九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Population genetics systematic studies of pinewood nematode, Bursaphelenchus xylophilus in Kyushu

張, 涵泳

https://hdl.handle.net/2324/2236304

出版情報：九州大学, 2018, 博士（農学）, 課程博士
バージョン：
権利関係：

Population genetic and systematic studies of pinewood nematode, *Bursaphelenchus xylophilus* in Kyushu

Dissertation

Zhang Hanyong

2019

Kyushu University

Contents

Chapter 1. General Introduction

1.1. Preface

Pine wilt disease (PWD) is one of the most serious threats to forests around the world. It originated in North America and spread to many other countries over the course of the 20th century. In the USA and Canada, PWD is not considered a primary pathogen of native pines, whereas in invasion areas, such as Japan, Korea, China, and Portugal, it exerts serious damage to forest ecosystems. It has now become a worldwide threat to forest ecosystems and international trade. PWD was first recorded in 1905 at Nagasaki Prefecture, in the northwestern part of the island of Kyushu in Japan (Yano, 1913), but the pinewood nematode (PWN), *Bursaphelenchus xylophilus*, was not identified as the causal agent of this disease until 1971. Since the pathogenicity of PWD was first identified by Kiyohara and Tokushige (1971), the disease has attracted more and more attention from researchers all over the world. Thanks to the investigations into PWD carried out by scientists around the world, we have a better understanding of the disease, including transmission routes around the world, diagnostics of the nematode, associated insect vectors, interactions between the nematode and its host tree, pathogenicity and associated bacteria, and population genetic variation for the nematode and development of disease-resistant pines. However, the many mechanisms behind the disease's symptoms remain unknown, and the spread of this terrible disease has not yet been prevented.

Understanding its mechanisms and pathways of transmission is believed to be effective in controlling the reproduction of PWNs, especially in uninfected areas. DNA analysis has been employed in several studies and molecular markers showing sufficient genetic polymorphism are wildly utilized in the population analysis of PWN via such methods as amplified fragment length polymorphism (AFLP), internal transcribed spacer

(ITS)-restriction fragment length polymorphism (RFLP) analysis, random amplified polymorphic DNA (RAPD)-PCR, and microsatellite markers (SSR).

All these DNA-based techniques provide an attractive solution for examining PWN populations, despite their limitations, as these DNA markers capture only mutations at specific sites on the sequence, resulting in a failure to comprehensively analyze variations. Furthermore, most of the DNA markers are only associated with the nuclear genome, so population studies performed with these methods can be seen as one-sided analyses of the nucleus. However, genetic material is not only present in the nuclei of eukaryotes, but also in mitochondria and chloroplasts. Therefore, we sought to develop a method to utilize all mutations in the nucleotide sequence and conduct an exhaustive population analysis from the two cell organelles that contain genetic material in PWN, the cell nuclear genome and mitochondria genome.

The development of next generation sequencing technologies has made it possible to obtain a large amount of sequence information in order to capture sufficient variation. Moreover, the entire nuclear genome sequence has already been determined for PWN by Kikuchi et al. (2011). In addition, numerous Expressed Sequence Tags (ESTs) have been registered in the DNA database based on gene expression analysis, making it possible to utilize Single Nucleotide Polymorphism (SNP) information in PWN. Furthermore, the whole mitochondrial genome sequence of PWN was reported by Sultana et al. (2013), thereby providing valid information for studying PWN populations' variability based on mitochondrial genome sequencing and analysis of its polymorphisms. All those previous studies enabled us to put our ideas into practice. We believe that this research can supply a large amount of novel genomic information on PWD, as well as valuable recommendations for population analysis of the PWN.

1.2. Literature review

1.2.1. Studies on nematodes

The phylum Nematoda is considered by many to be one of the most abundant and potentially speciose invertebrate groups in the world (Blaxter et al., 1998; Giblin-Davis et al., 2004). However, because of the their mostly microscopic size, conserved gross morphology, vastly inadequate inventories, and poor or uneven descriptions, nematodes are not as recognized as they deserve to be (Ye et al., 2007). Nevertheless, many researchers are still working hard on these organisms. For example, a molecular evolutionary framework for the phylum Nematoda has been clarified (Blaxter et al., 1998). Underwood and Bianco (1999) identified a molecular marker for the Y chromosome of *Brugia malayi*. Moreover, a single nucleotide polymorphism map of mitochondrial genome of the parasitic nematode *Cooperia oncophora* has been reported (Van der Veer and de Vries, 2004). Some researchers have used molecular markers to analyze the phylogenetics of some genera of nematodes, including *Pratylenchus* (Al-Banna et al., 1997), *Meloidogyne* (De Ley et al., 2002), *Heterorhabditidae* and *Steinernematidae* (Liu et al., 1997), and *Bactrocera* species (Smith et al., 2003).

Among the nematodes studied, one of the most popular species is a kind of plant parasite called PWN, *Bursaphelenchus xylophilus* (Steiner and Buhrer, 1934), *Nickle*, 1970 (Nematoda: Aphelenchoididae), which causes cell destruction leading to death of the host tree within a few months (Cardoso et al., 2012). This species was identified as the agent of the worldwide forest blight called PWD (Kiyohara and Tokushige, 1971).

1.2.2. Studies on PWD

PWD is thought to be native to North American countries (Dwinell, 1997), but is not considered a primary pathogen of native pines (Leal et al., 2013). In contrast, in areas where it is invasive, such as Korea, Japan, Portugal, and China, it afflicts serious damage to forest

ecosystems (Yun et al., 2012). PWD was first recorded in 1905 in Japan, but PWN was not identified as the causal agent of the disease until 1971 (Shi et al., 2013). In Japan, PWD has plagued pine forests for over a century, accounting for the loss of 700 000 m^3 of pine wood each year (Mamiya and Shoji, 2009). In China, ever since it was first found in 1982 in Sun Yat-sen's Mausoleum in Nanjing, Chinese PWD has rapidly expanded and spread to 14 provinces, including Anhui, Chongqing, Fujian, Guangdong, Guangxi, Guizhou, Hubei, Hunan, Jiangsu, Jiangxi, Shandong, Sichuan, Yunnan, and Zhejiang (Yang, 1995; Zhang and Luo, 2003; Wu, 2004). By 2013, it had killed more than 1 million ha of pine forests with losses of more than one hundred billion denominated in Chinese currency (Shi et al., 2013). In South Korea, PWD was first identified in the Gumjung Mountain Region of Busan in 1988 (Yi et al., 1989), and has spread widely, reaching 57 cities and prefectures by 2010 (Jung et al., 2010a). Outside of Asia, in 1999, PWN was first reported in Portugal in Europe (Mota et al., 1999), where over the course of approximately ten years, nearly 510,000 ha of the pine forests were destroyed (Valadas et al., 2012). By 2008, new areas for the disease were established in other regions of Portugal (Rodrigues, 2008), and more recently, PWN has been detected on Maderia, 1,000 km southwest of continental Portugal (Fonseca et al., 2012). Furthermore, it has also been detected in Spain (Abelleira et al., 2011; Robertson et al., 2011). Across the whole world, 44 species in the genus *Pinus* (23 species occurring in Japan, 20 in the USA, two in Korea and nine in China) have been reported to be infected by PWN under natural conditions (Shi et al., 2013). Obviously, PWD has expanded almost all over the world, and so far, the PWN has been listed as a quarantine pest in more than 40 countries (Mota et al., 1999; Schrader and Unger, 2003).

1.2.3. Studies concerning *Bursaphelenchus xylophilus*

The studies about *B. xylophilus* were mainly focus on three areas as follows:

Studies on B. xylophilus

Although PWD is the most threatening to pine trees, there are also other factors which can cause pines to wilt, such as physiological disturbances, drought, forest fires, or competition between plant species (Yun et al., 2012). Therefore, it is important to confirm exactly whether the tree was infected with PWN or not. Also, a species closely related and morphologically similar to *B. xylophilus* within genus *Bursaphelencus* has been found in pines, which is called *B. mucronatus* and is thought to be nonpathogenic to pines (Mamiya and Enda, 1979). For these reasons, the identification and distinction of PWN have attracted a lot of researchers. For example, a direct PCR-based method for detecting *B. xylophilus* in wood tissue of *Pinus massoniana* has been reported (Hu et al., 2011). Direct molecular detection of the *B. xylophilus* from pine wood, bark and insect vector has also been developed (Cardoso et al., 2012). Moreover, one paper reported a simple, modified PCR-RFLP method using an egg, a second-stage juvenile, or an adult of living *B. xylophilus* to provide an unambiguous identification (Iwahori et al., 2000), and satellite DNA as a target for TaqMan real-time PCR detection of *B. xylophilus* has also been reported (François et al., 2007). A SCAR (Sequence Characterized Amplified Region) molecular marker to distinguish *B. mucronatus* from *B. xylophilus* has also been developed (Chen et al., 2011). A multiplex one-step PCR method for the simultaneous identification of three species, *B. xylophilus*, *B. mucronatus* and *B. doui*, within the *xylophilus* group was reported by the team of Zhuo et al. (2011). In China, PWN has mainly been found in the south, as the climatic conditions are suitable there for the survival of PWN, but Shi et al. (2013) first reported the isolation of PWN from *P. tabuliformis* forests, which are mainly located in the north of China, indicating that PWD has expanded across China. For quarantine purposes, because only live organisms are regulated, a molecular diagnostic method that can distinguish between living and non-living nematodes present in wood would be preferred. For this reason, two reverse

transcription-PCR methods have been develop that target Hsp70 nucleotide sequences in order to differentiate between living and non-living *B. xylophilus* in wood (Leal et al., 2013). As the basis for identification, the method used to obtain *B. xylophilus* specimens is also an important factor. A staining method was developed for estimating the nematode mass in a 5 mg cross-section from a wood sample of *P. massoniana* (Wang et al., 2010). Because field sampling and the traditional Baermann funnel method (Mamiya, 1975) take much time and labor to extract nematodes from dead pine tree chip samples, Yun et al. (2012) developed a rapid diagnostic for PWN infection by use of host-tree volatiles, which is thought attractive to PWN. Furthermore, in a newly infested tree or stand, the population of PWN is very low (Cheng et al., 2009), and so PWN is difficult to detect using traditional Baermann funnel extractions of wood discs cut from trees at breast-height (Yang et al., 2003). Thus, a method to investigate the within-tree distribution of PWN and a way to attract propagative PWN, which encompasses two thirds of the annual life history of the nematode (April through November), has been reported, and is thought to be an early diagnostic approach for PWD (Zhao et al., 2009).

Studies on pathogenicity of bacteria associated with B. xylophilus

The pathogenicity of PWN is also a frequently discussed problem. Han et al. (2003) reported that inoculating aseptic black pine seedlings with aseptic PWNs or bacteria alone did not lead to browning or wilting, but inoculation with aseptic PWNs combined with the bacteria isolated from *B. xylophilus* resulted in the onset of severe symptoms (Han et al., 2003). Zhao et al. (2003) discovered that inoculation with bacteria alone did not lead to the development of disease symptoms, but a combination of axenic PWNs and bacteria led to disease, while seedlings exhibited no or only weak symptoms when inoculated with axenic PWNs or axenic PWNs combined with the non-pathogenic bacterium (Zhao et al., 2003). To

clarify these problems, pathogenicity of aseptic *B. xylophilus* has been studied (Zhu et al., 2012). In addition, pathogenicity testing of four *Bursaphelenchus* species on conifer seedlings under greenhouse conditions has also been reported (Dayi and Akbulut, 2012).

Bacteria are also considered to be an important factor, for their association with pathogenicity. It is said that PWN has a specific bacterial symbiont that confers stronger virulence. To address this problem, the diversity of bacteria associated with *B. xylophilus* and other nematodes isolated from *P. pinaster* with PWD has been investigated (Proença et al., 2010). Specific and functional diversity of entophytic bacteria from PWN with different levels of virulence has also been investigated (Wu et al., 2013).

Studies on genetic variation

With more and more countries reporting PWN, it could be said that PWD has become a worldwide forest disease. As previous noted, PWN originated in North America, with the first report of PWD coming out of Japan in the 1970s (Mamiya, 1984), followed by China in 1982 (Cheng et al., 1983), soon after in Korea in 1988 (Yi et al., 1989), and more recently detected in Europe in 1999 (Mota et al., 1999). PWN has been regarded as an invasive species in Asia and Europe, but not North America (Kiyohara and Tokushige, 1971; Yang and Wang, 1988; Chio et al., 2006). The questions of how PWN spread to Asia and Europe and subsequently expanded all over the northern hemisphere, and whether the PWN within these countries have any connection, are thus attractive targets for research. Indeed, many studies have been carried out to address these concerns. Cheng et al. (2008) studied the genetic variation in the invasive process of PWN in China and successfully inferred the possible spread routes in China. Moreover, Zhang et al. (2008) used three molecular markers to analyze PWN samples from every country harboring infection to infer the origin and spread of PWN in China and worldwide. The results showed that isolates from China, Japan and the US formed a

monophyletic group which was the sister group to the clade composed of Canadian isolates and confirmed the hypothesis that *B. xylophilus* originated in North America (Knowles et al., 1983; Dwinell, 1993; Mota et al., 1999; Kanzaki and Futai, 2002). This agrees with the studies by Iwahori et al. (1998a), who found that the ITS-RFLP maps of Japanese, Chinese and US isolates were highly homologous, and Lu et al. (2001), who reported that the restriction enzyme map of ITS regions of *B. xylophilus* isolates from China and Japan were the same, but different from the Canadian. Even though a younger infection area, the population in Portugal has already been investigated by RAPD-PCR (Vieira et al., 2007); the same with the study in Madeira island (Fonseca et al., 2012) 1,000 km south west of continental Portugal, where PWN was detected ten years after first being reported in Portugal. These two studies both revealed that the genetic variation of PWN populations in Portugal were low, a typical characteristic of invasive species. In its native area, the genetic diversity of PWN has also been studied by the use of polymorphic microsatellite loci (Mallez et al., 2013). In addition to these large-scale comparisons, some smaller-scale population studies have also been carried out. Zhou et al. (2007) used SSR markers to study the genetic structure of populations of PWN between and within pine forests. Takemoto and Futai (2007) reported the polymorphism of Japanese isolates of PWN in 29 populations sub-cultured in the laboratory and others collected from natural pine stands at heat-shock protein 70 locus.

From these population genetic variation studies of PWN, it is obvious that the population of PWN which directly related to the dispersal of this species was highly complex. Study of the transmission routes and mechanisms of dispersal is regarded as important for planning effective control of the expansion of PWD and preventing its spread to uninfected regions (Jung et al., 2010b). The transmission of PWN in nature is mainly dependent on its vector beetles, *Monochamus* spp. (*Insecta*: *Coleoptera*: *Cerambycidae*). Depending on the geographic location, different species of *Monochamus* beetles act as the primary vectors for

PWN, *M. alternatus* is native to and the primary vector in Asia, whereas *M. carolinensis* is the primary vector in North America (Zhao et al., 2014). But for long distance transmission, because of the beetle's limited flying capability, the transmission of PWN is almost always associated with the logs, lumber, and wood packaging material that are used in the transportation of commercial goods (Hu et al., 2013). As too many unrecognized factors involved in the spread of PWN, it is difficult to obtain detailed population information (Jung et al., 2010b). Molecular markers showing sufficient genetic polymorphism have been used in population analysis and it is thought they could be beneficial for understanding the epidemiology of PWD (Jung et al., 2010b). Hu et al. (2013) did research to investigate the factors that influence the genetic structure of *M. alternatus* in mainland China and to elucidate the relationship between the genetic structure of *M. alternatus* and both human activity and PWN dispersal in China using *cox1* and *cox2* gene sequence variation. Some DNA analyses of *B. xylophilus* have been reported. Iwahori et al. (1998b) compared the PCR-RFLP patterns among various isolates of *B. xylophilus* and *B. mucronatus* originating from Japan and other countries using twelve restriction enzymes. Aikawa et al. (2003) performed single pair reciprocal interbreeding between a virulent and an avirulent isolate of *B. xylophilus* and individually analyzed the PCR-RFLP pattern of the rDNA region containing the 5.8S gene, ITS1 and ITS2, and partial regions of 18S and 28S genes of the offspring, attempting to demonstrate the occurrence of population interbreeding between the two isolates. SSRs were another effective class of molecular marker for population genetic study, as these are distributed extensively in both coding and non-coding sequences of eukaryotic genomes (Tautz and Renz, 1984). Because SSRs are highly polymorphic, easy to detect (by PCR), specific, and exhibit codominance (Powell et al., 1996), they are widely used in studies of population genetics (Perera et al., 2000; Viard et al., 2001), molecular environmental genetics (Schwartz et al., 2003; Williams et al., 2003), and molecular adaptation (Saint-Laurent et al.,

2003; Storz, 2002). For PWN, Zhou et al. (2007) reported that microsatellite markers may be useful for studying PWD. However, adequate loci are needed for the identification of individuals, and for multi-locus analysis, a method must be designed to ensure sufficient amounts of template, which is difficult to obtain from small organisms. To overcome this problem, Jung et al. (2010a) used the linker-attached fragment amplification method, which may be successful in genotyping multiple microsatellite loci (more than ten) from one individual and can be used to study microsatellite genetic variation of organisms otherwise too tiny to amplify multiple loci to study the variation of PWN in South Korea, and in the same year they also did the amplified fragment length polymorphism (AFLP) analysis of PWN in South Korea. Vieira et al. (2014) analyzed sequence variability of the *MspI* satellite DNA family of the PWN at different geographic scales. Assessment of the geographic origins of PWN via SNP in effector genes has also been investigated, as SNP markers, which come from sequence information and are highly reproducible, have become popular molecular markers. SNP genotyping is an accurate, scalable, cost effective process for the simultaneous detection of hundreds of polymorphisms (Figueiredo et al., 2013).

For non-model organisms, Expressed Sequence Tag (EST) analysis has proven a resourceful data set (Beldade and Vision, 2006; Bouck et al., 2007). EST analysis has been a powerful tool for identification of plant parasitic nematode genes which have roles in the host–parasite interaction. Moreover, SSRs mined from ESTs have been found to be significantly more transferable across taxonomic boundaries than traditional SSRs (Sim et al., 2009; Ellis and Burke, 2007) and the levels of polymorphism of transcribed EST-SSR regions appear to be similar to those of genomic-SSR regions (Tehrani et al., 2009). As EST-SSRs generally have fewer null alleles, greater cross-species amplification, and less allelic variability than genomic SSRs, EST-SSRs have been used widely in plant genetic studies (Aggarwal et al., 2007; Park et al., 2005; Wang et al., 2007) and population genetic studies

(Ellis and Burke, 2007; Slate et al., 2007; Wang et al., 2006; Kim et al., 2008). This showed that other molecular markers mining from EST should also be considered as powerful tools. More than 120,000 ESTs have been sequenced from a variety of plant parasitic nematodes, and an EST project for PWN was performed by Kikuchi et al. (2007). From their investigation, 13,327 ESTs of PWN were generated. Clusters of *B. xylophilus* varied in size from a single EST (4,377 cases) to 251 ESTs (1 case). These data provided us a very favorable source for marker development from ESTs for future studies.

DNA-based techniques developed during the past decades provide an attractive solution to nematode identification. In spite of that, determination of intraspecific variability is still difficult (Valadas et al., 2013). Those markers introduced previously were all mined from the nucleus. Besides the nucleus, other cell organelles of metazoan containing genetic material can also be used, such as mitochondria. Due to their high copy number in individual cells, lack of recombination and strict maternal inheritance, the mitochondrial genome (mtDNA) is regarded as an excellent marker for the study of phylogenetic relationships (Valadas et al., 2013). In recent years, mitochondrial genes have been used as markers to study intraspecific variation (Madani et al., 2010) and the genetic relationships among Peruvian and Canadian populations of *Globodera pallida* in order to identify the origin of new populations (Picard et al., 2007; Plantard et al., 2008; Madani et al., 2010). For example, mtDNA cytochrome *b* (*cytb*) gene sequences allowed the determination of the origin of new *G. pallida* populations (Plantard et al., 2008), which makes this a possible marker to be tested in the study of *B. xylophilus* intraspecific variability. Furthermore, Sultana et al. (2013) reported the whole mitochondrial genome sequence of PWN in 2013, providing us a good opportunity to develop molecular markers from mitochondria or study the population structure of PWN using mtDNA. With their report of the complete mtDNA sequences of PWN, Sultana et al. (2013) did a comparative analysis among the major groups of *chromadoreans* and inferred their

phylogenetic relationship. In the same year, they compared the complete mitochondrial genomes of *B. xylophilus* and *B. mucronatus* and developed a molecular tool to identify these two species (Sultana et al., 2013). Similarly, Pereira et al. (2013) did an intraspecific phylogeny of *B. xylophilus* and *B. mucronatus* using three mtDNA gene regions, cytochrome coxidase subunit *I* (*cox1*), NADH dehydrogenase subunit 5 (*nad5*) mitochondrial and small subunit ribosomal RNA (*rrnS*) from isolates of different geographic regions. Prior to these studies, Ye et al. (2007) reported the phylogenetic relationships among *Bursaphelenchus* species inferred from nuclear ribosomal and mitochondrial DNA sequence data. Following these studies, mtDNA could serve as another efficient tool for use in population genetic studies of PWN in the future.

1.3. Objective of the present study

Genetic variation information is the basic premise for discriminating between native populations (Humphreys et al., 2005). Information on genetic variation greatly facilitates germplasm classification and monitoring of genetic shifts in plant populations (Humphreys et al., 2005). Furthermore, estimates of genetic diversity are important because maintenance of genetic diversity is required for populations to evolve to adapt to environmental changes and because a decrease in genetic diversity is often related to inbreeding and reduction in reproductive fitness (Frankham et al., 2002). To prevent loss of genetic variability in natural populations through habitat destruction, areas with high genetic diversity are to be determined for both conservation and breeding programs (Tehrani et al., 2009). It has been reported that Kyushu is the first area of Japan infected with PWN, so it is reasonable to expect that the PWN in Kyushu area have relatively high genetic diversity. Therefore, the present study is aim to develop efficient molecular markers to investigate the genetic diversity of PWN in Kyushu area.

As noted in the introduction, there are various types of molecular markers that have been used to study PWN. Therefore, the main challenge for the present study lies in selecting one or more of these markers according to their specific purposes. A number of factors should be considered when choosing between the various molecular markers: (a) Marker system availability (b) Complexity of the technique and time investment (c) Estimated polymorphism levels within the study population (d) Quantity and quality of available DNA available (e) Transferability between laboratories, populations, pedigrees and species (f) The size and structure of the population to be studied (g) Availability of skilled workers and equipment (h) Cost per data-point and funding availability (i) Method of marker inheritance (e.g., dominant vs. codominant) and the type of genetic information needed in the population (Staub et al., 1996; Karp et al., 1997; Wolfe and Liston, 1998; Mackay, 2001; Rungis et al., 2005). In principle, the ideal type of genetic marker should be highly polymorphic, show codominant inheritance and be evenly distributed throughout the genome (Miah et al., 2013). In addition, particular marker sequences should be easy to access and analyses should be low cost, high-throughput, reproducible, and transferable between laboratories, populations and/or species (Miah et al., 2013). Unfortunately, no marker type currently exists that meets all of these requirements (Miah et al., 2013).

Of the molecular markers which have been introduced, RFLP analyses are not easily scalable to high-throughput methods, and RAPD assays are often not reproducible or transferable between laboratories. Although both microsatellites and AFLPs can be used to efficiently identify polymorphisms, microsatellite-based methods are more readily automated (Shariflou et al., 2001). In addition, AFLP analysis is not always straightforward, as apparently solitary bands may actually be composed of multiple fragments (Shan et al., 1999), particularly when using large genomic templates. All those DNA-based techniques provide an attractive solution for examining PWN populations, despite their limitations: the resolution of

RAPD and RFLP markers is low; few of these markers can be applied to a single PWN individual due to its extremely small body size; and determination of intraspecific variability is difficult because sample sizes are too small (Valadas et al., 2013).

ESTs obtained by partial random sequencing of cDNA library are 300-500 nucleotide long single read mRNA sequences from many of the genes expressed in a sample from an organism and they represent a snapshot of gene expression in a specific organ or tissue at a specific developmental stage (Kalia et al., 2011). Genomic markers are derived from all regions of DNA, including both transcribed and non-transcribed regions, while EST markers are expected to be relatively more conserved as they are derived from transcribed regions of DNA (Scott et al., 2000). Saha et al. (2004) developed a set of EST-SSR markers and tested their transferability across seven grass species of four genera (Festuca, Lolium, Oryza L., Triticum L.) differing in mating system and ploidy level. This work revealed greater than 90% transferability to one or more of the target species. The surveyed loci revealed large levels of polymorphism for elucidating relationships amongst these species. EST–SSRs offer advantages over genomic SSRs because they detect variation in the expressed portion of the genome, so that gene tagging should give "perfect" marker-trait associations and once developed, these markers, unlike genomic SSRs, may be used across a number of related species (Gupta et al., 1999). However, microsatellite markers are not perfect either; frequently a small number of potential microsatellite loci are identified, polymerase slippage occurs when analyzing mono- and di-nucleotide repeats, and co-migrating fragments are not always homologous (Rakoczy-Trojanowska and Bolibok, 2004).

SNP refers to a single base change in a DNA sequence, typically involving two possible alternative nucleotides at a given position (Vignal et al. 2002). A SNP occurs when a single nucleotide (A, T, C, or G) in the genome or other shared sequence differs between members of a species or between paired chromosomes in an individual (Semagn et al., 2014). SNPs

have largely replaced SSRs in crop species that have been extensively sequenced, such as maize, and they are expected to replace other types of molecular markers in most species in the near future given the increased use of next-generation sequencing technologies for genotyping (Semagn et al., 2014). Because of their low assay cost, high genomic abundance, locus specificity, co-dominant inheritance, simple documentation, potential for high-throughput analysis, and relatively low genotyping error rates (Rafalski, 2002; Schlötterer, 2004), SNPs have emerged as powerful tools for many genetic applications, including germplasm characterization (genetic diversity, genetic relationship, and population structure), quality control (QC) analysis (genetic identity, genetic purity, and parentage verification), linkage mapping, linkage-based and linkage disequilibrium-based quantitative trait loci (QTL) mapping, allele mining, marker-assisted backcrossing (MABC), marker-assisted recurrent selection (MARS), and genomic selection (GS) (Semagn et al., 2014). For these reasons, SNP markers have received the most attention as third-generation molecular markers in recent years. (Vignal et al., 2002). Genome sequencing information is indispensable for the development of SNP markers, and entire genome sequences have already been determined for *B. xylophilus* (Kikuchi et al., 2011). Additionally, numerous ESTs registered in the DNA database from gene expression analyses have made it possible to utilize SNP information in *B. xylophilus*.

Furthermore, as noted above, beside nuclear molecular markers, the mitochondrial genome is also a good marker development source because of its high copy number in individual cells, lack of recombination, and strict maternal inheritance (Valadas et al., 2013). In recent years, numerous studies have demonstrated that mitochondrial genes can make excellent markers for the study of phylogenetic relationships (Madani et al., 2010; Picard et al., 2007; Plantard et al., 2008; Sultana et al., 2013; Pereira et al., 2013; Ye et al., 2007). Hence, the present study also develops molecular markers from mitochondria to analyze the

genetic diversity of a population of PWN. In addition to reporting the complete mitochondrial genome of PWN, Sultana et al. (2013) also compared the entire mitochondrial genomes of *B. xylophilus* and *B. mucronatus* Mamiya and Enda (1979) and developed a molecular tool to identify these two species. Similarly, Pereira et al. (2013) clarified the intraspecific phylogeny of *B. xylophilus* isolates from different world regions and of *B. mucronatus* isolates using three mitochondrial genes: *cox1*, *nad5*, and *rrnS*. A previous study also reported the phylogenetic relationships among *Bursaphelenchus* species inferred from mtDNA and nuclear ribosomal sequence data (Ye et al., 2007). All the above-mentioned studies suggest that mitochondrial genome information might be a new and efficient tool for examining population genetic variability in PWN. However, PWN mitochondrial genome diversity is not well understood, and the use of this information to evaluate PWN population diversity has not advanced to date.

As previously stated, we sought to develop a method to utilize all mutations in a nucleotide sequence, and perform an exhaustive population analysis from both the nuclear and mitochondrial genomes of PWN. The development of next generation sequencing technology makes it possible to obtain a large amount of sequence information to capture sufficient variations. Moreover, the entire nuclear genome sequence was already been determined for PWN by Kikuchi et al. (2011). Additionally, numerous ESTs have been registered in the DNA database from gene sequence expression analysis, making it possible to utilize SNP information in PWN. Furthermore, the whole mitochondrial genome sequence of PWN was reported by Sultana et al. (2013), thereby providing valid information for studying PWN populations' variability based on mitochondrial genome sequencing and analysis of its polymorphisms. Our work builds on the tools provided by those earlier studies. Therefore, the object of the present study is to use SNP information derived from the nuclear and mitochondrial genomes of PWN to elucidate the genetic structure of the nematode population

and the phylogenetic relationships between the regional populations in Kyushu, the first invasion site in Asia.

Chapter 2. Genetic diversity and genetic structure of pinewood nematode, *Bursaphelenchus xylophilus* populations in Kyushu

2.1. Introduction

Bursaphelenchus xylophilus ((Steiner et Buhrer) Nickle) is a plant parasitic nematode which is highly pathogenic and has afflicted a large number of trees with Pine wood disease (hereinafter referred to as "PWD") since 1971 (Kiyohara and Tokushige, 1971). Because of its rapid spread, it has caused tremendous damage to pine forests around the world (Suzuki 2002). In these affected countries, various control and quarantine systems have been implemented, but PWD has not stopped spreading (Hu et al., 2011).

Now, thanks to active international trade and economic exchanges, the risk of spreading PWD is increasing every day. Understanding the pathways and mechanisms of transmission is considered essential to control PWD expansion and prevent it reaching uninfected areas (Jung et al., 2010b). Various factors have been involved in the transmission of *B. xylophilus,* including its vector beetles, *Monochamus* spp. (*Insecta: Coleoptera: Cerambycidae*), at close range, while human activity accounts for most long-range dispersal (Jones et al. 2008). It seems that the nematode worm population is forming a complicated genetic structure. In order to explain this conundrum, it is necessary to conduct population genetic study of *B. xylophilus*.

DNA molecular markers are important tools for population genetic analysis (Sunnucks, 2000). In recent years, several molecular markers for genetic analysis of *B. xylophilus* have been developed. PCR-RFLP markers (Iwahori et al. 1998b; Takemoto and Futai 2007), RAPDs marker (Zhang et al. 2008) and SSR markers (Mallez et al. 2013; Jung et al. 2010a) have been developed and used for genetic studies of *B. xylophilus*. In addition, haplotype analysis using SNPs in the ITS region of rDNA has also been performed (Nose et al., 2009).

These DNA molecular markers are also used for population genetic studies of *B. xylophilus* in Japan. Zhou et al. (2007) used the SSR markers to elucidate genetic diversity in the Kanto region. In addition, Nose et al. (2009) conducted phylogenetic analysis of regional populations in the Kyushu region.

SNPs are widely distributed in the genome of the organism, which is rich in mutation sites. Markers that can take advantage of this function facilitate simple and accurate genotyping. For these reasons, SNP markers have received the most attention as third-generation molecular markers in recent years (Vignal et al. 2002). Genome sequencing information is indispensable for the development of SNP markers, and the entire genome sequences have already been determined for *B. xylophilus* (Kikuchi et al., 2011). Additionally, numerous Expressed Sequence Tags (ESTs) have been registered in the DNA database from gene sequence expression analyses, making it possible to utilize SNP information in *B. xylophilus*.

Due to the distribution of invasive organisms in Asia and Europe, it was predicted that genetic diversity of the *B. xylophilus* would gradually decrease over the invasion process, as is characteristic of invading organisms (Tsutsui et al., 2000; Sakai et al., 2001). In addition, as economic activities are also frequently involved in the expansion and distribution of *B. xylophilus*, it is predicted to form a different genetic structure than would be formed under a more natural expansion scenario. Therefore, this research attempted to take advantage of the nucleotide sequence polymorphism of EST loci reflecting a plurality of SNP information to evaluate the genetic diversity and to elucidate the genetic structure of *B. xylophilus* in the Kyushu area, which is the first invasion site in Asia.

