Genomic diversity of 39 *Pyropia* species grown in Japan

Yukio Nagano[1]*, Kei Kimura[2]*, Genta Kobayashi[2], and Yoshio Kawamura[2]

[1]Analytical Research Center for Experimental Sciences, Saga University, 1 Honjo-machi, Saga, 840-8502, Japan

[2]Faculty of Agriculture, Saga University, 1 Honjo-machi, Saga, 840-8502, Japan

*Contributed equally

Abstract

Pyropia species, such as nori (*P. yezoensis*), are important marine crops. We analysed the genome sequences of 39 *Pyropia* species grown in Japan. Analysis of organellar genomes showed that the chloroplast and mitochondrial genomes of *P. yezoensis* have different evolutionary histories. The genetic diversity of Japanese *P. yezoensis* used in this study is lower than that of Chinese wild *P. yezoensis*. Nuclear genome analysis revealed nearly 90% of the analysed loci as allodiploid in the hybrid between *P. yezoensis* and *P. tenera*. Although the genetic diversity of Japanese *P. yezoensis* is low, analysis of nuclear genomes genetically separated each accession. Accessions isolated from the sea were often genetically similar to those being farmed. Study of genetic heterogeneity within a single strain of *P. yezoensis* showed that accessions with frequent abnormal budding formed a single, genetically similar group. The results will be useful for breeding and the conservation of *Pyropia* species.

Introduction

The genus *Pyropia* is a marine red alga belonging to the family Bangiaceae. Within the genus *Pyropia*, *P. yezoensis* Ueda (Susabi-nori in Japanese) and *P. haitanensis* Chang et Zheng (tan-zicai in Chinese) are economically important marine crops that are consumed in many countries.[1] Other species are cultivated or wild, and some are occasionally used as edibles. For example, *P. tenera* Kjellman (Asakusa-nori in Japanese) was once cultivated in Japan but has now been replaced by *P. yezoensis* in many places.

Studying the genetic diversity of *Pyropia* species is necessary for planning the breeding and conservation of these seaweeds. This information is also important for the use of species that have not been utilised previously. These approaches are critical to addressing climate change.

Prior to the genomic era, various studies analysed the genetic diversity of *Pyropia* species. For example, studies used the methods based on simple sequence repeat,[2,3] DNA sequencing of several genes,[4-22] amplified fragment length polymorphism,[23,24] and polymerase chain reaction-restriction fragment length polymorphism.[25] These previous studies analysed small regions of the genome.

However, few genomic approaches were applied to study the genetic diversity of *Pyropia* species. For example, to study the genetic diversity of *P. yezoensis* in China, high-throughput DNA sequencing analysed variations in the organellar genome.[26] This study divided wild *P. yezoensis* from Shandong Province, China, into three clusters. Several methods, such as restriction-site associated DNA sequencing,[27] genotyping by sequencing,[28] and whole-genome resequencing[29,30] use high-throughput DNA sequencing and can be applied to the analysis of the nuclear genome. Among them, whole-genome resequencing is the most effective way to get a comprehensive view of the entire nuclear genome. In general, whole-genome resequencing requires short reads generated by high-throughput DNA sequencing. The short reads can be appropriate in another application, the *de novo* assembly of the chloroplast genome.[31] Hence, short reads are useful for research on both the nuclear and chloroplast genomes.

In this study, we conducted a phylogenetic analysis of 39 *Pyropia* accessions grown in Japan using chloroplast DNA sequences assembled from short reads. We also used whole-genome resequencing to analyse the nuclear genomes of 34 accessions of *P. yezoensis*, one accession of the closely related *P. tenera*, and one hybrid between *P. yezoensis* and *P. tenera*. Particularly, we focused on the analysis of some kinds of seaweed in the Ariake sound, Japan, where seaweed production is the most active.

Results

Samples used in this study

Table 1 summarises the samples used in this study, and Figure 1 shows the isolation site of each sample. In addition to 34 accessions of *P. yezoensis*, we used *P. haitanensis*, *P. dentata* Kjellman, *P. tenuipedalis* Miura, *P. tenera*, and the hybrid between *P. yezoensis* and *P. tenera* in this study. Many are pure accessions because they were cultured from a monospore (a single haploid cell).

Chloroplast and mitochondrial genomes

To elucidate the relationships among the 39 accessions, we analysed their chloroplast DNA sequences. Using the GetOrganelle[31] program, which is a toolkit for assembling the organelle genome from short reads, we attempted to determine the whole chloroplast DNA sequences. Of the 39 total cases, we successfully assembled the whole chloroplast genome of 30 samples, and 9 cases (Pyr_16, Pyr_18, Pyr_21, Pyr_29, Pyr_23, Pyr_25, Pyr_28, Pyr_33, and Pyr_34) were unsuccessful. The chloroplast genomes of *Pyropia* species contain two direct repeats carrying ribosomal RNA (rRNA) operon copies.[26] The circular chloroplast genome has a large single-copy section and a short single-copy section between two rRNA repeats. Mapping of the short reads to the assembled chloroplast genome detected two types of reads in the rRNA operons (Supplementary Figure 1) in all 39 cases, reflecting the presence of two non-identical direct repeats carrying two different rRNA operons.[26] The presence of two types of reads in the rRNA operons might cause the unsuccessful assembly of the whole chloroplast genome of the 9 samples. Indeed, we succeeded in assembling the long single section sequences of all 39 samples. We detected the presence of two chloroplast genomes, termed heteroplasmy, only in the accession Pyr_45 (*P. haitanensis*) (Supplementary Figure 2).

Previous research reported that wild varieties of *P. yezoensis* from Shandong Province, China, can be classified into three clusters.[26] We also assembled the chloroplast genome of one accession of each cluster. By using the assembled sequences of large single copy sections and publicly available chloroplast sequences, we created a phylogenetic tree based on the maximum likelihood method (Figure 2). The number of parsimony informative sites was 24,914. The chloroplast sequences of all Japanese accessions of *P. yezoensis* were very similar to each other. The Chinese wild varieties of *P. yezoensis* were genetically similar to but clearly separated from the Japanese varieties. The chloroplast sequence of a Chinese cultivar (MK695880) collected from the nori farm in Fujian province was identical to that of some Japanese samples but not identical to those of the Chinese wild varieties.

Among *Pyropia* species, *P. tenera* is the closest relative to *P. yezoensis* at the DNA level.[20,21,22] In addition, they can be crossed each other.[32] In fact, Pyr_19 (*P. tenera*) was present on a branch next to *P. yezoensis*. The nucleotide sequence of the hybrid (Pyr_27) between *P. yezoensis* and *P. tenera* was more similar to that of *P. tenera* than that of *P. yezoensis*. Examination of the sequence revealed that, as expected, Pyr_45 was *P. haitanensis*. The analysis showed that Pyr_35 (*P. dentata*) is similar to *P.*

haitanensis, as in previous studies.[22] Pyr_44 (*P. tenuipedalis*) belonged to a clade containing *P. yezoensis*, which is also consistent with previous molecular analyses.[22]

We mapped short reads of *P. yezoensis* accessions from Japan and China to the chloroplast genome and constructed the maximum likelihood tree using the mapped data (Supplementary Figure 3). As similar to Figure 2, Japanese and Chinese chloroplast genomes were clearly separated. In addition, chloroplast sequences of Chinese *P. yezoensis* did not clearly separate them into three clusters. Chloroplast sequences showed that the genetic diversity of Japanese *P. yezoensis* used in this study is slightly lower than that of the Chinese wild *P. yezoensis*. We applied a similar approach to the mitochondrial genome (Figure 3). Mitochondrial sequences separated Chinese *P. yezoensis* into three clusters as reported previously.[26] Thus, mitochondrial sequences are determinant for clustering in the previous report.[26] The analysis did not separate Japanese from Chinese *P. yezoensis*. Rather, Japanese *P. yezoensis* is similar to cluster 2 of Chinese *P. yezoensis*. Thus, the chloroplast and mitochondrial genomes had different evolutionary histories. In addition, mitochondrial sequences showed that the genetic diversity of Japanese *P. yezoensis* used in this study is significantly lower than that of Chinese wild *P. yezoensis*.

Nuclear genomes

The chloroplast and mitochondrial sequences did not clearly discriminate each accession of Japanese *P. yezoensis*. Therefore, comparisons based on the nuclear genome are important. We analysed the nuclear genome of only *P. yezoensis* from Japan because the Chinese *P. yezoensis* have small amounts of published reads. We mapped the short reads from 39 samples to the reference genome of *P. yezoensis*. Supplementary Table 1 summarises the results. The mean coverage of Pyr_35 (*P. dentata*), Pyr_44 (*P. tenuipedalis*), and Pyr_45 (*P. haitanensis*) was 2.7, 8.8, and 4.2, respectively. As described above, the analysis of the chloroplast genome showed that these three accessions were not similar to *P. yezoensis*, which was reflected in the low mean coverage. Therefore, we did not use the data from these three accessions for further analysis based on the mapping. In contrast, the mean coverage of Pyr_19 (*P. tenera*) and Pyr_27, the hybrid between *P. yezoensis* and *P. tenera*, was 47.4 and 25.7, respectively. Therefore, we used the data of these two accessions for the subsequent analysis. The mean coverages for *P. yezoensis* ranged from 14.6 to 83.4. Before DNA extraction, we did not remove the bacteria co-cultured with seaweed. Differences in the extent of the range of mean coverage reflect the amount of these bacteria.

We called genetic variants from the mapping data by DeepVariant.[33] After merging the variant data of 36 samples, we filtered out low-quality data (GQ (Conditional genotype quality) < 20). The data included 697,892 variant sites. Subsequently, we analysed variant information by MDS analysis (Supplementary Figure 4). Accessions of *P. yezoensis* were clustered together.

