

ORIGINAL RESEARCH

Transcriptomics and differential gene expression in *Whitmania pigra* (Annelida: Clitellata: Hirudinida: Hirudinidae): Contrasting feeding and fasting modes

Muhammad Salabat Khan¹  | De-Long Guan¹ | Sebastian Kvist^{2,3} | Li-Bin Ma¹ | Juan-Ying Xie⁴ | Sheng-Quan Xu¹

¹College of Life Sciences, Shaanxi Normal University, Xi'an, China

²Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada

³Department of Natural History, Royal Ontario Museum, Toronto, Ontario, Canada

⁴School of Computer Science, Shaanxi Normal University, Xi'an, China

Correspondence

Sheng-Quan Xu, College of Life Sciences, Shaanxi Normal University, Xi'an, China.
Email: xushengquan@snnu.edu.cn

Funding information

Shaanxi Normal University, Grant/Award Number: S2015YB03; Fundamental Research Funds for the Central Universities, Grant/Award Number: GK201702010 and GK201701006; the Key Research and Development Program of Shaanxi Province, Grant/Award Number: 2018ZDXMSF-079; NSERC Discovery Grant

Abstract

The medicinal utility of leeches has been demonstrated through decades of use in modern hospital settings, mainly as relievers of venous congestion following flap or digit replantation surgery. In the present study, we sequence and annotate (through BLAST- and Gene Ontology-based approaches) the salivary transcriptome of the nonblood feeding hirudinid *Whitmania pigra* and assess the differential gene expression of anticoagulation factors (through both quantitative real-time PCR [qRT-PCR] and in silico-based methods) during feeding and fasting conditions. This was done in order to evince the diversity of putative anticoagulation factors, as well as estimate the levels of upregulation of genes immediately after feeding. In total, we found sequences with demonstrated orthology (via both phylogenetic analyses and BLAST-based approaches) to seven different proteins that have previously been linked to anticoagulatory capabilities—eglin C, bdellin, granulin, guamerin, hyaluronidase, destabilase I, and lipocalin. All of these were recovered from leeches both in the fasting and in the feeding conditions, but all show signs of upregulation in the feeding leeches. Interestingly, our RNA-seq effort, coupled with a hypergeometric test, indicated that the differentially expressed genes were disproportionately involved in three main immunological pathways (endocytosis, peroxisome regulation, and lysosome regulation). The results and implications of the finding of anticoagulants in this nonblood feeding leech and the putative upregulation of anticoagulation factors after feeding are briefly discussed in an evolutionary context.

KEYWORDS

anticoagulants, differential gene expression, Hirudinea, Hirudinidae, metabolic pathways, transcriptomics, *Whitmania pigra*

1 | INTRODUCTION

Anticoagulation factors work directly, or in concert, to antagonize coagulation cascades that normally act to induce thrombus formation following trauma to the blood vessels. As a result of the anticoagulation capabilities, these factors are also commonly employed in hospital settings—for example, to prevent blood clot formation following stroke (Sanama & Guinet, 2011; Sandercock, Counsell, & Kane, 2015; Weitz & Bates, 2005)—and they are currently available in pill form or as intravenous medication (Garcia, Libby, & Crowther, 2010). These anticoagulants also occur naturally in blood feeding organisms, making the animals themselves of utility to modern medicine. For example, for millennia, leeches have been leveraged for their blood feeding nature and possession of potent anticoagulants (Siddall, Min, Fontanella, Phillips, & Watson, 2011 and references therein) and their frequency of use in medicine continues to increase (Mory, Mindell, & Bloom, 2000). Although hirudin, which binds bivalently to both the exosite and the catalytic site of thrombin thereby preventing the transformation of fibrinogen into fibrin, is arguably the most well-known leech anticoagulation factor, the salivary peptide repertoires of leeches may include over 20 compounds known to be involved in preventing coagulation of blood (Kvist, Brugler, Goh, Giribet, & Siddall, 2014). These include factor Xa inhibitors, thrombin inhibitors, inhibitors of platelet aggregation, protease inhibitors, and leukocyte elastase inhibitors, to name a few. In addition, some of these anticoagulation factors are extremely potent, active in picomolar concentrations (Dang & Cera, 1994), and some bind irreversibly to their target proteins, suggesting the selective advantage that they confer to the leech, to enable extended periods of feeding.

With an increase in evidence from independent investigations, we now believe that the ancestral leech was blood feeding, or at least possessed anticoagulation factors that presumably aided in the leeches' feeding biology (Siddall, Brugler, & Kvist, 2016; Siddall et al., 2011). If the ancestral leech was blood feeding, one could reasonably expect to find signatures of this ancestry in the transcriptomic makeup of nonblood feeding leeches, as well as the finding of orthologous anticoagulants in both leeches that share an ancient common ancestor and those that share more derived common ancestors. Indeed, both of these scenarios have been shown to be true for leeches, with the possession of leech antiplatelet proteins (LAPP) by the nonblood feeding leech, *Helobdella robusta* Shankland, Bissen & Weisblat, 1992 (Kvist, Sarkar, & Siddall, 2011), and the possession of hirudin orthologues in glossiphoniid and hirudinid leeches, the common ancestor of which existed relatively early in leech evolution (Siddall et al., 2016). However, nonblood feeding leeches from across the phylogeny need to be queried in order to robustly infer if these possess anticoagulation factors. This will both shed light on the diversity and evolution of leech anticoagulation factors more generally and will help elucidate the specific feeding habit of the ancestral leech. In addition, our knowledge regarding the expression of anticoagulants during

periods of digestion (i.e., fasting) and their potential upregulation during feeding is still almost nonexistent.

The Asian freshwater leech *Whitmania pigra* (Whitman, 1884) is nonblood feeding, despite the placement of this genus within the family Hirudinidae (Phillips & Siddall, 2009), which also includes the European medicinal leech, *Hirudo medicinalis* (Linnaeus, 1758) and several other blood feeding species. Instead, the leech is considered macrophagous, suggesting that it commonly swallows or takes bites out of prey sources (Sawyer, 1986). In traditional Chinese medicine, *W. pigra* is known to promote blood circulation and relieve blood clotting (Jia-Xin, Xing-Chang, & Qiu-Hai, 1994; Zhong, Yang, & Cui, 2008) and it can therefore be reasonably deduced that this leech may possess anticoagulation factors despite of its nonblood feeding nature. However, large-scale transcriptomic evidence for such a claim is sorely lacking, despite the previous sequencing of a transcriptome from an individual of *W. pigra* (Liu et al., 2018); that study focused on environmental adaptations of leeches, rather than the anticoagulant repertoire. To clarify whether or not this nonblood feeding leech secretes anticoagulants, which anticoagulants, and to shed light on the timing of their putative expression, the present study sequenced full body transcriptomes of *W. pigra*, both during feeding and during periods of digestion (fasting), followed by differential gene expression analyses.

2 | MATERIAL AND METHODS

2.1 | Sample preparation

Specimens of *W. pigra* (see Figure 1) were provided by the Institute of Zoology Lab, Shaanxi Normal University, China, with live snails delivered regularly as a food source. The specimens were divided into two test groups: fasting ($n = 3$) and feeding ($n = 3$). After allowing the leeches to acclimatize to their new environment for 1 month, specimens in the fasting group were deprived of food for 2 weeks, whereas specimens in the feeding group were continuously fed live snails. Prior to RNA extraction, all leeches were rinsed sequentially in a bleach solution and deionized water to reduce the risk of contamination by surface bacteria. The samples were stored in a liquid nitrogen facility until RNA extraction.

2.2 | RNA isolation and library preparation

Total RNA was extracted both from the whole body of each of the specimens and also directly from salivary tissues using the Qiagen RNA Kit (Qiagen Biotechnology, Hilden, Germany) following the manufacturer's instructions. RNA degradation and/or putative contaminations were checked by gel electrophoresis on a 1% agarose gel. Further, the purity, concentration, and integrity of the RNA samples were tested using a RNA Nano 6000 Assay Kit on an Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Libraries were constructed for samples that passed initial degradation tests. In total, 3 μg of total RNA for each sample was used for the RNA sample preparations. Briefly, eukaryotic mRNA was



FIGURE 1 Live habitus of *Whitmania pigra*. Photograph by M. Salabat Khan

enriched using poly-T oligo-attached magnetic beads and subsequently randomly fragmented in fragmentation buffer (Biomarker, Beijing, China). Then, mRNA was separated from the total RNA sample by divalent cations under high temperature and the mRNA was used as the template for cDNA synthesis using random hexamers. Second-strand cDNA fusion was achieved using DNA polymerase I with the addition of RNase, and double-stranded cDNA was thereafter purified using AMPure XP beads (Beckman Coulter, California, USA). Complementary DNA fragments of preferentially 150–200 bp length were selected with the AMPure XP system. Index-coded sample clustering was achieved on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina Inc., California, USA). After cluster formation, paired-end sequencing was performed on the Illumina HiSeq 4000 platform.

2.3 | Quality control

The raw data were principally treated with in-house perl scripts. At this stage, adapter sequences were removed and reads including poly-N regions, as well as low-quality reads (PHRED scores below 30), were eliminated from the dataset. Simultaneously, GC content of the clean data was calculated. All downstream analyses were based on the resulting, quality-controlled data.

