HRM analysis provides insights on the reproduction mode and the population structure of Gnomoniopsis castaneae in Europe

HRM analysis provides insights on the reproduction mode and the population structure of *Gnomoniopsis castaneae* in Europe

F. Sillo[a], L. Giordano[a,b], E. Zampieri[a], G. Lione[a], S. De Cesare[a], P. Gonthier[a]*

[a] Department of Agricultural, Forest and Food Sciences, University of Torino, Largo Paolo Braccini 2, I-10095 Grugliasco (TO), Italy.

[b] Centre of Competence for the Innovation in the Agro-Environment Field (AGROINNOVA), University of Torino, Largo Paolo Braccini 2, I-10095 Grugliasco (TO), Italy.

*: paolo.gonthier@unito.it

Running title: *Gnomoniopsis castaneae* population genetics

Keywords: *Gnomoniopsis castaneae, Gnomoniopsis smithogilvyi*, SSR, HRM, population genetics, genetic structure

Abstract

Gnomoniopsis castaneae ("*castanea*") is an emergent nut rot agent of chestnut in Southern Europe. To elucidate its population genetics, three Simple Sequence Repeat (SSR) and two hypervariable markers were developed and assessed through the High Resolution Melting (HRM) analysis on 132 isolates collected from ten sites in Italy, France and Switzerland. High allele diversity (ranging from 0.23 to 0.40 depending on site) and number of haplotypes (49) were observed. More than 70% of the molecular variance could be accounted among isolates within sites. Multilocus analysis showed absence of linkage disequilibrium, suggesting a predominant role played by sexual reproduction and random mating. Data analyses indicated the presence of at least two putative distinct subpopulations and this was confirmed by several approaches, including analysis of shared haplotypes, multivariate and Bayesian analyses. Based on data of allelic diversity, the possibility that the pathogen could have been introduced is discussed. This work assessed the genetic variability and the sexual strategies of *G. castaneae* in Europe, adding useful information on the epidemiology of this fungal plant pathogen.

Introduction

The European chestnut (*Castanea sativa* Mill.) faced several relevant diseases and pests in last centuries, including the ink disease caused by *Phytophthora* spp., the blight caused by *Cryphonectria parasitica* (Murrill) M.E. Barr and, more recently, the Asian gall wasp *Dryocosmus kuriphilus* Yasumatsu (Gonthier & Ferracini 2014). In this complex phytosanitary scenario, a relevant role is also played by fungi causing nut rot in pre- and/or post-harvest, resulting in yield and economic losses (Lione *et al*. 2015). The ascomycete *Gnomoniopsis castaneae* ("*castanea*") G. Tamietti (Visentin *et al*. 2012; Tamietti 2016) is an emergent nut rot agent in several areas of Europe, including Italy, France and Switzerland (Visentin *et al*. 2012; Maresi *et al.* 2013; Dennert *et al.* 2015; Pasche *et al.* 2016). Shuttleworth *et al.* (2015) recently demonstrated the synonymy between *G. castaneae* and *G. smithogilvyi* L.A. Shuttlew., E.C.Y. Liew & D.I. Guest, a fungus reported in Australia and New Zealand (Shuttleworth *et al.* 2012). The symptoms of the nut rot caused by *G. castaneae* include a chalky aspect of the nut kernel at ripening, turning to brown as soon as the mummification advances and the mycelium occupies the kernel tissues (Shuttleworth *et al.* 2012; Visentin *et al*. 2012). While the fungus was recently reported in association with bark cankers (Pasche *et al.* 2016), it can be easily isolated as an endophyte from buds, from the thin bark of young shoots (Visentin *et al*. 2012), and from the galls produced on chestnut by the Asian gall wasp, where fruiting bodies of the anamorphic stage in the form of acervuli may be found (Maresi et al. 2013). The perithecia producing teleomorphic stage can develop in the spring on the burrs (Visentin *et al*. 2012). Airborne infections may occur through flowers, as experimentally documented (Visentin *et al*. 2012). Disease incidence is better explained by warmer temperatures of months preceding harvesting rather than by rainfalls (Lione *et al.* 2015) and the disease is randomly distributed within orchards (Lione & Gonthier 2015). With the exception of this, very little is known about the infection biology and epidemiology of this pathogen, and even less on its population genetics. Moreover, the relative contributions of sexual and asexual reproduction to the

current distribution of *G. castaneae* populations are poorly studied. By combining haplotype network of the sequences of calmodulin and β-tubulin genes, Dennert *et al*. (2015) investigated the geographical distribution of different haplotypes of *G. castaneae* in Switzerland. They found five haplotypes, providing for the first time evidence of the genetic variation within the species. However, a broader sampling and molecular genetic markers with higher level of polymorphisms, easy to score and with better resolution, are still needed for studying in greater details the genetic structure of this fungal species. Microsatellites or Simple Sequence Repeats (SSRs) are single sequence motifs comprising no more than six bases that are tandem repeated (Leišová-Svobodová *et al.* 2014). They are popular as genetic markers due to their high reproducibility, multiallelic nature, codominant way of inheritance, abundance and wide genome coverage, and they have demonstrated to be helpful for genetic characterization of fungi (Garbelotto *et al*. 2013; Leišová-Svobodová *et al.* 2014; Gonthier *et al.* 2015). Two steps may be critical for the use of SSR markers in population genetics studies. First, loci harboring SSRs must be detected and isolated. When a genome or at least sequence libraries, such as Expressed Sequence Tags (ESTs), are available in public databases, this issue could be easily addressed; otherwise, isolation of *de novo* SSR loci is needed (Dutech *et al*. 2007). The second crucial step is represented by the assessment of the SSR polymorphisms, which generally requires the availability of capillary sequencers (Ganopoulus *et al.* 2011). Recently, a fast and sensitive polymorphism detecting method based on quantitative PCR (qPCR) technologies and called High Resolution Melting (HRM) analysis has been introduced as an alternative technique to investigate SSR polymorphisms (Mackay *et al.* 2008; Ganopoulus *et al*. 2011). HRM analysis is indeed very sensitive in genotype scanning and quickly finds small DNA sequence variants, *e.g*., Single Nucleotide Polymorphisms (SNPs), exploiting the fact that PCR products with different sequences have distinct melting profiles (Luchi *et al*. 2011). The sensitivity of HRM analysis has been deeply assessed (Reed & Wittwer 2004). HRM analysis has been applied in several different research areas, including the genotyping of plant pathogenic bacteria (Gori *et al.* 2012) and in studying the genetic variability in plants (Wu *et al*. 2008). Recently, this technique has

107 also been coupled with the analysis of SSRs identified in citrus (Distefano *et al.* 2012), in
108 grapevine/olive cultivars (Mackay *et al.* 2008) and in sweet cherry products (Ganopoulos *et al.*
109 2012), as well as in other organisms such as bacteria (Ricchi *et al.* 2011). To our knowledge, there
110 are no examples of HRM analysis in genotyping of fungal pathogens.
111 Here, we report on HRM analysis-based genotyping of *G. castaneae* isolates collected in Italy,
112 Switzerland and France in order to provide a first glimpse on the epidemiology of this pathogen. In
113 addition, we assessed the presence of linkage disequilibrium among analysed loci. The analysis of
114 linkage disequilibrium of alleles has been extensively used to infer the preferred mode of
115 reproduction (sexual/asexual) of pathogenic fungi (Linde *et al.* 2003; Stergiopoulos *et al.* 2007).
116 Here, we test the hypothesis that populations of *G. castaneae* may have a mixed reproduction mode
117 in nature, by exploring the balance between sexual and asexual reproduction. In particular, the
118 specific aims of this study were: I) to develop molecular markers for *G. castaneae*, for which the
119 complete genome sequence is not available; II) to assess their polymorphisms by using HRM
120 analysis in order to perform a population genetics analysis on isolates collected from the three
121 European countries where *G. castaneae* was reported; and III) to provide insights on the preferred
122 reproduction mode in populations of *G. castaneae* through analysis of linkage disequilibrium in
123 different sampling sites.
124
125 **Material and Methods**
126
127 **Samplings and isolation of *G. castaneae***
128 Forty nuts were randomly collected at the beginning of November 2011 from each of seven sites in
129 North-western Italy, one in South-eastern France and two in Southern Switzerland, and stored at
130 4°C before the subsequent analysis (Figure 1 and Table 1). In order to isolate *G. castaneae,* five
131 fragments *per* nut (approximately 1×1×2 mm in size) were excised and plated in 9 cm diameter
132 Petri dishes filled with malt extract agar (MEA; 31.3 g malt extract agar, 1 L distilled water) in

sterile conditions as previously described (Visentin *et al*. 2012). Putative colonies of *G. castaneae* were identified by examining macro- and micromorphological features including both the aspect of mycelium, acervuli, the shape and size of conidia, and by molecular identification through the specific primer set developed by Lione *et al.* (2015). Disease incidence, expressed as the percentage of infected nuts on the total number of nuts collected in each site, is reported in Table 1. One hundred and thirty two isolates of *G. castaneae* were obtained (Table 1; for the full list see Table S1).

