

Supplementary Information for

Evolutionary history and pan-genome dynamics of strawberry (*Fragaria* spp.)

Qin Qiao^{a,1}, Patrick P. Edger^{b,1,2}, Li Xue^{c,1}, La Qiong^{d,1}, Jie Lu^a, Yichen Zhang^a, Qiang Cao^a, Alan E. Yocca^{b,e}, Adrian E. Platts^b, Steven J. Knapp^f, Marc Van Montagu^{g,h,2}, Yves Van de Peer^{g,h,i,j,2}, Jiajun Lei^{c,2}, Ticao Zhang^{k,2}

Email: edgerpat@msu.edu, marc.vanmontagu@ugent.be, yves.vandeppeer@psb.vib-ugent.be, jjajunleisy@163.com or ticaozhang@126.com

This PDF file includes:

Figures S1 to S13
Tables S1 to S5

Supplementary Figures

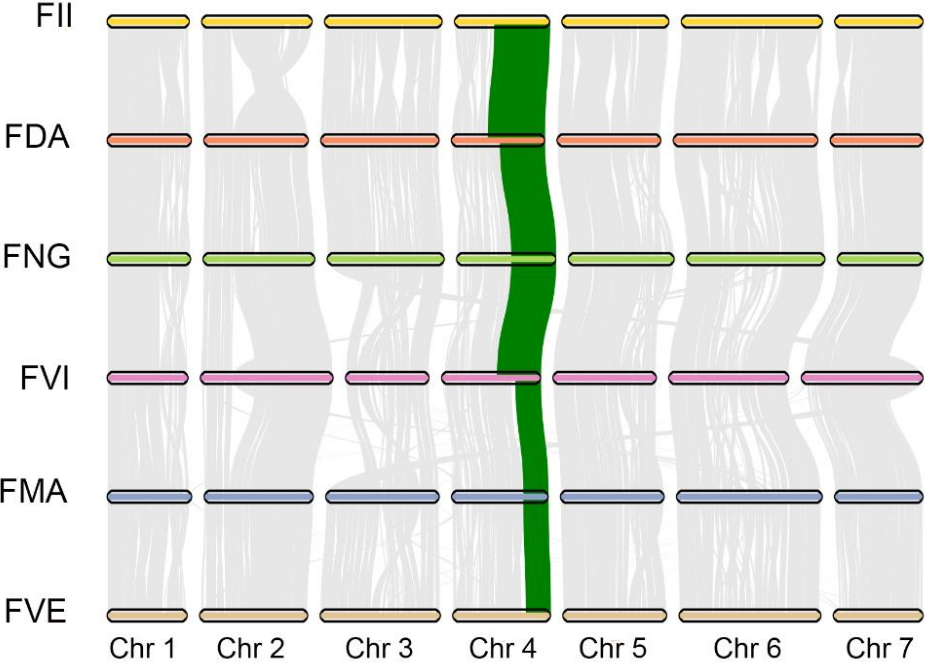


Fig. S1. Characterization of genomic variations in *Fragaria* spp. Synteny blocks between pair genomes are connected by gray lines. The 1:1 representative syntenic blocks are highlighted by green links.

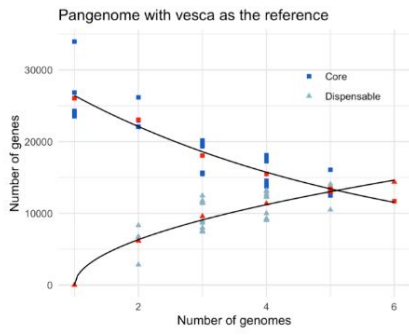
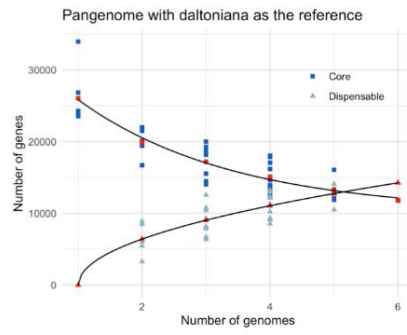
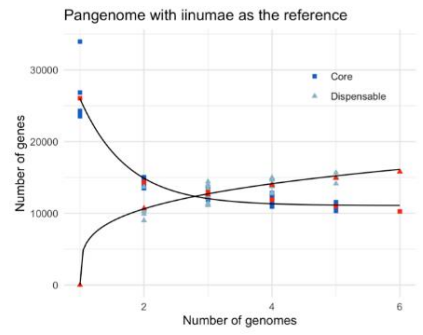
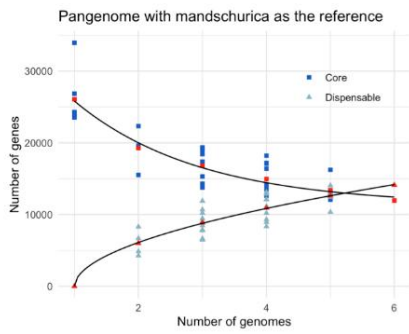
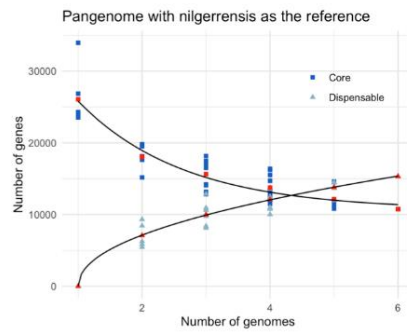
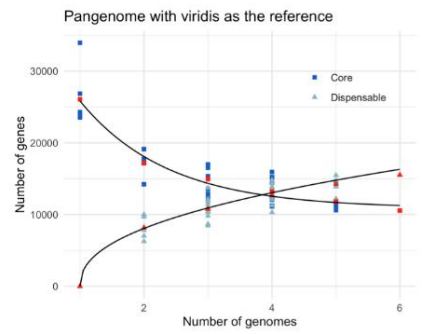
A**B****C****D****F****E**

Fig. S2. *Fragaria* pan-genome modeling with six species as the reference, respectively.

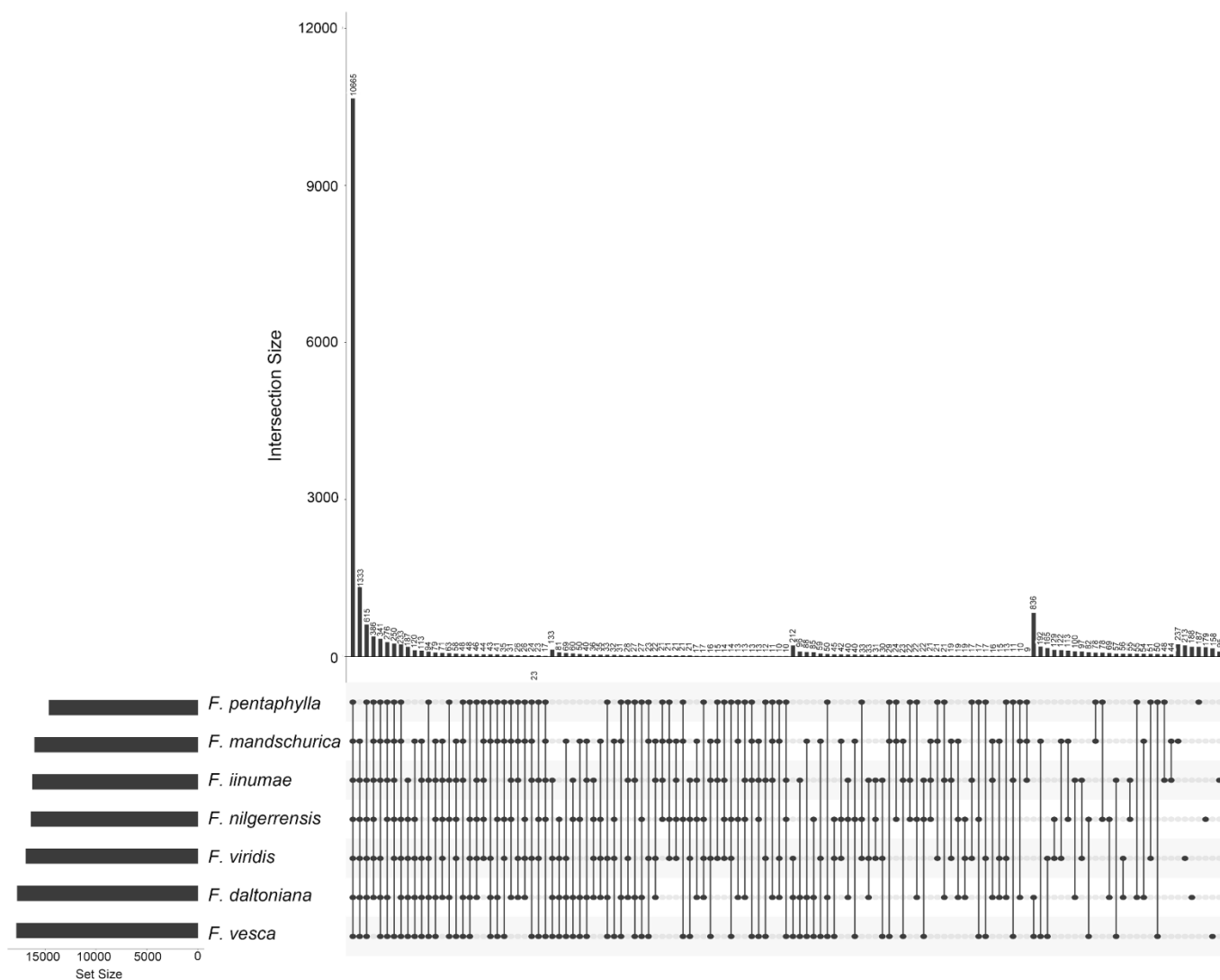


Fig. S3. Upset plot showing the unique and overlapping protein ortholog clusters across seven diploid *Fragaria* species. The intersection matrix is sorted in descending order. Connected dots represent intersections of overlapping orthogroups while vertical bars show the size of each intersection. The plot was generated using UpSetR package in R.

