

Supplementary Table S1. Genome assembly summary

	Primary*	Alternative*
HiFi Read coverage	Homozygous regions: 84, Heterozygous regions: 43	
Number of contigs	1,033	584
Contig N50 (bp)	17,751,673	15,995,289
Longest contigs	45,963,080	53,361,982
Number of scaffolds	970	497
Scaffold N50 (bp)	57,173,913	60,046,767
Size of final assembly (bp)	995,597,458	1,002,960,897
Base pair QV	50.004	53.3544
Base pair QV (both)		51.3697
k-mer completeness	79.8197%	80.0837%
k-mer completeness (both)	98.4765%	
BUSCO eukaryota_odb10 (255 genes)	C: 98.4% [S: 7.8%, D: 90.6%], F: 0.8%, M: 0.8%	C: 98.1% [S: 7.1%, D: 91.0%], F: 0.8%, M: 1.1%
BUSCO embryophyta_odb10 (1614 genes)	C: 99.1% [S: 11.0%, D: 88.1%], F: 0.2%, M: 0.7%	C: 99.3% [S: 12.3%, D: 87.0%], F: 0.1%, M: 0.6%
BUSCO fabales_odb10 (5366 genes)	C: 97.6% [S: 15.6%, D: 82.0%], F: 0.2%, M: 2.2%	C: 97.7% [S: 16.2%, D: 81.5%], F: 0.2%, M: 2.1%
Assigned chromosomes	94.529%	95.546%
# N's	26800	27500
# N's per 100 kbp	2.69	2.74
GC level	33.91 %	33.98 %
Repeated elements	60.51 %	60.18 %
Retroelements	29.07 %	30.40 %
LTR elements (Ty1/Copia)	13.32 %	13.58 %
LTR elements (Gypsy/DIRS1)	7.69 %	9.73 %
DNA transposons	1.11 %	1.38 %

*Representing the two haploid genomes of a diploid organism

Supplementary Table S2. Total alignment length (identity >95%, query length >10,000 bp, hit length >10,000 bp)

Alternative haplotig	25-32 (tp)	3,811,295,141	6,158,361,485
	17-24 (to)	5,817,545,634	3,764,652,819
		1-8 (to)	9-16 (tp)
Primary haplotig			

Supplementary Table S3. SNPs detected by direct Sanger sequencing of PCR products of Ac and Li genes in parents (bold) and F1 progeny (two-letter codes) in greenhouse crossing experiments					
Parental Accessions (bold) and F1 progeny (two letter codes)	Li phenotype	GenBank accession numbers	Li gene copy inferences from sequences (SNPs correspond to SNP locations in NCBI Popset alignment)	Estimated Li gene copies total (both alleles)	
PI 239977 PI 440745	Li+ li-	EF990455; PopSet: 1574 N/A	Model: 2 alleles, both with 2 (or more) gene copies differ N/A	4 (or more) 0	
WU A01	Li+		G/A @ 3540	2 (or more)	
WU A02	Li+		G/A @ 3540	2 (or more)	
WU A06	Li+		G/A @ 3540	2 (or more)	
WU A07	Li+		G/A @ 3540	2 (or more)	
WU A08	Li+		G/A @ 3540	2 (or more)	
WU B01	Li+		G/A @ 3540	2 (or more)	
WU B02	Li+		G/A @ 3540	2 (or more)	
WU B03	Li+		G/A @ 3540	2 (or more)	
WU B04	Li+		G/A @ 3540	2 (or more)	
WU B05	Li+		G/A @ 3540	2 (or more)	
WU B06	Li+		G/A @ 3540	2 (or more)	
WU B07	Li+		G/A @ 3540	2 (or more)	
WU B08	Li+		G/A @ 3540	2 (or more)	
WU B09	Li+		G/A @ 3540	2 (or more)	
WU B10	Li+		G/A @ 3540	2 (or more)	
WU B11	Li+		G/A @ 3540	2 (or more)	
PI 205062 PI 311494	li- Li+	N/A EF990459, EF990460; PopSet: 157416156	N/A Model: 2 alleles, at least one of which has 2 (or more) gene copies, different @ sites 3489	0 3 (or more)	
WU C02	Li+		A @ 583, A @ 3540	1 (or more)	
WU C03	Li+		A @ 583, G/A @ 3540	2 (or more)	
WU C04	Li+		A @ 583, A @ 3540	1 (or more)	
WU C06	Li+		A @ 583, A @ 3540	1 (or more)	
WU D01	Li+		A @ 583, A @ 3540	1 (or more)	
WU D02	Li+		A @ 583, A @ 3540	1 (or more)	
WU D03	Li+		A @ 583, A @ 3540	1 (or more)	
WU D04	Li+		T @ 583, G/A @ 3540	2 (or more)	
WU D05	Li+		A @ 583, A @ 3540	1 (or more)	
WU D06	Li+		T @ 583, G/A @ 3540	2 (or more)	
WU D07	Li+		T @ 583, G/A @ 3540	2 (or more)	
WU D08	Li+		T @ 583, G/A @ 3540	2 (or more)	
WU D09	Li+		A @ 583, A @ 3540	1 (or more)	
G = STL 2406 H = STL 1904 I = STL 2604	li- Li+ Li+	N/A PopSet: 157416156 PopSet: 157416156	N/A Model: hemizygote, 1 gene copy for allele with gene (G @ 3540) Model: hemizygote, 2 gene copies, different @ sites 3489	0 1 2	
GI F101	Li+		Identical to parent I	2 (Same as parent I)	
GI F102	Li+		Identical to parent I	2 (Same as parent I)	
GI F103	Li+		Identical to parent I	2 (Same as parent I)	
GI F104	li-		N/A	0	
GI F105	Li+		Identical to parent I	2 (Same as parent I)	
GI F106	Li+		Identical to parent I	2 (Same as parent I)	
GI F107	Li+		Identical to parent I	2 (Same as parent I)	
GI F108	Li+		Identical to parent I	2 (Same as parent I)	
IG F101	Li+		Identical to parent I	2 (Same as parent I)	
IG F102	Li+		Identical to parent I	2 (Same as parent I)	
IG F103	Li+		Identical to parent I	2 (Same as parent I)	
IG F104	Li+		Identical to parent I	2 (Same as parent I)	
IG F105	Li+		Identical to parent I	2 (Same as parent I)	
IG F106	li-		N/A	0	
IG F107	li-		N/A	0	
IG F108	Li+		Identical to parent I	2 (Same as parent I)	
IG F109	li-		N/A	0	
IG F110	li-		N/A	0	
HG F101	Li+		Identical to parent H	1 (Same as parent H)	
HG F102	Li+		Identical to parent H	1 (Same as parent H)	
HG F103	li-		N/A	0	
HG F104	li-		N/A	0	
HG F106	Li+		Identical to parent H	1 (Same as parent H)	
HG F107	Li+		Identical to parent H	1 (Same as parent H)	
HG F108	li-		N/A	0	
HG F109	li-		N/A	0	
J = WL 1007 K = MSJ 0710 L = LA 0410	Li+ Li+ Li+	PopSet: 157416156 PopSet: 157416156 PopSet: 157416156	Model: hemizygote, 2 gene copies, differ @ 3540 (G/A); s2 Model: hemizygote, 2 gene copies, differ @ 3540 (G/A); s2 Model: 2 alleles, both with 2 gene copies: all are T @ 703	2 2 4	
JK 101	li-		N/A	0	
JK 102	Li+		Identical to K	2 (Same as K)	
JK 103	Li+		J + K alleles	4 (J + K)	
JK 104	Li+		Identical to K	2 (Same as K)	
JK 105	Li+		Identical to K	2 (Same as K)	
JK 106	li-		N/A	0	
JK 107	Li+		Identical to K	2 (Same as K)	
JK 108	li-		N/A	0	
JK 109	li-		N/A	0	
JK 110	Li+		J + K alleles	4 (J + K)	
JK 111	li-		N/A	0	
JK 112	li-		N/A	0	
JK 113	Li+		Identical to K	2 (Same as K)	
JK 114	Li+		J + K alleles	4 (J + K)	
JK 115	Li+		Identical to K	2 (Same as K)	
JK 116	Li+		J + K alleles	4 (J + K)	
JK 117	li-		N/A	0	
JK 118	Li+		Identical to K	2 (Same as K)	
JK 120	li-		N/A	0	
JK 121	Li+		Identical to K	2 (Same as K)	
JK 122	Li+		Identical to K	2 (Same as K)	
JK 123	Li+		J + K alleles	4 (J + K)	
JK 124	Li+		J + K alleles	4 (J + K)	
JK 125	li-		N/A	0	
JK 126	Li+		Identical to K	2 (Same as K)	
JK 127	Li+		G/A @ 3540	(incomplete genotyping)	
JK 128	Li+		J + K alleles	4 (J + K)	
JK 129	li-		N/A	0	
JK 130	li-		N/A	0	
JK 131	Li+		Identical to J	2 (Same as J)	
JK 133	Li+		Identical to J	2 (Same as J)	
KJ 202	Li+		Identical to K	2 (Same as K)	
KJ 203	Li+		Identical to K	2 (Same as K)	
KJ 204	Li+		Identical to J	2 (Same as J)	
KJ 205	Li+		J + K alleles	4 (J + K)	
KJ 206	Li+		Identical to J	2 (Same as J)	
KJ 207	Li+		J + K alleles	4 (J + K)	
KJ 208	Li+		Identical to J	2 (Same as J)	
KJ 209	Li+		Identical to J	2 (Same as J)	
KJ 210	Li+		Identical to J	2 (Same as J)	
KJ 211	Li+		J + K alleles	4 (J + K)	
LJ 101	Li+		La + J	4	
LJ 102	Li+		Lb + J	4	
LJ 103	Li+		Lb	2	
LJ 104	Li+		La + J	4	
LJ 105	Li+		La	2	
LJ 106	Li+		Lb + J	4	
LJ 107	Li+		Lb	2	
LJ 201	Li+		Lb + J	4	
LJ 202	Li+		La + J	4	
LJ 203	Li+		Lb + J	4	
LJ 204	Li+		Lb	2	
LJ 205	Li+		La	2	
Q = TN 0506 R = MSG 1108	Li+ Li+	PopSet: 157416156 PopSet: 157416156	A @ 3271, C @ 3489, G @ 3540, C @ 3577 — hemizygote with two gene copies A/T @ 3271, A/C @ 3489, A @ 3540, C/T @ 3577 — hemizygote with two gene copies	2 (same as Q) 2 (same as Q)	
QR 101	li-		N/A	0	
QR 102	li-		N/A	0	
QR 103	li-		N/A	0	
QR 104	Li+		Identical to Q	2 (same as Q)	
QR 105	Li+		Identical to R	2 (same as R)	
RQ 201	Li+		Identical to R	2 (same as R)	
RQ 202	Li+		Q+R	4 (Q+R)	
RQ 203	Li+		Identical to R	2 (same as R)	
RQ 204	Li+		Q+R	4 (Q+R)	
RQ 205	li-		N/A	0	
RQ 206	li-		N/A	0	
RQ 207	Li+		Identical to Q	2 (same as Q)	
RQ 208	li-		N/A	0	
RQ 209	Li+		Identical to R	2 (same as R)	
RQ 210	Li+		Identical to Q	2 (same as Q)	
RQ 211	li-		N/A	0	
V = STL_HG_F101 X = MSG_1106	Li+ Li+	PopSet: 157416156 PopSet: 157416156	Identical to H: Model: hemizygote, 1 gene copy for allele 2 alleles, 2 gene copies apiece (based on TX crosses, where 2 per allele)	1 (Same as H) 4	
VX 101	Li+		V+Xb	3	
VX 102	Li+		V+Xb	3	
VX 103	Li+		V+Xb	3	
VX 104	Li+		Xb alone	2 (Xb)	
VX 105	Li+		Xb alone	2 (Xb)	
VX 106	Li+		Xa	2 (Xa)	
VX 107	Li+		Xa	2 (Xa)	
VX 108	Li+		V+Xb	3	
VX 109	Li+		Xa+V	3	
VX 110	Li+		Xa	2 (Xa)	
VX 112	Li+		Xa+V	3	
VX 113	Li+		Xa+V	3	
VX 114	Li+		Xa	2 (Xa)	
VX 115	Li+		V+Xb	3	
VX 116	Li+		Xa	2 (Xa)	
VX 117	Li+		Xa	2 (Xa)	
VX 118	Li+		Xa	2 (Xa)	
VX 119	Li+		Xa	2 (Xa)	
VX 120	Li+		V+Xb	3	
VX 121	Li+		V+Xb	3	
VX 122	Li+		Xa+V	3	
VX 123	Li+		Xa+V	3	
XV 201	Li+		Xa	2 (Xa)	
XV 202	Li+		V+Xb	3	
XV 203	Li+		V+Xb	3	
XV 204	Li+		Xa	2 (Xa)	
XV 205	Li+		V+Xb	3	
XV 206	Li+		Xb alone	2 (Xb)	
XV 207	Li+		Xa+V	3	
XV 208	Li+		Xa	2 (Xa)	
XV 209	Li+		Xa	2 (Xa)	
XV 210	Li+		V+Xb	3	
XV 211	Li+		Xa	2 (Xa)	
XV 212	Li+		V+Xb	3	
XV 213	Li+		Xa	2 (Xa)	
XV 214	Li+		Xb alone	2 (Xb)	
XV 215	Li+		Xa+V	3	
XV 216	Li+		V+Xb	3	
XV 217	Li+		Xa+V	3	
XV 218	Li+		V+Xb	3	
XV 219	Li+		Xb alone	2 (Xb)	
XV 220	Li+		Xb alone	2 (Xb)	
TX 101	Li+		Xa	2 (Xa)	
TX 102	Li+		Xb alone	2 (Xb)	
TX 103	Li+		Xb alone	2 (Xb)	
TX 104	Li+		Xa	2 (Xa)	
TX 105	Li+		Xa	2 (Xa)	
TX 106	Li+		Xb alone	2 (Xb)	
TX 107	Li+		Xb alone	2 (Xb)	
TX 108	Li+		Xb alone	2 (Xb)	
TX 109	Li+		Xa	2 (Xa)	
TX 110	Li+		Xa	2 (Xa)	
TX 111	Li+		Xb alone	2 (Xb)	
TX 112	Li+		Xa	2 (Xa)	
TX 113	Li+		Xb alone	2 (Xb)	
TX 114	Li+		Xb alone	2 (Xb)	
XT 201	Li+		Xa	2 (Xa)	
XT 202	Li+		Xb alone	2 (Xb)	
XT 203	Li+		Xa	2 (Xa)	
XT 204	Li+		Xb alone	2 (Xb)	
RT 101	Li+		Same as R	2 (same as R)	
RT 102	Li+		Same as R	2 (same as R)	
RT 103	Li+		N/A	0	
TR 201	Li+		Same as R	2 (same as R)	
TR 202	Li+		Same as R	2 (same as R)	
TR 203	Li+		Same as R	2 (same as R)	
Accession PI 239977 PI 440745	Ac+ Ac+ Ac+	EU370182, EU370183; PopSet: 1705 EU370196; PopSet: 170516368	Model: hemizygote with two gene copies, differing at 2 sites Model: hemizygote with one gene copy, characterized by 440745+239977	2 1	
WU A01	Ac+		440745+239977	3	
WU A02	Ac+		239977 alone	2 (Same as 239977)	
WU A06	Ac+		239977 alone	2 (Same as 239977)	
WU A07	Ac+		440745+239977	3	
WU A08	ac-		N/A	0	
WU B01	Ac+		239977 alone	2 (Same as 239977)	
WU B02	Ac+		440745 alone	1 (Same as 440745)	
WU B03	Ac+		239977 alone	2 (Same as 239977)	
WU B04	Ac+		440745+239977	3	
WU B05	Ac+		239977 alone	2 (Same as 239977)	
WU B06	ac-		N/A	0	
WU B07	Ac+		239977 alone	2 (Same as 239977)	
WU B08	Ac+		440745 alone	1 (Same as 440745)	
WU B09	Ac+		440745 alone	1 (Same as 440745)	
WU B10	Ac+		239977 alone	2 (Same as 239977)	
WU B11	ac-		N/A	0	
PI 205062 PI 311494	Ac+ Ac+	EU370170, EU370171; PopSet: 1705 EU370191, EU370192; PopSet: 1705	Model: hemizygote with one gene copy, characterized by 239977 alone Model: 2 alleles, one with 2 copies (=a' allele), the other with 1 copy (=a' allele)	1 4 or 5	
WU C02	Ac+		205062 + 311494a	3	
WU C03	Ac+		205062 + 311494a	3	
WU C04	Ac+		311494a alone	2	
WU C06	Ac+		205062 + 311494a	3	
WU D01	Ac+		311494a alone	2	
WU D02	Ac+		311494b alone	2 or 3	
WU D03	Ac+		205062 + 311494a	3	
WU D04	Ac+		(didn't sequence)	(didn't sequence)	
WU D05	Ac+		311494b alone	2 or 3	
WU D06	Ac+		311494a alone	2	
WU D07	Ac+		311494a alone	2	
WU D08	Ac+		205062 + 311494b	3 or 4	
WU D09	Ac+		311494a alone	2	
G = STL 2406 H = STL 1904 I = STL 2604	ac- Ac+ Ac+	N/A PopSet: 170516368 PopSet: 170516368	N/A Model: 2 alleles (Ha, Hb), 1 gene copy per allele, diff @ sites 497 Model: hemizygote, 2 gene copies, differ @ sites 497	N/A 2 2	
GI F101	ac-		N/A	0	
GI F102	Ac+		Identical to I	2 (Same as I)	
GI F103	Ac+		Identical to I	2 (Same as I)	
GI F104	Ac+		N/A	0	
GI F105	Ac+		Identical to I	2 (Same as I)	
GI F106	ac-		N/A	0	
GI F1					

