

Table S1. Gene contents in *E. cymosa* and *E. obtusifolia* chloroplast genomes

Category	Gene groups	Gene Names
RNA genes	Ribosomal RNA genes (rRNA)	<i>rrn4.5^a, rrn5^a, rrn16^a, rrn23^a</i>
	Transfer RNA genes (tRNA)	<i>trnA-UGC^{+,a}, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnI⁺-CAU, trnG-GCC, trnG-GCC⁺, trnG-UUC⁺, trnH-GUG, trnI-CAU^{+,a}, trnK-UUU⁺, trnL-CAA^a, trnL-GAU^a, trnL-UAA⁺, trnL-UAG, trnM-CAU, trnN-GUU^a, trnP-UGG, trnQ-UUG, trnR-ACG^a, trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC^a, trnV-UAC⁺, trnW-CCA, trnY-GUA</i>
Ribosomal proteins	Small sub-unit of ribosome	<i>rps2, rps3, rps4, rps7^a, rps8, rps11, rps12^{+,a}, rps14, rps15, rps16⁺, rps18, rps19</i>
Transcription Genes	Large sub-unit of ribosome	<i>rpl2^{+,a}, rpl14, rpl16⁺, rpl20, rpl22, rpl23^a, rpl32, rpl33, rpl36</i>
	DNA-dependent RNA polymerase	<i>rpoA, rpoB, rpoC1⁺, rpoC2</i>
Protein-coding genes	Photosystem I	<i>psaA, psaB, psaC, psaI, psaJ, ycf3⁺⁺</i>
	Photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
	Sub-unit of cytochrome	<i>petA, petB⁺, petD⁺, petG, petL, petN</i>
	Sub-unit of synthase	<i>atpA, atpB, atpE, atpF⁺, atpH, atpI</i>
	Large sub-unit of rubisco	<i>rbcL</i>
	NADH dehydrogenase	<i>ndhA⁺, ndhB^{+,a}, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	ATP-dependent proteasesubunit P	<i>clpP⁺⁺</i>
Other genes	Cp envelope membrane protein	<i>cemA</i>
	Maturase	<i>matK</i>
	Sub-unit acetyl-coA carboxylase	<i>accD</i>
	C-type cytochrome sythesis	<i>ccsA</i>
	Translation initiation factor	<i>infA</i>
	Hypothetical proteins	<i>ycf2^a, ycf4, ycf15^a</i>
	Component of TIC complex	<i>ycf1^a</i>

⁺Gene with one intron, ⁺⁺Gene with two introns and ^a Gene with multiple copies. *trnG-UCC* gene was absent in *E. cymosa*, while *trnG-GCC* gene was absent in *E. obtusifolia*.

Table S2. Exons and introns lengths in *E. cymosa* and *E. obtusifolia* chloroplast genomes.

Gene	Location	Exon I (bp)		Intron I (bp)		Exon II (bp)		Intron II (bp)		Exon III (bp)	
		<i>E. c</i>	<i>E. o</i>	<i>E. c</i>	<i>E. o</i>	<i>E. c</i>	<i>E. o</i>	<i>E. c</i>	<i>E. o</i>	<i>E. c</i>	<i>E. o</i>
<i>trnK-UUU</i>	LSC	34	34	2469	2475	36	36				
<i>rps16</i>	LSC	226	226	864	864	39	39				
<i>trnG-GCC/UCC</i>	LSC	22	22	685	685	47	47				
<i>atpF</i>	LSC	410	410	702	702	143	143				
<i>rpoC1</i>	LSC	1612	1610	767	767	450	431				
<i>ycf3</i>	LSC	154	154	783	777	227	227	735	733	123	123
<i>trnL-UAA</i>	LSC	34	34	496	496	49	49				
<i>trnV-UAC</i>	LSC	36	36	582	584	37	35				
<i>clpP1</i>	LSC	227	227	613	614	291	291	778	776	70	70
<i>petB</i>	LSC	5	5	778	784	641	641				
<i>petD</i>	LSC	7	7	674	674	474	474				
<i>rpl16</i>	LSC	398	398	1051	1049	8	8				
<i>rpl2</i>	IR	433	433	660	660	390	390				
<i>ndhB</i>	IR	755	757	678	678	776	774				
<i>rps12</i>	IR	25	25	535	535	231	231				
<i>trnI-GAU</i>	IR	36	36	947	947	34	34				
<i>trnA-UGC</i>	IR	37	37	806	806	34	34				
<i>ndhA</i>	SSC	538	538	1084	1086	552	552				

trnG-UCC gene was absent in *E. cymosa*, while *trnG-GCC* gene was absent in *E. obtusifolia*.

