Journal of Species Research 6(1):76-90, 2017

Plastid genome of *Aster altaicus* var. *uchiyamae* Kitam., an endanger species of Korean asterids

Jihye Park[1], Jaekyung Shim[1], Hyosig Won[2] and Jungho Lee[1,*]

[1]*Green Plant Institute, B-301, Heungdeok IT Valley, Giheung-gu, Yongin 16954, Republic of Korea*
[2]*Department of Biological Science, Daegu University, Gyeongsan, Gyeongbuk 38453, Republic of Korea*

**Correspondent: jlee@greenplant.re.kr*

Aster altaicus var. *uchiyamae* Kitam. is an endemic taxon of Korea and is protected by law as an endanger taxon. The genetic information of *A. altaicus* var. *uchiyamae* is unavailable in Genbank. Here we sequenced chloroplast genome of *A. altaicus* var. *uchiyamae*. The cp-genome of *Aster altaicus* var. *uchiyamae* was 152,446 bps in size: LSC was 84,240 bps, IR 25,005 bps, SSC 18,196 bps. The cp-genome contains 112 genes and 21 introns consisted of 79 protein coding genes (PCGs), 4 RNA genes, and 29 tRNA genes, with 20 group II introns and one group I intron. There were three pseudo-genes including *ψ-ycf1*, *ψ-rps19*, and *ψ-trnT_GGU*. Eighteen genes, five introns, and parts of two genes and an intron are found within the IR, which has two copies. The cp-DNA of *Aster altaicus* var. *uchiyamae* is distinguished from *A. spathulifolius*, only known cp-genome of the genus *Aster*, by 172 SNP in genic regions of 43 PCGs and 21 indels in 11 PCGs and *SSU*. The chloroplast genome sequence was deposited at GenBank (KX35265).

Keywords: *Aster altaicus*, Asteraceae, chloroplast genome, Compositae, indel, SNP, ψ-trnT_GGU

© 2017 National Institute of Biological Resources
DOI:10.12651/JSR.2017.6.1.076

Introduction

Aster altaicus is an Asian species, which widely occurs from China east to Korea and Mongolia, west to Afghanistan, Turkmenistan, Kazakhstan, Uzbekistan, and SW Asia (Iran) (Park, 2007; Chen *et al*., 2008), South to Nepal, Kashmir, Pakistan, and NW India, North to Russia (Siberia). In East Asia, seven varieties of the species occur including var. *altaicus*, var. *hirsutus*, var. *canescens*, var. *scaber*, var. *millefolius*, var. *taitoensis* and var. *uchiyamae*. *Aster altaicus* var. *taitoensis* Kitam only occurs in Taiwan as endemic taxon and *Aster altaicus* var. *uchiyamae* Kitam occurs as endemic in Korea. The other five varieties occur in mainland China, but they are not found in Taiwan and Korea.

Aster altaicus var. *uchiyamae* Kitam. was first collected by K. Uchiyama (Oct. 7, 1902) at Suanbo, Chungbuk province, and limitedly occurs in road side, sandy places of riverbank along the South Han River (Kitamura, 1937; personal observation). The *Aster altaicus* var. *uchiyamae* Kitam. is protected by Korean law as an endangered species and is included in Korean Red List of Threatened Species (Suh and Kim, 2014). The morphology of *Aster altaicus* var. *uchiyamae* have been reported compared with other 17 taxa of Aster s.l. on pappus (Chung and Kim, 1991a), receptacle (Chung and Kim, 1991b), ligulate florets (Chung and Kim, 1991c), captculum (Chung and Kim, 1993a), and anatomical characters (Chung and Kim, 1993b). Though the morphological information was available as listed above, the genetic information of *A. altaicus* var. *uchiyamae* is absent.

The genetic information on East Asian *Aster* is available by Ito *et al*. (1994; 1995; 1998) in 20 century, and that on Korea *Aster* is available by Hong *et al*. (2012). However, none of these included *A. altaicus* var. *uchiyamae* as research taxon. In *Aster*, plastid genomic information is very limited. The plastid genomic information is only available in *A. spathulifolius* occurring seaside in Korea and Japan. We characterized plastid genome of *A. altaicus* var. *uchiyamae* Kitam. in order to get the genetic information of the endangered species. Complete cp-genomic sequences have become more useful as genetic barcode of plants (Nock *et al*., 2011; Li *et al*., 2015). Here, we report the complete cp genome of *Aster altaicus* var. *uchiyamae* as genetic barcode.

Materials and Methods

Chloroplast DNA extraction, genome sequencing, assembly, and PCR-based validation

Aster altaicus var. *uchiyamae* was collected at the population from Jocheonri, Chungju city of ChungBuk province (N: 37°06′02″, E: 127°49′13″) in 2010, before dam was constructed. The photo-voucher of plant specimen was deposited at Daegu University Herbarium (Won6376-4). Fresh leaves of *A. altaicus* var. *uchiyamae* were collected from the population, and stored in liquid nitrogen until usage. Total DNA was extracted using the Qiagen DNeasy Plant Mini Kit (Qiagen, Germany), and DNA concentration and quality were determined using a Scandrop Nano-volume spectrophotometer (Analytik Jena, Germany). High quality DNA (concentration = 300 ng/μL, A260/280 ratio = 1.8-2.0, and A260/230 ratio = 1.7) was used for PCR and sequencing.

For Sanger sequencing, the whole cp-genome of *A. altaicus* var. *uchiyamae* was PCR-amplified in ~1-2 kb fragments, and cp-genome structure was verified using Long PCR, with ~5-28 kb fragments, as described by Lee and Manhart (2002a; 2002b). The detailed primer information is shown in Table S1 and Table S2. Only PCR products ranging from ~1-2 kb were sequenced using Bigdye (ver. 3.1) and ABI3730 at NICEM of Seoul National University. Assembled cp-sequences were verified using Sequencher ver. 5.0 (Gencode, USA).

Genome annotation, Genome comparison and Sequence Analysis

Protein coding and ribosomal RNA genes were annotated using DOGMA (http://dogma.ccbb.utexas.edu/; Wyman *et al*., 2004). The boundaries of each annotated gene were manually determined by comparison with orthologous genes from other known cp-genomes. Genes encoding tRNAs were first predicted using tRNAscan (http://lowelab.ucsc.edu/tRNAscan-SE; Lowe and Eddy, 1997) and ARAGORN, version 1.2 (http://130.235.46.10/ARAGORN/; Laslett and Canback, 2004), and were manually verified by predicting the tRNA secondary structure. Circular genome maps were drawn using GenomeVx (Conant and Wolfe, 2008), followed by manual modification. The sequencing data and gene annotation were submitted to GenBank with accession number KX35265. The mVISTA program in Shuffle-LAGAN mode (Fraser *et al*., 2004) was used to compare the cp-genome of *A. altaicus* var. *uchiyamae* with that of *A. spathulifolius* (Choi and Park, 2015).

Results and Discussion

The cp-genome of *Aster altaicus* var. *uchiyamae*, was determined (Fig. 1) and found to be 152,446 bp in length. It includes small and large single copy (SSC, LSC) regions of 18,196 bp and 84,240 bp, respectively, separated by a pair of 25,005 bp Inverted Repeats (IRs). A total of 112 genes were detected, including 79 protein coding genes, 29 tRNA genes, and four rRNA genes (Table 1, 2). The cp-genome of *A. altaicus* var. *uchiyamae* is slightly bigger than that of other asterids which are known to have *ψ-ycf1* and *ψ-rps19* in the borders of IR and Single copy regions. *A. altaicus* lacks *trnT_GGU* in its plastid genome but has *ψ-trnT_GGU* (Fig. 2). This cp-genome was also found to contain 21 different introns, including 20 group II introns and a group I intron with a cyanobacterial origin (Besendahl *et al*., 2000) found within the *trnL_uaa* gene. Three protein coding genes, including *clpP*, *rps12*, and *ycf3*, contain two group II introns (*clpP.i1*, *clpP.i2*, *rps12.i1*, *rps12.i2*, *ycf3.i1* and *ycf3.i2*), and 14 genes contain a single group II intron: *rpoC1.i*, *rpl2.i*, *rpl16.i*, *rps16.i*, *atpF.i*, *petB.i*, *petD.i*, *ndhA.i*, *ndhB.i*, *trnA_ugc.i*, *trnG_ucc.i*, *trnI_gau.i*, *trnK_uuu.i*, and *trnV_uac.i*. Among the 20 group II introns, the intron in *rps12*, between exons 1 and 2, is trans-splicing, while the other 19 group II introns are cis-splicing.

Eighteen genes, five introns, and parts of two genes and an intron are found within the IR, which has two copies. These 18 genes include seven protein-coding genes (*ndhB*, *rpl2*, *rps12*, *rpl23*, *rps7*, *rps19*, *ycf2*), all four rRNA genes (*16S*, *23S*, *4.5S*, *5S*), and seven tRNA genes (*trnA_ugc*, *trnI_cau*, *trnI_gau*, *trnL_caa*, *trnN_guu*, *trnR_acg*, *trnV_gac*). The five introns are *ndhB.i*, *rpl2.i*, *trnA_ugc.i*, *trnI_gau.i*, and *rps12.i2*. The IR also contains the 5′ end of *ycf1* at the border with the SSC, resulting in one intact *ycf1* and a 618-bp *ψ-ycf1* in the cp-genome. In addition, the IR contains parts of the *rps12* gene. This *rps12* gene consists of three exons, *rps12.e1*, *rps12.e2*, and *rps12.e3*; *rps12.e1* is in the LSC, but *rps12.e2* and *rps12.e3* are in the IR. Thus, the genome contains a single copy of *rps12.e1* but has two copies of *rps12.e2* and *rps12.e3*. A cis-splicing group II intron, *rps12.i2*, intervenes between *rps12.e2* and *rps12.e3*, but a trans-splicing intron, *rps12.i1t*, occurs between *rps12.e1* and *rps12.e2*. The *rps12.i1t* is split into two pieces, *rps12.i1t1* and *rps12.i1t2*, because the *rps12* gene is transcribed in two separate operons, the *clpP* operon (*clpP-rps12.e1-rps12.i1t1-rpl20*) and the *3′ rps12* operon (*rps12.i1t2-rps12.e2-rps12.i2-rps12.e3-rps7-ndhB*).

Direct comparison of chloroplast genomes of *A. altaicus* var. *uchiyamae* and *A. spathulifolius* (Choi and Park, 2015) using mVISTA program is shown in Fig. 3. Some of genes have length variation in 5′ end or 3′ end of the open reading frame. The 5′ end variation occurs in three protein coding genes (PCGs) including *atpB*, *ccsA* and *ycf4*. The 3′ end variation occurs in two PCGs



Fig. 1. Plastid genomic map of *Aster altaicus* var. *uchiyamae*.

including *rpl20* and *rpoC2*. There were two kinds of additional variations. One was Single Nucleotide Polymorphism (SNP, Table 3) and the other was indel (Table 4). SNPs were found at 172 sites in 43 protein coding genes.

The protein coding genes include 15 genetic system genes and 28 photosynthesis genes. As the genetic system genes, three conserved open reading frames (*ycf1*, *ycf2* and *ycf4*), maturase K gene (*matK*), four RNA polymerase genes (*rpoA*, *rpoB*, *rpoC1* and *rpoC2*), ribosomal protein small subunit genes (*rps3*, *rps8*, *rps12* and *rps16*), and three ribosomal protein large subunit gene (*rpl14*, *rpl20* and *rpl33*) are included. Among photosynthesis genes, seven genes of NADH dehydrogenase (*ndhA*, *ndhD*, *ndhF*, *ndhG*, *ndhH*, *ndhJ* and *ndhK*), three genes of ATP synthase (*atpA*, *atpB* and *atpE*), four photosystem I genes (*psaA*, *psaB*, *psaC* and *psaI*), five photosystem II genes (*psbA*, *psbB*, *psbC*, *psbD* and *psbE*), four cytochrome genes (*petA*, *petB*, *petD* and *petL*), a rubisco gene (*rbcL*), a membrane protein gene (*cemA*), cytochrome C biogenesis gene (*ccsA*), ATP-dependent Clp protease (*clpP*), and a Acetyl-CoA carboxylase gene (*accD*) were variable. There were no SNP in 29 tRNA and 4 rRNA genes.

