SHORT COMMUNICATION

Genetic diversity and population structure of *Raffaelea quercus-mongolicae*, a fungus associated with oak mortality in South Korea

By M.-S. Kim[1,5], P. A. Hohenlohe[2], K.-H. Kim[3], S.-T. Seo[3] and N. B. Klopfenstein[4]

[1]Department of Forestry, Environment and Systems, Kookmin University, Seoul 02707, South Korea; [2]Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID, USA; [3]Division of Forest Insect Pests and Diseases, Korea Forest Research Institutes, Seoul, South Korea; [4]USDA Forest Service, Rocky Mountain Research Station, Moscow, ID, USA; [5]E-mail: mkim@kookmin.ac.kr (for correspondence)

Summary

Raffaelea quercus-mongolicae is a fungus associated with oak wilt and deemed to cause extensive oak mortality in South Korea. Since the discovery of this fungus on a dead Mongolian oak (*Quercus mongolica*) in 2004, the mortality continued to spread southwards in South Korea. Despite continued expansion of the disease and associated significant impacts on forest ecosystems, information is lacking about the origin and genetic diversity of *R. quercus-mongolicae*. Restriction-site-associated DNA (RAD) sequencing was used to assess genetic diversity and population structure among five populations (provinces) of *R. quercus-mongolicae* in South Korea. In total, 179 single nucleotide polymorphisms (SNPs) were identified among 2,639 RAD loci across the nuclear genome of the 54 *R. quercus-mongolicae* isolates (0.0012 SNPs per bp), which displayed an overall low expected heterozygosity and no apparent population structure. The low genetic diversity and no apparent population structure among South Korean populations of this ambrosia beetle-vectored fungus support the hypothesis that this fungus was introduced to South Korea.

1 Introduction

Raffaelea quercus-mongolicae is a fungus associated with oak wilt mortality in South Korea (Kim et al. 2009). This fungus is vectored by a wood-boring ambrosia beetle, *Platypus koryoensis*, and it is phylogenetically distinct from *Raffaelea quercivora*, which causes a similar disease on oaks in Japan, and other *Raffaelea* spp. (Kim et al. 2009). It is a plausible cause of oak wilt disease although the pathogenicity of this fungus is not fully confirmed. After the discovery of this fungus on a dead Mongolian oak (*Quercus mongolica*) in 2004, the mortality was centred around Seoul and Gyeonggi Province in north-western South Korea, but the mortality has continued to spread southwards (Kim et al. 2009). The mortality rate (number of dead trees out of total infected trees) from this oak wilt disease is variable depending on the site and year, but it was 12.3% in 2013 (K.-H. Kim, personal communication). The primary host of *R. quercus-mongolicae* is Mongolian oak, which is widely distributed over central South Korea and plays an ecologically important role in South Korean forests. Despite continued expansion of the disease and associated negative impacts on forest ecosystems, information is lacking about the origin and genetic diversity of *R. quercus-mongolicae*.

Restriction-site-associated DNA sequencing (RAD-seq) has demonstrated diverse application for discovering 10 000s of single nucleotide polymorphisms (SNPs) in 100s of individuals, even in organisms for which few genomic resources presently exist (Davey et al. 2011). Although the cost of whole-genome sequencing is plummeting, the production of high-quality reference genomes is still a costly and time-consuming endeavour. In addition, genomic sequencing of every individual in a population is of only limited value for studying wild populations for which no reference genome is available. Presently, genomic resources for *R. quercus-mongolicae* are lacking, and only a few sequences of internal transcribed spacer of rDNA, large subunit rDNA, and β-tubulin regions are presently available in GenBank. Because a reference genome for *R. quercus-mongolicae* is not available, the RAD-seq markers are an optimal tool to assess genetic diversity and population structure of *R. quercus-mongolicae* across its range in South Korea.

Thus, the objective of this study was to use RAD-seq markers to assess genetic diversity and population structure of *R. quercus-mongolicae* across its range in South Korea, and provide insights as to whether this fungus is native or introduced to South Korea.

