AN ASSESSMENT OF THE GENETIC DIVERSITY AND ORIGIN OF THE INVASIVE WEED *CHROMOLAENA ODORATA* (L.) KING AND ROBINSON IN SOUTH AFRICA

A thesis submitted in fulfilment of the

requirements for the degree of

MASTER OF SCIENCE

of

RHODES UNIVERSITY

by

INGE VON SENGER

July 2002

Abstract

Chromolaena odorata (L.) King and Robinson is an alien invasive weed to most of the Old World tropical regions of the earth, including South Africa where it is morphologically distinct from most other *C. odorata* plants examined from both its native and invasive range. It is thought that these morphological differences are related to difficulties encountered in successful establishment of biological control agents on the South African population of *C. odorata*. It has been postulated that the source population of the South African population will harbour potential biocontrol agents that will be suited to successful establishment on the South African plants. Several morphological, cytological and isozyme studies have been attempted to identify the source population of the South African population, but these have failed to identify the origin of the South African population.

In this dissertation two PCR-based methods were attempted, in an investigation into whether the morphological differences and difficulties in establishment of biocontrol agents have a genetic basis. The two techniques attempted were: Inter Simple Sequence Repeat (ISSR) amplification, and DNA sequencing. Results could not be obtained using the ISSR method, and the reason for this was not discovered despite extensive trials. The internal transcribed spacer region and the external transcribed spacer region sequences were obtained from five samples, and compared. It was found that the ETS region gave more phylogenetic signal at the intraspecific level than the ITS region. However, due to difficulties in amplification of the external transcribed spacer region, work here focussed on obtaining Internal Transcribed Spacer sequences for 61 samples.

Each of the samples sequenced had a unique ITS sequence, displaying a high level of intraspecific genetic diversity. The degree of this diversity is discussed with reference to the possible influences of polyploidy and concerted evolution on genetic structure. The ITS data indicated that some of the physical traits used to define 'morphotypes' of *C. odorata* were not correlated to genotype. From discussion and comparison of morphological character distributions and the ITS-based phylogeography it is suggested that the geographical origin of the South African population is Greater Antilelan, rather than from the continents of North and South America, which is where the Australasian, West African and Mauritian infestations are suggested to have originated.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2,4-D	2,4-dichlorophenoxyacetic acid
AFLP	Amplified Fragment Length Polymorphism
ARC	Agricultural Research Council
cpDNA	chloroplast DNA
CTAB	hexadecyltrimethyl-ammonium bromide
DNA	deoxyribonucleic acid
dNTP	dinucleotide triphosphate
EDTA	Ethylenediamine tetra-acetic acid
ETS	External Transcribed Spacer
GA	Greater Antilles
IGS	Intergeninc Spacer
ISSR	Inter Simple Sequence Repeat
ITS	Internal Transcribed Spacer
KZN	KwaZulu-Natal
Mg^{2+}	magnesium ion
MST	Minimum Spanning Tree
mtDNA	mitochondrial DNA
ncpGS	Nuclear chloroplast-expressed Glutamine Synthetase
nDNA	nuclear DNA
nrDNA	nuclear ribosomal DNA
NTS	Nontranscribed Spacer
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PPRI	Plant Protection Research Institute
PVP	Polyvinylpyrrolidone
rDNA	Ribosomal DNA
SA	South Africa
SSLP	Simple Sequence Length Polymorphism
UBC	University of British Columbia
UCT	University of Cape Town
USA	United States of America
VNTR	Variable Number Tandem Repeat

Acknowledgements

I would like to express my gratitude to the many people who helped me with information or ideas for my thesis.

I am grateful to my supervisors, Dr. N. Barker and Dr. C. Zachariades for their invaluable input and advice, and to the NFR and ARC-PPRI for their financial support.

I am also grateful to all the students and staff at the Botany Department at Rhodes University for their support and help when I needed it. I am especially grateful to Irma Knevel for her sympathetic ear.

Finally, I am grateful to Karl Flowers for his neverending support and advice, and to my parents for being there when I needed them most.

CHAPTER 1

CHROMOLAENA ODORATA (ASTERACEAE) AS AN INVASIVE ALIEN WEED

The Approaching Biological-Extinction Event

Many biologists predict that coming decades will see the loss of large numbers of species (Hoch and Raven, 1995; Myers and Knoll, 2001) that will be as a result of humankind's influence in the environment. The primary effect that humankind will have is that species numbers will begin to dwindle in an extinction event that could be equal in magnitude to the last so-called Big Five extinction events that Earth has experienced. Although of very serious concern, present-day extinctions have not yet reached the intensities seen in the Big Five mass extinctions of the geological past, in which more than 50% of the subset of relatively abundant marine invertebrate genera were removed (Jablonski, 2001), though estimates suggest that at least one fifth of all species will disappear during this forthcomming extinction event (Hoch and Raven, 1995).

The magnitude of the effect that mass extinctions have on the earth's biota through the evolutionary opportunities they create is disproportionate to the relatively low number of species that are lost during the actual extinction event (less than 5% of all species in previous extinction events) (Erwin, 2001). Mass extinctions have been instrumental in creating the earth's biodiversity. Ways in which these extinctions achieved such biodiversity include differential survivorship, the disruption or preservation of evolutionary trends and ecosystem organisation, as well as the promotion of taxonomic and morphological diversifications (often in unexpected directions) after the destruction or reduction of previously dominant taxonomic groups (Jablonski, 2001).

It is not a well-recognised phenomenon that, in the longer term, these extinctions will have altered not only biological diversity but also the evolutionary processes by which diversity is generated (Myers and Knoll, 2001). Mooney and Cleland (2001) state that as a result of the changes that our existence have had on the environment, we are now causing the development of a whole new cosmopolitan assemblage of organisms across the surface of the Earth with large consequences not only for the functioning of ecosystems but also for the future evolutionary trajectory of life.

There will be several first-order effects stemming from the pending biotic crisis, some of which are alien invasions and the mixing of biotas [as is already evident - invasive species are thought to be one of the proximate causes of biodiversity loss worldwide (Czech and Krausman, 1997)] as well as progressive depletion and homogenisation of biotas (Myers and Knoll, 2001), such as that caused by invasive weeds.

The effects of this biotic homogenisation on evolution may depend in part on how it is achieved (Jablonski, 2001). Homogenisation via elimination of endemics will leave a residue of already widespread taxa that may be relatively resistant to geographical isolation and rapid diversification, whereas homogenisation via range expansion may more readily promote the origin and diversification of new endemic taxa (Jablonski, 2001).

Another effect of the biotic crisis could be that there is a proliferation of opportunistic species, which could lead to an ecology that is characterised as a "pest and weed" ecology (Myers and Knoll, 2001). Newsome and Noble (1986) state that we can expect a progressive simplification of the world's biota as superior species are transported across the world towards more suitable habitats and displace the local species. Though this simplification event will result in some extinctions, not all 'indigenous' species will be removed, and thus an increase in species richness may be expected (Newsome and Noble, 1986).

An issue to be considered is whether mankind knows enough about species interactions to mitigate the loss of biological diversity in the forthcoming 'crisis' (Myers and Knoll, 2001).

The Effect of Weeds on Ecology

Weedy plant species could play an important role in determining the degree of biodiversity that remains (or evolves) in the future. Thus, a weed can be defined as a plant, (native or introduced), that grows in a situation where it has detrimental effects on conservation areas, on mankind, or on his environment (Harley and Forno, 1992).

Before trans-continental travel became possible, biogeographic boundaries acted to isolate continental biotas for millions of years. However, since then there has been continual and regular breaching of these great biogeographic boundaries, allowing species to 'travel' to novel biogeographic realms (Mooney and Cleland, 2001). It has been estimated that over the past 500 years, invasive species have come to dominate three percent of the earth's ice-free surface (Mack, 1985, cited in Mooney and Cleland, 2001). Some countries have been more severely invaded than others – there are now as many alien established plants in New Zealand as there are native species (Mooney and

Cleland, 2001), and, of the 2 834 species listed in the New Flora of the British Isles, 1 264 (ca. 45%) are aliens (Abbott, 1992, cited in Mooney and Cleland, 2001).

According to Crawley (1983) the 'food limitation hypothesis' states that some herbivores (e.g., insects and mammals) regulate the abundance of some plant populations. Thus these herbivores help to limit the abundance of plant populations. However, when plants are introduced into a new place, their natural enemies are usually left behind (Harley and Forno, 1992) and when physical and biotic factors are favourable in the new environment, introduced plants (that are not accompanied by their natural enemies) will flourish and become much more abundant than they were in their native range because of the lack of natural enemies. Plants that have become successfully naturalised in new environments often have detrimental consequences to the new environment. Such plants are termed 'alien invasive weeds'.

In addition to greater numbers of species crossing borders and becoming immediately invasive, there is also a build up in the invasive potential of non-native species that have been established in a region for several years without becoming invasive. As more plants of the same species are (usually unknowingly) imported, the species' effective population size and genetic diversity increases. Introduced species may stay at a fairly low population size for years and then explode at some later date – the so called lag effect (Mooney and Cleland, 2001). In addition, environmental change (both biotic and abiotic) after establishment can catalyse invasiveness of a species. Other species may never become invasive.

Invaders themselves often evolve in response to their interactions with natives, as well as in response to the new abiotic environment. Flexibility in behaviour and mutualistic interactions can aid in the success of invaders in their new environment (Mooney and Cleland, 2001). One of the ways in which invaders appear to adapt to their new environments is by shifting their behaviour and/or traits, e.g. in populations of invasive species, the individuals are often larger in their new territory than in their native land. These size differences can be attributed to the consequences of natural selection for greater competitive capacity after release from herbivore attack and the need to produce defensive compounds (Mooney and Cleland, 2001).

Exotic species can impact on the abundance and evolutionary processes of native species in several ways. They competitively suppress or exclude native species, cause niche displacement and they can cause changes in disease incidences (Tilman and Lehman, 2001; Mooney and Cleland, 2001). Invasives can also cause changes in nutrient cycles, and induce changes in the physical habitat (Tilman and Lehman, 2001). An example can be found in the South African 'fynbos' vegetation, which is adapted to regular, cool fires. The introduction of invasive species such as bluegum, black wattle,

and *Pinus* spp., which create more aggressive fires than typical fynbos fires, prompted human intervention in fire prevention. As a result, seedling germination and nutrient recycling in the fynbos is adversely affected by the lack of fires. Any fires that do occur are hotter than those that the fynbos vegetation is adapted to, and thus cause severe damage to the flora as well as fauna (Tilman and Lehman, 2001).

Invasive species have been known to affect the evolutionary pathways of native species by hybridisation with them, as well as introgression, competition, predation (by invasive fauna) and, in some cases, even extinction (Mooney and Cleland, 2001). Exotic species are the second largest cause of native species of the United States being listed as endangered (Tilman and Lehman, 2001).

Environmental changes caused by mankind's activities are creating regional combinations of environmental conditions that, within the next 50 – 100 years, may fall outside the envelope within which many of the plants in existence today have evolved (Tilman and Lehman, 2001). These environmental modifications might become a greater cause of global species extinction than direct habitat destruction, as the changes could favour a few species that would competitively displace many of the other species from a region (Tilman and Lehman, 2001).

In South Africa itself, half of the land surface may experience a novel and far warmer and drier bioclimate in the future (Rutherford *et al*., 1999). This change in climate may result in the almost complete disappearance of the species-rich Succulent Karoo biome, and at least partial species loss to the fynbos biome (Rutherford *et al*., 1999). Most of the 16 centres of endemism in South Africa may lose as much as 60% of their current area (Rutherford, *et al*., 1999). In the long term, it is the weedy taxa that could become the dominant in the novel conditions imposed by global change, and they could become the progenitors of a series of new species that are progressively less weedy and diversify to produce new species that are suited to the new conditions (Tilman, and Lehman, 2001).

Weeds Collapse Mutualisims

Mutualism is an integral part of the proper functioning of an ecosystem. When invasive species disrupt these relationships, they affect the success of the species in the mutualistic relationship. Mutualisms would seem to be a barrier to the success of a single player of a partnership becoming an invasive species. There is some evidence for this in the fact that non-mycorrhizal plant taxa (such as the Brassicaceae and the Chenopodiaceae) are particularly successful weeds (Mooney and Cleland, 2001). Sometimes, however, the arrival of one non-native species is followed by the subsequent arrival of a co-evolved facilitator, thereby increasing the success of each in its new environment. This has happened with *Pinus* spp. and their mutualistic mycorrhizal fungi in the Southern Hemisphere

(Mooney and Cleland, 2001). The pines were introduced as timber trees to South Africa, and when they would not grow properly, soil from Europe was imported to South Africa, to be included in the soil when the saplings were planted. The imported soil contained mycorrhizae that are usually found in mutualism with pine trees. As a result of the unwitting import of mycorrhizae, the pines flourished, and subsequently became invasive. In other cases the tightness of mutualisms is not as great as supposed and other species in the new habitat can play the required role for the invader (e.g. pollination).

With the mixing of biota that is expected to occur in the forthcoming extinction event and thus new interaction potentials there is the great possibility of new kinds of mutualistic relationships evolving. One such example is the European pines being dispersed in South Africa by alien American squirrels (Richardson *et al.*, 2000, cited by Mooney and Cleland, 2001).

There are also instances of an invasive species disrupting mutualistic relationships. Native seed-harvesting ants disperse the seeds of certain *Protea* spp. in South Africa. These native ants have been displaced by Argentine ants that are not successful in dispersing the *Protea* seeds to suitable germination micro-sites, thus potentially leading to the extinction of rare and endemic *Protea* spp. (Mooney and Cleland, 2001).

Hybridisation of Weeds

In the case of invasive species, hybridisation with native species can cause a loss of fitness in the native species and even a threat of extinction (Rhymer and Simberloff, 1996, cited in Mooney and Cleland, 2001). McMillan and Wilcove (1994, cited in Mooney and Cleland, 2001) have documented that, of the 24 species that were listed as endangered in the United States and have subsequently become extinct, three were as a result of hybridisation with alien species.

It has been proposed that many alien weeds are not invasive when they first enter a new habitat, and that in many cases it is only after hybridisation with a local or other alien plant that invasiveness arises (Ellstrand and Schierenbeck, 2000). Over 10% of morphological characters in primary hybrids are novel, and over 30% are novel in later generations (Rieseberg, 1995). Such novel character states may increase a plant's fitness and thus enable it to become successfully invasive. This theory need not be limited to hybrids of different species but could also be applied to hybrids of plants from different populations within a species; introduction of distantly related individuals of the same species from different parts of its range may yield enough of an evolutionary stimulus to increase the fitness of the progeny, and induce invasiveness (Ellstrand and Schierenbeck, 2000).

Ellstrand and Schierenbeck also mention that if hybridisation among populations of the same taxa plays an important role in the evolution of invasiveness, then it is possible to expect certain correlates for the appearance of invasiveness. First, invasiveness would be expected to occur after multiple introductions of a species, because multiple introductions would be necessary for providing different genotypes from disparate sources (species that are intentionally introduced would have an advantage in this regard). Second, invasiveness would be expected to occur after a lag time, during which hybridisation and selection would act to create and increase invasive genotypes. Finally, if the evolution of invasiveness followed a bout of hybridisation between well-differentiated populations, then the resulting populations should be expected to be more genetically diverse than were their progenitors (Ellstrand and Schierenbeck, 2000).

It would thus appear that hybridisation is an important process in the evolution of invasiveness. However only a fraction of hybridisation events will lead to the evolution of invasiveness, and hybridisation is not the only evolutionary pathway to invasiveness and not all alien species have evolved invasiveness (Ellstrand and Schierenbeck, 2001).

Polyploidy

While hybridisation with invaders can be a threat to species integrity (particularly of native species), it can also be a source of new variation and a source of new species, through polyploidy (Grant, 1953). It is now commonly accepted that polyploidy is a highly effective evolutionary mechanism for introducing new plant species, promoting their persistence and survival, and ultimately increasing the diversity of plant species (Cook *et al.*,1998; Otto and Witton, 2000; Ramsey and Schemske, 1998; Soltis and Soltis, 1999; Wendel, 2000). It is estimated that between 20% and 70% of angiosperm species are polyploid.

Polyploidy is the presence of multiples of a basic chromosome number (Cook *et al.*, 1998), which arises via gametic non-reduction, somatic doubling in meristem tissue, and polyspermy (Ramsey and Schemske, 1998; Otto and Whitton, 2000). Allopolyploids were initially defined as arising through the process of interspecific hybridisation and chromosome doubling, while autopolyploids were defined as arise from intraspecific parents. At this time, chromosome pairing behaviour was believed to be a reliable indicator of chromosome homology, and therefore early workers took the frequency of multivalent formation at synapsis to be a criterion for distinguishing auto- and allopolyploidy. However, it has been recognised that some polyploids of known hybrid origin exhibit multivatent pairing, and that bivalent pairing is prevalent in some non-hybrid polyploids (Ramsey and Schemske, 1998).

In order to clarify ploidization, segmental allopolyploids were defined to arise from parents with partially divergent chromosome arrangements such that some chromosomal regions are homologous between the parents and others are homoeologous (Soltis and Soltis, 2000), or that they posess chromosome pairing characteristics of autopolyplooids (Ramsey and Schemske, 1998). Amphiploids were defined as all polyploids that combine chromosome compliments of distinct species and 'autopolyploid' was a term reserved for polyploids that arose within single populations or between ecotypes or races of a single species (Ramsey and Schemske, 1998).

Under traditional views, polyploidy was considered largely an evolutionary dead end; however a growing body of evidence suggests that recurrent polyploidisation is the rule rather than the exception (Cook *et al.*, 1998), and that polyploidy can confer evolutionary advantages on plant species (Doyle, *et al.*, 1999). Polyploids often occupy habitats different from those of their diploid parents, and have been proposed to be superior colonisers to diploids (Roose and Gotlieb, 1976; Soltis and Soltis, 2000).

As an example, *Spartina alterniflora* from the east coast of North America was introduced into Southampton in England in shipping ballast in the early 19th century (Mooney and Cleland, 2001). It subsequently hybridised with the comparatively less fit local species *S. maritima* (*S. alterniflora* has a higher pollen output and greater male fitness than the native species, and it occupies lower intertidal habitats). The hybrid is sterile, but subsequently underwent chromosome doubling to produce the new fertile allopolyploid species, *S. anglica*. This new species has become very aggressive and occupies large areas of the coastline of the British Isles while at the same time the original invader, *S. alterniflora*, and the native *S. maritima* have maintained limited distributions (Mooney and Cleland, 2001). The new polyploid evidently has characteristics that enable it to occupy bare tidal flats that were not available to the parents (Mooney and Cleland, 2001).

Most polyploid species of plants that have been examined with molecular markers have been shown to be polyphyletic, having arisen multiple times from the same diploid species (Soltis and Soltis, 2000). Such recurrent formation of a polyploid species has implications for the taxonomy of polyploids, the genetic diversity of polyploid 'species', and for an understanding of the ease with which and rate at which polyploidisation can occur (Soltis and Soltis, 2000). The concept of recurrent formation of polyploids forces one to consider polyploid species not as genetically uniform (as previous models of polyploidy formation imply) but as genetically variable. This concept calls into question the meaning and validity of the term "polyploid species" (Soltis and Soltis, 2000).

Some aspects of polyploid success have been attributed to their superior colonising ability, which may involve higher selfing rates than those of the diploid parents (Soltis and Soltis, 2000). Theoretical

models predict reduced inbreeding depression in polyploids relative to their diploid parents because of the buffering effect of additional genomes: deleterious alleles are masked by the extra genomes. Both allopolyploids and autopolyploids are expected to have reduced inbreeding depression and the magnitude of inbreeding depression is negatively correlated with selfing rates in diploid angiosperms and gymnosperms (Soltis and Soltis, 2000, and literature therein). With this lower inbreeding depression, highly selfing polyploids may be better colonisers, and have broader ecological aptitudes than their diploid progenitors, explaining the prevalence of polyploids on the list of the world's worst weeds (Soltis and Soltis, 2000). Another possible source of genetic novelty (and thus adaptability) in polyploids is genome rearrangements, as discussed in Soltis and Soltis (2000).

Control of Weeds

As a result of the evolutionary successfulness of weeds, they are dominant plants, competing successfully against agricultural crops and (in the case of invasive weeds) against indigenous plants, threatening the less successful and usually more desirable plants. For this reason, various methods of weed control have been devised.

The oldest way of controlling weeds is by mechanical means, where the weeds are removed by hand or with the help of tools. This is time-consuming and labour-intensive, and therefore faster and more efficient weed-control methods have been sought. It was the discovery of chemical herbicides such arsenic pentoxide and 2,4-D that were markedly selective that was particularly welcome at the time of the Second World War, when there was a shortage of farm labour, and wages were particularly high (Salisbury, 1964; Fryer and Evans, 1968). Farmers and land managers enthusiastically embraced these poisons. Unfortunately it was soon realised that the inorganic poisons were toxic to humans and livestock and so safer herbicides were sought. Advances in plant physiology and pesticide chemistry facilitated the development of selective herbicides that killed only certain plants while allowing others to live (Salisbury, 1964; Fryer and Evans, 1968; Harley and Forno, 1992).

In recent years, however, the benefits and advantages of modern herbicides have been reduced by public concern over the detrimental effects of herbicides on the environment as well as to the consumers. In developing countries (and subsistence agriculture communities) herbicide costs are often prohibitive; here the costs of weed control become even more critical (Wilson and McFadyen, 2000). Almost directly as a result of these concerns, agriculturalists have started to look for methods that reduce or eliminate the need for costly chemical herbicides (Harley and Forno, 1992).

Biocontrol

The most common non-chemical method of weed control is biological control (or biocontrol). It was defined by DeBach (1964) as 'the study and utilisation of parasites, predators and pathogens for the regulation of host population densities' (Harley and Forno, 1992). Put more simply, biological control is the control of one organism by another (Lucas *et al.*, 1992).

The main premise under which biological control functions is that any plant, animal, or other pest that is introduced (whether intentionally or unintentionally) into a new environment will flourish (to the detriment of the local indigenous fauna and flora) if none of the natural enemies of the introduced organism accompany it. Biological control programmes aim to introduce some of the natural enemies as 'biocontrol agents' to limit the pests' abundance. Biocontrol agents are usually arthropods or plant pathogens but other natural enemies, for example nematodes or fungi, can be used (Harley and Forno, 1992).

Biocontrol is usually not employed by itself, and it rarely results in sufficient control by itself, but is used as part of an integrated control programme. It reduces growth and reproduction to a level where other control methods become more cost-effective.

Although the introduction of an organism into a new environment always provides risks as to the impact it will have on other organisms, it is particularly disconcerting when organisms that are introduced to control the activities of an unwanted invader instead do collateral damage to other species, even driving them to extinction. This is apparently the case with the introduction of the rosy wolf snail, *Euglandia rosea*, which was imported into Hawaii in 1958 to control the giant African snail, *Achatina fulica*. Unfortunately, *E. rosea* did not restrict its predatory activity to the African snail but also attacked the rare native Hawaiian snails, apparently driving them to extinction. Between 1977 and 1987 *E. rosea* also pushed the endemic tree snails of the island of Moorea to extinction. There is another extinction crisis in the making with the movement of *Cactoblastis cactorum* from its point of introduction for the control of *Opuntia* in the Caribbean, to a trajectory that will bring it to a centre of diversity of *Opuntia* in Mexico (Mooney and Cleland, 2001, and literature therein).

Besides direct predation, unexpected associations can develop between native/alien plants and insects. Some plants develop secondary compounds that the arthropods find beneficial, e.g., *Zonocerus variegatus*, an indigenous insect to west Africa has become a pest since the introduction of *Chromolaena odorata*, as the weed is a readily available source of pyrrolizidine alkaloids that the insect uses in pheromone production (Boppré, 1991).

History of Biological Control of Weeds

One of the first uses of biological control principles was by the Chinese, who used the ant *Oecophylla smaragdina* to control caterpillars and large boring beetles in citrus groves (Harley and Forno, 1992). The first known establishment of a natural enemy that was purposely moved from its native country to another country is credited to de Maudave (cited in Harley and Forno, 1992), who introduced the mynah bird from India to Mauritius in order to control locusts, and in Europe a predacious pentatomid bug *Picromerus bidens* was introduced against bed bugs as early as 1776 (deBach, 1974, cited in Harley and Forno, 1992).

The first major programme for biological control of a weed was in 1902 when fruit- and flower-feeding insects collected in Mexico were introduced into Hawaii for the biological control of *Lantana camara* (Perkins and Swezey, 1924, cited in Harley and Forno, 1992). These insects effectively halted the spread of *L. camara* on drier parts of the islands of Hawaii, and were later sent to Fiji, Australia, India, East Africa, and South Africa. Australia later sent insects to Java (Goeden, 1978, cited in Harley and Forno, 1992).

By 1999, 949 biological control programmes had begun, aimed at controlling weedy plant species (Julien and Griffiths, 1999). By 1992, Australia, the United States of America, Canada and South Africa, together with the International Institute of Biological Control based in the United Kingdom, were foremost in the search for biological control organisms and the implementation of biological control (Harley and Forno, 1992).

Biocontrol in South Africa

Biological control of alien invasive weeds in South Africa began in 1913 when the cochineal insect *Dactylopius ceylonicus* was introduced to control the Brazilian cactus *Opuntia vulgaris* (Olckers, 1999). Since then, more than 85 species of biological control agents have been released onto ca. 47 weed species, making South Africa the third most active country in biological control after the United States of America and Australia (Olckers, 1999).

Most of South Africa's invasive alien weeds were introduced intentionally; some plants were initially intended, or were subsequently fostered for commercial uses, such as forestry (*e.g. Pinus* spp., *Eucalyptus* spp, *Acacia melanoxylon*, and *A. mearnsii*); agroforestry (*e.g. Prosopsis* spp. and *A. saligna*); horticulture (*e.g. Jacaranda mimosifolia*, *Melia azedarach* and *Lantana camara*); dune stabilisation (*e.g. A. cyclops*, *A. longifolia* and barrier plants such as *Pereskia aculeata* and *Caesalpinia*

decapetala) and fruit production (*e.g. Psidium guajava* and *Opuntia ficus-indica*) (Zimmermann and Neser, 1999).

The success rate of South African biological control compares favourably with that of other countries. At least 17% of the alien invasive weeds in South Africa are considered to be under complete biological control (where no other control measures are required to maintain the weed populations at acceptably low levels; Olckers, 1999). Thirty percent of the invasive weeds are under substantial control (where conventional control measures are still needed but at reduced rates) and approximately 8% are still under negligible control (where there has been virtually no reduction in conventional control methods, despite damage inflicted by biological control agents Olckers, 1999). For the remaining 45% (ca. 21 species), either the programmes have not yet been initiated, or releases of agents have been too recent to allow meaningful assessments (Olckers, 1999).

The only alien invasive plant to South Africa for which there has thus far been no success despite attempts in establishing biocontrol agents is *Chromolaena. odorata*.

Chaco and Narasimham (1988) state, "No species alone can be expected to suppress [*C. odorata*]. Introduction, simultaneously if possible, of several species causing different types of injury should be attempted." Liggitt (1983) also suggests that biological control agents alone will not control *C. odorata*, but will rather be used alongside control methods that are currently available for use in an integrated approach to the control of this weed.

C. odorata is regarded as the worst alien invasive plant species in the subtropical regions of South Africa, and the one which is increasing its range most rapidly (Liggitt, 1983). It is an aggressive, fast-growing weed (Liggitt, 1983), which is said to have one of the greatest impacts of all alien invasive plants in the Old World tropics and subtropics (Zachariades *et al.*, 1999). It is regarded as the worst alien invasive plant in the KwaZulu-Natal province (Coetsee, 1995). *C. odorata* has invaded much of the tropical and sub-tropical forests on the east coast of South Africa, as well as bushveld areas in the Mpumalanga and Swaziland (Coetsee, 1995). Already more than 50% of the coastal forests along the Eastern Cape (previously Transkei) coast have been invaded (Coetsee, 1995).

Plant diversity has been reduced in reserves, affecting all other dependant life forms (Coetsee, 1995), and the KwaZulu-Natal Wildlife Services lack sufficient resources to combat the plant (Sandberg, 2000).

Measures for the Control of Chromolaena *odorata*

At present, the most effective means of controlling *C. odorata* in South Africa is through mechanical weed eradication techniques, one of which is to slash the top-growth and uproot the subterranean portion by mattock or hoe. This 'slash-and-uproot' method, while being effective, has the disadvantages of disturbing the soil, increasing soil erosion, and exposing weed seeds for further germination. In addition, it requires twice as much labour (man-days/hectare) than the alternate 'slash and spray coppice' method (Goodall and Erasmus, 1996). In this second method, the top-growth is cut back and the coppice sprayed with an industrial herbicide. A third, promising (though costly) method involves the use of tractor-mounted fire-fighting equipment to spray the weed with herbicides (Goodall, 1997). All mechanical methods of control of *C. odorata* require follow-up procedures, involving slashing up to four times a year, to prevent re-growth and to kill any germinating seedlings (Sandberg, 2000).

The reported combustibility of *C. odorata* (Muniappan, 1996), even when green (Goodall and Erasmus, 1996), has led farmers and land managers to avoid the use of fire in controlling *C. odorata*. Recently, however, field tests have shown mortality rates of up to 94% on fire-treated *C. odorata* stands (Goodall and Erasmus, 1996; Goodall 2000), indicating that fire may be a relatively cost-effective means of bringing this weed under control in carefully managed areas. The use of fires to control and even eliminate *C. odorata* is limited to grasslands, which are adapted to, and can recover from fires. Burning in forests would only serve to damage the natural flora, and provide additional disturbed sites for invasion of *C. odorata* (Sandberg, 2000).

The use of mechanical and chemical clearing methods such as those mentioned above is time-consuming and therefore costly, and herbicides are also expensive. The only feasible means of bringing the weed *C. odorata* under control without draining South Africa's resources is through Integrated Management Planning (IMP), where conventional clearing practices are used to compliment alternative methods such as biological control (Zimmermann and Neser, 1999).

Over 240 arthropods have been recorded on *C. odorata* plants in Trinidad (Cruttwell, 1974). Of these only a few have been tested on the South African *C. odorata* for host specificity, and some of these were rejected because their host range was too wide. Only three species have passed host-specificity testing, and been released. Of these, one (*Paracheutes pseudoinsulata*) did not establish, but this was probably unrelated to incompatibility with the host plant. The second (*Paracheutes aurata aurata*) was taken from a different *Chromolaena* species, *C. jujuiensis*, and failed to establish, possibly due to host-incompatibility with the 'new association' species *C. odorata*. The third (*P.*

insulata) is in the process of being released, and it is too soon to tell what its success rate will be (Zachariades *et al*., 1999).

Taxonomy of *Chromolaena odorata*

The relationship between the host plant and control agent is central to the science of weed biological control. From a practical viewpoint, it is critical to be certain of the identity of the target weed as well as of the candidate agent (cf. O'Hanlon *et al*., 1999). Correct identification can be used to pinpoint where in the native range an introduced weed comes from and can thus focus the search for potential control agents. Failure to correctly identify a weed or potential agent can lead to lack of agent establishment as well as incomplete control because of high specificity between potential biocontrol agents and weeds, particularly in cases involving multiple biotypes of a single weed species (O'Hanlon *et al*., 1999).

Some of the problems encountered in the attempts to successfully establish biocontrol agents on the invasive species *C. odorata* have a taxonomic basis (O'Hanlon *et al*, 1999); in particular the precise identification of the weed invasive to South Africa that has been called *C. odorata* has in the past been questioned.

There was previously some concern that the *C. odorata* invasive in South Africa was a species other than *C. odorata.* This possibility was due to some morphological distinctions (Zachariades *et al*., 1999; see later discussion), or possibly even a hybrid of unknown parentage (see Ellstrand and Schierenbeck, 2000). However, H. Robinson of the Smithsonian Institute (Washington) has confirmed (pers. comm. to C. Zachariades, 2000) that the plant invasive in South Africa is definitely *C. odorata*.

The first taxonomic description of the plant now known as *Chromolaena odorata*, in written literature was by Pluckenet (1692, cited by Gautier, 1992), who called it "*Eupatoria Conyzoides folio molli incano, capitulus magnin, Americana"*. It was later called "*Eupatorium Odoratum hirsutum; foliis ovato-acuminatis, basim versus crenatis, opposites; floribus comosis"* by Browne (1756, cited by Gautier, 1992). Soon after, Linnaeus (1759, cited by Gautier, 1992) gave it the binomial name *Eupatorium odoratum*.

It was only recently that King and Robinson (1970) revised and split the genus *Eupatorium* and placed the taxa that were within the grouping variously known as *Osmia*, or Section Cylindrocephala (including the species *Eupatorium odoratum*) into the new genus *Chromolaena* (Robinson and King, 1985). *Eupatorium odoratum* was thus renamed *Chromolaena odorata* (L.) R. King and H. Robinson. However, it has been said that the elevation of infrageneric groups in *Eupatorium* by King and

Robinson is contentious, and is generally not accepted (Veldkamp, 1999). Other names for *C. odorata* are listed in Appendix 1.

Chromolaena belongs to the tribe Eupatorieae, of the Asteraceae. The Eupatorieae is a well-defined, mostly New World tribe with white, reddish or bluish flowers that lack ray florets (Robinson and King, 1977, cited by McFadyen, 1988a). The genus *Chromolaena* contains 129 species, all of which are native to South and Central America and the West Indies (King and Robinson, 1970).

As already mentioned, polyploidy may play an important role in the success of weedy species (Soltis and Soltis, 2000). The literature does not agree on a definite basal chromosome number and ploidy levels in the genus *Chromolaena* (or old *Eupatorium*) (Grant, 1953; Powell and King, 1969; Vos, 1989). The basic chromosome number for Asteraceae is x = 9 and for the genus *Chromolaena* n = 10 (x + 1) (Powell and King, 1969). The species *Chromolaena odorata* is recorded as having a chromosome number of n = ca. 40 by Powell and King (1969), but has also been reported to have chromosome numbers ranging from 2n = 51 to 2n = 60, and is even thought to be an allohexaploid, with n = 60 (see literature cited in Vos, 1989).

The ploidy levels of *C. odorata* may be a factor in its successful invasiveness, as polyploids are regarded to have several phenotypic differences to diploids, which are considered advantageous. Here, some of those advantages are briefly presented, but for a more thorough examination of the advantages and disadvantages of polyploidy to phenotype, see Otto and Whitton (2000). Polyploids frequentlly have larger seeds than diploids, and this can lead to higher rates of early development, and polyploidy in plants is sometimes associated with larger overall size (Otto and Whitton, 2000). Among the reproductive shifts that occur with polyploidy is apomixis, however while most apomictic plants are polyploids, not all polyploids are apomictic. *C. odorata* is considered to be at least facultatively apomictic (Vos, 1989; Rambuda, 2001). Polyploid plants are also frequently said to have broader ecological tolerances than their diploid progenitors (Otto and Whitton, 2000).

Native Distribution of Chromolaena odorata

C. odorata is native to the tropical and subtropical regions of the Neotropics. The exact extent of the native range of *C. odorata* is unclear, as its use as a fish poison resulted in an increased distribution through the Neotropics in Pre-Columbian times (H. Robinson, pers. comm. to C. Zachariades). Nevertheless, it is currently found from the south east USA and Mexico, through the Greater Antilles (GA) to South America, as far south as northern Argentina (Britton, 1965, cited in McFadyen, 1988b; King and Robinson, 1970; H. Robinson, pers. comm. to C. Zachariades, 2000).

Morphology of *Chromolaena odorata*

C. odorata is a perennial, multi-stemmed shrub, with stems branching freely and laterals that develop in pairs from the auxiliary buds (McFadyen and Skarratt, 1996). The older stems are brown and woody near the base, while tips and growing shoots are green and succulent. Old stems may die and be replaced by new shoots from ground level. The crown of each plant is single and does not divide even when many-stemmed, and suckering from the root does not occur (McFadyen, 1988a,b; McFadyen and Skarratt, 1996). *C. odorata* has two noticeable growth habits: the first is a bushy habit, where a very dense thicket two to three meters high is formed in the open or in almost pure stands. A large plant may have 20 or more stems of varying size, often bent under the weight of their branches, shading a ground area of ca. 3.5m^2. The second is a creeping habit [commonly observed in South Africa (C. Zachariades, pers. comm.)] where it climbs nearby vegetation and often reaches the top of a canopy [up to 25m heights (Sandberg, 2000)], where there is higher light intensity (McFadyen, 1988b; Gautier, 1992; McFadyen, and Skarratt, 1996; Binggeli, 1997). The added height gained by this clambering habit may also facilitate more efficient wind dispersal of the seeds.

C. odorata is heliophilous and has a fast growth rate due to efficient allocation of resources (Gautier, 1992). It possesses an underground organ at the base of the stem and considerable starch reserves in the crown which ensures the plant's survival through fire, drought or mechanical damage such as coppicing (McFadyen 1988a; Binggeli, 1997), though the underground organ is not obviously present in the South African *C. odorata*.

Sexual reproduction is first initiated when the plant is a year old. The terminal cymes bear ca. 70 insect-pollinated capitulae, and the capitula are borne in heads of 20 to 60 at the tips of all stems, branches, and auxiliary shoots, and the onset of flowering prevents further growth (McFadyen, 1988b; Binggeli, 1997). Each capitulum contains 17, 30 to 36, or 63 to 70 florets; although exhibiting this wide range between plants, the number is nearly constant for any one plant (McFadyen, 1988b). The flowers are white or pale bluish-lilac, and form masses covering the whole surface of the bush, making *C. odorata* conspicuous when in bloom (McFadyen, 1988b). Flowering initiation appears to be mainly related to the onset of the main dry season (McFadyen, 1988b; Binggeli, 1997). The small fruits (weight: 0.2mg) mature in about a month (Binggeli, 1997).

C. odorata is capable of producing vast quantities of seed, and the fruits (cypsela) are typically wind-dispersed as dry and windy weather is necessary for fruit release.

When flowering is over, most of the leaves wither and fall. New leaves and shoots grow from the old leaf axils, and the dead terminal parts of the stems drop off (McFadyen, 1988b). The extent of leaf-

fall and the rate of re-growth depend on the moisture available (McFadyen, 1988b). On the cleared slopes in the foothills of the Northern Range in Trinidad, bush fires are common at this season (when old leaves and stems have become dry and very flammable) and frequently destroy all above-ground growth. Where it is invasive, *C. odorata* appears to burn readily, and fires in a stand of this species can severely damage indigenous vegetation growing with or near to it.

As a result of a cycle of die-back and re-growth, old stems are gradually shaded by new growth above them. These old stems then suffer a die-back phase (McFadyen, 1988b). Old bushes thus form a tangled mass of old and new stems with green shoots and branches in all directions (McFadyen, 1988b). Witkowski and Wilson (2000) have shown that the density of *C. odorata* in monospecific stands decreases with age (especially after 15 years) (see also Slaats *et al*., 1998), possibly because the old growth prevents new growth, or perhaps because of allelopathic effects of the weed (see section on Review of the Invasive Character of *Chromolaena odorata*). The maximum lifespan of a *C. odorata* plant is not known (McFadyen, 1988b).

Morphological Variation

Some of the morphological features of *C. odorata* vary considerably over its native range (Neser, 1996, Zachariades *et al*., 1999), and as a result, plants that have invaded one area of the Paleotropics have a different morphology to plants that have invaded elsewhere. The differences are often clustered, and have led to the use of the terms 'form' or 'morphotype' to describe a group of morphological variations evident in a certain area. Within South Africa, *C odorata* is morphologically homogenous, and the plants in West Africa and Asia are morphologically identical to each other, yet distinct from the South African plants. Some of the differences are noted below, with particular reference to the morphology of the 'form' of *C. odorata* invasive to South Africa.

The two youngest pairs of leaves of the vegetative shoots of the West African and Asian forms of *C. odorata* have a characteristic purple pigment toward the bases (Sheldrick, 1968), while in the South African form the purple pigment is absent but the first and second pair of leaves may sometimes become red (C. Zachariades, pers. comm.).

The leaves and stems of *C. odorata* in South Africa are less pubescent than those from other countries (Scott *et al*., 1998), and as a result, the leaves appear to be a brighter green in South Africa (Zachariades, 1999). In general, the South African and Jamaican plants share leaf morphologies (C. Zachariades pers. comm.). The West African and Asian plants have bigger and darker leaves than the South African form of *C. odorata* (C. Zachariades, pers. comm.).

The odour of crushed leaves of South African *C. odorata* is different to the form common in West Africa and Asia (Zachariades *et al.*, 1999).

The South African form of *C. odorata* has whitish to cream-coloured capitula, while *C. odorata* from western Africa and Asia has pale mauve, pale blue, or whitish capitula (Holm *et al.*, 1977; Kluge, 1990, and literature therein). The flowers from Trinidad share a similarity with the South African *C. odorata* in that they have white or pale bluish-lilac capitulae (McFadyen, 1988b). In most neotropical countries it seems that flower colour varies from white to pale blue.

It appears that most forms of *C. odorata* have a deep taproot system that allows it to regenerate strongly (Holm *et al.*, 1977; Chaco and Narisimham, 1988). However, the South African *C. odorata* has a shallow, fibrous root system (Kluge, 1990; Leslie and Spotila, 2001), similar to plants in Trinidad, where the root system is fibrous, and does not penetrate the substrate beyond 20 to 30cm (McFadyen, 1988b).

The growth habit of the South African *C. odorata* is more upright, and is more scrambling than other forms (C, Zachariades, pers. comm.), except for some Jamaican plants, which are very similar in growth habit and foliar morphology to the South African *C. odorata* (Zachariades *et al.*, 1999). It is of importance to note here that several forms or morphotypes of *C. odorata* have been noted growing in Jamaica, ranging in morphologies from plants almost identical to the South African *C. odorata*, to plants more similar to those from Florida or West Africa (C. Zachariades, pers. comm.).

A summary of these morphological differences between *C. odorata* plants invasive to Australasia and West Africa versus South Africa is presented in Table 1.1.

It is possible that some of the difficulties encountered in attempts to establish biocontrol agents in South African *C. odorata* are related to the uniqueness of the South African *C. odorata*'s morphology. These morphological differences may be correlated with biochemical differences that would have additional detrimental effects on biocontrol attempts, as many insects and pathogens are sensitive to the chemistry as well as morphology of their host plant (Boppré, 1991).

Table 1.1 Summary of Morphological Variation within *Chromolaena odorata* in Invaded Regions. (Morphology of *C. odorata* in the Neotropics is very variable, and includes all character states mentioned here.)

Character	Australasia and West Africa	South(ern) Africa
Leaves	Purple pigment; duller, more pubescent; homologous odour	Reddish pigment; less pubescent, brighter leaves; different odour
Flowers	Pale mauve/pale blue/whitish	Whitish – cream
Roots	Deep taproot	Fibrous roots
Growth Habit	Very variable	Possibly more upright and scrambling

Ecological Tolerance of *Chromolaena odorata*

C. odorata's distribution can be explained with reference to its major bio-climatic limitations; it cannot withstand frost, and to a lesser extent, low rainfall (Goodall and Erasmus, 1996). *C. odorata* thrives in regions with rainfall of 1000 to 2000mm p.a. and within temperatures of 20°C and 37°C (Muniappan and Marutani, 1988; McFadyen and Skarratt, 1996). In more general terms, *C. odorata* is confined to within the latitudes 30° north and south, and survives at altitudes from sea level to 1000m (Muniappan and Marutani, 1988).

