

4 Characterization of the CWD Pathogen

*Ex Africa semper aliquid novi*¹

4.1 Main Findings

- Powerful molecular identification methods have been used to differentiate between the different strains of coffee wilt disease (CWD), which confirm the complexity of the disease.
- The strain of the disease currently affecting Robusta coffee in Democratic Republic of Congo (DRC), Uganda and Tanzania is identical with a strain isolated from DRC in 1960.
- There is no detectable genetic variation in the parent strains (MAT-1 and MAT-2) present in DRC, Uganda and Tanzania, making it very likely that the disease spread from a small initial outbreak in DRC, perhaps a single farm.
- It is likely, therefore, that CWD was never completely eradicated in DRC after the historical (mid-20th century) outbreak.
- The CWD found on Arabica in Ethiopia since 1957 is a different strain, which does not attack Robusta, and does not interbreed with the Robusta strain.
- It is suggested that the Arabica and Robusta strains of CWD are in fact separate diseases that arose independently, most likely from undetected disease forms on wild coffee species or even non-coffee species.
- It is, therefore, possible that a new outbreak could occur spontaneously at any time in the future.
- Hence we recommended that all African coffee zones be regularly monitored to detect new disease events so that they can be quickly contained.

4.2 Introduction

As with any disease, it is vital that we understand as much as possible about the pathogen that causes CWD. We need to be able to accurately and rapidly identify it and this has not been easy – even experts have found it difficult. Additionally, efforts to effectively manage CWD have been hindered by our lack of knowledge of the disease and how it behaves. This includes especially knowledge of its variability because this will affect breeding strategies to develop the durable resistance.

¹ 'Out of Africa, always something new'. (Pliny the Elder, from a Greek proverb)

Phiri N. and Baker, P.S. (2009) *Coffee Wilt in Africa Final Technical Report*. CAB International.

It is, therefore, important to quantify variation within the pathogen population by a number of approaches, which can include morphological, genetic, biological and pathogenic perspectives, to enable us to understand what we are dealing with and to support the development and introduction of suitable management strategies.

Since CWD was first discovered in 1927, it has become apparent that it exists in different forms or strains, as coffee species and even varieties of the same species can be resistant to one strain but susceptible to another. This immediately presents a problem for developing a resistant variety that would have to resist infection for many years, during which time it might come into contact with different races or even species of the disease.

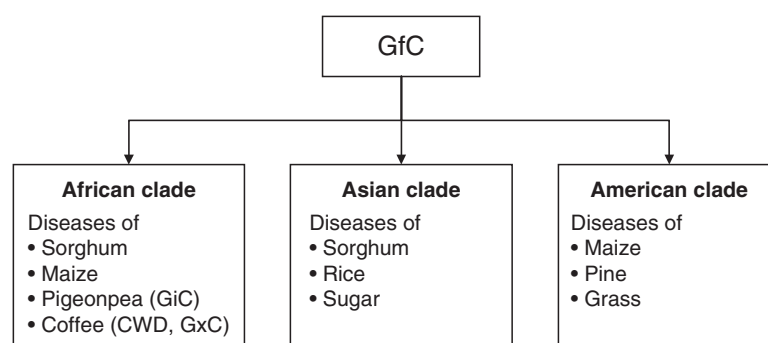
The total CWD strains so far discovered is lumped under the heading *Gibberella xylarioides sensu lato* (s.l., meaning *in the broad sense*, to include all known variants). However, this by itself is an unsatisfactory state of affairs because it lacks precision, and breeders especially need to know how similar or different are the strains present in the region they are responsible for and how likely others may appear, so that they know the scope of the challenge to develop a stable resistant variety.

G. xylarioides s.l. is also known as the *G. xylarioides* species complex (or GxC for short), which belongs to the *G. fujikuroi* species complex (GfC). This means that several plant diseases, about 50 at the last count, are lumped under the GfC heading because they closely resemble a disease of rice called *G. fujikuroi* which was historically the first of the diseases described. This collection of closely related diseases is subdivided into different clades (a clade is a group of individuals with a recognized common ancestor) – and the GxC is part of the African clade. A closely related disease pigeonpea wilt (*Fusarium udum*, sexual phase is called *G. indica*) is also a member of the African clade of GfC (Figure 4.1).

Hence CWD is a complex within a complex, and therein lies much of the problem of understanding this disease; even specialist mycologists have, over the years, made mistakes in the identification, description and classification of strains of CWD, confusing them with closely related diseases, so tiny are the apparent differences.

To the non-mycologist, things are further complicated by the way that the science of mycology differentially names the sexual and asexual forms of the disease: *F. xylarioides* is the asexual, and *G. xylarioides* the sexual or perfect form. Unlike some important *Fusarium* diseases, the sexual form turns out to be quite prevalent in infested fields of coffee and this implies that the disease can quickly develop new characteristics

Figure 4.1: Coffee wilt disease (CWD) is a member of the African clade of the *Gibberella fujikuroi* disease complex (GfC) that together comprises about 50 known plant diseases.



through outbreeding. To date, however, the amount of variability found in the field is very low, which is something of a surprise, but which might reflect its recent emergence from a very restricted genetic pool.