2.4 | Transcriptome assembly and annotation

A reference genome-guided assembly against the genome of *H. robusta* was performed for each of the paired-end transcriptomes using Trinity (Grabherr et al., 2011), and applying default settings. Transdecoder ver. 5 (<https://github.com/TransDecoder/TransDecoder>) was used to identify putative coding regions within transcripts. Initial transcript identities (some of which were later confirmed by phylogenetic analyses) were assessed by BLASTing (using the BLASTn algorithm) against the following databases: GenBank nr (nonredundant sequence database), Pfam, and Swiss-Prot. Gene ontology terms were then established using the topGO R

package (Aibar, Fontanillo, Droste, Rivas, & J., 2015), with analyses based on the Kolmogorov–Smirnov test (Antoneli, Passos, Lopes, & Briones, 2018).

2.5 | Identification of differentially expressed genes

Following the results from the RNA-seq data, in silico differential expression analysis of the two conditions (fasting and feeding) was performed using DESeq ver. 1.10.1 (Anders & Huber, 2010) in R (R Core Team, 2017) for the six transcriptomes resulting from the RNA sequencing. The Pearson correlation coefficient was used as an indicator of biological repeat correlation between the different samples. The screening criteria involved fold changes ≥ 2 or ≤ 0.5 , and false discovery rate (FDR) $< 10\%$ (see Anders & Huber, 2010), using normalized counts of transcripts as input. The resulting p -values were adjusted using the Benjamini and Hochberg (Benjamini & Hochberg, 1995) procedure for controlling for false discovery rates. Genes for which an adjusted p -value ≤ 0.05 was recovered were interpreted as differentially expressed.

As corroboration of the DESeq analysis, we performed SYBER GREEN-based, quantitative real-time PCRs (qRT-PCR) of select proteins, in order to robustly infer the differential gene expression that was proposed by the in silico method. Reverse transcription and cDNA synthesis used 300 ng of total mRNA. Primers (Supporting Information Table S1) were designed from the whole-body transcriptome data employing Primer-BLAST (Ye et al., 2012). Seven putative anticoagulant proteins were chosen for qRT-PCR validation (eglin C, bdellin, granulins, guamerin, hyaluronidase, destabilase I, and lipocalin). The qRT-PCR was run on a CFX96 real-time PCR detection system (Bio-Rad, California, USA), and statistical analysis of the results was performed using a two-way ANOVA. A total of six biological replicates (three for fasting leeches and three for recently fed leeches) were used to corroborate the results. Expression levels of target genes (the seven putative anticoagulation factors) were calculated with respect to mean RT-PCR, by quantifying the cycle threshold (C_T) value. The ΔC_T method performs pairwise, relative expression level comparisons of genes, and this method was used herein for normalization of qRT-PCR data.

A pathway annotation analysis of the differentially expressed genes was employed via a BLAST-based search against the Kyoto Encyclopaedia of Genes and Genomes (KEGG), and a hypergeometric test (Fisher's exact test; see, e.g., Goeman & Bühlmann, 2007) was used to identify any potential skew of the differentially expressed genes toward certain biological pathways (e.g., immunology).

2.6 | Amino acid alignments and phylogenetic analyses

Translated nucleotide sequences for each of the recovered anticoagulation factors were aligned together with a single “archetypal” anticoagulant. The online version of MAFFT ver. 7 (Kato & Standley, 2013) was employed with default settings, and the results were visualized using Jalview ver. 2.10.2 (Waterhouse, Procter, Martin,

TABLE 1 Sample details resulting from the RNA-seq approach. WP-1, WP-2, and WP-3 represent leeches in fasting mode, whereas WP-4, WP-5, and WP-6 represent leeches in feeding mode

Sample ID	Raw reads	Clean reads	Raw bases (Gb)	Clean bases (Gb)	Effective rate (%)	Q20 (%)	Q30 (%)	GC (%)
WP-1	46835484	44106896	7.03	6.62	94.17	98.17	95.36	43.71
WP-2	48362626	46625030	7.25	6.99	96.41	97.96	94.95	42.90
WP-3	51222768	49069134	7.68	7.36	95.80	98.01	95.07	42.73
WP-4	47602144	45290524	7.14	6.79	95.14	98.18	95.42	43.43
WP-5	46859934	44655484	7.03	6.70	95.30	98.12	95.33	42.91
WP-6	47742030	45942546	7.16	6.89	96.23	97.85	94.73	43.37

Note. The Q20 and Q30 columns indicate the percentage of overall base calls with a PHRED score ≥ 20 or ≥ 30 , respectively; the GC column indicates the GC content across all reads.

Clamp, & Barton, 2009). From this alignment, percent similarity was calculated and putative conservation of the cysteines, involved in disulfide bonds and protein folding, was estimated. The online platform for SignalP ver. 4.1 (Petersen, Brunak, Heijne, & Nielsen, 2011)

was used to estimate the presence of a signal peptide for each of the putative anticoagulant sequences.

In addition, phylogenetic analyses were performed for a larger subset of sequences (these were recovered from previous leech

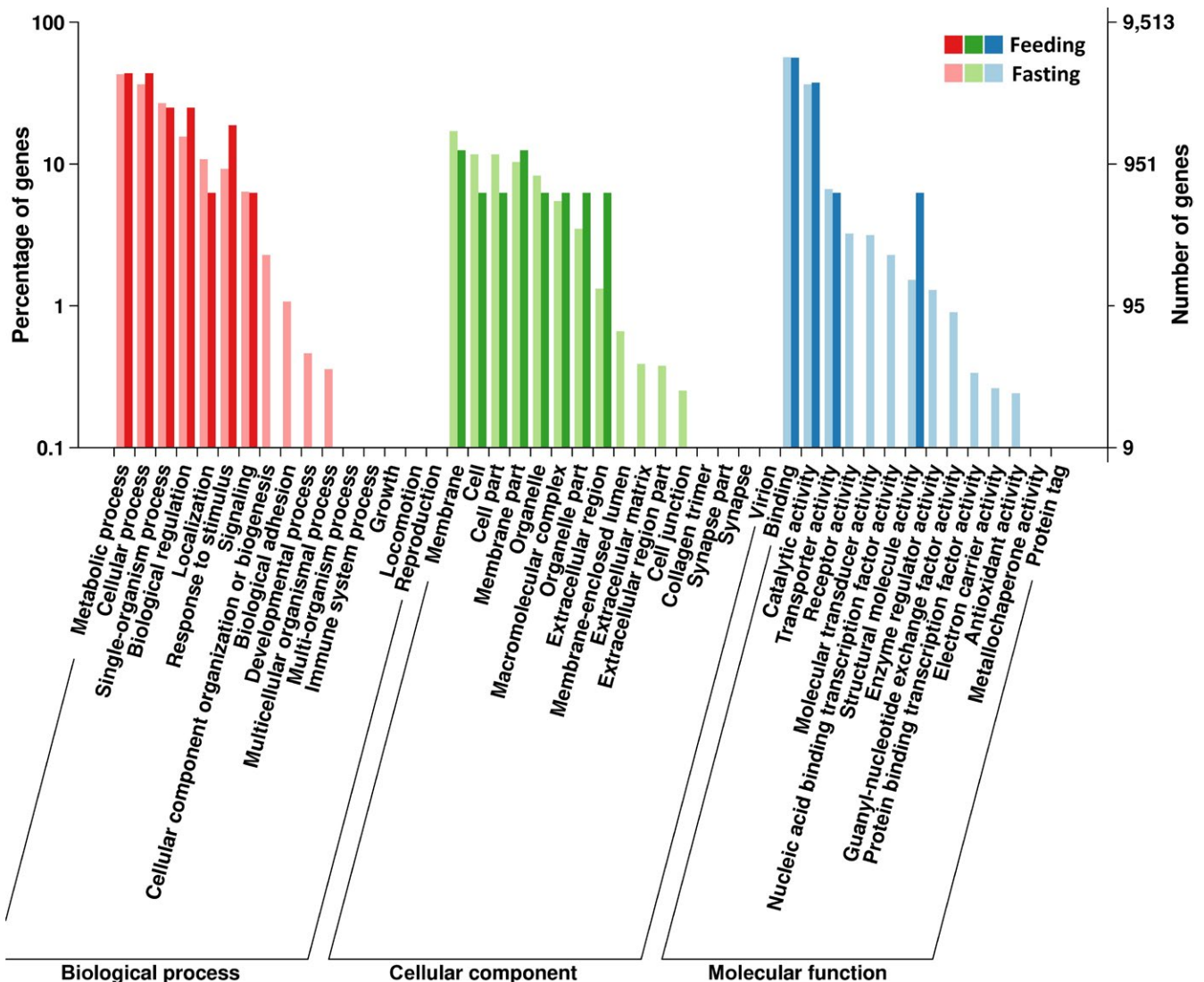


FIGURE 2 Results from the gene ontology analysis (examined using topGO), illustrating the log scale proportional devotion of the salivary transcripts to various processes regarding biological process, cellular component, and molecular function

anticoagulant studies (e.g., Min, Sarkar, & Siddall, 2010; Kvist, Min, & Siddall, 2013; Kvist et al., 2014; Siddall et al., 2016; Tessler, Marancik et al., 2018) including several putative orthologs beyond the archetypal sequences (the putative guamerin sequence from *W. pigra* was aligned with several known, putatively paralogous factor Xa inhibitors derived from leeches, in order to tease out the most likely ortholog for this sequence). Phylogenetic analyses were performed in order to robustly infer orthology between the newly derived sequences and the archetypal anticoagulants. Again, MAFFT ver. 7 was employed with default settings. To maximize homology and avoid detrimental effects by the partial nature of the comparative transcriptomic data, the alignments were cut to reflect the coverage of the specific genes for the *W. pigra* data. Phylogenetic analyses were performed in TNT ver. 1.5 (Goloboff, Farris, & Nixon, 2008) under the maximum parsimony criterion with gaps treated as missing data. For this purpose, 1,000 new technology replications with five rounds of ratcheting were employed, stipulating that the best tree should be found 10 times before terminating the search ("hits 10"). The results were then subjected to TBR branch swapping using the "bbreak" command. Support values were estimated using 100 rounds of standard bootstrapping. All trees were left unrooted due to the absence of an obvious outgroup for leech anticoagulants.

