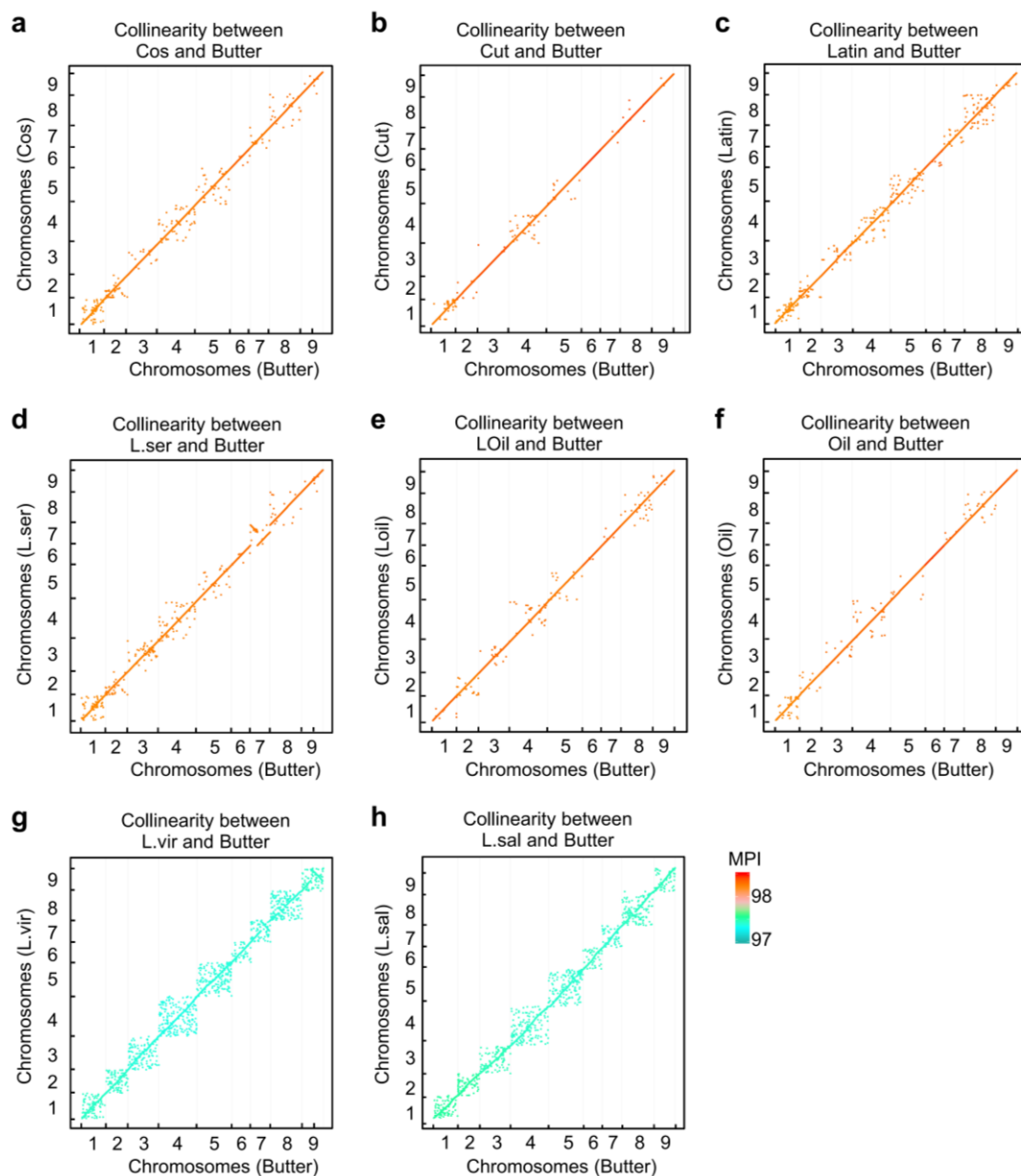