What follows is a synopsis of the work that has been carried out recently under the Regional Coffee Wilt Programme (RCWP) and EU COWIDI research programmes, aimed at more closely defining the differences between the strains. It relies heavily on the PhD thesis of Dr Pascale Lepoint (2006) to which the reader is referred for greater detail.

4.3 Identification Methods

There are three basic ways to distinguish between different individual fungi (accessions) collected in the field:

- Morphological – how they look to the naked eye or under the microscope.
- Biological – if the different individuals interbreed or not. If they do, then they are always regarded, by definition, as the same species.
- Molecular – one or more molecular tests to determine the degree of similarity in the molecular structure, either of a gene, or the DNA between genes.

4.3.1 Morphological differences

Morphological differences are the first to be studied by the mycologist, because if successful, they have the obvious advantages of speed and low cost. It is not possible, however, to distinguish variation between CWD strains by morphological differences of the fungus; spore shape, size, etc. are identical.

Despite claims from historical outbreaks that diverse strains give distinct field symptoms such as different colours of staining of wood fibres, as well as different spore morphologies under the microscope (e.g. Saccas, 1953), these are not now regarded as reliable ways to distinguish strains. As mentioned in Chapter 2, the disease can be present without the usually characteristic blue-black staining; hence, the variations seen in the field are more likely due to different responses by the plant under a range of environmental differences.

Girma (2004) and Lepoint (2006) conclude that only the Arabica and Robusta strains can be distinguished visually in the laboratory by growing the strains from each on a rich agar potato dextrose agar (PDA) medium. An orange pigmentation is apparent on young CWD cultures from Robusta grown on a PDA or cereal agar (CA) medium, which is absent in isolates from Arabica. But beyond this gross level, no visual separation is possible.

The conclusion is simple – morphology alone will not allow any reliable differentiation of CWD strains.

4.3.2 Biological species recognition

One way to resolve how closely strains are related is to mate them, if the sexual form of the disease is available, which is the case for CWD (*G. xylarioides*), which is heterothallic, meaning that there are two forms present, one with the MAT-1 version of the mating gene and the other with the MAT-2 version (see Section 4.3.3.3). If they

produce a viable offspring, it is reasonable to lump them under the same species heading, since this is the principal definition of a species. If they do not cross or they produce sterile offspring, it is reasonable to suggest that they are different species.

With this basic principle and taking advantage of the propensity of CWD to develop the sexual form, Lepoint carried out a series of crossings of different strains in the laboratory. Reproductive success was scored on a scale of 0 to 4: completely sterile (0), protoperithecia only (1), barren perithecia (2), perithecia containing unidentified structures (3) and fully fertile (4).

The results of the crossing experiments (Table 4.1) suggested the presence of at least three biological species (BS1 to BS3).

Table 4.1: Mating matrix for coffee wilt disease (CWD) strains. Reproductive success is scored on a scale of 0 to 4: completely sterile = 0; protoperithecia only = 1; barren perithecia = 2; perithecia containing unidentified structures = 3; fully fertile = 4.

			<i>Gibberella xylarioides</i> complex (GxC)			
<i>Coffea</i> , host of the CWD strain			<i>arabica</i>	<i>canephora</i>	<i>excelsa</i>	<i>canephora</i>
	Region		Ethiopia	DRC	CAR	Guinea, Côte d'Ivoire and CAR
		Biological species (BS)/ sterility group (SG)	BS1	BS2	BS3	SG4
<i>arabica</i>	Ethiopia	BS1	4	2-3	1	1-2
<i>canephora</i>	DRC	BS2		4	1-2	0
<i>excelsa</i>	CAR	BS3			4	0
<i>canephora</i>	Guinea, Côte d'Ivoire and CAR	SG4				0

BS1: The various accessions from Ethiopian Arabica all crossed with themselves (stage 4 fully fertile perithecia), but all other crossings between Arabica and either Robusta or Excelsa strains failed to attain stage 4. Some Robusta strains from Democratic Republic of Congo (DRC) and Uganda from the recent outbreaks did produce perithecia to stages 2 and 3, but this was not enough to count them as the same species. Ethiopian Arabica was thus designated as BS1.

BS2: The various strains collected from Uganda, Tanzania and DRC from the recent upsurge all crossed successfully to give stage 4 perithecia, hence all these strains were regarded as BS2.

BS3: Two strains only, one of unknown provenance (ATCC 15664) and one from Central African Republic (CAR) Excelsa isolated in 1955 (BBA 62457 = DSMZ 62457) successfully crossed, making these BS3.

