COLLETOTRICHUM GLOEOSPORIOIDES S.L. IN NORTH AMERICA: SEX, HOST, AND HABITAT-MEDIATED DIVERSITY IN A PLANT-ASSOCIATED ASCOMYCETE

by

Vinson P. Doyle

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the
requirements for the degree of Doctor of Philosophy, The City University of New York
2012

This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Dr. Amy Litt

Date

Chair of the Examining Committee

Dr. Laurel Eckhardt

Date

Executive Officer

Dr. Dwight Kincaid
Dr. Joseph Rachlin
Dr. Damon Little
Dr. Peter Oudemans
Dr. Peter E. Smouse
Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

Colletotrichum gloeosporioides s.l. in North America: sex, host, and habitat-mediated diversity in a plant-associated ascomycete

by

Vinson P. Doyle
Advisor: Dr. Amy Litt

Determining the factors that drive the evolution of pathogenic fungi is central to revealing the mechanisms of virulence and host preference, as well as developing effective disease control measures. Prerequisite to these pursuits is the accurate delimitation of species boundaries. *Colletotrichum gloeosporioides s.l.* is a species complex of plant pathogens and endophytic fungi for which reliable species recognition has only recently become possible through a multi-locus phylogenetic approach. Through intensive regional sampling that encompasses multiple hosts within and beyond agricultural zones associated with cranberry (*Vaccinium macrocarpon* Aiton), we have integrated North American strains of *Colletotrichum gloeosporioides s.l.* from these habitats into a broader phylogenetic framework and characterized some of the factors that influence species diversity. We have developed polymorphic microsatellite markers for *C. fructivorum*, a species determined to be responsible for cranberry fruit-rot in agricultural areas throughout North America, in order to understand the biotic and abiotic factors that shape populations within the species complex. These markers amplify across several species within the *C. gloeosporioides* species complex and some are variable within two species, *C. rhexiae* and *C. kahawae*, that are closely related to *C. fructivorum*. Broad geographical and fine-scale hierarchical sampling of *C. fructivorum* and *C. rhexiae* coupled with multilocus genotyping has allowed us to gain insight into the forces that shape populations of these species. Human-mediated dispersal is an important factor dissipating the population structure of *C. fructivorum* throughout its range in commercial cranberry bogs. In contrast, limited evidence suggests *C. rhexiae* is geographically structured within a more restricted range, implying distinct patterns of diversity between *Colletotrichum* species associated with wild versus agricultural hosts. We also investigate the reproductive mode of *C. fructivorum* using estimates of haploid disequilibrium and genotypic diversity, inferring a mixed (sexual and asexual) mode of reproduction in field populations. We discuss the importance of sexual and asexual reproduction on population dynamics and speciation within the *C. gloeosporioides* species complex.

ACKNOWLEDGEMENTS

The many hours that have gone into this body of work are dedicated, first and foremost, to three important people: my mother, Catherine Doyle; my father, Stan P. Doyle; and my best friend, Christina M. Mozzicato. My parents have consistently provided the intellectual, emotional, financial, and spiritual support that I needed to pursue lofty goals and the freedom to pursue my interests. Their hard work and sacrifice will always be inspirational. My best friend and partner, Christina Mozzicato, has made great sacrifices of her own, too numerable to mention, to help me reach this point. More than any other single person, Christina has helped with nearly every aspect of my life and this project over the last several years. There is not a doubt in my mind that this research would not have come to fruition were it not for the hard work of Christina. Thank you, Christina, for your smiles, laughs, encouragement, hard work, perseverance and friendship.

I also want to extend sincere gratitude to my thesis committee. First, thanks to my advisor, Dr. Amy Litt, for allowing me the room to pursue my interests and put together a project that would sustain me. I also want to thank Dr. Dwight Kincaid, Dr. Joseph Rachlin, Dr. Damon Little, Dr. Peter Oudemans, and Dr. Peter Smouse for dedicating their time to making sure I had the tools I needed to succeed.

There are many, many others that have helped me along the way. I will surely forget to mention some here, but even if they escape acknowledgement here they do not escape appreciation. I am grateful for the friendship and comradery of several former and current doctoral students at the New York Botanical Garden and/or the CUNY Graduate Center. I want to extend special thanks to Josh Simpson, who has been a great friend, a great collaborator, and a great inspiration. Thanks also to Natalia Pabón-Mora, William Perez, Ricardo Kriebel, and

Marcelo Reginato for bearing my presence as part of the “upstairs graduate student” contingent and engaging in lively and humorous conversation. Thanks to Michael Sundue for his continued friendship and collaboration. Thanks to Todd Osmundson for his undying optimism in the ability of others. And thanks to many others: Donald McClelland, Rachel Meyer, Nelson Salinas, Julian Aguirre, Fernando Matos, Aleja Vasco, Maria Alice-Neves, Marc Jeanson, Yayi Huang, Ryan Huish, Jack Henning, James Lyles, Kurt Reynertson, Cesar Marin, Ratnakar Vallabhaneni, Oren Tzfadia, James Lendemer, Terry Demos, and Aman Gill.

Several people at the Systematic Mycology and Microbiology Laboratory in Beltsville, Maryland were indispensable for my success, especially Dr. Stephen Rehner and Dr. Gary Samuels. I have found great mentors and friends in each of them. Many people from the Marucci Center for Cranberry and Blueberry Research in New Jersey have also been of great help along the way. In particular, this project would not have been possible if it were not for the help of Chris Constantelos, Jen Vaicunas, Lindsay Wells, and Donna Larsen. I am also grateful for the help of Dr. Jim Polashock and Dan Schiffauer.

Several individuals at Lehman College have been particularly supportive. I want to thank Dr. Dominick Basile and Mike Baxter for their friendship and assistance. Thanks to Dr. Miguel Cervantes-Cervantes for his help early in my graduate education. Thanks to Dolores Vitanza for her help at many difficult points.

I have made many friends at the New York Botanical Garden that I owe many thanks. In particular, I want to extend my appreciation to Dr. Barbara Thiers and Dr. Larry Kelly for making sure that I had a fellowship that would provide financial support throughout my graduate education. I also want to thank several curators that have provided help and humor along the way: Dr. Dennis Stevenson, Dr. Roy Halling, Dr. Larry Kelly, Dr. Lisa Campbell, Dr. Fabian

Michelangeli, Dr. Ken Karol, Dr. Barbara Ambrose, Dr. Jim Miller, Dr. Greg Plunkett, Dr. Doug Daly, Dr. Dick Harris, and Dr. Michael Nee. Thanks to Matt Sewell and Anna Nowogrodzki for their help in the lab. Many people in the herbarium have been great friends and coworkers: Benito Santos, Stephen Gottschalk, Kim Watson, Shannon Asencio, Nicole Tarnowsky, Nicole Tiernan, Nestor Perez-Moliere, Stella Sylva, and Mike Bevans. I also want to extend special thanks to Carmen Maldonado, Lisa Fruscella, and Myrna Alvarez for all of their help along the way.

Lastly, I want to thank all of my friends and family that have been encouraging and supportive. Thanks to the all of my friends in the United States and abroad. Thanks to the extended Doyle Family. Thanks to the Robertson Family and the Morrison Family. I also want to extend special thanks to Guy and Florette Mozzicato for bringing me into their family.

CONTENTS

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

APPENDIX 4B

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

APPENDIX 2

APPENDIX 4A

APPENDIX 4C

CHAPTER 1

INTRODUCTION AND BACKGROUND

INTRODUCTION

Several species of *Colletotrichum* Corda (teleomorph: *Glomerella* von Schrenk & Spaulding) are well known for their ability to cause necrotic lesions on herbaceous and woody plant species. *Colletotrichum gloeosporioides s.l.* (Penz.) Penz. & Sacc. is an aggregate of species, distributed worldwide and infecting a multitude of host plant species. The complex life history, multitude of host associations and morphological variability of *C. gloeosporioides* has inhibited understanding of its lifecycle and dispersal mechanisms, as well as of its phylogenetic diversity. Until the middle of the twentieth century, approximately 900 species were described within the genus *Colletotrichum,* defined primarily on the basis of host association (Sutton, 1992). In 1957, von Arx applied broad taxonomic concepts to reduce the number of recognized species to just eleven. *Colletotrichum gloeosporioides* was one of the taxa retained in von Arx's revision, for which he cited approximately 600 synonyms (Sutton, 1992). *Colletotrichum gloeosporioides s.l.* is a species complex with broad morphological variability that has caused confusion in distinguishing species within the complex (Sutton, 1992).

Since von Arx's revision, several efforts have been mounted to understand the phylogenetic diversity and host associations within *Colletotrichum*, and new species have been published (Cannon, Bridge, 2000; Crouch *et al.*, 2006; Rojas *et al.*, 2010). A recent description of accepted taxa was published in 1992, listing 39 species (Sutton, 1992). Sutton (1992: page 15) also recognized seven subspecific taxa within *C. gloeosporioides* and made the following statement about the future of the systematics of this species complex: "No progress in the systematics and identification of isolates belonging to this complex is likely to be made based on morphology alone." Species delineations among several of the taxa within *Colletotrichum* are unclear and often unreliable. An understanding of the molecular variation within the genus will be important for delimiting species boundaries as has recently been demonstrated (Rojas *et al.*, 2010; Silva *et al.*, 2012b), but an understanding of gene flow within and between populations of single species will also facilitate the development of a clear and reliable taxonomy for the genus, as well as providing important insight into developing effective management strategies for diseases caused by *Colletotrichum gloeosporioides* (Correll *et al.*, 2000).

Assessments of variation in host specificity, tissue/organ specificity, and pathogenicity among *C. gloeosporioides* populations have resulted in varying conclusions, relative to their utility in defining population structure and revealing cryptic/incipient species (Freeman *et al.*, 1996; González *et al.*, 2006; Guozhong *et al.*, 2004; Johnston, Jones, 1997; Kaufmann, Weidemann, 1996; MacKenzie, 2005). Due to their pathogenicity on several economically important crop plants, it is important to develop a better understanding of the factors that lead to population structure in *Colletotrichum*. It has been suggested, due to the lack of information regarding systematic relationships within the *C. gloeosporioides* complex, that regional sampling will be important in providing a greater understanding of the biology of this fungus (Cannon, Bridge, 2000). Intensive local sampling provides the means to determine if alternate host species serve as a reservoir for the subsequent infection of local crop species and allows for determining the effect of landscape features (e.g. irrigation networks and host species) on gene flow. Because *C. gloeosporioides* is a widespread and variable pathogen of cranberry (*Vaccinium macrocarpon*), an investigation of the factors contributing to its population structure is timely and important. In addition, management practices on cranberry farms are particularly conducive for testing hypotheses of dispersal patterns of this fungus. The object of this study was to characterize the genetic diversity of the *C. gloeosporioides* species complex that infect cranberry, primarily in eastern North America. Genotyping of fungal isolates were utilized to infer local and regional population structure within *C. gloeosporioides*, addressing questions of diversity, host range, dispersal and reproductive mode. An understanding of host range, at both local and regional scales, provides valuable information for future systematic studies, particularly in devising sampling strategies for ensuring inclusion of distinct lineages. Determining whether reproduction of *C. gloeosporioides* on cranberry is clonal or sexual may have implications for the development of more effective resistance breeding strategies and pathogen management practices, as well as shedding some light on the adaptive potential of this pathogen under field conditions. This study improves our understanding of the systematics, population dynamics, and natural history of *C. gloeosporioides,* using cranberry as the focal host population.

RESEARCH OBJECTIVES

Phylogenetic Diversity

I assessed the phylogenetic diversity of *C. gloeosporioides* from cultivated and wild *Vaccinium macrocarpon* populations and from sympatric host plant species.

I assessed the phylogenetic diversity of *C. gloeosporioides* on cranberry, using multiple molecular markers from nuclear loci (*nrITS*, *beta-tubulin*, *apn2,* and *apn2/matIGS*) demonstrated to be useful for inferring phylogenetic relationships within the *C. gloeosporioides* complex (Rojas *et al.*, 2010). Additional isolates of *C. acutatum* from cranberry and *C. gloeosporioides* from geographically distant populations were included in phylogenetic analyses. Molecular phylogenetic analysis was used to reveal the phylogenetic diversity of *Colletotrichum* species from cranberry and sympatric alternate host species. Based on this analysis, I evaluated a single clade for infraspecific population structure (see below). Identifying a clade that represents the majority of the strains isolated from cranberry improves the chances of identifying microsatellite loci useful for elucidating population structure of *C. gloeosporioides* isolates from cranberry. It has been demonstrated that the success rate in amplifying microsatellite loci tends to decrease as genetic distance increases, therefore developing microsatellite primers for a given group may be of limited utility for distantly related groups (Roa *et al*., 2000).

Local Population Substructure

I investigated the population substructure of *C. gloeosporioides* on *Vaccinium macrocarpon* and on sympatric alternate host species.

The genetic structure of a population reflects its evolutionary history. Analysis of the patterns of genetic variation allows for inferences of how dispersal mechanisms, reproductive modes and differences in lifestyle may affect population subdivision and spatial patterns. Comparing allele frequencies at polymorphic loci provides information on the extent of genetic isolation among subpopulations (Taylor *et al.*, 1999a). I assessed the population substructure of *C. gloeosporioides* on cranberry and sympatric alternate host species by determining the distribution of genetic diversity within and among populations. Using microsatellite markers, I examined genetic divergence across multiple spatial scales (regions, watersheds, subwatersheds, cultivars, hosts, etc.).

I hypothesized that populations of *C. gloeosporioides* from *Vaccinium macrocarpon* and other sympatric host species are panmictic across multiple spatial scales. Sampling was conducted using a hierarchical sampling strategy in order to detect structure at multiple spatial scales. Panmixia has been demonstrated for broadly distributed fungi and is therefore a reasonable null hypothesis (Taylor *et al.*, 1999b).

Reproductive Mechanisms

I determined whether outcrossing is evident in populations of *C. gloeosporioides* on *Vaccinium macrocarpon* and sympatric alternate host species.

Differences in reproductive strategies in natural populations can contribute to population structure and have consequences for disease management. Sexual populations tend to be more genetically diverse than asexual populations. However, for homothallic species, such as *C. gloeosporioides*, sexual recombination does not necessarily result in the production of genetically distinct progeny (Taylor *et al.*, 1999b).

I hypothesized that evidence of outcrossing is not present within populations of *C. gloeosporioides*. Reproductive strategies may differ among isolated populations, therefore understanding the reproductive strategy within distinct populations can have important implications for the development of effective management tools. Clonal populations are indicated by: (*i*) an overrepresentation of identical genotypes, (*ii*) a lack of recombinant genotypes, and (*iii*) non-random associations between independent genetic markers (Milgroom, 1996). Non-random association between independent genetic markers can be tested by examination of linkage disequilibrium between microsatellite loci.

BACKGROUND ON *C. GLOEOSPORIOIDES*: TELEOMORPH - *GLOMERELLA*

Population Structure

Colletotrichum species are well known among plant pathologists as being among the causal agents of a plant disease referred to as anthracnose, characterized by "limited necrotic lesions, necrosis and hyperplasia" (Kirk *et al*., 2001). *Colletotrichum* species have been reported from a wide range of hosts, including many economically important plant species of the Rosaceae, Anacardiaceae, Lauraceae, Ericaceae, and Fabaceae (Farr *et al*., 2008). In addition, *Colletotrichum* species have also been reported as endophytes, defined as inhabitants of asymptomatic plant tissues for some part of their lifecycle (Rodriguez, Redman, 2000). Members of the *C. gloeosporioides* complex are perhaps the most common species in the genus and are distributed worldwide as plant pathogens and endophytes. In fact, *C. gloeosporioides* is reported in many regional and host plant surveys as being among the most common endophytic fungi recovered (Rodriguez, Redman, 2000). Due to its broad distribution and economic importance as a plant pathogen, several studies have been carried out to determine host specificity and population structure on the basis of lifestyle (endophytic vs. pathogenic) and degree of pathogenicity. The results of these studies are variable. Some studies, based on ap-PCR and RFLP data (Freeman *et al.*, 1996), report clonal populations sharing several host species (suggesting a lack of host specificity), whereas others, using isozyme profiles, have reported *formae speciales* (an informal rank on the basis of host association), based on the observance of distinct populations from distinct host species (Kaufmann, Weidemann, 1996). The divergence between these studies makes it difficult to make generalized statements about the host specificity of pathogenic *C. gloeosporioides* strains.

Attempts to determine the host specificity of *C. gloeosporioides* in endophytic populations have been variable. One study on endophytic *C. gloeosporioides* from Guyana reported high levels of sexual recombination among isolates from 12 different host species, but did not detect host associations among isolates using RAPD profiles and rDNA ITS sequence data (Guozhong *et al*., 2004). However, another study, investigating both host and lifestyle specificity between endophytic and pathogenic isolates found that endophytic isolates from various hosts formed a group distinct from pathogenic isolates, on the basis of rDNA sequence analysis (Photita *et al*., 2005). A similar conclusion was reached in a more recent multilocus phylogenetic study suggesting that host-associated pathogens are phylogenetically distinct from endophytes (Rojas *et al*., 2010). Another notable indication of specialization among *C. gloeosporioides* isolates includes the detection of variability in protease production among endophytic and pathogenic isolates. Protease production among endophytic isolates of *C. gloeosporioides* was more efficient at alkaline pH, while protease production among pathogenic isolates was more efficient at acidic pH (Maccheroni *et al.*, 2004). These studies suggest that lifestyle and host specialization may be present among populations of *C. gloeosporioides*, but incongruent results among studies suggest that further investigation is warranted.

While *C. gloeosporioides* is suspected of being a complex of diverse species, the phylogenetic relationships within this complex have only recently begun to be addressed (Phoulivong *et al.*, 2010; Prihastuti *et al.*, 2009; Rojas *et al.*, 2010; Silva *et al.*, 2012b). The aforementioned studies have variously supported and falsified hypotheses of host specificity and the evolution of distinct endophytic and pathogenic lineages, leaving these questions unresolved. As previously mentioned, due to the paucity of information regarding systematic relationships within the *C. gloeosporioides* complex, regional sampling will be important for understanding the biology of this fungus (Cannon, Bridge, 2000). One cannot be sure that what is *C. gloeosporioides* on a given host in a given geographic region is part of the same metapopulation as what is present in another geographic region or on another host species. In addition, all of the available tools to assess population structure in these fungi have not been utilized. While molecular tools have been adopted in many studies of this fungus, including RAPD, RFLP, isozyme and sequence markers, microsatellite loci (short, tandem repeats of 1-6 nucleotides) have not been investigated to date. Microsatellites are advantageous in being highly polymorphic, codominant, and reproducible, due to locus-specific primers (Jarne, Lagoda, 1996). These loci may allow for more fine-scale analyses of infraspecific variation, such as the assignment of individuals to populations and inference of kin relationships within a population (Dutech *et al*., 2007).

Microsatellite loci may also provide higher resolution in the study of population structure and allow inferences regarding reproductive strategies, dispersal mechanisms and migration models.

Host Specificity

As mentioned previously, several studies have led to the conclusion that there are *formae speciales* within the *C. gloeosporioides* complex, based on host-specific clustering of isolates (Freeman *et al.*, 1996; Guozhong *et al.*, 2004; Kaufmann, Weidemann, 1996; Photita *et al.*, 2005). It has been demonstrated with RAPD and RFLP markers, however, that there has been local pathogenic adaptation to strawberry, drawn from a broadly distributed population of *C. gloeosporioides*, found on several native host plant species (MacKenzie *et al*., 2007). This phenomenon of locally-derived pathogenesis suggests that there is the potential for one to be misled into believing that host-specific races are present if sampling is done from different hosts in disparate geographic localities. That study also demonstrates the importance of determining the factors that contribute to population structure at a regional level, while trying to conduct broad scale analyses of host association.

Dispersal mechanisms

Dispersal mechanisms may contribute to population structure. *Colletotrichum* species are thought to be water dispersed (Nicholson, Moraes, 1980; Tu, 1992), so populations may be subdivided on the basis of irrigation networks that connect fields in an agricultural environment. Understanding the effect of management practices, such as irrigation, on population structure in *C. gloeosporioides* may help plant pathologists to predict pathogen outbreaks and assess pathogen diversity within agricultural systems. This information will also provide some guidance in designing a sampling strategy to detect host-specific forms within the species complex, by reducing the effect of population structure due to dispersal patterns. Developing a more refined understanding of host specificity within the species complex may have particular importance for determining if alternate host species may serve as inoculum reservoirs in agricultural landscapes.

Reproductive strategy

Colletotrichum species are known to reproduce by both sexual and asexual means (Sutton, 1992). An understanding of reproductive mechanisms within field populations is important for understanding the natural history of the fungus under study, but it also is important for developing management practices for plant pathogens. Sexual recombination enables a fungus to adapt more rapidly to selective pressures, such as those caused by application of fungicides and by plant host resistance genes, but asexual reproduction enables rapid spread of pathogenic genotypes during epidemic phases (McDonald, Linde, 2002a). Therefore, an understanding of reproductive strategies at work in a pathogenic population allows inference of the adaptive potential of the species. It also has important implications for the development of pathogen resistance breeding strategies and management practices (McDonald, Linde, 2002a). Genotyping isolates with microsatellite markers can provide inference on reproductive strategies within natural populations; the central idea is to analyze gametic associations between unlinked loci.

Tissue specificity

Colletotrichum gloeosporioides affects several plant organs, including fruits, flowers, leaves and stems of some of the host species on which they are known to be pathogens (Farr *et al.*, 2008). One study (González *et al.*, 2006), using morphological criteria to investigate the population structure of *Colletotrichum* species infecting apple cultivars, determined that a single type was capable of causing pathogenic symptoms on leaves and fruit; the other isolates recovered in the study were pathogenic on fruit only (González *et al.*, 2006). This suggests the ability of specific genotypes to infect multiple tissue types among *C. gloeosporioides* isolates on apples. *Colletotrichum gloeosporioides* on cranberry is known to be pathogenic to fruit, but it has recently emerged as the causal agent of stem canker as well. However, it is not known whether there is gene flow between strains of *Colletotrichum gloeosporioides* infecting stems and fruits. Genotyping isolates using polymorphic microsatellite loci should provide an appropriate means to determine whether specific genotypes are responsible for causing stem canker in cranberry and if the type of organ being colonized may be a factor contributing to population structure. In addition, this may provide valuable information for the development of better host breeding strategies and pathogen management practices.

Population Structure and Pathogenicity Phenotypes

Microsatellite markers developed for the *C. gloeosporioides* species complex may be useful for systematic mycologists, allowing for the detection of reproductively isolated populations, from which samples could be drawn for future studies to infer phylogenetic relationships within the species complex. Genotyping strains of plant pathogens within agricultural systems also provides valuable information on genetic diversity, such that representative genotypes can be utilized for the development of pathogen-resistant host cultivars and testing of new fungicides (McDonald, Linde, 2002a; Taylor *et al.*, 1999a). Furthermore, it has been recognized that hypovirulent isolates of some fungal pathogens can serve as potential biological control agents against conspecific pathogenic isolates (Boland, 2004). An investigation of the potential for hypovirulent isolates of *C. gloeosporioides* to reduce the incidence and severity of post-harvest fruit rot in mangos from Puerto Rico revealed that hypovirulent isolates were able to reduce the severity of damage caused by virulent isolates by more than 50% (Colon-Garay *et al*., 2002). Therefore, identifying reproductively isolated populations would provide a broad sample from which isolates can be selected for efficiently testing pathogenicity phenotypes.

Cranberry Fruit Rot

Cranberry is one of the few economically important crop plants native to North America (Oudemans *et al.*, 1998). Cranberry production in the United States was valued at over 307 million dollars in 2007, showing a slight increase over values for 2005 and 2006. This constitutes more than 2.5% of the total value of all non-citrus fruits produced in the United States in 2007 (USDA-NASS, 2008). However, a yield-limiting factor in cranberry production is fruit rot, caused by a complex of fungal species that includes *C. gloeosporioides* (Oudemans *et al.*, 1998). In 1998, it was estimated that fruit rot fungi caused losses ranging from $185 to $10,950 per hectare. Therefore, a better understanding of the population dynamics of fruit rot fungi will help to develop better management strategies to reduce losses incurred by cranberry growers.

LIFECYCLE OF COLLETOTRICHUM / GLOMERELLA

General Characteristics

Many ascomycetes, including *Colletotrichum* species, can be found in more than one physically independent form (pleomorph). In many cases, the two separate forms have not been connected because of their independent nature (Kendrick, 2000). Thus, for many species, two valid names have emerged for a single organism. Basic questions regarding the natural history of ascomycetes can be confounded by this pleomorphic lifecycle. The sexual form is referred to as the teleomorph and the asexual form as the anamorph. In the case where the anamorph has been connected with the teleomorph, it is apparent that sexual reproduction does not occur in all populations. In nature, plant pathogenic ascomycetes are most often observed in their anamorphic state (Taylor *et al.*, 1999b). *Glomerella cingulata* (Stoneman) Spauld & H. Schrenk is the teleomorph of *Colletotrichum gloeosporioides* (Penz.) Penz. & Sacc., which produces a sexual fruiting structure called a perithecium and sexual spores called ascospores or meiospores. Sexual reproduction in *Glomerella cingulata* follows the general pattern outlined and illustrated in Figure 1.

Mating System

The mating systems in fungi are not comparable to those of other organisms, because sexual reproduction can be controlled by both mating type loci and vegetative incompatibility loci. Furthermore, sexual reproduction does not always result in assortment and shuffling of genetic variation among loci, because there are homothallic species that are self-fertile, as well as heterothallic species that are cross-fertile. In *Glomerella,* some strains can be both self-fertile and cross-fertile, while other strains are strictly cross-fertile (Vaillancourt *et al.*, 2000b). Despite a series of papers on the mating system of *Glomerella cingulata* in the 1940s and 1950s (Chilton *et al.*, 1945; Chilton, Wheeler, 1949a; Chilton, Wheeler, 1949b; Lucas, 1946; Lucas *et al.*, 1944; McGahen, Wheeler, 1951; Wheeler, 1950; Wheeler, McGahen, 1952; Wheeler *et al.*, 1948), the complexity of the mating system has still not been entirely unraveled. In homothallic strains, three alleles at two loci result in nine homothallic genotypes, whose phenotypes are summarized in Table 1 (Wheeler, McGahen, 1952). The two loci, designated as ‘A’ and ‘B’ by Wheeler and McGahen (1952) are responsible for the production and arrangement of fruiting structures (conidial or perithecial), respectively. Two other loci were described by Wheeler and McGahen (1952), in which the presence of mutant alleles F^1 and st^1 at the F and st loci, respectively, result in self-sterile isolates, regardless of the alleles present at the A and B loci. In addition, the st mutation inhibits the formation of perithecia in cross-fertilization matings, when carried by the perithecial isolate. It has been inferred, based on mating studies and the segregation of RFLP haplotypes, that mating in self-sterile heterothallic isolates (probably representing genotypes 4, 5 & 6 in Table 1) is controlled by a single mating-type locus with several alleles (Cisar, TeBeest, 1999). However, it has also been suggested that the data presented by Cisar and TeBeest (1999) point to a multi-locus, multi-allelic mating system (Vailllancourt *et al*., 2000). Thus, it is apparent that more research is needed to determine the mating system of this economically important fungus. This includes investigating the reproductive strategy of *C. gloeosporioides* / *G. cingulata* in natural populations.

Dispersal mechanisms

Colletotrichum (*Glomerella*) species produce two types of dispersal units, conidia produced from acervuli (asexual reproductive structures embedded in host tissue) and directly from hyphae and ascospores produced in perithecia. These spores are immersed in a mucilaginous matrix composed of polysaccharides and glycoproteins (Bailey *et al*., 1992). Spores in this matrix can be dispersed by water droplets or, if the matrix dries, by wind as spore masses (Nicholson, Moraes, 1980), though it is commonly suggested that water is the primary dispersal mechanism (Tu, 1992). One of the major points of support for this conjecture is the fact that the gelatinous matrix inhibits spore germination until the concentration of the matrix is diluted. It is also apparent that the matrix maintains the viability of spores in a dessicated state (Nicholson, Moraes, 1980). Therefore, while water is proposed to be the primary dispersal mechanism in nature, it is possible that wind may serve as a long-distance dispersal mechanism. However, the role of dispersal vectors other than water has not been investigated.

Habitat and Lifestyle

Colletotrichum gloeosporioides is reported in the USDA fungal database to infect over 800 host plant species, distributed worldwide (Farr *et al*., 2008). The number of host species reflects the generalist infection strategy of *C. gloeosporioides*. However, the paucity of information regarding taxonomic boundaries within the *C. gloeosporioides* species complex does not instill confidence in the lack of host preference indicated by this extensive list of host plants.

Infection Strategies among *Colletotrichum spp.*

A range of infection strategies has been described, both within and among species of *Colletotrichum*. The first steps of infection appear to be consistent within the genus, beginning with adherence of spores to host tissue, followed by spore germination to produce germ tubes. These germ tubes differentiate into melanized appresoria (a swelling in the germ tube), which later produce infection pegs that penetrate the host cuticle. At this point in development, there is divergence of infection strategies among species. These strategies range from "intracellular hemibiotrophy to subcuticular/intramural necrotrophy" (O'Connell *et al*., 2000). Intracellular hemibiotrophy is characterized by an initial biotrophic phase, characterized by obtaining nutrients from the living cytoplasm of the host, after initial penetration. The fungus then spreads by large intracellular primary hyphae, but still does not induce the expression of symptoms. This stage is followed by the necrotrophic phase, initiating from less than 24 hours to more than 3 days after biotrophy. The necrotrophic phase produces narrow secondary hyphae that secrete cell wall degrading enzymes, ahead of the colonizing hyphae, which destroy host cells (O'Connell *et al*., 2000).

Subcuticular, intramural necrotrophy is characterized by the colonization of periclinal and anticlinal cell walls of the epidermal cells and does not result in symptomatic expression until 24 hours after penetration. Hyphae continue to proliferate asymptomatically and later penetrate host cells, resulting in necrotic lesions. The fungus spreads intra- and intercellularly. This strategy does not result in differentiated primary and secondary hyphae. Most species of *Colletotrichum* appear to be limited to a single infection strategy, while others - including *C. magna* on watermelon and *C. gloeosporioides*, employ both strategies in the same tissue (O'Connell *et al*., 2000). The ability of *C. gloeosporioides* to employ both strategies may be an important mechanism for infecting a diversity of host species and tissues.

THE AMERICAN CRANBERRY (*VACCINIUM MACROCARPON* AITON), AS A HOST

The American cranberry (*Vaccinium macrocarpon* Aiton) is native to the temperate regions of North America. It is distributed in natural stands throughout northeastern North America (USDA, NRCS, 2008). This member of the Ericaceae is an evergreen trailing shrub, flowering from June to August, in bogs throughout its range (Gleason, Cronquist, 1991). The cranberry was utilized by native Americans as a food source and introduced to European settlers as a treatment for scurvy and dysentery (Foster, Duke, 2000). The species was brought into cultivation in the early 19th century in Massachusetts and New Jersey, but has remained a relatively undomesticated crop. Many of the dominant varieties used today were selected prior to the Civil War from wild stands on Cape Cod (USDA-CSREES, 2001). Major production areas now include Massachusetts, Wisconsin, New Jersey, British Columbia, Washington and Oregon (Stiles, Oudemans, 1999).

Many of the cultivation methods developed in the early days of cranberry production are still employed today. Cranberry "beds" are engineered such that water flow into and among beds can be regulated. Beds are constructed of peat or sand, within a meter of the water table, and are surrounded by dikes and ditches that form a network connecting the beds in a single field. Water is employed throughout the season, via overhead irrigation for frost-protection, maintenance of growth, as well as during harvest. Cranberry beds are flooded in the fall when the fruits are ready for harvest. Fruits are then dislodged, allowing them to float to the surface of the water, where they can be collected. Fruits harvested in this way are sent for processing into juices and preserves, but those sold on the fresh market are typically dry harvested (Oudemans *et al.*, 1998). The National Agricultural Statistics service has estimated that 38,800 acres of cranberries were harvested in 2007, yielding slightly more than 6.5 million barrels (100 lbs/barrel), the majority of which is sent for processing (USDA-NASS, 2008). The estimated total value of the crop is over 300 million dollars a year.

Colletotrichum on cranberry

Two closely related species of *Colletotrichum*, *C. gloeosporioides* and *C. acutatum*, have been known to infect *Vaccinium macrocarpon* (cranberry) in areas of commercial cranberry production. The two species can be readily distinguished on the basis of growth rate, spore shape, and colony color (Sutton, 1992).

Colletotrichum gloeosporioides has been known as a fruit rot pathogen of cranberry since the early 1900's (Shear, 1907a). Cranberry fruit rot can be caused by several fungal species, however *C. gloeosporioides* was revealed as one of the most prevalent and widespread fungal species in New Jersey cranberry bogs causing fruit rot (Stiles, Oudemans, 1999). *Colletotrichum gloeosporioides* has also been implicated as the causal agent of stem canker of cranberry, a disease that has emerged recently in New Jersey (Oudemans, 2008). In addition, other species sympatric with cranberry are known to host *C. gloeosporioides* (Bills, Polishook, 1992; Farr *et al.*, 2008), and may serve as a reservoir of the cranberry pathogen. Gene flow between populations on alternate host species and populations in cranberry bogs can have important implications for management practices. If *C. gloeosporioides* on cranberries is part of a broader population on sympatric host species, management practices that focus solely on the eradication of *C. gloeosporioides* on cranberry populations may have limited effectiveness.

RESEARCH PLAN

The research presented here proceeded in two phases. The first phase consisted of acquiring *C. gloeosporioides* from wild and cultivated cranberry populations, as well as isolating strains from various other host plant species and geographic localities. This was followed by the generation of a multilocus molecular dataset to ascertain the phylogenetic diversity of *C. gloeosporioides* in and around cranberry bogs in eastern North America (Chapter 2). This allowed for the identification of a clade with which to carry out population level analyses. The first phase of research provided a set of isolates with which to carry out the second phase of research.

The second phase of research consisted of testing microsatellite primers designed from GenBank ESTs, sequence data, and a microsatellite enriched genomic library, and genotyping isolates to locate polymorphic loci in the isolated strains of *C. gloeosporioides.* The isolation of microsatellites and genotyping of the isolated strains of *C. gloeosporioides* allowed for determining the genetic structure and reproductive mechanisms in field populations. Isolates were also characterized for variation at the 3' end of *Cgret*, a long terminal repeat retrotransposon (LTR) that has been shown to be variable within populations of *C. gloeosporioides* on cranberry, to utilize as an additional marker for assessing population structure (Zhu, 1999; Zhu, Oudemans, 2000).

REFERENCES

Bailey JA, O'Connell RJ, Pring RJ, Nash C (1992) Infection strategies of *Colletotrichum* species. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 88-120. CAB International, Surrey, UK.

Bills G, Polishook JD (1992) Recovery of endophytic fungi from *Chamaecyparis thyoides*. *Sydowia*, **44**, 1-12.

Boland GJ (2004) Fungal viruses, hypovirulence, and biological control of *Sclerotinia* species. *Canadian Journal of Plant Pathology*, **26**, 6-18.

Cannon P, Bridge PD (2000) Linking the past, present and future of *Colletotrichum* systematics. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Dov Prusky, Freeman S, Dickman MB), pp. 1-20. American Phytopathological Society, St. Paul, Minnesota.

Chilton SJP, Lucas GB, Edgerton CW (1945) Genetics of *Glomerella* III: Crosses with a conidial strain. *American Journal of Botany*, **32**, 549-554.

Chilton SJP, Wheeler HE (1949a) Genetics of *Glomerella* VI: Linkage. *American Journal of Botany*, **36**, 270-273.

Chilton SJP, Wheeler HE (1949b) Genetics of *Glomerella* VII: Mutation and segregation in plus cultures. *American Journal of Botany*, **36**, 717-721.

Cisar CR, TeBeest DO (1999) Mating system of the filamentous ascomycete, *Glomerella cingulata*. *Current Genetics*, **35**, 127-133.

Colon-Garay J, Rivera-Vargas L, McGovern RJ, Rodriquez RdP (2002) Hypovirulent isolates of *Colletotrichum gloeosporioides* induce resistance to anthracnose in detached mango fruits and seedlings. *J. Agric. Univ. Puerto Rico*, **84**, 56-64.

Correll JC, Guerber JC, Wasliwa LA, Sherrill JF, Morelock TE (2000) Inter- and intra-species variation in *Colletotrichum* and mechanisms which affect population structure. In: *Colletotrichum: Host Specificity, Pathology and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 145-179. The American Phytopathogical Society Press, St. Paul, Minnesota.

Crouch JA, Clarke BB, Hillman BI (2006) Unraveling the evolutionary relationships among the divergent lineages of *Colletotrichum* causing anthracnose disease in turfgrass and corn. *Phytopathology*, **96**, 46-60.

Dutech C, Enjalbert J, Fournier E*, et al.* (2007) Challenges of microsatellite isolation in fungi. *Fungal Genetics and Biology*, **44**, 933-949.

Farr DF, Rossman AY, Palm ME, McCray EB (2008) Fungal Databases. Systematic Mycology and Microbiology Laboratory, ARS, USDA.

Fávaro LCdL, Araújo WLd, Azevedo JLd, Paccola-Meirelles LD (2005) The biology and potential for genetic research of transposable elements in filamentous fungi. *Genetics and Molecular Biology*, **28**, 804-813.

Foster S, Duke JA (2000) *A Field Guide to Medicinal Plants and Herbs of Eastern and Central North America*, 2nd edn. Houghton Mifflin, New York.

Freeman S, Katan T, Shabi E (1996) Characterization of Colletotrichum gloeosporioides isolates from avocado and almond fruits with molecular and pathogenicity tests. *Applied and Environmental Microbiology*, **62**, 1014-1020.

Gleason H, Cronquist A (1991) *Manual of vascular Plants of Northeastern United States and Adjacent Canada* New York Botanical Garden, New York.

González E, Sutton TB, Correll JC (2006) Clarification of the etiology of Glomerella Leaf Spot and Bitter Rot of apple caused by Colletotrichum spp. based on morphology and genetic, molecular, and pathogenicity tests. *Phytopathology*, **96**, 982-992.

Guozhong L, Cannon PF, Reid A, Simmons CM (2004) Diversity and molecular relationships of endophytic *Colletotrichum* isolates from the Iwokrama Forest Reserve, Guyana. *Mycological Research*, **108**, 53-63.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* **11**, 424-429.

Johnston PR, Jones D (1997) Relationships among *Colletotrichum* Isolates from fruit-rots assessed using rDNA sequences. *Mycologia*, **89**, 420-430.

Kaufmann PJ, Weidemann GJ (1996) Isozyme analysis of *Colletotrichum gloeosporioides* from five host genera. *Plant Disease*, **80**, 1289-1293.

Kendrick B (2000) *The Fifth Kingdom* Mycologue Publications, Sidney.

Kirk PM, Cannon PF, David JC, Stalpers JA (2001) Dictionary of the Fungi. CAB International, Surrey.

Lucas GB (1946) Genetics of *Glomerella* IV: Nuclear phenomena in the ascus. . *American Journal of Botany*, **33**, 802-806.

Lucas GB, Chilton SJP, Edgerton CW (1944) Genetics of *Glomerella* I: Studies on the behavior of certain strains. *American Journal of Botany*, **31**, 233-239.

Maccheroni WJ, Araújo WL, Azevedo JL (2004) Ambient pH-regulated enzyme secretion in endophytic and pathogenic isolates of the fungal genus *Colletotrichum*. *Scientia Agricola*, **61**, 298-302.

MacKenzie SJ (2005) *Population structure and pathogenicity of Colletotrichum gloeosporioides from strawberry and noncultivated hosts in Florida* Ph.D., University of Florida.

MacKenzie SJ, Seijo TE, Legard DE, Timmer LW, Peres NA (2007) Selection for pathogenicity to strawberry in populations of *Colletotrichum gloeosporioides* from native plants. *Phytopathology*, **97**, 1130-1140.

McDonald BA, Linde C (2002) The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica*, **124**, 163-180.

McGahen JW, Wheeler HE (1951) Genetics of *Glomerella* IX: Perithecial development and plasmogamy. . *American Journal of Botany*, **38**, 610-617.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, **34**, 457-477

Nicholson RL, Moraes WBC (1980) Survival of *Colletotrichum graminicola*: importance of the spore matrix. *Phytopathology*, **70**, 255-261.

O'Connell RJ, Perfect S, Hughes B*, et al.* (2000) Dissecting the cell biology of *Colletotrichum* infection processes. In: *Colletotrichum: Host Specificity, Pahtology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 57-77. American Phytopathological Society, St. Paul, Minnesota.

Oudemans PV (2008) Personal Communication.

Oudemans PV, Caruso FL, Stretch AW (1998) Cranberry fruit rot in the Northeast: A complex disease. *Plant Disease*, **82**, 1176-1184.

Photita W, Taylor PWJ, Ford R, Hyde KD (2005) Morphological and molecular characterization of Colletotrichum species from herbaceous plants in Thailand. *Fungal Diversity*, **18**, 117-133.

Phoulivong S, Cai L, Chen H*, et al.* (2010) *Colletotrichum gloeosporioides* is not a common pathogen on tropical fruits. *Fungal Diversity*, **44**, 33-43.

Prihastuti H, Cai L, Chen H, McKenzie EHC, Hyde KD (2009) Characterization of *Colletotrichum* species associated with coffee berries in northern Thailand. *Fungal Diversity*, **39**, 89-109.

Roa AC, Chavarriaga-Aguirre P, Duque MC*, et al.* (2000) Cross-species amplification of Cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. *American Journal of Botany*, **87**, 1647-1655.

Rodriguez RJ, Redman RS (2000) *Colletotrichum* as a model system for defining the genetic basis of fungal symbiont life styles. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 114-130. American Phytopathological Society, St. Paul, Minnesota.

Rojas EI, Rehner SA, Samuels GJ*, et al.* (2010) Colletotrichum gloeosporioides s.l. associated with Theobroma cacao and other plants in Panama´: multilocus phylogenies distinguish host-associated pathogens from asymptomatic endophytes. *Mycologia*, **102**, 1318-1338.

Shear CL (1907) Cranberry Disease (ed. USDA). Government Printing Office.

Silva DN, Talhinas P, Várzea V*, et al.* (2012) Application of the Apn2/MAT locus to improve the systematics of the Colletotrichum gloeosporioides complex: an example from coffee (Coffea spp.). *Mycologia*, **104**, 396-409.

Stiles CM, Oudemans PV (1999) Distribution of cranberry fruit-rotting fungi in New Jersey and evidence for nonspecific host resistance. *Phytopathology*, **89**, 218-225.

Sutton BC (1992) The genus *Glomerella* and its anamorph *Colletotrichum*. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 1-26. CAB International, Surrey, UK.

Taylor JW, Geiser DM, Burt A, Koufopanou V (1999a) The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews*, **12**, 126-146.

Taylor JW, Jacobson DJ, Fisher MC (1999b) The evolution of asexual fungi: Reproduction, speciation, and classification. *Annual Review of Phytopathology*, **37**, 197-245.

Tu JC (1992) *Colletotrichum lindethumianum* on bean: Population dynamics of the pathogen and breeding for resistance. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 203-224. CAB International, Surrey, UK.

USDA-CSREES (2001) Crop Profile for Cranberries in Massachusetts.

USDA-NASS (2008) Noncitrus Fruits and Nuts 2007 Summary

USDA, NRCS (2008) The PLANTS Database. National Plant Data Center, Baton Rouge, LA

Vaillancourt L, Wang J, Hanau R (2000) Genetic regulation of sexual compatability in *Glomerella graminicola*. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 29-44. American Phytopathological Society, St. Paul, Minnesota.

Vailllancourt L, Du M, Wang J, Rollins J, Hanau R (2000) Genetic analysis of cross fertility between two self-sterile strains of *Glomerella graminicola*. *Mycologia*, **92**, 430-435.

Wheeler HE (1950) Genetics of *Glomerella* VIII: A genetic basis for the occurrence of minus mutants. *American Journal of Botany*, **37**, 304-312.

Wheeler HE, McGahen JW (1952) Genetics of *Glomerella* X: Genes affecting sexual reproduction. *American Journal of Botany*, **39**, 110-119.

Wheeler HE, Olive LS, Ernest CT, Edgerton CW (1948) Genetics of *Glomerella* V: Crozier and ascus development. *American Journal of Botany*, **35**, 722-728.

Zhu P (1999) *Genetic diversity and systematics of Colletotrichum species on cranberry (Vaccinium macrocarpon)* Ph.D., Rutgers University.

Zhu P, Oudemans PV (2000) A long terminal repeat retrotransposon *Cgret* from the phytopathogenic fungus *Colletotrichum gloeosporioides* on cranberry. *Current Genetics*, **38**, 241-247.

Table 1: Genotype, phenotype, and fertility relationships at loci controlling sexual reproduction in *Colletotrichum gloeosporioides* / *Glomerella cingulata* (adapted from Wheeler and McGahen, 1952)

#	**Genotype**	**Phenotype**	**Self-Fertility**
1	A^+B^+	Clumped perithecial	High
2	A^+B^1	Scattered perithecial	Moderate
3	A^+B^2	Scattered perithecial	+/- self-sterile
4	A^1B^+	Clumped conidial	self-sterile
5	A^1B^1	Scattered conidial	self-sterile
6	A^1B^2	Scattered conidial	self-sterile
7	A^2B^+	Small clumped perithecial	High
8	A^2B^1	Black scattered perithecial	Low
9	A^2B^2	Black scattered perithecial	+/- self-sterile

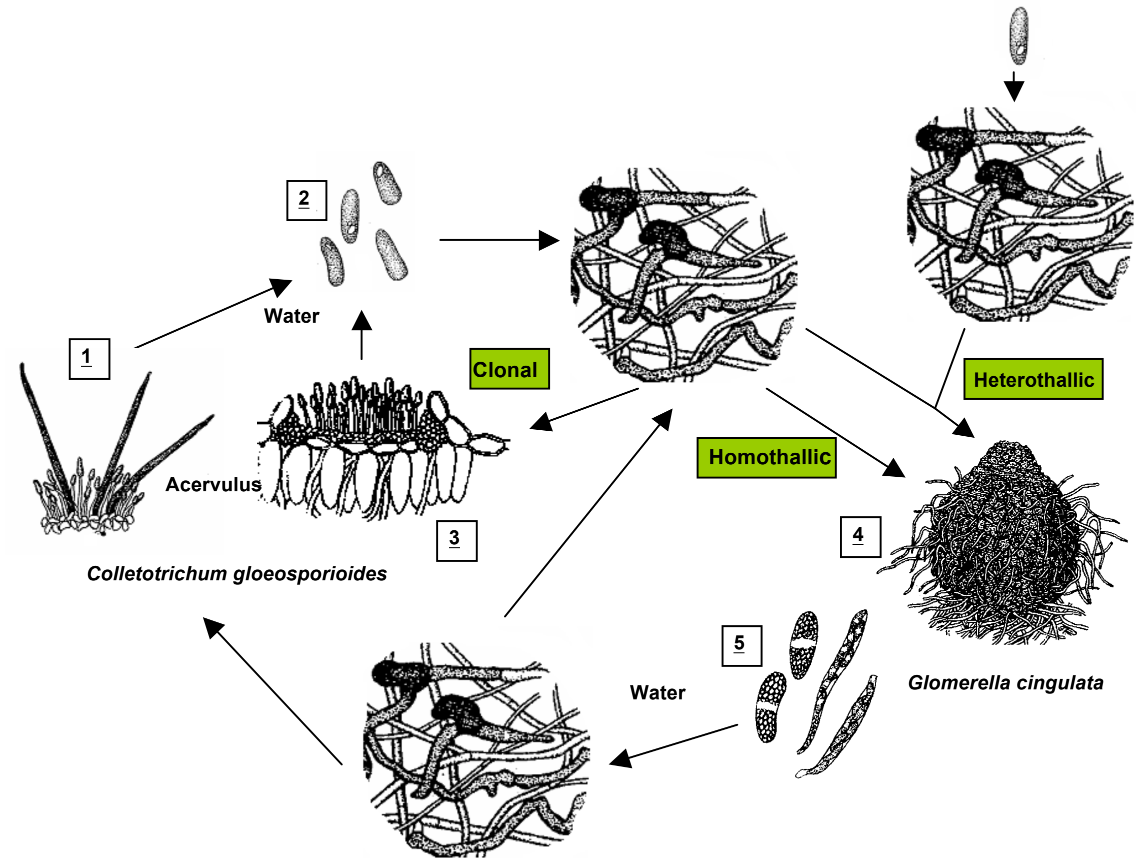


Figure 1: Generalized lifecycle of *Colletotrichum gloeosporioides/Glomerella cingulata*. The following descriptions correspond to numbers in the figure: (1) An acervulus is produced in host plant tissue from which (2) conidia are produced asexually and are putatively dispersed by water. A conidium germinates to produce a self-fertile or self-sterile thallus which will either (3) form a new acervulus and through mitosis produce more asexual spores or (4) undergo sexual reproduction through outcrossing (heterothallism) or self-fertilization (homothallism) to produce a perithecium, where (5) meiospores or ascospores are produced and dispersed by water to infect new plant tissue. Line drawings taken from (Shear, 1907a)

CHAPTER 2

HABITAT AND HOST INDICATE LINEAGE IDENTITY IN *COLLETOTRICHUM GLOEOSPORIOIDES S.L.* FROM WILD AND AGRICULTURAL LANDSCAPES IN NORTH AMERICA

Formatted for and submitted to PLoSONE

This work is not intended to represent valid publication of nomenclatural changes until published in PLoSONE or another peer-reviewed publication.

ABSTRACT

Understanding the factors that drive the evolution of pathogenic fungi is central to revealing the mechanisms of virulence and host preference, as well as developing effective disease control measures. Prerequisite to these pursuits is the accurate delimitation of species boundaries. *Colletotrichum gloeosporioides s.l.* is a species complex of plant pathogens and endophytic fungi for which reliable species recognition has only recently become possible through a multi-locus phylogenetic approach. By adopting an intensive regional sampling strategy encompassing multiple hosts within and beyond agricultural zones associated with cranberry (*Vaccinium macrocarpon* Aiton), we have integrated North America strains of *Colletotrichum gloeosporioides s.l.* from these habitats into a broader phylogenetic framework. We delimit species on the basis of genealogical concordance phylogenetic species recognition (GCPSR) and quantitatively assess the monophyly of recognized species at each of four nuclear loci and in the combined data set with the genealogical sorting index (*gsi*). Our analysis resolved two principal lineages within the species complex. Strains isolated from cranberry and sympatric host plants are distributed across both of these lineages and belong to seven distinct species or terminal clades. Strains isolated from *V. macrocarpon* in commercial cranberry beds belong to four species, three of which are described here as new, *C. fructivorum sp. nov.*, *C. temperatum sp. nov.*, and *C. melanocaulon sp. nov.* Another species, *C. rhexiae,* is epitypified. Intensive regional sampling has revealed a combination of factors, including the host species from which a strain has been isolated, the host organ of origin, and the habitat of the host species, as useful indicators of species identity. We have identified three broadly distributed temperate species, *C. fructivorum*, *C. rhexiae*, and *C. nupharicola*, which will be useful for understanding the microevolutionary forces that may lead to species divergence in this important genus of endophytes and plant pathogens.

INTRODUCTION

Delimiting species boundaries among fungi lays the groundwork for detailing the natural history and ecology of species and defines a robust framework from which further comparative studies can be designed [i.e. population genetics/genomics]. This is also prerequisite to providing targeted and effective disease control measures and identifying specific pathogens against which plant breeders can focus their efforts in developing and selecting disease resistant cultivars. Strictly agro-centric studies of plant pathogens risk sampling too narrowly, overlooking important adjacent (parapatric) niches driving pathogen evolution. Extensive sampling, both within and adjacent to agricultural landscapes, has the potential to provide a broader view of the natural history and evolution of plant pathogens. Uncovering reservoirs of pathogen diversity beyond agricultural zones can bring into focus selective forces that drive the evolution of pathogen traits important for survival in agricultural landscapes and expand our understanding of the ecology and population dynamics of these microorganisms (Morris *et al.*, 2009). Expanding the sampling of plant pathogens to wild habitats and non-agricultural host plants, coupled with objective molecular analytical approaches, should provide an essential foundation for the design of comparative population genetic and functional genomics studies that seek to clarify the evolution of host specificity versus generalism, pathogenicity, and virulence as well as other differential traits among closely related lineages.

Colletotrichum Corda is among the most important genera of plant-associated fungi, causing disease and occurring as asymptomatic endophytes on aerial organs of a broad range of host plants (Bailey, Jeger, 1992; Guozhong *et al.*, 2004; Latunde-Dada, 2001). *Colletotrichum gloeosporioides s.l.* represents an aggregate of species frequently reported as a dominant endophyte of tropical herbaceous plants and is known pan-globally as a field and post-harvest fruit pathogen of many economically important crops (Adaskaveg, Forster, 2000; Legard, 2000; Oudemans *et al.*, 1998; Sutton, 1992; von Arx, 1957). Morphological homoplasy and phenotypic plasticity have previously thwarted efforts to clearly define species boundaries within the species complex, necessary if we are to develop a greater understanding of the ecology and natural history of each lineage. However, recent development of molecular markers suitable for resolving species limits and phylogenetic relationships within this species aggregate have been proposed and validated (Crouch *et al.*, 2006; Du *et al.*, 2005; Rojas *et al.*, 2010; Silva *et al.*, 2012b). Therefore, it is now possible to examine the role of geography, host preference/specificity, the nature of host-pathogen/host-endophyte associations, and other niche specialization attributes that may underlie species divergence (Rojas *et al.*, 2010).

The current study uses a phylogenetic approach to infer species boundaries among North American members of the *C. gloeosporioides* species complex, particularly those in association with the large American cranberry (*Vaccinium macrocarpon* Aiton) and sympatric plant species. *Colletotrichum gloeosporioides s.l.* has been reported as a leaf and fruit pathogen of cranberry since the late 1800s (Halsted, 1889) and early 1900s (Shear, 1907b) and has recently been observed to colonize stem tissue (Stiles, Oudemans, 1997). Contemporary studies of fruit pathogens have confirmed the importance of *C. gloeosporioides s.l.* in agricultural ecosystems throughout the cultivated range of *V. macrocarpon* (Olatinwo *et al.*, 2003; Oudemans *et al.*, 1998; Polashock *et al.*, 2009; Zhu, 1999). However, given the fact that *C. gloeosporioides s.l.* is an aggregate of species that can be difficult or impossible to distinguish morphologically, it is not clear that the strains isolated from cranberry are conspecific throughout the cultivated range of cranberry. In addition, genetic studies have focused on isolates solely from fruit and have not investigated diversity from alternate host organs or sympatric host species, despite evidence that these represent potential reservoirs of diversity for the species complex in cranberry agricultural areas (Bills, Polishook, 1992; Ellis, Everhart, 1894; Johnson *et al.*, 1997; Shear, 1907b; Stiles, Oudemans, 1999).

Vaccinium macrocarpon is one of North America's few native, economically important crop species. Cranberry has undergone little selection from wild relatives and is cultivated in several regions alongside extant native populations (Rodriguez-Saona *et al.*, 2011). It can also be found in sympatry with the closely related species *V. oxycoccos* L., which has been used for inter-specific hybridization in breeding programs but is not cultivated for agricultural production (Vorsa, Polashock, 2005). In addition, many of the cultivars used in past breeding regimes remain available for further experimentation. These characteristics provide cranberry breeders with excellent resources for meeting the needs of cranberry growers to improve fruit production and reduce pathogen pressure through breeding for disease resistance (Vorsa, 1994). Providing a refined systematic understanding of the pathogens that are the source of disease pressure will help to focus the efforts of plant breeders.

Host specificity/preference has historically been a focal criterion for species delimitation in *Colletotrichum* (Sutton, 1992),(von Arx, 1957). Similarly, contemporary studies have indicated the presence of host specific lineages (Freeman *et al.*, 1996),(Crouch *et al.*, 2006). However, host-fungus associations within *Colletotrichum* are variable. It is clear that multiple species are capable of infecting single hosts and, conversely, some *Colletotrichum* species are capable of infecting multiple host species (Photita *et al.*, 2005),(Crouch *et al.*, 2006),(Rojas *et al.*, 2010). In addition, there is indication that different plant organ types may also act as selective forces leading to organ specialization in *Colletotrichum* species (González *et al.*, 2006). While *C. gloeosporioides s.l.* is established as a pathogen of a variety of important crop species worldwide (Agostini, Timmer, 1994; Chakraborty *et al.*, 2002; Colon-Garay *et al.*, 2002; Freeman *et al.*, 1996; Kaufmann, Weidemann, 1996; MacKenzie, 2005; Oudemans *et al.*, 1998; Rojas *et al.*, 2010; Silva *et al.*, 2012b), this study represents a unique perspective into the evolution of the species complex by investigating the genetic diversity of *C. gloeosporioides s.l.* in a single crop species and its surrounding habitat within the host's native range. The specific objectives of this study are to determine: (1) whether there are multiple sympatric lineages within the species complex that infect cranberry, 2) whether host preference/specificity is evident among sympatric lineages within the species complex, 3) whether host organ or nature of the fungus-host association is predictive of phylogenetic structure, and 4) whether we can identify lineages with broad geographical and/or host associations suitable for fine-scale landscape genetic analysis. We sampled horizontally across five sympatric host species in wild and commercial cranberry bogs in order to address questions related to host specificity, and sampled vertically among different plant organs within *V. macrocarpon* to target questions related to organ specialization (Table 1). Utilizing the recent development of molecular phylogenetic markers useful for distinguishing lineages within the species complex (Crouch *et al.*, 2006; Du *et al.*, 2005; Rojas *et al.*, 2010) and isolates from five major cranberry agricultural areas, wild cranberry bogs, fruit and stem of *V. macrocarpon*, and from five sympatric host species, we assess the lineage diversity among isolates from cranberry and surrounding habitats.

METHODS

Ethics Statement

All necessary permits were obtained for the described field studies from the Delaware Division of Parks and Recreation, the National Forest Service, and The Nature Conservancy. All other samples collected in this study were on private land and did not require permits. There were no endangered or protected species collected for this study.

Fungal isolation and culturing

Colletotrichum was isolated from symptomatic and asymptomatic tissue of several host species in North America with a focus on species sympatric with *Vaccinium macrocarpon* in wild and agricultural habitats. Sympatric host plant species from which *C. gloeosporioides s.l.* was isolated include *Vaccinium oxycoccos, Rhexia virginica* L.*, Chamaecyparis thyoides* (L.) Britton, Sterns, & Poggenb.*,* and *Nuphar lutea* (L.) Sm. (Table 1). *Colletotrichum gloeosporioides s.l.* was isolated from symptomatic and asymptomatic stem and fruit of *V. macrocarpon* (Table 1) in both wild and agricultural habitats. Both symptomatic and asymptomatic tissue was surface sterilized in 10% bleach (final concentration 0.6125% sodium hypochlorite) between 1.5 minutes and 5 minutes, depending on the permeability of the tissue, and plated on V8 juice agar, 2% malt extract agar (MEA: BD Diagnostics, Franklin Lakes, NJ, USA) or corn meal agar (CMA: BD Diagnostics), with the exception of *Nuphar lutea. Nuphar lutea* was incubated in a humid chamber under ambient light and conidia were transferred from anthracnose lesions to CMA and potato dextrose agar (PDA: BD Diagnostics) for purification. Isolates were characterized as endophytes if they were isolated from asymptomatic tissue after surface sterilization; all other isolates were considered pathogenic. Strains morphologically similar to *Colletotrichum gloeosporioides s.l.*, based on an assortment of characters including growth rate, colony color, hyphal morphology and/or conidial shape and size, were isolated by transferring conidia or hyphal tips to sterile media and preserved on CMA slants stored at 6°C, on CMA stored in 1.5 mL microcentrifuge tubes at 6°C, and in 10% glycerol at -80°C. Types, epitypes, and representative cultures were deposited at the Centraalbureau voor Schimmelcultures (CBS) with corresponding dried cultures deposited as vouchers at the U.S. National Fungus Collections (BPI).

DNA extraction and PCR amplification

Isolates were grown on potato dextrose broth (Difco) for 5-7 days before mycelium was harvested, blotted dry on sterile paper towels and dehydrated for 6-10 hours in a vacuum centrifuge on low heat. Approximately 20 mg of dried tissue was used for DNA extraction using the DNeasy Plant MiniKit (Qiagen Inc., Valencia, CA, USA) or a standard phenol-chloroform extraction method after tissue homogenization in the FastPrep FP120 (MP Biomedicals, LLC., Solon, OH, USA).

PCR amplification reactions of *nrITS*, *beta-tubulin* (*tub2*), *apn2*, and *apn2/matIGS* were performed in 25 µl reactions containing 9.3 µL of autoclaved ion-exchanged water, 2.5 µL of dNTP mixture (2.5 mM stock of each dNTP), 2.5 µL of bovine serum albumin (BSA; 0.25 µg µL−1 stock), 2.5 µL of buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 0.25 % (w/v) cresol red], 5 µL betaine (1.2 M stock), 1 µL of each primer (0.67 µM final concentration), 0.2 µL of *Taq* polymerase (GenScript USA Inc., Piscataway, NJ, USA) and 1 µL of DNA. The *nrITS* region was amplified and sequenced using primers ITS5 and ITS4 (White *et al.*, 1990), *tub2* with primers T1 and T224 (O'Donnell, Cigelnik, 1997), *apn2* with CgDL_R1 and ColDL_F3 (Rojas *et al.*, 2010), and *apn2mat/IGS* using primers CgDL_F6 and CgMAT1_F2 (Rojas *et al.*, 2010).

PCR reactions were run on an Eppendorf vapo.protect Mastercycler pro S with the following cycling parameters: *nrITS*—initial denaturing for 2 min at 94°C, followed by 38 cycles of 94°C for 1 min, 55°C for 30 s, and 72°C for 45 s, followed by a final extension at 72°C for 5 min; *tub2*— initial denaturing for 2 min at 94°C, followed by 30 cycles of 94°C for 35 s, 52°C for 55 s, and 72°C for 2 min, followed by a final extension at 72°C for 5 min; *apn2/matIGS*—initial denaturing for 5 min at 95°C, followed by 10 cycles of 95°C for 30 s, 62°C (decreasing by 1°C each cycle) for 30 s, and 72°C for 1 min, followed by 35 cycles of 95°C for 30 s, 52°C for 30 s, and 72°C for 1 min, followed by a final extension at 72°C for 10 min. PCR products were purified and sequenced at the High Throughput Genomics Unit, Department of Genome Sciences, University of Washington.

Contig assembly, sequence editing, and phylogenetic inference

Sequences were automatically assembled into contigs, and edited manually in Sequencher version 4.9 (GeneCodes Corp., Ann Arbor, Michigan). Alignments were carried out with the online version of the sequence alignment program MAFFT (Katoh *et al.*, 2005) using the iterative refinement option G-INS-i for each locus independently.

Independent gene trees were inferred under the maximum parsimony optimality criterion in TNT (Goloboff *et al.*, 2008) and maximum likelihood criterion in RAxML–HPC2 (7.2.8) implemented on the CIPRES Science Gateway portal (Miller *et al.*, 2010; Stamatakis, 2006; Stamatakis *et al.*, 2008). Phylogenetic analysis of the concatenated four gene matrix was inferred under maximum parsimony in TNT, maximum likelihood in RAxML–HPC2 (implemented on the CIPRES cluster), and Bayesian inference in MrBayes v3.1.2 (Huelsenbeck, Ronquist, 2001; Ronquist, Huelsenbeck, 2003). Phylogenetic analyses in TNT were carried out with parsimony ratchet tree searches and TBR branch swapping. Statistical support for the inferred nodes was determined with parsimony bootstrapping on 1000 replicate datasets. Rapid bootstrapping in RAxML was carried out implementing the GTRCAT model and the ML tree search under the GTRGAMMA model. The best-fit models for Bayesian analyses were selected with MrModeltest 2.3 (Nylander, 2004).