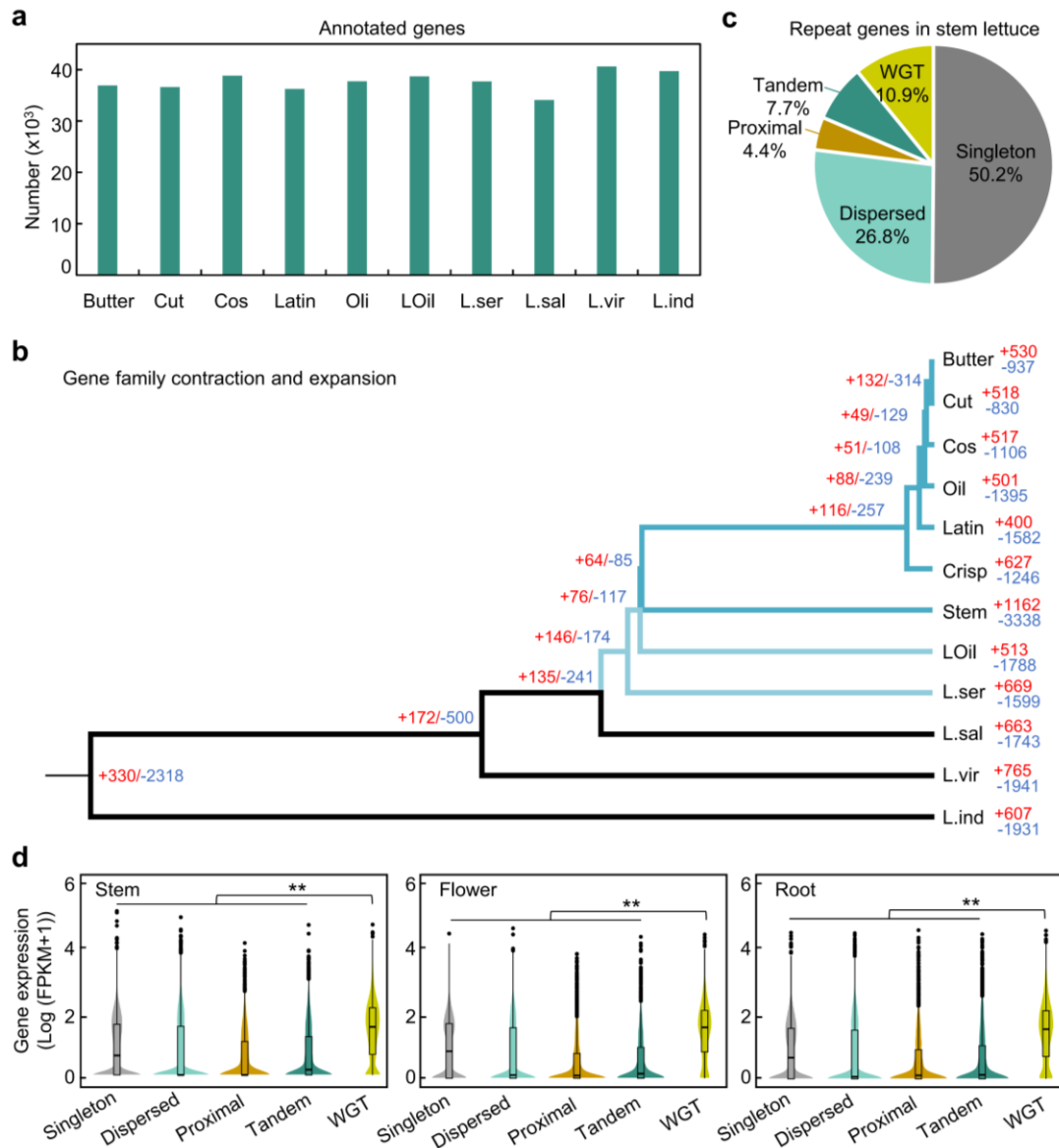


