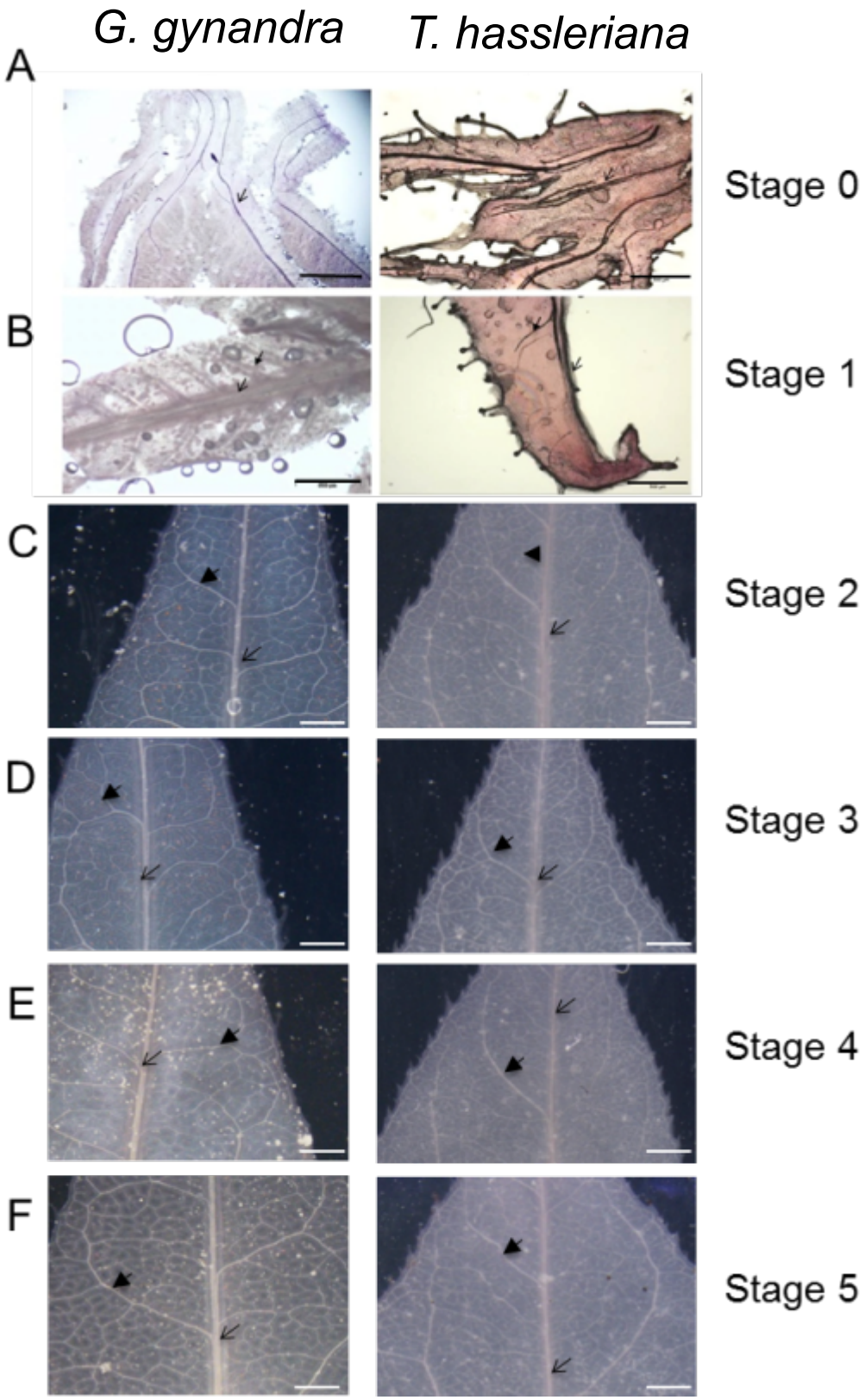
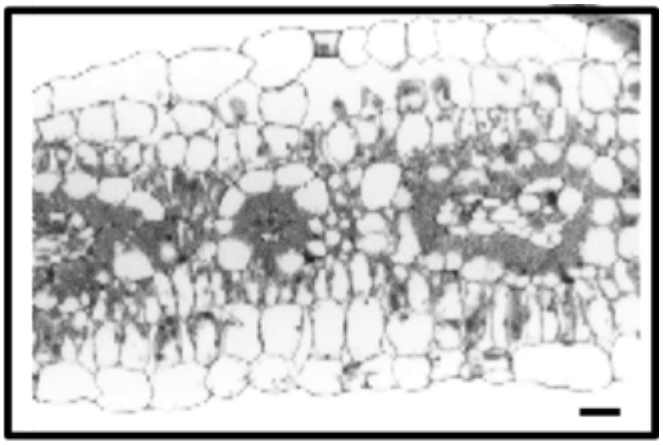
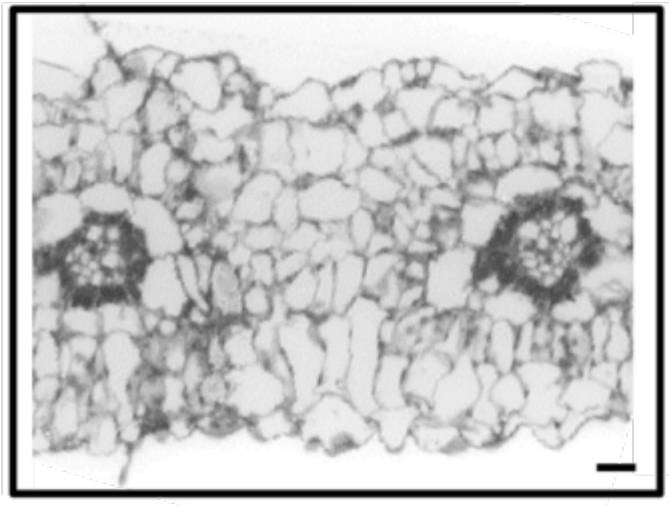


Supplemental Figure 1. Venation patterning during leaf development of *G. gynandra* and *T. hassleriana*.
(A-B) Cleared safranine stained leaves of stage 0 and 1 (n=3; scale bar 0.5 mm)
(C-F) Cleared leaves of stage 2, 3, 4 and 5 respectively (n=3; scale bar 1 mm)
Open arrows indicate the midvein (1°) and closed arrows the secondary vein (2°) localization

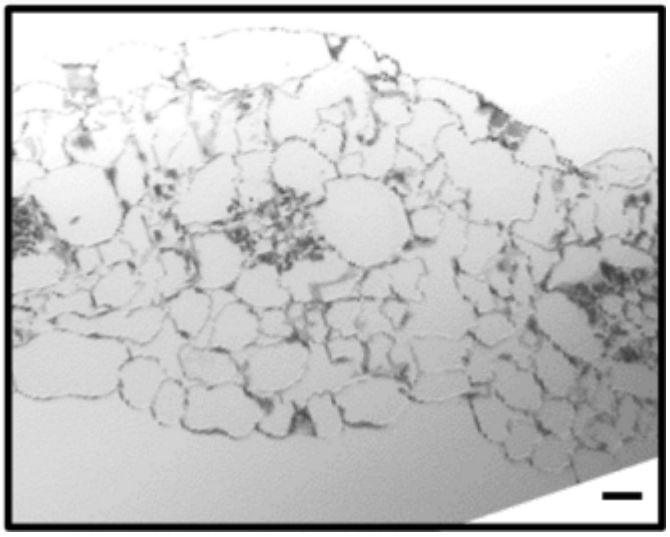
A



B



C



Supplemental Figure 2. *G. gynandra* cotyledon anatomy two, four and six days after germination (DAG). Semi-thin cross sections (3 µm) of *G. gynandra* cotyledons after two **(A)**; four **(B)**; six **(C)** DAG. Cross sections were stained with Toluidine Blue. (Scale bar 10 µm, n=3)

A

G. gynandra *T. hassleriana*

B

G. gynandra *T. hassleriana*

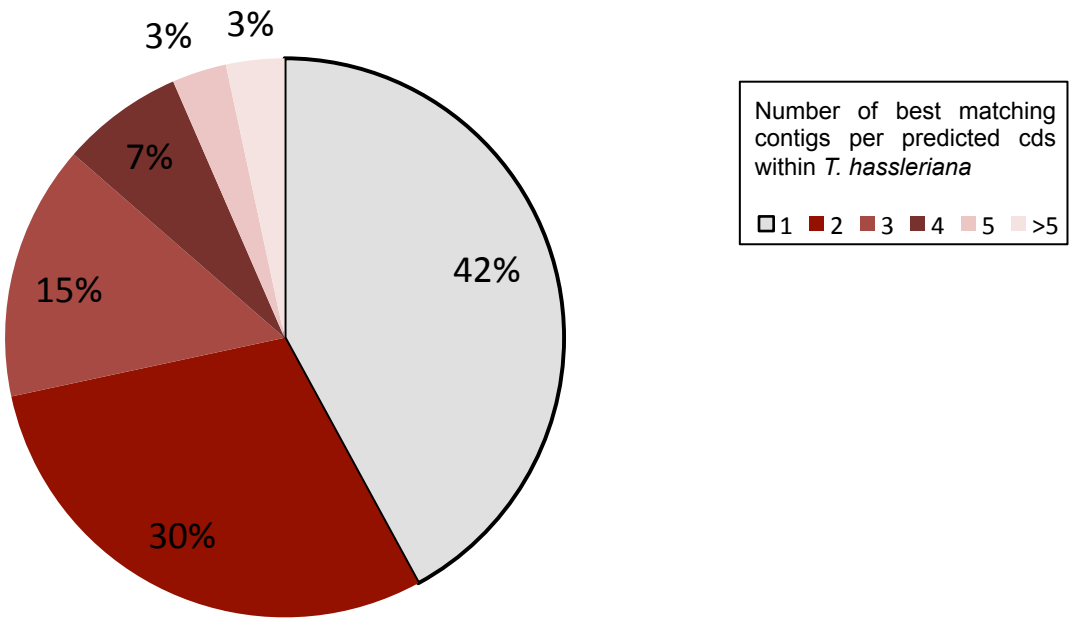
1

2

3

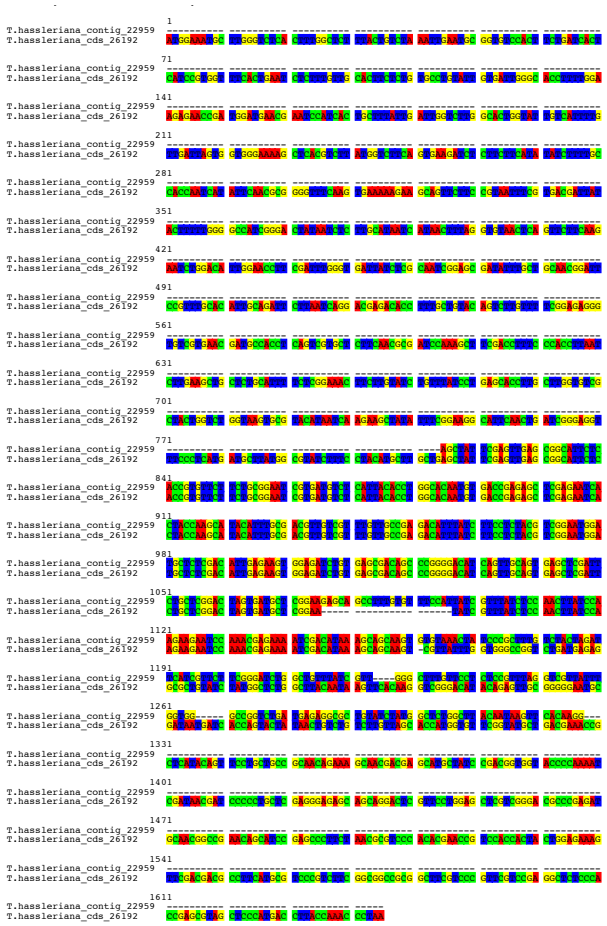
Supplemental Figure 3. Images of tissues harvested for RNA-seq in *G. gynandra* and *T. hassleriana*. (A) Photographic image of *G. gynandra* and *T. hassleriana* 8-week old plants, from which leaf gradient, stem and root system were harvested **(B) Seed coat development from harvested developmental seed gradient. (1)** young seed **(2)** semi-mature seed **(3)** mature seed. (Scale bar = 1cm)