The hybrid between *P. yezoensis* and *P. tenera*

The MDS analysis separated Pyr_27 from the Pyr_19 (*P. tenera*), although these two accessions were similar in chloroplast sequences. The analysis located the hybrid at approximately the middle position between two species, which supports the record that the hybrid was created by an interspecific cross between two species.

The previous analysis using a small number of genes reported that the hybrid between *P. yezoensis* and *P. tenera* is allodiploid.[34,35] We reanalysed this possibility using the whole genome data of Pyr_1 (*P. yezoensis*), Pyr_19 (*P. tenera*), and their hybrid Pyr_27. Both Pyr_1 and Pyr_19 should be homozygous diploid as they are pure accessions. The Pyr_27 is also a pure accession, but if it is an allodiploid, the genotype will be observed as heterozygous in the genome viewer. Figure 4a shows a representative example of a comparison of genotypes among three accessions. This figure showed that Pyr_27 has heterozygous-like data. To further analyse the tendency, we calculated the number of the types of loci in Pyr_27 when the Pyr_1 locus carries two reference alleles (homozygous locus) and when the Pyr_19 locus carries two alternate alleles (alternative homozygous locus) (Table 2). In about 90% of the cases, Pyr_27 had one reference allele and one alternate allele (heterozygous locus). Therefore, Pyr_27 must be an allodiploid. In the nuclear rRNA genes, the genotypes of Pyr_27 were identical to those of Pyr_19 (*P. tenera*) (Figure 4b). This result confirmed previous findings that this region became homozygous after conjugation.[34,35]

Genomic diversity of Japanese *P. yezoensis*

Of 34 accessions of *P. yezoensis*, 31 accessions were pure, so they should be homozygous. Therefore, we analysed these 31 accessions. From the variant data, we removed loci containing potentially heterozygous loci, which may be created due to the presence of non-identical repeats. This data included 69,918 variant sites. We then performed an MDS analysis (Figure 5) and an admixture analysis (Figure 6, Supplementary Figure 5) using this variant data. For the phylogenetic analysis, we created the tree using the data from 32 accessions, including 31 accessions of *P. yezoensis* and 1 accession of *P. tenera* as a root, all of which should be homozygous. After removal of heterozygous loci, the data included 681,115 variant sites, of which 33,985 were parsimony informative sites. We constructed a phylogenetic tree using the program SVDQuartet[36] (Figure 7). Under the coalescence model, this program assumes that intragenic recombination exists.

Although the analyses clearly separated each accession based on genetic distance, there were three closed clusters: cluster A containing 19 accessions (Pyr_1, 2, 3, 4, 7, 8, 10, 11, 12, 13, 14, 15, 17, 22, 23, 25, 28, 39, and 41), cluster B containing 2 accessions (Pyr_21 and 38), and cluster C containing 2 accessions (Pyr_16 .and 24) (Figure 5). The phylogenetic tree detected these closed clusters with high bootstrap supports (Figure 7). We will describe the results of the analysis of cluster A, containing 19 samples, in more detail later. The members of cluster B, Pyr_21 and 38, were independently isolated from an aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center,

but they originated in the same sound. In cluster C, Pyr_16 was isolated from an aquaculture farm at the Saga Prefectural Ariake Fisheries Research and Development Center, and Pyr_24 was provided as *P. tanegashimensis* previously. However, Pyr_24 was morphologically similar to *P. yezoensis.* For this reason, it appeared that the sample had been misplaced previously. Most of the samples used in this study were cultured or isolated in the Ariake sound, which is surrounded by Fukuoka, Saga, Nagasaki, and Kumamoto Prefectures (Figure 1), whereas Pyr_20, Pyr_24, Pyr_26, and Pyr_30 were from the sound in Hiroshima, Miyagi, Chiba, and Kagoshima Prefectures, respectively. In addition, Pyr_29 was from the Genkai Sea in Fukuoka Prefecture. (The Genkai Sea in northern Fukuoka Prefecture and the Ariake sound in southern Fukuoka Prefecture are separated by land.) Thus, the analyses did not clearly separate the seaweeds of other places from those of the Ariake sound. Pyr_20 was once marketed as *P. tenera*, but morphologically it is more likely to be *P. yezoensis*. Our genetic analysis confirmed that Pyr_20 is *P. yezoensis*.

Figure 6 and Supplementary Figure 5 show the results of the admixture analysis. Among the possible values of K (the number of ancestral populations), $K = 2$ was the most likely because $K = 2$ had the smallest cross-validation error (Supplementary Figure 5). Thus, the number of ancestral populations could be 2, with the 19 accessions shown in blue in Figure 6 belonging to one population and the six accessions shown in red in Figure 6 belonging to a second population. The former 19 accessions were cluster A, as described above. Six accessions, Pyr_18, 20, 21, 29, 33, and 38 might belong to an admixture population formed by the hybridisation between these two ancestral populations. However, it is also possible that K is 3 or 5 (Supplementary Figure 5). Because $K = 1$ had the highest cross-validation error, admixture has likely happened in these 39 accessions.

Because 19 samples belonging to cluster A were closely related, it was difficult to distinguish these samples in the above MDS analysis shown in Figure 5. Therefore, we analysed only these 19 accessions. The variant data from these 19 accessions included 36,459 variant sites. We used this data for MDS analysis (Figure 8). Twelve accessions (Pyr_1, 2, 3, 4, 7, 8, 10, 11, 12, 13, 14, and 15) are accessions isolated from the single strain S-18, which may contain heterogeneous cells. The S-18 strains have two subgroups: a subgroup containing seven accessions of Pyr_1, 10, 11, 12, 13, 14, and 15 and a subgroup containing five accessions of Pyr_2, 3, 4, 7, and 8. Phylogenetic analysis (Figure 7) supported this observation with high bootstrap values. Among the S-18 strain, Pyr_10, 11, 12, 13, 14, and 15 have an abnormal phenotype. In normal accessions, the rate of abnormal budding is about 15%, whereas in abnormal accessions it is about 40%. These 6 accessions formed the closely related cluster and were separated from Pyr_1, although these 7 accessions belonged to the former subgroup. In addition to the members of strain S-18, Pyr_17, 22, 23, 25, 28, 39, and 41 belonged to a clade containing 19 accessions. Of these, Pyr_17, 22, 23, 25, 39, and 41 were cultivated or isolated in the Ariake sound. Pyr_28 was from the sound in Ehime Prefecture.

To detect variants specific to abnormal accessions (Pyr_10, 11, 12, 13, 14, and 15), we extracted 69

candidate loci by analysing variant call data obtained with DeepVariant and then further selected the candidates by visual inspection (Supplementary Table 2). We also used the program FermiKit[37] to detect structural variants specific to abnormal accessions but failed to detect them. Furthermore, Pyr_36 is a green mutant. The similar strategies detected 104 candidate loci (Supplementary Table 3). In both cases, some of the loci were located within the gene or close to the gene. For example, deletion is detected in the calmodulin gene in the green mutant.

Discussion

A comparison of Japanese and Chinese accessions of *P. yezoensis* revealed that the chloroplast and mitochondrial genomes had different evolutionary histories (Figure 2, Figure 3, and Supplementary Figure 3). Incomplete lineage sorting[38] may cause this difference. However, we also need to consider the possibility that chloroplasts and mitochondria may have different inheritance mechanisms, which has not been studied in *P. yezoensis*. Therefore, hybridisation may have contributed to the occurrence of different evolutionary histories. However, further study is required to elucidate these two possibilities.

A comparison of Japanese and Chinese accessions of *P. yezoensis* revealed that the diversity of chloroplast and mitochondrial sequences in Japan is lower than that in China (Figure 2, Figure 3, and Supplementary Figure 3). This is especially evident in the case of mitochondrial DNA, which may be because most of the Japanese *P. yezoensis* used in this analysis were either cultivated or isolated from close to a nori farm. Therefore, it is important to investigate Japanese wild varieties, especially from sounds where farming is not performed, in the future. Since the farming of *P. yezoensis* has begun in Japan, studying their diversity is an attractive topic. In this study, we were able to distinguish the Japanese accessions by analysing the nuclear genomes.

We also examined the genetic changes that have occurred during an interspecific cross between *P. yezoensis* and *P. tenera*. Both chloroplast and mitochondrial genomes of the hybrid were similar to those of *P. tenera* (Figure 2, Figure 3, and Supplementary Figure 3). At least in this one conjugation event, a mechanism of maternal, paternal, or random DNA transmission might eliminate the chloroplast and mitochondrial genomes of *P. yezoensis*, one of two parental chloroplast genomes. However, there is the possibility that chloroplasts and mitochondria may have different inheritance mechanisms.

The changes that have occurred in the nuclear genome during an interspecific cross are more interesting. MDS analysis (Supplementary Figure 4) located the hybrid at approximately the middle position between two species. This is due to the allodiploid formation and not due to the DNA recombination between two species. We analysed the loci in the hybrid and found that when *P. yezoensis* had homozygous locus and the *P. tenera* had alternative homozygous locus, about 90% of the loci were allodiploid (Table 2). Thus, to our knowledge, the current study demonstrated the allodiploid formation at the genomic level for the first time, although the analysis of a small number of genes showed that the interspecific cross between *P. yezoensis* and *P. tenera* could produce an allodiploid.[34,35] The creation of interspecific hybrids is an important method in the breeding of seaweeds. The methods used in this study will be useful in the analysis of interspecific hybridisation after the conjugation because we can track changes in the nuclear genome. After the interspecific cross, the region of the nuclear rRNA became homozygous (Figure 4b). This was consistent with previous studies.[34,35] Thus, a mechanism to eliminate one of the two types of nuclear rRNA genes is present. In

contrast, there were two types of rRNA genes in the chloroplast genome.