***Lactuca* super-pangenome provides insights into lettuce genome
evolution and domestication**

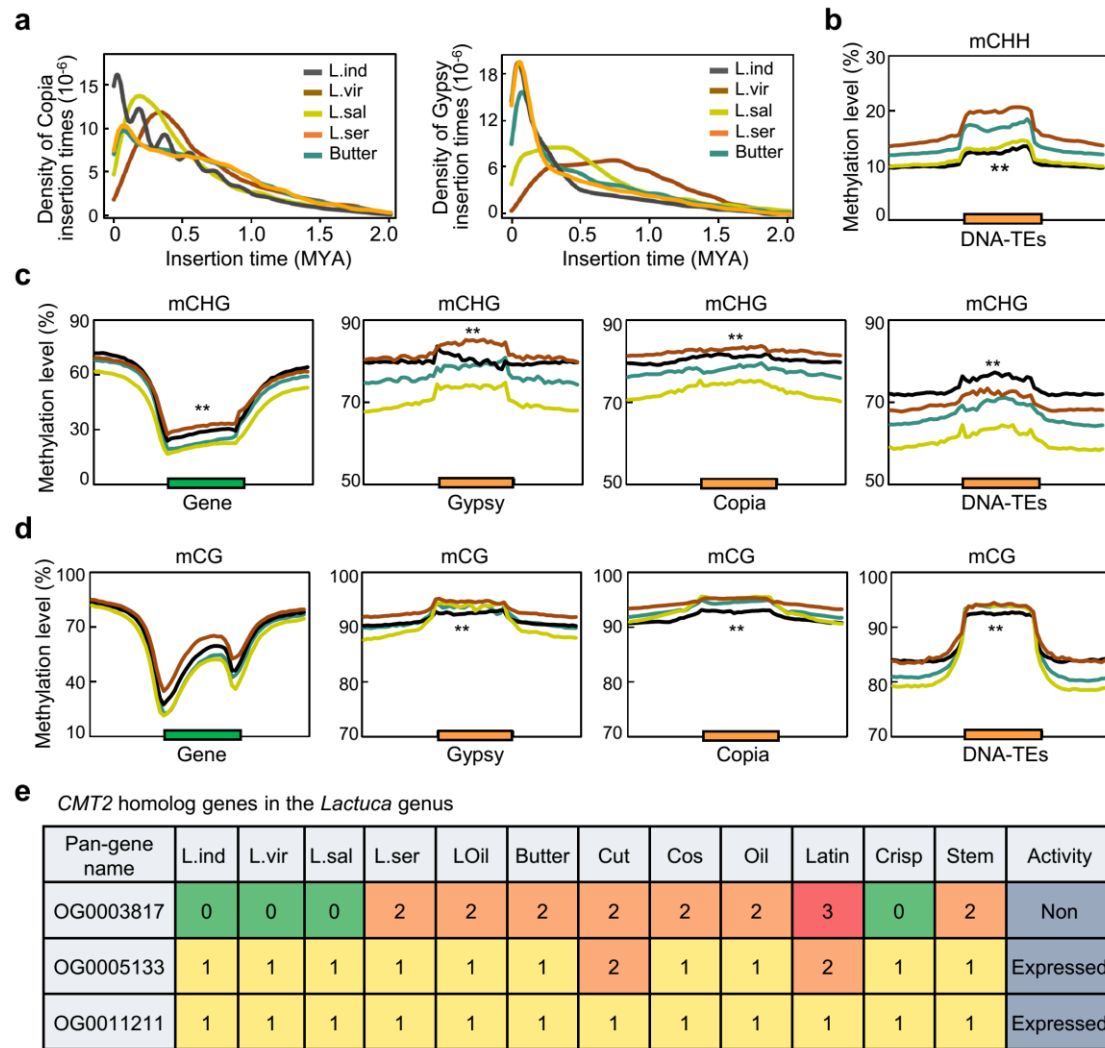
Cao *et al.*



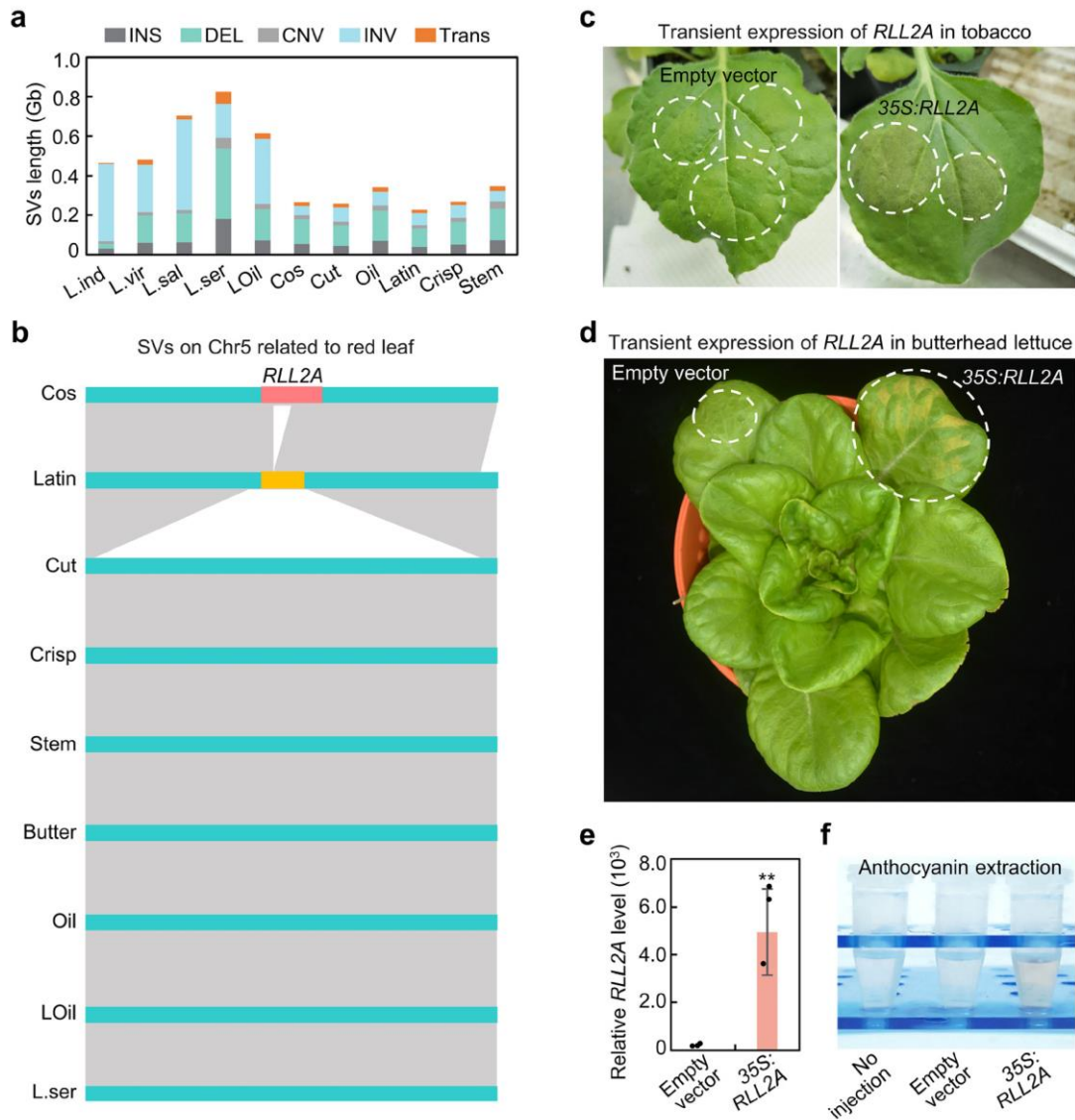
Supplementary Figure 1. Nucleotide alignment dot plots comparing the collinearities and similarities among *Lactuca* genomes. a-h Nucleotide alignment dot plots comparing the collinearities and similarities between butterhead lettuce and other *Lactuca* accessions, including cos lettuce (a), cutting lettuce (b), Latin lettuce (c), *L. serriola* (d), landrace (e), oilseed lettuce (f), *L. virosa* (g), and *L. saligna* (h). MPI indicates mean percent identify with a minimum nucleotide alignment length of 1 kb. Source data are provided as a Source Data file.



