Colletotrichum acutatum sensu lato:
from diversity study to genome analyses

By

Riccardo Baroncelli

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor
of Philosophy in Plant and Environmental Science
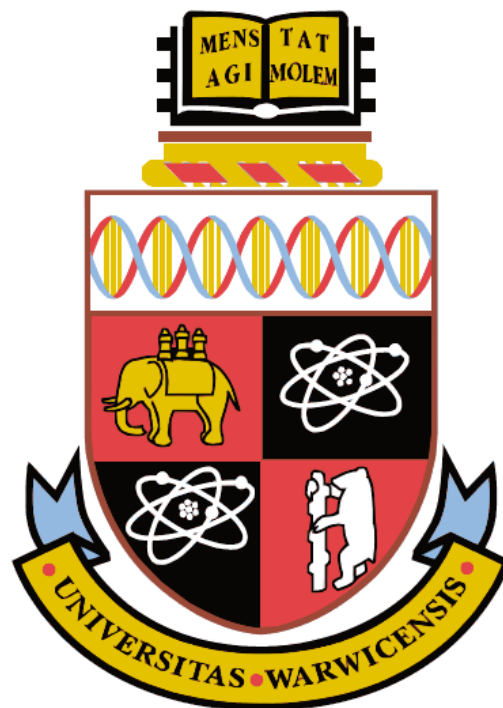
University of Warwick, School of Life Sciences

Date

October 2012

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

°C	Degree Celsius
5.8S	Ribosomal small subunit
AC	Acetate
ACP	Acyl carrier protein
AN°	Accession number (referred to NCBI)
AT	Acyl transferase
ATCC	American type cultures collection
BlastN	Nucleotide Basic local alignment search tool
bp	Base pair(s)
BSC	Biological species concept
CaITSdb	*Colletotrichum acutatum* internal transcripts spacer database
CAZy	Carbohydrate-Active Enzymes
CBM	Carbohydrate-binding motifs
CBS	The Centraalbureau voor Schimmelcultures
CDS	Coding DNA Sequence
CE	Carbohydrate esterase
cm	centimetre(s)
CSEP	Candidate secreted effector protein
CSL	Central Science Laboratory
CTAB	Cetrimonium bromide
cv	Cultivar
CZDA	Czapek–Dox agar medium
d.a.i.	Day(s) after inoculation
DEFRA	Department for Environment, Food and Rural Affairs
DH	Dehydratase
DMAT	Dimethylallyltryptophan synthase
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EPPO	European and Mediterranean plant protection organisation
ER	Enoyl reductase
ESC	Ecological species concept
EST	Expressed sequence tag
EtOH	Ethanol

FAO	Food and Agriculture Organization
FERA	Food and Environment Research Agency
g	Gram(s)
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase gene
GCPSR	Genealogical Concordance Phylogenetic Species Recognition
GH	Glycoside hydrolase
GT	Glycosyltransferase
HMG box	High Mobility Group box
ITS/1/2	Internal transcripts spacer
KR	Ketoreductase
KS	Ketoacyl CoA synthase
l	liter(s)
M	Molar
MAFF	Ministry of Agriculture, Fisheries and Food
MAT/1-1/1-2	Mating type gene(s)
MB	MrBayes
mg	Milligram(s)
MilliQ	Ultrapure water of type 1
min	Minute(s)
MJ	Median-joining
ML	Maximum Likelihood
ml	Milliliter(s)
mm	Millimiter(s)
mM	Millimolar(s)
MP	Maximum-parsimony
MSC	Morphological species concept
N°H	number of haplotypes
N°I	number of isolates
NGS	New Generation Sequencing
NJ	Neighbor-joining
NRPS	Nonribosomal peptide synthase
Ø	Diameter
ORF	Open reading frame
PCR	Polymerase chain reaction

PDA	Potato dextrose agar
PDB	Potato dextrose broth
PKS	Polyketide synthase
PL	Pectin lyase
PO	population assigned
PSC	Phylogenetic species concept
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
rRNA	Ribosomial RNA cluster
SDS	Sodium Dodecyl Sulphate
SDW	Sterile Distilled Water
sec	Second(s)
T	Tempeature
TAE	Tris-acetate-EDTA
TS	Terpenoid synthase
TUB	β-Tubulin gene
UPGMA	Unweighted pair group method with arithmetic mean
USDA	U.S. Department of Agriculture
V	Velocity
v	Volume
VCG	Vegetative compatibility groups
WA	Water agar medium
YpSs	Yeast Extract, Soluble Starch medium
Δ	Variability
μg	Microgram(s)
μl	Microliter(s)

ACKNOWLEDGMENT

I want to thank the following people without whom this study could not have been done.

First of all, I offer my gratitude to all the supervisors involved in this project:

- Professor E. Holub, The University of Warwick (from the 15th of August 2012 to present date)
- Professor S. Sreenivasaprasad, The University of Warwick (from the 1st of November 2008 to the 1st of August 2010* and from the 15th of August 2012 to present date)
- Dr Dez Barbaba, The University of Warwick (from the 20th of September 2010 to the 15th of July 2012**) who sadly is no longer with us
- Jeff Peterson, The Food and Environmental Research Agency part of DEFRA (from uncertain date to present date)
- Charles Lane, The Food and Environmental Research Agency part of DEFRA (from the 1st of November 2008 to uncertain date)

* made redundant by the University of Warwick and now at the University of Bedfordshire.

** sadly passed away the 15th of July 2012

Much gratitude to Michael Thon, Serenella Sukno and Gabriel Rech from the University of Salamanca for their probing minds and development of key ideas for the genome analyses. Special thanks to Mike and Serenella that have personally and constantly supported me many times and without whom my passion for this job could not have been maintained. Many thanks to my group colleague Andrew Armitage for his input helpful discussions. I could not have gotten as far as I have without the support of my family, Hannelore, Elisa, Daniela, Giuse, and some colleagues now firm friends, Ozge and Daniel, Ana, Vincenzo, Gianni and Sacha. Last but not least, thanks to my partner Alessia for the almost unlimited patience and support. This list is far from exhaustive, so I wish to thank those people not mentioned whose contributions have not been forgotten.

DECLARATION

I declare that this thesis and the research that is being contained within is the only work of the author, and that none of this work has been presented for other degrees. If the author collaborated with colleagues, used or adapted methodologies originally established by fellow academics, this is fully acknowledged in the relevant part of the text.

SIGNED   DATE

ABSTRACT

Colletotrichum acutatum sensu lato includes a number of important pathogens that cause economically significant losses of various crops. The *C. acutatum* species complex has a wide host range in both domesticated and wild plant species, and its capability to infect different types of hosts such as insects has also been described. Members belonging to this complex are able to develop three different types of interaction with plant hosts including biotrophic, necrotrophic and hemibiotrophic infections and are also capable of surviving on weeds and non-hosts without causing visible symptoms. They are mainly asexual, but some have a teleomorphic state called *Glomerella* and can be either homothallic or heterothallic. The sexual behaviour in *Glomerella* is more complicated than in most ascomycetes, and strains within the same species do not show a typical MAT1-1/2 system.

The overall aim of this study was to gain an improved understanding of the relationships between the genetic diversity of global populations, host association patterns, geographic distribution and biological and pathological attributes. A database (CaITSdb) containing more than 800 rRNA sequences deposited in GenBank was created along with key biogeographic information, and the data have been analysed in order to investigate genetic diversity and distribution of sub-populations and their evolutionary relationships. The combined evidence was used to assemble a core collection of 120 isolates that are representative of the diversity in host preference, geographic origin, mating behaviour and molecular genetic variation. A multi-locus sequencing approach (based on four neutral loci) has been used to evaluate phylogenetic relationships amongst the isolates in the core collection. A strong relationship was observed between various genetic groups distinguished and their mating behaviour, geographic distribution and host association patterns. Oceania has been identified as a likely geographic origin of this pathogen, as the highest level of variability and groups related to a hypothetical ancestral population are mainly distributed in these countries. All homothallic isolates capable of producing perithecia belonged to the same genetic group A7; whereas all self-sterile heterothallic isolates were classified as either A3 or A5. Isolates derived from the same host tend to cluster together into genetic groups or sub-populations. This evidence is generally in agreement with recent published work on taxonomic re-assessment of *Colletotrichum acutatum sensu lato*, which indicates at least fifteen new species. This study has provided the evidence for the occurrence of three distinct genetic groups on strawberry in the UK corresponding to three species reported in the literature namely, *C. nymphaeae*, *C. fioriniae* and *C. godetiae*. Isolates belonging to the genetic groups that correspond to *C. nymphaeae* and *C. fioriniae* appeared to be the most aggressive on strawberry, followed by *C. godetiae,* and *C. simmondsii* (not found in the UK). Representative isolates of other species were less aggressive.

The first whole genome sequence an isolate (A9 = *C. simmondsii*) from the *C. acutatum sensu* species complex was assembled and analysed using a range of bioinformatics algorithms. An isolate of *C. simmondsii* was chosen based on its wide host range including strawberry and the phylogeographic position. Genome analyses enabled prediction and annotation of the whole gene set at 13549 including 6 % unique to this species. The data also suggested an interesting expansion of several gene families, such as those encoding carbohydrate-active enzymes, secondary metabolites pathways and effectors which could be associated with the wide host range. The new knowledge and resources developed with the genome analyses along with the results of the population level diversity studies provide a platform for future comparative and functional genomics investigations to advance this research.

Resulting Publications and those in Preparation

Damm U, Baroncelli R, Cai L, Kubo Y, O'Connell R, Weir B, Yoshino K and Cannon PF. 2010. Colletotrichum: species, ecology and interactions. *IMA Fungus* 1(2): 161-165

Baroncelli R, Zapparata A, Vannacci G, Sreenivasaprasad S and Holub EB. *Working title*: Genetic characterization and pathogenicity of *Colletotrichum acutatum sensu lato* species associated with strawberry anthracnose in the UK. In preparation.

Oral and Posters Presentations

ECFG11 – European Conference on Fungal Genetics (Marburgh, GERMANY; 30 March – 2 April 2012).

- − Presentation of a poster on "*Colletotrichum acutatum* sensu lato genome project".

26th FGC – 26th Fungal Genetics Conference (Asilomar, CA, USA; 15/20 March 2011)

- − Presentation of a poster on "Molecular phylogenetics in the anthracnose pathogen *Colletotrichum acutatum sensu lato*".

IMC9 – 9th International Mycological Congress: The Biology of Fungi (Edinburgh, UK; 01/06 August 2010).

- − Invited oral presentation on "Molecular phylogenetic and evolutionary relationships in *Colletotrichum acutatum sensu lato*" - Colletotrichum: Species, ecology and interactions meeting
- − Presentation of a poster on "Molecular phylogenetic and evolutionary relationships in *Colletotrichum acutatum sensu lato*".

ECFG10 - European Conference on Fungal Genetics (Amsterdam, THE NETHERLANDS; 29 March – 1 April 2010).

- − Invited oral presentation on "Evolutionary relationships in *Colletotrichum acutatum* populations" - Colletotrichum Workshop
- − Presentation of a poster on "Evolutionary relationships *in Colletotrichum acutatum* populations".

CHAPTER 1

<u>GENERAL INTRODUCTION</u>

1.1. *Colletotrichum*: biology and pathogenicity

Many species belonging to the genus *Colletotrichum* are implicated in plant diseases, generally referred to anthracnose, on a wide range of hosts and these pathogens are characterized by a worldwide distribution. Virtually every crop grown in the world is susceptible to one or more species of *Colletotrichum* (Dean *et al.*, 2012). Common hosts include many dicotyledonous plants such as strawberry, apple, citrus, and stone fruits, and major cereals such as maize and sorghum. Serious diseases on leatherleaf fern and pines have also been reported (Peres *et al.*, 2005) Anthracnose symptoms include dark necrotic lesions, which are oval or angular. Plant parts can be superficially affected at all stages of maturity, from seedlings to mature plants. Various *Colletotrichum* species are also important post-harvest pathogens due to their ability to undergo a non-pathogenic phase.

Colletotrichum species are characterized by a distinctive hemibiotrophic lifestyle (also known to occur in other fungal species, e.g. *Magnaporthe*). Fungi belonging to this genus initially infect through a brief biotrophic phase, associated with large intracellular primary hyphae (some species such as *C. capsici* have been described as subcuticular). The fungus later switches to a necrotrophic phase, associated with narrower secondary hyphae that spread throughout the host tissue. Biomolecular processes that regulate this lifestyle have long been studied by the scientific community, especially those related to the switch from biotrophy to necrotrophy. Recent work has reported genome and transcriptome analyses of *C. higginsianum* infecting *Arabidopsis thaliana* and *Colletotrichum graminicola* infecting maize (O'Connell *et al.*, 2012).

Interestingly, several species of *Colletotrichum* have been reported to cause humans infections including *C. coccodes*, *C. crassipes*, *C. dematium*, *C. gloeosporioides*, and *C. graminicola* (Castro *et al.*, 2001; De Hoogs *et al.*, 2000; Fernandez *et al.*, 2002). These pathogens cause subcutaneous and systemic infections, most commonly occurring in immunosuppressed patients (Guarro *et al.*, 1998). For example, three cases of phaeohyphomycosis caused by *Colletotrichum*, were reported recently involving patients who were undergoing chemotherapy (O'Quinn *et al.*, 2000).

Generally, for experimental studies *Colletotrichum* spp. have the benefit of being haploid organisms that can be cultured axenically, and genetically transformed, which facilitates mutational analysis (Epstein *et al.*, 1998), and the critical assessment of gene function by targeted disruption (Redman and Rodriguez, 1994).

The recently completed *C. graminicola* (from Lisa Vaillancourt's lab at the Department of Plant Pathology, University of Kentucky, USA) and *C. higginsianum* (from Richard O'Connell's lab at the Department of Plant Breeding and Genetics, Max Planck Institute for Plant Breeding Research, Köln, Germany) genome sequences provide useful resources to increase our understanding of this fungus. For these reasons and peculiarities *Colletotrichum* has become one of the most studied plant fungal pathogens (Dean *et al.*, 2012).

Tode first described the genus *Vermicularia* in 1790. Later in 1837 Corda established it as *Colletotrichum* as reported by Sutton (Sutton, 1992). Despite the original description, the genus has been re-described under several different names and so far there are 17 (plus two which are dubiously included) acknowledged generic synonyms for *Colletotrichum* (Sutton, 1980). Most of these genera contained

very few species, but *Gloeosporium* and *Vermicularia* were described with hundreds (Sutton, 1992).

Both names *Colletotrichum* and *Vermicularia* were used during the 19th and early 20th centuries for a range of species now accepted as part of the genus *Colletotrichum*. In 1931 Clements and Shear distinguished *Colletotrichum* from *Vermicularia* based on the presence of marginal setae in *Colletotrichum,* as compared to setae dispersed throughout the conidiomata in *Vermicularia* (Clements and Shear, 1909). However, in 1928 Duke demonstrated that conidiomatal structure and the presence (or absence) of setae, is extremely variable and of no significance at the generic level (Duke, 1928). The revision of the genus *Colletotrichum* by von Arx in 1957 (von Arx, 1957) was a landmark in the classification in which around 750 species were reduced to 11 taxa. In 1980, based on similar criteria, Sutton identified 25 species, but later, in 1992, increased this to 37 on the basis of host specificity (Sutton, 1992). In 2000 the number of species was updated with more morphological, cultural and pathogenicity studies and around 40 were accepted (Cannon *et al.*, 2000). Despite significant developments, the taxonomy of *Colletotrichum* remains in a state of flux (Dean *et al.*, 2012) and few works are focusing on taxonomic re-assessment of this taxa (Shivas and Than, 2009; Damm *et al*, 2012).

1.2. *Colletotrichum acutatum*

Colletotrichum acutatum was identified by Simmonds in 1965 (Simmonds, 1965) and validated in 1968 (Simmonds, 1968). *C. acutatum* is an Ascomycete classified in the class Sordariomycetes, family *Glomerellaceae* (http://www.catalogueoflife.org).

Fungi classified in the genus *Glomerella* Spauld and H. Shrenk and in coelomycetes (anamorphic) in general, and in particular *Colletotrichum* have given one of the hardest challenges to taxonomists. While limits of the genus seem to be well established, the concept of species (such as *C. acutatum*) within this genus is not universally defined and accepted (Sutton, 1992).

C. acutatum and *C. gloeosporioides* are morphologically similar and, due to their overlapping host ranges and the huge variability that they show in culture, it has been very confusing to separate them by traditional taxonomic methods. For example, isolates both taxa can be distinguished by conidium shape and size when grown on potato dextrose agar, but the sizes overlap on pea straw agar demonstrating the low reliability of this characters (Forster and Adaskaveg, 1999). On the other hand, these species are now considered as complexes (or species group) rather than single species, due to the high genetic and morphological divergence and the wide range of hosts shown within intra-specific populations (Sutton, 1992).

In the case of *C. acutatum,* there has been discussion about further subdividing the complex into distinct species (Vinnere *et al.*, 2002). Eight groups (A1 to A8) have been suggested based on ITS and TUB sequences (Talhinhas *et al.*, 2005; Ladner *et al.*, 1999). Several research groups have recently shown that the global populations of *C. acutatum* comprise potentially nine distinct genetic groups, A1–A9 (Sreenivasaprasad and Talhinhas, 2005; Whitelaw *et al.*, 2007). Taxonomic rearrangement of *C. acutatum sensu lato* and new species names has been proposed. For example, A1 has been described as *C. lupini* due to the host specificity of the isolates contained in this group (Nirenberg *et al.*, 2002), A2 has been described as *C. simmondsii* (Shivas and Tan, 2009), A3 has been described as *C. fioriniae* (Shivas and Tan, 2009); and A4 has been described as *C. clavatum* (Faedda *et al.*, 2011)

1.2.1   Reproduction: Mating behavior, VCG and genetics

The life style of *Colletotrichum* species can include sexual (teleomorph *Glomerella*) and asexual stages. In nature, the sexual stages of *Colletotrichum* are rare or absent. Furthermore, sexual behavior in *Glomerella* is more complicated when compared to most ascomycetes; indeed some strains within the same species can be both self-fertile and cross-fertile, while others are strictly cross-fertile (heterothallic) (Chilton, 1949).

In vitro heterothallic mating capability of *C. acutatum* isolates has been reported (Gueber and Correll, 2001). In two cases the teleomorph has been found in nature and has been associated with *Glomerella acutata* (or associated taxa): the first case was on highbush blueberry (*Vaccinium corymbosum*) in Norway (Talgø *et al.*, 2007) and the second on Norway maple (*Acer platanoides*) close to Boston, USA (LoBuglio and Pfister, 2008).

The genetics of "mating system" have been studied principally on *G. cingulata* (Chilton, 1949). Data indicate that *Glomerella* does not have a simple bipolar mating system, or a tetrapolar mating system (Cisar and TeBeest, 1999). Wheeler (1954) also proposed that non-allelic self-fertile mutants of homothallic strains were the basis of heterothallism in this species hypothesizing that most *Glomerella* strains are basically homothallic, but that unbalanced heterothallism may occur as a result of mutations in genes involved in the mating process. The same idea can be applied to *G. acutata* because data obtained by Guerber indicate the two species seem to have the same behavior (Gueber and Correll, 2001). The presence of the mating type ideomorphs in relation to the sexual behavior has been studied *in G. cingulata* and *G. graminicola*, where both partners of a fertile cross carried the

MAT1-2 ideomorph (Vaillancourt 2000a, Rodriguez-Guerra *et al.*, 2005). To date the MAT1-1 ideomorph has never been reported from the genus *Glomerella* despite several attempts to amplify the alpha domain (Vaillancourt 2000a, Rodriguez-Guerra 2005). However, genetic bases of the unbalanced heterothallism have not been clarified yet.

Another system by which genetic diversity may be generated in *C. acutatum* populations (and a general phenomenon in many fungi) is through vegetative compatibility, when anastomosis occurs between genetically different strains. Since reproduction in many *Colletotrichum* species is mostly or exclusively vegetative, the only way of exchanging genetic material between two strains would be anastomosis and heterokaryosis. These processes occur between some *Colletotrichum* isolates but not others and, in some cases, seem to be restricted by the existence of vegetative incompatibility (Brooker *et al.*, 1991). Isolates that cannot form a heterokaryon with other isolates may be genetically isolated and create different populations. Isolates that can undergo anastomosis with others and form heterokaryons belong to the same vegetative-compatibility group (VCG). Those isolates may share gene pool, and are isolated from other strains or VCGs within the taxa (Katan, 2000). VCG analysis has been used widely to study population structures of a few *Colletotrichum* species including *C. gloeosporioides* and *C. acutatum* (Chacko *et al.*, 1994). Data from these studies suggest that genetics of sexual and vegetative compatibility in those two species are similar (Gueber and Correll, 2001).

1.2.2   *C. acutatum*: Pathogenicity and epidemiology

C. acutatum has been well documented from agricultural hosts and natural systems worldwide (Figure 1.1) and more recently, also from other systems such as sea turtles (*Lepidochelys kempi*) and insects (*Fiorinia externa*) (Manire *et al.*, 2002; Marcelino *et al.*, 2008)
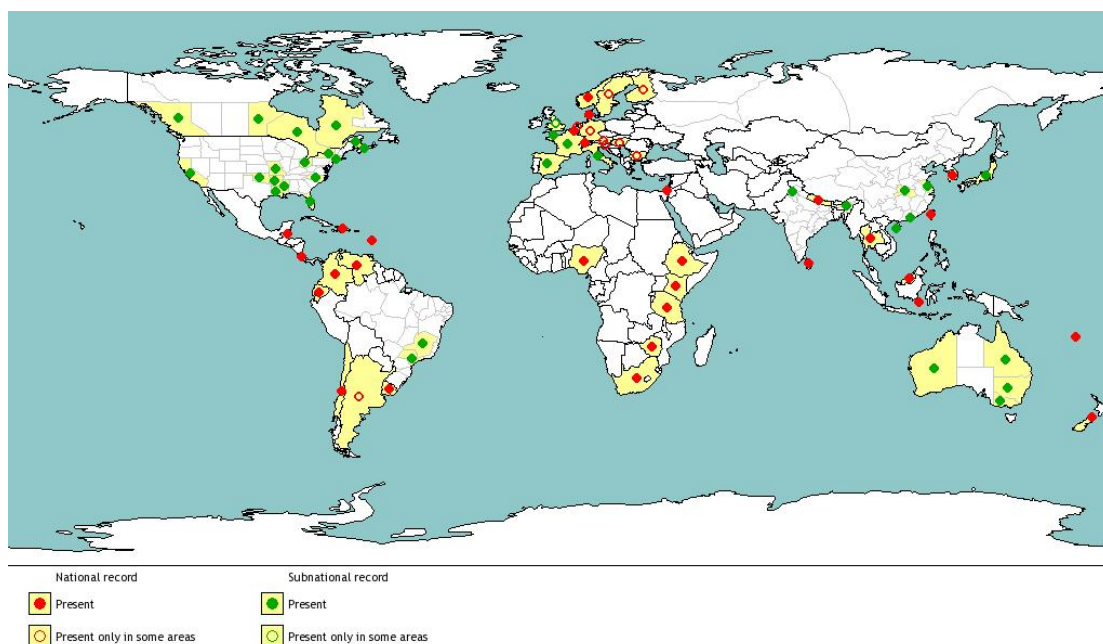
Figure 1.1. The world-wide distribution map of *Colletotrichum acutatum* in 2006. The map is based on data collected by EPPO and related to the presence of the pathogen at a country level. Source: http://www.eppo.org.

The species has a very wide host range (cultivated hosts include: *Anemone coronaria*, *Malus pumila* (apples), *Prunus dulcis* (almond), *Capsicum* spp. (chilli pepper), *Olea europea* (olives), *Carica papaya* and *Pinus spp.*); but it is economically most important on strawberries (*Fragaria* X *ananassa*) and olives (*Olea europaea*). *C. acutatum* can apparently affect almost any flowering plant,

especially in warm temperate or tropical regions, although its host range needs further clarification (Lardner *et al.*, 1999).

C. acutatum symptoms are predominantly necrosis including blights on different host tissues such as leaves, petioles, flowers, fruit, or even roots (quite unusual for this pathogen) on a wide range of hosts. However, this pathogen, like other *Colletotrichum* spp., is not a general necrotroph and it is highly specialized as to the tissue it infects on each host. For example: on orange (*Citrus × sinensis*) flowers are severely affected, but there are no symptoms formed on young leaves; on apples (*Malus domestica*) the fruit is affected, but no symptoms occur on leaves (Peres *et al.*, 2005).

On some fruits, such as apple and blueberry (*Vaccinium* spp.), *C. acutatum* is a pre-harvest pathogen and also a postharvest problem (Jones *et al.*, 1996; Milholland, 1995; Shi *et al.*, 1996). On these plants, quiescent infections that occur prior to harvest begin to develop as the fruit ripens. There are many other examples of postharvest decays on fruits such as peaches (*Prunus persica*), almonds (*Prunus dulcis*), avocado (*Persea americana*), mango (*Mangifera indica*), papaya (*Carica papaya*), and guava (*Psidium guajava*) that progress from quiescent infections occurred previously in the field (Afanador-Kafuri *et al.*, 2003, Agostini *et al.*, 1982, Freeman and Shabi, 1996; Smith, 1998).

Many studies have been conducted through pathology assays with *Colletotrichum*, demonstrating that most isolates are relatively host non-specific (Alahakoon *et al.*, 1994; Freeman *et al.*, 1998; Freeman and Shabi, 1996). However, few studies have focused on *C. acutatum* host specificity. Freeman and Shabi (Freeman and Shabi, 1996) inoculated nectarine, mango, avocado, almond, and apple

fruit with conidial suspensions of isolates of *C. acutatum* isolated from apple and peach. All fruits were susceptible to both isolates but wounding mango fruit was necessary to obtain infection. Peres *et al.*, (in Smith, 1998) found that a *C. acutatum* isolate from strawberry produced lesions on wounded, detached fruit of avocado, guava, papaya, mango, and passion fruit, but not banana.

All these studies demonstrate that *C. acutatum* has a broad host range among fruit crops, and is relatively non-specific in particular conditions (all of them were carried out in laboratory using detached and wounded fruit). Beside that some molecular groups often seem to be host specific (Cannon *et al.*, 2000). In fact isolates from certain hosts do show some specificity. Pathogenicity of *C. acutatum* isolates from lupin, pine, tomato, and persimmon fruit was evaluated on pine and lupin seedlings (Lardner *et al.*, 1999). All isolates from pine were non-pathogenic or weakly pathogenic to lupin and vice versa. The fruit isolates were not pathogenic on either pine or lupin.

Generally speaking we can assume that most isolates can be pathogenic to fruit due to the particular physiology of the host tissue or to cultivated systems, but a certain level of host-specificity can be found in others such as natural ecosystems. However, the level to which cross-infection occurs in nature, and the degree to which disease cycles on one crop depend on the ability of a given group of isolates to survive on a different crop is not known.

1.2.3   Biology and lifestyles

Although *C. acutatum* mostly produces necrotic symptoms, the interaction with hosts appears to have a longer biotrophic phase if compared to other species such as *C.*

higginsianum and *C. gloeosporiodes* (Freeman *et al.*, 2001B). *C. acutatum* survives and competes poorly as a saprophyte in most cases (Eastburn *et al.*, 1999). Nevertheless, *C. acutatum* has been isolated from many plant species and has not always been associated with disease. For example, this fungus may colonize plants as an epiphyte or endophyte on hosts and non-hosts system, without producing symptoms (Freeman *et al.*, 2000). Generally speaking *C. acutatum* is able to develop four different types of interaction with the host (Figure 1.2)
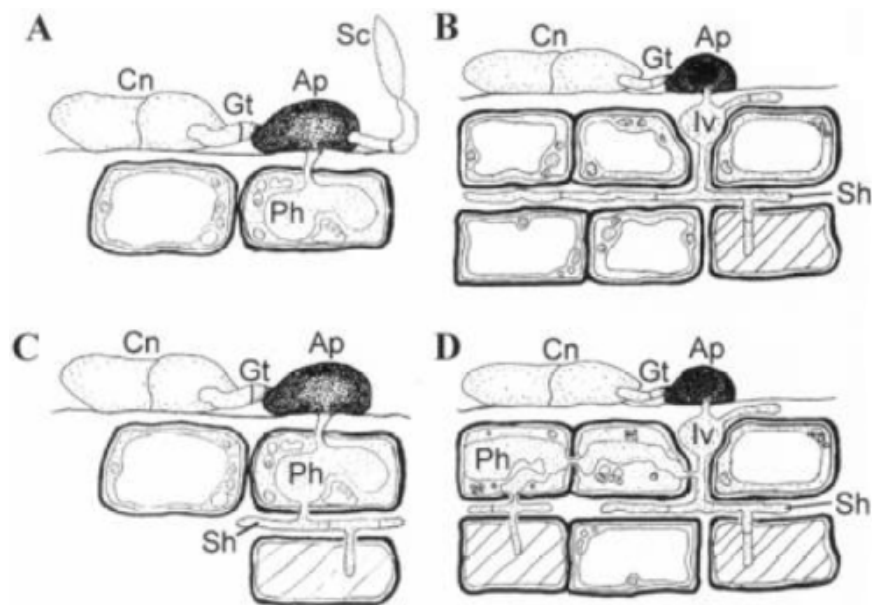
Figure 1.2. Different interactions and infection strategies of *Colletotrichum acutatum*: A - biotrophic infection, B - necrotrophic infection, C - hemibiotrophic infection with infection vesicle and broad primary hyphae within host cell, D - hemibiotrophic and subcuticular, intra- and intercellular development. Legend: Cn = conidium; Gt = germ tube; Ap = appressorium; Iv = infection vesicle; Ph = primary hyphae; Sh = secondary hyphae; Sc = secondary conidium. Source: Peres *et al.*, 2005.

1.3. Motivation for the study

The world population growth has prompted increasing agricultural production, which on the other hand is raising concern about sustainability. Furthermore agriculture has been described as being the most sensitive to climate change and its impact. However, one cannot discuss the impact of climate change on agriculture without taking into consideration the effect of change in crop protection. Crops' weakening is due to progressive selection for production and specific needs of appearance (in particular for soft fruit crops). The massive use of pesticides also poses the problem of pollution of ecosystems. In view of this, there are increasing calls for research on climate impacts involving plant disease, to focus on evolution of specific pathogen and the identification of genes involved in pathogenicity that could have a direct application in crop protection. While we can use knowledge and data to predict climatic changes, data on how this might affect interaction of crop plants with pathogens are still poor. However, it is understandable as the research in population studies and genetic bases of microbe/host interaction had in recent years a significant boost.

"Colletotrichum is one of the most common and important genera of plant-pathogenic fungi. Virtually every crop grown throughout the world is susceptible to one or more species of Colletotrichum" (Dean *et al.*, 2012).

Many species belonging to the genus *Colletotrichum* are implicated in plant diseases, generally referred to anthracnoses on a wide range of hosts and these pathogens are characterized by a worldwide distribution. All this characteristics, its

complexity and economic impact make *Colletotrichum acutatum sensu lato* a suitable system to study evolution and molecular fungal plant interaction as a tool for a better understanding of the processes involved in host/association patterns and speciation process.

1.4. Objectives and approach of the project

The overall goal of this work was to gain an improved understanding of the population structure and host association patterns in *Colletotrichum acutatum sensu lato* pathosystems. To achieve this, the main objectives of the project are:

- To investigate evolutionary relationships across *C. acutatum* species complex.

- Use the strawberry/*C. acutatum* pathosystems in order to evaluate correlations between genetic populations and host specificity.

- The genetic basis of host-interaction and preferences was investigated by comparative genomics

A bioinformatics approach based on data available on GenBank was used to evaluate diversity within the complex and to build a physical collection representative of the geographic distribution, host spectrum and mating behavior. Classical morphological characterization (MSC), multi-locus sequencing, phylogenetics and several bioinformatics tools (PSC) were used to evaluate speciation.

The strawberry anthracnose pathosystem (*C. acutatum*) related to UK was investigated with the following subobjectives: **1.** characterise the population structure of *C. acutatum* associated with the strawberry production systems in the UK using molecular markers; **2.** evaluate cultural behaviour of isolates on artificial media; and **3.** compare the pathogenic ability and aggressiveness of isolates from different molecular-typed groups. *C. acutatum* isolates from other hosts in the UK, as well as representative isolates from strawberry and other hosts from different geographic locations will also been analysed. This will enable comparison of UK *C.*

acutatum populations with wider examples and gain an understanding of how these fit with the global *C. acutatum* populations and their host-association patterns.

The genetic basis of host-interaction and preferences was investigated by inter- and intra-specific comparative analysis of gene content of three different *C. acutatum* strains (plus *C. graminicola* and *C. higginsianum* already available) to identify which genes are unique to *C. acutatum* and which genes are unique to each sequenced isolate will be carried out. The final objective focused on prediction of novel putative genes involved in pathogenicity and host association patterns specificities such as effectors genes. The expansion of gene families was also investigated, focusing on genes involved in secondary metabolites such as toxins. Genomic and genic data of *Colletotrichum* will be a unique and innovative approach for the identification of putative genes in this system. These resources and information could be used to find novel methods for diagnostics, disease control and for a better understanding of pathogens evolution.

CHAPTER 2

MATERIALS AND METHODS

2.1. Fungal strains and growth conditions

Routine cultures were maintained on PDA (Difco Laboratories, Detroit, Mich., USA) at 25°C for up to seven days. Stock cultures were maintained on PDA in sterile water under mineral oil at 18°C.

Two methods were used to grow mycelium when needed:

1.   Single spore–cultures were grown on PDA at 25°C in 9 cm Ø Petri dishes for 10 days.

2.   Small mycelium pieces from an actively growing culture (7-10 day old) on PDA were placed in three 500 ml flasks containing 250 ml PDB (Difco Laboratories, Detroit, Mich., USA). Liquid cultures were grown at 25°C for 5/7 days in shaking flasks.

2.1.1   A globally representative collection

After a preliminary bioinformatics analysis of more than 800 isolates for ITS sequence and host/geographic diversity, a subset of nearly 120 isolates of *C. acutatum sensu lato* has been used in the work. Strains were collected from different research groups and organizations world wide such as: Norwegian Institute for Agricultural and Environmental Research (Norway), Food and Environment Research Agency (as part of DEFRA – UK), Manaaki Whenua Landcare Research (New Zealand), Harvard University Herbaria (USA), American type culture collection (ATCC - USA), United States Department of Agriculture (USDA – USA), The Centraalbureau voor Schimmelcultures (CBS – The Netherlands), CABI Bioscience Centre Herbarium (UK), The Agricultural Research Organization as part of the Ministry of Agriculture and Rural Development (ISRAEL), National Collection of Fungi, Knoxfield Herbarium (AUSTRALIA), University of Florida

(USA) and The World Vegetable Centre (Taiwan). Isolates collected are: genetically representative of populations identified with the preliminary bioinformatics analyses, reference isolates characterized by different mating behavior and representative isolates of well studied model pathosystems have been included.

2.1.2   *Colletotrichum acutatum* from strawberry in UK

A set of 67 *C. acutatum* isolates from strawberry in United Kingdom used in this study were obtained from FERA (Food and Environment Research Agency part of the Department for Environment, Food and Rural Affairs; authorities responsible for Plant Health). 27 *C. acutatum* strain from strawberry worldwide and 9 strain of the pathogen isolated in UK from other hosts have been included in the set of isolates. In addiction a set of 37 isolates representative of the genetic groups identified in previous analyses have been included in this work.

2.1.3   *Colletotrichum* isolates as out-group

Phylogenetic relationships and taxonomy within the genus *Colletotrichum* are currently inaccurate, making the choice of a suitable out-group to root phylogenetic trees challenging. Likewise, molecular investigations of the group conducted to date failed to robustly infer evolutionary relationships between the *Colletotrichum* taxa, with unresolved, minimally supported topologies (Moriwaki *et al.*, 2002) from which no significant conclusions can be drawn. This was the main reason for choosing different species as out-groups.

Four isolates of *C. gloeosporioides* (associated with strawberry), since this specie is

morphologically closely related (and often confused) to *C. acutatum* have been included. In addition, two isolates of *C. spinaciae* (isolated in UK) were also included in the collection as out-group taxa because they are clearly distinct from *C. acutatum sensu lato* both morphologically and on the molecular level. When phylogenetic analysis was based only on rRNA sequences, genetic information regarding a *Colletotrichum* sp. genetically close to *C. acutatum sensu lato*, available on GenBank (AN°: AJ301980) has been used to root the analyses.

The *Colletotrichum graminicola* and *C. higginsianum* full genome sequences also give unlimited source of genetic information available at GenBank (respectively AN°: ACOD01000000 and CACQ02000000). For this reason, the two strains have been included in phylogenetic analyses.

2.2. Morphological and cultural studies

From the collection described above (Chapter 2.1.1) a subset of 49 isolates of *C. acutatum*, based on genetic diversity and host/geographic diversity, and two of *C. gloeosporioides* to form an out-group were used to carry out morphological and cultural studies.

2.2.1   Colony aspects and characteristics

Colony characteristics of the subset of isolates were recorded from cultures grown on PDA at 25°C in 9 cm Ø Petri dishes, under 12 hours photoperiod. The colony characters recorded after 15 days were colour, zonation, transparency aspect, presence of conidial masses or differentiating structures, and colour of the reverse side. For each isolate pictures of the upper side and reverse side were taken and recorded.

2.2.2   Growth rate

The growth rate of the subset of isolates on PDA at four different temperature (10°C, 15°C, 25°C and 30°C) in dark was measured using colonies initiated from 7 mm of diameter mycelia plugs excised from the margins of actively growing PDA cultures 7 days old cultures.

2.2.3   Evaluation of homothallic isolates

Mating assays on the subset of isolates used for morphological characterization for the identification of homothallic isolates have been carried on four different media:

- PDA Difco 39g/l

- WA 15g/l

- Modified CZDA (Guerber and Correll, 2001) (Table 2.1)

- Modified YpSs (Orr *et al.*, 1963) (Table 2.2)

Figure 2.1. Schematic picture showing the plate system used for inucing perithecial reproduction. Orange circle on the left represents the PDA plug used as inoculum and in dark brown distribution of toothpicks on the plate.

Sterilized toothpicks to increase the production of perithecia (Figure 2.1) were placed on the media surface in an "N" configuration. The cultures were incubated in a photoperiod 12 hrs dark and 12 hrs light at 20°C. Presence of perithecia was confirmed with a stereoscope.

Table 2.1. Modified CZDA medium composition used for the evaluation of homothallism of *Colletotrichum acutatum* strains (Guerber and Correll, 2001).

Modified CZDA medium	
$NaNO_3$	2g/l
K_2PO_4	1g/l
$MgSO_4 \cdot 7H_2O$	0,5g/l
KCl	0,5
$FeSO_4$	0.01
Agar	16g/l
Final pH 7.8	

Table 2.2. Modified YpSs medium composition used for the evaluation of homothallism of *Colletotrichum acutatum* strains (Orr *et al.*, 1963).

Modified YpSs medium	
Yeast extract	1g/l
K_2PO_4	1g/l
$MgSO_4 \cdot 7H_2O$	0,5g/l
Casamino acid	4g/l
Soluble starch	20g/l
agar	15g/l
Final pH 7.8	

2.3. DNA extraction

Different DNA extraction methods were used in order to obtain DNA for verious analyses. First two methods (modified Chelex® 100 and GenElute Plant Genomic DNA Miniprep Kit [Sigma-Aldrich]) were used to perform PCR reactions. CTAB

and modified phenol-chloroform DNA extraction were used to perform genome sequencing.

2.3.1   Modified Chelex® 100 (originally Sepp *et al.*, 1994)

1.5 ml microfuge tubes containing 10% Chelex 100 (0.1g Chelex 100 + 900µl water) were prepared and autoclaved. A "mustard seed" amount of mycelium was collected with a sterile toothpick avoiding agar from actively growing culture on PDA. Samples were vortexed for 15 sec and centrifuged at 14000 rpm for 15 sec. Tubes were incubated at 90°C for 20 min. Samples were vortexed for 15 sec and centrifuged at 14000 rpm for 1 minute. Supernatant (approximately 500 µl) was transferred into a new sterile microfuge tube.

2.3.2   GenElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich)

Mycelium grown in liquid culture was harvested by filtration through a double layer of sterile muslin cloth placed in a funnel and rinsed with sterile distilled water to remove traces of the medium. Excess water was drained by slightly squeezing the mycelium with a forceps and transferred into a 30ml sterile tube and the mycelium was placed in a freezer overnight. The frozen mycelium was lyophilised for 48-72 hours and stored at -20°C. Genomic DNA extraction was carried out following the supplier's protocol using 100 mg of the freeze-dried mycelium.

2.3.3   CTAB DNA extraction

This method was used for extracting genome-sequencing quality DNA that has been used for insert libraries. It was used to extract DNA from one isolate (CBS 122122) for the *Colletotrichum acutatum sensu lato* complete genomes project. Cut tips have been used for pipetting during the protocol to avoid DNA fragmentation. Genomic DNA was extracted based on a modified cetyltrimethylammonium-bromide (CTAB)

procedure (Schaafer and Wostmeyer, 1992). The mycelium (250 mg) was ground under liquid nitrogen using pre sterilised chilled mortar and pestle. The resultant powder was mixed with 15 ml of a preheated solution (60°C) containing 10% CTAB, 2M Tris-CI (pH 8.0), 0,5M EDTA, 1.4 M NaCI and 0,5% 2-mercaptoethanol. After incubation for 30 min at 60°C, proteins were removed twice with 15 ml volume of chloroformisoamyl alcohol 24:1 (v/v). The aqueous phase was transferred to a clean tube, and the nucleic acids were precipitated with 0.6 volume of cold 2-propanol. After two-hour incubation at room temperature, the samples were centrifuged for two minutes at 460 g. The pellet was washed twice with 66% (v/v) EtOH and 34% of 0.1M NaCl. Tubes were centrifuged at 1500g for ten minutes, Washing buffer (supernatant) was removed and pellet were air dried in the fume hood (approximately one hour). The pellet was resuspended in one ml of AB, left for few minutes, centrifuged for 5 min and supernatant (DNA) saved and pellet discarded.

2.3.4   Modified phenol-chloroform DNA extraction

This method has been used for extracting DNA of two strains (Pj7, CBS 607.94) for the *Colletotrichum acutatum* complete genomes project. Cut tips have been used for pipetting during the protocol to avoid DNA breaks. For each strain 300 mg of fungal tissue (wet weight) were placed in a 1.5 ml tube and lyophilize the tissue for 24 hrs. Biological material so prepared was treated to lyse the cells using presterilised mortar and pestle. 500 µl of extraction buffer (50 mM Tris + 50 mM EDTA + 2% SDS, pH=8.0) previously prepared was added to each tube and mix for 10 sec. 30 µl of 3 M NaO-AC have been added and gently mix by inversion. Tubes were incubated for 30 min in bath at 68˚C and centrifuged 13000 rpm for 20 min. After

this stage supernatant was transferred to a new tube, added 520 µL of phenol-chloroform and mixed two times for 8 sec with a one min interval. Tubes were centrifugeded at 13000 rpm for 15 min. Supernatant was transfered to a new tube and volume of each sample recorded. To each tube a volume of 1:1 chloroform:isoamylalcohol solution was added. Tubes were centrifuge for 15 min at 13000 rpm, supernatant transferred to a new tube and record volume. An equal volume of isopropanol to the supernatant was added and mixed by inverting. Tubes were centrifuged for 15 min at 13000 rpm; isopropanol discarded and left opened in the Fume hood to air dry the pellet.

2.4. DNA quantity and quality measurement

In order to estimate the concentrations of DNA extracts a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, DE, USA) available at the University of Warwick, School of Life Sciences laboratories – Wellesbourne campus was used to measure DNA concentrations. When DNA concentration was critical for further analysis also a Qubit® 2.0 Fluorometer has been used. Each batch of samples quantification has been preceded by two standard calibration of the machine as required by protocol provided by the supplier.

2.5. DNA purification

Method 1

Pellet obtained from DNA extraction was resuspended in 50µL of RNA free H_2O, 0.5 µl of 10 mg/ml RNase per 50 µl of DNA were added and the tubes placed

at 37˚C for 30 min. 1:10 volume of 3M NaO-AC (pH 5.2) and two volumes of 95%

EtOH was added to each sample, mixed gently and incubated at -20˚C for 20

min. Tubes were centrifuged for 10 min at 13000 rpm and followed by two washed

with 500 µl of 70% ethanol. Tubes were finally centrifuged for 5 min 13000 rpm for

each wash. Tubes were left opened in the Fume hood to air dry the pellet and when

needed DNA was resuspended.

Method 2

To remove the RNA, RNase plus II gel extraction/desalting Kit was used. 10

µl of RNase was added to the appropriate volume of DNA (250 µl) and left at room

temperature for 30 min. two volumes of water (500 µl) and three of QX1 (750 µl)

buffer were added to the sample turning the solution from hyaline to yellow. QIAEX

II desalting protocol provided by the supplier was followed using 10 µl of QIAEX

gel per 5 µg of DNA. Elution was carried at 50°C for 10 min.

2.6. PCR amplification

In order to amplify genomic loci chosen for molecular characterization PCR reaction

mixes (20 µl), contained one µl of DNA, one µl each primers (20 µM), 7 µl of H_2O

and 10 µl of ReadyMix RedTaq (Sigma). Loci analyzed (Figure 2.2) were:

- part of the rRNA region including ITS1, 5.8S and ITS2 (ITS)

- exons 3 through 6 (introns 2 through 4) , partial sequence of the beta-tubulin
  2 gene (TUB)

- partial sequence of the glyceraldehyde-3-phosphate dehydrogenase
  (GAPDH) gene

- partial sequence of the MAT1-2 gene (MAT), specifically the conserved HMGbox region.
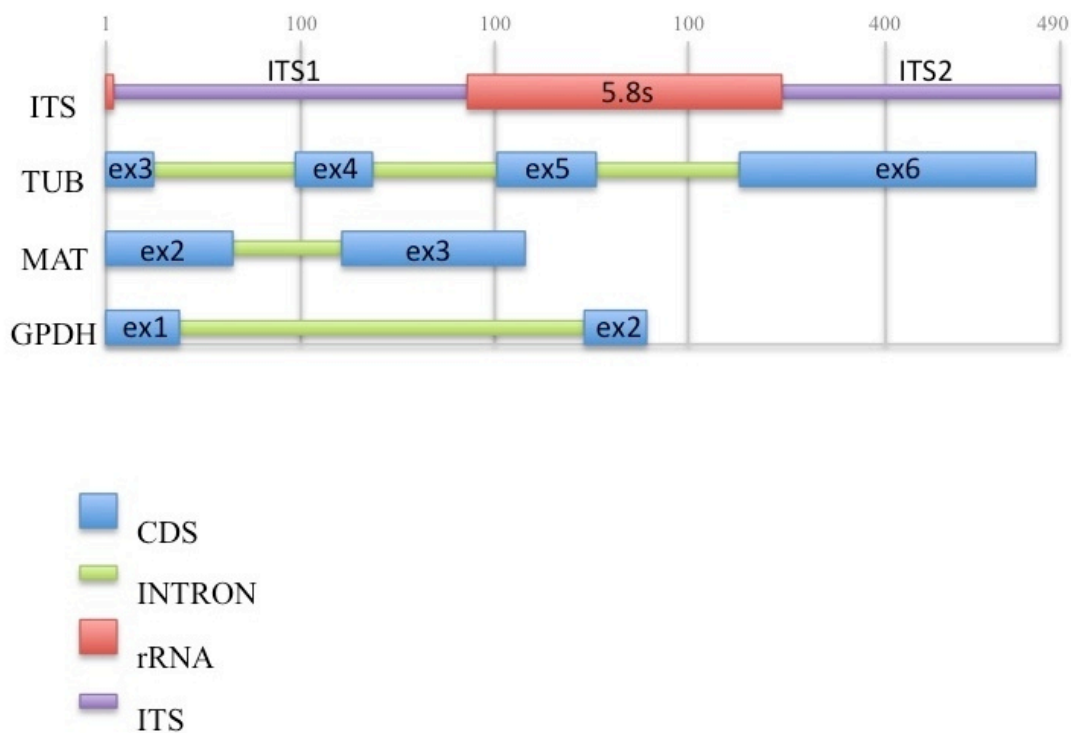
Figure 2.2. Scheme representing the four nucleotide sequences used in this study for phylogenetic relationships and genetic characterization of *Colletotrichum* strains. Coding regions have been represented in thick blue bars, slim green bars represent introns, thick red bars indicate rRNA subunit and slim purple bars indicate internal transcribed spacers (ITS1 and ITS2).

2.6.1 rRNA

Primers ITS1Ext (5'- GTAACAAGGTTTCCGTAGGTG -3') and ITS4Ext (5'- TTCTTTTCCTCCGCTTATTGATATGC -3') (Talhinhas *et al.*, 2002) were used for the amplification of rRNA ITS regions. Thermocycling parameters were as follows: one cycle of two min at 95°C, 30 cycles of one min at 94°C, one min at 55°C, and one min at 72°C, ending with one cycle of 5 min at 72°C.

2.6.2 <u>TUB</u>

Primers TB5 (5'-GGTAACCAGATTGGTGCTGCCTT-3') and TB6 (5'-GCAGTCGCAGCCCTCAGCCT-3') were used for the amplification and sequencing of the variable region 1, spanning the exons 3, 4, 5, and part of 6 (Talhinhas *et al.*, 2002). Thermocycling parameters were as follows: one cycle of two min at 95°C, 30 cycles of one min at 94°C, one min at 65°C, and one min at 72°C, ending with one cycle of 5 min at 72°*C*.

2.6.3 <u>MAT1-2</u>

A set of primers used and published by Du *et al.* (2005) were tested on a representative subset of isolates, we were able to get good amplification only for isolates belonging to some genetics groups, showing that this primer set of wasn't species-specific. Therefore, using published sequences in combination with the primers already published for the HMG-box of MAT1-2 gene from *C. gloeosporioides*, *C. acutatum* and *C. cereal* (Crouch *et al.*, 2009; Du *et al.*, 2005), two non-degenerate primers were designed in order to amplify and sequence variable regions of the target gene in *C. acutatum*. Primers used were HMGacuF2 (5'-CTCTACCGCAGTGACTACCAAGC-3') and HMGacuR (5'-TCTTGTTGTGGCGCTCCTTG-3'). Published primers HMGgloeF1 (5'-CCTAATGCGTACATTCTCTACC-3') and HMGgloeR1 (5'-TGGGATACATCATCAAGAGGC-3') were used to amplify the selected region in *C. gloeosporioides* (Du *et al.*, 2005).

Thermocycling parameters were as follows: denaturation at 95°C for 5 min was followed by 40 cycles of 95°C for one min; between 55°C and 48°C for 60 s; 72°C for 30 s. Final extension of 20 min at 72°C.

2.6.4   GAPDH

Primers GDF1 (5'-GCCGTCAACGACCCCTTCATTGA-3') and GDR1 (5'-GGGTGGAGTCGTACTTGAGCATGT-3') were used to amplify a 200-bp intron region of the GAPDH gene (Guerber *et al.*, 2003). Thermocycling parameters were as follows: one cycle of two min at 95°C, 35 cycles of denaturation at 94°C, annealing at 60°C for one min and extension at 72°C for 30 sec with final extension at 72°C for three min

For each PCR reaction a negative control using water instead of DNA was performed in order to evaluate any contaminations of the reagents.

2.7. Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine the presence or absence of PCR products and quantify the size (length of the DNA) of the product. 5 µl of each PCR reaction products were routinely analysed by electrophoresis on 1% agarose gels in 1× TAE buffer (Gibco) following standard procedures (Sambrook *et al.*, 1989). In each agarose gel electrophoresis HyperLadder™ I ready-to-use molecular weight marker was run to determine PCR product quantification and size.

2.8. PCR product purification

Depending on the purity of the PCR products, either agarose gel extraction (in order to set up new PCR and primers to amplify the variable region of MAT1-2 gene) or direct column-based PCR purification was applied to clean up the products. The QIAquick PCR kit based on silica-membrane-based purification of PCR products or

QIAquick gel extraction kit (Qiagen, West Sussex, UK) was used for this purpose.

2.9. Sequencing and alignment

Sequencing of PCR products was done at the University of Warwick, Genomics Center using ABI Prism 7900HT or ABI3100 sequence detection system (Applied Biosystems, UK). PCR products were cleaned up and then quantified by ladder. One to five microliter of each sample (depending on DNA concentration) was using in sequencing reaction with the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, UK). One microliter of each primer used for PCR reaction was used to sequence PCR products. Cycle sequencing reaction comprised one min at 96°C followed by 25 cycles of 10 sec at 96°C, 5 sec at 50°C and two min at 60°C. ABI trace files were analyzed and consensus sequences were generated using Geneious 5.6.5. All the sequence were aligned using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/) and MAFFT (http://mafft.cbrc.jp/alignment/software/) and when needed manually adjusted

2.10. Bioinformatics and phylogenetic analysis

Multiple sequence alignments were exported to MEGA5 (Tamura et al., 2011) where best-fit substitution models were calculated for each separate every sequence dataset.

When necessary, in order to evaluate whether the four loci sequenced were congruent and suitable for concatenation, tree topologies of 75% Neighbour-Joining bootstrap and maximum parsimony analysis (100,000 replicates) were performed on gene separately and compared visually (Mason-Gamer & Kellogg 1996). The multilocus concatenation alignment (ITS, TUB2, MAT1-2 and GAPDH) was

performed with Geneious 5.6.5 created by Biomatters available from http://www.geneious.com/.

A Markov Chain Monte Carlo (MCMC) algorithm was used to generate phylogenetic trees with Bayesian probabilities using MrBayes 3.2.1 (Ronquist and Huelsenbeck, 2003) for single set of data as well as combined sequence datasets. Models of nucleotide substitution for each gene determined by MEGA5 were included for each locus. The analyses of two MCMC chains were run from random trees for numbers of generations necessary to reach 0,01 and sampled every 100 generations. Details and further information will be provided in the specific chapters.

CHAPTER 3

COLLETOTRICHUM ACUTATUM SENSU LATO ITS SEQUENCES DATABASE:

CaITSdb

3.1. INTRODUCTION

3.1.1. Species concept, definition and criteria

Historically scientific debate of species concept has been focused on two main objectives: first, definition of "species"; and in second place, which are the best methods of species identification and delimitation.

Endler (1989) has delineated four different bipolar perspectives, resuming the kinds of species concept debate.

o *Taxonomic vs. evolutionary*: the taxonomic concept is based on the need to define species for practical reasons and without considering (or investigating) the evolutionary relationships of the target organisms of the study; in the evolutionary concept, on the other end, species designation is used as a tool to make hypothesis or to study evolutionary process and relationships.

o *Theoretical vs. operational*: in the first case the concept of "species" is strictly related to the meaning of the word itself and to the process that explores the origin of the species; theoretical concepts are more useful in terms of intellectual debate rather than in practical terms. Operational concepts usually include taxonomic studies, but they might also be focused on evolutionary processes. In this case, the objective of the concept resides in its applicability and practicality.

o *Contemporaneous vs. clade*: contemporaneous concept is related to those biologists that mainly focus their study on existing organisms, and rarely consider the concept of "species" as an on-going process. The second

definition refers to those studies more likely to view species within the context of clades and ancestor-descendant relationships.

o *Reproductive vs. Cohesive*: mainly separated by the consideration of genetic exchanges or the possibility that these events occur. The author classifies as "reproductive", those concepts that focus on the genetic exchange processes (such as reproduction) that maintain segregation between species. This concept contrasts with those focused on species as units with phenotypic and genetic cohesion.

