

Fig. S1. *A. rubens* **kisspeptin (KP)-type precursor (ArKPP)**. The nucleotide sequence (lowercase, 1038 bases) encoding the precursor protein (uppercase, 149 amino acid residues) is shown. The predicted signal peptide is represented in blue, two putative kisspeptin (KP)-type peptides (with cysteine (C) residues underlined) are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic/tribasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1095844 and has been deposited in GenBank under accession number KT601705.

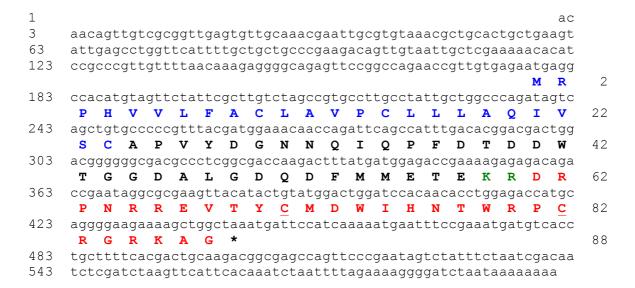


Fig. S2. A. rubens melanin-concentrating hormone (MCH)-type precursor (ArMCHP).

The nucleotide sequence (lowercase, 600 bases) encoding the precursor protein (uppercase, 88 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative melanin-concentrating hormone (MCH)-type peptide (with cysteine (C) residues underlined) is represented in red and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1059279 and has been deposited in GenBank under accession number KT601706.

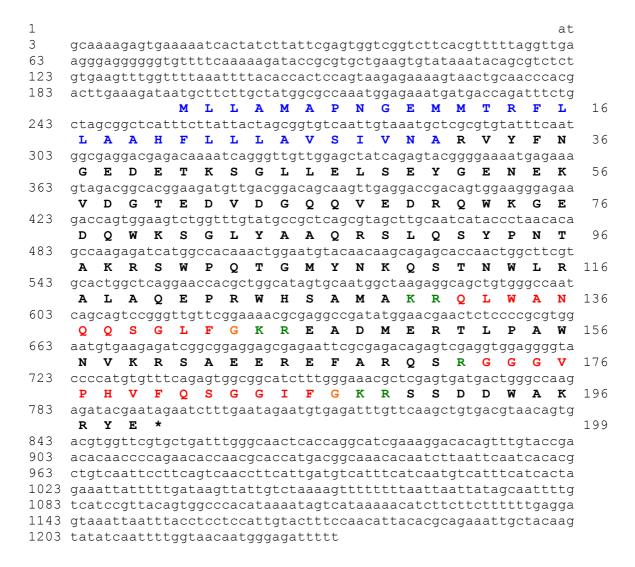


Fig. S3. *A. rubens* **tachykinin (TK)-type precursor (ArTKP).** The nucleotide sequence (lowercase, 1234 bases) encoding the precursor protein (uppercase, 199 amino acid residues) is shown. The predicted signal peptide is represented in blue, two putative tachykinin (TK)-type peptides are represented in red and putative monobasic/dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1103126 and has been deposited in GenBank under accession number KT601707.

```
1
    cttaacatcagcttcatggaattgggctctagtgatcgtttgatatgtttaccctaagtg
61
    aatagcctattaatatcgttttcttcttttaaatgttgtttgaagatttaaccgtgacaa
121
    attgatatgcaattgaaaaatcatcacccaatctcttggtctgtggctctctattttga
181
    gatgcaagtgagcgtgcaattgtctattgatgaggagcagagtttcgttccgctccctga
241
   tacagaccactgtagcaagccaaccccagctctcatcactgaaaccaaaccactgagat
301
   cacccaatccattttaaaacgtgcacttcgtgccagcgtggtatttaaaggttcactcaa
361
   421
    ctgcgcgcccatgcccacatccagtgatcacatcttgtcaccctcttacatcaactact
481
    541
    cactcactcaaccacgccgtttgaccattgaaatgacgcactgacgacatcgtgggctac
601
    aacagtttcttaagcagcqctccqqaqqttcacacatqqqcqqttacttcttccaa
661
    gcaggtccactgtacatcgactcgactggcgcgcgctaaaccatagactccattgctag
721
    gaagaaacccgtggatgcagagttatttgtgagttatacttgcgctccaactcaaaaact
781
   cgacaaaataataagtgaagacaagaaaaaaggggggttacgtggagcgcccattcctcc
841 tcctggagtcttcttaaagggggccttggtttttagtttcgtgaaactctacttctttta
901
    aacaccatgaggttctgttgttgttgttgttgttctaccgttgatctgtgtgctagctggg
          M R F C C V V V V L P L I C V L A
                                                         18
961
    {\tt tgcctagtggcccatgcagcaccgaggaggaggcggaggaggcaaccagcgacccgcggtgg}
    C L V A H A A P R R G G G N S D P R W
                                                          38
1021 aaaaggaatttctcgccaccgggaatgcagagctcgggcggcagctacaataaaggagat
     K R N F S P P G M Q S S G G S Y N K
                                                          58
1081 cttgttgaaaggatattgaacagattacaagaacgtctcctcgggaaagttgacctctct
                                                          78
     LVERILNRLQERLLGKVDLS
1141 caaaccaacacatggcatggaaatcagagtccaaaagaattagaccttcaacgatatagt
     Q T N T W H G N Q S P K E L D L Q R Y S
                                                          98
1201 gatcaagaggatgagtttatagacgatgatgatgaagtgcctaaccgcccggctattaag
     DQEDEFIDDDDEVPNRPAI
                                                         118
1261 aggaaatgtatcggtcggtttcaacctttctctatgccctgttagagttcacccagccta
                                                         132
     R K C I G R F Q P F S M P C
1321 tcagggctgtttggtctcataagaacctcccgtacaagattatttttaagggctgtgttt
1381 tgaaattgagcgtctttggctatggctgatgacgtca
```

Fig. S4. *A. rubens* **somatostatin (SS)-type precursor (ArSSP)**. The nucleotide sequence (lowercase, 1417 bases) encoding the precursor protein (uppercase, 132 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative somatostatin (SS)-type peptide (with cysteine (C) residues underlined) is represented in red and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1107850 and has been deposited in GenBank under accession number KT601708.

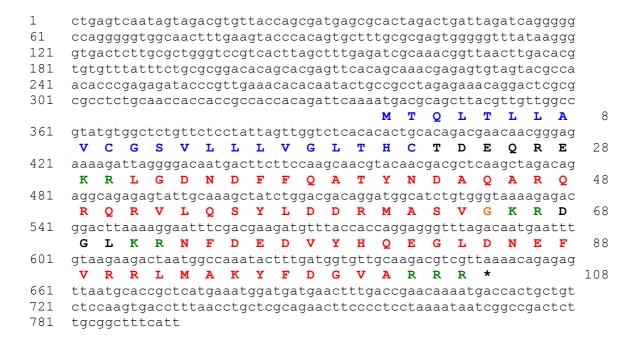


Fig. S5. *A. rubens* **pigment-dispersing factor (PDF)-type precursor (ArPDFP).** The nucleotide sequence (lowercase, 793 bases) encoding the precursor protein (uppercase, 108 amino acid residues) is shown. The predicted signal peptide is represented in blue, two putative pigment-dispersing factor (PDF)-type peptides (with cysteine (C) residues underlined) are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic/tribasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1081093 and has been deposited in GenBank under accession number KT601709.

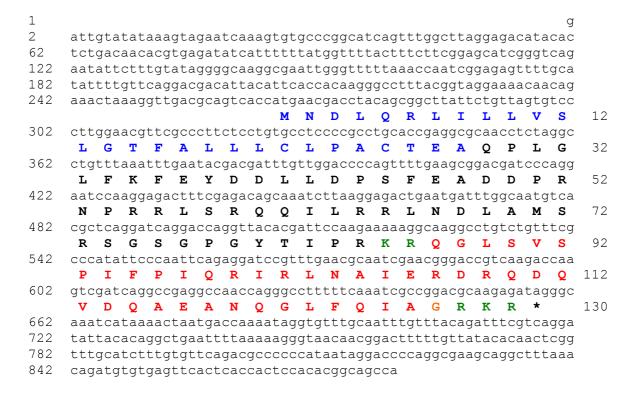


Fig. S6. A. rubens corticotropin-releasing hormone (CRH)-type precursor (ArCRHP).

The nucleotide sequence (lowercase, 880 bases) encoding the precursor protein (uppercase, 130 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative corticotropin-releasing hormone (CRH)-type peptide is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and putative dibasic/tribasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1087380 and has been deposited in GenBank under accession number KT601710.

```
1
2
    \verb|catgacgtg| tgtgtgtgtgcatgcccgacacacagtactacgatcagcagacgagatt|
62
    qaqqqacqatcqctqqttttttttqcqaqaqtacqcqattqataacqacttcatcatcaat
122
    tggctcagatacgtttattgaataacctgttgtgaagagaaaagttgtccaaggatattc
182
    \verb|cacagaacttccgtccggtggacgatgggcatgaagtcaatggtggcgctgtggactggg|
                          M G M K S M V A L W T G
                                                            12
242
    gtactggtcgcactgtgggttcaaagtcaagcatgtttagttcaagactgtccggaagga
     V L V A L W V Q S Q A C L V Q D C P E G
302
    qqaaaqaqqtccaqttataacacaatcaqacaqtqcctatcctqcqqccctqqtqqtttq
     G K R S S Y N T I R Q C L S C G P G G L
                                                            52
362
    qqacaatqtqtaqqttcaqctatatqctqcqqcaatactttcqqctqttttctcqqqaca
     G Q C V G S A I C C G N T F G C F L G T
                                                            72
422
    {\tt aaagaaaccttcgtgtgcagagaagaaagtcagctctccacaccttgtgaggttgttgga}
     KETFVCREESQLSTPCEVVG
                                                            92
482
    gagacatgtgaatctattactgacgggaaatgtgtttcaaacggcttctgttgcaatgag
     E T C E S I T D G K C V S N G F C C N E
                                                           112
   agaagctgctctttagacgtagcgtgcagagaaaccgatacagaacagagagacctcaaa
     R S C S L D V A C R E T D T E Q R D L K
                                                           132
602 aacagactcaaagagaggcttctggacgccctcttgcgtcaaccatgaacccagttacca
     NRLKERLLDALLRQP*
                                                           147
662 ctccctcaactttgttgataattttcacccacaatgcaatgtgatatggtcacttgtgac
722 accttcacggacacaatgttgcggtcctttatgttgtgacaatcttggtgaatttctgaa
782 tatgctcacgtatagacacaggacaccatacttttataggggcaccagaccattattata
842 gccaggctcggtttggtcacaaaacaaacccctatttttgtggatgataataatctttgg
902 gacacgtaactaactgtatcacgctgcattgtaggttaagagaagctttgaagatatcct
1022\ attctttgaaattaatcgtaaatttaaaaactatttaaatacacagcggtaccgacattg
1082 ggactaaagagcaatttagtttgtgtacatctttctgatgcaagtaaacggaacttcaaa
1142 tggggaaagatcagataaacaaacaggggacggatgcagttacttttaatgagacaaata
1262 ttgttaatgtcttaatgctttctagcacttacaaataatattacacaatttttgggtgcg
1322 ctaaactttgagcttttgacaagattctttcttgcattgcgtcatattttattatctttc
1382 qaaatgtcaggatcacagttgcaattatagccatatcgatgttcatcttgagcgaaataa
1442\ a cag caa aa aa caa tatttattttctgaa tatatagttga cag tcgttcacttttaa tata
1502\ tatatatttagattcaaattctcagcaattttagccaatattatagtacaataatgtgac
1562 tgctacgacagactttaaaaataaaaaaatgtaaataaaaatgttattttagatttagt
1622 taaattaacttttgaaaatcagctgcaaaactgcacgcaaatggccccaaaacaacaat
1682 aaatcaaatatatgttaacaagagcacactcttgtggaattttagaggggacaaaccggc
1742 \ tgtatgagacaggtgaactttctataagaagtccacactttaaaatttataatttaatct
1802 ttagagtacccagtaaaacaacattttctgtcgctttttattgtttaacgtcccatag
1862 aaataattgcacatgtagatttaacttagcagactttgactatacgatacttttcgactt
1922 tactgcgcaactcgtacatttcctgcgtataatagatacaattttgacctattttaaaag
1982\ {\tt ttcaatgtgaatttcttgttgtttcccaaaaatatgcagtcattctgtgtatcataatat}
2042 \ tgatgtgaagtggtttcattttagttgaattaaaaccccactgaacgaaaattatttcgt
2102 tacqaattqtacattqttaaaaqctctttctccatqqttaacccqtqtccataqtaqqcq
```