Supplementary Table S4. Genomic locations of cyanogenesis genes

Gene	Primary (this study)	Alternative (this study)	Santangelo et. al., 2023	Wang et. al., 2023
<i>CYP79D15</i>	drTriRepe4Chr2g112600.1 drTriRepe4Chr2g115700.1	drTriRepe4Chr18g122900.1 drTriRepe4Chr18g125500.1 drTriRepe4Chr18g128300.1	TrR_v6_g12417.t1 TrR_v6_g12456.t1	Chr10.g56734.m1 Chr10.g56705.m1
<i>CYP736A187</i>	drTriRepe4Chr2g111200.1 drTriRepe4Chr2g114400.1	drTriRepe4Chr18g121900.1 drTriRepe4Chr18g124100.1 drTriRepe4Chr18g127300.1	TrR_v6_g12398.t1 TrR_v6_g12430.t1	Chr10.g56752.m1 Chr10.g56786.m1 Chr10.g56724.m1
<i>UGT85K17</i>	drTriRepe4Chr2g110400.1 drTriRepe4Chr2g113800.1	drTriRepe4Chr18g121500.1 drTriRepe4Chr18g123600.1 drTriRepe4Chr18g126400.1	TrR_v6_g12398.t1 TrR_v6_g12430.t1	Chr10.g56793.m1 Chr10.g56732.m1 Chr10.g56760.m1
<i>Li</i>	drTriRepe4Chr12g338400.1 drTriRepe4Chr12g338600.1	drTriRepe4Chr28g340600.1 drTriRepe4Chr28g340900.1	TrR_v6_g38299.t1 TrR_v6_g38301.t1	Chr07.g39283.m2 Chr07.g39287.m1

Supplementary Table S5. Genomic locations of *Ac* gene cluster

Gene	chromosome	Transcript	Start	End	Strand
<i>UGT85K17</i>	drTriRepe4Chr2	drTriRepe4Chr2g110400.1	8,367,129	8,369,048	-
<i>CYP736A187</i>	drTriRepe4Chr2	drTriRepe4Chr2g111200.1	8,423,330	8,425,360	+
<i>CYP79D15</i>	drTriRepe4Chr2	drTriRepe4Chr2g112600.1	8,538,499	8,540,592	+
<i>UGT85K17</i>	drTriRepe4Chr2	drTriRepe4Chr2g113800.1	8,631,125	8,633,208	-
<i>CYP736A187</i>	drTriRepe4Chr2	drTriRepe4Chr2g114400.1	8,743,594	8,745,624	+
<i>CYP79D15</i>	drTriRepe4Chr2	drTriRepe4Chr2g115700.1	8,851,286	8,857,454	+
<i>UGT85K17</i>	drTriRepe4Chr18	drTriRepe4Chr18g121500.1	8,809,396	8,812,326	-
<i>CYP736A187</i>	drTriRepe4Chr18	drTriRepe4Chr18g121900.1	8,865,280	8,867,310	+
<i>CYP79D15</i>	drTriRepe4Chr18	drTriRepe4Chr18g122900.1	8,974,416	8,976,497	+
<i>UGT85K17</i>	drTriRepe4Chr18	drTriRepe4Chr18g123600.1	9,081,205	9,083,126	-
<i>CYP736A187</i>	drTriRepe4Chr18	drTriRepe4Chr18g124100.1	9,137,923	9,139,953	+
<i>CYP79D15</i>	drTriRepe4Chr18	drTriRepe4Chr18g125500.1	9,248,143	9,250,216	+
<i>UGT85K17</i>	drTriRepe4Chr18	drTriRepe4Chr18g126400.1	9,336,113	9,338,193	-
<i>CYP736A187</i>	drTriRepe4Chr18	drTriRepe4Chr18g127300.1	9,422,195	9,424,222	+
<i>CYP79D15</i>	drTriRepe4Chr18	drTriRepe4Chr18g128300.1	9,530,323	9,536,398	+

Supplementary Table S6. Genomic locations of *Li* genes

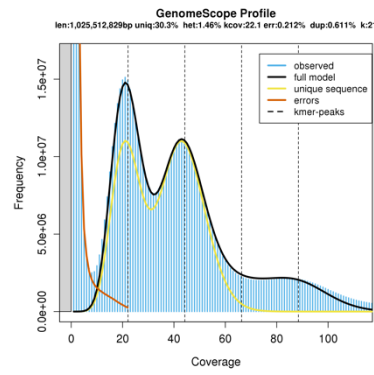
Gene	chromosome	Transcript	Start	End	Strand
<i>Li</i>	drTriRepe4Chr12	drTriRepe4Chr12g338400.1	28,886,861	28,890,925	-
<i>Li</i>	drTriRepe4Chr12	drTriRepe4Chr12g338600.1	28,910,546	28,915,335	-
<i>Li</i>	drTriRepe4Chr28	drTriRepe4Chr28g340600.1	28,844,916	28,849,166	-
<i>Li</i>	drTriRepe4Chr28	drTriRepe4Chr28g340900.1	28,869,253	28,875,856	-

Supplementary Table S7. Collection information and cyanogenesis gene CNVs for wild population samples									
Accession name	Collection Date (Year-Month)	Population abbreviation	Latitude	Longitude	Nearest City	Collector	Ac_CNV	Li_CNV	
AIA_001	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	2	2	
AIA_002	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	0	0	
AIA_003	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	0	0	
AIA_004	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	0	0	
AIA_007	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	4	0	
AIA_008	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	3	0	
AIA_009	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	0	0	
AIA_010	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	3	0	
AIA_011	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	0	1	
AIA_012	2015-10	AIA	42.046667	-93.621111	Ames, IA	Sara Wright	0	0	
ANC_001	2015-07	ANC	35.568944	-82.570833	Asheville, NC	Rebecca Thomson	5	2	
ANC_002	2015-07	ANC	35.568333	-82.568333	Asheville, NC	Rebecca Thomson	2	0	
ANC_004	2015-07	ANC	35.578056	-82.581667	Asheville, NC	Rebecca Thomson	0	0	
ANC_005	2015-07	ANC	35.578056	-82.581944	Asheville, NC	Rebecca Thomson	1	0	
ANC_006	2015-07	ANC	35.577778	-82.561389	Asheville, NC	Rebecca Thomson	1	0	
ANC_007	2015-07	ANC	35.57775	-82.55422	Asheville, NC	Rebecca Thomson	0	0	
ANC_008	2015-07	ANC	35.566389	-82.573333	Asheville, NC	Rebecca Thomson	0	0	
ATX_001	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	4	3	
ATX_002	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	4	4	
ATX_004	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	4	2	
ATX_005	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	5	2	
ATX_006	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	2	2	
ATX_007	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	5	4	
ATX_008	2017-04	ATX	30.285951	-97.734002	Austin, TX	Sara Wright	3	4	
ATX_009	2017-04	ATX	30.285706	-97.730922	Austin, TX	Sara Wright	3	4	
ATX_015	2017-04	ATX	30.285175	-97.729543	Austin, TX	Sara Wright	3	4	
ATX_016	2017-04	ATX	30.285175	-97.729543	Austin, TX	Sara Wright	1	3	
BAL_005	2015-07	BAL	34.056278	-86.762583	Hanceville, AL	Nicole Riddle	4	0	
BAL_013	2015-07	BAL	34.0725	-86.781944	Hanceville, AL	Nicole Riddle	0	4	
BAL_018	2015-07	BAL	33.501772	-86.7741557	Homewood, AL	Nicole Riddle	0	0	
BAL_019	2015-07	BAL	33.4760152	-86.8197731	Birmingham, AL	Nicole Riddle	0	0	
BAL_020	2015-07	BAL	33.487966	-86.817766	Birmingham, AL	Nicole Riddle	5	0	
BAL_021	2015-07	BAL	33.4931613	-86.7961006	Birmingham, AL	Nicole Riddle	3	2	
BAL_024	2015-07	BAL	33.438553	-86.837503	Birmingham, AL	Nicole Riddle	2	2	
BAL_027	2015-07	BAL	33.51721	-86.766378	Birmingham, AL	Nicole Riddle	0	0	
BAL_031	2015-07	BAL	33.500044	-86.807627	Birmingham, AL	Nicole Riddle	2	2	
BAL_043	2015-07	BAL	33.555556	-86.628056	Birmingham, AL	Nicole Riddle	5	0	
BID_001	2014-09	BID	43.596111	-116.222222	Boise, ID	Sara Wright	0	0	
BID_002	2014-09	BID	43.596111	-116.222222	Boise, ID	Sara Wright	0	1	
BID_003	2014-09	BID	43.596111	-116.222222	Boise, ID	Sara Wright	0	1	
BID_004	2014-09	BID	43.596111	-116.222222	Boise, ID	Sara Wright	5	0	
BID_005	2014-09	BID	43.596111	-116.222222	Boise, ID	Sara Wright	0	0	
BID_006-2	2014-09	BID	43.596111	-116.222222	Boise, ID	Sara Wright	0	0	
BID_007-2	2014-09	BID	43.595556	-116.222222	Boise, ID	Sara Wright	0	0	
BID_008-2	2014-09	BID	43.595556	-116.222222	Boise, ID	Sara Wright	0	0	
BID_010	2014-09	BID	43.595556	-116.222222	Boise, ID	Sara Wright	0	1	
BID_011	2014-09	BID	43.595556	-116.222222	Boise, ID	Sara Wright	0	0	
BIL_002	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	2	0	
BIL_003	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	0	0	
BIL_004	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	0	0	
BIL_005	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	4	2	
BIL_007	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	5	0	
BIL_008	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	0	1	
BIL_009	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	0	0	
BIL_012	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	0	0	
BIL_016	2015-08	BIL	40.517778	-88.953056	Bloomington-Normal	Sara Wright	0	0	
BKS_001	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	2	0	
BKS_003	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	0	0	
BKS_004	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	5	1	
BKS_005	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	2	0	
BKS_006	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	3	0	
BKS_007	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	0	0	
BKS_008	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	3	0	
BKS_009	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	1	0	
BKS_010	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	0	0	
BKS_012	2015-11	BKS	39.088611	-94.880556	Bonner Springs, KS	Sara Wright	3	1	
BMD_001	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	0	0	
BMD_002	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	0	0	
BMD_003	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	0	0	
BMD_005	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	4	2	
BMD_006	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	0	0	
BMD_007	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	3	0	
BMD_008	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	0	0	
BMD_009	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	0	1	
BMD_011	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	3	1	
BMD_012	2014-09	BMD	39.3313	-76.58768	Baltimore, MD	John McDonald	2	0	
BSD_001	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	4	0	
BSD_002	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	2	0	
BSD_003	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	3	0	
BSD_004	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	2	0	
BSD_005	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	3	0	
BSD_006	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	2	0	
BSD_007	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	0	0	
BSD_008	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	0	0	
BSD_009	2014-09	BSD	43.853611	-101.203611	Belvidere, SD	Sara Wright	2	0	
BWA_001	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	3	1	
BWA_002	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	0	2	
BWA_005	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	2	1	
BWA_006	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	0	1	
BWA_007	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	2	0	
BWA_008	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	4	0	
BWA_009	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	4	0	
BWA_010	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	2	2	
BWA_011	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	2	1	
BWA_012	2015-06	BWA	48.734	-122.486	Bellingham, WA	Jim Mullen	2	1	
CL_001	2014-08	CIL	41.962222	-87.633333	Chicago, IL	Sara Wright	0	0	
CL_003-2	2014-08	CIL	41.963056	-87.634444	Chicago, IL	Sara Wright	2	1	
CL_004	2014-08	CIL	41.963056	-87.635278	Chicago, IL	Sara Wright	0	4	
CL_005	2014-08	CIL	41.963056	-87.632778	Chicago, IL	Sara Wright	0	3	
CL_006	2014-08	CIL	41.963333	-87.631667	Chicago, IL	Sara Wright	3	0	
CL_007	2014-08	CIL	41.963333	-87.634167	Chicago, IL	Sara Wright	0	0	
CL_008	2014-08	CIL	41.963333	-87.634167	Chicago, IL	Sara Wright	3	2	
CL_009	2014-08	CIL	41.963611	-87.631389	Chicago, IL	Sara Wright	1	0	
CL_010	2014-08	CIL	41.963889	-87.632778	Chicago, IL	Sara Wright	4	1	
CL_011	2014-08	CIL	41.963889	-87.636611	Chicago, IL	Sara Wright	4	2	
CMO_004	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	4	0	
CMO_005	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	5	0	
CMO_006	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	2	2	
CMO_007	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	1	2	
CMO_008	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	2	0	
CMO_010	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	0	0	
CMO_012	2015-11	CMO	38.962766	-92.25064	Columbia, MO	Sara Wright	4	0	
CMO_013	2016	CMO	38.9379	-92.295	Columbia, MO	Felix Fritschi	5	0	
CMO_014	2016	CMO	38.8397	-92.295	Columbia, MO	Felix Fritschi	0	0	
CMO_015	2016	CMO	38.9395	-92.2954	Columbia, MO	Felix Fritschi	2	0	
CSC_001	2015-06	CSC	32.795278	-80.108889	Charleston, SC	Sandra Pelc	4	2	
CSC_005	2015-06	CSC	32.7925	-80.1125	Charleston, SC	Sandra Pelc	4	0	
CSC_006	2015-06	CSC	32.795278	-80.080278	Charleston, SC	Sandra Pelc	4	3	
CSC_007	2015-06	CSC	32.795556	-80.081944	Charleston, SC	Sandra Pelc	4	3	
CSC_008	2015-06	CSC	32.796111	-80.109444	Charleston, SC	Sandra Pelc	6	2	
CSC_009	2015-06	CSC	32.793056	-80.108056	Charleston, SC	Sandra Pelc	1	3	
CSC_010	2015-06	CSC	32.793611	-80.109722	Charleston, SC	Sandra Pelc	1	3	
CSC_011	2015-06	CSC	32.793333	-80.119244	Charleston, SC	Sandra Pelc	6	2	
CSC_012	2015-06	CSC	32.795833	-80.0825	Charleston, SC	Sandra Pelc	5	4	
CSC_014	2015-06	CSC	32.795278	-80.081944	Charleston, SC	Sandra Pelc	5	2	
CVA_001	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	2	2	
CVA_002	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	0	2	
CVA_003	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	0	0	
CVA_004	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	0	0	
CVA_005	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	5	0	
CVA_006	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	0	0	
CVA_007	2014-09	CVA	38.010242	-78.517095	Charlottesville, VA	Nic Kooyers	0	1	
CVA_010	2014-09	CVA	38.010242						