Independent analyses of *nrITS* and partial *beta-tubulin* were rooted in TNT and RAxML to *Coll57* and *Coll60*, strains of *Colletotrichum aff. acutatum* Simmonds, a species complex closely related to *Colletotrichum gloeosporioides s.l* (Phoulivong *et al.*, 2010; Simmonds, 1965; Than *et al.*, 2008). Strains *Coll57* and *Coll60* were identified based on morphological similarity and NCBI BLAST similarity searches of *nrITS* and partial *beta-tubulin*. Analyses of *apn2* and *apn2/matIGS* were rooted in TNT and RAxML to strain *4766*, which exploratory *nrITS* and *tub2* analyses and a previous study of the group (Rojas *et al.*, 2010) indicated is a suitable outgroup for inferring phylogenetic relationships within *C. gloeosporioides s.l.*. Statistical support for the inferred nodes was determined in RAxML–HPC2 by bootstrapping with the number of replicates for the independent gene trees determined by implementation of the extended majority rule (MRE) convergence criterion implemented in the CIPRES portal and the combined dataset with 1000 replicates. Node frequencies for both maximum parsimony and maximum likelihood bootstrap analyses were calculated with the SumTrees 3.1.0 program using the DendroPy Phylogenetic Computing Library version 3.7.1 (Sukumaran, Holder, 2010).

The best-fit models implemented in MrBayes for each of *nrITS*, *tub2*, *apn2* and *apn2/matIGS* respectively, were as follows: SYM+Γ, GTR+I, GTR+Γ, HKY+ Γ. Four parallel runs were conducted with one cold and three heated Markov chains per run for 10,000,000 generations sampling every 1,000 generations, with each of these models implemented for each locus in MrBayes. Posterior probabilities were calculated from the sampled trees after discarding the first 25% as burn-in. To assess convergence to a global optimum in both ML and Bayesian tree searches of the combined dataset, the log-likelihood scores (-lnL) for the best tree from the maximum likelihood and Bayesian tree searches were calculated in RAxML (Stamatakis, 2006).

Quantitative measures of genealogical sorting

The genealogical sorting index is used here to represent a quantitative assessment of the monophyly or genealogical exclusivity of a group of commonly labeled terminals in a set of trees. In order to determine the degree of exclusive ancestry of each of the well supported lineages inferred from the concatenated dataset across the independent gene trees, the bootstrap trees obtained from the maximum parsimony bootstrap analyses for each independent gene and the concatenated dataset were used to calculate the genealogical sorting index (*gsi*) for each lineage (Cummings *et al.*, 2008). The *gsi* was calculated for lineages of the *Colletotrichum gloeosporioides* species aggregate using the R package genealogicalSorting version 0.91 (Bazinet *et al.*, In preparation; R-Development-Core-Team, 2011) after rooting the trees with *Coll57* (*nrITS* & *tub2*) or *4766* (*apn2* & *apn2/matIGS*) then removing outgroup strains (*Coll57, Coll60, 4766, 4801, 4766*). The genealogical sorting index (*gsi*) reaches a maximum when a commonly labeled group (species or terminal lineage) reaches monophyly in a tree or set of trees and a minimum when all nodes on the tree are required to unite a group. P-values represent the probability that the calculated *gsi* from the inferred tree or trees would be observed by chance. The ensemble genealogical sorting index (gsi_T) is the sum of the *gsi* values for each topology weighted by the probability of a given topology based on the proportion of trees in the sample where it is represented. P-values are calculated by permutation of group labels on the terminals of a tree and determining the frequency distribution from the recalculated *gsi* value. P-values are estimated from 1,000 permutations of each dataset and represent the probability of observing *gsi* values ≥ to the reported *gsi* values by chance under the null hypothesis that labeled groups are of mixed ancestry.

Morphological Studies

Morphological observations were made from strains cultured on potato dextrose agar (PDA), corn meal agar (CMA), V8 agar, and Synthetischer nährstoffarmer agar (SNA) (Nirenberg, 1976) incubated at 22-25°C. Growth rates were determined from cultures grown in darkness at 25°C on PDA. Cultures were first grown on CMA and 10 mm plugs were taken from the expanding margin and transferred to 20 mL of PDA in 100 x 15 mm Petri dishes. Each strain was plated to three replicate plates. Three radial measurements were taken from the edge of the plug to the margin of the colony every 24-48 hours over the course of 5 days by marking the bottom of the plate at the margin of the colony using a stereomicroscope equipped with stage lighting, resulting in 9 radial measurements per strain. Each plate was then photographed and the distance between markings was measured.

Microscopic observations of conidia, phialides, and ascospores were made from specimens mounted in water. Hyphal appressoria were observed using slide cultures; a $10mm^2$ block of CMA was placed on a CMA plate, each of the four corners of the block was inoculated, covered with a sterile coverslip, and incubated at 25°C. Microscopic observation of perithecial development on cranberry fruit was made by surface sterilizing symptomatic field collected fruit, cutting in half transversely, placing face down on V8 agar, and incubating at room temperature (22°C) for approximately 3 weeks. The fruit was removed from the plate after perithecial development and fixed in FAA (3.7% formaldehyde, 5% glacial acetic acid, and 50% ethanol) before dehydration in an alcohol-xylene dehydration series and embedded in Paraplast X-TRA (Leica Microsystems, Buffalo Grove, IL, USA). The fruit was sectioned at 8 µm and stained with acid fuchsin and Cotton blue in lactic acid and mounted in Permount (Fisher Scientific, Pittsburgh, PA, USA).

Microscopic images were made with a Nikon DXM1200C digital camera attached to a Zeiss Axioplan compound microscope or with a Nikon SMZ1500 stereoscope equipped with a Nikon DXM1200F digital camera using Nikon ACT-1 software. Measurements were made using ImageJ 1.44p software (Abramoff *et al.*, 2004; Rasband, 1997-2008) and summary statistics calculated in R version 2.13.0. Growth rates are reported giving the minimum, average, maximum [(minimum-) average (-maximum)] and the standard deviation in millimeters per day. All other measurements additionally include the 1st quartile and the 3rd quartile with the following notation: (minimum-) 1st quartile – average – 3rd quartile (-maximum).

Nomenclature

New names contained in the validly published version of this work have been submitted to MycoBank from where they will be made available to the Global Names Index. The unique MycoBank number can be resolved and the associated information viewed through any standard web browser by appending the MycoBank number contained in the validly published version of this publication to the prefix http://www.mycobank.org/MycoTaxo.aspx?Link=T&Rec=.

RESULTS

Species assignments

In order to delimit novel species, we applied the criteria of genealogical concordance phylogenetic species recognition (GCPSR) (Dettman *et al.*, 2003; Taylor *et al.*, 2000). Briefly, following Dettman et al. (2003), novel species were recognized if they satisfied one of two criteria: genealogical concordance or genealogical non-discordance. Clades were genealogically concordant if they were inferred in three of the four single gene trees and genealogically non-discordant if they were strongly supported (MP≥75%; ML≥70%) in a single gene and not contradicted at or above this level of support in any other single gene tree. In addition, novel species were recognized if resolved with strong support in all analyses for the combined dataset (PP≥.95; ML≥70%; MP≥75%). All isolates were assigned to a species based on the combined analysis, with the exception of *GJS0857* and *Coll940,* which remain singletons.

Individual gene trees

In order to test the hypothesis that multiple sympatric lineages within the *C. gloeosporioides* species complex infect cranberry in the field, North American isolates need to be placed in a broader phylogenetic context. Therefore, sequence data generated in this study were combined with an earlier study of isolates of *C. gloeosporioides s.l.* from the New World tropics (Rojas *et al.*, 2010). Outgroup sampling was expanded from the aforementioned study to include two isolates of *C. aff. acutatum*. Individual locus data were analyzed separately to assess the topological congruence among datasets and the utility of each to resolve terminal lineages with robust statistical support for sister group relationships within the *C. gloeosporioides* species complex. Strain data are summarized in Table 1 and character and tree statistic data are summarized in Table 2.

Sequence data from *nrITS* has been widely used in fungal phylogenetic studies and has been proposed as a barcode locus for fungi (Seifert, 2009). However, our analysis indicates it neither provides adequate resolution for reliable species assignment, nor does it reliably assess phylogenetic relationships within the *C. gloeosporioides* species complex, as has been reported in previous studies of *Colletotrichum* (Crouch *et al.*, 2009a; Polashock *et al.*, 2009; Rojas *et al.*, 2010). Despite the low phylogenetic resolution inferred from *nrITS* data, six nodes within the species complex were supported in more than 75% of the maximum parsimony bootstrapped datasets (Figure S1) and eight nodes in more than 70% of the maximum likelihood bootstrapped datasets (Figure S5). In addition, *C. sp. indet. F* (*4766*, *3386*, *4801*) was determined to be closely related but peripheral to the *C. gloeosporioides* species complex, indicating this is a suitable outgroup, as previously suggested by Rojas et al. (2010).

Phylogenetic analysis of partial *tub2* sequence data largely supported the inferences made from *nrITS* data, but provided further resolution within the species complex, recovering 20 well-supported nodes in both maximum parsimony and maximum likelihood analyses (Figures S2 and S6). Isolates from cranberry (*V. macrocarpon* and *V. oxycoccos*) were distributed among 5 clades, including lineages originally described from tropical regions.

Phylogenetic analysis of the *apn2* locus provided greater resolution for terminal lineages than that achieved with either *tub2* or *nrITS*. Bootstrap analyses provided significant statistical support for 23 nodes and 20 nodes for MP (Figure S3) and ML (Figure S7) analyses, respectively, with six clades in both analyses containing isolates from *V. macrocarpon* and *V. oxycoccos*. While support for terminal lineages from *apn2* sequence data was greater than that from partial *tub2*, support for sister group relationships was not as robust.

A similar topology was recovered from the analysis of nucleotide data from *apn2/matIGS* , the intergenic spacer bridging the *apn2* and mating-type locus, with respect to the *nrITS*, *tub2*, and *apn2* gene trees while providing greater resolution than the other three datasets. Bootstrap analyses of *apn2/matIGS* recovered strong branch support for 25 and 29 nodes for MP (Figure S4) and ML (Figure S8), respectively, and in agreement with the apn2 analysis, six lineages include isolates from *V. macrocarpon* and *V. oxycoccos*. In addition to increased resolution of terminal clades within the species complex, compared with *apn2*, support for sister group relationships within the species complex is stronger than inferences based on partial *tub2*.

While statistical node support based on bootstrap resampling was somewhat inconsistent among maximum likelihood and maximum parsimony analyses, both sets of analyses were concordant, with differing levels of resolution. Similarly, the resolution of terminal clades among independent gene trees is consistent with a few exceptions. One strain, isolated as a leaf endophyte of *Persea americana* (*Coll11*), was relegated with strong support to a terminal clade that includes isolates designated as *Colletotrichum sp. indet. B* based on analysis of *nrITS*, but this isolate is strongly supported in both the partial *tub2* and *apn2/matIGS* tree as belonging to a lineage that includes isolates of *C. tropicale.* The resolution of this isolate in the *apn2* analysis was ambiguous. The phylogenetic placement of another strain, *GJS0857*, included in the study of New World *Colletotrichum* by Rojas et al. (2010), is inconsistent among gene trees. While *GJS0857* forms a sister group relationship with *C. sp. indet. D* in the *apn2* analyses, it is resolved as sister to a group that includes *C. melanocaulon, C. sp. indet. B, C. sp. indet. C,* and *C. sp. indet. D* with ≥ 99% bootstrap support in the *apn2/matIGS* analyses. Lastly, *Coll940* is a strain that is resolved as sister to a group that includes *C. ignotum, C.* ignotum 2*, C. nupharicola* and *C. sp. indet. E* in the *tub2* (ML) and *apn2* (MP) analyses, while its placement is switched with *C. sp. indet. E* in the *apn2/matIGS* analyses.

Phylogenetic analysis of the concatenated dataset

In order to infer organismal phylogenetic relationships and make species assignments, we relied on the combined dataset of 3,122 nucleotide characters from four nuclear loci and eighty-four terminals. The Bayesian consensus tree is presented in Figure 1 with node posterior probability and bootstrap proportion values from Bayesian, parsimony, and maximum likelihood analyses. The outgroup taxa, *C. aff. acutatum* and *C. sp. indet. F* have been trimmed from Figure 1 due to long branches between the species complex and the outgroup terminals. The relationships among clades in each of the three analyses converged on a topology, with varying levels of node support, identical to the Bayesian analysis presented in Figure 1. The log-likelihood of the tree with the best score (highest log-likelihood value) from the set of four Bayesian runs as calculated in RAxML under the GTRGAMMA model (-1298.97) was identical to the tree inferred under the maximum-likelihood criterion.

Thirty-six nodes were inferred with strong bootstrap and posterior probability support within the species complex. The multilocus analyses corroborated the majority of the single gene trees in recovering *C. theobromicola* as sister to two principal lineages within the species complex. The first lineage includes two previously described species, *C. kahawae* and *C. rhexiae,* and two newly described species, *C. temperatum* and *C. fructivorum.* The second lineage includes four previously described species, *C. tropicale, C. ignotum, C. nupharicola,* and *C. gloeosporioides*, one new species, *C. melanocaulon* and five undescribed sublineages (two of which, *C. sp. indet. A* and *C. sp. indet. B*, are being described elsewhere by S. Rehner). Added taxon sampling calls the monophyly of *C. ignotum* into question, the biphyletic status of which was discussed in Rojas et al (2010).

Genealogical sorting indices

The genealogical sorting index represents a quantitative assessment of the exclusive ancestry of each species in a set of bootstrap trees for each individual locus and in the combined dataset. Terminal labels for this analysis match the species assignments based on the combined four-locus analysis, as previously described. The results of these analyses are presented in Table 3 and the taxon assignments are presented in Table 4.

Colletotrichum gloeosporioides, C. sp. indet. A and *C. theobromicola* reached the maximum ensemble *gsi* (gsi_T) value of 1 when calculated across all bootstrapped trees from the *nrITS* data. Several species including *C. kahawae*, *C. tropicale, C. ignotum, C. nupharicola,* and *C. ignotum2* had low, but significant, gsi_T values based on the *nrITS* trees ranging from

0.064 to 0.221. Four species, *C. fructivorum, C. rhexiae, C. temperatum,* and *C. sp. indet. B* had moderate values of gsi_T ranging from 0.370 to 0.658. The remaining species, *C. melanocaulon, C. sp. indet. E, C. sp. indet. C,* and *C. sp. indet. D* had non-significant gsi_T values across the *nrITS* bootstrapped trees.

Unlike the gsi_T values of the *nrITS* trees, all values calculated from the *tub2* trees were significant with the exception of *C. kahawae. Colletotrichum temperatum, C. tropicale, C. sp. indet. A, C. nupharicola, C. ignotum 2, C. sp. indet. E, C. gloeosporioides,* and *C. theobromicola* reached the maximum gsi_T value of 1 when calculated across all bootstrapped trees. *Colletotrichum rhexiae, C. melanocaulon, C. sp. indet. B,* and *C. sp. indet. C,* had low but significant gsi_T values ranging from 0.114 to 0.220. *Colletotrichum fructivorum, C. sp. indet. D* and *C. ignotum* had moderate to high gsi_T values ranging from 0.466 to 0.751.

All gsi_T values of the *apn2* trees were significant. *Colletotrichum kahawae* had the lowest value of 0.11, while all others were moderate to high. *Colletotrichum rhexiae, C. sp. indet. A, C. melanocaulon, C. sp. indet. D, C. nupharicola, C. ignotum2, C. sp. indet. E, C. gloeosporioides,* and *C. theobromicola* had the maximum gsi_T value of 1 across all bootstrapped trees. *Colletotrichum fructivorum, C. temperatum, C. tropicale, C. sp. indet. B, C. sp. indet. C,* and *C. ignotum* had moderate to high gsi_T values ranging from 0.494 to 0.986.

Similarly, all gsi_T values of the *apn2mat/IGS* trees were significant. *Colletotrichum melanocaulon* had the lowest value of 0.241, while all others were moderate to high. *Colletotrichum fructivorum, C. rhexiae, C. tropicale, C. sp. indet. A, C. sp. indet. C, C. ignotum, C. nupharicola, C. gloeosporioides,* and *C. theobromicola* had gsi_T values of 1 when calculated across all bootstrapped trees, while *C. kahawae, C. temperatum, C. sp. indet. B, C. sp. indet. D, C. ignotum2,* and *C. sp. indet. E* had gsi_T values ranging from 0.447 to 0.901.

All gsi_T values of trees from the combined dataset were significant with high values ranging from 0.740 to 1. The only species that did not reach the maximum possible gsi_T value of 1 were *C. kahawae* (0.740) and *C. sp. indet. B* (0.952). All values based on the combined four-locus dataset represent strong measures of genealogical divergence across all 1000 bootstrap replicates.

Taxonomy and morphology

Multilocus phylogenetic analysis of *Colletotrichum gloeosporioides s.l.* strains isolated from *V. macrocarpon* and other sympatric host species revealed several distinct lineages within the species complex. Comparison of growth rates and conidial morphology indicates that there is significant morphological overlap between *Colletotrichum* species isolated from *V. macrocarpon* and sympatric host species as well as between strains representing additional species within the species complex. *Colletotrichum nupharicola* is exceptional, exhibiting a very slow growth rate, with significantly longer and wider conidia than all other species included in this study. A comparison of growth rates and conidial dimensions is presented in Figure 2 and Figure 3. Three new species are described below and *C. rhexiae* is epitypified. Conidial, ascospore, and colony morphology is shown in Figure 4 for each new species and *C. rhexiae*. Seta morphology is shown in Figure 5.

Colletotrichum fructivorum Doyle, Oudemans & Rehner, sp. nov.

Similar to *Colletotrichum rhexiae* Ellis & Everh. but setae less abundant, the interquartile range of the ascospore length to width ratio smaller (3.2-3.9 um) and the interquartile range of the conidial length to width ratio larger on CMA. Common fruit-rot pathogen of cranberry (*Vaccinium macrocarpon* Aiton) in commercial production.

Growth rate (3.0-) 5.4 (-6.8) mm per day with standard deviation of 1.1 mm on PDA at 25°C [n=54]; aerial mycelium floccose, white to greyish white, medium grey and brownish grey in some strains; sectoring common. Perithecia developing and maturing on V8 agar, clustered or solitary, dark brown to black, globose to obpyriform to papillate; ascospores allantoid, olive yellow, (15.7-) 16.6-18.2-18.5 (-26.9) X (3.8-) 4.5-5.1-5.6 (-7.1) µm with length/width ratio (2.6-) 3.2-3.7-3.9 (-5.6) µm [n=30]. On CMA and SNA perithecial fundaments with melanized hyphae radiating from the base but mature ascospores not observed. On SNA aerial mycelium flocculose; conidial masses flesh to light orange, (12.4-) 13.2-14.3-15.2 (-18.6) X (3.8) 4.3-4.9-5.6 (-6.2) µm with length/width ratio (2.0-) 2.4-3.0-3.5 (-4.3). On CMA mycelium barely visible; conidial masses usually abundant, embedded in the medium and on the surface, flesh to light orange; phialides tapering towards the tip, monoblastic, (9.0-) 10.5-12.9-14.2 (-18.1) X (2.9-) 3.7-3.9-4.1 (-4.9) µm [n=30] at the widest point; conidia subcylindrical to slightly tapered with obtuse apices, (12.4-) 13.6-14.3-14.9 (-19.1) X (3.7-) 4.7-5.0-5.2 (-6.2) µm with length/width ratio (2.3-) 2.7-2.9-3.1 (-4.6) µm [n=240]; hyphal appressoria melanized, clavate or irregular, shallowly lobed, terminal, (3.9-) 8.0-9.4-10.8 (-15.2) X (3.1-) 3.9-4.5-4.7 (-7.7) µm with length/width ratio (0.6-) 1.8-2.2-2.6 (-3.5) µm [n=30]

Habitat and Distribution: Commonly isolated as a fruit-rot pathogen and from asymptomatic tissue of *Vaccinium macrocarpon* in commercial cranberry beds throughout North America. Also isolated from *Rhexia virginica* from asymptomatic infections growing in commercial cranberry beds and symptomatic fruit of *V. oxycoccos* in a wild cranberry bog in Pennsylvania.

Etymology: The specific epithet, “fructivorum”, refers to the propensity of the species to be associated with fruit-rot of cranberry. From the Latin *fructus,* fruit, and *–vorous*, eating.

Holotype: USA. New Jersey: isolated as a fruit-rot pathogen of *Vaccinium macrocarpon,* Burlington County, 39.8128 N, 74.6399 W, Oct 2010, *V. Doyle, P.V. Oudemans, C. Constantelos Coll1414*.

Additional specimens examined: USA. New Jersey: isolated as a stem endophyte of *V. macrocarpon,* Burlington County, 39.9518 N, 74.5005 W, Oct 2010, *V. Doyle Coll1092*; isolated from *Rhexia virginica,* Burlington County, 39.7348 N, 74.5120 W, Sept 2009, *V. Doyle Coll864*; isolated as a fruit-rot pathogen of *V. macrocarpon*, Burlington County, 39.9518 N, 74.5005 W, Oct 2009, *V. Doyle, P.V. Oudemans Coll445*; isolated as a fruit-rot pathogen of *V. macrocarpon*, Burlington County, Fall 2007, *P.V. Oudemans* GLO7-17 (culture=*V. Doyle Coll21*); Collected by C.L. Shear, deposited to CBS in Apr 1922 (CBS124.22). Massachusetts: isolated as a fruit-rot pathogen of *V. macrocarpon*, Plymouth County, Nov 2010, *V. Doyle Coll1062*; isolated as a fruit-rot pathogen of *V. macrocarpon*, Plymouth County, Nov 2010, *V. Doyle Coll1081*; isolated as a fruit-rot pathogen of *V. macrocarpon*, *P.V. Oudemans* (culture *V.Doyle Coll116*). Wisconsin: isolated as a fruit-rot pathogen of *V. macrocarpon*, Monroe County, 44.0685 N, 90.4015 W, Sept 2010, *P.V. Oudemans, C. Constantelos, E. Zeldin* WI-1 (culture *V. Doyle Coll1164*); isolated as a fruit-rot pathogen of *V. macrocarpon*, Wood County, 44.3854 N, 90.0151 W, Sept 2010, *P.V. Oudemans, C. Constantelos, E. Zeldin* WI-53 (culture *V. Doyle Coll1216*); isolated as a fruit-rot pathogen of *V. macrocarpon*, Juneau County, 44.2072 N, 90.0995 W, Sept 2010, *P.V. Oudemans, C. Constantelos, E. Zeldin* WI-27 (culture *V. Doyle*

Coll1190); Delaware: isolated as a fruit-rot pathogen of *V. macrocarpon,* Kent County, 39.3130 N, 79.5555 W, Fall 2009, *P.V. Oudemans, C. Constantelos* GLO9-1 (culture *V. Doyle Coll873*). Pennsylvania: isolated as a fruit-rot pathogen of *V. oxycoccos*, Monroe County, Tannersville Cranberry Bog Preserve, 41.0376 N, 75.2648 N, *V. Doyle Coll886*. CANADA. British Columbia: isolated as a fruit-rot pathogen of *V. macrocarpon*, Oct 2010, *V. Doyle, P.V. Oudemans, B. Mouza Coll1002;* isolated as a fruit-rot pathogen of *V. macrocarpon*, Oct 2010, *V. Doyle, P.V. Oudemans, B. Mouza Coll1004;*

Notes: *Glomerella rufomaculans var. vaccinii* was described by C.L. Shear in 1907 (Shear, 1907c) from leaves of *Vaccinium macrocarpon* in New Jersey. The type material was a slide (slide *no. 1447A C.L.S.*) of a single ascospore isolate made from this collection deposited in the "pathological collection of the Department of Agriculture", now housed in the collections at the Systematic Mycology and Microbiology Laboratory in Beltsville, Maryland (BPI). While other slides designated by C.L. Shear as type material for new species published in the same protologue were located at BPI, slide *no. 1447A C.L.S.* was not located and is thought to be lost. However, a culture deposited by Shear in April, 1922 to Centraalbureau voor Schimmelcultures in the Netherlands was obtained, sequenced and found to be conspecific with other isolates described here. The nomenclature used by CBS in designating this isolate as *Glomerella rufomaculans-vaccinii,* however, seems to be based on a typographical error as C.L. Shear clearly intended recognition of this taxon at the varietal level and there is no indication in the literature that it has been formally elevated to species. Furthermore, *Glomerella rufomaculans* was subsumed into *Glomerella cingulata* (anamorph: *Colletotrichum gloeosporioides*) with the revisionary work of von Arx in 1957 (von Arx, 1957), necessitating the designation of the new specific epithet, *Colletotrichum fructivorum.* We have chosen a recently isolated strain as the ex-type strain due to the lack of sporulation observed in the culture deposited by Shear and the fact the holotype designated here is a single ascospore isolate for which we have morphological data for both the anamorph and the teleomorph (Figure 4).

Colletotrichum melanocaulon Doyle, Oudemans & Rehner sp. nov.

Similar to *Colletotrichum gloeosporioides* (Penz.) Sacc. but associated with stem canker of *Vaccinium macrocarpon* Aiton.

Growth rate (4.9-) 6.4 (-7.3) mm per day with standard deviation of 0.8 mm on PDA at 25°C [n=18]; aerial mycelium floccose, white to light grey, sectoring observed; fertile perithecia not observed; conidiomata forming abundantly in concentric rings in type strain, conidial masses orange. On SNA aerial mycelium flocculose; conidiomata abundant, conidial masses light orange; phialides tapering towards the tip, monoblastic, (8.6-) 11.5-13.6-15.3 (-19.0) X (2.1-) 2.8-3.3-3.9 (-4.8) µm [n=34] at the widest point; conidia subcylindrical to slightly tapered with obtuse apices, (11.3-) 12.6-13.1-13.4 (-14.8) X (4.4-) 5.1-5.5-5.9 (-6.3) µm with length/width ratio (2.1-) 2.3-2.4-2.5 (-3.0) µm [n=60]. On CMA mycelium barely visible, conidial masses pale orange to melon; phialides tapering toward the tip, monoblastic, (11.7-) 13.5-14.1-14.4 (-17.0) X (2.8-) 3.2-3.4-3.7 (-3.9) µm [n=10] at the widest point; conidia subcylindrical to slightly tapered with obtuse apices, (11.1-) 12.2-12.7-13.3 (-14.1) X (4.5-) 4.7-4.8-4.9 (-5.1) µm with length/width ratio (2.4-) 2.6-2.6-2.7 (-2.9) µm [n=30]; hyphal appressoria melanized, subglobose to clavate, not lobed, terminal, (4.4-) 5.8-6.3-7.2 (-8.2) X (3.3-) 4.1-4.5-4.9 (-5.9) µm with length/width ratio (1.0-) 1.1-1.4-1.6 (-1.9) µm [n=10].

Habitat and Distribution: Isolated from stem canker lesions of *Vaccinium macrocarpon* in commercial cranberry beds in New Jersey.

Etymology: The specific epithet, “melanocaulon”, refers to the brown or black stems from which the species was isolated. From the Greek, *melano*-, black or very dark, and *caulon*, stem, in agreement with the *neuter* generic name *Colletotrichum.*

Holotype: USA. New Jersey: isolated from cankered stems of *Vaccinium macrocarpon,* Burlington County, 39.7539 N, 74.5387 W, Aug 2008, *V. Doyle Coll131*.

Additional specimens examined: USA. New Jersey: isolated from cankered stems of *Vaccinium macrocarpon,* Burlington County, 39.7538 N, 74.5400 W, Aug 2008, *V. Doyle Coll126.*

Colletotrichum temperatum Doyle, Oudemans & Rehner, *sp. nov.*

Similar to *Colletotrichum rhexiae* Ellis & Everh. and *C. fructivorum* but ascospore length smaller, and ascospore length to width ratio smaller.

Growth rate (6.3-) 6.6 (-7.2) mm per day with standard deviation of 0.2 mm on PDA at 25°C [n=18]; aerial mycelium floccose, white to greyish white, sectoring observed. Mature ascospores not observed on PDA. On CMA aerial mycelium flocculose to barely visible; perithecia solitary to clustered, dark brown to black, subglobose to obpyriform, ascospores hemisphaeroid to reniform, olive brown, (13.5-) 14.3-14.7-15.3 (-17.1) X (4.6-) 5.3-5.5-5.7 (-5.8) µm with length/width ratio (2.3-) 2.6-2.7-2.8 (-3.3) µm [n=22]; conidial masses abundant on surface and embedded in agar, orange yellow; phialides tapering toward the tip, monoblastic, (7.7-) 11.3-13.9-15.7 (-22.6) X (2.2-) 3.3-3.4-3.7 (-4.2) µm [n=30] at the widest point; conidia subcylindrical to slightly tapered, rarely medially constricted, with obtuse apices, (13.5-) 14.5-14.9-15.2 (-17.2) X (3.7-) 4.5-4.6-4.7 (-4.9) µm with length/width ratio (2.9-) 3.1-3.3-3.4 (-4.0) µm [n=30]; hyphal appressoria melanized, irregular, shallowly to deeply lobed, terminal, (7.7-) 9.3-10.7-11.7 (-13.2) X (4.9-) 6.0-7.4-8.6 (-11.5) µm with length/width ratio (1.0-) 1.2-1.5-1.8 (-2.3) µm [n=50].

Habitat and Distribution: Isolated from rotten fruit of an ornamental cultivar, *Vaccinium macrocarpon* 'Hamilton', growing at the New York Botanical Garden and as a stem endophyte of *V. macrocarpon* in a commercial cranberry bog in New Jersey.

Etymology: The specific epithet, "temperatum", refers to the known distribution of the species. From the Latin, *temperatum*, temperate.

Holotype: USA. New York: isolated from rotten fruit of *Vaccinium macrocarpon,* Bronx County, The New York Botanical Garden, 40.8674 N, 73.8780 W, Nov 2009, *V. Doyle* & *C. Mozzicato Coll883*.

Additional specimens examined: USA. New Jersey: isolated as a stem endophyte of *Vaccinium macrocarpon,* Burlington County, 39.9514 N, 74.5008 W, Oct 2010, *V. Doyle Coll1103.*

Colletotrichum rhexiae Ellis & Everh. Proceedings of the Academy of Natural Sciences of Philadelphia 46: 372. 1894.

MycoBank MB178511

Growth rate (4.9-) 6.2 (-7.1) mm per day with standard deviation of 0.6 mm on PDA at 25°C [n=36]; aerial mycelium floccose, white to light grey to greyish brown, sectoring common; perithecia solitary to clustered, dark brown to black, , ascospores fusiform to lunate, greyish yellow (18.1-) 21.7-23.2-24.4 (-27.5) X (3.5-) 3.9-4.1-4.3 (-4.7) µm with length/width ratio (4.5-) 5.2-5.8-6.4 (-7.6) µm [n=60]; On SNA aerial mycelium flocculose; conidial masses flesh to light orange, phialides tapering toward the tip, monoblastic, (9.5-) 13.5-14.7-16.2 (-17.6) X (3.0-) 3.6-3.7-4.0 (-4.4) µm [n=15] at the widest point; conidia subcylindrical with obtuse apices, (11.7-) 13.4-14.0-14.3 (-27.9) X (3.9-) 5.4-5.8-6.4 (-7.4) µm with length/width ratio (1.8-) 2.1-2.5-2.6 (-4.4) µm [n=60] . On CMA mycelium barely visible; conidial masses abundant on the surface, flesh to light orange, phialides tapering towards the tip, monoblastic, (5.9-) 10.9-13.4-15.0 (-28.9) X (2.6-) 3.3-3.7-4.2 (-4.7) µm [n=30] at the widest point; conidia subcylindrical with obtuse apices, (9.1-) 13.1-13.7-14.4 (-21.7) X (4.4-) 5.1-5.5-5.8 (-6.7) µm with length/width ratio (1.9-) 2.4-2.5-2.6 (-3.9) µm [n=120]; hyphal appressoria melanized, clavate, obclavate to fusiform and irregular, shallowly to deeply lobed, terminal, (7.0-) 9.4-10.9-12.8 (-15.2) X (3.9-) 5.0.-6.1-7.1 (-9.7) µm with length/width ratio (1.0-) 1.3-1.9-2.2 (-3.4) µm [n=30].

Holotype: USA. Delaware: acervuli on several leaves of *Rhexia virginica*, Kimensi, Delaware, *A. Commons 2534*.

Epitype: isolated from lesioned stem tissue of *Rhexia virginica,* Sussex County, Delaware, Cape Henlopen State Park, 38.7778 N, 75.1062 W, Nov 2010, *V. Doyle Coll1026*;

Additional specimens examined: USA. Delaware: isolated from lesioned stem tissue of *Rhexia virginica,* Sussex County, Cape Henlopen State Park, 38.7772 N, 75.1075 W, Nov 2010, *V. Doyle Coll1038*; isolated from healthy fruit of *Vaccinium macrocarpon,* Sussex County, Cape Henlopen State Park, 38.7772 N, 75.1072 W, Nov 2009, *V. Doyle, P.V. Oudemans, C. Constantelos Coll877*; isolated from rotten fruit of *Vaccinium macrocarpon,* Sussex County, Cape Henlopen State Park, 38.7770 N, 75.1075 W, Nov 2010, *V. Doyle Coll1034*; isolated from rotten fruit of *Vaccinium macrocarpon,* Sussex County, Cape Henlopen State Park, 38.7772 N, 75.1075 W, Nov 2010, *V. Doyle Coll1041*. New Jersey: isolated from stem tissue of *R. virginica* in commercial cranberry bog, Burlington County, 39.7546 N, 74.5388 W, Sept 2010, *V. Doyle Coll952;* isolated from stem tissue of *R. virginica* in commercial cranberry bog, Burlington County, 39.7309 N, 74.5160 W, July 2010, *V. Doyle Coll1306*. Maryland: isolated from leaf anthracnose lesions of *R. virginica,* Prince George's County, 39.0308 N, 78.7886 W, Sept 2011, *V. Doyle* & *S.A. Rehner Coll1470*.

Notes: *Colletotrichum rhexiae* was described from leaves of *Rhexia virginica* collected in "Kimensi", Delaware in 1894. We attempted to find *R. virginica* in an area thought to be the type locality but were unable to find any host plants. The type material, consisting of *R.*

virginica leaves with circular lesions, is morphologically consistent *in vitro* with material collected in Cape Henlopen State Park in Delaware on *R. virginica* and is conspecific with material isolated from circular lesions on the same host in Maryland (Figure 6). We have decided to epitypify *C. rhexiae* with a living culture based on the necessity of molecular data for species delimitation in *Colletotrichum* and the need to evaluate morphological characters from a living culture.

DISCUSSION

Sympatric Lineages in North American cranberry bogs

Seven well-supported lineages of *C. gloeosporioides s.l.* were isolated from North American cranberry bogs, either from *V. macrocarpon, V. oxycoccos,* or other sympatric host plant species (Table 1 and Figure 1). Five of these lineages (*C. fructivorum*, *C. rhexiae*, *C. temperatum, C. melanocaulon*, and *C. ignotum*) were isolated from *V. macrocarpon* fruit and stems, either from diseased tissue or growing endophytically. However, all isolates from diseased fruit of *V. macrocarpon* in areas of agricultural production were restricted to the single lineage, *C. fructivorum*. *Colletotrichum melanocaulon* was isolated from stem cankers of *V. macrocarpon.* Strains isolated as stem endophytes from *V. macrocarpon* are members of three distinct lineages, *C. temperatum*, *C. ignotum,* and *C. fructivorum*.

Rhexia virginica L., an herbaceous perennial that grows in habitats with well-drained soils that are seasonally inundated, is often found sympatric with *Vaccinium macrocarpon* in wild and agricultural habitats in eastern North America. Strains isolated from *R. virginica* were resolved to three phylogenetically distinct lineages that were also isolated from *V. macrocarpon.* The first, *C. rhexiae,* includes isolates from wild *V. macrocarpon* populations and *R. virginica* in both wild and agricultural habitats. The second, *C. ignotum,* includes several tropical strains isolated from *Theobroma cacao* as well as three temperate strains, one from *V. macrocarpon* (*Coll1126*), another from *V. corymbosum* (*CollP1*), and a third from *R. virginica (Coll996).* The third lineage, *C. fructivorum,* is composed primarily of isolates from *V. macrocarpon* growing in commercial cranberry beds throughout the production range, but includes a single isolate from *R. virginica.*

Vaccinium oxycoccos and *Chamaecyparis thyoides* are additional species that have overlapping distributions with *V. macrocarpon. Vaccinium oxycoccos* is often sympatric with *V. macrocarpon* in the wild and *C. thyoides* is commonly found adjacent to commercial cranberry beds in New Jersey. Strains of *C. gloeosporioides s.l.* isolated from *V. oxycoccos* belong to two phylogenetically distinct lineages. *Coll886*, isolated from a diseased fruit of *V. oxycoccos* in Pennsylvania, is strongly supported as sister to isolates restricted to cranberry agricultural production zones and is conspecific with *C. fructivorum*. *Coll887*, isolated from a diseased fruit of *V. oxycoccos* in West Virginia, forms a phylogenetically distinct lineage (*C. sp. indet E*) with an endophytic isolate (*Coll920*) from *Chamaecyparis thyoides* growing as a weed in a commercial cranberry bed.

Colletotrichum nupharicola was previously described from *Nuphar lutea subsp. polysepala* (Engelm.) E.O. Beal in Washington and Idaho and *Nymphaea odorata* Aiton in Rhode Island. Our study indicates that this species is also present in irrigation reservoirs in agricultural cranberry beds in New Jersey, growing on *Nuphar lutea* (L.) Sm. *Colletotrichum nupharicola* is closely related to *C. ignotum* and *C. ignotum 2*, but sister group relationships among these lineages remain ambiguous.

Host preference, habitat range, and host organ specificity

Strains isolated from *V. macrocarpon* are distributed throughout the *C. gloeosporioides* species complex, with little indication of host specificity for most lineages. This study shows, however, that a composite of host, habitat, and host organ origin can be a useful indicator of lineage identity in some cases. Five lineages are represented by strains isolated from *Vaccinium macrocarpon* in cranberry production areas in New Jersey, but only two of those lineages, *C. temperatum* and *C. melanocaulon*, originate solely from *V. macrocarpon. Colletotrichum temperatum* has been isolated as a stem endophyte of *V. macrocarpon* and from rotten fruit in a horticultural variety (*V. macrocarpon* cv. Hamilton: not grown for agricultural production). The other species, *C. melanocaulon*, is associated with an emerging stem canker disease on cranberry where it has been isolated from affected stem tissue collected from field populations in commercial cranberry beds (personal observation).

As mentioned above, temperate *C. ignotum* has been isolated from healthy stem tissue of *V. macrocarpon* in a commercial cranberry bed in New Jersey, from *R. virginica* stem tissue, and from *V. corymbosum.* Likewise, isolates from *C. fructivorum* and *C. rhexiae* are not host specific; *C. fructivorum* has been isolated from *V. macrocarpon, V. oxycoccos* and *R. virginica*, while *C. rhexiae* has been isolated from *V. macrocarpon* and *R. virginica*. Nevertheless, all isolates originating from *V. macrocarpon* in wild habitats are conspecific with *C. rhexiae*. Likewise, all isolates from diseased fruit in cranberry agricultural production areas are conspecific with *C. fructivorum*.

The only North American lineage of *C. gloeosporioides s.l.* sympatric with cranberry with some indication of host preference over a broad geographical range is *C. nupharicola.* While *C. nupharicola* has been isolated from two aquatic host genera, *Nuphar* and *Nymphaea,* both of these genera are in the same family, Nymphaeaceae, indicating host specificity at the family level. This species is broadly distributed on Nymphaeaceae from Alaska through the Rocky Mountains to eastern North America (Dennis A. Johnson, pers. communication, and V. Doyle, unpublished data).

Geographic and host distribution of *Colletotrichum fructivorum*

This and other studies of *Colletotrichum gloeosporioides s.l.* have demonstrated the utility of delimiting species using multi-locus DNA sequence data. Given the cryptic nature of many species within the complex and the apparently broad diversity of species at small spatial scales, however, it is difficult to design an appropriate sampling scheme to capture the species diversity of *Colletotrichum* within a given host, habitat or geographic range. To develop a better understanding of scenarios that promote species divergence, it is necessary to understand the factors that lead to population substructure within a focal species. One of the objectives of this study was to identify one or more broadly distributed lineages within the *C. gloeosporioides* species complex that occur on cranberry to serve as a model system for investigation of the factors driving the divergence between populations. Given the widespread occurrence of *Colletotrichum gloeosporioides s.l.* as the causal agent of fruit rot in cranberry agricultural areas, the economic importance of cranberry production in North America, and the ease at which *C. gloeosporioides s.l.* can be isolated from diseased fruit in commercial cranberry production areas, we sampled diseased fruit across several of the major cranberry production areas in North America and Canada (Table 1). All strains of *C. gloeosporioides s.l.* isolated from fruit in commercial cranberry production areas were found to belong to *C. fructivorum* regardless of geographic origin. This reveals that *C. fructivorum* is broadly distributed in North America from Delaware to Massachusetts and west to Washington and British Columbia. In addition, there is evidence that strains of *C. fructivorum* are not organ specific (fruit and stems) and not host specific (*V. macrocarpon, V. oxycoccos,* and *R. virginica*). The combination of economic importance, broad geographic distribution, and diverse organ and host association makes this a suitable species for examining biotic and abiotic factors that influence microevolutionary processes in *Colletotrichum.*

Phylogenetic relationships and geographic distribution of temperate lineages

Fungal strains included in this study have been placed in a broader phylogenetic context to better understand the phylogenetic relationships among species within the *Colletotrichum gloeosporioides* species complex and to provide insight into possible phylogeographic patterns among New World temperate and tropical isolates. Combined analyses of four nuclear loci inferred strong support in ML and MP for *C. theobromicola* as sister to two major lineages within the species complex. The first lineage is represented primarily by temperate species, *C. temperatum, C. rhexiae,* and *C. fructivorum. Colletotrichum kahawae,* represented by the type strain originating in Kenya from *Coffea arabica* and three additional strains from *C. arabica* in Cameroon, is the only paleotropical species within this lineage.

The second lineage includes several species from both neotropical and nearctic ecozones. While temperate isolates, in most cases, do form clades that are reciprocally monophyletic with respect to closely related tropical species, there is no evidence for a single vicariance or dispersal event in or out of temperate regions. Temperate isolates are not monophyletic and one species, *Colletotrichum ignotum,* includes isolates from both temperate and tropical regions. However, the phylogenetic position of temperate *C. ignotum* strains (*Coll1126, Coll996,* and *CollP1*) is equivocal, likely due to inadequate phylogenetic signal in this part of the tree. The inadequacy of the data is further supported by the lack of resolution among *C. ignotum, C. nupharicola* and

C. ignotum 2. Despite the lack of obvious phylogeographic patterns among New World temperate and tropical isolates, additional geographical sampling and more powerful phylogenetic markers may provide the means to investigate more in-depth biogeographic hypotheses.

Colletotrichum rhexiae and *C. fructivorum* each present a unique opportunity to investigate the role of host and geography in shaping the microevolutionary patterns that may ultimately lead to species divergence. *Colletotrichum rhexiae* was described by Ellis and Everhart in the late 1800's (Halsted, 1889) from Delaware while *C. fructivorum* has been known from the same region since at least 1907 with its formal recognition by Shear (Shear, 1907b; Shear, 1907c) as *Glomerella rufomaculans var. vaccinii* and likely earlier as implicated by the work of Byron Halsted (Halsted, 1889). These species have apparently been sympatric for at least 105 years. While we do not have DNA evidence that *C. rhexiae* has been present for this extent of time, we have examined the original type material and it is not apparently distinct from our collections of *C. rhexiae* in New Jersey, Delaware and Maryland. In contrast, we do have a culture deposited to CBS in 1922 by Shear of what he considered to be *G. rufomaculans var. vaccinii* and it is identical in sequence across 4 loci to isolates collected from cranberry fruit in 5 US states and BC, from cranberry stem in New Jersey, and from *Rhexia virginica* growing as a weed in agricultural cranberry beds. The presence of sexual reproductive structures in both species suggests the potential for genetic exchange among conspecific individuals (unless they are obligately homothallic). However, despite their sympatry and ability to reproduce sexually, strong statistical support at the bifurcation between *C. rhexiae* and *C. fructivorum* indicates the presence of other biological factors reinforcing species limits. However, an objective test of species boundaries is better approached with more rapidly evolving markers variable within species, such as microsatellites, and using more appropriate non-phylogenetic algorithms to determine if there is any evidence of introgression. Further investigation using more rapidly evolving genetic markers could also lend the potential for inferring the reproductive strategy of these species and better understand the factors influencing species divergence within *Colletotrichum*. In addition, with the advancement of next-generation sequencing technology, it is possible to investigate these sister species at the genomic level to understand how genomic modifications (gene content or synteny) may influence pathogenicity on different hosts or in distinct organs.

CONCLUSIONS

Colletotrichum gloeosporioides s.l. has long been implicated as a pathogen of cranberry in commercial production areas, but a more refined understanding of the species responsible for fruit-rot of cranberry has not been possible due to the difficulty of species delimitation using morphological features. This study resolves this issue, identifying *C. fructivorum* as the species associated with fruit-rot in cranberry. This study also lays the groundwork for future studies regarding the natural history and ecology of members of *C. gloeosporioides s.l.* and should provide useful tools for plant breeders and plant pathologists, in their efforts to develop resistant cultivars for an industry that utilizes one of North America's few native crop species, *V. macrocarpon.* We have also determined that *C. fructivorum* is broadly distributed across North America and Canada in areas of commercial cranberry production and is capable of infecting alternate host species as well. This makes this species an appropriate model for addressing questions of population structure and dispersal at broad geographical and landscape level spatial scales. By virtue of a horizontal (among sympatric species) and vertical (among organs within a species) sampling scheme, we were able to uncover greater diversity of *C. gloeosporioides s.l.* in wild and commercial cranberry beds than has previously been suggested, revealing seven distinct lineages associated with *V. macrocarpon* and sympatric host species. Likewise, this level of sampling allowed us to determine that host specificity is not strongly implicated in determining macroevolutionary patterns among these species in temperate regions, with the exception of *C. nupharicola,* which has been isolated exclusively from Nymphaeaceae across a broad geographical area, yet has not been recovered from any other sympatric hosts. Similarly, despite the generalist nature of the species recovered in this study with respect to organ specificity, *C. fructivorum* does have an affinity for fruit of *V. macrocarpon* in commercial cranberry beds, to the exclusion of all other related species.

ACKNOWLEDGEMENTS

There are many people to thank for assistance with this work. Discussion with the following people helped to improve the manuscript: Peter Smouse, Larry Kelly, Drew Minnis, and Gary Samuels. The following people provided technical assistance: Natalia Pabon-Mora, Damon Little, Dennis Stevenson, Luiza Maciejak, and Josh Simpson. Chris Constantelos, Jen Vaicunas, Lindsay Wells, Donna Larsen, Patty McManus, Frank Caruso, Rob Hiller, Chris Bennett, Kent Karriker, Christine Worthington, Brian Mouza, Kim Patten, Eric Zeldin, Roger Spots, and Christina Mozzicato assisted with collections.

REFERENCES

1. Morris CE, Bardin M, Kinkel LL, Moury B, Nicot PC, et al. (2009) Expanding the Paradigms of Plant Pathogen Life History and Evolution of Parasitic Fitness beyond Agricultural Boundaries. PLoS Pathogens 5: 1-7.
2. Bailey JA, Jeger MJ, editors (1992) *Colletotrichum*: Biology, Pathology and Control: Wallingford, UK: CAB International.
3. Latunde-Dada AO (2001) *Colletotrichum*: tales of forcible entry, stealth, transient confinement and breakout. Molecular Plant Pathology 2: 187-198.
4. Guozhong L, Cannon PF, Reid A, Simmons CM (2004) Diversity and molecular relationships of endophytic *Colletotrichum* isolates from the Iwokrama Forest Reserve, Guyana. Mycological Research 108: 53-63.
5. Sutton BC (1992) The genus *Glomerella* and its anamorph *Colletotrichum*. In: Bailey JA, Jeger MJ, editors. Colletotrichum: Biology, Pathology and Control. Surrey, UK: CAB International. pp. 1-26.
6. Oudemans PV, Caruso FL, Stretch AW (1998) Cranberry fruit rot in the Northeast: A complex disease. Plant Disease 82: 1176-1184.

7. Adaskaveg JE, Forster H (2000) Occurrence and Management of Anthracnose Epidemics Caused by *Colletotrichum* Species on Tree Fruit Crops in California. In: Prusky D, Freeman S, Dickman MB, editors. *Colletotrichum*: Host Specificity, Pathology and Host-Pathogen Interaction. St. Paul, Minnesota: The American Phytopathogical Society Press. pp. 317-336.
8. Legard DE (2000) *Colletotrichum* diseases of Strawberry in Florida. In: Prusky D, Freeman S, Dickman MB, editors. *Colletotrichum*: Host Specificity, Pathology, and Host-Pathogen Interaction. St. Paul, Minnesota: American Phytopathological Society. pp. 292-299.
9. von Arx JA (1957) Die Arten der Gattung *Colletotrichum* Corda. Phytopathologische Zeitschrift 29: 414-468.
10. Du M, Schardl CL, Nuckles EM, Vaillancourt LJ (2005) Using mating-type gene sequences for improved phylogenetic resolution of *Colletotrichum* species complexes. Mycologia 97: 641-658.
11. Crouch JA, Clarke BB, Hillman BI (2006) Unravelling the evolutionary relationships among the divergent lineages of *Colletotrichum* causing anthracnose disease in turfgrass and corn. Phytopathology 96: 46-60.
12. Rojas EI, Rehner SA, Samuels GJ, Van Bael SA, Herre EA, et al. (2010) *Colletotrichum gloeosporioides* s.l. associated with *Theobroma cacao* and other plants in Panama´: multilocus phylogenies distinguish host-associated pathogens from asymptomatic endophytes. Mycologia 102: 1318-1338.
13. Silva DN, Talhinas P, Várzea V, Cai L, Paulo OS, et al. (2012) Application of the Apn2/MAT locus to improve the systematics of the *Colletotrichum gloeosporioides* complex: an example from coffee (*Coffea* spp.). Mycologia 104: 396-409.
14. Halsted B (1889) Some fungus diseases of the cranberry. New Jersey Agricultural College Experiment Station. pp. 1-40.
15. Shear CL (1907) Cranberry Diseases. In: USDA, editor: Government Printing Office.
16. Stiles CM, Oudemans PV (1997) Cranberry fruit rot pathogens isolated from various cranberry plant tissues over two growing seasons. Phytopathology 87: S94.
17. Zhu P (1999) Genetic diversity and systematics of *Colletotrichum* species on cranberry (*Vaccinium macrocarpon*) [Ph.D.]. New Brunswick: Rutgers University. 143 p.
18. Olatinwo RO, Hanson EJ, Schilder AMC (2003) A First Assessment of the Cranberry Fruit Rot Complex in Michigan. Plant Disease 87: 550-556.
19. Polashock JJ, Caruso FL, Oudemans PV, McManus PS, Crouch JA (2009) The North American cranberry fruit rot fungal community: a systematic overview using morphological and phylogenetic affinities. Plant Pathology 58: 1116-1127.
20. Ellis JB, Everhart BM (1894) New Species of Fungi from Various Localities Proceedings of the Academy of Natural Sciences of Philadelphia 46: 322-386.
21. Bills G, Polishook JD (1992) Recovery of endophytic fungi from *Chamaecyparis thyoides*. Sydowia 44: 1-12.
22. Johnson DA, Carris LM, Rogers JD (1997) Morphological and molecular characterization of *Colletotrichum nymphaeae* and *C. nupharicola* sp. nov. on water-lilies (*Nymphaea* and *Nuphar*). Mycological Research 101: 641-649.
23. Stiles CM, Oudemans PV (1999) Distribution of cranberry fruit-rotting fungi in New Jersey and evidence for nonspecific host resistance. Phytopathology 89: 218-225.
24. Rodriguez-Saona C, Vorsa N, Singh AP, Johnson-Cicalese J, Szendrei Z, et al. (2011) Tracing the history of plant traits under domestication in cranberries: potential

consequences on anti-herbivore defenses. Journal of Experimental Botany 62: 2633-2644.

25. Vorsa N, Polashock JJ (2005) Alteration of Anthocyanin Glycosylation in Cranberry Through Interspecific Hybridization. Journal of the American Society for Horticultural Science 130: 711-715.
26. Vorsa N. Breeding the American Cranberry. In: Roper TR, editor; 1994. Wisconsin State Cranberry Growers Association.
27. Freeman S, Katan T, Shabi E (1996) Characterization of *Colletotrichum gloeosporioides* isolates from avocado and almond fruits with molecular and pathogenicity tests. Applied and Environmental Microbiology 62: 1014-1020.
28. Photita W, Taylor PWJ, Ford R, Hyde KD (2005) Morphological and molecular characterization of *Colletotrichum* species from herbaceous plants in Thailand. Fungal Diversity 18: 117-133.
29. González E, Sutton TB, Correll JC (2006) Clarification of the etiology of *Glomerella* Leaf Spot and Bitter Rot of apple caused by Colletotrichum spp. based on morphology and genetic, molecular, and pathogenicity tests. Phytopathology 96: 982-992.
30. Agostini JP, Timmer LW (1994) Population dynamics and survival of strains of *Colletotrichum gloeosporioides* on Citrus in Florida. Phytopathology 84: 420-425.
31. Chakraborty S, Fernandes CD, Charchar MJdA, Thomas MR (2002) Pathogenic variation in *Colletotrichum gloeosporioides* infecting
Stylosanthes spp. in a center of diversity in Brazil. Phytopathology 92: 553-562.
32. Colon-Garay J, Rivera-Vargas L, McGovern RJ, Rodriquez RdP (2002) Hypovirulent isolates of *Colletotrichum gloeosporioides* induce resistance to anthracnose in detached mango fruits and seedlings. J Agric Univ Puerto Rico 84: 56-64.
33. Kaufmann PJ, Weidemann GJ (1996) Isozyme analysis of *Colletotrichum gloeosporioides* from five host genera. Plant Disease 80: 1289-1293.
34. MacKenzie SJ (2005) Population structure and pathogenicity of *Colletotrichum gloeosporioides* from strawberry and noncultivated hosts in Florida [Ph.D.]. United States -- Florida: University of Florida.
35. White TJ, Bruns T, LeeJ S, Taylor. JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR Protocols: A Guide to Methods and Applications. New York: Academic Press, Inc. pp. 315-322.
36. O'Donnell K, Cigelnik E (1997) Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. Molecular Phylogenetics and Evolution 7: 103-116.
37. Katoh K, Kuma K-i, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Research 33: 511-518.
38. Goloboff PA, Farris JS, Nixon KC (2008) TNT, a free program for phylogenetic analysis. Cladistics 24: 774-786.
39. Miller MA, Pfeifer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees; 2010; New Orleans, LA. pp. 1-8.
40. Stamatakis A (2006) RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. Bioinformatics 22: 2688-2690.
41. Stamatakis A, Hoover P, Rougemont J (2008) A Fast Bootstrapping Algorithm for the RAxML Web-Servers. Systematic Biology 57: 758-771.

42. Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. Bioinformatics 17: 754-755.
43. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.
44. Nylander JAA (2004) MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University. .
45. Than PP, Shivas RG, Jeewon R, Pongsupasamit S, Marney TS, et al. (2008) Epitypification and phylogeny of *Colletotrichum acutatum* J.H. Simmonds. Fungal Diversity 28: 97-108.
46. Phoulivong S, Cai L, Chen H, McKenzie EHC, Abdelsalam K, et al. (2010) *Colletotrichum gloeosporioides* is not a common pathogen on tropical fruits. Fungal Diversity 44: 33-43.
47. Simmonds JH (1965) A study of the species of *Colletotrichum* causing ripe fruit rots in Queensland. Queensland Journal of Agricultural and Animal Science 25: 437-459.
48. Sukumaran J, Holder MT (2010) DendroPy: A Python library for phylogenetic computing. Bioinformatics 26: 1569-1571.
49. Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. Evolution 62: 2411-2422.
50. Bazinet AL, Neel MC, Shaw KL, Cummings MP (In preparation) The genealogical sorting index: software and web site for quantifying exclusivity of lineages.
51. R-Development-Core-Team (2011) R: A Language and Environment for Statistical Computing.
52. Nirenberg H (1976) Untersuchungen u¨ ber die morphologische und biologische Differenzierung in der Fusarium-Sektion Liseola. Mitteilungen aus der Biologischen Bundesanstalt für Land-und Forstwirtschaft 169: 1-117.
53. Rasband WS (1997-2008) ImageJ. Bethesda, Maryland, USA: National Institutes of Health.
54. Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophotonics International 11: 36-42.
55. Seifert KA (2009) Progress towards DNA barcoding of fungi. Molecular Ecology Resources 9: 83-89.
56. Crouch JA, Clarke BB, Hillman BI (2009) What is the value of ITS sequence data in *Colletotrichum* systematics and species diagnosis? A case study using the falcate spored graminicolous *Colletotrichum* group. Mycologia 101: 648-656.
57. Dettman JR, Jacobson DJ, Taylor JW (2003) A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. Evolution 57: 2703-2720.
58. Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, et al. (2000) Phylogenetic Species Recognition and Species Concepts in Fungi. Fungal Genetics and Biology 31: 21-32.
59. Shear CL (1907) New species of fungi. The Bulletin of the Torrey Botanical Club 34: 305-317.

Table 1: Colletotrichum strains included in the phylogenetic analysis. Accession numbers for cultures deposited at Centraalbureau voor Schimmelcultures (CBS) are included for relevant strains.

Strain	**Locality**	**Host**	**Organ**	**Species**
Coll57	New Jersey	*Vaccinium macrocarpon*	Stem	*C. af. acutatum*
Coll60	New Jersey	*V. macrocarpon*	Fruit	*C. af. acutatum*
CBS124	New Jersey	*V. macrocarpon*	NA	*C. fructivorum*
Coll1002 = CBS133128	British Columbia	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1004	British Columbia	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1062	Massachussets	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1081 = CBS133121	Massachussets	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1092* = CBS133135	New Jersey	*V. macrocarpon*	Stem	*C. fructivorum*
Coll116	Massachussets	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1164 = CBS133126	Wisconsin	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1178	Wisconsin	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1190	Wisconsin	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1216	Wisconsin	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll1414 = CBS133125	New Jersey	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll445 = CBS133127	New Jersey	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll21 = CBS133133	New Jersey	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll864 = CBS133130	New Jersey	*Rhexia virginica*	Stem	*C. fructivorum*
Coll873	Delaware	*V. macrocarpon*	Fruit	*C. fructivorum*
Coll886 = CBS133124	Pennsylvania	*Vaccinium oxycoccos*	Fruit	*C. fructivorum*
Coll1026 = CBS133134	Delaware	*Rhexia virginica*	Stem	*C. rhexiae*
Coll1034 = CBS133136	Delaware	*V. macrocarpon*	Fruit	*C. rhexiae*
Coll1041	Delaware	*V. macrocarpon*	Fruit	*C. rhexiae*
Coll877* = CBS133132	Delaware	*V. macrocarpon*	Fruit	*C. rhexiae*

Table 1: continued

Coll1038	Delaware	*R. virginica*	Stem	*C. rhexiae*
Coll1306	New Jersey	*R. virginica*	Stem	*C. rhexiae*
Coll952 = CBS133131	New Jersey	*R. virginica*	Stem	*C. rhexiae*
Coll1470 = CBS133129	Maryland	*R. virginica*	Leaf	*C. rhexiae*
GJS08214	Cameroon	*Coffea arabica*	Fruit	*C. kahawae*
GJS08216	Cameroon	*C. arabica*	Fruit	*C. kahawae*
GJS08211	Cameroon	*C. arabica*	Fruit	*C. kahawae*
IMI319418	Kenya	*C. arabica*	Fruit	*C. kahawae*
Coll1103* = CBS133120	New Jersey	*V. macrocarpon*	Stem	*C. temperatum*
Coll883 = CBS133122	New York	*V. macrocarpon*	Fruit	*C. temperatum*
Coll11*	Florida	*Persea americana*	Leaf	*C. tropicale*
Coll918	Puerto Rico	*Terpsichore taxifolia*	Sporangium	*C. tropicale*
GJS0842	Panama	*Annona muricata*	Leaf	*C. tropicale*
4861	Panama	*T. cacao*	Leaf	*C. tropicale*
5101*	Panama	*T. cacao*	Leaf	*C. tropicale*
8401*	Panama	*T. cacao*	Leaf	*C. tropicale*
E1164*	Panama	*Trichilia tuberculata*	Leaf	*C. tropicale*
E2303	Panama	*Virola surinamensis*	Leaf	*C. tropicale*
E406	Panama	*Pentagonia macrophylla*	Leaf	*C. tropicale*
Q633*	Panama	*Cordia alliodora*	Leaf	*C. tropicale*
7423*	Panama	*Theobroma cacao*	Leaf	*C. tropicale*
Coll126 = CBS133123	New Jersey	*V. macrocarpon*	Stem	*C.melanocaulon*
Coll131 = CBSxxxxxx	New Jersey	*V. macrocarpon*	Stem	*C.melanocaulon*
Coll6	Florida	*Mangifera indica*	Leaf	*C. sp. indet. B*
1092	Panama	*Theobroma cacao*	Leaf	*C. sp. indet. B*

Table 1: continued

CollNC67	North Carolina	?	?	*C. sp. indet. B*
Coll54	New Jersey	*Prunus persica*	Fruit	*C. sp. indet. C*
CollNC60	North Carolina	?	?	*C. sp. indet. C*
GJS0852	Panama	*T. cacao*	Leaf	*C. sp. indet. D*
7767	Panama	*T. cacao*	Leaf	*C. sp. indet. D*
GJS0857	Panama	*P. americana*	Fruit	ND
Coll38	Florida	*M. indica*	Fruit	*C. sp. indet. A*
GJS08144	Panama	*M. indica*	Fruit	*C. sp. indet. A*
GJS08147	Panama	*M. indica*	Fruit	*C. sp. indet. A*
Coll1126*	New Jersey	*V. macrocarpon*	Stem	*C. ignotum*
Coll996	New Jersey	*R. virginica*	Stem	*C. ignotum*
CollP1	North Carolina	*V. corymbosum*	Leaf	*C. ignotum*
1087*	Panama	*T. cacao*	Leaf	*C. ignotum*
3589*	Panama	*T. cacao*	Leaf	*C. ignotum*
3679*	Panama	*T. cacao*	Leaf	*C. ignotum*
7574	Panama	*T. cacao*	Leaf	*C. ignotum*
E886	Panama	*T. cacao*	Leaf	*C. ignotum*
Coll878	New Jersey	*Nuphar lutea*	Leaf	*C. nupharicola*
cbs470	Washington	*N. lutea ssp. polysepala*	Leaf	*C. nupharicola*
Coll922	Washington	*N. lutea*	Leaf	*C. nupharicola*
cbs472	Rhode Island	*Nymphaea odorata*	Leaf	*C. nupharicola*
Coll919	Puerto Rico	*T. taxifolia*	Leaf seta	*C. ignotum2*
8395	Panama	*T. cacao*	Leaf	*C. ignotum2*
E183	Panama	*Genipa americana*	Leaf	*C. ignotum2*
Coll940	Oklahoma	*Juglans nigra*	Leaf	ND

Table 1: continued

Coll887	W. Virginia	*V. oxycoccos*	Fruit	*C. sp. indet. E*
Coll920	New Jersey	*Chamaecyparis thyoides*	Stem	*C. sp. indet. E*
Coll20*	Mexico	*Solanum sp.*	stem	*C.gloeosporioides*
IMI356878	Italy	*Citrus sinsensis*	Fruit	*C.gloeosporioides*
Coll914	Mexico	*Solanum sp.*	Leaf	*C.gloeosporioides*
GJS0848	Panama	*T. cacao*	Fruit	*C. theobromicola*
GJS0843	Panama	*T. cacao*	Leaf	*C. theobromicola*
GJS0850	Panama	*T. cacao*	Leaf	*C. theobromicola*
3386	Panama	*T. cacao*	Leaf	*C. sp. indet. F*
4801	Panama	*T. cacao*	Leaf	*C. sp. indet. F*
4766	Panama	*T. cacao*	Leaf	*C. sp. indet. F*

*, Endophyte isolates; ND, not designated; Type strains are in bold-faced type

Table 2: Tree statistics. %PIC: Percent of parsimony informative characters. #MPTs: number of maximum parsimony trees. MPT length: length of the maximum parsimony trees. CI: ensemble consistency index. RI: ensemble retention index. ML –ln L: ML score of best tree calculated in RAxML under the GTRGAMMA model option. BI –ln L: ML score of best tree from 4 runs of 10,000,000 generations inferred in MrBayes calculated in RAxML with GTRGAMMA model.