Within a bio-climatic region, light-intensity influences the distribution of *C. odorata*. It is not a shade-tolerant species, and as a result it is not capable of penetrating thick, undisturbed native forest (Muniappan and Marutani, 1988). This is beneficial to the continued existence of pristine forests in the tropical parts of the world where it has invaded. However it can invade forests where gaps in the forest canopy (from tree-fall) allow sufficient sunlight for *C odorata* germination and growth. *C. odorata* has become well adapted to the partial shade of forest margins (McFadyen and Skarratt, 1996), where it poses a considerable threat to the indigenous vegetation and general ecological health of these ecotones (Liggitt, 1983); *C. odorata* is highly flammable, and when burned can cause severe damage to fire-sensitive forest edges (Goodall, 2000). In its native habitat, it acts as a pioneer species, being shaded out by secondary vegetation, whereas in invaded areas it prevents the regeneration of this vegetation.

C. odorata grows best on well-drained sites (McFadyen 1988b; Chaco and Narasimham, 1988; Kluge, 1990) but it has been known to grow on a wide variety of substrates, including heavy clay (Liggitt, 1983) and chemically poor soils (De Rouw, 1991; Slaats *et al.*, 1998). Its presence appears to be reduced in water-logged soils (Gautier, 1996).

In its native distribution range, *C. odorata* is common in most habitats except in undisturbed rainforest. It is seldom weedy and is never the target of specific weed control measures. It does not invade pasture or compete successfully with plantation crops. This reduced aggressiveness is due to attack by a large complex of insects, other arthropods and diseases (fungal and bacterial), together with competition with related plants (Ambika and Jayachandra, 1990). However, due to the absence of these species-specific biotic factors, *C. odorata* has been able to thoroughly colonise parts of Africa and Asia where it is a very successful, noxious exotic weed (Ambika and Jayachandra, 1990).

In non-native areas, *C. odorata* appears to flourish in disturbed habitats, particularly areas with a slash and burn agriculture (Binggeli, 1997). It is an opportunistic plant, and will invade riparian ecologies where the soil has been disturbed by floods. It will also re-invade where land has been cleared of other invasive alien weeds – this secondarily invasive habit of *C. odorata* adds to the difficulty in controlling it.

Review of the Invasive Character of *Chromolaena odorata*

C. odorata is a very successful invasive weed: it has so quickly and efficiently colonised land in Southern Africa that its success is threatening the existence of ecosystems, and is severely reducing the efficiency of land-use in Southern Africa. The maximum rate of spread of *C. odorata* in South Africa recorded for the period 1975 – 1980 was in the region of 2000% (Liggitt, 1983). One could thus be tempted to declare that *C. odorata* is as perfect (as far as invasive potential is concerned) as any weed can be.

An ideal weed, according to Baker (1965, cited in Newsome and Noble, 1986) can be summarised as a plastic perennial which will germinate in a wide range of conditions, grow quickly, flower early, is self-compatible, produces many seeds which disperse widely, reproduces vegetatively and is a good competitor. See Table 1.2.

No plant will possess all the characters in Table 1.2, but it does not need to in order to be successful (Newsome and Noble, 1986). The mobility of modern society has probably changed the characteristics of the ideal invader; the ability to disperse great distances is no longer as important as

it was even a millennium ago, since many species will be translocated from one habitat to another at regular intervals intentionally or unintentionally by humans (Newsome and Noble, 1986).

Table 1.2 The characteristics of an ideal weed (from Baker, 1965, cited in Newsome and Noble, 1986), and a comparison of the characteristics of *Chromolaena odorata.* (√ = presence; X = absence of ideal weed characteristics)

Characteristics of an Ideal Weed	Corresponding Characteristics of Chromolaena as a Weed
Has no special environmental requirements for germination	X Needs light (De Rouw, 1991), and moist soil (Liggitt, 1983).
Has discontinuous germination (self-controlled) and great longevity of seed.	X Poor seed-bank longevity (Yadav and Tripathi, 1981).
Shows rapid seedling growth	√ Liggitt (1983).
Spends only a short period of time in the vegetative condition before beginning to flower	√ Liggitt (1983).
Maintains a continuous seed production for as long as growing conditions permit	X Flowers only in the dry season (McFadyen, 1988b; Binggeli, 1997).
Is self-compatible, but not obligatorily self-pollinated or apomictic	√ Apomictic (Vos, 1989; Rambuda, 2001).
When cross-pollinated, this can be achieved by a non-specialised flower visitor or by wind	√ Insect Pollinated (McFadyen, 1988b; Binggeli, 1997).
Has very high seed output in favourable environmental circumstances	√ Blackmore (1998).
Can produce some seed in a very wide range of environmental circumstances. Has high tolerance of (and plasticity in face of) climatic and edaphic variation.	√ Tolerant to edaphic variation (Liggitt, 1983; McFadyen, 1988b; Chaco and Narasimham, 1988; Kluge, 1990; De Rouw, 1991; Slaats *et al.*, 1998), but (X) not climatic (Muniappan and Marutani, 1988;McFadyen and Skarratt, 1996).
Has special adaptations for both long-distance and short-distance dispersal	√ Burrs → long distance on animals, humans, or machines (Blackmore, 1998). √ Wind-dispersal → short distances (Blackmore, 1998).
If perennial, has vigorous vegetative reproduction	X Branches rarely take root when soil is moist (Liggitt, 1983)
If perennial, has brittleness at the lower nodes or of the rhizomes or rootstocks	√ Old branches are woody at the base (McFadyen, 1988a,b; McFadyen and Skarratt, 1996).
If perennial, shows an ability to regenerate from severed portions of the root-stock	√ Liggitt, 1983; McFadyen 1988a; Binggelli, 1997).
Has ability to compete by special means: rosette formation, choking growth, exochrine production (but no fouling of soil for itself), etc.	√ Smothering habit (C. Zachariades, pers. comm.), and allelopathic (Ambika and Jayachandra, 1980; Yadav and Tripathi, 1981; Liggitt, 1983; Muniappan and Mauritani, 1988)

Table 1.2 would seem to indicate that *C. odorata* is indeed nearly a perfect weed. The traits mentioned above that make *C. odorata* such a successful weed are discussed in further detail below.

Adult *C. odorata* plants produce phytotoxins, mostly in the leaves (Ambika and Jayachandra, 1980). These toxins affect seedlings of other plants, as well as of *C. odorata*, by inhibiting the growth and germination of, and even killing, the seedlings (Yadav and Tripathi, 1981; Liggitt, 1983; Muniappan and Marutani, 1988). The mortality risk of plants subjected to the allelopathic effects of *C. odorata* decreases with age (Yadav and Tripathi, 1981).

It has been noted that seed production of adult plants decreases when other plants of the same species are close by (Yadav and Tripathi, 1981), and when a stand of *C. odorata* plants becomes old (>15 yr) (Witkowski and Wilson, 2000). Whether these observations are related to each other (i.e. competition between adjacent plants), or to the allelopathic effects of *C. odorata* has not been investigated.

Some typical control methods used when trying to clear stands of *C. odorata* are slashing or coppicing (Goodall and Erasmus, 1996) and burning (Goodall 2000). However, the large starch reserves in the root of *C. odorata* allow the plant to coppice again easily after slashing or fire (Liggitt, 1983; McFadyen, 1988a; Binggeli, 1997). As a result, continuous effort is required to keep the plants under sufficient control. Herbicides kill plants easily, but for resource-poor farmers who cannot afford them, mechanical control mechanisms like these are the only options available to them.

It appears that the South African *C. odorata* is less resistant to fire damage than elsewhere, as Goodall (2000) has found that a carefully regulated burning regime can clear *C. odorata* in South African grasslands. There is no such report of the control of *C. odorata* by fire from any other country where it has become a pest, except in the Ivory Coast, where regular burning appears to keep the weed at bay (Gautier, 1996). A common theme to the reports of Goodall (2000) and Gautier (1996) is that sufficient carbon reserves (in the form of other vegetation) must be present as fuel to allow the fire to burn at high enough temperatures to negatively affect *C. odorata*. Several papers mention that *C. odorata* is removed as a fallow plant by slash-and-burn practices, where the plants soon regenerate to become the dominant fallow plant species (*e.g.* De Rouw. 1991; Roder *et al.*, 1995; Slaats *et al.*, 1998), but whether this slash and burn regime reduces the weed's fitness is not noted.

C. odorata is a prolific seed producer.

"It is believed that the rapid spread of this species [*C. odorata*] is directly related to the extensive seed production and wind dispersal architecture of the seeds." (Blackmore, 1998)

It is estimated that in a year, a single plant can produce from 93 000 (Weerakoon, 1972, cited in Blackmore, 1998) to 1 600 000 (Wilson, 1995, cited in Blackmore, 1998) seeds. The seeds are thought to be formed apomictically (Vos, 1989; Rambuda, 2001), possibly facultatively, as butterflies

and other potentially pollinating insects have been observed on the flowers (C. Zachariades, pers. comm.). Strict apomixis would allow a single plant to effectively colonise an area by seed production, and facultative apomixis would facilitate the development of genetic diversity in the population.

The seeds of *C. odorata* are fire sensitive (Mbalo and Witkowski, 1997) and have little or no seed dormancy (Yadav and Tripathi, 1981; McFadyen, 1888a); Yadav and Tripathi (1981) found that there were no seeds in the soil below 2 cm (see also Ismail *et al.*, 1996), and two months before seed fall began the total seed count was zero. The seedlings germinate as soon as the soil is moist, and grow quickly to produce viable seed after a single season (Liggitt, 1983). The seedlings are weak competitors, and there is a high seedling mortality rate (Yadav and Tripathi, 1991). The seeds of *C. odorata* require light to germinate, and therefore will not germinate under the dense canopy of forests (De Rouw, 1991). The apparent sensitivity of *C. odorata* seeds and seedlings to fatalities does not appear to reduce the strong colonising ability of *C. odorata* due to the production of such high numbers of seeds.

Harmful Properties of *Chromolaena odorata*

C. odorata has different effects on different environments or communities. Some of the more deleterious effects that this weed has world wide are mentioned here.

The primary danger that *C. odorata* presents in South Africa is to the environment. The weed is such a successful competitor that it poses a serious threat to eight indigenous vegetation types of South Africa (Liggitt, 1983; Goodall and Erasmus, 1996). Two vegetation types are particularly vulnerable to infestation by the weed, namely Ngongoni Veld and Coastal Forest and Thornveld (particularly dune and coastal forests). *C. odorata* is also present in the following medium to dry bushveld types: Zululand Thornveld, Lowveld Sour Bushveld, Lowveld and Valley Bushveld, but infestations here are sporadic and vary in intensity. The weed also occurs in two arid bushveld types: Arid Lowveld and Mopani Veld, where it is restricted to river banks (Goodall and Erasmus, 1996). Part of the invasiveness of *C. odorata* is its aggressive pioneer habit that allows it to suppress the succession of desirable species in natural systems (Goodall and Erasmus, 1996). *C. odorata* reduces the species diversity in both forests and grasslands (Goodall and Erasmus, 1996), and is ranked as the alien invader posing the greatest threat to the floral diversity of the prominent Hluhluwe-Umfolozi Game Reserve Complex in KZN (Macdonald, 1983). Biodiversity is important for South Africa because of its importance in maintaining ecosystem function, its proven economic value, and its role in supporting subsistence lifestyles (Rutherford *et al.*, 1999). In general, *C. odorata* affects both the persistence of indigenous species and the appearance of the visual environment (Witkowski and Wilson, 2001).

The weed constitutes a fire hazard, even when green, and carries runaway fires into fire-sensitive forests, killing indigenous species (Goodall and Erasmus, 1996).

C. odorata reduces grazing and browsing for large herbivores in nature reserves, and is a problem to forestry during the establishment phase (Witkowski and Wilson, 2001).

Wilson and McFadyen (2000) point out that where *C. odorata* is used as a fallow in shifting agriculture, it causes the non-regeneration of secondary forests; in plantation agriculture it is easier to clear indigenous forests than *C. odorata*, resulting in very negative affects accruing on the biodiversity from agricultural practices that have included *C. odorata*. *C. odorata* is a nutrient-demandng early successional species, which uses high amounts of phosphorous and nitrogen in the soil (Witkowski and Wilson, 2001).

Its invasion of the river edges of the Greater St Lucia Wetland Park has reduced the number of viable nesting sites for the Nile crocodile (*Crocodylus niloticus*), and changed the conditions in the existing sites, so that its continued invasion poses a threat of extirpation to the Nile crocodile in the Lake St. Lucia ecosystem (Leslie and Spotila, 2001).

Not only is *C. odorata* a threat to conservation, it is also a potent weed where it invades potentially cultivatable land; because of its aggressive growth habits, the weed is difficult to suppress and increases the costs of maintaining infested farms (De Rouw, 1991; Timbilla and Braimah, 1996). It interferes with food production by reducing the carrying capacity of agricultural lands, and the allelopathic properties of the plant may adversely affect germination and growth of certain agricultural crops.

C. odorata is known to harbour and aid a number of insects and mites that are detrimental to commercial crops (Liggitt, 1983; Muniappan and Marutani, 1988), including the locust *Zonocerus variegatus* in Africa (Boppré, 1991; Timbilla and Braimah, 1996).

The leaves and young shoots of *C. odorata* have exceptionally high levels of nitrate (five to six times the toxic level), and are therefore poisonous to livestock, which are said to develop acute diarrhoea (Timbilla and Braimah, 1996), and lethal levels of anoxia after consumption of *C. odorata* (Sajise *et al.*, 1974, cited in Ambika and Jayachandra, 1990; Roder *et al.*, 1995; Goodall and Erasmus, 1996). Despite this, goats have been observed feeding on the weed where alternative food sources are scarce (Goodall and Erasmus, 1996).

It has been reported that hand weeding of *C. odorata* may cause skin allergies in people with allergic reactions, and that exposed stumps have caused "poisonous" wounds in farm workers that develop

into serious illnesses and sometimes cause paralysis (Ambika and Jayachandra, 1990; Timbilla and Brahima, 1996). Death from possible secondary infections has been reported in parts of Ghana (Timbilla and Braimah, 1996).

Benefits of *Chromolaena odorata*

Despite its weed status, *C. odorata* has been found to be of use in several countries.

In Ghana, it has several medicinal properties: the liquid extract is used primarily for treating fresh wounds, but it is also said to help blood clotting (Timbilla and Braimah, 1996). Old wounds and boils are also treated with the weed. Diseases like malaria and jaundice are also said to be cured by drinking the boiled extract of *C. odorata* (Timbilla and Braimah, 1996). Another important use of the weed is in the preservation (embalming) of bodies in the villages of Ghana (Timbilla and Braimah, 1996).

In the South-West Ivory Coast and Northern Laos as well as Ghana *C. odorata* is used as a fallow in semi-permanent crops. The non-indigenous plant is favoured because it colonises land quickly and is relatively easy to clear. It also has high biomass production, suppresses growth of other weeds and has a fast decomposition rate (Slaats, 1991; Roder *et al.*, 1995). *C. odorata* has also been alleged to improve the soil fertility of the soil where it is left as a fallow crop (De Foresta, 1991; Roder *et al.*, 1995; Slaats *et al.*, 1998); in general, fallow lands under *C. odorata* produce higher yields of crops (Timbilla and Braimah, 1996).

Another use for *C. odorata* is as a green manure crop for black pepper (Chandrashekar and Gajanana, 1996; Anwarulla and Chandrashekar, 1996), rice and cassava (Muniappan and Marutani, 1988). *C. odorata* can become a serious competitor for young rice plants when it regrows from rootstock, but the farmers find it relatively easy to remove the weeds by hand weeding (Roder *et al.*, 1995), and the weed has no negative impacts on rice crops when used as a fallow (Roder *et al.*, 1995).

Additionally, the weed is said to repel mosquitoes and snakes, and is used to preserve maize from rodents, and when its leaves are combined with other chemicals it is used as bait for trapping crabs (Timbilla and Braimah, 1996). *C. odorata* has also been used as a fish poison (H. Robinson, pers. comm. to C. Zachariades), and when used as a green manure in rice cultivations it is reported to kill fish (Muniappan and Marutani, 1988).

Its ability to suppress weeds is well known to Ghanaian farmers, and it suppresses grass growth (Timbilla and Braimah, 1996). Related to this last use, *C. odorata* has been reported to be of use in

suppressing grasses on several occasions. *C. odorata* was previously thought to be useful in controlling coarse grasses, mainly *Imperata* spp. (Binggeli, 1997), and it is possible that it was introduced into Asia and Ivory Coast for this reason (De Rouw, 1991). It was also previously thought that *C. odorata* was introduced into Ghana to suppress grasses, which are serious weeds in the area (Timbilla and Braimah, 1996).

A summary of the advantages and disadvantages of *C. odorata* is presented in Table 1.3.

Table 1.3 Summary of advantageous and disadvantageous effects of *Chromolaena odorata*

	Disadvantages	Advantages
Cultivation	Invades cultivable land; difficult and expensive to suppress; has allelopathic effects on crop plants	Fallow plant – quick coloniser, relatively easy to clear, fast decomposition, increased soil fertility, stabilises soils; possible suppresses weedy grasses; repels rodents from maize
Environment	Threatens ecosystems; suppresses desirable succession; fire hazard; shifting fallow use leads to deforestation; negative effects on crocodiles	
People	Skin allergies; toxic wounds which may lead to paralysis, secondary infections may lead to death	Medicines – wound-healing and clotting, malaria and jaundice, abortions; embalming agent; repels mosquitoes and snakes
Livestock	Toxic – acute diarrhoea, death by anoxia	
Fishing		Fish-poison; crab-bait

History of *Chromolaena odorata*'s spread

Since approximately 1870, *C. odorata* has increased its range from the Neotropics to the tropical and sub-tropical areas of the Paleotropics (Africa and Australasia) (McFadyen and Skarrat, 1996). Today there are few tropical or subtropical parts of Australasia and Africa that have not been invaded by *C. odorata*.

Australasia

The manner in which *C. odorata* first spread to the Old World is unclear. A previously accepted view was that it arrived in Singapore and Malaya (now peninsular Malaysia) in the 1920's via ballast in ships from the West Indies (Bennett and Rao, 1968, cited by McFadyen, 1988a). However, Hooker

stated in 1882 (cited in McFadyen, 1988a) that the plant was cultivated in India, and Prain (1903 and 1906, cited in McFadyen, 1988a) noted that it was occasionally cultivated in gardens in central and east Bengal and around Calcutta, and this could have been the start of the entire Asian invasion. It is thought likely that seed from the cultivated plants near Calcutta escaped and spread south into lower Burma and Malaysia, and north into Assam (Anon, 1967, cited in McFadyen, 1988a).

No matter how it was introduced, by 1912, the Director of the Java Botanical gardens had reported that a very vigorous weed of the genus *Eupatorium* was choking out other weeds at Deli in Sumatra (Johnstone and Tryon, 1914, cited in McFadyen, 1988a), and by 1940, *C. odorata* was recorded as a major weed in all the above-mentioned areas, as well as Sumatra (Biswas, 1934; Laan, 1940, both cited in McFadyen, 1988a).

The rapid spread of *C. odorata* from these initial areas of introduction to the entire south-east Asia was initially thought to have occurred primarily through long-distance wind-dispersal of the seeds, with a small amount of transport by humans and machinery (McFadyen, 1988a). It is now understood, however, that long-distance wind dispersal by *C. odorata* seeds is not as effective as previously thought (Blackmore, 1998).

Blackmore (1998) found that wind dispersal might only be possible over distances less than 80m. Wind dispersal over longer distances would appear to take place in a 'leap-frog' fashion, where each new generation's seed is deposited ca. 80 m from its parent plant (see De Rouw, 1991). The seed germinates, grows, and matures, to set seed that can travel a further 80m. Effectively, the short distance wind dispersal of *C. odorata* promotes the formation of dense, homogenous stands (Blackmore, 1998).

Blackmore (1998) has indicated that vehicles may be responsible for transporting a significantly higher number of seeds over greater distances than are carried by seed rain (see also Waterhouse and Zeimer, 2000). Thus the majority of the colonization of southeast Asia by *C. odorata* would most probably have been via man and machinery (however see De Rouw, 1991).

C. odorata was first discovered in Australia in 1994, in the Tully region on the East Coast of North Queensland. Major efforts have been taken to eradicate (not simply control) this weed. There are several hypotheses regarding the source of the Australian infestation: the seeds may have been introduced on bulldozers imported from infested areas in Texas (USA) in the 1960's; the seeds may have been introduced in contaminated clothes or baggage of personnel from neighbouring infested countries (the most probable theory); or seed may have contaminated imports of Stylo (*Stylosanthes*) seeds from Brazil in the late 1960s (Scott *et al.*, 1998).

Africa

Unlike Asia and Australia, where the whole invasion appears to have initiated from one site (Gautier, 1992), it seems that in Africa *C. odorata* was introduced on two separate occasions. The first African country to be infected was Nigeria, where *C. odorata* first appeared in the 1940's (McFadyen, 1988a; Gautier, 1992). This original introduction was probably via contaminated seed of *Gmelina arboreal*, a fast growing forestry tree from Sri Lanka (McFadyen, 1988a). Ten years later, after the Second World War, planters from Nigeria moved to Ivory Coast, taking the weed with them.

The second African locality which *C. odorata* infected was South Africa. *C. odorata* was first reported near Durban in 1947 (Hilliard, 1977, cited in Retief, 2001), from where it has spread rapidly. Binggeli (1997) reported that there is no time-lag (typical of some species when they enter a new environment) from the date of introduction to when *C. odorata* becomes invasive, and so it can be assumed that *C. odorata* arrived in South Africa soon before its first sighting in the late 1940's. The first plants may have been brought in to South Africa intentionally as ornamentals, or for horticultural reasons (Henderson and Anderson 1966; Gautier 1992; Binggeli 1997). Alternatively, the weed may have arrived accidentally in seed-contaminated packaging materials from the West Indies (Liggitt, 1983).

The country from which *C. odorata* came to South Africa, and the route by which it arrived is unknown (Vos, 1989; C. Erasmus, 1990, unpubl.). The initial spread of *C. odorata* in South Africa is thought to have been by the transport of its seeds along railways and roads from which it spread in 'leap-frog' pattern (Blackmore, 1998) to more remote areas. Figure 1.1 illustrates the recent pattern and rate of spread of *C. odorata*. The dispersal of *C. odorata* is thought to have been facilitated by the strong south westerly and north easterly winds, which are prevalent along the east coast of South Africa (Vos, 1989, pp 8/9).

C. odorata had spread from Durban Harbour to all the subtropical areas of the KwaZulu-Natal (KZN) province of South Africa by the 1980's (Goodall and Erasmus, 1996). It now extends south to Port St Johns in the Eastern Cape province of South Africa, and north to Mpumalanga and the Northern Province. It has crossed the South African borders, and is now found in Swaziland, Mozambique and possibly Zimbabwe (Vos, 1989; Gautier, 1992; Goodall and Erasmus, 1996; Neser, 1996).

C. odorata has the ability to invade all the semi-arid to mesic, frost-free areas of KwaZulu-Natal, the Eastern Cape, Swaziland, Mpumalanga, the Northern and North West Provinces of South Africa, Zimbabwe and Mozambique. The potential exists for *C. odorata* populations from West Africa and South Africa to merge (Sandberg, 2000).

Map of Spread and Distribution of *C. odorata* in South Africa

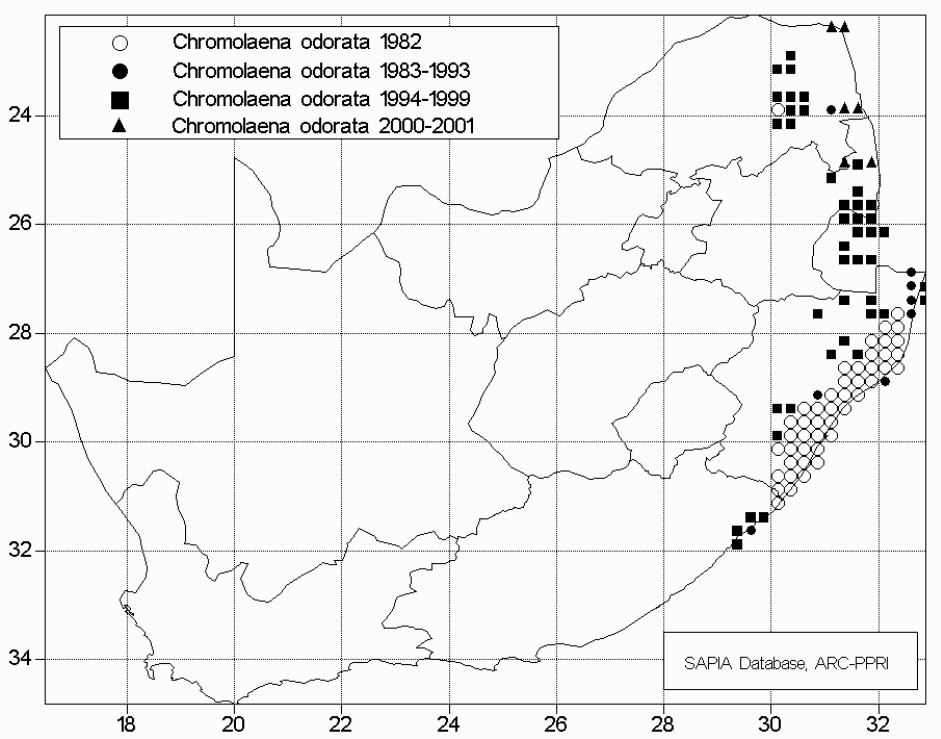

Figure 1.1 **Map showing spread and distribution of *Chromolaena odorata* in South Africa. (Drawn by L. Henderson, Plant Protection Research Institute, Pretoria.)**

Clues to the Origin of South Africa's *C. odorata*

It is thought that a population of *C. odorata* exists in its native range which is identical or very close in morphology and biochemistry to the South African *C. odorata*; i.e., that this is where the South African *C. odorata* was brought from (unless the South African *C. odorata* is a hybrid, or its original population is extinct in the neotropics). If the original population is not extinct, then it is believed that it harbours insects or pathogens that will become successful biocontrol agents on the South African *C. odorata*.

It has already been mentioned that several changes can take place within a population of an alien species before or while it becomes invasive. If only a small population was introduced into a new region, then the founder effect will influence the genetic structure of the population, but if a large population is introduced, or there are several introductions to the same place, then the genotype may remain unaffected, while the morphology (morphological plasticity) expressed is best suited to the immediate conditions (Mooney and Cleland, 2001).

Many techniques are available for use in investigating plant population structure. Some of them have been used in attempts to understand the differences between the South African and other populations of *C. odorata*. These techniques include morphology, cytology and cytogenetics, secondary metabolites, and isozymes and allozymes, and their use in investigating *C. odorata* are discussed below.

Morphology is a commonly used source of cladistic characters (Doyle, 1992), and has long been used to identify taxa from all ranks of living organisms from Kingdom level to varietal level (Morell *et al.*, 1995; Wiesing *et al.*, 1995b).

While morphological methods are very effective for many purposes, morphological comparisons may have limitations such as subjectivity in the analysis of the character, the influence of environmental or management practices on the character, limited diversity among cultivars with highly similar pedigrees (in cultivated plants), and confining of expression of some diagnostic character to a particular stage of development such as flowering or fruit ripening (Morell *et al.*, 1995).

Because morphological characters are so strongly influenced by the environment, special breeding programmes and experimental designs are needed to distinguish genotypic from phenotypic variation (Wiesing *et al.*, 1995b). Moreover, for small, non-flowering plant species (e.g. algae and mosses) as well as fungi, it is frequently difficult to find a sufficient number of morphological characters for a comprehensive systematic study (Wiesing *et al.*, 1995b).

The independence of morphological characters is another problem because of linkage, pleiotrophy, or the participation of numerous characters in adaptive syndromes such as self-pollination (Doyle, 1992). However, it is possible to estimate levels of variability for morphological characters, and their response to selection as well as their genetic background can be determined; genetic correlations and selection forces from the past can be inferred (Wiesing *et al.*, 1995b).

In general, morphological cladistics are often hampered by difficulties in defining characters; these problems are alleviated by the existence of more easily employed criteria in other techniques such as cytogenetic analyses; analysis of secondary metabolites, isozyme analysis and DNA profiling techniques (Doyle, 1992; Morell *et al.*, 1995).

A previous study (Vos, 1989) attempted to elucidate the origin of the South African *C. odorata* by looking at gross leaf morphology as well as leaf micromorphology (vestiture and venation) and floral features. The gross leaf morphology of the South African sample (from Durban, South Africa) was most similar to that of plants from Manaus (north Brazil), Thailand and Marica. Unfortunately, the precise locations of some of the samples that were examined by Vos (1989) were not made clear in

his thesis , so the exact locations cannot be given here, despite attempts to further define these localities.

The leaves from Thailand, Manaus, Durban (South Africa) and Medianeira (south Brazil) all had trichomes that were not found on leaves of plants from any other areas, and the trichomes on the Durban leaf were located in small pits, a feature not observed in any other leaves. These are thought to be a recent acquisition since the plant's establishment in South Africa (Vos, 1989). In floral morphology, the diameter of the capitulae of the plants from Goiania, Manaus and Durban (South Africa) were found to be most similar to each other than to other samples.

In a second, unpublished report (C. Erasmus, 1990), the gross morphology of South African samples was found to be most similar to that of plants from Guyana, Jamaica, Brazil and Colombia. (Plants from a wide range of North and South American as well as Caribbean countries were sampled.) Electrophoresis, chromosome numbers and trichomes were compared and the South African plants found to be most similar to those from Manaus (Brazil) and Thailand (in agreement with Vos's 1989 gross leaf morphology and trichome findings). The South African *C. odorata* has white flowers. White flowers are also recorded at eight locailities in Central America, five localities in northern South America and five localities in the Caribbean and Trinidad (C. Erasmus, 1990, unpubl.). Scanning electron microscopy of leaf glands and trichomes showed that the South African leaf material was most similar to those from Miami (USA), Medianeira (Brazil) and Itacuritiba (Brazil), and there was some similarity with material from Thailand and Florida. The report (C. Erasmus, 1990, unpubl.), based entirely on morphological data, found that no single country emerged as a convincing candidate for the origin of South African *C. odorata*.

H. Robinson (pers. comm. to C. Zachariades, 2001) has found that, based mainly on the appearance of involucral bracts, herbarium samples from the West Indies and (less frequently Central America) are most commonly similar to South African *C. odorata*. The morphology of herbarium specimens of Jamaican, Cuban and Guyanese origins was found to be similar to that of South African samples (A. Nicholas, unpubl.; J. Goodall, unpubl.), and similarities have also been observed between Puerto Rican and South African plants (M. Morris, pers. comm., to C. Zachariades), as well as between Bahaman and South African plants (I. Macdonald, pers. comm., to C. Zachariades.).

Cytogenetic analysis is another potentially useful method, however it has not been widely used and have the necessary resolving power for the interpretation of intraspecific relationships in only a few species, such as wheat, where the chromosomes have been thoroughly analysed (Morell *et al.*, 1995).

A study on the cytology of *C. odorata* has been attempted (Vos, 1989). However this study was inconclusive as the results indicated that the chromosome numbers of *C. odorata* is 2n = 60, 54 or

58. In a phylogenetic study of species and generic relationships of the Eupatorieae (Schmidt and Schilling, 2000), another species of *Chromoleana* (*C. sagittata*) fitted into a clade with a base chromosome number of 10. The multiple ploidy levels within *Chromolaena* (Sullivan, 1976) would seem to indicate that the chromosome number for *C. odorata* is most likely 60, however if the chromosome counts are correct, the aneuploid reduction (common in polyploids) has occurred, which could make intraspecific inferences from *C. odorata* chromosome counts more difficult to achieve.

The analysis of secondary metabolites, while potentially useful, is restricted to those plants that produce a suitable range of metabolites that can be rapidly analysed and are capable of distinguishing varieties (Morell *et al.*, 1995). This has not been done on *C. odorata*.

The methodology called 'allozyme' (shortened for alloenzyme) electrophoresis is used to examine variation at single discrete loci; allozymes are the different protein forms that are coded for by the various alleles at one locus (May, 1998). The term 'isozyme' or 'isoenzyme' is used to refer to a larger subset of protein forms, including the different protein products (intralocus and interlocus protein products) from all the genomic loci encoding a single type of enzyme [e.g., all the different forms of the enzyme malate dehydrogenase (MDH)] (May, 1998).

Traditionally, studies of population genetic structure have used allozyme markers (Peakall *et al.,* 1995). They are informative genetic markers with co-dominant inheritance that use straightforward laboratory procedures that are relatively rapid and inexpensive (Peakall *et al.,* 1995). Allozymes have been shown to provide important information at subspecific levels in plants (Peakall *et al.*, 1995), and in agricultural varieties (Morell *et al.*, 1995). Based on detailed analyses of ca. 450 plant species, it has been shown that an average of 50% of allozymes are polymorphic within plant populations, that widespread species exhibit greater genetic diversity than do narrowly distributed species, and that genetic diversity is greater at both the population and species level for outcrossers than for selfers (Peakall *et al.*, 1995).

There are some recognised limitations of allozymes: the detection of genetic variation is limited to protein coding loci, which may lead to underestimates of genetic diversity, and may not always be representative of the entire genome (Peakall *et al.*, 1995). A new allele will only be detected as a polymorphism if a nucleotide substitution has resulted in an amino acid substitution, which in its turn affects the electrophoretic mobility of the studied molecule. Because of the redundancy of the genetic code and the fact that not every amino acid replacement leads to a charge difference, only 30% of all nucleotide substitutions result in polymorphic fragment patterns in allozyme studies. Therefore, allozyme variation is often quite low (Reeve *et al.*, 1992), and its analysis underestimates the genetic variability of the genome. Using allozymes also restricts the study to those parts of the DNA that code

for stainable enzymes, and this is not necessarily a random sample of the genome. In general, the allelic proteins with similar electrophoretic mobilities may often be quite divergent at the DNA level (Doyle, 1995). Another drawback is that many plant species are polyploid, and the analysis of allozyme patterns of polyploids can be extremely difficult.

In a few cases it has been shown that allozymes differ in one or more physiological respect and, therefore, may not be evolutionarily neutral (Wiesing *et al*., 1995b). Allozymes are also tissue-specific in many cases, and protein expression can be environmentally responsive (Peakall *et al.*, 1995; Morell *et al*., 1995), similar to morphological characters; there are studies in both animals and plants suggesting the possibility of environmental selection for some allozyme loci, which can lead to biogeographic pattern discrepancies between allozymes and other genetic markers (Peakall *et al.*, 1995). Isozymes are also restricted in their utility by the number of enzyme systems that can be visualised and by the possibility that environmental conditions or management practices (Morell *et al*. 1995) can influence their expression.

A previous allozyme analysis on *C. odorata* (Vos, 1989) used peroxidases, and found that there was no variation recorded between plants within a population, and over a distributional range covering several populations within South Africa, yet some variation was detected between localities from around the world. Banding patterns from Thailand and a South African sample were found to be most similar, whilst the banding patterns exhibited by South American plants were more similar to each other than to the South African and Thailand plants. Unfortunately this work was never published in a peer reviewed journal, and therefore one cannot make inferences on the reliability of its results.

In addition to the allozyme, morphological, and cytological investigations into the origin of the South African *C. odorata*, a bioassay has been performed using pathogens (Morris *et al*., 1999). None of the investigations thus far have been able to elucidate the origin of the South African *C. odorata*.

Motivation for DNA-Based Study

Because of the failure of all studies thus far to find the origin of the South African *C. odorata*, it was felt that a DNA-based study on the genetic diversity of *C. odorata* was needed.

It is possible that the South African population of *C. odorata* is genetically identical (although the morphology might be altered) to a population in the native distribution range. Alternatively, the genetic structure may have been altered by founder effect, and in this case, the genotype of the South African population will be unique from other populations of *C. odorata*, but close similarity would still be detectable. A study on the genotype of *C. odorata* might also indicate whether or not

the morphological differences displayed between populations of *C. odorata* are genetically determined, or a phenotypic response.

Recent developments in molecular techniques have given biological control practitioners a selection of powerful tools with which to resolve the identity and relationships of both target weed and candidate agent (O'Hanlon *et al.*, 2000). Characterisation of genetic variability of an invasive weed species, as done by Amsellem *et al.* (2000a), could facilitate greater focus in a research project that is searching for biological agents specialised for a particular strain of host-plant (Hassan, 1972). The variety of PCR (polymerase chain reaction)-based molecular techniques available today almost ensures the accurate identification of potential biocontrol agents and their hosts, so as to avoid unwanted incompatibility or host-shift problems. Molecular techniques have been used to study the extent of differentiation among populations, ecotypes, forms and subspecies (Wolff and Morgan-Richards, 1998), as well as at the species level, for taxonomic, evolutionary and ecological studies (Richardson *et al.*, 1995).

AIMS OF THIS STUDY

It is believed that the morphological uniqueness of the South African *C. odorata* reflects physical and possibly chemical differences to most other populations, causing compatibility problems with potential biological control agents thus far assayed to date. It is possible that the genetic structure of this species will correlate with the morphological differences and geographical distribution.

Because previous direct morphological, cytological, isozyme and pathogen-bioassay studies have all thus far failed in determining the geographic origin of the South African *C. odorata*, it was decided that a genetic profiling study of the species would be undertaken. The aim of this study was therefore to use PCR-based molecular techniques to obtain genetic fingerprints (Chapter 2) and DNA sequences (Chapter 3) of as many *C. odorata* samples from as wide a range of localities as possible, particularly in the native region. It is thus hoped that an exact or near match (or matches, if there were multiple intrductions of *C. odorata*) to the South African 'genotype' will indicate the region of origin of the South African form(s), or that the pattern of phylogrography of *C. odorata* will be elucidated. With this information, it may be possible to find a genotype of *C. odorata* in its native habitat that is similar enough to the South African genotype(s) to harbour an insect that will establish in South Africa. It is believed that this will aid in the search for successful, compatible biocontrol agents for the South African *C. odorata*.

A secondary objective was to investigate the genetic variability within the species *C. odorata*, particularly within South Africa, in order to assess the number of invasion events of *C. odorata* into South Africa.

CHAPTER 2

DNA FINGERPRINTING OF CHROMOLAENA ODORATA

PCR-Based Molecular Markers

Analysis of DNA variation is rapidly becoming the method of choice in taxonomic studies because of its higher potential sensitivity than other morphological and biochemical methods, as well as the vast number of loci to be tapped, the high probability of variation being neutral, and its independence of environmental effects (Slade *et al.*, 1993; Richardson *et al.*, 1995; Morell *et al.*, 1995). DNA polymorphism is also probably a less biased estimator of genetic variation than gene-product-level variation such as allozymes (Stewart and Excoffier, 1996, cited in Martín *et al.*, 1999). A further advantage of molecular techniques is that the presence of the same DNA in every living cell of the plant allows tests on any tissue at any stage of growth (Morell *et al.*, 1995). These markers provide a data set that is independent of the morphological characters (Van Heusden and Bachmann, 1992a).

In the past decade, several key advances in molecular genetics have greatly increased the impact of population genetics on biology. Most important have been (1) the development of PCR; (2) the application of evolutionarily conserved sets of PCR primers; (3) the discovery of hypervariable microsatellite loci; and (4) the advent of routine and relatively simple DNA sequencing (Sunnucks, 2000). These innovations, coupled with the recent developments in analytical methods and relatively user-friendly software, have meant that much of the information available from molecular genetic data can now be tapped (Sunnucks, 2000).

Many techniques have been devised that make use of the above-mentioned advances in technology. The most common of these, as well as their potential application to the current study are briefly presented below.

Random Amplified Polymorphic DNA PCR

Random amplified polymorphic DNA (RAPD) PCR is by far the most commonly used variant of single-primer, random-amplification protocols. RAPD PCR employs one (random) 10 – 12 bp oligonucleotide (though longer primers are sometimes used) to amplify random segments of DNA. A 10-mer will

anneal to its random sequence approximately every million base pairs (Hoelzel and Green, 1998). Since a very large number of potential primers of ± 10 base pairs can be used, the amount of detail in a DNA fingerprint based on random amplified polymorphic DNA markers (RAPDs) is limited mainly by practical considerations (Van Heusden and Bachmann, 1992b).

RAPD PCR works because a small number of fragments (usually 5 – 10) will be amplified when the oligonucleotide anneals on each strand over a length range that can be readily amplified by PCR (usually less than 3 – 4 Kb).

Polymorphisms detected by the RAPD technology can theoretically result from several types of events: (1) the insertion of a large piece of DNA between the two annealing sites which may render the original fragment too large to be amplified, resulting in loss or a change in fragment size; (2) the deletion of a DNA fragment carrying one of the two primer annealing sites also results in the loss of a fragment; (3) a nucleotide substitution may affect the annealing of one of the two primers at a given site because of changes in homology, which can lead to a presence/absence of polymorphism; or (4) the insertion or deletion of a small piece of DNA (such as a tandem repeat number) can lead to a change in size of the amplified fragment (Wiesing *et al.*, 1995b; Hoelzel and Green, 1998). Under carefully controlled reaction conditions, these factors are likely to explain most of the observed variation in RAPD banding patterns (Hoelzel and Green, 1998). In practice, however, size changes are rarely observed (Wiesing *et al.*, 1995b). Instead a particular RAPD fragment is usually present (allele A) or absent (allele a) (Wiesing *et al.*, 1995b).

This allele distribution is typical for a dominant marker. A fragment is seen in the homozygous (AA) as well as in the heterozygous (Aa) situation, and only the absence of the fragment reveals the underlying genotype (aa) (Lynch and Milligan, 1994; Wiesing *et al.*, 1995b). The dominance of RAPD markers is one of the downfalls of the RAPD technique. Provided that there is only a single amplifiable allele per locus, this does not prevent the estimation of allele frequencies necessary for population genetic analysis, but it does reduce the accuracy of such estimation relative to analysis with codominant markers (Lynch and Milligan, 1994). Backeljau *et al.* (1995) found that RAPD data from diploid organisms cannot be directly subjected to parsimony methods, because of the non-codominant nature of the data, as well as other factors.

RAPD markers provide a powerful tool for the investigation of genetic variation in natural and domesticated populations (Huff *et al.*, 1993). Where a true phylogenetic relationship between units is expected, a cladistic treatment (Swofford, 1991, cited in Bachmann, 1997) is a more reliable indicator of the branching pattern than a distance-based approach. Since parsimony methods are designed as an objective approach to data sets containing both homoplasies and truly identical character states, they are ideal for RAPD bands (Bachmann, 1997). Treating RAPD bands as individual characters in a cladistic treatment of representative inbred lines should result in a strongly supported cladogram if

intraspecific variation is the result of evolution within genetically isolated lines. However, the cladogram should collapse into a general polymorphism if there is intraspecific gene flow (Bachmann, 1997). For this reason, the use of RAPDs in the intraspecific study of *C. odorata* would not be suitable.

Advantages of RAPDs

There are many advantages to the use of RAPDs for detecting genetic variation, amongst which are: relatively low cost; no radioactivity; low requirement of DNA sample quality and quantity; and the high frequency of polymorphic bands revealed (Heaton *et al.*, 1999). Another advantage is that it is not necessary to have containers of dry ice or liquid nitrogen in the field for sample tissue preservation. Instead, it is possible to desiccate samples using silica gel, which preserves DNA of high enough quality for RAPD analysis for several weeks (Heaton *et al.*, 1999).

RAPD analysis is simple and has been successfully used in several phylogeographic studies (Hagen *et al.*, 1999). RAPDs are of potential interest to evolutionary ecologists, because they may enable active pursuit of previously intractable questions including parentage analysis and investigations of clonal structure (Peakall *et al.*, 1995), as well as hybrid detection (Arnold, 1993; Crawford *et al.*, 1993; Marsolais *et al.*, 1993; Barker *et al.*, 1996; Padget *et al.*, 1998; Kuehn *et al.*, 1999).