Supplementary Figure 2. Characterization of gene annotation across the *Lactuca* genomes. **a** Number of total annotated genes across the *Lactuca* genomes. **b** Gene family contraction and expansion in the *Lactuca* genome. **c** Percentage of WGT, small-scale duplicated (tandem, proximal, and dispersed), and single-copy (singleton) genes in total annotated genes in stem lettuce. **d** Expression levels of WGT, small-scale duplicated (tandem, proximal, and dispersed), and singleton genes in stem, flower, and root. The numbers of genes with PAV and without PAV are 3291 and 18293. And in boxplots, the 25% and 75% quartiles are shown as lower and upper edges of boxes, respectively, and central lines denote the median. The whiskers extend to 1.5× the interquartile range. Asterisks indicate significant differences (** $P < 0.01$, two-sided Wilcoxon signed-rank test). The exact P values (**d**) are provided in Supplementary Data 4. Source data are provided as a Source Data file.

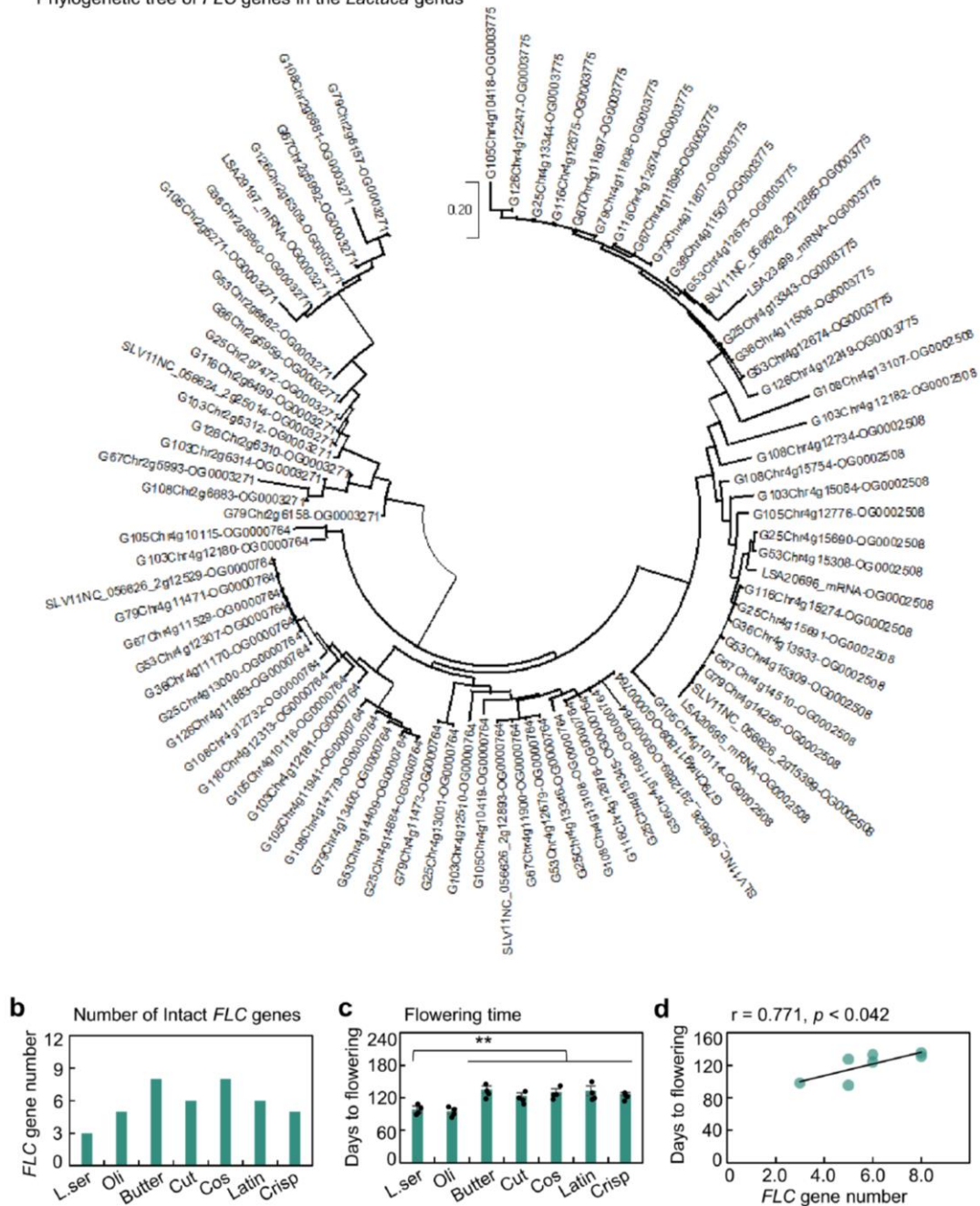


Supplementary Figure 3. Characterization of transposons. **a** Proliferation periods of the main retrotransposon superfamilies of Copia (left) and Gypsy (right) in the *Lactuca* species. **b-d** DNA methylation levels of CHH (**b**), CHG (**c**), and CHH (**d**) across different genomic features, including gene regions, Copia, Gypsy, and DNA TEs. Asterisks indicate significance differences (** $P < 0.01$, two-sided Wilcoxon signed-rank test) between *L. indica* and butterhead lettuce. The lines representing different accessions are color-coded as shown in **a**. **e** Number of CMT2 homolog genes across the *Lactuca* genus. The exact P values (**c** and **d**) are provided in Supplementary Data 4. Source data are provided as a Source Data file.

