

SUPPLEMENTARY MATERIALS – Jenke *et al.*

Table S 1: Species names, taxonomic affiliations, and database accession numbers for each plastid genome analyzed. Both NCBI Nucleotide and SRA accession numbers are provided.

Order	Family	Species	NCBI Nucl.	NCBI SRA
ANGIOSPERMS				
Angiosperms – Eudicots – Rosids				
Asterales	Asteraceae	<i>Artemisia annua</i>	NC_034683	SRR5602595
		<i>Echinacea angustifolia</i>	NC_034324	SRR5602579
		<i>Echinacea atrorubens</i>	NC_034323	SRR5602578
		<i>Echinacea laevigata</i>	NC_034322	SRR5602581
		<i>Echinacea pallida</i>	NC_034321	SRR5602580
		<i>Echinacea paradoxa</i>	NC_034320	SRR5602575
		<i>Echinacea purpurea</i>	NC_034327	SRR5602574
		<i>Echinacea sanguinea</i>	NC_034328	SRR5602577
		<i>Echinacea speciosa</i>	NC_034325	SRR5602576
		<i>Echinacea tennesseensis</i>	NC_034326	SRR5602587
		<i>Helianthus annuus</i>	NC_007977	SRR825830
		<i>Helianthus argophyllus</i>	NC_030275	SRR2155086
		<i>Helianthus debilis</i>	NC_030173	SRR5907791
		<i>Carthamus tinctorius</i>	NC_030783	SRR2154065
		<i>Centaurea diffusa</i>	NC_024286	SRR2729212
		<i>Taraxacum kok-saghyz</i>	NC_032057	SRR8185394
		<i>Lactuca sativa</i>	NC_007578	SRR8736654
Ericales	Actinidiaceae	<i>Actinidia chinensis</i>	NC_026690	DRR083750
	Ebenaceae	<i>Diospyros lotus</i>	NC_030786	SRR1560932
	Ericaceae	<i>Vaccinium macrocarpon</i>	NC_019616	SRR1276173
	Primulaceae	<i>Primula veris</i>	NC_031428	SRR1660449
Gentianales	Rubiaceae	<i>Mitragyna speciosa</i>	NC_034698	SRR5602600
		<i>Coffea canephora</i>	NC_030053	ERR321701
Lamiales	Lamiaceae	<i>Premna microphylla</i>	NC_026291	SRR6940036
		<i>Tectona grandis</i>	NC_020098	SRR6940065
		<i>Haplostachys haplostachya</i>	NC_029819	SRR3170741
		<i>Stachys byzantina</i>	NC_029825	SRR3170744
		<i>Stachys chamissonis</i>	NC_029822	SRR3170745
		<i>Stachys coccinea</i>	NC_029823	SRR3170746
		<i>Stachys sylvatica</i>	NC_029824	SRR3170747
		<i>Perilla frutescens</i>	NC_030756	SRR6940083
		<i>Ocimum basilicum</i>	NC_035143	SRR6940087
		<i>Scutellaria baicalensis</i>	NC_027262	SRR6940088
		<i>Ajuga reptans</i>	NC_023102	SRR6940062
	Oleaceae	<i>Jasminum sambac</i>	NC_034694	SRR5602611
		<i>Jasminum tortuosum</i>	NC_034691	SRR5602601
		<i>Olea europaea</i> subsp. <i>europaea</i>	NC_015401	ERR375848
	Phrymaceae	<i>Erythranthe lutea</i>	NC_030212	SRR5307620
	Plantaginaceae	<i>Digitalis lanata</i>	NC_034688	SRR5602573
	Verbenaceae	<i>Aloysia citrodora</i>	NC_034695	SRR5602597
Solanales	Convolvulaceae	<i>Ipomoea batatas</i>	NC_026703	SRR7868503
		<i>Ipomoea trifida</i>	NC_034670	SRR6667669
	Solanaceae	<i>Nicotiana otophthora</i>	NC_032724	SRR1171700
		<i>Solanum berthaultii</i>	NC_034951	SRR5349620
		<i>Solanum commersonii</i>	NC_028069	SRR5349609
		<i>Solanum melongena</i>	NC_030207	DRR014074

Continued on next page

Table S 1 – continued from previous page

Order	Family	Species	NCBI Nucl.	NCBI SRA
Angiosperms – Eudicots – Asterids				
Brassicales	Brassicaceae	<i>Arabidopsis thaliana</i>	NC_023367	SRR7880736
		<i>Brassica napus</i>	NC_016734	SRR6856349
		<i>Brassica nigra</i>	NC_030450	SRR2054762
		<i>Arabidopsis halleri</i>	NC_034366	SRR8040824
		<i>Arabidopsis lyrata</i>	NC_034365	SRR8157563
		<i>Arabidopsis thaliana</i>	NC_000932	SRR3340908
		<i>Capsella bursa-pastoris</i>	NC_009270	SRR5412136
		<i>Capsella grandiflora</i>	NC_028517	SRR1508428
		<i>Capsella rubella</i>	NC_027693	ERR636124
		<i>Thlaspi arvense</i>	NC_034362	SRR1034659
Cucurbitales	Cucurbitaceae	<i>Cucumis melo</i> subsp. <i>melo</i>	NC_015983	ERR246536
Fabales	Fabaceae	<i>Cicer arietinum</i>	NC_011163	SRR5183095
		<i>Vicia sativa</i>	NC_027155	ERR413103
		<i>Glycyrrhiza glabra</i>	NC_024038	SRR8690419
		<i>Lotus japonicus</i>	NC_002694	SRR8115522
		<i>Cajanus cajan</i>	NC_031429	SRR073338
		<i>Glycine soja</i>	NC_022868	SRR7725010
		<i>Phaseolus vulgaris</i>	NC_009259	SRR5807695
		<i>Vigna unguiculata</i>	NC_018051	SRR7125688
		<i>Wisteria floribunda</i>	NC_027677	SRR1265941
		<i>Betula nana</i>	NC_033978	ERR2026268
Geraniales	Betulaceae	<i>Ostrya rehderiana</i>	NC_028349	SRR8302715
	Juglandaceae	<i>Juglans regia</i>	NC_028617	SRR2057822
	Francoaceae	<i>Melanthus villosus</i>	NC_023256	SRR576532
Geraniales	Geraniaceae	<i>Pelargonium citronellum</i>	NC_031194	SRR576534
	Chrysobalanaceae	<i>Chrysobalanus icaco</i>	NC_024061	SRR1179655
Malpighiales	Chrysobalanaceae	<i>Couepia guianensis</i>	NC_024063	SRR1179648
		<i>Hirtella physophora</i>	NC_024066	SRR1179646
		<i>Hirtella racemosa</i>	NC_024060	SRR1179649
		<i>Licania alba</i>	NC_024064	SRR1179652
		<i>Licania sprucei</i>	NC_024065	SRR1179651
		<i>Parinari campestris</i>	NC_024067	SRR1179645
		<i>Euphorbia esula</i>	NC_033910	SRR6355713
	Euphorbiaceae	<i>Populus alba</i>	NC_008235	SRR3678826
		<i>Populus tremula</i>	NC_027425	ERR1735633
		<i>Populus tremula</i> x <i>alba</i>	NC_028504	SRR1653109
		<i>Populus trichocarpa</i>	NC_009143	SRR5467392
Malvales	Cytinaceae	<i>Cytinus hypocistis</i>	NC_031150	ERR964904
Malvales	Malvaceae	<i>Abelmoschus esculentus</i>	NC_035234	SRR5812498
		<i>Althaea officinalis</i>	NC_034701	SRR5602596
		<i>Gossypium herbaceum</i>	NC_023215	SRR2012721
		<i>Hibiscus syriacus</i>	NC_026909	SRR1265942
		<i>Punica granatum</i>	NC_035240	SRR5812494
Myrtales	Lythraceae	<i>Pimenta dioica</i>	NC_034684	SRR5602585
Rosales	Myrtaceae	<i>Cannabis sativa</i>	NC_026562	SRR7285294
		<i>Humulus lupulus</i>	NC_028032	SRR8690418
	Cannabaceae	<i>Ficus carica</i>	NC_035237	SRR5678803
		<i>Ficus racemosa</i>	NC_028185	SRR1405699
	Moraceae	<i>Prunus dulcis</i>	NC_034696	SRR5602582
		<i>Prunus kansuensis</i>	NC_023956	SRR3138168
		<i>Prunus mume</i>	NC_023798	SRR8240066
	Rosaceae			

Continued on next page

Table S 1 – continued from previous page

Order	Family	Species	NCBI Nucl.	NCBI SRA
Sapindales	Anacardiaceae	<i>Prunus yedoensis</i>	NC_026980	DRR169775
		<i>Malus prunifolia</i>	NC_031163	SRR3571175
		<i>Fragaria chiloensis</i>	NC_019601	SRR1612828
		<i>Fragaria vesca subsp. bracteata</i>	NC_018766	SRR866156
		<i>Fragaria virginiana</i>	NC_019602	SRR5602605
		<i>Anacardium occidentale</i>	NC_035235	SRR5812497
		<i>Mangifera indica</i>	NC_035239	SRR5812496
		<i>Pistacia vera</i>	NC_034998	SRR4453367
		<i>Boswellia sacra</i>	NC_029420	SRR5602594
		<i>Citrus limon</i>	NC_034690	SRR5602593
Vitales	Sapindaceae	<i>Litchi chinensis</i>	NC_035238	SRR5812499
	Vitaceae	<i>Vitis amurensis</i>	NC_031383	SRR5891950
		<i>Vitis rotundifolia</i>	NC_023790	SRR5627788
Angiosperms – Eudicots – Other				
Caryophyllales	Cactaceae	<i>Carnegieia gigantea</i>	NC_027618	SRR5036293
	Chenopodiaceae	<i>Chenopodium quinoa</i>	NC_034949	SRR5938314
	Droseraceae	<i>Aldrovanda vesiculosa</i>	NC_035416	SRR7072768
		<i>Dionaea muscipula</i>	NC_035417	SRR7072324
		<i>Drosera erythrorhiza</i>	NC_035241	SRR7072766
		<i>Drosera regia</i>	NC_035415	SRR7072322
		Polygonaceae	<i>Fagopyrum tataricum</i>	NC_027161
Ranunculales	Berberidaceae	<i>Epimedium koreanum</i>	NC_029943	SRR8534306
	Ranunculaceae	<i>Hydrastis canadensis</i>	NC_034702	SRR5602606
Angiosperms – Monodicots				
Arecales	Arecaceae	<i>Elaeis guineensis</i>	NC_017602	ERR276848
		<i>Podococcus barteri</i>	NC_027276	SRR2120220
		<i>Phoenix dactylifera</i>	NC_013991	SRR2518264
Asparagales	Amaryllidaceae	<i>Allium cepa</i>	NC_024813	SRR1686960
	Asparagaceae	<i>Asparagus officinalis</i>	NC_034777	DRR056675
	Orchidaceae	<i>Apostasia odorata</i>	NC_030722	SRR6037860
		<i>Cymbidium ensifolium</i>	NC_028525	SRR6117755
		<i>Cymbidium kanran</i>	NC_029711	SRR6117756
		<i>Cymbidium lancifolium</i>	NC_029712	SRR6117757
		<i>Cymbidium macrorhizon</i>	NC_029713	SRR6117754
		<i>Dendrobium aphyllum</i>	NC_035322	SRR3932123
		<i>Dendrobium chrysanthum</i>	NC_035336	SRR3932147
		<i>Dendrobium chrysotoxum</i>	NC_028549	SRR3932141
		<i>Dendrobium crepidatum</i>	NC_035331	SRR3932124
		<i>Dendrobium denneanum</i>	NC_035324	SRR3932144
		<i>Dendrobium falconeri</i>	NC_035326	SRR3932139
		<i>Dendrobium nobile</i>	NC_029456	SRR3932122
		<i>Dendrobium parishii</i>	NC_035339	SRR3932128
		<i>Dendrobium pendulum</i>	NC_029705	SRR3932138
		<i>Dendrobium primulinum</i>	NC_035321	SRR3932120
		<i>Dendrobium wardianum</i>	NC_035329	SRR3932135
		<i>Ludisia discolor</i>	NC_030540	SRR3484539
		Dioscoreales	Dioscoreaceae	<i>Dioscorea rotundata</i>
<i>Dioscorea villosa</i>	NC_034686			SRR5602590
Liliales	Liliaceae	<i>Lilium tsingtauense</i>	NC_027675	SRR1265940
Poales	Poaceae	<i>Phyllostachys edulis</i>	NC_015817	SRR8245858
		<i>Oryza barthii</i>	NC_027460	SRR7341625

