

Complete chloroplast genomes of *Cerastium alpinum*, *C. arcticum* and *C. nigrescens*: genome structures, comparative and phylogenetic analysis

Sylwia E. Milarska¹, Piotr Androsiuk^{1*}, Łukasz Pauksto², Jan P. Jastrzębski¹, Mateusz Maździarz², Keith Larson³ and Irena Giełwanowska¹

¹ Department of Plant Physiology, Genetics and Biotechnology, Faculty of Biology and Biotechnology, University of Warmia and Mazury in Olsztyn, ul. M. Oczapowskiego 1A, 10-719 Olsztyn, Poland;

² Department of Botany and Nature Protection, Faculty of Biology and Biotechnology, University of Warmia and Mazury in Olsztyn, Pl. Łódzki 1, 10-721 Olsztyn, Poland;

³ Climate Impacts Research Centre, Department of Ecology and Environmental Sciences, Umeå University, 901 87, Umeå, Sweden.

* corresponding author – piotr.androsiuk@uwm.edu.pl

Table S2A. List of repeated sequences in the chloroplast genome of *Cerastium alpinum*

Repeat length (bp)	Strat site of repeat A	Repeat A location	Repeat A region	Strat site of repeat B	Repeat B location	Repeat B region	Repeat type
30	63,333	<i>trnS-GCU</i>	LSC	97,992	<i>trnS-GGA</i>	LSC	P
30	58,549	<i>matK</i>	LSC	58,549	<i>matK</i>	LSC	P
30	88,609	<i>trnS-UGA</i>	LSC	97,992	<i>trnS-GGA</i>	LSC	P
30	91,702	<i>psaB</i>	LSC	93,926	<i>psaA</i>	LSC	F
30	100,208	<i>trnL-UAA</i> (intron)	LSC	100,208	<i>trnL-UAA</i> (intron)	LSC	P
30	113,132	IGS (<i>ycf4-cemA</i>)	LSC	113,132	IGS (<i>ycf4-cemA</i>)	LSC	R
30	1,460	IGS (<i>rps12-rrn16</i>)	IR	96,827	<i>ycf3</i> (intron 1)	LSC	F
30	63,333	<i>trnS-GCU</i>	LSC	88,609	<i>trnS-UGA</i>	LSC	F
30	64,936	<i>trnG-UCC</i>	LSC	89,510	IGS (<i>psbZ-rps14</i>)	LSC	F
31	139,955	<i>ycf2</i>	IR	139,976	<i>ycf2</i>	IR	F
32	55,946	IGS (<i>trnH-GUG-psbA</i>)	LSC	55,984	IGS (<i>trnH-GUG-psbA</i>)	LSC	P
34	65,218	IGS (<i>trnR-UCU-atpA</i>)	LSC	67,647	<i>atpF</i> (intron)	LSC	P
35	95,382	IGS (<i>psaA-ycf3</i>)	LSC	112,340	IGS (<i>psaI-ycf4</i>)	LSC	P
36	82,993	IGS (<i>petN-psbM</i>)	LSC	82,993	IGS (<i>petN-psbM</i>)	LSC	P
37	23,345	<i>ndhA</i> (intron)	SSC	96,815	<i>ycf3</i> (intron 1)	LSC	F
39	1,446	IGS (<i>rps12-rrn16</i>)	IR	23,343	<i>ndhA</i> (intron)	SSC	F
39	23,343	<i>ndhA</i> (intron)	SSC	42,182	IGS (<i>rrn16-rps12</i>)	IR	P
39	1,448	IGS (<i>rps12-rrn16</i>)	IR	96,815	<i>ycf3</i> (intron 1)	LSC	F
46	20,047	IGS (<i>ndhD-psaC</i>)	SSC	20,047	IGS (<i>ndhD-psaC</i>)	SSC	P
73	95,310	IGS (<i>psaA-ycf3</i>)	LSC	112,375	IGS (<i>psaI-ycf4</i>)	LSC	P

IGS (*ycf4-cemA*) means spacer between *ycf4* and *cemA*, P means palindromic match, F means forward (direct) match and R means reverse match

Table S2B. List of repeated sequences in the chloroplast genome of *Cerastium arcticum*