The species concept is crucial to biologist and has received extensive debate, yet a general definition of "a species" has not found final agreement. Plant and animal populations are groups of diploid organisms with defined growth that reproduce through sexuality; fungi, instead, are generally haploid, with unlimited growth and capable of broad range of reproductive strategies that may complicate the discussion. Fungi may reproduce sexually (genetic recombination), asexually (clonal propagation) and also through a distinctive process of these organisms: the parasexual cycle (genetic recombination takes place within three steps: diploidization, mitosis and haploidization). All these reasons, and the higher complexity of this kingdom, explain why much of the discussion has been focused on animals and plants, while fungi have generally been marginal to the debate.

However, defining species across the fungal kingdom is essential and crucial from two points of view:

o Applied: plant and human pathologists must be able to diagnose, name, and communicate about the organisms they face.

- o Theoretical: fungi are organisms that due to their complexity could be considered as model systems useful to study evolution in a broad sense of the term (ecological adaptation, evolution of reproductive behavior, origin of organisms interaction, *etc.*)

Correct species definitions in fungal taxonomy are critical in the establishment of quarantine regulations, identification of plants resistant to pathogens, preservation of biodiversity, description of organisms for biochemical production and, in a less scientific view, for patent applications. Another crucial point is recognizing that organisms are dynamic entities; a useful species concept should define an organism at a single period in time and, generally speaking help as a resource of communication among biologists.

Species criteria, generally speaking are related to recognition and delimitation of species (Taylor *et al.*, 2000). Species criteria have more recently been subdivided into:

- o The Biological Species Concept (BSC) for instance is mainly based on genetic isolation

- o The Morphological Species Concept (MSC) emphasizes morphological divergence

- o The Ecological Species Concept (ESC) emphasizes adaptation to a particular environment

- o The Phylogenetic Species Concept (PSC) emphasizes genetic divergence

These criteria match with different events that occur during population divergence and speciation process. Historically, the criterion most frequently used for fungi has

been the MS*C*. However, many new species have been determined using the Genealogical Concordance Phylogenetic Species Recognition (GCPSR), which is a bioinformatics approach of the PSC (Taylor *et al.*, 2000). The GCPSR has been extremely useful in many cases with fungi, because it is more closely discriminating than other criteria and due to less variability of the practical approach is currently the most widely used within the fungal kingdom (Giraud *et al.*, 2008).

3.1.2. Speciation process

Speciation is the evolutionary event in which one species splits into two or more. Biologically it's also the process by which biodiversity is generated. There are four geographic types of speciation occurring in nature (Figure 3.1), based on the extent to which speciating populations are geographically isolated from one another:

o allopatric: a population splits into two isolated populations due to formation of a geographic barrier; the two sub-populations then undergo genotypic and phenotypic divergence as they become subjected to different selective pressures or they independently undergo genetic drift.

o peripatric: in this case new species are formed in isolated peripheral populations; this kind of speciation is like allopatric but different because the two subpopulations are unbalanced with one population isolated and not representative of the original (founder effect)

o parapatric: similar to allopatric but in this case there is only a partial separation of the areas of two diverging populations

o sympatric: process in which populations diverge while inhabiting the same geographic location
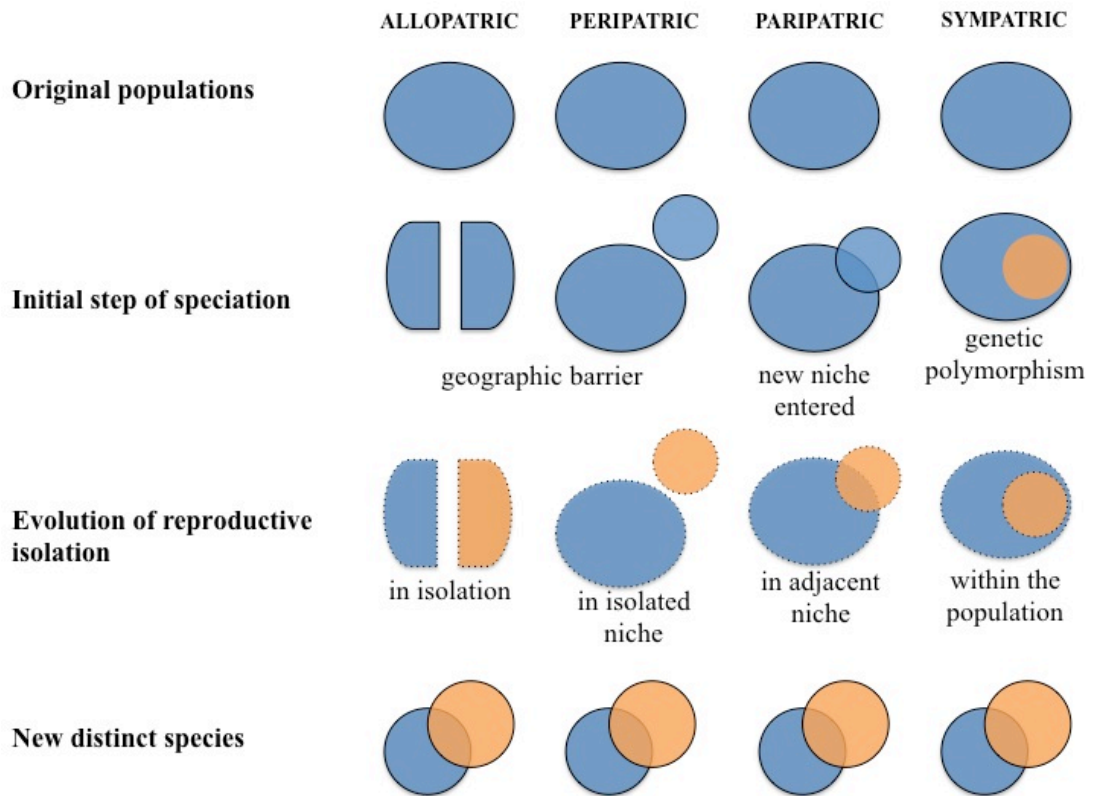
Figure 3.1. Graphical representation of four different speciation process that can occur in nature

Important studies have been carried out on speciation process in fungi, and they have described valid biological models. Across scientific community, generally speaking, it has long been assumed that species originate mostly through allopatric divergence (Mayr, 1963). However, fungi might be an exception as they have been considered ubiquitous and not characterized by specific geographic areas (Finlay, 2002). The reason for this assumption is related to airborne fungi (and their capability to

disperse spores over long distances) and plant pathogens affecting cultivated hosts that have been distributed worldwide.

Nevertheless, recent studies have shown allopatric divergence using GCPSR in different fungal systems (Taylor *et al.*, 2006). Another reason why scientists mainly refer to allopatric speciation is due to the difficulty of defining sympatric speciation in parasites. For instance, plant pathogenic fungi that evolved specialization to different hosts have been considered allopatric (Giraud *et al.*, 2008). Kondrashov in 1986 gave a definition of sympatry specifying that the probability of mating between two individuals should depend only on their genotypes and not on physical barriers (Kondrashov, 1986).

3.2. <u>OBJECTIVES AND APPROACH</u>

The target of this research is to study the evolution of *C. acutatum* species complex, using an integrated bioinformatics approach to investigate related aspects such as origin, spread, host association patterns and speciation of this pathogen. Based on ITS sequence (the most used as universal barcode in fungi) this chapter will focus on the identification of genetic populations within *C. acutatum sensu lato*.

These analyses will provide initial information on the evolutionary relationships of this taxa and the results will inform the assembly of a collection of strains representative of the global diversity.

To achieve this, the focus is on:

- Collection of data available in GenBank related to isolates of *C. acutatum sensu lato*.

- Analysis of data available and published for a better understanding of global population structure, geographic distribution, genetic diversity and host association patterns.

- Analysis of the genetic variability of *C. acutatum* global population.

3.3. <u>RESULTS</u>

3.3.1. <u>*C. acutatum* rRNA sequences database</u>

A database of comparable sequences of *C. acutatum* and related species has been built using an integrated bioinformatics approach based on sequence similarity.

To build the CaITSdb (***C****olletotrichum* ***a****cutatum* **ITS d**ata**b**ase), reference sequences representative of the genetic groups identified before have been used as a starting point (Table 3.1).

Table 3.1. rRNA reference sequences of representative strains belonging to the different genetic groups within *C. acutatum sensu lato* identified in previous work (Sreenivasaprasad and Talhinhas, 2005; Whitelaw *et al.*, 2007). These sequences have been used to build the CaITSdb.

Strain	Country	Date	Host	Group	AN°
PT30	Portugal	1999	*Lupinus albus*	A1	AJ300561
CA 397	USA		*Fragaria x ananassa*	A2	AF411765
CR46	Portugal	2000	*Vitis vinifera*	A3	AJ300563
NI90	UK		*Fragaria x ananassa*	A4	AF411766
PT227	Portugal	2003	*Olea europaea*	A5	AJ749694
PT250	Portugal	2003	*Olea europaea*	A6	AJ749700
MP3	USA	2006	*Acer Platanoides*	A7	EU622052
Tom-21	Colombia	1998	*Cyphomandra betacea*	A8	AF521196
DAR76919	Australia	1998	*Vitis vinifera*	A9	DQ991733

These nucleotide sequences have been used to run Netblast (BlastN based webtool), searches for sequences similar in non-redundant nucleotides database at the National Centre for Biotechnology Information (NCBI). For each query sequence the software found the 1000 most similar nucleotide sequences available in GenBank. The cut-off of results has been evaluated manually for each query. Batch Entrez

(www.ncbi.nlm.nih.gov/sites/batchentrez) has been used to retrieve the data in a batch mode to avoid duplication of this large number of sequences. Information such as: taxonomic classification, strain codes, ITS accession number, length of the sequence, host, location, year and reference of isolation have been collected for each isolate. The database has been created to help develop a wider view of *C. acutatum* populations, a better understanding of the taxonomy of this taxa and to link ITS haplotypes with reproductive behavior, host association and geographic distribution patterns. The CaITSdb has also been used to build a physical collection of strains representative of the genetic populations identified.

3.3.2. Global population structure

Using a bioinformatics approach, more than 800 comparable rRNA sequences of *C. acutatum* and related species has been built (Appendix p.193). All the sequences were then aligned using MAFFT and MUSCLE and manually checked in order to increase the accuracy of the alignment. For population analyses of *C. acutatum*, ITS nucleotide sequences data obtained for the CaITSdb was reduced using ALTER (http://sing.ei.uvigo.es/ALTER/) to collapse sequences to single haplotypes. This resulted in 152 unique haplotypes and 520 (including gaps) comparable characters for *C. acutatum sensu lato*. To visualize intraspecific evolutionary relationships between individuals and genetic populations (Figure 3.2 and 3.3), the Median-joining network algorithm (Bandelt et al., 1999) was used to build the evolutionary lineage for the set of haplotypes using the software Network 4.6 (www.fluxus-engineering.com/ netwinfo.htm). In this work the term "population" will be used as

synonimus of "genetic group", and both related the genetic clusterization of the analysis carried out.

The Network representing evolutionary relationship (Figure 3.3) showed different genetic populations (or groups); some of them can be related to the genetic groups identified previously (Sreenivasaprasad and Talhinhas 2005). However, using a wider set of data, clustering of the strains were appears less compactly related showing a higher complexity and variability compared to previous work based on restricted set of isolates related to specific hosts or geographic areas (Sreenivasaprasad and Talhinhas, 2005; Whitelaw *et al.*, 2007). From the genetic network represented in the picture it also shows how isolates belonging to A5 group and related populations (H7, H44 and H46) seem to be closely related to a central theoretical genetic bridge; in fact, these groups link the other *C. acutatum* strains with the out-groups. From this observation, the first hypothesis is that isolates belonging to this genetic group could be the closest to a theoretical ancestral population. What can be seen from this data is the complexity in the taxonomy and the limit of the information stored in GenBank. Overall, 16 sequences have been deposited and annotated as unknown or a different genus (e.g. *Fusarium phormii* designation has been used for *Colletotrichum* strains isolated from *Phormium* spp.). In the case of species designation it is even more complicated, usually due to the lack of specific morphological characters useful for the identification, the tendency to associate pathogens to their host and the tendency to subdivide the *C. acutatum* species complex into new host-related species (for example: *C. carthami* as pathogen of different species of *Carthamus*, *G. miyabeana* as pathogen of *Salix*, *C. phormii* as pathogen of *Phormium, G. fioriniae* as pathogen of the insect *Fiorinia externa*, *C. lupini* as pathogen of *Lupinus, C. simmondsii, etc.*).
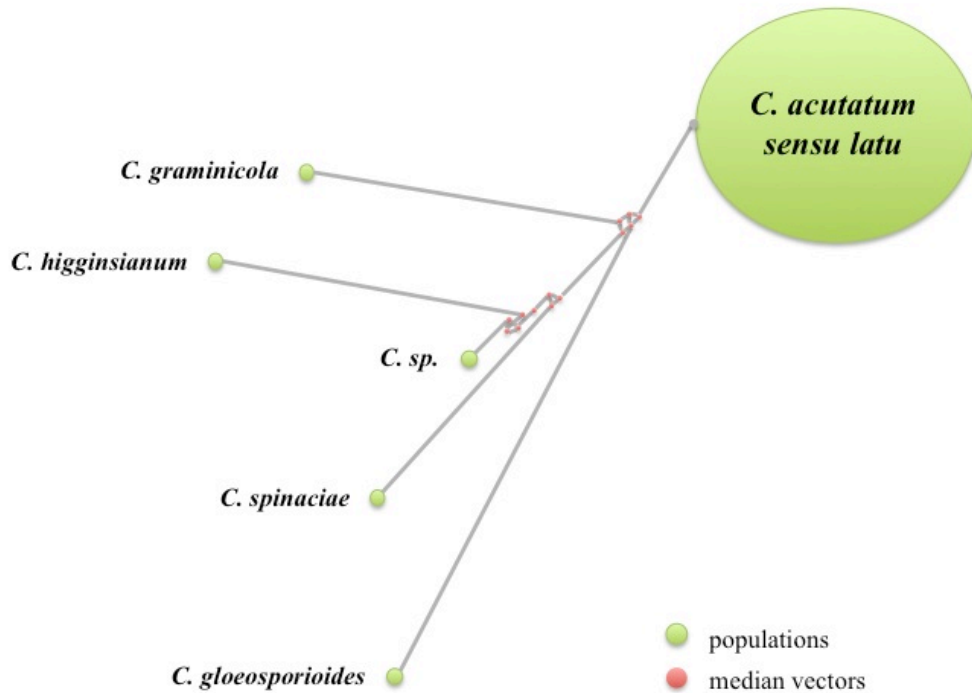
Figure 3.2. Median-joining network of 152 rRNA sequences haplotypes of *Colletotrichum acutatum* and: one *C. higginsianum* (AB105955), one *C. spianciae* (GU227848), one *Colletotrichum graminicola* (DQ126256), one *Colletotrichum gloeosporioides* (FJ755268) and one *Colletotrichum sp.* (AJ301980) genetically close to *C. acutatum* species complex as out-groups. Circles areas are proportional to the genetic variability of the populations and length of lines is proportional to the genetic distances between populations.
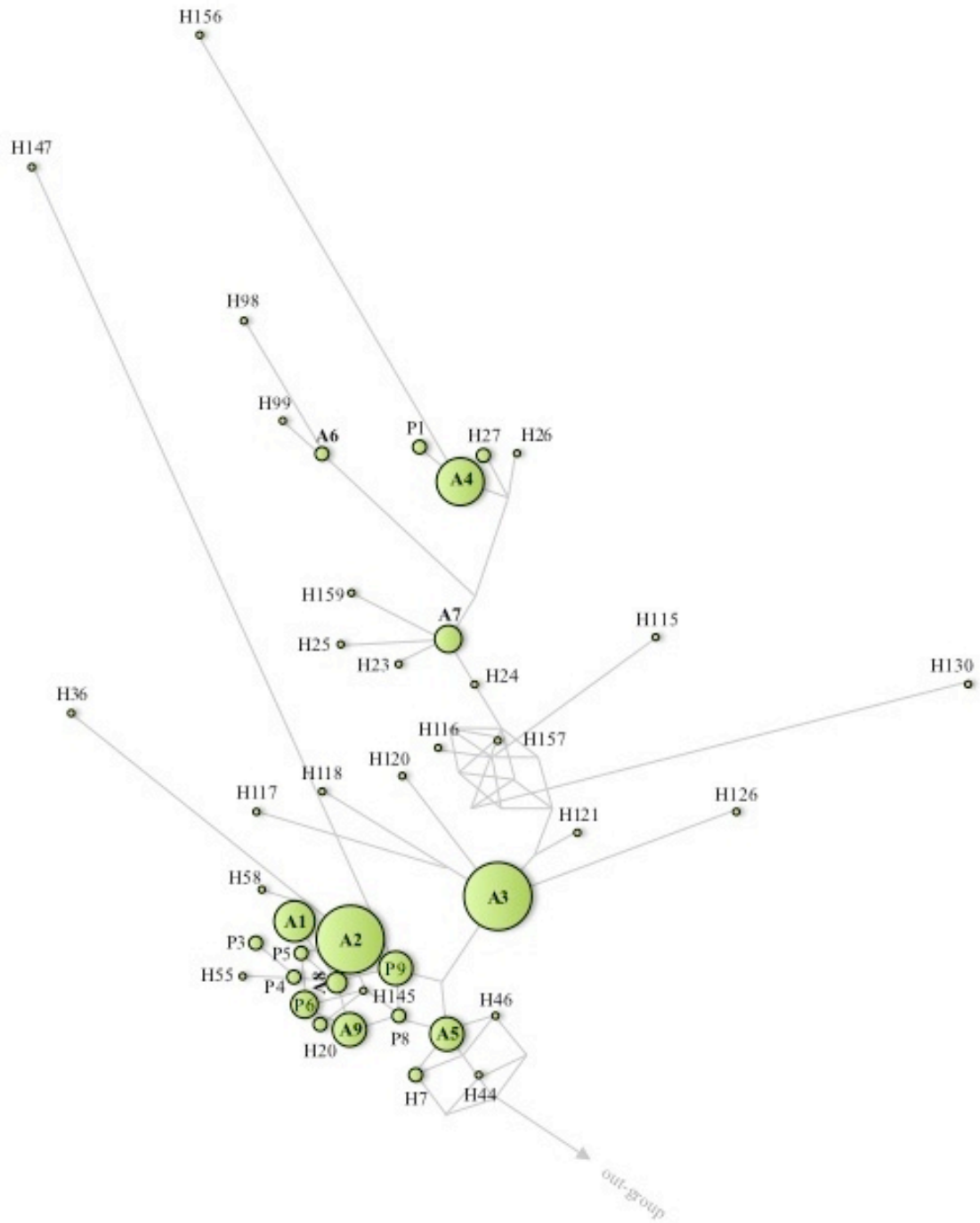
Figure 3.3. Median-joining network showing evolutionary relationships of 152 rRNA haplotypes sequences of *Colletotrichum acutatum* species complex. Populations corresponding to genetic groups previously identified have been labelled (A1, A2, A3, A4, A5, A6, A8 and A9). Further populations have been labelled as Pn (P1, P2, P3, P4, P5, P8 and P9) and single haplotypes as Hn. Circles areas are proportional to the genetic variability of the populations and length of lines is proportional to the genetic distances between populations.

3.3.3. Host association patterns

In relation to host association patterns, the data collected showed that *C. acutatum* strains have been isolated from more than 90 genera of plants (either in crops and in natural ecosystems), two insect species (*Fiorinia externa* and *Orthezia praelonga*) and in a couple of cases also with marine ecosystems (Table 3.2). Most of the hosts do not show any strong relationship with genetic groups/populations of *C. acutatum*, this is true especially in cultivated fruit systems such as strawberry, olive, *etc.* Nevertheless some hosts show a specific correlation with the genetic groups/populations and in few cases also with the ITS haplotypes suggesting a progressive trend (at different evolutionary levels) of specialization of some genetically related isolates to a particular host. For example, almost all the strains belonging to group A1 have been isolated from a single host, in fact more than 90% of isolates belonging to this group have been isolated from *Lupinus* infected tissues. However, strains belonging to different genetic groups, genetically very close to A1 (A2, A3, A8 and P9), have also been isolated from infected tissues of this host; the incidence and the evolutionary distances between those populations suggest an increasing host-specificity over time of those strains. Similarly isolates belonging to A6 and groups genetically close (H98 and H99) have been isolated from *Rhododendron*. In other cases the relationship is at a deeper level; for example all the isolates associated with *Hakea* have the same ITS sequence (haplotypes H7) and all the strains with that sequence have been isolated from *Hakea spp.*. Furthermore these isolates form a specific genetic group. Another situation showing a specific pattern between host and population is represented by strains isolated from *Citrus* hosts; all strains associated to this host (15 from 5 different countries in Asia and America) are

represented by two populations (P8 and A8) genetically close to each other. There is also a relationship in different systems such as insects. Entomopathogenic strains isolated from *Orthezia* belong to genetic group A2 (originally described as *Colletotrichum gloeosporioides*), the one from *Fiorinia* to A3; in this case, strains are characterized by haplotypes that are also associated with plant hosts. In these cases, it seems that the capability of infecting insect species by *C. acutatum* has not led to the differentiation of specific sub-populations. The hypotheses that follow these observations could be two: the capability of infecting insect has been acquired relatively recently by two distinct populations or the process has been acquired by an ancestral organism, has been lost by some populations and has not led to any evolutionary specialization. In fact these strains have also maintained the capability of interacting with plants. The use of higher resolution loci suitable for genetic characterization and pathogenicity assays to explore intra-specific differences in aggressiveness should be used on a wide range of isolates and hosts for a better understanding of relationships that lead the evolutionary process in this organism.

Table 3.2. Host range, percentage of occurrence, number of isolates and genetic variability of different sub-population identified in the CaITSdb.

PO	N° I	N° H	Δ	Hosts and incidence
A1	76	4	0.40%	92% *Lupinus*; 1% *Cinnamomium*; 1% *Fragaria*; 1% *Coffea*; 1% *Urtica*
A2	189	37	0.90%	42% *Fragaria*; 11% *Olea*; 5% *Malus*; 3% *Anemone*; 3% *Carica*; 3% *Persea*; 3% *Prunus*; 3% *Lupinus*; 3% *Protea*; 2% *Citrus*; 2% *Vitis*; 2% *Mangifera*; 2% *Nymphaea*; 1% *Litchi*; 1% *Solanum*; 1% *Orthezia**; 1% *Apium*; 1% *Calluna*; 1% *Capsicum*; 1% *Cirsium*; 1% *Eriobotrya*; 1% *Matthiola*; 1% *Murraya*; 1% *Rubus*; 1% *Vaccinium*; 1% *Vigna*; 1% *Ziziphus*; 1% *Phaseolus*; 1% *Photinia*; 1% *Actinidia*
A3	179	25	0.90%	26% *Vaccinium*; 10% *Fragaria*; 5% *Malus*; 5% *Vitis*; 4% *Olea*; 4% *Fiorinia**; 4% *Prunus*; 3% *Pyrus*; 2% *Solanum*; 1% *Anemone*; 1% *Capsicum*; 1% *Huperzia*; 1% *Ixodes*; 1% *Magnolia*; 1% *Marine ecosystem***; 1% *Tulipa*; 1% *Acacia*; 1% *Acer*; 1% *Ailanthus*; 1% *Apium*; 1% *Calamagrostis*; 1% *Carya*; 1% *Costanea*; 1% *Cinnamomum*; 1% *Cotinus*; 1% *Diospyros*; 1% *Eriobotrya*; 1% *Fagus*; 1% *Hydrangea*; 1% *Liriodendron*; 1% *Lupinus*; 1% *Mangifera*; 1% *Myrica*; 1% *Nandina*; 1% *Origanum*; 1% *Persea*; 1% *Primula*
A4	78	15	0,5%	29% *Rhododendron*; 14% *Olea*; 12% *Fragaria*; 10% *Prunus*; 6% *Vaccinium*; 4% *Solanum*; 1% *Alnus*; 1% *Bergenia*; 1% *Camellia*; 1% *Ceanothus*; 1% *Cydonia*; 1% *Eriobotrya*; 1% *Fagus*; 1% *Fraxinus*; 1% *Hepatica*; 1% *Hordeum*; 1% *Juglans*; 1% *Malus*; 1% *Primula*; 1% *Quercus*; 1% *Rubus*; 1% *Sanguisorba*; 1% *Vitis*

PO	N° I	N° H	Δ	Hosts and incidence
A5	38	8	0.5%	16% *Carica*; 16% *Prunus*; 11% *Olea*; 11% *Pinus*; 8% *Anemone*; 5% *Boronia*; 5% *Phlox*; 3% *Coffea*; 3% *Cucumis*; 3% *Fragaria*; 3% *Leucadendron*; 3% *Mangifera*; 3% *Pistacia*; 3% *Protea*; 3% *Ranunculus*; 3% *Rosa*; 3% *Statica*; 3% *Vitis*
A6	6	2	0.20%	67% *Rhododendron*; 17% *Olea*
A7	21	1	0.00%	29% *Phormium*; 14% *Salix*; 14% *Fragaria*; 10% *Capsicum*; 5% *Syringa*; 5% *Pyrus*; 5% *Prunus*; 5% *Populus*; 5% *Acer*
A8	17	6	0.70%	35% *Citrus*; 18% *Cyphomandra*; 12% *Malus*; 12% *Solanum*; 6% *Coffea*; 6% *Lupinus*
A9	39	6	0.5%	41% *Vitis*; 28% *Fragaria*; 13% *Cyclamen*; 5% *Carica*; 3% *Cyphomandra*; 3% *Olea*; 3% *Vaccinium*
P1	3	3	0.90%	33% *Fagus*; 33% *Fragaria*; 33% *Vaccinium*
P3	6	4	0.20%	67% *Litchi*; 17% *Nephelium*; 17% *Vaccinium*
P4	5	3	0.5%	75% *Litchi*; 25% *Vitis*
P5	6	2	0.20%	50% *Eriobotrya*; 17% *Averrhoa*; 17% *Coffea*; 17% *Vitis*
P6	30	6	0.5%	47% *Fragaria*; 17% *Mangifera*; 10% *Hevea*; 10% *Vitis*; 3% *Chrysanthemum*
P8	9	1	0.90%	56% *Citrus*; 11% *Annona*; 11% *Coffea*; 11% *Mangifera*
P9	50	6	0.40%	28% *Rumohra*; 28% *Capsicum*; 18% *Lupinus*; 4% *Coffea*; 4% *Carica*; 2% *Vitis*; 2% *Syzygium*; 2% *Mangifera*; 2% *Annona*
H99	1	1	0	100% *Rhododendron*

PO	N° I	N° H	Δ	Hosts and incidence
H98	1	1	0	100% *Rhododendron*
H7	10	1	0	100% *Hakea*
H58	1	1	0	100% *Fragaria*
H55	1	1	0	100% *Litchi*
H46	1	1	0	100% *Fragaria*
H44	1	1	0	100% *Leucospermum*
H36	1	1	0	100% *Kurtovska*
H27	8	1	0	38% *Rhododendron*; 25% *Sambucus*; 13% *Rubus*; 13% *Fragaria*
H26	2	1	0	100% *Prunus*
H25	1	1	0	100% *Fragaria*
H24	1	1	0	100% *Phormium*
H23	2	1	0	100% *Salix*
H20	5	1	0	25% *Vitis*; 25% *Solanum*; 25% *Capsicum*; 25% *Persea*
H120	1	1	0	100% *Vaccinium*
H126	1	1	0	100% Marine ecosystem**
H159	1	1	0	100% *Salix*
H157	1	1	0	100% *Nothofagus*

PO	N° I	N° H	Δ	Hosts and incidence
H156	1	1	0	100% *Cornus*
H147	1	1	0	Unknown
H145	1	1	0	100% *Hevea*
H130	1	1	0	100% *Vaccinium*
H121	1	1	0	100% *Acacia*
H117	1	1	0	100% *Fragaria*
H118	1	1	0	100% *Fragaria*
H116	1	1	0	100% *Prunus*
H115	1	1	0	100% *Rubus*
total	791	151	2.10%	92 genera of plants and two of insects

In first column has been reported the population (PO) in second column number of isolates (N° I) belonging to the specific population in third number of haplotypes (N° H) and in fourth nucleotide variations of sequences analysed. In the last column host range and percentage of occurrence of the populations have been reported. * Indicate insect systems and ** marine ecosystems.

3.3.4. Phylogeography

The data collected in the CaITSdb also showed a wider geographic distribution (Figure 3.4) of *C. acutatum* compared to the data published by the European and Mediterranean plant protection organisation (EPPO).

Considering the Figure 3.5 there is no clear evidence of connections between genetic groups or populations and origin. However, different geographic areas do show particular trends in population distribution. For example, isolates that belong to A5 are mainly distributed in Oceania and South Africa. Strains belonging to group A4, A6 and associated populations are mainly found in Europe. A8 and P8 populations seem to be quite specific to North and South America and also P9 is predominantly associated with North and South American countries. A9 population shows a high geographic specificity, in fact 85% of the isolates of this genetic group have been isolated in Oceania.

More than 50% of the isolates belonging to group A7 have also been isolated in Oceania; this geographic area is also showing the highest rate of diversity through the species complex; in fact almost all populations have representative strains isolated from Australia or New Zealand. This final evidence, and in particular the distribution of *C. acutatum* populations, is concordant with the results obtained with the phylogenetic analysis and suggests Oceania as possible geographic origin of the pathogen.
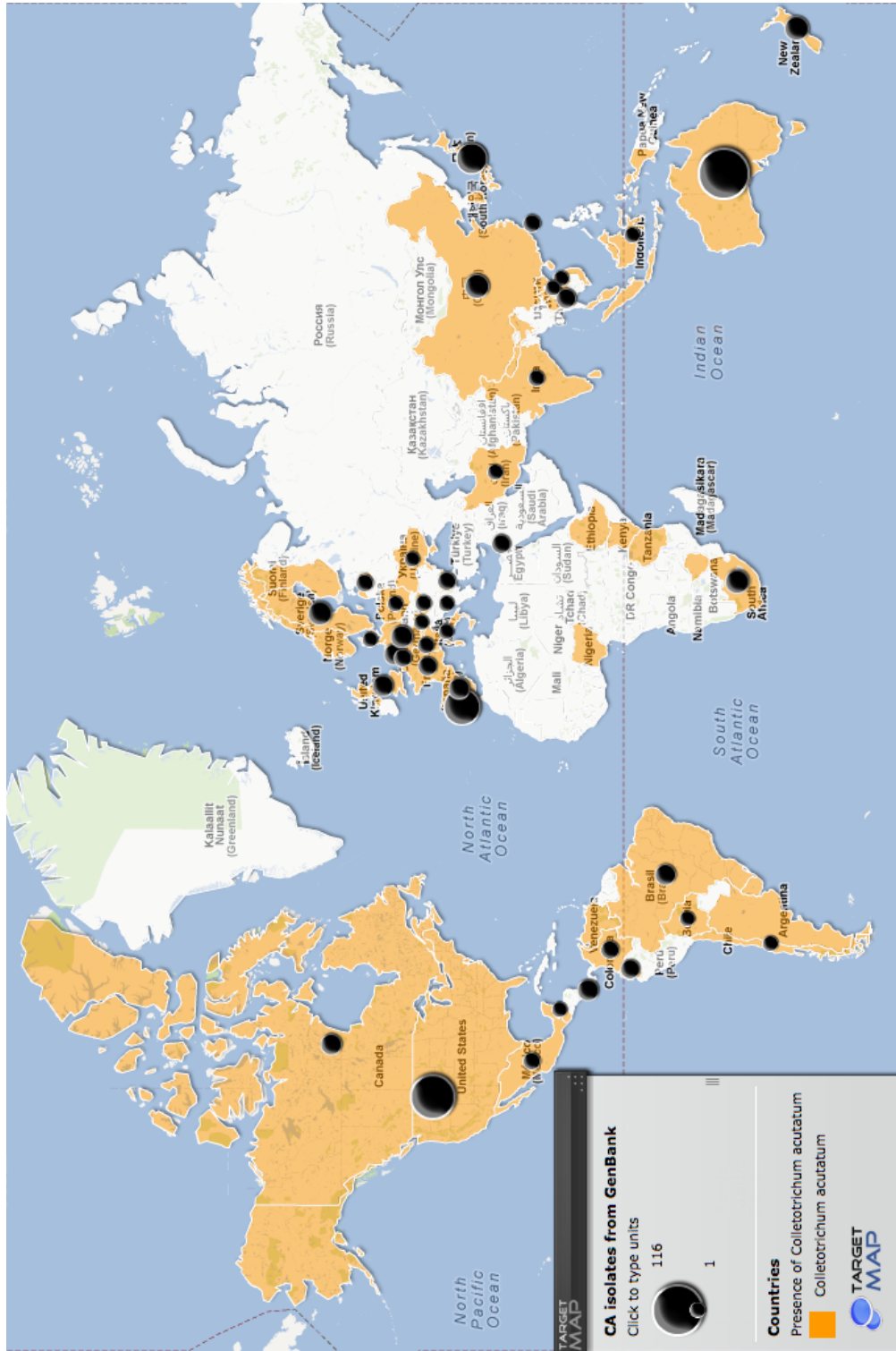
Figure 3.4. Map showing distribution of the isolates analysed in the CaITSdb and of *Colletotrichum acutatum* combining information with the EPPO report in 2006. Black spots are related to the information of the isolates used in the CaITSdb. Circles are proportional to the number of strains isolated in a specific country. The map has been made using the free web tool Target Map (https://www.targetmap.com/viewer.aspx?reportId=17213).
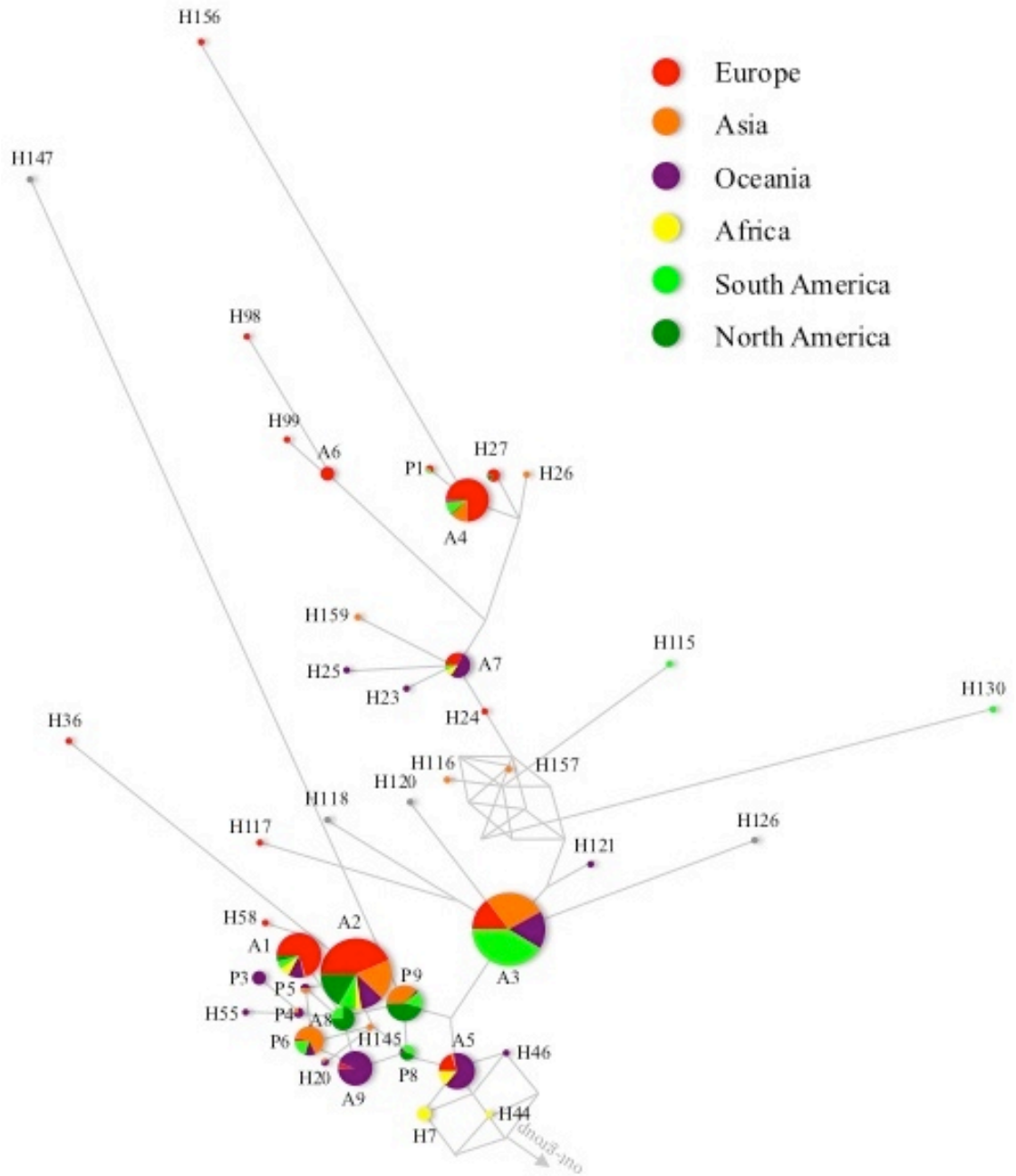
Figure 3.5. Median-joining network showing the geographic distribution of populations belonging to the species complex of *Colletotrichum acutatum* rRNA sequences. Circles areas are proportional to the genetic variability of the populations and length of lines is proportional to the genetic distances between populations.

3.4. Discussion

Colletotrichum acutatum is an important pathogen causing economically significant losses of crops. *C. acutatum* has a wide plant host range in both crops and natural ecosystems, and its capability to infect different types of host, such as insects, has also been described. *C. acutatum* is able to develop different types of interaction with plant host: covering biotrophic, necrotrophic and hemibiotrophic infections. It is also capable of growing as a non-pathogen. The life styles of *C. acutatum* species complex can include sexual, both homothallic and heterothallic (teleomorph *Glomerella*) and asexual states. Furthermore, sexual behavior in *Glomerella* is more complicated than in most ascomycetes and strains within the same species do not show a typical MAT1-1/2 system. All this evidence and complexity suggest *C. acutatum* is an important model system for studying evolution and speciation process in fungi and in particular plant pathogens. The importance of studying the evolution of this organism is also based in the actual limits of reliable diagnostic tools. For example, in the United Kingdom, *C. acutatum* has been considered a quarantine pathogen since early 80s but at least one isolate belonging to *C. acutatum sensu lato*, deposited and described as *Glomerella phacidiomorpha*, was isolated in the country in 1935 (Farr *et al.*, 2006).

Results obtained in the initial bioinformatics analysis provided useful information regarding the evolution of *C. acutatum*. Evolutionary network of more than 800 rRNA sequences deposited in GenBank provided an initial view of genetic distribution of sub-populations and their evolutionary relationships. These results confirmed the sub-division of these taxa in several genetics groups.

The evolutionary analyses also suggested that isolates belonging to A5 are relatively close to a hypothetical ancestor. Populations also showed differences in mating behavior; all known heterothallic isolates belong to group A5 and to the closest related group A3 (par. 1.2.1 and 2.4.7). The network also suggests a progressive loss of mating capability. In fact heterothallic isolates belonged to two distinct groups: A3 and A5 (Guerber and Correll 2001) are evolutionarily linked to strains belonging to A7 (*Glomerella salicis*, syn. *Glomerella myabeana*) characterized by homothallic isolates. Further more all the other populations seem to have lost the mating capability. As reviewed by Lee (Lee *et al.*, 2010), sexual self-compatibility seems to be the ancestral mating style in fungi. However, unbalanced heterothallism is conserved, unique and spread across this genus and in *Colletotrichum* organisms it is still uncertain which mating behavior is ancestral. A lack of knowledge of the genetic basis of the mating behaviour of this organism makes prediction of the origin of mating capability in this organism more complicated. Geographic distribution of isolates analysed also suggest Oceania as possible origin of this pathogen. This region showed the highest level of variability and groups related to an hypothetical ancestral population are mainly distributed in these countries.

CHAPTER 4

EVOLUTIONARY RELATIONSHIPS

IN *COLLETOTRICHUM ACUTATUM*

SENSU LATO

4.1 <u>INTRODUCTION</u>

In the genus *Colletotrichum* (teleomorph *Glomerella*, Phylum Ascomycota) the species concepts of both the anamorphic and teleomorphic form are poorly defined. Identification of *Colletotrichum* spp. has been conventionally performed using classical mycological methods based on morphological characters (MSC) such as shape and size of conidia, setae, appressoria and sclerotia together with origin (ESC) both geographically and related to host association patterns. Using this system around 900 species were assigned to *Colletotrichum* (reviewed by Sutton, 1992). These criteria alone are not always sufficient to differentiate species due to the variations in morphology and phenotype among species under different environmental conditions (Than *et al.*, 2008).

Since 1991, a number of molecular methods has been developed and broadly implemented for more accurate characterization and discrimination between *Colletotrichum* spp. These include isoenzyme comparisons, restriction fragment length polymorphisms (RFLP) analysis of mitochondrial DNA, amplified fragment length polymorphisms (AFLP), AT-rich analysis, random amplified polymorphic DNA (RAPD), *Colletotrichum* genus-specific and species-specific PCR primers and enzyme-linked immunosorbent assay (Garrido *et al.*, 2008). Multigene phylogenetic analysis has been shown to be informative in *Colletotrichum* species delineation, even if the species concept in its pure sense has not been appropriately applied to this system due to a lack of knowledge in evolutionary pathways, reproductive strategies and host association patterns.

DNA sequence analyses have been used to differentiate species and clarify the taxonomic complications of some fungal genera, e.g. *Fusarium* (O'Donnell *et*

al., 1998) and *Colletotrichum* (Sreenivasaprasad *et al.*, 1996). Analysis of the nucleic acid should provide the most reliable structure to build a classification of *Colletotrichum* because DNA characters are not directly influenced by environmental factors (Cannon *et al.*, 2000). A combination of molecular diagnostic tools and traditional methods such as morphological characterization and pathogenicity tests is a suitable and consistent approach for studying species complexes in *Colletotrichum* (Cannon *et al.*, 2000).

Many studies have been carried out on *C. acutatum*; it has been accepted as a species complex (or group species) rather than single species. An integrated approach on BSC, MSC and molecular data (PSC and GCPSR) showed the presence of distinct genetic groups. Many groups have started to characterize and describe different genetic sub-groups of *C. acutatum* as cryptic species. The first subgroup belonging to *C. acutatum* species complex described as a different species has been A1 as *C. lupini*. The fungus has been characterized morphologically, physiologically as well as by RAPD-PCR and DNA-sequencing and described as two different forms, *Colletotrichum lupini* Nirenberg, Feiler & Hagedorn, comb. nov. var. *lupini* and *Colletotrichum lupini* var. *setosum* Nirenberg, Feiler & Hagedorn var. nov. (Nirenberg *et al.*, 2002). More recently, Shivas and Tan published a study on taxonomic re-assessment of *Colletotrichum acutatum,* introducing *C. fioriniae* comb. et stat. nov. for isolates belonging to A3 group and *C. simmondsii* sp. nov. for isolates belonging to A2 group. Last publication focused on the taxonomy of *C. acutatum sensu lato* has been by Faedda *et al.* (2011) where isolates belonging to A4 have been described as *C. clavatum* sp. nov. (Faedda *et al.*, 2011). Despite the attention focused on

taxonomy and some re-assessment of the taxonomic entities (Damm *et al.*, 2012)

the evolutionary relationships of this taxa remains a largely unexplored area.

4.2 OBJECTIVES AND APPROACH

The target of this research is to improve the knowledge of the evolutionary relationships of *C. acutatum* species complex. The approach is based on an integrated application of bioinformatics, molecular biology and mycological tools.

To achieve this, the focus is on:

- Assembling a collection of isolates from different hosts and geographic locations, with different mating behavior and genetically representative of the global populations based on results obtained in the previous chapter

- Cultural analyses to evaluate reliability of this tools for species characterization

- Analysis of the genetic variability of *C. acutatum* global population through a multi-locus sequencing to evaluate levels of speciation process

- Investigate the evolutionary relationships to link above data with geographic origin and host association patterns.

4.3 RESULTS

4.3.1 Collection of isolates

4.3.1.1 Isolates representative of global genetic diversity

Where possible, for each genetic group and populations previously identified a set of strains isolated from different hosts and in different geographic locations has been collected. In total the subset of strains have been isolated from 38 different genera of plant hosts collected from 28 different countries in 5 different continents (Table 4.1).

4.3.1.2 Isolates showing different mating behavior

The only two strains characterized as homothallic in nature and that have been associated with *Glomerella acutata* (or associated taxa) have been collected from the authors and included in the collection. One strain, 9178, was isolated from highbush blueberry (*Vaccinium corymbosum*) in Norway (Talgø *et al.*, 2007) and the second, MP1, on Norway maple (*Acer platanoides*) close to Boston, USA (LoBuglio and Pfister, 2008). Also 6 strains (PJ4, PJ8, PJ7, ATCC 56813, ATCC MYA-662, ATCC MYA-663) have been identified as self-sterile and used as mating testers based on their capability to cross with each other and with others strains (Guerber and Correll, 2001). Three of them were obtained from Dr. Peter Johnson (PJ), the scientist who first isolated them and three were bought from the American type culture collection (ATCC).

Table 4.1. Strains of *Colletotrichum acutatum sensu lato* and related information used in this study.

Original code	Genus	complex	Country of Origin	Area	Year	Host
PT227[4]	*Colletotrichum*	*acutatum*	Portugal	Faro	2003	*Olea europaea*
PD90-443[4]	*Colletotrichum*	*acutatum*	Netherlands	N/A	N/A	*Phlox sp.*
PJ8[2]	*Colletotrichum*	*acutatum*	New Zealand	Nelson	1987	*Pyrus pyrifolia*
Ca 287[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Statice sp.*
PT712[4]	*Colletotrichum*	*acutatum*	Portugal	N/A	2004	*Olea europaea*
PT796[4]	*Colletotrichum*	*acutatum*	Portugal	N/A	2005	*Olea europaea*
CBS 112759[6]	*Colletotrichum*	*acutatum*	South Africa	Eastern Cape	N/A	*Hakea sericea*
CBS 112980[6]	*Colletotrichum*	*acutatum*	South Africa	Southern Cape	N/A	*Pinus radiata*
CBS 112996[6]	*Colletotrichum*	*acutatum*	Australia	N/A	N/A	*Carica papaya*
CBS 113008[6]	*Colletotrichum*	*acutatum*	South Africa	Western Cape	N/A	*Hakea sericea*
CBS 144.29[6]	*Colletotrichum*	*acutatum*	Sri Lanka	N/A	N/A	*Capsicum annuum*
CBS 370.73[5]	*Colletotrichum*	*acutatum*	New Zealand	Tokoroa	N/A	*Pinus radiata*
CBS 292.67[6]	*Colletotrichum*	*acutatum*	Australia	Queensland	N/A	*Capsicum annuum*
PD85-694[4]	*Colletotrichum*	*acutatum*	Netherlands	N/A	N/A	*Chrysanthemum sp.*
CBS 211.78[6]	*Colletotrichum*	*acutatum*	Costa Rica	Turrialba	N/A	*Coffea spp.*
PT170[4]	*Colletotrichum*	*acutatum*	Portugal	Torres Vedras	2002	*Olea europaea*
CR46[4]	*Colletotrichum*	*acutatum*	Portugal	LPVVA Lisbon	2000	*Vitis vinifera*
AC54[4]	*Colletotrichum*	*acutatum*	Portugal	LPVVA Lisbon	N/A	*Vitis vinifera*

Original code	Genus	complex	Country of Origin	Area	Year	Host
PJ4[2]	*Colletotrichum*	*acutatum*	New Zealand	Bay of Planty	1988	*Actinida deliciosa*
PJ7[2]	*Colletotrichum*	*acutatum*	New Zealand	Auckland	1988	*Fragaria x ananassa*
ATCC 56813[5]	*Colletotrichum*	*acutatum*	Australia	N/A	N/A	*Persea americana*
ATCC MYA-662[5]	*Colletotrichum*	*acutatum*	USA	Arkansas	N/A	*Malus domestica*
ATCC MYA-663[5]	*Colletotrichum*	*acutatum*	USA	Virginia	N/A	*Malus domestica*
JC51[4]	*Colletotrichum*	*acutatum*	UK	N/A	2003	*Tulipa sp.*
Ca 302a[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Nandina domestica*
Ca 473[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Liriodendron tulipifera*
Ca 318[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Magnolia sp.*
IMI345578[7]	*Colletotrichum*	*acutatum*	New Zealand	Auckland, Papakura	1988	*Fragaria ananassa*
IMI346585[7]	*Colletotrichum*	*acutatum*	N/A	N/A	N/A	N/A
Ca 1049[1]	*Colletotrichum*	*acutatum*	France	N/A	1988	*Fragaria x ananassa*
CBS 797.72[6]	*Colletotrichum*	*acutatum*	New Zealand	N/A	N/A	*Pinus radiata*
JG05[4]	*Colletotrichum*	*acutatum*	France	Paris	N/A	*Ceanothus sp.*
TN47[4]	*Colletotrichum*	*acutatum*	Portugal	LPVVA Lisbon	2000	*Eriobotrya japonica*
CBS 193.32[6]	*Colletotrichum*	*acutatum*	Italia	N/A	1932	*Olea europaea*
Ca 1053[1]	*Colletotrichum*	*acutatum*	Netherlands	N/A	N/A	*Fragaria x ananassa*
RB-Ap-3*	*Colletotrichum*	*acutatum*	UK	N/A	2008	*Malus domestica*

Original code	Genus	complex	Country of Origin	Area	Year	Host
RB-Ap-4*	Colletotrichum	acutatum	UK	N/A	2008	Malus domestica
Ca 1430[1]	Colletotrichum	acutatum	Norway	N/A	N/A	Fragaria vesca
Ca 1432[1]	Colletotrichum	acutatum	Norway	Østfold county	1999	Fragaria x ananassa
RB-Vi-1*	Colletotrichum	acutatum	UK	N/A	2010	Vitis spp.
JL198[4]	Colletotrichum	acutatum	Serbia	N/A	2003	Olea europaea
PT743[4]	Colletotrichum	acutatum	Spain	N/A	2004	Olea europaea
IMI345026[7]	Colletotrichum	acutatum	Spain	N/A	1991	Fragaria x ananassa
ALM-NRB-30K[10]	Colletotrichum	acutatum	Isreal	Southern Israel	N/A	Prunus dulcis
ALM-BZR-9A[10]	Colletotrichum	acutatum	Isreal	Southern Israel	N/A	Prunus dulcis
CBS 126527[6]	Colletotrichum	acutatum	UK	Harpenden Herts	N/A	Prunus avium
CBS 198.35[6]	Colletotrichum	acutatum	UK	N/A	N/A	Phormium spp.
CBS 112989[6]	Colletotrichum	acutatum	India	N/A	N/A	Hevea brasiliensis
PT30[4]	Colletotrichum	acutatum	Portugal	Azores	1999	Lupinus albus
CMG12[4]	Colletotrichum	acutatum	Portugal	Lisbon	1996	Cinnamomium zeylanicum
Ca 1294[1]	Colletotrichum	acutatum	UK	York	1998	Lupinus polyphyllus
G52[4]	Colletotrichum	acutatum	Germany	N/A	1995	Lupinus albus
96A649[4]	Colletotrichum	acutatum	Australia	West Australia	1996	Lupinus polyphyllus
HY09[4]	Colletotrichum	acutatum	Canada	N/A	N/A	Lupinus albus

Original code	Genus	complex	Country of Origin	Area	Year	Host
C3[4]	Colletotrichum	acutatum	Poland	N/A	1997	*Lupinus lateus*
SHK788[4]	Colletotrichum	acutatum	South Africa	N/A	N/A	*Lupinus albus*
70555[4]	Colletotrichum	acutatum	Chile	N/A	N/A	*Lupinus albus*
70399[4]	Colletotrichum	acutatum	Austria	N/A	N/A	*Lupinus albus*
PT702[4]	Colletotrichum	acutatum	Spain	N/A	N/A	*Olea europaea*
IMI350308[7]	Colletotrichum	acutatum	UK	Kent	1991	*Lupinus spp.*
CBS 109221[6]	Colletotrichum	acutatum	Germany	N/A	N/A	*Lupinus albus*
CBS 109225[6]	Colletotrichum	acutatum	Ukraine	N/A	N/A	*Lupinus albus*
Ca 2897[1]	Colletotrichum	acutatum	Australia	Weat Australia	1996	*Fragaria x ananassa*
Ca 397[1]	Colletotrichum	acutatum	USA	N/A	N/A	*Fragaria x ananassa*
B88[4]	Colletotrichum	acutatum	UK	N/A	1989	*Fragaria x ananassa*
PT135[4]	Colletotrichum	acutatum	Portugal	Elvas	2001	*Olea europaea*
Ca 1079[1]	Colletotrichum	acutatum	UK	N/A	N/A	*Fragaria x ananassa*
Ca 892[1]	Colletotrichum	acutatum	UK	Nottinghamshire	N/A	*Fragaria x ananassa*
Ca 455[1]	Colletotrichum	acutatum	UK	N/A	N/A	*Photinia sp.*
IMI299103[7]	Colletotrichum	acutatum	UK	Kent	1985	*Fragaria vesca*
Ca 1020[1]	Colletotrichum	acutatum	Kenia	Nairobi	1986	*Fragaria vesca*
IMI311743[7]	Colletotrichum	acutatum	USA	Ohio	1986	*Fragaria x ananassa*

Original code	Genus	complex	Country of Origin	Area	Year	Host
IMI335544[7]	Colletotrichum	acutatum	Italy	Trento	1989	Fragaria x ananassa
Ca 1005[1]	Colletotrichum	acutatum	France	N/A	1991	Fragaria x ananassa
IMI345028[7]	Colletotrichum	acutatum	Colombia	N/A	1991	Fragaria x ananassa
IMI345029[7]	Colletotrichum	acutatum	Costa Rica	N/A	1991	Fragaria x ananassa
Ca 1034[1]	Colletotrichum	acutatum	Costa Rica	N/A	1991	Fragaria x ananassa
IMI345031[7]	Colletotrichum	acutatum	Italy	N/A	1991	Fragaria x ananassa
Ca 1090[1]	Colletotrichum	acutatum	USA	Louisiana	1985	Fragaria x ananassa
IMI348177[7]	Colletotrichum	acutatum	USA	Missisipi	1981	Fragaria x ananassa
IMI348490[7]	Colletotrichum	acutatum	France	N/A	1984	Fragaria x ananassa
Ca 1086[1]	Colletotrichum	acutatum	France	N/A	N/A	Fragaria x ananassa
IMI360928[7]	Colletotrichum	acutatum	Swiss	Zurich	1993	Fragaria x ananassa
CBS 125973[6]	Colletotrichum	acutatum	UK	Harpenden Herts	N/A	Fragaria x ananassa
CBS 526.77[6]	Colletotrichum	acutatum	Netherlands	Kortenhoefse	N/A	Nymphaea alba, leaf
AR3787	Colletotrichum	acutatum	South Africa	N/A	N/A	Phormium spp.
AR3546	Colletotrichum	acutatum	Germany	N/A	N/A	Phormium spp.
AR3389	Colletotrichum	acutatum	New Zealand	N/A	N/A	Phormium spp.
CBS 102054[6]	Colletotrichum	acutatum	New Zealand	Auckland	N/A	Phormium
PT250[4]	Colletotrichum	acutatum	Portugal	Mirandela	2003	Olea europaea

Original code	Genus	complex	Country of Origin	Area	Year	Host
9178[12]	Colletotrichum	acutatum	Norway	Vest-Agder county	2004	Vaccinium corymbosum
MP1[3]	Colletotrichum	acutatum	USA	Boston	2006	Acer Platanoides
MP3[3]	Colletotrichum	acutatum	USA	Boston	2006	Acer Platanoides
SS1[3]	Colletotrichum	acutatum	USA	Boston	2006	Acer Platanoides
VPRI 32545[11]	Colletotrichum	acutatum	Australia	Tumut River, NSW	2005	Salix fragilis
CBS 180.97[6]	Colletotrichum	acutatum	Netherlands	Bergeijk	N/A	Populus canadensis
Ca 1044[1]	Colletotrichum	acutatum	New Zealand	N/A	N/A	Fragaria x ananassa
CBS 607.94[6]	Colletotrichum	acutatum	Netherlands	Z. Flevolan	1994	Salix spp.
CBS 192.56[6]	Colletotrichum	acutatum	Germany	West-Germany	N/A	Salix spp.
PD89-582[4]	Colletotrichum	acutatum	Netherlands	N/A	N/A	Cyclamen sp.
Ca 1046[1]	Colletotrichum	acutatum	Australia	Brisbane	1967	Fragaria x ananassa
CBS 122122[6]	Colletotrichum	acutatum	Australia	Queensland	N/A	Carica papaya
CBS 294.67[6]	Colletotrichum	acutatum	Australia	Queensland	N/A	Carica papaya
CBS 101611[6]	Colletotrichum	acutatum	Costa Rica	N/A	N/A	Fern
OCO-ARC-4[9]	Colletotrichum	acutatum	USA	Florida, Arcadia	N/A	Citrus x sinensis
STF-FTP-10[9]	Colletotrichum	acutatum	USA	Florida, Frostproof	N/A	Citrus x sinensis
Coll-25[8]	Colletotrichum	acutatum	Taiwan	Tainan	1998	Capsicum annum
Coll-154[8]	Colletotrichum	acutatum	Taiwan	Tainan	1992	Capsicum annum

Original code	Genus	complex	Country of Origin	Area	Year	Host
Ca 593[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Spinacia oleracea*
Ca 739[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Tom-21[4]	*Colletotrichum*	*acutatum*	Colombia	Antioquia	1998	*Cyphomandra betacea*
Tom-12[4]	*Colletotrichum*	*acutatum*	Colombia	Antioquia	1998	*Cyphomandra betacea*
Tom-009[4]	*Colletotrichum*	*acutatum*	Colombia	Antioquia	1998	*Cyphomandra betacea*
Ca 1500[1]	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Dianthus sp.*

Isolates references:

1 - Charles Lane, Food and Environment Research Agency - UK
2 - Peter Johnston, Manaaki Whenua Landcare Research - New Zealand
3 - Katherine Lobuglio, Harvard University Herbaria - USA
4 - available in the department, The University of Warwick - UK
5 - bought from the American type coltures collection - USA
6 - Ulrike Damm, The Centraalbureau voor Schimmelcultures - The Netherlands
7 - Paul Cannon, CABI Bioscience Centre Herbarium - UK
8 - Zong-ming Sheu, The World Vegetable Center - Taiwan
9 - Peres Natalia, University of Florida - USA
10 - Stanley Freeman, The Volcani Center - Israel
11 - James Cunnington, National Collection of Fungi, Knoxfield Herbarium - Australia
12 - Gunn Mari Stromeng, Bioforsk Plant Health and Plant Protection Division - Norway
* - isolated in this study

4.3.2   Morphological and cultural studies

4.3.2.1 <u>Colony aspects and characteristics</u>

Isolates were chosen on the basis of genetic variability (based on CaITSdb, Chapter 3) and distribution. Characteristics such as variation in the colour of pigments within the agar, amount and colour of the aerial mycelium, and the presence or absence of differentiated structures and teleomorphic perithecia have been investigated and recorded.
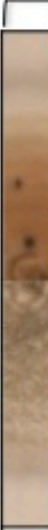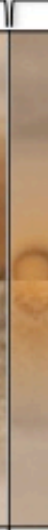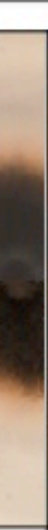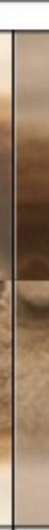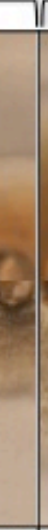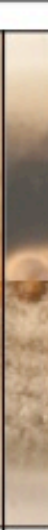
All isolates observed were representative of *C. acutatum sensu lato*, as defined in the literature (Lardner *et al.*, 1999). In the subset of isolates studied, conidia commonly developed within the aerial mycelium, conidia were typically elliptical, or if cylindrical then acute at one or both ends, and tapering towards one end, and setae were generally absent and rare when present.