The hybrid is slightly more similar to *P. yezoensis* than to *P. tenera* in the MDS analysis (Supplementary Figure 4). Furthermore, we analysed the loci in the hybrid and found that when the *P. yezoensis* had homozygous locus and the *P. tenera* had alternative homozygous locus, about 10% of the loci were identical to those of *P. yezoensis* (Table 2). In order to explain this observation, we must consider two possibilities. One possibility is that this study did not use the real parents of the hybrid. The *P. tenera* used in the cross might be a more genetically similar accession to *P. yezoensis*. Related to this possibility, the mitochondrial sequence of the hybrid was similar to but not identical to that of *P. tenera* used in this study (Figure 3). The second possibility is that recombination or other mechanisms eliminated the sequence of *P. tenera* drastically. Considering the two possibilities, it is important to analyse the genomes of the trio, the hybrid and its biological parents, immediately after performing an interspecific cross.

Our analysis revealed that genetically similar seaweeds were frequently used in the Ariake sound (Figure 5, Figure 8), although this fact had been unknown previously. Furthermore, it turns out that genetically similar seaweed had been repeatedly isolated by researchers. Cluster A is a typical example of the genomic similarity because cluster A contains accessions that could not be considered an S18 strain. Of the 19 ascensions of cluster A, all except one were cultivated in the Ariake sound or isolated from the Ariake sound. This may be a problem for the conservation of *P. yezoensis*. We need to examine whether the genetic diversity of *P. yezoensis* is maintained in sounds where farming is not performed.

Of the 19 ascensions of cluster A, Pyr_28 was isolated in the sound in Ehime Prefecture. It is possible that this was due to human activity. Seaweed farming began in the Ariake sound later than other areas. Therefore, ancestors of cluster A members may have been artificially brought into the Ariake sound from the other areas such as Ehime Prefecture. Seaweed farming is now very active in the Ariake sound. Therefore, we do not exclude the possibility that Pyr_28 was artificially brought into Ehime Prefecture from the Ariake sound and isolated in Ehime Prefecture. In any case, the actual situation will not be known until we analyse the genomes of various seaweeds in various parts of Japan.

The accessions in the S-18 strains were very similar. Despite the similarities, we were able to detect genetic differences among these. In other words, we were able to detect heterogeneity in the strain successfully. It is an interesting finding, for example, that the accessions having abnormal budding are in tight clusters. Six accessions with frequent abnormal budding formed a single, genetically similar closed cluster. These abnormal accessions probably diverged after an event in which a single mutation occurred. The heterogeneity of somatic cells has been studied in cancer research.[39] Similar studies are now possible in seaweed.

In this study, we detected variants specific to abnormal accessions (Supplementary Table 2) or to a green mutant (Supplementary Table 3). However, a number of candidate loci were present. Further

studies are needed to identify the responsible genes. Targeted deletion by genome editing is one of the ways to identify them.

In summary, we used high-throughput sequencing to examine the genomic diversity of *Pyropia* species. To our knowledge, this is the first study to examine the genomic diversity of *Pyropia* species at the genome level. The information obtained in this study could be used to develop a breeding and conservation plan. A variety of *Pyropia* species grow wild or are cultivated in Japan and around the world. It is essential to perform genomic studies of these seaweeds.

Methods

Materials

Blades (a mixture of four types of haploid cell) were cultured in the water of Ariake sound supplemented with modified nutrients SWM-III[40] at 17–22 ºC under 100 μmol/m^2/s (11-13:13-11 h LD). A monospore (a single haploid cell) was obtained from each blade according to a previous method.[20] Conchocelis (homozygous diploid cells) derived from a single monospore were cultured in the water of Ariake sound supplemented with modified nutrients SWM-III at 18 °C under 30 μmol/m^2/s (11:13 h LD).

DNA sequencing

Total nucleic acid, including genomic DNA and RNA, was extracted from the conchocelis of each strain using the DNAs-ici!-F (Rizo, Tsukuba, Japan) according to the instructions of the manufacturer, and total RNA was digested by using RNase A (NIPPON GENE, Tokyo, Japan). The quality of the isolated genomic DNA was checked by 1% agarose gel electrophoresis. The concentration of DNA was determined by Qubit dsDNA BR Assay Kit (Thermo Fisher, Foster City, CA, USA). The libraries were sequenced with 150 bp paired-end reads using NovaSeq 6000 (Illumina, San Diego, CA, USA) by Novogene (Beijing, China). Low-quality bases and adapter sequences from paired reads were trimmed using the Trimmomatic[41] (version 3.9) (ILLUMINACLIP:adapter_sequence:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:5:20 MINLEN:50).

Assembly-based analysis of chloroplast genome

Chloroplast DNA sequences were assembled from filtered reads or public reads using the GetOrganelle[31] program (version 1.6.4). For visual confirmation of the assembled result, filtered reads were aligned with the assembled genome by using the short read alignment program bowtie2[42] (version 2.3.5.1). Samtools[43,44] (version 1.9) was used to process aligned data. The aligned data were visually inspected by the Integrative Genomics Viewer[45] (version 2.5.0). Multiple alignments of the assembled chloroplast genome were performed by MAFFT[46] (version 7.455). A phylogenetic tree based on the maximum likelihood method was constructed by the program RAxML[47] (version 8.2.12) (-f = a, -x = 12,345, -p = 12,345, -N (bootstrap value) = 1,000, and -m = GTRGAMMA).

Mapping-based analysis of nuclear, chloroplast, and mitochondrial genome

The reference nuclear genome data of *P. yezoensis* were downloaded from the National Center for Biotechnology Information (NCBI; Assembly number: GCA_009829735.1 ASM982973v1). The reference genome sequence of chloroplast was the assembled sequence of Pyr_1, one of the *P. yezoensis* accessions used in this study. The reference mitochondrial genome data of *P. yezoensis* were

downloaded from the NCBI (accession number: MK695879).[26] Filtered reads of our data were aligned with the reference genome by using the program bowtie2.[42] For the analysis of the chloroplast and mitochondrial genomes, public data of wild varieties in China (NCBI Sequence Read Archive under the BioProject: PRJNA55033) were also used. Samtools[43,44] was used to process aligned data. The quality of aligned data was analysed using the Qualimap[48] program (version 2.2.1). From aligned data, DeepVariant[33] (version 0.9.0) was used to call the variants to make vcf files.[49] Vcf files were merged using GLnexus[50] (version 1.2.6) (--config DeepVariantWGS). To process and analyse vcf files, bcftools[44] (version 1.9), bgzip implemented in tabix[44] (version 0.2.5), vcffilter implemented in vcflib[51] (version 1.0.0_rc3), and "grep" and "awk" commands of Linux were used. Low-quality data (GQ (Conditional genotype quality) < 20) were filtered out by vcffilter. The aligned data were visually inspected by the Integrative Genomics Viewer.[45] FermiKit[37] (version 0.14.dev1) was used to detect the structural variations. Multidimensional scaling (MDS) analysis was conducted based on this vcf file using the SNPRelate[52] (version 1.20.1) program. For the admixture analysis, vcf file was processed by the program PLINK[53] (version 2-1.90b3.35) (--make-bed --allow-extra-chr --recode --geno 0.1). The resulting files were used to perform the admixture analysis using the program admixture[54] (version 1.3.0).

For the phylogenetic analysis of the chloroplast, the rRNA regions were removed from the vcf file. For the phylogenetic analysis of the chloroplast and mitochondrial genomes, the vcf files were converted to the fasta format file using the VCF-kit[55] (version 0.1.6), and a phylogenetic tree based on the maximum likelihood method was constructed by the program RAxML[47] (version 8.2.12) (-f = a, -x = 12,345, -p = 12,345, -N (bootstrap value) = 1,000, and -m = GTRGAMMA).

For the phylogenetic analysis of the nuclear genome, vcf files were converted to the fasta format file using the VCF-kit[55] (version 0.1.6) and then were converted to NEXUS format file using MEGA X[56] (version 10.0.5). Phylogenetic tree analysis was performed using SVDquartets[51] implemented in the software PAUP[57,58] (version 4.0a) using a NEXUS file. The parameters used for SVDquartets were as follows: quartet evaluation; evaluate all possible quartets, tree inference; select trees using the QFM quartet assembly, tree model; multi-species coalescent; handling of ambiguities; and distribute. The number of bootstrap analyses was 1,000 replicates. Pyr_19 was used as a root.

Public RNA sequencing data (NCBI Sequence Read Archive under accession numbers: SRR5891397, SRR5891398, SRR5891399, SRR5891400, and SRR6015124) were analysed using HISAT2[56] (version 2.2.0) and StringTie[60] (version 2.1.1) to determine the transcribed regions of the nuclear genome.

Data availability

Sequences are available at the DNA Data Bank of Japan Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/index_e.shtml; Accession no. DRA010178).

