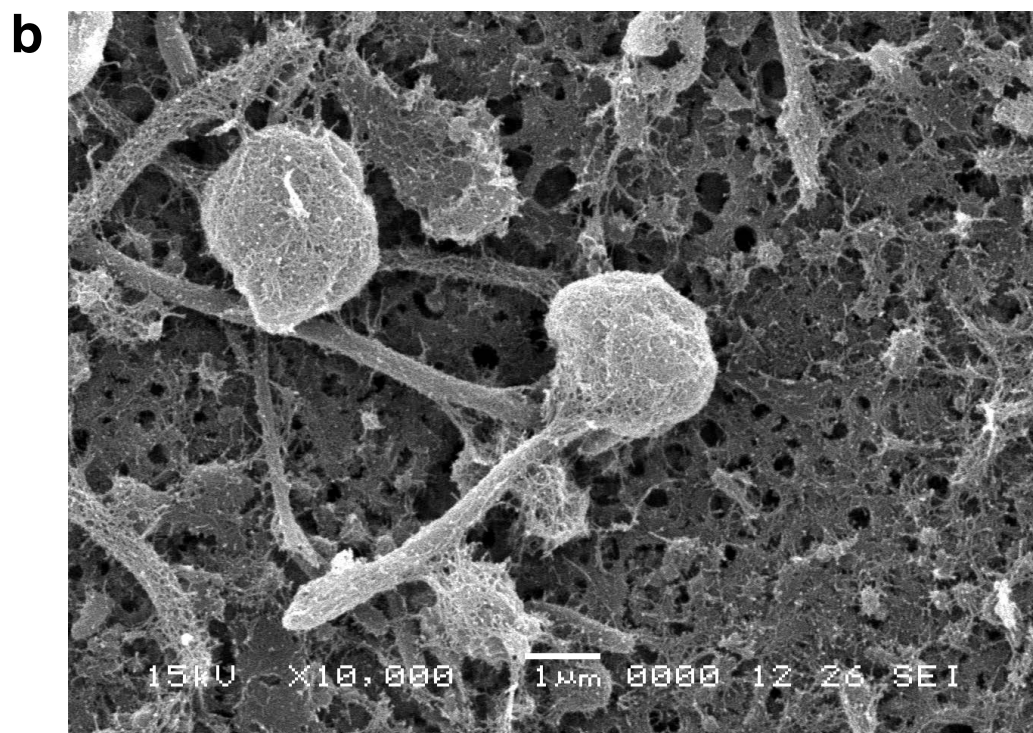
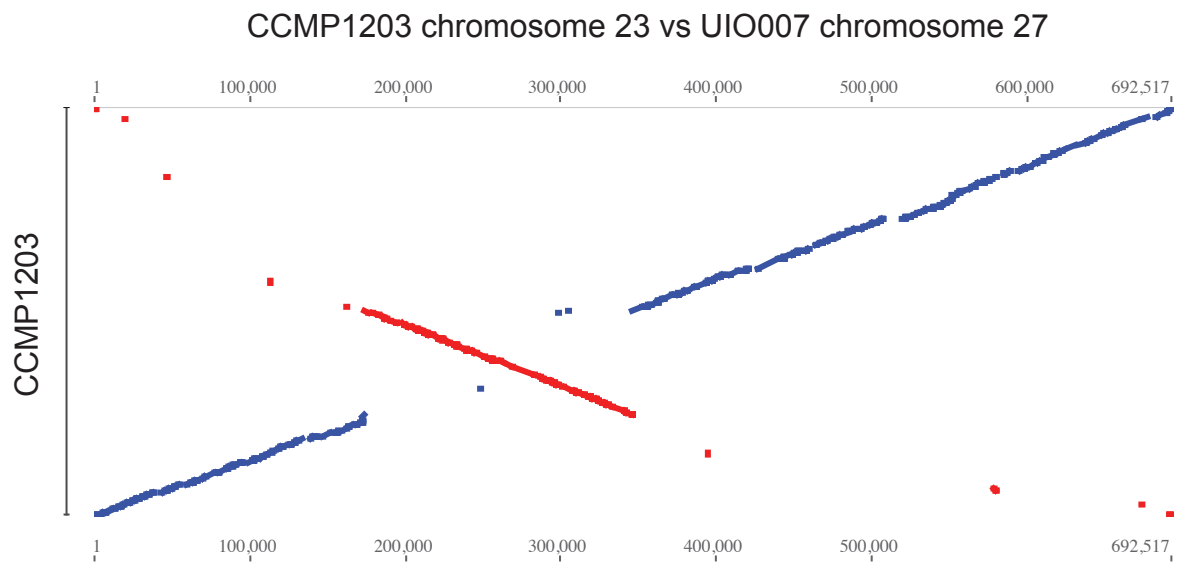
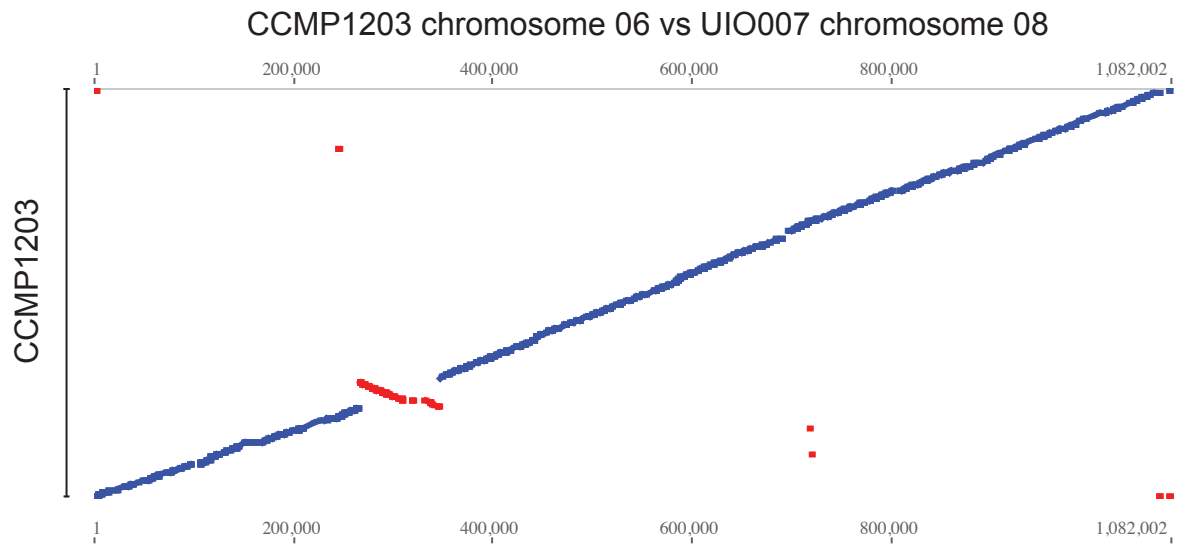
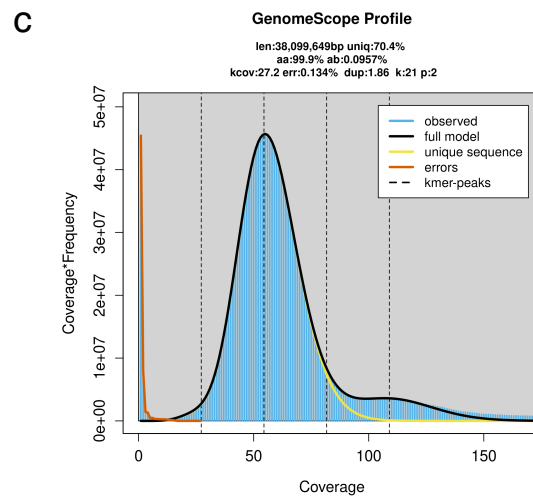
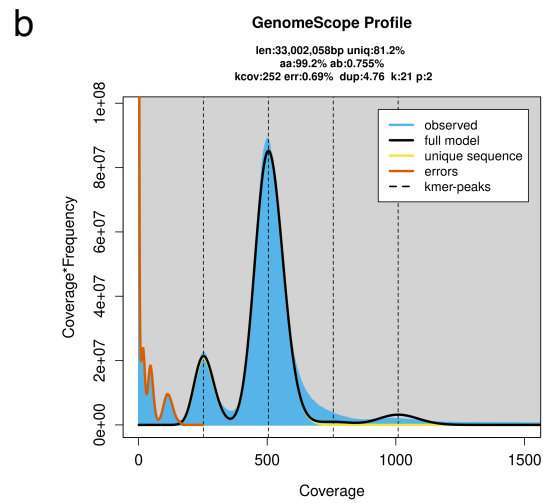
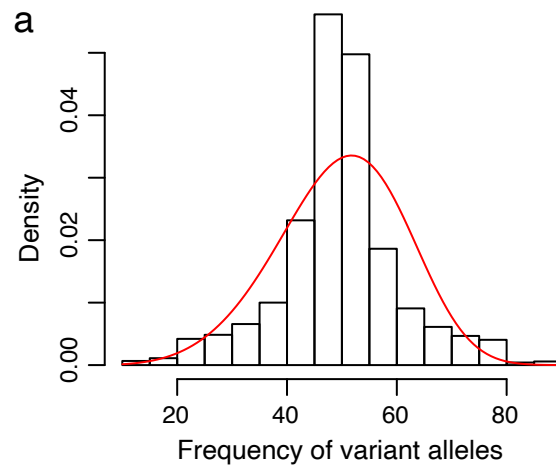




Supplementary Fig. 1. Scanning electron micrographs of *P. marina* strains CCMP1203 (**a**) and UIO007 (**b**). Scale bars = 1 μ m.

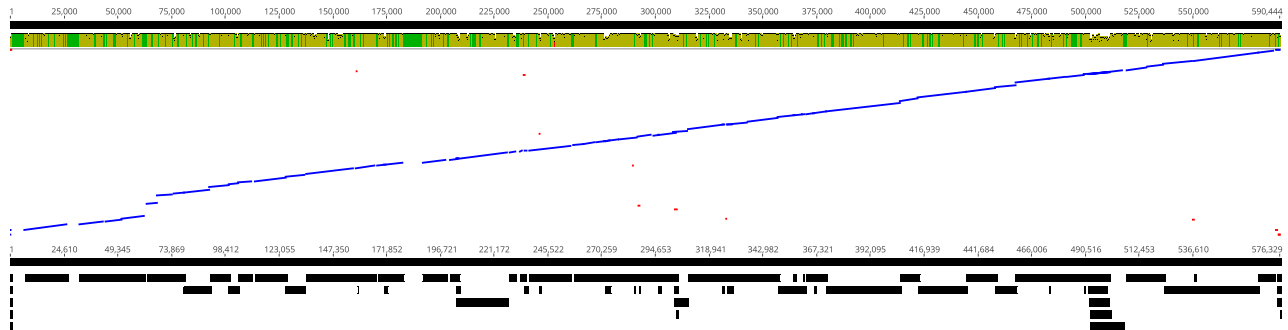


Supplementary Fig. 2. Alignments of two pairs of homologous *P. marina* UIO007 and CCMP1203 chromosomes that were rearranged by sequence inversions. These alignments were performed using the LASTZ plugin in Geneious Prime 2025.1.2 (Dotmatics, Boston, MA), with default options.

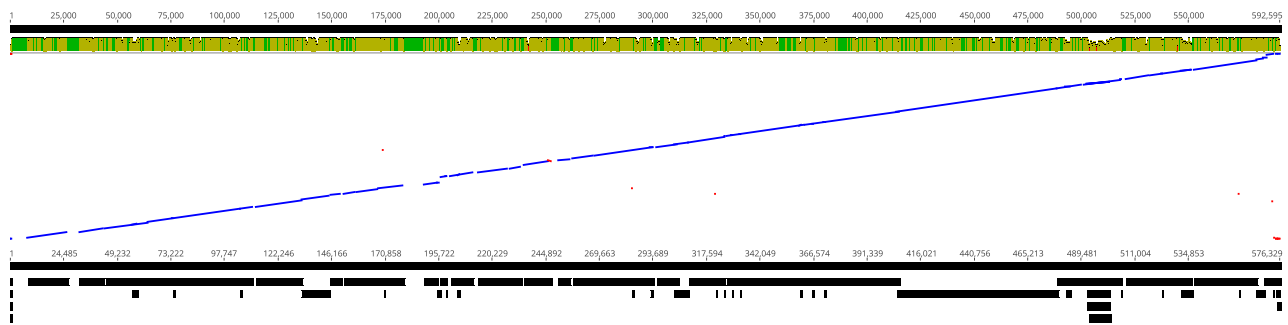


Supplementary Fig. 3. Ploidy analyses of *P. marina* genomes. **(a)** Allelic frequency distribution of SNPs in the *P. marina* CCMP1203 genome. **(b)** GenomeScope analysis of the k-mer coverage profile of the *P. marina* CCMP1203 genome. **(c)** GenomeScope analysis of the kmer coverage profile of the *P. marina* UIO007 genome.

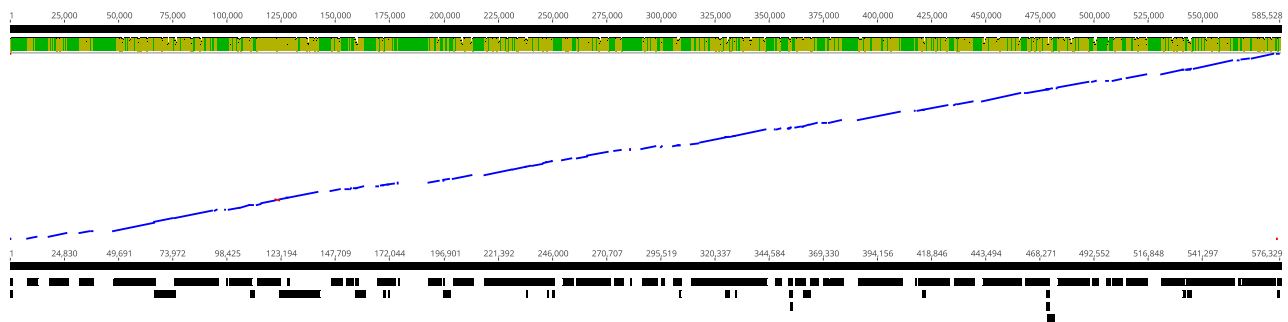
CCMP1203 (CP150576) vs UIO007 (CP150411)



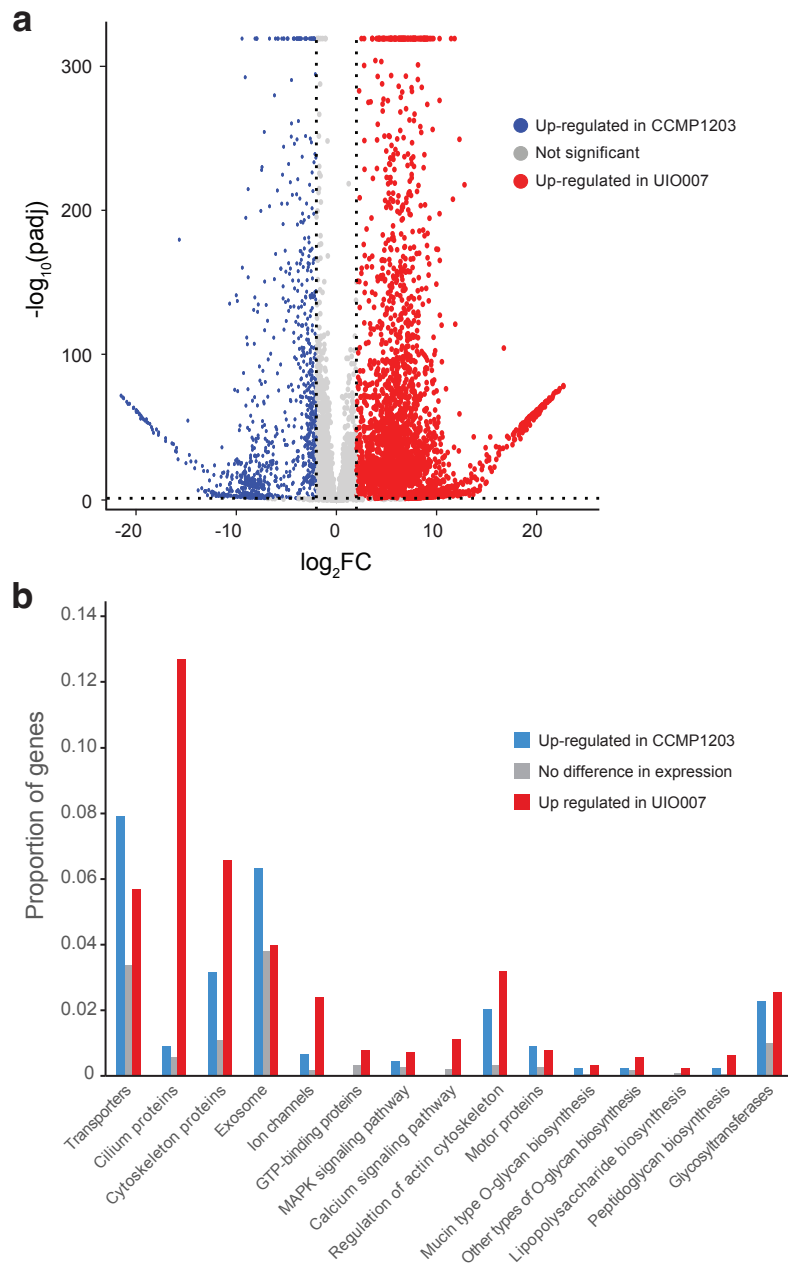
CCMP1203 (CP150576) vs CCAP 190/2 (OW568889)



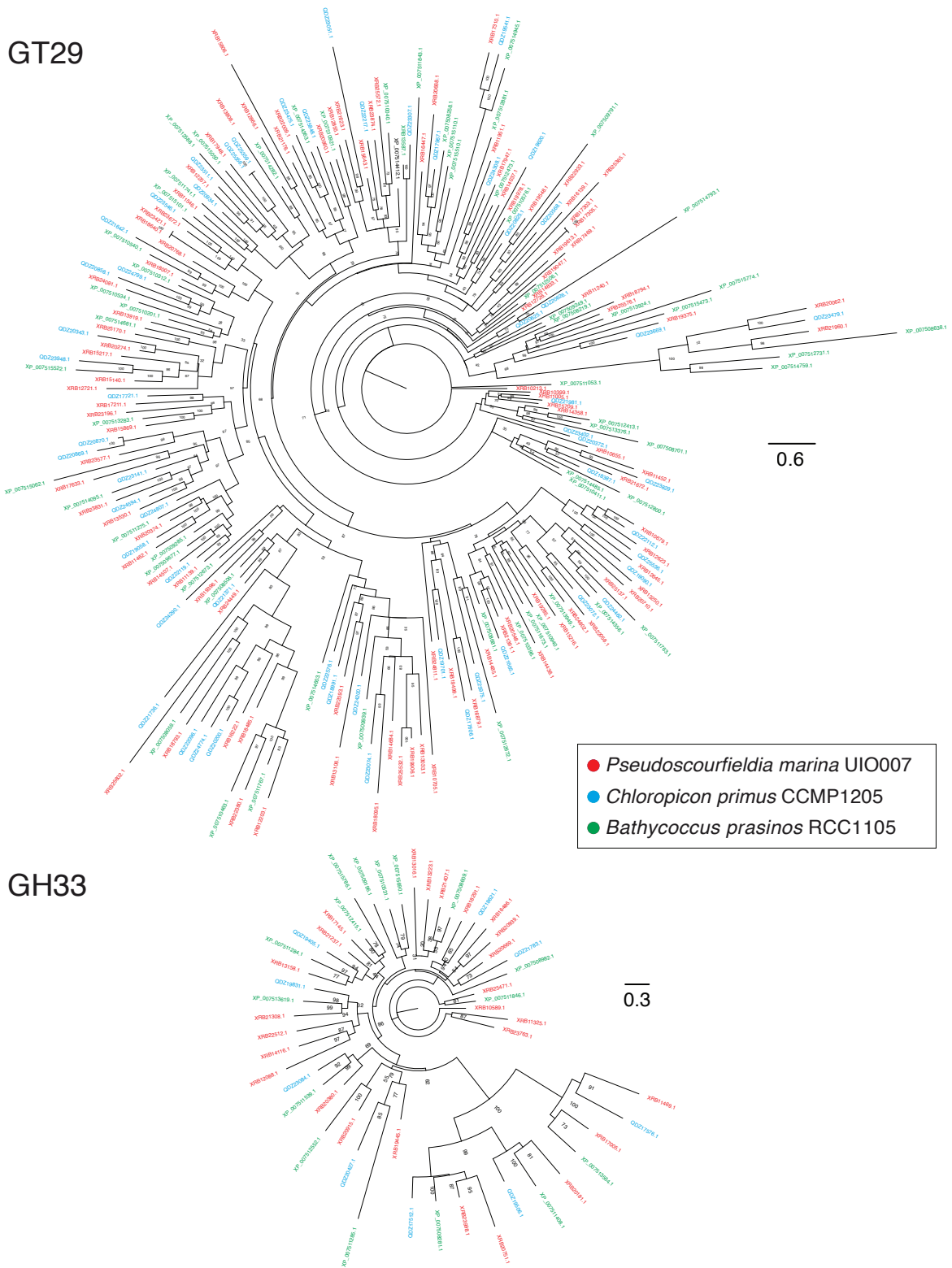
CCMP1203 (CP150576) vs NIES-2893 (BNJQ01000030)



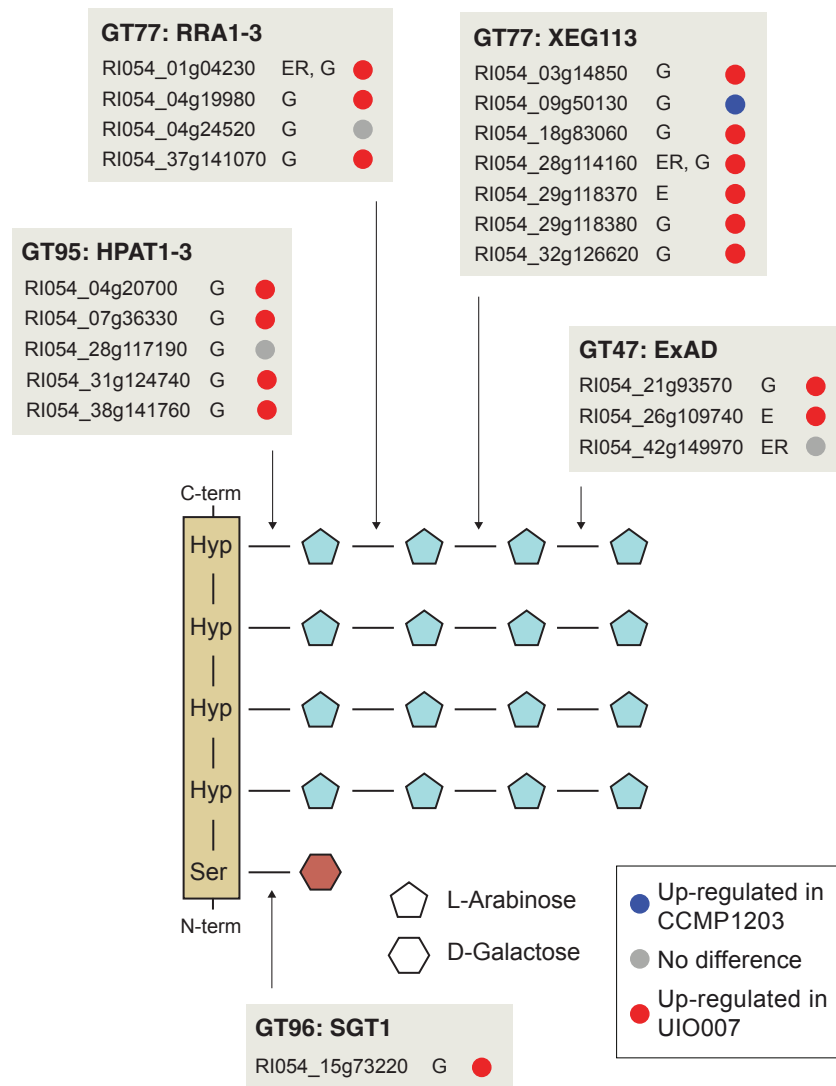
Supplementary Fig. 4. Alignments of three pairs of *P. marina* chromosomes homologous to CCMP1203 chromosome 31 (reference sequence). The accession numbers of the compared CCMP1203, UIO007, CCAP 190/2 and NIES-2893 chromosomes are indicated in parentheses. The alignments were performed using the LASTZ plugin in Geneious Prime 2025.1.2 (Dotmatics, Boston, MA), with default options.