3 | RESULTS

In total, we recovered 41.35 Gb of clean data from all specimens combined and predicted a total of 22,632 genes with mean length of 3,397 bp and with 95.14% of transcripts showing PHRED scores above 30. The mean GC content for the combined pool of fasting and feeding individuals was 43.17%. Sample details are provided in Table 1, and detailed description of transcript lengths is shown in Supporting Information Figure S1.

3.1 | GO analysis

Functional categories (cellular component, molecular function, and biological process) for gene ontology statements were assigned to

each of the 22,632 gene regions (see Figure 2). In the biological process category, the four most common processes were metabolic process, cellular process, single-organism process, and biological regulation. Within the cellular component category, the most enriched components were membrane, cell, cell part, organelle, and membrane part. In the molecular function category, the most highly expressed functions were binding activity, catalytic activity, transporter activity, and receptor activity.

3.2 | Anticoagulant alignments

Despite the nonblood feeding nature of *W. pigra*, we found transcripts that show significant BLAST hits (e-values below E^{-5}) against seven different proteins that have previously been linked to anticoagulation. These were eglin C, bdellin, granulin, guamerin, hyaluronidase, destabilase I, and lipocalin (Table 2). Each of these transcripts was reciprocally BLASTed against GenBank nr to ensure that they did not match any other protein at a lower e-value.

For eglin C, the amino acid alignment (Figure 3a) of the newly acquired sequence with the archetypal variant (GenBank accession number O905140A; Knecht et al., 1983), originally isolated from an unidentified leech, showed 61% similarity between shared amino acid positions. Both sequences are devoid of cysteine residues, suggesting the nonfolding nature of the protein.

The amino acid alignment for bdellin (Figure 3b) indicated that the sequence derived from *W. pigra* shares 70% similarity (in positions without gaps) with the archetypal anticoagulant, originally isolated from salivary tissue of *H. medicinalis*; GenBank accession number P09865 (Fink et al., 1986). Of particular import, the six cysteine residues of the mature peptide (i.e., downstream from the signal peptide) that are responsible for the disulfide bonds and, therefore, the folding capability of the protein are fully conserved between the two sequences.

The amino acid alignment for our granulin sequence (Figure 4a) together with the archetypal anticoagulant (GenBank accession number ABB17975; see He & Bateman, 2003) suggested that only a partial sequence had been generated for the present study; we recovered only 294 out of the estimated 1,405 amino acids. Regardless, for shared amino acid sites, our sequence shows 80%

TABLE 2 Summary statistics for the seven putative anticoagulation factors found in the transcriptome of *Whitmania pigra*, as well as the details of the differential expression between feeding and fasting modes

Name	Biological actions	Normalized count (Fasting)	Normalized count (Feeding)	Log2 fold change	Up/Down
Lipocalin	Platelet aggregation inhibitor	2.357	749.58	2.64447	UP
Granulin	Thrombin inhibitor, Elastase inhibitor	35.365	171.441	1.71898	UP
Hyaluronidase	Hyaluronic acid cleavage	10.548	29.450	1.2631	UP
Bdellin	Trypsin and plasmin inhibitor	46.207	83.891	0.30256	UP
Eglin	Inhibitor of elastase/chymotrypsin	2.866	37.452	3.09321	UP
Destabilase	Fibrin depolymerase factor	5,062.76	18,880.8	1.34167	UP
Guamerin	Elastase inhibitor	175.43	1,145.28	1.2631	UP

Note. DESeq was used to calculate the normalized counts for transcripts, and these counts were used to estimate the Log2 fold change.