Results showed some common characteristics of strains belonging to the same genetic group for the subset analysed (Table 4.2). The intra-group variability is relatively high, which makes it difficult to categorize isolates into different groups based on this aspect alone. Based on the genetic groups identified some characters seem to be typical of different clusters.

For example only, all the strains belonging to groups A3 and A5 are able to produce red pigments in the media and large spore masses organized in orange concentric bands. A7 isolates are able to produce perithecia, and on plates the strains look darker compared with others. Nevertheless, colony aspect is closely related to factors such as media, temperature, photoperiod, *etc.* and can vary even using same growth conditions. All this evidence, combined with the high rate of
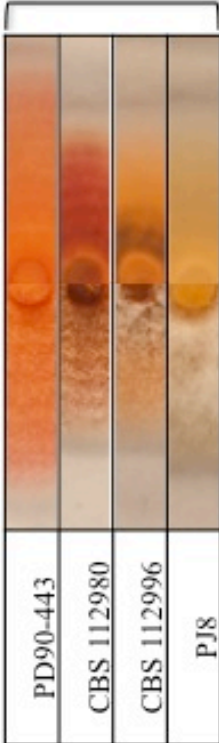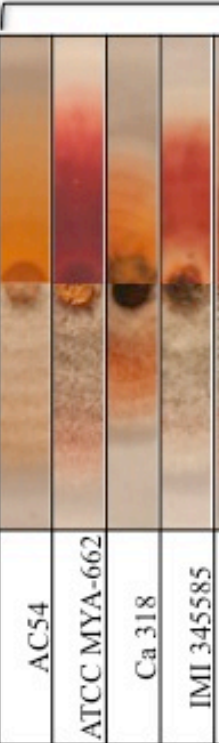
intra-genetic group variability makes it difficult to use colony morphology as a useful tool for identification. Descriptions from cultures on PDA for each morphological group and representative pictures are illustrated below (Table 4.2).

Table 4.2. Morphological variability of *Colletotrichum acutatum sensu lato* strains

Strain	Code	Description
PT250	A6	Grey or white cottony aerial mycelium, becoming darker from the inoculum with age; In reverse from light yellow to dark brown with age. Light yellowish conidial masses on the colony surface
CBS 198.35	H24	White cottony aerial mycelium; In reverse from white to light yellow. Light yellowish conidial masses on the colony surface.
AR3546	A7	Very variable. White or grey cottony aerial mycelium; light brown, dark brown and pink. In four isolates analysed, presence of perithecia on the colony surface and inside the media.
AR3787		
AR3389		
MP3		
IMI 345585		
9178		
CBS 180.97		
CBS 607.94		
VPRI 32545	H23	White cottony aerial mycelium; light brown to pink.
RB-Ap-3	H27	White cottony aerial mycelium; In reverse from grey to dark brown with age
ALM-NRB-30K	H26	White cottony aerial mycelium; In reverse from white to dark orange with age
CBS 193.32	A4	White thick cottony aerial mycelium. In reverse from light brown to dark brown with age Light orange conidial masses on the colony surface and presence of dark melanised structures similar to acervuli.
NI90		
Ca 1053		
JL198		
CBS 144.29		
PT227	A5	White\pinkish cottony aerial mycelium. In reverse, yellow or from light red to dark red with age. Orange conidial masses in big drops on the colony surface in concentric band. Presence of dark melanised structures similar to acervuli.

Table 4.2. Morphological variability of *Colletotrichum acutatum sensu lato* strains

Strain	Group	Description
PD90-443	A5	White\pinkish cottony aerial mycelium. In reverse, yellow or from light red to dark red with age. Orange conidial masses in big drops on the colony surface in concentric band. Presence of dark melanised structures similar to acervuli.
CBS 112980		
CBS 112996		
PJ8		
AC54	A3	White\pinkish thick cottony aerial mycelium. In reverse from yellow to red. Orange conidial masses in big drops on the colony surface in concentric band. Presence of dark melanised structures similar to acervuli.
ATCC MYA-662		
Ca 318		
IMI 345585		
CR46		
CBS 797.72		
OCO-ARC-4	P8	White thick cottony aerial mycelium; In reverse from white to grey with age
CBS 101611	P9	White aerial mycelium; In reverse from white to orange.
Tom-12	A8	White thick aerial mycelium; In reverse from white to yellow. Light yellowish conidial masses on the colony surface.
Tom-21		
CBS109225	A1	White cottony aerial mycelium, becoming grey\brownish from the inoculum with age; In reverse from light brown to dark brown with age. Light orange conidial masses on the colony surface and presence of dark melanised structures similar to acervuli
HY09		
PT30		
PT702		

Table 4.2. Morphological variability of *Colletotrichum acutatum sensu lato* strains

Strain	Image	Group	Description
CBS 122122			
CBS 294.67		A9	White and grey thick cottony aerial mycelium; In reverse from yellow to dark brown with age.
PD89- 582			
CBS 112989		P6	White aerial mycelium; In reverse from white to brown or dark gray
PD85-694			
Coll-25		P9	White aerial mycelium; In reverse from white to brown
IMI335544			
PT135			White cottony aerial mycelium; In reverse from yellow to dark brown with age. Light orange conidial masses on the colony surface and presence of dark melanised structures similar to acervuli.
CA455		A2	
CBS 526.77			
CA397			
IMI345028			

in first column is reported the strain code, pictures on the left are upper side of the colony and on the right reverse side Genetic groups and description of the colony are on the right of the table

4.3.2.2 <u>Growth rates</u>

Growth rate at 15, 20, 25 and 30°C of the set of isolates has been measured. Growth was expressed as a rate of mm/day (Table 4.3). The growth rate of each fungal strain at each temperature was calculated for each replicate as the slope of a line as determined by linear regression using GraphPad Prism version 5. The time interval considered was the phase in which growth obseved in a constant linear manner.

Table 4.3. Growth rate of *Colletotrichum acutatum sensu lato* strains

C. acutatum genetic groups

Strain	15°C			20°C			25°C			30°C			Population
	Growth rate	Average	Stand error	Growth rate	Average	Stand error	Growth rate	Average	Stand error	Growth rate	Average	Stand error	
PT250	0.09	0.09	-	0.14	0.14	-	0.11	0.11	-	0.00	0.00	-	A6
CBS 198.35	0.08	0.08	-	0.11	0.11	-	0.08	0.08	-	0.00	0.00	-	H24
AR3546	0.11			0.15			0.14			0.00			
AR3787	0.11	0.11	0.00	0.17	0.16	0.01	0.14	0.14	0.00	0.00	0.00	0.00	A7b
AR3389	0.11			0.15			0.13			0.00			
MP3	0.10			0.14			0.18			0.01			
Ca 1044	0.08			0.15			0.18			0.04			
9178	0.09	0.10	0.01	0.16	0.15	0.00	0.20	0.18	0.01	0.04	0.05	0.01	A7a
CBS 180.97	0.11			0.16			0.18			0.08			
CBS 607.94	0.10			0.14			0.16			0.06			
VPRI 32545	0.12	0.12	-	0.18	0.18	-	0.19	0.19	-	0.05	0.05	-	H23
Apple3	0.10	0.10	-	0.13	0.13	-	0.13	0.13	-	0.00	0.00	-	H27
ALM-NRB-30K	0.06	0.06	-	0.08	0.08	-	0.10	0.10	-	0.00	0.00	-	H26
CBS 193.32	0.10			0.13			0.15			0.00			
NI90	0.10	0.11	0.00	0.16	0.15	0.01	0.20	0.17	0.01	0.01	0.01	0.00	A4
Ca 1053	0.11			0.17			0.18			0.00			
JL198	0.11			0.15			0.16			0.01			
CBS 144.29	0.05			0.16			0.21			0.11			
PT227	0.08			0.14			0.17			0.08			
PD90-443	0.07	0.06	0.01	0.16	0.14	0.01	0.20	0.17	0.01	0.09	0.07	0.02	A5
CBS 112980	0.03			0.09			0.15			0.04			
CBS 112996	0.08			0.13			0.17			0.06			
PJ8	0.06			0.13			0.12			0.05			

C. acutatum genetic groups   C. gloeosporioides

Strain	15°C Growth rate	15°C Average	15°C Stand error	20°C Growth rate	20°C Average	20°C Stand error	25°C Growth rate	25°C Average	25°C Stand error	30°C Growth rate	30°C Average	30°C Stand error	Population
AC54	0.06	0.07	0.01	0.15	0.14	0.01	0.16	0.18	0.01	0.07	0.09	0.02	A3
MYA-663	0.08			0.16			0.20			0.03			
Ca 318	0.03			0.09			0.14			0.09			
IMI345578	0.10			0.16			0.20			0.14			
CR46	0.08			0.12			0.19			0.14			
RB189	0.08			0.14			0.18			0.09			
CBS 797.72	0.09	0.09	-	0.14	0.14	-	0.18	0.18	-	0.08	0.08	-	P8
CBS 101611	0.09	0.07	0.03	0.13	0.12	0.01	0.14	0.13	0.01	0.04	0.05	0.01	P9
Coll-25	0.04			0.11			0.12			0.06			
Tom-12	0.07	0.07	0.00	0.13	0.12	0.01	0.16	0.15	0.01	0.01	0.01	0.00	A8
Tom-21	0.07			0.11			0.14			0.01			
CBS 109225	0.08	0.09	0.00	0.12	0.13	0.01	0.15	0.16	0.01	0.05	0.06	0.00	A1
HY09	0.09			0.13			0.13			0.06			
PT30	0.10			0.15			0.18			0.06			
PT702	0.08			0.13			0.16			0.07			
CBS 122122	0.03	0.05	0.01	0.06	0.10	0.02	0.08	0.13	0.02	0.05	0.09	0.02	A9
CBS 294.67	0.06			0.11			0.16			0.10			
PD89-582	0.05			0.12			0.14			0.12			
CBS 112989	0.06	0.06	0.01	0.11	0.10	0.02	0.16	0.14	0.02	0.12	0.11	0.02	P6
PD85-694	0.05			0.08			0.11			0.09			
IMI335544	0.08	0.08	0.00	0.14	0.13	0.01	0.17	0.17	0.01	0.12	0.06	0.01	A2
PT135	0.08			0.14			0.18			0.05			
Ca 455	0.10			0.15			0.19			0.08			
CBS 526.77	0.07			0.09			0.11			0.02			
Ca 397	0.08			0.14			0.17			0.05			
IMI345028	0.09			0.14			0.18			0.06			
Cg 311	0.15	0.13	0.03	0.23	0.21	0.02	0.28	0.27	0.02	0.24	0.22	0.02	Cg
Cg 386	0.10			0.19			0.25			0.20			

From the table above it is clear that generally strains of *C. gloeosporioides* (Cg311 and Cg386) grow faster than those belonging to *C. acutatum sensu lato*. Within *C. acutatum* species complex growth rate showed particular trends. Most *C. acutatum* strains showed an optimum growth temperature around 25°C. However, even if most of the isolates have a higher growth rate at 25, strains belonging to A6, H24 and A7b grow better at 20°C. This observation raises the possibility to subdivide group A7 into: A7a isolates with an optimum growth closer to 25°C and A7b isolates with an optimum closer to 20°C. Strains belonging to group A3 showed high intra-group growth variability compared to the other groups. Isolates belonging to A6, H24, A7b, and A4 did not grow at 30°C. Exploring the growth trend of isolates at different temperatures might provide useful clues as to their adaptive potential to different environmental conditions.

4.3.2.3 Investigation of homothallic isolates

All isolates used for morphological characterization have also been tested for their sexual reproductive capability. Across the sub-set of strains only few were able to produce perithecia: 9178, MP3, CBS 180.97, Ca 1044 and CBS 607.94.

Three of the four isolates were able to produce perithecia on every media, but isolate CBS 180.97 and Ca 1044 produced them only on WA and modified Czapek-Dox agar media. All the perithecia analysed were brown to black, globose, ovoid, ampulliform or obpyriform (Figure 4.1). No asci or ascospores were seen in any isolates, even in those characterized as *Glomerella acutata* by other authors on the basis that they could produce either asci or ascospores. This

may be due to the progressive loss of the capability to undergo a full reproductive cycle after prolonged maintenance in culture and/or several subcultures on artificial media. For this reason we have tested unsuccessfully strains re-isolated after passing them on different hosts.

Tests to evaluate cross-fertility of the set of strains using standard mating tester strains did not lead to the production of any perithecia. Data reported on self-sterile strains able to produce perithecia by mating crosses have been based on data published in "Characterization of *Glomerella acutata* the teleomorph of *Colletotrichum acutatum*" by Guerber and Correll in 2001. The mating testers used in the present study were from the above published work, and these strains have been included in further molecular analysis. Most of the isolates that did not differentiate perithecia produced different sized dark melanised structures similar to acervuli, from where often spore masses were rejected.
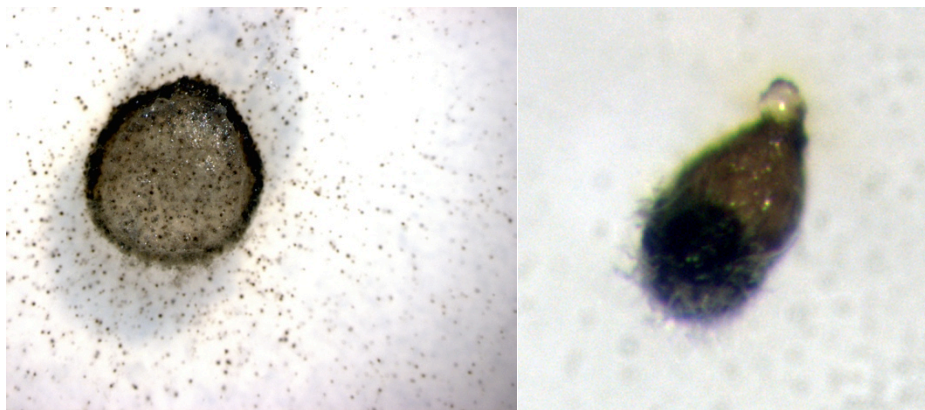
Figure 4.1. Pictures showing *Colletotrichum acutatum sensu lato* perithecia production by CBS 607.94 on a water agar plate (left) and a single perithecium (right).

4.3.3   DNA extraction and PCR amplification

Initially two DNA extraction methods described in 2.2. (10% Chelex and Sigma kit) were compared using a limited number of isolates. The quality and quantity of DNA prepared using the SIGMA kit was compared to the Chelex-method. The result of the comparison, based on a single copy locus (Beta-tubulin gene: tub2) PCR amplification, did not show significant differences between the two protocols (data not showed). Therefore, the 10% Chelex method was more useful for the purposes as it is rapid and less expensive. This method has been successfully used to extract DNA from the 120 isolates used in this study.

DNA samples were used for PCR-based analysis and to check successful amplification of four different nuclear loci:

- rRNA region (partial 18S, complete ITS1, 5.8S and ITS2 and partial 28S) - ITS

- exons 3 through 6 (introns 2 through 4) , partial sequence of the beta-tubulin 2 gene (TUB)

- part of the glyceraldehyde-3-phosphate dehydrogenase gene (intron 1 and partial CDS) - GAPDH.

- part of the mating-type gene 1-2 (intron 2 breaking the conserved domain HMG-box and partial CDS) – MAT1-2

PCR amplification products of all *Colletotrichum* isolates were assessed by gel electrophoresis and appropriate quantities were used as templates for sequencing.

4.3.4. Sequencing and alignment

Sequence data of the ITS, TUB, MAT1-2 and GAPDH region were generated from the whole set of *C. acutatum sensu lato* isolates (120 strains) as well as two stains of *C. spinaciae* and four *C. gloeosporioides* plus data related to *C. graminicola* and *C. higginsianum* obtained from available genomes published online. Comparative analysis by BlastN confirmed in all the cases the original classification of the isolates and the targeted locus. Before the phylogenetic analysis, all the sequences obtained were aligned by MAFFT and MUSCLE. After this the sequence ends were trimmed manually to obtain comparable data in all the isolates:

ITS: 430 nucleotides for *C. graminicola* and *C. higginsianum*, 417 for *C. gloeosporioides*, 418 for *C. spinaciae* and from 427 to 428 for isolates of *C. acutatum sensu lato*

TUB: 485 nucleotides for *C. graminicola*, 484 for *C. higginsianum*, 475 for *C. gloeosporioides*, 476 for *C. spinaciae* and from 475 to 479 for isolates of *C. acutatum sensu lato*

MAT1-2: 216 nucleotides for *C. graminicola*, 217 for *C. higginsianum*, 214 for *C. gloeosporioides*, 215 for *C. spinaciae* and isolates of *C. acutatum sensu lato*

GAPDH: 278 nucleotides for *C. graminicola*, 180 for *C. higginsianum*, 264 for *C. gloeosporioides*, 254 for *C. spinaciae* and from 254 to 267 for isolates of *C. acutatum sensu lato*

Information regarding the characteristics and resolution of loci analysed (such as alignment length, variability in terms of identical sites and pairwise identity and

GC content) are summarised in the table below (Table 4.4). Data reported refers

to isolates belonging to *C. acutatum sensu lato* with and without out-group.

Table 4.4. Characteristics of loci used for phylogenetic analyses.

Locus	*C. acutatum*	+ out group	*C. acutatum*	+ out group	*C. acutatum*	+ out group	*C. acutatum*	+ out group
	Alignment Length		**Identical sites**		**Pairwise identity**		**GC content**	
ITS	429	441	394 (91.8%)	347 (78.7%)	98.60%	97.30%	55.90%	5420%
TUB	481	492	400 (832%)	298 (60.6%)	96.60%	94.60%	57,50%	56.10%
MAT1-2	215	218	168 (78.1%)	74 (33.9%)	94.10%	90.30%	48.10%	47,50%
GAPDH	280	295	164 (58.6%)	70 (23.7%)	90.70%	86.00%	45.40%	4320%
CONC	1405	1446	1126 (80.1%)	789 (54.6%)	95.70%	93.10%	53.10%	51.60%

Data related to 120 *C. acutatum sensu lato* sequences (columns on the left) as well as the out-groups: two strains of *C. spinaciae*, four *C. gloeosporioides*, one *C. graminicola* and one *C. higginsianum* (columns on the right)

4.3.5 Phylogenetic analysis and evolutionary relationships

Phylogenetic analysis has been carried out using information of four loci: ITS,

TUB, GAPDH and MAT1-2 from 120 isolates. Phylogenetic analyses by

MrBayes (MB), Maximum Likelihood (ML), Maximum-parsimony (MP),

neighbor-joining (NJ) and unweighted pair group method with arithmetic mean

(UPGMA) produced similar groups and tree topologies. Therefore, only the

MrBayes results are shown (Figure 4.2 and 4.3). Branches corresponding to

partitions reproduced in less than 50% trees were collapsed. The percentage of

parsimonious trees in which the associated taxa clustered together is shown next

to the branches. Bayesian trees were obtained using distances produced with the

Jukes-Cantor model with equal rates for ITS (Jukes and Cantor, 1969) and the

K2 with gamma rates for the other loci. The tree is drawn to scale and there were

a total of 1446 positions in the final dataset.
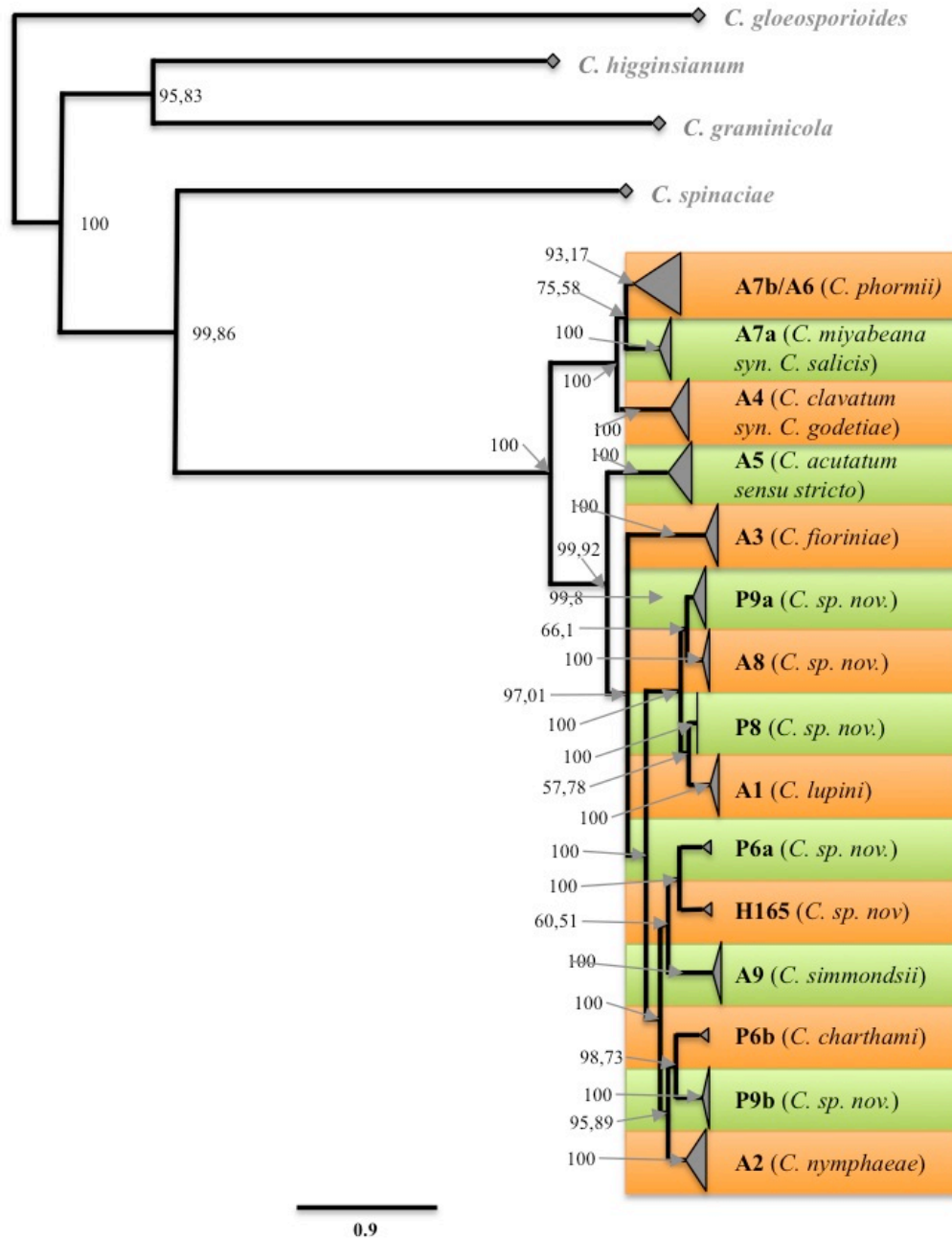
Figure 4.2. Bayesian MCMC analysis tree constructed from an alignment based on the concatenation of rRNA, TUB, MAT1-2 and GPDH partial sequences of 120 *Colletotrichum* isolates used in this study showing genetic groups identified. Numbers on nodes are bootstrap values. Tree was rooted with data for respective loci from the *C. graminicola* and *C. higginsianum* genomes and sequences of *C. gloeosporioides* and *C. spinaciae* were obtained experimentally. Potential species designations reported in literature/GenBank are shown on the right.
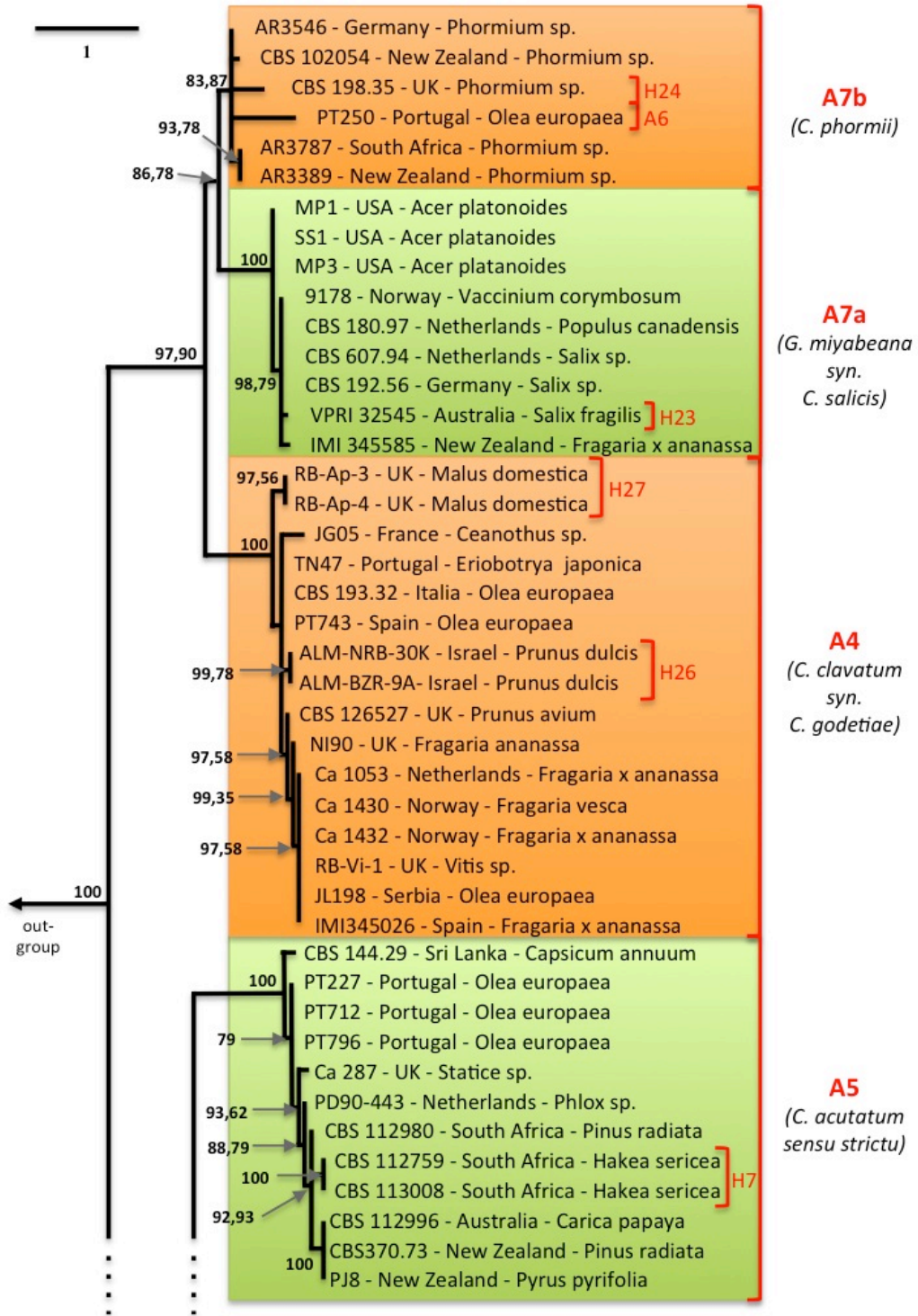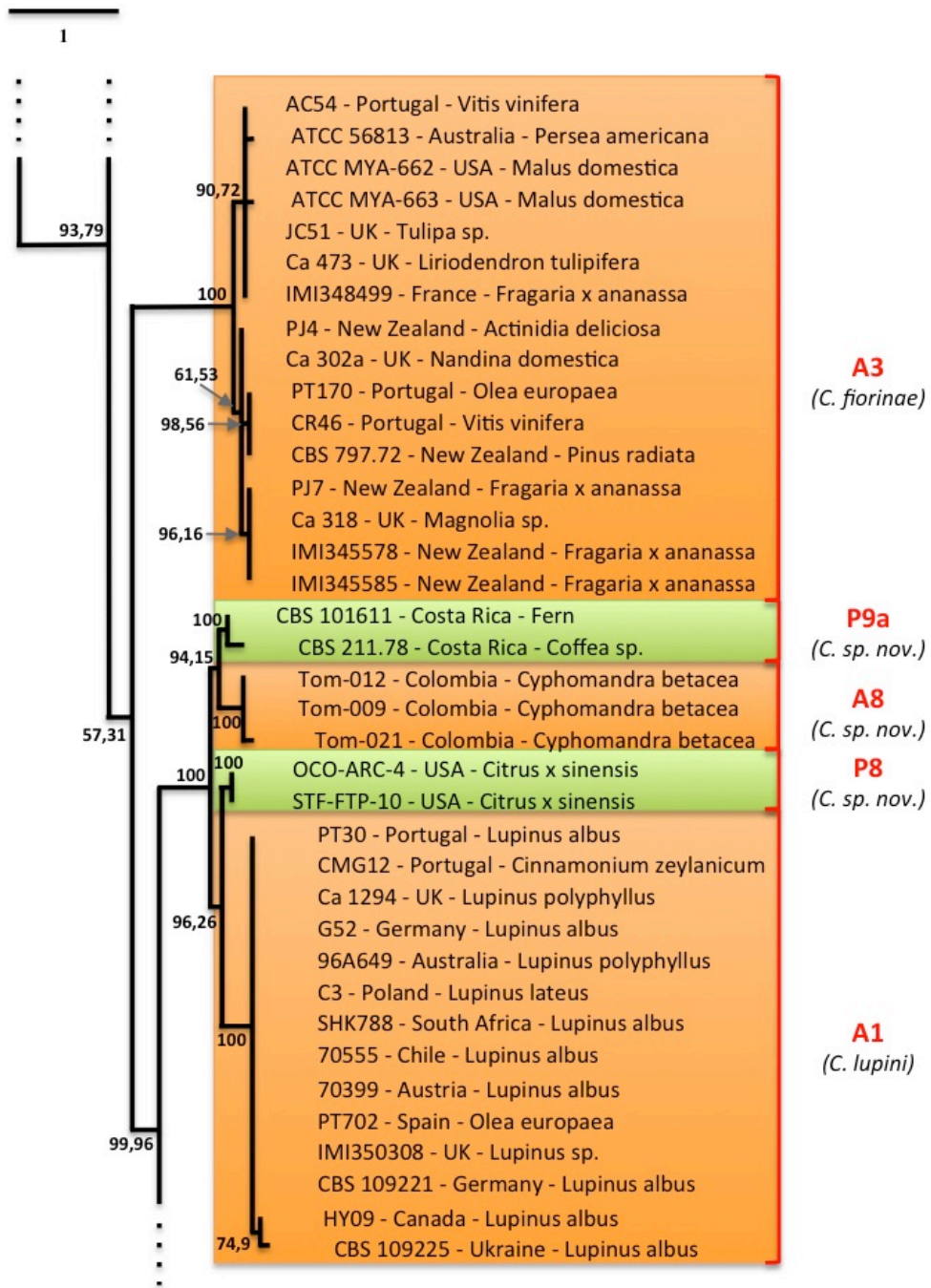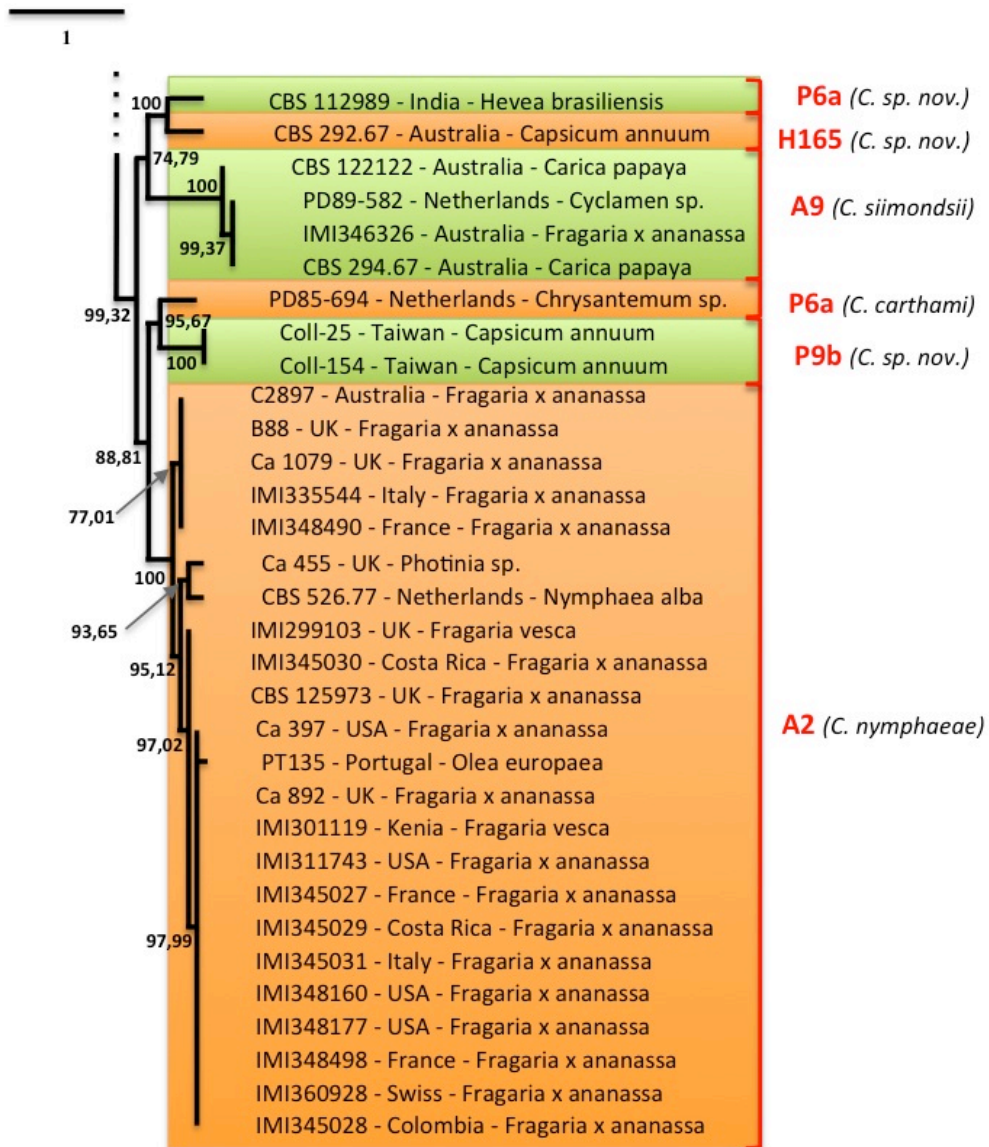
Figure 4.3. Bayesian MCMC analysis tree constructed from an alignment based on the concatenation of rRNA, TUB, MAT1-2 and GPDH partial sequences of 120 *Colletotrichum* isolates used in this study. Numbers on nodes are bootstrap values. Tree was rooted with data for respective loci from the *C. graminicola* and *C. higginsianum* genomes and sequences of *C. gloeosporioides* and *C. spinaciae* were obtained experimentally. Potential species designations reported in literature/GenBank are shown on the right.

Several distinct groups were identified within *C. acutatum* based on genetic variability. Using loci with higher resolution (compared to ITS database analysis in Chapter 3), the subsets of isolates show a higher complexity than the nine genetic groups A1 to A9 described previously (Sreenivasaprasad and Talhinhas, 2005; Whitelaw *et al.*, 2007). For example, confirmation of two different sub groups of A7 (A7a and A7b); those strains showing identical ITS sequence and different growth rate clusters. Furthermore, genetic groups showed particular trends associated with mating behavior, geographic distribution and host association patterns.

All isolates able to produce perithecia belonged to the same genetic group A7a. On the other hand, all self-sterile isolates capable of heterothallic mating (cross-fertile) (based on Gueber and Correll 2001) were spread in two different genetic groups A3 and A5. Although cross-fertilization has been demonstrated and described *in vitro*, no biological or genetic evidence that this process occurs in nature has been reported. This strong relationship between mating behavior and phylogenetic clustering, and the lack of evidence of mating crosses in nature, fit with the hypothesis of an on-going speciation process even in those cases where the strains retain the capability to exchange share genetic material through sexuality.

Regarding host association patterns, the first observation is that, by increasing the extent of genetic information (different loci) isolates from the same hosts tended to cluster together better compared to a single locus such as ITS. Using the set of data based on four loci (in this kind of study mainly ITS, and recently a second locus, have been used for genetic characterization), isolates associated to the same host such as *Phormium spp.*, *Chyphomandra betacea*, *Lupinus spp.* and

Citrus x sinensis could be differentiated into genetics groups. However, the level of host-specificity of these groups needs further analysis through cross infection assays on different hosts. Also in this case, the process that leads to host-specificity could be seen as an on-going process and as a key factor closely related to the evolution of the populations.

According to published work on taxonomic re-assessment of *Colletotrichum acutatum sensu lato,* results obtained confirm the presence of cryptic species, or in a wider evolutionary view, a clear speciation process. For example isolates belonging to A7b correspond to what has been described as *C. phormii* (Farr *et al.*, 2006). Mainly based on morphological characterization, Farr *et al.*, (2006) showed how *C. phormii* constitutes a distinct and host-specific species within the *C. acutatum* lineage. In the present study, isolates described previously as A6 (and confirmed by ITS analysis) appear to descend from *C. phormii*. Data collected showed a high host-specificity of those strains; in fact, most of the strains have been isolated from *Rhododendron* species. Unfortunately, these strains are no longer available from the University of Uppsala where the original studies were carried out and no further investigations have been done. A7a group instead has a longer history in literature. Strains belonging to this group have been firstly described as *Physalospora miyabeana* (Fukushi, Ann. phytopath. Soc. Japan 1: 100 1921) and then associated to *Glomerella miyabeana* by Arx (1957). In parallel, another organism was taxonomically evolving: firstly described as *Sphaeria salicis* (Auersw. ex Fuckel in 1870), then renamed *Physalospora salicis* (Auersw. ex Fuckel. in 1882); *Anisostomula salicis* (Auersw. ex Fuckel) Petr.in 1925 and finally as *Glomerella salicis* (Auersw. ex Fuckel) L. Holm in 2000 (http://www.speciesfungorum.org). All these strains

belong to the same genetic group within the *C. acutatum* lineage and are associated with infections on willow (*Salix* spp.) black canker. Strains belonging to A4 group have been historically described with different names such as *Colletotrichum godetiae* (by Neerg. in 1943) due to the capability to infect *Godetia hybrid*. This group has been described as *Colletotrichum clavatum*: a new species of *C. acutatum sensu lato* (Faedda *et al.*, 2001). Description has been based on morphological and genetic characterization of strains isolated from infected tissue of Olives (*Olea europea*) in Italy. Several other groups have already been described as new species within the *C. acutatum* lineage: *Colletotrichum lupini* for isolates belonging to group A1 and specific for *Lupinus* spp., *C. fioriniae* comb. et stat. nov. for isolates belonging to A3 group, *C. simmondsii* sp. nov. for isolates belonging to A2. Recently published work proposed a new combination, *Colletotrichum carthami*, for the anthracnose pathogen of the asteraceous plants (*Carthamus tinctorius, Chrysanthemum coronarium* var. spatiosum, and *Calendula officinalis*), characterizing the species, which differs from the closely related *Colletotrichum simmondsii*, based on pathogenicity and molecular characteristics (Uematsu *et al.*, 2012).

Despite the fact that recent research has focused on the reassessment of *C. acutatum* sensu lato taxonomy, *Colletotrichum nymphaeae* (Pass. Aa 1978) and as *C. simmondsii* sp. nov (Shivas and Tan 2009) have been both described as overlapping with the A2 genetic group. Results obtained using a set of higher resolution loci than ITS and TUB, have suggested that those strains belonged to two different genetic groups. In the present work *Colletotrichum nymphaeae* is associated to isolates belonging to A2 and *C. simmondsii* for isolates belonging to A9. The main reason for this is the position in the phylogram of isolate CBS

122122 (BRIP 28519) used as epitype by Shivas and Tan 2009; this isolate does not cluster with the majority of A2 isolates forming a separate cluster.

Results showed the presence of at least 15 genetic groups within *C. acutatum sensu lato* and 9 of these correspond to species that are beginning to be recognised as cryptic species and are being assigned new names in the literature:

- A1 genetic group: *Colletotrichum lupini* (Nirenberg *et al.*, 2002)

- A2 genetic group: *Colletotrichum nymphaeae* (Pass. Aa 1978)

- A3 genetic group: *Colletotrichum fioriniae* (Shivas and Tan, 2009)

- A4 genetic group: *Colletotrichum godetiae* (Neerg, 1943) syn. *C. clavatum* (Faedda *et al.*, 2011)

- A5 genetic group: *Colletotrichum acutatum sensu strictu* teleomorph *Glomerella acutata* (Guerber and Correll, 2001)

- A7b genetic group: *Colletotrichum phormii* (Farr and Rossman, 2006)

- A7a genetic group: *Colletotrichum salicis* teleomorph *Glomerella salicis* (Holm, 2000); syn. *Colletotrichum miyabeana* teleomorph *Glomerella miyabeana* (Arx, 1957)

- A9 genetic group: *Colletotrichum simmondsii* (Shivas and Tan, 2009)

- P6 genetic group: *Colletotrichum carthami* (Uematsu *et al.*, 2012)

4.4 Discussion

Based on previous results (Chapter 3) a collection of 120 isolates representative of the global diversity of *Colletotrichum acutatum sensu lato* populations has been assembled. The set of isolates has been choosen based on host association, geographic distribution, phylogenetic relationships and biological diversity from ITS sequence analysis (CaITSdb).

Morphological analysis carried out on the sub-set of isolates representative of the genetic groups showed variations in appearance in culture and growth rate at different temperatures. These characters are often difficult to describe reliably, and can change following sub culturing or based on the length and type of storage. The present study highlighted the difficulties in using morphology to distinguish sub-groups within *Colletotrichum acutatum sensu lato* and the need for other methods to understand their relationships.

Multi-locus genetic characterization distinguished genetic groups some of which correspond to various species previously recognised based on host association patterns and classical analysis such as *C. phormii*, *G. miyabeana* and *C. carthami*. Using loci with higher resolution (such as TUB, GAPDH and MAT), host association patterns of other groups also showed particular trends. Isolates belonging to "ex A8" group seem to be specific to *Cyphomandra betacea*. Mainly all *G. miyabeana* strains have isolated from tree-hosts such as *Acer*, *Salix* and *Populus*.

The present study has also confirmed that all isolates belonging to group A1 have been associated with *Lupinus* anthracnose, and all the strains isolated from this host belong to A1; evidence supporting the nomenclature of *Colletotrichum*

lupini for the strains belonging to this taxa (Nirenberg *et al.*, 2002). It is the same situation for strains that belong to group A6; mainly all of them have been isolated from *Rhododendron* spp. (Vinnere *et al.*, 2002). All isolates characterized as *C. phormii* have been associated with *Phormium* anthracnose in Oceania; the genus *Phormium* (belonging to monocotyledons) comprises native species of New Zealand. Present results also showed a comparable distance between isolates belonging to group A7 and A5 to a hypothetical ancestral population. Evolutionary analysis carried out on ITS sequences confirmed the hypothesis that Oceania could be the geographic origin of this organism. However the data emerging from the present study highlighted the complexities in fully understanding the evolutionary relationships in *C. acutatum sensu lato*.

The general evolutionary trend can be seen as a tendency of the pathogen to exhibit variations in their host range (some populations with a specific host range adapting to wider range of hosts and populations with a wide range of hosts adapting to specific hosts). Furthermore, comparing host association patterns of distinct genetics groups suggests a fundamental role of heterothallism in host adaptation. Considering isolates A5 and A3, which are evolutionarily close to each other; and described as cross fertile, the occurrence of heterothallism (or switch in mating behavior) seems to clash with an increase of host range. Also isolates capable of self-reproduction seems to have a narrow range of hosts. On the other end, populations evolutionarily less close to A3 and A5 that seems to have lost any kind of sexual reproduction evolved host preference/specificity (A8, A1, etc.). In conclusion, the capacity of populations to share genetic information could be responsible for an increase in the host spectrum of this pathogen. The phylogenetic analysis did not show the occurrence of genetic

exchange in recent time and this observation is confirmed by the absence of heterothallic sexual structures in nature.

This data combined with some recent research focused on the re-assessment of *Colletotrichum* taxonomy also suggest the presence of *combinatio nova* species. However, speciation processes driving the evolution of this organism are still not easy to clarify; results obtained so far do not fully support the hypothesis that host-interaction as the sole driver e. The fact that isolates belonging to different and evolutionarily close genetic groups (or sub species) could infect the same host, and are present in same geographic region (for example isolates from A3 and A5 both are present in Portugal on Olives) but don't show evidence of genetic exchange reflect a sympatric speciation. In other cases, allopatric speciation is the best model to describe this process. For example, strains genetically very close (not identical) but clustering apart have been isolated from different hosts in Costa Rica. Strains belonging to A8 have been isolated from the same host in the same country and in this case, the speciation process could be described (based on geographic isolation) with an allopatric or parapatric or peripatric model.

CHAPTER 5

CHARACTERIZATION OF MORPHOLOGICAL, GENETIC AND PATHOGENIC VARIABILITY AMONG ISOLATES OF *COLLETOTRICHUM ACUTATUM SENSU LATO* FROM STRAWBERRY IN THE UK

5.1.   INTRODUCTION

Strawberry anthracnose in the UK and worldwide

5.1.1.   Strawberry: an important crop worldwide

Fragaria × *ananassa* (common name: strawberry), is a hybrid species belonging to Rosaceae family, that is cultivated worldwide for its fruit, the. FAO statistics service shows that, the world production of strawberry has increased from 1961 to 2010 at 259.10% and the cultivated area has increased by 578.73%. The total world production and cultivation are estimated in 2010 as about 4366662 tons and 243907 hectares, respectively. The three biggest producers were USA, Turkey and Spain. According to import data, the UK is 14th in the list of the most important countries. Since the 1990s the British strawberry industry has expanded rapidly representing a significant component of UK fruit production (Bech and Simpson, 1989). In the last decade home production in the UK grew up from 35054 tonnes in 1961 to 102900 tonnes in 2010 (FAO STAT) (Figure 5.1).
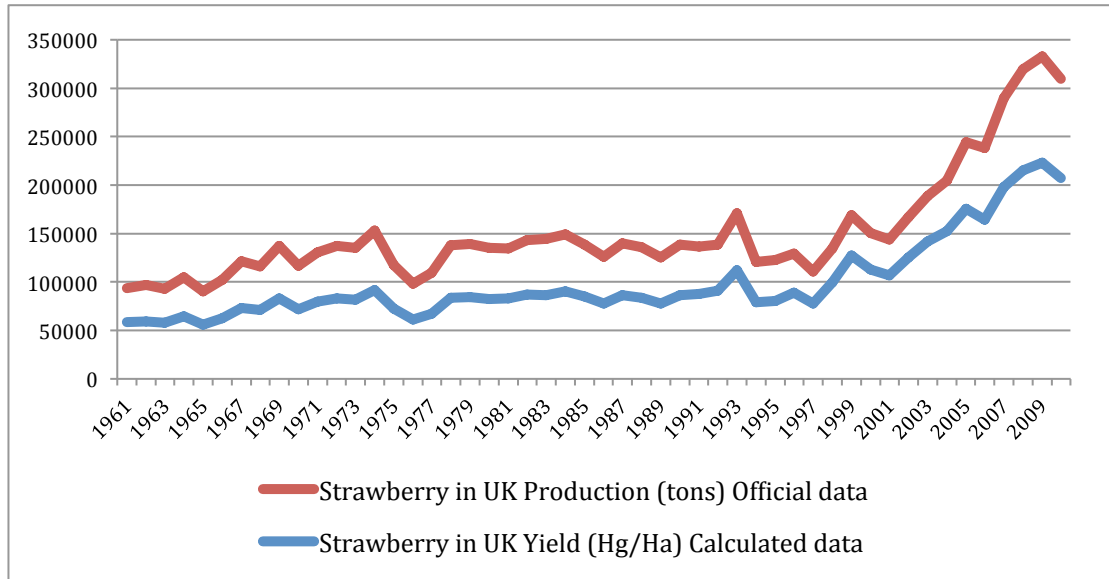
Figure 5.1. Graph showing Strawberry production in UK since 1961 (data obtained from http://faostat.fao.org/)

Diseases are a key factor limiting strawberry production, both the crop and fruit field worldwide, and diseases are often difficult to control. A lack of detailed understanding of the pathogen populations and their dynamics and consequentially, problems in accurate and early diagnosis, make it difficult to implement appropriate measures to control those diseases.

5.1.2. Anthracnose: a major strawberry disease

A number of diseases caused by fungi affect different tissues in strawberry such as fruit, leaves, roots and crowns. Several fungal species have been associated with strawberry diseases; most important are: grey mold caused by *Botrytis cinerea*, anthracnose caused by *Colletotrichum* spp., verticillium wilt caused by *Verticillium dahliae*, powdery mildew caused by *Sphaerotheca macularis*,

strawberry red core caused by *Phytophthora fragariae*, leaf scorch caused by *Marssonina fragariae*, etc.

Colletotrichum causes extensive losses in strawberry fruit production (Sreenivasaprasad & Talhinhas, 2005). Two species complex, *C. acutatum* and *C. gloeosporioides*, have been reported as causal agents of strawberry anthracnose, which is a major disease of the cultivated strawberry. Originally *C. gloeosporioides* was considered a separate species from *C. fragaria*; von Arx (1957) assigned *C. fragariae* as synonymous of *C. gloeosporioides*, but researchers have generally retained the use of the name *C. fragariae* when the pathogen was associated with strawberries disease (Howard and Albregts, 1984; Mass and Howard, 1985; Sutton, 1992). In this thesis strawberry antrachnose pathogens are considered as two species complex including the *C. acutatum* species complex (equivalent to *C. acutatum sensu lato*).

C. acutatum in particular has been considered economically the second most important pathogen after *Botrytis cinerea* worldwide (Calleja *et al.*, 2012). In Europe, in fact, *C. acutatum* is the most prevalent species causing anthracnose, whereas *C. gloeosporioides* is found only occasionally (Buddie *et al.*, 1999; Hemelrijck *et al.*, 2010).