DNA extraction

All isolates were grown in the dark and in agitation in flasks containing malt extract 2% liquid medium, at 25°C for seven days. Subsequently, mycelia were collected using a vacuum pump, frozen at -80°C and dry lyophilized overnight. Total DNA extraction was performed using a modified CTAB-based method. Briefly, lyophilized mycelium (approximately 50 mg) was homogenized in a 1.5 ml microcentrifuge tube with a pestle, resuspended in 600-800 µL of a CTAB extraction buffer (100 mM TrisCl pH 8.4, 1.4 M NaCl, 25 mM EDTA pH 8.0, 2% CTAB) and incubated at 65°C for 30 minutes. After the extraction step, an equal volume of chloroform/isoamyl alcohol (24:1) was added to each tube, vortexed, and then spinned for 10 minutes at 12300 g in a microcentrifuge. The upper phase was transferred to a 1.5 ml microcentrifuge tube where the DNA was precipitated by the addition of 600 µL of cold isopropanol and by a centrifugation at 12300 g for 5 minutes. Subsequently, the supernatant was discarded and the DNA pellet was gently washed with 70% ethanol and resuspended in 100 µL TE buffer (10 mM TrisCl pH 8.0.1 mM EDTA pH 8.0) by heating at 65°C for about 30 minutes.

SSR isolation

Genomic libraries enriched in SSRs were prepared using a MAL (Microsatellite Amplified Library) approach as described by Acquadro *et al.* (2005). Five SSR motifs (AAG, AG, AC, CCG and AAA)

were chosen to perform SSR enrichment on the basis of their relevance in fungal genomes as previously reported (Karaoglu *et al*. 2005). One genomic library enriched in SSRs was generated for each motif. Briefly, the DNA (2.5 µl) of the isolate BOF25 (Table S1) was digested with *MseI* (0.25 µl) at 37°C for 120 minutes, followed by 20 minutes at 65°C. The digested product was then purified by E.Z.N.A. Gel Extraction kit (Omega Bio-Tek). The *MseI/SalI* adaptors were incubated at 98°C for 5 minutes and then kept at room temperature for 90 minutes. The adaptors were ligated to the digested DNA (2.5 µl) by T4 ligase (2 µl, Invitrogen) at 14°C for 24 hours. In order to isolate the SSRs, five PCR reactions were carried out with a common primer (Primer *MseI/SalI*) and a primer specific for each microsatellite (Table S2), following the programme: 94°C for 3 minutes, 94°C for 1 minute, 55°C for 1 minute, 72°C for 1 minute, repeated for 45 cycles, and the 72°C for 10 minutes. The reaction mix was made up of: water (8.875 µl), 5X buffer (4 µl), 15mM $MgCl_2$ (1 µl), 2 mM dNTPs (2 µl), 5 µM specific primer forward and common reverse primer (1 µl), 6.25 U GoTaq Polymerase (0.125 µl, Promega) and DNA (2 µl).

The PCR products were purified by E.Z.N.A. Gel Extraction kit (Omega Bio-Tek) and digested with *SalI* (1 µl) at 37°C for 120 minutes, followed by 20 minutes at 65°C. After this step, the digested amplified products were purified again with the above kit and ligated by T4 Ligase (Invitrogen, 0.5 µl) to the plasmid vector pUC19 (8 µl) at 25°C for 60 minutes. The ligated products were cloned in Top10 chemically competent cells of *Escherichia coli* (Life Technologies) following the manufacturer's instructions. Bacterial colonies were screened by PCR using M13FW and M13RV primers, following the programme: 94°C for 5 minutes, 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 45 seconds, repeated for 35 cycles, and 72°C for 7 minutes. All amplified products were run on agarose gel to select amplicons to sequence. The amplified products were digested by ExoSAP-IT (Affymetrix) at 37°C for 15 minutes and then at 80°C for 15 minutes. The products were sequenced by BMR Genomics S.R.L. (Padua, Italy).

The design of forward primers specific for each SSR was performed by using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/). The primers were tested in PCR with the common reverse primer (*i.e.,* Primer *MseI/SalI*) on the ligated *MseI/SalI* DNA, with the PCR mix and the programme described above. The PCR products were digested by ExoSAP-IT (Affymetrix) as previously described and sequenced. Specific reverse primers were designed for each sequenced SSR locus by using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/). Each primer pair specific for one SSR was tested in PCR on DNA of three isolates of *G. castaneae* (VF5, CHB2, BSD14; Table S1). The presence of amplification products was checked by agarose electrophoresis. All primers used in this work are reported in Table S2.

HRM analysis

The primer pairs were tested in HRM analysis on all isolates to identify their polymorphisms. The qPCR for the HRM analysis was carried out with Connect™ Real-Time PCR Detection System (Bio-Rad Laboratories). Each PCR reaction was conducted on a total volume of 10 µl, containing 1 µl diluted DNA (dilution 1:50), 5 µl Sso Fast Eva Green Supermix (Bio-Rad Laboratories), 0.3 µl of each primer (3 µM) and 3.4 µl of water, using a 96 well plate. The following PCR programme, which includes the calculation of a melting curve, was used: 98°C for 2 minutes, 45 cycles of 98°C for 5 seconds, 60°C for 10 seconds, ramp from 65°C to 95°C with a temperature increment of 0.1°C and a read plate every 10 seconds. Melting curves were analysed by using the Precision Melt Analysis™ Software from Bio-Rad, setting the Tm difference threshold = 0.15 and the Melt curve shape sensitivity = 50. This software allowed to group the melting curves in different clusters, representing different alleles of analysed SSR loci. Clusters identified by HRM analysis were confirmed by sequencing of the PCR products of two representative samples *per* cluster. The PCR products were digested by ExoSAP-IT (Affymetrix) at 37°C for 15 minutes and subsequently at 80°C for 15 minutes. The PCR products were sequenced by BMR Genomics S.R.L. (Padua, Italy).