Indels were found at 20 sites in eight PCGs including *accD*, *ccsA*, *matK*, *rpl22*, *rpoB*, *rpoC2*, *ycf1* and *ycf2* (Table 4), and a single bp indel was found in *SSU*. Among Indels, five Small Simple Repeats (SSRs) were in three protein coding genes including *matK*, *accD* and *ycf1*. Five SSRs include one 1-bp SSR, two 3-bp SSRs and two 6-bp SSRs.

Conclusion

This is the second report of chloroplast genome in the

Table 1. Gene list in chloroplast genome of *Aster altaicus* var. *uchiyamae*.

	Gene								
Genetic system genes									
Conserved orf	ycf1	ycf2 ×2	ycf3**	ycf4					
maturase K	matK								
RNA polymerase	rpoA	rpoB	rpoC1*	rpoC2					
Ribosomal protein									
Large subunits	rpl2* ×2	rpl14	rpl16*	rpl20	rpl22	rpl23 ×2	rpl32	rpl33	rpl36
Small subunits	rps2	rps3	rps4	rps7 ×2	rps8	rps11	rps12**$^{\alpha}$ ×2	rps14	rps15
	rps16*	rps18	rps19						
Photosynthesis genes									
Acetyl-CoA carboxylase	accD								
ATP-dependent Clp protease	clpP**								
ATP synthase	atpA	atpB	atpE	atpF*	atpH	atpI			
Cytochrome b	petB*								
Cytochrome b/f	petD*	petG	petL	petN					
Cytochrome f	petA								
Cytochrome C biogenesis	ccsA								
Membrane protein	cemA								
NADH dehydrogenase	ndhA*	ndhB* ×2	ndhC	ndhD	ndhE	ndhF	ndhG	ndhH	ndhI
	ndhJ	ndhK							
Photosystem I	psaA	psaB	psaJ	psaC	psaI				
Photosystem II	psbA	psbB	psbC	psbD	psbE	psbF	psbH	psbI	psbJ
	psbK	psbL	psbM	psbN	psbT	psbZ			
Rubisco	rbcL								
Translation initiation factor 1	InfA								
Ribosomal RNA	rrn16S ×2	rrn23S ×2	rrn4.5S ×2	rrn5S ×2					
Transfer RNA	trnA_UGC* ×2	trnC_GCA	trnD_GUC	trnE_UUC	trnF_ GAA	trnfM_CAU	trnG_GCC	trnG_UCC*	trnH_ GUG
	trnI_CAU ×2	trnI_GAU* ×2	trnK_UUU*	trnL_CAA ×2	trnL_ UAA*	trnL_UAG	trnM_CAU	trnN_GUU ×2	trnP_ UGG
	trnQ_UUG	trnR_ACG ×2	trnR_UCU	trnS_GCU	trnS_ GGA	trnS_UGA	trnT_UGU	trnV_GAC ×2	trnV_ UAC*
	trnW_CCA	trnY_GUA							
Pseudo gene	ψ-ycf1	ψ-rps19	ψ-trnT_GGU						

*intron; $^{\alpha}$cp genome contains a copy of *rps12* exon 1 in LSC and two copies of *rps12* exon 2 and 3 in IR

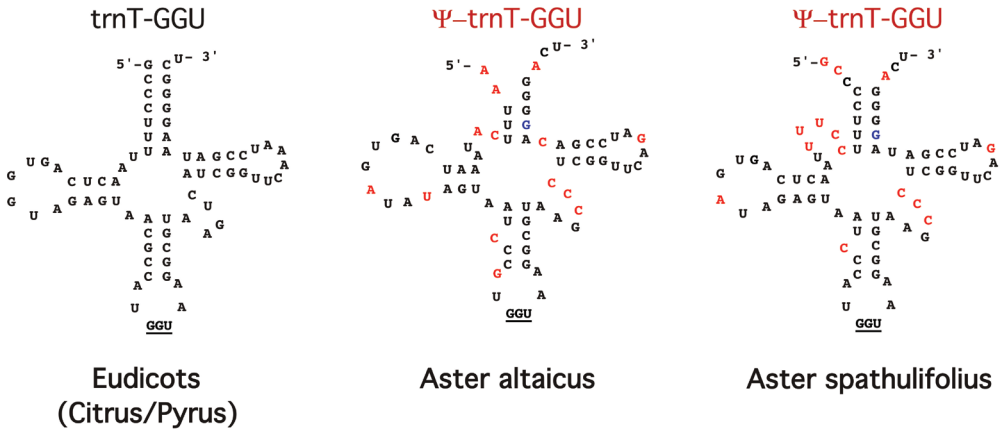

Fig. 2. The secondary structure of *ψ-trnT_GGU* of *Aster altaicus* var. *uchiyamae* and *A. spathulifolius* compared to that of normal *trnT_GGU* gene. Variable sequences were marked in red.

Table 2. Chloroplast genome of *Aster altaicus* var. *uchiyamae* Kitam. and their relatives.

Species	Genome size (bp)	LSC size (bp)	IR size (bp)	SSC size (bp)	ACC number	Reference
Aster altaicus var. *uchiyamae*	152,446	84,240	25,005	18,196	NCBI KX35265	This study
Aster spathulifolius	149,555	81,999	24,796	17,972	NCBI NC_027434	Choi and Park, 2015
Dendranthema sp. K247003	151,175	82,886	24,971	18,347	NABIC NG-0482	Park *et al*., 2015
Helianthus annuus	151,104	83,530	25,295	17,576	NCBI NC_007977	Timme *et al*., 2007

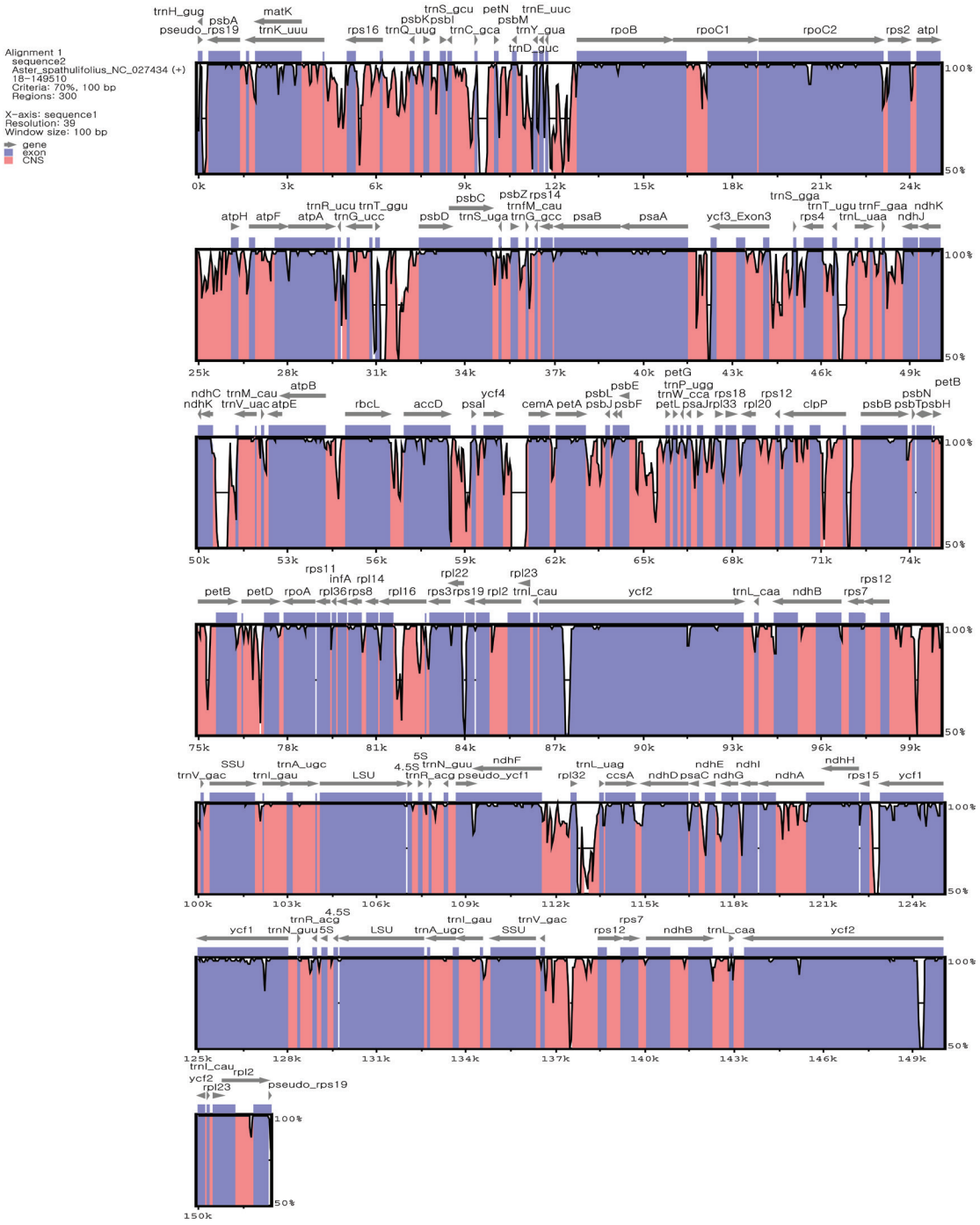

Fig. 3. Comparison of chloroplast genomes of *Aster altaicus* var. *uchiyamae* and *A. spathulifolius* using mVISTA program. Grey arrows and thick black lines above the alignment indicate genes with their orientation and the position of the IRs, respectively. The Y-scale represents the percent identity between 50-100%. Genome regions are color-coded: Coding regions in blue; noncoding sequences (CNS) in red.

Table 3. SNPs found in plastid protein coding genes of *Aster altaicus* var. *uchiyamae* and *A. spathulifolius*.

	Gene name	Size (*A. altaicus*/*A. spathulifolius*)	No of SNP		Gene name	Size (*A. altaicus*/*A. spathulifolius*)	No of SNP
1	psbA	1062	2	23	petL	96	2
2	matK	1515/1512	9	24	rpl33	201	1
3	rps16	267	1	25	rpl20	381/363 [β]	1
4	rpoB	3186/3183	5	26	rps12	372	1
5	rpoC1	2070	3	27	clpP	591	1
6	rpoC2	4146/4153 [β]	11	28	psbB	1527	3
7	atpA	1527	6	29	petB	648	3
8	psbD	1062	3	30	petD	483	2
9	psbC	1422	1	31	rpoA	1008	5
10	psaB	2205	4	32	rps8	405	1
11	psaA	2253	4	33	rpl14	369	1
12	ndhJ	477	1	34	rps3	657	1
13	ndhK	678	1	35	ycf2	6840/6651	8
14	atpE	402	1	36	ndhF	2226	9
15	atpB	1482/1380 [α]	5	37	ccsA	969/990 [α]	4
16	rbcL	1479	3	38	ndhD	1503	4
17	accD	1485/1497	7	39	psaC	246	2
18	psal	111	1	40	ndhG	531	1
19	ycf4	633/555 [α]	2	41	ndhA	1092	4
20	cemA	690	3	42	ndhH	1236	3
21	petA	963	3	43	ycf1	5085/5073	38
22	psbE	252	1				

α: 5′ end variation; β: 3′ end variation

Table 4. Indels found in plastid genes of *Aster altaicus* var. *uchiyamae* and *A. spathulifolius*.