Continued on next page

Table S 1 – continued from previous page

Order	Family	Species	NCBI Nucl.	NCBI SRA
		<i>Oryza glumipatula</i>	NC_027461	DRR057991
		<i>Oryza longiglumis</i>	NC_034763	DRR056658
		<i>Oryza longistaminata</i>	NC_027462	DRR058031
		<i>Oryza meridionalis</i>	NC_016927	DRR058012
		<i>Oryza meyeriana</i>	NC_034765	DRR056670
		<i>Oryza officinalis</i>	NC_027463	DRR000604
		<i>Oryza punctata</i>	NC_027676	SRR1264539
		<i>Oryza ridleyi</i>	NC_034764	DRR056663
		<i>Oryza rufipogon</i>	NC_017835	SRR6220521
		<i>Oryza sativa</i>	NC_031333	ERR2696318
		<i>Avena sativa</i>	NC_027468	SRR6056489
		<i>Dactylis glomerata</i>	NC_027473	SRR5236602
		<i>Deschampsia antarctica</i>	NC_023533	SRR1158316
		<i>Aegilops sharonensis</i>	NC_024816	ERR359708
		<i>Arundo plinii</i>	NC_034652	SRR4319202
		<i>Eragrostis tef</i>	NC_029413	SRR1463402
		<i>Sporobolus michauxianus</i>	NC_029416	SRR4434178
		<i>Miscanthus sacchariflorus</i>	NC_028720	SRR559245
		<i>Miscanthus sinensis</i>	NC_028721	SRR4028761
		<i>Sorghum timorense</i>	NC_023800	SRR424217
		<i>Alloteropsis angusta</i>	NC_027951	SRR7528995
		<i>Alloteropsis cimicina</i>	NC_027952	SRR7529015
		<i>Alloteropsis paniculata</i>	NC_032078	SRR4051980
		<i>Alloteropsis semialata</i>	NC_027824	SRR7528994
		<i>Echinochloa crus-galli</i>	NC_028719	SRR5920285
		<i>Echinochloa oryzicola</i>	NC_024643	SRR5902661
		<i>Cenchrus americanus</i>	NC_024171	SRR5204424
		<i>Panicum capillare</i>	NC_030493	SRR485871
Angiosperms – Basal Angiosperms				
Austrobaileyales	Schisandraceae	<i>Illicium anisatum</i>	NC_034703	SRR5602608
		<i>Illicium floridanum</i>	NC_034685	SRR5602609
		<i>Illicium henryi</i>	NC_034699	SRR5602610
		<i>Illicium verum</i>	NC_034689	SRR5602607
Laurales	Lauraceae	<i>Cinnamomum verum</i>	NC_035236	SRR5812493
		<i>Laurus nobilis</i>	NC_034700	SRR5602602
Magnoliales	Magnoliaceae	<i>Magnolia biondii</i>	NC_034687	SRR5602588
Piperales	Piperaceae	<i>Piper auritum</i>	NC_034697	SRR5602592
		<i>Piper nigrum</i>	NC_034692	SRR5602591
GYMNOSPERMS				
Cupressales	Cupressaceae	<i>Juniperus cedrus</i>	NC_028190	SRR1145775
		<i>Juniperus communis</i>	NC_035068	ERR268423
	Taxaceae	<i>Taxus baccata</i>	NC_035066	ERR268425
Gnetales	Gnetaceae	<i>Gnetum gnemon</i>	NC_026301	ERR268420
Pinales	Pinaceae	<i>Picea glauca</i>	NC_028594	SRR869482
		<i>Picea abies</i>	NC_021456	ERR1727007
		<i>Picea asperata</i>	NC_032367	ERR1735493
		<i>Picea crassifolia</i>	NC_032366	ERR1735613
		<i>Pinus massoniana</i>	NC_021439	SRR7666034
		<i>Pinus sylvestris</i>	NC_035069	ERR268429
		<i>Pinus taeda</i>	NC_021440	SRR1049756

Table S 2: Metadata, sequencing depth, E-score, and assembly quality metrics of each plastid genome analyzed. For each genome, the plastome-wide E-score, WRSD counts by genome partition and coding status, assembly quality metrics, sequencing platform, and assembly software are displayed. Dashes indicate missing or unresolvable values. Abbreviations: Asm. = Assembler/Assembly; Cod. = Coding; GA = Genome Analyzer; Mism. = Mismatch; Non. = Non-coding; Seq. = Sequencing.

Sample	Metadata		E-score	WRSD						Asm. quality	
	Seq. platform	Asm. software		LSC	IR _B	SSC	IR _A	Cod.	Non.	Ns	Mism.
NC_000932	HiSeq 1500	-	0.97	-	-	-	-	0.52	0.34	0	-
NC_002694	NextSeq 500	-	0.92	0.61	0.36	0.66	0.52	0.51	0.90	0	0
NC_007578	HiSeq X Ten	-	0.85	0.59	0.24	0.60	0.20	0.41	0.56	0	0
NC_007977	GA II	-	0.88	0.60	1.19	0.72	1.11	0.39	0.59	0	0
NC_008235	HiSeq 2000	-	0.94	0.64	1.06	0.55	1.02	0.61	0.27	0	0
NC_009143	HiSeq 2500	-	0.96	0.59	1.13	0.55	0.62	0.55	0.40	0	0
NC_009259	HiSeq 4000	-	0.95	0.63	0.38	0.52	0.38	0.69	0.73	0	0
NC_009270	MiSeq	-	0.94	0.54	0.92	0.62	0.69	0.58	0.83	0	0
NC_011163	HiSeq 2000	-	0.88	-	-	-	-	0.52	1.01	0	-
NC_013991	HiSeq 2500	-	0.91	0.53	0.63	0.75	0.22	0.34	0.63	0	0
NC_015401	MiSeq	-	0.95	0.50	0.79	0.17	0.59	0.43	0.21	0	0
NC_015817	HiSeq X Ten	-	0.84	0.51	0.42	0.79	0.55	0.29	0.30	0	0
NC_015983	GA IIx	Newbler	0.91	0.59	1.05	0.28	1.01	0.51	0.50	0	0
NC_016734	HiSeq 3000	-	0.78	0.62	0.81	0.68	0.77	0.41	0.65	0	1
NC_016927	HiSeq 2500	-	0.97	0.59	0.68	0.74	0.24	0.98	0.16	0	0
NC_017602	HiSeq 2000	-	0.03	3.97	3.35	4.02	3.42	3.70	3.98	0	0
NC_017835	HiSeq X Ten	-	0.97	0.64	0.68	0.74	0.29	0.48	0.24	0	0
NC_018051	HiSeq 4000	MIRA	0.90	0.57	0.64	0.41	0.38	0.42	0.87	0	31
NC_018766	HiSeq 2000	YASRA	0.76	-	-	-	-	0.08	0.19	206	-
NC_019601	GA IIx	-	0.90	0.54	0.66	0.50	0.66	0.56	0.66	0	0
NC_019602	MiSeq	-	0.95	0.61	0.78	0.67	0.39	0.65	0.27	1	0
NC_019616	GA IIx	GS_denovo	0.69	0.33	0.00	0.33	0.00	0.00	0.01	0	0
NC_020098	HiSeq 4000	-	0.95	0.61	0.76	0.52	0.60	0.56	0.30	0	0
NC_021439	HiSeq X Ten	-	0.96	-	-	-	-	0.38	0.26	0	-
NC_021440	MiSeq	-	0.91	-	-	-	-	0.88	1.01	0	-
NC_021456	HiSeq 2500	-	0.85	-	-	-	-	0.87	0.92	0	-
NC_022868	HiSeq 2000	CLC	0.95	0.48	0.35	0.57	0.24	0.52	0.76	0	0
NC_023102	HiSeq 2500	Velvet	0.95	0.50	0.47	0.59	0.31	0.53	0.12	0	0
NC_023215	MiSeq	-	0.93	0.49	0.79	0.55	0.91	0.70	0.84	0	0
NC_023256	HiSeq 2000	Velvet	0.93	0.39	0.76	0.45	0.72	0.51	0.41	0	0
NC_023367	HiSeq 2500	-	0.89	0.70	0.76	0.62	0.57	0.49	0.92	0	0
NC_023533	HiSeq 2000	-	0.93	0.74	0.90	0.56	0.90	0.60	0.17	0	0
NC_023790	HiSeq 2500	SOAPdenovo	0.95	0.60	0.88	0.58	0.61	0.51	0.42	0	0
NC_023798	HiSeq 4000	SOAPdenovo	0.90	0.63	0.46	0.48	0.23	0.44	1.13	0	0
NC_023800	HiSeq 2000	SOAPdenovo	0.94	0.46	0.59	0.16	0.50	0.76	0.37	1	1
NC_023956	HiSeq 2000	SOAPdenovo	0.90	0.60	0.96	0.58	0.99	0.48	0.80	0	0
NC_024038	NextSeq 500	Velvet	0.84	-	-	-	-	0.57	1.06	1	-
NC_024060	HiSeq 2000	-	0.93	0.67	0.83	0.61	0.83	0.65	0.40	0	0
NC_024061	HiSeq 2000	-	0.94	0.58	0.90	0.56	0.75	0.66	0.29	0	0
NC_024063	HiSeq 2000	-	0.94	0.62	0.64	0.68	0.68	0.52	0.25	0	0
NC_024064	HiSeq 2000	-	0.93	0.57	0.79	0.72	0.75	0.56	0.24	0	0
NC_024065	HiSeq 2000	-	0.92	0.61	0.90	0.62	0.90	0.63	0.27	0	0
NC_024066	HiSeq 2000	-	0.95	0.53	0.56	0.81	0.45	0.47	0.43	0	0
NC_024067	HiSeq 2000	-	0.92	0.56	0.83	0.31	0.75	0.56	0.44	0	0
NC_024170	MiSeq	MIRA	0.10	1.62	2.03	1.55	1.39	1.86	1.11	1	26
NC_024171	HiSeq 2500	MIRA	0.90	0.51	0.88	0.48	0.77	0.66	0.53	6	0