Locus	# taxa	# chars	% PIC	#MPTs	MPT length	CI	RI	ML -ln L	BI -ln L
nrITS	84	570	13.3	28	93	84	95	-1324.92	NA
tub2	84	900	31.2	12	403	87	97	-3389.23	NA
apn2	82	756	27.3	14	339	79	96	-3024.04	NA
apn2/MAT12IGS	82	893	52.1	1	723	83	97	-4531.11	NA
Combined	84	3122	33	1	1585	82	96	-12981.97	-12981.97

Table 3: Genealogical sorting indices of 1000 bootstrap trees. Calculated using genealogical Sorting R package with 1,000 permutations.

	combined		nrITS		Tub2		APN2		APN2MAT	
Taxon Assignment	gsiT	p	gsiT	p	gsiT	p	gsiT	p	gsiT	p
kahawae	0.740	***	0.155	***	0.073	NS	0.110	*	0.572	***
fructivorum	1.000	***	0.370	***	0.466	***	0.703	***	1.000	***
rhexiae	1.000	***	0.381	***	0.179	**	1.000	***	1.000	***
temperatum	1.000	**	0.404	**	1.000	***	0.494	**	0.901	***
tropicale	1.000	***	0.149	***	1.000	***	0.810	***	1.000	***
CspindetA	1.000	***	1.000	***	1.000	**	1.000	***	1.000	***
melanocaulon	1.000	**	0.022	NS	0.187	*	1.000	***	0.241	*
CspindetB	0.952	***	0.658	***	0.220	**	0.658	***	0.487	**
CspindetC	1.000	***	0.022	NS	0.114	*	0.986	***	1.000	***
CspindetD	1.000	**	0.022	NS	0.679	***	1.000	**	0.812	***
ignotum	1.000	***	0.221	***	0.751	***	0.757	***	1.000	***
nupharicola	1.000	***	0.198	***	1.000	***	1.000	***	1.000	***
ignotum2	1.000	***	0.064	*	1.000	***	1.000	***	0.447	***
CspindetE	1.000	**	0.024	NS	1.000	***	1.000	***	0.786	**
gloeosporioides	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***
theobromicola	1.000	***	1.000	***	1.000	***	1.000	***	1.000	***

***, p-value below 0.001; **, p-value below 0.01; *, p-value below 0.05; NS, p-value greater than 0.05

Table 4: Taxon assignments for the genealogical sorting index analysis.

Strain	Assignment	Strain	Assignment
CBS124	fructivorum	3679	ignotum
Coll1002	fructivorum	7574	ignotum
Coll1004	fructivorum	E886	ignotum
Coll1062	fructivorum	Coll919	ignotum2
Coll1081	fructivorum	8395	ignotum2
Coll1092	fructivorum	E183	ignotum2
Coll116	fructivorum	GJS08211	kahawae
Coll1164	fructivorum	GJS08214	kahawae
Coll1178	fructivorum	GJS08216	kahawae
Coll1190	fructivorum	IMI319418	kahawae
Coll1216	fructivorum	Coll38	CspindetA
Coll1414	fructivorum	GJS08144	CspindetA
Coll21	fructivorum	GJS08147	CspindetA
Coll445	fructivorum	Coll878	nupharicola
Coll864	fructivorum	Coll922	nupharicola
Coll873	fructivorum	CBS470	nupharicola
Coll886	fructivorum	CBS472	nupharicola
Coll1103	temperatum	Coll1026	rhexiae
Coll883	temperatum	Coll1034	rhexiae
Coll126	melanocaulon	Coll1041	rhexiae
Coll131	melanocaulon	Coll1306	rhexiae
Coll887	CspindetE	Coll877	rhexiae
Coll920	CspindetE	Coll952	rhexiae

Table 4: continued

Coll940	Coll940	Coll1038	rhexiae
Coll6	CspindetB	Coll1470	rhexiae
CollNC67	CspindetB	GJS0843	theobromicola
1092	CspindetB	GJS0848	theobromicola
Coll54	CspindetC	GJS0850	theobromicola
CollNC60	CspindetC	Coll11	tropicale
GJS0852	CspindetD	Coll918	tropicale
7767	CspindetD	GJS0842	tropicale
GJS0857	GJS0857	4861	tropicale
Coll20	gloeosporioides	5101	tropicale
Coll914	gloeosporioides	7423	tropicale
IMI356878	gloeosporioides	8401	tropicale
Coll1126	ignotum	E1164	tropicale
Coll996	ignotum	E2303	tropicale
CollP1	ignotum	E406	tropicale
1087	ignotum	Q633	tropicale
3589	ignotum		

Table 5: Genbank accession numbers for sequence data generated in phylogenetic study of *C. gloeosporioides s.l.*

Strain	*nrITS*	*Tub2*	*APN2*	*APN2MAT*
CBS124	JX145125	JX145176	JX145228	JX145278
CBS470	JX145173	JX145225	JX145275	JX145319
CBS472	JX145174	JX145226	JX145276	JX145320
Coll1002	JX145126	JX145177	JX145229	JX145279
Coll1004	JX145127	JX145178	JX145230	JX145280
Coll1026	JX145128	JX145179	JX145231	JX145290
Coll1034	JX145129	JX145180	JX145232	JX145291
Coll1038	JX145168	JX145220	JX145270	JX145295
Coll1041	JX145130	JX145181	JX145233	JX145292
Coll1062	JX145131	JX145182	JX145234	JX145281
Coll1081	JX145132	JX145183	JX145235	JX145282
Coll1092	JX145133	JX145184	JX145236	JX145283
Coll11	JX145134	JX145185	JX145237	JX145306
Coll1103	JX145135	JX145186	JX145238	JX145297
Coll1126	JX145136	JX145187	JX145239	JX145315
Coll116	JX145137	JX145188	JX145240	JX145299
Coll1164	JX145138	JX145189	JX145241	JX145284
Coll1178	JX145139	JX145190	JX145242	JX145285
Coll1190	JX145140	JX145191	JX145243	JX145286
Coll1216	JX145141	JX145192	JX145244	JX145287
Coll126	JX145142	JX145193	JX145245	JX145309

Table 5: continued

Coll1306	JX145143	JX145194	JX145246	JX145293
Coll131	JX145144	JX145195	JX145247	JX145313
Coll1414	JX145145	JX145196	JX145248	JX145300
Coll1470	JX145146	JX145197	JX145249	JX145304
Coll20	JX145148	JX145199	JX145251	JX145326
Coll21	JX145149	JX145200	JX145252	JX145301
Coll38	JX145150	JX145201	JX145253	JX145308
Coll445	JX145151	JX145202	JX145254	JX145288
Coll54	JX145152	JX145203	JX145255	JX145311
Coll57	JX145124	JX145204	-	-
Coll6	JX145153	JX145205	JX145256	JX145314
Coll60	JX145154	JX145206	-	-
Coll864	JX145155	JX145207	JX145257	JX145305
Coll873	JX145156	JX145208	JX145258	JX145289
Coll877	JX145157	JX145209	JX145259	JX145302
Coll878	JX145158	JX145210	JX145260	JX145323
Coll883	JX145159	JX145211	JX145261	JX145298
Coll886	JX145160	JX145212	JX145262	JX145303
Coll887	JX145147	JX145198	JX145250	JX145321
Coll914	JX145161	JX145213	JX145263	JX145327
Coll918	JX145162	JX145214	JX145264	JX145307
Coll919	JX145163	JX145215	JX145265	JX145317
Coll920	JX145164	JX145216	JX145266	JX145322
Coll922	JX145172	JX145224	JX145274	JX145318

Table 5: continued

Coll940	JX145165	JX145217	JX145267	JX145325
Coll952	JX145166	JX145218	JX145268	JX145294
Coll996	JX145167	JX145219	JX145269	JX145324
CollNC60	JX145169	JX145221	JX145271	JX145312
CollNC67	JX145170	JX145222	JX145272	JX145310
CollP1	JX145171	JX145223	JX145273	JX145316
IMI319418	JX145175	JX145227	JX145277	JX145296

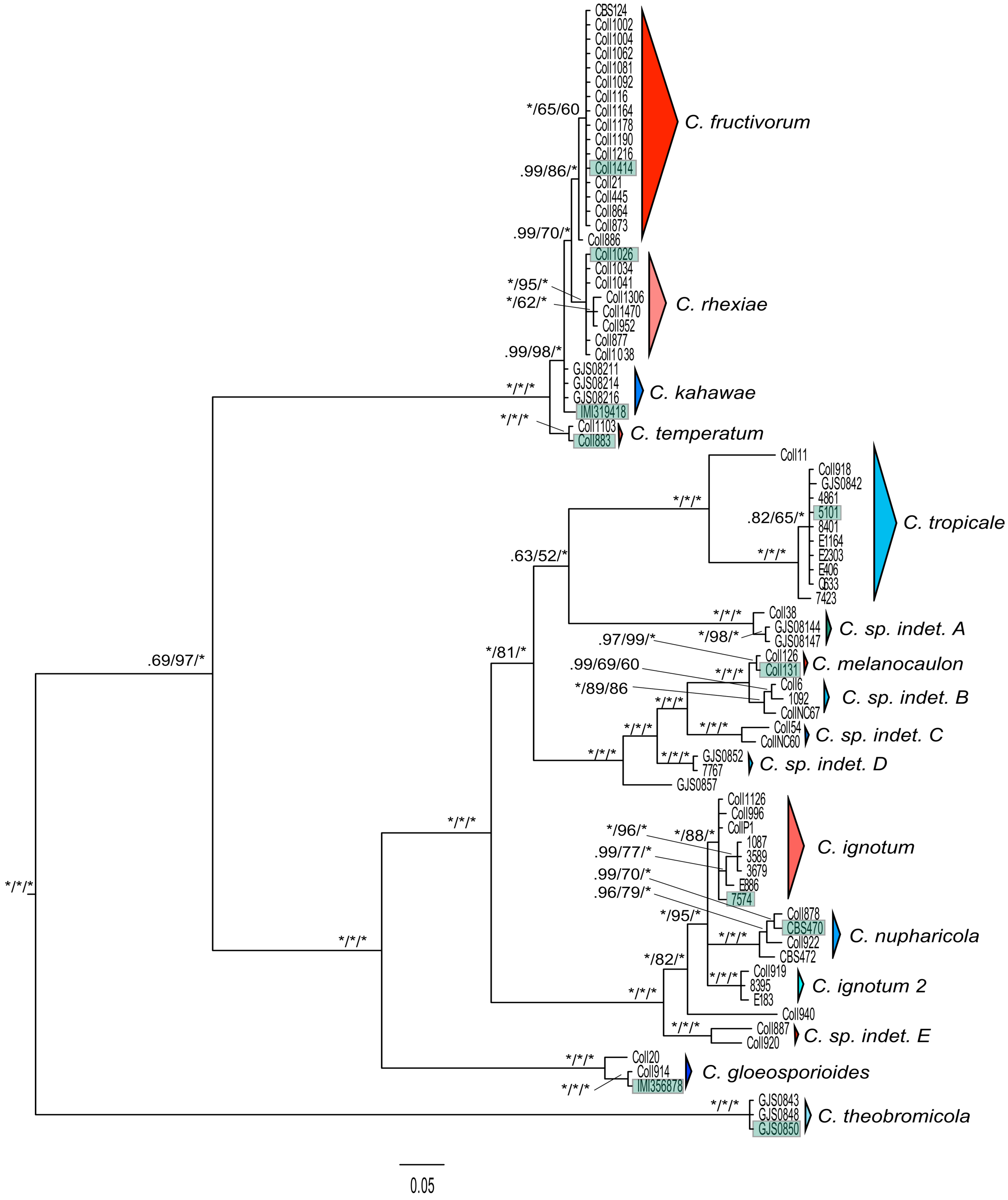


Figure 1: legend on next page

Figure 1: Bayesian majority rule consensus tree with support values (PP/ML-BS/MP-BS) for the combined four-gene analysis. Outgroup taxa (*Coll57, Coll60, 3386, 4801, 4766*) have been trimmed from the phylogram. Support values of 1.0 or 100% are represented with "*". Terminals enclosed in shaded boxes represent sequence data from ex-type strains.

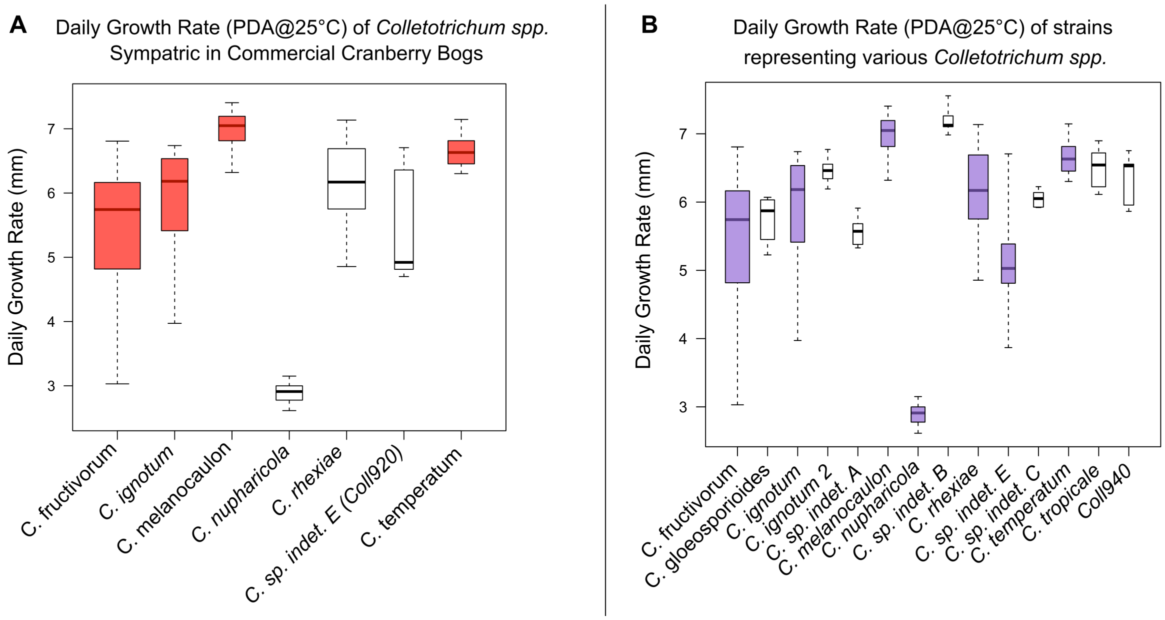


Figure 2: Box plots of daily growth rates representing various species in the *Colletotrichum gloeosporioides* species complex. A) Box plot of species isolated from *Vaccinium macrocarpon* and other sympatric host species in commercial cranberry beds. Red boxes represent species where strains have been isolated from *V. macrocarpon* in commercial beds only. Non-colored boxes represent species isolated only from sympatric host species. B) Box plot of several strains representing species within the *C. gloeosporioides* species complex. Purple boxes represent species that contain strains isolated from host plants found in wild and commercial cranberry bogs in North America. Non-colored boxes represent species not found in these habitats. Black bars represent the median and the whiskers cover the data extremes.

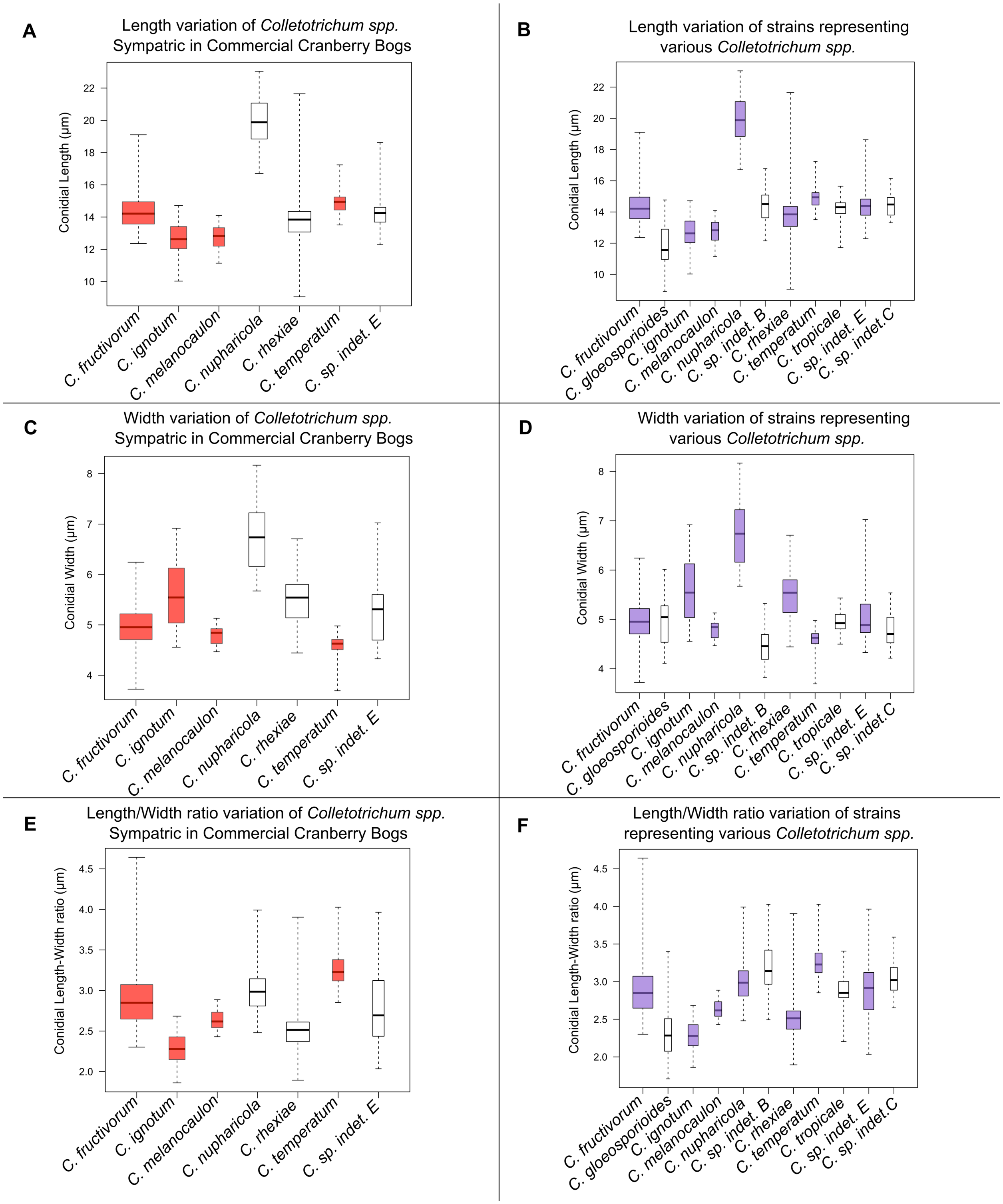


Figure 3: legend on next page

Figure 3: Box plots of conidial sizes representing various species in the *Colletotrichum gloeosporioides* species complex. A, C, E) Box plot of species isolated from *Vaccinium macrocarpon* and other sympatric host species in commercial cranberry beds. Red boxes represent species where strains have been isolated from *V. macrocarpon* in commercial beds only. Non-colored boxes represent species isolated only from sympatric host species. B, D, F) Box plot of several strains representing species within the *C. gloeosporioides* species complex. Purple boxes represent species that contain strains isolated from host plants found in wild and commercial cranberry bogs in North America. Non-colored boxes represent species not found in these habitats. Black bars represent the median and the whiskers cover the data extremes.

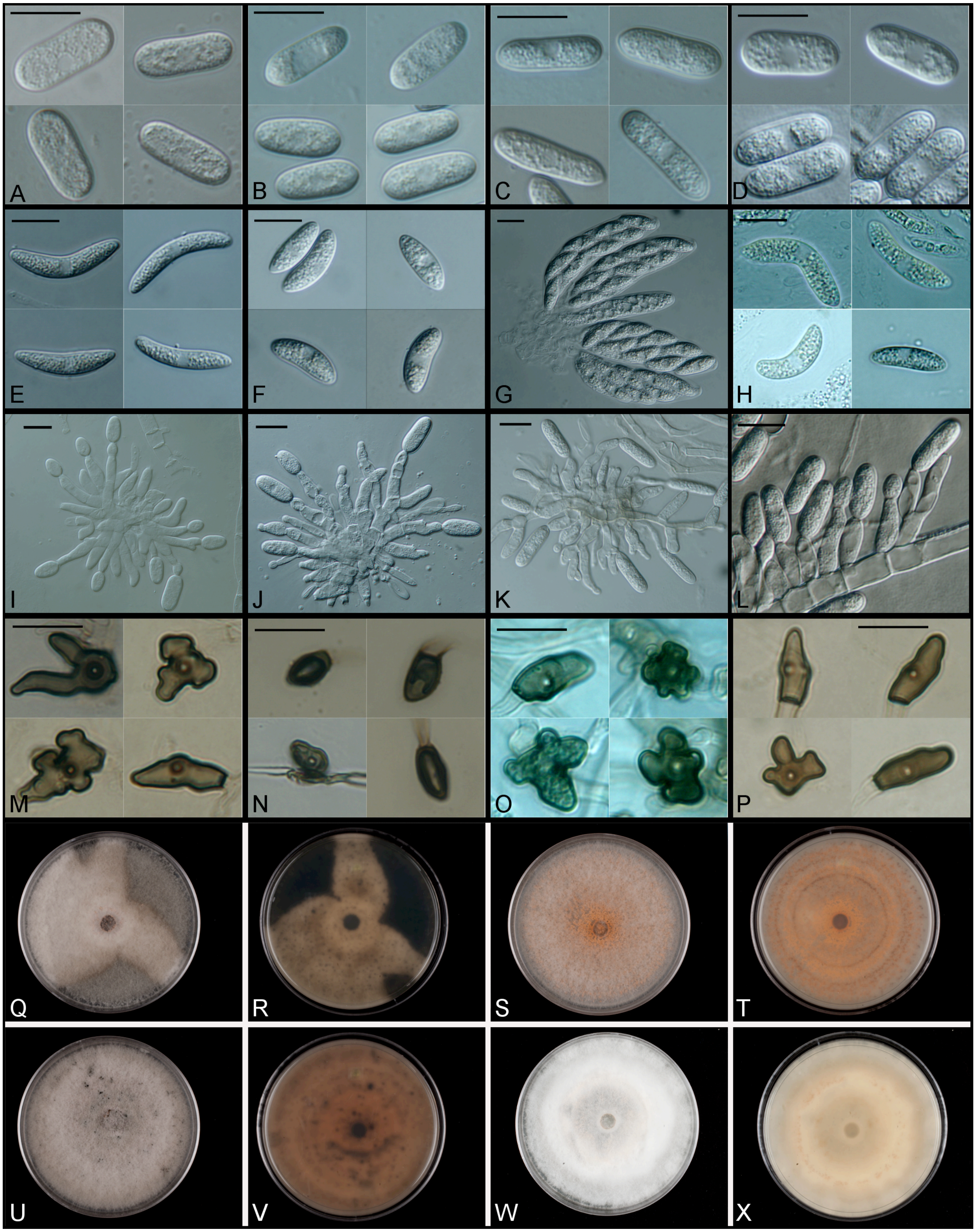


Figure 4: legend on next page

Figure 4: New and epitypified species of *Colletotrichum* from North American cranberry bogs. A-D. Conidia. A) *C. rhexiae* on CMA. B) *C. melanocaulon* on CMA. C) *C. temperatum* on CMA. D) *C. fructivorum* on CMA. E-H. Asci and ascospores. E) *C. rhexiae* on PDA. F) *C. temperatum* on CMA. G) *C. temperatum* on CMA. H) *C. fructivorum* on V8 agar. I-L. Phialides. I) *C. rhexiae* on CMA. J) *C. melanocaulon* on SNA. K) *C. temperatum* on CMA. L) *C. fructivorum* on CMA. M-P. Hyphal appressoria. M) *C. rhexiae.* N) *C. melanocaulon.* O) *C. temperatum*. P) *C. fructivorum*. Q-X. Colonies on PDA after 10 days at 25°C in the dark (100 X 15 mm plates). Q) *C. rhexiae* obverse. R) *C. rhexiae* reverse. S) *C. melanocaulon* obverse. T) *C. melanocaulon* reverse. U) *C. temperatum* obverse V) *C. temperatum* reverse. W) *C. fructivorum* obverse. X) *C. fructivorum* reverse. All scale bars = 10 µm

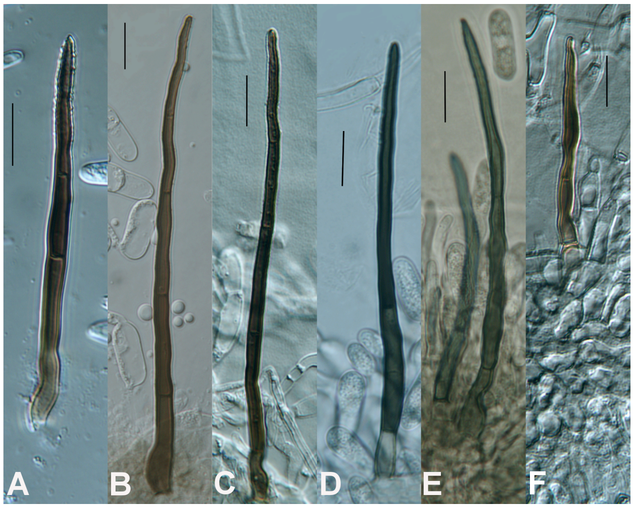


Figure 5: Setae of *Colletotrichum melanocaulon, C. rhexiae, C. fructivorum, C. temperatum* and *C. sp. indet. C*. A) *Colletotrichum melanocaulon* on SNA. B) *Colletotrichum rhexiae* on CMA. C) *Colletotrichum rhexiae* on SNA. D) *Colletotrichum fructivorum* on CMA. E) *Colletotrichum temperatum* on CMA. F) *Colletotrichum sp. indet. C* on SNA. Scale bars = 10 µm.

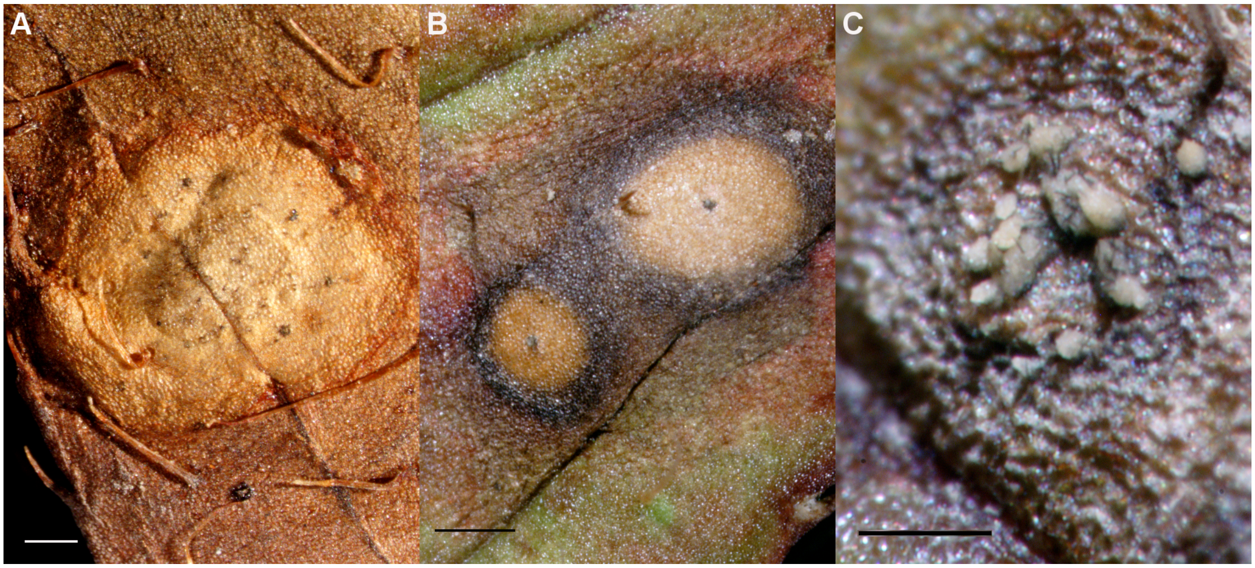


Figure 6: Acervuli of *Colletotrichum rhexiae* on *Rhexia virginica.* A) Acervulus of the holotype, *A. Commons 2534*, on leaf. scale bar = 500µm B) Acervuli on leaves of *R. virginica.* scale bar = 250µm C) Acervulus with setae and conidial masses from which *V. Doyle Coll1470* was isolated. scale bar = 250µm

CHAPTER 3

MICROSATELLITE MARKERS FOR MOLECULAR ECOLOGY STUDIES OF *COLLETOTRICHUM GLOEOSPORIOIDES S.L.*

ABSTRACT

Colletotrichum gloeosporioides s.l. is a diverse complex of species that are globally distributed pathogens and endophytes of a wide range of host plant species. The ecology of these species, however, is poorly understood in large part because of the broad morphological variation within the complex. The low levels of genetic variation in the markers that have been developed to date prohibit fine-scale analyses of population structure and limit the potential for inferring migration patterns and reproductive mechanisms. We utilized several strategies to develop polymorphic microsatellite markers for *C. gloeosporioides s.l.* to provide tools for molecular ecology and population genetic studies of the species complex. Microsatellite primer sets were developed from EST and genomic sequence data deposited in Genbank, membrane and bead-enriched genomic libraries, and 454 sequencing. We focused on a principal lineage within the complex that includes two recently described species, *C. fructivorum* and C. *temperatum*, as well as *C. rhexiae* and *C. kahawae* and report 42 primer sets that are variable within or among these species. We also tested these markers for cross-amplification, and included markers that were not variable among the aforementioned species. We compared the efficiency of the methods employed for microsatellite development and found that 454 sequencing of a microsatellite-enriched genomic library offers the best option for polymorphic marker development.

INTRODUCTION

Colletotrichum Corda (teleomorph: *Glomerella* von Schrenk & Spaulding) is a genus primarily composed of plant pathogens and endophytes that are globally distributed. *Colletotrichum gloeosporioides s.l.* is an aggregate of species that has, until recently, eluded reliable species delimitation due to broad morphological variability and homoplasy. Recent molecular phylogenetic studies, however, have provided the tools necessary to develop reliable taxonomic assessments and thus delve deeper into the ecology and natural history of particular species (Crouch *et al.*, 2006; Du *et al.*, 2005; Rojas *et al.*, 2010; Silva *et al.*, 2012b). A recent study of *C. gloeosporioides s.l.* in North American cranberry bogs revealed two principal lineages within the species complex, one of which contains an economically important and broadly distributed species, *Colletotrichum fructivorum*, pathogenic on cranberry, *Vaccinium macrocarpon* (Chapter 2). The aim of this study was to develop variable microsatellite markers useful for inferring population parameters at a fine-scale in *C. fructivorum*. In addition, we aimed to develop markers that would cross-amplify in the other species of this lineage (*C. rhexiae, C. kahawae,* and *C. temperatum*) in the hope that these would allow for an independent test of species boundaries among these lineages. Furthermore, we tested these markers for amplification and fragment length variation among several species in the diverse lineage that includes *C. gloeosporioides sensu stricto*, including the following taxa: *C. gloeosporioides s.s., C. ignotum, C. nupharicola, C. tropicale* and *C. melanocaulon*. The first approach for marker development involved screening sequence data of "*Colletotrichum gloeosporioides*" available from Genbank for simple sequence repeat motifs. When preliminary evaluation of the markers developed from Genbank data indicated low variability, we developed additional markers using a series of approaches starting with membrane-hybridization genomic enrichment, followed by magnetic bead-hybridization genomic enrichment, and finally 454 pyrosequencing of the microsatellite-enriched genome resulting from the magnetic bead-hybridization. Here we describe the approaches we used for microsatellite marker development and present the results of testing these markers for variability with a focus on *C. fructivorum*.

MICROSATELLITE MARKER DEVELOPMENT

Screening Expressed Sequence Tags for SSR motifs

Expressed sequence tags and gene sequence data of "*Colletotrichum gloeosporioides"* were downloaded from Genbank and screened for microsatellite motifs with read2Marker (Fukuoka *et al.*, 2005) and flanking primers designed with Primer3 (Rozen, Skaletsky, 2000).

DNA extraction, digestion and linker ligation

Microsatellite-enriched genomic libraries were constructed from genomic DNA of a single strain of *Colletotrichum fructivorum* (*Coll21*) collected in New Jersey, USA. Genomic DNA was extracted from a single conidium isolate using the DNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA, USA). Genomic DNA was quantified on a NanoDrop 1000 spectrophotometer (Thermo Scientific). Between 100 and 200 ng of genomic DNA was ethanol precipitated and digested with Sau3AI (New England Biolabs, Ipswich, MA, USA) in 100 µL reaction for 18 hours at 37 ºC followed by 20 minutes at 65 ºC. Restriction digested DNA was dephosphorylated with 1 U shrimp alkaline phosphatase (Promega) at 37 ºC for 1 hour followed by deactivation at 65 ºC for 15 minutes. Restriction fragments were ligated to Sau3AI linkers (2 µM linkers; 15 units Roche T4 ligase), Linker A (5' GCG GTA CCC GGG AAG CTT GG 3') and Linker B (5' pGAT CCC AAG CTT CCC GGG TAC CGC 3') (Carleton *et al.*, 2002; Moraga Amador *et al.*, 1998), in a 115 µL reaction for 18 hours at 16 ºC.

Membrane enrichment

The nylon hybridization membranes (Roche Nylon Cat# 11417240) were prepared following the protocol of Edwards et al. (1996) with a slight modification, substituting the *Mlu*I

21-mer with linker A. The nylon membrane was probed with the following oligomers: $GA_{(15)}$, $GT_{(15)}$, $AT_{(15)}$, $GC_{(15)}$, $CAA_{(10)}$, $CATA_{(10)}$, $ATT_{(10)}$, $GATA_{(10)}$, $GCC_{(10)}$, and $ATAG_{(10)}$. Primers were diluted to 100μM in EB (10mM Tris pH 8.0). Membranes were spotted with 20μL spotting mixture (390μL 3x standard saline citrate [SSC], 10μL primer mix (1μL each)), in order to yield approximately 50 ng of oligonucleotides per membrane. The nylon membranes were air-dried for ~2.5hrs, fixed to the membrane under a UV transilluminator [2000 x 100MJ/cm3] for 30 seconds, and rinsed multiple times in 10mL of hybridization buffer solution (50% Formimide, 3x SSC, 25mM Na-phosphate, 0.5% SDS) in order to wash off any weakly bound oligonucleotides. The membranes were washed at 45°C while constantly rotating at medium speed (Fisher Scientific Isotemp Hybridization Incubator); changing the wash buffer every 24 hours for 4 days, for a total of 5 washes.

Ligated DNA fragments were PCR amplified in 100 μL reactions containing 56.5 μL of autoclaved ion-exchanged water, 8 μL of dNTP (2.5 mM stock of each dNTP), 10 μL of bovine serum albumin (BSA; 0.25 μg μL−1 stock), 10 μL of buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 0.25 % (w/v) cresol red], 10 μL of Linker A, 0.5 μL of *Taq* polymerase and 5 μL of ligated DNA for 20 cycles at 95°C for 30 seconds; 59°C for 30 seconds; 72°C for 2 min; and a final extension at 72°C for 10 minutes prior to membrane hybridization.

The DNA fragments were hybridized to the probed membrane in two separate 500 μL reactions; 1) with 110 μL of the original ligation and 2) with 95 μL of the preamplified ligation (2 μL Linker A, 250 μL formamide, 75 μL 20x SSC, 5.28 μL monosodium phosphate, 7.12 μL disodium phosphate, 20 μL 10% SDS, with 25.6 μL or 40.6 μL of H2O, respectively). The hybridization reactions were heated to 100°C for 5 min followed by incubation for 24 hours at

37°C. The membranes were then rinsed 20 times in 450 µL 0.5x SSC heated to 65°C. The hybridized DNA was eluted in 200 µL autoclaved ion-exchanged water at 95°C for 5 min.

In order to increase the concentration of DNA, the eluted fragments were amplified in a 25 µL PCR reaction containing 10.4 µL autoclaved ion-exchanged water, 2 µL of dNTP (2.5 mM stock of each dNTP), 10 µL bovine serum albumin (BSA; 0.25 µg µL−1 stock), 2.5 µL buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 0.25 % (w/v) cresol red], 2.5 µL of Linker A, 0.125 µL of *Taq* polymerase and 5 µL of the eluted fragments with same cycling parameters as previously mentioned. The resulting PCR product was then run through the hybridization protocol an additional time, eluted, reamplified, and ethanol precipitated prior to cloning.

Dynabead enrichment

Dynabead enrichment followed the protocol of Glenn and Schable (2005) using the following biotin labeled oligonucleotides: $GTT_{(8)}$, $AAG_{(8)}$, $CA_{(10)}$, and $AT_{(12)}$. Linker ligated DNA fragments were hybridized to the biotin labeled oligonucletoides in a 50µL reaction containing 25 µL 2X hybridization solution (1.8 M NaCl, 0.18 M sodium citrate, 0.2% SDS), 10 µL pooled biotin-labeled oligonucleotides (1 µM) and 10 µL linker ligated DNA and 5 µL of autoclaved ion-exchanged water. The hybridization reaction was run on an Eppendorf vapo.Protect Mastercycler pro S with the following cycling parameters: denaturation at 95°C for 5 mintues, ramping to 70°C with a step down of 0.2°C every 5 seconds for 99 cycles down to 50.2°C, holding at 50°C for 10 minutes and rapid ramping down to 15°C. Subsequently, the DNA + oligonucleotide probe mixture was hybridized to magnetic streptavidin beads (Invitrogen Cat# 112-05D). The streptavidin beads were prepped by rinsing twice in TE [10 mM Tris pH8, 2 mM EDTA] and 1x Hyb solution (0.9 M NaCl, 0.09 M sodium citrate, 0.1% SDS) and resuspending in 150 µL 1x Hyb solution before adding the DNA + oligonucleotide probe mixture and incubating at room temperature while slowly rotating for 1 hour. The magnetic beads were then captured with a magnetic particle concentrator (Invitrogen) and rinsed twice in 400 µL of 2x Wash solution (0.3 M NaCl, 0.03 sodium citrate, 0.1% SDS) and 2 times in 1x Wash solution (0.15 M NaCl, 0.015 M sodium citrate, 0.1% SDS). The beads were rinsed two final times in 1x Wash solution, heating the solution to 47°C, resuspended in 200 µL TLE (10 mM Tris pH 8, 0.2 mM EDTA) and incubated at 95°C for 5 minutes. The beads were captured a final time and the supernatant was quickly pipetted to a new tube, ethanol precipitated, and resuspended in 25 µL TLE.

In order to increase the amount of microsatellite-enriched DNA, fragments were recovered via PCR in 25 µL reactions containing 13.0 µL of autoclaved ion-exchanged water, 1.5 µL of dNTP mixture (2.5 mM stock of each dNTP), 2.5 µL of bovine serum albumin (BSA; 0.25 µg µL−1 stock), 2.5 µL of buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 0.25 % (w/v) cresol red], 2.0 µL $MgCl_2$ [25 mM], 1.3 µL of Linker A, 0.2 µL of *Taq* polymerase and 2 µL of eluted DNA fragments. The PCR cycling parameters were as follows: 95°C for 2 min, followed by 25 cycles at 95°C for 20 s, 60°C for 20 s, 72°C for 1.5 min, followed by a final extension at 72°C for 30 min.

Cloning and Sequencing

The microsatellite-enriched products from the membrane and bead enrichments were cloned with the pCR 2.1 vector (Invitrogen) following the supplier's recommendations. White colonies were inoculated into 100 μL of LB broth (Sigma) amended with kanamycin (50 μg/mL) and incubated at 37°C for 1 hour. Colonies were PCR-screened using M13 forward and reverse primers (95°C 10 min; 35 cycles: 95°C 30 s, 55°C 30s, 72°C 1 min; 72°C 10 min) in a 15 μL reaction containing the following reagents: 1.5 μL buffer, 0.2 mM dNTPs, 0.025 μg/μL BSA, 3 μL Q solution (Qiagen), 0.4 μM M13 primers, 0.1 units Taq (Qiagen), and 1 μL bacterial culture. PCR products were run on a 1% agarose gel alongside a 1 Kb ladder (New England Biolabs) and products containing inserts in the 500-800 bp range were sequenced. PCR products were cleaned with ExoSAP-IT and sequenced with M13 primers on an Applied Biosystems 3730 sequencer (High-Throughput Genomics Unit, University of Washington, Seattle, WA).

Sequence assembly, microsatellite motif detection and primer development

Sequences were edited and assembled in Sequencher 4.9 (Gene Codes). The assembled contigs were screened for microsatellite motifs using read2Marker (Fukuoka *et al.*, 2005) and WebSat (Martins *et al.*, 2009) and flanking primers designed with Primer3 (Rozen, Skaletsky, 2000). In order to compare the efficiency of the membrane enrichment versus the bead enrichment, read2Marker was used to screen contigs for microsatellite motifs with a core repeat sequence greater than 12 bp in length. Plots of the frequency distribution of the maximum continuous repeat length (perfect SSR) for both enrichment methods was done to compare the stringency of each method to select longer repeat motifs. Histograms were produced using R version 2.13.0.

Pyrosequencing of an enriched genomic library

The streptavidin bead-enriched genomic library was sent for parallel pyrosequencing to the United States Geological Survey lab in Leetown, West Virginia. Fragments from the SSR-enriched library were tagged with a barcode sequence and sequenced in a 1/16th run pooled with samples unrelated to this project. The 454 sequencing was done on the Roche GS Junior System with GS Junior Titanium Series reagents. Sequences were assembled with CLC Bio Workbench software. Both mapped reads and unmapped reads were screened for microsatellite motifs in read2Marker (Fukuoka *et al.*, 2005).

Polymorphism Screening

In order to find microsatellite markers polymorphic within *C. fructivorum,* 28 strains were selected from 5 regional populations (Delaware: *n*=4; Massachusetts: *n*=4; British Columbia: *n*=4; Wisconsin: *n*=4; New Jersey: *n*=11) isolated from cultivated cranberry (*Vaccinium macrocarpon*) populations and one strain isolated from *Vaccinium oxycoccos* in Pennsylvania. Strains representing other species from the *Colletotrichum gloeosporioides* species complex were also screened for cross-amplification (*C. temperatum, C. melanocaulon; C. gloeosporioides s.s.; C. ignotum; C. ignotum 2;* and *C. theobromicola*) and polymorphism (*C. rhexiae*: *n*=2; C. *nupharicola*: *n*=2; *C. kahawae*: *n*=2; *C. mangivora*: *n*=2; and *C. tropicale*: *n*=2). In addition, strains representing undescribed species (*C. sp. indet. E*, *C. sp. indet C*) and an unnamed singleton, *Coll940,* were screened for cross-amplification.

Isolates were grown on potato dextrose broth (Difco) for 5-7 days before mycelium was harvested, blotted dry on sterile paper towels and dehydrated for 6-10 hours in a vacuum centrifuge on low heat. Approximately 20 mg of dried tissue was used for DNA extraction using the DNeasy Plant MiniKit (Qiagen Inc., Valencia, CA, USA) or a standard phenol-chloroform extraction method after tissue homogenization in the FastPrep FP120 (MP Biomedicals, LLC., Solon, OH, USA). Primers selected for polymorphism screening were synthesized by Bioneer (Alameda, CA). PCR amplification reactions were performed in 10 uL reactions containing 4.37 µL of autoclaved ion-exchanged water, 0.8 µL of dNTP mixture (2.5 mM stock of each dNTP), 1 μL of bovine serum albumin (BSA; 0.25 μg μL−1 stock), 1 μL of buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 20 μg/mL bovine serum albumin (BSA)], 2 μL betaine (1.2 M stock), 100 nM 5' M13-tailed forward primer, 200 nM reverse primer, 200 nM universal fluorescent-labeled M-13 primer, 0.08 μL of *Taq* polymerase and 0.5 μL of DNA.

PCR reactions were run on an Eppendorf vapo.Protect Mastercycler pro S with the following cycling parameters: initial denaturing for 2.5 min at 94°C, followed by 28-30 cycles of 94°C for 30 s, 55-66°C for 30 s (depending on the optimal annealing temperature of the primers), and 64°C for 45 s, followed by a final extension at 64°C for 20 min. PCR fragments were sized on the Beckman Coulter CEQ 8800 along with the DNA Size Standard Kit 400 (Beckman Coulter PN 608098) following the Frag-3 separation method (denaturation at 90°C for 2 min; injection voltage of 2 kV at 50°C for 30 s; separation voltage of 6.0 kV at 50°C for 35 minutes). In order to accurately estimate fragment sizes, the fragment analysis parameters were adjusted to PA ver. 1 to calibrate for the phosphoramidite linked dye primers.

RESULTS AND DISCUSSION

Read2Marker identified 48 contigs with microsatellite motifs from 257 sequenced clones (18.7%), of which 33 were suitable for primer design, from the membrane-enriched library. The majority of the perfect repeat microsatellites (uninterrupted) were less than ten tandem repeats with only 4 fragments (8.3 % of the fragments with microsatellite motifs) containing ten or more. In contrast, the magnetic bead enrichment yielded 68 microsatellite containing contigs from 155 sequenced clones (43.9%), of which 50 were suitable for primer design. Fragments containing microsatellites with ten or more perfect tandem repeat units were found in 27.9 % of the contigs with microsatellite motifs (Figure 1). In contrast to all other approaches, the EST and gene sequence data obtained from Genbank yielded 104 microsatellite containing fragments from 1701 sequences (6.1%), of which none contained ten or more repeat units.

A total of 30,248,506 nucleotides were sequenced from the bead-enriched genomic library on the 454 sequencing platform from 99,351 reads. The average read length for all the sequenced fragments was 304 base pairs. Sequence contigs were assembled from 93,941 reads to yield 4,308 contigs with an average length of 380 bp, leaving 5,410 unmatched reads. Screening the contigs with read2Marker yielded 918 unique microsatellite containing sequences of which 689 were suitable for primer design with Primer3. The unmatched reads consisted of 525 microsatellite containing fragments of which 322 were suitable for primer design. Fragments containing microsatellites with ten or more perfect tandem repeat units were found in 13.4 % (194) of all sequenced fragments from the 454 run with microsatellite motifs (Figure 1).

Primers from each of the aforementioned search methods were screened for polymorphism within and across several *Colletotrichum* species within the *C. gloeosporioides* species complex. Primers that vary across species are reported in two separate tables. Table 1 represents eight markers that are only variable among *C. fructivorum, C. rhexiae, C. temperatum* and *C. kahawae*, four closely related species, three of which are known to be sympatric. Table 2 represents 34 markers that are variable within *C. fructivorum, C. rhexiae, C. temperatum* or *C. kahawae.* Thirty-one of these are polymorphic in *C. fructivorum* and three are monomorphic. *Colletotrichum kahawae* is the only other species for which any of the reported markers are polymorphic. Eleven of the thirty-four markers are polymorphic in *C. kahawae* where more than a single strain was tested and thirty-two markers (94%) cross-amplify. While no markers are reported as polymorphic for *C. rhexiae* or *C. temperatum*, all markers cross-amplify in *C. rhexiae* and thirty (88%) cross-amplify in *C. temperatum.* Gene diversity is only reported for *C.*

fructivorum because it is the only species in this group for which more than two strains were tested for polymorphism. The gene diversity for the thirty-one markers reported as polymorphic within *C. fructivorum* ranges from 0.0689 to 0.6862 with an average of 0.1996. However, many of the markers are polymorphic due to the inclusion of a single strain, *V. Doyle Coll886,* unique in being the only strain of *C. fructivorum* originating from a wild (rather than cultivated) cranberry bog and from *Vaccinium oxycoccos* (rather than *V. macrocarpon*). Excluding markers where *V. Doyle Coll886* carried one of only two alleles, the number of polymorphic markers is reduced to fourteen and the average gene diversity is increased to 0.3403, with a range from 0.1352 to 0.6862. Table 3 represents twenty-nine markers that are variable across at least two species among *C. theobromicola* and ten species in the principal lineage of the *Colletotrichum gloeosporioides* species complex that includes *C. gloeosporioides sensu stricto*. Allele size has not been determined for all species represented in Table 3; however, cross-amplification has been determined for all primer sets. Six of these markers are also variable within species where allele size was estimated for two strains, including *C. tropicale* (Locus AWFJM)*, C. mangivora* (Locus EG359924, EG360161, and EG359900) and *C. nupharicola* (Locus M93427 and 10A2). All primer sequences and optimal annealing temperatures are listed in Table 4.

The microsatellite markers reported here are the first that have been developed for the *C. gloeosporioides* species complex. These markers are a useful addition to the limited molecular tools that have thus far been developed for the aggregate of species. The low level of genetic variation reported for *C. gloeosporioides s.l.* makes it difficult to make inferences regarding population structure and offers limited power to resolve questions concerning its ecology (Rojas *et al.*, 2010; Silva *et al.*, 2012b). However, at least with respect to *C. fructivorum*, we can now begin to address more detailed questions regarding the influence of biotic and abiotic factors on species evolution and test for evidence of interspecific hybridization. In addition, these markers will provide plant pathologists working in cranberry agricultural systems, where *C. fructivorum* is an important pathogen, with the means to track clonal lineages and to assess the evolutionary potential of the species by investigating genetic diversity within and among populations. Furthermore, the markers reported here may provide an opportunity to unravel the complexity of the mating system of *Colletotrichum* (Chilton *et al.*, 1945; Chilton, Wheeler, 1949a; Chilton, Wheeler, 1949b; Cisar, TeBeest, 1999; Edgerton, 1914; Edgerton *et al.*, 1945; Vaillancourt *et al.*, 2000a; Vaillancourt *et al.*, 2000b; Wheeler, 1950; Wheeler *et al.*, 1959; Wheeler, McGahen, 1952; Wheeler *et al.*, 1948).

REFERENCES

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. *Biophotonics International*, **11**, 36-42.

Adaskaveg JE, Forster H (2000) Occurrence and Management of Anthracnose Epidemics Caused by *Colletotrichum* Species on Tree Fruit Crops in California. In: *Colletotrichum: Host Specificity, Pathology and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 317-336. The American Phytopathogical Society Press, St. Paul, Minnesota.

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

Agostini JP, Timmer LW (1994) Population dynamics and survival of strains of Colletotrichum gloeosporioides on Citrus in Florida. *Phytopathology*, **84**, 420-425.

Alberto F (2009) MsatAllele_1.0: An R package to visualize the binning of microsatellite alleles. *Journal of Heredity*, **1003**, 394-397.

Antonovics J (2006) Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, **97**, 33-37.

Bailey JA, Jeger MJ (1992) Colletotrichum: Biology, Pathology and Control. Wallingford, UK: CAB International.

Bailey JA, O'Connell RJ, Pring RJ, Nash C (1992) Infection strategies of *Colletotrichum* species. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 88-120. CAB International, Surrey, UK.

Bazinet AL, Neel MC, Shaw KL, Cummings MP (In preparation) The genealogical sorting index: software and web site for quantifying exclusivity of lineages.

Bills G, Polishook JD (1992) Recovery of endophytic fungi from *Chamaecyparis thyoides*. *Sydowia*, **44**, 1-12.

Boland GJ (2004) Fungal viruses, hypovirulence, and biological control of *Sclerotinia* species. *Canadian Journal of Plant Pathology*, **26**, 6-18.

Brasier CM (1995) Episodic selection as a force in fungal microevolution, with special reference to clonal speciation and hybrid introgression. *Canadian Journal of Botany*, **73**, S1213-S1221.

Cannon P, Bridge PD (2000) Linking the past, present and future of *Colletotrichum* systematics. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Dov Prusky, Freeman S, Dickman MB), pp. 1-20. American Phytopathological Society, St. Paul, Minnesota.

Carleton KL, Streelman JT, Lee BY*, et al.* (2002) Rapid isolation of CA microsatellites from the tilapia genome. *Animal Genetics*, **33**, 140-144.

Chakraborty S, Fernandes CD, Charchar MJdA, Thomas MR (2002) Pathogenic variation in *Colletotrichum gloeosporioides* infecting

Stylosanthes spp. in a center of diversity in Brazil. *Phytopathology*, **92**, 553-562.

Chilton SJP, Lucas GB, Edgerton CW (1945) Genetics of *Glomerella* III: Crosses with a conidial strain. *American Journal of Botany*, **32**, 549-554.

Chilton SJP, Wheeler HE (1949a) Genetics of *Glomerella* VI: Linkage. *American Journal of Botany*, **36**, 270-273.

Chilton SJP, Wheeler HE (1949b) Genetics of *Glomerella* VII: Mutation and segregation in plus cultures. *American Journal of Botany*, **36**, 717-721.

Cisar CR, TeBeest DO (1999) Mating system of the filamentous ascomycete, *Glomerella cingulata*. *Current Genetics*, **35**, 127-133.

Cole RA, Lynn J (1996) Spatial and temporal population structure of *Brevicoryne brassicae*. *Entomologia Experimentalis et Applicata*, **78**, 121-127.

Colon-Garay J, Rivera-Vargas L, McGovern RJ, Rodriquez RdP (2002) Hypovirulent isolates of *Colletotrichum gloeosporioides* induce resistance to anthracnose in detached mango fruits and seedlings. *J. Agric. Univ. Puerto Rico*, **84**, 56-64.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.

Correll JC, Guerber JC, Wasliwa LA, Sherrill JF, Morelock TE (2000) Inter- and intra-species variation in *Colletotrichum* and mechanisms which affect population structure. In: *Colletotrichum: Host Specificity, Pathology and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 145-179. The American Phytopathogical Society Press, St. Paul, Minnesota.

Couch BC, Fudal I, Lebrun M-H*, et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, **170**, 613-630.

Crouch JA, Clarke BB, Hillman BI (2006) Unraveling the evolutionary relationships among the divergent lineages of *Colletotrichum* causing anthracnose disease in turfgrass and corn. *Phytopathology*, **96**, 46-60.

Crouch JA, Clarke BB, Hillman BI (2009a) What is the value of ITS sequence data in *Colletotrichum* systematics and species diagnosis? A case study using the falcate spored graminicolous *Colletotrichum* group. *Mycologia*, **101**, 648-656.

Crouch JA, Tredway LP, Clarke BP, Hillman BI (2009b) Phylogenetic and population genetic divergence correspond with habitat for the pathogen *Colletotrichum cereale* and allied taxa across diverse grass communities. *Molecular Ecology*, **18**, 123-135.

Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution*, **62**, 2411-2422.

Dean R, Kan JALV, Pretorius ZA*, et al.* (2012) The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology*, **13**, 414-430.

Dettman JR, Anderson JB, Kohn LM (2008) Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus *Neurospora*. *BMC Evolutionary Biology*, **8**, 35.

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

Dettman JR, Sirjusingh C, Kohn LM, Anderson JB (2007) Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature*, **447**, 585-588.

Dilmaghani A, Gladieux P, Gout L*, et al.* (2012) Migration patterns and changes in population biology associated with worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. *Molecular Ecology*, **21**, 2519-2533.

Du M, Schardl CL, Nuckles EM, Vaillancourt LJ (2005) Using mating-type gene sequences for improved phylogenetic resolution of Colletotrichum species complexes. *Mycologia*, **97**, 641-658.

Dutech C, Enjalbert J, Fournier E*, et al.* (2007) Challenges of microsatellite isolation in fungi. *Fungal Genetics and Biology*, **44**, 933-949.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Eck P (1990) *The American Cranberry* Rutgers University Press, New Brunswick, NJ.

Edgerton CW (1914) Plus and minus strains in the Genus Glomerella. *American Journal of Botany*, **1**, 244-254.

Edgerton CW, Chilton SJP, Lucas GB (1945) Genetics of *Glomerella* II: Fertilization between strains. *American Journal of Botany*, **32**, 115-118.

Ellis JB, Everhart BM (1894) New Species of Fungi from Various Localities *Proceedings of the Academy of Natural Sciences of Philadelphia*, **46**, 322-386.

Evanno G, Regnaut S, Goudet J (2005) Detecting th enumber of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587.

Farr DF, Rossman AY, Palm ME, McCray EB (2008) Fungal Databases. Systematic Mycology and Microbiology Laboratory, ARS, USDA.

Farr DF, Rossman AY, Palm ME, McCray EB (2011) Fungal Databases. http://nt.ars-grin.gov/fungaldatabases/ Systematic Mycology and Microbiology Laboratory, ARS, USDA

Foster S, Duke JA (2000) *A Field Guide to Medicinal Plants and Herbs of Eastern and Central North America*, 2nd edn. Houghton Mifflin, New York.

Freeman S, Katan T, Shabi E (1996) Characterization of Colletotrichum gloeosporioides isolates from avocado and almond fruits with molecular and pathogenicity tests. *Applied and Environmental Microbiology*, **62**, 1014-1020.

Fukuoka H, Nunome T, Minamiyama Y*, et al.* (2005) Read2Marker: a data processing tool for microsatellite marker development from a large data set. *Biotechniques*, **39**, 472, 474, 476.

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

Gladieux P, Guerin F, Giraud T*, et al.* (2011) Emergence of fungal pathogens by ecological speciation: importance of the reduced viability of imigrants. *Molecular Ecology*, **20**, 4521-4532.

Gleason H, Cronquist A (1991) *Manual of vascular Plants of Northeastern United States and Adjacent Canada* New York Botanical Garden, New York.

Goloboff PA, Farris JS, Nixon KC (2008) TNT, a free program for phylogenetic analysis. *Cladistics*, **24**, 774-786.

Gomez A, Carvalho GR (2000) Sex, parthenogenesis and genetic structure of rotifers: microsatellite analysis of contemporary and resting egg bank populations. *Molecular Ecology*, **9**, 203-214.

González E, Sutton TB, Correll JC (2006) Clarification of the etiology of Glomerella Leaf Spot and Bitter Rot of apple caused by Colletotrichum spp. based on morphology and genetic, molecular, and pathogenicity tests. *Phytopathology*, **96**, 982-992.

Guozhong L, Cannon PF, Reid A, Simmons CM (2004) Diversity and molecular relationships of endophytic *Colletotrichum* isolates from the Iwokrama Forest Reserve, Guyana. *Mycological Research*, **108**, 53-63.

Halsted B (1889) Some fungus diseases of the cranberry, pp. 1-40. New Jersey Agricultural College Experiment Station.

Hang VTT, Holmes EC, Veasna D*, et al.* (2010) Emergence of the Asian 1 Genotype of Dengue Virus Serotype 2 in Viet Nam: In Vivo Fitness Advantage and Lineage Replacement in South-East Asia. *PLOS Neglected Tropical Diseases*, **4**, e757.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633-1638.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, **290**, 516-518.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754-755.

Jakobsson M, Rosenberg N (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801-1806.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* **11**, 424-429.

Johnson-Cicalese J, Vorsa N, Polashock J (2009) Breeding for fruit rot resistance in *Vaccinium macrocarpon.* In: Proceedings of the IXth International Symposium on *Vaccinium*.

Johnson DA, Carris LM, Rogers JD (1997) Morphological and molecular characterization of *Colletotrichum nymphaeae* and *C. nupharicola* sp. nov. on water-lilies (*Nymphaea* and *Nuphar*). *Mycological Research*, **101**, 641-649.

Johnston PR, Jones D (1997) Relationships among *Colletotrichum* Isolates from fruit-rots assessed using rDNA sequences. *Mycologia*, **89**, 420-430.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Jombart T, Pontier D, Dufour AB (2009) Genetic markers in the playground of multivariate analysis. *Heredity*, **102**.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015-4026.

Jost L (2009) Reply: D vs G'st: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology*, **18**, 2088-2091.

Kareiva P, Watts S, McDonald R, Boucher T (2007) Domesticated Nature: Shaping Landscape and Ecosystems for Human Welfare. *Science*, **316**, 1866-1869.

Katoh K, Kuma K-i, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, **33**, 511-518.

Kaufmann PJ, Weidemann GJ (1996) Isozyme analysis of *Colletotrichum gloeosporioides* from five host genera. *Plant Disease*, **80**, 1289-1293.

Kendrick B (2000) *The Fifth Kingdom* Mycologue Publications, Sidney.

Kirk PM, Cannon PF, David JC, Stalpers JA (2001) Dictionary of the Fungi. CAB International, Surrey.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of poulation differentiation. *Conservation Genetics* **7**, 295-302.

Latunde-Dada AO (2001) Colletotrichum: tales of forcible entry, stealth, transient confinement and breakout. *Molecular Plant Pathology*, **2**, 187-198.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, **11**, 2445-2449.

Legard DE (2000) *Colletotrichum* diseases of Strawberry in Florida. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 292-299. American Phytopathological Society, St. Paul, Minnesota.

Lucas GB (1946) Genetics of *Glomerella* IV: Nuclear phenomena in the ascus. . *American Journal of Botany*, **33**, 802-806.

Lucas GB, Chilton SJP, Edgerton CW (1944) Genetics of *Glomerella* I: Studies on the behavior of certain strains. *American Journal of Botany*, **31**, 233-239.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998a) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238-247.

Luikart G, Cornuet JM, Allendorf FW (1999) Temporal changes in allele frequencies provide estimates of population bottleneck size. *Conservation Biology*, **13**.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998b) Usefullness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**.

Maccheroni WJ, Araújo WL, Azevedo JL (2004) Ambient pH-regulated enzyme secretion in endophytic and pathogenic isolates of the fungal genus *Colletotrichum*. *Scientia Agricola*, **61**, 298-302.

MacKenzie SJ (2005) *Population structure and pathogenicity of Colletotrichum gloeosporioides from strawberry and noncultivated hosts in Florida* Ph.D., University of Florida.

MacKenzie SJ, Seijo TE, Legard DE, Timmer LW, Peres NA (2007) Selection for pathogenicity to strawberry in populations of *Colletotrichum gloeosporioides* from native plants. *Phytopathology*, **97**, 1130-1140.

Martins WS, Lucas DCS, Neves KFS, Bertioli DJ (2009) WebSat - A web software for microsatellite marker development. *Bioinformation*, **3**, 282-283.

Maruyama T, Fuerst PA (1985) Population bottlenecks and nonequilibrium models in population genetics. II. Number of Alleles in a small population that was formed by a recent bottleneck. *Genetics*, **111**, 675-689.