Table S3. Codon-anticodon recognition patterns and codon usage of the *Ehretia cymosa* chloroplast genome

Codon	Amino Acid	RSCU	tRNA	Codon	Amino Acid	RSCU	tRNA
UUU(F)	Phe	1.2	trnF-GAA	UAU(Y)	Tyr	1.3	trnY-GUA
UUC(F)	Phe	0.8		UAC(Y)	Tyr	0.7	
UUA(L)	Leu	1.19	trnL-UAA	UAA(*)	Stop	0.99	
UUG(L)	Leu	1.38	trnL-CAA	UAG(*)	Stop	0.94	
CUU(L)	Leu	1.2	trnL-UAG	CAU(H)	His	1.33	trnH-GUG
CUC(L)	Leu	0.76		CAC(H)	His	0.67	
CUA(L)	Leu	0.85		CAA(Q)	Gln	1.3	trnQ-UUG
CUG(L)	Leu	0.63		CAG(Q)	Gln	0.7	
AUU(I)	Ile	1.27	trnI-GAU	AAU(N)	Asn	1.3	trnN-GUU
AUC(I)	Ile	0.85		AAC(N)	Asn	0.7	
AUA(I)	Ile	0.88	trnI-CAU	AAA(K)	Lys	1.25	trnK-UUU
AUG(M)	Met	1	trnM-CAU	AAG(K)	Lys	0.75	
GUU(V)	Val	1.3	trnV-GAC	GAU(D)	Asp	1.35	trnD-GUC
GUC(V)	Val	0.79		GAC(D)	Asp	0.65	
GUA(V)	Val	1.07		GAA(E)	Glu	1.29	trnE-UUC
GUG(V)	Val	0.84	trnV-UAC	GAG(E)	Glu	0.71	
UCU(S)	Ser	1.44	trnS-GGA	UGU(C)	Cys	1.17	trnC-GCA
UCC(S)	Ser	0.91		UGC(C)	Cys	0.83	
UCA(S)	Ser	1.29		UGA(*)	Stop	1.07	
UCG(S)	Ser	0.86	trnS-UGA	UGG(W)	Trp	1	trnW-CCA
CCU(P)	Pro	1.06	trnP-UGG	CGU(R)	Arg	0.66	trnR-ACG
CCC(P)	Pro	0.87		CGC(R)	Arg	0.3	trnR-UCU
CCA(P)	Pro	1.13		CGA(R)	Arg	1.02	
CCG(P)	Pro	0.94		CGG(R)	Arg	0.81	
ACU(P)	Thr	0.96		AGA(R)	Arg	1.96	
ACC(P)	Thr	1		AGG(R)	Arg	1.25	
ACA(T)	Thr	1.24	trnT-GGU	AGU(S)	Ser	0.8	trnS-GCU
ACG(T)	Thr	0.8	trnT-UGU	AGC(S)	Ser	0.7	
GCU(A)	Ala	1.19	trnA-UGC	GGU(G)	Gly	0.91	trnG-GCC
GCC(A)	Ala	0.85		GGC(G)	Gly	0.61	
GCA(A)	Ala	1.12		GGA(G)	Gly	1.29	
GCG(A)	Ala	0.85		GGG(G)	Gly	1.19	trnG-UCC

Table S4. Codon-anticodon recognition patterns and codon usage of the *Ehretia obtusifolia* chloroplast genome