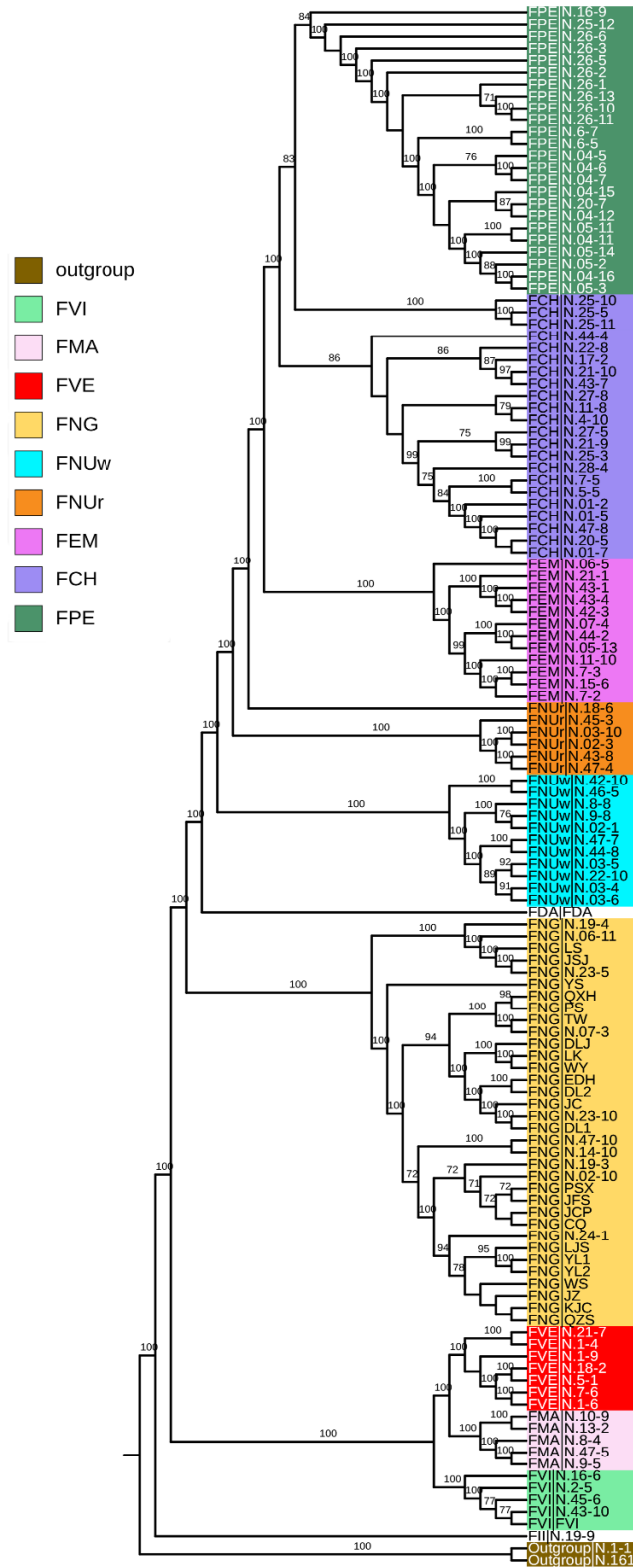


Fig. S4. The Maximum likelihood (ML) tree of diploid species in *Fragaria* with two outgroups based on whole-genome polymorphisms (the same as Fig.3A). *F. vesca* (FVE), *F. iinumae* (FII), *F. viridis* (FVI), *F. mandschurica* (FMA), *F. daltoniana* (FDA), *F. pentaphylla* (FPE), and *F. nilgerrensis* (FNG), *F. nubicola* (FNUw, white fruit type; FNUr, red fruit type).

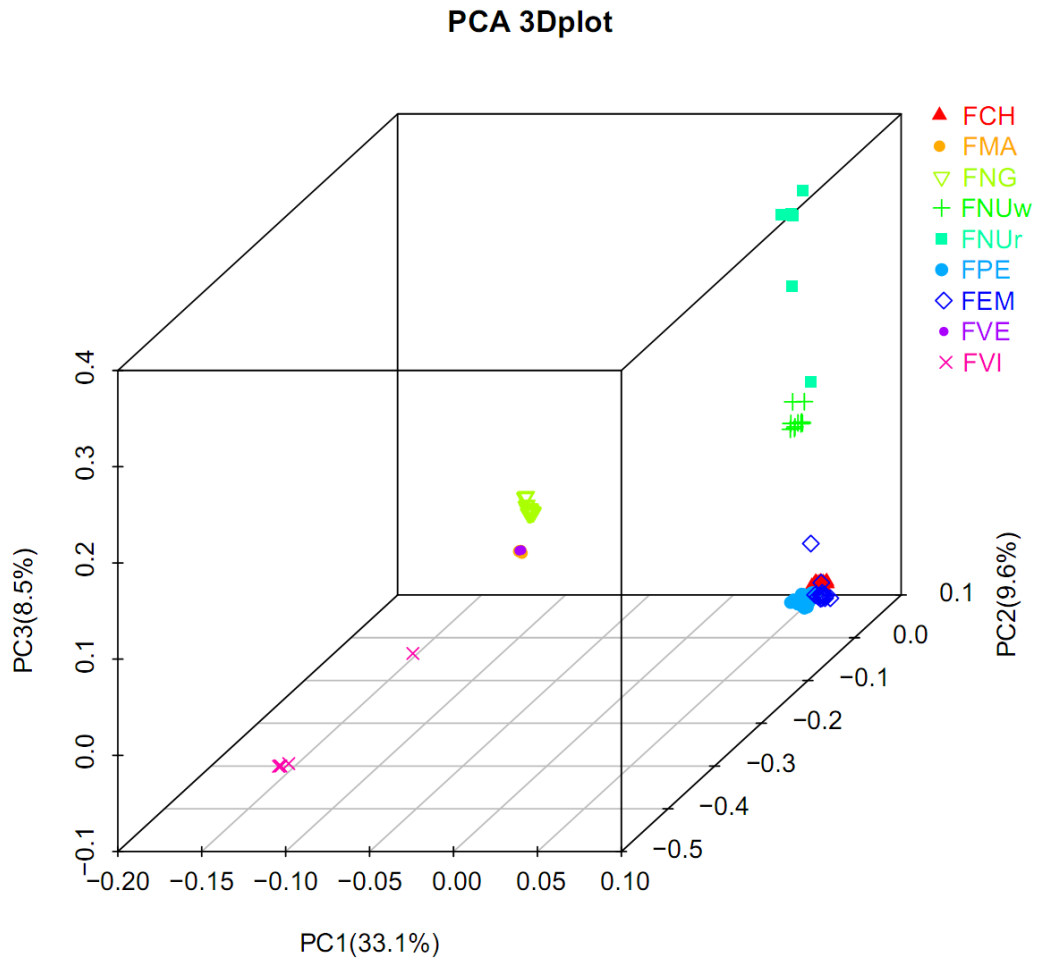


Fig. S5. Principle components analysis (PCA) of key species in *Fragaria* based on whole-genome polymorphisms.

