the putative venom glands (RQ = 1). Relative GLTx expression is significantly different between both putative venom glands and pharyngeal lobes

apparatus resembles to a degree the venom system of carnivorous cone snails. Cone snails are able to produce two types of venom in distinct parts of the venom duct [63]. A defensive toxin cocktail containing paralytic peptides and neurotoxins from the proximal part of the duct, and a less complex predatory venom cocktail from the distal part of the duct [55]. The cone snails are able to secrete these venom types selectively depending on whether predatory or defensive stimuli are received. Since the GLTx expressing lobes are innervated by the nervous system (Additional file 4: Figure S4), it is possible that bloodworms are capable of a rapid stimulus-evoked (defensive or predatory) secretion of GLTx. It is also possible that neurotoxin-rich secretions are only selectively added to the venom mix synthesized and stored in the putative venom glands. Concordantly, *Glyceria alba* is supposed to deplete venom glands incompletely during the first bite which evokes discoordination rather than death and paralysis of the prey, whereas quantitatively more venom seems to be delivered after a firm grip [64]. The remarkable variability in GLTx expression levels revealed by the qPCR experiments of 10 studied specimens (Additional file 4: Figure S5) may indicate that these specimens are in distinct phases of venom replenishment. These results highlight the possibility that Glyceridae are able to meter their venom stocks.

Further research is necessary to test if the neurotoxin GLTx is unique in being differentially expressed in glycerid putative venom glands and pharyngeal lobes, or whether the expression of other toxins also corresponds to this anatomical differentiation of the venom apparatus. To address this question, comparative transcriptomic and proteomic analyses of other venom toxins as well as detailed histological studies of the pharyngeal structures are required. The recent transcriptomic study of von Reumont et al. [54] focused exclusively on glycerid venom gland tissue, hence their results need to be reassessed in view of the role of the pharyngeal lobes in the synthesis of venom toxins reported here. Our current results clearly demonstrate that multidisciplinary analyses are invaluable for understanding the glycerid venom apparatus in particular, and venom systems in general.

## Conclusions

In this work, we report the full sequence of glycerotoxin (GLTx), a neurotoxin known to act specifically as a  $Ca_v2.2$  agonist. GLTx represents a toxin family comprising at least three different paralogs with uncertain evolutionary origin. Moreover, our data show that GLTx likely functions as a dimer with the subunits being held together by non-covalent bonds. GLTx transcripts are expressed in two locations in the glycerid venom apparatus, the putative venom glands and pharyngeal lobes, a

previously unrecognized component of the venom system. GLTx protein is restricted to the pharyngeal lobes and to the lumen of the putative venom glands. Furthermore, GLTx is expressed 500 to 1,300 times higher in the pharyngeal lobes than in the putative venom glands. Our results overturn more than a century of textbook consensus, suggesting that a fundamental revision of our understanding on the functional organization of the venom system in bloodworms is urgently needed.

## Methods

### Protein studies

#### Protein sequencing and characterization

For its initial characterization, GLTx was purified on an 8% SDS-PAGE, silver stained [13], cut out and after destaining, in gel digested with trypsin. Samples were analyzed by nano-LC MS/MS using a quadrupole time-of-flight mass spectrometer (Micromass) and *de novo* sequencing visually inspected by the Cancer Research UK London Research Institute mass spec facility.

Lyophilized venom of *Glyceria tridactyla* Schmarida, 1861 (Annelida, Glyceridae) was further dissolved in ultrapure water to a concentration of 5 mg/ml, and 50  $\mu$ g separated by SDS-PAGE using a 12.5% Tris-glycine gel under reducing conditions. Bands were visualized by staining with colloidal Coomassie followed by destaining of the gel by 1% (*vol/vol*) acetic acid. Individual bands were dissected, digested with trypsin, and tryptic peptides eluted as described previously [65]. Proteins were identified by analyzing the tryptic peptides by LC-ESI-MS/MS and matching the resulting fragment spectra with sequences obtained by translated tissue transcriptomes (see below). LC-MS/MS experiments were carried out on an AB Sciex 5600 TripleTOF equipped with a nano-source heated to 150 °C. Venom was fractionated on a Shimadzu Prominence nano-HPLC with a 1  $\mu$ m internal diameter 100 mm Agilent 3  $\mu$ m 90 Å C18 reverse phase column at a flow of 500 nl/min and a gradient of 2–40% solvent B (0.1% formic acid (FA), 90% acetonitrile) in 0.1% FA over 10 min. MS1 scans were acquired at 350–1800 m/z with an accumulation time of 250 ms. MS2 scans were acquired on up to 20 ions per cycle that were of 80–1400 m/z with 2–5 charges and intensity greater than 120 counts per second, accumulating ions for 100 ms. Spectra were searched against a pooled-tissue transcriptomic sequence database with ProteinPilot v.5.0 (AB Sciex, Mt Waverley, Victoria, AUS) using thorough search settings and allowing for biological modifications. Decoy-based false discovery rates (FDR) were estimated by ProteinPilot, and only protein identifications ranked above the 1% local FDR threshold were considered.

## Transcriptome sequencing and GLTx characterization

### *Specimen collection and tissue preparation*

Specimens of *Glycera tridactyla* Schmarda, 1861 (Annelida, Glyceridae) were obtained from the Roscoff marine biological station (Station Biologique Roscoff, France) in February 2015. To minimize influences of stress, the animals were maximally kept for 6 days in seawater aquaria. The small-sized pharyngeal lobes (four per specimen) were left attached to the inner pharynx epithelium to ensure their complete removal during dissection. The putative venom glands (four per specimen) were cut below the basis of the teeth to exclude ducts connecting the pharyngeal lobes and teeth (Additional file 4: Figure S6). As reference tissue, a part of the posterior body wall was dissected and the gut and parapodia were removed (Additional file 4: Figure S6). Dissected tissues were immediately homogenized in TRIzol<sup>®</sup> LS Reagent (Life Technologies, Darmstadt, Germany), and stored at -20 °C before proceeding to RNA isolation.

### *RNA extraction, library reconstruction and Illumina sequencing*

Total RNA was extracted from the pharyngeal lobes, the putative venom glands, and a posterior part of the body wall (Additional file 4: Figure S6) using TRIzol<sup>®</sup> LS Reagent (Life Technologies, Darmstadt, Germany). To remove genomic DNA residues from the samples, a DNA digestion step using DNase I (Roche, Mannheim, Germany) was carried out in a RNase-free environment before purification with the RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. RNA concentration and quality were determined on a NanoDrop 2000 (Thermo Scientific, Wilmington, DE) and an Agilent 2100 Bioanalyzer (High Sensitivity RNA Chip, Agilent Technologies, Santa Clara, CA).

Transcriptome libraries were constructed for a single specimen (SR21, putative venom glands; SR23, pharyngeal lobes; and SR25, posterior body wall), and for pooled samples comprising RNA of three individuals (SR22, putative venom glands; SR24, pharyngeal lobes; and SR26, posterior body wall). For purification of mRNA out of total RNA, the Dynabeads<sup>®</sup> mRNA Purification Kit (Invitrogen, Carlsbad, CA) was used for higher concentrated samples and Sera-Mag Oligo(dT) Beads (DistriLab, Leusden, Netherlands) were used for lower concentrated samples (putative venom glands, SR21 and SR22). First strand cDNA synthesis reactions, implementing an 8 min (at 85 °C) fragmentation step, were performed with random hexamer primers (Thermo Fisher Scientific, Wilmington, DE) and SuperScript<sup>®</sup> III reverse transcriptase (Invitrogen, Carlsbad, CA). Subsequently, second strand cDNA synthesis reactions were performed with DNA

polymerase I and ribonuclease H (Life Technologies, Carlsbad, CA), and reaction products were purified with the QIAquick<sup>®</sup> PCR Purification Kit (Qiagen). Starting from the blunt-end repair, Illumina libraries were processed according to the Illumina multiplex protocol of Meyer, Kircher [66] using double indexed library adapters [67]. Libraries were sequenced together on one lane of the HiSeq 2500 (Illumina, San Diego, CA) at the Max Planck Institute for Evolutionary Anthropology (Leipzig, Germany). Afterwards, Illumina paired-end reads (140 bp) were sorted according their indices, adapters were clipped and base calling was conducted with freebis [68]. Reads with false paired indices were discarded, and overlapping paired-end reads were trimmed and merged to a single sequence [69].