The RAPD technique is relatively easy to apply to a wide array of plant and animal taxa, and the number of loci that can be examined is essentially unlimited, whereas the available battery of allozyme markers has already been effectively exhausted (Peakall *et al.*, 1995). RAPDs provide a powerful tool for studying population and regional differences; however, essentially the same results are available with the less numerous allozyme markers (Peakall *et al.*, 1995).

Disadvantages of RAPDs

The problem of reproducibility amongst laboratories is an issue of relevance to RAPD analyses. Running the same program on different thermal cyclers may result in different amplification patterns (MacPherson *et al.*, 1993, cited in Wiesing *et al.*, 1995b). This phenomenon is most likely caused by different temperature profiles in the reaction tubes (Wiesing *et al.*, 1995b). Some differences in DNA profiles have been found among laboratories, yet it was discovered that if the overall temperature profiles of the PCR reactions are identical among laboratories, then the RAPD fragments are reproducible for appropriately chosen primers. Some differences can still be expected between different thermal cyclers (Penner *et al.,* 1993, cited in Morell *et al.*, 1995).

The use of PCR to generate large amounts of desired product can potentially be both beneficial and detrimental to a study, particularly with multilocus markers. Failure to amplify at optimum conditions can lead to the generation of multiple undefined and unwanted products, even to the exclusion of the

desired product. At the other extreme, no product may be amplified (Roux, 1995). Optimisation of RAPD patterns is laborious since many reaction components as well as any part of the PCR program can be changed with quite unpredictable effects (Wiesing *et al.*, 1995b). Other factors that may also affect reproducibility are the quality and concentration of the DNA, the specific polymerase used, and the cycling conditions. With regard to the last variable, the annealing temperature is most important (Roux, 1995). The situation is further complicated by the fact that some of the variables in PCR are quite interdependent. For example, because dinucleotide triphosphates (dNTPs) directly chelate a proportional number of magnesium (Mg^{2+}) ions, an increase in the concentration of dNTPs decreases the concentration of free Mg^{2+} available to influence polymerase function (Roux, 1995). Many studies have demonstrated a marked influence of Mg^{2+} concentration on the obtained RAPD patterns (Wiesing *et al.*, 1995b). While strong and reproducible bands are obtained over a wide range of Mg^{2+} concentrations, a change in concentration often results in a qualitative change of fragment patterns (Williams *et al.*, 1993b, Wolff *et al.*, 1993, both cited in Wiesing *et al.*, 1995b). Wiesing *et al.* (1995b) suggest that Mg^{2+} concentrations are tested from about 1.5 up to 10 mM, and that two mM generally seems to be a good starting point.

Yield and specificity of PCR reactions can be increased by incorporating various additives such as dimentyl sulfoxide (DMSO, 1- 10%), Polyethylene glycol-PEG 6000 (5 – 15%), glycerol (5 – 20%), non-ionic detergents, formamide (1.25 – 10%), and bovine serum albumin (19 – 100µg/mL) into the reaction, and in fact, some reactions may amplify only in the presence of such additives (Roux, 1995).

Different polymerase enzymes often give rise to different RAPD products (Sobral *et al.*, 1993 and Wolff, 1991; both cited in Wiesing *et al.*, 1995b); switching to another type of enzyme is likely to render comparisons with previous experiments impossible (Wiesing *et al.*, 1995b).

Because prior DNA sequence information is not required for RAPD analyses, they enable the easy study of anonymous genomes. However, one of the inevitable trade-offs with the RAPD technique is that amplification is often performed under conditions of low stringency. Consequently, some of the products are produced from weak complexes between primer and template, and this can result in poor reproducibility for some primers and bands (Morell *et al.*, 1995).

Smith *et al.* (1994, cited in Rieseberg, 1996) point out that the use and analysis of multiple related PCR products could lead to overestimation of relatedness because the same character could be scored more than once. Genomic mapping in *Helianthus* revealed that approximately 13% of "homologous" RAPD bands actually represent paralogous loci and not orthologous genes (Rieseberg, 1996), and the use of paralogous loci could bias estimates of phylogenetic relationships (Doyle, 1992, cited in Wolfe and Liston 1998).

Several authors have demonstrated the occurrence of competition among RAPD fragments (Rieseberg, 1996, and literature therein), in that some polymorphic bands appear to represent better

matches to the primer than others resulting in reduced amplification of poorly matched fragments. In turn, this may lead to the mis-scoring of homologous fragments because the presence or absence of a particular band may be due to competition with other polymorphic fragments rather than a change in primer site. The problem of band competition appears to be most pronounced in situations where few fragments are amplified (Rieseberg, 1996).

Another problem reported for RAPD analyses is a low incidence of non-inherited bands, which are probably PCR artefacts (Morell *et al.*, 1995). The great majority of RAPD bands are inherited as Mendelian markers, so care is needed when drawing conclusions based on a small number of band differences.

As is the case of other DNA fingerprinting methods where multiple markers appear on the same gel, there can be uncertainty in assigning markers to specific loci in the absence of a preliminary pedigree analysis (Lynch and Milligan, 1994).

Rieseberg (1996) investigated an assumption often used in RAPD assessments; that fragment size is a dependable indicator of homology. It was found that of 220 pairwise comparisons, only 174 (79.1%) were appropriate for comparative genetic studies (i.e. orthologous rather than paralogous) (Rieseberg, 1996). However, it was concluded that the similarity of fragment size is a good predictor of homology, at least among closely related populations or species (Rieseberg, 1996). If fragment size is used as an indicator of homology, then gel resolution becomes critical for the correct classification of fragments (Rieseberg, 1996).

RAPD patterns in plants should not be compared unless there are some obvious shared bands that indicate the relatedness of the plants in the sample (Bachmann, 1997). It is possible that products of different loci will have similar molecular weights, and therefore be indistinguishable on a gel (because of comigration) (Lynch and Milligan, 1994).

As a consequence of dominance, RAPD and other multilocus profiles provide less genetic data than profiles for single locus, codominant markers such as simple sequence length polymorphisms (SSLPs) (Morell *et al.*, 1995), and reliable estimates of within-population heterozygosity are generally unattainable with RAPDs.

Studies in Plants

In a study on the implications of allozyme and RAPD variation in populations of *Buchloë dactyloides*, Peakall *et al.* (1995) found that the qualitative results were the same for both sets of markers. Allozymes exhibited a larger fraction of the total variation among individuals within populations and less among regional biotypes than the RAPDs, and the RAPDs revealed greater variation among regional biotypes than allozymes, but less variation among individuals within populations than the

allozymes. The among-populations component within a region was about the same for both markers (Peakall *et al.*, 1995).

The findings of Peakall *et al.* (1995) are consistent with qualitative reports that RAPDs detect more variation than allozymes in a wide variety of plant species (see Peakall *et al.*, 1995), though the essential 'genetic structure' and conclusions drawn from the two methods are the same.

Ayres and Ryan (1997) attempted to evaluate the genetic variability, assess clonal diversity, and delimit populations of *Wyethia reticulata*, (a slow-growing, long-lived herbaceous perennial that spreads by underground rhizomes - asexual reproduction) using allozyme and RAPD techniques. They found that the genetic similarity matrices derived from RAPD data and allozyme data were not correlated; RAPDs and allozymes each revealed a different pattern of variation when their similarity matrices were evaluated using multivariate analyses. Additionally, no consensus was found between RAPD-based and allozyme-based dendrograms when they were examined for identical clusters. In general, allozyme markers revealed more geographically concordant groupings than RAPD markers, but completely accurate geographical groupings resulted only when the allozyme and RAPD markers were combined.

In cultivar identification, RAPD analysis is more efficient than RFLP analysis (Fang and Roose, 1997), yet because RAPD analysis does not target those rapidly evolving sequences that may be most likely to differ between mutationally derived cultivars, the detection of polymorphic RAPD markers may require PCR amplifications with many different primers (Fang and Roose, 1997).

Amplified Fragment Length Polymorphisms

Amplified Fragment Length Polymorphisms (AFLPs) (Vos *et al.*, 1995) are PCR-based markers for the rapid screening of genetic diversity (Mueller and Wolfenbarger, 1999). They are essentially an intermediate technique between RFLPs and PCR (Karp *et al.*, 1996).

The AFLP technique is based on the selective amplification of sets of restriction fragments from genomic DNA (Hoelzel and Green, 1998). DNA is cut with restriction enzymes, and double-stranded adaptors are then ligated to the fragments. PCR primers designed to match the sequence of the adaptors are used to amplify the fragments, and the number of fragments amplified is intentionally limited by including nucleotides at the 3' end of the primer that extend into the unique sequence of the DNA fragment (by at least three base pairs on each primer for the complex genomes of plants and animals), so only a subset of the fragments will be amplified (those matching the designed 3' primer extension at the 5' end of the amplified sequence) (Hoelzel and Green, 1998). The number of amplified fragments can further be limited by using two restriction enzymes; one that cuts frequently

and one that cuts rarely (Hoelzel and Green, 1998). AFLP markers are highly variable and dominant (Mes, 1998), and have large genome coverage (Karp *et al.*, 1996).

Advantages

AFLPs allow rapid screening of genetic diversity (Mueller and Wolfenbarger, 1999). They generate hundreds of highly reproducible markers, and are easily applied to DNA of any organism (Mueller and Wolfenbarger, 1999). AFLPs produce a greater number of polymorphic markers than RAPDs for any single experiment (Gerber *et al.*, 2000), and they allow high-resolution genotyping of fingerprinting quality (Mueller and Wolfenbarger, 1999).

The time, cost and efficiency of AFLPs are superior or equal to those of other markers, such as allozymes, RAPDs, RFLPs and microsatellites (Mueller and Molfenbarger, 1999).

Disadvantages

AFLPs are dominant, rather than codominant markers (Mueller and Wolfenbarger, 1999). In many studies, particularly with polyploids, the potential of AFLPs in population genetic studies and analyses of breeding systems is reduced due to this dominant nature (Mes, 1998).

AFLPs are technically more demanding and require more DNA than RAPDs (Karp *et al.*, 1996). The effectiveness of AFLPs is reduced when DNA quality is poor, presumably because of interfering plant metabolites and/or DNA degradation (McLenachan *et al.*, 2000). It has been found that AFLPs are generally only successful with DNA extractions from fresh leaf material (pers. comm., C. van Heerden, Department of Genetics, University of Stellenbosch, South Africa, to N. Barker).

Studies in Plants

In a study of the intra- and inter-area genetic diversity of the weedy species *Rubus alceifolius*, Amsellem *et al.*, (2000b) used AFLPs, because this technique was capable of providing information on the history of introduction and relationships between native and introduced populations of *R. alceifolius*. The AFLP markers successfully differentiated between native and introduced populations of *R. alceifolius*.

O'Hanlon *et al.*, (1999) used AFLPs to investigate the relationships between invasive species of *Onopordum* in Australia, which were thought to have hybridised. Analysis of the AFLP markers revealed that Australia contained *O. acanthium*, *O. illyricum* and a full range of genetic intermediates between these species. They concluded that the range of genetic diversity expressed by the AFLP markers could best be explained by a combination of processes, including multiple introduction of

seed (including hybrid material), and continuous dispersal in Australia, leading to an increase in the contact among hybridising taxa.

Variable Number Tandem Repeats

Variable number tandem repeats (VNTRs) are the basis of several DNA fingerprinting and profiling techniques.

VNTRs refer to a DNA sequence consisting of a reiterated simple sequence motif (Warner, 1998), and include both micro- and minisatellites (Chambers and MacAvoy, 2000). Minisatellites are moderately repeated segments of 10 to 100nt forming more or less uniform tracts, while microsatellites are highly polymorphic regions of DNA containing simple repeats of 2-5 bp motifs of di-, tri-, and tetra-nucleotides (Morell *et al.*, 1995; Godwin *et al.*, 1997; Chambers and MacAvoy, 2000). Microsatellites are exceptionally useful for population genetics and they have been used in mapping studies (Schaal and Leverich, 2001). Microsatellites are highly discriminating, bi-parentially inherited co-dominant markers that are particularly suitable for discerning relationships between individuals (Chambers and MacAvoy, 2000).

Classical DNA fingerprinting (Jeffreys *et al.*, 1995a,b) uses laboratory procedures similar to RFLPs except that the probes are designed to hybridise with VNTRs, resulting in a complex profile or fingerprint (Morell *et al.*, 1995; Richardson *et al.*; Hoelzen and Green, 1998).

Microsatellites are highly abundant in eukaryotic genomes, but also occur in prokaryotes at lower frequencies (Godwin *et al.*, 1997; Schlőtterer, 1998). They are comparatively rare in organellar DNA (Wang *et al.*, 1994, cited in Weising and Gardner, 1999). VNTRs almost uniformly distributed over the entire genome, which is what makes them so useful for genome mapping projects (Godwin *et al.*, 1997, and literature therein; Schlőtterer, 1998, and literature therein). It would appear that most microsatellites are embedded in single-copy DNA, and thus facilitate unambiguous scoring of alleles (Schlőtterer, 1998). Database surveys have shown that microsatellites are comparatively rare in organellar DNA

Various functional roles have been attributed to VNTRs; for example, hot spots of recombination (Kobori *et al.*, 1986; Bullock *et al.*, 1986, both cited in Gupta *et al.*, 1994); the recognition and expression of genes (Haamada *et al*, 1984b; Shafer *et al.*, 1986; Murphy *et al.*, 1989, all cited in Gupta *et al.*, 1994); sex determination (Singh *et al.*, 1980, cited in Gupta *et al.*, 1994; Huijser *et al.*, 1987, cited in Wiesing *et al.*, 1995b); and the control of biological rhythms (Wharton *et al.*, 1985, cited in Wiesing *et al.*, 1995b).

VNTRs are often characterised by high meiotic mutation rates that mainly concern the number of repeats; individual microsatellites often show a 'variable number of tandem repeats' with alleles showing considerable variation in the number of iterations when more than about 20 nucleotides are present in a microsatellite (Wiesing *et al.*, 1995b; Warner, 1998). Microsatellites seldom include more than about 70 repeat units (Schlőtterer, 1998). The simple-sequence repeats are known to have a high rate of gaining and losing repeat units due to DNA slippage (Schlötterer, 1998, cited in Camacho and Liston, 2001). The most common changes are those of a single repeat unit, which allow microsatellite mutations to be interpreted as a very good approximation of a stepwise mutation process (Schlőtterer, 1998). Interestingly, mutation rates appear to be positively correlated with the total size of the VNTR array (Caskey *et al.*, 1992, Wayne *et al.*, 1991, both cited in Wiesing *et al.*, 1995b). The repeats might remain invariant for long periods of time if they contain a few repeat motifs only. However, as soon as the tandem copy number exceeds a certain threshold, the chance for further mutation is greatly enhanced (Wiesing *et al.*, 1995b). In concordance with these observations, high molecular weight bands within a multilocus VNTR fingerprint are often more variable than bands occurring in a low molecular weight range (Wiesing *et al.*, 1995b).

Internal heterogeneity of the repeated units (Jeffreys *et al.*, 1990, cited in Wiesing *et al.*, 1995b) as well as somatic mutations are also observed (Armour, *et al.*, 1989, Kelly *et al.*, 1989, both cited by Wiesing *et al.*, 1995b). Together with the accumulation of point mutations within repeat units, the intermingling of different types of repeats may result in DNA sequences which are 'cryptically simple' (Tauz *et al.*, 1986, cited in Wiesing *et al*, 1995b), i.e., different VNTR sequences occur intermingled with each other in a particular stretch of DNA (Weising *et al.*, 1995b).

VNTRs in Plants

Wiesing and Gardner (1999) attempted to use a set of primers to study mononucleotide repeat variation in chloroplast DNA of angiosperms. The conserved nature of the chloroplast genome (Wolfe *et al.*, 1987, cited in Weising and Gardner, 1999) makes it unsurprising that they failed to detect any intraspecific polymorphisms.

Compared with animals, VNTRs are poorly known in the nuclear (and chloroplast) genomes of plants and extensive studies of variability within species are lacking. Although little is known about plant microsatellites, searches of GenBank by several groups have shown that microsatellites are abundant in plants, although their distribution and type differ from those in animals, with the most common motif being $(AT)_n$ in plants (Morell *et al.*, 1995, and literature therein). Several recent studies, however, show that VNTRs can be used to fingerprint both plant cultivars (Nybom *et al.*, 1990; Beyermann *et al.*, 1992, both cited in Peakall *et al.*, 1995) and genets within natural populations (Neuhaus *et al.*, 1993; Antonius and Nybom, 1994, both cited in Peakall *et al.*, 1995).

Inter Simple Sequence Repeats

Inter simple-sequence repeat (ISSR) amplification was first described by Ziętkiewicz *et al.* (1994). The approach relies on the existence of inversely repeated microsatellite regions. When the inverted microsatellite regions or VNTRs are located close to each other (i.e., within about 5kb), the inter-repeat region may be amplified by PCR with a single primer based on the microsatellite sequence (Fang and Roose, 1997). See Figure 2.1.

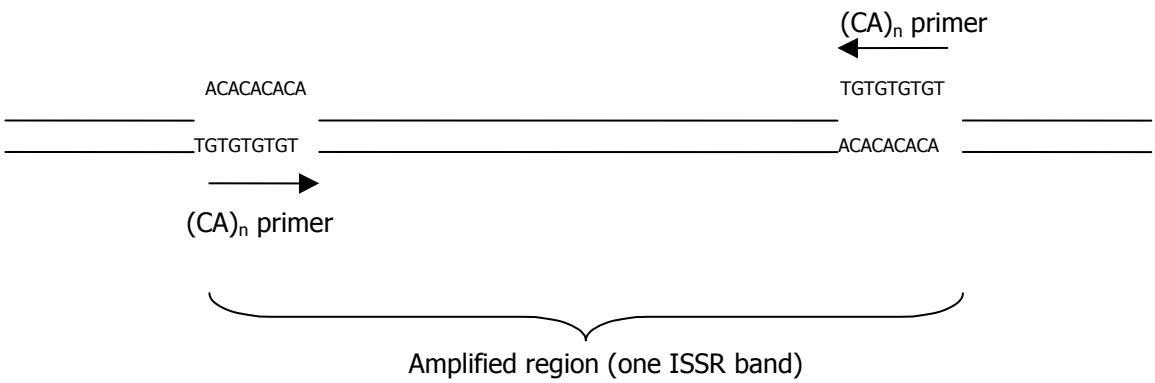

Figure 2.1 ISSR PCR. A single primer targeting a $(CA)_n$ repeat is used to amplify genomic sequence flanked by two inversely oriented $(CA)_n$ elements (from Ziętkiewicz *et al.*, 1994).

Since the copy number of the repeated motif in the target DNA is usually higher than that of the primer, pairing may occur at different registers. This is true for both ends of a given amplicon. Therefore, the initial amplification products of a particular inter-repeat region are not necessarily fragments of distinct length, but rather comprise a population of fragments of slightly different size. However, in the course of 40 PCR cycles, internal priming results in the successive shortening of a given fragment, with the shortest possible product (primed by the innermost repeat) predominating in the end. Cloning and sequencing of PCR products obtained by simple sequence- or minisatellite-primed PCR verified that each end is 100% complementary to the primer sequence and does not contain additional repeats (Fang and Roose, 1997). In a variation of the technique, single fragments of defined length may be obtained when "anchored" primers are used, i.e., primers carrying one or more unique or degenerate nucleotides in addition to the repeat (Wiesing *et al.*, 1995b; Fang and Roose, 1997).

Bands amplified by simple sequence primers are basically generated by two mechanisms: (1) RAPD-like bands that result from binding of the primer to cryptically simple sequences in the template, allowing mismatch pairing during the first cycles. Depending on the extent of mismatch, these bands do not occur at elevated annealing temperatures. (2) True inter-repeat bands that result from 100% match priming to template regions containing inversely repeated perfect simple sequences. These bands are still generated using annealing temperatures that considerably exceed those calculated by

the Wallace rule (Wiesing *et al.,* 1995). These higher annealing temperatures are the reason that the profiles produced by the ISSR technique are more reliable and repeatable than methods like RAPDs (Wolfe and Liston, 1998, cited in Camacho and Liston, 2001).

ISSR markers are inherited as dominant (or sometimes codominant) genetic markers in a usually Mendelian fashion (Gupta *et al*., 1994; Tsumura *et al*., 1996a, both cited in Wolfe and Liston 1998). Tsumura *et al.* (1996, cited in Camacho and Liston, 2001) found that most of their ISSR bands (96%) segregated according to Mendelian expectations. The fact that not all bands segregate according to Mendelian fashion suggests either a high mutation rate from generation to generation (Camacho and Liston, 2001) or that some of the bands are organellar in origin (Wolfe *et al*., 1998).

The bands are interpreted as diallelic with alleles designated as "band present" or "band absent". Presumably, the absence of a band means that divergence has occurred at one or both of the primer sites. Other possibilities include the loss of an SSR site (one of the ISSR primer annealing sites) or a chromosomal structural rearrangement (Wolfe and Liston, 1998). Variation could be due to mutations in length of the simple-sequence repeat, to mutations that alter anchored nucleotides, or to insertion or deletion mutations in the sequence between the primer sites (Fang and Roose, 1997). Chromosomal structural rearrangements have also been suggested as a source of ISSR variation (Camacho and Liston, 2001), as well as the mechanisms suggested for VNTR variation.

Advantages

ISSR PCR generates reliable and reproducible polymorphic patterns (Gupta *et al*., 1994; Wolff and Morgan-Richards, 1998) that are effective multilocus markers for applications such as diversity analysis, fingerprinting and genome mapping (Godwin *et al*., 1997). It combines the advantages of DNA fingerprinting with highly polymorphic probes on one hand with those of the PCR technique on the other (Wiesing *et al*., 1995b). Fang and Roose (1997) found that ISSR markers are highly reproducible.

By virtue of greater numbers of bands amplified and detected per primer (over 100 in some cases, Godwin *et al*, 1997), ISSR analysis is quicker to apply than other methodologies with the possible exception of AFLP analysis (Godwin *et al*, 1997).

ISSRs are more reliable and robust than RAPD markers, due mainly to the method of detection, and possibly to the fact that the primers are longer, allowing for more stringent annealing temperatures (Wolfe and Liston, 1998, cited in Camacho and Liston, 2001). These higher annealing temperatures apparently provide a higher reproducibility of bands than RAPDs (Nagaoka and Ogihara, 1997; Wolfe *et al*., 1998; both cited in Chamcho and Liston, 2001). Yang *et al.* (1996, cited in Godwin *et al*., 1997)

found that 10% of RAPD bands in pairwise sorghum genotype comparisons were unreliable, whereas the error rate for ISSR bands was less than half of that (Godwin *et al.*, 1997).

Coupled with the separation of amplification products on a polyacrylamide gel, ISSR amplification can reveal a much larger number of fragments per primer than RAPD analysis (Gupta *et al.*, 1994; Wolf *et al.*, 1995), because in contrast to arbitrary primers, there is a high degree of homology between simple-sequence primers and their binding sites (Wiesing *et al.*, 1995b).

Disadvantages

ISSR markers have not proven to be as polymorphic or efficient as AFLP analysis in wheat, with its large, complex genome (Nadella *et al.* 1996, cited in Godwin *et al*, 1997). Nevertheless, in most plant species tested, ISSR markers can be suitably applied to most situations, and have been shown to be particularly useful in genetic fingerprinting and diversity analysis (Godwin *et al.*, 1997).

Because microsatellites are found in all eukaryotes (and even some prokaryotes), microsatellite-based primers can amplify DNA from virtually any eukaryotic organism, and therefore contamination of a leaf sample with insect or perhaps fungal DNA can cause false results (Fang and Roose, 1997). In citrus, careful washing of leaf samples minimises this problem (Fang and Roose, 1997).

Studies in Plants

Because of the great many advantages over RAPDs and AFLPs, there was support for the use of ISSRs as the fingerprinting technique in the study of genetic variation of *Chromolaena odorata*. A brief section is presented here with a selection from the available literature on the use of ISSRs, with the findings and applicability of these findings to the current study on *Chromolaena odorata*.

Just as with RAPDs, PCR with minisatellite and simple sequence-based primers provides a convenient method of genetic identification and differentiation in plants and fungi at various levels ranging from species to individuals (Wiesing *et al.*, 1995a). The ISSR technique has been shown to work in a range of plant and animal species (Morell *et al.*, 1995), and can rapidly differentiate between closely related individuals (Zietkiewicz *et al.*, 1994). Though quite new, the ISSR technique has yielded distinctive, variable, multi-fragment profiles from all fungal and plant species tested so far (Wiesing *et al.*, 1995a and literature therein). Doyle *et al.* (1998) demonstrated that chloroplast microsatellites in species of the B-genome of *Glycine* subgenus *Glycine* show considerable intraspecific and even intra-haplotype polymorphisms.

Blair *et al.* (1999) found that most of the primers that they used gave the same information and therefore could be interchanged without any loss of power to distinguish varietal diversity. This would indicate that any primer that produces sufficient numbers of polymorphic bands in the study of the

genetic diversity of *C. odorata* could be used with confidence, i.e., that information that may be presented by other primers is not being lost.

Unfortunately, no data are yet available on mutation rates of tandem repeats in plants or fungi (Wiesing *et al.*, 1995b). If polymorphism and the number of alleles per locus in plants and fungi are different from those in animal species, precautions should be taken when applying equations originally developed for human and animal fingerprints to plants and fungi (Wiesing *et al.*, 1995b).

Godwin *et al* (1997) claim to have confirmed previous reports that ISSR markers have useful applications to plant genetics and breeding. ISSR markers have previously compared favourably with RFLP and RAPD markers with applications to DNA fingerprinting and diversity analysis in sorghum (Yang *et al.*, 1886, cited in Godwin *et al.*, 1997), finger millet (Salimath *et al.*, 1995, cited in Godwin *et al.*, 1997) and maize (Kantety *et al.*, 1995, cited in Godwin *et al.*, 1997).

Although ISSR loci can potentially distinguish many individuals, they unfortunately do not measure true heterozygosity due to their generally dominant inheritance (Wolfe and Liston, 1998, cited in Camacho and Liston, 2001). For this reason, the level of inbreeding cannot be determined with ISSR data alone (Camacho and Liston, 2001), though Wolfe *et al.* (1998) used ISSR data to successfully establish the hybrid status of natural *Penstemon* populations.

Camacho and Liston (2001) found (using ISSR markers) that most of the genetic diversity of the fern *Botrychium pumicola* was within populations, and that there is little among-population differentiation, yet they found the ISSR markers useful for examining the significance of asexual reproduction and the population structure of the fern. They suggested that ISSR loci could potentially distinguish individuals, even though they don't measure true heterozygosity due to their generally dominant nature (Camacho and Liston, 2001). Their ability to distinguish between asexual and sexual populations indicates that it could be possible to determine to a degree whether *C. odorata* is reproducing apomictically within a population.

In contrast to the recent findings of Camacho and Liston (2001), Wiesing *et al.* (1995) found that there were a considerable number of intraspecific and interspecific ISSR polymorphisms in the *Actinidia* spp that they were studying. The similarity of banding patterns in the *Actinidia* spp. was higher within the species than between them. The results of Camacho and Liston (2001) indicate that there is a possibility that ISSRs may not be sufficiently variable within certain species for a population genetics study, as they found more interspecific than intraspecific variation.

In yet another possible drawback of ISSRs, Wolfe *et al.* (1998) warn that the hypervariability of ISSR markers might make them too variable (as opposed to not variable enough) and therefore unusable for studies of natural populations or for species-level studies. The literature mentioned above appears to refute this suggestion, yet the possibility remains that in some families or genera the ISSR loci may

be too variable to yield informative loci at certain levels. As an example, Van Houten *et al.,* (1991) found that ISSR fingerprint loci showed too much variation for a reconstruction of genetic differentiation among populations of *Microseris pygmaea* (Asteraceae, Latuaceae) (cited in van Heusden and Bachmann, 1992a).

In general, more ISSR studies have been done on cultivated plants than natural populations (Wolfe and Liston, 1998). For example, Blair *et al*. (1999) found ISSR markers useful in determination of closely related cultivars within a single species (*Oryza sativa*), and Wolfe *et al*. (1998) found them to be useful population- and species-specific markers for *Penstemon* species. The relative lack of studies of natural populations with ISSRs indicates that the protocols for the analysis of the natural *C. odorata* populations may not be readily available. The use of ISSR loci in phylogeographic analyses (particularly on a panmictic scale) is not common. Thus the study of ISSR loci in *C. odorata* with the aim of investigating phylogeography of this species would be unusual.

In an example of a study on a naturally occurring plant species, Wolff and Morgan-Richards, (1998) found that within the species *Plantago major*, substantial polymorphism was revealed using three RAPD and three ISSR primers. Most of the variation was between subspecies and between countries of origin, as can be expected from a highly inbred plant species (Wolff and Morgan-Richards, 1998). The two groups were clearly clustered in different branches of a neighbour joining tree, and thus it appeared that the variation within each group was smaller than the variation between the two groups (Wolff and Morgan-Richards, 1998).

The fact that most polymorphic RAPD and ISSR fragments are almost subspecies-specific did not coincide with the results found using allozyme electrophoresis on the two subspecies of *P. major* (Wolff and Morgan-Richards, 1998). The allozymes showed no subspecies-specific alleles and only two allele-frequency differences between Dutch populations of the two subspecies. The RAPD and ISSR results are also not totally in concordance with the DNA fingerprinting results where it was shown that the two subspecies had very dissimilar DNA fingerprint patterns (Wolff and Morgan-Richards, 1998). The RAPD and ISSR results take an intermediate position between the allozyme and fingerprint results. It is not known to what extent the RAPD and ISSR fragments generated consist of repetitive DNA (Wolff and Morgan-Richards, 1998), but if a substantial number of them amplify repetitive DNA this may explain the discordance with the allozymes (encoded by functional non-repetitive DNA) and the fingerprint data (solely based on tandemly repeated motifs).

The differentiation of two subspecies for many RAPD and ISSR bands, but not for allozymes and DNA fingerprinting can be explained in three ways (Wolff and Morgan-Richards, 1998): Firstly, the two subspecies may have diverged relatively recently and differentiation for the different markers has occurred at different rates, leaving some markers that are still shared between the two subspecies. Secondly, functional DNA, like allozyme loci and undoubtedly some RAPD and inter-SSR amplicons, do

not allow high mutation rates as there is strong selection against non-synonymous substitutions due to the lowered functionality of the resulting mutant alleles. On the other hand it is known that repetitive DNA can diverge relatively fast between species, subspecies and even populations (Wolff and Morgan-Richards, 1998). Thirdly, in nature, some hybrids are formed, and in these hybrids recombination takes place between some parts of the genomes of the two subspecies (Wolff and Morgan-Richards, 1998). The latter two explanations are favoured by Wolff and Morgan-Richards (1998).

Materials and Methods

Unfortunately, despite the efforts outlined below, ISSR amplification was only sporadically successful in this study. There were no immediately obvious reasons for the frequent failure of ISSR amplification and a long series of trials were carried out in an effort to isolate and solve the source(s) of the troubles.

What follows is a description of all the trials that were performed in an effort to obtain useable and repeatable ISSR PCR product.

Four main components of the ISSR process were investigated; extraction protocols, PCR reagents, thermal cycling profiles, and gel visualisation. The variations of these components that were tried are presented below. In the section following this (Results and Discussion), the relative successes and failures of each component are presented and possible explanations for the observed results offered.

Gel Visualisation

Detection methods of ISSR markers on banana cultivars using agarose or polyacrylamide gels and ethidium bromide, silver staining or radioactive labelling were compared by Godwin *et al.* (1997). The least sensitivity was recorded on agarose with ethidium bromide, intermediate on a polyacrylamide mini-gel with silver staining and most sensitive on polyacrylamide sequencing gel with radiolabelling. Hence it is likely that useful polymorphisms will be lost if ISSR products are resolved on agarose gels with ethidium bromide (Godwin *et al.*, 1997; Wolfe *et al.*, 1998).

ISSR amplification products in this dissertation were initially run on a 1% agarose gel that contained ethidium bromide. 20µL of PCR product was mixed with a drop (ca. 4µL) of the loading buffer (Bromphenol Blue).

Later, a silver staining of a polyacrylamide system was used, to see if it would show bands that were not visible in agarose and ethidium bromide visualisation. The polyacrylamide gel electrophoresis (PAGE) system used was Mightly Small Miniature Slab Gel Electrophoresis unit (Hoefer Scientific Instruments). A 5% stacking and 10% resolving polyacrylamide gel was used. Ten microlitres of each sample was mixed with a small drop (ca. 2µL) of 40% glycerol loading buffer. Samples were then electrophoresed at 150V using a PS 500X DC Power Supply (Hoefer Scientific Instruments). After 4 hours the loading buffer had run off the end of the gel, and the gel was then stained according to the protocol in Appendix 2.

Radiolabelling facilities were not available to attempt this most sensitive gel visualisation technique.

Comparisons of DNA Extraction Protocols

In total, four DNA extraction protocols were tested; three of these protocols are based on the CTAB (hexadecyltrimethyl-ammonium bromide) DNA extraction protocol of Doyle and Doyle (1987), and the fourth is a commercially available extraction kit. The extraction protocols are as follows:

Procedure 1: 'Basic' CTAB extraction with ethanol rinse (Doyle and Doyle, 1987, with modification by Scott *et al*., 1998) (Appendix 3)

Procedure 2: CTAB extraction after grinding in ethanol (Adams *et al*., 1999)

Procedure 3: CTAB extraction after grinding in ethanol, with polyvinylpyrrolidone (PVP or Povidine, Sigma) added to the CTAB extraction buffer (Kim *et al*., 1997)

Procedure 4: Qiagen DNeasy Plant Mini Kit

CTAB-Based Extraction Protocols

DNA was initially extracted from leaf samples with a modified hot CTAB method (Doyle and Doyle, 1987). Details of this initial extraction protocol are listed in Appendix 3. A new addition to this extraction protocol is an initial rinse of the leaf tissue in absolute ethanol, to reduce surface contamination and assist in rehydration according to Scott *et al*. (1998), who sequenced the first internal transcribed spacer (ITS-1) region of *C. odorata* plants. It was decided that, because Scott *et al*. (1998) were successful in DNA extraction and amplification from *C. odorata* tissue, their modification would be employed from the start (Procedure 1).

Adams *et al*. (1999) suggested a further modification of the CTAB extraction protocol that proposes to irreversibly inactivate DNases in the DNA extractions. During DNA extractions the EDTA (ethylenediamine tetra-acetic acid) present in CTAB buffer and many other extraction buffers (Adams *et al*, 1999) is thought to chelate Mg^{2+} ions, which are often needed by DNases (Ogawa and Kuroiwa, 1985, cited in Adams *et al*., 1999). Thus it is assumed that by extraction with EDTA will render any present DNases inactive after DNA extraction. However, the assumption that all DNases can be inhibited by EDTA is incorrect (Adams *et al*., 1999). For example, in *Chlamydomonas reinhardtii* six DNases were found and each of these required Ca^{2+} for activation. The DNases were little affected by the amount of Mg^{2+} ions (Ogawa and Kuroiwa, 1985, cited in Adams *et al*., 1999). In tobacco, the DNases did not appear to need any specific ions for activity and were inhibited by Mg^{2+} (Zilberstein *et al*., 1987, cited in Adams *et al*., 1999). Two DNases were found in wheat seedlings: one required Mg^{2+} and the other was activated by EDTA (Jones and Boffey, 1984, cited in Adams *et al*., 1999). Thus it is possible that any tissue extracted by means of the CTAB extraction method, or any other

method employing EDTA is not necessarily free of DNases, which could present troubles in PCR amplification.

Preservation in silica gel does not irreversibly inactivate DNases, and because different plant species apparently produce different kinds of DNases, it seems that a more general method for DNase inactivation is needed (Adams *et al.*, 1999). Through trial and error, Adams *et al.* (1999) found that the grinding of plant material in a small quantity of ethanol, before grinding in the extraction buffer, would seem to be a general method for the inactivation of DNases, regardless of their requirements for Mg^{2+}, other ions, or no ions.

Based on these findings, samples of *C. odorata* were ground in a small (<500µL) amount of 100% ethanol, which was allowed to evaporate to a smaller volume, before further grinding in CTAB extraction buffer, and normal extraction was followed thereafter (Procedure 2).

In a further modification of the CTAB extraction protocol, 1% PVP was added to the CTAB extraction buffer (Procedure 3). Inclusion of PVP in the extraction buffer may reduce the amount of polyphenolics in the DNA extraction (Kim *et al.*, 1997).

Commercial DNA Extraction Kit

The final DNA extraction protocol (Procedure 4) follows Camacho and Liston (2001), who successfully amplified ISSR bands from fern samples, from which the DNA had been extracted using the Qiagen DNeasy Plant Mini Kit. DNeasy-purified DNA is reported to have A_{260}/A_{280} ratios of 1.7 to 1.9 and absorbance scans are reported to show a symmetric peak at 260nm, confirming high DNA purity. High concentrations of RNA are known to suppress PCR amplifications (Pikaart and Villeponteau, 1993, cited in Nickrent, 1994), and DNA purified using the DNeasy Plant procedure is reported to be free of RNA contamination since an RNase digestion step is included in the procedure. This kit was obtained for this dissertation, and DNA extracted from leaf tissue as per manufacturer's instructions.

The only difficulty encountered in using this kit was in preventing samples that had been ground in liquid nitrogen from thawing while other samples were being ground in nitrogen. This was partially overcome by dipping an Eppendorf tube into the liquid nitrogen just before the ground sample was decanted into the tube. The tube (containing frozen, ground leaf tissue) was then immediately put into a deep-freezer to prevent the eppendorf tube and sample from warming up. Despite these measures it was still possible that samples thawed before the next stage, which was cautioned against by the Qiagen DNeasy Plant Mini Kit manufacturers.

Quantification of DNA Extract

An attempt was made to quantify and qualify the DNA of the initial samples extracted with this 'basic' CTAB extraction protocol, using the UV $A^{280/}_{260}$ absorption ratio. Measurement of absorbance at 260nm is a well-established method for the estimation of concentration of nucleic acids and oligonucleotides in solution. Measurement at more than one wavelength, e.g., 260 and 280nm, from which a ratio can be determined, is a check of the validity of the A_{260} reading. It is also a means of establishing whether the nucleic acid is contaminated with protein of other materials (Manchester, 1996).

A Philips PU 8679 Vis/NIR spectrophotometer from the Plant Physiology Laboratory of the Botany Department at Rhodes University was used for these measurements. The samples were diluted 1:20 with distilled water, as the 1mL cuvettes were too large to take readings of 200µL of undiluted extract.

Initial PCR Thermal Cycling Conditions

A prior protocol for ISSR amplification had not been determined for the Molecular Systematics Laboratory at the Botany Department of Rhodes University, so the initial reagent concentrations and volumes were taken from Professor Andi Wolfe's web page (www.biosci.ohio-state.edu/`awolfe/ISSR/protocols.ISSR.htm) (See Point 1, Appendix 4). Volumes and concentrations were later adjusted slightly for easier pipetting (Point 3, Appendix 4). When it became evident that the adjusted ISSR protocol was not ideal, the original volumes and concentrations (as according to Professor Andi Wolfe) were used (point 14, Appendix 4), but the results were no different to those with the protocol with easy-to-pipette volumes.

PCR Reagents used were: dNTPs (Promega deoxynucleotide triphosphates), *Taq* polymerase (Promega *Taq* DNA Polymerase in Storage Buffer A), magnesium (Promega magnesium chloride solution, 25mM), Dilution Buffer (Promega Thermophilic DNA Polymerase 10X Buffer, Magnesium Free), water (Sigma Double-Processed Tissue Culture water), and primers. A set of 100 primers was purchased from the University of British Columbia (UBC). The UBC Primer Set #9 was obtained with the kind help of Dr. John Hobbs (NAPS Unit, Biotechnology Laboratory, UBC, Vancouver, Canada). The dried primer pellets were all resuspended in 100µL dH_20 to give a 15 µM primer solution (Dr. John Hobbs, pers. comm.).

PCR Reagents

The reagents within the PCR mix can all be varied, with many different consequences. Wiesing *et al.* (1995) showed that ISSR banding patterns were affected by primer, template and Mg^{2+}

concentrations as well as annealing temperatures. In the hope that changes in one or more of the reagent concentrations would improve the quality and frequency of ISSR success, a series of trials were performed to test this possibility.

Primers

Primer Screening

Primers were initially screened according to their reported abundance in plant genomes (Zietkiewicz *et al.*, 1994; Blair *et al.*, 1999; Wolfe and Liston, 1998). This was done with the PCR reagents prepared with Andi Wolfe's reagent volumes (i.e., difficult to pipette), and then later these volumes were adjusted to be easier to pipette. The primers screened were primers 841, 821, 856, 962, 867, 888, 891, and 841 of the UBC ISSR primer set #9.

Primer Volumes

The primer concentration was tested (using primer 841). It has already been mentioned that the first ISSR reactions were made up according to Wolfe's recipe and later adjusted. These recipes used 0.33µL (0.8pM) and 1.0µL (2.7pM) of the 15µM primer solution respectively, in a 25µL total reaction volume, and were tested against each other in a trial set up specifically to reduce all other variation (Point 14, Appendix 4). Primer volumes of 1, 2, 4, and 8µL (1.3, 2.7, 5.3, and 10.7pM respectively) per 50µL reaction volume were tested.

Primer Degradation

The possibility that the primers had degraded was considered, and therefore new primers were made up (according to the sequences provided by UBC) by the University of Cape Town (Oligonucleotide Synthesis Facility, Department of Biochemistry, UCT, Rondebosch).

Optimisation of Magnesium Concentration

As a matter of routine, a magnesium range of 1 to 4 mM $MgCl_2$ was run for most ISSR PCR reactions.

DNA Polymerase Concentration

Many tubes of *Taq* polymerase (Promega) were used up during the trials, however the make of polymerase used was a variable that had not been changed.

Taq polymerase is known to decrease in its activity after many PCR cycles, and as the ISSR protocol required 40 cycles it was thought that the *Taq* may not be present in sufficient quantities to amplify

sufficient ISSR PCR product for gel visualisation. Therefore the amount of *Taq* used was doubled from 1 unit per 50µL total reaction volume to 2 units per 50µL reaction volume (Point 20, Appendix 4).

dNTPs

The dNTP concentration in the PCR reaction mix was not changed; the dNTP solution was made up, and used at a constant concentration throughout the trials.

Additives

Many practitioners have found that additives to the PCR reaction mixture improve the frequency of success and/or quality of PCR product (Roux, 1995; Henke *et al.*, 1997). A solution of Betaine [(carboxymethyl)trimethylammonium] was used following a publication (Henke *et al.*, 1997) that Betaine improved PCR quality. Although Betaine is only reported to improve GC rich PCR, it was added to an ISSR PCR reaction (10µL of a 5M Betaine solution was added to a 50µL PCR reaction volume).

Template Concentration

Several different DNA template concentrations were tried in the PCR mix to see if the template was responsible for the troubles encountered. Template volumes of 4, 6, and 8µL per 50µL reaction volume were tested.

Tissue Preservation

Most of the tissue extracted in the trials mentioned thus far had been preserved in silica gel. It was considered that the silica gel in which the leaf tissue has been preserved was affecting the quality of the final DNA extracts, since silica gel drying may result in DNA degradation (Nickrent, 1994). Many of the samples were collected from South America and the West Indies, where immediate extraction of fresh leaf material was not possible, so an alternative tissue preservation method was sought: preservation in a saturated NaCl/CTAB buffer (Rogstad, 1992; Štorchová *et al.*, 2000). This was compared to silica-dried and fresh leaf tissue.