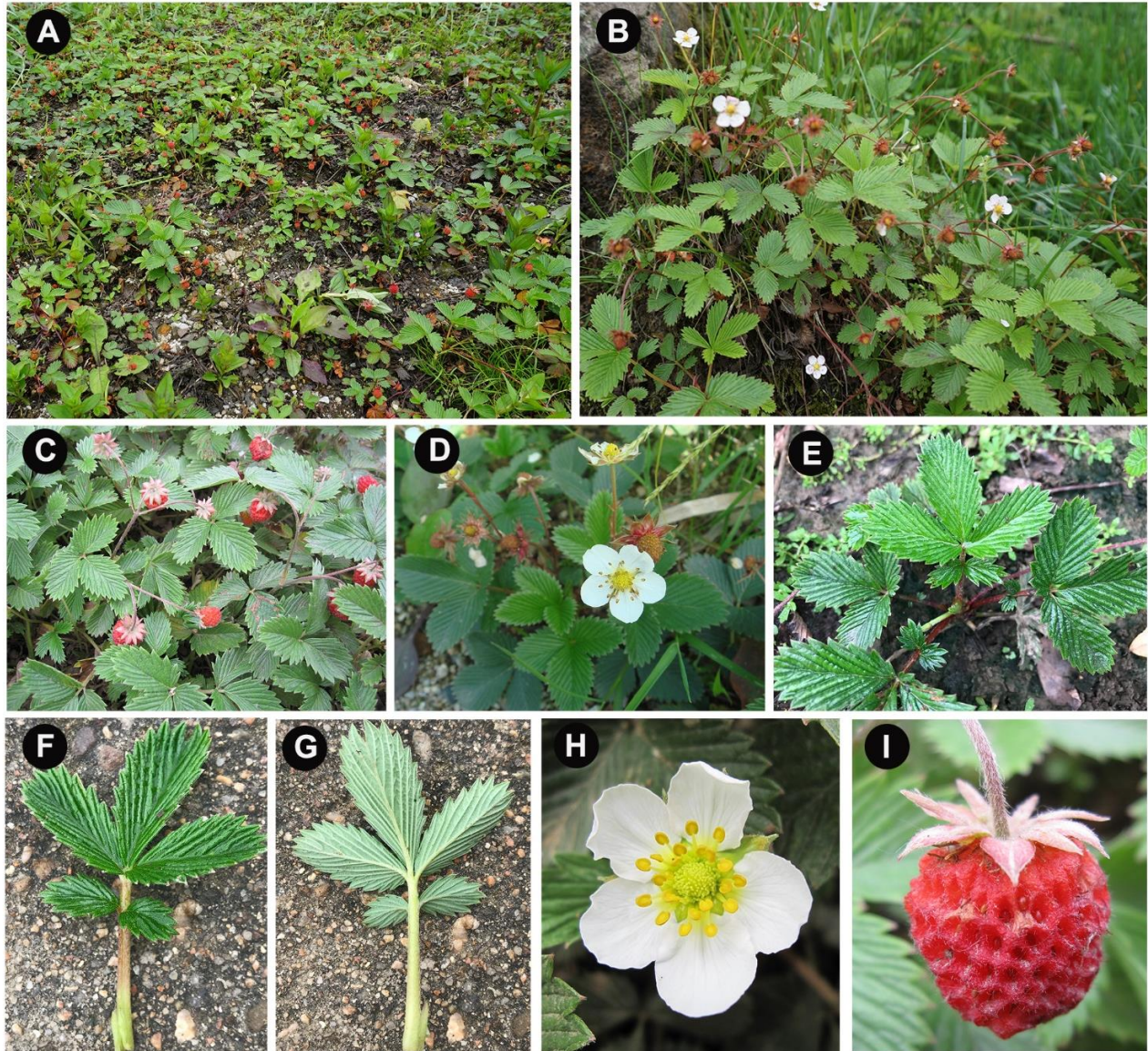


Fig. S6. Morphological features of the new species *Fragaria emeiensis* Jia J. Lei. (A, B) Habitat. (C-E) Plants. (F, G) Leaves. (H). Flower. (I) Fruit. (Photos by Jia J. Lei and Li Xue).

Taxonomic treatment

Fragaria emeiensis Jia J. Lei, *sp. nov.* (Figs. S6 and S7). **Type:** China, Sichuan Province, Jiulong Ping, Emei Mountain, 29° 32' 38.11" N, 103° 19' 23.07" E; altitude, 2243 m; 21 June 2017; Jia J. Lei EM35 (holotype: YUKU [02074626]!, isotype: YUKU [02074625]!).

Etymology: The specific epithet *emeiensis* is derived from the type locality of the new species, the Emei Mountain.

Diagnosis: *Fragaria emeiensis* is most similar to *F. nubicola*, *F. chinensis*, and *F. pentaphylla* in morphology. However, *F. emeiensis* has unique and stable morphological features and can be

distinguished from related species. Plants of *F. emeiensis* are robust with thick petiole, stolons, and leaves are also thick with deep veins, which separate it from its closely related species *F. nubicola*. Different from *F. chinensis* with trifoliolate leaves, *F. emeiensis* has pinnately quinquefoliolate leaves. In addition, petioles, stolons and peduncles of *F. emeiensis* are covered with appressed hairs, while these of *F. pentaphylla* are covered with spreading hairs. Another trait differentiating *F. emeiensis* from *F. pentaphylla* is the color of the abaxial leaf. The abaxial leaf of *F. emeiensis* is light green, while that of *F. pentaphylla* usually is green to purplish.

Plants perennial, 5~25 cm tall. **Leaves** usually pinnately quinquefoliolate, rarely trifoliolate, adaxially dark green, abaxially light green, with prominent veins; leaflets ovate, nearly sessile, or only central one shortly petiolulate. **Petioles** green or slightly light red, thick, with glabrous or sparsely appressed hairs. **Stolons** red, monopodial, with sparsely appressed hairs. **Inflorescence** corymbiform, higher than leaves, 2~6 flowered, peduncle and pedicel with sparsely appressed hairs. **Flowers** bisexual; sepals wide lanceolate, epicalyx lanceolate, apex occasionally 2-3 lobed; petals white, suborbicular or obovate, base not prominently clawed, apopetalous or slightly overlapping; stamens ca. 22, anthers relatively big, filament unequal in length, obviously longer than pistils. **Fruits** red, irregular oval or short conical, little juice, slightly acid or tasteless without fragrance; achenes red and sunken; persistent sepals spreading or slightly reflexed. $2n = 2x = 14$.

Distribution and habitat: *F. emeiensis* is only found in Emei Mountain of Sichuan Province, Southwest China. It usually grows with *F. nilgerrensis* Schlecht on hillside, grassland, and ditch, at altitudes of 1800–3000 m.

Phenology: Flowering from May to July, fruiting from June to August.

There are 14 recognized wild *Fragaria* species containing nine diploids (including *F. emeiensis* Jia J. Lei) and five tetraploids distributed in China. According to the ploidy examination and taxonomic observation, a revised key to the known *Fragaria* species of China is given as follows:

The dichotomous key of wild *Fragaria* species found in China

- 1 Stolon sympodial
 - 2 Plant hermaphroditic ($2n=2x=14$)
 - 3 Petioles, stolons and peduncles covered spreading hairs
 - 4 Fruit white; achenes extremely sunken; calyx clasping*F. nilgerrensis*
 - 4 Fruit red; achenes raised; calyx spreading or lightly reflexed*F. mandschurica*
 - 3 Petioles covered spreading hairs, runners and peduncles covered appressed hairs
 - 5 Leaf green, non-coreacious, non-shiny; fruit long conic; calycle with entire apex.....*F. vesca*
 - 5 Leaf dark green, coreacious, shiny; fruit spindly; calycle with 2 or 3 lobed apex.....*F. daltoniana*
 - 2 Plant dioecious, rarely sub-dioecious ($2n=4x=28$).....*F. orientalis*
- 1 Stolon monopodial
 - 6 Plant hermaphroditic ($2n=2x=14$)
 - 7 Leaf trifoliolate

- 8 Fruit light green to reddish-green, large, hard; calyx clasping; achenes raised.....*F. viridis*
- 8 Fruit light red to red, small, soft; calyx spreading; seed sunken.....*F. chinensis*
- 7 Leaf pinnately quinquefoliate, rarely trifoliate
 - 9 Petioles and runners covered spreading hairs, but sparse hairs on peduncles; calyx reflexed...*F. pentaphylla*
 - 9 Petioles, stolons and peduncles covered appressed hairs; calyx clasping or spreading
 - 10 Petiole, stolons, leaf relative thin; leaf green; calyx clasping.....*F. nubicola*
 - 10 Petiole, stolons, leaf obvious thick; leaf deep green; calyx spreading.....*F. emeiensis*
- 6 Plant dioecious, rarely sub-dioecious ($2n=4x=28$)
 - 11 Leaf trifoliate.....*F. gracilis*
 - 11 Leaf pinnately quinquefoliate, rarely trifoliate
 - 12 Petioles, stolons and peduncles covered appressed hairs.....*F. tibetica*
 - 12 Petioles, stolons and peduncles covered spreading hairs
 - 13 Hair long and sparse.....*F. corymbosa*
 - 13 Hair short and thick.....*F. moupinensis*



Fig. S7. Type specimen of the new species *F. emeiensis* Jia J. Lei. Left, holotypus; right, isotypus.

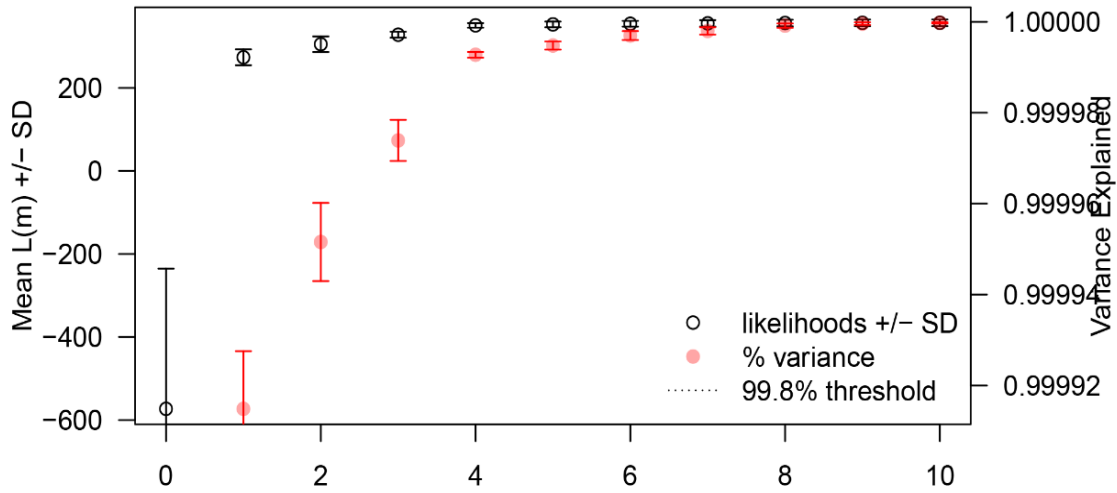


Fig. S8. The OptM R package (<https://cran.r-project.org/web/packages/OptM>) was used to determine the optimal number of migration edges. To induce enough variation to assess an optimal model, TreeMix was run with 0-10 migration edges in 20 iterations, starting from $-k = 200$ to 4000, with an increment of 200 per iteration. The output files from TreeMix were used as input for OptM, and we used the Evanno method to estimate the proportion of variance explained by different number of migration edges. The results showed that the repeatability of $-m = 4$ was high and the standard errors of migration weights was small.