### *Processing of sequencing data*

Illumina raw reads were trimmed (10 bp) at both ends and single sequences shorter than 60 bp were removed using cutadapt v.1.8.1 [70], respectively. Afterwards, Illumina sequences were filtered with ConDeTri v.2.2 [71] and only reads of which 95% of the nucleotides have a PHRED score [72, 73] above 15 were kept for further analyses (Additional file 11). The processed (using cutadapt v.1.8.1 and ConDeTri v.2.2) Illumina reads were assembled *de novo* using IDBA-tran v.1.1.1 [74]. IDBA-tran assemblies are constructed using an initial k-mer size of 20, an iteration size of 5, and a maximum k-mer size of 120 (Additional file 11).

### *Identification of GLTx in transcriptome libraries and cDNA*

Assemblies were screened for putative GLTx transcripts through BLAST-searches (tblastn) v.2.2.28+ [75] using short amino acid sequences of the GLTx protein (see Material and methods section "Protein sequencing" and Additional file 1) as reference.

A nearly full-length GLTx transcript (around 3600 bp; using SR23, contig 5772 as reference) was amplified in second strand cDNA (primer pair full-transF/full-transR; Additional file 12) prior to cloning (for details see Material and methods section "Morphological analyses, cloning"). Finally, three GLTx clones (Cons\_clone\_5A, 6A, and 7) were analyzed through primer walking (Additional file 12). Amplicons were sequenced at the GATC Biotech AG (Constance, Germany).

### *Annotation of GLTx full-length transcripts*

GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) and three nearly full-length GLTx clones (Cons\_clone\_5A, 6A, and 7; Additional file 3) were annotated using common online tools. Signal peptides were identified through the SignalP v.4.1 server [76]. Identification of signaling domains was carried out using Pfam v.30.0 database [77] and SMART, the simple

modular architecture research tool [78, 79]. Similarity to published sequence data was analyzed through BLAST-searches (blastp) v.2.6.0+ in NCBI GenBank. Molecular weight was calculated based on the full-length transcripts (without signal peptide) using CLC Main Workbench v.7.7 (CLCbio, Qiagen, Aarhus, Denmark; www.clcbio.com).

#### **GLTx paralog screening**

To screen for putative GLTx paralogs, a Maximum likelihood analyses v.8.2.8 [80] of different clones was performed (raxmlHPC-PTHREADS-AVX, GTR + GAMMA + I, 1,000 pseudoreplicates). The phylogenetic analysis comprises the same 746 bp GLTx gene fragment (Additional file 6) analyzed in 49 GLTx clones (primer pair 2 F/4R; see Material and methods section “Morphological analyses, cloning”), three GLTx clones (primer pair full-transF/full-transR; see Material and methods section “Identification of GLTx in cDNA”), and three GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318; see Material and methods section “Identification of GLTx in transcriptome libraries”). The unrooted paralog tree was visualized and edited with iTOL v.3.2.4 [81–83].

#### **Genomic structure of the GLTx gene**

The genomic structure of the GLTx gene was analyzed in genomic DNA of *G. tridactyla*, collected in April 2011 nearby the Roscoff marine station. DNA was extracted using the NucleoSpin® Tissue Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer’s protocols. Based on transcriptome sequence information, primers were designed using NetPrimer (PREMIER Biosoft, Palo Alto CA; Additional file 12), and PCR experiments carried out. Unknown gene parts that are adjacent to known gene regions were determined using the Genome-Walker™ Universal Kit (Clontech Laboratories, Inc., Takara Bio Company, Mountain View, CA) [84]. Amplicons were purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer’s protocols. Sanger sequencing was performed by the GATC Biotech AG (Constance, Germany).

#### **Morphological analyses on the glycerid venom apparatus**

##### **Specimen collection and fixation**

For *in situ* hybridization experiments and antibody staining, adult specimens of *Glycera tridactyla* Schmarida, 1861 (Annelida, Glyceridae) were collected

intertidally from muddy areas of the rocky shore nearby Roscoff marine biological station in spring 2013. For fluorescence *in situ* hybridization (FISH) experiments and antibody staining, adult individuals of *G. tridactyla* were collected from sandy beach sections of the intertidal zone nearby the Wimereux marine biological station (Station Marine de Wimereux, Université de Lille, France) in March 2014.

Specimens used for *in situ* and FISH were fixed in 4% paraformaldehyde (PFA) in 0.1 M phosphate buffered saline (PBS) for 4 h at 4 °C, washed 5 min in a solution equally proportioned PBS and methanol, before transferred in 100% methanol, and stored at –20 °C. Samples used for antibody staining were fixed in 4% PFA in 0.1 M PBS overnight at 4 °C, washed 3 times for at least 2 h in PBS at 4 °C, and stored in PBS containing 0.005% sodium azide (NaN<sub>3</sub>) at 4 °C.

#### **RNA isolation, amplification of a GLTx gene fragment, cloning and probe construction**

Total RNA was extracted from whole pharynx tissue of an adult *G. tridactyla* using TRIzol® Reagent (Invitrogen, Carlsbad, CA), and purified through the RNeasy MinElute Cleanup Kit (Qiagen) according to the manufacturer’s protocols. The specimen was collected nearby the Roscoff marine biological station in December 2010, and fixed in RNAlater (Ambion, Darmstadt, Germany). First strand cDNA synthesis was performed using random hexamer primers (Fermentas, St. Leon-Rot, Germany) and SuperScript™ III reverse transcriptase (Invitrogen, Carlsbad, CA). Second strand synthesis was carried out with ribonuclease H (Invitrogen, Carlsbad, CA) and DNA polymerase I (Invitrogen, Carlsbad, CA), and the second strand cDNA product subsequently purified through the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer’s protocols.

A 746 bp exonic gene fragment of GLTx identified in *G. tridactyla* (exon 3, Additional file 4: Figure S2) was amplified through the primer pair 2 F/4R (Additional file 12) in second strand cDNA and genomic DNA, and subsequently purified through the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel). The purified PCR products were cloned into the pGEM® -T Vector (pGEM®-T Vector System I, Promega Corporation, Madison, WI) and transformed in *E. coli* JM109. Finally, 49 clones were transferred in HPLC-H<sub>2</sub>O, frozen (–20 °C), defrost, amplified through the M13 primer pair (M13F/M13R), and sequenced at the GATC Biotech AG (Constance, Germany). Two clones were used for preparation of digoxigenin-labeled RNA probes through the DIG RNA Labeling Kit SP6/T7 (Roche) according to the manufacturer’s protocols.



### **In situ hybridization**

Expression studies were performed on pharynx tissues of *G. tridactyla*. Per specimen, the pharynx was dissected into two halves.

