Aligned precursor candidates

Alignments of candidate miRNA precursor sequences obtained from ranked clusters, showing the predicted miRNA sequence as identified by findMiRNA (Yellow highlight). Suggested extension to full miRNA length that is based partly on target site optimization is shown (Grey highlight). Precursor alignments were done using T-COFFEE Version_1.37

cmr3

```
AT1G69800 ------------------CTAGAGTTCTCCTGAACACTTCATTGGAAA-TTTGTTATT-------------CAGTAAGCTAACAGTTAATTCCACTGAAGTGTTTGGGGGAACTCCCG------------------
AT1G27000 TTCTTGGTTGGATGTCTCCTAGAGTTCCTCTGAGCACTTCATTGGGGA-TACAATTTTTC--------------------TAAATGATTATCCACTGAAGTGTTTGGGGGAACTCCCGGACCCATTCAACGGATTT
AT1G69800 ------------------CTAGAGTTCCTCTGAGCACTTCATTGGAGA-TACAATTTT-----------------TTATAAAATAGTTTTCTACTGAAGTGTTTGGGGGAACTCCCG------------------
AT1G69800 --TAAAGTCAAATGTCCCCTTGAGTTCCCTTAAACGCTTCATTGTTCA-TACTTTGTTATCATCTATCGATCGATCAATCAATCTGATGAAC-ACTGAAGTGTTTGGGGGGACTCTAGGTGACAT--CTTGGCACC
AT1G27000 -------------GTCCCCATGAGTTCCCTTTAACGCTTCATTGTTAAATACT-------------CAAAGCCACATTGGTTTGTATACAAC-ACTGAAGTGTTTGGGGGGACTCTTGGTGT--------------
AT1G27000 -AGAAAGTCAAATGTCCACATGAGTTCCCTTTAACGCTTCATTGTTGAATACT-------------CAAAGCCACATTGGTTTGTATATAAC-ACTGAAGTGTTTGGGGGGACTCTTGGTGTCATTCTGGCGTT--
                          *  ******   * * * ********   * *                                     * ***************** ****  *
```

cmr5

```
AT1G73360 ----ATAG-AAGTCAATGCCCGAACCATA-------------C-------------------------------------AAACAAA---GTCAAAGTTGCA---TGGTGCGTTCTTGACTTCC---AG--
AT5G46890 CCAAGTAGAAAGTCAATACCCGAACCAGCTACGCTTTTCTGTCAAACCGAAAATCGGTTTAAGACCGATAATCCGGTTCAGAAAACCAAAAGGTCAACTTACCAACCTGGTGCGCTCTTGACTTCCTCTAATT
          *** ******** *********              *                                   **** **   *****  *  **   ******* ***********   *
```

cmr6

```
AT2G34210 ---------TATGCATTAC-AGGGCAAGATCACCATTGGCAGAGATCTATTACTT--CATTCTT-GCAT---CATATGCATAAAT-GTTTGTGGT---------------GAGCTCTCTGCCAAAGGAGATTTGCCCGGTAATTCTCT----------
AT1G29270 ---------AATGCATTAC-AGGGTAAGATCTCTATTGGCAGGAAACCATTACTT--AGATCTTTGCATCTCTTTATGCATTGCTTTTAATTAGT---------------GAGTTATCTGCCAAAGGAGATTTGCCCTGTAATTCTTC----------
AT1G63010 --------CTCACTAGTTTTAGGGCG-CCTCTCCATTGGCAGG--TCCTTTACTTCCAAATATACACATACATATATGAAT--ATCGAAAATTTCCGATGATCGATTTATAAATGACCTGCCAAAGGAGAGTTGCCCTGAAACTGGTTCA--------
AT2G34210 -----------------AC-TGGGCGAATACTCCTATGGCAGATCGC-ATTGGCT--AGA--------------TATGCAAGTAAAAT-----------------------GCTTCTCTGCCAAAGGAGATTTGCCCCG------------------C
AT2G34210 -----------------AC-AGGGCGAATCCTCTATTGGCAGTGGAAGTTGATG---------------------ACCCTTATAT-GTTATTTTCTCATCAT---------TTTCCTCTGCCAAAGGAGATTTGCCTCG------------------C
AT5G62165 ACGATAAGCGGAGCAGTAATAGGGC-ATCTTTCTATTGGCAGGCGAC----------------TTGGCTATTTGTAT-CTTTTGT--------GTTCTTGACTATTGGCTATGTCACTTGCCAAAGGAGAGTTGCCCTGTCACTGCTTCCGCTTAAAC
                            ***        *   ******                                *                                          ************ *****  *
```