SG4: There were also a number of completely incompatible strains. All these were Central and West African strains that produced no fertile perithecia, and hence were lumped conservatively into a sterile group SG4, on the basis that they originate from the same host and region.

It should be understood that BS3 and SG4 were collected in the 1950s and 1960s, and hence they may have deteriorated somewhat over long storage and subculturing, and this may have affected their fertility.

Hence, Lepoint distinguishes three species by biological species recognition (BSR) that correspond closely to Arabica, Robusta and Excelsa strains, as well as a residual group of incompatible strains which are lumped as a further species, making four in all. The residual group displays heterogeneity in MAT loci (see below), suggesting that with further research these could be further separated. Further collections are needed in countries such as DRC, Central African Republic (CAR) and Guinea to try to find any traces of the strains of diseases that once flourished there.

4.3.3 Molecular methods

Molecular methods offer the greatest possibilities for differentiation between strains. The main problem is to find parts of the genome to analyse that might correspond to meaningful divisions – i.e. not too coarse-grained that little differentiation is seen, but not so fine-grained that every strain collected is deemed unique.

The story of molecular studies of CWD is one of gradually increasing resolution of what initially looked like an almost identical collection of strains. The most meaningful distinguishing markers are those that relate closely to the ability of putative species to interbreed, but finding such reliable markers can be very difficult.

4.3.3.1 Random amplified polymorphic DNA (RAPD) analysis

RAPD was one of the first molecular techniques used to look for differences between organisms. The technique analyses randomly chosen parts of the genome for differences between genomes of two or more organisms, and can successfully distinguish between species but is now often regarded as too unreliable when it comes to determining between strains of the same species.

Similar approaches were taken by Bieysse (2007) at CIRAD and by Rutherford (2005) at CAB International. Single-enzyme, agarose gel-based amplified fragment length polymorphism (AFLP) analysis and amplification and enzymatic digestion of the intergenic spacer (IGS) region, generally revealed little genetic variation among the CWD isolates, but showed that two distinguishable pathogen populations exist, those collected from Arabica and those from Robusta in Uganda, DRC and Tanzania.

These findings were confirmed by a polymerase chain reaction (PCR) analysis of ISSR DNA sequences (Figure 4.2) and by an analysis of mitochondrial DNA RFLPs (Figure 4.3). As seen in Figure 4.3, the few historical isolates available for molecular analysis were found to differ slightly from those of either the Arabica group or the Robusta group. These isolates, accessions BBA 62457 = DSMZ 62457 (Lepoint's BS3), CBS 25852, CBS74979 (SG4) and ATCC 15664 (BS3), were obtained in the 1950s and 1960s from Robusta and Excelsa trees.

CBS 25852 and CBS 74979 from Côte d'Ivoire and Guinea, respectively, appear to be genetically identical and, according to the mitochondrial DNA analysis, are more closely related to ATCC 15664 (geographic origin unknown) than to DSMZ (BBA)

Figure 4.2: Unweighted pair group method (UPGMA) dendrogram showing the percentage of genetic similarity between *Fusarium* isolates obtained from coffee affected by coffee wilt disease (CWD). The dendrogram was based on mitochondrial DNA restriction fragment length polymorphisms (RFLP) data and was derived using Nei and Li's genetic distance coefficient. Abbreviations: Fx, *F. xylarioides*; Fo, *F. oxysporum*; Fso, *F. solani*; Fst, *F. stilboides*; Fl, *F. lateritium*; Ca, *C. arabica*; Cc, *C. canephora*; Ce, *C. excelsa*; Et, Ethiopia; Ug, Uganda; DRC, Democratic Republic of Congo; Tz, Tanzania; Gu, Guinea; CAR, Central African Republic; IC, Côte d'Ivoire; FEA, French East Africa.