The pharynx samples were rehydrated stepwise by washing at room temperature for 5 min in mixtures of a serial dilution (see Additional file 13: Protocol S1) of methanol and PTW (1 × PBS + 0.1% Tween-20), followed by 4 × washing in 100% PTW. Samples were digested with proteinase K (0.01 mg/ml in PTW) for 5 min without shaking, and stopped by washing twice for 5 min in glycine/PTW (2 mg/ml). The pharynx samples were then washed in 1% triethanolamine in PTW, and glacial acetic acid was added twice with an incubation time of 5 min each to permeabilize the cells. Samples were washed twice in 100% PTW for 5 min, and then re-fixed through 60 min incubation in 4% PFA in PTW, at room temperature on a shaker. Afterwards, the pharynx samples were again washed 5 × 5 min in PTW, transferred in new 2 ml non-sticky tubes filled with PTW, incubated for 5 min on a shaker before heated to 80 °C for 10 min without shaking. After removing liquids, the samples were incubated in hybridization buffer for 10 min at room temperature. Liquids were discarded, 65 °C pre-warmed hybridization buffer added, and pre-hybridization was carried out overnight at 65 °C. For hybridization, digoxigenin-labeled RNA probes (SP6/T7, concentration: 1 ng/μl in hybridization buffer) were denatured by heating for 10 min at 80 °C without shaking. Per analyzed specimen, one pharynx half was transferred in SP6 probe, the other in T7 probe (sense and antisense), and hybridization was performed for 72 h at 65 °C. After removing probes, the pharynx tissues were washed at 65 °C for 5 min and 20 min in 65 °C pre-warmed hybridization buffer. The pharynx samples were then washed at 65 °C stepwise in a serial dilution of 65 °C pre-warmed mixtures of hybridization buffer and 2 × SSC (sodium saline citrate) for 10 min, followed by washing 10 min in 65 °C pre-warmed 100% 2 × SSC, and 2 × 30 min in 65 °C pre-warmed 0.02 × SSC. At room temperature, additional 5-min wash steps in mixtures of a serial dilution of 0.02 × SSC and PTW were carried out, followed by washing 6 × 5 min in 100% PTW. For visualization, the samples were blocked for 1 h at room temperature on a shaker in blocking buffer (5% normal goat serum in PTW). The Anti-Digoxigenin-AP Fab fragments antibody (Roche, Mannheim, Germany) diluted at 1:5,000 in blocking buffer was added, and incubation was carried out overnight at 4 °C on a shaker. After washing 8 × 10 min in PTW at room temperature, followed by washing 3 × 5 min in AP staining buffer, color staining was initiated by adding the NBT/BCIP staining solution (Carl Roth, Karlsruhe, Germany). Light-sensitive staining reaction was kept in

the dark without shaking, and stopped after 45 min–3 h through 4% PFA in PTW. After 60 min incubation, the pharynx samples were washed once in PTW, placed overnight at 4 °C on a shaker, and washed again 2 × 2 h in PTW at 4 °C. Until imaging, the pharynx samples were stored in the dark at 4 °C without shaking. After imaging, the pharynx tissues were transferred in 0.1 M PBS containing 0.005% NaN<sub>3</sub>, and stored in the dark at 4 °C.

To detect signal, the pharynx samples were analyzed under a stereomicroscope (Leica WILD M10, Leica Microsystems, Wetzlar, Germany) equipped with a color digital camera (SensiCam, 12 bit cooled imaging, PCO AG, Kelheim, Germany), and the CamWare v.3.11 software. Final panels were designed with Adobe Photoshop CS5.1 and Adobe Illustrator CS5.1.

### **Anti-GLTx staining**

Antibody staining (see Additional file 13: Protocol S2) was carried out on pharynx tissues of *G. tridactyla*. The everted pharynx was dissected into two equally sized halves.

For tissue permeabilization, pharynx samples were incubated overnight at room temperature in a solution of 0.1 M PBS containing 0.1% Triton X-100 (PTA), 0.1% NaN<sub>3</sub>, and 6% normal goat serum (Sigma-Aldrich, St. Louis, MO, USA) (block-PTA). The primary antibody monoclonal mouse anti-GLTx (4G9, [13]; diluted 1:500), was applied for 48–72 h at 4 °C on a shaker. Samples were then washed 3 × 2 h at room temperature in 0.1 M PBS containing 0.1% Triton X-100 (PTA), and block-PTA. The secondary fluorochrome conjugated antibody (goat anti-mouse Alexa Fluor 488; Invitrogen, Carlsbad, CA; diluted 1:500) was added and incubation was carried out in the dark for 72 h at 4 °C on a shaker. Subsequently, tissue samples were washed 2 × 1.5 h–2 h in 0.1 M PBS, followed by 2-h incubation in phalloidin-rhodamine (5 μl phalloidin stock solution per 500 μl PBS; Invitrogen, Darmstadt, Germany) for additional staining of muscle tissue. At last, tissue samples were dehydrated in an ascending series of isopropanol, treated for 10 min in Murray's clearing solution (benzyl alcohol + benzyl benzoate, at a ratio of 1:2), and finally mounted between two coverslips in dibutyl phthalate xylene (DPX; Sigma-Aldrich).

Specimens were analyzed with the confocal laser scanning microscope Leica TCS STED (Leica Microsystems, Wetzlar, Germany), and confocal image stacks were processed with Leica AS AF v.2.3.5 (Leica Microsystems) and Imaris v.6.3.1 (Bitplane AG, Zurich, Switzerland). Final panels were designed with Adobe Photoshop CS5.1 and Adobe Illustrator CS5.1.



**Antibody staining (II) against GLTx**

Antibody staining was further performed on sections of putative venom glands of *G. tridactyla* [13], following classical paraffin-embedded tissue sections, treated with blocking buffer (0.1 M PBS, 3% normal goat serum, and 0.1% Triton X-100). The primary antibody monoclonal mouse anti-GLTx (4G9, see Material and methods section “Anti-GLTx staining”) was applied overnight at 4 °C, washed several times with blocking buffer by addition of the secondary antibody (goat anti-mouse Alexa Fluor 488; Invitrogen, Carlsbad, CA). Sections were examined with a Zeiss LSM 510 confocal microscope (Carl Zeiss, Jena, Germany).

**Antibody staining against serotonin**

Antibody staining against serotonin was performed on bisected pharynges of *G. tridactyla* as described in the Material and methods section “Anti-GLTx staining”, using the primary antibody polyclonal rabbit anti-serotonin (INCSTAR, Stillwater, USA; diluted 1:500), and the secondary fluorochrome conjugated antibody goat anti-rabbit Alexa Fluor 633 (Invitrogen, Carlsbad, CA; diluted 1:500). The samples were embedded, analyzed and processed as described in the Material and methods section “Anti-GLTx staining”.

**Fluorescence in situ hybridization (FISH) coupled with antibody staining against GLTx**

Double-staining (see Additional file 13: Protocol S3) was performed on an entire pharynx of *G. tridactyla*.

The first part of protocol is concordant with the *in situ* hybridization protocol (Additional file 13: Protocol S1), with modifications regarding the blocking reagent, the anti-digoxigenin antibody, and color staining solution. Before blocking, the pharynx sample was washed five times for 5 min in TNT, instead of PTW. Furthermore, blocking was carried out for 3 h in TNB blocking buffer (0.5% blocking reagent, PerkinElmer, Beaconsfield, MA), and incubated overnight with the Anti-Digoxigenin-POD, Fab fragments antibody (Roche, Mannheim, Germany), diluted at 1:100 in TNB blocking buffer. Other than in the *in situ* protocol, color staining was performed in FITC-tyramide staining solution, diluted at 1:50 in 1 × Plus Amplification Diluent (TSA™ Plus Fluorescein System, PerkinElmer), and stopped after 30 min through washing for 5 min, and 10 min in TNT. After color staining, the pharynx sample was kept in the dark. Next, the sample was washed for 5 min at room temperature in mixtures of a serial dilution of TNT/0.1 M PBS, followed by washing 3 × for 5 min in 0.1 M PBS and proceeding with a modified version of the anti-GLTx staining protocol (Additional file 13: Protocol S2). As modification, the tissue was permeabilized for only

30 min, and the primary antibody and the secondary fluorochrome conjugated antibody (goat anti-mouse Alexa Fluor 568; Invitrogen, Carlsbad, CA; diluted 1:500) were added overnight, instead of 72 h. As a further modification, additional nuclear counterstaining was carried out through 1-h incubation with TO-PRO<sup>®</sup>-3 Iodide (Life Technologies, Darmstadt, Germany). The sample was analyzed and processed as described at the end of the anti-GLTx staining protocol (see Material and methods section above).