	Gene name	Size (*A. altaicus*/*A. spathulifolius*)	INDEL #	Memo [*A. altaicus*/*A. spathulifolius*]
1	matK	1515/1512	3	9-bp[(ATTTGATTC)/−], 12-bp[−/(GGAGCATCTTGC)], 6-bp SSR[$(TCAAAA)^2$/(TCAAAA)]
2	rpoB	3186/3183	1	3-bp[(GTC)/−]
3	rpoC2	4146/4153	1	9-bp[(TCAGAAATC)/−]
4	accD	1485/1497	3	3-bp[−/(AAC)], 6-bp[−/(CGAGAA)], 3-bp SSR[$(CTT)/(CTT)^2$]
5	rpl22	462/465	2	9-bp[(ACCTAAACG)/−], 12-bp[−/(TTAAATAAGAGA)]
6	ycf2	6840/6651	2	9-bp[(CGATATTGA)/−], 180-bp
7	ccsA	969/990	1	9-bp[(ATAATGAAA)/−]
8	ycf1	5085/5073	7	6-bp[(AACTTT)/−], 9-bp[−/(CAATGCCAT)], 9-bp−/(TCAAAGTAA)], 21-bp[(TGGAAGAGGAAGAAGAAGATA)/−], 1-bp SSR[$(A)^6/(A)^3$], 3-bp SSR[$(GAA)^5/(GAA)^7$], 6-bp SSR[$(TCAAGA)^2$/(TCAAGA)]
9	SSU	1490/1491	1	1-bp[−/C]

genus *Aster*, Asteraceae, and the first reported plastid genome confirmed by Sanger method in *Aster*. As genetic barcode of *A. altaicus* var. *uchiyamae*, an endanger species, 152,446 bps of chloroplast genomic sequence was registered to NCBI (KX35265). The chloroplast genome is distinguished from that of *A. spathulifolius*, by 127 SNPs and 21 Indels in coding regions. The high variation of coding genes in *Aster* suggests that, as suggested by Li *et al*. (2015), not plastid IGS region but chloroplast genomic information in coding region is useful for genetic barcode in *Aster*. As suggested by Dong *et al*. (2015) in land plants, *ycf1* would be useful for plant identification as having 38 SNPs and 7 Indels in the comparison of *A. altaicus* var. *uchiyamae* and *A. spathulifolius*, which belong to *Aster* sensu stricto. Loss of *atpB* in *A. spathulifolius*, reported by (Choi and Park, 2015), and the presence of *ψ-trnT_GGU* in *Aster altaicus* var. *uchiyamae* and *A. spathulifolius* in this study have enough to draw our interests in addition to high variation of coding genes between *Aster* plastids.

Acknowledgements

This work was supported by a grant from the National Institute of Biological Resources under the Ministry of Environment, Republic of Korea.

References

Besendahl, A., Y.L. Qiu, J. Lee, J.D. Palmer and D. Bhattacharya. 2000. The endosymbiotic origin and vertical evolution of the plastid tRNA Leu group I intron. Curr. Genet. 37:12-23.

Chen, Y., L. Brouillet and J.C. Semple. 2008. *Aster* Linnaeus. Flora of China. In eFloras (2008). Published on the Internet http://www.efloras.org

Choi, K.S. and S. Park. 2015. The complete chloroplast genome sequence of *Aster spathulifolius* (Asteraceae); genomic features and relationship with Asteraceae. Gene 572:214-221.

Chung, G.Y. and Y.S. Kim. 1991a. A taxonomic study of pappus morphology on the genus Aster. L., and its allied groups in Korea. Kor. J. Plant Tax. 21(3):197-209.

Chung, G.Y. and Y.S. Kim. 1991b. A taxonomic study of the receptacle morphology on the genus Aster. L., and its allied groups in Korea. Kor. J. Plant Tax. 21(4):229-238.

Chung, G.Y. and Y.S. Kim. 1991c. A study on epidermal cells of the ligulate florets of the genus Aster. L., and its allied groups in Korea. Kor. J. Plant Tax. 21(4):239-250.

Chung, G.Y. and Y.S. Kim. 1993a. A study on the capitulum morphology of the genus Aster. L., and its allied taxa in Korea. Kor. J. Plant Tax. 23(3):105-118.

Chung, G.Y. and Y.S. Kim. 1993b. A taxonomic study of Korean Aster. L., and its allied taxa on the anatomical characters. Kor. J. Plant Tax. 23(3):119-130.

Conant, G.C. and K.H. Wolfe. 2008. GenomeVx: simple web-based creation of editable circular chromosome maps. Bioinformatics 24:861-862.

Dong, W., C. Xu, C. Li, J. Sun, Y. Zuo, S. Shi, T. Chang, J. Guo and S. Zhou. 2015. ycf1, the most promising plastid DNA barcode of land plants. Sientific Report 5:8348.

Fraser, K.A., L. Pachter, A. Poliakov, E.M. Rubin and I. Dubchak. 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Research 32:W273-W279.

Hong, S.-Y., K.-S. Cho and K.-O. Yoo. 2012. Phylogenetic Analysis of Korean Native *Aster* Plants Based on Internal Transcribed Spacer (ITS) Sequences. Kor. J. Hort. Sci. Technol. 30(2):178-184.

Ito, M., A. Soejima and T. Nishino. 1994. Phylogeny and speciation of Asian Aster. Kor. J. Plant Tax. 24(3):133-143.

Ito, M., A. Soejima, M. Hasebe and K. Watanabe. 1995. A Chloroplast-DNA Phylogeny of *Kalimeris* and *Aster*, with Reference to the Generic Circumscription. J. Plant Res. 108:93-96.

Ito, M., A. Soejima and K. Watanabe. 1998. Phylogenetic Relationships of Japanese *Aster* (Asteraceae, Astereae) sensu lato Based on Chloroplast-DNA Restriction Site Mutations. J. Plant Res. 111:217-223.

Kitamura, S. 1937. *Aster altaicus* Willd., var. *Uchiyamae* Kitamura var. nov. Memoirs of the College of Science, Kyoto Imperial University, Series B. pp. 367-368.

Laslett, D. and B. Canback. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 32:11-16.

Lee, J. and J.R. Manhart. 2002a. Four embryophyte introns and *psbB* operon explains *Chlorokybus* as a basal lineage of streptophytes. Algae 17(1):53-58.

Lee, J. and J.R. Manhart. 2002b. The Chloroplast *rpl23* gene cluster of *Spirogyra maxima* (Charophyceae) shares many similarities with the angiosperm *rpl23* operon. Algae 17(1):59-68.

Lee, M., J. Park, H. Lee, S.-H. Sohn and J. Lee. 2015. Complete chloroplast genomic sequence of *Citrus platymamma* determined by combined analysis of Sanger and NGS data. Hortic. Environ. Biotechnol. 56(5):704-711.

Li, X., Y. Yang, R.J. Henry, M. Rosseto, Y. Wang and S. Chen. 2015. Plant DNA barcoding: from gene to genome. Biol. Rev. 90:157-166.

Lowe, T.M. and S.R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955-964.

Park, C.-W. 2007. The Genera of Vascular Plants of Korea. Academy Publishing Co., Seoul.

Park, J.H. and J. Lee. 2016. The complete plastid genome of *Scopolia parviflora* (Dunn.) Nakai (Solanaceae). Korean J. Pl. Taxon. 46(1):70-74.

Park, S.K., S.-J. Kwon, J. Park, M. Lee, S.Y. Won, Y.C. Kim, Y.-J. Hwang, S.-H. Sohn and J. Lee. 2015. Chloroplast genome of white wild chrysanthemum, *Dendranthema* sp. K247003, as genetic barcode. Journal of Species Research 4(2):152-158.

Suh, Y.B. and Y.J. Kim. 2014. Korean Red List of Threatened Species. 2nd ed. National Institute of Biological Resources, Korea.

Timme, R.E., V.K. Jennifer, L.B. Jefferey and K.J. Robert. 2007. A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: Identification of divergent regions and categorization of shared repeats. American Journal of Botany 94(3):302-312.

Wyman, S.K., R.K. Jansen and J.L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20:3252-3255.

Submitted: July 12, 2016
Revised: September 30, 2016
Accepted: February 7, 2017

Table S1. Short primer list used in this study.

Primer	Sequence	Primer	Sequence	Primer	Sequence
An_S01_00	CTGGCTCAGGATGAACGC	S01_NW0R	GCTCGAACTGATGACTTCC	AP02A_13RF	GCTAAACCTCTCTTAATGTC
An_S01_01	TTTCCAGYTGTTGTYCCCCTCCC	S01RA	CGCAGTAGAGTTTAACCACC	AP02A_F	CTTTGAAAAGCCCACATACG
An_S01_02	GCTAACGCGTTAAGTATCCCGCC	1YF_04F20	GTTTGAAAAACCTCTTYTGACC	AP02A_F1	GCAGTTCCAAAAAAACGTAC
An_S01_03	TGACGACAGCCATGCACCACCTG	1YF_04FAp	CTTCTTCCTGATACTGATTACC	AP02A_R	TCTACCGATTTCACCATAGC
An_S01_04	CTTGTACACACCGCCCGTCACAC	1YF_04FF20	GTCTCCTTCTTCCTGATACTG	AP02A_R1	TGACAACGCTTTTAAGGCTG
An_S01_05	CTTGTTACGACTTCACTCCAGTC	1YF_06GF	GGTCATACAAACTTATTACCG	AP02A_R2	GAACAGACATATCAATATGC
An_S01_06	AGCTCAGYGGTAGAGCGCGCCCC	1YF_06GFR	CGGTAATAAGTTTGTATGACC	AP02B_01	GTTGAATTTCAAGTAGTGAG
An_S01_07	GTCTCTGGTTCAAGTCCAGGATG	1YF_12RAp	CTAGACGTCGTAGAGAATTGAG	AP02B_11	GTATGTGTTCGTGTATGTCC
An_S01_08	CTCGAACCGCTGACATCCGCCAC	1YF_14F23	CTAGACCATCTAATCTGAG	AP02B_11F	ACTTTCTACATATGATCGTC
An_S01_09	CTCTACCAACTGAGCTATATCCC	1YF_14FAp	GATRATGACCTCTTAACCAG	AP02B_11FN	RAYGATTACACGACTCGAAC
An_S01_10	GGATACCTAGGCACCCAGAGACG	1YF_22GR	CCAGTTTTTCCATTGCTTCG	AP02B_11FN	GTTCGAGTCGTGTAATCRTY
An_S01_11	GCTCCCATTTCGCTCGCCGCTAC	1YF_An20_G1F	CTACTACCAGAAATTCAATGCG	AP02B_11N	CACATGCAAGTCGTTTACCTG
An_S01_12	CCACTCGAACCCAGAGCTAGCTG	1YF_An20_G1R	CCGTTTCATCGATTTTGTCCG	AP02B_14F	GTAATACCACTCCCAGACCTC
An_S01_13	CTCTCGACGATGAAGCTTATCCC	1YF_Ap4790R	CCATCGTCTATTATCCATTG	AP02B_14N	TCCCCAACGGTCCAAAGATG
An_S01_14	GAGAATCCAATGCCCCGAAAACC	1YF_AST_F	KGTGTCTACTATTTAGGCAG	AP02B_25F	ATCATTCCCTTTTCTGACG
An_S01_15	CCTACCYACCTGTGTYGGTTTCG	1YF_AST_F1	GAAGAGATCCGAGTGAATG	AP02C_2000F	CGTCGAGAAYTCATCAAAGG
An_S01_16	GCGGACTACCTGCACCTGGACAG	1YF_AST_F2	CKTGCCTAAAATATATGATCC	AP02C_2000R	CCTTTGATGARTTCTCGACG
An_S01_17	CCGAGGAAACCTTTGCDCGCCTC	1YF_AST_F2N23	CTTGCCTAAAATGTATGATCC	AP02C_3220F	AATAAAAAGAAGAAGCACTCG
An_S01_18	CCCACCCCAAGATGAGTGCTCTC	1YF_AST_F2R23	GGATCATACATTTTAGGCAAG	AP02C_3220R	GATTAATCGAGTGCTTCTTC
An_S01_19	GCTGCATACATCRCTGCACTTCC	1YF_AST_F8	GGCCTAATTATCGATTAGAAG	AP02C_3710F	GGAAAARCTGGTTAAGAGGTC
An_S01_20	GTGCTCTAATCCTCTGAGC	1YF_AST_R	CTTCTAATCGATAATTAGGCC	AP02C_4200R	TTMTATTACTGTTATAATCC
IR_S01F1	GATTTTCAGTCCTCTGCTCTACC	1YF_AST_R2	GKAMTTTGCTTGGTATAATCC	C97_S2C0F	CAAGTACAAGTATGAGTAGC
S01_00R	AGCGTTCATCCTGAGCCAG	1YF_AST_R3	GATAATTTGTAAAATACATATGC	C97_S2C2F	CTAGACGTCGTAGAGAATTGAG
S01_01N	GGTATTAGCAGCCGTTTCCAG	1YF_AST_R4	TTGAACACCRTTATARGTAC	C97_S2C2R	GGTCTCATCTTATTCCGCATC
S01_01NR	CTGGAAACGGYTGCTAATACC	1YF_AST_R4N20	CTCTTCTTCTTCCCTTTGTCTC	CYM02_G1F	CCGAATCAATCCATATGCTC
S01_03N	GACGACAGCCATGCACCACC	1YF_AST_R4N23	TTGAACACCATTATAGGTGC	CYM02A_06RF	CTCATATTTCYTTTTCGATC
S01_04N	GCCTTGTACACACCGCCCGTC	1YF_AST_R6	CGTAATGGBYGATTCCATCG	CYM02A_F	AGAATGAATCAAATCTCCCC
S01_05FC17	CTTGCGAACGTACTCCCCAG	1YF_CYM137F	CTCTTATCTCTTCCTTCTCCG	CYM02A_R	GCGTGTCTACCAATTTCACC
S01_06R	GCGCGCTCTACCRCTGAGC	1YF_CYM551R	GAATAATACCAATCCAACCC	CYM02B_16F	CGTTCATGAACATCACAGC
S01_07F18T	GCCTCTATGGTAGAATCAGTCG	1YF_IR5523F	GCTTGTATGAATCGGTATTGG	CYM02B_27F	CATTAGCCCATACTGCTCCAG
S01_07N	GTGGTAGAGCGCGCCCCTG	1YF_LY_1370F	CTTGCCAAAAATGTATGATCC	Iris_MK02R	CTGTGTCTACTATTTAGGC
S01_07NF18	GAGTTGTATGCTGCGTTCGG	1YF_MF1N06	TTACCTTCATTGATAATAGTC	PH02A_03R	CGATATGAACATCATGAGCTG
S01_07R	CATCCTGGACTTGAACCAGAGAC	1YF_MF1N22	CCTTCATTGATAATAGTTAAAAAC	PH02A_06RF1R	TTCCTAATAGATTGATCCC
S01_07R18	GAACCAGAGACCTCGCCCGTG	1YF_MF4	GCMTATGTATTTTACAAATTATCAC	PH02A_F	TCCCTCCTCGCCCACAACCG
S01_07R18F	GTACGATCCCTCCGTCACCC	1YF_MF4R22	GATAATWTGTAAAATACATATGC	PH02A_F1	AGAATGAATCWAATCTCCCC
S01_08R	GTGGCGGATGTCAGCGGTTCGAG	1YF03	TATAAACGATGGAATCGHC	PH02A_N01F	TAACAGCCAACCGCTCTACC
S01_10FE	GAGAGTCCAGTAGCCGAAAGC	1YF03AL	GATTATAAACGATGGAACCGC	PH02A_R1	CATCTCTTATGTCAATGACAC
01_10N09F	GGTTCGAACAGGAGGAGTACGC	1YF11B	GSCCTAATTATCGATTAGAAG	S01_18F06	CTCTCCCACTTCACACCTCG
S01_10R18	GCTTATTTTCAACTCCCCGAAGC	4CYM02B_11	GTATGTGTTCGCGTATGTCC	S01RAR	GGTGGTTAAACTCTACTGCG
S01_14N	GTGGACGAAGCGGAAGCGAG	An20_S2CF0	CKGTGTCTACTATTTAGGCAG	S02A_00	GAGCGGTYGGCTGTTAAC
S01_15N	GCCGAGTTCCTTAGAGAGAG	An22_S02CF	GATTCAAAAGAACCAAAAGG	S02A_00F06	CTCTACTATTCTGCCCCAGCC
S01_16R	CTGTCCAGGTGCAGGTAGTCCGC	An22_S02CF1	GTGTACGAATCCGGCGAAAC	S02A_01	CGACCAATCAGTTAACAGC
S01_17N	GTCTTGCAGTCAAGCTCCC	An22_S02CR	AAGTGCACTTTTTAGAGTCG	S02A_02	AATTCGGTCGTTDTGGTCGG
S01_20R	CAGAGGATTAGAGCACGTGG	An23_116047F	GCATATTGATATGTATGTTCC	S02A_03	GTGACCAATTAWCCAACC
S01_NW0	GTTCGAGCCTGATTATCCC	An23_118153R	GCGTGTCTACCGATTTCACC	S02A_03A	GTGMCCAATTAACCARCC