209

210 **Population genetics analysis**

211 By analysing the melting curves, alleles were assigned to all isolates for each locus and a matrix
212 including all the allelic data was prepared (Table S1). The data matrix was analysed using
213 POPGENE version 1.32 software package
214 (https://www.ualberta.ca/~fyeh/popgene_download.html). Number of alleles *per* locus, allele
215 frequencies *per* locus and number of private alleles *per* site were estimated using GenAlEx version
216 6.5 (Peakall & Smouse 2012). The total gene/allele diversity (Ht), the genetic diversity within sites
217 (Hs) and Nei's coefficient of genetic differentiation among sites (Gst) were calculated, as well as
218 pairwise population matrix of Nei genetic identity. The matrix was used as input to construct a
219 dendrogram by using DendroUPGMA (http://genomes.urv.cat/UPGMA/).
220 The linkage disequilibrium was assessed by using both POPGENE and MultiLocus 1.2
221 (http://www.bio.ic.ac.uk/evolve/software/multilocus/). The observed linkage disequilibrium was
222 compared to expected distributions of linkage among loci from 1,000 permutations using the index
223 of association (I_A).
224 Analysis of molecular variance (AMOVA) was conducted to estimate the distribution of genetic
225 variation among and within sites, using GenAlEx version 6.5 (Peakall & Smouse 2012). In order to
226 explore the genetic relationships among isolates of each site, Principal Coordinate Analysis (PcOA)
227 and Discriminant Analysis (DA) were performed by using the software XLSTAT
228 (http://www.xlstat.com/en/home/). Bayesian analysis as implemented by the software
229 STRUCTURE v2.3.4 (Pritchard *et al.* 2000) was used to investigate the population structure and to
230 assign individuals into subpopulations. Successive K value (number of populations) from 1 to 10
231 was used to obtain the distinct clusters and to estimate the number of subpopulations. Twenty runs
232 each for K = 1–10 with 750000 MCMC repetitions after a burn-in period of 500000 repetitions
233 were performed with the "Admixture Model" option and without any prior information as the origin
234 (location) of individual samples. The K value that best represented the observed data under the

model implemented was inferred by determining the ΔK based on the highest likelihood of the data (LnP(D)) (Evanno *et al.* 2005). In order to confirm the status as single species, two randomly selected isolates *per* subpopulation *per* site were sequenced in their Internal Transcribed Regions (ITS), elongation factor 1-α and β-tubulin by using primers reported in Table S2. The sequences were aligned by using MEGA v. 6.0 (Tamura *et al.* 2013). Clonal fractions were calculated using the following formula: clonal fraction (%)=[1−(number of haplotypes / total number of isolates in sampling site)]×100 (Stukenbrock *et al.* 2006). In addition, in order to explore the relationships among genetic and geographical distances, both Spearman's rank correlation test and Mantel test were performed. The Spearman's rank correlation test was also used to assess the association between geographical distance and allelic diversity for each subpopulation, starting from sites harbouring the highest allelic diversity.

Results

SSR isolation

Both morphological and molecular identification performed with taxon-specific primers on putative colonies confirmed their status as *G. castaneae*. Nineteen SSR loci were isolated in *G. castaneae* through the MAL approach. Of these, fifteen resulted PCR-amplifiable, and among them, five were found to be polymorphic. In detail, polymorphic loci were Gc_AAA60, Gc_AAG8, Gc_AAG57, Gc_AG26 and Gc_CCG42. A list of primers to amplify the identified loci is reported in Table S2.

HRM analysis

Based on the analysis of the melting curves, the locus Gc_AAA60 resulted in six different alleles, Gc_AAG8 in four alleles, Gc_AAG57 in five alleles, and Gc_AG26 in four alleles. The locus Gc_CCG42 resulted in four alleles based on HRM analysis and the further sequencing. Sequencing of alleles of Gc_AAA60 and Gc_CCG42 showed that they did not contain pure SSR motifs but

rather homopolymeric tracts that were considered as hypervariable loci. Melting curve difference plots (after normalization and overlay) for the five polymorphic loci are displayed in Figure 2.

Population genetics analysis

Genotyping detected 49 haplotypes and allowed to assess their allelic diversity. The two most frequent haplotypes included about 28% of the isolates, whereas 32 out of 49 haplotypes were unique and found in specific sites (Figure 3). The overall mean number of alleles *per* locus was 4.6, with a total of 23 alleles identified across all five loci (Table 1). Frequencies of private alleles were very low, ranging from 0.0 to 0.20 in Donato and Cadenazzo (Table 1). The overall Nei's gene diversity (Ht) *per* locus ranged from 0.32 (Gc_AAG57) to 0.68 (Gc_CCG42), with an average value of 0.52 (± 0.03). The values of Hs for each locus in each site are reported in Table 2. The coefficient of genetic differentiation among all sites (Gst) was 0.37 (Table 2). Pairwise matrix of Nei genetic identity (Table 3) showed that the highest similarity was between Biasca and Donnas (0.99) while the lowest was between Donnas and Mattie (0.29). The dendrogram based on this matrix allowed to distinguish two different clusters (Fig. S1).

Concerning linkage disequilibrium analysis, by using POPGENE the percentage of locus combinations that were significantly different from equilibrium expectations ($\chi 2$ test, p-value < 0.05) ranged from a minimum of 0.47% (1/210 allelic pairs) in Donato and Cadenazzo to a maximum of 3.8% (8/210 allelic pairs) in Peveragno. Five sites (Biasca, Boves, Donnas, Sisteron, Villarfocchiardo) showed an absence of significant linkage disequilibrium (p-value > 0.05) (Table 1). Multilocus analysis showed absence of significant linkage disequilibrium in all sites (p-value > 0.05) (Figure S2). The I_A calculated for each site ranged from -0.44 in Donnas to 0.55 in Borgo San Dalmazzo (Table 1).

AMOVA results including all loci indicated that the genetic variability within-site was about 71% (variance 0.90, degree of freedom 122), while the genetic variance among sites was about 29% (variance 0.35, degree of freedom 5). Both PCoA and DA clearly showed two distinct groups, one

287 including Biasca, Boves, Cadenazzo, Donnas, Peveragno and Sisteron, and the other comprising
288 Borgo San Dalmazzo, Donato, Mattie and Villarfocchiardo (Fig. 4). Based on STRUCTURE
289 analysis, the optimal number of populations (K) as inferred by evaluating the ΔK was two,
290 suggesting that all isolates, with the exception of isolate BOF22 from Boves and isolate CHC16
291 from Cadenazzo, fell into one of the two genetically distinct clusters hereafter referred to as
292 subpopulations (Figure 5; STRUCTURE runs with K=3 and K=4 are shown in Figure S3). The first
293 subpopulation (Subpopulation 1) included Biasca, Boves, Cadenazzo, Donnas, Peveragno, Sisteron,
294 while the second subpopulation (Subpopulation 2) included the remaining sites. No significant
295 correlation between pairwise Gst and geographical distance among sites was detected (Spearman's
296 rank correlation coefficient = -0.17, *p*-value > 0.05), and this was confirmed by the Mantel test
297 between the two matrices (*p*-value < 0.05). On the basis of the STRUCTURE analysis, a
298 Spearman's rank correlation was used to test the association between geographical distance from
299 Peveragno and Borgo San Dalmazzo, the two sites with the highest allelic diversity, to the other
300 sites belonging to the same putative subpopulation. Hs was significantly correlated with distance in
301 Subpopulation 1 (rho = -0.89, *p*-value was 0.02), but not in Subpopulation 2 (rho = -0.80, *p*-value
302 was 0.20). The overall allelic diversity (Hs) was 0.33 (± 0.06) and 0.38 (± 0.08) for Subpopulation 1
303 and Subpopulation 2, respectively. The Gst value for Subpopulation 1 was 0.06, while for
304 Subpopulation 2 was 0.16. The status as single species was confirmed by analysing the three loci in
305 twenty isolates from both the Subpopulations. No polymorphisms were detected in the alignment of
306 ITS and elongation factor 1-α sequences, while four SNPs were detected in only one isolate
307 (VDA24) in the aligned β-tubulin sequences. Two representative sequences *per* locus were
308 deposited at EMBL - European Nucleotide Archive (ENA) under the accession numbers
309 LN999963-LN999983. The clonal fraction among *G. castaneae* isolates was over 50% in Donnas
310 only. By contrast, Sisteron, Borgo San Dalmazzo and Mattie showed a reduced clonal fraction (less
311 than 20%) (Table 1).
312

Discussion

The nut rot caused by *G. castaneae* represents a severe threat for sweet chestnut orchards, requiring a better understanding of its ecology, epidemiology, biogeography and infection biology (Lione *et al.* 2015). This research was mainly focused on elucidating population genetics and genetic structure of this fungal pathogen in all European countries where *G. castaneae* was reported so far.