Fig. S7. A. rubens vasopressin (VP)/oxytocin (OT)-type (asterotocin) precursor. The nucleotide sequence (lowercase, 2209 bases) encoding the precursor protein (uppercase, 147 amino acid residues) is shown. The predicted signal peptide is represented in blue, putative asterotocin peptide is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site is represented in green. The C-terminal region of the precursor comprises a neurophysin domain, with fourteen cysteine (C) residues that are a characteristic and conserved feature of neurophysins, which is underlined. The asterisk shows the position of the stop codon. This sequence was determined from contig 1119045 and has been deposited in GenBank under accession number KT601711.

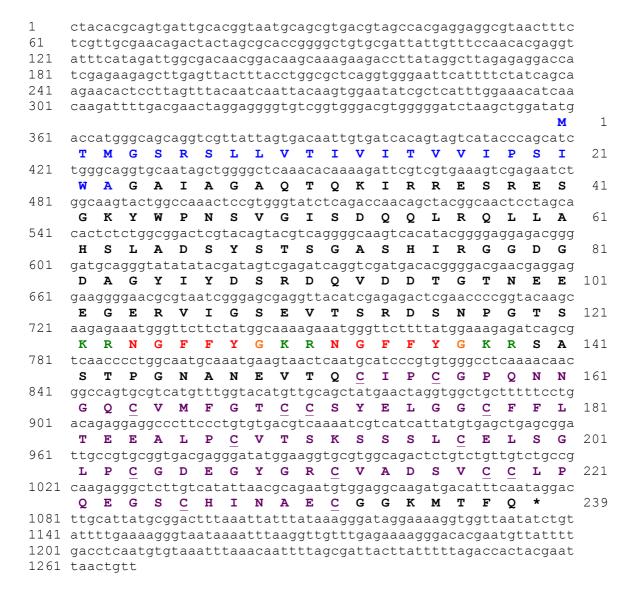


Fig. S8. A. rubens NGFFYamide precursor. The nucleotide sequence (lowercase, 1268 bases) encoding the precursor protein (uppercase, 239 amino acid residues) is shown. The predicted signal peptide is represented in blue, two tandem copies of the putative NGFFYamide peptide are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The C-terminal region of the precursor comprises a neurophysin domain, with fourteen cysteine (C) residues that are a characteristic and conserved feature of neurophysins, which is underlined. The asterisk shows the position of the stop codon. This sequence was determined from contig 1104160 and has been deposited in GenBank under accession number KC977457.

```
1
   aaaaactgatttgcttaagctggaagtaacaaaatattgatcgattgtcttgaaagatga
61
   acctgtcccagttcattcttcggagacattactctgtattaagatcaatccacgtgtgat
121
   gaaggttacggatacgctgagacgtcaattgatgtcactggagtcgacaaatgcgtgcac
181
   gccgattagttgctagactttttagccaatggcgcactgttgctcagctcacgtgataat
241
   accgtcaactgtccccttttcactcatttcggttccttgttttaaagcatcagacgtgac
301
   ctcttgagattgatcaaaaagagctttatatgttcagtgtcaactcattatttcagcaga
361 taaagaccgtcagagtgatttttcacacagtatttacgtttatagtcatcgttaatagtt
421
   cacqtcacqaccactqqqctacattaacacccatctataqatccqcqtaacactcccaaq
481
   541
   cacqcqacacaqcqtqcaaqtctcaqttcqttctcaqttaqcaaccaaqaaataqtqtaa
601
   qcqcttcataaqqaaaactqtaaqaqaaqaacacaqqaqaqtcactqqaqttaaqaaqcc
661
   10
                             MADMRMLTLT
   agcgtattagtctctctactcttcatggcagaaattcaaagatgccaagggcagatacat\\
721
    S V L V S L L F M A E I Q R C Q G Q I
                                                      30
781
   taca aga at cct ggat ggggacct ggt ggt aa aag gagt tcaca cat gact ggt ag caat \\
    Y K N P G W G P G G K R S S H M T G S
                                                      50
841
   {\tt V} L R K R H W R V E S D Q M G T D S M Q
                                                      70
901
   aaaqaacqaaacttqatcatqcttcaaqaaattqcaaaatctttqqcaaaqcaactqqta
    KERNLIMLQEIAKSLAKQL
                                                      90
961
   gtaccaacgagtgaggacgacacagtcctggaccaattaacggtcgaccaatggcggcag
    V P T S E D D T V L D Q L T V D Q W R Q
1021 gaagcagacgagataaatgacaacggttggaattaagcgggaaaagctctgaaatttgac
    EADEINDNGWN*
                                                     121
1081\ aacaattatttagaatcaggaagaactgaacaacttgatacaggttcatatgtgtttgta
1141 ttgctttcttttt
```

Fig. S9. *A. rubens* gonadotropin-releasing hormone (GnRH)-type peptide 1 precursor (ArGnRH1P). The nucleotide sequence (lowercase, 1154 bases) encoding the precursor protein (uppercase, 121 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative gonadotropin-releasing hormone (GnRH)-type peptide is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1100532 and has been deposited in GenBank under accession number KT601712.

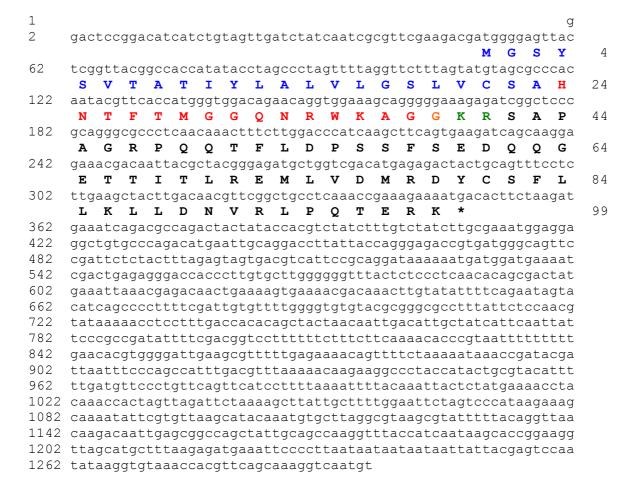


Fig. S10. *A. rubens* gonadotropin-releasing hormone (GnRH)-type peptide 2 precursor (ArGnRH2P). The nucleotide sequence (lowercase, 1297 bases) encoding the precursor protein (uppercase, 99 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative gonadotropin-releasing hormone (GnRH)-type peptide is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1104992 and has been deposited in GenBank under accession number KT601713.

```
1
    aaaaaacatacacatagaaaagaacgggcattctcaacttgctgtagatcacagccgt
61
    aataagtgactggctataacaaccgatacaatttatagggactagccaatcaaagctcct
121
   cactttgcggtggtacgctgtgtgtggtgcaagtgataagacagcgtgctgcgatacggt
181 gcttgacagtaaagcgagaacacgctagcgaacgggacacgatactttttatttctctc
241 tctcttactctctatttctctctgatgaacaagcagtcaaccatcatcgagatcatgctt
301 atgcgtgtctgattacacattgtgcttgttcttaatttttgtcctgcagagttgttactt
361 caaqcaqtcaaatctcttaagttgtgccggccgacaccaagcaaccgataaagccaagtc
   actccacggcaagggacgagcaccctgttattttqacccqqqccqqtaqqccqttaqcqt
481
    ccagtcactataatcatgtcggttacgagaaacagcggttttctacttgttactctactg
                  M S V T R N S G F L L V T L L
                                                           15
541
    ttcacttqqqtqqtttqtcqcqctqaqttqqcaqatttcqtqqaqaatqctqaqqtqqcq
     F T W V V C R A E L A D F V E N A E V A
                                                           35
601
    aaagaagtatctaatgaaattgaaggagtagaagcagagcaatggcagagggacgaagac\\
     KEVSNEIEGVEAEQWQRDED
                                                           55
661
    aaaagacaatacccaggaggggctcccattggtctagatggtaaaagacaatggtacacc
     KRQYPGGAPIGLDGKRQWYT
                                                           75
721
    \tt ggcaagcggcaatggtatacgggtaaacgggacgctgaagactcgcccgcattattagca
     G K R Q W Y T G K R D A E D S P A L L A
                                                           95
781
    qaaaacqataaacqccaatqqtacactqqcaaacqqaqcqqcaacqaaqaacaqcaqccc
     ENDKRQWYTGKRSGNEEQQP
                                                          115
841
    qacqaqqcqaacaaqaqacaqtqqtataccqqcaaacqqcaatqqtacaccqqcaaqcqa
     D E A N K R Q W Y T G K R Q W Y T G K R
                                                          135
    ggggacgaagacagggtgctggacgacgatgccgtgaactctctcaaaacgccagtggtac
901
     G D E D R V L D D D A V N S L K R Q W Y
    accqqaaaacqccaqtqqtacaccqqtaaqcqaaqcqqqqtcqaqcaaqccqacqqc
     T G K R Q W Y T G K R S G V E Q A D D G
1021 gacttggagcaacaatacaacaacggcaatggtataccggtaaaagggccgacgacctt
     D L E Q Q Y N K R Q W Y T G K R A D D L
                                                          195
1081 gctgatgctgctgacctagaaaagcggcagtggtacaccgggaaaaggcagtggtacacc
     ADAADLEKR<mark>QWYTG</mark>KR<mark>QWYT</mark>
                                                          215
1141 ggtaaaagacaatggtataccggtagacgttaatcaaccaccaccactccttaatgaact
     GKRQWYTGRR*
                                                          225
1201 gtcaaaatatttcctggatattgcggggctattaaagactttttgttttcaaacacaaac
1261 ggttgacgaatccctccccgaatccagttgcattattgaatggaagaaacggtgtaagat
1321 aacattaaaattatagttttaagttttgcataaattattatggattttttgtttttc
1381 tcgaacaaagaccgaagtcacactatagtcaaaatattattttcactaccttttctgaag
1441 gttcaaagcaatgaaaacattttaggcctagcacatttgttttgaaacgataaggtaact
1501 catgatcttcgtaaggtgacactgacaaaacgaaatccaccccatattttttaaatttaa
1561 taattaacgattaatacggaagccccaaactgccattcaagatatacgtatatcgtatag
1621 taatacagaagtttggtatagtttatctatatgttcttcaatttcttacatgacgaacgt
1681 tttgacgtcactgttacgtcactacggcgagggctgtcccagtgacgatgataaggtgtc
1741 acgaatattgtcggtgaccgtttacaagttcaatgtctgtgtttatggcgagattgtgtc
1861 ttcaaatcttgatttcttgtatctaagaacttgttattattcagggctacagtggccaat
1921 cccgatgaaaacaatacaacgaccacatt
```