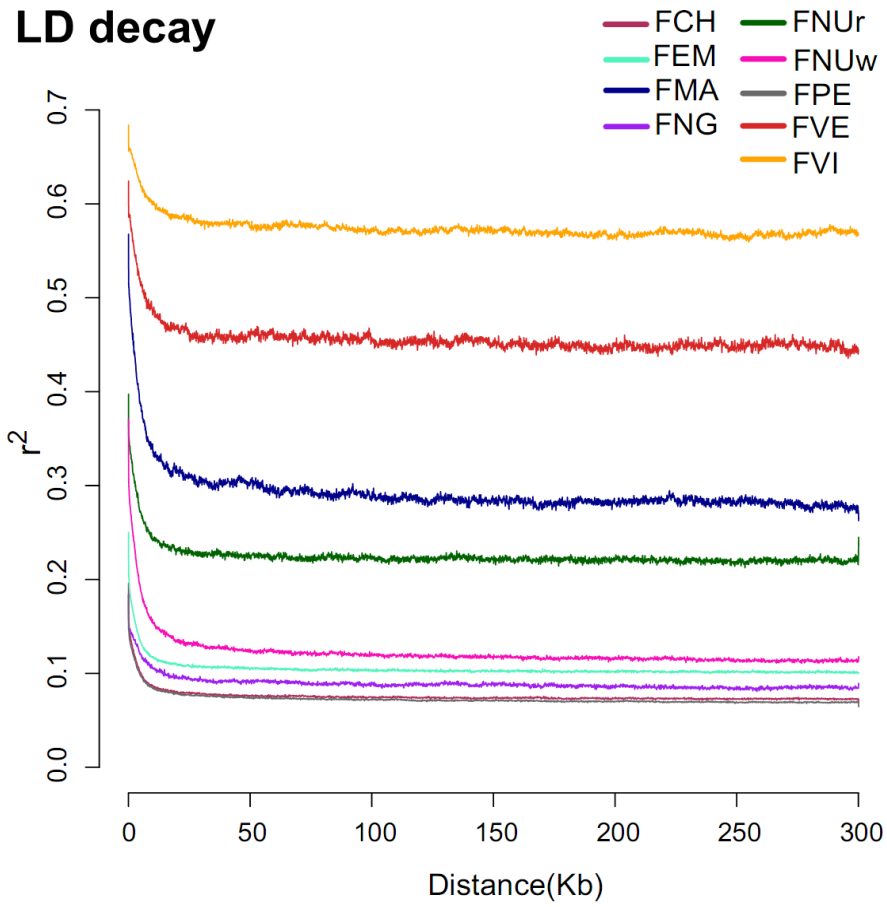


Fig. S9. Linkage disequilibrium decay of eight *Fragaria* diploid species (two types of *F. nubicola* were measured respectively).

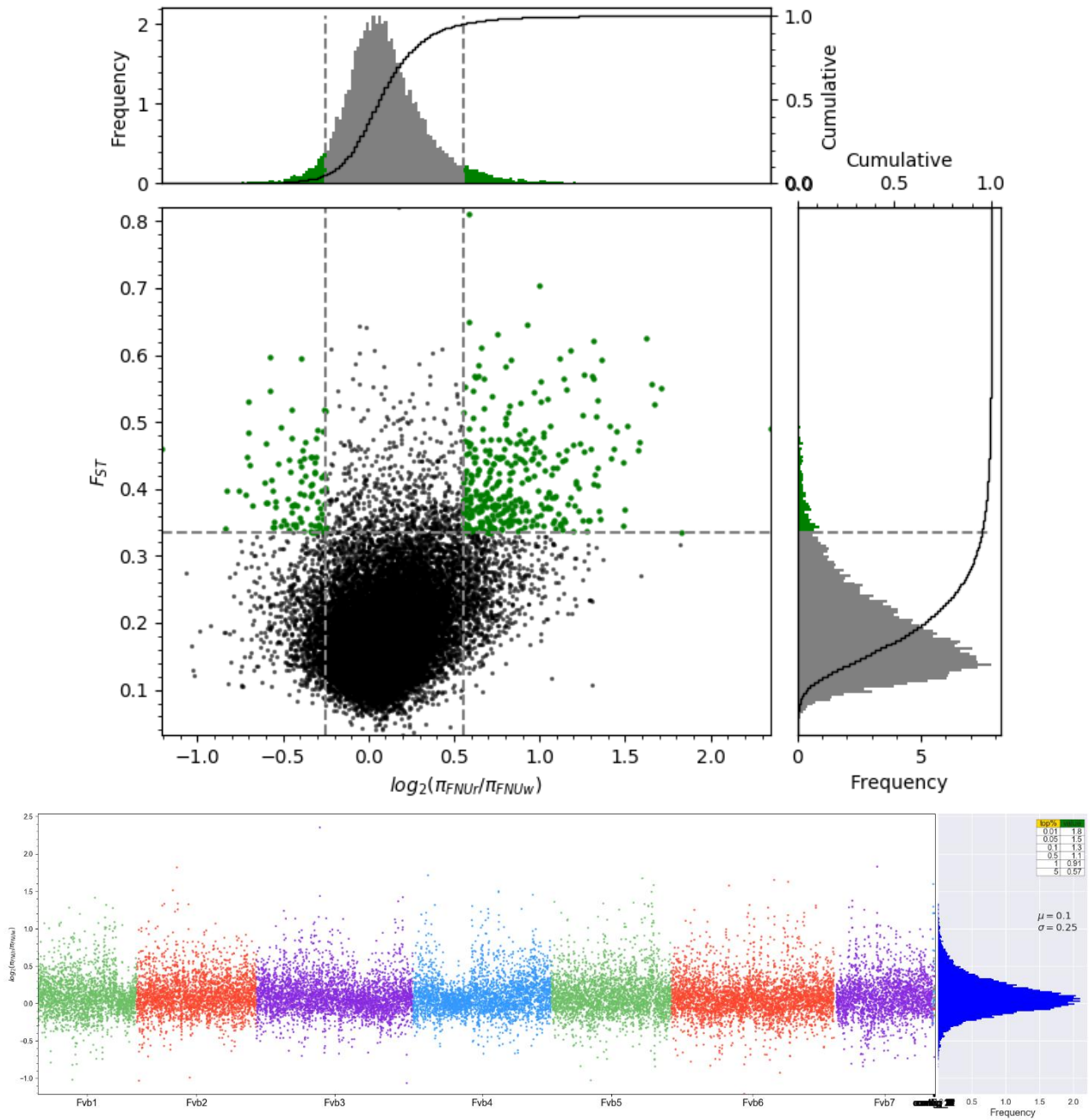


Fig. S10. The selective sweep analysis between two fruit types of *F. nubicola*. Up: The vertical and horizontal dashed lines correspond to the 5% right tails of the F_{ST} and θ_π value distribution, respectively. Below: Manhattan plots of θ_π ratios between red and white fruit types of *F. nubicola*.