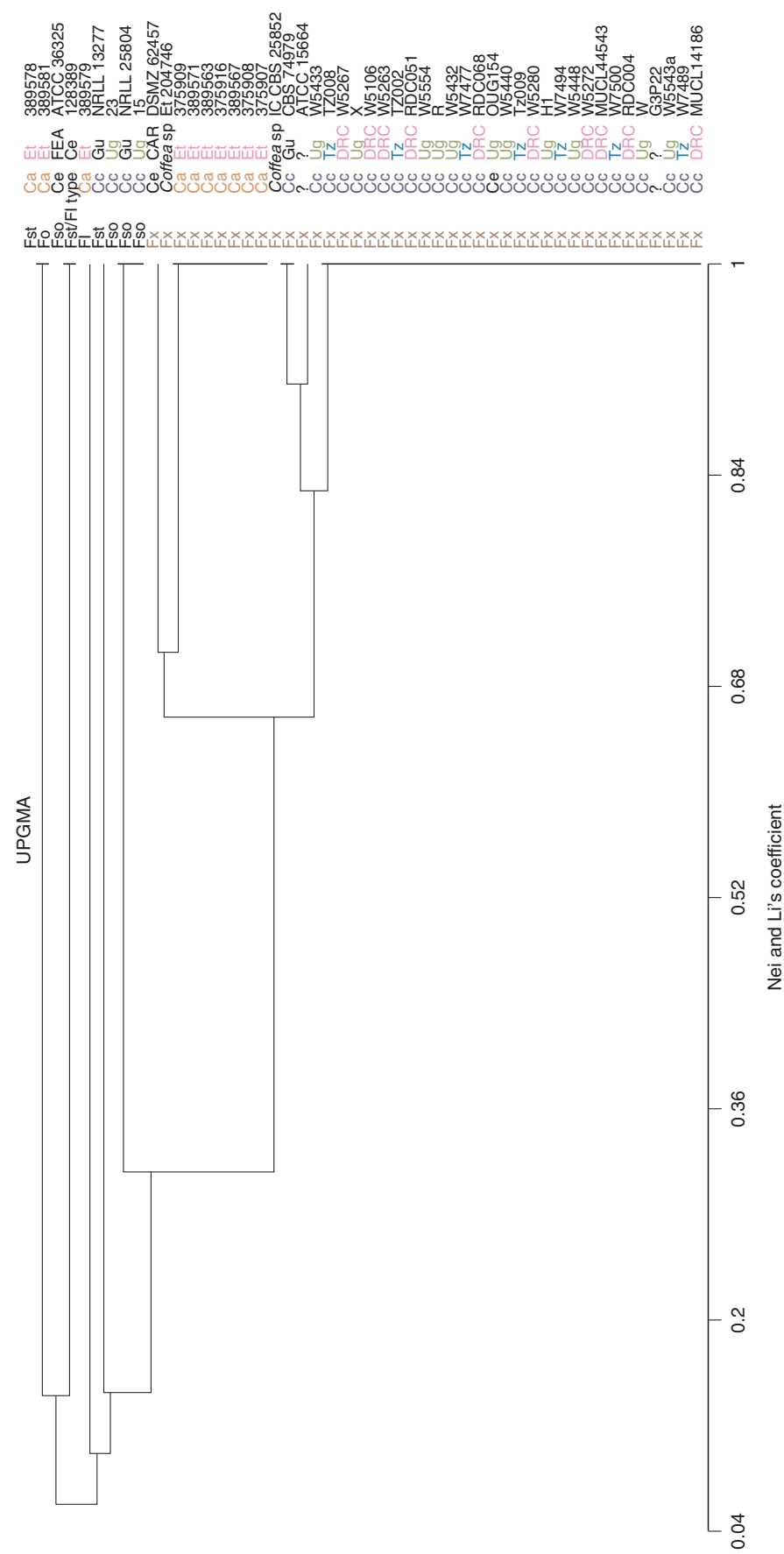
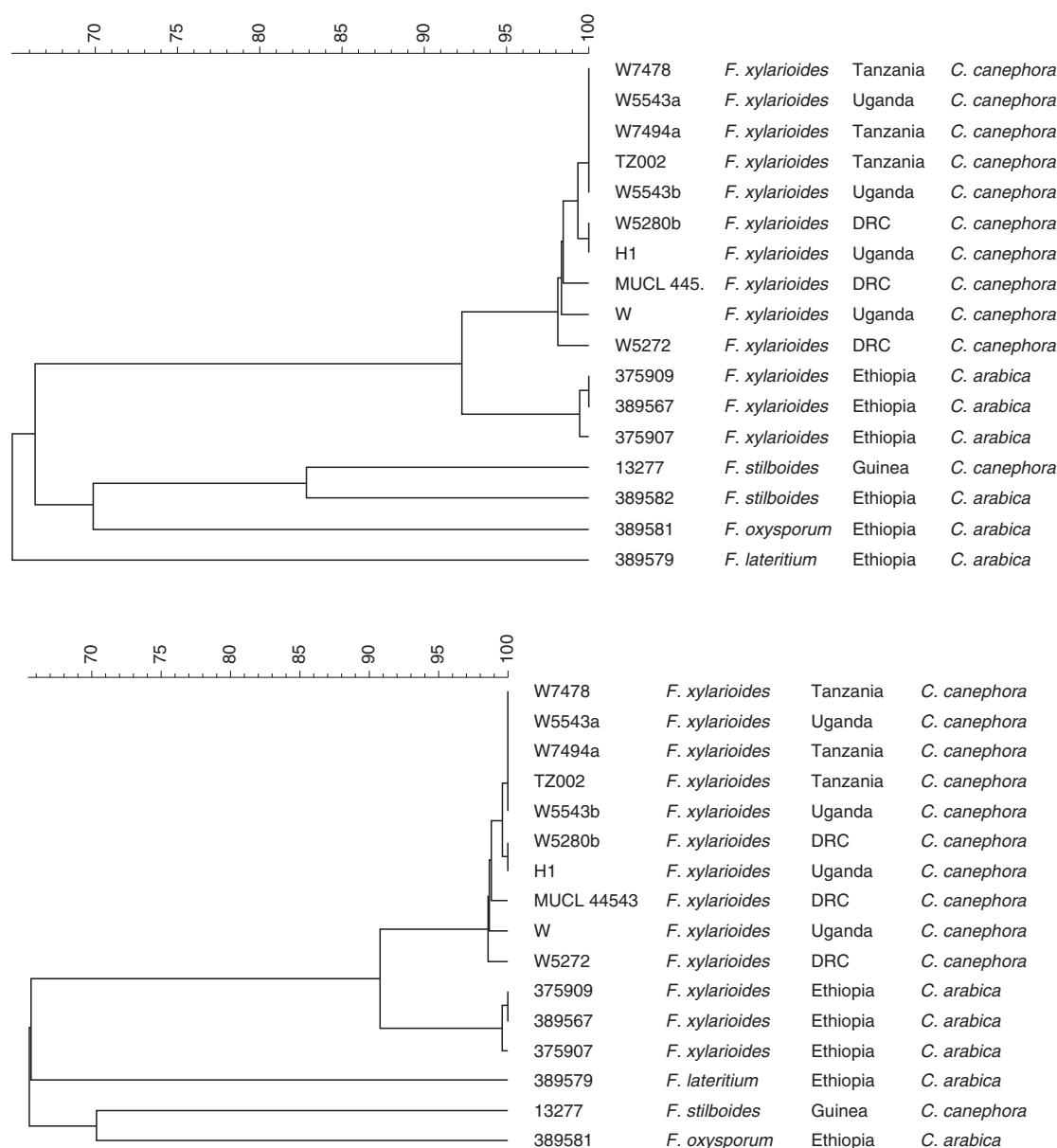


