# Epigenetics aspects of floral homeotic genes in relation to sexual dimorphism in the dioecious plant *Mercurialis annua*

Janardan Khadka[1], Narendra Singh Yadav[1†], Micha Guy[1], Gideon Grafi[1*] and Avi Golan-Goldhirsh[1*]

[1]French Associates Institute for Agriculture and Biotechnology of Drylands, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben Gurion 84990, Israel. [†]Present address: Department of Biological Sciences, University of Lethbridge, AB T1K 3M4, Canada.

[*]To whom correspondence should be addressed
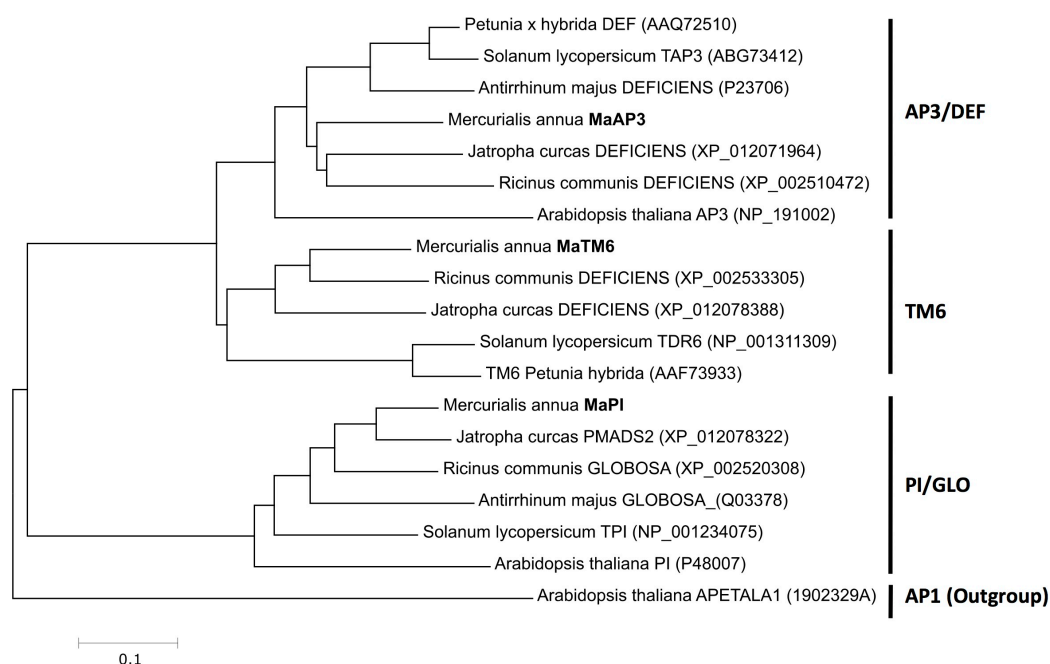
## Supplementary Figures