McDonald BA, Linde C (2002a) The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica*, **124**, 163-180.

McDonald BA, Linde CC (2002b) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349-379.

McGahen JW, Wheeler HE (1951) Genetics of *Glomerella* IX: Perithecial development and plasmogamy. . *American Journal of Botany*, **38**, 610-617.

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differntiation measure. *Evolution*, **60**, 2399-2402.

Meirmans PG, Hedrick PW (2011) Assessing population structure: Fst and related measures. *Molecular Ecology Resources*, **11**, 5-18.

Meirmans PG, Tienderen PHV (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity in asexual organsims. *Molecular Ecology Notes*, **4**, 792-794.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, **34**, 457-477

Miller MA, Pfeifer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees, 1-8.

Moraga Amador D, Farmerie B, Holland K*, et al.* (1998) Tools for Developing Microsatellite Markers: Microsatellite Libraries and AFLPs. Workshop Manual. University of Florida, Gainesville, Florida.

Morris CE, Bardin M, Kinkel LL*, et al.* (2009) Expanding the Paradigms of Plant Pathogen Life History and Evolution of Parasitic Fitness beyond Agricultural Boundaries. *PLoS Pathogens*, **5**, 1-7.

Nei M, Mauyama T, Chakraboty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.

Nicholson RL, Moraes WBC (1980) Survival of *Colletotrichum graminicola*: importance of the spore matrix. *Phytopathology*, **70**, 255-261.

Nirenberg H (1976) Untersuchungen u¨ ber die morphologische und biologische Differenzierung in der Fusari-

um-Sektion Liseola. *Mitteilungen aus der Biologischen Bundesanstalt für Land-und Forstwirtschaft*, **169**, 1-117.

Nylander JAA (2004) MrModeltest v2, Program distributed by the author. Evolutionary Biology Centre, Uppsala University. .

O'Connell RJ, Perfect S, Hughes B*, et al.* (2000) Dissecting the cell biology of *Colletotrichum* infection processes. In: *Colletotrichum: Host Specificity, Pahtology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 57-77. American Phytopathological Society, St. Paul, Minnesota.

O'Donnell K, Cigelnik E (1997) Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. *Molecular Phylogenetics and Evolution*, **7**, 103-116.

Olatinwo RO, Hanson EJ, Schilder AMC (2003) A First Assessment of the Cranberry Fruit Rot Complex in Michigan. *Plant Disease*, **87**, 550-556.

Oudemans PV (2008) Personal Communication.

Oudemans PV, Caruso FL, Stretch AW (1998) Cranberry fruit rot in the Northeast: A complex disease. *Plant Disease*, **82**, 1176-1184.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (Glycine max) simple sequence repeats (SSRs) within the genus and other legume genera: implications for transferability of SSRs in plants. *Molecular Biology and Evolution*, **15**, 1275-1287.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetics software for teaching and research. *Molecular Ecology Resources*, **6**, 288-295.

Photita W, Taylor PWJ, Ford R, Hyde KD (2005) Morphological and molecular characterization of Colletotrichum species from herbaceous plants in Thailand. *Fungal Diversity*, **18**, 117-133.

Phoulivong S, Cai L, Chen H*, et al.* (2010) *Colletotrichum gloeosporioides* is not a common pathogen on tropical fruits. *Fungal Diversity*, **44**, 33-43.

Polashock JJ, Caruso FL, Oudemans PV, McManus PS, Crouch JA (2009) The North American cranberry fruit rot fungal community: a systematic overview using morphological and phylogenetic affinities. *Plant Pathology*, **58**, 1116-1127.

Prihastuti H, Cai L, Chen H, McKenzie EHC, Hyde KD (2009) Characterization of *Colletotrichum* species associated with coffee berries in northern Thailand. *Fungal Diversity*, **39**, 89-109.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

R-Development-Core-Team (2011) R: A Language and Environment for Statistical Computing.

Ramakrishnan U, Hadly EA, Mountain JL (2005) Detecting past population bottlenecks using temporal genetic data. *Molecular Ecology*, **14**, 2915-2922.

Rasband WS (1997-2008) ImageJ. National Institutes of Health, Bethesda, Maryland, USA.

Riccioni G, Landi M, Ferrara G*, et al.* (2010) Spatio-temporal population structuring and genetic diversity retention in depleted Atlantic Bluefin tuna of the Mediterranean Sea. *Proceedings of the National Academy of Sciences*, **107**, 2102-2107.

Roa AC, Chavarriaga-Aguirre P, Duque MC*, et al.* (2000) Cross-species amplification of Cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. *American Journal of Botany*, **87**, 1647-1655.

Rodriguez-Saona C, Vorsa N, Singh AP*, et al.* (2011) Tracing the history of plant traits under domestication in cranberries: potential consequences on anti-herbivore defences. *Journal of Experimental Botany*, **62**, 2633-2644.

Rodriguez RJ, Redman RS (2000) *Colletotrichum* as a model system for defining the genetic basis of fungal symbiont life styles. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 114-130. American Phytopathological Society, St. Paul, Minnesota.

Rojas EI, Rehner SA, Samuels GJ*, et al.* (2010) Colletotrichum gloeosporioides s.l. associated with Theobroma cacao and other plants in Panama´: multilocus phylogenies distinguish host-associated pathogens from asymptomatic endophytes. *Mycologia*, **102**, 1318-1338.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Rosenberg N (2004) DISTRUCT: a program for the grpahical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ.

Ryman N, Leimar O (2009) Gst is still a useful measure of genetic differentiation - a comment on Jost's D. *Molecular Ecology*, **18**, 2084-2087.

Seifert KA (2009) Progress towards DNA barcoding of fungi. *Molecular Ecology Resources*, **9**, 83-89.

Shear CL (1907a) Cranberry Disease (ed. USDA). Government Printing Office.

Shear CL (1907b) Cranberry Diseases (ed. USDA). Government Printing Office.

Shear CL (1907c) New species of fungi. *The Bulletin of the Torrey Botanical Club*, **34**, 305-317.

Silva DN, Talhinas P, Cai L*, et al.* (2012a) Host-jump drives rapid and recent ecological speciation of the emergent fungal pathogen *Colletotrichum kahawae*. *Molecular Ecology*, **21**, 2655-2670.

Silva DN, Talhinas P, Várzea V*, et al.* (2012b) Application of the Apn2/MAT locus to improve the systematics of the Colletotrichum gloeosporioides complex: an example from coffee (Coffea spp.). *Mycologia*, **104**, 396-409.

Silvertown J, Servaes C, Biss P, Macleod D (2005) Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity*, **95**, 198-205.

Simmonds JH (1965) A study of the species of *Colletotrichum* causing ripe fruit rots in Queensland. *Queensland Journal of Agricultural and Animal Science*, **25**, 437-459.

Spencer CC, Neigel JE, Leberg PL (2000) Experimental evaluation of the usefullness of microsatellite DNA for detecting demographic bottlenecks. *Molecular Ecology*, **9**, 1517-1528.

Stamatakis A (2006) RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics*, **22**, 2688-2690.

Stamatakis A, Hoover P, Rougemont J (2008) A Fast Bootstrapping Algorithm for the RAxML Web-Servers. *Systematic Biology*, **57**, 758-771.

Stiles CM, Oudemans PV (1997) Cranberry fruit rot pathogens isolated from various cranberry plant tissues over two growing seasons. *Phytopathology*, **87**, S94.

Stiles CM, Oudemans PV (1999) Distribution of cranberry fruit-rotting fungi in New Jersey and evidence for nonspecific host resistance. *Phytopathology*, **89**, 218-225.

Stockwell CA, Mulvey M, Vinyard GL (1996) Translocations and the preservation of allelic diversity. *Conservation Biology*, **10**, 1133-1141.

Stoddart JA (1983) A genotypic diversity estimate. *Journal of Heredity*, **74**, 489-490.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75-100.

Sukumaran J, Holder MT (2010) DendroPy: A Python library for phylogenetic computing. *Bioinformatics*, **26**, 1569-1571.

Sutton BC (1992) The genus *Glomerella* and its anamorph *Colletotrichum*. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 1-26. CAB International, Surrey, UK.

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498-2504.

Taylor JW, Geiser DM, Burt A, Koufopanou V (1999a) The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews*, **12**, 126-146.

Taylor JW, Jacobson DJ, Fisher MC (1999b) The evolution of asexual fungi: Reproduction, speciation, and classification. *Annual Review of Phytopathology*, **37**, 197-245.

Taylor JW, Jacobson DJ, Kroken S*, et al.* (2000) Phylogenetic Species Recognition and Species Concepts in Fungi. *Fungal Genetics and Biology*, **31**, 21-32.

Than PP, Shivas RG, Jeewon R*, et al.* (2008) Epitypification and phylogeny of *Colletotrichum acutatum* J.H. Simmonds. *Fungal Diversity*, **28**, 97-108.

Tu JC (1992) *Colletotrichum lindethumianum* on bean: Population dynamics of the pathogen and breeding for resistance. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 203-224. CAB International, Surrey, UK.

USDA-CSREES (2001) Crop Profile for Cranberries in Massachusetts.

USDA-NASS (2008) Noncitrus Fruits and Nuts 2007 Summary

USDA, NRCS (2008) The PLANTS Database. National Plant Data Center, Baton Rouge, LA

Vaillancourt L, Du M, Wang J, Rollins J, Hanau R (2000a) Genetic analysis of cross fertility between two self-sterile strains of *Glomerella graminicola*. *Mycologia*, **92**, 430-435.

Vaillancourt L, Wang J, Hanau R (2000b) Genetic regulation of sexual compatability in *Glomerella graminicola*. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 29-44. American Phytopathological Society, St. Paul, Minnesota.

Vailllancourt L, Du M, Wang J, Rollins J, Hanau R (2000) Genetic analysis of cross fertility between two self-sterile strains of *Glomerella graminicola*. *Mycologia*, **92**, 430-435.

Viard F, Justy F, Jame P (1997) Population dynamisc inferred from temporal variation at microsatellite loci in the selfing snail *Bulinus truncatus*. *Genetics*, **146**, 973-982.

von Arx JA (1957) Die Arten der Gattung *Colletotrichum* Corda. *Phytopathologische Zeitschrift*, **29**, 414-468.

Vorsa N (1994) Breeding the American Cranberry **5**.

Vorsa N, Polashock JJ (2005) Alteration of Anthocyanin Glycosylation in Cranberry Through Interspecific Hybridization. *Journal of the American Society for Horticultural Science*, **130**, 711-715.

Wheeler HE (1950) Genetics of *Glomerella* VIII: A genetic basis for the occurrence of minus mutants. *American Journal of Botany*, **37**, 304-312.

Wheeler HE, Driver CH, Campa C (1959) Cross- and self-fertilization in *Glomerella*. *American Journal of Botany*, **46**, 361-365.

Wheeler HE, McGahen JW (1952) Genetics of *Glomerella* X: Genes affecting sexual reproduction. *American Journal of Botany*, **39**, 110-119.

Wheeler HE, Olive LS, Ernest CT, Edgerton CW (1948) Genetics of *Glomerella* V: Crozier and ascus development. *American Journal of Botany*, **35**, 722-728.

White TJ, Bruns T, LeeJ S, Taylor. JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds. Innis MA, Gelfand DH, Sninsky JJ, White TJ), pp. 315-322. Academic Press, Inc., New York.

Whitlock M (2011) G'st and D do not replace Fst. *Molecular Ecology*, **20**, 1083-1091.

Zaffarano PL, McDonald BA, Linde CC (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution*, **62**, 1418-1436.

Zhu P (1999) *Genetic diversity and systematics of Colletotrichum species on cranberry (Vaccinium macrocarpon)* Ph.D., Rutgers University.

Zhu P, Oudemans PV (2000) A long terminal repeat retrotransposon *Cgret* from the phytopathogenic fungus *Colletotrichum gloeosporioides* on cranberry. *Current Genetics*, **38**, 241-247.

Table 1: Eight microsatellite markers variable among *C. fructivorum, C. rhexiae, C. temperatum* and *C. kahawae*.

		C. fructivorum		*C. rhexiae*		*C. temperatum*		*C. kahawae*	
Locus	Motif	Allele size (bp)	N	Allele size (bp)	N	Allele size	N	Allele size	N
4H11	(tg)3gg(tg)5	262	28	262	2	252	1	264	2
13H1	(ag)5	164	28	164	1	163	1	164	2
CGU18758	(tg)3c(gt)3	206	21	205	2	206	1	206	2
EG359574	(gca)5	229	22	229	2	230	1	229	2
E2	(ac)3(ca)10gagaaacagaaacagaaacagacac(ag)12	233	28	231	2	216	1	216	2
EG359924	(cca)5	267	26	267	2	265	1	267	2
EG359976	(tc)6	306	24	306	2	302	1	302	1
DQ812968	(caa)6	286	15	286	2	287	1	289	2

N, number of strains tested

Table 2: Thirty-four microsatellite markers that cross-amplify and are variable within *C. fructivorum, C. rhexiae, C. temperatum* or *C. kahawae*.

		C. fructivorum				*C. rhexiae*		*C. temperatum*		*C. kahawae*	
Locus	Motif	Allele size (bp)	N_a	N	H_e	Allele size (bp)	N	Allele size (bp)	N	Allele size (bp)	N
5E5**	(ag)3gcgtgg(tg)7	207-209	2	27	0.0713	209	2	182	1	209	2
8B11	(tc)10	268-274	3	23	0.1626	268	2	263	1	268	1
10F5	(cgc)6c(gct)5	303	1	27	0	303	1	303	1	303-306	2
12B1**	(ctc)3(cgc)5	150-158	2	28	0.0689	155	1	150	1	158	2
EG360161**	(gca)6	157-162	2	28	0.0689	162	2	162	1	162	2
EG359900**	(ac)3g(ac)4	270-272	2	23	0.0832	272	2	272	1	272	2
EG360108**	(tc)3at(ac)3(tc)4	284-286	2	25	0.0768	286	1	NA	NA	286	1
EF408244	(tc)4ac(tc)3	351-355	3	23	0.1626	370	1	350	1	354-356	2
A4**	(ca)12	174-176	2	28	0.0689	178	2	167	1	167	2
AWFJM	(gtt)13	224-227	2	26	0.142	324	2	397	1	279	1
B2**	(ttc)10	403-411	2	26	0.074	396	1	389	1	NA	NA
C1**	(ac)12	191-195	2	28	0.0689	197	2	193	1	191	2
C6**	(tg)13	319-355	2	27	0.0713	334	1	309	1	307-345	2
C8**	(caa)9	135-141	2	28	0.0689	135	1	132	1	138-141	2
C10**	(ca)14	259-260	2	26	0.074	259	1	246	1	262-266	2

N_a, number of alleles; N, number of strains tested; H_e, gene diversity; *, allele size 495 approximated from 1% agarose gel; **, markers where one of two alleles is carried by *Coll886*; NA, marker did not amplify

Table 2: Continued

		C. fructivorum				*C. rhexiae*		*C. temperatum*		*C. kahawae*	
Locus	Motif	Allele size (bp)	N_a	N	H_e	Allele size (bp)	N	Allele size (bp)	N	Allele size (bp)	N
C11	(tg)17	334-352	3	28	0.1352	336	2	NA	NA	342	2
C2101	(cgt)3gattactattatcgtcatcgtcacttt(cgt)13	234-249	3	27	0.1399	237	2	237	1	231	2
C2574**	(gtg)15	220-242	2	26	0.074	221	1	201	1	214	2
C2324	(ctgctt)10	312-374	3	24	0.2882	312	1	319	1	321	1
C416**	(aca)14	358-495*	2	25	0.0768	344	2	NA	NA	NA	NA
D8	(gaa)10	222	1	17	0	231	1	231	1	222-291	2
D9**	(tg)15	288-323	2	28	0.0689	304	2	293	1	274	2
E5**	(tc)3ttcctgtggatgtctccccg(ac)11	363-375	2	27	0.0713	369	2	364	1	358	2
F5**	(tc)13actccctc(ac)5tc(ac)4	151-153	2	28	0.0689	165	2	156	1	190-213	2
G4	(ac)12tgagacacc(cg)3	397	1	27	0	397	2	397	1	411-415	2
G11**	(tg)7ta(tg)8	384-399	2	28	0.0689	383	2	397	1	383-385	2
C4168	(gtgcaa)12	202-267	6	28	0.5357	160	2	222	1	168	2
14F4	(at)3ggagata(tg)50(cg)4	233-346	8	28	0.676	257	2	249	1	260	2
10D10	(tg)34	235-282	7	28	0.4745	218	2	228	1	218-220	2
C1942	(gtt)25	290-321	7	28	0.6862	249	2	NA	NA	365	1
C2090	(agcaga)14	314-486	5	28	0.4133	247	2	273	1	272-295	2
B5-B4	(caa)12	366-376	3	28	0.1964	353	2	NA	1	352	2
2C1	(tg)3ta(tg)26ta(tg)3	206-217	4	28	0.3112	206	2	162	1	188	2
F9	(ac)25	260-287	4	28	0.44	248	2	263	1	248	2

N_a, number of alleles; N, number of strains tested; $H_{e,}$ gene diversity; *, allele size 495 approximated from 1% agarose gel; **, markers where one of two alleles is carried by *Coll886*; NA, marker did not amplify

Table 3: Twenty-nine markers that cross-amplify across several species within the *C. gloeosporioides* species complex.

	C. tropicale	C. tropicale	C. mangivora	C. mangivora	C. sp. indet. C	C. melanocaulon	C. ignotum
E2	203	203	NA	NA	208	208	193
C4168	143	143	143	143	153	153	153
C2324	A	A	A	A	A	A	A
C11	A	A	345	345	A	352	A
C1	182	182	201	201	182	182	183
AWFJM	196	193	NA	NA	A	A	NA
10A2*	225	225	A	225	225	A	225
10F5*	290	290	290	290	290	290	275
16F10*	197	197	194	194	192	192	188
AY357890*	218	218	218	218	218	218	220
CGU18758	A	203	204	A	204	204	204
EG359986*	A	300	300	300	300	300	A
EG360257*	175	175	175	175	175	175	175
EG359924	A	255	211	261	270	270	A
EG359694*	A	244	244	A	245	244	245
EG359827*	A	303	302	302	302	302	A
EG359976	A	305	303	303	307	A	A
DQ812968	A	A	312	312	286	286	A
M93427*	A	A	A	368	360	A	A
EG360161	166	166	173	195	163	163	159
EG359574	A	A	A	218	215	215	A
DQ413024*	267	267	A	266	266	A	266
AY386203*	A	A	289	289	289	289	A
EG359900	A	272	271	272	272	A	A
EG360108	288	288	286	286	286	286	286
EF408244	299	299	298	298	298	298	298
8B11	A	NA	258	A	NA	NA	264
14F4	154	154	154	154	154	154	154
2C1	A	A	231	231	233	233	216

NA, no amplification; A, amplification success at 55ºC but fragment length not determined; *, loci not present in Table 1 or Table 2

Table 3: Continued with the remainder of the species that were screened for cross-amplification.

	C. ignotum 2	C. sp. indet. E	Coll940	C. theobromicola	C. gloeosporioides	C. nupharicola	C. nupharicola
E2	197	199	NA	NA	234	NA	NA
C4168	153	NA	153	134	154	153	153
C2324	A	NA	NA	A	A	A	A
C11	329	333	341	NA	354	340	340
C1	NA	183	187	NA	185	187	187
AWFJM	A	NA	A	NA	A	A	A
10A2*	A	227	225	210	225	225	223
10F5*	NA	297	289	291	291	293	293
16F10*	188	188	188	183	194	188	188
AY357890*	A	223	223	NA	227	220	220
CGU18758	A	204	204	223	204	204	204
EG359986*	A	A	A	300	300	300	300
EG360257*	A	175	175	175	177	175	175
EG359924	A	A	A	276	261	250	250
EG359694*	A	245	NA	246	244	245	245
EG359827*	A	A	A	302	302	303	303
EG359976	A	A	A	A	301	308	308
DQ812968	A	A	A	A	283	NA	330
M93427*	A	A	A	A	363	325	315
EG360161	A	159	156	155	NA	160	160
EG359574	A	A	A	A	215	215	215
DQ413024*	A	266	266	264	266	266	266
AY386203*	A	A	A	A	289	289	289
EG359900	A	A	272	273	272	272	272
EG360108	A	286	286	NA	289	286	286
EF408244	A	298	298	NA	294	298	298
8B11	NA	306	286	263	260	NA	NA
14F4	154	154	A	159	158	154	A
2C1	229	A	NA	A	237	225	225

NA, no amplification; A, amplification success at 55ºC but fragment length not determined; *, loci not present in Table 1 or Table 2

Table 4: Microsatellite primer sequences and optimal annealing temperatures.

	Forward Primer	Reverse Primer	T_a (°C)
5E5	ccttttcgcacttgttttcgtttc	tacctgtcccggcaatatctgtct	55
8B11	cacagctccaagacatc	gctgtcacggtcggtaatccac	55
10F5	agagcatgtgccctttatccaaaa	cttgactctcggctcatcccttt	55
12B1	ggactcgtaggatccagcatcttc	attcgacggagctcccttctactc	55
EG360161	gttcagcactgcatccaacaggt	acaaatgaactgcaacgccagata	55
EG359900	gagaagtccgagatcgaccagaag	ccccgtaatcttcaaagtcaaccc	55
EG360108	agcggtaccttagcctatccccat	gccgcgctgactacactacctact	55
EF408244	tgcctagtatggtgtgtcggatgt	aacggcaaagggaaagtacttggt	55
A4	tctgtctggtagtggcatgttc	ctgatgatgattccctccctc	55
AWFJM	atcctcactcccaaccgc	ccagaagcgaaagaagaagaaa	55
B2	atgcctgataggtgagaccagt	gaatcgtttgagaatagcgacc	55
C1	aaatggaggtcaacaagacgac	catctctcctgcttgcataagg	55
C6	gtcatatttcggaccagagacc	catggaaacagaacagagcaga	55
C8	gtaggtcgagcagcacttacag	tctgtggtacgatggaagctat	55
C10	agatgaagtcaaaaggggttga	cttactccgtccacaaacaggt	55
C11	tgatgggtgtgtgtacgtgag	cagaaaggctagattggggtaa	55
C2101	tattgctatcttcgtcgtcgtg	tcatatacctagctgcgcttga	55
C2574	aaggtagccatgcaatatgagg	taggctctcctccagacaagaa	55
C2324	gatctccttgagcgacatgc	gctccttcttctttcccgtc	55
C416	agggtttgtgaaaggcaactaa	tgggatcgtgtattccatcata	55
D8	agaaaacaccaaatccacctgt	attaccaaaacgggtcgtctc	55
D9	cctggtatcatgtgtactgcgt	ttttcgtaccttgggctaccta	55
E5	gtaggcagcaggaccctaaac	gcgggaagaagagacataaaca	55
F5	agctaaagttaccgagccgaag	agttttggattttggaaggagg	55
G4	accttcaaaagcacccatgtag	gatcccgatgtgagtggg	55

Table 4: Continued

	Forward Primer	Reverse Primer	T_a (°C)
G11	gttgtgtccgttgtgacttcag	aaatgttcggggaaaacacc	55
C4168	ctggaaatagctgagacgggac	gtgggaaggacgagcctg	59
14F4	ctcaaaggaaaccgggggttgt	actctgcacatacctgcctctgtg	55
10D10	gagcgagttgatgcttcctc	gagcgagttgatgcttcctc	66
C1942	actgaccctcgaagaccttg	gacgagaacgatgacgagc	66
C2090	tgaagaaggcaagagaacacaa	cagacgggaaggtgaagtaaac	62
B5-B4	cgttcaaggagaagcagaagat	gaccagaaagtgacgatgacag	55
2C1	gcatcactccaacacatgct	tcatttgggggagacaaatc	55
F9	aatctgaccaccttgagagacg	agacgatgccaatcttttcg	55
10A2	gggattcgagacctagtgcgagta	tgccttaccaaccagaccaagact	55
10F5	agagcatgtgccctttatccaaaa	cttgactctcggctcatccctttt	55
16F10	tcccctgttgatctt	tcgcgacactcacagctact	55
AY357890	ttacaaatgattgcacccaaccac	ttttccctacaggcaacggatcta	55
EG359986	catcccgaagccctcagattc	caagatggagacggaaacatcgt	55
EG360257	aaggggcacgtcttttatagggtc	agtcacgatgtctccatagccctc	55
EG359694	gtttccagtcccaggagaacttgg	taatctgggggttgacatattggg	55
EG359827	cacccgcgtaaaataactacccct	acggagatggtcaggtagacgaag	55
M93427	accaggtacctccctaacttgcgt	aagatggatcgaggtgtggaagag	55
DQ413024	gcgattgcatacatgaacaccatt	tgctcctgtctacgttatcctccc	55
AY386203	tgaggtttttgagaagttccctcg	tcgtacttgttccagatgacgagc	55
4H11	ccctttccagaagacgagatcaag	gtcctcccacacgaccttcaac	55
13H1	cttgagcgacaacaga	tggtgtcatggtcttcaagc	55
203	ggcactattctcctctttggctga	acaagacgaagagtggcactgatg	55
1035	ctgcaacgtctttaccacctcctt	cggtagttggagacataagacccg	55
E2	cgactagccagacggaggttg	cacatctcagaacgcacgac	55
213	actcctccggttcttctcaggtct	ggagggttcgatagtctcttcggt	55
1021	tcttggtagtcgggtgccactaat	aggaacggtactgagactggatgg	55
1025	aaaacaccagcagatgtatcccgt	ctgctgttgaggagcgtgtaacat	55

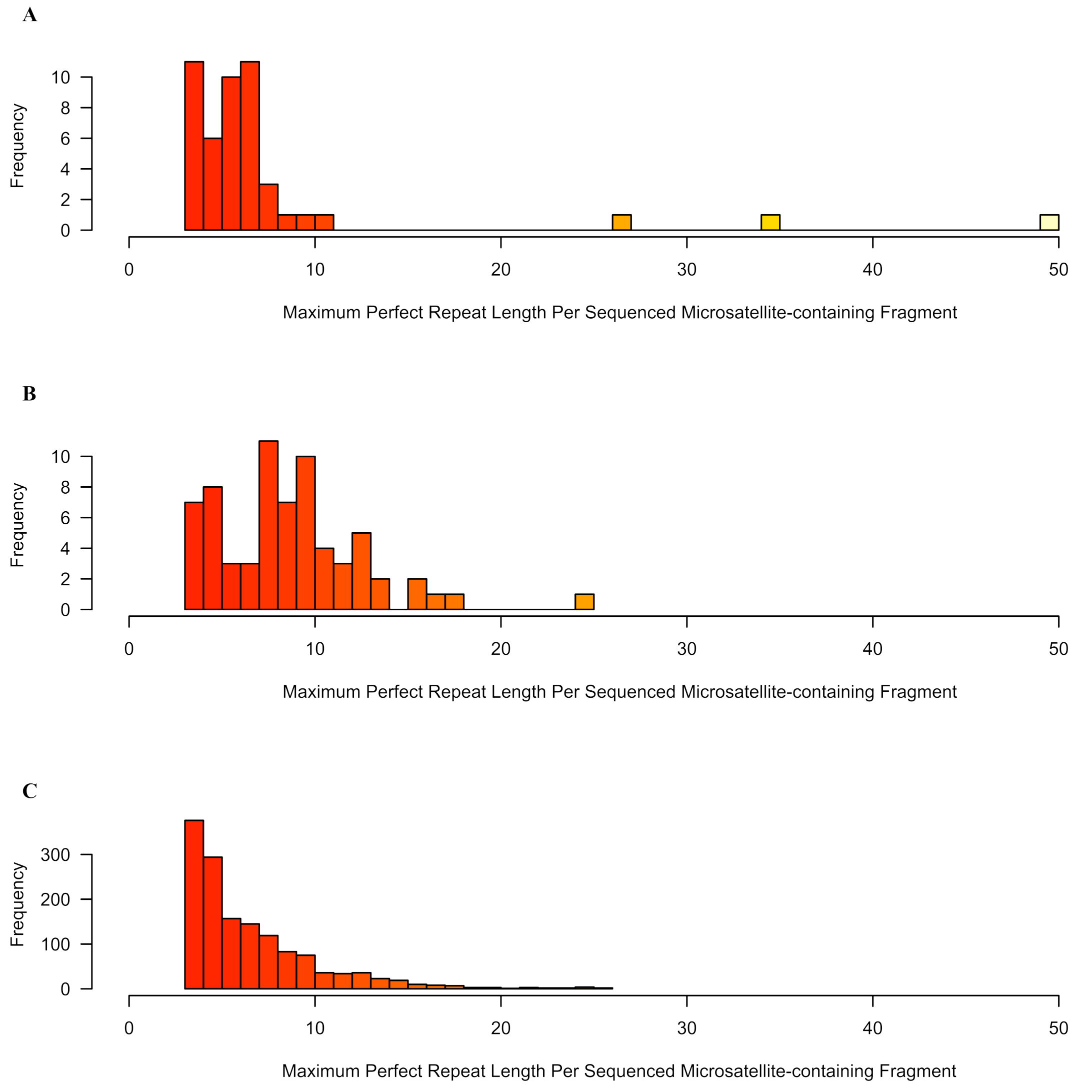


Figure 1: Frequency distribution plots of the maximum perfect repeat length found by read2marker per sequenced microsatellite-containing fragment. A) Membrane-hybridization enrichment B) Streptavidin bead-hybridization enrichment C) 454 sequencing of streptavidin bead-hybridization enriched library (note the change in scale of the Y-axis).

CHAPTER 4

PATTERNS OF GENETIC VARIATION WITHIN AND AMONG POPULATIONS OF *COLLETOTRICHUM FRUCTIVORUM* AND *COLLETOTRICHUM RHEXIAE*

ABSTRACT:

Understanding the influence of biotic and abiotic factors on the evolutionary trajectory of a species is central to many pursuits in evolutionary biology. We use a population genetics approach, sampling at several hierarchical levels in two closely related fungal species, *Colletotrichum fructivorum* and *Colletotrichum rhexiae*, to assess the role of extrinsic biotic and abiotic factors on intraspecific diversity and structure. We also investigate the role of intrinsic biological features of *C. fructivorum,* such as outcrossing and clonal reproduction, in shaping genetic diversity. We propose a hypothesis for the recent emergence of this recently described species, one that may also apply to other species within the *C. gloeosporioides* species complex. Factors related to host and habitat have probably influenced the divergence between *C. fructivorum* and *C. rhexiae,* as genetic isolation is almost complete despite sympatry in agricultural habitats. Based on a comparison with *C. rhexiae*, extrapolation to other species within the complex may have to be limited to those in association with agricultural crops. Clonal reproduction and human-mediated dispersal are important forces dissipating population structure of *C. fructivorum* over broad geographical scales. The data are inadequate to provide robust inferences on the population biology of *C. rhexiae*, but there is evidence that *C. rhexiae* is highly structured within a more restricted geographical range than is *C. fructivorum*. Additional and more variable markers coupled with broader sampling will need to be applied to make more robust inferences with respect to the population structure of *C. rhexiae*. We provide a foundation that will be useful for plant breeders to develop resistant cranberry cultivars, plant pathologists to diminish the yield losses caused by *C. fructivorum*, and mycologists interested in the evolution of the *C. gloeosporioides* species complex.

INTRODUCTION

Colletotrichum Corda is among the most important genera of plant pathogenic fungi, recognized for its significant economic impact on a broad range of plant hosts (Dean *et al.*, 2012; Farr *et al.*, 2011). Species of the *C. gloeosporioides* complex are distributed worldwide as plant pathogens and are among the most common taxa recovered in surveys of fungal endophytes (Rodriguez, Redman, 2000). Despite the abundance and economic importance of the complex, understanding the factors that influence its microevolutionary trajectory and species diversity have been compromised by the lack of phylogenetic information available for delimiting species. Recent work to delimit species within *C. gloeosporioides s.l.* using more informative phylogenetic markers, however, has laid the groundwork for further comparative studies (Chapter2), (Rojas *et al.*, 2010; Silva *et al.*, 2012b). We are now in a position to develop a refined perspective of the ecology and natural history of individual species and the species complex as a whole.

Several fungal species are responsible for fruit-rot in commercial cranberry production; a yield-limiting factor that is responsible for significant economic losses in the industry (Oudemans *et al.*, 1998). *Colletotrichum fructivorum* Doyle, Oudemans & Rehner is one such species, broadly distributed throughout agricultural production areas as a fruit-rot pathogen of the cultivated cranberry (*Vaccinium macrocarpon* Aiton) (Chapter 2). It is also capable of persisting latently in stem tissue as an endophyte and colonizing an alternate sympatric host species, *Rhexia virginica* L. (Chapter 2),(Stiles, Oudemans, 1997).

Colletotrichum rhexiae Ellis & Everh. is the sister species of *C. fructivorum,* and colonizes both *V. macrocarpon* and *Rhexia virginica.* However, *C. rhexiae* has not been found on *V. macrocarpon* in commercial cranberry production areas in New Jersey, but does colonize

R. virginica in those populations where *C. fructivorum* and *C. rhexiae* are sympatric (Chapter 2). Previous research indicates that these species have been sympatric for more than a century (Ellis, Everhart, 1894; Halsted, 1889; Shear, 1907c). *Colletotrichum rhexiae* was described in 1894 from Delaware as a pathogen of *R. virginica* and C.L. Shear deposited a strain of *C. fructivorum* to the Centraalbureau voor Schimmelcultures in 1922 (Chapter 2), (Ellis, Everhart, 1894). Their close phylogenetic affinity and sympatric distribution suggests the boundary between species may be porous, however their respective host associations in agricultural habitats indicates they remain ecologically distinct.

The three primary objectives of this study were: (*i*) to understand the distribution of genetic diversity within and among populations of *C. fructivorum* and *C. rhexiae,*(*ii*) to determine if there is evidence of admixture between these species where their distributions overlap, and (*iii*) to determine if there is evidence of outcrossing in field populations of *C. fructivorum.* The distribution of genetic diversity within a species can reveal the factors that have shaped populations and may also yield insight into the potential of species to adapt to environmental changes. This is of particular importance for plant pathogens, where understanding the evolutionary potential of the pathogen population will provide useful information for optimizing disease management and breeding for host-resistance (McDonald, Linde, 2002b). It is also evident from studies in a diversity of distantly related organisms that ecological speciation can occur over short time scales and that adaptive divergence is probably more pronounced among sympatric populations than allopatric populations (Antonovics, 2006; Dettman *et al.*, 2008; Dettman *et al.*, 2007; Hendry *et al.*, 2000; Silvertown *et al.*, 2005). Understanding the biotic and abiotic forces that influence populations within a species may provide valuable information about the factors influencing speciation. This is particularly important for mycologists interested in understanding the breadth of species diversity within a genus like *Colletotrichum*, known for having a cryptic lifestyle. Similarly, assessing admixture between two closely related species, such as *C. fructivorum* and *C. rhexiae,* provides an objective test of species delimitation, while providing some insight into the factors maintaining genetic isolation.

The economic impact of *C. gloeosporioides s.l.* on a diversity of crop species has led to several studies investigating its population structure, particularly with respect to lifestyle differences (endophyte versus pathogen) and host specificity (Freeman *et al.*, 1996; Guozhong *et al.*, 2004; Kaufmann, Weidemann, 1996). These studies are variable in their conclusions. Clonal and sexual lineages have been reported to share multiple host species (Freeman *et al.*, 1996; Guozhong *et al.*, 2004; MacKenzie *et al.*, 2007), while other studies suggest the presence of host-specific populations (Kaufmann, Weidemann, 1996). Given the phylogenetic diversity within *C. gloeosporioides s.l.,* even within a single ecosystem (Chapter 2), it is important to first define phylogenetic lineages that may constitute a metapopulation in order to infer the ongoing processes that account for observed patterns of diversity. Otherwise, patterns that result from historical divergence can be confused with present day selective forces. The recent development and validation of suitable phylogenetic markers for delimiting species boundaries enables us to move forward, assessing the factors that drive infraspecific differentiation (Chapter 2), (Rojas *et al.*, 2010).

Host and habitat are important factors influencing the distribution of individual species within *C. gloeosporioides s.l.* (Chapter 2). Thus, it is important to evaluate host population ecology in order to determine the factors that may be influencing fungal population structure. The American cranberry (*Vaccinium macrocarpon* Aiton) is native to temperate areas of North

America and and is one of North America's few native crop plants (USDA, NRCS, 2008). It is distributed in natural stands throughout the northeast, and was brought into cultivation in the early 19th century in Massachusetts and New Jersey, but is still a relatively undomesticated crop. Many of the dominant varieties used today, such as *cv.* Early Black, were selected from wild stands, prior to the Civil War (USDA-CSREES, 2001). Major production areas now include Massachusetts, Wisconsin, New Jersey, British Columbia, Washington and Oregon (Stiles, Oudemans, 1999). Given the history and cultivation practices of the cranberry industry, there are several biotic and abiotic factors that may contribute to the partitioning of genetic diversity in *C. fructivorum* across broad geographical scales and at the landscape level. Cranberry farms in commercial growing regions may effectively represent islands of evolution for the species, as there is apparently no suitable habitat providing a contiguous dispersal corridor between these areas. There are several factors that may contribute to population structure within a single growing region, given what is known about the biology of *Colletotrichum,* the biology of the host species, and the agricultural practices employed during cranberry production. For example, cranberry beds are engineered such that water flow into and among beds can be regulated by dikes and ditches that form a network connecting the beds in a single field and across farms in a single irrigation network. During harvest on commercial cranberry farms, beds are flooded and the fruit is dislodged from the plant, allowing it to rise to the surface where it can be collected. This harvesting method may influence the distribution of genetic variation of *Colletotrichum* species, particularly fruit pathogens such as *C. fructivorum*, as evidence from laboratory studies suggest that *Colletotrichum* is water dispersed (Nicholson, Moraes, 1980). Similarly, the planting regimen of most cranberry growers and the demands of the commercial market lead to the separation of cranberry cultivars among beds. Selection criteria for early cultivar development was based in part on evidence of resistance to fruit rot; descriptions of existing cultivars note variation in storage quality and susceptibility to storage rot; and recent work confirms the presence of heritable resistance in cranberry cultivars (Eck, 1990; Johnson-Cicalese *et al.*, 2009). All of this suggests variation in disease resistance among cranberry cultivars, and selection of fungal haplotypes that can overcome resistance mechanisms in the host can lead to fine-scale population structure in the pathogen.

The tools for species delimitation have only recently become available for *C. gloeosporioides s.l.* (Du *et al.*, 2005; Rojas *et al.*, 2010; Silva *et al.*, 2012b). This makes it difficult to compare previous studies investigating the influence of biotic forces on the evolution of the species complex. Nonetheless, it is clear that there is a dynamic relationship between host-fungus biology that can lead to the partitioning of fungal populations or species, including chemical, structural, or other differences among host species and organs that can lead to niche specialization(González *et al.*, 2006; Guozhong *et al.*, 2004; Kaufmann, Weidemann, 1996; MacKenzie *et al.*, 2007). *Colletotrichum fructivorum* is known to occur on both stem and fruit of *V. macrocarpon*, and it can also be isolated from sympatric *R. virginica* (Chapter 2), but it is not clear how these factors may influence infraspecific gene flow and selection.

Sampling from the same population at different time points can provide valuable information with respect to population dynamics and demography, as has been demonstrated in a wide range of organisms (Cole, Lynn, 1996; Hang *et al.*, 2010; Ramakrishnan *et al.*, 2005; Riccioni *et al.*, 2010; Viard *et al.*, 1997). Pronounced changes have been observed in the species composition of fruit rot fungi and disease severity in the same commercial cranberry beds sampled from one year to the next (Stiles, Oudemans, 1999). In addition, an active research program in cranberry pathology has led to major improvements in disease management

(Oudemans *et al.*, 1998). Changes in disease management may have a profound impact on pathogen populations, leading to significant changes in genetic diversity over time (McDonald, Linde, 2002b). Conversely, understanding how pathogen diversity has changed over time can provide insight into how disease management strategies impact individual pathogen species and their potential to adapt to new selective pressures.

Under a proposed framework to predict the evolutionary potential of pathogen populations, McDonald and Linde (2002) suggest pathogens with a mixed mating system pose the greatest risk. Sexual recombination can produce novel genotypes with increased fitness, and asexual reproduction leads to the amplification of this genotype during an epidemic spread. While the overall potential of a species to adapt to selective forces imposed by disease management strategies is a composite of reproductive strategy, the potential for gene flow, effective population size, and mutation rate, understanding the reproductive mode is an important step in assessing its evolutionary potential (McDonald, Linde, 2002a; McDonald, Linde, 2002b; Milgroom, 1996). *Colletotrichum* is capable of reproducing both sexually and asexually, but sexual reproduction does not always result in assortment and shuffling of genetic variation among loci. There are both self-fertile (homothallic) and cross-fertile (heterothallic) species. *Colletotrichum fructivorum* is known to form sexual reproductive structures, but it is not known whether it outcrosses in field populations. It is necessary to confirm the ability to reproduce through selfing with laboratory studies as well as determine the reproductive mechanisms that prevail under field conditions by looking for evidence of genetic reassortment among genotyped individuals from field collections.

The presence of sexual reproductive structures in *C. fructivorum* and *C. rhexiae* suggests the potential for genetic exchange among conspecific individuals (unless they are obligately homothallic). However, despite their overlapping distribution and ability to reproduce sexually, phylogenetic analysis indicates the presence of other biological factors reinforcing species boundaries (Chapter 2). But an objective test of species boundaries is better approached with more rapidly evolving markers variable within species, such as microsatellites, and using appropriate non-phylogenetic algorithms to determine if there is any evidence of introgression.

Colletotrichum fructivorum and *C. rhexiae* can both be isolated from *R. virginica*, but they occupy different ecological niches and only come into contact in agricultural areas (Chapter 2). In contrast to *C. fructivorum, C. rhexiae* has not been isolated from cultivated cranberry and therefore gene flow among regional populations is not likely to be mediated by anthropogenic activity. These distinct profiles provide a unique opportunity for a comparative study of the population genetics of each species and the role of human-mediated dispersal on gene flow in *Colletotrichum.*

MATERIALS AND METHODS

Sampling

The focus of the sampling efforts for this study was on *C. fructivorum* and *C. rhexiae. Colletotrichum fructivorum* was collected at multiple spatial scales ranging from multiple isolates within a 50 cm^2 sampling area (sample point) to collections distributed across six regional cranberry cultivation areas. We employed a hierarchical sampling strategy for isolates from Delaware and Washington and the majority of the isolates from New Jersey. We used a hierarchical sampling approach within New Jersey to address the influence of biotic and abiotic factors on the distribution of genetic diversity over limited spatial scales. Samples were isolated from *V. macrocarpon* across watersheds, subwatersheds, irrigation networks, beds, host-plant organs, and cultivars as well as from *R. virginica*. Spatially hierarchical sampling was conducted by sampling symptomatic fruit from up to five sample points within a bed, with each sampling point separated by a minimum of approximately 20 meters. The GPS coordinates were recorded for each sampling point. Collections of *C. fructivorum* from a study by Stiles and Oudemans (1997) were included to investigate evidence of temporal structure and to infer demographic changes. Specific GPS coordinates are not available for these collections; however, isolates could be assigned to specific beds. These isolates were included in analyses of regional population structure and temporal structure, but were excluded from all other analyses of population structure. A bulk sampling approach was employed for collections from Massachusetts, British Columbia, and the majority of isolates from Wisconsin. Symptomatic fruit originating from a single bed were pooled, so that the origin of isolates from these collections was attributable to distinct beds within a region. Collections of *C. fructivorum* from *R. viriginica* were isolated from host plants growing at the edge of irrigation ditches adjacent to and within cranberry beds.

Colletotrichum rhexiae was collected in three regional populations; two wild populations (Delaware and Maryland) and one agricultural population (New Jersey). Samples were isolated from *R. virginica* in all three regions and from both *R. virginica* and *V. macrocarpon* in Delaware. The spatial distribution of isolates was dependent on the patchy distribution of host plants. Collections from Delaware were isolated from host plants separated by no more than 150 meters and the maximum distance between two collections in Maryland was less than 50 meters. In contrast, collections of *C. rhexiae* from New Jersey were separated by up to 3,450 meters.

Fungal isolation and culturing

Symptomatic and asymptomatic tissues were surface sterilized in 10% bleach (final concentration 0.6125% sodium hypochlorite) for between 3-5 minutes, depending on the permeability of the tissue, and plated on V8 juice agar. Fungi were isolated from fruit, stem, and leaf tissue of *V. macrocarpon,* fruit of *V. oxycoccos,* and stem and leaf tissue of *R. virginica*. Strains morphologically similar to *C. fructivorum* or *C. rhexiae* were isolated by transferring conidia or hyphal tips to sterile media and preserved on corn meal agar (CMA:BD Diagnostics) slants stored at 6 C, on CMA blocks stored in 1.5 mL microcentrifuge tubes at 6 C, and in 10% glycerol at -80 C. In order to determine the potential for selfing in *C. fructivorum*, single ascospores were plated on V8 juice agar and potato dextrose agar (PDA: BD Diagnostics) and observed over the course of four weeks for the production of fertile perithecia.

DNA extraction, PCR amplification, and Fragment Sizing

Isolates were grown on potato dextrose broth (Difco) for 5-7 days before mycelium was harvested, blotted dry on sterile paper towels and dehydrated for 6-10 hours in a vacuum centrifuge on low heat. Approximately 20 mg of dried tissue was homogenized in the FastPrep FP120 (MP Biomedicals, LLC., Solon, OH, USA) and used for DNA extraction using the DNeasy Plant MiniKit (Qiagen Inc., Valencia, CA, USA) or a standard phenol-chloroform extraction method. PCR amplification reactions were performed in 10 ul reactions containing 4.37 μL of autoclaved ion-exchanged water, 0.8 μL of dNTP mixture (2.5 mM stock of each dNTP), 1 μL of bovine serum albumin (BSA; 0.25 μg μL−1 stock), 1 μL of buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 20 μg/mL bovine serum albumin (BSA)], 2 μL betaine (1.2 M stock), 100 nM 5’

M13-tailed forward primer, 200 nM reverse primer, 200 nM universal fluorescent-labeled M-13 primer, 0.08 µL of *Taq* polymerase and 0.5 µL of DNA.

PCR reactions were run on an Eppendorf vapo.Protect Mastercycler pro S with the following cycling parameters: initial denaturing for 2.5 min at 94°C, followed by 28-30 cycles of 94°C for 30 s, 55-66°C for 30 s (depending on the optimal annealing temperature of the primers), and 64°C for 45 s, followed by a final extension at 64°C for 20 min. PCR fragments were sized on the Beckman Coulter CEQ 8800 along with the DNA Size Standard Kit 400 (Beckman Coulter PN 608098) following the Frag-3 separation method (denaturation at 90°C for 2 min; injection voltage of 2 kV at 50°C for 30 s; separation voltage of 6.0 kV at 50°C for 35 minutes). In order to size PCR products, the fragment analysis parameters were adjusted to PA ver. 1 to calibrate for the phosphoramidite-linked dye primers. A proportion of individuals ranging from 9-26% of all individuals were genotyped at each locus more than once to assess the reliability of fragment size estimation. All isolates, including those of *C. fructivorum, C. rhexiae, C. kahawae,* and *C. temperatum*, were genotyped at 8 microsatellite loci previously shown to be polymorphic in *C. fructivorum* (L2C1, L10D10, L14F4, LB5B4, LC2090, LC4168, LC1942, and LF9) (Chapter 3). Isolates were also characterized by deletion length variation at the 3' end of a long terminal repeat (LTR) retrotransposon, *Cgret*, using primers P23ad-F (5'ATAGCGCCTATCTCATACGGCGAAACA) and P27ad-R (5'GTTACGGTCAGCTATCCGACAGTAGGG) (Zhu, 1999; Zhu, Oudemans, 2000). PCR amplification reactions of *Cgret* were done in 25 µl reactions containing 9.3 µL of autoclaved ion-exchanged water, 2.5 µL of dNTP mixture (2.5 mM stock of each dNTP), 2.5 µL of bovine serum albumin (BSA; 0.25 µg µL−1 stock), 2.5 µL of buffer [200 mm Tris pH 8.8, 100 mM KCl, 100 mM $(NH_4)_2SO_4$, 20 mm $MgSO_4 \cdot 7H_2O$, 1 % (v/v) Triton X-100, 50 % (w/v) sucrose, 0.25 % (w/v) cresol red], 5 µL betaine (1.2 M stock), 1 µL of each primer (0.67 µM final concentration), 0.2 µL of *Taq* polymerase (GenScript USA Inc., Piscataway, NJ, USA) and 1 µL of DNA. PCR reactions were run with the following cycling parameters: initial denaturing for 3 min at 94°C, followed by 30 cycles of 93°C for 1 min, 60°C for 50 s, and 64°C for 1.5 min, followed by a final extension at 64°C for 5 min. Polymorphic states were scored on 1% agarose gels with ethidium bromide staining. Three different polymorphisms were scored among the isolates included in this study: 1) a single full length 1.8 kb amplicon at the 3' end of the LTR region, 2) a single full length 1.8 kb amplicon plus an additional fragment with an approximately 800 bp deletion, and 3) an amplicon with a 1.2 kb deletion (Zhu, 1999).

Data Analysis

Allele Binning

In order to minimize scoring error based on microsatellite electropherograms, we used the R package 'MsatAllele_1.03' to visualize the distribution of raw allele sizes, compute summary statistics for each bin, and create whole number allele codes from a database of raw allele sizes (real numbers) based on migration time on the CEQ8800 (Alberto, 2009). The bin sizes for each allele assignment were determined based on the distribution of raw allele sizes rather than user-defined arbitrary bin size classes (Alberto, 2009). Allele assignments were then exported from MsatAllele for downstream analysis.

Data Sets

We used the eight microsatellite loci and the retrotransposon locus, *Cgret* (Zhu, 1999; Zhu, Oudemans, 2000), to genotype a total of 341 isolates from four closely related species, *C. fructivorum* (*n* = 303)*, C. rhexiae* (*n* = 32)*, C. kahawae* (*n* = 4)*,* and *C. temperatum* (*n* = 2).

Due to hierarchical sampling within *C. fructivorum,* the *n* = 303 isolates were partitioned for comparisons at multiple spatial scales. Clone correction was performed on all datasets in GenoDive 2.0b22 (Meirmans, Tienderen, 2004) by removing identical haplotypes that occurred more than once within a population. Clone correction mediates complications in the analysis of population structure in pathogen populations for which the amplification of clones through asexual reproduction is known to be common (Milgroom, 1996).

Regional Data - Colletotrichum fructivorum – The genetic structure among 5 regional populations of *C. fructivorum*, all isolated from cultivated cranberry, was investigated *for n* = 301 isolates; Delaware (*n* = 21), New Jersey (*n* = 193), Massachusetts (*n* = 34), Washington/British Columbia (*n* = 16), and Wisconsin (*n* = 37). *Coll886* and *CBS124* were excluded from most analyses of these data because *Coll886* is a singleton from an isolated wild cranberry population and *CBS124* was deposited in 1922 and is thus temporally isolated from all other recent collections (1994-2011) included in this study. In order for balanced comparison of resampled data sets, regional populations were partitioned, if possible, into subpopulations of a minimum of *n* samples, where *n* is equal to the smallest regional population sample. This led to the partitioning of the Wisconsin and New Jersey samples into two and six subpopulations, respectively. The Wisconsin population was subdivided based on geographic proximity into the following two subpopulations: WI_1 (*n* = 20) and WI_2 (*n* = 17). The New Jersey population was subdivided on the basis of watershed and by irrigation network into the following six subpopulations: NJ_1 (*n* = 33), NJ_2 (*n* = 29), NJ_3 (*n* = 30), NJ_4 (*n* = 16), NJ_5 (*n* = 35), and NJ_6 (*n* = 33) (Refer to Figure 1 in Appendix 4A). Subpopulations with sample sizes less than 16 were combined if they were adjacent, but isolated from other subpopulations. Some individuals from isolated drainages with a small sample size could not be reasonably and unequivocally grouped with a proximate drainage and thus were excluded. This formulation produced a dataset of 284 individuals from 5 regions and 11 subpopulations. Clone correction by subpopulation resulted in a dataset of 204 individuals with the following population sizes: Delaware (*n* = 16), NJ_1 (*n* = 21), NJ_2 (*n* = 22), NJ_3 (*n* = 22), NJ_4 (*n* = 14), NJ_5 (*n* = 31), and NJ_6 (*n* = 28), Massachusetts (*n* = 25), Washington/British Columbia (*n* = 12), WI_1 (*n* = 8) and, WI_2 (*n* = 5). The two Wisconsin subpopulations were combined for subsequent analysis of the clone corrected data, to yield a population of 13 haplotypes. An additional regional dataset was constructed for the discriminant analysis of principal components that was clone corrected at the level of region in order to limit the effect of clonal lineages. This further reduced the regional dataset to 194 haplotypes.

Regional Data - Colletotrichum rhexiae - Regional structure among *C. rhexiae* populations was investigated with 32 genotyped isolates among two host species (*Vaccinium macrocarpon* and *Rhexia virginica*) and three regional populations; Delaware (*n* = 18), Maryland (*n* = 6), and New Jersey (*n* = 8). Six of the nine loci used to genotype *C. fructivorum* were variable within *C. rhexiae*. These loci (L10D10, L14F4, L2C1, LC1942, LC2090, LC4168) were used to infer population structure among regional populations. Clone correction would reduce the dataset to 6 haplotypes (Delaware: *n* = 1; Maryland: *n* = 1; and New Jersey: *n* = 4), but to provide a quantitative estimate of the differentiation among these populations we retained all haplotypes for further analyses.

Fine-scale Data - Colletotrichum fructivorum - Fine-scale population structure was investigated within New Jersey, using various biotic and abiotic features to subdivide the n = 193 isolates. Temporally separated isolates from 1994-1997 were excluded from all analyses with the exception of inferences of temporal structure. Isolates were also excluded if the data necessary to assign them to the respective populations was not available. Clone correction was done by population, prior to estimating differentiation.

In order to assess genetic differentiation among watersheds, we genotyped individuals from the Rancocas and Mullica watershed, an adjacent pair of major watersheds which encompass the commercial cranberry farms in New Jersey. Clone correction yielded a dataset of 134 individuals ($n = 29$ & $n = 105$, from Rancocas and Mullica drainages, respectively). We assessed differentiation among subwatersheds with collections distributed among 6 subwatersheds. Clone correction by subwatershed produced a dataset of 148 individuals: Batsto River (BR: $n = 9$), Mullica River above the Batsto River (MRABR: $n = 25$), Oswego River (OR: $n = 18$), North Branch of Rancocas Creek (RCNB: $n = 23$), South Branch of Rancocas Creek (RCSB: $n = 8$), and West Branch of the Wading River (WBWR: $n = 65$). Isolates were assigned to a watershed or subwatershed using the NJDEP 14-digit hydrologic unit code delineation (DEPHUC14) for New Jersey (available: http://www.nj.gov/dep/gis/digidownload/metadata/statewide/dephuc14.htm).

We genotyped individuals from various cultivars, but with a focus on two widely planted cultivars that have been in use for several decades, cv. Early Black and cv. Stevens, to assess the influence of host cultivar on population structure. Clone correction yielded a dataset *of* $n = 93$ isolates ($n = 42$) from Early Black (EB) and ($n = 51$) from Stevens (ST).

Differentiation among isolates on the basis of host organ was assessed with 119 isolates from stem and fruit, (n = 11) for stems and (n = 108) for fruit, respectively. Isolates of *C. fructivorum* from two different host species, *V. macrocarpon* and *Rhexia virginica*, were also genotyped to evaluate the role of host species on population divergence. Twelve (n = 12) isolates of *C. fructivorum* from *R. virginica* and (n = 107) haplotypes from *V. macrocarpo*n were included in the analysis.

We investigated temporal changes in *C. fructivorum* populations by estimating genetic differentiation among isolates collected from 1994 to 1997 (Stiles, Oudemans, 1997) and those collected from 2008 to 2010 with a clone-corrected dataset of 139 haplotypes, (n = 24) from 1994-1997 and (n = 115) from 2008-2010, respectively.

Data for assessment of reproductive mode, demography and isolation by distance - There are many potential causes of linkage disequilibrium including: population structure, linkage, natural selection, and asexual reproduction (Agapow, Burt, 2001; Milgroom, 1996) 1). In order to test for random mating and assess the reproductive mode of *C. fructivorum* while reducing the influence of population structure, we constructed a restricted dataset retaining isolates for which specific GPS coordinates were available in New Jersey. Clone correction by sample point yielded a dataset of 129 haplotypes from 103 sample points. In addition, we constructed a second dataset to infer reproductive mode by retaining only unique haplotypes across all the GPS-referenced isolates, reducing the dataset to 101 haplotypes. The dataset of n=129 isolates was also used to infer the demographic history of *C. fructivorum* in New Jersey and to assess the correlation between genetic distance and geographic distance.

Diversity summary

A summary of the genetic diversity within and among populations is reported in Table 1 for the regional dataset, divided into eleven subpopulations. We report the sample size (n), number of haplotypes after clone correction (n_h), clonal fraction (CF), defined as 1-[(number of distinct multilocus haplotypes)/(total number of isolates)] (Dilmaghani *et al.*, 2012). We report genotypic diversity (G), defined as the probability that two randomly selected genotypes are different, as calculated in Multilocus v1.3b (Agapow, Burt, 2001), and the unbiased gene diversity (H_e), calculated as the average per locus, using Genalex 6.1(Peakall, Smouse, 2006). We also report the frequency of the most common genotype (MF), allelic richness (A_{R}) averaged over loci, and private allelic richness (P_{AR}), averaged across loci for each subpopulation. We estimated (A_{R}) and (P_{AR}), correcting for sample size differences with a rarefaction procedure in ADZE (Szpiech *et al.*, 2008), standardizing to the smallest sample size in the collection. This estimate of the number of alleles (or private alleles) that would occur in smaller samples of individuals is based on the frequency distribution of alleles at a locus, and offers similar precision to making estimates from multiple subsamples (Leberg, 2002).

Measures of Genetic Differentiation and Population Structure

A test of the null hypothesis of no significant differentiation among populations of *C. fructivorum* was carried out with an analysis of molecular variance (AMOVA), computing ϕ_{PT} values (an analogue of F_{ST}) for each pair of populations with 1000 permutations in GENALEX v.6.1 (Excoffier *et al.*, 1992; Peakall, Smouse, 2006). Jost (2008) argues that G_{ST} and its analogues (among them ,ϕ_{PT} and F_{ST}) do not measure actual population differentiation, particularly for high levels of within population variation, as expected with microsatellite markers and has developed another metric, D_{est}, that does. Meanwhile, Meirmans (2006) developed a method to scale the measure of ϕ_{PT} to the [0,1] interval, gauging ϕ_{PT} by the maximum value it could achieve (ϕ_{PTmax}); ϕ'_{PT} ($=\phi_{\mathrm{PT}}/\phi_{\mathrm{PTmax}}$) ranges from 0 (no divergence) to 1 (maximum divergence).

The utility and application of traditional measures, G_{ST} and F_{ST}, and various replacement metrics, ϕ'_{PT}, G'_{ST}, G''_{ST}, and D_{est}, continue to be debated, and a consensus has yet to emerge (Hedrick, 2005; Jost, 2008; Jost, 2009; Meirmans, 2006; Meirmans, Hedrick, 2011; Ryman, Leimar, 2009; Whitlock, 2011). Rather than choose among these metrics, we report multiple measures of population structure. For population comparisons with significant ϕ_{PT} values, we used two additional measures of population differentiation, ϕ'_{PT} and D_{est}, to characterize the degree of differentiation among populations (Jost, 2008; Meirmans, 2006). ϕ_{PTmax} was estimated in GENALEX v.6.1, with a dataset recoded using RecodeData v.0.1 (http://www.bentleydrummer.nl/software/) and ϕ'_{PT} was calculated in Excel. D_{est} was calculated with the R packages 'mmod' v.0.31 and 'adegenet' v.1.3.1, using 95% confidence intervals of the mean estimated from 1000 bootstrap replicates (Jombart, 2008). In addition, to assess the impact of variation in sample size among populations on inferences of population structure, we resampled 100 new data sets, using a custom R script, such that all population sizes were equal, with the population size determined by the sample size of the smallest population sample in the original dataset. AMOVA was then conducted on each resampled dataset with 1000 permutations, using the R packages adegenet v.1.3.1 and pegas v.0.4-1. The global and single locus estimates of subpopulation gene diversity (H_{S}), total gene diversity (H_{T}), G_{ST}, and D_{est} were determined for the resampled datasets of *C. fructivorum* in the R package mmod and adegenet. Pairwise comparsions of population structure for resampled datasets with more than two populations were made with D_{est} in mmod. The mean and 95% confidence interval are reported for each differentiation measure for the resampled data. A test of the null hypothesis of no significant differentiation among populations of *C. rhexiae* was carried out using an analysis of molecular variance (AMOVA) with 1000 permutations in Genalex v. 6.1. D_{est} among populations of *C. rhexiae* was calculated with 95% confidence intervals of the mean estimated from 1000 bootstrap replicates. We also tested for isolation by distance among individuals of *C. fructivorum* collected in New Jersey by evaluating the correlation between a genetic distance matrix, with the difference between individuals determined by the number of distinct alleles, and a geographic distance matrix in Genalex 6.1 with a Mantel test.

Principal Components Analysis

Multivariate analyses are a useful approach to summarizing multivariate genetic information and do not require strong assumptions, such as the absence of linkage disequilibrium (Jombart *et al.*, 2009). In order to explore and visualize a summary of the genetic diversity of individual genotypes among sampled populations, a principal components analysis (PCA) was implemented in 'adegenet'. A PCA was also carried out with a dataset that included *C. fructivorum* and *C. rhexiae*, as well as *C. kahawae* and *C. temperatum*, the latter two of which were not included in the genetic differentiation estimates, due to small sample sizes.

We also used PCA to evaluate the relative placement of two isolates, CBS124 collected by C.L. Shear and deposited to the Centraalbureau Voor Schimmelcultures in 1922 and *Coll886* sampled from a wild population of *V. oxycoccos* in Pennsylvania. Both isolates were previously determined to be *C. fructivorum* (Chapter 2), but were not included in the genetic differentiation estimates.

Cluster Analysis and Discriminant Analysis of Principal Components

Inferring population structure on the basis of *a priori* defined populations can obscure biologically relevant differentiation if inappropriate factors are used to define populations. Bayesian cluster analysis has been used to detect population differentiation at low levels of divergence (Latch *et al.*, 2006). We used the software STRUCTURE 2.3.2.1 to assess admixture between regional populations and to determine whether there was a signature of population structure not expected on the basis of *a priori* population assignment (Falush *et al.*, 2003; Pritchard *et al.*, 2000). We used the admixture model with correlated allele frequencies to assess the probability of the data for numbers of populations (values of *k*) ranging from 1 to 18 with burn-in set to 30,000, post burn-in steps to 60,000, and 10 iterations for each value of *k*. STRUCTURE HARVESTER, which implements the method of Evanno et al. (2005), was used to evaluate the optimal number of populations (*k*) indicated by the data (Earl, vonHoldt, 2012; Evanno *et al.*, 2005). The analysis was then rerun with extended burn-in at 100,000 and post burn-in to 500,000 steps for values of *k* ranging from 1 to 12, and reanalyzed with STRUCTURE HARVESTER. The assignment of individuals to populations was evaluated for various values of *k*.