References

1. Mumford, T. F. & Miura, A. Porphyra as food: cultivation and economics in *algae and human affairs* (eds. Lemby, C.A. & Walland, J. R.) 87-117 (Cambridge University Press, 1988).
2. Bi, Y. H., Wu, Y. Y. & Zhou, Z. G. Genetic diversity of wild population of *Pyropia haitanensis* based on SSR analysis. *Biochem. Syst. Ecol.* **54**, 307-312 (2014).
3. Xu, Y. *et al.* Developing a core collection of *Pyropia haitanensis* using simple sequence repeat markers. *Aquaculture* **452**, 351-356 (2016).
4. Koh, Y. H. & Kim, M. S. DNA barcoding reveals cryptic diversity of economic red algae, *Pyropia* (Bangiales, Rhodophyta): description of novel species from Korea. *J. Appl. Phycol.* **30**, 3425-3434 (2018).
5. Dumilag, R. V. & Aguinaldo, Z. Z. A. Genetic differentiation and distribution of *Pyropia acanthophora* (Bangiales, Rhodophyta) in the Philippines. *Eur. J. Phycol.* **52**, 104-115 (2017).
6. Guillemin, M. L. *et al.* The bladed Bangiales (Rhodophyta) of the South Eastern Pacific: molecular species delimitation reveals extensive diversity. *Mol. Phylogenet. Evol.* **94**, 814-826. (2016).
7. Ramirez, M. E. *et al. Pyropia orbicularis sp. nov.* (Rhodophyta, Bangiaceae) based on a population previously known as *Porphyra columbina* from the central coast of Chile. *Phytotaxa* **158**, 133-153 (2014).
8. Milstein, D., Medeiros, A. S., Oliveira, E. C. & Oliveira, M. C. Native or introduced? A re-evaluation of *Pyropia* species (Bangiales, Rhodophyta) from Brazil based on molecular analyses. *Eur. J. Phycol.* **50**, 37-45 (2015).
9. Kavale, M. G., Kazi, M. A., Sreenadhan, N. & Singh, V. V. Morphological, ecological and molecular characterization of *Pyropia vietnamensis* (Bangiales, Rhodophyta) from the Konkan region, India. *Phytotaxa* **224**, 45-58 (2015).
10. Harden, L. K., Morales, K. M. & Hughey, J. R. Identification of a new marine algal species *Pyropia nitida* sp. nov. (Bangiales: Rhodophyta) from Monterey, California. *Mitochondrial DNA Part A* **27**, 3058-3062 (2016).
11. Lindstrom, S. C. An undescribed species of putative Japanese *Pyropia* first appeared on the central coast of British Columbia, Canada, in 2015. *Mar. pollut. Bull.* **132**, 70-73 (2018).
12. Kucera, H. & Saunders, G. W. A survey of Bangiales (Rhodophyta) based on multiple molecular markers reveals cryptic diversity. J. Phycol. **48**, 869-882 (2012).
13. Hwang, I. K. *et al.* Intraspecific variation of gene structure in the mitochondrial large subunit ribosomal RNA and cytochrome c oxidase subunit 1 of *Pyropia yezoensis* (Bangiales, Rhodophyta). *Algae-Seoul* **33**, 49-54 (2018).
14. Vergés, A., Comalada, N., Sánchez, N. & Brodie, J. (2013). A reassessment of the foliose Bangiales (Rhodophyta) in the Balearic Islands including the proposed synonymy of *Pyropia*

olivii with *Pyropia koreana*. Botanica Marina, 56(3), 229-240.

15. Dumilag, R. V. & Monotilla, W. D. Molecular diversity and biogeography of Philippine foliose Bangiales (Rhodophyta). *J. Appl. Phycol.* **30**, 173-186 (2018).
16. Xie, Z. Y., Lin, S. M., Liu, L. C., Ang, P. O. & Shyu, J. F. Genetic diversity and taxonomy of foliose Bangiales (Rhodophyta) from Taiwan based on *rbc*L and *cox*1 sequences. *Bot. Mar.* **58**, 189-202 (2015).
17. Dumilag, R. V. *et al.* Morphological and molecular confirmation of the occurrence of *Pyropia tanegashimensis* (Bangiales, Rhodophyta) from Palaui Is., Sta. Ana, Cagayan, Philippines. *Phytotaxa* **255**, 83-90. (2016).
18. Vergés, A., Sánchez, N., Peteiro, C., Polo, L. & Brodie, J. *Pyropia suborbiculata* (Bangiales, Rhodophyta): first records from the northeastern Atlantic and Mediterranean of this North Pacific species. *Phycologia* **52,** 121-129 (2013).
19. Choi, S. J., Kim, Y. & Choi, C. Chloroplast genome-based hypervariable markers for rapid authentication of six Korean *Pyropia* species. *Diversity* **11**, 220; 10.3390/d11120220 (2019).
20. Niwa, K., Kato, A., Kobiyama, A., Kawai, H. & Aruga, Y. Comparative study of wild and cultivated *Porphyra yezoensis* (Bangiales, Rhodophyta) based on molecular and morphological data. *J. Appl. Phycol.* **20**, 261-270 (2008).
21. Yang, L. E. *et al.* A molecular phylogeny of the bladed Bangiales (Rhodophyta) in China provides insights into biodiversity and biogeography of the genus *Pyropia*. *Mol. Phylogenet. Evol.* **120**, 94-102 (2018).
22. Sutherland, J. E. *et al*. A new look at an ancient order: generic revision of the Bangiales (Rhodophyta). *J. Phycol*. **47**, 1131-1151 (2011).
23. Cao, Y. *et al.* AFLP fingerprints of *Pyropia yezoensis* (Bangiales, Rhodophyta) populations revealed the important effect of farming protocols on genetic diversity. *Bot. Mar.* **61**, 141-147 (2018).
24. Niwa, K., Kikuchi, N., Iwabuchi, M. & Aruga, Y, Morphological and AFLP variation of *Porphyra yezoensis* Ueda form, narawaensis Miura (Bangiales, Rhodophyta). *Phycol. Res.* **52**, 180-190 (2004).
25. Abe, M. *et al.* Use of PCR-RFLP for the discrimination of Japanese *Porphyra* and *Pyropia* species (Bangiales, Rhodophyta). *J. Appl. Phycol.* **25**, 225-232 (2013).
26. Xu, K., Yu, X., Tang, X., Kong, F. & Mao, Y. Organellar genome variation and genetic diversity of Chinese *Pyropia yezoensis*. *Front. Mar. Sci.* **6**, 756; 10.3389/fmars.2019.00756 (2019).
27. Baird, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3**, e3376; 10.1371/journal.pone.0003376 (2008).
28. Elshire, R. J. *et al.* A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379; 10.1371/journal.pone.0019379 (2011).

29. Li, R. *et al.* SNP detection for massively parallel whole-genome resequencing. *Genome Res*. **19**, 1124-1132 (2009).
30. Huang, X. *et al*. High-throughput genotyping by whole-genome resequencing. *Genome Res*. **19**, 1068-1076 (2009).
31. Jin, J. J. *et al.* GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Preprint at https://www.biorxiv.org/content/10.1101/256479v4 (2018).
32. Ohme, M. Cross experiments of the color mutants in *Porphyra yezoensis* Ueda. *Jpn. J. Phycol.* **34**, 101-106 (1986).
33. Poplin, R. *et al.* A universal SNP and small-indel variant caller using deep neural networks. *Nat. Biotechnol*. **36**, 983-987 (2018).
34. Niwa, K., Kobiyama, A. & Sakamoto, T. Interspecific hybridization in the haploid blade-forming marine crop *Porphyra* (Bangiales, Rhodophyta): occurrence of allodiploidy in surviving F_1 gametophytic blades. *J Phycol* **46**, 693-702 (2010).
35. Niwa, K. & Sakamoto, T. Allopolyploidy in natural and cultivated populations of *Porphyra* (Bangiales, Rhodophyta). *J Phycol* **46**, 1097-1105 (2010).
36. Chifman, J. & Kubatko, L. Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**, 3317–3324 (2014).
37. Li, H. FermiKit: assembly-based variant calling for Illumina resequencing data. *Bioinformatics* **31**, 3694-3696 (2015).
38. Maddison, W. P. & Knowles, L. L. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol*. **55**, 21-30 (2006).
39. Fisher, R., Pusztai, L. & Swanton, C. Cancer heterogeneity: implications for targeted therapeutics. *Br. J. Cancer* **108**, 479-485 (2013).
40. Ogata, E. On a new algal culture medium SWM-III. *Bull. Jap. Soc. Phycol.* **18**, 171-173 (1970)
41. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30**, 2114–2120 (2014).
42. Langmead, B. & Salzberg, S. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357-359 (2012).
43. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078-2079 (2009).
44. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993 (2011).
45. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26 (2011).
46. Katoh, K. & Standley D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772-780 (2013).

47. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
48. García-Alcalde, F. *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678-2679 (2012).
49. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
50. Yun, T., Li, H., Chang, P. C., Lin, M. F., Carroll, A. & McLean, C. Y. Accurate, scalable cohort variant calls using DeepVariant and GLnexus. Preprint at https://www.biorxiv.org/content/10.1101/2020.02.10.942086v1 (2020).
51. Garrison E. Vcflib: A C++ library for parsing and manipulating VCF files. https://github.com/ekg/vcflib (2012).
52. Zheng, X. *et al.* A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012).
53. Chang, C. C. *et al.,* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, s13742-015; 10.1186/s13742-015-0047-8 (2015).
54. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
55. Cook, D. E. & Andersen, E. C. VCF-kit: assorted utilities for the variant call format. *Bioinformatics*, **33**, 1581-1582 (2017).
56. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol*. **35**, 1547–1549 (2018).
57. Swofford, D. L. PAUP* Phylogenetic Analysis Using Parsimony (* and Other methods) Version 4. (Sinauer Associates, 1998).
58. Swofford, D. L. PAUP* (*Phylogenetic Analysis Using PAUP), http://paup.phylosolutions.com/ (2020).
59. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907-915 (2019).
60. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019).

Acknowledgements

We would like to thank H. Matsuo for technical assistance with the experiments. We would also like to thank Daiichi Seimo Co. Ltd., Saga Prefectural Ariake Fisheries Research and Development Center, Ariake Regional Laboratory of Fukuoka Fisheries and Marine Technology Research Center, Saga Prefectural Ariake Fishery Cooperative, and the National Federation of Nori & Shellfish-fishers cooperative associations for providing us with the accessions of *Pyropia* species used in this study. This study was supported by the “Projects for sophistication of production and utilisation technology

supporting local agriculture and marine industry" from Saga University and KAKENHI (18K19235) from the Ministry of Education, Culture, Science, Sports and Technology (MEXT) of Japan. We would like to thank Editage (www.editage.com) for English language editing.