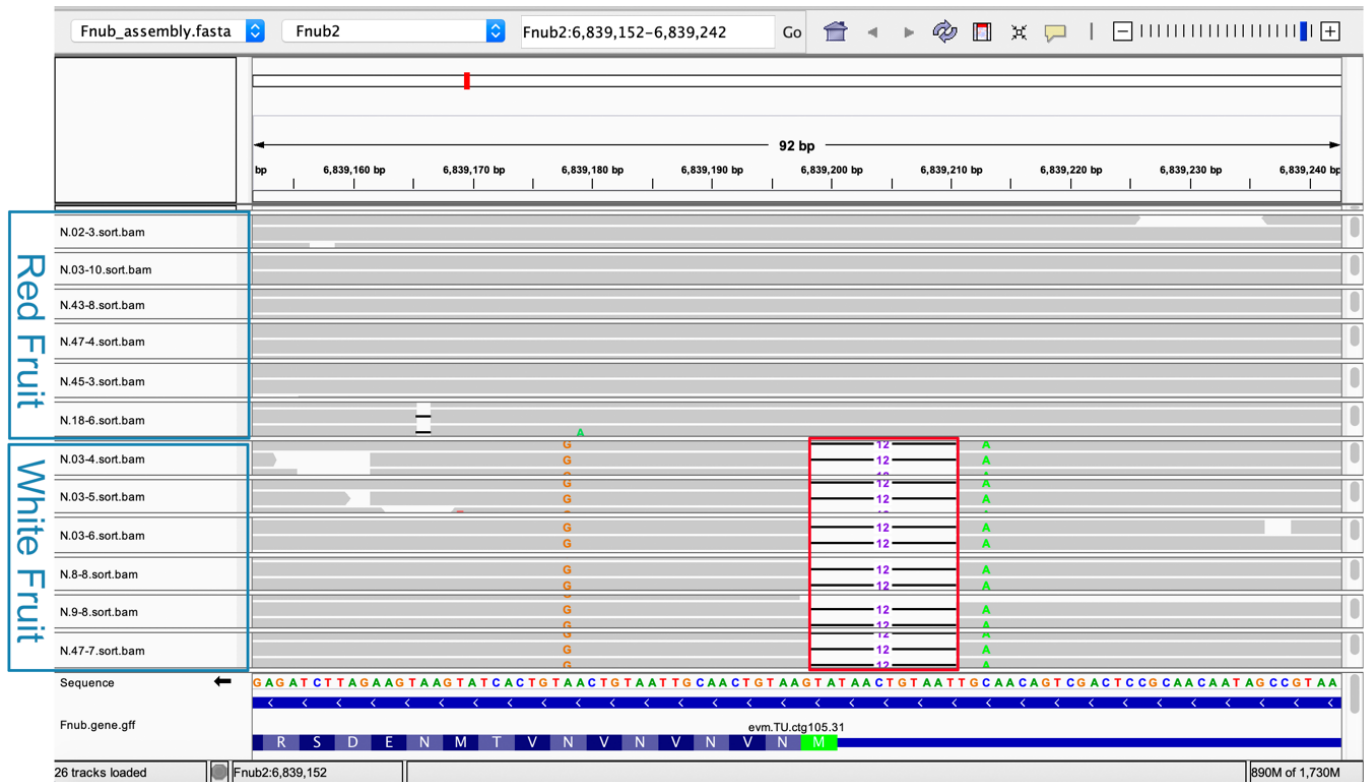


Fig. S11. Integrative genomics viewer (IGV) images of the mutation in *FY* gene between white and red fruit accessions of *F. nubicola*. The 12 nucleotide deletion of *FY* in all the white fruit individuals of *F. nubicola* that altered the original start codon ATG>GTG, resulted in the subsequent 10 amino acids are not translated.

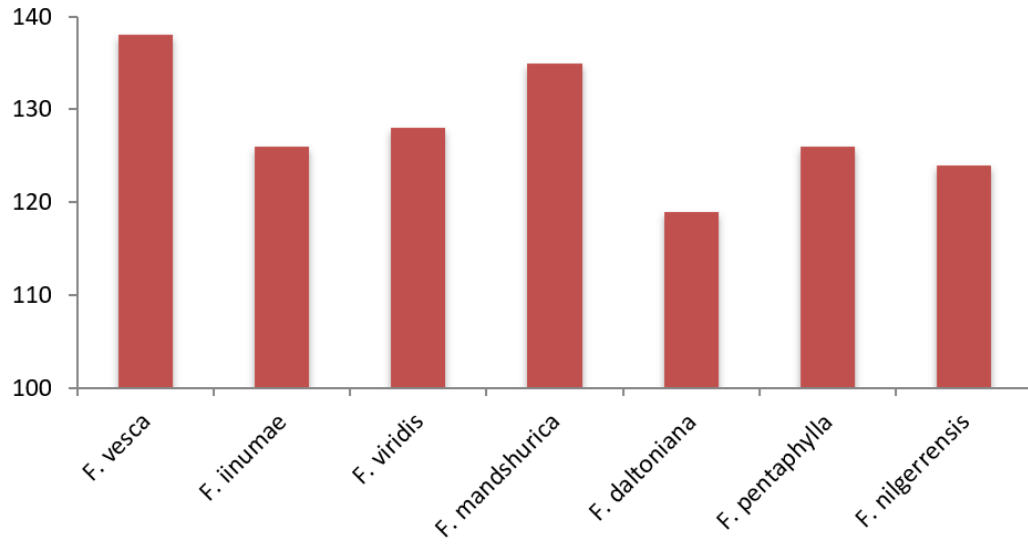


Fig. S12. Number of R2R3-MYB gene family in diploid wild species of *Fragaria*.



Fig. S13. Integrative genomics viewer (IGV) images present the polymorphisms in coding regions of *MYB10* between white and red fruits phenotypes for each species using re-sequencing data. A. The presence of mutation Trp12Ser in white fruit accessions of *F. vesca* (Noting the reference genome of *F. vesca* is Hawaii4, a white fruit type). B. The presence of mutation Val122Gly in white fruit accessions of *F. nubicola*. C. The

mutation R82S was exclusively discovered in all accessions of *F. nilgerrensis* comparing with other five species. D. The presence of two heterozygous mutations in red fruit accessions of *F. pentaphylla* (Considering the overall fragmented state of the assembly of the *F. pentaphylla* genome, the short reads were mapping to the reference genome of *F. vesca*).

Supplementary Tables

Table S1. Non-coding RNA, repeat elements and gene features of five newly sequenced species and *F. iinumae*.

Non-coding RNA	<i>F. iinumae</i> (Num.)	<i>F. nilgerrensis</i> (Num.)	<i>F. daltoniana</i> (Num.)	<i>F. mandshurica</i> (Num.)	<i>F. pentaphylla</i> (Num.)	<i>F. viridis</i> (Num.)
miRNA	280	364	225	386	454	411
tRNA	461	489	786	534	601	651
rRNA	128	301	102	33	67	54
snRNA	545	516	619	474	511	400
Repeat elements	<i>F. iinumae</i> (Mb/percent)	<i>F. nilgerrensis</i> (Mb/percent)	<i>F. daltoniana</i> (Mb/percent)	<i>F. mandshurica</i> (Mb/percent)	<i>F. pentaphylla</i> (Mb/percent)	<i>F. viridis</i> (Mb/percent)
DNA transposon	12.41/5.16%	40.43/14.79%	40.490/14.01%	27.1/11.30%	33.98/12.17%	13.46/6.04%
LINE retrotransposon	3.59/1.49%	4.44/1.62%	0.56/1.02%	4.39/1.83%	6.31/2.26%	2.25/1.01%
SINE retrotransposon	0.012/0.005%	0.12/0.043%	0.04/0.002%	0.04/0.02%	0.28/0.10%	0.13/0.06%
LTR retrotransposon	81.24/33.77%	72.17/26.40%	65.08/22.52%	53.04/22.11%	67.61/24.22%	61.89/27.74%
Unclassified transposon	3.82/1.59%	7.12/2.60%	20.2/7.00%	7.37/3.07%	10.36/3.71%	9.35/4.19%
Total transposable elements	99.81/41.49%	118.91/43.50%	140.85/48.74%	88.68/36.97%	115.05/41.21%	84.62/37.93%
Total repeat elements	102.72/42.70%	122.16/44.69%	140.86/48.74%	91.69/38.23%	118.41/42.42%	87.23/39.10%
Gene features	<i>F. iinumae</i>	<i>F. nilgerrensis</i>	<i>F. daltoniana</i>	<i>F. mandshurica</i>	<i>F. pentaphylla</i>	<i>F. viridis</i>
Gene Number	23,665	24,491	28,131	25,411	23,853	24,779
Average gene length (bp)	2,686.96	2,651.89	2,918.41	2,648.36	2,772.43	2,802.77
Average CDS length	1,254.27	1,164.75	1,121.58	1,140.46	1,189.48	1,169.24
Average exons per gene	5.09	5	4.88	5.02	5.14	5.08
Average exon length	246.48	232.72	229.99	227.34	231.38	230.31

Table S2. Changes in orthogroup (gene family) size of diploid species in *Fragaria*.

Species change	Expansions			Contractations			No changes	Avg. Expansion
	Families	Genes gained	Genes/Expansion	Families	Genes Lost	Genes/Contraction		
<i>F. iinumae</i>	873 (46)	1418	1.62	1982 (20)	2281	1.15	13502	-0.0528
<i>F. daltoniana</i>	1536 (188)	2849	1.85	1426 (14)	1537	1.08	13395	0.0802
<i>F. pentaphylla</i>	2751 (32)	3465	1.26	3771 (108)	4714	1.25	9835	-0.0764
<i>F. vesca</i>	950 (221)	2205	2.32	1291 (44)	1413	1.09	14116	0.0484
<i>F. nilgerrensis</i>	1177 (50)	1629	1.38	2070 (56)	2580	1.25	13110	-0.0581
<i>F. mandschurica</i>	2476 (60)	2930	1.18	2246 (206)	2948	1.31	11635	-0.0011
<i>F. viridis</i>	1066 (90)	1658	1.56	1730 (96)	2419	1.4	13561	-0.0465

Table S3. The enriched gene functional categories of PSGs in key species.