Continued on next page

Table S 2 – continued from previous page

Sample	Metadata		E-score	WRSD						Asm. quality	
	Seq. platform	Asm. software		LSC	IR _B	SSC	IR _A	Cod.	Non.	Ns	Mism.
NC_024286	HiSeq 2000	Ray	0.97	0.53	0.20	0.99	0.36	0.62	0.74	0	2
NC_024643	HiSeq 2500	CLC	0.93	0.58	0.40	0.72	0.36	0.43	0.68	0	0
NC_024813	HiSeq 2000	Newbler	0.92	0.66	0.92	0.74	0.88	0.60	0.41	0	0
NC_024816	HiSeq 2000	Consed	0.95	0.64	0.65	0.55	0.23	0.42	0.66	0	0
NC_026291	HiSeq 2500	-	0.95	0.36	0.43	0.29	0.43	0.55	0.62	0	0
NC_026301	HiSeq 2000	Velvet	0.95	0.51	0.70	0.61	0.55	0.38	0.58	0	0
NC_026562	NextSeq 500	SOAPdenovo	0.63	0.23	0.31	0.79	0.27	0.01	0.00	0	1
NC_026690	HiSeq 4000	Velvet	0.95	0.63	0.75	0.74	0.54	0.60	0.39	0	0
NC_026703	HiSeq 2500	-	0.96	0.35	0.72	0.25	0.56	0.59	0.23	0	32
NC_026909	HiSeq 2000	GS_denovo	0.94	0.68	1.06	0.61	0.94	0.71	0.95	0	-
NC_026980	MiSeq	Geneious	0.97	0.49	0.57	0.63	0.15	0.52	0.64	0	0
NC_027155	HiSeq 1000	-	0.88	-	-	-	-	0.50	0.31	0	-
NC_027161	HiSeq 2500	CLC	0.94	0.63	0.91	0.54	0.85	0.50	0.25	0	2
NC_027262	HiSeq 2500	-	0.95	0.43	0.55	0.58	0.75	0.53	0.69	0	1
NC_027276	MiSeq	-	0.50	2.02	2.29	2.12	1.91	1.93	2.34	91	0
NC_027425	HiSeq 2500	CLC	0.19	3.16	2.67	3.83	2.52	2.90	3.52	0	0
NC_027460	HiSeq 2500	-	0.93	0.69	1.11	0.74	0.68	0.83	0.49	0	0
NC_027461	HiSeq 2500	-	0.97	0.45	0.48	0.82	0.19	0.63	0.40	0	0
NC_027462	HiSeq 2500	-	0.97	0.66	0.58	0.74	0.19	0.59	0.31	0	0
NC_027463	GA II	-	0.83	0.47	0.82	0.25	0.73	0.59	0.33	0	0
NC_027468	MiSeq	Velvet	0.97	0.50	0.37	0.56	0.05	0.35	0.33	0	0
NC_027473	HiSeq X Ten	Velvet	0.89	0.68	0.66	0.74	0.52	0.58	0.66	0	0
NC_027618	MiSeq	SOAPdenovo	0.91	-	-	-	-	0.88	0.71	0	-
NC_027675	HiSeq 2000	CLC	0.93	0.71	0.91	0.75	0.80	0.68	0.53	0	0
NC_027676	HiSeq 2000	CLC	0.95	-	-	-	-	0.69	0.29	0	-
NC_027677	HiSeq 2000	-	0.92	-	-	-	-	0.59	0.26	0	-
NC_027693	GA II	CLC	0.91	0.54	0.92	0.74	0.88	0.54	0.75	0	0
NC_027824	HiSeq 2500	-	0.96	0.22	0.37	0.32	0.23	0.38	0.51	0	1
NC_027951	HiSeq 2500	-	0.96	0.53	0.39	0.16	0.09	0.57	0.25	0	0
NC_027952	HiSeq 2500	-	0.97	0.39	0.27	0.72	0.18	0.32	0.26	0	1
NC_028032	NextSeq 500	SPAdes	0.91	0.66	0.70	0.69	0.74	0.52	0.60	0	0
NC_028069	HiSeq 2500	-	0.95	0.67	1.06	0.77	0.87	0.53	0.33	0	0
NC_028185	GA	SPAdes	0.73	0.24	0.08	0.60	0.08	0.39	0.31	2	1
NC_028190	HiSeq 2000	-	0.96	-	-	-	-	0.40	0.61	0	-
NC_028349	HiSeq 2500	-	0.89	0.63	0.39	0.70	0.55	0.44	0.76	0	4
NC_028504	HiSeq 2000	CLC	0.96	0.58	0.61	0.51	0.32	0.50	0.54	0	0
NC_028517	HiSeq 2000	CLC	0.94	0.66	0.88	0.62	0.92	0.65	0.53	0	0
NC_028525	MiSeq	-	0.90	0.53	0.22	1.03	0.22	0.37	0.78	0	0
NC_028549	HiSeq 4000	-	0.87	0.71	0.42	0.61	0.34	0.50	0.92	0	0
NC_028594	HiSeq 2000	-	0.87	-	-	-	-	0.61	0.45	0	-
NC_028617	HiSeq 2500	-	0.97	0.42	0.70	0.44	0.31	0.54	0.25	0	0
NC_028719	HiSeq 2000	-	0.93	0.70	0.31	0.64	0.40	0.53	0.90	0	0
NC_028720	GA II	-	0.95	0.48	0.44	0.48	0.26	0.64	0.67	0	0
NC_028721	GA II	-	0.93	0.69	0.75	0.80	0.62	0.78	1.06	0	2
NC_029413	HiSeq 2000	Geneious	0.87	0.78	0.38	0.88	0.43	0.43	1.21	0	0
NC_029416	HiSeq 2000	Velvet	0.94	0.61	0.73	0.56	0.68	0.61	0.53	0	0
NC_029420	MiSeq	-	0.92	0.40	0.83	0.59	0.75	0.60	0.90	0	0
NC_029456	HiSeq 4000	-	0.92	0.55	0.27	0.48	0.30	0.44	0.78	0	12830
NC_029705	HiSeq 4000	-	0.89	0.56	0.34	0.83	-	0.48	1.06	0	-
NC_029711	MiSeq	Geneious	0.96	0.60	0.67	0.73	0.67	0.65	0.39	0	0
NC_029712	MiSeq	Geneious	0.96	0.50	0.51	0.66	0.43	0.74	0.33	0	0
NC_029713	MiSeq	Geneious	0.44	0.00	0.00	0.29	0.00	0.00	0.00	0	0

Continued on next page

Table S 2 – continued from previous page

Sample	Metadata		E-score	WRSD						Asm. quality	
	Seq. platform	Asm. software		LSC	IR _B	SSC	IR _A	Cod.	Non.	Ns	Mism.
NC_029819	MiSeq	YASRA	0.89	-	-	-	-	0.47	0.54	130	-
NC_029822	MiSeq	YASRA	0.97	-	-	-	-	0.39	0.29	0	-
NC_029823	MiSeq	YASRA	0.91	-	-	-	-	0.50	0.45	102	-
NC_029824	MiSeq	YASRA	0.93	-	-	-	-	0.44	0.28	56	-
NC_029825	MiSeq	YASRA	0.90	-	-	-	-	0.55	0.79	551	-
NC_029943	NextSeq 500	CLC	0.87	0.67	0.64	0.83	0.64	0.58	0.71	0	0
NC_030053	unspecified	Newbler	0.92	0.65	0.62	0.61	0.58	0.46	0.36	0	38
NC_030173	HiSeq 2000	SOAPdenovo	0.95	0.49	0.84	0.46	0.57	0.50	0.46	0	0
NC_030207	HiSeq 2000	-	0.78	0.56	0.35	0.61	0.28	0.52	0.57	1	6
NC_030212	HiSeq 2000	-	0.92	0.52	0.67	0.74	0.47	0.51	0.51	0	2
NC_030275	HiSeq 2000	SOAPdenovo	0.96	0.51	0.86	0.83	0.86	0.52	0.36	0	2
NC_030450	HiSeq 2500	CLC	0.90	0.55	0.62	0.69	0.50	0.70	0.77	0	0
NC_030493	-	Velvet	0.93	0.60	0.90	0.80	0.75	0.72	0.19	0	0
NC_030540	MiSeq	-	0.94	0.38	0.72	0.35	0.15	0.61	0.31	0	23
NC_030722	HiSeq 4000	-	0.86	0.64	0.52	0.70	0.56	0.56	0.92	0	0
NC_030756	HiSeq 2500	-	0.96	0.59	0.79	0.69	0.83	0.62	0.23	0	0
NC_030783	HiSeq 1500	CLC	0.93	0.57	1.04	0.60	0.88	0.59	0.36	0	1
NC_030786	HiSeq 2000	FLASH	0.93	0.51	0.35	0.77	0.50	0.48	0.85	0	1
NC_031150	HiSeq 2000	OBITools	0.88	-	-	-	-	0.87	0.60	0	-
NC_031163	HiSeq 2500	-	0.90	0.55	0.96	0.74	0.96	0.59	0.53	0	0
NC_031194	HiSeq 2000	Velvet	0.90	0.59	0.84	0.59	0.84	0.56	0.36	0	0
NC_031333	MiSeq	CLC	0.96	0.49	0.87	0.74	0.05	0.72	0.31	0	0
NC_031383	HiSeq 4000	-	0.96	0.62	1.11	0.63	0.84	0.48	0.44	0	0
NC_031428	HiSeq 2000	-	0.84	0.64	0.55	0.45	0.67	0.34	1.21	0	55
NC_031429	GA II	SPAdes	0.84	0.56	0.43	0.29	0.43	0.64	0.78	0	0
NC_032057	MiSeq	Velvet	0.86	0.62	0.66	0.88	0.41	0.43	0.91	0	2
NC_032078	HiSeq 2500	-	0.95	0.67	0.05	0.64	0.32	0.62	0.35	0	8
NC_032366	HiSeq 2500	-	0.89	-	-	-	-	0.67	0.51	0	-
NC_032367	HiSeq 2500	-	0.91	-	-	-	-	0.81	0.83	0	-
NC_032724	HiSeq 2500	-	0.94	0.55	0.31	0.75	0.35	0.37	0.52	0	284
NC_033910	HiSeq 2000	PriceTI	0.89	0.74	0.88	0.83	0.84	0.58	0.26	0	0
NC_033978	NextSeq 500	-	0.90	-	0.49	-	-	0.45	0.92	0	-
NC_034320	MiSeq	-	0.90	0.64	0.76	0.73	0.60	0.61	0.51	0	0
NC_034321	MiSeq	-	0.79	0.57	0.84	0.84	0.48	0.70	0.57	0	0
NC_034322	MiSeq	-	0.92	0.72	0.52	0.56	0.48	0.45	0.54	0	0
NC_034323	MiSeq	-	0.92	0.57	0.72	0.73	0.52	0.72	0.31	0	0
NC_034324	MiSeq	-	0.84	0.57	0.76	0.67	0.52	0.48	0.38	0	0
NC_034325	MiSeq	-	0.91	0.61	0.73	0.61	0.56	0.48	0.38	0	3
NC_034326	MiSeq	-	0.86	0.75	0.92	0.56	0.72	0.55	0.55	0	0
NC_034327	MiSeq	-	0.88	0.69	0.72	0.61	0.60	0.50	0.67	0	0
NC_034328	MiSeq	-	0.90	0.54	0.88	0.73	0.72	0.57	0.40	0	0
NC_034362	MiSeq	-	0.94	0.64	0.27	0.52	0.15	0.47	0.18	0	2
NC_034365	HiSeq 4000	-	0.91	0.66	0.76	0.74	0.65	0.62	0.65	0	1
NC_034366	HiSeq 2500	-	0.97	0.45	0.84	0.34	0.46	0.53	0.24	0	0
NC_034652	MiSeq	SPAdes	0.80	0.64	0.53	0.33	0.48	0.11	0.51	0	1
NC_034670	HiSeq 2500	-	0.92	0.54	0.59	0.59	0.59	0.52	0.71	0	2
NC_034683	MiSeq	-	0.92	0.69	0.57	0.55	0.41	0.54	0.46	0	0
NC_034684	MiSeq	-	0.92	0.62	0.50	0.71	0.42	0.63	0.66	0	0
NC_034685	MiSeq	-	0.94	0.63	0.64	0.50	0.09	0.60	0.34	2	0
NC_034686	MiSeq	-	0.94	0.64	0.71	0.64	0.71	0.69	0.50	0	0
NC_034687	MiSeq	-	0.93	0.66	0.72	0.75	0.57	0.75	0.40	2	0
NC_034688	MiSeq	-	0.94	0.55	0.63	0.63	0.55	0.68	0.30	0	0

