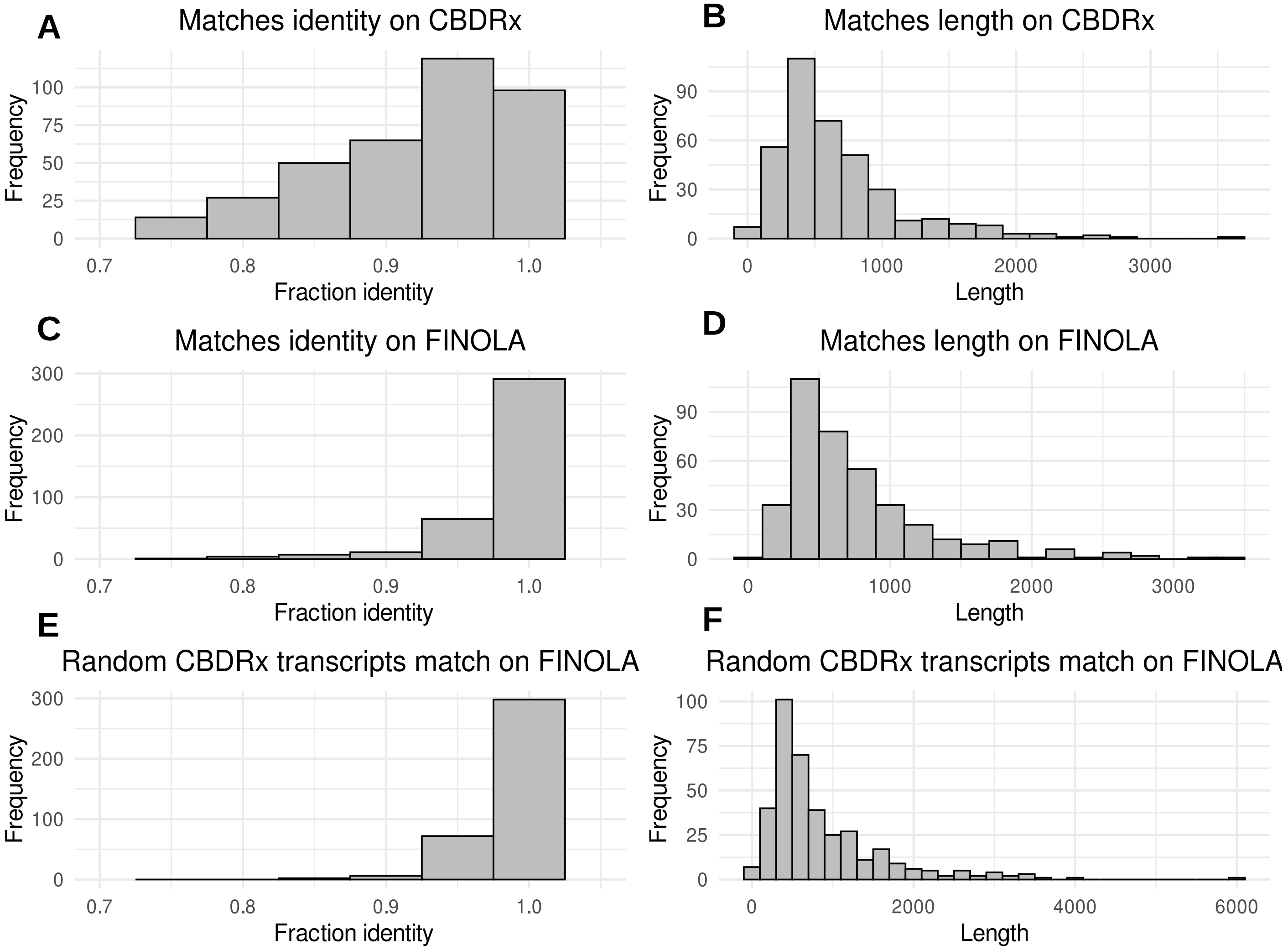
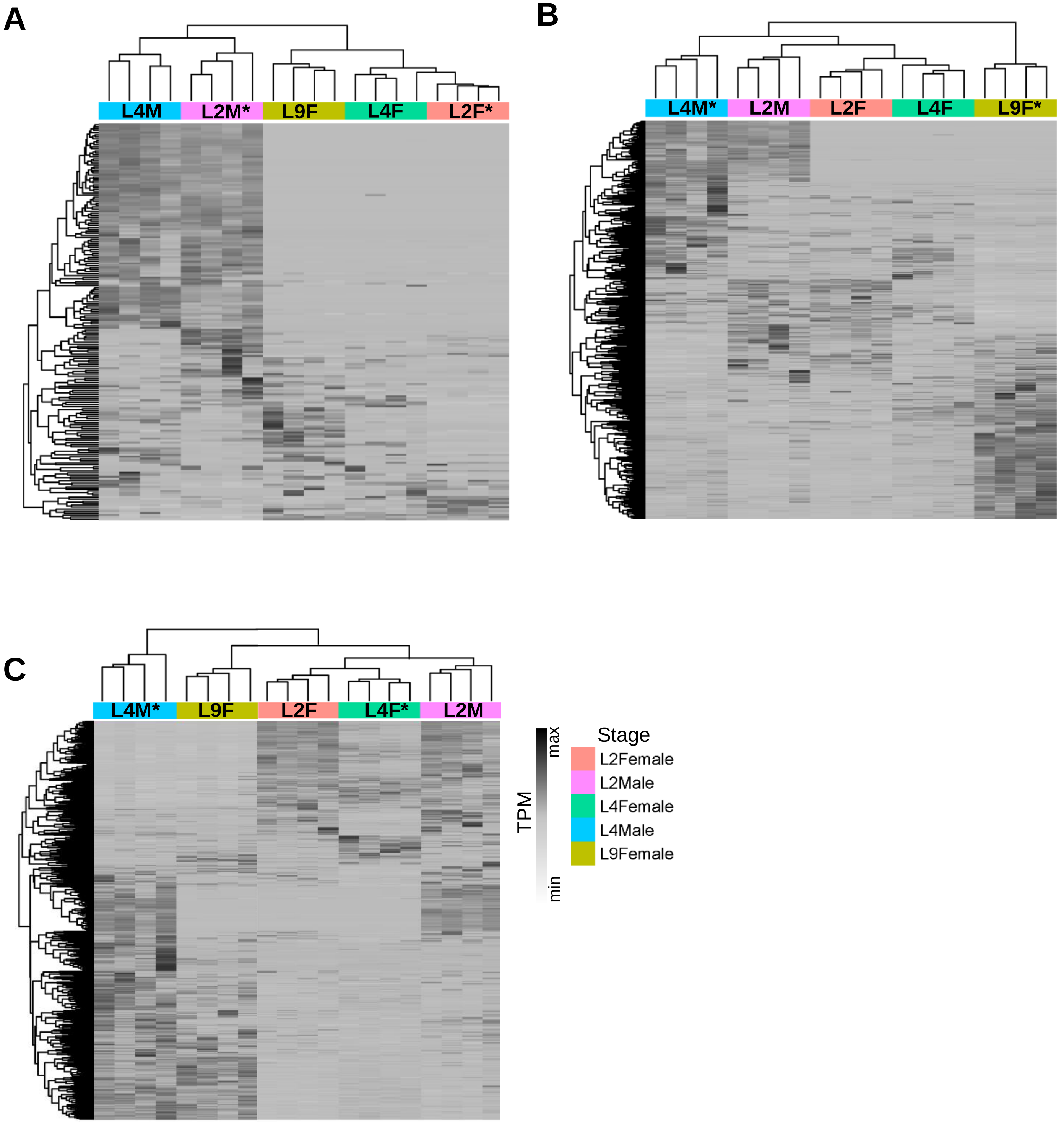
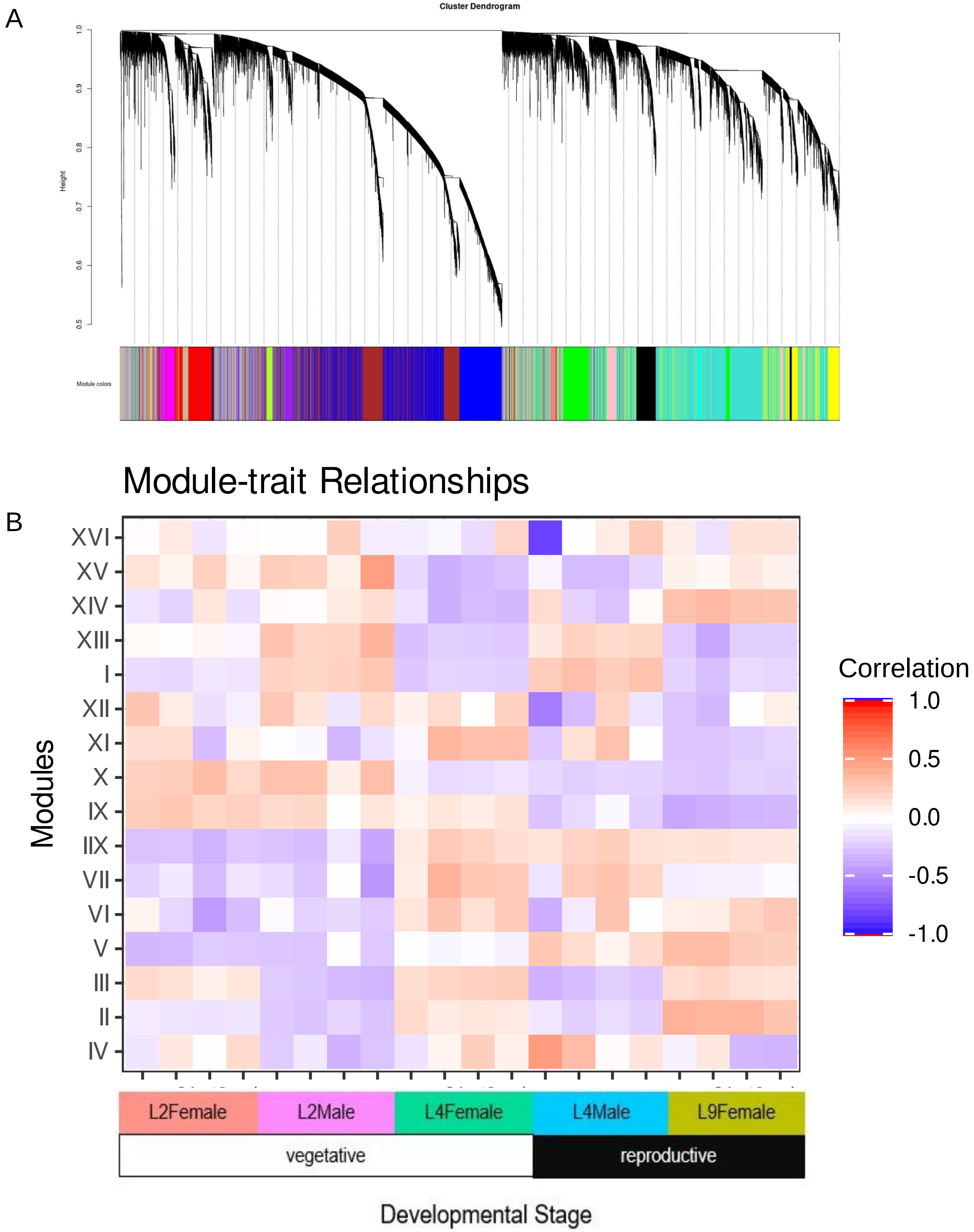


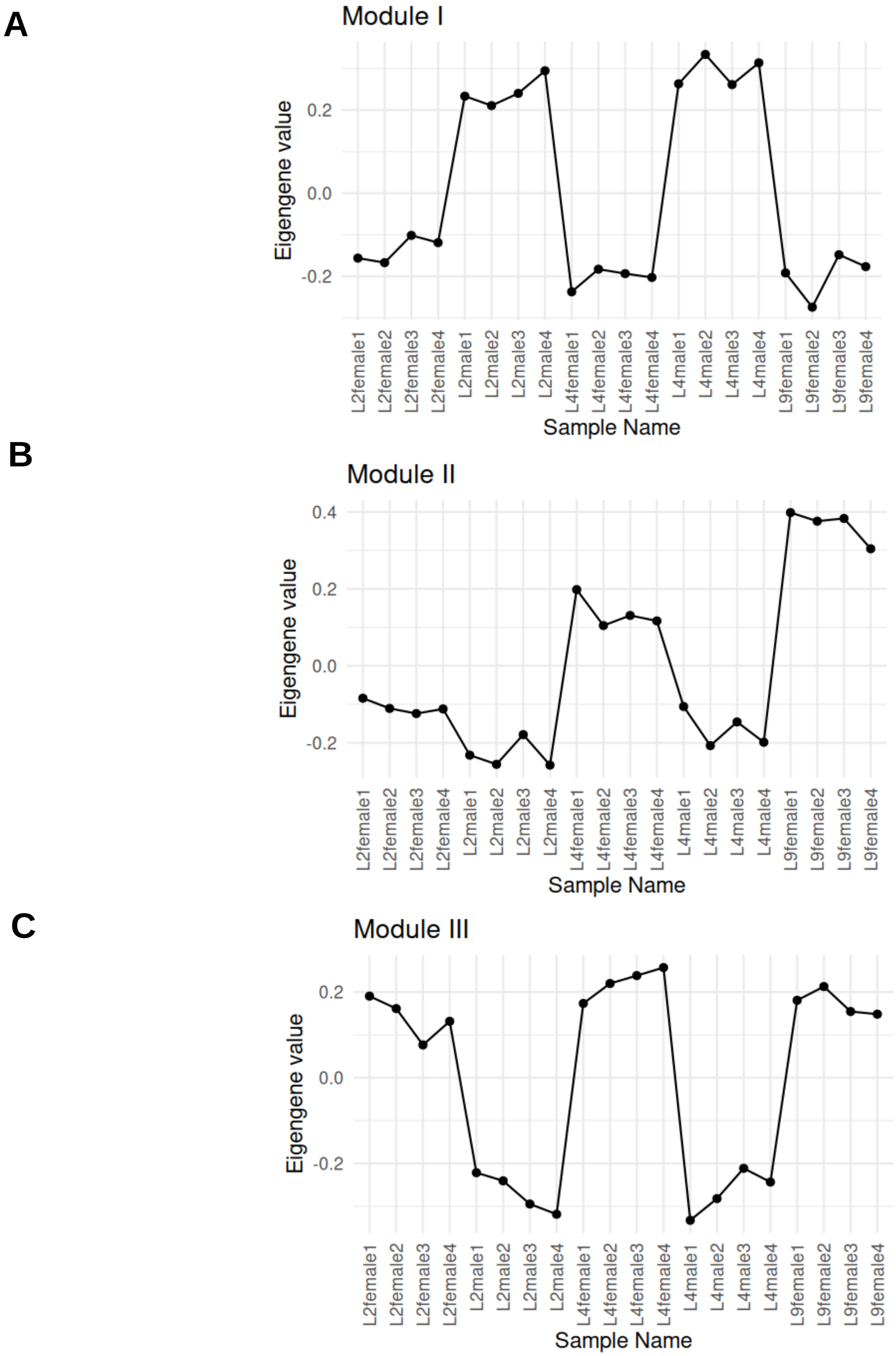
Supplementary Fig. S1. Sequence comparison of ChrY transcripts to 'CBDRx' and 'FINOLA' genomes. Putative ChrY transcripts were compared against the female *C. sativa* reference genome 'CBDRx' (GCF_900626175.2) (Grassa et al., 2021) using BLAST, with sequence identities as fractions (A) and the match lengths (B) in base pairs are displayed as bar charts. Subsequently, putative ChrY transcripts were compared against the male genome of the *C. sativa* hemp cultivar 'FINOLA' (ASM341772v2) (Laverty et al., 2019), and sequence identity as fractions (C) and match lengths in base pairs (D) are presented as bar charts. To establish a baseline for similarity between 'CBDRx' and 'FINOLA' genomes, a set of randomised 'CBDRx' transcripts was BLASTed against the 'FINOLA' genome, with identity as fractions (E) and match lengths in base pairs (F) being displayed as bar charts.