Table S1. Continued.

Primer	Sequence	Primer	Sequence	Primer	Sequence
S02A_03N02	CAGCTCATGATGTTCATATCG	S02A_MK12D	GAACACATTCCKACTAATTCCC	S02B_23FR02	GAATAGCTATCCCAATGCC
S02A_03R02N	ACCTCATACAATAACTGTCC	S02A_MK12DE	GCATATTGATATGTCTGTTCC	S02B_25	CTKACGTGCATYATTAACC
S02A_04	GGTTGGTTAATTGGKCAC	S02A_N01	CCCCAAGTAGGATTCGAACC	S02B_25FC97	GAGATGCTCTAGCACTGCTTCC
S02A_04F06	GTTTGAAAAACCTCTTCTGACC	S02A_N22R17	GTATCCATAGAACATCTGACG	S02B_25N11	CTTACGTGCATCATTAACCAC
S02A_04I	GGTATTATTCTGGATACGGC	S02B_01	CTTCACATTTGGAAYTGCAC	S02B_25N97	CTKACRTGCATCATTAACCAC
S02A_06	CACTGCTTCCTAAGAACAG	S02B_01A	TACTTCMCATTTAGAATTGC	S02B_27	CCAAATTAKAAGAAADCC
S02A_06F11R11	GCAGTTCCAAAAAAACGTAC	S02B_01N11	AGACTCACTTCACATTTGG	S02B_27FN17	GTATGCCGCCACTCGGACTCG
S02A_06F1E	CCTTTTACCTTTGTTAATCC	S02B_01R06	GTGCAATTCCAAATGTGAAG	S02B_27N23	CCATAACTATGTAACCCTATTCC
S02A_06FF06	GAAAGACTACTCCCCTCCC	S02B_01RLL	ATACTTGAAATTCAACAGAYCC	S02B_27N97	CCATAACTATGKAACCCWATTCC
S02A_06FFEN	CATTAGTACAACCTTTTACC	S02B_02	TCAATGCAYGGYGTTCTTCG	S02B_28	CTTAGGAAGCAGTGCTAG
S02A_06N22	CACTGCTTCCTAAGARCAG	S02B_03	ATTCKKACTAAATAACGAG	S02B_29	GCTTGTATGAATCGCTAYTGG
S02A_06R	CTGYTCTTAGGAAGCAGTG	S02B_04	CATAGTVGCTAAATAATCCC	S02B_29A	CTTGYATGAATCGNTATTGG
S02A_06RF06	GCCATTTGAGCACAATAACTGC	S02B_05	GTTAAAAGAATGAAATTGGC	S02B_29B	GCTTGTATGAATCGYTAYTGG
S02A_07	GTYCCAAARAAACGTACTTC	S02B_05I02	GATGTTATTGACTGTGAACC	S02B_30	TCTAATCGATRATTRGGCC
S02A_07R02	GTACGTTTCTTTGGAACCGCC	S02B_05I05R	TTTCAACGATCAACTTCTCC	S02B_30A	CTTCTAATCGATAATTAGGSC
S02A_07R06	CCGGTAGAAAGAGATTTCCCC	S02B_05IR_17	CTTGAAATTCAACAGACCC	S02B_31	CCATCYGCTADAGCTTGAAG
S02A_07R22	GAGATTTCCCTAATGACAACGC	S02B_06	CTTCAAGCTHTAGCRGATGG	S02B_I11	GTATGTGTTCGTGTATGTCC
S02A_07R22N	GCTGCCCAATATCCCCTCCC	S02B_06A	CTTCAAGCTHTAGCRGACGG	S02B_I16	GCCCAATAGCTGCTTCAGC
S02A_07R97	GTACGTTTYTTTGGRACTGC	S02B_07	CWATTAGTTATGAAATACC	S02B_I22	CCAATAAATCCATGAGAGAG
S02A_08	TTTTGAYATAGAAGTACG	S02B_07E	CACATCCAACTATAACAACG	S02B_I27	CCATAACTATGTAGACCTATTCC
S02A_08E	GATTTCCCTAATGACAACGC	S02B_09	CAACYGTACTTRAACTGTTAG	S02B_I28F	GGTATTCTTTCAGGAGCAG
S02A_08FE	CCATATCTCTAGAATTCGC	S02B_09N17	CAACTGTACTTGAACTGTTCG	S02B_N07F22	GCAGACAGAATTCCATTGTC
S02A_09RE	GACATAATAGAAGTAAGTGG	S02B_09R22	CTAACAGTTTAAGTACAGTTG	S02B_N07R06	TCAACGAATCGCACGTAGAG
S02A_10F1E	CTTTTCTCAAATCATGTATC	S02B_10	CAACAAGATGGACTYTACC	S02B_N13R22	AGACCCGATCCCAGAAAGAC
S02A_10F1NE	GAAACGCTATAATATCTGAC	S02B_10R	GGTARAGTCCATCTTGTTG	S02B_N14	CCRTTCATAAACATCACAGC
S02A_10F22	CCACGCATATTGATATGTCTG	S02B_11	GATAAATGYATYGCTTGTG	S02B_N14R22	GCTGTGATGTTYATGAATGG
S02A_10F23	CCACGCATATTGATATGTMTG	S02B_11R22A	CTATAGGACATACACGAACAC	S02B_N15	GAAGCAGCTATTGGACCAGC
S02A_10FE_19	AAAGAAGAAGTCCCGCCCC	S02B_14	GTCCAAATAATHCCRTACC	S02B_N15R22	CTAGTCCAATAGCTGCTTCAG
S02A_10FE_20	GAAAAAGAAGAAGTCCCGCC	S02B_14R	GGTAYGGDATTATTTGGAC	S02B_N22	CAGATGTTCTATGGATACAGG
S02A_12MKE	CACATTCCTACCAATTCCC	S02B_15	AGYCGAAAYATGGTTAGAGC	S02C_22MK1F	CACCATCAGAATTAAATACAG
S02A_12MKFE	TGATATGACCAACTTGCTC	S02B_15A	CGAGYCGAAATATGGTTMGG	S02C_22MK1R	GTGCATATGATCTAAGGCC
S02A_13	CCYCTYGCTTGTTTTTGGTC	S02B_16N11	GCCCAATAGCTGCTTCMGC	S02C_22RE	CATAAAAAGCTCCTTCTACCC
S02A_13R	GACCAAAAACAAGCRAGRGG	S02B_16R22	CGCTGAAGCAGCTATTGGAC	S02C_2456FE	TAGCCCAATATCCCGAATG
S02A_13R100	GACCAAAAACARGCGAGAGG	S02B_17	TCAGTAAAGATTTATGATAC	S02C_3091RE	GTGCTTGTTCTATCAGTCAG
S02A_14	AAYTGGTCATAYAATCGYGG	S02B_17FR97	CTTAGCTTTACATCCGTCCC	S02C_4100FE	GGTGATCTTATGTATATGC
S02A_14F06	GGAATTACTAATGGWGTTGGTG	S02B_18	GATTCRCATCTYTTACAACC	S02C_5259FE	GAGTGTGGATAAAAATCGAG
S02A_14R06	TAAGAAGARATGCGCCCGCCAC	S02B_18N22	GACTCACATCTCTTACAACC	2YF_1908FC55B	CTCTATCAATAACCGAGCCG
S02A_22MK2F	CCATAAAAACCATATTCTGAC	S02B_18R17_20	GTTGTAAGAGATGTGAATCC	An23_89873F	GTATAAATGTATCGCCATGC
S02A_22MK2R	GTTCACACGGCCTGTTGCTG	S02B_19	GGATTYATCACYACTTTAGC	An23_90320R	CTACAAAGAGTATTACACCC
S02A_D03	GAACATYATGAGCTGYCCCG	S02B_19N02N	GGMTTCATCACTACTTTAGC	An23C_89639F	CCTAGTTGATCCTGATTCGAC
S02A_H13F	GATTGGTCATATAATCGTGG	S02B_19N23	GGATTCATYACTACTTTAGC	S03A_01	CACATTTTTTTCCAYCAAGARCG
S02A_I01	CGACCAATCAGTTAACAGCCG	S02B_19R22	GCTAAAGTRGTGATRAATCC	S03A_03	CCTATTGATATYGGRAGAC
S02A_I03	GAACATCATGAGCTGCCCCG	S02B_21	TAATCTATGCAGCTTYAAC	S03A_07	RATACCGTCTCAATTCATCC
S02A_I11	CTACAATGTCTCAATTGG	S02B_22	TATTYATTATCTATGTYACG	S03A_07N23	AATACCGTYTCAATTCATCC
S02A_I12	GAGCTAACATAATATAACCC	S02B_23	TTCKCGYGGTCCAGAATC	S03A_08	TTCTCGAACYGAGAGATSC

Table S1. Continued.