This study reported for the first time the development and the assessment of a set of useful molecular markers in *G. castaneae*. Only five SSR loci out of 19 developed were found to be polymorphic. In addition, two of the polymorphic loci did not contain a typical SSR motif but rather small homopolymeric traits. This finding is in agreement with Dutech *et al.* (2007), which highlighted the low abundance of polymorphic SSRs useful as molecular markers in fungal genomes. It should be noted that isolation of SSRs without *a priori* knowledge of genome sequences is challenging. A whole genome sequencing approach, *i.e.*, Restriction site-associated DNA sequencing, could be pivotal to identify other SSR loci as well as to verify the real ratio between polymorphic/monomorphic SSR loci in *G. castaneae*.

To the best of our knowledge, this is the first article describing the application of HRM analysis for genotyping purpose in fungal pathogens. The genotyping with HRM is a robust and reproducible method, and may be an appropriate alternative to capillary sequencer when loci show a reduced number of alleles (Ricchi *et al*. 2011). In our work, HRM analysis allowed to detect both SSR length and SNP-based polymorphisms in SSR, SSR flanking regions and homoploymeric traits, overcoming the issue related to allelic size homoplasy. The high mutational instability of SSR units coupled with their tendency to allelic homoplasy, which might be caused by convergent or parallel evolution, mutations in the repeat units and Insertions/Deletions (InDels) in the sequence flanking the SSR, has been reported to lead mistakes in molecular genotyping of fungi (McEwen *et al*. 2000; Dettman & Taylor 2004). Moreover, regions flanking SSR markers may be a rich source of polymorphisms in plants and fungi (Mogg *et al*. 2002; Dettman & Taylor 2004). In the current

339 work, although only few alleles *per* locus (on average four alleles) were identified, the detection of
340 small polymorphisms in the analysed loci, i.e. SNPs, demonstrated the sensitivity of this technique
341 for genotyping purposes in fungi. In particular, for four loci, homoplasic alleles characterized by
342 SNPs or InDels were observed by analysing the melting curves and the related sequences.
343 Sequencing of representative samples *per* cluster identified through HRM analysis can thus be
344 considered a fast approach to score sequence polymorphisms.
345 Multilocus genotyping of the five polymorphic loci enabled to detect 49 haplotypes. This finding
346 demonstrates the high resolution of the approach we used, especially if evaluated in comparison to
347 the previous studies on haplotype diversity that detected only five haplotypes by analysing three
348 loci (Dennert *et al.* 2015). Allele diversities were moderately high (0.23-0.71, average 0.32) among
349 isolates within sites. AMOVA results confirmed that the majority of the variation (more than 70%)
350 could be accounted among isolates within sites. Inter-site identity was high, with pairwise
351 differences among the sites ranging from 0.29 to 0.99 (average 0.68). Two main factors may have
352 contributed to increase allele and haplotype diversity: recurrent gene flow among sites leading to
353 arrival of new alleles from neighbouring sites and/or sexual recombination (McDonald & Linde
354 2002). This last factor not only has a great impact on haplotype diversity, but can also increase
355 allele diversity through intragenic recombination that can create new alleles. In the current work,
356 sites with high allele diversity (H_s > 0.40) were Borgo San Dalmazzo and Peveragno, the
357 southernmost sites sampled in Italy. The presence of a high number of haplotypes could be
358 consistent with the hypothesis that *G. castaneae* is native to Europe, but for reasons yet unknown
359 may have recently re-emerged. However, a high number of haplotypes may also be expected in the
360 case of introduced invasive fungal plant pathogens, provided they reproduce sexually (Garbelotto *et*
361 *al.* 2013).
362 Linkage disequilibrium was non-significant or very low in most of the sampling sites, indicating
363 that random mating sexual reproduction may occur. Multilocus analysis confirmed this finding.
364 Overall genetic diversity, absence of linkage disequilibrium and high differentiation within

365 populations might be indicative of a dominant role played by sexual reproduction in *G. castaneae.*
366 This hypothesis was also supported by the 1:1 ratio of presence/absence of one mating type
367 idiomorph (MAT1-2) in each sampling site (De Cesare 2013). On the other hand, the occurrence of
368 repeated haplotypes in some sites like Donnas, may indicate a persistence of clonal propagation.
369 Asexual reproduction in this site might have been favoured by high numbers of *D. kuriphilus* galls
370 where acervuli are reported to develop (Maresi *et al.* 2013). A definitive explanation for the reason
371 why asexual reproduction is operating at some sites requires further investigations.
372 Interestingly, STRUCTURE analysis showed that, with only a few exceptions, each isolate may be
373 assigned to one of two different subpopulations, and that subpopulations are mutually exclusive
374 present in the sampled sites: one subpopulation is present in Borgo San Dalmazzo, Donato, Mattie
375 and Villarfocchiardo, while the other one in the remaining sites. The distinction in two
376 subpopulations was strongly supported by multivariate analysis i.e., DA and PCoA. The result was
377 also confirmed by converting the pairwise identity matrix into a dendrogram, which clearly allowed
378 to separate the two clusters. STRUCTURE analysis also showed that one isolate from Boves and
379 one from Cadenazzo could not be clearly assigned to one of the two subpopulations, but rather they
380 seemed admixed genotypes. The presence of these two putatively admixed isolates might suggest
381 that recombination between the two subpopulations is possible and is occurring in nature. However,
382 further analyses on additional loci would be needed to clearly and definitely identify admixed
383 isolates, since Bayesian analyses on admixed populations have been reported to be affected by the
384 number of analysed loci (Hansen & Mensberg 2009).
385 It could be hypothesized that the basal area of the two putative subpopulations includes two Italian
386 southernmost sampled sites, Borgo San Dalmazzo and Peveragno. At our genotyping resolution,
387 these sites harboured the highest allele diversity, expected in the epicentre of epidemic processes
388 (Tsui *et al*. 2012). As assessed by Spearman's rank correlation test at least for the Subpopulation 1,
389 moving away from Peveragno the allele diversity decreased, an event which might be expected for a
390 plant pathogenic species invading areas starting from the most diverse sites (Garbelotto *et al*. 2013).

391 In addition, Gst value of Subpopulation 1 was lower than that of Subpopulation 2, suggesting that
392 the isolates of the former might be more closely related. According to this scenario, two
393 subpopulations of the fungus might be present in Europe, one of which (i.e., Subpopulation 1) could
394 have been introduced and subsequently spread from Northwestern Italy to areas as far as France and
395 Switzerland. Alternatively, taking into account that the disease emerged recently and that there is no
396 evidence of the presence of *G. castaneae* before 2005 (Visentin *et al*. 2012), it could be speculated
397 that both subpopulations might have been introduced at different times, Subpopulation 2 before
398 Subpopulation 1.
399 The clear separation between different subpopulations in closely located sites (for instance only a
400 few kilometers separate Borgo San Dalmazzo and Peveragno) and an almost total absence of
401 population admixture may be indicative of recent introduction events. Moreover, the co-occurrence
402 of the same haplotypes in Italy and in Switzerland may support the hypothesis that Italy could be a
403 dissemination source to Switzerland, as suggested by Pasche *et al.* (2016). Although intriguing, a
404 definitive evidence for the introduction of a subpopulation requires a broader sampling and further
405 studies including information on the phylogeography of *G. castaneae*. In fact, the results may have
406 been biased by the reduced number of polymorphisms of analysed loci in the current study.
407 In conclusion, the current work allowed the development and the validation through a powerful
408 technique, *i.e.* HRM analysis, of five molecular markers in *G. castaneae* useful for assessing the
409 genetic diversity. This work highlights the importance of population genetics studies for this
410 pathogen, and further research should be focused on the development of additional molecular
411 markers, also through Next Generation Sequencing approaches. The population genetics analyses
412 on isolates from three different countries in Europe showed high allele and haplotype diversity.
413 Several approaches, such as analysis of shared haplotypes and multivariate/Bayesian analyses,
414 showed the presence of at least two putative distinct subpopulations.
415 Results related to linkage disequilibrium suggest a predominant role played by sexual reproduction
416 in *G. castaneae* populations. Pathogens with sexual reproductive mode have great potential for