A



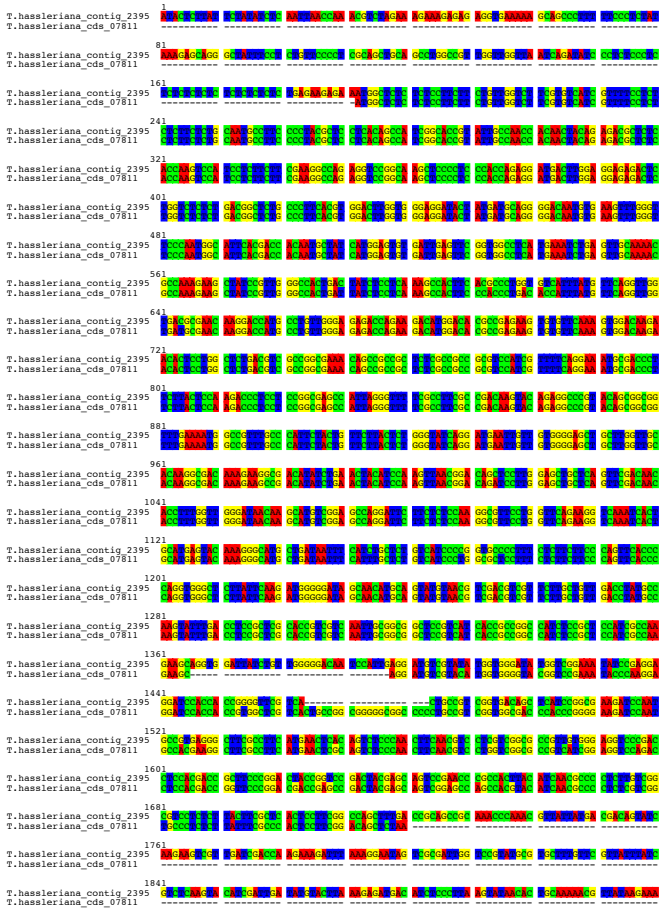
B

Fragmented contig aligned to *T. hassleriana* cds



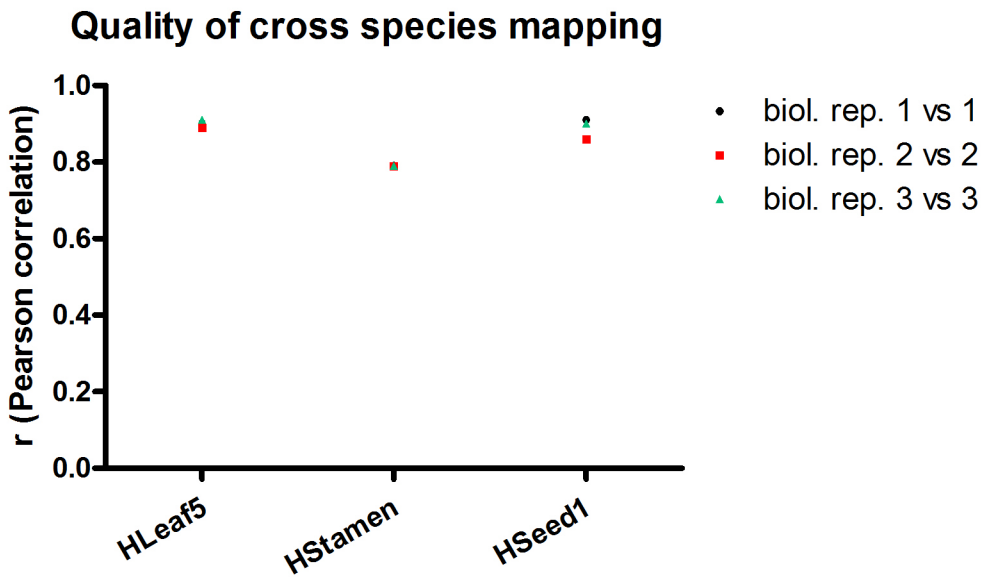
C

Fused/Hybrid contig aligned to *T. hassleriana* cds

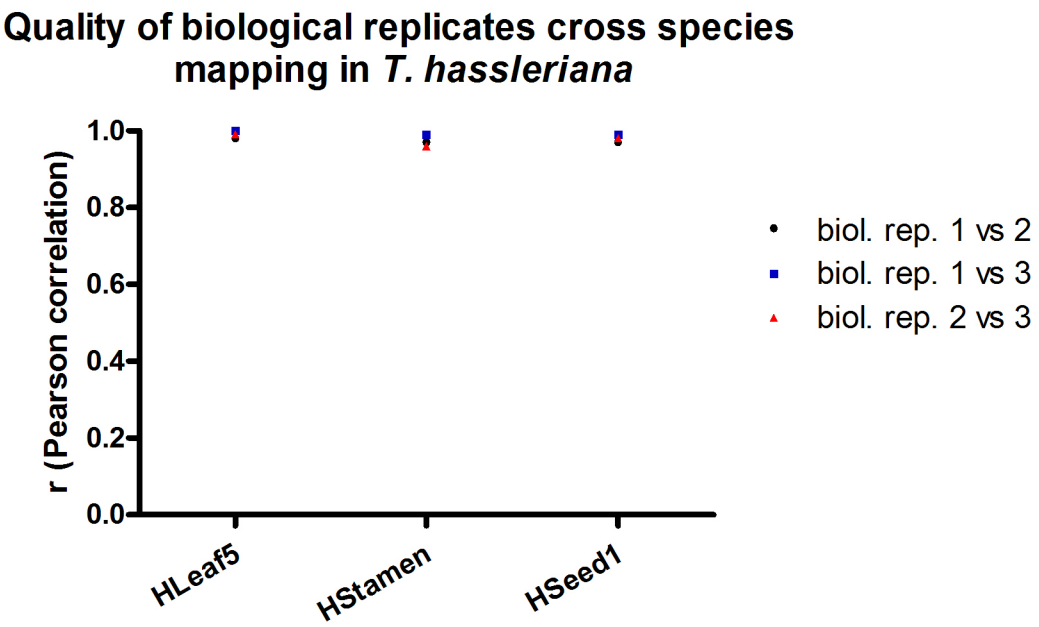


Supplemental Figure 4. Quality assessment of Velvet/OASES assembled *T. hassleriana* contigs against predicted corresponding cds from *T. hassleriana* genome.
(A) Percentage of contig number per predicted cds (Cheng et al., 2013) showing redundancy in assembled contigs.
(B) ClustalW alignment of fragmented contig (top) with corresponding cds (below).
(C) ClustalW alignment of fused contig (top) with corresponding cds (below).

A

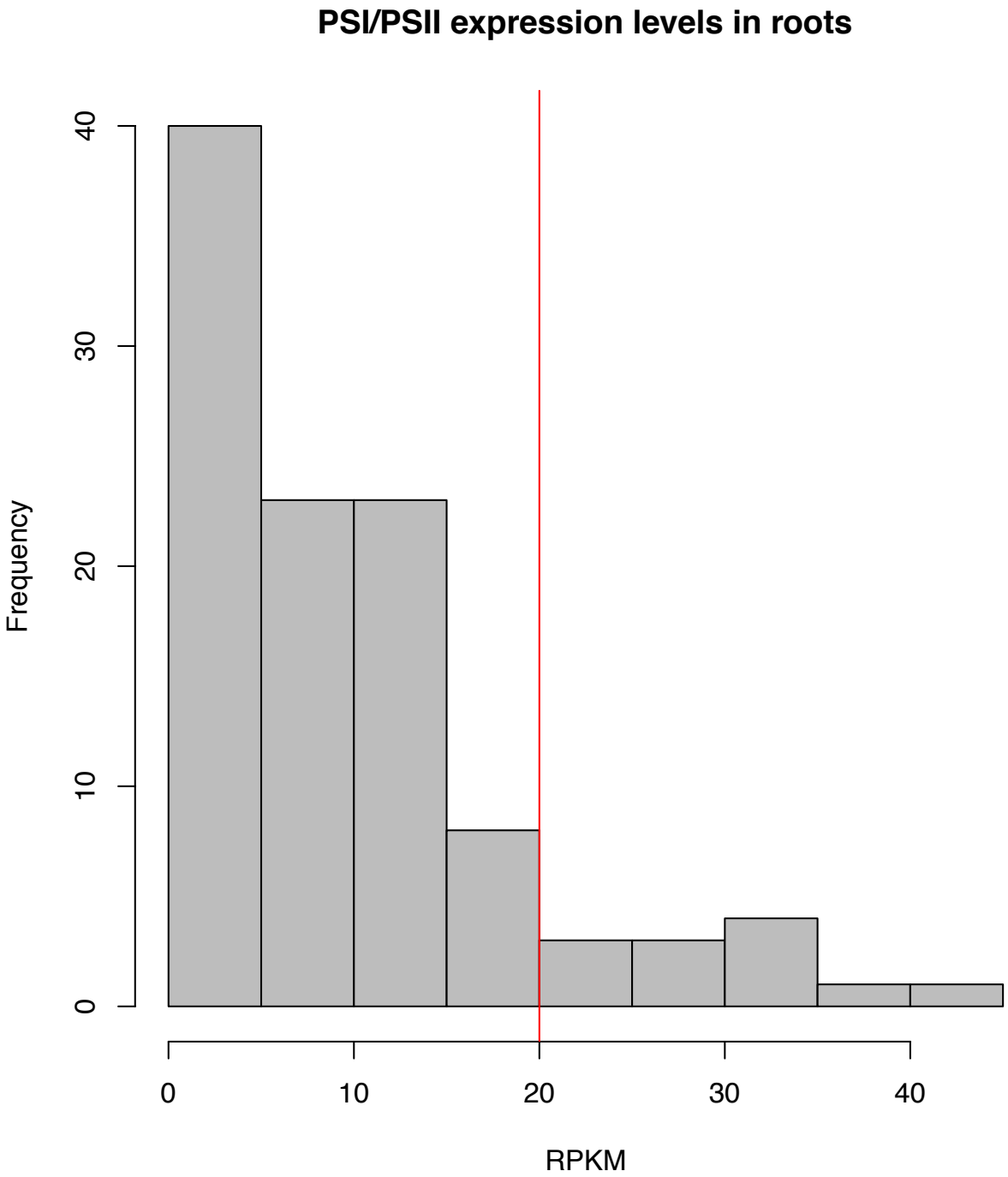


B



Supplemental Figure 5. Quality assessment of the biological replicates of *T. hassleriana* libraries mapped to *A. thaliana* and mapping similarity of *T. hassleriana* libraries mapped to *A. thaliana* and to its own cds.

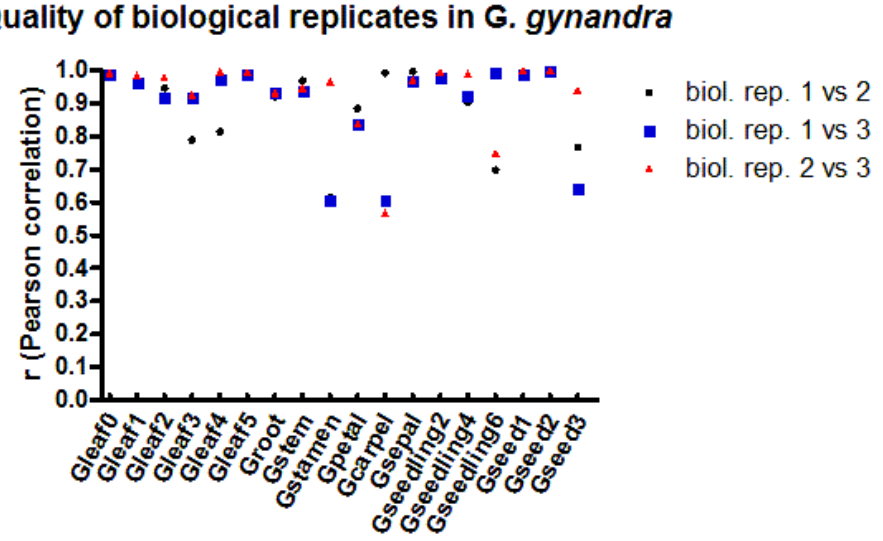
(A) Pair-wise Pearson's correlation (*r*) was calculated for all three pairs of biological replicates for each tissue in *T. hassleriana* mapped to *A. thaliana.* **(B)** Pair-wise Pearson's correlation (*r*) between leaf 5, stamen and seed 1 in (n=3) of *T. hassleriana* mapped to its own coding sequence and *A. thaliana*.



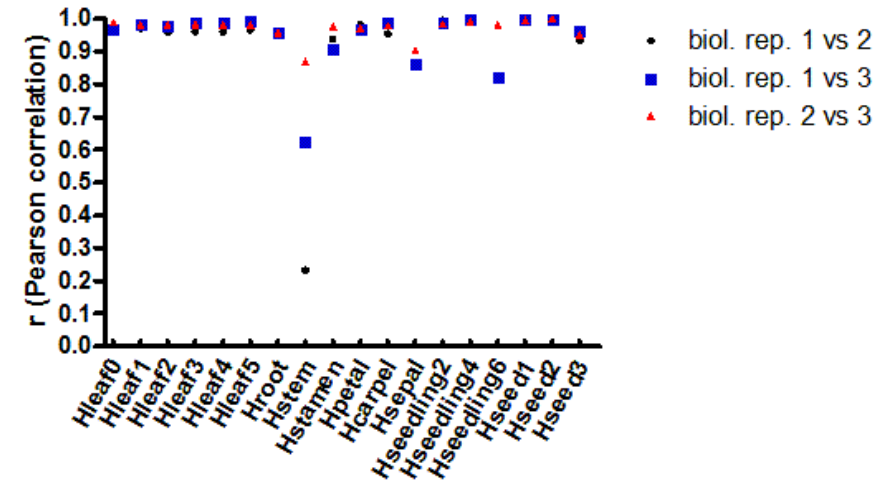
Supplemental Figure 6. Determination of base line gene expression via a histogram of photosystem (PS) I and II transcript abundances reads per mappable million (RPKM) in the *G. gynandra* root.

Y- axis shows frequency and Y- axis depicts RPKM level of PSI and PSII transcript abundance. Red line indicates where threshhold of base line expression was set.