2.2. Materials and Methods

B. xylophilus from eight damage regions in Kyushu area of Karatsu (Saga prefecture),

Matsuura (Nagasaki prefecture), Amakusa (Kumamoto prefecture), Shintomi, Miyazaki and Nichinan (Miyazaki prefecture), Ibusuki and Sendai (Kagoshima prefecture) have been collected in 2012-2014. The *B. xylophilus* population isolated from one damaged tree was regarded as one isolate (hereinafter referred to as a damaged tree group). The number of damaged tree populations in each regional population is shown in Table 2-1. In total, 107 damage tree populations of 8 regional populations have been collected. The *B. xylophilus* collection was performed from one place (about 20 cm × 20 cm) in the breast height part of the dead tree. After the bark was taken out from the height of the chest (about 20 cm × 20 cm), the sawdust was collected from 5 points of the damaged tree with a 16 mm diameter electric drill, and then thoroughly mixed and extracted. *B. xylophilus* was extracted by Baermann funnel technique (Iwahori and Futai, 1993), and then cultured on *Botrytis cinerea* Pars., grown on barley culture medium and incubated at 25 °C for 10~20 days. Regarding the preparation of the barley medium, distilled water of the same amount (W / W) as that of barley was placed in a tube and sterilized in an autoclave (121 °C for 20 min). The successfully cultured *B. xylophilus* were separated from the culture medium, then suspended in distilled water and stored at 4 °C until use. DNA isolation and purification were carried out according to Nose et al. (2009). A PCR template DNA was obtained from a population isolated from a damaged tree.

Nucleotide sequences information of the *B. xylophilus* EST（Kikuchi et al., 2007）obtained from DDBJ database（https://www.ddbj.nig.ac.jp/index.html）was used to design primer pairs for PCR. Multiplex PCR was performed using primer pairs of the EST loci shown in Table 2-2. The multiplex PCR were performed with Takara Multiplex Ver 2.0 in a final volume of 25 μl containing 50 μl of genomic DNA, 5 μl of primer mix (1 μM), 12.5 μl of 2× Multiplex PCR Buffer (Mg^{2+}, dNTP plus), and Multiplex PCR Enzyme Mix 0.125 μl. PCR was performed using Agilent SureCycler 8800 (Agilent Technologies) at 94 °C for 1

min, followed by 30 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s, finally extended at 72 °C for 10 min. The PCR product was electrophoresed on 1.2% agarose gel, a DNA fragment of 100 to 300 bp was excised and purified using MagExtractor-PCR & Gel clean up (TOYOBO) to prepare a library. All libraries were mixed in equal amounts, and used for analysis in the next generation sequencer (Illumina MiSeq System (Illumina)). The analytical sample used for the next-generation sequencer was prepared using TruSeq Nano DNA Library Prep Kit (Illumina). Sequencing was performed according to the operation manual of Illumina.

The nucleotide sequence data (only Q20 or more) obtained from the sequencing was first classified for each library and further classified according to the EST loci. Allele frequencies for each library (isolate) and EST locus were determined, and the genetic diversity (H_T, H_S), gene differentiation coefficient (G_{ST}) (Nei, 1987) and genetic distance (Nei, 1972) were calculated from the allele frequencies.

2.3. Results

With primer pairs designed using the nucleotide sequences information of reported *B. xylophilus*（Kikuchi et al., 2007）and sequences obtained from a next-generation sequencer, ten EST loci acquired sufficient information (more than 50 reads per locus· population) and are shown in Table 2-2. Table 2-2 also shows the number of sequences and the number of alleles per locus of the damaged tree population (the number of alleles with a frequency of 1% or higher and the total number of alleles). The heterozygosity of each subpopulation in ten polymorphic loci is shown in Appendix (Table S2-1), and detailed information of the allele frequencies in ten polymorphic loci is shown in Appendix (Table S2-2). Moreover, nucleotide sequences of alleles in ten polymorphic loci are shown in Appendix (Table S2-3). The average value and the standard deviation of the number of sequences of each locus obtained per

population were 116.0 ± 33.8 (pX070) ~ 201.2 ± 63.9 (pX144). In addition, as shown in Table 2-3 (H_T), the genetic diversity of each locus in the entire Kyushu region ranged from 0.487 (pX021) to 0.790 (pX155) with an average of 0.628.

Genetic diversity in eight regional populations of Kyushu (Table 2-1) was evaluated by Nei (1987) using the data set of ten polymorphic loci (Fig. 2-1 (A)). The expected heterozygosity value (H_S in Table 2-3) of each tested damage tree within each regional population of *B. xylophilus* was taken as representative of the genetic diversity of the population. The genetic diversity of the regional population ranged from 0.123 (Amakusa) to 0.421 (Matsuura), and the average value was 0.296. The highest genetic diversity was 0.421 ± 0.196 in Matsuura, then 0.414 ± 0.092 in Karatsu. Shintomi (0.371 ± 0.136) and Sendai (0.338 ± 0.135) also showed relatively high diversity. On the other hand, the lowest genetic diversity values were obtained for Amakusa (0.123 ± 0.018) and Miyazaki (0.173 ± 0.022). A bottleneck test was performed for these two populations using Tajima's D. The results showed that seven out of the eight effective loci in Amakusa had a significant negative bias (six loci were at the 0.1% level and one locus was at the 5% level), while eight out of the ten effective loci in Miyazaki showed significant negative bias (all of the loci were at the 0.1% level). These negative biases were observed in the population-processed bottleneck. The total heterozygosity (H_T in Table 2-3) of the regional population, which was calculated from the average allele frequency of the damaged tree population, ranged from 0.123 (Amakusa) to 0.590 (Sendai).

The highest polymorphism in the population of damaged trees was observed in Matsuura, with a frequency of 0.660. In this regional population, six out of 16 trees exceeded 0.6, which was confirmed to contain a large number of *B. xylophilus* population that has a high diversity. In addition, one damaged tree population exceeding 0.6 was also identified in Shintomi. On the other hand, the diversity of each damaged tree population of Amakusa was

0.091 to 0.142, and all six of the tested tree populations showed extremely low values. The diversity of Miyazaki ranged from 0.134 to 0.212, and was low in all 18 tree populations. Values of six tree populations in Karatsu were 0.330 to 0.541, and all of the populations showed similar diversity. In order to evaluate the difference in genetic composition between damaged tree populations in each regional population, the genetic distances (Nei, 1972) of all pairs of the damaged tree populations were calculated and the average value was used as a measure of the genetic difference of the population (Fig. 2-1 (B)). The highest was confirmed in Sendai (0.647 ± 0.660), followed by Shintomi (0.0.488 ± 0.458). On the other hand, the average genetic distance between Amakusa (0.001 ± 0.000) and Miyazaki (0.003 ± 0.002) was extremely small, indicating that there was no difference in the genetic composition of the damaged tree populations. Fig. 2-2 shows the relationship between average gene diversity and mean genetic distance. Naturally, in regional populations with extremely low genetic diversity (Amakusa, Miyazaki, Ibusuki), the average genetic distance was also extremely small (a significant positive correlation among eight populations (r^2= 0.781*)). However, in regional populations with relatively high genetic diversity (Matsuura, Karatsu, Shintomi, Sendai), a tendency for genetic differences between damaged tree populations to increase as genetic diversity of regional populations increases was not confirmed (a significant negative correlation among 4 populations (r^2= -0.991*)).

In order to evaluate the genetic structure in individual regional populations, the gene differentiation coefficient (G_{ST}: Nei, 1987) was calculated from the total gene diversity (H_T) and the average value (H_S) of genetic diversity of the damaged tree population (Table 2-3). The G_{ST} for Sendai was the largest at 0.428, followed by 0.346 for Shintomi. On the other hand, extremely low G_{ST} values of 0.006 and 0.015 were observed in Amakusa and Miyazaki, respectively.

2.4. Discussion

In this study, all sequence variations in EST loci obtained by the next generation sequencer were captured and alleles were detected from the polymorphism. As a result, ten loci were developed for *B. xylophilus* with high variability (average 0.628 (0.487 - 0.790), Table 2-3). Compared to the DNA markers, such as SNPs and SSRs, that have been frequently used in many studies which only captured mutations at specific sites on the nucleotide sequences (Tautz and Renz 1984; Brookes, 1999), the method in the present study made it possible to utilize all of the mutations on the nucleotide sequences. It is considered to be an effective method when only the allele frequency of the investigated individual is obtained in this study. Since abundant EST nucleotide sequence information has already been reported (Kikuchi et al., 2007), it is not difficult to increase the number of loci. It is hoped that this innovation will lead to more detailed findings in future population genetic study of *B. xylophilus*.

Genetic diversity was used to assess the genetic diversity of regional populations (Nei, 1987). As a result, it was observed that the regional population differed greatly (0.123-0.590) (H_T in Table 2-3). Four populations were divided into three groups according to genetic diversity (Karatsu (0.554), Matsuura (0.564), Shintomi (0.567), Sendai (0.590)) with high diversity. Nichinan (0.413) with moderate diversity, and three populations ((Amakusa: 0.123, Miyazaki: 0.176, Ibusuki: 0.267) with low diversity. A significant difference at the 1% level among these three groups was revealed through the Kruskal-Wallis test (Kruskal and Wallis, 1952). Zhou et al. (2007) evaluated the genetic diversity of three populations of Kanto (Tanashi, Tsukuba, Chiba) using four SSR markers. They reported two populations (Tsukuba, Chiba) rich in diversity and one population of Tanashi with extremely poor diversity in the results. It can be seen from these facts that populations of high diversity and populations of extremely low diversity exist quite normally in the regional population of *B. xylophilus*.

Based on the genetic differentiation coefficient (G_{ST}), the genetic structure of the entire Kyushu population was analyzed. It was found that more than half (G_{ST} = 0.529) of the total genetic diversity (H_T = 0.628) was present in the regional population, and there were mutations between the regional populations. This confirmed that there is a large genetic difference between the regional populations. Kishi (1988) reported the propagation path of *B. xylophilus* of Kyushu based on past damage records. According to the report, there are four routes invaded from four places as Nagasaki, Sasebo, Nichinan and Aira, which is supported by polymorphism analysis of rDNA by Nose et al. (2009). The high G_{ST} value revealed that recent studies have shown that systematically different *B. xylophilus* populations from North America have invaded Kyushu and their influence is still strong.

In the highly diverse regional populations (Karatsu, Matsuura, Shintomi, Sendai), there is also a large difference between the damaged tree populations (Fig. 2-1 (A)). Furthermore, the genetic distances among the damaged tree populations (Karatsu: 0.370, Matsuura: 0.340, Shinnito: 0.488, Sendai: 0.647) are large (Table 2-4, Fig. 2-1 (B)). Obviously, the genetic composition of damaged tree species is very different. Among the four regional populations that maintain high diversity, 25 to 43% of all genetic mutations were present among damaged tree populations (Table 2-3), and the genetic distances among damaged tree populations were also large (Table 2-4). The investigation in Kanto by Zhou et al. (2007) also revealed that within the population of Tsukuba and Chiba with high diversity, the genetic distance was also large. Judging by the beetles which are vectors for *B. xylophilus*, the genetic composition of *B. xylophilus* in the tree that holding the beetle before emergence was strongly reflected by the genetic composition of the *B. xylophilus* in the next injured tree. Therefore, it is hypothesized that *B. xylophilus* populations adhering to each beetle also have various levels of genetic diversity. This is evident from the study of the *B. xylophilus* population carried by each beetle. In addition, the number of carriers that multiply on a single pine tree and the

difference in postprandial behavior (Zhou et al., 2007) also have a complex but important relationship; thus it is understandable that large genetic composition differences have developed among the damaged tree populations.

On the other hand, in Amakusa, Miyazaki and Ibusuki with low diversity, G_{ST} was extremely small (0.006, 0.015, 0.092, respectively) and genetic variation in almost all regional populations exists in individual damaged tree populations. In addition, the genetic distance between damaged tree populations was also small (Amakusa: 0.001, Miyazaki: 0.003, Ibusuki: 0.034) (Table 2-4, Fig. 2-1 (B)), and the genetic composition of the damaged tree populations was also similar in these three populations. In the area destroyed by PWD in the past, the number of *B. xylophilus* drastically decreased due to the decline of the pine forest. It seems that the small surviving population (Bottleneck effect; Bonnell and Selander, 1974) would proliferate again with the regeneration of the pine forest. On the other hand, we believe that in the artificial expansion, the genetic diversity of newly introduced *B. xylophilus* populations in pine forest was originally low (Founder effect; Mayr, 1954). Even in the areas damaged by PWD, it is believed that the founder effect would occur when *B. xylophilus* was newly transmitted naturally to nearby pine forests by a very small number of vectors. Zhou et al. (2007) reported the existence of a population with extremely low diversity (Tanashi) among the three regional populations of the investigated Kanto, and they inferred that it was influenced by the founder effect and genetic drift according to the progress of fragmentation of the pine forest by urbanization. As a result of examining the presence or absence of a significant bottleneck has been observed using Tajima 'D (Tajima, 1989) for two populations (Amakusa, Miyazaki) in which the diversity was low in this study. Since both the Kanto and Kyushu revealed a very low-diversity population, the impact of the bottleneck/founder effect seems to have played a significant role in the formation of regional populations in Japan. The Amakusa sampling site in Amakusa, which showed the least diversity this time is the same

area as Amakusa Kamishima (Amakusa) in the study of Nose et al. (2009) where sampling was done from 2006 to 2007. At that time, four of the seven rDNA haplotypes mainly present in Kyushu were detected and were diverse populations of Kyushu. In this regard, it has suffered serious damage from PWD since 2009. It is presumed that the population size of *B. xylophilus* drastically decreased due to the large-scale annihilation of pine forest, but the bottlenecked population has nonetheless survived.

It is clear that in Kyushu, *B. xylophilus* populations with rich genetic diversity and *B. xylophilus* populations lacking genetic diversity were both exciting. As with the Amakusa phenomenon mentioned above, even in those who have maintained diversity in the past, there may be a decline in the diversity of some populations. In particular, Kyushu was the first region in which *B. xylophilus* was found to be invasive. It has already been a century since 1905, when then species invaded Nagasaki (Yano, 1913). The period over which the pine forests have been exposed to the depredations of PWD is longer in Kyushu than in other areas. It can be argued that the history of damage caused by repeated PWD over the years has revealed the polarization of diversity possessed by a regional population of *B. xylophilus*.

To date, little quantitative evaluation of the genetic diversity and genetic structure of *B. xylophilus* populations has been performed. In this study, the genetic population analysis method was used to quantify the genetic diversity harbored at each level: the entirety of Kyushu, regional populations, and damaged tree populations. Large differences in genetic characteristics among regional populations of *B. xylophilus* were also revealed by the population structure analysis. Nowadays, PWD has spread throughout Japan, excepting Hokkaido. In order to elucidate the genetic diversity and structure of the population in the process of *B. xylophilus* expansion, it is necessary to conduct quantitative surveys nationwide.

Tables

Table 2-1. The sampling sites of pinewood nematode and the number of damaged pine trees used in this study

Regional population (Code)	Locality	No. of pine trees
Karatsu (Ka)	Karatsu, Saga Pref.	6
Matsuura (Ma)	Matsuura, Nagasaki Pref.	15
Amakusa (Am)	Kamiamakusa, Kumamoto Pref.	6
Shintomi (Sh)	Shintomi, Miyazaki Pref.	14
Miyazaki (Mi)	Miyazaki, Miyazaki Pref.	18
Nichinan (Ni)	Nichinan, Miyazaki Pref.	17
Sendai (Se)	Satsumasendai, Kagoshima Pref.	15
Ibusuki (Ib)	Ibusuki, Kagoshima Pref.	16
Total		107

Table 2-2. The PCR primer sequences of the ten putative loci, the numbers of the effective sequences, and the detected alleles in pinewood nematode

Locus	Sequences of primer pair	No. of sequences*	No. of alleles**	Accession No. of DDBJ
pX021	Fw: TTGTAACCTCCTGACAACTCGCTCAC Rv: TTATGGTCGGATCTTATCTTGTCCTC	146.3±48.3	4 (64)	CJ982155
pX043	Fw: TTATGCCAACATTCTCATAAACAGACTTCG Rv: TTCCCAACCAACGTAGAGGAGACTC	137.4±40.4	6 (58)	CJ984741
pX047	Fw: TTGAGTCATTTGGCAAGAGGACTTCA Rv: TTGTACTTGAAGTCAAACTTTCCACGAG	140.3±38.7	5 (32)	CJ983640
pX070	Fw: TTGTCCAAATCGGAAACGCCTG Rv: TTCCTTTCCGGTGATCAGTTGCTC	116.0±33.8	6 (42)	CJ984816
pX079	Fw: TTAAAGGACGGCACCAACCCACA Rv: TTGCTCGTCTCGGTACTGACCTG	134.9±39.6	9 (16)	CJ984251
pX130	Fw: TTGCGGTTCAACTGATCCTACATCGAC Rv: TTATCTCCAGATCCCGAGCCAC	157.5±47.3	4 (51)	CJ984817
pX143	Fw: TTAGTTGGTCAAAGCAAACGCTCA Rv: TTAATGACGCCACGACAGCAC	143.5±51.3	3 (41)	CJ983461
pX144	Fw: TTATCGTTCATTATGAGATTGCCGTTG Rv: TTACAATCTGATTCGCCAAGTCCA	201.2±63.9	7 (41)	CJ983946
pX155	Fw: TTGACAAGCCCGACCAACAGC Rv: TTACCTGAATAGTTCCAGCCGAGA	181.2±53.6	7 (36)	CJ982187
pX194	Fw: TTGCTAGTTTGTTCCTCCAATCCGA Rv: TTGTCCCTTCGAGGTCTTCCGA	132.7±41.7	6 (43)	CJ979751

* Values indicate the average number and the standard deviation of effective sequences per subpopulation.
** Value indicates the number of alleles with a frequency of 1% or more, while the value in the parentheses indicates the number of all alleles.

Table 2-3. Genetic diversity (H_T, H_S) and genetic population structure (G_{ST}) of pinewood nematode estimated from ten putative loci

Locus		Regional population								Whole Kyushu
		Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX021	H_T	0.377	0.345	0.245	0.408	0.237	0.302	0.400	0.664	0.487
	H_S	0.342	0.325	0.242	0.295	0.234	0.272	0.209	0.521	0.305
pX043	H_T	0.558	0.628	0.257	0.572	0.227	0.547	0.763	0.138	0.623
	H_S	0.457	0.439	0.256	0.380	0.224	0.293	0.421	0.130	0.325
pX047	H_T	0.540	0.477	0.005	0.496	0.129	0.346	0.615	0.126	0.668
	H_S	0.248	0.289	0.005	0.289	0.128	0.231	0.270	0.118	0.197
pX070	H_T	0.493	0.507	0.000	0.495	0.152	0.324	0.757	0.167	0.686
	H_S	0.232	0.400	0.000	0.399	0.150	0.277	0.409	0.157	0.253
pX079	H_T	0.780	0.514	0.023	0.707	0.014	0.451	0.723	0.131	0.744
	H_S	0.571	0.354	0.023	0.437	0.014	0.234	0.484	0.126	0.280
pX130	H_T	0.364	0.541	0.180	0.616	0.204	0.433	0.482	0.199	0.584
	H_S	0.277	0.458	0.180	0.447	0.202	0.360	0.280	0.190	0.299
pX143	H_T	0.521	0.539	0.140	0.517	0.169	0.344	0.366	0.319	0.524
	H_S	0.420	0.442	0.139	0.282	0.168	0.226	0.200	0.300	0.272
pX144	H_T	0.596	0.685	0.164	0.529	0.143	0.356	0.675	0.374	0.614
	H_S	0.563	0.516	0.163	0.314	0.141	0.257	0.400	0.350	0.338
pX155	H_T	0.752	0.724	0.064	0.798	0.304	0.670	0.526	0.194	0.790
	H_S	0.573	0.480	0.063	0.542	0.293	0.471	0.260	0.190	0.359
pX194	H_T	0.555	0.680	0.156	0.537	0.179	0.359	0.597	0.357	0.561
	H_S	0.461	0.505	0.156	0.322	0.176	0.247	0.446	0.342	0.332
Average	H_T	0.554a*	0.564a*	0.123c*	0.567a*	0.176c*	0.413b*	0.590a*	0.267c*	0.628
	H_S	0.414	0.421	0.123	0.371	0.173	0.287	0.338	0.243	0.296
Coefficient of gene differentiation (G_{ST})		0.252	0.254	0.006	0.346	0.015	0.306	0.428	0.092	0.529

H_T: total gene diversity of population; H_S: average gene diversity wthin population

* a, b, and c indicate significant differences of three groups at p <0.01 by Kruskal-Wallis test.

Table 2-4. Average pairwise genetic distances (Ds) between and within eight regional populations of pinewood nematode

Population	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib
Ka	0.370 ± 0.321							
Ma	0.386 ± 0.277	0.340 ± 0.299						
Am	0.664 ± 0.242	0.608 ± 0.254	0.001 ± 0.000					
Sh	0.418 ± 0.222	0.478 ± 0.297	0.689 ± 0.261	0.488 ± 0.456				
Mi	0.375 ± 0.221	0.528 ± 0.342	0.563 ± 0.014	0.413 ± 0.457	0.003 ± 0.002			
Ni	0.362 ± 0.218	0.485 ± 0.319	0.613 ± 0.186	0.404 ± 0.442	0.227 ± 0.285	0.262 ± 0.334		
Se	1.024 ± 0.511	1.153 ± 0.657	1.083 ± 0.373	1.058 ± 0.610	1.540 ± 0.821	1.242 ± 0.730	0.647 ± 0.661	
Ib	1.214 ± 0.371	1.269 ± 0.920	1.686 ± 0.249	1.380 ± 0.868	2.581 ± 0.534	1.913 ± 0.777	0.947 ± 0.588	0.034 ± 0.029

Values indicate the average and the standard deviation of the pairwise genetic distances. Diagonal element indicates the distance within a population, and element below the diagonal indicates the distance between populations.

Figures

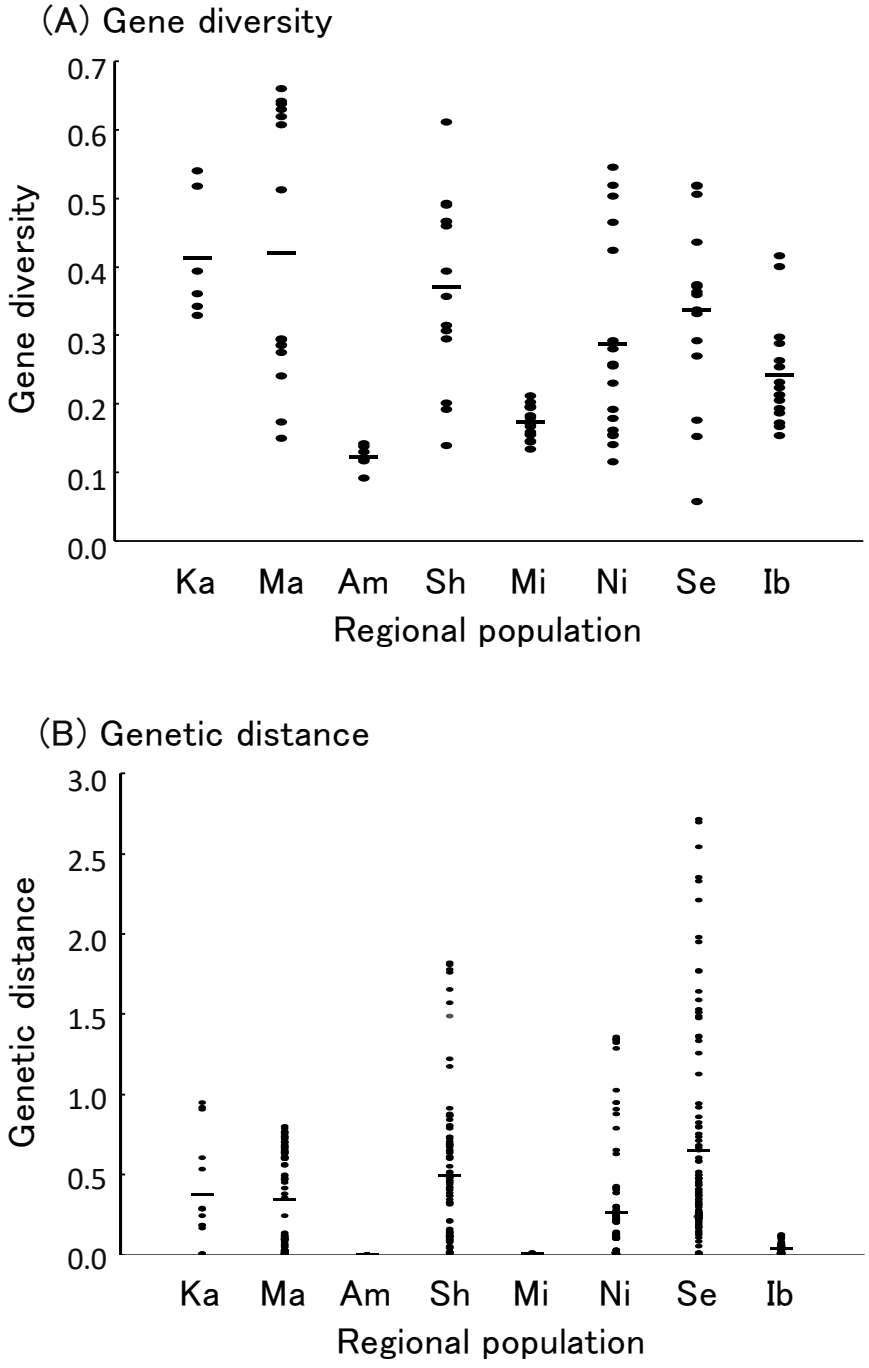

Fig. 2-1. The gene diversity of the populations within damaged trees (subpopulation) ((A) Gene diversity) and the pairwise genetic distances among the subpopulations ((B) Genetic distance) in eight regional populations of pinewood nematode

"-" indicates the average gene diversity (A) or the average pairwise genetic distance (B) in each regional population.

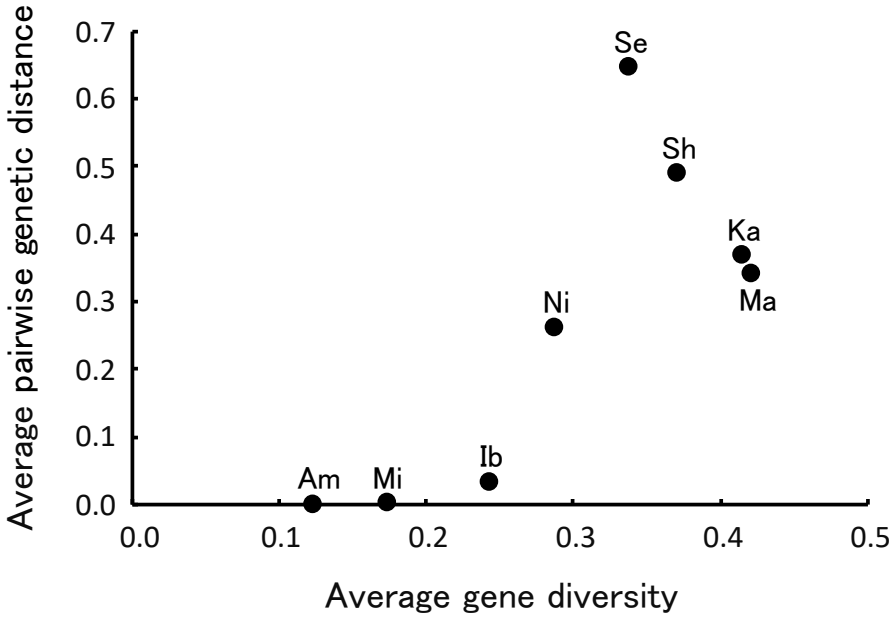

Fig. 2-2. Relationship between the average gene diversity and the average pairwise genetic distance among subpopulations in eight regional populations of pinewood nematode

Chapter 3. High mitochondrial genome diversity and intricate population structure of pinewood nematode, *Bursaphelenchus xylophilus* in Kyushu

3.1. Introduction

Pine wilt disease (PWD), one of the most serious forest problems worldwide, originated in North American countries (Dwinell, 1997) spreading to many others during the 20th century. In USA and Canada, PWD is not considered a primary pathogen of native pines (Leal et al., 2013), whereas in invasion areas, such as Japan, Korea, China, and Portugal, it exerts serious damage to forest ecosystems (Yun et al., 2012).

In Japan, PWD was first recorded at Nagasaki prefecture, northwestern part of Kyushu, in 1905. During the last century, this disease has been responsible for the yearly loss of 700,000 m^3 of pinewood (Mamiya and Shoji, 2009). In China, PWD was first observed in Nanjing in 1982, but it rapidly expanded (Yang, 1995; Zhang and Luo, 2003; Wu, 2004) and over one million hectares of pine forests have died due to PWD (Zhao, 2008). In South Korea, PWD was first identified in the Gumjung Mountain Region of Busan in 1988 (Yi et al., 1989), but it widely spread to 57 cities and prefectures until 2010 (Jung et al., 2010a). In Europe, PWD was first reported in Portugal in 1999 (Mota et al., 1999). During the next decade, nearly 510,000 ha of Portuguese pine forests were destroyed due to PWD (Valadas et al., 2012). By 2008, the disease spread across the country (Rodrigues, 2008) and, more recently, the PWD was detected in Madeira Island, 1,000 km southwest of mainland Portugal (Fonseca et al., 2012). PWD has been detected also in Spain (Abelleira et al., 2011; Robertson et al., 2011), and, throughout the world, 44 *Pinus* species have been infected by PWD under natural conditions (Shi et al., 2013).

The pinewood nematode (PWN), *Bursaphelenchus xylophilus* (Steiner and Buhrer, 1934), Nickle, 1970 (Nematoda: Aphelenchoididae), which is a kind of plant parasite, is the

causal agent of PWD (Kiyohara and Tokushige, 1971). This nematode can survive in healthy and in dead trees. In healthy host trees, the PWN causes cell destruction leading to host death in a few months (Cardoso et al., 2012). Although it has been listed as a quarantine pest in more than 40 countries (Mota et al., 1999; Schrader and Unger, 2003), the lack of an effective method to inhibit its expansion is still a serious problem worldwide. Because the PWN is present in the logs, lumber, and wooden packaging material used in commercial transportation (Hu et al., 2013), it is difficult to inhibit PWN expansion, except for the transportation performed by its vectors, the beetles within the genus *Monochamus*. Understanding its transmission routes and dispersal mechanisms could be effective for controlling the propagation of the PWN, especially to uninfected regions (Jung et al., 2010b).

Molecular markers showing sufficient genetic polymorphism are thought to help understanding the epidemiology of PWD (Jung et al., 2010b), and many studies have focused on PWN population analysis. Cheng et al. (2008) studied the genetic variation during the invasive process of PWN in China and successfully inferred its possible spread routes using amplified fragment length polymorphism (AFLP). Iwahori et al. (1998a) used an internal transcribed spacer (ITS)-Restriction Fragment Length Polymorphism (RFLP) map to identify the relationship among Japanese, Chinese, Canadian, and US PWN isolates. Random amplified polymorphic DNA (RAPD)-PCR and ITS-RFLP have been used to examine PWN populations in mainland Portugal (Vieira et al., 2007) and in Madeira Island (Fonseca et al., 2012), respectively. Zhou et al. (2007) reported that microsatellite markers might be useful for studying PWD, and they used such markers to evaluate the genetic structure of PWN populations among and within pine forests. The polymorphism of Japanese PWN isolates from 29 populations sub-cultured in the laboratory and collected from natural pine stands was analyzed based on variations at the heat-shock protein 70 locus (Takemoto and Futai, 2007). All these DNA-based techniques provide an attractive solution for examining PWN

populations, despite their limitations: the resolution level of RAPD and RFLP markers is low; few of these markers can be applied to a single PWN individual due to its extremely small body size; and determination of intraspecific variability is difficult because sample sizes are too small (Valadas et al., 2013).