417 rapid evolution and should be considered high-risk (McDonald & Linde 2002). From a practical
418 perspective, this finding highlights the importance of management guidelines focused on removing
419 the fallen burrs in chestnut orchards, where the teleomorphic stage has been reported to develop.
420
421 **Acknowledgements**
422 This study was supported by a grant of Regione Piemonte through the activity of the Chestnut
423 Growing Centre. The authors are grateful to the anonymous reviewers that with their suggestions
424 significantly improved the manuscript.
425
426 **References**
427
428 Acquadro A, Portis E, Lee D, Donini P, Lanteri S, 2005. Development and characterization of
429 microsatellite markers in *Cynara cardunculus* L. *Genome* **48**, 217–25.
430
431 Carbone I, Kohn LM, 1999. A method for designing primer sets for speciation studies in
432 filamentous ascomycetes. *Mycologia* **91,** 553–556.
433
434 De Cesare S, 2013. Sviluppo di marcatori molecolari e risultati preliminari sulla genetica di
435 popolazione del fungo fitopatogeno *Gnomoniopsis castanea*. Master thesis, University of Torino,
436 pp. 80.
437
438 Dennert FG, Broggini GAL, Gessler C, Storari M, 2015. *Gnomoniopsis castanea* is the main agent
439 of chestnut nut rot in Switzerland. *Phytopathologia Mediterranea* **54**, 199–211.
440
441 Dettman JR, Taylor JW, 2004. Mutation and evolution of microsatellite loci in *Neurospora*.
442 *Genetics* **168.3**, 1231–1248.

Distefano G, Caruso M, La Malfa S, Gentile A, Wu SB, 2012. High Resolution melting Analysis Is a More Sensitive and Effective Alternative to Gel-Based Platforms in Analysis of SSR – An Example in *Citrus*. *PLoS ONE* **7**, e44202.

Dutech C, Enjalbert J, Fournier E, Delmotte F, Barrès B, Carlier J, Tharreau D, Giraud T, 2007. Challenges of microsatellite isolation in fungi. *Fungal Genetics and Biology* **44**, 933–49.

Evanno G, Regnaut S, Goudet J, 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14,** 2611–2620.

Ganopoulos I, Argiriou A, Tsaftaris A, 2011. Microsatellite high resolution melting (SSR-HRM) analysis for authenticity testing of protected designation of origin (PDO) sweet cherry products. *Food Control* **22**, 532–541.

Garbelotto M, Guglielmo F, Mascheretti S, Croucher PJP, Gonthier P, 2013. Population genetic analyses provide insights on the introduction pathway and spread patterns of the North American forest pathogen *Heterobasidion irregulare* in Italy. *Molecular Ecology* **22**, 4855–4869.

Gardes M, Bruns TD, 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology* **2,** 113–118.

Glass NL, Donaldson GC, 1995. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Applied Environmental Microbiology* **61,** 1323–1330.

469 Gonthier P, Ferracini C, 2104. How invasive pathogens and pests may threaten a multipurpose tree
470 species: the European chestnut as a case study. *International Forestry Review* **16**, 335.
471
472 Gonthier P, Sillo F, Lagostina E, Roccotelli A, Cacciola OS, Stenlid J, Garbelotto M, 2015.
473 Selection processes in simple sequence repeats suggest a correlation with their genomic location:
474 insights from a fungal model system. *BMC Genomics* **16**, 1107.
475
476 Gori A, Cerboneschi M, Tegli S, 2012. High-resolution melting analysis as a powerful tool to
477 discriminate and genotype *Pseudomonas savastanoi* pathovars and strains. *PLoS One* **7**, e30199.
478
479 Hansen MM, Mensberg KLD, 2009. Admixture analysis of stocked brown trout populations using
480 mapped microsatellite DNA markers: indigenous trout persist in introgressed populations. *Biology*
481 *Letters* **5**, 656–659.
482
483 Karaoglu H, Lee CM, Meyer W, 2005. Survey of simple sequence repeats in completed fungal
484 genomes. *Molecular Biology Evolution* **22**, 639–49.
485
486 Leišová-Svobodová L, Minaříková V, Matušinsky P, Hudcovicová M, Ondreičková K, Gubiš J,
487 2014. Genetic structure of *Pyrenophora teres* net and spot populations as revealed by microsatellite
488 analysis. *Fungal Biology* **118**, 180–92.
489
490 Linde CC, Zala M, Ceccarelli S, McDonald BA, 2003. Further evidence for sexual reproduction in
491 *Rhynchosporium secalis* based on distribution and frequency of mating-type alleles. *Fungal*
492 *Genetics and Biology* **40**, 115–125.
493

494 Lione G, Giordano L, Sillo F, Gonthier P, 2015. Testing and modelling the effects of climate on the
495 incidence of the emergent nut rot agent of chestnut *Gnomoniopsis castanea*. *Plant Pathology* **64**,
496 852–863.
497
498 Lione G, Gonthier P, 2015. A permutation-randomization approach to test the spatial distribution of
499 plant diseases. *Phytopathology* **106**, 19–28.
500
501 Luchi N, Pratesi N, Simi L, Pazzagli M, Capretti P, Scala A, Slippers B, Pinzani P, 2011. High-
502 resolution melting analysis: a new molecular approach for the early detection of *Diplodia pinea* in
503 Austrian pine. *Fungal Biology* **115**, 715–23.
504
505 Mackay JF, Wright CD, Bonfiglioli RG, 2008. A new approach to varietal identification in plants
506 by microsatellite high resolution melting analysis: application to the verification of grapevine and
507 olive cultivars. *Plant Methods* **4**, 8.
508
509 Maresi G, Oliveira Longa CM, Turchetti T, 2013. Brown rot on nuts of *Castanea sativa* Mill: an
510 emerging disease and its causal agent. *iForest* **6,** 294–301.
511
512 McDonald BA, Linde C, 2002. Pathogen population genetics, evolutionary potential, and durable
513 resistance. *Annual Review of Phytopathology* **40**, 349–379.
514
515 McEwen JG, Taylor JW, Carter D, Xu J, Felipe MS, Vilgalys R, Mitchell TG, Kasuga T, White T,
516 Bui T, Soares CM, 2000. Molecular typing of pathogenic fungi. *Medical Mycology* **38**, 189–97.
517

518 Mogg R, Batley J, Hanley S, Edwards D, O'Sullivan H, Edwards KJ, 2002. Characterisation of the
519 flanking regions of *Zea mays* microsatellites reveals a large number of useful sequence
520 polymorphisms. *Theoretical and Applied Genetics* **105**, 532–543.
521
522 Pasche S, Calmin G, Auderset G, Crovadore J, Pelleteret P, Mauch-Mani B, Barja F, Paul B,
523 Jermini M, Lefort F, 2016. *Gnomoniopsis smithogilvyi* causes chestnut canker symptoms in
524 *Castanea sativa* shoots in Switzerland. *Fungal Genetics and Biology* **87**, 9–21.
525
526 Peakall R, Smouse PE, 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software
527 for teaching and research-an update. *Bioinformatics* **28**, 2537–2539.
528
529 Pritchard JK, Stephens M, Donnelly P, 2000. Inference of population structure using multilocus
530 genotype data. *Genetics* **155**, 945–959.
531
532 Ricchi M, Barbieri G, Cammi G, Garbarino CA, Arrigoni N, 2011. High-resolution melting for
533 analysis of short sequence repeats in *Mycobacterium avium* subsp. *paratuberculosis*. *FEMS*
534 *Microbiology Letters* **323**, 151–4.
535
536 Shuttleworth LA, Liew ECY, Guest DI, 2012. *Gnomoniopsis smithogilvyi* sp. nov. Fungal Planet
537 Description Sheets 108. *Persoonia* **28**, 138–182.
538
539 Shuttleworth LA, Walker DM, Guest DI, 2015. The chestnut pathogen *Gnomoniopsis smithogilvyi*
540 (Gnomoniaceae, Diaporthales) and its synonyms. *Mycotaxon* **130**, 929–940.
541 Stergiopoulos I, Groenewald M, Staats M, Lindhout P, Crous PW, De Wit PJ, 2007. Mating-type
542 genes and the genetic structure of a world-wide collection of the tomato pathogen *Cladosporium*
543 *fulvum*. *Fungal genetics and biology* **44**, 415–429.