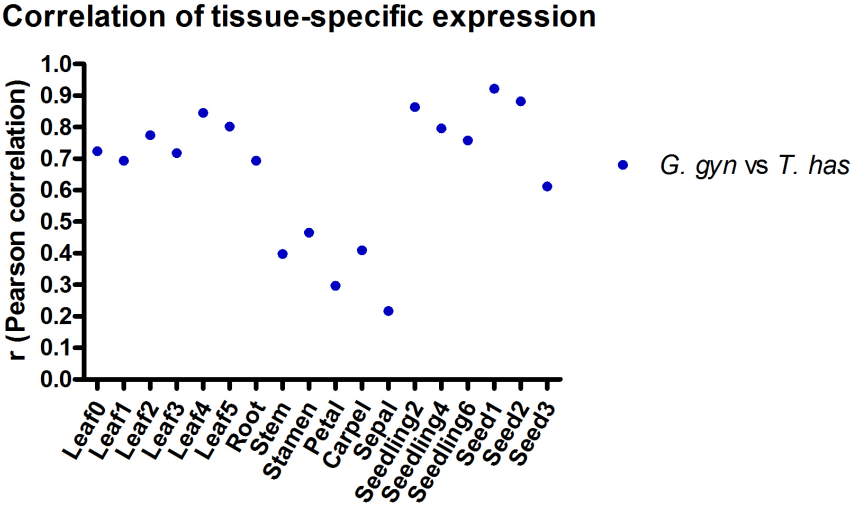
A



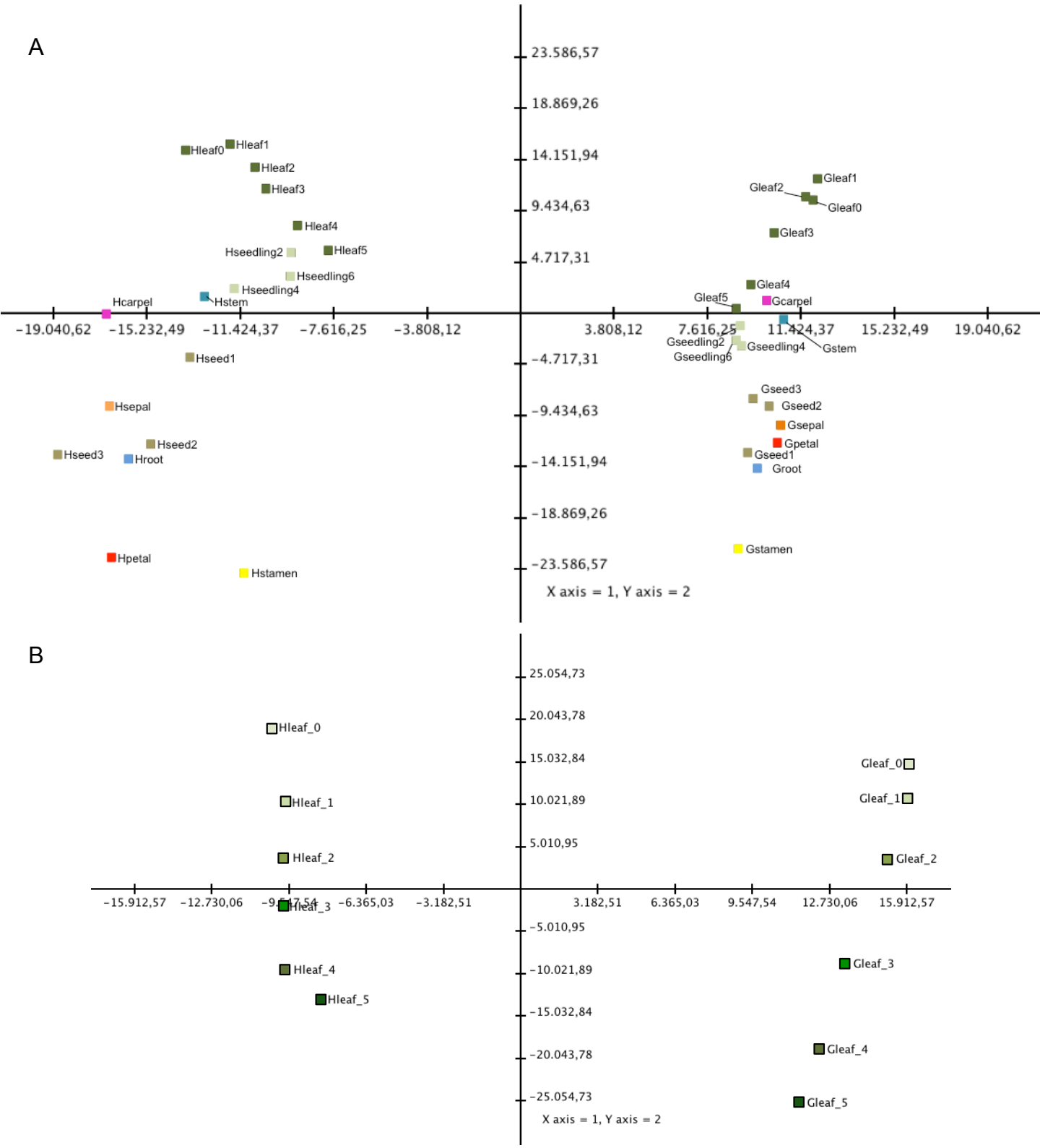
B



C



Supplemental Figure 7. Quality assessment of the biological replicates within each species and tissue similarity between *G. gynandra* and *T. hassleriana*. (A) Pair-wise Pearson's correlation (*r*) was calculated for all three pairs of biological replicates for each tissue (n=3) in *G. gynandra*. **(B)** Pair-wise Pearson's correlation (*r*) was calculated for all three pairs of biological replicates for each tissue (n=3) in *T. hassleriana*. **(C)** Pair-wise Pearson‘s correlation between individual tissues of *T. hassleriana* and *G. gynandra*.

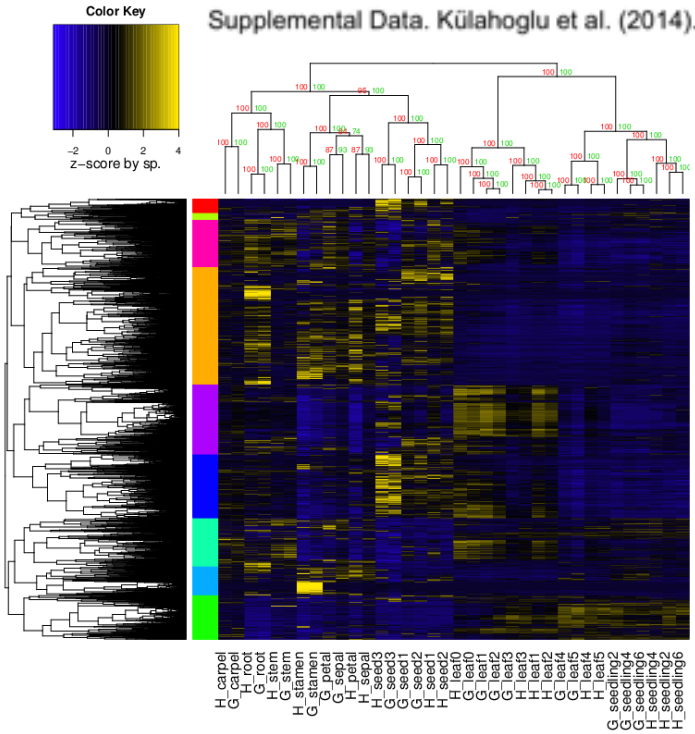


Supplemental Figure 8. Principle component analysis between *G. gynandra* and *T. hassleriana.*

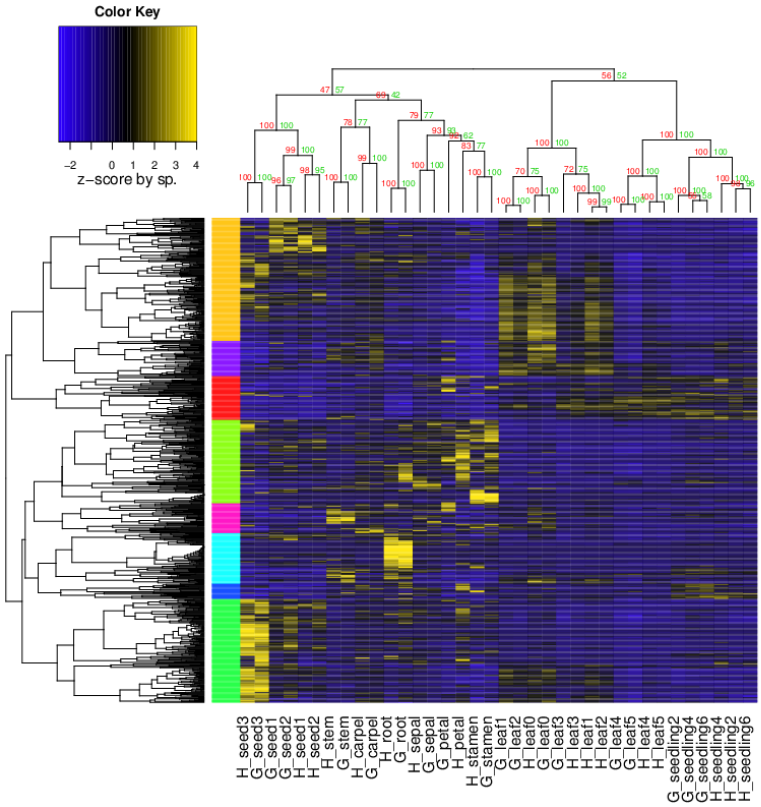
(A) Plot shows all averaged tissues from *G. gynandra* (G) and *T. hassleriana* (H) sequenced (n=3). The first component describes 15% of all data variablility seperating both species. The second component (14%) separates samples by tissue identity within each species. Tissues are indicated by color key (left).

(B) Averaged leaf gradient samples (n=3) from *G. gynandra* (G) and *T. hassleriana* (H) were analysed. First component decribes 44 % and second component describes 29% of variability.