**Expression studies****Specimen collection and tissue preparation**

Quantitative real-time PCR (qPCR) experiments were performed on *Glycera tridactyla* Schmarda, 1861 (Annelida, Glyceridae) specimens obtained from the Roscoff marine biological station in February 2015. Three different tissue types (pharyngeal lobes, putative venom glands, and posterior body wall) were dissected as described in detail for library preparation (see according Material and methods section “Transcriptome sequencing”).

**RNA isolation and first strand cDNA synthesis**

Total RNA (600 ng each) obtained from pharyngeal lobes, putative venom glands, and a posterior part of the body wall (for details see Material and methods section “Transcriptome sequencing”) was used in first strand cDNA synthesis reactions carried out with random hexamer primers (Thermo Fisher Scientific) and SuperScript<sup>®</sup> III reverse transcriptase (Invitrogen, Carlsbad, CA). The first strand cDNA synthesis products were used for quantitative real-time PCR (qPCR) experiments.

**Quantitative real-time PCR with SYBR green**

Quantitative real-time PCR (qPCR) studies were performed with Platinum<sup>®</sup> SYBR<sup>®</sup> Green qPCR SuperMix-UDG with ROX (Invitrogen, Carlsbad, CA) using half the sample size (25 µl each) as described in the manufacturer’s protocol. Ten biological samples (= specimens) were analyzed with two technical replicates per studied amplicon, each around 210 bp in size. Analyses were performed on the same exonic gene region (primers see Additional file 10) of three GLTx paralogs (see Results section “Characterization of glycerotoxin”) and two adjacent gene regions amplified with one primer spanning an intron-exon boundary (primers see Additional file 10). Amplification of the normalizer genes (reference genes/endogenous control) — the single copy ribosomal protein genes *rps3* and *rps15a* — was also carried out with an intron-exon spanning primer. Samples were run on the Applied Biosystems 7300 Real-Time

PCR System (Applied Biosystems, Darmstadt, Germany) under the following cycling conditions: 1 cycle at 50 °C/2 min, 1 cycle of denaturation at 95 °C/2 min, followed by 40 two-segment cycles of amplification (95 °C/15 s, 60 °C/30 s), and a final dissociation cycle. Baseline and threshold were adjusted automatically and Ct values were determined for each sample using the accompanying 7300 System SDS v.1.4 software (Applied Biosystems). Delta-Ct values ( $\Delta Ct$ ) used to normalize gene expression were calculated as follows:  $\Delta Ct = \text{average } Ct_{\text{target gene}} - \text{Normalization Factor (NF)}$ , where NF is the mean Ct of both endogenous controls used (*rps3* and *rps15a*) as described by Vandesompele et al. [85]. The average  $Ct_{\text{target gene}}$  comprises of two technical replicates per gene for each studied specimen. The relative quantification shown as Fold Change (RQ) between biological groups (putative venom glands, pharyngeal lobes, and body tissue from all analyzed specimens) were calculated automatically with the DataAssist™ v.3.01 software (Applied Biosystems) according to the following formula:  $RQ = \text{geometric mean } 2^{(-\Delta Ct)} / \text{geometric mean } 2^{(-\Delta Ct_{\text{reference}})}$ . Thereby, the replicates of all biological samples of a group contributed to their respective geometric mean calculation. RQ significance was assessed by a two-sample, two-tailed Student's *t*-test comparing the  $2^{(-\Delta Ct)}$  values of the groups and p-values were adjusted using Benjamini-Hochberg False Discovery Rate [86] using the DataAssist™ v.3.01 software (Applied Biosystems).

#### Expression level estimates based on transcriptome libraries

GLTx expression was estimated from *G. tridactyla* transcriptome libraries (for details on library preparation see Material and methods section “Transcriptome sequencing”) of three different tissue types (putative venom glands, pharyngeal lobes, and posterior body wall). Before mapping, the processed and filtered Illumina reads of a single specimen library and the referring pooled library were merged (SR21 + SR22, putative venom glands; SR23 + SR24, pharyngeal lobes; SR25 + SR26, posterior body wall; Additional file 11). Merged Illumina reads were mapped against three GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) and three nearly full-length GLTx clones (Cons\_clone\_5A, 6A, and 7). Mapping was carried out using segemehl v.0.2.0 [87, 88] and an identity score of 95% ( $A = 95$ ). Mapping results were visualized using Tablet v.1.15.09.01 [89]. To estimate the Fold Changes of GLTx transcription between the pharyngeal lobes and putative venom glands, the relative numbers of matched reads per contig from either of these libraries were calculated and thus put in relation to the other (matched reads<sub>lobes</sub> divided by matched reads<sub>gland</sub>; Table 1 and Additional file 9).

## Additional files

**Additional file 1:** Position of short amino acid sequences alongside the GLTx full-length transcripts. Short amino acid sequences were recovered through protein sequencing of purified glycerotoxin extracted out of the venom gland cocktail in *G. tridactyla*. Short amino acid sequences varying in length between 7–18 amino acids served as reference in BLAST-searches to identify the neurotoxin in transcriptome libraries of *G. tridactyla*. (TXT 30 kb)

**Additional file 2:** Nucleotide alignment of GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) and three nearly full-length GLTx clones amplified from second strand cDNA (Cons\_clone\_5A, 6A, and 7) in *G. tridactyla*. (TXT 23 kb)

**Additional file 3:** Protein-translated alignment of GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) and three nearly full-length GLTx clones amplified from second strand cDNA (Cons\_clone\_5A, 6A, and 7) in *G. tridactyla*. (TXT 7 kb)

**Additional file 4: Figure S1.** SDS-PAGE of reduced *G. tridactyla* venom. **Figure S2.** Genomic organization of the glycerotoxin gene analyzed in *G. tridactyla*. **Figure S3.** Immunolocalization of GLTx in a cross section through a putative venom gland embedded in paraffin. **Figure S4.** Anti-serotonin (5-HT) staining and phalloidin–rhodamine counterstaining on everted *G. tridactyla* pharynges cut into two halves. **Figure S5.** Quantitative real-time PCR (qPCR) expression levels (shown as Fold Change, RQ) between biological groups (putative venom glands [A], pharyngeal lobes [B], and posterior body wall [C]) per analyzed specimen ( $n = 10$ , biological samples). **Figure S6.** Tissues analyzed in comparative GLTx expression studies (qPCR experiments and transcriptome analyses) on *G. tridactyla*. (PDF 2931 kb)

**Additional file 5:** MS-MS best hits of the 150 kDa fractions (sheet a and b) of *G. tridactyla* venom. Spectra were searched against a pooled-tissue transcriptomic sequence database of the pharyngeal lobes (SR23, single specimen and merged SR23 + SR24, pooled multiple specimen) with ProteinPilot v.5.0. (XLSX 14 kb)

**Additional file 6:** Alignment comprising a 746 bp gene fragment of 52 GLTx clones and three GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318). The alignment was used for the maximum likelihood analysis shown in Fig. 1b. (TXT 42 kb)

