



HAL
open science

Genetic isolation between two recently diverged populations of a symbiotic fungus

Sara Branco, Pierre P. Gladieux, Christopher Ellison, Alan Kuo, Kurt Labutti, Anna Lipzen, Igor V I. V. Grigoriev, Liao Hui-Ling, Rytas Vilgalys, Kabir G. Peay, et al.

► **To cite this version:**

Sara Branco, Pierre P. Gladieux, Christopher Ellison, Alan Kuo, Kurt Labutti, et al.. Genetic isolation between two recently diverged populations of a symbiotic fungus. *Molecular Ecology*, 2015, 24 (11), pp.2747-2758. 10.1111/mec.13132. hal-01303143

HAL Id: hal-01303143

<https://hal.science/hal-01303143v1>

Submitted on 11 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Genetic isolation between two recently diverged
populations of a symbiotic fungus**

Journal:	<i>Molecular Ecology</i>
Manuscript ID:	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Branco, Sara; University of California, Berkeley, Plant and Microbial Biology Gladieux, Pierre; University Paris Sud, Laboratoire d'Ecologie, Systematique et Evolution Ellison, Chris; University of California, Berkeley, Integrative Biology Kuo, Allan; Department of Energy, Joint Genome Institute LaButii, Kurt; Department of Energy, Joint Genome Institute Lipzen, Anna; Department of Energy, Joint Genome Institute Grigoriev, Igor; Department of Energy, Joint Genome Institute Liao, Hui-Ling; Duke University, Biology Vilgalys, Rytas; Duke University, Biology Peay, Kabir; Stanford University, Biology Taylor, John; University of California, Berkeley, Plant and Microbial Biology Bruns, Thomas; University of California Berkeley, Plant and Microbial Biology; University of California, Berkeley, Plant and Microbial Biology
Keywords:	Adaptation, Fungi, Natural Selection and Contemporary Evolution, Population Genetics - Empirical



1 **Genetic isolation between two recently diverged populations of a symbiotic fungus**

2 Sara Branco¹, Pierre Gladieux², Christopher E. Ellison³, Alan Kuo⁴, Kurt LaButii⁴, Anna
3 Lipzen⁴, Igor V. Grigoriev⁴, Hui-Ling Liao⁵, Rytas Vilgalys⁵, Kabir G. Peay⁶, John W.
4 Taylor¹, Thomas D. Bruns¹

5

6 ¹ Department of Plant and Microbial Biology, University of California, Berkeley,
7 California, United States of America

8 ² Laboratoire d'Ecologie, Systematique et Evolution, Universite Paris Sud, Orsay, France

9 ³ Department of Integrative Biology, University of California, Berkeley, California,
10 United States of America

11 ⁴ Department of Energy, Joint Genome Institute, Walnut Creek, California, United States
12 of America

13 ⁵ Department of Biology, Duke University, Durham, North Carolina, United States of
14 America

15 ⁶ Department of Biology, Stanford University, Stanford, California, United States of
16 America

17

18 Keywords – Adaptation; population genomics; mycorrhizal fungi; *Suillus brevipes*

19

20 Corresponding author – Sara Branco; sara.mayer.branco@gmail.com

21

22 Running title – Recent fungal population divergence

23

24 **Genetic isolation between two recently diverged populations of a symbiotic fungus**

25 Sara Branco, Pierre Gladieux, Christopher E. Ellison, Alan Kuo, Kurt LaButti, Anna

26 Lipzen, Igor V. Grigoriev, Hui-Ling Liao, Rytas Vilgalys, Kabir G. Peay, John W.

27 Taylor, Thomas D. Bruns

28

29 **Abstract**

30 Fungi are an omnipresent and highly diverse group of organisms, making up a
31 significant part of eukaryotic diversity. Little is currently known about the drivers of
32 fungal population differentiation and subsequent divergence of species, particularly in
33 symbiotic fungi. Here we investigate the population structure and environmental
34 adaptation in *Suillus brevipes* (Peck) Kuntze, a wind-dispersed soil fungus that is
35 symbiotic with pine trees. We assembled and annotated the reference genome for *S.*
36 *brevipes* and re-sequenced the whole genomes of 28 individuals from coastal and
37 montane sites in California. We detected two highly supported populations with very low
38 divergence and no evidence for migration. However, we detected highly differentiated
39 genomic regions, most notably a region that contains a gene encoding for a membrane
40 Na⁺/H⁺ exchanger known for enhancing salt tolerance in plants and yeast. Our results are
41 consistent with a very recent split between the montane and coastal *S. brevipes*
42 populations, with few small genomic regions under positive selection and a pattern of
43 dispersal and/or establishment limitation. This is the first population genomics study
44 comprising a member of the phylum Basidiomycota or of a mycorrhizal fungus. It
45 documents the least differentiated fungal populations currently known and identifies a

46 putatively adaptive gene that motivates further functional tests to link genotypes and
47 phenotypes and shed light on the genetic basis of adaptive traits.

48

49 **Introduction**

50 The way organisms are distributed across the landscape is the result of their
51 ability to reach, establish and persist in suitable habitats. Factors such as dispersal and
52 establishment limitation influence these abilities, shaping the genetic diversity of
53 populations in nature. Fungi are no exception. They have simple morphologies allied to
54 well-defined ecological niches, comprise considerable diversity and tend to have small
55 genomes. These features make them good models for studying population differentiation
56 and adaptation in nature (Gladieux *et al.*, 2014). Unlike most animals and plants, the vast
57 majority of fungi produce copious amounts of microscopic airborne propagules and are
58 often assumed to have exceptional long-range dispersal capabilities (Brown *et al.*, 2013)
59 leading to wide distributions and large population sizes (Green & Bohannan, 2006). This
60 idea has however been recently challenged by ecological evidence that dispersal
61 limitation at the landscape scale affects community structure and richness (Peay *et al.*,
62 2012) and that there is widespread fungal regional endemism (Taylor *et al.*, 2006,
63 Amend *et al.*, 2010, Talbot *et al.*, 2014). Even though fungi are good models for
64 understanding the genetic basis of evolutionary change, it is not yet clear how fungal
65 adaptive divergence occurs and whether it is linked to dispersal, natural selection, or
66 both.

67 Environmental factors are known to drive fungal adaptation, with temperature as a
68 recurrent feature associated with wild fungal population differentiation even in host-

69 associated fungi (Vacher *et al.*, 2008, Frenkel *et al.*, 2010, Ellison *et al.*, 2011, Mboup *et*
70 *al.*, 2012). However, fungal population genomics and adaptation have been studied in a
71 very small set of fungi (Liti *et al.*, 2009, Neafsey *et al.*, 2010, Ellison *et al.*, 2011,
72 Rosenblum *et al.*, 2013). This limited sample leaves out an immense diversity of fungal
73 groups and life histories, including all symbiotic basidiomycete mushroom-forming
74 fungi. The phylum Basidiomycota is one of the largest fungal phyla, comprising about
75 30,000 species (James *et al.*, 2006) and including the majority of mycorrhizal species,
76 which are soil fungi living in obligate symbiotic associations with vascular plants (Smith
77 & Read, 2010). In particular, ectomycorrhizal (EM) fungi are an important group
78 associated with a wide range of plant families including tree species dominant in many
79 biomes. EM fungi play a crucial role in forest ecosystems by providing water and
80 nutrients to their host trees, for which they receive carbohydrates in return (Smith &
81 Read, 2010). Although there have been attempts to understand EM fungal population
82 genetics (Douhan *et al.*, 2011), little is currently known about the patterns of population
83 structure in these fungi or on the relative role of dispersal limitation and local adaptation
84 in shaping fungal diversity.