Figure 4.3: Dendrograms showing the genetic similarity between selected isolates of *F. xylarioides* obtained from *C. canephora* and *C. arabica* in Ethiopia, Uganda, DRC, Tanzania, Guinea and other *Fusaria* based on dual enzyme amplified fragment length polymorphisms (AFLP) analysis. The upper and lower dendrograms were generated using the combined AFLP data for three (AFLP ATACC + AFLP ATCTC + AFLP CGCTC) and four (AFLP ATACC + AFLP ATCTC + AFLP CGACC + AFLP CGCTC) primer sets, respectively. Data were analysed using GelComparII software (Applied Maths, Sint-Martens-Latem, Belgium) and similarity matrices calculated using Dice's coefficient. The dendrograms were constructed using the unweighted pair generated method of arithmetic averages (UPGMA). Percentage similarity is shown along the top of each dendrogram.



62457, from CAR. The former three isolates also have more genetic similarity with the Robusta group than DSMZ 62457, which shows almost the same level of dissimilarity to the Arabica group as to the Robusta group.

Furthermore, DSMZ 62457 has also been shown to be uniquely capable of causing CWD on both Arabica and Robusta seedlings (see Chapter 5). These findings are intriguing,

as they suggest that the Arabica group, the Robusta group and the historical isolates could have evolved independently, perhaps due to selection pressures related to geographic isolation or the coffee type cultivated. Alternatively, they may have evolved sequentially, with one or more genetic forms emerging from another.

Hence, through this molecular analysis, sufficient evidence was found to be able to separate strains into three groups. Later, Lepoint (2006) by using RAPD decamer A17 found four haplotypes corresponding to the BS delimited through mating tests as described above.

4.3.3.2 Individual gene studies

The subjectivity of determining potential species boundaries within the GxC can be reduced by the analysis of several genes. Lepoint chose three different genes to look for differences: the translation elongation factor 1- α (*tef* gene), the calmodulin gene and the histone gene. These are all so-called nuclear house-keeping genes that are permanently switched on and are vital for the correct functioning of the cell, and hence differences found are much less likely to be due to the random genetic variation and much more likely to represent significant functional differences.

By analysing these three genes in different strains and using a statistical technique called maximum parsimony analysis with the help of a phylogenetics software package (Swofford, 2001), Lepoint found the following differences:

Translation elongation factor tef 1- α : Based on *tef* analysis, three groupings (clades) were identified, corresponding to (i) BS3, (ii) SG4 and (iii) a third group containing both BS1 and BS2.

Calmodulin CL: Based on a 609bp calmodulin fragment analysis, GxC divides into two clades, clade 1 contains BS1 strains while clade 2 contains BS2, BS3 and SG4 strains.