**Additional file 7:** Alignment of amplicons generated in genomic DNA of *G. tridactyla*. Sanger sequences revealed an intron-exon-structure for the GLTx gene (see Additional file 4: Figure S2). (TXT 68 kb)

**Additional file 8:** Immunolocalization of GLTx in a pharynx of *G. tridactyla* which was cut into two halves. Confocal maximum projection. Anti-GLTx staining occurs in clear defined pharyngeal lobes. GLTx-IR staining revealed a canalized GLTx transport from the lobes to the teeth, where it becomes delivered through a series of pores. (AVI 117177 kb)

**Additional file 9:** Comparative GLTx expression analyses carried out on transcriptome libraries of three different tissue types (putative venom glands (pvg), pharyngeal lobes, and posterior body wall) originating from four pooled *G. tridactyla* specimens. Filtered Illumina reads were mapped against three GLTx full-length transcripts (SR23, contig 5772, contig 4317, and contig 4318) and three full-length GLTx clones (Cons\_clone\_5A, 6A, and 7). Note the remarkably higher number of matched reads in the pharyngeal lobes in comparison to the putative venom glands and body tissue. (XLSX 11 kb)

**Additional file 10:** Raw data and summary of qPCR experiments performed on three tissue types (putative venom glands [group A], pharyngeal lobes [group B], and posterior body wall [group C]) of *G. tridactyla*. sheet a Ct values determined per analyzed gene and technical replicate for 10 biological samples using the 7300 System SDS v.1.4 software (Applied Biosystems). sheet b and c The Fold Change (RQ, relative quantification) between biological groups (putative venom glands [A], pharyngeal lobes [B], and posterior body wall [C]) and p-values were calculated with the DataAssist™ v.3.01 software (Applied Biosystems). Two technical replicates per analyzed biological sample ( $n = 10$ ) contribute to the calculations of one group. The single copy ribosomal genes *rps3* and *rps15a* were used as normalizer genes. Software parameters: Maximum allowable Ct value, 40.0; Include max Ct values in calculations, Yes; Exclude outliers among replicates, Yes; Adjust p-values using Benjamini-Hochberg False Discovery Rate, Yes; Normalization method, Endogenous Control; Selected

controls, rps15a and rps3. **sheet b** RQ-values and p-values calculated for the putative venom glands [A] and pharyngeal lobes [B] in comparison to the body tissue [C]. **sheet c** RQ-values and p-values calculated for the pharyngeal lobes [B] in comparison to the putative venom glands [A]. **sheet d** and **e** The Fold Change (RQ, relative quantification) between biological groups (putative venom glands [A], pharyngeal lobes [B], and posterior body wall [C]) per analyzed specimen (n = 10, biological samples). Two technical replicates were performed per analyzed specimen. The single copy ribosomal genes rps3 and rps15a were used as normalizer genes. **sheet f** Primer sets used for qPCR amplification of five GLTx specific gene fragments and two normalizer genes (rps3, rps15a). (XLSX 63 kb)

**Additional file 11: sheet a** Summary statistics of filtering steps (using cutadapt v.1.8.1 and ConDeTri v.2.2) performed on raw Illumina sequencing data of three different tissue types (putative venom glands, pharyngeal lobes, and posterior body wall) analyzed in *G. tridactyla*. Transcriptome libraries were constructed for a single individual (SR21, putative venom glands; SR23, pharyngeal lobes; and SR25, posterior body wall), and based on pooled samples comprising RNA of three individuals (SR22, putative venom glands; SR24, pharyngeal lobes; and SR26, posterior body wall). Abbreviations: PE, paired-end reads; SR, single reads. **sheet b** Summary statistics of IDBA-tran v.1.1.1 assemblies performed on processed Illumina sequencing data originating from *G. tridactyla*. Assemblies were done for three different tissue types of a single individual (SR21, putative venom glands; SR23, pharyngeal lobes; and SR25, posterior body wall), and for Illumina sequencing data of three pooled individuals (SR22, putative venom glands; SR24, pharyngeal lobes; and SR26, posterior body wall). In addition, the filtered reads of single and pooled libraries were merged (SR21 + SR22, putative venom glands; SR23 + SR24, pharyngeal lobes; SR25 + SR26, posterior body wall) to perform comparative expression studies by mapping them against GLTx reference sequences. (XLSX 17 kb)

**Additional file 12:** Set of primers used to analyze the molecular organization of the GLTx gene. **sheet a** Primer used to amplify GLTx fragments before cloning into the pGEM<sup>®</sup>-T Vector (pGEM<sup>®</sup>-T Vector System I, Promega Corporation, Madison, WI). **sheet b** Primer used for primer walking (Cons\_clone\_5A, 6A, and 7). **sheet c** Primer used to analyze the intron-exon-structure in genomic DNA of *G. tridactyla*. GW, Genome walking. (XLSX 14 kb)

**Additional file 13: Protocol S1–S3.** Protocol S1. Protocol: *In situ* hybridization on pharynx tissue of *G. tridactyla* (Glyceridae, Annelida). Protocol S2. Protocol: Anti-GLTx staining on pharynx tissue of *G. tridactyla* (Glyceridae, Annelida). Protocol S3. Protocol: Fluorescence *in situ* hybridization (FISH) coupled with antibody staining against GLTx on pharynx tissue of *G. tridactyla*. (PDF 590 kb)

## Acknowledgements

We are grateful to Martin Schlegel (University of Leipzig, Germany) for providing working facilities and support. We gratefully acknowledge the Roscoff marine station (CNRS, UPMC, Station Biologique Roscoff, France) and the Wimereux marine station (Station Marine de Wimereux, Université de Lille, France) to use their facilities during our research stays. We are thankful to the Max Planck Institute for Evolutionary Anthropology Leipzig for Illumina sequencing, especially Anne Weigert, Birgit Nickel, Matthias Meyer, and Svante Pääbo. We further thank the group of Paul A. Stevenson (University of Leipzig, Germany) for providing the stereomicroscope and confocal laser scanning microscope. We are thankful to the group of Georg Mayer (University of Kassel, Germany) to provide the TO-PRO<sup>®</sup>-3 reagents. Furthermore, we thank Markus Böggemann (University of Vechta, Germany) for helpful discussions on the manuscript. CB is a "Ramon y Cajal" fellow supported by the Spanish Ministry of Science and Education (MEC) (RYC-2014-15615). CH is financed by a personal research fellowship from the German Research Foundation (DFG; grant HE 7224/1-1). EABU acknowledges support from the Australian Research Council (Discovery Early Career Researcher Award DE160101142). RAJ gratefully acknowledges support from the Biotechnology and Biological Sciences Research Council (Grant BB/K003488/1). This work was supported by Cancer Research UK [GS], a Wellcome Trust Senior Investigator Award (107116/Z/15/Z) [GS], and University College London [GS]. This work was further supported by the German Research Foundation (DFG; grant BL787/7-1) and an EU ASSEMBLE grant (No. 227799; <http://www.assemblemarine.org>) to CB. We acknowledge support from the German Research Foundation (DFG) and Universität Leipzig within the program of Open Access Publishing.