Continued on next page

Table S 2 – continued from previous page

Sample	Metadata		E-score	WRSD						Asm. quality	
	Seq. platform	Asm. software		LSC	IR _B	SSC	IR _A	Cod.	Non.	Ns	Mism.
NC_034689	MiSeq	-	0.94	0.59	0.46	0.50	0.64	0.61	0.75	0	0
NC_034690	MiSeq	-	0.95	0.57	0.60	0.66	0.45	0.52	0.40	2	0
NC_034691	MiSeq	-	0.93	0.75	0.48	0.61	0.54	0.59	0.52	0	0
NC_034692	MiSeq	-	0.95	0.62	0.82	0.66	0.33	0.55	0.38	0	0
NC_034694	MiSeq	-	0.94	0.55	0.54	0.70	0.61	0.51	0.27	87	4
NC_034695	MiSeq	-	0.94	0.59	0.90	0.64	0.74	0.77	0.31	2	0
NC_034696	MiSeq	-	0.95	0.62	0.69	0.80	0.50	0.62	0.45	2	0
NC_034697	MiSeq	-	0.93	0.62	0.71	0.55	0.52	0.72	0.50	2	0
NC_034698	MiSeq	-	0.96	-	-	-	-	0.46	0.20	0	-
NC_034699	MiSeq	-	0.90	0.62	0.65	0.71	0.84	0.53	0.46	0	0
NC_034700	MiSeq	-	0.95	0.61	0.60	0.11	0.10	0.50	0.27	0	0
NC_034701	MiSeq	-	0.94	0.58	0.60	0.72	0.56	0.65	0.34	0	0
NC_034702	MiSeq	-	0.94	0.61	0.52	0.59	0.30	0.70	0.49	0	0
NC_034703	MiSeq	-	0.95	0.65	0.75	0.70	0.65	0.63	0.49	0	0
NC_034763	HiSeq 2500	-	0.96	0.75	0.39	0.57	0.15	0.66	0.18	0	0
NC_034764	HiSeq 2500	-	0.96	0.62	0.58	0.49	0.19	0.52	0.21	0	0
NC_034765	HiSeq 2500	-	0.97	0.65	0.29	0.49	0.24	0.69	0.11	0	0
NC_034777	HiSeq 2500	-	0.98	0.31	0.61	0.27	0.04	0.64	0.38	0	0
NC_034949	HiSeq 2500	CLC	0.94	0.60	0.76	0.67	0.72	0.66	0.27	0	0
NC_034951	HiSeq 2500	CLC	0.94	0.68	1.10	0.61	0.98	0.50	0.37	0	0
NC_034998	HiSeq 2000	-	0.97	0.41	0.42	0.63	0.27	0.40	0.35	0	0
NC_035066	HiSeq 2000	ABYSS	0.95	-	-	-	-	0.42	0.64	0	-
NC_035068	HiSeq 2000	ABYSS	0.95	-	-	-	-	0.61	0.30	0	-
NC_035069	HiSeq 2000	ABYSS	0.95	-	-	-	-	0.53	0.24	0	-
NC_035143	HiSeq 2500	Velvet	0.97	0.43	0.94	0.57	0.79	0.56	0.27	0	0
NC_035234	HiSeq 2500	Velvet	0.92	-	-	-	-	0.52	0.85	0	-
NC_035235	HiSeq 2500	Velvet	0.95	0.57	0.71	0.74	0.74	0.59	0.63	0	0
NC_035236	HiSeq 2500	Velvet	0.94	0.64	0.35	0.64	0.30	0.54	0.74	0	0
NC_035237	MiSeq	Velvet	0.86	0.50	0.51	0.60	0.39	0.48	0.86	0	0
NC_035238	HiSeq 2500	Velvet	0.91	0.71	0.44	0.73	0.47	0.44	0.68	0	0
NC_035239	HiSeq 2500	Velvet	0.93	0.71	0.50	0.44	0.38	0.41	0.76	0	0
NC_035240	HiSeq 2500	Velvet	0.92	0.73	0.36	0.71	0.40	0.40	0.59	0	0
NC_035241	HiSeq 1500	-	0.95	0.62	0.73	0.49	0.73	0.78	0.61	0	0
NC_035321	HiSeq 4000	-	0.92	0.50	0.27	0.80	0.23	0.29	0.63	0	0
NC_035322	HiSeq 4000	-	0.91	0.51	0.27	0.49	0.27	0.31	0.57	0	0
NC_035324	HiSeq 4000	-	0.93	0.48	0.34	0.28	0.31	0.38	0.77	0	0
NC_035326	HiSeq 4000	-	0.91	0.58	0.15	0.35	0.27	0.32	0.74	0	0
NC_035329	HiSeq 4000	-	0.87	0.69	0.38	0.70	0.38	0.51	0.83	0	2
NC_035331	HiSeq 4000	-	0.91	0.54	0.38	0.28	0.27	0.40	0.79	0	2
NC_035336	HiSeq 4000	-	0.92	0.45	0.23	0.35	0.27	0.37	0.72	0	0
NC_035339	HiSeq 4000	-	0.91	0.46	0.38	0.92	0.38	0.32	0.50	0	0
NC_035415	HiSeq 1500	-	0.97	0.39	0.60	0.60	0.43	0.61	0.33	0	0
NC_035416	HiSeq 1500	-	0.95	0.64	0.78	0.60	0.67	0.72	0.48	0	0
NC_035417	HiSeq 1500	-	0.97	-	-	-	-	0.53	0.44	0	-

Table S 3: Pairwise Wilcoxon rank-sum test results regarding differences in sequencing evenness. Displayed are the p-values and effect sizes for each comparison. Prefix removal and abbreviations follow Table 1, significance and effect size indicators Table 3 of the main text.

State 1	State 2	n1	n2	p.adj	d_c
GA II	HiSeq 1500	6	6	0.035 *	2.198 ^l
GA II	HiSeq 2000	6	41	0.087	0.634 ^m
GA II	HiSeq 2500	6	44	0.035 *	0.779 ^m
GA II	HiSeq 4000	6	19	0.365	0.471 ^s
GA II	HiSeq X Ten	6	5	0.741	0.335 ^s
GA II	MiSeq	6	52	0.142	0.483 ^s
GA II	NextSeq 500	6	5	0.821	0.222 ^s
HiSeq 1500	HiSeq 2000	6	41	0.055	0.711 ^m
HiSeq 1500	HiSeq 2500	6	44	0.343	0.352 ^s
HiSeq 1500	HiSeq 4000	6	19	0.030 *	1.352 ^l
HiSeq 1500	HiSeq X Ten	6	5	0.209	1.141 ^l
HiSeq 1500	MiSeq	6	52	0.052	0.672 ^m
HiSeq 1500	NextSeq 500	6	5	0.030 *	2.928 ^l
HiSeq 2000	HiSeq 2500	41	44	0.098	0.437 ^s
HiSeq 2000	HiSeq 4000	41	19	0.092	0.542 ^m
HiSeq 2000	HiSeq X Ten	41	5	0.680	0.172
HiSeq 2000	MiSeq	41	52	0.680	0.127
HiSeq 2000	NextSeq 500	41	5	0.035 *	0.828 ^l
HiSeq 2500	HiSeq 4000	44	19	0.030 *	0.839 ^l
HiSeq 2500	HiSeq X Ten	44	5	0.349	0.345 ^s
HiSeq 2500	MiSeq	44	52	0.052	0.505 ^m
HiSeq 2500	NextSeq 500	44	5	0.030 *	0.876 ^l
HiSeq 4000	HiSeq X Ten	19	5	0.787	0.160
HiSeq 4000	MiSeq	19	52	0.209	0.370 ^s
HiSeq 4000	NextSeq 500	19	5	0.216	0.657 ^m
HiSeq X Ten	MiSeq	5	52	0.680	0.150
HiSeq X Ten	NextSeq 500	5	5	0.841	0.199
MiSeq	NextSeq 500	52	5	0.068	0.618 ^m
CLC	SOAPdenovo	14	7	1.000	0.000
CLC	Velvet	14	19	0.922	0.269 ^s
CLC	YASRA	14	5	0.922	0.481 ^s
SOAPdenovo	Velvet	7	19	0.934	0.125
SOAPdenovo	YASRA	7	5	0.922	0.534 ^m
Velvet	YASRA	19	5	0.934	0.219 ^s

Table S 4: Results of linear regression analyses on sequencing depth.***Without covariates included in the model***

```
> tidy(seqDepthWOCovar_lm_fit)
  term                                estimate std.error statistic  p.value
  <chr>                                <dbl>    <dbl>    <dbl>    <dbl>
1 (Intercept)                        13.4      0.609     22.0    1.11e-91
2 StructQuadriIRb                     -0.577    0.859     -0.672  5.02e- 1
3 StructQuadriLSC                      1.79     0.864      2.07  3.87e- 2
4 StructQuadriSSC                     -7.09     0.859     -8.25  3.78e-16
5 CodingStatusnoncoding               -10.1     0.862    -11.8   1.73e-30
6 StructQuadriIRb:CodingStatusnoncoding 0.261    1.22      0.215  8.30e- 1
7 StructQuadriLSC:CodingStatusnoncoding 4.89     1.22      4.01   6.52e- 5
8 StructQuadriSSC:CodingStatusnoncoding 5.75     1.22      4.72   2.57e- 6

> summary(seqDepthWOCovar_lm_fit$fit)
Call: stats::lm(formula = N_lowCovWin ~ . + StructQuadri * CodingStatus, data = data)
Residuals:    Min       1Q   Median       3Q      Max
      -15.210   -5.210   -0.958    4.042   44.669

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      13.4207    0.6094   22.025 < 2e-16 ***
StructQuadriIRb   -0.5774    0.8592   -0.672  0.5017
StructQuadriLSC    1.7891    0.8644    2.070  0.0387 *
StructQuadriSSC   -7.0894    0.8592   -8.252 3.78e-16 ***
CodingStatusnoncoding -10.1463    0.8618  -11.774 < 2e-16 ***
StructQuadriIRb:CodingStatusnoncoding  0.2608    1.2150    0.215  0.8301
StructQuadriLSC:CodingStatusnoncoding  4.8935    1.2215    4.006 6.52e-05 ***
StructQuadriSSC:CodingStatusnoncoding  5.7479    1.2169    4.724 2.57e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 7.804 on 1307 degrees of freedom
Multiple R-squared:  0.2872,    Adjusted R-squared:  0.2834
F-statistic: 75.24 on 7 and 1307 DF,  p-value: < 2.2e-16
```

With covariates included in the model

```
> tidy(seqDepthWITHCovar_lm_fit)
  term                                estimate std.error statistic  p.value
  <chr>                                <dbl>    <dbl>    <dbl>    <dbl>
1 (Intercept)                        5.69     1.71      3.32  9.12e- 4
2 StructQuadriIRb                     -0.816    1.14     -0.714 4.75e- 1
3 StructQuadriLSC                     -13.5     2.80     -4.82  1.63e- 6
4 StructQuadriSSC                     -4.99     1.19     -4.18  3.10e- 5
5 PartitionLen                        0.000601 0.0000878 6.84   1.19e-11
6 CodingStatusnoncoding               -4.60     1.43     -3.23  1.29e- 3
7 avgReadLen                         -0.00336 0.00204   -1.65  1.00e- 1
8 StructQuadriIRb:CodingStatusnoncoding 0.581     1.62     0.359 7.20e- 1
9 StructQuadriLSC:CodingStatusnoncoding 12.1      2.00      6.06  1.81e- 9
10 StructQuadriSSC:CodingStatusnoncoding 5.59      1.62      3.44  5.98e- 4

> summary(seqDepthWITHCovar_lm_fit$fit)
Call: stats::lm(formula = N_lowCovWin ~ . + StructQuadri * CodingStatus, data = data)
Residuals:    Min       1Q   Median       3Q      Max
      -24.423   -4.964   -0.927    3.595   165.451

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.694e+00  1.713e+00   3.324 0.000912 ***
StructQuadriIRb -8.164e-01  1.144e+00  -0.714 0.475475
StructQuadriLSC -1.350e+01  2.803e+00  -4.816 1.63e-06 ***
StructQuadriSSC -4.989e+00  1.193e+00  -4.181 3.10e-05 ***
PartitionLen     6.009e-04  8.781e-05   6.843 1.19e-11 ***
CodingStatusnoncoding -4.601e+00  1.426e+00  -3.226 0.001285 **
avgReadLen     -3.364e-03  2.044e-03  -1.646 0.100062
StructQuadriIRb:CodingStatusnoncoding  5.815e-01  1.620e+00   0.359 0.719679
StructQuadriLSC:CodingStatusnoncoding  1.211e+01  1.999e+00   6.057 1.81e-09 ***
StructQuadriSSC:CodingStatusnoncoding  5.587e+00  1.624e+00   3.441 0.000598 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 10.38 on 1306 degrees of freedom (8 observations deleted due to missingness)
Multiple R-squared:  0.2442,    Adjusted R-squared:  0.239
F-statistic: 46.89 on 9 and 1306 DF,  p-value: < 2.2e-16
```