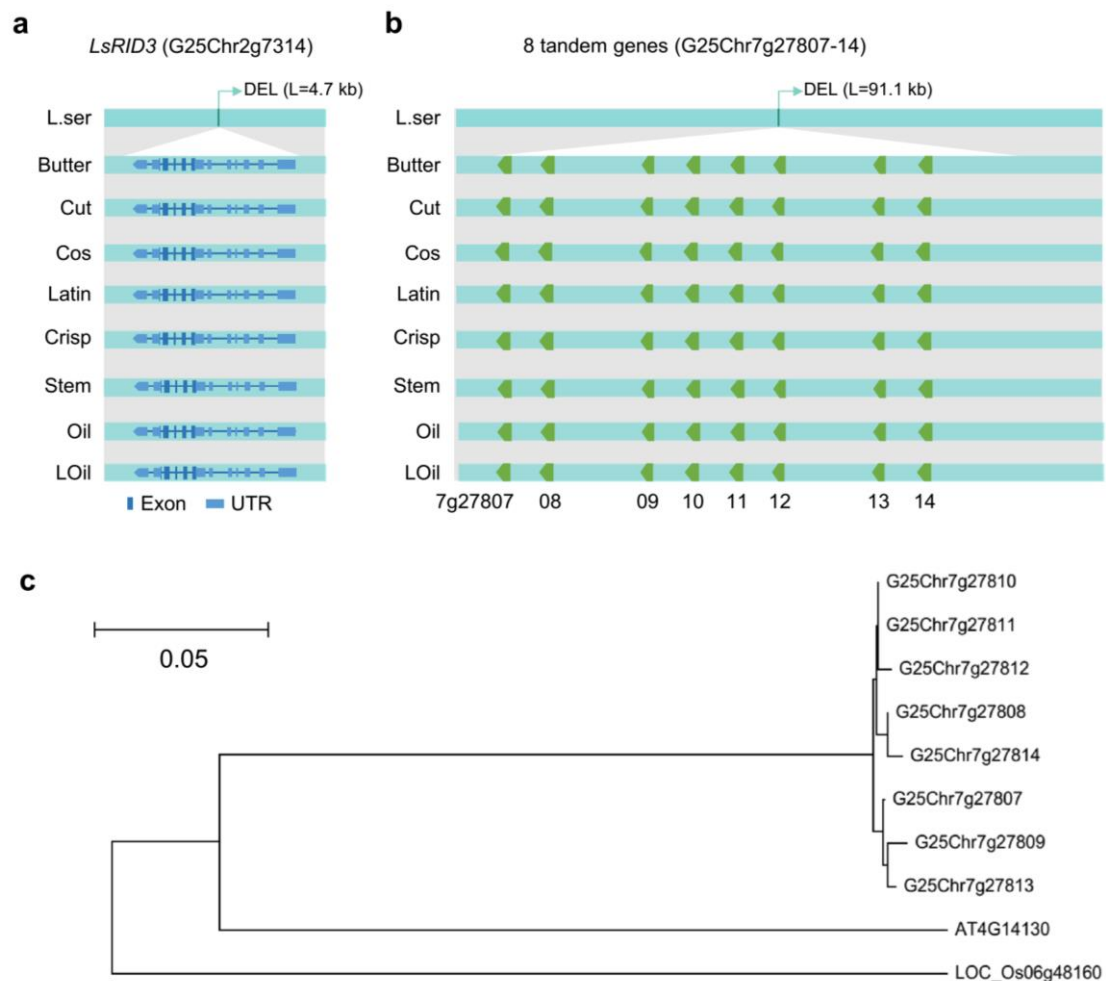


Supplementary Figure 4. The SVs related to red leaf. **a** Length of different types of SVs in each genome compared to the butterhead lettuce genome. **b** SVs on Chr5 involving *RLL2A* across the *Lactuca* genomes. The genomic locus of *RLL2A* contains multiple deletions in Latin lettuce as compared to cos lettuce and is shown in orange color. **c,d** Transient overexpression of *RLL2A* from cos lettuce in tobacco (**c**) and butterhead lettuce (**d**) results in red leaf color. **e** Relative expression of *RLL2A* in butterhead lettuce following transient expression of *RLL2A*. Data are presented as mean $\pm$ s.d. ($n = 3$ biological replicates). An asterisk indicates a significant difference (** $P < 0.01$, two-tailed Student's t test). The exact P values are provided in Supplementary Data 4. **f** Anthocyanin extraction from butterhead lettuce following transient overexpression of *RLL2A*. Source data are provided as a Source Data file.

a Phylogenetic tree of *FLC* genes in the *Lactuca* genus