GO category	P value	PSG Num.	Num. in Category	Term
<i>F. daltoniana</i>				
GO:0009451	4.40E-05	81	564	RNA modification
GO:0036503	0.000198957	13	46	ERAD pathway
GO:0030433	0.000976759	10	35	ubiquitin-dependent ERAD pathway
GO:0034765	0.001187331	9	30	regulation of ion transmembrane transport
GO:0090305	0.001538638	80	621	nucleic acid phosphodiester bond hydrolysis
GO:0009856	0.001616335	32	200	pollination
GO:0044706	0.001616335	32	200	multi-multicellular organism process
GO:0016554	0.002250761	6	16	cytidine to uridine editing
GO:0000245	0.002419511	8	27	spliceosomal complex assembly
GO:0034762	0.002482007	9	33	regulation of transmembrane transport
GO:0006376	0.003122093	5	12	mRNA splice site selection
GO:0016553	0.00320772	6	17	base conversion or substitution editing
GO:0034976	0.003444628	14	68	response to endoplasmic reticulum stress
GO:0099111	0.003808174	7	23	microtubule-based transport
GO:0043269	0.004564754	11	49	regulation of ion transport
GO:0043266	0.004688621	5	13	regulation of potassium ion transport
GO:0044257	0.005150422	62	482	cellular protein catabolic process
GO:0048588	0.005431138	22	133	developmental cell growth
GO:0051865	0.006125678	10	44	protein autoubiquitination
GO:0016567	0.006384216	63	496	protein ubiquitination
GO:0051603	0.006885812	60	470	proteolysis involved in cellular protein catabolic process
GO:0048868	0.007727489	22	137	pollen tube development
GO:0006511	0.008011923	53	409	ubiquitin-dependent protein catabolic process
GO:0019941	0.008799367	53	411	modification-dependent protein catabolic process
GO:0010498	0.009087424	40	294	proteasomal protein catabolic process
GO:0071218	0.009349734	5	15	cellular response to misfolded protein
GO:0051240	0.009460825	24	156	positive regulation of multicellular organismal process
GO:0031330	0.009759555	9	40	negative regulation of cellular catabolic process
<i>F. nilgerrensis</i>				
GO:0006084	0.001124497	4	10	acetyl-CoA metabolic process
GO:0019287	0.001243766	3	5	isopentenyl diphosphate biosynthetic process, mevalonate pathway
GO:0030520	0.002628838	2	2	intracellular estrogen receptor signaling pathway
GO:0000729	0.003384263	4	13	DNA double-strand break processing
GO:0002097	0.004028184	3	7	tRNA wobble base modification
GO:0071902	0.005088508	7	42	positive regulation of protein serine/threonine kinase activity
GO:0055071	0.006200822	3	8	manganese ion homeostasis
GO:0006097	0.00761758	2	3	glyoxylate cycle
GO:0046901	0.00761758	2	3	tetrahydrofolylpolyglutamate biosynthetic process
GO:0001934	0.008979087	10	82	positive regulation of protein phosphorylation
GO:0033865	0.009378272	5	26	nucleoside bisphosphate metabolic process
GO:0033875	0.009378272	5	26	ribonucleoside bisphosphate metabolic process
GO:0034032	0.009378272	5	26	purine nucleoside bisphosphate metabolic process
GO:0006760	0.009566861	4	17	folic acid-containing compound metabolic process
<i>F. pentaphylla</i>				
GO:0042762	0.002044865	7	20	regulation of sulfur metabolic process
GO:1902074	0.003864553	5	12	response to salt
GO:0043069	0.00605271	10	42	negative regulation of programmed cell death
GO:0043476	0.007969851	3	5	pigment accumulation

GO:0043648	0.008204364	8	31	dicarboxylic acid metabolic process
GO:0006435	0.009517335	2	2	threonyl-tRNA aminoacylation
GO:0033356	0.009517335	2	2	UDP-L-arabinose metabolic process
GO:0043482	0.009517335	2	2	cellular pigment accumulation
GO:0043485	0.009517335	2	2	endosome to pigment granule transport
GO:0048757	0.009517335	2	2	pigment granule maturation
GO:0090548	0.009517335	2	2	response to nitrate starvation
GO:1901333	0.009517335	2	2	positive regulation of lateral root development
GO:1900618	0.011438352	5	15	regulation of shoot system morphogenesis
GO:0060688	0.011706365	4	10	regulation of morphogenesis of a branching structure
GO:2000032	0.011706365	4	10	regulation of secondary shoot formation
GO:0000447	0.014788585	3	6	endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA
GO:0002444	0.014788585	3	6	myeloid leukocyte mediated immunity
GO:0002446	0.014788585	3	6	neutrophil mediated immunity
GO:0043299	0.014788585	3	6	leukocyte degranulation
GO:0043312	0.014788585	3	6	neutrophil degranulation
GO:2000023	0.014788585	3	6	regulation of lateral root development
GO:0055086	0.01485758	29	196	nucleobase-containing small molecule metabolic process
GO:0019674	0.016995994	4	11	NAD metabolic process
GO:0043473	0.016995994	4	11	pigmentation
GO:0006417	0.018960871	16	94	regulation of translation
<i>F. vesca</i>				
GO:0042891	0.000158033	4	8	antibiotic transport
GO:0040011	0.00040459	12	95	locomotion
GO:0006020	0.000444534	4	10	inositol metabolic process
GO:0044273	0.001526767	5	22	sulfur compound catabolic process
GO:0015746	0.001609209	2	2	citrate transport
GO:0033869	0.001609209	2	2	nucleoside bisphosphate catabolic process
GO:0034031	0.001609209	2	2	ribonucleoside bisphosphate catabolic process
GO:0034034	0.001609209	2	2	purine nucleoside bisphosphate catabolic process
GO:0071260	0.001609209	2	2	cellular response to mechanical stimulus
GO:0006655	0.001995552	3	7	phosphatidylglycerol biosynthetic process
GO:1901606	0.003274366	6	37	alpha-amino acid catabolic process
GO:0010183	0.003967356	5	27	pollen tube guidance
GO:0050918	0.003967356	5	27	positive chemotaxis
GO:0002213	0.004027235	4	17	defense response to insect
GO:0005984	0.004036444	7	51	disaccharide metabolic process
GO:0006842	0.0046989	2	3	tricarboxylic acid transport
GO:0009154	0.0046989	2	3	purine ribonucleotide catabolic process
GO:0009261	0.0046989	2	3	ribonucleotide catabolic process
GO:0035674	0.0046989	2	3	tricarboxylic acid transmembrane transport
GO:0050848	0.0046989	2	3	regulation of calcium-mediated signaling
GO:0070059	0.0046989	2	3	intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress
GO:0071963	0.0046989	2	3	establishment or maintenance of cell polarity regulating cell shape
GO:0007283	0.005015776	4	18	spermatogenesis
GO:0009063	0.005533221	6	41	cellular amino acid catabolic process
GO:0010020	0.006154713	4	19	chloroplast fission
GO:0043572	0.006154713	4	19	plastid fission
GO:0046471	0.00625145	3	10	phosphatidylglycerol metabolic process
GO:0000070	0.006344754	5	30	mitotic sister chromatid segregation

GO:0043414	0.007764244	14	168	macromolecule methylation
GO:0006305	0.007852791	6	44	DNA alkylation
GO:0006306	0.007852791	6	44	DNA methylation
GO:0000819	0.008761657	6	45	sister chromatid segregation
GO:0048870	0.008761657	6	45	cell motility
GO:0051674	0.008761657	6	45	localization of cell
GO:0055114	0.008814111	50	886	oxidation-reduction process
GO:0001817	0.008920003	4	21	regulation of cytokine production
GO:0010155	0.009148056	2	4	regulation of proton transport
GO:0031032	0.009148056	2	4	actomyosin structure organization
GO:0045691	0.009148056	2	4	regulation of embryo sac central cell differentiation
GO:0045697	0.009148056	2	4	regulation of synergid differentiation
GO:0050900	0.009148056	2	4	leukocyte migration
GO:0019751	0.009576487	5	33	polyol metabolic process
<i>F. viridis</i>				
GO:0007098	0.000181572	4	7	centrosome cycle
GO:0051298	0.000181572	4	7	centrosome duplication
GO:0010165	0.001106151	3	5	response to X-ray
GO:0031023	0.001460849	4	11	microtubule organizing center organization
GO:0000076	0.002131324	3	6	DNA replication checkpoint
GO:0006808	0.002131324	3	6	regulation of nitrogen utilization
GO:0090307	0.002391481	5	20	mitotic spindle assembly
GO:0009850	0.003073856	7	40	auxin metabolic process
GO:0001894	0.003593637	3	7	tissue homeostasis
GO:0000959	0.003667676	8	52	mitochondrial RNA metabolic process
GO:0007052	0.004229431	6	32	mitotic spindle organization
GO:0044242	0.004370494	10	77	cellular lipid catabolic process
GO:0000373	0.004592958	5	23	Group II intron splicing
GO:0060249	0.004677593	7	43	anatomical structure homeostasis
GO:1902850	0.004958345	6	33	microtubule cytoskeleton organization involved in mitosis
GO:0051415	0.005540426	3	8	microtubule nucleation by interphase microtubule organizing center
GO:0051418	0.005540426	3	8	microtubule nucleation by microtubule organizing center
GO:1900864	0.005775401	6	34	mitochondrial RNA modification
GO:0016042	0.006281769	10	81	lipid catabolic process
GO:0000226	0.00658199	13	121	microtubule cytoskeleton organization
GO:0006641	0.006686214	5	25	triglyceride metabolic process
GO:0031570	0.006686214	5	25	DNA integrity checkpoint
GO:0010896	0.007040695	2	3	regulation of triglyceride catabolic process
GO:0018175	0.007040695	2	3	protein nucleotidylation
GO:0018177	0.007040695	2	3	protein uridylylation
GO:0018871	0.007040695	2	3	1-aminocyclopropane-1-carboxylate metabolic process
GO:0042218	0.007040695	2	3	1-aminocyclopropane-1-carboxylate biosynthetic process
GO:0042776	0.007040695	2	3	mitochondrial ATP synthesis coupled proton transport
GO:0045103	0.007040695	2	3	intermediate filament-based process
GO:0045104	0.007040695	2	3	intermediate filament cytoskeleton organization
GO:0048478	0.007040695	2	3	replication fork protection
GO:0048873	0.007040695	2	3	homeostasis of number of cells within a tissue
GO:0060052	0.007040695	2	3	neurofilament cytoskeleton organization
GO:0090293	0.007040695	2	3	nitrogen catabolite regulation of transcription
GO:0006638	0.00769738	6	36	neutral lipid metabolic process
GO:0006639	0.00769738	6	36	acylglycerol metabolic process
GO:0051225	0.007949352	5	26	spindle assembly