Fig. S11. A. rubens thyrotropin-releasing hormone (TRH)-type precursor (ArTRHP).

The nucleotide sequence (lowercase, 1949 bases) encoding the precursor protein (uppercase, 225 amino acid residues) is shown. The predicted signal peptide is represented in blue, twelve putative thyrotropin-releasing hormone (TRH)-type peptides are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contigs 1062625 [1-42] and 1105818 [43-225] and has been deposited in GenBank under accession number KT601714.

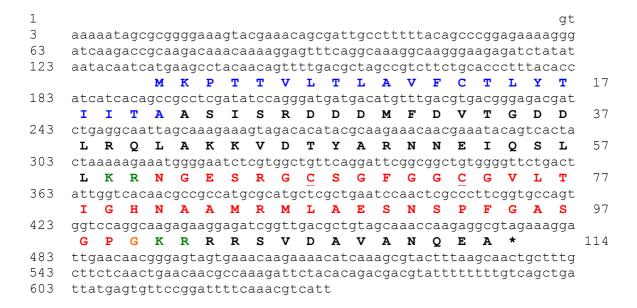


Fig. S12. *A. rubens* calcitonin (CT)-type precursor (ArCTP). The nucleotide sequence (lowercase, 633 bases) encoding the precursor protein (uppercase, 114 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative calcitonin (CT)-type peptide (with cysteine (C) residues underlined) is represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1064235 and has been deposited in GenBank under accession number KT601715.

```
1
3
       ggcttcaaacttcaacatctcttgttatgctcgtcagctatcgtggttctaaaatcgcat
       actcttacaacqttccqqctqcttacatctctccccaattccqtctqcatctaaccaaca
123
       aagggacctcctctgttaatttgttcacttttttgatcgaattgatctgattttcagttt
183
      tgattcccactaggaacacgaacttgacgtgtttaaaacgaactgtcggtattttggttc
243
       \verb|cttcttgctgagctgcaaggggccacatcgctactgtttctctcgcattgcttgtatttt|\\
303
       gactgattgaacttatcaacacctatccgtatataggaagttctgctcgagatcataatg
363
       agtagttggcttacagtcgccatagcaactgtgacatgccttttgctttcgccaatcacg
        S S W L T V A I A T V T C L L L S P I T
                                                                                                   21
423
       tgcctgcctcttcatgacgtagccgacggtaaggaaaggcgggaactcctgcacagcacg
        C L P L H D V A D G K E R R E L L H S T
                                                                                                   41
483
       tggttagacccgtccggttcaacaggtcaaggcacgaggaattggccgagacgagcaag
        W L D P S G S T G Q G T E E L A E T S K
                                                                                                   61
543
       cgactacttggggataacaacagggactcgggcattattgacctccttgtagcactgcga
        R L L G D N N R D S G I I D L L V A L R
      gacacaaacacaaacccgagagatctttatcttcacggcaacacagagacagctcgtaaa
        D T N T N P R D L Y L H G N T E T A R K
                                                                                                  101
663
       cqaaqacaatccaaqqtqqatqactacqqccatqqtctattctqqqqcaaqaqqqtcc
        R R Q S K V D D Y G H G L F W G K R G S
723
       aactggtcagaccacggggtacgagcaatgacagataaggacaccaagagggggggtgat
        N W S D H G V R A M T D K D T K R G G D
       gatcaatatggctttggcttatttttttggcaagcgaaatgaagaagactacgaagacttt
        D Q Y G F G L F F G K R N E E D Y E D F
       acgttgtagattgtttagcaaataaggatacttaaaagctctaaagaattttggcaaagta
                                                                                                  163
      atgttttgtaaataggcatagtctttatgataagttaacgcaatatcaaactaatataac
      tcgacacttcgttgggggttaagtctcaatagagtgtccaactatcggccaaatacaatt
1023 gtgcaaaaacgtctcgaaatcatttacaaataattcacacaaaccatgggtacttggta
1083 attgtgtaatgcgtttgaacgtgcgatgtatcgagttaattagactgataaatacggata
1143\ a at catg taa agc cag ta cag ct cac actg caa at taca a cattag ctt tag ttt a agt tag can be a catger of the category of
1263 ataggaatcattgttcaaatgacgtttaccgcaagtggtattacacctctctttctgaac
1323 \ {\tt tacatgacaggatcaaaatgacagtttagcagaagtattttgtgttatgcggttttataa}
1383 atataaaccccatgtttaagaaaccttgtattttagaacgcaagcagttgttgacgcctt
1443 ttggaagtcttgcaatctgtaaatgtatgtgggtaccattgaagctatatgccatctcaa
1503 tcactgcttctctatcataacctttttggaaaaacacacattttataggcaaaatagtaa
1563 \ {\tt agttatgagttaaccacaacgttgcgttgcaaatcggagcgctttcaatccgtcgatgtt}
1623 ttattggtttgaactgactgttgtttctcaataataacaaatctgtgcaactactgtgga
1683 tttcgttgcgtcattctgaatattgtttcttttgtttaaactccaatgttgacaaactta
1743 acqttaatccattqtaaattattttatqcataaccqtcttttqaqtcaacttcqtaaqcc
1803 cgatgttttatttttgttaagtgtatataccttgaacatgaactagcagaaaagagttac
1863 ttagcacgatttgcagatttgaatcatgtaggaagtataacacatttatatagttctcaa
1923 caatgctttactaatctatctaaggaaagttgccgttctgaaaagaagcaggcttgagat
1983 ggaggtatttcgtttcctactcacaacctctccttcttgttgtacagattaatataacat
2043 agttgtcaacctacattattaaaccgatcgtgtaatttgattggcggaacacgcgtcacg
2163 gagttctatattattaaatattgtacgcttttgacgtttcgtaacttcgctttgagattt
2223 acattgagaatcgtttgcattgattgcgctccaggacaggttaggctatggacgccgaac
2283 gtgctttaaatagcatacttgtttatacacagctttatcgatcatttcatacagtaatgt
2403 aaggatttatccatgggcactcacatatatactgcattttattcataataaagcagatca
2463 caatccctagttqaaatctqaactqttttatttaacttqaqctcqtqattctaatatqac
2523 actctttacgcatcacacaatatttgtatcgtagcccctgacgcaaatgaatacaagtgt
2583 cttatgtattatgcaaattatagttgagttgttatccaaagattagttaacattattttg
2643 tcctatatcattagcgctaaaatgtaacatttactttgaagtgataaaacgtgtctctat
2703 tcctttctgactgagcacttgggtagatacagtgcgcttaaatttcaacgtgctgactat
2763 gacattaaacaataacgacttgcttaaattaagttttcaggtgataaatgcatcaccaaa
```

Fig. S13. *A. rubens* **cholecystokinin (CCK)-type precursor (ArCCKP)**. The nucleotide sequence (lowercase, 3434 bases) encoding the precursor protein (uppercase, 163 amino acid residues) is shown. The predicted signal peptide is represented in blue, two putative cholecystokinin (CCK)-type peptides are represented in red and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1124413 and has been deposited in GenBank under accession number KT601716.

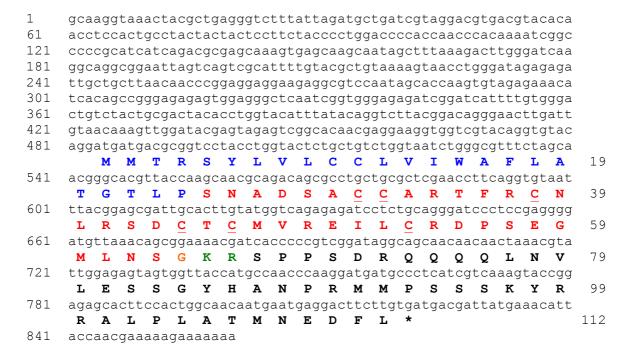


Fig. S14. A. rubens orexin (OX)-type precursor 1 (ArOXP1). The nucleotide sequence (lowercase, 860 bases) encoding the precursor protein (uppercase, 112 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative orexin (OX)-type peptide (with cysteine (C) residues underlined) is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1086098 and has been deposited in GenBank under accession number KT601717.

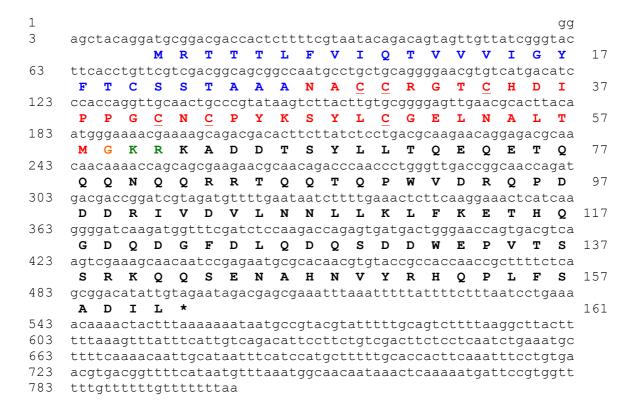


Fig. S15. A. rubens orexin (OX)-type precursor 2 (ArOXP2). The nucleotide sequence (lowercase, 802 bases) encoding the precursor protein (uppercase, 161 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative orexin (OX)-type peptide (with cysteine (C) residues underlined) is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1081859 and has been deposited in GenBank under accession number KT601718.