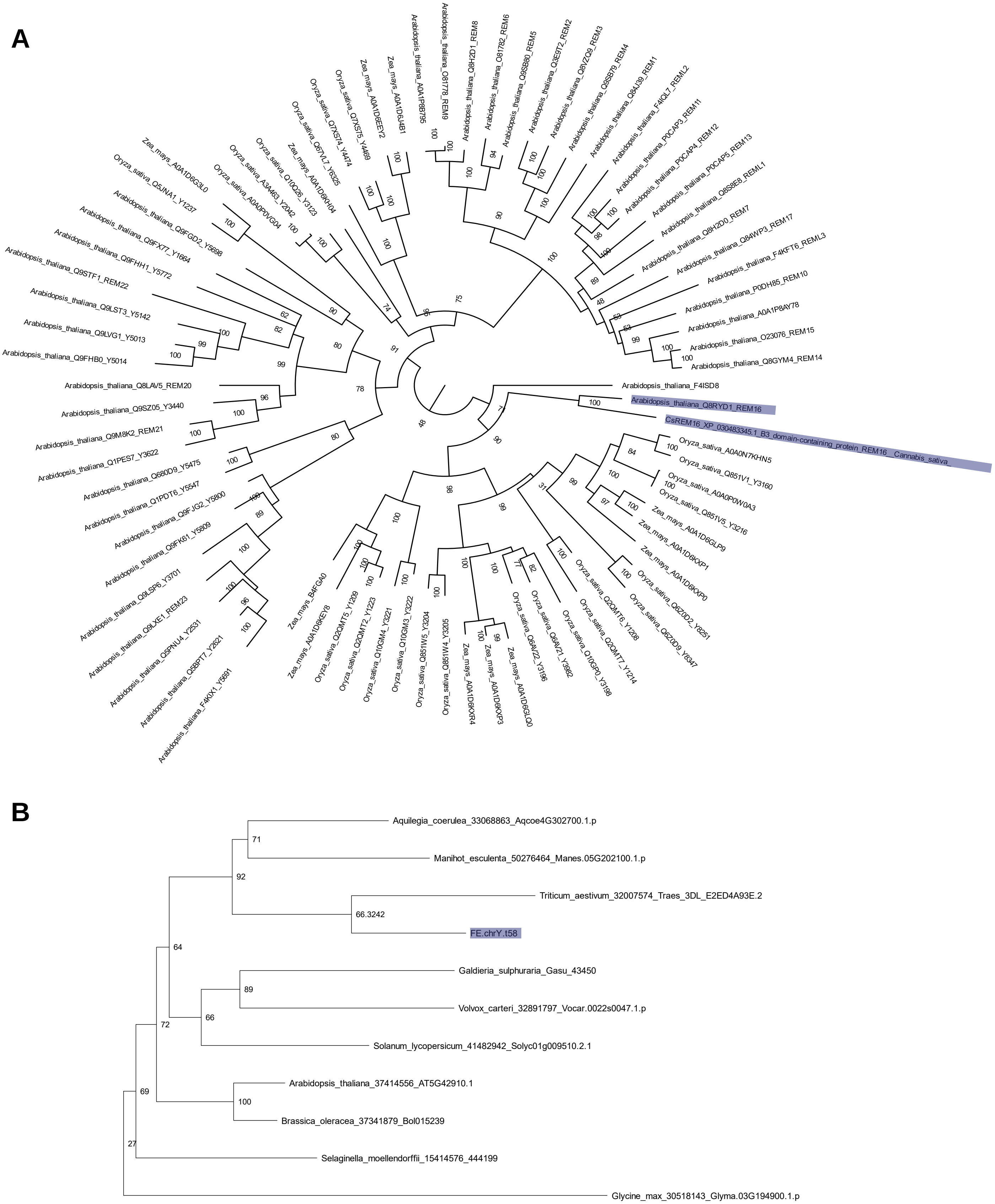
Supplementary Fig. S2. Hierarchical clustering of differentially expressed genes (DEGs) in hemp. DEGs identified in a comparison of L2M vs L2F (A), L4M vs L9F (B), L4M vs L4F (C), with bi-directional clustering based on the complete method performed according to developmental stage and expression level. Asterisks (*) indicate the stages used to determine DEGs. Heatmaps present the relative expression levels of each transcript (row) in each plant sample (column). The expression values were normalised using TPM values. The grey scale, based on a relative color scheme, represents high expression (dark black) and low expression (light grey) of the transcripts across samples. The lightest and darkest colours are converted using the minimum and maximum values in each row.



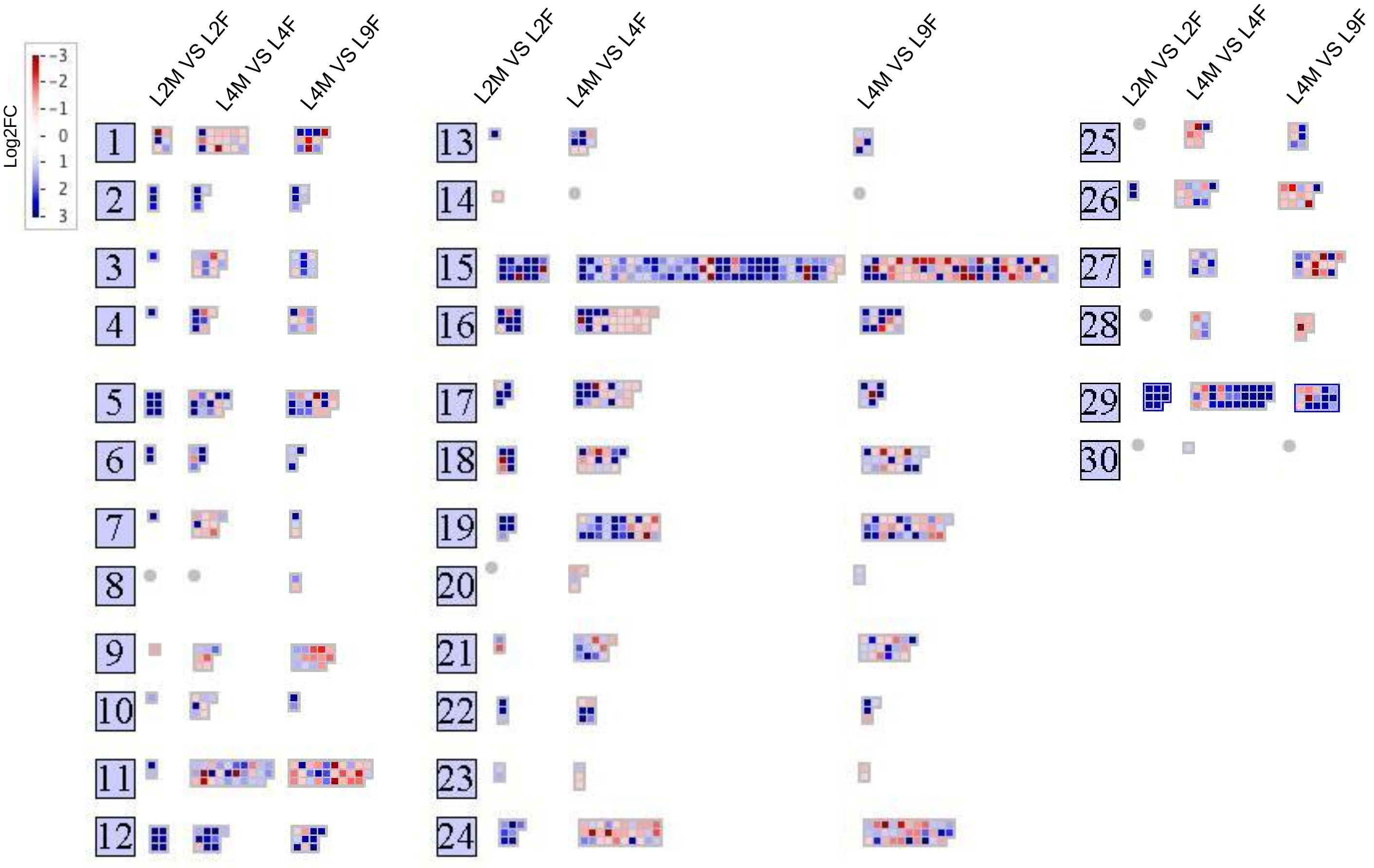
Supplementary Fig. S3. Modules of co-expressed genes identified by Weighted Gene Co-expression Network Analysis (WGCNA). A cluster dendrogram of all modules identified with WGCNA (A). The colour row underneath the dendrogram shows the determined module assignment. A heatmap of average expression per sample in all modules (B). The blue colour represents a negative correlation, while the red represents a positive correlation. A complete list of all genes can be found in Supplementary Table S5.