Table S 5: Results of linear regression analyses on sequencing evenness. Rows containing only NA values were omitted.**Without covariates included in the model**

```
> tidy(seqEvenWOCovar_lm_fit)
  term                estimate std.error statistic  p.value
  <chr>                <dbl>      <dbl>      <dbl>      <dbl>
1 (Intercept)          0.977      0.0656     14.9      2.63e-29
2 SeqPlatformHiSeq 1500 -0.0508    0.0809    -0.627    5.32e- 1
3 SeqPlatformHiSeq 2000 -0.435     0.115     -3.79     2.39e- 4
4 SeqPlatformHiSeq 2500 -0.00943   0.0498    -0.190    8.50e- 1
5 SeqPlatformHiSeq 4000 -0.0494    0.0524    -0.943    3.48e- 1
6 SeqPlatformHiSeq X Ten -0.00316   0.0222    -0.143    8.87e- 1
7 SeqPlatformMiSeq     -0.0355    0.0519    -0.684    4.95e- 1
8 SeqPlatformNextSeq 500 0.00382    0.0272     0.141    8.88e- 1
9 N_count              -0.0276    0.0450    -0.612    5.41e- 1
10 IR_mismatch         -0.0543    0.0736    -0.738    4.62e- 1 ## 22 more rows

> summary(seqEvenWOCovar_lm_fit$fit)
Call: stats::lm(formula = E_score ~ . + SeqPlatform * N_count * IR_mismatch, data = data)
Residuals:    Min       1Q   Median       3Q      Max
 -0.126890 -0.011477  0.003032  0.019756  0.077232
Coefficients: (14 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.977107   0.065612  14.892 < 2e-16 ***
SeqPlatformHiSeq 1500 -0.050766   0.080913  -0.627 0.531551
SeqPlatformHiSeq 2000 -0.434786   0.114860  -3.785 0.000239 ***
SeqPlatformHiSeq 2500 -0.009434   0.049778  -0.190 0.849990
SeqPlatformHiSeq 4000 -0.049400   0.052407  -0.943 0.347724
SeqPlatformHiSeq X Ten -0.003163   0.022194  -0.143 0.886916
SeqPlatformMiSeq   -0.035534   0.051926  -0.684 0.495058
SeqPlatformNextSeq 500 0.003820   0.027182   0.141 0.888461
N_count           -0.027574   0.045020  -0.612 0.541353
IR_mismatch       -0.054345   0.073623  -0.738 0.461829
SeqPlatformHiSeq 2000:N_count 0.370687   0.111882   3.313 0.001212 **
SeqPlatformHiSeq 1500:IR_mismatch 0.114639   0.089580   1.280 0.203046
SeqPlatformHiSeq 2000:IR_mismatch 0.915775   0.260399   3.517 0.000613 ***
SeqPlatformHiSeq 2500:IR_mismatch 0.063694   0.054970   1.159 0.248825
SeqPlatformHiSeq 4000:IR_mismatch 0.077229   0.057981   1.332 0.185328
SeqPlatformMiSeq:IR_mismatch 0.065839   0.056907   1.157 0.249532
N_count:IR_mismatch -0.004214   0.051475  -0.082 0.934888
SeqPlatformHiSeq 2000:N_count:IR_mismatch -0.809538   0.258534  -3.131 0.002175 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.03139 on 123 degrees of freedom (24 observations deleted due to missingness)
Multiple R-squared:  0.33, Adjusted R-squared:  0.2374
F-statistic: 3.563 on 17 and 123 DF, p-value: 1.966e-05
```

With covariates included in the model

```
> tidy(seqEvenWITHCovar_filter_lm_fit)
  term                estimate std.error statistic  p.value
  <chr>                <dbl>      <dbl>      <dbl>      <dbl>
1 (Intercept)          1.02      0.0810     12.6      8.18e-24
2 PartitionLen         -0.000000499 0.000000310 -1.61     1.10e- 1
3 N_count              -0.0285    0.0444    -0.642    5.22e- 1
4 IR_mismatch          -0.0475    0.0722    -0.657    5.12e- 1
5 SeqPlatformHiSeq 1500 -0.0196    0.0802    -0.244    8.08e- 1
6 SeqPlatformHiSeq 2000 -0.401     0.113     -3.54     5.69e- 4
7 SeqPlatformHiSeq 2500 -0.00478   0.0490    -0.0976   9.22e- 1
8 SeqPlatformHiSeq 4000 -0.0395    0.0515    -0.767    4.45e- 1
9 SeqPlatformHiSeq X Ten -0.00852   0.0218    -0.390    6.97e- 1
10 SeqPlatformMiSeq     -0.0455    0.0517    -0.881    3.80e- 1 ## 24 more rows

> summary(seqEvenWITHCovar_filter_lm_fit$fit)
Call: stats::lm(formula = E_score ~ . + SeqPlatform * N_count * IR_mismatch, data = data)
Residuals:    Min       1Q   Median       3Q      Max
 -0.118047 -0.013021  0.003608  0.017288  0.074227
Coefficients: (14 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.022e+00  8.103e-02  12.619 < 2e-16 ***
PartitionLen   -4.992e-07  3.100e-07  -1.610 0.109935
N_count        -2.848e-02  4.435e-02  -0.642 0.521958
IR_mismatch     -4.749e-02  7.224e-02  -0.657 0.512182
SeqPlatformHiSeq 1500 -1.956e-02  8.018e-02  -0.244 0.807636
SeqPlatformHiSeq 2000 -4.012e-01  1.133e-01  -3.540 0.000569 ***
SeqPlatformHiSeq 2500 -4.779e-03  4.899e-02  -0.098 0.922451
SeqPlatformHiSeq 4000 -3.952e-02  5.154e-02  -0.767 0.444695
SeqPlatformHiSeq X Ten -8.520e-03  2.185e-02  -0.390 0.697202
SeqPlatformMiSeq   -4.555e-02  5.169e-02  -0.881 0.380021
SeqPlatformNextSeq 500 1.065e-02  2.679e-02  0.398 0.691690
avgReadLen       7.657e-05  3.725e-05  2.056 0.041978 *
N_count:SeqPlatformHiSeq 2000 3.408e-01  1.105e-01  3.085 0.002521 **
IR_mismatch:SeqPlatformHiSeq 1500 7.648e-02  8.900e-02  0.859 0.391840
IR_mismatch:SeqPlatformHiSeq 2000 8.608e-01  2.564e-01  3.357 0.001052 **
IR_mismatch:SeqPlatformHiSeq 2500 5.757e-02  5.395e-02  1.067 0.288064
IR_mismatch:SeqPlatformHiSeq 4000 6.892e-02  5.691e-02  1.211 0.228216
IR_mismatch:SeqPlatformMiSeq 5.911e-02  5.583e-02  1.059 0.291820
N_count:IR_mismatch -3.561e-03  5.051e-02  -0.071 0.943910
N_count:IR_mismatch:SeqPlatformHiSeq 2000 -7.507e-01  2.545e-01  -2.949 0.003827 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.03076 on 121 degrees of freedom (24 observations deleted due to missingness)
Multiple R-squared:  0.367, Adjusted R-squared:  0.2676
F-statistic: 3.692 on 19 and 121 DF, p-value: 5.529e-06
```

Table S 6: Fit of regression tree model on sequencing depth to study data.

Without covariates included in the model

```
> seqDepthWOCovar_TREE_fit %>% collect_metrics()

  .metric .estimator .estimate .config
<chr>   <chr>      <dbl> <chr>
1 rmse    standard    7.53 Preprocessor1_Model1
2 rsq     standard    0.312 Preprocessor1_Model1
```

With covariates included in the model

```
> seqDepthWITHCovar_TREE_fit %>% collect_metrics()

  .metric .estimator .estimate .config
<chr>   <chr>      <dbl> <chr>
1 rmse    standard    7.95 Preprocessor1_Model1
2 rsq     standard    0.259 Preprocessor1_Model1
```

Table S 7: Fit of regression tree model on sequencing evenness to study data.

Without covariates included in the model

```
> seqEvenWOCovar_TREE_fit %>% collect_metrics()

  .metric .estimator .estimate .config
<chr>   <chr>      <dbl> <chr>
1 rmse    standard    0.0298 Preprocessor1_Model1
2 rsq     standard    0.126 Preprocessor1_Model1
```

With covariates included in the model

```
> seqEvenWITHCovar_TREE_fit %>% collect_metrics()

  .metric .estimator .estimate .config
<chr>   <chr>      <dbl> <chr>
1 rmse    standard    0.0281 Preprocessor1_Model1
2 rsq     standard    0.229 Preprocessor1_Model1
```