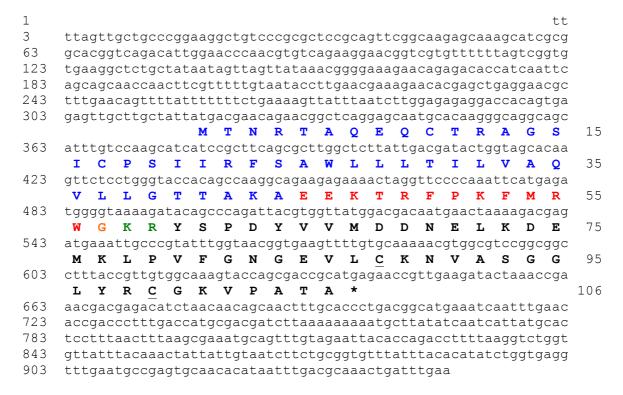


Fig. S16. *A. rubens* **luqin (LQ)-type precursor (ArLQP)**. The nucleotide sequence (lowercase, 948 bases) encoding the precursor protein (uppercase, 149 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative luqin (LQ)-type peptide is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and putative monobasic and dibasic cleavage sites are represented in green. Two cysteine (C) residues characteristic of the C-terminal region of LQ-type precursors are underlined. The asterisk shows the position of the stop codon. This sequence was determined from contig 1091466 and has been deposited in GenBank under accession number KT601719.

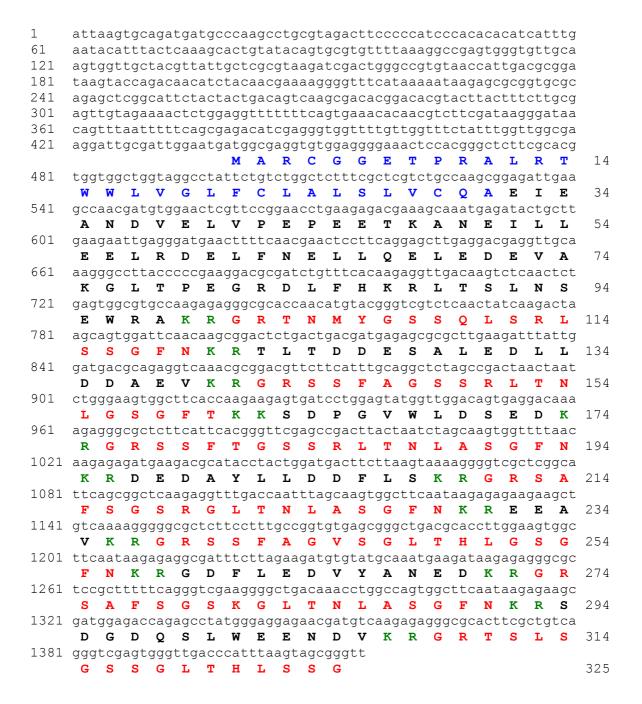


Fig. S17. *A. rubens* **pedal peptide (PP)/orcokinin-type precursor (ArPPLNP).** The nucleotide sequence (lowercase, 1415 bases) encoding a partial precursor protein sequence (uppercase, 325 amino acid residues) is shown. The predicted signal peptide is represented in blue, seven putative pedal-peptide (PP)-type peptides are represented in red and putative dibasic cleavage sites are represented in green. This sequence was determined from contig 1107802 and has been deposited in GenBank under accession number KT601720.

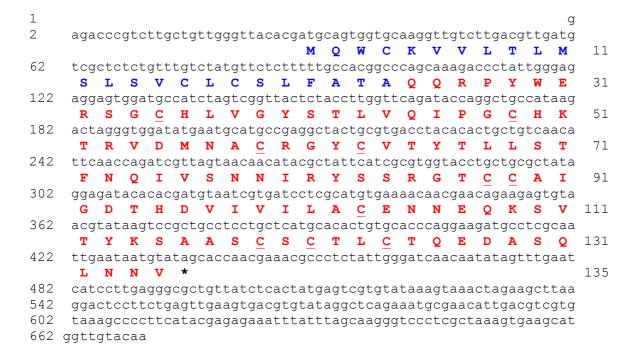


Fig. S18. A. rubens glycoprotein hormone α -2 (GPA2)-type precursor 1 (ArGPA2-1).

The nucleotide sequence (lowercase, 671 bases) encoding the precursor protein (uppercase, 135 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative glycoprotein hormone α -2-type (GPA2)-type subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1069119 and has been deposited in GenBank under accession number KT601721.

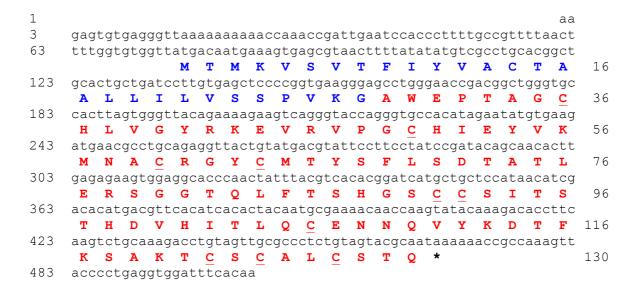


Fig. S19. A. rubens glycoprotein hormone α -2 (GPA2)-type precursor 2 (ArGPA2-2).

The nucleotide sequence (lowercase, 504 bases) encoding the precursor protein (uppercase, 130 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative glycoprotein hormone α -2-type (GPA2)-type subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1041026 and has been deposited in GenBank under accession number KT601722.

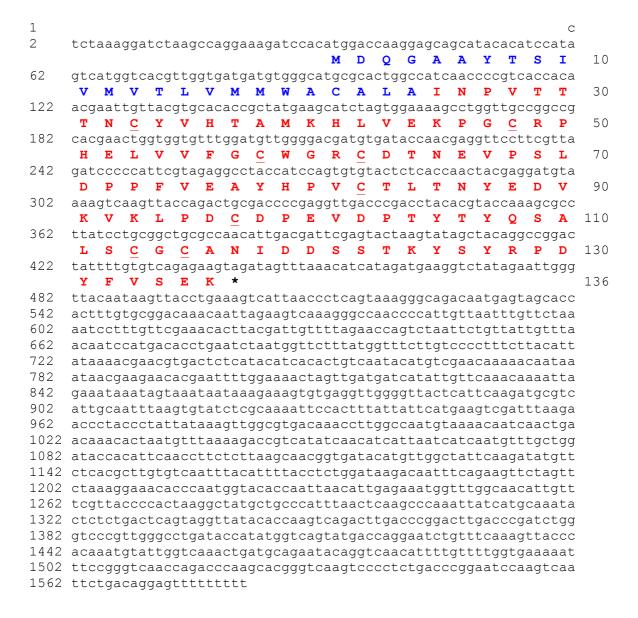


Fig. S20. A. rubens glycoprotein hormone β-5-type (GPB5)-type precursor 1 (ArGPB5-

1). The nucleotide sequence (lowercase, 1582 bases) encoding the precursor protein (uppercase, 136 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative glycoprotein hormone β -5-type (GPB5)-type subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1111214 and has been deposited in GenBank under accession number KT601723.

```
1
   61
   ccgccatcttaagagtttagtgcgaataacgtttatatatgtagtcaatttggatcgtta
121
   aaagacacgtccgtcctgctccaacgagaacttctatttccatcgggtgctaaagcggaa
181 cgacatacacgtcgcacaatgttcttcaagtgtcaagcgaggcacagatggactgccacg
                  MFFKCQARHRWTAT
                                                   14
241
   \verb|ttcctgcctcttttgcattggctctcctcattggcaccatcgttgagggcgctcgttac|
    F L P L F A L A L L I G T I V E G A R Y
                                                   34
   caaaaacaqaqcatqccctttttqqacqtaaqttqcaqaqttcqqqaqtacacaaaqtat
    Q K Q S M P F L D V S C R V R E Y T K Y
                                                   54
361
   gaggctaagctgcccggatgtatggacgaagtagtgccggcaagagggtgctatggtcgg
                                                   74
    E A K L P G C M D E V V P A R G C Y G R
421
   tgccagtcgttcgaggtaccagttttgctaccaccgcacaaagcgtcaaagtcacaagatg
    94
481
   tgcttggtcgaggaaatagagcttcgtagcgtggagttgtcggattgcctaccgggggtc\\
    C L V E E I E L R S V E L S D C L P G V
                                                  114
541
   {\tt aacaggacctttgtctatcagagtgccgtacgatgcagatgcaagaaatgcatcgagtct}
    NRTFVYQSAVRCRCKKCIES
                                                  134
601
   NTFCARN*
                                                  141
661
   qtqtatatatttqttaaaacaaatttcaaactqqttttcqctttqqttttaqaqqqqttta
721
   caaaatcaaacgtatttaacaaacattttataggtatagtatatgaatcaattccaaata
781
   atgctagatgcatctacccgcgctgttcttgattaca
```

Fig. S21. A. rubens glycoprotein hormone β-5-type (GPB5)-type precursor 2 (ArGPB5-

2). The nucleotide sequence (lowercase, 817 bases) encoding the precursor protein (uppercase, 141 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative glycoprotein hormone β -5-type (GPB5)-type subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1083001 and has been deposited in GenBank under accession number KT601724.

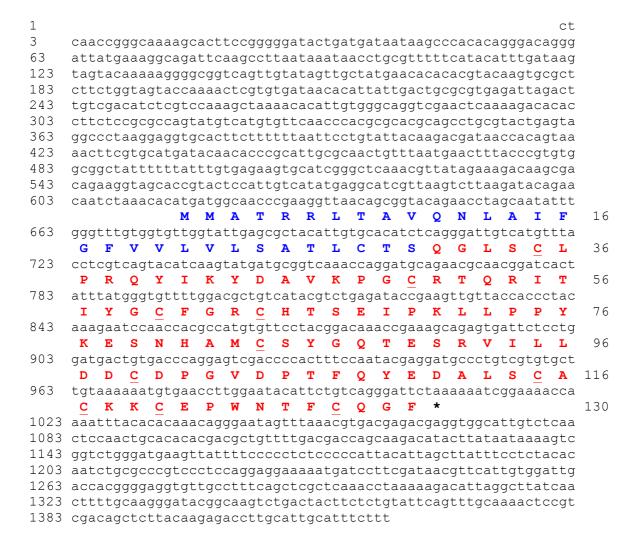


Fig. S22. A. rubens glycoprotein hormone β-5-type (GPB5)-type precursor 3 (ArGPB5-

3). The nucleotide sequence (lowercase, 1420 bases) encoding the precursor protein (uppercase, 130 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative glycoprotein hormone β -5-type (GPB5)-type subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1107934 and has been deposited in GenBank under accession number KT601725.