Author Contributions

Y.N., K.K. and Y.K. designed the research. K.K and Y. K. performed the experiments. Y.N. performed the bioinformatic analysis. Y.N., K.K., G.K. and Y.K. analysed the data. Y.N. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Figure Legends

Figure 1. Positions of the isolation sites. The base map was produced using SimpleMappr (https://www.simplemappr.net/).

Figure 2. Phylogenetic tree of *Pyropia* species using long single section sequences of chloroplast DNA based on maximum likelihood reconstruction. The numbers at the nodes indicate bootstrap values (% over 1000 replicates). The scale bar shows the number of substitutions per site. The sequence of *Porphyra umbilicalis* was used as a root. The large figure is a phylogram showing the relationships among all the data used in this analysis. The small figure in the upper left is a cladogram showing the relationships for *P. yezoensis* only. Colours were used to distinguish between Chinese and Japanese *P. yezoensis*.

Figure 3. Phylogenetic tree of *Pyropia yezoensis* accessions using the mitochondrial DNA sequences based on maximum likelihood reconstruction. The scale bar shows the number of substitutions per site. The sequences of Pyr_19 (*P. tenera*) and the hybrid (Pyr_27) between *P. yezoensis* and *P. tener*a were used as roots. Colours were used to distinguish between Japanese *P. yezoensis* and 3 clusters of Chinese *P. yezoensis*.

Figure 4. Comparison of the nuclear genotypes of Pyr_1 (*Pyropia yezoensis*), Pyr_27 (the hybrid between *P. yezoensis* and *P. tener*a), and Pyr_19 (*P. tenera*) visualised by the Integrative Genomics Viewer. a) a representative example of allodiploid formation. b) nuclear rRNA regions.

Figure 5. Multidimensional scaling representation using nuclear DNA data of 31 *Pyropia yezoensis* accessions. Two-dimensional data were obtained in this analysis. Colours were used to show 3 clusters, cluster A, B, and C, found in this study.

Figure 6. Admixture analysis using nuclear DNA data of 31 *Pyropia yezoensis* accessions. The number of populations (K) was set to 2. Colours of the accession names were used to show 3 clusters, cluster A, B, and C.

Figure 7. Phylogenetic tree constructed using the SVDquartets with PAUP using nuclear DNA data of 31 *Pyropia yezoensis* accessions and 1 *P. tenera*. The numbers at the nodes indicate bootstrap values (% over 1,000 replicates). The data of Pyr_19 (*P. tenera*) was used as a root. Colours of the accession

546 names were used to show 3 clusters, cluster A, B, and C.

547

548 **Figure 8.** Multidimensional scaling representation using nuclear DNA data of 19 *Pyropia yezoensis*
549 accessions belonging to the cluster A. Two-dimensional data were obtained in this analysis. Members
550 of the strain S-18 were shown in red.

551

Table 1. List of *Pyropia* accessions

Accession name	Purification to became homozygous	Species	Isolation site	Isolation year	Notes
Pyr_1	+	*P. yezoensis*, accession 1	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, variety recommended by Saga Prefecture Fishery Cooperative Federation named as 'Shin Saga 4 gou'.
Pyr_2	+	*P. yezoensis*, accession 2	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, variety recommended by Saga Prefecture Fishery Cooperative Federation named as 'Shin Saga 4 gou'. Preservation place is different from that of Pyr_1.
Pyr_3	+	*P. yezoensis*, accession 3	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, reisolated from Pyr_2 in 2017.
Pyr_4	+	*P. yezoensis*, accession 4	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, reisolated from Pyr_2 in 2017.
Pyr_7	+	*P. yezoensis*, accession 7	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, reisolated from Pyr_2 in 2017.
Pyr_8	+	*P. yezoensis*, accession 8	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, reisolated from Pyr_2 in 2017.
Pyr_9		*P. yezoensis*, accession 9	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, normal phenotype, reisolated from Pyr_2 in 2017.
Pyr_10	+	*P. yezoensis*, accession 10	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, abnormal phenotype, reisolated from Pyr_2 in 2017.
Pyr_11	+	*P. yezoensis*, accession 11	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, abnormal phenotype, reisolated from Pyr_2 in 2017.
Pyr_12	+	*P. yezoensis*,	Ariake sound, Saga	2010	S-18 strain, abnormal phenotype, reisolated from Pyr_2 in 2017.

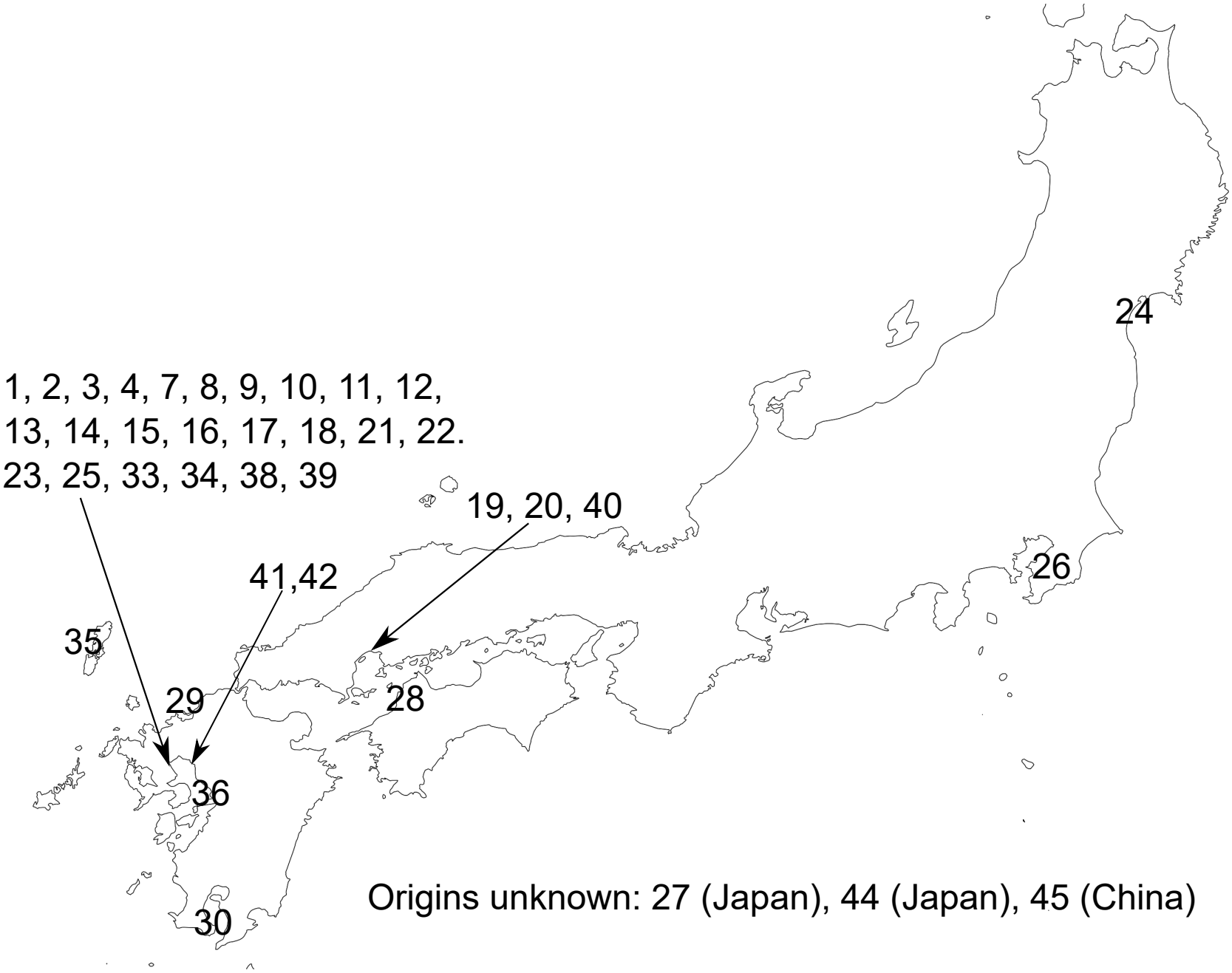
		accession 12	Prefecture, Japan		
Pyr_13	+	*P. yezoensis*, accession 13	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, abnormal phenotype, reisolated from Pyr_2 in 2017.
Pyr_14	+	*P. yezoensis*, accession 14	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, abnormal phenotype, reisolated from Pyr_2 in 2017.
Pyr_15	+	*P. yezoensis*, accession 15	Ariake sound, Saga Prefecture, Japan	2010	S-18 strain, abnormal phenotype, reisolated from Pyr_2 in 2017.
Pyr_16	+	*P. yezoensis*, accession 16	Ariake sound, Saga Prefecture, Japan	2015	Isolated from the aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center.
Pyr_17	+	*P. yezoensis*, accession 17	Ariake sound, Saga Prefecture, Japan	1999	Isolated as probable low-temperature resistant strain.
Pyr_18	+	*P. yezoensis*, accession 18	Ariake sound, Saga Prefecture, Japan	2017	Isolated from the aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center.
Pyr_19	+	*P. tenera*	Hiroshima Prefecture, Japan	1978	Provided from Saga Prefecture Fishery Cooperative Federation.
Pyr_20	+	*P. yezoensis*, accession 20	Hiroshima Prefecture, Japan	1978	Marketed as *P. tenera*, but morphologically similar to *P. yezoensis*.
Pyr_21	+	*P. yezoensis*, accession 21	Ariake sound, Saga Prefecture, Japan	2017	Isolated from the aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center.
Pyr_22	+	*P. yezoensis*, accession 22	Ariake sound, Saga Prefecture, Japan	1997	Variety recommended by Saga Prefecture Fishery Cooperative Federation ‘Shin Saga 1 gou’.
Pyr_23	+	*P. yezoensis,* accession 23	Ariake sound, Saga Prefecture, Japan	2014	Isolated from the aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center.