Mitochondrial DNA (mtDNA) is an excellent marker for the study of phylogenetic relationships due to its high copy number in a cell (Valadas et al., 2013). In recent years, mitochondrial genes have been used as markers for intraspecific variation studies in nematodes (Madani et al., 2010). They were employed to investigate genetic relationships among Peruvian and Canadian populations of *Globodera pallida* (Picard et al., 2007; Plantard et al., 2008; Madani et al., 2010), and the mitochondrial gene cytochrome *b* (*cyt*b) successfully confirmed the origin of new populations of this species (Plantard et al., 2008). These results suggested that mitochondrial genes might be useful for studying intraspecific variability in the genus *Bursaphelenchus*. Furthermore, Sultana et al. (2013) reported the whole genome sequence of PWN, thereby providing valid information for studying PWN populations' variability based on mitochondrial genome sequencing and analysis of its polymorphisms. In addition to reporting the complete mitochondrial genome of PWN, Sultana et al. (2013) also compared the entire mitochondrial genomes of *B. xylophilus* and *B. mucronatus* Mamiya and Enda, 1979, and developed a molecular tool to identify these two species. Similarly, Pereira et al. (2013) clarified the intraspecific phylogeny of *B. xylophilus* isolates from different world regions and of *B. mucronatus* isolates using three mitochondrial genes: cytochrome *c* oxidase subunit *I* (*cox1*), NADH dehydrogenase subunit 5 (*nad5*), and small subunit ribosomal RNA (*rrnS*). A previous study also reported the phylogenetic relationships among *Bursaphelenchus* species inferred from mtDNA and nuclear ribosomal sequence data (Ye et al., 2007).

All the above-mentioned studies suggest that mitochondrial genome information might

be a new and efficient tool for examining population genetic variability in PWN. However, PWN mitochondrial genome diversity is not well understood, and using this information to evaluate PWN population diversity has not advanced to date. Thus, in the present study, the partial mitochondrial genome was sequenced for individual PWNs, and the sequence polymorphism and genetic population structure in PWN were investigated in detail.

3.2. Materials and methods

Nematode collection and preparation

The nematode samples used in this study were collected from 12 different forests in Kyushu, Japan where PWD was reported from 2012 to 2014 (Fig. 3-1). Pine chips were obtained from five points in the trunk, using a drill with a diameter of 16 mm. Nematodes were extracted from the chips using the Baermann funnel method (Iwahori and Futai, 1993), cultured on *Botrytis cinerea* Pers. (1794) grown on barley culture medium, and incubated at 25 °C for 10–20 days. After successful rearing, nematodes were removed from the culture medium, and stored in 5-ml tubes containing distilled water, at 4 °C until use. Mitochondrial DNA (mtDNA) sequencing was performed for each individual nematode. The number of nematode individuals in each population for which mtDNA sequences were successfully obtained and used in further analysis are listed in Table 3-1. The number of trees in each population that became the source of the sequenced nematodes is also listed.

PCR amplification and DNA sequencing

A primer pair for long PCR was designed in Oligo7 (Molecular Biology Insights), based on the complete mitochondrial genome sequence of *B. xylophilus* (14,778 bp; NCBI Accession No GQ332424). The forward (5'-TCCTCCATTAAGAACTTTAGGGCAT-3') and reverse (5'-TACAGTCAAAGCAATAGGACGAGA-3') primers designed produced an

amplicon about 8 kb-long that covered more than 50% of the mitochondrial genome.

Each nematode was used for direct amplification of mtDNA. Under a binocular stereomicroscope, each individual was separated from the nematode suspension using a needle and divided into two halves using a scalpel. Both halves were then transferred to a 25-μl PCR mixture [1× Gflex PCR Buffer (Takara), 0.43 μM each primer, 0.625 units Tks Gflex DNA Polymerase (Takara)]. The PCR reaction was carried out in the SureCycler 8800 (Agilent Technologies) device and comprised an initial denaturation at 94 °C for 1 min, followed by 30 cycles at 98 °C for 10 s and 68 °C for 8 min. The resulting amplicon was electrophoresed on 1% agarose gel with ethidium bromide, and the ~8 kb-long fragment obtained was excised from the gel with the MagExtractor®-PCR & Gel clean up (Toyobo), following the manufacturer's instructions. Using the excised fragment as DNA template, a sequencing library was prepared using the TruSeq Nano DNA Library Prep Kit (Illumina) and mtDNA sequencing was carried out in the Illumina MiSeq (Illumina) platform, following the manufacturer's protocols.

Sequence data analyses

The quality of the obtained sequences was evaluated. Only the sequences with quality over Q20 were extracted and were used for assembly into the mtDNA sequence (ca. 8 kb) on GS Reference Mapper Software (Roche). The corresponding region of the complete mitochondrial genome sequence of *B. xylophilus* (GQ332424) was used as reference. Contigs with an average depth of 451 for the 285 individuals were used for sequence alignment and for the identification of single nucleotide polymorphisms (SNPs) and insertions and deletions (indels) in Sequencher (Gene Codes). The exact coding region of each gene was determined based on the annotation listed for GQ332424 and on basic local alignment search tool (BLAST) analysis conducted on the National Center for Biotechnology Information (NCBI)

web platform (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Mitochondrial DNA polymorphism analysis

The evaluation of genetic polymorphisms, including the number of polymorphic sites, nucleotide substitutions, and transitions and transversions without indels, was carried out in DnaSP (Librado and Rozas, 2009). This software was also used to determine the number of haplotypes (Nh) and nucleotide and haplotype diversities. Haplotype diversity (Hd: Nei and Tajima, 1981) and Nh were computed to evaluate mitogenomic diversity in each population. MEGA (Kumar et al., 2016) was used to translate protein-coding genes based on the invertebrate mitochondrial codon usage table, and synonymous and non-synonymous substitutions in protein-coding regions were identified. The number of variations, including nucleotide substitutions in genes and noncoding regions, and variations in amino acid sequences, were also determined. The significant differences in SNP densities between genes or regions were examined by Student's *t*-test in Excel (Gosset, 1908).

Phylogenetic analysis

Haplotype distribution and diversity (Hd) in each population were investigated independently. Haplotype sequence data was imported into MEGA to build a phylogenetic tree of haplotypes using the maximum likelihood (ML) method based on the Kimura 2-parameter model (Kimura, 1980), with 1000 bootstrap replications. Phylogenetic relationships among the 12 populations in Kyushu were analyzed using haplotype frequencies in populations. Haplotype composition and frequency of haplotypes in each population were used to calculate genetic differentiation (G_{ST}), gene flow (Nm), and genetic distances (Ds) among populations. A phenogram was also drawn for the 12 populations based on the neighbor-joining (NJ) method (Saitou and Nei, 1987).

3.3. Results

Sequence statistics

A fragment of about 8,060 bp, corresponding to 54.5% of the PWN mitochondrial genome (ca. 14.8 kb), was sequenced. In the PWN mitochondrial genome, there is an extremely AT-rich and extensive non-coding region of about 1.7 kb (13,129–14,778 of GQ332424), while all mitochondrial genes are placed in the remaining 13.1 kb (Sultana et al., 2013). The sequence decoded in the present study corresponds to 61.4% of the latter fragment.

Analysis of SNPs

One hundred and fifty-eight SNPs were detected in mtDNA sequences, corresponding to one SNP per 51 bp, on average (Table 3-2). Twenty-two genes (eight protein-coding genes, 12 transfer RNA (tRNA) genes, and two rRNA genes) and six non-coding regions were confirmed in the 8,060 bp sequence. The number of SNPs in each gene/region is shown in Table 3-2. Nine tRNA genes (*trnD*, *trnG*, *trnH*, *trnA*, *trnP*, *trnV*, *trnW*, *trnE*, and *trnY*) showed no SNPs. Thus, the 158 SNPs corresponded to 139 SNPs from protein-coding genes, three SNPs from the remaining three tRNA genes, 13 SNPs from two rRNA genes, and three SNPs from non-coding regions.

Eighty-eight percent (139/158) of the SNPs were found in the protein-coding genes, leading to an average occurrence of one SNP/41 bp. Most of these SNPs (40 SNPs, 1 SNP/39 bp) were found in *nad5*, but *cox1* showed the highest average occurrence of SNPs (39 SNPs, 1 SNP/30 bp). Genes coding for NADH dehydrogenase subunits 1 and 6 (*nad1* and *nad6*) presented high number and average occurrence of SNPs (22 SNPs, 1 SNP/40 bp and 13 SNPs, 1 SNP/33 bp, respectively). Genes coding for NADH dehydrogenase subunit 3 (*nad3*) and ATP synthase subunit 6 (*atp6*) showed a relatively low average occurrence of SNPs, as only

six (1 SNP/57 bp) and seven (1 SNP/60 bp) SNPs were found in each sequence, respectively. Cytochrome *c* oxidase subunit 2 (*cox2*) was the most conservative among the eight protein-coding genes (seven SNPs, 1 SNP/99 bp). Because only three SNPs were detected in three of the 12 tRNAs for which nucleotide sequences were decoded (i.e., *trnC*, *trnM*, and *trnS*), the average occurrence of SNPs in tRNA genes was extremely low (1 SNP/223 bp). According to the predicted secondary structure of mitochondrial tRNA genes, 20 of the 22 tRNA genes identified in PWN have a unique structure; a loop of variable size (TV-replacement loop; 6–12 bp) is thus displayed instead of the TΨC arm loop (Sultana et al., 2013). In the present study, the SNPs of *trnC* and *trnS* occurred in the TV-replacement loop and TΨC arm loop, respectively; the SNP of *trnM* occurred close to its 5' extremity. The two rRNA genes (*rrnS* and large subunit rRNA gene (*rrnL*)) found here contained six and seven SNPs, respectively, thereby showing a lower average occurrence of SNPs than the protein-coding genes (1 SNP/116 bp for *rrnS* and 1 SNP/134 bp for *rrnL*). The total length of the six non-coding regions identified here was 50 bp (1–20 bp in each region). Generally, non-coding regions are the most variable in the genome (Van der Veer and De Vries, 2004), which was also revealed for the partial mitochondrial genome of PWN. For example, the region comprising *cox1-trnC* was only 15 bp but contained two SNPs, thereby presenting an extremely high occurrence rate of SNPs (1 SNP/7.5 bp). In the *trnD-trnG* region, one SNP was detected in three base pairs. The differences in SNP densities between protein-coding genes and tRNA genes or rRNA genes were significant ($P \leq 0.01$) according to the *t*-test.

The 158 SNPs corresponded to 122 transitions (Ts) and 36 transversions (Tv), leading to a Ts/Tv ratio of 3.4 (Table 3-2). One hundred and fourteen of the 122 Ts were found in the protein-coding genes. These genes presented Ts/Tv ratios of 2.5–8.8 (average 4.6), which indicated that Ts were more frequent that Tv. The *cox1* gene contained most Ts and showed the highest Ts/Tv ratio (8.8). *Nad*5, which harbored the largest number of SNPs, held more Tv

than *cox1* and therefore showed a lower Ts/Tv ratio (3.0). Only one Tv was identified among the five SNPs observed in NADH dehydrogenase subunit 4L gene (*nad4L*). On the other hand, no Tv appeared in the *cox2*. The Ts/Tv ratio of the two rRNA genes was 1.2, which means that the number of Ts and Tv was almost identical. Only Tv were detected in all of the tRNA genes, while in the non-coding region there were two Tv in *cox1-trnC* and one Ts in *trnD-trnG* (Table 3-2).

Amino acid variation was examined for all SNPs detected in the protein-coding genes. The numbers of synonymous and non-synonymous substitutions (Syn and Non-syn, respectively) in amino acids were 118 and 20, respectively (Table 3-2). Within the 118 Syn, 104 were Ts and 14 were Tv. Ninety-seven percent (115) of the Syn were due to mutations at the third position of the codon, while the remaining three Syn were caused by mutations at the first codon position (C/T substitution in *nad5*, C/T substitution in *nad6*, and C/T substitution in *nad1*). Ten (50%) of the 20 Non-syn were caused by mutations at the second codon position, seven were due to mutations at the first codon position, and two were due to mutations at the third codon position. One Non-syn recognized in *nad5* corresponded to simultaneous changes in the second and the third positions of the codon. The overall Syn/Non-syn ratio was 5.9. Thus, Syn occurred about six times more frequently than Non-syn, which means that most variations in the partial mitochondrial genome of PWN could be considered neutral (Kimura, 1983). However, the Syn/Non-syn ratio in *nad3* (2.0) and *nad5* (3.3) were lower than in other genes, suggesting that relatively high selective effects might have occurred in these two genes.

Analysis of indels

Because the mitochondrial genome of PWN is extremely AT-rich (Sultana et al., 2013), it is difficult to accurately estimate the length variation of long homopolymer (poly-A and

poly-T) regions (Linnertz et al., 2012). Therefore, a PCR was employed in the present study to allow the occurrence of artificial length variation due to possible PCR slippage in simple sequence repeats (Hauge and Litt, 1993; Murray et al., 1993). All indels (length variation) recognized in homopolymer regions were excluded due to the low quality of the identification of nucleobases (i.e., Q score) of their last parts in many individuals. Only two indels were detected outside homopolymer regions and no indels were found in the eight protein-coding genes (Table 3-2). Only one of the 12 tRNA genes (*trnY*) displayed one-base deletion in its anticodon stem, which was found in one of the 285 nematodes investigated. In the *cox1-trnC* inter-genic spacer region, one four-base deletion was observed (Table 3-2).

Haplotype distribution and diversity

Haplotype analysis was performed for the 285 PWNs sampled, and 30 haplotypes (hts-01–30 in Table 3-3) were detected based on 160 polymorphic sites that consisted of 158 SNPs and two indels. Polymorphic sites of haplotypes and the position of each SNP in relation to the reference sequence (GQ 332424) are displayed in Appendix (Table S3).

The frequency of each haplotype detected in each of the 12 investigated populations is shown in Table 3-3. The most frequent haplotype in Kyushu was ht-21 (34.4%, 98 nematodes), although ht-01 (15.1%, 43 nematodes) and ht-13 (14.7%, 42 nematodes) also showed relatively high frequency. On the other hand, 14 haplotypes (ht-05, hts-08–09, hts-11–12, hts-16–19, ht-23, hts-25–26, and hts-29–30) were identified as extremely rare, as they were detected in only one nematode. Twenty of 30 haplotypes were population-specific, while the remaining ten haplotypes were distributed in multiple populations. The most frequent haplotype (ht-21) was observed in seven of the 12 populations, including all six populations in the southern region of Kyushu (Shintomi, Miyazaki, North Nichinan, South Nichinan, Sendai, and Ibusuki) and in the central western region (Amakusa). Haplotypes 10, 13, and 14

were found in the northern region (Karatsu, Itoshima, Yukuhashi, and Chikujo) and central western region (Amakusa), and four haplotypes (hts-01–04) were detected in the northern and southern Kyushu regions. Therefore, haplotype composition notably differed among regions.

Haplotype diversity in the 12 populations varied from 0.30 to 0.83, and its average was 0.55 (Table 3-3); in eight populations it was above 0.5. Particularly, populations from Amakusa, Karatsu, and Ibusuki showed Hd above 0.8, which is considered high haplotype polymorphism, while populations from Sendai, Miyazaki, and Itoshima showed Hd below 0.4 and Nh below 4. The haplotype diversity for the entire Kyushu region was 0.83.

The haplotypes found in three populations from northeastern Kyushu (Itoshima, Yukuhashi, and Chikujo) were similar to each other, and common haplotypes were also detected in four populations from southeastern Kyushu (Shintomi, Miyazaki, North Nichinan, and South Nichinan).

Phylogenic analysis of haplotype data

A phylogenetic tree for the 30 haplotypes was constructed based on the ML method, using the 158 SNPs identified in the present study (Fig. 3-2). The closely related *B. mucronatus* (Accession GU177865) was employed as outgroup to determine the root of the phylogenetic tree.

The dendrogram clustered the 30 haplotypes into two Clades (Clade I and Clade II). A major phylogenic difference was observed between the two clades; there was a minimum difference of 58 SNP sites (between ht-26 of Clade I and ht-25 of Clade II) and a maximum difference of 105 sites (between ht-27 of Clade I and hts-02/03/21 of Clade II). In contrast, no significant difference emerged among haplotypes within each clade. Among the 24 haplotypes within Clade I, 91 sites differed between hts-10 and 30, while only four sites differed between hts-25 and 28, among the six haplotypes within Clade II. Clade I was further divided into two

subclades (Subclades I-1 and I-2) and three single haplotypes (hts-17, 26, and 10). The sequences of these three haplotypes differed from that of the other 21 haplotypes contained in Clade I. Subclades I-1 and I-2 consisted of 11 and 10 haplotypes, respectively. However, the maximum disagreement within Clade I consisted of only 16 SNP sites between ht-30 (Subclade I-1) and ht-12 (Subclade I-2), and no significant difference was identified between the two subclades. Most of the haplotypes within Subclade I-1 were distributed in the northern Kyushu region, haplotypes within Clade II were mainly distributed in southern Kyushu region, and haplotypes in Subclade I-2 were detected in both northern and southern Kyushu regions.

Phylogenetic relationships among the mitogenomic sequences of six isolates (Ibaraki, Japan: AP017463; Korea: GQ332424 and NC023208; and Portugal: JQ429761, JQ514067, and JQ514068) were also analyzed. No sequence polymorphisms were identified for isolates within the same area in Korea and Portugal and, therefore, a phylogenetic tree was built after aligning three mitogenomic sequences (Ibaraki: AP017463, Korea: GQ332424, and Portugal: JQ429761) and the 30 haplotype sequences from Kyushu (hts-01–30). As a result, the Ibaraki (East Japan) sequence was assigned to Subclade I-1, whereas sequences from Korea and Portugal were allocated to Clade II. Sequences from Ibaraki and Portugal were highly similar to ht-13 and hts-21/23, respectively. The Korean sequence was not closely related to seven of the haplotypes within Clade II (hts-02, 03, 21, 23, 25, 28, and JQ429761). Comparing the Korean sequence with the other 32 sequences (haplotype and mitochondrial sequences) revealed 35 SNPs and eight indels in two rRNA genes, and no significant differences for the eight protein-coding genes and 12 tRNA genes.

Genetic relationships among populations

Genetic distances among populations were 0.14–0.92, based on haplotype frequency in

each population (Table 3-4). The maximum distance was obtained between Miyazaki (southeastern region) and Karatsu (northwestern region), and no common haplotypes were detected. On the other hand, the Ds among three populations in northeastern Kyushu (Itoshima, Yukuhashi, and Chikujo) and the Ds among four populations in southeastern Kyushu (Shintomi, Miyazaki, North Nichinan, and South Nichinan) were 0.14–0.37 and 0.14–0.35, respectively. Geographical distances among populations within each group were very small.

The NJ tree constructed for the 12 populations based on Ds (Fig. 3-3) revealed that the three populations in northeastern Kyushu and the four populations in southeastern Kyushu formed cohesive clades. Three populations, located in the northwestern region (Matsuura and Karatsu) and southernmost region (Ibusuki), formed a loose clade. The other two populations (Amakusa in the central western region and Sendai in the southwestern region) were clade-independent.

The G_{ST} and Nm among populations were 0.33 and 1.01, respectively.

3.4. Discussion

In the present study, a method for performing long-PCR using single-nematodes directly was developed for individually sequencing each mitogenome. Sequences of about 8 kb, corresponding to 55% of the whole mitochondrial genome of PWN, were obtained for 285 nematode individuals using this method. One hundred and sixty polymorphic sites were detected in these sequences, leading to a frequency of polymorphism as high as one polymorphic site per 51 bp. However, a low SNP frequency was detected in tRNA genes, although this is consistent with tRNAs gene sequences being more conservative than that of other genes (Van der Veer and De Vries, 2004). The mitogenomic polymorphisms of the PWN have been mainly used in interspecies comparative studies (Sultana et al., 2013; Pereira et al.,

2013). Their application to intraspecific studies is limited, and a lack of polymorphisms' evaluation is evident.

Mitogenome sequences (three complete and three incomplete) of six PWN isolates, one from Japan, two from Korea, and three from Portugal, have been reported so far. No variations were detected between the two Korean sequences (14,788 bp) or among the three Portuguese sequences (about 12 kb). However, 171 polymorphic sites containing 150 SNPs were confirmed among the sequences from Ibaraki, Korea, and Portugal, whereas sequences from Kyushu alone showed 158 SNPs.

Sequences of four PWN mitochondrial genes (*cox1*, *cyt*b, *nad5*, and *rrnS*) have been revised to date. In *cox1*, 52 SNPs were found among 36 isolates from Japan, Korea, China, Portugal, Canada, USA, and Mexico. Eighteen SNPs were identified in *cyt*b sequences of 20 isolates obtained from Japan, Korea, China, Portugal, and USA, and 11 SNPs were observed in *nad5* sequences from 17 isolates derived from these five countries. Thirteen polymorphic sites containing two SNPs and 11 indels were observed in *rrnS* sequences from six isolates. Except for the *cox1*, which has been sufficiently investigated on relatively large sample sizes, sequence diversity in mitochondrial genes is generally not high. However, the present study revealed 39 SNPs in *cox1*, 40 SNPs in *nad5*, and six SNPs in *rrnS* in sequences from the Kyushu region alone, which can be considered a very limited area compared to the areas examined in the studies mentioned above.

Past genetic studies on PWN employed exclusively DNA extracted from one isolate. As an isolate consists of many nematodes, several genotypes are included and sequencing based on isolates masks many low-frequency variations. Therefore, it is presumed that only high-frequency variations could be treated as sequences of the isolate, which seriously underestimates variation. Sequencing based on the method employed here allowed reading sequences of individual mitogenomic sequences, thereby detecting rare alleles/haplotypes. A

considerably large sample size (285 individuals) might also contribute to correctly estimate diversity. Unlike the worldwide PWN collections performed in previous studies, the present study targeted the Kyushu region only, and used a localized collection of PWN within an extremely small area. Nevertheless, it was evident that extremely high sequence diversity was preserved in the mitochondrial genome of PWN. Sequencing based on individuals instead of isolates is therefore a powerful approach for population genetic studies.

In the phylogenetic tree produced for haplotypes, 12 (ht-01, ht-04–14) out of the 14 haplotypes distributed in northern Kyushu belonged to Clade I. On the contrary, there was no evident trend in the haplotypes from southern Kyushu (Table 3-3, Fig. 3-2). Sasebo (Nagasaki prefecture) was the first invasion point of PWN in Asia (Yano, 1913) from which PWN gradually spread to Shikoku and Honshu (Futai, 2008). It was then disseminated from Japan to Korea and China (Mamiya and Shoji, 2009). Furthermore, PWN populations appearing in Portugal in 1999 were assumed to have originated from East Asia (Metge and Bürgermeister, 2006). Based on the method presented here, the propagation route of PWN could be further examined using mitochondrial haplotypes from invaded countries (Japan, Korea, China, and Portugal) and native countries (USA and Canada).

Based on the past records of pine wilting and timber import and movement, Kishi (1988) estimated three possible routes (Sasebo route, Nichinan route, and Aira route) through which PWN invaded and expanded into Kyushu. In the Sasebo route, PWN invaded from Sasebo, northwestern Kyushu in 1905, and extensively dispersed into the northern region of Kyushu. Because ht-13 was the major haplotype in the three populations from northeastern Kyushu (Itoshima, Yukuhashi, and Chikujo), this haplotype might characterize populations dispersed through the Sasebo route. Haplotype-13 was also relatively frequent in Amakusa in central western Kyushu, it was likely held with high frequency in the PWNs of the Sasebo route and then transmitted to northern Kyushu and Amakusa regions. According to Kishi (1988),

pinewoods withered by PWD in Nagasaki prefecture were transferred to a pulp mill in Yatsushiro (Kumamoto prefecture), extending PWD to the central western region of Kyushu. A PWN group that irrupted from Nichinan in Miyazaki prefecture (Nichinan route) in 1939 expanded north and south along the east coast of Kyushu (Kishi, 1988). The PWN collections from four populations (Shintomi, Miyazaki, North Nichinan, and South Nichinan) in the coastal area of Miyazaki prefecture showed similar haplotype compositions, and ht-21, which presented an extremely high frequency, might be considered as the specific haplotype of the Nichinan route. The PWNs expanding through the Aira route (third route) invaded from Aira (Kagoshima prefecture) in 1942 and dispersed to the western region of Kagoshima prefecture (Sendai and Ibusuki) (Kishi, 1988). Although four haplotypes appeared in the Sendai population, most nematodes (84%) were ht-21. Because this haplotype might be considered specific to the Nichinan route and haplotype composition in Sendai was somewhat similar to that of populations dispersing through the Nichinan route, the Sendai population might have been recently influenced by the Nichinan route. However, the three low-frequency haplotypes (hts-24–26) that were not observed in other populations might persist as a vestige of the Aira route. The Ibusuki population, which was also regarded as dispersed through the Aira route, shared only one haplotype (ht-21) with Sendai and its frequency was quite low. The Ibusuki population, however, shared three haplotypes (hts-02–04) with Matsuura, northwestern Kyushu, which is geographically distant (Table 3-3, Fig. 3-3), suggesting that artificial long-range migration might have occurred. Thus, in the region occupied by Aira route PWNs, profound changes in haplotype composition might be rapidly progressing by natural migration or artificial transfer from other nematode strains (Fig. 3-1).

Analysis of the genetic structure of the PWN population in Kyushu revealed high genetic differentiation among the 12 populations (G_{ST} = 0.331). While 33.1% of the total genetic variability was due to variation among populations, 66.9% was due to variation within

populations (Mouhaddab et al., 2015). The Nm among populations was nearly 1 (1.01) evidencing that PWN migration among populations had little effect on the genetic differentiation within the Kyushu region (Wright, 1951; Lowe et al., 2004). In the NJ tree of the 12 populations (Fig. 3-3) based on Ds (Table 3-4), the three populations within northeastern Kyushu (Itoshima, Yukuhashi, and Chikujo) and the four populations within southeastern Kyushu (Shintomi, Miyazaki, North Nichinan, and South Nichinan) formed cohesive clades. Average G_{ST} values for the three populations of northeastern Kyushu and four populations of southeastern Kyushu were 5.7% and 1.9%, and their Nm values were 8.3 and 24.4, respectively. Thus, genetic differentiation among populations within these areas was much smaller than within the whole Kyushu area. Additionally, due to the large value of Nm, it could be inferred that genetic differentiation within the referred local regions occurred due to genetic drift (Wright, 1951; Lowe et al., 2004).

When the three population in northeastern Kyushu and four populations in southeastern Kyushu were grouped as metapopulations and analyzed with the remaining five populations (Matsuura, Karatsu, Amakusa, Sendai, and Ibusuki), average G_{ST} and Nm for these seven populations were recalculated as 32.1% and 1.06, respectively. Because Nm was nearly 1, the influence of gene flow on the genetic structure of those areas did not appear to be substantial (Wright, 1951; Lowe et al., 2004). In nature, PWD was expanded because PWN was propagated by vector insects. Due to the severe decline of the damaged pine forests in the past, the fragmentation of PWN populations has progressed and this geographical separation might have decreased the natural long-distance gene flow. Matsuura and Karatsu populations in the northwestern region and the Ibusuki population in the southernmost region formed a loose clade in the NJ tree (Fig. 3-3). Matsuura and Karatsu are geographically close, whereas Ibusuki is geographically distant from the other two (Fig. 3-1). As Ds among the three populations were not too large (Table 3-4), a long-range migration of PWN might have

occurred by human activity.

According to Nose et al. (2009), the two populations of Ibusuki and Sendai in southwestern Kyushu were similarly affected by the PWN expanded from Aira in Kagoshima prefecture, as these two populations presented similar haplotype composition and shared a common dominant haplotype that was not very frequent in other areas. However, in the present study, the haplotype composition of the two populations was quite distinct. Eighty-four percent of the individuals in Sendai showed the haplotype (ht-21) appearing in the four populations (Shintomi, Miyazaki, North Nichinan, and South Nichinan) of the southeastern region. In Ibusuki, which is located in southernmost area of Kyushu, the only haplotype common to Sendai was ht-21, but its frequency was quite low.

Population size of nematodes invading an area seems to drastically decrease following the decline of the pine forest, and small populations are maintained or nearly extinguish if the pine forest is extensively destroyed. Therefore, through migration from other areas, the genetic composition of these populations will drastically change. In extremely small populations, drastic changes might occur by random drift (Miko et al., 2009), which might be responsible by the dynamics observed in the distribution of PWN strains in the Kyushu region. In the present study, PWN sampling was performed in 2012–2014, whereas that of Nose et al. (2009) was performed in 2006–2007. Therefore, population composition changes seem to have occurred within a single decade. Because there is a distance limit for vector insects' natural travel, human activities might be heavily involved in genetic dynamics.

Investigations on PWN genetic variations have been conventionally concentrated on the nuclear genome. In recent years, the number of studies on the mtDNA of PWN gradually increased, although its utilization mainly involved comparing *B. xylophilus* with close relative species (Pereira et al., 2013; Sultana et al., 2013; Moreira et al., 2014). In Nematode, it has been widely confirmed that there are intraspecific mutations in mitogenomes (Thomas and

Wilson, 1991), and intraspecific mutations are also recognized in some genes of PWNs (Valadas et al., 2013). In the present research, intraspecific variation surveys quantitatively evaluated by the mitochondrial genome scale.

In previous intraspecific variation studies, a few mitochondrial genes were used, and sample (isolate) sizes were small. Furthermore, there was no analysis of individual nematodes. All these facts hampered the detection of detailed intraspecific variation (Valadas et al., 2013). In the present study, we clarified the high diversity of PWN mitochondrial genome by using a wide range of nucleotide sequence information and sufficient number of individuals (285 individuals) that enabled genetic population analysis based on haplotype analysis. Furthermore, the present study evidenced that the haplotype analysis of a single mitochondrial genome sequence is a valid approach for several aspects of PWN studies, including population phylogenetic analysis, genetic diversity evaluation, artificial migration possibility, and invasion and expansion pathways' elucidation.

