MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 2581–2602 doi: 10.1111/j.1365-294X.2011.05109.x

Gene genealogies reveal cryptic species and host preferences for the pine fungal pathogen *Grosmannia clavigera*

SEPIDEH M. ALAMOUTI,* VINCENT WANG,* SCOTT DiGUISTINI,* DIANA L. SIX,† JÖRG BOHLMANN,‡§ RICHARD C. HAMELIN,§ NICOLAS FEAU¶ and COLETTE BREUIL*

**Department of Wood Science, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, †Department of Ecosystem and Conservation Sciences, University of Montana, Missoula, MT 59812, USA, ‡Michael Smith Laboratory, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, §Department of Forest Science, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, ¶UMR 1202 BIOGECO, INRA 69 Route d'Arcachon, 33612 Cestas Cedex, France*

Abstract

***Grosmannia clavigera* is a fungal pathogen of pine forests in western North America and a symbiotic associate of two sister bark beetles: *Dendroctonus ponderosae* and *D. jeffreyi*. This fungus and its beetle associate *D. ponderosae* are expanding in large epidemics in western North America. Using the fungal genome sequence and gene annotations, we assessed whether fungal isolates from the two beetles inhabiting different species of pine in epidemic regions of western Canada and the USA, as well as in localized populations outside of the current epidemic, represent different genetic lineages. We characterized nucleotide variations in 67 genomic regions and selected 15 for the phylogenetic analysis. Using concordance of gene genealogies and distinct ecological characteristics, we identified two sibling phylogenetic species: Gc and Gs. Where the closely related *Pinus ponderosa* and *P. jeffreyi* are infested by localized populations of their respective beetles, Gc is present. In contrast, Gs is an exclusive associate of *D. ponderosae* mainly present on its primary host-tree *P. contorta*; however, in the current epidemic areas, it is also found in other pine species. These results suggest that the host-tree species and the beetle population dynamics may be important factors associated with the genetic divergence and diversity of fungal partners in the beetle-tree ecosystems. Gc represents the original *G. clavigera* holotype, and Gs should be described as a new species.**

Keywords: cryptic species, fungal pathogen, *Grosmannia clavigera*, host adaptation, mountain pine beetle, pine tree

Received 29 November 2010; revision received 23 February 2011; accepted 16 March 2011

Introduction

Because of global trade, and environmental and climate changes, phytophagous insects and insect-vectored fungi that are pathogenic to trees have the potential to undergo rapid population expansion and cause substantial ecological changes (Anderson *et al.* 2004). A key aspect of estimating risks to ecosystems because of the spread of native or introduced pathogenic species involves defining species boundaries and genetic diversity. A growing number of fungal pathogens that were originally reported as dispersed generalists are now described as collections of populations or sister species adapted to new hosts or environments (Burnett 2003; Giraud *et al.* 2008). However, like many other organisms, it is difficult to define species boundaries in fungi. Molecular approaches, such as phylogenetic species recognition by genealogical concordance (PSR), can be more effective than traditional concepts (Taylor *et al.* 2000; Dettman *et al.* 2003). While it can be challenging to identify a genetic threshold that defines a species

Correspondence: Colette Breuil, Fax: (+1) 604 822 8645; E-mail: Colette.Breuil@ubc.ca

boundary, it is becoming increasingly practical to generate genomic sequence data for delimiting species with many independent gene genealogies (Knowles & Carstens 2007).

Native bark beetles and their fungal associates, which evolve within coniferous trees, are among the most damaging forest pests in North America (Harrington 2005). The current *Dendroctonus ponderosae* [mountain pine beetle (MPB)] outbreak is the largest epidemic in recorded history. It has affected more than 16 million hectares of *Pinus contorta* forest in western Canada (http://www.for.gov.bc.ca/hfp/mountain_pine_beetle), leading to major impacts on ecosystem dynamics and associated economies (Kurz *et al.* 2008). MPB normally remains at low population levels within pine forests for many decades, but can rapidly erupt into large-scaled outbreaks, killing large areas of susceptible host trees (Raffa 1988). Climate change and large areas of susceptible host trees likely contribute to the epidemic expanding northward and into high-elevation pine forests, beyond the MPB's recorded historical range (Safranyik *et al.* 2010). Population studies of both beetles and fungal associates (i.e., *Grosmannia clavigera*) confirmed population expansion in the northern part of the beetle/fungal species range, where outbreak activity is currently increasing (Lee *et al.* 2007; Mock *et al.* 2007; Roe *et al.* 2011). Further, if conditions continue to be suitable for MPB in its current geographic range, there is a risk that the outbreak will expand eastward into the boreal forests via *P. banksiana* (Logan & Powell 2001; Bentz *et al.* 2010; Safranyik *et al.* 2010).

One of the most common characteristics of bark beetles is their association with the wood-colonizing filamentous ascomycetes grouped as ophiostomatoid (Six & Wingfield 2011). *Grosmannia clavigera* is an ophiostomatoid tree pathogen that forms a symbiotic association with MPB (Robinson-Jeffrey & Davidson 1968) and its sister species *D. jeffreyi* [jeffrey pine beetle (JPB); Six & Paine 1997]. While MPB and JPB have only subtle phenotypic and genetic differences, they inhabit different host trees. JPB is highly specialized, infests only *P. jeffreyi*, and has no history of large-scaled epidemics, whereas MPB inhabits its primary host *P. contorta* and 22 other pine species, but not *P. jeffreyi* (Wood 1982; Safranyik *et al.* 2010). These bark beetles carry similar mycoflora, and their geographic distributions overlap in some regions of the USA (Wood 1982; Six & Paine 1997; Kelley & Farrell 1998). *G. clavigera* is one of the most important fungal associates of MPB and JPB (Robinson-Jeffrey & Davidson 1968; Six & Paine 1997; Lee *et al.* 2006; Rice *et al.* 2007), and a central component of current MPB epidemics. Vectored fungi benefit from the association because the beetles carry them through the tree bark into a new host's tissues (Six & Wingfield 2011). The benefits reported for the beetle and its progeny include the fungi providing a suitable substrate for brood development, participating in weakening tree defences, and serving as a source of nutrients (Raffa & Berryman 1983; Harrington 2005; Bleiker & Six 2007; Lieutier *et al.* 2009; DiGuistini *et al.* 2011). While both beetle and fungi contribute to tree death, *G. clavigera* can kill trees without the beetle when manually inoculated into the host at a certain density (Solheim & Krokene 1998; Lee *et al.* 2006).

While species identification is important to understanding the ecology and biology of organisms, boundaries between closely related species often lack clear limits and diagnostic characteristics. In *G. clavigera,* the sexual state (teleomorph) has been rarely found in nature and is not produced in the laboratory (Lee *et al.* 2003). The fungus is predominantly haploid through its life cycle and is mainly known to reproduce asexually (Six & Paine 1999; Lee *et al.* 2007). Because teleomorphs are rare, morphological identification relies on asexual structures, which occur in a variety of forms including the anamorph *Leptographium* (Jacobs & Wingfield 2001; Six *et al.* 2003). Morphology in *G. clavigera* shows large variations in conidiophores and conidia shapes and sizes and can become confusingly indistinguishable from those of close relatives in the genus *Grosmannia* and *Leptographium* (Tsuneda & Hiratsuka 1984; Six *et al.* 2003). Distinguishing *G. clavigera* from morphologically similar species using molecular tools is also challenging (Zambino & Harrington 1992; Six *et al.* 2003; Lim *et al.* 2004). This fungus is part of a complex of closely related species that are associated with bark beetles and can infest the same pine-host trees (Six *et al.* 2003; Lim *et al.* 2004). Multigene phylogenies using ribosomal DNA and the housekeeping genes (e.g., actin, elongation factor 1, alpha and beta tubulin) have been useful for distinguishing *G. clavigera* from its most closely related species, with the exception of a generalist fungus *L. terebrantis* (Six *et al.* 2003; Lim *et al.* 2004; Roe *et al.* 2010). MPB- and JPB-associated *G. clavigera* are shown to represent a single species using molecular and morphological data, and the species is characterized with a low level of gene diversity within both beetle-associated populations (Six & Paine 1999; Six *et al.* 2003; Lim *et al.* 2004; Lee *et al.* 2007). AFLP markers suggested the presence of two genetically distinct groups within *G. clavigera* associated with MPB epidemic in Rocky Mountain of Alberta and the northern USA (Lee *et al.* 2007); however, these groups have not been supported by phylogenetic analysis of multiple loci (Roe *et al.* 2011) or by microsatellite markers (Tsui C, Roe A, El-Kassaby Y, Rice A, Massoumi Alamouti S, Sperling F, Cooke J, Bohlmann J, Breuil C, Hamelin R, unpublished data).

Beetle-associated plant pathogens like *G. clavigera* depend on beetle vectors and host trees to complete their life cycles (Harrington 2005). Therefore, detecting genetic isolation in relation to the degree of host specialization or evolution of symbiosis is relevant to this group of fungi. Although *G. clavigera* is considered to be a single species, it develops a tight ecological and/or biological association with different species of pines, as well as with the two distinct sibling beetle species MPB and JPB. Paine & Hanlon (1994) and Six & Paine (1998) suggested that *G. clavigera* isolated from MPB or JPB had different tolerance to host defence metabolites. These beetles segregate in different ecological niches; as such we hypothesize that this segregation might lead to genetic divergence in their fungal associates.

Defining species boundaries is essential for understanding the potential adaptive variations and the ecological and/or biological traits that may impact the evolution of beetle-associated fungi. Theoretical models that incorporate adaptation and divergence among pathogens are applicable to risk assessment and to developing control measures, and detailed genetic information on evolving species should improve such models (Giraud *et al.* 2010). Currently, information on genetic variation in *G. clavigera* is limited to few protein-coding genes and noncoding markers (Six & Paine 1999; Lee *et al.* 2007; Tsui *et al.* 2009; Roe *et al.* 2011). Here, we screened nucleotide polymorphisms in 67 loci and applied PSR using a subset of 15 protein-coding genes to assess whether genetically isolated lineages occur within *G. clavigera*, and whether host beetle and/or tree specialization may influence the evolution of these fungi. We combined the sequence data from the 15 loci to clarify how the species recognized by PSR were related to each other. We show evidence of recombination in these apparently asexual fungi. Finally, we conclude that *G. clavigera* consists of Gc and Gs lineages and discuss the ecology and biology of these fungi.

Materials and methods

Samples

We examined 166 isolates of *G. clavigera* and eight additional isolates of its four closely related species *G. aurea*, *L. longiclavatum*, *L. terebrantis* and *L. wingfieldii* (Tables 1 and S1, Supporting information). The *G. clavigera* isolates were collected from the two beetle associates MPB and JPB and their host trees at different sites in Canada and the USA. In California, where both MPB and JPB are present (i.e., sympatric regions), we collected 30 and 25 isolates from *P. jeffreyi* and *P. contorta*, respectively, as well as a few isolates from *P. ponderosa*. We also included *G. clavigera* from locations where only MPB is present (i.e., allopatric regions). Sixty-seven isolates were from *P. contorta* in Canada and the USA, 29 isolates from *P. ponderosae* in South Dakota and British Columbia (BC), and a limited number of isolates from other MPB-host species, including *P. albicaulis* in BC and *P. strobiformis* in Arizona. Our fungal collection provides a comprehensive coverage of the beetles' geographic distribution. This included samples from current MPB epidemics in Canada, Idaho and Montana as well as from previous outbreaks in the 1960s and 1980s. It also included samples from small, geographically isolated outbreak populations in South Dakota, California and Arizona. In the work described here, we refer to such populations as 'localized'. Figure 2a shows the sampling locations.

Isolation

Fungal isolations from beetle exoskeletons or mycangia, or from galleries in infested trees, were carried out following the methods described by Six & Paine (1997) and Massoumi Alamouti *et al.* (2007). Identification and molecular analyses were carried out from single-spore isolates. All cultures are maintained at the Breuil culture collection (University of British Columbia, Canada). Morphological features were determined from colonies grown on 2% MEA (20 g Difco malt extract, 10 g Difco agar and 1 l distilled water) or from fungi taken from beetle galleries. The *G. clavigera* reproductive structures were examined and compared to those described by Robinson-Jeffrey & Davidson (1968), using light microscopy.