Pyr_24	+	*P. yezoensis*, accession 24	Miyagi Prefecture, Japan	2018	Provided as *P. tanegashimensis*, but morphologically similar to *P. yezoensis*.
Pyr_25	+	*P. yezoensis*, accession 25	Ariake sound, Saga Prefecture, Japan	1984	Variety recommended by Saga Prefecture fishery cooperative federation named as 'Hagakure'.
Pyr_26	+	*P. yezoensis*, accession 26	Chiba Prefecture, Japan	1983	Considered as *P. yezoensis* f. narawaensis.
Pyr_27	+	*P. tenera* × *P. yezoensis*	Japan	2005	Crossed by National Federation of Nori & Shellfish-fishers cooperative Associations named as 'Gyoko'.
Pyr_28	+	*P. yezoensis*, accession 28	Ehime Prefecture, Japan	1976	Isolated by Nagasaki University. Used in Fukuoka Prefecture as 'Ariake 1 gou'.
Pyr_29	+	*P. yezoensis*, accession 29	Genkai Sea, Fukuoka Prefecture, Japan	1975	Variety recommended by Saga Prefecture Fishery Cooperative Federation named as 'Saga 1 gou'.
Pyr_30	+	*P. yezoensis*, accession 30	Kagoshima Prefecture, Japan	1978	Variety recommended by Saga Prefecture Fishery Cooperative Federation named as 'Saga 6 gou'.
Pyr_33	+	*P. yezoensis*, accession 33	Ariake sound, Saga Prefecture, Japan	1982	Variety recommended by Saga Prefecture Fishery Cooperative Federation named as 'Saga 10 gou'.
Pyr_34	+	*P. yezoensis*, accession 34	Ariake sound, Saga Prefecture, Japan	2017	Isolated from the aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center.
Pyr_35		*P. dentata*,	Tsushima Island, Nagasaki Prefecture, Japan	2009	Isolated by the Saga Prefectural Ariake Fisheries Research and Development Center.
Pyr_36		*P. yezoensis*, accession 36	Ariake sound, Kumamoto Prefecture,	2004	Isolated from Amakusa district in Kumamoto Prefecture by Saga Prefectural Ariake Fisheries Research and Development Center.

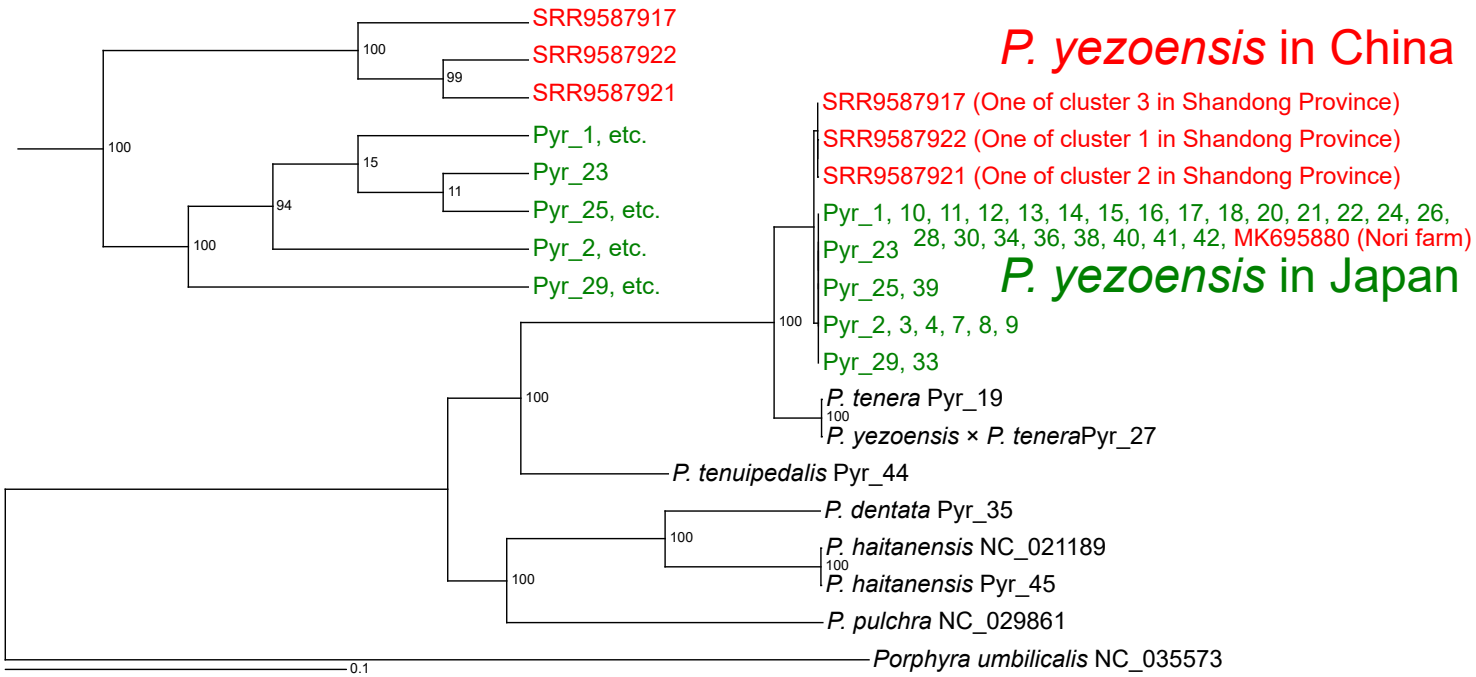
			Japan		
Pyr_38	+	*P. yezoensis*, accession 38	Ariake sound, Saga Prefecture, Japan	2000	Green mutant, isolated from the aquaculture farm of the Saga Prefectural Ariake Fisheries Research and Development Center.
Pyr_39	+	*P. yezoensis*, accession 39	Ariake sound, Saga Prefecture, Japan	2009	Variety recommended by Saga Prefecture fishery cooperative federation named as 'Shin Saga 3 gou'.
Pyr_40		*P. yezoensis*, accession 40	Hiroshima Prefecture, Japan	1978	Marketed as *P. tenera*, but morphologically similar to *P. yezoensis*.
Pyr_41	+	*P. yezoensis*, accession 41	Ariake sound, Fukuoka Prefecture, Japan	2002	Variety named as 'Fukuoka Ariake 1 gou'.
Pyr_42	+	*P. yezoensis*, accession 42	Ariake sound, Fukuoka Prefecture, Japan	1981	Variety named as 'Fukuoka 1 gou'.
Pyr_44	+	*P. tenuipedalis*	Japan	2000	Provided by Daiichi Seimo Co., Ltd.
Pyr_45	+	*P. haitanensis*	China	2000	Provided from Daiichi Seimo Co., Ltd.

Table 2. Number of loci in Pyr_27, when Pyr_1 locus carries two reference alleles and Pyr_19 locus carries two alternate alleles

Two reference alleles	One reference allele/one alternate allele	Two alternate alleles
18,484	139,259	141

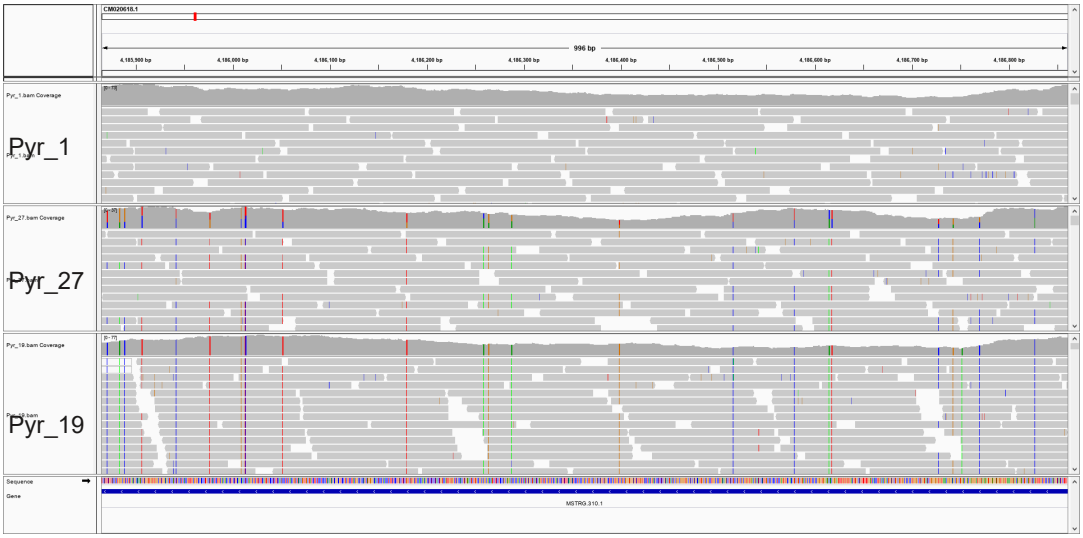


Origins unknown: 27 (Japan), 44 (Japan), 45 (China)