Tables

Table 3-1. Summary information for regional populations of pinewood nematode in Kyushu

Population	Code	N[a]	No. of trees[b]	Location	Prefecture
Matsuura	Ma	23	16	33.3°N, 129.7°E	Nagasaki
Karatsu	Ka	27	10	33.5°N, 129.9°E	Saga
Itoshima	It	25	25	33.6°N, 130.2°E	Fukuoka
Yukuhashi	Yu	14	14	33.7°N, 131.0°E	Fukuoka
Chikujo	Ch	12	12	33.7°N, 131.1°E	Fukuoka
Amakusa	Am	28	10	32.6°N, 130.4°E	Kumamoto
Shintomi	Sh	24	20	32.1°N, 131.5°E	Miyazaki
Miyazaki	Mi	26	20	31.9°N, 131.4°E	Miyazaki
North Nichinan	Ni-N	28	4	31.6°N, 131.4°E	Miyazaki
South Nichinan	Ni-S	27	20	31.6°N, 131.4°E	Miyazaki
Sendai	Se	25	20	31.8°N, 130.2°E	Kagoshima
Ibusuki	Ib	26	20	31.2°N, 130.6°E	Kagoshima
Total		285			

[a] Number of individuals analyzed in each population.
[b] The number of trees in each population that became the source of the sequenced nematodes

Table 3-2. Sequence polymorphism and substitutions identified in the mitochondrial genome of pinewood nematode

Gene / region		Length[a] (bp)	Sequence polymorphism				Substitution					
							Nucleotide			Amino acid residue		
			No. Indels	Indels/bp[b]	No. SNPs	SNPs/bp[b]	No. Ts[c]	No. Tv[c]	Ts[c]/ Tv[c] ratio	No. Syns[d]	No. Non-syns[d]	Syn[d]/ Non-syn[d] ratio
Protein-coding gene	*cox1*	1156			39	1/30	35	4	8.8	34	5	6.8
	cox2	690			7	1/99	7	0	–	6	1	6.0
	nad3	342			6	1/57	5	1	5.0	4	2	2.0
	nad5	1569			40	1/39	30	10	3.0	30	9	3.3
	nad6	435			13	1/33	11	2	5.5	11	2	5.5
	nad4L	234			5	1/47	4	1	4.0	5	0	–
	nad1	873			22	1/40	17	5	3.4	21	1	21.0
	atp6	417			7	1/60	5	2	2.5	7	0	–
	subtotal	5716			139	1/41	114	25	4.6	118	20	5.9
tRNA	*trnC*	54			1	1/54	0	1	–			
	trnM	54			1	1/54	0	1	–			
	trnD	55										
	trnG	54										
	trnH	55										
	trnA	55										
	trnP	57										
	trnV	57										
	trnW	55										
	trnE	55										
	trnS	55			1	1/55	0	1	–			
	trnY	63	1	1/63								
	subtotal	669	1	1/669	3	1/223	0	3	–			
rRNA	*rrnL*	937			7	1/134	4	3	1.3			
	rrnS	697			6	1/116	3	3	1.0			
	subtotal	1634			13	1/126	7	6	1.2			
Non-coding region	*cox1-trnC*	15	1	1/15	2	1/7.5	0	2	–			
	trnC-trnM	7										
	trnM-trnD	20										
	trnD-trnG	3			1	1/3	1	0	–			
	nad3-nad5	1										
	rrnS-trnS	4										
	subtotal	50	1	1/50	3	1/17	1	2	0.5			
Total		8060[e]	2	1/4030	158	1/51	122	36	3.4	118	20	5.9

[a] The nucleotide sequence length for that gene or region

[b] The average occurrence of single nucleotide polymorphisms (SNPs) or insertions/deletions (Indels) in that specific part of the sequenced mitochondrial genome

[c] Ts, transitions; Tv, transversions

[d] Syn, synonymous substitution; Non-syn, non-synonymous substitution

[e] The total length contains 9 bp overlapping genes or regions

Blank cells indicate no value; – means the value can't be calculated.

Table 3-3. Haplotype distribution within in each population of pinewood nematode

Haplotype (Ht-)	Regional population												Total	Frequency (%)
	Ma	Ka	It	Yu	Chi	Am	Sh	Mi	Ni-N	Ni-S	Se	Ib		
01	16						8	5	4	10			43	15.1
02	3											9	12	4.2
03	1	8										1	10	3.5
04	3	3										1	7	2.5
05		1											1	0.4
06		8											8	2.8
07		2											2	0.7
08		1											1	0.4
09		1	1										1	0.4
10		3	1	6	3								13	4.6
11													1	0.4
12			20		1								1	0.4
13			3	7	8	7							42	14.7
14				1		4							8	2.8
15						2							2	0.7
16						1							1	0.4
17						1							1	0.4
18						1							1	0.4
19						1							1	0.4
20						9	1						10	3.5
21						2	14	21	21	15	21	4	98	34.4
22							1		3	1			5	1.8
23										1			1	0.4
24											2		2	0.7
25											1		1	0.4
26											1		1	0.4
27												7	7	2.5
28												2	2	0.7
29												1	1	0.4
30			4									1	1	0.4
Nh[a]	4	8	4	3	3	9	4	2	3	4	4	8		
Hd[b]	0.50	0.82	0.36	0.60	0.53	0.83	0.57	0.32	0.42	0.57	0.30	0.80	0.83	

Numbers indicate the number of individuals present in each haplotype. Blank cells indicate that no individual presented the haplotype.

[a] Number of haplotypes detected in the population

[b] Haplotype diversity, $Hd = (1 - \sum xi^2) n / (n - 1)$, where xi is the frequency of the ith haplotype and n is the sample size (Nei and Tajima, 1981)

Table 3-4. Pairwise genetic distances between the 12 populations of pinewood nematode examined based on haplotype frequency

Population	Ma	Ka	It	Yu	Ch	Am	Sh	Mi	Ni-N	Ni-S	Se
Matsuura (Ma)											
Karatsu (Ka)	0.570										
Itoshima (It)	0.693	0.665									
Yukuhashi (Yu)	0.703	0.738	0.144								
Chikujo (Ch)	0.703	0.714	0.367	0.288							
Amakusa (Am)	0.773	0.671	0.549	0.598	0.706						
Shintomi (Sh)	0.560	0.771	0.693	0.703	0.703	0.465					
Miyazaki (Mi)	0.570	0.916	0.752	0.714	0.714	0.752	0.347				
North Nichinan (Ni-N)	0.549	0.809	0.703	0.693	0.693	0.673	0.144	0.203			
South Nichinan (Ni-S)	0.560	0.752	0.693	0.703	0.703	0.652	0.223	0.347	0.144		
Sendai (Se)	0.693	0.752	0.693	0.703	0.703	0.662	0.560	0.570	0.703	0.560	
Ibusuki (Ib)	0.464	0.560	0.752	0.809	0.809	0.678	0.665	0.811	0.714	0.665	0.665

Figures

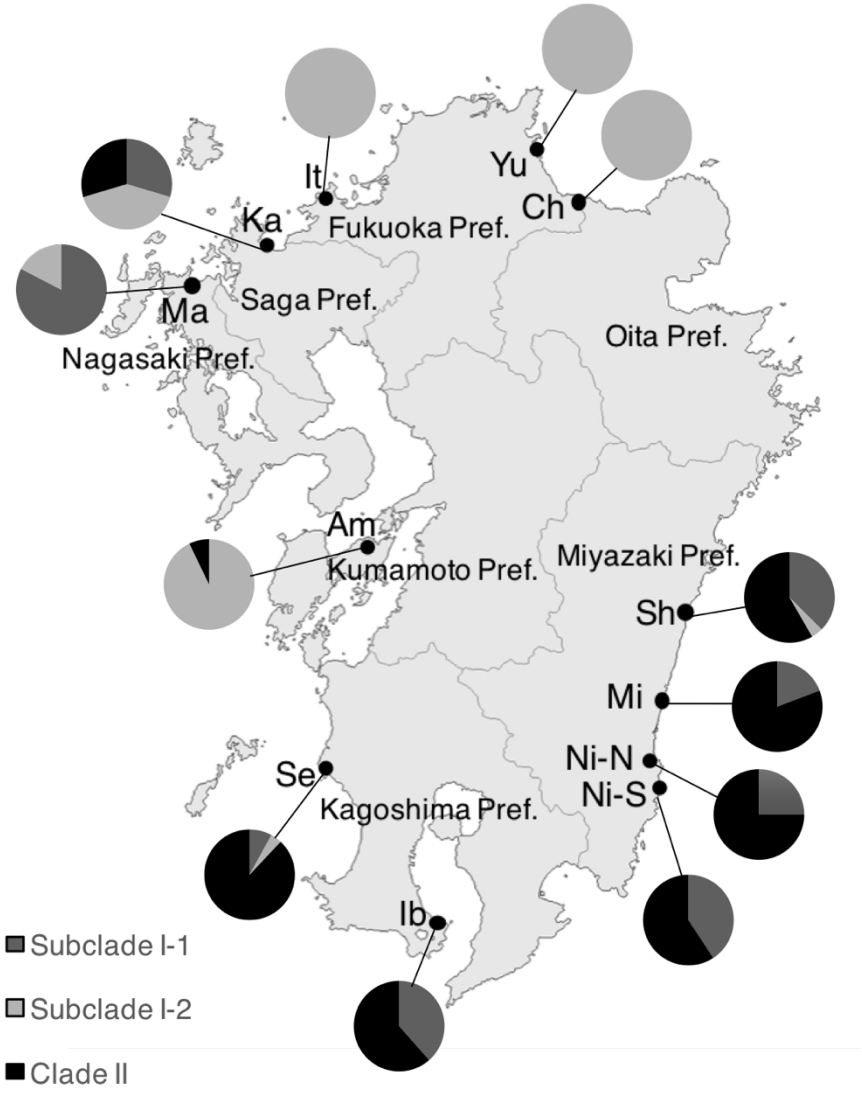

Fig. 3-1. Location of the studied populations of pinewood nematode (black dots) in Kyushu and the group of haplotype distribution for each population

The names of populations are written in code as described in Table 1.
The classification of haplotypes is described as in Fig. 3-2.

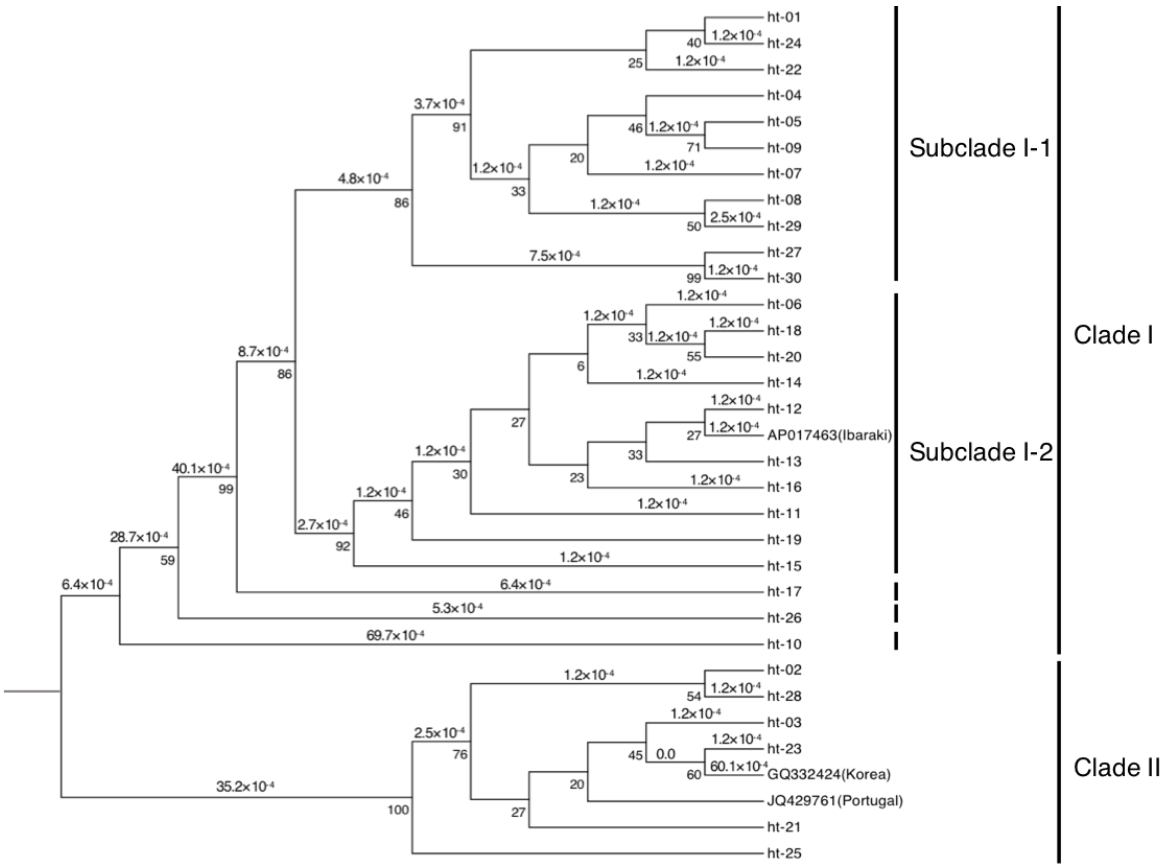

Fig. 3-2. Phylogenetic tree for 33 mitochondrial haplotypes of pinewood nematode containing 30 from Kyushu and 3 from previous studies using the Maximum Likelihood (ML) method based on the Kimura 2-parameter model

Numbers under the branches represent support values from bootstrap replications of 1,000.
Evolutionary distances greater than 0.0001 are shown above branches.
Wide bars indicate the major clades clustering.

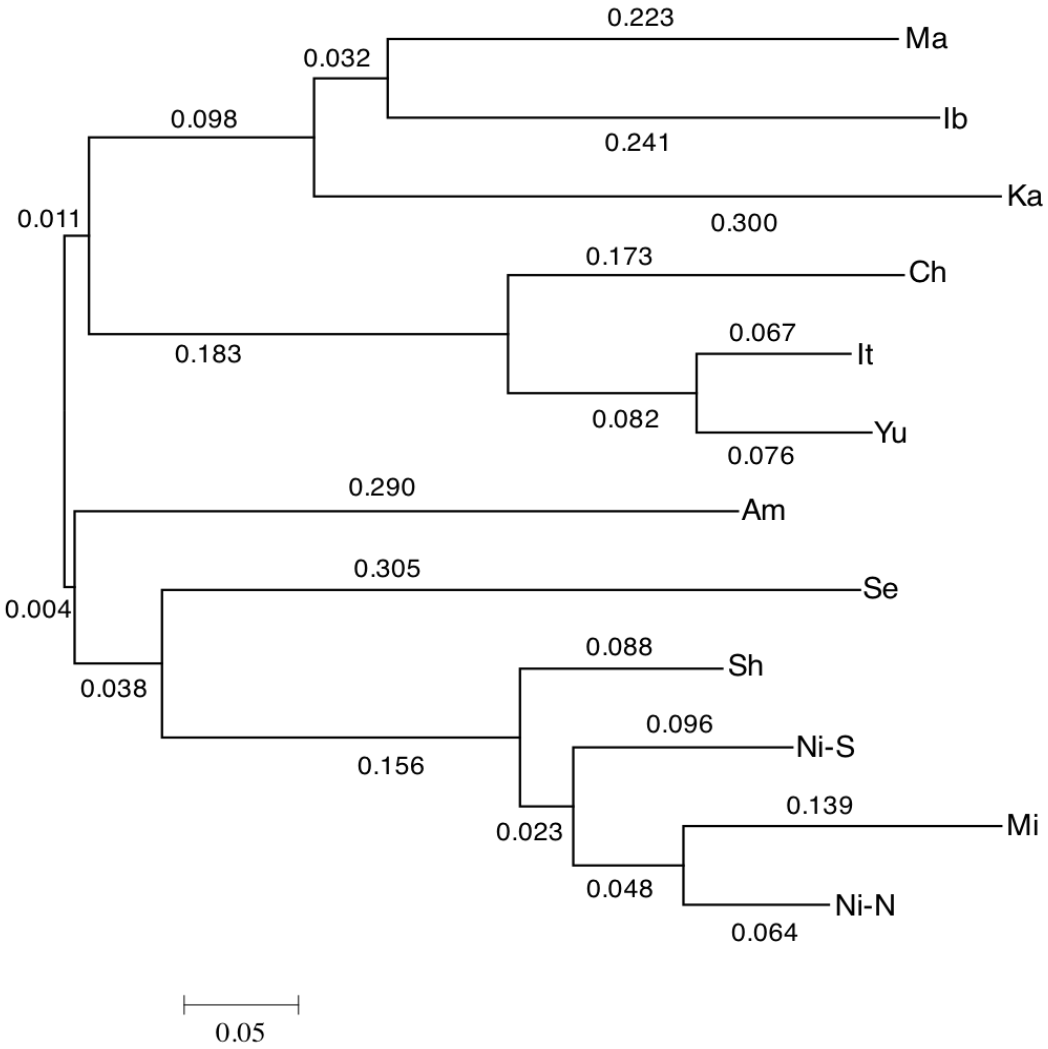

Fig. 3-3. Neighbor-joining (NJ) tree diagram showing the genetic relationship that was determined using genetic distances among 12 populations of pinewood nematode

Numbers next to corresponding nodes indicate the branch lengths in the same units of the genetic distances that are computed based on the haplotype frequency for each population.

Chapter 4. General Discussion

Genetic diversity is a measure of the genetic differences that exist within a population, which is often called genetic variation (Brooker et al., 2017). It facilitates the long-term survival of a population by enabling some individuals to adapt to changes in environment (Brooker et al., 2017). As the agent of pine wilt disease (PWD), clarifying the genetic variation of pinewood nematode (PWN) may help us to understanding how this kind of invasive organism has survived for such a long time. Genetic variation might also be used to infer the transmission route of PWN to understand how it be expanded throughout the world, enabling us to take measures to prevent the spread of PWN to uninfected regions.

The present study evaluated the genetic diversity from two cellular organelles which contain genetic material: the nucleus and mitochondria. Owing to the spread of next generation sequencers in recent years making it possible to collect sequence information with high throughput regardless of species, the present study utilized next generation sequencing to excavate the high genetic diversity of PWN in Kyushu area.

For the nuclear genome analysis, ten loci were identified from EST information of *Bursaphelenchus xylophilus* to assess the nuclear genome polymorphism of *B. xylophilus* from eight populations of Kyushu. With high-throughput sequencing, the average and standard deviation of the number of sequences obtained per subpopulation of each locus were 116.0 ± 33.8 (pX070) to 201.2 ± 63.9 (pX144). The genetic variability of each locus was 0.487 (pX021) to 0.790 (pX155) for all of Kyushu, leading to an average of 0.628 indicative of relatively high genetic diversity existing in the nuclear genome of *B. xylophilus* in Kyushu. The gene diversity (H_t) for eight populations ranged from 0.12 to 0.59. Sendai, Shintomi, Matsuura, and Karatsu populations were evaluated as rich in gene diversity (0.59, 0.57, 0.56, and 0.55) while Amakusa and Miyazaki populations showed extremely low gene diversity

(0.12 and 0.18). In order to evaluate the genetic structure in individual regional populations, the gene differentiation coefficient (G_{ST}) was calculated from the total gene diversity (H_T) and the average value of the genetic diversity of the subpopulation (H_S). As a result, considerable genetic differences among regional populations were observed, with a G_{ST} of 0.53 showing that more than half of the total gene diversity was possessed among regional populations. Populations with high gene diversity (Sendai, Shintomi, Matsuura, and Karatsu) also revealed high G_{ST} (0.43, 0.35, 0.25, and 0.25) indicating that the genetic compositions were notably different among the populations within damaged trees. On the contrary, in Amakusa and Miyazaki, populations with extremely low gene diversity and small G_{ST} (0.01 and 0.02) were confirmed, which showed that little genetic difference existed among subpopulations. It seemed that bottleneck effect or founder effect had a great impact on the formation of these regional populations. From the nuclear genome analysis, it is obvious that the genetic diversity of the regional populations was polarized in Kyushu. As an invasive species in Asia and Europe, previous reports consistently reported a lack of genetic diversity in *B. xylophilus* (Vieira et al., 2007; Cheng et al., 2008; Zhang et al., 2008; Fonseca et al., 2012). Nevertheless, the variability of the ten loci was consistently high in this study. This is because DNA markers such as SNPs, SSRs often used in many studies capture only mutations at specific sites on the sequence (Tautz and Renz 1984; Brookes 1999), whereas all mutations on the sequence can be utilized with the present method. Furthermore, it is considered to be an effective method when only the allele frequency of the surveyed individual is acquired, as in this study. Since already abundant EST nucleotide sequence information has been reported (Kikuchi et al. 2007), it is easy to increase the number of loci, thus it seems more detailed findings can be found in population genetic studies of *B. xylophilus* in the future.

For the mitochondrial analysis, a method for performing long-PCR using single-nematodes and sequencing nematode mitochondrial genomes individually was

developed. With the next generation sequencing, about 8 kb (~55%) of the complete mitochondrial genome was successfully obtained from 285 individuals collected from 12 populations. As a result, 163 polymorphic sites containing 158 SNPs and 5 indels were detected in mtDNA sequences corresponding to one SNP per 51 bp on average. Twenty-two genes (eight protein-coding genes, 12 transfer RNA (tRNA) genes, and two rRNA genes) and six non-coding regions were confirmed in the 8060 bp sequence. Eighty-eight percent (139/158) of the SNPs were found in the protein-coding genes, and the differences in SNP densities between protein-coding genes and tRNA genes or rRNA genes was significant ($P \leq 0.01$) according to the *t-test*. Haplotype analysis was performed for the 285 *B. xylophilus* and 30 haplotypes were detected. Haplotype diversity in the 12 populations varied from 0.30 to 0.83, with an average value of 0.55; in eight populations it was above 0.5. Haplotype composition notably differed among regions. The haplotype diversity for the entire Kyushu region was 0.83, confirming the remarkable high genetic diversity found in the mitochondria of *B. xylophilus* in Kyushu. Thirty haplotypes were clearly classified into two clades through phylogenic analysis of the haplotype data. Twelve of the 14 haplotypes that appeared in the regional population of northern Kyushu formed the same clade, but no clear trend was found in the haplotypes that appeared in the southern population. Genetic differentiation (G_{ST}) and genetic distance among the 12 populations was calculated from the haplotype frequency. The high genetic differentiation (0.331) might be due to past invasion and expansion routes of PWN in northeastern and southeastern Kyushu. Genetic distances among populations ranged from 0.14 to 0.92, and the NJ tree of the 12 populations based on genetic distances showed that the three populations within northeastern Kyushu (Itoshima, Yukuhashi, and Chikujo) and the four populations within southeastern Kyushu (Shintomi, Miyazaki, North Nichinan, and South Nichinan) formed cohesive clades. The distinct genetic composition of populations within the northwestern, central western, and southwestern Kyushu seem to be mostly related

to the extinction of pine forests and long-range migration of *B. xylophilus* due to human activity. In addition, it is suggested that due to the gene flow a dynamic has occurred during the distribution of the *B. xylophilus* among populations as compared with the reported information on the invasion propagation route of *B. xylophilus* in Kyushu. Before this study, the mitochondrial polymorphisms *B. xylophilus* had been reported as low (Metge and Bürgermeister, 2006; Valadas et al., 2013). This could be attributed to the use of pooled DNA of isolates causing only high-frequency variations to be treated as sequences of the isolate, which seriously underestimates variation. Sequencing based on the method employed here allowed reading of individual mitogenomic sequences, thereby detecting rare alleles/haplotypes. A large sample size (285 individuals) might also contribute to correct estimates of diversity, so the present study targeted the Kyushu region alone while confirming that extremely high sequence diversity was preserved in the mitochondrial genome of PWN. Direct long-PCR and sequencing of single nematode individuals is an effective method for investigating mitochondrial polymorphisms, and together these form an effective tool for PWN population genetics and other intraspecific studies.

On the other hand, this study also revealed differences in genetic diversity between the nuclear genome and the mitochondrial genome. For example, the Amakusa population that was assessed to have low diversity in the nuclear genome had high haplotype diversity in the mitochondrial genome. The same contradiction was observed in the Sendai population, which was evaluated to have high diversity in the nuclear genome, but low haplotype diversity in the mitochondrial genome. The most plausible explanation is based on the difference in the inheritance patterns of these two kinds of genome. This contradiction was also found in other reports. Johnson et al. (2003) reported contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. El Mokhefi et al. (2016) reported the contrasting patterns between mitochondrial and nuclear markers for the

genetic differentiation of the pine processionary moth. Differences in effective population size resulting from maternal inheritance of the mitochondrial may lead to different patterns, and the gender dispersion bias could also be a factor (Johnson et al., 2003; El Mokhefi et al., 2016). As differences exist between nuclear and mitochondrial results, many studies have committed to compare the effectiveness of nuclear and mitochondrial genome analysis (FitzSimmons et al., 1997; Schrey and Heist, 2003; Chappell et al., 2004; Canino et al., 2010; Ji et al., 2011; Croucher et al., 2011; Bagda et al., 2012; Durand et al., 2013; Bracken et al., 2015; Rašić et al., 2015 Rabone et al., 2015). In these reports, some prefer to use nuclear marker because it reveals higher genetic diversity and strong genetic structure (FitzSimmons et al., 1997; Canino et al., 2010; Durand et al., 2013; Bracken et al., 2015), whiles others held the opposite opinion (Schrey and Heist, 2003; Chappell et al., 2004; Ji et al., 2011; Croucher et al., 2011; Bagda et al., 2012; Rašić et al., 2015; Rabone et al., 2015). The present study confirmed high genetic diversity from nuclear genome and mitochondrial genome while indicating that both the nuclear and mitochondrial genome analyses could be effective for genetic studies. The combination of all genomic information with different genetic patterns might significantly improve future genetic studies.

Summary

Pine wilt disease (PWD) is one of the most serious forest problems in the world. It originated in North America and spread to many other countries in the 20th century. It has now become a worldwide threat to forest ecology and international trade. Understanding its transmission pathways and transmission mechanisms is believed to be effective in controlling the reproduction of pinewood nematode (PWN), especially in uninfected areas. DNA-based technique has been employed in several studies and molecular markers showing sufficient genetic polymorphism are wildly utilized in analysis of the PWN, while the reported DNA markers are limited in capturing only mutations at specific sites on the sequence resulting in a failure to comprehensively analyze variations, are only associated with the nuclear genome. However, genetic material is not only present in nuclei of eukaryotes, but also in mitochondria and chloroplasts. The development of next generation sequencing technologies has made it possible to obtain a large amount of sequence information to capture sufficient variations. Moreover, the reported entire nuclear genome sequence and entire mitochondria genome sequence make it possible to utilize SNP information in the PWN. Therefore, the object of the present study is to use SNP information derived from the nuclear and mitochondrial genomes of the PWN to elucidate the genetic structure of the nematode population and the phylogenetic relationships between the regional populations in Kyushu, the first invasion site in Asia.

In the first part of this study, genetic diversity and genetic structure of PWN in eight populations of Kyushu region were elucidated using the nucleotide polymorphism of ten EST loci. Considerable genetic differences among regional populations were observed with the gene differentiation (G_{ST}) of 0.53, which showed that more than half of the total gene diversity (H_T = 0.63) was possessed among regional populations. The H_T values of eight regional

populations were between 0.12 and 0.59. Sendai, Shintomi, Matsuura, and Karatsu populations were rich in gene diversity (0.59, 0.57, 0.56, and 0.55), and their high G_{ST} (0.43, 0.35, 0.25, and 0.25) indicated that the genetic compositions were notably different among the populations within damaged trees (subpopulations). On the contrary, in Amakusa and Miyazaki populations, extremely low gene diversity (0.12 and 0.18) and small G_{ST} (0.01 and 0.02) were confirmed, which showed that little genetic difference existed among subpopulations. It seemed that bottleneck effect or founder effect had a great impact on the formation of these regional populations. The genetic diversity of the regional populations was polarized in Kyushu.

In another study, mitogenomic diversity and genetic population structure of the PWN inhabiting Kyushu were analyzed. A method for performing long-PCR using single-nematodes and sequencing nematode mitochondrial genomes individually is presented here. About 8 kb (~55%) of the complete mitochondrial genome was successfully obtained from 285 individuals collected from 12 populations. The 158 single nucleotide polymorphisms detected corresponded to 30 haplotypes, clearly classified into two clades. Haplotype diversity was 0.83, evidencing a remarkable high diversity within Kyushu. The high genetic differentiation among the 12 populations (0.331) might be due to past invasion and expansion routes of PWN in Kyushu. The distinct genetic composition of populations within the northwestern, central western, and southwestern Kyushu seem to be mostly related to the extinction of pine forests and long-range migration of PWN due to human activity. Direct long-PCR and sequencing of single nematode individuals is an effective method for investigating mitochondrial polymorphisms, and these are an effective tool for PWN population genetics and other intraspecific studies.

The present study evaluated the genetic diversity from two cellular organelles which contain genetic material: the nucleus and mitochondria. The present study confirmed high

genetic diversity from nuclear genome and mitochondrial genome while indicating that both the nuclear and mitochondrial genome analyses could be effective for genetic studies. Moreover, the effectiveness of utilize the next generation sequencing to analyze the population genetic studies of the PWN was revealed in this study. On the other hand, this study also revealed differences in genetic diversity between the nuclear genome and the mitochondrial genome confirmed that the combination of all genomic information with different genetic patterns might significantly improve future genetic studies.

（和文要旨）

マツ材線虫病は世界で最も深刻な森林被害の一つである。北米に起源をもつこの病気は、20 世紀に他地域の多くの国々に広がり、現在では森林生態系や国際貿易に世界的脅威をおよぼしている。その感染経路や伝播メカニズムを解明することで、特に未感染地域でのマツノザイセンチュウ（以下、材線虫）の伝播抑制に効果的であると考えられる。DNA 技術は多くの研究で導入され、高い遺伝的多型を示す分子マーカーが材線虫で広く利用されてきた。一方、現在までに報告された DNA マーカーの多くは、配列上の特定部位の変異のみを捉えることに限定されており、変異の包括的な分析には至らず、核ゲノムの変異に限定されている。真核生物では、遺伝物質は細胞核に存在するだけでなく、ミトコンドリアや葉緑体にも存在する。次世代シークエンシング技術が進歩し、十分な変異を捉えるための大量の配列情報の取得が可能となった。さらに、核とミトコンドリアゲノムの全塩基配列が解読され、材線虫においても SNP（一塩基多型）情報の利用が容易になった。そこで、本研究では、材線虫の核およびミトコンドリアゲノムにある SNP 情報を使用し、アジアにおける最初の侵入地である九州に生息する材線虫集団の遺伝的構造と地域集団間の系統分類学的関係を解明した。

最初に、九州の 8 地域の材線虫集団の遺伝的多様性と遺伝的構造を 10 個の EST 遺伝子座の塩基配列多型を用いて解明した。九州全域の遺伝子分化係数（G_{ST}）は 0.53 で、全遺伝子多様度（H_T = 0.63）の半分以上が地域集団間に存在し、集団間に大きな差異があった。8 地域集団の H_T は 0.12〜0.59 であり、多様性に富んでいたのは、川内、新富、松浦、唐津（0.59, 0.57, 0.56, 0.55）で、地域集団内における G_{ST}（0.43, 0.35, 0.25, 0.25）も高く、被害木内集団（亜集団）間に大きな差異があった。一方、多様性が特に低いのは、天草、宮崎（0.12, 0.18）で、その G_{ST} も小さく（0.01, 0.02）、亜集団間の違いは極めて小さかった。これらの 2 集団の形成には、ボトルネックもしくは創始者効果が影響したことが示唆された。九州では、地域集団が保有する多様性の二極化が進行していた。

次に、九州に生息する材線虫集団のミトコンドリアゲノムにおける遺伝的多様性と遺伝的集団構造を分析した。まず、材線虫 1 個体を用いて直接ロング PCR し、ミトコンドリアゲノム配列を個別に取得する方法を開発した。12 の地域集団から集めた 285 個体から約 8 kbの塩基配列（全ミトコンドリアゲノム配列の 55%）を取得した。検出された 158 個の SNP により、30 種類のハプロタイプが確認され、これらは 2 つのクレードに明確に分類された。ハプロタイプの多様性は 0.83 であり、九州において極めて高い多様性があることが示された。 12 集団間の高い遺伝的分化係数（0.331）は、九州における過去の材線虫の侵入や拡大経路に起因していると推測された。九州の北西部、中西部、南西部の地域間に認められた遺伝的構成の明確な差異は、主にマツ林の消長と人間活動に伴う材線虫の長距離移動が関与したと考えられる。材線虫 1 個体をダイレクト・ロング PCR し、シークエンシングするこの分析系は、ミトコンドリア多型の検出に効果的な手法であるとともに、今後、材線虫の集団遺伝学や他の種の種内変異の研究における有効な手段と考える。

本研究では、核とミトコンドリアの 2 つの細胞小器官がもつ遺伝情報から遺伝的多様性を評価した。核とミトコンドリアの両ゲノムにおいて高い遺伝的多様性が確認され、両ゲノムを併用することが遺伝学的研究において有効であることを示した。さらに、材線虫の集団遺伝学的研究において次世代シークエンシング技術が有効であることを明らかにした。一方、本研究では、核ゲノムとミトコンドリアゲノムの遺伝的多様性にはゲノム間で違いのあることが示唆された。異なる遺伝様式をもつゲノムの遺伝情報を合わせて解析することにより、遺伝学的研究がさらに進展すると思われる。

Acknowledgments

On the accomplishment of this research, I would like to express my sincere gratitude to Emeritus Professor Susumu Shiraishi (Kyushu University) and Professor Shigejiro Yoshida (Kyushu University) who gave me guidance thoroughly and gently from Master course to Doctoral course for five years. Especially, I really would like express my gratitude to Professor Shiraishi for his support to get the Doctoral degree. He did not only teach me about knowledge but also teach me how to be a scientist and what the scientist spirit should be. It will be the most valuable asset in my research career. I also thank Professor Naruto Furuya (Kyushu University) for critical reading of my thesis.

