

Supplementary Information

Combined transcriptomic and proteomic analysis of the posterior salivary gland from the southern blue-ringed octopus and the southern sand octopus

Brooke Whitelaw¹, J.M. Strugnelli¹, Pierre Faou², Rute R. da Fonseca⁴, Nathan Hall^{5,2}, Mark Norman⁵, Julian Finn⁵, Ira R. Cooke^{2,6}

- Table S1. Orthogroup sequences of the top 20 proteins with the highest abundance isolated from the posterior salivary gland (PSG) of *Octopus kaurna*. Families include AC (Actin), CP (CAP), HS (Histone), SP (Serine Protease), TR (Tropomyosin) and proteins with no annotated family were classified as UK (Unknown). Family was manually identified through observation of InterPro annotations. pp S3-S4
- Table S2. Orthogroup sequences of the top 20 proteins with the highest abundance isolated from the posterior salivary gland (PSG) of *Haplochlada maculosa*. Families include AC (Actin), CP (CAP), HS (Histone), SP (Serine Protease), TR (Tropomyosin) and proteins with no annotated family were classified as UK (Unknown). Family was manually identified through observation of InterPro annotations. pp S4-S5
- Table S3. Number of sequenced paired-end reads (in millions) for each tissue library of *Haplochlada maculosa* and *Octopus kaurna*. pp S6
- Fig S1. Maximum likelihood tree of the phylogenetic relationships between homologous CAP putative proteinaceous toxins from *Octopus kaurna* (orange) and *Haplochlada maculosa* (blue). Bootstrap values ≥ 50 shown above nodes. Tree was built with the VT +G +I model using amino acid sequences. pp S7

Excel SI files

“Supplementary table 4”

- Table S4. Sequences, which exhibit “toxin-like” features [short (<150 amino acids), cysteine rich (>3% cysteines) and secreted (SignalP)] from the posterior salivary gland of *Haplochlada maculosa* and *Octopus kaurna*.

“Supplementary tables 5&6”

- Table S5. Putative toxins isolated from the PSG of *Octopus kaurna* with their corresponding Trinotate IDs.
- Table S6. Putative toxins isolated from the PSG of *Hapalochlaena maculosa* with their corresponding Trinotate IDs.

“Supplementary tables 7&8”

- Table S7. All proteins isolated from *Octopus kaurna* posterior salivary gland. Includes protein IDs, sequence, cysteine count, length, cysteine %, indistinguishable proteins, #proteins, #peptides, #unique peptides, sequence coverage, molecular weight, score and iBAQ mean.
- Table S8. All proteins isolated from *Hapalochlaena maculosa* posterior salivary gland. Includes protein IDs, sequence, cysteine count, length, cysteine %, indistinguishable proteins, #proteins, #peptides, #unique peptides, sequence coverage, molecular weight, score and iBAQ mean.

“Supplementary table 9&10”

- Table S9. All protein sequences from *Octopus kaurna* with homology to a transcript isolated in a previous study. Putative toxin IDs are highlighted in yellow.
- Table S10. All protein sequences from *Hapalochlaena maculosa* with homology to a transcript isolated in a previous study. Putative toxin IDs are highlighted in yellow.

Table S1: Orthogroup sequences of the top 20 proteins with the highest abundance isolated from the posterior salivary gland (PSG) of *Octopus kaurna*. Families include AC (Actin), CP (CAP), HS (Histone), SP (Serine Protease), TR (Tropomyosin) and proteins with no annotated family were classified as UK (Unknown). Family was manually identified through observation of InterPro annotations.