Polymorphism detection

We identified polymorphic loci using two approaches. The first approach involved sequencing 28 candidate genes that were available from the *G. clavigera* genome sequence and EST-supported gene predictions (DiGuistini *et al.* 2007, 2009, 2011), followed by polymorphism discovery and verification. The second approach involved sequencing 39 putative polymorphic loci identified from an expressed sequence tag (EST) library obtained by pooling mRNA from eight *G. clavigera* isolates (DiGuistini *et al.* 2009) that were characterized as distinct haplotypes (i.e., unique sequence type) using the first approach. The target loci were identified from the genomic resource using CLCbio Genomics Workbench (CLC) 3.7.1 (Aarhus, Denmark). To discover polymorphisms, we sequenced the 67 loci across nine *G. clavigera* isolates (Table 1) chosen from distinct ecological and geographical sources. The sequences were aligned and analysed for polymorphisms using CLC. We validated the novel polymorphisms in 15 genes (Table 2) selected for further characterization in an

Table 1 Fungal isolates used in this study

Fungal species	Beetle associate	Host tree	Collection site (map no.*)	No. isolates†	Source‡	ID§	Collector (date sampled)
***Grosmannia* sp. (Gs clade)**	*Dendroctonus ponderosae*	*Pinus contorta*	Canada, BC, Riske Creek	1	UAMH 4585	B01	Whitney (1982)
			Canada, BC, Terry Fox Creek	1	NOF 1280	B02	Hiratsuka & Maruyama (1987)
			Canada, BC, Houston (1)	10	UAMH (11153)	B03	Lee (2003)
			Canada, BC, Tweedsmuir park (2)	1	CB SLA11	B04	
			Canada, BC, Williams Lake (3)	2	CB W14	B05	
			Canada, BC, Kamloops (4)	2	UAMH (11150)	B06	Lee (2003)
					UAMH (11151)	B07	Lee (2004)
			Canada, BC, Kelowna (5)	**2**	CB KDW4	B08	M. Alamouti (2007)
					UAMH (11152)	B09	
			Canada, BC, Manning Park (6)	5	CB M6	B10	Lee (2003)
			Canada, Alberta, Westcastle	1	UAMH 4818	A01	Tsuneda (1983)
			Canada, Alberta, Carbondale	1	NOF 842	A02	
			Canada, Alberta, Blairmore	1	NOF 2893	A03	Unknown (1983)
			Canada, Alberta, Banff (7)	15	CB B20	A04	Lee (2003)
					UAMH (11154)	A05	
					CB B6	A06	
					CB BW26	A07	
					CB B14	A08	
					UAMH (11155)	A09	
			Canada, Alberta, Cypress Hills (8)	**5 (2 trees)**	UAMH (11347)	A10	M. Alamouti (2007)
					UAMH 11355	A11	
					UAMH 11356	A12	
			USA, Montana, Hidden Valley (9)	10	**UAMH 11156**	M01	Six (2003)
					UAMH 11357	M02	
			USA, Idaho, Hell Roaring (10)	10	UAMH 11359	I01	Six (2002)
					UAMH 11360	I02	
			USA, California, Sierra Nevada (11)	**2**	DLS 1061	C01	Six (1995)
					DLS 1037	C02	
				23 (5 trees)	UAMH 11349	C03	M. Alamouti (2009)
					CB 23G23	C04	
					UAMH 11350	C05	
					UAMH 11361	C06	
					UAMH 11362	C07	

Table 1 (*Continued*)

Fungal species	Beetle associate	Host tree	Collection site (map no.*)	No. isolates†	Source‡	ID§	Collector (date sampled)
	D. ponderosae	*P. albicaulis*	Canada, BC, Nelson (12)	**5**	UAMH 11363 UAMH 11364	B11 B12	Blaiker (2007)
	D. ponderosae	*P. strobiformis*	USA, Arizona, Pinaleño mountains (13)	**7**	UAMH 11365 UAMH 11366	Z01 Z02	Six (2009)
	D. ponderosae	*P. ponderosae*	Canada, BC, Kamloops (4)	**8 (4 trees)**	UAMH 11367 UAMH 11368	B13 B14	M. Alamouti (2007)
			Canada, BC, Kelowna (5)	**5 (5 trees)**	CB KGW5	B15	
***G. clavigera* (Gc clade)**			Canada, BC, Cache Creek	1	ATCC (18086)	B16	Robinson-J. (1968)
			USA, South Dakota, Black Hills (14)	**15 (5 trees)**	UAMH 11369 UAMH 11370 UAMH 11371	D01 D02 D03	Bleiker & Six (2007)
			USA, California, Sierra Nevada	**1**	UAMH 11372	C08	Six (1993)
			USA, California, Lassen	**1**	UAMH 11373	C09	
	D. ponderosae		(15)	**1**	UAMH 11374	C10	
	D. jeffreyi	*P. jeffreyi*	USA, California, Sierra Nevada	10	C 843 UAMH 11375 DLS 776 UAMH 11376 DLS 681 DLS 771	C11 C12 C13 C14 C15 C16	Harrington (2005) Six (1999)
			USA, California, Lassen	10	UAMH 11377 DLS 210 UAMH 11378	C17 C18 C19	
			USA, California, San Bernardino (16)	1 **9**	DLS 52 DLS 1560 DLS 1565 UAMH (11351) UAMH 11352 DLS 1595	C20 C21 C22 C23 C24 C25	Six (1993) Hansen (2006)
Total number of isolates				166	62		
G. aurea	*Dendroctonus* sp.	*P. contorta*	Canada, BC, Invermere	1	CBS 438.69	UB	Davidson (1963)
Leptographium longiclavatum	*D. ponderosae*	*P. contorta*	Canada, BC, Kamloops	1	CB SLKW1436	LB	Lee *et al.* (2003)
	D. jeffreyi	*P. jeffreyi*	USA, California, Sierra Nevada	**1**	[C 845]	LC	Harrington (1999)

Table 1 (*Continued*)

Fungal species	Beetle associate	Host tree	Collection site (map no.*)	No. isolates†	Source‡	ID§	Collector (date sampled)
L. terebrantis	*D. ponderosae*	*P. contorta*	Canada, BC, Kamloops	**2**	CB 878AW1-2	TB1	Kim (2004)
					CB LPKRLT-3	TB2	Kim (2003)
	D. brevicomis	*P. ponderosae*	USA, California, Sierra Nevada	1	C 418	TC	Harrington (2003)
L. wingfieldii	*Tomicus piniperda*	*P. sylvestris*	France, Orléans	2	CBS 645.89	WF	Morelet (1984)
	NA	*P. brutia*	Greece, Thessaloniki		CBS 648.89	WG	Skarmoutsos (1987)

*Generalized map location of collection sites corresponding to Fig. 2a

†Number of isolates analysed for the ecological assessment using single-locus sequencing; Samples isolated and/or identified in this study are bolded; Isolates from the same locality are originated from different sources (i.e., from beetles and/or galleries collected from different tree individuals); otherwise number of isolation sources are shown in parentheses.

‡Isolates selected for 15 single-gene phylogenies; Source of isolates: UAMH, University of Alberta Microfungus Collection & Herbarium, Canada; NOF, Culture Collection of Northern Forestry Centre, Canada; ATCC, American Type Culture Collection, USA; CBS, Centraalbureau voor Schimmelcultures, Netherlands; Isolates beginning with CB, DLS and C are from culture collections of C. Breuil, University of British Columbia, Canada; D.L. Six, University of Montana, USA; and T.C. Harrington, Iowa State University, USA; respectively; Nine isolates chosen for the polymorphism discovery are shown in parentheses; The isolate re-identified (i.e., one example of misidentified cultures in the literature) using our new data set is shown in bracket; Isolates identified as distinct AFLP group by Lee *et al.* 2007 are bolded.

§Letters indicate the location, and colours indicate the host trees corresponding to Fig. 2a; Numbers indicate the number of isolates from each location.

additional 53 *G. clavigera* isolates and eight isolates of four closely related species (Table 1). Genomic locations and gene descriptions of the 67 *G. clavigera* loci screened for polymorphisms are listed in Table S2 in Supporting information, and concatenated alignment of these data sets is deposited in TreeBASE (TB2: http://www.treebase.org).

DNA extraction, primer design and sequencing

We followed DNA extraction method by Möller *et al.* (1992) for mycelia grown on 2% MEA (33 g Oxoid malt extract agar, 10 g Technical agar No. 3, and 1 l distilled water) plates overlaid with cellophane (gel dry grade, BioRad). Primer pairs were designed with optimal melting temperatures of 58–62 °C, using CLC (Table S2, Supporting information). PCR amplifications were performed following standard methods (Lim *et al.* 2004). Amplicons were purified and sequenced at the Sequencing and Genotyping Platform, CHUL Research Center (Québec, Canada). Sequence data were collected from one strand, except for new haplotypes, which were all confirmed by sequencing both strands. All sequences are available at GenBank (accession nos. HQ633073–HQ634118, Table S4, Supporting information).

Sequence alignments and analyses

Sequences were edited and aligned using Geneious 5.1 (Biomatters Ltd, New Zealand). Coding, intronic and untranslated (UTR) regions were determined based on alignment of DNA sequences to the *G. clavigera* genome sequence and gene prediction models. Genetic diversity indices and divergence analyses were assessed using DnaSP 5.10 (Librado & Rozas 2009). Net nucleotide divergence (Dxy) (Nei 1987) was calculated with the Tamura-Nei gamma correction model using Mega 4.0 (Tamura *et al.* 2007).

Gene trees and concatenated data phylogeny

Phylogenetic analyses were conducted using maximum parsimony (MP) and Bayesian inference of each of the 15 genes, as well as the combined data set of these genes. The best-fit model of sequence evolution for each gene was determined using the Akaike information criterion (AIC) implemented in JModelTest 0.1.1 (Posada 2008). MP trees were identified using PAUP* 4.0b10 (Swofford 2003) by heuristic searches and 100 random sequence additions. Gaps were treated as missing data, and no weighting was introduced in single-gene analysis. Confidence was examined using bootstrapping (BS) with 1000 replicates and the heuristic option (Felsen-