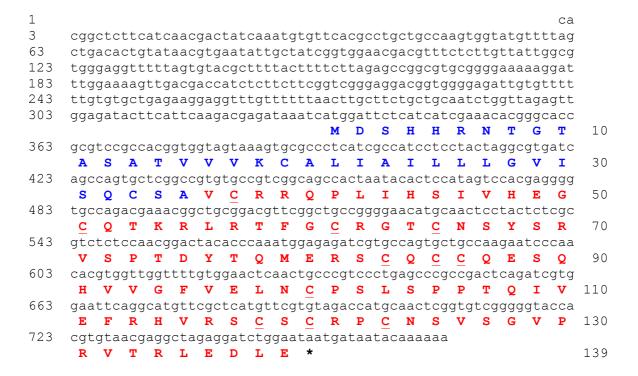


Fig. S23. *A. rubens* bursicon α -type precursor (ArBAP). The nucleotide sequence (lowercase, 767 bases) encoding the precursor protein (uppercase, 139 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative bursicon- α subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1079029 and has been deposited in GenBank under accession number KT601726.

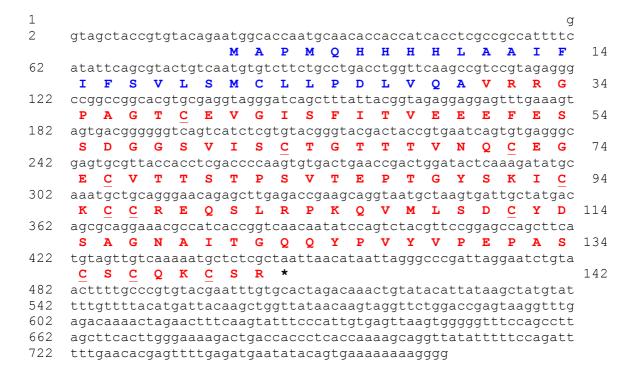


Fig. S24. *A. rubens* bursicon β-type precursor (ArBBP). The nucleotide sequence (lowercase, 766 bases) encoding the precursor protein (uppercase, 142 amino acid residues) is shown. The predicted signal peptide is represented in blue and the putative bursicon-β subunit (with cysteine (C) residues underlined) is represented in red. The asterisk shows the position of the stop codon. This sequence was determined from contig 1078917 and has been deposited in GenBank under accession number KT601727.

```
tcccgttttgattttgtggctgtttgttatcgcttgctactatcgacaactcagacgctt
61
    acatcccttcaaacacggtaaaaaatggcaaactaccgtctcatcctcgaggccacctgt
                          M A N Y R L I L E A T C
    ctccttgtcctccttataaacaccgccctctacgccgaggctgctgaaaaatactgcgat
121
     L L V L L I N T A L Y A E A A E K Y C D
181
    gaggatttccacatggccgtctacagaacgtgtacagagcacaagcgcagtgggagatcc
     E D F H M A V Y R T C T E H K R S G R S
    gccttcagcttgaacgactttttccgaagtaactcaaaacggaccgctggatctccccga
     A F S L N D F F R S N S K R T A G S P R
    ccaqacqacqactttttcctqactatqcaqaaqaqacctqaqacttacqttqqaatqqqa
301
     P D D D F F L T M Q K R P E T Y V G M G
    tcctactqctqtttqqtqqqqtqcacacqtqaccaactatcacaaqtctqctaaaactta
     S Y C C L V G C T R D Q L S Q V C *
421
   aagaaatcccccattaccactttttgaaaccccttgacaacggaacacacccaacactcc
481 agttctattaatagctttgtcatgaacaaatttaagcatagcccaaccacatggttggaa
541 cagccccaaaataagtaatttttgcaatttctattgtccctgctctatgctctatgataa
601 taagcgacgtataaggatatttagttgactatttcatgggtggctttttatcagagatgc
aatgaagtacttatcgaccccttcccaatggcgtcacccttgcccagtgccctcactgtg
721
   781
    qtqtacqtqacctctqttacqttttacatqtatacaaqqqqtctattqttcatcaacqaa
841
   ccctatctcttgttaaaaagagggttaagaaaagacaactgtcagaaacgggaagtaag
901 \quad \texttt{gaatttgactctcaaataaataaattacgaaaaaaatataattactcgctttaaaatgctt}
1021 agcaacaaatagtttatggcctgtatgcatagagctatgcctgtatgacgttttgaccca
1081 agcacaatccatcctttctcggtcatttagccttcgaaaaaagactctgctatatagggt
1141 attttttgtaaatacttaagttcaactataatttaaaaacatgacgtcgctgaccaccat
1201 ctttggtgtacaatggaaaagtgcacgcgagtacgcgctgcacacaatcagccaataata
1261 gcctttcgcacgtgcacaatatcaaagatggcggtcaaggtctacacagcattttctcat
1321 tgaatgctccccctatgtgaacagagacgttgtataccaaacagatctgagacctttgtt
1381 tgtaacctgtacgaaaacgctttgctggaatttgaggtcgagtgataagatcgtggagtg
1441\ {\tt tttttgtgggggcgagtgcaatgacacgccagtgttaaatgattaggctcttccaaacca}
1501 cggcttgggcttaggcttccggcttaggcttaggcttagcactttctctcacacgtcttc
1561 aaacagacagaggttatcaagcctgagccacaagccacctgggacagggcttcggg
1621 agccgaagcaaaagctgtggttttgaaactatcagggtctatgttgagtgaagatggagg
1681 cacactgtgacagttctccacttttaggcttttggtacaatcttaaaacttccgggtcag
1741 aattgatccacattcagagtccagtgcctggtgtctgatgcctgacccatttctgggtca
1801 tactqacccaqaaaatgagtcttgatgacccagaaaagggtcaagtacacactgagttaa
1861 gaccccaattttaaaccatgtttcttggccaacctgacctagaaatgggtcagctccatc
1921 acccggaatccgggtcaatattattctgacccaggagattttagagtgcctcttgttatg
1981 taggctatgcaggtatcattaagtaaaatcatgctgaactatgtcttgactgttattcca
2041 tttttqtttcaatttcacatccctqaccqtqattqtaaqqataatccqaqctqtctctct
2161 tgtgcaaaggcttgaccaagttttgtcgtacctctcgtatgagtgtgagatgcttctgga
2221\ {\tt aaagtgtgtccataggaataatttgcctagtcagaagaaggtggtgaaaagcttccac}
2281 caactgctgtggacgttgggtgtggggtagggggaaattaactccatcgagatacagt
2341 agtatacgggatggatgtatttcgtaaatgtatatacttttcaacatctttgaattatga
2461 tgtagaatatttcttcattatgcagagtagtaaatagtaggaatataagtttttaagttt
2521 gaggagtaaaaaatctagttatttcaaaattgcattccaaatcatggtttatgttaactt
2581 taaaaacatcttgtgttcaacagaaagtaggctttggctttgaagaaaccaataaaagga
2641 aaataaatgaacttttcatcgtaaatgcatctgttattactatagactgtgtggcaggac
2701 atgttgtgcccagtgatggcgaactgtgtacaaacctcttctattatagagcgttctttc
2761 gaagcaacctcttccccatcgatgtaaatttcccatgagggactgaaatttcgagggaaa
2881 ttttcgcgcttctcggttgaggcgcggtgacgatg
```

32

72

92

109

Fig. S25. A. rubens relaxin-like gonad-stimulating peptide (RGP) precursor (ArRGPP).

The nucleotide sequence (lowercase, 2915 bases) encoding the precursor protein (uppercase, 109 amino acid residues) is shown. The predicted signal peptide is represented in blue, two putative relaxin-like gonad stimulating peptides (RGP) (with cysteine (C) residues underlined) are represented in red, a putative C-peptide is represented in black and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1122961 and has been deposited in GenBank under accession number KT601728.