Supplementary Fig. 5. (a) Volcano plot of differential gene expression between *P. marina* CCMP1203 and UIO007. Genes were classified as differentially expressed if they exhibited a log₂FC ≥ 2 or ≤ -2 and padj value ≤ 0.05 . **(b)** Proportion of genes from each KEGG functional category within three expression groups defined by log₂FC intervals (≥ 2.00 , ≤ -2.00 , and intermediate), calculated relative to the total number of genes in each group.

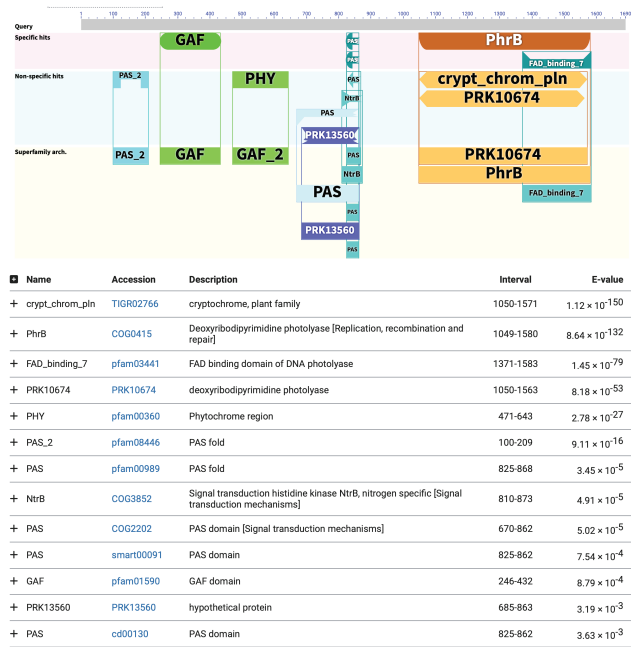


Supplementary Fig. 6. Phylogenetic relationships among the GT29 glycosyltransferases and GH33 glycosyl hydrolases identified in the genomes of three prasinophyte algae with expanded sets of these CAZymes.

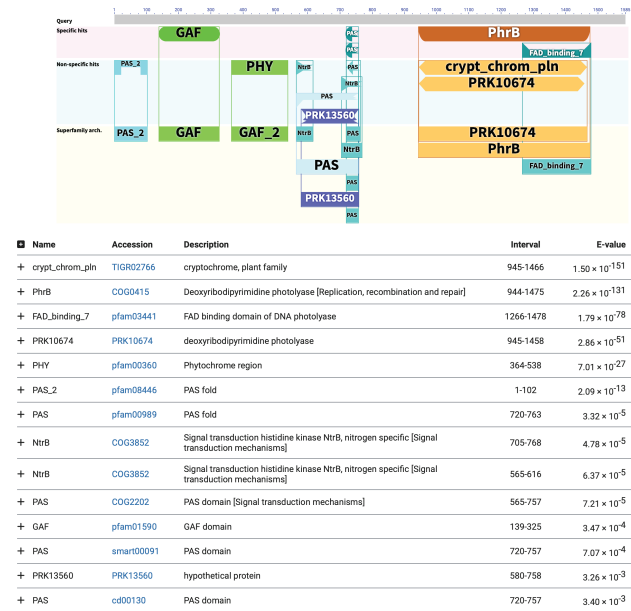


Supplementary Fig. 7. Sites of galactose and arabinose additions by glycosyltransferases from families GT47, GT77, GT95, and GT96 on the pentapeptide repeat of extensins. The corresponding *P. marina* UIO007 gene locus tags encoding these CAZymes are indicated. For each gene, the predicted subcellular localization of the encoded enzyme is shown: E, extracellular; ER, endoplasmic reticulum; G, Golgi apparatus. Colored circles represent differential gene expression levels between *P. marina* UIO007 and CCMP1203.

NIES-2893
GHP08179



CCMP1203
XRB04897



UIO007
XRB20694



Supplementary Fig. 8. Compared size and domain composition of the proteins encoded by the DUC1 locus of the *P. marina* NIES-2893, CCMP1203, and UIO007 strains. These analyses were performed on the NCBI BlastP webserver.

Supplementary Table 1. Size, G+C content and Telomere to Telomere (T2T) status of homologous *P. marina* CCMP1203 and UIO007 chromosomes.

CCMP1203				UIO007			
Chromosome	Chromosome size (bp)	T2T status	G+C (%)	Chromosome	Chromosome size (bp)	T2T status	G+C (%)
1	1,402,211	complete	57.37	2	1,556,407	complete	56.77
2	1,369,921	complete	57.84	1	1,566,147	complete	57.37
3	1,272,376	complete	57.88	3	1,344,335	complete	57.81
4	1,191,027	complete	57.75	4	1,239,368	complete	57.19
5	1,143,350	complete	57.80	5	1,238,128	5' telomere missing	57.51
6	1,042,048	complete	57.69	8	1,082,002	complete	57.53
7	1,002,356	complete	57.31	7	1,098,183	complete	56.85
8	973,802	complete	57.36	9	1,034,441	complete	56.98
9	946,859	complete	58.00	6	1,105,354	complete	57.35
10	872,757	complete	57.30	10	1,019,862	complete	57.01
11	848,505	complete	57.32	12	888,493	3' telomere missing	56.91
12	833,413	complete	57.33	11	996,019	complete	57.08
13	797,734	complete	57.65	14	857,328	complete	56.97
14	727,747	complete	57.57	13	878,076	complete	57.35
15	725,789	complete	57.46	20	767,324	complete	57.15
16	715,766	complete	57.86	19	781,729	complete	57.49
17	714,248	complete	57.54	17	819,774	complete	56.64
18	704,394	complete	57.42	18	794,485	complete	56.91
19	698,004	complete	57.65	23	743,287	complete	57.33
20	686,436	complete	57.26	15	824,404	complete	56.79
21	684,535	complete	57.56	21	766,343	complete	57.35
22	678,412	complete	57.12	25	721,224	complete	56.70
23	642,279	complete	57.35	27	692,517	complete	57.04
24	621,204	complete	57.18	24	742,331	complete	56.72
25	619,061	complete	58.12	22	760,274	5' telomere missing	57.23
26	613,715	complete	57.30	31	664,872	complete	57.09
27	605,788	complete	58.80	30	666,418	complete	58.51
28	593,600	complete	57.52	29	680,506	complete	56.88
29	590,613	complete	57.15	32	646,615	complete	56.81
30	590,029	complete	57.42	28	690,700	complete	56.69
31	576,329	complete	57.42	26	704,272	complete	57.05
32*	565,097	complete	49.20	16*	820,622	complete	49.77
33	528,593	complete	56.89	34	595,476	complete	56.64
34	527,594	complete	57.48	33	612,006	complete	56.59
35	522,243	complete	57.51	36	563,668	complete	56.95
36	515,615	complete	57.57	35	592,680	complete	56.71
37	489,857	complete	57.55	37	539,112	complete	57.02
38	480,670	complete	57.08	39	537,455	5' telomere missing	56.71
39	452,092	complete	57.07	38	538,673	complete	56.13
40	417,301	complete	57.20	40	484,222	complete	56.45
41	377,776	complete	57.54	41	416,919	complete	57.09
42	316,225	complete	56.76	45	298,190	5' telomere missing	56.37
43	314,157	complete	57.04	43	333,995	complete	56.62
44	285,835	complete	57.09	42	343,902	complete	56.48
45	242,694	complete	55.81	44	304,599	complete	54.93
Complete genome:		31,520,057		Complete genome:		35,352,737	

* Outlier chromosome.

Supplementary Table 2. Structural differences between the *P. marina* outlier chromosomes and their respective nuclear genomes.

	G+C content (%)	Gene density (no. /kb)	Repeats (%)
UIO007			
Nuclear genome	56.88	0.44	16.79
Chromosome 16	49.77	0.38	47.13
CCMP1203			
Nuclear genome	57.36	0.46	9.50
Chromosome 32	49.20	0.27	22.18

Supplementary Table 3. Annotations and expression patterns of genes encoding proteins with identifiable functions/domains on the *P. marina* UIO007 and CCMP1203 outlier chromosomes.

Gene product	Protein ID		% Identity	DeSeq2 analyses		UIO007 RPKM	CCMP1203 RPKM
	UIO007	CCMP1203		log2FC	padj		
Sugar metabolism/Transporters							
TDP-4-oxo-6-deoxy-D-glucose transaminase	RI054_16g76450	NFJ02_32g81890	92.59	21.38	0.000	5229	0
dTDP-4-amino-4,6-dideoxygalactose transaminase	RI054_16g76490	NFJ02_32g81510	93.26	22.67	0.000	22272	0
GDP-mannose/fucose transporter	RI054_16g77010	NFJ02_32g81230	73.14	20.30	0.000	5041	5
dTDP-4-amino-4,6-dideoxygalactose transaminase	RI054_16g77020	NFJ02_32g82200	80.75	20.27	0.000	3616	0
GDP-D-mannose dehydratase/GDP-fucose synthetase	RI054_16g77090	NFJ02_32g81330	91.24	14.38	0.000	11907	0
CMP-sialic acid transporter	RI054_16g74940	—	—	—	—	8929	—
dTDP-4-amino-4,6-dideoxygalactose transaminase	RI054_16g75020	—	—	—	—	12532	—
CMP-sialic acid transporter	RI054_16g75060	—	—	—	—	4928	—
peptidoglycan/xylan/chitin deacetylase	RI054_16g75090	—	—	—	—	8393	—
sugar O-acyltransferase	RI054_16g75250	—	—	—	—	7699	—
CMP-sialic acid transporter	RI054_16g75740	—	—	—	—	4291	—
GDP-mannose/fucose transporter	RI054_16g75850	—	—	—	—	3577	—
CMP-sialic acid transporter	RI054_16g75940	—	—	—	—	5810	—
O-fucosyltransferase family protein	RI054_16g75960	—	—	—	—	4351	—
sugar O-acyltransferase	RI054_16g76270	—	—	—	—	19886	—
UDP-rhamnose/UDP-galactose transporter	RI054_16g76520	—	—	—	—	8479	—
dTDP-4-dehydroharmnose reductase	RI054_16g77080	—	—	—	—	8316	—
UDP-sugar transporter	RI054_16g77240	—	—	—	—	5878	—
CMP-sialic acid transporter	RI054_16g77300	—	—	—	—	6247	—
dTDP-4-amino-4,6-dideoxygalactose transaminase	—	NFJ02_32g81150	—	—	—	—	0
dTDP-4-amino-4,6-dideoxygalactose transaminase	—	NFJ02_32g81240	—	—	—	—	0
dTDP-4-amino-4,6-dideoxygalactose transaminase	—	NFJ02_32g81340	—	—	—	—	0
trifunctional UDP-glucose 4,6-dehydratase/UDP-4-keto-6-deoxy-D-glucose 3,5-epimerase/UDP-4-keto-L-rhamnose-reductase RHM1	—	NFJ02_32g81460	—	—	—	—	9
bifunctional dTDP-4-dehydroharmnose 3,5-epimerase/dTDP-4-dehydroharmnose reductase	—	NFJ02_32g81550	—	—	—	—	0
nucleoside-diphosphate-sugar epimerase WcaG	—	NFJ02_32g81650	—	—	—	—	0
dTDP-4-amino-4,6-dideoxygalactose transaminase	—	NFJ02_32g81680	—	—	—	—	0
GDP-D-mannose dehydratase/GDP-fucose synthetase	—	NFJ02_32g81790	—	—	—	—	0
nucleoside-diphosphate-sugar epimerase WcaG	—	NFJ02_32g81840	—	—	—	—	0
TDP-4-oxo-6-deoxy-D-glucose transaminase	—	NFJ02_32g81880	—	—	—	—	0
dTDP-4-amino-4,6-dideoxygalactose transaminase	—	NFJ02_32g81900	—	—	—	—	3
TDP-4-oxo-6-deoxy-D-glucose transaminase	—	NFJ02_32g81940	—	—	—	—	0
UDP-glucose 6-dehydrogenase	—	NFJ02_32g82020	—	—	—	—	0
dTDP-4-amino-4,6-dideoxygalactose transaminase/GDP-fucose synthetase	—	NFJ02_32g82230	—	—	—	—	0
Glycosyltransferases							
glycosyltransferase family 32 protein	RI054_16g74900	NFJ02_32g82290	99.17	1.14	0.000	2429	4
O-fucosyltransferase family protein	RI054_16g75170	NFJ02_32g81490	94.82	12.74	0.000	3310	0
glycosyltransferase family 31 protein	RI054_16g75640	NFJ02_32g81780	99.56	19.35	0.000	3573	15
glycosyltransferase family 61 protein	RI054_16g76430	NFJ02_32g81390	86.32	10.64	0.001	6744	0
glycosyltransferase family 10 protein	RI054_16g76900	NFJ02_32g82510	65.25	20.57	0.000	2184	0
glycosyltransferase family 90 protein	RI054_16g77280	NFJ02_32g81200	87.78	21.17	0.000	7877	0
glycosyltransferase family 47 protein	RI054_16g74910	—	—	—	—	1988	—
glycosyltransferase family 61 protein	RI054_16g75080	—	—	—	—	2836	—
glycosyltransferase family 8 protein	RI054_16g75310	—	—	—	—	3515	—
glycosyltransferase family 47 protein	RI054_16g75340	—	—	—	—	4261	—
glycosyltransferase family 29 protein	RI054_16g75380	—	—	—	—	7162	—
glycosyltransferase family 29 protein	RI054_16g75560	—	—	—	—	4380	—
glycosyltransferase family 90 protein	RI054_16g75760	—	—	—	—	5686	—
glycosyltransferase family 15 protein	RI054_16g76090	—	—	—	—	9851	—
glycosyltransferase family 8 protein	RI054_16g76280	—	—	—	—	8195	—
glycosyltransferase family 10 protein	RI054_16g76580	—	—	—	—	6746	—
fucosyltransferase	RI054_16g76870	—	—	—	—	1422	—
O-fucosyltransferase family protein	RI054_16g77110	—	—	—	—	10522	—
glycosyltransferase family 61 protein	RI054_16g77200	—	—	—	—	5302	—
glycosyltransferase family 43 protein	RI054_16g77500	—	—	—	—	4159	—
glycosyltransferase family 23 protein	RI054_16g77580	—	—	—	—	4088	—
glycosyltransferase family 23 protein	RI054_16g77590	—	—	—	—	2682	—
glycosyltransferase family 11 protein	RI054_16g77670	—	—	—	—	2024	—
glycosyltransferase family 69 protein	—	NFJ02_32g81360	—	—	—	—	0
Methyltransferases							
class I SAM-dependent methyltransferase	RI054_16g75180	NFJ02_32g82320	98.21	2.62	0.000	5954	68
class I SAM-dependent methyltransferase	RI054_16g75660	NFJ02_32g82170	98.19	10.90	0.000	2931	1
methyltransferase, FkbM family	RI054_16g75470	—	—	—	—	6727	—
methyltransferase, FkbM family	RI054_16g75590	—	—	—	—	539	—
SAM-dependent methyltransferase	RI054_16g75900	—	—	—	—	6124	—
methyltransferase, FkbM family	RI054_16g76080	—	—	—	—	7014	—
methyltransferase, FkbM family	RI054_16g76420	—	—	—	—	749	—
class I SAM-dependent methyltransferase	RI054_16g76560	—	—	—	—	8472	—
methyltransferase, FkbM family	RI054_16g76620	—	—	—	—	4887	—
class I SAM-dependent methyltransferase	RI054_16g76710	—	—	—	—	18587	—
class I SAM-dependent methyltransferase	RI054_16g76820	—	—	—	—	14640	—
methyltransferase, FkbM family	RI054_16g76840	—	—	—	—	1525	—
class I SAM-dependent methyltransferase	RI054_16g77380	—	—	—	—	12250	—
SAM-dependent methyltransferase	RI054_16g77520	—	—	—	—	2235	—
methyltransferase, FkbM family	—	NFJ02_32g81530	—	—	—	—	0
class I SAM-dependent methyltransferase	—	NFJ02_32g81760	—	—	—	—	1
methyltransferase, FkbM family	—	NFJ02_32g82250	—	—	—	—	0
methyltransferase, FkbM family	—	NFJ02_32g82270	—	—	—	—	0
class I SAM-dependent methyltransferase	—	NFJ02_32g82540	—	—	—	—	0
class I SAM-dependent methyltransferase	—	NFJ02_32g82610	—	—	—	—	4
Other functions							
protease	RI054_16g75460	NFJ02_32g82590	93.81	6.98	0.000	2559	13
protease	RI054_16g76040	NFJ02_32g82590	97.98	6.65	0.000	1921	13
DUF616 domain-containing protein	RI054_16g76740	NFJ02_32g81220	68.30	19.37	0.000	2786	0
FAD dependent oxidoreductase domain-containing protein	RI054_16g75980	—	—	—	—	6975	—
DUF616 domain-containing protein	RI054_16g75990	—	—	—	—	6970	—
3-oxoacyl-[acyl-carrier-protein] synthase III	RI054_16g76260	—	—	—	—	1197	—
protease	RI054_16g76720	—	—	—	—	543	—
sulfotransferase family protein	RI054_16g76850	—	—	—	—	348	—
sulfotransferase family protein	RI054_16g77040	—	—	—	—	9668	—
DUF616 domain-containing protein	RI054_16g77250	—	—	—	—	12211	—
DNA polymerase A	RI054_16g77340	—	—	—	—	7767	—
Cu2+-containing amine oxidase	RI054_16g77360	—	—	—	—	5205	—
histone deacetylase	RI054_16g77700	—	—	—	—	1053	—
acetyl/propionyl-CoA carboxylase, alpha subunit	RI054_16g77790	—	—	—	—	72	—
acetohydroxyacid synthase	—	NFJ02_32g81130	—	—	—	—	0
acetaldehyde dehydrogenase	—	NFJ02_32g81160	—	—	—	—	3
acetylacetyl synthase large subunit	—	NFJ02_32g81170	—	—	—	—	3
4-hydroxy 2-oxovalerate aldolase	—	NFJ02_32g81180	—	—	—	—	0
deoxyuridine 5'-triphosphate nucleotidohydrolase	—	NFJ02_32g81280	—	—	—	—	0
PIN domain-containing protein	—	NFJ02_32g81430	—	—	—	—	1338
zinc transporter ZupT	—	NFJ02_32g82100	—	—	—	—	24
DUF268 domain-containing protein	—	NFJ02_32g82350	—	—	—	—	0
acetohydroxyacid synthase	—	NFJ02_32g82400	—	—	—	—	0

Supplementary Table 4. Annotations and expression patterns of *P. marina* UIO007 and CCMP1203 genes encoding flagellar proteins.