In addition, we used a non-parametric approach of population clustering, *k-means*, after transforming the data using principal components analysis and retaining all the principal components (PC axes) for the clustering algorithm. The number of clusters that best describe the data was evaluated with Bayesian Information Criterion (BIC). BIC values were plotted for increasing values of *k* and the lowest BIC score was used to select the optimal value for *k*. Discriminant analysis of principal components (DAPC) was performed on the clusters inferred in the *k-means* clustering. DAPC constructs synthetic variables from the genetic data that maximizes among-group variance while minimizing within-group variance (Jombart *et al.*, 2010). The number of PC axes that explained approximately 90% of the cumulative variance was retained for the discrimant analysis and all of the discriminant functions were retained. In addition, the level of differentiation among the clusters inferred in the DAPC was evaluated in STRUCTURE 2.3.2.1, with D_{est}, and with AMOVA using the inferred cluster assignments as factors.

We also used the software STRUCTURE 2.3.2.1 to infer admixture between *C. fructivorum* and *C. rhexiae* and to assign two individuals previously determined to be *C. fructivorum* (Chapter 2) to a population. The data consisted of all of the *C. fructivorum* haplotypes, including *Coll886* and CBS124, and all of the *C. rhexiae* haplotypes. The analysis was then run at k=2 (five iterations) with the burnin set to 100,000 and post-burnin to 500,000.

The output from STRUCTURE of independent iterations for each value of *k* were combined using CLUMPP 1.1.2, and the graphical output was produced with DISTRUCT 1.1 (Jakobsson, Rosenberg, 2007; Rosenberg, 2004).

Demography

Estimates of allelic richness (A_{R}) and heterozygosity (H_{e}) have been proposed as genetic signatures of population expansion and population bottlenecks (Leberg, 2002; Luikart *et al.*, 1998a; Luikart *et al.*, 1999; Luikart *et al.*, 1998b; Maruyama, Fuerst, 1985; Stockwell *et al.*, 1996). During a bottleneck, the number of alleles in a population is reduced at a faster rate than is gene diversity (H_{e}), leading to an excess of gene diversity, relative to that expected at mutation-drift equilibrium (H_{eq}) (Nei *et al.*, 1975). We have thus compared H_{e} with H_{eq} to infer demographic processes in both New Jersey and Delaware, two populations where we sampled hierarchically. The expected gene diversity at equilibrium was calculated based on the number of alleles (*A*) under the assumption of a population at equilibrium and compared with H_e at each locus as implemented in the program Bottleneck 1.2.02 (Cornuet, Luikart, 1996; Luikart *et al.*, 1998a). In order to determine whether a significant number of loci show an excess of gene diversity we used the standardized differences test and the Wilcoxon sign-ranked test as implemented in Bottleneck. An alternative method of inferring demographic history was approached using allelic diversity estimates calculated with a rarefaction procedure in the program ADZE (Szpiech *et al.*, 2008).

We tested for a temporal change in allelic richness using population samples collected in 1994-1997 and those collected between 2008 and 2010 using the Wilcoxon signed-rank test on the mean allelic richness across loci per subsample. Using the same approach, we tested for the signature of a bottleneck that may have resulted from a founder event during the establishment of cranberry beds in Delaware by comparing with Wisconsin, the source of the host plants for this population.

Signatures of clonality and recombination

Colletotrichum species are known to reproduce asexually, particularly in agricultural environments where host homogeneity may contribute to epidemic phases of asexual population growth. However, many *Colletotrichum* species are also known to produce sexual structures (perithecia), including *C. fructivorum*. We calculated the index of association (I_A) and $\bar{r}_d$, a similar measure that attempts to remove the dependency of I_A on the number of loci, in Multilocus 1.3b (Agapow, Burt, 2001) to test for random mating. These tests are based on the variance of the number of loci for which two individuals are distinct compared to that expected in the absence of linkage disequilibrium. A significance test is performed by randomizing alleles as would be expected under random mating to determine if the observed variance is significantly different than the randomized variance. We also determined the number of genotypes and the effective number of genotypes, G_o (1/(sum of the squared genotype frequencies)) as a genotypic diversity estimate (Gomez, Carvalho, 2000; Stoddart, 1983). The effective number of genotypes can take a range of values from 1 to *n*, where *n* is the number of individuals genotyped. To determine if the genotypic structure of *C. fructivorum* departs from that expected as a result of sexual recombination, we assessed clonal diversity by comparing G_o from the same measure of a randomized dataset in GenoDive 2.0b22.

RESULTS

Population Differentiation and Diversity

Colletotrichum fructivorum - As previously described, clone-correction was applied at the population level in order to estimate levels of differentiation at various hierarchical levels. Identical haplotypes were not removed when they occurred in two different populations. However, clone assignment on the original regional dataset of 284 individuals identified 171 unique clonal haplotypes. Twenty-five of these haplotypes occurred across regions or subpopulations. Fifteen were distributed across subpopulations within a region, while 10 of the 25 were distributed among regions. Five of the haplotypes distributed among regions were shared between MA and NJ and two were shared between WABC, NJ, and MA. The remaining three were shared between NJ and WABC, NJ and WI, and DE and WI, respectively.

The analysis of molecular variance (AMOVA) of the clone corrected datasets indicated significant population structure among regions and subpopulations ($\phi_{PT} = 0.052$, $\phi'_{PT} = 0.085$, P

< 0.001) with 5% of the variation found within populations and several significant pairwise comparisons (Table 2 & Table 3). A similar result was found among the resampled regional datasets with a mean of 6% of the variation found within populations (Supplementary Table 1). The null hypothesis of panmixis could not be rejected among the other population comparisons of clone-corrected frequencies (Supplemental Tables 2-7). However, while the global ϕ_{PT} among subwatersheds within New Jersey was not significant, a pairwise comparison between the Rancocas Creek (North Branch) and the Mullica River (above the Batsto River) subwatersheds was significant ($\phi_{PT} = 0.024$, $\phi'_{PT} = 0.042$, $P < 0.047$), albeit with a very low level of differentiation (Table 4 & Table 5). Pairwise comparisons of population structure among populations in the regional dataset reveals low to moderate levels of population structure ($0.029 < \phi_{PT} > 0.220$, $0.047 < \phi'_{PT} > 0.341$, $P < 0.047$) in statistically significant comparisons), both among subpopulations within regions and between regions (Table 2 and Table 3). The highest levels of differentiation were consistently found between Delaware and all other subpopulations ($0.141 < \phi_{PT} > 0.220$, $0.235 < \phi'_{PT} > 0.341$, $P < 0.001$), with the exception of its separation from Wisconsin ($\phi_{PT} = 0.053$, $\phi'_{PT} = 0.084$, $P < 0.033$). Subpopulations within New Jersey, defined on the basis of irrigation networks, are differentiated at low levels ($0.039 < \phi_{PT} > 0.061$, $0.063 < \phi'_{PT} > 0.101$, $P < 0.024$). With the exception of Delaware, we found similar levels of differentiation among regions ($0.029 < \phi_{PT} > 0.083$, $0.047 < \phi'_{PT} > 0.134$, $P < 0.047$).

We also compared differentiation among populations with D_{est}, which does not depend on within population variation (Jost, 2008). Gerlach et al. (2010) and Pennings et al. (2011) have shown that *P*-values do not depend on the choice of criterion, thus the *P*-values associated with D_{est} are reported from the AMOVA permutation test. Due to the effect of the number of demes on estimates of D_{est} (Jost, 2008) each pairwise comparison was done independently, but all comparisons are presented together in Table 6. Estimates of D_{est} can be affected by differences in sample size among populations, so pairwise estimates were obtained by resampling on the original data to a standardized population size. These estimates were comparable to the point estimates from the original data and are presented in Table 7 for comparison. Consistent with both ϕ_{PT} and ϕ'_{PT}, D_{est} values indicate low to moderate levels of differentiation between Delaware and all other populations (barring Wisconsin) and the rank orders of divergence are similar for all three measures. However, given the wide confidence intervals around the mean for all pairwise comparisons, levels of differentiation between the other populations is very low to low ($0.012< D_{est} > 0.117$).

D_{est} can be affected by the variation in mutation rate among loci, therefore individual locus estimates were calculated for all datasets and compared with estimates of G_{ST} and unbiased estimates of subpopulation and total gene diversity (Tables 8-9; Supplementary Tables 8-13) (Meirmans, Hedrick, 2011; Ryman, Leimar, 2009). This illustrates the broad range in gene diversity among loci, the relatively low levels of genetic diversity in total, and the fact that much of the signature of differentiation among populations comes from just a few loci. In addition, as Meirmans and Hedrick (2011) have suggested, estimates of G_{ST} are higher in cases where subpopulation gene diversity is less than 0.5 and D_{est} is higher when subpopulation gene diversity is greater than 0.5 (such as L14F4 and LC1942; Table 8).

Comparisons between clone corrected datasets and data without clone correction were made to determine the impact of clone correction. It is common practice to restrict analyses to clone corrected datasets. However, for organisms that are known to reproduce clonally, clone correction may actually obscure the differentiation among populations and diminish the contribution of single haplotypes to the composition of populations. The results of AMOVA as well as single locus, pairwise and global estimates of D_{est} are presented in supplementary tables (Supplementary Tables 14-28). But we rely on estimates of differentiation from the clone corrected data because the sampling strategy in each regional population was not identical and therefore it is difficult to assess the reason for differences in the frequency of individual haplotypes in each population.

We genotyped 146 individuals with precise spatial information in New Jersey, however, removing identical haplotypes when they originated from the same sampling point [50 cm^2] reduced the dataset to 129 haplotypes. The correspondence between a genetic distance matrix based on the number of distinct alleles between two isolates and the geographic distance matrix among these collections was not significant in a mantel test ($R^2 = 0.003$, $P = 0.093$).

Colletotrichum rhexiae - *Colletotrichum rhexiae* had much lower levels of diversity at the loci used to characterize populations of *C. fructivorum.* Six of the nine genotyped loci were polymorphic among individuals of *C. rhexiae* with a total of 18 alleles compared with 84 alleles across 9 polymorphic loci in *C. fructivorum.* Only two loci were variable within New Jersey (L14F4: 4 alleles; LC2090: 2 alleles). Despite the low levels of diversity, differentiation measures were very high (($0.667 < (D_{\mathrm{est}} = 0.765) > 0.879$), P=0.001; ($0.845 < (G_{\mathrm{ST}} = 0.929) > 1.000$), P=0.001. The Maryland and Delaware collections were fixed for separate alleles at all loci while New Jersey shared alleles with both populations, though there were no shared haplotypes among populations. The pairwise estimates of D_{est} were 1.000 between Delaware and Maryland, 0.819 between Delaware and New Jersey, and 0.453 between New Jersey and Maryland.

Colletotrichum temperatum & *C. kahawae* - Both of the known isolates for *C. temperatum* were successfully genotyped at 7 of the 9 markers. There were missing data at a single locus for a single individual and two loci did not amplify. The two isolates differed at a single locus. A single allele was shared between three individuals of *C. kahawae* and a single individual of *C. temperatum*. Four isolates of *C. kahawae*, including the ex-type, representing two African populations (Cameroon and Kenya) were successfully genotyped at all nine loci for a total of 13 alleles. Three loci, LC2090, L10D10, and LC1942, were polymorphic. Alleles are shared between *C. kahawae* and *C. rhexiae* at three microsatellite loci (L10D10, LC2090, and LF9). There are no shared alleles between *C. kahawae* and *C. fructivorum.*

Principal Components Analysis & Allelic Rarefaction

Principal components analysis and rarefied private allelic richness estimates are presented together because they both provide an alternative to differentiation metrics for exploring genetic diversity among populations, particularly when sample sizes may not be large enough to be useful for statistical testing. While principal components analysis offers a qualitative visualization of the distribution of genetic data, estimates of allelic richness through rarefaction can offer both a quantitative and qualitative perspective on the genetic diversity within and among populations. PCA of each dataset was carried out separately to visualize the data with respect to the various sampling hierarchies. Each analysis is plotted with populations as factors, identified with unique color labels and a text label at the center of the inertia ellipse for each population. A barplot of the eigenvalues, indicating the amount of variance summarized by each principal component, is overlaid on each plot with the black bars representing the plotted axes.

Colletotrichum fructivorum - The PCA plot of the regional data illustrates the broad distribution of closely related haplotypes across all of the sampling regions and considerable overlap of the inertia ellipses among regions (Figure 1). Although slight, there is indication from the PCA plots of axes 1 and 2 and axes 2 and 3 that WI_2 (Wisconsin-subpopulation 2) and DE (Delaware) are somewhat isolated from the other populations based on location of the center of the inertia ellipses and the amount of overlap with the other populations (Figure 1A and 1C). However, comparison of all other datasets revealed little evidence of the influence of any of the biotic factors or landscape features, such as subwatersheds, on population structure (Figure 2 and Supplementary Figures 1-5).

Plotting the results of rarefied private allelic richness identified some trends that were not obvious, based on other approaches to characterize the structure among sampled populations. Private alleles, rarefied to a sample size equal to the number of stem haplotypes ($n = 10$) in the comparison of isolates from fruit and stem of cranberry, shows there is a higher number of private alleles among individuals from stem tissue (Figure 3A). In contrast, the private allelic richness of *C. fructivorum* isolates is much lower on *Rhexia virginica* than those isolated from *Vaccinium macrocarpon* when standardized to a sample size of 12, the number of haplotypes isolated from *R. virginica* (Figure 3B). A comparison of the rarefied private allelic richness between regional subpopulations shows declining private allelic richness as the number of sampled haplotypes increases in the Washington-British Columbia (WABC) and Delaware (DE) populations (Figure 4A and 4B), indicating a lack of compositional difference between these and other regional populations. Figure 4A illustrates the obscuring of differences in genetic diversity that has occurred by combining WI_1 and WI_2 due to the large reduction in sample size in WI_2 after clone correction. WI_2 has lower private allelic richness than WI_1 and is comparable to that of WABC and DE.

Colletotrichum fructivorum, C. rhexiae, C. kahawae, & *C. temperatum* - A PCA of all haplotypes from *C. fructivorum*, *C. rhexiae, C. kahawae*, and *C. temperatum* show a clear separation among species along axis 1 (Figure 5). The addition of two singleton isolates, geographically and temporally separated from all other isolates of *C. fructivorum*, fall between *C. fructivorum* and *C. rhexiae*, the two species with which they share alleles (Figure 6). Figure 7 illustrates the genetic differentiation found among populations of *C. rhexiae,* separating the New Jersey and Delaware populations along axis 1 and separating New Jersey and Maryland along axis 2.

Clustering and Discriminant Analysis of Principal Components

Colletotrichum fructivorum - We used two different clustering algorithms to uncover population structure in the regional data that may not be evident from other types of analyses. A Bayesian clustering algorithm that minimizes linkage disequilibrium and maximizes Hardy-Weinberg equilibrium was implemented to determine the optimal number of populations (*k*) that explains the data. Each isolate is assigned a probability of belonging to each population, and is interpreted as inferring admixture among populations. We used the method of Evanno et al. (2005), as implemented in STRUCTURE HARVESTER, to choose optimal values of *k* to evaluate in subsequent analyses. The three values of *k* with the largest Δ_k from the exploratory analyses in STRUCTURE were k=5 (Δ_k =5.9), k=9 (Δ_k =10.0), and k=11(Δ_k =16.7). An alternate approach using non-parametric *k*-means clustering and evaluation of increasing values of *k* with BIC indicated *k*=9 as the optimal number of haplotype clusters (Figure 8), with all but cluster 6 composed of isolates from multiple regions. Evaluation of these clusters using discriminant analysis of principal components (DAPC) shows four non-overlapping clusters of haplotypes, with the inertia ellipses of clusters 2 and 7 not overlapping with any adjacent groups (Figure 9).

The assignment probability of individuals to the optimal values of k determined in Structure Harvester were evaluated further with increased MCMC sampling in STRUCTURE and were plotted, ordering the individuals on the basis of the *a priori* regional population assignments (Figure 10: k=5, k=9, and k=11). Subtle levels of differentiation can be distinguished at k=5, with DE and MA being distinctive from all other *a priori* population assignments. There is little improvement with increasing *k* and the MA population becomes less distinctive. In concordance with other analyses, geographical origin is not indicative of population structure. High levels of gene flow between populations is suggested by the large proportion of admixed individuals per population.

Comparison of the Bayesian clustering in STRUCTURE with the clusters inferred from the non-parametric *k*-means, by ordering haplotypes in the bar plots on the basis of the cluster assignments, illustrates some congruence between the approaches. Bayesian clustering clearly distinguished cluster 2 at each level of k (Figure 10: **k*=5, **k*=7, and **k*=9), while cluster 7 was somewhat distinguishable. Cluster 2 is primarily composed of individuals from New Jersey, with a single individual from Wisconsin. The majority of haplotypes in cluster 7 originate from Delaware, but it also includes isolates from New Jersey and Wisconsin (subpopulation WI_2). The mean global measure of D_{est} among the clusters was 0.267 (0.261< 95% CI > 0.274) and more than 37% of the variation was found within clusters. This indicates differentiation among the clusters inferred from this analysis, but the cluster assignments do not seem to be congruent with any of the factors (biotic or abiotic) that we have used to define populations.

Colletotrichum rhexiae & *C. fructivorum* - A test for admixture between *C. fructivorum* and *C. rhexiae* revealed strong differentiation between the two species with very low levels of admixture (Figure 11). Only three individuals showed any level of admixture, *Coll886* (PA wild), CBS 124 (Shear 1922), and an isolate of *C. fructivorum* from *R. virginica*. *Coll886* was assigned to *C. fructivorum* on the basis of phylogenetic analysis (Chapter 2), however a significant proportion of its genome appears to be shared with *C. rhexiae*. CBS 124 was collected in the 1920s by C.L. Shear, presumably in New Jersey where he did much of his work, and shows a moderate level of admixture with *C. rhexiae*. A single isolate of *C. fructivorum* that was isolated from *Rhexia virginica* adjacent to a cultivated cranberry bed shares a single allele with isolates of *C. rhexiae*.

Demography

An excess of alleles is expected in a population that has been expanding in size for several generations (Maruyama, Fuerst, 1985), but heterozygosity deficiency can also result from a selective advantage or disadvantage conferred by a given allele. Therefore, inferences of demographic history should be based on the average of multiple loci within a population (Cornuet, Luikart, 1996). There is an excess of alleles (H_e deficiency) in 7 of 9 loci under the assumption that all alleles fit the IAM (each mutation produces a unique allele) and 8 of 9 under the assumption that all alleles fit the SMM (alleles evolve in a stepwise fashion and can be homoplasious) as illustrated by negative standard deviates for the New Jersey population (Table 10). Both the standardized differences test and the Wilcoxon sign-rank test (non-parametric) indicate a significant number of loci with an excess of alleles under both mutation models ($P<0.05$) and negative standard deviates (IAM: -3.458; SMM: -16.218). Both theoretical predictions and the analysis of empirical data from populations with a known demographic history (Cornuet, Luikart, 1996; Maruyama, Fuerst, 1985) suggest that our results indicate population expansion. This is concordant with expectations of an expansion in host population size concomitant with the expansion of the cranberry industry.

The cranberry growing regions representing the collection sites in this study have been well established for many years, with the exception of the Delaware population, which was only established within the last 15 years. The Delaware beds were originally established by transplanting cuttings from a nursery in Wisconsin, where *C. fructivorum* is present. It is suspected that *C. fructivorum* migrated to Delaware with plant material and represents a founder population, which is known to leave a signature of a bottleneck. A test of heterozygosity indicates six loci with a deficiency in H_e and two loci with excess H_e. Three of the four tests (standardized differences test and the Wilcoxon sign-rank test under each of the two mutation models) reject the hypothesis of heterozygosity excess and detect significant H_e deficiency ($P<0.05$), but the magnitude of deficiency is much lower than in the New Jersey population, as indicated by the standard deviates (Table 10: IAM: -1.121; SMM: -2.978) . The standardized differences test under the infinite allele model supports neither heterozygosity excess or deficiency, suggesting that the population is at mutation-drift equilibrium ($P < 0.13$). Evidence for a population bottleneck based on heterozygosity excess is expected to be detectable only for recent bottleneck events (Cornuet, Luikart, 1996). It may be that the there has been sufficient time for the population to remove any signature of a bottleneck with the low levels of heterozygosity deficiency indicating a transition from excess to deficiency as the population expands. It also may be that the small sample size in Delaware ($n = 12$) limits the power to reject the null hypothesis of a population at mutation-drift equilibrium.

It has been suggested that allelic richness is more sensitive in detecting the signature of a bottleneck than estimates of heterozygosity excess (Leberg, 2002).

As an alternate means of inferring the demographic history of *C. fructivorum*, we estimated the mean allelic richness (A_R) across loci through rarefaction (Leberg, 2002; Szpiech *et al.*, 2008). To determine whether richness has declined over time within the New Jersey population, we compared the allelic richness of the clone corrected 1994-1997 population sample ($n = 24$) with contemporary collections, clone-corrected by sample point and tested the null hypothesis of no difference with a Wilcoxon signed-rank test (Spencer *et al.*, 2000). The allelic richness of the contemporary population sample was significantly greater than that of the collections from 1994-1997 ($A_R = 3.86$ and $A_R = 3.66$, $P < 0.01$, respectively) (Figure 12). The difference in mean allelic richness is small between the two population samples, so concluding that the population has been expanding since 1994-1997 is tenuous. However, we can reject the hypothesis that there has been a population bottleneck over the last 15 years. Following the same rationale, we tested the Delaware ($A_R = 2.42$) population for evidence of a reduction in allelic richness, when compared with the Wisconsin population ($A_R = 3.00$, $P < 0.01$), from which it was likely derived, suggesting that there has been some loss of alleles (Figure 13). This reduction in allelic richness may have resulted from a founder event at the time of planting, known to have occurred within the last 15 years.

Reproductive system

We limited our inferences on modes of reproduction to the New Jersey population of *C. fructivorum* because this is the only population for which we have employed an appropriate sampling strategy and have large enough sample sizes to infer outcrossing. The same dataset that was used for the mantel test, clone-corrected by sample site, was used to infer reproductive mode. The fraction of clones distributed among these spatially separated haplotypes was 21.7%

and the probability of randomly selecting two unique haplotypes was 99.1% (Table 11). Multiple individuals were genotyped from 37 of the 103 sample points distributed across New Jersey and 35% of these contained identical haplotypes. We could not reject the null hypothesis of random mating based on the clonal diversity estimated from the observed number of unique haplotypes (n_{obs}=101) compared with that expected by randomizing alleles over individuals (n_{exp}=106.64, P=0.086, 999 randomizations) (Table 11; Figure 14). However, the null hypothesis of linkage equilibrium as expected under random mating was rejected by both indices of association, I_A and $\bar{r}_d$ (I_A=0.159, $P<0.05$; $\bar{r}_d$=0.021, $P<0.05$) (Table 11). To further explore the possibility that there is a random association of alleles among loci, we removed all identical haplotypes from the dataset resulting in 101 haplotypes (as opposed to only those within a single sample point). The null hypothesis of linkage equilibrium among loci within this group of haplotypes could not be rejected by both indices of association, I_A and $\bar{r}_d$ (I_A=-0.028, P=0.648; $\bar{r}_d$=-0.004, P=0.648) (Table 11; Figure 14). The disparity among these tests highlights the diversity of reproductive strategies that are likely implemented within populations of *C. fructivorum.* We can not infer the magnitude of clonal reproduction without explicit sampling to address that question. However, the results of the genotypic diversity and linkage disequilibrium tests at different levels of clone-correction do suggest that clonality is an important component of the diversity within populations of *C. fructivorum*, but also indicate that outcrossing has at least played a historical role in reassorting alleles.

DISCUSSION

We used a population genetics approach and sampling at several hierarchical levels to characterize the distribution of genetic diversity in *Colletotrichum fructivorum*, an important pathogen of the American cranberry (*Vaccinium macrocarpon* Aiton). With these data we are able to make inferences regarding the influence of biotic and abiotic factors on population divergence, the demographic and dispersal history, and the mode of reproduction employed in field populations. Comparing the distribution of genetic data within *C. fructivorum* to the distribution of genotypes within *C. rhexiae* provides insight into the influence of anthropogenic factors in shaping pathogen populations. Multilocus genotyping of both species in parapatric populations indicates nearly complete genetic isolation, suggesting that ecological divergence maintains the species boundary. The samples for *C. kahawae* and *C. temperatum* were limited; more markers and expanded sampling will be needed, but the data available provide support for the recognition of *C. kahawae* and *C. temperatum* as distinct species.

Evidence for human-mediated dispersal of *Colletotrichum fructivorum*

The distribution of genetic diversity within *Colletotrichum fructivorum* is indicative of a pathogen with relatively high levels of gene flow among local and regional populations, with no significant differentiation on the basis of host species or cultivar, or host organ. Pairwise estimates of population structure among most regional populations reveal low levels of differentiation. Similarly, the principal components analysis, Bayesian and non-parametric *k*-means clustering, and the discriminant analysis of principal components reveal high levels of gene flow among regional populations. Given the geographic isolation of these populations and the fact that *Colletotrichum* is thought to be water-dispersed, seemingly not capable of long distance dissemination, this lack of structure is indicative of other means of dispersal. Cranberry beds are established from cuttings rather than from seed, and are frequently transferred among growing regions to establish new beds. Thus, it is likely that the lack of differentiation among growing regions is the result of wide-ranging and human-mediated dispersal via the colonization of transplants.

The genetic differentiation observed between Delaware and all other regional populations, with the exclusion of Wisconsin offers support for the hypothesis of human-mediated dispersal. The farm in Delaware from which these isolates were collected was recently established with plant material originating in Wisconsin, suggesting migration from Wisconsin to Delaware as endophytes/latent pathogens colonizing stems. The transfer was most likely associated with a founder event, as the cranberry growing area in Delaware is much smaller than that of Wisconsin.

Both the Bayesian and non-parametric *k-means* cluster analysis evaluated with the discriminant analysis of principal components differentiated a population composed of individuals primarily from Delaware and some isolates from Wisconsin (WI_2). Similarly, the principal components analysis of the regional data before clone-correction shows subpopulation WI_2 grouping with the Delaware population. This suggests a more specific origin for the isolates from Delaware. The collective evidence suggests that *C. fructivorum* in WI_2 and Delaware share the same source population or that plant material was moved from WI_2 to Delaware. Given the bulk sampling approach employed for Wisconsin and the large fraction of clones (CF: Table 1), particularly in WI_2, hierarchical sampling across Wisconsin is needed to test this hypothesis.

A pattern of genetic diversity congruent with water dispersal should fit an isolation-by-distance model in a landscape with multiple isolated irrigation networks. However, the lack of correlation between geographic distance and genetic distance across New Jersey, with collections spanning several irrigation networks, subwatersheds and two major watersheds, indicates an alternate mode of dispersal. This is also supported by the ranking in levels of differentiation among subpopulations within New Jersey and distant geographical regions. For example, genetic differentiation is no greater between Washington and British Columbia (WABC) than it is between irrigation networks within New Jersey (e.g. NJ_3 vs. NJ_4). Given the movement of plant material among geographical regions in the establishment of new cranberry beds, endophytic colonization of stems prior to transport seems the most likely vector for the transfer of haplotypes among regions and subpopulations.

There appears to be sufficient power to distinguish among unique genets with the microsatellite markers used in this study as indicated by the high genotypic diversity estimates. Nonetheless, there are a several clonal haplotypes distributed among regional subpopulations and across regions. The inference of clonality in this case may be the result of homoplasic mutation as might be expected for highly mutable microsatellite markers evolving under a stepwise mutation model, but it could also be interpreted as evidence for the anthropogenic transport of clonal haplotypes over long distances. On balance, it appears that clonal haplotypes can migrate over long distances.

While biotic factors such as host species, host cultivar, and host organ can not be implicated as factors in the differentiation of *C. fructivorum* populations, the level of sampling has to be considered for these comparisons. Permutation tests indicate there is no significant differentiation among isolates originating from distinct host species or distinct host organs, but the small sample sizes would make it difficult to detect significant differentiation, even if it were present. Private allelic richness estimates indicate, with increased sampling, differentiation may be detected between fruit and stem isolates, of which the latter have a higher number of private alleles. This supports the idea that not only is the species diversity of *Colletotrichum* higher in stem tissue (Chapter 2), the diversity of haplotypes within *C. fructivorum* may also be higher in stem tissue. Similarly, the higher private allelic richness in the population on cranberry suggests that the isolates drawn from *R. virginica* may simply be a subset of haplotypes from the broader source population. There is, however, no evidence that *R. virginica* serves as a reservoir of *C. fructivorum* for subsequent infection of cranberry.

Comparing the distribution of genetic diversity of *C. fructivorum* with that of *C. rhexiae* provides additional insight into the role of human-mediated dispersal. The microsatellite markers used in this study were developed for *C. fructivorum*. While they did cross amplify with *C. rhexiae*, they were much less variable in this species. As others have reported difficulty in the cross-amplification of microsatellite markers, outside of the target species for which they were developed, it may be that the diversity among species is not directly comparable (Peakall *et al.*, 1998; Roa *et al.*, 2000). However, the levels of polymorphism among these loci is higher in New Jersey where the geographic range of the host population (*Rhexia virginica*) is highest, so the variation may be limited by the small population size of the host, likely limiting the population size of *C. rhexiae.* Despite the low levels of variation, *C. rhexiae* appears to be highly differentiated geographically, with alternate haplotypes fixed in different populations. *Rhexia virginica*, both in wild populations and as a weed in agricultural populations, is colonized primarily by *C. rhexiae*. *Rhexia virginica* is not utilized for agricultural purposes; plant material is not anthropogenically transferred among regions. These populations are very likely more structured than *V. macrocarpon*, an agricultural crop plant. Investigating the host genetic structure within *Rhexia virginica* might shed some light on the causes of genetic structure within its pathogen, *C. rhexiae.* The contrast between the high levels of population structure within *C. rhexiae* and the low levels of population structure within *C. fructivorum* further implicates human agricultural practices as an important means of dispersal for *C. fructivorum.*

Reproductive mode of *Colleotrichum fructivorum*

Colletotrichum fructivorum is known to reproduce asexually. We have also confirmed through laboratory study that it is capable of homothallic sexual reproduction. The genetic outcome of these two different modes of reproduction is identical in haploid organisms, so we refer to either homothallic sexual or asexual reproduction as ‘clonal’ reproduction. Populations dominated by clonal reproduction exhibit a paucity of recombinant haplotypes and a non-random association between independent genetic markers (Milgroom, 1996). The rejection of random mating in two of the four tests for multilocus linkage disequilibrium, suggests that clonality plays an important role in the life history of *C. fructivorum*, inference that is further supported by the broad distribution of clonal haplotypes. However, high levels of genotypic diversity and distinct genotypes being found within 65% of 50 cm^2 sample points where more than a single individual was genotyped, suggest that outcrossing plays a role in shuffling alleles within the population. It is worthy of note, however, that in *C. fructivorum* 70% of the haplotypes differ by only a single allele. This illustrates the fact that populations of *C. fructivorum* may consist of several clonal lineages, but with new haplotypes arising through mutation rather than via sexual recombination. Taken as a whole, we propose that *C. fructivorum* employs a mixed mating system in field situations, suggesting that it has a high evolutionary potential to adapt to changes in selective pressure and amplify adaptive genotypes through clonal reproduction (McDonald, Linde, 2002b).

Ecological Speciation and Sex in the Evolution of *Colletotrichum fructivorum*

Ecological speciation via host shift has been proposed for several species of fungi, e.g. (Couch *et al.*, 2005; Gladieux *et al.*, 2011; Zaffarano *et al.*, 2008), including most recently for

Colletotrichum kahawae (Silva *et al.*, 2012a). An alternate mode of ecological speciation through episodic selection resulting from an ecosystem change has been suggested for others (Crouch *et al.*, 2009b; MacKenzie *et al.*, 2007). There are several inherent features of fungi that enable rapid ecological speciation through episodic selection including haploidy, short generation time, the propensity for asexual reproduction, large numbers of offspring (asexual spores), and within-host mating (Brasier, 1995; Giraud *et al.*, 2010). The sudden change from a native, heterogeneous ecosystem to a monocultural ecosystem, as happens in the establishment of agricultural crop systems, provides suitable conditions for episodic selection (Kareiva *et al.*, 2007; Stukenbrock, McDonald, 2008). There are several examples of such habitat changes leading to the emergence of new pathogens [reviewed in (Stukenbrock, McDonald, 2008)]. The genetic uniformity of *V. macrocarpon* in agricultural environments, and agricultural management practices that promote the spread of propagules between infected and uninfected plants within beds, promotes the habitat-specialization that has emerged in *C. fructivorum*.

We propose two hypothetical pathways for the emergence of *Colletotrichum fructivorum* as a widespread pathogen of cranberry. Both models are suggestive of ecological speciation and divergence through episodic selection and amplification of adaptive genotypes through 'clonal' reproduction. However, one pathway implies speciation in sympatry, while the other implies allopatric speciation and expansion to the point that sister species are now parapatric. Under the first scenario, the ancestor of *C. fructivorum* and *C. rhexiae* was present in a heterogeneous environment and capable of infecting both host species, *V. macrocarpon* and *R. virginica*, although preferentially colonizing *R. virginica,* much like *C. rhexiae* seems to do today. With the emergence of cranberry cultivation in the mid-1800s, the ecosystem changed from a heterogeneous environment with a patchy susceptible host distribution to a homogenous monoculture of low host genetic diversity. Sexual reproduction in the ancestral population generated the diversity needed for selective forces to act. But given an adaptive genotype capable of efficiently colonizing cranberry, the ability to increase the frequency of adaptive alleles through clonal reproduction, and management practices that promote dispersal within and among planted beds, a new pathogen population emerged that is well-adapted for colonizing cranberry. The differential fitness between the two species would maintain the divergence. The relatively recent emergence of *C. fructivorum* as a pathogen within an agricultural ecosystem is supported by its absence in wild populations and its sister relationship with *C. rhexiae* (Chapter 2), a species with an overlapping distribution and a shared host range. Further support comes from the multilocus genotype of an individual (CBS 124) deposited in 1922 by C.L. Shear as *C. rufomaculans* var. *vaccinii* (now *C. fructivorum*). While the origin of this isolate is unknown, C.L. Shear was a cranberry pathologist stationed in New Jersey where he described *C. rufomaculans* var. *vaccinii* in 1907, thus this strain was probably isolated from cranberry in New Jersey. Multilocus genotyping of this individual indicates that it shares a proportion of its genome with *C. rhexiae* and a portion with *C. fructivorum*, but also carries several private alleles. If this can be taken to be representative of the makeup of the population in the early 1900s, this would suggest that *C. fructivorum* was still in the process of diverging several decades after the start of cranberry cultivation. Of course, this could also represent an anomalous individual that is a hybrid between two species that were already diverged. Tests of cross-species pathogenicity on distinct hosts could provide an opportunity to test this hypothesis and evaluate the role of ecological adaptation and influence of selection against migrants in maintaining distinct gene pools.

In an alternate scenario, *C. rhexiae* and *C. fructivorum* represent two parapatric species that diverged in allopatry. This implies that the ancestral population was widespread and a host shift through ecological speciation may have occurred twice in separate environments. This is supported by the presence of an individual from a wild population of *V. oxycoccos* (a close relative of *V. macrocarpon*) that shares a similar proportion of its genome with both *C. fructivorum* and *C. rhexiae.* Additional and intensive sampling outside of agricultural ecosystems and from additional host species might identify such a population. These proposals are models for the development of additional studies to evaluate the reasons for divergence between *C. fructivorum* and *C. rhexiae*.

The results of this study provide a foundation upon which mycologists, plant pathologists, and plant breeders can develop a better understanding of the biotic and abiotic factors that drive speciation, lead to the emergence of new pathogens, and affect host susceptibility. Using the tools we have built to characterize this pathogen system, plant pathologists and plant breeders can work together to develop a more effective strategy for combating fruit-rot in commercial cranberry populations. This includes the characterization of gene-gene and genotype-genotype interactions between host and pathogen and expanded hierarchical sampling to better understand the role of management practices on the distribution of genetic diversity. Furthermore, the ability to identify genets using the markers validated in this study may allow mycologists to investigate the mating system of these species in detail and test the potential for hybridization among species.

CONCLUSIONS

The work presented here represents a significant advance in understanding the biology of *Colletotrichum gloeosporioides s.l.* and the factors that influence species diversity and divergence within the complex. It is evident from the phylogenetic study, sampling across a broad geographical range from a single host species and across multiple spatial scales and host species within a narrow regional distribution, that a composite of factors including host and habitat can be indicative of lineage identity within the complex. However, it appears that host specificity is not prevalent. Furthermore, both pathogenic and non-pathogenic (endophytic) phenotypes can be conspecific, indicating that pathogenic species can persist as latent pathogens or endophytes in alternate host organs. A comparative population genetic study of two closely related species, *Colletotrichum fructivorum* and *C. rhexiae*, implicates anthropogenic factors as important forces in the evolution of *Colletotrichum*. Human-mediated dispersal maintains the genetic cohesiveness of *C. fructivorum* populations across a broad geographical range, while it appears that the emergence of cranberry cultivation in eastern North America may have led to the divergence between *C. fructivorum* and *C. rhexiae*. These findings have important implications for disease management practices, plant and pathogen quarantine regulations, future systematic studies of *Colletotrichum gloeosporioides s.l.,* and efforts to develop a better understanding of the biology of the species complex as a whole.

The determination that several sympatric, cryptic lineages of *Colletotrichum gloeosporioides s.l.* are associated with a single crop species highlights the importance of carrying out a detailed phylogenetic study before attempting to characterize “populations” of pathogenic or endophytic fungi. This also illustrates the need for additional studies that employ both vertical (among host organs) and horizontal (among host species) sampling approaches within a clearly delimited regional area to uncover the wide range of diversity within the complex. The range of lifestyles, from endophytic to pathogenic, exhibited by conspecific isolates should be a note of caution when characterizing communities of pathogenic or endophytic fungi. Studies by plant pathologists or fungal endophyte ecologists that focus solely on a particular portion of the continuum between pathogen and endophyte, such as choosing to focus on isolates only from tissue where disease pathology is expressed, may be mischaracterizing these communities and misunderstanding the ecology and life history of the component species. A more complete understanding of the life history has important implications for developing effective disease management protocols and restricting the movement of pathogens. For example, the understanding that *C. fructivorum* is being transported long distances as a stem endophyte suggests that cultivation practices limiting the colonization of plant material used for the establishment of new farms may limit the diversity and abundance of pathogenic species. Understanding the geographical and host distribution of pathogenic *Colletotrichum* species also provides the information needed to develop more targeted quarantine regulations that will limit the movement of pathogens into new areas.

The detailed sampling strategy employed in this study also provides insight into the factors that shape diversity within the species complex that can be capitalized upon to continue to unravel the complex history of *Colletotrichum gloeosporioides s.l.* The high levels of diversity within the complex that can be recovered from a single host species highlights the importance of sampling multiple plant parts to uncover the diversity at the level of host species. However, sampling across host species and among wild and cultivated habitats sheds light on the importance of both host and habitat in shaping diversity within the complex. A sampling strategy that accounts for these factors is not only likely to recover a greater proportion of the diversity within the species complex, but may provide an opportunity to place pathogenic species in a broader evolutionary context and develop a more refined understanding of the emergence of pathogenic lineages.

We have begun to develop a more detailed picture of the emergence and spread of one of the most prominent cranberry pathogens, *Colletotrichum fructivorum*, and its sister species, *Colletotrichum rhexiae*. However, additional studies are needed to understand the evolutionary history of these species and capitalize on the foundational work provided by the research presented here. Broader geographical and host sampling, comparative pathological studies, and comparative genomics are necessary to understand the evolutionary pathway that led to the divergence of *C. fructivorum*, the differential fitness that maintains genetic isolation between two sympatric species, *C. fructivorum* and *C. rhexiae*, and the genetic modifications that accompany host shifts and lifestyle changes in plant-associated fungi.

REFERENCES

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

Alberto F (2009) MsatAllele_1.0: An R package to visualize the binning of microsatellite alleles. *Journal of Heredity*, **1003**, 394-397.

Antonovics J (2006) Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, **97**, 33-37.

Brasier CM (1995) Episodic selection as a force in fungal microevolution, with special reference to clonal speciation and hybrid introgression. *Canadian Journal of Botany*, **73**, S1213-S1221.

Cole RA, Lynn J (1996) Spatial and temporal population structure of *Brevicoryne brassicae*. *Entomologia Experimentalis et Applicata*, **78**, 121-127.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.

Couch BC, Fudal I, Lebrun M-H*, et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, **170**, 613-630.

Crouch JA, Tredway LP, Clarke BP, Hillman BI (2009) Phylogenetic and population genetic divergence correspond with habitat for the pathogen *Colletotrichum cereale* and allied taxa across diverse grass communities. *Molecular Ecology*, **18**, 123-135.

Dean R, Kan JALV, Pretorius ZA*, et al.* (2012) The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology*, **13**, 414-430.

Dettman JR, Anderson JB, Kohn LM (2008) Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus *Neurospora*. *BMC Evolutionary Biology*, **8**, 35.

Dettman JR, Sirjusingh C, Kohn LM, Anderson JB (2007) Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature*, **447**, 585-588.

Dilmaghani A, Gladieux P, Gout L*, et al.* (2012) Migration patterns and changes in population biology associated with worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. *Molecular Ecology*, **21**, 2519-2533.

Du M, Schardl CL, Nuckles EM, Vaillancourt LJ (2005) Using mating-type gene sequences for improved phylogenetic resolution of Colletotrichum species complexes. *Mycologia*, **97**, 641-658.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Eck P (1990) *The American Cranberry* Rutgers University Press, New Brunswick, NJ.

Ellis JB, Everhart BM (1894) New Species of Fungi from Various Localities *Proceedings of the Academy of Natural Sciences of Philadelphia*, **46**, 322-386.

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.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587.

Farr DF, Rossman AY, Palm ME, McCray EB (2011) Fungal Databases. http://nt.ars-grin.gov/fungaldatabases/ Systematic Mycology and Microbiology Laboratory, ARS, USDA

Freeman S, Katan T, Shabi E (1996) Characterization of Colletotrichum gloeosporioides isolates from avocado and almond fruits with molecular and pathogenicity tests. *Applied and Environmental Microbiology*, **62**, 1014-1020.

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

Gladieux P, Guerin F, Giraud T*, et al.* (2011) Emergence of fungal pathogens by ecological speciation: importance of the reduced viability of imigrants. *Molecular Ecology*, **20**, 4521-4532.

Gomez A, Carvalho GR (2000) Sex, parthenogenesis and genetic structure of rotifers: microsatellite analysis of contemporary and resting egg bank populations. *Molecular Ecology*, **9**, 203-214.

González E, Sutton TB, Correll JC (2006) Clarification of the etiology of Glomerella Leaf Spot and Bitter Rot of apple caused by Colletotrichum spp. based on morphology and genetic, molecular, and pathogenicity tests. *Phytopathology*, **96**, 982-992.

Guozhong L, Cannon PF, Reid A, Simmons CM (2004) Diversity and molecular relationships of endophytic *Colletotrichum* isolates from the Iwokrama Forest Reserve, Guyana. *Mycological Research*, **108**, 53-63.

Halsted B (1889) Some fungus diseases of the cranberry, pp. 1-40. New Jersey Agricultural College Experiment Station.

Hang VTT, Holmes EC, Veasna D*, et al.* (2010) Emergence of the Asian 1 Genotype of Dengue Virus Serotype 2 in Viet Nam: In Vivo Fitness Advantage and Lineage Replacement in South-East Asia. *PLOS Neglected Tropical Diseases*, **4**, e757.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633-1638.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, **290**, 516-518.

Jakobsson M, Rosenberg N (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801-1806.

Johnson-Cicalese J, Vorsa N, Polashock J (2009) Breeding for fruit rot resistance in *Vaccinium macrocarpon.* In: Proceedings of the IXth International Symposium on *Vaccinium*.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Jombart T, Pontier D, Dufour AB (2009) Genetic markers in the playground of multivariate analysis. *Heredity*, **102**.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015-4026.

Jost L (2009) Reply: D vs G'st: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology*, **18**, 2088-2091.

Kareiva P, Watts S, McDonald R, Boucher T (2007) Domesticated Nature: Shaping Landscape and Ecosystems for Human Welfare. *Science*, **316**, 1866-1869.

Kaufmann PJ, Weidemann GJ (1996) Isozyme analysis of *Colletotrichum gloeosporioides* from five host genera. *Plant Disease*, **80**, 1289-1293.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of poulation differentiation. *Conservation Genetics* **7**, 295-302.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, **11**, 2445-2449.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998a) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238-247.

Luikart G, Cornuet JM, Allendorf FW (1999) Temporal changes in allele frequencies provide estimates of population bottleneck size. *Conservation Biology*, **13**.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998b) Usefullness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**.

MacKenzie SJ, Seijo TE, Legard DE, Timmer LW, Peres NA (2007) Selection for pathogenicity to strawberry in populations of *Colletotrichum gloeosporioides* from native plants. *Phytopathology*, **97**, 1130-1140.

Maruyama T, Fuerst PA (1985) Population bottlenecks and nonequilibrium models in population genetics. II. Number of Alleles in a small population that was formed by a recent bottleneck. *Genetics*, **111**, 675-689.

McDonald BA, Linde C (2002a) The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica*, **124**, 163-180.

McDonald BA, Linde CC (2002b) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349-379.

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differntiation measure. *Evolution*, **60**, 2399-2402.

Meirmans PG, Hedrick PW (2011) Assessing population structure: Fst and related measures. *Molecular Ecology Resources*, **11**, 5-18.

Meirmans PG, Tienderen PHV (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity in asexual organsims. *Molecular Ecology Notes*, **4**, 792-794.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, **34**, 457-477

Nei M, Mauyama T, Chakraboty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.

Nicholson RL, Moraes WBC (1980) Survival of *Colletotrichum graminicola*: importance of the spore matrix. *Phytopathology*, **70**, 255-261.

Oudemans PV, Caruso FL, Stretch AW (1998) Cranberry fruit rot in the Northeast: A complex disease. *Plant Disease*, **82**, 1176-1184.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (Glycine max) simple sequence repeats (SSRs) within the genus and other legume genera: implications for transferability of SSRs in plants. *Molecular Biology and Evolution*, **15**, 1275-1287.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetics software for teaching and research. *Molecular Ecology Resources*, **6**, 288-295.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

Ramakrishnan U, Hadly EA, Mountain JL (2005) Detecting past population bottlenecks using temporal genetic data. *Molecular Ecology*, **14**, 2915-2922.

Riccioni G, Landi M, Ferrara G*, et al.* (2010) Spatio-temporal population structuring and genetic diversity retention in depleted Atlantic Bluefin tuna of the Mediterranean Sea. *Proceedings of the National Academy of Sciences*, **107**, 2102-2107.

Roa AC, Chavarriaga-Aguirre P, Duque MC*, et al.* (2000) Cross-species amplification of Cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. *American Journal of Botany*, **87**, 1647-1655.

Rodriguez RJ, Redman RS (2000) *Colletotrichum* as a model system for defining the genetic basis of fungal symbiont life styles. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 114-130. American Phytopathological Society, St. Paul, Minnesota.

Rojas EI, Rehner SA, Samuels GJ*, et al.* (2010) Colletotrichum gloeosporioides s.l. associated with Theobroma cacao and other plants in Panama´: multilocus phylogenies distinguish host-associated pathogens from asymptomatic endophytes. *Mycologia*, **102**, 1318-1338.

Rosenberg N (2004) DISTRUCT: a program for the grpahical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Ryman N, Leimar O (2009) Gst is still a useful measure of genetic differentiation - a comment on Jost's D. *Molecular Ecology*, **18**, 2084-2087.

Shear CL (1907) New species of fungi. *The Bulletin of the Torrey Botanical Club*, **34**, 305-317.

Silva DN, Talhinas P, Cai L*, et al.* (2012a) Host-jump drives rapid and recent ecological speciation of the emergent fungal pathogen *Colletotrichum kahawae*. *Molecular Ecology*, **21**, 2655-2670.

Silva DN, Talhinas P, Várzea V*, et al.* (2012b) Application of the Apn2/MAT locus to improve the systematics of the Colletotrichum gloeosporioides complex: an example from coffee (Coffea spp.). *Mycologia*, **104**, 396-409.

Silvertown J, Servaes C, Biss P, Macleod D (2005) Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity*, **95**, 198-205.

Spencer CC, Neigel JE, Leberg PL (2000) Experimental evaluation of the usefullness of microsatellite DNA for detecting demographic bottlenecks. *Molecular Ecology*, **9**, 1517-1528.

Stiles CM, Oudemans PV (1997) Cranberry fruit rot pathogens isolated from various cranberry plant tissues over two growing seasons. *Phytopathology*, **87**, S94.

Stiles CM, Oudemans PV (1999) Distribution of cranberry fruit-rotting fungi in New Jersey and evidence for nonspecific host resistance. *Phytopathology*, **89**, 218-225.

Stockwell CA, Mulvey M, Vinyard GL (1996) Translocations and the preservation of allelic diversity. *Conservation Biology*, **10**, 1133-1141.

Stoddart JA (1983) A genotypic diversity estimate. *Journal of Heredity*, **74**, 489-490.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75-100.

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498-2504.

USDA-CSREES (2001) Crop Profile for Cranberries in Massachusetts.

USDA, NRCS (2008) The PLANTS Database. National Plant Data Center, Baton Rouge, LA

Viard F, Justy F, Jame P (1997) Population dynamisc inferred from temporal variation at microsatellite loci in the selfing snail *Bulinus truncatus*. *Genetics*, **146**, 973-982.

Whitlock M (2011) G'st and D do not replace Fst. *Molecular Ecology*, **20**, 1083-1091.

Zaffarano PL, McDonald BA, Linde CC (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution*, **62**, 1418-1436.

Zhu P (1999) *Genetic diversity and systematics of Colletotrichum species on cranberry (Vaccinium macrocarpon)* Ph.D., Rutgers University.

Zhu P, Oudemans PV (2000) A long terminal repeat retrotransposon *Cgret* from the phytopathogenic fungus *Colletotrichum gloeosporioides* on cranberry. *Current Genetics*, **38**, 241-247.

Table 1: Summary of the diversity within regional populations of *Colletotrichum fructivorum*

Region/Subpop.	ID	n	n_h	CF	G	MF	H_e	$*H_e$	A_R	P_{AR}	$*A_R$	$*P_{AR}$
Delaware	DE	21	16	0.238	0.971	3	0.303 ± 0.053	0.341 ± 0.061	2.404 ± 0.263	0.133 ± 0.092	2.365 ± 0.249	0.133 ± 0.091
Massachusetts	MA	34	25	0.265	0.970	5	0.321 ± 0.092	0.367 ± 0.095	2.810 ± 0.519	0.346 ± 0.126	2.820 ± 0.515	0.369 ± 0.132
New Jersey 1	NJ_1	33	21	0.364	0.922	9	0.313 ± 0.072	0.399 ± 0.084	2.674 ± 0.388	0.219 ± 0.112	2.746 ± 0.397	0.191 ± 0.095
New Jersey 2	NJ_2	29	22	0.241	0.975	4	0.314 ± 0.091	0.358 ± 0.094	2.786 ± 0.57	0.248 ± 0.136	2.727 ± 0.532	0.246 ± 0.135
New Jersey 3	NJ_3	30	22	0.267	0.979	3	0.425 ± 0.087	0.423 ± 0.088	3.077 ± 0.541	0.212 ± 0.155	3.002 ± 0.538	0.224 ± 0.165
New Jersey 4	NJ_4	16	14	0.125	0.983	2	0.352 ± 0.081	0.371 ± 0.085	2.778 ± 0.494	0.379 ± 0.182	2.617 ± 0.440	0.329 ± 0.158
New Jersey 5	NJ_5	35	31	0.114	0.992	3	0.413 ± 0.081	0.438 ± 0.084	3.569 ± 0.636	0.450 ± 0.186	3.285 ± 0.563	0.39 ± 0.161
New Jersey 6	NJ_6	33	28	0.152	0.989	3	0.383 ± 0.104	0.411 ± 0.109	3.481 ± 0.787	0.545 ± 0.238	3.240 ± 0.697	0.423 ± 0.185
Washington-British Columbia	WABC	16	12	0.250	0.950	3	0.356 ± 0.081	0.401 ± 0.082	3.000 ± 0.553	0.187 ± 0.111	3.000 ± 0.553	0.193 ± 0.114
Wisconsin 1	WI_1	20	8	0.600	0.742	10	0.268 ± 0.066	0.403 ± 0.068	2.418 ± 0.337	0.226 ± 0.118	2.889 ± 0.410	0.397 ± 0.233
Wisconsin 2	WI_2	17	5	0.706	0.507	12	0.141 ± 0.049	#	1.647 ± 0.225	0.120 ± 0.120	#	#

ID, Population identifier; n, sample size; n_{h}, number of unique multilocus haplotypes; CF, clonal fraction; G, genotypic diversity; MF, frequency of the most frequent multilocus haplotype; H_{e}, mean unbiased gene diversity across loci (± standard error); A_{R}, mean allelic richness across loci (± standard error); P_{AR}, mean private allelic richness across loci (± standard error); *, statistic calculated from clone-corrected dataset; #, statistic could not be calculated because clone correction reduced the sample size in WI2 such that WI1 and WI2 were combined for the clone-corrected data

Table 2: ϕ_{PT} values of regional population structure. ϕ_{PT} values are below the diagonal and p-values based on 999 permutations above the diagonal. Significant ϕ_{PT} values shaded in gray. Refer to Table 1 for population IDs.

	DE	**MA**	**NJ$_1$**	**NJ$_2$**	**NJ$_3$**	**NJ$_4$**	**NJ$_5$**	**NJ$_6$**	**WABC**	**WI**	
DE	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.033	**DE**
MA	0.220	0.000	0.099	0.249	0.011	0.007	0.057	0.034	0.136	0.003	**MA**
NJ$_1$	0.168	0.021	0.000	0.441	0.160	0.259	0.444	0.310	0.213	0.195	**NJ$_1$**
NJ$_2$	0.171	0.009	0.000	0.000	0.021	0.240	0.424	0.017	0.389	0.190	**NJ$_2$**
NJ$_3$	0.181	0.044	0.018	0.052	0.000	0.024	0.010	0.099	0.047	0.014	**NJ$_3$**
NJ$_4$	0.215	0.079	0.012	0.016	0.055	0.000	0.139	0.009	0.060	0.174	**NJ$_4$**
NJ$_5$	0.141	0.020	0.000	0.000	0.042	0.018	0.000	0.094	0.258	0.307	**NJ$_5$**
NJ$_6$	0.195	0.029	0.006	0.039	0.020	0.061	0.014	0.000	0.042	0.011	**NJ$_6$**
WABC	0.184	0.024	0.018	0.002	0.051	0.050	0.009	0.042	0.000	0.039	**WABC**
WI	0.053	0.083	0.019	0.017	0.072	0.028	0.006	0.060	0.063	0.000	**WI**
	DE	**MA**	**NJ$_1$**	**NJ$_2$**	**NJ$_3$**	**NJ$_4$**	**NJ$_5$**	**NJ$_6$**	**WABC**	**WI**	

Table 3: Standardized ϕ'_{PT} values of regional population structure. ϕ'_{PT} values are below the diagonal and p-values based on 999 permutations of the original data above the diagonal. Significant ϕ'_{PT} values shaded in gray. Refer to Table 1 for population IDs.

	DE	**MA**	**NJ$_1$**	**NJ$_2$**	**NJ$_3$**	**NJ$_4$**	**NJ$_5$**	**NJ$_6$**	**WABC**	**WI**	
DE		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.033	**DE**
MA	0.341		0.099	0.249	0.011	0.007	0.057	0.034	0.136	0.003	**MA**
NJ$_1$	0.268	0.034		0.441	0.160	0.259	0.444	0.310	0.213	0.195	**NJ$_1$**
NJ$_2$	0.263	0.014	0.000		0.021	0.240	0.424	0.017	0.389	0.190	**NJ$_2$**
NJ$_3$	0.295	0.072	0.031	0.085		0.024	0.010	0.099	0.047	0.014	**NJ$_3$**
NJ$_4$	0.333	0.125	0.020	0.025	0.092		0.139	0.009	0.060	0.174	**NJ$_4$**
NJ$_5$	0.235	0.034	0.000	0.000	0.074	0.030		0.094	0.258	0.307	**NJ$_5$**
NJ$_6$	0.316	0.047	0.009	0.063	0.035	0.101	0.024		0.042	0.011	**NJ$_6$**
WABC	0.291	0.038	0.030	0.003	0.087	0.082	0.015	0.071		0.039	**WABC**
WI	0.084	0.134	0.032	0.027	0.122	0.045	0.011	0.101	0.105		**WI**
	DE	**MA**	**NJ$_1$**	**NJ$_2$**	**NJ$_3$**	**NJ$_4$**	**NJ$_5$**	**NJ$_6$**	**WABC**	**WI**	

Table 4: ϕ_{PT} values of subwatershed population structure. ϕ_{PT} values are below the diagonal and p-values based on 999 permutations above the diagonal. Significant ϕ_{PT} values shaded in gray.

	BR	MRABR	OR	RCNB	RCSB	WBWR	
BR	0.000	0.077	0.367	0.064	0.118	0.208	**BR**
MRABR	0.035	0.000	0.413	0.047	0.548	0.104	**MRABR**
OR	0.009	0.000	0.000	0.272	0.445	0.437	**OR**
RCNB	0.040	0.024	0.008	0.000	0.171	0.056	**RCNB**
RCSB	0.057	0.000	0.000	0.024	0.000	0.259	**RCSB**
WBWR	0.016	0.011	0.000	0.020	0.011	0.000	**WBWR**
	BR	**MRABR**	**OR**	**RCNB**	**RCSB**	**WBWR**	

BR, Batsto River; MRABR, Mullica River above the Batsto River; OR, Oswego River; RCNB, Rancocas Creek North Branch; RCSB, Rancocas Creek –South Branch; WBWR, West Branch Wading River.

Table 5: Standardized ϕ'_{PT} values of subwatershed population structure. ϕ'_{PT} values are below the diagonal and p-values based on 999 permutations of the original data above the diagonal. Significant ϕ'_{PT} values shaded in gray.

	BR	**MRABR**	**OR**	**RCNB**	**RCSB**	**WBWR**	
BR		0.077	0.367	0.064	0.118	0.208	**BR**
MRABR	0.061		0.413	0.047	0.548	0.104	**MRABR**
OR	0.016	0.000		0.272	0.445	0.437	**OR**
RCNB	0.070	0.042	0.014		0.171	0.056	**RCNB**
RCSB	0.090	0.000	0.000	0.039		0.259	**RCSB**
WBWR	0.027	0.019	0.000	0.034	0.018		**WBWR**
	BR	**MRABR**	**OR**	**RCNB**	**RCSB**	**WBWR**	

BR, Batsto River; MRABR, Mullica River above the Batsto River; OR, Oswego River; RCNB, Rancocas Creek North Branch; RCSB, Rancocas Creek –South Branch; WBWR, West Branch Wading River.

Table 6: *D*$_{est}$ values of regional population structure. *D*$_{est}$ values with 95% confidence intervals estimated from 1000 bootstrapped datasets are below the diagonal and *p*-values based on 999 permutations from the AMOVA analysis are above the diagonal.

	DE	**MA**	**NJ$_1$**	**NJ$_2$**	**NJ$_3$**	**NJ$_4$**	**NJ$_5$**	**NJ$_6$**	**WABC**	**WI**
DE		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.033
MA	0.168 (0.125−0.212)		0.099	0.249	0.011	0.007	0.057	0.034	0.136	0.003
NJ$_1$	0.135 (0.09−0.179)	0.027 (0.008-0.05)		0.441	0.16	0.259	0.444	0.31	0.213	0.195
NJ$_2$	0.124 (0.082−0.165)	0.017 (0.001-0.039)	0.01 (-0.006-0.031)		0.021	0.24	0.424	0.017	0.389	0.19
NJ$_3$	0.154 (0.111−0.199)	0.043 (0.018−0.074)	0.029 (0.006-0.058)	0.049 (0.019−0.083)		0.024	0.01	0.099	0.047	0.014
NJ$_4$	0.167 (0.118−0.221)	0.066 (0.032−0.108)	0.027 (0-0.057)	0.026 (-0.001-0.06)	0.058 (0.026−0.095)		0.139	0.009	0.06	0.174
NJ$_5$	0.124 (0.082−0.17)	0.026 (0.008-0.048)	0.004 (-0.009-0.020)	0.004 (-0.007-0.02)	0.047 (0.02−0.08)	0.031 (0.008-0.064)		0.094	0.258	0.307
NJ$_6$	0.164 (0.119−0.208)	0.031 (0.012−0.055)	0.018 (0.001-0.042)	0.038 (0.015−0.067)	0.029 (0.007-0.056)	0.06 (0.028−0.103)	0.023 (0.005-0.046)		0.042	0.011
WABC	0.15 (0.103−0.201)	0.033 (0.012-0.061)	0.034 (0.012-0.062)	0.02 (0.004-0.041)	0.06 (0.025−0.103)	0.057 (0.027-0.096)	0.028 (0.011-0.051)	0.05 (0.02−0.086)		0.039
WI	0.053 (0.017−0.102)	0.073 (0.04−0.112)	0.034 (0.007-0.071)	0.029 (0.004-0.063)	0.076 (0.04−0.117)	0.041 (0.008-0.087)	0.025 (0-0.058)	0.063 (0.03−0.1)	0.071 (0.038−0.11)	
	DE	**MA**	**NJ$_1$**	**NJ$_2$**	**NJ$_3$**	**NJ$_4$**	**NJ$_5$**	**NJ$_6$**	**WABC**	**WI**

Table 7: D_{est} values of regional population structure from clone-corrected, resampled regional datasets. D_{est} values with 95% confidence intervals estimated from 100 datasets resampled to a standardized sample size corresponding to the sample size of the smallest clone-corrected population (WABC: *n*=12). Continued on the next page.

	NJ$_1$	NJ$_2$	NJ$_3$	NJ$_4$	NJ$_5$	NJ$_6$
NJ$_1$	0					
NJ$_2$	0.024 (0.021-0.026)	0				
NJ$_3$	0.044 (0.039-0.048)	0.064 (0.059-0.069)	0			
NJ$_4$	0.035 (0.032-0.038)	0.033 (0.03-0.036)	0.07 (0.065-0.075)	0		
NJ$_5$	0.019 (0.016-0.022)	0.021 (0.018-0.023)	0.067 (0.061-0.072)	0.043 (0.039-0.046)	0	
NJ$_6$	0.036 (0.033-0.04)	0.056 (0.052-0.061)	0.046 (0.042-0.05)	0.075 (0.07-0.08)	0.043 (0.04-0.047)	0
WABC	0.041 (0.038-0.043)	0.025 (0.023-0.027)	0.069 (0.065-0.073)	0.06 (0.058-0.062)	0.038 (0.036-0.04)	0.062 (0.059-0.065)
WI	0.037 (0.035-0.04)	0.035 (0.033-0.037)	0.083 (0.078-0.088)	0.044 (0.042-0.046)	0.035 (0.032-0.037)	0.073 (0.07-0.077)
MA	0.04 (0.036-0.043)	0.029 (0.026-0.032)	0.058 (0.054-0.061)	0.074 (0.07-0.079)	0.043 (0.039-0.046)	0.049 (0.045-0.053)
DE	0.138 (0.133-0.143)	0.131 (0.128-0.134)	0.163 (0.157-0.17)	0.171 (0.168-0.174)	0.135 (0.131-0.139)	0.175 (0.169-0.18)

Table 7: continued from previous page.