Repeat length (bp)	Strat site of repeat A	Repeat A location	Repeat A region	Strat site of repeat B	Repeat B location	Repeat B region	Repeat type
30	63,359	<i>trnS-GCU</i>	LSC	98,903	<i>trnS-GGA</i>	LSC	P
30	58,564	<i>matK</i>	LSC	58,564	<i>matK</i>	LSC	P
30	89,326	<i>trnS-UGA</i>	LSC	98,903	<i>trnS-GGA</i>	LSC	P
30	92,419	<i>psaB</i>	LSC	94,643	<i>psaA</i>	LSC	F
30	113,900	IGS (<i>ycf4-cemA</i>)	LSC	113,900	IGS (<i>ycf4-cemA</i>)	LSC	R
30	1,451	IGS (<i>rps12-rrn16</i>)	IR	97,732	<i>ycf3</i> (intron 1)	LSC	F
30	63,359	<i>trnS-GCU</i>	LSC	89,326	<i>trnS-UGA</i>	LSC	F
30	64,942	<i>trnG-UCC</i>	LSC	90,227	IGS (<i>psbZ-rps14</i>)	LSC	F
31	65,242	IGS (<i>trnR-UCU-atpA</i>)	LSC	67,677	<i>atpF</i> (intron)	LSC	P
31	140,727	<i>ycf2</i>	IR	140,748	<i>ycf2</i>	IR	F
31	97,369	<i>ycf3</i> (intron 2)	LSC	97,383	<i>ycf3</i> (intron 2)	LSC	F
33	55,959	IGS (<i>trnH-GUG-psbA</i>)	LSC	55,998	IGS (<i>trnH-GUG-psbA</i>)	LSC	P
34	100,973	<i>trnL-UAA</i> (intron)	LSC	100,977	<i>trnL-UAA</i> (intron)	LSC	P
36	83,701	IGS (<i>petN-psbM</i>)	LSC	83,701	IGS (<i>petN-psbM</i>)	LSC	P
37	23,356	<i>ndhA</i> (intron)	SSC	97,720	<i>ycf3</i> (intron 1)	LSC	F
39	1,437	IGS (<i>rps12-rrn16</i>)	IR	23,354	<i>ndhA</i> (intron)	SSC	F
39	23,354	<i>ndhA</i> (intron)	SSC	42,199	IGS (<i>rrn16-rps12</i>)	IR	P
39	1,439	IGS (<i>rps12-rrn16</i>)	IR	97,720	<i>ycf3</i> (intron 1)	LSC	F
43	97,333	<i>ycf3</i> (intron 2)	LSC	97,360	<i>ycf3</i> (intron 2)	LSC	F
46	20,055	IGS (<i>ndhD-psaC</i>)	SSC	20,055	IGS (<i>ndhD-psaC</i>)	SSC	P
61	68,941	IGS (<i>atpH-atpI</i>)	LSC	69,002	IGS (<i>atpH-atpI</i>)	LSC	F
72	96,199	IGS (<i>psaA-ycf3</i>)	LSC	113,109	IGS (<i>psaI-ycf4</i>)	LSC	P
162	96,109	IGS (<i>psaA-ycf3</i>)	LSC	113,109	IGS (<i>psaI-ycf4</i>)	LSC	P

IGS (*ycf4-cemA*) means spacer between *ycf4* and *cemA*, P means palindromic match, F means forward (direct) match and R means reverse match

Table S2C. List of repeated sequences in the chloroplast genome of *Cerastium nigrescens*