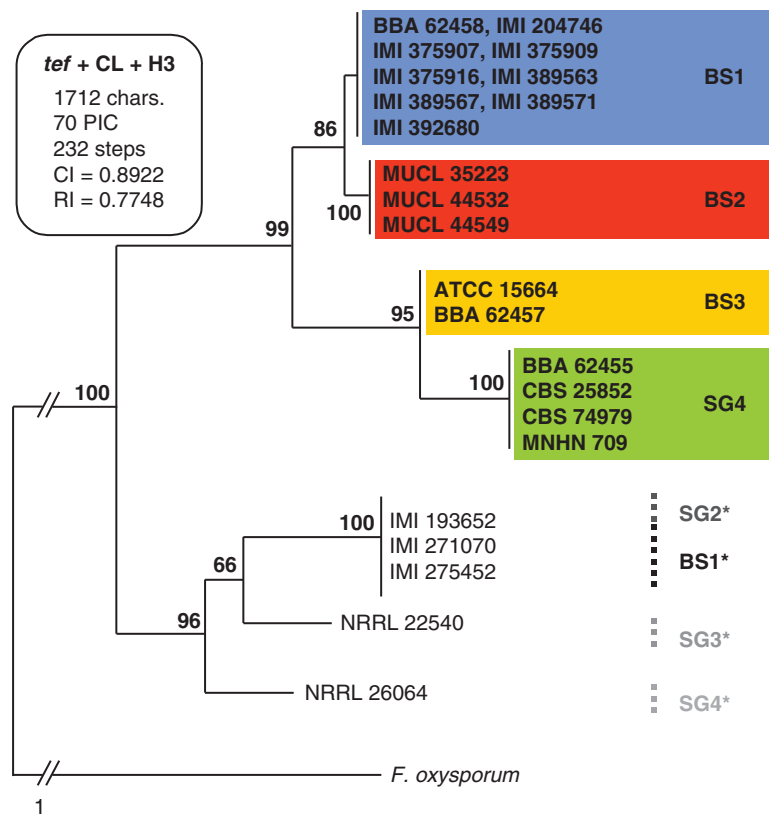
Histone 3 (H3): Phylogenetic analysis of the 459bp H3 sequence revealed six single nucleotide polymorphisms (SNPs). Three clades could be distinguished within GxC, clade 1 containing BS2 strains, clade 2 including all SG4 strains while clade 3 containing BS1 and BS3 strains.

Combining the three: Using all the above, the maximum parsimony analysis suggests that four phylogenetic species (PS) are recognizable within GxC and are consistent with the four BS groups defined in Section 4.3.2, with a further subdivision so that BS1 and BS2 form a sister clade, whereas BS3 and SG4 form another. With additional analysis of strains from GiC, it was also confirmed that GxC and GiC (the *G. indica* complex that contains pigeonpea wilt *F. udum*) are closely related sister clades (Figure 4.4).

4.3.3.3 The mating gene MAT

A further possible means of distinguishing between species of fungi is to study differences in the mating gene (MAT) that is known to control the fungal sexual reproduction. It is likely that differences in the MAT structure between strains are related to their ability to cross and hence is an interesting place to look for taxonomic markers. Indeed, previous work on a non-*Fusarium* species has shown that a single mutation in MAT can greatly reduce sexual compatibility (e.g. Ferreira *et al.*, 1998).

Figure 4.4: Maximum-parsimony phylogram based on the combined autosomal (*tef* + CL + H3) data set of representative strains of the *G. xylarioides* (GxC) and *G. indica* (GiC) complexes from diverse geographical and host origins. Biological species (BS) and sterility groups (SG) defined in carrot agar crosses are indicated by coloured boxes for the GxC and by a dotted line for the GiC next to terminal clades resolved. Trees were generated with PAUP v.4.0 b10 (Swofford, 2001) using *F. oxysporum* as out-group and available NCBI sequences for closely related species belonging to the GfC African clade (O'Donnell *et al.*, 1998). Bootstrap values based on 1000 replications are indicated in percentages at internodes when replication frequencies exceed 50%. (Courtesy of P. Lepoint.)



A single MAT gene of two different forms (alleles), MAT1-1 and MAT1-2, has been previously described for the *Fusarium* genome (Coppin *et al.*, 1997; Kronstad and Staben, 1997). Each individual of heterothallic species such as GxC has either the MAT1-1 or MAT1-2 version of the gene. Perithecia will only be formed when crossing takes place between individuals that have the opposite MAT1-1 and MAT1-2 versions.

The MAT1-1 gene consists of three subsections or open-reading frames: MAT1-1-1, MAT1-1-2 and MAT1-1-3, whereas the MAT1-2 gene has only one frame called MAT1-2-1.

Lepoint, therefore, looked at the structure of these genes in different strains of GxC to find out if differences might correspond to those found from other taxonomic methods described above.

4.3.3.4 Results of differences in the MAT gene for GxC strains identified

A phylogenetic species (PS1) was identified, containing all the Ethiopian Arabica strains. This could be further subdivided into two groups based on differences in MAT1-1-1 and MAT1-1-3 frames.

Robusta associated strains from Uganda and DRC from the recent upsurge formed a second PS2 group.

PS3 includes only two strains, a CAR strain from Excelsa coffee from the 1950s (BBA 62457) and a strain of unknown provenance collected in 1964.