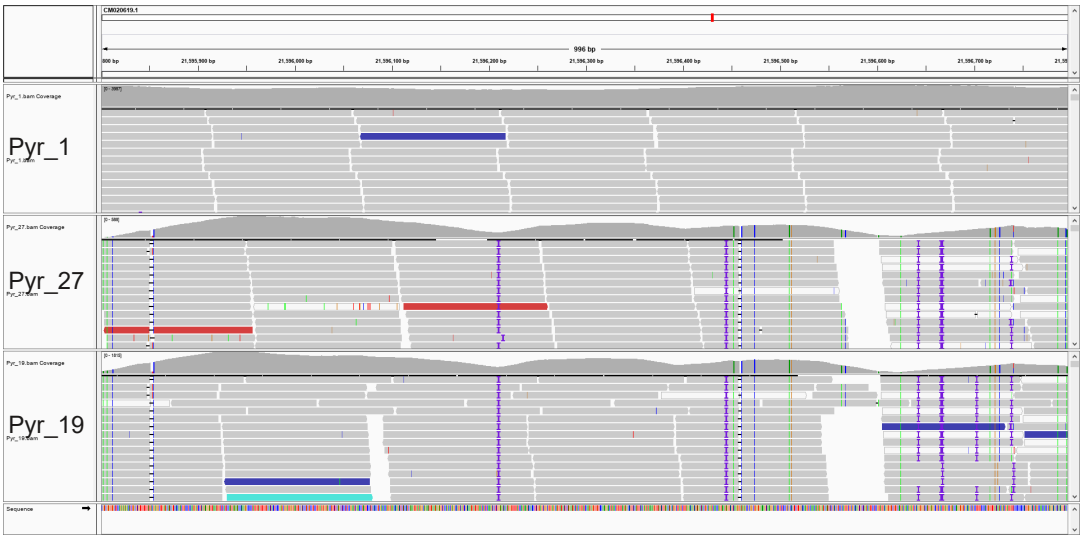

SRR9587917
SRR9587922
SRR9587921
Pyr_1, etc.
Pyr_23
Pyr_25, etc.
Pyr_2, etc.
Pyr_29, etc.
100
99
100
15
11
94
100
P. yezoensis in China
SRR9587917 (One of cluster 3 in Shandong Province)
SRR9587922 (One of cluster 1 in Shandong Province)
SRR9587921 (One of cluster 2 in Shandong Province)
Pyr_1, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 24, 26, 28, 30, 34, 36, 38, 40, 41, 42, MK695880 (Nori farm)
Pyr_23
Pyr_25, 39
Pyr_2, 3, 4, 7, 8, 9
Pyr_29, 33
P. yezoensis in Japan
100
P. tenera Pyr_19
100
P. yezoensis × P. teneraPyr_27
100
P. tenuipedalis Pyr_44
P. dentata Pyr_35
100
P. haitanensis NC_021189
100
P. haitanensis Pyr_45
100
P. pulchra NC_029861
Porphyra umbilicalis NC_035573
0.1

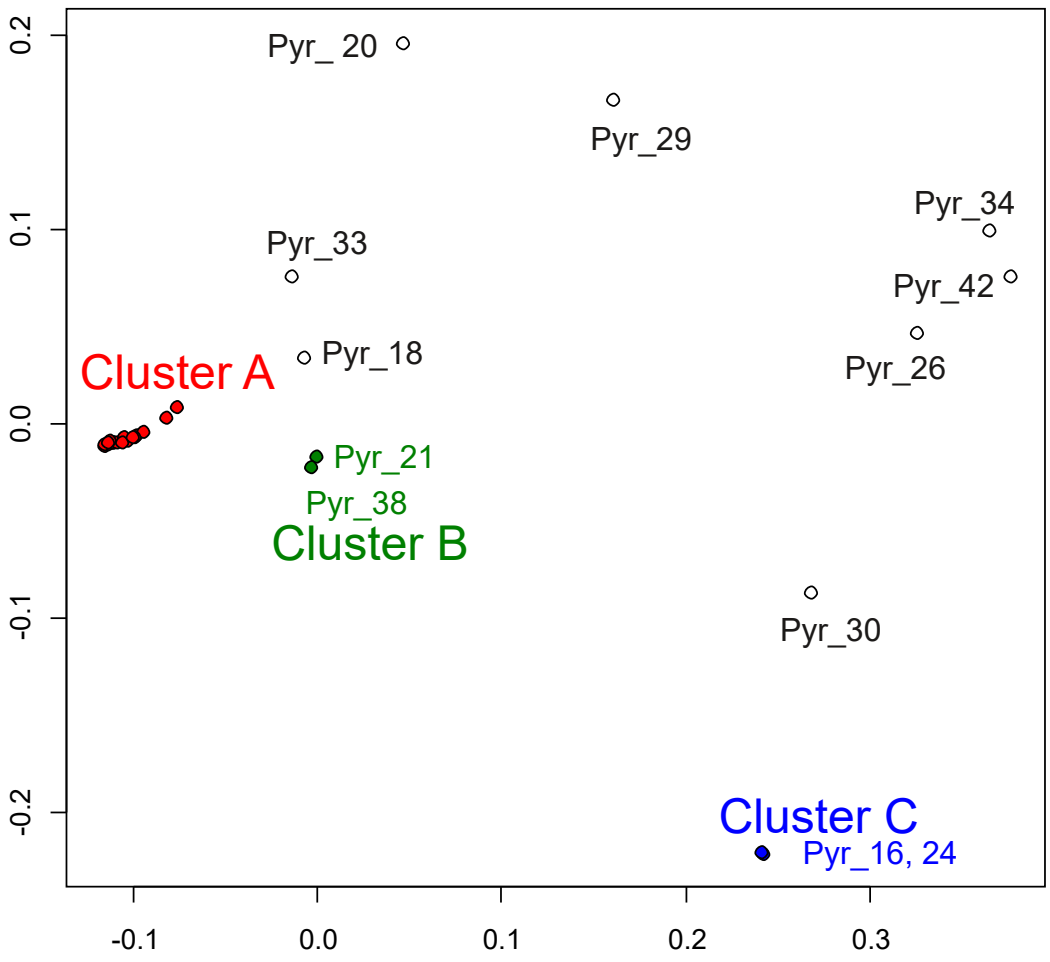
Cluster 3 in China
Cluster 1 in China
Cluster 2 in China
Cluster in Japan
Pyr 19
Pyr 27
0.1

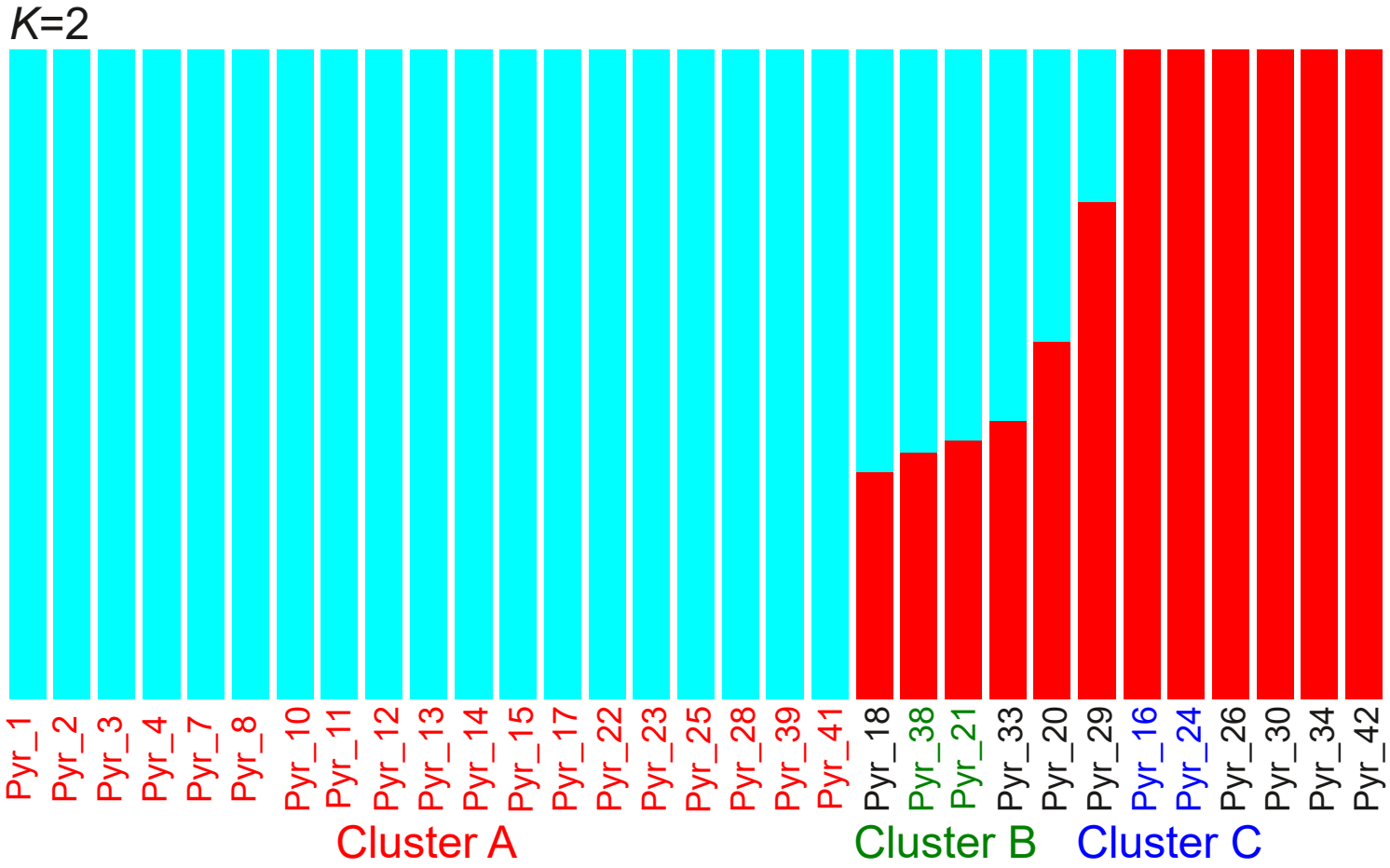
a)