Primer	Sequence	Primer	Sequence	Primer	Sequence
S03A_08N22	GATTGGTCTGAGGTTATCGAC	S03B_07	GTCACGYCGAGGTACTGC	NS11FA	GGATYAAGGYAGTGGATTGTG
S03A_09	TCTTKTTTGTCGATARCC	S03B_08	GCTTTTCCTTGTGTRGATCC	NS11FAR	CACAATCCACTRCCTTRATCC
S03A_10	TACRTATGGACAATTYCTC	S03B_09	CTCATARAATGGCAGAGGC	S04_01R	CTGCRTATTTGATTCCATCC
S03A_10RAL14	GAGRAATTGTCCATAYGTAAG	S03B_10	GCCCTMTTGTTCCGATGGAG	S04_04N22A	TTGATTCTTCGTCGCCGTAG
S03A_10RR	CTTACRTATGGACAATTYCTC	S03B_10S	GCCCTMTTGTTCCGATGG	S04_04RF20	CTACTTCAACCGATATGCCC
S03A_13	CTGTCTGTAACTCAYTAKAG	S03B_11	CCTGACGTATACGAAGGAGTG	S04_05F	GTTGGCTAGGTAAGCGTCCTG
S03A_13RE	CTMTARTGAGTTACAGACAG	S03B_12	AAAGCAACGACTGGAGTGG	S04_05FF20	CGTCGCCGTAGTAAATAGGAG
S03A_14	MCTTACGATAYTTAGTTG	S03B_13	CTCTRACTCTCCCACTCCAG	S04_05FF20N	GAACCCTGTAGACCATCCAC
S03A_14N18	ACTTACGATACTTAGTTGAC	S03B_14	GCCCTATGTCTCYTATCCC	S04_05FR	GACGCTTACCTAGCCAAC
S03A_14R23	GTCAACTAAGTATCGTAAG	S03B_15	RTCATACTCAATTCGATGG	S04_05NF20	TTGATTCTTCGTCGCCGTAG
S03A_15	GGAAKATCCGTYTTTCTRC	S03B_16	GAGATTTTGAGTCTCGCG	S04_07	TTGATTCTTCGTCGCCGTAG
S03A_16	RAARCAATGTCTCCTTGC	S03B_16R	CGCGAGACTCAAAATCTC	S04_08A	GCAAATTCTCCCAATTTGTG
S03A_17	GGGAAKAKGAGTCGAAGC	S03B_D05F	CCTTCCCGATACCTCYCAAG	S04_08B	GTGAAAATTCTCCCAATTTGTG
S03A_17R22	GTAGCACTTACCAATGAGGCCC	S03B_I04F	CTCTATGTATTCCTCTATCCG	S04_08FAL14	CYCTAGCTAATTGTCCACC
S03A_17RC97	GCTTCGACTCATATTCCCC	S03B_I04R	CGATCCTTTACTCCGACAGC	S04_09	GAACTCATGCCTTATMGAGC
S03A_17RFC97	GATCCCGCTCTAATAGCTCC	S03B_I07F	GTAACTCCCGATATAGCAG	S04_09R	GCYCGATAAGGCATGAGTTCC
S03A_18	TCGTYGAGAAYGATTCTG	S03B_I13R	GAGATTCCCCAATATCATGC	S04_09R06	GCKCGATAAGGCATKAGTTC
S03A_20	CCGTTCTAATACTCYATCCG	S03B_I14F	GTGATAGCATCTACTATACC	S04_10R06	GTTGTACCAAGTCTGAAACC
S03A_21	CATTTCATCYGGGAAAAGCC	S03B_IN01F	GCTTCGATTTGAATTATCCG	S04_12N06	CGAAAAGCAATGAAAAAGGC
S03A_I00	CGATTGCTTGTTGAACCCTG	S03B_IN01F2	CATAACCCTATTCCGGTCTG	S04_13	ATGCTTAGTGTGYGACTC
S03A_I07	AACCGAGCCGGATCTGGTG	S03B_IN03F	CGATCCTTTACTCCGACAGC	S04_13F06	AGATGCGGATAAATGGAAGG
S03A_I08	TCTTGATCAATGGAGGAAC	S03B_IN05R	GTAACTCCCGATATAGCAG	S04_13F06R	CCCTTCCATTTATCCGCATC
S03A_I12	GCTCTGGAACAATTAGGAG	S03B_IN06R	CCAGGAATATCAATGAACCC	S04_13F20	GATCCAAAGAACCAGTCAAG
S03A_I12R	CTCCTAATTGTTCCAGAGC	S03B_IN07R	CTCACCAAAGCACGAAAGCC	S04_13FF20	GAGCTTCGAGCCAATAAAGAC
S03A_I15	GTGGAATAAGAAGCCTTCG	S03B_N01	CTACCGCTGAGTTATATCCC	S04_13R11	GAGTCRCACACTAAGCATAGC
S03A_I18	GTAATTGCCCCATAGATCC	S03B_N02	AAGATGATACGGATCAACCC	S04_14	TAACCCNAAAAGAACCAG
S03A_I22	GATTTCCCATTCCTTAGCCG	S03B_N04	CTCTGGATTTGAAATCACTGC	S04_14N11	CTCATCCAGCTCCTCGCG
S03A_N02	CGATATATAAGGACCGCTAC	S03B_N04R	GCAGTGATTTCAAATCCAGAG	S04_15A	GAACCCATACGTGTTTCCG
S03A_N06	GCTCGGTTATTGATAGAGTG	S03B_N05	CGATCCTTTACTCCGACAGC	S04_15F22	CACCCTTCCATTTATCCGC
S03A_N07	TCCATCTATTCCATGACCG	S03B_N07	GTCACGYCGAGGTACTGCAG	S04_17	GTACGBACAATTACAGCTC
S03A_N09F23	CTCCTTTCCTATTTGATCCC	S03B_N10R	ATGTACTCTACGGATAGAGG	S04_17N17	GTAGATGCTATTCGAACCG
S03A_N17	AGATCTCGTGCATTGGAACC	S03B_N11F	CAAAACTCCATGGACATGCAG	S04_17N22	GTAGATGCTATTCGAACCRCCC
S03A_N19N22	CCCYTCTTCAAACTCCGATTCG	S03B_N13	GCCTGCCTGCCATCCACACC	S04_17N23	GTAGATGCTATTCGAACCAC
S03A_N19R	GAATCGGAGTTTGAAGAAGG	S03B_N14	TTCTGGTGTGGATGGCAGGC	S04_17N23R	GYGGTTCGAATAGCATCTAC
S03AR	GTGAAATGGTAGACACGCGAG	S03B_NN02F	GAGATTCCCCAATATCATGC	S04_17R23	GTAGATGCTATTCGAACCRCTC
S03B_01	GATGACTTCCACCACGTC	S03B_NN04R	TTAGTAGATCTGTTCCGCCG	S04_20	GCTTCYGTTCGTAARATTTGTG
S03B_02	GCAGAACCAAGTCAAGATG	S04R02	TTGATACAGAGCCAATTCC	S04_20R	CACAAATYTTACGAACRGAAGC
S03B_02N09	CAAGATGATAYGGATCAMCCC	AN23_89873F	GTATAAATGTATCGCCATGC	S04_20R23	GAGYGGTTCGAATAGCATCTAC
S03B_03N10	CTCTATGTATTCCTCTATCCG	An23_90320R	CTACAAAGAGTATTACACCC	S04_21	GGTATTRTTGAAACTYGC
S03B_04	GTAAAGTTGCCAGAGTACG	An23C_89639F	CCTAGTTGATCCTGATTCGAC	S04_21F06	ACACCTAATCTTTCGAATCC
S03B_05	CAATGTGATATCTCACACC	AP04_01RF	CTATTTCTTAGATTCATGGAC	S04_22	CCTATGCCRCATAATGGMTG
S03B_05F09	CCTTCCCGATACCTCCCAAG	AP04_05F	TGATTCTTCGTCGCYGTAG	S04_22R	CAKCCATTATGYGGCATAGG
S03B_05F23	CCTTCCCGATAYCTCCCAAG	AP04_11	CCYTCTCTGATCCATTCGAC	S04_23	CACTACAGTGGAARTGTGTTG
S03B_05FF02	GACTTGGTTCTGCTTCCTCTC	NS04F	CATGGCTGAATGGTTAAAGC	S04_23R	CAACACAYTTCCACTGTAGTG
S03B_06F09	GTTTCTAGCGATTCACATGG	NS04FAL14	ACTTCTAATGTCGAATCAG	S04_24	GAAATATGGACRAAYGGAAG

Table S1. Continued.