GO:0080156	0.007949352	5	26	mitochondrial mRNA modification
GO:0006297	0.008008771	3	9	nucleotide-excision repair, DNA gap filling
GO:0007368	0.008008771	3	9	determination of left/right symmetry
GO:0050994	0.008008771	3	9	regulation of lipid catabolic process
GO:0051103	0.008008771	3	9	DNA ligation involved in DNA repair
GO:0090693	0.009038767	12	112	plant organ senescence
GO:0019725	0.009264552	16	169	cellular homeostasis
<i>F. mandschurica</i>				
GO:0006376	0.009775373	5	17	mRNA splice site selection
GO:0043484	0.009587093	9	45	regulation of RNA splicing
GO:0043543	0.009560081	12	69	protein acylation
GO:0140027	0.009048594	3	6	establishment of contractile vacuole localization
GO:0090239	0.009048594	3	6	regulation of histone H4 acetylation
GO:0009812	0.008659075	11	60	flavonoid metabolic process
GO:0022412	0.008597158	17	111	cellular process involved in reproduction in multicellular organism
GO:0010431	0.007619504	11	59	seed maturation
GO:0006284	0.007583165	7	29	base-excision repair
GO:0060193	0.006689093	2	2	positive regulation of lipase activity
GO:0060191	0.006689093	2	2	regulation of lipase activity
GO:0043496	0.006689093	2	2	regulation of protein homodimerization activity
GO:0010518	0.006689093	2	2	positive regulation of phospholipase activity
GO:0010517	0.006689093	2	2	regulation of phospholipase activity
GO:0006473	0.005899916	12	65	protein acetylation
GO:0055046	0.005833069	11	57	microgametogenesis
GO:0009226	0.005499729	6	21	nucleotide-sugar biosynthetic process
GO:0046368	0.004816544	3	5	GDP-L-fucose metabolic process
GO:0042350	0.004816544	3	5	GDP-L-fucose biosynthetic process
GO:0000245	0.004276285	8	33	spliceosomal complex assembly
GO:0045490	0.004017607	4	9	pectin catabolic process
GO:0043968	0.004017607	4	9	histone H2A acetylation
GO:0006475	0.002988689	12	60	internal protein amino acid acetylation
GO:0007276	0.002797601	18	108	gamete generation
GO:0018394	0.002578709	12	59	peptidyl-lysine acetylation
GO:0018393	0.002578709	12	59	internal peptidyl-lysine acetylation
GO:1901605	0.002484685	25	168	alpha-amino acid metabolic process
GO:0048235	0.002425939	9	37	pollen sperm cell differentiation
GO:0046283	0.002409172	7	24	anthocyanin-containing compound metabolic process
GO:0048232	0.002405073	16	90	male gamete generation
GO:0009649	0.002385811	4	8	entrainment of circadian clock
GO:0010288	0.002051649	3	4	response to lead ion
GO:0050684	0.002036291	10	43	regulation of mRNA processing
GO:0044703	0.001765554	29	200	multi-organism reproductive process
GO:0016573	0.001614203	12	56	histone acetylation
GO:0035264	0.001275373	4	7	multicellular organism growth
GO:0043967	0.001060948	8	27	histone H4 acetylation
GO:0019953	0.000841401	28	182	sexual reproduction
GO:0048024	0.000613673	9	31	regulation of mRNA splicing, via spliceosome
GO:0032504	0.000247319	34	221	multicellular organism reproduction
GO:0048609	6.39E-05	33	198	multicellular organismal reproductive process
<i>F. iinumae</i>				
GO:0019438	0.000784541	40	625	aromatic compound biosynthetic process
GO:0060142	0.001433367	2	2	regulation of syncytium formation by plasma membrane fusion

GO:0060143	0.00143367	2	2	positive regulation of syncytium formation by plasma membrane fusion
GO:1901739	0.00143367	2	2	regulation of myoblast fusion
GO:1901741	0.00143367	2	2	positive regulation of myoblast fusion
GO:1990089	0.00143367	2	2	response to nerve growth factor
GO:1990090	0.00143367	2	2	cellular response to nerve growth factor stimulus
GO:2001135	0.00143367	2	2	regulation of endocytic recycling
GO:2001137	0.00143367	2	2	positive regulation of endocytic recycling
GO:0034654	0.00146494	28	403	nucleobase-containing compound biosynthetic process
GO:0043966	0.001788665	5	24	histone H3 acetylation
GO:0030091	0.001992034	4	15	protein repair
GO:1901362	0.002388004	42	706	organic cyclic compound biosynthetic process
GO:0044271	0.00326089	45	783	cellular nitrogen compound biosynthetic process
GO:0032774	0.004069594	17	219	RNA biosynthetic process
GO:0010831	0.004192805	2	3	positive regulation of myotube differentiation
GO:0051149	0.004192805	2	3	positive regulation of muscle cell differentiation
GO:0051155	0.004192805	2	3	positive regulation of striated muscle cell differentiation
GO:0061512	0.004192805	2	3	protein localization to cilium
GO:0006351	0.004526412	16	203	transcription, DNA-templated
GO:0018130	0.004841192	34	563	heterocycle biosynthetic process
GO:0006366	0.004884649	12	134	transcription by RNA polymerase II
GO:0016444	0.005318851	3	10	somatic cell DNA recombination
GO:0070534	0.005318851	3	10	protein K63-linked ubiquitination
GO:0009867	0.005718381	8	72	jasmonic acid mediated signaling pathway
GO:0097659	0.005977742	16	209	nucleic acid-templated transcription
GO:0006352	0.007427447	6	46	DNA-templated transcription, initiation
GO:0010830	0.008175309	2	4	regulation of myotube differentiation
GO:0051153	0.008175309	2	4	regulation of striated muscle cell differentiation
GO:0009059	0.009637075	45	832	macromolecule biosynthetic process

Table S4. Statistics of re-sequenced data of 128 samples in eight diploid species.

Species	Sample code	Clean reads	Mapped reads	Mapping rate	Average depth	Coverage rate (>4X)
<i>F. chinensis</i>	N.43-7	39785268	30073665	0.7559	19.22	0.7181
	N.44-4	36297886	30768905	0.8477	19.68	0.7211
	N.47-8	33597464	28606498	0.8514	18.62	0.7131
	N.11-8	47234470	39158980	0.829	25.66	0.7449
	N.20-5	36476624	30347517	0.832	19.5	0.7194
	N.01-2	33640024	25821525	0.7676	16.69	0.7101
	N.01-7	39767218	31837641	0.8006	21.05	0.7262
	N.4-10	75023254	65179979	0.8688	41.01	0.7681
	N.21-10	75364944	64621269	0.8574	40.71	0.7733
	N.21-9	38391858	33750949	0.8791	21.8	0.7283
	N.7-5	66848436	59360816	0.888	38.71	0.7634
	N.22-8	37451080	33050105	0.8825	21.73	0.7293
	N.5-5	56975470	50110059	0.8795	31.46	0.7513
	N.25-3	35440670	30551680	0.8621	19.22	0.7153
	N.27-5	48264510	40909933	0.8476	25.12	0.7436
	N.27-8	46309484	38604052	0.8336	23.77	0.7275
	N.28-4	46612642	39802476	0.8539	24.6	0.7387
	N.01-5	38768748	33939682	0.8754	21.37	0.7197
	N.17-2	42796480	37451983	0.8751	24.43	0.7372
	N.25-5	38598296	33930116	0.8791	21.21	0.7204
	N.25-10	37866098	32667830	0.8627	20.82	0.7276
	N.25-11	41398386	35986508	0.8693	22.14	0.7285
<i>F. mandshurica</i>	N.8-4	34892472	32626173	0.935	20.32	0.867
	N.9-5	41141298	38172695	0.9278	23.11	0.8743
	N.10-9	58534850	54968249	0.9391	32.04	0.9023
	N.13-2	49539654	46933944	0.9474	27.79	0.8962
	N.47-5	50351426	46088347	0.9153	27.51	0.8818
<i>F. nilgerrensis</i>	N.14-10	78375310	68186897	0.87	41.33	0.7539
	DLJ	42778212	32084872	0.75	20.56	0.7042
	LS	39324568	33358908	0.8483	21.5	0.7041
	JSJ	35663028	29529998	0.828	19.2	0.6958
	WY	37136678	30213810	0.8136	20.2	0.7057
	LJS	57166260	46068176	0.8059	29.66	0.7364
	QXH	31095718	24472271	0.787	16.14	0.6905
	PS	42379748	33803444	0.7976	21.76	0.7126
	PSX	45994308	39843816	0.8663	25.42	0.7217
	JC	42943708	37886905	0.8822	24.18	0.7203
	WS	50670144	43805773	0.8645	28.01	0.731
	YS	43834506	37855582	0.8636	24.01	0.7284
	JCP	46411142	40158547	0.8653	25.75	0.7199
	JZ	41729494	35440644	0.8493	23	0.7194