Meanwhile, I want to give my deep thanks to Associate Professor Koichiro Gyokusen, Assistant Professor Eiji Gotoh, and Assisitant Professor Kotaro Sakuta, Faculty of Agriculture, Kyushu University，who gave me valuable advice on my Doctoral degree paper.

In addition, I would like to express my sincere gratitude to Erika Okii and Miho Yamashita for their cooperation during sampling pinewood nematodes and laboratory analyses. I thank Fumihiko Miyahara (Institute of Agricultural and Forest Resources, Fukuoka Agriculture and Forestry Research Ctr., Japan) and Koji Matsunaga (Kyushu Regional Breeding Office, Forest Tree Breeding Center, Forestry and Forest Products Research Institute, Japan) for providing the technology necessary to collect pinewood nematodes in the forest, and for isolating and rearing them in the laboratory. I also thank Junji Miyazaki (Saga Prefectural Forestry Exp. Stn., Japan), Hajime Maeda (Nagasaki Agricultural and Forestry Exp. Stn., Japan), Hideo Furusawa (Miyazaki Prefectural Forest Exp. Stn., Japan) and Manabu Miyazato (Kagoshima Prefecture Forest Exp. Stn., Japan) for assisting to collect the samples of PWNs.

Furthermore, I want to thank Kyushu University that gives me sufficient resource to

study and finish my research. I really thank Kyushu University to give me the chance to get the doctorate. Also, I would like to thank all of the professors in Kyushu University and all of the classmates in Silviculture laboratory who helped me smooth learning in Japan.

Finally, I would like to express my great gratitude to my parents who always supported me to complete my studies in Japan. They were patient with me to study abroad and let me have chance to understand the world far from my homeland. I am extremely grateful to them for supporting me to realize my ideal of life.

References

Abelleira, A., Picoaga, A., Mansilla, J. P., and Aguin, O. 2011. Detection of *Bursaphelenchus xylophilus*, causal agent of pine wilt disease on *Pinus pinaster* in North-western Spain. Plant Disease 39(6): 776.

Aggarwal, R. K., Hendre, P. S., Varshney, R. K., Bhat, P. R., Krishnakumar, V., and Singh, L. 2007. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. Theoretical and Applied Genetics 114(2): 359–372.

Aikawa, T., Kikuchi, T., and Kosaka, H. 2003. Demonstration of interbreeding between virulent and avirulent populations of *Bursaphelenchus xylophilus* (Nematoda: Aphelenchoididae) by PCR-RFLP method. Applied Entomology and Zoology 38(4): 565–569.

Al-Banna, L., Williamson, V. M., and Gardner, S. L., 1997. Phylogenetic analysis of nematodes of the genus *Pratylenchus* using nuclear 26S rDNA. Molecular Phylogenetics and Evolution 7(1): 94–102.

Bagda, E., Bardakci, F., and Turkozan, O. 2012. Lower genetic structuring in mitochondrial DNA than nuclear DNA among the nesting colonies of green turtles (*Chelonia mydas*) in the Mediterranean. Biochemical Systematics and Ecology 43: 192–199.

Beldade, P., Rudd, S., Gruber, J. D., and Long, A.D. 2006. A wing expressed sequence tag resource for *Bicyclus anynana* butterflies, an evo-devo model. BMC Genomics 7: 130.

Blaxter, M. L., De Ley, P., Garey, J. R., Liu, L. X., Scheldeman, P., Vierstraete, A., Vanfleteren, J. R., Mackey, L. Y., Dorris, M., Frisse, L. M., Vida, J. T., and Thomas, K. 1998 A molecular evolutionary framework for the phylum Nematoda. Nature 392(6671): 71–75.

Bonnell, M. L., and Selander, R. K. 1974. Elephant seals: Genetic variation and near extinction. Science 184(4193): 908–909.

Bouck, A., and Vision, T. 2007. The molecular ecologist's guide to expressed sequence tags. Molecular Ecology 16(50): 907–924.

Bracken, F. S. A., Hoelzel, R., Hume, J. B., and Lucas, M. C. 2015. Contrasting population genetic structure among freshwater-resident and anadromous lampreys: the role of demographic history, differential dispersal and anthropogenic barriers to movement. Molecular Ecology 24(6): 1188–1204.

Brooker, R., Widmaier, E. P., Graham, L., Stiling, P. 2017. Principles of Biology. McGraw-Hill Education.

Brookes, A. 1999. The essence of SNPs. Gene 234(2): 177–186.

Canino, M. F., Spies, I. B., and Lowe, S. A., and Grant, W. S. 2010. Highly discordant nuclear and mitochondrial DNA diversities in Atka Mackerel. Marine and Coastal Fisheries:

Dynamics, Management, and Ecosystem Science 2(1): 375–387.

Cardoso, J. M. S., Fonseca, L., and Abrantes, I. 2012. Direct molecular detection of the pinewood nematode, *Bursaphelenchus xylophilus*, from pine wood, bark and insect vector. European Journal of Plant Pathology 133(2): 419–425.

Chappell, D. E., Bussche, R., Krizan, J., and Patterson, B. 2004. Contrasting levels of genetic differentiation among populations of wolverines (*Gulo gul*o) from northern Canada revealed by nuclear and mitochondrial loci. Conservation Genetics 5(6): 759–767.

Chen, F. M., Negi, S., and Ye, J. R. 2011. A SCAR molecular marker to distinguish *Bursaphelenchus mucronatus* from the pinewood nematode, *B. xylophilus*. Forest Pathology 41(5): 376–381.

Cheng, H. R., Lin M. S., and Li, W. Q. 1983. The occurrence of a pine wilting disease caused by a nematode found in Nanjing. Forest Pest and Disease (4):1–5.

Cheng, X. Y., Cheng, F. X., Xu, R. M., and Xie, B. Y. 2008. Genetic variation in the invasive process of *Bursaphelenchus xylophilus* (Aphelenchida: Aphelenchoididae) and its possible spread routes in China. Heredity 100(4): 356–365.

Cheng, X. Y., Xie, P. Z., Cheng, F. X., Xu, R. M., and Xie, B. Y. 2009. Competitive displacement of the native species *Bursaphelenchus mucronatus* by an alien species *Bursaphelenchus xylophilus* (Nematoda: Aphelenchida: Aphelenchoididae): a case of successful invasion. Biological Invasions 11(2): 205–213.

Chio, I. H., Park, J. Y., Shin, S. C., and Park, I. K. 2006. Nematicidal activity of medicinal plant extracts and two cinnamates isolated from *Kaempferia galangal* L. (Proh Hom) against the pinewood nematode, *Bursaphelenchus xylophilus*. Nematology 8(3):359–365.

Croucher, P. J. P., Oxford, G. S., and Gillespie, R. G. 2011. Population structure and dispersal in a patchy landscape: nuclear and mitochondrial markers reveal area effects in the spider *Theridion californicum* (Araneae: Theridiidae). Biological Journal of the Linnean Society 104(3): 600–620.

Dayi, M., and Akbulut, S. 2012. Pathogenicity testing of four *Bursaphelenchus* species on conifer seedlings under greenhouse conditions. Forest Pathology 42(30): 213–219,.

De Ley, I. T., De Ley, P., Vierstraete, A., Karssen, G., Moens, M., and Vanfleteren, J. 2002. Phylogenetic analyses of *Meloidogyne* small subunit rDNA. Journal of Nematology 34(4): 319–327.

Durand, J. D., Blel, H., Shen, K. N., Koutrakis, E. T., and Guinand, B. 2013. Population genetic structure of *Mugil cephalus* in the Mediterranean and Black Seas: a single mitochondrial clade and many nuclear barriers. Marine Ecology Progress Series 474: 243–261.

Dwinell, L. D. 1993. First record of pinewood nematode (*Bursaphelenchus xylophilus*) in Mexico. Plant Disease 77(8): 846.

Dwinell, L. D. 1997. The pinewood nematode: Regulation and mitigation. Annual Review of

Phytopathology 35(1): 153–166.

EL Mokhefi, M., Kerdelhué, C., Burban, C., Battisti, A., Chakali, G., and Simonato, M. 2016. Genetic differentiation of the pine processionary moth at the southern edge of its range: contrasting patterns between mitochondrial and nuclear markers. Ecology and Evolution 6(13): 4274–4288.

Ellis, J. R., and Burke, J. M. 2007. EST-SSRs as a resource for population genetic analyses. Heredity 99: 125–132.

Figueiredo, J., Simões, M. J., Gomes, P., Barroso, C., Pinho, D., Conceição, L., Fonseca, L., Abrantes, I., Pinheiro, M., and Egas, C. 2013. Assessment of the geographic origins of pinewood nematode isolates via single nucleotide polymorphism in effector genes. PLoS ONE 8(12): e83542.

FitzSimmons, N. N., Moritz, C., Limpus, C. J., Pope, L., and Princef, R. 1997. Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian Green Turtle Populations and Male-Biased Gene Flow. Genetics 147(4): 1843–1854.

Fonseca, L., Cardoso, J. M. S., Lopes, A., Pestana, M., Abreu, F., Nunes, N., Mota, M., and Abrantes, I. 2012. The pinewood nematode, *Bursaphelenchus xylophilus*, in Madeira Island. Helminthologia 49(2): 96–103.

François, C., Castagnone, C., Boonham, N., Tomlinson, J., Lawson, R., Hockland, S., Quill, J., Vieira, P., Mota, M., and Castagnone-Sereno, P. 2007. Satellite DNA as a target for TaqMan real-time PCR detection of the pinewood nematode, *Bursaphelenchus xylophilus*. Molecular Plant Pathology 8(6): 803–809.

Frankham, R., Ballou, J. D., and Briscoe, D. A. 2002. Introduction to conservation genetics. Cambridge, USA: Cambridge University Press.

Futai, K. 2008. Pine wilt in Japan, from first incidence to the present. Pp. 5–6 in B. G. Zhao, K. Futai, J. R. Sutherland, and Y. Takeuchi, eds. Pine wilt disease. Tokyo, Japan: Springer.

Giblin-Davis, R. M., Davies, K. A., Taylor, G. S., and Thomas, W. K. 2004. Entomophilic nematode models for studying biodiversity and cospeciation. Pp. 493–540 in Z. Chen, S. Chen, D.W. Dickson, eds. Nematology: Advances and perspectives. Beijing, China: Tsinghua University Press. Publishing.

Gosset, W. S. 1908. The probable error of a mean. Biometrika 6(1): 1–25.

Gupta, P. K., Varshney, R. K., Sharma, P. C., and Ramesh, B. 1999. Molecular markers and their applications in wheat breeding. Plant Breeding 118(5): 369–390.

Han, Z. M., Hong, Y. D., and Zhao, B. G. 2003. A study on pathogenicity of bacteria carried by pinewood nematodes. Journal of Phytopathology 151(11–12): 683–689.

Hauge, X. Y., and Litt, M. 1993. A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. Human Molecular Genetics 2(4): 411–415.

Hu, S. J., Ning, T., Fu, D. Y., Haack, R. A., Zhang, Z., Chen, D. D., Ma, X. Y., and Ye, H. 2013. Dispersal of the Japanese pine sawyer, *Monochamus alternatus* (Coleoptera:

Cerambycidae), in mainland China as inferred from molecular data and associations to indices of human activity. PLoS ONE 8(2): e57568.

Hu, Y. Q., Kong, X. C., Wang, X. R., Zhong, T. K., Zhu, X. W., Mota, M. M., Ren, L. L., Liu, S., and Ma, C. 2011. Direct PCR-based method for detecting *Bursaphelenchus xylophilus*, the pine wood nematode in wood tissue of *Pinus massoniana*. Forest Pathology 4(2): 165–168.

Humphreys, M. W., Yadav, R. S., Cairns, A. J., Turner, L. B., Humphreys, J., and Skøt, L. 2005. A changing climate for grassland research. New Phytologist 169: 9–26.

Iwahori, H., and Futai, K. 1993. Lipid peroxidation and ion exudation of pine callus tissues inoculated with pinewood nematodes. Japanese Journal of Nematology 23(2): 79–89.

Iwahori, H., Kanzak, N., and Futai, K. 1998a. Phylogenetic relationship among several isolates of *Bursaphelenchus xylophilus* and *B. mucronatus* based on their ribosomal DNA sequences. Sustainability of Pine Forests in Relation to Pine Wilt and Decline: 35–38.

Iwahori, H., Kanzak, N., and Futai, K. 2000. A simple, polymerase chain reaction–restriction fragment length polymorphism-aided diagnosis method for pine wilt disease. Forest Pathology 30(3): 157–164.

Iwahori, H., Tsuda, K., Kanzaki, N., Izui, K., and Futai, K. 1998b. PCR-RFLP and sequencing analysis of ribosomal DNA of *Bursaphelenchus* nematodes related to pine wilt disease. Fundamental and Applied Nematology 21(6): 655–666.

Ji, Y. Q., Wu, D. D., Wu, G. S., Wang, G. D., and Zhang, Y. P. 2011. Multi-Locus analysis reveals a different pattern of genetic diversity for mitochondrial and nuclear DNA between wild and domestic pigs in east Asia. PLoS ONE 6(10): e26416.

Johnson, J. A., Toepfer, J. E., and Dunn, P. O. 2003. Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. Molecular Ecology 12(12): 3335–3347.

Jones, J. T., Moens, M., Mota, M., Li, H. M., and Kikuchi, T. 2008. *Bursaphelenchus xylophilus*: Opportunities in comparative genomics and molecular host-parasite interactions. Molecular Plant Pathology 9(3): 357–368.

Jung, J., Han, H., Ryu, S. H., and Kim, W. 2010a. Microsatellite variation in the pinewood nematode, *Bursaphelenchus xylophilus* (Steiner and Buhrer) Nickle in South Korea. Genes and Genomics 32(2): 152–158.

Jung, J., Han, H., Ryu, S. H., and Kim, W. 2010b. Amplified fragment length polymorphism analysis and genetic variation of the pinewood nematode *Bursaphelenchus xylophilus* in South Korea. Animal Cells and Systems 14(1): 31–36.

Kalia, R.K., Rai, M.K., and Kalia, S., Singh, R., Dhawan, A. K. 2011. Microsatellite markers: an overview of the recent progress in plants. Euphytica 177(3): 309–334.

Kanzaki, N., and Futai, K. 2002. A PCR primer set for determination of phylogenetic relationships of *Bursaphelenchus* species within the *xylophilus* group. Nematology: International Journal of Fundamental and Applied Nematological Research 4(1): 35–41.

Karp, A., Kresovich, S., Bhat, K. V., Ayad, W. G., and Hodgkin, T. 1997. Molecular tools in plant genetic resources conservation: A guide to the technologies. Roman: International Plant Genetic Resources Institute.

Kikuchi, T., Aikawa, T., Kosaka, H., Pritchard, L., Ogura, N., and Jones J. T. 2007. Expressed sequence tag (EST) analysis of the pine wood nematode *Bursaphelenchus xylophilus* and *B. mucronatus*. Molecular and Biochemical Parasitology 155(1): 9–17.

Kikuchi, T., Cotton, J. A., Dalzell, J. J., Hasegawa, K., Kanzaki, N. McVeigh, P., Takanashi, T., Tsai, I. J., Assefa, S. A., Cock, P. J. A., Otto, T. D., Hunt, M., Reid, A. J., Sanchez-Flores, A., Tsuchihara, K., Yokoi, T., Larsson, M. C., Miwa, J., Maule, A. G., Sahashi, N., Jones, J. T. and Berriman, M. 2011. Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. PLoS Pathogens 7(9): e1002219.

Kim, K. S,. Ratcliffe, S. T., French, B. W., Liu, L., and Sappington, T. W. 2008. Utility of EST-derived SSRs as population genetic markers in a beetle. Journal of Heredity 99(2): 112–124.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16(2): 111–120.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge, USA: Cambridge University Press.

Kishi, Y. 1988. Pine wilt disease (Matsu kui mushi), (translation by authors). Tokyo, Japan: Thomas Company, (In Japanese).

Kiyohara, T., and Tokushige, Y. 1971. Inoculation experiments of a nematode, *Bursaphelenchus* sp., onto pine trees. Journal of the Japanese Forest Society 53(7): 210–218, (In Japanese).

Knowles, K., Beaubien, Y., Wingfield, M. J., Bark, F. A., and French, D. W. 1983. The pinewood nematode new in Canada. The Forestry Chronicle 59: 40.

Kruskal, W. K., and Wallis, W. A. 1952. Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association 47(260): 583–621.

Kumar, S., Stecher, G., and Tamura, K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution 33(7): 1870–1874.

Leal, I., Foord, B., Allen, E., Campion, C., Rott, M., and Green, M. 2013. Development of two reverse transcription-PCR methods to detect living pinewood nematode, *Bursaphelenchus xylophilus*, in wood. Forest Pathology 43(2): 104–114.

Librado, P., and Rozas, J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452.

Linnertz, C., Saunders, A. M., Lutz, M. W., Crenshaw, D. M., Grossman, I., Burns, D. K., Whitfield, K. E., Hauser, M. A., McCarthy, J. J., Ulmer, M., Rand Allingham, R., Welsh-Bohmer, K. A., Roses, A. D., and Chiba-Falek, O. 2012. Characterization of the poly-T

variant in the TOMM40 gene in diverse populations. PLoS ONE 7(2): e30994.

Liu, J., Berry, R. E., and Moldenke, A.F. 1997. Phylogenetic relationships of entomopathogenic nematodes (Heterorhabditidae and Steinernematidae) inferred from partial 18S rRNA gene sequences. Journal of Invertebrate Pathology 69(3): 246–252.

Lowe, A., Harris, S., and Ashton, P. 2004. Ecological genetics: design, analysis, and application, UK: Blackwell Publishing.

Lu, W., Chen, X. X., Zheng, J. W., Wu, H., and Zhang, L. X. 2001. A comparative study on ITS region of *Bursaphelenchus xylophilus* and *B. mucronatus* by PCR-RFLP. Journal of Agricultural Biotechnology 9: 387–390.

Mackay, T. F. C. 2001. The genetic architecture of quantitative traits. Annual Review of Genetics 35: 303–339.

Madani, M., Subbotin, S. A., Ward, L. J., Li, X., and De Boer, S. H. 2010. Molecular characterization of Canadian populations of potato cyst nematodes, *Globodera rostochiensis* and *G. pallida* using ribosomal nuclear RNA and cytochrome *b* genes. Canadian Journal of Plant Pathology 32(2): 252–263.

Mallez, S., Castagnone, C., Espada, M., Vieira, P., Eisenback, J. D., Mota, M., Guillemaud, T., and Castagnone-Sereno, P. 2013. First insights into the genetic diversity of the pinewood nematode in its native area using new polymorphic microsatellite loci. PLoS ONE 8(3): e59165.

Mamiya, Y. 1975. Efficiency of Baermann funnel method for the extraction of pinewood nematodes. Forest Pest 24: 115–119.

Mamiya, Y. 1984. The pine wood nematode. Pp. 589–626 in W. R. Nickle, ed. Plant and insect nematodes. New York, USA: Marcel Dekker.

Mamiya, Y., and Enda, N. 1979. *Bursaphelenchus mucronatus* n. sp. (Nematoda: Aphelenchoididae) from pine wood and its biology and pathogenicity to pine trees. Nematologica 25(3): 353–361.

Mamiya, Y., and Shoji, T. 2009. Pathogenicity of the pinewood nematode, *Bursaphelenchus xylophilus*, to Japanese larch, *Larix kaempferi*, seedlings. Journal of Nematology 41(2): 157–162.

Mayr, E. 1954. Change of genetic environment and evolution. Pp. 157–180 in J. Huxley, A. C. Hardy, and E. B. Ford, eds. Evolution as a process. London, UK: Allen and Unwin.

Metge, K., and Bürgermeister, W. 2006. Intraspecific variation in isolates of *Bursaphelenchus xylophilus* (Nematoda: Aphelenchoididae) revealed by ISSR and RAPD fingerprints. Journal of Plant Diseases and Protection 113(6): 275–282.

Miah, G., Rafii, M. Y., Ismail, M. R., Puteh, A. B., Rahim, A. H., Islam, K. N., and Latif, M. A. 2013. A review of microsatellite markers and their applications in rice breeding programs to improve blast disease resistance. International Journal of Molecular Sciences 14(11): 22499–22528.

Miko, I., LeJeune, L., Chial, H., Drovdlic, C., Koopman, M., Nelson, S. C., Spivey, A., and Smith, R. 2009. The genetic variation in a population is caused by multiple factors. Pp. 125 in I. Miko. and I. LeJeune, eds. Essentials of genetics. Cambridge, USA: NPG Education.

Moreira, C., Asch, B., Fonseca, L., Pereira-Castro, I., Silva, R., Azevedo, L., Mota, M., Abrantes, I., Amorim, A and Pereira, F. 2014. The mitochondrial genome of the pinewood nematode (*Bursaphelenchus xylophilus*) lineage introduced in Europe. Mitochondrial DNA 25(6): 420–421.

Mota, M. M., Braasch, H., Bravo, M. A., Penas, A. C., Burgermeister, W., Metge, K., and Sousa, E. 1999. First report of *Bursaphelenchus xylophilus* in Portugal and in Europe. Nematology: International Journal of Fundamental and Applied Nematological Research 1(7): 727–734.

Mouhaddab, J., Ait Aabd, N., Achtak, H., Msanda, F., Zahidi, A., Filali-Maltouf, A, Ferradouss, A., EL Modafar, C., Nejmeddine, M., and EL Mousadik, A. 2015. Patterns of genetic diversity and structure at fine scale of an endangered Moroccan endemic tree (*Argania spinosa* L. Skeels) based on ISSR polymorphism. Notulae Botanicae Horti Agrobotanici Cluj-Napoca 43(2): 528–535.

Murray, V., Monchawin, C., and England, P. R. 1993. The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. Nucleic Acids Research 21(10): 2395–2398.

Nei, M. 1972. Genetic distance between populations. The American Naturalist 106(949): 283–292.

Nei, M. 1987. Molecular evolutionary genetics. New York, USA: Columbia University Press.

Nei, M., and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97(1): 145–163.

Nose, M., Shiraishi, S., Miyahara, F., Ohira, M., Matsunaga, K., Tobase, M., Koyama, T., and Yoshimoto, K. 2009. Ribosomal DNA haplotype distribution of *Bursaphelenchus xylophilus* in Kyushu and Okinawa islands, Japan. Journal of Nematology 41(3): 194–202.

Park, Y. H., Alabady, M. S., Ulloa, M., Sickler, B., Wilkins, T. A., Yu, J., Stelly, D. M., Kohel, R. J., El-Shihy, O. M., and Cantrell, R. G. 2005. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. Genomics 274(4): 428–441.

Pereira, F., Moreira, C., Fonseca, L., Van Asch, B., Mota, M., Abrantes, I., and Amorim, A. 2013. New insights into the phylogeny and worldwide dispersion of two closely related nematode species, *Bursaphelenchus xylophilus* and *Bursaphelenchus mucronatus*. PLoS ONE 8(2): e56288.

Perera, L., Russell, J. R., Provan, J., and Powell, W. 2000. Use of microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (Cocos nucifera L.). Genome 43(1): 15–21.

Picard, D., Sempere, T., and Plantard, O. 2007. A northward colonisation of the Andes by the

potato cyst nematode during geological times suggests multiple host-shifts from wild to cultivated potatoes. Molecular Phylogenetics and Evolution 42(2): 308–316.

Plantard, O., Picard, D., Valette, S., Scurrah, M., Grenier, E., and Mugniéry D. 2008. Origin and genetic diversity of western European populations of the potato cyst nematode (*Globodera pallida*) inferred from mitochondrial sequences and microsatellite loci. Molecular Ecology 17(9): 2208–2218.

Powell, W., Machray, G. C., and Provan, J. 1996. Polymorphism revealed by simple sequence repeats. Trends in Plant Science 1(7): 215–222.

Proença, D. N., Francisco. R., Santos, C. V., Lopes, A., Fonseca, L., Abrantes, I. M. O., and Morais, P. V. 2010. Diversity of bacteria associated with *Bursaphelenchus xylophilus* and other nematodes isolated from *Pinus pinaster* trees with pine wilt disease. PLoS ONE 5(12): e15191.

Rabone, M., Lavery, S. D., Little, A., and Clements, K. D. 2015. Discordance between nuclear and mitochondrial DNA analyses of population structure in closely related triplefin fishes (*Forsterygion lapillum* and *F. capito*, F. Tripterygiidae) supports speciation with gene flow. Marine Biology 162(8): 1611–1624.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. Current opinion in plant biology 5(2): 94–100.

Rakoczy-Trojanowska, M., and Bolibok, H. 2004. Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. Cellular and Molecular Biology Letters 9(2): 221–238.

Rašić, G., Schama, R., Powell, R., Maciel-de Freitas, R., Endersby-Harshman, N. M., Filipović, I., Sylvestre, G., Máspero, R. C., and Hoffmann, A. A. 2015. Contrasting genetic structure between mitochondrial and nuclear markers in the dengue fever mosquito from Rio de Janeiro: implications for vector control. Evolutionary Applications 8(9): 915–950.

Robertson, L., Arcos, S. C., Escuer, M., Merino, R. S., Esparrago, G., Abellera, A., and Navas, A. 2011. Incidence of the pinewood nematode *Bursaphelenchus xylophilus* Steiner & Buhrer, 1934 (Nickle, 1970) in Spain. Nematology: International Journal of Fundamental and Applied Nematological Research 13(6): 755–757.

Rodrigues, J. M. 2008. National eradication programme for the pinewood nematode. Pp. 5–44 in M. M. Mota and P. Vieira, eds. Pine wilt disease: A worldwide threat to forest ecosystems. Heidelberg, Germany: Springer.

Rungis, D., Hamberger, B., Bérubé, Y., Wilkin, J., Bohlmann, J., and Ritland, K. 2005. Efficient genetic mapping of single nucleotide polymorphisms based upon DNA mismatch digestion. Molecular Breeding 16(3): 261–270.

Saha, M. C., Mian, M. A. R., Eujayl, I., Zwonitzer, J. C., Wang, L., and May, G. D. 2004. Tall fescue EST–SSR markers with transferability across several grass species. Theoretical and Applied Genetics 109(4): 783–791.

Saint-Laurent, R., Legault, M., and Bernatchez, L. 2003. Divergent selection maintains

adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). Molecular Ecology 12(2): 315–330.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4(4): 406–425.

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., Ellstrand, N. C., McCauley, D. E., O'Neil, P., Parker, I. M., Thompson, J. N. and Weller, S. G. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics 32: 305–332.

Schlötterer, C. 2004. The evolution of molecular markers—just a matter of fashion? Nature Reviews Genetics 5: 63–69.

Schrader, G., and Unger, J. G. 2003. Plant quarantine as a measure against invasive alien species: the framework of the international plant protection convention and the plant health regulations in the European Union. Biological Invasions 5(4): 357–364.

Schrey, A. W., and Heist, E. J. 2003. Microsatellite analysis of population structure in the shortfin mako (*Isurus oxyrinchus*). Canadian Journal of Fisheries and Aquatic Sciences 60(6): 670–675.

Schwartz, M. K., Mills, L. S., Ortega, Y., Ruggiero, L. F., and Allendorf, F. W. 2003. Landscape location affects genetic variation of Canada lynx (*Lynx canadensis*). Molecular Ecology 12(7): 1807–1816.

Scott, K. D., Eggler, P., Seaton, G., Rossetto, M., Ablett, E. M., Lee, L. S., and Henry, R. J. 2000. Analysis of SSRs derived from grape ESTs. Theoretical and Applied Genetics 100(5): 723–726.

Semagn, K., Babu, R., Hearne, S., and Olsen, M. 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. Molecular Breeding 33(1): 1–14.

Shan, X., Blake, T. K., and Talbert, L. E. 1999. Conversion of AFLP markers to sequence-specific PCR markers in barley and wheat. Theoretical and Applied Genetics 98(6–7): 1072–1078.

Shariflou, M. R., Hassani, M. E., and Sharp, P. J. 2001. A PCR-based DNA marker for detection of mutant and normal alleles of the *Wx-D1* gene of wheat. Plant Breeding 120(2): 121–124.

Shi, J., Chen, F., Luo, Y. Q., Wang, Z., and Xie, B. Y. 2013. First isolation of pine wood nematode from *Pinus tabuliformis* forests in China. Forest Pathology 43(1): 59–66.

Sim, S. C., Yu, J. K., Jo, Y. K., Sorrells, M. E., and Jung, G. 2009. Transferability of cereal EST-SSR markers to ryegrass. Genome 52(5): 431–437.

Slate, J., Hale, M. C., and Birkhead, T. R. 2007. Simple sequence repeats in zebra finch (*Taeniopygia guttata*) expressed sequence tags: a new resource for evolutionary genetic studies of passerines. BMC Genomics 8: 52.

Smith, P. T., Kambhampati S., and Aarmstrong, K. 2003. Phylogenetic relationships among *Bactrocera* species (Diptera: Tephritidae) inferred from mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 26(1): 8–17.

Staub, J. E., Serquen, F. C., and Gupta, M. 1996. Genetic markers, map construction, and their application in plant breeding. HortScience 31(5): 729–741.

Steiner, G., and Buhrer, E. M. 1934. Aphelenchoides *xylophilus* n. sp., a nematode associated with blue-stain and other fungi in timber. Journal of Agricultural Research 48(1): 949–951.

Storz, J. F. 2002. Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. Molecular Ecology 11(12): 2537–2551.

Sultana, T., Han, H., and Park, J. K. 2013. Comparison of complete mitochondrial genomes of pine wilt nematode *Bursaphelenchus xylophilus* and *Bursaphelenchus mucronatus* (Nematoda: Aphelenchoidea) and development of a molecular tool for species identification. Gene 520(1): 39–46.

Sunnucks, P. 2000. Efficient genetic markers for population biology. Trends in Ecology and Evolution 15(5): 199–203.

Suzuki, K. 2002. Pine wilt disease - a threat to pine forest in Europe. Dendrobiology 48: 71–74.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Takemoto, S., and Futai, K. 2007. Polymorphism of Japanese isolates of the pinewood nematode, *Bursaphelenchus xylophilus* (Aphelenchida: Aphelenchoididae), at heat-shock protein 70A locus and the field detection of polymorphic populations. Applied Entomology and Zoology 42(2): 247–253.

Tautz, D., and Renz, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Research 12(10): 4127–4138.

Tehrani, M. S., Mardi, M., Sahebi, J., Catalán, P., and Díaz-Pérez, A. 2009. Genetic diversity and structure among Iranian tall fescue populations based on genomic-SSR and EST-SSR marker analysis. Plant Systematics and Evolution 282(1–2): 57–70.

The Japan Nematology Society. 2004. Experimental method of nematode, (translation by authors). Japanese academy of nematodes, (In Japansese).

Thomas, W. K., and Wilson, A. C. 1991. Mode and tempo of molecular evolution in the nematode Caenorhabditis: cytochrome oxidase II and calmodulin sequences. Genetics 128(2): 269–279.

Tsutsui, N. D., Suarez, A. V., Holway, D. A. and Case, T. J. 2000. Reduced genetic variation and the success of an invasive species. Proceeding of the National Academy Science of the United States of America 97(11): 5948–5953.

Underwood, A. P., and Bianco, A. E. 1999. Identification of a molecular marker for the Y chromosome of *Brugia malayi*. Molecular and Biochemical Parasitology 99(1): 1–10.

Valadas, V., Laranjo, M., Barbosa, P., Espada, M., Mota, M., and Oliveira, S. 2012. The pine wood nematode, *Bursaphelenchus xylophilus*, in Portugal: possible introductions and spread routes of a serious biological invasion revealed by molecular methods. Nematology 14(8): 899–911.