## Funding

Funder	Grant reference number	Author
German Research Foundation (DFG)	HE 7224/1-1	Conrad Helm
Australian Research Council (Discovery Early Career Researcher Award)	DE160101142	Eivind A. B. Undheim
Biotechnology and Biological Sciences Research Council	BB/K003488/1	Ronald A. Jenner
Cancer Research UK	LF4605	Giampietro Schiavo
University College London	506303	Giampietro Schiavo
Wellcome Trust Senior Investigator Award	107116/Z/15/Z	Giampietro Schiavo
EU ASSEMBLE	227799	Christoph Bleidorn
German Research Foundation (DFG)	BL787/7-1	Christoph Bleidorn
Spanish Ministry of Science and Education (MEC)	RYC-2014-15615	Christoph Bleidorn

## Availability of data and materials

The data sets supporting the conclusions of this article are included within the article and its additional files. The GLTx clones have been deposited at GenBank under the accession numbers KY464001, KY464002, and KY464003. The Illumina short reads were submitted to the Sequence Read Archive (SRA) of NCBI (accession numbers SRR5167052, SRR5167051, SRR5167050, SRR5167049, SRR5167048, and SRR5167047). The morphological data sets are accessible in the MorphDBase repository, <https://www.morphdbase.de/> [90], under the accessions [www.morphdbase.de/?C\\_Helm\\_20170119-M-5.1](http://www.morphdbase.de/?C_Helm_20170119-M-5.1) (Fig. 3a), [www.morphdbase.de/?C\\_Helm\\_20170119-M-7.1](http://www.morphdbase.de/?C_Helm_20170119-M-7.1) (Fig. 3b), [www.morphdbase.de/?C\\_Helm\\_20170119-M-6.1](http://www.morphdbase.de/?C_Helm_20170119-M-6.1) (Fig. 3c), [www.morphdbase.de/?C\\_Helm\\_20170116-M-2.1](http://www.morphdbase.de/?C_Helm_20170116-M-2.1) (Fig. 3d and Fig. 5), and [www.morphdbase.de/?C\\_Helm\\_20170116-M-4.1](http://www.morphdbase.de/?C_Helm_20170116-M-4.1) (Fig. 4).

## Authors' contributions

CB, GS, SR designed the study; SR, CB, GS, RAJ, FAM analyzed and interpreted the data; SR, CH, LH conducted *in situ*, FISH and antibody stainings; SR, LH, SHD conducted qPCR experiments; GS, FAM, EABU conducted proteomic work; SR prepared transcriptomic libraries; SR, LIC performed the bioinformatic analyses. SR wrote the first draft of the manuscript. All authors contributed in revising the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Author details

<sup>1</sup>Institute of Biology - Molecular Evolution and Systematics of Animals, University of Leipzig, Talstraße 33, D-04103 Leipzig, Germany. <sup>2</sup>German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, D-04103 Leipzig, Germany. <sup>3</sup>Department of Life Sciences, Natural History Museum, Cromwell Rd, London SW7 5BD, UK. <sup>4</sup>Sars



International Centre for Marine Molecular Biology, University of Bergen, Thormøhlensgate 55, N-5008 Bergen, Norway. <sup>5</sup>Clem Jones Centre for Ageing Dementia Research, Queensland Brain Institute, University of Queensland, St. Lucia, Brisbane 4072, QLD, Australia. <sup>6</sup>Institute of Biology - Department of Zoology, University of Kassel, Heinrich-Plett-Straße 40, D-34132 Kassel, Germany. <sup>7</sup>The European Bioinformatics Institute (EMBL-EBI) - Wellcome Genome Campus, Hinxton, Cambridge CB10 1SD, UK. <sup>8</sup>Centre for Advanced Imaging, University of Queensland, St. Lucia, Brisbane 4072, QLD, Australia. <sup>9</sup>Sobell Department of Motor Neuroscience & Movement Disorders, UCL Institute of Neurology, University College London, Queen Square, London WC1N 3BG, UK. <sup>10</sup>Museo Nacional de Ciencias Naturales, Spanish National Research Council (CSIC), Calle José Gutierrez Abascal 2, 28006 Madrid, Spain.