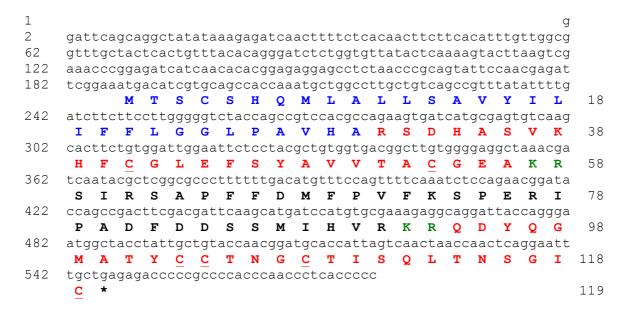


Fig. S26. *A. rubens* relaxin-like peptide (RLP) precursor 2 (ArRLPP2). The nucleotide sequence (lowercase, 578 bases) encoding the precursor protein (uppercase, 119 amino acid residues) is shown. The predicted signal peptide is represented in blue, two putative relaxin-like peptides (RLP) (with cysteine (C) residues underlined) are represented in red, a putative C-peptide is represented in black and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1055767 and has been deposited in GenBank under accession number KT601729.

```
1
2
    \verb| aattttacgcgtatataaaagaaagtgtgtgactcctttctcgatccaatcagtctctgg| \\
62
    tggtcattgtgagctcgtgatcttagatgtctcaatagtttcttaaaaaacggcaaggag
122
   aaacattcttcatcagatgctgagaactcagtttatttaagacttggacattacggagct
182
   agttatagacaaagaagcttcatcagtactactacaagaagcttgtgtatagccaagatc
242
   \verb|caagacg|| ttttcaaggtgtggacatgttcagaaacacatcgacaatgagggcgcttctc||
302
                        MFRNTSTMRALL
                                                      12
362
    ctacttgacgtcattttcgttgccttggttctaccaatcacggcatggccgaagatctgc
    LLDVIFVALVLPITAWPKIC
                                                      32
422
    qqaqaqcaactqqtqqaqacaqtqtctctaqtttqcaqtacacqaqqtttctacaqccat
    GEQLVETVSLVCSTRGFYSH
                                                      52
482
    agagattcaaaacgagatgttgaggttttccagaatgaacgagcggctaagagcttcctt
                                                      72
    RDSKRDVEVFQNERAAKSFL
542
    ggttcccgcatcggctcgaggcagcgacggcggacgggtagaatcgccacggagtgctgc
    G S R I G S R Q R R R T G R I A T E C C
                                                      92
602
   gacagaatatgtagcttcgacatcgtggagagttactgtaacccttggcccgtggcgata
    DRICSFDIVESYCNPWPVAI
                                                     112
662
    qaqtcqcqaqacccqcctctatctcccqtcqcaccqqqcaqqqtacqqqaaqataaatct
    E S R D P P L S P V A P G R V R E D K S
722
    \tt gcggacgtggactacatgtacaacccggatgttgtggacgtggaagaagccaactcggtc
    A D V D Y M Y N P D V V D V E E A N S V
782
    atacagcgcgaggaagatcttatcgatgacatcgagacgcaagaacaggaaatcgaacaa
    I Q R E E D L I D D I E T Q E Q E I E Q
    qatqaqqaacaqaacatqcaaactctacccqaaqaaqacqctqaaqacacaqatatacqa
    D E E Q N M Q T L P E E D A E D T D I R
                                                     192
902
    qaaccaqagqacqttqaaqaatctttcccaqttccaqtcccqacqaaaaaqagqagaaaq
    E P E D V E E S F P V P V P T K K R R K
                                                     212
    gtcgagggaaggagtcgaaagagagcaagaacaaaggtggaaaatcggaagggaaaaac
    V E G R R S K E S K N K G G K S E G K N
                                                     232
1022 aagaagagaagcgggagcagggaaggggacggtcttctcgtcggagtcgtggaaaaagc
    K K R S G S R E G G R S S R R S R G K S
1082 tcccgatcgaagaagcagagggatggccgggagagaagcaagagatgggaggggcttgat
     S R S K K Q R D G R E R S K R W E G L D
1142 acatcgcacccagtaaaggaacctactgcaaggagtgttttaggtcgagtcgataccaga
     T S H P V K E P T A R S V L G R V D T R
                                                     292
1202 ccattccgtaactttctgtacaaccgctacactgtagacgagaagcgagacactgaaaga
     P F R N F L Y N R Y T V D E K R D T E R
1262 gaaagctatcgagcggtggcgcccttactgggtacaattcacaccgagggggatcccaa
    E S Y R A V A P L T G Y N S H R G G S Q
1322 cctgacaaccacccaacactagcggcgctatacaaccttgcagttaaactcgccaaggga
     P D N H P T L A A L Y N L A V K L A K G
1382 cttcaacattgacgaataatagtacaagaggaatttatatgaaagtacagcgcctctaa
                                                     355
     L Q H *
1442 cgtggttataattataacctaaaaggggtcatgatagttcaatttttgtccagttgtgata
1562 gtttataattcaattttgtttcaacttaattaaatccactgccgatctttaaaagaatca
1622 ccaacattgatattataactttttgaatctcattaattgacttcatgtatcaccattgga
1682 aacgcaatgcttactgtctaaaacagtttgtttctaaagttatgtgttcaacaaatcttc
1802 gccacctagcgtt
```

Fig. S27. *A. rubens* insulin-like growth factor (IGF)-type precursor 1 (ArIGFP1). The nucleotide sequence (lowercase, 1814 bases) encoding the precursor protein (uppercase, 355 amino acid residues) is shown. The predicted signal peptide is represented in blue, putative B-domain and A-domains (with cysteine (C) residues underlined) are represented in red, a putative C-peptide is represented in black, a putative D-domain is represented in olive, a putative E-domain is represented in maroon and dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1114876 and has been deposited in GenBank under accession number KT601730.

```
1
    agattattcaagatgaaccaatatcagctgatagtactattcgaagtgttggctcacgcc
              M N Q Y Q L I V L F E V L A H A
                                                    16
    {\tt tctatgcta} {\tt aactatgcctctccagtccagctgtgtggaagggagttaacggagacgctt}
    36
    aggtcaatctgtggtgatagaggttactattcacctggacaaagtttttcacgtcgagcg
121
    R S I C G D R G Y Y S P G Q S F S R R A
                                                    56
   ccgacacacgacggcattgcgacaagatgttgccagtccctttgcgagtcttcaatcctc
    P T H D G I A T R C C Q S L C E S S I L
   qaqacctactqcaatctaccaqcaccqccatcqcaaacaaccatccacaqcaqcccca
    E T Y C N L P A P P S Q T Q P S T A A P
301
    acaacaacaacaaaaatggcgccattaacaqaaqaccgacggacgaaagatgtggtcgtc
    T T T K M A P L T E D R R T K D V V 116
    \tt gactactcagaccagctagccaccgaaggctcccaaatgtcgcgtgtggacggcgtactc
    D Y S D Q L A T E G S Q M S R V D G V L
                                                    136
421
    acgcatgacaccgtgacgaacagaagtaaaaccacgacagagtcgaacgagggcagttat
    T H D T V T N R S K T T T E S N E G S Y
481
   \tt gacaacgaagaggcgcaccttatgataaaccagacgatagcagtccatcagaacgagga
    541
    qaaaqcatacaqqatqaaqataacqaaqtaaataaaccqqaaccaaacaacattaqqqac
    E S I Q D E D N E V N K P E P N N I R D
601
   aattcaaaagagcgaggtagaaataggacccacaagggggttagctcggagagaagagca
    N S K E R G R N R T H K G V S S E R R A
661
    aataacagcaggaggagggtctaagcagtgaaaggaggatcaagcagcagtagaagg
    N N S R R R G L S S E R R G S S S S R R
    gaagagaagcttagaaggagacgacagagacatcgggaaaagggagctgagagaacaacgg
    EEKLRRRQRHRERELREQR
    K Q S N S K R K S K G D K K D H S V A A 276
    \verb|actactccactcgcagtacaagagcgccctctaaagaatggtggtcggaacagcacatcc|
    T T P L A V Q E R P L K N G G R N S T S 296
901
    GEHSSVNGTETDTAGAGSPE
   gtcaagaaagatgatctgatcaccacgattacggccgtgctcagtgatatgataggcttc
    V K K D D L I T T I T A V L S D M I G F
1021 caaccagacaatggtaatcgctgatcaccccatcaacgagtcgagaactagacgacttta
    Q P D N G N R *
                                                    343
1081 acccaatgatttaaagttcagtttgttagtattaactttcactgctcatcactgatcgat
1141 ttcaaactgacttttataaaaactgggagcttggatttggaaagcaacggaaatatttat
1201 tgatatttattatcccagaccccttggactatggataggtaaagacaatgagagatagag
1261 agaggggaaggtgtaattagacatgatgagaaagctatatagcgatgtgacgatttgtgt
1381 accaaatgtacaatttacgaaataatttaaatggttgttaactcttagggcggatgaagc
1441 gctctttcattgatcagatggaagttctcttcattttatacactttagaaggaaattttc
1501 ttcccaagatcaaaaggacgtgttcttcccacgatcatgctcatagcgtccaaaaaactt
1561 acagatacgaaatagcgctcttcctacaaggaattgctcacagtaaagtttataagggca
1621 cccagtctgaaaaacctttttttccacaggcagctgacg
```

Fig. S28. A. rubens insulin-like growth factor (IGF)-type precursor 2 (ArIGFP2). The nucleotide sequence (lowercase, 1659 bases) encoding the precursor protein (uppercase, 343 amino acid residues) is shown. The predicted signal peptide is represented in blue, putative B-domain and A-domains (with cysteine residues underlined) are represented in red, a putative C-peptide is represented in black, a putative E-domain is represented in maroon and dibasic cleavage sites are represented in green The asterisk shows the position of the stop codon. This sequence was determined from contig 1112524 and has been deposited in GenBank under accession number KT601731.

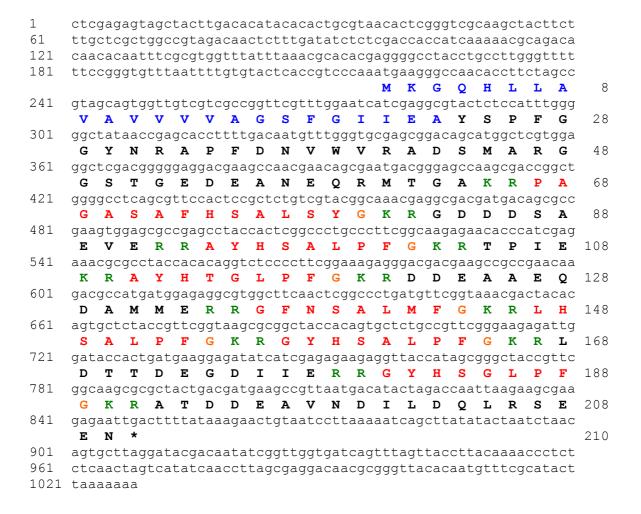


Fig. S29. *A. rubens* **L-type SALMFamide precursor.** The nucleotide sequence (lowercase, 1028 bases) encoding the precursor protein (uppercase, 210 amino acid residues) is shown. The predicted signal peptide is represented in blue, seven putative L-type SALMFamide peptides are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contigs 1073975 [1-46], 192719 [47-61], 104468 [62-74] and 1057915 [75-210] and has been deposited in GenBank under accession number KT601732.