Valadas, V., Laranjo, M., Mota, M., and Oliveira, S. 2013. Molecular characterization of Portuguese populations of the pinewood nematode *Bursaphelenchus xylophilus* using cytochrome *b* and cellulase genes. Journal of Helminthology 87(4): 457–466.

Van der Veer, M., and de Vries, E. 2004. A single nucleotide polymorphism map of the mitochondrial genome of the parasitic nematode *Cooperia oncophora*. Parasitology 128(4): 421–431.

Viard, F., El-Kassaby, Y. A., and Ritland, K. 2001. Diversity and genetic structure in populations of *Pseudotsuga menziesii* (Pinaceae) at chloroplast microsatellite loci. Genome 44(3): 336–344.

Vieira, P., Burgermeister, W., Mota, M., Metge, K., and Silva, G. 2007. Lack of genetic variation of *Bursaphelenchus xylophilus* in Portugal revealed by RAPD-PCR analyses. Journal of Nematology 39(2): 118–126.

Vieira, P., Castagnone, C., Mallez, S., Espada, M., Navas, A., Mota, M., and Castagnone-Sereno, P. 2014. Sequence variability of the *Msp*I satellite DNA family of the pinewood nematode *Bursaphelenchus xylophilus* at different geographic scales. Molecular Phylogenetics and Evolution 70: 120–129.

Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. 2002. A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution 34 (3): 275–305.

Wang, H. Y., Wei, Y. M., Yan, Z. H., and Zheng, Y. L, 2007. EST-SSR DNA polymorphism in durum wheat (*Triticum durum* L.) collections. Journal of Applied Genetics 48(1): 35–42.

Wang, M. L., Mosjidis, J. A., Morris, J. B., Dean, R. E., Jenkins, T. M., and Pederson, G. A. 2006. Genetic diversity of *Crotalaria* germplasm assessed through phylogenetic analysis of EST-SSR markers. Genome 49(6): 707–715.

Wang, X. R., Kong, X. C., Jia, W. H., Zhu, X. W., Ren, L. L., and Mota, M. M. 2010. A rapid staining-assisted wood sampling method for PCR-based detection of pine wood nematode *Bursaphelenchus xylophilus* in *Pinus massoniana* wood tissue. Forest Pathology 40(6): 510–520.

Williams, B. L., Brawn, J. D., and Paige, K. N. 2003. Landscape scale genetic effects of habitat fragmentation on a high gene flow species: *Speyeria idalia* (Nymphalidae). Molecular Ecology 12(1): 11–20.

Wolfe, A. D., and Liston, A. 1998. Contribution of PCR-based methods to plant systematics and evolutionary biology. Molecular Systematics of Plants II: 43–86.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenetics 15(4): 323–354.

Wu, J. 2004. Forest invasive species and management in China. Science and Technology Review 4: 41–44.

Wu, X. Q., Yuan, W. M., Tian, X. J., Fan, B., Fang, X., Ye, J. R., and Ding, X. L. 2013. Specific and functional diversity of endophytic bacteriafrom pine wood nematode *Bursaphelenchus xylophilus* with different virulence. International Journal of Biological Sciences 9(1): 34–44.

Yang, B. J. 1995. The pine wood nematode in China. Pp. 207–210 in B. J. Yang, ed. International symposium on pine wilt disease caused by pine wood nematode. Beijing: Chinese Society of Forestry.

Yang, B. J., and Wang, Q. L. 1988. Distribution of *Bursaphelenchus xylophilus* in our country. Research of Forestry Sciences 1: 450–451.

Yang, B. J., Pan, H. Y., Tang, J., Wang, Y. Y., Wang, L. F., and Wang, Q. 2003. Vectors of pine wood nematode, Pp. 61–62 in: B. J. Yang, H. Y. Pan, J. Tang, Y. Y. Wang, L. F. Wang, eds. *Bursaphelenchus xylophilus*. Beijing, China: Chinese Forestry Press.

Yano, S. 1913. Investigation on the cause of pine mortality in Nagasaki prefecture. Sanrinkoho 4(Suppl): 1–14.

Ye, W. M., GiblinDavis, R. M., Braasch, H., Morris, K., and Thomas, W. K. 2007. Phylogenetic relationships among *Bursaphelenchus* species (Nematoda, Parasitaphelenchidae) inferred from nuclear ribosomal and mitochondrial DNA sequence data. Molecular Phylogenetics and Evolution 43(3): 1185–1197.

Yi, C. K., Byun, B. H., Park, J. D., Yang, S. I., and Chang, K. H. 1989. First finding of the pinewood nematode, *Bursaphelenchus xylophilus* (Steiner & Buhrer) Nickle and its insect vector in Korea. Seoul: Research Reports of the Forestry Research Institute 38: 141–149.

Yun, J. E., Kim, J., and Park, C. G. 2012. Rapid diagnosis of the infection of pine tree with pine wood nematode (*Bursaphelenchus xylophilus*) by use of host-tree volatiles. Journal of Agricultural and Food Chemistry 60(30): 7392–7397.

Zhang, K. Y., Liu, H., Sun, J., Liu, J. R., Fei, K., Zhang, C. X., Xu, M. X., Sun, J., Ma, X. Y., Lai, R., Wu, Y. D., and Lin, M. S. 2008. Molecular phylogeny of geographical isolates of *Bursaphelenchus xylophilus*: Implications on the origin and spread of this species in China and worldwide. Journal of Nematology 40(2): 127–137.

Zhang, X. Y., and Luo, Y. Q. 2003. Major forest diseases and insect pests in China. Beijing, China: Chinese Forestry Publishing.

Zhao, B. G. 2008. Pine wilt disease in China. Pp. 18–25 in B. G. Zhao, K. Futai, J. R. Sutherland, and Y. Takeuchi, eds. Pine wilt disease. Tokyo, Japan: Springer.

Zhao, B. G., Wang, H. L., Han, S. F., and Han, Z. M. 2003. Distribution and pathogenicity of bacteria species carried by *Bursaphelenchus xylophilus* in China. Nematology 5(6): 899–906.

Zhao, L. L., Jiang, P., Humble, L. M., and Sun, J. H. 2009. Within-tree distribution and attractant sampling of propagative pinewood nematode, *Bursaphelenchus xylophilus*: An early

diagnosis approach. Forest Ecology and Management 258(9): 1932–1937.

Zhao, L. L., Mota, M., Vieira, P., Butcher, R. A., and Sun, J. H. 2014. Interspecific communication between pinewood nematode, its insect vector, and associated microbes. Trends in Parasitology 30(6): 299–308.

Zhou, Z. H., Sakaue, D., Wu, B. Y., and Hogetsu, T. 2007. Genetic structure of populations of the pinewood nematode *Bursaphelenchus xylophilus*, the pathogen of pine wilt disease, between and within pine forests. Phytopathology 97(3): 304–310.

Zhu, L. H., Ye, J. R., Negi, S., Xu, X. L., Wang, Z. L., and Ji, J. Y. 2012. Pathogenicity of aseptic *Bursaphelenchus xylophilus*. PLoS ONE 7(5): e38095.

Zhuo, K., Luo, M., Cui, R. Q., and Liao, J. L. 2011. A multiplex one-step PCR method for the simultaneous identification of *Bursaphelenchus xylophilus*, *B. mucronatus* and *B. doui* – three species within the *xylophilus* group. Forest Pathology 41(1): 66–69.

Appendices

Table S2-1. Heterozygosity of each subpopulation in ten polymorphic loci of pinewood nematode

Subpopulation		Locus										Average
		pX021	pX043	pX047	pX070	pX079	pX130	pX143	pX144	pX155	pX194	
Ka	Ka-01	0.483	0.601	0.395	0.440	0.675	0.475	0.611	0.621	0.489	0.619	0.541
	Ka-02	0.361	0.657	0.604	0.626	0.576	0.538	0.422	0.461	0.450	0.481	0.518
	Ka-03	0.522	0.572	0.017	0.000	0.000	0.095	0.464	0.576	0.410	0.642	0.330
	Ka-04	0.260	0.378	0.256	0.163	0.731	0.168	0.380	0.559	0.702	0.338	0.393
	Ka-05	0.231	0.281	0.085	0.047	0.717	0.195	0.280	0.566	0.690	0.330	0.342
	Ka-06	0.192	0.253	0.129	0.113	0.724	0.192	0.365	0.595	0.695	0.357	0.362
Ma	Ma-01	0.324	0.318	0.073	0.186	0.000	0.149	0.520	0.576	0.248	0.465	0.286
	Ma-02	0.210	0.203	0.173	0.151	0.025	0.149	0.553	0.465	0.382	0.541	0.285
	Ma-03	0.306	0.142	0.155	0.140	0.000	0.230	0.471	0.576	0.453	0.475	0.295
	Ma-04	0.277	0.256	0.168	0.156	0.000	0.452	0.111	0.068	0.068	0.175	0.173
	Ma-05	0.217	0.110	0.176	0.125	0.000	0.425	0.149	0.129	0.000	0.173	0.150
	Ma-06	0.439	0.722	0.498	0.698	0.775	0.750	0.549	0.657	0.635	0.574	0.630
	Ma-07	0.416	0.730	0.554	0.783	0.771	0.726	0.582	0.745	0.609	0.687	0.660
	Ma-08	0.471	0.725	0.478	0.722	0.740	0.768	0.425	0.661	0.728	0.667	0.639
	Ma-09	0.426	0.715	0.467	0.714	0.758	0.743	0.499	0.660	0.620	0.591	0.619
	Ma-10	0.461	0.770	0.465	0.700	0.766	0.734	0.519	0.693	0.687	0.626	0.642
	Ma-11	0.249	0.597	0.394	0.523	0.717	0.488	0.456	0.432	0.739	0.535	0.513
	Ma-12	0.189	0.286	0.069	0.168	0.000	0.219	0.385	0.559	0.412	0.472	0.276
	Ma-13	0.485	0.669	0.442	0.723	0.738	0.725	0.509	0.560	0.624	0.603	0.608
	Ma-14	0.198	0.154	0.147	0.205	0.020	0.178	0.544	0.481	0.600	0.410	0.294
	Ma-15	0.203	0.191	0.075	0.000	0.000	0.126	0.356	0.479	0.400	0.584	0.241
Am	Am-01	0.279	0.279	0.031	0.000	0.000	0.199	0.055	0.169	0.053	0.127	0.119
	Am-02	0.295	0.269	0.000	0.000	0.023	0.166	0.142	0.192	0.070	0.216	0.137
	Am-03	0.302	0.218	0.000	0.000	0.045	0.191	0.170	0.142	0.058	0.175	0.130
	Am-04	0.177	0.224	0.000	0.000	0.000	0.145	0.094	0.104	0.053	0.117	0.091
	Am-05	0.194	0.260	0.000	0.000	0.023	0.169	0.160	0.185	0.048	0.124	0.116
	Am-06	0.206	0.286	0.000	0.000	0.047	0.207	0.213	0.186	0.099	0.174	0.142
Sh	Sh-01	0.203	0.275	0.272	0.435	0.712	0.702	0.421	0.542	0.593	0.516	0.467
	Sh-02	0.321	0.172	0.379	0.363	0.717	0.716	0.384	0.550	0.682	0.383	0.467
	Sh-03	0.173	0.286	0.237	0.197	0.415	0.336	0.313	0.251	0.585	0.359	0.315
	Sh-04	0.246	0.112	0.251	0.530	0.116	0.516	0.268	0.132	0.606	0.171	0.295
	Sh-05	0.180	0.640	0.390	0.551	0.624	0.553	0.426	0.377	0.706	0.459	0.491
	Sh-06	0.393	0.493	0.250	0.575	0.524	0.261	0.210	0.245	0.658	0.332	0.394
	Sh-07	0.107	0.204	0.110	0.105	0.158	0.200	0.165	0.135	0.124	0.084	0.139
	Sh-08	0.291	0.412	0.130	0.470	0.099	0.479	0.135	0.178	0.663	0.215	0.307
	Sh-09	0.224	0.551	0.480	0.431	0.600	0.602	0.434	0.399	0.700	0.514	0.493
	Sh-10	0.649	0.633	0.429	0.648	0.560	0.074	0.054	0.057	0.329	0.136	0.357
	Sh-11	0.571	0.525	0.476	0.653	0.748	0.808	0.444	0.701	0.592	0.607	0.612
	Sh-12	0.198	0.218	0.192	0.093	0.218	0.161	0.249	0.195	0.254	0.237	0.202
	Sh-13	0.328	0.306	0.088	0.110	0.030	0.151	0.129	0.143	0.504	0.137	0.193
	Sh-14	0.249	0.497	0.367	0.427	0.590	0.702	0.322	0.492	0.591	0.364	0.460
Mi	Mi-01	0.278	0.180	0.251	0.052	0.000	0.228	0.099	0.196	0.107	0.360	0.175
	Mi-02	0.282	0.237	0.081	0.150	0.000	0.293	0.144	0.080	0.095	0.194	0.155
	Mi-03	0.198	0.178	0.088	0.101	0.087	0.160	0.155	0.187	0.386	0.202	0.174
	Mi-04	0.204	0.163	0.148	0.167	0.019	0.241	0.136	0.163	0.329	0.142	0.171
	Mi-05	0.218	0.281	0.142	0.197	0.000	0.122	0.156	0.188	0.247	0.250	0.180
	Mi-06	0.175	0.235	0.205	0.197	0.000	0.122	0.160	0.087	0.380	0.108	0.167
	Mi-07	0.259	0.371	0.083	0.161	0.000	0.273	0.183	0.121	0.217	0.289	0.196
	Mi-08	0.323	0.268	0.142	0.107	0.000	0.196	0.214	0.244	0.333	0.136	0.196
	Mi-09	0.292	0.214	0.094	0.189	0.031	0.236	0.199	0.137	0.428	0.202	0.202

Table S2-1 (continued)

Subpopulation		Locus										Average
		pX021	pX043	pX047	pX070	pX079	pX130	pX143	pX144	pX155	pX194	
Mi	Mi-10	0.194	0.190	0.110	0.110	0.000	0.269	0.166	0.105	0.144	0.170	0.146
	Mi-11	0.168	0.163	0.126	0.176	0.012	0.175	0.153	0.063	0.374	0.145	0.156
	Mi-12	0.269	0.178	0.150	0.154	0.000	0.233	0.211	0.147	0.124	0.124	0.159
	Mi-13	0.159	0.189	0.124	0.086	0.032	0.197	0.129	0.118	0.284	0.022	0.134
	Mi-14	0.217	0.237	0.108	0.213	0.022	0.259	0.256	0.195	0.496	0.117	0.212
	Mi-15	0.300	0.351	0.117	0.148	0.036	0.105	0.096	0.121	0.415	0.251	0.194
	Mi-16	0.254	0.213	0.055	0.109	0.000	0.223	0.211	0.138	0.393	0.111	0.171
	Mi-17	0.187	0.274	0.103	0.233	0.000	0.171	0.178	0.166	0.420	0.100	0.183
	Mi-18	0.244	0.114	0.170	0.156	0.014	0.137	0.172	0.089	0.102	0.247	0.145
Ni	Ni-01	0.246	0.295	0.102	0.350	0.011	0.478	0.249	0.163	0.615	0.076	0.259
	Ni-02	0.212	0.234	0.100	0.483	0.065	0.425	0.074	0.152	0.618	0.196	0.256
	Ni-03	0.181	0.390	0.120	0.350	0.055	0.498	0.189	0.168	0.619	0.238	0.281
	Ni-04	0.253	0.194	0.118	0.338	0.012	0.404	0.130	0.096	0.624	0.134	0.230
	Ni-05	0.222	0.571	0.335	0.485	0.596	0.605	0.362	0.375	0.720	0.389	0.466
	Ni-06	0.385	0.338	0.179	0.170	0.207	0.222	0.325	0.322	0.530	0.244	0.292
	Ni-07	0.410	0.494	0.414	0.377	0.464	0.423	0.355	0.348	0.528	0.438	0.425
	Ni-08	0.541	0.546	0.518	0.701	0.713	0.473	0.261	0.506	0.542	0.232	0.503
	Ni-09	0.273	0.191	0.538	0.239	0.819	0.595	0.557	0.652	0.735	0.596	0.519
	Ni-10	0.232	0.198	0.537	0.366	0.862	0.659	0.524	0.646	0.797	0.632	0.545
	Ni-11	0.205	0.228	0.220	0.115	0.017	0.214	0.044	0.154	0.363	0.227	0.179
	Ni-12	0.152	0.206	0.023	0.040	0.019	0.233	0.153	0.064	0.174	0.092	0.116
	Ni-13	0.181	0.263	0.154	0.208	0.000	0.249	0.094	0.125	0.155	0.182	0.161
	Ni-14	0.243	0.250	0.142	0.258	0.040	0.080	0.128	0.183	0.441	0.152	0.192
	Ni-15	0.282	0.137	0.120	0.088	0.100	0.142	0.152	0.114	0.197	0.078	0.141
	Ni-16	0.255	0.231	0.206	0.070	0.000	0.200	0.058	0.182	0.197	0.157	0.156
	Ni-17	0.343	0.222	0.101	0.072	0.000	0.221	0.186	0.112	0.154	0.129	0.154
Se	Se-01	0.334	0.345	0.261	0.446	0.476	0.537	0.278	0.318	0.281	0.435	0.371
	Se-02	0.200	0.668	0.642	0.726	0.767	0.486	0.270	0.593	0.196	0.654	0.520
	Se-03	0.556	0.689	0.485	0.546	0.764	0.526	0.466	0.539	0.031	0.583	0.519
	Se-04	0.595	0.694	0.494	0.417	0.681	0.528	0.453	0.546	0.023	0.635	0.507
	Se-05	0.017	0.127	0.261	0.039	0.332	0.013	0.084	0.299	0.132	0.225	0.153
	Se-06	0.000	0.000	0.047	0.065	0.069	0.011	0.024	0.068	0.088	0.196	0.057
	Se-07	0.000	0.658	0.467	0.198	0.702	0.000	0.081	0.597	0.075	0.590	0.337
	Se-08	0.374	0.272	0.338	0.464	0.445	0.530	0.165	0.344	0.289	0.519	0.374
	Se-09	0.417	0.333	0.246	0.458	0.374	0.566	0.314	0.329	0.233	0.373	0.364
	Se-10	0.000	0.664	0.000	0.504	0.513	0.000	0.049	0.528	0.440	0.626	0.332
	Se-11	0.081	0.664	0.266	0.675	0.625	0.153	0.177	0.606	0.589	0.532	0.437
	Se-12	0.000	0.652	0.000	0.518	0.540	0.108	0.048	0.545	0.552	0.631	0.359
	Se-13	0.196	0.110	0.101	0.205	0.053	0.205	0.200	0.093	0.474	0.125	0.176
	Se-14	0.336	0.383	0.151	0.144	0.137	0.190	0.336	0.325	0.458	0.234	0.270
	Se-15	0.036	0.052	0.286	0.734	0.779	0.342	0.059	0.267	0.040	0.327	0.292
Ib	Ib-01	0.648	0.071	0.026	0.051	0.034	0.257	0.340	0.313	0.281	0.301	0.232
	Ib-02	0.705	0.063	0.052	0.034	0.040	0.121	0.336	0.278	0.080	0.345	0.206
	Ib-03	0.567	0.176	0.105	0.128	0.121	0.182	0.200	0.299	0.074	0.278	0.213
	Ib-04	0.238	0.070	0.083	0.048	0.057	0.118	0.193	0.294	0.252	0.180	0.153
	Ib-05	0.420	0.112	0.184	0.146	0.230	0.276	0.198	0.241	0.190	0.245	0.224
	Ib-06	0.681	0.069	0.071	0.067	0.181	0.080	0.245	0.315	0.211	0.216	0.213
	Ib-07	0.286	0.024	0.122	0.097	0.000	0.253	0.245	0.208	0.196	0.243	0.167
	Ib-08	0.516	0.275	0.215	0.282	0.296	0.336	0.549	0.661	0.306	0.566	0.400
	Ib-09	0.178	0.434	0.378	0.379	0.233	0.402	0.558	0.735	0.202	0.662	0.416

Table S2-1 (continued)

Subpopulation		Locus										Average
		pX021	pX043	pX047	pX070	pX079	pX130	pX143	pX144	pX155	pX194	
Ib	Ib-10	0.665	0.000	0.070	0.000	0.049	0.000	0.228	0.287	0.129	0.291	0.172
	Ib-11	0.536	0.176	0.165	0.180	0.154	0.346	0.298	0.258	0.126	0.298	0.254
	Ib-12	0.577	0.275	0.219	0.316	0.281	0.216	0.318	0.297	0.185	0.296	0.298
	Ib-13	0.547	0.173	0.000	0.305	0.101	0.104	0.252	0.329	0.324	0.498	0.263
	Ib-14	0.622	0.167	0.174	0.420	0.221	0.349	0.202	0.277	0.073	0.377	0.288
	Ib-15	0.605	0.000	0.000	0.042	0.021	0.000	0.286	0.399	0.239	0.345	0.194
	Ib-16	0.547	0.000	0.019	0.026	0.000	0.000	0.352	0.415	0.175	0.335	0.187
All		0.519	0.651	0.662	0.686	0.732	0.627	0.547	0.647	0.781	0.581	0.643

Table S2-2. Allele frequencies in ten polymorphic loci of pinewood nematode

Locus pX021

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX021_a01	0.779	0.804	0.869	0.765	0.874	0.834	0.218	0.466	0.701
pX021_a02	0.129	0.088	0.000	0.057	0.001	0.044	0.743	0.030	0.137
pX021_a03	0.000	0.000	0.000	0.054	0.000	0.000	0.000	0.336	0.049
pX021_a04	0.000	0.000	0.000	0.009	0.000	0.000	0.001	0.048	0.007
pX021_a05	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.046	0.006
pX021_a06	0.004	0.003	0.001	0.002	0.003	0.005	0.001	0.001	0.002
pX021_a07	0.004	0.003	0.002	0.001	0.003	0.006	0.000	0.001	0.002
pX021_a08	0.003	0.003	0.001	0.001	0.001	0.003	0.002	0.003	0.002
pX021_a09	0.001	0.004	0.004	0.002	0.002	0.004	0.002	0.000	0.002
pX021_a10	0.001	0.002	0.002	0.003	0.003	0.004	0.001	0.002	0.002
pX021_a11	0.003	0.004	0.004	0.003	0.004	0.000	0.002	0.001	0.003
pX021_a12	0.000	0.001	0.005	0.004	0.003	0.002	0.000	0.002	0.002
pX021_a13	0.001	0.002	0.005	0.004	0.003	0.002	0.001	0.001	0.003
pX021_a14	0.002	0.001	0.001	0.005	0.002	0.002	0.001	0.001	0.002
pX021_a15	0.004	0.003	0.004	0.002	0.003	0.004	0.000	0.001	0.002
pX021_a16	0.000	0.003	0.003	0.003	0.002	0.003	0.000	0.000	0.002
pX021_a17	0.000	0.005	0.004	0.002	0.001	0.002	0.000	0.001	0.002
pX021_a18	0.003	0.001	0.003	0.002	0.004	0.002	0.001	0.001	0.002
pX021_a19	0.004	0.000	0.005	0.004	0.002	0.002	0.000	0.003	0.002
pX021_a20	0.001	0.002	0.002	0.002	0.004	0.001	0.000	0.001	0.002
pX021_a21	0.003	0.002	0.002	0.001	0.001	0.004	0.001	0.001	0.002
pX021_a22	0.003	0.002	0.005	0.004	0.004	0.002	0.002	0.000	0.003
pX021_a23	0.001	0.003	0.004	0.002	0.003	0.003	0.000	0.001	0.002
pX021_a24	0.000	0.003	0.004	0.002	0.003	0.001	0.001	0.000	0.002
pX021_a25	0.003	0.001	0.002	0.002	0.001	0.004	0.001	0.001	0.002
pX021_a26	0.001	0.002	0.002	0.002	0.001	0.002	0.001	0.003	0.002
pX021_a27	0.000	0.002	0.000	0.004	0.003	0.002	0.000	0.000	0.001
pX021_a28	0.003	0.002	0.005	0.000	0.003	0.002	0.000	0.001	0.002
pX021_a29	0.002	0.001	0.000	0.004	0.001	0.001	0.000	0.001	0.001
pX021_a30	0.000	0.001	0.003	0.003	0.003	0.002	0.000	0.000	0.001
pX021_a31	0.004	0.002	0.000	0.001	0.005	0.002	0.000	0.002	0.002
pX021_a32	0.000	0.001	0.002	0.003	0.001	0.001	0.001	0.005	0.002
pX021_a33	0.000	0.004	0.001	0.002	0.001	0.002	0.001	0.001	0.002
pX021_a34	0.001	0.000	0.005	0.003	0.002	0.002	0.002	0.002	0.002
pX021_a35	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.012	0.001
pX021_a36	0.000	0.001	0.004	0.000	0.003	0.002	0.001	0.002	0.002
pX021_a37	0.004	0.001	0.002	0.001	0.003	0.002	0.000	0.001	0.002
pX021_a38	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.000	0.002
pX021_a39	0.000	0.003	0.003	0.002	0.001	0.001	0.000	0.001	0.002
pX021_a40	0.000	0.002	0.001	0.001	0.003	0.002	0.000	0.002	0.001
pX021_a41	0.003	0.003	0.002	0.001	0.000	0.004	0.000	0.001	0.002
pX021_a42	0.001	0.002	0.000	0.002	0.001	0.002	0.000	0.000	0.001
pX021_a43	0.004	0.001	0.000	0.002	0.002	0.002	0.000	0.001	0.001
pX021_a44	0.004	0.003	0.006	0.001	0.001	0.000	0.000	0.000	0.002
pX021_a45	0.002	0.003	0.000	0.001	0.003	0.003	0.001	0.000	0.002
pX021_a46	0.002	0.001	0.000	0.001	0.003	0.001	0.001	0.001	0.001
pX021_a47	0.002	0.000	0.001	0.004	0.002	0.001	0.000	0.000	0.001
pX021_a48	0.002	0.003	0.004	0.002	0.001	0.001	0.000	0.000	0.002

Table S2-2 (continued)

pX021_a49	0.001	0.000	0.002	0.001	0.003	0.002	0.000	0.003	0.002
pX021_a50	0.001	0.000	0.000	0.001	0.002	0.002	0.000	0.001	0.001
pX021_a51	0.004	0.000	0.002	0.001	0.001	0.002	0.000	0.002	0.002
pX021_a52	0.001	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001
pX021_a53	0.003	0.002	0.000	0.001	0.003	0.002	0.000	0.000	0.001
pX021_a54	0.000	0.001	0.004	0.002	0.002	0.002	0.000	0.000	0.001
pX021_a55	0.000	0.002	0.002	0.001	0.002	0.002	0.002	0.001	0.001
pX021_a56	0.001	0.000	0.000	0.002	0.001	0.001	0.000	0.001	0.001
pX021_a57	0.001	0.001	0.005	0.000	0.001	0.001	0.001	0.000	0.001
pX021_a58	0.000	0.001	0.000	0.000	0.002	0.002	0.001	0.000	0.001
pX021_a59	0.000	0.002	0.001	0.000	0.004	0.001	0.000	0.001	0.001
pX021_a60	0.000	0.000	0.002	0.002	0.000	0.004	0.000	0.001	0.001
pX021_a61	0.000	0.003	0.002	0.000	0.002	0.001	0.000	0.000	0.001
pX021_a62	0.000	0.001	0.005	0.001	0.002	0.001	0.001	0.001	0.002
pX021_a63	0.000	0.002	0.000	0.003	0.001	0.000	0.000	0.003	0.001
pX021_a64	0.000	0.001	0.002	0.002	0.001	0.002	0.000	0.001	0.001

Locus pX043

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX043_a01	0.626	0.570	0.862	0.599	0.879	0.612	0.219	0.019	0.548
pX043_a02	0.220	0.187	0.000	0.260	0.000	0.279	0.292	0.928	0.271
pX043_a03	0.019	0.013	0.000	0.017	0.002	0.013	0.280	0.010	0.044
pX043_a04	0.002	0.099	0.000	0.000	0.001	0.000	0.157	0.013	0.034
pX043_a05	0.027	0.012	0.006	0.025	0.001	0.016	0.005	0.009	0.013
pX043_a06	0.000	0.027	0.000	0.001	0.000	0.000	0.013	0.006	0.006
pX043_a07	0.000	0.019	0.000	0.000	0.000	0.000	0.010	0.005	0.004
pX043_a08	0.005	0.002	0.004	0.003	0.005	0.002	0.000	0.000	0.003
pX043_a09	0.004	0.002	0.002	0.003	0.004	0.003	0.000	0.000	0.002
pX043_a10	0.001	0.002	0.005	0.002	0.002	0.002	0.001	0.001	0.002
pX043_a11	0.000	0.002	0.003	0.001	0.003	0.004	0.001	0.000	0.002
pX043_a12	0.002	0.003	0.001	0.003	0.005	0.001	0.000	0.000	0.002
pX043_a13	0.002	0.002	0.004	0.002	0.002	0.002	0.000	0.000	0.002
pX043_a14	0.000	0.004	0.000	0.000	0.000	0.000	0.005	0.000	0.001
pX043_a15	0.000	0.002	0.001	0.003	0.004	0.001	0.000	0.000	0.001
pX043_a16	0.000	0.002	0.003	0.005	0.003	0.002	0.000	0.000	0.002
pX043_a17	0.002	0.000	0.003	0.002	0.000	0.002	0.001	0.001	0.002
pX043_a18	0.002	0.001	0.009	0.002	0.003	0.002	0.001	0.000	0.002
pX043_a19	0.002	0.002	0.005	0.003	0.004	0.002	0.000	0.000	0.002
pX043_a20	0.000	0.002	0.002	0.006	0.003	0.003	0.001	0.000	0.002
pX043_a21	0.002	0.001	0.003	0.002	0.004	0.001	0.001	0.000	0.002
pX043_a22	0.004	0.001	0.004	0.000	0.002	0.003	0.002	0.000	0.002
pX043_a23	0.008	0.004	0.004	0.001	0.002	0.002	0.001	0.000	0.003
pX043_a24	0.000	0.001	0.003	0.001	0.004	0.003	0.000	0.000	0.001
pX043_a25	0.004	0.001	0.004	0.003	0.001	0.003	0.000	0.000	0.002
pX043_a26	0.000	0.004	0.006	0.001	0.002	0.002	0.000	0.000	0.002
pX043_a27	0.004	0.001	0.005	0.002	0.003	0.001	0.002	0.001	0.002
pX043_a28	0.002	0.002	0.002	0.002	0.004	0.002	0.000	0.000	0.002
pX043_a29	0.000	0.002	0.000	0.002	0.005	0.002	0.001	0.000	0.001
pX043_a30	0.000	0.001	0.001	0.001	0.003	0.002	0.001	0.000	0.001
pX043_a31	0.004	0.001	0.000	0.003	0.003	0.002	0.001	0.000	0.002