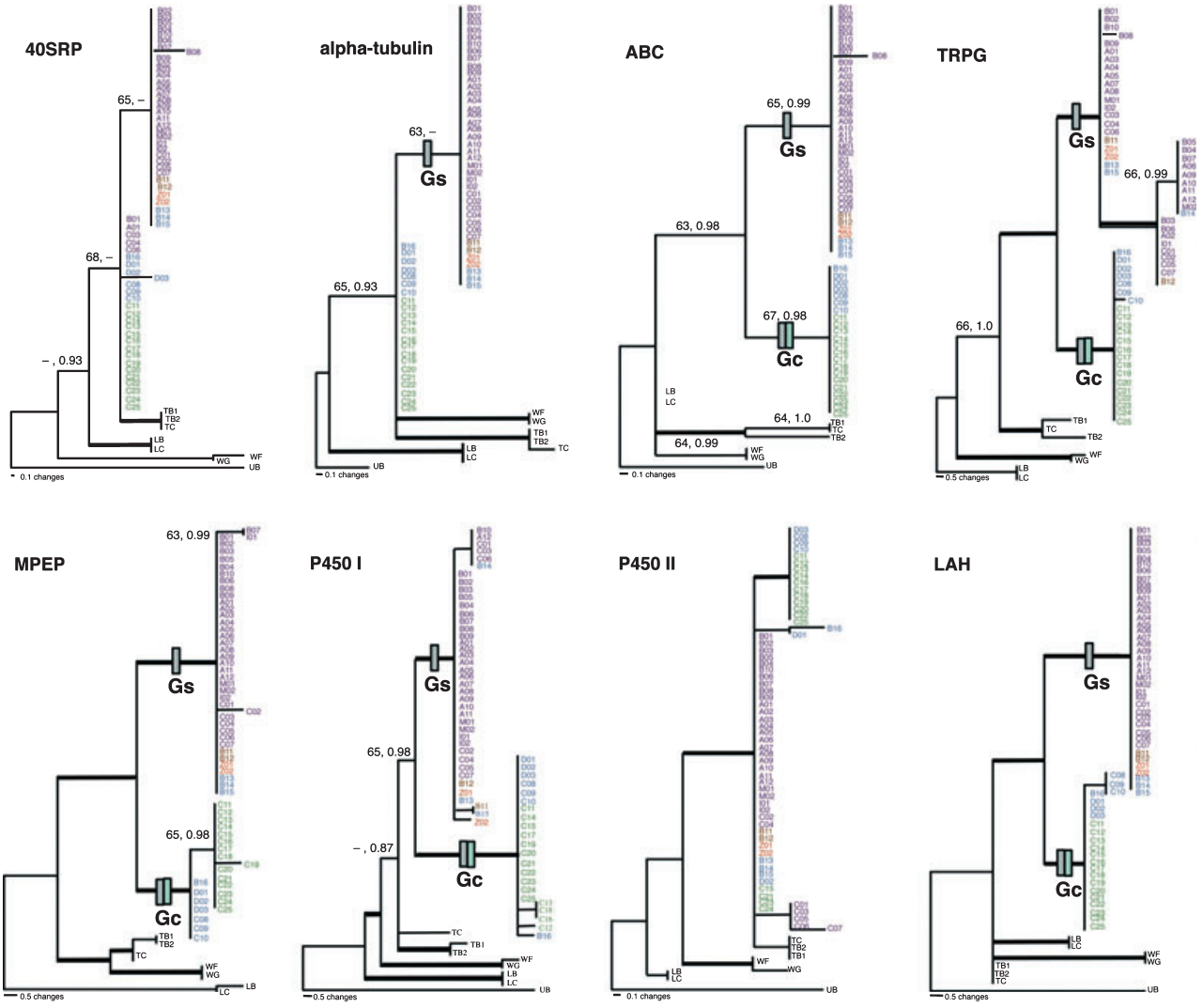

Fig. 1 Single-locus phylogenies of 15 genes studied in *G. clavigera* and its four closely related species. Bootstrap (BS > 50) and posterior probabilities (PP > 0.8) from MP and Bayesian analyses are shown along the branches. Thick branches indicate nodes with PP ≥ 95 and BS ≥ 70. The two bars indicate the *G. clavigera* monophyletic clades colour-coded according to their beetle associates: mountain pine beetle (grey) and jeffrey pine beetle (green). Trees are rooted with *G. aurea*, except for TRPG and MPEP that miss the outgroup taxon and therefore are midpoint rooted. Refer to Fig. 2a and Table 1 for colour codes and labels.

stein 1985). Bayesian analyses were run using MrBayes 3.2 (Ronquist & Huelsenbeck 2003), under the best-fit substitution model. Each run consisted of four incrementally heated Markov chains, with default heating values. The chains were initiated from a random tree and were run for 2 million generations with sampling every 1000 generations. Posterior probabilities (PP) were inferred with a 50% majority-rule consensus tree sampled after the likelihood scores had converged. The 15 nuclear genes were concatenated to conduct partitioned maximum likelihood (ML) analysis (with 1000 nonparametric replicates bootstrap) using RAxML-VI-HPC 7.0.4 (Stamatakis 2006) and partitioned Bayesian analysis. The partitioned-ML and Bayesian analysis utilized the substitution models selected by the AIC in JModelTest for each gene locus. The combined data set was also analysed with weighted parsimony, with the weighting inversely proportional to the number of parsimony informative characters at each locus. Weighting allowed each locus to contribute equally to the combined data tree. All phylogenetic trees were rooted using *G. aurea* as outgroup (Massoumi Alamouti *et al.* 2007). Monophylies supported by both BS ≥70% and PP ≥95% were considered significant.

Constraints on topologies were applied in PAUP*, and the Wilcoxon signed-rank (WSR; Templeton 1983) test was employed to assess significant differences among topologies. For this test, up to 100 MPTs recov-

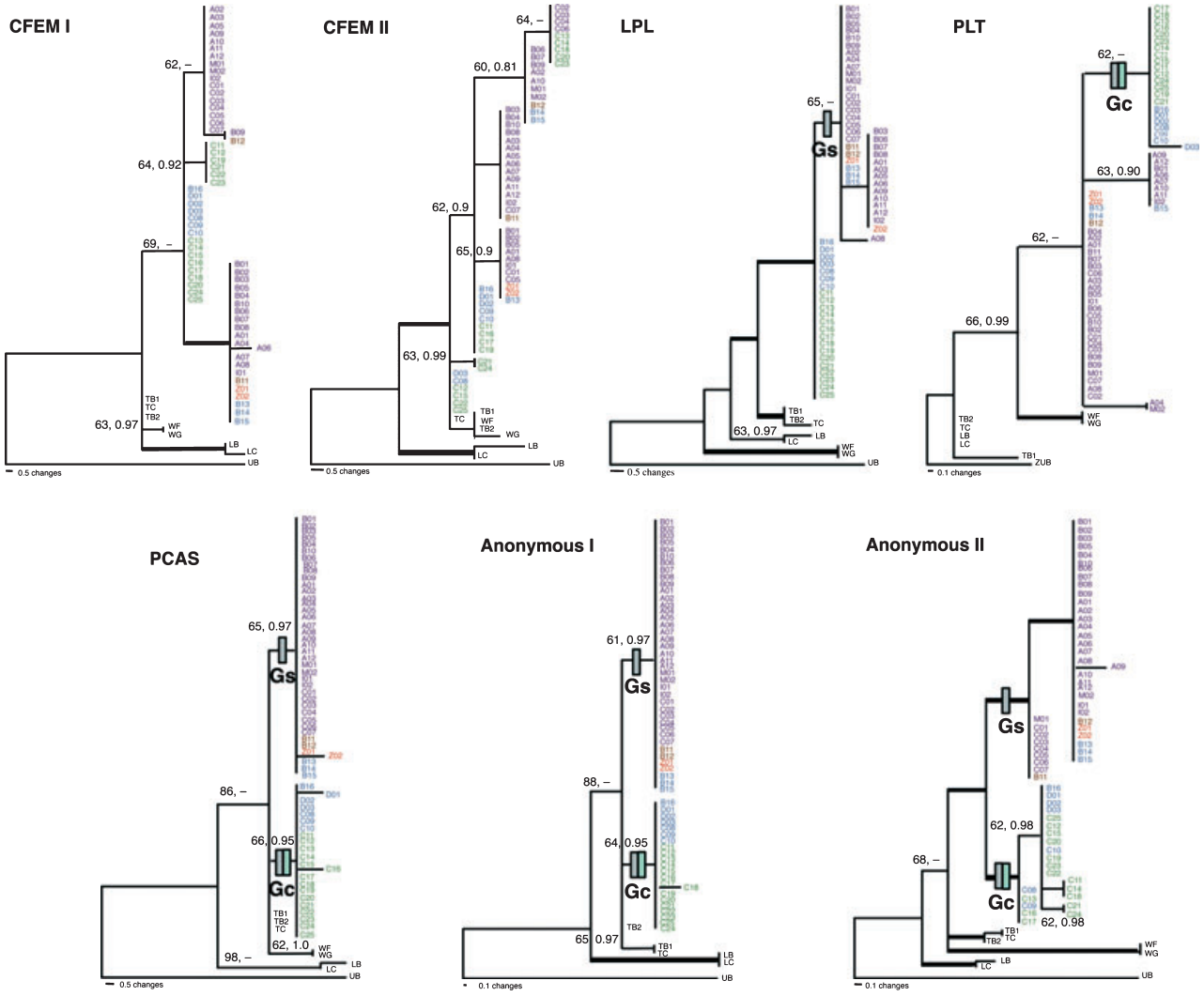

Fig. 1 (Continued).

ered were used as constraint topologies. When testing the constraint of lineage-specific monophyly, the lack of significance in the WSR tests indicates that nonmonophyly could be the result of insufficient phylogenetic signal.

Network approaches and evidence for recombination in G. clavigera

For each of the 15 gene data sets, we generated parsimony networks of *G. clavigera* haplotypes, which is described in the supporting information Appendix S1. We applied three approaches to detect the presence or absence of recombination in *G. clavigera*. First, we applied the index of association (I_A) to estimate the extent of clonality in *G. clavigera*, using the program Multilocus 1.3b (Agapow & Burt 2001). I_A determines to what extent individuals that are the same at one locus are more likely than random to be the same at other loci. We used 10 000 randomizations on the subset of polymorphic sites that showed the most balanced distribution of alleles in each gene (i.e., excluding the uninformative sites). The test assumes an infinite amount of recombination, so significant departure ($P < 0.05$) from simulated recombined data sets suggests the presence of clonality (Maynard Smith *et al.* 1993). Second, we used the NeighborNet algorithm (Bryant & Moulton 2004) for decomposition analysis with SplitsTree 4.10 (Huson & Bryant 2006) to visualize the incongruence generated by recombination from the pairwise distance matrix of the *G. clavigera* concatenated sequence data set estimated under the GTR model. Third, we estimated the pairwise homoplasy index (PHI; Bruen *et al.* 2006) in SplitsTree. Using a 100-bp window, compatibil-

Table 2 Primer sequences and gene descriptions for loci used in the phylogenetic and population genetic analysis

Primer sequence 5′ → 3′			*G. clavigera* sequence length in bp*				
Forward	Reverse	Gene description (abbreviation)	Total	Exon	Intron	UTR	GenBank accession no.
TCACGCCCACCGTTACCGACA	TGGAAATGGTCGGTGCCGAGGT	40S ribosomal protein S3 (40SRP)	742	585	0	157	HQ633911 – HQ633980
TCCAGACGAACCTGGTGCCGT	CAGGCGTCATCGAGCAAGCGA	alpha-tubulin	640	489	59	92	HQ633073 – HQ633142
ATGTGCAGGGTGGCGAGCGAA	GAATACCGCTCCGCTCGCACA	ATP-binding-cassette multidrug transporter (ABC)	549	549	0	0	HQ633143 – HQ633212
TGATTCGACTTTCCCCCT	CGTCGAACACAAACTCCT	Anthranilate synthase (TRPG)	1925	1925	0	0	HQ633981 – HQ634049
GGAGTTTGTGTTCGACGAG	GAATGACAAGGCTATGAAGGGA						
TAAGGAAAGGGAGGGCGGT	TGGGTGCGTGATGAGCGA	Metallo-peptidase (MPEP)	1672	1386	215	71	HQ634050 – HQ634118
ATTCCCCTCCCCTACTCC	CTTCCATGTCCTCCTTCC						
GACATTGTAGAGGGCAGC	AGATGGGAGGTTGGAGAG	Cytochrome P450 (P450 I)	1596	1440	113	43	HQ633213 – HQ633282
AGTAGAACACCGCCGACAG	CCGACCAAACACACCGCA						
TGCAGCAATGGGACCGGATGA	TCGTCACGTTCTCCCAGCGCT	P450 II	710	710	0	0	HQ633283 – HQ633352
CACACGGACCAACGACGA	CTCTCCTGCCCCTCTTCTC	Lipid acyl hydrolase (LAH)	1123	1123	0	0	HQ633353 – HQ633422
CTCTTCTTTGCCGGCCTTGCTGT	CGCAACGCAAACGCCAGAAGA	Fungal extracellular membrane protein (CFEM I)	667	510	58	99	HQ633423 – HQ633492
GCGTCCATTGATCGGCGTGATGT	AACCGCCAACATGGCAACGG	CFEM II	491	427	64	0	HQ633493 – HQ633562
TGCTGTCGAGAACTGGAGGCGT	CGGCAGGACCTGGAACAGGAA	Lysophospholipase (LPL)	568	443	125	0	HQ633563 – HQ633632
CGGTCGCCCGCTCTACATTGA	CTCAGCCTCTAAGCCGTTGCCT	Phosphatidylinositol transferase (PLT)	570	570	0	0	HQ633633 – HQ633702
TGCCGACAAGGTGGCCAAGTTC	GCGCAGCGCAACATTGACGACT	Peroxisomal-coenzyme A synthetase (PCAS)	685	117	24	544	HQ633703 – HQ633772
CACGACGACGAACTCCTCTCCCA	CAGGATGCCCTCGGCCTCTAAC	Anonymous I	455	296	3	156	HQ633773 – HQ633840
ACGCCGGCAAGACCTACACCA	TGCCAGACTGGTCCACATCTGCA	Anonymous II	805	240	61	504	HQ633841 – HQ633910
			13 198	10 810	722	1666	HQ633073 – HQ634118

*Base pair.

ity among sites was calculated and, assuming no recombination, significance was determined with a permutation test.