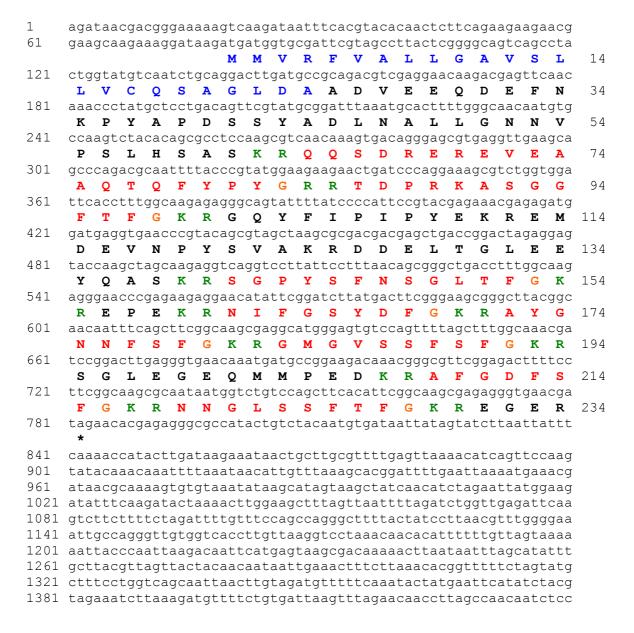


Fig. S30. *A. rubens* **F-type SALMFamide precursor**. The nucleotide sequence (lowercase, 1440 bases) encoding the precursor protein (uppercase, 233 amino acid residues) is shown. The predicted signal peptide is represented in blue, seven putative F-type SALMFamide peptides and one putative L-type SALMFamide peptide are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1108394 and has been deposited in GenBank under accession number KP330476.

```
acacttccatctcggcatcaacccggagcagacgaaaggaattgtacatcatctttttgc
61
    tgcgttacttttcccgccagtattcacgaccctgactgaaacattggactacttgttgtt
121
    181
    atcttgcaaccagagcttgaaatggcaagattggacttgattctacttcttatcggtg
                        \begin{smallmatrix} M & A & R & L & D & L & I & L & L & L & S & V \\ \end{smallmatrix} 
                                                           13
241
    \verb|gtagtggctgcagtggtgctccagtgccgggctgacgatgaaaacaccgatgggcaaggc|
     V V A A V V L Q C R A D D E N T D G Q G
    qccqccqatqtaqacqaaqttctaaaacaactatatqcaqaaqatqaqaatqatqacqat
     AADVDEVLKQLYAEDENDD
                                                           53
361
    aaacqcqccaaccactacqctaqtqqacqccaacqtacaaqqqqqaaqqccqqctaccqt
     K R A N H Y A S G R Q R T R G K A G Y R
                                                           73
421
    cgggtgactaaatccgacatggacgcagtagccggactagaagacatggatgaggagaaa
     R V T K S D M D A V A G L E D M D E E K
                                                           93
    \verb|cgggctaactaccgcgcctcagtcgtggtggcaagaagggattccgtagatacacc|\\
481
     RANYRAS V SRGGKKGFRRYT
                                                          113
    K S D G D E E E I P E L E E M D E E K R
                                                          133
601
    \verb|gccaactacagagccactatgaacggcggacgtaaaccccaccgcttcaaccgattcacc|
     ANYRAT MNGGRKPHRFNRFT
661
    aaggccgacggagacggaggaggaggaaggaccggaggatctggagggggaggatata
     K A D G D G E E E E G P E D L E G E D I
721
    gacgaagataaacgagccaactaccacgcccgaggtggcaagcctagaggtggattccgc
     D E D K R A N Y H A R G G K P R G G F R
781
    aggtacacqaagtcggacqaggatcaagatgctatggagcaggcccctgccgaggaaatg
     RYTKSDEDQDAMEQAPAEEM
                                                          213
    D E D K R A N Y H A R G G K P R G G F R
                                                          233
901
    aggtacacaaagtcagacgaagatctcgccatgggggaagcagcagccgatgaactggac
     R Y T K S D E D L A M G E A A A D E L D
                                                          253
    qaggaaaaqcqcqccaattaccacqcttcacqcqqqaqacaqaqqqqcaacqqqcqtcqa
     E E K R A N Y H A S R G R Q R G N G R R
1021 aattagattgtgacaaaaagtaaaatatcagttacacaacagtttaaactctttttcat
     N *
                                                          274
1081 tatctggtctggcggaaatcatgcagagggggaaggggggacacacatcaaaacaaaatc
1141 taagcatcttaactttcagtaagttttcatttttgacctagggatttattgccgcaatgtc
1201 tttgacttaaattatcgctttagtgtattcaatcacgaaagctcgtttaaaccactcccc
1261 aatttttagtggaaaacattcactcatggccagagaaattaatattgtcagatattatgt
1321 ctttatctaaagacgactgacaaggtggcagcaacatggctaaaacaaattatgttttca
1381 aaaaattggctttacacagacacccattgagctttgaagcaagtgttggcgaaaacagaa
1441 acaaaacaaaccaataaaagatgaaaagttaaagtgaccttggacaaagacaagattttt
1501 cggcaattataccagctcgacgtgatttgtggtatagagtatctgcgccacaatattgat
1561 ggaggctcttgtgattagattgtatttttcttaagttcataaacaaatgtgcattatgta
1621\ taacttcacatgcatttggttaccgattactctgtaacattaatcattgtattttataag
1681 tgtaaatatcctcatcattctatatttgtaagtgttgctggtgttgctcttcaaaaacaa
1741 aacacaaaggcgaagaatgaaagaaaaagagtaacaaaacgcccataaactgagattcaa
1801 attggttgcttggtgcattctggtgcattttagctgcaggtataaaagaaaactgagtg
1861 tacataaaagcagaaaatagatgattgcccggagttggttttttctttaagtggaaaacc
1921 ttctatttctttcgagggaaattttggacacttgatctggaggaaggctagattcagaag
1981 ttctaacaaacaacaagacctttcaagcttttaacacacaaactataaaagcgcgcaaa
2041 aacaagctcggcataggcttacttattattatgaagtgcacttcaattagcaaccatgtc
2101 accatacatacttattcaaatttatacatgcttattcaaaatttaatcttggaaaaatct
2161 ggtaaaaataaatgtattagtttgccaatgcttaaaggaagaaagggtactgtgcgacaa
2221 ctttgactggtaactaatggttaattagtggttgaattcgaaaagtctatggtttcattg
2281 tgtggtaactaattttaatcaactgatcaaccatgttgacaaagaaggcgattgtcaaac
2341 tataatcaaccqqaqttaccaqtccaqcttqctcqtaatqctattttttcqtccqttaqt
2401 gtaaaccacattctctcgatttaacagtagtttaaacatgaaacgcacctgagaattaag
2461 tgatgagtagccgtaacagctgcaaattatgtggcttttcttgtcatcgaaagcaaaata
2521 aagaagtagtcaaagttaaaccaaaaaaaaaa
```

Fig. S31. *A. rubens* **AN peptide-type precursor (ArANPP)**. The nucleotide sequence (lowercase, 2553 bases) encoding the precursor protein (uppercase, 274 amino acid residues) is shown. The predicted signal peptide is represented in blue, six putative AN-type peptides are represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1121231 and has been deposited in GenBank under accession number KT601733.

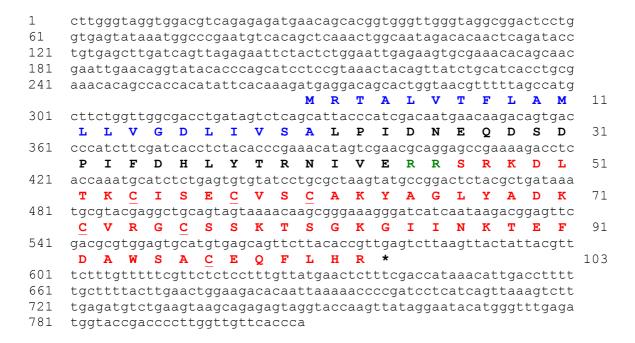


Fig. S32. *A. rubens* **Arnp11 precursor**. The nucleotide sequence (lowercase, 808 bases) encoding the precursor protein (uppercase, 103 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative Arnp11 peptide (with cysteine (C) residues underlined) is represented in red and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1082355 and has been deposited in GenBank under accession number KT601734.

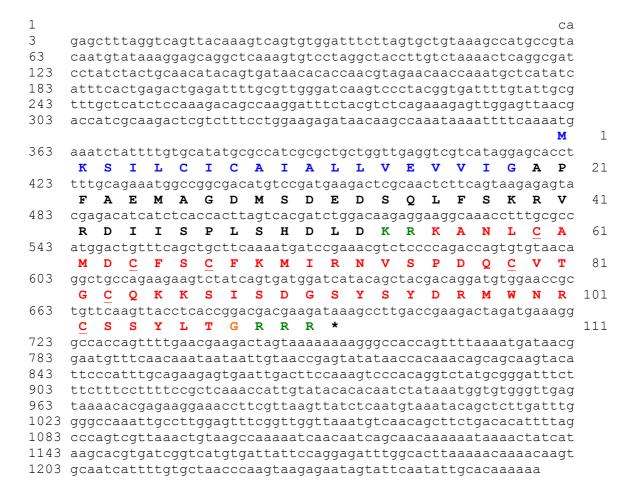


Fig. S33. *A. rubens* **Arnp15a precursor**. The nucleotide sequence (lowercase, 1258 bases) encoding the precursor protein (uppercase, 111 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative Arnp15 peptide (with cysteine (C) residues underlined) is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and putative dibasic/tribasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1103866 and has been deposited in GenBank under accession number KT601735.

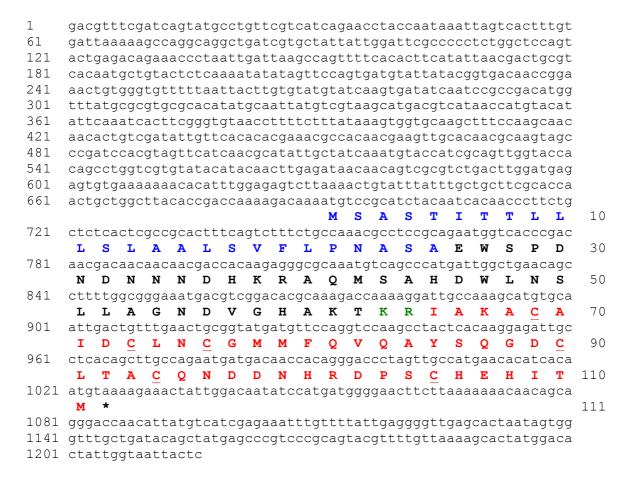


Fig. S34. *A. rubens* **Arnp15b precursor**. The nucleotide sequence (lowercase, 1216 bases) encoding the precursor protein (uppercase, 111 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative Arnp15 peptide (with cysteine (C) residues underlined) is represented in red and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1102592 and has been deposited in GenBank under accession number KT601736.