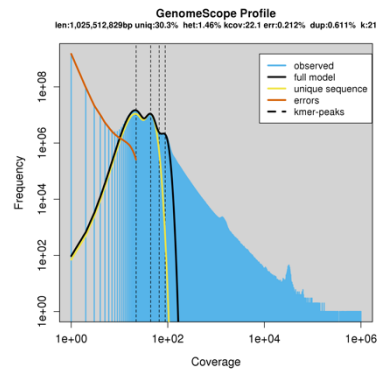
Supplementary Table S8. Primer list

Name	Sequence	Amplicon length	Note
CYP79D15_q03Fa (exon2) CYP79D15_q03R (exon2)	CAA TGT GGT TGG CAA AGA GAG GCG CAA GCT CTA ACA AAT GG	72	For CNV and gene expression
CYP736_q05F (exon1) CYP736_q05R (exon2)	GCT TGT TCA TGA AGT CCT GC AAA TTC TTT GTG CGC TTT CTT G	61+39	For CNV and gene expression
CYP736_q06F (exon1) CYP736_q05R (exon2)	ATG AAG TCC TGC ACC TCA TG AAA TTC TTT GTG CGC TTT CTT G	53+39	For CNV and gene expression
CYP736_q02F (intron) CYP736_q01R (intron)	ACG TAC GCA ATA AAC TTC TTT CC TGT CCT TGG TAT AAG GTC TTT CAC A	135	For CNV
UGT85_q02F (exon1) UGT85_q02R (exon2)	GTG GCT TCG TCG GAT ATT TG GTT GTA TCC AAG GTG CCA TC	63+37	For CNV and gene expression
UGT85_q03F (exon1) UGT85_q02R (exon2)	TGC AAT TCG ATG AAC TTG CC GTT GTA TCC AAG GTG CCA TC	45+37	For CNV and gene expression
UGT85_q01F (intron) UGT85_q01R (exon2)	TTG CTC AAT CTA TTT CGG ATG TC GTC GCA TAT TTT TCA AGC CAG AG	110	For CNV
Li_q04Fa (intron) Li_q04Rb (exon3)	GCA AGT TCC ATT CAT TTT ATT TTT AGT T CTT GTA CCG ATG ATA TTC GTC AAT	183	For CNV
Ef1-alpha_01F Ef1-alpha_01R	TCG AGA AGG AAG CTG CTG AAA CCC AGG CAT ACT TGA ATG ACC T		[1]

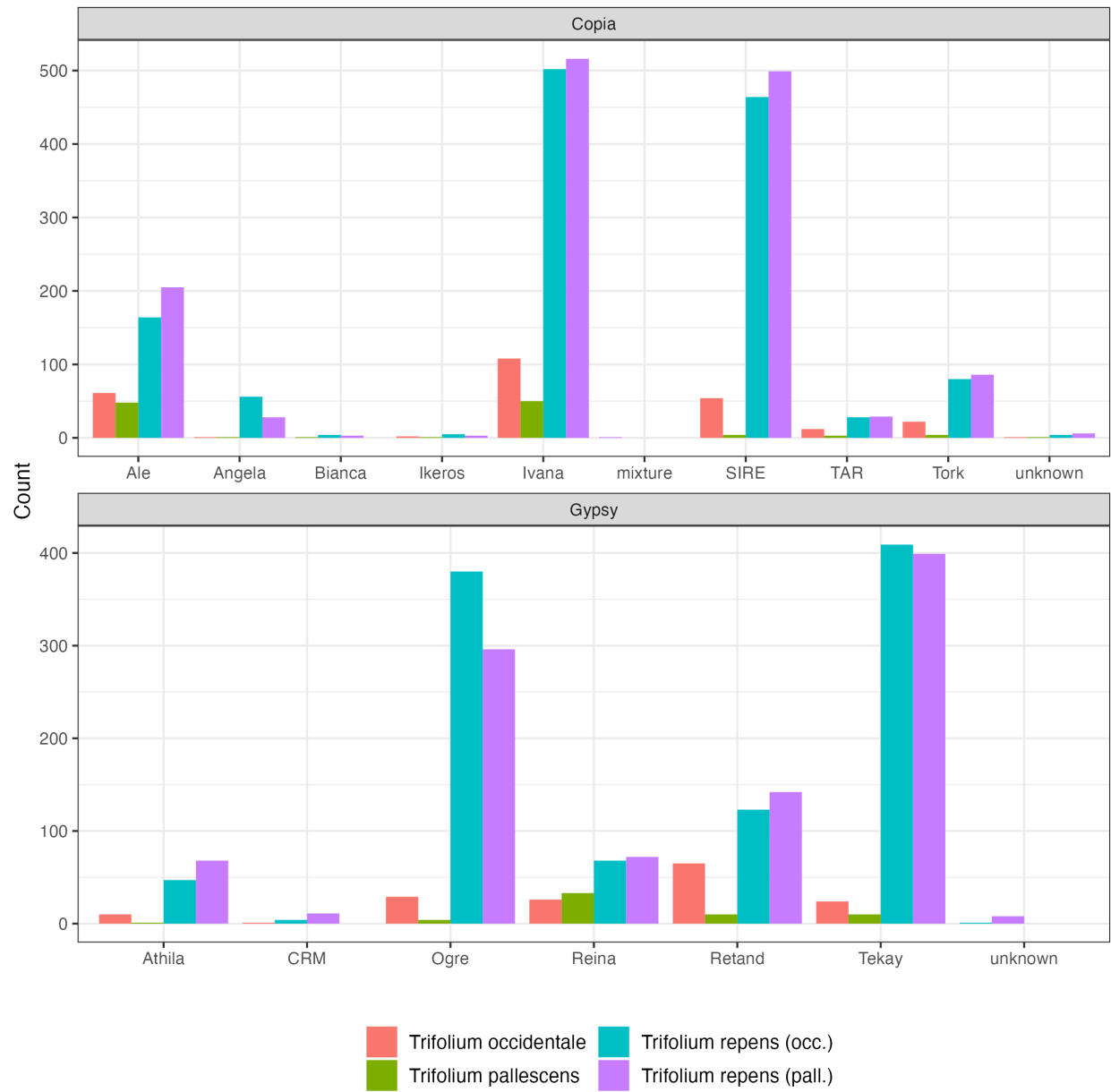
(A)



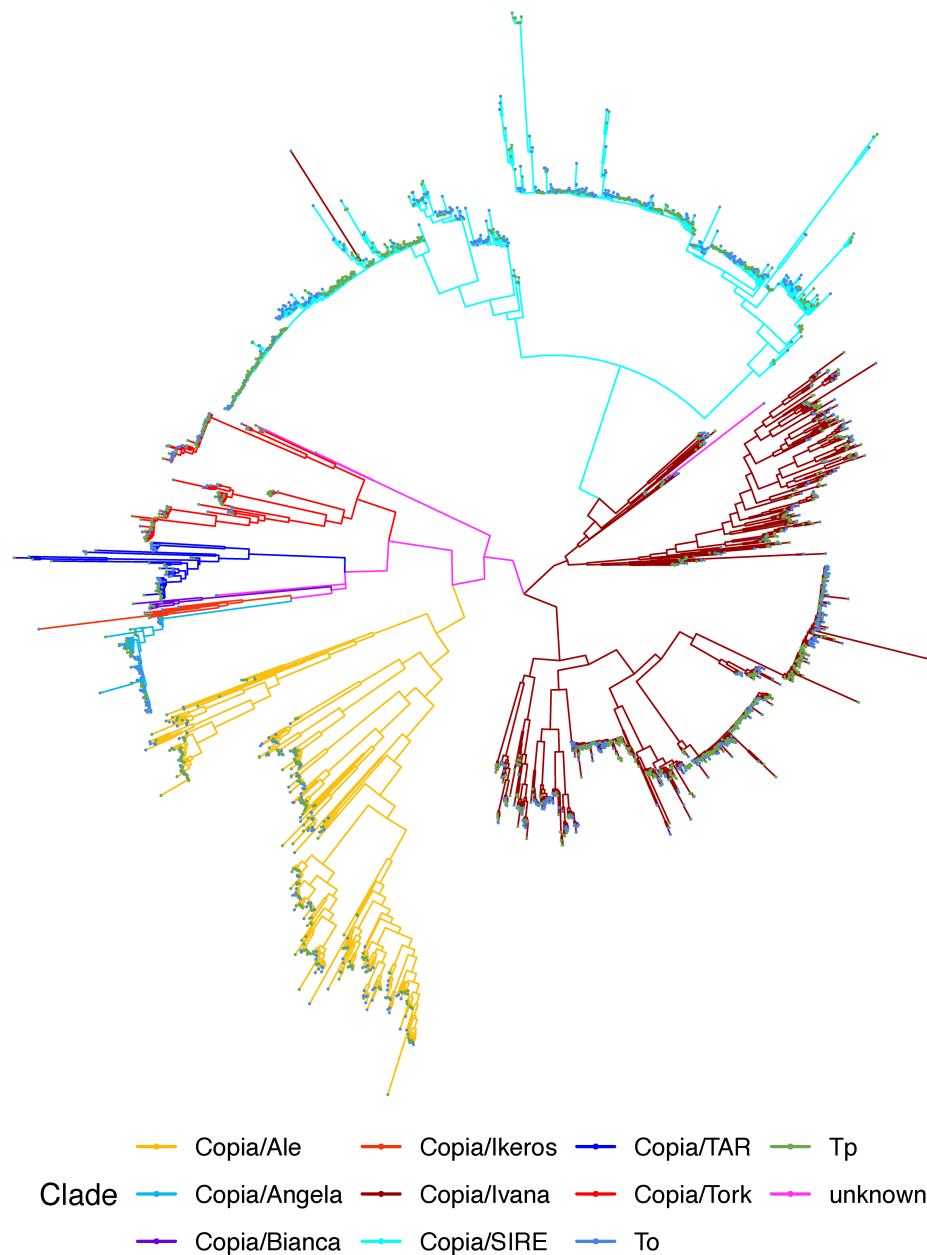
(B)