544

545 Stukenbrock E, Banke S, McDonald BA, 2006. Global migration patterns in the fungal wheat
546 pathogen *Phaeosphaeria nodorum*. *Molecular Ecology* **15**, 2895–904.

547

548 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular Evolutionary
549 Genetics Analysis Version 6.0. *Molecular Biology Evolution* **30**, 2725–2729.

550

551 Tamietti G, 2016. On the fungal species *Gnomoniopsis castaneae* ("*castanea*") and its synonym *G.*
552 *smithogilvyi*. *Journal of Plant Pathology*, doi: 10.4454/JPP.V98I2.001.

553

554 Tsui CK, Roe AD, El-Kassaby YA, Rice AV, Alamouti SM, Sperling FA, Cooke JE, Bohlmann J,
555 Hamelin RC, 2012. Population structure and migration pattern of a conifer pathogen, *Grosmannia*
556 *clavigera*, as influenced by its symbiont, the mountain pine beetle. *Molecular Ecology* **21**, 71-86.

557

558 Visentin I, Gentile S, Valentino D, Gonthier P, Tamietti G, Cardinale F, 2012. *Gnomoniopsis*
559 *castanea* sp. *nov*. (Gnomoniaceae, Diaporthales) as the causal agent of nut rot in sweet chestnut.
560 *Journal of Plant Pathology* **94**, 411–9.

561

562 White TJ, Bruns T, Lee S, Taylor J, 1990. *Amplification and direct sequencing of fungal ribosomal*
563 *RNA genes for phylogenetics. PCR Protocols. A Guide to Methods and Applications.* San Diego,
564 USA, Academic Press, Innis MA, Gelfand DH, Sninsky JJ eds, 315–322.

565

566 Wu SB, Wirthensohn MG, Hunt P, Gibson JP, Sedgley M, 2008. High resolution melting analysis
567 of almond SNPs derived from ESTs. *Theoretical and Applied Genetics* **118**, 1–14.

568

Figure legends

Figure 1. Map of the sites sampled in the current work.

Figure 2. Melting curve difference plots after normalization and overlay for the five *G. castaneae* polymorphic SSR loci. (a) Gc_AAA60; (b) Gc_AAG8; (c) Gc_AAG57; (d) Gc_AG26; (e) Gc_CCG42. Differences in terms of relative fluorescence were obtained using the PRECISION MELT ANALYSIS™ software.

Figure 3. Distribution of haplotypes in the sites. Histograms are divided into sections, which represent the frequency (%) of each haplotype in each site. Shared haplotypes among sites share similar pattern. Unique haplotypes are blank.

Figure 4. Principal coordinate analysis (PCoA) (a) and Discriminant Analysis (DA) (b) results. In (b), each site is represented by an ellipsoid. The percentage of variability represented by the first two axes is 94.4%.

Figure 5. Analysis of genetic structure. (a) Result of STRUCTURE analyses, where ΔK values (ΔK=mean ((|L″(K)|)/SD(L(K)); Evanno *et al.* 2005) are plotted for values of K from 2 to 10. In x-axis the number of inferred K were plotted, while in y-axis the values of ΔK were shown. (b) Bar plots for K=2 showing the assignment values for *G. castaneae* isolates from ten sites as inferred by STRUCTURE analysis.

Supporting information for online publication

Table S1. List of isolates used in the current study with allele identity. For each locus, allele number represents the corresponding cluster as detected by HRM analysis. Null alleles are indicated with "0". All isolates are available at DISAFA fungal collection.

Table S2. List of the primers used in the current work. A brief description of the use of each primer set is included.

Figure S1: Dendrogram based on pairwise population matrix of Nei genetic identity.

Figure S2. Graphs representing the expected distribution of I_A based on 1000 permutations of each considered site. The frequencies of I_A are represented as histograms. The observed I_A for each site is represented by a dotted line (grey). For each site, there is no evidence of linkage disequilibrium among loci ($p > 0.05$), suggesting the presence of sexual recombination.

Figure S3. Analysis of genetic structure as inferred by STRUCTURE. (a) Bar plots for K=3 showing the assignment values for *G. castaneae* isolates from ten sites. (b) Bar plots for K=4.