	WABC	WI	MA	DE
NJ_1				
NJ_2				
NJ_3				
NJ_4				
NJ_5				
NJ_6				
WABC	0			
WI	0.071 (0.07-0.073)	0		
MA	0.039 (0.037-0.042)	0.078 (0.074-0.081)	0	
DE	0.151 (0.148-0.155)	0.056 (0.054-0.058)	0.173 (0.168-0.178)	0

Table 8: Global and single locus estimates of D_{est} on clone-corrected data. 95% confidence intervals based on 1000 bootstrapped datasets.

	Mean	2.5%	97.5%
Cgret	0.027	-0.005	0.069
L10D10	0.049	0.023	0.089
L14F4	0.088	0.063	0.177
L2C1	0.003	-0.016	0.000
LB5B4	0.023	-0.010	0.038
LC1942	0.321	0.285	0.411
LC2090	-0.006	-0.021	0.016
LC4168	0.174	0.110	0.226
LF9	0.042	0.005	0.072
global	0.0584	0.047	0.071

Table 9: Global and single locus estimates of D_{est} and G_{ST} from clone-corrected, resampled regional datasets. Each of the 100 clone-corrected datasets were constructed by resampling to the size of the smallest clone corrected population (WABC: n=12) in the regional dataset.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.403 (0.399-0.407)	0.422 (0.418-0.426)	0.045 (0.041-0.05)	0.035 (0.032-0.039)
L10D10	0.428 (0.421-0.434)	0.46 (0.453-0.467)	0.07 (0.066-0.075)	0.064 (0.059-0.068)
L14F4	0.64 (0.635-0.644)	0.677 (0.672-0.682)	0.055 (0.051-0.058)	0.115 (0.107-0.123)
L2C1	0.137 (0.132-0.142)	0.141 (0.136-0.146)	0.025 (0.02-0.029)	0.005 (0.004-0.006)
LB5B4	0.105 (0.102-0.108)	0.124 (0.12-0.127)	0.148 (0.137-0.158)	0.023 (0.021-0.025)
LC1942	0.716 (0.712-0.72)	0.805 (0.801-0.808)	0.11 (0.107-0.113)	0.347 (0.339-0.355)
LC2090	0.386 (0.38-0.391)	0.388 (0.383-0.393)	0.006 (0.003-0.008)	0.004 (0.002-0.005)
LC4168	0.349 (0.342-0.355)	0.454 (0.448-0.461)	0.233 (0.226-0.239)	0.18 (0.175-0.185)
LF9	0.208 (0.203-0.212)	0.241 (0.236-0.246)	0.137 (0.129-0.144)	0.047 (0.044-0.05)
global	0.375 (0.373-0.376)	0.412 (0.411-0.414)	0.092 (0.09-0.093)	0.067 (0.066-0.068)

Table 10: Summary statistics per locus calculated for the estimation of demographic history of *C. fructivorum* in the Delaware population.

	observed			I.A.M.		S.M.M.	
locus	*n*	*A*	H_e	H_{eq}	DH/sd	H_{eq}	DH/sd
Cgret	21	2	0.257	0.273	-0.099	0.31	-0.346
L10D10	21	3	0.186	0.449	-1.73	0.53	-2.811
L14F4	21	4	0.486	0.58	-0.775	0.659	-1.978
L2C1	21	2	0.257	0.263	-0.04	0.321	-0.413
LB5B4	21	3	0.495	0.456	0.274	0.529	-0.279
LC1942	21	3	0.452	0.447	0.034	0.528	-0.615
LC2090	21	2	0.257	0.267	-0.065	0.32	-0.416
LC4168	21	3	0.338	0.454	-0.772	0.527	-1.565
LF9	21	1	NA	NA	NA	NA	NA

n, number of haplotypes; *A*, number of alleles; H_e, expected heterozygosity (Nei)/gene diversity; H_{eq}, expected heterozygosity under the null model of mutation-drift equilibrium; DH/sd, Standard Deviates (H_e-H_{eq})/(standard deviation of the mutation-drift equilibrium of the gene diversity); I.A.M, infinite alleles model; S.M.M., stepwise mutation model

Table 11: Summary of the diversity indices for collections with precise global positioning system coordinates in New Jersey.

n	n_{h}	$\$n_{\mathrm{h}}$	$\$n_{\mathrm{h}[-1]}$	CF	G	MF	$*H_{\mathrm{e}}$	$**H_{\mathrm{e}}$	$*I_{A}$	$*\bar{r}_{\mathrm{d}}$	$**I_{A}$	$**\bar{r}_{\mathrm{d}}$
146	129	101	30	0.217	0.991	9	0.389 ± 0.081	0.428 ± 0.085	0.159	0.021	-0.028	-0.004

n, sample size; n_{h}, number of unique multilocus haplotypes when clone corrected by sample point; $\$n_{\mathrm{h}}$, number of unique haplotypes across the entire collection; $\$n_{\mathrm{h}[-1]}$, the number of unique haplotypes when all haplotypes differing by a single mutation under the infinite alleles model are removed; CF, clonal fraction :1-($\$n_{\mathrm{h}}/n_{\mathrm{h}}$); G, genotypic diversity calculated from data clone corrected by sample point; MF, frequency of the most frequent multilocus haplotype; H_{e}, mean unbiased gene diversity across loci (± standard error); I_{A}, index of association; $\bar{r}_{\mathrm{d}}$, estimate of index of association independent of the number of loci; *, statistic calculated from clone-corrected dataset n_{h}; **, statistic calculated from clone-corrected dataset $\$n_{\mathrm{h}}$;

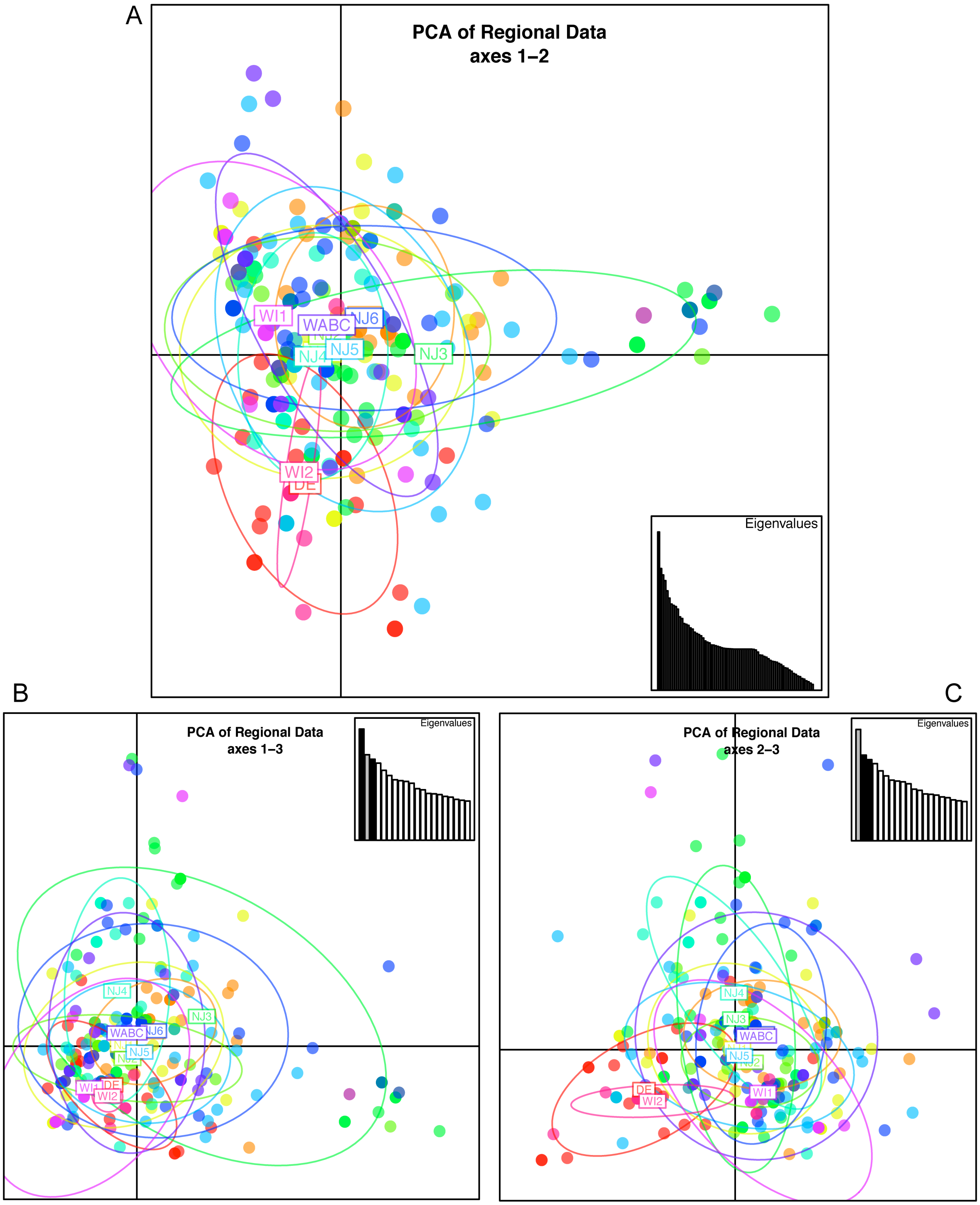


Figure 1: legend on next page

Figure 1: Principal components analysis of the regional data with regional subpopulations as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.

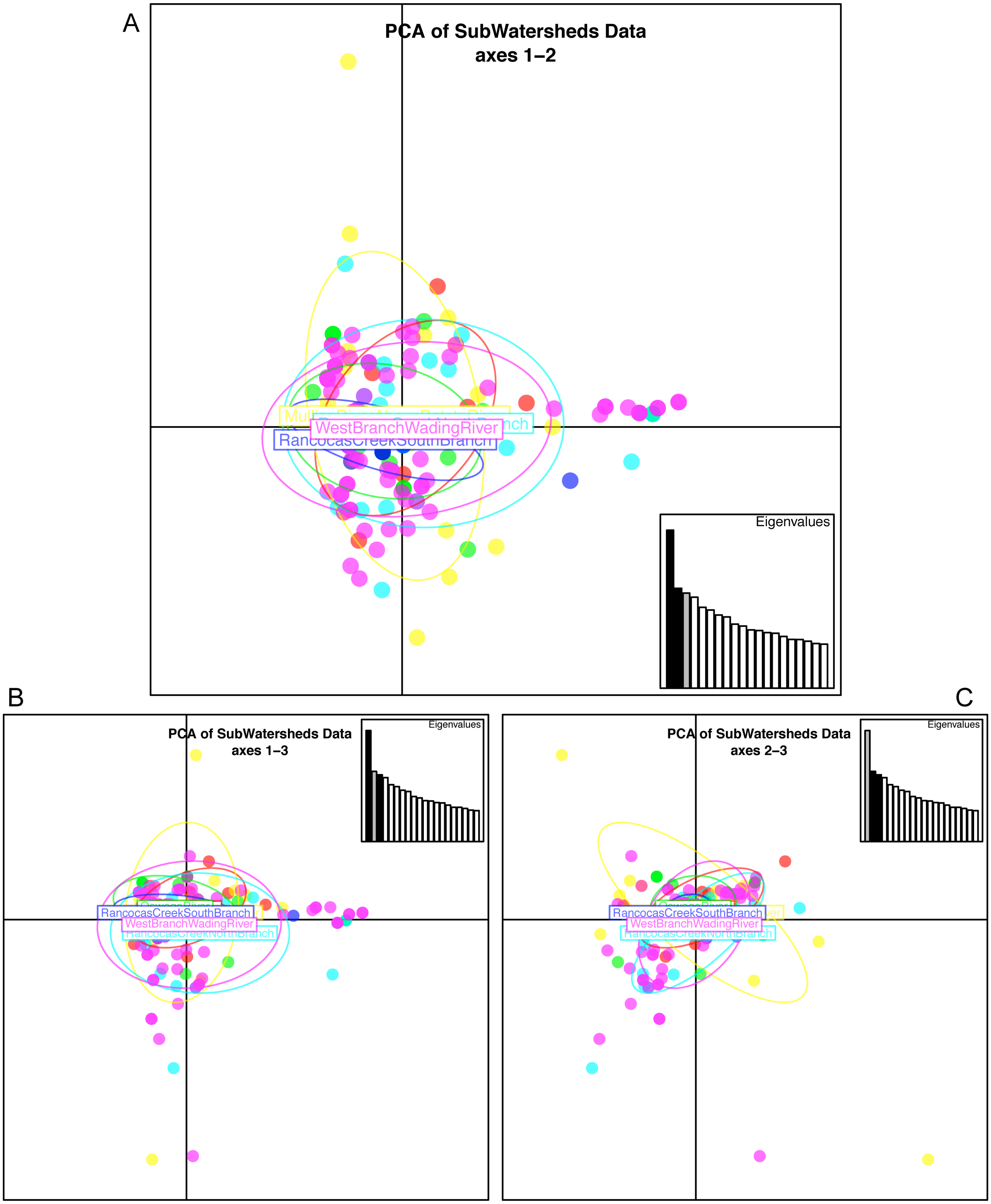


Figure 2: legend on next page

Figure 2: Principal components analysis of the subwatershed data with subwatersheds as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.

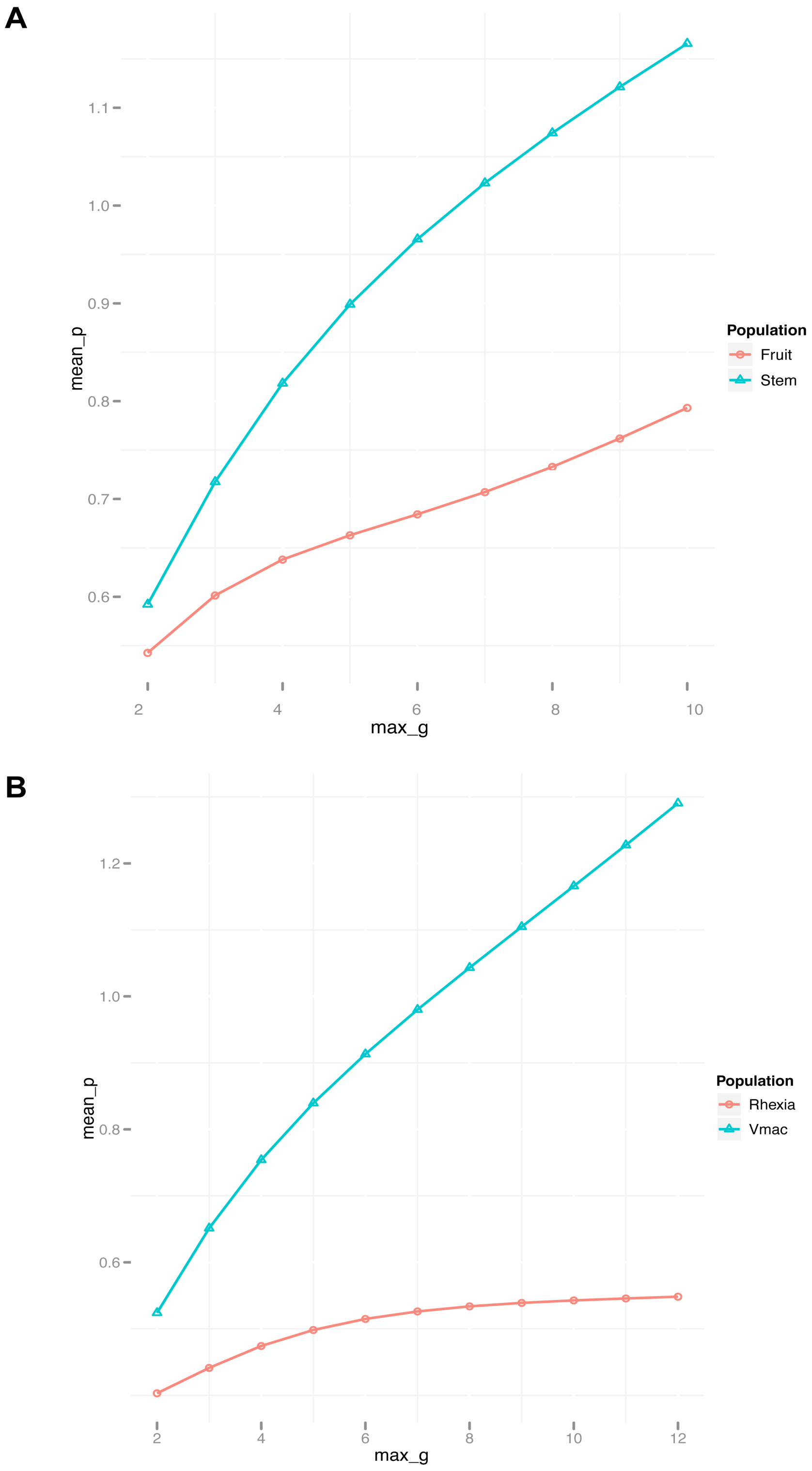


Figure 3: legend on next page

Figure 3: Comparison of rarefied private allelic richness among host species and among host organs. Mean_p is the private allelic richness averaged over loci for each number of sampled genotypes (max_g). A) Private allelic richness among host organs. B) Private allelic richness among host species.

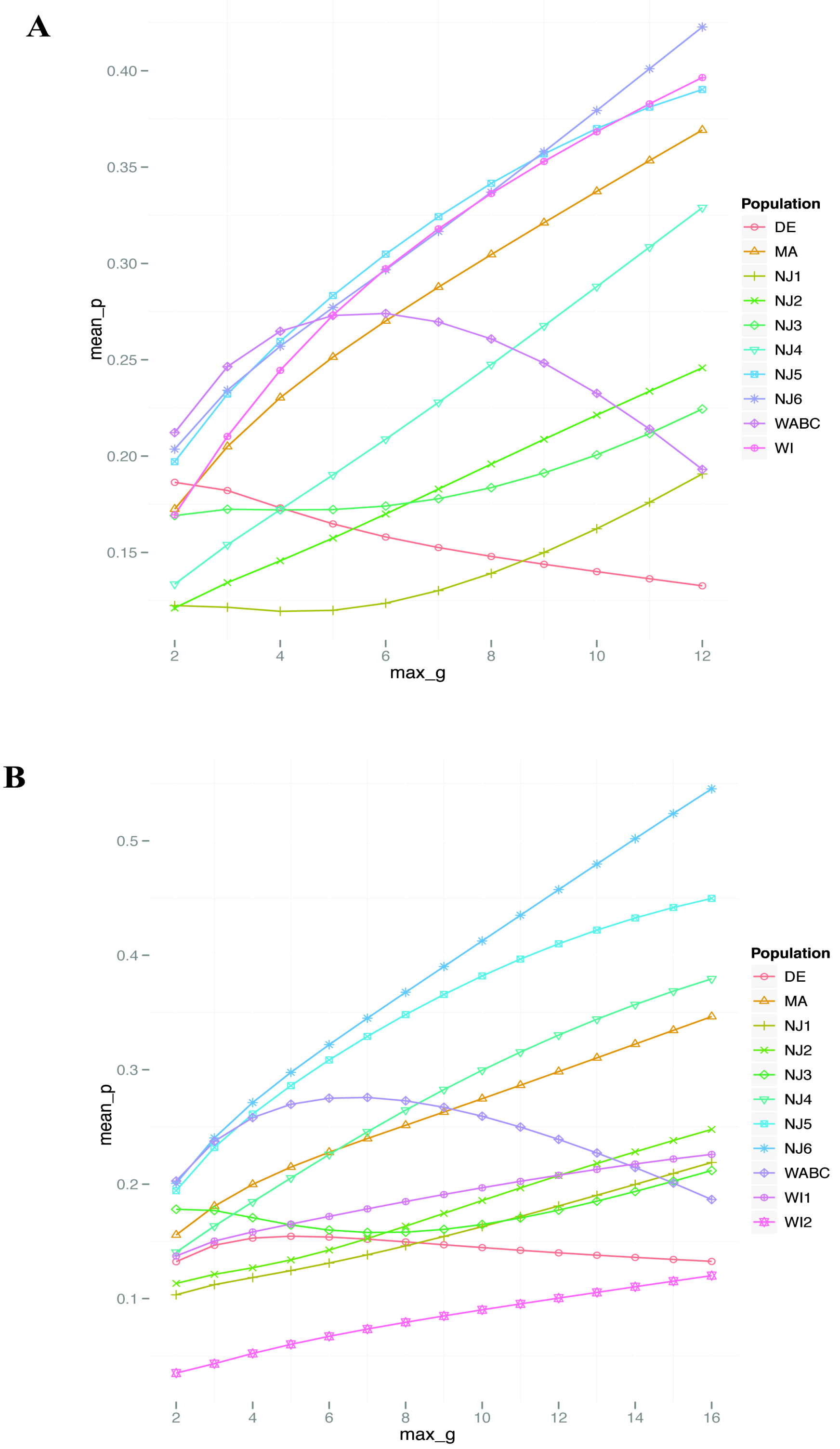


Figure 4: legend on next page

Figure 4: Comparison of private allelic richness among clone-corrected regional subpopulations and without clone-correction. Mean_p is the private allelic richness averaged over loci for each number of sampled genotypes (max_g). A) Private allelic richness among clone-corrected regional subpopulations. B) Private allelic richness among regional subpopulations without clone correction.

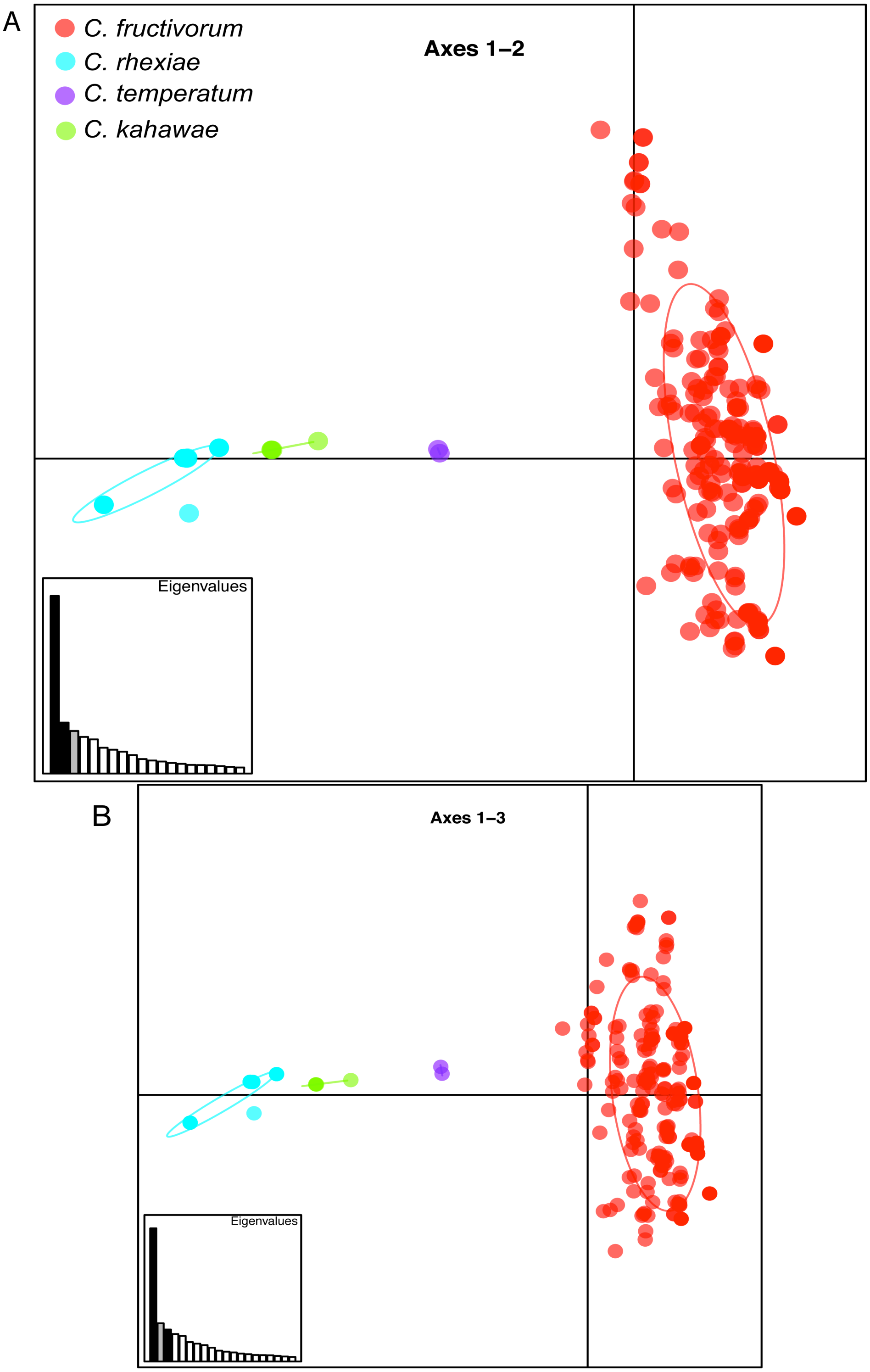


Figure 5: legend on next page

Figure 5: Principal components analysis of haplotypes from *C. fructivorum, C. rhexiae, C. temperatum,* and *C. kahawae* with species as factors. Each species is represented by a unique color. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3.

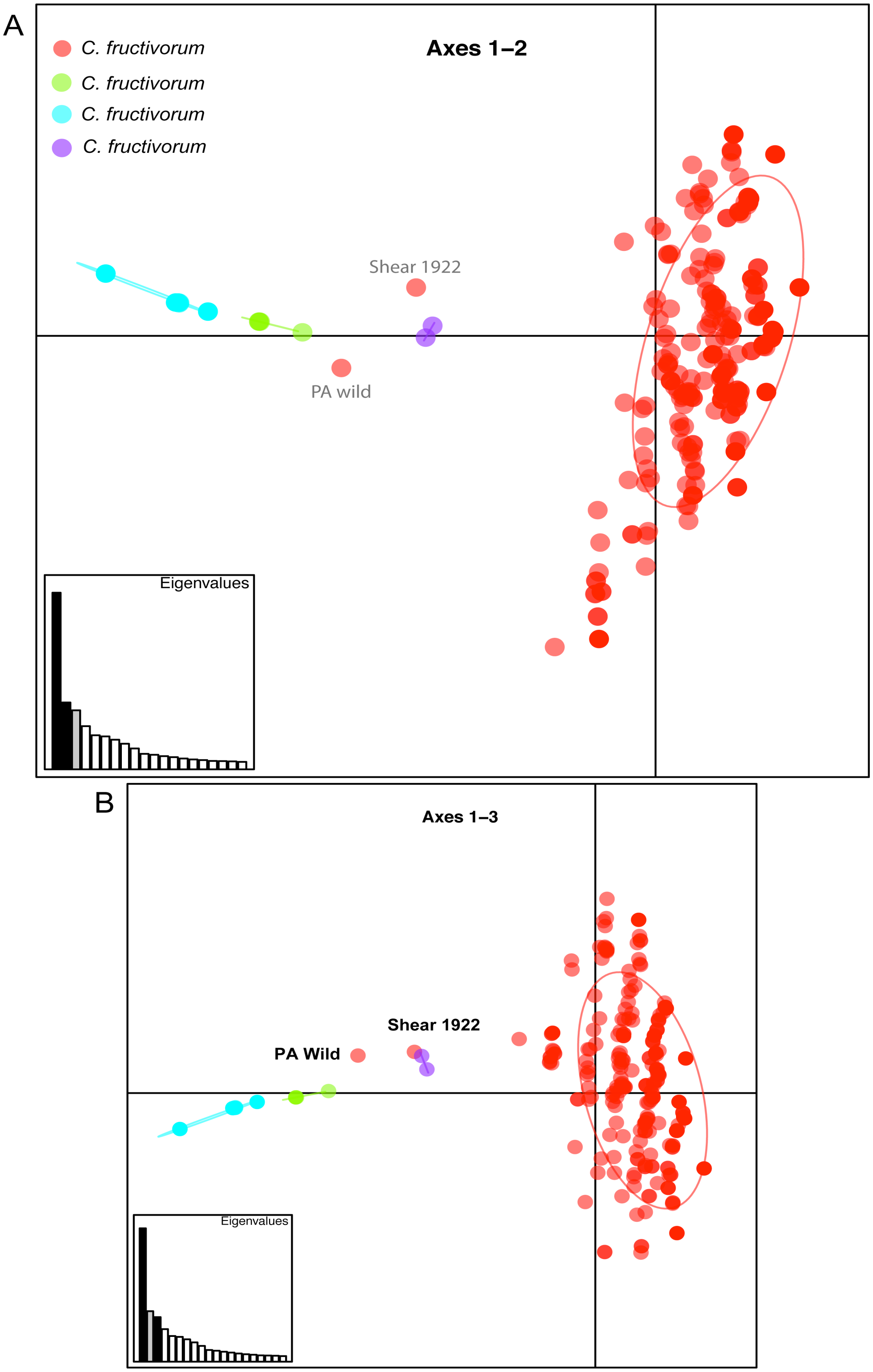


Figure 6: legend on next page

Figure 6: Principal components analysis of haplotypes from *C. fructivorum, C. rhexiae, C. temperatum,* and *C. kahawae* species as factors and two additional isolates that were singletons and thus excluded from other analyses. Each species is represented by a unique color. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3.

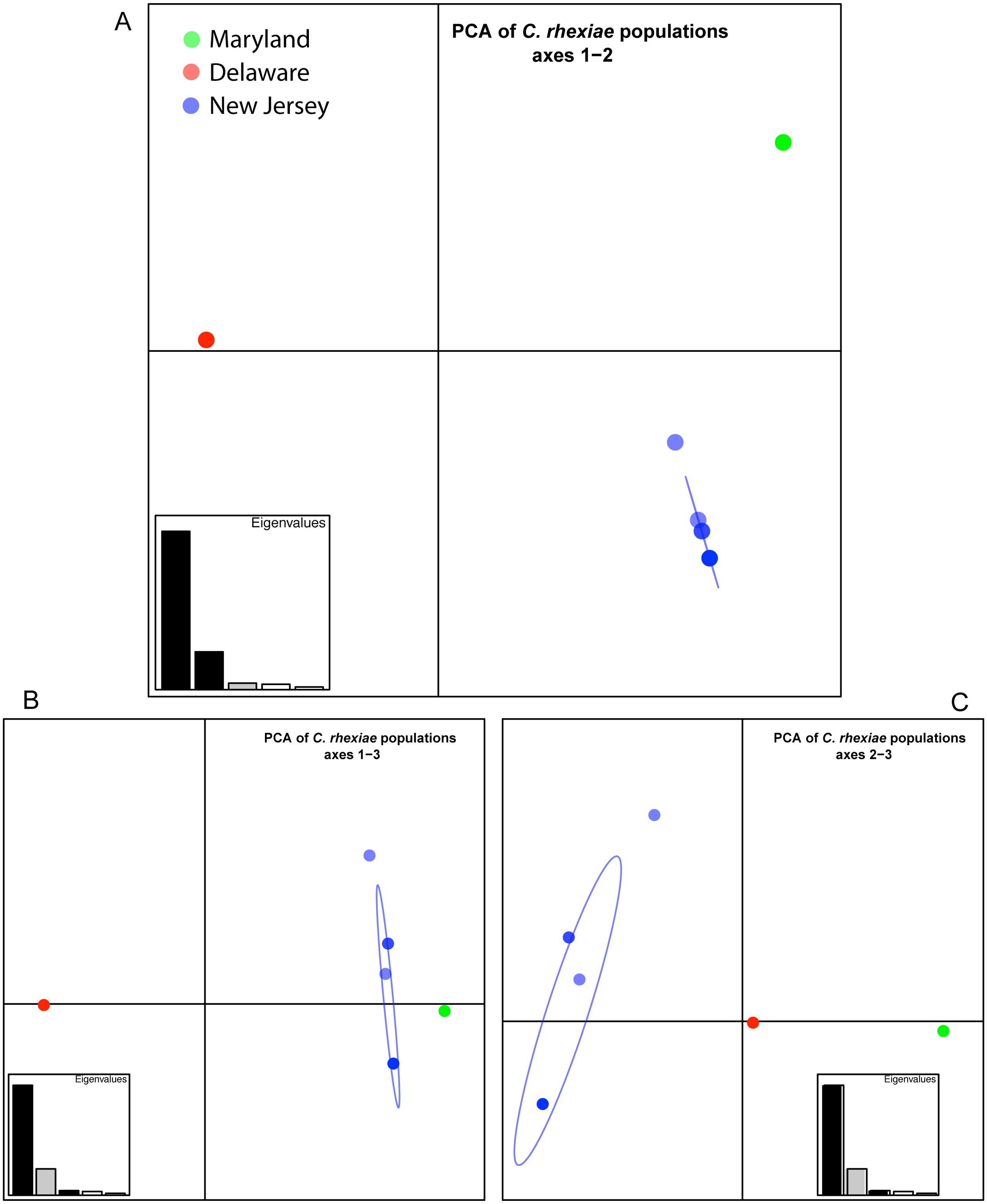


Figure 7: legend on next page

Figure 7: Principal components analysis of *C. rhexiae* with regional populations as factors. Each population is represented by a unique color. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.

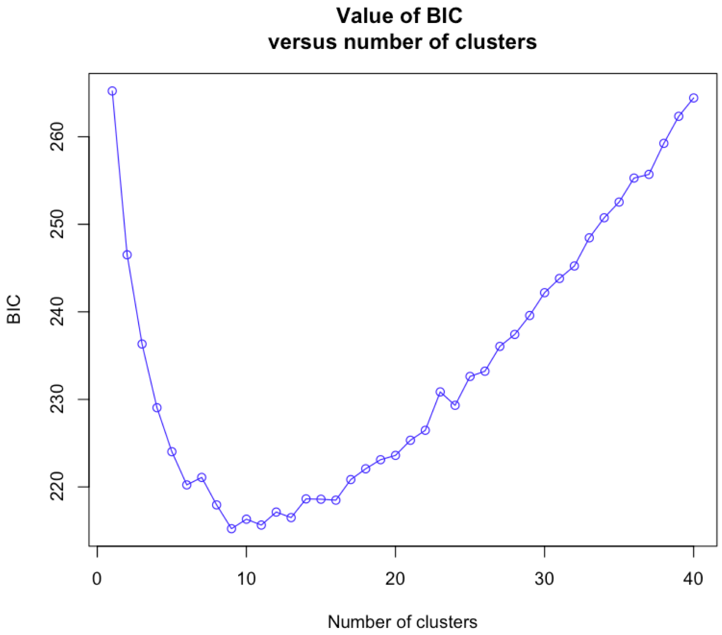


Figure 8: Value of BIC plotted against the number of clusters inferred with non-parametric *k*-means clustering.

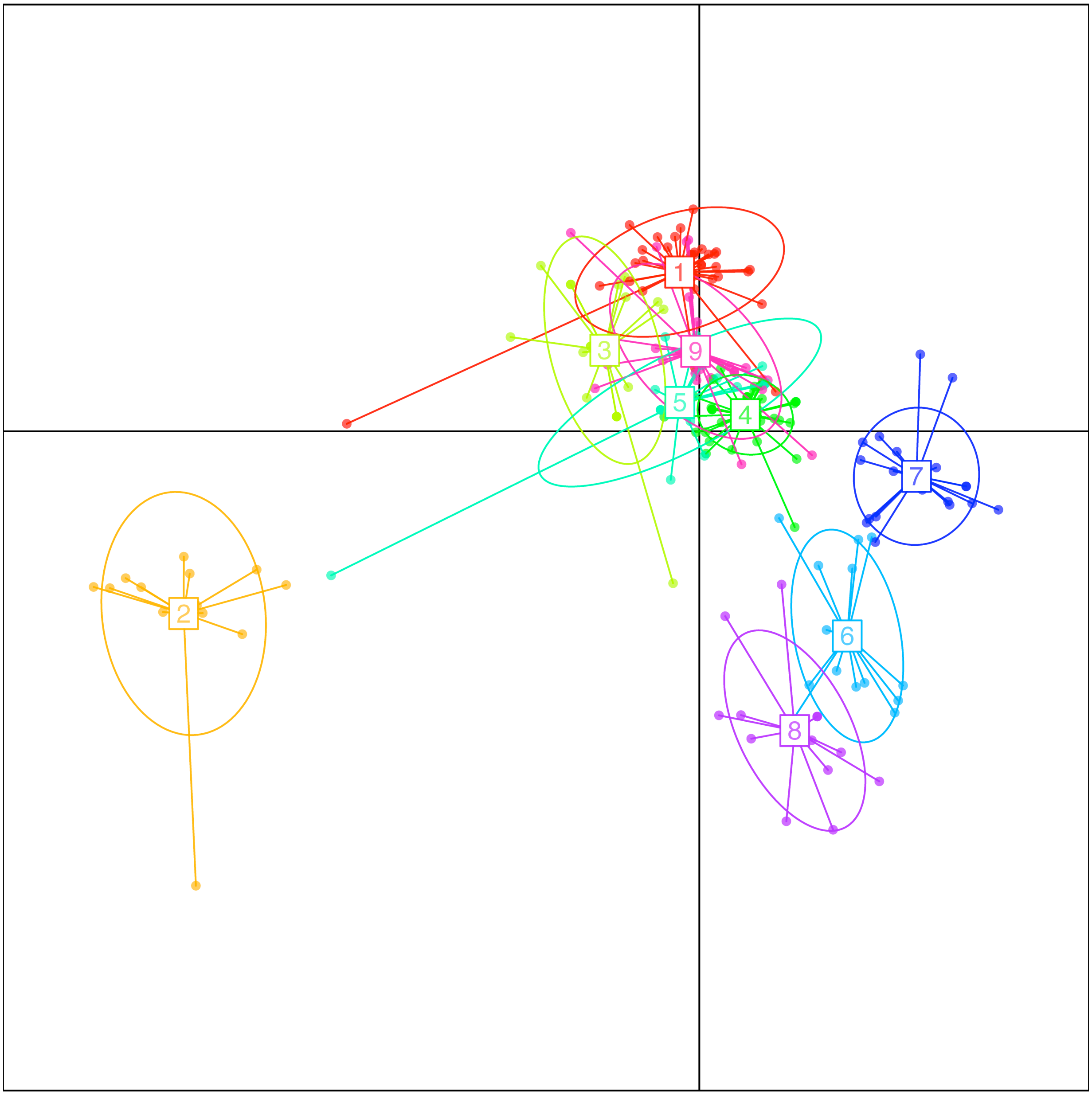


Figure 9: Scatterplot of haplotype clusters inferred from the non-parametric *k*-means clustering characterized by DAPC. Individual haplotypes are represented by dots and clusters are represented by inertia ellipses.

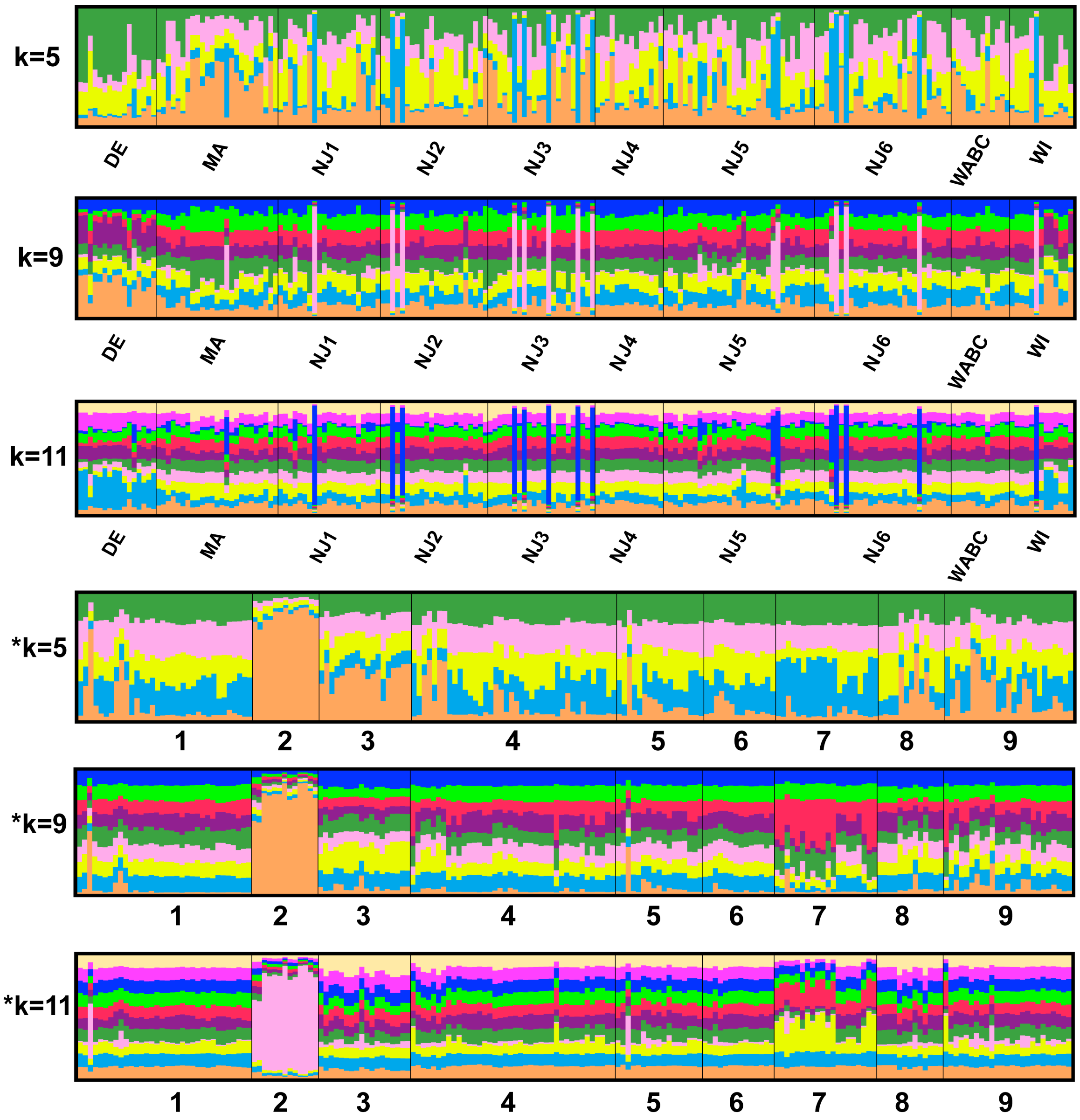


Figure 10: Comparison of *C. fructivorum* a priori defined populations and *DAPC defined populations with population assignment probabilities using STRUCTURE for k=5, k=9, and k=11. Each vertical bar represents a single individual and the color partitions within the bars represent the cluster assignment probabilities of each individual to each population (k).

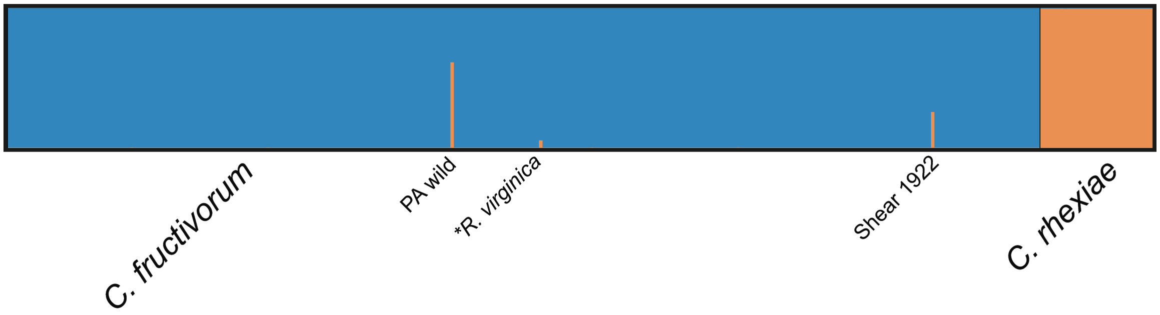


Figure 11: Levels of admixture between Colletotrichum fructivorum and C. rhexiae inferred with Bayesian clustering implemented in STRUCTURE. Each vertical bar represents a single individual and the color partitions within the bars represent the cluster assignment probabilities of each individual to each population. PA wild (Coll886), *R. virginica, and Shear 1922 (CBS 124) are the only individuals with some signature of admixture.

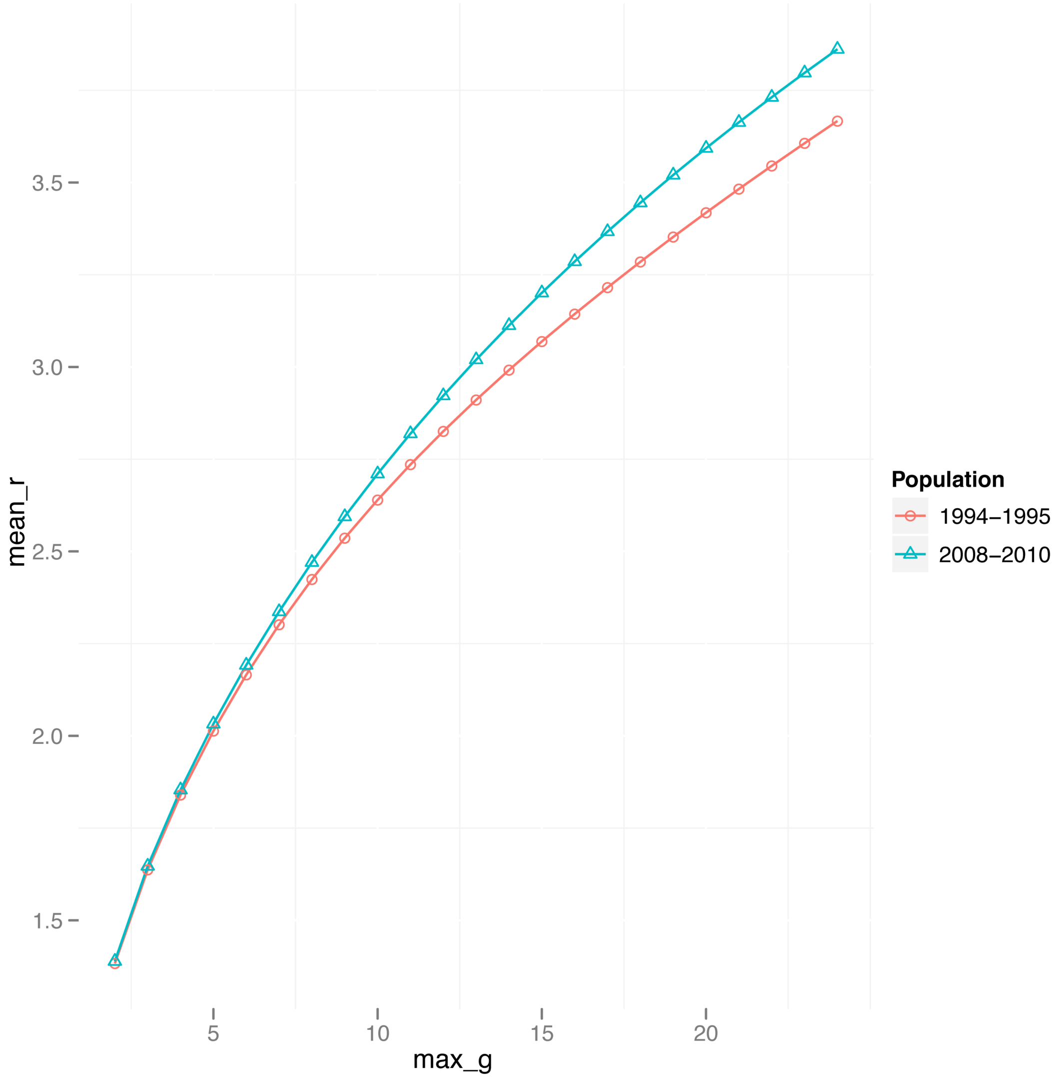


Figure 12: Comparison of allelic richness among temporally isolated collections. The red line represents the allelic richness of collections from 1994 to 1997. The blue line represents the allelic richness of isolates collected between 2008 and 2010. Mean_r is the allelic richness averaged over loci for each number of sampled genotypes (max_g).

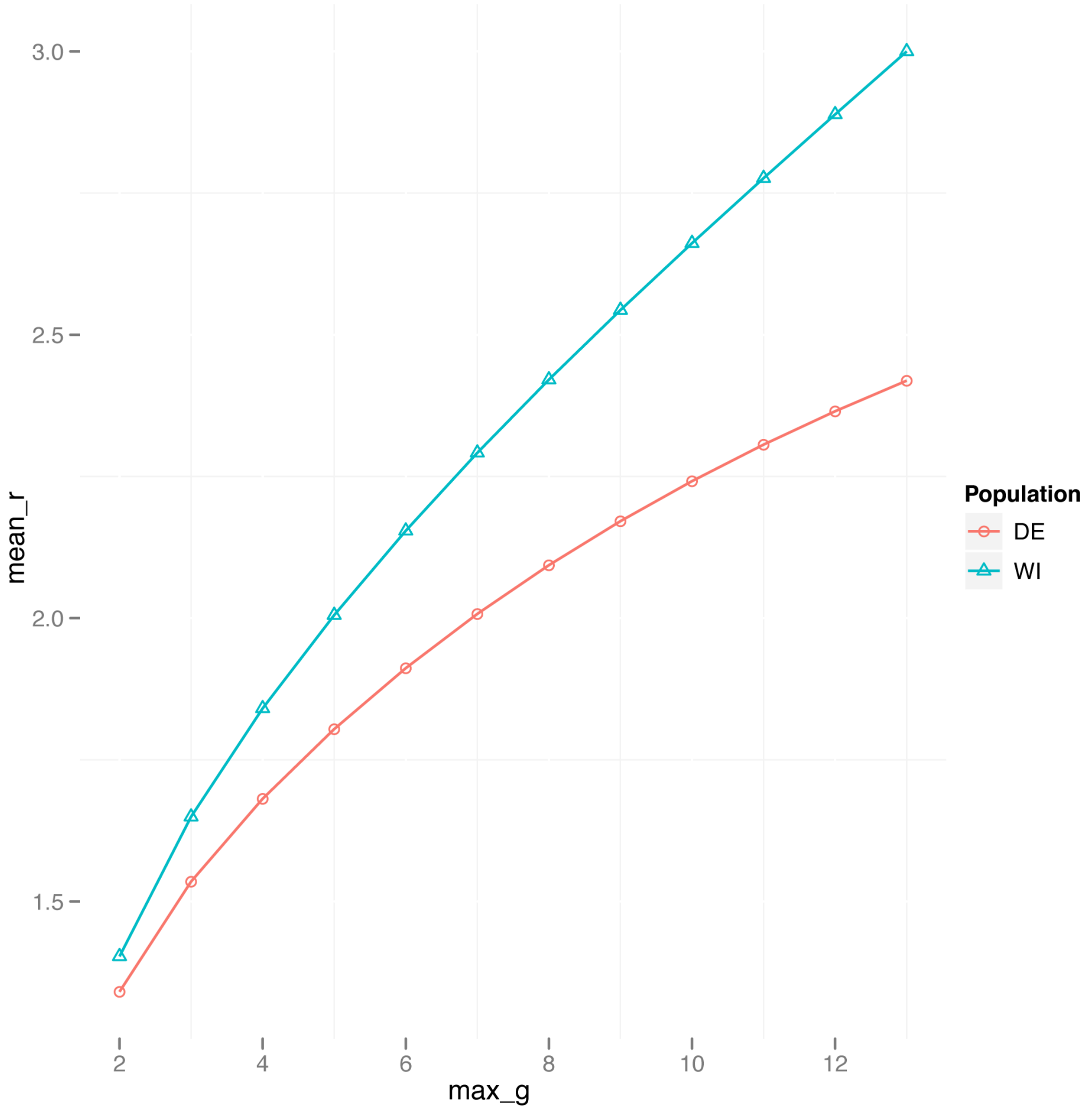


Figure 13: Comparison of rarefied allelic richness between WI and DE. The red line represents the allelic richness of collections from Delaware. The blue line represents the allelic richness of isolates collected from Wisconsin. Mean_r is the allelic richness averaged over loci for each number of sampled genotypes (max_g).

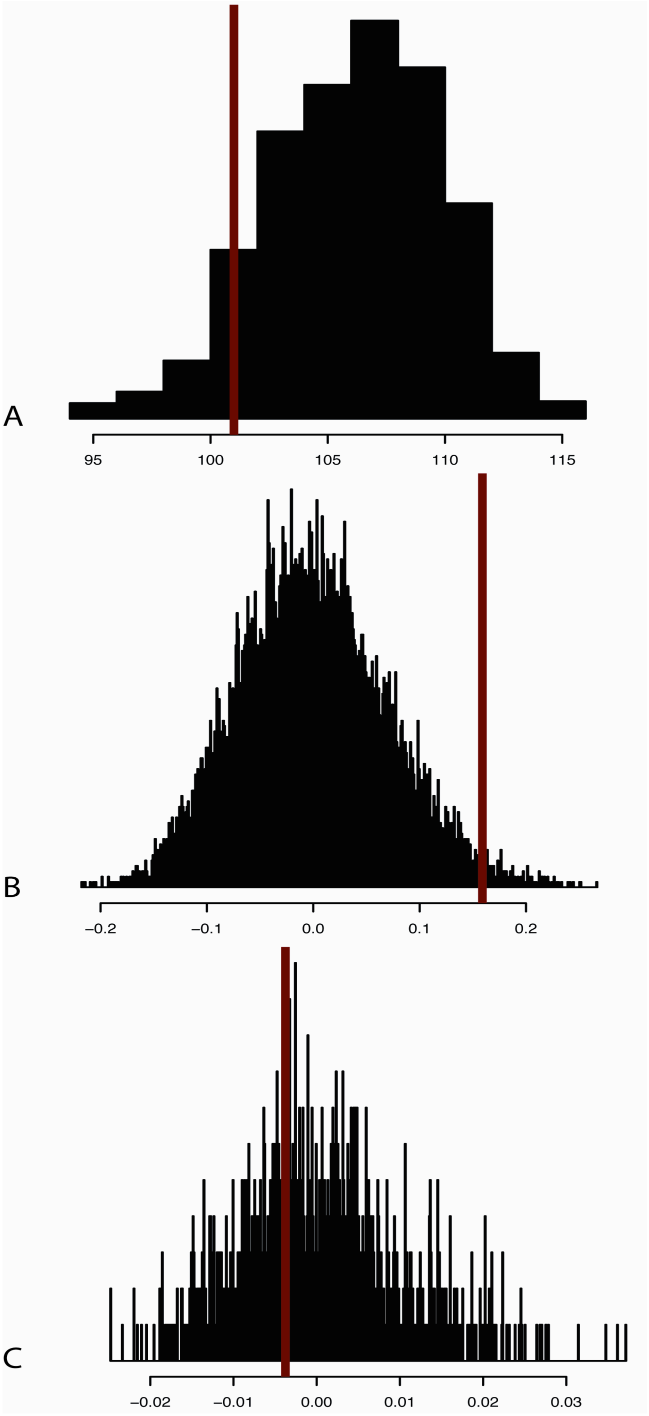


Figure 14: legend on next page

Figure 14: Tests for sexual recombination in *C. fructivorum*. A) Null distribution of the number of genotypes expected under random mating compared with observed data clone corrected by sample point. B) Null distribution of $\bar{r}_d$ expected under random mating compared with observed data clone-corrected by sample point. C) Null distribution of $\bar{r}_d$ expected under random mating compared with observed data after removing all identical haplotypes. Observed data represented by the red vertical line.

APPENDIX 2

SUPPLEMENTARY FIGURES FOR CHAPTER 2

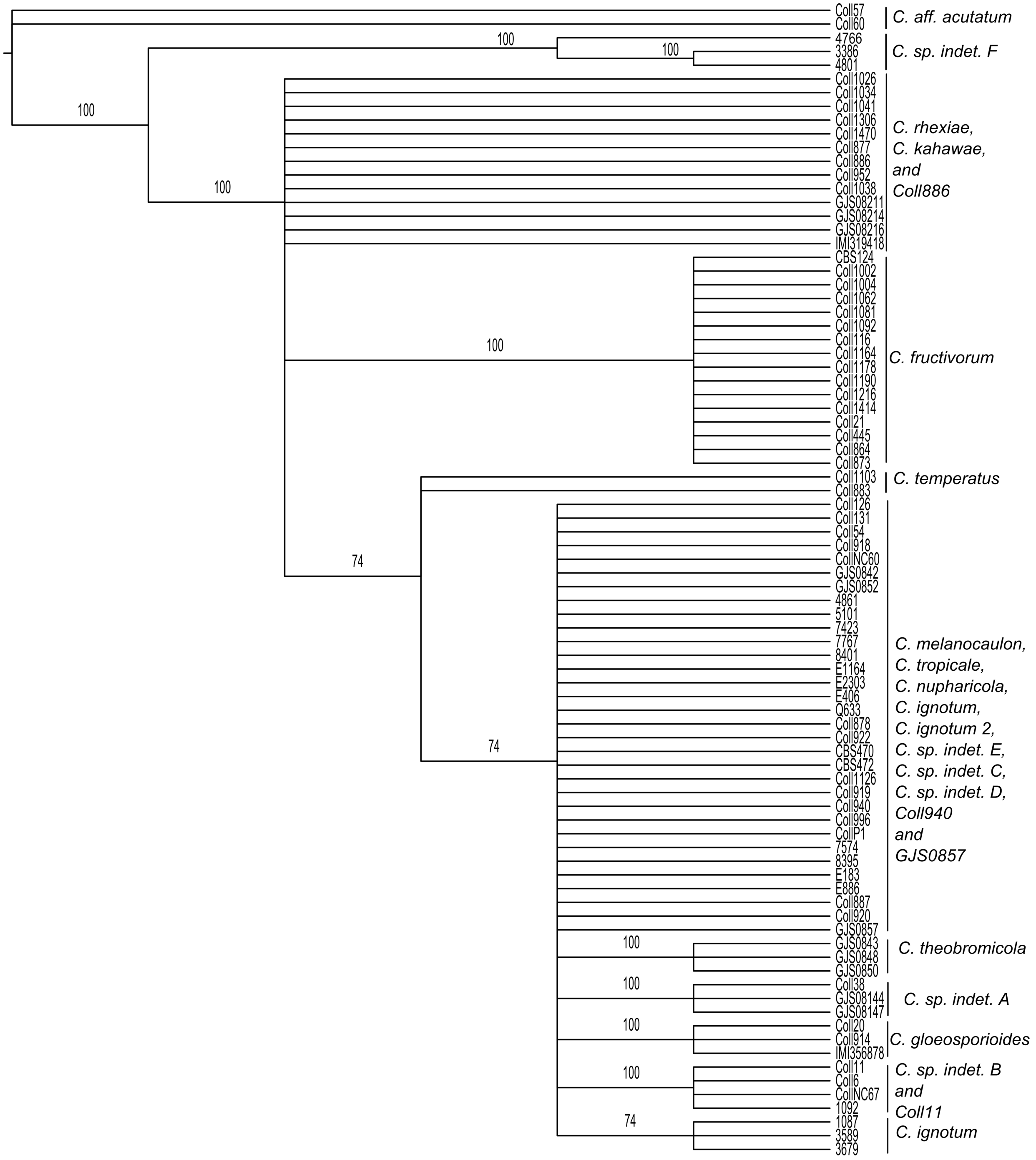


Figure S1: legend on next page

Figure S1: Nelsen consensus from maximum parsimony analysis of *nrITS* gene. Boostrap support values shown above branches.

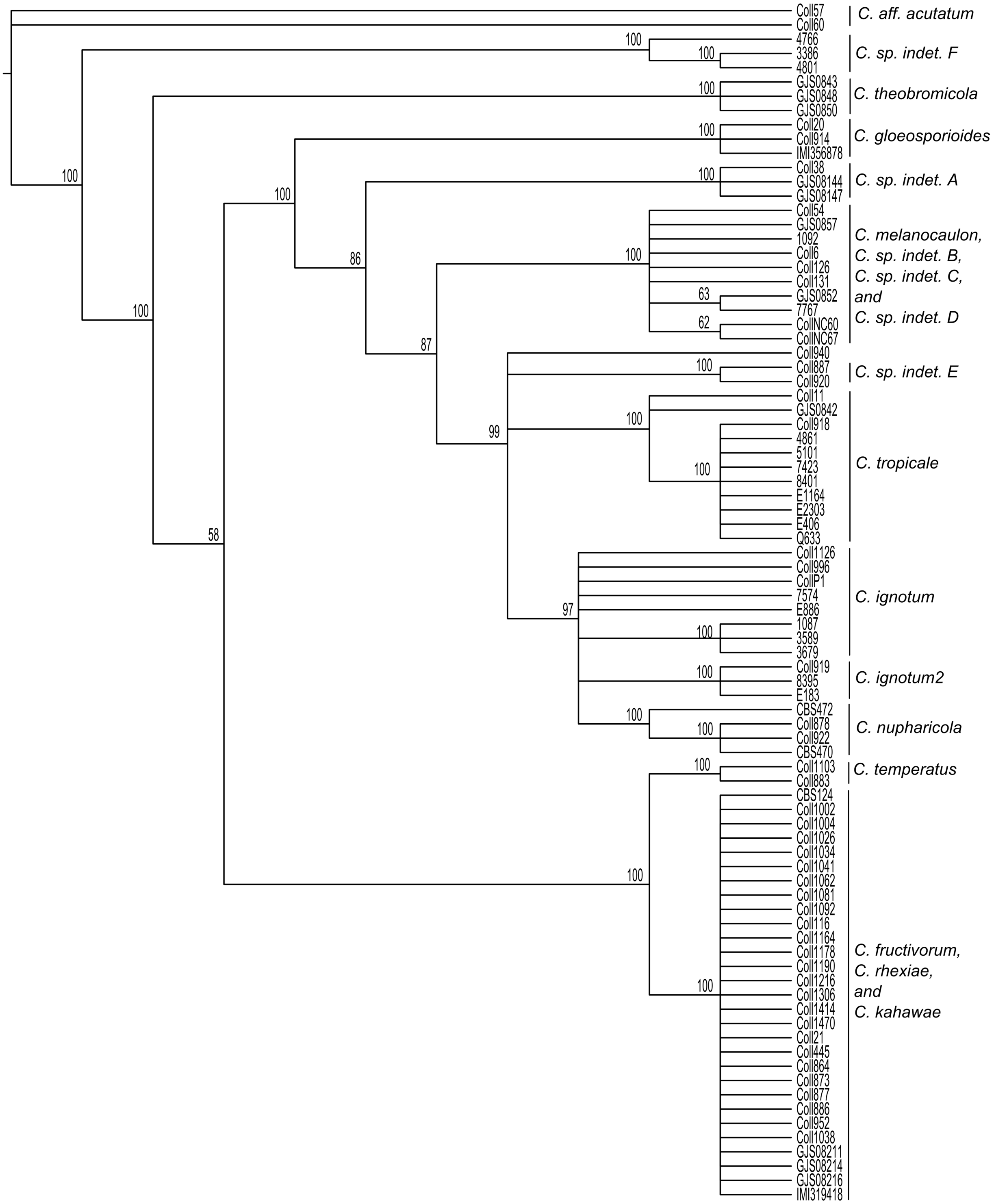


Figure S2: legend on next page

Figure S2: Nelsen consensus from maximum parsimony analysis of partial *beta-tubulin* gene.

Boostrap support values shown above branches.