	DL1	48086264	39386159	0.8191	25.51	0.7213
	DL2	47415276	35653201	0.7519	23.14	0.7168
	JFS	40961160	30114620	0.7352	19.5	0.7032
	QZS	30981602	26006370	0.8394	17.33	0.69
	YL1	38056904	30860573	0.8109	20.22	0.7071
	YL2	37725236	30108086	0.7981	19.78	0.7032
	LK	43816348	37807000	0.8629	24.38	0.7157
	EDH	44215554	37566878	0.8496	24.17	0.7215
	KJC	43582332	33325519	0.7647	21.67	0.7123
	TW	42998194	31600056	0.7349	20.96	0.7098
	CQ	43164728	37410128	0.8667	23.78	0.7134
	N.19-3	34283436	29772735	0.8684	19.14	0.6978
	N.19-4	36109690	31493307	0.8722	20.29	0.7023
	N.23-5	40859990	35905368	0.8787	23.26	0.716
	N.23-10	37093166	32184417	0.8677	20.6	0.7019
	N.24-1	44764680	39076538	0.8729	24.81	0.7225
	N.47-10	34160720	29799361	0.8723	19.66	0.7001
	N.02-10	39804846	34278871	0.8612	21.45	0.7076
	N.06-11	34968156	30591203	0.8748	20.05	0.7051
	N.07-3	33552960	29399674	0.8762	19.07	0.6999
<i>F. nubicola</i> (white fruits)	N.03-4	85538028	73766908	0.8624	45.61	0.77
	N.46-5	73043922	64622350	0.8847	39.24	0.7619
	N.03-5	71298428	63330978	0.8883	40.03	0.7635
	N.02-1	70871502	62797470	0.8861	39.63	0.7581
	N.42-10	33864284	29366822	0.8672	19.15	0.711
	N.8-8	42741472	38109838	0.8916	24.11	0.7361
	N.9-8	33936314	29832741	0.8791	19.36	0.7141
	N.47-7	79873334	70011503	0.8765	45.32	0.7698
	N.22-10	41230812	36386330	0.8825	24.83	0.732
	N.44-8	36847384	32159557	0.8728	20.66	0.7238
	N.03-6	41346762	36203566	0.8756	23.25	0.7279
<i>F. nubicola</i> (red fruits)	N.47-4	33615490	29532178	0.8785	18.81	0.7148
	N.43-8	36219304	22161268	0.6119	14.37	0.6871
	N.45-3	34545260	30222454	0.8749	19.61	0.7178
	N.02-3	46940926	41601586	0.8863	26.44	0.7412
	N.03-10	33783392	28374692	0.8399	18.39	0.7125
	N.18-6	38361302	33875618	0.8831	21.38	0.7422
<i>F. pentaphylla</i>	N.20-7	38144006	33832176	0.887	21.9	0.7289
	N.26-5	39889040	34735969	0.8708	21.79	0.7263
	N.05-3	33391020	28962371	0.8674	18.41	0.7073
	N.6-5	34770726	30625907	0.8808	19.52	0.7208
	N.26-1	33628430	29441993	0.8755	18.73	0.7149
	N.26-10	38365126	33293868	0.8678	21.11	0.7219
	N.04-6	33789212	29555238	0.8747	18.84	0.7123

	N.04-12	38722230	34218562	0.8837	21.62	0.7317
	N.05-14	38664466	34063193	0.881	21.75	0.7289
	N.04-15	42691656	37520966	0.8789	23.86	0.7436
	N.25-12	34266092	29506382	0.8611	18.52	0.7103
	N.26-11	36371010	31647042	0.8701	19.65	0.7169
	N.04-7	34322796	30094283	0.8768	19.4	0.7174
	N.04-11	34748256	30571607	0.8798	19.44	0.7185
	N.05-2	45245306	40159485	0.8876	25.49	0.739
	N.26-6	35421022	30732008	0.8676	18.38	0.7078
	N.26-2	37641140	32964583	0.8758	20.76	0.7263
	N.04-5	42473986	37501745	0.8829	23.47	0.743
	N.05-11	37715050	33059230	0.8766	20.56	0.7329
	N.6-7	72208704	63266676	0.8762	38.8	0.7684
	N.16-9	42369292	37047831	0.8744	23.03	0.7398
	N.26-3	34967020	29332618	0.8389	18.85	0.7168
	N.26-13	45346532	39510684	0.8713	23.99	0.735
	N.04-16	47798798	41630384	0.871	26.96	0.7485
<i>F. emeiensis</i>	N.44-2	37226020	32331463	0.8685	20.96	0.7239
	N.05-13	58133324	50696297	0.8721	32.18	0.7628
	N.07-4	36144692	31955820	0.8841	21.4	0.7253
	N.21-1	40375032	35707942	0.8844	23.41	0.7258
	N.7-3	71491026	62494434	0.8742	39.05	0.7648
	N.11-10	73871992	65271644	0.8836	39.26	0.7762
	N.06-5	88461122	77218000	0.8729	44.52	0.7928
	N.42-3	38886248	33152776	0.8526	21.27	0.7209
	N.43-4	38958468	33756856	0.8665	21.88	0.7223
	N.7-2	34291594	30226746	0.8815	19.89	0.7211
	N.15-6	40084844	33696785	0.8406	21.41	0.727
	N.43-1	48783676	42928501	0.88	27.1	0.7441
<i>F. vesca</i>	N.5-1	39853518	37280806	0.9354	20.92	0.9597
	N.18-2	42238218	40296269	0.954	22.41	0.9616
	N.1-6	39126078	37610608	0.9613	21.09	0.9909
	N.7-6	36734502	35065566	0.9546	19.43	0.9754
	N.1-4	37132820	35102865	0.9453	19.91	0.9779
	N.1-9	45661048	42490156	0.9306	23.72	0.9689
	N.21-7	34318592	32185167	0.9378	18.27	0.9563
<i>F. viridis</i>	N.2-5	43987946	38516055	0.8756	24.22	0.7557
	N.43-10	35719334	31583963	0.8842	19.76	0.7419
	N.45-6	40020866	35331045	0.8828	21.96	0.7509
	N.16-6	98021718	90262146	0.9208	49.87	0.9276
	FVI	2.64E+08	2.38E+08	0.9032	158.09	0.8446
Outgroup	N.161	90387462	17766801	0.1966	44.5	0.1262
Outgroup	N.1-1	89333736	17510146	0.196	30.56	0.1962

Table S5. Population genetic diversity and differentiation of eight key diploid strawberry species

Species	Num.	θ_w	θ_π	Tajima's D	Pairwise comparison	F_{ST}
<i>F. viridis</i>	5	0.0245	0.024	-0.215	<i>F. viridis</i> vs. <i>F. vesca</i>	0.640
<i>F. nubicola</i>	17	0.0193	0.0196	-0.0003	<i>F. nubicola</i> vs. <i>F. emeiensis</i>	0.203
<i>F. nilgerrensis</i>	34	0.0177	0.0194	0.303	<i>F. nilgerrensis</i> vs. <i>F. nubicola</i>	0.539
<i>F. emeiensis</i>	12	0.0197	0.0188	-0.216	<i>F. emeiensis</i> vs. <i>F. pentaphylla</i>	0.164
<i>F. nubicola</i> (red)	6	0.0183	0.0187	0.105	<i>F. nubicola</i> (red) vs. <i>F. nubicola</i> (white)	0.177
<i>F. pentaphylla</i>	24	0.0185	0.0185	-0.075	<i>F. pentaphylla</i> vs. <i>F. chinensis</i>	0.157
<i>F. chinensis</i>	22	0.0189	0.0184	-0.146	<i>F. chinensis</i> vs. <i>F. emeiensis</i>	0.151
<i>F. nubicola</i> (white)	11	0.0147	0.0172	0.682	<i>F. viridis</i> vs. <i>F. nilgerrensis</i>	0.595
<i>F. mandschurica</i>	5	0.0121	0.0132	0.386	<i>F. mandschurica</i> vs. <i>F. viridis</i>	0.565
<i>F. vesca</i>	7	0.0071	0.0071	0.086	<i>F. vesca</i> vs. <i>F. mandschurica</i>	0.461