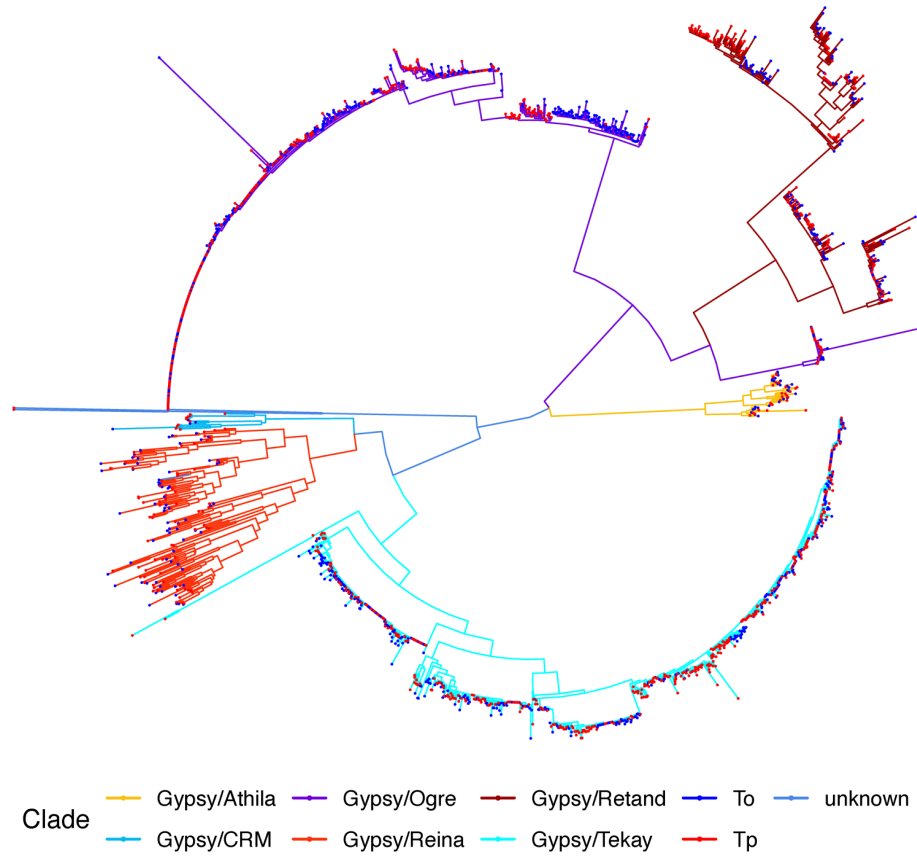
Supplementary Fig. S1. (A-B), GenomeScope genome size and heterozygosity assessment.



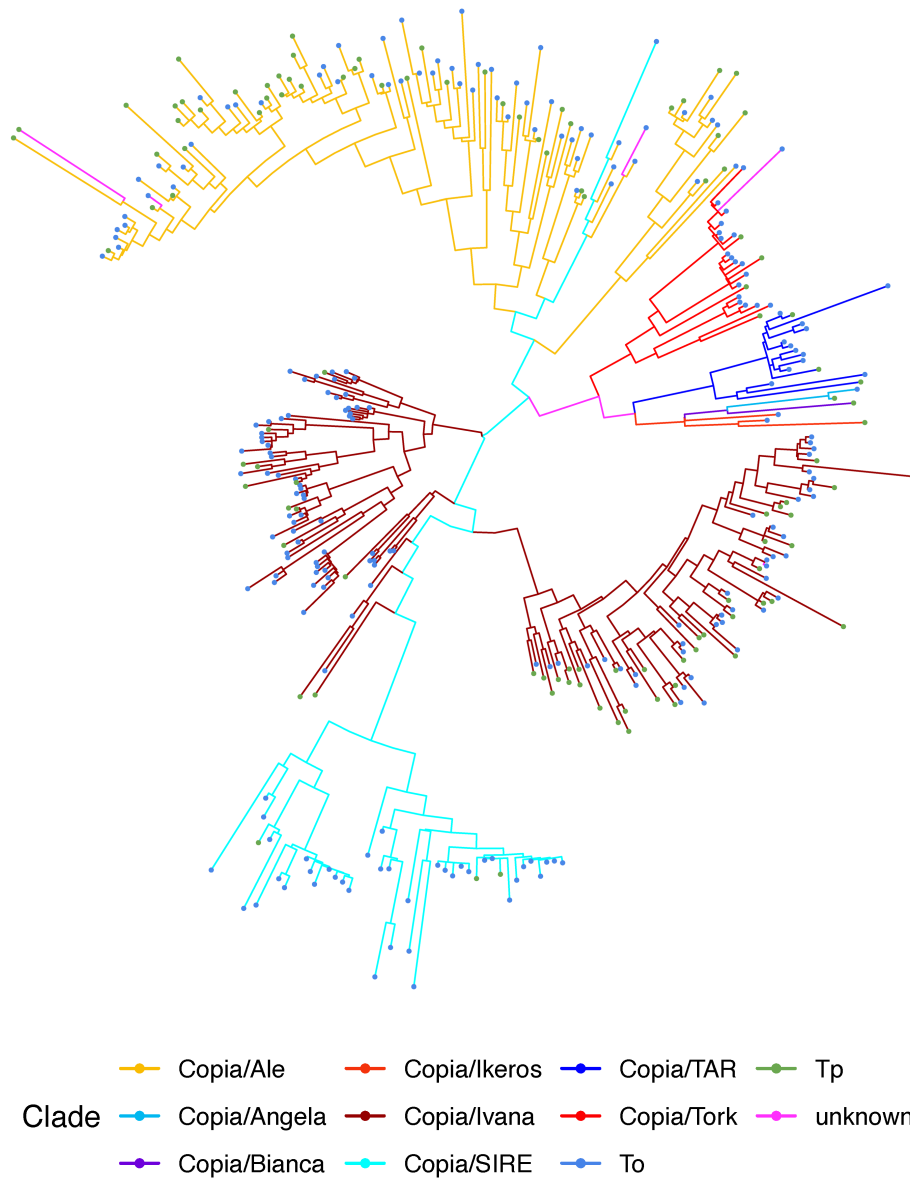
Supplementary Fig. S2. Categorization of retroelements (LTR-RT) of the white clover genome (haplotig1, this study) and the two extant diploid progenitors from [2].



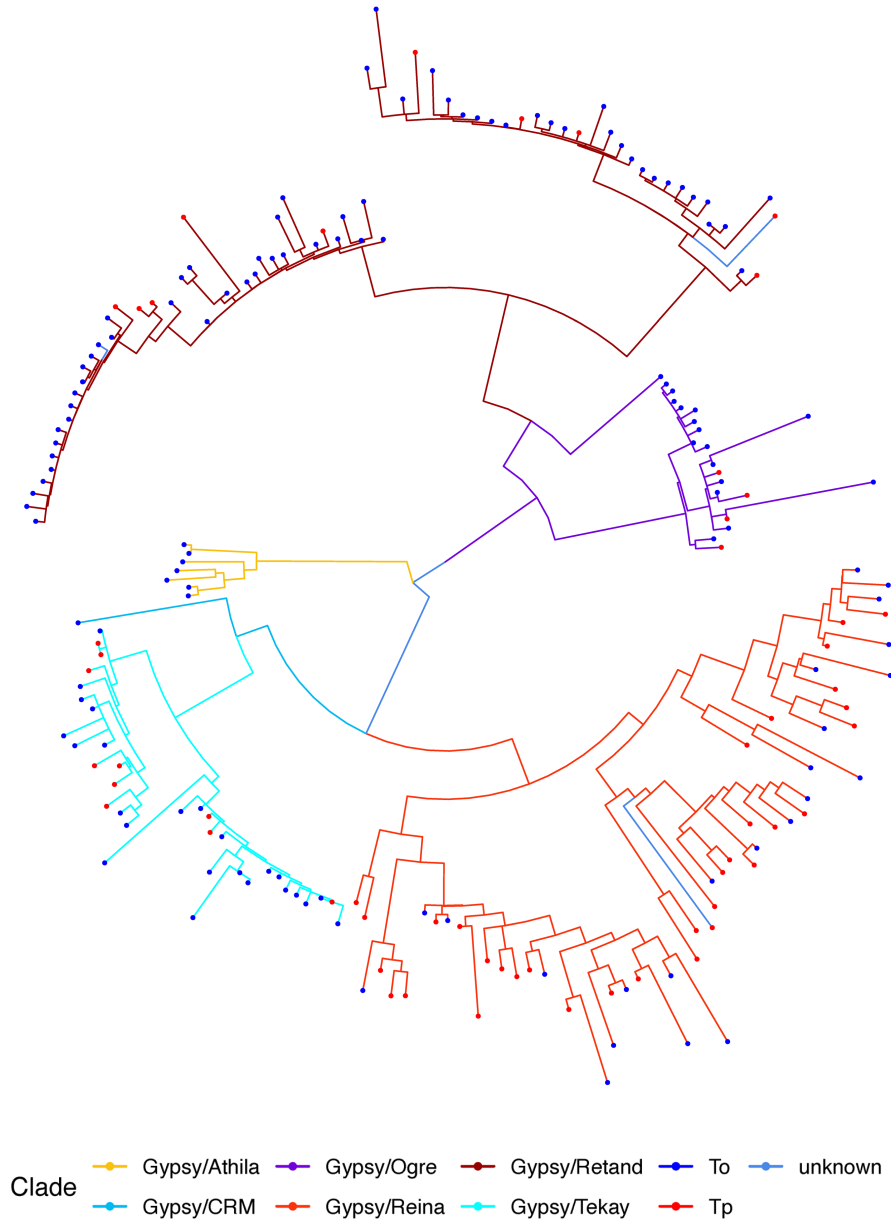
Supplementary Fig. S3. Phylogeny of the reverse transcriptase (RT) amino acid sequences of the *Copia* super-family in the white clover genome (hap1) by maximum likelihood method. The tree branches are colored by clades. The tree tips are color by the locations of the LTR-RT in *Trifolium repens* (occ., To in the legend) or *Trifolium repens* (pall., Tp in the legend) subgenome.



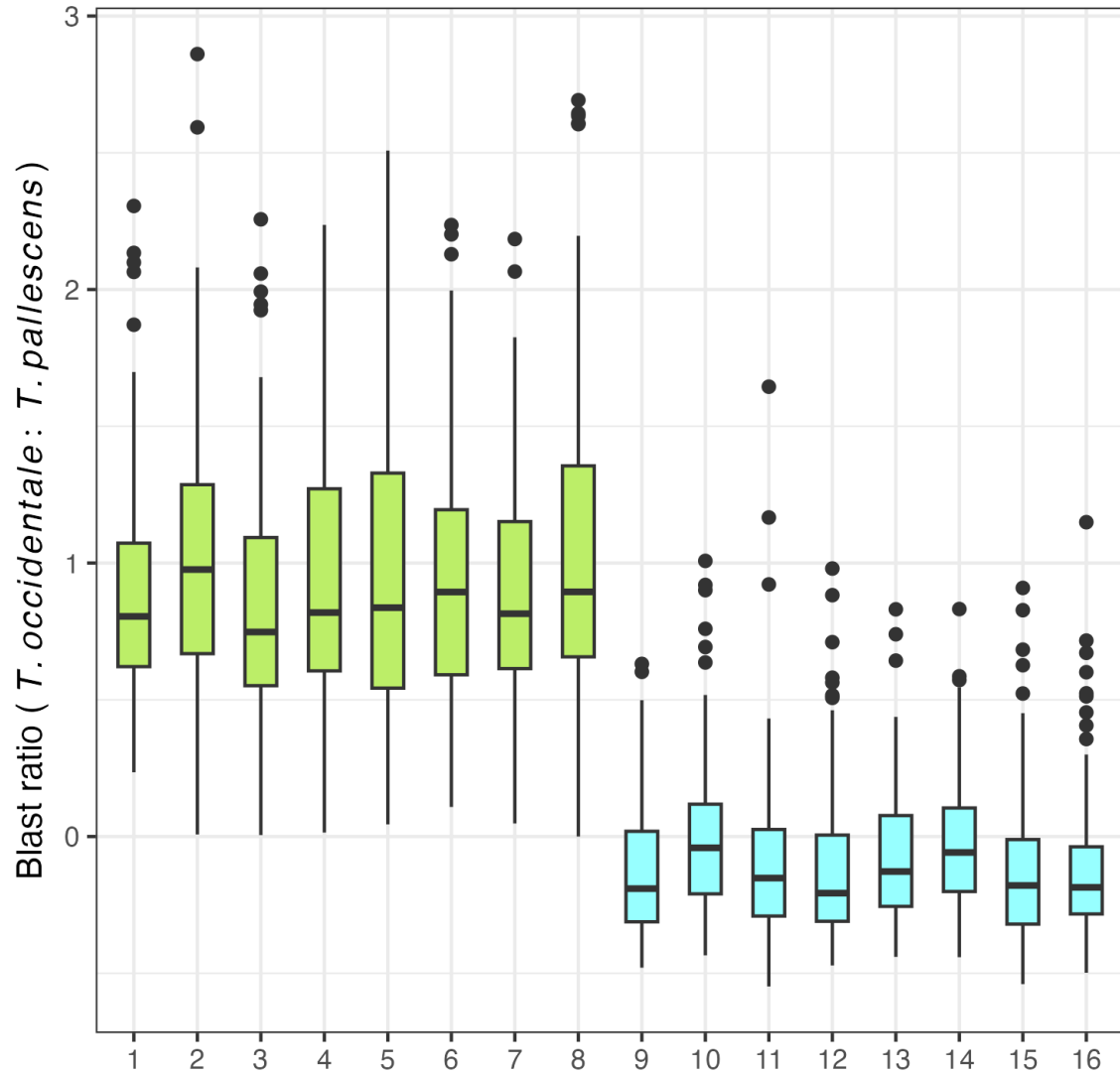
Supplementary Fig. S4. Phylogeny of the reverse transcriptase (RT) amino acid sequences of the *Gypsy* super-family in the white clover genome (hap1) by maximum likelihood method. The tree branches are colored by clades. The tree tips are color by the locations of the LTR-RT in *Trifolium repens* (occ., To in the legend) or *Trifolium repens* (pall., Tp in the legend) subgenome.



Supplementary Fig. S5. Phylogeny of the reverse transcriptase (RT) amino acid sequences of the *Copia* super-family in *T. occidentale* (To) and *T. pallescens* (Tp) genomes by maximum likelihood method. The tree branches are colored by clades. The tree tips are color by the locations of the LTR-RT in *T. occidentale* (To) or *T. pallescens* (Tp) species.