**Fig. S1**. Phylogenetic analysis of class B genes from *M. annua*, *A. thaliana* and various taxonomic groups. The complete translated sequences of *MaPI* (211 amino acids), *MaAP3* (224 amino acids), and *MaTM6* (225 amino acids) were used. The *AP1* gene of *A. thaliana* was used as an outgroup. The phylogenetic tree was constructed by neighbor joining (NJ) method with 500 bootstrap replications using MEGA6.0 package. The results confirm the placement of *MaPI* in the PI/GLO clade; *MaAP3* in the euAP3/DEF clade and *MaTM6* gene in the TM6 clade.
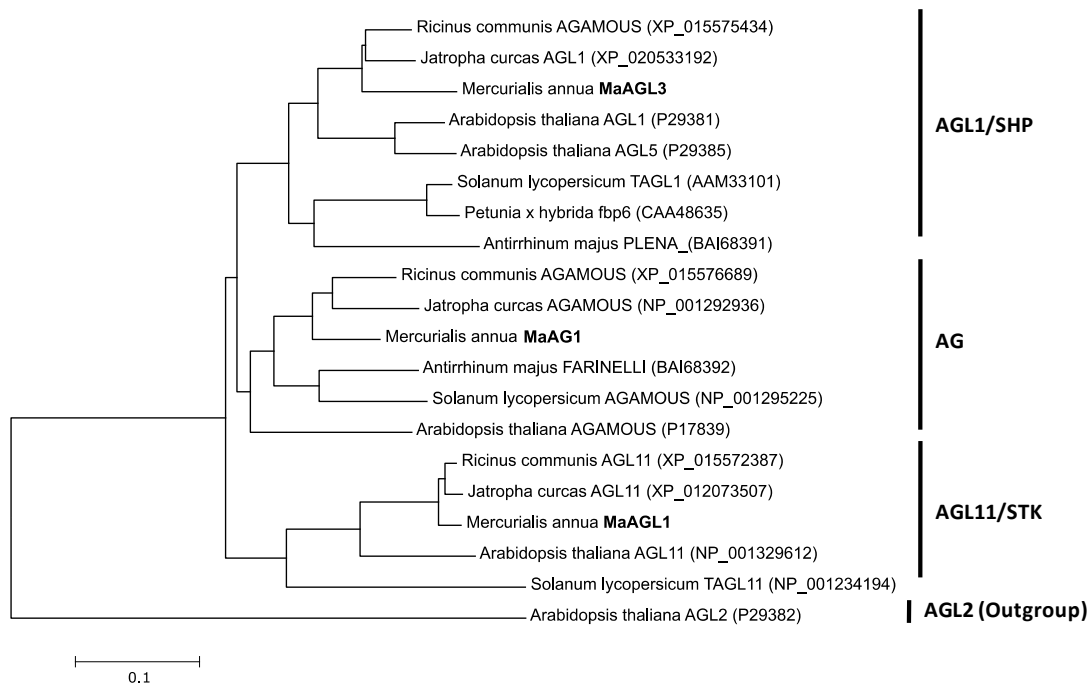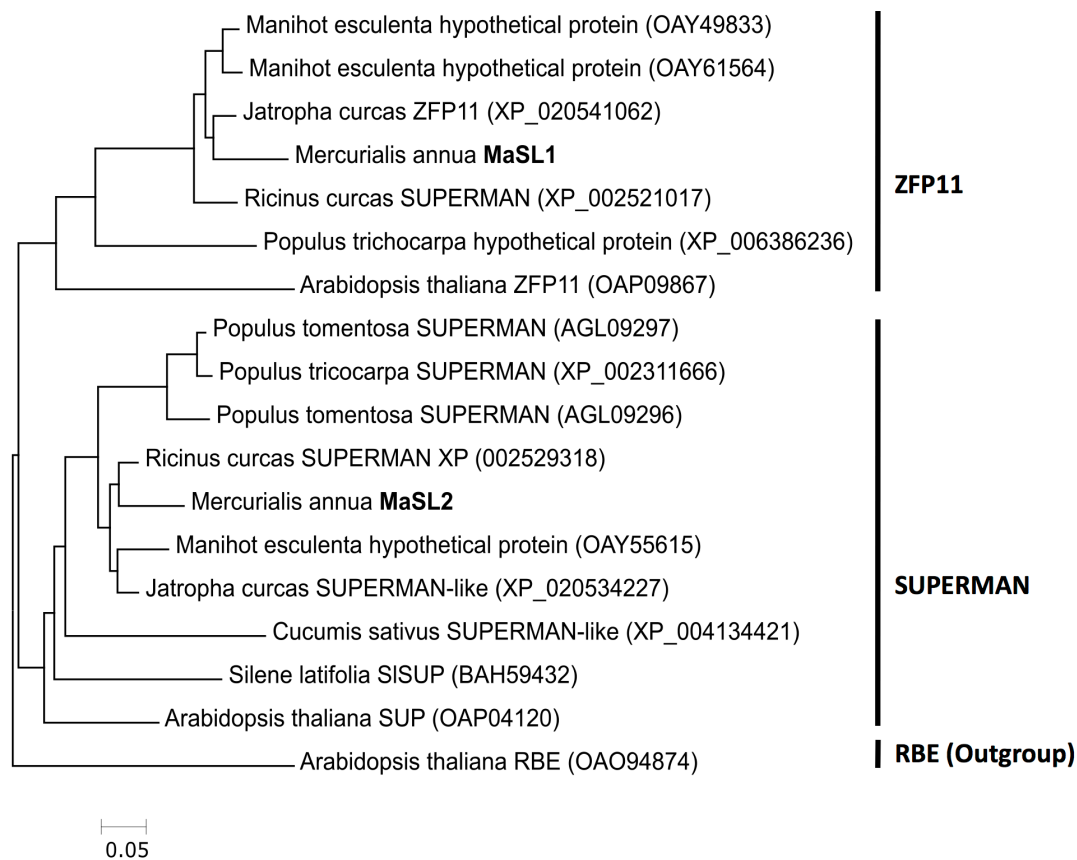
**Fig. S2**. Phylogenetic analysis of *AG*-like genes from *M. annua*, *A. thaliana* and various taxonomic groups. The complete translated sequences of *MaAG1* (241 amino acids), *MaAGL1* (222 amino acids) and *MaAGL3* (255 amino acids) were used. The *Arabidopsis AGL2* was used as an outgroup. The phylogenetic tree was constructed by a neighbor-joining method with 500 bootstrap replications using MEGA6.0 package. The dendrogram placed the *MaAGL1* in the STK clade (class D), *MaAG1* in the AG clade (class C) and *MaAGL3* gene in the SHP clade (class D).