Fresh Leaf Tissue

Fresh leaf samples of *C. odorata* were obtained, and extracted in the same manner as silica-dried tissue. Extracts from fresh and silica-dried tissue were then run in parallel in an ISSR PCR reaction together.

Saturated NaCl/CTAB Buffer

DNA was extracted from the samples collected in the Caribbean region and preserved in a saturated NaCl/CTAB buffer (Rogstad, 1992; Štorchová *et al.*, 2000). DNA was extracted from these samples with exactly the same procedure as the silica-dried tissue (Procedure 3), but with a rinse in distilled water in the beginning of the procedure to remove residues of the preservation buffer. Simultaneous PCR reactions were run with both silica-dried and NaCl/CTAB-preserved tissue.

Test for Inhibitory Compounds

The leaves and stems of *C. odorata* have a characteristic aroma (which has led to one of its common names: "paraffin weed"), which indicates that unusual secondary compounds are present in the tissue that could interfere with PCR reactions. Thus a possible reason for the lack of consistent success in amplifying ISSR bands from *C. odorata* samples was that secondary compounds that have not been removed during the DNA extraction process could have been inhibiting PCR amplification.

Two trials were run to test this hypothesis of secondary compounds:

- ISSR PCR with samples other than *C. odorata*
- Further purification of DNA extracts by ammonium acetate precipitation

Non – *C. odorata* Samples

A PCR trial was attempted with *Zea mays*, to test whether the difficulties encountered were as a result of the species examined. An extract from *Zea mays* that had been giving consistently positive PCR amplification for other students in the Molecular Systematics Laboratory of the Botany Department, at Rhodes University was run with other *C. odorata* samples (Point 68, Appendix 4).

Ammonium Acetate Precipitation

In a final attempt to remove any inhibitory compounds that might be present in the DNA extracts, three extractions were 'cleaned up' by ammonium acetate precipitation (Appendix 5).

Thermal Cycling Profiles

PCR was initially performed with 40 cycles of the following thermal cycling conditions

Denaturing	95°C for 45 seconds
Annealing	55°C for 45 seconds
Extension	72°C for 180 seconds, all followed by 10min at 72°C

Annealing temperatures of 50ºC and 48ºC were also tested.

Thermal Cyclers

It is well known that different thermal cyclers may have different temperature regimes, even though they may be running the same program (MacPherson *et al.*, 1993, cited in Wiesing *et al.*, 1995b; Penner *et al.,* 1993, cited in Morell *et al.*, 1995).

ISSR PCR reactions were performed on three thermal cyclers to test if the temperature regime of one of the thermal cyclers would not be better suited for ISSR amplification:

- Hybaid PCR Sprint Thermal Cycler, SPRT 001 (Hybaid Limited, UK)
- PC-960G Gradient Thermal Cycler (Corbett Research)
- GeneAmp PCR System 9700 (PE Applied Biosystems)

Results and Discussion

Gel Visualisation

In a comparison between the usefulness of agarose gels with ethidium bromide and polyacrylamide gels with silver staining, it was found that, with a particular ISSR product, the polyacrylamide could resolve clear ISSR bands (Figure 2.2) where only smears were visualised with ethidium bromide in agarose gels. It was therefore decided that all ISSR PCR products would be visualised with the Mighty Small polyacrylamide and silver staining system.

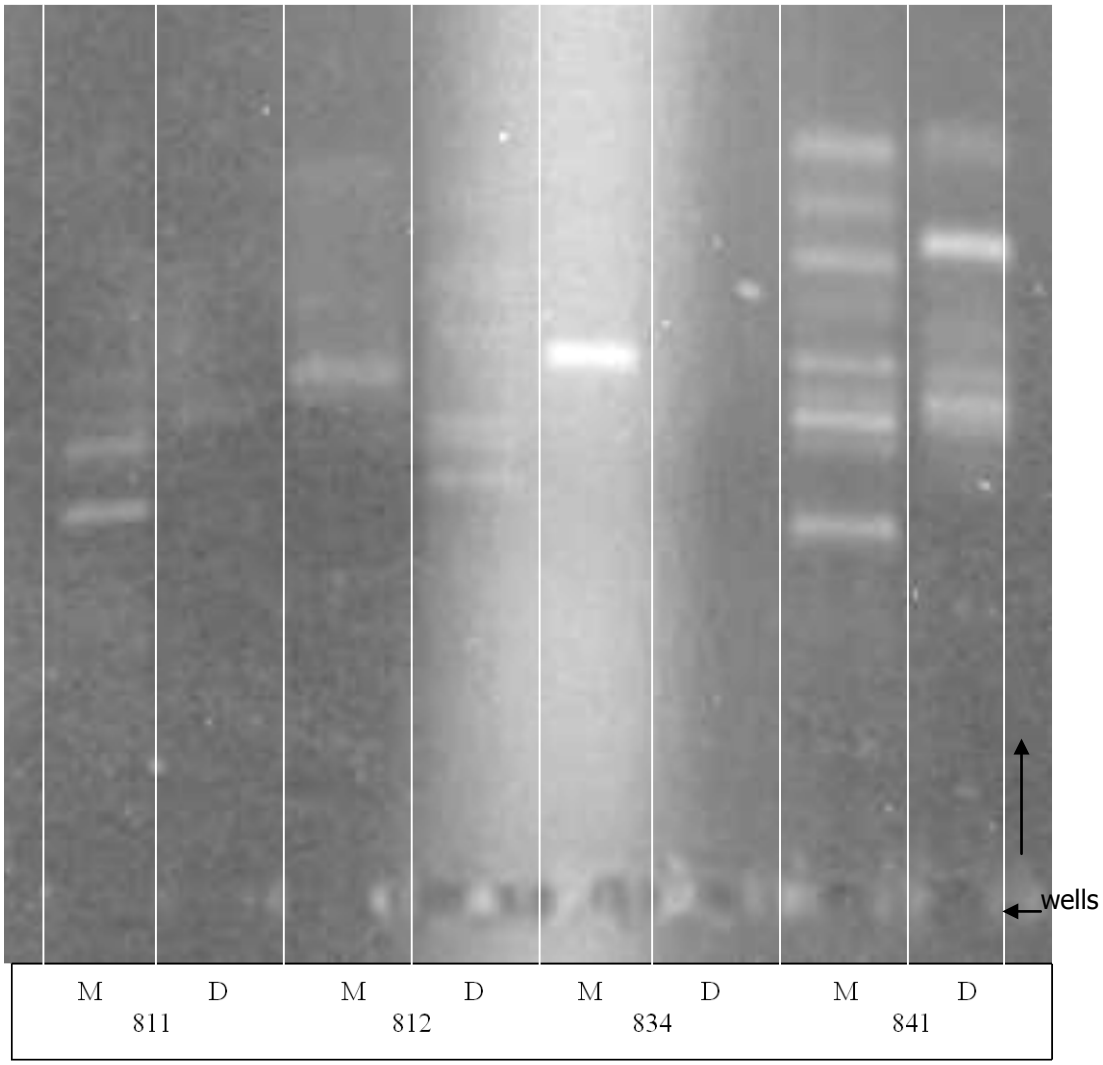

Figure 2.2: Initial successful ISSR PCR amplification. Numbers refer to primers [each with two samples: Mexico10 (M) and SA_Dbn13 (D)]. Vertical arrow shows gel direction.

Comparisons of DNA Extraction Protocols

CTAB-Based Extraction Protocols

It has been noted that the quality of the DNA preparation has little influence on the amplified polymorphic bands (Wiesing *et al.*, 1995; Wolfe *et al.*, 1998), yet Fang and Roose (1997) found through a systematic study of factors affecting fingerprint patterns that high-quality DNA template was essential to obtain a large number of well-resolved fragments.

RNA in DNA template preparations is suggested to interfere with the PCR reaction and reduce the number of fragments that can be scored (Fang and Roose, 1997). The preservation of plant specimens by silica gel may have detrimental effects on sample DNA (Adams *et al.*, 1999), yet it was found here that *C. odorata* samples preserved with silica gel produced better quality DNA extractions than sample preserved in the saturated NaCl/CTAB buffer (Rogstad, 1992; Štorchová *et al.*, 2000), and were only of very slightly poorer quality than DNA extractions from fresh leaf tissue. Chase and Hillis (1991) found that samples dried in the field contained DNA that was intact enough to be useful in restriction site studies and of more than sufficient quality for PCR amplification in gene sequencing studies. This study corroborates the findings of Chase and Hillis (1991), rather than those of Adams *et al.* (1999).

Of the three CTAB extraction protocols, the third (with PVP and ethanol-grinding) appeared to be the best for ISSR results, i.e., in this dissertation the samples extracted by the third method gave the darkest smears or smudges or most often produced clear bands on a polyacrylamide gel with silver staining.

UV Absorbance Values

No meaningful UV absorbance values were obtained, and it was concluded that the DNA concentrations were too low to be detected accurately, as many of the readings were negative. The spectrophotometer used was not a DNA-specific spectrophotometer, and the smallest volume cuvette that it uses is 1mL. Therefore, the DNA extractions had to be diluted to make up a volume of 1mL. Because of this dilution, the DNA concentrations were too low to be accurately read by the spectrophotometer.

Commercial DNA Extraction Kit

Extraction with the Qiagen DNeasy Plant Extraction Kit was more labour-intensive than any of the permutations of the CTAB extraction protocol tried. The extractions from these kits did not produce amplification product that was better than any of the CTAB-based extracts, and it was found that the

Qiagen extracts produced less PCR product than the CTAB-based method with PVP and ethanol-grinding. Extracts using the Qiagen kit never produced distinct ISSR bands.

Observations from Multiple Extractions

Simultaneously with the ISSR trials, these initial extractions were also being used for PCR of the ITS region (Chapter 3). The samples that produced consistently good ITS product were utilised as positive controls (or extractions most likely to work) in the ISSR trials. As a result the 200µL extraction volumes were quickly used up while extracts were still needed. Thus, for most samples in this dissertation (particularly the samples used in trials) several separate extractions were used.

During these sequential multiple extractions, it was found that the quality of PCR product differed from one extraction to the next, despite attempts to keep extraction conditions as homogenous as possible. In addition, it was found that some samples consistently produced high quality extractions (i.e., extractions that gave consistently positive results). The reason(s) that these samples produced good extracts could not be discovered; but it is possible that the leaves sampled were younger (or older) than most other samples, or that secondary compounds common in other leaves were not present in the sampled plants (due to phenotypic or genotypic differences).

Because the DNA concentrations in the DNA extractions were too low for meaningful DNA concentration and purity determination with available equipment, it was not possible to confirm whether the extractions that gave consistently good results had high DNA concentrations or high purity.

PCR Reagents

Primers

Primer Screening

From the primer screening process, it was found that four primers gave polymorphic ISSR bands. They are primers 841, 812, 888, 891. Primer 841 produced the most bands (see Figure 2.2).

Through the screening process, several PCR reactions failed to produce ISSR product or produced poor product. These reactions were repeated (without changing any conditions or reagents) and very often produced improved PCR product in the repeat (e.g., Point 3, Appendix 4) for no determinable reason.

Primer Volumes

It was found that the primer volumes tested of 1 and 2µL per 50µL reaction volume (1.3 and 2.7pM respectively) gave some ISSR PCR product (though there was no difference between them), while the 4 and 8µL volumes gave no product. It was therefore decided that primer volumes of 2µL per 50µL would be used, as there was less likely to be pipetting error with the larger volume. The failure to amplify ISSR PCR product with primer volumes of 4 and 8µL per 50µL reaction volume indicates that at higher concentrations the primer inhibits ISSR PCR.

Control Against Primer Degradation

Tests with both the UBC and re-manufactured new primers (Points 16 and 17, Appendix 4) both failed to produce PCR product. It was decided, though, that future reactions would be performed with the new, and therefore presumably more reliable primers from the University of Cape Town.

A high homology between simple sequence-primers and their binding sites can be assumed (Wiesing *et al.*, 1995). Gupta *et al*. (1994) found that tetranucleotide repeats were most effective for a variety of eukaryotic genomes. They (Gupta *et al*., 1994) found that the dinucleotide-based primers amplified smears and not patterns; the tetranucleotide-based primers in Gupta *et al.'*s (1994) study were most effective in amplifying the polymorphic patterns. This finding is supported by Wiesing *et al*. (1995) who found that the most distinct amplification product bands were produced by all but one of the tetranucleotide-based primers that they used. Wiesing *et al*. (1995) also achieved the same level of product amplification with all of the GC-rich trinucleotide-based primers that they used. Forty five percent of Gupta *et al*.'s (1994) trinucleotide-based primers amplified multiband patterns from most species, however, only a few species exhibited polymorphism with each of the trinucleotide-based primers. Wiesing *et al*. (1995) found that all dinucleotide repeats and some AT-rich trinucleotide repeats resulted in a smear rather than discrete bands, and that primers with 100% AT content did not yield any products under the chosen PCR conditions. Primers based on AT and GC repeats should be avoided, as they are self-annealing due to sequence complementarity (Blair *et al*., 1999).

Primer 841 used in this study was a 3'-anchored AG (dinucleotide) repeat. The literature cited above suggests that a trinucleotide or tetranucleotide repeat might have more successfully produced clear amplification bands. This was not attempted in this study, as the first PCR attempts with the AG-based primer resulted in clearly defined bands. Subsequent PCR amplifications yielded very little and, unable to predict where the troubles lay, attempts were made to repeat unsuccessful amplification, and occasionally succeeded. It was thought that changing to a different primer would increase the number of variables to be tested, and was not necessary as it had been shown that ISSR amplification with this primer was possible.

Optimisation of Magnesium Concentration

It was found that the samples that were run at magnesium concentrations of 2 and 3mM gave the most polymorphic ISSR bands, but it was found that the banding pattern of the profiles differed between the two magnesium concentrations for the same samples. Thus, most ISSR PCR reactions were done at a magnesium concentration of 2.5mM, in the hope that this intermediate concentration would allow amplification of all the ISSR bands from both the 2 and 3mM reactions.

DNA Polymerase Concentrations

The increased *Taq* concentrations did not improve the success rate of ISSR amplification, though it was later found that 1.25 units of *Taq* polymerase per 50µL total reaction volume appeared to improve ISSR PCR amplification slightly.

Additives

The addition of Betaine was not found to improve ISSR amplifications (Points 41 to 44, Appendix 4).

Fang and Roose (1997) found that the inclusion of formamide in the PCR reaction was essential for repeatable amplification of ISSR bands, and for reduced background and smearing on gels. Formamide may influence primer-template annealing and melting temperatures (Tsumura *et al.*, 1996, cited in Fang and Roose, 1997). If formamide was excluded from the reaction mixture, no or fewer fragments were detected (Fang and Roose, 1997). However, formamide concentrations of 3% or above inhibited amplification completely; 2% formamide in the reaction mixture generally gave good results (Fang and Roose, 1997). This PCR additive was never tested in this dissertation; a future investigation into the ISSR region of *C. odorata* may benefit from the use of formamide.

Template Concentration

Of all the template volumes tested (concentrations could not be determined; see previous mention of UV absorbance), 4µL per 50µL reaction volume most often produced results. The failure to amplify ISSR PCR with higher volumes of template may indicate that something in the DNA extractions may have inhibited ISSR PCR amplification. It is possible that secondary compounds present in the leaf tissue, or preservation medium, were carried over in all four DNA extraction protocols.

Tissue Preservation

Fresh Leaf Tissue

The DNA extracts for fresh leaf tissue produced slightly better quality ISSR PCR product than that of silica-dried samples, but once again the results were not repeatable.

This would seem to indicate that there is a small degree of DNA degradation in silica-dried leaf tissue of *C. odorata*, but not enough to be indicated as the cause of the troubles encountered in attempts to successfully amplify ISSR PCR product.

If the DNA extracted from the fresh leaf tissue had successfully produced reproducible ISSR bands, then fresh tissue from within South Africa would have been collected and immediately extracted for use in a study on the population structure of *C. odorata* in South Africa according to ISSRs.

Saturated NaCl/CTAB Buffer

The samples preserved in the saturated NaCl/CTAB buffer failed to produce ISSR PCR product (Point 30, Appendix 4), indicating that the preservation of leaf tissue of *C. odorata* for DNA analysis is better with silica gel than with the saturated NaCl/CTAB buffer.

Test for Inhibitory Compounds

In the tests to determine whether inhibitory compounds were preventing ISSR amplification, no ISSR product was obtained for any of the samples. Therefore, because ISSR product was not obtained for the *Zea mays* extract, nor the ammonium acetate-purified extract, nor the normal CTAB-based extracts (Procedure 3), it was not possible to compare these results, and thus to determine whether inhibitory compounds were preventing ISSR amplification.

Some extracts that were used in the ISSR trials successfully amplified the ITS region for sequencing with relatively few troubles (Chapter 3). This would seem to indicate that there are no compounds inhibiting PCR, and that the source of the difficulties in amplifying ISSRs lies elsewhere, possibly in the ISSR protocol itself.

Thermal Cycling Profiles

Reduction in the annealing temperature to 50°C, resulted in clear, scoreable ISSR bands being produced (Figure 2.2). Later, when ISSR PCR amplification failed almost constantly, the annealing temperature was further reduced to 48ºC, but there was still no product amplification at this lower temperature.

Thermal Cyclers

No differences in quantity or quality of ISSR PCR product were detected between any of the three PCR thermo cyclers tested. This seemed to indicate that the thermal cyclers were not at fault during the ISSR amplification trials.

Molecular Weight of Occasionally-Produced ISSR Bands

Estimates from molecular-weight markers from the occasional successes (Figure 2.3) indicate that the molecular weight range of the ISSR bands was between ca. 1000 and 300bp.

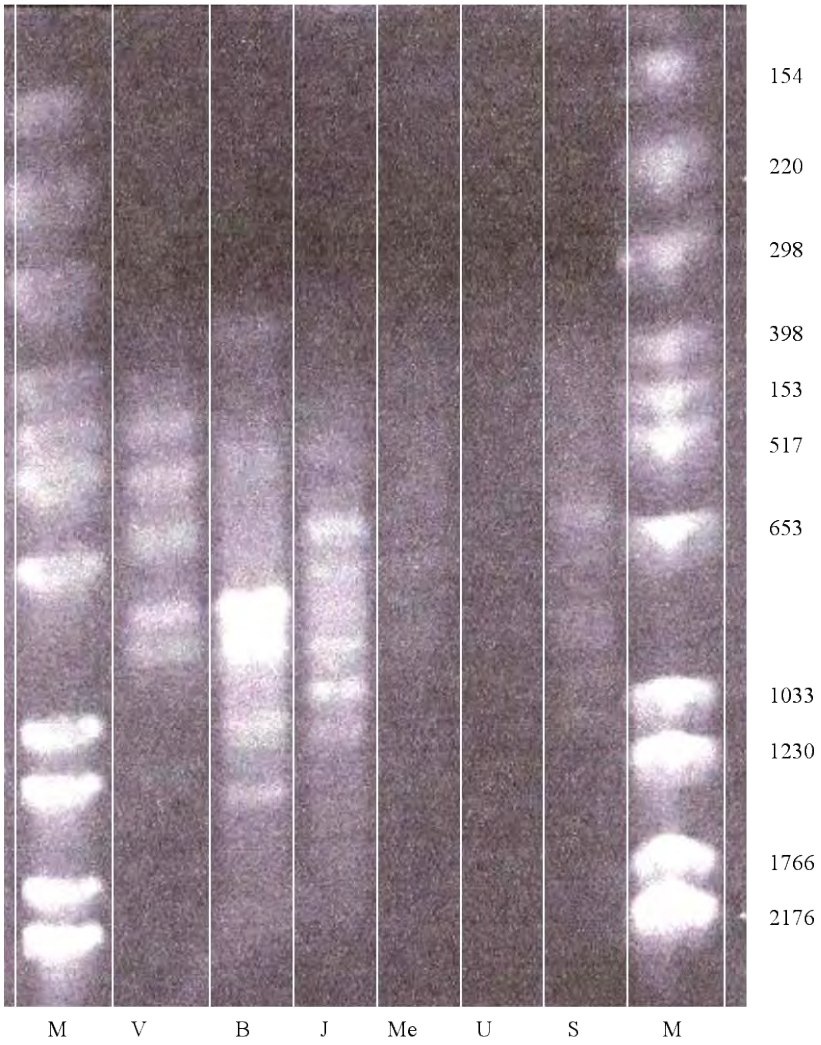

Figure 2.3 1% Agarose gel with ethidium bromide showing ISSR PCR amplification (with background smears). Molecular markers on the right hand side indicate that the ISSR bands are between 1000 and 300bp long. Sample names are: M=(molecular markers); V= (Venezue2); B= (Brazil_5); J= (Jamiaca8); Me= (Mexico10); U= (USA_Fl12); S= (SA_Dbn13);.

Summary and Future Recommendations

The ISSR fingerprinting technique was chosen for the genetic profiling of *C. odorata* because of its reported high reproducibility, ease of use, and low running costs. The first few ISSR PCR reactions produced clear, polymorphic ISSR bands. However, the frequency with which ISSR product was produced decreased thereafter, as did the quality of the product, and without any apparent reason.

Therefore, the series of trials outlined in this chapter were embarked upon, in which the variables in PCR amplification were tested. The reason for the failure to amplify ISSR bands was never found , but dubious quality Taq DNA polymerase is suspected, as this proved to be cause of PCR failure in other studies in the same laboratory.

Though occasionally clear ISSR markers were obtained (e.g., Figures 2.2 and 2.3), these results were not repeatable. On the occasions where ISSR bands were clear enough to score, the samples run were often optimisations of single samples, and therefore preliminary scoring of these un-repeatable gels was not possible.

The number of polymorphic bands that are visible in Figure 2.2 range from 1 to 9, which is similar to the number of ISSR bands found by Camacho and Liston (2001) (using ethidium bromide staining of 1.2% agarose gels); they had 1 to 7 bands per marker. This would seem to indicate that the primers screened (particularly primers 841 and 812) have the potential to produce phylogenetically informative bands.

CHAPTER 3

SEQUENCING

Introduction

One of the primary attractions of nucleic acid sequencing is the fact that the characters (nucleotides) are the basic unit of information available for any and that the potential sizes of informative data sets are immense (Hillis *et al.*, 1990). DNA sequencing of PCR products has become the method of choice for most studies of DNA sequence variation because of the large number of sites surveyed and the additional information provided about the location and nature of the sequence variants (*e.g.*, coding, non-coding, replacement, silent) (Aquadro, *et al.,* 1998).

Although possibly more difficult, DNA sequencing affords a much finer level of resolution than isozymes by detecting not only all synonymous and nonsynonymous substitutions in coding regions, but also variation in noncoding sequences (Doyle, 1995). Another important advantage of nuclear DNA sequencing over isozymes is that it is always possible to obtain more detailed information on markers, and eventually to trace a polymorphism to the level of DNA sequence comparisons (Bachmann, 1997).

Although there are a number of strategies for obtaining sequence data, all methods have four steps in common. First, a particular target sequence must be identified that contains an appropriate amount of variation across species or individuals for the problem that is to be addressed. Second, large numbers of copies of the target sequence must be isolated and purified from each individual to be examined. Third, the purified DNA or RNA must be sequenced. Finally, homologous sequences must be aligned and analysed (Hillis *et al.*, 1990).

The methods differ primarily in how the nucleic acid is isolated: "direct" methods involve either directly amplifying the target DNA or isolating abundant RNA transcripts; while indirect cloning methods involve the preparation and isolation of viral and/or bacterial vectors that contain copies of the sequence of interest (Hillis *et al.*, 1990). Direct sequencing with PCR is employed in this study.

Intraspecific Molecular Variation in Plants

Gene genealogies offer great promise for furthering current understanding of plant evolution (Schaal *et al.*, 1998), and When genetic variation is organised into a genealogy, the resulting analysis of the

overlay of geography on genealogy has been called "intraspecific phylogeography" (Templeton *et al.*, 1995). In a phylogeographic analysis, genealogical lineages are used to make inferences about the principles and processes that have resulted in geographical distribution (Avise, 1998; Schaal *et al.*, 1998; Schaal and Olsen, 2000), in a process that tracks geographical divergences along a phylogenetic tree (Smouse, 1998). The success of a phylogeographic analysis is dependent on the integrity of the phylogenetic tree, as well as the interpretation of congruence between the phylogeny and geographical distribution (Schaal *et al.*, 1998).

Pictorial overlays of haplotype network upon geography are currently the standard inference tool for intraspecific phylogeography (Templeton *et al.*, 1995). Such pictorial representations usually find a strong association between the geographic location of haplotypes and their evolutionary position within a gene tree, but the demonstration of such an association *per se* does not reveal the causes of the association (Templeton *et al.*, 1995).

The ability of phylogeographic methods to detect geographic associations depends upon there being mutational resolution in the haplotype tree, and historical events cannot be older than the coalescence time for the gene region being investigated (Templeton *et al.*, 1995). However, in order to construct gene trees (for phylogeographic use), significant genetic variation must occur at the appropriate level (*i.e.* among the populations or taxonomic units being investigated) (Schaal *et al.*, 1998).

In the search for highly informative markers within a single species, genetic markers with recurrent mutation rates high enough to yield multiple mutations over the time-frame of interest are being chosen (Smouse, 1998). This recurrent mutation, coupled with recombination, can create homoplasy in the phylogenetic tree on which phylogenetic analyses are based (Smouse, 1998).

High-resolution nuclear markers such as RAPDs and AFLPs are historically unordered, leading to phylogenetic ambiguity in the phylogenetic tree (Smouse, 1998; Schaal and Olsen, 2000). An alternative source of variation, the non-coding regions of 'single-copy' (low copy number) nuclear genes, could potentially provide multiple, unlinked allele genealogies at the intraspecific level (Olsen and Schaal, 1999), but they have not yet been extensively studied in plants. An example of a potentially useful single-copy nuclear gene is the nuclear chloroplast Glutamine Synthetase (ncpGS) gene, as described by Emshwiller and Doyle (1999).

Phylogeographic studies in plants have lagged behind those of animal studies, primarily because of difficulties in finding ordered, neutral intraspecific variation required for constructing gene trees (Olsen and Schaal, 1999; Schaal and Olsen, 2000). The detection of this phylogenetically informative intraspecific variation is probably the single most difficult problem facing plant population biologists interested in using the techniques of phylogeography (Schaal *et al.*, 1998).

Most attempts to detect intraspecific variation in plants have relied on the chloroplast genome (Olsen and Schaal, 1999), which has low rates of sequence evolution (Doyle, 1995), but with only varying success (Olsen and Schaal, 1999).

In plants, non-recombining chloroplast DNA (cpDNA) and mitochondrial DNA (mtDNA) sequences are much more conserved than animal mtDNA sequences, and rates of sequence evolution in plant mitochondrial genomes are too low for phylogeographic use (Doyle, 1995; Olsen and Schaal, 1999), a serious limitation for intraspecific and phylogeographic studies (Mousadik and Petit, 1996). Although several small regions of the chloroplast genome (such as some intergenic spacers) show potential for phylogenetic analysis, attempts by Schaal *et al*. (1998) indicate that single cpDNA loci are only occasionally useful at the intraspecific level. Approximately half of all cpDNA mutations are short indels (1-10 bases), located primarily in the noncoding regions of the chloroplast genome (Zurawski *et al.*, 1984, and Kanno and Hirai, 1992, both cited in Mousadik and Petit, 1996), and in many cases, the chloroplast spacer regions that have been informative for some species show little or no intraspecific variation in other plant species (Schaal and Olsen, 2000). However, studies of noncoding chloroplast DNA sequences have shown that insertions/deletions of more than two bases that do not belong to tandem repetitions are good phylogenetic markers (Gielly and Taberlet, 1994a, 1994b, 1996, cited in Mousadik and Petit 1996).

The effective population size of a nuclear gene is four times that of an organelle gene, because it is diploid and bi-parentally inherited (Schaal and Olsen, 2000). This larger effective population size results in larger coalescence times, because nuclear DNA (nDNA) evolves more slowly than organellar DNA (Palumbi *et al.*, 2001), which in turn increases the likelihood of encountering ancestral polymorphisms (Schaal and Olsen, 2000). The utility of nuclear gene (*e.g.* intron) sequences in intraspecific phylogenetic analyses appears to be limited by this increased coalescence time (and associated variance) of nuclear genes as compared to mitochondrial and chloroplast genes, and the potential for reticulate evolution among nuclear alleles due to recombination (Bermingham and Moritz, 1998).

Several factors contribute towards an increase in genetic structure for organelle genes in comparison to nuclear genes (Dumolin-Lapègue *et al.*, 1997). Firstly, effective gene flow is limited to seeds for maternally inherited genomes (Petit *et al.*, 1993b, cited in Dumolin-Lapègue *et al.*, 1997). Secondly, in hermaphroditic species such as the oaks, the flowering and fruiting pattern result in an effective number of trees that contribute towards the next generation as females is less than the effective number of trees acting as males (Demesure *et al.*, 1996; Dow and Ashley, 1996, both cited in Dumolin-Lapègue *et al.*, 1997), and thirdly, drift is twice as strong for a haploid genome as compared to a diploid one (Dumolin-Lapègue *et al.*, 1997). This is because of the diploid nature of the nuclear gene.

It would be desirable to consider the multiple gene trees that may be provided by both the chloroplast and nuclear genomes (Schaal *et al.*, 1998). However, Doyle (1997) warns that when the history of the organellar genome is different from that of the nuclear genome, every nucleotide in every one of the organellar genome's genes will give the "wrong" phylogenetic pattern for those taxa affected (*e.g.*, in lineage sorting or introgression).

Because of the difficulty in finding genealogically informative markers, many plant studies have been phylogeographic only in the broad sense, meaning that they detect an association between patterns of genetic variation and geography. Such studies do not incorporate a genealogical perspective (Schaal and Olsen, 2000).

Plant molecular phylogenetic studies at the lower taxonomic levels are limited by the availability of sequences with sufficient variation to be suitable for the construction of well-supported trees (Doyle *et al.*, 1996). Single sequences are of limited value when recombining diploid genomes in populations are being investigated (Bachmann, 1997). This limitation can be overcome by sampling sequence polymorphisms at many points throughout the nuclear genome in order to obtain multilocus genotypes; various molecular methods have been designed to this end (Bachmann, 1997). Ultimately the full potential of phylogeography will be realised when multiple loci are considered together (Schaal *et al.*, 1998).

The growing examples in recent literature have demonstrated that low-copy nuclear genes have a great potential to compensate cpDNA and nrDNA for the improvement of resolution and robustness of plant phylogenetic reconstruction. However, the phylogenetic utility of low-copy nuclear genes has been confounded by the complex evolutionary dynamics of nuclear gene families. For example, gene duplication and selection can potentially lead to reconstruction of gene duplication events (parology) rather than speciation events (orthology) (Sang, 2002). The extra effort required to disentangle orthology and parology apparently has discouraged the common application of nuclear genes in plant phylogenetics (Sang, 2002).

A generalisation emerging from recent data on genome sequencing and mapping is that most 'single-copy' genes belong to larger gene families, even in putatively diploid species. Many angiosperm genomes have experienced several cycles of polyploidisation at various times in the past. The more ancient of these past genome doubling events may be difficult to discern, due to potentially rapid evolutionary restoration of diploid-like chromosomal behaviour. Most angiosperms are therefore considered to have 'paleoplolyploid' genomes (Wendel, 2000).

Chloroplast DNA restriction site map data meet this requirement (sampling at many points) and have been used for this purpose (low-taxon studies). However, it is generally accepted that DNA sequence data have numerous technical advantages over restriction mapping (Doyle *et al.*, 1996). The use of nucleotide sequences has become an even more powerful approach than restriction site analysis of cpDNA (Käss and Wink, 1995), and efforts have been made to identify relatively rapidly evolving chloroplast sequences for use in DNA sequence-based phylogenetic studies.

The chloroplast *rbcL* gene was one of the first widely sequenced genes, as was the *ndhF* chloroplast gene (which provides approximately three times more phylogenetic information than *rbcL*) (Kim and Jansen, 1995; Käss and Wink, 1997). More recently the number of highly variable nuclear genes that are suitable for sequence studies at lower taxonomical levels have had increased, such as the Histone H3 intron sequences (Doyle *et al.*, 1996) and the chloroplast- expressed Glutamine Synthetase gene (ncpGS) (Emshwiller and Doyle, 1999). The ncpGS gene is a nuclear encoded gene containing several introns, and appears to be single copy in most taxa. Levels of variation among the ncpGS sequences compare favourably with those of the Internal Transcribed Spacer (ITS; Emshwiller and Doyle, 1999). The ITS region of the 18S – 36S ribosomal repeat is one of the most commonly sequenced regions in plant systematics, however, its use at the intraspecific level is debatable (Baldwin, 1993; Kim and Jansen, 1995; Schaal *et al.*, 1998; Cullings and Vogler, 1998).

Despite the uncertainty surrounding the use of ITS sequences at intraspecific levels, it was decided that the ITS region would be sequenced as part of this study, because of its ease of use and because it was still thought that there was some potential for its phylogenetic use at the intraspecific level. In addition to the ITS region, another part of the ribosomal repeat unit, the external transcribed region (ETS) was sequenced.

The 18-26S ribosomal repeat (Figure 3.1) is attractive for phylogeny reconstruction because of its high copy number [there are hundreds of tandemly repeated nuclear ribosomal DNA (nrDNA) copies located on one or more chromosomes], rapid concerted evolution, and diverse rates of evolution within and among component subunits and spacers (Baldwin, 1993, and literature therein).

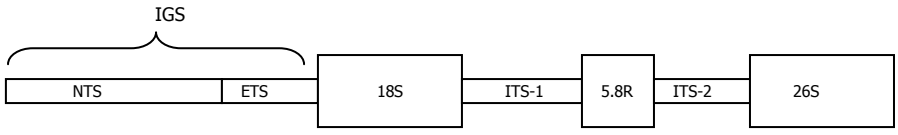

Figure 3.1 The nuclear ribosomal 18S to 26S repeat unit, showing positions of the ETS, and ITS-1 & ITS-2. According to Baldwin (1993) and Baldwin and Markos (1998). Gene sizes are not shown to scale. (ITS = Internal Transcribed Spacer, ETS = External Transcribed Spacer, NTS = Non-Transcribed Spacer, IGS = Intergenic Spacer)

Variation in 18-26S nrDNA among closely related plant species has been detected largely within the intergenic spacer (IGS) and internal transcribed spacer (ITS) regions of the 18-26S nrDNA (Baldwin, 1992 and literature cited therein).

The multiple copies of this ribosomal gene are subject to concerted evolution (Weising *et al.*, 1995b; Bachmann, 1997; Schaal *et al.*, 1998) which explains some cases of incongruence between ITS data and other characters (Bachmann, 1997), as not all molecular markers are subject to the homogenisation effect of concerted evolution. The phylogenetic interpretation of sequences from the 18S – 26S repeat are considerably less straightforward than those of chloroplast DNA polymorphisms, and results from ribosomal DNA (rDNA) analysis should be treated with caution as concerted evolution and gene conversion may mask hidden phylogenetic signal (Weising *et al.*, 1995b).

Intergenic Spacer Region

The nrDNA IGS (intergenic spacer) consists of two parts, the non-transcribed spacer (NTS), and the ETS, which is adjacent to the 5'–end of the 18S gene (Volkov *et al.*, 1996). The IGS contains different regulatory elements (which are necessary for transcription initiation and termination) evolves far more rapidly than the nrRNA coding regions, and may thus differ between species, populations and even individuals within a population (Volkov *et al.*, 1996; Cullings and Vogler, 1998, and literature therein).

External Transcribed Spacer Regions

The ETS of the IGS of angiosperms is longer than ITS-1 and ITS-2 combined and, based on restriction site data, appears to evolve at least as rapidly as the ITS at the nucleotide level (Baldwin and Markos, 1998, and literature therein). The level of subrepeat identity within the ETS region in the Solanaceae family ranges between 57% to 92% (Volkov *et al.*, 1996). The essential differences in ETS length in different species in the genus *Nicotiana* may be due to variation of the number of subrepeats in the ETS (Volkov *et al.*, 1996).

The primary barrier to sequencing the ETS in plants is the lack of a highly conserved region for primer design flanking the 5' end of the spacer, (Baldwin and Markos, 1998). Although the highly conserved 18S gene offers various options for primer sites downstream from the 3' end of the ETS, the highly variable nontranscribed spacer (NTS) borders the 5' end of the ETS and is too rapidly evolving in sequence and length to provide an universal primer site for plants. Baldwin and Markos (1998) overcame this limitation by using long-distance PCR to amplify the entire IGS (NTS + ETS) using universal primers that bind to the flanking and highly conserved 18S and 26S sequences. From the resultant PCR products, they found a relatively invariable site from which they produced a primer suitable for PCR amplification of a shorter region of the ETS.

Internal Transcribed Spacer Regions

At present, the internal transcribed spacer regions are technically the most convenient and universally accessible nuclear DNA sequences with sufficient variation to distinguish species of a genus or even populations of a species (Bachmann, 1997). The ease with which ITS sequence information can now be obtained makes it a promising nuclear DNA resource for comparison with cpDNA phylogenetic data (Baldwin, 1993).

One of the advantages of sequencing the ITS region is that it is noncoding and therefore contains a relatively high level of variability. Another advantage to sequencing the ITS region is that it is flanked by highly conserved regions from which universal primers can be obtained (White *et al.*, 1990; Schmidt and Schilling, 2000) and that can serve as reference points for sequence alignment (Baldwin *et al.*, 1995, cited in Schmidt and Schilling, 2000).

The ITS region is part of the transcriptional unit of nrDNA, but the spacer segments of the transcript are not incorporated into mature ribosomes. Instead, ITS-1 and ITS-2 regions of the nrDNA transcript appear to function, at least in part, in the maturation of nrRNA's. It seems probable that ITS-1 and ITS-2 are under some evolutionary constraint in structure and sequence, as suggested by size and GC content comparisons among angiosperms (Baldwin *et al.*, 1995). However that they are not used in any part of the ribosomal molecule suggests that they are under minimal evolutionary constraints, and therefore phylogenetically useful at lower taxonomic levels.

The ITS region of 18S-26S nuclear ribosomal DNA includes three components: the 5.8S subunit, an evolutionarily highly conserved sequence, and two spacers designated ITS-1 and ITS-2 (Baldwin *et al.*, 1995). It would seem that the combination of highly conserved primer sites with highly variable sequences is rare (Bachmann, 1997), and is extremely useful for PCR amplification.

The ITS region has been shown to be evolutionarily conservative in length, as opposed to the IGS, which varies extensively in length. The ITS region in plants is also small enough such that relatively few restriction sites occur within it, yet sequencing of the ITS region has potential as a source of nuclear DNA characters for phylogenetic reconstruction in plants (Baldwin, 1992). Preliminary indications of ITS length conservation and high ITS nucleotide sequence variability suggest that DNA sequences of these spacers might be readily alignable across closely related taxa, yet sufficiently variable to allow resolution of lower-level phylogenetic questions in angiosperms (Baldwin *et al.*, 1995).

Use of ITS in Asteraceae

ITS sequences have been studied in many groups of Asteraceae (Clevinger and Panero, 2000; Bain and Golden, 2000; Konishi *et al.*, 2000; Urbatsch, 2000; Oberprieler, 2001; Markos and Baldwin, 2001; Fernández *et al.*, 2001; Lowrey *et al.*, 2001) and have proven to be a useful source of information at the generic level to resolve phylogenetic relationships (Schmidt and Schilling, 2000, and literature therein). Schmidt and Schilling (2000) found that the mutation rates estimated for the Eupatorieae ITS region (1-3% per million years) were in general very high relative to rates reported for this gene region in other plant groups. They suggest that life history traits, such as shorter generation time, may partially explain higher mutation rates in the ITS region of the Eupatorieae.

Baldwin (1993) found that the ITS region in *Calycadenia* and *Osmadenia* has evolved primarily by point mutations. Such structural conservatism of ITS sequences was attributed to their role in the production of mature rRNAs from primary transcripts, *i.e.*, in formation of secondary RNA structures that bring the ends of the 18S, 5.8S, and 26S regions into close proximity for processing (Baldwin, 1993, and literature cited therein).

Due to close agreement between the ITS consensus tree and parsimony-based interpretations of cytological and morphological data from *Calycadenia*, Baldwin (1993) suggested that the ITS region could prove to be useful for addressing phylogenetic questions among closely related plant species.

Potential for ITS in Intraspecific and Phylogeographic Use

The ITS region has shown its greatest utility in phylogeny reconstruction at the interspecific and intergenic levels, providing a source of data comparable to variation in cpDNA restriction sites (Mayer and Soltis, 1999). Because cpDNA variation, with a relatively slow rate of sequence divergence (Sang, 2002) has in many cases provided phylogenetic insight within species, it seems likely that ITS variation should hold similar promise for intraspecific phylogenetic investigations. Several studies have suggested this potential (Sytsma and Schaal, 1990; Walker and Paris, 1997, both cited in Mayer and Soltis, 1999; and Baldwin, 1993; Shaw, 2000; Vanderpoorten *et al.*, 2001; Comes and Abbott, 2001; Kropof *et al.*, 2002), however analyses of ITS variation among intraspecific populations are still uncommon (Mayer and Soltis, 1999).

The chloroplast genome typically is predominantly uni-parentially inherited (usually maternal in Angiosperms), thereby useful for tracking haplotype lineages and discriminating between maternal and paternal contributions to offspring, though it is unable to provide direct evidence for hybridisation (Sang, 2002). It also In contrast, nuclear ribosomal genes, such as the ITS, are recombining, biparentially inherited markers, thereby allowing simultaneous observation of maternal and parental

contributions to offspring, exposing gene flow and hybridisation events that many not be revealed by analysis of an organellar genome or single genealogy alone (Mayer and Soltis, 1999).

The potential of ITS for intraspecific phylogeny estimation also carries with it pitfalls, for micro evolutionary processes such as gene flow and lineage sorting can result in reticulate patterns of relationship among populations (Mayer and Soltis, 1999). Furthermore, if gene flow leads to heterozygosity of ITS types within an individual, subsequent recombination or partial gene conversion can disrupt phylogenetic signal in the ITS genealogy.

Low levels of sequence divergence reported among some species of long-lived plants has likely contributed both to warranted and unwarranted reluctance to employ ITS at the intraspecific level (Mayer and Soltis, 1999, and literature therein).

Schaal *et al*. (1998) suggest that the internal transcribed spacer (ITS) region of ribosomal DNA is generally not considered conducive to phylogeographic study. First, for most species examined, they claim that intraspecific variation has not been detected in this region (Schaal *et al*., 1998). Furthermore, as part of a multicopy gene family, the ITS region is subject to the poorly understood process of concerted evolution, which confounds interpretation of sequence polymorphism at the intraspecific level.

In a study on the *Calycadenia* (Compositeae), Baldwin (1993) found that ITS sequence divergence could be found both between and within species. This study found that the ITS sequences offer potential for the resolution of intraspecific relationships among disjunct populations in *Calycadenia*. Up to 3.7% combined ITS-1 and ITS-2 nucleotide sequence divergence was found between populations of a single species of *Calycadenia* (Baldwin, 1993).

In a study to confirm the biogeographic origin of an invasive alien species to Australia, *Senecio madagascarensis*, Scott *et al*. (1998, cited in Radford *et al.,* 2000) found that the ITS1 data was too invariable to allow them to differentiate between two species: *S. madagascarensis* and *S. inaequidens*, and isozyme data proved to be more informative than ITS nuclear data (Radford *et al*., 2000).

One of the primary limitations of ITS data for phylogenetic studies in angiosperms is the small number of characters from these short spacers. Simulations by Huelsenbeck and Hillis (1993, cited in Baldwin *et al*., 1995) suggest that sequences of such short length are, under most conditions and types of analysis, less effective for accurate tree reconstruction than longer sequences. Useful variation must be more highly concentrated within a set of ITS-1 and ITS-2 sequences than in longer DNA regions in order to achieve the same level of phylogenetic resolution and support (Baldwin *et al*., 1995).