ChlamyFPv5 Annotation	Protein ID			DeSeq2 analyses			Chlamydomonas ortholog
	KO identifier	UIO007	CCMP1203	% identity	log2FC	padj	
Microtubules: Tubulin							
TUBA1, tubulin alpha	K07374	R054_12g161480	NF02_11g062080	100.00	10.47	0.000	Cro31.919050.11.0
TUBA1, tubulin alpha	K07374	R054_12g162340	NF02_11g170600	100.00	10.00	0.000	Cro31.919050.11.0
TUBA1, tubulin alpha	K07374	R054_38g138570	NF02_35g09770	100.00	9.64	0.000	Cro31.919050.11.0
TUBB2, tubulin beta	K07375	R054_24g102540	NF02_24g45780	99.77	11.51	0.000	Cro31.919050.11.0
TUBB2, tubulin beta	K07375	R054_28g110770	NF02_30g78450	100.00	-7.19	0.000	Cro12.64550.11.0
TUBD1 (UH3), Tubulin delta 1	K10390	R054_41g147910	NF02_41g107010	99.03	-0.10	0.623	Cro31.878750.11.0
TUBE1 (BLD2), Tubulin epsilon 1	K10391	R054_07g07020	NF02_07g131060	100.00	-0.42	0.058	Cro31.72810.11.0
TUBG1, Tubulin gamma 1	K10389	R054_28g116370	NF02_36g77790	99.43	-0.84	0.000	Cro6.299300.11.0
Microtubules: Tubulin polyglutamylase and associated proteins							
FAP234 (TPG2), Flagellar Associated Protein 234	—	R054_31g123490	NF02_26g06020	99.49	5.35	0.000	Cro11.923550.11.1
LRRCA49, leucine-rich repeat-containing protein 49	K18606	R054_03g19550	NF02_03g105270	96.57	9.18	0.000	—
PGS22, tubulin polyglutamylase complex subunit 2	K18605	R054_04g12120	NF02_04g115290	99.66	10.27	0.000	—
TLL1, tubulin polyglutamylase	K18599	R054_08g5210	NF02_08g120690	99.59	4.97	0.000	—
TLL11, tubulin polyglutamylase	K18604	R054_25g107850	NF02_22g06010	99.65	3.42	0.000	—
TLL3, Tubulin Monoglycylase	K18608	R054_05g09920	NF02_05g123540	100.00	7.02	0.037	Cro31.919050.11.0
TLL3, Tubulin Monoglycylase	K18608	R054_29g110290	NF02_29g06900	99.25	18.95	0.000	Cro31.919050.11.0
TLL3, Tubulin Monoglycylase	K18608	R054_30g120090	NF02_27g06280	96.95	4.98	0.000	Cro31.919050.11.0
TLL4, tubulin polyglutamylase	K18601	R054_04g147470	NF02_40g106620	95.45	8.11	0.014	—
TLL4, tubulin polyglutamylase	K18601	R054_04g147470	NF02_40g106620	95.45	8.11	0.014	—
TLL6, tubulin polyglutamylase	K18582	R054_01g12360	NF02_02g17160	99.61	3.21	0.000	—
TLL8, tubulin polyglutamylase	K18582	R054_03g16120	NF02_03g10140	99.58	20.45	0.000	—
TLL9 (FAP267) (TPG1), Tubulin-Tyrosine Ligase 9	—	R054_10g14560	NF02_10g13690	100.00	7.94	0.000	Cro17.899950.11.2
Axoneme: Outer Dynein Arms							
DHC13 (ODA-DHC1), Outer Arm Dynein Heavy Chain alpha 13	K10408	R054_01g16350	NF02_02g17140	99.83	9.10	0.000	Cro31.914527.11.1
DHC13 (ODA-DHC1), Outer Arm Dynein Heavy Chain beta 14	K10408	R054_03g120070	NF02_27g06280	99.83	8.37	0.000	Cro9.940380.11.2
DHC15 (ODA-DHC1), Outer Arm Dynein Heavy Chain gamma 15	K10408	R054_20g08860	NF02_15g02110	99.94	9.13	0.000	Cro11.8476050.11.0
DI1 (ODA-IC1), Outer Arm Dynein Intermediate Chain 1	K10409	R054_03g15360	NF02_03g101290	99.85	8.54	0.000	Cro12.636550.11.0
DI2 (ODA-IC2), Outer Arm Dynein Intermediate Chain 2	K11143	R054_01g13880	NF02_01g12550	99.83	8.55	0.000	Cro12.65900.11.0
DI3 (ODA-IC3), Calmodulin-like Outer Arm Dynein Light Chain 4	—	R054_37g13940	NF02_37g09160	100.00	12.73	0.000	Cro12.65900.11.0
DI4 (ODA-IC8) (RSP22), Dynein Light Chain LC8; LC8 component of outer arm dynein, inner arm 11/1 and RSP22 of radial spokes	K10418	R054_04g19810	NF02_04g11820	100.00	9.64	0.000	Cro12.65900.11.0
DI4 (ODA-IC8) (RSP22), Dynein Light Chain LC8; LC8 component of outer arm dynein, inner arm 11/1 and RSP22 of radial spokes	K10418	R054_29g0670	NF02_19g13340	97.40	-7.87	0.000	Cro12.65900.11.0
DI2 (ODA-LC6), Outer Arm Dynein Light Chain 6; LC6 homolog	K10418	R054_01g16350	NF02_01g10180	100.00	8.86	0.000	Cro31.81720.11.1
DI4 (ODA-LC10), Outer Arm Dynein light chain (DNAL4 in mammals)	K10412	R054_09g27310	NF02_05g121090	100.00	8.88	0.000	Cro12.65900.11.0
DI1 (ODA-L7), Outer Arm Dynein Light Chain 7; Outer Arm Dynein Light Chain 7	K10419	R054_01g14050	NF02_01g16420	100.00	11.39	0.000	Cro12.65900.11.0
DI1 (ODA-L7), Tctex-1-like Dynein Light Chain LC9; LC9 component of outer arm dynein	K10420	R054_03g16660	NF02_03g10140	100.00	10.40	0.000	Cro12.628750.11.0
DI2 (ODA-LC2), Tctex2-like Outer Arm Dynein Light Chain 2	K23866	R054_26g14450	NF02_26g76150	100.00	12.00	0.000	Cro12.627750.11.0
DI1 (ODA-LC1), Outer Arm Dynein Light Chain 1	K10411	R054_38g14440	NF02_38g05290	100.00	12.11	0.000	Cro12.627750.11.0
DL1 (ODA-LC3), Outer Arm Dynein Light Chain 3	—	R054_02g12070	NF02_29g06440	100.00	10.09	0.000	Cro12.6528850.11.1
DI2 (ODA-LC5), Threonase-like Flagellar Outer Arm Dynein 14 kDa Light Chain LC5	—	R054_24g10120	NF02_24g45440	98.21	7.32	0.000	Cro17.614250.11.0
ODA10, Outer Arm Dynein Assembly ODA10	—	R054_41g14810	NF02_41g107240	100.00	12.25	0.000	Cro6.298120.11.2
Axoneme: Inner Dynein Arms							
DHC11 (IA1-DHC1), Axonemal Inner Arm Dynein 11/1 Heavy Chain alpha	K10408	R054_02g12460	NF02_01g10700	100.00	7.62	0.000	Cro12.6842450.11.1
DHC10 (IA1-DHC1), Axonemal Inner Arm Dynein 11/1 Heavy Chain beta 10	K10408	R054_18g13210	NF02_18g10790	99.85	6.01	0.000	Cro12.642950.11.1
DHC2, Axonemal Inner Arm Dynein e Heavy Chain 2	K10408	R054_11g15770	NF02_12g08890	99.91	6.56	0.000	Cro9.92822.11.1
DHC5, Axonemal Inner Arm Dynein b Heavy Chain 5	K10408	R054_38g14340	NF02_38g09150	99.22	8.27	0.000	Cro12.642950.11.1
DHC6, Axonemal Inner Arm Dynein g Heavy Chain 6	K10408	R054_01g16720	NF02_02g17160	98.84	5.06	0.000	Cro5.244250.11.0
DHC7, Axonemal Inner Arm Dynein g Heavy Chain 7	K10408	R054_27g11280	NF02_23g06270	99.71	5.13	0.000	Cro14.6275.11.1
DHC7, Axonemal Inner Arm Dynein g Heavy Chain 7	K10408	R054_27g11370	NF02_23g13600	99.32	5.40	0.000	Cro14.6275.11.1
DHC8, Axonemal Inner Arm Dynein e Heavy Chain 8	—	R054_01g16420	NF02_02g17160	99.45	6.22	0.000	Cro16.895450.11.0
DHC8, Axonemal Inner Arm Dynein e Heavy Chain 8	K10408	R054_18g16350	NF02_18g14210	99.79	8.08	0.000	Cro12.641005.11.1
DHC9, Axonemal Inner Arm Dynein c Heavy Chain 9	K10408	R054_39g14490	NF02_38g16130	99.69	8.43	0.000	Cro12.641005.11.1
DI3 (IDA1-IC1), Dynein Intermediate Chain IC10; Flagellar inner dynein arm 11 intermediate chain IC10	K24722	R054_05g08790	NF02_05g12240	99.74	10.20	0.000	Cro16.74515.11.1
DI4 (IDA-IC2), Dynein Intermediate Chain IC10; Flagellar inner dynein arm 11 intermediate chain	K24723	R054_02g12110	NF02_01g10420	99.61	7.57	0.000	Cro12.620950.11.1
DI1 (IDA-IC8), Dynein Light chain p28; Flagellar inner dynein light chain p28	K10410	R054_05g13710	NF02_10g10820	100.00	8.84	0.000	Cro12.649400.11.1
DI2 (FAP146) (p38) (ZMYND12), Dynein Light Chain p38	—	R054_27g11140	NF02_23g11510	100.00	6.19	0.000	Cro12.649400.11.1
DI3, TPR-domain Dynein Light Chain	K23769	R054_06g0480	NF02_06g13760	100.00	21.87	0.000	Cro6.74855.11.1
DI3, TPR-domain Dynein Light Chain; Subunit of axonemal inner dynein d with TPR domains	K25474	R054_05g0360	NF02_05g12300	100.00	13.76	0.000	Cro6.74855.11.1
DI6 (FAP94) (DWA7), Dynein Intermediate Chain IC97; Flagellar inner Arm Dynein 11/1 intermediate chain	K17580	R054_13g6730	NF02_14g18550	100.00	7.23	0.000	Cro12.6950.11.2
DI2 (PL2) (Caltractin) (Centrin), 20 kD Calcium-Binding Protein Centrin; component of flagellar inner arm dyneins and basal body/centriole-associated structures	K16485	R054_08g1910	NF02_08g12540	100.00	8.53	0.000	Cro12.6950.11.2
DI2 (PL2) (Caltractin) (Centrin), 20 kD Calcium-Binding Protein Centrin; component of flagellar inner arm dyneins and basal body/centriole-associated structures	K16485	R054_18g1010	NF02_18g12540	99.75	7.20	0.000	Cro12.6950.11.2
DI2 (PL2) (Caltractin) (Centrin), 20 kD Calcium-Binding Protein Centrin; component of flagellar inner arm dyneins and basal body/centriole-associated structures	K16485	R054_26g10970	NF02_31g19740	100.00	-0.93	0.000	Cro12.6950.11.2
OL3, Text-1-like Dynein Light Chain; Present in DA11/1 and IFT Cytoplasmic Dynein 1b; Flagellar inner arm dynein light chain	K10420	R054_12g0430	NF02_11g10480	100.00	5.54	0.000	Cro12.694250.11.0
OL14 (Tctex20), Dynein Light Chain; Associated with DA11/1 and IFT Cytoplasmic Dynein 1b; Associated with Inner Arm Dynein 11/1 and Cytoplasmic Dynein 1b	K23866	R054_01g10200	NF02_02g17160	100.00	9.17	0.000	Cro12.694250.11.0
FAP57, WD-domain Coiled-coil Flagellar Associated Protein 57	K24729	R054_03g13540	NF02_38g16190	99.83	10.32	0.000	Cro12.694250.11.0
FAP43, Flagellar Associated Protein 43	K24223	R054_04g12440	NF02_04g115220	99.36	7.88	0.000	Cro12.694250.11.0
FAP44, Flagellar Associated Protein 44	K24224	R054_22g16410	NF02_25g17110	98.90	8.43	0.000	Cro12.694250.11.0
Axoneme: Dynein assembly factors							
DMAF1 (ODA7) (DAU1) (LRRCS5), Dynein Assembly Factor 1	K19750	R054_15g17420	NF02_20g11720	99.26	7.30	0.001	Cro11.809150.11.2
DMAF1 (ODA7) (DAU1) (LRRCS5), Dynein Assembly Factor 1	K19750	R054_25g10610	NF02_22g49180	98.67	5.51	0.000	Cro12.62915.11.1
DMAF1 (ODA7) (DAU1) (LRRCS5), Dynein Assembly Factor 1	K19750	R054_38g141570	NF02_38g09150	99.31	4.73	0.000	Cro12.601350.11.0
DMAF10 (WDR62), Axonemal Dynein Assembly Factor 10	K24761	R054_04g17190	NF02_04g11570	98.36	5.90	0.000	Cro16.82000.11.2
DMAF11 (MOT47) (LRRCS8) (Seahorse), Dynein Assembly Factor 11	K19753	R054_22g17780	NF02_25g02320	98.61	5.93	0.000	Cro17.788650.11.2
DMAF12 (OL2) (ODA8) (MOT37) (LRRCS6), Outer Arm Dynein Assembly Factor	K25425	R054_41g14800	NF02_41g107170	99.22	7.45	0.000	Cro12.649400.11.1
DMAF16 (DAB2) (CPAP29) (FBF18) (C22ORF59), Dynein Assembly Factor 16	K24229	R054_08g12080	NF02_18g12540	100.00	8.63	0.000	Cro16.884450.11.2
DMAF17 (CPAP30) (FBF8), Dynein Assembly Factor 17	K24230	R054_15g17200	NF02_20g12350	100.00	3.56	0.000	Cro12.65900.11.0
DMAF18 (DAW1) (CDA16) (WDR69), Dynein Assembly Factor 18; Outer Arm Dynein-IFT Adaptor	K19760	R054_05g27500	NF02_05g121260	100.00	4.09	0.000	Cro12.65900.11.0
DMAF19 (CCDC103), Dynein Assembly Factor 19	K27371	R054_10g15790	NF02_10g10520	98.67	8.24	0.000	Cro6.23304.11.1
DMAF2 (DAP1) (PF13) (Iku), Dynein Assembly Factor 3	K19751	R054_05g12350	NF02_10g11720	94.93	5.83	0.000	Cro9.941450.11.2
DMAF2 (DAP1) (PF22), Dynein Assembly Factor 3	K19752	R054_24g10420	NF02_24g56270	99.24	0.54	0.001	Cro12.601057.11.1
DMAF4 (DYX1C1) (PF23), Dynein Assembly Factor 4	K19758	R054_17g19060	NF02_17g12060	97.67	7.58	0.016	Cro12.649750.11.1
DMAF5 (HEATR1), Dynein Assembly Factor 5	K19759	R054_13g16750	NF02_14g17550	99.52	0.28	0.124	Cro9.99550.11.0
DMAF6 (TNM1) (PHD13) (Twister), Dynein Assembly Factor 6	K24263	R054_03g13910	NF02_38g02210	100.00	10.02	0.000	Cro17.33580.11.1
DMAF7 (ZMYND10), Dynein Assembly Factor 7	K24030	R054_03g13780	NF02_03g09660	99.84	8.74	0.000	Cro12.649750.11.1
DMAF9, Dynein Assembly Factor 9	K25424	R054_02g12080	NF02_02g12220	99.69	6.81	0.000	Cro11.84755.11.1
RUBL1 (Purin), Nucleoside-triphosphatase RuB-like Protein	K04499	R054_01g15150	NF02_02g12050	99.72	0.11	0.234	Cro9.941450.11.2
RUBL2 (Reptin), Nucleoside-triphosphatase RuB-like Protein	K11338	R054_30g12120	NF02_27g06410	99.59	0.91	0.000	Cro12.44770.11.2
Axoneme: Modifier of Inner arms (MIA) complex							
FAP73, Flagellar Associated Protein 73; MIA2 subunit of the modifier of inner arms (MIA) complex	K25440	R054_03g15350	NF02_03g101220	99.75	11.05	0.000	Cro16.889800.11.2
FAP73, Flagellar Associated Protein 73; MIA2 subunit of the modifier of inner arms (MIA) complex	K25440	R054_18g16650	NF02_18g10240	99.70	7.65	0.000	Cro16.889800.11.2
FAP73, Flagellar Associated Protein 100; MIA1 subunit of the modifier of inner arms (MIA) complex	K25439	R054_41g148210	NF02_41g107310	99.43	6.52	0.000	Cro12.83950.11.0
Axoneme: Nexin-dynein regulatory complex							
DRC1, Dynein Regulatory Complex 1	K19754	R054_43g151870	NF02_43g110750	98.50	9.41	0.000	AFM1554.1
DRC2 (FAP250), Dynein Regulatory Complex 2	K27378	R054_25g10580	NF02_22g49800	100.00	9.89	0.000	Cro12.667750.11

FAP252, EF Hand Flagellar Associated Protein 252
FAP252, EF Hand Flagellar Associated Protein 252
FAP306 (RB2.1), Flagellar Associated Protein 306
FAP306 (RB2.1), Flagellar Associated Protein 306
FAP363, Stabilizer Of Axonemal Microtubules (SAVO)-like, HSPT70-like Flagellar Associated protein 363

Axoneme: Central apparatus associated proteins (C1 microtubule)

CAM1, Calmodulin 1
CAM1, Calmodulin 1
CAM1, Calmodulin 1
CAM1, Calmodulin 1
CAM1, Calmodulin 1
CPC1, Central Pair 1
FAP42, Flagellar Associated Protein 42
FAP47, Flagellar Associated Protein 47
FAP54, Flagellar Associated Protein 54
FAP69, Flagellar Associated Protein 69
FAP74, Flagellar Associated Protein 74
FAP76, Flagellar Associated Protein 76
FAP81, Flagellar Associated Protein 81
FAP89, Flagellar Associated Protein 79
FAP99, Flagellar Associated Protein 99
FAP101, Flagellar Associated Protein 101
FAP114, Coiled-Coil Flagellar Associated Protein 114
FAP119, Flagellar Associated Protein 119
FAP174, MYCBP-like Flagellar Associated Protein 174
FAP194, Flagellar Associated Protein 194
FAP227, Flagellar Associated Protein 227
FAP246, Leucine Rich Repeat Flagellar Associated Protein 246
FAP266, Flagellar Associated Protein 266
FAP276, Flagellar Associated Protein 276
FAP306 (MO17), Spata17-like Flagellar Associated Protein 305
HYD3, Hydrocephalus 3-like Protein Hyd3
PF 16, Central Apparatus Protein PF16

Axoneme: Central apparatus associated proteins (C2 microtubule)

FAP20, Flagellar Associated Protein 20
FAP65, Flagellar Associated Protein 65
FAP70, Flagellar Associated Protein 70
FAP147, Flagellar Associated Protein 147
FAP178, Spata4-like Flagellar Associated Protein 178
FAP196, Flagellar Associated Protein 196
FAP213, Coiled-coil Flagellar Associated Protein 213
FAP221, Coiled-Coil Flagellar Associated Protein 221
FAP225, Flagellar Associated Protein 225
FAP238, Coiled-coil Flagellar Associated Protein 239
FAP388, Flagellar Associated Protein 388
PF 20, Central Apparatus Associated WD-domain Protein

Intraflagellar transport: complex A

IFT43, Intraflagellar Transport Protein 43
IFT121 (FAP118) IFT122B, Intraflagellar Transport Protein 121
IFT122 (FAP80), Intraflagellar Transport Protein 122
IFT139 (FAP60), Intraflagellar Transport Protein 139
IFT140, Intraflagellar Transport Protein 140
IFT144 (FAP60), Intraflagellar Transport Protein 144

Intraflagellar transport: complex B

IFT20, Intraflagellar Transport Protein 20
IFT21 (FAP5), Intraflagellar Transport Protein 22
IFT25 (FAP23), Intraflagellar Transport Protein 25
IFT27 (FAP156), Intraflagellar Transport Protein 27
IFT38 (FAP22 [Oly-D] Dy-3), Intraflagellar Transport Protein 38
IFT40 (FAP2), Intraflagellar Transport Protein 46
IFT52, Intraflagellar Transport Protein 52
IFT54 (FAP118 [Dy-11] [Eliapa]), Intraflagellar Transport Protein 54
IFT56 (Dy-13) [C26b], Intraflagellar Transport Protein 56
IFT57, Intraflagellar Transport Protein 57
IFT70 (FAP259 [Dy-1] [Fleer]), Intraflagellar Transport Protein 70
IFT74, Intraflagellar Transport Protein 74
IFT80 (FAP167), Intraflagellar Transport Protein 80
IFT81, Intraflagellar transport protein 81
IFT88, Intraflagellar transport protein 88
IFT172, Intraflagellar Transport Protein 172

Intraflagellar Transport: Dyneins

DHC16 (DHC1b), Intraflagellar Transport Dynein Heavy Chain 1b, Cytoplasmic Dynein 1b/2 Heavy Chain
DIC5 (FAP133) [D1BC2] [WD34], WD-domain Intermediate Chain of Dynein 1b
DIC9 (FAP163) [D1BC1] [Mammalian WD80b], Dynein 1b Intermediate Chain IC1
DUI (D13ALC), Dynein 1b Light Intermediate Chain D13ALC

Intraflagellar Transport: Other associated proteins

ARMC2 (PF27), IFT Adaptor for Radial Spokes
CEP19, Centrosomal Protein 19
RABL2, Rab-like GTP Binding Protein 2
TLP1, Tubby-like Protein 1
TLP1, Tubby-like Protein 1

Intraflagellar transport: Kinesin-2 and associated proteins

FLA8 (KIN2A), Kinesin-2 Motor Protein 2A
FLA10 (KIN2B), Kinesin-2 Motor Protein 2B
KAP, Kinesin-2 Associated Protein 1

Intraflagellar Transport: BBSome

Btsp1, BBSome Interacting Protein 1
BBS1, Bardet-Biedl syndrome 1
BBS2, Bardet-Biedl syndrome 2
BBS3 (ARL8) [HFA1A], Bardet-Biedl Syndrome 3
BBS4, Bardet-Biedl syndrome 4
BBS5, Bardet-Biedl syndrome 5
BBS7, Bardet-Biedl syndrome 7
BBS8, Bardet-Biedl syndrome 8
BBS9, Bardet-Biedl syndrome 9
Cep115 (A21), Centrosomal Protein 121
LZTFL1, Leucine Zipper Transcription Factor Like 1

Intraflagellar Transport: BBS-CCT complex

CCT2, T-complex beta Subunit
CCT2, T-complex protein 1 subunit beta
CCT3, T-complex gamma Subunit
CCT4, T-complex delta Subunit 4
CCT5, T-complex epsilon Subunit
CCT8, T-complex theta Subunit
MKKS, Kluver-Bucay-Kaufman syndrome protein

Transition zone: MKS complex

BSD1, B9-domain Transition Zone Protein 1
BSD2A, BSD2-like Transition Zone Protein
CCT20A (MKS8), Coiled-coil C2-domain Flagella Associated Protein
MKSL, Centriole Proteome Protein
Retinitis pigmentosa GTPase regulator-interacting protein 1
TC7N1, Tectonic Transition Zone Protein
TMEM107, transmembrane protein 107
TMEM216, Transmembrane Protein 216
TMEM231, Transmembrane protein 231
TMEM67, Transition Zone Protein 67

Transition zone: Nephronophthisis-related proteins

NHP4, Nephrocytin-4

Other flagella associated proteins

ARL13B, ARL12L1 ADP-ribosylation factor-like protein 13B
ARL2BP, ARL2 binding protein 2
ARMC9
CEP20, Centrosomal Protein 20
CEP44, centrosomal protein CEP44
CH232, Tugami-like Flagellar Protein
DIP13, Deflagellation Inducible Protein, 13KD
DZP1L, DAZ Interacting Protein 1 Like
FAP36, Flagellar Associated Protein 36
FAP189, Coiled-coil SMC-domain Flagellar Associated Protein 189
FAP206, Flagellar Associated Protein 206
FAP296 (CEP104) [KCC22], Ciliary Tip-localized Flagellar Associated Protein 256
FAP299 (CFAP298) [FBB1], Flagellar Associated Protein 299
FAP313, Ar13-like Flagellar Associated Protein 313
FAP343, Flagellar Associated Protein 343
FAP403, Kinesin-like Flagellar Associated Protein 403
FAP410, Flagellar Associated Protein 410
FB810, Flagellar and Basal Body 10
FB811, Flagellar and Basal Body Protein 11
FB87, Flagellar and Basal Body Protein 7
MAPK7, Mitogen-Activated Protein Kinase 7
POC1, Proteome of Centriole Protein 1
POC16, WD-domain Centriolar Protein
RP2, protein RP2
Spn1-like Protein
STB36, Serine/Threonine Kinase 36