Primer	Sequence	Primer	Sequence	Primer	Sequence
S04_24N20	GAAATATGGACAAATGGMAG	S05_18R	ACCACATAGTTCCVGCAAC	S06A_H03RF	TCAACCCACCCTTAGTACCG
S04_24R20	CTKCCATTTGTCCATATTTC	S05_19	CAACVGCTATTCCATCYCC	S06A_H04R	GCTTCRTCGATGTTACCTAC
S04_25R22	TCTAGAATGCCCAATATCTG	S05_19R	CGGRGATGGAATAGCBGTTG	S06A_I00	CTTATACTCATTATGCCTAGC
S04_D05	CTACTTCAACCGATATGCCC	S05_19RFE	CATTCCTCTTAGTCTCGACTC	S06A_I00R	GCTAGGCATAATGAGTATAAG
S04_D15	GAACCCATACGTGTTTCTG	S05_20	CAGAGATGTTTTTGCYGGT	S06A_I03R	GTCCCTCCCTACAACTCATG
S04_I10	GTTTCAGACTTGGTACAACCC	S05_20N20	GTTTTTGCTGGTATTGACCC	S06AB_NF	CCGAGTAACTCCTCAACCTG
S04_I21	CGTCCATTCTTACGTGAACC	S05_20R	ACCRGCAAAAACATCTCTG	S06AB_NFR55B	CMGGTTGAGGAGTTACTCGG
S04_IN01F06	TAGGAGTAAGCTGTGACACG	S05_22	CAAARGCGGTATATAAGGC	S06AB_NR	CCATATATGTTCTGATCGCC
S04_IN07RR	GGCTTTCGTGCTTTGGTGAG	S05_22RN	GCCTTATATACCGCYTTTG	S06B_01	GTCACCAMAAACAGAGACTAAAGC
S05_34R23	GCACTGTTAATAATCCCGCAG	S05_24F06	CAAAGGTTCTTACTCAGAGAC	S06B_02	GCAATGCATGCRGTTATTGATAGAC
AP05_03	AAGAATCGKCAYAACACGCC	S05_24N06	CACTGAATTCGGAATATGG	S06B_02F1N06	GAAATGCACCCGGTGCCGTAGC
AP05_14	GTCGCACTATACGTCAACCC	S05_24NR	CCATATTCCGAATTCARTG	S06B_02FF06	GTGCCGTAGCTAATCGAGTAG
AP05_29	CACTTTTGTATTGCCGCTTC	S05_27	CCTATCTCAATAAAGTHTATG	S06B_02FN06	CTGCGTGATGATTTCATTGCC
PH05_16	CTGTTGTATTGAATGATCCCG	S05_27R	CATADACTTTATTGAGATAGG	S06B_02R	GTCTATCAATAACYGCATGCATTGC
S04_24E	GCACTTCATGAAGCCTCCCG	S05_28	CGRTAGTAAAARGTCATAG	S06B_05	ATGATACTMAATWTAGTTGG
S05_00	GTACCCAAAACAAACGCGC	S05_28F22	GATTCTCTTGTTAAGACCC	S06B_08	GGCGATCAGAACRTATATGG
S05_00R	GCGCGTTTGTTTTGGGTAC	S05_29N23	CTAATGATACGTAAACARGG	S06B_08R22	CCATATATGTTCTGATCGAC
S05_03A	TGGAATGTACCGVTTGTG	S05_29R06	CCAGAAATGCCTTGTTTACG	S06B_10	TATMGTMTTTTTGCCCTGG
S05_03FAP	AGCGACCTTTTCTTAAATCC	S05_30F22	GATAAAGACAACTGAGCCAAC	S06B_14	CACTATATTTCTGACCVGG
S05_03R22	CCAAAATAACCGTTATTCG	S05_30N20	GGAATATGAGTGTGTGACTTG	S06B_14FR06	GCCATTTAGCTAATAAGCCCGTG
S05_04	CAAGCGATCTTTTCGTARGC	S05_31N20	CAAGTCACACACTCATATTCC	S06B_14R	CCBGGTCAGAAATATAGTG
S05_04F55B	CTATTTGTGCAAGTATTTTACG	S05_32	ATGATCCTGTATTAAGAGC	S06B_15F20	GATGACTACCTGAAAGAACCGC
S05_05	CGACTAATTAAACTCATG	S05_33	CCYGAYTTGAATGATCCTG	S06B_16	TTCCMGGRAATCCCCAACG
S05_06	GAGTGATCCACAAACGRCG	S05_33R22	TACAGGATCRTTCAARTCRGG	S06B_16FAp	CATGTTATACCTTCTTCCAC
S05_06FE	CGAATAACTAATATATTTCC	S05_34R23	GCACTGTTAATAATCCCGCAG	S06B_16R11	CATTCCCTTGACTCGTACTG
S05_06R97	TCGYCGTTTGTGGATCACTC	S05_34RF22	GCGAACCAGCCTATCCTCTG	S06B_17FR97	GACATAAGTCCCTCCCTACAAC
S05_07	GCTCATTCAAGRCTTACTCG	S05_35N22	CCTATTGATAAATCCCTAAC	S06B_17R06	GAAGCCATTGCCATTGCCG
S05_08	CACCCTAGTACATGTTCC	S05_36B	GTAATGTTGCACCAATMCC	S06B_17R06R	CGGCAATGGCAATGGCTTC
S05_08F23	CAATGACCAGAATTAGACGAG	S05_A10FR	CCAAACGTCTAGCATTCCCC	S06B_18	GATATCYCCCAACATACC
S05_09	ATGCCAACTATTAARCAACT	S05_C10F	CCAAACGTCTAGCATTCCCTC	S06B_18F06	TGAGACTCAAGATAAATGTC
S05_09FE	GGAAATATATTAGTTATTCG	S05F_10F23	GATCAAATTACCACCCTTCC	S06B_19	CAATTTTCRCATTGAACCC
S05_10F22	GTAATTTCTCCTCCGGCCAAG	S05FR11	GCGCTCTTAGTTCAGTTCGG	S06B_19R17	CAATGYGAAAATTGTTATGG
S05_12F06A	GACAAATGCTTCCATTACCGC	55B06B_01R	GACATAAGTCCCTCCCTAC	S06B_22	TCATTACGAGCTTGTACAC
S05_12F22	GACAAATGCTYCCATTACCAC	55B06B_02F	ATGCCTGCYCTGACYGAGATC	S06B_23	CCATACTTGTTCAAYTTATCTC
S05_12FR22	GGTAATGGRAGCATTTGTC	AP06B_06	CCAATTCTTACATCTCCTAC	S06B_23F06	CCGCGGCCCCCGCTTCTTCAG
S05_13F100	CGTCTAGCATTCCCTCACGC	AP06B_17R	GACAACTTTCGACTTTCCCTC	S06B_23F23	CACCAGCTTTGAATCCAACAC
S05_13F22	GGATTCCAGTTCAGCGTCAC	AP06C_03	CTGACGAATAACCAACCCGC	S06B_23R	GAGATAARTTGAACAAGTATGG
S05_13FFCT	GACTATGCCTTCGCCATATG	AP06C_05	AGTCGTATCTTRCTCAGACC	S06B_O18F	CAATTYTCACATTGAACCC
S05_14	GTCGCACTATAYGTCAACCC	AP06C_06	TCTACTAGCGGCTTTAACGG	S06C_01	CCGATCACARGAATRCCAG
S05_16	GGTTTRCCTTGGTATCGTG	S06A_00R	GAAAGCAATACTCTAACCAC	S06C_01A	GAATACCAGYTACAGTACC
S05_16R	CACGATACCAAGGYAAACC	S06A_00RC97	GCYAGGCATAATGAGTATAAG	S06C_02FRC97	TAGCATTACTATACCTTCCC
S05_16RE	GTATGAACACGATACCAAG	S06A_02N06	TTCTTTCTTGTAACGTACCC	S06C_03F22	GTTGTATTAGATCGAGATC
S05_16RF97	CTCGATAAGTACCAATACGC	S06A_03	CAATCACTTCCATYCCTCTC	S06C_03F23	GTTGTATTAGATCGAGACC
S05_17R22	GTTTCTTGGCAGCTATCTGGC	S06A_03R	GAGAGGRATGGAAGTGATTG	S06C_04	GTTTTTYTATTTGGAATCGT
S05_18	GTTGCBGGAACTATGTGGT	S06A_05	CCGGGAGCTCGTATGAGAG	S06C_05FF22	CCAATGACCATAGATCGAACC

Table S1. Continued.

Primer	Sequence	Primer	Sequence	Primer	Sequence
S06C_07	GCAATGCAGTTCATCCAACG	S07_YCF3Ap3F	GTAAATGCCTCTTTTTCTCC	S07C_07RE	TCCATAACGATCAAARTCG
S06C_08	ATGATACTACTGGAAGGATTC	S07_YCF3Ap3R	CGAATAATTCCGACAACTTC	S07C_08B	TTGGCATTCCATTGCTGTC
S06C_08F23	GAACAGCCCCTACATTCAAAGC	S07A_06R	GCTTGGRCAGGRCATTTAG	S07C_08F02	GAAATACAAGATGCTCATTG
S06C_08FR06A	GGAATAGAACGGATAAATGCG	S07A_06RF23A	GTTATTACTTCCTTGGTAGC	S07C_08F22	GAATATGTATATGCGTATGAGG
S06C_08N18	CGATACTACTGGAAGGATTC	S07A_06RFF23A	CAGTTTAATGAATCTTCCAC	S07C_09	GACACGAGGATTTTCAGTC
S06F	GGRCTATACGGAYTCGAACC	S07A_10F06	CTTYAGTTATTCACTAGAGC	S07F	GATTTGAACCAATGACTCCC
S06FA	CTCAGTGGTTAGAGTRTTGC	S07A_10F1E	CCCTACGAAATACCCAGAC	S78_03F	GAGCCCAATGCACAAAAATGC
S06FAR	GCAAYACTCTAACCACTGAG	S07A_10F22	GTCCTCCTCTTTCCGGACAAC	SAA1	GATTATTCGTTCRCCGGAAYC
S07C_05F22	CTCCGGAAACGAAAAATGC	S07A_10FRE	ATATCCGAACACTTGCCCC	SAA2	CAACTTTATCCNAGTTTTGC
An23_49470R	GTGAGCCCGCTTAGCTCAGAG	S07A_11N06	GTGTTCTTGAGAAATGACCC	SAA4	GTTGCTCCTGCTAYYCAGCC
An23_54693R	GTTCGAGTCCGTATAGCCC	S07A_15R17	CCAAATACGACGAGTAGTG	SAA5	GTTGTGGCAATTCCRCCCAG
AP07A_07F	TGTGCCCATAAGAAATCGC	S07A_17FAL	GCTTTAGATGCACGTGGTTC	SAA6	ACATGRATYGTAAATGCATG
AP07A_07FR	CAGGCATCCCARGTAATTC	S07B_01	GGCYTGTTTCCAATACTC	SAA7	AGAGMTAATTGAGCATGCC
AP07A_09R	TGGCATGCTCAATTATCTC	S07B_01F23	GCACGTAATGACAGATCACG	SAA8	CCYGCCCAAGARAGAGACCC
AP07B_01	ACGCCTCGTGATCTTCAACC	S07B_05	GTTBAAGGCGTAGCATTGG	SAB1	CAGGACCCCACTACTCGTCG
AP07B_01F	CTTGCTCTATCAATTAGATCC	S07B_05N97	AATGCCTYTTTTTCTCCTG	SAB2	AATCTTTATGCYCAAAACYC
AP07B_01F1R	GTCATTACGTGCGACTATCTCC	S07B_06FE	GAATCGAACCGATGACCATC	SAB3	GACTTTACTACTCAAGCYGC
AP07B_06	GTTTTGTTCMTCCCTCGCCG	S07B_06FEN	GCGAAGTTCACGGAAGTAGTG	SAB4	CAYATYACATTATGGCARGG
AP07C_04	CCCYGAAATTGTACATGCTC	S07B_07	GGTTAGAGCATCGCATTTGT	SAB6	CCHAGAGCAATAGCATGATG
AP07C_06F	TTTCCGGAGTKTTAGCCCC	S07B_08	GGATATGGCGAAATYGGTAGAC	SAB7	GCTTGAGTAGTAAAGTCTTG
BR_LFF0	CGARAGTTGGRTAGGRTGGC	S07B_09	CCGATGACCATCGCATTAC	SAB8	CTAAATGYCCTGYCCAAGC
BR_LFF1	ATTTGAACTGGTGACACGAG	S07B_11	CTACATAGCAKTTCCAATGC	Tugu_FSP	GTAATGCGATGGTCATCGG
BR_LFF2	GGGATAGAGGGACTTGAACC	S07B_12F06	GACAACAAACATCCATCTCG	Tugu_RSP	ATGCGAWGCTCTAACCTCTG
BR_LFF4	GAGCCCGCTTAGCTCAGAGG	S07B_12RAp	GTAAATGCCTCTTTTTCTCC	An23C_36128F	CTGTATTCCTGATTTATCCAC
BR_LFR1	CTCTAACCTCTGAGCTAAGCG	S07B_13N22	GTAACGGATGTCRGCTCAATC	An23C_36623R	CTTGAGCCGCCATCCAAGCACG
BR_LFR3	GGTTCAAGTCCCTCTATCCC	S07B_13R09	GATTGAGCYGACATCCGTTAC	AP08A_01F	GAGATATTTCAAGGATTCCC
BR_LFR4	CTGAAAATCCTCGTGTCACC	S07B_14	GTGATCTGYCATTACGTG	AP08A_02F	CACTGTTGATGAACTAATCCC
CTR_3YCF3R	CTCCGCAATTTCAGAATCTCCC	S07B_14F06	CCGAGATAGTTACCTTTCAG	AP08A_04	CTAGTTCCTACTGCTTTTCTGC
CTR_ndhCR	GATATTATTAGAAATGCCCAG	S07B_15A	GCTTGGTTYGATCAAGCCGC	AP08A_04F	CTTCGTGTCGATCAATCTAC
CTR_psaAR	GTAGGTTCCAGATCCAAGTG	S07B_18FAp	GATGATCATATAAAGCCCC	AP08A_04FR	GATCGACACGAAGTCCCTGTC
fCAUR	GACCTCAAGGTTATGAGCC	S07B_A04F	TCGAATCCCTCTCTTTCCG	AP08A_04R	GATAAGCAGAAAAGCAGTAGG
gccR	TGGCAAAGAGAAATTTTACC	S07C_00F06	TCATATTCATTATGACTAGC	AP08F	GACTCACGAGAATTTATGAGTG
IR_S01F1	GATTTTCAGTCCTCTGCTCTACC	S07C_01	GTTCAAGTCCCTCTATCC	DEN_TS1A_F0	CCCTTTCACTCGTAGTATGG
IR_S01R2	GATATGGCGAAATTGGTAGACGC	S07C_01F22	GACTCCCCATTTGAACGAGTC	DEN_TS1A_RB	CGAGCCCCGTCAGTCCCGACG
Lcaa_AR	CTACCAATTTCGCCATATCC	S07C_01R	GGATAGAGGGACTTGAAC	S08A_01	GGATATAGTAAGTCTYGCTTGG
PH07B_11	AACCACTCGGCCATCTCTCC	S07C_01RF1E	CCCGCACCCGCATTTGGTAC	S08A_01F23	GTTGACGAGACTCAGATTGC
S05_10F22	GTAATTTCTCCTCCGGCCAAG	S07C_01RF22	CATTGAGTCTCTGCACCTATCC	S08A_01F23A	GGAATAGATGCCGATAATGC
S05_12F22	GACAAATGCTYCCATTACCAC	S07C_02	GAGCAGAGGACTGAAAATC	S08A_02F23	GTGAATTATGTCCCCTATCTC
S05_17R22	GTTTCTTGGCAGCTATCTGGC	S07C_04	CGAACRGTACTATAAGAATC	S08A_I02	CCATTAAAGCAGCCCAAGCG
S05_33R22	TACAGGATCRTTCAARTCRGG	S07C_04N23	CGAAYAGTACTATAAGAATC	S08A_M01	GGATATAGTAAGTYTCGCTTGG
S07_08R	CCATGCATTTACRATYCATG	S07C_05F22	CTCCGGAAACGAAAAATGC	S08A_N01	GGATATAGTAAGTCTCGCTTGG
S07_YCF3Ap1F	CGAATAATTCCGACAACTTC	S07C_06	GTTTCTGCTTYACGAATATGAT	S08A_N01RR	CCATTAAAGCAGCCCAAGCG
S07_YCF3Ap1R	GTAAATGCCTCTTTTTCTCC	S07C_06F23	GAGTTTCGATGTATTAGGTG	S08A_N02	CTTCCCCATACTACGAGTG
S07_YCF3Ap2F	CGTAACGGATGTCRGCTCAATC	S07C_06FR20	CACCTAATACATCGAAACTC	S08B_01	GACGGACTGTAAATTCGTTG
S07_YCF3Ap2R	CGTTACTTGTGTGGATAAGTC	S07C_07	CGAYTTTGATCGTTATGG	S08B_02	CTGGATTTGAACCAGCGTAG