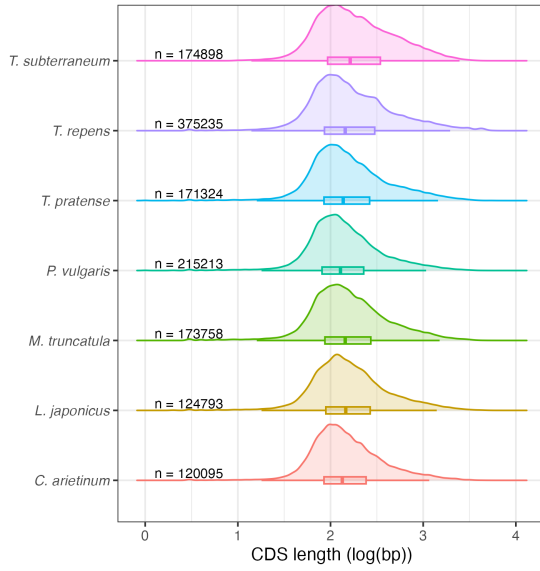


Supplementary Fig. S6. Phylogeny of the reverse transcriptase (RT) amino acid sequences of the *Gypsy* super-family in *T. occidentale* (To) and *T. pallescens* (Tp) genomes by maximum likelihood method. The tree branches are colored by clades. The tree tips are color by the locations of the LTR-RT in *T. occidentale* (To) or *T. pallescens* (Tp) species.

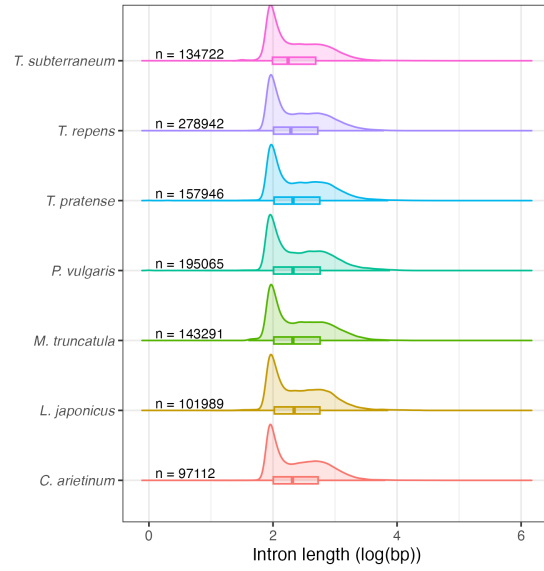


Supplementary Fig. S7. Chromosome categorized into *T. occidentale* (to) or *T. pallescens* (tp) subgenome. The Blast ratio is defined as $\sum (To \text{ alignment length})_i - \sum (Tp \text{ alignment length})_k / \text{Bin size (0.5 Mb)}$, where i and k are the filtered hits to either the *T. occidentale* (To) or *T. pallescens* (Tp) subgenome, respectively. This measure creates an index where a value of -1 is completely *T. pallescens*-like and +1 is completely *T. occidentale*-like.

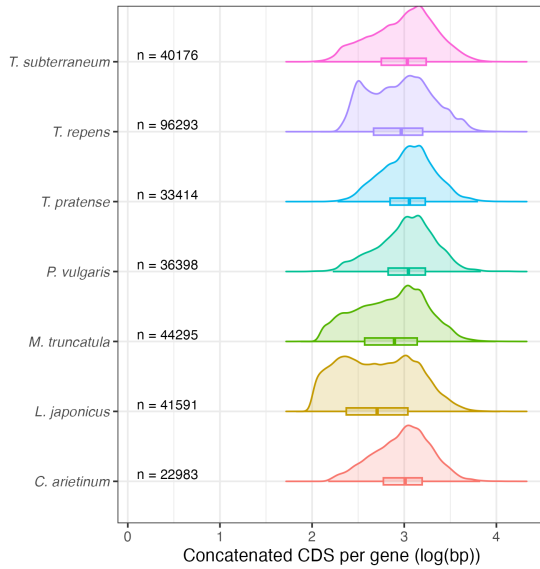
(A)



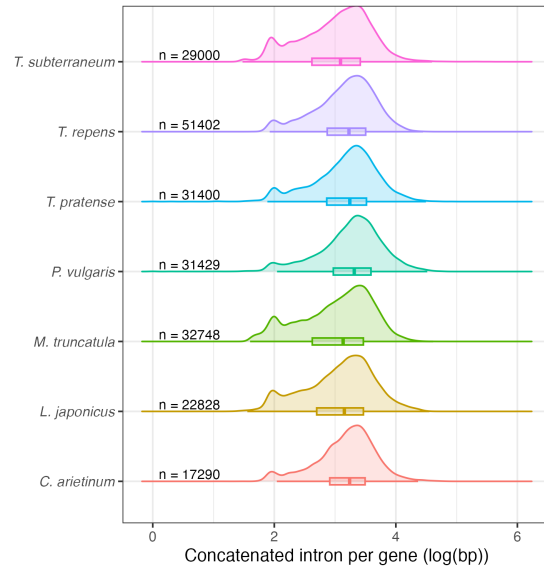
(B)



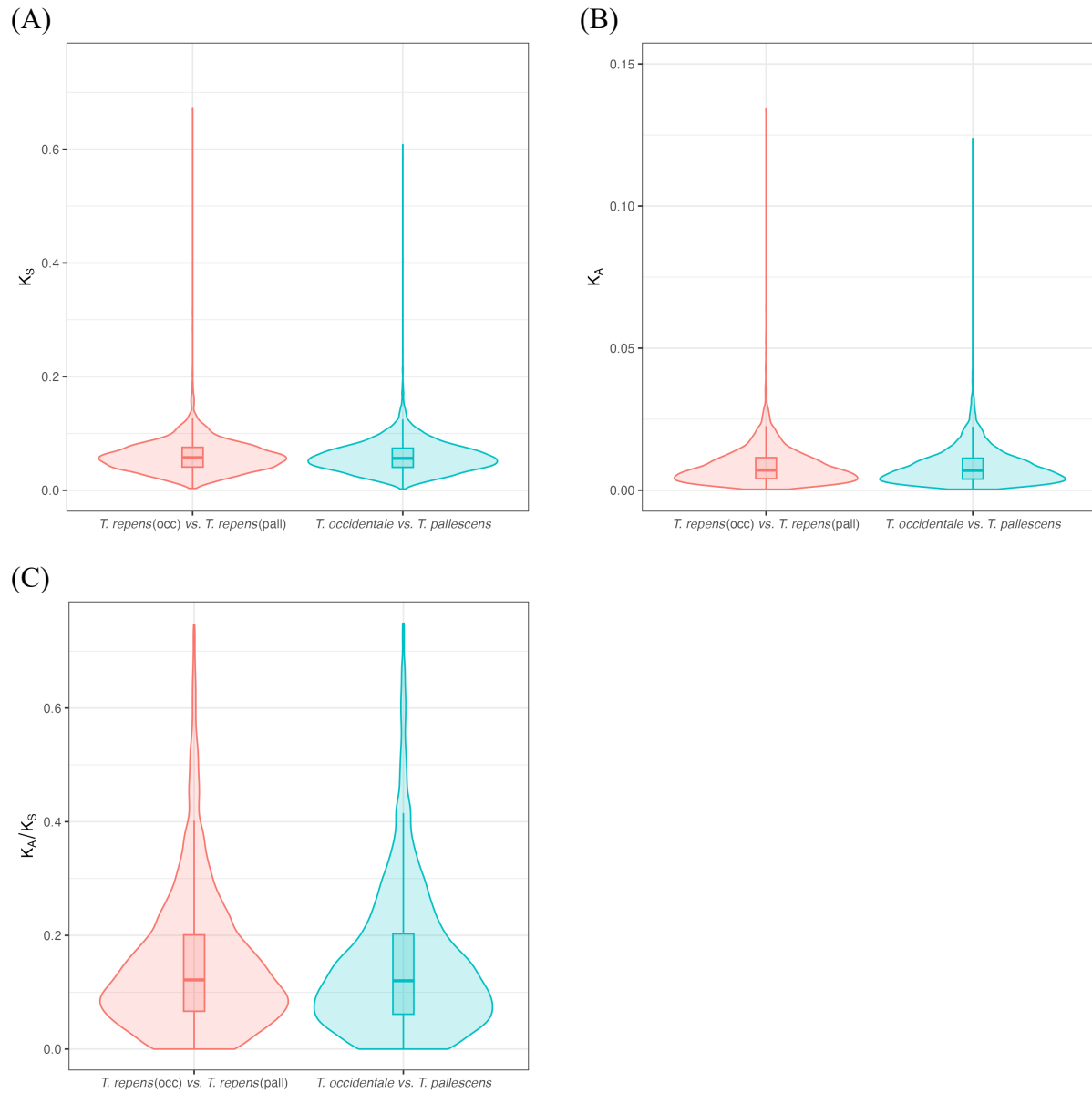
(C)



(D)

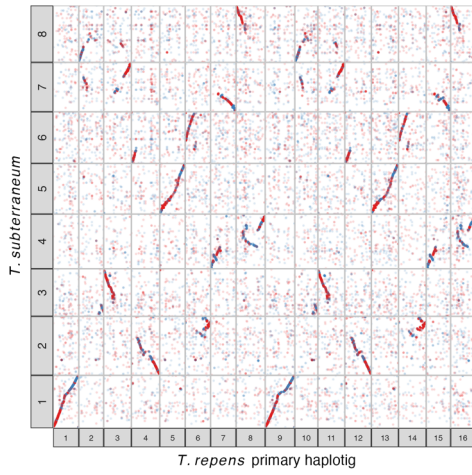


Supplementary Fig. S8. Coding sequence (CDS) length and intron length distribution. (A-B), the CDS or the introns are pooled and counted individually. (C-D), the CDS or the introns are concatenated per gene before counting.

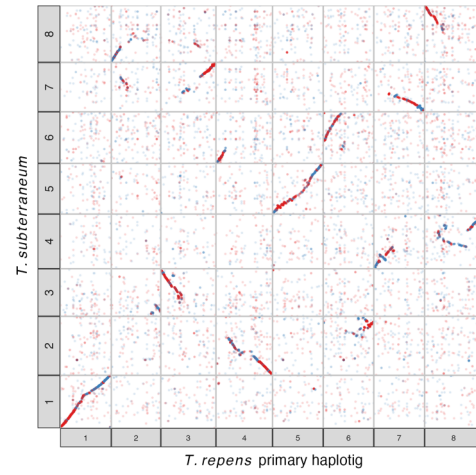


Supplementary Fig. S9. Pairwise nonsynonymous mutation rate (K_A) and synonymous mutation rate (K_S) between the white clover subgenomes and the diploid progenitors' genomes. No significant differences were detected by Student's t test.

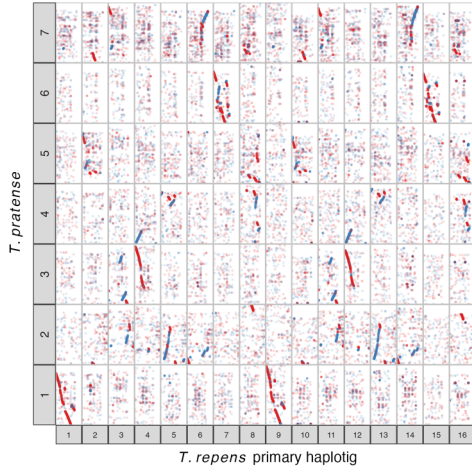
(A)



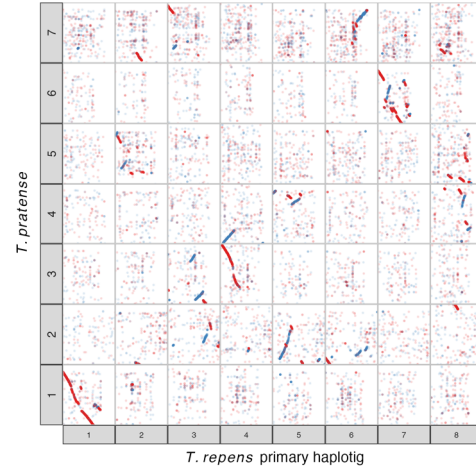
(B)



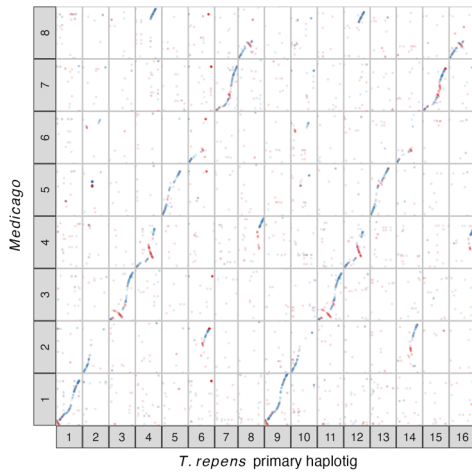
(C)



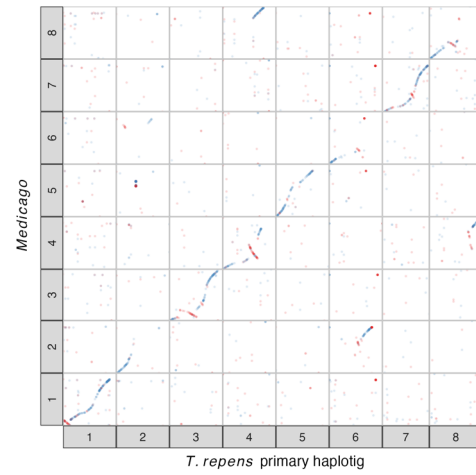
(D)



(E)