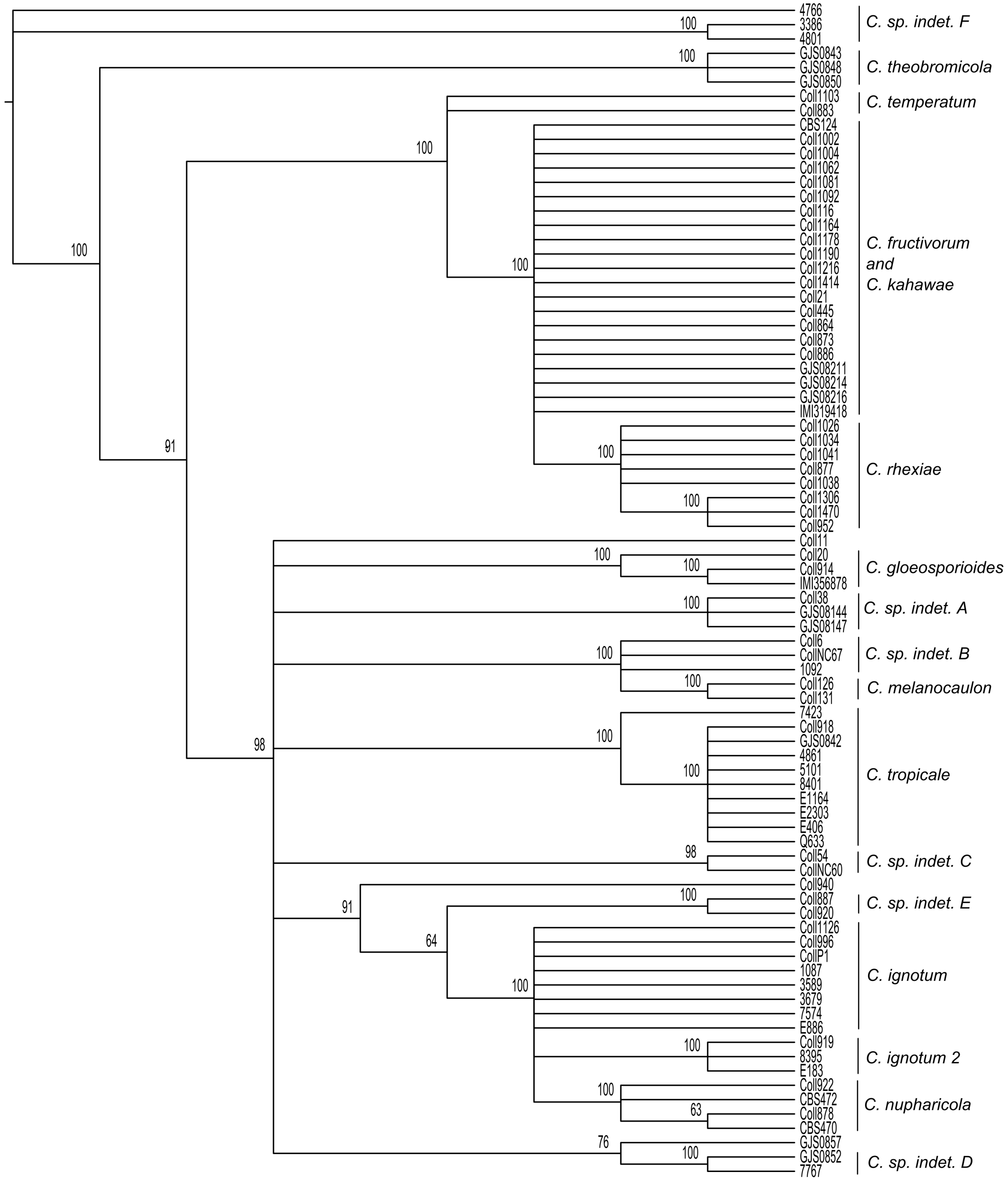


Figure S3: legend on next page

Figure S3: Nelsen consensus from maximum parsimony analysis of *apn2* gene. Boostrap support values shown above branches.

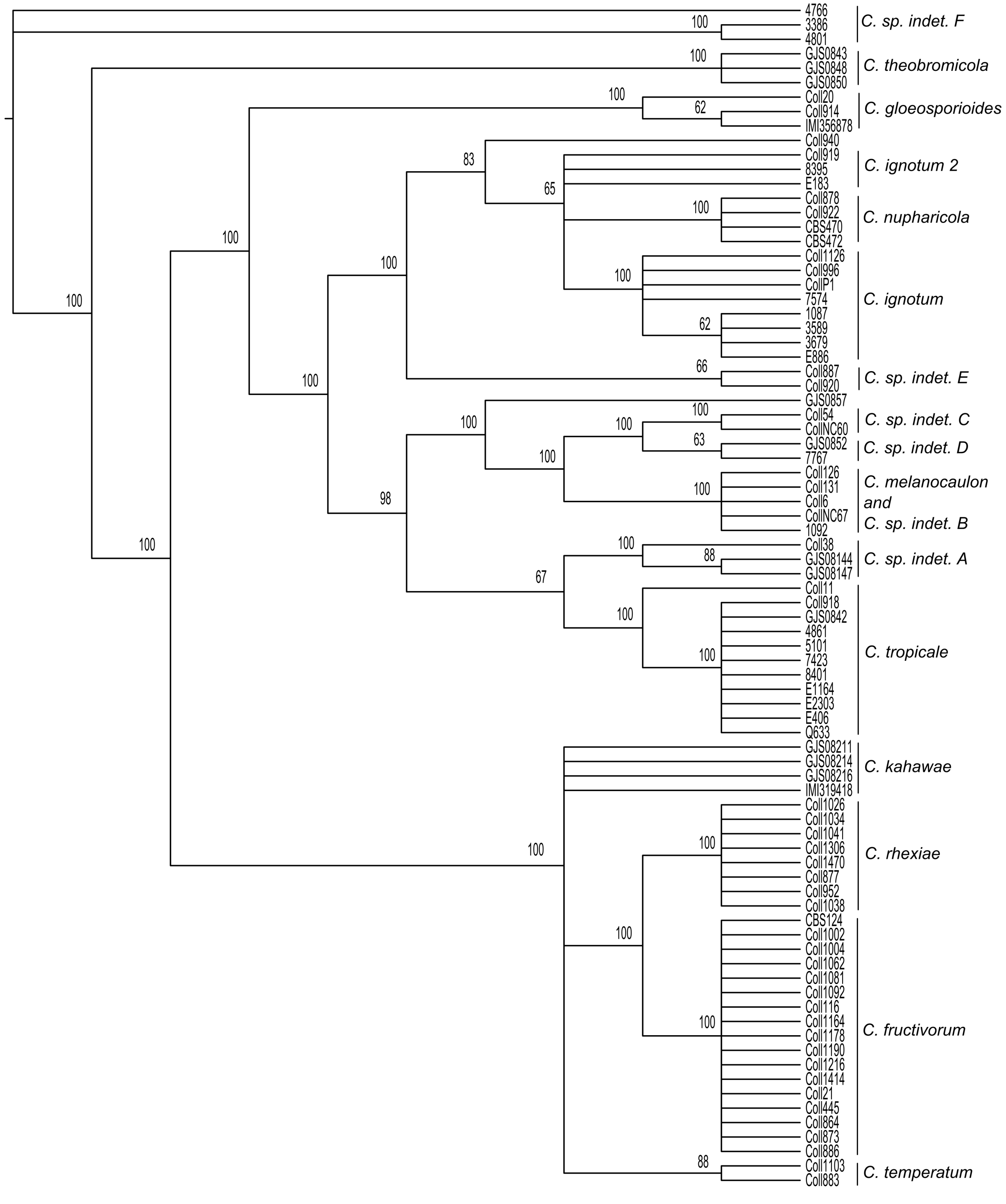


Figure S4: legend on next page

Figure S4: Nelsen consensus from maximum parsimony analysis of *apn2/matIGS* gene.

Boostrap support values shown above branches.

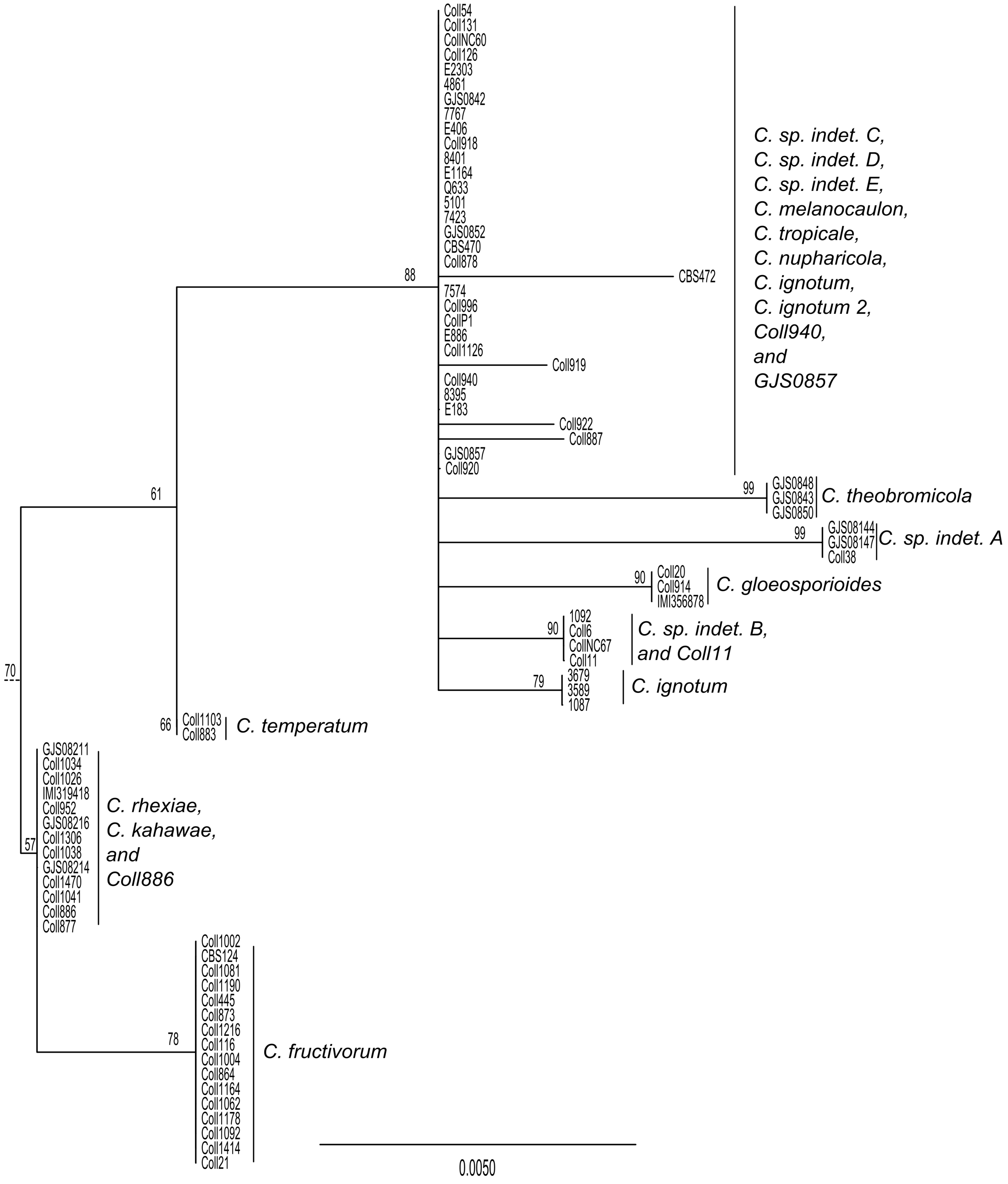


Figure S5: legend on next page

Figure S5: Maximum-likelihood bootstrap consensus tree of *nrITS* gene. Boostrap support values shown above branches. Outgroups (*C. aff. acutatum* and strains 4766, 3386, and 4801) have been trimmed from the tree.

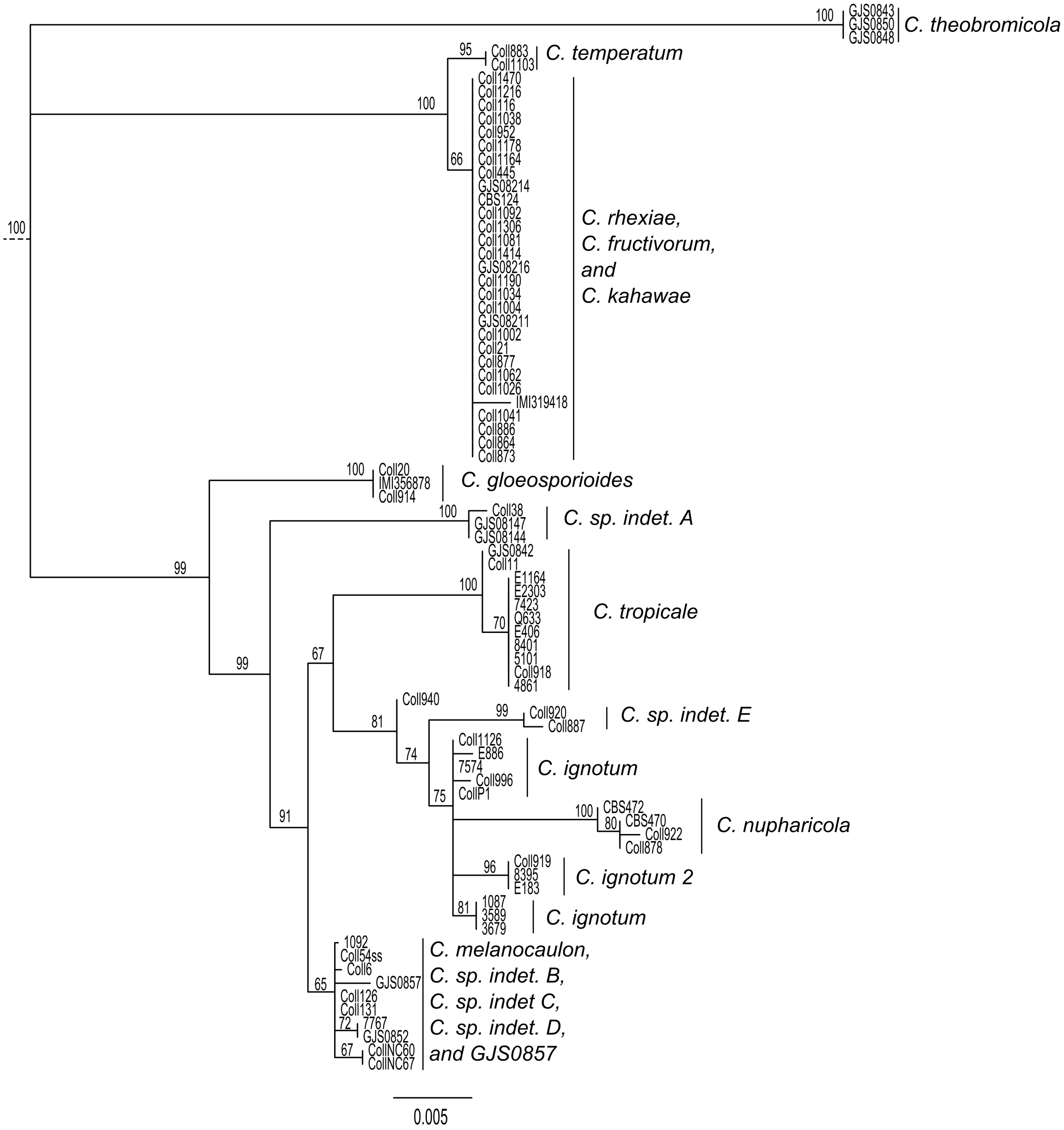


Figure S6: legend on next page

Figure S6: Maximum-likelihood bootstrap consensus tree of partial *beta-tubulin* gene. Boostrap support values shown above branches. Outgroups (*C. aff. acutatum* and strains 4766, 3386, and 4801) have been trimmed from the tree.

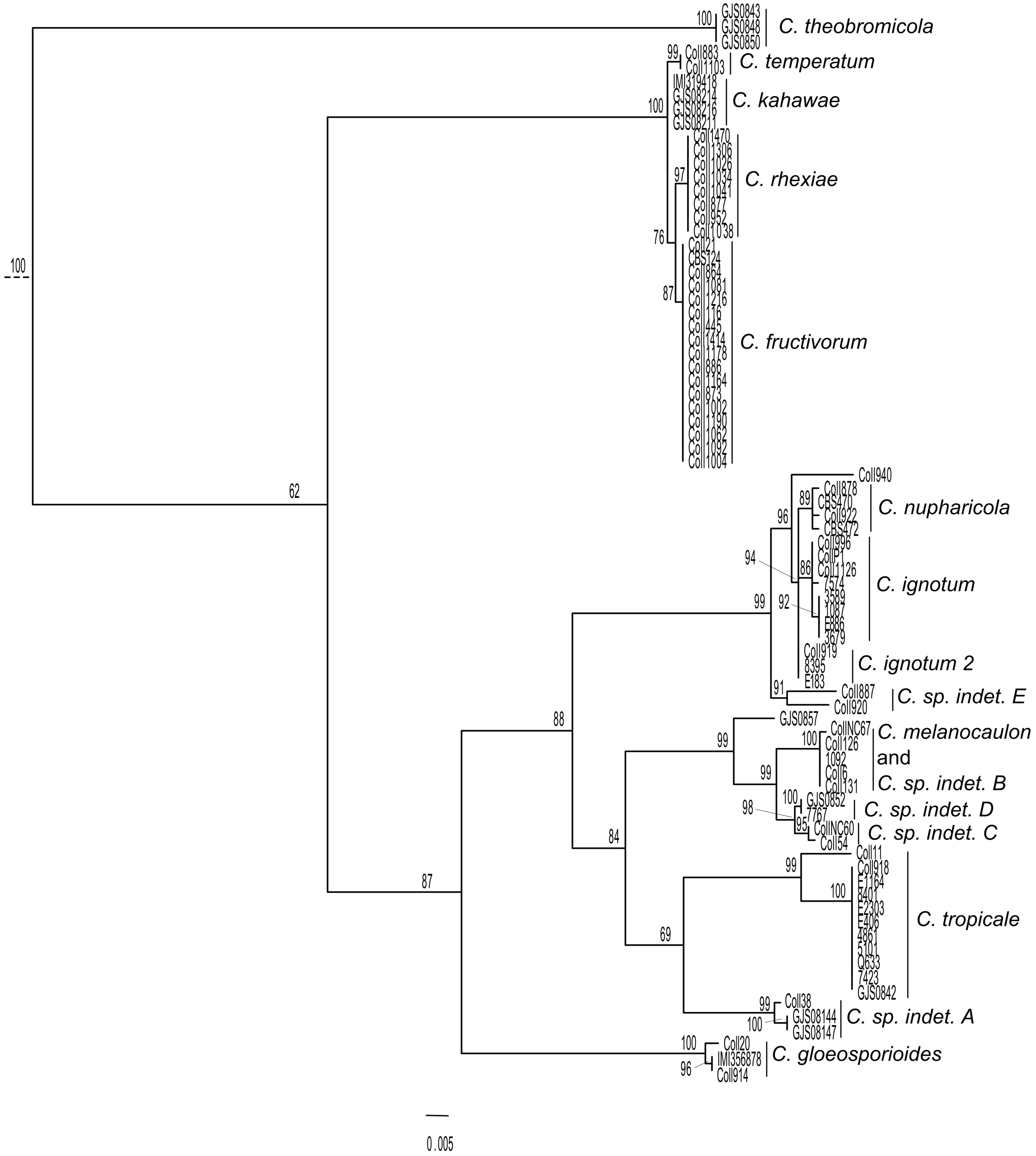


Figure S7: legend on next page

Figure S7: Maximum-likelihood bootstrap consensus tree of *apn2* gene. Boostrap support values shown above branches. Outgroups (*C. aff. acutatum* and strains 4766, 3386, and 4801) have been trimmed from the tree.

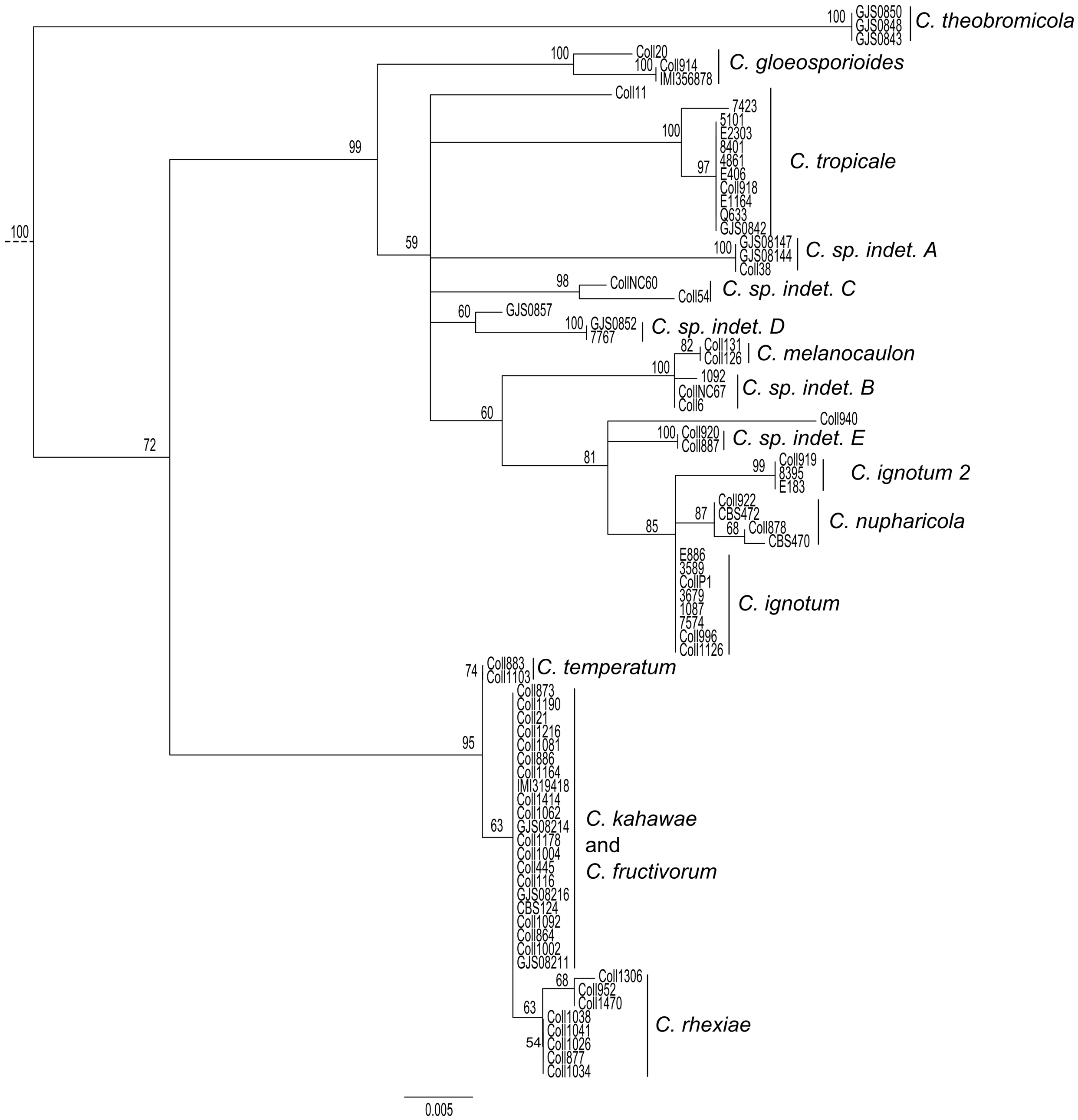


Figure S8: legend on next page

Figure S8: Maximum-likelihood bootstrap consensus tree of *apn2/matIGS* gene. Boostrap support values shown above branches. Outgroups (*C. aff. acutatum* and strains 4766, 3386, and 4801) have been trimmed from the tree.

APPENDIX 4A

MAPS

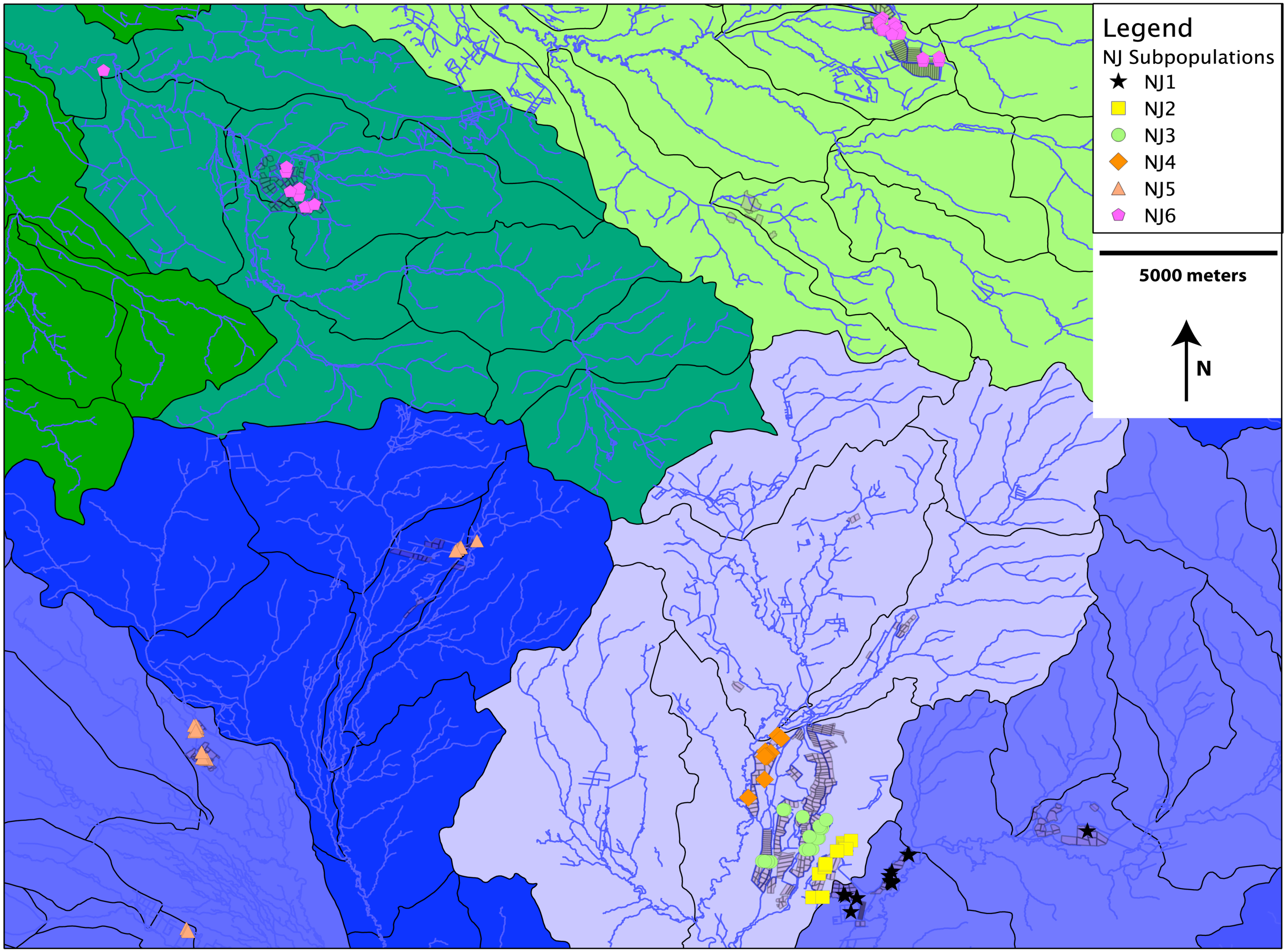

Legend
NJ Subpopulations
★ NJ1
NJ2
NJ3
NJ4
NJ5
NJ6
5000 meters
N

Figure 1: Spatial distribution of New Jersey subpopulations across two major watersheds. Green hued sections represent the Rancocas watershed. Blue hued sections represent the Mullica watershed. NJ_2 through NJ_4 are each part of different irrigation networks despite being within the same subwatershed.

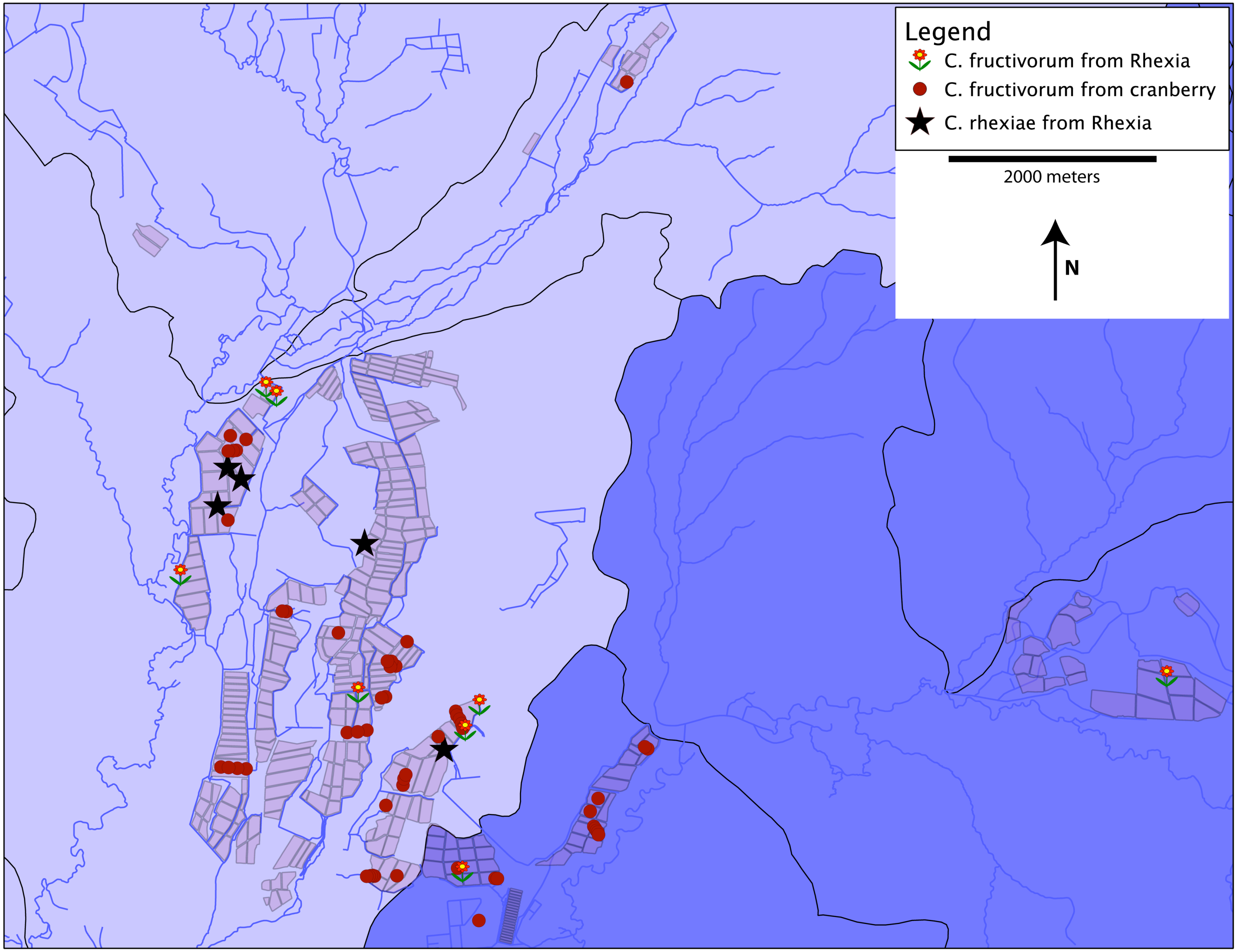

Legend
C. fructivorum from Rhexia
C. fructivorum from cranberry
C. rhexiae from Rhexia
2000 meters
N

Figure 2: Spatial distribution of *Colletotrichum fructivorum* and *C. rhexiae* where their distribution overlaps in New Jersey. Red dots represent sample points from which *C. fructivorum* was isolated from *Vaccinium macrocarpon.* Flower symbols represent points from which *C. fructivorum* was isolated from *Rhexia virginica*. Black stars represent points from which *C. rhexiae* was isolated from *R. virginica*. Multiple isolates were genotyped from some individual sample points.

APPENDIX 4B

SUPPLEMENTARY TABLES

Supplementary Table 1: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by regional subpopulation. Each of the 100 clone-corrected regional datasets was constructed by resampling to the size of the smallest clone-corrected population and subpopulation used as a factor in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	56.1 (55.3-56.89)	6.23 (6.14-6.32)	9	0.23 (0.22-0.23)	0.00002	6.13
Error	387.44 (385.77-389.11)	3.52 (3.51-3.54)	110	3.52 (3.51-3.54)		93.87
Total	443.54 (441.72-445.36)	3.73 (3.71-3.74)	119			

Supplementary Table 2: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by cultivar. Each of the 100 clone-corrected cultivar datasets was constructed by resampling to the size of the smallest clone-corrected population and cultivar used as a factor in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	4.64 (4.59-4.69)	4.64 (4.59-4.69)	1	0.02 (0.02-0.02)	0.23498	0.52
Error	325.84 (325.57-326.1)	3.79 (3.79-3.79)	86	3.79 (3.79-3.79)		99.48
Total	330.48 (330.18-330.77)	3.8 (3.8-3.8)	87			

Supplementary Table 3: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by host organ. Each of the 100 clone-corrected host organ datasets was constructed by resampling to the size of the smallest clone-corrected population and host organ used as a factor in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	3.54 (3.33-3.76)	3.54 (3.33-3.76)	1	-0.05 (-0.07--0.03)	0.59667	0
Error	81.37 (80.63-82.1)	4.07 (4.03-4.11)	20	4.07 (4.03-4.11)		100
Total	84.91 (84.24-85.58)	4.04 (4.01-4.08)	21			

Supplementary Table 4: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by subwatershed. Each of the 100 clone-corrected subwatershed datasets was constructed by resampling to the size of the smallest clone-corrected population and subwatershed used as factors in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	19.81 (19.33-20.29)	3.96 (3.87-4.06)	5	0.05 (0.04-0.06)	0.30659	1.39
Error	149.28 (148.05-150.5)	3.55 (3.53-3.58)	42	3.55 (3.53-3.58)		98.61
Total	169.09 (167.79-170.38)	3.6 (3.57-3.63)	47			

Supplementary Table 5: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by time period. Each of the 100 clone-corrected temporal datasets was constructed by resampling to the size of the smallest clone-corrected population and time period used as a factor in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	3.32 (3.14-3.49)	3.32 (3.14-3.49)	1	-0.01 (-0.02--0.01)	0.56493	0
Error	167.02 (166-168.04)	3.63 (3.61-3.65)	46	3.63 (3.61-3.65)		100
Total	170.34 (169.26-171.41)	3.62 (3.6-3.65)	47			

Supplementary Table 6: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by host species. Each of the 100 clone-corrected host species datasets was constructed by resampling to the size of the smallest clone-corrected population and host species used as a factor in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	4.37 (4.14-4.59)	4.37 (4.14-4.59)	1	0.08 (0.06-0.1)	0.27543	2.31
Error	74.37 (73.7-75.04)	3.38 (3.35-3.41)	22	3.38 (3.35-3.41)		97.69
Total	78.73 (77.98-79.49)	3.42 (3.39-3.46)	23			

Supplementary Table 7: Analysis of molecular variance (AMOVA) on clone-corrected datasets resampled by watershed. Each of the 100 clone-corrected watershed datasets was constructed by resampling to the size of the smallest clone-corrected population and watershed used as a factor in the analysis.

	SSD	MDS	df	σ^2	*p*-value	% variation
Pops	5.65 (5.43-5.86)	5.65 (5.43-5.86)	1	0.06 (0.06-0.07)	0.13553	1.56
Error	211.67 (210.62-212.72)	3.78 (3.76-3.8)	56	3.78 (3.76-3.8)		98.44
Total	217.32 (216.31-218.33)	3.81 (3.79-3.83)	57			

Supplementary Table 8: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from clone-corrected datasets resampled by cultivar. Each of the 100 datasets was resampled to the size of the smallest population among cultivars.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.414 (0.414-0.415)	0.416 (0.415-0.417)	0.003 (0.002-0.003)	0.004 (0.004-0.005)
L10D10	0.505 (0.503-0.506)	0.506 (0.505-0.508)	0.003 (0.003-0.003)	0.006 (0.006-0.007)
L14F4	0.717 (0.717-0.718)	0.719 (0.718-0.72)	0.003 (0.003-0.003)	0.014 (0.013-0.015)
L2c1	0.186 (0.184-0.187)	0.186 (0.185-0.188)	0.003 (0.003-0.003)	0.001 (0.001-0.002)
LB5B4	0.045 (0.044-0.045)	0.044 (0.044-0.045)	-0.006 (-0.006--0.005)	-0.001 (-0.001--0.001)
LC1942	0.76 (0.76-0.761)	0.764 (0.763-0.765)	0.005 (0.005-0.005)	0.031 (0.03-0.032)
LC2090	0.409 (0.408-0.41)	0.408 (0.407-0.409)	-0.002 (-0.002--0.002)	-0.003 (-0.003--0.003)
LC4168	0.38 (0.379-0.381)	0.38 (0.379-0.382)	0.002 (0.001-0.002)	0.002 (0.002-0.002)
LF9	0.329 (0.328-0.33)	0.352 (0.351-0.353)	0.065 (0.064-0.066)	0.068 (0.068-0.069)
global	0.416 (0.416-0.416)	0.42 (0.419-0.42)	0.008 (0.008-0.008)	0.012 (0.012-0.012)

Supplementary Table 9: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from clone-corrected datasets resampled by host organ. Each of the 100 datasets was resampled to the size of the smallest population among stem and fruit isolates.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.357 (0.346-0.369)	0.367 (0.353-0.381)	0.019 (0.01-0.029)	0.033 (0.018-0.048)
L10D10	0.653 (0.639-0.667)	0.684 (0.674-0.694)	0.048 (0.04-0.056)	0.16 (0.138-0.181)
L14F4	0.786 (0.775-0.798)	0.797 (0.789-0.806)	0.015 (0.01-0.02)	0.079 (0.058-0.101)
L2C1	0.067 (0.055-0.079)	0.07 (0.057-0.083)	0.024 (0.018-0.029)	0.008 (0.005-0.012)
LB5B4	0.026 (0.017-0.034)	0.026 (0.018-0.035)	0.008 (0.005-0.010)	0.002 (0.001-0.003)
LC1942	0.794 (0.785-0.802)	0.804 (0.796-0.811)	0.013 (0.009-0.017)	0.089 (0.063-0.114)
LC2090	0.467 (0.453-0.481)	0.473 (0.46-0.486)	0.016 (0.01-0.022)	0.02 (0.011-0.028)
LC4168	0.334 (0.318-0.349)	0.336 (0.32-0.351)	0.007 (0.001-0.012)	0.008 (0.001-0.014)
LF9	0.395 (0.382-0.408)	0.394 (0.382-0.406)	0 (-0.004-0.005)	-0.004 (-0.008-0)
global	0.431 (0.427-0.435)	0.439 (0.436-0.443)	0.019 (0.016-0.021)	0.028 (0.024-0.032)

Supplementary Table 10: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from clone-corrected datasets resampled by subwatershed. Each of the 100 datasets was resampled to the size of the smallest population among subwatersheds.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.374 (0.366-0.382)	0.396 (0.389-0.403)	0.056 (0.046-0.065)	0.042 (0.035-0.048)
L10D10	0.511 (0.5-0.521)	0.571 (0.561-0.582)	0.106 (0.098-0.114)	0.149 (0.138-0.161)
L14F4	0.664 (0.657-0.672)	0.713 (0.706-0.72)	0.068 (0.063-0.073)	0.174 (0.162-0.186)
L2C1	0.182 (0.174-0.189)	0.191 (0.183-0.198)	0.047 (0.041-0.053)	0.013 (0.011-0.015)
LB5B4	0.015 (0.011-0.019)	0.016 (0.011-0.021)	0.018 (0.014-0.024)	0.001 (0.001-0.001)
LC1942	0.73 (0.724-0.737)	0.772 (0.765-0.779)	0.054 (0.049-0.059)	0.185 (0.168-0.202)
LC2090	0.301 (0.29-0.312)	0.327 (0.315-0.34)	0.079 (0.072-0.086)	0.046 (0.041-0.052)
LC4168	0.33 (0.319-0.342)	0.338 (0.327-0.35)	0.024 (0.019-0.03)	0.015 (0.011-0.018)
LF9	0.213 (0.203-0.222)	0.235 (0.224-0.246)	0.092 (0.082-0.101)	0.035 (0.03-0.04)
global	0.369 (0.366-0.372)	0.396 (0.392-0.399)	0.067 (0.065-0.07)	0.051 (0.049-0.053)

Supplementary Table 11: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_{S}), and total gene diversity (H_{T}) from clone-corrected datasets resampled by time period. Each of the 100 datasets was resampled to the size of the smallest population among temporally isolated populations.

	H_{S}	H_{T}	G_{ST}	D_{est}
Cgret	0.392 (0.384-0.4)	0.393 (0.384-0.401)	0.001 (-0.002-0.004)	0.001 (-0.003-0.005)
L10D10	0.48 (0.468-0.491)	0.484 (0.472-0.496)	0.008 (0.006-0.011)	0.019 (0.014-0.025)
L14F4	0.728 (0.722-0.734)	0.739 (0.731-0.746)	0.014 (0.011-0.016)	0.083 (0.066-0.1)
L2C1	0.112 (0.104-0.121)	0.113 (0.104-0.121)	0.002 (0-0.004)	0.001 (0-0.002)
LB5B4	0.013 (0.009-0.018)	0.014 (0.009-0.018)	0.004 (0.003-0.005)	0 (0-0.001)
LC1942	0.803 (0.798-0.808)	0.813 (0.809-0.817)	0.012 (0.01-0.014)	0.097 (0.084-0.11)
LC2090	0.384 (0.375-0.392)	0.384 (0.377-0.392)	0.003 (0.001-0.005)	0.002 (0-0.005)
LC4168	0.353 (0.344-0.361)	0.353 (0.344-0.362)	0.002 (0-0.003)	0.002 (0-0.004)
LF9	0.29 (0.281-0.3)	0.295 (0.284-0.306)	0.013 (0.009-0.017)	0.015 (0.009-0.02)
global	0.395 (0.393-0.397)	0.399 (0.396-0.401)	0.009 (0.008-0.01)	0.012 (0.011-0.013)

Supplementary Table 12: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from clone-corrected datasets resampled by host species. Each of the 100 datasets was resampled to the size of the smallest population among host species.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.429 (0.418-0.439)	0.428 (0.419-0.437)	0.002 (-0.005-0.009)	-0.003 (-0.01-0.005)
L10D10	0.292 (0.279-0.305)	0.343 (0.324-0.361)	0.135 (0.125-0.145)	0.152 (0.132-0.171)
L14F4	0.672 (0.663-0.68)	0.683 (0.673-0.694)	0.016 (0.011-0.02)	0.077 (0.054-0.1)
L2C1	0.154 (0.142-0.166)	0.156 (0.143-0.168)	0.011 (0.008-0.015)	0.005 (0.002-0.007)
LB5B4	0.014 (0.008-0.021)	0.015 (0.008-0.021)	0.004 (0.002-0.006)	0.001 (0-0.001)
LC1942	0.81 (0.802-0.818)	0.839 (0.834-0.845)	0.035 (0.031-0.04)	0.287 (0.26-0.314)
LC2090	0.291 (0.277-0.305)	0.304 (0.287-0.321)	0.034 (0.026-0.041)	0.041 (0.031-0.052)
LC4168	0.259 (0.245-0.274)	0.268 (0.251-0.284)	0.025 (0.019-0.031)	0.025 (0.018-0.033)
LF9	0.312 (0.299-0.325)	0.312 (0.298-0.325)	-0.001 (-0.005-0.003)	0 (-0.005-0.005)
global	0.359 (0.356-0.362)	0.372 (0.368-0.376)	0.034 (0.032-0.037)	0.04 (0.037-0.043)

Supplementary Table 13: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from clone-corrected datasets resampled by watershed. Each of the 100 datasets was resampled to the size of the smallest population among watersheds.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.327 (0.321-0.333)	0.339 (0.331-0.346)	0.031 (0.026-0.037)	0.036 (0.029-0.043)
L10D10	0.622 (0.613-0.632)	0.652 (0.645-0.659)	0.047 (0.042-0.051)	0.152 (0.143-0.161)
L14F4	0.766 (0.76-0.772)	0.774 (0.769-0.779)	0.011 (0.009-0.013)	0.065 (0.053-0.076)
L2C1	0.108 (0.101-0.116)	0.109 (0.101-0.117)	0.005 (0.003-0.007)	0.002 (0.001-0.003)
LB5B4	0.015 (0.011-0.019)	0.015 (0.011-0.019)	0.004 (0.003-0.006)	0 (0-0.001)
LC1942	0.827 (0.823-0.831)	0.831 (0.827-0.834)	0.004 (0.003-0.006)	0.037 (0.026-0.047)
LC2090	0.432 (0.424-0.44)	0.432 (0.424-0.439)	-0.001 (-0.003-0.001)	-0.003 (-0.005--0.001)
LC4168	0.346 (0.338-0.354)	0.348 (0.339-0.357)	0.005 (0.003-0.006)	0.006 (0.004-0.009)
LF9	0.276 (0.268-0.285)	0.284 (0.274-0.294)	0.023 (0.018-0.027)	0.022 (0.017-0.027)
global	0.413 (0.411-0.415)	0.42 (0.418-0.422)	0.017 (0.016-0.018)	0.024 (0.022-0.025)

Supplementary Table 14: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from datasets without clone correction resampled by regional subpopulation. Each of the 100 datasets was resampled to the size of the smallest population among regional subpopulations.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.38 (0.376-0.383)	0.412 (0.409-0.415)	0.078 (0.074-0.083)	0.057 (0.054-0.06)
L10D10	0.366 (0.361-0.37)	0.45 (0.445-0.455)	0.187 (0.181-0.192)	0.146 (0.141-0.15)
L14F4	0.529 (0.526-0.532)	0.585 (0.581-0.589)	0.094 (0.091-0.098)	0.13 (0.124-0.136)
L2C1	0.115 (0.111-0.118)	0.12 (0.116-0.124)	0.045 (0.04-0.049)	0.007 (0.006-0.008)
LB5B4	0.08 (0.078-0.082)	0.094 (0.092-0.097)	0.148 (0.14-0.157)	0.017 (0.016-0.018)
LC1942	0.596 (0.593-0.6)	0.737 (0.734-0.74)	0.191 (0.187-0.195)	0.384 (0.377-0.391)
LC2090	0.329 (0.325-0.333)	0.336 (0.331-0.34)	0.02 (0.018-0.022)	0.011 (0.01-0.012)
LC4168	0.291 (0.287-0.296)	0.473 (0.469-0.477)	0.386 (0.379-0.392)	0.283 (0.279-0.286)
LF9	0.161 (0.158-0.164)	0.198 (0.195-0.202)	0.19 (0.184-0.195)	0.049 (0.048-0.051)
global	0.316 (0.315-0.318)	0.378 (0.377-0.38)	0.164 (0.162-0.166)	0.1 (0.099-0.101)

Supplementary Table 15: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from datasets without clone correction resampled by cultivar. Each of the 100 datasets was resampled to the size of the smallest population among cultivars.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.369 (0.367-0.37)	0.367 (0.366-0.369)	-0.004 (-0.004--0.004)	-0.005 (-0.005--0.005)
L10D10	0.549 (0.547-0.551)	0.568 (0.566-0.57)	0.033 (0.032-0.034)	0.083 (0.082-0.084)
L14F4	0.628 (0.626-0.63)	0.645 (0.644-0.647)	0.027 (0.026-0.028)	0.094 (0.092-0.097)
L2c1	0.141 (0.14-0.141)	0.146 (0.145-0.146)	0.034 (0.034-0.035)	0.012 (0.011-0.012)
LB5B4	0.034 (0.033-0.034)	0.034 (0.033-0.034)	-0.004 (-0.004--0.004)	0 (0-0)
LC1942	0.733 (0.732-0.735)	0.763 (0.762-0.764)	0.039 (0.038-0.04)	0.223 (0.22-0.225)
LC2090	0.34 (0.337-0.342)	0.344 (0.342-0.345)	0.012 (0.011-0.013)	0.012 (0.011-0.013)
LC4168	0.295 (0.293-0.297)	0.295 (0.293-0.297)	0.001 (0.001-0.002)	0.001 (0.001-0.001)
LF9	0.267 (0.265-0.269)	0.274 (0.272-0.276)	0.026 (0.025-0.027)	0.02 (0.019-0.02)
global	0.373 (0.372-0.373)	0.382 (0.381-0.382)	0.024 (0.023-0.024)	0.029 (0.028-0.029)

Supplementary Table 16: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from datasets without clone correction resampled by time period. Each of the 100 datasets was resampled to the size of the smallest population among temporally isolated collections.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.361 (0.353-0.369)	0.362 (0.354-0.37)	0.002 (-0.001-0.005)	0.004 (0-0.007)
L10D10	0.453 (0.443-0.464)	0.456 (0.446-0.467)	0.006 (0.004-0.008)	0.012 (0.008-0.016)
L14F4	0.677 (0.669-0.684)	0.682 (0.674-0.689)	0.007 (0.006-0.009)	0.031 (0.025-0.036)
L2C1	0.094 (0.087-0.102)	0.095 (0.087-0.103)	0.002 (0-0.004)	0.001 (0-0.002)
LB5B4	0.014 (0.01-0.017)	0.014 (0.01-0.018)	0.004 (0.002-0.005)	0 (0-0)
LC1942	0.791 (0.786-0.797)	0.802 (0.798-0.807)	0.014 (0.012-0.016)	0.099 (0.086-0.113)
LC2090	0.362 (0.352-0.372)	0.365 (0.356-0.374)	0.01 (0.007-0.014)	0.01 (0.007-0.013)
LC4168	0.337 (0.328-0.346)	0.337 (0.329-0.346)	0.003 (0.001-0.005)	0.002 (0-0.003)
LF9	0.235 (0.225-0.244)	0.237 (0.226-0.247)	0.007 (0.004-0.01)	0.006 (0.004-0.008)
global	0.369 (0.366-0.372)	0.372 (0.369-0.375)	0.008 (0.007-0.009)	0.009 (0.008-0.01)

Supplementary Table 17: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from datasets without clone correction resampled by host organ. Each of the 100 datasets was resampled to the size of the smallest population among stem and fruit collections.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.342 (0.331-0.353)	0.345 (0.332-0.358)	0.004 (-0.004-0.011)	0.011 (0-0.022)
L10D10	0.626 (0.607-0.645)	0.665 (0.653-0.678)	0.065 (0.053-0.076)	0.183 (0.16-0.207)
L14F4	0.734 (0.721-0.746)	0.758 (0.749-0.767)	0.034 (0.027-0.041)	0.158 (0.134-0.182)
L2C1	0.064 (0.051-0.077)	0.068 (0.054-0.081)	0.022 (0.018-0.026)	0.008 (0.005-0.011)
LB5B4	0.015 (0.008-0.022)	0.016 (0.008-0.023)	0.005 (0.002-0.008)	0.001 (0-0.003)
LC1942	0.772 (0.762-0.782)	0.785 (0.778-0.793)	0.018 (0.013-0.023)	0.103 (0.078-0.127)
LC2090	0.437 (0.422-0.451)	0.446 (0.434-0.458)	0.026 (0.019-0.034)	0.03 (0.021-0.038)
LC4168	0.304 (0.289-0.32)	0.306 (0.291-0.322)	0.009 (0.003-0.014)	0.006 (0.001-0.011)
LF9	0.357 (0.342-0.371)	0.359 (0.347-0.372)	0.013 (0.006-0.02)	0.006 (0.001-0.012)
global	0.406 (0.401-0.41)	0.417 (0.413-0.42)	0.027 (0.023-0.03)	0.036 (0.032-0.041)

Supplementary Table 18: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_{S}), and total gene diversity (H_{T}) from datasets without clone correction resampled by subwatershed. Each of the 100 datasets was resampled to the size of the smallest population among subwatershed.

	H_{S}	H_{T}	G_{ST}	D_{est}
Cgret	0.352 (0.343-0.362)	0.377 (0.369-0.386)	0.067 (0.059-0.075)	0.045 (0.04-0.05)
L10D10	0.483 (0.471-0.494)	0.553 (0.542-0.563)	0.127 (0.118-0.137)	0.161 (0.15-0.173)
L14F4	0.624 (0.616-0.633)	0.673 (0.665-0.682)	0.072 (0.068-0.077)	0.157 (0.147-0.167)
L2C1	0.17 (0.162-0.177)	0.181 (0.173-0.188)	0.062 (0.055-0.069)	0.016 (0.014-0.018)
LB5B4	0.017 (0.013-0.021)	0.018 (0.014-0.022)	0.020 (0.016-0.025)	0.001 (0.001-0.001)
LC1942	0.692 (0.684-0.7)	0.746 (0.74-0.753)	0.073 (0.067-0.079)	0.208 (0.193-0.224)
LC2090	0.272 (0.262-0.282)	0.296 (0.285-0.308)	0.078 (0.071-0.086)	0.041 (0.036-0.046)
LC4168	0.293 (0.283-0.303)	0.302 (0.292-0.313)	0.029 (0.024-0.035)	0.016 (0.012-0.019)
LF9	0.187 (0.178-0.196)	0.207 (0.197-0.218)	0.095 (0.084-0.105)	0.03 (0.026-0.035)
global	0.343 (0.34-0.347)	0.373 (0.369-0.376)	0.078 (0.075-0.082)	0.053 (0.051-0.056)

Supplementary Table 19: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_{S}), and total gene diversity (H_{T}) from datasets without clone correction resampled by host species. Each of the 100 datasets was resampled to the size of the smallest population among host species.

	H_{S}	H_{T}	G_{ST}	D_{est}
Cgret	0.418 (0.408-0.429)	0.418 (0.41-0.427)	0.003 (-0.003-0.009)	-0.002 (-0.009-0.004)
L10D10	0.269 (0.254-0.285)	0.312 (0.291-0.332)	0.120 (0.109-0.130)	0.124 (0.106-0.143)
L14F4	0.631 (0.619-0.642)	0.633 (0.622-0.645)	0.004 (0.001-0.007)	0.017 (0.004-0.03)
L2C1	0.137 (0.125-0.149)	0.138 (0.126-0.15)	0.009 (0.006-0.013)	0.002 (0.001-0.003)
LB5B4	0.009 (0.004-0.014)	0.009 (0.004-0.014)	0.002 (0.001-0.004)	0 (0-0.001)
LC1942	0.794 (0.784-0.803)	0.828 (0.822-0.834)	0.042 (0.037-0.048)	0.308 (0.28-0.336)
LC2090	0.243 (0.225-0.26)	0.251 (0.231-0.27)	0.022 (0.014-0.029)	0.025 (0.016-0.035)
LC4168	0.222 (0.207-0.236)	0.228 (0.212-0.244)	0.021 (0.015-0.028)	0.019 (0.012-0.026)
LF9	0.282 (0.269-0.294)	0.282 (0.269-0.295)	0.001 (-0.004-0.006)	0 (-0.004-0.004)
global	0.334 (0.329-0.338)	0.344 (0.339-0.35)	0.03 (0.027-0.033)	0.032 (0.029-0.036)

Supplementary Table 20: Global and single locus estimates of D_{est}, G_{ST}, subpopulation gene diversity (H_S), and total gene diversity (H_T) from datasets without clone correction resampled by watershed. Each of the 100 datasets was resampled to the size of the smallest population among regional subpopulations.

	H_S	H_T	G_{ST}	D_{est}
Cgret	0.334 (0.329-0.339)	0.344 (0.338-0.351)	0.027 (0.022-0.031)	0.031 (0.025-0.036)
L10D10	0.569 (0.559-0.579)	0.597 (0.589-0.605)	0.049 (0.044-0.053)	0.127 (0.119-0.134)
L14F4	0.695 (0.689-0.701)	0.712 (0.708-0.717)	0.025 (0.022-0.028)	0.11 (0.099-0.12)
L2C1	0.095 (0.088-0.102)	0.096 (0.089-0.103)	0.005 (0.003-0.007)	0.002 (0.001-0.002)
LB5B4	0.014 (0.011-0.018)	0.015 (0.011-0.018)	0 (0-0)	0 (0-0)
LC1942	0.784 (0.778-0.789)	0.787 (0.782-0.791)	0.004 (0.003-0.005)	0.026 (0.019-0.033)
LC2090	0.358 (0.35-0.367)	0.359 (0.351-0.367)	0.002 (0-0.004)	0.001 (0-0.003)
LC4168	0.289 (0.281-0.297)	0.29 (0.282-0.298)	0.004 (0.002-0.005)	0.004 (0.002-0.005)
LF9	0.237 (0.23-0.245)	0.242 (0.234-0.251)	0.018 (0.014-0.022)	0.015 (0.011-0.019)
global	0.375 (0.372-0.378)	0.382 (0.38-0.385)	0.019 (0.018-0.021)	0.024 (0.022-0.025)

Supplementary Table 21: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by regional subpopulation. Each of the 100 regional datasets was constructed by resampling to the size of the smallest clone-corrected population and subpopulation used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	112.68 (111.36-114)	11.27 (11.14-11.4)	10	0.52 (0.51-0.53)	0	15.03
Error	485.16 (482.97-487.35)	2.94 (2.93-2.95)	165	2.94 (2.93-2.95)		84.97
Total	597.84 (595.6-600.08)	3.42 (3.4-3.43)	175			

Supplementary Table 22: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by cultivar. Each of the 100 datasets was constructed by resampling to the size of the smallest clone-corrected population and cultivar used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	11.12 (10.99-11.25)	11.12 (10.99-11.25)	1	0.13 (0.13-0.14)	0.00017	3.70
Error	385.38 (384.77-386)	3.38 (3.38-3.39)	114	3.38 (3.38-3.39)		96.30
Total	396.51 (395.99-397.03)	3.45 (3.44-3.45)	115			

Supplementary Table 23: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by host organ. Each of the 100 regional datasets was constructed by resampling to the size of the smallest clone-corrected population and host organ used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	3.96 (3.71-4.21)	3.96 (3.71-4.21)	1	0.01 (-0.01-0.04)	0.46315	0.26
Error	76.59 (75.68-77.5)	3.83 (3.78-3.88)	20	3.83 (3.78-3.88)		99.74
Total	80.55 (79.8-81.3)	3.84 (3.8-3.87)	21			

Supplementary Table 24: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by subwatershed. Each of the 100 regional datasets was constructed by resampling to the size of the smallest clone-corrected population and subwatershed used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	21.92 (21.34-22.51)	4.38 (4.27-4.5)	5	0.12 (0.11-0.14)	0.09586	3.53
Error	157.45 (155.88-159.02)	3.28 (3.25-3.31)	48	3.28 (3.25-3.31)		96.47
Total	179.37 (177.78-180.96)	3.38 (3.35-3.41)	53			

Supplementary Table 25: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by time period. Each of the 100 regional datasets was constructed by resampling to the size of the smallest clone-corrected population and time period used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	3.15 (2.99-3.31)	3.15 (2.99-3.31)	1	-0.01 (-0.01-0)	0.52767	0
Error	182.75 (181.19-184.32)	3.38 (3.36-3.41)	54	3.38 (3.36-3.41)		100
Total	185.9 (184.37-187.44)	3.38 (3.35-3.41)	55			

Supplementary Table 26: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by host species. Each of the 100 regional datasets was constructed by resampling to the size of the smallest clone-corrected population and host species used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	3.78 (3.53-4.04)	3.78 (3.53-4.04)	1	0.05 (0.03-0.07)	0.35006	1.57
Error	69.09 (68.14-70.04)	3.14 (3.1-3.18)	22	3.14 (3.1-3.18)		98.43
Total	72.88 (71.77-73.98)	3.17 (3.12-3.22)	23			

Supplementary Table 27: Analysis of molecular variance (AMOVA) on datasets without clone correction resampled by watershed. Each of the 100 regional datasets was constructed by resampling to the size of the smallest clone-corrected population and watershed used as a factor in the analysis.

	SSD	MDS	df	σ^2	p-value	% variation
Pops	6.34 (6.07-6.6)	6.34 (6.07-6.6)	1	0.09 (0.08-0.09)	0.0663	2.56
Error	225.93 (224.34-227.52)	3.42 (3.4-3.45)	66	3.42 (3.4-3.45)		97.44
Total	232.27 (230.83-233.7)	3.47 (3.45-3.49)	67			

Supplementary Table 28: D_{est} values of regional population structure on datasets without clone correction resampled by regional subpopulation. D_{est} values with 95% confidence intervals estimated from 100 datasets resampled to a standardized sample size corresponding to the sample size of the smallest population (WABC: n=16). Continued on the next page.

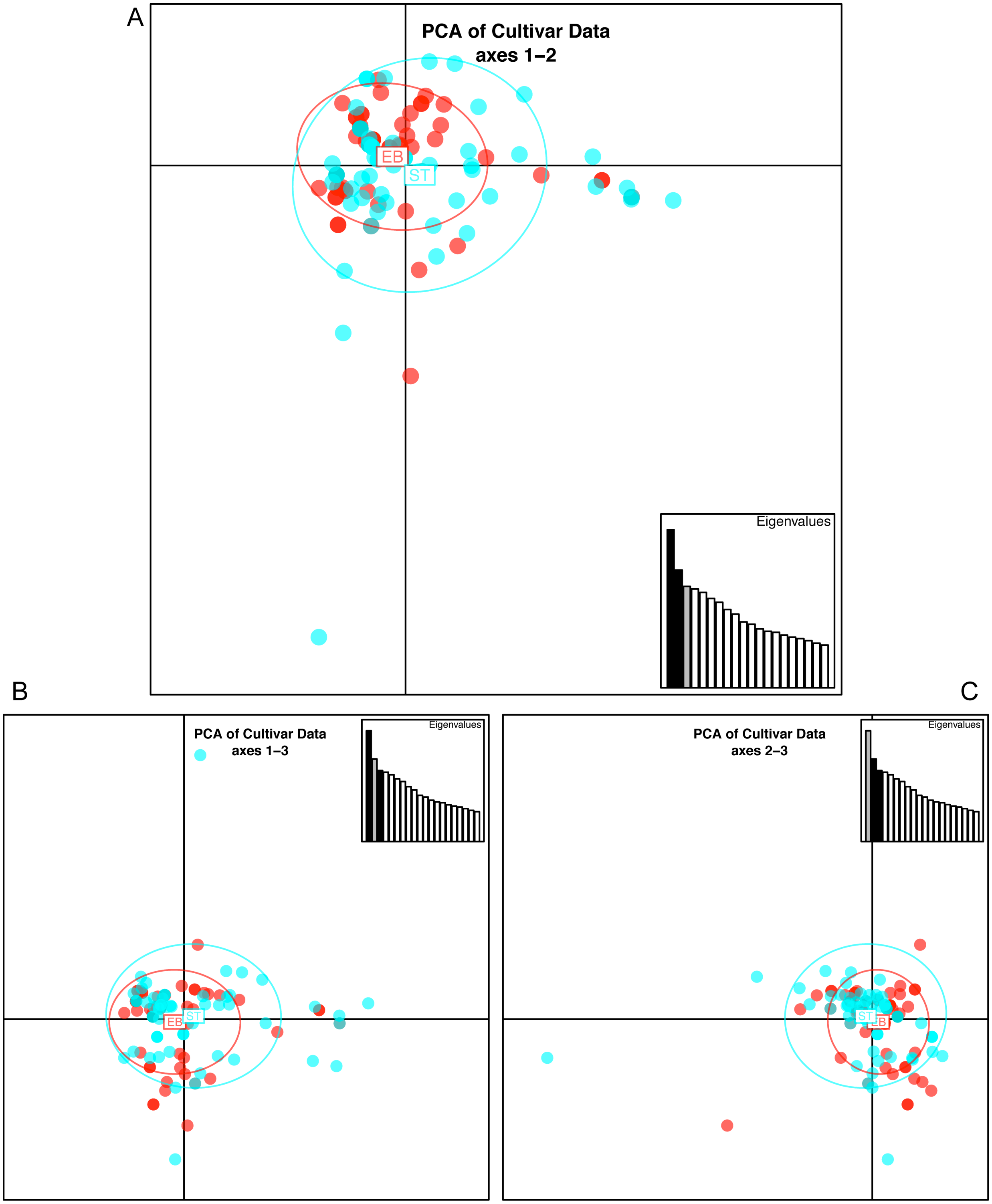
	NJ_1	NJ_2	NJ_3	NJ_4	NJ_5	NJ_6
NJ_1	0					
NJ_2	0.024 (0.022-0.026)	0				
NJ_3	0.059 (0.054-0.065)	0.074 (0.069-0.079)	0			
NJ_4	0.034 (0.032-0.036)	0.048 (0.046-0.05)	0.085 (0.08-0.091)	0		
NJ_5	0.019 (0.017-0.021)	0.023 (0.021-0.025)	0.064 (0.06-0.069)	0.049 (0.046-0.051)	0	
NJ_6	0.028 (0.025-0.031)	0.046 (0.042-0.049)	0.042 (0.038-0.045)	0.061 (0.058-0.064)	0.025 (0.022-0.028)	0
WABC	0.054 (0.052-0.056)	0.027 (0.026-0.028)	0.094 (0.089-0.098)	0.080 (0.080-0.080)	0.051 (0.049-0.053)	0.07 (0.066-0.073)
WI_1	0.079 (0.076-0.083)	0.101 (0.098-0.104)	0.173 (0.165-0.18)	0.088 (0.086-0.09)	0.09 (0.086-0.094)	0.118 (0.112-0.123)
WI_2	0.114 (0.111-0.117)	0.137 (0.135-0.14)	0.188 (0.182-0.193)	0.152 (0.151-0.153)	0.142 (0.138-0.145)	0.171 (0.167-0.175)
MA	0.052 (0.048-0.056)	0.042 (0.038-0.045)	0.06 (0.057-0.063)	0.095 (0.092-0.099)	0.047 (0.044-0.05)	0.046 (0.043-0.05)
DE	0.126 (0.122-0.129)	0.145 (0.142-0.148)	0.177 (0.172-0.182)	0.182 (0.18-0.184)	0.14 (0.137-0.144)	0.166 (0.162-0.17)

Supplementary Table 28: continued from the previous page.