85 Here we report on the first population genomics study of an EM basidiomycete
86 fungus. We compile a new fungal reference genome, document fungal population
87 structure and investigate candidate genomic regions for local adaptation. We focused on
88 *Suillus brevipes*, a widespread North American EM fungus associated with pine trees
89 (Smith & Thiers, 1964). *S. brevipes* is a soil fungus and reproduces by forming fruiting
90 bodies that produce and disperse in the order of 10^8 airborne spores per mushroom
91 (Dahlberg & Stenlid, 1994). It plays an important ecological role as a pioneer species,

92 colonizing pine seedlings and young trees (Peay *et al.*, 2012) and contrary to most EM
93 species, is amenable for *in vitro* cultures, making it possible to document population
94 differentiation and to unveil its causal mechanisms. We assembled and annotated the *S.*
95 *brevipes* reference genome and re-sequenced the genomes of isolates collected from two
96 contrasting environments in California, coastal mixed *Pinus muricata* and *P. contorta*
97 and mountain *P. contorta* forests separated by 300 km of inhospitable habitat, lacking
98 either plant host. Given the inherent differences between coastal and montane
99 environments, we expected to find clearly differentiated and locally adapted populations,
100 with strong signatures of positive selection in genomic regions related to factors known
101 to be relevant for fungal ecology and evolution, such as host tree, climate, and/or soil
102 type. We also hypothesized ongoing gene flow between the coast and mountains,
103 facilitated by the massive amounts of airborne spores produced by *S. brevipes*.
104 Surprisingly, we detected distinct yet very little diverged coastal and montane
105 populations, with only 155 reciprocally fixed SNPs. Demographic modeling pointed to
106 isolation as the most likely model, with the absence of gene flow between the two
107 populations. We also identified a small set of highly differentiated genomic regions,
108 including a gene involved in salt tolerance, indicating positive selection plays a role in
109 shaping these populations and suggesting the emergence of local adaptation in *S.*
110 *brevipes*.

111

112 **Materials and Methods**

113 *Isolate collection, culturing, and DNA extraction and sequencing*

114 *Suillus brevipes* (Peck) Kuntze individual sporocarps were collected from pine
115 forests from the coast and mountains of Northern California (Table S1). We harvested 11
116 coastal individuals in Mendocino County, Salt Point and Point Reyes National Seashore
117 from forests dominated by *Pinus muricata* with residual *P. contorta* in the Mendocino
118 site (under the Scientific Research and Collecting Permit # PORE-2011-SCI-0014).
119 These sites are characterized by a Mediterranean climate with cool summers and mild
120 winters with temperatures rarely reaching below freezing. We also collected 17 *S.*
121 *brevipes* individuals in the Sierra Nevada mountains (Yosemite National Park; under the
122 Scientific Research and Collecting Permit # YOSE-2013-SCI-0071) from *P. contorta*
123 forests, around 300 km East from the coastal sites. These mountains also offer a
124 Mediterranean climate but winters include below freezing temperatures and most
125 precipitation arrives as snowfall. Sporocarps were collected at least 10 meters apart to
126 ensure selection of different genets (Bergemann & Miller, 2002, Dahlberg & Stenlid,
127 1994), with some mountain individuals separated by up to 30 miles and coastal
128 individuals separated by over 50 miles. The coastal and mountain collection sites are
129 separated by dry, non-forested and inhospitable Californian Central Valley habitat.

130 We obtained pure cultures from each individual by placing a piece of each sporocarp
131 in full strength Modified Melin-Norkrans (MMN) agar plates. Pure mycelium was
132 harvested after no more than two weeks of growth, flash frozen in liquid N, and stored at
133 -80C. Genomic DNA was obtained for each isolate using a CTAB-based extraction
134 protocol. Genomic libraries were generated with Apollo 324 NGS Library Prep System
135 (Wafergen Biosystems) and samples were sequenced in 3 lanes of Illumina HiSeq2000,
136 100bp paired-end with a 350bp insert size at the QB3 Vincent J. Coates Genomics

137 Sequencing Laboratory at the University of California, Berkeley. The raw reads were
138 deposited in the Short Read Archive and the accession numbers are listed in Table S1.
139 One mountain isolate (Sb2) was selected as the reference genome and deposited in the
140 CBS Collection of Fungi, with accession number CBS 139023. DNA was extracted as
141 above, however library preparation, sequencing, genome assembly and annotation were
142 carried out by the US DOE Joint Genome Institute under the 1000 Fungal Genomes
143 project (Grigoriev *et al.*, 2013). The genome was sequenced using Illumina HiSeq
144 platform. Paired-end reads from 270bp insert size standard Illumina library were filtered
145 for artifact/process contamination and subsequently assembled as a diploid using
146 AllPathsLG release version R43839 (Gnerre *et al.*, 2011). The final assembly of 1550
147 scaffolds has an N50 of 159 kbp and an L50 of 83. The assembly was repeat-masked and
148 annotated using the JGI Annotation Pipeline (Kuo *et al.*, 2014). Briefly, the Pipeline
149 computes multiple gene models at each locus using a variety of RNASeq-based, protein-
150 based, and *ab initio* gene predictors, characterizes those models by predicting functional
151 domains and signal peptides, and selects the highest quality model at each locus based on
152 expression, similarity to known proteins, and domain content. In addition, the Pipeline
153 compares the draft gene set to those of other genomes and builds a draft set of gene
154 families. Finally, the Pipeline exports the assembly and annotation to an interactive Portal
155 in MycoCosm (Grigoriev *et al.*, 2013). The *Suillus brevipes* Sb2 Portal can be found at
156 http://genome.jgi.doe.gov/Suillus_brevipes/

157

158 *Read filtering, mapping and variant calling*

159 Isolate re-sequencing produced over 60 million sequences. Reads were filtered based
160 on quality and all duplicates and reads with more than 10% Ns, and with at least one base
161 with phred score smaller than 20 were discarded. Read pairs were re-grouped and orphan
162 reads were also discarded. All good quality paired reads were then aligned to the
163 reference genome using Bowtie2. 2.1.0 (Langmead & Salzberg, 2012). Alignments were
164 sorted with SAMtools (Li *et al.*, 2009) and duplicate and poor quality alignments
165 removed with Picard (<http://picard.sourceforge.net>) and BamTools (Barnett *et al.*, 2011).
166 Realignment, indexing and SNP calling was performed with the Genome Analyses
167 Toolkit (GATK; McKenna *et al.*, 2010), using unified genotyper. We discarded
168 singletons, poor quality SNPs (<30) and SNPs including Ns using VCFtools (Danecek *et*
169 *al.*, 2011) and in-house scripts. Per individual genome coverage was calculated using
170 GATK.