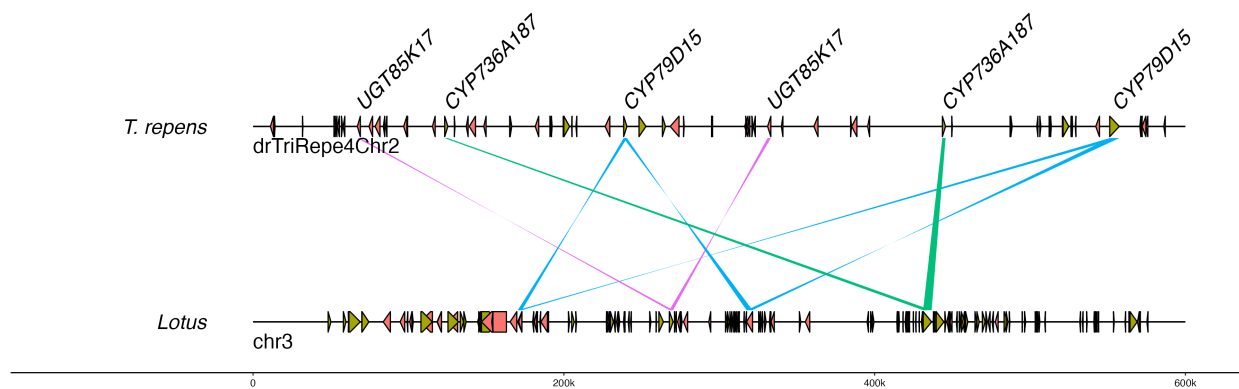


(F)

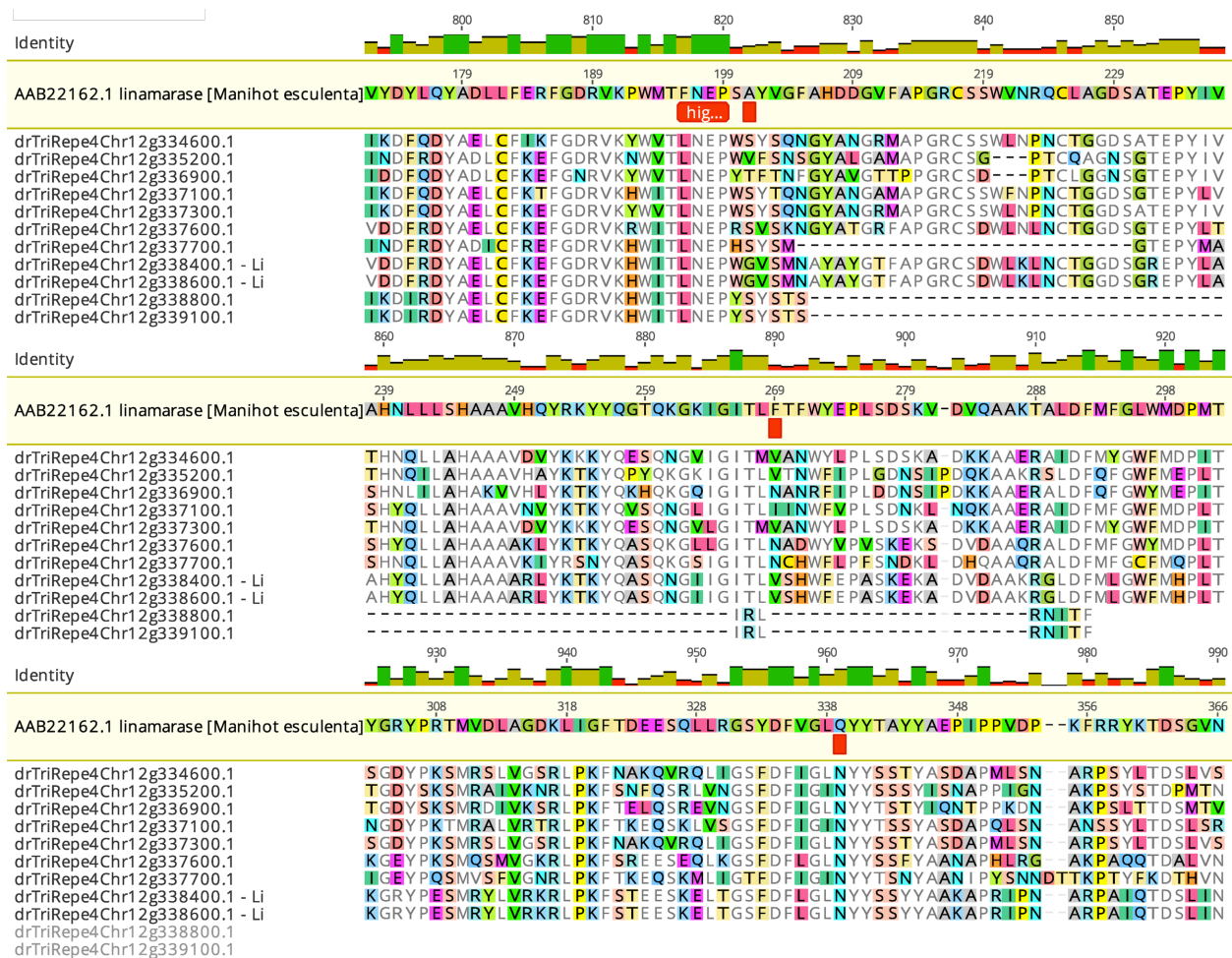


Supplementary Fig. S10. Whole genome nucleotide alignment of the white clover (this study) and related species. Blue, the pairwise alignment is in the same direction. Red, the pairwise

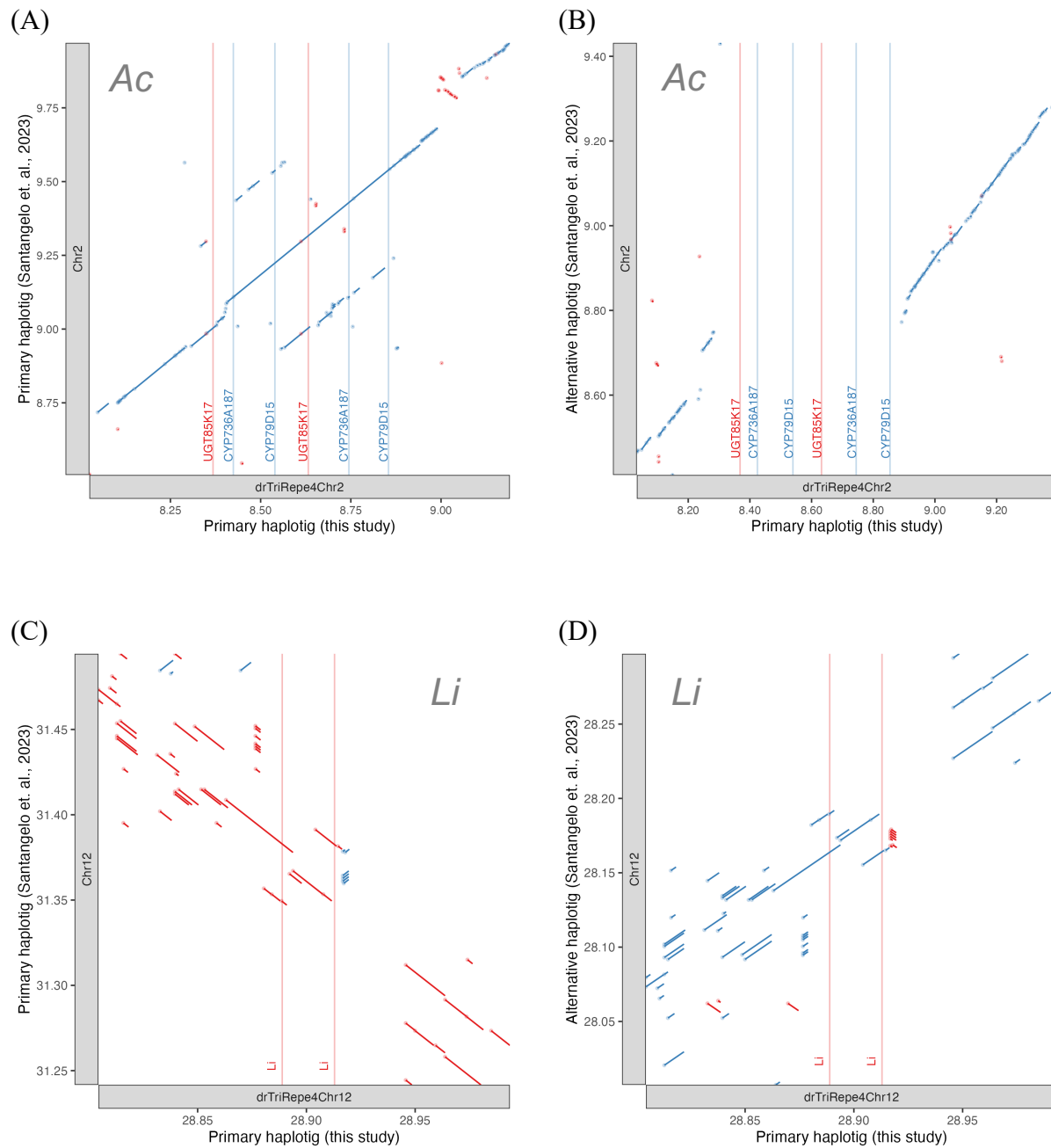
alignment is in the opposite directions. (B), (D), (F), only the chr 1 – chr 8 are shown. Only the alignments of the length >1000 bp and the identity >90% are shown.



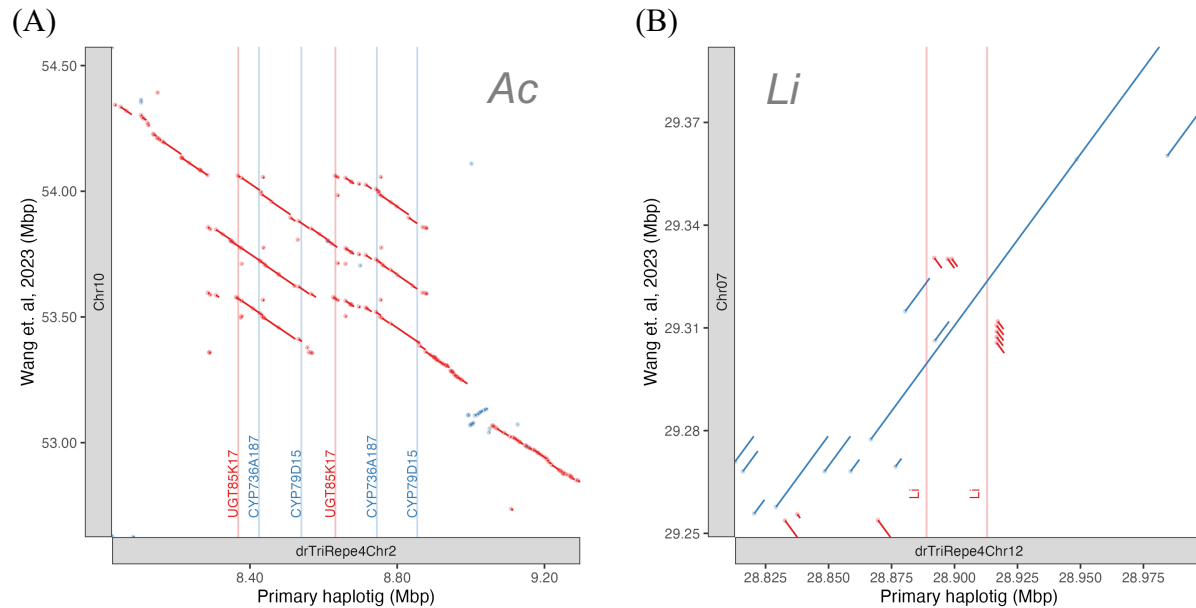
Supplementary Fig. S11. *Ac* gene cluster. Gene synteny between white clover chr 2 (primary haplotig) and *Lotus japonica* chr 3.



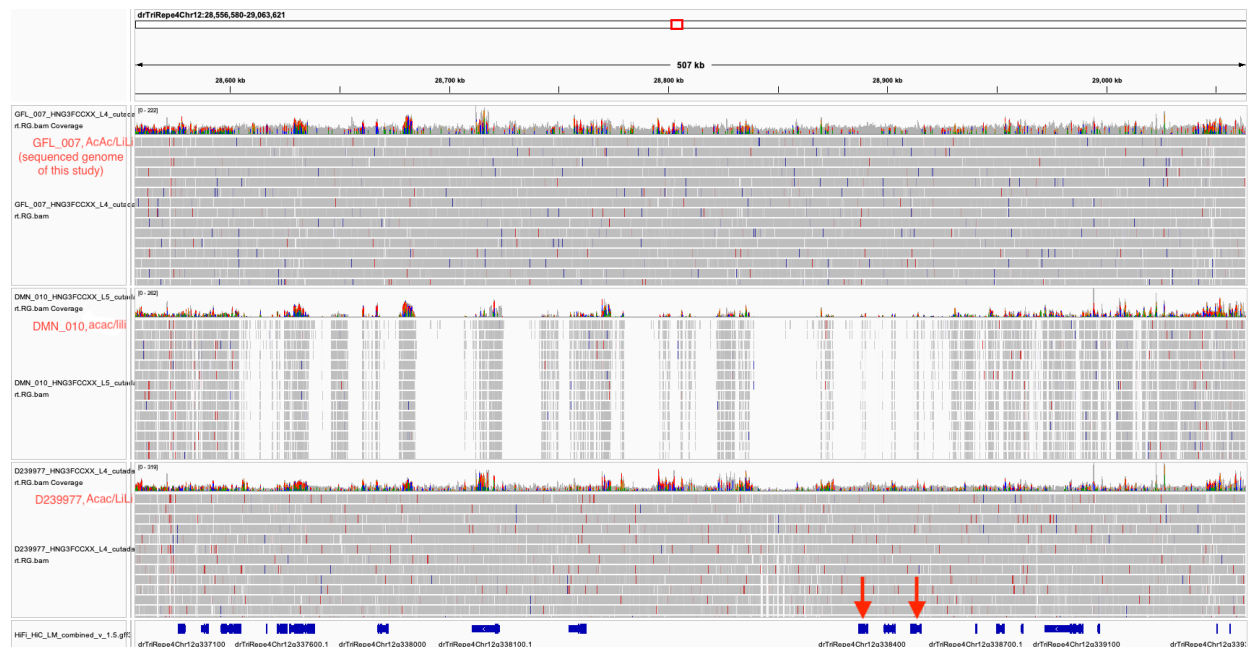
Supplementary Fig S12. Nucleotide alignment of the *Li* gene (cyanogenic glucosidase) and other glucosidase gene, that are predicted to have different substrate specificities than linamarase, at the *Li* locus. Only partial alignment, including the enzyme active center (the red annotation in the cassava linamarase gene), is shown. The amino acid composition around the active center has been predicted to affect the substrate specificity [3].