—	R054_11f66960	NF02_12g10010	99.55	8.86	0.000	Cre10.g468650.11.2
—	R054_23g100800	NF02_19g49410	100.00	7.20	0.000	Cre10.g468650.11.2
K25475	R054_01g04410	NF02_02g73960	99.70	10.30	0.000	Cre03.g178500.11.1
K25475	R054_10g53170	NF02_10g22220	99.01	8.63	0.000	Cre03.g178500.11.2
K04043	R054_04g24940	NF02_04g118760	99.44	10.18	0.001	Cre12.g45850.11.2
K02183	R054_03g15680	NF02_03g101490	100.00	3.21	0.000	Cre03.g178150.11.1
K02183	R054_07g78960	NF02_07g131850	100.00	7.24	0.000	Cre03.g178150.11.1
K02183	R054_22g7590	NF02_22g49110	100.00	8.19	0.000	Cre03.g178150.11.1
K02183	R054_43g15090	NF02_43g110900	100.00	6.00	0.000	Cre03.g178150.11.1
K02183	R054_48g14500	NF02_48g109250	100.00	7.70	0.000	Cre03.g178150.11.1
K25615	R054_03g15700	NF02_03g101510	99.83	22.67	0.000	Cre03.g183200.11.2
—	R054_01g00930	NF02_02g70380	99.63	9.67	0.000	Cre12.g19960.11.1
K25552	R054_21g10140	NF02_21g00130	99.84	8.83	0.000	Cre17.g704300.11.1
—	R054_12g81910	NF02_12g06650	99.47	8.19	0.000	Cre12.g183550.11.1
K24227	R054_03g13990	NF02_03g09880	98.76	9.07	0.000	Cre03.g186200.11.2
—	R054_18g51570	NF02_18g22740	99.31	9.38	0.000	Cre06.g710150.11.1
—	R054_19g5950	NF02_19g23560	99.93	8.33	0.000	Cre09.g38768.11.1
—	R054_17g9130	NF02_17g26840	99.48	7.63	0.000	Cre06.g26850.11.2
K25509	R054_04g20360	NF02_04g114290	99.43	8.37	0.000	Cre14.g24400.11.2
K55608	R054_30g14210	NF02_30g07970	99.48	9.28	0.000	Cre14.g24400.11.2
K25390	R054_17g80340	NF02_17g28200	99.51	7.89	0.000	Cre02.g121100.11.1
K25547	R054_09g50220	NF02_09g13890	100.00	8.53	0.000	Cre09.g88262.11.1
K25547	R054_08g4350	NF02_08g12110	99.84	7.73	0.000	Cre06.g26450.11.2
—	R054_04g24970	NF02_04g118790	94.12	9.52	0.000	Cre02.g105950.11.2
K25532	R054_17g19880	NF02_17g28790	98.68	10.07	0.000	Cre12.g221150.11.1
—	R054_27g11180	NF02_27g13180	100.00	5.12	0.000	Cre17.g728950.11.2
—	R054_27g12450	NF02_27g151790	99.35	10.16	0.000	Cre14.g618750.11.1
K25458	R054_17g77950	NF02_17g25760	99.43	7.09	0.000	Cre16.g60450.11.1
—	R054_31g23200	NF02_31g06320	99.77	5.00	0.000	Cre06.g26850.11.2
K25546	R054_34g130970	NF02_34g26300	98.51	6.22	0.000	Cre11.g482300.11.2
K17570	R054_35g13420	NF02_35g06030	99.49	8.19	0.000	Cre01.g252400.11.2
K25532	R054_01g02010	NF02_02g71390	100.00	11.39	0.000	Cre09.g94251.11.1
K25470	R054_14g70860	NF02_14g15390	100.00	4.50	0.000	Cre07.g351650.11.1
K4226	R054_13g66410	NF02_14g18290	99.74	9.21	0.000	Cre07.g354551.11.2
K24932	R054_19g8600	NF02_19g23010	99.20	7.50	0.000	Cre17.g45400.11.2
K2548	R054_26g10200	NF02_26g01900	99.12	6.40	0.000	Cre14.g22450.11.1
K25616	R054_27g11180	NF02_27g151700	99.57	20.31	0.000	Cre10.g418150.11.2
—	R054_08g41410	NF02_08g125150	99.84	8.07	0.000	Cre17.g78650.11.1
—	R054_28g41710	NF02_28g07950	99.02	7.65	0.000	Cre06.g26450.11.2
K25610	R054_14g68570	NF02_14g13090	99.46	9.47	0.004	Cre11.g476376.11.1
—	R054_39g144700	NF02_39g59430	98.96	9.31	0.000	Cre17.g351050.11.2
—	R054_24g12320	NF02_24g36480	99.31	9.44	0.000	Cre03.g162400.11.2
—	R054_43g12320	NF02_43g111110	99.33	8.74	0.000	Cre14.g12250.11.2
K4735	R054_17g80660	NF02_17g28440	98.07	9.96	0.000	Cre04.g227900.11.2
K18675	R054_40g146820	NF02_40g106260	95.80	4.19	0.000	Cre06.g251200.11.2
K18674	R054_21g12260	NF02_21g05700	99.76	8.12	0.000	Cre11.g47320.11.1
K18656	R054_29g19760	NF02_29g06820	100.00	7.51	0.000	Cre01.g06822.11.1
K18673	R054_23g101200	NF02_19g54240	99.85	8.00	0.000	Cre07.g26880.11.2
K18672	R054_31g23200	NF02_31g06320	98.28	6.01	0.000	Cre06.g26880.11.2
K18671	R054_19g86650	NF02_19g24410	99.93	9.81	0.000	Cre06.g72700.11.2
K16473	R054_39g145650	NF02_39g06880	99.48	6.83	0.000	Cre02.g089950.11.2
K07935	R054_39g120740	NF02_27g36300	100.00	6.02	0.000	Cre01.g30320.11.2
K13969	R054_30g142130	NF02_30g07970	99.25	7.12	0.000	Cre10.g450150.11.2
K07934	R054_13g66160	NF02_14g18040	100.00	20.09	0.000	Cre01.g047960.11.1
K19684	R054_14g68820	NF02_13g13310	100.00	9.99	0.000	Cre17.g71250.11.2
K19682	R054_15g14260	NF02_15g24470	100.00	11.17	0.000	Cre05.g244637.11.1
K19681	R054_29g105600	NF02_22g46660	100.00	8.83	0.000	Cre04.g12520.11.2
K19680	R054_23g68310	NF02_19g26220	99.41	9.11	0.000	Cre11.g467739.11.1
K19685	R054_03g10040	NF02_03g100520	100.00	11.29	0.000	Cre11.g467610.11.1
K04839	R054_17g19150	NF02_17g26960	100.00	9.60	0.000	Cre10.g467000.11.2
K19683	R054_11g55050	NF02_12g08220	99.85	13.24	0.000	Cre07.g342200.11.2
K19679	R054_09g8750	NF02_09g13710	99.83	12.07	0.000	Cre01.g27950.11.2
K19678	R054_11g55040	NF02_12g08110	100.00	8.16	0.000	Cre07.g23600.11.2
K19677	R054_10g4750	NF02_02g73900	99.85	10.24	0.000	Cre17.g23600.11.2
K16474	R054_35g135670	NF02_36g1990	98.98	6.79	0.000	Cre07.g33750.11.1
K19676	R054_32g123340	NF02_29g7130	100.00	4.58	0.999	Cre17.g703900.11.2
K10414	R054_30g120250	NF02_27g26270	99.49	3.85	0.000	Cre06.g250300.11.2
K22868	R054_33g130650	NF02_34g7350	99.66	6.73	0.000	Cre02.g110950.11.1
K22699	R054_12g61280	NF02_11g7690	99.28	7.31	0.001	Cre10.g28664.11.1
K10417	R054_11g53570	NF02_02g72660	99.28	7.86	0.000	Cre09.g38862.11.1
K24123	R054_03g19180	NF02_03g104880	97.07	7.49	0.000	Cre12.g559300.11.2
K18801	R054_15g1740	NF02_20g19170	100.00	10.24	0.000	Cre12.g487350.11.2
K07931	R054_39g144880	NF02_39g06820	99.57	20.32	0.000	Cre17.g72390.11.2
K19600	R054_05g19890	NF02_05g12100	99.68	6.10	0.000	Cre12.g26230.11.2
K19600	R054_35g134120	NF02_36g09480	99.57	11.07	0.000	Cre12.g50650.11.2
K20196	R054_05g25280	NF02_05g110900	99.59	7.67	0.000	Cre12.g52550.11.2
—	R054_04g49310	NF02_06g73200	99.11	5.41	0.000	Cre17.g73950.11.2
K25388	R054_03g17530	NF02_03g10340	100.00	5.52	0.000	Cre10.g46250.11.2
K19399	R054_19g85770	NF02_18g23400	100.00	8.42	0.000	AV035245
K16748	R054_23g100170	NF02_19g54430	99.85	8.67	0.000	Cre17.g741950.11.1
K16747	R054_04g2180	NF02_04g117150	99.87	7.55	0.000	Cre06.g57250.11.1
K07951	R054_15g7280	NF02_20g24550	99.43	7.01	0.000	Cre16.g64500.11.2
K16531	R054_35g135720	NF02_36g02010	100.00	9.05	0.000	Cre12.g458650.11.1
K16748	R054_04g24380	NF02_04g113260	95.93	6.85	0.000	Cre06.g76550.11.2
K16749	R054_07g7790	NF02_07g13190	99.64	7.65	0.000	Cre01.g244760.11.1
K16781	R054_24g101900	NF02_24g54280	100.00	11.60	0.000	Cre16.g66500.11.2
K19398	R054_04g25290	NF02_04g112680	99.67	6.86	0.000	Cre04.g15700.11.1
K16540	R054_05g12490	NF02_05g25500	99.59	11.17	0.000	Cre09.g39473.11.1
K19400	R054_04g22860	NF02_04g116820	99.07	20.86	0.000	Cre02.g262500.11.2
K09493	R054_23g107180	NF02_22g50000	99.82	-0.15	0.438	Cre10.g491100.11.2
K09494	R054_07g36970	NF02_07g133070	99.81	1.12	0.000	Cre09.g161750.11.2
K09495	R054_13g65860	NF02_14g71720	100.00	-0.45	0.000	Cre10.g445250.11.2
K09496	R054_04g23350	NF02_04g117320	99.63	0.31	0.004	Cre05.g234639.11.1
K09497	R054_11g6600	NF02_12g09650	99.81	1.20	0.000	Cre03.g156750.11.2
K09500	R054_03g20840	NF02_03g09440	99.09	1.00	0.000	Cre03.g164540.11.2
K09492	R054_04g2490	NF02_04g116950	99.79	3.54	0.000	—
K16774	R054_04g0530	NF02_04g14520	100.00	7.55	0.000	Cre17.g74627.11.1
K16745	R054_21g5320	NF02_21g24630	99.45	6.07	0.000	Cre03.g196050.11.2
K16746	R054_21g5320	NF02_21g24630	99.45	6.07	0.000	Cre03.g196050.11.2
K15332	R054_37g14510	NF02_37g43110	91.03	6.65	0.000	Cre12.g485350.11.1
K16754	R054_09g46640	NF02_09g13530	100.00	4.55	0.000	Cre14.g23050.11.2
K15382	R054_04g2180	NF02_04g117150	99.59	7.55	0.000	Cre06.g57250.11.1
K22670	R054_18g1770	NF02_18g0330	100.00	11.21	0.000	Cre12.g45400.11.1
K19385	R054_01g0770	NF02_02g70230	99.66	5.72	0.000	Cre08.g77850.11.2
K19382	R054_04g420	NF02_02g72490	100.00	7.53	0.000	Cre14.g21250.11.2
K19381	R054_04g420	NF02_02g72490	99.74	6.73	0.000	Cre13.g49401.11.1
K16478	R054_27g12530	NF02_27g32460	99.35	9.96	0.000	Cre12.g31400.11.1
K07962	R054_26g10680	NF02_26g013890	99.51	6.63	0.000	Cre09.g39151.11.1
K17472	R054_04g2180	NF02_04g117150	99.55	7.55	0.000	Cre06.g57250.11.1
K22684	R054_11g6940	NF02_11g45780	99.90	5.71	0.000	Cre05.g10951.11.1
K22684	R054_23g9710	NF02_19g4000	99.89	2.99	0.000	Cre05.g237100.11.1
K18485	R054_20g12010	NF02_20g43070	99.23	6.27	0.000	—
K18486	R054_20g12010	NF02_20g43070	99.23	6.27	0.000	—
K16780	R054_20g3860	NF02_20g15490	94.00	8.15	0.000	Cre07.g74550.11.2
K16740	R054_05g7070	NF02_05g123310	99.22	12.69	0.000	Cre01.g04400.11.1
K25395	R054_05g7070	NF02_05g123310	99.22	12.69	0.000	Cre01.g04400.11.1
K25394	R054_05g7070	NF02_05g123310	99.22	12.69	0.000	Cre01.g04400.11.1
K25354	R054_32g11340	NF02_27g3650	100.00	8.00	0.000	Cre13.g48400.11.2
K2541	R054_19g2650	NF02_20g27130	99.85	8.46	0.000	Cre04.g23040.11.1
K15458	R054_04g2180	NF02_04g117150	99.85	7.55	0.000	Cre06.g57250.11.1
K25611	R054_14g6860	NF02_14g13010	100.00	3.71	0.000	Cre11.g42200.11.2
K17044	R054_02g2380	NF02_04g116390	100.00	15.01	0.000	Cre04.g18250.11.2
K15462	R054_04g2180	NF02_04g117150	99.85	7.55	0.000	Cre06.g57250.11.1
K08874	R054_01g0180	NF02_01g0180	99.91	3.47	0.000	Cre01.g001300.11.1
K25478	R054_11g8910	NF02_11g11880	99.32	4.70	0.000	Cre14.g25550.11.2
K25475	R054_30g11540	NF02_30g0670	99.18	5.02	0.000	Cre07.g78500.11.2
K17078	R054_04g2180	NF02_04g117150	99.51	7.55	0.000	Cre06.g57250.11.1
K28229	R054_22g9570	NF02_22g56910	99.82	9.83	0.000	Cre01.g14827.11.1
K16789	R054_01g4590	NF02_02g72740	99.62	11.59	0.000	Cre16.g161100.11.2

Supplementary Table 5. Annotations and expression patterns of *P. marina* UIO007 and CCMP1203 genes encoding Kdo biosynthesis and transport proteins.

KEGG definition	KO identifier	Protein ID		% identity	DeSeq2 analyses		
		UIO007	CCMP1203		log2FC	padj	DeepLoc2.1 localization
Kdo biosynthesis proteins							
kdsD, kpsF; arabinose-5-phosphate isomerase [EC:5.3.1.13]	K06041	RI054_02g09940	NFJ02_01g38410	100.00	8.26	0.000	Cytoplasm (0.6720),
kdsA; 2-dehydro-3-deoxyphosphooctonate aldolase (KDO 8-P synthase) [EC:2.5.1.55]	K01627	RI054_03g14560	NFJ02_03g100440	100.00	7.95	0.000	Cytoplasm (0.7371)
kdsB; 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase) [EC:2.7.7.38]	K00979	RI054_19g86100	NFJ02_16g23690	96.69	8.23	0.000	Mitochondrion (0.6266)
kdsB; 3-deoxy-manno-octulosonate cytidylyltransferase (CMP-KDO synthetase) [EC:2.7.7.38]	K00979	RI054_01g01720	NFJ02_02g71070	100.00	5.68	0.000	Cytoplasm (0.5687)
Putative CMP-Kdo Transporters							
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_08g43440	NFJ02_06g127130	99.70	6.91	0.000	Golgi apparatus (0.8588)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_11g54990	no homolog	—	1.46	0.000	Golgi apparatus (0.8183)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_12g62420	NFJ02_11g07110	98.49	-2.43	0.000	Golgi apparatus (0.6291)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_16g74940	no homolog	—	21.04	0.000	Golgi apparatus (0.7777)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_16g75060	no homolog	—	19.78	0.000	Golgi apparatus (0.7083)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_16g75740	no homolog	—	19.93	0.000	Golgi apparatus (0.7823)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_16g75940	no homolog	—	20.41	0.000	Golgi apparatus (0.7663)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_16g77300	no homolog	—	20.49	0.000	Golgi apparatus (0.7537)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_17g80490	NFJ02_17g28300	100.00	6.80	0.000	Golgi apparatus (0.7445)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_19g85940	NFJ02_16g23550	100.00	14.05	0.000	Golgi apparatus (0.8250)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_35g135330	NFJ02_36g91640	96.63	5.95	0.001	Endoplasmic reticulum (0.7586)
SLC35A1_2_3; solute carrier family 35 (UDP-sugar transporter), member A1/2/3	K15272	RI054_42g150430	NFJ02_44g112070	95.68	9.81	0.000	Golgi apparatus (0.8008)

Supplementary Table 6. Annotations and expression patterns of *P. marina* UIO007 and CCMP1203 genes encoding selected CAZymes.

CAZyme family	Protein ID		% identity	DeSeq2 analyses		DeepLoc 2.1 localisation
	UIO007	CCMP1203		log2FC	padj	
GT14	RI054_18g84080	NFJ02_18g31480	82.68	7.69	0.002	Golgi apparatus (0.8397)
	RI054_21g92350	NFJ02_21g45490	98.60	6.21	0.000	Golgi apparatus (0.7110)
GT23	RI054_09g46090	NFJ02_08g134760	99.74	5.50	0.000	Extracellular (0.5234)
	RI054_16g77110	NFJ02_22g48840	41.11	11.07	0.000	Extracellular Golgi apparatus
	RI054_16g77590	NFJ02_08g134760	35.29	20.70	0.000	Extracellular (0.5758)
	RI054_21g91440	NFJ02_21g44660	99.22	8.76	0.009	Golgi apparatus (0.8004)
	RI054_25g105780	NFJ02_22g48840	99.80	8.28	0.000	Golgi apparatus (0.8877)
	RI054_27g113380	NFJ02_23g53270	99.79	7.27	0.000	Extracellular (0.739899993)
	RI054_29g118110	NFJ02_28g65270	97.51	6.76	0.000	Extracellular (0.472099999)
GT29	RI054_01g00260	NFJ02_02g69860	99.47	6.46	0.000	Golgi apparatus (0.7551)
	RI054_01g02120	NFJ02_02g71510	98.00	20.60	0.000	Extracellular Golgi apparatus
	RI054_01g04690	NFJ02_02g73840	99.43	5.42	0.000	Golgi apparatus (0.5837)
	RI054_01g04930	NFJ02_02g74050	99.48	6.16	0.000	Golgi apparatus (0.7871)
	RI054_01g05190	NFJ02_02g74320	100.00	7.29	0.000	Extracellular Golgi apparatus
	RI054_02g08200	NFJ02_01g37000	99.82	4.37	0.000	Golgi apparatus (0.7398)
	RI054_02g09540	NFJ02_01g38060	99.17	5.50	0.000	Golgi apparatus (0.8176)
	RI054_02g10550	NFJ02_01g39050	99.22	8.90	0.000	Golgi apparatus (0.8458)
	RI054_02g12850	NFJ02_01g41080	99.11	6.43	0.000	Extracellular (0.8761)
	RI054_02g12950	NFJ02_01g41180	98.77	19.53	0.000	Golgi apparatus (0.8695)
	RI054_03g13790	NFJ02_03g99660	93.96	8.21	0.000	Golgi apparatus (0.6317)
	RI054_03g17850	NFJ02_03g103510	99.29	6.38	0.000	Extracellular (0.5593)
	RI054_04g20390	NFJ02_04g114310	99.51	4.27	0.000	Extracellular (0.6813)
	RI054_04g20930	NFJ02_04g114870	99.77	5.40	0.000	Extracellular (0.4548)
	RI054_04g24590	NFJ02_04g118420	98.54	-0.12	0.435	Golgi apparatus (0.7727)
	RI054_04g24810	NFJ02_04g118630	99.77	7.65	0.000	Golgi apparatus (0.7089)
	RI054_05g25570	NFJ02_05g119400	97.52	0.19	0.137	Extracellular Golgi apparatus
	RI054_05g25620	NFJ02_05g119450	99.24	2.62	0.000	Golgi apparatus (0.8104)
	RI054_05g26910	NFJ02_05g120690	99.60	6.39	0.000	Golgi apparatus (0.9140)
	RI054_05g27020	NFJ02_05g120800	99.52	1.59	0.000	Extracellular (0.6410)
	RI054_05g28700	NFJ02_05g122360	99.14	5.99	0.090	Endoplasmic reticulum (0.5434)
	RI054_05g29430	NFJ02_05g123080	100.00	7.74	0.000	Extracellular (0.6674)
	RI054_05g30870	NFJ02_05g124460	98.30	3.63	0.000	Extracellular Golgi apparatus
	RI054_06g33680	NFJ02_09g141660	99.74	7.36	0.001	Golgi apparatus (0.8208)
	RI054_06g34250	NFJ02_09g142220	99.65	8.17	0.000	Golgi apparatus (0.8634)
	RI054_07g36460	NFJ02_07g130360	99.27	5.55	0.000	Golgi apparatus (0.8286)
	RI054_07g37580	NFJ02_07g131610	100.00	7.01	0.002	Extracellular Golgi apparatus
	RI054_07g38760	NFJ02_07g132770	99.89	6.12	0.000	Golgi apparatus (0.7383)
	RI054_08g41970	NFJ02_06g125560	99.69	5.41	0.000	Golgi apparatus (0.7859)
	RI054_08g42770	NFJ02_06g126450	99.27	5.39	0.000	Golgi apparatus (0.7630)
	RI054_08g43250	NFJ02_06g126890	99.86	5.48	0.000	Golgi apparatus (0.7560)
	RI054_08g43470	NFJ02_06g127160	99.72	3.22	0.000	Golgi apparatus (0.7871)
	RI054_08g45240	NFJ02_06g128990	99.83	5.71	0.000	Extracellular Golgi apparatus
	RI054_09g46730	NFJ02_08g135420	97.84	-0.41	0.004	Extracellular (0.7695)
	RI054_09g49800	NFJ02_08g138440	99.85	5.87	0.000	Golgi apparatus (0.8302)
	RI054_10g50560	NFJ02_10g00020	99.14	5.50	0.000	Golgi apparatus (0.6711)
	RI054_10g50570	NFJ02_10g00030	99.28	3.40	0.000	Golgi apparatus (0.7146)
	RI054_10g54830	NFJ02_10g03890	95.87	6.23	0.000	Golgi apparatus (0.8157)
	RI054_11g55510	NFJ02_12g08580	100.00	6.25	0.000	Extracellular Golgi apparatus
	RI054_11g56480	NFJ02_12g09550	100.00	6.20	0.001	Extracellular (0.6324)
	RI054_11g57110	NFJ02_12g10140	99.81	5.36	0.000	Extracellular (0.5096)
	RI054_12g59820	NFJ02_11g04290	98.15	2.77	0.000	Golgi apparatus (0.7880)
	RI054_12g60650	NFJ02_11g05170	99.61	8.19	0.000	Extracellular (0.6335)
	RI054_12g62900	NFJ02_11g07570	99.52	8.37	0.000	Golgi apparatus (0.8210)
	RI054_13g67230	NFJ02_14g19090	100.00	4.55	0.000	Golgi apparatus (0.8214)
	RI054_14g70560	NFJ02_13g15130	99.47	9.20	0.004	Golgi apparatus (0.9052)
	RI054_15g71480	NFJ02_20g41750	98.40	4.96	0.008	Extracellular Golgi apparatus
	RI054_15g71500	NFJ02_20g41750	98.56	4.79	0.000	Golgi apparatus (0.8235)
	RI054_15g71550	NFJ02_20g41800	99.46	21.29	0.000	Golgi apparatus (0.8847)
	RI054_15g73140	NFJ02_20g43180	100.00	6.50	0.000	Extracellular (0.6595)
	RI054_15g74780	NFJ02_20g44620	98.79	5.72	0.000	Golgi apparatus (0.6491)
	RI054_17g77920	NFJ02_17g25730	98.76	6.47	0.000	Golgi apparatus (0.7652)
	RI054_17g77930	NFJ02_17g25740	98.73	6.66	0.003	Extracellular (0.7412)
	RI054_17g78520	NFJ02_17g26340	98.87	8.78	0.006	Extracellular (0.6779)
	RI054_17g79400	NFJ02_17g27190	99.23	5.87	0.000	Extracellular Golgi apparatus
	RI054_18g83310	NFJ02_18g30890	99.36	11.42	0.000	Extracellular Golgi apparatus
	RI054_18g84860	NFJ02_33g84430	99.41	1.98	0.000	Cytoplasm (0.2319)
	RI054_19g86390	NFJ02_16g23980	99.82	4.19	0.000	Golgi apparatus (0.5440)
	RI054_19g86400	NFJ02_16g23990	96.93	7.41	0.000	Extracellular (0.5694)
	RI054_20g88930	NFJ02_15g20060	100.00	17.64	0.000	Extracellular (0.7087)
	RI054_20g91310	NFJ02_15g22430	95.36	5.90	0.000	Golgi apparatus (0.8317)
	RI054_21g92210	NFJ02_21g45360	100.00	5.50	0.000	Golgi apparatus (0.6573)
	RI054_21g92420	NFJ02_21g45560	100.00	6.94	0.000	Golgi apparatus (0.6222)
	RI054_21g93450	NFJ02_21g46520	99.74	7.61	0.000	Golgi apparatus (0.7948)
	RI054_21g93940	NFJ02_21g47010	98.99	4.20	0.000	Extracellular (0.5996)
	RI054_21g94590	NFJ02_21g47550	99.80	8.37	0.000	Golgi apparatus (0.6368)
	RI054_22g97530	NFJ02_25g59050	97.96	0.66	0.041	Extracellular (0.8727)
	RI054_23g101210	NFJ02_19g35430	98.96	0.86	0.000	Extracellular Golgi apparatus
	RI054_23g98250	NFJ02_19g32540	99.22	4.28	0.000	Extracellular (0.7801)
	RI054_23g99090	NFJ02_19g33380	99.70	5.73	0.000	Golgi apparatus (0.7579)
	RI054_24g102210	NFJ02_24g54490	99.05	7.54	0.000	Golgi apparatus (0.8999)
	RI054_24g103970	NFJ02_24g56120	98.59	-0.64	0.008	Golgi apparatus (0.8677)
	RI054_25g105170	NFJ02_22g48220	99.44	0.79	0.000	Golgi apparatus (0.8178)
	RI054_25g105590	NFJ02_22g48650	98.28	5.43	0.003	Golgi apparatus (0.8511)
	RI054_25g106170	NFJ02_22g49230	100.00	9.17	0.004	Extracellular Golgi apparatus
	RI054_26g108090	NFJ02_31g78670	99.78	6.66	0.000	Golgi apparatus (0.8485)
	RI054_26g110270	NFJ02_31g80430	99.01	3.80	0.000	Cell membrane (0.4061)
	RI054_27g112400	NFJ02_23g51840	100.00	8.54	0.000	Extracellular Golgi apparatus
	RI054_28g114720	NFJ02_30g76310	100.00	6.85	0.000	Golgi apparatus (0.9165)
	RI054_28g115210	NFJ02_30g76740	99.50	5.06	0.000	Extracellular (0.7358)
	RI054_29g118100	NFJ02_28g65240	98.52	6.07	0.000	Extracellular (0.5992)
	RI054_29g118590	NFJ02_28g65720	97.01	6.14	0.000	Golgi apparatus (0.7289)