Limitations imposed by the small size of the ITS region can be overcome, in part by combination of ITS data with sequence data from the external transcribed spacer (ETS) region of 18-25S nrDNA (Baldwin, 1993). While the majority of the data presented here is ITS, the ETS region was sequenced from five samples in an attempt to complement the potentially limited phylogenetic signal at the intraspecific level in *C. odorata*.

Materials and Methods

Both the ETS and ITS were successfully sequenced in this study, however only five samples were sequenced for the ETS, as difficulties were encountered. Focus was thus placed on sequencing the ITS region. The results of the five ETS sequences are compared with the ITS sequences of the same five samples in order to investigate the relative variability and phylogenetic usefulness of these two regions. The complete ITS intraspecific analysis is also presented and discussed in this chapter.

Sampling Strategy

All samples (Table 3.1) were collected, silica dried (Chase and Hillis, 1991) and supplied by the Weeds Division of ARC-PPRI, or collected by the author.

Samples were collected from regions representing the entire distribution (native and invasive) of *C. odorata*, and specific emphasis was placed on samples of Caribbean origin because the various morphological studies (Vos, 1989; Erasmus, 1990, unpubl.; Morris *et al.*, 1999) indicated that similarities were evident between South African and Caribbean populations of *C. odorata*. Samples within localities or countries were collected to be representative of the morphological variation found within regions. Samples from within and outside the genus *Chromolaena* (but still within the tribe Eupatorieae) were obtained as outgroups.

Samples Venezue2, Brazil_5, Jamaica8, Mexico 10, USAFl_12 and SA_Dbn13 (Table 3.1) were used to test for phylogenetically informative genes at the intraspecific level, and to optimise the PCR parameters.

DNA Isolation, PCR Amplification, and Sequencing

DNA was extracted using the modified CTAB extraction protocol (Doyle and Doyle, 1987), with PVP and ethanol-grinding as described in Chapter 2.

PCR was initially performed with Promega's *Taq* polymerase. However, it was found that some samples produced insufficient PCR product for sequencing. These samples were found to amplify more successfully with Promega's proof-reading *Pfu* polymerase.

PCR Reactions were performed in 100µL reactions, using 6µL DNA extract (concentrations unknown, see Chapter 2), 10X PCR buffer (Promega), 1 to 5µM $MgCl_2$ (not with *Pfu* enzyme), 0.25mM each dNTP, 2 units *Taq* or *Pfu* polymerase, and 0.2µM of each primer. Primers used for ITS amplification were "ITS-1" (5'-TCCGTAGGTGAACCTGCGG -3') and "ITS-4" (5'-TCCTCCGCTTATTGATATGC-3';

White *et al.*, 1990). Primers used for ETS amplification were "18S-ETS" (5'-ACTTACACATGCATGGCTTAATCT-3') and "ETS-Hel-1 (5'-GCTCTTTGCTTGCGCAACAACT3'; Baldwin and Markos, 1998).

Table 3.1 Sample list for *Chromolaena odorata* and outgroups.

Sample name	Sample No.	Country of Origin	Locality	Characteristics of sampled C odorata plant.	Accession code *
Aadenoph	**141**	**South Africa (invasive)**	**Cedara culture**	***Ageratina adenophora***	**-**
Ariparia	**139**	**South Africa (invasive)**	**Cedara culture**	***Ageratina riparia***	**-**
Cborinqu	**120**	**Puerto Rico**	**Between Arecibo and Utado**	***Chromolaena borinquensis***	**Ced. WP 071**
Ccollina	**86**	**Mexico**	**Senora**	***Chromolaena collina***	**Univ. Texas, Austin**
Cmacroce	**140**	**South Africa (invasive)**	**Cedara culture**	***Campuloclinium macrocephalum***	
Csqualid	**80**	**Australia (invasive)**	**Caravan Hill, Tulley**	***Chromolaena squalida***	**BMW 6085**
Pclemati	**83**	**Australia (invasive)**	**Centre for Tropical Agriculture, Mareeba**	***Praxelis clematidea***	**BMW 6088**
Austra81	81	Australia (invasive)	Potted plant, Cardwell Shire Nursery, Tulley	Type 2, duplicate	BMW 6086
Austra84	84	Australia (invasive)	Indian Ocean territory of Cocos Island	Asian/Florida form	
Brazil_5	5	Brazil	Salvador	Similar to Florida form, base of leaf a bit rounded	Ced. Int. Coll. V4
Brazil51	51	Brazil	North of Nitero, Rio de Janero	*C. maximiliani*?	Ced. Int. Coll. I2
CostRic6	6	Costa Rica	?	Leaves & stems smoother than Florida form, more reddish; leaves broad (may not be *C. odorata*?)	Ced. Int. Coll. AH2
Guatem52	52	Guatemala	?	C. odorata or C. maximiliani?	Ced. Int. Coll. AC4
Guatema7	7	Guatemala	?	Harder leaves, darker stems than Florida form (may not be *C. odorata*?)	Ced. Int. Coll. AB1
India_16	16	India (invasive)	Bangalore	Similar to Florida form	Ced. Int. Coll. U2
Jamaic54	54	Jamaica	Runaway/Discovery Bay	South African form	Ced. Int. Coll. AG1
Jamaic56	56	Jamaica	Gordon's Town, St Andrew	Similar to Florida form	CZ 99, wp 032, P1
Jamaic57	57	Jamaica	Gordon's Town, St Andrew	Similar to South African form	CZ 99, wp 032, P2
Jamaic58	58	Jamaica	Gordon's Town to Guava Ridge, St Andrew	Similar to Florida form	CZ 99, wp 033, P1
Jamaic59	59	Jamaica	Gordon's Town to Guava Ridge, St Andrew	Similar to South African form	CZ 99, wp 033, P2
Jamaic62	62	Jamaica	Gordon's Town to Guava Ridge, St Andrew	Similar to South Africa form	CZ 99, wp 033, P5
Jamaic63	63	Jamaica	Guava Ridge to Mavis Bank, St Andrew	spiky	CZ 99, wp 034, P1
Jamaic69	69	Jamaica	West of Stuart Town, Trelawny	Spiky South African form	CZ 99, wp 043, P1
Jamaic71	71	Jamaica	West of Stuart Town, Trelawny	Florida form	CZ 99, wp 043, P3
Jamaica8	8	Jamaica	Irish Town	Identical to South African form	Ced. Int. Coll. AE2
Jamaica9	9	Jamaica	Guava Ridge	Similar to Florida form	Ced. Int. Coll. C1
Maurit15	15	Mauritius (invasive)	?	Similar to Florida form	Ced. Int. Coll. W4
Mexico10	10	Mexico	Nr Actopan	Similar to South African form, but hairier, with darker stems and leaves	Ced. Int. Coll. X2
NSumat38	38	Sumatra (invasive)	Marihat	Similar to Florida form	-
Prico 102	102	Puerto Rico	Between Arecibo and Utado	shiny spiky SA form	CZ 01, WP 069 P2
PRico100	100	Puerto Rico	Outside Isabella	Identical to South African form	CZ 01, WP 076 P1

SA_Dbn13	13	South Africa (invasive)	Durban area	South African form	-
SA_Esh26	26	South Africa (invasive)	Eshowe	South African form	IvS4
SA_Hlu42	42	South Africa (invasive)	Hluhluwe, Hilltop to Gate 3	South African form	-
SA_Hlu44	44	South Africa (invasive)	Hluhluwe, Hilltop to Gate 5	South African form	-
SA_Mtu27	27	South Africa (invasive)	Mtunzini	South African form	IvS5
SA_Pal88	88	South Africa (invasive)	Phalaborwa - Hans Merenski Country Club	South African form	P1
SA_Pal92	92	South Africa (invasive)	Phalaborwa - Phalaborwa Mining Company	South African form	PMC1
SA_PEd23	23	South Africa (invasive)	Port Edward	South African form	IvS1
SA_PSh28	28	South Africa (invasive)	Umtentweni - Old St Faiths road	South African form	IvS6
SA_PSh32	32	South Africa (invasive)	Umtentweni - Link Road, Spar end	South African form	IvS9
SA_PSh36	36	South Africa (invasive)	Umtentweni - Stapleton Road exit	South African form	IvS13
SA_PSJ45	45	South Africa (invasive)	Port St John's, Military Base 1	South African form	-
SA_PSJ48	48	South Africa (invasive)	Port St John's, Silaka Nature Reserve, W. of 3rd beach	South African form	-
SA_PSJ50	50	South Africa (invasive)	Port St John's, Southern-most limit - Prof. Lubke	South African form	-
SA_Tza18	18	South Africa (invasive)	Tzaneen	South African form	-
SA_Tza95	95	South Africa (invasive)	Tzaneen - Mamathola forestry station	South African form	Tz1
SA_Tza96	96	South Africa (invasive)	Tzaneen - Tzaneen-Lydenberg rd	South African form	Tz2
SA_Tza97	97	South Africa (invasive)	Tzaneen - Entrance rd to Tzaneen dam	South African form	Tz3
SA_Tza99	99	South Africa (invasive)	Tzaneen - Mike Amm's farm	South African form	Tz5
Thaila11	29	Thailand (invasive)	?	Similar to Florida form	Stell. 018/012
Trinid_11	11	Trinidad	Blanchisuisse	Similar to Florida form	Ced. Int. Coll. P1
Trinid55	`	Trinidad	Waller	narrow leaves, maybe C. odorata or C. ivifolium?	Ced. Int. Coll. *Q3
USA_Fl12	12	USA	Rolling Hills, Florida	Florida form	Ced. Int. Coll. L1
Venezue1	1	Venezuela	Nr Trujillo	Typical Venezuelan form (similar to Florida form)	CZ 98, wp 096, P1
Venezue2	2	Venezuela	Nr Puerto la Cruz	Similar to above	CZ 98, wp 117, P1
Venezue3	3	Venezuela	Nr Trujillo	Typical Venezuelan form (similar to Florida form)	CZ 99, wp 011, P1
Venezue4	4	Venezuela	Nr Trujillo	Leaves redder, with deep serrations	CZ 99, wp 011, P3
WAfric17	17	West Africa (invasive)	?	Similar to Florida form	Ced. Int. Coll. L3

* Voucher Specimen Key: **BMW** – Dr. B. M. Waterhouse, AQUIS, Australia; Ced. – ARC-PPRI at Cedara, South Africa. **IvS** - Author (specimens kept at Schonland Herbarium, Rhodes University, Grahamstown, South Africa). **CZ** – Dr. Costas Zachariades, ARC-PPRI, Cedara, South Africa.

The thermal cycling profile used for the ITS was: 40 cycles of 95°C for 45s, 52°C for 45s and 72°C for 150s, followed by 10min at 72°C. A different thermal cycling program was used with the *Pfu* enzyme: 40 cycles of 95°C for 45s, 52°C for 45s, and 75°C for 180s, followed by 10 minutes at 75°C.

The thermal cycling profile for the ETS was: 10 cycles of 95°C for 30s, 57°C for 30s and 72°C for 120s, followed by 10 cycles of 95°C for 45s, 55°C for 45s and 72°C for 150s, followed by 15 cycles of 95°C for 60s, 55°C for 60s and 72°C for 180s. All of this was followed by 10min at 72°C.

PCR products were visualised by agarose gel electrophoresis and purified using the QIAgen QIAquick PCR Purification Kit.

The sequencing reaction was carried out using an ABI PRISM BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (ver. 2.0, later 3.0), with AmpliTaq® DNA Polymerase, FS, and with the ABI PRISM 5X Sequencing Dilution Buffer as per manufacturers' instructions. Final sequencing reaction volumes were 20µL.

Sequencing of the ITS was done in both directions using the flanking primers used for PCR amplification, and internal primers "danth 5.8F" (5' – GACTCTCGGCAACGG-3') and "chromo 5.8R" (5'-GATTCTGCAATTCAC-3', both designed by N. Barker, Rhodes University). The ETS was sequenced with the primers used for PCR amplification.

The sequencing product was precipitated according to Ethanol Precipitation Protocol 1 of the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit. Sequences were analysed either on an ABI 377 Automated DNA Sequencer (Department of Genetics, University of Stellenbosch, South Africa) or on an ABI 3100 Genetic Analyzer (DNA Sequencing Facility, Rhodes University, South Africa).

Sequence Alignment and Phylogenetic Reconstructions

Sequences were checked and assembled using Sequencher™ ver. 3.1.1, and were then aligned manually using DAPSA (DNA And Protein Sequence Alignment; written by E. H. Harley, Dept. Chemical Pathology, University of Cape Town Medical School, Observatory, 7935, South Africa). The sequence alignments are provided in Appendix 6. Phylogenetic relationships were analysed using three different software packages: PAUP* (Phylogenetic Analysis Using Parsimony) 4.0b2a (Swofford 1999), TCS (Clement *et al.*, 2000), and Arlequin ver. 2.000 (Schneider *et al.*, 2000).

Choice of Analytical Methods

Each of the standard tree-making algorithms in common usage has its own particular criteria, and each algorithm makes assumptions about the nature of the evolutionary process affecting a species' development; assumptions that that can only be poorly checked (Smouse, 1998). In practice, the best possible tree is constructed according to the data and whatever criteria are explicit (or implicit) in the chosen tree-making algorithm (Smouse, 1998). Conversion of data into a tree represents an extrapolation from the data; alternative trees are alternative extrapolations, and analysis of any tree is an analysis of the corresponding extrapolation (Smouse, 1998). To the extent that a tree is an extrapolation beyond the data, analysis of that tree is unavoidably analysis of any (erroneous) assumptions (Smouse, 1998). A possible method of depicting trees with minimal homoplasy is by using reticulating trees. However, lineage sorting, which is the primary stochastic source of incongruence, is not reticulation *per se* (Doyle, 1997). Minimum spanning trees (MSTs) or minimum spanning networks (MSNs) have OTUs (haplotypes) as both nodes and branch tips, as opposed to Steiner trees (which are usually used) that have OTUs only as branch tips, with the nodes representing OTUs that are now extinct and ancestral (Smouse, 1998). The MST which least increases the variation among haplotypes is that tree which is thought to represent minimal extrapolation from the raw data (Smouse, 1998). For interspecific work, the ancestral intermediates are commonly absent from the data set, and Steiner trees provide adequate information. For intraspecific work, however, the intermediates are often still present, and the 'trees' resemble become web-like (Smouse, 1998; Schaal and Olsen, 2000).

Four basic analytical methods or protocols were employed in this study, and were chosen to display a range of possible algorithmic methods and interpretations of the data at hand. They are (with the associated software):

Analysis 1: Maximum Parsimony (PAUP)

Analysis 2: Neighbour Joining analysis of a Distance matrix, with Jukes Cantor correction (PAUP)

Analysis 3: Statistical Parsimony (TCS)

Analysis 4: Minimum Spanning Tree with Pairwise Difference (Arlequin)

The Maximum parsimony analysis is a hierarchical method (that produces Steiner trees), and is useful at interspecific level and above, but not necessarily at intraspecific levels. Maximum parsimony will not be able to resolve reticulate relationships that are commonly found within

a population or species. In contrast, the TCS analytical parsimony utilises statistical parsimony to produce MSNs, and is t hus designed to resolve reticulate, intraspecific relationships.

The Jukes Cantor Neighbour Joining distance tree is expected to better resolve relationships between intraspecific taxa than maximum parsimony because it reconstructs relationships based on distances, rather than haplotypes. The Arlequin program is able to produce MSTs based on pairwise difference matrices. These MSTs are specific to reticulate relationships found within a species. Thus it is expected that the Arlequin-based phylogenetic reconstructions will be more informative about the intraspecific relationships within *C. odorata* than the Neighbour Joining distance reconstruction.

With regards to how one would analyse the resulting trees from the different analyses, Excoffier and Smouse (1994) are of the opinion that one should not view any one solution with a large degree of comfort. Rather they advocate the need for molecular systematists to begin to think about classes of acceptable solutions that allow one to bracket estimations and ignorance.

It may be misleading to view some gene trees as 'agreeing' and other gene trees as 'disagreeing' with a species tree; rather, if one viewed all of the possible gene trees as part of the species tree, which can be visualised as a fuzzy statistical distribution, or a cloud of gene histories (Maddison, 1997). Conventional tree-like phylogenies can be regarded as simplifications of reality (similar to maps); *i.e.*, generalised representations of the information with events selectively deleted according to the level and nature of the detail required (Avise and Wollenberg, 1997). Phylogeny has a variance that is represented by the diversity of trees of different genes, and thus a simple phylogenetic tree diagram with stick-like branches represents only the mean, or mode, of a distribution (Maddison, 1997). This variance does not represent uncertainty due to ignorance or measurement error; it is an intrinsic part of phylogeny's nature.

Maddison (1997) uses an analogy from physics to describe the variance present in a phylogeny: electrons in an atom are said to be diffuse, not because of any uncertainty about the position of the electron, but because it is in 'more than one place at a time'. In a similar manner, phylogenetic history is in more than one place at once; it is a composite of all the varied histories of all the genes, some of which might place species A next to B, others might place A next to C, *etc.*, and thus, just as an electron can be depicted as a cloud, phylogeny can be viewed as a diffuse cloud of gene histories. This cloud-phylogeny still has some form, and in many cases it will have a central tendency that will take the form of a tree (Maddison, 1997), allowing inferences to be made about the phylogeny.

Perhaps a method that uses the concepts proposed by Maddison (1997) and Avise and Wollenberg (1997) would reduce the need for multiple analyses with different assumptions and algorithms.

Maximum parsimony

The maximum parsimony approach was used with the PAUP default settings, except that Multistate Characters were considered as "polymorphisms". Two hundred Random Entry analyses were conducted, followed by a Heuristic search of all the shortest trees retained in the memory. Maxtrees was set to 10,000. A strict consensus tree was calculated from all of the most parsimonius trees. Bootstrap values were obtained from 1000 fast bootstrap replicates. .

The non *C. odorata* taxa (outgroups) were used according to availability from the ARC-PPRI, and from the University of Texas, Section on Integrative Biology (Panero, pers. comm.), and are indicated in bold text in Table 3.1. The *Ageratina* species were used as outgroups, and this allowed the remaining non-*C. odorata* taxa to be placed within the cladogram relative to *C. odorata*.

Neighbour Joining Distance Analysis with Jukes Cantor Correction

A Neighbour Joining tree was constructed from a pairwise distance matrix (Appendix 7) obtained using the Jukes Cantor correction. Any gaps were ignored for each pairwise comparison. Bootstrap values were obtained from 1 000 neighbour joining replicates.

Statistical Parsimony

TCS is a computer program that implements the estimation of gene genealogies from DNA sequences as described by Templeton *et al*. (1992). The program collapses sequences into parsimony informative haplotypes and calculates the frequencies of the haplotypes in the sample. The cladogram estimation method utilised in this analysis is also known as statistical parsimony (Clement *et al*., 2000).

In this analysis, gaps were set as fifth character states (the resolution of the MSTs collapsed completely when only four character states are recognised), and nodes and branches were selected.

Distance (defined according to Arlequin)

Arlequin (Schneider *et al.*, 2000) is software that calculates many indices and variables that have been described by various authors. In this analysis the Minimum Spanning Tree option was used, with Pairwise Difference molecular distances.

Results and Discussion

Comparison of ETS and ITS Phylogeny Reconstructions

A five-taxon comparison of ETS and ITS data and resultant phylogenies is given below. Maximum parsimony analysis produced only a single tree of length 28 and 18 steps for the ETS and ITS trees respecively. Figure 3.2 shows the maximum parsimony trees, with bootstrap values >50% indicated. Table 3.2 shows the distribution of variable and parsimony uninformative characters for parsimony use.

Table 3.2 Character distribution for parsimony analysis.

Character Information	ETS	% of ETS Total	ITS	% of ITS total
Total characters	647*	100%	765	100%
Constant characters	624	96.45%	747	97.65%
Variable characters	23	3.56%	18	2.35%
Variable, parsimony-informative characters	6	0.93%	2	0.26%

* Partial ETS sequence data. (Complete ETS sequence not available.)

These data show that there is more variation within the ETS than ITS sequence, even though the portion of ETS region sequenced is shorter than the ITS sequence. Figure 3.3 shows the Jukes-Cantor Neighbour-Joining distance tree, with bootstrap values >50% indicated. The TCS MSTs for ETS and ITS gave no information on the respective variability and usefulness of the ETS and ITS regions and is therefore not included. The distance-based MSTs produced by Arlequin ver 2.000 are identical for the ITS and ETS regions. Table 3.3 gives the statistical values for the data sets used.

Table 3.3 Statistical values obtained from Arlequin ver. 2.000.

Statistical Values	ETS	ITS
No. of loci	647	765
No. of useful loci (with <5% missing)	598 (92.42% of total)	660 (86.27% of total)
No. of polymorphic sites	19 (2.94% of total)	6 (0.78% of total)
Observed transitions	8 (1.24% of total)	3 (0.39% of total)
Observed transversions	11 (1.70% of total)	3 (0.39% of total)

Discussion of Relative Usefulness of ETS and ITS in Intraspecific Studies

The character distributions for the ETS and ITS data sets for the Parsimony analysis indicate that more variable loci were present in the ETS region than the ITS region, as 0.93% of the characters in the ETS data set were parsimony informative, while a lesser 0.26% of ITS characters were parsimony informative (Table 3.2). The statistical values from the Arlequin analysis (Table 3.3) confirm that the ETS is more variable than the ITS: 2.94% of the ETS sequence had informative polymorphic sites, while 0.78% of the ITS sequence were polymorphic.

Further support for the stronger phylogenetic signal produced by the ETS region as opposed to the ITS region can be found in Figure 3.3, the Jukes Cantor Neighbour Joining distance reconstruction, where the ETS region provides enough signal for two strongly supported clades, with bootstrap values of 88% and 77%, yet there appears to be insufficient signal in the ITS region to provide bootstrap support values greater than 50%.

One can therefore say that the character distribution and informativeness of these two regions indicates that the ETS region is more variable than the ITS, and would therefore be more suitable for an intraspecific study than the ITS region. This supports findings by Linder *et al*. (2000) who found that the ETS was consistently more phylogenetically informative than the ITS.

Although the ETS region provided a high degree of informativeness within the species *C. odorata*, difficulties were encountered during PCR with primer mismatch, due to the variable nature of the region. The difficulties could not be easily overcome, and it was therefore decided that the ETS region would not be sequenced for the entire data set, and that emphasis would be placed on the ITS region, which was working consistently.

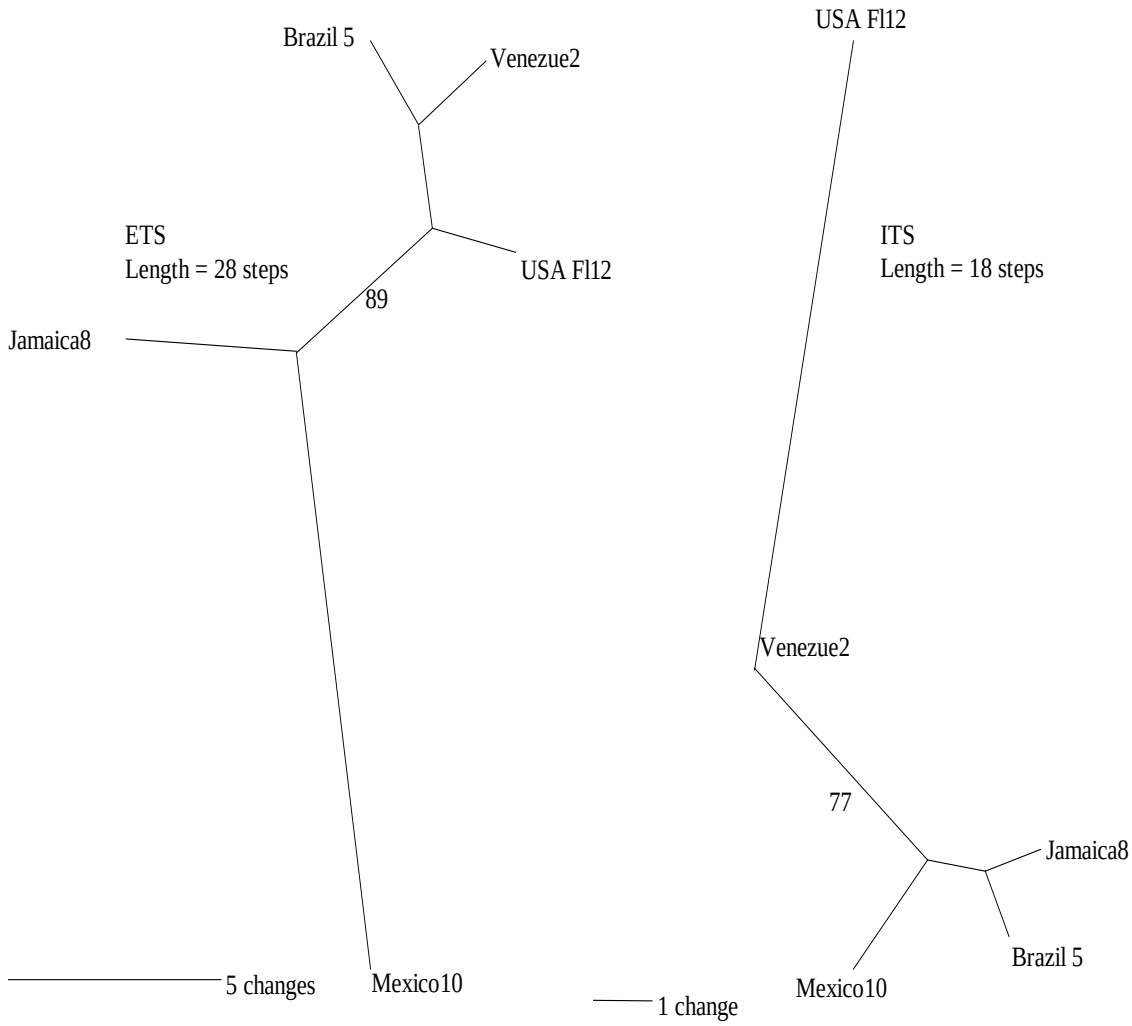

Figure 3.2 Unrooted ETS and ITS maximum parsimony trees with bootstrap values.

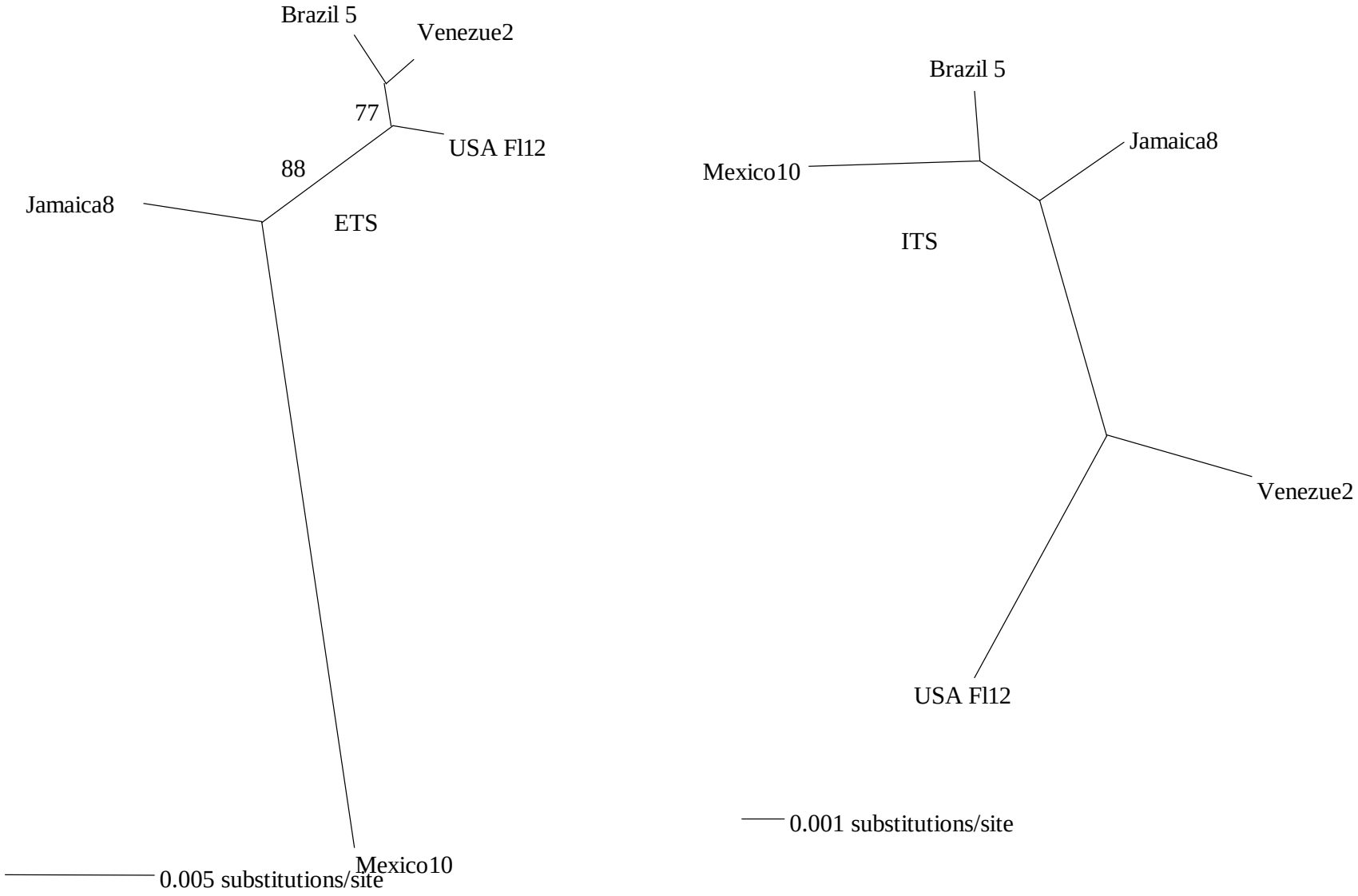

Figure 3.3 Unrooted Neighbor Joining Jukes Cantor distance trees, with bootstrap values > 50% (ITS did not have bootstrap values >50%).

Phylogenetic Analysis of ITS Sequence Data

The following section begins by describing the topologies of the ITS genealogies. It then draws some conclusions on the relative informativeness of morphology and ITS nuclear data based on observations from the non-South African samples. Thereafter, these observations are extended to make inferences on the structure of the South African population. The trends found in the South African population are then discussed, and attempts are made to understand them in terms of concerted evolution and polyploidy.

In total, 61 ITS sequences were obtained; 52 ingroup taxa, seven outgroup taxa, and two sequences of the ITS-1 region from Australian samples of *C. odorata* (Scott *et al*., 1998). The results for the complete 61-taxon ITS data set are presented below. Not all of the samples are complete. In these samples, the missing nucleotides in the sequences were indicated by question marks (Appendix 6). Analyses were done with and without the samples that had more than 5% missing nucleotides, and it was found that the topology of the phylogenetic reconstructions did not change significantly when the taxa with missing data were excluded. Therefore, the trees with the full data set are included and discussed below. Samples with more than 5% missing nucleotide data are: AusITS1A and AusITS1B (missing 64.6%), SA_Hlu44 (missing 32.0%), SA_Tza99 (missing 25.2%), Pclemati (missing 23.4%), SA_Dbn13 (missing 18.8%), SA_Tza18 (missing 8.7%), Venezue2 (missing 7.5%), Mexico10 (missing 6.3%), and Venezue1 (missing 5.6%).

A colour code for morphotypes of *C. odorata* (see Chapter 1) has been used in this chapter. Due to software limitations, the colour codes could not be kept consistent, but keys are included in all relevant captions.

Maximum Parsimony Analysis

Figure 3.4 shows the strict consensus tree of two most parsimonious trees of length 673 and bootstrap values greater than 50% are included. This figure shows a surprising amount of resolution for a hierarchical analysis at the intraspecific level, yet it contains few bootstrap values greater than 50%. The outgroups used in this and all other analyses were *Ageratina adenophora* and *A. riparia.*

In this hierarchical analysis, the positions of the non-*C. odorata* taxa are well-resolved and strongly supported with high bootstrap values (from 96 to 100%). The samples Brazil51, Guatem52, Guatem57, and Trinid55 (Table 3.1) are clearly placed within the *C. odorata* clade, indicating that they are of the species *C. odorata*, despite misgivings noted in Table 3.1.

It was thought that a higher level of resolution in the Maximum Parsimony tree would be achieved by using a species more similar to *C. odorata* than *Ageratina adenophora*, such as *Chromolaena squalida* (McFadyen, pers. comm.) Therefore, a maximum parsimony analysis with *C. squalida* as the only outgroup was done, however the resolution of this tree was very poor. There were several 2-taxon clades, but no further resolution. Because of the lack of resolution, the tree is not included in this dissertation. There were only 19 parsimony informative characters in this tree; possibly the reason for the poor resolution of the tree.

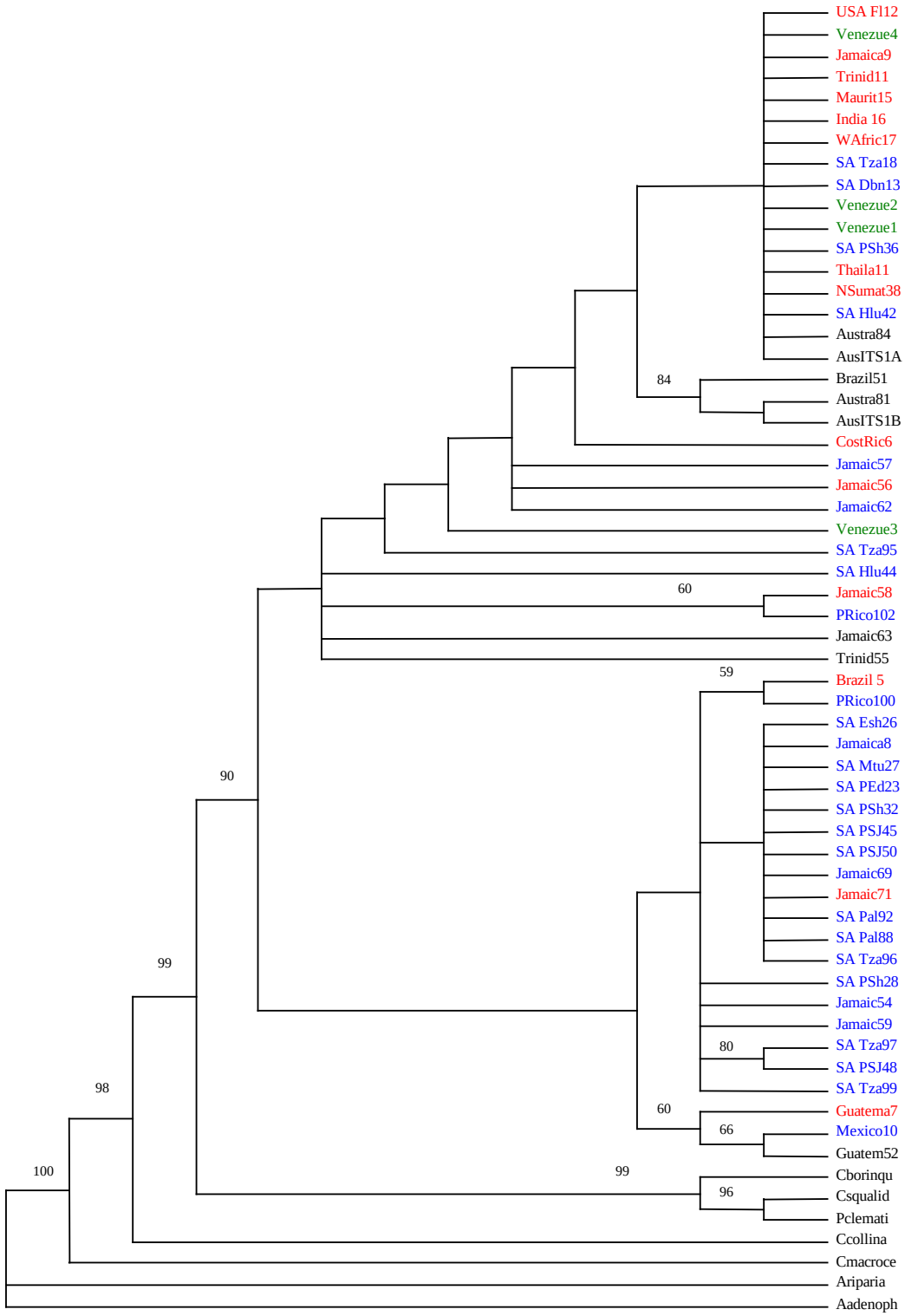

Figure 3.4 Strict consensus parsimony tree for two most parsimonious trees of 673 steps. Consistency Index (CI) = 0.932; Retention Index (RI) = 0.881 Colours indicate morphotypes: Blue = South African, Red = Floridean, Green = Venezuelan, Black = other/unknown.

Statistical Parsimony Analysis

Figure 3.5 is a TCS-constructed MST. Each box or circle contains one or more samples of a single haplotype according to statistical parsimony (Clement *et al.*, 2000). Haplotypes that are considered probable outgroups are contained in boxes rather than circles, and the size of the box or circle is indicative of the number of samples in that haplotype. Small, open circles represent missing haplotypes. An analysis of this data without the outgoup taxa does not increase the resolution of the ingroup haplotypes, and is therefore not presented here. The analysis with only the ingroup taxa shows a slight rearrangement of the haplotypes and samples; however, the trends and associations are essentially the same as in Figure 3.5

Neighbour Joining Distance Analysis

Figure 3.6 is a neighbour-joining Jukes-Cantor distance tree. The numbers on the tree are neighbour-joining bootstrap values greater than 50%. Figure 3.7 is the same as Figure 3.6, but with only a single outgroup - the one closest to the ingroup in Figure 3.6 so that intraspecific relationships can be more clearly discerned. Figure 3.7 shows more clearly the intraspecific relationships of *C. odorata*.

Pairwise Differences Analysis

Figure 3.8 is an MST created by Arlequin ver. 2.000. Arlequin found 61 original haplotypes from 61 samples, so each sequence for each sample is unique: this magnitude of sequence divergence was not expected. Of 765 loci, 474 had less than 5% missing data through the entire data set, and there were 157 polymorphic sites. There were 101 observed transitions and 90 observed transversions. The relative nucleotide composition is shown in Table 3.4:

Table 3.4 Relative nucleotide composition of ITS sequences for 61 samples, as estimated by Arlequin ver. 2.000.

Nucleotide	Relative Percentage
C	25.30
T	24.97
A	22.07
G	27.67

Figure 3.9 is the same as Figure 3.8, but with the outgroups removed so that the intraspecific relationships are more clearly discerned. Figure 3.10 is a rooted distance phylogram of the Arlequin MST (rooted on sample Venezue3, as it was the sample closest to the outgroup taxa in Figure 3.8), to show the relationships of the taxa that are clustered very close together in the centre of Figures 3.8 and 3.9.

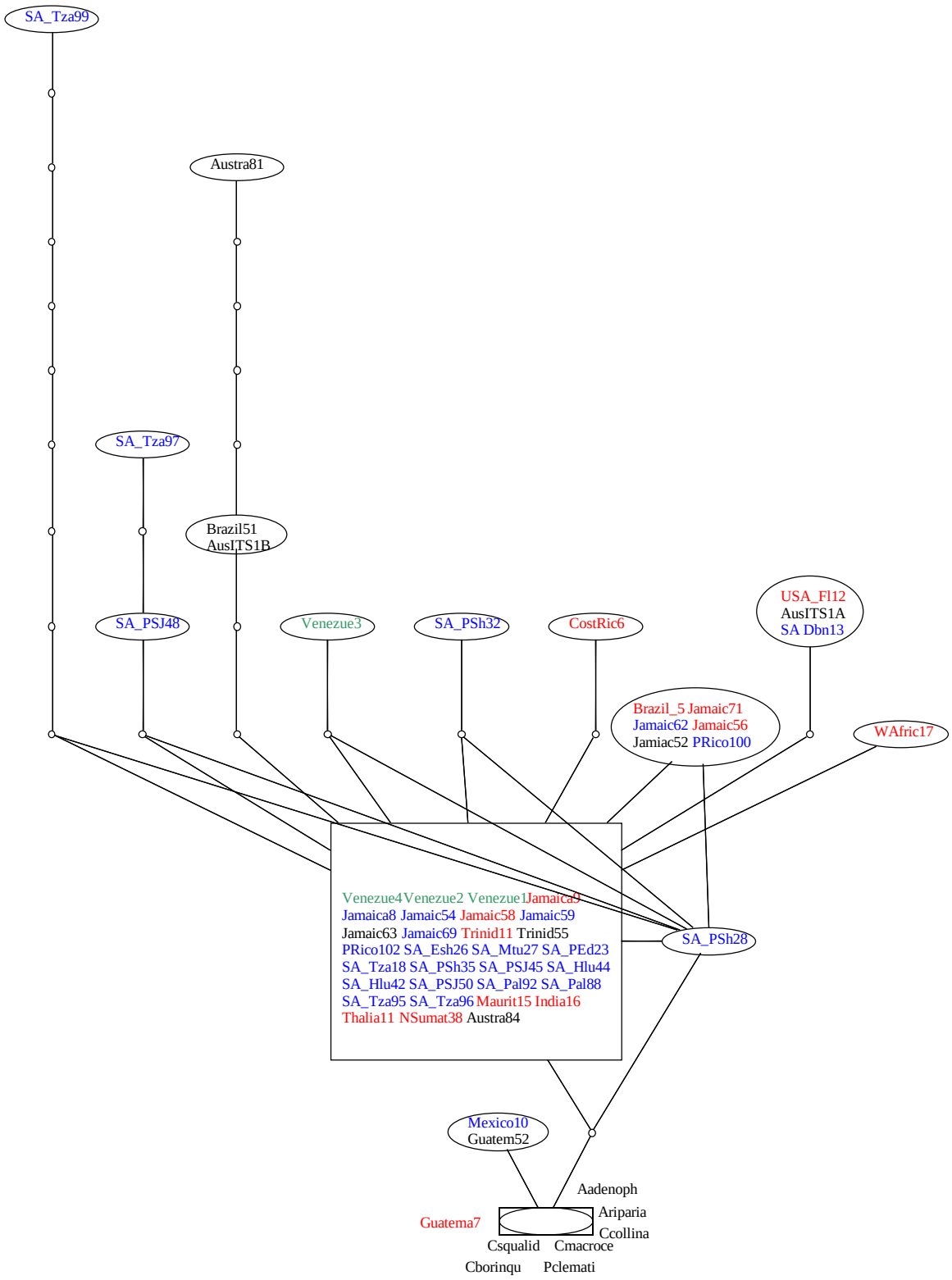

Figure 3.5 TCS Minimum spanning network for complete ITS data set. Empty circles indicate missing haplotypes, and rectangles indicate probable outgroups. Colours indicate morphotypes: Blue = South African, Red = Floridean, Green = Venezuelan, Black = other/unknown.