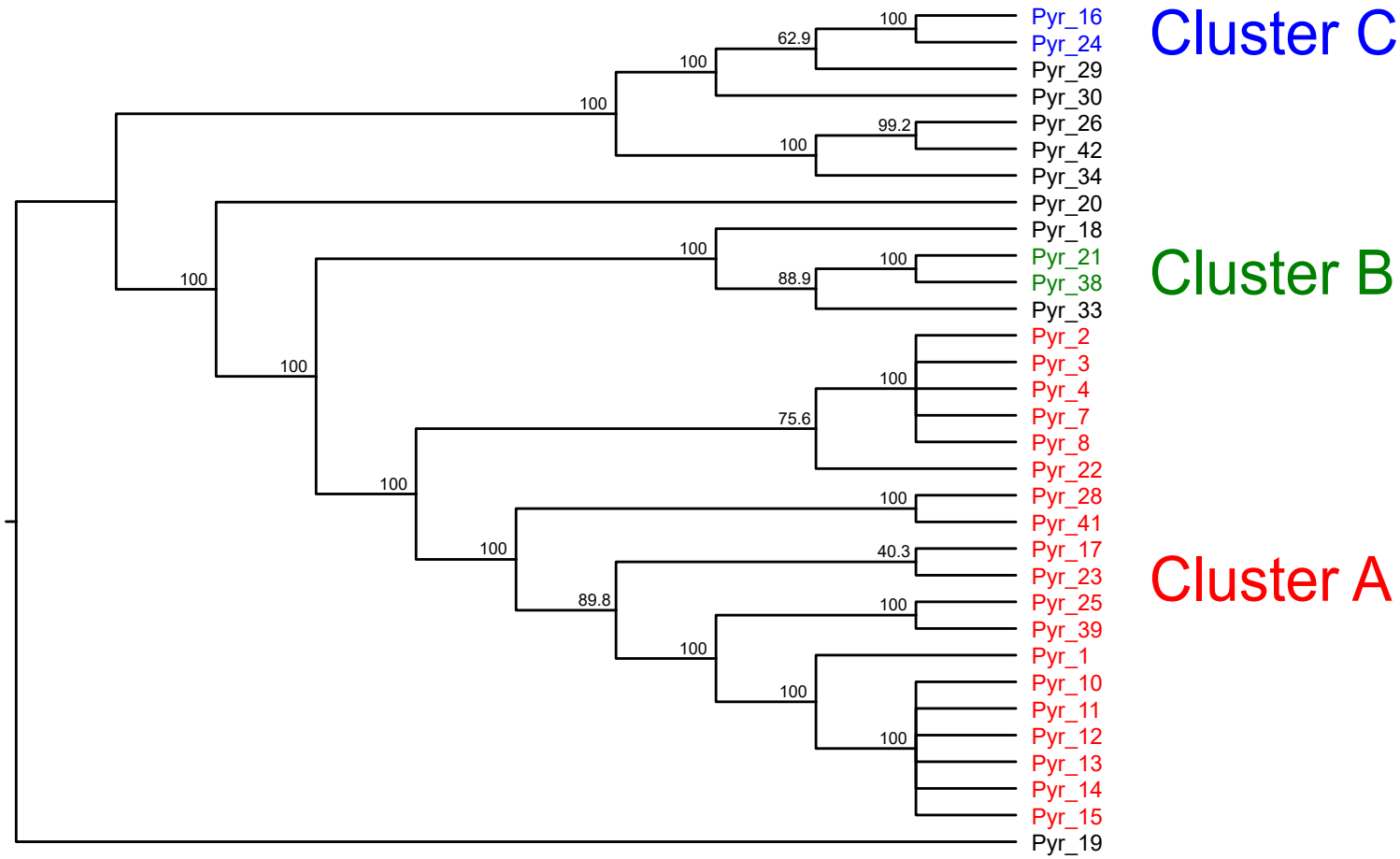

CM020610.1
996 bp
Pyr_1
Pyr_27
Pyr_19
Sequence
Gene
MSTRG.310.1

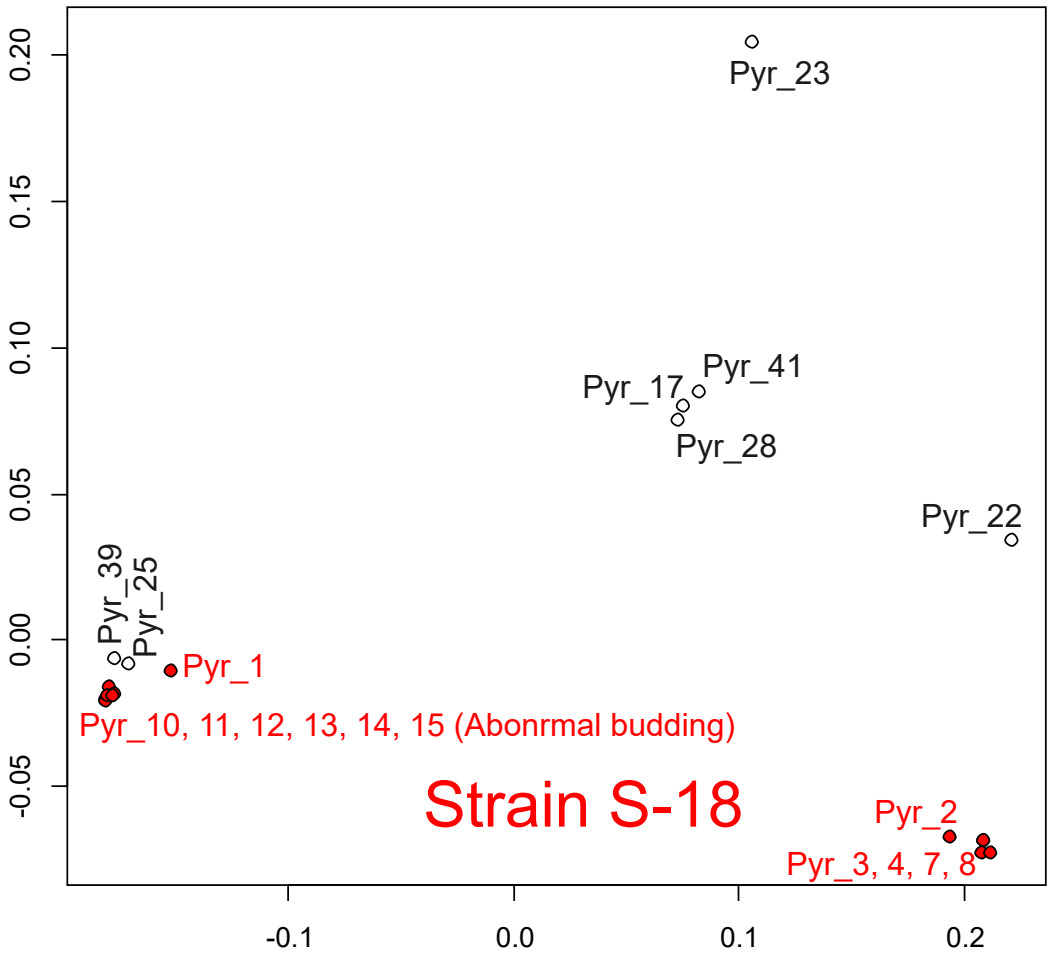
b)


CM020610.1
996 bp
Pyr_1
Pyr_27
Pyr_19
Sequence


Pyr_ 20
Pyr_29
Pyr_34
Pyr_33
Pyr_42
Pyr_18
Pyr_26
Cluster A
Pyr_21
Pyr_38
Cluster B
Pyr_30
Cluster C
Pyr_16, 24
0.2
0.1
0.0
-0.1
-0.2
-0.1
0.0
0.1
0.2
0.3


K=2
Pyr_1
Pyr_2
Pyr_3
Pyr_4
Pyr_7
Pyr_8
Pyr_10
Pyr_11
Pyr_12
Pyr_13
Pyr_14
Pyr_15
Pyr_17
Pyr_22
Pyr_23
Pyr_25
Pyr_28
Pyr_39
Pyr_41
Pyr_18
Pyr_38
Pyr_21
Pyr_33
Pyr_20
Pyr_29
Pyr_16
Pyr_24
Pyr_26
Pyr_30
Pyr_34
Pyr_42
Cluster A
Cluster B
Cluster C


Pyr_16
Pyr_24
Pyr_29
Pyr_30
Pyr_26
Pyr_42
Pyr_34
Pyr_20
Pyr_18
Pyr_21
Pyr_38
Pyr_33
Pyr_2
Pyr_3
Pyr_4
Pyr_7
Pyr_8
Pyr_22
Pyr_28
Pyr_41
Pyr_17
Pyr_23
Pyr_25
Pyr_39
Pyr_1
Pyr_10
Pyr_11
Pyr_12
Pyr_13
Pyr_14
Pyr_15
Pyr_19
Cluster C
Cluster B
Cluster A
100
62.9
100
100
99.2
100
100
100
100
88.9
100
100
100
75.6
100
100
100
40.3
89.8
100
100
100
100


Pyr_23
Pyr_41
Pyr_17
Pyr_28
Pyr_22
Pyr_39
Pyr_25
Pyr_1
Pyr_10, 11, 12, 13, 14, 15 (Abonrmal budding)
Strain S-18
Pyr_2
Pyr_3, 4, 7, 8
0.20
0.15
0.10
0.05
0.00
-0.05
-0.1
0.0
0.1
0.2