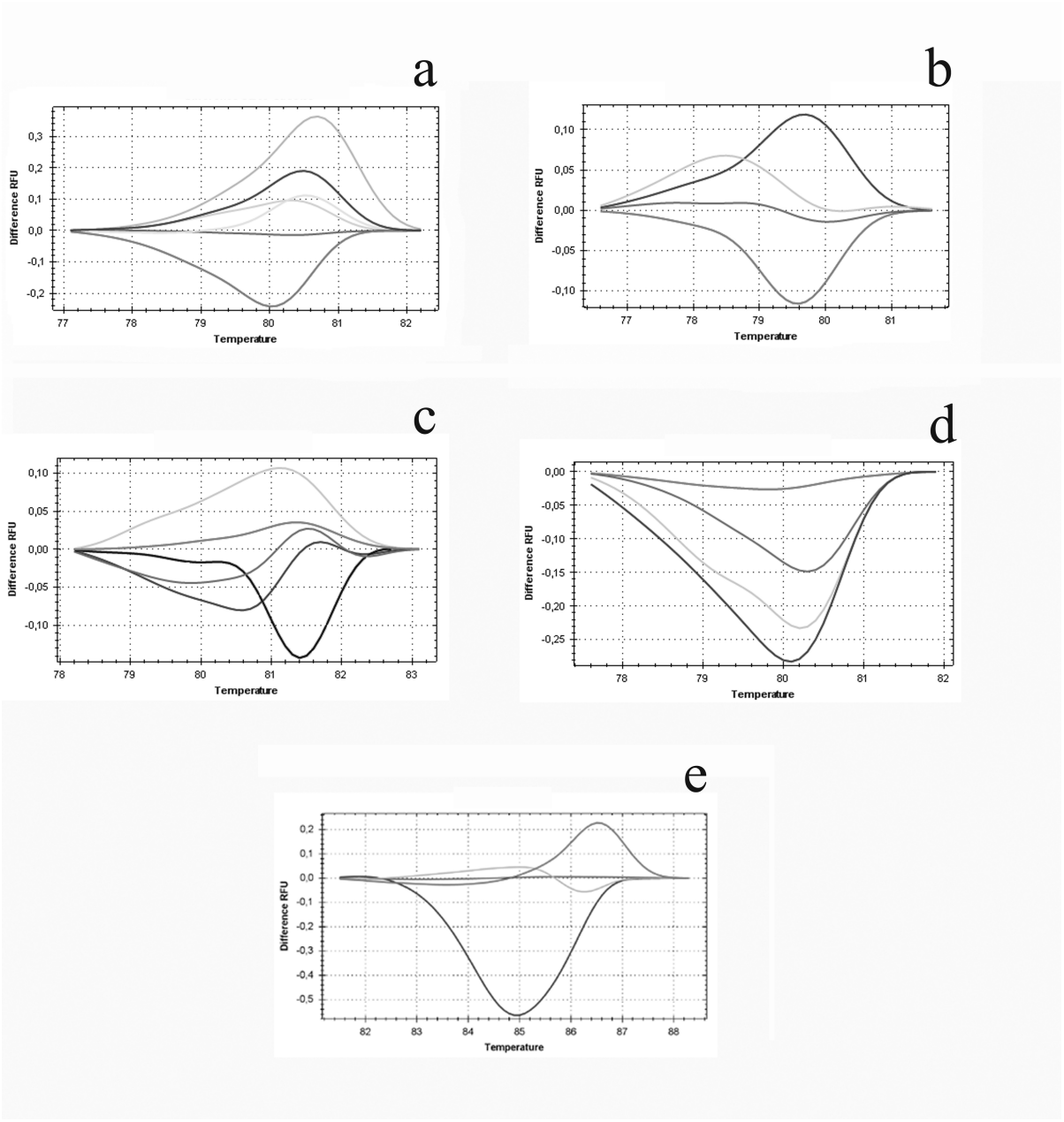
603 **Fig. 1**

604

605

606

607 **Fig. 2**



608

609

610

611

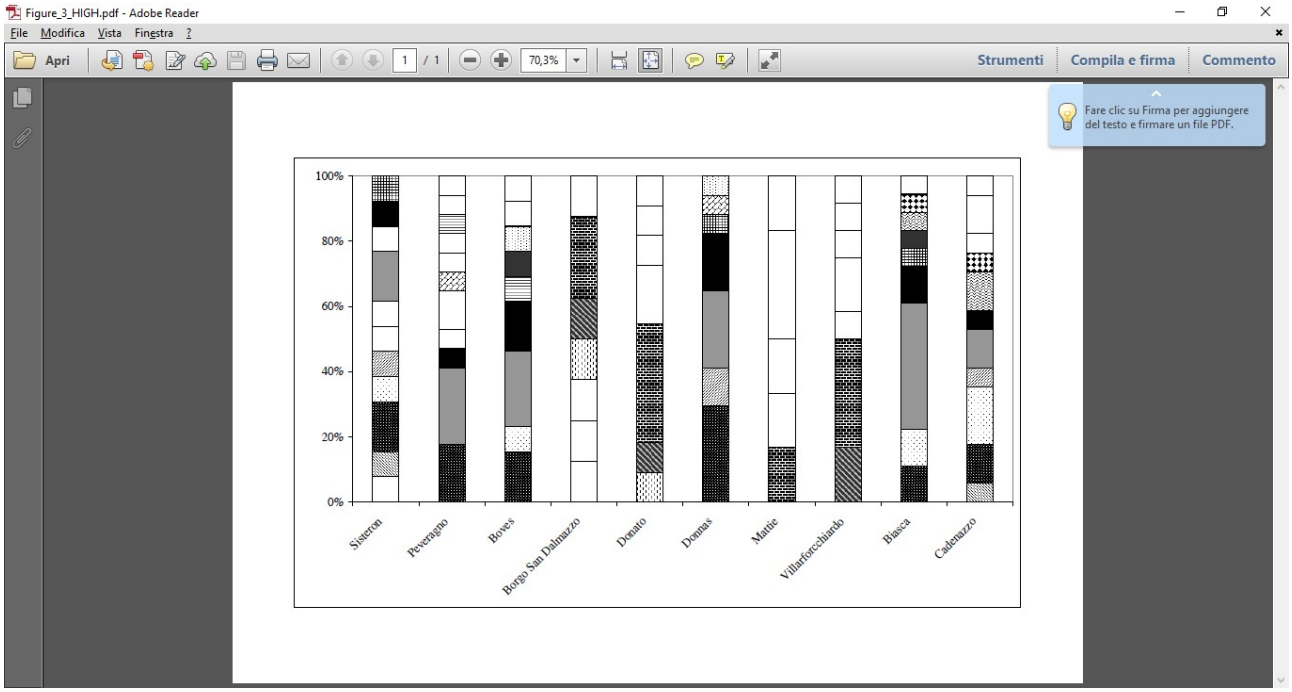
612

613

614

615

616 **Fig. 3**

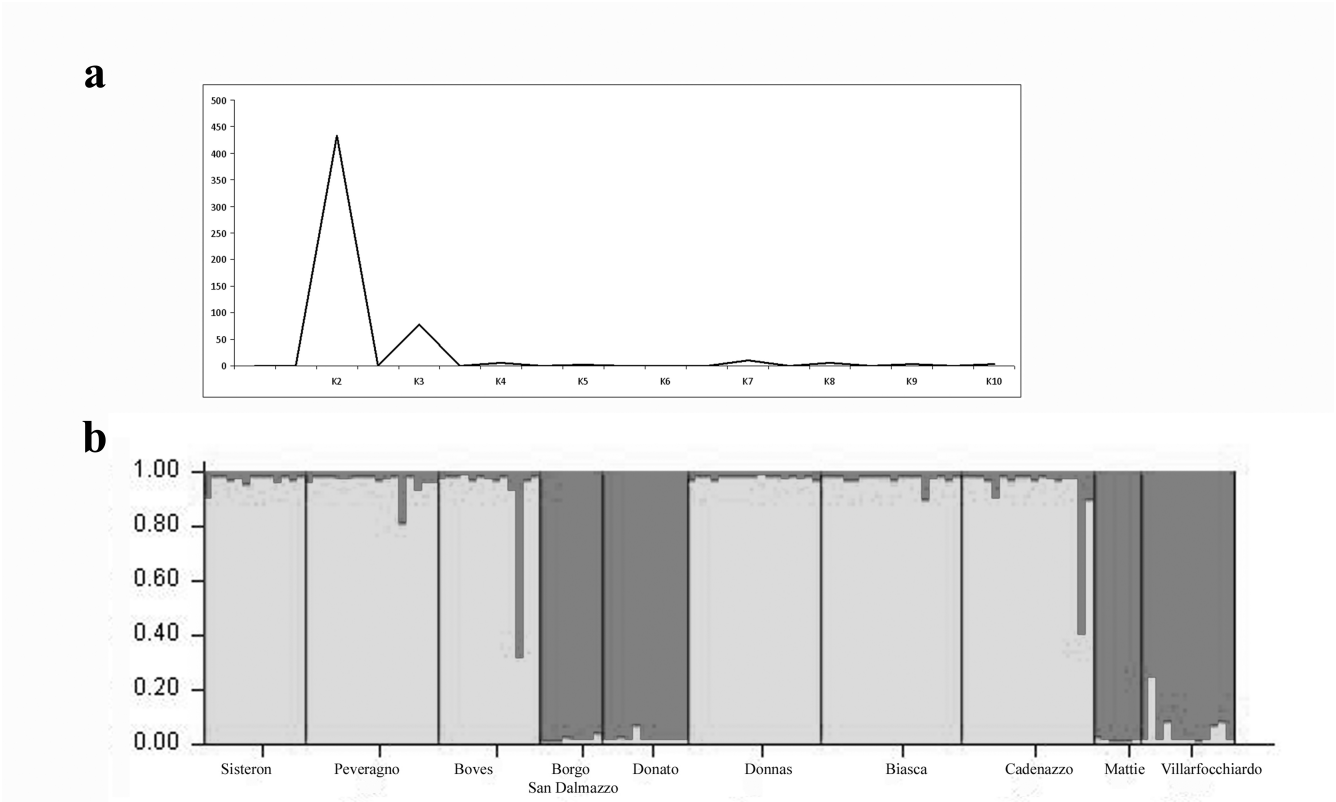


617

618

619

620 **Fig. 4**



621

622

623

624

625 **Table 1**. List of sampling sites, *G. castanea* incidence *per* site and number of obtained isolates.
626

627

Country	Site	UTM WGS84 coordinates (m)	Altitude (m a.s.l.)	*G. castanea* incidence (%)	Number of *G. castanea* isolates
Italy	Borgo San Dalmazzo	E 378203.3 N 4909837.6	655	85.0	8
	Boves	E 385186.1 N 4907245.0	783	69.2	13
	Donato	E 414851.2 N 5043995.9	1011	55.0	11
	Donnas	E 402474.5 N 5048801.2	848	59.5	17
	Mattie	E 351141.2 N 4995572.5	1170	20.0	6
	Peveragno	E 389871.2 N 4907514.9	680	80.0	17
	Villar Focchiardo	E 359474.5 N 4995073.5	1150	45.0	12
France	Sisteron	E 256967.6 N 4897070	485	-	13
Switzerland	Biasca	E 500159 N 5132793	301	-	18
	Cadenazzo	E 496405 N 5110698	247	-	17

628

629

630

631 **Table 2.** List of the primers used in the current work. A brief description of the use of each primer set was included.