| Orthogroup | Trinotate id | Length | Cysteine # | iBAQ | Family |
|------------|----------------------------|--------|------------|-------------|-----------------|
| AC20K4 | comp26684_c2_seq3_m.83259 | 375 | 4 | 44609.5 | Actin |
| AC20K1 | comp20288_c2_seq1_m.31146 | 376 | 6 | 140076091.4 | Actin |
| AC20K2 | comp26684_c2_seq4_m.83264 | 375 | 4 | 2777700.571 | Actin |
| AC20K3 | comp26684_c2_seq1_m.83246 | 375 | 4 | 648321.5714 | Actin |
| CP10K1 | comp27308_c0_seq1_m.101997 | 264 | 19 | 54290685.71 | CAP |
| HS20K1 | comp22121_c0_seq1_m.37985 | 121 | 0 | 51131857.14 | Histone |
| HS40K1 | comp24911_c0_seq2_m.55852 | 103 | 0 | 86875714.29 | Histone |
| HS50K1 | comp14212_c0_seq1_m.16916 | 125 | 0 | 174003700 | Histone |
| HY10K1 | comp27648_c0_seq1_m.102959 | 436 | 9 | 897990.8 | Hyaluronidase |
| SP10K1 | comp27296_c0_seq1_m.101938 | 207 | 8 | 68514174.29 | Serine Protease |
| SP20K1 | comp11207_c0_seq1_m.11947 | 255 | 11 | 31414057.14 | Serine Protease |
| SP30K1 | comp20357_c3_seq3_m.31402 | 257 | 11 | 78796428.57 | Serine Protease |
| SP40K1 | comp25408_c0_seq1_m.60376 | 254 | 12 | 73267714.29 | Serine Protease |
| SP50K1 | comp27295_c0_seq1_m.101936 | 254 | 11 | 55028785.71 | Serine Protease |
| SP60K1 | comp27297_c0_seq1_m.101939 | 274 | 11 | 55408457.14 | Serine Protease |
| SP80K1 | comp14350_c1_seq1_m.17313 | 253 | 10 | 74678285.71 | Serine Protease |
| TR10K1 | comp22305_c0_seq2_m.38864 | 284 | 1 | 52355328.57 | Tropomyosin |
| TR20K1 | comp27320_c0_seq1_m.102043 | 174 | 1 | 56695971.43 | Tropomyosin |

| | | | | | |
|----------------|----------------------------|-----|---|-------------|---------|
| UK100K1 | comp27294_c0_seq1_m.101933 | 266 | 5 | 889025714.3 | Unknown |
| UK110K1 | comp27298_c0_seq1_m.101943 | 287 | 1 | 12290585.71 | Unknown |
| UK120K1 | comp14353_c0_seq1_m.17320 | 300 | 1 | 116809428.6 | Unknown |
| UK30K1 | comp20288_c1_seq1_m.31142 | 79 | 1 | 110599666.7 | Unknown |
| UK40K1 | comp20288_c1_seq2_m.31143 | 99 | 1 | 113375428.6 | Unknown |
| UK50K1 | comp20288_c1_seq3_m.31145 | 71 | 1 | 163112928.6 | Unknown |
| UK60K1 | comp11211_c3_seq1_m.11958 | 294 | 2 | 18247185.71 | Unknown |
| UK70K1 | comp11247_c1_seq2_m.12024 | 86 | 1 | 10587100 | Unknown |
| UK80K1 | comp24658_c0_seq1_m.53671 | 307 | 2 | 97936800 | Unknown |
| UK90K1 | comp27293_c1_seq1_m.101930 | 322 | 2 | 311743571.4 | Unknown |

Table S2. Orthogroup sequences of the top 20 proteins with the highest abundance isolated from the posterior salivary gland (PSG) of *Haplochroma maculosa*. Families include AC (Actin), CP (CAP), HS (Histone), SP (Serine Protease), TR (Tropomyosin) and proteins with no annotated family were classified as UK (Unknown). Family was manually identified through observation of InterPro annotations.