A further species group, PS4, contains strains from CAR, Côte d'Ivoire and Guinea all isolated in the 1950s and 1960s, but the designation of this group is less certain, because of conflicting tendencies from analyses of MAT1-2, MAT1-1-2, MAT1-1-1 and MAT1-1-3. This area needs further analysis with more material from these countries, if it could be found.

4.4 Conclusions

It is now clear that the various strains of the disease (i.e. those separated in space and time) are distinguishable by various techniques, even though these differences are quite small and invisible to the eye, even with a microscope. The overall lack of variation found by molecular studies suggests a recent evolution of this disease.

A mystery has been as to why there is so little variation within the current outbreak in DRC, Uganda and Tanzania, given that the sexual form of the disease is evident. It suggests that the disease has emerged from a single location that escaped eradication and that maybe through severe selection pressure due to resistant varieties, only one variant has survived. This corresponds closely to what we know about the history of the recent outbreak and hence gives us confidence that the molecular techniques are robust.

It also suggests that as long as this low variability of the current epidemic persists, it will help the development of durable resistance through breeding, since there is less likelihood for the narrow genetic base of the disease to find combinations to overcome the resistance of new coffee cultivars that are developed. Conversely, it also means that we should closely monitor the disease in the field to detect if new strains may be emerging, which will tend to happen when resistant varieties become widely grown, since this will raise selection pressure on the disease.

Some of the genes studied were very similar between strains and species and, not surprisingly, the MAT genes (used for mating) seem to show the most difference between species, the acid test for belonging to a species being if the putative strains of the disease are found to be sexually compatible. Hence, distinguishing MAT genes combined with mating tests seems the best way to distinguish between different species of the cryptic GxC, and more generally the GfC.

However, the techniques covered above confirm that the late 20th century epidemic in DRC, Uganda and Tanzania is identical to the strain MUCL 14186 isolated in 1960 in DRC by Meyer, and the strain is sexually compatible with the current strain. In addition, it seems certain that the disease has remained there at a low level, but to what

extent that strain was responsible for all the damage of the initial outbreaks is not clear because the original type specimen from the 1927 outbreak (Steyaert, 1948) is lost, so we will probably never know the true identity of that original detection.

Conversely, the new techniques also show that two species isolated from collections are no longer found in the field – they could have been eradicated, or quite possibly, as in the case of the genesis of the recent outbreak which started in DRC, they might still be present at currently undetected levels – or as Lepoint puts it ‘just waiting to be discovered’.

The sexual state can be produced in the laboratory by confronting two compatible strains, but not all strains are compatible, leading to the conclusion that there are at least three separate diseases present and possibly more. Indeed, based on all the methods covered above, Lepoint *et al.* (2005) and Lepoint (2006) propose that CWD may effectively consist of four different species:

- *Gibberella abyssiniae* – responsible for the Ethiopian outbreaks on Arabica.
- *Gibberella congoensis* – responsible for the current outbreaks in DRC, Uganda and Tanzania.
- *Fusarium guineensis* – the asexual form, responsible for the original mid-20th-century outbreaks in Guinea, Côte d’Ivoire and possibly CAR.
- *Gibberella xylarioides* – the originally described disease from the CAR *Coffea excelsa* strain BBA 62457 (= DSMZ 62457) and strain ATCC 15664 collected during the first CWD epidemic.

Significantly, Lepoint (2006) makes it clear that the re-emergence of CWD in DRC in the 1980s, and its subsequent spread to Uganda and Tanzania, is not a new form, but identical to the strain (now called *G. congoensis*) found there in the 1960s.

However, as Lepoint (2006) points out:

Species concepts or criteria delimiting a species are very controversial and much debated. Morphological and biological species recognition have been widely applied to a large set of fungi; however, their pertinence for speciation has been questioned and superseded with the advent of molecular biology and phylogenetic species recognition.

Hence the suggestion to define so many new species is unlikely to be easily accepted by some scientists. In a way, it depends on how you define a species, but the classical way must surely count – i.e. that if two strains do not produce viable offspring then they belong to different species. And when this is supported by molecular differences, then the case becomes more substantial.

As it stands then, we are far from the last word on the matter, but the research has marked a substantial advance in our understanding of CWD. Quite apart from the scientific debate about whether CWD is a mix of species or merely strains, the possibility that we are dealing with several diseases is a different way of looking at the problem, which may have some practical advantages.