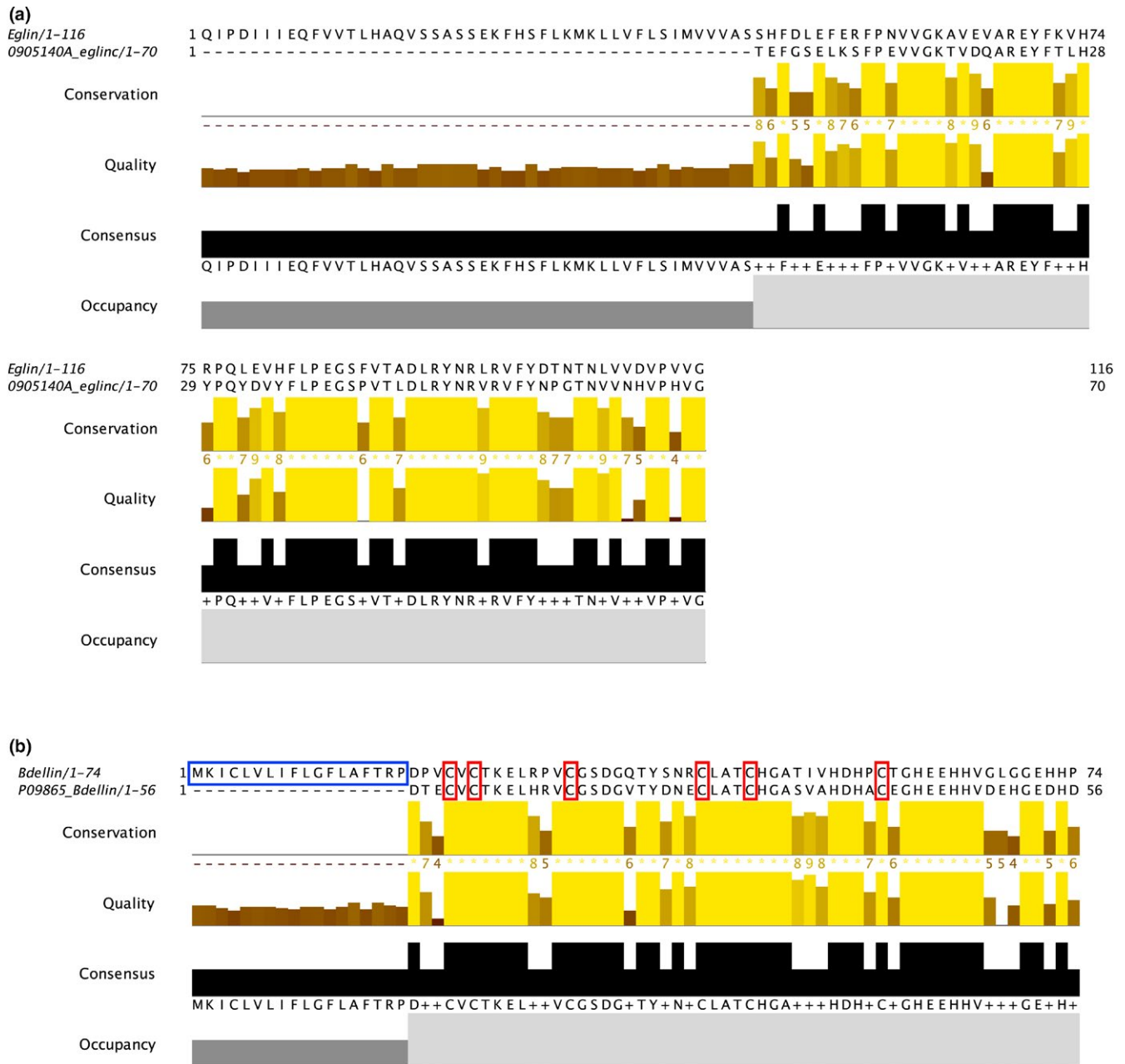


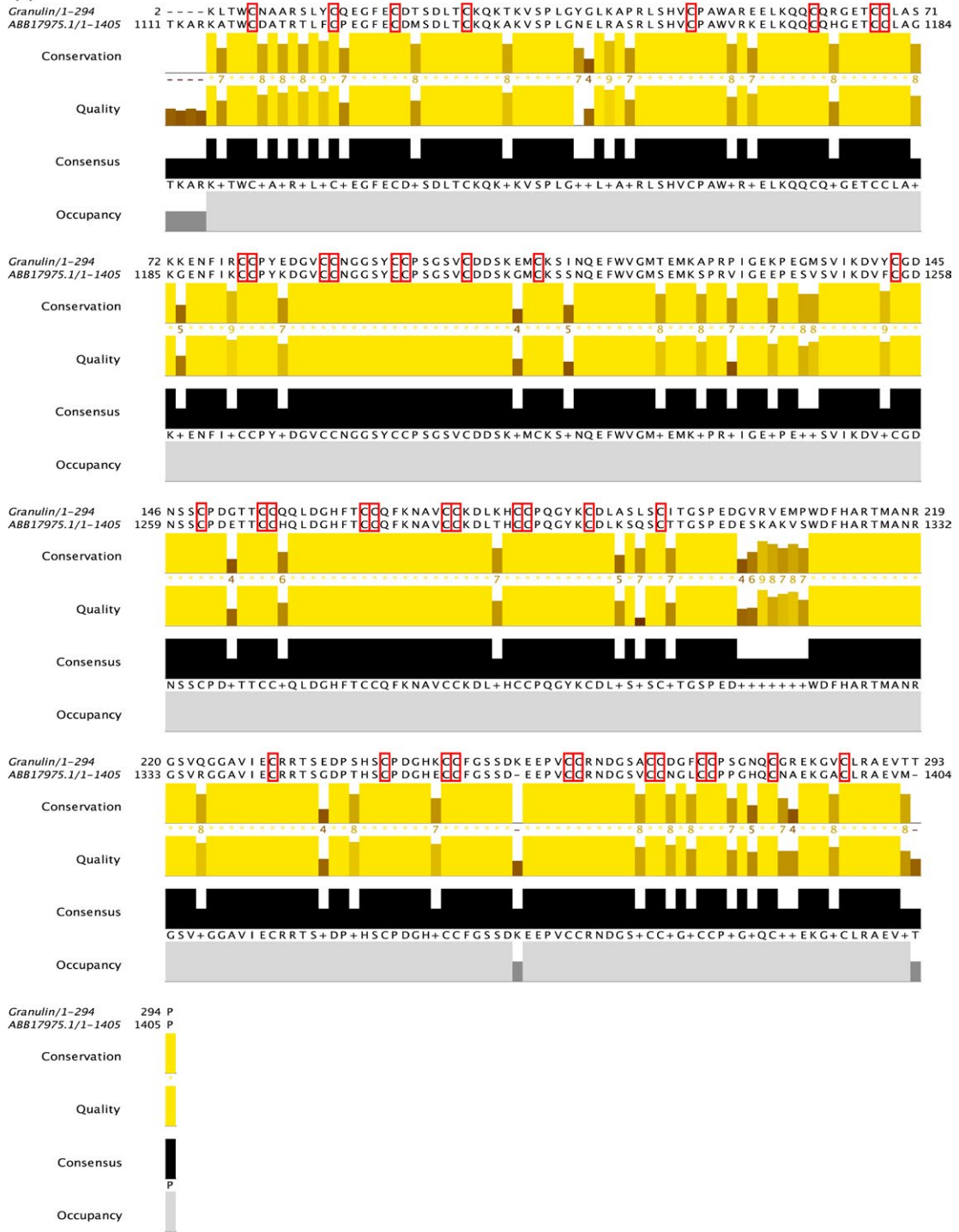
FIGURE 3 MAFFT-based alignment of the amino acid sequences for (a) eglin C and (b) bdellin putative orthologues. The newly acquired, *Whitmania pigra*-derived sequence is shown above, and the “archetypal” sequence below in the alignments. Blue boxes denote the predicted signal peptide region, and red boxes denote fully conserved cysteine residues

sequence similarity with progranulin derived from *H. medicinalis*. Within the homologous region of the mature peptide, the 40 cysteine residues (including 12 sets of two consecutive cysteines) are fully conserved between the sequences.

For guamerin, the amino acid alignment (Figure 4b) shows 43% amino acid identity between the sequence from *W. pigra* and the archetypal sequence, derived from *Hirudo nipponia* (GenBank accession number AAD09442; Jung, Kim, Ha, Joe, & Kang, 1995). Despite of this relatively low similarity, 10 out of the 12 cysteine residues show conservation and a relatively long stretch of homopolymers exist in the alignment (5'-CPNGFKVDENGCEYPCTC-3').

For hyaluronidase (Figure 5), the archetypal anticoagulant sequence for manillase, recovered from the US Patent application US7, 049, 124 B1, was aligned with the newly generated sequence. Manillase has demonstrated endoglucuronidase/hyaluronidase activity and is one of the few leech-derived proteins of this family of genes. Oddly, the alignment includes two large regions of gaps, suggesting either that the assembly may have inadvertently created chimeric sequences or that two insertion events have happened for these sequence (one in the archetypal sequence and one in the newly generated sequence). Regardless, for shared amino acid positions, the two sequences show 40% similarity. Of

(a)



(b)

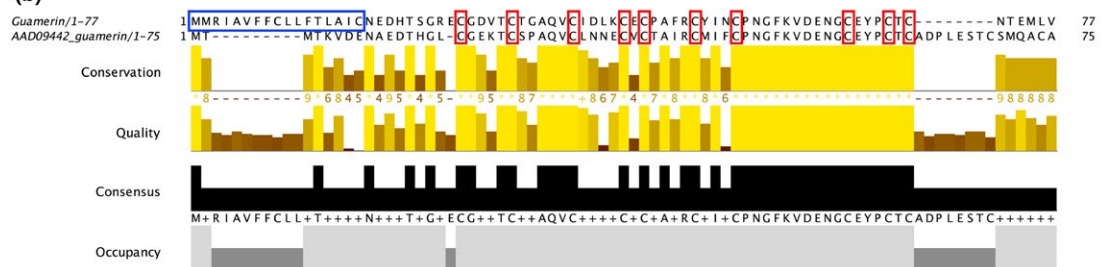


FIGURE 4 MAFFT-based alignment of the amino acid sequences for (a) granulin and (b) guamerin putative orthologues. The newly acquired, *Whitmania pigra*-derived sequence is shown above, and the “archetypal” sequence below in the alignments. Blue boxes denote the predicted signal peptide region, and red boxes denote fully conserved cysteine residues

the four cysteine residues present in the archetypal sequence, only a single residue is fully conserved (note that two cysteines occur in the region of gaps). Previous studies have found similar values of amino acid similarity within this relatively long molecule (Kvist et al., 2014).

Our newly generated sequence of putative destabilase I shows 62% amino acid identity with the archetypal sequence (GenBank accession number AAA96144; Zavalova et al., 1996) isolated from *H. medicinalis* (Figure 6a). The 14 cysteine residues that are present in both sequences show full conservation.

Finally, in the amino acid alignment for lipocalin (Figure 6b), including our newly generated sequence and an archetypal sequence isolated from *H. robusta* (GenBank accession number ESO09778; Simakov et al., 2013), the two sequences show 68% amino acid identity. In addition, two out of the three cysteine residues present in the archetypal sequence (note the uneven number of residues) are conserved in the alignment and relatively long stretches of homopolymers are present, suggesting orthology between the sequences.

3.3 | Phylogenetic trees

The phylogenetic trees resulting from the analyses of the newly sequenced putative anticoagulant sequences and their potential allies are presented in Figure 7a–g.

For the eglin C tree (Figure 7a; length: 171, consistency index [CI]: 0.971, retention index [RI]: 0.792), the newly derived sequence forms a clan (equivalent to a monophyletic group in a rooted tree; see Wilkinson, McInerney, Hirt, Foster, & Embley, 2007) with the archetypal sequence and a sequence derived from *H. medicinalis* (bootstrap support [BS] = 100%). Although the length of the branch leading to the two allies is relatively long, this result suggests orthology between the new sequence and true eglin C.

Much the same scenario is recovered for the tree resulting from the bdellin alignment (Figure 7b; length: 220, CI: 0.879, RI: 0.797) in which the new sequence forms a clan (BS = 100%) with the archetypal sequence, as well as several bdellin orthologues derived from *Haemadipsa interrupta* (Moore, 1935), *Macrobodella decora* Say, 1824, *H. medicinalis*, and *H. nipponia*. Again, the length of the branches within this clan is relatively long, but still portrays an image of orthology between the newly derived sequence and bdellin *sensu stricto*.

Due to the paucity of comparative data, only four sequences were included in the granulin alignment. Despite this, our newly derived sequence forms a clan with the archetypal sequence (BS = 100%) in the tree (Figure 7c; length: 173, CI: 1.00, RI: 1.00). This strongly suggests that our newly derived sequence is indeed an ortholog of granulin.

The tree resulting from the factor Xa inhibitor dataset (Figure 7d; length: 1,093, CI: 0.837, RI: 0.649) shows a tight grouping between the newly derived sequence and the archetypal sequence for

guamerin, as well as a putative ortholog derived from *Aliolimnatis fenestrata* (Moore, 1939; BS = 100%). This suggests that the sequence derived from *W. pigra* is indeed a guamerin ortholog and, because of the inclusion of several other factor Xa inhibitors, we can rule out an affinity to any of the other known anticoagulants (e.g., antistasin, therostasin, ghilanten).

The hyaluronidase dataset included exclusively sequences labeled as mannase. In the resulting tree (Figure 7e; length: 906, CI: 0.959, RI: 0.825), the newly acquired sequence forms a clan with the archetypal sequence (BS = 100%), with a relatively short branch length. This strongly suggests orthology between the two sequences.

A potentially different scenario is recovered for the destabilase I tree (Figure 7f; length: 521, CI: 0.769, RI: 0.618) wherein the newly acquired sequence nests with two putative orthologous sequences derived from *M. decora* and this clan does not include the archetypal sequence (BS < 50%). However, this clade is the adjacent group (equivalent to sister group relationship in a rooted tree; see Wilkinson et al., 2007) to a clan including the archetypal sequence. The clan constructed from these two subclans shows 100% bootstrap support. It therefore seems likely that all of the sequences included in this larger clan are orthologous with destabilase I. This is further supported by the strong sequence similarity seen in the alignments mentioned above.

Finally, the lipocalin tree (Figure 7g; length: 199, CI: 1.00, RI: 1.00) included only four putatively orthologous sequences, due to the dearth of comparative data for leeches. In fact, no “archetypal sequence” was included in the dataset, because this protein has seldom been sequenced for leeches. Our newly acquired sequence forms a tight clan (BS = 100%) with a hypothetical protein derived from *H. robusta*. Upon further inspection of this sequence, the protein product is labeled “lipocalin”, and this annotation is likely derived from sequence similarity with human- or mouse-derived versions of this protein. The tree suggests orthology between the new sequence and the *H. robusta* version, but it is still unclear whether or not these are indeed lipocalin orthologs.