Strawberry anthracnose symptoms produced by the two species complexes of *Colletotrichum* are similar and all can be found on all parts of the plant (Buddie *et al.*, 1999). Flower blight and fruit rot are common symptoms in the field (Howard *et al.*, 1992) while lesions on stolons, petioles, and leaves are mainly damaging in plant nurseries (Freeman *et al.*, 2001). Crown symptomatology is characterized by brown-reddish necrotic areas (Urena-Padilla *et al.*, 2002), and in

some cases stunting and chlorosis has been associated with root necrosis (Freeman *et al.*, 2001). Classic *Colletotrichum* symptoms are illustrated in Figure 5.2.
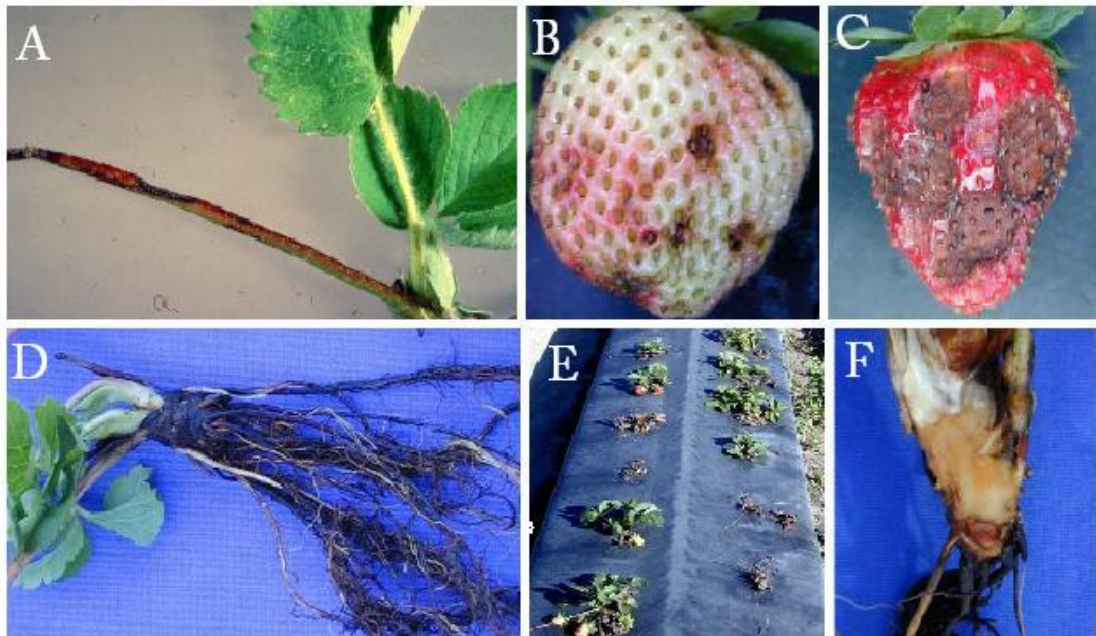
Figure 5.2. Symptoms caused by *C. acutatum* on strawberry: **A**, lesions on stolons; **B**, fruit lesions on unripe fruit; **C**, lesions on ripe fruit; **D**, Root necrosis symptoms; **E**, Stunted plants due to root necrosis; **F**, Basal crown rot. (Sources: A -http://www.forestryimages.org/browse/detail.cfm?imgnum=1263080; B and C - http://strawberry.ifas.ufl.edu/plantpathfiles/PP-col-26full.htm; D, E and F. Source: Mertely *et al.*, 2005.

Disease cycle of *C. acutatum* on strawberry is shown in Figure 5.3. In this case, the pathogen is primarily a necrotroph on strawberry tissues; the biotrophic phase is usually very short in infected leaves, petioles, and runners. Furthermore the pathogen may not show symptoms for some time.

Transplants are frequently produced far from commercial fields (Freeman *et al.*, 1998); thus, the fungus must survive as asymptomatic or quiescent infections on transplants that will serve as inoculum in the commercial fields. During warm,

wet periods, lesions formed on propagation material produce secondary conidiation that can occur on the surface of vegetative tissues, and this can serve to augment inoculum levels to infect flowers and fruit. Those few conidia that are formed on the leaves without causing symptoms are dispersed by water splash and harvesting operations from the leaves to flowers and fruit. Primary inoculum for dispersal within the field is mostly dependent on the formation of spores in acervuli on petioles and fruit tissues (Peres *et al.*, 2005). *C. acutatum* may overwinter as vegetative mycelium on different parts of the host (Wharton & Dieguez-Uribeondo, 2004). Inoculum of *Colletotrichum* is not known to survive in buried plant remains between seasons (Urena-Padilla *et al.*, 2001). However, due to its wide host range the pathogen is able to overwinter colonizing other crops (such as tomato, pepper and aubergine, etc.), weeds and native species and could survive for few months (Freeman *et al.*, 2001). This lifestyle of the pathogen, spreading from strawberry plants to other species may lead to reinfestation of new strawberry fields the following year (Peres *et al.*, 2005). This pathogen may also survive on symptomless leaves through the presence of secondary conidia and appressoria (Sreenivasaprasad & Talhinhas, 2005).
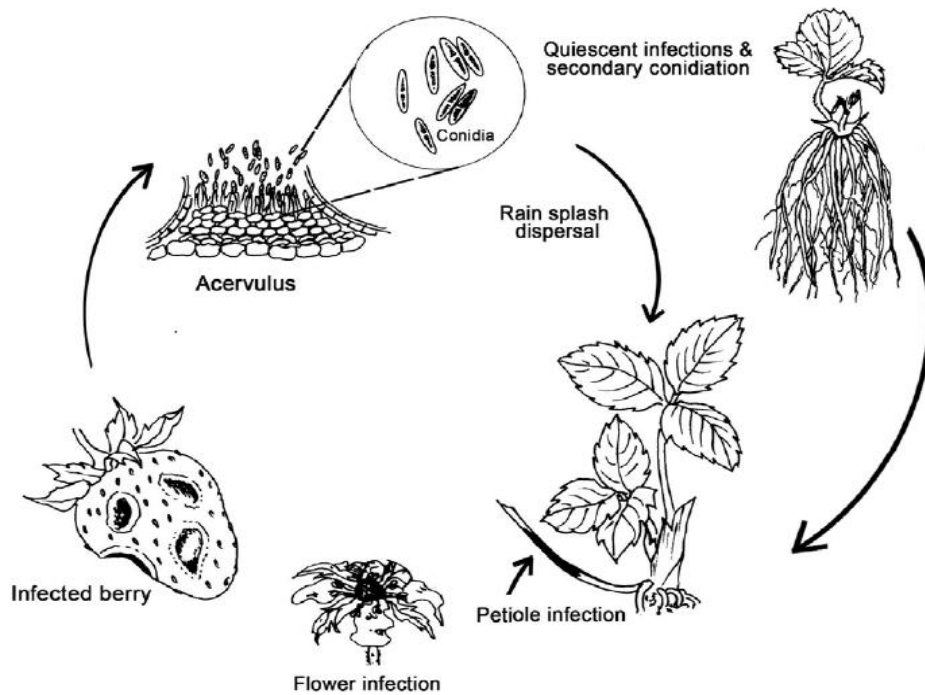
Figure 5.3. Disease cycle of anthracnose fruit rot of strawberries caused by *Colletotrichum acutatum.* Source: Peres *et al.*, 2005.

Research has been carried out to characterize *Colletotrichum acutatum* populations from strawberry in specific geographic areas such as Israel, France, Bulgaria, Spain, Belgium and other countries in Europe (Freeman and Katan, 1997; Martinez-Culebras *et al.*, 2000; Denoyes-Rothan *et al.*, 2003, Jelev *et al.*, 2008; Garrido *et al.*, 2008; Van Hemelrijck *et al.*, 2010), and also from some regions of the USA (Urena-Padilla *et al.*, 2002). Other researches have tried to characterize *C. acutatum* from strawberry using isolates collected worldwide (Garrido *et al.*, 2008). All these projects have been carried out using genomic finger printing (such as RFLP, apPCR, etc.) or using sequence analysis based on the ITS region. Results have shown the presence of one representative "clonal" population usually suggesting that they originate from a single genetic source

and consequentially that the disease is spread through infected propagation material.

5.1.3. *Colletotrichum acutatum* in the UK

Colletotrichum acutatum was identified for the first time as strawberry pathogen in California in 1983 (Smith & Black, 1986). Since this first identification, *C. acutatum* has spread worldwide including the UK through runners and propagating material of strawberry (Sreenivasaprasad *et al.*, 1996; Freeman *et al.*, 2001; Denoyes-Rothan *et al.*, 2003; Peres *et al.*, 2005; Sreenivasaprasad & Talhinhas, 2005).

The exact date of the first record of *C. acutatum* in the UK is still uncertain; the main reason is the lack of knowledge related to the taxonomy of this taxa. Evidence collected with the CaITSdb (see chapter 2.4.1.) showed the presence of at least one isolate belonging to *C. acutatum* species complex in Great Britain in 1938 associated with *Phormium* spp. (common name "New Zealand flax"). In fact, in the very early 19th century the quality of rope materials made from New Zealand flax was already widely known internationally and used for spars and masts, and the Royal Navy (branch of the British Armed Forces) was one of the largest customers; this evidence may suggest the entry of the pathogen through this host. However, because no other detections have been reported for a few decades it is probable that the pathogen did not spread in the country. On the other hand, published data refer to the first record of *C. acutatum* 50 years later in 1978 isolated from *Anemone* sp. (Jones & Baker, 2007; Calleja *et al.*, 2012). In 1983, the first instance of disease in strawberries caused by *C. acutatum* was

recorded in the UK; this occurrence was attributed to the importation of infected strawberry runners from California that same year (Simpson *et al.*, 1994).

Lovelidge in 1993 supposed that the continued introduction of infected strawberry material from abroad was so common that the disease was destined to become endemic in the UK (Sreenivasaprasad *et al.*, 1996). In fact, in the next few years, further outbreaks have been reported caused by the importation of infected propagation material from mainland Europe and *C. acutatum* started to become a serious problem causing important losses.

5.2.   OBJECTIVES AND APPROACH

The main target of this chapter is to investigate the complexity of host specificity of *C. acutatum sensu lato* related to UK strawberry production and use this information to further understanding the evolution of this pathogen and its population. To achieve this the focus is on:

o collection of strains related to this pathogen associated to strawberry infections in the UK held by the authorities responsible for Plant Health during the last three decades (plus isolates from other geographic regions and hosts)

o characterization of populations of *C. acutatum* species complex populations responsible for strawberry anthracnose in United Kingdom based on cultural studies and molecular techniques (multi-locus phylogeny)

o investigation of differences in aggressiveness and the relationships with various hosts and genetic groups

In addition to its economic impact, the choice of this model has been useful based on the results obtained in the evolutionary analyses. In fact strawberry plants seem to be susceptible to most genetic groups identified.

5.3. MATERIALS AND METHODS

5.3.1. Pathogenicity assays on strawberry fruits

Representative *Colletotrichum* isolates from each group identified in the genetic characterization of *C. acutatum* from strawberry in the UK were used for pathogenicity tests, together with reference isolates of populations associated with other hosts. Conidial suspension was prepared by flooding 10 days old PDA cultures with sterile deionized water, scratching with a spatula and filtering the suspension with one layer of cheesecloth. Spore concentration was determined using a haemocytometer and diluted to 10^5 spores/ml. cultivar Elsanta were chosen to carry out pathogenicity test due to it susceptibility to *Colletotrichum* infection. Strawberry fruit were chosen at a specific phenological stage: fruit turning white-pink (Denoyes-Rothan *et al.*, 1999) and were surface sterilized for 5 min in a solution of NaClO (1% active chlorine) in 50% EtOH and washed three times in sterilized water and blotted dry. Fruits so prepared were placed in a tray with moist sand on the bottom (Figure 5.4) to avoid any movement of the fruits during further procedures

Figure 5.4. Moist chamber set up to carry out pathogenicity assays on strawberry fruit. Sets of two trays were considered as a single replication.

After inoculation, the fruits were incubated at $25 \pm 1°C$ in 12h photoperiod. The capability of single isolates to produce symptoms on strawberry fruit was evaluated by inoculating fruits with a 5μl drop of a conidial suspension. This was done with three replications per isolate; each replication consisted of four fruits.

Lesion development was evaluated 7 days after inoculation (d.a.i.) and scored as: 0, no symptoms; 1, symptoms less than 33% of fruit surface; 2, symptoms surface 33-66%; and 3, when symptoms were more than 66% of the surface. Two fruits with sterile distilled water (SDW) as well as two fruit untouched for each replicate served as negative controls. Four replicates of three fruits for each isolate have been evaluated. Each block consists in randomized fruits each of them inoculated with spores suspension derived from different cultures and the replicates have been set up in different times.

5.3.2. Pathogenicity assays on strawberry plants

Cryopreserved plants cv. Elsanta were used to evaluate the ability of the isolates to produce symptoms characteristic of *Colletotrichum* crown rot. Prior to inoculation strawberry plants replicates were grown in 25 cm pots in an unheated/ventilated polyethylene greenhouse during July in Northern Italy. After three months, plants fully developed changing from vegetative to reproductive growth (Figure 5.5).

Figure 5.5. Strawberry plants cv. Elsanta after three months from revival and prior to inoculation tests

Aggressiveness of *Colletotrichum* species isolates was evaluated by injecting the crowns of the greenhouse-grown strawberry plants with 0.2 ml conidial suspension (10^6 spores/ml) using a syringe (Garrido *et al.*, 2008). This was done with three replications per isolate; each replication consisted of three plants. Each block consists in three randomized plants each of them inoculated with spores suspension derived from different cultures and the replicates were carried out in three different area of the same greenhouse. At 24 d.a.i., plants were evaluated for wilting or collapse of the plant: symptoms characteristic of *Colletotrichum* crown rot. Lesion development was evaluated and scored as: 0, no symptoms; 1, symptoms on crown but no collapse; 2, symptoms of wilting or collapse of part of the plant; and 3, dead of the pant. In addition, the crowns of all plants were sectioned and examined for the presence of red-brownish lesions. Two plant crowns injected with SDW as well as two untouched plants served as negative controls. Crown infection was confirmed by re-isolation of the pathogen.

5.4.   RESULTS

5.4.1. Isolate collection

A set of 67 *C. acutatum sensu lato* strains from strawberry production in the United Kingdom used in this study has been obtained mainly from FERA (Food and Environment Research Agency part of the Department for Environment, Food and Rural Affairs; authorities mainly responsible for Plant Health in the UK). These strains were isolated from infected material and collected by the Ministry of Agriculture, Fisheries and Food (MAFF) in the 1980s and later by the Central Science Laboratory (CSL) in York (now Fera) until 2002 as standard protocol for quarantine pathogens. Organism's identity was confirmed by the agency through morphological characterization. 27 *C. acutatum* strains from strawberry worldwide (from: Australia, Colombia, Costa Rica, France, Italy, Kenya, Netherlands, New Zealand, Norway, Portugal, Spain, Swiss, USA) and 9 strains of the pathogen isolated in UK from other hosts (*Malus domestica*, *Vitis* sp., *Photinia* sp., *Tulipa* sp., *Nandina domestica*, *Liriodendron tulipifera*, *Magnolia* sp., *Phormium* sp., *Prunus avium*) have been included in the set of isolates. In order to assist genetic comparative analysis, in addition to the isolates mentioned above, a set of 37 isolates representative of the pathogen global population diversity (based on results obtained in Chapter 3) has been included in the present work. Complete list of isolates used in this study is reported in Table 5.1.

Table 5.1. Reference isolates of Colletotrichum used in this study .

Original code	Genus	compex	Country of Origin	Area	Year	Host
			Strains isolated from strawberry in UK			
B88	*Colletotrichum*	*acutatum*	UK	N/A	1989	*Fragaria x ananassa*
NI90	*Colletotrichum*	*acutatum*	UK	Northern Ireland	N/A	*Fragaria x ananassa*
Ca 1079	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 2546*	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 899*	*Colletotrichum*	*acutatum*	UK	Cambridge	N/A	*Fragaria x ananassa*
Ca 310	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 915*	*Colletotrichum*	*acutatum*	UK	Nottinghamshire	N/A	*Fragaria x ananassa*
Ca 886	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 919*	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 916*	*Colletotrichum*	*acutatum*	UK	Suffolk	N/A	*Fragaria x ananassa*
Ca 918*	*Colletotrichum*	*acutatum*	UK	Gloucestershire	N/A	*Fragaria x ananassa*
Ca 917	*Colletotrichum*	*acutatum*	UK	Norfolk	N/A	*Fragaria x ananassa*
Ca 223	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 224	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 225	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 255	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 256*	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 258	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 456	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria vesca*
Ca 493	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 494	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria vesca*
Ca 604	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 607	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 608	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 872	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*

Strains isolated from strawberry in UK

Original code	Genus	compex	Country of Origin	Area	Year	Host
Ca 903	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1001	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1258	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1259*	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1260	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1261	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1262*	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1305	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Fragaria x ananassa*
Ca 1376	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1377	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1378	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1379	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1380	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1381	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1382	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1383	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1384	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1385	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1386	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1387	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1388	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1389	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1390	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1391	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1392	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1393	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1394	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1395	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*
Ca 1396	*Colletotrichum*	*acutatum*	UK	Isle of White	N/A	*Fragaria x ananassa*

Original code	Genus	compex	Country of Origin	Area	Year	Host
Strains isolated from strawberry in UK						
Ca 1397	Colletotrichum	acutatum	UK	Isle of White	N/A	Fragaria x ananassa
Ca 1398	Colletotrichum	acutatum	UK	Isle of White	N/A	Fragaria x ananassa
Ca 1429	Colletotrichum	acutatum	UK	N/A	N/A	Fragaria x ananassa
Ca 1441	Colletotrichum	acutatum	UK	Isle of White	N/A	Fragaria x ananassa
Ca 1442	Colletotrichum	acutatum	UK	Isle of White	N/A	Fragaria x ananassa
Ca 1443	Colletotrichum	acutatum	UK	Isle of White	N/A	Fragaria x ananassa
Ca 1444	Colletotrichum	acutatum	UK	Isle of White	N/A	Fragaria x ananassa
Ca 1449	Colletotrichum	acutatum	UK	N/A	N/A	Fragaria x ananassa
Ca 2064	Colletotrichum	acutatum	UK	N/A	N/A	Fragaria x ananassa
Ca 1002	Colletotrichum	acutatum	UK	N/A	N/A	Fragaria x ananassa
Ca 892	Colletotrichum	acutatum	UK	Nottinghamshire	N/A	Fragaria x ananassa
IMI299103	Colletotrichum	acutatum	UK	Kent	1985	Fragaria vesca
CBS 125973	Colletotrichum	acutatum	UK	Harpenden Herts	N/A	Fragaria x ananassa
Strains isolated from strawberry worldwide						
C2897	Colletotrichum	acutatum	Australia	Weat Australia	1996	Fragaria x ananassa
Ca 397	Colletotrichum	acutatum	USA	N/A	N/A	Fragaria x ananassa
Ca 1053	Colletotrichum	acutatum	Netherlands	N/A	N/A	Fragaria x ananassa
Ca 891	Colletotrichum	acutatum	Portugal	N/A	N/A	Fragaria sp.
Ca 511	Colletotrichum	acutatum	France	N/A	N/A	Fragaria x ananassa
Ca 729	Colletotrichum	acutatum	Swiss	N/A	N/A	Fragaria x ananassa
Ca 1430	Colletotrichum	acutatum	Norway	N/A	N/A	Fragaria vesca
Ca 1432	Colletotrichum	acutatum	Norway	Østfold county	1999	Fragaria x ananassa
PJ7	Colletotrichum	acutatum	New Zealand	Auckland	1988	Fragaria x ananassa
IMI301119	Colletotrichum	acutatum	Kenia	Nairobi	1986	Fragaria vesca
IMI311743	Colletotrichum	acutatum	USA	Ohio	1986	Fragaria x ananassa
IMI335544	Colletotrichum	acutatum	Italy	Trento	1989	Fragaria x ananassa
IMI345026	Colletotrichum	acutatum	Spain	N/A	1991	Fragaria x ananassa
IMI345027	Colletotrichum	acutatum	France	N/A	1991	Fragaria x ananassa
IMI345028	Colletotrichum	acutatum	Colombia	N/A	1991	Fragaria x ananassa

Original code	Genus	compex	Country of Origin	Area	Year	Host
Strains isolated from strawberry worldwide						
IMI345029	*Colletotrichum*	*acutatum*	Costa Rica	N/A	1991	*Fragaria x ananassa*
IMI345030	*Colletotrichum*	*acutatum*	Costa Rica	N/A	1991	*Fragaria x ananassa*
IMI345031	*Colletotrichum*	*acutatum*	Italy	N/A	1991	*Fragaria x ananassa*
IMI345578	*Colletotrichum*	*acutatum*	New Zealand	Auckland, Papakura	1988	*Fragaria ananassa*
IMI346326	*Colletotrichum*	*acutatum*	Australia	Brisbane	1967	*Fragaria x ananassa*
Ca 1044	*Colletotrichum*	*acutatum*	New Zealand	N/A	N/A	*Fragaria x ananassa*
IMI348160	*Colletotrichum*	*acutatum*	USA	Louisiana	1985	*Fragaria x ananassa*
IMI348177	*Colletotrichum*	*acutatum*	USA	Missisipi	1981	*Fragaria x ananassa*
IMI348490	*Colletotrichum*	*acutatum*	France	N/A	1984	*Fragaria x ananassa*
IMI348498	*Colletotrichum*	*acutatum*	France	N/A	N/A	*Fragaria x ananassa*
IMI348499	*Colletotrichum*	*acutatum*	France	N/A	1988	*Fragaria x ananassa*
IMI360928	*Colletotrichum*	*acutatum*	Swiss	Zurich	1993	*Fragaria x ananassa*
Strains isolated from different hosts in UK						
Apple3	*Colletotrichum*	*acutatum*	UK	N/A	2008	*Malus domestica*
Apple4	*Colletotrichum*	*acutatum*	UK	N/A	2008	*Malus domestica*
Ca 1294	*Colletotrichum*	*acutatum*	UK	York	1998	*Lupinus polyphyllus*
Ca 287	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Statice sp.*
RBXXX	*Colletotrichum*	*acutatum*	UK	N/A	2010	*Vitis spp.*
Ca 455*	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Photinia sp.*
JC51	*Colletotrichum*	*acutatum*	UK	N/A	2003	*Tulipa sp.*
Ca 302a	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Nandina domestica*
Ca 473	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Liriodendron tulipifera*
Ca 318	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Magnolia sp.*
IMI350308	*Colletotrichum*	*acutatum*	UK	Kent	1991	*Lupinus spp.*
CBS 198.35*	*Colletotrichum*	*acutatum*	UK	N/A	N/A	*Phormium spp.*
CBS 126527	*Colletotrichum*	*acutatum*	UK	Harpenden Herts	N/A	*Prunus avium*
Strains isolated from different host worldwide and used as references for genetics groups						
PT250	*Colletotrichum*	*acutatum*	Portugal	Mirandela	2003	*Olea europaea*
PT135*	*Colletotrichum*	*acutatum*	Portugal	Elvas	2001	*Olea europaea*

Original code	Genus	compex	Country of Origin	Area	Year	Host
			Strains isolated from different host worldwide and used as references for genetics groups			
PD85-694	Colletotrichum	acutatum	Netherlands	N/A	N/A	Chrysanthemum sp.
PD89-582*	Colletotrichum	acutatum	Netherlands	N/A	N/A	Cyclamen sp.
PT227*	Colletotrichum	acutatum	Portugal	Faro	2003	Olea europaea
Tom-21	Colletotrichum	acutatum	Colombia	Antioquia	1998	Cyphomandra betacea
Tom-12	Colletotrichum	acutatum	Colombia	Antioquia	1998	Cyphomandra betacea
CBS 193.32	Colletotrichum	acutatum	Italia	N/A	1932	Olea europaea
PT30*	Colletotrichum	acutatum	Portugal	Azores	1999	Lupinus albus
CR46	Colletotrichum	acutatum	Portugal	LPVVA Lisbon	2000	Vitis vinifera
9178*	Colletotrichum	acutatum	Norway	Vest-Agder county	2004	Vaccinium corymbosum
MP3*	Colletotrichum	acutatum	USA	Massachusetts, Boston	2006	Acer Platanoides
PJ8*	Colletotrichum	acutatum	New Zealand	Nelson	1987	Pyrus pyrifolia
ATCC MYA-663*	Colletotrichum	acutatum	USA	Virginia	N/A	Malus domestica
HY09*	Colletotrichum	acutatum	Canada	N/A	N/A	Lupinus albus
JL198	Colletotrichum	acutatum	Serbia	N/A	2003	Olea europaea
AR3787	Colletotrichum	acutatum	South Africa	N/A	N/A	Phormium spp.
CBS 607.94	Colletotrichum	acutatum	Netherlands	Z. Flevoland	1994	Salix spp.
ALM-NRB-30K	Colletotrichum	acutatum	Isreal	Southern Israel	N/A	Prunus dulcis
CBS 101611	Colletotrichum	acutatum	Costa Rica	N/A	N/A	Fern
CBS 109225*	Colletotrichum	acutatum	Ukraine	N/A	N/A	Lupinus albus
CBS 112980*	Colletotrichum	acutatum	South Africa	Southern Cape	N/A	Pinus radiata
CBS 112989	Colletotrichum	acutatum	India	N/A	N/A	Hevea brasiliensis
CBS 122122*	Colletotrichum	acutatum	Australia	Queensland, Yandina	N/A	Carica papaya
CBS 211.78	Colletotrichum	acutatum	Costa Rica	Turrialba	N/A	Coffea spp.
CBS 292.67	Colletotrichum	acutatum	Australia	Queensland; Brisbane	N/A	Capsicum annuum
CBS 294.67*	Colletotrichum	acutatum	Australia	Queensland; Brisbane	N/A	Carica papaya
ATCC 38896	Colletotrichum	acutatum	Netherlands	Kortenhoefse Plassen	N/A	Nymphaea alba
CBS 797.72*	Colletotrichum	acutatum	New Zealand	N/A	N/A	Pinus radiata
OCO-ARC-4*	Colletotrichum	acutatum	USA	Florida, Arcadia	N/A	Citrus x sinensis
STF-FTP-10	Colletotrichum	acutatum	USA	Florida, Frostproof	N/A	Citrus x sinensis

Original code	Genus	compex	Country of Origin	Area	Year	Host
Coll-24	Colletotrichum	acutatum	Taiwan	Tainan	1998	Capsicum annum
Coll-54	Colletotrichum	acutatum	Taiwan	Tainan	1992	Capsicum annum
Colletotrichum strains used as out groups						
Cg 311*	Colletotrichum	gloeosporioides	USA	N/A	N/A	Fragaria x ananassa
Cg 386*	Colletotrichum	gloeosporioides	USA	N/A	N/A	Fragaria x ananassa
Cg 780*	Colletotrichum	gloeosporioides	UK	N/A	N/A	Fragaria x ananassa
Cg 869*	Colletotrichum	gloeosporioides	UK	N/A	N/A	Fragaria x ananassa

Isolates marked with asterisk (*) were used in pathogenicity tests.

5.4.2. Morphological characterization

Isolates collected from FERA and from strawberry infections in the UK could be divided into three groups based on colony aspect (Figure 5.6). Considering the high variability of this organism, this evidence leads the hypothesis that potentially three different populations were responsible for strawberry anthracnose in the United Kingdom.
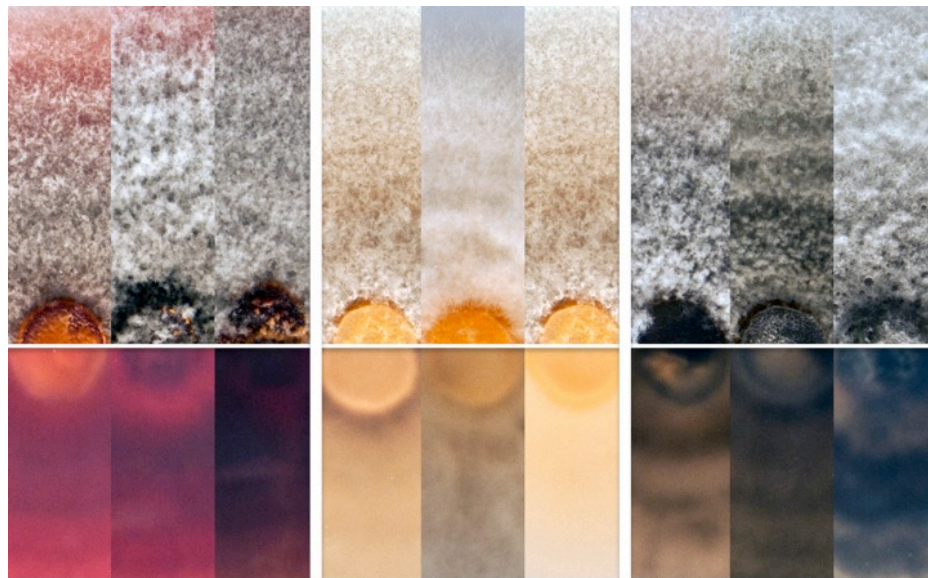
Figure 5.6. Colony aspect of the three groups of strains identified in the subset of isolates associated to strawberry anthracnose in the UK. Three examples for each group based on itra-group variability. In the top of the figure are pictures of the upper side of the culture and on the bottom the reverse side.

The three groups were characterized by different colony aspect. Most of them developed white cottony aerial mycelium, light orange conidial masses on the inoculum surface and presence of dark melanised structures similar to acervuli; colour of the colony was from dark grey to dark brown. Around ten isolates were characterized by white aerial mycelium and from yellow to white pigmentation in reverse. Just 6 isolates were dark red in the reverse with orange conidial masses in big drops on the colony surface.

5.4.3. Genetic characterization and variability

Based on the four loci sequences and results obtained previously (see Chapter 4), the 67 *Colletotrichum* isolates from strawberry production system in the United Kingdom, belonged to three distinct genetic groups linked to new species: **A3**, *Colletotrichum fioriniae* comb. et stat. nov. (Marcelino and Gouli 2008; Shivas and Tan 2009); **A4**, *Colletotrichum godetiae* (Neerg 1950) and as *Colletotrichum clavatum* sp. nov. (Faedda *et al*., 2011); **A2**, *Colletotrichum nymphaeae* (Pass. Aa 1978). Incidence of these genetic groups/species in the fields is reported on the chart below (Figure 5.7)
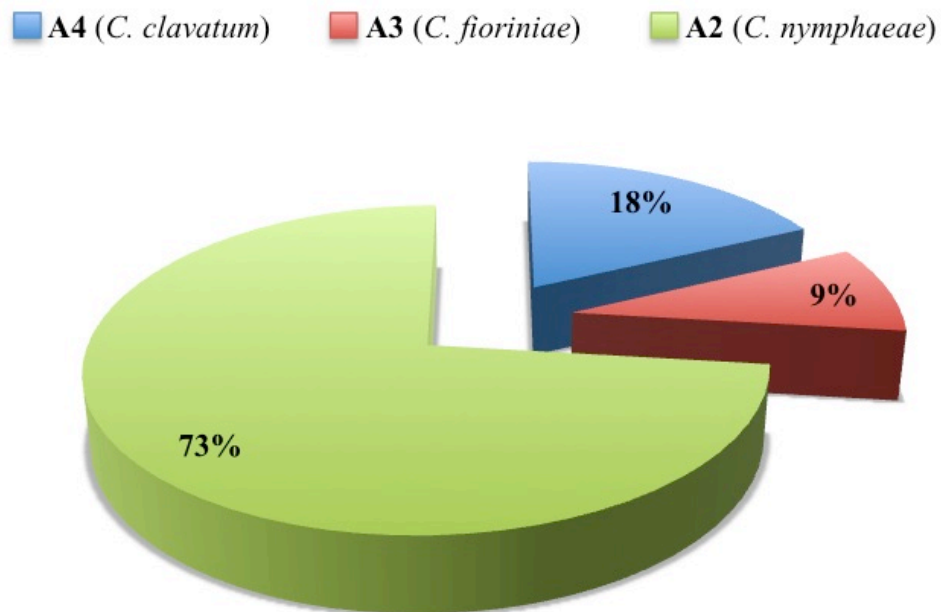
Figure 5.7. Chart showing the percentage incidence of *Colletotrichum acutatum sensu lato* species identified among 67 strains isolated from strawberry in UK.
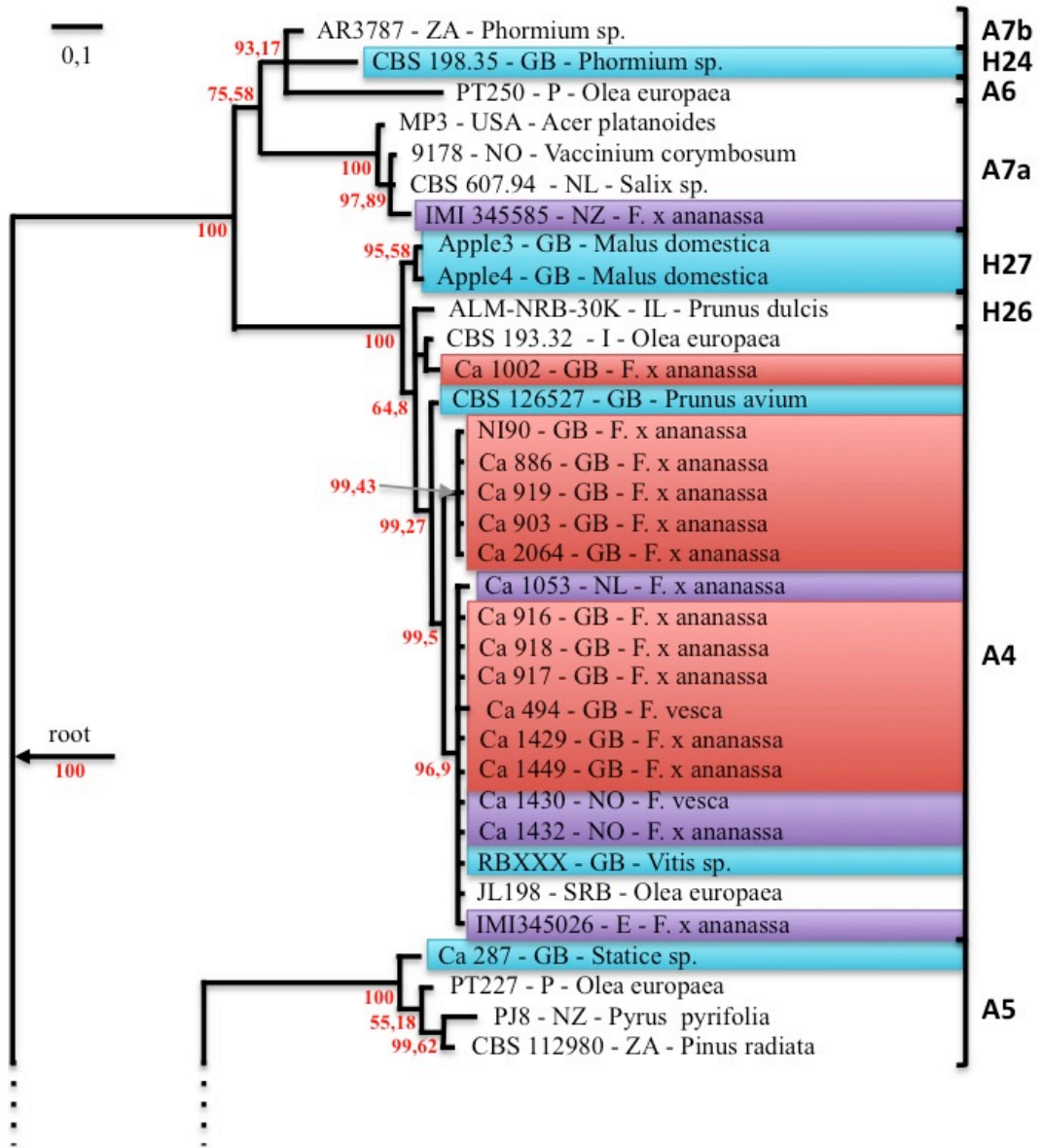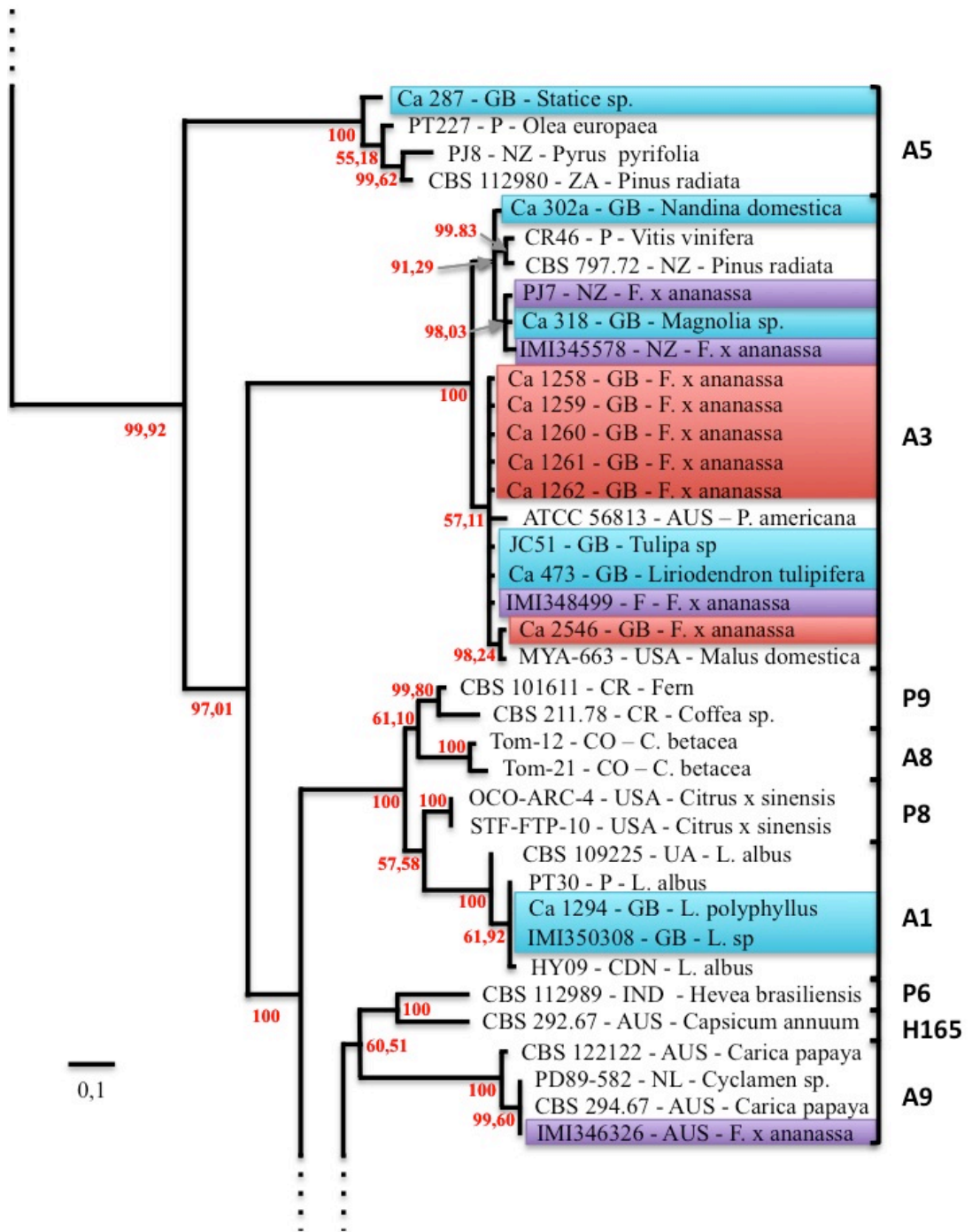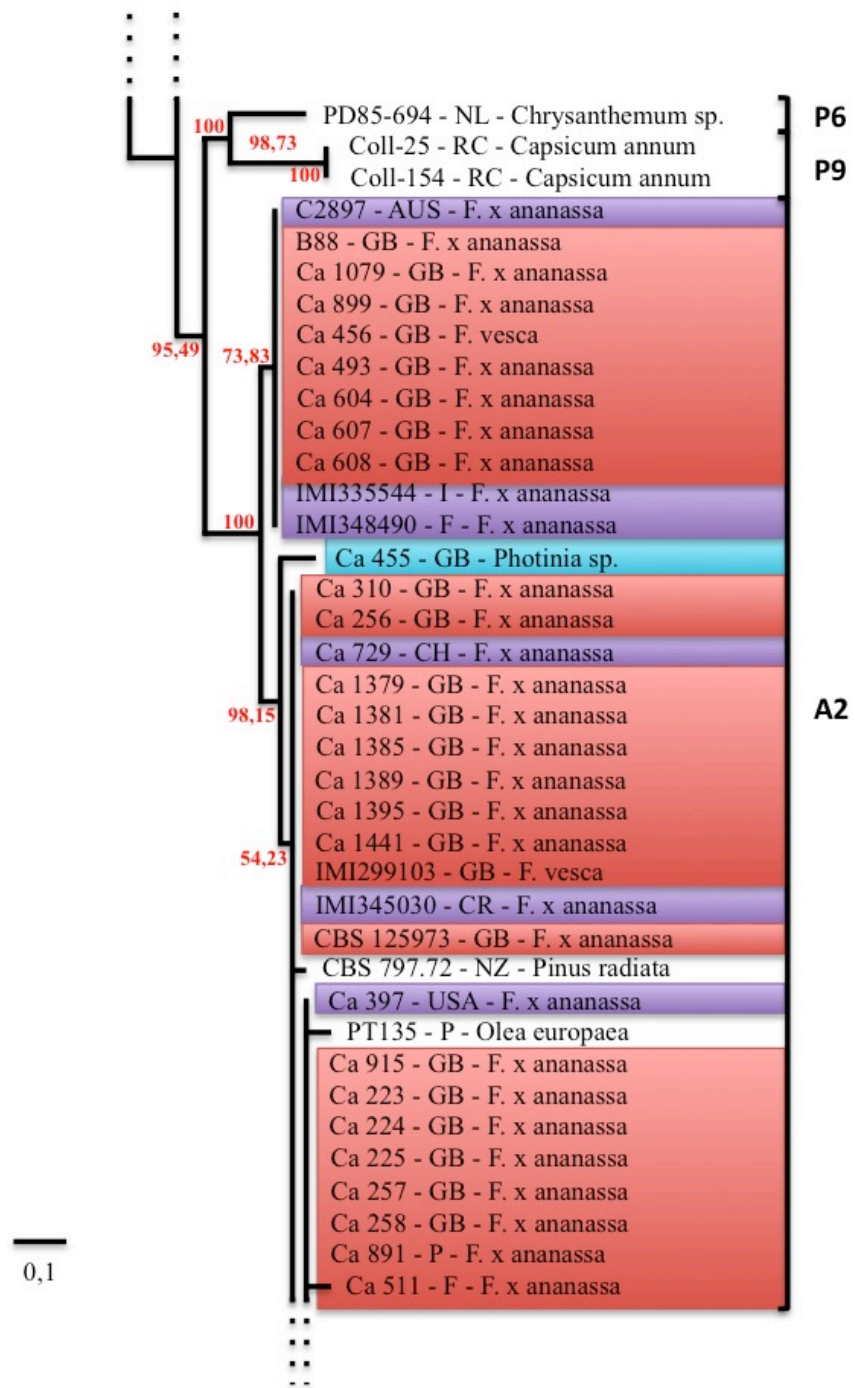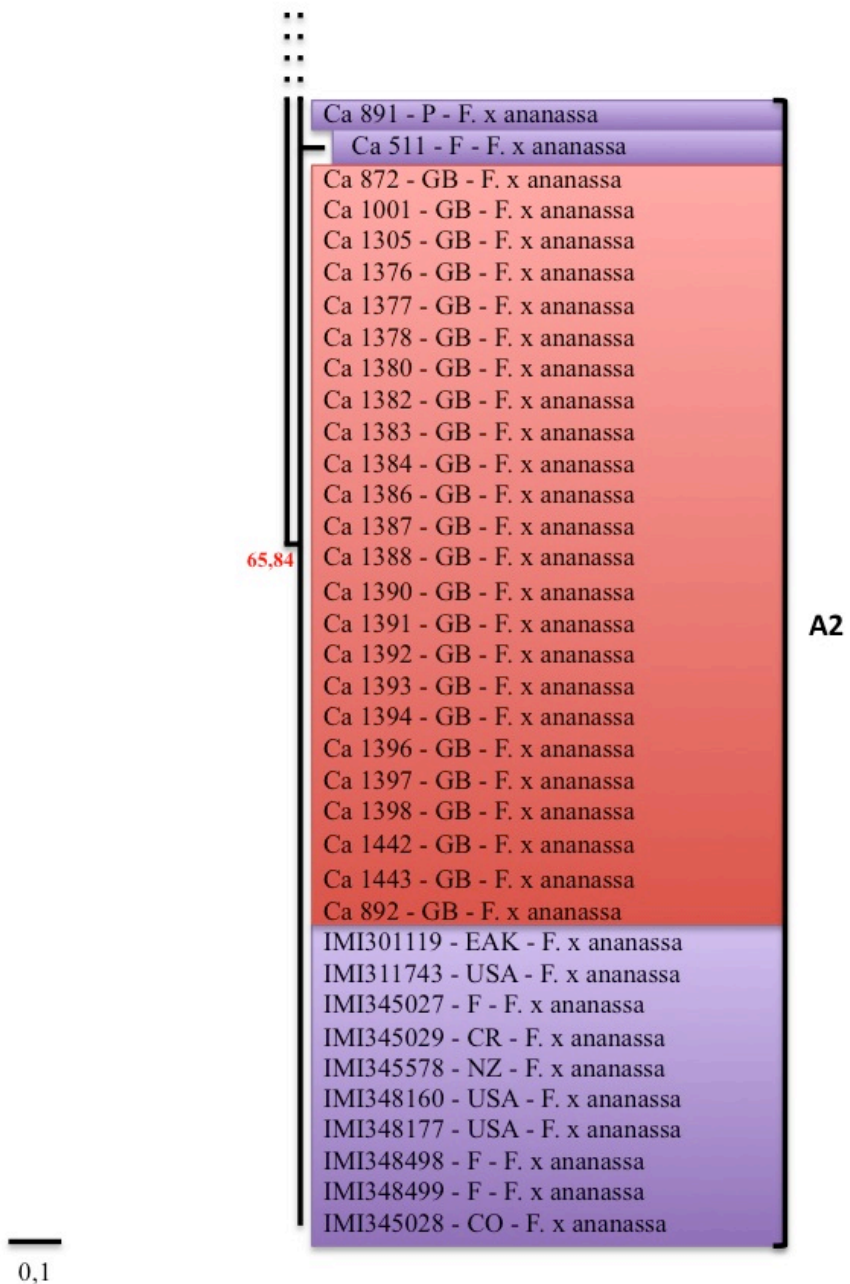
Figure 5.8. Bayesian MCMC analysis tree constructed from the alignment based on the concatenation of rRNA, TUB, MAT1-2 and GPDH partial sequences alignment of 140 *Colletotrichum acutatum sensu lato* isolates used in this study. Numbers on nodes are bootstrap values. Tree was rooted with data from *C. graminicola* and *C. higginsianum* genomes and sequences of four *C. gloeosporioides* obtained experimentally. Genetic groups designation is reported on the right.
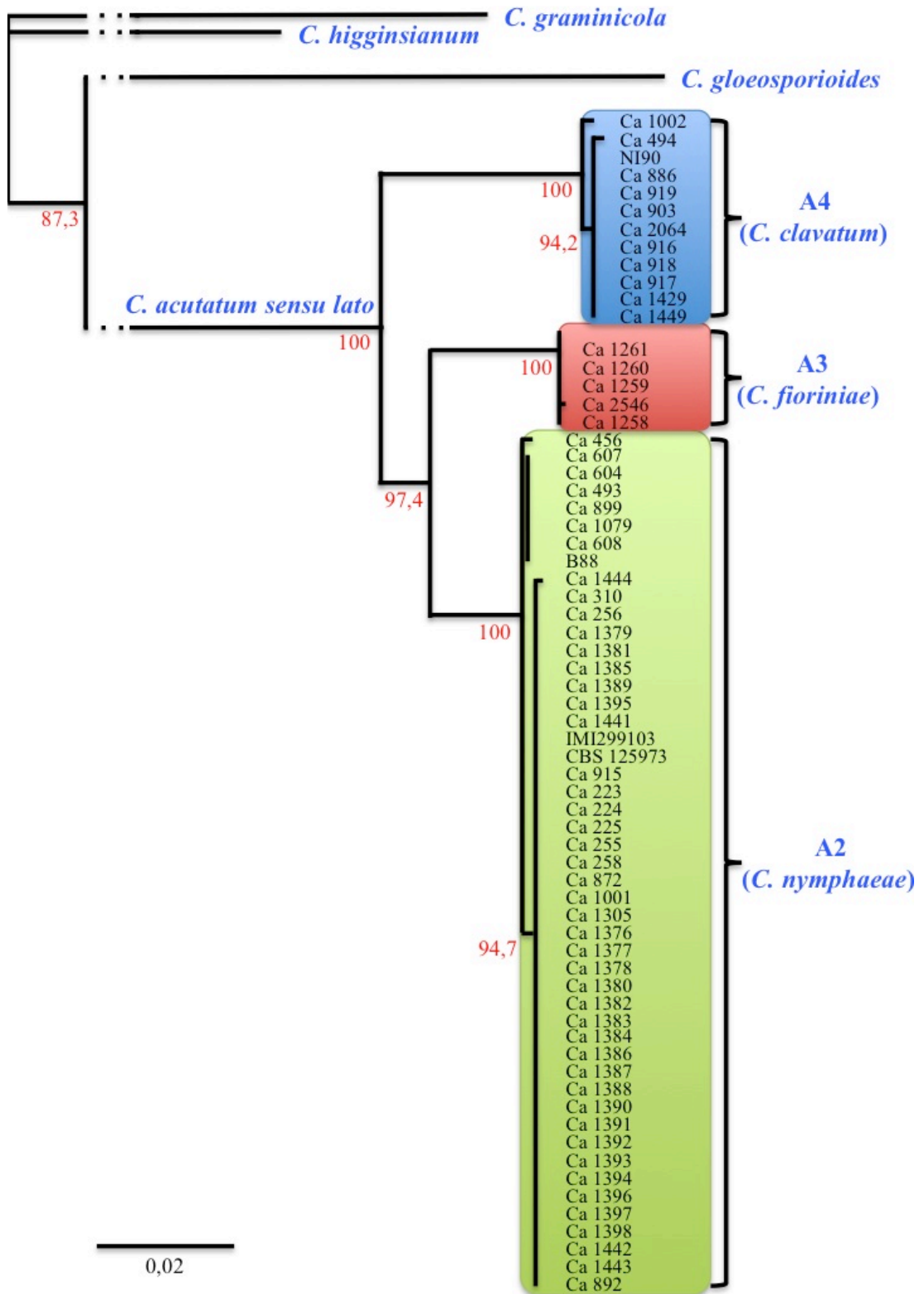
Figure 5.9. Bayesian MCMC analysis tree constructed from the alignment based on the concatenation of rRNA, TUB, MAT1-2 and GPDH partial sequences alignment of 67 *Colletotrichum acutatum sensu lato* isolates from strawberry in UK. Numbers on nodes are bootstrap values. Tree was rooted with data from *C. graminicola* and *C. higginsianum* and *C. gloeosporioides*. Groups and specie designation is reported on the right.

Based on the genetic characterization, the *Colletotrichum* isolates could be subdivided into three groups (Figure 5.8 and 5.9) corresponding to three different species. The majority of the *C. acutatum* isolates (49/67) belonged to A2 genetic group (or *C. nymphaeae*) as clustering with high bootstrap value with isolates used as references for this taxa such as CBS 797.72, PT135, IMI345028, etc. and genetically close (Identical Sites = 1422/1438 [98.9%]; Pairwise % Identity = 99.9%).

A smaller proportion of strains (12/67) belonged to A4 genetic group (corresponding to *C. godetiae* or *C. clavatum*) as genetically close to reference strains such as ALM-NRB-30K, CBS 193.32, JL198, etc. (Identical Sites = 1411/1438 [94.6%]; Pairwise % Identity = 99.4%) forming one unique solid cluster. Only 6 strains belonged to *C. fioriniae* (Figure 5.7) as genetically very close to the reference isolate ATCC 56813 (Identical Sites = 1.436 /1443 [99.5%]; Pairwise % Identity = 99.9%).

The genetic relationship of 67 *Colletotrichum* species isolates linked to the strawberry production systems in the UK suggest a multiple introduction of the pathogen. Data collected also showed the presence of further genetic groups in the UK such as A5 (*C. acutuatum sensu stricto*), A7a (*C. salicis;* syn. *C. miyabeana*); isolates belonging to these species have been isolated from strawberry infected tissues in other countries of Europe or worldwide. Further experiments on variability in aggressiveness of *Colletotrichum* spp. on strawberry plants could explain the spread of these specific groups and the absence of other groups present in various countries.

5.4.4   Pathogenic variability

36 *C. acutatum sensu lato* and four *C. gloeosporioides* isolates were chosen for pathogenicity tests (Table 5.2). Three representatives strains from each of the three

populations associated to strawberry antrachnose in the UK were chosen to test differences in aggressiveness. The set of isolates included one or more strains representatives of all major cryptic species belonging to *C. acutatum sensu lato*. The four *C. gloeosporioides* strains used as reference were isolated from strawberry infected tissues from the UK and USA.

Crown rot occurrence on strawberry has been calculated on the percentage of plants showing symptoms (drop off, or crown necrosis) 24 d.a.i. Aggressiveness factor has been calculated scoring as: 0, no symptoms; 1, symptoms on crown but no collapse; 2, symptoms of wilting or collapse of part of the plant; and 3, dead of the pant; calculating the average for each isolate on 6 plants for three replicates. To evaluate the presence of necrosis each crown has been sectioned (Figure 5.10) and results recorded.

Figure 5.10. Picture showing strawberry crown section with typical *Colletotrichum* symptoms. Infected crown were cut open lengthwise, the inside surfaces was reddish-brown and firm, or shown reddish-brown streaks.