2 Materials and methods

Fifty-four isolates of *R. quercus-mongolicae* were collected from five provinces of South Korea (Seoul/Gyeonggi – 18 isolates, Gangwon – nine isolates, Chungcheong – eight isolates, Jeolla – 10 isolates and Gyeongsang – nine isolates) during 2005–2010 (Fig. 1a). Nine fungal isolates were collected prior to 2009 from infected Mongolian oaks and the remaining isolates (45 isolates) were directly isolated from the insect vector, *P. koryoensis*. The purity of the *R. quercus-mongolicae* cultures was ensured by single-spore isolation, and identification of the *R. quercus-mongolicae* was confirmed by morphological and DNA analyses described by Kim et al. (2009). Fungal isolates were cultured on potato dextrose agar media, and DNA was extracted using DNeasy Plant Maxi DNA extraction kits (Qiagen Inc., Valencia, CA, USA). A total of 3.0 μg (30 ng/μl) DNA was obtained for generating RAD-seq markers. RAD-seq library construction and DNA sequencing were carried
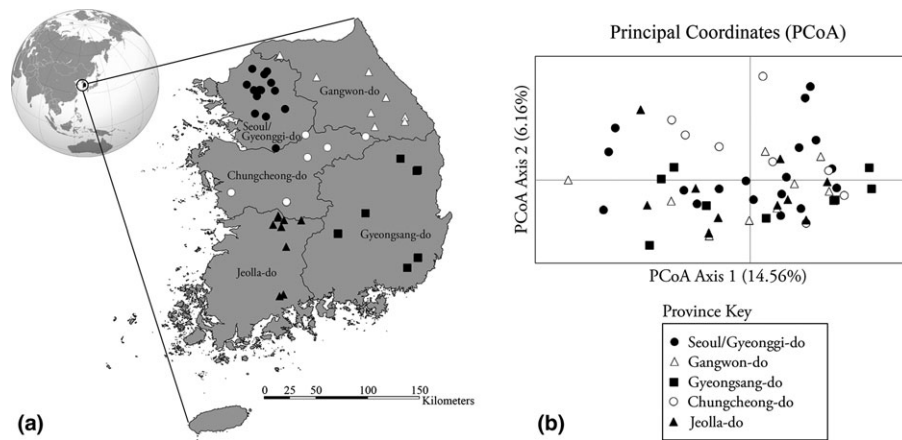
Fig. 1. (a) Geographic origins of *Raffaelea quercus-mongolicae* isolates used in this study. (b) Principal coordinates analysis of 54 *R. quercus-mongolicae* isolates sampled from different populations (provinces) throughout South Korea. The first two axes explain 21% of the observed variation. Populations of *R. quercus-mongolicae* are described in Table 1.

out by Floragenex, Inc. (Eugene, OR, USA), using the restriction enzyme *Sbf*I (a standard enzyme for RAD-seq when a reference genome is unavailable) and single-end, 100-bp Illumina HiSeq sequencing (Etter et al. 2011).

Following removal of the barcode, we trimmed reads to 90 bp and identified a reference set of putative RAD loci from a single individual using custom scripts designed by Floragenex, Inc., as follows. Using the individual with the highest coverage (isolated from Gangwon), we identified clusters of all reads within one nucleotide mismatch. Consensus sequences were retained for clusters with a minimum coverage of 4× and maximum of 700× and no more than 10 haplotypes, and these sequences were used as the reference set of putative RAD loci. Raw reads were aligned against this reference using bowtie (Langmead et al. 2009), allowing up to three mismatches and requiring unique alignment. We used samtools to designate genotypes, requiring coverage of at least 6× and a minimum minor allele frequency of 0.075 to call a SNP. We retained SNPs where at least 80% of samples were genotyped.

Single nucleotide polymorphism data were analysed with GᴇɴAʟEx 6.5 (Peakall and Smouse 2006). The following genetic parameters were determined: percentage of polymorphic loci (PPL), number of SNPs, observed heterozygosity (H_O), expected heterozygosity (H_E) and inbreeding coefficient (F_{is}). In addition, Nei's unbiased genetic distance and pairwise F_{st} for all populations were also estimated. To test the correlations among genetic distances and geographic distances of populations, Mantel's (1967) tests (permutations = 999) were conducted. Principal coordinate analysis (PCoA), based on a covariance matrix with data standardization, was conducted. The analysis of molecular variance (ᴀᴍᴏᴠᴀ) was performed to determine whether hierarchical partitioning of genetic variation among populations exists.