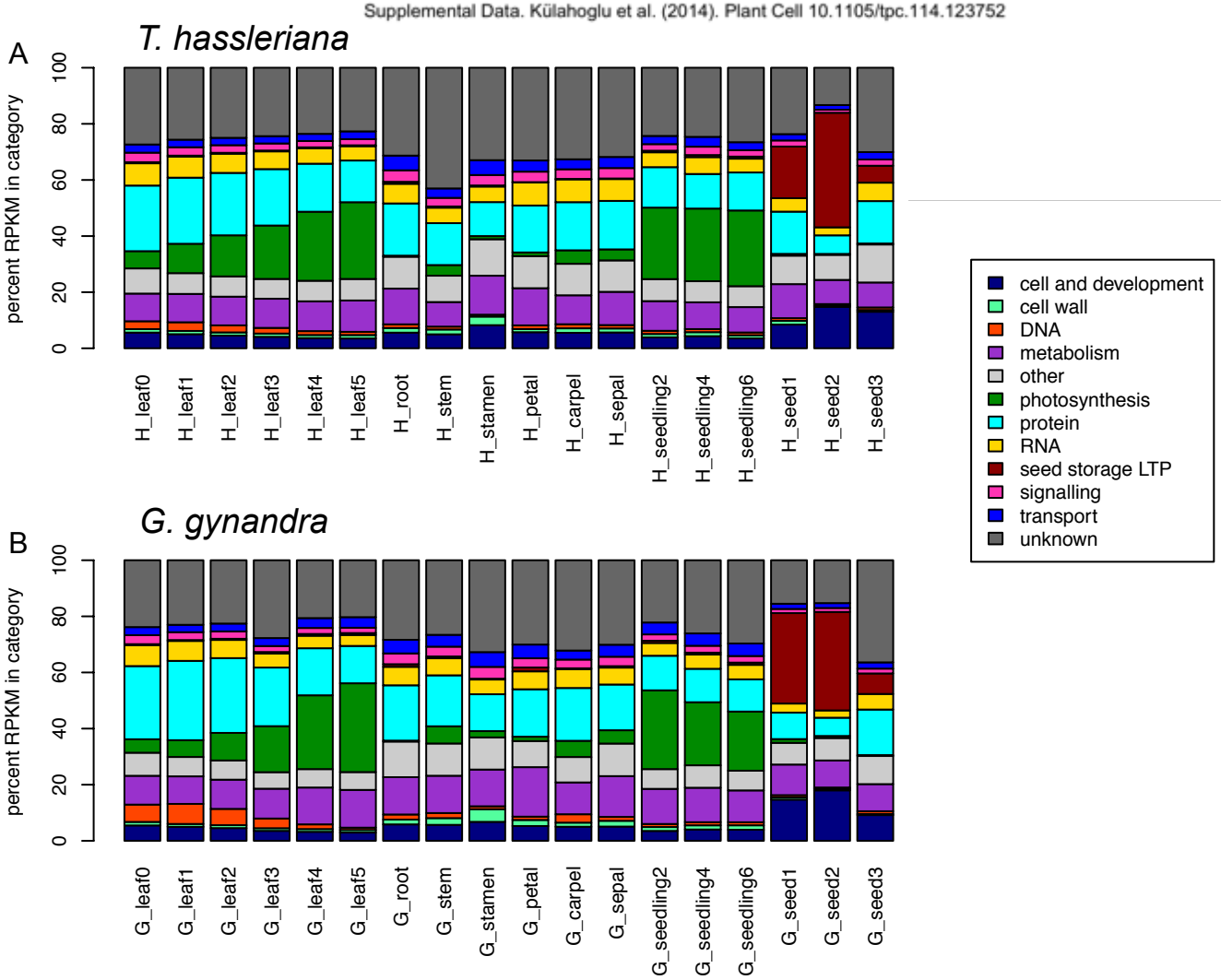
A



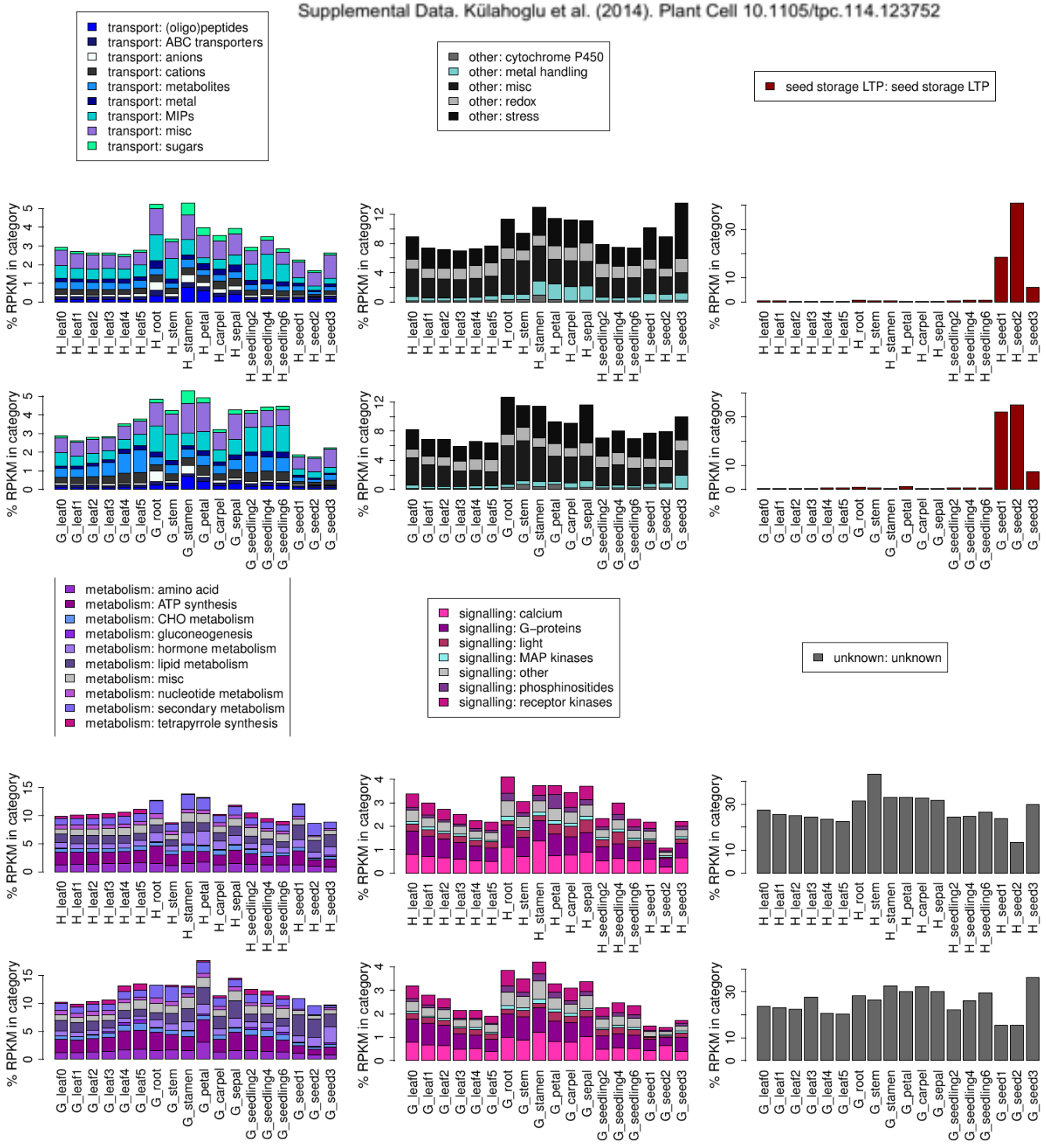
B



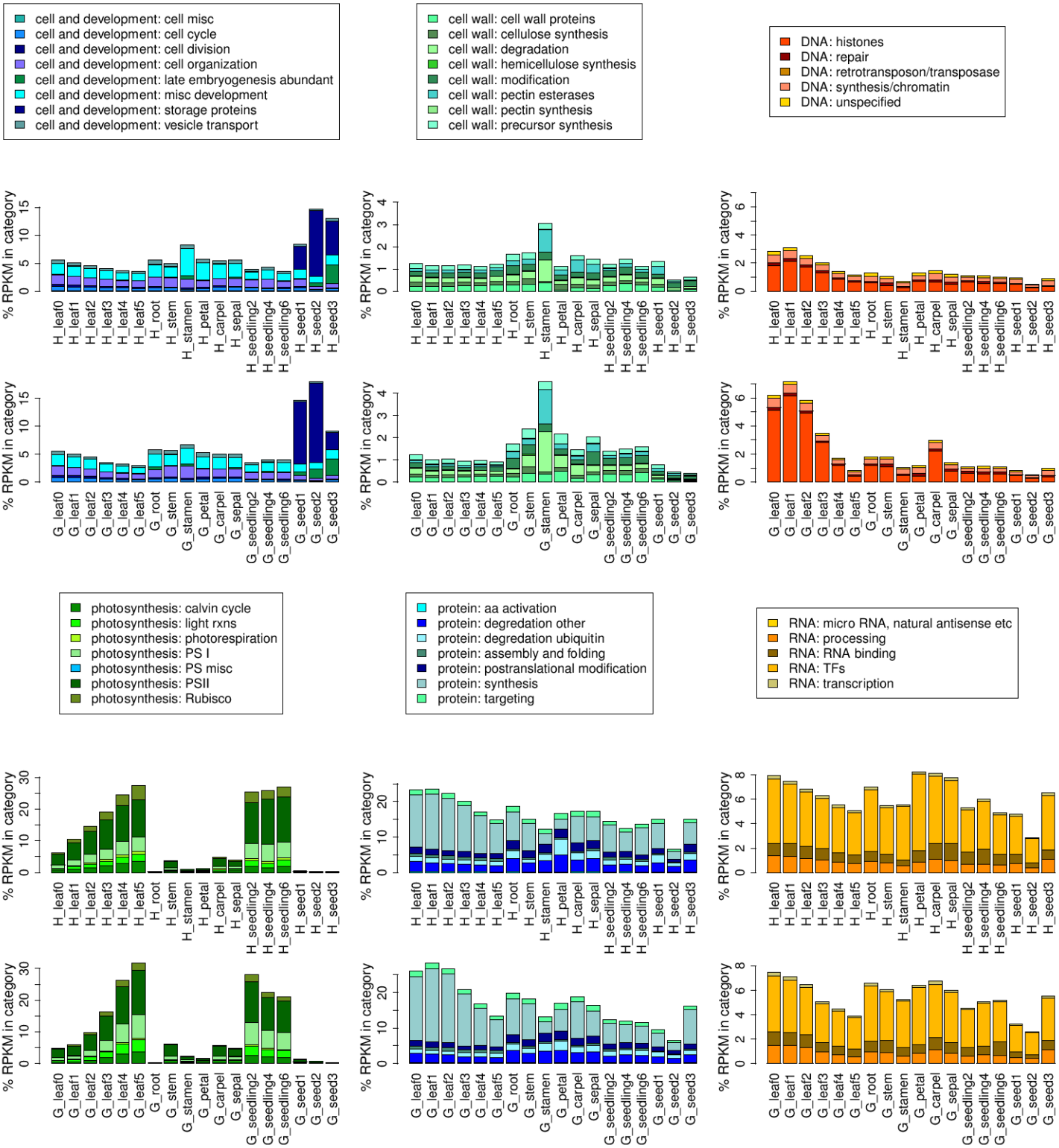
Supplemental Figure 9. Hierarchical cluster analysis with bootstrapped samples of *G. gynandra* and *T. hassleriana*. Numbers above the nodes show the approximately unbiased p-value (red) and the bootstrap probability (green). Blue is lowest expression and yellow highest expression. Left-hand vertical bars denote major clusters in the dendrogram by color. **(A)** Clustering of all over 20 RPKM expressed genes in all averaged samples (n=3). Sample averages were clustered as species scaled Z-scores with Pearson‘s Correlation. **(B)** Hierarchical Clustering of all transcriptional regulators expressed in all tissues sequenced in *G. gynandra* and *T. hassleriana*. Sample averages (n=3) were clustered as species-scaled Z-scores with Pearson‘s Correlation.



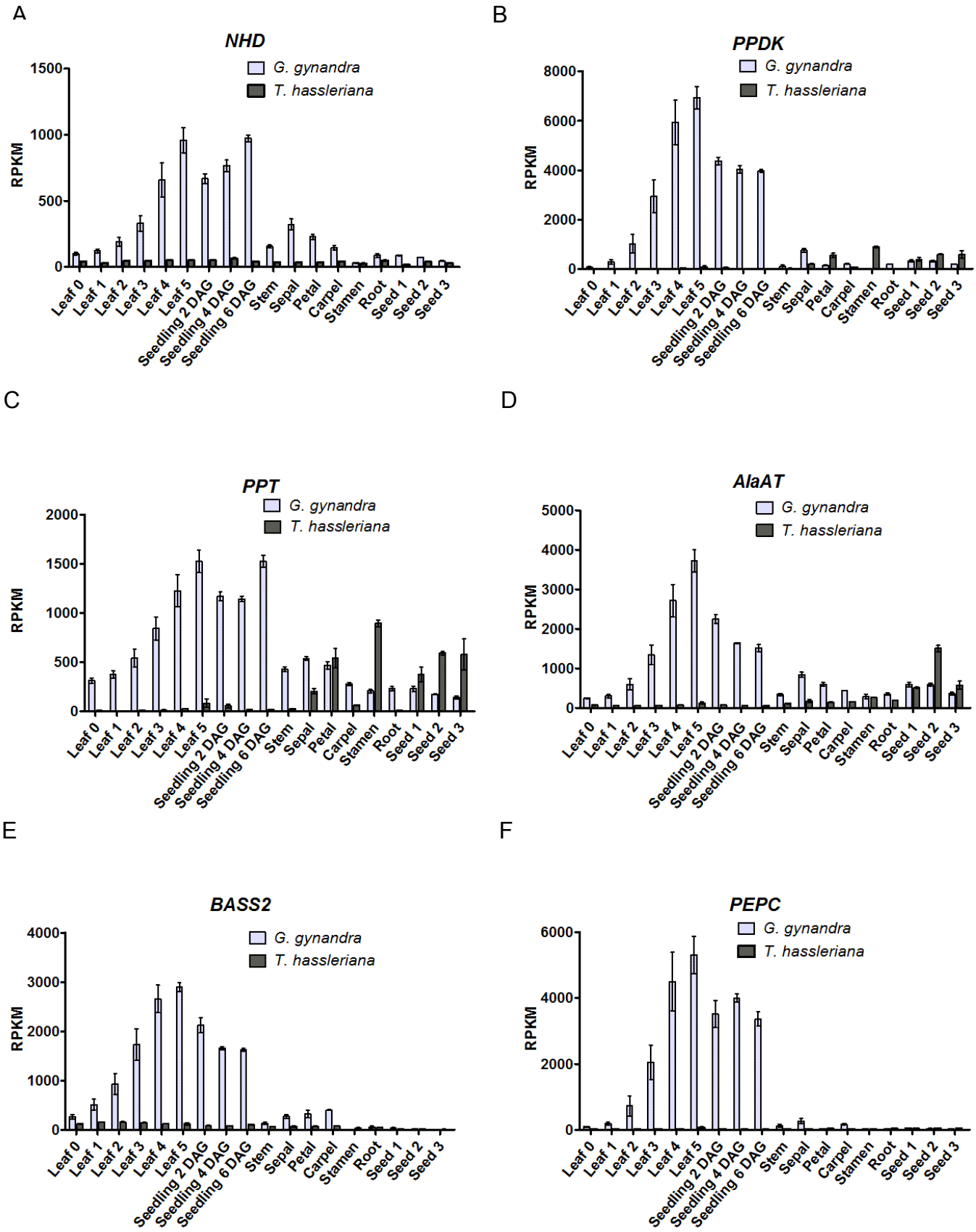
Supplemental Figure 10. Transcriptional investment of each tissue compared in both species. Cumulative average RPKMs in percent for basal Mapman categories for each tissue in *G. gynandra* and *T. hassleriana*.



Supplemental Figure 11.1. Transcriptional investment at secondary Mapman category of each tissue compared in both species (Part 1). Distribution of the Mapman categories in each tissue in *G. gynandra* and *T. hassleriana*. Plot shows percent of average RPKMs of the 12 customized secondary Mapman bins for each tissue.

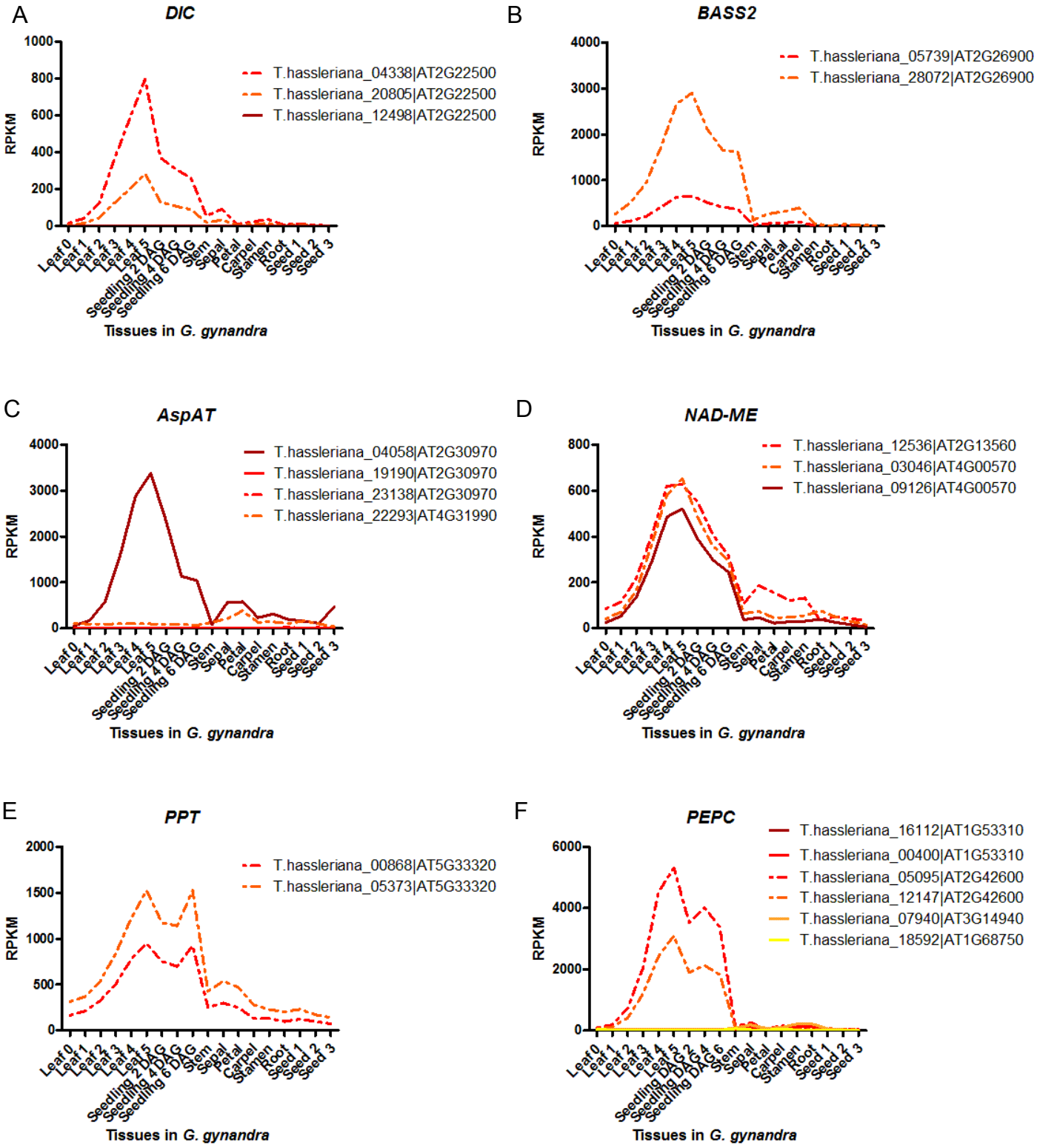


Supplemental Figure 11.2. Transcriptional investment at secondary Mapman category of each tissue compared in both species (Part 2). Distribution of the Mapman categories in each tissue in *G. gynandra* and *T. hassleriana*. Plot shows percent of average RPKMs of the 12 customized secondary Mapman bins for each tissue.

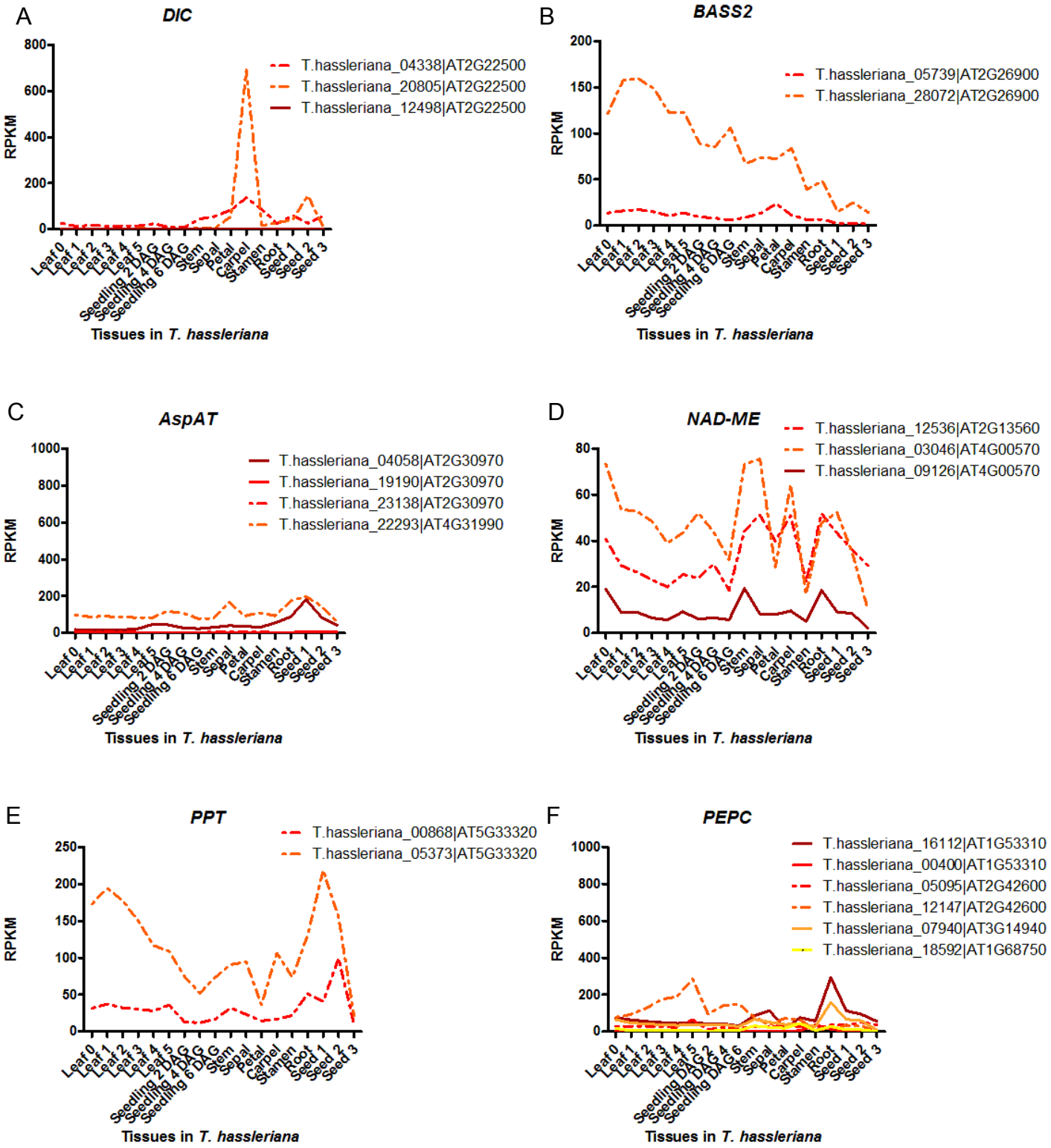


Supplemental Figure 12. Comparison of gene expression dynamics within the leaf gradient of both species.