**S3**. Phylogenetic analysis of *SUPERMAN-LIKE* genes from *M. annua*, *A. thaliana* and other various taxonomic groups. The conserved region between translated sequences of *MaSL1* (168 amino acids) and *MaSL2* (182 amino acids) were used because N-terminal regions of these sequences showed very low similarity, while C2H2 domain and EAR-motif of *MaSL1* and *MaSL2* were highly conserved. The *AtRBE* gene, a C2H2 zinc-finger protein that acts as spatial sepal-petal boundaries regulator was used as an outgroup. The phylogenetic tree was constructed by neighbor joining (NJ) method with 500 bootstrap replications in the MEGA6.0 package. It was found that *MaSL1* is an ortholog of *AtZFP11,* while *MaSL2* is an ortholog of *AtSUP*.

**Supplementary table**

**Table S1**: List of primers used in this study

| Gene | Primer | Sequence (5'--> 3') | Used for |
|---|---|---|---|
| Agamous-like | MADSF | AARATHSARATHAARMGNATHGARAA | Gene sequence isolation |
| | MADSR | ARNSMRACHKBDGCRTCRCANARNAC | |
| | AGLF | GTNCANTTYTGYAARMGNMGNAAYGG | |
| | AGLR | GCYAAYAARTAYCTNTCDATNGA | |
| Pistillata | PILF | ATGGGNMGNGGNAARATHGARATHAA | |
| | PIRc | ARRTTTGGYTGDATNGGYTGNAC | |
| | PIF2 | ATHGAYMGNGTNAARAARGARAAYGA | |
| | PIR2 | TCTTCTCCNTTCARRTGCCKNAG | |
| Apetala3 | AP3F2 | GATGCWAAGGTTTSNATHATHATG | |
| | EAP5F | CCAAGGTTTCCATTATCATGTTYTC | |
| | EAP3R | CTCAAGCAAAGGGTARGTKGTRAG | |
| TM6 | PAP5F | AACTCACWGTTCTYTGTGATGC | |
| | PAP3R | TCATGCAAGGCGRAGYTYATGGG | |
| | PAP5R | CATAATCYCCTTCATTSWCCACTA | |
| | AP3_TM6_F | GGTCGTGGAAAGATTGAGAT | |
| | AP3_TM6_R | TAGAGGTTAGAAGCCCCATTTG | |
| Superman-like | SM_F1 | AGACAATTCAACTCTGCTCAAG | |
| | SM_F2 | GCTCTTGGTGGTCACATGAA | |
| | SM_R | CCAAGTCGTAGCTCCAAATCCAA | |
| | SM_S | GGAGACGATTATCTTGGTGGGTTTTC | |
| | SM_A | CTAAGAGTAWCCCARTCGAAGTTCC | |
| Actin | ACTF | TGGTTCCACTATGTTCCCTG | |
| | ACTR | GATCCTCCAATCCAGACACT | |
| | MaAct_F | ATCTCACTGATGCTTTGATGA | |
| | MaACT_R | ACTTTCTTTCAGGTGGTGCGA | |
| MaAGL1 | 3RF1 | GATAGAGAACACCACAAATCGTCAG | 3' RACE |
| | 3RF2 | TGCAAGCGGAGAAGTGGC | |
| | RACE1 | CCAGTGAGCAGAGTGACGAGGACTCGAG CTCAAGC(T)$_{17}$ | |
| | RACE2 | CCAGTGAGCAGAGTGACG | |
| | RACE3 | GAGGACTCGAGCTCAAGC | |
| MaAGL1 | AGL1_RT | (P)TATCAGGCAGGAAC | 5' RACE |
| | AGL1_S2 | CGGATGGGGAGCAATGGAG | |
| | AGL1_A2 | ACACGCCTTCTTGTATCTATC | |
| | AGL1R5 | TACTCAATTTCAGCCAGCAACAA | |
| MaAG1 | AGL2_RT | (P)TGAAGGGGCATCTG | |
| | AGL2_S2 | CCAACTAATCACTACCCACATC | |
| | AG2R2 | GTATTTGATGTGTCCGAGGATGC | |