Received: 23 October 2016 Accepted: 6 February 2017

Published online: 04 March 2017

## References

- Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JDA, King GF, et al. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu Rev Genomics Hum Genet.* 2009;10:483–511. doi:10.1146/annurev.genom.9.081307.164356.
- Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol Evol.* 2013;28(4):219–29. doi:10.1016/j.tree.2012.10.020.
- von Reumont BM, Campbell LI, Jenner RA. Quo vadis venomics? A roadmap to neglected venomous invertebrates. *Toxins.* 2014;6(12):3488–551.
- Kalia J, Milescu M, Salvatierra J, Wagner J, Klint JK, King GF, et al. From foe to friend: using animal toxins to investigate ion channel function. *J Mol Biol.* 2015;427(1):158–75. doi:10.1016/j.jmb.2014.07.027.
- Terlau H, Olivera BM. *Conus* venoms: a rich source of novel ion channel-targeted peptides. *Physiol Rev.* 2004;84(1):41–68. doi:10.1152/physrev.00020.2003.
- Pringos E, Vignes M, Martinez J, Rolland V. Peptide neurotoxins that affect voltage-gated calcium channels: a close-up on  $\omega$ -Agatoxins. *Toxins.* 2011; 3(1):17–42.
- Schmidtko A, Lötsch J, Freyhagen R, Geisslinger G. Ziconotide for treatment of severe chronic pain. *Lancet.* 2010;375(9725):1569–77. doi:10.1016/S0140-6736(10)60354-6.
- Catterall WA. Voltage-gated Calcium channels. *Cold Spring Harb Perspect Biol.* 2011;3(8):1–23. doi:10.1101/cshperspect.a003947.
- Brini M, Cali T, Ottolini D, Carafoli E. Neuronal calcium signaling: function and dysfunction. *Cell Mol Life Sci.* 2014;71(15):2787–814. doi:10.1007/s00018-013-1550-7.
- Zamponi GW, Striessnig J, Koschak A, Dolphin AC. The physiology, pathology, and pharmacology of voltage-gated calcium channels and their future therapeutic potential. *Pharmacol Rev.* 2015;67(4):821–70. doi:10.1124/pr.114.009654.
- Yang S, Liu Z, Xiao Y, Li Y, Rong M, Liang S, et al. Chemical punch packed in venoms makes Centipedes excellent predators. *Mol Cell Proteomics.* 2012; 11(9):640–50. doi:10.1074/mcp.M112.018853.
- Undheim EAB, Fry BG, King GF. Centipede venom: recent discoveries and current state of knowledge. *Toxins.* 2015;7(3):679–704.
- Meunier FA, Feng Z-P, Molgó J, Zamponi GW, Schiavo G. Glycerotoxin from *Glycera convoluta* stimulates neurosecretion by up-regulating N-type  $Ca^{2+}$  channel activity. *EMBO J.* 2002;21(24):6733–43.
- Böggemann M. *Glyceridae* Grube, 1850. In: Westheide W, Purschke G, editors. *Handbook of Zoology Online. Annelida: Polychaetes.* Berlin: De Gruyter; 2014. <http://www.degruyter.com/view/db/zoology>. Accessed 18 Feb 2015.
- Ehlers EH. Die Borstenwürmer (Annelida Chaetopoda) nach systematischen und anatomischen Untersuchungen. Leipzig: Verlag von W. Engelmann; 1868.
- Pleijel F. *Glyceriformia* Fauchald, 1977. In: Rouse GW, Pleijel F, editors. *Polychaetes.* New York: Oxford University Press; 2001. p. 111–14.
- Wolf G. Kieferorgane von *Glyceriden* (Polychaeta) - ihre Funktion und ihr taxonomischer Wert. *Senckenbergiana marit.* 1977;9(5/6):261–83.
- Fauchald K, Rouse G. *Polychaete systematics: Past and present.* *Zool Scr.* 1997;26(2):71–138.
- Michel C. Mâchoires et glandes annexes de *Glycera convoluta* (Keferstein) Annelide Polychète *Glyceridae*. *Cah Biol Mar.* 1966;7(4):367–73.
- Michel C. Rôle physiologique de la trompe chez quatre annélides polychètes appartenant aux genres: *Eulalia*, *Phyllodoce*, *Glycera* et *Notomastus*. *Cah Biol Mar.* 1970;XI:209–28.
- Richter S, Schwarz F, Hering L, Böggemann M, Bleidorn C. The utility of genome skimming for phylogenomic analyses as demonstrated for glycerid relationships (Annelida, *Glyceridae*). *Genome Biol Evol.* 2015;7(12):3443–62. doi:10.1093/gbe/evw224.
- Bon C, Saliou B, Thieffry M, Manaranche R. Partial purification of  $\alpha$ -glycerotoxin, a presynaptic neurotoxin from the venom glands of the polychaete annelid *Glycera convoluta*. *Neurochem Int.* 1985;7(1):63–75. doi:10.1016/0197-0186(85)90009-9.
- Böggemann M. Revision of the *Glyceridae* Grube 1850 (Annelida: Polychaeta). *Abhandlungen der Senckenbergischen Naturforschenden Gesellschaft Frankfurt am Main*, vol 555. Stuttgart: E. Schweizerbart'sche Verlagsbuchhandlung; 2002.
- Lichtenegger HC, Schöberl T, Bartl MH, Waite H, Stucky GD. High abrasion resistance with sparse mineralization: copper biomineral in worm jaws. *Science.* 2002;298(5592):389–92. doi:10.1126/science.1075433.
- Michel C, Keil B. Biologically active proteins in the venomous glands of the polychaete annelid, *Glycera convoluta* Keferstein. *Comp Biochem Physiol, B.* 1975;50(1):29–33. doi:10.1016/0305-0491(75)90294-1.
- Manaranche R, Thieffry M, Israel M. Effect of the venom of *Glycera convoluta* on the spontaneous quantal release of transmitter. *J Cell Biol.* 1980;85(2): 446–58. doi:10.1083/jcb.85.2.446.
- Thieffry M, Bon C, Manaranche R, Saliou B, Israël M. Partial purification of the *Glycera convoluta* venom components responsible for its presynaptic effects. *J Physiol.* 1982;78(4):343–7.
- Morel N, Thieffry M, Manaranche R. Binding of a *Glycera convoluta* neurotoxin to cholinergic nerve terminal plasma membranes. *J Cell Biol.* 1983;97(6):1737–44.
- Schenning M, Proctor DT, Ragnarsson L, Barbier J, Lavidis NA, Molgó JJ, et al. Glycerotoxin stimulates neurotransmitter release from N-type  $Ca^{2+}$  channel expressing neurons. *J Neurochem.* 2006;98(3):894–904. doi:10.1111/j.1471-4159.2006.03938.x.
- Meunier FA, Nguyen TH, Colasante C, Luo F, Sullivan RKP, Lavidis NA, et al. Sustained synaptic-vesicle recycling by bulk endocytosis contributes to the maintenance of high-rate neurotransmitter release stimulated by glycerotoxin. *J Cell Sci.* 2010;123(7):1131–40. doi:10.1242/jcs.049296.
- Hargreaves AD, Swain MT, Hegarty MJ, Logan DW, Mulley JF. Restriction and recruitment—gene duplication and the origin and evolution of snake venom toxins. *Genome Biol Evol.* 2014;6(8):2088–95. doi:10.1093/gbe/evu166.
- Longenecker HE, Hurlbut WP, Mauro A, Clark AW. Effects of black widow spider venom on the frog neuromuscular junction: effects on end-plate potential, miniature end-plate potential and nerve terminal spike. *Nature.* 1970;225(5234):701–3.
- Frontali N, Ceccarelli B, Gorio A, Mauro A, Siekevitz P, Tzeng MC, et al. Purification from black widow spider venom of a protein factor causing the depletion of synaptic vesicles at neuromuscular junctions. *J Cell Biol.* 1976; 68(3):462–79. doi:10.1083/jcb.68.3.462.
- Garb JE, Hayashi CY. Molecular evolution of  $\alpha$ -Latrotoxin, the exceptionally potent vertebrate neurotoxin in black widow spider venom. *Mol Biol Evol.* 2013;30(5):999–1014. doi:10.1093/molbev/mst011.
- Ouanounou G, Malo M, Stinnakre J, Kreger AS, Molgó J. Trachynilysin, a neurosecretory protein isolated from stonefish (*Synanceia trachynis*) venom, forms nonselective pores in the membrane of NG108-15 cells. *J Biol Chem.* 2002;277(42):39119–27. doi:10.1074/jbc.M203433200.
- Colasante C, Meunier FA, Kreger AS, Molgó J. Selective depletion of clear synaptic vesicles and enhanced quantal transmitter release at frog motor nerve endings produced by trachynilysin, a protein toxin isolated from stonefish (*Synanceia trachynis*) venom. *Eur J Neurosci.* 1996;8(10):2149–56. doi:10.1111/j.1460-9568.1996.tb00736.x.
- Meunier FA, Mattei C, Chameau P, Lawrence G, Colasante C, Kreger AS, et al. Trachynilysin mediates SNARE-dependent release of catecholamines from chromaffin cells via external and stored  $Ca^{2+}$ . *J Cell Sci.* 2000;113(7):1119–25.
- Selander-Sunnerhagen M, Ullner M, Persson E, Telemann O, Stenflo J, Drakenberg T. How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the  $NH_2$ -terminal EGF-like domain in coagulation factor X. *J Biol Chem.* 1992;267(27):19642–9.
- O'Leary JM, Bromek K, Black GM, Uhrinova S, Schmitz C, Wang X, et al. Backbone dynamics of complement control protein (CCP) modules reveals mobility in binding surfaces. *Protein Sci.* 2004;13(5):1238–50. doi:10.1110/ps.03582704.
- Reid KBM, Day AJ. Structure-function relationships of the complement components. *Immunol Today.* 1989;10(6):177–80. doi:10.1016/0167-5699(89)90317-4.