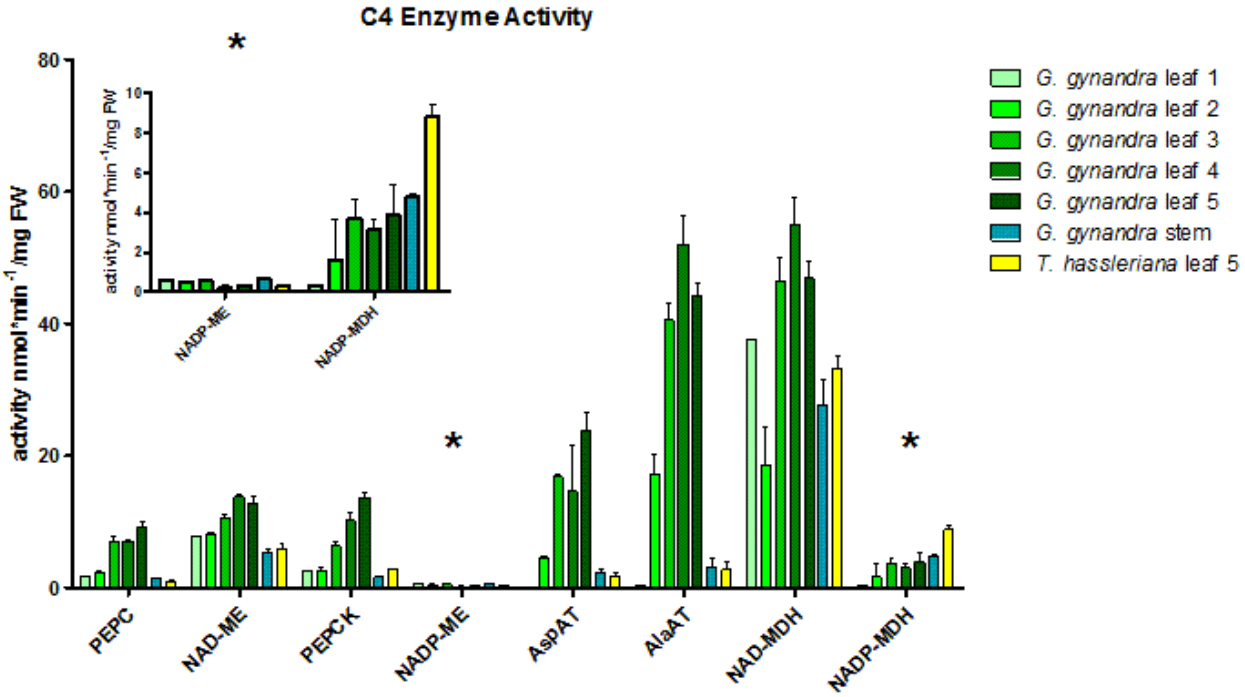
(A-F) Average expression pattern of highest abundant putative ortholog of C_4 cycle genes (*NHD, PPDK, PPT, AlaAT, BASS2, PEPC*) in photo- and heterotrophic tissues in *G. gynandra* (light grey) and *T. hassleriana* (dark grey); (n=3 ± SE, standard error)



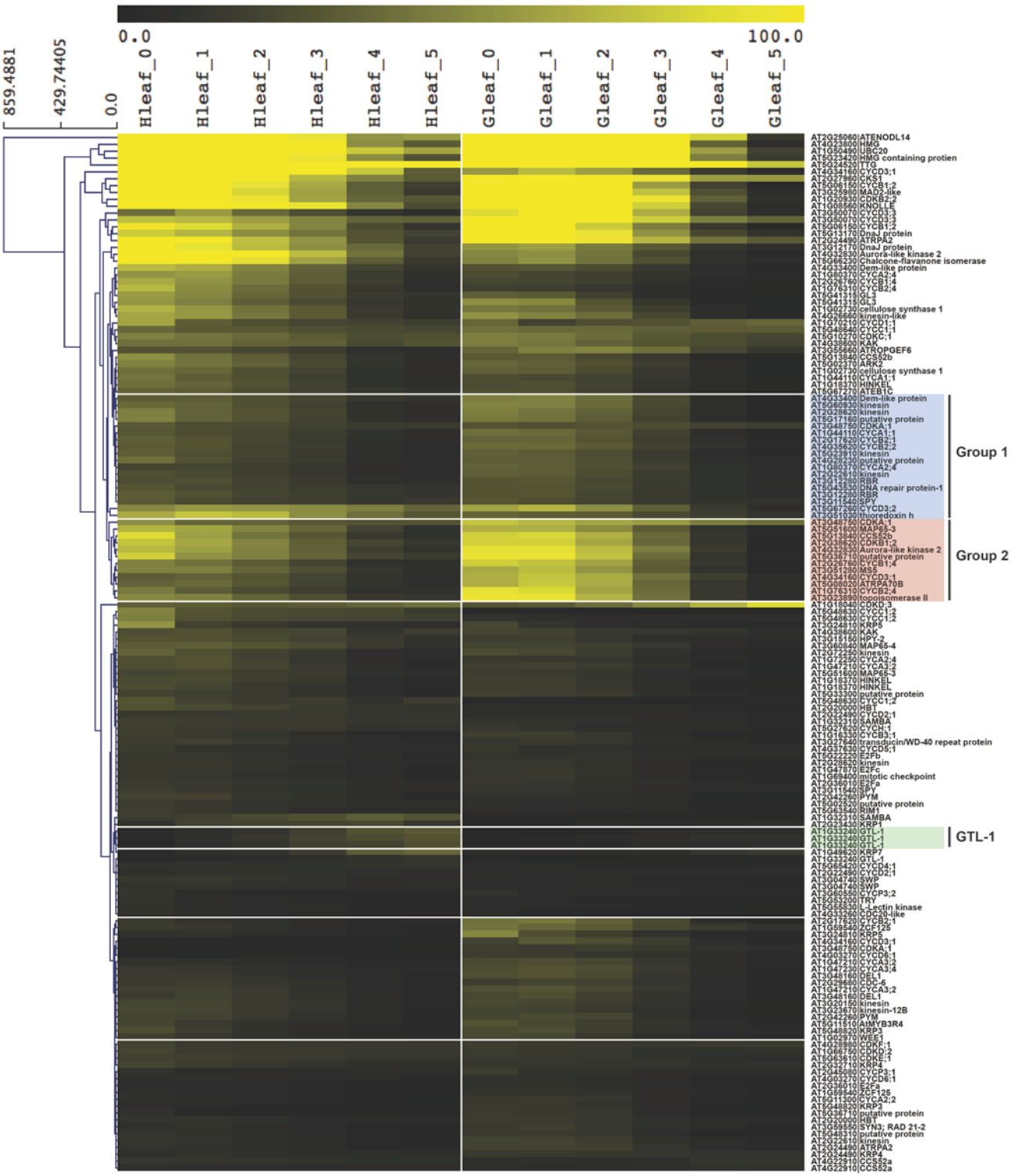
Supplemental Figure online 13. Plot of all C_4 gene putative orthologs expression pattern (RPKM) in *G. gynandra,* that were annotated as C_4 genes with AGI identifier and respective *T. hassleriana* ID. **(A-F)** Average expression pattern of putative ortholog of C_4 cycle genes (*DIC, BASS2, AspAT, NAD-ME, PPT, PEPC*) in photo- and heterotrophic tissues in *G. gynandra* (n=3).



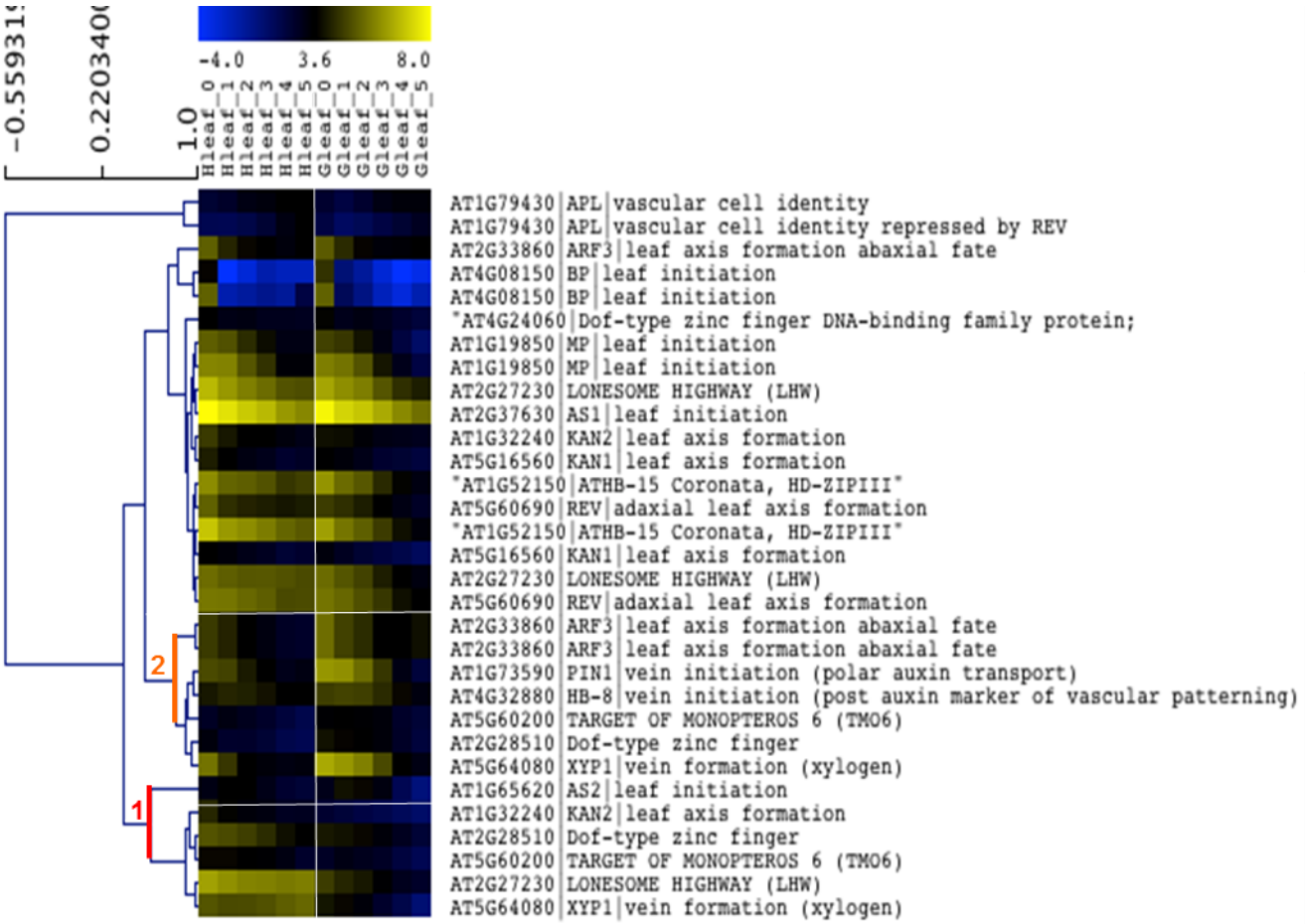
Supplemental Figure online 14. Plot of all C_4 gene putative orthologs expression pattern (RPKM) in *T. hassleriana,* that were annotated as C_4 genes with AGI identifier and respective *T. hassleriana* ID. **(A-F)** Average expression pattern of putative ortholog of C_4 cycle genes (*DIC, BASS2, AspAT, NAD-ME, PPT, PEPC*) in photo- and heterotrophic tissues in *T. hassleriana* (n=3).