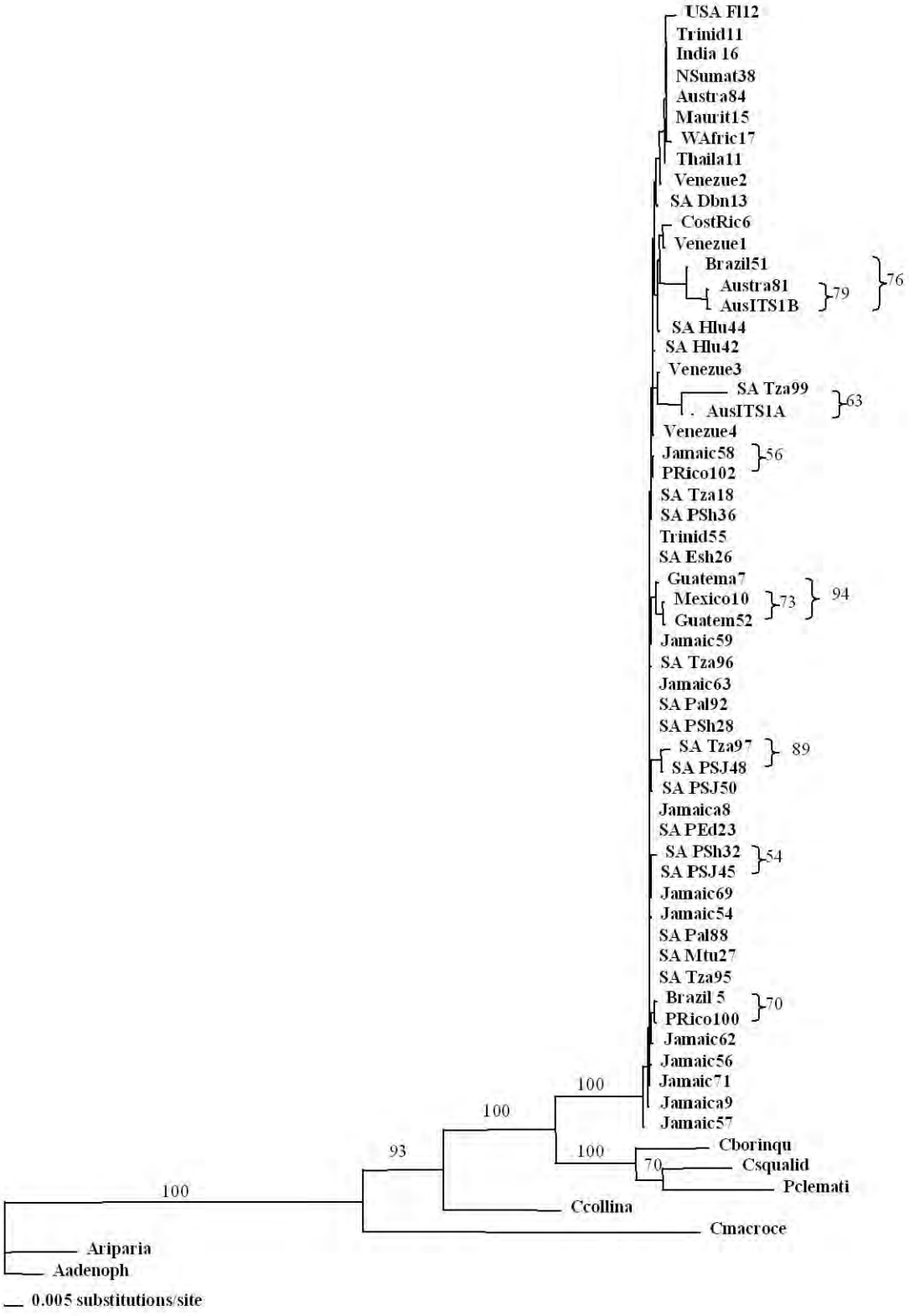

Figure 3.6 Rooted Jukes Cantor Neighbour Joining distance tree. Numbers behind brackets indicate bootstrap values for bracketed taxa.

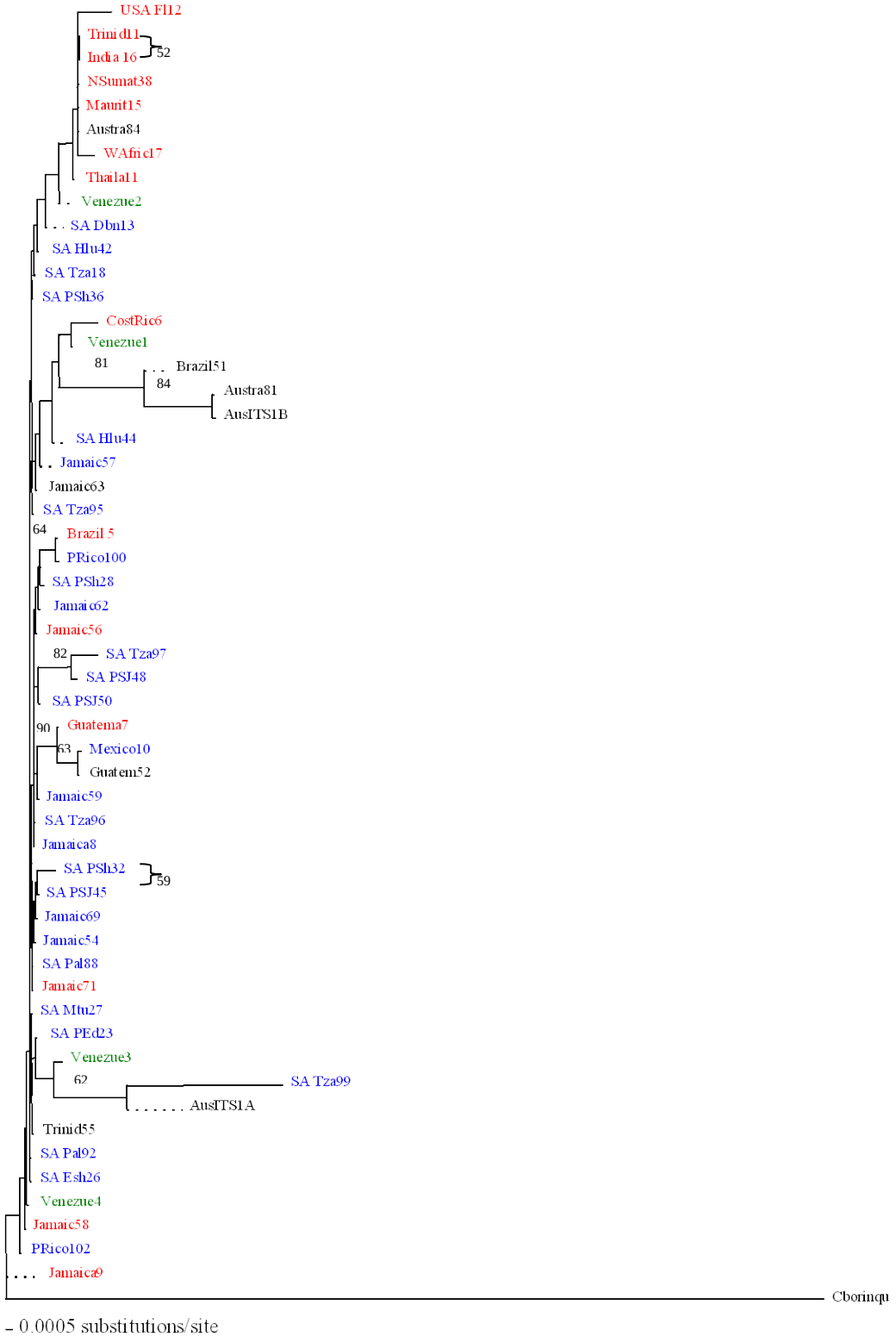

Figure 3.7 The same Jukes Cantor Neighbour Joining distance trees as Figure 3.6, but showing only the outgroup closest to the ingroup, to allow better resolution of ingroup relationships. Colours indicate morphotypes: Blue = South African, Red = Floridean, Green = Venezuelan, Black = other/unknown.

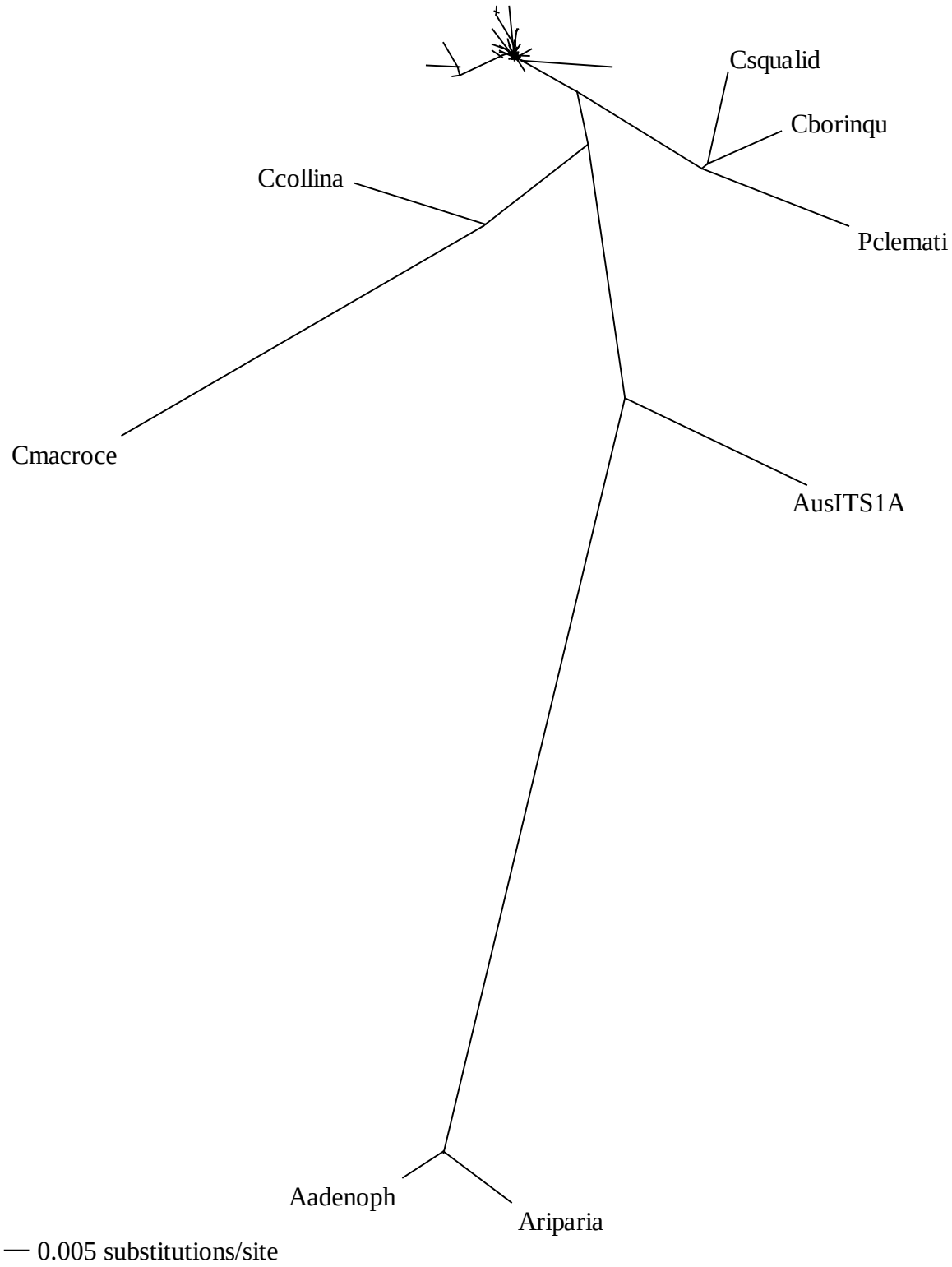

Figure 3.8 Arlequin Minimum Spanning Tree, showing relative positions of outgroup taxa.

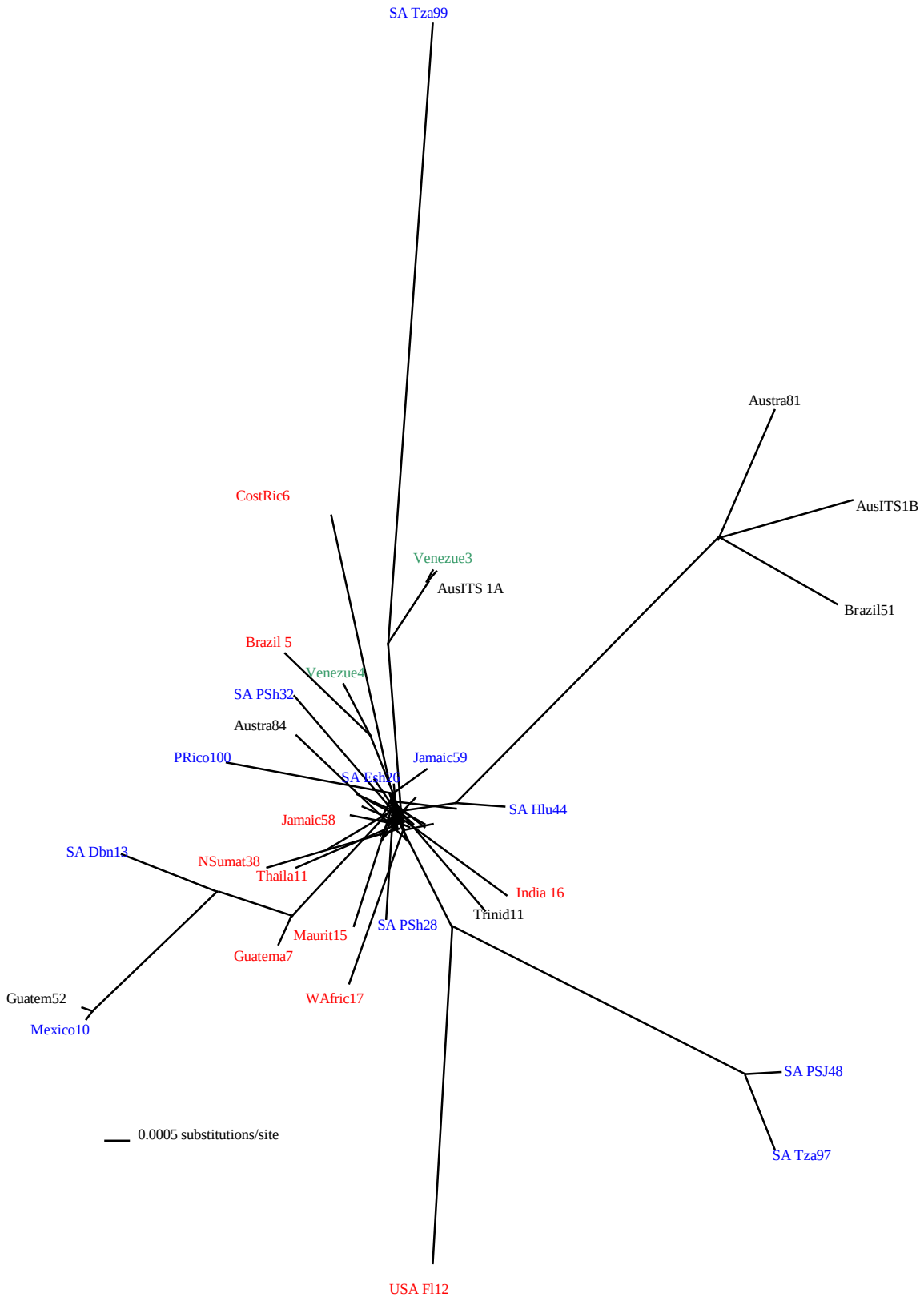

Figure 3.9 The same MST from Arlequin as in Figure 3.8, but with outgroups removed, for better resolution of ingroup taxa. Taxa in the centre of the MST are excluded. See Figure 3.10 for detail of relationships between these central taxa. Colours indicate morphotypes: Blue = South African, Red = Floridean, Green = Venezuelan, Black = other/unknown.

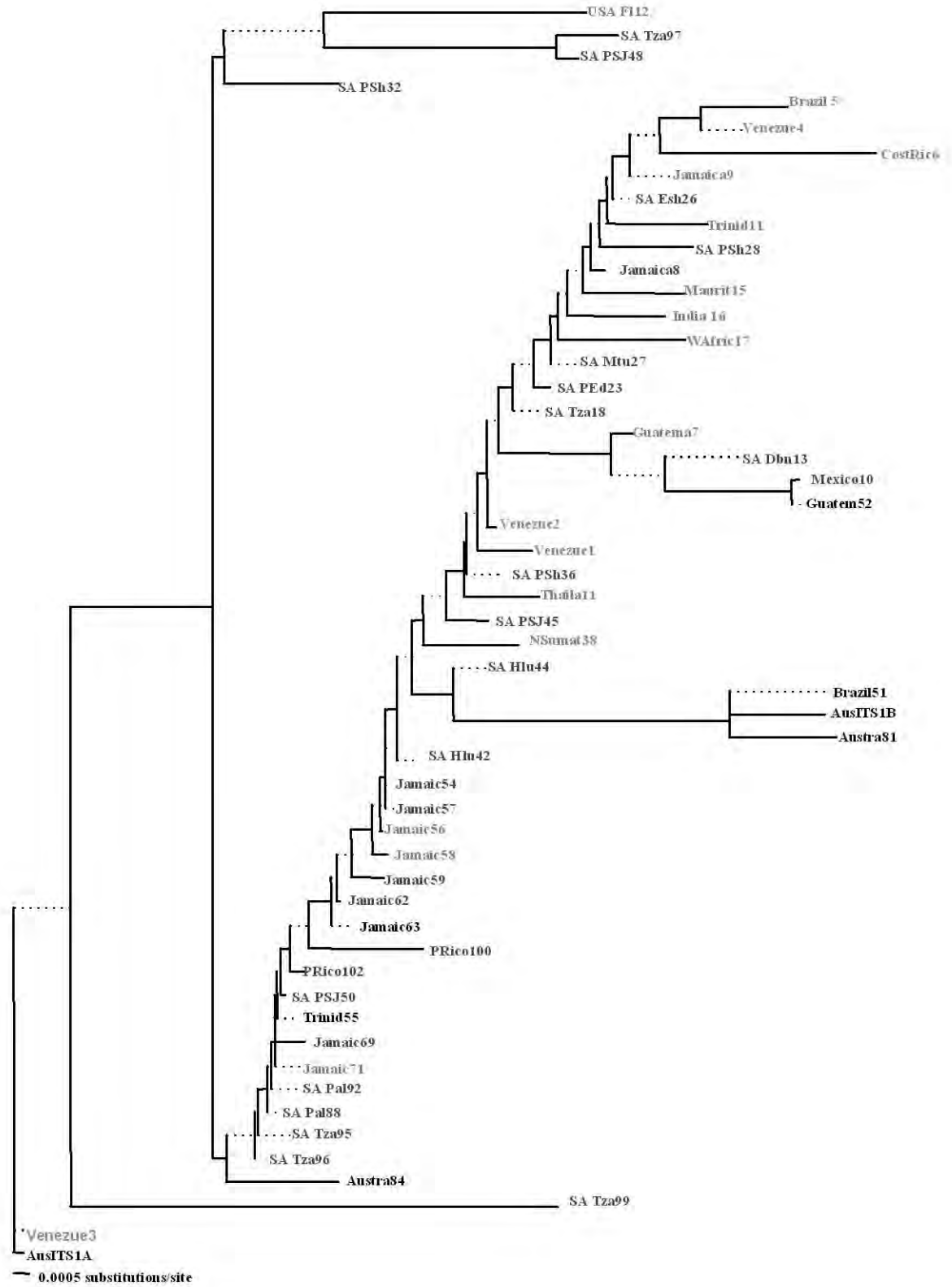

Figure 3.10 Arlequin MST arranged as a rooted phylogram (on sample Venezue3, as it was the sample closest to the outgroup taxa in Figure 3.8), to show relationships of taxa closely arranged in Figures 3.8 and 3.9 Colours indicate morphotypes: Blue = South African, Red = Floridean, Green = Venezuelan, Black = other/unknown.

Phylogeography of *C. odorata*

Phylogeography has been hailed as the conceptual bridge linking population-level processes to macroevolutionary relationships (Olsen and Schaal, 1999). The following section draws together the morphological and genetic similarities of the samples of *C. odorata*, in a phylogeographic context (Figure 3.11)

Over all the trees obtained here, several groupings or associations have remained constant. Samples Brazil51 and Austra81 are grouped together in all four methods, as are SA_Tza97 and SA_PSJ48. Samples Guatem52 and Mexico10 are found together, in a position descendant from Guatema7 under all four analyses. Samples Brazil5 and PRico100 as well as Jamaica58 and PRico102 are grouped by distance, TCS and Maximum Parsimony. A large group comprising of samples (Trinid11, India16, Nsumat38, Maurit15, Austra84, Wafric17 and Thalia11) is found by Distance and TCS analyses. These common associations are presented in Figure 3.11.

Other than these groupings, the remaining samples are distributed without apparent pattern through the different trees or phylogenies. The South African samples are not clustered together, and nor are the multiple samples from other countries such as Jamaica and Venezuela, indicating that there is a high degree of variance within the ITS gene of *C. odorata.*

Chromolaena odorata can be roughly divided into three "morphotypes", here denoted as 'Venezuelan', 'Floridean' and 'South African' (Zachariades, pers. comm.). Most of the ingroup taxa have been identified as having one of these three morphologies, and are shaded in Figures 3.4, 3.5, 3.7, 3.9, and 3.10 according to their 'morphotype'. The correlation (or lack thereof) of the distribution of morphotypes of *C. odorata* and the ITS genealogy are discussed below.

Australia and Other Regions

It seems that all of the introduced populations of *C. odorata* to Australasia, West Africa and Mauritius (not including sample Austra81) are associated with (or have) the 'Florida' morphology, and have genetic similarities with each other and with sample Trinid11. This suggests that all Australasian, West African and Mauritian infestations originated this region (Trinidad). These samples can be correlated with the *C. odorata* ITS1-A genotype found by Scott *et al*. (1998), which is most commonly found in North Queensland and Asia.

Place phylogeog map here

International Biogeography of *Chromolaena odorata*

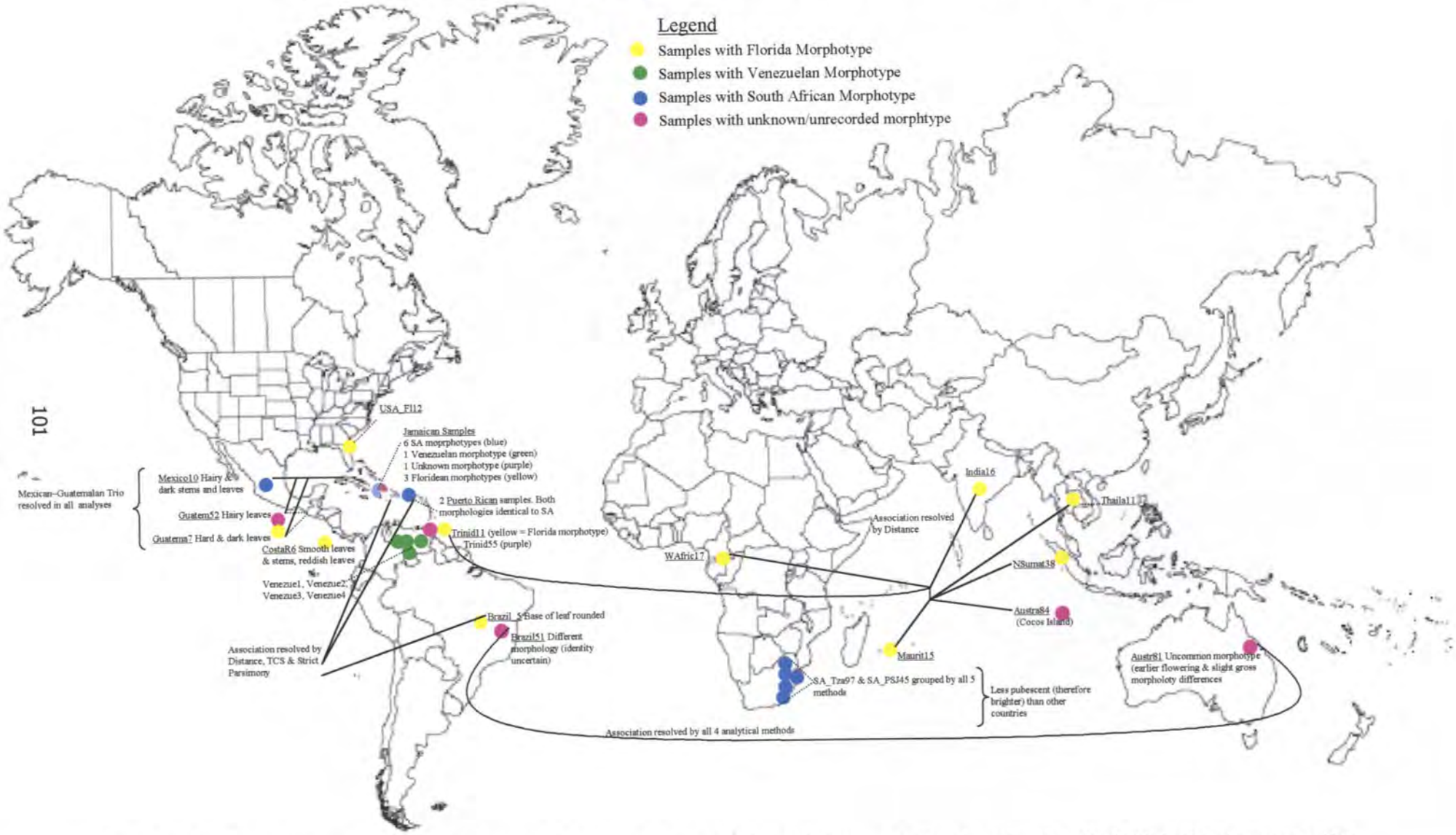

Figure 3.11 Phylogeography of *Chromolaena odorata*. Sample positions and morphotypes are show by coloured dots (with broken lines conneting where necessary). Genetic associations are indicated by solid lines. (SA = South Africa)

Scott *et al*. (1998) found that this ITS1-A genotype had a biogeographic origin in central America, which agrees with the findings of this dissertation. Samples Maurit15, India16, and WAfric17 are all thought to have come from a single origin [Calcutta Botanical Gardens (1800's), McFadyen, 1988], and this would give support for the genotypic signal from the ITS genealogy.

Austra81 is an uncommon morphotype in the Australian infestation. This Australian morphotype is noted as having an earlier flowering time, as well as having slight gross morphological differences to the most common morphotype in the Australian infestation (Waterhouse, pers. comm.). This morphotype can be correlated with the uncommon ITS1-B genotype found by Scott *et al*. (1998) (early flowering time). All four analytical methods in this dissertation have placed AusITS-B in a clade with Brazil, giving support for phylogeographic origin of this sample as being Brazil, in agreement with the results of Scott *et al.* (1998).

The separate phylogeographic origins of the two morphotypes in North Queensland, agreed upon by both this dissertation and the results of Scott *et al*. (1998), indicate that there were two separate introductions of *C. odorata*; the most common morphotype was most probably introduced from neighbouring Asian countries, and the second morphotype appears to be a direct introduction from Brazil.

Phylogeography of *C. odorata* in the Americas

The two Guatemalan and the Mexican samples (Guatema7, Guatem52 and Mexico10) all have similar morphological traits (that are not found in any other samples), which agree with the genetic signal: The Mexican sample has dark and hairy stems and leaves, the sample Guatema7 has dark and hard leaves, while the Guate52 sample has hairy leaves. Although the 'morphotypes' of these samples are not the same (Mexico10 = South African; Guatema7 = Florida; and Guate52 = unknown), the common dark and hairy stems and leaves as well as the ITS phylogeny group them together. This raises the question of whether the morphological traits that have been used to classify samples according to morphotype are uninformative, or that only some of the characters used are informative. This example of the Guatemalan and Mexican samples suggests that pubescence and the darkness of leaves and stems agree with the ITS genealogy, and may therefore be a good set of characters for subdivision of *C. odorata* populations or morphotypes, assuming that the ITS genealogy follows *C. odorata* phylogeny.

Studies involving genetic markers and morphological differences have found that there is often no correlation between phenotypic differences between populations and variation in

genetic markers (Heaton *et al.*, 1999). In a comparison of phylogenetic usefulness of morphology, ITS and *trnL*-F sequence data of *Doronicum*, Fernández *et al.* (2001) found that the morphological data set contained poor phylogenetic signal. They advocated that morphological characters need to be tested for homology by congruence with molecular data. Similarly, Aguilar *et al.* (1999a) found that the morphological characters for subspecies of *Armeria villosa* conflict strongly with the ITS genealogy, and that the genealogy showed greater congruence with the phylogeography of the plants than the morphology. It is appears that at least some of the morphological characters that have been used in determining morphotype in *C. odorata* are incongruent with the ITS data, and that other morphological characters (pubescence and darkness/lightness of leaves and stems) are more congruent with the ITS data. It is therefore suggested that the morphological traits of pubescence and darkness/lightness of leaves and stems are informative morphological characters for population studies of *C. odorata*.

A lack of congruence between gene trees and taxonomical categories has been explained by Olsen and Schaal (1999) to have potentially arisen either by gene flow among the taxa [*i.e.*, introgression - the gradual infiltration of the genetic material of one species into another as a consequence of hybridization and repeated backcrossing; Anderson (1949, cited in Aguilar, 1999)] or by the persistence and sorting of ancestral polymorphisms that predate the divergence of the taxonomic lineages (lineage sorting). Within *C. odorata* it is more possible that the morphological traits used are environmentally affected, and therefore not good for phylogenetic use. However, the possibility of introgression of populations of *C. odorata* with closely related species is not excluded. It is possible that such introgression is occurring in the Mexico-Guatemala region, and could explain the positioning of these samples with the outgroup taxa in the TCS minimum spanning network.

The leaves and stems of *C. odorata* in South Africa are less pubescent than those from other countries (Scott *et al.*, 1998), and therefore appear brighter. The leaves of sample CRica6 are also smooth. If pubescence and brightness/darkness are indeed informative characters (as suggested above) then it can be strongly suggested that the South African *C. odorata* did not originate from Guatemala and Mexico, but that there is a similarity between South African and Costa Rican populations of *C. odorata*. The ITS genealogies corroborate this suggestion; it is only in the Arlequin distance-based MST (Figures 3.9 and 3.10) that a South African sample is associated with the Mexican-Guatemalan trio, and it is a sample with 18.8% of the sequence data missing (SA_Dbn13), so its position in the ITS tree is not strongly supported. In none of the other genealogies is this trio associated in any way with the South African samples, and in the TCS MST (Figure 3.5) they are even associated with the outgroups. The parsimony and distance analyses (Figures 3.4 and 3.6) have very strong bootstrap support for the positioning of the outgroup species, so the possibility that this Mexico-Guatemala trio

is a different species to *C. odorata* is low as the grouping is still resolved within the tree structure. However, it is possible that they may be differentiated (or differentiating) into a subspecies or race of *C. odorata* or have not been subject to outbreeding with other populations (i.e. have been reproductively isolated for some time).

Phylogeography of South African *C. odorata*

Because there is no direct genetic association to be found between South African and other samples, the Jukes Cantor distance matrix (Appendix 7) was studied directly. It was found that many of the samples shared multiple pairwise difference values of 0.00000. These data havebeen modified in Table 3.5, where the samples sharing distance values of 0.00000 have been listed. Where samples have no pairwise distance values of 0.00000, the lowest pairwise value for that taxon has been given, followed by the name(s) of the sample(s) with which that taxon shares the distance value (in red).

From Table 3.5, it can be seen that the South African samples (marked in green) share either a high number of pairwise distance values of 0.00000 (40 to 21 such associations) or have no pairwise associations of 0.00000 with any taxa at all (like the non-*C. odorata* samples), shown at the bottom of Table 3.5. The South African samples have low (0.00000) distance values with many other South African samples, as well as Jamaican and Venezuelan samples. Some of the South African samples have low distance values with Australian, Indian, Northern Sumatran and Mauritian samples, but none with the AusITS1-B sample, and only one sample, SA Hlu42, shared a low distance value with the unusual Austra81 sample. Only SA Dbn13 showed 0.00000 distance values with USA Fl12, Wafric17 and Guatema7 samples. None of the South African samples shared 0.00000 distance values with any Mexican, Brazillian, Puerto Rican or Costa Rican samples, suggesting that the origin of the South African *C. odorata* population would not be in any of these countries. As the *C. odorata* populations in Australia, India, Northern Sumatra and Mauritius are invasive (and therefore not the original origin of the South African population), the distance matrix values would seem to indicate Jamaica or Venezuela as possible countries of origin of the South African population of *C. odorata*.

The data from Table 3.5 has been used to create Table 3.6, where the frequency of pairwise differences of 0.00000 are summarised by country. In other words, Table 3.6 counts the number of times each country has a pairwise difference of 0.00000 with another country in Table 3.5. From this table, it can be seen that the South African sample has the highest number of pairwise differences of 0.00000 with the Jamaican samples (skewed by the number of Jamaican samples in the study), followed by the Venezuelan and Trinidad samples.

Table 3.5: Taxa with pairwise Jukes Cantor distance values of 0.00000, or closest to 0.00000.

SA Dbn13	USA Fl12	Guatema7	Venezue4	Jamaica9	SA Esh26	Trinid11	SA PSh28	Jamaica8	Maurit15	India 16
Jamaica9	Venezue4	SA Esh26	Jamaica8	SA PSh28	Jamaica8	Maurit15	India 16	SA Mtu27	SA PEd23	SA Tza18
SA Tza18	Venezue4	Jamaica9	SA Esh26	Trinid11	SA PSh28	Jamaica8	Maurit15	India 16	SA Mtu27	SA PEd23
SA PSh36	Venezue4	Jamaica9	SA Esh26	Trinid11	Maurit15	India 16	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13
Venezue4	Jamaica9	SA Esh26	Trinid11	Jamaica8	Maurit15	India 16	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13
Jamaic71	Brazil 5	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13
SA Tza95	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2
SA Mtu27	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA PEd23	SA Tza18	SA Dbn13	Venezue2	SA PSh36
SA Hlu42	Venezue4	Jamaica9	SA Esh26	Jamaica8	Maurit15	India 16	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13
SA Pal92	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2
SA Esh26	Venezue4	Jamaica9	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2	SA PSh36
SA PEd23	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA Tza18	SA Dbn13	Venezue2	Venezue1
Jamaic63	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue1
Trinid55	Venezue4	Jamaica9	SA Esh26	Jamaica8	Maurit15	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2
Jamaica8	Venezue4	Jamaica9	SA Esh26	SA PSh28	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2	SA PSJ45
Jamaic54	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36
Jamaic69	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2
SA Pal88	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36
Jamaic62	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36
SA Tza96	Venezue4	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36
SA PSJ50	Venezue4	Jamaica9	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36	SA PSJ45
AusITS1A	USA Fl12	Venezue4	Jamaica9	SA Esh26	Trinid11	Maurit15	India 16	WAfric17	SA Mtu27	SA PEd23
Jamaic59	Venezue4	Jamaica9	SA Esh26	Jamaica8	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36
SA PSh28	Jamaica9	SA Esh26	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36	Jamaic54	Jamaic57
Venezue2	Venezue4	Jamaica9	SA Esh26	Trinid11	Jamaica8	Maurit15	India 16	SA Mtu27	SA PEd23	SA Tza18
SA Hlu44	Venezue4	Jamaica9	SA Esh26	Jamaica8	SA Mtu27	SA Tza18	SA Dbn13	SA PSh36	SA PSJ45	SA Hlu42
Jamaic56	Brazil 5	Jamaica9	SA Esh26	SA PSh28	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36
SA PSJ45	Venezue4	Jamaica9	SA Esh26	Jamaica8	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	SA PSh36	SA Hlu44
Venezue1	Venezue4	Jamaica9	Trinid11	Maurit15	India 16	SA Mtu27	SA PEd23	SA Tza18	SA Dbn13	Venezue2
Thaila11	Venezue4	Jamaica9	SA Esh26	Trinid11	Maurit15	India 16	SA Mtu27	SA Tza18	SA Dbn13	Venezue2
NSumat38	Venezue4	Jamaica9	Trinid11	Maurit15	India 16	SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36
Trinid11	Venezue4	Jamaica9	Maurit15	India 16	SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11
Maurit15	Venezue4	Jamaica9	Trinid11	India 16	SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11
India 16	Venezue4	Jamaica9	Trinid11	Maurit15	SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11
Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	PRico100	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92
Austra84	Jamaica9	Trinid11	SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11	NSumat38	SA Hlu42
PRico100	Jamaic57	Jamaic56	Jamaic62	Jamaic71						
Brazil 5	Jamaic57	Jamaic56	Jamaic71							
USA Fl12	SA Dbn13	AusITS1A								
WAfric17	SA Dbn13	AusITS1A								
Guatema7	SA Dbn13									
Mexico10	Guatem52									
Venezue3	AusITS1A									
Jamaic58	PRico102									
PRico102	Jamaic58									
Brazil51	AusITS1B									
Guatem52	Mexico10									
Austra81	AusITS1B									
AusITS1B	Brazil51	Austra81								
CostRic6	0.00136	Jamaic57								
SA PSh32	0.00136	SA PEd23	SA PSJ45	Jamaic69						
SA PSJ48	0.00273	Jamaic59	SA Tza97							
SA Tza97	0.00275	SA PSJ48								
SA Tza99	0.00686	AusITS1A								
Ariparia	0.02869	Aadenoph								
Aadenoph	0.02869	Ariparia								
Cborinqu	0.03936	Csqualid								
Csqualid	0.04773	Pclemati								
Pclemati	0.04773	Csqualid								
Ccollina	0.07660	SA Hlu44								
Cmacroce	0.12642	Ccollina								

Table 3.5: Taxa with pairwise Jukes Cantor distance values of 0.00000, or closest to 0.00000 (contd.)

SA Dbn13	WAfric17	SA Mtu27	SA PEd23	SA Tza18	Venezue2	Venezue1	SA PSh36	Thaila11	SA PSJ45	NSumat38
Jamaica9	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11	SA PSJ45	NSumat38	SA Hlu44	SA Hlu42	Jamaic54
SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11	SA PSJ45	NSumat38	SA Hlu44	SA Hlu42	Jamaic54
SA PSh36	Venezue2	Venezue1	Thaila11	SA PSJ45	NSumat38	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56
Venezue4	Venezue2	Venezue1	SA PSh36	Thaila11	SA PSJ45	NSumat38	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57
Jamaic71	Venezue2	SA PSh36	Thaila11	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic59	Jamaic62
SA Tza95	Venezue1	SA PSh36	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62
SA Mtu27	Thaila11	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63
SA Hlu42	Venezue2	Venezue1	SA PSh36	Thaila11	SA PSJ45	NSumat38	SA Hlu44	Jamaic54	Jamaic63	SA PSJ50
SA Pal92	SA PSh36	Thaila11	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62
SA Esh26	Thaila11	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63
SA PEd23	SA PSJ45	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55
Jamaic63	SA PSh36	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	SA PSJ50
Trinid55	SA PSh36	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic59	Jamaic62	Jamaic63	SA PSJ50
Jamaica8	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55
Jamaic54	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55
Jamaic69	SA PSh36	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic59	Jamaic62	Jamaic63	SA PSJ50
SA Pal88	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50
Jamaic62	SA Hlu44	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic63	PRico100	SA PSJ50	Trinid55	Jamaic69
SA Tza96	SA PSJ45	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic62	Csqualid	SA PSJ50	Trinid55
SA PSJ50	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	Trinid55	Jamaic69	Jamaic71
AusITS1A	SA Tza18	SA Dbn13	Venezue2	Venezue1	SA PSh36	Thaila11	NSumat38	SA Hlu42	Venezue3	Trinid55
Jamaic59	Jamaic54	Jamaic57	Jamaic56	Jamaic62	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92
SA PSh28	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88
Venezue2	SA Dbn13	Venezue1	SA PSh36	Thaila11	NSumat38	SA Hlu42	Trinid55	Jamaic71	SA Pal92	SA Tza95
SA Hlu44	Jamaic54	Jamaic57	Jamaic56	Jamaic62	Jamaic63	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88
Jamaic56	SA Hlu44	Jamaic54	Jamaic57	Jamaic59	Jamaic62	Jamaic63	PRico100	SA PSJ50	SA Pal92	SA Pal88
SA PSJ45	SA Hlu42	Jamaic54	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95
Venezue1	SA PSh36	Thaila11	NSumat38	SA Hlu42	Jamaic57	Jamaic63	SA Tza95	Austra84	AusITS1A	
Thaila11	Venezue1	SA PSh36	NSumat38	SA Hlu42	Jamaic71	SA Pal92	Austra84	AusITS1A		
NSumat38	Thaila11	SA Hlu42	Austra84	AusITS1A						
Trinid11	NSumat38	Austra84	AusITS1A							
Maurit15	NSumat38	SA Hlu42	AusITS1A							
India 16	NSumat38	SA Hlu42	AusITS1A							
Jamaic57	SA Pal88	SA Tza95	SA Tza96							
Austra84	AusITS1A									
PRico100										
Brazil 5										
USA Fl12										
WAfric17										
Guatema7										
Mexico10										
Venezue3										
Jamaic58										
PRico102										
Brazil51										
Guatem52										
Austra81										
AusITS1B										

Table 3.5: Taxa with pairwise Jukes Cantor distance values of 0.00000, or closest to 0.00000 (contd.)

SA Dbn13	SA Hlu44	SA Hlu42	Jamaic54	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55
Jamaica9	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92
SA Tza18	Jamaic57	Jamaic56	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92
SA PSh36	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95
Venezue4	Jamaic59	Jamaic62	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95
Jamaic71	Jamaic63	PRico100	SA PSJ50	Trinid55	Jamaic69	SA Pal92	SA Pal88	SA Tza95	SA Tza96	AusITS1A
SA Tza95	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96	AusITS1A
SA Mtu27	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96	AusITS1A	
SA Hlu42	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	Austra81	Austra84	AusITS1A	
SA Pal92	Jamaic63	SA PSJ50	Trinid55	Jamaic69	Jamaic71	SA Pal88	SA Tza95	SA Tza96	AusITS1A	
SA Esh26	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96	AusITS1A		
SA PEd23	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96	AusITS1A			
Jamaic63	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96			
Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96	AusITS1A			
Jamaica8	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96				
Jamaic54	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96				
Jamaic69	Trinid55	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96				
SA Pal88	Trinid55	Jamaic69	Jamaic71	SA Pal92	SA Tza95	SA Tza96				
Jamaic62	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96					
SA Tza96	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95					
SA PSJ50	SA Pal92	SA Pal88	SA Tza95	SA Tza96						
AusITS1A	Jamaic71	SA Pal92	SA Tza95	Austra84						
Jamaic59	SA Pal88	SA Tza95	SA Tza96							
SA PSh28	SA Tza95	SA Tza96								
Venezue2	Austra84	AusITS1A								
SA Hlu44	SA Tza95	SA Tza96								
Jamaic56	SA Tza95	SA Tza96								
SA PSJ45	SA Tza96									
Venezue1										
Thaila11										
NSumat38										
Trinid11										
Maurit15										
India 16										
Jamaic57										
Austra84										
PRico100										
Brazil 5										
USA Fl12										
WAfric17										
Guatema7										
Mexico10										
Venezue3										
Jamaic58										
PRico102										
Brazil51										
Guatem52										
Austra81										
AusITS1B										

Table 3.5: Taxa with pairwise Jukes Cantor distance values of 0.00000, or closest to 0.00000 (contd.)