Codon	Amino Acid	RSCU	tRNA	Codon	Amino Acid	RSCU	tRNA
UUU(F)	Phe	1.15	trnF-GAA	UAU(Y)	Tyr	1.35	trnY-GUA
UUC(F)	Phe	0.85		UAC(Y)	Tyr	0.65	
UUA(L)	Leu	1.23	trnL-UAA	UAA(*)	Stop	0.94	
UUG(L)	Leu	1.34	trnL-CAA	UAG(*)	Stop	1.04	
CUU(L)	Leu	1.09	trnL-UAG	CAU(H)	His	1.28	trnH-GUG
CUC(L)	Leu	0.67		CAC(H)	His	0.72	
CUA(L)	Leu	0.97		CAA(Q)	Gln	1.22	trnQ-UUG
CUG(L)	Leu	0.69		CAG(Q)	Gln	0.78	
AUU(I)	Ile	1.14	trnI-GAU	AAU(N)	Asn	1.32	trnN-GUU
AUC(I)	Ile	0.86		AAC(N)	Asn	0.68	
AUA(I)	Ile	1	trnI-CAU	AAA(K)	Lys	1.21	trnK-UUU
AUG(M)	Met	1	trnM-CAU	AAG(K)	Lys	0.79	
GUU(V)	Val	1.32	trnV-GAC	GAU(D)	Asp	1.32	trnD-GUC
GUC(V)	Val	0.74		GAC(D)	Asp	0.68	
GUA(V)	Val	1.12		GAA(E)	Glu	1.3	trnE-UUC
GUG(V)	Val	0.82	trnV-UAC	GAG(E)	Glu	0.7	
UCU(S)	Ser	1.38	trnS-GGA	UGU(C)	Cys	1.17	trnC-GCA
UCC(S)	Ser	0.9		UGC(C)	Cys	0.83	
UCA(S)	Ser	1.36		UGA(*)	Stop	1.02	
UCG(S)	Ser	0.89	trnS-UGA	UGG(W)	Trp	1	trnW-CCA
CCU(P)	Pro	0.96	trnP-UGG	CGU(R)	Arg	0.61	trnR-ACG
CCC(P)	Pro	0.86		CGC(R)	Arg	0.33	trnR-UCU
CCA(P)	Pro	1.22		CGA(R)	Arg	0.93	
CCG(P)	Pro	0.96		CGG(R)	Arg	0.85	
ACU(P)	Thr	1.1		AGA(R)	Arg	2.03	
ACC(P)	Thr	0.83		AGG(R)	Arg	1.24	
ACA(T)	Thr	1.22	trnT-GGU	AGU(S)	Ser	0.82	trnS-GCU
ACG(T)	Thr	0.85	trnT-UGU	AGC(S)	Ser	0.64	
GCU(A)	Ala	1.24	trnA-UGC	GGU(G)	Gly	0.97	trnG-GCC
GCC(A)	Ala	0.8		GGC(G)	Gly	0.58	
GCA(A)	Ala	1.08		GGA(G)	Gly	1.26	
GCG(A)	Ala	0.88		GGG(G)	Gly	1.19	trnG-UCC

Table S5. Predicted RNA editing site in the *Ehretia cymosa* and *Ehretia obtusifolia* chloroplast genome.

Gene	Nucleotide position	Amino acid position	Triplet position within codon	Base conversion	Codon change	Amino acid conversion
matK	655	219	1	C→U	CAU→UAU	H→Y
atpF	92	31	2	C→U	CCA→CUA	P→L
rps2	248	83	2	C→U	UCA→UUA	S→L
rpoB	338	113	2	C→U	UCU→UUU	S→F
	551	184	2	C→U	UCA→UUA	S→L
	566	189	2	C→U	UCG→UUG	S→L
	2432	811	2	C→U	UCA→UUA	S→L
psbZ	50	17	2	C→U	UCA→UUA	S→L
rps14	80	27	2	C→U	UCA→UUA	S→L
	149	50	2	C→U	CCA→CUA	P→L
accD	830	277	2	C→U	UCG→UUG	S→L
psbE	214	72	1	C→U	CCU→UCU	P→S
petB	611	204	2	C→U	CCA→CUA	P→L
rpoA	200	67	2	C→U	UCU→UUU	S→F
rpl23	89	30	2	C→U	UCA→UUA	S→L
ndhB	53	18	2	C→U	UCA→UUA	S→L
	59	20	2	C→U	UCA→UUA	S→L
	95	32	2	C→U	UCA→UUA	S→L
	149	50	2	C→U	UCA→UUA	S→L
	467	156	2	C→U	CCA→CUA	P→L
	586	196	1	C→U	CAU→UAU	H→Y
	704	235	2	C→U	CCA→CUA	P→L
	737	246	2	C→U	CCA→CUA	P→L
	746	249	2	C→U	UCU→UUU	S→F
ndhF	290	97	2	C→U	UCA→UUA	S→L
ndhD	2	1	2	C→U	ACG→AUG	T→M
	383	128	2	C→U	UCA→UUA	S→L
	674	225	2	C→U	UCG→UUG	S→L
	878	293	2	C→U	UCA→UUA	S→L
ndhG	50	17	2	C→U	UCG→UUG	S→L
ndhA	341	114	2	C→U	UCA→UUA	S→L