Table 5.2. Pathogenicity on strawberry fruit and crowns of representative *Colletotrichum acutatum sensu lato* isolates and *C. gloeosporioides* isolates as out group

Isolate	Fruit assays						Plant assays					
	Symptoms incidence			Aggressiveness factor			Symptoms incidence			Aggressiveness factor		
	Value	group mean	SE	Value	group mean	SE	Value	group mean	SE	Value	group mean	SE
C. gloeosporioides												
Cg 311	91.67%	75.00%	7.61%	2.50	2.10	0.24	90.00%	42.50%	16.52%	1.4	0.7	0.27
Cg 386	58.33%			1.50			33.33%			0.2		
Cg 780	83.33%			2.50			33.33%			0.5		
Cg 869	66.67%			1.92			13.33%			0.6		
A5 (*C. acutatum sensu stricto*)												
PT227	50.00%	52.78%	12.11%	1.42	1.44	0.36	26.67%	30.00%	12.62%	0.8	0.6	0.17
PJ8	75.00%			2.08			53.33%			0.7		
CBS 112980	33.33%			0.83			10.00%			0.3		
A3 (*C. fioriniae*) from strawberry in UK												
Ca 2546	100.00%	86.11%	10.02%	2.67	2.44	0.26	43.33%	53.33%	10.00%	0.8	0.8	0.12
Ca 1259	91.67%			2.75			43.33%			0.7		
Ca 1262	66.67%			1.92			73.33%			1.1		
A3 (*C. fioriniae*) from other hosts												
ATCC MYA-663	75.00%	63.89%	11.11%	2.00	1.75	0.34	50.00%	28.89%	11.60%	0.8	0.5	0.12
CBS 797.72	41.67%			1.08			26.67%			0.5		
CR46	75.00%			2.17			10.00%			0.3		
A4 (*C. clavatum*) from strawberry in UK												
Ca 919	83.33%	77.78%	2.78%	2.08	2.06	0.07	46.67%	31.11%	7.78%	0.6	0.5	0.08
Ca 916	75.00%			1.92			23.33%			0.4		
Ca 918	75.00%			2.17			23.33%			0.4		

Isolate	Fruit assays						Plant assays					
	Symptoms incidence			Aggressiveness factor			Symptoms incidence			Aggressiveness factor		
	Value	group mean	SE	Value	group mean	SE	Value	group mean	SE	Value	group mean	SE
A4 (*C. clavatum*) from other hosts												
CBS 193.32	25.00%	25.00%	0.00%	0.75	0.58	0.17	3.33%	5.56%	4.01%	0.3	0.2	0.10
JL198	25.00%			0.75			13.33%			0.3		
ALM-NRB-30K	25.00%			0.25			0.00%			0.0		
A1 (*C. lupini*)												
PT30	16.67%	16.67%	4.81%	0.50	0.39	0.15	0.00%	0.00%	0.00	0.0	0.0	0.00
HY09	8.33%			0.08			0.00%			0.0		
CBS 109225	25.00%			0.58			0.00%			0.0		
A2 (*C. nymphaeae*) from strawberry in UK												
Ca 899	100.00%	86.11%	7.35%	3.00	2.53	0.26	56.67%	55.56%	4.84%	0.9	0.6	0.12
Ca 915	75.00%			2.08			46.67%			0.6		
Ca 256	83.33%			2.50			63.33%			0.4		
A2 (*C. nymphaeae*) from other hosts												
PT135	83.33%	50.00%	17.35%	1.67	1.14	0.29	23.33%	28.89%	2.94%	0.9	0.5	0.18
Ca 455	41.67%			1.08			30.00%			0.5		
CBS 526.77	25.00%			0.67			33.33%			0.3		
A7b (*C. phormii*)												
PT250	33.33%	27.78%	2.78%	0.67	0.63	0.02	3.33%	1.11%	1.11%	0.4	0.1	0.14
AR3787	25.00%			0.58			0.00%			0.0		
CBS 198.35	25.00%			0.63			0.00%			0.0		

Isolate	Fruit assays						Plant assays					
	Symptoms incidence			Aggressiveness factor			Symptoms incidence			Aggressiveness factor		
	Value	group mean	SE	Value	group mean	SE	Value	group mean	SE	Value	group mean	SE
A7a (*C. salicis*)												
9178	25.00%	19.44%	5.56%	0.50	0.44	0.15	10.00%	4.44%	2.94%	0.3	0.2	0.10
MP3	8.33%			0.17			3.33%			0.3		
CBS 607.94	25.00%			0.67			0.00%			0.0		
A9 (*C. simmondsii*)												
PD89-582	66.67%	44.44%	14.70%	1.83	1.08	0.46	26.67%	24.44%	5.88%	0.4	0.5	0.10
CBS 122122	16.67%			0.25			13.33%			0.3		
CBS 294.67	50.00%			1.17			33.33%			0.6		
P9												
CBS 101611	16.67%	16.67%	-	0.33	0.33	-	0.00%	0.00%	-	0.0	0	-
P8												
OCO-ARC-4	25.00%	25.00%	-	0.67	0.67	-	0.00%	0.00%	-	0.0	0	-

Colletotrichum isolates information in Table 5.1

When inoculated on strawberry fruits, after three d.a.i., all isolates produced symptoms characteristic of anthracnose fruit rot (Figure 5.11).

Different kinds of lesion could be distinguished on the strawberry fruits. From brown lesion containing orange drops of conidia to lesion entirely covered with aerial mycelium. All inoculated *Colletotrichum* isolates were re-isolated from infected fruits and transferred onto PDA plate to confirm colony aspect and to check the presence of any other microorganisms.
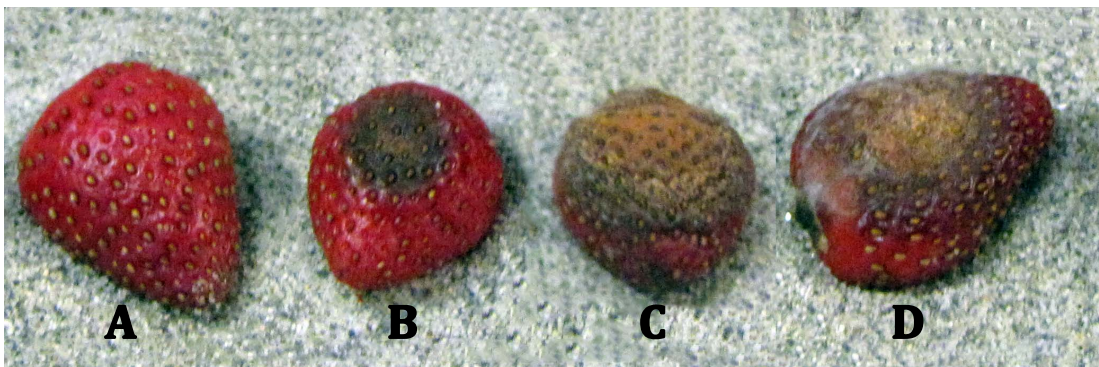
Figure 5.11. Different kinds of lesions characteristic of *C. acutatum sensu lato* on Strawberry fruits. **A** - no symptoms; **B** - brown lesion with absence of aerial mycelium; **C** - brown lesion containing orange drops of conidia and aerial mycelium on the edge of the lesion; **D** - lesion entirely covered with aerial mycelium

Isolates of the *C. acutatum sensu lato* showed a strong relationship between genetic groups (cryptic species) and aggressiveness on strawberry tissues. Within the set of strains tested there is also a strong relationship between aggressiveness on fruits and on crowns. Strains isolated from strawberry are more aggressive compared to strains from the same taxa but isolated from other hosts. A2 (*C. nymphaeae*) and A3 (*C. fioriniae*) isolates appeared to be the most aggressive, followed by A5 (*C. acutatum sensu stricto*) and A4 (*C. godetiae*; syn *C. clavatum*) strains isolated from

strawberry. Also A9 isolates (*C. simmondsii*) are capable of infecting strawberry plants. All the other groups/species: A1 (*C. lupini*), A7b (*C. phormii*), A7a (*C. salicis*), P8 and P9 even if capable of infecting strawberry fruits are much less aggressive compared to the others: in fact average of fruit infection incidence is lower than 25% and average aggressiveness score is lower than 0.7. Those species have an incidence on crown infection lower than 5% with aggressiveness score lower than 0.2. *C. gloeosporioides* isolates were as aggressive as the most aggressive *C. acutatum sensu lato* on strawberry plants and fruits.

5.5 <u>Discussion</u>

Since the 1990s the British strawberry industry has expanded rapidly representing a significant component of the UK fruit production (Beech and Simpson, 1989) and *Colletotrichum* causes extensive losses in this sector (Sreenivasaprasad & Talhinhas, 2005). This part of the project reports for the first time a detailed biomolecular characterization of *Colletotrichum acutatum sensu lato* strains from strawberry in the UK.

The morphological characteristics, based on colony aspect, of the *Colletotrichum* strains collected were a useful tool to subdivide the set of strains into three different populations and each of them were characterized by a specific colony colour.

Colletotrichum acutatum sensu lato was the main focus of this research because work carried out in France, Israel, UK, Bulgaria and Spain reported this taxa as the major problem; other species such as *C. gloeosporioides* occurred only sporadically in France and the UK, respectively (Freeman & Katan, 1997; Buddie *et al.*, 1999; Denoyes-Rothan *et al.*, 2003; Jelev *et al.*, 2008; Garrido *et al.*, 2008). The *C. acutatum sensu lato* strains were assigned the genetic groups indentified in Chapter 3 and 4 and, when possible, the taxonomy reported in literature, based on four loci sequence analysis: ITS, TUB, MAT1-2 and GAPDH. *Colletotrichum* A2 isolates (*C. nymphaeae*) is the most common taxa with A4 (*C. clavatum or C. godetiae*) also often reported on strawberry in Europe and America (Sreenivasaprasad & Talhinhas, 2005). These two groups were also the most representative in our dataset of strains from strawberry in UK. A3 strains (*C. fioriniae*) have a worldwide distribution, and are quite common on strawberry but only just a few isolates were detected and this group was not commonly present in

the fields in the UK. A9 (*C. simmondsii*), A5 (*C. acutatum sensu stricto*) and A7a (*C. salicis* and *C. miyabeana)* are common on strawberry in the Oceania and have only been found sporadically in Europe. No isolates belonging to these taxa have been detected in the UK from the present work. The variability observed within the UK *C. acutatum sensu lato* species fit in part with previous reports of *C. acutatum* on strawberry within specific geographic regions. For example, in France, Israel, Bulgaria and Spain, the majority of *C. acutatum* isolates clustered in the same molecular group (*C. nymphaeae*, A2) and almost no intraspecies diversity was observed within each country (Freeman & Katan, 1997; Denoyes-Rothan *et al.*, 2003; Jelev *et al.*, 2008; Garrido *et al.*, 2008). A different situation has been observed on Belgian isolates, where the population represented: 33% strains belonging to A2 (*C. nymphaeae*), 5% to A3 (*C. fioriniae*), 50% to A4 (*C. godetiae* or *C. clavatum*), 3% to *C. acutatum sensu strictu* and 6% to *C. salicis.* A possible explanation to *C. acutatum sensu lato* status in the UK might be due to recent introduction from a limited number of sources (early 80s). Despite this evidence, the reason for these differences in occurrences of these populations remains unclear.

Pathogenicity tests revealed a strong correlation between *C. acutatum sensu lato* species and aggresivness. Isolates representative of the UK populations (A2, *C. nymphaeae*; A3, *C. fioriniae*; A4, *C. godetiae* or *C. clavatum*) with A5 (*C. acutatum sensu stricto*) appeared to be the most aggressive. Also isolates belonging to A9 (*C. simmondsii*) were capable of infecting strawberry plants and fruits severely. Strains isolated from strawberry were more aggressive compared to strains from the same taxa but isolated from other hosts clearly indicating a degree of host-preference. Isolates belonging to other groups were much less agressive on strawberry fruit compared to the others and almost non-pathogenic on plants.

CHAPTER 6

THE FIRST *COLLETOTRICHUM*

ACUTATUM SENSU LATO GENOME

SEQUENCE

6.1 <u>INTRODUCTION</u>

6.1.1   Fungal genomes: an overview

As pathogens of crop plants, fungi have a massive impact on human welfare, remaining major problems in all areas of the world either by destroying valuable crops as devastating pathogens or producers of mycotoxins. Better understanding of the molecular basis of plant-microbes interactions could be essential for the progress of disease control strategies.

In the last two decades, researchers have begun to reveal the genetic bases of fungal pathogenicity using molecular genetic approaches involving loss-of-function and gain-of-function mutations to identify causal pathogenicity genes. For example, molecular identification of the first pathogen effector genes (called avirulence factors) in *Pseudomonas syringae* pv. *glycinea* (Staskawicz *et al.*, 1984) provided a major advance in plant pathology. This seminal work demonstrated that a single gene could explain why a pathogen is unable to cause disease in a resistant host that contains a matching resistance gene. Shortly afterwards the first fungal effector genes were identified (Walton, 1987). However, the first host resistance genes that encode receptor-like proteins were not molecularly characterized until the mid-1990s (Staskawicz *et al.*, 1995).

Genome sequence technology has provided a powerful means to compare whole genomes of pathogen isolates with differing pathogenicity attributes. Recent progress in sequencing technologies (NGS - Next Generation Sequencing) has led to a huge increase in the number of projects involved in fungal genomes sequencing and analyses. For example, genome analyses have been successfully used to study

fungi in the Order Sordariomycetes such as *Colletotrichum* (O'Connell *et al* 2012), *Trichoderma* (Kubiceck *et al.* 2011), *Fusarium* (Ma *et al.* 2010) and *Magnaporthe* (Xue *et al.,* 2012). Together, comparative genomics and functional genetics studies offer great promise to improve our understanding of host-pathogen interactions.

The implicit hypothesis of earlier studies was that pathogenic organisms diverged from closely related saprophytic species by acquiring a set of key genes that confer pathogenicity (Oliver, 2012). In fungal pathogens, the situation appears to be more complicated than in bacterial pathogens. There are also striking contrasts in sets of genes that distinguish trophic lifestyles of fungi: biotrophs which rely on effector proteins to suppress host defense to keep the host cell alive before further development and reproduction of the pathogen, and necrotrophs which often rely on expression of nonspecific toxins (or secondary metabolites) and enzymes involved in degradation of host biopolymers to actively cause disruption of host cell membranes and ultimately host cell death. Hemibiotrophic fungi are also common, which initiate the infection process as a biotroph but then switch gene expression to a necrotrophic phase (Oliver, 2012).

6.1.2   Genes involved in fungal virulence

The pathogenicity to plant appears to have arisen multiple times during fungal evolution (van der Does and Rep, 2007). Sometimes the ability to induce disease in particular hosts is based on specific genes that distinguish pathogenic fungi from closely related non-pathogenic organisms. These genes encode "virulence factors," such as small, secreted proteins and toxins. However others sets of genes are involved in plant pathogenicity such as nonspecific toxins (or secondary metabolites)

and enzymes involved in degradation of host biopolymers that cause disruption of host cell wall and membranes and ultimately host cell death.

6.1.2.1 Carbohydrate-Active Enzymes (CAEs).

So far only a few cell wall-degrading enzymes (CWDEs) have been reported as having an important role of pathogenicity (ten Have *et al*., 2002), probably due to the genetic redundancy of these genes. Among CAEs, specific pectin lyases (PL) have been shown to be important pathogenicity enzymes, although they also represent an expanded family of genes (Lara-Márquez *et al.*, 2011). Across all structurally-related catalytic enzymes that degrade glycosidic bonds, the most studied are the endo-polygalacturonases. In the model species *Botrytis cinerea*, deletion of a gene encoding an endo-polygalacturonase reduced virulence on tomato, although at least five other endo PG genes are present (ten Have *et al.,* 1998). Plants also have genes that encode polygalacturonase-inhibiting proteins (PGIPs), which are extracellular used in defense by inhibiting fungal endopolygalacturonases (De Lorenzo *et al.,* 2001).

The critical role of pectin degrading enzymes has been further studied in several systems as well as *Colletotrichum*. Yakoby *et al.* (2001) revealed that heterologous expression of one PL from *C. gloeosprioides* in *C. magna* increased virulence of transformants on watermelon. Wey *et al.* (2002) studied the expression of two PL genes in *C. gloeosporioides* f. sp. *malvae* and proposed that the encoded proteins may be required for fungal growth in host tissue during both biotrophic and necrotrophic phases of infection. A comparison between pathogenic and non-pathogenic races of *C. lindemuthianum* revealed significant differences in terms of growth and production of extracellular pectin lyase activity on different sources

(Hernández-Silva *et al.*, 2007). The most fully studied system regarding CWDE has been *Cochliobolus carbonum*. Disruption of a gene encoding a protein kinase involved in carbon-catabolite repression resulted in a significant reduction of expression of several CWDE-genes and in a drastically reduced number of spreading lesions, showing an active role of this class of genes in pathogenicity (Tonukari *et al.*, 2000).

6.1.2.2 Toxins and secondary metabolites related genes

Fungi produce a wide range of secondary metabolites (SM) with important functions. These organisms encode four key groups of SM such as polyketides produced by polyketide synthases (PKS), peptides produced by nonribosomal peptide synthases (NRPS), alkaloids produced by dimethylallyl tryptophan synthases (DMATS) and terpenes produced by terpene synthases (TS). Some genes can also include both PKS and NRPS function, and they are described as PKS-NRPS hybrids (Bergmann *et al.*, 2007). Fungal pathogens acquire nutrients from their host plants. Some pathogens can weaken plant defence responses to provide such nutrition for their growth and colonisation. Secondary metabolite toxins produced by these organisms often play a role in eliciting these reactions. The production of toxins that kill or weaken plant cells prior the penetration is a widespread phenomenon in plant pathogenic fungi such as *Cochliobolus, Fusarium* and *Alternaria* (Panaccione, 2003; Scott, 2012; Tsuge *et al.*, 2012). These compounds can be either host non-specific (effective on numerous unrelated host-plants) or host-specific (restricted to a certain species genotype). For example, functional analyses have shown that the HT-toxin produced by *Coc. carbonum* Race 1 (Tox2+) is required for virulence in the maize (Walton,

1996). The toxin genes in this fungus are organized in a single cluster (locus TOX2) of seven genes involved in biosynthesis, export and regulation of the toxin (Ahn *et al.,* 2002). Clustering of genes involved in toxin biosynthesis seems to be a common feature in fungal pathogens (Akagi *et al.*, 2009).

In *Alternaria alternata*, different pathotypes produce host-specific toxins, and the role of these toxins as pathogenicity factors on pear (Tanaka *et al.,* 1999) and on apple (Johnson *et al.,* 2000) has been confirmed by gene's knock out transformants.

In the *Fusarium* complex, comparative genomic (including *F. oxysporum, F. graminearum* and *F. verticillioides*) analysis revealed lineage-specific genomic regions in *F. oxysporum* that include four entire chromosomes. Lineage-specific regions contain genes related to pathogenicity. Ma *et al.* (2010) demonstrated that two lineage-specific chromosomes can be transferred from a pathogenic isolate of *F. oxysporum* to a non-pathogenic isolate, and convert a non-pathogenic strain into a pathogen. Within the host non-specific toxins, the trichothecenes produced by various *Fusarium* pathogens have been analysed in detail. Inactivation of a gene related to the first step of trichotecenes biosynthesis (*tox*5/*tri*5) caused the reduction of virulence of *Gibberella pulicaris* on parsnip and of *G. zeae* on wheat (Desjardins *et al.* 1993; Proctor *et al.* 1995).

6.1.2.3 Effectors

The gene-for-gene hypothesis was proposed by Harold Henry Flor (1947) who was working with flax rust caused by *Melampsora lini* in *Linum usitatissimum* (Flor, 1947). This theory states that for every avirulence (Avr) gene in the pathogen there is a related resistance (R) gene in the host, and the interaction between the two gene-

products activate host defense, such as the hypersensitive response. Since then, many plant pathologists have focused their research on molecular and biochemical evidence of the gene-for-gene model. The molecular cloning and characterization of the first fungal *Avr* gene was in 1991 (van Kan *et al.*, 1991), a race-specific peptide elicitor from *Cladosporium fulvum* that induced a hypersensitive response on Cf9 tomato genotypes. With recent advances in genomics and comparative genomics much attention has been directed to candidate secreted effector proteins (CSEPs) (Spanu *et al.*, 2010; O'Connell *et al.*, 2012). By definition CSEPs are predicted secreted proteins that do not have BLAST sequence similarity hits outside the taxa considered (e.g. genera-specific CSEPs, species-specific, etc.)

6.1.3   The first *Colletotrichum* genomes – what did they reveal?

A recent work on lifestyle transitions in *Colletotrichum* species decrypted by genome and transcriptome analyses has been published. O'Connell *et al.* (2012) used genome sequencing and transcriptome analyses to compare two *Colletotrichum* species: the brassica pathogen *C. higginsianum*, which can also infect *Arabidopsis thaliana*, and the maize pathogen *C. graminicola*. Both fungi have large sets of genes related to pathogenicity compared to other organisms. However, gene families encoding secreted effectors, pectin-degrading enzymes, secondary metabolism enzymes, transporters and peptidases are more expanded in *C. higginsianum*. Transcriptome profiling revealed genes with different expression patterns that were associated with the biotrophic or necrotrophic stage of pathogen development. This includes effectors and secondary metabolism proteins that are induced before

penetration and during the biotrophic phase, and hydrolases and transporters that are up-regulated during the necrotrophic phase.

6.2 OBJECTIVES

The final part of my thesis research focuses on genome sequencing and analysis of a single isolate of *C. simmondsii*, and comparison of this genome with reference genomes of *C. graminicola* and *C. higginsianum*.

After an extensive study of *C. acutatum* population structure we found interesting evolutionary relationships that characterize, and link different populations with biological processes. Phylogenetic analyses suggest a hypothetical heterothallic ancestral population and the advent of two different events: the first related to the acquisition of homothallic capability in a specific subpopulation and the other the loss of mating behavior in other populations. The genetic bases of sexual behavior in *Glomerella* species have never been revealed. Genome comparison of isolates with different mating behavior has been suggested as an effective approach to identify genetic bases of the unbalanced heterothallism (Menat J, *et al.*, 2012). Also through the population study, and thanks to previous work, we were able to identify host-specific populations. Genome comparisons between isolates with different characteristics could reveal the genetic bases of diversity in host preference. Pathogenicity assays have been carried out to evaluate level of aggressiveness and host specificity using *C. acutatum/ Fragaria* x *ananassa* − [common name: strawberry] as a model system.

The work in this chapter was comprised of the following steps:

1. Paired-end sequencing of genomic DNA from a *C. simmondsii* isolate.

2. *De novo* sequence assembly using different software and techniques.

3. Structural and functional annotation analysis of the genome assembly using different gene annotation tools such as AUGUSTUS and Genemark.

4. Comparison of genes involved in mating behavior relative to published examples.

5. Comparison of several genes families involved in plant interaction, by manual protein annotation using BLAST sequence similarity and InterProScan to assign functional categories to each protein.

6.   Comparison of gene content in the *C. simmondsii* isolate with reference genomes of *C. graminicola* and *C. higginsianum*.

Step 5 focused on prediction of novel putative genes involved in pathogenicity and host specificity such as effector genes. Another approach was based on the investigation of expansion of gene families related to host interactions such as carbohydrate-active enzymes coding genes or genes involved in secondary metabolites such as toxins.

Genomic data were produced and predicted genes were identified in order to provide a platform that will be useful for future sequencing and genome comparison with other members of the *C. acutatum* species complex. This information could be used to find novel methods for diagnostics, disease control and for a better understanding of pathogen evolution.

6.3 <u>MATERIALS AND METHODS</u>

6.3.1   Genome sequencing and assembly

DNA extraction for genome sequencing has been carried out as described in chapter 2.4.4. 50μl of DNA 20μg/μl of genomic DNA in AB buffer have been submitted for sequencing. Genomic libraries with an average insertion size of 260bp were constructed using TruSeqTM RNA and DNA Sample Preparation Kits provided by Illumina ® Sequencing. Genomic DNA has been sequenced using Illumina GAII 76bp (X2) mate-pair reads at the School of Life Sciences of the University of Warwick.

Geneious 5.6.5 has been used to perform an assisted assembly of the Illumina raw data against two *Colletotrichum* genomes available on the Broad Institute website.

De novo assembly has been performed using four different software packages: **1.** SOAPdenovo (V1.05, Released 22-02-2011) - part of the Short Oligonucleotide Analysis Package, expressly designed to assemble Illumina GA short reads; **2.** ABYSS Assembly By Short Sequences (V1.3.4, Released 30-03-2012) a parallelized sequence assembler (Simpson, *et al.*, 2009); **3.** IDBA (V0.17 Released 08-2010) *De Novo* Assembler (Peng, *et. al.*, 2010); and **4.** VELVET (V 1.2.08, Released 04-2011) - *de novo* genomic assembler specially designed for short read sequencing technologies (Zerbino, 2010).

Assembly has been carried out under the guidance of Michael Thon and his research group at Centro Hispano-Luso de Investigaciones Agrarias, Universidad de

Salamanca (Spain) using bioinformatics facilities available at the SuperComputing and BioInformatics Center of the University of Malaga (Spain).

6.3.2   Genes structure annotation

Different bioinformatc tool were used for gene structure annotation. GeneMark.hmm-E (V2.0 - Ter-Hovhannisyan, *et al*., 2008), trained with EST evidence from *C. graminicola* (provided by Mike Thon's group), was used to predict transcripts across *C. simmondsii* genome. AUGUSTUS (V 2.6.1) gene prediction program for eukaryotes (Stanke, et al. 2004) has then run, incorporating hints on the gene structure coming from a *C. acutatum sensu lato* 955 EST library sequences available in GenBank. Only sequences showing nucleotide BLAST evidence with the *de novo* genome assembly were used to train AUGUSTUS. The performance of each method was compared based on predicted proteins features, BLAST evidence and InterproScan (V39.0 - Quevillon *et al*., 2005) predicted domains. The one showing most reliable results (based on number of genes, length and gene integrity) have been used for further investigations.

6.3.3   Gene function prediction

A preliminary gene function prediction has been performed using the Blast2GO (Contesa *et al*., 2005) Pipeline Version (B2G4Pipe – database version b2g_aug12) based on BLAST and InterProScan results (http://www.blast2go.com). For each specific gene family analyzed, gene function have been improved using different software tools and manually.

Genes encoding putative carbohydrate-active enzymes were identified manually using the CAZy resources (Cantarel *et al*., 2009) available at

http://www.cazy.org/. Identification has been carried out using an integrated approach of BLAST evidence and PFAM domains using InterProScan (V39.0 - Quevillon *et al.*, 2005).

Web-based tool SMURF "Secondary Metabolite Unknown Region Finder" (Khaldi *et al.*, 2010) was used to predict secondary metabolism gene clusters and genes such as polyketide synthases (PKSs), nonribosomal peptide synthetases (NRPS), and dimethylallyltryptophan synthases (DMATs). SMURF finds secondary metabolite biosynthesis genes and pathways in fungal genomes based on based on PFAM and TIGRFAM domain content as well as on a gene's chromosomal position (Khaldi et al., 2010). Terpenoid synthase have been identified using an integrated approach of BLAST evidence and specific conserved domains.

The secretome of *C. simmondsii* was predicted using WoLF-PSORT (Horton *et al.,* 2007). Candidate secreted effector proteins (CSEPs) were identified as described by O'Connell *et al.* (2012) as extracellular proteins with no significant BLAST sequence similarity (expect value $<1 \times 10^{-3}$) to sequences in the UniProt database (Swiss-Prot and TrEMBL components database relaeased in May 2012). Homologs of proteins from outside the genus *Colletotrichum* were excluded.

All candidate genes identified using automated searches were inspected manually, including protein sequence similarity to known enzymes conserved domain through InterproScan database.

6.3.4   Genomes comparison

Two *Colletotrichum* genomes were publicly available for this project including the cereal pathogen *C. graminicola* and the brassica pathogen *C. higginsianum*. Eight

genomes from other ascomycete fungi with different lifestyles were also available for further comparisons: *Aspergillus nidulans* and *Neurospora crassa* as saprotrophs; *Fusarium oxysporum*, *Botryotinia fuckeliana*, *Sclerotinia sclerotiorium* and *Verticillium dahliae* as necrotroph; *Magnaporthe grisea* as hemibiotroph) and *Ustilago maydis* (basidiomycetes) as out-group (Table 6.1).

Table 6.1 Complete list and source information of fungal genomes used in the study for the genome comparison.

Organism	Strain	BioProject	NCBI ID
Colletotrichum simmondsii	CBS 122122	-	-
Colletotrichum graminicola	M1.001	PRJNA37879	2138
Colletotrichum higginsianum	IMI 349063	PRJNA47061	11306
Verticillium dahliae	VdLs.17	PRJNA28529	832
Fusarium oxysporum	4286	PRJNA18813	707
Magnaporthe grisea	70-15	-	13132
Neurospora crassa	OR74A	PRJNA132	19
Sclerotinia sclerotiorum	1980	PRJNA20263	487
Botryotinia fuckeliana	T4	PRJNA64593	494
Aspergillus nidulans	FGSC A4	PRJNA40559	17
Ustilago maydis	521	PRJNA29393	70

Data and information related to genes encoding putative carbohydrate-active enzymes have been obtained by the recent publication on "Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses" (O'Connell *et al.*, 2012).

Secondary metabolism gene clusters and genes have been identified as done for the *C. simmondsii* genome, based on web-based tool SMURF and using an integrated approach of BLAST evidence and specific conserved domains. Genome comparison related to secreted protein and candidate secreted effector proteins have

been focused only on *Colletotrichum* genomes focusing on the identification of candidate genes involved in host specificity.

6.4 <u>RESULTS</u>

6.4.1   Selection of isolates

The member of the *Colletotrichum acutatum* species complex sequenced in this work is strain CBS 122122 (also known as BRIP 28519; HKUCC 10928; ICMP 17298; KACC 43258). The strain was collected during May of 1987 by L.M. Coates in Queensland, Australia from infected fruit tissues of papaya [sn: *Carica papaya*].

Based on previous results presented in Chapter 4, and a recent publication on the reassessment of *C. acutatum* species complex, this isolate, belonging to A9 genetic group, has been named as *C. simmondsii*, and it has been nominated as the holotype of the species (Damm *et al.*, 2012). The species does not show a specific host range, because strains belonging to this specie have been isolated from infected tissues of different hosts such as *Carica* spp. (Caricaceae), *Cyclamen* spp. (Myrsinaceae), *Fragaria* spp. (Rosaceae), *Mangifera* spp. (Anacardiaceae), *Protea* spp. (Proteaceae). This strain was selected for sequencing because it is commonly used in research laboratories as a reference for evolutionary analyses and phylogenetics. Furthermore this strain isolated from papaya has been tested for pathogenicity against strawberry and different groups across Europe are testing it against different non-fruit crops systems. A sexual state has not been reported.

Genomic DNA and libraries were produced also for two other isolates for further genome sequences: PJ7 - genetic group A3 (*Colletotrichum fioriniae;*

teleomorph: *Glomerella fioriniae*), isolated from strawberry (sn: *Fragaria* x *ananassa*) infected tissues in New Zealand by Peter R. Johnson and described as heterothallic mating tester based on his capability of undergo sexual crosses with most other strains (Guerber *et al.*, 2001). And CBS 607.94 - genetic group A7a (*Colletotrichum salicis;* teleomorph: *Glomerella salicis*), isolated from leaf spot of salix (sn: *Salix* sp.) in the Salix Forest near Blocq van Kuffeler in the Netherlands by H.A. van der Aa and described as homothallic (Chapter 3).

The three species used in this study have also been chosen based on different mating behaviour; the set include one strain capable of self-fertilization [CBS 607.94], one is self-sterile/cross-fertile [PJ7] and one is supposed to be asexual [CBS 122122] (sexual state has never been observed). Furthermore the strains show different host association patterns: two strains belong to species showing a wide host range [CBS 122122, PJ7] and one to a species that shows high host specificity to *Salix*. We also chose the isolates based on their phylogenetic distances, trying to cover the entire species complex making the genomes suitable reference for further projects.

To position the strains described above together with those sequenced in other projects, a phylogenetic tree of 29 *Colletotrichum* strains (plus *Monilochaetes infuscans* as outgroup) was generated on five loci (ITS, CHS-1, HIS3, ACT, TUB2 partial sequences) sequenced in previous projects and available on-line (Figure 6.1).
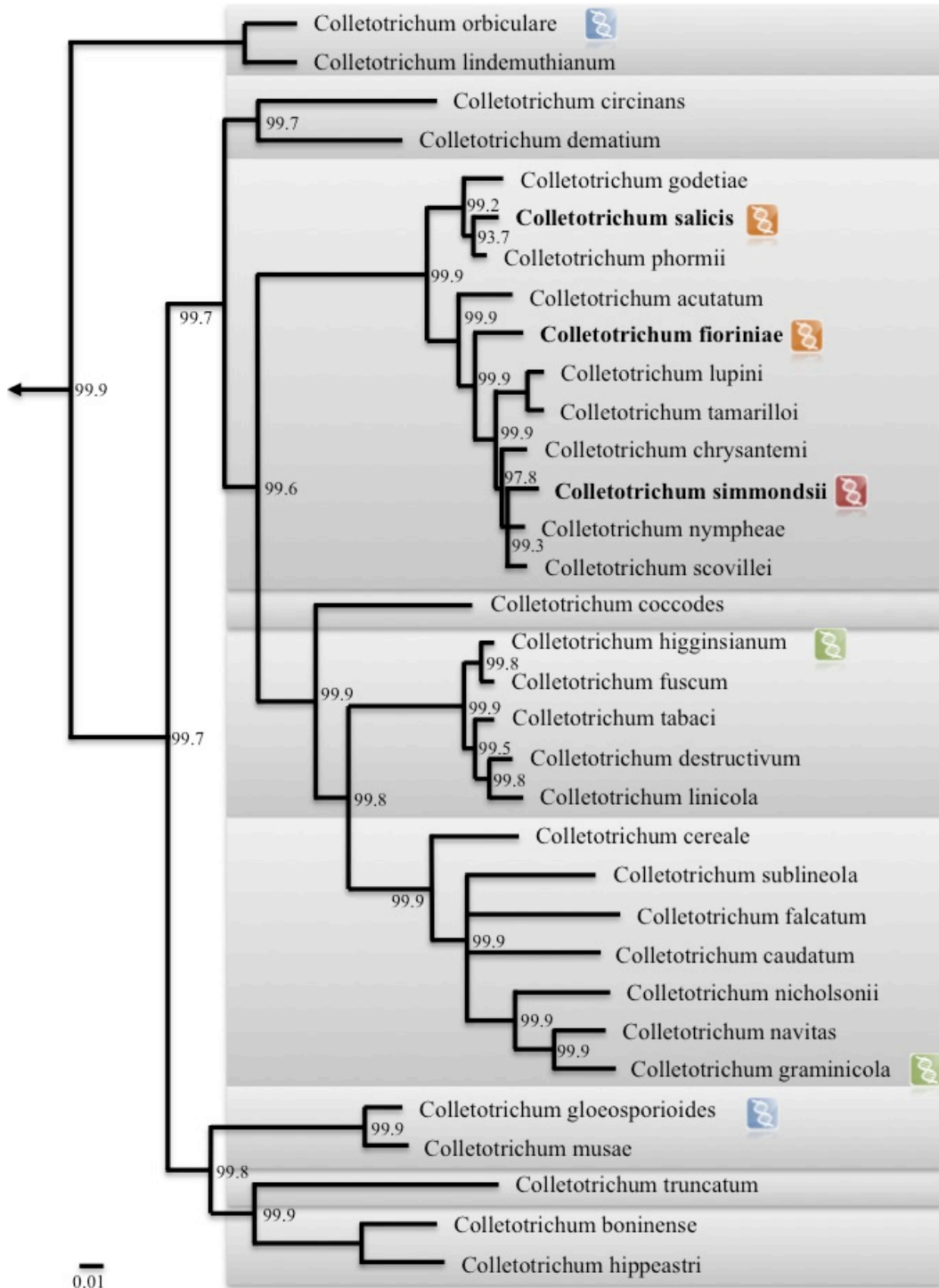
Figure 6.1 Evolutionary relationship of 29 *Colletotrichum* species subdivided in species complexes obtained from a Markov Chain Monte Carlo (MCMC) algorithm used to generate phylogenetic trees with Bayesian probabilities using MrBayes 3.2.1 (Ronquist and Huelsenbeck, 2003) on a multi-locus alignment (ITS, CHS-1, HIS3, ACT, TUB2). Bootstrap support values (10000 replicates) above 90 % are shown at the nodes. *Monilochaetes infuscans* was used as an outgroup. DNA logos indicate genome sequence projects: green for genome sequence completed and available, blue completed but not released, red sequenced and analyzed in this work and orange for works in progress. In bold are organisms used in this work

6.4.2   Sequencing and Assembly Statistics

The genome sequence of *C. simmondsii* was generated by Genomics Facility at the School of Life Sciences of the University of Warwick (Coventry, United Kingdom) using Illumina GAII pair-end reads (260bp insert lenght). Output of Illumina sequencing resulted in 28,125,902 sequences 70bp long (after tags trimmed) for a total of almost 2×10^9 nucleotides. Considering a hypothetical genome size of 50 Mb expected average coverage was calculated to be 40X.

Since an important step of the *de novo* assembly process is to generate a set of read-read alignments, errors introduced in this step could have a major effect on the final product (Gnerre at al., 2009). For this reason, assisted assembly against two *Colletotrichum* genomes available on-line was unsuccessfully performed confirming the high divergence and diversity of the three genomes. Less than 1.5 million (4%) reads were aligned to the *Colletotrichum* genomes. Therefore *de novo* genome assembly was performed using different software (Table 6.2) by Michael Thon and his research group at Centro Hispano-Luso de Investigaciones Agrarias, Universidad de Salamanca (Spain).

Table 6.2. Four different software packages were used to perform *de novo* assembly of *Colletotrichum simmondsii* genome sequence.

Program	Obs.	Assembly Size	NumContigs	N50	N90
SOAP	31 k-mer	49583274	3391	48765	13658
ABYSS	31 k-mer	49790207	972	183466	58816
IDBA	31 k-mer	50956265	3583	52445	11515
VELVET	-ins length 200	50389248	1114	206154	49764
VELVET	-ins length 260	50481459	1134	198936	48118
VELVET	-ins length AUTO	50494453	1129	198969	48125

The best assembly based on total supercontigs size, number of supercontigs and N50/N90 value (weighted median statistic such that 50 and 90% of the entire assembly is contained in contigs or scaffolds equal to or larger than the value) (shown in Table 6.2) and total contigs length was given from Velvet using auto insertion length option. Assembly comprised 1.129 supercontigs with total length of 50.5 Mb (contig N50 length = 198.97 kb) (Table 6.3). The best Velvet assembly has been used to perform gene prediction and for further analyses. Results were compared to the one related to the other *Colletotrichum* genome sequencing projects carried out using different and integrated technology (Table 6.3).

Table 6.3. Assembly statistics related to three *Colletotrichum* genomes.

Genome features	*C. graminicola*	*C. higginsianum*	*C. simmondsii*
Genome physical size (Mb)	57.44	53.35	Unknown
Total contig length (Mb)	50.87	49.08	50.49
Contig number	654	10.235	1.129
Average contig length	78966.8	4795.8	44724.9
standard deviation	218190.4	3779	93457.5
minimum	2000	921	200
maximum	1824042	49362	669049
Average base coverage (Fold)	9.1	101	38.99
Sanger	7.9	0.2	-
Roche 454	1.2	25	-
Illumina	-	76	38.99
N$_{50}$ contig (kb)	228.96	6.15	198.97
GC-content (%)	49.12	55.10	51.6

Data related to *C. graminicola* and *C. higginsianum* have been either calculated on available data downloaded or obtained from the Colletotrichum Sequencing Project, Broad Institute of Harvard and MIT (http://www.broadinstitute.org/).

Considering the different approaches used to sequence the three *Colletotrichum* species we were able to reach a good and comparable assembly using only Illumina GAII sequencing approach. This observation is in contrast with the conclusion reached by O'Connell et al (2012) suggesting that *C. higginsianum*

assembly shows a higher fragmentation than that of *C. graminicola* and because compiled from short-read (Roche 454 sequencing technology) data only. Having said that, differences in assembly results might be due to the nature and biology of the organism and its DNA organization. Regarding the genome size further statistical analyses are needed in order to independently estimate the total *C. simmondsii* genome size as shown by Baxter *et al.* (2010).

6.4.3   Predicted Genes and Statistics

A *de novo* gene prediction of the genome sequence using GeneMark and AUGUSTUS gave similar results to *C. graminicola* data. GeneMark was able to predict 14244 genes, whereas AUGUSTUS predicted 13549 genes (Figure 6.2). Differences in length patterns indicate that GeneMark predicted a large number of proteins (816) shorter than 30 amino acids, whereas only one was predicted by AUGUSTUS (Figures 6.3). AUGUSTUS seems to be more reliable than GeneMark because there is no evidence of peptides shorter than 30 amino acids (with the exception of CH063_16038T0 - 29 amino acids) across the two publically available *Colletotrichum* genomes.

Lengths statistic:
Mean: 39.1   Std Dev: 61.6
Minimum: 3   Maximum: 653
100% no blast hits

1216   13028 / 13278   271

Lengths statistics:
Mean: 204.3   Std Dev: 142.1
Minimum: 50   Maximum: 1038
100% no blast hits

○ **Genemark (14244)**   ○ **Augustus (13549)**

blastp e-value cut-off= 1e-1
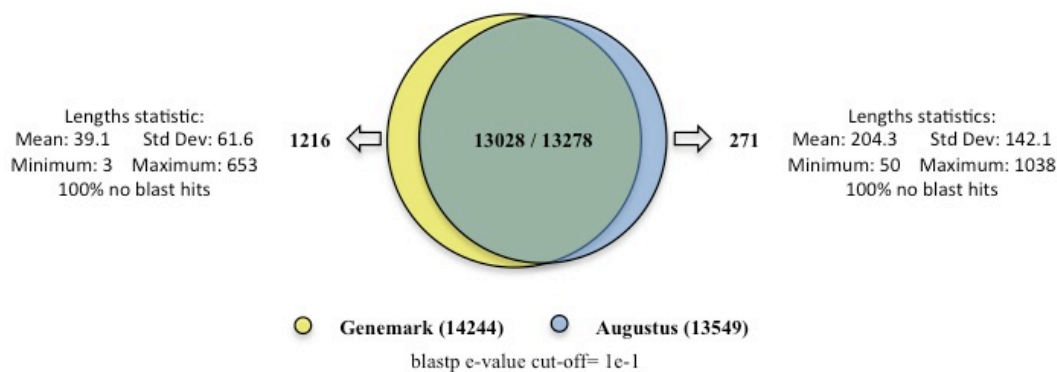
Figure 6.2. Venn diagram showing similar putative proteins predicted by the two software packages used in this study to predict genes in *Colletotrichum simmondsii* genome and features of unshared protein.

GeneMark

AUGUSTUS
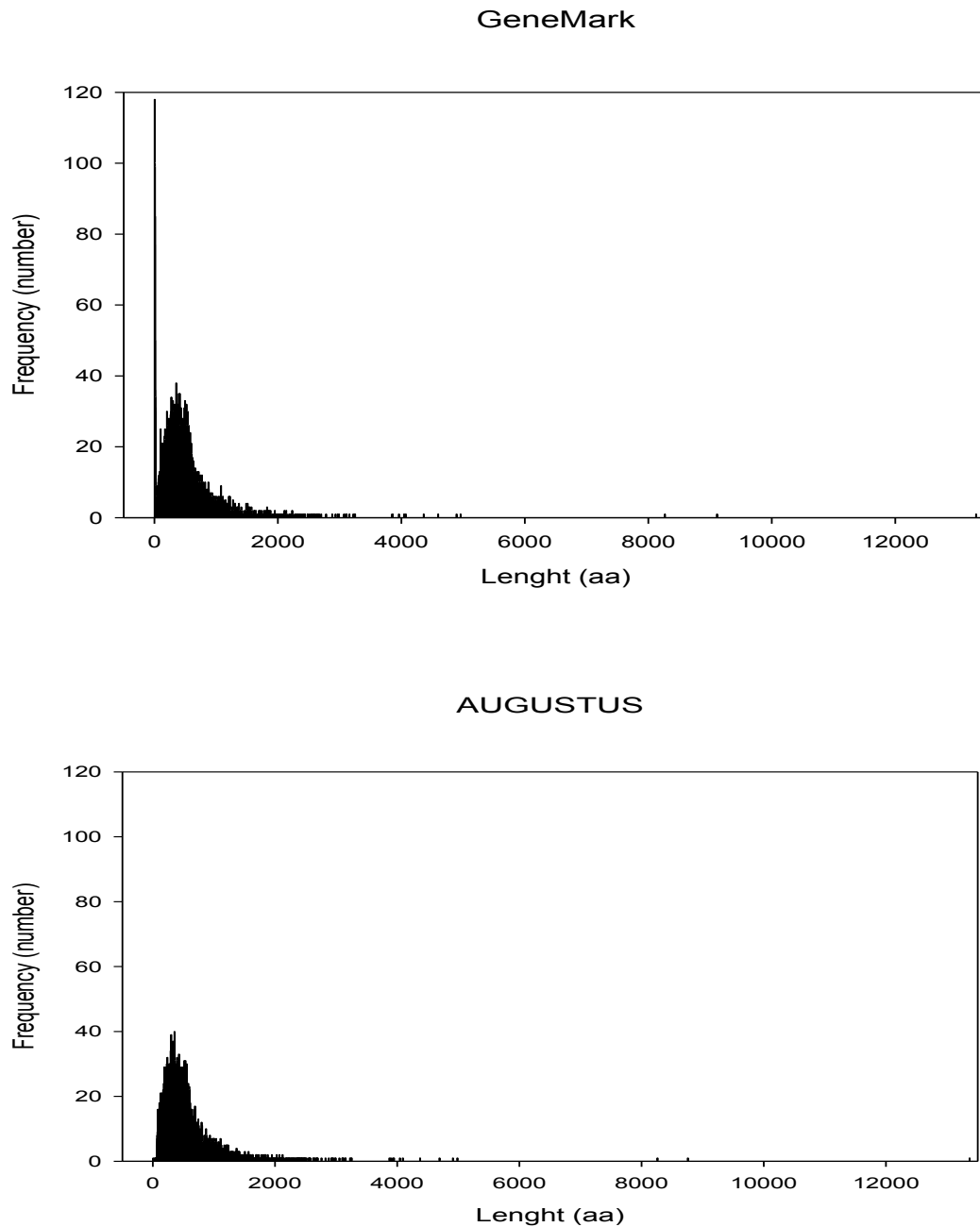
Figure 6.3.   Graphs showing length patterns and features of *Colletotrichum simmondsii* predicted proteins by the two software packages. On the top the graph related to 14344 proteins predicted by GeneMark and on the bottom graph related to 13549 proteins predicted by AUGUSTUS.

Furthermore *C. simmondsii* predicted proteins were compared against annotated proteins of *C. higginsianum* and *C. graminicola* and against protein

available in Swiss-Prot (manually annotated and reviewed section of the UniProtKB). Based on this comparison, the decision was made to use AUGUSTUS instead GeneMark for further analyses.

6.4.4 *Colletotrichum* genome comparison

Gene features obtained for *C. simmondsii* were compared with the data related to *C. graminicola* and *C. higginsianum* (Table 6.4).

Table 6.4. Table showing gene statistics related to 3 *Colletotrichum* genomes.

Genes features	*C. graminicola*	*C. higginsianum*	*C. simmondsii*
Protein-coding genes			
Number	12006	16172	13549
Mean transcript length (bp)	1397	1095	1497
Number of exons	32967	39537	35492
Mean number of introns/gene	2.7	2.4	1.62
Percentage coding	32.51	36.08	40.1
GC-content (%)			
Exons	58.36	59.33	56.2
Introns	50.99	51.63	47.6
Intergenic regions	44.22	52.78	48.61
Protein			
Mean protein length (aa)	465.8	365	498.2
standard deviation	380.4	294.3	384.1
minimum	32	30	27
maximum	8936	4140	13369

Data related to *C. graminicola* and *C. higginsianum* have been either calculated on available data downloaded or obtained from the Colletotrichum Sequencing Project, Broad Institute of Harvard and MIT (http://www.broadinstitute.org/).

C. simmondsii seems to have a large number of genes compared to *C. graminicola* but not to *C. higginsianum*. Based on results of further analyses, approximately 5.2% of the gene models of *C. higginsianum* represent genes that

were split into two or more models, and 4.0% are truncated representations of the true gene structure (O'Connell *et al.*, 2012). Based on observations and also confirmed by statistics related to gene structures features (such as: mean transcripts/protein length), there were a few cases of two genes that are very close to each other but were recognized as a single copy. A manual analysis of the gene prediction may increase the number of genes in the future.
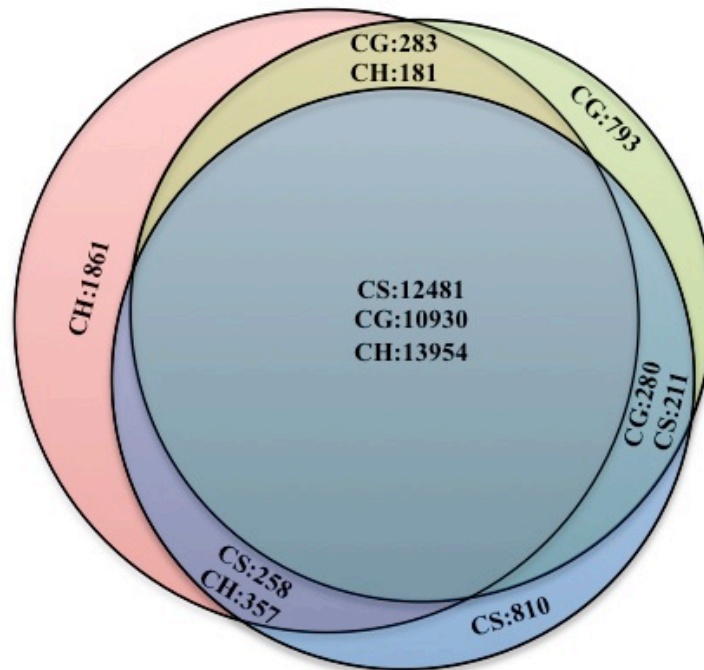
Figure 6.4. Venn diagram showing numbers of the genes that are unique to each isolate, specific to two isolates, and common to all three isolates. Predicted genes of *Colletotrichum simmondsii* (CS), *C. graminicola* (CG), and *C. higginsianum* (CH) are represented with circles colored in yellow, blue, and red, respectively.

Based on protein similarity analyses (e-value threshold of 1e-1), 1861, 793, and 810 genes, respectively, were unique to *C. higginsianum*, *C. graminicola* and *C. simmondsii* (Figure 6.4). Furthermore *C. simmondsii* is characterized by a number of unique genes comparable to those of *C. graminicola* but shares more genes with *C.*

higginsianum. The diagram also shows gene families expansions across the three species and across two species sharing a common set of genes. Overall, a huge difference could be found in the spatial predicted localization (extracellular, citoplasmatic, mitochondrial, etc…) of the species-specific proteins. A comparison of these genes that encode secreted proteins indicated a frequency of 11.60% of *C. graminicola*, 12.47% of *C. higginsianum* and 21.98 % of *C. simmondsii* (see Chapter 6.4.6.3).

6.4.5   Mating type locus and genes

Sexual reproduction is exceptional in some *Colletotrichum* species and absent (or not identified yet) in most of them. Previous results suggest a strong relationship between phylogenetics (or novel species) and mating behavior. The strain used in this study does not appear to be able to undergo sexual reproduction, either on its own or with other strains. Sexual reproduction in all *Colletotrichum* species studied is different from the normal mating system in other filamentous (Wheeler, 1954). *C. simmondsii*, like *C. graminicola* and *C. higginsianum*, has a single Mat1-2 gene containing a characteristic high mobility - DNA binding group domain (HMG box). On the other side, across the three genomes there is no evidence for a Mat1-1 gene. Only one example for the presence of a gene with classic structure of the Mat1-1 has been reported in one *Colletotrichum* species (*C. musae*). Unfortunately this organism was wrongly classified and has recently been reassigned to *Gibberella avenacea*. Analysis of synteny in the region of the Mat1-2 locus in *C. simmondsii* revealed a conserved group of 14 genes (Cia30, Apc5, Cox13, Apn2, Mat1, Sla2, L21e, S49, Slu7, Rev3, Tex2, Ami1, Cwc24 and Atg3) shared with *C. graminicola*

and other Sordariomycete fungi. However, analysis carried out manually on the locus revealed few ORFs and hypothetical proteins that might be involved in the genetics of mating behavior or might be evolutionary derived by an ancestral gene (Figure 6.5).
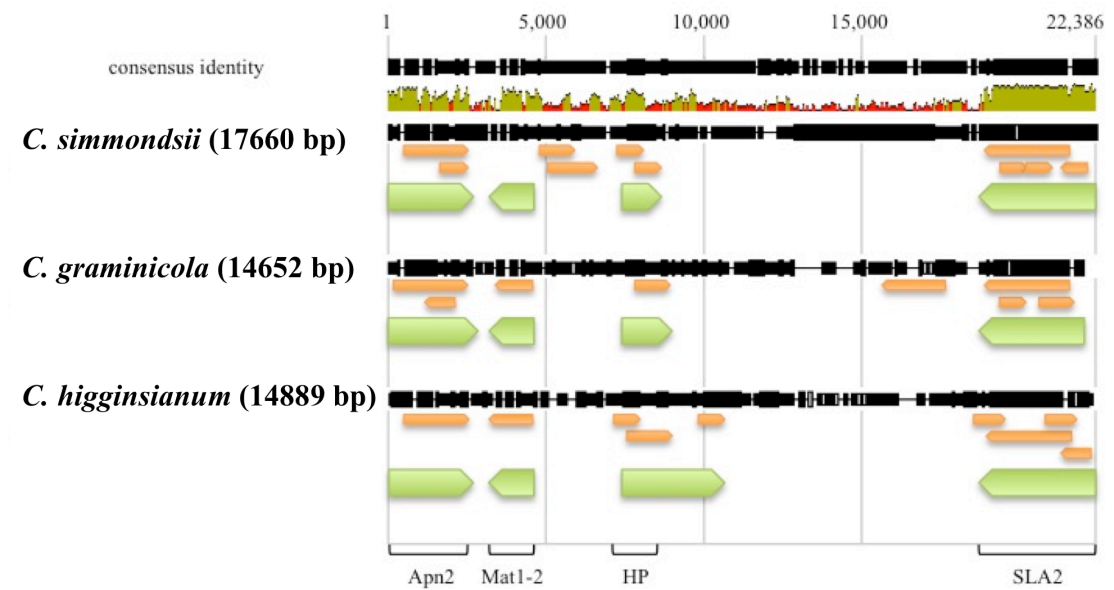
Figure 6.5. Alignment performed with MAUVE (improved by Geneious) of the mat locus of the three *Colletotrichum* species sequenced. Green bars indicate genes predicted running AUGUSTUS using only the genomic portion of the MAT1-2 gene and flanking regions; orange bars indicate ORFs starting with a methionine and longer than 500 bp identified.

The locus in *C. simmondsii* between the *apn2* and *Sla2* genes is 3kb longer compared to the same region in the other *Colletotrichum* genomes. A hypothetical protein was predicted running AUGUSTUS trained with *Fusarium graminearum* (closest related species showing a sexual behaviour) with the genomic locus including Apn2/Mat1/Sla2 (running the gene prediction on complete genome sequences those genes were not predicted). Hypothetical proteins that were predicted

do not show any similarity against non-redundant protein available in GenBank or conserved domain when analyzed with InterProScan. The set of data, even showing some interesting patterns doesn't revile the genetic bases of the mating behavior. Further comparative analyses, based on genome sequences related to strains evolutionary related showing different mating behavior, could be a valid approach for an improvement understanding of genetics (as well as the evolution) of sexual reproduction in this genus.

6.4.6   Genes involved in pathogenesis

In order to identify genes involved in pathogenicity, attention was focused on three major classes: **1.** families of structurally related catalytic and carbohydrate-binding modules (or functional domains) of enzymes that degrade, modify, or create glycosidic bonds; **2.** genes involved in secondary metabolism such as polyketide synthases (PKS), non-ribosomal peptide synthetases (NRPS), terpenoid synthases (TS), dimethylallyl diphosphate tryptophan synthases (DMATS) and their clusters in the genome; and **3**, candidate effectors proteins (CEPs).