Repeat length (bp)	Strat site of repeat A	Repeat A location	Repeat A region	Strat site of repeat B	Repeat B location	Repeat B region	Repeat type
30	63,336	<i>trnS-GCU</i>	LSC	98,004	<i>trnS-GGA</i>	LSC	P
30	585,48	<i>matK</i>	LSC	58,548	<i>matK</i>	LSC	P
30	88,612	<i>trnS-UGA</i>	LSC	98,004	<i>trnS-GGA</i>	LSC	P
30	91,705	<i>psaB</i>	LSC	93,929	<i>psaA</i>	LSC	F
30	100,215	<i>trnL-UAA</i> (intron)	LSC	100,215	<i>trnL-UAA</i> (intron)	LSC	P
30	113,134	IGS (<i>ycf4-cemA</i>)	LSC	113,134	IGS (<i>ycf4-cemA</i>)	LSC	R
30	1,460	IGS (<i>rps12-rrn16</i>)	IR	96,839	<i>ycf3</i> (intron 1)	LSC	F
30	63,336	<i>trnS-GCU</i>	LSC	88,612	<i>trnS-UGA</i>	LSC	F
30	64,939	<i>trnG-UCC</i>	LSC	89,513	IGS (<i>psbZ-rps14</i>)	LSC	F
31	139,950	<i>ycf2</i>	IR	139,971	<i>ycf2</i>	IR	F
32	55,945	IGS (<i>trnH-GUG-psbA</i>)	LSC	55,983	IGS (<i>trnH-GUG-psbA</i>)	LSC	P
34	65,221	IGS (<i>trnR-UCU-atpA</i>)	LSC	67,650	<i>atpF</i> (intron)	LSC	P
35	95,394	IGS (<i>psaA-ycf3</i>)	LSC	112,342	IGS (<i>psaI-ycf4</i>)	LSC	P
36	82,996	IGS (<i>petN-psbM</i>)	LSC	82,996	IGS (<i>petN-psbM</i>)	LSC	P
37	23,345	<i>ndhA</i> (intron)	SSC	96,827	<i>ycf3</i> (intron 1)	LSC	F
39	1,446	IGS (<i>rps12-rrn16</i>)	IR	23,343	<i>ndhA</i> (intron)	SSC	F
39	23,343	<i>ndhA</i> (intron)	SSC	42,181	IGS (<i>rrn16-rps12</i>)	IR	P
39	1,448	IGS (<i>rps12-rrn16</i>)	IR	96,827	<i>ycf3</i> (intron 1)	LSC	F
46	20,047	IGS (<i>ndhD-psaC</i>)	SSC	20,047	IGS (<i>ndhD-psaC</i>)	SSC	P
73	95,322	IGS (<i>psaA-ycf3</i>)	LSC	112,377	IGS (<i>psaI-ycf4</i>)	LSC	P

IGS (*ycf4-cemA*) means spacer between *ycf4* and *cemA*, P means palindromic match, F means forward (direct) match and R means reverse match

Table S2D. List of repeated sequences in the chloroplast genome of *Cerastium glomeratum*

Repeat length (bp)	Strat site of repeat A	Repeat A location	Repeat A region	Strat site of repeat B	Repeat B location	Repeat B region	Repeat type
30	63,304	<i>trnS-GCU</i>	LSC	98,815	<i>trnS-GGA</i>	LSC	P
30	89,290	<i>trnS-UGA</i>	LSC	98,815	<i>trnS-GGA</i>	LSC	P
30	92,395	<i>psaB</i>	LSC	94,619	<i>psaA</i>	LSC	F
30	113,867	IGS (<i>ycf4-cemA</i>)	LSC	113,867	IGS (<i>ycf4-cemA</i>)	LSC	R
30	63,304	<i>trnS-GCU</i>	LSC	89,290	<i>trnS-UGA</i>	LSC	F
30	64,873	<i>trnG-UCC</i>	LSC	90,194	<i>trnG-GCC</i>	LSC	F
31	140,629	<i>ycf2</i>	IR	140,650	<i>ycf2</i>	IR	F
31	100,965	<i>trnL-UAA</i> (intron)	LSC	100,965	<i>trnL-UAA</i> (intron)	LSC	P
33	55,897	IGS (<i>trnH-GUG-psbA</i>)	LSC	55,936	IGS (<i>trnH-GUG-psbA</i>)	LSC	P
36	58,506	<i>matK</i>	LSC	58,506	<i>matK</i>	LSC	P
36	83,674	IGS (<i>petN-psbM</i>)	LSC	83,674	IGS (<i>petN-psbM</i>)	LSC	P
39	1,446	IGS (<i>rps12-trnV-GAC</i>)	IR	20,343	<i>ndhA</i> (intron)	SSC	P
39	20,343	<i>ndhA</i> (intron)	SSC	42,115	IGS (<i>trnV-GAC-rps12</i>)	IR	F
46	23,632	IGS (<i>psaC-ndhD</i>)	SSC	23,632	IGS (<i>psaC-ndhD</i>)	SSC	P
85	96,157	IGS (<i>psaA-ycf3</i>)	LSC	113,077	IGS (<i>psaI-ycf4</i>)	LSC	P
170	96,072	IGS (<i>psaA-ycf3</i>)	LSC	113,077	IGS (<i>psaI-ycf4</i>)	LSC	P

IGS (*ycf4-cemA*) means spacer between *ycf4* and *cemA*, P means palindromic match, F means forward (direct) match and R means reverse match