Table S6. Repeat sequences present in the *Ehretia cymosa* chloroplast genome.

SN	Repeat Size	Repeat Position 1	Repeat Type	Repeat Location 1	Repeat Position 2	Repeat Location 2	E-Value
1	44	76685	P	IGS	76685	IGS	0 2.22e-17
2	41	40237	F	psaB	42461	psaA	0 1.42e-15
3	29	8682	P	trnS-GCU	46755	trnS-GGA	0 2.38e-08
4	29	65460	P	IGS	110054	IGS	0 2.38e-08
5	29	65460	F	IGS	132819	IGS	0 2.38e-08
6	26	90191	P	ycf2	90191	ycf2	0 1.53e-06
7	26	90191	F	ycf2	152685	ycf2	0 1.53e-06
8	26	152685	P	ycf2	152685	ycf2	0 1.53e-06
9	24	127356	P	ycf1	127356	ycf1	0 2.44e-05
19	22	8919	P	IGS	44989	ycf3 - Intron	0 3.91e-04
11	22	10255	F	trnG-GCC	38027	trnG-GCC	0 3.91e-04
12	22	32385	P	IGS	32385	IGS	0 3.91e-04
13	22	96276	P	IGS	96302	IGS	0 3.91e-04
14	22	96276	F	IGS	146578	IGS	0 3.91e-04
15	22	96302	F	IGS	146604	IGS	0 3.91e-04
16	22	116011	R	IGS	116011	IGS	0 3.91e-04
17	22	117844	P	IGS	117844	IGS	0 3.91e-04
18	22	146578	P	IGS	146604	IGS	0 3.91e-04
19	21	8687	F	trnS-GCU	37104	trnS-UGA	0 1.56e-03
20	21	37104	P	trnS-UGA	46758	trnS-GGA	0 1.56e-03
21	21	38237	F	trnM-CAU	69019	trnP-UGG	0 1.56e-03
22	21	93831	F	ycf2	93849	ycf2	0 1.56e-03
23	21	93831	P	ycf2	149032	ycf2	0 1.56e-03
24	21	93849	P	ycf2	149050	ycf2	0 1.56e-03
25	21	100862	F	IGS	122634	ndhA - Intron	0 1.56e-03
26	21	122168	R	ndhA	122168	ndhA	0 1.56e-03
27	21	122634	P	ndhA - Intron	142019	IGS	0 1.56e-03
28	21	149032	F	ycf2	149050	ycf2	0 1.56e-03
29	20	51915	R	ndhC	51915	ndhC	0 6.25e-03
30	20	53899	P	trnV-UAC	105576	trnA-UGC	0 6.25e-03
31	20	53899	F	trnV-UAC	137306	trnA-UGC	0 6.25e-03
32	20	115985	F	IGS	123274	ndhA - Intron	0 6.25e-03
33	19	7431	R	IGS	7431	IGS	0 2.50e-02
34	19	15661	R	atpI	15661	atpI	0 2.50e-02
35	19	67949	F	IGS	84330	rpl16	0 2.50e-02
36	19	68508	R	IGS	68508	IGS	0 2.50e-02
37	19	82952	R	IGS	82952	IGS	0 2.50e-02
38	18	263	P	IGS	310	IGS	0 1.00e-01
39	18	4213	F	IGS	71586	IGS	0 1.00e-01
40	18	7240	P	IGS	67949	IGS	0 1.00e-01
41	18	7240	P	IGS	84330	rpl16 - Intron	0 1.00e-01
42	18	7241	F	IGS	73944	clpP1 - Intron	0 1.00e-01
43	18	8752	F	trnS-GCU	37174	trnS-UGA	0 1.00e-01
44	18	16935	P	IGS	16935	IGS	0 1.00e-01
45	18	25440	C	rpoB	29434	IGS	0 1.00e-01
46	18	31414	C	IGS	62888	IGS	0 1.00e-01
47	18	33040	R	IGS	33040	IGS	0 1.00e-01