Supplemental Figure 15. Enzyme activity measurement of soluble C_4 cycle enzymes. Enzyme activities of PEPC, NAD-ME, PEPCK, NADP-ME, AspAT, AlaAT, NAD-MDH and NADP-MDH were measured along the developing *G. gynandra* leaf (stage 1-5) with the mature *T. hassleriana* leaf (stage 5) as C_3 control. (FW: fresh weight; n=3 ±SE, standard error; biological replicates with each 3 technical replicates)

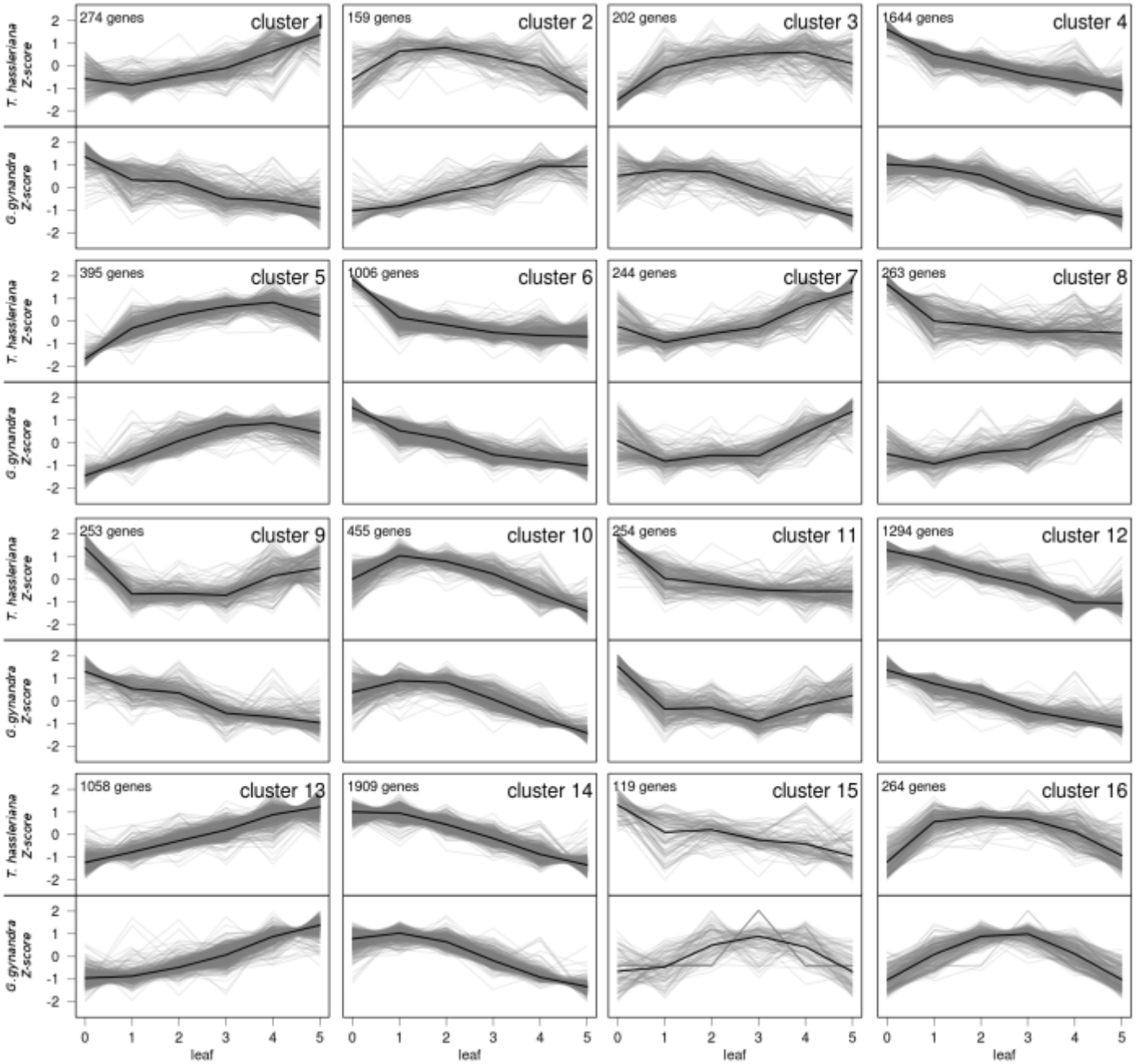


Supplemental Figure 16. Hierarchical clustering of average RPKM with Euclidean distance of core cell cycle genes in *T. hassleriana* and *G. gynandra*.
Core cell cycle genes were extracted from (Vandepoele et al., 2002; Beemster et al., 2005). Deregulated cluster of interest are marked with blue and red boxes. *GTL1* cluster is highlighted with green box.

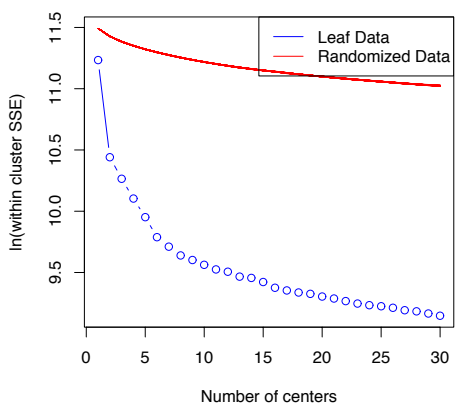


Supplemental Figure 17. Hierarchical clustering with Pearson's correlation of leaf developmental factors. Averaged transcript abundances (RPKM) of leaf gradient sample of transcriptional regulators involved in axial and vasculature fate determination were clustered. Group 1 (orange) and group 2 (red) show genes that are altered between *T. hassleriana* (H) and *G. gynandra* (G).

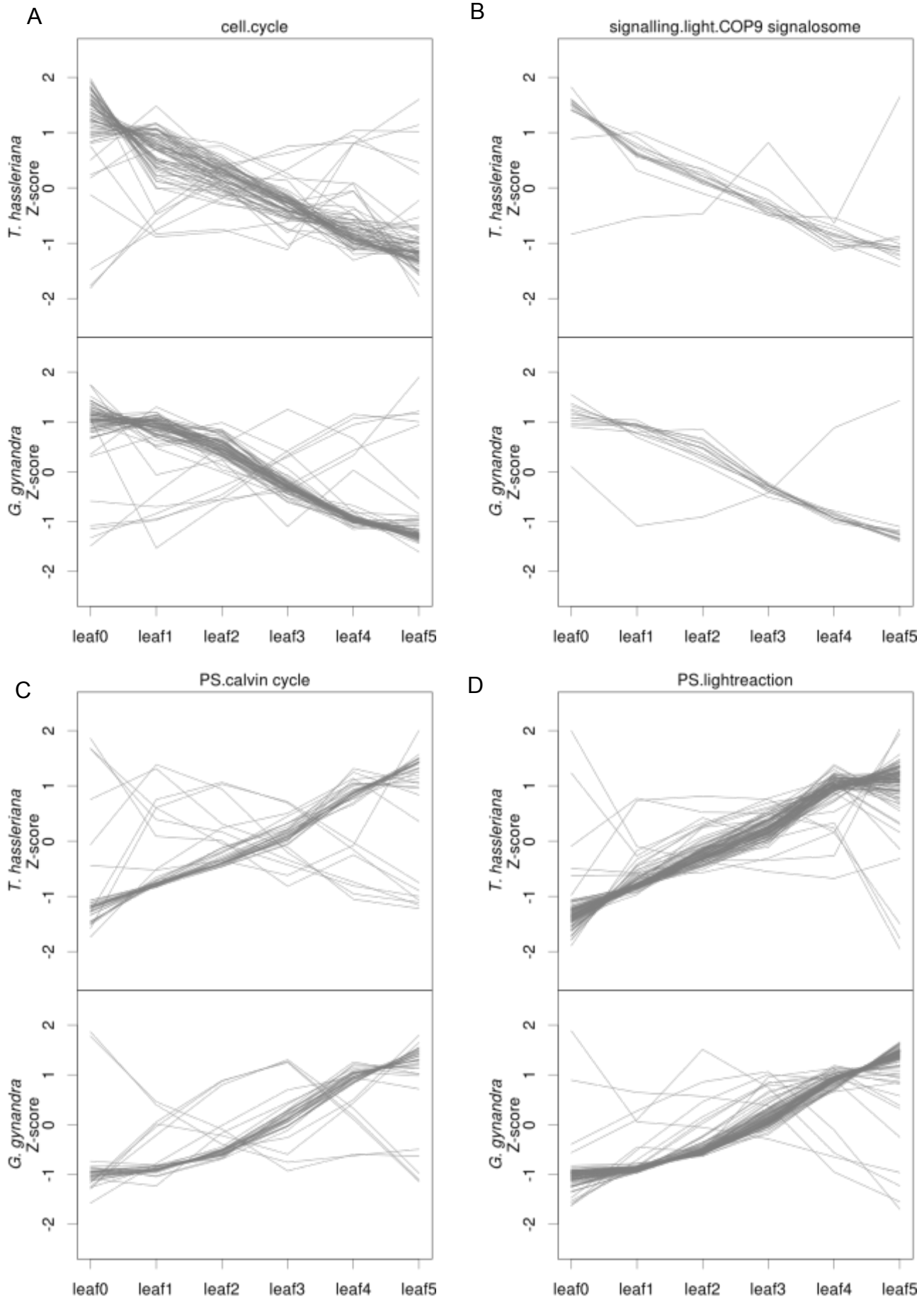
A



B

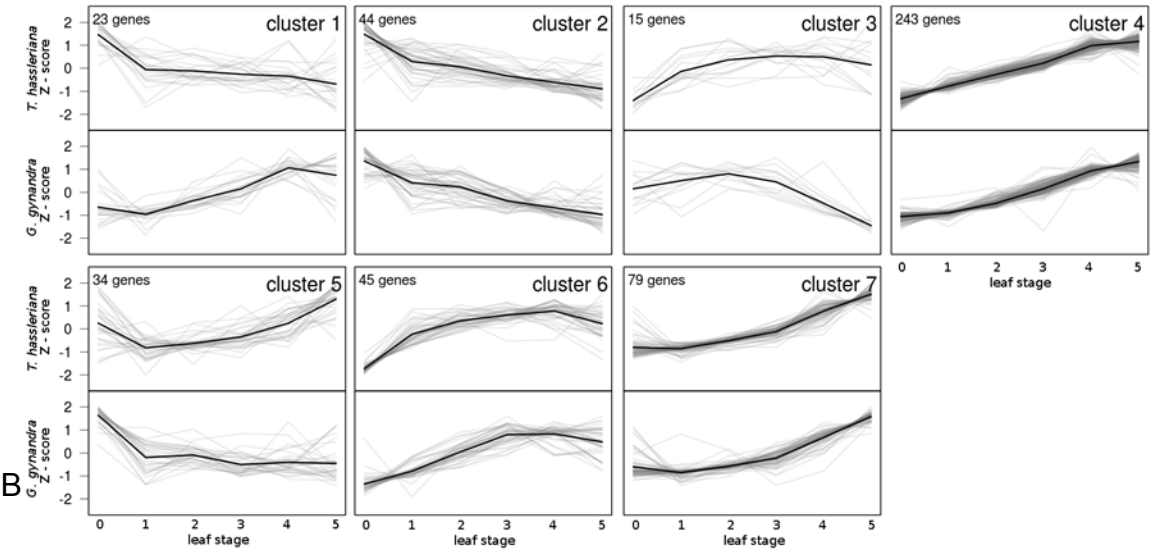


Supplemental Figure 18. *K*-means clustering of leaf gradient expression data and quality assessment. **(A)** *K*-means clustering of transcript abundances (RPKM) of leaf stage averages (*n*=3) between *T. hassleriana* and *G. gynandra* shown as species-scaled Z-scores. Size of each cluster is indicated in each cluster box. **(B)** Ln of the sum of the squared euclidean distance (SSE) between each gene and the center of it's cluster across various numbers of clusters calculated with a *K*-means algorithm for the leaf gradient data (blue) compared to the average of 250 scrambled datasets (red).

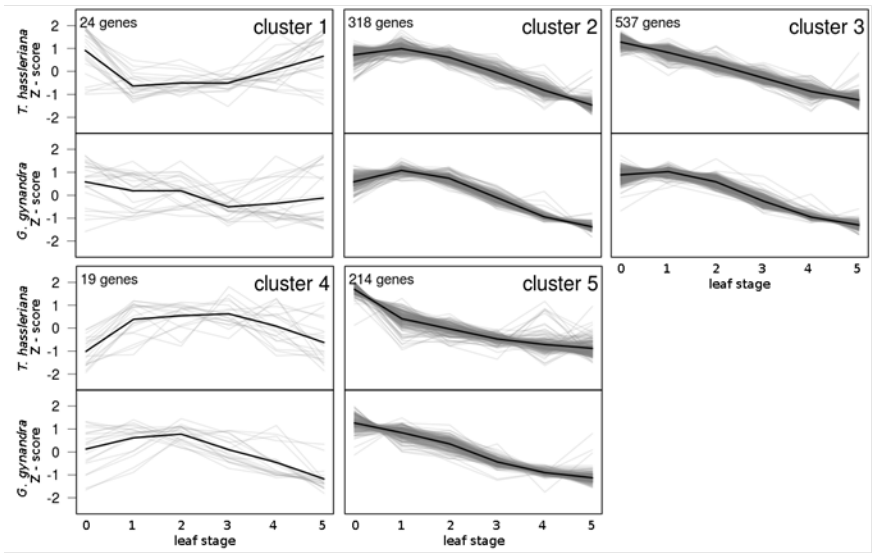


Supplemental Figure 19. Z-score plots of enriched mapman categories in the shifted clusters. Species scaled Z-scores from averaged transcript abundances (RPKM) for each leaf stage per species (n=3). **(A,B)** shifted enriched categories from cluster 4. **(C,D)** shifted enriched categories from cluster 13. Number in brackets are the respective Mapman category bin codes.