3.4 | Differential gene expression

The DESeq *in silico* analysis identified strong evidence for the differential expression of proteins relating to several immunological pathways, but three pathways showed a distinctly higher number of differentially expressed genes: endocytosis, lysosome regulation, and peroxisome regulation (Supporting Information Table S2). Each of these pathways was recovered as statistically significantly ($p < 0.05$) upregulated when compared to other pathways. In total, 58 genes were found to be differentially expressed between the fasting and feeding modes, 31 of which were upregulated and 27 downregulated in the feeding mode of the leech. Seven of the 58 upregulated genes have been directly linked

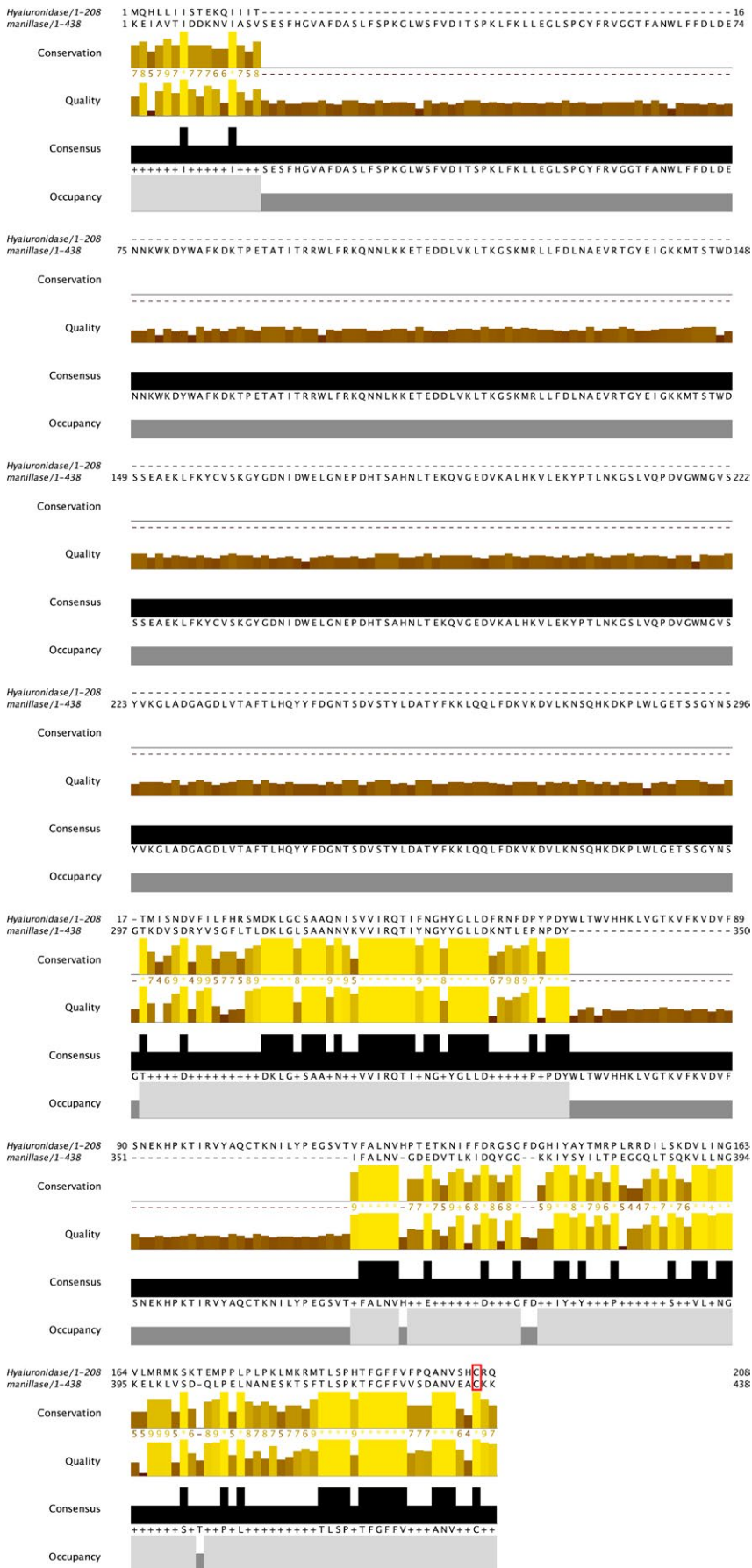


FIGURE 5 MAFFT-based alignment of the amino acid sequences for the hyaluronidase manillase. The newly acquired, *Whitmania pigra*-derived sequence is shown above, and the “archetypal” sequence below in the alignment. The red boxes denote the only fully conserved cysteine residue. Note the long gapped region toward the 5′-end of the sequence (see text for further details)

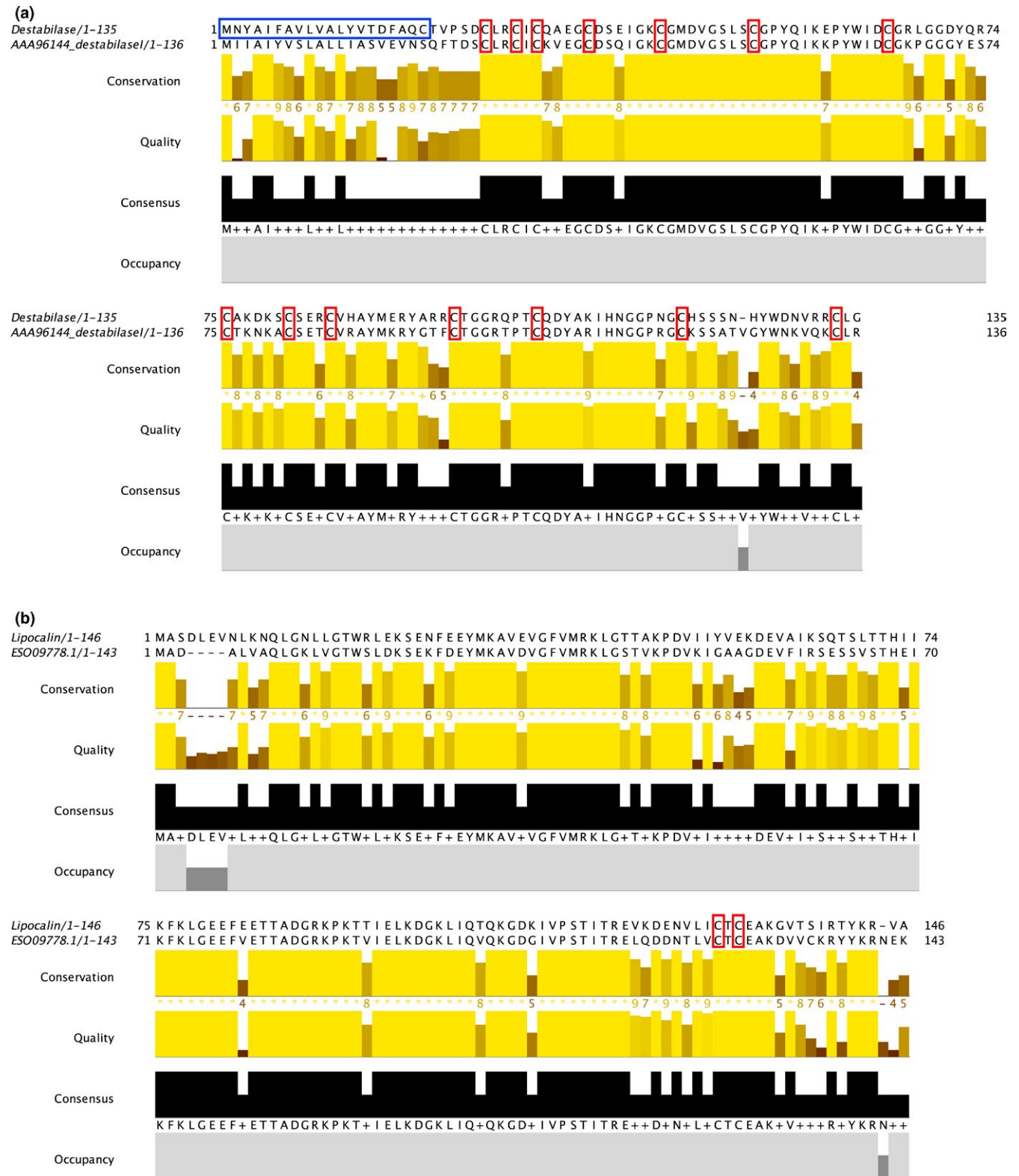


FIGURE 6 MAFFT-based alignment of the amino acid sequences for (a) destabilase I and (b) lipocalin putative orthologues. The newly acquired, *Whitmania pigra*-derived sequence is shown above, and the “archetypal” sequence below in the alignments. Blue boxes denote the predicted signal peptide region, and red boxes denote fully conserved cysteine residues

to anticoagulation and the expression levels of these were therefore further scrutinized using qRT-PCR. For each of the putative anticoagulation proteins, leeches that were allowed to feed continuously show a

marked increase in expression levels in the qRT-PCR analysis (Table 2; Figure 8). The normalized log2 fold (see Love, Huber, & Anders, 2014) expression level change between the anticoagulation factors from