171

172 *Population structure*

173 Population structure was inferred using both clustering and phylogenetic inference.
174 We used 567192 SNPs (excluding singletons) to compute a neighbor-joining tree with
175 PHYLIP (Felsenstein, 2002) including all the *S. brevipes* isolates. We also used FRAPPE
176 to compute global ancestry from maximum likelihood estimates of model parameters
177 (Tang *et al.*, 2005) and assess population structure. Two coastal isolates, Sb7 and Sb10,
178 were collected just 10m apart in Mendocino. These individuals shared 97% SNPs, so we
179 considered them to be clones and removed Sb7 from further analyses.

180

181 *Polymorphism and genome scans for divergence*

182 Gene and genome wide polymorphism were assessed using egglib (Mita & Siol,
183 2012) and in-house scripts. We calculated the average 5kb window and gene number of
184 segregating sites and the whole genome and genic nucleotide diversity (π), including
185 silent and non-silent π in each population. We also calculated Tajima's D to test for
186 neutrality and Zeng's E to compare for high- and low-frequency variants (Zeng *et al.*,
187 2006). In addition, we calculated the average number of nucleotide substitutions per site
188 between populations (Dxy) for 5Kb windows and measured population differentiation
189 due to genetic structure by computing the average genome-wide Weir Cockerham's Fst
190 and average Fst per 5Kb and 1Kb windows with VCFtools (Danecek *et al.*, 2011). We
191 calculated fixed and exclusive differences as well as expected and observed
192 heterozygosity across the two populations with in-house scripts.

193

194 *Demography*

195 Population demography was analyzed using a diffusion-based approach implemented
196 in the program $\partial a \partial i$ (Gutenkunst *et al.*, 2009). Demographic parameters were estimated
197 from the joint mountain and coastal *S. brevipes* allele frequency spectra polarized with
198 the *S. luteus* genome sequence. Given the risk for potential misidentification of ancestral
199 states, we also fitted the model using folded allele frequency spectra. Conventional
200 bootstrapping (100 replicates) was calculated to estimate parameter uncertainties.

201

202 *Natural selection*

203 To detect signals of positive selection we identified outlier Dxy and Fst values
204 within each population, performed McDonald Kreitman tests and computed a selective
205 sweep analysis. We selected the top 1% Dxy values for 5Kb windows and top 0.1% Fst
206 values for 5Kb and 1Kb windows to search for differentiated regions across the two
207 populations and respective genes. The functions of these highly differentiated genes were
208 inferred by retrieving gene ontology categories from MycoCosm (Grigoriev *et al.*, 2013).

209 We also identified genes showing an excess of amino acid substitutions using the
210 MacDonal-Kreitman test (McDonald & Kreitman, 1991), with *S. luteus* as the outgroup.
211 We filtered the results to exclude genes that showed limited variability and corrected for
212 multiple hypothesis testing using the Benjamini-Hochberg method (Benjamini &
213 Hochberg, 1995).

214 We searched for selective sweeps by measuring haplotype homozygosity using an in-
215 house script (Wilson & Messer, unpublished). This program calculates an H statistics to
216 detect hard and soft sweeps by measuring the length of pairwise homozygosity stretches
217 in a population sample generated by elevated linkage disequilibrium around adaptive
218 sites.

219

220 **Results**

221 *Suillus brevipes* genomes

222 The *S. brevipes* reference genome (obtained from Sb2, a mountain individual
223 included in this study) was assembled into 1550 scaffolds totaling 51.7 Mbp and
224 annotated. The 22,453 predicted genes, represent 2313 distinct protein domains, and

225 cluster into 2844 multigene families. Details on the genome assembly and annotation can
226 be found in Table S2.

227 The average sequence coverage per individual varied from 7X to 26X (Table S1).
228 We recovered a total of 1254275 single nucleotide polymorphisms (SNPs) across the 28
229 *S. brevipes* genomes, with 2.4% of the genome as variable sites.

230

231 *Population subdivision*

232 Using a neighbor-joining approach we found two distinct and highly supported
233 *Suillus brevipes* montane and coastal populations (Fig. 1a). The same two populations
234 were found with a maximum likelihood population structure analysis using FRAPPE,
235 which emphasized the shared ancestry of the two populations (Fig. 1b). All individuals
236 collected from montane Yosemite National Park clustered together, and individuals from
237 coastal Mendocino, Point Reyes, and Salt Point formed another cluster with no evidence
238 of geographic substructure.

239

240 *Genomic diversity, differentiation and divergence*

241 Genomic summary statistics were estimated using 5kb sliding windows (6387
242 windows for the mountain population and 6301 for the coast population) and the genic
243 regions (22453 genes) (Table 1). The two populations presented comparable levels of
244 genetic variation, and each population showed close to 2% polymorphism across the
245 genome and 0.2% in the genic regions. The average observed heterozygosity was similar
246 in both populations ($H_{O_{Coast}} = 0.32 \pm 0.19$, $H_{O_{Mountain}} = 0.28 \pm 0.18$), and only slightly
247 higher than the heterozygosity expected under Hardy Weinberg equilibrium ($H_{e_{Coast}} =$

248 0.20 ± 0.21 ; $H_{e\text{Mountain}} = 0.21 \pm 0.21$). Nucleotide diversity (π) was also similar between
249 coast and mountain (Table 1). Tajima's D was slightly negative, indicating a genome-
250 wide excess of low frequency polymorphism compared to neutral expectations, however
251 this was not observed when including only genic regions (Table 1). Zeng's E statistics,
252 which measures positive selection, was negative in both coastal and montane populations,
253 indicating an excess of high frequency derived polymorphisms in both populations. In
254 fact, testing the fit between the observed site frequency spectra and spectra generated
255 with a constant size model using the PYTHON package $\partial a \partial i$ (Gutenkunst *et al.*, 2009)
256 detected departures from the standard neutral model in both populations (mountain
257 chisquare, d.f. = 32, $P = 0$; coast chisquare, d.f. = 20, $P = 0$).

258 Most of the SNPs were shared between the two populations (44%) and only 155
259 (0.01%) of the sites were reciprocally fixed for different alleles in the two populations.
260 The remaining SNPs were variable in one population and fixed in the other (Table 2).
261 This distribution of polymorphism and divergence suggests a recent split between the two
262 populations, the prevalence of gene flow, or both. The mean pairwise nucleotide
263 difference between populations (D_{xy}) for 5kb windows was 0.008, consistent with the
264 results of the clustering analyses.

265

266 *Demography*

267 Demographic inference of coastal and montane populations indicated isolation
268 (absence of gene flow) between the mountain and coastal populations as the most likely
269 demographic model (Table 3). Inference using folded site frequency spectra also pointed
270 to the isolation model as the best fit to the data (isolation likelihood = -32595.1;

271 migration likelihood = -239130.9), suggesting that model choice was not biased by
272 multiple mutations along the outgroup branch.