	RI054_29g119080	NFJ02_28g66150	98.17	5.08	0.000	Golgi apparatus (0.7347)
	RI054_29g119440	NFJ02_28g66510	92.96	4.51	0.000	Golgi apparatus (0.8325)
	RI054_31g122320	NFJ02_26g59390	98.74	5.89	0.000	Golgi apparatus (0.7460)
	RI054_32g127850	NFJ02_29g69490	99.25	5.67	0.002	Golgi apparatus (0.7278)
	RI054_33g130480	NFJ02_34g87210	99.20	7.52	0.001	Golgi apparatus (0.6930)
	RI054_34g132730	NFJ02_33g84430	96.50	-0.61	0.001	Extracellular (0.6033)
	RI054_35g134310	NFJ02_36g90680	99.40	0.07	0.754	Extracellular Golgi apparatus
	RI054_35g135260	NFJ02_36g91570	100.00	2.40	0.000	Golgi apparatus (0.7635)
	RI054_36g136860	NFJ02_35g88070	97.23	4.34	0.000	Golgi apparatus (0.7720)
	RI054_36g137290	NFJ02_35g88500	99.56	7.27	0.000	Golgi apparatus (0.5855)
	RI054_37g139370	NFJ02_37g93110	99.45	4.13	0.000	Golgi apparatus (0.7754)
	RI054_38g143050	NFJ02_39g98790	99.26	6.83	0.000	Extracellular Golgi apparatus
	RI054_40g146580	NFJ02_40g105860	99.08	0.72	0.000	Extracellular Golgi apparatus
	RI054_40g146670	NFJ02_40g105980	99.53	9.25	0.000	Golgi apparatus (0.8658)
	RI054_42g149950	NFJ02_44g111680	99.39	4.19	0.031	Golgi apparatus (0.7471)
	RI054_42g150280	NFJ02_44g111940	99.25	7.13	0.000	Golgi apparatus (0.5747)
	RI054_43g152230	NFJ02_43g111050	98.53	6.93	0.000	Extracellular Golgi apparatus
	RI054_45g153900	NFJ02_42g108910	98.95	6.14	0.000	Extracellular (0.5727)
	RI054_45g154300	NFJ02_42g109270	96.97	6.92	0.000	Golgi apparatus (0.8503)
	RI054_45g154340	NFJ02_42g109300	98.68	4.13	0.000	Extracellular (0.7445)
	RI054_45g154600	NFJ02_42g109560	98.92	3.59	0.000	Golgi apparatus (0.7955)
GT43	RI054_01g01580	NFJ02_02g70950	65.56	-8.95	0.000	Golgi apparatus (0.9011)
	RI054_16g77500	NFJ02_02g70950	99.04	20.11	0.000	Golgi apparatus (0.8113)
	RI054_18g84020	NFJ02_18g31410	99.13	-8.59	0.000	Golgi apparatus (0.7964)
GT47	RI054_12g82880	NFJ02_11g07560	98.94	4.45	0.000	Golgi apparatus (0.9193)
	RI054_13g83970	NFJ02_14g16030	98.92	6.04	0.000	Extracellular (0.7121)
	RI054_16g74910	NFJ02_32g82320	89.71	3.96	0.000	Extracellular Golgi apparatus
	RI054_16g75340	no homolog	—	20.73	0.000	Golgi apparatus (0.8636)
	RI054_21g93570	NFJ02_21g46650	99.72	5.58	0.000	Golgi apparatus (0.7317)
	RI054_22g95100	NFJ02_25g56840	98.99	2.57	0.000	Endoplasmic reticulum Golgi apparatus
	RI054_23g98060	NFJ02_19g32390	99.63	9.71	0.000	Extracellular (0.7272)
	RI054_24g104440	NFJ02_24g56500	98.84	0.17	0.401	Extracellular (0.6211)
	RI054_26g109740	NFJ02_31g79970	99.63	4.67	0.000	Extracellular (0.4395)
	RI054_41g148750	NFJ02_41g107780	99.10	6.85	0.003	Endoplasmic reticulum Golgi apparatus
	RI054_42g149970	NFJ02_44g111700	93.73	-0.96	0.649	Endoplasmic reticulum (0.6075)
GT61	RI054_03g14490	NFJ02_03g100370	85.17	-1.53	0.000	Cell membrane (0.5972)
	RI054_17g78120	NFJ02_17g25940	98.99	6.20	0.000	Extracellular (0.6041)
	RI054_24g103590	NFJ02_24g55790	97.75	7.82	0.000	Extracellular (0.8056)
GT77	RI054_01g04230	NFJ02_02g73400	98.60	7.20	0.000	Endoplasmic reticulum Golgi apparatus
	RI054_03g14850	NFJ02_03g100720	99.43	4.31	0.000	Golgi apparatus (0.9071)
	RI054_04g19980	NFJ02_04g113820	100.00	4.57	0.000	Golgi apparatus (0.7849)
	RI054_04g24520	NFJ02_04g118370	99.03	1.75	0.000	Golgi apparatus (0.8580)
	RI054_09g50130	NFJ02_08g138750	98.14	-8.66	0.000	Golgi apparatus (0.8021)
	RI054_18g83060	NFJ02_18g30630	99.86	6.01	0.000	Golgi apparatus (0.9024)
	RI054_28g114160	NFJ02_30g75900	98.87	3.27	0.000	Endoplasmic reticulum Golgi apparatus
	RI054_29g118370	NFJ02_28g55530	99.28	8.13	0.000	Extracellular (0.7919)
	RI054_29g118380	NFJ02_28g55530	99.69	2.63	0.000	Golgi apparatus (0.8477)
	RI054_32g126620	NFJ02_29g58270	99.72	4.75	0.000	Golgi apparatus (0.9160)
	RI054_37g141070	NFJ02_37g94820	99.72	4.01	0.000	Golgi apparatus (0.7561)
GT95	RI054_04g20700	NFJ02_04g114680	98.73	3.89	0.000	Golgi apparatus (0.8282)
	RI054_07g36330	NFJ02_07g130230	98.38	6.23	0.000	Golgi apparatus (0.9210)
	RI054_28g117190	NFJ02_30g78540	99.71	1.13	0.000	Golgi apparatus (0.8800)
	RI054_31g124740	NFJ02_26g51730	100.00	6.67	0.000	Golgi apparatus (0.8677)
	RI054_38g141760	NFJ02_39g97590	97.93	6.08	0.001	Golgi apparatus (0.9329)
GT96	RI054_15g73220	NFJ02_20g43230	97.99	6.64	0.000	Golgi apparatus (0.7391)
GH33	RI054_01g04020	NFJ02_02g73180	99.52	1.48	0.000	Golgi apparatus (0.8472)
	RI054_02g11400	NFJ02_01g39700	99.25	3.65	0.000	Golgi apparatus (0.8220)
	RI054_02g13020	NFJ02_01g41230	99.35	7.36	0.031	Endoplasmic reticulum Golgi apparatus
	RI054_03g19240	NFJ02_03g104920	98.81	4.26	0.000	Extracellular Lysosome/Vacuole Golgi apparatus
	RI054_05g28520	NFJ02_05g122150	99.81	3.25	0.000	Golgi apparatus (0.7811)
	RI054_05g29950	NFJ02_05g123570	99.81	10.04	0.000	Golgi apparatus (0.7227)
	RI054_05g30600	NFJ02_05g124170	99.48	10.67	0.000	Golgi apparatus (0.8821)
	RI054_07g39550	NFJ02_07g133540	99.33	7.80	0.000	Endoplasmic reticulum Golgi apparatus
	RI054_12g63300	NFJ02_11g07980	99.39	8.53	0.000	Extracellular Golgi apparatus
	RI054_14g68500	NFJ02_13g13020	98.88	-0.99	0.255	Cytoplasm (0.5692)
	RI054_14g69900	NFJ02_13g14490	99.47	19.98	0.000	Golgi apparatus (0.7651)
	RI054_17g81370	NFJ02_17g29000	99.63	6.60	0.000	Golgi apparatus (0.6605)
	RI054_21g92910	NFJ02_21g45980	99.05	10.07	0.000	Golgi apparatus (0.7658)
	RI054_23g100280	NFJ02_19g34500	99.11	5.71	0.000	Golgi apparatus (0.8235)
	RI054_24g102270	NFJ02_24g54540	99.76	7.75	0.000	Extracellular (0.8654)
	RI054_25g105180	NFJ02_22g48230	99.81	7.33	0.001	Golgi apparatus (0.8131)
	RI054_25g106000	NFJ02_22g49040	99.11	8.64	0.000	Cytoplasm Lysosome/Vacuole
	RI054_25g107640	NFJ02_22g50760	99.44	6.53	0.057	Cytoplasm (0.6944)
	RI054_25g107880	NFJ02_22g50940	98.92	0.15	0.331	Golgi apparatus (0.8659)
	RI054_26g110860	NFJ02_31g81020	99.41	5.29	0.000	Golgi apparatus (0.6672)
	RI054_27g111570	NFJ02_23g51530	98.85	7.20	0.001	Endoplasmic reticulum Golgi apparatus
	RI054_27g112560	NFJ02_23g52500	99.53	8.57	0.000	Golgi apparatus (0.8223)
	RI054_31g123640	NFJ02_26g50640	66.21	6.74	0.000	Golgi apparatus (0.7922)
GH43	RI054_36g136180	NFJ02_35g87460	99.10	5.07	0.000	Golgi apparatus (0.6442)
	RI054_36g138540	NFJ02_35g89740	99.29	4.52	0.000	Golgi apparatus (0.8618)
	RI054_44g153290	NFJ02_45g113230	98.00	1.69	0.000	Extracellular (0.8333)
GH141	RI054_02g11190	NFJ02_01g39590	99.18	5.51	0.000	Extracellular (0.8248)
	RI054_09g47500	NFJ02_08g136140	99.53	4.64	0.000	Extracellular (0.8576)
	RI054_43g152000	NFJ02_43g110870	99.26	5.24	0.000	Extracellular (0.55835)
	RI054_01g03130	NFJ02_02g72500	94.43	7.16	0.000	Extracellular (0.7794)
	RI054_15g73700	NFJ02_20g43690	99.61	3.50	0.000	Lysosome/Vacuole (0.5868)
	RI054_20g90210	NFJ02_15g21310	99.47	5.85	0.000	Cell membrane (0.5777)
	RI054_29g119010	NFJ02_28g56090	99.33	3.69	0.065	Extracellular Cell membrane
	RI054_40g147510	NFJ02_40g106660	98.54	7.80	0.000	Cell membrane Lysosome/Vacuole
	RI054_40g147530	NFJ02_40g106670	95.33	5.89	0.001	Extracellular (0.7880)

Supplementary Table 7. Annotations and expression patterns of *P. marina* UIO007 and CCMP1203 genes encoding life-cycle-related proteins.

Gene	Product	Protein ID		% Identity	DeSeq2 analyses		Additional notes
		UIO007	CCMP1203		log2FC	padj	
Core meiotic genes (meiosis-specific genes are denoted in bold)							
Spo11	meiotic recombination protein SPO11	R054_41g149150	NFJ02_41g108270	100.00	-0.18	0.322	Double-strand break: creates DNA double-strand break in homologous chromosomes
Hse11	double-strand break repair protein HSE11	R054_22g07950	NFJ02_25g59070	97.21	-1.12	0.000	Double-strand break: trims back broken DNA ends and hairpins
Rad50	DNA repair protein RAD50	R054_28g114330	NFJ02_30g76030	99.73	0.41	0.000	Double-strand break: holds broken DNA ends together while Hse11 trims
Hop1	HORMA domain-containing protein	R054_13g63890	NFJ02_14g15990	99.30	-0.75	0.019	Homologous alignment and synaptonemal complex: forms axial and lateral elements of the synaptonemal complex
Pch2	pachytene checkpoint protein 2	R054_05g26360	NFJ02_05g120170	98.28	1.44	0.000	Homologous alignment and synaptonemal complex: initiation of meiotic recombination
Dmc1	meiotic recombination protein DMC1	R054_09g49770	NFJ02_08g138410	99.73	-1.35	0.000	Recombination: promotes interhomolog recombination
Hop2	homologous-pairing protein 2	R054_25g106700	NFJ02_22g49770	98.99	-12.51	0.000	Recombination: ensures accurate and efficient homology searching during pachytene stage of meiotic prophase
Mnd1	meiotic nuclear division protein 1	R054_42g150010	NFJ02_44g111750	99.13	-5.55	0.000	Recombination: required for stable heteroduplex formation
Rad51	DNA repair protein RAD51	R054_02g10270	NFJ02_01g39220	100.00	0.26	0.032	Recombination: catalyzes homologous DNA pairing and strand exchange
Exo1	exonuclease 1	R054_01g06680	NFJ02_02g75570	99.37	-0.51	0.000	Cross-over resolution pathway: involved in crossing-over and its resolution
Mer3	ATP-dependent DNA helicase HFM1/MER3	R054_14g69880	NFJ02_13g14470	98.01	-2.87	0.000	Cross-over resolution pathway: promotes Holliday junction resolution with crossover interference
Mhr1	DNA mismatch repair protein MHR1	R054_09g47290	NFJ02_08g135950	99.07	-0.86	0.000	Cross-over resolution pathway: mismatch repair of dinucleotide and trinucleotide sequences
Msh4	DNA mismatch repair protein MSH4	R054_28g115560	NFJ02_30g77070	100.00	-2.72	0.000	Cross-over resolution pathway: directs Holliday junction resolution towards crossover with interference
Msh5	DNA mismatch repair protein MSH5	R054_13g66180	NFJ02_14g18070	99.72	-1.28	0.000	Cross-over resolution pathway: directs Holliday junction resolution towards crossover with interference
Mus81	crossover junction endonuclease MUS81	R054_02g10090	NFJ02_01g38580	99.35	-1.30	0.000	Cross-over resolution pathway: involved in crossing-over and its resolution
Msh2	DNA mismatch repair protein MSH2	R054_15g72440	NFJ02_20g42530	99.09	-1.44	0.000	Gene conversion: forms a heterodimer with Msh3 or Msh6
Msh3/6	DNA mismatch repair protein MSH6	R054_04g21080	NFJ02_04g115080	99.10	-0.11	0.413	Gene conversion: forms a heterodimer with Msh2, binds base-base mismatches
Pms1	DNA mismatch repair protein PMS1	R054_33g128260	NFJ02_34g85290	89.70	-0.35	0.189	Gene conversion: mismatch repair
Rec8/Rad21	cohesin complex subunit SCC1	R054_08g41170	NFJ02_06g124910	100.00	-0.56	0.000	Cohesin complex: holds Smc1 and Smc3 heads together by binding N-terminal domain to Smc3 and C-terminal domain to Smc1, thus holding sister chromatids together during mitosis and meiosis
Rec8/Rad21	cohesin complex subunit SCC1	R054_30g121890	NFJ02_27g54060	99.73	-0.29	0.182	Cohesin complex: holds Smc1 and Smc3 heads together by binding N-terminal domain to Smc3 and C-terminal domain to Smc1, thus holding sister chromatids together during mitosis and meiosis
Scc3	cohesin complex subunit SA-1/2	R054_11g58120	NFJ02_12g10910	99.67	-1.57	0.000	Cohesin complex: interacts with Smc1, Smc3 and Rec8/Rad21 in holding cohesin ring together
Smc1	structural maintenance of chromosome 1	R054_41g149310	NFJ02_41g108430	99.69	-1.37	0.000	Cohesin complex: forms a heterodimer with Smc3 to form core sister chromatid cohesin subunits
Smc3	structural maintenance of chromosome 3	R054_24g103350	NFJ02_24g55540	99.85	-1.93	0.000	Cohesin complex: forms a heterodimer with Smc1 to form core sister chromatid cohesin subunits, with ring shape around sister chromatids
Pds5	sister chromatid cohesin protein PDS5	R054_03g17950	NFJ02_03g103610	99.23	-1.26	0.000	Cohesin complex: important for maintenance of sister chromatid cohesion in late prophase
Smc2	structural maintenance of chromosome 2	R054_17g68040	NFJ02_17g28260	99.52	-1.24	0.000	Condensin complex: forms a heterodimer with Smc4 to form core condensin subunits, ring shape, essential for chromosome assembly and segregation
Smc4	structural maintenance of chromosome 4	R054_25g10780	NFJ02_22g50930	99.84	-1.03	0.000	Condensin complex: forms a heterodimer with Smc3 to form core sister chromatid cohesin subunits
Smc5	structural maintenance of chromosomes protein 5	R054_01g01300	NFJ02_02g70680	99.45	-1.56	0.000	SMC5-SMC6 complex: forms a heterodimer with Smc6 (Rad18) and is involved in DNA repair and checkpoint responses
Smc6	structural maintenance of chromosomes protein 6	R054_29g119420	NFJ02_28g66480	97.76	-1.68	0.000	SMC5-SMC6 complex: forms a heterodimer with Smc5 and is involved in DNA repair and checkpoint responses
Smc6	structural maintenance of chromosomes protein 6	R054_23g101390	NFJ02_19g35630	99.74	-1.29	0.000	SMC5-SMC6 complex: forms a heterodimer with Smc5 and is involved in DNA repair and checkpoint responses
Rad18	DNA repair protein RAD18	R054_03g16340	NFJ02_03g102180	99.26	-0.65	0.000	SMC5-SMC6 complex: involved in repairing double-strand
Genes encoding TALE homeobox transcription factors and syngamy proteins							
Gsm1	gamete-specific minus 1 protein	R054_03g17760	NFJ02_03g103430	98.49	5.58	0.001	GSM1 heterodimerizes with GSP1 to form the TALE homeobox transcription factor that initiates the development of the zygote after gamete fusion
Gsp1	gamete-specific plus 1 protein	R054_02g08940	NFJ02_01g37500	99.29	7.06	0.036	GSP1 heterodimerizes with GSM1 to form the TALE homeobox transcription factor that initiates the development of the zygote after gamete fusion
Hap2	hapless 2 male gamete fusion factor	R054_29g117830	NFJ02_28g54950	99.52	2.25	0.000	Mediates membrane fusion between gametes in a broad range of eukaryotes
Gex1	protein gamete expressed 1	R054_05g27710	NFJ02_05g121470	98.91	-0.83	0.000	Facilitates the fusion of gamete nuclei
Putative MID transcription factors of the RWP-RK family							
	RWP-RK domain-containing protein of 236 aa	R054_02g13170	NFJ02_01g41400	98.86	6.24	0.000	
	RWP-RK domain-containing protein of 546 aa	R054_05g29630	NFJ02_05g123240	94.96	1.07	0.002	
	RWP-RK domain-containing protein of 575 aa	R054_05g30440	NFJ02_05g124020	99.65	10.11	0.000	
	RWP-RK domain-containing protein of 627 aa	R054_06g30890	NFJ02_09g139240	99.84	9.09	0.000	
	RWP-RK domain-containing protein of 195 aa	R054_11g59180	NFJ02_12g11920	100.00	10.99	0.000	
	RWP-RK domain-containing protein of 384 aa	R054_13g66260	NFJ02_14g18140	98.18	7.33	0.001	
	RWP-RK domain-containing protein of 187 aa	R054_14g68950	NFJ02_13g13410	98.18	2.70	0.000	
	RWP-RK domain-containing protein of 535 aa	R054_17g78100	NFJ02_17g25920	98.69	6.75	0.000	
	RWP-RK domain-containing protein of 288 aa	R054_37g140540	NFJ02_37g94340	98.61	16.29	0.000	

Supplementary Table 8. NCBI accession numbers for SRA data of *P. marina* strains CCMP1203 and UIO007.