Table S2-2 (continued)

pX043_a32	0.004	0.001	0.002	0.001	0.002	0.002	0.000	0.000	0.002
pX043_a33	0.002	0.000	0.004	0.005	0.000	0.003	0.000	0.000	0.002
pX043_a34	0.000	0.004	0.004	0.001	0.002	0.001	0.001	0.000	0.002
pX043_a35	0.004	0.001	0.000	0.001	0.004	0.000	0.000	0.000	0.001
pX043_a36	0.002	0.002	0.000	0.003	0.002	0.003	0.001	0.000	0.002
pX043_a37	0.000	0.001	0.005	0.002	0.005	0.001	0.000	0.000	0.002
pX043_a38	0.002	0.001	0.003	0.001	0.002	0.002	0.000	0.000	0.001
pX043_a39	0.000	0.001	0.004	0.002	0.002	0.002	0.001	0.000	0.001
pX043_a40	0.002	0.002	0.001	0.002	0.003	0.001	0.001	0.000	0.001
pX043_a41	0.000	0.003	0.003	0.003	0.002	0.002	0.001	0.000	0.002
pX043_a42	0.002	0.001	0.002	0.002	0.003	0.001	0.001	0.000	0.001
pX043_a43	0.000	0.000	0.000	0.002	0.004	0.002	0.001	0.001	0.001
pX043_a44	0.004	0.001	0.005	0.000	0.003	0.002	0.000	0.000	0.002
pX043_a45	0.002	0.002	0.004	0.001	0.002	0.000	0.000	0.000	0.001
pX043_a46	0.007	0.001	0.001	0.003	0.000	0.001	0.001	0.000	0.002
pX043_a47	0.008	0.001	0.002	0.001	0.002	0.001	0.000	0.000	0.002
pX043_a48	0.002	0.001	0.003	0.002	0.003	0.000	0.000	0.000	0.001
pX043_a49	0.009	0.000	0.004	0.001	0.001	0.001	0.000	0.000	0.002
pX043_a50	0.002	0.004	0.000	0.003	0.001	0.002	0.000	0.000	0.001
pX043_a51	0.002	0.003	0.002	0.002	0.001	0.001	0.000	0.000	0.001
pX043_a52	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.003	0.001
pX043_a53	0.002	0.001	0.000	0.000	0.002	0.001	0.000	0.000	0.001
pX043_a54	0.000	0.001	0.001	0.002	0.002	0.001	0.001	0.000	0.001
pX043_a55	0.000	0.001	0.003	0.003	0.000	0.000	0.000	0.000	0.001
pX043_a56	0.000	0.001	0.002	0.001	0.001	0.000	0.000	0.001	0.001
pX043_a57	0.000	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.000
pX043_a58	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Locus pX047

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX047_a01	0.624	0.672	0.000	0.636	0.933	0.802	0.168	0.030	0.483
pX047_a02	0.260	0.267	0.000	0.315	0.001	0.092	0.009	0.934	0.235
pX047_a03	0.000	0.013	0.997	0.000	0.001	0.047	0.563	0.030	0.206
pX047_a04	0.060	0.000	0.000	0.000	0.000	0.000	0.193	0.000	0.032
pX047_a05	0.008	0.001	0.000	0.000	0.000	0.000	0.057	0.000	0.008
pX047_a06	0.003	0.001	0.003	0.000	0.000	0.002	0.006	0.000	0.002
pX047_a07	0.000	0.003	0.000	0.002	0.005	0.002	0.000	0.000	0.001
pX047_a08	0.002	0.002	0.000	0.003	0.003	0.002	0.000	0.000	0.001
pX047_a09	0.000	0.003	0.000	0.002	0.002	0.002	0.000	0.000	0.001
pX047_a10	0.005	0.002	0.000	0.003	0.004	0.002	0.000	0.000	0.002
pX047_a11	0.002	0.003	0.000	0.001	0.003	0.004	0.000	0.000	0.002
pX047_a12	0.003	0.002	0.000	0.002	0.001	0.006	0.000	0.000	0.002
pX047_a13	0.001	0.002	0.000	0.002	0.003	0.002	0.000	0.000	0.001
pX047_a14	0.002	0.003	0.000	0.005	0.002	0.002	0.000	0.000	0.002
pX047_a15	0.002	0.004	0.000	0.004	0.003	0.000	0.000	0.000	0.002
pX047_a16	0.003	0.000	0.000	0.001	0.004	0.001	0.001	0.001	0.001
pX047_a17	0.002	0.002	0.000	0.003	0.002	0.003	0.000	0.000	0.001
pX047_a18	0.000	0.002	0.000	0.003	0.003	0.002	0.000	0.000	0.001
pX047_a19	0.000	0.002	0.000	0.001	0.004	0.003	0.000	0.000	0.001
pX047_a20	0.002	0.001	0.000	0.001	0.002	0.004	0.000	0.000	0.001

Table S2-2 (continued)

pX047_a21	0.002	0.001	0.000	0.002	0.002	0.004	0.000	0.000	0.001
pX047_a22	0.002	0.002	0.000	0.002	0.004	0.001	0.000	0.000	0.001
pX047_a23	0.006	0.001	0.000	0.002	0.003	0.001	0.000	0.000	0.002
pX047_a24	0.003	0.002	0.000	0.002	0.002	0.004	0.000	0.000	0.002
pX047_a25	0.002	0.003	0.000	0.001	0.002	0.004	0.000	0.000	0.001
pX047_a26	0.000	0.001	0.000	0.002	0.003	0.002	0.000	0.000	0.001
pX047_a27	0.005	0.000	0.000	0.002	0.003	0.002	0.000	0.000	0.001
pX047_a28	0.002	0.002	0.000	0.001	0.004	0.002	0.000	0.000	0.001
pX047_a29	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.003	0.001
pX047_a30	0.000	0.001	0.000	0.001	0.000	0.001	0.001	0.001	0.001
pX047_a31	0.001	0.003	0.000	0.001	0.000	0.000	0.000	0.001	0.001
pX047_a32	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Locus pX070

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX070_a01	0.657	0.683	0.000	0.701	0.921	0.819	0.274	0.006	0.508
pX070_a02	0.000	0.150	0.000	0.097	0.001	0.000	0.265	0.912	0.178
pX070_a03	0.000	0.000	1.000	0.001	0.000	0.000	0.025	0.000	0.128
pX070_a04	0.273	0.011	0.000	0.040	0.000	0.061	0.309	0.008	0.088
pX070_a05	0.018	0.051	0.000	0.026	0.000	0.023	0.023	0.023	0.021
pX070_a06	0.000	0.004	0.000	0.023	0.000	0.007	0.030	0.000	0.008
pX070_a07	0.000	0.028	0.000	0.013	0.001	0.000	0.024	0.007	0.009
pX070_a08	0.019	0.004	0.000	0.017	0.000	0.015	0.003	0.000	0.007
pX070_a09	0.000	0.005	0.000	0.022	0.000	0.006	0.011	0.000	0.006
pX070_a10	0.000	0.002	0.000	0.007	0.000	0.005	0.004	0.000	0.002
pX070_a11	0.000	0.000	0.000	0.000	0.000	0.000	0.013	0.018	0.004
pX070_a12	0.000	0.002	0.000	0.003	0.004	0.004	0.001	0.001	0.002
pX070_a13	0.002	0.003	0.000	0.004	0.004	0.003	0.001	0.000	0.002
pX070_a14	0.000	0.001	0.000	0.001	0.005	0.004	0.000	0.000	0.001
pX070_a15	0.002	0.002	0.000	0.001	0.005	0.003	0.000	0.000	0.002
pX070_a16	0.000	0.002	0.000	0.002	0.004	0.007	0.001	0.000	0.002
pX070_a17	0.002	0.006	0.000	0.004	0.002	0.003	0.002	0.000	0.002
pX070_a18	0.000	0.005	0.000	0.004	0.003	0.001	0.001	0.000	0.002
pX070_a19	0.004	0.002	0.000	0.001	0.003	0.003	0.002	0.001	0.002
pX070_a20	0.002	0.000	0.000	0.002	0.003	0.006	0.001	0.000	0.002
pX070_a21	0.000	0.004	0.000	0.001	0.004	0.002	0.000	0.000	0.001
pX070_a22	0.002	0.005	0.000	0.003	0.002	0.003	0.000	0.000	0.002
pX070_a23	0.008	0.003	0.000	0.000	0.004	0.002	0.000	0.000	0.002
pX070_a24	0.000	0.001	0.000	0.002	0.006	0.003	0.001	0.000	0.002
pX070_a25	0.000	0.004	0.000	0.005	0.001	0.004	0.000	0.000	0.002
pX070_a26	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.015	0.002
pX070_a27	0.004	0.003	0.000	0.003	0.002	0.002	0.001	0.000	0.002
pX070_a28	0.000	0.002	0.000	0.000	0.003	0.004	0.001	0.000	0.001
pX070_a29	0.000	0.002	0.000	0.002	0.004	0.003	0.001	0.000	0.001
pX070_a30	0.000	0.003	0.000	0.001	0.003	0.001	0.000	0.000	0.001
pX070_a31	0.003	0.001	0.000	0.002	0.003	0.000	0.001	0.000	0.001
pX070_a32	0.000	0.003	0.000	0.002	0.004	0.001	0.001	0.000	0.001
pX070_a33	0.002	0.001	0.000	0.003	0.003	0.001	0.000	0.000	0.001
pX070_a34	0.000	0.001	0.000	0.000	0.003	0.001	0.001	0.000	0.001
pX070_a35	0.002	0.002	0.000	0.000	0.001	0.002	0.000	0.000	0.001

Table S2-2 (continued)

pX070_a36	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.004	0.001
pX070_a37	0.000	0.001	0.000	0.003	0.002	0.002	0.000	0.000	0.001
pX070_a38	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.004	0.001
pX070_a39	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000
pX070_a40	0.000	0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.000
pX070_a41	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000
pX070_a42	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Locus pX079

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX079_a01	0.169	0.000	0.000	0.483	0.993	0.730	0.360	0.011	0.343
pX079_a02	0.344	0.140	0.000	0.189	0.000	0.047	0.360	0.932	0.252
pX079_a03	0.222	0.674	0.988	0.139	0.001	0.112	0.007	0.004	0.268
pX079_a04	0.108	0.075	0.000	0.043	0.000	0.008	0.006	0.007	0.031
pX079_a05	0.104	0.075	0.000	0.036	0.001	0.008	0.004	0.011	0.030
pX079_a06	0.009	0.000	0.000	0.018	0.000	0.022	0.078	0.005	0.016
pX079_a07	0.002	0.000	0.000	0.020	0.001	0.015	0.061	0.012	0.014
pX079_a08	0.000	0.008	0.000	0.000	0.000	0.000	0.089	0.011	0.013
pX079_a09	0.006	0.000	0.004	0.027	0.001	0.025	0.004	0.000	0.008
pX079_a10	0.002	0.001	0.006	0.022	0.001	0.021	0.002	0.000	0.007
pX079_a11	0.015	0.010	0.000	0.006	0.000	0.001	0.008	0.005	0.006
pX079_a12	0.012	0.012	0.002	0.009	0.000	0.001	0.003	0.001	0.005
pX079_a13	0.005	0.005	0.000	0.002	0.000	0.002	0.005	0.001	0.002
pX079_a14	0.000	0.000	0.000	0.001	0.000	0.003	0.010	0.001	0.002
pX079_a15	0.000	0.000	0.000	0.003	0.002	0.004	0.002	0.000	0.001
pX079_a16	0.000	0.000	0.000	0.004	0.002	0.002	0.001	0.000	0.001

Locus pX130

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX130_a01	0.785	0.658	0.905	0.583	0.892	0.737	0.318	0.013	0.611
pX130_a02	0.140	0.009	0.002	0.185	0.001	0.154	0.646	0.041	0.147
pX130_a03	0.000	0.154	0.000	0.087	0.000	0.000	0.001	0.894	0.142
pX130_a04	0.000	0.000	0.000	0.041	0.000	0.012	0.000	0.000	0.007
pX130_a05	0.000	0.030	0.000	0.001	0.000	0.000	0.001	0.007	0.005
pX130_a06	0.000	0.015	0.000	0.013	0.001	0.007	0.000	0.001	0.005
pX130_a07	0.000	0.025	0.000	0.002	0.000	0.000	0.001	0.011	0.005
pX130_a08	0.002	0.012	0.003	0.002	0.002	0.002	0.000	0.005	0.003
pX130_a09	0.000	0.012	0.000	0.001	0.000	0.000	0.000	0.015	0.004
pX130_a10	0.000	0.017	0.000	0.002	0.000	0.000	0.001	0.006	0.003
pX130_a11	0.000	0.000	0.000	0.013	0.000	0.007	0.000	0.000	0.002
pX130_a12	0.002	0.001	0.003	0.005	0.004	0.003	0.002	0.000	0.002
pX130_a13	0.000	0.007	0.003	0.002	0.004	0.002	0.002	0.000	0.002
pX130_a14	0.002	0.003	0.003	0.001	0.005	0.004	0.001	0.001	0.002
pX130_a15	0.004	0.003	0.002	0.001	0.004	0.005	0.000	0.000	0.002
pX130_a16	0.001	0.002	0.005	0.004	0.003	0.001	0.001	0.000	0.002
pX130_a17	0.004	0.000	0.003	0.002	0.005	0.003	0.001	0.000	0.002
pX130_a18	0.005	0.004	0.000	0.003	0.002	0.002	0.000	0.001	0.002
pX130_a19	0.003	0.002	0.000	0.004	0.003	0.001	0.003	0.000	0.002

Table S2-2 (continued)

pX130_a20	0.002	0.003	0.002	0.003	0.002	0.002	0.002	0.000	0.002
pX130_a21	0.001	0.002	0.004	0.003	0.001	0.002	0.000	0.000	0.002
pX130_a22	0.000	0.003	0.005	0.000	0.004	0.003	0.000	0.000	0.002
pX130_a23	0.005	0.001	0.004	0.002	0.002	0.003	0.000	0.000	0.002
pX130_a24	0.003	0.002	0.002	0.000	0.003	0.004	0.000	0.000	0.002
pX130_a25	0.003	0.001	0.004	0.002	0.004	0.002	0.001	0.000	0.002
pX130_a26	0.005	0.001	0.001	0.002	0.002	0.002	0.000	0.000	0.002
pX130_a27	0.001	0.002	0.006	0.001	0.004	0.001	0.001	0.000	0.002
pX130_a28	0.002	0.003	0.004	0.002	0.003	0.002	0.000	0.000	0.002
pX130_a29	0.002	0.001	0.003	0.002	0.002	0.002	0.001	0.000	0.002
pX130_a30	0.000	0.001	0.003	0.002	0.003	0.003	0.001	0.000	0.002
pX130_a31	0.006	0.002	0.001	0.001	0.002	0.001	0.001	0.000	0.002
pX130_a32	0.003	0.001	0.004	0.001	0.003	0.001	0.000	0.000	0.002
pX130_a33	0.000	0.001	0.004	0.002	0.002	0.002	0.001	0.000	0.002
pX130_a34	0.005	0.001	0.003	0.001	0.003	0.003	0.001	0.000	0.002
pX130_a35	0.003	0.002	0.002	0.002	0.002	0.001	0.000	0.000	0.002
pX130_a36	0.001	0.001	0.001	0.001	0.002	0.002	0.000	0.001	0.001
pX130_a37	0.004	0.001	0.002	0.002	0.001	0.003	0.002	0.000	0.002
pX130_a38	0.002	0.001	0.001	0.003	0.003	0.001	0.000	0.000	0.001
pX130_a39	0.002	0.001	0.001	0.002	0.004	0.002	0.000	0.000	0.001
pX130_a40	0.000	0.002	0.002	0.001	0.002	0.002	0.000	0.000	0.001
pX130_a41	0.002	0.002	0.000	0.001	0.000	0.000	0.002	0.006	0.002
pX130_a42	0.000	0.003	0.000	0.002	0.001	0.002	0.000	0.000	0.001
pX130_a43	0.000	0.001	0.003	0.001	0.004	0.001	0.001	0.000	0.001
pX130_a44	0.000	0.003	0.000	0.000	0.003	0.003	0.001	0.000	0.001
pX130_a45	0.000	0.001	0.002	0.001	0.002	0.000	0.000	0.001	0.001
pX130_a46	0.001	0.001	0.001	0.002	0.003	0.003	0.000	0.000	0.001
pX130_a47	0.001	0.002	0.003	0.001	0.002	0.002	0.001	0.000	0.002
pX130_a48	0.002	0.001	0.002	0.002	0.003	0.002	0.000	0.000	0.002
pX130_a49	0.000	0.001	0.001	0.002	0.002	0.000	0.002	0.000	0.001
pX130_a50	0.000	0.001	0.000	0.001	0.001	0.001	0.001	0.000	0.001
pX130_a51	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Locus pX143

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX143_a01	0.626	0.584	0.927	0.608	0.911	0.798	0.198	0.160	0.602
pX143_a02	0.296	0.346	0.003	0.336	0.001	0.136	0.771	0.809	0.337
pX143_a03	0.001	0.005	0.000	0.002	0.004	0.003	0.001	0.000	0.002
pX143_a04	0.000	0.001	0.006	0.003	0.004	0.003	0.002	0.001	0.002
pX143_a05	0.007	0.003	0.000	0.000	0.003	0.003	0.000	0.000	0.002
pX143_a06	0.000	0.002	0.006	0.001	0.002	0.003	0.000	0.000	0.002
pX143_a07	0.003	0.002	0.003	0.005	0.002	0.002	0.000	0.001	0.002
pX143_a08	0.006	0.001	0.002	0.001	0.003	0.002	0.001	0.000	0.002
pX143_a09	0.003	0.001	0.004	0.003	0.004	0.003	0.001	0.000	0.002
pX143_a10	0.000	0.004	0.002	0.001	0.002	0.002	0.002	0.001	0.002
pX143_a11	0.002	0.002	0.000	0.002	0.003	0.003	0.000	0.001	0.002
pX143_a12	0.002	0.002	0.000	0.003	0.003	0.000	0.000	0.000	0.001
pX143_a13	0.000	0.003	0.003	0.002	0.004	0.002	0.000	0.000	0.002
pX143_a14	0.003	0.002	0.005	0.000	0.005	0.001	0.000	0.000	0.002
pX143_a15	0.005	0.001	0.004	0.003	0.002	0.001	0.000	0.000	0.002

Table S2-2 (continued)

pX143_a16	0.000	0.001	0.003	0.001	0.004	0.001	0.000	0.001	0.001
pX143_a17	0.002	0.001	0.002	0.000	0.003	0.003	0.000	0.000	0.001
pX143_a18	0.002	0.003	0.000	0.002	0.003	0.000	0.000	0.000	0.001
pX143_a19	0.000	0.002	0.000	0.002	0.002	0.002	0.000	0.001	0.001
pX143_a20	0.000	0.003	0.000	0.001	0.003	0.002	0.000	0.000	0.001
pX143_a21	0.002	0.003	0.000	0.001	0.003	0.002	0.001	0.000	0.001
pX143_a22	0.000	0.000	0.005	0.002	0.001	0.004	0.001	0.000	0.002
pX143_a23	0.003	0.001	0.002	0.004	0.002	0.002	0.000	0.000	0.002
pX143_a24	0.002	0.002	0.000	0.002	0.000	0.000	0.003	0.002	0.001
pX143_a25	0.003	0.003	0.000	0.000	0.002	0.002	0.000	0.001	0.001
pX143_a26	0.002	0.001	0.002	0.001	0.002	0.002	0.000	0.000	0.001
pX143_a27	0.000	0.002	0.000	0.001	0.000	0.001	0.003	0.001	0.001
pX143_a28	0.002	0.001	0.000	0.001	0.000	0.000	0.005	0.002	0.001
pX143_a29	0.005	0.002	0.000	0.001	0.001	0.001	0.000	0.001	0.002
pX143_a30	0.000	0.001	0.004	0.001	0.003	0.003	0.001	0.001	0.002
pX143_a31	0.000	0.002	0.000	0.000	0.000	0.001	0.001	0.005	0.001
pX143_a32	0.005	0.001	0.004	0.001	0.004	0.000	0.000	0.001	0.002
pX143_a33	0.005	0.001	0.000	0.001	0.003	0.002	0.000	0.000	0.002
pX143_a34	0.000	0.001	0.004	0.003	0.003	0.003	0.000	0.000	0.002
pX143_a35	0.002	0.001	0.002	0.002	0.002	0.003	0.000	0.000	0.001
pX143_a36	0.002	0.004	0.000	0.001	0.000	0.000	0.001	0.003	0.001
pX143_a37	0.003	0.000	0.005	0.002	0.004	0.001	0.000	0.000	0.002
pX143_a38	0.002	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.001
pX143_a39	0.001	0.001	0.000	0.000	0.000	0.000	0.002	0.005	0.001
pX143_a40	0.006	0.001	0.003	0.000	0.001	0.001	0.001	0.000	0.002
pX143_a41	0.000	0.001	0.000	0.000	0.000	0.000	0.001	0.001	0.000

Locus pX144

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX144_a01	0.499	0.475	0.914	0.655	0.926	0.798	0.194	0.034	0.562
pX144_a02	0.391	0.244	0.000	0.095	0.000	0.049	0.458	0.782	0.252
pX144_a03	0.000	0.000	0.000	0.182	0.000	0.066	0.277	0.011	0.067
pX144_a04	0.000	0.167	0.000	0.013	0.001	0.008	0.002	0.113	0.038
pX144_a05	0.025	0.030	0.002	0.001	0.000	0.002	0.008	0.018	0.011
pX144_a06	0.039	0.023	0.000	0.000	0.000	0.001	0.014	0.006	0.010
pX144_a07	0.000	0.013	0.000	0.000	0.000	0.001	0.007	0.015	0.004
pX144_a08	0.001	0.005	0.000	0.000	0.000	0.000	0.006	0.012	0.003
pX144_a09	0.000	0.000	0.000	0.007	0.000	0.007	0.006	0.000	0.003
pX144_a10	0.008	0.003	0.000	0.000	0.000	0.001	0.006	0.005	0.003
pX144_a11	0.001	0.003	0.002	0.001	0.003	0.003	0.001	0.000	0.002
pX144_a12	0.002	0.002	0.000	0.002	0.004	0.003	0.001	0.000	0.002
pX144_a13	0.002	0.001	0.005	0.004	0.003	0.002	0.001	0.000	0.002
pX144_a14	0.003	0.002	0.001	0.001	0.003	0.004	0.000	0.000	0.002
pX144_a15	0.001	0.001	0.008	0.000	0.003	0.003	0.000	0.001	0.002
pX144_a16	0.001	0.001	0.004	0.001	0.003	0.003	0.000	0.000	0.002
pX144_a17	0.000	0.001	0.004	0.000	0.003	0.003	0.001	0.000	0.002
pX144_a18	0.001	0.001	0.001	0.003	0.003	0.003	0.000	0.000	0.002
pX144_a19	0.001	0.002	0.006	0.002	0.002	0.001	0.001	0.000	0.002
pX144_a20	0.001	0.002	0.003	0.001	0.003	0.002	0.001	0.000	0.002
pX144_a21	0.000	0.000	0.000	0.002	0.002	0.002	0.002	0.000	0.001

Table S2-2 (continued)

pX144_a22	0.000	0.000	0.003	0.003	0.002	0.003	0.001	0.000	0.002
pX144_a23	0.002	0.004	0.000	0.001	0.002	0.001	0.001	0.000	0.001
pX144_a24	0.000	0.001	0.002	0.002	0.004	0.002	0.000	0.000	0.001
pX144_a25	0.001	0.002	0.002	0.001	0.002	0.002	0.001	0.000	0.002
pX144_a26	0.000	0.002	0.001	0.001	0.001	0.003	0.001	0.001	0.001
pX144_a27	0.002	0.000	0.004	0.002	0.003	0.003	0.000	0.000	0.002
pX144_a28	0.001	0.001	0.005	0.001	0.003	0.002	0.001	0.000	0.002
pX144_a29	0.002	0.001	0.004	0.002	0.004	0.002	0.000	0.000	0.002
pX144_a30	0.002	0.001	0.003	0.001	0.001	0.004	0.001	0.000	0.002
pX144_a31	0.000	0.003	0.001	0.001	0.003	0.001	0.000	0.000	0.001
pX144_a32	0.001	0.001	0.003	0.002	0.002	0.002	0.001	0.000	0.002
pX144_a33	0.000	0.001	0.007	0.002	0.002	0.001	0.000	0.000	0.002
pX144_a34	0.004	0.002	0.002	0.001	0.002	0.001	0.001	0.000	0.002
pX144_a35	0.004	0.001	0.001	0.001	0.003	0.003	0.000	0.000	0.002
pX144_a36	0.002	0.001	0.004	0.002	0.002	0.003	0.001	0.000	0.002
pX144_a37	0.002	0.001	0.002	0.003	0.001	0.001	0.001	0.000	0.001
pX144_a38	0.000	0.000	0.004	0.001	0.003	0.001	0.000	0.000	0.001
pX144_a39	0.001	0.001	0.000	0.000	0.000	0.000	0.002	0.001	0.001
pX144_a40	0.001	0.002	0.000	0.000	0.000	0.000	0.001	0.001	0.001
pX144_a41	0.001	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000

Locus pX155

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX155_a01	0.017	0.214	0.968	0.253	0.001	0.066	0.664	0.896	0.385
pX155_a02	0.002	0.000	0.000	0.068	0.827	0.007	0.001	0.000	0.113
pX155_a03	0.120	0.000	0.000	0.319	0.000	0.528	0.124	0.001	0.136
pX155_a04	0.222	0.309	0.000	0.145	0.000	0.179	0.128	0.010	0.124
pX155_a05	0.378	0.366	0.000	0.067	0.000	0.089	0.000	0.004	0.113
pX155_a06	0.027	0.000	0.000	0.076	0.107	0.076	0.034	0.000	0.040
pX155_a07	0.000	0.005	0.000	0.012	0.000	0.005	0.003	0.048	0.009
pX155_a08	0.201	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.025
pX155_a09	0.000	0.002	0.001	0.005	0.016	0.002	0.002	0.001	0.004
pX155_a10	0.002	0.008	0.000	0.007	0.000	0.005	0.006	0.002	0.004
pX155_a11	0.000	0.000	0.000	0.002	0.018	0.000	0.000	0.000	0.003
pX155_a12	0.000	0.008	0.000	0.005	0.000	0.002	0.005	0.002	0.003
pX155_a13	0.000	0.001	0.005	0.002	0.009	0.000	0.002	0.002	0.003
pX155_a14	0.003	0.003	0.000	0.004	0.000	0.007	0.005	0.002	0.003
pX155_a15	0.004	0.020	0.000	0.002	0.000	0.000	0.000	0.001	0.003
pX155_a16	0.004	0.003	0.000	0.004	0.000	0.012	0.001	0.000	0.003
pX155_a17	0.000	0.000	0.000	0.001	0.015	0.000	0.000	0.000	0.002
pX155_a18	0.002	0.000	0.000	0.001	0.000	0.008	0.007	0.000	0.002
pX155_a19	0.000	0.000	0.000	0.005	0.000	0.006	0.004	0.000	0.002
pX155_a20	0.005	0.011	0.000	0.004	0.000	0.001	0.000	0.001	0.003
pX155_a21	0.009	0.012	0.000	0.002	0.000	0.001	0.000	0.000	0.003
pX155_a22	0.000	0.000	0.000	0.004	0.007	0.000	0.000	0.000	0.001
pX155_a23	0.000	0.004	0.003	0.001	0.000	0.001	0.001	0.004	0.002
pX155_a24	0.000	0.000	0.004	0.000	0.000	0.000	0.002	0.006	0.002
pX155_a25	0.000	0.000	0.004	0.001	0.000	0.001	0.002	0.005	0.001
pX155_a26	0.000	0.002	0.006	0.001	0.000	0.000	0.001	0.004	0.002
pX155_a27	0.000	0.014	0.000	0.000	0.000	0.001	0.000	0.001	0.002

Table S2-2 (continued)

pX155_a28	0.000	0.001	0.001	0.004	0.000	0.000	0.001	0.003	0.001
pX155_a29	0.000	0.004	0.000	0.002	0.000	0.000	0.003	0.002	0.001
pX155_a30	0.000	0.003	0.000	0.001	0.000	0.000	0.003	0.001	0.001
pX155_a31	0.000	0.000	0.008	0.000	0.000	0.000	0.001	0.004	0.002
pX155_a32	0.002	0.008	0.000	0.002	0.000	0.002	0.000	0.000	0.002
pX155_a33	0.002	0.002	0.000	0.000	0.000	0.001	0.000	0.000	0.001
pX155_a34	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
pX155_a35	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
pX155_a36	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Locus pX194

Allele	Regional population								Kyushu
	Ka	Ma	Am	Sh	Mi	Ni	Se	Ib	
pX194_a01	0.624	0.465	0.918	0.618	0.906	0.792	0.492	0.030	0.605
pX194_a02	0.230	0.308	0.000	0.284	0.000	0.116	0.399	0.793	0.266
pX194_a03	0.000	0.087	0.000	0.001	0.000	0.000	0.000	0.113	0.025
pX194_a04	0.043	0.025	0.007	0.017	0.003	0.009	0.033	0.016	0.019
pX194_a05	0.038	0.024	0.000	0.012	0.000	0.013	0.025	0.010	0.015
pX194_a06	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.011	0.003
pX194_a07	0.000	0.015	0.000	0.000	0.000	0.000	0.000	0.012	0.003
pX194_a08	0.005	0.002	0.005	0.005	0.003	0.003	0.002	0.000	0.003
pX194_a09	0.005	0.002	0.002	0.002	0.002	0.006	0.003	0.000	0.003
pX194_a10	0.002	0.000	0.008	0.001	0.003	0.004	0.002	0.000	0.003
pX194_a11	0.000	0.008	0.000	0.000	0.000	0.000	0.001	0.006	0.002
pX194_a12	0.002	0.003	0.000	0.002	0.003	0.003	0.002	0.000	0.002
pX194_a13	0.000	0.007	0.000	0.000	0.000	0.000	0.000	0.005	0.001
pX194_a14	0.010	0.001	0.000	0.004	0.001	0.001	0.002	0.000	0.002
pX194_a15	0.002	0.002	0.001	0.004	0.002	0.002	0.002	0.000	0.002
pX194_a16	0.000	0.002	0.003	0.001	0.005	0.001	0.003	0.000	0.002
pX194_a17	0.002	0.004	0.000	0.001	0.002	0.002	0.003	0.000	0.002
pX194_a18	0.000	0.003	0.000	0.004	0.004	0.005	0.000	0.000	0.002
pX194_a19	0.002	0.001	0.003	0.002	0.002	0.003	0.001	0.000	0.002
pX194_a20	0.002	0.001	0.003	0.001	0.004	0.001	0.001	0.000	0.002
pX194_a21	0.004	0.002	0.005	0.002	0.001	0.003	0.001	0.000	0.002
pX194_a22	0.003	0.000	0.003	0.001	0.004	0.001	0.001	0.000	0.002
pX194_a23	0.002	0.002	0.000	0.000	0.004	0.002	0.002	0.000	0.001
pX194_a24	0.000	0.000	0.004	0.003	0.004	0.002	0.001	0.000	0.002
pX194_a25	0.001	0.002	0.002	0.001	0.003	0.001	0.002	0.000	0.001
pX194_a26	0.004	0.002	0.003	0.001	0.003	0.002	0.000	0.000	0.002
pX194_a27	0.003	0.002	0.000	0.001	0.003	0.003	0.002	0.000	0.002
pX194_a28	0.002	0.001	0.001	0.001	0.003	0.002	0.002	0.000	0.001
pX194_a29	0.000	0.002	0.000	0.001	0.000	0.001	0.002	0.003	0.001
pX194_a30	0.001	0.002	0.002	0.001	0.003	0.002	0.000	0.000	0.002
pX194_a31	0.000	0.000	0.001	0.003	0.002	0.003	0.002	0.000	0.001
pX194_a32	0.000	0.001	0.000	0.003	0.002	0.002	0.001	0.000	0.001
pX194_a33	0.005	0.000	0.003	0.004	0.002	0.003	0.000	0.000	0.002
pX194_a34	0.000	0.001	0.001	0.002	0.004	0.001	0.001	0.000	0.001
pX194_a35	0.000	0.002	0.004	0.003	0.001	0.000	0.003	0.000	0.002
pX194_a36	0.000	0.000	0.003	0.003	0.004	0.001	0.001	0.001	0.002
pX194_a37	0.005	0.000	0.003	0.002	0.002	0.001	0.002	0.000	0.002
pX194_a38	0.000	0.001	0.000	0.002	0.004	0.001	0.002	0.001	0.001

Table S2-2 (continued)

pX194_a39	0.003	0.001	0.000	0.002	0.001	0.002	0.001	0.000	0.001
pX194_a40	0.000	0.002	0.005	0.002	0.001	0.001	0.001	0.000	0.001
pX194_a41	0.002	0.000	0.001	0.001	0.003	0.003	0.001	0.000	0.002
pX194_a42	0.000	0.003	0.003	0.002	0.003	0.000	0.001	0.000	0.002
pX194_a43	0.000	0.000	0.005	0.002	0.002	0.003	0.001	0.000	0.002