cmr7

```
AT5G35360 AGGTAATACCCGAACCCGAACCGATACCCACGGGTACACTTACATAGGACCCATCAGTGTAAAATCCTTATACTCGTACCCGCCCTGAATCGTTTTTTTTTGTTCGGTTCCGGTTCGGGTATGCGAG
AT5G35340 ----AATAACCGAACCCGAACCGAT------TTG------------------------CTCAAATC-------------------------------TTGGTTCGGTT-CGGTTCGG---------
              **** ****************        *                        * *****                                ** ******** ********
```

cmr75

```
AT5G08740 TTCCTGTAACAACAATACGACGTTGTTTTG-TCG--TAACGATGAAACTAAATTCCAAATTTCAAAGTTTCGTTGTTATTACC----------AAACGACGTCGTTTTGTTGTTAGAGAAA
AT4G05520 -----GA---ATCAAAACGACGACGTTTCTATCGATTATGGCT-------------------------CATTTTGTCATTATAG---------AAACGATGTCGTTTTGCT--------AA
AT5G37820 ---TTCT---ATCAAAACGACGTCGTTTT-AT-GAGTTCCGTTTACAAAAAAATC--TAAAGCAACGTCGTTTTGATTTAACAGAAAATTCTAAAACGACGTCGTTTTGAT------TTAA
                  * *** ******  ****   * *  *   * *                            ***   * *            ****** ********* *        **
```

cmr84

```
AT3G45850  AAACAGGGGCGGATCTATGTTGAGACTAGGTAGGGCACGCGCCCTATACTAATTGTTAAAGTAGTTTTAATTGATCACTTCTTATTGATATCTCTTGGCTCATTTGGCATAGAGTGCCCC
AT5G34853  -----------------------------------TG------TAA-----------------ATAAT-----------------------TAG--------------TGGTGCCCG
                                               *       *                  ****                      * *               ******
(cmr84 cont)
AT3G45850  AGGTTAATATTGTAAATTTTAATATGTGCCCCATACAAATTTTGTTTCTAGATCCGCCACTGGGT---------
AT5G34853  AGTAGAATA-------------------------------ACG-TTCTAGATCCGCCACTGATTTTTTTTTTT
           **   ****                               * ****************  *
```

cmr165

```
AT3G30220  -ATTATCCAGATCAGGATTGGG-----CCT--TGCATGGTATT--TTTGTTTTCCTGAGAGGATTGGAC--ATC-----ATCACA-------GATCGGATCTCGGATCCTGGATCC-T
AT1G32230  CATAATCCGGATCCGTATCCGGATCCGCCGGATTC--GTAAATGCTT--TACCCGTATCCGGATCCGTCGCACCCGGTTATTCGGTTTTCTGGATCGGATCACGGATCC-GGATTCGA
            ** **** **** * **  **     **   * *  *  * *  **  *   * *    ****  * *  * *     **           ********* ******* **** *
```

cmr195

```
AT1G19900 ------------------------------------------------------------ATAGTCACTTTTCACTGTTTGATTCGTACACTTGGATT-GTACATCACTTTTCTTATA
AT2G12480 ------------------------------------------------------------------TTCACTGTT-GATTCGTACA-----------------CTTTTAGTGTA
AT1G11680 AATTAGTCACTTTTTCAAACTTAATTAGAAAAGTAGTCACTTTACTATTCATAATTTGCAAAATAGTCATTTTTCACTGTTTGATTTGTAAATTTAGATT-GTACATCACTTTTAGTGTA
AT4G15280 ------------------------------------------------------------------TTCACTGTTTGATTCGTATACTTGGATT-GTACATCATTTTTAGTGTA
AT2G39795 ----------------------------AAAATAGTCACTTTACTATTCATAATTTGCAAAATAGTCACTTTTCACTGTTTGATTCGTACACTTGGATT-GTACATCACTTTTAGTGTA
AT2G18540 ------------------------------------------------------------------TTATTTTTACTGTTTTATTTGTACACTTAAATTAGTACAACACTTTTAATGTC
                                                                                ** ******  *** *** *                  ****  * *
(cmr195 cont)
AT1G19900 CAAATGTA---AAGTGTACGCTGTGTATTTTT-ATATATATTTGTACATTAAAAGTGGTGTACA-ATCTAAGTGTACGTATCAAACAGTGAAAAGTATCAAA------------------------------------------------------------
AT2G12480 TAGATGTATGTAAATATACACAGCGTACACTT-----ATATTTGTACACTAAAAATATTGTACAGATTCAAGTGTACGAATCAAATAGTAAA-------------------------------------------------------------------------
AT1G11680 CAAATATATATAAAAATACACAGTGTACACTT----TACATTTGTACACTAAAAGTGATGTACA-ATCCAAGTGTACGAATCAAACAGTGAAAAGTGATTGTTTTGCAAATTATGAATAGTAAAGTGACTACTTTCCTAATTAAGTTTGAAAAAGTGACTATTT
AT4G15280 CAAATGTA---AAGTGTACGCTGTGTATTTTTTATATACATATGTACACTAAAAGTGGTGTACT-ATCGAAGTGTACGAATCAAACAGTAAA-------------------------------------------------------------------------
AT2G39795 CAAATATA---AAATGTACGCTGTGTATTTTT-ATATATATTTGTACACTAAAAGTGATGTACA-ATCCAAGTGTACAAATCAAACAGTGAAAA----------------------------AGTGACTATTTTGCAAATTATGAATAGTAAAGTGACTACTT
AT2G18540 CAGATATATGT---------------A------ATCTA-------ACACTAAAAATGTTATACAGATCTAAGTGTATCAATCAAACAGCGAAAAGTA------------------------------------------------------------
           * ** **                 *         *       *** ***** *  * ***  **  *******   ****** **  **
```