Table S7. Repeat sequences present in the *Ehretia obtusifolia* chloroplast genome.

SN	Repeat Size	Repeat Position 1	Repeat Type	Repeat Location 1	Repeat Position 2	Repeat Location 2	E-Value
1	44	76319	P	IGS	130163	ycf1	0 2.21e-17
2	41	39891	F	psaB	76319	IGS	0 1.41e-15
3	32	61289	P	IGS	42115	psaA	0 3.71e-10
4	29	8701	P	IGS	61289	IGS	0 2.37e-08
5	29	65149	P	IGS	46410	trnS-GGA	0 2.37e-08
6	29	65149	F	IGS	109682	IGS	0 2.37e-08
7	26	89824	P	ycf2	132461	IGS	0 1.52e-06
8	26	89824	F	ycf2	89824	ycf2	0 1.52e-06
9	26	152322	P	ycf2	152322	ycf2	0 1.52e-06
19	24	115652	R	IGS	152322	ycf2	0 2.43e-05
11	24	126998	P	ycf1	115652	IGS	0 2.43e-05
12	22	10278	F	trnG-UCC	126998	ycf1	0 3.89e-04
13	22	14780	F	IGS	37679	trnG-GCC	0 3.89e-04
14	22	32034	P	IGS	14802	IGS	0 3.89e-04
15	22	95911	P	IGS	32034	IGS	0 3.89e-04
16	22	95911	F	IGS	95937	IGS	0 3.89e-04
17	22	95937	F	IGS	146213	IGS	0 3.89e-04
18	22	117485	P	IGS	146239	IGS	0 3.89e-04
19	22	146213	P	IGS	117485	IGS	0 3.89e-04
20	21	8706	F	trnS-GCU	146239	IGS	0 1.56e-03
21	21	36754	P	trnS-UGA	36754	trnS-UGA	0 1.56e-03
22	21	37891	F	IGS	46413	trnS-GGA	0 1.56e-03
23	21	93464	F	ycf2	68676	trnP-UGG	0 1.56e-03
24	21	93464	P	ycf2	93482	ycf2	0 1.56e-03
25	21	93482	P	ycf2	148669	ycf2	0 1.56e-03
26	21	100497	F	IGS	148687	ycf2	0 1.56e-03
27	21	121808	R	ndhA	122274	ndhA - Intron	0 1.56e-03
28	21	122274	P	ndhA - Intron	121808	ndhA	0 1.56e-03
29	21	148669	F	ycf2	141654	IGS	0 1.56e-03
30	20	28435	F	IGS	148687	ycf2	0 6.22e-03
31	20	37781	R	IGS	28445	IGS	0 6.22e-03
32	20	51578	R	ndhC	37781	IGS	0 6.22e-03
33	20	53562	P	trnV-UAC	51578	ndhC	0 6.22e-03
34	20	53562	F	trnV-UAC	105204	trnA-UGC	0 6.22e-03
35	20	115626	F	IGS	136948	trnA-UGC	0 6.22e-03
36	19	7449	R	IGS	122916	ndhA - Intron	0 2.49e-02
37	19	15726	R	atpI	7449	IGS	0 2.49e-02
38	18	288	P	IGS	15726	atpI	0 9.96e-02
39	18	8771	F	trnS-GCU	335	IGS	0 9.96e-02
40	18	8938	P	IGS	36824	trnS-UGA	0 9.96e-02
41	18	10342	R	IGS	44642	ycf3 - Intron	0 9.96e-02
42	18	25498	C	rpoB	10342	IGS	0 9.96e-02
43	18	32689	R	IGS	29512	IGS	0 9.96e-02
44	18	58588	P	IGS	32689	IGS	0 9.96e-02
45	18	62955	P	IGS	83127	IGS	0 9.96e-02
46	18	68167	R	IGS	120745	IGS	0 9.96e-02
47	18	68914	R	IGS	68167	IGS	0 9.96e-02
48	18	104122	P	trnI-GAU	68914	IGS	0 9.96e-02
49	18	104122	F	trnI-GAU	110592	trnN-GUU	0 9.96e-02