273 The maximum-likelihood parameter estimates for the unfolded isolation model are
274 reported in Table 4. The coastal population is estimated to have a smaller effective
275 population size compared to the montane population and the estimated divergence time is
276 25000 generations. Even though this is the most likely of all tested demographic models,
277 the unfolded modeled site frequency spectrum did not fit the observed data well
278 (chisquare test mountain: d.f. = 33, $P = 0$; chisquare test coast: d.f. = 21, $P = 0$). The poor
279 fit was due to a deficit of high frequency derived alleles in simulated site frequency
280 spectra compared to the observed dataset (Fig S1).

281

282 *Genome scans for natural selection*

283 In order to test for natural selection, we investigated genomic regions with a high
284 proportion of reciprocally fixed SNPs, regions identified as outliers for Dxy and Fst
285 (Noor & Bennett, 2009, Cruickshank & Hahn, 2014). We also used McDonald-Kreitman
286 tests and selective sweep analysis to identify signals of positive selection across the
287 genome.

288 Out of the detected 155 reciprocally fixed differences, 30 were in genic regions.
289 Scaffold 16 hosted 30% of these, with 5 non-synonymous fixed differences, and scaffold
290 12 included 20%, of which 4 were non-synonymous. Three of these genes had known
291 functions. Scaffold 16 includes gene 669775, which codes for a protein kinase. Scaffold
292 12 includes three adjacent genes with six fixed differences. 826957 has no known
293 function, 826959 codes for a cysteine-type endopeptidase and 826958 is involved in pH

294 regulation. The latter is invariable in the coastal individuals (Fig. 2) and codes for a
295 Na^+/H^+ exchanger, which is a membrane protein involved in salt tolerance in
296 *Sacchromyces cerevisiae* and *Arabidopsis* (Nass *et al.*, 1997, Apse *et al.*, 1999). The *S.*
297 *cerevisiae* homolog *Nha1* encodes for a membrane protein involved in Na^+ and K^+ efflux
298 needed for yeast to tolerate alkali cation tolerance at acidic pH (Bañuelos *et al.*, 1998).
299 Gene 826958 will hereafter be referred to as *Nha1-like*.

300 We used Dxy outliers as an absolute divergence metric to scan the genome for
301 regions under divergent selection (Noor & Bennett, 2009, Cruickshank & Hahn, 2014).
302 Table S3 reports the genes located in the top 1% 5kb window Dxy values. These encode
303 proteins with a wide variety of functions, with 20% involved in transporter activity.
304 *Nha1-like* and flanking genes in scaffold 12 (coding for an endopeptidase and another
305 unknown protein) were included in the Dxy outlier list.

306 We also investigated Fst across the genome. On average Fst values per SNP were
307 low (average Fst = 0.053, s.d. = 0.115), as expected given the low level of differentiation
308 between populations. Fst was also calculated for 5 Kb and 1Kb genomic sliding windows
309 and compared to the list of Dxy outliers. The majority of windows showed low Fst values
310 yet some regions had high relative divergence. Fig. 3 depicts the 1Kb window Fst values,
311 showing 807 windows displaying Fst values above 0.5. We found 24 genes associated
312 with the top 0.1% of 1kb windows with highest Fst values (Table 5). These are involved
313 in a variety of cellular functions, namely transporter, catalytic, oxidoreductase activities
314 and DNA binding. Four Dxy outliers were also on this list, including one gene in scaffold
315 66 and the three Dxy outlier genes in scaffold 12 (these were not listed as Fst outliers

316 when scanning the genomes with 5Kb windows). Gene *Nha1-like* fell in the window with
317 the maximum F_{st} value ($F_{st} = 1$).

318 Using McDonald Kreitman tests, we found four genes displaying significant
319 differences between ratios of non-synonymous to synonymous substitutions within and
320 between species, which are consistent with positive selection (Table S4). These are
321 involved in heterokaryon incompatibility (also listed as a Dxy outlier), carbohydrate
322 metabolic processes, and post-translational processes in organelles.

323 We did not find signals of recent selective sweeps across the genome. We quantified
324 haplotype homozygosity using an H statistics that measures the length of pairwise
325 homozygosity stretches in a population sample generated by elevated linkage
326 disequilibrium around adaptive sites (Wilson and Messer, unpublished). The highest H
327 values are listed in Table S5 and show no evidence for selective sweeps.

328

329 **Discussion**

330 Our study revealed emerging differentiation of *Suillus brevipes* populations in two
331 contrasting environments at a relatively small spatial scale. We found a very recent split
332 between montane and coastal populations in California with no evidence for post-
333 divergence migration, as indicated by low population divergence, the presence of very
334 few reciprocally fixed differences and genetic isolation as the most likely demographic
335 model. Despite the low overall genomic divergence, we detected a small set of highly
336 differentiated and positively selected genes across the two populations. These results are
337 consistent with the early stages of environmental adaptation.

338 Our results point to the existence of dispersal and/or colonization limitation in *Suillus*
339 *brevipes*, which can result from constraints in physically moving spores or challenges in
340 spore germination and mycelium persistence in new habitats. Although *Suillus* is known
341 to be a prolific disperser, producing annual spore loads on the order of trillions of spores
342 per km², spore loads decrease exponentially with distance from their source, rendering
343 colonization of new host trees and establishment in new habitats less effective than
344 expected (Peay *et al.*, 2012). This finding makes spore dispersal limitation a good
345 candidate explanation for the lack of gene flow between coastal and montane populations
346 in California. These habitats are separated by the Central Valley, which comprises an
347 approximate 300 km non-forested gap that based on our results appears to represent an
348 effective geographic barrier to migration between the coastal and montane populations of
349 *S. brevipes*.

350 Despite demographic modeling pointing to two allopatric *S. brevipes* populations as
351 the most likely demographic model, this model had a poor fit to the data due to an excess
352 of high frequency derived alleles in both populations (Fig. S1). This pattern could be an
353 artifact derived from misidentification of ancestral alleles or a true biological pattern,
354 explained by more complex demographic scenarios (including bottlenecks) or genome-
355 wide positive selection, such as selective sweeps (Ellison *et al.*, 2011, Caicedo *et al.*,
356 2007). We had access to a single outgroup (*S. luteus*, the sister species to *S. brevipes*),
357 making it challenging to control for errors in calling ancestral alleles. However, results
358 from demographic modeling without an outgroup (using a folded site frequency spectra)
359 also pointed to absence of migration as the most probable scenario. Furthermore, we

360 failed to detect selective sweeps indicating genome-wide positive selection is probably
361 not the cause for the non-monotonic observed site frequency spectra.

362 The few available studies investigating wild EM fungal population structure report
363 mixed trends, ranging from panmixis to strong population differentiation (Douhan *et al.*,
364 2011). Microsatellite approaches conducted in several EM fungal species point to
365 examples of lack of population structure, extensive gene flow and little genetic
366 differentiation both at local and regional spatial scales (Zhou *et al.*, 2001, Kretzer *et al.*,
367 2005, Bergemann *et al.*, 2006, Muller *et al.*, 2007, Roy *et al.*, 2008, Rivera *et al.*, 2014).
368 However, most of these studies were based on a very small number of markers that might
369 have been insufficient to detect emerging population differentiation. In fact, several other
370 examples point to high levels of population differentiation in EM fungi (Bergemann &
371 Miller, 2002, Grubisha *et al.*, 2007, Carriconde *et al.*, 2008, Amend *et al.*, 2010, Dunham
372 *et al.*, 2013). In this regard, population genomics of the saprobe *Neurospora*, detected
373 structure hidden in microsatellite analysis (Ellison *et al.*, 2011). Unfortunately, the causes
374 underlying population differentiation are not always clear.