The choice of gene classes has been based on results shown by O'Connell *et al.* (2012) in the recent publication focused on *C. graminicola* and *C. higginsianum* genomes and transcriptomics comparison. Nine additional fungal genomes were also included the analyses (Table 6.5), which are representative of saprophytic, biotrophic, necrotrophic and hemibiotrophic lifestyles. Evolutionary relationships of organisms included in further analases are shown in Figure 6.6.

Table 6.5. List of fungal genomes (genome statistics and lifestyles) used in the study for the genome comparison.

Organism	Size* (Mb)	Genes	GC (%)	Lifestyle
Colletotrichum simmondsii	50.49	13,549	51.60	Hemibiotroph
Colletotrichum graminicola	51.64	12,006	49.12	Hemibiotroph
Colletotrichum higginsianum	49.08	16,172	55.10	Hemibiotroph
Verticillium dahliae	33.83	10,535	55.85	Necrotroph
Verticillium albo-atrum	32.83	10,221	56.06	Necrotroph
Fusarium oxysporum	61.36	17,708	48.40	Necrotroph
Magnaporthe grisea	41.70	11,074	51.57	Hemibiotroph
Neurospora crassa	41.04	9,733	48.25	Non pathogenic
Sclerotinia sclerotiorum	38.33	14,503	41.84	Necrotroph
Botryotinia fuckeliana	42.66	16,448	43.06	Necrotroph
Aspergillus nidulans	30.07	10,560	50.32	Non pathogenic
Ustilago maydis	19.68	6,522	54.03	Biotroph

In bold is highlighted the organism sequenced in this work. Data related to the other genomes are available at the Broad Institute of Harvard and MIT (http://www.broadinstitute.org/). *Size refer to total contigs leght.
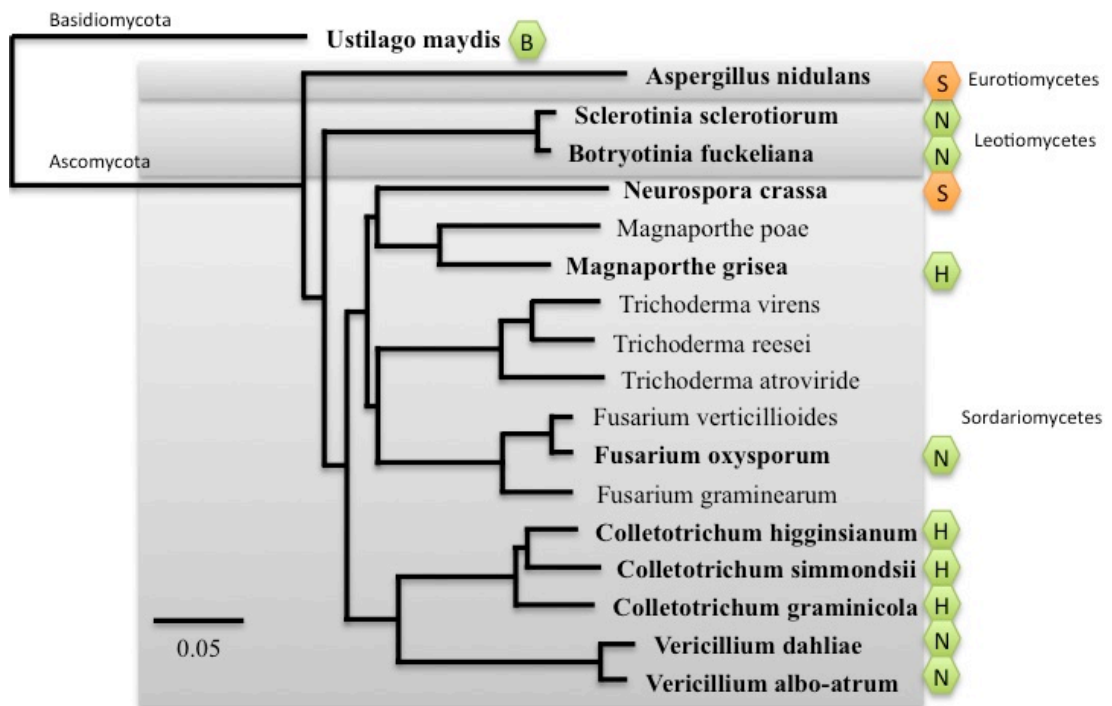
Figure 6.6. Evolutionary relationship and taxonomic references of 18 fungal species subdivided in species complexes obtained from a Bayesian analysis based on the rRNA locus. In bold are highlighted the genomes used in this study for comparison. Green hexagons indicate plant pathogens and orange hexagons saprotrophic fungi; letters inside indicate the lifestyle (B = biotroph, N = necrotroph, H = hemibiotroph and S = saprotroph).

6.4.6.1 <u>CAEs - Genes encoding carbohydrate-active enzymes</u>

Genes encoding putative carbohydrate-active enzymes have been searched in the genome of *C. simmondsii* using a manual approach based on BLAST sequence similarity and conserved domains search using the CAZy database (http://www.cazy.org). Results were compared to those belonging to fungi representing a range of different life-styles. The numbers of enzymes belonging to each CAZy class are shown in Figure 6.7. Data for enzymes acting on pectin, hemicellulose and chitin, as well as proteins with CBM18 and CBM50 carbohydrate-binding motifs, were manually curated.
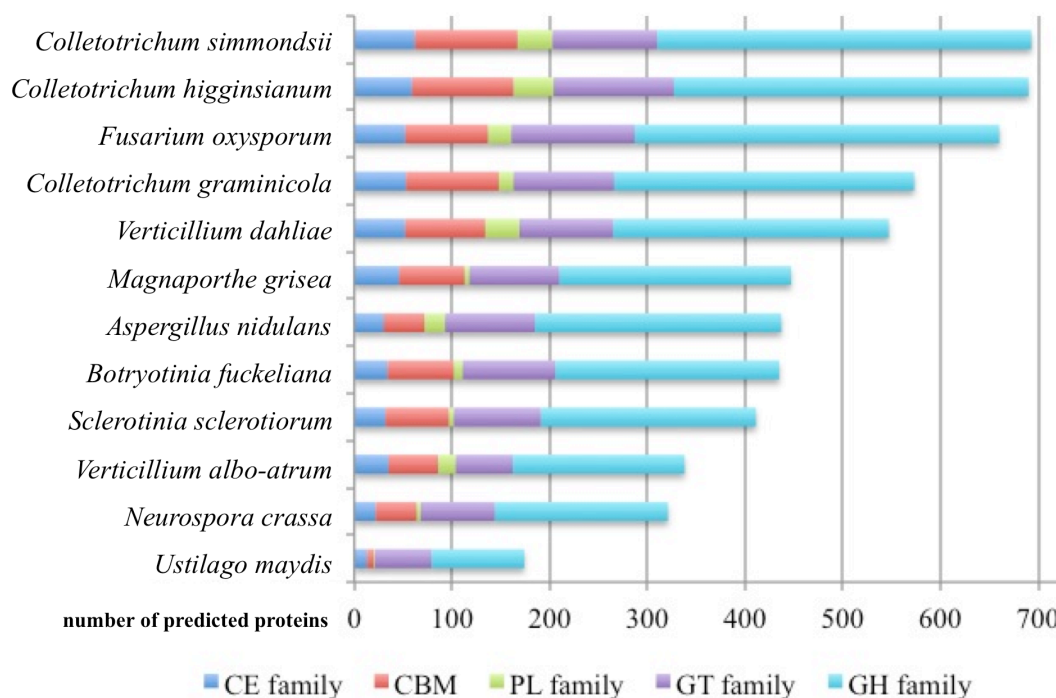
Figure 6.7. Bar diagram showing specific carbohydrate-active enzyme gene classes expansions across 12 fungal genomes showing different behavior. Genes families (group of genes that show sequence similarity and share important biological characteristics) analyzed have been choosen based on the CAZy database (CE, CBM, PL, GT, and GH).

Organisms scanned encode similar numbers of glycosyltransferases (GTs) as these genes are involved in basal activities of the fungal cell. These results are consistent with with those of O'Connell *et al.* (2012). *C. simmondsii* genome contains a total of 693 genes encoding putative carbohydrate-active enzymes, more than all other fungi examined in this study. Families most expanded in this organism compared to others are:

- ✓ 62 genes encoding carbohydrate esterases (CEs) that catalyze the de-O or de-N-acylation of substituted saccharides.

- ✓ 383 genes encoding enzymes that hydrolyse the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety (GHs). The genomes of *C. higginsianum* and *C. graminicola* encode a total of 361 and 307 GHs, respectively.

- ✓ 105 genes characterized by one or more carbohydrate-binding module (CBM). *C. higginsianum* encode 104 CBMs and *C. graminicola* 95.

The large resource of sugar-cleaving enzymes is further extended by polysaccharide lyases (PLs) enzymes encoded by the *C. simmondsii* genome (Figure 6.8b). The total number of PLs in *C. simmondsii* is 36 and it is the second organism with the most expanded PL family after *C. higginsianum* (39 genes). A much more detailed analysis has been carried out to identify which specific sub-class of genes was expanded in the *C. simmondsii* genome compared to of *C. higginsianum* and *C. graminicola* (Figure 6.8a and 6.8b). The aim was to identify the subclass of genes more expanded in *C. simmondsii* and to look at the specific functions.
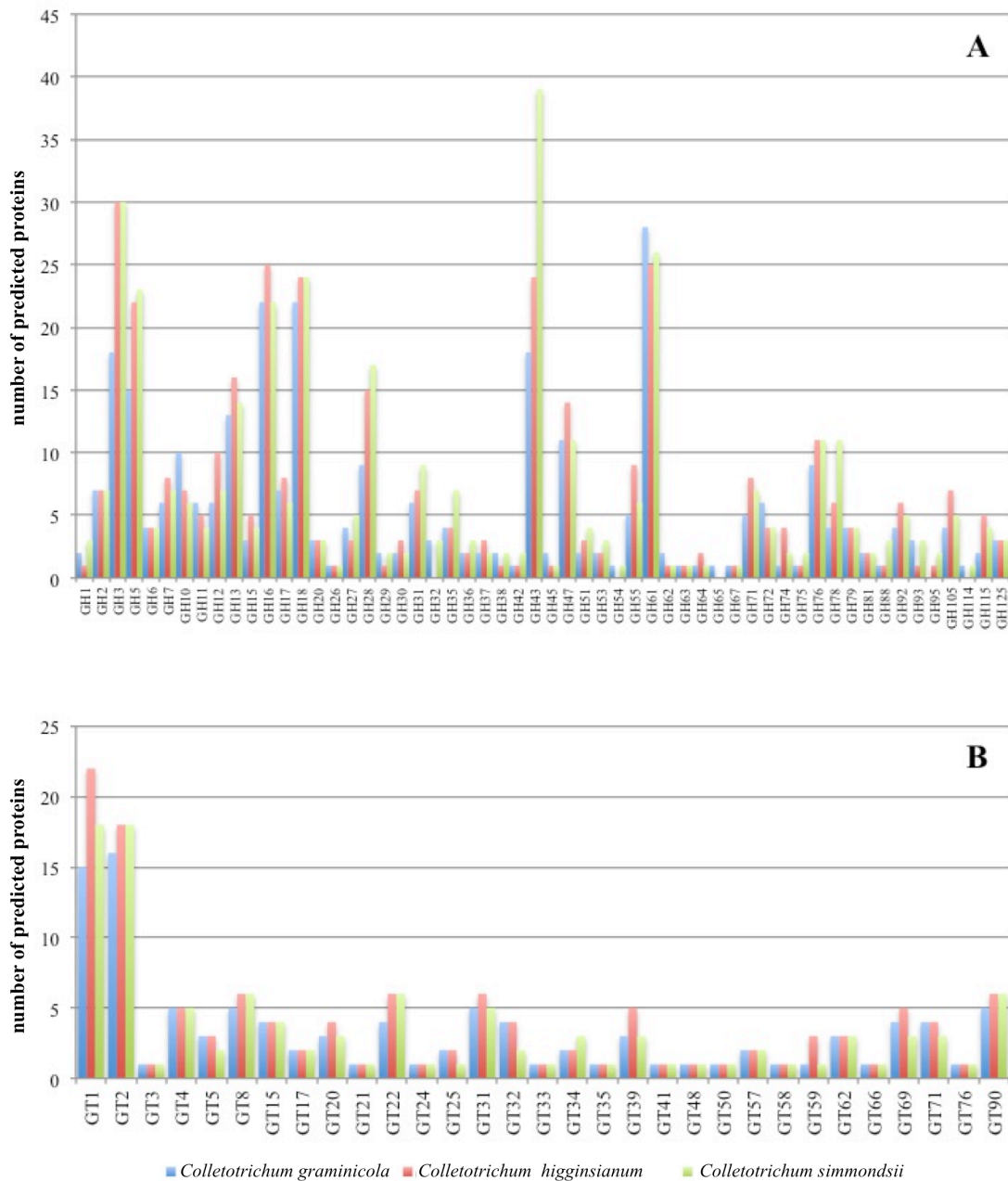
Figure 6.8a. Bar diagrams showing the number of genes related to CAZy subclasses across the three *Colletotrichum* genomes. Only classes with representative genes have been reported. A graph represents the number of genes encoding enzymes that hydrolyse the glycosidic bond between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety (GHs). B graph report the number of genes encoding glycosyltransferases (GTs).
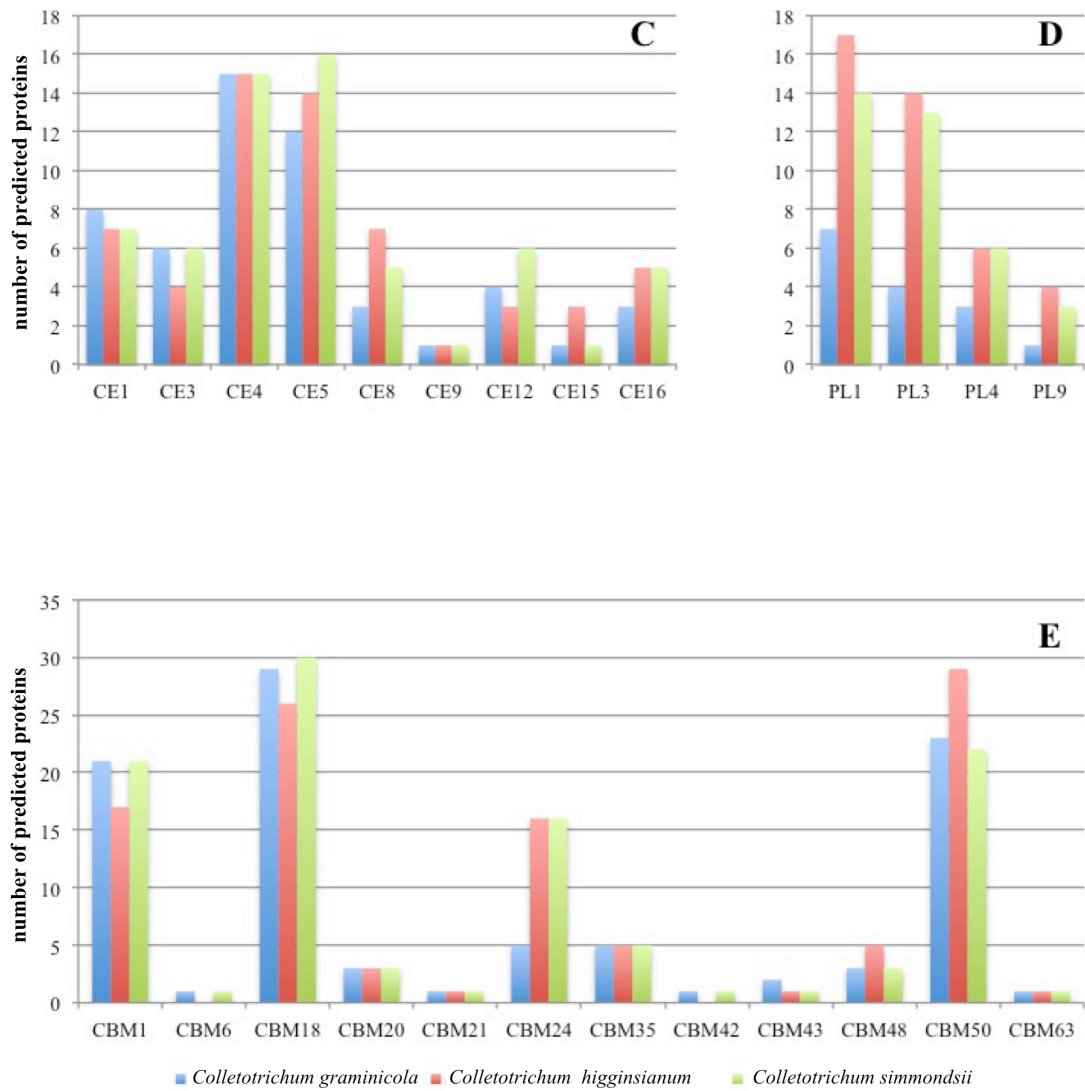
Figure 6.8b - Bar diagrams showing the number of genes related to CAZy sub-classes across the three *Colletotrichum* genomes. Only classes with representative genes have been reported. C graph represents the number of genes encoding carbohydrate esterases (CEs). D graph report the number of genes encoding polysaccharide lyases (PLs) enzymes. E graph shows the number of carbohydrate-binding modules (CBMs).

As shown in the charts above (Figure 6.8a and 6.8b) the enzyme classes more expanded in the new genome are: CE5, CE12, GH1, GH5, GH27, GH28, GH31, GH35, GH36, GH42, GH43, GH51, GH53, GH75, GH78, GH88, GH95. Amongst these, GH43 class has double of the genes (39) compared to the average of the other

two *Colletotrichum* species (24 in *C. higginsianum* and 18 in *C. graminicola*), and is the most expanded gene class across the fungi analyzed.

6.4.6.2 <u>Putative genes related to secondary metabolism</u>

Fungi produce a wide range of secondary metabolites with important functions. These organisms encode four key groups of SM investigated in this chapter:

- ✓ polyketides produced by polyketide synthases (PKS)

- ✓ peptides produced by nonribosomal peptide synthases (NRPS)

- ✓ alkaloids produced by dimethylallyl tryptophan synthases (DMATS)

- ✓ terpenes produced by terpene synthases (TS)

Some genes can also include both PKS and NRPS function, and they are described as PKS-NRPS hybrids. The Secondary Metabolite Unknown Region Finder (SMURF) program (http://jcvi.org/smurf/index.php) was applied to the genomes used in this study in order to identify PKS, NRPS, PKS-NRPS hybrids and DMATS gene. PKS and NRPS genes were manually inspected for conserved domain using InterProScan plugin available for Geneious® Pro 5.6.5 (www.geneious.com) and developed by Michael R. Thon (http://michaelrthon.com/). SMURF was also used for the identification of SM clusters across genomes. TS and P450 genes were identified by BLAST searches against the NCBI databases, and InterProScan analysis. Cytochrome P450 monooxygenases also play an important role in SM modification and are often included in SM gene clusters and therefore they were included in the comparison.
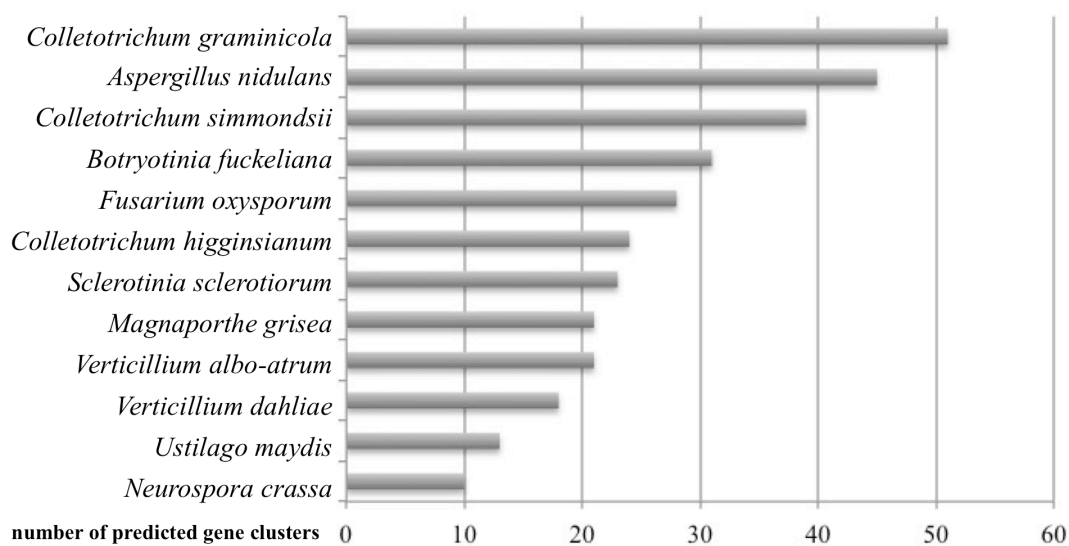
Figure 6.9. Bar diagram showing the number of secondary metabolites related gene clusters identified by SMURF across twelve fungal genomes showing different behavior.

In order to predict secondary metabolites gene clusters we used the web-based prediction tool SMURF. Thirty-nine clusters were identified in the genome of *C. simmondsii,* 51 in *C. graminicola* and 24 in *C. higginsianum* (Figure 6.9).

C. simmondsii clusters identified were very variable in number of genes. The biggest cluster includes 21 genes and some of them only two. Over half (63%) of the backbone genes were placed into clusters. Considering the number of genes involved in secondary metabolism (including the number of P450) in the three *Colletotrichum* genomes, it appears that the numbers of predicted clusters are strongly influenced by the quality of genome assembly (e.g. in *C. higginsianum* gene clusters may be fragmented in more supercontigs). Therefore, the number of clusters in *C. higginsianum* and *C. simmondsii* are most likely underestimated by SMURF. However, the number of clusters in *C. graminicola* may also be underestimated since it included neither the melanin nor the carotenoid biosynthesis gene clusters

(O'Connell *et al.*, 2012). Like for the genes encoding putative carbohydrate-active enzymes, *C. simmondsii* appears to have a wide arsenal of genes involved in secondary metabolism (Figure 6.10). Based on our analyses it's the second filamentous fungus, after *C. higginsianum*, with the largest set of SM genes.
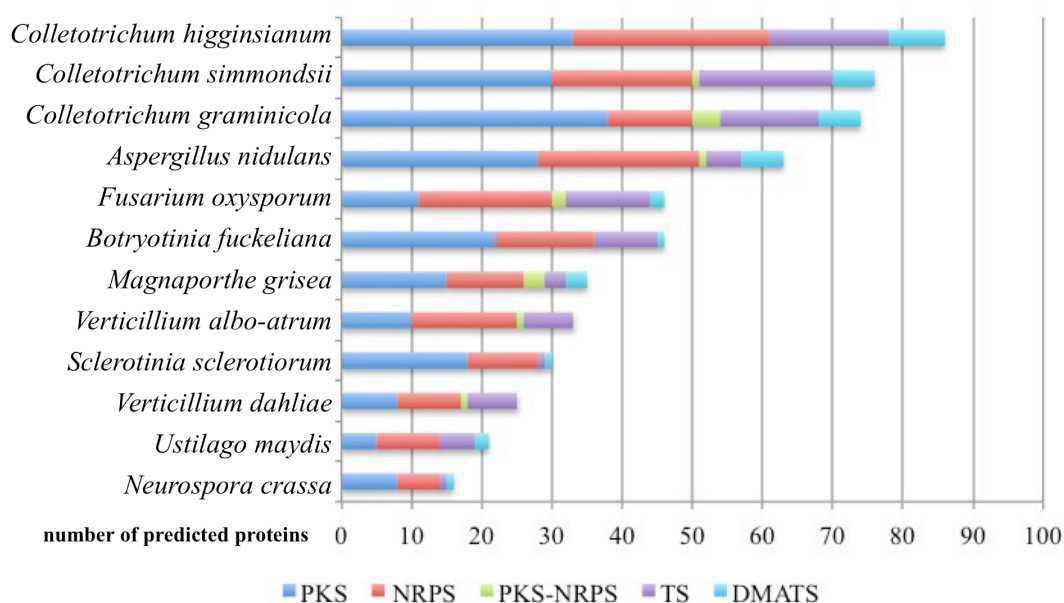
Figure 6.10. Bar diagram showing the number of specific SM genes identified by SMURF across 12 fungal genomes showing different behavior. Classes analyzed have been: polyketide synthases (PKS), peptide synthases (NRPS), PKS-NRPS hybrids, dimethylallyl tryptophan synthases (DMATS) and terpene synthases (TS)

Across the gene families analyzed *C. simmondsii* shows the highest number of genes involved in terpenes production. There are 19 putative TS genes in this organism, followed by *C. higginsianum* with 18 and 14 in *C. graminicola*. All these putative genes are characterized by the terpenoid synthase (IPR008949) conserved domain. However, this set of genes shows different PFAM domain patterns in the three genomes. Four of *C. simmondsii* and *C. graminicola* and nine of *C. higginsianum* genes are characterized by a conserved domain (PFAM ID: PF00348)

characteristic of polyprenyl synthetase enzymes that catalyze a 1'4-condensation between five carbon isoprene units. Three genes in each genome are carrying domain typical of squalene and phytoene synthase (PFAM ID: PF00494). *C. simmondsii* shows an expansion on genes characterized by TRI5 domains (PFAM ID: PF06330) encoding by enzyme trichodiene synthase, which has been shown to catalyse the first step of the trichothecene pathways (Trapp *et al.*, 1998) with a total of three genes against only one in the other two species.

Another gene expansion appears on genes encoding a terpene synthase family, metal binding domain (PFAM ID: PF03936), six genes are present in *C. simmondsii*, three in *C. higginsianum* and four in *C. graminicola*. Furthermore, three, two and one genes respectively of *C. simmondsii*, *C. higginsianum* and *C. graminicola* did not show any conserved PFAM domain. Polyketide synthases genes contain three essential modues: acyl transferase (AT); acyl carrier (ACP); and ketoacyl CoA synthase (KS). And, they can contain up to three optional modules encoding ketoreductase (KR), dehydratase (DH) or enoyl reductase (ER) domains.

Based on this evidence, all genes identified by SMURF were manually checked for presence or absence of PKS specific domains (Figure 6.11). All *Colletotrichum* genomes analyzed show a very high number of PKS encoding genes. However, *C. simmondsii* seems to encode fewer PKS genes compared to the other two species; 30 genes were identified in this genome, 33 in *C. higginsianum* and 38 in *C. graminicola*. Due to the vast variety of biological process involving secondary metabolites, and the lack of biological evidence of the role of these genes, it is not easy to draw conclusions from this data and further functional analyses are required.

Contig	Gene ID	Conserved domains	SMURF Cluster
NODE 3773	g10207.t1		
NODE 3885	g10298.t1		Cluster1
NODE 12523	g1069.t1		Cluster3
NODE 13476	g1566.t1		Cluster12
NODE 14976	g1927.t1		Cluster13
NODE 14976	g1939.t1		Cluster13
NODE 15460	g2074.t1		Cluster14
NODE 15636	g2153.t1		Cluster15
NODE 18266	g3623.t1		Cluster17
NODE 19933	g4641.t1		Cluster18
NODE 19933	g4648.t1		Cluster18
NODE 20461	g4982.t1		Cluster20
NODE 21374	g5358.t1		Cluster21
NODE 21780	g5457.t1		Cluster22
NODE 22626	g5685.t1		Cluster23
NODE 22626	g5688.t1		Cluster23
NODE 23405	g6001.t1		Cluster26
NODE 11710	g662.t1		
NODE 25855	g6729.t1		Cluster27
NODE 27466	g7991.t1		
NODE 27466	g8003.t1		
NODE 2884	g8576.t1		Cluster33
NODE 2884	g8581.t1		Cluster33
NODE 29543	g8818.t1		Cluster34
NODE 29996	g8933.t1		Cluster35
NODE 12157	g922.t1		Cluster36
NODE 32059	g9522.t1		Cluster38
NODE 33067	g9824.t1		Cluster39
NODE 18684	g4001.t1		
NODE 26166	g7027.t1		Cluster28

PF00501	AMP-binding enzyme
PF00698	Acyl transferase domain
PF00550	Phosphopantetheine attachment site
PS50075	Acyl carrier protein-like
PF00109	Beta-ketoacyl synthase, N-terminal domain
PF02801	Beta-ketoacyl-ACP synthase
PF08659	Beta-keto reductase domain
SM00826	Polyketide synthase, dehydratase domain
SM00829	Polyketide synthase, enoylreductase
PF00975	Thioesterase domain
PF08242	Methyltransferase domain
PF00668	Condensation domain
PF13745	HxxPF-repeated domain

Figure 6.11. Summary of protein structures in the PKS genes identified by SMURF in the *Colletotrichum simmondsii* genome. Results have been obtained using InterProScan. Conserved domains identified are shown with different colors; functions are reported in the legend (bottom left). For each gene, contig (first column) and SM cluster resulting from SMURF (last column) have been reported.

Dimethylallyl tryptophan synthases (DMATS) are a set of genes involved alkaloids biosynthesis starting from the tryptophan molecule *via* the indole pathway. There are six putative DMATS genes in *C. simmondsii*, seven in *C. graminicola* and eleven in *C. higginsianum*. Also in this case, like for PKS genes, all *Colletotrichum* genomes analyzed show a very high number of DMATS compared to the other fungi analyzed, this evidence could reflect a higher number of indole alkaloids in *Colletotrichum* species.

Nonribosomal peptides are a class of peptide secondary metabolites synthesized by nonribosomal peptide synthetases (NRPS), which are independent of messenger RNA. Each NRPS can synthesize only one type of peptide. Each NRPS contains several conserved domains including an AMP-binding domain, a peptide carrier domain with attached phosphopantetheine, and a condensation domain forming the amide bond. SMURF identified 12 putative NRPS encoded by *C. graminicola*, 20 by *C. simmondsii* and 28 by *C. higginsianum*. In this case, putative genes were manually checked for presence or absence of specific domains (Figure 6.12). Most of the genes identified by SMURF did not show evidence related to a condensation domain. As already reported by O'Connell *et al* (2012), NRPS genes are often fragmented by gene prediction software due to their large size and modular nature. However, a manual annotation of these genes could be the best approach in order to avoid mistakes. The same principle could be applied for PKS-NRPS hybrids. SMURF was able to identify one hybrid gene encoded by *C. simmondsii* genome and four by *C. graminicola*; none were identified in *C. higginsianum*. Also in this case further analyses in order to manually predict and annotate this class of genes are needed before final conclusions can be reached.

Contig	Gene ID	Conserved domains				SMURF Cluster
NODE_33783	g10022.t1					
NODE_4354	g10507.t1					Cluster2
NODE_4354	g10515.t1					Cluster2
NODE_5198	g10950.t1					Cluster5
NODE_12523	g1099.t1					Cluster6
NODE_12574	g1124.t1					Cluster7
NODE_6581	g11290.t1					Cluster8
NODE_9551	g13172.t1					Cluster9
NODE_18144	g3326.t1					Cluster16
NODE_18175	g3555.t1					
NODE_20285	g4840.t1					Cluster19
NODE_21374	g5357.t1					Cluster21
NODE_23139	g5814.t1					
NODE_11615	g585.t1					Cluster25
NODE_26436	g7224.t1					Cluster29
NODE_27234	g7911.t1					Cluster30
NODE_28536	g8542.t1					Cluster32
NODE_12157	g914.t1					
NODE_33067	g9795.t1					
NODE_3315	g9845.t1					

PF00501	AMP-binding enzyme
PF00550	Phosphopantetheine attachment site
PS50075	Acyl carrier protein-like
PF00668	Condensation domain

Figure 6.12. Pictures showing the structures of NRPS genes identified by SMURF in the *Colletotrichum simmondsii* genome. Results have been obtained using InterProScan. Conserved domains identified are shown with different colors; functions are reported in the legend (bottom left). For each gene, contig (first column) and SM cluster resulting from SMURF (last column) have been reported.

Due to the importance of genes encoding Cytochrome P450 monooxygenase in modifying secondary metabolites and the evidence that they are often included in SM gene cluster we searched for these genes across *C. simmondsii* genome. Like for other classes *C. simmondsii* (with *C. higginsianum*) shows the highest expansion of this class (Figure 6.13). Furthermore, with a total of 230 genes encoding Cytochrome P450 monooxygenase, the two *Colletotrichum* species are the ascomycetes with the

highest number of representative genes with this function

(http://p450.riceblast.snu.ac.kr/species.php).
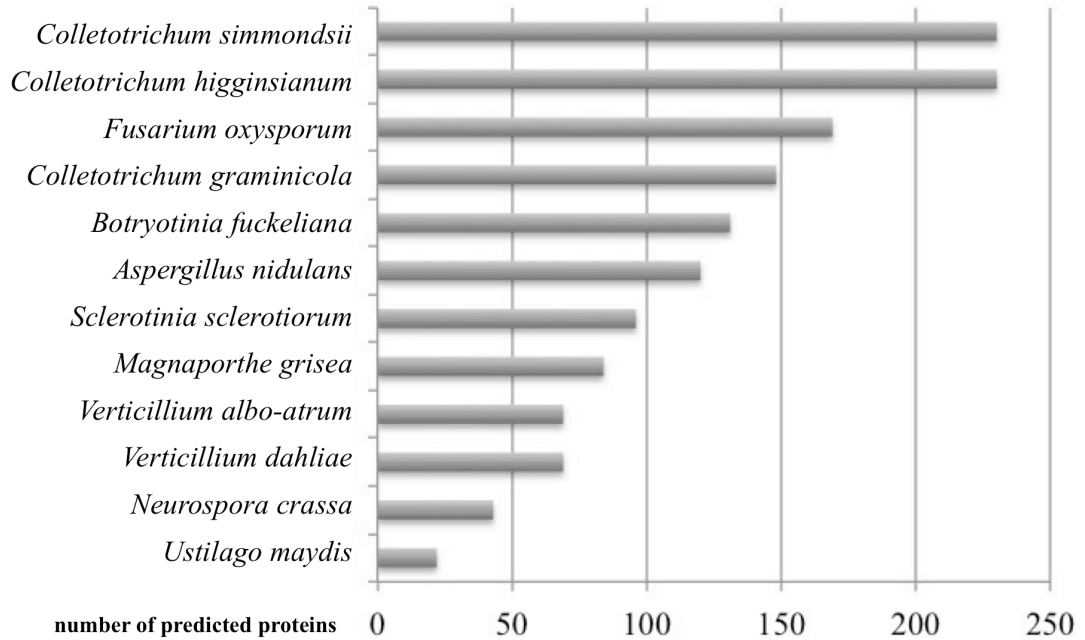
Figure 6.13. Bars chart showing the number genes encoding Cytochrome P450 monooxygenase manually identified across twelve fungal genomes compared in this study.

6.4.6.3 <u>Secretome and candidate secreted effectors proteins (CSEPs)</u>

In order to identify *C. simmondsii* secretome we scanned all proteins predicted in the genome with WoLF-PSORT (http://wolfpsort.org/). WoLF-PSORT was chosen for analysis based on results reported by O'Connell et al (2012), which report it as the most accurate tools able to identify correctly the highest number of secreted proteins. 2182 *C. simmondsii* predicted proteins were identified as secreted. The secretome in this organism represent the 16.21% of the entire proteome. The proportion on proteins secreted is much bigger compared to the other *Colletotrichum* organisms: *C. graminicola* have a set of 1650 secreted proteins over a total of 12006 (13.74%) and 2142 proteins are secreted in *C. higginsianum* over 16150 (13.26%). Therefore the number of secreted proteins in *C. simmondsii* is higher compared to the other *Colletotrichum* sequenced (Figure 6.14).

Candidate secreted effector proteins (CSEPs) were identified as secreted proteins with no significant BLAST evidence to sequences in the UniProt databases using an e-value threshold of 1e-3 and excluding sequence belonging the genus *Colletotrichum* (O'Connell *et al.*, 2012; Spanu *et al.*, 2010). Same set were used to find *C. simmondsii* species specific CSEPs based on no-BLAST sequence similarity with the other *Colletotrichum* proteins.
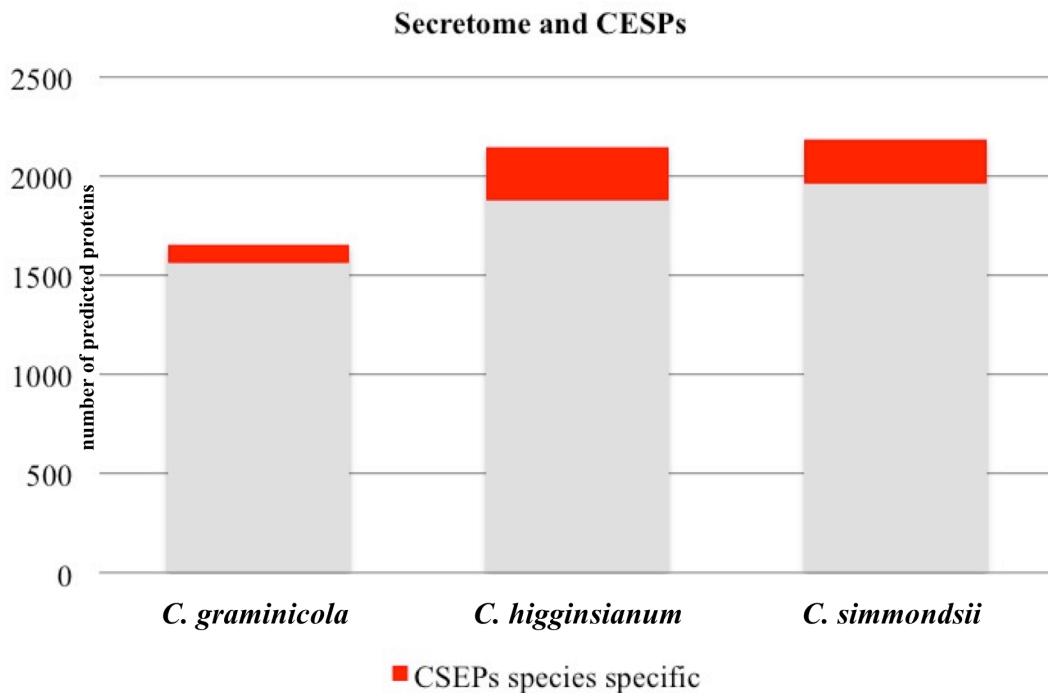
Figure 6.14. Bars diagram showing the number secreted proteins (full bars) and candidate secreted effector proteins (red portions) identified in the three *Colletotrichum* genomes.

Putative CSEPs predicted share properties that are coherent with known fungal protein effectors. They are small proteins, with an average length of 224 amino acids, compared to the average sizes of the proteome that is 498 amino acids. They are cysteine-rich 2.0%, with CSEPs less than 200 amino acids in length having approximately three times (3.8%) the cysteine content of the whole proteomes (1.2%) and with CSEPs less than 100 amino acids in length having 6%.

A total of 224 CSEPs specific for *C. simmondsii* were identified in the genome; this number was comparable to the one identified in the *C. higginsianum* genome (264) but was much higher compared to the CSEPs predicted in *C. graminicola* (84).

6.5 <u>DISCUSSION</u>

This study has generated the first example of a complete genome sequence from a member of the *C. acutatum* species complex. An isolate of *C. simmondsii* from this complex was used to assemble the genome for comparison with previous public genomes representing *C. higginsianum* and *C. graminicola*. The choice of the isolate was based on biological evidence such as evolutionary relationships and host range as well as scientific relevance and impact of the organism.

Whole-genome sequencing was carried out using Illumina GAII 76bp (X2) mate-pair reads. To carry out the assembly on Illumina data we tested different software such as SOAP, ABYSS, IDBA and VELVET, in order to get the best performance and to provide a useful tool for further *Colletotrichum* genome sequencing. We were able to rich a good assembly (50.5 Mb on 1,129 supercontigs) using the combination of NGS approach based on Illumina technology with an average coverage of 40x and VELVET. This result is in contrast with the hypothesis that the *C. higginsianum* assembly, having been assembled from short-read data only, is more fragmented than that of *C. graminicola* (O'Connell *et al.*, 2012). The observed fragmentation could be related to the repetitive nature of the genome itself or to genomic rearrangement.

The number of protein-coding genes in *C. simmondsii* (13,549) was similar to the number reported in *C. graminicola* (12,006) and smaller than *C. higginsianum* (16,172)(*Colletotrichum* Sequencing Project, Broad Institute of Harvard and MIT - http://www.broadinstitute.org/). Genome comparison shows that 94% of putative proteins have BLAST sequence similarity to proteins inside the genus *Colletotrichum* and 6% are unique to *C. simmondsii*.

An interesting observation is the expansion of the sub-family of glycoside hydrolases that show a α-L-arabinofuranosidases function. L-arabinosyl residues are commonly distributed as side chains in cellulose, hemicelluloses, lignin and pectins. The presence of these residues restricts the enzymatic hydrolysis of hemicelluloses and pectins (Rahman *et al.*, 2003). The α-L-arabinofuranosidases (α-L-AFases) are accessory enzymes that cleave α-L-arabinofuranosidic linkages and act synergistically with other hemicellulases and pectic enzymes for the complete hydrolysis of hemicelluloses and pectins (Margolles-Clark *et al.*, 1996).

Another class widely expanded in *C. simmondsii* is the GH78, the new genome encode eleven genes compared to *C. graminicola* and *C. higginsianum* that encode four and six genes, respectively. The main function of GH78 class is α-L-rhamnosidase and these enzymes catalyse the hydrolysis of terminal non-reducing α-L-rhamnose residues in α-L-rhamnosides. L-Rhamnose is a common constituent of glycolipids and glycosides, such as plant pigments, pectic polysaccharides and gums. The largest set of this class of genes, which potentially degrade the plant cell wall (van den Brink *et al.*, 2011) and modify the fungal cell wall, belongs to *C. simmondsii* genome. This may be related to the range of host of this organism, which has been suggested by differences reported within other genera of Sordariomycete fungi. For example, *Fusarium oxysporum* has a wide host range and has a total of 660 carbohydrate-active enzyme genes, whereas *F. graminearum* is a specialized cereal pathogen (wheat and barley) and has 487 of these genes (data from O'Connell *et al.*, 2012). Similarly, *Verticillium dahliae* is a vascular wilt pathogen in many plant species and has a set of 547 genes belonging to this family; whereas *V. albo-atrum* has a limited host range and has 338 of these genes (Bhat and Subbarao, 1999).

C. simmondsii, like the other *Colletotrichum* species, encode significantly more enzymes involved in secondary metabolism than other ascomycetes. In particular genes encoding terpenoid synthase are widely expanded in *C. simmondsii*. Nineteen putative TS genes were predicted in this genome, compared with 18 in *C. higginsianum* and 14 in *C. graminicola* (O'Connell *et al.*, 2012). Despite the fact that terpenoid products have drawn significant interest in different organisms (Rynkiewicz *et al.*, 2001), the complexity of secondary metabolites resulting from the enzymatic reaction of TS proteins still needs further investigation before final conclusions can be reached about their functional role in plant-*Colletotrichum* interactions. *C. simmondsii* shows an expansion on TS genes and in particular those characterized by TRI5 domains (encoding trichodiene synthase), with a total of three genes against only one in the other two species. These genes encode for sesquiterpene cyclase that catalyzes the formation of trichodiene in the biosynthesis of antibiotics and mycotoxins (Rynkiewicz *et al.*, 2001); and have been shown to have a role in pathogenicity in *Gibberella pulicaris* on parsnip and of *G. zeae* on wheat (Desjardins *et al*. 1993; Proctor *et al*. 1995).

These predicted proteins from *C. simmondsii* were also screened for predicted cellular localization in order to identify genes in the fungal secretome. *C. simmondsii* compared to the other *Colletotrichum* species appear capable of more secretable proteins, with 16.2% of the entire proteome containing signal peptide for secretion outside the cell, compared with less than 14% in *C. graminicola* (13.7%) and *C. higginsianum* (13.3%). *C. simmondsii* appears to have 224 genes that encode candidate effectors proteins (CEPs) with no similarity to proteins from other fungal genera (O'Connell *et al.*, 2012: Spanu *et al.*, 2010). A similar number of CEPs was found in *C. higginsianum* (264), and much fewer in *C. graminicola* (85). This

translates to 1.7% of the proteome as candidate effectors in *C. simmondsii,* 1,6% in *C. higginsianum* and 0.7% in *C. graminicola*.

In conclusion, there have been important expansions of gene families in *C. simmondsii* including examples that encode carbohydrate-active enzymes with specific functions, proteins involved in secondary metabolites pathways, and candidate effector proteins that could be associated to wider range of hosts.

CHAPTER 7

Final conclusions and future prospects

The overall aim of this project was to gain an improved understanding of the evolutionary relationships in *Colletotrichum acutatum sensu lato* pathosystems. This aim was addressed through a set of defined objectives. Based on the results several conclusions can be drawn also providing future perspectives to advance this research area.

The first part of the study (Chapter 3) focused on gaining an overall understanding of the genetic diversity reported within Colletotrichum acutatum *sensu lato* populations in relation to their biological diversity and geographic distribution. A database of more than 800 rRNA sequences deposited on GenBank and related information about these strains such as host, geographic area, etc. was built based on ITS sequence similarity (CaITSdb - *Colletotrichum acutatum* **ITS d**atab*a*se). This approach aimed to use the data available in the public domain in order to gain a wider view of population structure of the taxa. Evolutionary network of more than 800 rRNA sequences deposited in GenBank provided an initial assessment of the genetic diversity of sub-populations and their evolutionary relationships. These results confirmed the sub-division of these taxa into several genetics groups. However, using a larger set of data, clustering of the strains appeared less compact showing a higher complexity and variability compared to previous work based on restricted set of isolates related to specific hosts or geographic areas. Results obtained and strain information have been analysed in order to investigate host association patterns and phylogeographic relationships. This has revealed that *C. acutatum sensu lato* strains have been associated with infection on more than 90 genera of plants (either crops or plants in natural ecosystems, monocotyledons and dicotyledons), two insect species (*Fiorinia externa* and *Orthezia praelonga*) and in few cases also with marine ecosystems. Most of the hosts do not show any strong

relationship with genetic groups, this is true especially in cultivated fruit systems such as strawberry, olive, *etc*. Nevertheless several hosts show a specific correlation with a genetic group, a sub-population and in few cases also with the ITS haplotypes suggesting a progressive trend in of specialization of some genetically related populations to a particular host. The data collected in the CaITSdb also showed a wide geographic distribution of *C. acutatum sensu lato*. Strains belonging to these taxa are present in diverse every climatic areas worldwide. Even if different geographic areas do show particular trends in population distribution we did not find any strong connection between genetic groups or populations and their distribution. However results obtained suggest Oceania as possible origin of this pathogen; this region showed the highest level of variability and groups closely related to a hypothetical ancestral population are mainly distributed in these countries.

Next stage in this study focused on the genetic variability of *C. acutatum sensu lato* global populations through multi-locus sequencing to evaluate the speciation process (Chapter 4). Based on the results obtained in Chapter 3, a collection of 120 isolates representative of the global populations of this organism was assembled. The set of isolates has been chosen based on host association, geographic distribution, mating behaviour and genetic diversity. Cultural studies carried out on the subset of isolates (colony aspect and growth rates) highlighted the unreliability of these classic mycological methods to accurately distinguish and identify *C. acutatum sensu lato* populations, particularly as these characters can vary change following sub-culturing or prolonged storage. However, a multi-locus sequencing approach has proved useful and the genetic groups, sub-populations and unique haplotypes identified were congruent with previous results. All isolates able to produce perithecia belonged to the same phylogenetic cluster previously

characterized as *Glomerella miyabeana* [syn. *G. salicis*]. On the other hand all self-sterile, cross-fertile isolates capable of heterothallic mating capability (based on Gueber and Correll 2001) were represented in two different genetic groups A3 and A5 corresponding to recently proposed species *C. fioriniae* and *C. acutatum sensu stricto*. Although heterothallic mating has been demonstrated and described in culture, no biological or genetic evidence that this process occurs in nature has been reported. This strong relationship between mating behavior and phylogenetic clustering, and the lack of evidence of mating cross in nature, fit with the hypothesis of an on-going speciation process even in those cases where the strains are hypothetically able of exchange genetic material through sexuality. Based on this approach we were able to distinguish a number of genetic groups, sub-populations and unique haplotypes and several of these correspond to some of the proposed species such as *C. phormii, G. miyabeana, C. lupini* and *C. carthami*. Furthermore, some groups also showed a particular trend in host specialisation; for example isolates belonging to "ex A8" group seem to be specific to *Cyphomandra betacea* and the same situation is reflected by some of the other genetic clusters. However the level of host-specificity of these groups needs further analysis including cross infection assays. Based on these results, the general evolutionary trend emerging in *C. acutatum* species complex appears to be distinct populations undergoing clear changes in their host-association pattern. . This also suggests a role for heterothallism in host adaptation particularly considering isolates characterized as A3 and A5 (corresponding to *C. fioriniae* and *C. acutatum sensu strict*) and described as cross fertile. The occurrence of heterothallism seems to influence the host range diversity. In contrast, isolates capable of homothallic (self-fertile) seems to have a narrow range of hosts. On the other end, Populations evolutionarily distant

to a hypothetical ancestral population seem to have lost any kind of sexual reproduction capability and this event seems associated with a host preference/specificity.   In conclusion, the results suggest that the capacity of populations to exchange genetic information could lead to an increase in the host spectrum of this pathogen. This data viewed along with the recent research on the re-assessment of *Colletotrichum* taxonomy suggests that *C. acutatum sensu lato* encompasses several cryptic species including at least six *combinatio nova* species.

Next phase of the project focused on investigating the complexity of *C. acutatum sensu lato* populations associated with the UK strawberry production systems.   Generally, strawberry seem to be susceptible to almost all genetic groups (corresponding to a number of species being proposed) within the *C. acutatum* species complex. A set of 67 *C. acutatum sensu lato* strains related to strawberry production in the United Kingdom has been used in this study. Genetic characterization based on multi-locus sequencing showed the occurrence of three distinct genetic groups (A2, A3 and A4) associated with strawberry in the UK and these correspond to   *C. nymphaeae*, *C. fioriniae* and *C. godetiae*. Pathogenicity tests revealed a strong correlation between these *C. acutatum sensu lato* populations and aggressiveness. Specifically, genetic groups corresponding to *C. nymphaeae* and *C. fioriniae* appeared to be the most aggressive, followed by genetic types corresponding to *C. acutatum sensu stricto* and *C. godetiae*. Also, isolates belonging to the genetic group corresponding to *C. simmondsii* seem capable of infecting strawberry plants. All other genetic groups corresponding to the other species: *C. lupini*, *C. phormii*, *C. salicis* and the two undescribed *Colletotrihcum spp.* even if capable of infecting strawberry fruits, seems to be much less aggressive compared to the others and are unlikely to be associated with natural infections currently. Isolates

belonging to the same genetic group (or a taxonomic entity) originating from a particular host are aggressive on that host. These results show that even in susceptible hosts such as strawberry, *C. acutatum sensu lato* species are characterized by host-preference patterns. Furthermore, host specificity of *C. acutatum* complex on strawberry could reflect an ongoing specialisation process, where adaptation to a host representing a specific ecological niche could overlap with the speciation process.

The last part of the project has focused on generating and analysing the first example of a complete genome sequence from a member of the *C. acutatum* species complex. An isolate belonging to the genetic group (A9) corresponding to *C. simmondsii* from within this species complex was used to assemble the genome for comparison with previous publicly available genomes representing *C. higginsianum* and *C. graminicola*. The choice of the isolate was based on biological evidence such as evolutionary relationships and host range as well as scientific relevance and impact of the organism. The main objective of this part of the study was the prediction of novel putative genes widely recognised as associated with pathogenicity and host specificity such as effector genes and the investigation of the expansion of gene families related to host interactions such as Carbohydrate-Active Enzymes (CAZymes) coding genes or genes involved in secondary metabolites such as toxins. This work led to a first level genome assembly using the combination of NGS approach based on Illumina technology with an average coverage of 40x and VELVET. The number of protein-coding genes in *C. simmondsii* (13,549) was similar to the number reported in *C. graminicola* (12,006) and smaller than *C. higginsianum* (16,172)(*Colletotrichum* Sequencing Project, Broad Institute of Harvard and MIT - http://www.broadinstitute.org/). Genome comparison showed

that 94% of putative proteins have BLAST sequence similarity to proteins within the genus *Colletotrichum* and 6% are unique to *C. simmondsii*. In addition to the *C. simmondsii* genome sequenced in this work, the two *Colletotrichum* genomes publicly available, and data related to 11 fungal genomes showing different types of interactions with hosts have been included in the analyses. *C. simmondsii* possesses the highest number of putative carbohydrate-active enzymes compared to the other genomes analysed. This could be related to the adaptation of this pathogen to diverse hosts. This hypothesis is reinforced by differences in other Sordariomycete fungi such as *Fusarium* and *Verticillium*. *F. oxysporum* has a wide host range and has a total of 660 carbohydrate-active enzyme genes, *F. graminearum* is a specialized cereal pathogen (wheat and barley) and has 487 of these genes (data from O'Connel *et al.*, 2012). *V. dahliae*, is a vascular wilt pathogen of many plant species and has a set of 547 genes belonging to this family; whereas *V. albo-atrum* has a limited host range and has 338 of these genes (Bhat and Subbarao, 1999). Furthermore, *C. simmondsii*, like the other *Colletotrichum* species, contains significantly more number of genes encoding enzymes involved in secondary metabolism than other ascomycetes. In particular, genes encoding terpenoid synthase were widely expanded in *C. simmondsii* and in particular the one involved in encoding trichodiene synthase enzyme that catalyzes the formation of trichodiene in the biosynthesis of metabolites such as antibiotics and mycotoxins (Rynkiewicz *et al.*, 2001). Similarly, *C. simmondsii* appears to have 224 genes that encode for CEPs, proportionally more than the other two *Colletotrichum* genomes; this translates to 1.7% of the proteome as candidate effectors in *C. simmondsii,* 1,6% in *C. higginsianum* and 0.7% in *C. graminicola*. In conclusion, the expansions in gene families such as those encoding carbohydrate-active enzymes, secondary metabolites pathways and CEPs could be

associated with the wider range of hosts. This data provides a range of new resources such as the genome sequence data and predicted genes and would serve as an useful platform for further research in the field.

Overall, the diversity studies mainly focused on the evolutionary relationships in *C. acutatum species complex*, using a holistic approach of bioinformatics, molecular biology, biology and pathology to successfully investigate and integrate genetic diversity data with other key attributes such as the origin, spread, host association patterns and speciation processes in this pathogen. Genome analyses study led to the sequencing, assembly and annotation of the first genome representing (*C. simmondsii*) the C. acutatum sensu lato complex. This has revealed various interesting feature including novel genes and expansion of gene families associated pathogenicity and host interaction. The new knowledge generated and the resources developed in this study are likely to provide impetus for comparative and functional genomics studies in *C. acutatum sensu lato*. Advances in next generation sequencing methods and bioinformatics tools for genome analyses provide significant scope to advance our knowledge of the genetic basis of host-microbe interactions and pathogen evolution in these *Colletotrichum* species. Comparative genomics of several strains representing the diverse life-styles such as mating behavior and host range of *C. acutatum sensu lato* could be a successful approach in order to identify and link genome wide differences to evolutionary adaptations.