3 Results and discussion

Sequencing the RAD-tag libraries generated 143 696 855 reads using Illumina HiSeq. We identified 2639 putative RAD loci, of which 1606 were genotyped in at least 44 of the 54 samples (80%). On these RAD loci, we identified 179 SNPs across the nuclear genome of the 54 *R. quercus-mongolicae* isolates (0.0012 SNPs per bp) with overall low H_E (0.0364–0.0428), indicating a low level of genetic diversity across these samples (Table 1; Data S1). The level of genetic diversity, as indicated by H_E, did not differ significantly among populations (p = 0.514), ranging from 0.0364 (JL: Jeolla Province) to 0.0428 (GW: Gangwon Province) (Table 1). The inbreeding coefficient (F_{is}) levels observed were low, and ranged from 0.08 (JL: Jeolla Province) to 0.14 (SG: Seoul/Gyeonggi Province) (Table 1). The SG population had the highest genetic diversity values for PPL (19.1%) and number of single nucleotide polymorphisms (91 SNPs), while the remaining populations had similar genetic diversity values for genetic diversity parameters (Table 1).

Table 1. Genetic diversity of *Raffaelea quercus-mongolicae* populations (provinces) in South Korea.

Population	N	PPL (%)	SNPs	H_O	H_E	F_{is}
Seoul/Gyeonggi (SG)	16.95	19.1	91	0.0107	0.0427	0.1398
Gangwon (GW)	8.46	13.6	65	0.0114	0.0428	0.0958
Gyeongsang (GS)	8.56	13.2	63	0.0095	0.0385	0.0953
Chungcheong (CC)	7.58	14.9	66	0.0113	0.0472	0.0978
Jeolla (JL)	9.41	12.6	60	0.0125	0.0364	0.0775
Mean	10.19	14.7	69	0.0111	0.0415	0.1012

N, average number of individuals genotyped at each locus; PPL, percentage of polymorphic loci; SNPs, # of single nucleotide polymorphisms; H_O, observed heterozygosity; H_E, expected heterozygosity; F_{is}, inbreeding coefficient.

Table 2. The Nei's unbiased genetic distance (above diagonal) and pairwise F_{st} (below diagonal) for all populations (provinces) of *Raffaelea quercus-mongolicae*.

	SG[1]	GW	GS	CC	JL
SG	–	0.002	0.005	0.004	0.004
GW	0.045	–	0.004	0.003	0.004
GS	0.059	0.072	–	0.003	0.003
CC	0.055	0.068	0.064	–	0.004
JL	0.055	0.074	0.061	0.070	–

[1]Populations of *R. quercus-mongolicae* are described in Table 1.

To assess the genetic relationships among populations, the average F_{st} for pairwise comparisons and Nei's unbiased genetic distance were calculated (Table 2). Pairwise F_{st} values ranged from 0.045 between SG and GW provinces to 0.074 between JL and GW provinces (Table 2). In addition, PCoA showed a low degree of genetic differentiation (Fig. 1b). The first two axes explained 21% of total variation, with no apparent genetic relationships among populations. Furthermore, a Mantel test showed a non-significant correlation between genetic and geographic distances for the entire data set ($r = -0.012$, p = 0.649). The AMOVA showed that the majority (99%) of variation exists within populations ($\phi_{PT} = 0.006$, p = 0.269), which provides no evidence of population differentiation.

The sudden emergence of oak wilt disease in South Korea is perhaps attributable to the ageing of host populations, increasing density of host trees and/or influences of climate change (Choi 2011), which could also be coupled with reduced resistance of the populations. Other possible explanations are that the fungal pathogen could have evolved to overcome resistance of the native hosts or that the fungus is non-pathogenic, but it has the ability to trigger a suicidal overreaction (the exaggerated response hypothesis) of native trees that it contacts (Hulcr and Dunn 2011). A recent study showed that *R. quercus-mongolicae* can colonize sapwood and disrupt sap flow, but pathogenicity of the fungus was not confirmed (Torii et al. 2014). That recent study might support the exaggerated response hypothesis; however, further inoculation studies are needed to verify the pathogenicity of the fungus. The cause of sudden oak mortality can be also explained by increasing populations of the ambrosia beetle, *P. koryoensis*, that have vectored *R. quercus-mongolicae* over the years and/or associated mass attacks (J.-H. Park, personal communication). Under this scenario, the considered hypothesis is that the fungus is a native (or naturalized) organism – a long coexisting, but unrecognized, pathogen. Other supporting evidence for the native fungus hypothesis includes that its symbiotic vector beetle, *P. koryoensis*, has been reported early in Korea (Murayama 1930) and is distributed across wide geographic areas in Korea. Furthermore, the existence of some trees that are relatively resistant to the disease may reflect historical exposure to *R. quercus-mongolicae*.