Results

Polymorphism discovery

Sixty-seven loci, which represented 50 nuclear protein-coding genes with predicted functions, were sequenced and screened for polymorphisms (Table S2, Supporting information). Some genes were constitutively expressed, e.g., housekeeping genes; others were differentially expressed in specific growth conditions and were potentially involved in growth, metabolic processes or host-tree pathogenicity (DiGuistini *et al.* 2007, 2009, 2011). A number of genes lacked significant homology with proteins or domains of known functions. We obtained approximately 50 kb of high-quality sequence data for these genes in nine *G. clavigera* isolates (http://purl.org/phylo/treebase/phylows/study/TB2:S11355) and identified 128 polymorphic sites (i.e., substitutions) across 33 genes. The majority of variations (63%), i.e., 81 single nucleotide polymorphisms (SNP) in 31 genes, separated the seven isolates representing the MPB associates at the epidemic sites from the two other isolates: *G. clavigera* holotype (ATCC 18086) and JPB associate (DLS1575). A subset of 18 informative (i.e., shared by two or more isolates) SNP (14%) in 12 genes were exclusive polymorphisms that segregated only within the seven epidemic isolates. The rest of polymorphisms were substitutions that were unique to one isolate (i.e., singletons).

Polymorphism validation

For further analysis, we selected 15 genes (Table 2) that showed different levels and patterns of variation in the polymorphism-discovery panel and sequenced them in an additional 53 *G. clavigera* isolates (Table 1). These isolates were selected to represent the beetle associates MPB and JPB, and their respective primary host trees *P. contorta* and *P. jeffreyi*, as well as a few other MPB pine-host species. Within *G. clavigera* isolates, we identified a total of 86/13 198 (0.65%) base substitutions and two indels in the concatenated 15-gene data set. No site had more than two alleles (biallelic). The most polymorphic gene region was CFEM II ($\pi = 0.0039$), and the least polymorphic was alpha-tubulin ($\pi = 0.00073$). Of the 86 polymorphic sites, 68 were informative and 18 were singletons. Eighteen of the changes were predicted in noncoding locations (i.e., intronic and UTR), and, for the 68 that were in coding regions, 33 were synonymous and 35 were nonsynonymous. The vast majority of variants were identified either as fixed SNPs ($n = 33$) or as exclusive polymorphisms ($n = 49$) that segregate only within one of the two potential *G. clavigera* lineages. The remaining four SNPs were the only shared polymorphisms found. The number of sites in the different classes of polymorphisms for each gene as well as for the concatenated data set is shown in Tables 3 and S3, Supporting information.

Single-gene phylogenies, phylogenetic species recognition and concatenated phylogeny

Using 15 gene phylogenies, we assessed *G. clavigera* species boundaries and phylogenetic relationships with related pine-infesting fungi: *G. aurea, L. longiclavatum, L. terebrantis* and *L. wingfieldii*. A summary of the phylogenetic data and model parameters inferred for each locus and the combined data set are presented in Table 4. The target genes were amplified in all species, except for TRPG and MPEP in the outgroup taxon *G. aurea* and anonymous I in *L. wingfieldii*. MP and Bayesian consensus trees inferred similar topologies that are only shown for MP trees (Fig. 1). MP analyses yielded one to five trees for each locus, which mainly differed in the branching orders of two close relatives *L. terebrantis* and *L. wingfieldii*. The majority of gene trees (10/15) resolved the pathogen *G. clavigera* into two distinct clades. We referred to these clades as Gs with 40 isolates and Gc with 22 isolates (Fig. 1). The Gs clade contained all isolates from epidemic MPBs, as well as those from localized populations except for those collected from *P. ponderosa* trees. This clade was significantly (BS ≥70% and PP ≥0.95) supported by five loci (TRPG, MPEP, P450-I, LAH, anonymous II). The Gc clade encompassed all JPB associates, as well as isolates from MPB that were infesting *P. ponderosa* trees in sympatric (California) and allopatric (South Dakota) regions. The *G. clavigera* holotype (ATCC18086; Robinson-Jeffrey & Davidson 1968) was also placed within Gc. This clade was significantly supported by the same subset of loci that supported the Gs clade. Clades in gene trees that did not agree with this partitioning were either not fully resolved (Fig. 1 40SRP, P450-II) and/or not significantly supported (Fig. 1 CFEM II: PP ≤0.95 and/or BS ≤70%). While one additional group showed a high level of support (Fig. 1: BS = 100% and PP = 1.0) in the TRPG and another in the CFEM I (Fig. 1: BS = 75% and PP = 0.95) phylogenies, we considered neither clade to be an independent lineage, because their partitions contradicted each other and neither was supported in the concatenated phylogeny (Fig. 2b). In the constraint analysis forcing the monophyly of Gs and Gc, WSR results were significant ($P = 0.04$) for only CFEM II, indicating that incongruence from the constraint phylogeny is only significant in 1/15 of the loci.

Table 3 Fixed and shared polymorphisms between the two monophyletic clades in *G. clavigera*

	Fixed polymorphisms				Shared polymorphisms			Exclusive polymorphisms to		Genetic differentiation*	
Locus	Total	Noncoding	Synonymous	Replacement	Total	SNPs	Indels (base pair)	Gs	Gc	Dxy (10^{-3})	SD of Dxy (10^{-4})
40SRP	0	0	0	0	0	0	0	2	1	1.27	3.4
alpha-tubulin	1	0	1	0	0	0	0	0	0	1.56	4.1
ABC	2	0	0	2	0	0	0	1	0	3.69	9.4
TRPG	7	0	3	4	0	0	0	6	1	4.79	9.5
MPEP	4	2	0	2	0	0	0	2	2	3.44	7.8
P450 I	6	3	2	1	0	0	0	3	3	4.48	9.3
P450 II	0	0	0	0	0	0	0	2	3	1.36	3.6
LAH	4	0	0	4	0	0	0	0	1	3.33	8.7
CFEM I	0	0	0	0	0	0	0	5	1	2.81	5.6
CFEM II	0	0	0	0	3	3	0	1	2	4.18	6.9
LPL	1	0	0	1	0	0	0	2	0	2.43	5.5
PLT	1	0	1	0	0	0	0	2	1	2.32	5.1
PCAS	2	2	0	0	0	0	0	1	2	3.10	7.7
Anonymous I	2	0	2	0	0	0	0	0	1	4.51	11.6
Anonymous II	3	3	0	0	1	1	0	2	2	7.00	14.2
Concatenated data set	33	10	9	14	4	4	0	29	20	3.71	5.7

*Dxy, net nucleotide divergence for the pairwise comparison of the two monophyletic clades in *G. clavigera;* SD, standard deviation.

Table 4 Information on phylogenetic data set sequenced from *G. clavigera* and its close relatives

Locus	Sample size	Total number of characters	Variable sites	Parsimony informative characters (PI)	Number of tree steps (TS)	Number of MP trees	Homoplasy level		Nucleotide substitution model
							PI/TS	CI	
40SRP	70	742	23	11	23	1	2.09	1.00	TrN
alpha-tubulin	70	640	10	8	10	1	1.25	1.00	HKY
ABC	70	549	10	6	10	1	1.67	1.00	HKY
TRPG	69	1925	37	33	45	3	1.36	0.89	TPM1uf + I
MPEP	69	1672	30	28	31	2	1.11	0.96	HKY + I
P450 I	70	1597	53	29	55	4	1.90	0.98	TrN
P450 II	70	710	17	8	18	3	2.25	0.94	TRN
LAH	70	1123	25	15	26	2	1.73	1.00	TPM1uf + I
CFEM I	70	673	31	12	33	4	2.75	0.96	TIM1 + I
CFEM II	70	491	27	12	29	5	2.42	0.93	TIM1 + I
LPL	70	569	28	13	30	2	2.31	0.97	HKY + I
PLT	70	570	10	6	10	1	1.67	1.00	HKY
PCAS	70	706	27	9	27	1	3.00	1.00	HKY
Anonymous I	68	467	22	9	22	1	2.44	1.00	HKY
Anonymous II	70	805	46	27	47	2	1.74	0.98	GTR
Concatenated data set	70	13 239	402	226	503	100	2.23	0.81	GTR

Information with the outgroup taxon.

Of the 15 genes, TRPG, MPEP and anonymous II showed the highest resolving power for species boundaries, supporting five monophyletic groups: Gs, Gc, *L. terebrantis*, *L. wingfieldii* and *L. longiclavatum*. While species-level clades were strongly supported by a number of single-gene phylogenies, relationships between species were difficult to resolve. For example, *L. terebrantis* showed a nonrobust phylogenetic placement among trees, and it was collapsed into polytomy in at least ten single-gene phylogenies. Although positioning of some ingroup taxa varied among gene trees, TRPG, MPEP, LAH and anonymous II significantly supported a sister-group relationship between Gs and Gc.

The concatenated matrix of 15 gene sequences (TB2:S11355) resulted in 13 239 bp of aligned nucleotide positions, 402 variable sites and 226 informative characters (Table 4). MP, partitioned ML and partitioned Bayesian analyses resulted in similar topologies that had only minor differences in the placement of terminal taxa (Fig. 2b, ML tree). The topology of concatenated phylogeny was consistent with the single-gene tree partitions resolving the *G. clavigera* isolates into two monophyletic clades, and with the sister-group relationship between Gs and Gc (MP and ML BS = 100%, PP = 1.0).

Finally, we challenged our phylogenetic results by testing whether the polymorphism distribution of *G. clavigera* into two groups was because of independent evolutionary histories or to random sorting of genetic variations. The probability of observing different groups that, by chance, do not share polymorphisms was tested by random shuffling the 15 data set across (nonpartitioned data set) and within (partitioned data set) the two phylogenetic species. For the randomization, the association of polymorphic sites within each gene was left intact (i.e., each gene was randomized as blocks). In 1000 such randomizations, we found no partition that would create groups with no shared polymorphisms. Among 62 *G. clavigera* isolates, the shortest trees acquired from the nonpartitioned, randomized data set were significantly ($P < 0.001$) longer (510–570 steps) than trees obtained from the randomized data set considering the Gs and Gc partitions (61–185 steps).