Species	BioProject	Biosample	SRA Accessions	Title	No. of Bases
<i>Pycnococcus provasolii</i> CCMP1203	PRJNA849025	SAMN29039285	SRB28475444	Random Oxford Nanopore sequencing of the <i>Pycnococcus provasolii</i> CCMP1203 genome (MinION_1)	2.6G
	PRJNA849025	SAMN29039285	SRB28475445	Random Illumina NovaSeq sequencing of the <i>Pycnococcus provasolii</i> CCMP1203 genome	22.6G
	PRJNA849025	SAMN29039285	SRB28475442	RNAseq Illumina NovaSeq sequencing of <i>Pycnococcus provasolii</i> CCMP1203 cDNAs	14.1G
	PRJNA849025	SAMN29039285	SRB28475443	Random Oxford Nanopore sequencing of the <i>Pycnococcus provasolii</i> CCMP1203 genome (MinION_2)	3.5G
	PRJNA849025	SAMN29039285	SRB33318468	Replica S1 of RNAseq AVITI sequencing of <i>Pycnococcus provasolii</i> CCMP1203 cDNAs	9G
	PRJNA849025	SAMN29039285	SRB33318467	Replica S2 of RNAseq AVITI sequencing of <i>Pycnococcus provasolii</i> CCMP1203 cDNAs	11.4G
	PRJNA849025	SAMN29039285	SRB33318466	Replica S3 of RNAseq AVITI sequencing of <i>Pycnococcus provasolii</i> CCMP1203 cDNAs	9.9G
	PRJNA1014779	SAMN37341342	SRB28476104	RNAseq Illumina NovaSeq sequencing of <i>Pseudoscourfieldia marina</i> K0017 cDNAs	17.9G
<i>Pseudoscourfieldia marina</i> UIO007	PRJNA1014779	SAMN37341342	SRB28476105	Random Oxford Nanopore sequencing of the <i>Pseudoscourfieldia marina</i> K0017 genome	2.1G
	PRJNA1014779	SAMN37341342	SRB28476106	Random Illumina NovaSeq sequencing of the <i>Pseudoscourfieldia marina</i> K0017 genome	2.5G
	PRJNA1014779	SAMN37341342	SRB33319090	Replica S1 of RNAseq AVITI sequencing of <i>Pseudoscourfieldia marina</i> K-0017 cDNAs	10.2G
	PRJNA1014779	SAMN37341342	SRB33319089	Replica S2 of RNAseq AVITI sequencing of <i>Pseudoscourfieldia marina</i> K-0017 cDNAs	3.7G
	PRJNA1014779	SAMN37341342	SRB33319088	Replica S3 of RNAseq AVITI sequencing of <i>Pseudoscourfieldia marina</i> K-0017 cDNAs	6.6G

Supplementary Report. MultiQC report summarizing the BUSCO results, as well as the dataset and assembly metrics.



General Statistics

Copy table

Configure columns

Scatter plot

Violin plot

Export as CSV...

Showing 18/18 rows and 10/15 columns.

Summarize table

Sample Name	N50 (Kbp)	Assembly Length (Mbp)	% Duplication	Reads After Filtering	GC content	% PF	% Adapter	Dups	GC	Seqs
CCMP1203_PE_R1			5.3 %	63.7 M	57.1 %	42.5 %	12.2 %	35.3 %	56.0 %	74.9 M
CCMP1203_PE_R2								33.2 %	56.0 %	74.9 M
CCMP1203_S1_R1.RNASeq			68.5 %	48.3 M	58.8 %	77.8 %	0.1 %	82.7 %	58.0 %	31.0 M
CCMP1203_S1_R2.RNASeq								79.9 %	58.0 %	31.0 M
CCMP1203_S2_R1.RNASeq			66.4 %	64.0 M	58.9 %	81.8 %	0.1 %	81.2 %	58.0 %	39.1 M
CCMP1203_S2_R2.RNASeq								75.9 %	58.0 %	39.1 M
CCMP1203_S3_R1.RNASeq			77.8 %	48.4 M	58.6 %	67.6 %	0.4 %	86.3 %	58.0 %	35.8 M
CCMP1203_S3_R2.RNASeq								84.6 %	58.0 %	35.8 M
CCMP1203_pr	715.8 Kbp	31.5 Mbp								
Pseudoscourf	819.8 Kbp	35.4 Mbp								
UIO007_PE_R1			12.6 %	7.5 M	56.7 %	41.6 %	0.9 %	24.8 %	55.0 %	9.0 M
UIO007_PE_R2								24.4 %	55.0 %	9.0 M
UIO007_S1_R1.RNASeq			66.5 %	59.6 M	59.4 %	86.2 %	0.0 %	78.6 %	59.0 %	34.6 M
UIO007_S1_R2.RNASeq								74.1 %	58.0 %	34.6 M
UIO007_S2_R1.RNASeq			61.8 %	18.9 M	59.4 %	73.2 %	0.1 %	70.1 %	59.0 %	12.9 M
UIO007_S2_R2.RNASeq								72.7 %	58.0 %	12.9 M
UIO007_S3_R1.RNASeq			81.2 %	31.5 M	59.6 %	65.1 %	0.1 %	89.3 %	59.0 %	24.2 M
UIO007_S3_R2.RNASeq								87.3 %	59.0 %	24.2 M

Long read data

Metrics calculated with keep_longest_reads.pl (<https://github.com/PombertLab>).

Copy table

Configure columns

Scatter plot

Violin plot

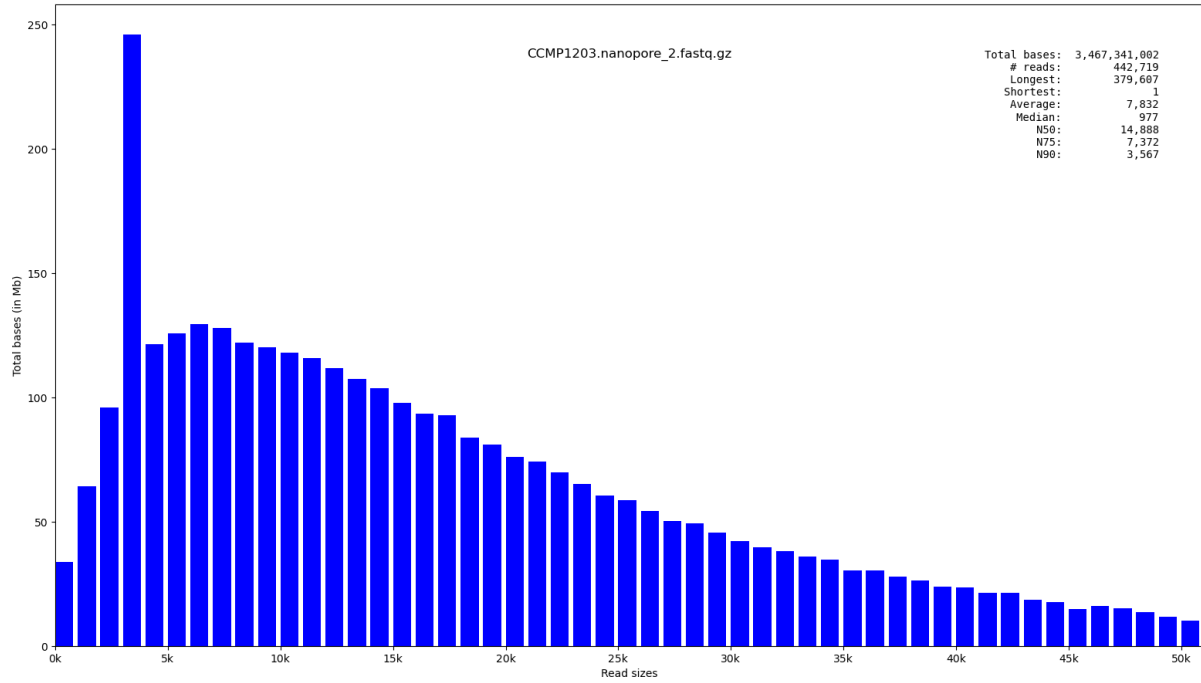
Export as CSV...

Showing 2/2 rows and 9/9 columns.

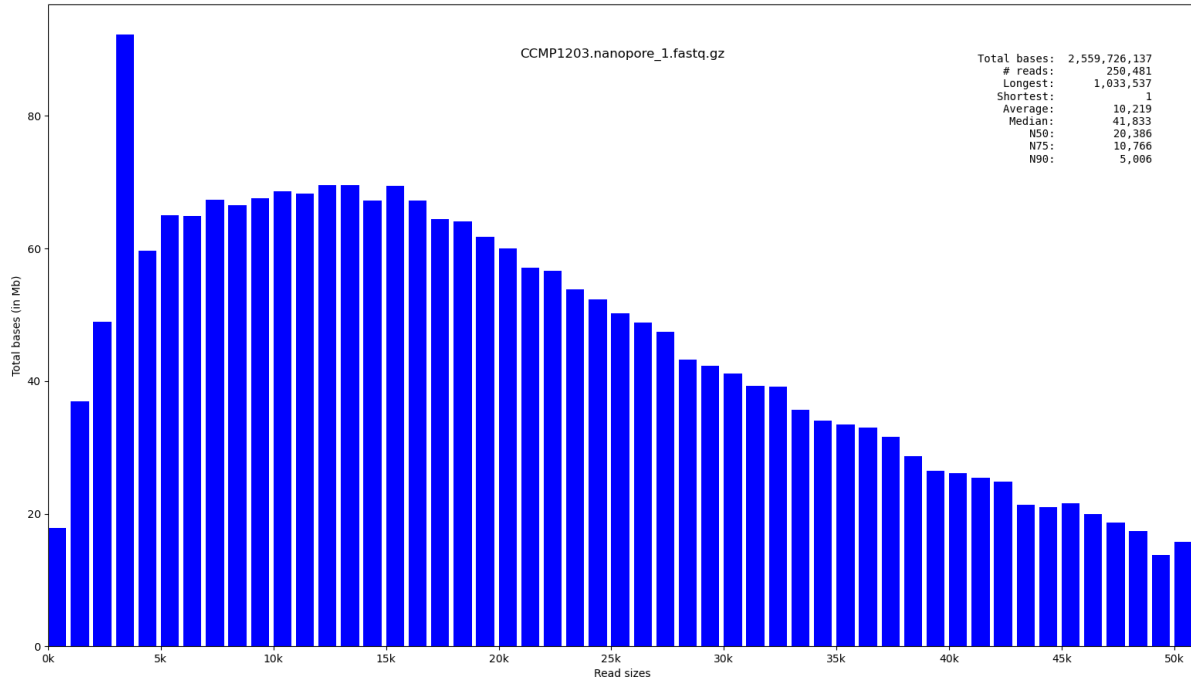
Summarize table

Sample Name	Number of reads	Total number of bases	Longest read	Shortest read	Average read size	Median read size	N50	N75	N90
CCMP1203.nan	442 719	3 467 341 002	379 607	1	7 832	4 192	14 888	7 372	3 567
UIO007.nanopo	162 303	2 134 051 402	310 039	4 000	13 149	8 724	17 461	9 358	5 882

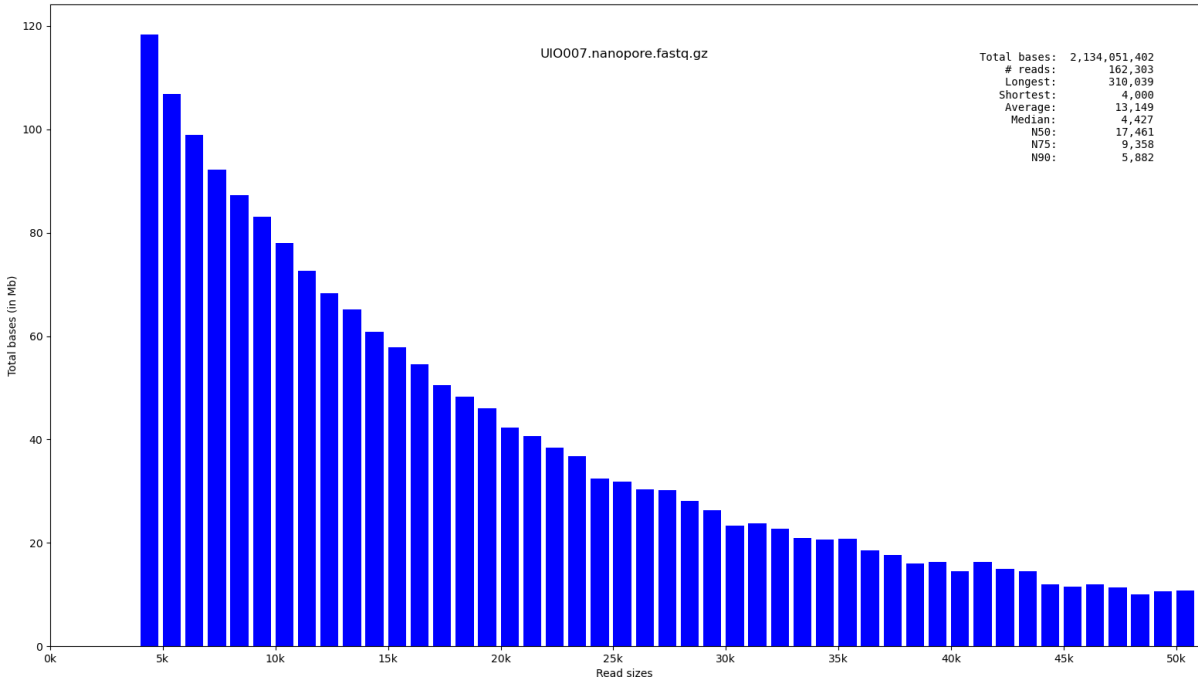
CCMP1203 nanopore 2



CCMP1203 nanopore 1



UIO007 nanopore



QUAST

Quality assessment tool for genome assemblies. URL: <http://quast.bioinf.spbau.ru> DOI: 10.1093/bioinformatics/btt086

Assembly Statistics

Copy table

Configure columns

Scatter plot

Violin plot

Export as CSV...

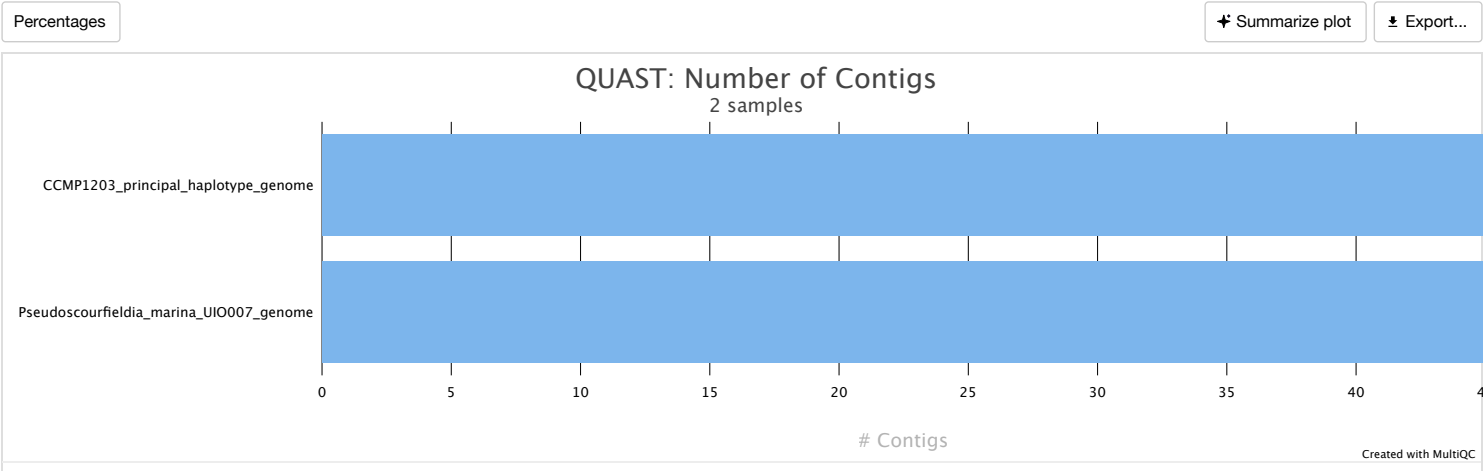
Showing $2\frac{1}{2}$ rows and $4\frac{1}{4}$ columns.

Summarize table

Sample Name	N50 (Kbp)	L50 (K)	Largest contig (Kbp)	Length (Mbp)
CCMP1203_principal_haplotype	715.8 Kbp	0.0 K	1 402.2 Kbp	31.5 Mbp
Pseudoscourfieldia_marina_UIC	819.8 Kbp	0.0 K	1 566.1 Kbp	35.4 Mbp

Number of Contigs

This plot shows the number of contigs found for each assembly, broken down by length.

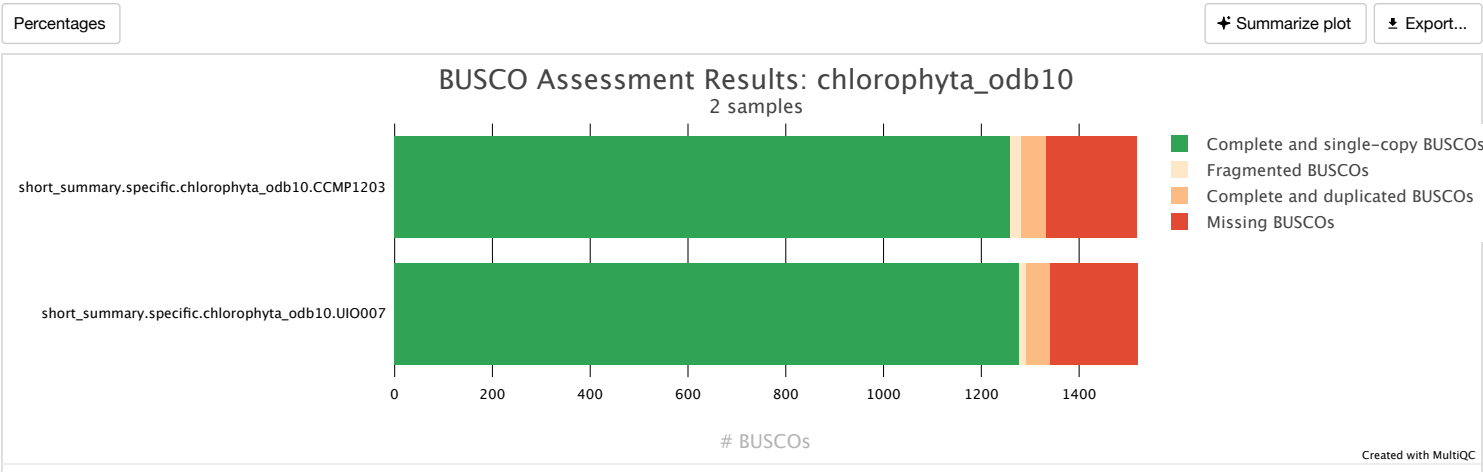


BUSCO Version: 5.8.2

Assesses genome assembly and annotation completeness. URL: <http://busco.ezlab.org> DOI: 10.1093/bioinformatics/btv351

BUSCO v2 provides quantitative measures for the assessment of genome assembly, gene set, and transcriptome completeness, based on evolutionarily-informed expectations of gene content from near-universal single-copy orthologs selected from OrthoDB v9.

Lineage: chlorophyta_odb10



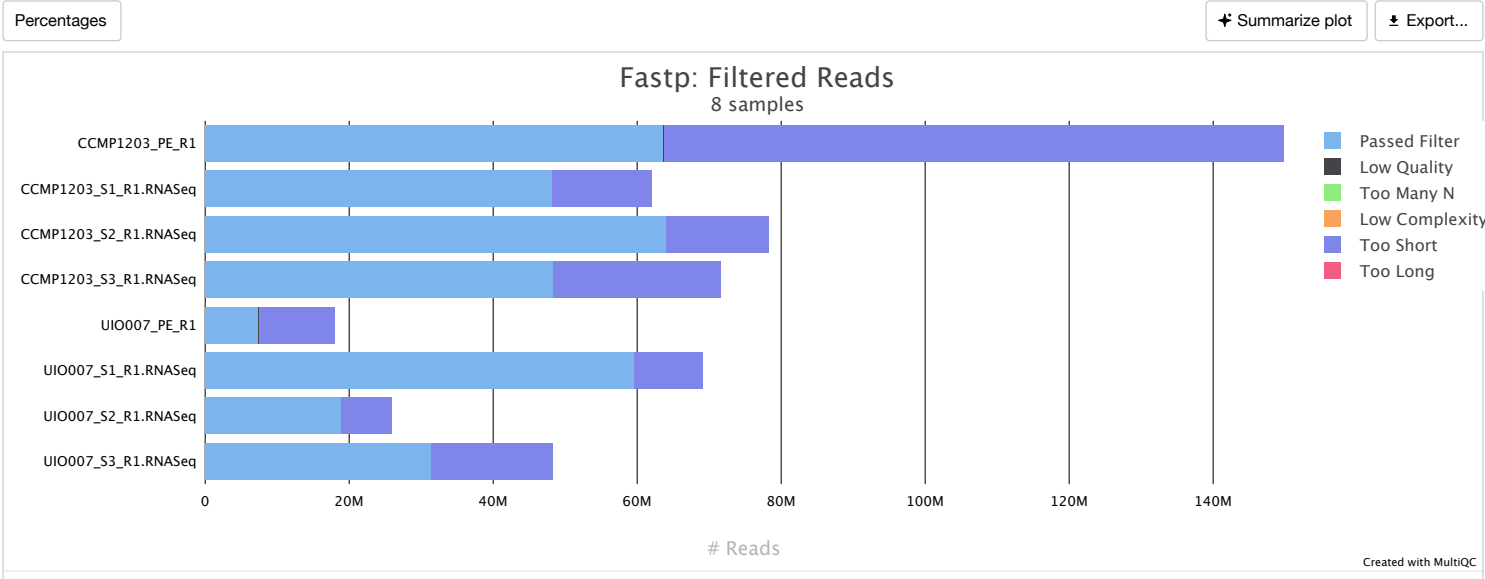
fastp Version: 0.23.4

All-in-one FASTQ preprocessor (QC, adapters, trimming, filtering, splitting...). URL: <https://github.com/OpenGene/fastp> DOI: 10.1093/bioinformatics/bty560

Fastp goes through fastq files in a folder and perform a series of quality control and filtering. Quality control and reporting are displayed both before and after filtering, allowing for a clear depiction of the consequences of the filtering process. Notably, the latter can be conducted on a variety of parameters including quality scores, length, as well as the presence of adapters, polyG, or polyX tailing.