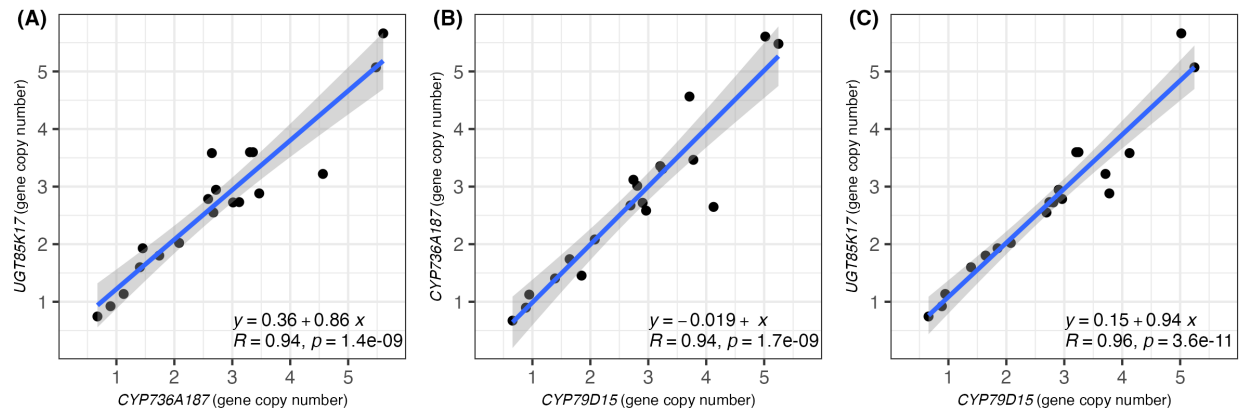
Supplementary Fig S13. Nucleotide alignment of the cyanogenesis gene copy number variation regions between the white clover reference genome of the present study and that of Santangelo *et al.* [4]. (A), there are two *Ac* gene clusters in the primary haplotig of this study and of Santangelo *et al.* [4] (the existence of the genes has been confirmed by BLAST). (B), whole *Ac* gene cluster deletion in the alternative haplotig of Santangelo *et al.* [4]. (C)-(D), there are two copies of *Li* gene in each haplotig of Santangelo *et al.* [4].



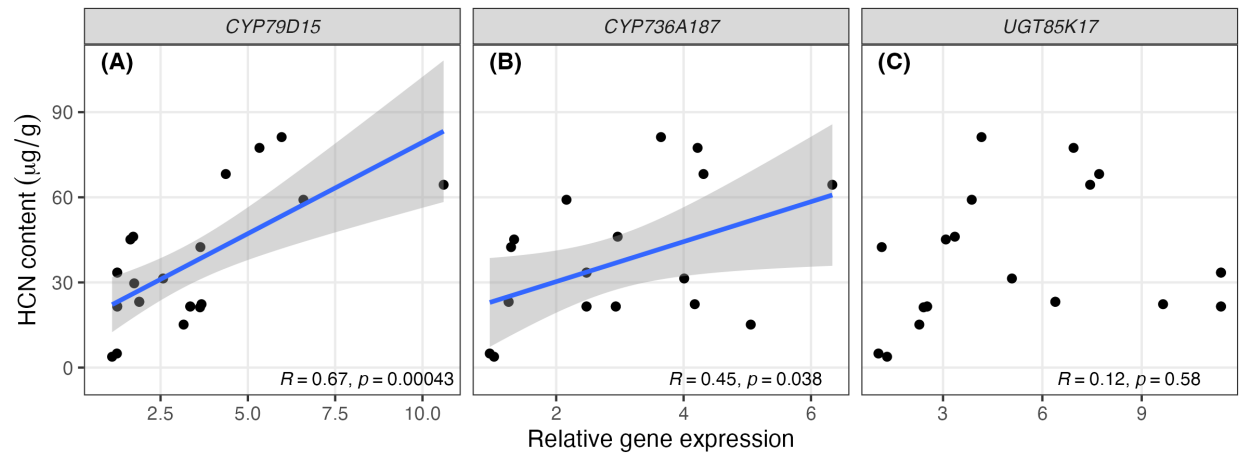
Supplementary Fig S14. Nucleotide alignment of the cyanogenesis gene copy number variation regions between the white clover reference genome of the present study and Wang *et al.* [5]. (A), there are two *Ac* gene clusters in the primary haplotig of this study and three *Ac* gene clusters in the genome of Wang *et al.* [5]. However, only two *CYP79D15* genes can be BLAST found based on their gene model annotations (See **Table S4**). (B), there are two copies of *Li* gene in the primary haplotig of this study and in the genome of Wang *et al.* [5].



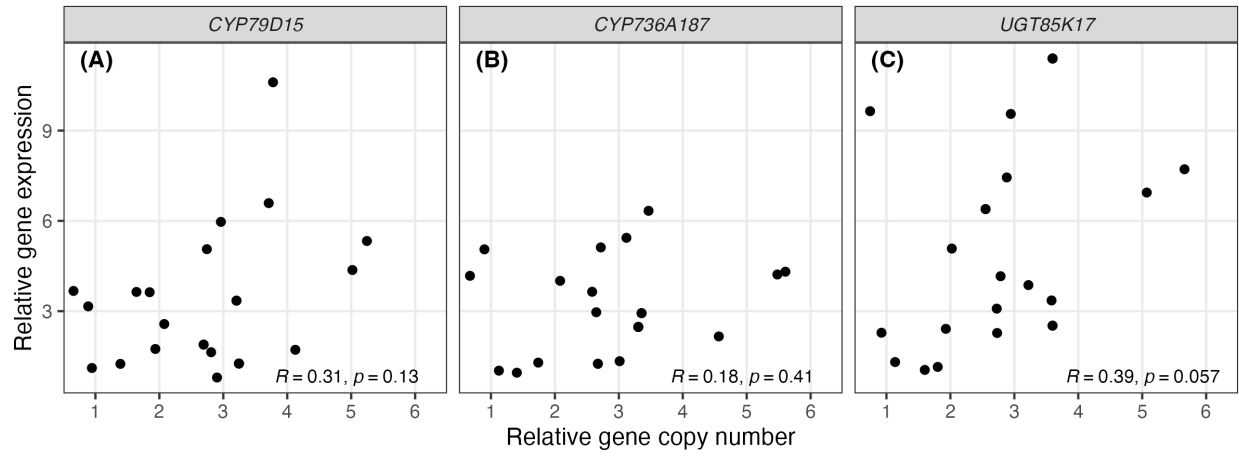
Supplementary Fig. S15. Mapping result of the whole genome re-sequencing of the accessions GFL_007 (NCBI BioSample: SAMN37329216) (*AcAc/LiLi*, sequenced genome of this study), DMN_010 (NCBI BioSample: SAMN34157026) (*acac/lili*), and PI239977 (NCBI BioSample: SAMNXXXXXXX) (*Acac/LiLi*). Note that in the accession that is homozygous for the *Li* locus deletion (*lili*), there is a low mapping window up to 500 kbp. Red arrows, the *Li* genes in the primary haplotig of the present study.



Supplementary Fig. S16. Pairwise gene copy number correlation of the genes in *Ac* gene cluster.

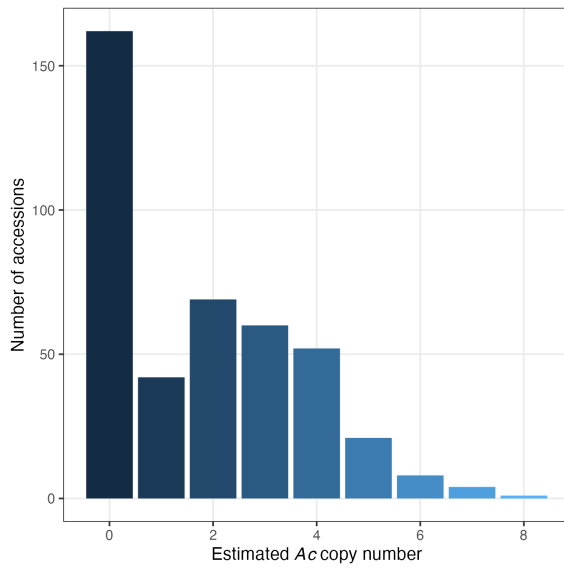


Supplementary Fig. S17. Relative gene expression and cyanide (HCN) content.

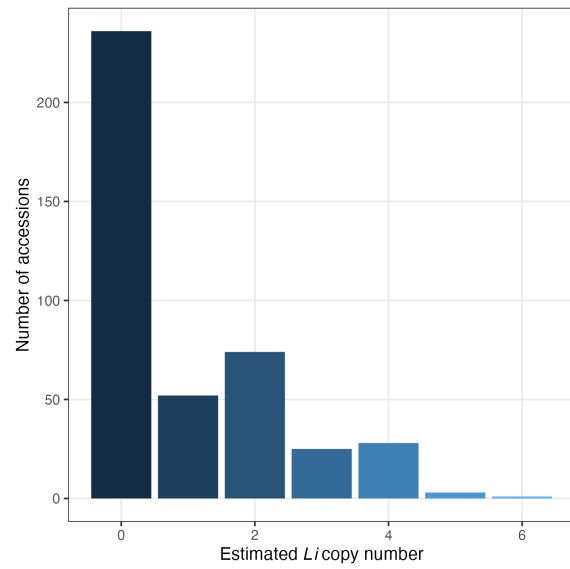


Supplementary Fig. S18. Gene copy number variation and relative gene expression. (A-C), Ac gene cluster.

(A)

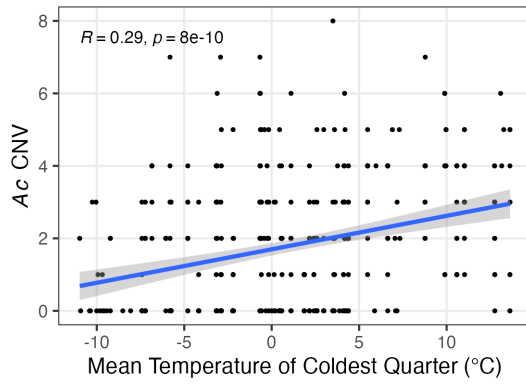


(B)

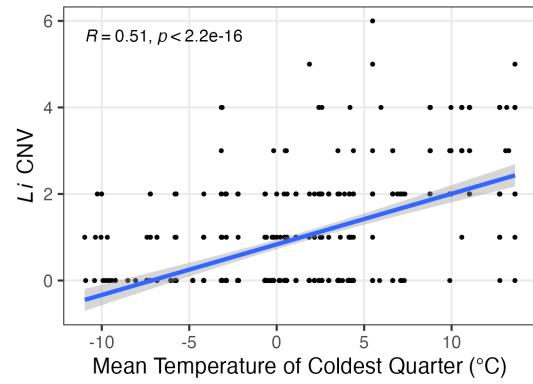


Supplementary Fig. S19. Estimated *Ac* (A) and *Li* (B) copy number of 419 wild accessions by genomic qPCR.

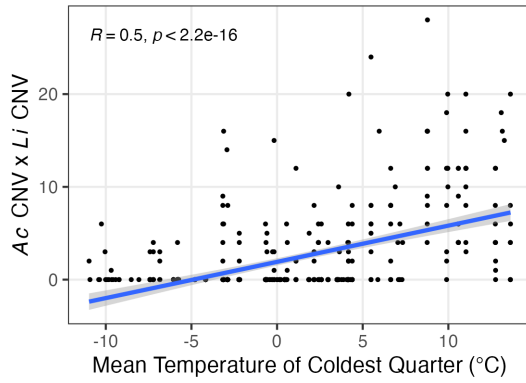
(A)



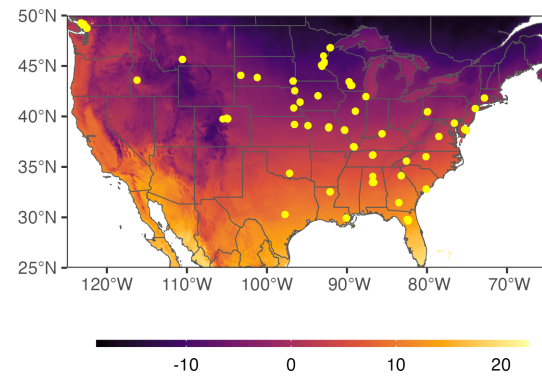
(B)



(C)

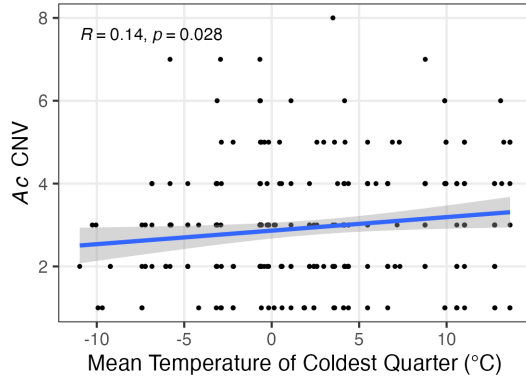


(D)

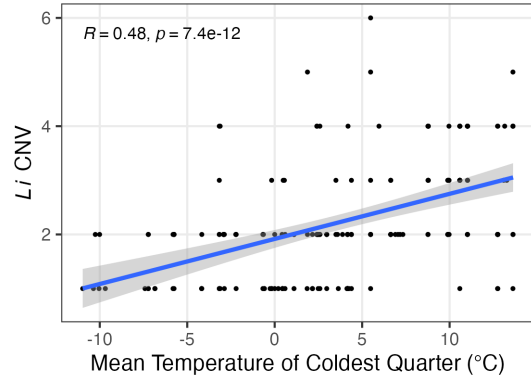


Supplementary Fig. S20. Association between cyanogenesis gene copy number variation (CNV) and mean temperature of coldest quarter (MTCQ). (A), *Ac* CNV. (B), *Li* CNV. (C), the product of (*Ac* CNV) * (*Li* CNV). (D), map of 419 accessions (43 populations) across North America. The map is shaded by MTCQ.