41. Kirkitadze MD, Barlow PN. Structure and flexibility of the multiple domain proteins that regulate complement activation. *Immunol Rev.* 2001;180(1): 146–61. doi:10.1034/j.1600-065x.2001.1800113.x.
42. Schmidt CQ, Herbert AP, Hocking HG, Uhrin D, Barlow PN. Translational mini-review series on complement factor H: Structural and functional correlations for factor H. *Clin Exp Immunol.* 2008;151(1):14–24. doi:10.1111/j.1365-2249.2007.03553.x.
43. Colinet D, Deleury E, Anselme C, Cazes D, Poulain J, Azema-Dossat C, et al. Extensive inter- and intraspecific venom variation in closely related parasites targeting the same host: The case of Leptopilina parasitoids of *Drosophila*. *Insect Biochem Mol Biol.* 2013;43(7):601–11. doi:10.1016/j.ibmb.2013.03.010.
44. Poirié M, Colinet D, Gatti J-L. Insights into function and evolution of parasitoid wasp venoms. *Curr Opin Insect Sci.* 2014;6:52–60. doi:10.1016/j.cois.2014.10.004.
45. Ponting CP, Hofmann K, Bork P. A latrophilin/CL-1-like GPS domain in polycystin-1. *Curr Biol.* 1999;9(16):R585–R8. doi:10.1016/s0960-9822(99)80379-0.
46. Qian F, Wei W, Germino G, Oberhauser A. The nanomechanics of polycystin-1 extracellular region. *J Biol Chem.* 2005;280(49):40723–30. doi:10.1074/jbc.M509650200.
47. Verna J, Lodder A, Lee K, Vagts A, Ballester R. A family of genes required for maintenance of cell wall integrity and for the stress response in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci.* 1997;94(25):13804–9.
48. Levin DE. Cell wall integrity signaling in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev.* 2005;69(2):262–91. doi:10.1128/mmr.69.2.262-291.2005.
49. Cohen-Kupiec R, Broglie KE, Friesem D, Broglie RM, Chet I. Molecular characterization of a novel  $\beta$ -1,3-exoglucanase related to mycoparasitism of *Trichoderma harzianum*. *Gene.* 1999;226(2):147–54. doi:10.1016/S0378-1119(98)00583-6.
50. Gravier C. Contribution à l'étude de la trompe des Glycériens. *Bull sci Fr Bel.* 1898;31:421–48.
51. Oppenheimer A. Certain sense organs of the proboscis of the polychaetous annelid *Rhynchobolus Dibranchiatus*. *Proc Am Acad Arts Sci.* 1902;37(21): 553–62. doi:10.2307/20021708.
52. Raphaël C. Étude de la trompe des Glycères et de son organe excréteur d'hémoglobine. *Travaux de la Station Biologique de Roscoff.* 1933;11:5–18.
53. Michel C. Comparaison des masses prépharyngiennes de la trompe d'*Eulalia viridis* (Mueller) (Phyllocladidae) et des languettes de la trompe de *Glycera convolvata* (Keferstejn) (Glyceridae). *Annélides Polychètes Errantes. Bull Soc Zool France.* 1969;94(2):331–40.
54. von Reumont BM, Campbell LI, Richter S, Hering L, Sykes D, Hetmank J, et al. A polychaete's powerful punch: venom gland transcriptomics of *Glycera* reveals a complex cocktail of toxin homologs. *Genome Biol Evol.* 2014;6(9):2406–23. doi:10.1093/gbe/evu190.
55. Dutertre S, Jin A-H, Vetter I, Hamilton B, Sunagar K, Lavergne V, et al. Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails. *Nat Commun.* 2014;5:1–9. doi:10.1038/ncomms4521.
56. Undheim EAB, Hamilton BR, Kurniawan ND, Bowlay G, Cribb BW, Merritt DJ, et al. Production and packaging of a biological arsenal: Evolution of centipede venoms under morphological constraint. *Proc Natl Acad Sci.* 2015;112(13):4026–31. doi:10.1073/pnas.1424068112.
57. Morgenstern D, King GF. The venom optimization hypothesis revisited. *Toxicon.* 2013;63:120–8. doi:10.1016/j.toxicon.2012.11.022.
58. Gennaro JF, Lorincz JAE, Brewster HB. The anterior salivary gland of the octopus (*Octopus vulgaris*) and its mucous secretion. *Ann N Y Acad Sci.* 1965;118(24):1021–5. doi:10.1111/j.1749-6632.1965.tb40168.x.
59. Undheim EAB, Georgieva DN, Thoen HH, Norman JA, Mork J, Betzel C, et al. Venom on ice: first insights into Antarctic octopus venoms. *Toxicon.* 2010; 56(6):897–913. doi:10.1016/j.toxicon.2010.06.013.
60. Ruder T, Sunagar K, Undheim EAB, Ali SA, Wai T-C, Low DHW, et al. Molecular phylogeny and evolution of the proteins encoded by Coleoid (Cuttlefish, Octopus, and Squid) posterior venom glands. *J Mol Evol.* 2013; 76(4):192–204. doi:10.1007/s00239-013-9552-5.
61. Whitelaw BL, Strugnell JM, Faou P, da Fonseca RR, Hall NE, Norman M, et al. Combined transcriptomic and proteomic analysis of the posterior salivary gland from the southern blue-ringed octopus and the southern sand octopus. *J Proteome Res.* 2016;15(9):3284–97. doi:10.1021/acs.jproteome.6b00452.
62. Fry BG, Roelants K, Norman JA. Tentacles of venom: toxic protein convergence in the kingdom animalia. *J Mol Evol.* 2009;68(4):311–21. doi:10.1007/s00239-009-9223-8.
63. Hu H, Bandyopadhyay PK, Olivera BM, Yandell M. Elucidation of the molecular envenomation strategy of the cone snail *Conus geographus* through transcriptome sequencing of its venom duct. *BMC Genomics.* 2012; 13(1):1–12. doi:10.1186/1471-2164-13-284.
64. Ockelmann KW, Vahl O. On the biology of the polychaete *Glycera alba*, especially its burrowing and feeding. *Ophelia.* 1970;8(1):275–94. doi:10.1080/00785326.1970.10429564.
65. Undheim EAB, Jones A, Clauser KR, Holland JW, Pineda SS, King GF, et al. Clawing through evolution: toxin diversification and convergence in the ancient lineage Chilopoda (Centipedes). *Mol Biol Evol.* 2014;31(8):2124–48. doi:10.1093/molbev/msu162.
66. Meyer M, Kircher M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb Protoc.* 2010; 2010(6):pdb.prot5448. doi:10.1101/pdb.prot5448.
67. Kircher M, Sawyer S, Meyer M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 2012; 40(1):e3. doi:10.1093/nar/gkr771.
68. Renaud G, Kircher M, Stenzel U, Kelso J. freebis: an efficient basecaller with calibrated quality scores for Illumina sequencers. *Bioinformatics.* 2013;29(9): 1208–9. doi:10.1093/bioinformatics/btt117.
69. Renaud G, Stenzel U, Kelso J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* 2014;42(18):e141. doi:10.1093/nar/gku699.
70. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal.* 2011;17(1):10–2.
71. Smeds L, Künstner A. ConDeTri - a content dependent read trimmer for Illumina data. *PLoS ONE.* 2011;6(10):e26314. doi:10.1371/journal.pone.0026314.
72. Ewing B, Green P. Base-calling of automated sequencer traces using *Phred*. II. Error probabilities. *Genome Res.* 1998;8(3):186–94. doi:10.1101/gr.8.3.186.
73. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using *Phred*. I. Accuracy assessment. *Genome Res.* 1998;8(3):175–85. doi:10.1101/gr.8.3.175.
74. Peng Y, Leung HCM, Yiu S-M, Lv M-J, Zhu X-G, Chin FYL. IDBA-tran: a more robust de novo de Bruijn graph assembler for transcriptomes with uneven expression levels. *Bioinformatics.* 2013;29(13):i326–i34. doi:10.1093/bioinformatics/btt219.
75. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389–402. doi:10.1093/nar/25.17.3389.
76. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8(10):785–6. doi:http://www.nature.com/nmeth/journal/v8/n10/abs/nmeth.1701.html#supplementary-information.
77. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 2016;44(D1):D279–D85. doi:10.1093/nar/gkv1344.
78. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci.* 1998;95(11):5857–64.
79. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res.* 2015;43(D1):D257–D60. doi:10.1093/nar/gku949.
80. Stamatakis A. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014. doi:10.1093/bioinformatics/btu033.
81. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics.* 2007;23(1):127–8. doi:10.1093/bioinformatics/btl529.
82. Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* 2011;39 suppl 2:W475–W8. doi:10.1093/nar/gkr201.
83. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 2016. doi:10.1093/nar/gkw290.
84. Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res.* 1995; 23(6):1087–8. doi:10.1093/nar/23.6.1087.
85. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002;3(7):1–12. doi:10.1186/gb-2002-3-7-research0034.
86. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 1995;57(1):289–300.



87. Hoffmann S, Otto C, Doose G, Tanzer A, Langenberger D, Christ S, et al. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biol.* 2014;15(2):R34.
88. Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, Vogel J, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol.* 2009;5(9):e1000502. doi:10.1371/journal.pcbi.1000502.
89. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, et al. Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinformatics.* 2013;14(2):193–202. doi:10.1093/bib/bbs012.
90. Grobe P, Vogt L. Morph.D.Base 2.0: A public data base for morphological data, metadata, and phylogenetic matrices. <http://www.morphdbase.de>. 2009. Accessed 22 Jan 2017.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