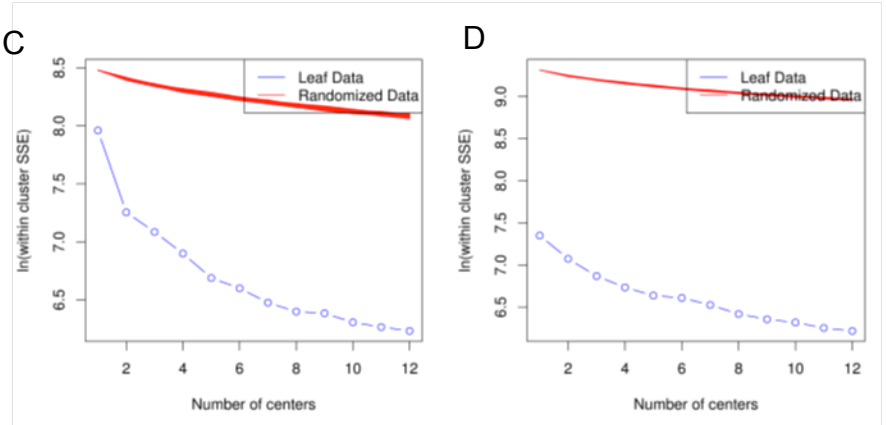
A



B

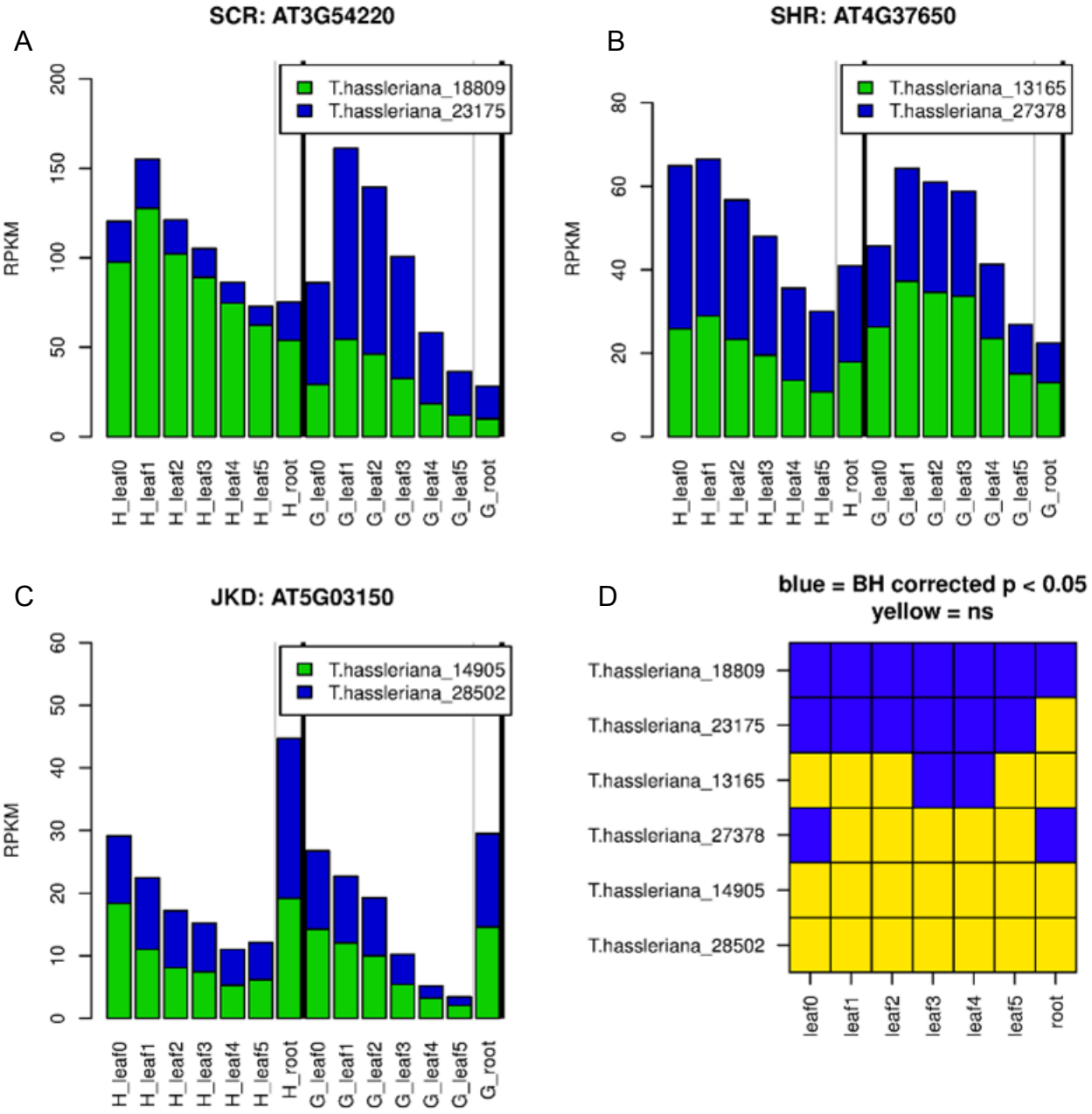


C

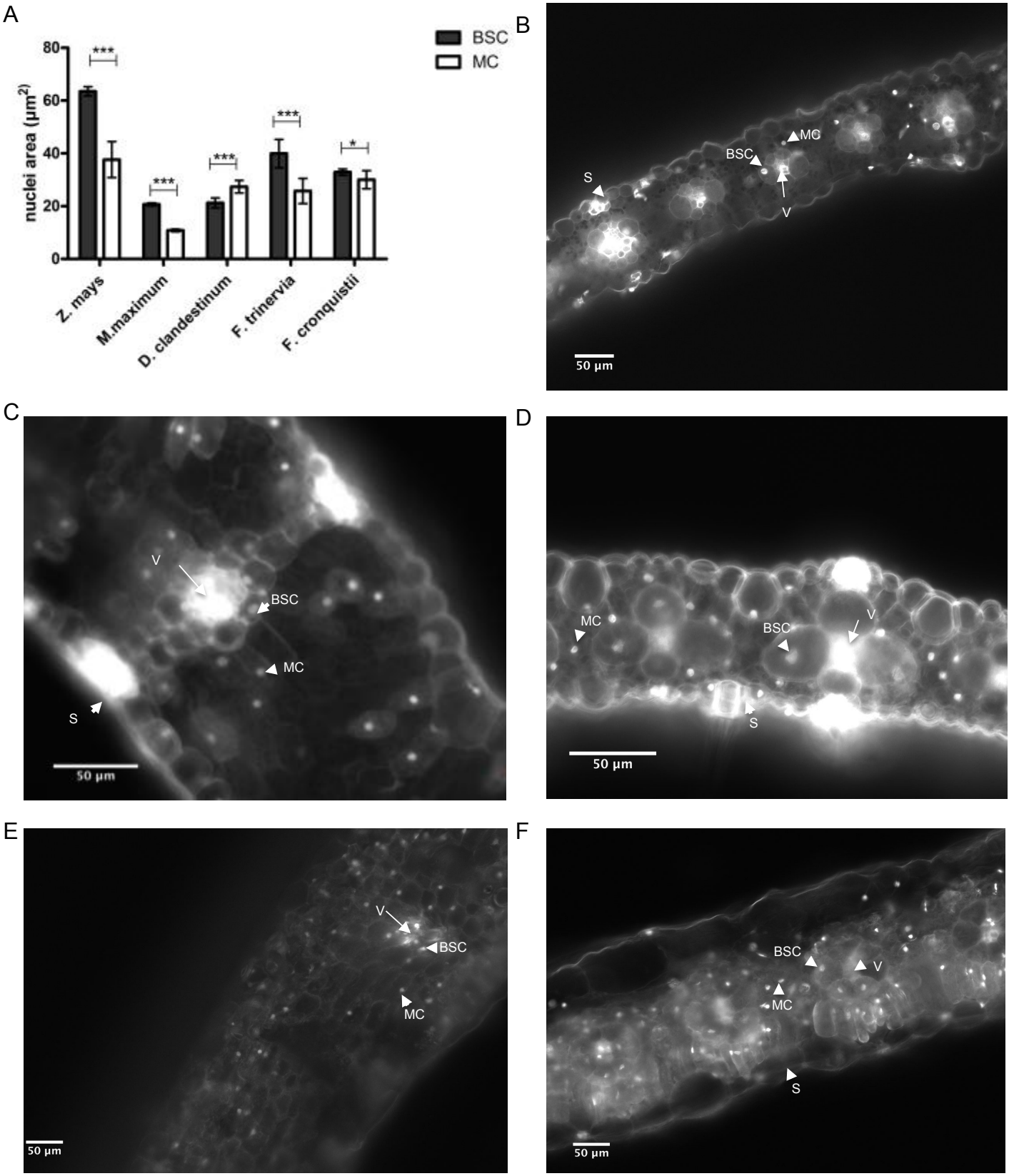


D

Supplemental Figure 20. K-means clustering of genes differentially regulated during the transition from proliferation to enlargement. (A,B) K-means clustering of *T. hassleriana* and *G. gynandra* homologs of gene set that is significantly up-regulated **(A;** p-value<0.05**)** or down-regulated (**B**; p-value<0.05) between day 9 and 10 day in developing *A. thaliana* leaves (Andriankaja et al., 2012). Per species scaled Z-scores from averaged transcript abundances (RPKM) for each leaf stage per species (n=3). **(C,D)** Ln of the sum of the squared Euclidean distance (SSE) between each gene and the center of its clusters across various numbers of clusters calculated with a K-means algorithm for the leaf gradient data (blue) compared to the average of 250 scrambled datasets (red) for **(C)** up- and **(D)** down-regulated.



Supplemental Figure 21. Transcript abundances of *SCARECROW* and *SHORTROOT* homologs in *G. gynandra* (G) and *T. hassleriana* (H) leaf and root.
(A-C) Expression pattern (average RPKM; n=3) of all homologs of *SCARECROW* (***SCR*; A**); *SHORTROOT* (***SHR*; B)** and *JACKDAW* (***JKD*; C)** in both species. **(D)** Dual color map of significant (blue; FWE corrected p-Value<0.05) or non significant (yellow; n.s) expressed transcripts of *SCR*, *SHR* and *JKD*.



Supplemental Figure 22. Nuclei area and images of C_4 and C_3 species.
(A) Quantification of BSC and MC nuclei area of mature leaves of monocotyledonous (*Zea mays*; *Megathyrsus maximus*; *Dichantelium clandestinum*) and dicotyledonous (*Flaveria trinervia*; *Flaveria cronquistii*) C_4 and C_3 species cross sections (error bars ±SD; n=3). Area of nuclei is given as µm^2 with at least 100 nuclei analyzed per cell type per species. Asterisks indicate statistically significant differences between BSC and MC (*** p-value<0.001; * p-value<0.05). **(B-F)** Microscopic fluorescence images of propidium iodide stained mature leaf cross sections of *Zea mays*,C_4 (**B**); *Dichantelium clandestinum*;C_3 (**C**); *Megathyrsus maximus*, C_4 (**D**); *Flaveria cronquistii*, C_3 (**E**); *Flaveria trinervia,* C_4 (**F**). Scale bar: 50 µm; closed arrows pointing to nuclei of indicated cell type. BSC: bundle sheath cell; MC: mesophyll cell; V: vein; S: stomata.

Supplemental Table 1 online. Velvet/OASES assembly stats from *G. gynandra* and *T. hassleriana* paired end reads. Backmapping of paired end reads was performed with TopHat standard settings. Annotation via blastp against TAIR10 proteome.

	G. gynandra (C_4)	*T. hassleriana* (C_3)
k-mer	31	31
N50 contig	1916	1996
unigenes	59471	52479
total transcripts	176850	163456
Backmapping %	60	63
Annotation of TAIR10 %	86	87

Supplemental Table 2 online. Cross species mapping results. *T. hassleriana* Leaf 5, Seed 1, Stamen (n=3) was mapped to *A. thaliana* via blat in translated protein (A) mode to assess sensitvity of cross species mapping. Results of mapping were normalized as RPKM and collapsed on 1 AGI per multiple identifier in *T. hassleriana* Pearson's correlation *r* values of collapsed *T. hassleriana* Leaf 5, Seed 1 and Stamen (n=3) mapped to *A. thaliana* (B) and to itself were calculated (C).

A

Species	Sample	Total number of cleaned reads	Total number of mapped reads	Mapping efficiency against A.thaliana reference	Number of genes >20 RPKM	Number of genes >1000 RPKM
T. hassleriana	**Hleaf5_1**	41085063	23502678	57.20492141	5825	151
	Hleaf5_2	26393836	22289304	84.44889936	5675	122
	Hleaf5_3	67907227	43184738	63.59372913	5684	146
	Hstamen_1	46237107	27726175	59.96520284	5923	48
	Hstamen_2	48025041	28220020	58.76105343	5950	47
	Hstamen_3	17855771	14433105	80.83159781	5467	60
	Hseed1_1	38620315	21654259	56.06960741	6253	39
	Hseed1_2	28792149	17462026	60.64856777	6301	48
	Hseed1_3	25372947	14217549	56.03428329	6107	42

B

	collapsed expression by mapping to own cds vs to A. thaliana	1vs1	2vs2	3vs3	average
Hleaf5	r	0.90	0.89	0.91	0.90
	r2	0.81	0.80	0.82	0.81
Hstamen	r	0.79	0.79	0.79	0.79
	r2	0.62	0.62	0.62	0.62
Hseed1	r	0.91	0.86	0.9	0.89
	r2	0.83	0.74	0.81	0.79

C

	T. hassleriana mapped to *A. thaliana*	1vs2	1vs3	2vs3	average
Hleaf5	r	0.98	1.00	0.98	0.99
	r2	0.97	0.99	0.96	0.97
Hstamen	r	0.97	0.96	0.98	0.97
	r2	0.94	0.92	0.96	0.94
Hseed1	r	0.97	0.99	0.98	0.98
	r2	0.94	0.98	0.96	0.96

Supplemental Table 3 online. Pearson's correlation (r) of each individual replicate per tissue in G. gynandra and T. hassleriana respectively (A). Pearson's correlation between G. gynandra and T. hassleriana individual tissues (B).

A

#	Species	Tissue	Pearson correlation *r* between biological replicates 1 vs 2	1 vs 3	2 vs 3
1	*G. gynandra*	**Gleaf0**	0.98	0.99	0.99
2		**Gleaf1**	0.97	0.96	0.98
3		**Gleaf2**	0.95	0.92	0.98
4		**Gleaf3**	0.79	0.92	0.93
5		**Gleaf4**	0.81	0.97	1.00
6		**Gleaf5**	0.99	0.99	0.99
7		**Groot**	0.92	0.93	0.93
8		**Gstem**	0.97	0.94	0.95
9		**Gstamen**	0.61	0.61	0.97
10		**Gpetal**	0.88	0.84	0.84
11		**Gcarpel**	0.99	0.61	0.57
12		**Gsepal**	1.00	0.97	0.97
13		**Gseedling2**	0.99	0.98	0.99
14		**Gseedling4**	0.90	0.92	0.99
15		**Gseedling6**	0.70	0.99	0.75
16		**Gseed1**	0.99	0.99	1.00
17		**Gseed2**	1.00	1.00	1.00
18		**Gseed3**	0.77	0.64	0.94
19	*T. hassleriana*	**Hleaf0**	0.97	0.97	0.99
20		**Hleaf1**	0.97	0.98	0.98
21		**Hleaf2**	0.96	0.98	0.98
22		**Hleaf3**	0.96	0.99	0.98
23		**Hleaf4**	0.96	0.99	0.98
24		**Hleaf5**	0.97	0.99	0.98
25		**Hroot**	0.95	0.96	0.96
26		**Hstem**	0.23	0.62	0.87
27		**Hstamen**	0.94	0.91	0.98
28		**Hpetal**	0.98	0.97	0.97
29		**Hcarpel**	0.95	0.99	0.98
30		**Hsepal**	0.87	0.86	0.90
31		**Hseedling2**	0.99	0.99	0.98
32		**Hseedling4**	0.99	1.00	0.99
33		**Hseedling6**	0.82	0.82	0.98
34		**Hseed1**	0.99	1.00	0.99
35		**Hseed2**	1.00	1.00	1.00
36		**Hseed3**	0.93	0.96	0.95

Supplemental Table 3 online. Pearson's correlation (r) of each individual replicate per tissue in G. gynandra and T. hassleriana respectively (A). Pearson's correlation between G. gynandra and T. hassleriana individual tissues (B).

B

Pearson Correlation *r* between G. *gynandra* and *T. hassleriana*		
#	**Tissue**	**r**
1	**Leaf0**	0.723369664
2	**Leaf1**	0.693967315
3	**Leaf2**	0.774414647
4	**Leaf3**	0.718280077
5	**Leaf4**	0.845767325
6	**Leaf5**	0.801946455
7	**Root**	0.693418487
8	**Stem**	0.397920288
9	**Stamen**	0.465027959
10	**Petal**	0.296842384
11	**Carpel**	0.409336161
12	**Sepal**	0.216833607
13	**Seedling2**	0.864093832
14	**Seedling4**	0.79602302
15	**Seedling6**	0.757896499
16	**Seed1**	0.922002838
17	**Seed2**	0.882400443
18	**Seed3**	0.612106172

Supplemental Table 4 online. Number of significatly up- or downregulated genes in *G. gynandra* compared to *T. hassleriana* within the different tissues. Differential expressed gene p-Values were calculated via EdgeR and Bonferroni-Holms corrected, genes with p<0.05 were classified as differential regulated.