	WABC	WI_1	WI_2	MA	DE
NJ_1					
NJ_2					
NJ_3					
NJ_4					
NJ_5					
NJ_6					
WABC	0				
WI_1	0.139 (0.136-0.141)	0			
WI_2	0.188 (0.187-0.19)	0.138 (0.135-0.141)	0		
MA	0.06 (0.058-0.063)	0.153 (0.148-0.158)	0.194 (0.19-0.199)	0	
DE	0.181 (0.178-0.184)	0.166 (0.162-0.17)	0.026 (0.025-0.028)	0.191 (0.186-0.195)	0

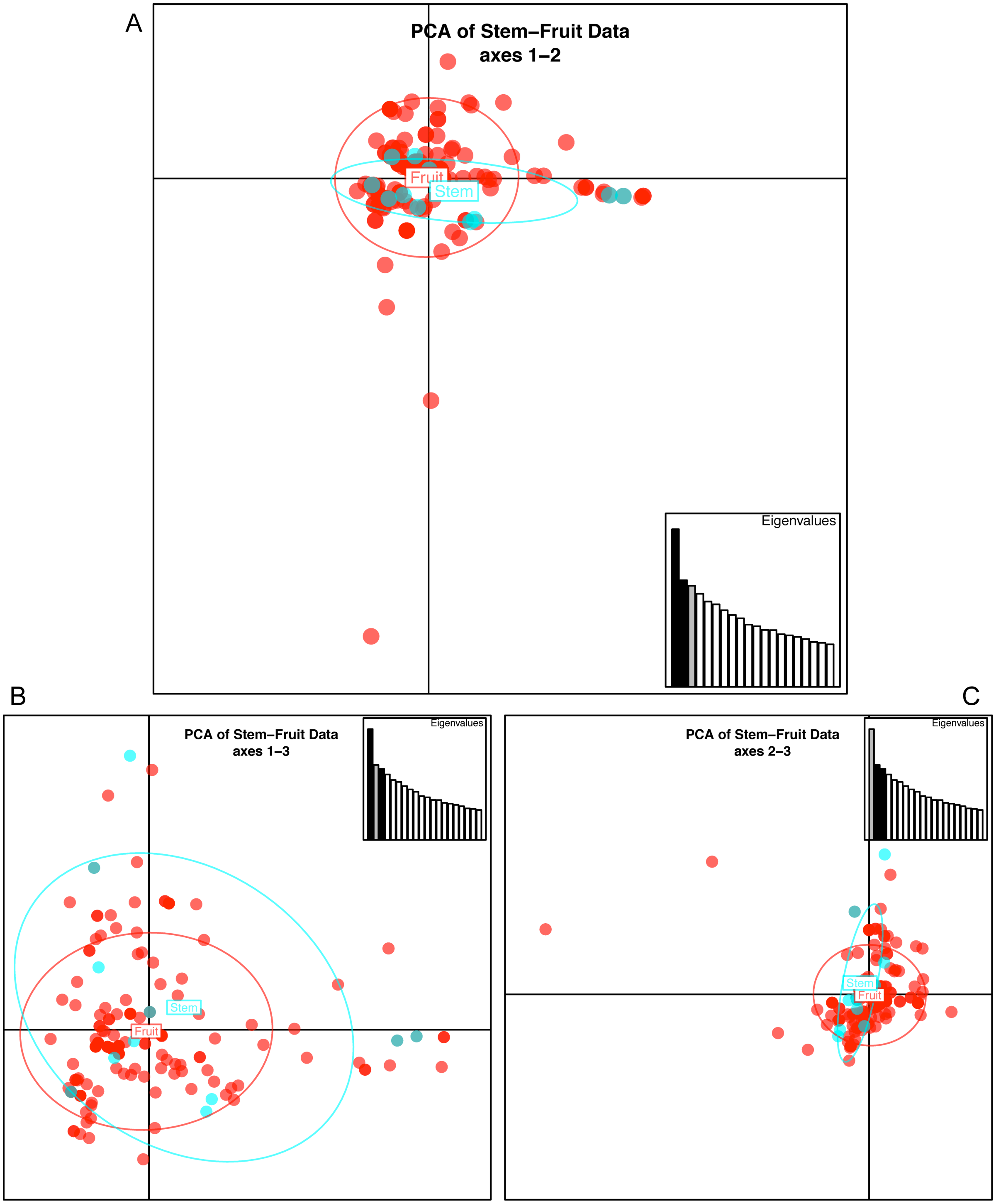
APPENDIX 4C

SUPPLEMENTARY FIGURES



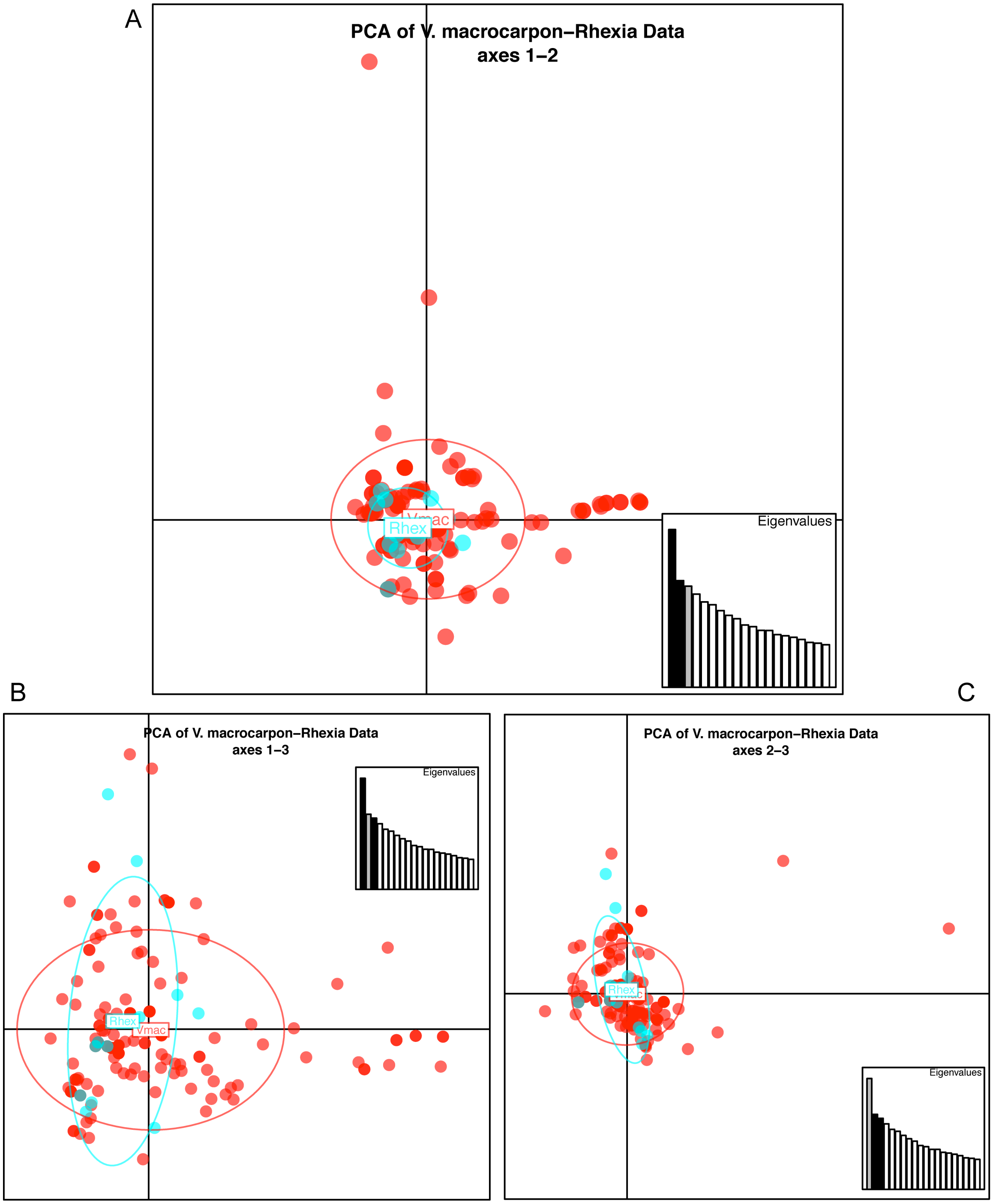
Supplemental Figure 1: legend on next page

Supplemental Figure 1: Principal components analysis of the cultivar data with cultivars as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.



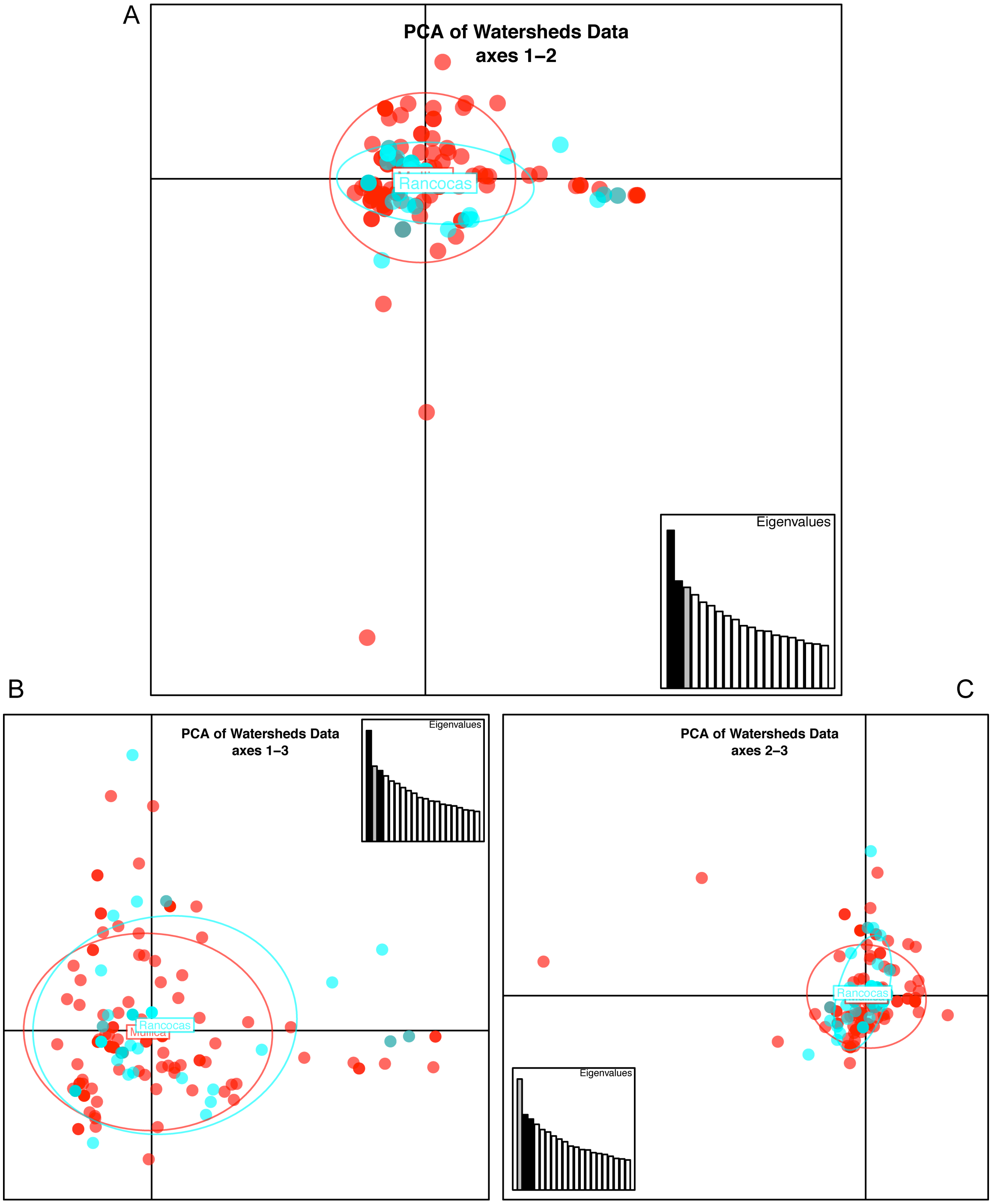
Supplemental Figure 2: legend on next page

Supplemental Figure 2: Principal components analysis of isolates from stem and fruit with host organ as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.



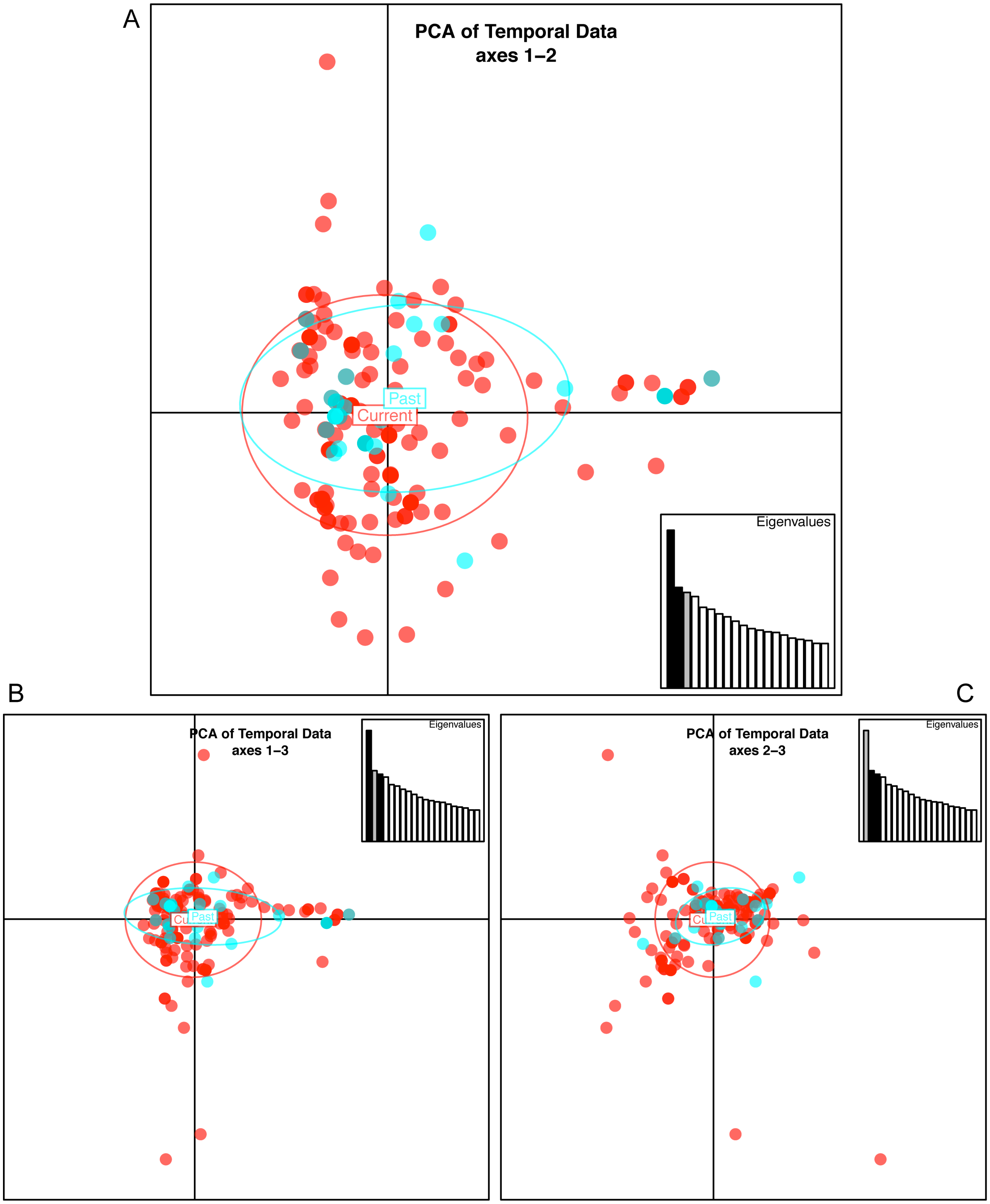
Supplemental Figure 3: legend on next page

Supplemental Figure 3: Principal components analysis of isolates from *Vaccinium macrocarpon* and *Rhexia virginica* with host species as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.



Supplemental Figure 4: legend on next page

Supplemental Figure 4: Principal components analysis of isolates from the Mullica and Rancocas watersheds with watersheds as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.



Supplemental Figure 5: legend on next page

Supplemental Figure 5: Principal components analysis of isolates from 1994-1997 and 2008-2010 with collection period as factors. Each population is represented by a unique color and a text label at the center of the inertia ellipse. Barplots represent the amount of variance accounted for by each PC. Black bars indicate which PC axes are plotted. A) Plot of PC axes 1 and 2. B) Plot of PC axes 1 and 3. C) Plot of PC axes 2 and 3.

APPENDIX 4D

R CODE

The following script is an exemplar that randomly samples rows from a table of genotypes based on the population identifier.

```
library(genetics)
library(fpc)
library(adegenet)
library(ade4)
library(pegas)
library(mmod)

##the following script randomly samples rows from a table of genotypes based on population identifier (in this case
##in the column labeled "Region") and combines the subsets into a table
##be sure to clear workspace before running this

A<-read.table("/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/DataFiles/AdeGenet/FructivorumSubPopulations.tab") #reads in a table
FstReps<-data.frame()
AmovaReps<-data.frame()
DSReps<-data.frame()
pairDReps<-data.frame()
for(i in 1:100){
B<-subset(A,Region =="NJ1")
C<-sample(nrow(B),16)
NJ1<-B[C,]
D<-subset(A,Region=="NJ2")
E<-sample(nrow(D),16)
NJ2<-D[E,]
F<-subset(A,Region=="NJ3")
G<-sample(nrow(F),16)
NJ3<-F[G,]
H<-subset(A,Region=="NJ4")
I<-sample(nrow(H),16)
NJ4<-H[I,]
J<-subset(A,Region=="NJ5")
K<-sample(nrow(J),16)
NJ5<-J[K,]
L<-subset(A,Region=="NJ6")
M<-sample(nrow(L),16)
NJ6<-L[M,]
N<-subset(A,Region=="WABC")
O<-sample(nrow(N),16)
```

```
WABC<-N[O,]
P<-subset(A,Region=="WI1")
Q<-sample(nrow(P),16)
WI1<-P[Q,]
R<-subset(A,Region=="WI2")
S<-sample(nrow(R),16)
WI2<-R[S,]
T<-subset(A,Region=="MA")
U<-sample(nrow(T),16)
MA<-T[U,]
V<-subset(A,Region=="DE")
W<-sample(nrow(V),16)
DE<-V[W,]
Resampled<-rbind(NJ1,NJ2,NJ3,NJ4,NJ5,NJ6,WABC,WI1,WI2,MA,DE)
ResampledRegions<-df2genind(Resampled[,-10], sep=NULL, ncode=3,
pop=as.factor(Resampled[,10]),ploidy=1, type="codom")
ResampledDist<-dist(truenames(ResampledRegions)$tab)
Pops<-factor(ResampledRegions$pop)
YY<-pegas::amova(ResampledDist~Pops, nperm=1000, is.squared=FALSE) #1-level AMOVA
matFstResampledRegions <-pairwise.fst(ResampledRegions)
FstReps<- rbind(FstReps,matFstResampledRegions)
print(FstReps)
Y<-
as.vector(c(YY[1]$tab[1,1],YY[1]$tab[1,2],YY[1]$tab[1,3],YY[1]$tab[2,1],YY[1]$tab[2,2],YY[
1]$tab[2,3],YY[1]$tab[3,1],YY[1]$tab[3,2],YY[1]$tab[3,3],YY[3]$varcomp[1,1],YY[3]$varco
mp[1,2],YY[3]$varcomp[2,1],YY[2]$varcoef[1]))
AmovaReps<-rbind(AmovaReps,Y)
print(AmovaReps)
DS<-diff_stats(ResampledRegions)
DSbind<-
as.vector(c(DS[[1]][1,],DS[[1]][2,],DS[[1]][3,],DS[[1]][4,],DS[[1]][5,],DS[[1]][6,],DS[[1]][7,],D
S[[1]][8,],DS[[1]][9,],DS[[2]]))
DSReps<-rbind(DSReps,DSbind)
pairD<-pairwise_D(ResampledRegions)
pairDReps<-rbind(pairDReps,pairD)
}
###
MeansFst<-colMeans(FstReps)
MeansAmov<-colMeans(AmovaReps)
MeansDS<-colMeans(DSReps)
sdFst<-numeric(0)  ##10 in 1:10 below is based on the number of columns, so if the number of
populations changes, the number 10 will have to change - get this value with ncol(FstReps)
for(i in 1:55){
SD<-sd(FstReps[,i])
sdFst<- append(sdFst,SD)
}
```

```
sdAmov<-numeric(0)
for(i in 1:13){
SDa<-sd(AmovaReps[,i])
sdAmov<- append(sdAmov,SDa)
}

sdDS<-numeric(0)
for(i in 1:51){
SDb<-sd(DSReps[,i])
sdDS<- append(sdDS,SDb)
}

sdpairD<-numeric(0)  ##10 in 1:55 below is based on the number of columns, so if the number
of populations changes, the number 10 will have to change - get this value with ncol(FstReps)
for(i in 1:55){
SDD<-sd(pairDReps[,i])
sdpairD<- append(sdpairD,SDD)
}


FstRepsTable<-rbind(FstReps,MeansFst,sdFst)
AmovRepsTable<-rbind(AmovaReps,MeansAmov,sdAmov)
DSRepsTable<-rbind(DSReps,MeansDS,sdDS)
PairDRepsTable<-rbind(pairDReps,MeansPairD, sdpairD)
write.table(FstRepsTable,
file="/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/FstRepsTable.txt")
write.table(AmovRepsTable,
file="/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/AmovRepsTable.txt")
write.table(DSRepsTable,
file="/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/DSRepsTable.txt")
write.table(PairDRepsTable,
file="/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/PairDRepsTableCCRegi
onsSubpops.txt")
```

The following script is an exemplar that calculates Jost's D with bootstrap confidence intervals for global comparison and builds dataframes for pairwise comparison.

```
library(genetics)
library(fpc)
library(adegenet)
library(ade4)
library(pegas)
library(mmod)

A<-
read.table("/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/DataFiles/AdeGenet/Clone
CorrectedFructivorumSubPopulations.txt") #reads in a table

AllRegions<-df2genind(A[,-10], sep=NULL, ncode=3, pop=as.factor(A[,10]),ploidy=1,
type="codom")

##global dataset
obs_DAllRegions<-D_Jost(AllRegions)
bsAllRegions<-chao_bootstrap(AllRegions,nreps=1000)
bs_DAllRegions<-summarise_bootstrap(bsAllRegions,D_Jost)
biasAllRegions<-bs_DAllRegions$summary.global.het[1]-obs_DAllRegions$global.het
summaryAllRegions_D<-bs_DAllRegions$summary.global.het-biasAllRegions
summaryAllRegions_D


X <- split(A, A$Region)  #split A into separate dataframes - now a list of dfs

DE<-X$DE  #split the list into separate dataframes
DE<-droplevels(DE)  ##remove all levels except DE
MA<-X$MA
MA<-droplevels(MA)
NJ1<-X$NJ1
NJ1<-droplevels(NJ1)
NJ2<-X$NJ2
NJ2<-droplevels(NJ2)
NJ3<-X$NJ3
NJ3<-droplevels(NJ3)
NJ4<-X$NJ4
NJ4<-droplevels(NJ4)
NJ5<-X$NJ5
NJ5<-droplevels(NJ5)
NJ6<-X$NJ6
NJ6<-droplevels(NJ6)
WABC<-X$WABC
```

```
WABC<-droplevels(WABC)
WI<-X$WI
WI<-droplevels(WI)

####Delaware with all others

#MA-DE
RegionsMADE<-rbind(MA,DE)
RegMADE<-df2genind(RegionsMADE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsMADE[,10]),ploidy=1, type="codom")

#NJ1-DE
RegionsNJ1DE<-rbind(NJ1,DE)
RegNJ1DE<-df2genind(RegionsNJ1DE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsNJ1DE[,10]),ploidy=1, type="codom")


#NJ2-DE
RegionsNJ2DE<-rbind(NJ2,DE)
RegNJ2DE<-df2genind(RegionsNJ2DE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsNJ2DE[,10]),ploidy=1, type="codom")

#NJ3-DE
RegionsNJ3DE<-rbind(NJ3,DE)
RegNJ3DE<-df2genind(RegionsNJ3DE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsNJ3DE[,10]),ploidy=1, type="codom")

#NJ4-DE
RegionsNJ4DE<-rbind(NJ4,DE)
RegNJ4DE<-df2genind(RegionsNJ4DE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsNJ4DE[,10]),ploidy=1, type="codom")

#NJ5-DE
RegionsNJ5DE<-rbind(NJ5,DE)
RegNJ5DE<-df2genind(RegionsNJ5DE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsNJ5DE[,10]),ploidy=1, type="codom")

#NJ6-DE
RegionsNJ6DE<-rbind(NJ6,DE)
RegNJ6DE<-df2genind(RegionsNJ6DE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsNJ6DE[,10]),ploidy=1, type="codom")

#WABC-DE
RegionsWABCDE<-rbind(WABC,DE)
RegWABCDE<-df2genind(RegionsWABCDE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsWABCDE[,10]),ploidy=1, type="codom")
```

```
#WI-DE
RegionsWIDE<-rbind(WI,DE)
RegWIDE<-df2genind(RegionsWIDE[,-10], sep=NULL, ncode=3,
pop=as.factor(RegionsWIDE[,10]),ploidy=1, type="codom")

####Delaware with all others


obs_DMADE<-D_Jost(RegMADE)
bsRegMADE<-chao_bootstrap(RegMADE,nreps=1000)
bs_DRegMADE<-summarise_bootstrap(bsRegMADE,D_Jost)
biasRegMADE<-bs_DRegMADE$summary.global.het[1]-obs_DMADE$global.het
summaryRegMADE_D<-bs_DRegMADE$summary.global.het-biasRegMADE
summaryRegMADE_D

obs_DNJ1DE<-D_Jost(RegNJ1DE)
bsRegNJ1DE<-chao_bootstrap(RegNJ1DE,nreps=1000)
bs_DRegNJ1DE<-summarise_bootstrap(bsRegNJ1DE,D_Jost)
biasRegNJ1DE<-bs_DRegNJ1DE$summary.global.het[1]-obs_DNJ1DE$global.het
summaryRegNJ1DE_D<-bs_DRegNJ1DE$summary.global.het-biasRegNJ1DE
summaryRegNJ1DE_D

obs_DNJ2DE<-D_Jost(RegNJ2DE)
bsRegNJ2DE<-chao_bootstrap(RegNJ2DE,nreps=1000)
bs_DRegNJ2DE<-summarise_bootstrap(bsRegNJ2DE,D_Jost)
biasRegNJ2DE<-bs_DRegNJ2DE$summary.global.het[1]-obs_DNJ2DE$global.het
summaryRegNJ2DE_D<-bs_DRegNJ2DE$summary.global.het-biasRegNJ2DE
summaryRegNJ2DE_D

obs_DNJ3DE<-D_Jost(RegNJ3DE)
bsRegNJ3DE<-chao_bootstrap(RegNJ3DE,nreps=1000)
bs_DRegNJ3DE<-summarise_bootstrap(bsRegNJ3DE,D_Jost)
biasRegNJ3DE<-bs_DRegNJ3DE$summary.global.het[1]-obs_DNJ3DE$global.het
summaryRegNJ3DE_D<-bs_DRegNJ3DE$summary.global.het-biasRegNJ3DE
summaryRegNJ3DE_D

obs_DNJ4DE<-D_Jost(RegNJ4DE)
bsRegNJ4DE<-chao_bootstrap(RegNJ4DE,nreps=1000)
bs_DRegNJ4DE<-summarise_bootstrap(bsRegNJ4DE,D_Jost)
biasRegNJ4DE<-bs_DRegNJ4DE$summary.global.het[1]-obs_DNJ4DE$global.het
summaryRegNJ4DE_D<-bs_DRegNJ4DE$summary.global.het-biasRegNJ4DE
summaryRegNJ4DE_D

obs_DNJ5DE<-D_Jost(RegNJ5DE)
bsRegNJ5DE<-chao_bootstrap(RegNJ5DE,nreps=1000)
```

```
bs_DRegNJ5DE<-summarise_bootstrap(bsRegNJ5DE,D_Jost)
biasRegNJ5DE<-bs_DRegNJ5DE$summary.global.het[1]-obs_DNJ5DE$global.het
summaryRegNJ5DE_D<-bs_DRegNJ5DE$summary.global.het-biasRegNJ5DE
summaryRegNJ5DE_D

obs_DNJ6DE<-D_Jost(RegNJ6DE)
bsRegNJ6DE<-chao_bootstrap(RegNJ6DE,nreps=1000)
bs_DRegNJ6DE<-summarise_bootstrap(bsRegNJ6DE,D_Jost)
biasRegNJ6DE<-bs_DRegNJ6DE$summary.global.het[1]-obs_DNJ6DE$global.het
summaryRegNJ6DE_D<-bs_DRegNJ6DE$summary.global.het-biasRegNJ6DE
summaryRegNJ6DE_D

obs_DWABCDE<-D_Jost(RegWABCDE)
bsRegWABCDE<-chao_bootstrap(RegWABCDE,nreps=1000)
bs_DRegWABCDE<-summarise_bootstrap(bsRegWABCDE,D_Jost)
biasRegWABCDE<-bs_DRegWABCDE$summary.global.het[1]-obs_DWABCDE$global.het
summaryRegWABCDE_D<-bs_DRegWABCDE$summary.global.het-biasRegWABCDE
summaryRegWABCDE_D

obs_DWIDE<-D_Jost(RegWIDE)
bsRegWIDE<-chao_bootstrap(RegWIDE,nreps=1000)
bs_DRegWIDE<-summarise_bootstrap(bsRegWIDE,D_Jost)
biasRegWIDE<-bs_DRegWIDE$summary.global.het[1]-obs_DWIDE$global.het
summaryRegWIDE_D<-bs_DRegWIDE$summary.global.het-biasRegWIDE
summaryRegWIDE_D

########GET THE STATS YOU NEED
###############################################################################
#########

summaryRegMADE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DMADE                    ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                    ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                    ####bootstrap locus estimates
bs_DRegMADE                         ##bootstraps estimates for each locus

###########
#########
##########
########
##########
#########
######
```

```
summaryRegNJ1DE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DNJ1DE                  ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                  ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                  ####bootstrap locus estimates
bs_DRegNJ1DE                      ##bootstraps estimates for each locus

###########
#########
##########
########
##########
#########
######

summaryRegNJ2DE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DNJ2DE                  ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                  ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                  ####bootstrap locus estimates
bs_DRegNJ2DE                      ##bootstraps estimates for each locus

###########
#########
##########
########
##########
#########

summaryRegNJ3DE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DNJ3DE                  ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                  ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                  ####bootstrap locus estimates
bs_DRegNJ3DE                      ##bootstraps estimates for each locus
###########
#########
##########
########
```

```
##########
#########
######
summaryRegNJ4DE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DNJ4DE                    ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                    ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                    ####bootstrap locus estimates
bs_DRegNJ4DE                        ##bootstraps estimates for each locus
###########
#########
##########
########
##########
#########
######

summaryRegNJ5DE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DNJ5DE                    ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                    ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                    ####bootstrap locus estimates
bs_DRegNJ5DE                        ##bootstraps estimates for each locus
###########
#########
##########
########
##########
#########
######
summaryRegNJ6DE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DNJ6DE                    ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                    ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                    ####bootstrap locus estimates
bs_DRegNJ6DE                        ##bootstraps estimates for each locus

###########
#########
##########
```

```
########
##########
#########
######
summaryRegWABCDE_D   ##the global value with confidence intervals based on 1000
bootstraps - the one to talk about
obs_DWABCDE                          ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                     ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                     ####bootstrap locus estimates
bs_DRegWABCDE                        ##bootstraps estimates for each locus
###########
#########
##########
########
##########
#########
######
summaryRegWIDE_D   ##the global value with confidence intervals based on 1000 bootstraps -
the one to talk about
obs_DWIDE                      ###observed values for each locus - to get locus confidence
intervals subtract the difference between
                                     ####the bootstrap mean - subtract this value from the lower
bound and add to the upper bound of the
                                     ####bootstrap locus estimates
bs_DRegWIDE                          ##bootstraps estimates for each locus
###########
```

The following script is an exemplar that performs Principle Components Analysis of multilocus haplotypes.

```
library(ade4)
library(adegenet)

A<-
read.table("/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/DataFiles/AdeGenet/Fructiv
orumSubPopulations.tab") #reads in a table


        #reads in a table
        #of genotypes with rows as individuals and columns as markers with the last column
assigning each individual to a population
A[1:10,] #look at the table to make sure you have the correct import
FructivorumNineLocusRegion<-df2genind(A[,-10], sep=NULL, ncode=3,
pop=as.factor(A[,10]),ploidy=1, type="codom")
        #converts dataframe into a genind object and assigns individuals to populations by
reading in the
        #first 9 columns of the dataframe A[,-9] as the dataset and using the last column to assign
        #individuals to populations with pop=as.factor(A[,10], sets the ploidy to haploid with
ploidy=1 and
        #the data type to codominant
        #for the output, use the next command

FructivorumNineLocusRegion$pop #check to see that populations have been assigned - warning!
this will list populations as P1,P2,etc.
FructivorumNineLocusRegion$pop.names #check that the actual names have been assigned to
each population, which


genind2df(FructivorumNineLocusRegion, sep = "|") #check to see that conversion was correct to
genind object
any(is.na(FructivorumNineLocusRegion$tab))    #are there missing values?
sum(is.na(FructivorumNineLocusRegion$tab))    #how many missing values?

X <- scaleGen(FructivorumNineLocusRegion, missing = "mean") #another way to replace
missing values
pca1 <- dudi.pca(X,cent=FALSE,scale=FALSE,scannf=FALSE,nf = 3)
col<-rainbow(length(levels(pop(FructivorumNineLocusRegion))))
s.class(pca1$li,pop(FructivorumNineLocusRegion),xax=1,yax=2,col=transp(col,
0.6),axesell=FALSE,cstar=0,cpoint=3,grid=FALSE)
```

```
###color coded points with regional ellipses of PC1 and 2
pdf("/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/AdeGenet/ScaleGenRegi
onsalSubPopsPCAaxes1-2ellipses.pdf")
col<-rainbow(length(levels(pop(FructivorumNineLocusRegion))))
s.class(pca1$li,pop(FructivorumNineLocusRegion),xax=1,yax=2,col=transp(col,
0.6),axesell=FALSE,cstar=0,cpoint=3,grid=FALSE)
title("PCA of Regional Data\naxes 1-2")
add.scatter.eig(pca1$eig[1:length(pca1$eig)],nf=3,xax=1,yax=2,posi="bottomright") #barplot of
eigenvalues overlaid with PC used in plot in bold
dev.off()

###color coded points with regional ellipses of PC1 and 3
pdf("/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/AdeGenet/ScaleGenRegi
onsalSubPopsPCAaxes1-3ellipses.pdf")
col<-rainbow(length(levels(pop(FructivorumNineLocusRegion))))
s.class(pca1$li,pop(FructivorumNineLocusRegion),xax=1,yax=3,col=transp(col,
0.6),axesell=FALSE,cstar=0,cpoint=3,grid=FALSE)
title("PCA of Regional Data\naxes 1-3")
add.scatter.eig(pca1$eig[1:20],nf=3,xax=1,yax=3,posi="topright") #barplot of eigenvalues
overlaid with PC used in plot in bold
dev.off()

###color coded points with regional ellipses of PC2 and 3
pdf("/Users/vinson/Dropbox/Manuscripts/PopulationGenetics/Routput/AdeGenet/ScaleGenRegi
onsalSubPopsPCAaxes2-3ellipses.pdf")
col<-rainbow(length(levels(pop(FructivorumNineLocusRegion))))
s.class(pca1$li,pop(FructivorumNineLocusRegion),xax=2,yax=3,col=transp(col,
0.6),axesell=FALSE,cstar=0,cpoint=3,grid=FALSE)
title("PCA of Regional Data\naxes 2-3")
add.scatter.eig(pca1$eig[1:20],nf=3,xax=2,yax=3,posi="topright") #barplot of eigenvalues
overlaid with PC used in plot in bold
dev.off()
```

BIBLIOGRAPHY

Abramoff MD, Magelhaes PJ, Ram SJ (2004) Image Processing with ImageJ. *Biophotonics International*, **11**, 36-42.

Adaskaveg JE, Forster H (2000) Occurrence and Management of Anthracnose Epidemics Caused by *Colletotrichum* Species on Tree Fruit Crops in California. In: *Colletotrichum: Host Specificity, Pathology and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 317-336. The American Phytopathogical Society Press, St. Paul, Minnesota.

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

Agostini JP, Timmer LW (1994) Population dynamics and survival of strains of Colletotrichum gloeosporioides on Citrus in Florida. *Phytopathology*, **84**, 420-425.

Alberto F (2009) MsatAllele_1.0: An R package to visualize the binning of microsatellite alleles. *Journal of Heredity*, **1003**, 394-397.

Antonovics J (2006) Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, **97**, 33-37.

Bailey JA, Jeger MJ (1992) Colletotrichum: Biology, Pathology and Control. Wallingford, UK: CAB International.

Bailey JA, O'Connell RJ, Pring RJ, Nash C (1992) Infection strategies of *Colletotrichum* species. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 88-120. CAB International, Surrey, UK.

Bazinet AL, Neel MC, Shaw KL, Cummings MP (In preparation) The genealogical sorting index: software and web site for quantifying exclusivity of lineages.

Bills G, Polishook JD (1992) Recovery of endophytic fungi from *Chamaecyparis thyoides*. *Sydowia*, **44**, 1-12.

Boland GJ (2004) Fungal viruses, hypovirulence, and biological control of *Sclerotinia* species. *Canadian Journal of Plant Pathology*, **26**, 6-18.

Brasier CM (1995) Episodic selection as a force in fungal microevolution, with special reference to clonal speciation and hybrid introgression. *Canadian Journal of Botany*, **73**, S1213-S1221.

Cannon P, Bridge PD (2000) Linking the past, present and future of *Colletotrichum* systematics. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Dov Prusky, Freeman S, Dickman MB), pp. 1-20. American Phytopathological Society, St. Paul, Minnesota.

Carleton KL, Streelman JT, Lee BY*, et al.* (2002) Rapid isolation of CA microsatellites from the tilapia genome. *Animal Genetics*, **33**, 140-144.

Chakraborty S, Fernandes CD, Charchar MJdA, Thomas MR (2002) Pathogenic variation in *Colletotrichum gloeosporioides* infecting

Stylosanthes spp. in a center of diversity in Brazil. *Phytopathology*, **92**, 553-562.

Chilton SJP, Lucas GB, Edgerton CW (1945) Genetics of *Glomerella* III: Crosses with a conidial strain. *American Journal of Botany*, **32**, 549-554.

Chilton SJP, Wheeler HE (1949a) Genetics of *Glomerella* VI: Linkage. *American Journal of Botany*, **36**, 270-273.

Chilton SJP, Wheeler HE (1949b) Genetics of *Glomerella* VII: Mutation and segregation in plus cultures. *American Journal of Botany*, **36**, 717-721.

Cisar CR, TeBeest DO (1999) Mating system of the filamentous ascomycete, *Glomerella cingulata*. *Current Genetics*, **35**, 127-133.

Cole RA, Lynn J (1996) Spatial and temporal population structure of *Brevicoryne brassicae*. *Entomologia Experimentalis et Applicata*, **78**, 121-127.

Colon-Garay J, Rivera-Vargas L, McGovern RJ, Rodriquez RdP (2002) Hypovirulent isolates of *Colletotrichum gloeosporioides* induce resistance to anthracnose in detached mango fruits and seedlings. *J. Agric. Univ. Puerto Rico*, **84**, 56-64.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.

Correll JC, Guerber JC, Wasliwa LA, Sherrill JF, Morelock TE (2000) Inter- and intra-species variation in *Colletotrichum* and mechanisms which affect population structure. In: *Colletotrichum: Host Specificity, Pathology and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 145-179. The American Phytopathogical Society Press, St. Paul, Minnesota.

Couch BC, Fudal I, Lebrun M-H*, et al.* (2005) Origins of host-specific populations of the blast pathogen *Magnaporthe oryzae* in crop domestication with subsequent expansion of pandemic clones on rice and weeds of rice. *Genetics*, **170**, 613-630.

Crouch JA, Clarke BB, Hillman BI (2006) Unraveling the evolutionary relationships among the divergent lineages of *Colletotrichum* causing anthracnose disease in turfgrass and corn. *Phytopathology*, **96**, 46-60.

Crouch JA, Clarke BB, Hillman BI (2009a) What is the value of ITS sequence data in *Colletotrichum* systematics and species diagnosis? A case study using the falcate spored graminicolous *Colletotrichum* group. *Mycologia*, **101**, 648-656.

Crouch JA, Tredway LP, Clarke BP, Hillman BI (2009b) Phylogenetic and population genetic divergence correspond with habitat for the pathogen *Colletotrichum cereale* and allied taxa across diverse grass communities. *Molecular Ecology*, **18**, 123-135.

Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution*, **62**, 2411-2422.

Dean R, Kan JALV, Pretorius ZA*, et al.* (2012) The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology*, **13**, 414-430.

Dettman JR, Anderson JB, Kohn LM (2008) Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus *Neurospora*. *BMC Evolutionary Biology*, **8**, 35.

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

Dettman JR, Sirjusingh C, Kohn LM, Anderson JB (2007) Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature*, **447**, 585-588.

Dilmaghani A, Gladieux P, Gout L*, et al.* (2012) Migration patterns and changes in population biology associated with worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. *Molecular Ecology*, **21**, 2519-2533.

Du M, Schardl CL, Nuckles EM, Vaillancourt LJ (2005) Using mating-type gene sequences for improved phylogenetic resolution of Colletotrichum species complexes. *Mycologia*, **97**, 641-658.

Dutech C, Enjalbert J, Fournier E*, et al.* (2007) Challenges of microsatellite isolation in fungi. *Fungal Genetics and Biology*, **44**, 933-949.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Eck P (1990) *The American Cranberry* Rutgers University Press, New Brunswick, NJ.

Edgerton CW (1914) Plus and minus strains in the Genus Glomerella. *American Journal of Botany*, **1**, 244-254.

Edgerton CW, Chilton SJP, Lucas GB (1945) Genetics of *Glomerella* II: Fertilization between strains. *American Journal of Botany*, **32**, 115-118.

Ellis JB, Everhart BM (1894) New Species of Fungi from Various Localities *Proceedings of the Academy of Natural Sciences of Philadelphia*, **46**, 322-386.

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.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587.

Farr DF, Rossman AY, Palm ME, McCray EB (2008) Fungal Databases. Systematic Mycology and Microbiology Laboratory, ARS, USDA.

Farr DF, Rossman AY, Palm ME, McCray EB (2011) Fungal Databases. http://nt.ars-grin.gov/fungaldatabases/ Systematic Mycology and Microbiology Laboratory, ARS, USDA

Foster S, Duke JA (2000) *A Field Guide to Medicinal Plants and Herbs of Eastern and Central North America*, 2nd edn. Houghton Mifflin, New York.

Freeman S, Katan T, Shabi E (1996) Characterization of Colletotrichum gloeosporioides isolates from avocado and almond fruits with molecular and pathogenicity tests. *Applied and Environmental Microbiology*, **62**, 1014-1020.

Fukuoka H, Nunome T, Minamiyama Y*, et al.* (2005) Read2Marker: a data processing tool for microsatellite marker development from a large data set. *Biotechniques*, **39**, 472, 474, 476.

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

Gladieux P, Guerin F, Giraud T*, et al.* (2011) Emergence of fungal pathogens by ecological speciation: importance of the reduced viability of imigrants. *Molecular Ecology*, **20**, 4521-4532.

Gleason H, Cronquist A (1991) *Manual of vascular Plants of Northeastern United States and Adjacent Canada* New York Botanical Garden, New York.

Goloboff PA, Farris JS, Nixon KC (2008) TNT, a free program for phylogenetic analysis. *Cladistics*, **24**, 774-786.

Gomez A, Carvalho GR (2000) Sex, parthenogenesis and genetic structure of rotifers: microsatellite analysis of contemporary and resting egg bank populations. *Molecular Ecology*, **9**, 203-214.

González E, Sutton TB, Correll JC (2006) Clarification of the etiology of Glomerella Leaf Spot and Bitter Rot of apple caused by Colletotrichum spp. based on morphology and genetic, molecular, and pathogenicity tests. *Phytopathology*, **96**, 982-992.

Guozhong L, Cannon PF, Reid A, Simmons CM (2004) Diversity and molecular relationships of endophytic *Colletotrichum* isolates from the Iwokrama Forest Reserve, Guyana. *Mycological Research*, **108**, 53-63.

Halsted B (1889) Some fungus diseases of the cranberry, pp. 1-40. New Jersey Agricultural College Experiment Station.

Hang VTT, Holmes EC, Veasna D*, et al.* (2010) Emergence of the Asian 1 Genotype of Dengue Virus Serotype 2 in Viet Nam: In Vivo Fitness Advantage and Lineage Replacement in South-East Asia. *PLOS Neglected Tropical Diseases*, **4**, e757.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633-1638.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, **290**, 516-518.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics*, **17**, 754-755.

Jakobsson M, Rosenberg N (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801-1806.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* **11**, 424-429.

Johnson-Cicalese J, Vorsa N, Polashock J (2009) Breeding for fruit rot resistance in *Vaccinium macrocarpon.* In: Proceedings of the IXth International Symposium on *Vaccinium*.

Johnson DA, Carris LM, Rogers JD (1997) Morphological and molecular characterization of *Colletotrichum nymphaeae* and *C. nupharicola* sp. nov. on water-lilies (*Nymphaea* and *Nuphar*). *Mycological Research*, **101**, 641-649.

Johnston PR, Jones D (1997) Relationships among *Colletotrichum* Isolates from fruit-rots assessed using rDNA sequences. *Mycologia*, **89**, 420-430.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403-1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Jombart T, Pontier D, Dufour AB (2009) Genetic markers in the playground of multivariate analysis. *Heredity*, **102**.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015-4026.

Jost L (2009) Reply: D vs G'st: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology*, **18**, 2088-2091.

Kareiva P, Watts S, McDonald R, Boucher T (2007) Domesticated Nature: Shaping Landscape and Ecosystems for Human Welfare. *Science*, **316**, 1866-1869.

Katoh K, Kuma K-i, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, **33**, 511-518.

Kaufmann PJ, Weidemann GJ (1996) Isozyme analysis of *Colletotrichum gloeosporioides* from five host genera. *Plant Disease*, **80**, 1289-1293.

Kendrick B (2000) *The Fifth Kingdom* Mycologue Publications, Sidney.

Kirk PM, Cannon PF, David JC, Stalpers JA (2001) Dictionary of the Fungi. CAB International, Surrey.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of poulation differentiation. *Conservation Genetics* **7**, 295-302.

Latunde-Dada AO (2001) Colletotrichum: tales of forcible entry, stealth, transient confinement and breakout. *Molecular Plant Pathology*, **2**, 187-198.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, **11**, 2445-2449.

Legard DE (2000) *Colletotrichum* diseases of Strawberry in Florida. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 292-299. American Phytopathological Society, St. Paul, Minnesota.

Lucas GB (1946) Genetics of *Glomerella* IV: Nuclear phenomena in the ascus. . *American Journal of Botany*, **33**, 802-806.

Lucas GB, Chilton SJP, Edgerton CW (1944) Genetics of *Glomerella* I: Studies on the behavior of certain strains. *American Journal of Botany*, **31**, 233-239.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998a) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, **89**, 238-247.

Luikart G, Cornuet JM, Allendorf FW (1999) Temporal changes in allele frequencies provide estimates of population bottleneck size. *Conservation Biology*, **13**.

Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998b) Usefullness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**.

Maccheroni WJ, Araújo WL, Azevedo JL (2004) Ambient pH-regulated enzyme secretion in endophytic and pathogenic isolates of the fungal genus *Colletotrichum*. *Scientia Agricola*, **61**, 298-302.

MacKenzie SJ (2005) *Population structure and pathogenicity of Colletotrichum gloeosporioides from strawberry and noncultivated hosts in Florida* Ph.D., University of Florida.

MacKenzie SJ, Seijo TE, Legard DE, Timmer LW, Peres NA (2007) Selection for pathogenicity to strawberry in populations of *Colletotrichum gloeosporioides* from native plants. *Phytopathology*, **97**, 1130-1140.

Martins WS, Lucas DCS, Neves KFS, Bertioli DJ (2009) WebSat - A web software for microsatellite marker development. *Bioinformation*, **3**, 282-283.

Maruyama T, Fuerst PA (1985) Population bottlenecks and nonequilibrium models in population genetics. II. Number of Alleles in a small population that was formed by a recent bottleneck. *Genetics*, **111**, 675-689.

McDonald BA, Linde C (2002a) The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica*, **124**, 163-180.

McDonald BA, Linde CC (2002b) Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, **40**, 349-379.

McGahen JW, Wheeler HE (1951) Genetics of *Glomerella* IX: Perithecial development and plasmogamy. . *American Journal of Botany*, **38**, 610-617.

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differntiation measure. *Evolution*, **60**, 2399-2402.

Meirmans PG, Hedrick PW (2011) Assessing population structure: Fst and related measures. *Molecular Ecology Resources*, **11**, 5-18.

Meirmans PG, Tienderen PHV (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity in asexual organsims. *Molecular Ecology Notes*, **4**, 792-794.

Milgroom MG (1996) Recombination and the multilocus structure of fungal populations. *Annual Review of Phytopathology*, **34**, 457-477

Miller MA, Pfeifer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees, 1-8.

Moraga Amador D, Farmerie B, Holland K*, et al.* (1998) Tools for Developing Microsatellite Markers: Microsatellite Libraries and AFLPs. Workshop Manual. University of Florida, Gainesville, Florida.

Morris CE, Bardin M, Kinkel LL*, et al.* (2009) Expanding the Paradigms of Plant Pathogen Life History and Evolution of Parasitic Fitness beyond Agricultural Boundaries. *PLoS Pathogens*, **5**, 1-7.

Nei M, Mauyama T, Chakraboty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.

Nicholson RL, Moraes WBC (1980) Survival of *Colletotrichum graminicola*: importance of the spore matrix. *Phytopathology*, **70**, 255-261.

Nirenberg H (1976) Untersuchungen u¨ ber die morphologische und biologische Differenzierung in der Fusari-

um-Sektion Liseola. *Mitteilungen aus der Biologischen Bundesanstalt für Land-und Forstwirtschaft*, **169**, 1-117.

Nylander JAA (2004) MrModeltest v2, Program distributed by the author. Evolutionary Biology Centre, Uppsala University. .

O'Connell RJ, Perfect S, Hughes B*, et al.* (2000) Dissecting the cell biology of *Colletotrichum* infection processes. In: *Colletotrichum: Host Specificity, Pahtology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 57-77. American Phytopathological Society, St. Paul, Minnesota.

O'Donnell K, Cigelnik E (1997) Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. *Molecular Phylogenetics and Evolution*, **7**, 103-116.

Olatinwo RO, Hanson EJ, Schilder AMC (2003) A First Assessment of the Cranberry Fruit Rot Complex in Michigan. *Plant Disease*, **87**, 550-556.

Oudemans PV (2008) Personal Communication.

Oudemans PV, Caruso FL, Stretch AW (1998) Cranberry fruit rot in the Northeast: A complex disease. *Plant Disease*, **82**, 1176-1184.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (Glycine max) simple sequence repeats (SSRs) within the genus and other legume genera: implications for transferability of SSRs in plants. *Molecular Biology and Evolution*, **15**, 1275-1287.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetics software for teaching and research. *Molecular Ecology Resources*, **6**, 288-295.

Photita W, Taylor PWJ, Ford R, Hyde KD (2005) Morphological and molecular characterization of Colletotrichum species from herbaceous plants in Thailand. *Fungal Diversity*, **18**, 117-133.

Phoulivong S, Cai L, Chen H*, et al.* (2010) *Colletotrichum gloeosporioides* is not a common pathogen on tropical fruits. *Fungal Diversity*, **44**, 33-43.

Polashock JJ, Caruso FL, Oudemans PV, McManus PS, Crouch JA (2009) The North American cranberry fruit rot fungal community: a systematic overview using morphological and phylogenetic affinities. *Plant Pathology*, **58**, 1116-1127.

Prihastuti H, Cai L, Chen H, McKenzie EHC, Hyde KD (2009) Characterization of *Colletotrichum* species associated with coffee berries in northern Thailand. *Fungal Diversity*, **39**, 89-109.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.

R-Development-Core-Team (2011) R: A Language and Environment for Statistical Computing.

Ramakrishnan U, Hadly EA, Mountain JL (2005) Detecting past population bottlenecks using temporal genetic data. *Molecular Ecology*, **14**, 2915-2922.

Rasband WS (1997-2008) ImageJ. National Institutes of Health, Bethesda, Maryland, USA.

Riccioni G, Landi M, Ferrara G*, et al.* (2010) Spatio-temporal population structuring and genetic diversity retention in depleted Atlantic Bluefin tuna of the Mediterranean Sea. *Proceedings of the National Academy of Sciences*, **107**, 2102-2107.

Roa AC, Chavarriaga-Aguirre P, Duque MC*, et al.* (2000) Cross-species amplification of Cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. *American Journal of Botany*, **87**, 1647-1655.

Rodriguez-Saona C, Vorsa N, Singh AP*, et al.* (2011) Tracing the history of plant traits under domestication in cranberries: potential consequences on anti-herbivore defences. *Journal of Experimental Botany*, **62**, 2633-2644.

Rodriguez RJ, Redman RS (2000) *Colletotrichum* as a model system for defining the genetic basis of fungal symbiont life styles. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 114-130. American Phytopathological Society, St. Paul, Minnesota.

Rojas EI, Rehner SA, Samuels GJ*, et al.* (2010) Colletotrichum gloeosporioides s.l. associated with Theobroma cacao and other plants in Panama´: multilocus phylogenies distinguish host-associated pathogens from asymptomatic endophytes. *Mycologia*, **102**, 1318-1338.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572-1574.

Rosenberg N (2004) DISTRUCT: a program for the grpahical display of population structure. *Molecular Ecology Notes*, **4**, 137-138.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ.

Ryman N, Leimar O (2009) Gst is still a useful measure of genetic differentiation - a comment on Jost's D. *Molecular Ecology*, **18**, 2084-2087.

Seifert KA (2009) Progress towards DNA barcoding of fungi. *Molecular Ecology Resources*, **9**, 83-89.

Shear CL (1907a) Cranberry Disease (ed. USDA). Government Printing Office.

Shear CL (1907b) Cranberry Diseases (ed. USDA). Government Printing Office.

Shear CL (1907c) New species of fungi. *The Bulletin of the Torrey Botanical Club*, **34**, 305-317.

Silva DN, Talhinas P, Cai L*, et al.* (2012a) Host-jump drives rapid and recent ecological speciation of the emergent fungal pathogen *Colletotrichum kahawae*. *Molecular Ecology*, **21**, 2655-2670.

Silva DN, Talhinas P, Várzea V*, et al.* (2012b) Application of the Apn2/MAT locus to improve the systematics of the Colletotrichum gloeosporioides complex: an example from coffee (Coffea spp.). *Mycologia*, **104**, 396-409.

Silvertown J, Servaes C, Biss P, Macleod D (2005) Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity*, **95**, 198-205.

Simmonds JH (1965) A study of the species of *Colletotrichum* causing ripe fruit rots in Queensland. *Queensland Journal of Agricultural and Animal Science*, **25**, 437-459.

Spencer CC, Neigel JE, Leberg PL (2000) Experimental evaluation of the usefullness of microsatellite DNA for detecting demographic bottlenecks. *Molecular Ecology*, **9**, 1517-1528.

Stamatakis A (2006) RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics*, **22**, 2688-2690.

Stamatakis A, Hoover P, Rougemont J (2008) A Fast Bootstrapping Algorithm for the RAxML Web-Servers. *Systematic Biology*, **57**, 758-771.

Stiles CM, Oudemans PV (1997) Cranberry fruit rot pathogens isolated from various cranberry plant tissues over two growing seasons. *Phytopathology*, **87**, S94.

Stiles CM, Oudemans PV (1999) Distribution of cranberry fruit-rotting fungi in New Jersey and evidence for nonspecific host resistance. *Phytopathology*, **89**, 218-225.

Stockwell CA, Mulvey M, Vinyard GL (1996) Translocations and the preservation of allelic diversity. *Conservation Biology*, **10**, 1133-1141.

Stoddart JA (1983) A genotypic diversity estimate. *Journal of Heredity*, **74**, 489-490.

Stukenbrock EH, McDonald BA (2008) The origins of plant pathogens in agro-ecosystems. *Annual Review of Phytopathology*, **46**, 75-100.

Sukumaran J, Holder MT (2010) DendroPy: A Python library for phylogenetic computing. *Bioinformatics*, **26**, 1569-1571.

Sutton BC (1992) The genus *Glomerella* and its anamorph *Colletotrichum*. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 1-26. CAB International, Surrey, UK.

Szpiech ZA, Jakobsson M, Rosenberg NA (2008) ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, **24**, 2498-2504.

Taylor JW, Geiser DM, Burt A, Koufopanou V (1999a) The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews*, **12**, 126-146.

Taylor JW, Jacobson DJ, Fisher MC (1999b) The evolution of asexual fungi: Reproduction, speciation, and classification. *Annual Review of Phytopathology*, **37**, 197-245.

Taylor JW, Jacobson DJ, Kroken S*, et al.* (2000) Phylogenetic Species Recognition and Species Concepts in Fungi. *Fungal Genetics and Biology*, **31**, 21-32.

Than PP, Shivas RG, Jeewon R*, et al.* (2008) Epitypification and phylogeny of *Colletotrichum acutatum* J.H. Simmonds. *Fungal Diversity*, **28**, 97-108.

Tu JC (1992) *Colletotrichum lindethumianum* on bean: Population dynamics of the pathogen and breeding for resistance. In: *Colletotrichum: Biology, Pathology and Control* (eds. Bailey JA, Jeger MJ), pp. 203-224. CAB International, Surrey, UK.

USDA-CSREES (2001) Crop Profile for Cranberries in Massachusetts.

USDA-NASS (2008) Noncitrus Fruits and Nuts 2007 Summary

USDA, NRCS (2008) The PLANTS Database. National Plant Data Center, Baton Rouge, LA

Vaillancourt L, Du M, Wang J, Rollins J, Hanau R (2000a) Genetic analysis of cross fertility between two self-sterile strains of *Glomerella graminicola*. *Mycologia*, **92**, 430-435.

Vaillancourt L, Wang J, Hanau R (2000b) Genetic regulation of sexual compatability in *Glomerella graminicola*. In: *Colletotrichum: Host Specificity, Pathology, and Host-Pathogen Interaction* (eds. Prusky D, Freeman S, Dickman MB), pp. 29-44. American Phytopathological Society, St. Paul, Minnesota.

Vailllancourt L, Du M, Wang J, Rollins J, Hanau R (2000) Genetic analysis of cross fertility between two self-sterile strains of *Glomerella graminicola*. *Mycologia*, **92**, 430-435.

Viard F, Justy F, Jame P (1997) Population dynamisc inferred from temporal variation at microsatellite loci in the selfing snail *Bulinus truncatus*. *Genetics*, **146**, 973-982.

von Arx JA (1957) Die Arten der Gattung *Colletotrichum* Corda. *Phytopathologische Zeitschrift*, **29**, 414-468.

Vorsa N (1994) Breeding the American Cranberry **5**.

Vorsa N, Polashock JJ (2005) Alteration of Anthocyanin Glycosylation in Cranberry Through Interspecific Hybridization. *Journal of the American Society for Horticultural Science*, **130**, 711-715.

Wheeler HE (1950) Genetics of *Glomerella* VIII: A genetic basis for the occurrence of minus mutants. *American Journal of Botany*, **37**, 304-312.

Wheeler HE, Driver CH, Campa C (1959) Cross- and self-fertilization in *Glomerella*. *American Journal of Botany*, **46**, 361-365.

Wheeler HE, McGahen JW (1952) Genetics of *Glomerella* X: Genes affecting sexual reproduction. *American Journal of Botany*, **39**, 110-119.

Wheeler HE, Olive LS, Ernest CT, Edgerton CW (1948) Genetics of *Glomerella* V: Crozier and ascus development. *American Journal of Botany*, **35**, 722-728.

White TJ, Bruns T, LeeJ S, Taylor. JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: *PCR Protocols: A Guide to Methods and Applications* (eds. Innis MA, Gelfand DH, Sninsky JJ, White TJ), pp. 315-322. Academic Press, Inc., New York.

Whitlock M (2011) G'st and D do not replace Fst. *Molecular Ecology*, **20**, 1083-1091.

Zaffarano PL, McDonald BA, Linde CC (2008) Rapid speciation following recent host shifts in the plant pathogenic fungus *Rhynchosporium*. *Evolution*, **62**, 1418-1436.

Zhu P (1999) *Genetic diversity and systematics of Colletotrichum species on cranberry (Vaccinium macrocarpon)* Ph.D., Rutgers University.

Zhu P, Oudemans PV (2000) A long terminal repeat retrotransposon *Cgret* from the phytopathogenic fungus *Colletotrichum gloeosporioides* on cranberry. *Current Genetics*, **38**, 241-247.