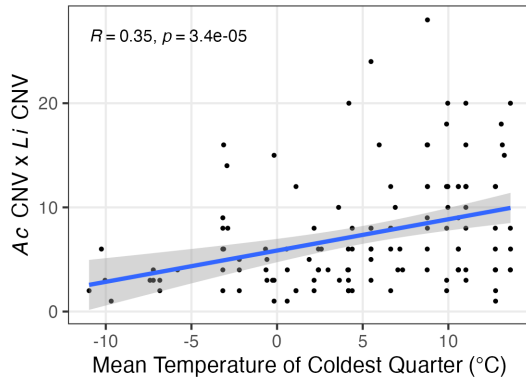
(A)



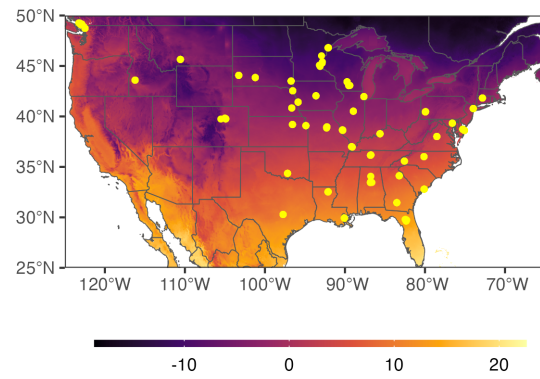
(B)



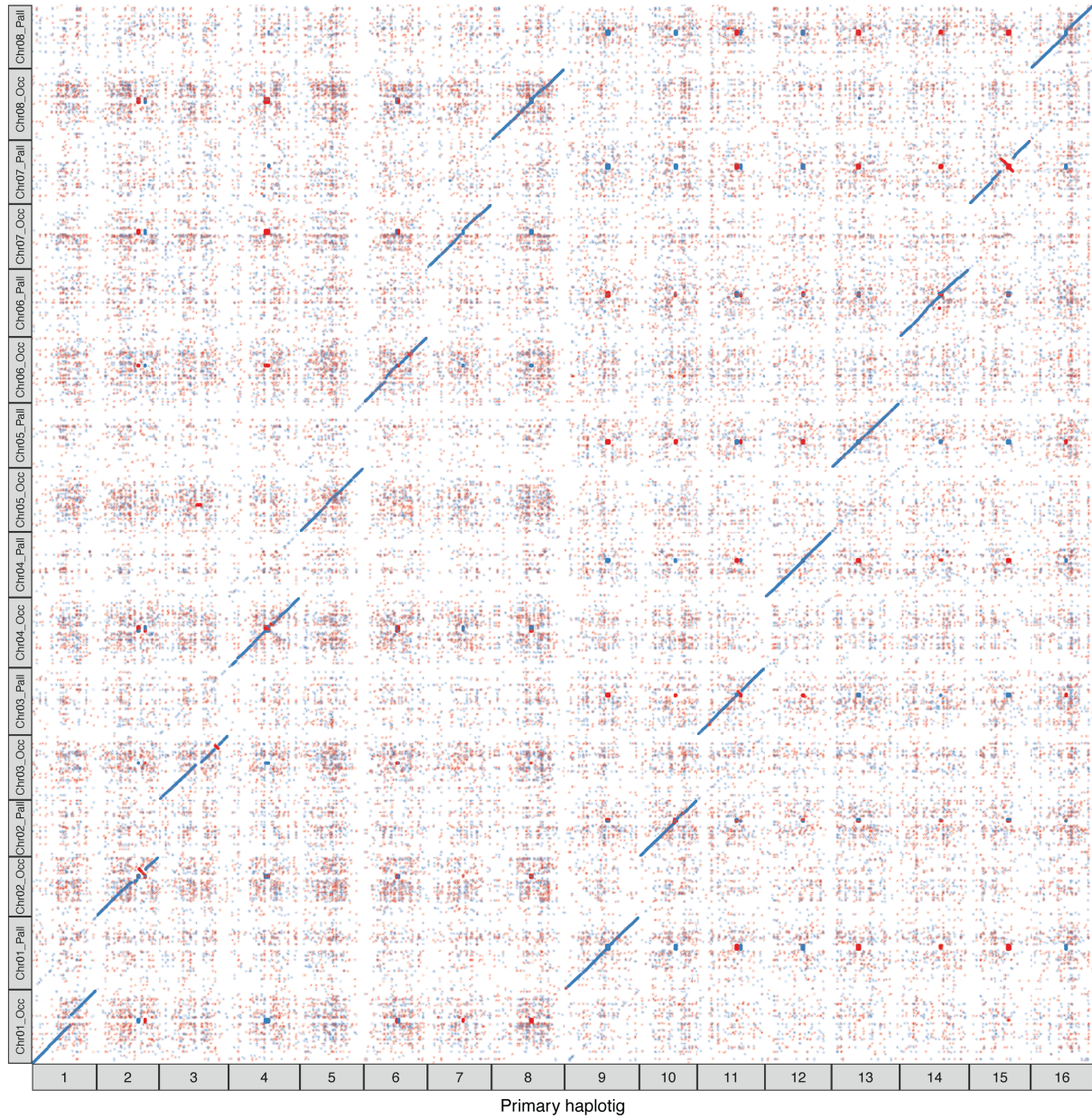
(C)



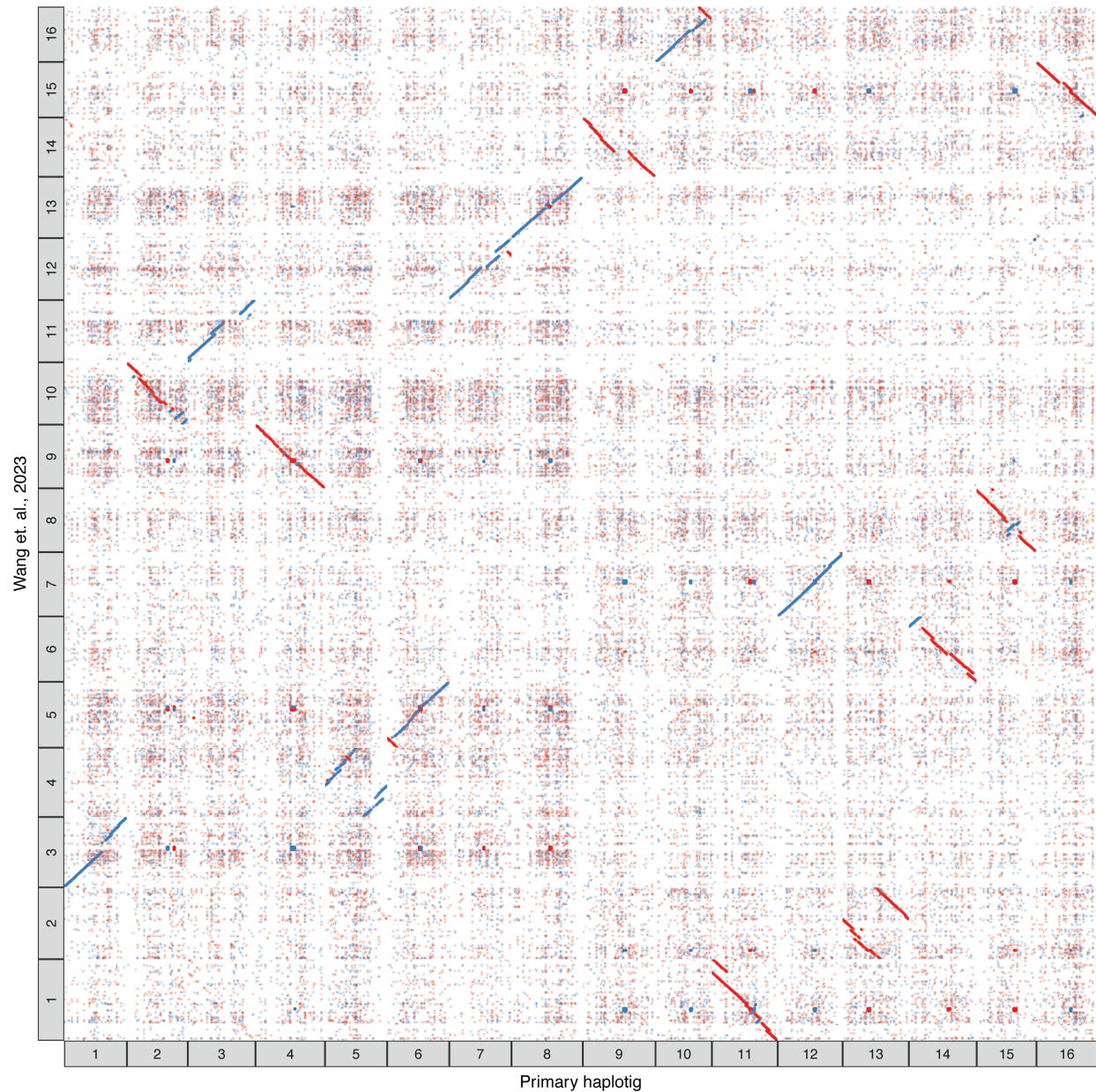
(D)



Supplementary Fig. S21. Association between cyanogenesis gene copy number variation (CNV) and mean temperature of coldest quarter (MTCQ). **Only the data with CNV ≥ 1 are shown.** (A), *Ac* CNV. (B), *Li* CNV. (C), the product of (*Ac* CNV)*(*Li* CNV). (D), map of 419 accessions (43 populations) across North America. The map is shaded by MTCQ.

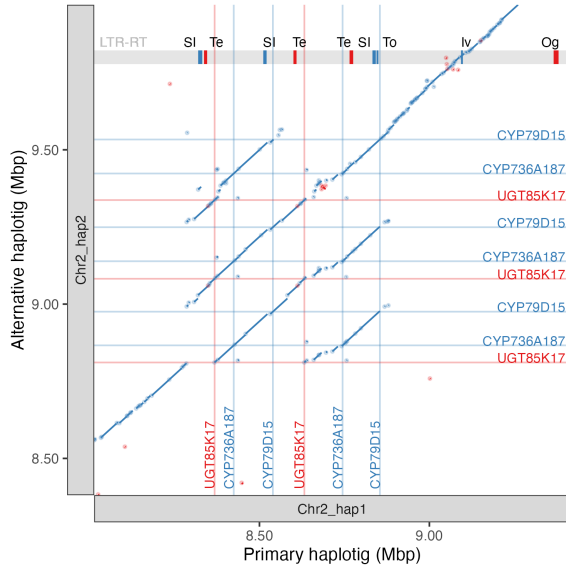


Supplementary Fig. S22. Whole genome nucleotide alignment (identity >95%, query length >10,000 bp, hit length >10,000 bp) between the primary haplotig (this study) and Santangelo *et al.* [4]. Linkage map assisted scaffolding was conducted in both genomes. Note that the two genomes are highly colinear to each other, except for a few scaffolds around the centromeric regions.

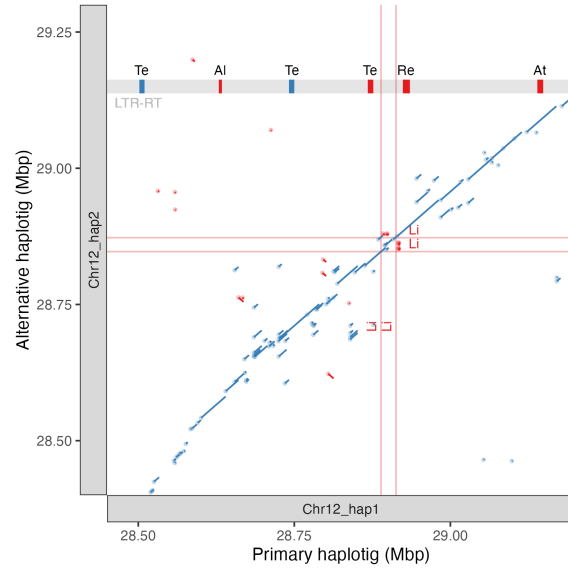


Supplementary Fig. S23. Whole genome nucleotide alignment (identity >95%, query length >10,000 bp, hit length >10,000 bp) between the primary haplotig (this study) and Wang *et al.* [5]. Linkage map assisted scaffolding was only conducted in the primary haplotig (this study). Note that Wang *et al.* [5] shows chromosome inversions or duplications at the corresponding chr 2, chr 5, chr 6, chr 10, chr 11, chr 13, chr 14 of the primary haplotig (this study).

(A)



(B)



Supplementary Fig. S24. Nucleotide alignment and gene synteny of the white clover cyanogenesis genes with LTR-RT labeled. **(A-B)**, nucleotide alignment of the regions around the *Ac* and the *Li* loci. Blue, the pairwise alignment is in the same direction. Red, the pairwise alignment is in the opposite directions. SI, *Copia/SIRE*. Te, *Gypsy/Tekay*. To, *Copia/Tork*. Iv, *Copia/Ivana*. Og, *Gypsy/Ogre*. Al, *Copia/Angela*. Re, *Gypsy/Reina*. At, *Gypsy/Athila*.

References

1. Abeynayake SW, Panter S, Chapman R, Webster T, Rochfort S, Mouradov A, Spangenberg G. **Biosynthesis of Proanthocyanidins in White Clover Flowers: Cross Talk within the Flavonoid Pathway** *Plant Physiology* 2011, **158**:666-678.
2. Griffiths AG, Moraga R, Tausen M, Gupta V, Bilton TP, Campbell MA, Ashby R, Nagy I, Khan A, Larking A, et al. **Breaking free: The genomics of allopolyploidy-facilitated niche expansion in white clover.** *Plant Cell* 2019, **31**:1466-1487.
3. Kongsaree PT, Ratananikom K, Choengpanya K, Tongtubtim N, Sujiwattanasat P, Porncharoenop C, Onpium A, Svasti J. **Substrate specificity in hydrolysis and transglucosylation by family 1 β -glucosidases from cassava and Thai rosewood.** *Journal of Molecular Catalysis B: Enzymatic* 2010, **67**:257-265.
4. Santangelo JS, Battlay P, Hendrickson BT, Kuo WH, Olsen KM, Kooyers NJ, Johnson MTJ, Hodgins KA, Ness RW. **Haplotype-Resolved, Chromosome-Level Assembly of White Clover (*Trifolium repens* L., Fabaceae).** *Genome Biology and Evolution* 2023, **15**:2023.2006.2006.543960.
5. Wang H, Wu Y, He Y, Li G, Ma L, Li S, Huang J, Yang G. **High-quality chromosome-level de novo assembly of the *Trifolium repens*.** *BMC Genomics* 2023, **24**:326.