Supplementary Figure 5. Increase in *FLC* copy numbers may contribute to delayed flowering traits during lettuce domestication. **a** Phylogenetic tree based on protein sequences encoded by *FLC* genes in the *Lactuca* genus. **b** Number of intact *FLC* genes detected in each genome annotations of the *Lactuca* genus. **c** Flowering time indicated by days to flowering of each *Lactuca* accession. Data are presented as mean $\pm$ s.d. (n = 4 biological replicates). Asterisks indicate significant differences ($**P < 0.01$, two-tailed Student's *t* test). The exact *P* values are provided in Supplementary Data 4. **d** Correlation between flowering time and *FLC* genes (*Spearman's* rank correlation coefficient). Source data are provided as a Source Data file.



Supplementary Figure 6. Additional examples of genes related to core domestication-associated PAVs. a SVs in *G25Chr2g7314* between wild (*L. serriola*) and cultivated lettuce. **b** A > 90 kb deletion encompassing 8 genes identified in wild (*L. serriola*) compared to cultivated lettuce. **c** Phylogenetic tree based on protein sequences encoded by the 8 tandemly repeated lettuce genes of lettuce, along with homologs from *Arabidopsis* and rice. Source data are provided as a Source Data file.

Supplementary Table 1. Organization of repetitive sequences.