375 The few highly differentiated genomic regions are consistent with the very early
376 steps in local adaptation in *Suillus brevipes*. In particular, the reduced nucleotide diversity
377 among coastal individuals at the *NhaI-like* gene is consistent with recent positive
378 selection in the coastal population, making it tempting to hypothesize that the coastal
379 allele of this gene spread by natural selection for adaptation to the stress of saline soil.
380 Consistent with this hypothesis, soil analyses from at least the Point Reyes area confirm
381 high Na content, with an average of 170 ppm, and acidic pH (average pH = 5.1; Peay *et al.*
382 *et al.*, 2010). Basidiomycetes are in general relatively salt-intolerant and saline soils are

383 known to host poorly diverse EM fungal communities (Schechter & Branco, 2014 and
384 references therein). Further research will be needed to fully understand the role of gene
385 *Nha1-like* for salt tolerance in *Suillus brevipes*.

386 Our study is one of the few available fungal population genomic studies so far and
387 provides the earliest example of divergence and emergence of environmentally adapted
388 fungal populations. Other studies also reported evidence of adaptive evolution, revealing
389 more complex patterns of genomic differentiation, demographics and positive selection.
390 A pioneer population genomics study on wild and domesticated yeasts (*Saccharomyces*
391 *cerevisiae* and *S. paradoxus*) found worldwide population differentiation correlated to
392 phenotypic variation (Liti *et al.*, 2009). Ellison *et al.*, 2011 found evidence of cryptic
393 population differentiation and local adaptation in *Neurospora crassa*. Their genome scans
394 revealed distinct populations at different latitudes, with up to 9% genomic divergence,
395 asymmetrical migration and areas of genomic differentiation including genes related to
396 temperature and circadian rhythm, showing environmental factors shape fungal
397 evolution. *Batrachochytrium dendrobatidis*, the amphibian-killing chytrid fungus, was
398 found to include high genetic variation across populations and a complex evolutionary
399 history, including differences in chromosome copy numbers and heterozygosity across
400 isolates as well as signatures of selection in genes involved in amphibian skin infection
401 (Rosenblum *et al.*, 2013). Whole genome data has also revealed population structure,
402 strong genomic divergence (~50%) and signals of introgression in the human pathogens
403 *Coccidioides immitis* and *C. posadasii* (Neafsey *et al.*, 2010). Comparing to these studies,
404 *Suillus brevipes* is an example of a much more recent population split, as demonstrated

405 by the very low amount of reciprocally fixed alleles (0.01%), low Dxy and low overall
406 Fst.

407 This study documents the first steps of fungal population differentiation, reporting
408 the most recent split between fungal populations to date. Our results corroborate the
409 existence of dispersal limitation in fungi and show positive selection associated with
410 abiotic factors as a driver for population differentiation in fungi, including a gene
411 involved in salt tolerance not previously hypothesized to be relevant in EM fungi. Further
412 functional tests can clarify the importance of this gene and shed light on the genetic basis
413 of adaptive traits in fungi. Furthermore, a broader geographic sampling, including
414 samples from other environments will aid in understanding the evolution of *S. brevipes* in
415 particular and of EM fungi in general.

416

417 **Acknowledgements**

418 We thank Holly Edes for technical assistance, the Computational Genomics Resource
419 Laboratory at the University of California, Berkeley for assistance with computational
420 analysis, Joey Spatafora and the 1000 Fungal Genomes Project for the *Suillus brevipes*
421 reference genome and annotation, Ben Wilson and Philipp Messer for sharing scripts to
422 perform selective sweep analysis, and Megan Phifer-Rixey for comments on an earlier
423 draft. Financial support was supported by National Science Foundation grant DBI
424 1046115. The work conducted by the U.S. Department of Energy Joint Genome Institute,
425 a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-
426 05CH11231.

427

428 **References**

- 429 Amend A, Garbelotto M, Fang Z *et al.* (2010) Isolation by landscape in populations of a
430 prized edible mushroom *Tricholoma matsutake*. *Conserv Genet*, **11**, 795–802.
- 431 Amend A, Seifert KA, Samson R *et al.* (2010) Indoor fungal composition is
432 geographically patterned and more diverse in temperate zones than in the tropics.
433 *Proc Natl Acad Sci*, **107**, 13748–13753.
- 434 Apse MP, Aharon GS *et al.* (1999) Salt tolerance conferred by overexpression of a
435 vacuolar Na⁺/H⁺ antiport in *Arabidopsis*. *Science*, **285**, 1256–1258.
- 436 Bañuelos MA, Sychrová H, Bleykasten-Grosshans C *et al.* (1998) The Nha1 antiporter of
437 *Saccharomyces cerevisiae* mediates sodium and potassium efflux. *Microbiol Read*
438 *Engl*, **144**, 2749–2758.
- 439 Barnett D, Garrison E, Quinlan A *et al.* (2011) BamTools: a C++ API and toolkit for
440 analyzing and managing BAM files. *Bioinformatics*, **17**, 1691–1692.
- 441 Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and
442 Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol*, **57**, 289–
443 300.
- 444 Bergemann SE, Douhan GW, Garbelotto M *et al.* (2006) No evidence of population
445 structure across three isolated subpopulations of *Russula brevipes* in an oak/pine
446 woodland. *New Phytol*, **170**, 177–184.
- 447 Bergemann SE, Miller SL (2002) Size, distribution, and persistence of genets in local
448 populations of the late-stage ectomycorrhizal basidiomycete, *Russula brevipes*.
449 *New Phytol*, **156**, 313–320.
- 450 Brown AMV, Huynh LY, Bolender CM *et al.* (2013) Population genomics of a symbiont
451 in the early stages of a pest invasion. *Mol Ecol*, **23**, 1516–1530.
- 452 Caicedo AL, Williamson SH, Hernandez RD *et al.* (2007) Genome-Wide Patterns of
453 Nucleotide Polymorphism in Domesticated Rice. *PLoS Genet*, **3**, e163.
- 454 Carriconde F, Gryta H, Jargeat P *et al.* 2008. High sexual reproduction and limited
455 contemporary dispersal in the ectomycorrhizal fungus *Tricholoma scalpturatum*:
456 new insights from population genetics and spatial autocorrelation analysis. *Mol*
457 *Ecol*, **17**, 4433–4445.
- 458 Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of
459 speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol*, **23**,
460 3133–3157.
- 461 Dahlberg A, Stenlid J (1994) Size, distribution and biomass of genets in populations of
462 *Suillus bovinus* (L.: Fr.) Roussel revealed by somatic incompatibility. *New Phytol*,
463 **128**, 225–234.
- 464 Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools.
465 *Bioinformatics*, **27**, 2156–2158.
- 466 Douhan GW, Vincenot L, Gryta H *et al.* (2011) Population genetics of ectomycorrhizal
467 fungi: from current knowledge to emerging directions. *Fungal Biol*, **115**, 569–
468 597.
- 469 Dunham SM, Mujic AB, Spatafora JW *et al.* (2013) Within-population genetic structure
470 differs between two sympatric sister-species of ectomycorrhizal fungi,
471 *Rhizopogon vinicolor* and *R. vesiculosus*. *Mycologia*, **105**, 814–826.