The results and resources from this study have led to collaborative links with various European groups such as:

- **Gene and metabolite discovery of *Colletotrichum acutatum* - strawberry interaction.** Daniel Buchvaldt Amby, Thomas Sundelin and Birgit Jensen.

University of Copenhagen, Faculty of Life Science, Department of Plant Biology and Biotechnology. Denmark

- **Pathogenicity and epidemiology of *Colletotrichum carthami*: an emerging pathogen in Italy**. Sabrina Sarrocco, Antonio Zapparata and Giovanni Vannacci. University of Pisa, Department of Sciences of Agriculture, Food and the Agricultural Environment, Italy – reflecting the future direction of research developments in this area.

REFERENCES

Afanador-Kafuri L, Minz D, Maymon M, *et al*. 2003. Characterization of *Colletotrichum* isolates from tamarillo, passiflora, and mango in Colombia and identification of a unique species from the genus. *Phytopathology* 93:579-587.

Agostini JP, Timmer LW and Mitchell DJ. 1992. Morphological and pathological characteristics of strains of *Colletotrichum gloeosporioides* from citrus. *Phytopathology* 82:1377-1382.

Ahn JH, Cheng YQ and Walton JD. 2002. An extended physical map of the TOX2 locus of *Cochliobolus carbonum* required for biosynthesis of HC-toxin. *Fungal Genetics and Biology* 35(1):31-38.

Akagi Y, Akamatsu H, Otani H, *et al*. 2009. Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. *Eukaryot Cell* 8(11):1732-8.

Alahakoon PW, Brown AE and Sreenivasaprasad S. 1994. Cross infection potential of genetic groups of *Colletotrichum gloeosporioides* on tropical fruits. *Physiological and Molecular Plant Pathology* 44:93-103.

Baxter L, Tripathy S, Ishaque N, *et al.* 2010. Signatures of adaptation to obligate biotrophy in the *Hyaloperonospora arabidopsidis* genome. *Science.* 10;330(6010):1549-51

Bandelt HJ, Forster P and Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1):37-48.

Beech MG and Simpson DW. 1989. Strawberry production in the United Kingdom. *ISHS Acta Horticulturae* 265:693-696. http://www.actahort.org/books/265/265_107.htm

Bergmann S, Schümann J, Scherlach K, *et al*. 2007. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nature Chemical Biology* 3(4):213-7.

Brooker NL, Leslie JF and Dickman MB. 1991. Nitrate non-utilizing mutants of *Colletotrichum* and their use in studies of vegetative compatibility and genetic relatedness. *Phytopathology* 81(6):672-677.

Buddie AG, Martinez-Culebras P, Bridge PD, *et al*. 1999. Molecular characterization of *Colletotrichum* strains derived from strawberry. *Mycological Research* 103:385-394.

Calleja EJ, Ilbery B, Spence NJ, *et al*. 2012. The effectiveness of phytosanitary controls in preventing the entry of *Colletotrichum acutatum* in the UK strawberry sector. *Plant Pathology* DOI: 10.1111/j.1365-3059.2012.02647.x

Cannon PF, Bridge PD and Monte E. 2000. *Colletotrichum*: Host Specificity, Pathology, and Host-Pathogen Interaction. (eds. D. Prusky, S. Freeman and M.B. Dickman). APS Press, St Paul, Minnesota, USA: 1-20.

Cantarel BL, Coutinho PM, Rancurel C, *et al*. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* 37:D233-238

Castro MLG, da Silva Lacaz C, Guarro J, *et al*. 2001. Phaeohyphomycotic cyst caused by *Colletotrichum crassipes*. *Journal of Clinical Microbiology* 39: 2321–2324.

Chacko RJ, Weidemann GJ, Tebeest DO, *et al*. 1994. The use of vegetative compatibility and heterokaryosis to determine potential asexual gene exchange in *Colletotrichum gloeosporioides*. *Biological Control* 4(4):382-389.

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

Cisar CR and TeBeest DO. 1999. Mating system of the filamentous ascomycete, *Glomerella cingulata*. *Current Genetics* 35(2):127-133.

Clements FE and Shear CL. 1909. Genera of Fungi. Hafner Publishing Co Ltd HW Wilson Co., Minneapolis, USA.

Conesa A, Götz S, García-Gómez JM, *et al*. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 15;21(18):3674-6.

Crouch JA, Tredway LP, Clarke BB, *et al*. 2009. Phylogenetic and population genetic divergence correspond with habitat for the pathogen *Colletotrichum cereale* and allied taxa across diverse grass communities. *Molecular Ecology* 18(1):123-135.

Damm U, Cannon P, Woudenberg J, Johnston P, Weir B, Tan Y, Shivas R and P Crous. 2012. The *Colletotrichum boninense* species complex 73:1-36.

De Hoog GS, Guarro J, Gene J, *et al*. 2000. Atlas of clinical fungi. Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands.

De Lorenzo G, D'Ovidio R and Cervone F. 2001. The role of polygalacturonase-inhibiting proteins (PGIPs) in defense against pathogenic fungi. *Annual Review of Phytopathology* 39:313-35.

Dean R, Van Kan JAl, Pretorius ZA, *et al*. 2012. The Top 10 fungal pathogens in molecular plant pathology. *Molecular Plant Pathology* 13(4): 414–430.

Denoyes-Rothan B, Guerin G, Delye C, Smith B, Minz D, Maymon M and Freeman S. 2003. Genetic diversity and pathogenic variability among isolates of *Colletotrichum* species from strawberry. *Phytopathology* 93(2):219-28.

Desjardins AE, Hohn TM and McCormick SP. 1993. Trichothecene biosynthesis in *Fusarium* species: chemistry, genetics, and significance. *Microbiolology and Molecular Biology Reviews* 57(3):595-604.

Du MZ, Schardl CL, Nuckles EM, *et al*. 2005. Using mating-type gene sequences for improved phylogenetic resolution of *Colletotrichum* species complexes. *Mycologia* 97(3):641-658.

Duke MM. 1928. The genera *Vermicularia* Fr. and *Colletotrichum*. *Transactions of the British Mycological Society* 13:156-184.

Eastburn DM and Gubler WD. 1990. Strawberry anthracnose: Detection and survival of *Colletotrichum acutatum* in soil. *Plant Disease* 74:161-163.

Epstein L, Lusnak K and Kaur S. 1998. Transformation-mediated developmental mutants of *Glomerella graminicola* (*Colletotrichum graminicola*). *Fungal Genetics and Biology* 23(2):189-203.

Faedda R, Agosteo GE, Schena L, *et al*. 2011. *Colletotrichum clavatum sp. nov*. identified as the causal agent of olive anthracnose in Italy. *Phytopathologia Mediterranea* 50(2): 283-302.

Farr DF, Aime MC, Rossman AY, *et al*. 2006. Species of *Colletotrichum* on *Agavaceae*. *Mycological Research* 110(12):1395-408.

Fernandez V, Dursun D, Miller D, *et al*. 2002. *Colletotrichum* keratitis. *American Journal of Ophthalmology* 134(3):435–438.

Finlay B. 2002. Global dispersal of free-living microbial eukaryote species. *Science* 296:1061–1063.

Flor HH. 1947. Inheritance of reaction to rust in flax. *Journal of Agricultural Research* 74: 241–262.

Forster H and Adaskaveg JE. 1999. Identification of subpopulations of *Colletotrichum acutatum* and epidemiology of almond anthracnose in California. *Phytopathology* 89(11):1056-1065.

Freeman S and Katan T. 1997. Identification of *Colletotrichum* Species Responsible for Anthracnose and Root Necrosis of Strawberry in Israel. *Phytopathology* 87(5):516-21.

Freeman S and Shabi E. 1996. Cross-infection of subtropical and temperate fruits by *Colletotrichum* species from various hosts. *Physiological and Molecular Plant Pathology* 49:395-404.

Freeman S, Horowitz S and Sharon A. 2001. Pathogenic and non pathogenic lifestyles in *Colletotrichum acutatum* from strawberry and other plants. *Phytopathology* 91:986-992. B

Freeman S, Katan T and Shabi E. 1998. Characterization of *Colletotrichum* species responsible for anthracnose diseases of various fruits. *Plant Disease* 82(6):596-605.

Freeman S, Minz D, Maymon M, *et al*. 2001. Genetic diversity within *Colletotrichum acutatum sensu* Simmonds. *Phytopathology* 91(6):586-592.

Garrido C, Carbu M, Fernandez-Acero FJ, *et al*. 2008. Isolation and pathogenicity of *Colletotrichum spp*. causing anthracnose of strawberry in south west Spain. *European Journal of Plant Pathology* 120(4):409-415.

Giraud T, Refregier G, Le Gac M, *et al*. 2008. Speciation in fungi. *Fungal Genetics and Biology* 45(6):791-802.

Gnerre S, Lander ES, Lindblad-Toh K, Jaffe DB. 2009. Assisted assembly: how to improve a de novo genome assembly by using related species. *Genome Biology* 10(8):R88.

Guarro J, Svidzinski TE, Zaror L, *et al*. 1998. Subcutaneous hyalohyphomycosis caused by *Colletotrichum gloeosporioides*. *Journal of Clinical Microbiology* 36:3060–3065.

Guerber JC and Correll JC. 2001. Characterization of *Glomerella acutata*, the teleomorph of *Colletotrichum* acutatum. *Mycologia* 93:216-219.

Guerber JC, Liu B, Correll JC, *et al*. 2003. Characterization of diversity in *Colletotrichum acutatum* sensu lato by sequence analysis of two gene introns, mtDNA and intron RFLPs, and mating compatibility. *Mycologia* 95(5):872-895.

Guindon S and Gascuel O. 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology* 52(5):696–704.

Hernández-Silva L, Piñón-Escobedo C, Cano-Camacho H, *et al*. 2007. Comparison of fungal growth and production of extracellular pectin lyase activity by pathogenic and non-pathogenic races of *Colletotrichum lindemuthianum* cultivated under different conditions. *Physiological and Molecular Plant Pathology* 70:88-95.

Horton P, Park KJ, Obayashi T, *et al*. 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Research 35:585-7

Howard CM and Albregts EE. 1984. Anthracnose. In: Compendium of Strawberry Diseases. American Phytopathological Society: St Paul, ed. JL Maas USA: 85–87.

Howard CM, Maas JL, Chandler CK, *et al*. 1992. Anthracnose of strawberry caused by the *Colletotrichum* complex in Florida. *Plant Disease* 76(10):976-981.

Jelev ZJ, Bobev SG, Minz D, *et al*. 2008. Characterization of *Colletotrichum* Species Causing Strawberry Anthracnose in Bulgaria. *Journal of Phytopathology* 156:668–677.

Johnson RD, Johnson L, Itoh Y, *et al*. 2000. Cloning and characterization of a cyclic peptide synthetase gene from *Alternaria alternate* apple pathotype whose product

is involved in AM-toxin synthesis and pathogenicity. *Molecular Plant-Microbe Interactions* 13(7):742-53.

Jones AL, Ehret GR, Meyer MP, *et al*. 1996. Occurrence of bitter rot on apple in Michigan. *Plant Disease* 80:1294-1297.

Jones D and Baker R. 2007. Introductions of non-native plant pathogens into Great Britain, 1970–2004. *Plant Pathology*. 56:891–910.

Khaldi N, Seifuddin FT, Turner G, *et al*. 2010. SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal genetics and biology*. 47(9):736-741.

Katan T. Vegetative compatibility in populations of *Verticillium* - an overview. 2000. In: Advances in Verticillium Research and Disease Management (Eds.: E. Tjamos, R.C. Rowe, J.B. Heale, and D. Fravel). The American Phytopathological Society, St. Paul, MN, USA: 69–86.

Kondrashov A. 1986. Sympatric speciation: when is it possible? *Biological Journal of the Linnean Society* 27:201–223.

Kubicek CP, Herrera-Estrella A, Seidl-Seiboth V, *et al*. 2011. Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma. Genome biology*, 12:R40.

Lara-Márquez A, Zavala-Páramo MG, López-Romero E, *et al*. 2011. Cloning and characterization of a pectin lyase gene from *Colletotrichum lindemuthianum* and comparative phylogenetic/structural analyses with genes from phytopathogenic and saprophytic/opportunistic microorganisms. *BMC Microbiology* 9:11-260.

Lardner R, Johnston PR, Plummer KM, *et al*. 1999. Morphological and molecular analysis of *Colletotrichum acutatum* sensu lato. *Mycological Research* 103:275-285.

Lee SC, Ni M, Li W, *et al*. 2010. The Evolution of Sex: a Perspective from the Fungal Kingdom. *Microbiology and Molecular Biology Reviews*, 74(2):298-340.

LoBuglio KF and Pfister DH. 2008. A *Glomerella* species phylogenetically related to *Colletotrichum acutatum* on Norway maple in Massachusetts. *Mycologia* 100(5):710-715.

Lubbe CM, Denman S, Cannon PF, *et al*. 2004. Characterization of *Colletotrichum* species associated with diseases of Proteaceae. *Mycologia* 96(6):1268-1279.

Ma LJ, van der Does HC, Borkovich KA, *et al*. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464:367-373.

Maas JL and Howard CM. 1985. Variation of several anthracnose fungi in virulence to strawberry and apple. *Plant Disease* 69:164-166.

Manire CA, Rhinehart HL, Sutton DA, *et al*. 2002. Disseminated mycotic infection caused by *Colletotrichum acutatum* in a Kemp's ridley sea turtle (*Lepidochelys kempi*). *Journal of Clinical Microbiology* 40(11):4273-80.

Marcelino J, Giordano R, Gouli S, *et al*. 2008. *Colletotrichum acutatum var. fioriniae* (teleomorph : *Glomerella acutata var. fioriniae var. nov.*) infection of a scale insect. *Mycologia* 100(3):353-374.

Martínez-Culebras PV, Barrio E, García MD, *et al*. 2000. Identification of *Colletotrichum* species responsible for anthracnose of strawberry based on the internal transcribed spacers of the ribosomal region. *FEMS Microbiol Letters* 189(1):97-101.

Mayr E. 1963.Animal Species and Evolution. Harvard University Press, Cambridge, MA, USA.

Menat J, Cabral AL, Vijayan P, *et al*. 2012. *Glomerella truncata*: another *Glomerella* species with an atypical mating system. *Mycologia*. 104(3):641-9.

Mertely JC and Peres NA. 2005. Root Necrosis of Strawberries Caused by *Colletotrichum acutatum.* In: vol. PP-211. Plant Pathology Department, Florida Cooperative Extension Service, Institute of Food and Agricultural, USA.

Milholland RD. Anthracnose fruit rot (ripe rot). 1995. In: *Compendium of Blueberry and Cranberry Diseases.*. L. Caruso and D. C. Ramsdell, eds. American Phytopathological Society, St. Paul, MN, USA.

Nair J and Corbin JB. 1981. Histopathology of *Pinus radiata* seedlings infected by *Colletotrichum acutatum* f. sp. *pinea. Phytopathology* 71(8):777-783.

Nei M and Kumar S. 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York, USA.

Nirenberg HI, Feiler U and Hagedorn G. 2002. Description of *Colletotrichum lupini comb. nov* in modern terms. *Mycologia* 94(2):307-320.

O'Connell RJ, Thon MR, Hacquard S, *et al*. 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature genetics* 44(9):1060-1065.

O'Donnell K, Cigelnik E and Nirenberg HI. 1998. Molecular systematics and phylogeography of the *Gibberella fujikuroi* species complex. *Mycologia* 90(3):465-493.

O'Quinn RP, Hoffman JL and Boyd AS. 2000. *Colletotrichum* species as emerging opportunistic fungal pathogens: a report of 3 cases of phaeohyphomycosis and review. *Journal of the American Academy of Dermatology* 45:56–61.

Oliver R. Genomic tillage and the harvest of fungal phytopathogens. 2012. *New Phytologist* doi: 10.1111/j.1469-8137.2012.04330.x

Orr GF, Kuehn HH and Plunkett OA. 1963. The genus *Myxotrichum* Kunze. *Canadian Journal of Botany* 41:1457-1480.

Panaccione DG. 1993. The fungal genus *Cochliobolus* and toxin-mediated plant disease. *Trends in Microbiology* 1(1):14-20.

Peng Y, Leung HCM, Yiu SM, et al. 2010. IDBA – A Practical Iterative de Bruijn Graph De Novo Assembler. Research in Computational Molecular Biology - Lecture Notes in Computer Science 6044:426-440.

Peres NA, Timmer LW, Adaskaveg JE, *et al*. 2005. Life styles of *Colletotrichum acutatum*. *Plant Disease* 89:784- 796.

Proctor RH, Hohn TM and McCormick SP. 1995. Reduced virulence of *Gibberella zeae* caused by disruption of a trichothecene toxin biosynthetic gene. . *Molecular Plant-Microbe Interactions* 8(4):593-601.

Quevillon E, Silventoinen V, Pillai S, *et al*. 2005. InterProScan: protein domains identifier. *Nucleic Acids Research* 33:116-120

Redman RS and Rodriguez RJ. 1994. Factors affecting the efficient transformation of *Colletotrichum* species. *Experimental Mycology* 18(3):230-246.

Rodriguez-Guerra R, Ramirez-Rueda MT, Cabral-Enciso M, *et al*. 2005. Heterothallic mating observed between Mexican isolates of *Glomerella lindemuthiana*. *Mycologia* 97:793–803.

Ronquist F and Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19(12):1572-1574.

Rynkiewicz MJ, Cane DE and Christianson DW. 2001. Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. *Proceedings of the National Academy of Sciences USA* 98(24):13543-13548.

Sambrook J, Fritsch EF and Maniatis T. 1989. Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press New York, USA :1659.

Schafer C and Wostemeyer J. 1992. Random primer dependent PCR differentiates aggressive from non-aggressive isolates of the oilseed rape pathogen *Phoma lingam* (*Leptosphaeria maculans*). *Journal of Phytopathol*ogy 136:124-136.

Scott PM. 2012. Recent research on fumonisins: a review. *Food Additives & Contaminants: Part A* 29(2):242-248.

Sepp R, Szabo I, Uda H, *et al*. 1994. Rapid techniques for DNA extraction from routinely processed archival tissue for use in PCR. *Journal of Clinical Pathology* 47(4):318-323.

Shi YL, Correll JC and Guerber JC. 1996. Frequency of *Colletotrichum* species causing bitter rot of apple in the south-eastern United States. *Plant Disease* 80:692-696.

Shivas RG and Tan YP. 2009. A taxonomic re- assessment of *Colletotrichum acutatum*, introducing *C. fioriniae comb. et stat. nov*. and *C. simmondsii sp. nov*. *Fungal Diversity* 39: 111-122.

Simmonds JH. 1965. A study of the species of *Colletotrichum* causing ripe fruit rots in Queensland. *Queensland Journal of Agricultural and Animal Science* 22:437-459.

Simmonds JH. 1968. Type specimens of *Colletotrichum gloeosporioides var. minor* and *Colletotrichum acutatum. Queensland Journal of Agricultural and Animal Science* 25:178a.

Simpson DW, Winterbottom CQ, Bell JA, *et al*. 1994. Resistance to a single UK isolate of *Colletotrichum acutatum* in strawberry germplasm from Northern Europe. *Euphytica* 77:161-164.

Simpson JT, Wong K, Jackman SD, *et al*. 2009. ABySS: a parallel assembler for short read sequence data. *Genome research* 19(6):1117-1123.

Smith BJ and Black LL. 1986. First report of *Colletotrichum acutatum* on strawberry in the United States. *Plant Disease* 70:1074.

Smith BJ. 1998. Compendium of Strawberry Diseases. J. L. Mass, ed. American Phytopathological Society, St. Paul, MN, USA.

Spanu PD, Abbott JC, Amselem J, *et al*. 2010. Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* 330(6010):1543-1546.

Sreenivasaprasad S and Talhinhas P. 2005. Genotypic and phenotypic diversity in *Colletotrichum acutatum*, a cosmopolitan pathogen causing anthracnose on a wide range of hosts. *Molecular Plant Pathology* 6(4):361-378.

Sreenivasaprasad S, Sharada K, Brown AE, *et al*. 1996. PCR-based detection of *Colletotrichum acutatum* on strawberry. *Plant Pathology* 45(4):650-655.

Staskawicz BJ, Ausubel FM, Baker BJ, *et al*. 1995. Molecular genetics of plant disease resistance. *Science* 268(5211):661-667.

Stanke M, Steinkamp R, Waack S, *et al*. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Research* 32:309-312

Staskawicz BJ, Dahlbeck D and Keen NT. 1884. Cloned avirulence gene of *Pseudomonas syringae* pv. *glycinea* determines race-specific incompatibility on *Glycine max* (L.) Merr. *Proceedings of the National Academy of Sciences USA* 81(19):6024-6028.

Sutton BC. 1980. The Coelomycetes: Fungi Imperfecti With Pycnidia Acervuli and Stromata: Kew, Surrey, Commonwealth Mycological Institute, UK.

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

Talgø V, Aamot HU, Strømeng GM, *et al*. 2007. *Glomerella acutata* on Highbush Blueberry (*Vaccinium corymbosum* L.) in Norway. *Plant Management Network* doi:10.1094/PHP-2007-0509-01-RS.

Talhinhas P, Sreenivasaprasad S, Neves-Martins J, *et al*. 2002. Genetic and morphological characterization of *Colletotrichum acutatum* causing anthracnose of lupins. *Phytopathology* 92(9):986-996.

Talhinhas P, Sreenivasaprasad S, Neves-Martins J, *et al*. 2005. Molecular and phenotypic analyses reveal association of diverse *Colletotrichum acutatum* groups and a low level of *C. gloeosporioides* with olive anthracnose. *Applied and Environmental Microbiology* 71(6):2987-2998.

Tamura K, Peterson D, Peterson N, *et al*. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28(10):2731-2739.

Tanaka A, Shiotani H, Yamamoto M, *et al*. 1999. Insertional mutagenesis and cloning of the genes required for biosynthesis of the host-specific AK-toxin in the Japanese pear pathotype of *Alternaria alternata*. *Molecular Plant Pathology* 12(8):691-702.

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

Taylor JW, Turner E, Townsend JP, *et al*. 2006. Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom fungi. *Philosophical Transactions of the Royal Society B: Biological Sciences. B* 361:1947–1963.

ten Have A, Mulder W, Visser J, *et al*. 1998. The endopolygalacturonase gene Bcpg1 is required for full virulence of *Botrytis cinerea*. *Molecular Plant-Microbe Interactions* 11(10):1009-1016.

ten Have A, Tenberge KB, Benen JAE, *et al*. 2002. The contribution of the cell wall degrading enzymes to pathogenesis of fungal plant pathogens. *The Mycota*

XI:341–358.

Ter-Hovhannisyan V, Lomsadze A, Chernoff Y, *et al*. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research* 18(12):1979-1990.

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

Tonukari NJ, Scott-Craig JS and Walton JD. 2000. *The Cochliobolus carbonum* SNF1 gene is required for cell wall-degrading enzyme expression and virulence on maize. *Plant Cell* 12:237-247.

Trapp SC, Hohn TM, McCormick S, *et al*. 1998. Characterization of the gene cluster for biosynthesis of macrocyclic trichothecenes in *Myrothecium roridum*. *Molecular & General Genetics* 257:421-432

Tsuge T, Harimoto Y, Akimitsu K, *et al*. 2012. Host-selective toxins produced by the plant pathogenic fungus *Alternaria alternata*. *FEMS Microbiology Reviews* doi: 10.1111/j.1574-6976.2012.00350.x

Ureña-Padilla AR, MacKenzie SJ, Bowen BW, *et al*. 2002. Etiology and population genetics *of Colletotrichum spp.* causing crown and fruit rot of strawberry. *Phytopathology* 92:1245–1252.

Ureña-Padilla AR, Mitchell DJ and Legard DE. 2001. Oversummer survival of inoculum for *Colletotrichum* crown rot in buried strawberry crown tissue. *Plant Disease* 85:750–754.

van der Does HC and Rep M. 2007. Virulence genes and the evolution of host specificity in plant-pathogenic fungi. *Molecular Plant-Microbe Interactions* 20(10):1175-82.

Van Hemelrijck W, Debode J, Heungens K, *et al*. 2010. Phenotypic and genetic characterization of *Colletotrichum* isolates from Belgian strawberry fields. *Plant Pathology* 59(5):53–861.

van Kan JA, van den Ackerveken GF and de Wit PJ. 1991. Cloning and characterization of cDNA of avirulence gene avr9 of the fungal pathogen

Cladosporium fulvum, causal agent of tomato leaf mold. *Molecular Plant-Microbe Interactions* 4(1):52-59.

Vinnere O, Fatehi J, Wright SAI, *et al*. 2002. The causal agent of anthracnose of Rhododendron in Sweden and Latvia. *Mycological Research* 106:60-69.

von Arx JA. 1957. Die Arten der Gattung *Colletotrichum*. *Phytopathologische Zeitrift* 29:413-468.

Walton JD. 1987. Two enzymes involved in biosynthesis of the host-selective phytotoxin HC-toxin. *Proceedings of the National Academy of Sciences USA* 84:8444–8447.

Walton JD. 1996. Host-selective toxins: agents of compatibility. *Plant Cell* 8(10):1723-33.

Wei Y, Shih J, Li J, *et al*. 2002. Two pectin lyase genes, *pnl*-1 and *pnl*-2, from *Colletotrichum gloeosporioides* f. sp. *malvae* differ in a cellulose-binding domain and in their expression during infection of *Malva pusilla*. Microbiology 148(Pt 7):2149-57.

Wharton PS and Dieguez-Uribeondo J. 2004. The biology of *Colletotrichum acutatum*. *Anales del Jardՙın Botanico de Madrid* 61:3–22.

Wheeler HE. 1954. Genetics and evolution of heterothallism in *Glomerella*. *In. Phytopathology*, 44:342–345.

Wheeler HE. 1954. Genetics and evolution of heterothallism in *Glomerella*. *Phytopathology* 44:342–345.

Whitelaw-Weckert MA, Curtin SJ, Huang R, Steel CC, *et al*. 2007. Phylogenetic relationships and pathogenicity of *Colletotrichum acutatum* isolates from grape in subtropical Australia. *Plant Pathology* 56(3):448-463.

Xue M, Yang J, Li Z, *et al*. 2012. Comparative analysis of the genomes of two field isolates of the rice blast fungus *Magnaporthe oryzae*. *PLoS Genetics* doi:10.1371/journal.pgen.1002869.

Yamamoto N, Matsumoto T and Ishibashi Y. 2001. Fungal keratitis caused by *Colletotrichum gloeosporioides*. *Cornea* 20:902-903.

Zerbino DR. 2010. Using the Velvet de novo assembler for short-read sequencing technologies. Curr Protoc Bioinformatics. Chapter 11:Unit 11.5.

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

	Genus	species	strain	Accession N	Host	Location	Year
1	Glomerella	acutata	MAFF 306282	AB042300.1	*Fragaria x ananassa*	Japan	
2	Glomerella	acutata	MAFF 306406	AB042301.1	*Eriobotrya japonica*	Japan	
3	Colletotrichum	carthami	MAFF 239370	AB042306.1	*Carthamus tinctorius*	Japan (Yamagata)	
4	Colletotrichum	carthami	MAFF 239362	AB042307.1	*Chrysanthemum coronarium*	Japan (Chiba)	
5	Colletotrichum	carthami	MAFF 239355	AB042312.1	*Calendula officinalis*	Japan (Chiba)	
6	Glomerella	acutata	MAFF305138	AB219020.1	*Prunus persica*	Japan	
7	Glomerella	acutata	MAFF305596	AB219021.1	*Eriobotrya japonica*	Japan, Chiba	
8	Glomerella	acutata	MAFF410044	AB219022.1	*Prunus salicina*	Japan, Nagano	
9	Glomerella	acutata	GC-AK-1	AB219023.1	*Akèbia spp.*	Japan, Nagano	
10	Glomerella	acutata	AL-5	AB219024.1	*Fragaria x ananassa*	Japan, Nagasaki	
11	Glomerella	acutata	AL-9	AB219025.1	*Fragaria x ananassa*	Japan, Nagasaki	
12	Glomerella	acutata	FPeCG-9301	AB219026.1	*Prunus persica*	Japan, Fukuoka	
13	Glomerella	acutata	GC-M-1	AB219027.1	*Castanea crenata*	Japan, Nagano	
14	Glomerella	acutata	GC-PR-8	AB219028.1	*Acacia spp.*	Japan, Nagano	
15	Glomerella	acutata	GC-B-1	AB219029.1	*Vaccinium spp.* (blueberry)	Japan, Nagano	
16	Glomerella	acutata	Na91-016	AB219030.1	*Fragaria x ananassa*	Japan, Tochig	
17	Glomerella	acutata	GC-P-38	AB219031.1	*Prunus domesticus*	Japan, Nagano	
18	Glomerella	acutata	C1-40	AB219032.1	*Vitis vinifera*	Japan, Akita	
19	Glomerella	acutata	C1-19	AB219033.1	*Vitis vinifera*	Japan, Akita	
20	Glomerella	acutata	GC1-2	AB219034.1	*Pyrus pyrifolia*	Japan, Akita	
21	Glomerella	acutata	H-44	AB219035.1	*Malus domestica*	Japan, Akita	
22	Glomerella	acutata	T-22	AB219036.1	*Malus domestica*	Japan, Akita	
23	Glomerella	acutata	B-23	AB219037.1	*Vitis vinifera*	Japan, Akita	
24	Glomerella	acutata	GC1-1	AB219038.1	*Pyrus pyrifolia*	Japan, Akita	
25	Glomerella	acutata	YH1-9	AB219039.1	*Vitis vinifera*	Japan, Akita	
26	Glomerella	acutata	M-82	AB219040.1	*Malus domestica*	Japan, Akita	
27	Glomerella	acutata	G5-1-7	AB219041.1	*Vitis vinifera*	Japan, Akita	
28	Glomerella	acutata	MAFF 306630	AB233341.1	*Malus pumila Mill. var. Domestica*	Japan, Yamagata	2002
29	Glomerella	acutata	MAFF 306611	AB233344.1	*Vaccinium corymbosum*	Japan, Ibaraki, Tsukuba	1997
30	Glomerella	acutata	MAFF 306650	AB233347.1	*Vaccinium corymbosum*	Japan, Yamagata	2002
31	Glomerella	acutata	MAFF 306651	AB233348.1	*Prunus domestica*	Japan, Yamagata	2002
32	Glomerella	acutata	MAFF 306677	AB233349.1	*Prunus salicina*	Japan, Yamagata	2002
33	Glomerella	acutata	MAFF 306673	AB269935.1	*Vaccinium corymbosum*	Japan, Ibaraki	2002

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

34	Glomerella	acutata	MAFF 306682	AB269940.1	*Fragaria x ananassa Buch*	Japan, Fukushima	2002
35	Glomerella	acutata	MAFF 306647	AB269942.1	*Fragaria x ananassa Buch*	Japan, Fukushima	2001
36	Glomerella	acutata	MAFF 306610	AB269943.1	*Vaccinium corymbosum*	Japan, Ibaraki	1999
37	Glomerella	acutata	MAFF 306609	AB269944.1	*Vaccinium corymbosum*	Japan, Ibaraki	2000
38	Glomerella	acutata	MAFF 306648	AB269945.1	*Vaccinium corymbosum*	Japan, Ibaraki	2002
39	Glomerella	acutata	CAB03	AB273192.1	*Vitis vinifera*		
40	Glomerella	acutata	CAP01	AB273193.1	*Diospyros sp.*		
41	Glomerella	acutata	GCP17	AB273194.1	*Malus domestica*		
42	Glomerella	acutata	00G74	AB273195.1	*Prunus domesticus*		
43	Glomerella	acutata	B-1	AB305160.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
44	Glomerella	acutata	B-2	AB305161.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
45	Glomerella	acutata	B-3	AB305162.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
46	Glomerella	acutata	Op-2	AB305163.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
47	Glomerella	acutata	Op-3	AB305164.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
48	Glomerella	acutata	T-1	AB305165.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
49	Glomerella	acutata	QL74	AB369503.1	*Olea europea*	Japan, Kagawa, Shyodoshima island	
50	Glomerella	acutata	Ya499	AB443950.1	*Cotinus coggygria*	Japan, Yamagata, Shonai	2007
51	Glomerella	acutata	Ya543	AB444085.1	*Matthiola incana*	Japan, Yamagata, Sakata	2007
52	Glomerella	acutata	MAFF 240289	AB458663.1	*Sanguisorba officinalis*	Japan, Aomori	2006
53	Glomerella	acutata	MAFF 240237	AB458666.1	*Prunus sp. (fallen leaf of a cherry tree)*	Japan, Hokkaido	
54	Glomerella	acutata	MAFF 240192	AB458671.1	*Origanum vulgare*	Japan, Ibaraki	2006
55	Colletotrichum	gloeosporioides	TS08-97-1	AB470867.1	*Quercus liaotungensis*	China, Gansu, Tianshui	2008
56	Glomerella	acutata	P-1	AB548282.1	*Pirus communis*	Iran, Guilan province	
57	Glomerella	acutata	GL-118-9411	AF081292.1	*Olea europea*	Spain, Tarragona	
58	Colletotrichum	acutatum	IMI345028	AF090853.1	*Fragaria x ananassa*	Colombia	
59	Colletotrichum	sp.	ALM-KSH-10	AF207791.1	*Prunus dulcis*	Israel, North	
60	Colletotrichum	acutatum	U.S.-ALM-4	AF207793.1	*Prunus dulcis*	USA, California	
61	Colletotrichum	acutatum	TUT-5954	AF207794.1	*Fragaria x ananassa*	Israel	1994
62	Glomerella	acutata	ANE NL12A	AF272781.1	*Anemone coronaria*	Netherlands	1989
63	Glomerella	acutata	ANE HV83C	AF272782.1	*Anemone coronaria*	Israel	1979
64	Glomerella	acutata	IMI223120	AF272783.1	*Anemone coronaria*	Australia	1978
65	Glomerella	acutata	STR3	AF272784.1	*Fragaria x ananassa*	USA, Florida	
66	Glomerella	acutata	IMI348494	AF272785.1	*Fragaria x ananassa*	France	
67	Glomerella	acutata	PCN5	AF272786.1	*Carya illinoinensis*	USA, Alabama	
68	Glomerella	acutata	APL2	AF272787.1	*Malus domestica*	USA, South Carolina	
69	Glomerella	acutata	PCH8	AF272788.1	*Prunus persica*	USA, South Carolina	
70	Glomerella	acutata	IMI345026	AF272789.1	*Fragaria x ananassa*	Spain	

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

71	Glomerella	acutata	G1	AF411697.1	*Rhododendron ponticum*	Sweden, Gothenburg
72	Glomerella	acutata	G2	AF411698.1	*Rhododendron ponticum*	Sweden, Gothenburg
73	Glomerella	acutata	G4	AF411699.1	*Rhododendron ponticum*	Sweden, Gothenburg
74	Glomerella	acutata	IMI 117617	AF411700.1	*Carica papaya*	Australia, Ormiston
75	Glomerella	acutata	IMI 117619	AF411701.1	*Carica papaya*	Australia, Ormiston
76	Glomerella	acutata	L1	AF411702.1	*Rhododendron x catawbiense*	*Latvia, Babite*
77	Glomerella	acutata	L2	AF411703.1	*Rhododendron x catawbiense*	*Latvia, Babite*
78	Glomerella	acutata	L3	AF411704.1	*Rhododendron x catawbiense*	*Latvia, Babite*
79	Glomerella	acutata	L4	AF411705.1	*Rhododendron x catawbiense*	*Latvia, Babite*
80	Glomerella	acutata	L5	AF411706.1	*Rhododendron x catawbiense*	*Latvia, Babite*
81	Glomerella	acutata	L6	AF411707.1	*Rhododendron x catawbiense*	*Latvia, Babite*
82	Glomerella	acutata	P1	AF411708.1	*Rhododendron x catawbiense*	*Sweden, Pålsjö*
83	Glomerella	acutata	S1	AF411709.1	*Rhododendron x catawbiense*	Sweden, Sofiero
84	Glomerella	acutata	S10	AF411710.1	*Rhododendron x orbiculare*	Sweden, Sofiero
85	Glomerella	acutata	S11	AF411711.1	*Rhododendron calophytum*	Sweden, Sofiero
86	Glomerella	acutata	S12	AF411712.1	*Rhododendron insigne*	Sweden, Sofiero
87	Glomerella	acutata	S13	AF411713.1	*Rhododendron x catawbiense*	Sweden, Sofiero
88	Glomerella	acutata	S14	AF411714.1	*Rhododendron brachycarpum*	Sweden, Sofiero
89	Glomerella	acutata	S16	AF411715.1	*Rhododendron japonicum*	Sweden, Sofiero
90	Glomerella	acutata	S17	AF411716.1	*Rhododendron luteum*	Sweden, Sofiero
91	Glomerella	acutata	S18	AF411717.1	*Rhododendron spp.*	Sweden, Sofiero
92	Glomerella	acutata	S19	AF411718.1	*Rhododendron degronianum*	Sweden, Sofiero
93	Glomerella	acutata	S2	AF411719.1	*Rhododendron x orbiculare*	Sweden, Sofiero
94	Glomerella	acutata	S20	AF411720.1	*Rhododendron brachycarpum*	Sweden, Sofiero
95	Glomerella	acutata	S21	AF411721.1	*Rhododendron x orbiculare*	Sweden, Sofiero
96	Glomerella	acutata	S22	AF411722.1	*Rhododendron luteum*	Sweden, Sofiero
97	Glomerella	acutata	S23	AF411723.1	*Rhododendron luteum*	Sweden, Sofiero
98	Glomerella	acutata	S24	AF411724.1	*Rhododendron brachycarpum*	Sweden, Sofiero
99	Glomerella	acutata	S25	AF411725.1	*Rhododendron degronianum*	Sweden, Sofiero
100	Glomerella	acutata	S3	AF411726.1	*Rhododendron x orbiculare*	Sweden, Sofiero
101	Glomerella	acutata	S4	AF411727.1	*Rhododendron japonicum*	Sweden, Sofiero
102	Glomerella	acutata	S5	AF411728.1	*Rhododendron x catawbiense*	Sweden, Sofiero
103	Glomerella	acutata	S6	AF411729.1	*Rhododendron spp.*	Sweden, Sofiero
104	Glomerella	acutata	S7	AF411730.1	*Rhododendron calophytum*	Sweden, Sofiero
105	Glomerella	acutata	S8	AF411731.1	*Rhododendron japonicum*	Sweden, Sofiero
106	Glomerella	acutata	S9	AF411732.1	*Rhododendron japonicum*	Sweden, Sofiero
107	Glomerella	acutata	397	AF411765.1	*Fragaria x ananassa*	USA

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

108	Glomerella	acutata	NI90	AF411766.1	*Fragaria x ananassa*	UK, Northern Ireland	
109	Glomerella	acutata	Clemson SF21	AF411768.1	*Prunus persica*		
110	Glomerella	acutata	Nantana A	AF411772.1	*Vitis vinifera*	USA	
111	Glomerella	acutata	IMI 383015	AF488778.1	*Hevea brasiliensis*		
112	Glomerella	acutata	120V.2II	AF489556.1	*Rubus spp.* (raspberry)	USA	
113	Glomerella	acutata	Coll.15a	AF489557.1	*Fragaria x ananassa*	Switzerland	
114	Glomerella	acutata	Cooley2	AF489558.1	*Fragaria x ananassa*		
115	Glomerella	acutata	Nantana A1	AF489559.1	*Vitis spp.*	USA	
116	Glomerella	acutata	Tomato ARK	AF489560.1	*Lycopersicon esculentum*	USA	
117	Glomerella	acutata	F3e	AF489561.1	*Fragaria x ananassa*	France	
118	Glomerella	acutata	1267b	AF489562.1	*Fragaria x ananassa*	France	
119	Glomerella	acutata	Coll.14a	AF489563.1	*Rubus spp.* (raspberry)	Switzerland	
120	Glomerella	acutata	Cha-s	AF489564.1	*Fragaria x ananassa*	France	
121	Glomerella	acutata	BJS Tomato	AF489565.1	*Lycopersicon esculentum*		
122	Glomerella	acutata	F3c	AF489566.1	*Fragaria x ananassa*	France	
123	Glomerella	acutata	Myrtille.a	AF489567.1	*Vaccinium myrtillus*	France	
124	unclassified	unclassified	its358	AF502861.1	*Fagus sylvatica*	Germany	
125	Glomerella	acutata	TOM-21	AF521196.1	*Cyphomandra betacea*	Colombia, Antioquia, San Vicente	
126	Glomerella	acutata	TOM-9	AF521205.1	*Cyphomandra betacea*	Colombia, Antioquia, Rio Negro	1998
127	Glomerella	acutata	TOM-12	AF521210.1	*Cyphomandra betacea*	Colombia, Antioquia, Santa Rosa	
128	Glomerella	acutata	M 7	AF534467.1	*Coffea arabica*	Colombia	
129	Colletotrichum	sp.	JG05	AJ300557.1	*Ceanothus spp.*	France, Paris	
130	Colletotrichum	sp.	C2897	AJ300558.1	*Fragaria x ananassa*	Australia, West Australia	
131	Colletotrichum	sp.	CR20	AJ300558.1	*Fragaria x ananassa*	Portugal	1998
132	Colletotrichum	sp.	HO01	AJ300558.1	*Lupinus albus*	Portugal, Azores Is.	
133	Colletotrichum	sp.	JR03	AJ300558.1	*Lupinus albus*	Portugal, Azores Is., Terceira	
134	Colletotrichum	sp.	70354	AJ300561.1	*Lupinus luteus*	Germany	
135	Colletotrichum	sp.	70399	AJ300561.1	*Lupinus albus*	Austria	
136	Colletotrichum	sp.	70555	AJ300561.1	*Lupinus albus*	Chile	
137	Colletotrichum	sp.	96A4	AJ300561.1	*Lupinus albus*	Australia, West Australia	
138	Colletotrichum	sp.	96A649	AJ300561.1	*Lupinus spp.*	Australia, West Australia	
139	Colletotrichum	sp.	C3	AJ300561.1	*Lupinus luteus*	Poland	1997
140	Colletotrichum	sp.	CMG12	AJ300561.1	*Cinnamonium zeylanicum*	Portugal, Lisbon	1996
141	Colletotrichum	sp.	CR02	AJ300561.1	*Lupinus albus*	Portugal, Montemor-o-Velho	
142	Colletotrichum	sp.	CSL 1179	AJ300561.1	*Lupinus albus*	Germany	1998
143	Colletotrichum	sp.	CSL 1294	AJ300561.1	*Lupinus x polyphyllus*	UK, York	1998
144	Colletotrichum	sp.	G52	AJ300561.1	*Lupinus albus*	Germany, Rastatt	1995

145	Colletotrichum	sp.	IMI350308	AJ300561.1	*Lupinus spp.*		1991
146	Colletotrichum	sp.	JR11	AJ300561.1	*Lupinus albus*	Portugal, Azores Is.	
147	Colletotrichum	sp.	JR15	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Terceira	1997
148	Colletotrichum	sp.	JR16	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., S. Jorge	1997
149	Colletotrichum	sp.	JR17	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Faial	1997
150	Colletotrichum	sp.	KH48	AJ300561.1	*Lupinus x polyphyllus*	UK, York	1998
151	Colletotrichum	sp.	KH49	AJ300561.1	*Lupinus albus*	Germany	1999
152	Colletotrichum	sp.	PT22	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Terceira	1999
153	Colletotrichum	sp.	PT23	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Terceira	2000
154	Colletotrichum	sp.	PT24	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Pico, Lajes	2001
155	Colletotrichum	sp.	PT25	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Faial	2002
156	Colletotrichum	sp.	PT26	AJ300561.1	*Lupinus luteus*	Portugal, Azores Is., Faial	2003
157	Colletotrichum	sp.	PT27	AJ300561.1	*Lupinus luteus*	Portugal, Azores Is., Faial	2004
158	Colletotrichum	sp.	PT28	AJ300561.1	*Lupinus luteus*	Portugal, Azores Is., Faial	2005
159	Colletotrichum	sp.	PT29	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2006
160	Colletotrichum	sp.	PT30	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2007
161	Colletotrichum	sp.	PT31	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2008
162	Colletotrichum	sp.	PT32	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2009
163	Colletotrichum	sp.	PT33	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2010
164	Colletotrichum	sp.	PT34	AJ300561.1	*Lupinus luteus*	Portugal, Azores Is., São Miguel	2011
165	Colletotrichum	sp.	PT35	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2012
166	Colletotrichum	sp.	PT36	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., São Miguel	2013
167	Colletotrichum	sp.	PT37	AJ300561.1	*Lupinus albus*	Portugal, Azores Is., Santa Maria	2014
168	Colletotrichum	sp.	PT38	AJ300561.1	*Lupinus albus*	Portugal, Vagos, Sosa, Lavandeira	2015
169	Colletotrichum	sp.	PT39	AJ300561.1	*Lupinus albus*	Portugal, Vagos, Sosa, Lavandeira	2016
170	Colletotrichum	sp.	PT40	AJ300561.1	*Lupinus albus*	Portugal, Vagos, Sosa, Lavandeira	2017
171	Colletotrichum	sp.	PT41	AJ300561.1	*Lupinus albus*	Germany, Rastatt	2018
172	Colletotrichum	sp.	SHK1033	AJ300561.1	*Lupinus albus*	South Africa	
173	Colletotrichum	sp.	SHK788	AJ300561.1	*Lupinus albus*	South Africa	
174	Colletotrichum	sp.	TNOS4747	AJ300562.1	*Eriobotrya japonica*	Portugal	2000
175	Colletotrichum	sp.	TNOS4646	AJ300563.1	*Vitis vinifera*	Portugal	2000
176	Colletotrichum	acutatum	BBA 68396	AJ301905.1	*Vaccinium corymbosum*		
177	Colletotrichum	acutatum	BBA 69645	AJ301906.1	*Primula spp.*		
178	Colletotrichum	acutatum	BBA 70338	AJ301910.1	*Tulipa spp.*		
179	Colletotrichum	acutatum	BBA 70339	AJ301911.1	*Vaccinium corymbosum*	Netherlands	
180	Colletotrichum	acutatum	BBA 70341	AJ301913.1	*Fragaria spp.*		
181	Colletotrichum	acutatum	BBA 70342	AJ301914.1	*Fragaria x ananassa*	UK	

182	Colletotrichum	acutatum	BBA 70343	AJ301915.1	*Primula spp.*	Netherlands
183	Colletotrichum	lupini	BBA 70344	AJ301916.1	*Lupinus spp.*	Netherlands
184	Colletotrichum	acutatum	BBA 70345	AJ301917.1	*Prunus cerasus*	Netherlands
185	Colletotrichum	lupini	BBA 70346	AJ301918.1	*Lupinus spp.*	Netherlands
186	Colletotrichum	acutatum	BBA 70348	AJ301920.1	*Capsicum annuum*	Indonesia
187	Colletotrichum	acutatum	BBA 70349	AJ301921.1	*Capsicum annuum*	Indonesia
188	Colletotrichum	acutatum	BBA 70350	AJ301922.1	*Anemone spp.*	Netherlands
189	Colletotrichum	lupini	BBA 70352	AJ301923.1	*Lupinus albus*	Germany
190	Colletotrichum	acutatum	BBA 62124	AJ301924.1	*Coffea spp.*	
191	Glomerella	cingulata	BBA 65797	AJ301925.1	*Syringa vulgaris*	Germany
192	Colletotrichum	acutatum	BBA 67875	AJ301926.1	*Cyclamen persicum*	Germany
193	Colletotrichum	lupini	BBA 70073	AJ301927.1	*Lupinus x polyphyllus*	Germany
194	Colletotrichum	lupini	BBA 70317	AJ301928.1	*Lupinus albus*	Germany
195	Colletotrichum	lupini	BBA 63879	AJ301930.1	*Lupinus mutabilis*	Bolivia
196	Colletotrichum	gloeosporioides	BBA 67435	AJ301931.1	*Sambucus nigra*	Germany
197	Colletotrichum	acutatum	BBA 67859	AJ301932.1	*Fragaria x ananassa*	
198	Colletotrichum	lupini	BBA 70358	AJ301933.1	*Lupinus albus*	Germany
199	Colletotrichum	lupini	BBA 68334	AJ301934.1	*Lupinus spp.*	Germany
200	Colletotrichum	lupini	BBA 70385	AJ301935.1	*Lupinus angustifolius*	Germany
201	Colletotrichum	acutatum	BBA 70486	AJ301936.1	*Bergenia sp.*	
202	Colletotrichum	lupini	BBA 70884	AJ301948.1	*Lupinus albus*	Ukraine
203	Colletotrichum	acutatum	BBA 70886	AJ301949.1	*Cyclamen spp.*	
204	Colletotrichum	acutatum	BBA 67866	AJ301950.1	*Fragaria x ananassa*	Germany
205	Colletotrichum	acutatum	BBA 70093	AJ301951.1	*Fragaria spp.*	
206	Glomerella	cingulata	BBA 70991	AJ301952.1	*Salix sp.*	
207	Colletotrichum	acutatum	BBA 70820	AJ301956.1	*Hepatica acutiloba*	Germany
208	Colletotrichum	lupini	BBA 71249	AJ301959.1	*Lupinus albus*	Canada
209	Colletotrichum	acutatum	BBA 71286	AJ301963.1	*Lycopersicon spp.*	
210	Colletotrichum	acutatum	BBA 71292	AJ301964.1	*Lupinus albus*	Portugal
211	Colletotrichum	lupini	BBA 71310	AJ301968.1	*Lupinus luteus*	Poland
212	Colletotrichum	sp.	BBA 71320B	AJ301969.1	*Hordeum vulgare*	
213	Colletotrichum	acutatum	BBA 71331	AJ301971.1	*Prunus cerasus*	
214	Colletotrichum	gloeosporioides	BBA 71332	AJ301972.1	Sambucus sp."	
215	Colletotrichum	lupini	BBA 71330	AJ301975.1	*Urtica dioica*	
216	Colletotrichum	acutatum	BBA 71370	AJ301981.1	*Cyclamen spp.*	
217	Colletotrichum	acutatum	BBA 71371	AJ301982.1	*Cyclamen spp.*	
218	Colletotrichum	acutatum	BBA 71383	AJ301983.1	*Juglans spp.*	

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

219	Colletotrichum	acutatum	BBA 71427	AJ301987.1	*Camellia spp.*		
220	Colletotrichum	sp.	HY09	AJ311391.1	*Lupinus albus*	Canada	
221	Colletotrichum	acutatum	IMI 345027	AJ536199.1	*Fragaria x ananassa*	France	
222	Colletotrichum	acutatum	IMI 348160	AJ536200.1	*Fragaria x ananassa*	USA	
223	Colletotrichum	acutatum	IMI 360928	AJ536201.1	*Fragaria x ananassa*	Switzerland	
224	Colletotrichum	acutatum	IMI 351587	AJ536202.1	*Fragaria x ananassa*	USA	
225	Colletotrichum	acutatum	IMI 345030	AJ536203.1	*Fragaria x ananassa*	Costa Rica	1991
226	Colletotrichum	acutatum	IMI 299103	AJ536204.1	*Fragaria x ananassa*	UK	
227	Colletotrichum	acutatum	IMI 360086	AJ536205.1	*Fragaria x ananassa*	Japan	
228	Colletotrichum	acutatum	IMI 345033	AJ536206.1	*Fragaria x ananassa*	Australia	
229	Colletotrichum	acutatum	IMI 345034	AJ536207.1	*Fragaria x ananassa*	Australia	
230	Colletotrichum	acutatum	IMI 346326	AJ536208.1	*Fragaria x ananassa*	Australia	
231	Colletotrichum	acutatum	IMI 345026	AJ536209.1	*Fragaria x ananassa*	Spain	
232	Colletotrichum	acutatum	IMI 367466	AJ536210.1	*Fragaria vesca*	Netherlands	
233	Colletotrichum	acutatum	IMI 351255	AJ536211.1	*Fragaria x ananassa*	UK	
234	Colletotrichum	acutatum	IMI 345581	AJ536212.1	*Fragaria x ananassa*	New Zealand	
235	Colletotrichum	acutatum	IMI 345585	AJ536213.1	*Fragaria x ananassa*	New Zealand	
236	Colletotrichum	acutatum	IMI 324993	AJ536214.1	*Fragaria x ananassa*	USA	
237	Colletotrichum	acutatum	IMI 345575	AJ536215.1	*Fragaria x ananassa*	New Zealand	
238	Colletotrichum	acutatum	IMI 345577	AJ536216.1	*Fragaria x ananassa*	New Zealand	
239	Colletotrichum	acutatum	IMI 348489	AJ536217.1	*Fragaria x ananassa*	France	
240	Colletotrichum	acutatum	IMI 348494	AJ536218.1	*Fragaria x ananassa*	France	
241	Colletotrichum	acutatum	IMI 345576	AJ536219.1	*Fragaria x ananassa*	New Zealand	
242	Colletotrichum	acutatum	IMI 348499	AJ536220.1	*Fragaria x ananassa*	France	
243	Colletotrichum	acutatum	CA302a	AJ749670.1	*Nandina domestica*	UK	
244	Colletotrichum	acutatum	PD90-443	AJ749671.1	*Phlox spp.*	Netherlands	
245	Colletotrichum	acutatum	CA473	AJ749672.1	*Liriodendron tulipifera*	UK	
246	Colletotrichum	acutatum	CA287	AJ749673.1	*Statice spp.*	UK	
247	Colletotrichum	acutatum	CA546	AJ749674.1	*Lupinus spp.*		
248	Colletotrichum	acutatum	PD85-694	AJ749675.1	*Chrysanthemum spp.*	Netherlands	
249	Colletotrichum	acutatum	CA455	AJ749676.1	*Photinia spp.*	UK	
250	Colletotrichum	acutatum	CA318	AJ749677.1	*Magnolia spp.*	UK	
251	Colletotrichum	acutatum	PD89-582	AJ749678.1	*Cyclamen spp.*	Netherlands	
252	Colletotrichum	acutatum	PD88-673	AJ749679.1	*Anemone spp.*	Netherlands	
253	Colletotrichum	acutatum	JC51	AJ749680.1	*Tulipa spp.*	UK	2003
254	Colletotrichum	acutatum	PT108	AJ749681.1	*Olea europaea ssp. europaea 'Galega'*	Portugal, Vila Viçosa	2001
255	Colletotrichum	acutatum	PT135	AJ749683.1	*Olea europaea ssp. europaea*	Portugal, Elvas	2001