| | | | |
|---|---|---|---|
| MaAGL3 | AGL23_A2 | TCAGCATCACAAAGAACAGAC | 5' RACE |
| | AGL3_RT | (P)AGGTTTACAGGGAAG | |
| | AG3R2 | TTTGTGGAATCAGAACAGGCTT | |
| MaPI | PI_RT | (P)ACCCTTTGTTGATG | |
| | PI_A1 | TGTCACGAACACCAGCAAG | |
| | PI_S1 | GAAAAGATGTTGGAGGATG | |
| | PI_S2 | GGCAATGGAAGAGAATGTGAGA | |
| | PI_A2 | AGTGCTTCTTCTAAACCCATCAG | |
| MaAP3 | AP3_RT | (P)AGATGGTTAGGGTTC | |
| | AP3_A1 | TACGCAGATTCCTATTCACA | |
| | AP3_S2 | GGAGGGGATTATGAGTCTGTG | |
| | AP3_A2 | TCTGCTTAGTTGTGATGGAA | |
| MaSL1 | SL_RT | (P)CACAACTTCTTTAGG | |
| | SL_A1 | GAGAAGTAATGTGAGGTGGA | |
| | SL_S1 | GAGGTTTTGCAGAGTAATGAAG | |
| | SL_A2 | GGTTAGGGTTAGGGTTTGGATG | |
| | SL_S2 | GCTTGGAGTTGGAAATAGGGTT | |
| MaAGL1 | AGL1W1 | GCGACTGGAGAAGACGATTAGG | Promoter isolation |
| | AGL1W2 | ATTTGTGGTGTTCTCTATCCT | |
| | AGL1W3 | CATCCCTGGGCGTTGTACT | |
| MaAG1 | AG2W1 | CTAGAGAAGACGATGAGAGC | |
| | AG2W2 | TCCTCCCTCGAATCGCCTTG | |
| | AG2W3 | CTGATGAGTAGAAAGAAAGAGATGC | |
| MaAGL3 | AG3W1 | ACGAGTGGAGAAGATAATGAGAGA | |
| | AG3W2 | GCCTTCCGATTCTTGATTAG | |
| | AG3W3 | TGGTAAGAAAACAAACAAGG | |
| MaAP3 | AP3W1 | GTGATGGAAGGGCTCATGTATT | |
| | AP3W2 | TAGCCTTCTTGAACAGACCAT | |
| | AP3W3 | AGTCACCTGTCGATTTGTAGA | |
| MaPI | PIW1 | GTAGAAGGGCTGCAATACTCAT | |
| | PIW2 | GCATCGCATAAAACTGTAATCTCC | |
| | PIW3 | AGTGACTTGCCTGTTGCTTG | |
| MaSL1 | SL_A1 | GAGAAGTAATGTGAGGTGGA | |
| | SL_A2 | GGTTAGGGTTAGGGTTTGGATG | |
| MaSL2 | SMWF1 | GTGAGCTTGATTTAGTGGCAAGA | |
| | SMWF2 | AGTAAGTAGTGACTAAAACATA | |
| | SMWR1 | GGTTAAGAATGGAGCACTGAC | |
| MaPI | MPF | ACAAGAAATGGCAATGGAAGAG | RT and qPCR |
| | MPR | TGTAGTCCCTCACCCTTTGTTG | |
| MaAP3 | MaAP3_F | CAAGAGAAGAAGATCCGCAGTATG | |
| | MaAP3_R | GTTAGGGTTCATCCGCAGAG | |
| | | | |