Within Gs ($n = 40$), we identified 36 distinct haplotypes that were characterized by 33 base substitutions across 12 polymorphic genes (Table S3, Fig. S1, Supporting information). Of the 33/13 198 (0.25%) polymorphic sites, 23 were informative. The number of haplotypes ranged from 1 to 4 among the genes. Gene/haplotype diversity (H) ranged from 0.0 to 0.73 in CFEM II. The diversity over the combined data set showed a high value of 0.99; however, genetic differentiation within the isolates was low, resulting in minor nucleotide diversity ($\pi = 0.00068$). Similar haplotypes did not cluster based on either geographic locations or the pine-host species (Fig. S1, Supporting information). Within Gs, we found seven isolates representing three identical haplotypes, two from the same (BC and Alberta) and one from different (BC and Arizona) localities. The probability of identical haplotypes (i.e., isolates sharing the same sequence type at all 12 poly-

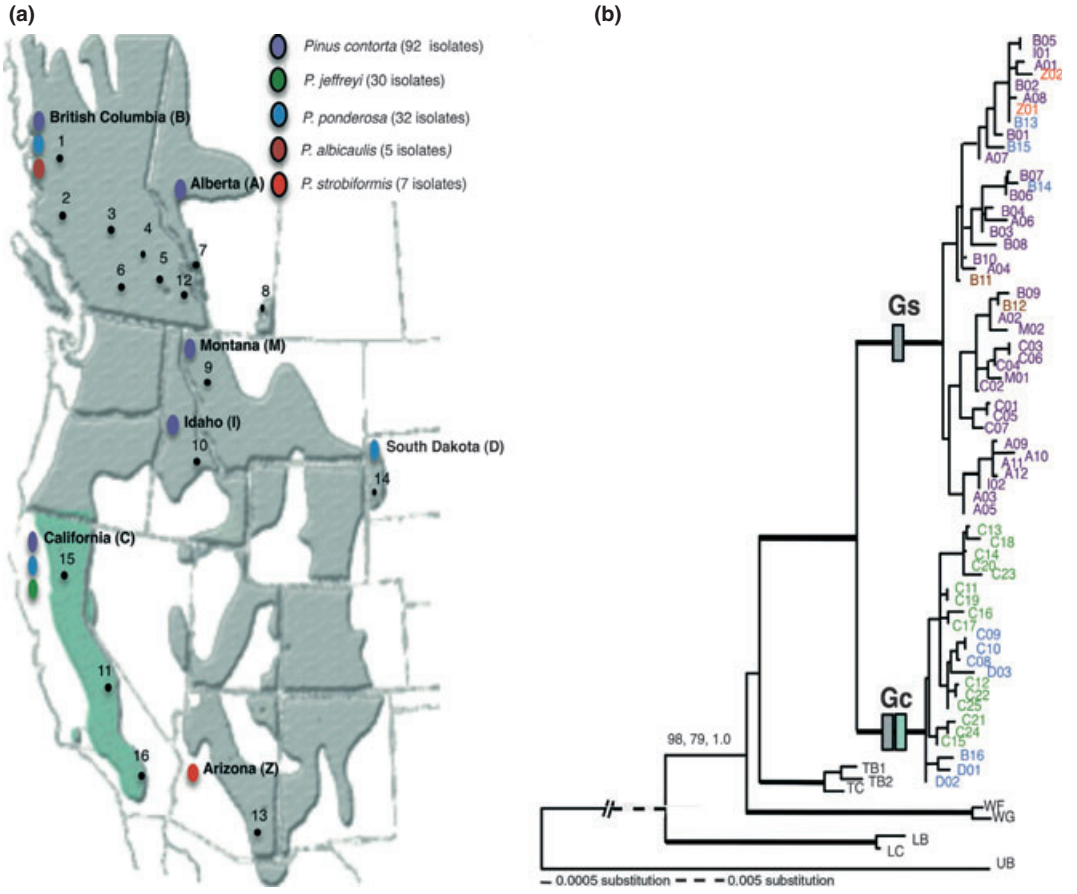

Fig. 2 a) Map of western North America showing fungal collection sites where only one [mountain pine beetle (MPB):grey] or two [MPB & jeffrey pine beetle (JPB): green] *G. clavigera* beetle associates are present. Host-tree species are colour-coded, and the number of fungal isolates from each tree species is shown in parentheses. b) ML analysis of 15-gene combined data set showing how the species recognized by phylogenetic species recognition are related to each other and to other closely related species. Thick branches indicate nodes with ML and MP BS values of 100 and the Bayesian PP of 1.0. Gs and Gc monophyletic clades are labelled with bars colour-coded according to beetle associates: mountain pine beetle (grey) and JPB (green). Letters indicate the collection localities, and colours indicate host-tree species corresponding to the map and Table 1. Dashed line indicates an adjustment of scale.

morphic loci) by recombination was small (4.2×10^{-3}–1.8×10^{-6}), suggesting that they represent epidemic clones from the asexual reproductions. In comparison, the Gc isolates ($n = 22$) showed a similar pattern but with a slightly lower level of nucleotide diversity (Table S3, Fig. S1, Supporting information). They represented a collection of 22 unique haplotypes (i.e., H = 1.0) that consisted of a total of 24 base substitutions across 12 polymorphic genes. Of the 24/13 198 (0.18%) polymorphic sites, 14 were informative. As was the case for the Gs group, CFEM II showed the highest level of both haplotype and nucleotide diversity. However, some genes that showed a higher level of variation within Gs (e.g., TRPG, LPL, PLT) showed almost no polymorphisms in Gc. Haplotypes did not correlate with the host beetle/tree species, except for one allele in MPEP that was only found for MPB/*P. ponderosa* associates; however, this partition was not statistically supported.

Evidence of recombination

For Gs, we evaluated I_A for all isolates, as well as for the reduced-by-haplotype data set in which we excluded identical haplotypes. When all isolates were included ($n = 40$), the $I_A = 1.6$ was significantly ($P = 0.007$) different from the values obtained for the simulated recombined data set, leading us to reject the null hypothesis of recombination. However, the $I_A = 1.3$ for the unique haplotypes ($n = 36$) was indistinguishable ($P = 0.09$) from that expected for a recombinant population (Fig. 3b–Gs). Within Gc, the I_A–Gs 0.5 ($P = 0.3$) also suggested recombination, both when all 22 isolates were included or only those from JPB

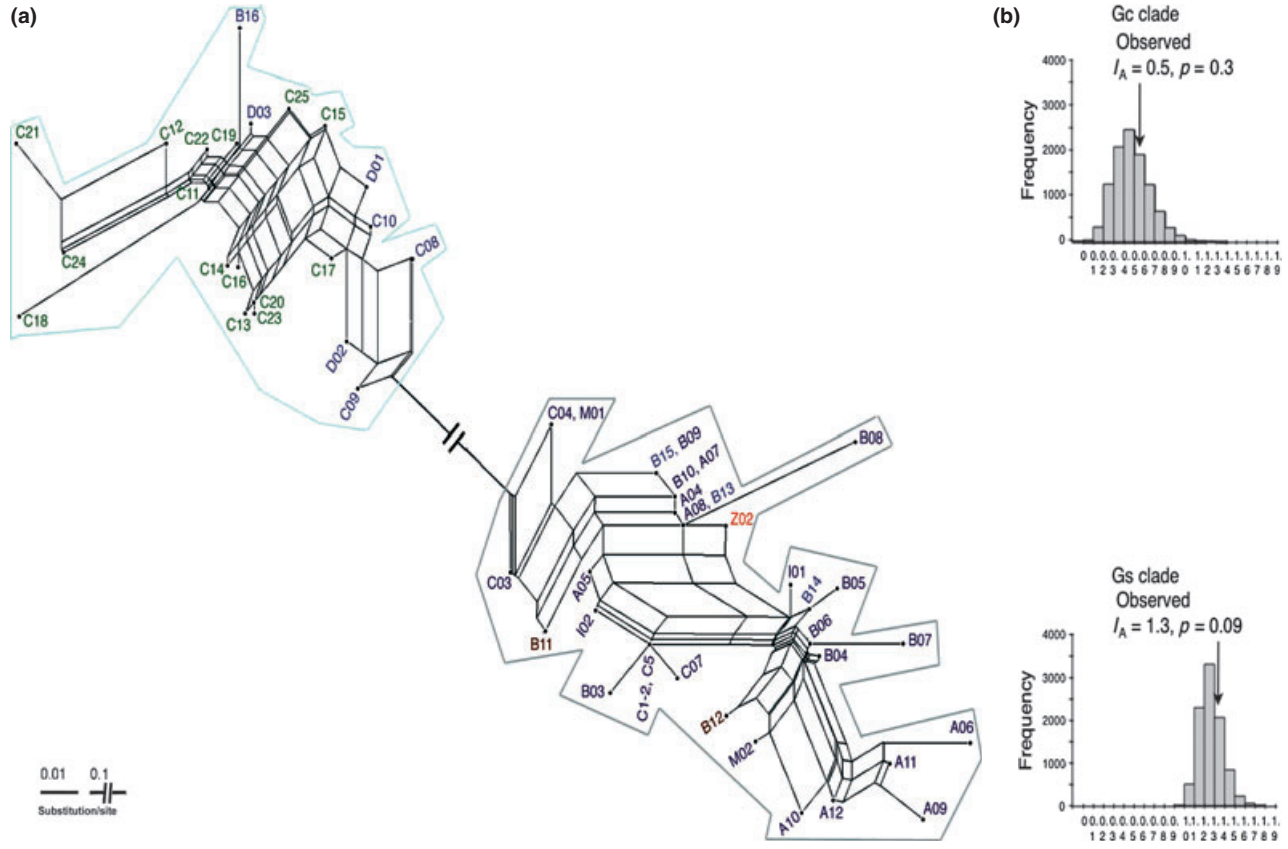

Fig. 3 Recombination analysis. a) Split decomposition analysis of the 15-gene combined data set. Coloured boxes represent the two *G. clavigera* monophyletic clades: Gc (green box) and Gs (grey box). The interconnected networks are suggestive of recombination within both Gc and Gs clades. The labels refer to *G. clavigera* isolates listed in Table 1. b) The I_A values for Gs and Gc clades are shown and compared against histograms of I_A values for 10 000 simulated recombined data set.

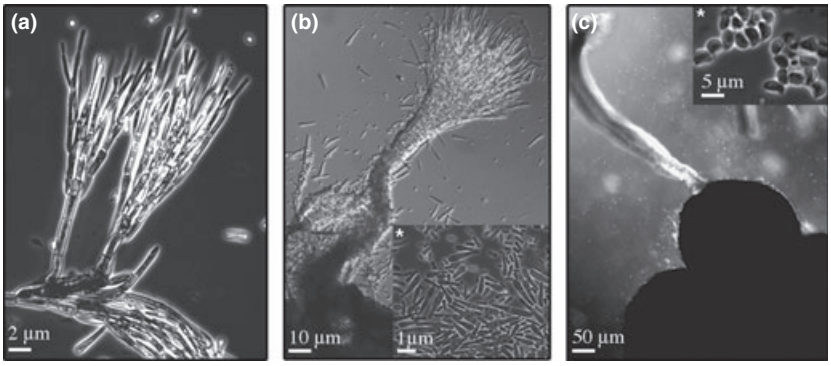

Fig. 4 Reproductive structures in Gs. Light micrographs of asexual stage characterized with mononematous (a) and synnematous (b) conidiophores reproducing conidia (*). Light micrograph of sexual structure (c) characterized by a spherical ascocarp oozing ascospores (*).

(Fig. 3b–Gc). Split decomposition analysis also provided evidence for network relationships, giving a graphical support for the presence of recombination within both Gs and Gc (Fig. 3a). Finally, PHI provided another significant evidence (P = 0.00006) of recombination.

Ecological and morphological characteristics

To assess the host and distribution ranges of the *G. clavigera* lineages in more detail, we sequenced a single informative locus, PCAS (Table 2), in an additional 104 isolates (Table S1 in Supporting information, Fig. 2a).

Locus PCAS contains two fixed SNPs (2/685 = 0.29%) that differentiate the two *G. clavigera* lineages and possess exclusive polymorphisms that are not shared between the two fungi. This locus has been tested against a large number of other closely related species and has been used as target-specific PCR-primers to detect and differentiate microbial communities associated with the MPB (Khadempour *et al.* 2010). We generated the data for *P. contorta*-associated isolates (n = 67) from BC, Alberta, Montana and Idaho, as well as isolates from *P. ponderosa* trees in BC (n = 13). Consistent with results from 15-gene phylogenies, the deeper single-locus sampling showed that the Gc group was not present in the epidemic populations of MPB; instead, this fungus largely represented isolates from *P. ponderosa* (n = 18) and *P. jeffreyi* (n = 30) trees attacked by the localized populations of respective beetle associates MPB and JPB in South Dakota and California. In contrast, Gs (n = 117) occurred on MPB in epidemic populations of the beetle and its pine-host species in western Canada and the USA, as well as in localized populations infesting *P. contorta* in California and *P. strobiformis* in Arizona. Both Gs and Gc were found in MPB and JPB localized populations in California, where the two beetle associates live in sympatry on *P. contorta, P. jeffreyi* and *P. ponderosa*. In South Dakota, where the localized population of MPB infests *P. ponderosa* trees, we found Gc (n = 15) but no evidence of Gs.