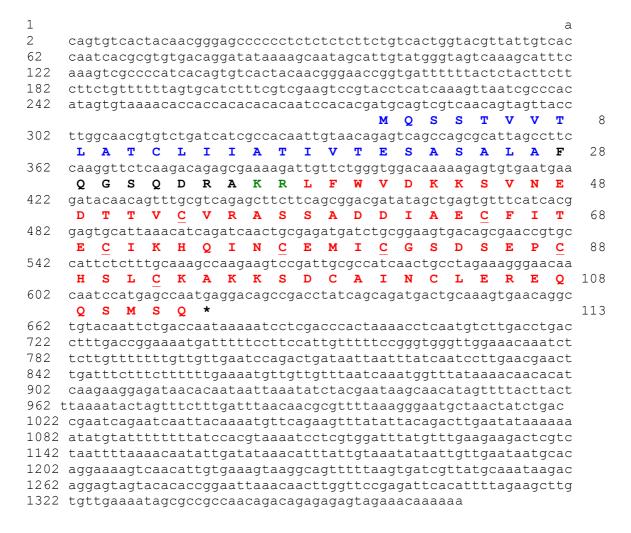


Fig. S35. *A. rubens* **Arnp18 precursor**. The nucleotide sequence (lowercase, 1368 bases) encoding the precursor protein (uppercase, 113 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative Arnp18 peptide (with cysteine (C) residues underlined) is represented in red and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1106730 and has been deposited in GenBank under accession number KT601737.

```
atgcagagaaaaaatgaaaaaacaaacacacaaacagagaaaaaatgaataaaatgggt
61
    gcagtaaacaaatgggtttattggagaagaggtctgaatgggctcactgtgtccgcaagt
121
   tcatctctctagttctctcccttccactctgggtcattctgctcagtcactgcactaaat
181
    ccctgtaagttattggattcaggtgtatccacaagatcgctgagtttgacattgtccagt
241
    ctctggatgatggcttcaagttcgtcttcatcttcctcaaagtggcatcccgtgtaggca
301
    ttattctccatacagtaacgtaggaaattgttccagttgttgtccttgcatagcatgtat
361
    gggttacccactaaaaccaacaacgccttcgctcgtgtgactgcaacattcaatcttttg
    gggttcttgagaaagccgagttggtagtcaatgtccatctggatgtgttctgctttagtg
481 ctgcggactgtggacacgatgatgaccaacctctcctgaccttgaaactcttcaacggac
541
   cccaccttgatgttcttatggtgacgcttctgtaggactttctgtagctttttgaacctgt
601
    cqacqatatqqtqaaatqatcccaatqtcaqactccttaatttqtctcccccctcqctta
661
    ttcagcaaatcctccacatattccacaacgatcgtcacctcggctaggttgaagaaggac
721 ggactatgtccctcacgctcatccttcccttctaccccatgaaacaaaatgggaaagttc
781 ttctgaggtagcctgtcccaactgcacagcgacgtcctcacagcttcatctgcatgaacc
841 tccaqctcatttqcataqaacatctqattqqqcaqtttqaqqataqcaqqqttqaqaqcqa
901 tagttacgcagtagcttggtcaggactctgctgtcgtaatgcggctcagcctcctgacgt
961 tgatagatatcgcattcagtcatcagtcgctcaagtaaagataaacctagtccatgttca
1021 atagcaagaggtgacctcagaatgggaccaagttgcttggggtcaccggccagaaccact
1081 tgacctcctttggtgttctttggattcaggagattagacagggctataatggcttccggt
1141 tctactqcatqaccaqcctcatcaatqaatacatqaqtqaaatqattcqaqqqaaaqtca
1201 gctgtagacagtctgcctgcagtacacagtgtggtcaccaggattcttttttccatcagc
1261 tcttctttagaagggtagaaggtttccctggttccccgcttatcaaagttacacacatcc
1321 ttctccttcaggatttcatcaacagtcttccaatctcgactctgtgcgttgattctcatg
1381 aggctagtcttggcaatgggcgtgccccggtcagaaggcgtgtcgcgatgaggtcagct
1441 gcactattagatggcgcgcacgcaaggatataacttccacgaagttgatggtatgtctgc
1501 ttgatggcctctacaatggtgactgtcttgcctgtgcctggtgggccaaagactaggtaa
1561 ggactagggcgagatgacccacaaacaatgtgcttaacggctttgatctgttccacgttg
                               MCLTALICSTL
                                                                11
1621 gtggcgagtttcttgtcaaagaacagtctggcaaaatcactgtcatcacctggcacggtg
      V A S F L S K N S L A K S L S S P G T V
                                                                31
1681 cagagtttctttgtcgttccgatgttggatcggtcgggaaagaggagtttacccagatca
      Q S F F V V P M L D R S G K R S L P R S
                                                                51
1741 ttcttgagggcgtcctctatagcattgtgctgtagtcgcaacggcaaacggttgaacgta
      F L R A S S I A L C C S R N G K R L N V
                                                                71
1801 aactcaatgtcaaacttgagtccttgggtaaacctattcaccaaatctttggagaatccc
     N S M S N L S P W V N L F T K S L E N P
                                                                91
1861 aagaccacttcttccatttccacccggtggacataacccttgtaggtcttgccatcttca
      KTTSSISTRWT
                                                               102
1921 gagtcagtactgaagcgggcaaacagatgatcacctttgaggaccgagggacggttctca
1981 gcaagacctggaattctaagtcgtaggaatttgccgtcctgctgcatgggcactcccttc
2041 atatcqtacctcttqatqtcaacttccatctqqcactcctctacqtaqaqaaqqttqqat
2101 aacttactgacgtagttcacccaagccagatcaggcttaatggtttttctgacttcctct
2161\ attttgtgtttattgttcagcatccagcggagatggtctggaattgggtaatacttggct
2221 \ {\tt tgtctcatcttcagtttgctcttggaggactgtggtaggggcactccttcaacagtctct}
2281 cctttccqttqccttcttqccacccttqqtqqqcqttqqtaaqqaacaqacqqqqqtaqa
2341 tccttcgctacctcgctctccacctccacagtcacaaagcgcagaatatgaaatacctcg
2401 ttggatggtactttcttgaaggcaagtatgagtgggctgctgagaacacccacggagtct
2461 cccactacagccacctgaatacagtactcactccctggtttgaggtggaccactttctta
2581 caagataggaatgcggtagttaagtcatcacttacattcttgatggtaatcctgaaggtt
2641 \ tgactggcactccgcgccagtttcattgcaatggtcccatcctgaccatcatgctcactg
2701 gcaacctctacgccattcttattttcaatgaactttgacctgctcttcttgacacatttg
2761 aagatcatttqctqttqatqtctactcccttqqaqatqtqcattaaaaqactqatqaqaa
2821 ttcccctggacgccacactcctcgcagtaaaacgggtacggctccggcccgttggacttg
2881 cctcctttacctqqqctctqcttcattqcaccqqccctcqttqcaccqqccctcqttqta
2941 ctggccctcgttgtgtctttctggtcaagggcatgcttctttagggtggtgaactgatcc
3001 acagaaacaaaactgttctgaacttggactcggccatctttgatgttgatcttgccattg
3061 gtcttcaaagcatagagtatgcttgaaaagttgaaacctattcctcctttctctaatctt
```

Fig. S36. *A. rubens* **Arnp21 precursor**. The nucleotide sequence (lowercase, 3488 bases) encoding the precursor protein (uppercase, 102 amino acid residues) is shown. The predicted signal peptide is represented in blue, putative Arnp21 peptides (with cysteine (C) residues underlined) is represented in red, C-terminal glycine (G) residues that are putative substrates for amidation are represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1124521 and has been deposited in GenBank under accession number KT601738.

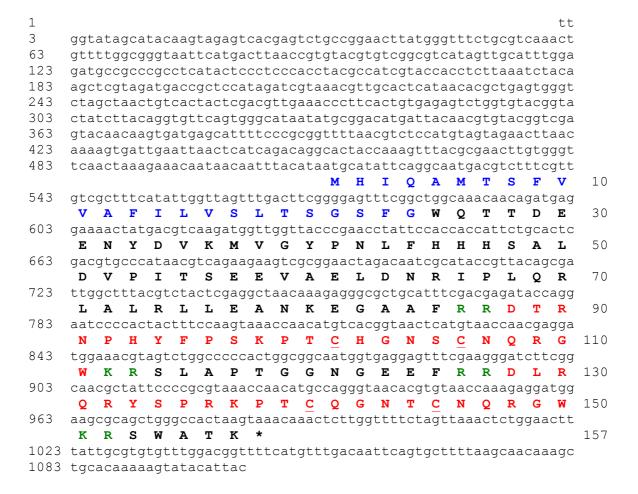


Fig. S37. *A. rubens* **Arnp22 precursor**. The nucleotide sequence (lowercase, 1103 bases) encoding the precursor protein (uppercase, 157 amino acid residues) is shown. The predicted signal peptide is represented in blue, putative Arnp22 peptides (with cysteine (C) residues underlined) is represented in red and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1098506 and has been deposited in GenBank under accession number KT601739.

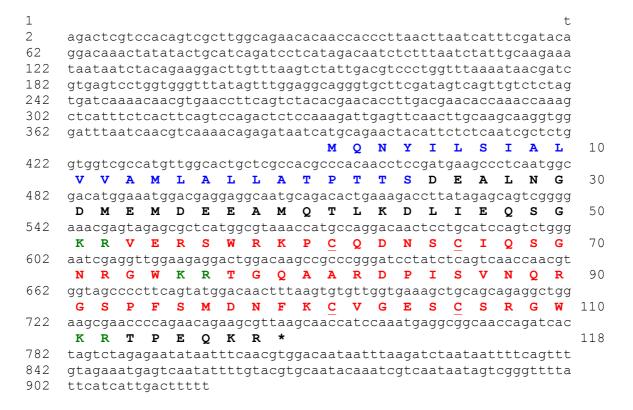


Fig. S38. *A. rubens* **Arnp23 precursor**. The nucleotide sequence (lowercase, 918 bases) encoding the precursor protein (uppercase, 118 amino acid residues) is shown. The predicted signal peptide is represented in blue, putative Arnp23 peptides (with cysteine (C) residues underlined) are represented in red and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1089690 and has been deposited in GenBank under accession number KT601740.

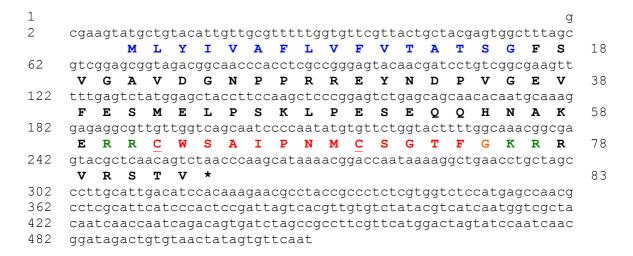


Fig. S39. *A. rubens* **Arnp24 precursor**. The nucleotide sequence (lowercase, 510 bases) encoding the precursor protein (uppercase, 83 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative Arnp24 peptide (with cysteine (C) residues underlined) is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and putative dibasic cleavage sites are represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1042300 and has been deposited in GenBank under accession number KT601741.

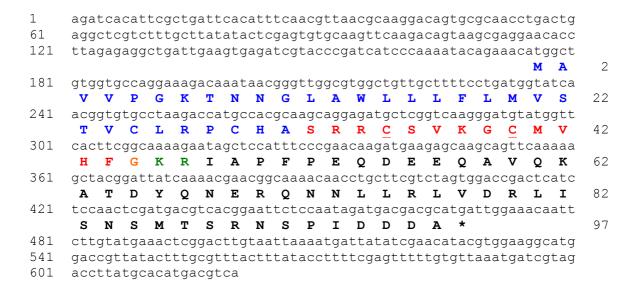


Fig. S40. *A. rubens* **Arnp25 precursor**. The nucleotide sequence (lowercase, 620 bases) encoding the precursor protein (uppercase, 97 amino acid residues) is shown. The predicted signal peptide is represented in blue, a putative Arnp25 peptide (with cysteine (C) residues underlined) is represented in red, a C-terminal glycine (G) residue that is a putative substrate for amidation is represented in orange and a putative dibasic cleavage site is represented in green. The asterisk shows the position of the stop codon. This sequence was determined from contig 1062339 and has been deposited in GenBank under accession number KT601742.