256	Colletotrichum	acutatum	PT166	AJ749684.1	*Olea europaea ssp. europaea*	Portugal, Vila Velha de Ródão	2002
257	Colletotrichum	acutatum	PT169	AJ749685.1	*Olea europaea ssp. europaea*	Portugal, Sabrosa, Qta. da Cavadinha	2002
258	Colletotrichum	acutatum	PT170	AJ749686.1	*Olea europaea ssp. europaea*	Portugal, Torres Vedras, Dois Portos	2002
259	Colletotrichum	acutatum	PT186	AJ749687.1	*Olea europaea ssp. europaea*	Portugal, Alter do Chão	2003
260	Colletotrichum	acutatum	CBS193.32	AJ749688.1	*Olea europaea ssp. europaea*	Italy	1932
261	Colletotrichum	acutatum	JL198	AJ749689.1	*Olea europaea ssp. europaea*	Serbia and Montenegro	2003
262	Colletotrichum	acutatum	JL199	AJ749690.1	*Olea europaea ssp. europaea*	Serbia and Montenegro	2003
263	Colletotrichum	acutatum	PT201	AJ749691.1	*Olea europaea ssp. europaea 'Galega'*	Portugal, Vila Viçosa	2003
264	Colletotrichum	acutatum	PT227	AJ749694.1	*Olea europaea ssp. europaea*	Portugal, Vila Real de Sto. António	2003
265	Colletotrichum	acutatum	PT231	AJ749695.1	*Olea europaea ssp. europaea*	Portugal, Mirandela, Bouça	2003
266	Colletotrichum	acutatum	PT232	AJ749696.1	*Olea europaea ssp. europaea*	Portugal, Valpaços	2003
267	Colletotrichum	acutatum	PT247	AJ749697.1	*Olea europaea ssp. europaea*	Portugal, Lisbon	2003
268	Colletotrichum	acutatum	PT248	AJ749698.1	*Olea europaea ssp. europaea*	Portugal, Valpaços	2003
269	Colletotrichum	acutatum	PT249	AJ749699.1	*Olea europaea ssp. europaea*	Portugal, Torres Vedras, Dois Portos	2003
270	Colletotrichum	acutatum	PT250	AJ749700.1	*Olea europaea ssp. europaea*	Portugal, Mirandela	2003
271	Glomerella	acutata	SM956	AM404275.1	*Olea europaea ssp. europaea*	Portugal, Avis, Ervedal	
272	Glomerella	acutata	SM955	AM404276.1	*Olea europaea ssp. europaea*	Portugal; Elvas	
273	Glomerella	acutata	SM954	AM404277.1	*Olea europaea ssp. europaea*	Portugal, Vila Vicosa	
274	Glomerella	acutata	SM953	AM404278.1	*Olea europaea ssp. europaea*	Portugal; Evora, Sao Mancos	
275	Glomerella	acutata	SM61	AM404279.1	*Vitis vinifera*	Portugal, Torres Vedras, Dois Portos	
276	Glomerella	acutata	SM967	AM404280.1	*Olea europaea ssp. europaea*	Portugal, Torres Vedras, Dois Portos	
277	Glomerella	acutata	SM966	AM404281.1	*Olea europaea ssp. europaea*	Portugal, Azambuja	
278	Glomerella	acutata	SM965	AM404282.1	*Olea europaea ssp. europaea*	Portugal, Azambuja	
279	Glomerella	acutata	SM60	AM404283.1	*Rubus sp.*	Portugal, Rio Maior, Assentiz	
280	Glomerella	acutata	SM963	AM404284.1	*Olea europaea ssp. europaea*	Portugal, Rio Maior, Assentiz	
281	Glomerella	acutata	SM962	AM404285.1	*Olea europaea ssp. europaea*	Portugal, Rio Maior, Assentiz	
282	Glomerella	acutata	SM961	AM404286.1	*Olea europaea ssp. europaea*	Portugal, Alcanena, Malhou	
283	Glomerella	acutata	SM59	AM404287.1	*Prunus persica*	Portugal, Tomar, Castelo do Bode	
284	Glomerella	acutata	SM959	AM404288.1	*Olea europaea ssp. europaea*	Portugal, Tomar, Castelo do Bode	
285	Glomerella	acutata	SM957	AM404289.1	*Olea europaea ssp. europaea*	Portugal, Tomar, Castelo do Bode	
286	Glomerella	acutata	PT811	AM991131.1	*Olea europea*	Portugal:Silves, Portela de Messines	2005
287	Glomerella	acutata	06-228	AM991132.1	*Olea europea*	Portugal, Loule, Andrezes	2006
288	Glomerella	acutata	PT794	AM991133.1	*Olea europea*	Portugal, Loule, Almansil	2005
289	Glomerella	acutata	07-101b	AM991134.1	*Olea europea*	Portugal, Silves	2007
290	Glomerella	acutata	PT715	AM991135.1	*Olea europea*	Portugal, Tavira	2004
291	Glomerella	acutata	06-133	AM991136.1	*Olea europea*	Portugal, Vila Nova de Foz Coa	2004
292	Glomerella	acutata	06-222	AM991137.1	*Olea europea*	Portugal, Olhao, Pereiro	2006

ANNEX 1 - *Colletotrichum acutatum* **rRNA sequences database: CaITSdb**

293	unclassified	unclassified	agrAP4244	AM992163.1	*Calluna vulgaris*	Germany
294	Glomerella	acutata	NZ 10	AY177327.1	*Persea americana*	New Zealand
295	Glomerella	acutata	NZ 18	AY177328.1	*Persea americana*	New Zealand
296	Glomerella	acutata	G2	AY266405.1	*Euphatorium thymifolia*	Thailand
297	Glomerella	acutata	STE-U 5122	AY376497.1	*Leucospermum spp.*	South Africa
298	Glomerella	acutata	STE-U 164	AY376498.1	*Pinus radiata*	South Africa
299	Glomerella	acutata	STE-U 160	AY376499.1	*Pinus radiata*	South Africa
300	Glomerella	acutata	STE-U 162	AY376500.1	*Pinus radiata*	South Africa
301	Glomerella	acutata	STE-U 4448	AY376501.1	*Leucadendron spp.*	South Africa
302	Glomerella	acutata	STE-U 4460	AY376502.1	*Protea cynaroides*	South Africa
303	Glomerella	acutata	STE-U 4452	AY376503.1	*Protea magnifica*	South Africa
304	Glomerella	acutata	STE-U 4456	AY376504.1	*Protea repens*	South Africa
305	Glomerella	acutata	STE-U 4457	AY376505.1	*Protea spp.*	South Africa
306	Glomerella	acutata	STE-U 4458	AY376506.1	*Protea spp.*	South Africa
307	Glomerella	acutata	STE-U 4459	AY376507.1	*Protea spp.*	South Africa
308	Glomerella	acutata	STE-U 5303	AY376508.1	*Hevea brasiliensis*	India
309	Glomerella	acutata	STE-U 5287	AY376509.1	*Malus domestica*	USA
310	Glomerella	acutata	STE-U 5292	AY376510.1	*Carica papaya*	Australia
311	Glomerella	acutata	STE-U 4471	AY376511.1	*Hakea sericea*	South Africa
312	Glomerella	acutata	STE-U 4467	AY376512.1	*Hakea sericea*	South Africa
313	Glomerella	acutata	STE-U 4466	AY376513.1	*Hakea sericea*	South Africa
314	Glomerella	acutata	STE-U 4470	AY376514.1	*Hakea sericea*	South Africa
315	Glomerella	acutata	STE-U 4465	AY376515.1	*Hakea sericea*	South Africa
316	Glomerella	acutata	STE-U 4462	AY376516.1	*Hakea sericea*	South Africa
317	Glomerella	acutata	STE-U 4463	AY376517.1	*Hakea gibbosa*	South Africa
318	Glomerella	acutata	STE-U 4461	AY376518.1	*Hakea sericea*	South Africa
319	Glomerella	acutata	STE-U 4469	AY376519.1	*Hakea sericea*	South Africa
320	Glomerella	acutata	STE-U 4468	AY376520.1	*Hakea sericea*	South Africa
321	Glomerella	acutata	424	AY513765.1		
322	Glomerella	acutata	554	AY513766.1		
323	Glomerella	acutata	565	AY518543.1		
324	Glomerella	acutata	562	AY518544.1		
325	Glomerella	acutata	427	AY518545.1		
326	Glomerella	acutata	511	AY518546.1		
327	Glomerella	acutata	WAC 12421	AY714051.1	*Boronia megastigma*	Australia
328	Glomerella	acutata	CUSCA02	AY769517.1	*Cuscuta campestris*	Taiwan, Miao-Li
329	Glomerella	acutata	M. Schiller S2.2	AY770553.2	*Fragaria x ananassa*	Costa Rica

330	Glomerella	acutata	M. Schiller S4.10	AY770554.2	*Fragaria x ananassa*	Costa Rica	
331	Glomerella	acutata	M. Schiller F7.8	AY770555.2	*Rumohra adiantiformis*	Costa Rica	
332	Glomerella	acutata	M. Schiller F8.10	AY770556.2	*Rumohra adiantiformis*	Costa Rica	
333	Glomerella	acutata	L.F. Arauz N1	AY770557.2	*Citrus sinensis*	Costa Rica	
334	Glomerella	acutata	SA 0-1	AY818361.1	*Fragaria x ananassa*	Denmark	
335	Glomerella	acutata	GRAY	AY826765.1	*Vitis spp. Marquis*	USA, Michigan, Onondaga	
336	Glomerella	acutata	TUT5954	DQ003101.1	*Fragaria chiloensis*	Israel	1994
337	Glomerella	acutata	ALM-KSH-10	DQ003102.1	*Prunus communis*	Israel, North	
338	Glomerella	acutata	216	DQ003119.1	*Fragaria chiloensis*	USA	
339	Glomerella	acutata	5.7.52	DQ003120.1	*Pyrus malus*	USA, Arkansas	
340	Glomerella	acutata	ALM-9-US	DQ003121.1	*Prunus communis*	USA, California	
341	Glomerella	acutata	Mil-1	DQ003122.1	*Fragaria chiloensis*	USA, Michigan	
342	Glomerella	acutata	TUT137A	DQ003123.1	*Fragaria chiloensis*	Israel	
343	Glomerella	acutata	1.4.57	DQ003124.1	*Pyrus malus*	USA, Rhode Island	
344	Glomerella	acutata	2.7.15	DQ003125.1	*Fragaria chiloensis*	New Zealand	
345	Glomerella	acutata	APPY3	DQ003126.1	*Pyrus malus*	USA, Kentucky	
346	Glomerella	acutata	ALM-IKS-7Q	DQ003127.1	*Prunus communis*	Israel	
347	Glomerella	acutata	ALM-BZR-9A	DQ003128.1	*Prunus communis*	Israel	
348	Glomerella	acutata	ALM-NRB-30K	DQ003129.1	*Prunus communis*	Israel	
349	Glomerella	acutata	ANE-27A	DQ003130.1	*Anemone coronaria*	Israel	
350	Glomerella	acutata	ANE-4	DQ003131.1	*Anemone coronaria*	Israel	
351	Glomerella	acutata	ANE-NL12	DQ003132.1	*Anemone coronaria*	Netherlands	
352	Glomerella	acutata	ANE-25A	DQ003133.1	*Anemone coronaria*	Israel	
353	Glomerella	acutata	S1.7 M. Schiller	DQ018736.1	*Fragaria x ananassa*	Costa Rica	
354	Glomerella	acutata	S1.2 M. Schiller	DQ018737.1	*Fragaria x ananassa*	Costa Rica	
355	Glomerella	acutata	S2.8 M. Schiller	DQ018738.1	*Fragaria x ananassa*	Costa Rica	
356	Glomerella	acutata	S2.9 M. Schiller	DQ018739.1	*Fragaria x ananassa*	Costa Rica	
357	Glomerella	acutata	S3.2 M. Schiller	DQ018740.1	*Fragaria x ananassa*	Costa Rica	
358	Glomerella	acutata	S3.7 M. Schiller	DQ018741.1	*Fragaria x ananassa*	Costa Rica	
359	Glomerella	acutata	S3.9 M. Schiller	DQ018742.1	*Fragaria x ananassa*	Costa Rica	
360	Glomerella	acutata	S4.3 M. Schiller	DQ018743.1	*Fragaria x ananassa*	Costa Rica	
361	Glomerella	acutata	F5.2 M. Schiller	DQ018744.1	*Rumohra adiantiformis*	Costa Rica	
362	Glomerella	acutata	F5.9 M. Schiller	DQ018745.1	*Rumohra adiantiformis*	Costa Rica	
363	Glomerella	acutata	F6.4 M. Schiller	DQ018746.1	*Rumohra adiantiformis*	Costa Rica	
364	Glomerella	acutata	F6.6 M. Schiller	DQ018747.1	*Rumohra adiantiformis*	Costa Rica	
365	Glomerella	acutata	F7.4 M. Schiller	DQ018748.1	*Rumohra adiantiformis*	Costa Rica	
366	Glomerella	acutata	F8.4 M. Schiller	DQ018749.1	*Rumohra adiantiformis*	Costa Rica	

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

367	Glomerella	acutata	unspecified	DQ062670.1	*Cornus florida*	USA, Florida	
368	Colletotrichum	lupini	SHK 788	DQ174692.1	*Lupinus spp.*	South Africa, Free State Province	
369	Colletotrichum	lupini	SHK 1033	DQ174693.1	*Lupinus spp.*	South Africa, Western Cape Province	
370	Colletotrichum	lupini	SHK 2148	DQ174694.1	*Lupinus spp.*	South Africa, Western Cape Province	
371	Glomerella	acutata	Ca.SC.CO-34.04	DQ177875.1	*Prunus persica*	USA, South Carolina	
372	Colletotrichum	lupini	AR2820	DQ286117.1	*Lupinus spp.*	USA, Utah	
373	Colletotrichum	lupini	AR2826	DQ286119.1	*Lupinus spp.*	USA	
374	Glomerella	acutata	ATCC MYA-662	DQ286121.1	*Mauls domestica*	USA, Arkansas	
375	Glomerella	acutata	ATCC MYA-664	DQ286123.1	*Mauls domestica*	USA, Arkansas	
376	Glomerella	acutata	MEP1323	DQ286124.1	*Vaccinium spp.*	New Zealand	
377	Glomerella	acutata	MEP1325	DQ286126.1	*Vaccinium spp.*	New Zealand	
378	Glomerella	acutata	MEP1322	DQ286128.1	*Vaccinium spp.*	New Zealand	
379	Glomerella	acutata	MEP1534	DQ286130.1	*Vaccinium spp.*	New Zealand	
380	Glomerella	acutata	ATCC 56816	DQ286132.1	*Carica papaya*	Australia	
381	Fusarium	phormii	AR3389	DQ286134.1	*Phormium spp.*	New Zealand	
382	Fusarium	phormii	AR3546	DQ286136.1	*Phormium spp.*	Germany	
383	Fusarium	phormii	AR3410	DQ286138.1	*Phormium cookianum*	South Africa	
384	Fusarium	phormii	MEP1334	DQ286140.1	*Phormium spp.*	New Zealand	
385	Fusarium	phormii	CBS199.35	DQ286142.1	*Phormium spp.*	UK, England	
386	Fusarium	phormii	CBS198.35	DQ286144.1	*Phormium spp.*	UK, England	
387	Fusarium	phormii	AR3787	DQ286146.1	*Phormium spp.*	South Africa	
388	Colletotrichum	sp.	ID03	DQ300347.1	*Capsicum chinense*		
389	Colletotrichum	sp.	SM03	DQ300348.1	*Capsicum chinense*		
390	Colletotrichum	sp.	SM01	DQ300349.1	*Capsicum chinense*		
391	Glomerella	acutata	Coll 154	DQ410028.1	*Capsicum annuum*	Taiwan	1992
392	Glomerella	acutata	Coll 153	DQ410029.1	*Capsicum annuum*	Taiwan	1992
393	Glomerella	acutata	Coll 25	DQ410030.1	*Capsicum annuum*	Taiwan	1988
394	Glomerella	acutata	Coll 279	DQ410031.1	*Capsicum annuum*	Taiwan	1997
395	Glomerella	acutata	Mj2	DQ454006.1	*Capsicum annuum*	Thailand	
396	Glomerella	acutata	Mj3	DQ454007.1	*Capsicum annuum*	Thailand	
397	Glomerella	acutata	Mj4	DQ454008.1	*Capsicum annuum*	Thailand	
398	Glomerella	acutata	Mj5	DQ454009.1	*Capsicum annuum*	Thailand	
399	Glomerella	acutata	Mj6	DQ454010.1	*Capsicum annuum*	Thailand	
400	Glomerella	acutata	Mj9	DQ454011.1	*Capsicum annuum*	Thailand	
401	Glomerella	acutata	Mj10	DQ454012.1	*Capsicum annuum*	Thailand	
402	Glomerella	acutata	S2	DQ454018.1	*Fragaria spp.*	Thailand	
403	Glomerella	acutata	S3	DQ454019.1	*Fragaria spp.*	Thailand	

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

404	Glomerella	acutata	S4	DQ454020.1	*Fragaria spp.*	Thailand
405	Glomerella	acutata	S5	DQ454021.1	*Fragaria spp.*	Thailand
406	Glomerella	acutata	S6	DQ454022.1	*Fragaria spp.*	Thailand
407	Glomerella	acutata	S7	DQ454023.1	*Fragaria spp.*	Thailand
408	Glomerella	acutata	VNML	DQ463362.1		
409	Glomerella	acutata	05-16	DQ839609.1	*Myrica cerifera*	USA, Florida
410	Glomerella	acutata	DAR76915	DQ991713.1	*Vitis vinifera*	Australia, NSW, Hastings valley
411	Glomerella	acutata	DAR76884	DQ991714.1	*Vitis vinifera*	Australia, NSW, Hastings valley
412	Glomerella	acutata	DAR76922	DQ991715.1	*Vitis vinifera*	Australia, NSW, Hastings valley
413	Glomerella	acutata	DAR76928	DQ991716.1	*Vitis vinifera*	Australia, NSW, Hastings valley
414	Glomerella	acutata	DAR76929	DQ991717.1	*Vitis vinifera*	Australia, NSW, Hastings valley
415	Glomerella	acutata	DAR77282	DQ991718.1	*Vitis vinifera*	Australia, NSW, Hastings valley
416	Glomerella	acutata	DAR77423	DQ991719.1	*Vitis vinifera*	Australia, NSW, Hastings valley
417	Glomerella	acutata	DAR76923	DQ991720.1	*Lupinus spp.*	Australia, NSW, Junee
418	Glomerella	acutata	DAR76924	DQ991721.1	*Lupinus spp.*	Australia, NSW, Junee
419	Glomerella	acutata	DAR76907	DQ991722.1	*Vitis vinifera*	Australia, NSW, Shoalhaven
420	Glomerella	acutata	DAR76908	DQ991723.1	*Vitis vinifera*	Australia, NSW, Shoalhaven
421	Glomerella	acutata	DAR76909	DQ991724.1	*Vitis vinifera*	Australia, NSW, Shoalhaven
422	Glomerella	acutata	DAR24831a	DQ991725.1	*Persea americana*	Australia, NSW, Murwilumbah
423	Glomerella	acutata	DAR68512	DQ991726.1	*Lycopersicon esculentum*	Australia, NSW, Baulkham hills
424	Glomerella	acutata	DAR76931	DQ991727.1	*Vaccinium myrtillus*	Australia, NSW, Tumbarumba
425	Glomerella	acutata	DAR76932	DQ991728.1	*Vaccinium myrtillus*	Australia, NSW, Tumbarumba
426	Glomerella	acutata	DAR76930	DQ991729.1	*Capsicum spp.*	Australia, NSW, Wagga Wagga
427	Glomerella	acutata	DAR69982	DQ991730.1	*Lycopersicon esculentum*	Australia, Tasmania, Taroota
428	Glomerella	acutata	DAR76933	DQ991731.1	*Capsicum spp.*	Australia, NSW, Wagga Wagga
429	Glomerella	acutata	DAR72407	DQ991732.1	*Prunus dulcis*	Australia, South Australia
430	Glomerella	acutata	DAR76919	DQ991733.1	*Vitis vinifera*	Australia, NSW, Shoalhaven
431	Glomerella	acutata	DAR76887	DQ991737.1	*Vitis vinifera*	Australia, NSW, Hastings valley
432	Glomerella	acutata	DAR75574	DQ991738.1	*Vitis vinifera*	Australia, NSW, Hastings valley
433	Glomerella	acutata	DAR77284	DQ991739.1	*Vitis vinifera*	Australia, Queensland, Kingaroy
434	Glomerella	acutata	DAR28076	DQ991740.1	*Mangifera indica*	Australia, NSW, Alstonville
435	Glomerella	acutata	DAR76886	DQ991741.1	*Vitis vinifera*	Australia, NSW, Tenterfield
436	Glomerella	acutata	DAR76889	DQ991742.1	*Vitis vinifera*	Australia, NSW, Hastings valley
437	Glomerella	acutata	DAR76888	DQ991743.1	*Vitis vinifera*	Australia, NSW, Hastings valley
438	Glomerella	acutata	DAR76925	DQ991744.1	*Vaccinium myrtillus*	Australia, NSW, Corrindi
439	Glomerella	acutata	DAR32068	DQ991745.1	*Fragaria x ananassa*	Australia, NSW, Lismore
440	Glomerella	acutata	DAR76896	DQ991746.1	*Vitis vinifera*	Australia, NSW, Hastings valley

ANNEX 1 - *Colletotrichum acutatum* **rRNA sequences database: CaITSdb**

441	Glomerella	acutata	DAR76900	DQ991747.1	*Vitis vinifera*	Australia, NSW, Hastings valley
442	Glomerella	acutata	DAR76901	DQ991748.1	*Vitis vinifera*	Australia, NSW, Hastings valley
443	Glomerella	acutata	DAR76913	DQ991749.1	*Vitis vinifera*	Australia, NSW, Hastings valley
444	Glomerella	acutata	DAR76926	DQ991750.1	*Mangifera indica*	Australia, NSW, northern coast
445	Glomerella	acutata	DAR76921	DQ991751.1	*Olea europaea*	Australia, NSW, Hunter valley
446	Glomerella	acutata	Cg 5	EF025968.1	*Vitis vinifera*	
447	Colletotrichum	carthami	Cg 33	EF025973.1	*Vitis vinifera*	
448	unclassified	unclassified	LM440	EF060746.1	see water	USA, Hawaii
449	Glomerella	acutata	BRIP 4703a	EF143971.1	*Fragaria x ananassa*	Australia
450	Glomerella	acutata	BRIP 4704a	EF143972.1	*Fragaria x ananassa*	Australia
451	Glomerella	acutata	BRIP 28517a	EF143973.1	*Carica papaya*	Australia
452	Glomerella	acutata	BRIP 11086a	EF143974.1	*Fragaria x ananassa*	Australia
453	Glomerella	acutata	BRIP 28519a	EF143975.1	*Carica papaya*	Australia
454	Glomerella	acutata	CIAD/GAQ-34	EF175780.1	*Persea americana*	Mexico, Nuevo Parangaricutiro
455	Glomerella	acutata	CIAD/GAQ-01	EF221831.1	*Persea americana*	Mexico, Tancitaro, Michoacan
456	Glomerella	acutata	CIAD/GAQ-03	EF221832.1	*Persea americana*	Mexico, Tancitaro, Michoacan
457	Glomerella	miyabeana	VPRI 32545	EF452724.1	*Salix fragilis*	Australia
458	Glomerella	miyabeana	VPRI 32575	EF452725.1	*Salix fragilis*	Australia
459	Glomerella	miyabeana	VPRI 32546	EF452726.1	*Salix alba var. vitellina*	Australia
460	Glomerella	miyabeana	VPRI 32547	EF452727.1	*Salix alba var. vitellina*	Australia
461	Colletotrichum	acutatum	EHS36	EF464591.1	*Fiorinia externa*	USA, New York, Mohonk
462	Colletotrichum	sp.	EHS41	EF464592.1	*Fiorinia externa*	USA, New York, Mohonk
463	Colletotrichum	acutatum	EHS48	EF464593.1	*Fiorinia externa*	USA, New York, Bayberry Lane
464	Colletotrichum	acutatum	EHS58	EF464594.1	*Fiorinia externa*	USA, New York
465	Colletotrichum	acutatum	EHS61	EF464595.1	*Fiorinia externa*	USA, New York
466	Glomerella	cingulata	qdh-1	EF501982.1	*Capsicum sp.*	China
467	Colletotrichum	acutatum	EHS51	EF593369.1	*Fiorinia externa*	USA, New York, Esopus
468	Colletotrichum	acutatum	EHS52	EF593370.1	*Fiorinia externa*	USA, New York, Esopus
469	Colletotrichum	gloeosporioides	ARSEF4360	EF593371.1	*Orthezia praelonga*	Brazil, Rio de Janeiro
470	Colletotrichum	gloeosporioides	EMA26	EF593372.1	*Orthezia praelonga*	Brazil, Sao Paulo
471	Colletotrichum	gloeosporioides	Cg2	EF608052.1	*Mangifera indica*	Taiwan
472	Colletotrichum	gloeosporioides	Cg9	EF608055.1	*Carica papaya*	Taiwan
473	Colletotrichum	sp.	C-sp1	EF608061.1	*Syzygium samarangense*	Taiwan
474	Glomerella	acutata	UCA1018	EF622177.1	*Fragaria x ananassa*	Spain
475	Glomerella	acutata	UCA1026	EF622178.1	*Fragaria x ananassa*	Spain
476	Glomerella	acutata	UCA1072	EF622179.1	*Fragaria x ananassa*	Spain
477	Glomerella	acutata	UCA1084	EF622180.1	*Fragaria x ananassa*	Spain

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

478	Glomerella	acutata	UCA1100	EF622181.1	*Fragaria x ananassa*	Switzerland
479	Glomerella	acutata	UCA1101	EF622182.1	*Fragaria x ananassa*	Switzerland
480	Glomerella	acutata	UCA1102	EF622183.1	*Fragaria x ananassa*	New Zealand
481	Glomerella	acutata	UCA1103	EF622184.1	*Fragaria x ananassa*	Portugal
482	Glomerella	acutata	UCA1105	EF622186.1	*Fragaria x ananassa*	Portugal
483	Glomerella	acutata	UCA1109	EF622187.1	*Fragaria x ananassa*	Germany
484	Glomerella	acutata	UCA1113	EF622188.1	*Rosa spp.*	Netherlands
485	Glomerella	acutata	UCA1117	EF622189.1	*Fragaria x ananassa*	USA
486	Glomerella	acutata	UCA1118	EF622190.1	*Fragaria x ananassa*	USA
487	Glomerella	acutata	UCA1120	EF622191.1	*Lupin spp.*	Germany
488	Glomerella	acutata	UCA1121	EF622192.1	*Lupin spp.*	Netherlands
489	Glomerella	acutata	UCA1122	EF622193.1	*Lupin spp.*	Australia
490	Glomerella	acutata	UCA1123	EF622194.1	*Lupin spp.*	Australia
491	Glomerella	acutata	UCA1125	EF622195.1	*Lupin spp.*	UK
492	Glomerella	acutata	UCA1126	EF622196.1	*Lupin spp.*	UK
493	Glomerella	acutata	UCA1127	EF622197.1	*Lupin spp.*	UK
494	Glomerella	acutata	UCA1128	EF622198.1	*Lupin spp.*	Netherlands
495	Glomerella	acutata	UCA1135	EF622199.1	*Fragaria x ananassa*	New Zealand
496	Glomerella	acutata	UCA1136	EF622200.1	*Phlox spp.*	Netherlands
497	Glomerella	acutata	UCA-ST8	EF622202.1	*Fragaria x ananassa*	UK
498	Glomerella	acutata	UCA-ST17	EF622203.1	*Fragaria x ananassa*	UK
499	Glomerella	acutata	UCA1160	EF622205.1	*Fragaria x ananassa*	UK
500	Glomerella	cingulata	DAOM 233253	EF672241.1	*Vaccinium sp. (cranberry)*	
501	Colletotrichum	sp.	Vega633	EF687919.1	*Coffea arabica (*crown endophyte)	Colombia
502	Colletotrichum	sp.	Vega389	EF687920.1	*Coffea arabica (*crown endophyte)	Colombia
503	Colletotrichum	sp.	Vega007	EF687921.1	*Coffea arabica (*leaf endophyte)	Mexico
504	Colletotrichum	sp.	Vega418	EF687922.1	*Coffea arabica (*crown endophyte)	Colombia
505	Glomerella	acutata	UCA1005	EF694673.1	*Fragaria x ananassa*	Spain
506	Glomerella	acutata	UCA1015	EF694674.1	*Fragaria x ananassa*	Spain
507	Glomerella	acutata	UCA1025	EF694675.1	*Fragaria x ananassa*	Spain
508	Glomerella	acutata	UCA1028	EF694676.1	*Fragaria x ananassa*	Spain
509	Glomerella	acutata	UCA1070	EF694677.1	*Fragaria x ananassa*	Spain
510	Glomerella	acutata	UCA1076	EF694678.1	*Fragaria x ananassa*	Spain
511	Glomerella	acutata	UCA1077	EF694679.1	*Fragaria x ananassa*	Spain
512	Glomerella	acutata	UCA1078	EF694680.1	*Fragaria x ananassa*	Spain
513	Glomerella	acutata	IMI 348487	EF694681.1	*Fragaria x ananassa*	France
514	Glomerella	acutata	IMI 348489	EF694682.1	*Fragaria x ananassa*	France

515	Glomerella	acutata	M1	EU008828.1	*Malus domestica*	Brazil	
516	Glomerella	acutata	M2	EU008829.1	*Malus domestica*	Brazil	
517	Glomerella	acutata	M31	EU008857.1	*Malus domestica*	Brazil	
518	Glomerella	acutata	M32	EU008858.1	*Malus domestica*	Brazil	
519	Glomerella	acutata	M33	EU008859.1	*Malus domestica*	Brazil	
520	Glomerella	acutata	M34	EU008860.1	*Malus domestica*	Brazil	
521	Glomerella	acutata	M35	EU008861.1	*Malus domestica*	Brazil	
522	Glomerella	acutata	M36	EU008862.1	*Malus domestica*	Brazil	
523	Glomerella	acutata	M37	EU008863.1	*Malus domestica*	Brazil	
524	Glomerella	acutata	M38	EU008864.1	*Malus domestica*	Brazil	
525	Glomerella	acutata	M39	EU008865.1	*Malus domestica*	Brazil	
526	Glomerella	acutata	M40	EU008866.1	*Malus domestica*	Brazil	
527	Glomerella	acutata	C2	EU008878.1	*Citrus spp.*	Brazil	
528	Glomerella	acutata	MICH-ZZ	EU016517.1	*Rubus spp.*	Mexico	
529	Glomerella	acutata	489	EU109500.1	*Solanum melongena*		
530	Glomerella	acutata	UCA1083	EU109737.1	*Fragaria x ananassa*	Spain	
531	Glomerella	acutata	UCA1088	EU109738.1	*Fragaria x ananassa*	Spain	
532	Glomerella	acutata	UCA1089	EU109739.1	*Fragaria x ananassa*	Spain	
533	Glomerella	acutata	UCA1090	EU109740.1	*Fragaria x ananassa*	Spain	
534	Glomerella	acutata	UCA1091	EU109741.1	*Fragaria x ananassa*	Spain	
535	Glomerella	acutata	Stb-20	EU131874.1	*Fragaria x ananassa* cv. Senga Sengana	Bulgaria, Pchelin, Sofia	2004
536	Glomerella	acutata	Stb-27	EU131875.1	*Fragaria x ananassa* cv. Pokahontas	Bulgaria, Berkovitsa	2004
537	Glomerella	acutata	Stb-32	EU131876.1	*Fragaria x ananassa* cv. Siabella	Bulgaria, Velingrad, Pazardzhik	2004
538	Glomerella	acutata	Cir-49	EU131877.1	*Cirsium arvense*	Bulgaria, Izbegli, Plovdiv	2004
539	Glomerella	acutata	Stb-51	EU131878.1	*Fragaria x ananassa* cv. Selva	Bulgaria, Skutare, Plovdiv	2004
540	Glomerella	acutata	Stb-61	EU131879.1	*Fragaria x ananassa* cv. Marmolada	Bulgaria, Yagodovo, Plovdiv	2005
541	Glomerella	acutata	Stb-101	EU131880.1	*Fragaria x ananassa* cv. Cambridge	Bulgaria, Zhrebichko, Plovdiv	2005
542	Glomerella	acutata	Stb-91	EU131881.1	*Fragaria x ananassa* cv. Elsanta	Bulgaria, Krichim, Plovdiv	2005
543	Glomerella	acutata	Tm-106	EU131882.1	*Solanum lycopersicum* cv. Florida	Bulgaria, Cheshnigirovo, Plovdiv	2005
544	Glomerella	acutata	Pep-115	EU131883.1	*Kurtovska kapija*	Bulgaria: Cheshnigirovo, Plovdiv	2005
545	Glomerella	acutata	ALB-IND-25	EU168901.1	*Citrus x sinensis*	USA, Florida, Indiantown	
546	Glomerella	acutata	MTR-KLA-A1	EU168902.1	*Citrus aurantifolia*	Belize	
547	Glomerella	acutata	HM-1	EU168903.1	*Citrus aurantifolia*	USA, Florida, Homestead	
548	Glomerella	acutata	Gc-13-1	EU168904.1	*Vaccinium sp.* (blueberry)	USA, North Carolina	
549	Glomerella	acutata	cmf-04	EU200457.1	*Fragaria x ananassa*	China, Shangai	
550	Glomerella	acutata	CIAD/GAQ-37	EU301722.1	*Persea americana*	Mexico, Nuevo Parangaricutiro	
551	Glomerella	acutata	CIAD/GAQ-36	EU301723.1	*Persea americana*	Mexico, Nuevo Parangaricutiro	

552	Glomerella	acutata	CA14-1	EU391655.1	*Fragaria x ananassa* cv. AGF-80	Brazil, Bragança Paulista	
553	Colletotrichum	higginsianum	DAOM225478	EU400147.1			
554	Glomerella	acutata	DAOM214715	EU400153.1			
555	Glomerella	acutata	DAOM214992	EU400154.1			
556	Colletotrichum	gloeosporioides	NW551b	EU520094.1	*Paulownia tomentosa*	China, Luoyang	
557	Colletotrichum	gloeosporioides	NW721	EU520113.1	*Populus sp.*	China, Luochuan	
558	Colletotrichum	gloeosporioides	NW551	EU520250.1	*Populus sp.*	China, Luochuan	
559	Glomerella	acutata	PCF230	EU523537.1	*Fragaria x ananassa*	Belgium	
560	Glomerella	acutata	PCF229	EU523538.1	*Fragaria x ananassa*	Belgium	
561	Glomerella	acutata	PCF459	EU523539.1	*Fragaria x ananassa*	Belgium	
562	Colletotrichum	lupini	A-1	EU589451.1	*Lupinus albus*	Chile, Gorbea	
563	Glomerella	sp.	MP3	EU622052.1	*Acer Platanoides*	USA, Massachusetts, Boston	2006
564	Glomerella	acutata	May-88	EU647299.1	*Vaccinium spp.* (blueberry fruit)	Usa, Florida, Dover	
565	Glomerella	acutata	05-148	EU647300.1	*Vaccinium spp.* (blueberry fruit)	USA, North Carolina	
566	Glomerella	acutata	05-197	EU647301.1	*Vaccinium spp.* (blueberry fruit)	USA, Georgia	
567	Glomerella	acutata	02-163	EU647302.1	*Fragaria x ananassa* (petiole)	USA, Florida, Floral City	
568	Glomerella	acutata	02-179	EU647303.1	*Fragaria x ananassa* (fruit)	USA, Florida, Plant City	
569	Glomerella	acutata	Mar-32	EU647304.1	*Fragaria x ananassa* (crown)	USA, Florida, Thonotosassa	
570	Glomerella	acutata	OCO-ARC-4	EU647305.1	*Citrus x sinensis*	USA, Florida, Arcadia	
571	Glomerella	acutata	STF-FTP-10	EU647306.1	*Citrus x sinensis*	USA, Florida, Frostproof	
572	Glomerella	acutata	Ss	EU647307.1	*Citrus aurantifolia*	USA, Florida, Sarasota	
573	Glomerella	acutata	KLA-Anderson	EU647308.1	*Citrus aurantifolia*	USA, Florida, Lake Alfred	
574	Glomerella	acutata	05-155	EU647309.1	*Rumohra adiantiformis*	USA, Florida, Pierson	
575	Glomerella	acutata	05-161	EU647310.1	*Rumohra adiantiformis*	USA, Florida, Crescent City	
576	Glomerella	acutata	05-200	EU647311.1	*Rumohra adiantiformis*	USA, Florida, Seville	
577	Glomerella	acutata	CSL-1690	EU670079.1	*Prunus dulcis*	Australia, Angle Vale	
578	Glomerella	acutata	CSL-1689	EU670080.1	*Prunus dulcis*	Australia, Willunga	
579	Glomerella	acutata	W16	EU670081.1	*Prunus dulcis*	Australia, Willunga	
580	Glomerella	acutata	CSL-1318	EU670083.1	*Prunus dulcis*	Australia, Angle Vale	
581	Glomerella	acutata	CSL-1688	EU670084.1	*Prunus dulcis*	Australia, Angle Vale	
582	Glomerella	acutata	W153	EU670085.1	*Prunus dulcis*	Australia, Victoria, Nangiloc	
583	Glomerella	acutata	US-1813	EU670086.1	*Prunus dulcis*	Australia, Willunga	
584	Glomerella	acutata	CHY-1	EU727317.1	*Hevea spp.*	China, Yunnan	
585	Glomerella	acutata	CHH-1	EU727318.1	*Hevea spp.*	China, Hainan	
586	Glomerella	acutata	GM59a	EU734581.1	*Annona muricata*	Colombia	
587	Glomerella	acutata	GM77	EU734582.1	*Annona muricata*	Colombia	
588	Glomerella	acutata	ICMP 16981	EU770245.1	*Vitis spp.* (grapewine)	New Zealand	

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

589	Glomerella	acutata	ICMP 16982	EU770251.1	*Vitis spp.* (grapewine)	New Zealand
590	Glomerella	acutata	SS5	EU886753.1	*Ixodes scapularis*	
591	Glomerella	acutata	MM1	EU886755.1	*Ixodes scapularis*	
592	unclassified	unclassified	P1921A	EU977229.1		
593	Colletotrichum	gloeosporioides	NJ-GACA	EU979125.1	*Calamagrostis x acutiflora*	USA, New Jersey, Barrington
594	unclassified	unclassified	M4-3151	FJ025324.1		
595	unclassified	unclassified	O10-2171	FJ025352.1		
596	Glomerella	acutata	Gilan	FJ185786.1	*Phaseolus sp.* bean	Iran
597	Glomerella	acutata	c.Traminette	FJ189505.1	*Vitis vinifers* cult. traminette	USA, North Carolina
598	Glomerella	acutata	r8.1	FJ228184.1	*Fraxinus excelsior*	Sweden
599	Colletotrichum	sp.	EXMY-22	FJ233191.1	*Vitis vinifera*(red-grape)	
600	Glomerella	acutata	D/161	FJ372571.1	*Fragaria x ananassa*	Hungary
601	unclassified	unclassified	1071	FJ449924.1	*Dendrobium spp.*	
602	Glomerella	acutata	HT27	FJ455521.1		
603	Glomerella	acutata	JF9	FJ455522.1		
604	Glomerella	acutata	JF17	FJ455523.1		
605	Glomerella	acutata	WZ46	FJ455524.1		
606	Glomerella	acutata	152	FJ478047.1	*Vaccinium macrocarpon*	USA, New Jersey
607	Glomerella	acutata	153	FJ478048.1	*Vaccinium macrocarpon*	USA, New Jersey
608	Glomerella	acutata	154	FJ478049.1	*Vaccinium macrocarpon*	USA, New Jersey
609	Glomerella	acutata	155	FJ478050.1	*Vaccinium macrocarpon*	USA, New Jersey
610	Glomerella	acutata	156	FJ478051.1	*Vaccinium macrocarpon*	Canada, British Columbia
611	Glomerella	acutata	157	FJ478052.1	*Vaccinium macrocarpon*	Canada, British Columbia
612	Glomerella	acutata	159	FJ478053.1	*Vaccinium macrocarpon*	Canada, British Columbia
613	Glomerella	acutata	160	FJ478054.1	*Vaccinium macrocarpon*	Canada, British Columbia
614	Glomerella	acutata	161	FJ478055.1	*Vaccinium macrocarpon*	USA, Wisconsin
615	Glomerella	acutata	162	FJ478056.1	*Vaccinium macrocarpon*	USA, Wisconsin
616	Glomerella	acutata	163	FJ478057.1	*Vaccinium macrocarpon*	USA, Wisconsin
617	Glomerella	acutata	164	FJ478058.1	*Vaccinium macrocarpon*	USA, Wisconsin
618	Glomerella	acutata	165	FJ478059.1	*Vaccinium macrocarpon*	USA, Wisconsin
619	Glomerella	acutata	166	FJ478060.1	*Vaccinium macrocarpon*	USA, Wisconsin
620	Glomerella	acutata	167	FJ478061.1	*Vaccinium macrocarpon*	USA, Wisconsin
621	Glomerella	acutata	168	FJ478062.1	*Vaccinium macrocarpon*	USA, Wisconsin
622	Glomerella	acutata	169	FJ478063.1	*Vaccinium macrocarpon*	USA, Wisconsin
623	Glomerella	acutata	170	FJ478064.1	*Vaccinium macrocarpon*	USA, Wisconsin
624	Glomerella	acutata	172	FJ478065.1	*Vaccinium macrocarpon*	USA, Wisconsin
625	Glomerella	acutata	173	FJ478066.1	*Vaccinium macrocarpon*	Canada, British Columbia

626	Glomerella	acutata	178	FJ478067.1	*Vaccinium macrocarpon*	USA, Massachusetts
627	Glomerella	acutata	180	FJ478068.1	*Vaccinium macrocarpon*	USA, Massachusetts
628	Glomerella	acutata	181	FJ478069.1	*Vaccinium macrocarpon*	USA, Massachusetts
629	Glomerella	acutata	182	FJ478070.1	*Vaccinium macrocarpon*	USA, New Jersey
630	Glomerella	acutata	183	FJ478071.1	*Vaccinium macrocarpon*	USA, New Jersey
631	Glomerella	acutata	184	FJ478072.1	*Vaccinium macrocarpon*	USA, New Jersey
632	Glomerella	acutata	185	FJ478073.1	*Vaccinium macrocarpon*	USA, New Jersey
633	Glomerella	acutata	186	FJ478074.1	*Vaccinium macrocarpon*	USA, New Jersey
634	Glomerella	acutata	201	FJ478075.1	*Vaccinium macrocarpon*	Canada, British Columbia
635	Glomerella	acutata	202	FJ478076.1	*Vaccinium macrocarpon*	Canada, British Columbia
636	Glomerella	acutata	203	FJ478077.1	*Vaccinium macrocarpon*	Canada, British Columbia
637	Glomerella	acutata	205	FJ478078.1	*Vaccinium macrocarpon*	Canada, British Columbia
638	Glomerella	acutata	206	FJ478079.1	*Vaccinium macrocarpon*	Canada, British Columbia
639	Glomerella	acutata	207	FJ478080.1	*Vaccinium macrocarpon*	Canada, British Columbia
640	Glomerella	acutata	210	FJ478081.1	*Vaccinium macrocarpon*	Canada, British Columbia
641	Glomerella	acutata	211	FJ478082.1	*Vaccinium macrocarpon*	Canada, British Columbia
642	Glomerella	acutata	212	FJ478083.1	*Vaccinium macrocarpon*	Canada, British Columbia
643	Glomerella	acutata	213	FJ478084.1	*Vaccinium macrocarpon*	Canada, British Columbia
644	Glomerella	acutata	214	FJ478085.1	*Vaccinium macrocarpon*	Canada, British Columbia
645	Glomerella	acutata	WY 21	FJ608645.1	*Fragaria x ananassa*	China
646	Glomerella	acutata	WY 22	FJ608646.1	*Fragaria x ananassa*	China
647	Glomerella	acutata	WY 23	FJ608647.1	*Fragaria x ananassa*	China
648	Glomerella	acutata	WY 24	FJ608648.1	*Fragaria x ananassa*	China
649	Glomerella	acutata	WY 25	FJ608649.1	*Fragaria x ananassa*	China
650	Glomerella	acutata	WY 26	FJ608650.1	*Fragaria x ananassa*	China
651	Glomerella	acutata	WY 27	FJ608651.1	*Fragaria x ananassa*	China
652	Glomerella	acutata	WY 28	FJ608652.1	*Fragaria x ananassa*	China
653	Glomerella	acutata	WY 29	FJ608653.1	*Fragaria x ananassa*	China
654	Glomerella	acutata	WY 30	FJ608654.1	*Fragaria x ananassa*	China
655	Glomerella	acutata	WY 31	FJ608655.1	*Fragaria x ananassa*	China
656	Glomerella	acutata	ATCC MYA-4396	FJ746689.1	*Citrus blossom*	USA, Florida
657	Glomerella	acutata	ATCC MYA-4397	FJ746690.1	*Citrus aurantifolia*	USA, Florida
658	Glomerella	acutata	ATCC MYA-4398	FJ746691.1	*Vaccinium spp.* (blueberry)	USA, North Carolina
659	Glomerella	acutata	16633D	FJ788417.1	*Carica papaya*	Australia
660	Glomerella	acutata	ATCC MYA-4516	FJ810511.1	*Rumohra adiantiformis*	USA, Florida
661	Glomerella	acutata	ATCC MYA-4517	FJ810512.1	*Rumohra adiantiformis*	USA, Florida
662	Glomerella	acutata	ATCC MYA-4518	FJ810513.1	*Rumohra adiantiformis*	USA, Florida

663	Glomerella	acutata	ATCC MYA-4519	FJ810514.1	*Fragaria spp.*	USA, Florida	
664	Glomerella	acutata	ATCC MYA-4520	FJ810515.1	*Vaccinium spp.* (blueberry)	USA, Florida	
665	Glomerella	acutata	T181	FJ938293.1		China	
666	Glomerella	acutata	T182	FJ938294.1		China	
667	Glomerella	acutata	213(1)	FJ968598.1	*Coffea arabica*	North Vietnam	
668	Glomerella	acutata	BMT(HL)19	FJ968601.1	*Coffea arabica*	South Vietnam	
669	Glomerella	acutata	BRIP28519	FJ972601.1	*Carica papaya*	Australia	
670	Glomerella	acutata	CBS 29467	FJ972610.1	*Carica papaya*	Australia	
671	Colletotrichum	acutatum	MU-2009 38	FN548156.1	*Fagus sylvatica*	Germany, City of Greifswald	
672	Glomerella	acutata	Soskut1	FR716517.1	*Prunus cerasus*	Hungary	
673	Glomerella	acutata	Hajdudorog1	FR716518.1	*Prunus cerasus*	Hungary	
674	Colletotrichum	sp.	SOD107	GQ119341.1	*Ailanthus altissima*	China	
675	Colletotrichum	caudatum	XXXXXXXXX	GQ369598.1			
676	Colletotrichum	caudatum	XXXXXXXXX	GQ379684.1			
677	Glomerella	acutata	PCF602	GQ861450.1	*Fragaria spp.*	Belgium	
678	Glomerella	acutata	PCF492	GQ861451.1	*Fragaria spp.*	Belgium	
679	Glomerella	acutata	PCF194	GQ861452.1	*Fragaria spp.*	Belgium	
680	Glomerella	acutata	PCF982	GQ861453.1	*Fragaria spp.*	Belgium	
681	Glomerella	acutata	PCF989	GQ861454.1	*Fragaria spp.*	Belgium	
682	Glomerella	acutata	TDARES35	GQ889269.1	*Vigna radiata*	Taiwan	
683	Glomerella	acutata	FDC31a08	GQ994099.1	*Murraya paniculata*	Brazil	
684	unclassified	unclassified	mh3496.1	GQ996079.1	*Acer barbatum*	USA: Duke Forest, Orange County	
685	Glomerella	acutata	Ca3	GU045506.1	*Mangifera indica*	USA, Florida	
686	Glomerella	acutata	Ca13	GU045507.1	*Mangifera indica*	USA, Florida	
687	Glomerella	acutata	Ca26	GU045508.1	*Mangifera indica*	USA, Florida	
688	Glomerella	acutata	Ca33	GU045509.1	*Mangifera indica*	USA, Florida	
689	Glomerella	acutata	Ca60	GU045510.1	*Mangifera indica*	USA, Florida	
690	Glomerella	acutata	HNZJ001	GU059863.1	*Capsicum sp. (hot pepper)*	China	
691	Glomerella	sp.	I325	GU062296.1	*Alnus incana*	Latvia	
692	Glomerella	acutata	FDC89a08	GU120390.1	*Citrus sp.*	Brasil	
693	Colletotrichum	simmondsii	BRIP4684	GU183315.1	*Capsicum frutescens*	Australia,QLD, Brisbane	1955
694	Colletotrichum	simmondsii	BRIP4703	GU183316.1	*Fragaria x ananassa*	Australia, QLD, Townsville	1971
695	Colletotrichum	simmondsii	BRIP4704	GU183317.1	*Fragaria x ananassa*	Australia, QLD, Forest Glen	1972
696	Colletotrichum	simmondsii	BRIP11086	GU183318.1	*Fragaria x ananassa*	Australia, QLD, Nanbour	1965
697	Colletotrichum	simmondsii	BRIP19776	GU183319.1	*Carica papaya*	Australia, QLD, Yandina	1987
698	Glomerella	fioriniae	BRIP20127	GU183320.1	*Persea americana*	Australia,QLD, Brisbane	1989
699	Colletotrichum	simmondsii	BRIP24124	GU183321.1	*Nephelium lappaceum*	Australia,QLD, Kamerunga	1989

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

700	Colletotrichum	simmondsii	BRIP24191	GU183322.1	*Actinidia chinensis*	Australia, QLD, Mt Tamborine	1991
701	Colletotrichum	simmondsii	BRIP24197	GU183323.1	*Actinidia chinensis*	Australia, QLD, Mt Tamborine	1991
702	Colletotrichum	simmondsii	BRIP24243	GU183324.1	*Litchi chinensis*	Australia, QLD, Atherton Tableland	1992
703	Colletotrichum	simmondsii	BRIP24246	GU183325.1	*Litchi chinensis*	Australia, QLD, Atherton Tableland	1992
704	Glomerella	acutata	BRIP27048	GU183326.1	*Mangifera indica*	Australia, QLD, Ayr	1993
705	Colletotrichum	simmondsii	BRIP28420	GU183327.1	*Cyphomandra betacea*	Australia, QLD, Mt Tamborine	1987
706	Colletotrichum	simmondsii	BRIP28487	GU183328.1	*Averrhoa carambola*	Australia, QLD	1987
707	Colletotrichum	simmondsii	BRIP28517	GU183329.1	*Carica papaya*	Australia, QLD, Yandina	1987
708	Colletotrichum	simmondsii	BRIP28518	GU183330.1	*Carica papaya*	Australia, QLD, Yandina	1987
709	Colletotrichum	simmondsii	BRIP28519	GU183331.1	*Carica papaya*	Australia, QLD, Yandina	1987
710	Colletotrichum	simmondsii	BRIP28533	GU183332.1	*Persea americana*	Australia, QLD	1986
711	Glomerella	fioriniae	BRIP28761	GU183333.1	*Mangifera indica*	Australia, QLD, Yarwun	1994
712	Colletotrichum	simmondsii	BRIP28832	GU183334.1	*Mangifera indica*	Australia, QLD, Ayr	1993
713	Glomerella	fioriniae	BRIP29284	GU183335.1	*Persea americana*	Australia, QLD, Mt Tamborine	2002
714	Glomerella	fioriniae	BRIP29285	GU183336.1	*Persea americana*	Australia, QLD, Mt Tamborine	2002
715	Colletotrichum	simmondsii	BRIP39473	GU183337.1	*Litchi chinensis*	Australia, NSW, Byron Bay	2003
716	Colletotrichum	simmondsii	BRIP48724	GU183338.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
717	Colletotrichum	simmondsii	BRIP48726	GU183339.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
718	Colletotrichum	simmondsii	BRIP48729	GU183340.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
719	Colletotrichum	simmondsii	BRIP48731	GU183341.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
720	Colletotrichum	simmondsii	BRIP48734	GU183342.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
721	Colletotrichum	simmondsii	BRIP48737	GU183343.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
722	Colletotrichum	simmondsii	BRIP48742	GU183344.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
723	Colletotrichum	simmondsii	BRIP48761	GU183345.1	*Litchi chinensis*	Australia, NSW, Mena Creek	2003
724	Glomerella	fioriniae	BRIP52335	GU183346.1	*Persea americana*	Australia, WA, Pemberton	2008
725	Glomerella	fioriniae	BRIP52336	GU183347.1	*Persea americana*	Australia, WA, Pemberton	2008
726	Colletotrichum	simmondsii	BRIP52651	GU183348.1	*Vaccinum corymbosum*	Australia, VIC, Knoxfield	1987
727	Glomerella	acutata	BRIP52652	GU183349.1	*Ranunculus sp.*	Australia, VIC, Clayton South	1989
728	Glomerella	acutata	BRIP52653	GU183350.1	*Anemone sp.*	Australia, VIC, Geelong	1976
729	Colletotrichum	simmondsii	BRIP52654	GU183351.1	*Fragaria x ananassa*	Australia, VIC, Scoresby	1976
730	Colletotrichum	simmondsii	BRIP52655	GU183352.1	*Fragaria x ananassa*	Australia, VIC, Silvan	1955
731	Glomerella	acutata	BRIP52656	GU183353.1	*Anemone sp.*	Australia, VIC, Geelong	
732	Colletotrichum	simmondsii	BRIP52657	GU183354.1	*Lycopersicon esculentum*	Australia, NSW, Tweed Heads	1980
733	Glomerella	acutata	BRIP52690	GU183355.1	*Pistacia vera*	Australia	1989
734	Glomerella	acutata	BRIP52691	GU183356.1	*Fragaria x ananassa*	Australia, WA, Wanneroo	1988
735	Glomerella	acutata	BRIP52692	GU183357.1	*Olea europaea*	Australia, WA, Kalamunda	1991
736	Colletotrichum	simmondsii	BRIP52693	GU183358.1	*Fragaria x ananassa*	Australia	1992

ANNEX 1 - *Colletotrichum acutatum* rRNA sequences database: CaITSdb

737	Glomerella	acutata	BRIP52695	GU183360.1	*Boronia megastigma*	Australia, SA, Mt Baker	2004
738	Glomerella	fioriniae	BRIP52697	GU183362.1	*Actinidia chinensis*	Australia	1991
739	Glomerella	acutata	DLEN2008024	GU244527.1	surface of marine organism	China, Liaoning, Dalian	
740	unclassified	unclassified	UPSC_F5_65	GU565008.1	*Populus tremula*	Sweden	
741	Colletotrichum	sp.	LF17	GU951761.1	*Huperzia serrata*	China, Jiangxi province	
742	Colletotrichum	sp.	LF43	GU951766.1	*Huperzia serrata*	China, Jiangxi province	
743	Colletotrichum	sp.	MFU09 0619	HM038360.1	*Capsicum sp.* (chilli)	Laos, Vientiane	
744	Colletotrichum	sp.	MFU09 0628	HM038361.1	*Mangifera indica*	Thailand, Bankok	
745	Colletotrichum	sp.	MFU09 0624	HM038362.1	*Zizyphus mauritiane*	Thailand, Bankok	
746	Glomerella	acutata	YN-01	HM575267.1	*Citrus spp.*	China	
747	Glomerella	acutata	YN-02	HM575268.1	*Citrus spp.*	China	
748	Glomerella	acutata	YN-3-16	HM575269.1	*Citrus spp.*	China	
749	Glomerella	sp.	E7024b	HQ003925.1	Rain forest plant		
750	Glomerella	sp.	NY7955a	HQ007248.1	*Pyrus fauriei*	Ecuador	
751	Glomerella	acutata	Ca262 HP	HQ330982.1	*Capsicum annuum* (chili)	India	