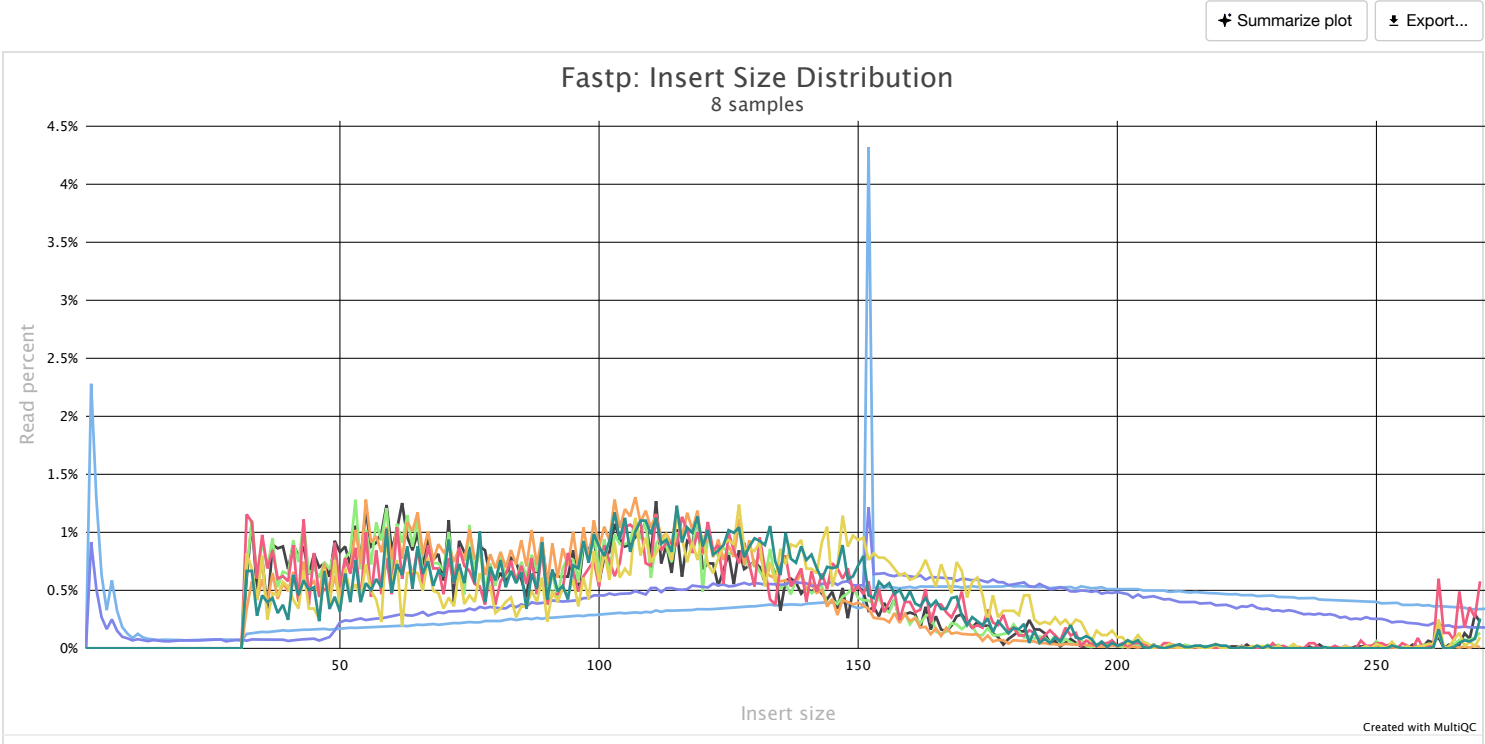
Filtered Reads

Filtering statistics of sampled reads.



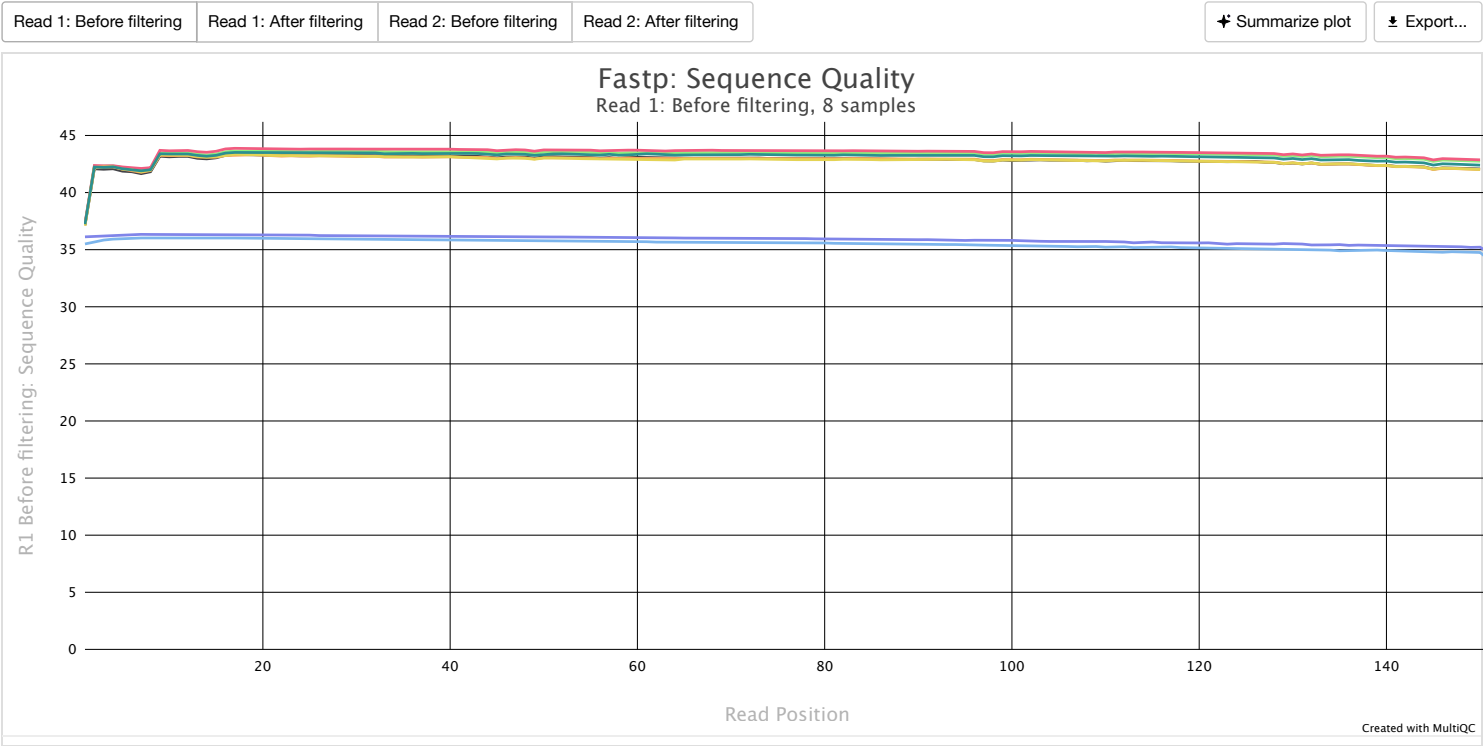
Insert Sizes

Insert size estimation of sampled reads.



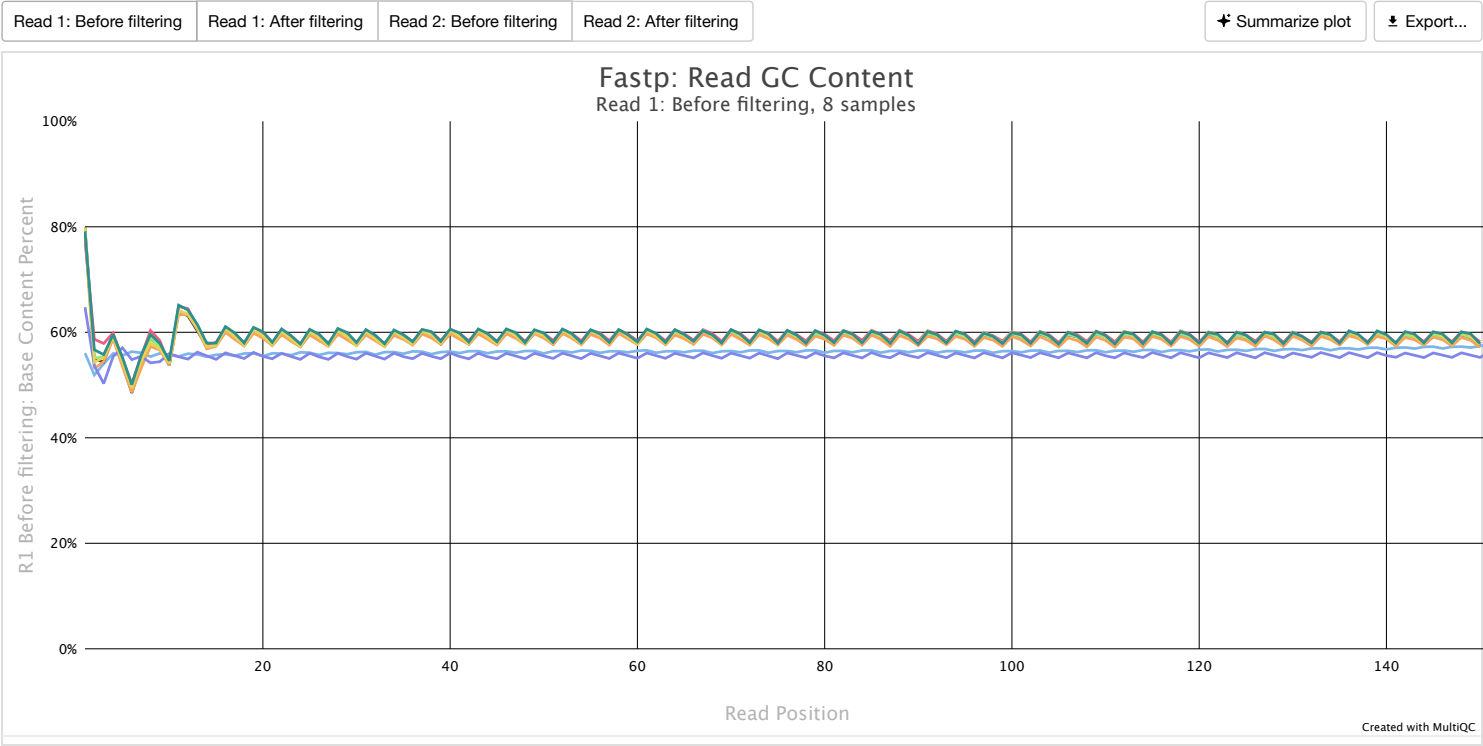
Sequence Quality

Average sequencing quality over each base of all reads.



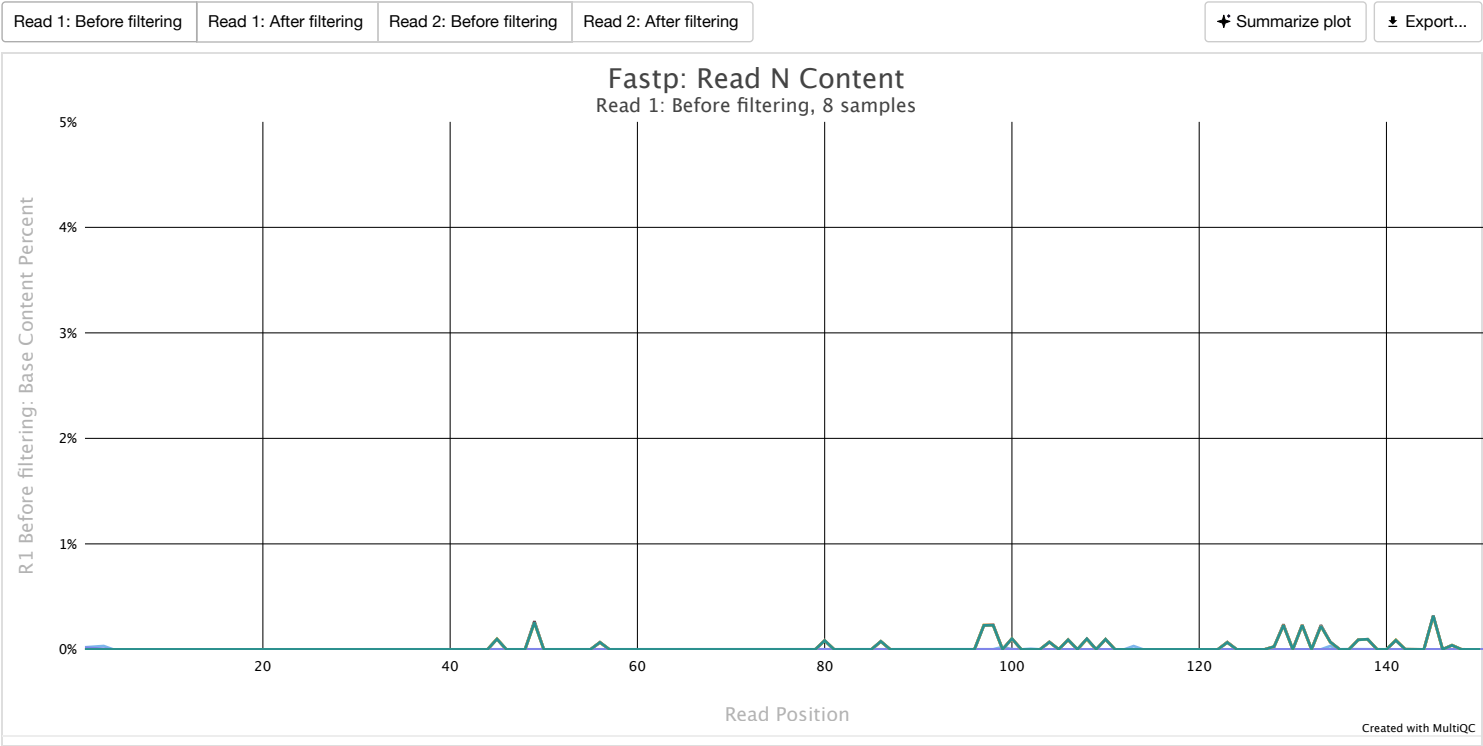
GC Content

Average GC content over each base of all reads.



N content

Average N content over each base of all reads.



FastQC

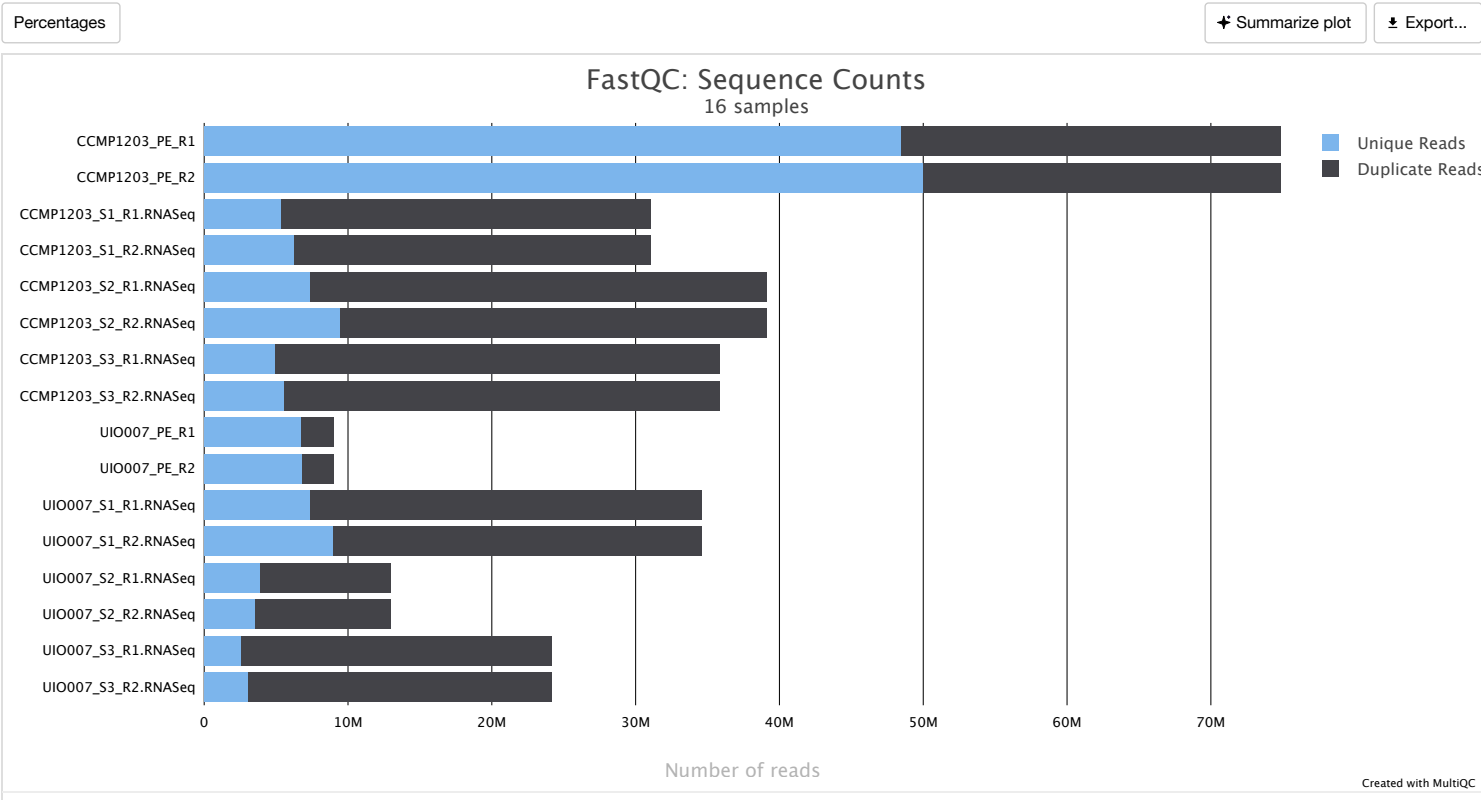
Version: 0.12.1

Quality control tool for high throughput sequencing data. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>

Sequence Counts

Help

Sequence counts for each sample. Duplicate read counts are an estimate only.



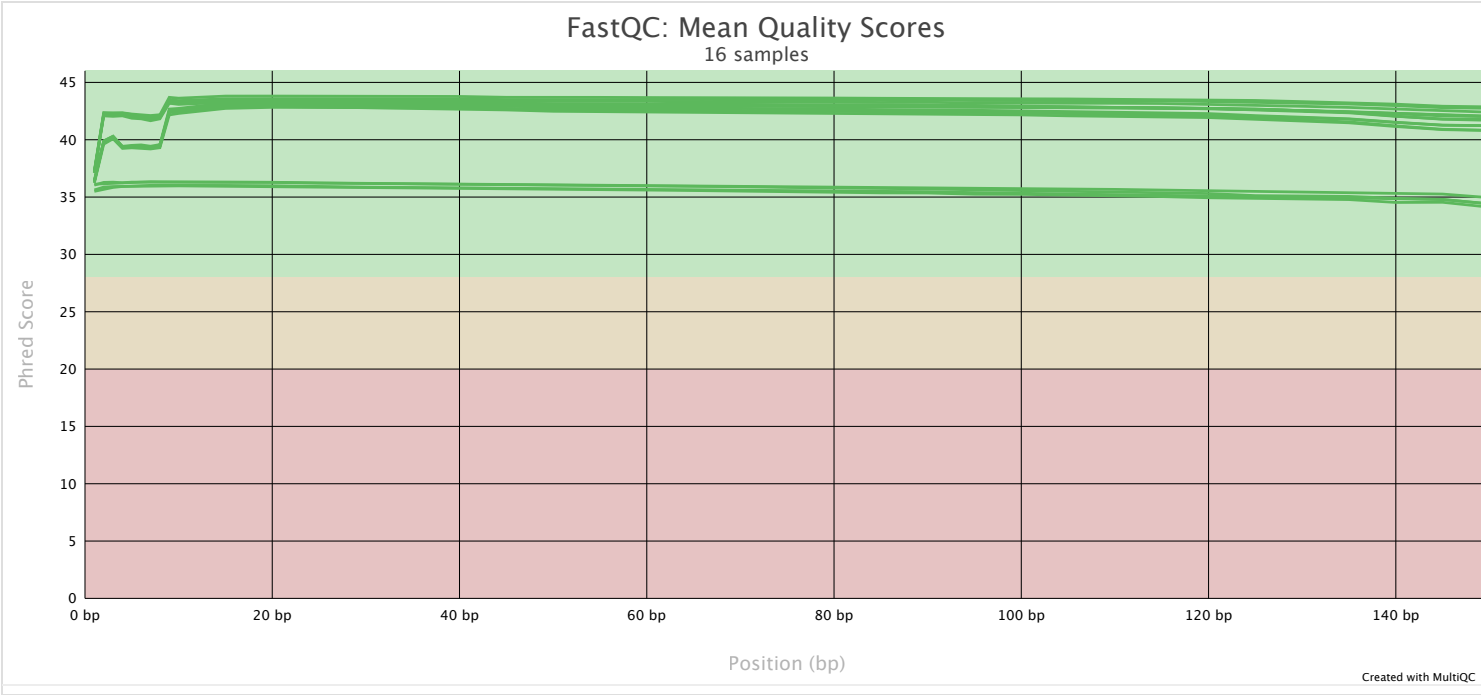
Sequence Quality Histograms

16

Help

The mean quality value across each base position in the read.

Summarize plot Export...



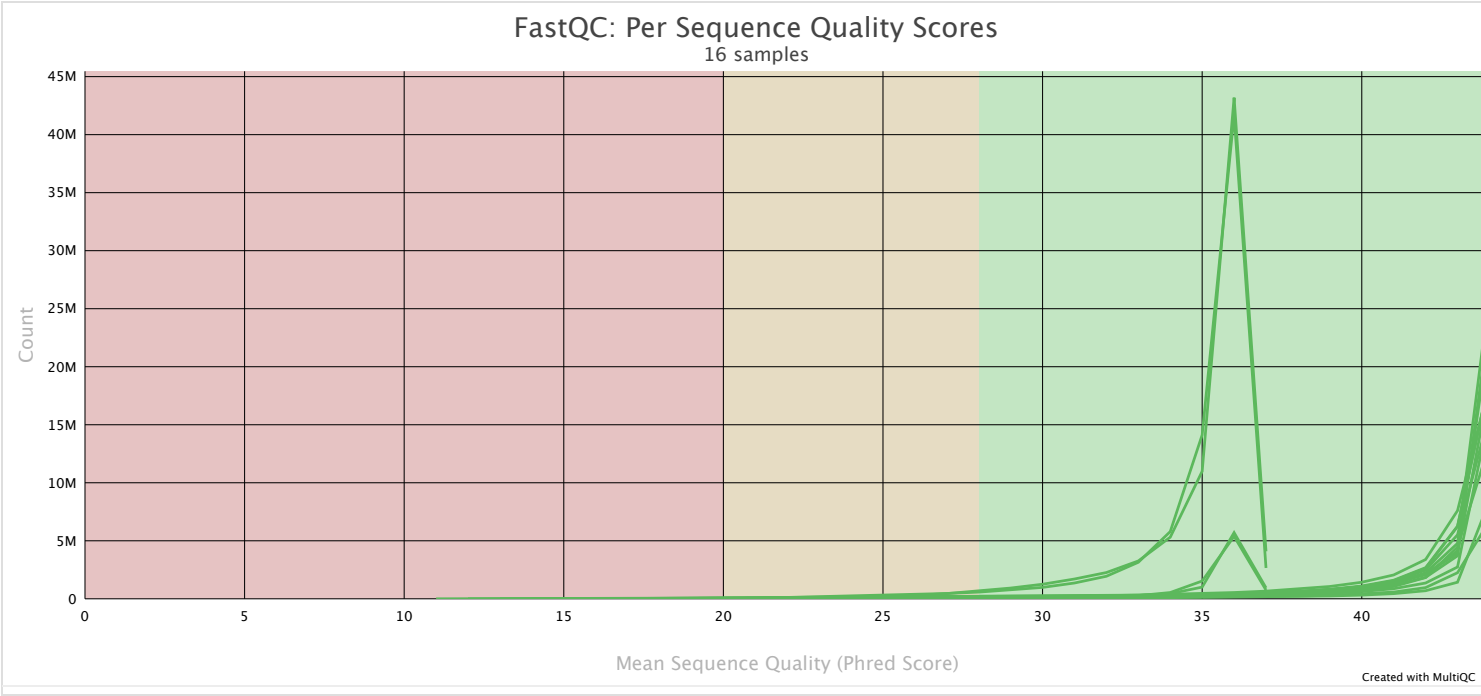
Per Sequence Quality Scores

16

Help

The number of reads with average quality scores. Shows if a subset of reads has poor quality.

Summarize plot Export...