Table S1. Continued.

Primer	Sequence	Primer	Sequence	Primer	Sequence
S08B_03	GACATCTCTCTTTCAAGGC	AP09A_05F	ATTCCAATAGAAGTCCTTC	S09A_12FFF23	GGTGAAGGAAGTATGACTGC
S08B_03N17	GGACATCTCTCTTTCAAGG	AP09A_07	CCCTACAGCCAATCCAGCAG	S09A_12N22	AAGTGATTTATTAGATAATCG
S08B_04	TCGAATCCCCGYTGCCTCC	AP09A_08	AATACCGATATTTTAGCAAC	S09A_12R23	CGAATAGTATTCAAGATCCTC
S08B_04F22N	GTCGAATTGATAGGTGTAC	AP09A_09	GTATTAAACCCGAAACTCCC	S09A_13	TTCAGAGTTCCTAGAGCYC
S08B_05	AGTAACGCCATGGTAAGG	AP09A_11	GTATCCATTGACTCGAAACTG	S09A_14	CAAGACCRTGAATACGAGC
S08B_05FC17	CTGACCCATTGAATCATGAC	AP09A_11F	GAATGTAATGAGCCTATCCTC	S09A_14R	GCTCGTATTCAYGGTCTTG
S08B_05FE	TTGTCGATATTTCCAATTCC	AP09A_12	CAGTTTCGAGTCAATGGATAC	S09A_15	GYGTTTATGTAGCTATTGG
S08B_06	CCGATGACTTACGCCTTACC	AP09A_12F	TAATTGAATCTCCTGCTCCG	S09A_16	GCVAGRTATTGTAATGTAGC
S08B_07	GATATTATGGATGAYTGG	AP09A_16F	GAATCAATAGCAATAAGCCC	S09A_17	CTCTTCTATTAMGAAGACC
S08B_08	CGGTTTATGGATGAGYGC	AP09A_18	TGAATTGCTTAAACAATCCC	S09A_18	CAGAATCAATTGGCAAGAGG
S08B_09	CGYGCAYTTGTATGGTCTGG	AP09A_18RF	GCAGTCATACTTCCTTCACC	S09A_18F23	GCTATTCAGGAACAGATGGAG
S08B_10	TTGAAAAAGGAATYGATCG	AP09A_19F	ACGAATCACACTTTTACCAC	S09A_18FR22	CAGAAAGAAGACCCCTTTGC
S08B_10N05	GGTAAGGCGTAAGTCATCGG	AP09A_19FN	ACGTCTCTAATTCAAAACCG	S09A_18R	CCYCTTGCCAATTGATTCTG
S08B_11	GCTAACCAACTTCTAGGAG	AP09A_19FR	AGTGGTAAAAGTGTGATTCG	S09A_19	GAAGACCTMTGTCCTATCC
S08B_11N17	GCTARCCAACTTCTAGGAG	AP09ABF	TCGAAGAGTGCACCAATCCC	S09A_19F02	TGAATYAAACCGAAAGACCCC
S08B_12	CGTCYACACTAACAATCC	AP09ABR	GCRATCTGTTACCGAGCCGTC	S09A_19R	TGGATAGGACAKAGGTCTTC
S08B_13	CCATTAAAGAGCGTTTCCAC	AP09B_03R	GTTTACACCCATTATCTCAAG	S09A_19RF	GAAGACCTMTGTCCTATCC
S08B_13R	GTGGAAACGCTCTTTAATGG	AP09B_04	AGTYATGTATCAACGTTCC	S09A_I05F	GAGTTGGTCAAGGTACTGC
S08B_14	TTGAACAGATCGAGCAAG	AP09B_08	GAAGTAGATACCTGAATGC	S09A_I06	CACTGATTTCTGCCGCTTCC
S08B_14F06	CACCAGTTGATCCTCAGACC	AP09B_11R	ATCACTCACATATTCTTCTG	S09A_N16R_21	GCGTTTATGTAGCTATTGGTC
S08B_14F22	GTGAACCAGCCTCCTAAAGCG	AP09B_18	TCTTGTACCAATTTCTCCTG	S09B_00	CGACACCCRGATTTGAACTGG
S08B_14F22N	CATGAAGTTACAAAGGTTGTAC	AP09B_22F	GACTGACCCTTTTAATCCAG	S09B_01	ATCAGATMCAATTTGARGG
S08B_14FE	ATATTTCCCCTTACCTTACG	AP09B_32	CATCCGCTAATACACGACC	S09B_01R	CCYTCAAATTGKATCTGAY
S08B_14FEF	GACGACTTATCATCCTACTG	AP10B_G01F	ATTTAATYCTTTACCTCTC	S09B_01R11	CCTTCAAATTGTATCTGAC
S08B_14RN17	CTGTATTCCTGATTTATCC	PH09_ES1F	GTCCACTCAGCCATCTCTCC	S09B_03	CGAGATATCCATCCTAGTC
S08B_16	GCTCCGGTCTTGAAAACC	S09A_01	SRTTTTTGATGCAGCAAG	S09B_03N20	CGAGATATYCAYCCTAGTC
S08B_17	CTATGCCGGTTTTCAAGAC	S09A_01RF22	GTCATTTCTCCCCACACCCC	S09B_03R	GACTAGGATGGATATCTCG
S08B_17F06	GGAGGATTAAAAGCATGAC	S09A_02F06	GCAGATATTTCGATTCCAGC	S09B_04	CATTGGTTATGTATCARCGTTC
S08B_18	GCTTTCCAATTRGCTGTTTTTGC	S09A_02FF06	CGCTTAGTTTTCGACTTTTCG	S09B_06	GCACKTGTTACACAACAACC
S08B_18F	GTTCGATTCCCGCTAYCCG	S09A_02FF22	GTTAGCCGCACCGACGAATG	S09B_06R	GGTTGYTGTGTAACAMGTGC
S08B_18F55B	TCTTAATTCTCTCATCTCTTG	S09A_04	GCAATTACRACCCAAGAAG	S09B_09F2N22	GGATCCCCGCCTACACAAGC
S08B_19	GAAAGACTAAYCCAATCC	S09A_04F20	CTGTCATATATTTAATCCCGC	S09B_09F3N22	CTTAGATTCGAACCCATAGCC
S08B_CT11F	CAAGTAAGCCTCYCCAGACC	S09A_05	AGCCGCGGCYTATATAGG	S09B_09FAp	CCCATTCCAAATTTCATTCC
S08B_I05	GGTAAGGCGTAAGTCATCGG	S09A_05R17	CACCTATATARGCCGCGG	S09B_11	GATTCATAYACAGCATCYC
S08B_I11F	CAAGTAAGCCTCTCCAGACC	S09A_05R18	GATTCCCCTATATAAGCCGCG	S09B_11F06A	GACATTCCCTCATTTCCATCTC
S08FA	CTCGAACCCGCAGCTTCCGC	S09A_06	GTTATTGCTGCTGGATTRGC	S09B_12R23	CCAATTCGGAGTTGTTGATG
S08FAF02	CAACGAATTTACAGTCCGTC	S09A_07	GTTAAAGCTTCCATAAAAGC	S09B_13F20	GTAAATTCCTTATGAATTCC
S08FN	GACTGCAAATCCTTTHTCCC	S09A_08	CCGATATTTTAGCAACAAATC	S09B_13FR22	GTCCCGACTTCACCTCGAC
S09B_09F2N22	GGATCCCCGCCTACACAAGC	S09A_08F22	GGTCAGAAGAGTCCTCCGAC	S09B_14F100	CTATTACCTGTTCTTCCTCC
S09B_09F3N22	CTTAGATTCGAACCCATAGCC	S09A_08FR22	GATAGAGAATCAAAGCGG	S09B_16F23	CTGACCCTTTTAATCCAGTCC
S09RF23	CCGATTCAAGATCATTGAGAG	S09A_09	CAACTCGCACACACTCCC	S09B_17A	TATGACCAACAGTDGTTCG
04_S09A_12FF	CTACATTACAATACCTCGCTCC	S09A_11	ATTCATTTGGCTCTCACG	S09B_21	CCACAAGCACAAATTCCRCT
An20_23850F	CTTGGTAACTAACTATTTGC	S09A_12	GTGATTTATTAGATAATCG	S09B_21R	AGYGGAATTTGTGCTTGTGG
An20_24301R	CCTCTTTTCGAGATTCATCG	S09A_12F22	GTAACCATTCGAGCCGACG	S09B_22F22	CTGACCCTTTTAATCCAGTAC
AP09A_03	CTAGCTGTTCGGAACCCAC	S09A_12FF22	GGTTTGCTGATACAAGAAGG	S09B_22R22	GTACCTCCATATTCCGCTG

Table S1. Continued.