cmr201

```
AT5G06700 -TAATCGATCTCTCTAATAGAGTTGGTGAAGTACGTTTGAAAGAAAATAAA----------GGC------CGACGTTTAGAGAGTCGAAAT
AT2G19180 TTCA--GATCTCTCTAATTGAGTGA------------TTAAAGAGAAACATTTTTGTTTCTGGGTTTGTTCGTAGAATAGAGA--TCTAGC
           * *  ************ ****             * ***** **  *          **       **  *  ******    *
```

cmr214

```
AT3G55735 CACAGCAACTAGAGAAAGGATCCAAAGGGATCGCATTGATCCTAATTAAGCTGATTTATTCCCCAATAATTGTTTTTTTTTCCTTCTCAATCGAAAGATGGAAGAAAAACAAATTCCAAACATTTTGCTTACTTTTCCGGATCATGCGATCTCTTTGGATTCATTCTTTGGTAGCTTCC
AT2G39890 CATAGCAACTAGAGGAAGGATCCAAAGGGATCGCATTGATCCTAATTAAGGTGAATT-CTCCCCATATTTTCTTTATAATTGGCAAATAAATC--------------ACAAAAA----------TTTGCTTGGTTTT--GGATCATGCTATCTCTTTGGATTCATCCTTCGGTAGCTTCC
          ** *********** ********************************** *** **  ******    ** *** *  **  *   * ****             * * ***         *******  ****  ********* *************** *** **********
```

cand 1

```
AT5G58470 -TAAAGCTCAGGAGGGATAGCGCCATGGCTCACCAGTGCTGTATGTTTTGTATATGCGTACATGTATATCTGTTGGCGCTATCCATCCTGAGTTCCA-
AT2G38330 GTAAAGCTCAGGAGGGATAGCGCCATG---------------ATGATC-----A-CATTCGTTATCTATTTTTTGGCGCTATCCATCCTGAGTTTCAT
           *************************               *** *      *    *   * * *** * ********************** **
```

cand 2

```
AT1G20380 -------TTGACAAA-GATTTTCTTACAGTCATCTTTGGCATTCTGTCCACCTCCT---TCTATACATATATGCA-TGTGTATATATATATGCG-----TTTCGTGTGAAAGAAGGAGGTGGGTATACTGCCAATAGAGATCTGTTAGGGCTTCTTCGTAAA--------
AT1G76140 TGCGGGTTTGACAAAAGAGTCTCTTACAGAGATCTTTGGCATTCTGTCCACCTCCTCTCTCTATATTTATGTGTAATAAGTGTACGTATCTACGGTGTGTTTCGT-----AAGAGGAGGTGGGCATACTGCCAATAGAGATCTGTTAGGGTTTCTTTGTAAAACCCCTCA
                 ******** ** * ********  ************************   ******  *** ** * *  ** **  *** * **     ******     *  ********** ************************** ***** *****
```

cand 3 (anti-MIR398)

```
AT5G14550  AAAGAGCTCAGCAGGGGTGACCTGAGAACACATGAAAACGAGAGCTGTAACAATG-ACGTAGCGTCATTAC-AGCCGTTGATTACTCGTGTGTTCTCATATCAACCCTGTCGAGATCCAC
AT5G14570  AAAGAGCTCAGCAGGGGTGACCTGAGAACACATG-AAACGAGAGCTGTAACAATG-ACGTAGCGTCGTTAT-AGCCGTTGATTGCTCGTGTGTTCTCATATCAACCCTGTCGAGATCCAC
AT2G03450  ----AGATTCAAAGGGGTGACCTGAGAACACA---AAACTC-AATGGTTTCAATGGAAATT-TGAAGAAACCATAGGAT-A-------TGTGTTCACATGCCACTCCTTTGAAATT----
               ** *    *******************   ****   *   **  ***** *  *   *     *  *   * * *       ******* ***  **  *** *  *  *
```