| Orthogroup | Trinotate id | Length | Cysteine # | iBAQ | Family |
|-------------------|----------------------------|---------------|-------------------|-------------|---------------|
| AC1HM1 | comp70911_c2_seq1_m.207412 | 246 | 4 | 127436984 | Actin |
| AC2HM1 | comp69190_c3_seq1_m.168612 | 376 | 6 | 523065000 | Actin |
| CP1HM1 | comp74924_c0_seq1_m.388877 | 273 | 18 | 51893750 | CAP |
| EF1HM1 | comp71908_c0_seq1_m.232960 | 111 | 2 | 1423950000 | EF-hand |
| HS1HM1 | comp70358_c0_seq2_m.195124 | 95 | 1 | 147085000 | Histone |
| HS3HM1 | comp65470_c0_seq4_m.126516 | 125 | 0 | 264161750 | Histone |
| HS4HM1 | comp17737_c0_seq1_m.14626 | 103 | 0 | 132967000 | Histone |
| HS5HM1 | comp20968_c0_seq1_m.20388 | 125 | 0 | 140135200 | Histone |
| HY1HM1 | comp17608_c0_seq1_m.14275 | 432 | 9 | 107868750 | Hyaluronidase |

| | | | | | |
|----------------|-----------------------------|-----|----|-----------|-----------------|
| SP2HM1 | comp29613_c0_seq1_m.29982 | 263 | 11 | 149594250 | Serine Protease |
| SP3HM1 | comp48800_c0_seq1_m.60195 | 318 | 18 | 19253900 | Serine Protease |
| SP4HM1 | comp75768_c0_seq1_m.390801 | 262 | 11 | 1488650 | Serine Protease |
| SP5HM1 | comp29651_c0_seq1_m.30075 | 254 | 11 | 122141800 | Serine Protease |
| SP6HM1 | comp29671_c0_seq1_m.30130 | 302 | 18 | 89623175 | Serine Protease |
| SP8HM1 | comp74737_c0_seq1_m.388209 | 259 | 12 | 233434750 | Serine Protease |
| SP8HM2 | comp74766_c0_seq1_m.388320 | 254 | 11 | 18721525 | Serine Protease |
| TR1HM1 | comp66058_c0_seq11_m.132840 | 284 | 1 | 49818750 | Tropomyosin |
| TR2HM1 | comp74829_c0_seq2_m.388572 | 174 | 0 | 106943000 | Tropomyosin |
| UK10HM1 | comp17775_c0_seq1_m.14749 | 281 | 4 | 296415925 | Unknown |
| UK11HM1 | comp17707_c3_seq1_m.14548 | 288 | 1 | 99312750 | Unknown |
| UK12HM1 | comp64608_c0_seq1_m.121541 | 317 | 2 | 288257500 | Unknown |
| UK1HM1 | comp35214_c0_seq1_m.40617 | 74 | 5 | 91959500 | Unknown |
| UK2HM1 | comp64558_c0_seq2_m.121377 | 76 | 1 | 91020000 | Unknown |
| UK6HM1 | comp17707_c0_seq1_m.14529 | 285 | 2 | 153052250 | Unknown |
| UK7HM1 | comp74735_c0_seq1_m.388207 | 83 | 0 | 119114700 | Unknown |
| UK9HM1 | comp74972_c0_seq1_m.388993 | 309 | 2 | 94628750 | Unknown |

Table 3: Number of sequenced paired-end reads (in millions) for each tissue library.

| Species | Tissue | Number of PE Reads (Million) |
|--------------------|--------------------------------|---|
| <i>O. kaurna</i> | Salivary gland | 23 |
| <i>O. kaurna</i> | Posterior mantle | 22 |
| <i>H. maculosa</i> | Skin of Dorsal mantle | 24 |
| <i>H. maculosa</i> | Eyeballs | 24 |
| <i>H. maculosa</i> | Gills | 25 |
| <i>H. maculosa</i> | Systemic heart | 24 |
| <i>H. maculosa</i> | Posterior salivary gland | 24 |
| <i>H. maculosa</i> | Ventral mantle muscle and skin | 27 |
| <i>H. maculosa</i> | Brain | 24 |
| <i>H. maculosa</i> | Anterior salivary glands. | 24 |
| <i>H. maculosa</i> | Branchial hearts | 22 |
| <i>H. maculosa</i> | Renal appendages | 24 |
| <i>H. maculosa</i> | Male reproductive tract | 20 |
| <i>H. maculosa</i> | Digestive gland | 23 |