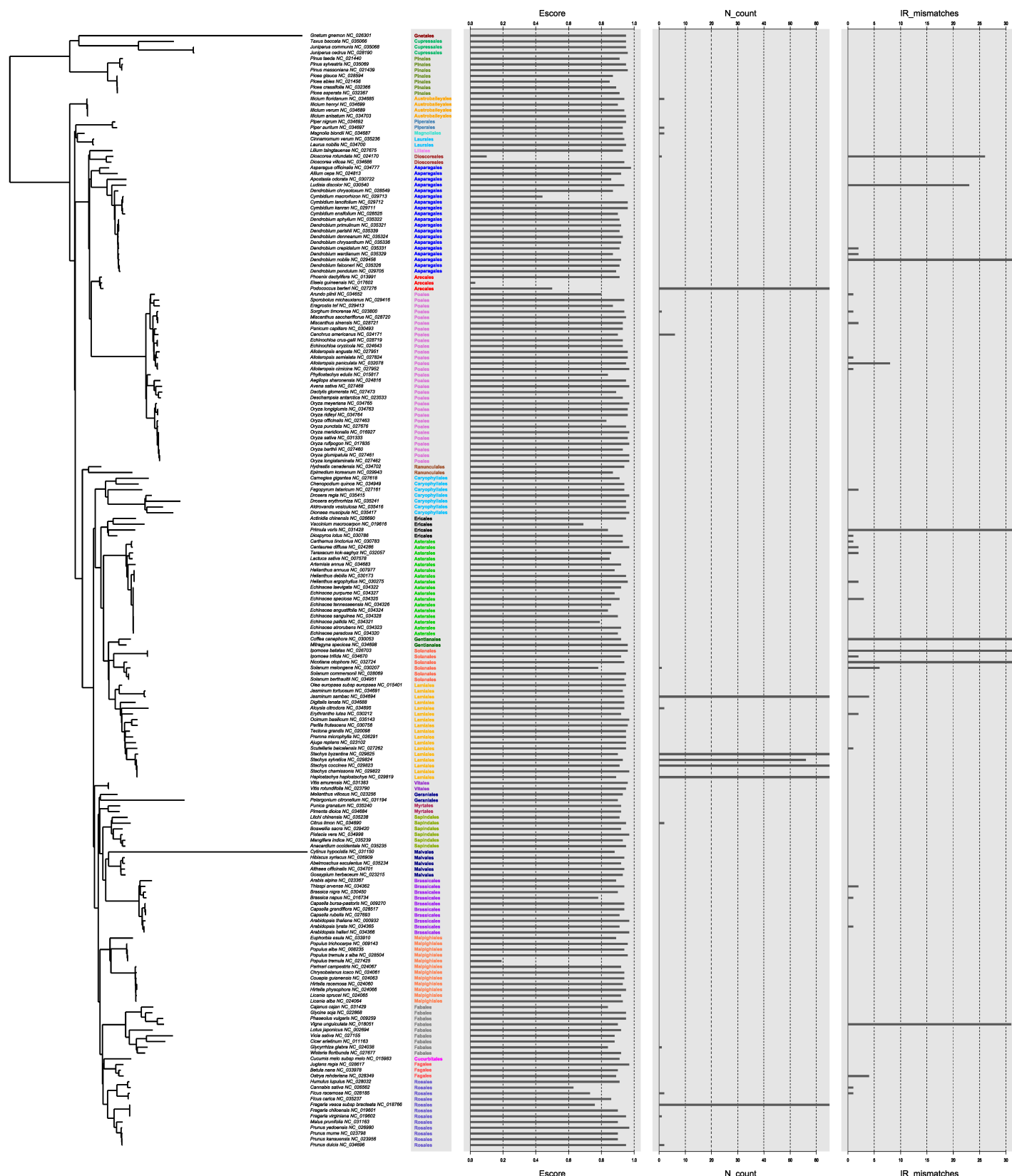
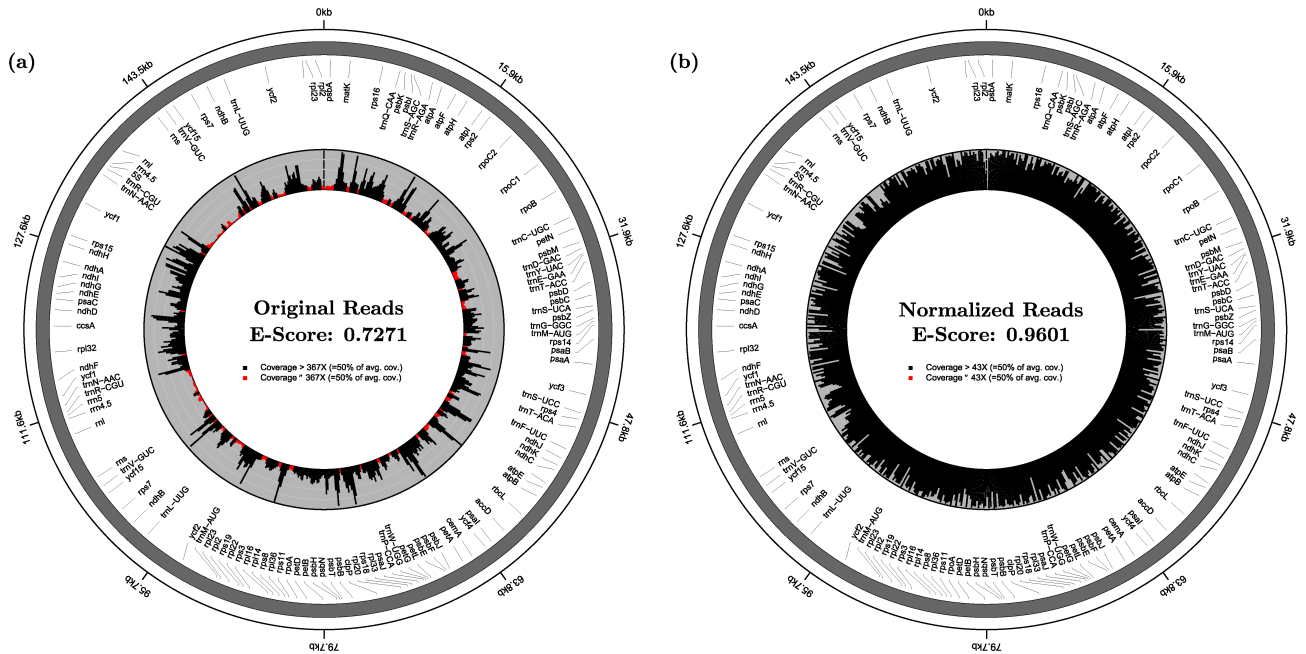


Figure S 1: E-score, ambiguous nucleotide count, and IR mismatch count for each of the 194 plastid genomes under study on a phylogenetic tree that displays the evolutionary relationships among the genomes. Plant orders are labeled in bold and assigned unique colors. Trait values are visualized as horizontal bar charts, with ambiguous nucleotide and IR mismatch counts capped at 65 and 35, respectively, for easier viewing.

Ficus racemosa chloroplast, complete genome. NC_028185



Solanum melongena plastid, complete genome. NC_030207

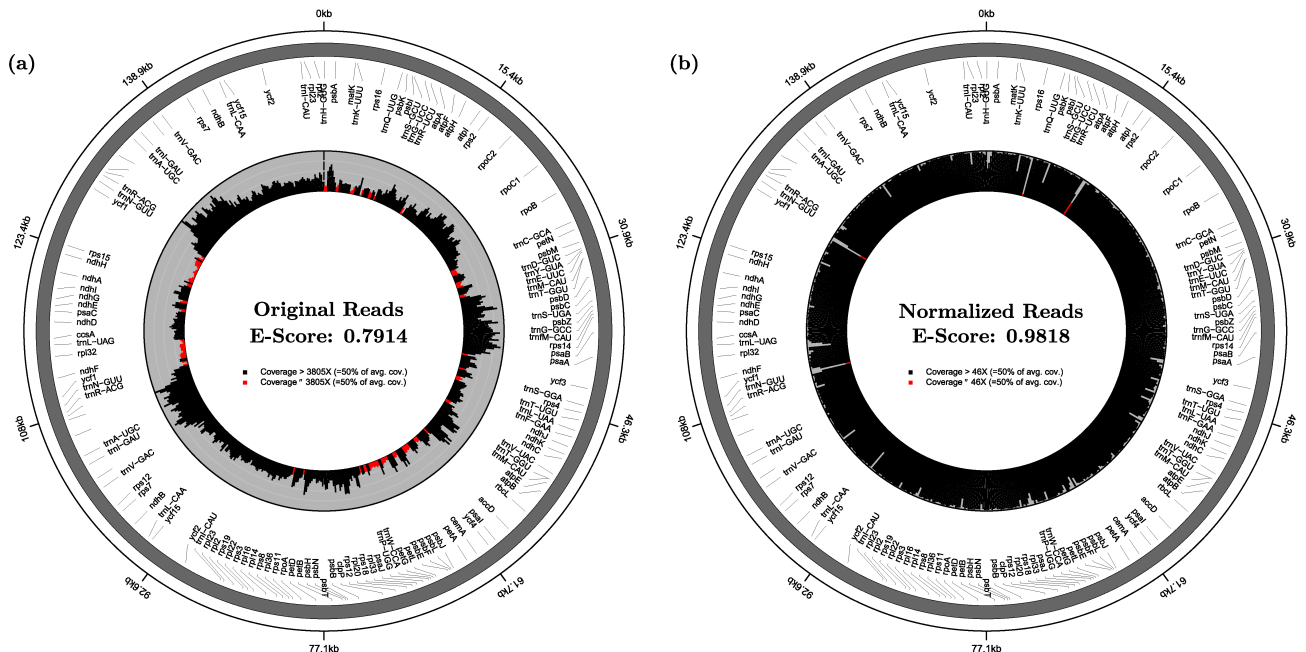
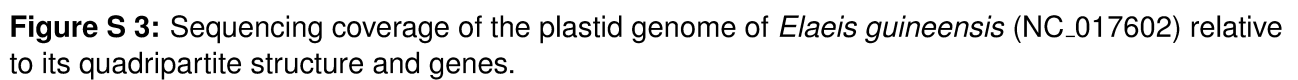


Figure S 2: Sequencing coverage of the plastid genomes of *Ficus racemosa* (NC_028185; top) and *Solanum melongena* (NC_030207; bottom) before (a) and after (b) read normalization.



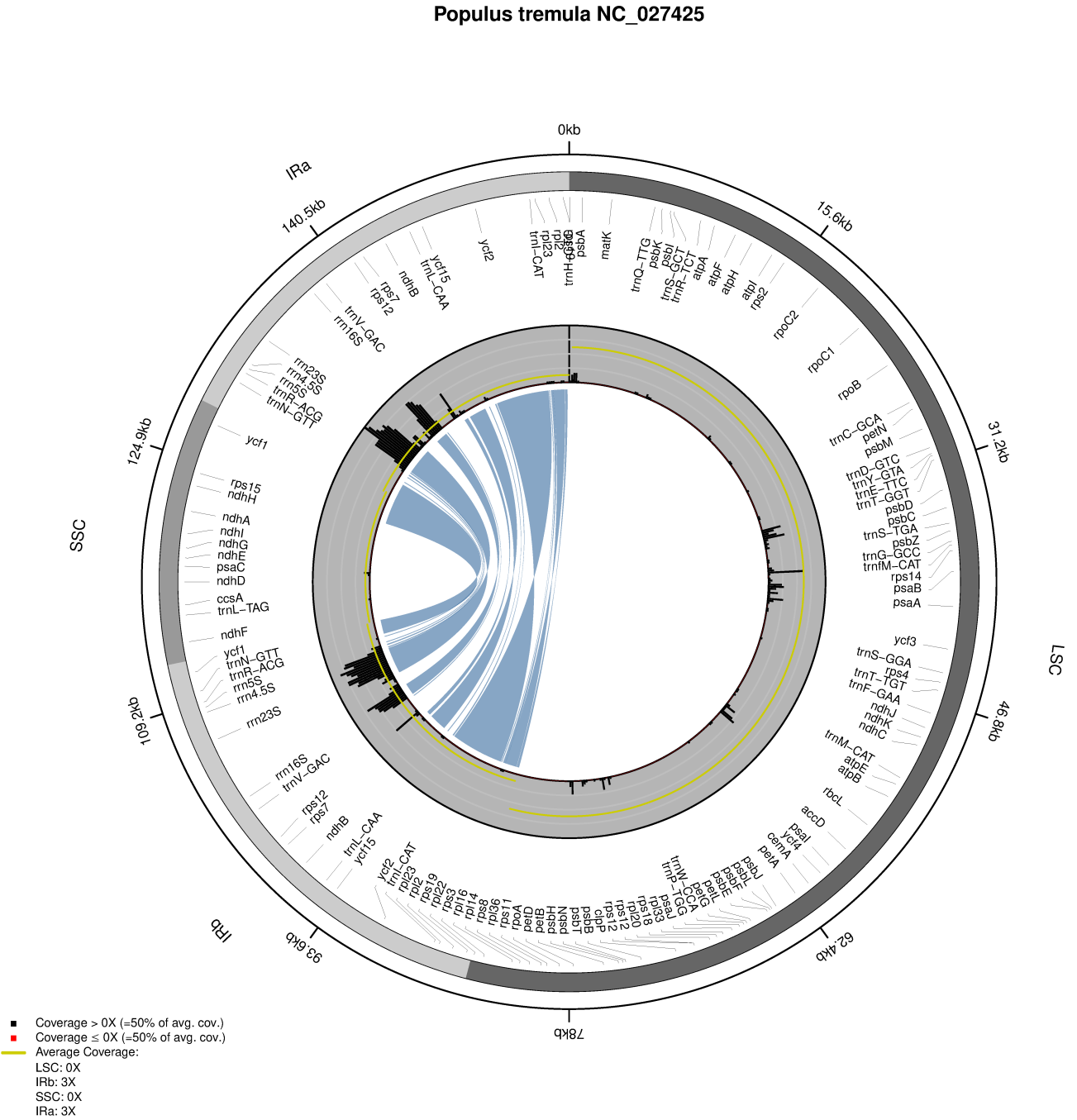


Figure S 5: Sequencing coverage of the plastid genome of *Populus tremula* (NC_027425) relative to its quadripartite structure and genes.