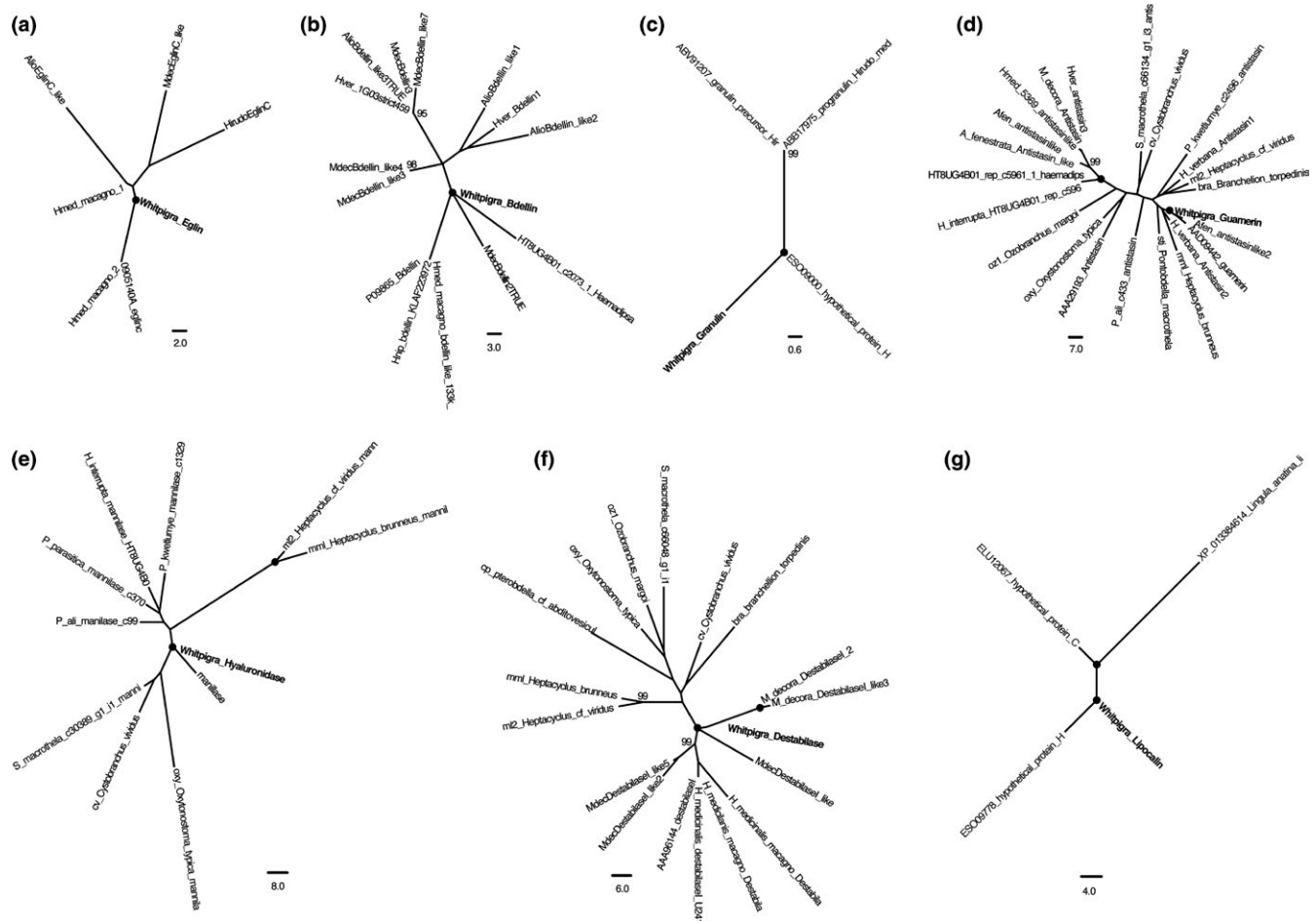


FIGURE 7 Single most parsimonious trees for leech-derived putative orthologues of (a) eglin C (length: 171, CI: 0.971, RI: 0.792), (b) bdellin (length: 220, CI: 0.879, RI: 0.797), (c) granulin (length: 173, CI: 1.00, RI: 1.00), (d) antistatin (length: 1,093, CI: 0.837, RI: 0.649), (e) manillase (length: 906, CI: 0.959, RI: 0.825), (f) destabilase (length: 521, CI: 0.769, RI: 0.618), and (g) lipocalin (length: 199, CI: 1.00, RI: 1.00). The newly acquired sequences are denoted in bold font. CI: consistency index; RI: retention index.

feeding and fasting modes range between 0.3 and 3.1, suggesting an upper limit of 800% upregulation of the expression.

4 | DISCUSSION

Through whole-body transcriptomics and qRT-PCR of known anticoagulation factors, our results indicate that (a) anticoagulation factors are expressed and secreted by the nonblood feeding leech *W. pigra* and (b) expression of several genes, most dramatically the anticoagulation factors, seems to be up- and/or downregulated depending on if the leech is in feeding or fasting mode. We found sequences with putative orthology to seven anticoagulation factors in our transcriptomic dataset—eglin C, bdellin, granulin, guamerin, hyaluronidase, destabilase I, and lipocalin. These factors antagonize the coagulation cascade in diverse ways, suggesting that *W. pigra* is capable of relatively effective anticoagulation. For example, eglin C is an inhibitor of elastase and cathepsin G (Baici & Seemüller, 1984; Seemüller, Fritz, & Eulitz, 1981), which are both involved in inflammatory processes and blood clotting disorders. Guamerin and bdellin act as both elastase inhibitors and as inhibitors of trypsin-like serine protease inhibitors,

thereby inhibiting the conversion of prothrombin to thrombin (Jung et al., 1995; Kim, Hong, Ha, Joe, & Kang, 1996; Salzet, 2001). Previous studies have shown putative orthology between guamerin and other strictly factor Xa inhibitors (e.g., Kvist et al., 2014), which is why this sequence was included in the factor Xa dataset used herein. With much the same result, destabilase I is capable of cleaving fibrin cross-links (Salzet, 2001). By contrast, little is known about leech hyaluronidases, except that they are capable of cleaving hyaluronic acids (Kordowicz, Gussow, Hofmann, Pacuszka, & Gardas, 2006) and are on occasion mentioned as possible treatments for glaucoma (Min et al., 2010). In vitro studies of hyaluronic acid have demonstrated its potential of binding to collagen, the result of which is an increased coagulatory capability of the complex (Wu et al., 2008). Its specific utility in the salivary cocktail of leeches has yet to be realized, but it is likely that hyaluronidase performs similar tasks in leech saliva. Calin acts antagonistically against the coagulation cascade by blocking von Willebrand factor-mediated binding of platelets to subendothelial collagen, thereby preventing the secretion of a granular matrix by collagen (Siddall et al., 2011). This is very similar to the function of leech antiplatelet proteins (LAPPs) and saratin (Kvist et al., 2011; Salzet, 2001), both of which have often been recorded from leech salivary

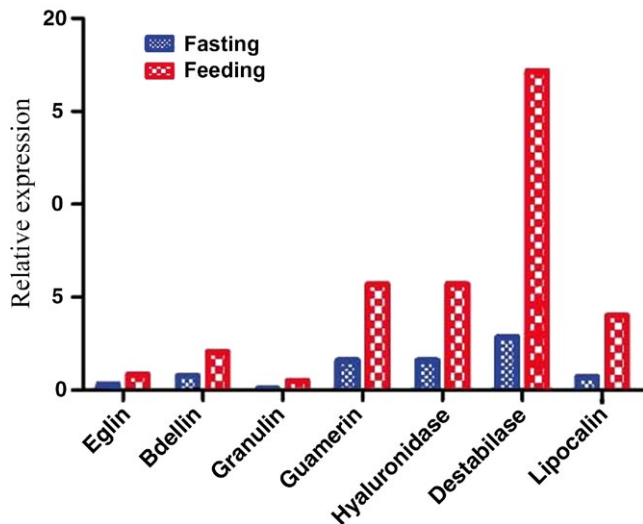


FIGURE 8 Relative expression of the seven putative anticoagulation factors as evinced by the quantitative real-time PCR analysis. Red color indicates the expression of genes in leeches that were allowed to feed, and blue color indicates the expression of genes in fasting leeches

extracts. Granulin, much like hirudin, directly inhibits thrombin and has previously been recovered from the hirudinid leech *H. nipponia* (Hong & Kang, 1999). The molecule is rather rich in cysteine residues ($n = 12$), suggesting a more complex folding structure than hirudin ($n = 6$). It is still unknown whether or not granulin mimics the irreversible effects of hirudin on the fibrin synthesis pathway by binding to both the catalytic site and the exosite of thrombin. During a normal thrombus formation, collagen, ADP, and TXA_2 act in concert to activate signaling pathways and generating secondary messengers, all of which leads to the expression of integrins (Francischetti, Ribeiro, Champagne, & Andersen, 2000). Lipocalin, which was also recovered in our analyses of *W. pigra*, specifically binds to ADP, thereby halting the ADP-induced pathways of platelet aggregation (Francischetti et al., 2000). Lipocalin has previously been recovered from a variety of blood feeding and nonblood feeding organisms (Francischetti, 2010; Koh & Kini, 2008).