Sequence	Name	Description
AGCGGGCATGCCTGTTCGAG	Gc1f	Lione *et al.* 2015
ACGGCAAGAGCAACCGCCAG	Gc1r	
CTTGGTCATTTAGAGGAAGTAA	ITS1-F	Gardes & Bruns 1993
TCCTCCGCTTATTGATATGC	ITS4	White *et al.* 1990
CATCGAGAAGTTCGAGAAGG	EF1-728F	Carbone & Kohn 1999
TACTTGAAGGAACCCTTACC	EF1-986R	
GGTAACCAAATCGGTGCTGCTTTC	Bt2a	Glass & Donaldson 1995
ACCCTCAGTGTAGTGACCCTTGGC	Bt2b	
CGTGTCGACGATGAGTCCTGAG	MseI/SalI_Adaptor1	Adapters after digestion with MseI (Acquadro *et al.* 2005)
TACGTGTCGACGATGAGTCCTGAG	MseI/SalI_Adaptor2	
CGTGTCGACGATGAGTCCTGAG	MseI/SalI_F	Primer used to generate the SSR enriched libraries (Acquadro *et al.* 2005)
GCGGTCGACACACACACACACACACAC	AC_SalI	Primers used to generate the SSR enriched libraries

GCGGTCGACAGAGAGAGAGAGAGAGAG	AG_SalI	(This work)
GCGGTCGACAAGAAGAAGAAGAAGAAGAAGAAGAAG	AAG_SalI	
GCGGTCGAAAAAAAAAA	A_SalI	
GCGGTCGCCGCCGCCGCCGCCGCCGCCGCCGCCG	CCG_SalI	
CGCCAGGGTTTTCCCAGTCACGAC	M13-FW	Universal primers used to screen the libraries (Promega)
CAGGAAACAGCTATGAC	M13-RV	
CATTGCCGGATTCTTCACTT	Gc_AAG8_F	Primers for polymorphic SSR loci (this work)
CGACGAACTGAGACGGATGT	Gc_AAG8_R	
GCGAGAACCACGTCCTTTCT	Gc_AAG57_F	
AGTACGAGCGGGCAAATTGA	Gc_AAG57_R	
AGGACATTCATCCGTCCAAC	Gc_AG26_F	
AGAGCAAGAGAGAGTTTCCGA	Gc_AG26_R	
CCCAGGGAATATCATCAAGC	Gc_AAA60_F	
GTGGTTTGAGTTGAGCGGC	Gc_AAA60_R	
ACGAGGCGTTCAATGATACC	Gc_CCG42_F	
AACGAAAAGCGCTAGAACCA	Gc_CCG42_R	

CTCCCAAGATCGGCAAAAT	Gc_AAG54_F	Primers for monomorphic SSR loci (this work)
GAATTTGTGCGATGGTGGGC	Gc_ AAG54_R	
CCCTGATTCAGATCGAAAAGA	Gc_ AC37_F	
CACGACCACCTACCATCCAC	Gc_ AC37_R	
CGCCAGAAAAGAGAGCAATG	Gc_AC38_F	
GCTAGGCGGGTAATACAGAC	Gc_ AC38_R	
GTATCTCGATGCGACGACAA	Gc_AC39_F	
CGACCGCGTGAAGTCTATAGT	Gc_ AC39_R	
CGTTCATTCACCTTCTGACTAGG	Gc_AC42_F	
GAGTGAGTGTGGGGATGACG	Gc_AC42_R	
CCTGGGCTTCGACTTCAATA	Gc_AG8_F	
CCGTTCTCACACCGTCTTGA	Gc_AG8_R	
AATTAGCGCGGAGACACACT	Gc_AG15_F	
GGGCTACGGTGGCAGATATC	Gc_AG15_R	
GCTTCTTGCAGGTTTGCAGT	Gc_AG19_F	
AAGCGGGTCCCACTCAATG	Gc_AG19_R	

AAGTTTGATGGATCGCCTTG	Gc_AG29_F	
CATGCTTCCAACGCCCAAAA	Gc_AG29_R	
CCTTCAGCTCGTCATCATCA	Gc_AAG52_F	
ACAGACCGATGAAGAAGACGAC	Gc_AAG52_R	
TGACACAGAAAAGGGGCAGT	Gc_AAG55_F	Primers for SSR loci (not amplifiable) (this work)
ACGATTTGTCCTGAGTAATACGG	Gc_AAG55_R	
CGATGGGCATAGTGAGGTTC	Gc_AC44_F	
GGGGAGGAGTGTCGAGTGAA	Gc_AC44_R	
CTTTGCCCTTCCGTGTTAGA	Gc_AC45_F	
ATCAAGCAAATAATTCTAATATAACCC	Gc_AC45_R	
AATTGGTCAAGTCGGTCCAG	Gc_CCG57_F	
AACATGCGTCACCTACCAGA	Gc_CCG57_R	
CCYCGYCCYCCYAAYGCNTAYAT	NcHMG1	Degenerate primers used to isolate HMG domain of MAT-1-2-1 of *Ascomycetes* (Arie et al. 1997)
CGNGGRTTRTARCGRTARTNRGG	NcHMG2	
CATCCTGAGGCGTCCAACAA	GcHMGF	Primers for HMG domain of MAT-1-2-1 of

CCTTGATCTCGGCAGCCTTA	GcHMGR	*G. smithogilvy* (This work)

632

633

Table 3. Summary of allele variation, linkage disequilibrium and haplotype diversity statistics *per* site.

	France	Italy						Switzerland	
	Sisteron	**Peveragno**	**Boves**	**Borgo San Dalmazzo**	**Donato**	**Donnas**	**Mattie**	**Villarfocchiardo**	**Biasca**	**Cadenazzo**
Number of alleles (average)	2.200	2.800	2.200	2.400	2.200	2.200	2.000	2.000	2.000	2.400
Number of effective alleles	1.689	1.733	1.635	1.791	1.472	1.356	1.751	1.652	1.392	1.583
Frequency of private alleles	0.000	0.000	0.000	0.000	0.200	0.000	0.000	0.000	0.000	0.200
Number of loci in LD[a] (*p*-value < 0.05)	0/210	8/210	0/210	4/210	1/210	0/210	2/210	0/210	0/210	1/210
Unique haplotypes (n°)	4	5	1	3	4	0	3	3	1	3
Shared haplotypes (n°)	7	5	6	2	2	6	1	2	6	6
Clonal fraction (%)	15.38	70.59	46.15	37.50	45.45	64.71	33.33	16.67	61.11	47.06

[a] Linkage disequilibrium

640 **Table 4.** Summary of allele diversity statistics for all loci analyzed *per* site.

	Hs_a									Ht_b	Gst_c
	France	**Italy**						**Switzerland**			
Locus	**Sisteron**	**Peveragno**	**Boves**	**Borgo San Dalmazzo**	**Donato**	**Donnas**	**Mattie**	**Villarfocchiardo**	**Biasca**	**Cadenazzo**		
Gc_AAA60	0.2604	0.2076	0.1420	0.4444	0.4600	0.2076	0.0000	0.3750	0.0000	0.3806	0.6128	0.5957
Gc_AAG8	0.4970	0.4567	0.4734	0.5938	0.5124	0.2907	0.5000	0.6528	0.3457	0.3599	0.6461	0.2753
Gc_AAG57	0.1420	0.3945	0.4615	0.5312	0.0000	0.1107	0.6111	0.0000	0.1049	0.1107	0.3224	0.2349
Gc_AG26	0.5444	0.4922	0.0000	0.2188	0.3140	0.2145	0.4444	0.3750	0.2076	0.1244	0.3256	0.0985

641

Gc_CCG42	0.4260	0.4844	0.5562	0.2188	0.0000	0.4152	0.2778	0.1528	0.5123	0.5952	0.6835	0.4677
Average	0.3740	0.4071	0.3266	0.4014	0.2573	0.2478	0.3667	0.3111	0.2341	0.3142	0.5181	0.3746
SD	0.1685	0.1179	0.2415	0.1750	0.2459	0.1133	0.2377	0.2484	0.2013	0.2018	0.0320	

642
643
644 [a] Nei's gene diversity within each site.
645 [b] Nei's gene diversity including all sites.
646 [c] Fixation index for population differentiation.
647
648
649

650