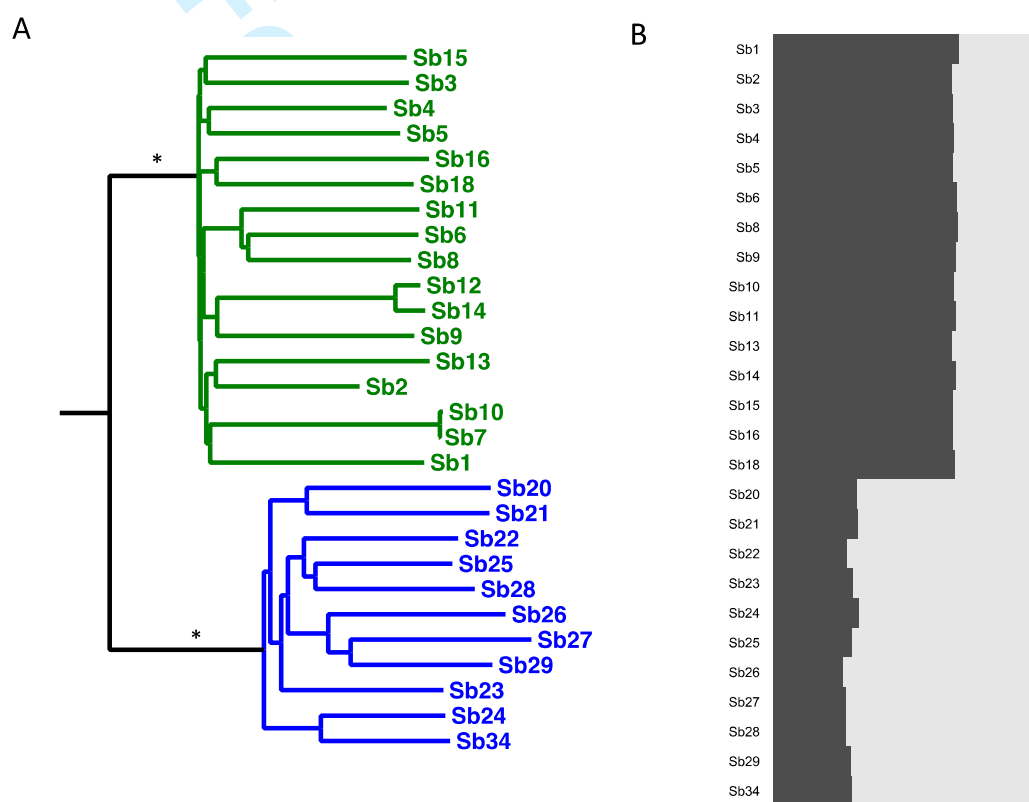
- 472 Ellison CE, Hall C, Kowbel D *et al.* (2011) Population genomics and local adaptation in
473 wild isolates of a model microbial eukaryote. *Proc Natl Acad Sci*, **108**, 2831-
474 2836.
- 475 Felsenstein, J (2002) {PHYLIP} (Phylogeny Inference Package) version 3.6a3.
- 476 Frenkel O, Peever TL, Chilvers MI *et al.* (2010) Ecological Genetic Divergence of the
477 Fungal Pathogen *Didymella rabiei* on Sympatric Wild and Domesticated *Cicer*
478 spp. (Chickpea). *Appl Environ Microbiol*, **76**, 30–39.
- 479 Gladieux P, Ropars J, Badouin H *et al.* (2014) Fungal evolutionary genomics provides
480 insight into the mechanisms of adaptive divergence in eukaryotes. *Mol Ecol*, **23**,
481 753–773.
- 482 Gnerre S, MacCallum I, Przybylski D *et al.* (2011) High-quality draft assemblies of
483 mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci*,
484 **108**, 1513–1518.
- 485 Green J, Bohannan BJM (2006) Spatial scaling of microbial biodiversity. *Trends Ecol*
486 *Evol*, **21**, 501–507.
- 487 Grigoriev IV, Nikitin R, Haridas S *et al.* (2013) MycoCosm portal: gearing up for 1000
488 fungal genomes. *Nucleic Acids Res*, **42**, D699–D704.
- 489 Grubisha LC, Bergemann SE, Bruns TD. 2007. Host islands within the California
490 Northern Channel Islands create fine-scale genetic structure in two sympatric
491 species of the symbiotic ectomycorrhizal fungus *Rhizopogon*. *Mol Ecol*, **16**,
492 1811–1822.
- 493 Gutenkunst RN, Hernandez RD, Williamson SH *et al.* (2009) Inferring the Joint
494 Demographic History of Multiple Populations from Multidimensional SNP
495 Frequency Data. *PLoS Genet*, **5**, e1000695.
- 496 James TY, Kauff F, Schoch CL *et al.* (2006) Reconstructing the early evolution of Fungi
497 using a six-gene phylogeny. *Nature*, **443**, 818–822.
- 498 Kretzer AM, Dunham S, Molina R *et al.* (2005) Patterns of vegetative growth and gene
499 flow in *Rhizopogon vinicolor* and *R. vesiculosus* (Boletales, Basidiomycota). *Mol*
500 *Ecol*, **14**, 2259–2268.
- 501 Kuo A, Bushnell B, Grigoriev IV (2014) Fungal genomics: sequencing and annotation.
502 *In: Martin F, editor. Fungi. Advances in Botanical Research*, **70**, Elsevier, 1–52.
- 503 Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat*
504 *Methods*, **9**, 357–359.
- 505 Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and
506 SAMtools. *Bioinforma Oxf Engl*, **25**, 2078–2079.
- 507 Liti G, Carter DM, Moses AM *et al.* (2009) Population genomics of domestic and wild
508 yeasts. *Nature*, **458**, 337–341.
- 509 Mboup M, Bahri B, Leconte M *et al.* (2012) Genetic structure and local adaptation of
510 European wheat yellow rust populations: the role of temperature-specific
511 adaptation. *Evol Appl*, **5**, 341–352.
- 512 McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in
513 *Drosophila*. *Nature*, **351**, 652–654.
- 514 McKenna A, Hanna M, Banks E *et al.* (2010) The Genome Analysis Toolkit: a
515 MapReduce framework for analyzing next-generation DNA sequencing data.
516 *Genome Res*, **20**, 1297–1303.