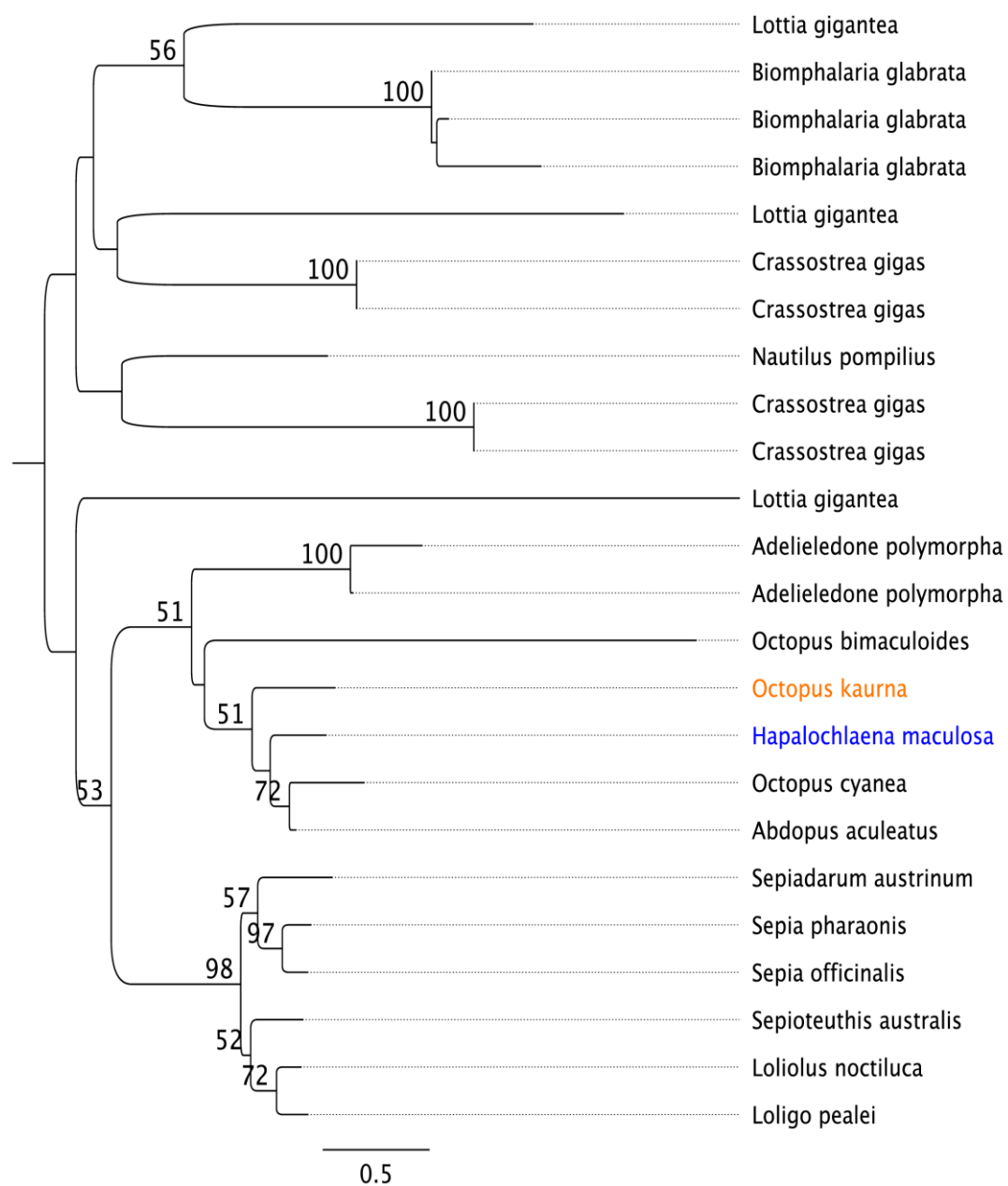


Fig S1. Maximum likelihood tree of the phylogenetic relationships between homologous CAP putative proteinaceous toxins from *Octopus kaurna* (orange) and *Hapalochlaena maculosa* (blue). Bootstrap values ≥ 50 shown above nodes. Tree was built with the VT +G +I model using amino acid sequences.