Nevertheless, the results from this study raise a critical issue about whether this pathogen was introduced to South Korea. The low genetic diversity and no apparent population structure among South Korean populations of *R. quercus-mongolicae* provide supporting evidence that *R. quercus-mongolicae* is a non-native fungus, which was introduced to South Korea. A similar oak wilt pathogen, *R. quercivora*, also displayed uniformity of several DNA markers across Japan (Matsuda et al. 2010), supporting the hypothesis that *R. quercivora* was introduced to Japan (Hulcr and Dunn 2011). A recently introduced fungus might be expected to have such low genetic diversity and minimal population structure, as suggested in another oak wilt pathogen, *Ceratocystis fagacearum* (Juzwik et al. 2008). To date, this newly described ambrosia beetle-vectored fungus has only been found in South Korea (Kim et al. 2009) and the origin of *R. quercus-mongolicae* remains unknown.

Understanding population structure is essential to examine potential differences in virulence of this fungus among geographic regions, evaluate pathways of spread and develop more efficient disease prediction and management methods. Continued studies are critical to confirm our conclusion that *R. quercus-mongolicae* was introduced to South Korea and determine its potential sources, evolutionary origin and pathogenic role.

Acknowledgements

References

Choi, W. I., 2011: Influence of global warming on forest coleopteran communities with special reference to ambrosia and bark beetles. J. Asia Pac. Entomol. **14**, 227–231.

Davey, J. W.; Hohenlohe, P. A.; Etter, P. D.; Boone, J. Q.; Catchen, J. M.; Blaxter, M. L., 2011: Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev. Genet. **12**, 499–510.

Etter, P. D.; Bassham, S.; Hohenlohe, P. A.; Johnson, E. A.; Cresko, W. A., 2011: SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: Molecular methods for evolutionary genetics. Ed. by Orgogozo, V.; Rockman, M. V. New York: Humana Press, pp. 157–178.

Hulcr, J.; Dunn, R. R., 2011: The sudden emergence of pathogenicity in insect-fungus symbioses threatens native forest ecosystems. Proc. R. Soc. B **278**, 2866–2873.

Juzwik, J.; Harrington, T. C.; MacDonald, W. L.; Appel, D. N., 2008: The origin of *Ceratocystis fagacearum*, the oak wilt fungus. Annu. Rev. Phytopathol. **46**, 13–26.

Kim, K.-H.; Choi, Y. J.; Seo, S.-T.; Shin, H.-D., 2009: *Raffaelea quercus-mongolicae* sp. nov. associated with *Platypus koryoensis* on oak in Korea. Mycotaxon **110**, 189–197.

Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S. L., 2009: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. **10**, R25.

Mantel, N., 1967: The detection of disease clustering and a generalized regression approach. Cancer Res. **27**, 209–220.

Matsuda, Y.; Kimura, K.; Ito, S.-I., 2010: Genetic characterization of *Raffaelea quercivora* isolates collected from areas of oak wilt in Japan. Mycoscience **51**, 310–316.

Murayama, J., 1930: Révisions des Familles des *Ipides* et des *Platypides* de Corée. J. Chosen Nat. Hist. Soc. **10**, 6–38.

Peakall, R.; Smouse, P. E., 2006: Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes **6**, 288–295.

Torii, M.; Matsuda, Y.; Seo, S. T.; Kim, K. H.; Ito, S.-I.; Moon, M. J.; Kim, S. H.; Yamada, T., 2014: The effect of *Raffaelea quercus-mongolicae* inoculations on the formation of non-conductive sapwood of *Quercus mongolica*. Mycobiology **42**, 210–214.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Genotypes for 54 samples of *Raffaelea quercus-mongolicae* at 179 single nucleotide polymorphisms in STRUCTURE format.