- 517 Mita SD, Siol M (2012) EggLib: processing, analysis and simulation tools for population
518 genetics and genomics. *BMC Genet*, **13**, 27.
- 519 Muller LAH, Vangronsveld J, Colpaert JV (2007) Genetic structure of *Suillus luteus*
520 populations in heavy metal polluted and nonpolluted habitats. *Mol Ecol*, **16**,
521 4728–4737.
- 522 Nass R, Cunningham KW, Rao R (1997) Intracellular sequestration of sodium by a novel
523 Na⁺/H⁺ exchanger in yeast is enhanced by mutations in the plasma membrane
524 H⁺-ATPase. *J Biol Chem*, **272**, 26145–26152.
- 525 Neafsey DE, Barker BM, Sharpton TJ *et al.* (2010) Population genomic sequencing of
526 *Coccidioides* fungi reveals recent hybridization and transposon control. *Genome*
527 *Res*, **20**, 938–946.
- 528 Noor MAF, Bennett SM (2009) Islands of speciation or mirages in the desert? Examining
529 the role of restricted recombination in maintaining species. *Heredity*, **103**, 439–
530 444.
- 531 Peay KG, Garbelotto M, Bruns TD (2010) Evidence of dispersal limitation in soil
532 microorganisms: Isolation reduces species richness on mycorrhizal tree islands.
533 *Ecology*, **91**, 3631–3640.
- 534 Peay KG, Schubert MG, Nguyen NH *et al.* (2012) Measuring ectomycorrhizal fungal
535 dispersal: macroecological patterns driven by microscopic propagules. *Mol Ecol*,
536 **21**, 4122–4136.
- 537 Rivera Y, Burchhardt KM, Kretzer AM (2014) Little to no genetic structure in the
538 ectomycorrhizal basidiomycete *Suillus spraguei* (Syn. *S. pictus*) across parts of
539 the northeastern USA. *Mycorrhiza*, **24**, 227–232.
- 540 Rosenblum EB, James TY, Zamudio KR *et al.* (2013) Complex history of the amphibian-
541 killing chytrid fungus revealed with genome resequencing data. *Proc Natl Acad*
542 *Sci*, **110**, 9385–9390.
- 543 Roy M, Dubois M-P, Proffit M *et al.* (2008) Evidence from population genetics that the
544 ectomycorrhizal basidiomycete *Laccaria amethystina* is an actual multihost
545 symbiont. *Mol Ecol*, **17**, 2825–2838.
- 546 Schechter S, Branco S (2014) The Ecology and Evolution of Mycorrhizal Fungi in
547 Extreme Soils. In: Rajakaruna N, Boyd R, Harris T (eds). *Plant Ecology and*
548 *Evolution in Harsh Environments*. Nova Publishers.
- 549 Smith AH, Thiers HD (1964) A contribution towards a monograph of North American
550 species of *Suillus*. Published by the authors, Ann Harbor, MI.
- 551 Smith SE, Read DJ (2010) *Mycorrhizal Symbiosis*. Academic Press.
- 552 Talbot JM, Bruns TD, Taylor, JW *et al.* (2014) Endemism and functional convergence
553 across the North American soil mycobiome. *Proc Natl Acad Sci*, **111**, 6341–6346.
- 554 Tang H, Peng J, Wang P *et al.* (2005) Estimation of individual admixture: Analytical and
555 study design considerations. *Genet Epidemiol*, **28**, 289–301.
- 556 Taylor JW, Turner E, Townsend JP *et al.* (2006) Eukaryotic microbes, species
557 recognition and the geographic limits of species: examples from the kingdom
558 Fungi. *Philos Trans R Soc B Biol Sci*, **361**, 1947–1963.
- 559 Vacher C, Vile D, Helion E *et al.* (2008) Distribution of parasitic fungal species richness:
560 influence of climate versus host species diversity. *Divers Distrib*, **14**, 786–798.
- 561 Zeng K, Fu Y-X, Shi S *et al.* (2006) Statistical Tests for Detecting Positive Selection by
562 Utilizing High-Frequency Variants. *Genetics*, **174**, 1431–1439.

563 Zhou Z, Miwa M, Hogetsu T (2001) Polymorphism of simple sequence repeats reveals
 564 gene flow within and between ectomycorrhizal *Suillus grevillei* populations. *New*
 565 *Phytol*, **149**, 339–348.

566
 567
 568
 569
 570
 571
 572
 573
 574
 575

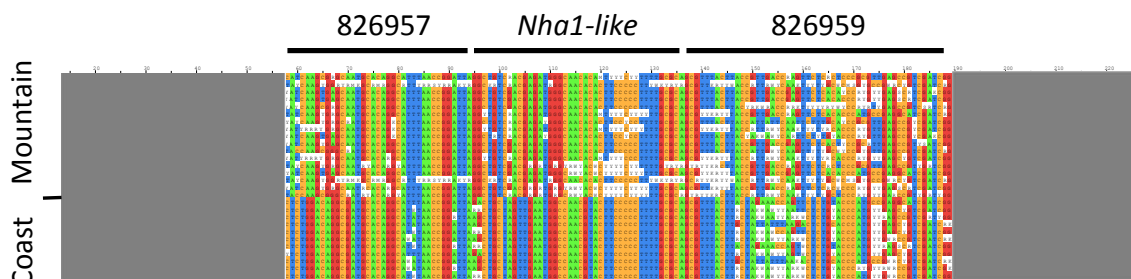
Figures and figure legends:



576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586

Figure 1 – Evidence for population structure between the coastal and montane *Suillus brevipes* populations. a) Neighbor-joining tree based on 567192 SNPs after singleton exclusion; green branches depict individuals from the mountains (Yosemite National Park, CA) and blue branches refer to coastal individuals (Sb20 was collected from Pt Reyes, Sb21 from Salt Point, and all others from Mendocino Co, CA); Sb7 and Sb10 are clones. * indicates bootstrap support values of 100. b) FRAPPE diagrams showing genetic differentiation between coastal and montane individuals.

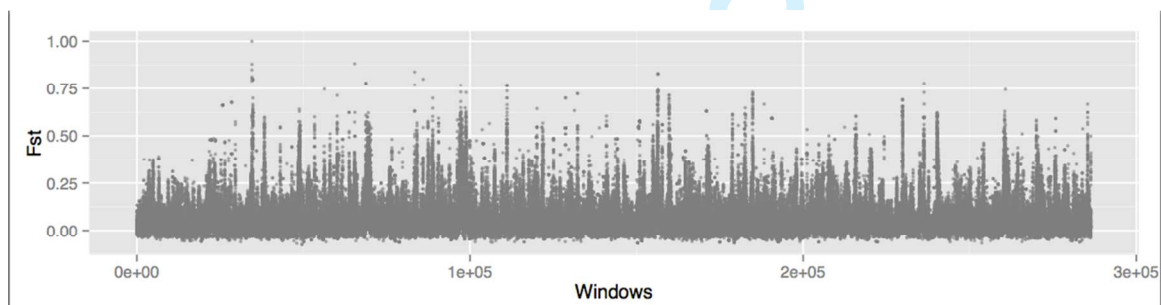
587
588
589
590
591
592



593
594

595 Figure 2 – Genomic regions of divergence genotype matrices. Columns depict
596 polymorphic sites and rows the genotypes for the different individuals. The shaded areas
597 represent flanking regions surrounding the divergence outliers and non-colored bases are
598 heterozygous sites. The location of introns is not shown. Individuals are grouped by
599 population of origin (coast and mountain). The matrix shows three adjacent genes
600 including 20% of genic reciprocally fixed SNPs in scaffold 12. 826959 codes for a
601 cysteine-type endopeptidase, *Nha1-like* (826958) for a Na⁺/H⁺ exchanger, and 826957
602 has an unknown function. These three genes are both Dxy and Fst outliers (see text for
603 details).