Classification		<i>Lactuca indica</i>	<i>Lactuca virosa</i>	<i>Lactuca saligna</i>	<i>Lactuca serriola</i>	<i>Lactuca serriola</i>	Butterhead	Cutting	Cos	Oilseed	Latin
Total genome length		5,548,885,818	3,382,219,856	2,156,369,586	2,629,607,914	2,597,086,913	2,599,952,138	2,593,728,029	2,603,406,097	2,610,155,122	2,601,801,946
LTR retroelements	Ty1/Copia	2,448,906,009	606,482,797	377,381,835	591,235,638	423,598,306	362,693,134	460,046,927	435,877,024	483,194,397	337,202,502
	Gypsy/DIRS1	1,096,577,537	1,140,643,350	783,970,968	906,880,052	925,736,876	992,920,433	887,305,267	915,116,054	944,898,869	913,777,964
	other LTR	332,267,696	770,207,648	399,197,285	397,804,761	596,517,368	562,095,156	529,510,050	547,730,921	504,896,544	587,844,868
Other retroelements		31,007,670	19,544,382	8,551,616	4,701,524	12,911,243	23,040,001	11,582,170	13,232,508	10,254,581	12,275,740
DNA transposons		132,713,051	61,746,659	43,105,567	43,365,614	51,342,477	48,767,821	61,786,152	51,489,232	74,676,855	49,749,880
Unclassified		905,704,657	354,596,598	211,214,016	322,639,733	222,023,888	242,786,521	269,836,622	272,231,045	223,810,328	333,584,524
Other repeats	Small RNA	8,930,575	7,430,728	20,977,845	16,226,341	6,507,260	10,975,635	16,286,953	8,989,420	16,431,147	10,853,095
	Satellites	6,419,902	8,559,865	716,810	534,179	3,103,151	4,081,334	2,280,843	4,900,741	1,067,947	1,595,285
	Simple repeats	65,657,603	29,818,094	23,614,680	22,784,528	28,851,212	26,349,972	26,733,896	27,629,663	27,668,537	27,983,542
	Low complexity	2,406,365	1,947,174	1,550,655	1,559,548	1,674,061	1,704,334	1,666,796	1,671,977	1,650,562	1,663,785
	Norepeat	509,724,194	378,083,304	282,529,161	303,822,757	321,828,042	320,961,579	320,773,132	319,383,624	318,027,511	321,120,413

Supplementary Table 2. Verification of SVs between butterhead lettuce and the wild relative *L. saligna*. Source data are provided as a Source Data file.

SVs Loci	Primers	Confirmed by PCR analysis
Chr1:3372716-3373544	PDE_1F/R	YES
Chr1:4180700-4181277	PDE_2F/R	YES
Chr1:12495943-12496480	PDE_3F/R	YES
Chr1:25967323-25967938	PDE_4F/R	YES
Chr2:33250730-33251246	PDE_5F/R	YES
Chr2:102222127-102222769	PDE_6F/R	YES
Chr2:109879153-109879871	PDE_7F/R	YES
Chr3:51484712-51485433	PDE_8F/R	YES
Chr4:17068311-17068983	PDE_9F/R	YES
Chr4:20470972-20471681	PDE_10F/R	YES
Chr5:26218108-26218698	PDE_11F/R	YES
Chr5:6854254-6855076	PDE_12F/R	YES
Chr6:3918532-3919124	PDE_13F/R	YES
Chr6:14452916-14453424	PDE_14F/R	YES
Chr7:2503846-2504433	PDE_15F/R	YES
Chr7:11332483-11333049	PDE_16F/R	YES
Chr8:28113834-28114417	PDE_17F/R	NO
Chr8:81856112-81856794	PDE_18F/R	NO
Chr9:9934808-9935329	PDE_19F/R	NO
Chr9:10471393-10472233	PDE_20F/R	YES

Supplementary Table 3. CT values of qPCRs.

Gene	Cos	Latin	Butterhead	Cutting
<i>RLL2A</i>	23.33	38.77	NA	NA
	24.15	NA	NA	NA
	23.56	38.12	NA	NA
<i>LsActin</i>	18.32	18.43	18.54	19.33
	18.28	18.22	17.68	18.85
	18.31	18.61	17.89	18.37