| | | | |
|---|---|---|---|
| MaAGL1 | AG1LF1 | CCAAGATAGCAGAAGTGGAAAGG | |
| | AG1LR1 | CTCCCCATCCGTTTAGCCAAG | |
| | AG1SLF | TGCGGCAACAGATTCAGATG | |
| | AG1SR | TCAAGCTCAATCTCCTTTTTGG | |
| MaAG1 | AG2F1 | CAAGATTGCTGAGAATGAAAGG | |
| | AG2R1 | CTTGATGTGGGTAGTGATTAGTTGG | |
| MaAGL3 | MaAG3F | TGAAAATGAGAGAGCACAACAGC | |
| | MaAG3R | ACGGAACAGATTCGTACACAGAAC | |
| MaActin | ACTNF | TGGTTCCACTATGTTCCCTG | |
| | ACTNR | GATCCTCCAATCCAGACACT | |
| MaTM6 | AP3TM6_F | GCAACAACGACGAAGAAGAT | |
| | AP3TM6_R | GCGGCGTGATGAATGGTAG | |
| MaSL1 | SL_RT_F | ATGTTTAGGAAGTGAAGGAGACAAG | |
| | SL_A1 | GAGAAGTAATGTGAGGTGGA | |
| MaSL2 | SM_RT_F | AGCACTTGGGGGTCACAT | |
| | SM_RT_R | ATCTCAGCTTTCTTCAACAACT | |
| MaAGL1-Promoter | AGL1CPF | GTGAATTGCTTTTGTATTTCTGG | Chop-PCR/ MNase-PCR |
| | AGL1CPR | ACCTGGGCGTTGTACTCC | |
| MaAP3 promoter | BS_AP3C_S | GAAGGATGAAGTGAATATTTGTTGG | |
| | AP3W1 | GTGATGGAAGGGCTCATGTATT | |
| MaPI Promoter | PIPF1 | GAAAAAAAGCAGTTGCCTTGCC | |
| | PIW1 | GTAGAAGGGCTGCAATACTCAT | |
| MaSL1 promoter | SL1_S | ACAACACTCATCACTCTAACTCTC | |
| | SL1_A1 | CTTGAATTCTTGTCTCCTTCACTTCC | |
| MaSL2 promoter | SMP_BSF1 | AAAACAAGGGGGTGAGCAAATA | |
| | SMP_BSR2 | CCACCTACATAATCTTCTCCAC | |
| MaActin | MaACT_F | ATCTCACTGATGCTTTGATGA | |
| | MaACT_R | ACTTTCTTTCAGGTGGTGCGA | |
| MaAP3 genebody | AP3GBF1 | ATTCAAAAGGAGGAGGAGTT | Bisulfite-seq |
| | AP3GBR1 | CATTTAAACAATCGCCCATCC | |
| | AP3GBF2 | ATTAGGGTTATGGCGAGAGG | |
| | AP3GBR2 | CCTATTCACATCTTTCAATTTCC | |
| MaAP3 promoter | BS_AP3C_S | GAAGGATGAAGTGAATATTTGTTGG | |
| | BS_AP3_A1 | CTTGCCTCTCGCCATAACCCTAATTT | |
| | BS_AP3_A2 | TTTTGAAAACTCCTCCTCCTTT | |
| MaSL1 genebody | S1GBF1 | AATGGAGAGAATATGTTTAGGAAGTG | |
| | S1GBR1 | TCTAAACTCCAAATCCACAACTTCT | |
| | S1GBF2 | GAAGGAGATAAGAATTTAAGTGG | |
| | S1GBR2 | TAAACCCTATTTCCAACTCCAA | |
| MaSL1 promoter | SLCT_S1 | GTAAACGTTAAAAAATGATGGAAG | |
| | SL1_A1 | CTTGAATTCTTGTCTCCTTCACTTCC | |
| | SL1_A2 | CATGTTCTCTCCATTTTTTTCTTTTA | |