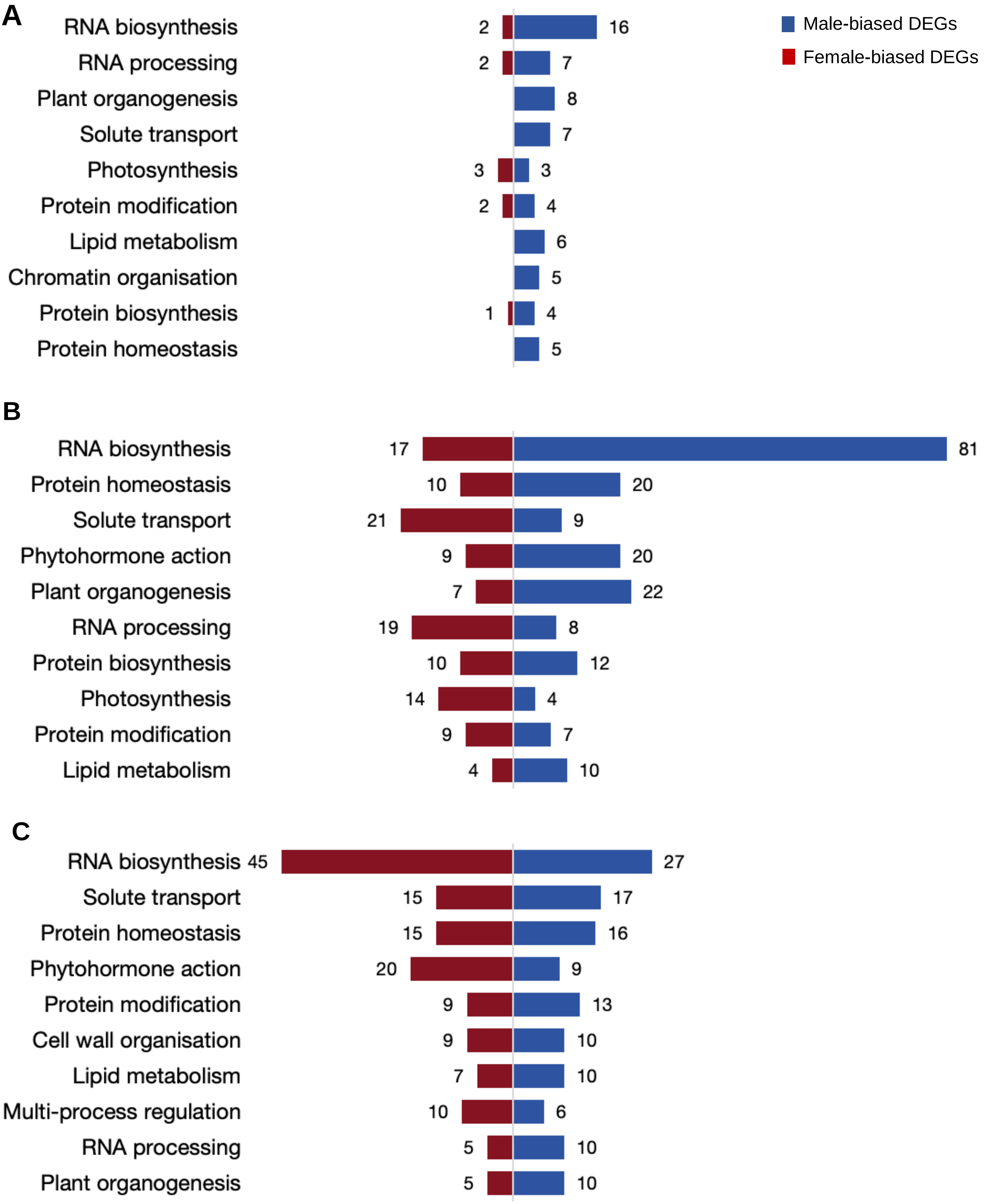
Supplementary Fig. S4. The eigengene values for the sex-related modules. Plot of eigengene values summarising the activity of the three modules with clear sex correlation, obtained by taking the first principal component of expression level of genes in that module. Module I (A) shows clear activity in male samples. While modules II (B) and III (C) show correlation with late stage and all stages female samples respectively.



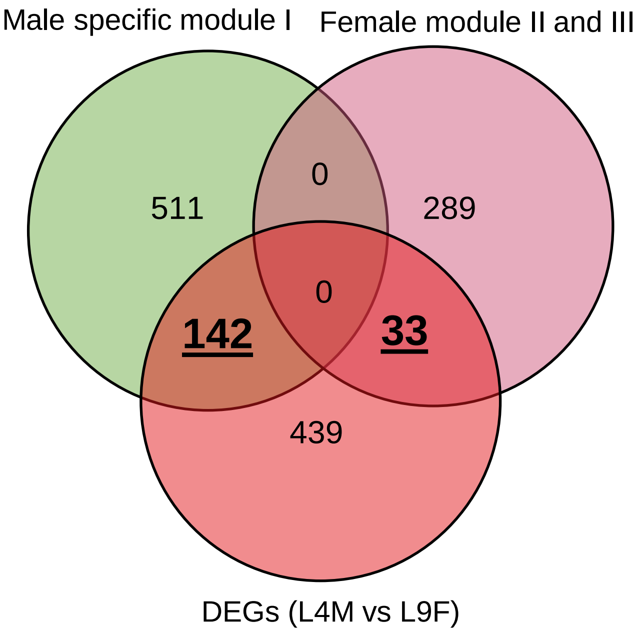
Supplementary Fig. S5. SHOOT.bio homology association. Phylogeny of *CsREM16* (LOC115699937), shows that *CsREM16* putative homolog in *A. thaliana* is Q8RYD1 (At4g33280) named *AtREM16* (A). Phylogeny of FE.chrY.t58, shows that its putative homolog in *A. thaliana* is AT5G42910, a gene in the same bZIP group A as *FD* (B).