In combination with our knowledge of leech phylogenetics (e.g., Siddall et al., 2011; Tessler, de Carle et al., 2018), our results suggest that nonblood feeding leeches possess anticoagulation factors by virtue of their presence in a common ancestor. This hypothesis is gaining traction among hirudinologists and our study greatly increases our knowledge of the salivary components of nonblood feeding leeches; only very few studies have investigated the salivary peptide cocktails of nonblood feeding leeches (e.g., Kvist et al., 2011; Liu et al., 2018). The phylogenetic placement of *Whitmania* inside Hirudinidae and, further, rendering the genus *Hirudo* paraphyletic by virtue of nesting within (de Carle, D., Oceguera-Figueroa, A., Tessler, M, Phillips, A. J., Siddall, M. E. & Kvist, S. Unpublished data), may be suggestive of a relatively recent transition from blood feeding to a macrophagous lifestyle. To this end, one could argue that *W. pigra* should maintain a higher expression level of anticoagulants than, for example, the distantly related *H. robusta* (see Kvist et al., 2011). This

hypothesis is corroborated by the present study, due to the finding of seven putative anticoagulation factors as compared to only a single anticoagulant (LAPP) robustly inferred to be expressed by *H. robusta*.

Our resulting amino acid alignments and phylogenetic analyses suggest a close affinity between the seven anticoagulants found in *W. pigra* and the “archetypal” anticoagulants isolated from a range of other leech species. The phylogenetic analyses are vital in uncovering potential paralogues among the sequences, yet it is troublesome that no obvious outgroups exist for leech anticoagulation factors, such that no root can be applied. Therefore, it is difficult to decipher the direction of the evolution of the amino acid structures within leeches. However, it seems safe to say that, from our phylogenetic analyses, the anticoagulation factors expressed and secreted by *W. pigra* are orthologous with their archetypal counterparts.

Beyond the recovered anticoagulation factors, we also found significant matches against proteins that are utilized in important immunological pathways. Results from our *in silico* analyses of differential expression suggest that a disproportionate amount of genes relating to the following pathways show a distinct upregulation in all leeches, both feeding and fasting individuals: endocytosis, lysosome activity, and peroxisome activity. Lysosome plays an important role in the cytoplasm, where it is used to digest unwanted material and defective organelles. Whereas extracellular waste material is transported into the cell by endocytosis, unwanted intracellular components are transported to the lysosome via autophagy. In essence, lysosome acts as the cells immune system, removing unwanted or detrimental molecules, and is also valuable for cell homeostasis (Roczniak-Ferguson et al., 2012). The foremost function of peroxisome is the disintegration of long-chain fatty acids through beta-oxidation. In animal cells, long-chain fatty acids are changed into medium-chain fatty acids and then transported to mitochondria where they are converted into carbon dioxide and water (Mannaerts, Debeer, Thomas, & De Schepper, 1979 and references therein).

To our knowledge, this is the first study to target differential expression of anticoagulant proteins in feeding and fasting leeches. Our results provide insight into the upregulation of anticoagulation factors during feeding and, in addition, provide some of the first data on the expression and secretion of anticoagulation factors by the nonblood feeding hirudinid *W. pigra*. Given our results, future investigations into the anticoagulation repertoires of leeches would benefit from including recently fed leeches, in order to recover as substantial a part of the expressed salivary transcriptome as possible.

ACKNOWLEDGMENTS

MSK and D-LG contributed equally to this work. This work was financially supported by the Excellent Doctor Innovation Project of Shaanxi Normal University (S2015YB03), Fundamental Research Funds for the Central Universities (GK201702010; GK201701006), and the Key Research and Development Program of Shaanxi Province (2018ZDXMSF-079). SK was funded by a NSERC Discovery Grant. We thank the Willi Hennig Society for making TNT freely available, and the editor and two anonymous reviewers for their constructive comments that greatly benefitted the paper.

CONFLICT OF INTEREST

The authors declare that there are no competing interests.

AUTHOR CONTRIBUTIONS

MSK and SQX designed the experiments. MSK and GDL performed the experiments. JYX, LBM, and SK analyzed the data. MSK wrote the manuscript draft. MSK and SK revised the manuscript.


DATA AVAILABILITY

All clean data for each of the six *Whitmania pigra* specimens (fasting: $n = 3$; feeding: $n = 3$), and annotated data (CDS, SNP, etc.) are deposited in the NCBI Short Read Archive (Bioproject ID PRJNA523919).

ETHICS APPROVAL

Not applicable.

ORCID

Muhammad Salabat Khan  <https://orcid.org/0000-0002-5740-2516>

REFERENCES

- Aibar, S., Fontanillo, C., Droste, C., Rivas, D. L., & J., (2015). Functional Gene Networks: R/Bioc package to generate and analyse gene networks derived from functional enrichment and clustering. *Bioinformatics*, 31(10), 1686–1688. <https://doi.org/10.1093/bioinformatics/btu864>
- Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, 11(10), R106. <https://doi.org/10.1186/gb-2010-11-10-r106>
- Antoneli, F., Passos, F. M., Lopes, L. R., & Briones, M. R. S. (2018). A Kolmogorov-Smirnov test for the molecular clock based on Bayesian ensembles of phylogenies. *PLoS ONE*, 13(1), e0190826. <https://doi.org/10.1371/journal.pone.0190826>
- Baici, A., & Seemüller, U. (1984). Kinetics of the inhibition of human leucocyte elastase by eglin from the leech *Hirudo medicinalis*. *Biochemical Journal*, 218(3), 829–833. <https://doi.org/10.1042/bj2180829>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Dang, Q. D., & Di Cera, E. (1994). A simple activity assay for thrombin and hirudin. *Journal of Protein Chemistry*, 13(4), 367–373. <https://doi.org/10.1007/BF01901692>
- Fink, E., Rehm, H., Gippner, C., Bode, W., Eulitz, M., & Macleaidt, W. (1986). The structure of bdellin B-3 from the leech *Hirudo medicinalis*. *Biological Chemistry*, 367, 1235–1242.
- Francischetti, I. M., Ribeiro, J. M., Champagne, D., & Andersen, J. (2000). Purification, cloning, expression, and mechanism of action of a novel platelet aggregation inhibitor from the salivary gland of the blood-sucking bug, *Rhodnius prolixus*. *Journal of Biological Chemistry*, 275(17), 12639–12650.
- Francischetti, I. M. (2010). Platelet aggregation inhibitors from hematophagous animals. *Toxicon*, 56(7), 1130–1144. <https://doi.org/10.1016/j.toxicon.2009.12.003>
- Garcia, D., Libby, E., & Crowther, M. A. (2010). The new oral anticoagulants. *Blood*, 115(1), 15–20. <https://doi.org/10.1182/blood-2009-09-241851>
- Goeman, J. J., & Bühlmann, P. (2007). Analyzing gene expression data in terms of gene sets: Methodological issues. *Bioinformatics*, 23(8), 980–987. <https://doi.org/10.1093/bioinformatics/btm051>
- Goloboff, P. A., Farris, J. S., & Nixon, K. C. (2008). TNT, a free program for phylogenetic analysis. *Cladistics*, 24(5), 774–786. <https://doi.org/10.1111/j.1096-0031.2008.00217.x>
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... Regev, A. (2011). Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology*, 29(7), 644. <https://doi.org/10.1038/nbt.1883>
- He, Z., & Bateman, A. (2003). Progranulin (granulin-epithelin precursor, PC-cell-derived growth factor, acrogranin) mediates tissue repair and tumorigenesis. *Journal of Molecular Medicine*, 81(10), 600–612. <https://doi.org/10.1007/s00109-003-0474-3>
- Hong, S. J., & Kang, K. W. (1999). Purification of Granulin-like Polypeptide from the Blood-Sucking Leech, *Hirudo nipponia*. *Protein Expression and Purification*, 16(2), 340–346. <https://doi.org/10.1006/prep.1999.1077>
- Jia-Xin, D., Xing-Chang, O., & Qiu-Hai, Z. (1994). Comparison among five different extracting methods of *Whitmania pigra*, Whitman. *Chinese Journal of Integrated Traditional and Western Medicine*, 3, 165–166.
- Jung, H. I., Kim, S. I., Ha, K.-S., Joe, C. O., & Kang, K. W. (1995). Isolation and characterization of guamerin, a new human leukocyte elastase inhibitor from *Hirudo nipponia*. *Journal of Biological Chemistry*, 270(23), 13879–13884. <https://doi.org/10.1074/jbc.270.23.13879>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kim, D. R., Hong, S. J., Ha, K.-S., Joe, C. O., & Kang, K. W. (1996). A cysteine-rich serine protease inhibitor (Guamerin II) from the non-blood sucking leech *Whitmania edentula*: Biochemical characterization and amino acid sequence analysis. *Journal of Enzyme Inhibition*, 10(2), 81–91. <https://doi.org/10.3109/14756369609020161>
- Knecht, R., Seemüller, U., Liersch, M., Fritz, H., Braun, D., & Chang, J. (1983). Sequence determination of eglin C using combined micro-techniques of amino acid analysis, peptide isolation, and automatic Edman degradation. *Analytical Biochemistry*, 130(1), 65–71. [https://doi.org/10.1016/0003-2697\(83\)90650-4](https://doi.org/10.1016/0003-2697(83)90650-4)
- Koh, C. Y., & Kini, R. M. (2008). Anticoagulants from hematophagous animals. *Expert Reviews in Hematology*, 1(2), 135–139. <https://doi.org/10.1586/17474086.1.2.135>
- Kordowicz, M., Gussow, D., Hofmann, U., Pacuszka, T., & Gardas, A. (2006). *Hyaluronidase from the Hirudinaria manillensis isolation, purification and recombinant method of production*. US Patent no. US7,049,124 B1.
- Kvist, S., Brugler, M. R., Goh, T. G., Giribet, G., & Siddall, M. E. (2014). Pyrosequencing the salivary transcriptome of *Haemadipsa interrupta* (Annelida: Clitellata: Haemadipsidae): Anticoagulant diversity and insight into the evolution of anticoagulation capabilities in leeches. *Invertebrate Biology*, 133(1), 74–98.
- Kvist, S., Min, G. S., & Siddall, M. E. (2013). Diversity and selective pressures of anticoagulants in three medicinal leeches (Hirudinida: Hirudinidae, Macrobdellidae). *Ecology and Evolution*, 3(4), 918–933. <https://doi.org/10.1002/ece3.480>
- Kvist, S., Sarkar, I. N., & Siddall, M. E. (2011). Genome-wide search for leech antiplatelet proteins in the non-blood-feeding leech *Helobdella robusta* (Rhynchobdellida: Glossiphoniidae) reveals evidence of secreted anticoagulants. *Invertebrate Biology*, 130(4), 344–350. <https://doi.org/10.1111/j.1744-7410.2011.00238.x>
- Linnaeus, C. (1758). *Systema Naturae Regnum Animale*, 10th ed. Stockholm: Salvius.