SA Dbn13	Jamaic69	Jamaic71	SA Pal92	SA Pal88	SA Tza95	SA Tza96	Austra84	AusITS1A
Jamaica9	SA Pal88	SA Tza95	SA Tza96	Austra84	AusITS1A			
SA Tza18	SA Pal88	SA Tza95	SA Tza96	Austra84	AusITS1A			
SA PSh36	SA Tza96	Austra84	AusITS1A					
Venezue4	SA Tza96	AusITS1A						
Jamaic71								
SA Tza95								
SA Mtu27								
SA Hlu42								
SA Pal92								
SA Esh26								
SA PEd23								
Jamaic63								
Trinid55								
Jamaica8								
Jamaic54								
Jamaic69								
SA Pal88								
Jamaic62								
SA Tza96								
SA PSJ50								
AusITS1A								
Jamaic59								
SA PSh28								
Venezue2								
SA Hlu44								
Jamaic56								
SA PSJ45								
Venezue1								
Thaila11								
NSumat38								
Trinid11								
Maurit15								
India 16								
Jamaic57								
Austra84								
PRico100								
Brazil 5								
USA Fl12								
WAfric17								
Guatema7								
Mexico10								
Venezue3								
Jamaic58								
PRico102								
Brazil51								
Guatem52								
Austra81								
AusITS1B								

Table 3.6 Number of pairwise differences of 0.00000, by country.

	Venezuela	Jamaica	Trinidad	Australia	Thailand	N Sumatra	Mauritius	India	USA	Puerto Rico	Brazil	W Africa	Guatemala	Mexico	Costa Rica
South Africa	29	130	25	13	7	4	4	4	1	0	0	1	1	0	0
Venezuela		13	4	6	3	3	3	3	0	0	0	0	0	0	0
Jamaica			9	3	2	1	1	1	0	5	0	0	0	0	0
Trinidad				3	1	1	1	1	0	0	0	0	0	0	0
Australia					2	2	1	1	1	0	1	1	0	0	0
Thailand						1	1	1	0	0	0	0	0	0	0
N Sumatra							1	1	0	0	0	0	0	0	0
Mauritius								1	0	0	0	0	0	0	0
India									0	0	0	0	0	0	0
USA										0	0	0	0	0	0
Puerto Rico											0	0	0	0	0
Brazil												0	0	0	0
W Africa													0	0	0
Guatemala														1	0
Mexico															0

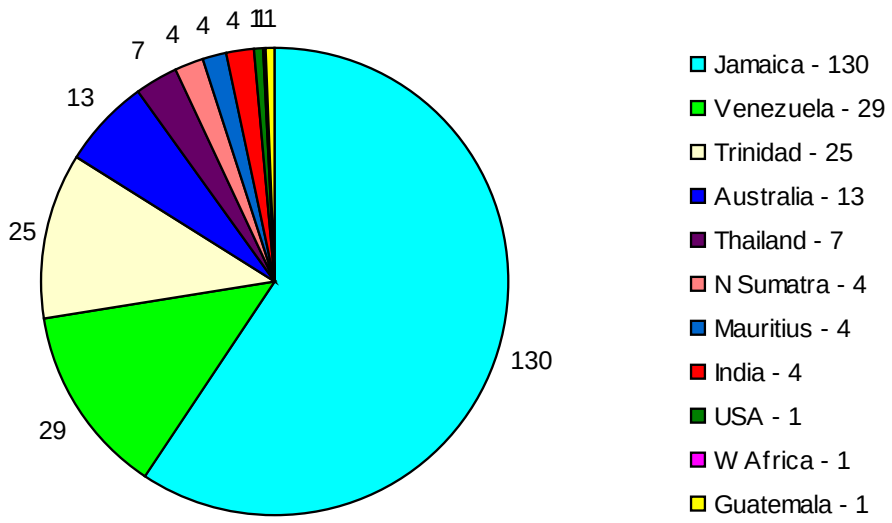

Figure 3.12 Proportion of pairwise differences of 0.00000 to South African samples, by country, as taken from Table 3.5.

Theory predicts that colonisation of new areas will be associated with population bottlenecks that reduce within-population genetic diversity and increase genetic differentiation among populations (Amsellem *et al.*, 2000b). This should be especially true for weedy plant species, which are often characterised by self-compatible systems and vegetative propagation (Amsellem *et al.*, 2000b).

Morphologically, the samples most similar to the South African population are from the Greater Antilles (GA). It may be possible that the South African population originated from the GA and has since diversified genetically subsequent to a bottleneck to the population. However, the South African population has only been isolated for approximately 60 years (*ca.* 60 generations), which is a relatively short period of time for a population to emerge from a bottleneck, so it is possible that the South African population is currently within a population bottleneck.

The other invasive populations sampled (Australasia, West Africa, Mauritius) show no genetic association with any samples from the GA. It could thus be possible that all of the invasions of *C. odorata* (not including South Africa) are from regions other than the GA, and that the South African population is GA in origin. It is possible that the GA populations are distinct from the "mainland" populations, and have different plant-predator relationships. The faunal biogeography of the GA, as presented by Liebherr (1988) and Donnelly (1988) supports the recent isolation and thus potential modification of plant-predator systems on the GA relative to South and Central American mainlands. This would lead to high plant-predator specificity, and could account for the difficulties encountered in getting biocontrol agents that are successful on other invasive populations (here postulated not to originate in the GA) to

establish on the South African population (postulated to be of GA origin). The GA island that has been considered most likely to be the origin of the South African population of *C. odorata* is Jamaica (Zachariades, pers. com.). This theory is supported by the data in this dissertation (Figure 3.12), which shows that the South African samples have a very high number of pairwise difference values of 0.00000 with Jamaican samples. Additionally, a comparison of the frequency at which a Jamaican sample was found in a position sister to a South African sample was carried out. (This comparison was only possible with the Maximum Parsimony and Neighbour Joining Distance trees.) It was found that 72% of the South African samples were in a position sister to the Jamaican samples in the Maximum parsimony analysis, but only 53% were found sister to Jamaican samples in the NJ distance analysis. Therefore, it is suggested that there is a genetic link between the South African and Jamaican samples. This is supported by the pairwise Jukes Cantor values (Table 3.5) which indicate Jamaica as one of the possible origins of the South African *C. odorata*.

Within South Africa itself, several regions were sampled to investigate the degree of genetic divergence within and between populations of *C. odorata* in South Africa. Several individuals sampled within each of the regions; Port Shepstone (PSh); Port St. Johns, (PSJ); Hluhluwe (Hlu); Tzaneen (Tza) and Phalaborwa (Pha). In all four of the genealogical reconstructions the samples from each region are scattered throughout the tree, and not grouped according to location (or population), as might be expected. Thus it can be said that there is no population structure within the South African population of *C. odorata* based on ITS data.

Significance of Differences in Flower Colour and Root System

The South African *C. odorata* has whitish to cream capitula, while *C. odorata* from West Africa and Asia has pale mauve, pale blue, or whitish capitula (Holm *et al.*, 1977; Kluge, 1990). The flowers from Trinidad are whitish or pale blue in colour (McFadyen, 1988b), and another similarity between the South African and Trinidad samples is that both have shallow, fibrous root systems, while *C. odorata* elsewhere has a deep taproot system. The Trinidad samples have been genetically associated with the regions of the world where *C. odorata* has become naturalised (excluding South Africa) and a fibrous root system of *C. odorata* in Australasia, Mauritius, and West Africa has not been mentioned in the literature. Thus, one must conclude that the characters of root system and flower colour are not congruent with the ITS genealogy, and are possibly not 'good' or informative characters. A fibrous root system may be a function of climate or substrate (*i.e.*, morphologically plastic).

Distribution of Morphotypes

The ingroup in the Maximum Parsimony tree (Figure 3.4) is weakly divided into two clades, one of which contains mostly samples with a South African morphology, as well as the Guatema7-Mexico10-Guatem53 trio. This weak division could indicate that the ITS gene of *C. odorata* is diverging, and that not all of the gene copies have completed this divergence possibly due to incomplete concerted evolution processes.

In the Neighbour Joining Distance tree, as in the strict consensus maximum parsimony tree, the morphotypes are not grouped together in clades on the gene tree, however there are two clades that contain only samples of South African morphotype: (Jamaic59, Jamaica8, SA_Pal88, and SA_Mtu27) and (SA_Psh32, SA_PSJ45, Jamaic69, Jamaic54, SA_Tza96, SA_Pal92, and SA_Tza95). This concentration of samples with South African morphotypes may indicate that the morphology is indicative of a genetic divergence process, where samples with the South African morphotype are diverging from samples with Venezuelan and Floridean morphotypes.

The large central haplotype in Figure 3.5 could be interpreted in two possible ways: Firstly, in a minimum spanning tree, such as in Figures 3.5 and 3.9, the most frequent haplotypes in any population are probably the oldest haplotypes, and should therefore cluster relatively close together, near the centroid of the phylogenetic space or MST, whereas the rarer variants should occupy the outer fringes of the tree (Smouse, 1998; Schaal *et al*, 1998 and literature therein). Good trees generally have the common haplotypes near the centre of mass, where they define the central branching order, with the rare haplotypes being relegated to the periphery (Smouse, 1998). In the minimum spanning network (MSN) from TCS (Figure 3.5), the most common haplotype (and thus presumably the oldest) contains samples from all three 'morphotypes'. This would indicate that there is a lack of congruence between morphology and statistical parsimony-defined haplotypes. However, in some studies, the pattern of haplotype distribution reveals that a sizeable portion of a species range contains little or no genetic variation relative to the rest of the species range (Schaal *et al*., 1998), as would be the second interpretation of the large, central haplotype in Figure 3.5. Such a pattern is consistent with a rapid range expansion in that region, which may indicate rapid historical dispersal caused by humans, as in the legume *Gliricidia* (Schaal *et al*., 1989).

Thus, the existence of a large, central haplotype in the MST could either be the oldest and most common haplotype, or it may simply be diagnostic of human-caused range dispersal.

Very little can be said about the arrangements of the samples with various morphologies in Figure 3.9, because the arrangement at the centre of the tree is so obscured. In the rooted

phylogram of this reconstruction (Figure 3.10), it can be seen that there are no general groupings or associations formed between samples with similar morphologies. However, it is noted that the samples in Figure 3.9 that are distant from the central region of the minimum spanning tree are the same haplotypes that are not included in the large central haplotype in the statistical parsimony-based minimum spanning network in Figure 3.5. This would suggest that these central samples are older, more common haplotypes, as suggested by the first interpretation of the large central haplotype in Figure 3.5.

The relative lack of population genetic structure is not limited to South African populations. It would be expected that populations that have been geographically separated for longer than South Africa's 60 years, such as the Jamaican and Venezuelan samples, would show some differentiation from each other. In addition, it would have been expected that populations that are diverging from one another would show more similarity within than between populations, and therefore each population would be resolved as distinct clades on a gene tree. In particular, it would have been expected that the relatively isolated South African population would have resolved or would be resolving into a clade of its own. However, the samples from these countries appear randomly distributed throughout the genealogies, as with the South African samples. Thus it appears that *C. odorata* is not in the process of population differentiation. This could either be as a result of continuous and multiple re-introductions of the species to both native and invaded areas, or it could be an indication that this plant is able to alter its genetic composition with relative ease and speed. The precise endemic region of *C. odorata* is unknown (Robinson, pers. comm. to Zachariades) as human intervention has been aiding its geographic range expansion for many centuries. It is therefore not possible to say whether the GA populations are as a result of recent introductions and therefore subject to their own bottlenecks, or whether they were separated from the mainland populations by vicariance or arrived on the islands via relatively recent natural dispersal events.

Possibility of Finding Consensus between Previous and Current Findings

In a study of the genetic variability of *Wyethia reticulata*, Ayres and Ryan (1997) found that the RAPD and allozyme data sets each revealed a different pattern of variation (with multivariate similarity matrices), and that no consensus was found in cluster arrangements in dendrograms for these two methods. Despite these differences, they found that only when the allozyme and RAPD markers were combined were completely geographically correlated groupings obtained. Therefore it is possible that a combination of all analyses done on the

origin of the South African *C. odorata* would prove more informative than any single analysis on its own.

In previous studies aimed at finding the origin of South African *C. odorata*, Thailand, Brazil and Jamaica are most frequently mentioned as having similarities to South African *C. odorata*. From field observations, the country cited most often as having plants morphologically similar to South African *C. odorata* is Jamaica, though samples from Puerto Rico have also been mentioned as being identical to South Africa (Zachariades, pers. comm.). Jamaica and Venezuela have genetic links to the South African *C. odorata* in the ITS genealogy through the pairwise distance comparison.

The previous allozyme study of *C. odorata* (Vos, 1989) indicated similarities to the South African populations with Thailand samples, and that the South American samples were more similar to each other than to the South African and Thailand plants. The distance matrix data shows that at lease some of the South African samples have a distance value of 0.00000 with the Thail11 sample. Therefore, the ITS data agrees with the allozyme findings that there are similarities between some of the South African and Thai samples.

The manner in which most of the South African samples of *C. odorata* appear scattered throughout the various ITS gene genealogies can be interpreted to mean several things: (1) that the ITS region is not sufficiently variable to detect intraspecific signal, (2) that the signal is mixed due to incomplete concerted evolution, or (3) that the probably man-mediated dispersal of *C. odorata* throughout its existing 'native' range has resulted in introgression, removing any genetic structure that was initially present. The ITS region has a higher mutation rate in Asteraceae than most other plant taxa (Schmidt and Schilling, 2000), and thus it could be expected to be more informative about the intraspecific relationships than the intraspecific study of the ITS within Crucifereae (Mayer and Soltis, 1999), which shows that the ITS is sufficiently variable for intraspecific use. With the higher mutation rates of the ITS in Asteraceae (as commpared to families such as Crucifereae), there must be little doubt that the ITS region in *C. odorata* is insufficiently variable to detect intraspecific variation. Any useful intraspecific phylogenetic signal of the ITS region for *C. odorata* must be masked by (2) or (3) above. Either incomplete concerted evolution is resulting in multiple paralogues within each population (as discussed by Mayer and Soltis, 1999), or continued man-mediated dispersal is preventing the homogenisation of individual populations and divergence between populations. Therefore the groupings that are found consistently throughout most analyses and which agree with some morphological trends suggest that (due to this high mutation rate) the ITS region could be beginning to reflect population divergences, or have lost them due to introgression, or are reflecting vestigial differences that were present prior to human aided dispersal and subsequent introgression.

With reference to the sequences obtained, several nucleotide polymorphisms were scored in most samples (Figure 3.13), and it was thought that samples that had become naturalised (or were in the process of naturalisation) would have the highest number of polymorphic sites, as an indication of emergence from population bottlenecks and thus genetic adaptation to the new environment(s) through outcrossing or possibly polyploidisation. (This would assume multiple introductions of the species, or at least the introduction of several individual plants or successfully germinating seeds.) From Figure 3.13, it is evident that many of the populations that have naturalised - Australasia, West Africa and Mauritius have relatively low numbers of polymorphic sites. All have two or less polymorphic sites, except for sample Austra81 (which the ITS genealogy indicates is from a separate origin) that has four polymorphic sites. The samples from the South African population of *C. odorata* have a range in the number of polymorphic sites, from none (SA_PSh28) to 15 (SA_Hlu42). In addition, the samples from the regions where *C. odorata* is considered indigenous also have a range in polymorphism frequencies from zero to 11 over the entire ITS sequenced region. Thus, it appears that the frequency of polymoprhic sites in the ITS region of *C. odorata* cannot be correlated with population divergence or emergence from bottleneck processes that might be expected to have occurred in regions where it has recently been introduced. It is still possible, though, that the frequency of polymorphic sites is correlated with mutation rates, and that would indicate that *C. odorata* is undergoing rapid genetic mutation in some of its native regions as well as in some of the regions where it has been introduced as a result of human-mediated dispersal.

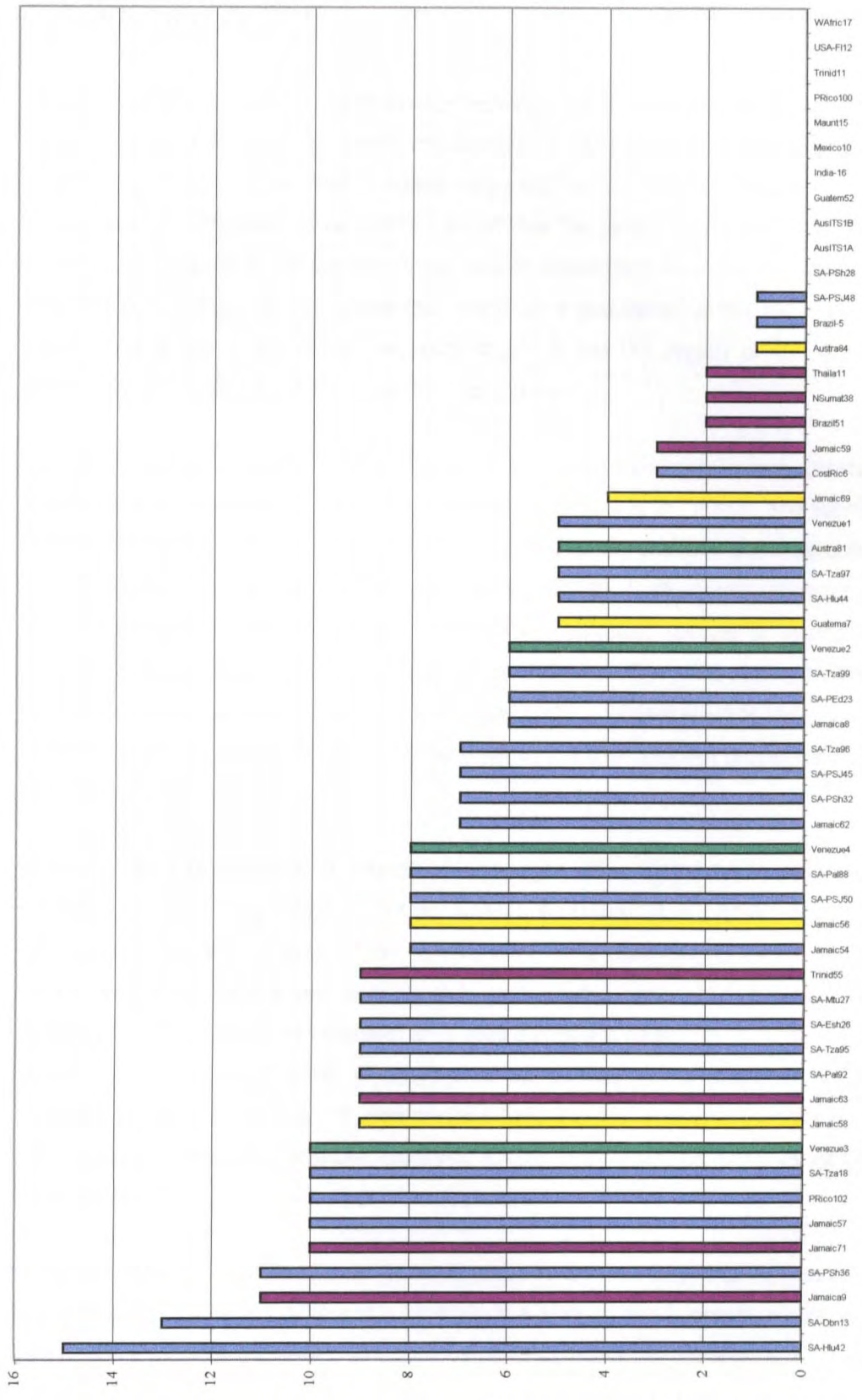

Figure 3.13 Frequency of polymorphic sites in ITS sequences of Chromolaena odorata ingroup samples. Colours indicate morphotypes (Blue = South African, Green = Venezuelan, Mauve = Floridean.

Significance of Genetic Diversity of the South African *C. odorata* population.

This is not the first study to use molecular techniques to characterise the genetic diversity of an invasive weed and compare it with the diversity of other samples of the same species from other areas of introduction and its native range with an aim towards improving a biological control project. Amsellem *et al.* (2000a) found that the genetic variability of RAPD markers among individuals of *Rubus* spp within the area of introduction was low, and they suggested that this meant that the individuals that made up a population within each area sampled represented a clone; the genetic diversity found in the ITS region of the South African population of *C. odorata* indicates that it is not a clone.

The genus *Rubus* is usually hermaphroditic, thus, while outcrossing occurs, plants are also potentially self-pollinating, and selfing is frequent (Amsellem *et al.,* 2000b, and literature cited therein). *Rubus* spp. propagate vegetatively very vigorously, enabling clonal spread of single individuals in a patch of habitat (Amsellem *et al.*, 2000b). Furthermore, apomixis occurs in the two subgenera most important in horticulture: polyploid species of *Rubus* and, very rarely, *Idaeobatus* (from polyploid species deviant of normally diploid species) (Amsellem *et al.*, 2000b). Some similarities between *Rubus* spp and *C. odorata* are that both *Rubus* and *C. odorata* are usually apomictic and polyploids, yet *C. odorata* does not propagate vegetatively very easily.

Within areas of introduction, *R. alceifolius* is strongly differentiated from populations within the native range in Asia (Amsellem *et al.*, 2000b). Amsellem *et al.* (2000b) found that their data agreed with the statement that genetic diversity is much higher in the native range, where geographical areas are discriminated even at the scale of intra-area diversity. In contrast, in most areas of introduction, intra-area diversity is negligible, and there is very little genetic variation among areas (Amsellem *et al.*, 2000b). In a single invaded locality, Madagascar, genetic variability is intermediate between the level of variability observed in other areas of introduction and that observed among localities in the native range (Amsellem *et al.*, 2000b).

In Madagascar, *R. alceifolius* is not especially invasive. In contrast, in La Réunion, Mauritius, Mayotte, and Queensland, where this bramble is a serious weed, genetic variation is much lower. There may be a causal connection between diversity and invasiveness, in the sense that a particular well-adapted genotype to a specific biotype may propagate through an area very quickly via asexual reproduction (Amsellem *et al.*, 2000b).

Thus, it is possible that if a weed is genetically suited to an environment, then there is no reason for it to change or mutate its genetic identity, and it will reproduce via means that do not involve genetic change, *e.g.*, vegetative propagation. However, if a weed is ill-suited to an environment, then it may resort to sexual reproduction rather than vegetative, to increase the genetic diversity and thus adaptability of the population.

Is it possible that the high genetic variability of the *C. odorata* population within South Africa is due to a process of mutation whereby it is becoming better adapted to the South African climate. It is tempting to consider here that the relatively high frequency of polymorphic sites in ITS sequences from South African samples is related to its adaptation to the new climate however, this high frequency of polymorphic sites is not limited to the South African (or even any non-indigenous) samples. Thus, it seems that the high genetic variability in the South African population of *C. odorata* is related to the high mutation rate in the species as a whole, as opposed to any processes of adaptation.

Possible Roles of Polyploidy and Concerted Evolution in Genetic Diversity of *C. odorata*

The basal chromosome number for *Chromolaena* is n = 10. The species *C. odorata* has either a chromosome number of n = 40 (Powell and King, 1969) or n = 60 (Vos, 1989, and literature therein). Thus, *C. odorata* can be a tetraploid or a hexaploid, however it is unknown whether the tetraploid and hexaploid samples came from the same or different localities.

Assumptions of random mating and random synapsis of homologous chromosomes may not be the case in allotetraploids (Roelofs *et al.*, 1997), so the possibility exists that different variations of the ITS gene are present in the multiple chromosome sets of *C. odorata*, and that different copies of these variations are dominant and therefore sequenced (direct sequencing) in each sample, resulting in the range of sequences displayed. [The relatively high number of polymorphic sites in the ITS region (Figure 3.11) are evidence that several different copies of the ITS region are present in most of the samples.]

The high frequency of polymorphic sites in the ITS region of *C. odorata* does not appear to be common among the Asteraceae. In an investigation into the tetraploid nature of *Microseris*, Roelofs *et al*. (1997) found that two tetraploid *Microseris* species contained only a single ITS sequence each instead of two different ones – one from each of the two parental genomes. This was thought to be the result of concerted evolution (sequence homogenisation) including recombination between the two parental ITS sequences combined in the original hybrid. Similarly, in a study of the relationships of *Armeria* at the interspecific

and intraspecific levels, Aguilar *et al*. (1999a) found that the ITS in *Armeria* shows a great sequence homogeneity: only two indels were recorded in the 625 bp alignment.

Thus it appears that in at least two genera of the Asteraceae (of which *C. odorata* is a member) the multiple chromosome copies of the ITS region are rapidly homogenised, most probably by concerted evolution, making it debatable whether the genetic variety observed in the ITS of *C. odorata* is due to the presence of multiple ploidy levels. However, it may be possible that concerted evolution is not functioning or is functioning at a retarded rate in *C. odorata* resulting in high sequence divergence. Or, it maybe possible that concerted evolution has not been completed in the *ca.* 60 generations since the introduction of the *C. odorata* population to South Africa since the polyploidisation event.

Both spacers of the ITS region have a nucleotide substitution rate high enough to generate intra- and interspecific variability. However concerted evolution (sequence homogenisation) of the gene usually results in homogenisation of the sequences of ribosomal DNA, resulting in lower intraspecific variability than interspecific (Aguilar, 1999a). As an example, Aguilar (1999b) found that concerted evolution can occur as quickly as one generation after the combination of two parental ITS types. They also suggest that the fast rate of homogenisation in (artificial) hybrids has important implications for detecting past or recent hybrid events in natural populations. Concerted evolution promotes intragenomic uniformity of repeat-units (Baldwin *et al*., 1995), and has been invoked as the mechanism responsible for ITS homogenisation after hybridisation events.

However, a lack of concerted evolution has been used to explain the presence of polymorphic repeats in some ITS sequences where an absence of sexual recombination and presence of nrDNA loci on nonhomologous chromosomes seems to slow or preclude concerted evolution after hybridisation events (Campbell *et al*., 1997; Sang *et al*., 1995; Walters and Schaal, 1996, all cited in Aguilar *et al*., 1999b). The forces of concerted evolution were found to be insufficient to homogenise the arrays of 5S RNA genes in the Triticeae (Kellogg and Appels, 1995).

The possibly apomictic nature of *C. odorata* could have two possible effects on the ITS genome, irrespective of polyploidy. Firstly, it would retain the copy number and different copies within the genome, as concerted evolution would not be able to operate effectively through normal meiosis. If apomixis and sexual reproduction are both occurring within the species, then the apomictic reproduction could be functioning to reduce the rate at which concerted evolution is occurring, thereby retaining diverse copies within any single genome. The second effect that facultative apomixis could have is to retard the rate of increase (or decrease) in frequency of a particular set of alleles (*i.e.*, ITS copy) through the population

over time, as sexual reproduction would be required for allele frequency to change over time. The effect of polyploidy (not including autopolyploidy) on the ITS region would be to potentially introduce additional copies of the ITS region to the genome.

Should concerted evolution in the ITS region of *C. odorata* be retarded, it is possible that several different copies of the gene (possibly introduced as a result of polyploidy) are present, allowing not only for a degree of false variation within a genotype, but also for the production of false (chimeric) genotypes for some samples and absence of phylogeographic signal.

With the occurrence of both apomixis and sexual reproduction in *C. odorata*, one would expect geographically isolated populations to have distinct (and therefore detectable) gene pools. A null hypothesis that was assumed in this dissertation is that there are/were distinct genetic populations of *Chromolanea odorata* in the native habitat, and therefore that there is a possibility of identifying the population from which an invasion/invasions occurred. What was not known is that there are no distinct genetic populations of *C, odorata* in its native habitat, and therefore the probability of tracing the 'parental' population of the South African (or any other invasive population) would be unlikely.

The results of this investigation have shed some light in the possible geographic origins of the South African population of *c. odorata*. Nevertheless, the results are by no means conclusive, and have raised some intriguing questions:

- How variable are the chromosome numbers within the species of *C. odorata*? A study correlating chromosome counts with ITS sequences might resolve the uncertainty surrounding the polyploidy (and polymorphisms) within the species.

- How many ITS paralogues are there in each sample? An investigation and absence of phylogeographic signal could answer this question.

- To what extent is concerted evolution functional in *C. odorata*? Greenhouse crosses of known ITS paralogues could help in elucidating the functionality of concerted evolution in *C. odorata*.

FINAL CONCLUSIONS

The aims of this study were to use PCR-based molecular techniques to obtain and compare genetic fingerprints and DNA sequences of samples of *C. odorata* in order to locate the geographic origin of the South African population. Despite many attempts and a rigorous troubleshooting programme, the ISSR fingerprints were never obtained, and the study was then based on sequence data.

Comparisons were made of ITS and ETS sequences, and though initial studies showed that the ETS region contained a higher variability and informativeness than the ITS, the ETS region was difficult to amplify, and therefore the study focussed on comparisons of ITS sequences.

To the author's knowledge, this is one of the first studies to use the ITS region at an intraspecific level, particularly of this scale. Other intraspecific studies using the ITS region have focussed on defining relationships both between and within species, and this is the first study that the author knows of that has concentrated on the use of the ITS region for resolution of relationships of populations of only a single species. This is also the only study known to the author that uses methods other than parsimony for resolution of the phylogenetic relationships within a species. Methods used in this study were: strict parsimony, statistical parsimony to resolve reticulate haplotypic relationships (TCS), Jukes Cantor Neighbour Joining distance (including Jukes Cantor distance matrix) and Arlequin distance analysis to resolve distance-based reticulate relationships.

Each ITS sequence obtained in this dissertation was unique, indicating that the ITS region is sufficiently variable for intraspecific use within *C. odorata*. However, there was very little differentiation of genetic populations into clades or nodes. A possible reason for the lack of phylogenetic signal in the ITS region may be that incomplete or impaired concerted evolution has been operating, or that man-mediated dispersal of *C. odorata*, has resulted in introgression, thus confusing any signal indicative of population isolation, or both processes. However, the lack of genetic structure that can be correlated to geographic origin, is as a result of a lack of a period of isolation where population divergence would have occurred. There appear to be no distinct genetic populations of *C. odorata* in either the native or invaded habitats, indicating that gene flow is, or has been, occurring on a continual basis throughout its distribution range, hindering the divergence of populations into distinct genotypic populations.

Comparisons of molecular data and previous morphological observations led to the suggestion that the population of *C. odorata* that is common in Australia, West Africa and Mauritius is g enetically and morphologically similar to samples from Trinidad, which is therefore suggested as the origin of the invasive *C. odorata* in these three countries.

The less common morphotype of *C. odorata* in the Tulley region of Australia is genetically similar to a sample from Brazil, indicating that this sample represents a second infestation of a different genotype from Brazil.

Despite the apparent lack of genetic correlation of South African samples to any other population, a comparison of genetic data as well as morphological observations led to the suggestion that the South African population of *C. odorata* originated in either Jamaica or Venezuela.

A surprisingly high amount of genetic variation was found within the South African population of *C. odorata*. Not one of the sequences (from South Africa or otherwise) were identical, and 18 of the 19 South African samples show polymorphic sites within the ITS region.

A further investigation into the genetic distances between samples gave another indication of the range of genetic variation within the ITS of the South African population: 15 of the 19 South African samples shared 40 to 21 pairwise distance values of 0.00000 (close genetic similarity), while four had no distance values of 0.00000 at all (*i.e.*, they were relatively distinct from other sequenced samples).

This apparent genetic variability within the South African population of *C. odorata* may be the reason that difficulties have been encountered in successfully establishing biological control agents on the species within South Africa. If the potential biocontrol agents that are currently undergoing field trials on *C. odorata* do not establish, a possible future approach could involve regarding the weed in South Africa as being comprised of distinct genetic populations.

Although there is little genetic (or morphological) differentiation on the South African population into distinct, geographically located populations, further genetic studies (possibly involving cpDNA and other sources of nDNA information) might show that the South African population of *C. odorata* can be divided into several geographically overlapping, yet genetically distinct, populations. These potential populations may be the result of separate introductions into South Africa that have not yet begun crossing with the other populations. It may thus be found that the use of several biocontrol agents – one for each potential genetic population of *C. odorata* in South Africa – would be successful in the control of the weedy pest where single biocontrol agents have thus far failed.

LITERATURE CITED

ADAMS, R. P., M. ZHONG, AND Y. FEI. 1999. Preservation of DNA in plant specimens: inactivation and re-activation of DNases in field specimens. Molecular Ecology 8: 681 – 684.

AGUILAR, J. F., J. A. ROSSELLÓ, AND G. N. FELINER. 1999a. Molecular evidence for the Compillospecies model for reticulate evolution in *Armeria* (Plumbaginaceae). Systematic Biology 48: 735 – 754.

_____., _____., and _____. 1999b. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Molecular Ecology 8: 1341 – 1346.

AMBIKA, S. R., AND X. JAYACHANDRA. 1980. Suppression of plantation crops by *Eupatorium* weed. Current Science 20: 874 – 875.

_____., and _____. 1990. The problem of *Chromolaena* weed. *Chromolaena odorata* Newsletter 3: 1 - 6.

AMSELLEM. L., T. LE BOURGEOIS, J. L. NOYER, AND M. HOSSAERT-MCKEY. 2000a. Comparison of genetic diversity of the weed *Rubus alceifolius* in its introduced and native areas. In Proceedings of the X International Symposium on Biological Control of Weeds, Montana, USA.

_____., L., J. L. NOYER, T. LE BOURGEOIS, AND M. HOSAERT-MCKEY. 2000b. Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Molecular Ecology 9: 443 – 455.

ANWARULLA, M. S., AND S. C. CHANDRASHEKAR. 1996. A novel approach for combating Chromolaena problem: possibilities of its use as a green manure. In Proceedings of the Fourth International Workshop on Biological Control and Management of *C. odorata,* Bangalore, India.

AQUADRO, C. F., W. A. NOON, D. J. BEGUN, AND B. N. DANFORTH. 1998. RFLP analysis using heterologous probes. In A. R. Hoelzel (ed.), Molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

ARNOLD, M. L. 1993. *Iris nelsonii* (Iridaceae): origin and genetic composition of a homoploid hybrid species. American Journal of Botany 80: 577 - 591

AVISE, J. C. 1998. The history and purview of phylogeography: a personal reflection. Molecular Ecology 7: 371 – 379.

_____., AND K. WOLLENBERG. 1997. Phylogenetics and the origin of species. Proceedings of the National Academy of Sciences, USA 94: 7748 – 7755.

_____., J. ARNOLD, R. M. BALL, E. BIRMINGHAM, T. LAMB, J. E. NEIGEL, C. A. REEB, N. C. SAUNDERS. 1987. Intraspecifc phylogeography: the mitochondrial DNA bridge between populations genetics and systematics. Annual Review of Ecology and Systematics 18: 489 – 522.

AYRES, D. R., AND F. J. RYAN. 1997. The clonal and population structure of a rare endemic plant, *Wyethia reticulata* (Asteraceae): allozyme and RAPD analysis. Molecular Ecology 6: 761 – 772.

BACHMANN, K. 1997. Nuclear DNA markers in plant biosystematic research. Opera Botanica 132: 137 – 148.

BACKELJAU, T., L. DE BRUYN, H. DE WOLF, K. JORDAENS, S. VAN DONGEN, R. VERHAGEN, AND B. WINNEPENNINCKX. 1995. Random amplified polymorphic DNA (RAPD) and parsimony methods. Cladistics 11: 119 – 130.

BAIN, J. F., AND J. L. GOLDEN. 2000. A phylogeny of *Packera* (Senecioneae; Asteraceae) based on internal transcribed spacer region sequence data and a broad sampling of outgroups. Molecular Phylogenetics and Evolution 16: 331 - 338.

BALDWIN, B. G. 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: and examples from the Compositae. Molecular Phylogenetics and Evolution 1: 3 – 16.

_____. 1993. Molecular phylogenetics of *Calycadenia* (Compositeae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. American Journal of Botany 80: 222 – 238.

_____., M. J. SANDERSON, J. M. PORTER, M. F. WOKCIECHOWSKI, C. S. CAMPBELL, AND M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Annals of the Missouri Botanical Garden 82: 247 - 277.

_____., AND S. MARKOS. 1998. Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). Molecular Phylogenetics and Evolution 10: 449 – 463.

BARKER, N. P., L. Y. KRUGER, J. H. VLOK, AND E. H. HARLEY. 1996. RAPD data suggest the Gamkapoort Aloe to be of hybrid origin. South African Journal of Botany 62: 292 – 295.

BERMINGHAM, E., AND C. MORITZ. 1998. Comparative phylogeography: concepts and applications. Molecular Ecology 7: 367 – 369.

BINGGELI, P. 1997. *Chromolaena odorata* (L.) King & Robinson (Asteraceae). http://members.tripod.co.uk/WoodyPlantEcology/docs/web-sp4.htm Last accessed on 24/06/2002.

BLACKMORE, A. C. 1998. Seed dispersal of *Chromolaena odorata* reconsidered. In Proceedings of the Fourth International Workshop on Biological Control and Management of *C.odorata,* Bangalore, India.

BLAIR, M. W., O. PANAUD, S. R. MCCOUCH. 1999. Inter-simple sequence repeat (ISSR) amplification for analysis of micro-satellite motif frequency and fingerprinting in rice. Theoretical and Applied Genetics 98: 780 – 792.

BOPPRÉ, M. 1991. A non-nutritional relationship of *Zonocerus* (Orthoptera) to *Chromolaena* (Asteraceae) and general implications for weed management. In Proceedings of the Second International Workshop on Biocontrol and Management of *Chromolaena odorata*, Bogor, Indonesia.

BROOKFIELD, J. F. Y. 1992. DNA fingerprinting in clonal organisms. Molecular Ecology 1: 21 – 26.

CAMACHO, F. J., AND A. LISTON. 2001. Population structure and genetic diversity of *Botrychium pumicola* (Ophioglossaceae) based on inter-simple sequence repeats (ISSR). American Journal of Botany 88: 1065 – 1070.

CHACO, M. J., AND A. U. NARASIMHAM. 1988. Biocontrol attempts against *Chromolaena odorata* in India – a review. In Proceedings of the First International Workshop on Biological Control of *C. odorata*. Bangkok, Thailand.

CHAMBERS, G. K., AND E. S. MACAVOY. 2000. Microsatellites: consensus and controversy. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 126: 455 - 476

CHANDRASHEKAR, S. C., AND G. N. GAJANANA. 1996. Exploitation of *Chromolaena odorata* (L.) King and Robinson as green manure for paddy. In Proceedings of the Fourth International Workshop on Biological Control and Management of *C. odorata*, Bangalore, India.

CHASE, M. W., AND H. H. HILLIS. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. Taxon 40: 215 – 220.

CLEMENT, M., D. POSADA, AND K. A. CRANDALL. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657 – 1660.

CLEVINGER, J. A., AND J. L. PANERO. 2000. Phylogenetic analysis of *Silphium* and Subtribe Engelmanniinae (Asteraceae: Hellianthae) based on ITS and ETS sequence data. American Journal of Botany 87: 565 – 572.

COETSEE, J. 1995. Paraffienbos sluk SA in. Farmers Weekly 915: 58 – 59.

COMES, H. P., AND R. J. ABBOTT. 2001. Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). Evolution 55: 1943 – 1962.

COOK, L. M., P. S. SOLTIS, S. J. BRUNSFELD, AND D. E. SOLTIS. 1998. Multiple independent formations of *trogopogon* tetraploids (Asteraceae): evidence from RAPD markers. Molecular Ecology 7: 1293 – 1302.

CRAWFORD, D. J., S. BRAUNER, M. B. COSNER, AND T. F. STUESSY. 1993. Use of RAPD markers to document the origin of the intergeneric hybrid x*Margyracaena skottsbergii* (Rosaceae) on the Juan Fernandez islands. American Journal of Botany 80: 89 – 92.

CRAWLEY, M. J. 1983. Herbivory: the dynamics of animal-plant interactions. Blackwell Scientific Publications, London.

CRUTTWELL, R. E. 1974. Insects and mites attacking *Eupatorium odoratum* in the neotropics. 4. An annotated list of the insects and mites recorded from *Eupatorium odoratum* L., with a key to the types of damage found in Trinidad. Commonwealth Institute of Biological Control Technical Bulletin 17: 87 – 125.

CULLINGS, K. W., AND D. R. VOGLER. 1998. A 5.8S nuclear ribosomal RNA gene sequence database: applications to ecology and evolution. Molecular Ecology 7: 919 – 923.

CZECH, B., AND P. R. KRAUSMAN. 1997. Distribution and causation of species endangerment in the United States. Science 277: 1116.

DE FORESTA, H. 1991. *Chromolaena odorata*: calamité ou Chance pour l'Afrique Tropicale? (*Chromolaena odorata*: Disaster or Hope for Tropical Africa?) In Proceedings of the Third International Workshop on Biological Control and Management of *C. odorata*, Abidjan, Ivory Coast.

DE ROUW, A. 1991. The invasion of *Chromolaena ododrata* (L.) King & Robinson (*ex Eupatorium odoratum*) and competition with the native flora, in a rain forest zone, south-west Côte d'Ivore. Journal of Biogeography 18: 13 – 23.

DONNELLY, T. W. 1988. Geologic constraints on Caribbean biogeography. In J. K. Liebherr (ed.), Zoogeography of Caribbean insects. Cornell University Press, New York, USA.

DOYLE, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Systematic Botany 17: 144 – 163.

_____. 1995. The irrelevance of allele tree topologies for species delineation, and non-topological alternatives. Systematic Botany 20: 574 – 588.

_____. 1997. Within trees: genes and species, molecules and morphology. Systematic Biology 46: 537 – 553.

_____., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11 – 15.

_____., V. KANAZIN, AND R. C. SCHOEMAKER. 1996. Phylogenetic utility of Histone H3 intron sequences in the perennial relatives on the soybean (*Glycine*: Subg. *Glycine*: *Legumonosae*). American Journal of Botany 83: 434.

_____., M. MORGANTE, S. V. TINGEY, AND W. POWELL. 1998. Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine*). Molecular Biology and Evolution 15: 215 – 218.

_____., J. L. DOYLE, AND A. H. D. BROWN. 1999. Origins, colonization, and lineage recombination in a widespread perennial soybean complex. Proceedings of the National Academy of Sciences, USA 96: 10741 – 10745.

DUMOLIN-LAPÈGUE, S., B. DEMESURE, S. FINESCHI, V. LE CORRE, AND R. J. PETIT. 1997. Phylogeographic structure of white oaks throughout the European continent. Genetics 146: 1475 – 1487.

DYER, A. F. 1979. Investigating Chromosomes. Edward Arnold Ltd., London, UK.

ELLSTRAND, N. C., AND K. A. SCHIERENBECK. 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? Proceedings of the National Academy of Sciences, USA 97: 7043 – 7050.

EMSHWILLER, E., AND J. J. DOYLE. 1999. Chloroplast-expressed Glutamine Synthetase (ncpGS): potential utility for phylogenetic studies with an example from *Oxalis* (Oxalidaceae). Molecular Phylogenetics and Evolution 12: 310 – 319.

ERASMUS, C. 1990. Attempts to determine the origin of local *Chromolaena odorata*. Internal Report, ARC-PPRI, unpublished.

ERWIN, D. H. 2001. Lessons from the past: biotic recoveries from mass extinctions. Proceedings of the National Academy of Sciences, USA 98: 5399 – 5403.

EXCOFFIER, L., P. E. SMOUSE, J. M. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479 – 491.

_____., AND P. SMOUSE. 1994. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. Genetics 136: 343 – 359.

FANG, D. Q., AND M. L. ROOSE. 1997. Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theoretical and Applied Genetics 95: 408 – 417.

FERNÁNDEZ, I. A., J. F. AGUIAR, AND J. L. PANERO. 2001. A phylogenetic analysis of doronicum (Asteraceae, Senecioneae) based on morphological, nuclear ribosomal (ITS) and chloroplast *trnL*-F evidence. Molecular Phylogenetics and Evolution 20: 41 – 64.

FRYER, J. D., AND S. A. EVANS. 1968. Weed Control Handbook, Volume 1 - Principles. Blackwell Scientific, Oxford.

FU, Y.-X., AND W.-H. LI. 1999. Coalescing into the 21st century: an overview and prospects of coalescent theory. Theoretical Population Biology 56: 1 – 10.