We compared the reproductive structures of five isolates representing the Gs group to those of the *G. clavigera* holotype, which was included in our analysis and represented the Gc group. The anamorph (conidia and conidiophore) and teleomorph (i.e., ascocarp and ascospores) morphologies of Gs (Fig. 4) representatives agreed with the formal *G. clavigera* holotype description by Robinson-Jeffrey & Davidson (1968). The conidiophores and conidia sizes varied among isolates, but all measurements (Table S5, Supporting information) agreed with the *G. clavigera* original descriptions (Robinson-Jeffrey & Davidson 1968; Six & Paine 1997), as well as with descriptions of *P. contorta* associates (Lee *et al*. 2003).

Discussion

We generated the first comprehensive data set of protein-coding gene variability in the bark-beetle symbiont and pine pathogen *G. clavigera*. We used this data set to characterize patterns of DNA polymorphism and divergence within the pathogen and among four close relatives that also inhabit pine trees. In contrast to the current taxonomy, our results show species diversity and ecological complexity with respect to host species. Paine & Hanlon (1994) and Six & Paine (1998) showed that the *G. clavigera* isolates of JPB were more tolerant to host oleoresin than those of MPB, suggesting some potential physiological differences between these two types of isolates. Here, we suggest that the genetic divergence and diversity in *G. clavigera* isolates may have resulted from the fungus adapting to particular pine species and to extensive expansion of the epidemic.

Our phylogenetic analyses identified two distinct lineages in *G. clavigera*. While the combined data set of nuclear ribosomal DNA and the protein-coding genes have improved the phylogenetic positioning of *G. clavigera* (Lim *et al.* 2004; Roe *et al.* 2010), these loci failed to distinguish the two lineages identified in this study. These results indicate that the sequences currently available for the phylogenetic inference of ophiostomatoid fungi provide inadequate data for defining species and inferring evolutionary relationships in the genus *Grosmannia*. We demonstrated that sequencing more genomic regions is more effective for inferring species boundaries. Given this, care should be taken when interpreting ecological characteristics of this group of fungi. The literature suggests speculative evolutionary processes (Six *et al.* 2003; Lim *et al.* 2004; Roe *et al.* 2011) that rely on data that are insufficient for identifying species and on an imperfectly known phylogeny. Six *et al.* (2003) and Lim *et al.* (2004) suggested that *G. clavigera* is a recently diverged morphological variant of the generalist fungus *L. terebrantis*. Our results show that *L. terebrantis* is a distinct species separated from both *G. clavigera* lineages; we also found that some isolates assigned as *L. terebrantis* were genetically different from the *L. terebrantis* holotype isolated from *D. terebrans* (Six DL & Massoumi Alamouti S, unpublished data), suggesting that this fungus represents a complex of closely related species that need to be taxonomically and ecologically re-assessed.

Below, we provide two main lines of evidence to show that *G. clavigera* lineages represent two distinct species: (i) they are evolutionary independent and (ii) they are ecologically distinguishable. Because a lineage can represent a species, a clone or a divergent group within a population, we will discuss these two concepts and discuss the evidence of recombination and ecological significance in each species. Current concepts agree that species correspond to 'segments of separately evolving lineages' (De Queiroz 2007); however, different characteristics (e.g., morphological, reproductive and nucleotide divergence) are used to infer boundaries for species, clones and divergent groups. Such characteristics do not arise at the same time during the process of speciation, and so each type of evidence can lead to different conclusions regarding species boundaries (Avise 2004).

Evolutionarily independent lineages

Phylogenetic species recognition by genealogical concordance (Taylor *et al.* 2000; Dettman *et al.* 2003) stipulates that when lineages are separated for long periods of time relative to population size, genealogies from the majority of loci should be congruent. This criterion considers a clade to be an independent evolutionary lineage and a phylogenetic species if it is present in the majority of single-locus phylogenies (Dettman *et al.* 2003). Here, the concordance of ten genealogies defines *G. clavigera* lineages as two sibling phylogenetic species and suggests genetic isolation—even when the lineages occur in the same geographic region, as in California.

In general, the pattern of gene genealogies and the level of polymorphism depend on the timing of speciation event, historical population sizes, mode of reproduction, extent of hybridization and natural selection (Avise 2004). Enforcing topological constrains for the monophyly of Gs and Gs showed that only one (CFEM II) of the 15 genomic regions exhibits significant (at the $\alpha = 0.05$ level) incongruent pattern, and therefore, the lack of reciprocal monophyly and/or lack of nodal supports in other loci is the result of insufficient phylogenetic signal. For CFEM II, we found no evidence for significant departure from neutrality (data not shown), intragenic recombination or paralogous in the *G. clavigera* genome's predicted gene models (DiGuistini *et al.* 2009, 2011). Therefore, none of these mechanisms can explain the incongruent pattern. Introgression can occur when interspecific hybridization results in the transfer of genetic material from one species into another, which leads to paraphyly of recipient species; alternatively, incomplete lineage sorting can result in incongruent genealogies if species divergence occurred too recently for ancestral polymorphisms to have sorted into reciprocal monophyly (Avise 2004).

Distinguishing between interspecific hybridization and lineages sorting is difficult, because both result in the same pattern of incongruence (Hey & Nielsen 2004). While we could not estimate the divergence time of *G. clavigera* lineages with certainty, because of the lack of fossils and the great variance in fungal nucleotide substitution rates (Kasuga *et al.* 2002), two observations suggest that these fungi diverged recently. First, the low interspecific nucleotide divergence ($0.0037 \pm 5.7 \times 10^{-4}$) and the unresolved species phylogeny suggest that not all loci have reached reciprocal monophyly. Second, when we compared ingroup and outgroup taxa of two or four species, a large number of ancestral polymorphisms appeared to predate divergence, consistent with the speciation event being so recent that ancestral polymorphisms were retained.

Evidence of recombination

Because classical phylogenetic trees can give only a snapshot of the actual complex relationships that can be encountered when intraspecific details are considered, we describe *G. clavigera* population structure with modified phylogenies using split decomposition analysis. In this, network relationships account for recombination within both Gs and Gc that, in agreement with our gene phylogenies, are separated into two distinct groups. I_A values not significantly different from artificially recombined data sets, and large numbers of unique sequence types suggest that recombination within each phylogenetic species occurred frequently enough to create many different combinations of alleles. While these results can also be explained by convergent or parallel mutations, the very low sequence divergence and lack of multiple alleles observed for each polymorphic site, even when compared against other close relatives, indicate that the most likely explanation is recombination.

We also observed direct evidence for clonal propagation in Gs with the occurrence of the same haplotype over a wide geographic area. In this species, applying the I_A test for all isolates and for the reduced-by-haplotype data set suggested the existence of epidemic clonality (Maynard Smith *et al.* 1993). Overall, the recombination component appears greater in Gc (100% unique haplotypes and lower I_A); however, concordant with the fungal asexual reproductions in natural environments (Six & Paine 1997), $I_A > 0$ still suggests some deviation from complete panmixia.

These results agree with the genomic analysis of *G. clavigera sensu lato*; both suggested that this fungus is a heterothallic sexual species (Tsui *et al.* 2009; DiGuistini *et al.* 2011). Consistent with this, *G. clavigera* ascocarps have been occasionally reported at epidemic sites in 1-year-old MPB galleries (Robinson-Jeffrey & Davidson 1968; Lee *et al.* 2003). There has been no report of either the sexual state of *G. clavigera* associated with JPB, or of their sexual reproduction under experimental conditions. Our morphological observations, finding the teleomorph in epidemic MPB galleries, and molecular results suggest that recombination is ongoing. However, sexual reproduction seems to occur in older galleries when competition and predation increases and when environmental variables change. The asexual state is abundant in the galleries and pupal chambers during the active life cycle of the two beetles; as well it is abundant on artificial media used for fungal isolations. Systematic investigations with more isolates from different phases of the beetle cycles would allow differentiating the relative contribution of clonal versus sexual reproductive modes in these fungi.

Ecologically distinguishable

Evidence for host-specific differentiation between the two *G. clavigera* lineages is as follows. While we expected that the Gc and Gs would be specific to beetle vectors, our ecological data indicate that one lineage (Gc) occurs on both beetle vectors (MPB and JPB) infesting respective host trees *P. jeffreyi* and *P. ponderosa*, whereas the other (Gs) is exclusively associated with MPB. Gc was only isolated from two geographically distinct and localized US populations, one of which was populated with *P. contorta*, *P. jeffreyi* and *P. ponderosa* and the other only with *P. ponderosa*. In contrast, Gs was associated with epidemic and localized populations of MPB inhabiting *P. contorta*, as well as other pine species, but not with *P. jeffreyi* and the localized *P. ponderosa* supporting the Gc clade. Further, our phylogenetic data showed that *G. clavigera* from the same host species in different geographic areas are genetically closer than those collected from different host species occurring in the same geographic region (e.g., California). While our data in some geographic areas were limited, preventing us from assessing the role of geographical isolation in speciation, overall, the data indicate that both a beetle vector's preference for a host-tree species and the geographic isolation of the host species can contribute to progressive differentiation of the vectored fungal species.

Grosmannia clavigera lineages develop all phases of their life cycles on host trees and are dispersed by their respective beetle vectors via a specific association (Harrington 2005). Between beetle generations, these fungi are protected and maintained inside the specialized beetle structures called mycangia. Given this, the fates of the mutualistic fungus and beetle partners are linked, and mating is more likely to occur between fungi within the specific host tree. Such a degree of inherent isolation has been suggested to facilitate adaptive differentiation in a large number of fungal plant pathogens recognized as complexes of specialized sibling species (Giraud *et al.* 2006). The frequent asexual reproduction and sexual recombination in fungi can also promote ecological divergence by creating new combinations of alleles and rapid reproduction of those combinations that favour host adaptation (Giraud *et al.* 2010).

During the early phases of a massive attack by a beetle—fungal complex, healthy standing pine trees release constitutive or induce defence chemicals such as oleoresin (Boone *et al.* 2011). To survive in such hostile and toxic environments, beetle—fungal complexes must have mechanisms for modifying or metabolizing tree defence compounds (DiGuistini *et al.* 2009, 2011). While pine species have similar chemical defence systems, there are quantitative and specific chemical differences among pine species and even between populations of the same host species (Forrest 1980; Gerson *et al.* 2009). For example, β phellandrene is the most abundant monoterpene in *P. contorta* while heptane is the major volatile chemicals in *P. jeffreyi* (Mirov & Hasbrouck 1976; Smith 2000). Heptane has been found at moderate concentration in the hybrid between *P. jeffreyi* and *P. ponderosae*, but has not been reported in *P. ponderosae*. However, tree chemical data are limited, especially for *P. ponderosae*, which needs to be systematically characterized across its range in western North America. Given this, specific association of the fungal pathogen with a host tree may also be maintained by the ability of pathogen to overcome and adapt to a tree's chemical defence systems.

Concordant with our results showing a distinct phylogenetic separation between *P. jeffreyi* (Gc) and *P. contorta* (Gs) associates, Six & Paine (1998) showed that *G. clavigera* from *P. contorta* exhibit a poor growth in *P. jeffreyi*. They also indicated that JPB associates were tolerant to a wider range of host chemicals. These differences might be because of the pathogen adapting to tree's chemical defence compounds, e.g., β phellandrene being at higher concentration in *P. contorta* than either *P. jeffreyi* and *P. ponderosae*. Further, molecular phylogeny of *Pinus* species is concordant with the monophyly of *G. clavigera* from localized populations of *P. ponderosae* and *P. jeffreyi*, and with the separation of the Gc from these two pine species from the Gs of *P. contorta*. *P. jeffreyi* and *P. ponderosae* are genetically and morphologically close relatives; they can hybridize and are classified in the *Pinus* subsection *Ponderosa* (Gernandt *et al.* 2009), while *P. contorta* is phylogenetically distinct and is classified in the subsection *Contorta* (Krupkin *et al.* 1996). Similarly, MPB genetic divergence related to host trees has been also reported (Stock & Amman 1980; Stock *et al.* 1984; Sturgeon & Mitton 1986; Langor *et al.* 1990; Kelley *et al.* 2000). Hopkins (1909) described MPB as two species, *D. ponderosae* and *D. monticola*. Although these species were synonymized by Wood (1982), they attack and breed in different pine species (Stock *et al.* 1984). Genetic studies using allozyme and AFLP markers have reported contradictory results: host-dependent (e.g., *P. contorta* vs. *P. ponderosa*) differentiation between localized beetle populations for allozymes and no host-dependent differences between MPB populations for AFLP (Mock *et al.* 2007). However, tree species and geographic areas vary between these studies, and it will be necessary to sample additional populations in the eastern and southern portion of the MPB range, and from different host trees including *P. contorta*, *P. ponderosa* and *P. flexilis* to resolve these contradictory results.