604
605
606
607



608
609

610 Figure 3 – Average Fst values for 1kb windows across the *S. brevipes* genome.

611
612
613
614
615
616
617
618

619
620
621
622
623
624
625
626
627
628
629
630

Tables

Table 1 - Genomic summary statistics for the montane and coastal *Suillus brevipes* populations with number of segregating sites, nucleotide diversity (π), π for silent sites, Tajima's D, and Zeng's E for 5Kb genomic windows and genes.

	Mountain (5kb windows)	Mountain (genes)	Coast (5kb windows)	Coast (genes)
Segregating sites	52.2	10.7	48.8	12.7
π	2.7×10^{-3}	2.1×10^{-3}	3.2×10^{-3}	2.0×10^{-3}
π Silent	-	5.3×10^{-3}	-	5.0×10^{-3}
π non-Silent	-	1.2×10^{-3}	-	1.1×10^{-3}
Tajima's D	-0.15	0.31	-0.4	0.36
Zeng's E	-	-2.01	-	-1.94

631
632
633
634
635

Table 2 – Number and percentage of polymorphic and fixed SNP between the coastal and mountain populations. (Total number of SNPs = 1254275).

	Number of SNP	Percentage
Fixed in both populations	155	0.01%
Polymorphic in both populations	701640	44.0%
Fixed in mountain only	410696	32.7%
Fixed in coast only	141784	11.3%

636
637
638
639
640
641
642
643

Table 3 - Optimized log-likelihoods for demographic models as provided by $\partial a \partial i$ (using the unfolded site frequency spectra). Tested models included isolation, migration, migration and exponential growth, migration with two epochs, and migration with two epochs and exponential growth. The isolation model is the most likely model and is highlighted in bold.

Model	Optimized log-likelihood	Theta
Isolation	-32595.1	135137.6

Migration	-76454.7	112079.5
Migration/Exponential	-62192.6	17137.5
Migration 2 epochs	-141888.1	20751.7
Migration 2 epochs/ Exponential	-117173.2	21888.5

644

645

646 Table 4 - Maximum-likelihood demographic parameter estimates for the isolation model
 647 to the joint allele frequency spectrum for the mountain and coastal *S. brevipes*
 648 populations, with respective means and standard deviation (SD). $\mu=10^{-9}$; divergence time
 649 in generations.

650

Parameter	No Migration model		
	Estimate	Mean	SD
Ancestral N_e	6.53×10^5	4.1×10^5	608.1
Mountain N_e after split	3.87×10^5	2.3×10^5	3824.1
Coast N_e after split	1.16×10^5	7.3×10^4	1353.9
Divergence time	2.5×10^4	1.6×10^4	291.5

651

652 Table 5 – 1kb window Fst gene 0.1% outliers, with localization in genome and protein
 653 function (with respective description) if available.

654

Scaffold	Protein ID	Function	Description	Obs
2	763737	Translation elongation factor		
2	948321	Proton-transporting ATPase activity, rotational mechanism		
4	713130	Transporter activity	MFS general substrate transporter	
8	814720	Purine ribonucleoside monophosphate biosynthetic process	Adenosine/AMP deaminase	
8	853122	GTPase activity	Translation elongation factor EFG/EF2, C-terminal	
8	814728	Catalytic activity	Alcohol dehydrogenase, zinc-binding	
9	860915	Sequence-specific DNA binding transcription factor activity	Zinc finger, GATA-type	

12	826957	-		Dxy outlier; gene with fixed differences
12	826958	Regulation of pH, solute:hydrogen antiporter	Sodium/hydrogen exchanger	Dxy outlier; gene with fixed differences
12	826959	Cysteine-type endopeptidase	Peptidase C1B, bleomycin hydrolase	Dxy outlier; gene with fixed differences
14	718701	Glucose-6-phosphate dehydrogenase		
16	795305	Nucleotide binding, nucleotide-excision repair		
22	862126	Oxidoreductase activity	Pyridine nucleotide- disulphide oxidoreductase, class I	
32	830321	tRNA processing		
32	830324	Oxidoreductase activity	Aldo/keto reductase	
41	831287	Catalytic activity	Phospholipase D/Transphosphatidylase	
41	855962	Protein modification process	Ubiquitin	
42	831382	Catalytic activity	Serralysin-like metalloprotease, C- terminal	
63	733034	Transporter activity	MFS general substrate transporter	
66	804723	L-serine ammonia- lyase activity	Iron-sulphur-dependent L-serine dehydratase single chain form	Dxy outlier
66	779351	DNA ligase (ATP) activity, DNA replication, DNA repair		
66	733526	Nucleic acid binding, zinc ion binding		
85	834749	Two-component response regulator activity		

123 783492 Transferase activity
nitrogenous groups

655

656

657

658 **Data Accessibility**

659

660 DNA sequences : Short Read Archive accessions in table S1

661 *Suillus brevipes* reference genome and annotation:

662 http://genome.jgi.doe.gov/Suillus_brevipes/

663 *S. brevipes* raw single nucleotide polymorphism list : Dryad doi XXXXX

664 *S. brevipes* neighbor-joining tree : Dryad doi XXXXX

665

666

667 **Author Contributions**

668

669 SB, TDB and JWT designed the study and collected samples. SB, PG and CEE

670 performed analyses. AK, KL, AL and IVG assembled and annotated the *Suillus brevipes*

671 reference genome. SB, PG, CEE, H-L L, RV, KGP, JWT and TDB wrote the paper.

672

673

674 **Supporting Information**

675

676 Table S1 - *Suillus brevipes* isolate location, habitat, pine host, coordinates and
677 elevation, number of raw and high quality reads, % alignment rate, genome coverage, and
678 Short Read Archive accession numbers. * Sb2 was the isolate used to generate the
679 reference genome for the species (<http://genome.jgi-psf.org/Suibr1/Suibr1.home.html>).
680 See text for details.

681

682 Table S2 - 5kb window Dxy gene 1% outliers, with localization in genome and
683 protein function and description. Genes in scaffold 12 include 20% of the reciprocally
684 fixed differences falling in genic regions and gene 828720 was found to under positive
685 selection by McDonald Kreitman test.

686

687 Table S3 - Genes under positive selection as inferred by McDonald-Kreitman tests
688 on genes using *S. luteus* as the outgroup. *P* values were adjusted using the Benjamini-
689 Hochberg correction.

690

691 Table S4 – Selective sweep analysis; top five H value peaks for coastal and mountain
692 *Suillus brevipes* populations.

693

694 Figure S1 – Observed and estimated site frequency spectra under the isolation model
695 from coastal and montane *Suillus brevipes* populations.

696

697