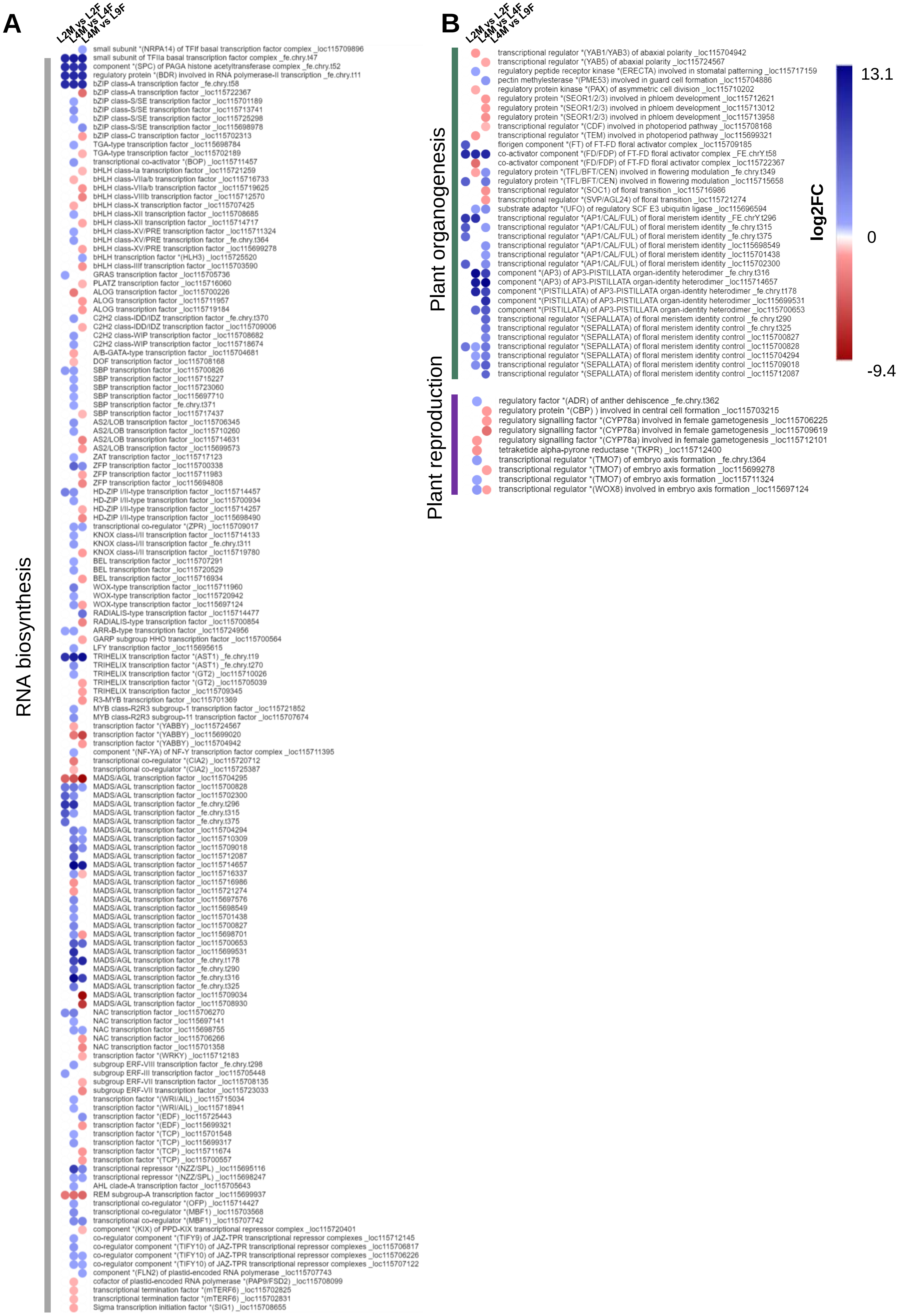
Supplementary Fig. S6. Summary of MapMan metabolic pathways from the DEGs. Differentially expressed genes identified through comparison of L2M vs L2F, L4M vs L4F and L4M vs L9F were analysed using MapMan and assigned to metabolic pathway bins: 1- photosynthesis, 2- cellular respiration, 3- carbohydrate metabolism, 4- amino acid metabolism, 5- lipid metabolism, 6- nucleotide metabolism, 7- coenzyme metabolism, 8- polyamine metabolism, 9- secondary metabolism, 10- redox homeostasis, 11- phytohormone action, 12- chromatin organization, 13- cell cycle organization, 14- DNA damage response, 15- RNA biosynthesis, 16- RNA processing, 17- protein biosynthesis, 18- protein modification, 19- protein homeostasis, 20- cytoskeleton organization, 21- cell wall organization, 22- vesicle trafficking, 23- protein translocation, 24- solute transport, 25- nutrient uptake, 26- external stimuli response, 27- multi-process regulation, 28- plant reproduction, 29- plant organogenesis, 30- clade-specific metabolism. Expression values are log2fold changes (log2FC) with colour scales of red (female-biased), blue (male-biased), and white indicating log2fold change values close to 0.



Supplementary Fig. S7. The top 10 MapMan metabolic pathways from the DEGs in hemp. To better identify the functional pathway response to sex determination, the top ten metabolic pathway bins, which contain the largest number of genes, are presented in butterfly plots for DEGs from L2M vs L2F (A), L4M vs L4F (B), and L4M vs L9F (C), respectively.



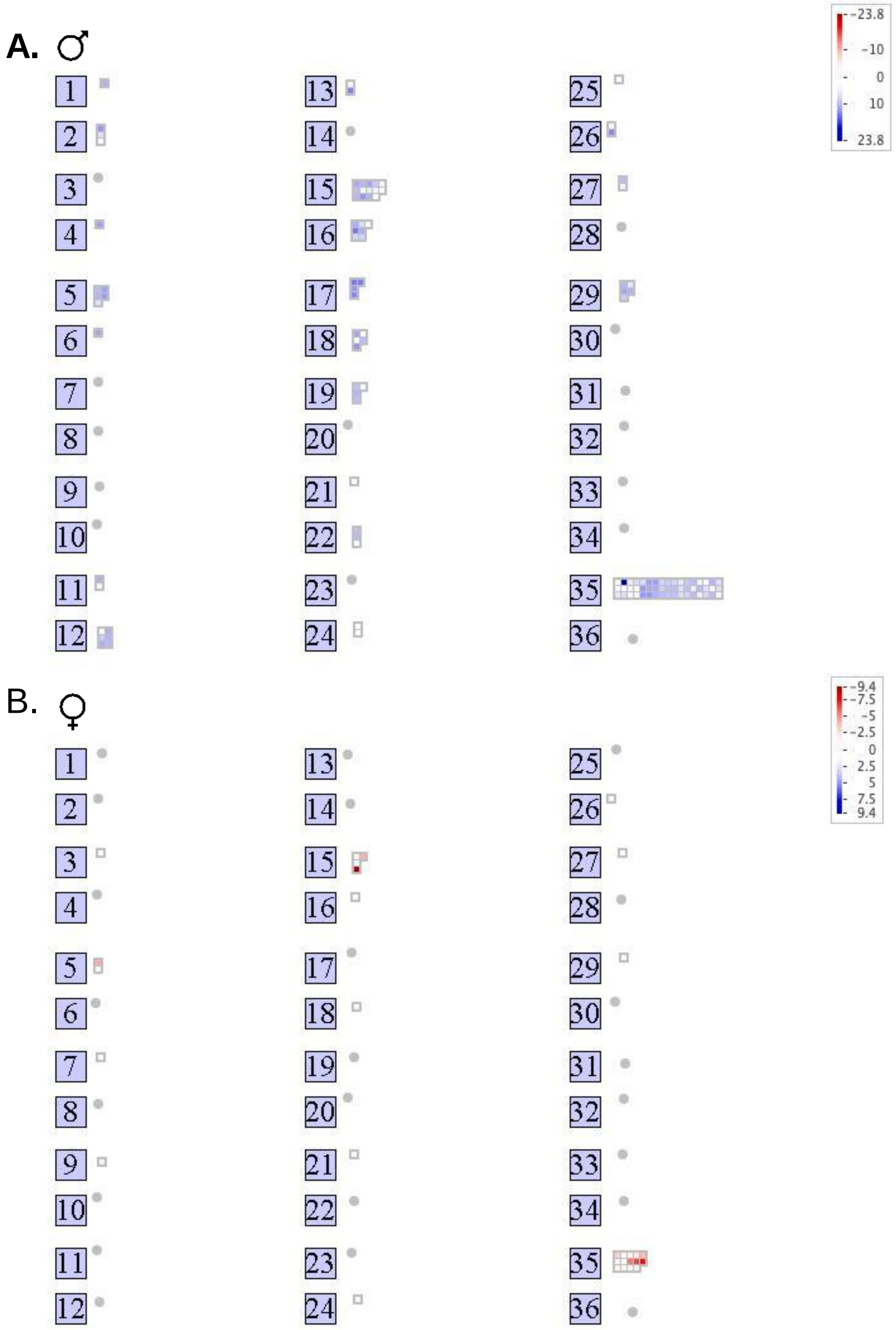
Supplementary Fig. S8. Identification of putative candidate genes for flower development in hemp. The Venn diagram illustrates the number of overlapping genes between DEGs from the L4M vs L9F comparison and the male-specific or female-specific module from WGCNA results.