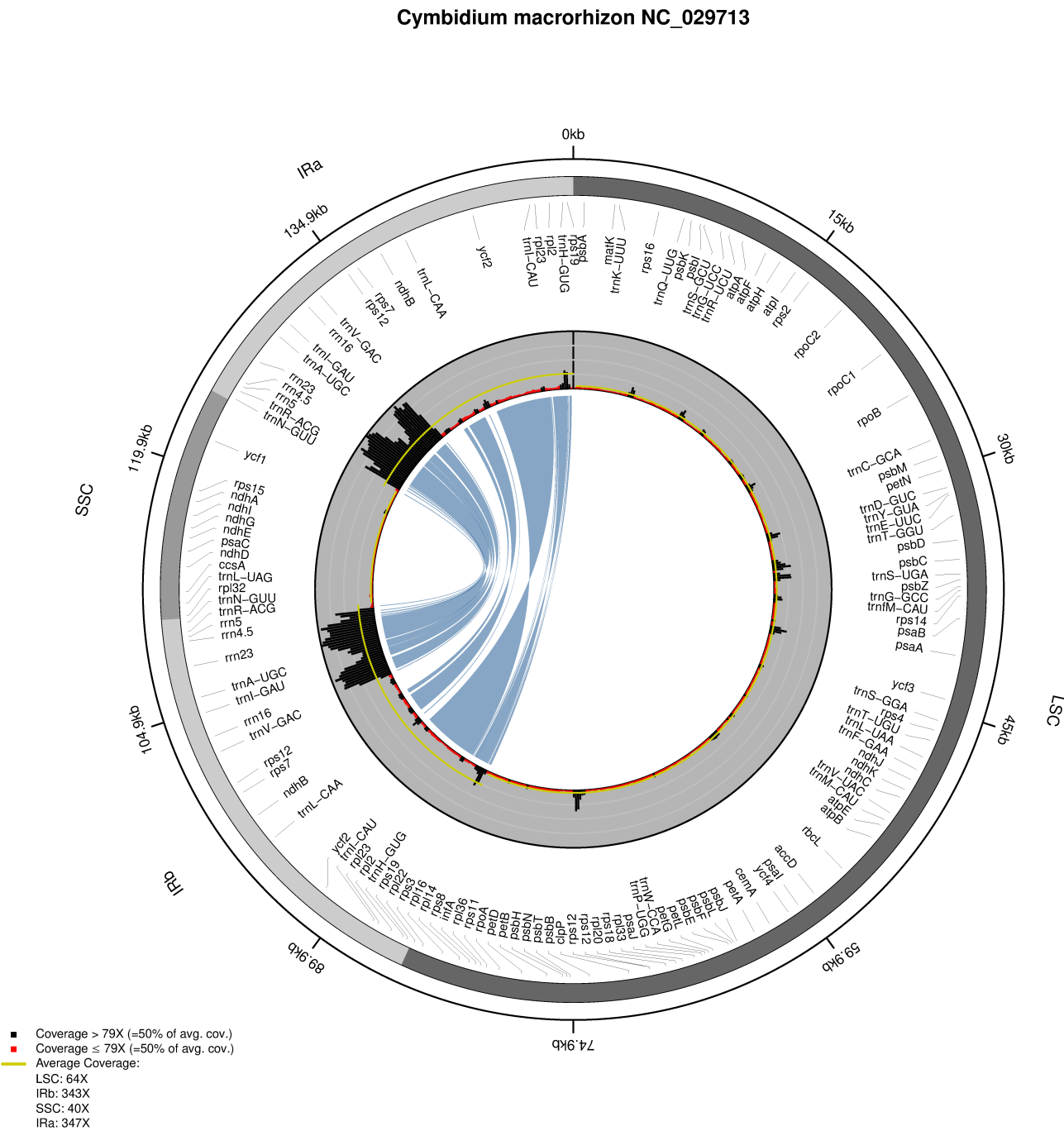
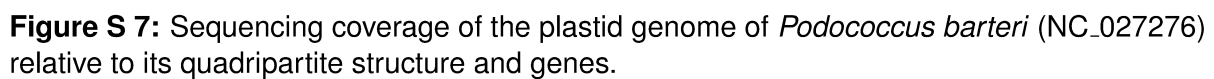
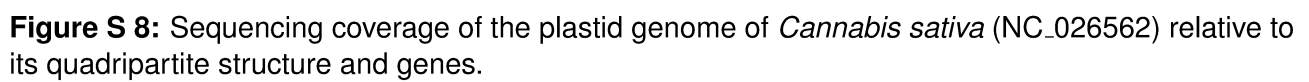


Figure S 6: Sequencing coverage of the plastid genome of *Cymbidium macrorhizon* (NC_029713) relative to its quadripartite structure and genes.





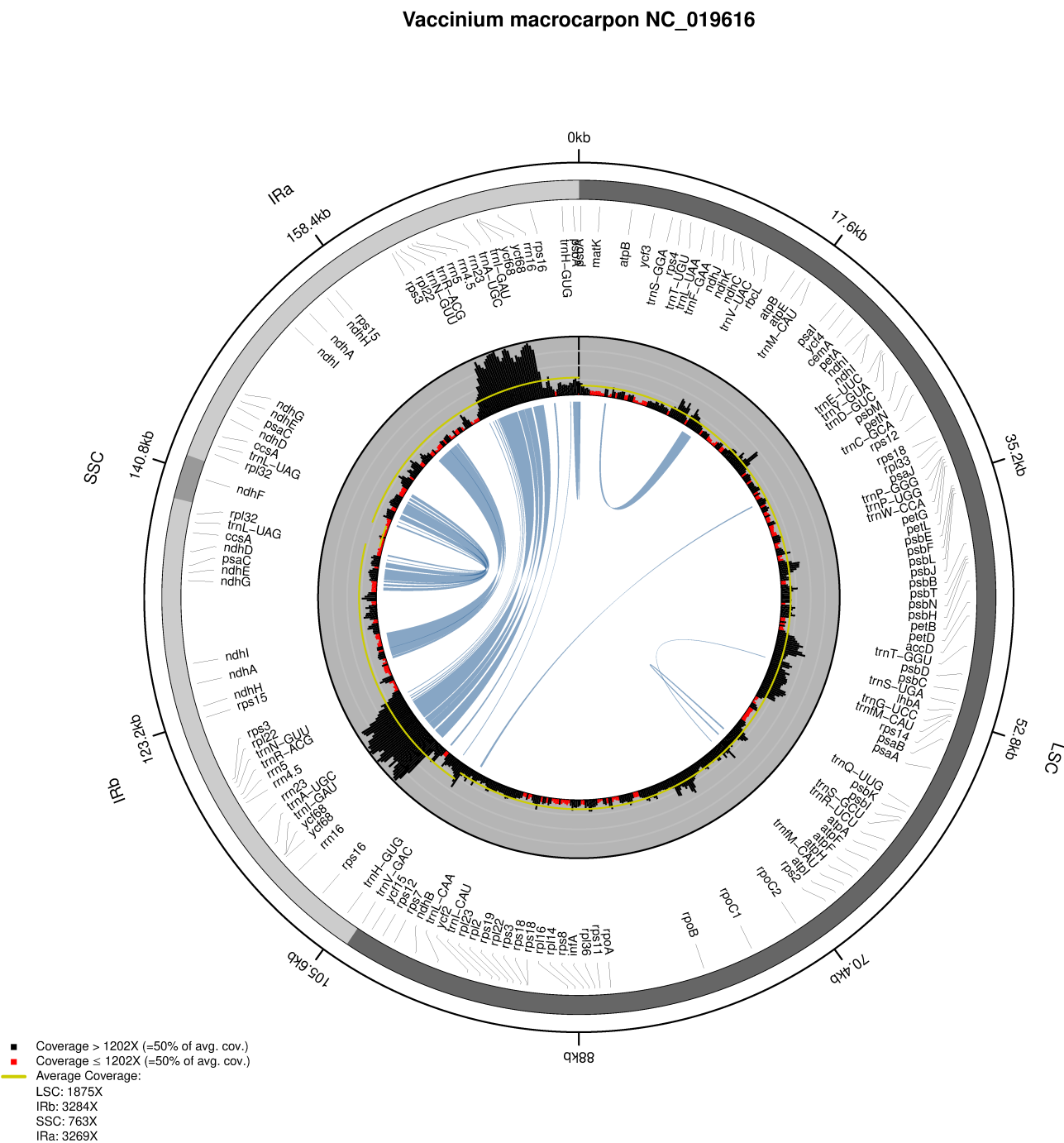
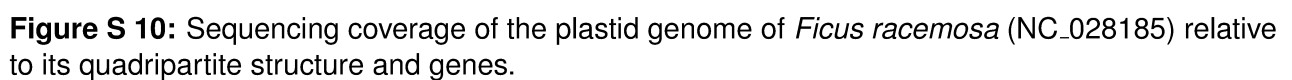


Figure S 9: Sequencing coverage of the plastid genome of *Vaccinium macrocarpon* (NC_019616) relative to its quadripartite structure and genes.



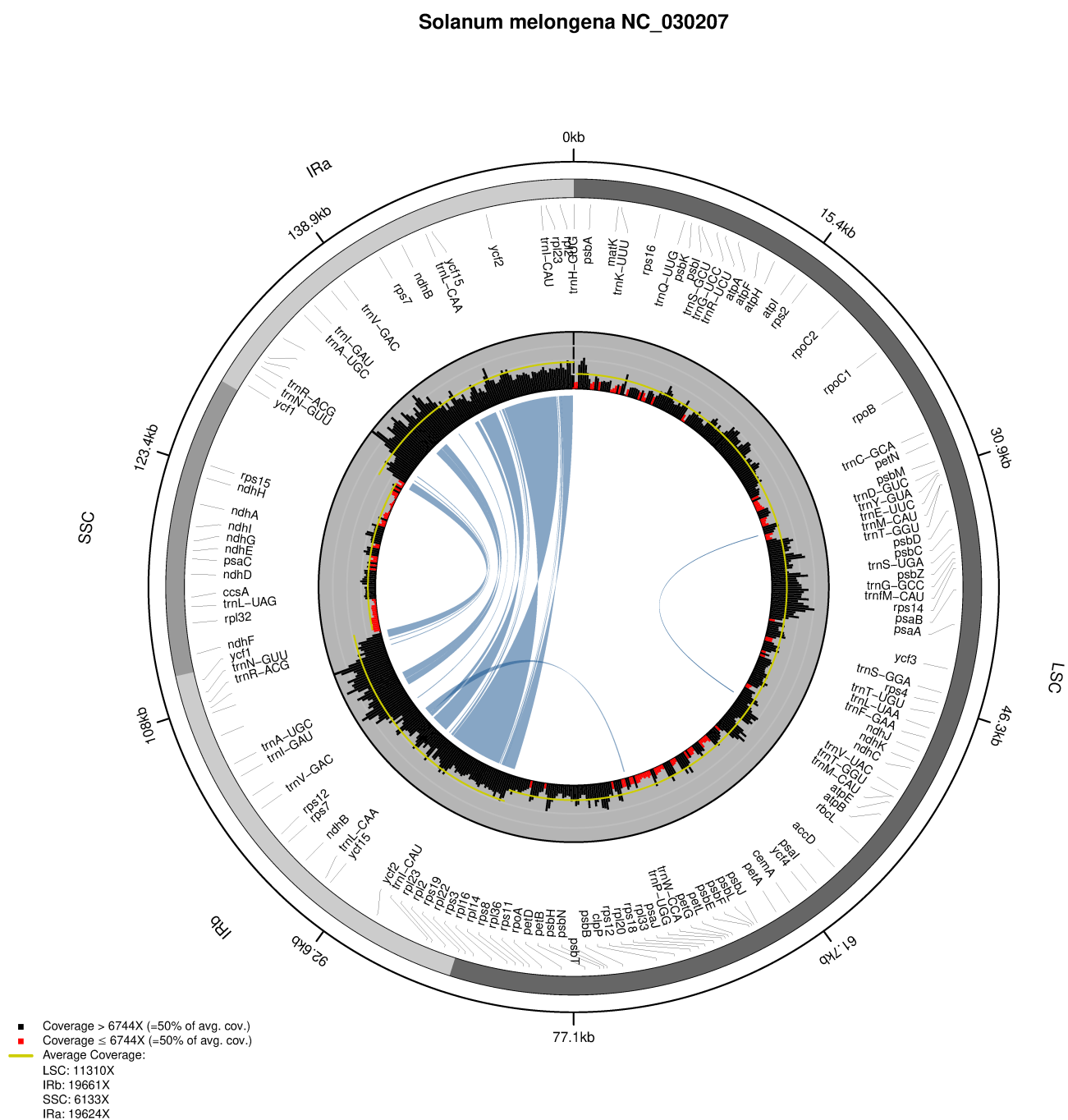


Figure S 11: Sequencing coverage of the plastid genome of *Solanum melongena* (NC_030207) relative to its quadripartite structure and genes.