Tissue	UP p< 0.05	UP p< 0.01	UP p< 0.001	DOWN p< 0.05	DOWN p< 0.01	DOWN p< 0.001
leaf0	5435	5061	4539	6076	5696	5237
leaf1	5197	4841	4391	5914	5529	5026
leaf2	4234	3894	3443	5047	4644	4204
leaf3	4646	4283	3833	5484	5070	4576
leaf4	3250	2911	2511	3774	3399	2979
leaf5	3236	2894	2447	4133	3716	3191
root	4343	3973	3511	5151	4755	4254
stem	7835	7497	7123	8462	8129	7698
stamen	4545	4116	3652	5388	4976	4451
petal	4445	4063	3613	5122	4751	4317
carpel	3718	3352	2929	3640	3274	2894
sepal	5650	5276	4780	6422	6023	5539
seedling2	4012	3644	3186	4354	3981	3546
seedling4	4113	3684	3202	4416	4043	3569
seedling6	2874	2534	2180	3542	3154	2714
seed1	4116	3764	3321	4457	4083	3591
seed2	6600	6270	5807	7075	6727	6276
seed3	6108	5725	5307	7088	6674	6190
mean	4686.5	4321.222222	3876.388889	5308.055556	4923.555556	4458.444444
max	7835	7497	7123	8462	8129	7698

Supplemental Table 5 online. List of genes present in root to shoot recruitment module.

T. hassleriana cds ID (Cheng et al., 2013)	Arabidopsis homologue	Coexpressed with TF	TAIR short annotation
T.hassleriana_10164	AT1G70410		beta carbonic anhydrase 4
T.hassleriana_20805	AT2G22500		uncoupling protein 5
T.hassleriana_17885	**AT5G61590**	**ERF**	**Integrase-type DNA-binding superfamily protein**
T.hassleriana_27615	**AT1G04250**	**Aux/IAA**	**AUX/IAA transcriptional regulator family protein**
T.hassleriana_13599	**AT5G13180**	**VND-I2**	**NAC domain containing protein 83**
T.hassleriana_07159	AT4G12730	Aux/IAA	FASCICLIN-like arabinogalactan 2
T.hassleriana_22160	AT5G57560		Xyloglucan endotransglucosylase/hydrolase family protein
T.hassleriana_03276	AT1G11545	Aux/IAA	xyloglucan endotransglucosylase/hydrolase 8
T.hassleriana_11774	AT1G43670		Inositol monophosphatase family protein
T.hassleriana_19959	AT5G19140	ERF	Aluminium induced protein with YGL and LRDR motifs
T.hassleriana_13658	AT1G25230	ERF	Calcineurin-like metallo-phosphoesterase superfamily protein
T.hassleriana_11758	AT3G14690	VND-I2	cytochrome P450, family 72, subfamily A, polypeptide 15
T.hassleriana_00726	AT5G46900		Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily
T.hassleriana_13312	AT3G22120		cell wall-plasma membrane linker protein
T.hassleriana_18867	AT3G54110		plant uncoupling mitochondrial protein 1
T.hassleriana_22110	AT1G14870		PLANT CADMIUM RESISTANCE 2
T.hassleriana_13333	AT5G19190		
T.hassleriana_11698	AT3G13950		
T.hassleriana_01980	AT5G25265		
T.hassleriana_04483	AT5G62900		
T.hassleriana_21987	AT1G13700	ERF	6-phosphogluconolactonase 1
T.hassleriana_15837	AT1G05000		Phosphotyrosine protein phosphatases superfamily protein
T.hassleriana_08797	AT5G23750	Aux/IAA	Remorin family protein
T.hassleriana_08517	AT5G36160		Tyrosine transaminase family protein
T.hassleriana_12936	AT5G25980		glucoside glucohydrolase 2
T.hassleriana_04639	AT2G01660		plasmodesmata-located protein 6
T.hassleriana_22812	AT4G21870	ERF	HSP20-like chaperones superfamily protein
T.hassleriana_10363	AT3G11660	VND-I2	NDR1/HIN1-like 1
T.hassleriana_19882	AT3G04720		pathogenesis-related 4
T.hassleriana_27070	AT2G15220		Plant basic secretory protein (BSP) family protein
T.hassleriana_05312	AT2G37170		plasma membrane intrinsic protein 2
T.hassleriana_05313	AT2G37170		plasma membrane intrinsic protein 2
T.hassleriana_12285	AT2G36830	Aux/IAA	gamma tonoplast intrinsic protein
T.hassleriana_12284	AT2G36830		gamma tonoplast intrinsic protein
T.hassleriana_14369	AT1G11670	Aux/IAA	MATE efflux family protein
T.hassleriana_08980	N.A.		
T.hassleriana_07000	N.A.		

Supplemental Table online 6. List of clustered general leaf developmental and vasculature regulating genes along both leaf gradients.

T. hassleriana cds ID (Cheng et al., 2013)	AGI	Annotation based on TAIR10	Function in vascular development
T.hassleriana_16883	AT1G19850	MONOPTEROS (MP)	leaf initiation
T.hassleriana_08823	AT1G19850	MONOPTEROS (MP)	leaf initiation
T.hassleriana_08424	AT1G32240	KANADI 2 (KAN2)	leaf axis formation
T.hassleriana_09176	AT1G32240	KANADI 2 (KAN2)	leaf axis formation
T.hassleriana_20498	AT1G52150	ATHB-15	neg reg of vasc cell diff
T.hassleriana_09793	AT1G52150	ATHB-15	neg reg of vasc cell diff
T.hassleriana_06450	AT1G65620	ASYMMETRIC LEAVES 2 (AS2)	leaf initiation
T.hassleriana_19648	AT1G73590	PIN-FORMED 1 (PIN1)	vein initiation (polar auxin transport)
T.hassleriana_01843	AT1G79430	ALTERED PHLOEM DEVELOPMENT (APL)	vascular cell identity repressed by REV
T.hassleriana_19440	AT1G79430	ALTERED PHLOEM DEVELOPMENT (APL)	vascular cell identity repressed by REV
T.hassleriana_27016	AT2G13820	Bifunctional inhibitor/lipid-transfer protein	vein formation (xylogen)
T.hassleriana_27989	AT2G27230	LONESOME HIGHWAY (LHW)	transcription factor-related
T.hassleriana_09087	AT2G27230	LONESOME HIGHWAY (LHW)	transcription factor-related
T.hassleriana_15265	AT2G27230	LONESOME HIGHWAY (LHW)	transcription factor-related
T.hassleriana_15152	AT2G28510	Dof-type zinc finger DNA-binding family protein	Dof-type zinc finger DNA-binding family protein
T.hassleriana_27908	AT2G28510	Dof-type zinc finger DNA-binding family protein	Dof-type zinc finger DNA-binding family protein
T.hassleriana_06822	AT2G33860	ETTIN (ETT)	leaf axis formation abaxial fate
T.hassleriana_23279	AT2G33860	ETTIN (ETT)	leaf axis formation abaxial fate
T.hassleriana_23086	AT2G37630	ASYMMETRIC LEAVES 1 (AS1)	leaf initiation
T.hassleriana_18733	AT4G08150	KNOTTED-like from Arabidopsis thaliana (KNAT1)	leaf initiation
T.hassleriana_09854	AT4G08150	KNOTTED-like from Arabidopsis thaliana (KNAT1)	leaf initiation
T.hassleriana_25576	AT4G24060	Dof-type zinc finger DNA-binding family protein	Dof-type zinc finger DNA-binding family protein
T.hassleriana_22410	AT4G32880	homeobox gene 8 (HB-8)	vein initiation (post auxin marker of vascular patterning)
T.hassleriana_28697	AT5G16560	KANADI (KAN)	leaf axis formation abaxial; neg reg of PIN1
T.hassleriana_19776	AT5G16560	KANADI (KAN)	leaf axis formation abaxial; neg reg of PIN1
T.hassleriana_18288	AT5G60200	TARGET OF MONOPTEROS 6 (TMO6)	TARGET OF MONOPTEROS 6
T.hassleriana_16642	AT5G60200	TARGET OF MONOPTEROS 6 (TMO6)	TARGET OF MONOPTEROS 6
T.hassleriana_18265	AT5G60690	REVOLUTA (REV)	adaxial leaf axis formation
T.hassleriana_19132	AT5G60690	REVOLUTA (REV)	adaxial leaf axis formation
T.hassleriana_17767	AT5G64080	XYP1	vein formation (xylogen)
T.hassleriana_26861	AT5G64080	XYP1	vein formation (xylogen)

Supplemental Methods

Leaf clearings and safranine staining (Supplemental Figure 1)

For leaf clearings *T. hassleriana* and *G. gynandra* leaves of stage 0 to 5 were destained in 70% EtOH with 1% glycerol added for 24 hrs and cleared in 5% NaOH until they appeared translucent and rinsed with H_2O_{dest}. Leaves were imaged under dark field settings with stereo microcope SMZ1500 (Nikon, Japan). Prior safranine staining, leaves were destained with increasing EtOH series until 100% EtOH and stained for 5 -10 min with 1% safranine (1g per 100ml 96% EtOH). After destaining leaves were analyzed with bright field microscope (Zeiss, Germany). Vein orders were determined by width and position as described by (McKown and Dengler, 2009) for Flaveria species.

Contig assembly and annotation (Supplemental Figure 4, Table 1 and Dataset 3)

Cleaned and filtered paired end (PE) reads were used to create a reference transcriptome for each species. The initial *de novo* assembly was optimized by using 31-kmer using Velvet (v1.2.07) and Oases (v0.2.08) pipeline (Zerbino and Birney, 2008; Schulz et al., 2012). For quality purposes the longest assembled transcript was selected with custom made perl scripts if multiple contigs were present (Schliesky et al., 2012) resulting in 59,471 *G. gynandra* and 52,479 *T. hassleriana* contigs. For quality assessment PE reads were aligned again to the respective contigs for each species via TopHat standard settings with over 60% backmapping efficiency in both species. Assembled longest transcripts were annotated using BLASTX mapping against TAIR10 proteome database (cut-off $1e^{-10}$). The best blastx hits were filtered by the highest bitscore. For quality assessment of contigs, *T. hassleriana* contigs were aligned with BLASTN against *T. hassleriana* predicted cds (Cheng et al., 2013). Multiple matching contigs to one cds identifier were filtered with customized perl script.

Cross species mapping sensitivity assessment (Supplemental Figure 5; Table 2)

All three biological replicates of leaf stage 5, stamen and young seed from *T. hassleriana* were mapped with BLAT V35 in dnax mode (nucleotide sequence of query and reference are translated in six frames to protein) with default parameters to both, the *T. hassleriana* gene models and the *A. thaliana* TAIR10 representative gene models. Subsequently, the BLAT output was filtered for the best match per read based on the highest score. RPKMs were calculated based on mappable reads per million (RPKM). The RPKM expression data was collapsed to single *A. thaliana* AGIs (RPKM were added) to avoid multiple assigned *T. hassleriana*'s IDs to the same AGI. Pearson's correlation r was calculated between the mapped *T. hassleriana* replicates mapped on *A. thaliana* gene models among each other. Also Pearson's correlation r was calculated between cross species mapped *T. hassleriana* leaf5, stamen and seed1 replicates and the replicates of Leaf5 mapped to its own cds in *T. hassleriana*.

Principal component analysis (Supplemental Figure 8)

Principal component analyses (PCA,Yeung and Ruzzo, 2001) was carried out with MULTI EXPERIMENT VIEWER VERSION 4 (MEV4, (Saeed et al., 2003; Saeed et al.,

2006) on gene row SD normalized averaged RPKMs with median centering.

Enzyme Assays (Supplemental Figure 15)

From *G. gynandra* leaf stage 2 to 5, enzymatic activities of known C_4 enzymes were determinedas summarized by Ashton et al. (1990) in three biological replicates.

Comparison of Cleomaceae leaf gradients to *A. thaliana* leaf differentiation (Supplemental Figure 19)

Examination of Cleomaceae expression patterns of genes differentially regulated during the transition from cell proliferation to expansion in *A. thaliana*.
Andriankaja et al. (2012) observed that the transition between cell proliferation and expansion occurred between days 9 and 10. They defined two sets of genes significantly differentially expressed between day 9 and 10, one up-regulated and one down-regulated. The expression of the *T. hassleriana* and *G. gynandra* homologues of these genes were analyzed. The sum of standard error (SSE) was taken as a quality control to determine an appropriate number of clusters. The number of cluster centers chosen was 7 and 5 for up-regulated and down-regulated genes, respectively. The *K*-means clustering was performed the same as before, except that genes were not previously filtered by expression level and genes were only binned once into clusters.

Supplemental References

Andriankaja, M., Dhondt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., Muehlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., and Inze, D. (2012). Exit from Proliferation during Leaf Development in Arabidopsis thaliana: A Not-So-Gradual Process. Dev. Cell 22, 64-78.

Ashton A.R., Burnell J.N., Furbank R.T., Jenkins C.L.D., Hatch M.D. (1990). The enzymes in C4 photosynthesis. In Enzymes of Primary Metabolism. Methods in Plant Biochemistry, P.M. Dey and J.B. Harborne, eds (London: Academic Press), pp. 39–72

Cheng, S., van den Bergh, E., Zeng, P., Zhong, X., Xu, J., Liu, X., Hofberger, J., de Bruijn, S., Bhide, A.S., Kuelahoglu, C., Bian, C., Chen, J., Fan, G., Kaufmann, K., Hall, J.C., Becker, A., Braeutigam, A., Weber, A.P.M., Shi, C., Zheng, Z., Li, W., Lv, M., Tao, Y., Wang, J., Zou, H., Quan, Z., Hibberd, J.M., Zhang, G., Zhu, X.-G., Xu, X., and Schranz, M.E. (2013). The T arenaya hassleriana Genome Provides Insight into Reproductive Trait and Genome Evolution of Crucifers. Plant Cell **25,** 2813-2830.

McKown, A.D., and Dengler, N.G. (2009). Shifts in leaf vein density through accelerated vein formation in C-4 Flaveria (Asteraceae). Annals of Botany **104,** 1085-

1098.

Saeed, A.I., Hagabati, N.K., Braisted, J.C., Liang, W., Sharov, V., Howe, E.A., Li, J., Thiagarajan, M., White, J.A., and Quackenbush, J. (2006). TM4 microarray software suite. In DNA Microarrays, Part B: Databases and Statistics, A. Kimmel and B. Oluver, eds, pp. 134.

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003). TM4: A free, open-source system for microarray data management and analysis. Biotechniques **34,** 374.

Schliesky, S., Gowik, U., Weber, A.P.M., and Brautigam, A. (2012). RNA-Seq Assembly - Are We There Yet? Front Plant Sci **3,** 220-220.

Schulz, M.H., Zerbino, D.R., Vingron, M., and Birney, E. (2012). Oases: robust de novo RNA- seq assembly across the dynamic range of expression levels. Bioinformatics **28,** 1086- 1092.