A
L2M vs L2F
L4M vs L4F
L4M vs L9F
RNA biosynthesis
small subunit *(NRPA14) of TFIf basal transcription factor complex _loc115709896
small subunit of TFIIa basal transcription factor complex _fe.chry.t47
component *(SPC) of PAGA histone acetyltransferase complex _fe.chry.t52
regulatory protein *(BDR) involved in RNA polymerase-II transcription _fe.chry.t11
bZIP class-A transcription factor _fe.chry.t58
bZIP class-A transcription factor _loc115722367
bZIP class-S/SE transcription factor _loc115701189
bZIP class-S/SE transcription factor _loc115713741
bZIP class-S/SE transcription factor _loc115725298
bZIP class-S/SE transcription factor _loc115698978
bZIP class-C transcription factor _loc115702313
TGA-type transcription factor _loc115698784
TGA-type transcription factor _loc115702189
transcriptional co-activator *(BOP) _loc115711457
bHLH class-Ia transcription factor _loc115721259
bHLH class-VIIa/b transcription factor _loc115716733
bHLH class-VIIa/b transcription factor _loc115719625
bHLH class-VIIIb transcription factor _loc115712570
bHLH class-X transcription factor _loc115707425
bHLH class-XII transcription factor _loc115708685
bHLH class-XII transcription factor _loc115714717
bHLH class-XV/PRE transcription factor _loc115711324
bHLH class-XV/PRE transcription factor _fe.chry.t364
bHLH class-XV/PRE transcription factor _loc115699278
bHLH transcription factor *(HLH3) _loc115725520
bHLH class-IIIf transcription factor _loc115703590
GRAS transcription factor _loc115705736
PLATZ transcription factor _loc115716060
ALOG transcription factor _loc115700226
ALOG transcription factor _loc115711957
ALOG transcription factor _loc115719184
C2H2 class-IDD/IDZ transcription factor _fe.chry.t370
C2H2 class-IDD/IDZ transcription factor _loc115709006
C2H2 class-WIP transcription factor _loc115708682
C2H2 class-WIP transcription factor _loc115718674
A/B-GATA-type transcription factor _loc115704681
DOF transcription factor _loc115708168
SBP transcription factor _loc115700826
SBP transcription factor _loc115715227
SBP transcription factor _loc115723060
SBP transcription factor _loc115697710
SBP transcription factor _fe.chry.t371
SBP transcription factor _loc115717437
AS2/LOB transcription factor _loc115706345
AS2/LOB transcription factor _loc115710260
AS2/LOB transcription factor _loc115714631
AS2/LOB transcription factor _loc115699573
ZAT transcription factor _loc115717123
ZFP transcription factor _loc115700338
ZFP transcription factor _loc115711983
ZFP transcription factor _loc115694808
HD-ZIP I/II-type transcription factor _loc115714457
HD-ZIP I/II-type transcription factor _loc115700934
HD-ZIP I/II-type transcription factor _loc115714257
HD-ZIP I/II-type transcription factor _loc115698490
transcriptional co-regulator *(ZPR) _loc115709017
KNOX class-I/II transcription factor _loc115714133
KNOX class-I/II transcription factor _fe.chry.t311
KNOX class-I/II transcription factor _loc115719780
BEL transcription factor _loc115707291
BEL transcription factor _loc115720529
BEL transcription factor _loc115716934
WOX-type transcription factor _loc115711960
WOX-type transcription factor _loc115720942
WOX-type transcription factor _loc115697124
RADIALIS-type transcription factor _loc115714477
RADIALIS-type transcription factor _loc115700854
ARR-B-type transcription factor _loc115724956
GARP subgroup HHO transcription factor _loc115700564
LFY transcription factor _loc115695615
TRIHELIX transcription factor *(AST1) _fe.chry.t19
TRIHELIX transcription factor *(AST1) _fe.chry.t270
TRIHELIX transcription factor *(GT2) _loc115710026
TRIHELIX transcription factor *(GT2) _loc115705039
TRIHELIX transcription factor _loc115709345
R3-MYB transcription factor _loc115701369
MYB class-R2R3 subgroup-1 transcription factor _loc115721852
MYB class-R2R3 subgroup-11 transcription factor _loc115707674
transcription factor *(YABBY) _loc115724567
transcription factor *(YABBY) _loc115699020
transcription factor *(YABBY) _loc115704942
component *(NF-YA) of NF-Y transcription factor complex _loc115711395
transcriptional co-regulator *(CIA2) _loc115720712
transcriptional co-regulator *(CIA2) _loc115725387
MADS/AGL transcription factor _loc115704295
MADS/AGL transcription factor _loc115700828
MADS/AGL transcription factor _loc115702300
MADS/AGL transcription factor _fe.chry.t296
MADS/AGL transcription factor _fe.chry.t315
MADS/AGL transcription factor _fe.chry.t375
MADS/AGL transcription factor _loc115704294
MADS/AGL transcription factor _loc115710309
MADS/AGL transcription factor _loc115709018
MADS/AGL transcription factor _loc115712087
MADS/AGL transcription factor _loc115714657
MADS/AGL transcription factor _loc115716337
MADS/AGL transcription factor _loc115716986
MADS/AGL transcription factor _loc115721274
MADS/AGL transcription factor _loc115697576
MADS/AGL transcription factor _loc115698549
MADS/AGL transcription factor _loc115701438
MADS/AGL transcription factor _loc115700827
MADS/AGL transcription factor _loc115698701
MADS/AGL transcription factor _loc115700653
MADS/AGL transcription factor _loc115699531
MADS/AGL transcription factor _fe.chry.t178
MADS/AGL transcription factor _fe.chry.t290
MADS/AGL transcription factor _fe.chry.t316
MADS/AGL transcription factor _fe.chry.t325
MADS/AGL transcription factor _loc115709034
MADS/AGL transcription factor _loc115708930
NAC transcription factor _loc115706270
NAC transcription factor _loc115697141
NAC transcription factor _loc115698755
NAC transcription factor _loc115706266
NAC transcription factor _loc115701358
transcription factor *(WRKY) _loc115712183
subgroup ERF-VIII transcription factor _fe.chry.t298
subgroup ERF-III transcription factor _loc115705448
subgroup ERF-VII transcription factor _loc115708135
subgroup ERF-VII transcription factor _loc115723033