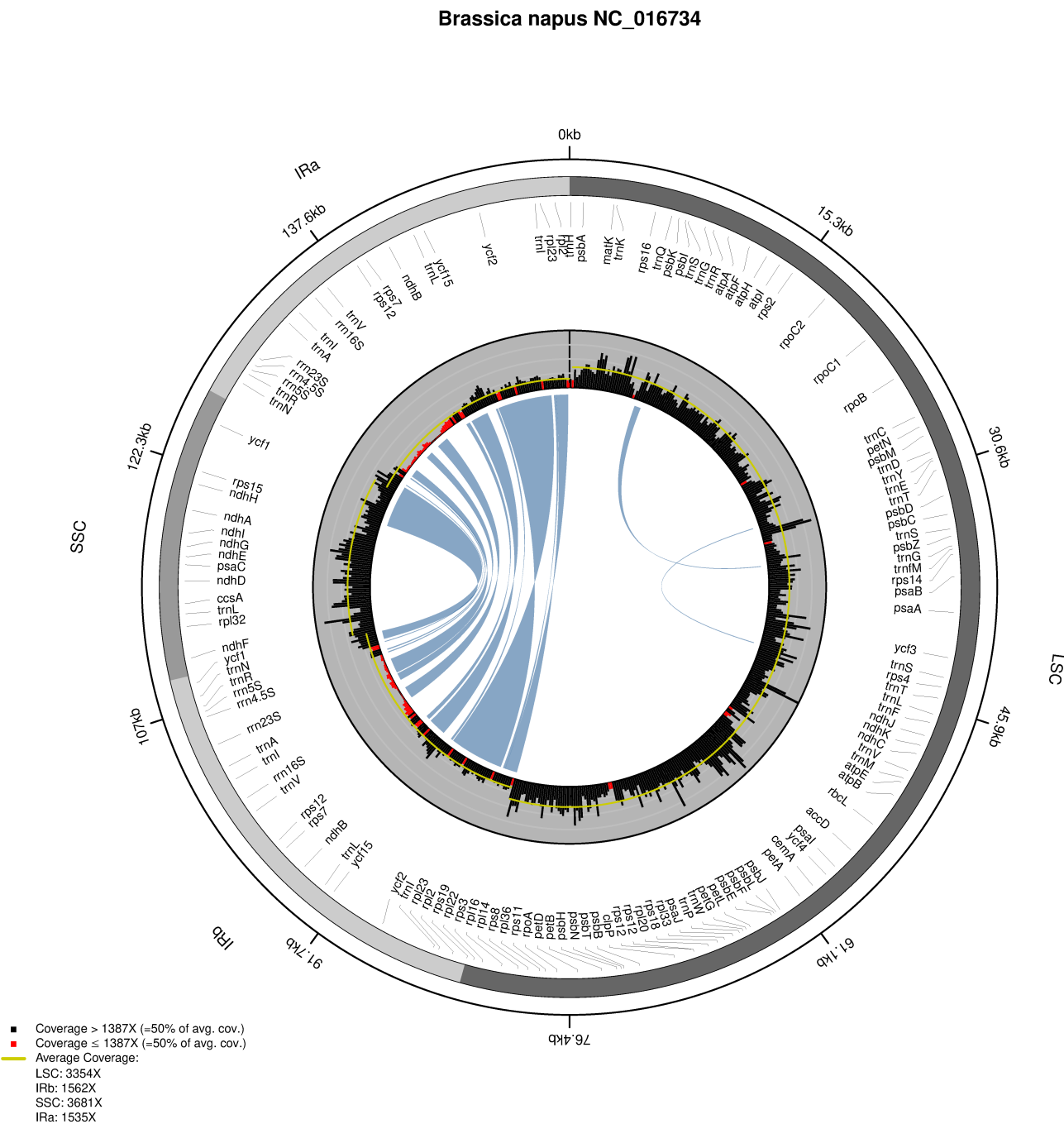


Figure S 12: Sequencing coverage of the plastid genome of *Brassica napus* (NC_016734) relative to its quadripartite structure and genes.

Fragaria vesca subsp. *bracteata* NC_018766

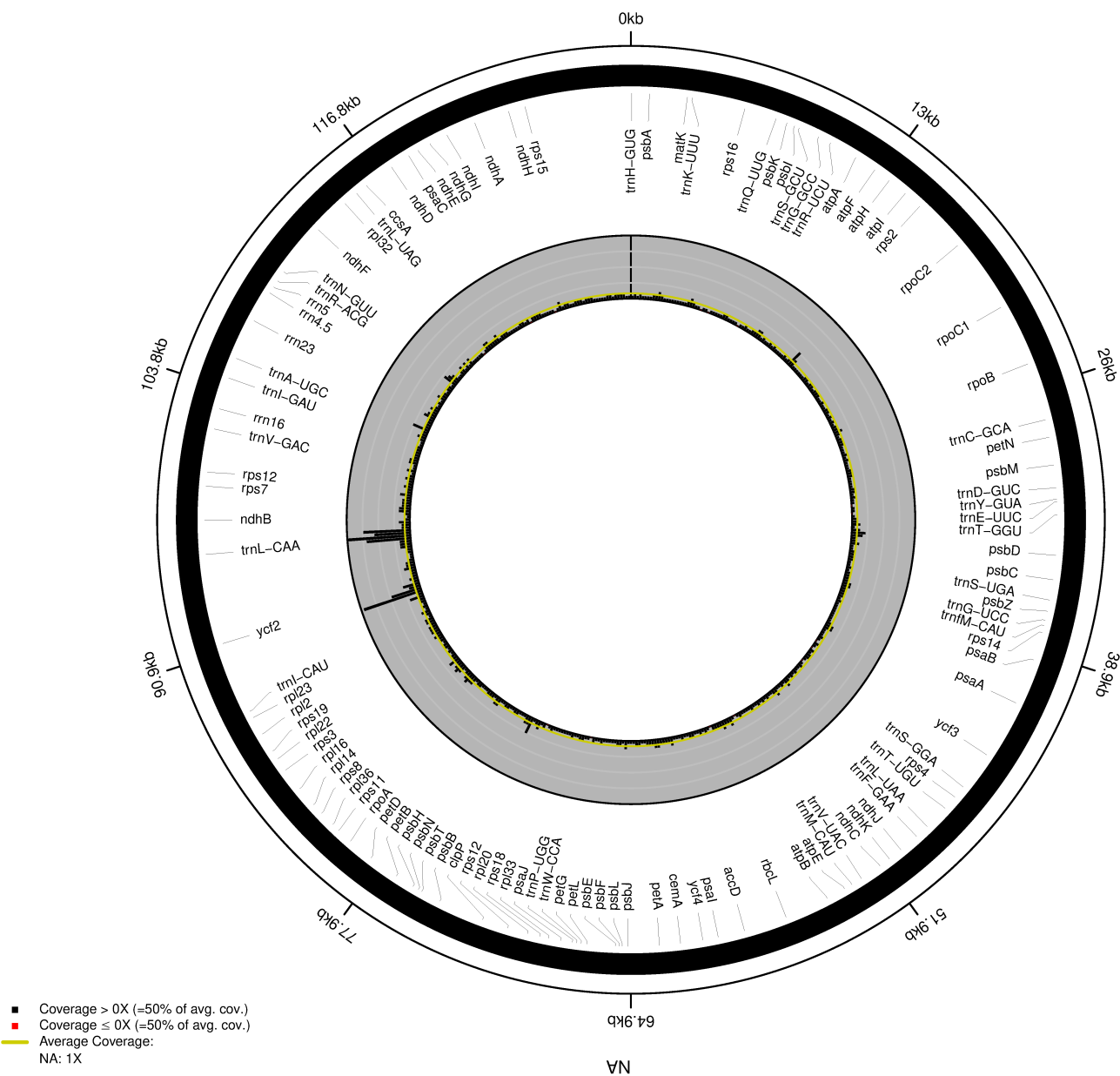


Figure S 13: Sequencing coverage of the plastid genome of *Fragaria vesca* subsp. *bracteata* (NC_018766) relative to its genes; the structural quadripartition was not annotated in the original record.

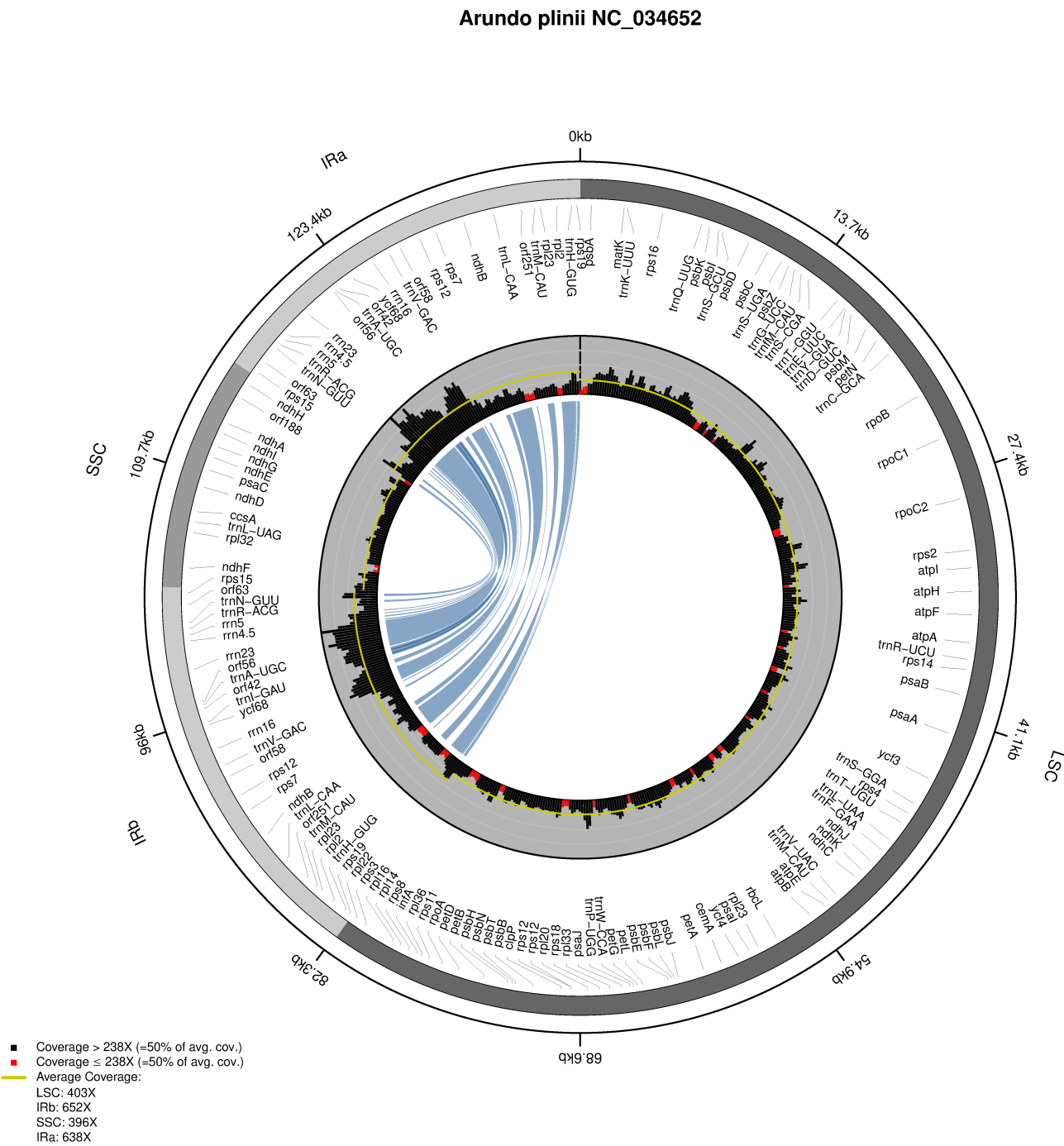


Figure S 14: Sequencing coverage of the plastid genome of *Arundo plinii* (NC_034652) relative to its quadripartite structure and genes.