While MPB can attack and breed in different pine species, it is important to note that localized populations of MPB prefer one host pine species, even when that species is intermixed with other species that MPB could colonize (Wood 1982; Langor *et al.* 1990). A combination of events may contribute to the accumulation of host-adapted genes in MPB localized populations; for example, selective pressures on developing broods imposed by different tree species, host preferences by the beetle, differences among trees and allochronic separation of beetles' emergence from different hosts (Sturgeon & Mitton 1982; Borden 1984; Langor *et al.* 1990). Localized populations are also characterized by temporary small outbreaks that are often initiated by secondary bark beetles attacking stressed trees (Smith *et al.* 2010); beetle populations in such regions may maintain a stable diversity of fungal species for extended periods of time. In contrast, epidemic populations of MPB often contain a high number of beetles relative to the preferred pine species in a given geographic range and so attack other pine species (Wood 1963; Logan & Powell 2001; Bentz *et al.* 2010; Safranyik *et al.* 2010). Furthermore, during extensive outbreaks, MPBs have been reported occasionally as attacking and reproducing in nonhost pine such as *Picea* when faced with a shortage of host trees (Huber *et al.* 2009). Consequently, the spread of epidemics, which is affected by host tree's susceptibility, availability and continuity on large geographic regions (Safranyik *et al.* 2010), may dilute or replace older fungal populations that have become host adapted during the nonepidemic phases (Sturgeon & Mitton 1982; Langor *et al.* 1990). Such a population change was suggested by AFLP analysis of both MPB and *G. clavigera* populations (Lee *et al.* 2007; Mock *et al.* 2007; Roe *et al.* 2011). Lee *et al.* (2007) reported two genetically distinct groups of *G. clavigera* associated with *P. contorta* in the epidemic regions; the major group contains 166 individuals from BC and the Rocky Mountains, and the second group includes nine individuals from the Rocky Mountains. They suggested that the latter might represent the original population of the Rocky Mountains that was mixed with the larger group that was introduced into the region by the eastward expansion of MPB epidemic. Although representative isolates were included in our data set, we found no evidence of these two MPB-associated *G. clavigera* groups. While support of distinct lineages based on independent gene genealogies would indicate more ancient divergence among these fungi, microsatellite makers have also not supported such a distinction (Tsui *et al.* unpublished data).

Although the data from localized populations (i.e., California and South Dakota) suggested that *P. ponderosa* might not be a preferred host of the Gs lineage, this tree species was found hosting Gs in the epidemic regions (BC and Rocky mountains). This might be the result of the current rapid expansion of MPB and the pathogen (Gs) from their primary preferred host *P. contorta* to other pine species, including *P. ponderosa*. The holotype (ATCC 18086; Robinson-Jeffrey & Davidson 1968) is the only remaining isolate from *P. ponderosa*-infested trees before the current epidemic in BC. It clusters genetically with other current localized *P. ponderosa* associates, and not with Gs isolates from epidemic regions; this is consistent with the MPB rapidly expanding its population and geographic range in the epidemic. While no other historical isolates of *G. clavigera* are available, we would expect to find additional evidence for host-tree preferences among *G. clavigera* lineages by sampling populations from different infested-tree species in the eastern and southern portion of the MPB range, i.e., in areas that have not been reached by the current epidemics. If fungal lineages are adapted to host species, then lineages should correlate with host species locations; however, this assumes, simplistically (Thompson 1994), that ecological constraints or genetic structure of host beetles/trees and pathogen is the same in different geographic regions. But they are not; both the beetle and host trees vary genetically and phenotypically between different geographic regions (Krupkin *et al.* 1996; Richardson 2000; Mock *et al.* 2007; Gernandt *et al.* 2009). And there are significant chemical differences between trees at different geographic locations and with environmental conditions that need to be further characterized (Mirov 1948; Latta *et al.* 2003).

While the nomenclatural name *G. clavigera* is tied to the species that is genetically and ecologically represented by the holotype (Robinson-Jeffrey & Davidson 1968), we showed that the fungus consists of Gs and Gc lineages. These are distinct sibling species that should be recognized taxonomically. Gc should retain the name *G. clavigera*, while Gs should be described as a new species. In the future, we can anticipate that Gc genetic variation will evolve slowly while Gs might go through further genetic variation, and we outline two scenarios. In BC, we already observed a postepidemic phase in which the MBP population is decreasing, and we anticipate that this population will collapse in the near future because of the lack of mature *P. contorta*. In the first scenario, only a small number of Gs haplotypes survive the MPB collapse and are maintained through the endemic cycle of the beetle until young pine trees reach maturity. At that point, in a future outbreak, the population and the fungal genetic diversity will increase, as it did in the current epidemic, leading to an array of closely related new haplotypes. In a second scenario that is potentially a shorter-term concern, a subset of

the current large population in Alberta succeeds in becoming established in a new host-tree species and, with its fungal symbionts, adapts to the new physical and chemical environment presented by this host. There is evidence that this may already be occurring, as it has recently been shown that the beetle can successfully reproduce in the wild, in hybrids between *P. contorta* and *P. banksiana*. Significantly, *P. banksiana* occurs across the northern Canadian boreal forest. While *P. banksiana* is more closely related to *P. contorta* than to *P. ponderosa* or *P. jeffreyi*, landscape and environmental conditions prevailing in the boreal forest would lead the symbiotic partners to evolve as the MPB spread across the boreal forest. Extending the work described above could characterize how Gc and Gs populations are evolving and so help to assess threats related to the above scenarios. Even if MPB does not become established on *P. banksiana*, in the near future, climate change will affect geographic distributions of trees and beetles, and populations of fungal associates will evolve with vectors and hosts. Similar work on other MPB-host trees or other beetle systems could establish accurate species diversity and provide a foundation for understanding ecological interactions of the ophiostomatoid group that includes the most common fungal symbionts associated with bark beetles.

Acknowledgements

We thank A.G. Robertson, K. Ritland and W. Maddison for valuable discussions that improved the manuscript, B. Bentz, M. Hansen, J. Vandygriff and Willis C. Schaupp for their help with the field works, and K. Bleiker, T.C. Harrington and A. Rice for sharing their fungal cultures. This work was supported with a Strategic Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC; to JB and CB), and funds from Genome Canada, Genome BC and Genome Alberta (grant to JB, CB, RCH) in support for the Tria project (http://www.thetriaproject.ca). Salary support for JB came in part from an NSERC Steacie award and the UBC Distinguished Scholars Program S. Massoumi Alamouti was the recipient of NSERC Graduate Fellowship.

References

Agapow P, Burt A (2001) Indices of multiloci linkage disequilibrium. *Molecular Ecology Notes*, **1**, 101–102.

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change, agrotechnology drivers. *Trends in Ecology & Evolution*, **19**, 535–544.

Avise JC (2004) *Molecular Markers, Natural History, and Evolution*. Sinauer Associates, Sunderland, MA.

Bentz BJ, Régnière J, Fettig CJ *et al.* (2010) Climate change and bark beetles of the western United States and Canada: direct and indirect effects. *BioScience*, **60**, 602–613.

Bleiker K, Six DL (2007) Dietary benefits of fungal associates to an eruptive herbivore: potential implications of multiple associates on host population dynamics. *Environmental Entomology*, **36**, 1384–1396.

Boone CK, Aukema BH, Bohlmann J, Carroll AL, Raffa KF (2011) Efficacy of tree defense physiology varies with bark beetle population density: a basis for positive feedback in eruptive species. *Canadian Journal of Forest Research*, in press.

Borden JH (1984) Semiochemical-mediated aggregation and dispersal in the Coleoptera. In: *Insect Communication* (ed. Lewis T). pp. 123–149, Academic Press, London.

Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics*, **172**, 2665–2681.

Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, **21**, 255–265.

Burnett J (2003) *Fungal Population and Species*. Oxford University Press, Oxford, UK.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879–886.

Dettman JR, Jacobson DJ, Taylor JW (2003) A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote Neurospora. *Evolution*, **57**, 2703–2720.

DiGuistini S, Ralph S, Lim YW *et al.* (2007) Generation and annotation of lodgepole pine and oleoresin-induced expressed sequences from the blue-stain fungus *Ophiostoma clavigerum*. *FEMS Microbiology Letters*, **267**, 151–158.

DiGuistini S, Liao NY, Platt D *et al.* (2009) De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biology*, **10**, R94.

DiGuistini S, Wang Y, Liao NY *et al.* (2011) Genome and transcriptome analyses of the mountain pine beetle fungal symbiont Grosmannia clavigera, a lodgepole pine pathogen pathogen. *PNAS*, **108**, 2504–2509.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Forrest G (1980) Genotypic variation among native scots pine populations in Scotland based on monoterpene analysis. *Forestry*, **53**, 101–128.

Gernandt DS, Hernández LS, Salgado HE, Pérez de LR, Jorge A (2009) Phylogenetic Relationships of *Pinus* Subsection *Ponderosae* Inferred from Rapidly Evolving cpDNA Regions. *Systematic Botany*, **34**, 481–491.

Gerson EA, Kelsey RG, St Clair JB (2009) Genetic variation of piperidine alkaloids in *Pinus ponderosa*. *Annals of Botany*, **103**, 447–457.

Giraud T, Villaréal LM, Austerlitz F, Le Gac M, Lavigne C (2006) Importance of the life cycle in sympatric host race formation and speciation of pathogens. *Phytopathology*, **96**, 280–287.

Giraud T, Refrégier G, Le Gac M, de Vienne DM, Hood ME (2008) Speciation in fungi. *Fungal Genetics and Biology*, **45**, 791–802.

Giraud T, Gladieux P, Gavrilets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution*, **25**, 387–395.

Harrington T (2005) Ecology and evolution of mycophagous bark beetles and their fungal partners. In: *Ecological and Evolutionary Advances in Insect–Fungal Associations* (eds Vega F, Blackwell M). pp. 257–291, Oxford University Press, New York, NY.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hopkins AD (1909) Contributions toward a monograph of the Scolytid beetles. The Genus *Dendroctonus*. US Government Printing Office, US Bureau of Entomology Technical Series Washington. No. 17.

Huber DPW, Aukema BH, Hodgkinson RS, Lindgren BS (2009) Successful colonization, reproduction, and new generation emergence in live interior hybrid spruce *Picea engelmannii* x *glauca* by mountain pine beetle *Dendroctonus ponderosae*. *Agricultural and Forest Entomology*, **11**, 83–89.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Jacobs K, Wingfield MJ (2001) *Leptographium Species: Tree Pathogens, Insect Associates, and Agents of Blue-stain*. American Phytopathological Society Press, St. Paul, MN.

Kasuga T, White TJ, Taylor JW (2002) Estimation of nucleotide substitution rates in eurotiomycete fungi. *Molecular Biology and Evolution*, **19**, 2318–2324.

Kelley S, Farrell B (1998) Is specialization a dead end? The phylogeny of host use in Dendroctonus bark beetles (Scolytidae). *Evolution*, **52**, 1731–1743.

Kelley ST, Farrell BD, Mitton JB (2000) Effects of specialization on genetic differentiation in sister species of bark beetles. *Heredity*, **84**, 218–277.