Table S2-3. Nucleotide sequences of alleles in ten polymorphic loci of pinewood nematode

Locus pX021

pX021_a01
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a02
GTCAGTAATGGGACTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a03
GTCAGTAATGGGACTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGGATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAATGTC
GAGACTGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a04
GTCAGTAATGGGACTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGGATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a05
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAATGTC
GAGACTGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a06
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACACCTCGTTCACAGTTTGGGATGTCG

pX021_a07
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGCCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a08
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGGACATCTCGTTCACAGTTTGGGATGTCG

pX021_a09
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAGCAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a10
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCCTTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a11
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGGCGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a12
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGCTCACAGTTTGGGATGTCG

pX021_a13
GTCAGTAATGGGCCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a14
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GGGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a15
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCCTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a16
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACGACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

Table S2-3 (continued)

pX021_a17
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTCTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a18
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGGGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a19
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAGGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a20
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAGGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a21
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAGACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a22
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGGCCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a23
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTCCACAGTTTGGGATGTCG

pX021_a24
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGGCTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a25
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTCTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a26
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGGATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a27
GTCAGTAATGGGTCCAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a28
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGACCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a29
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACGAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a30
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGC
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a31
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGGTGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a32
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCCGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a33
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGACGTCG

Table S2-3 (continued)

pX021_a34
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGGTGTCG

pX021_a35
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGGATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAATGTC
GAGACTGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a36
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCCCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a37
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGCGACGACCATTCCCACCATTGGCTTTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a38
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
GCGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a39
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATCCCCACCATTGGCTTTAACGT
CGAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a40
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAGCTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a41
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCGCCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a42
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCGACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a43
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCCCGTTCACAGTTTGGGATGTCG

pX021_a44
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTCTGGGATGTCG

pX021_a45
GTCAGTAATGGGTCTAGCCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a46
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGGAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a47
GTCAGTAATGGGTCTAACCATCCCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a48
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGGCCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a49
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAACCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a50
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGCCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

Table S2-3 (continued)

pX021_a51
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACCACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a52
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGCGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a53
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCCCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a54
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACGGTTTGGGATGTCG

pX021_a55
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACGAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a56
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGGTTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a57
GTCAGTAATGGGTCTAACCACCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a58
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAGCATCTCGTTCACAGTTTGGGATGTCG

pX021_a59
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAGCGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a60
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCGTTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a61
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCGCAGTTTGGGATGTCG

pX021_a62
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTCGGGATGTCG

pX021_a63
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCCGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACAAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

pX021_a64
GTCAGTAATGGGTCTAACCATCTCCAGCTTGCTGGGCCGTCTCTTCGGCAACAAGGCTGTGCGAATCCTGATGGTTGGGCTCG
ACGCTGCCGGTAAAACAACGATCCTTTACGAACTGAAGCTGGGAGAGATTGTGACGACCATTCCCACCATTGGCTTTAACGTC
GAGACCGTCGACTACAAGAACATCTCGTTCACAGTTTGGGATGTCG

Locus pX043

pX043_a01
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a02
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTCGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTCT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

Table S2-3 (continued)

pX043_a03
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTCGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a04
TTATTGTGCTGCTGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a05
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTCT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a06
TTATTGTGCTGCTGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTCGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTCT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a07
TTATTGTGCTGCTGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTCT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a08
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGCTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a09
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTTT
GTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a10
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACGCGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a11
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGGAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a12
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTCTTCAGGAAAACGCCCGTTTGGTGT

pX043_a13
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAACTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a14
TTATTGTGCTGCTGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTCGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a15
TTATTGTGCCGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a16
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTACTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a17
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGGGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a18
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGGAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a19
TTATTGTGCTGCCGATTACAAGTACGATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

Table S2-3 (continued)

pX043_a20
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAGAACGCCCGTTTGGTGT

pX043_a21
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCCCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a22
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGCGT

pX043_a23
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGCTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a24
TTATCGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a25
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTCCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a26
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCGGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a27
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAGAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a28
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTC
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a29
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
CGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a30
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
GTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a31
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTCTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a32
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAGGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a33
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCCTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a34
TTACTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a35
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCACGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a36
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGGTCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

Table S2-3 (continued)

pX043_a37
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGA

pX043_a38
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAGACGCCCGTTTGGTGT

pX043_a39
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAGCGCCCGTTTGGTGT

pX043_a40
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCCGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a41
CTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a42
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGCGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a43
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAAGTTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a44
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGCGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a45
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGGTTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a46
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGCTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a47
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCGGGAAAACGCCCGTTTGGTGT

pX043_a48
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCGGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a49
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCCTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a50
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGGAAACGCCCGTTTGGTGT

pX043_a51
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGGACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a52
TTATTGTGCTGCCGATTGCAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTCGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTCT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a53
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTGGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

Table S2-3 (continued)

pX043_a54
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACCGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a55
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGCCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a56
TTATTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTGATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a57
TTATTGTGCTGCCGTTTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAACTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

pX043_a58
TTACTGTGCTGCCGATTACAAGTACAATTATGGCGAGCCTATTCCATGTGAAGAATTGGTTCAGTCGCTGTGTAATGAGAAGC
AACGTTACACGCAGATCGGAGGTAAAAAATTGCTTTTTCTCTTTTTGAGATTTAGGTTAGGTGTATGAGTCTGGAAATATTTTT
TGTGTTTTCAGGAAAACGCCCGTTTGGTGT

Locus pX047

pX047_a01
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a02
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATCGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCAC
TTCCGGAAGAATCCATA

pX047_a03
ACAGTTTGGACACCAGCAAGTGGAATTTTGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a04
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGCTGGAGTGTTACACGA
CCAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCA
CTTCCGGAAGAATCCATA

pX047_a05
ACAGTTTGGACACCAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGA
CCAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCACGTTCA
CTTCCGGAAGAATCCATA

pX047_a06
ACAGTTTGGACACCAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGA
CCAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCA
CTTCCGGAAGAATCCATA

pX047_a07
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAGCAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a08
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAGCGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a09
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGCTATACGA
CCAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCA
CTTCCGGAAGAATCCATA

pX047_a10
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTCCAC
TTCCGGAAGAATCCATA

pX047_a11
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCGGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

Table S2-3 (continued)

pX047_a12
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAGACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a13
ACAGTTTGGACACTAGCAAGTGGAATCTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGA
CCAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCA
CTTCCGGAAGAATCCATA

pX047_a14
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAGCAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a15
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAGTCCATA

pX047_a16
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGGGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a17
ACAGTTTGGACACTAGCAAGTGGAATTCCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGA
CCAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCA
CTTCCGGAAGAATCCATA

pX047_a18
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGGATCCATA

pX047_a19
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGGGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a20
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAGACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a21
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGGGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a22
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAGGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a23
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGGGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a24
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCCCGTTCAC
TTCCGGAAGAATCCATA

pX047_a25
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGGTGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a26
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGGACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a27
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGGGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a28
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCGTA

Table S2-3 (continued)

pX047_a29
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATTGAGGCTAGAAGAGAGAATGTGGATGGATGCTCGTTCACT
TCCGGAAGAATCCATA

pX047_a30
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAGACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATCGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCAC
TTCCGGAAGAATCCATA

pX047_a31
ACAGTTTGGACACTAGCAAGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGGGAACGGCAACTTGGTGATCGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCAC
TTCCGGAAGAATCCATA

pX047_a32
ACAGTTTGGACACTAGCACGTGGAATTTCGAGACCGGAAACAACAATGGATGGGGAAACTCGGAGTTGGAGTGTTATACGAC
CAGCAGCAATAATGTGAGAGTTGAGAACGGCAACTTGGTGATCGAGGCAAGAAGAGAGAATGTGGATGGATGCTCGTTCAC
TTCCGGAAGAATCCATA

Locus pX070

pX070_a01
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a02
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGGCTGTTCCATCCC

pX070_a03
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGCGGCACCCAGGCTGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCACGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTTTCATTTTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a04
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAGATTCGCACTGGAACCTACAAGCGTCTGTTCCATCCC

pX070_a05
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a06
CTGGGAATTGTACTGCCTGGAACACGGCATCCAACCCGATGGCCAACTGCCCAGCGGCACCCAGGCTGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCACGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTTTCATTTTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a07
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGGCTGTTCCATCCC

pX070_a08
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAGATTCGCACTGGAACCTACAAGCGTCTGTTCCATCCC

pX070_a09
CTGGGAATTGTACTGCCTGGAACACGGCATCCAACCCGATGGCCAACTGCCCAGCGGCACCCAGGCTGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCACGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a10
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTTTCATTTTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a11
CTGGGAATTGTACTGTCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a12
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGGCCTACAAGCGTCTGTTCCATCCC

Table S2-3 (continued)

pX070_a13
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGCTCCATCCC

pX070_a14
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGGCCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a15
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGCTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a16
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCCTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a17
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCGACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a18
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACCGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a19
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGGCGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a20
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGCCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a21
CTGGGAATTGTACTGCCTGGAACATGGAGTCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a22
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCGGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a23
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTCCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a24
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTCTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a25
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGCCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a26
CTGGGAATTGTACTGTCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGGCTGTTCCATCCC

pX070_a27
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGCCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a28
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGCTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a29
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCCTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

Table S2-3 (continued)

pX070_a30
CTGGGAGTTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a31
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATCGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a32
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGCCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a33
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCGTTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a34
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAACTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a35
CTGGGAATCGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a36
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGGCGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGGCTGTTCCATCCC

pX070_a37
CTGGGAATTGCACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGACAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGTCTGTTCCATCCC

pX070_a38
CTGGGAATTGTACTGCCTGGAACATGGAGTCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGGCTGTTCCATCCC

pX070_a39
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATCTTCTTTGCCATTGTCTTCTGTTTTTAGACGAAATTCGCACTGGGACCTACAAGCGGCTGTTCCATCCC

pX070_a40
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTCGCCATTGTCTTCTGTTTTTAGACGAGATTCGCACTGGAACCTACAAGCGTCTGTTCCATCCC

pX070_a41
CCGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGATTCC
TTTTCGACCTTCTTCAGCGAGACCGGCACTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAGATTCGCACTGGAACCTACAAGCGTCTGTTCCATCCC

pX070_a42
CTGGGAATTGTACTGCCTGGAACATGGAATCCAACCCGATGGCCAACTGCCCAGTGGCACCCAGGCCGGAACGGACGGTTCC
TTTTCGACCTTCTTCAGCGAGACCGGCCCTGGTCACTATGTTCCGCGCGCCGTCTTGGTCGATCTGGAGCCAACCGTCATTGGT
ATGTGGATAATTTTCTTTGCCATTGTCTTCTGTTTTTAGACGAGATTCGCACTGGAACCTACAAGCGTCTGTTCCATCCC

Locus pX079

pX079_a01
GCAGCATGGCAAAACCTCCGAACAAGTTCTTGCAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATCGAA
CATGATTATCTGCCGTCTCTCACATTTTC

pX079_a02
GCAACATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTTGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAAC
ATCATTATCTGCCGTCTCTCACATTTTC

pX079_a03
GCAGCATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAA
CATGATTATCTGCCGTCTCTCACATTTTC

Table S2-3 (continued)

pX079_a04
GCAGCATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAA
CATCATTATCTGCCGTCTCTCACATTTTC

pX079_a05
GCAACATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTTGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAAC
ATGATTATCTGCCGTCTCTCACATTTTC

pX079_a06
GCAGCATGGCAAAACCTCCGAACAAGTTCTTGCAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAA
CATCATTATCTGCCGTCTCTCACATTTTC

pX079_a07
GCAACATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTTGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATCGAAC
ATGATTATCTGCCGTCTCTCACATTTTC

pX079_a08
GCAGCATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTGGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTTGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAAC
ATCATTATCTGCCGTCTCTCACATTTTC

pX079_a09
GCAGCATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATCGAA
CATGATTATCTGCCGTCTCTCACATTTTC

pX079_a10
GCAGCATGGCAAAACCTCCGAACAAGTTCTTGCAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAA
CATGATTATCTGCCGTCTCTCACATTTTC

pX079_a11
GCAGCATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTTGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAAC
ATCATTATCTGCCGTCTCTCACATTTTC

pX079_a12
GCAACATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAA
CATGATTATCTGCCGTCTCTCACATTTTC

pX079_a13
GCAACATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATTGAA
CATCATTATCTGCCGTCTCTCACATTTTC

pX079_a14
GCAACATGGCAAAACCTCCGAACAAGTTCTTACAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATCGAA
CATGATTATCTGCCGTCTCTCACATTTTC

pX079_a15
GCAGCATGGCAAAACCTCCGAACAAGTTCTTGCAGAACTTGATCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCCTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATCGAA
CATGATTATCTGCCGTCTCTCACATTTTC

pX079_a16
GCAGCATGGCAAAACCTCCGAACAAGTTCTTGCAGAACTTGGTCCATTCTGGAGTAGTCTTGGGTGGAGAAAGGGCGTTTGG
TCCATCTCTGGCCAACACTTCTTCGACCTTAGCTGTGGTCAGACCCTAAAATAATATTTATTAAATCAGCCACATGAATCGAA
CATGATTATCTGCCGTCTCTCACATTTTC

Locus pX130

pX130_a01
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a02
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCTGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a03
CGAGTCAAGTTACTCACCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGCGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
GGATGAACTCAACGATGAGTCTGGATCTGGTA

Table S2-3 (continued)

pX130_a04
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTTTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGATCTGGTA

pX130_a05
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
GGATGAACTCAACGATGAGTCTGGATCTGGTA

pX130_a06
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGATCTGGTA

pX130_a07
CGAGTCAAGTTACTCACCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGCGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a08
CGAGTCAAGTTACTCACCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a09
CGAGTCAAGTTACTCACCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGCGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
GGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a10
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGCGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
GGATGAACTCAACGATGAGTCTGGATCTGGTA

pX130_a11
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTTTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a12
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGACCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a13
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGCGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a14
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGGACTCAACGATGAGTCTGGCTCTGGTA

pX130_a15
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACCCAACGATGAGTCTGGCTCTGGTA

pX130_a16
CGAGTCAAGTTACTCATCAGGCACTGGGAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a17
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCCGGCTCTGGTA

pX130_a18
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
GGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a19
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGCCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a20
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGCCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

Table S2-3 (continued)

pX130_a21
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGGTCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a22
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAGGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a23
CGAGTCAAGTTACTCATCAGGCGCTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a24
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAGCTCAACGATGAGTCTGGCTCTGGTA

pX130_a25
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGGCCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a26
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCCCTGGTA

pX130_a27
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGCTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a28
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAGGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a29
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAGGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a30
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGGAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a31
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCCGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a32
CGAGTCAAGTTACCCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a33
CGAGTCAAGTTACTCATCAGGCACTGGAGGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a34
CGAGTCAAGCTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a35
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCCGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a36
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGGAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a37
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCCTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

Table S2-3 (continued)

pX130_a38
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTGTGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a39
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGGTGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a40
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGGAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a41
CGAGTCAAGTTACTCACCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCTGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a42
CGAGCCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a43
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGCCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a44
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCGACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a45
CGAGTCAAGTTACTCATCGGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a46
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGCCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a47
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGACGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a48
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGACGAGTCTGGCTCTGGTA

pX130_a49
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCCATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

pX130_a50
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAACTGAAAGCTCCGGAAGTGGTTCTGGATCCGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCGGG
AGATGAACTCGACGATGAGTCTGGCTCTGGTA

pX130_a51
CGAGTCAAGTTACTCATCAGGCACTGGAAGTGGGGCTTCTCCCTATGTTGCTGAATCGACTGGACCGGTGTATGTGGATCCTT
ATGTCCCAGCTGAAAGCTCCGGAAGTGGTTCTGGATCTGGAGAAGTGGTTGGGTCTGGGTCCGGAAGTGGATCAGGGTCAGG
AGATGAACTCAACGATGAGTCTGGCTCTGGTA

Locus pX143.

pX143_a01
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a02
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

Table S2-3 (continued)

pX143_a03
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATCCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a04
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGGGTGGAATGACAAGTCAGCCATCGATGG

pX143_a05
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGCCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a06
CATCTTGGCCATCAAGGATATGGCCGGCGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a07
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTCGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a08
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCGGCCATCGATGG

pX143_a09
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCCGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a10
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAGTGACAAGTCAGCCATCGATGG

pX143_a11
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGCGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a12
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACGAGTCAGCCATCGATGG

pX143_a13
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGACGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a14
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAACCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a15
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGGTGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a16
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCCCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a17
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGGTCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a18
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGGCATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a19
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGGTTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

Table S2-3 (continued)

pX143_a20
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGCTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a21
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGCTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a22
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCCCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a23
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCCGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a24
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACACCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

pX143_a25
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGGGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a26
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGGTCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a27
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTCGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

pX143_a28
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGGCACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

pX143_a29
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAGGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a30
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGCCAGCCATCGATGG

pX143_a31
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGGTCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

pX143_a32
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCGAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a33
CACCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a34
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATCCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a35
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGGGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a36
CATCTTGGCCATCAGGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

Table S2-3 (continued)

pX143_a37
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCGTGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a38
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGCGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

pX143_a39
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCCCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTTGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCGTCGATGG

pX143_a40
CATCTTGGCCATCAAGGATATGGCCGGTGTTTTAAAGCCTCAGGCGGCCAAGTTGTTGATCTCTGCCCTTAGAGATCGCCATC
CAGACATCCCAATCCATGTCCATTCTCATGACACTGCTGGAGCTGCTGTTGCCTCGATGTCGGAGTGCGCCAAAGCTGGTGCT
GATGTTGTTGATGTGGCTGTGGATTCGATGAGTGGAATGACAAGTCAGCCATCGATGG

pX143_a41
CGGGTGAGAATCTGCTGGCCGGCATGCTCACGGTCGCCGACGGGGCGGTGGTGCTCGGAGAAGGCGGCTCACTGCTGAACAT
GACCGACGTCACACTCGATGCCGGCACATCGCTCGATGTCAGCGCCGCCACAGCACCTACGATCGGCGCACTCAACGGGGGA
TCGGGCAGTGCCCTTCGCGTCGGCGCAACCGCGATAACGCTCGACAGTGGCAACAACGGCACGTTCGGTGGAACGCTGAGTG
GCTCGGGAGGCATCGTCAAACAAGGCAGCGGTACGCAAACGCTCTCGGGCGGTATG

Locus pX144

pX144_a01
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a02
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAATGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a03
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGTTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAAGATATGT

pX144_a04
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAAGATATGT

pX144_a05
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAATGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a06
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a07
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAAGATATGT

pX144_a08
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAATGTCTTTTAGGTGCCAAACCAGAAGATATGT

pX144_a09
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGTTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a10
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAATGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a11
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCCTTTAGGTGCCAAACCAGAGGATATGT

pX144_a12
AACATGACATTGAGCATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a13
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGCCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a14
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAGCGTCTTTTAGGTGCCAAACCAGAGGATATGT

Table S2-3 (continued)

pX144_a15
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCCTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a16
AACATGGCATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a17
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGGAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a18
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGCGCCAAACCAGAGGATATGT

pX144_a19
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTCCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a20
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTCGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a21
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCCTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a22
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGGGGATATGT

pX144_a23
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a24
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTCCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a25
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAGATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a26
AACATGACGTTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a27
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAACCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a28
AACGTGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a29
AACATGACATTGAACGTATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a30
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a31
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGCCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a32
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGC

pX144_a33
AACATGACATTGAACATATTGTCGACTTCTGGCGCGGAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a34
AACATGACATTGAACATATCGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a35
AACATGACATCGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a36
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATCGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

Table S2-3 (continued)

pX144_a37
AACATGACATTGAACATATTGTCGACCTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a38
AACATGACATTGAACATATTGTCGACTTCTGGCGCGAAAATATTGGTGATCAATCACCTCTTCTTCGATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAACGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a39
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAATCTGAAATGTCTTTTAGGTGCCAAACCAGGGGATATGT

pX144_a40
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTTTTGTAGTCTGAAATGTCTTTTAGGTGCCAAACCAGAGGATATGT

pX144_a41
AACATGACATTGAACATATTGTCGACTTCTGGCGTGAAAATATTGGTGACCAATCACCTCTTCTTCAATCGATTGGTAGGTCA
TCCTCTTCTGGGTTATAGCCGATTGAAGAGTCTTGTAATCTGAAATGTCTTTTAGGTGCCAAACCAGAGGATATGT

Locus pX155

pX155_a01
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a02
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACCGGTTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAGTATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a03
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGGGGGGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a04
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCTGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTACAGATCTTGGAAACATTGT

pX155_a05
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAGGAAGCTGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACACTGT

pX155_a06
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a07
CTTTGATCTTCTCACCTTGGCGAAGGCTGAAAACCAACCCGACCGGTTCCTTCATGTGAACCCTGAAGATGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCGAATATCAAAGAAGCCGCCAAGAGAGGAGTGGGATATTATCATGAGGGAACAGCAA
GAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a08
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACCGGTTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ACTTGGTGAACGCTATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGAA
GAAGGATGTTCAGATCTTGGAAACATTAT

pX155_a09
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a10
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTACAGATCTTGGAAACATTGT

pX155_a11
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACCGGTTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAGTATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a12
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCTGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a13
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAGTATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

Table S2-3 (continued)

pX155_a14
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCTGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTACAGATCTTGGAAACATTGT

pX155_a15
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAGGAAGCTGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a16
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTACAGATCTTGGAAACATTGT

pX155_a17
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACCGGTTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a18
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGGGGGGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a19
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGGGGGGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACCTTGT

pX155_a20
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAGGAAGCTGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACATTGT

pX155_a21
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCTGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACACTGT

pX155_a22
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACATTGT

pX155_a23
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACACTGT

pX155_a24
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATCTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a25
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGGAACATTGT

pX155_a26
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAGGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a27
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTTCAGATCTTGGAAACATTGT

pX155_a28
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTCCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a29
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a30
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGCGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

Table S2-3 (continued)

pX155_a31
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGGGGGAACGGCGA
AGAAGGATGTCCAGATCTTGGAAACATTGT

pX155_a32
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCATGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAGGAAGCTGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACACTGT

pX155_a33
CTTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCCGACAGATTCCTCCACGTGAACCCGGAAGACGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCAAATATCAAGGAAGCTGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACGGCGA
AGAAGGATGTTCAGATCTTGGAAACACTGT

pX155_a34
CTTTGATCTTCTCACCTTGGCGAAGGCTGAAAACCAACCCGACCGGTTCCTTCATGTGAACCCTGAAGATGAAGATTTCCAGA
ATTTGGTGAACGCGATTGCTGATCCGAATATCAAAGAAGCCGCCAAGAGAGGAGTGGGATATTATCATGAGGGAACAGCAA
GAAGGACGTTCAGATCTTGGAAACCTTGT

pX155_a35
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAGTCAACCTGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCAGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCATGAGGGAACAGCGA
AGAAGGATGTACAGATCTTGGAAACATTGT

pX155_a36
ATTTGATCTTCTCACTTTGGCGAAGGCTGAAAATCAACCTGACAGATTCCTCCATGTGAACCCGGAAGATGAAGATTTCCGGA
ACTTGGTGAACGCGATTGCTGATCCAAATATTAAAGAAGCCGCCAAAAGAGGAGTGGGATATTATCGTGAGGGAACAGCGA
AGAAGGATGTACAGATCTTGGAAACATTGT

Locus pX194

pX194_a01
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a02
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GACTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a03
GAACTAATAACTCTATAATGCAAATGCCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGATGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a04
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a05
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GACTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a06
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGATGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a07
GAACTAATAACTCTATAATGCAAATGCCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GACTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a08
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGACCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a09
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCGAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a10
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCCTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

Table S2-3 (continued)

pX194_a11
GAACTAATAACTCTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GACTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a12
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAGTCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a13
GAACTAATAACACTATAATGCAAATGCCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGATGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a14
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAGAACCTGAGGTCGA

pX194_a15
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCCTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a16
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAGCCTGAGGTCGA

pX194_a17
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTCCCTGGAGGGAAAACCTGAGGTCGA

pX194_a18
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGC
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a19
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAGACCTGAGGTCGA

pX194_a20
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGACGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a21
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGCCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a22
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAGGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a23
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGGCCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a24
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCCTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a25
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGGGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a26
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGGGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a27
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCGGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

Table S2-3 (continued)

pX194_a28
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATCCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a29
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGCCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GACTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGGAAACCTGAGGTCGA

pX194_a30
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAGGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a31
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGGGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a32
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGGACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a33
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGGGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a34
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTCGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a35
GAACTAATAACACTATAATGCAGATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a36
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCCGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a37
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGGATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a38
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCCGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a39
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATCTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a40
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGGCTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a41
GAACTAATAACGCTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a42
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTCCCCTTTCCTGGAGGGAAAACCTGAGGTCGA

pX194_a43
GAACTAATAACACTATAATGCAAATACCGACCACTTCCACTGATCCCTTTCCAGAATCGCTGGTCGATTTGCAAGCAATCACA
TCTCTGGTCATGCCAATTCGCTCTGCCTGGACTAAGTCGTCGACTGGAACTGGAGCTGATGCGGTGCCGGACCTGAGGTGAGT
GCCTCAGCCCTGATTCTGACCTGGCTTTAGTGAATCCGAGTTCCCTCTCCTGGAGGGAAAACCTGAGGTCGA

Table S3. Nucleotide variations at 160 polymorphic sites in the 30 haplotypes of pinewood nematode detected in Kyushu

Haplotype (ht-)		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
*CX*1	40	A	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	43	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	61	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	.	.
	175	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	202	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	229	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	231	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	T	.	.	.	.	.	.	.	.	.	.
	253	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	274	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	289	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	G	.	G	.	.
	292	T	C	C	.	.	C	.	.	.	C	C	C	C	C	C	C	C	C	C	C	C	.	C	.	C	C	.	C	.	.
	340	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	367	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	394	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	A	.	.	.	A	.	A	.	A	A	.	A	.	.
	472	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	490	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	511	T	C	C	.	.	.	.	.	.	C	.	.	.	.	.	.	C	.	.	.	C	.	C	.	C	.	.	C	.	.
	518	G	.	.	.	A	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	544	G	A	A	.	.	T	.	.	.	A	.	.	.	.	.	.	A	.	.	.	A	.	A	.	A	.	.	A	.	.
	556	T	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	628	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	634	G	.	.	.	A	.	.	.	.	.	A	A	A	A	A	A	A	A	A	A	.	.	.	.	.	.	A	.	.	A
	682	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	703	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	724	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	796	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	823	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	829	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	868	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	.	.	T	.	.
	877	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	919	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	.	.	T	.	.
	934	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	940	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	961	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	1024	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	1072	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	1099	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	1116	A	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.
	1122	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	A
*CX*1-*trn*C	1160	A	–	–	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.	.	–	.	–	.	–	.	.	–	.	.
	1161	T	–	–	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.	.	–	.	–	.	–	.	.	–	.	.
	1162	A	–	–	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.	.	–	.	–	.	–	.	.	–	.	.
	1163	A	–	–	.	.	.	.	.	.	.	.	.	.	.	.	-	.	.	.	.	–	.	–	.	–	.	.	–	.	.
	1166	T	.	.	.	A	.	.	.	.	A	A	A	A	A	A	.	A	A	A	.	.	.	.	.	.	.	.	.	.	.
	1168	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.	.
*trn*C	1216	A	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	T	.	T	.	T	.	.	T	.	.
*trn*M	1233	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.
*trn*D-*trn*G	1364	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
*CX*2	1586	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.

Table S3 (continued)

	ht-	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
CX2	1688	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	G
	1754	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	1825	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	2015	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	2021	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	.	.
	2036	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
*rrn*L	2179	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	2182	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	2324	T	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	2328	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	2411	T	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	A	.
	2450	C	T	T	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	T	.
	2842	T	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
ND3	3152	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3199	A	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	3221	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	.	G	.	G	.	G	.	.	G	.	.
	3231	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	3242	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	A	.	.	.	A	.	A	.	A	A	.	A	.	.
	3419	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
ND5	3492	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3507	G	T	T	.	.	.	T	.	.	.	.	.	.	T	T	.	.	T	.	.	T	T	T	.	T	T	.	T	.	T
	3515	A	.	.	.	.	T	.	.	.	.	T	T	T	T	.	T	.	T	.	T	.	.	.	.	.	.	.	.	.	.
	3516	T	.	.	.	.	A	.	.	.	.	.	A	A	A	.	A	.	A	A	A	.	.	.	.	.	.	.	.	.	.
	3549	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	.	.	.	C	.	.
	3643	C	T	T	.	.	T	.	.	.	T	T	T	T	T	T	T	T	T	T	T	T	.	T	.	T	T	T	T	T	T
	3651	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3706	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	3810	A	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	3864	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	.	G	.	G	.	G	G	.	G	.	.
	3921	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	3972	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	.	.
	3975	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	3993	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	4002	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	4041	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4140	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	4164	A	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	4197	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4235	G	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4323	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4329	A	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	4374	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4416	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4449	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	T	.	T	.	T	T	.	T	.	.
	4455	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	.	.
	4514	A	T	T	.	.	T	.	.	.	T	T	T	T	T	T	T	T	T	T	T	T	.	T	.	T	T	T	T	.	T
	4524	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	A	.	.	.	A	.	A	.	A	A	.	A	.	.
	4587	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	4648	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	4653	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.

Table S3 (continued)

	ht-	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ND5	4716	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4760	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	C
	4771	A	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	4776	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	.	.	.
	4785	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	4836	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	.	.	T	.	.
	4937	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	.	T
	4944	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	4962	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
ND6	5238	G	A	A	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	5268	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	5271	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	5293	T	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	5331	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	5373	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	5388	G	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	5389	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	5442	A	.	.	.	.	.	.	.	.	.	G	G	G	G	G	G	.	.	G	.	.	.	.	.	.	.	.	.	.	.
	5451	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	5454	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	.	.	.	A	.	.
	5545	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	5559	T	.	.	C	C	.	C	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.
ND4L	5651	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	.	.	A	.	.
	5738	G	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	5756	C	T	T	.	.	T	.	.	.	T	T	T	T	T	T	T	T	T	T	T	T	.	T	.	T	.	.	T	.	.
	5768	G	.	.	.	.	A	.	.	.	.	A	A	A	A	A	A	A	A	A	A	.	.	.	.	.	.	.	.	.	.
	5774	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
*rrn*S	6028	G	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	.	.	T	.	.
	6130	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	6188	G	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	.	.	T	.	.
	6343	T	.	.	.	.	A	.	.	.	.	A	A	A	A	A	A	A	A	A	A	.	.	.	.	.	.	.	.	.	.
	6389	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	G
	6557	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
*trn*S	6665	T	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
*trn*Y	6748	G	.	.	.	.	.	.	.	–	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
ND1	6779	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	6802	A	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	6814	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	6847	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	6889	G	A	A	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	.	.
	6934	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	6940	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	A
	6943	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7009	T	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7108	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	7150	A	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	7159	G	A	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	A	.	A	A	.	A	.	.
	7189	T	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7201	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7306	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.

Table S3 (continued)

	ht-	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
*ND*1	7309	A	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
	7359	C	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7399	G	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7435	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	C	.	C	.	.
	7534	A	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	G	.	.	.	G	.	G	.	G	.	.	G	.	.
	7567	C	T	T	.	.	.	.	.	.	T	.	.	.	.	.	.	T	.	.	.	T	.	T	.	T	T	.	T	.	.
	7579	C	T	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	T	T	.	T	.	.
*ATP*6	7760	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7772	T	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	7844	T	C	C	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	7850	T	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	C	.	C	.	.	C	.	.
	7865	T	G	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	G	.	G	.	G	.	.	G	.	.
	7904	A	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	7985	C	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Number indicates the position of the SNPs and indels in the total sequence nucleotide in this study.

“.” indicates the same nucleotide with ht-01 at the polymorphic site.

“-” indicates deletion.