Per Base Sequence Content 23

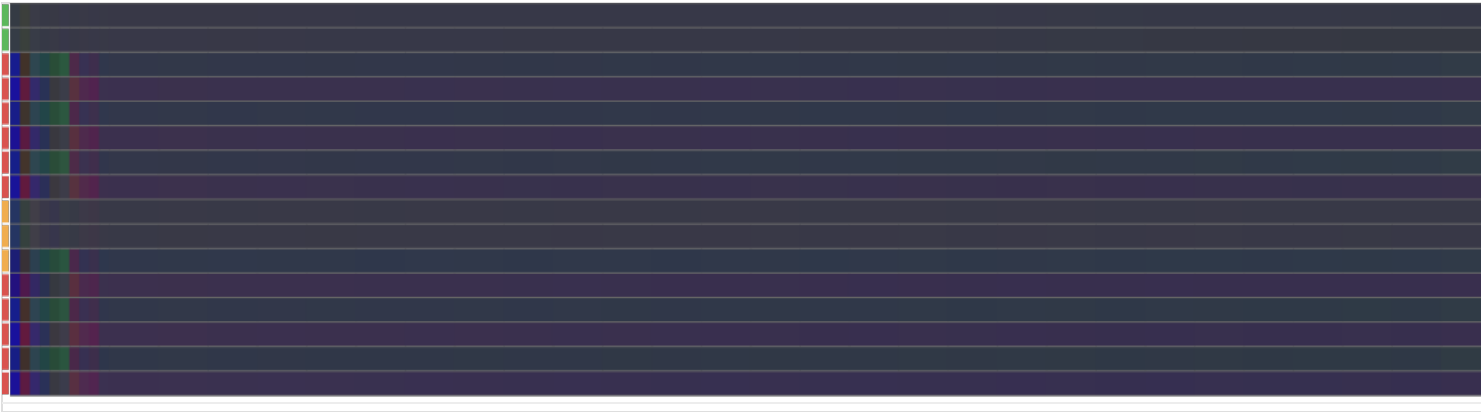
Help

The proportion of each base position for which each of the four normal DNA bases has been called.

Click a sample row to see a line plot for that dataset.

Rollover for sample name

Position: - %T: - %C: - %A: - %G: -



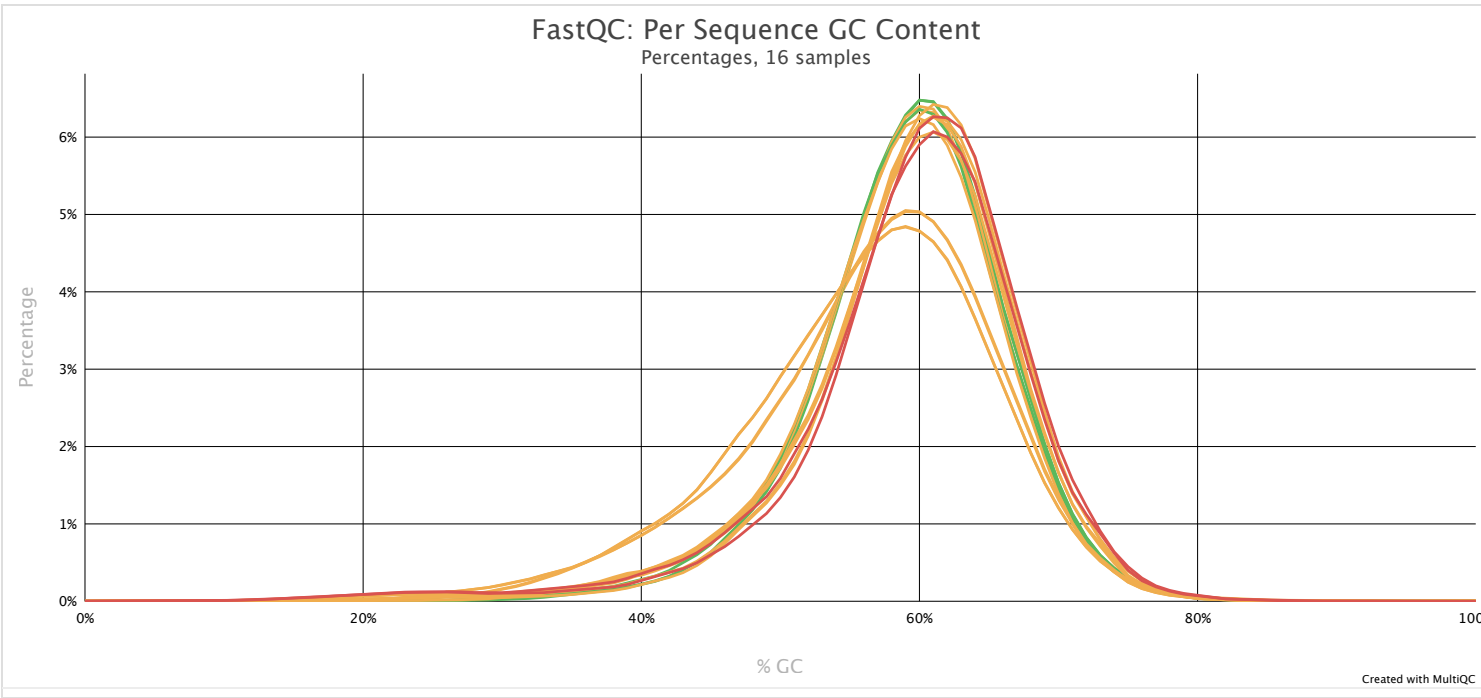
Per Sequence GC Content 311

Help

The average GC content of reads. Normal random library typically have a roughly normal distribution of GC content.

Percentages Counts

Summarize plot Export...

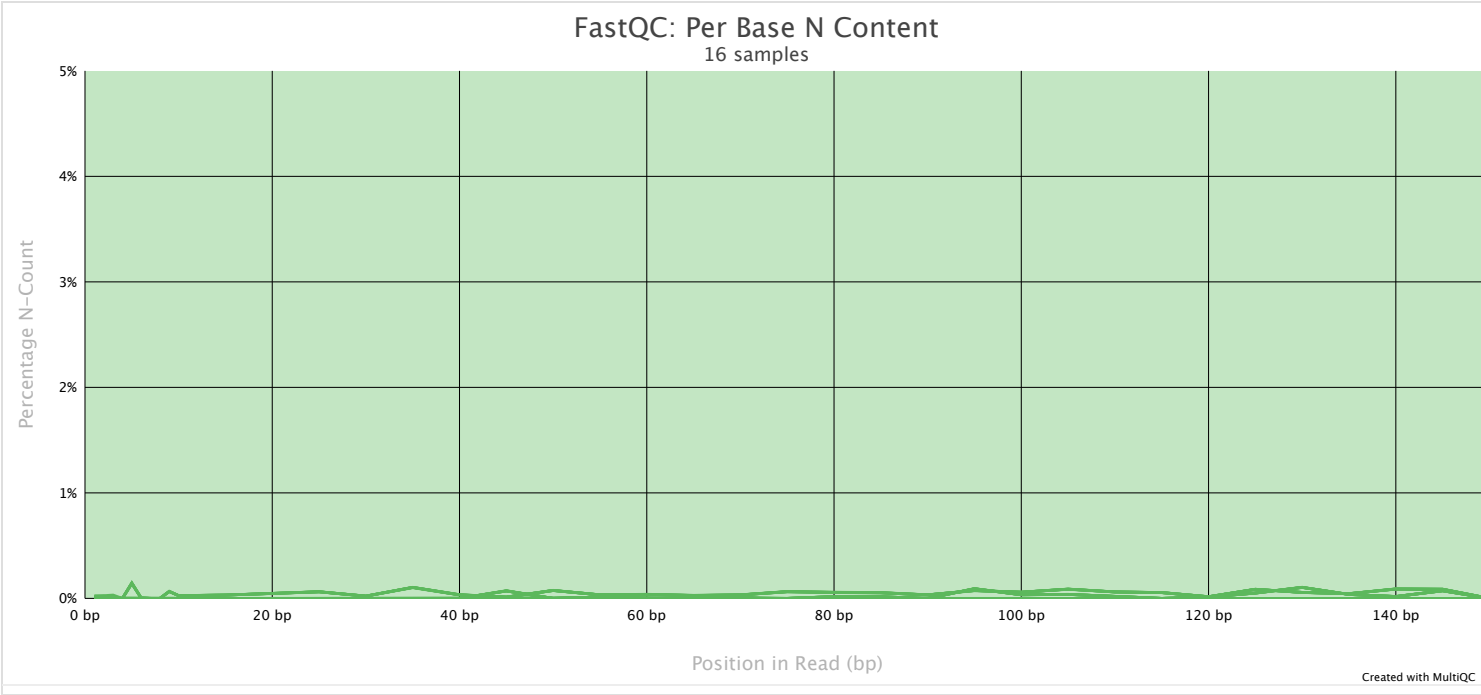


Per Base N Content 16

Help

The percentage of base calls at each position for which an N was called.

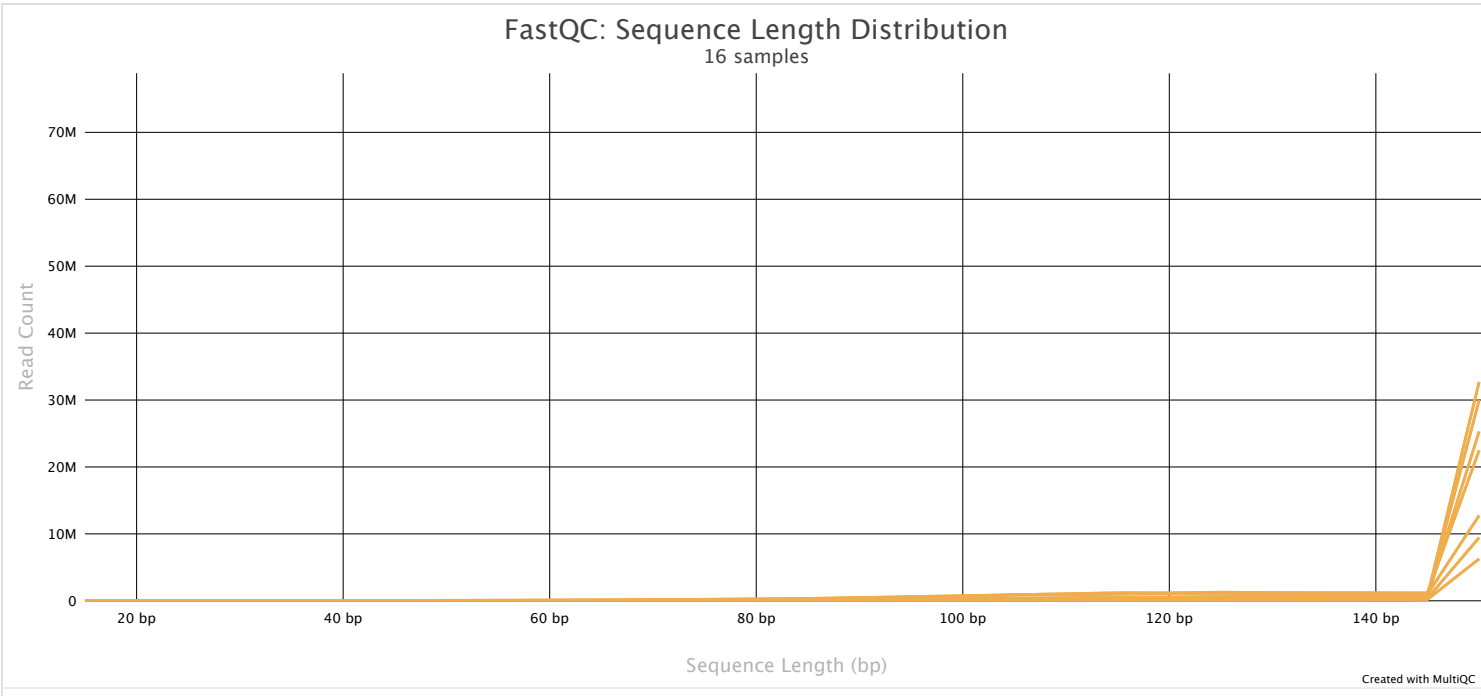
Summarize plot Export...



Sequence Length Distribution 2 14

The distribution of fragment sizes (read lengths) found. See the FastQC help

Summarize plot Export...

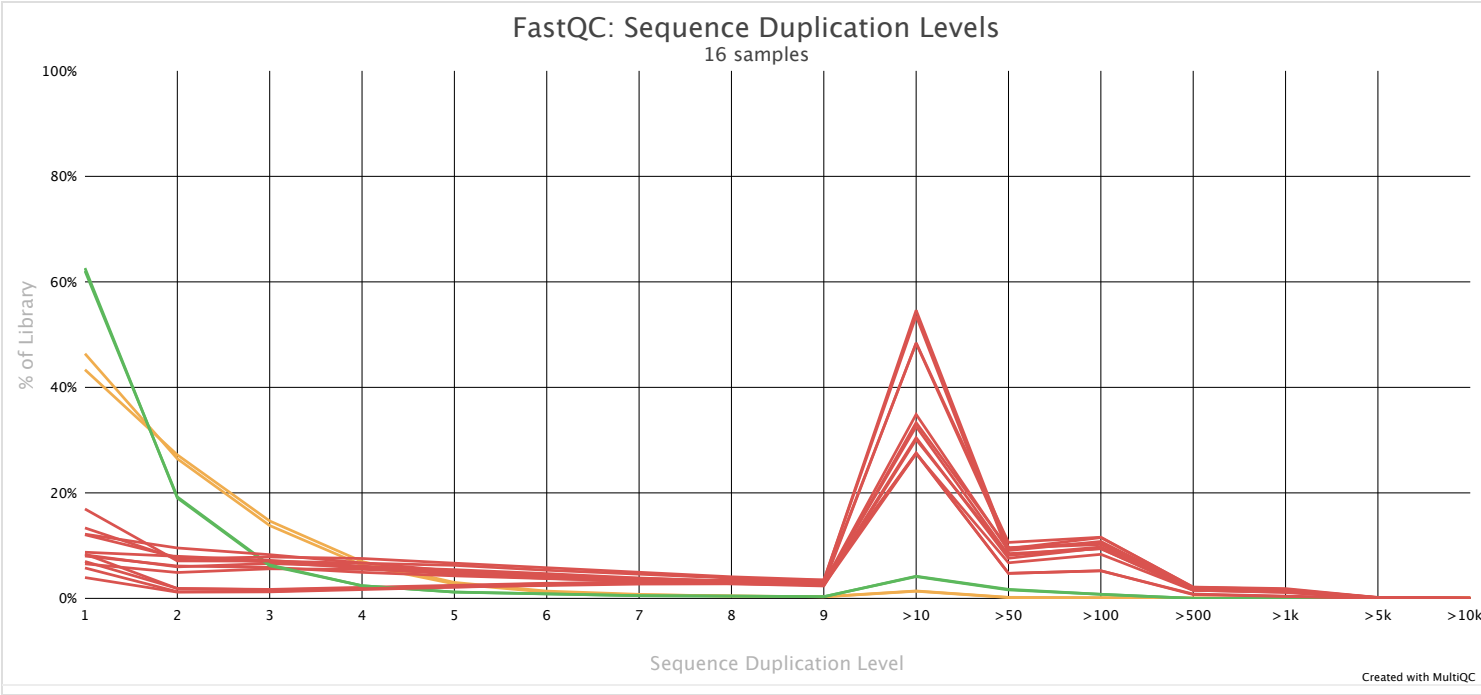


Sequence Duplication Levels 22

Help

The relative level of duplication found for every sequence.

Summarize plot Export...



Overrepresented sequences by sample 16

Help

The total amount of overrepresented sequences found in each library.

16 samples had less than 1% of reads made up of overrepresented sequences

Top overrepresented sequences

Top overrepresented sequences across all samples. The table shows 20 most overrepresented sequences across all samples, ranked by the number of samples they occur in.

Adapter Content

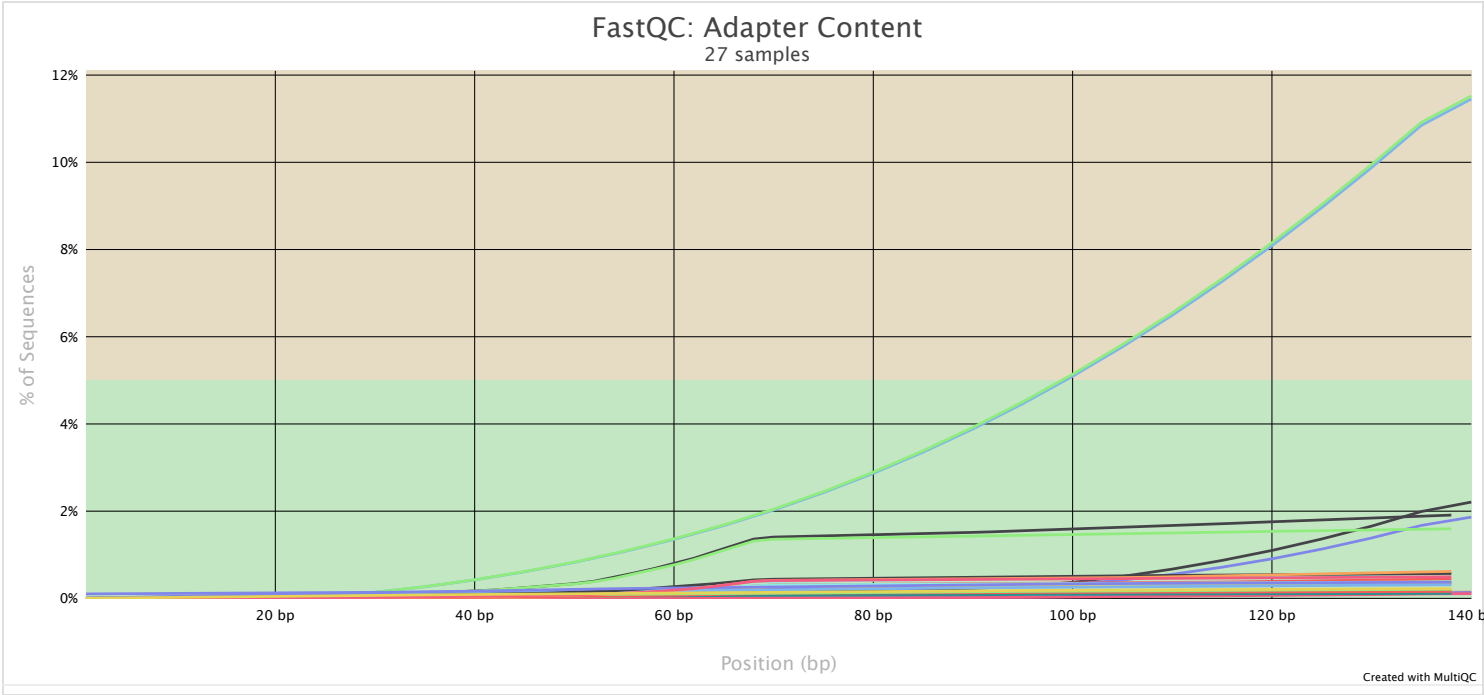
14

Help

The cumulative percentage count of the proportion of your library which has seen each of the adapter sequences at each position.

Summarize plot

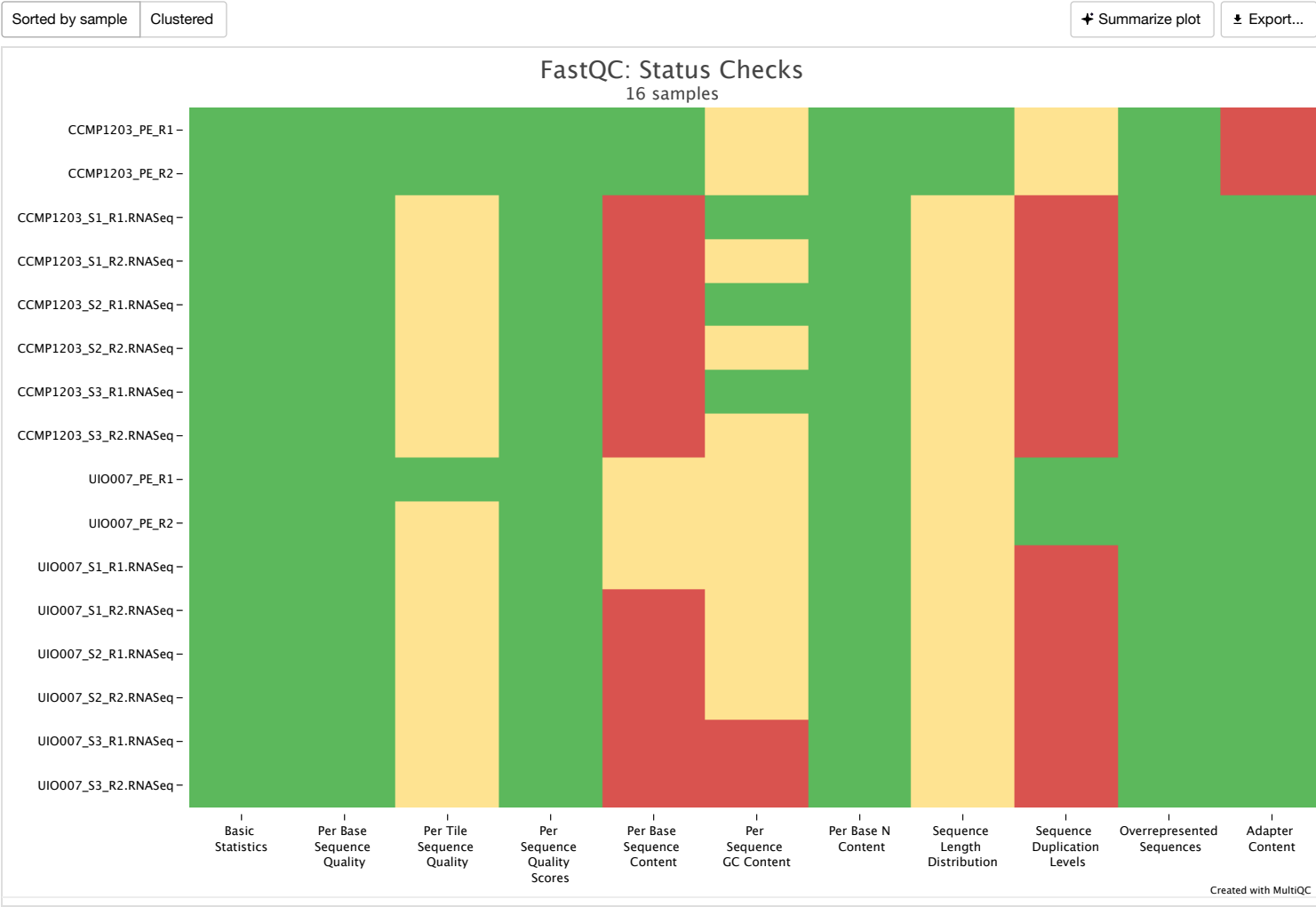
Export...



Status Checks

Help

Status for each FastQC section showing whether results seem entirely normal (green), slightly abnormal (orange) or very unusual (red).



Software Versions

Software Versions lists versions of software tools extracted from file contents.

Copy table

Software	Version
BUSCO	5.8.2
FastQC	0.12.1
fastp	0.23.4