- Liu, Z., Zhao, F., Tong, X., Liu, K., Wang, B., Yang, L., ... Wang, D. (2018). Comparative transcriptomic analysis reveals the mechanism of leech environmental adaptation. *Gene*, 664(2018), 70–77. <https://doi.org/10.1016/j.gene.2018.04.063>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Mannaerts, G. P., Debeer, L. J., Thomas, J., & De Schepper, P. J. (1979). Mitochondrial and peroxisomal fatty acid oxidation in liver homogenates and isolated hepatocytes from control and clofibrate-treated rats. *Journal of Biological Chemistry*, 254(11), 4585–4595.
- Min, G. S., Sarkar, I. N., & Siddall, M. E. (2010). Salivary transcriptome of the North American medicinal leech, *Macrobdella decora*. *Journal of Parasitology*, 96(6), 1211–1221. <https://doi.org/10.1645/GE-2496.1>
- Moore, J. P. (1935). Leeches from Borneo and the Malay Peninsula. *Bulletin of the Raffles Museum*, 10, 67–78.
- Moore, J. P. (1939). Additions to our knowledge of African leeches (Hirudinea). *Proceedings of the Academy of Natural Sciences of Philadelphia*, 90, 297–360.
- Mory, R. N., Mindell, D., & Bloom, D. A. (2000). The leech and the physician: Biology, etymology, and medical practice with *Hirudinea medicinalis*. *World Journal of Surgery*, 24(7), 878–883. <https://doi.org/10.1007/s002680010141>
- Petersen, T. N., Brunak, S., von Heijne, G., & Nielsen, H. (2011). SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10), 785.
- Phillips, A. J., & Siddall, M. E. (2009). Poly-paraphyly of Hirudinidae: Many lineages of medicinal leeches. *BMC Evolutionary Biology*, 9(1), 246. <https://doi.org/10.1186/1471-2148-9-246>
- R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roczniak-Ferguson, A., Petit, C. S., Froehlich, F., Qian, S., Ky, J., Angarola, B., ... Ferguson, S. M. (2012). The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Science Signaling*, 5(228), ra42. <https://doi.org/10.1126/scisignal.2002790>
- Salzet, M. (2001). Anticoagulants and inhibitors of platelet aggregation derived from leeches. *FEBS Letters*, 492(3), 187–192. [https://doi.org/10.1016/S0014-5793\(01\)02212-8](https://doi.org/10.1016/S0014-5793(01)02212-8)
- Sanama, M. M., & Guinet, C. (2011). Laboratory assessment of new anticoagulants. *Clinical Chemistry and Laboratory Medicine*, 49(5), 761–772. <https://doi.org/10.1515/CCLM.2011.134>
- Sandercock, P. A. G., Counsell, C., & Kane, E. J. (2015). Anticoagulants for acute ischaemic stroke. *Cochrane Database of Systematic Reviews*, 3, 1–71. <https://doi.org/10.1002/14651858.CD000024.pub4>
- Say, T. (1824). D. Class Vermes. Order Cryptobranchia. Hirudo. In: Keating, W. H. (ed.) *Narrative of an expedition to the source of St. Peter's River, Lake Winnepeek, Lake of the Woods, performed in the year 1823, by order of the Hon. J. C. Calhoun, Secretary of War, under the command of Stephen H. Long, Major U.S.T.E.: Compiled from the notes of Major Long, Messrs. Say, Keating and Calhoun*. Volume 2. Philadelphia: H.C. Carey and I. Lea, 266–268.
- Sawyer, R. T. (1986). *Leech biology and behaviour* (Vol. II). Oxford, UK: Clarendon Press.
- Seemüller, U., Fritz, H., & Eulitz, M. (1981). Eglin: Elastase–Cathepsin G inhibitor from leeches. *Methods in Enzymology*, 80, 804–816. [https://doi.org/10.1016/s0076-6879\(81\)80062-6](https://doi.org/10.1016/s0076-6879(81)80062-6)
- Shankland, M., Bissen, S. T. & Weisblat, D.A. (1992). Description of the Californian leech *Helobdella robusta* sp. nov., and comparison with *Helobdella triserialis* on the basis of morphology, embryology, and experimental breeding. *Canadian Journal of Zoology*, 70(6), 1258–1263.
- Siddall, M. E., Brugler, M. R., & Kvist, S. (2016). Comparative transcriptomic analyses of three species of *Placobdella* (Rhynchobdellida: Glossiphoniidae) confirms a single origin of blood feeding in leeches. *Journal of Parasitology*, 102(1), 143–150. <https://doi.org/10.1645/15-802>
- Siddall, M. E., Min, G. S., Fontanella, F. M., Phillips, A. J., & Watson, S. C. (2011). Bacterial symbiont and salivary peptide evolution in the context of leech phylogeny. *Parasitology*, 138(13), 1815–1827. <https://doi.org/10.1017/S0031182011000539>
- Simakov, O., Marletaz, F., Cho, S.-J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., ... Rokhsar, D. S. (2013). Insights into bilaterian evolution from three spiralian genomes. *Nature*, 493(7433), 526. <https://doi.org/10.1038/nature11696>
- Tessler, M., de Carle, D., Voiklis, M. L., Gresham, O. A., Neumann, J. S., Cios, S., & Siddall, M. E. (2018). Worms that suck: Phylogenetic analysis of Hirudinea solidifies the position of Acanthobdellida and necessitates the dissolution of Rhynchobdellida. *Molecular Phylogenetics and Evolution*, 127, 129–134. <https://doi.org/10.1016/j.ympev.2018.05.001>
- Tessler, M., Marancik, D., Champagne, D., Dove, A., Camus, S., Siddall, M. E., & Kvist, S. (2018). Marine leech anticoagulant diversity and evolution. *Journal of Parasitology*, 104(3), 210–220. <https://doi.org/10.1645/17-64>
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25(9), 1189–1191. <https://doi.org/10.1093/bioinformatics/btp033>
- Weitz, J. I., & Bates, S. M. (2005). New anticoagulants. *Journal of Thrombosis and Haemostasis*, 3(8), 1843–1853. <https://doi.org/10.1111/j.1538-7836.2005.01374.x>
- Whitman, C. O. (1884). The external morphology of the leech. *Proceedings of the American Academy of Arts and Science*, 20, 76–87.
- Wilkinson, M., McInerney, J. O., Hirt, R. P., Foster, P. G., & Embley, T. M. (2007). Of clades and clans: Terms for phylogenetic relationships in unrooted trees. *Trends in Ecology & Evolution*, 22(3), 114–115. <https://doi.org/10.1016/j.tree.2007.01.002>
- Wu, B., Zheng, Q., Guo, X., Wu, Y., Wang, Y., & Cui, F. (2008). Preparation and ectopic osteogenesis in vivo of scaffold based on mineralized recombinant human-like collagen loaded with synthetic BMP-2-derived peptide. *Biomedical Materials*, 3(4), 044111. <https://doi.org/10.1088/1748-6041/3/4/044111>
- Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13(1), 134. <https://doi.org/10.1186/1471-2105-13-134>
- Zavalova, L., Lukyanov, S., Baskova, I., Snezhkov, E., Akopov, S., Berezhnoy, S., ... Sverdlov, E. (1996). Genes from the medicinal leech (*Hirudo medicinalis*) coding for unusual enzymes that specifically cleave endo- ϵ (γ -Glu)-Lys isopeptide bonds and help to dissolve blood clots. *Molecular and General Genetics MGG*, 253(1–2), 20–25. <https://doi.org/10.1007/s004380050291>
- Zhong, S., Yang, D., & Cui, Z. (2008). Studies on anticoagulant constituents in dried *Whitmania pigra*. *Zhongguo Zhong Yao Za Zhi = Zhongguo Zhongyao Zazhi = China Journal of Chinese Materia Medica*, 33(23), 2781–2784.

SUPPORTING INFORMATION

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

How to cite this article: Khan MS, Guan D-L, Kvist S, Ma L-B, Xie J-Y, Xu S-Q. Transcriptomics and differential gene expression in *Whitmania pigra* (Annelida: Clitellata: Hirudinida: Hirudinidae): Contrasting feeding and fasting modes. *Ecol Evol*. 2019;9:4706–4719. <https://doi.org/10.1002/ece3.5074>