Table S8. Simple sequence repeats in the chloroplast genome of *Ehretia cymosa*

Repeat	Length (bp)	Number	Start position
A	8	26	7444; 8903; 9016; 9251; 22954; 37344; 37815; 38868; 43793; 48493; 49116; 56243; 61905; 64759; 66204; 66764; 70336; 70845; 73599; 73619; 95830; 113111; 116974; 117926; 119498; 121104
	9	19	4568; 4776; 8662; 14142; 15992; 19288; 38146; 44402; 46227; 65808; 72706; 80271; 92017; 99226; 137933; 141668; 142967; 146877; 156271
	10	5	4217; 30216; 37815; 45971; 71518
	11	7	4217; 29208; 56825; 56848; 65808; 83491; 115208
	12	3	67955; 84336; 116240
	17	1	16936
C	8	1	5082
	12	1	38146
T	8	28	1561; 2177; 8012; 8214; 9016; 17245; 26852; 30866; 43793; 48359; 49116; 50320; 51730; 56701; 62140; 67468; 69317; 70845; 71518; 72968; 76456; 77682; 86557; 117306; 118511; 128279; 130289; 147066
	9	23	16936; 43793; 52477; 53365; 58933; 58984; 86312; 86557; 86632; 96018; 99928; 101227; 104962; 112382; 115208; 117805; 121104; 123027; 125556; 128508; 143669; 150878
	10	9	9016; 14703; 31557; 56243; 60860; 62788; 80516; 116018; 128882
	11	3	33459; 65808; 68513
	12	3	7242; 61576; 67468
	13	2	73944; 82956
	15	1	19145
AT	5	4	15177; 20518; 87252; 155642
TTC	4	1	36754
AAAC	3	1	70336
TTTA	3	2	10317; 23263
TTGGA	3	1	67468

Table S9. Simple sequence repeats in the chloroplast genome of *Ehretia obtusifolia*

Repeat	Length (bp)	Number	Start position
A	8	26	7462; 8922; 9040; 9274; 23012; 36994; 37465; 38522; 43447; 45874; 48148; 48771; 55906; 61594; 64448; 65497; 66425; 69995; 70504; 73253; 73273; 95463; 112753; 116615; 117567; 119139
	9	19	4248; 4599; 4807; 14188; 16057; 19346; 37802; 45874; 65497; 71177; 72359; 79911; 91650; 98861; 137575; 141310; 142602; 155908; 155951
	10	8	8680; 30298; 45618; 56488; 83128; 83972; 114850; 115883; 8680
	11	2	29286; 146512
	12	3	4248; 37465; 67614
	13	1	56488
C	9	1	5113
	10	1	37802
T	8	30	1586; 1796; 2208; 8030; 8232; 9040; 17303; 26910; 30948; 43447; 48014; 48771; 49983; 51393; 53028; 53504; 56364; 61829; 67129; 70504; 71177; 76090; 77316; 85946; 86190; 116947; 118152; 127921; 129931; 146703
	9	18	9040; 10457; 52140; 58597; 72621; 86190; 86265; 99563; 100855; 104590; 114850; 117446; 120818; 125198; 128150; 128229; 143304; 150515
	10	11	43447; 55906; 60526; 62477; 65497; 67129; 68172; 68974; 80156; 82596; 128524
	11	4	7261; 58597; 95651; 122667
	12	4	17001; 61242; 73598; 115659
	13	1	19203
	14	1	33106
AT	5	4	15242; 20576; 86885; 155279
TTC	4	1	36404
AAAC	3	1	69995
TTTA	3	2	10340; 23321
TTGGA	3	1	67129