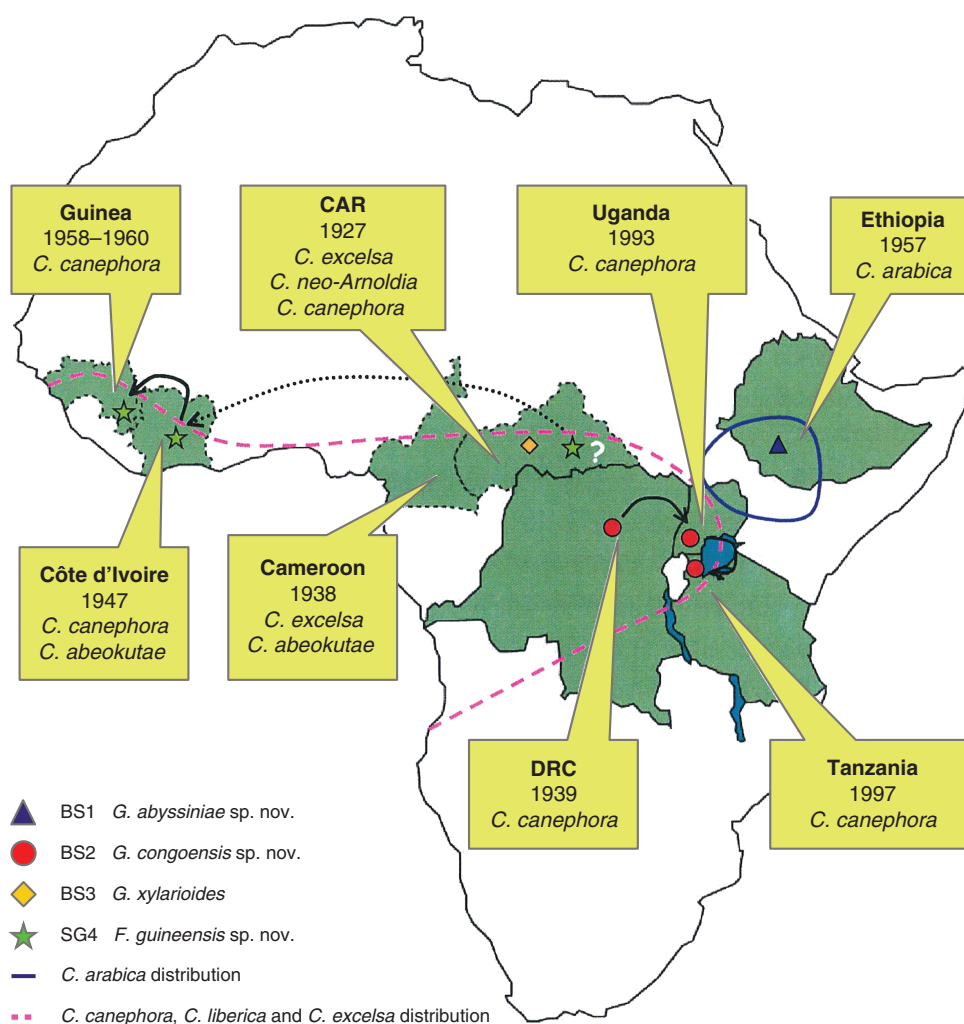
For example, if there are three or four diseases, however closely they are related, then it is much easier to believe that they are unlikely to have evolved from a common ancestor recently (i.e. within the last 100 years) and so their discovery

on coffee has not been due to a single event that has spread, but rather a series of separate events. A synoptic map of the current distribution of the diseases can be seen in Figure 4.5.

Perhaps the most likely scenario is that these diseases have existed on wild coffee species (or even non-coffee species) in natural conditions for a long time, and that as coffee production expanded these previously well-contained diseases had the opportunity to break out. Similar events are known to have occurred in the past, most notoriously for instance for human diseases such as Ebola and Marburg disease.

It also suggests that such an event could happen again, independent from any of the previous outbreaks. It is very likely that other strains do exist, so there is the potential for greater variation in the future that could break down the rather narrow resistant base of the Robusta strains now being developed.

Figure 4.5: Synoptic map of the development of coffee wilt disease (CWD) across Africa since first discovery in 1927 (modified after Lepoint, 2006). Countries with broken line borders have reported no CWD since the 1950s. Countries currently affected are shown with full line borders. For each of the eight countries, date of first appearance and coffee species affected appear under the country's name. Symbols indicate names of the wilt diseases according to Lepoint (2006).



A case in point for a new outbreak might be a country like Angola that is trying to rebuild its coffee industry after years of strife – disturbing long-abandoned plots or deforesting new areas, that have reverted to forest, could conceivably bring newly planted coffee into contact with a new disease form. Regular surveys of Angolan coffee zones, therefore, could spot a problem early and avoid much unnecessary destruction.

The upshot of the research presented in this chapter, therefore, is that there should be regular surveys of coffee-producing regions to gauge the overall health of coffee stock and look for unusual occurrences. It is possible that other pockets of non-epidemic infestation exist in Cameroon, CAR, Côte d'Ivoire, etc. that need to be monitored. The monitoring should include regular molecular screening, which would easily now detect unusual strains that might need immediate and drastic action in the field.

This is then a significant advance in the otherwise sad history of CWD: molecular tests give accurate information that if acted upon decisively, could in future help to halt the advance of a new disease before it becomes an epidemic.