transcription factor *(WRI/AIL) _loc115715034
transcription factor *(WRI/AIL) _loc115718941
transcription factor *(EDF) _loc115725443
transcription factor *(EDF) _loc115699321
transcription factor *(TCP) _loc115701548
transcription factor *(TCP) _loc115699317
transcription factor *(TCP) _loc115711674
transcription factor *(TCP) _loc115700557
transcriptional repressor *(NZZ/SPL) _loc115695116
transcriptional repressor *(NZZ/SPL) _loc115698247
AHL clade-A transcription factor _loc115705643
REM subgroup-A transcription factor _loc115699937
transcriptional co-regulator *(OFP) _loc115714427
transcriptional co-regulator *(MBF1) _loc115703568
transcriptional co-regulator *(MBF1) _loc115707742
component *(KIX) of PPD-KIX transcriptional repressor complex _loc115720401
co-regulator component *(TIFY9) of JAZ-TPR transcriptional repressor complexes _loc115712145
co-regulator component *(TIFY10) of JAZ-TPR transcriptional repressor complexes _loc115706817
co-regulator component *(TIFY10) of JAZ-TPR transcriptional repressor complexes _loc115706226
co-regulator component *(TIFY10) of JAZ-TPR transcriptional repressor complexes _loc115707122
component *(FLN2) of plastid-encoded RNA polymerase _loc115707743
cofactor of plastid-encoded RNA polymerase *(PAP9/FSD2) _loc115708099
transcriptional termination factor *(mTERF6) _loc115702825
transcriptional termination factor *(mTERF6) _loc115702831
Sigma transcription initiation factor *(SIG1) _loc115708655
B
L2M vs L2F
L4M vs L4F
L4M vs L9F
Plant organogenesis
transcriptional regulator *(YAB1/YAB3) of abaxial polarity _loc115704942
transcriptional regulator *(YAB5) of abaxial polarity _loc115724567
regulatory peptide receptor kinase *(ERECTA) involved in stomatal patterning _loc115717159
pectin methylesterase *(PME53) involved in guard cell formation _loc115704886
regulatory protein kinase *(PAX) of asymmetric cell division _loc115710202
regulatory protein *(SEOR1/2/3) involved in phloem development _loc115712621
regulatory protein *(SEOR1/2/3) involved in phloem development _loc115713012
regulatory protein *(SEOR1/2/3) involved in phloem development _loc115713958
transcriptional regulator *(CDF) involved in photoperiod pathway _loc115708168
transcriptional regulator *(TEM) involved in photoperiod pathway _loc115699321
florigen component *(FT) of FT-FD floral activator complex _loc115709185
co-activator component *(FD/FDP) of FT-FD floral activator complex _FE.chrY.t58
co-activator component *(FD/FDP) of FT-FD floral activator complex _loc115722367
regulatory protein *(TFL/BFT/CEN) involved in flowering modulation _fe.chry.t349
regulatory protein *(TFL/BFT/CEN) involved in flowering modulation _loc115715658
transcriptional regulator *(SOC1) of floral transition _loc115716986
transcriptional regulator *(SVP/AGL24) of floral transition _loc115721274
substrate adaptor *(UFO) of regulatory SCF E3 ubiquitin ligase _loc115696594
transcriptional regulator *(AP1/CAL/FUL) of floral meristem identity _FE.chrY.t296
transcriptional regulator *(AP1/CAL/FUL) of floral meristem identity _fe.chry.t315
transcriptional regulator *(AP1/CAL/FUL) of floral meristem identity _fe.chry.t375
transcriptional regulator *(AP1/CAL/FUL) of floral meristem identity _loc115698549
transcriptional regulator *(AP1/CAL/FUL) of floral meristem identity _loc115701438
transcriptional regulator *(AP1/CAL/FUL) of floral meristem identity _loc115702300
component *(AP3) of AP3-PISTILLATA organ-identity heterodimer _fe.chry.t316
component *(AP3) of AP3-PISTILLATA organ-identity heterodimer _loc115714657
component *(PISTILLATA) of AP3-PISTILLATA organ-identity heterodimer _fe.chry.t178
component *(PISTILLATA) of AP3-PISTILLATA organ-identity heterodimer _loc115699531
component *(PISTILLATA) of AP3-PISTILLATA organ-identity heterodimer _loc115700653
transcriptional regulator *(SEPALLATA) of floral meristem identity control _fe.chry.t290
transcriptional regulator *(SEPALLATA) of floral meristem identity control _fe.chry.t325
transcriptional regulator *(SEPALLATA) of floral meristem identity control _loc115700827
transcriptional regulator *(SEPALLATA) of floral meristem identity control _loc115700828
transcriptional regulator *(SEPALLATA) of floral meristem identity control _loc115704294
transcriptional regulator *(SEPALLATA) of floral meristem identity control _loc115709018
transcriptional regulator *(SEPALLATA) of floral meristem identity control _loc115712087
Plant reproduction
regulatory factor *(ADR) of anther dehiscence _fe.chry.t362
regulatory protein *(CBP)) involved in central cell formation _loc115703215
regulatory signalling factor *(CYP78a) involved in female gametogenesis _loc115706225
regulatory signalling factor *(CYP78a) involved in female gametogenesis _loc115709619
regulatory signalling factor *(CYP78a) involved in female gametogenesis _loc115712101
tetraketide alpha-pyrone reductase *(TKPR) _loc115712400
transcriptional regulator *(TMO7) of embryo axis formation _fe.chry.t364
transcriptional regulator *(TMO7) of embryo axis formation _loc115699278
transcriptional regulator *(TMO7) of embryo axis formation _loc115711324
transcriptional regulator *(WOX8) involved in embryo axis formation _loc115697124
13.1
log2FC
0
-9.4

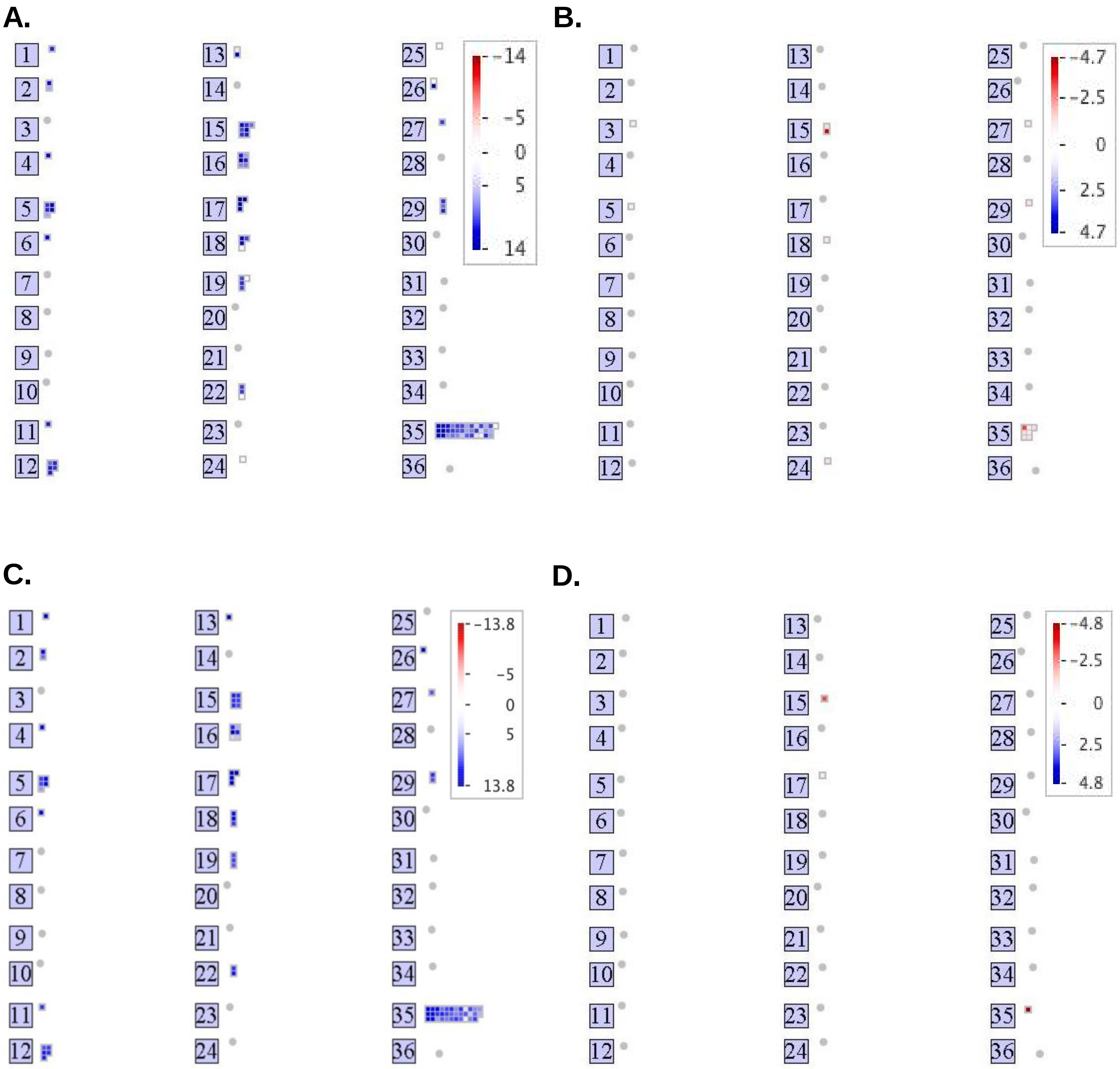
Supplementary Fig. S9. Differentially expressed genes potentially involved in sex determination and floral development. Differential expression between L2F vs L2F, L4M vs L4F and L4MF vs L9F is indicated for genes linked to RNA biosynthesis (A), or and plant organogenesis and plant reproduction (B) as identified through MapMan analysis. The relative expression of each gene (row) in each comparison (column) is shown. Expression values are log2fold changes (log2FC) with colour scales of red (female-biased) and blue (male-biased).