Primer	Sequence	Primer	Sequence	Primer	Sequence
S09B_23F23	CAAGTGATTGCTGAGATTCGC	S10A_01	GGTATGGAAGTWATGCAYG	S10B_04R	GGTTTAGACATTACTTCGG
S09B_23FR22	GCGAATCTCAGCAATCACTTG	S10A_01E	CTCATAATTTCCCTCTAGAC	S10B_04R11	GTCCTTCAAGTCGCACGTTGC
S09B_23R09	KGCCATAACATAATTGACAG	S10A_02	GCATWACTTCCATACCAAG	S10B_05	CGTTTRAAACGATGTGGTAG
S09B_24F22	AACTACAAGAAGCCGAGTAG	S10A_02F20	GTGCCAATCCAACACAAGYCC	S10B_06	CCAAAACCCGNTGCCTTACC
S09B_27	CGTATYGATTGGTTGAAAGG	S10A_02F22	CCGCCGAATACACCAGCTACG	S10B_06F18	CTTCGCCAAATTGCCYGAG
S09B_27F22	GGAACATCGATTTGGAAGAG	S10A_03	GTGCTATGCATGGTTCYTTGG	S10B_07	AGGCANCGGGTTTTGGTCC
S09B_27R	CCTTTCAACCAATCRATACG	S10A_03R	CCAARGAACCATGCATAGCAC	S10B_08	GTKTGGCAAGCTGCTGYAAG
S09B_28	GATATKGAATCAATYGAACG	S10A_04	GATCATYAAAACACCRAACC	S10B_08F22	CAAGTCATCAACAAAAACGG
S09B_28N09	GATATTGAAYCAATCGAACG	S10A_04F22	CGAATCATCCAATTCATGACTG	S10B_08R06	CGAAAACTTACAGCAGCTTG
S09B_28N100	GAACAATTARAACTTGACCGG	S10A_05A	GAGAGACGCGAAAGYGAAAG	S10B_09	CGAANAGTTTGAGAGTAAGC
S09B_28N23	GAACAATTAGAACTTGACCCG	S10A_06F22	GATAACCAGCACCGAAAACCG	S10B_09F22	GATATTAAGCATAACAGACG
S09B_30A	CTTATCTGTYTCTCCAGG	S10A_07	CTCGAACCCGGAACTARTC	S10B_09F23	AGATATTAAGCATAACAGAC
S09B_30F22	CTAAATTCCTTGACTCCTCG	S10A_07F06	CTTCCAATAACCCCGAGCCC	S10B_09R	GCTTACTCTCAAACTNTTYG
S09B_31	TTTGATCYTGATCCTTGTG	S10A_10	GATTTTCCTTGATACCTRAC	S10B_09RFF02	GTCCACTCAGCCATCTCTCC
S09B_32	GTTCGTACAGCAGTATCYAC	S10A_10A	GATTTTCCTTGATAYCKAAC	S10B_09RFR17	CCTCTGTCCTATCCATTAGAC
S09B_32F20	GAGTCTTTACTTGATCCAG	S10A_10F06	GAAGAAGCGTCTTTTACCCAG	S10B_10	CGCACGTTGCYTTCTACCAC
S09B_32R	GTRGATACTGCTGTACGAAC	S10A_10R20	GCATTATGTTAGGTATCAAGG	S10B_10FR06	CCAATGACTATTCATGATTCC
S09B_I01R	TCAGAAGAATATGTGAGCG	S10A_12N06	TTTAAACCATGATCATGCGC	S10B_Ap_F	GACCAAGCTMCTTAAAAACC
S09B_I20F	CACATCGCAGTATACTAGGCC	S10A_13B	CTTTCAGGAGTATATTTATG	S10B_Ap_R	CGAAGTAATGTCTAAACCCAC
S09B_I20FR	GGCCTAGTATACTGCGATGTG	S10A_14	GCTAACTCAACGGTAGAG	S10B_G00	RTCCACTCAGCCATCTCTCC
S09R	GTCAGAGCACCGCCCTGTC	S10A_14NPHA	GTTGCTAACTCAAYGGTAGAG	S10B_G01	GTAGCGGGTATAGTTTAGTGG
S09RA	GCAGTCCCCCGCCTTACCAC	S10A_15	CCGCACTTAAAAGCCGAG	S10B_G01F55B	AGCTGACATAGATGTTATGG
S09RAF06A	CCATTAACTATTGACTCCTG	S10A_CH12	GTGCTAGAACTTTGGCTCG	S10B_G01FRC17	CCATAACATCTATGTCAGC
S10B_G02F22	TGAATCAAACCGAAAGACCCC	S10A_CH12F22	CATGTAAATGTAAATTGTCAC	S10B_G02	TACCACTAAACTATACCCGC
AP10B_08RN	ATAAGATCTACGATTGGATTC	S10A_H12	AGYGCATAAATATACTCCCG	S10B_G02F17F	GAAAAGCGTCCATTGTCTAATG
AP10B_G01	GGTAAAAGTGTGATTCGTTC	S10A_H12F	TTTTYTCAACCCAATCGCTC	S10B_G02F22	TGAATCAAACCGAAAGACCCC
AP10B_G01F	GAGCTGACATAGATGTTATGG	S10A_H14F	CAAAATCAAAAGAGCGATTGG	S10B_G03	GGAGAGATGGCTGAGTGGAC
AP10B_G03	GTTCGAATCCCTCTCTTTCC	S10B_01	CATCAATTGCAAYGATTCG	S10B_G03R97	CCACTCAGCCATCTCTCCCC
Gucc1R_02	TACCACTAAACTATACCCGC	S10B_02	CGAATCGTTGCAATTGATG	S10B_I10	CTCGAGCTCCATCATGTAC
PH10A_01FRN	GCATCTACCATTATACCCAC	S10B_02F06	TGCCAATCCAACACAAGCCC	S10B_I10R23	GTASATGATGGAGCTCGAG
rps16_MF	ATGGTAAAACTTCGTTTGAAAC	S10B_02N02	ATCGTTGCAATTGATRTTC	S11_01	GGATGGAATCAAATAYGCAG
rps16_MF1	GAAACGATGTGGTAGAAAGCAAC	S10B_03	GTGGTAGAARGCAACGTG	S11_03F	KAAGAAGCTTGTACAGTTTGG
rps16_MR1	GAACATCAATTGCAACGATTCG	S10B_04	CCGAAGTAATGTCTAAACC	S11_N05F	GAACCCWGTAGACCATCCCC

Table S2. Long PCR primer list used in this study.

Primer	Sequence	Primer	Sequence	Primer	Sequence
An_L01F	GGAAGTCATCAGTTCGAGCCTGATTATCCC	An_L10B_F	CCGAGTACTCTACCGTTGAGTTAGCAACCC	L07G_An06_F1	GGTCTGGCGCGAATAATATTCTACGACTAG
An_L01FA	GAGAGTTCGATCCTGGCTCAGGATGAACGC	An_L10F	CTGCCTTRATCCACTTGGCYACATCCGCC	L07G_An06_R	CTAGTCGTAGAATATTATTCGCGCCAGACC
An_L01R	TTCGTAGCCACGTGCTCTAATCCTCTGAGC	An_L10R	CRCRCTTTTACCACTAAACTATACCCGC	L07G_An22_FQ	CCTCGTGATCTTCAACCAATTATGCGCTTC
An_L01RA	GTATCGTCACCGCAGTAGAGTTTMACCACC	An_L10RA	GGAGAGATGGCTGAGTGGAYTAARGCG	L07G_An22_FQ2	CGCACCATCTCTGTAATAGGTAAATGCC
An_L02AR	CACTGCTTCCTAAGAGCAGCGTGTCTACC	An_L11FA	CAAGTGGATYAAGGYAGTGGATTGTG	L07G_An22_FQ3	CCATTAACCGAAGTGCAAAAATAGACCC
An_L02BF	GGTAGACACGCTGYTCTTAGGAAGCAGTG	An_LRF	GGCGTAGAGGAACMACACCAATCCATCCCG	L07G_An22_RQ	GATCCCTATGATCGAAGTTATATACTC
An_L03BF	CCTTGAAGAGGAYTCGAACCTCCACGCTC	An_LRFA	GAGCACGTGGCTACGAACCACGGTGTCGG	L07G_An22_RQ1	GATTCTGAAATTGCGGAGGCTTGGTTCGAC
An_L03F	TTATGAGTTGGGCGCTTTAACCATTCAGC	An20_L2B03F	CAAGCCATAACAAATGCGAAGCTATACGAC	L08ABN_FR23	CTTACTATATCCATGTGAATTATGTCCCC
An_L03R	CGTTCATCCTGAGCCAGGATCRAACTCTCC	An20_L4GF1	GAAATCGGCCACATTAAAATTACCTTCTG	L08ABN_R	GTCTACGCTGGTTCAAATCCAGCTCGGCCC
An_L04F	GCATCCATGGCTGAATGGTTAAAGCGCCC	An22_L2B31F	CACTTGAATATTGCCTAACTGCTCTGGTGC	L08B_14F22	GTTGTACCTGTGAACCAGCCTCCTAAAGCG
An_L04R	GAATTCACCTATCCCAATAACAAARAAACC	An22_L2CR	GTGCACTTTTTAGAGTCGAAATAAAGTG	L08B_17F20	CAATTGGCTGTTTTTGCATTAATTGCTAC
An_L05FA	GTGACATTTTGTACCCAAAACAAACGCGC	An22_L2CR1	ACTCTTAAAAGTATATCCAACTCGATTC	L08B_17FAp	CTACTTCATCAATCTTATTAATTGGCG
An_L05R	TATTTGAAATACGGGAAAGAARTACCATTC	An22_L4GF	CCCCTTCTTTCACCCTTCCATTTATCCGC	L08FAF22	GATAGGTGTACATGTATCAATCAAACG
An_L06R	GCAGCTTGGTAGCGCGTTTGTTTTGGGTAC	An22_L4GR	AAAGAAGAATCTACTTCAACCGATATGCCC	L09FCF23	AACTCCAAAATRGCATTTTCTTTTGACCC
An_L07FA	GAACCYGTRACCTCAAGGTTATGAGCC	An23_L4GR	GAGGTCCGTTTGATATCCAAAAACTGCTC	L09RR23	GAATTCCATTTTTGTATGTGCGCTTGAG
An_L07R	TACCTACTATTGGATTTGAACCAATGAC	C97_LuucR	GAAGTCGAATCCCCGTTGCCTCCTTG	L10B_08FF20	GCTTACTCTCAAACTTTTCGTTTACACAG
An_L08F	GGACTGCAAATCCTTTHTCCCCRGTTC	C97_SSC_G3R	GAATTAAACCATCGTCTATTATCCATTG	L22_MKF	TCCACGCATATTGATATGTCTGTTCC
An_L08FA	CTCGAACCCGCAGCTTCCGCCTTGACAG	L02B_27F17	GGACTCGAACCGAGATGCTCTAGCAC	L22_MKR	CACATCCATATCTCTAGAATTCGCTG
An_L08R	GTAGAGCAGTTTGGTAGCTCGCAAGGCTC	L06B_GF06	GAAATGCACCCGGTGCCGTAGCTAATCG	L2C_G20F1	GTCTCCTTCTTCCTGATACTGATTACC
An_L09F	GHCGCYTTAGTCCACTCAGCCATCTCTCC	L06B_R	CYGATACTACTGGAAGGATTCCTCTTTGG	L2C_G20R1	GATTCTTTCTTTTCCATCACTTCGGCATG
An_L09F20	CGCYTTMGTCCACTCAGCCATCTCTCC	L06NF	GCAGCATTCCGAGTAACTCCTCAACC	L7B_GR22	GTAAACCCCAGAACATAAATTGTTACAC
An_L09FA	GTRTAGTTTAGTGGTAAAAGTGYGATTCG	L07B_01F22	CAGCTTCCATTGAGTCTCTGCACCTATCC	L89_GF22	GTAAACCCCAGAACATAAATTGTTACACAG
An_L09FB	GGATAGGACAKAGGTCTTCTAAACCTTTGG	L07B_F	GAGCATCRGCATGTAGGTTCCAGATCCAAG	L89_GR20	CAATCTCTAATTCTAATCTTCCTCCCC
An_L09FBR	CYAAAGGTTTAGAAGACCTMTGTCCTATCC	L07B_R	TGGGGATAGAGGGACTTGAACCCTCA	L89_GR22	TCTTCCGTCAAGCCCTGATCCATGAACC
An_L09R	AATTGGTCAGAGCACCGCCCTGTCAAGG	L07B_RA	AGCGTCTACCRATTTCGCCATATCCCC	LC97_G1NF	GACGCTTTAGTCCACTCAGCCATCTCTCC
An_L09RA	AAGGATTTGCAGTCCCCCGCCTTACCAC	L07B_RAF22	CTTTACGAATCAGAGCGCAATAATGTCTC	LC97_G2NR	GATTTGCAGTCCCCCGCCTTACCACTTG