Khadempour L, Massoumi Alamouti S, Hamelin RC, Bohlmann J, Breuil C (2010) Target-specific PCR primers can detect and differentiate ophiostomatoid fungi from microbial communities associated with the mountain pine beetle *Dendroctonus ponderosae*. *Fungal biology*, **114**, 825–833.

Knowles L, Carstens B (2007) Delimiting species without monophyletic gene trees. *Systematic Biology*, **56**, 887–895.

Krupkin AB, Liston A, Strauss SH (1996) Phylogenetic Analysis of the hard pines (*Pinus, Pinaceae*) from Chloroplast DNA Restriction Site Analysis. *American Journal of Botany*, **83**, 489–498.

Kurz WA, Dymond CC, Stinson G *et al.* (2008) Mountain pine beetle and forest carbon feedback to climate change. *Nature*, **452**, 987–990.

Langor DW, Spence JR, Pohl GR (1990) Host effects on fertility and reproductive success of *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae). *Evolution*, **44**, 609–618.

Latta RG, Linhart YB, Snyderc MA, Lundquist M (2003) Patterns of variation and correlation in the monoterpene composition of xylem oleoresin within populations of ponderosa pine. *Biochemical Systematics and Ecology*, **31**, 451–465.

Lee S, Kim J, Fung S, Breuil C (2003) A PCR-RFLP marker distinguishing *Ophiostoma clavigerum* from morphologically similar *Leptographium* species associated with bark beetles. *Canadian Journal of Botany*, **81**, 1104–1112.

Lee S, Kim J, Breuil C (2006) Pathogenicity of *Leptographium longiclavatum* associated with *D. ponderosae* to *Pinus contorta*. *Canadian Journal of Forest Research*, **36**, 2864–2872.

Lee S, Hamelin RC, Six DL, Breuil C (2007) Genetic diversity and the presence of two distinct groups in *Ophiostoma clavigerum* associated with *Dendroctonus ponderosae* in BC and northern Rocky Mountains. *Phytopathology*, **97**, 1177–1185.

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

Lieutier F, Yart A, Salle A (2009) Stimulation of tree defences by Ophiostomatoid fungi can explain attack success of bark beetles on conifers. *Annals of Forest Science*, **66**, 801–823.

Lim Y, Massoumi Alamouti S, Kim JJ, Lee S, Breuil C (2004) Multigene phylogenies of *Ophiostoma clavigerum* and closely related species from bark beetle-attacked *Pinus* in North America. *FEMS Microbiology Letters*, **237**, 89–96.

Logan JA, Powell JA (2001) Ghost forests, global warming, and the mountain pine beetle (Coleoptera: Scolytidae). *American Entomologist*, **47**, 160–173.

Massoumi Alamouti S, Kim J, Humble M, Uzunovic A, Breuil C (2007) Ophiostomatoid fungi associated with the northern spruce engraver, *Ips perturbatus*, in western Canada. *Antonie van Leeuwenhoek*, **91**, 19–34.

Maynard Smith J, Smith NH, O'Rourke M, Spratt BG (1993) How clonal are bacteria? *Proceedings of the National Academy of Sciences of the USA*, **90**, 4384–4388.

Mirov NT (1948) The terpenes (in relation to the biology of genus *Pinus*). *Annual Review of Biochemistry*, **17**, 521–540.

Mirov NT, Hasbrouck J (1976) *The Story of Pines*. Indiana University Press, Bloomington, IN.

Mock KE, Bentz BJ, O'Neill EM, Chong JP, Orwin J, Pfrender ME (2007) Landscape-scale genetic variation in a forest outbreak species, the mountain pine beetle (*Dendroctonus ponderosae*). *Molecular Ecology*, **16**, 553–568.

Möller EM, Bahnweg G, Sandermann H, Geiger HH (1992) A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. *Nucleic Acids Research*, **20**, 6115–6116.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.

Paine TD, Hanlon CC (1994) Influence of oleoresin constituents from *Pinus ponderosa* and *Pinus jeffreyi* on growth of mycangial fungi from *Dendroctonus ponderosae* and *Dendroctonus jeffreyi*. *Journal of Chemical Ecology*, **20**, 2551–2563.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Raffa K (1988) The mountain pine beetle in western North America. In: *Dynamics of Forest Insect Populations Patterns, Causes, Implications* (ed. Berryman A). pp. 506–550, Plenum Press, New York, NY.

Raffa K, Berryman A (1983) Physiological-aspects of lodgepole pine wound responses to a fungal symbiont of the mountain pine-beetle *Dendroctonus ponderosae* (Coleoptera Scolytidae). *Canadian Entomologist*, **115**, 723–734.

Rice AV, Thormann MN, Langor DW (2007) Mountain pine beetle associated blue-stain fungi cause lesions on jack pine, lodgepole pine, and lodgepole × jack pine hybrids in Alberta. *Canadian Journal of Botany*, **85**, 307–315.

Richardson DM (2000) *Ecology and Biogeography of Pinus*. Cambridge University Press, UK.

Robinson-Jeffrey RC, Davidson RW (1968) Three new *Europhium* species with *Verticicladiella* imperfect states on blue-stained pine. *Canadian Journal of Botany*, **46**, 1523–1527.

Roe AD, Rice AV, Bromilow SE, Cooke JEK, Sperling FAH (2010) Multilocus species identification and fungal DNA barcoding: insights from blue stain fungal symbionts of the

mountain pine beetle. *Molecular Ecology Resources*, **10**, 946–959.

Roe AD, Rice AV, Coltman DW, Cooke JEK, Sperling FAH (2011) Comparative phylogeography, genetic differentiation and contrasting reproductive modes in three fungal symbionts of a multipartite bark beetle symbiosis. *Molecular Ecology*, **20**, 584–600.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Safranyik L, Carroll AL, Régnière L *et al.* (2010) Potential for range expansion of mountain pine beetle into the boreal forest of North America. *Canadian Entomologist*, **142**, 415–442.

Six DL, Paine TD (1997) *Ophiostoma clavigerum* is the mycangial fungus of the jeffrey pine beetle, *Dendroctonus jeffreyi*. *Mycologia*, **89**, 858–866.

Six DL, Paine TD (1998) Effects of mycangial fungi and host tree species on progeny survival and emergence of *Dendroctonus ponderosae* (Coleoptera: Scolytidae). *Environmental Entomology*, **27**, 1393–1401.

Six DL, Paine TD (1999) Allozyme diversity and gene flow in *Ophiostoma clavigerum* (Ophiostomatales: Ophiostomataceae), the mycangial fungus of *Dendroctonus jeffreyi*. *Canadian Journal of Forest Research*, **29**, 324–331.

Six DL, Wingfield MJ (2011) The role of phytopathogenicity in bark beetle-fungus symbioses: a challenge to the classic paradigm. *Annual Review of Entomology*, **56**, 255–272.

Six DL, Harrington TC, Steimel J, McNew D, Paine TD (2003) Genetic relationships among *Leptographium terebrantis* and the mycangial fungi of three western *Dendroctonus* bark beetles. *Mycologia*, **95**, 781–792.

Smith RH (2000) *Xylem Monoterpenes of Pines: Distribution, Variation, Genetics, Function*. USDA, Berkeley, CA.

Smith GD, Carroll AL, Lindgren BS (2010) Facilitation in bark beetles: endemic mountain pine beetle gets a helping hand. *Agricultural and Forest Entomology*, **13**, 37–43.

Solheim H, Krokene P (1998) Growth and virulence of mountain pine beetle associated blue stain fungi, *Ophiostoma clavigerum* and *Ophiostoma montium*. *Canadian Journal of Botany*, **76**, 561–566.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

Stock MW, Amman GD (1980) Isozymes differentiation among mountain pine beetle populations from lodgepole pine and ponderosa pine in northeast Utah. *Annals of the Entomological Society of America*, **73**, 472–478.

Stock MW, Amman GD, Higby PK (1984) Genetic variation among mountain pine beetle (*Dendroctonus ponderosae*) (Coleoptera; Scolytidae) populations from seven western states. *Annals of the Entomological Society of America*, **77**, 760–764.

Sturgeon KB, Mitton JB (1982) Evolution of bark beetle communities. In: *Bark Beetles in North American Conifers* (eds Mitton JB, Sturgeon KB). pp. 350–384, Texas University Press, Austin, TX.

Sturgeon KB, Mitton JB (1986) Allozyme and morphological differentiation of mountain pine beetles, *Dendroctonus ponderosae* Hopkins (Coleoptera: Scolytidae), associated with host tree. *Evolution*, **40**, 290–302.

Swofford DL (2003) *PAUP*, Phylogenetic Analysis using Parsimony (*and other methods)*. Sinaer Associates, Sunderland, MA.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Taylor JW, Jacobson DJ, Kroken S *et al.* (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology*, **31**, 21–32.

Templeton AR (1983) Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution*, **37**, 221–244.

Thompson JN (1994) *The Coevolutionary Process*. University of Chicago Press, Chicago, IL.

Tsui CKM, Feau N, Ritland C *et al.* (2009) Characterization of microsatellite loci in the fungus, *Grosmannia clavigera*, a pine pathogen associated with the mountain pine beetle. *Molecular Ecology Resources*, **9**, 1500–1503.

Tsuneda A, Hiratsuka Y (1984) Sympodial and annellidic conidiation in *Ceratocystis clavigera*. *Canadian Journal of Botany*, **62**, 2618–2624.

Wood SL (1963) A revision of the bark beetle genus *Dendroctonus* Erichson (Coleoptera: Scolytidae). *Great Basin Naturalist*, **23**, 1–117.

Wood SL (1982) The bark and ambrosia beetles of North and Central America (Coleoptera, Scolytidae): a taxonomic monograph. *Great Basin Naturalist*, **6**, 1–1359.

Zambino PJ, Harrington TC (1992) Correspondence of isozyme characterization with morphology in the asexual genus *Leptographium* and taxonomic implications. *Mycologia*, **84**, 12–25.

S.M.A. is a doctoral candidate with C.B. at the university of British Columbia (UBC). Her research interests focus on the systematics, molecular evolution and ecological adaptations of ophiostomatoid fungi. V.W. is a research assistant in Breuil's lab with interests in molecular biology. Dr S.D. is a UBC graduate with research interests in plant-pathogen interactions, fungal genetics and genomics. D.S., a professor in forest entomology/pathology, has a research program on the ecology and evolution of bark beetle-fungus symbioses and the effects of climate change on insect-microbial symbioses. J.B. is a professor of plant biology with a research program on forest health genomics and plant biochemistry. Dr R.H. is an adjunct professor at UBC with research interests in forest pathology and fungal genomics. Dr N.F. has research interests in the use of genomic and bioinformatic approaches to address questions on the evolution of fungal-plant interactions. C.B. is a professor in microbiology has a research program on the physiology, ecology and functional genomics of wood-colonizing fungi including tree pathogens.

Data accessibility

List of isolates, sampling locations, collection resources and deposited culture collections: Table S1 in Supporting information.

List of primer sequences, gene descriptions and their coordinates on the *G. clavigera* genome (NCBI accession ACXQ00000000): Table S2 in Supporting information.

Polymorphism summaries and diversity indices: Table S3 in Supporting information.

DNA sequences: GenBank (accession nos. HQ633073–HQ634118, list of isolates with related GenBank numbers: Table S4 in Supporting information).

Morphological data: Table S5 in Supporting information.

67-loci alignment (i.e., polymorphism-detection matrix) and 15-gene phylogeny: TreeBASE study accession no. S11355.

Haplotype network methods and figures: Appendix S1 and Fig. S1 in Supporting information.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Parsimony haplotype networks for 15 independent gene fragments studied in the 62 isolates of *G. clavigera*.

Table S1 Fungal isolates used in this study

Table S2 Primer sequences, scaffold identities and gene descriptions for 67. *G. clavigera* loci screened for polymorphisms

Table S3 Polymorphisms summaries and diversity indices with in the two monophyletic clades in *G. clavigera*

Table S4 Funal isolates used in this study and GenBank numbers for sequences

Table S5 Characteristics of Gs isolates compared with those of the *G. clavigera* holotype

Appendix S1 Methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.