A
Module I
component *(NRPD2) of RNA polymerase IV complex _fe.chry.t126
component *(NRPE2) of RNA polymerase V complex _fe.chry.t126
small subunit of TFIIa basal transcription factor complex _fe.chry.t47
subunit beta of TFIIe basal transcription factor complex _loc115709141
component *(SCS1/SCS2) of HAT histone acetyltransferase module _loc115721375
component *(ADA1) of SAGA transcription co-activator complex _fe.chry.t40
component *(ADA1) of SAGA transcription co-activator complex _loc115715332
component *(SPC) of PAGA histone acetyltransferase complex _fe.chry.t52
component *(MED18) of head module of MEDIATOR transcription co-activator complex _loc115700285
regulatory protein *(BDR) involved in RNA polymerase-II transcription _fe.chry.t11
bZIP class-A transcription factor _fe.chry.t58
bZIP class-A transcription factor _loc115702933
bZIP class-S/SE transcription factor _loc115698978
bZIP class-S/SE transcription factor _loc115725298
bHLH class-IVa transcription factor _loc115712227
bHLH class-IIIa transcription factor _loc115718651
BSD-type transcription factor _loc115712016
C2H2 class-WIP transcription factor _loc115709101
C2H2 class-WIP transcription factor _loc115721385
DOF transcription factor _loc115697076
AS2/LOB transcription factor _loc115705047
AS2/LOB transcription factor _loc115705389
zf-HD-type transcription factor _loc115700466
HSF transcription factor _loc115716259
HSF transcription factor _loc115722192
TRIHELIX transcription factor *(AST1) _fe.chry.t19
MYB class-R2R3 subgroup-1 transcription factor _loc115710324
MYB class-R2R3 subgroup-13 transcription factor _loc115711659
MADS/AGL transcription factor _loc115700653
MADS/AGL transcription factor _loc115714657
NAC transcription factor _loc115709772
NAC transcription factor _loc115723252
VOZ transcription factor _loc115699585
subgroup ERF-VIII transcription factor _loc115723530
subgroup ERF-III transcription factor _loc115708230
transcription factor *(WRI/AIL) _loc115718941
transcriptional repressor *(NZZ/SPL) _loc115695116
transcriptional repressor *(NZZ/SPL) _loc115698247
AHL clade-A transcription factor _loc115704801
transcriptional co-regulator *(MBF1) _loc115707742
co-regulator component *(TIFY10) of JAZ-TPR transcriptional repressor complexes _loc115707122
regulatory protein *(LAP3) of exine formation _loc115724889
fatty acyl-CoA reductase *(MS2) _loc115699209
regulatory protein *(ASL19/20) of xylem differentiation _loc115705047
cleavage and polyadenylation specificity factor *(FY) _fe.chry.t69
florigen component *(FT) of FT-FD floral activator complex _loc115700781
co-activator component *(FD/FDP) of FT-FD floral activator complex _fe.chry.t58
regulatory protein *(TFL/BFT/CEN) involved in flowering modulation _loc115700781
component *(AP3) of AP3-PISTILLATA organ-identity heterodimer _loc115714657
component *(PISTILLATA) of AP3-PISTILLATA organ-identity heterodimer _loc115700653
RNA biosynthesis
Plant reproduction
Plant organogenesis
kWithin
1.2
129.2
B
Module II
bZIP class-C transcription factor _loc115702313
TGA-type transcription factor _loc115699157
bHLH class-X transcription factor _loc115702403
bHLH class-XII transcription factor _loc115712856
DBP phosphatase-domain transcription factor _loc115724355
ALOG transcription factor _loc115697612
C2H2 class-IDD/IDZ transcription factor _loc115709006
A/B-GATA-type transcription factor _loc115704484
ZFP transcription factor _loc115722966
HD-ZIP I/II-type transcription factor _loc115700801
zf-HD-type transcription factor _loc115700881
transcription factor *(ARID3/4/6) _loc115711694
MADS/AGL transcription factor _loc115704295
FRS/FRF transcription factor _loc115704634
transcription factor *(WRKY) _loc115700154
subgroup ERF-V transcription factor _loc115703125
transcription factor *(TCP) _loc115706624
REM subgroup-A transcription factor _loc115699925
transcriptional co-regulator *(TIFY) _loc115699142
component *(LUG/LUH) of transcriptional co-repressor complex _loc115702276
adapter component *(SEU/SLK) of transcriptional co-repressor complex _loc115710310
regulatory protein kinase *(PAX) of asymmetric cell division _loc115710202
C
Module III
component *(SSL2/XPB) of TFIIh basal transcription factor complex _loc115703047
component *(ADA1) of SAGA transcription co-activator complex _loc115700092
C-GATA-type transcription factor _loc115697696
C-GATA-type transcription factor _loc115697707
TRIHELIX transcription factor *(AST1) _loc115710018
transcription factor *(WRKY) _loc115706387
REM subgroup-A transcription factor _loc115699937
cleavage and polyadenylation specificity factor *(FY) _loc115724500

Supplementary Fig. S10. Sex-related genes identified by WGCNA with potential function in RNA biosynthesis, plant organogenesis, and plant reproduction. Genes included in module I (A), module II (B) and module III (C) were assigned to either RNA biosynthesis, plant organogenesis, and plant reproduction using MapMan. The colour indicates the kWithin value derived from the WGCNA. Light grey presents the lower value, and dark green presents the higher value which means a stronger correlation in the module.



Supplementary Fig. S11. Functional annotation of putative candidate genes for flower development. The MapMan functional annotation of the putative candidate genes for flower development. Genes that were significantly differentially expressed between L4M vs LF9, and highly correlated in WGCNA were considered putative candidates (see Figs. 6, 7). Candidate genes for flower development in male plants(A), candidate genes for flower development in female plants (B) were sorted into MapMan function bins. The colour bar represents the expression values as log2 fold changes (log2FC) of female-related genes (red) and male-related genes (blue).



Supplementary Fig. S12. Identification of putative candidate genes for early sex determination. Genes that were significantly differentially expressed between L2M vs L2F, highly correlated in WGCNA and located on sex chromosomes were considered putative candidates (refer to Fig. 7). Putative gene for male sex identity in the flowering stage (A). Putative gene for female sex identity in the flowering stage (B). Putative gene for male sex identity in the vegetative stage (C). Putative genes for female sex identity in the vegetative stage (D). The colour bar represents the expression values as log2 fold changes (log2FC) of female-related genes (red) and male-related genes (blue).