GAUTIER, L. 1992. Taxonomy and distribution of a tropical weed: *Chromolaena odorata* (L.) R. King & H. Robinson. Candollea 47: 645 – 662.

_____. 1996. Establishment of *Chromolaena odorata* in a savanna protected from fire: an example from Lamto, Central Cote 'Ivore. In Proceedings of the Third International Workshop on Biological Control and Management of *C. odorata*, Abidjan, Ivory Coast.

GERBER, S., S. MARIETTE, R. STREIFF, C. BODENES, AND A. KREMER. 2000. Comparison of microsatellites and AFLP markers for parentage analysis. Molecular Ecology 9: 1037 – 1048.

GLASEL, J. A. 1995. Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. BioTechniques 18: 62 – 63.

GODWIN, I. D., E. A. B. AITKEN, AND L. W. SMITH. 1997. Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18: 1524 – 1528.

GOODALL, J. M. 1996. J. Goodall attends Second International Weed Control Congress, Copenhagen, Denmark. Internal Report, ARC-PPRI. Unpublished.

_____. 1997. Rapid treatment for *Chromolaena odorata* invading forest margins on sugar cane plantations in South Africa using tractor-mounted sprayers. International Journal of Pest Management 43: 85 – 86.

_____. 2000. Managing *Chromolaena odorata* (chromolaena) in subtropical grasslands in KwaZulu-Natal, South Africa. In Proceedings of the Fifth International Workshop on Biological Control and Management of *C. odorata*, South Africa, Durban.

_____., AND D. J. ERASMUS. 1996. Review of the status and integrated control of the invasive alien weed, *Chromolaena odorata*, in South Africa. Agriculture, Ecosystem and Environment 56: 151 – 164.

GRANT, W. F. 1953. A cytotaxonomic study in the genus *Eupatorium*. American Journal of Botany 40: 729 – 742.

GUPTA, M., Y.-S. CHYI, J. ROMERO-STEVENSON, AND J. L. OWEN. 1994. Amplification of DNA markers from evolutionary diverse genomes using single primers of simple-sequence repeats. Theoretical and Applied Genetics 89: 998 – 1006.

HAGEN, A. R., H. GIESE, AND C. BROCHMANN. 2001. Trans-Atlantic dispersal and phylogeography of *Cerastium arcticum* (Caryophyllaceae) inferred from RAPD and SCAR markers. American Journal of Botany 88: 103 – 112.

HARLEY, K. L. S., AND I. W. FORNO. 1992. Biological Control of Weeds, a handbook for practitioners and students. Inkata Press, Australia.

HARRISON, R. G. 1991. Molecular changes at speciation. Annual Review of Ecology and Systematics 22: 281 – 308.

HASSAN, S. 1972. Specificity and host specialization of *Puccinia chondrillina*. Annuals of Applied Biology 72: 257 – 263.

HEATON, H. J., R. WHITKUS, AND A. GÓMEZ-POMPA. 1999. Extreme ecological and phenotypic differences in the tropical tree chicozapote (*Manilkara zapota* (L.) P. Royen) are not matched by genetic divergence: a random amplified polymorphic DNA (RAPD) analysis. Molecular Ecology 8: 627 – 632.

HENDERSON, M., AND J. G. ANDERSON. 1966. Common weeds in South Africa. Memoirs of the Botanical Survey of South Africa. N° 37 Botanical Research institute, Pretoria.

HENKE, W., K. HERDEL, K. JUNG, D. SCHNORR, SA. KOENING. 1997. Betaine improves the PCR amplification of GC-rich DNA sequences. Nucleic Acids Research 25: 3957 – 3958.

HILLIS, D. M., A. LARSON, S. K. DAVIS, AND E. A. ZIMMER. 1990. Nucleic acids III: sequencing. In D. M. Hillis, and C. Moritz (eds.), Molecular systematics. Sinauer Associates, Sunderland, Massachusetts, USA.

HOCH, P. C., AND P. R. RAVEN. 1995. Introduction: perspectives in biosystematics. In Hoch, P. C., and A. G. Stephenson (eds.), Experimental and molecular approaches to plant biosystematics. Missouri Botanical Gardens, USA.

HOELZEL, A. R., AND D. R. BANKROFT. 1998. Statistical analyses of variation. In A. R. Hoelzel (ed.), Molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

_____., AND A. GREEN. 1998. PCR protocols and population analysis by direct DNA sequencing and PCR-based DNA fingerprinting. In A. R. Hoelzel (ed.), molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

HOLM, L. G., D. L. PLUCKNETT, AND J. V. PANCHO. 1977. Worlds' Worst Weeds: Distribution and Biology. University Press of Hawaii, Honolulu, Hawaii.

HTTP://WWW.BIOSCI.OHIO-STATE.EDU/~AWOLFE/ISSR/PROTOCOLS.DNA.HTML Last accessed on 25/07/2002

HUFF, D. R., R. PEAKALL, AND P. E. SMOUSE. 1993. RAPD variation within and among natural populations of outcrossing buffalo grass [*Buchloë dactyloides* (Nutt.) Engelm.]. Theoretical and Applied Genetics 86: 927 – 934.

ISMAIL, B. S., B. J. ROSMINI, AND K. SAMIAH. 1996. Factors affecting germination of Siam weed (*Chromolaena odorata* (L.) King & Robinson) seeds. Plant Protection Quarterly 11: 2 – 5.

JABLONSKI, D. 2001. Lessons from the past: evolutionary impacts of mass extinctions. Proceedings of the National Academy of Sciences, USA 98: 5393 – 5398.

JEFFREYS, A. J., V. WILSON, AND S. L. THEIN. 1985a. Hypervariable 'minisatellite' regions in human DNA. Nature 314: 67 – 73.

_____., _____., AND _____. 1985b. Individual-specific 'fingerprints' of human DNA. Nature 316: 76 – 79.

JULIEN, M. H. AND M. W. GRIFFITHS. 1999. Biological control of weeds: a world catalogue of agents and their target weeds, 4th Edition. CABI Publishing, Wallingford, UK.

KARP, A., O. SEBERG, AND M. BUIATTI. 1998. Molecular techniques in the assessment of botanical diversity. Annals of Botany 78: 143 – 149.

KÄSS, E., AND M. WINK. 1997. Phylogenetic relationships in the Papilionoideae (family Leguminosae) based on nucleotide sequences of cpDNA (rbcL) and ncDNA (ITS1 and 2). Molecular Phylogenetics and Evolution 8: 65 – 88.

KELLOGG, E. A., AND R. APPELS. 1995. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics 140: 325 – 343.

KIM, Y.-D., AND R. K. JANSEN. 1995. Phylogenetic relationships pf chloroplast DNA variatioin in the Berberidaceae. Plant Systematics and Evolution 9: 341 – 349.

_____., C. H. LEE, J. S. SHIN, Y. S. CHUNG, AND N. I. HYUNG. 1997. A simple and rapid method for isolation of high quality genomic DNA from fruit trees and conifers using PVP. Nucleic Acids Research 25: 1085 – 1086.

KING, R. M., AND H. ROBINSON. 1970. Studies in the Eupatorieae (Compositae). XXIX. The Genus *Chromolaena*. Phytologia 20: 196 - 209.

KLUGE, R. L. 1990. Prospects for the biological control of triffid weed, *Chromolaena odorata*, in Southern Africa. South African Journal of Science 86: 229 – 230.

KONISHI, N., K. WATANABE, AND K. KOSUGE. 2000. Molecular systematics of Australian *Podolepsis* (Asteraceae, Gnaphlaieae): chloroplast *matK* gene. Australian Systematic Botany 13: 709 – 727.

KROPF, M., J. W. KADEREIT, AND H. P. COMES. 2002. Late Quaternary distributional stasis in the subterranean mountain plant *Anthyllis montana* L. (Fabaceae) inferred from ITS sequences and amplified fragment length polymorphism markers. Molecular Ecology 11: 447 – 463.

KUEHN, M. M., J. E. MINOR, AND B. N. WHITE. 1999. An examination of hybridization between the cattail species *Typha latifolia* and *Typha augustifolia* using RAPD and cpDNA markers. Molecular Ecology 8: 1981 –1990.

LESLIE, A. J., AND J. R. SPOTILA. 2001. Alien plant threatens Nile crocodile (*Crocodylus niloticus*) breeding in Lake St. Lucia, South Africa. Biological Conservation 98: 347 – 355.

LIEBHERR, J. K. 1988. The Caribbean: fertile ground for zoogeography, In J. K. Liebherr (ed.), Zoogeography of Caribbean insects. Cornell University Press, New York, USA.

LIGGITT, B. 1983. The invasive alien plant *Chromolaena odorata*, with regard to its status and control in Natal. Institute of Natural Resources, Monograph 2. Pietermaritzburg.

LINDER, C. R., L. R. GOERTZEN, B. VANDEN HEUVEL, J. FRANCISCO-ORTEGA, AND R. K. JANSEN. 2000. The complete external transcribed spacer of 18S-26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. Molecular Phylogenetics and Evolution 14: 285 – 303.

LOWREY, T., C. J. QUINN, R. K. TAYLOR, R. CHAN, R. KIMBALL, AND J. C. DE NARDI. 2001. A reassessment of relationships within the *Vittadinia* group of Asteraceae (Asteraceae). American Journal of Botany 88: 1279 – 1289.

LUCAS, G. B., C. L. CAMPBELL, AND L. T. LUCAS. 1992. Biotechnology. In Introduction to plant diseases, identification and management. Chapman & Hall, New York, USA

LYNCH, M., AND B. G. MILLIGAN. 1994. Analysis of population genetic structure with RAPD markers. Molecular Ecology 3: 91 – 99.

MACDONALD, I. A. W. 1983. Alien trees, shrubs and creepers invading the Hluhluwe-Umfolozi game reserve complex in Natal. Bothalia 14: 949 – 959.

MADDISON, W. P. 1997. Gene trees in species trees. Systematic Biology 46: 523 – 536.

MANCHESTER, K. L. 1996. Use of UV methods for measurement of protein and nucleic acid concentrations. BioTechniques 20: 968 – 970.

MARKOS, S., AND B. G. BALDWIN. 2001. Higher-level relationships and major lineages of *Lessingia* (Compositeae, Asteraceae) nased on nuclear rDNA internal and external transcribed spacer (ITS and ETS) sequences. Systematic Botany 26: 168 – 183.

MARSOLAIS, J. V., J. S. PRINGLE, AND B. N. WHITE. 1993. Assessment of random amplified polymorphic DNA (RAPD) as genetic markers for determining the origin of interspecificc lilac hybrids. Taxon 42: 531 - 537.

MARTÍN, C., M. E. GONZÁLEZ-BENITO, AND J. M. IRIONDO. 1999. The use of genetic markers in the identification and characterization of three recently discovered populations of a threatened plant species. Molecular Ecology 8: S31 – S40.

MAY, B. 1998. Starch gel electrophoresis of allozymes. In A. R. Hoelzel (ed.), Molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

MAYER, M. S., AND P. S. SOLTIS. 1999. Intraspecific phylogeny analysis using ITS sequences: Insights from studies of the *Streptanthus glandulosus* complex (Cruciferae). Systematic Botany 24:47 – 61.

MBALO, B. A., AND E. T. F. WITKOWSKI. 1997. Tolerance to soil temperatures experienced during and after the passage of fire in seeds of *Acacia karoo*, *A. tortillis* and *Chromolaena odorata*: a laboratory study. South African Journal of Botany 63: 421 –425.

MCFADYEN, R. E. 1988a. History and distribution of *Chromolaena odorata* (L.) R.M. King and H. Robinson. In Proceedings of the First International Workshop on Biological Control of *C. odorata,* Bangkok, Thailand.

_____. 1988b. Ecology of *Chromolaena odorata* in the Neotropics. In Proceedings of the First International Workshop on Biological Control of *C. odorata,* Bangkok, Thailand.

_____. 1998. Biological control of weeds. Annual Review of Entomology 43: 369 – 393.

_____., AND B. SKARRATT. 1996. Potential distribution of *Chromolaena odorata* (siam weed) in Australia, Africa and Oceania. Agriculture, Ecosystems and Environment 59: 89 – 96.

MCLENACHAN, P.A., K. STOECKLER, R. C. WINKWORTH, K. MCBREEN, S. ZAUNER, AND P. J. LOCKHART. 2000. Sequence specific markers derived from AFLP gels for plant ecology and evolution studies. Molecular Ecology 9: 1899 – 1903.

MES, T. H. M. 1998. Character compatibility of molecular markers to distinguish asexual and sexual reproduction. Molecular Ecology 8: 1719 – 1728.

MILLIGAN, B. G. 1998. Total DNA isolation. In A. R. Hoelzel (ed.), Molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

MOONEY H. A., AND E. E. CLELAND. 2001. The evolutionary impact of invasive species. Proceedings of the National Academy of Science, U.S.A. 98: 5446 – 5451.

MORELL, M. K., R. PEAKALL, R. APPELS, L. R. PRESTON, AND H. L. LLOYD. 1995. DNA profiling techniques for plant variety identification. Australian Journal of Experimental Agriculture 35: 807 – 819.

MORRIS, M. J., A. R. WOOD, AND A. DEN BREEŸEN. 1999. Plant pathogens and biological control of weeds in South Africa: a review of projects ad progress during the last decade. African Entomology Memoir 1: 129 - 138.

MOUSADIK, A., AND R. J. PETIT. 1996. Chloroplast DNA phylogeny of the argan tree of Morocco. Molecular Ecology 5: 547 – 555.

MUELLER, Y. G., AND L. L. WOLFENBARGER. 1999. AFLP genotyping and fingerprinting. Tree 14: 389 – 393.

MUNIAPPAN, R. 1996. Biological control of *Chromolaena odorata*. In Hind, D. J. N.(ed.), Compositae: biology and utilization. Proceedings of the International Compositae Conference. Whitstable Litho Printers, UK.

_____., AND M. MARUTANI. 1988. Ecology and Distribution of Chromolaena odorata in Asia and the Pacific. In Proceedings of the First International Workshop on Biological Control of *C. odorata,* Bangkok, Thailand.

_____., AND J. BAMBA. 2000. Biological control of *Chromolaena odorata*: successes and failures. In Proceedings of the X International Symposium on Biological Control of Weeds, Montana, USA.

MYERS, N., AND A. H. KNOLL. 2001. The biotic crisis and the future of evolution. Proceedings of the National Academy of Sciences, USA 98: 5389 – 5392.

NESER, S. 1996. Country report - South Africa. In Proceedings of the Fourth International Workshop on Biological Control and Management of *C. odorata*, Bangalore, India .

NEWSOME, A. E., AND I. R. NOBLE. 1986. Ecological and physiological characters of invading species. In Groves, R. H., and J. J. Burdon (eds.), Ecology of biological invasions: an Australian perspective, Australian Academy of Science, Canberra, Australia.

NICHOLAS, A. 1990. Origin of the southern African noxious weed: *Chromolaena odorata*. Report, Royal Botanical Gradens, Kew, for Department of Agriculture and Water Supply: Natal region, South Africa. Unpublished.

NICKRENT, D. L. 1994. From field to film: Rapid sequencing methods for field-collected plant species. BioTechniques 16: 470 – 475.

OBERPRIELER, C. 2001. Phylogenetic relationships in *Anthemis* L. (Compositeae, Anthemideae) based in nrDNA ITS sequence variation. Taxon 50: 745 – 762.

O'HANLON, P. C., R. PEAKALL, AND D. T. BRIESE. 1999. Amplified fragment length polymorphism reveals introgression in weedy *Onopordum* thistles: hybridization and invasion. Molecular Ecology 8: 1239 – 1246.

_____., D. T. BRIESE, AND R. PEAKALL. 2000. Know your enemy: the use of molecular ecology in the *Onopordum* biological control project. In Proceedings of the X International Symposium on Biological Control of Weeds, Montana, USA.

OLCKERS, T. 1999. Introduction. African Entomology Memoir 1: 1 – 2.

OLSEN, K. M, AND B. A. SCHAAL. 1999. Evidence on the origin of cassava: phylogeography of *Manihot esculenta*. Proceedings of the National Academy of Sciences, USA 96: 5586 – 5591.

OTTO, S. P., AND J. WITTON. 2000. Polyploid incidence and evolution. Annual Review of Genetics 34: 401 – 437.

PADGET, D. J., D. H. LES, AND G. E. CROW. 1998. Evidence for the hybrid origin of N*uphar* x *Rubrodisca* (Nymphaeaceae). American Journal of Botany 85: 1468 – 1476.

PALUMBI, S. R., F. CIPRIANO, AND M. HARE. 2001. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. Evolution 55: 859 – 868.

PEAKALL, R., P. E. SMOUSE, AND D. R. HUFF. 1995. Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloë dactyliodes*. Molecular Ecology 4: 135 – 147.

POWELL, A. M., AND KING, R. M. 1969. Chromosome numbers in the Compositae: West Indian species. SIDA 3: 319 – 320.

RADFORD, I. J., P. MULLER, S. FIFFER, AND P. W. MICHAEL. 2000. Genetic relationships between Australian fireweed and South African and Madagascan populations of *Senecio madagascarensis* Poir. and closely related *Senecio* species. Australian Systematic Botany 13: 409 – 423.

RAMBUDA, T. D. 2001. Pollination and Breeeding Systems of Alien invasive plants in KwaZulu-Natal in South Africa. Masters Thesis. University of Natal, Pietermaritzburg, South Africa.

RAMSEY, J., AND SCHEMSKE, D. W. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecological Systematics 29: 467 – 501.

REEVE, H. K., D. F. WESTNEAT, AND D. C. QUELLER. 1992. Estimating average within-group relatedness from DNA fingerprints. Molecular Ecology 1: 223 – 232.

RETIEF, E. 2001. Fifth international workshop on biological control and management of *Chromolaena odorata*. SABONET News 16: 14 – 15.

RICHARDSON, T., S. CATO, J. RAMSER, G. KAHL, AND K. WEISING. 1995. Hybridization of microsatellites to RAPD: a new source of polymorphic markers. Nucleic Acids Research 23: 3798 – 3799.

RIESEBERG, L. H. 1995. The role of hybridization in evolution: old wine in new skins. American Journal of Botany 82: 944 – 953.

_____.1996. Homology among RAPD fragments in interspecific comparisons. Molecular Ecology 5: 99 – 105.

ROBINSON, H., AND R. M. KING. 1985. Comments on the generic concepts in the Eupatorieae. Taxon 34: 11 - 16.

RODER, W., S. PHENGCHANH, B. KEOBOUALAPHA, AND S. MANIPHONE. 1995. *Chromolaena odorata* in slash-and-burn rice systems of Northern Laos. Agroforestry Systems 31: 79 – 92.

ROELOFS, D., J. VAN VELZEN, P., KUPERUS, AND K. BACHMANN. 1997. Molecular evidence for an extinct parent of the tetraploid species of *Microseris accuminata* and *M. campestris* (Asteraceae, Latuceae). Molecular Systematics 6: 641 – 650.

ROGSTAD, S. H. 1992. Saturated NaCl-CTAB solution as a means of field preservation of leaves for DNA analyses. Taxon 41: 701 – 708.

ROOSE, M. L., AND L. D. GOTTLIEB. 1976. Genetic and biochemical consequences of polyploidy in *Tragopogon*. Evolution: 30: 818 – 830.

ROUX, K. H. 1995. Optimisation and troubleshooting in PCR. In Dieffenbach C. W., and G. S. Dveksler (eds.), A laboratory manual. Cold Spring Harbor Press, New York, USA.

RUTHERFORD, M. C., G. F. MIDGLEY, W. J. BOND, L. W. POWRIE, R. ROBERTS, AND J. ALLSOPP. 1999. South African country study on climate change. Plant biodiversity: vulnerability and adaptation assessment. Unpublished.

Salisbury, E. 1964. Weeds and Aliens. Collins Clear-Type Press, London.

SANDBERG, T. 2000. Chromolaena – the silent killer. Farmer's Weekly 29 September: 26 – 32.

SANG, T. 2002. Utility of low-copy nuclear gene sequences in plant phylogenetics. Critical Reviews in Biochemistry and Molecular Biology 37: 121 - 147.

SCHAAL, B. A., D. A. HAYWORTH, K. M. OLSEN, J. T. RAUSCHER, AND W. A. SMITH. 1998. Phylogeographic studies in plants: problems and prospects. Molecular Ecology 7: 465 – 474.

_____., AND K. M. OLSEN. 2000. Gene genealogies and population variation in plants. Proceedings of the National Academy of Sciences, USA 97: 7024 – 7029.

_____., AND W. J. LEVERICH. 2001. Plant population biology and systematics. Taxon 50: 679 – 695.

SCHLÖTTERER, C. 1998. Microsatellites. In A. R. Hoelzel (ed.), molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

SCHMIDT, G. J., AND E. E. SCHILLING. 2000. Phylogeny and biogeography of *Eupatorium* (Asteraceae: Eupatorieae) based on nuclear ITS sequence data. American Journal of Botany 87: 716 – 726.

SCHNEIDER, S., D. ROESSLI, AND L. EXCOFFIER. 2000. Arlequin ver. 2.000: a software for population genetics analysis. Genetics and Biometry Laboratory, University of Genva, Switzerland.

SCOTT, L. J., C. L. LANGE, G. C. GRAHAM, AND D. K. YEATES. 1998. Genetic diversity and origin of Siam weed (*Chromolaena odorata*) in Australia. Weed Technology 12: 27 – 31.

SHAW, A. 2000. Molecular phylogeography and cryptic speciation in the mosses *Mielichhoferia elongata* and *M. mielichhoferiana* (Bryaceae). Molecular Ecology 9: 595 – 608.

SHELDRICK, R. D. 1968. The control of Siam weed (*Eupatorium odoratum* Linn.). Journal of the Nigerian Institute for Oil Palm Research 5: 7 – 19.

SLAATS, J. J. P. 1991. The Use of *Chromolaena odorata* as Fallow in a Semi-permanent Cropping System in south-west Côte d'Ivoire. In Proceedings of the Third International Workshop on Biological Control and Management of *C. odorata,* Abidjan, Ivory Coast.

_____., B. H. JANSEN, AND M. WESSEL. 1998. Crop production in relation to cultural practices in the *Chromolaean odorata* fallow system in south-west Côte d'Ivoire. Netherlands Journal of Agricultural Science 46: 305 – 317.

SLADE, R. W., C. MORITZ, A. HEIDEMAN, AND P.T. HALE. 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. Molecular Ecology 2: 359 – 373.

SMOUSE, P. E. 1998. To tree or not to tree. Molecular Ecology 7: 399 – 412.

SOLTIS, P. S., AND D. E. SOLTIS. 1999. Polyploidy: recurrent formation and genome evolution. TREE 14: 348 – 352.

_____., AND _____. 2000. The role of genetic and genomic attributes in the success of polyploids. Proceedings of the National Academy of Sciences, USA 97: 7051 – 7057.

STEEL, M., AND D. PENNY. 2000. Parsimony, likelihood, and the role of models in molecular phylogenetics. Molecular Biology and Evolution 17: 839 – 850.

STEPHENS, C. AND J. P. SLATTERY. 1998. Computational resources for population analysis. In A. R. Hoelzel (ed.), Molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

ŠTORCHOVÁ, H., R. HRDLIČKOVÁ, J. CHRTEK, JR., M. TETERA, D. FITZEE AND J. FEHRER. 2000. An improved method of DNA isolation from plants collected in the field and conserved in saturated NaCl/CTAB solution. Taxon 49: 79 – 84.

SULLIVAN, V. I. 1976. Diploidy, polyploidy, and agamospermy among species of *Eupatorium* (Compositae). Canadian Journal of Botany 54: 2907 – 2917.

SUNNUCKS, P. 2000. Efficient genetic markers for population biology. Tree 15: 199 – 203.

TEMPLETON, A. R. 2001. Using phylogeographic analyses of gene trees to test species status and processes. Molecular Ecology 10: 779 – 791.

_____., K. A. CRANDALL, AND C. F. SING. 1992. A cladistic analysis of phenotype assocaitions with halotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132: 619 – 633.

_____., E. ROUTMAN, AND C. A. PHILLIPS. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum.* Genetics 140: 767 – 782.

TILMAN, D., AND C. LEHMAN. 2001. Human-caused environmental change: impacts on plant diversity and evolution. Proceedings of the National Academy of Sciences, USA 98: 5433 – 5440.

TIMBILLA, J. A., AND H. BRAIMAH. 1996. A survey of the introduction, distribution and spread of *Chromolaena odorata* on Ghana. In Proceedings of the Third International Workshop on Biological Control and Management of *C. odorata*, Abidjan, Ivory Coast.

TOOTILL, E. (ed.) 1984. The Penguin Dictionary of Botany. Penguin Books Ltd., England, UK.

TORRELL, M., N. GARCIA-JACAS, A. SUSANNA, AND J. VALLÈS. 1999. Phylogeny in *Artemisia* (Asteraceae, Anthemideae) inferred from nuclear ribosomal DNA (ITS) sequences. Taxon 48: 721 – 736.

URBATSCH, L. E., B. G. BALDWIN, AND M. J. DONOGHUE. 2000. Phylogeny of the coneflowers and relatives (Heliantheae: Asteraceae) based in nuclear rDNA ITS sequences and cpDNA restriction site data. Systematic Botany 25: 539 – 565.

VANDERPOORTEN, A., A. J. SHAW, AND B. GOFFINET. 2001. Testing controversial alignments in *Amblystegium* and related genera (Mblystegiaceae: Bryopsida). Evidence from rDNA ITS sequences. Systematic Botany 28: 470 – 479.

VAN HEUSDEN, A. W., AND B. BACHMANN. 1992a. Genetic differentiation of *Microseris pygmaea* (Asteraceae, Latuceae) studied with DNA amplification from arbitrary primers (RAPDs). Acta Botanica Neerlandica 41: 385 – 395.

_____., AND _____. 1992b. Genotype relationships in *Microseris elegans* (Asteraceae, Latuaceae) revealed by DNA amplification from arbitrary primers (RAPDs). Plant Systematics and Evolution 179: 221 – 233.

VELDKAMP, J. F. 1999. *Eupatorium catarium*, a new name for *Eupatorium clematideum* Griseb., *non* Sch.Bip. (Compositae), a South American species naturalized and spreading in south-east Asia and Queensland, Australia. Gardens' Bulletin Singapore 51: 119 – 124.

VOLKOV, R., S. KOSTISHIN, F. EHRENDORFER, D. SCHWEIZER. 1996. Molecular organization and evolution of the external transcribed rDNA spacer region in two diploid relatives of *Nicotiana tabacum* (Solanaceae). Plant Systematics and Evolution 201: 117 – 129.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407 - 4414.

VOS, W. T. 1989. The Status and Origin of *Chromolaena odorata* (L.) R. M. King and H. Robertson in Natal. Honours Thesis. University of Natal, Pietermaritzburg.

WARNER, J. P. 1998. Genomic libraries and the development of species-specific probes. In A. R. Hoelzel (ed.), Molecular genetic analysis of populations, a practical approach. Oxford University Press, New York, USA.

WATERHOUSE, B. M., AND O. ZEIMER. 2000. "On the Brink": the status of *Chromolaena odorata* in Northern Austrlia. In Proceedings of the Fifth International Workshop on Biological Control and Management of *Chromolaena odorata,* South Africa, Durban.

WEISING, K., R. G. ATKINSON, AND R. C. GARDNER. 1995b. Genomic finger-printing by microsatellite-primed PCR: a critical evaluation. PCR Methods and Applications 4: 249 – 255.

_____., NYBOM, H., WOLFF, K., AND W. MEYER. 1995a. DNA fingerprinting in plants and fungi. CRC Press, Inc. Florida, USA.

_____., AND R. C. GARDNER. 1999. A set of conserved primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42: 9 – 19.

WENDEL, J. F. 2000. Genome evolution in plants. Plant Molecular Biology 42: 225 – 249.

WHEELER, W. 2001. Homology and the Optimization of DNA sequence data. Cladistics 17: S3 - S11.

WHITE, T. J., T. BRUNS, S. LEE, AND J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis, M. A., D. H. Gelfand, J. J. Sninsky, and T. J. White (eds.) PCR Protocols: a guide to methods and applications. Academic Press, San Diego, California, USA.

WILSON, C.G. AND R. E. C. MCFADYEN. 2000. Biological control in the developing world: safety and legal issues. In Proceedings of the X International Symposium on Biological Control of Weeds, Montana, USA.

WITKOWSKI, E. T. F., AND M. WILSON. 2000. Changes in Density, biomass, seed production and soil seed banks of the non-native invasive plant, *Chromolaena odorata*, along a 15-year chronosequence. In

Proceedings of the Fifth International Workshop on the Biological Control and Management of *Chromolaena odorata*, South Africa, Durban.

_____., AND _____. 2001. Changes in density, biomass, seed production and soil seed banks of the non-native invasive plant, *Chromolaena odorata*, along a 15 year chronosequence. Plant Ecology 152: 13 – 27.

WOLFE, A. D., AND A. LISTON. 1998. Contributions of PCR-based methods to plant systematics and evolutionary biology. In Soltis, D. E., P. J. Soltis, and J. J. Doyle (eds.), Plant molecular systematics II. Kluwer Academic Publishers, Boston, Massachusetts, USA.

_____., Q.-Y. XIANG, AND S. KEPHART. 1998. Assessing hybridization in natural populations of *Penstemon* (Scrophulariaceae) using hypervariable inter-simple sequence repeat markers. Molecular Ecology 7: 1107 – 1125.

WOLFF, K., AND M. MORGAN-RICHARDS. 1998. PCR markers distinguish *Plantago major* subspecies. Theoretical and Applied Genetics 96: 287 – 293.

YADAV, A. S., AND R. S. TRIPATHI. 1981. Population dynamics of the ruderal weed *Eupatorium odoratum* and its natural regulation. Oikos 36: 355 – 361.

ZACHARIADES, C., L. W. STRATHIE-KORRÛBEL, AND R. L. KLUGE. 1999. The South African programme on the biological control of *Chromolaena odorata* (L.) King & Robinson (Asteraceae) using insects. African Entomology Memoir 1: 89 – 102.

ZIĘTKIEWICZ, E., A. RAFALSKI, AND D. LABUDA. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20: 176 – 183.

ZIMMERMANN, H. G., AND S. NESER. 1999. Trends and Prospects for Biological Control of Weeds in South Africa. African Entomology Memoir 1: 165 – 173.

APPENDICES

Appendix 1

Alternative names for *Chromolaena odorata*

Eupatorium odoratum
Agonoi (Philippines)
Archangel
Armstrong's weed
Asloke lata (Sikkim)
Bitter bush
Chimuyo (El Salvador)
Christmas rose
Co hoi (Vietnam)
Crucito (Honduras
Hemp agrimony
Hierbe de chiva (Panama)
Huluhagonoi
Jack-in-the-bush (Jamaica)
Kingsweed
Paleca
Paraffin weed / paraffienbos
Pokok Tjerman (Malaysia)
Rey del todo (Honduras)
Rompesaraguey (Dominican Republic)
Sam-solokh (India)
Santa Maria (Puerto Rico)
Sap sua (Thailand)
Siam weed
Tontrean khet (Cambodia)
Triffid weed
Ya-su'a-mop (from Liggitt, 1983).

Appendix 2

Silver Staining Protocol

1. Remove the gel carefully from the plates.

2. Soak in fixing solution (10% ethanol and 0.1% acetic acid v/v in dH_20) for 6 minutes.

3. Discard the buffer and immerse the gel in silver nitrate (0.1%w/v in dH_20).

4. Stain for 10 minutes. Save the silver nitrate.

5. Remove excess silver ions by washing the gel twice with dH_20.

6. Develop the gel in a fresh solution of 1.5% (w/v) NaOH, 0.01% (w/v) sodium borohydride and 0.15% (v/v) formaldehyde. Do not stain for more than 40 minutes.

7. Stop the development by washing for 30 seconds in fixing solution (10% ethanol and 0.1% acetic acid v/v in dH_20).

Appendix 3

Modified CTAB Plant DNA Extraction Protocol

1. Rinse *ca.* 4mm^2 piece of leaf tissue in absolute ethanol to reduce surface contamination and assist rehydration (Scott *et al.*, 1998). Allow excess ethanol to evaporate.

2. Place the tissue in a mortar with 1000µL CTAB extraction buffer and 1 drop β-mercaptoethanol

3. Decant into a labelled 1.5mL eppendorf tube. Incubate the tube in a 60°C water bath for 10 minutes. Shake the tube occasionally.

4. Add 500µL of CIA (chloroform:isoamyl alcohol, 24:1) to this tube. Close and shake vigorously.

5. Centrifuge for 1 minute at 13 000 rpm.

6. Remove 600µL of the clear aqueous phase (top layer, above the band of cell debris) and transfer to a clean, labelled 1.5mL eppendorf tube.

7. Add 400µL isopropanol, shake well, and leave on ice for at least 2 hours.

8. Centrifuge at 13 000 rpm for 10 minutes.

9. Carefully pour off all the liquid, leaving a small grey-white pellet of DNA behind.

10. Add 750µL of 75% ethanol, and gently invert the tube a few times before pouring the ethanol off. (Some argue that DNA goes into solution with 70% ethanol, from Wolfe (2002) http://www.biosci.ohio-state.edu/~awolfe/ISSR/protocols.DNA.html. Last accessed on 25/07/2002)

11. Using a small piece of paper towel, gently remove as much of the traces of ethanol as possible from the tube (without wiping the DNA pellet out), and then leave the tube open for the rest of the ethanol to evaporate away.

12. Resuspend the DNA in 200µL ultrapure tissue culture water, and shake or tap the tube to aid in the redissolving of the DNA.

Appendix 4

ISSR Protocol Variations

Unless otherwise stated, all volumes are given in µL. Concentrations are as in Chapter 2.

Extraction procedure 1. CTAb extraction, resuspended in TE.

Primers = UBC, Gel = Agarose + Ethidium Bromide (EtBr)

1) Primers 811, 812, 835, 841. 25µL, 55°C Ann. T° No ISSR product produced.

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
18.2	2.5	0.33	1	2	0.1	1

2) Repeat above at 50°C Annealing Temp

ISSR Product Produced.

3) Primers 867, 868, 891, 841
Poor ISSR product produced.
Repeat, ISSR Product Produced.

H_2O	10x	Primer	dNTP	Mg	*Taq*	Templ
17.5	2.5	1	1	2	0.1	1

4) Primers 888, 891 (Mg^{2+} range of (3 mM))

No ISSR Product produced

5) Repeat above

No ISSR product produced

Mg	H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
1	18.5	2.5	1	1	1	0.1	1
2	17.5	2.5	1	1	2	0.1	1
3	16.5	2.5	1	1	3	0.1	1
4	15.5	2.5	1	1	4	0.1	1

6) Primers 888, 891, 864, 812, 841

Partial ISSR product produced

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
17	2.5	1	1	2.5	0.1	1

7) Repeat above, with primers 841, 852

2 X *C. odorata* samples Partial ISSR product produced with primer 841, 4mM Mg

8) Repeat above, Mg^{2+} 4mM, Template volume/25µL = 2, 4, 6 ea. 1 X *C. odorata*.

Partial ISSR product produced.

9) Repeat (1), with 1X *C. odorata*. Partial ISSR product produced

10) 2 Annealing T°'s: 50°C and 55°C
Faint ISSR product produced. (best at 50°C, 4µL Templ.)

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
12½/ 10½	2.5	1	1	4	0.1	4 or 6

11) Repeat above, Mg^{2+} 1mM to 4mM. Partial ISSR product produced

12) Comparison of Ethidium bromide (EtBr) and SYBR Green visualisation.

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
17	2.5	1	1	2.5	0.1	3

Partial ISSR product produced -
Some samples worked, other samples failed. (EtBr better)

13) Repeat above to compare EtBr with silver stain (big gel)

EtBr – faint, indistinct ISSR product produced

Silver stain – Laborious procedure, and faint ISSR bands produced, only slightly migrated after 7hr

(Comment: Should have run a single, large sample (all in a single reaction volume), and run that sample in all three staining methods.)

14) Primer = 1µL/25µL reaction volume

Partial ISSR product produced (smeared bands with 1µL primer)

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
1.5/ 17.5	2.5	0.33 /1.0	1	2	0.1	1

15) Repeat above, with primer = 1, 2µL/15µL reaction volume – indistinct ISSR product produced

CTAB buffer incorrectly made (incorrect weights of contents)

16) Newly made primers (still UBC set), 841, 856, 812, 864.

No ISSR product produced

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
17.5	2.5	1	2	2	0.1	1

New UCT primers arrive.

17) Repeat (16) with new UCT primers. Oil in Corbett wells (better thermal conductivity)

No ISSR product produced

18) No ISSR product produced

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
17.7	2.5	1	1	2	0.1	1

19) Old UBC primers

No ISSR produce produced

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
17.5	2.5	1	1	2	0.1	1

20) Repeat above, with double *Taq* — No ISSR Product produced

New CTAB Buffer made up correctly

21) With new extractions

No ISSR product produced

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
16	2.5	1	1	3	0.1	1

22) Partial ISSR product produced – not all samples worked.

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
16.5	2.5	1	1	3	0.2	1

23) Repeat above, with Mg^{2+} 1mM- 4mM. 2µL/25µL reaction template.

Partial ISSR product produced – samples that worked at Mg^{2+} 1mM + 4mM did not work at Mg^{2+} 2mM + 3mM and vice versa.

24) Repeat above with newly made up primer (812), 2µL template.

No ISSR product produced

25) Same Recipe as above, primer 841, partial ISSR product produed (slightly smeared)

26) Same recipe as above, Mg^{2+} 1 to 4, primer 856, 25µL reaction volume, 2µL template. No ISSR product produced

Fresh leaf tissue extracted

27) Same recipe as above, Mg^{2+} 1 to 4, primer 841, Fresh vs. Dried material.

2µL template, 25µL reaction volume

2 of 5 fresh samples – partial ISSR produce produced

all dried samples – no ISSR produce produced

28) Repeat above, 3µL template, other(dried) samples no ISSR product produced

29) No ISSR product produced

H_2O	10x	Primer	dNTP	Mg	*Taq*	Templ
17.6/ 6.4	2.6	1	1	2 or 3	0.1	3ea

Saturated NaCl/CTAB –stored samples extracted

30) With NaCl/CTAB extractions

No ISSR product produced

H_2O	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
18.5	2.5	1	1	1	01	1
17.5				2		
16.5				3		
15.5				4		

(Decision: don't change recipe – works with other student's samples.)

31) Mg^{2+} 1mM to 4mM, 3 extractions (dried), Annealing T° = 45°C (Andi Wolfe's Program) No ISSR produce produced

32) Repeat above, 40 cycles, 1 X fresh, 2 X dried extractions. Hybaid

No ISSR product produced

33) Repeat above, Corbett No ISSR product produced

34) Compare Hybaid and Corbett. 2 X *C. odorata*, 2 X Seranne's samples. Mg^{2+} 2mM to 4mM, 50µL volumes, primer 856

Hybaid Partial ISSR product produced

Corbett Faint ISSR product produced

35) Repeat above, with 2 X *C. odorata* extractions. Hybaid, 40 cycles. 53°C annealing temp, 50 µL volumes No ISSR product produced

QIAgen extractions: C62, C63, worked well for ITS; C51, C52, C80, C85, C87 used below.

36) Use above extractions, Mg^{2+} 2mM to 3mM, primers 841, 856. No ISSR product produced

37) Repeat above, Mg^{2+} 1mM to 5mM No ISSR product produced

38) Repeat above, Mg^{2+} 1mM to 5mM, only 3 (not 5) sample. Primer 841.

No ISSR product produced

39) Repeat above, double Template, Corbett. Newly made primers (UCT).

No ISSR product produced

40) Repeat above, Hybaid (change one sample) No ISSR product produced

41) Repeat above, with less Template (1µL/25µL reaction volume)
No ISSR product produced

42) Repeat above with Betain Smeared ISSR product produced

43) Repeat above, with 2 µL template/25µL reaction volume

Partial ISSR product produced (2 visible bands)

44) Repeat one sample from above, Mg^{2+} 1mM to 5mM, with and without Betain, to run on silver stainied acrylamide gel. ISSR product produced (Mg^{2+} 2mM to 5mM)

Extractions (Procedure 3, and Procedure 4): C69, C70, C71, C88, C92, C95

Extractions (Procedure 3): C51, C52, C55, C78, C79, C80, C81, C83, C86, C87

45) Trial at Mg^{2+} 2mM, 50µL reaction volumes. 4µL template, both extraction types
No SSR product produced

46) Mg^{2+} 2mM + 3mM, 1X sample, 3 X primers, 831, 867, 811
No ISSR product produced

47) Repeat above, primer 841, 831, 867, 811, Mg^{2+} 1½mM to 2½mM. Same sample as above.
ISSR product produced! Only at primer 841, Mg^{2+} 2½mM and 2mM.

H_2O	B	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
23	10	5	2	2	3	0.25	4ea
21					4		
19					5		

48) Same sample as above + 2 more. Extractions from Protocol 3

H_2O	B	10x	Primer	dNTP	Mg^{2+}	*Taq*	Templ
23	10	5	2	2	4	0.25	4
22	10	5	2	2	5	0.25	4
21	1	5	2	2	6	0.25	4
21	15.5	5	2	2	4	0.25	4

ISSR product produced

49) Repeat above, with different samples (C69, C70, C71, C88, C92, C95)

ISSR product produced with some samples

50) Repeat above, with QIAgen extractions No ISSR product produced

Extractions (Protocol 3)

51) Repeat (50), with above extractions (newly made primer)

Partial ISSR product produced (one dominant band)

Extractions (Protocol 3)

52) Repeat (51) with above extractions Partial ISSR product produced (one dominant band)

Pooled successful product from previous successes used up on large gel run, and product was smeared.

53) Repeat of previous ISSR's to replace used up product

Partial ISSR produce produced (one dominant band)

54) Repeat of ISSR's that have not 'worked' at all - No ISSR produce produced

55) Compare UCT + UBC primers

UBC - no ISSR product produced

UCT - partial ISSR product produced (one dominant band)

56) Repeat above (new dNTP's) Partial ISSR product produced (one dominant band)

57) Comparison PVP+CTAB extractioni vs. EtOH+PVP+CTAB extraction

PVP+CTAB 1 of 2 ISSR product produced

EtOH+PVP+CTAB 2 of 2 ISSR product produced (one faint)

58) Repeat above (two extractions) with range of other samples with +ve control from (57)

PVP+CTAB 2 of 5 partial ISSR product produced (one dominant band)

EtOH+PVP+CTAB 5 of 8 partial ISSR product produced (one dominant band)

positive control no ISSR product produced

59) Range of samples (two extractions again), with PCR machine in sequencing lab. Mg^{2+} 2mM + 3mM Partial ISSR product produced (faint) – Mg^{2+} 2mM only.

60) Try Hybiad AGS gold DNA polymerase with sample from (57) that worked best.
No ISSR product produced.

61) Range of Primer concentrations: 1 & 2µL/50µL reaction volume

Partial ISSR product produced, faint bands, but no difference between primer concentrations

62) Primer concentration range: 4 & 8µL/50µL reaction volume

No ISSR product produced – gel trouble

63) Template range No ISSR product produced

64) Further template range No ISSR product produced - smeared

65) Range of samples, with a positive control (C51), primer concentration 2µL/50µL

No ISSR product produced

66) Template concentration range: 4 & 8µL/50µL reaction volume

PARTIAL SUCCESS – C52 & C132 @ Mg^{2+} 3mM. Faint ISSR product produced, and all others smeared.

67) Repeat above, with different samples — No ISSR product produced

68) Non-C. odorata sample, *Zea mays* extract; Procedure 3 extracts; and ammonium acetate-cleaned extracts — No ISSR product produced

Appendix 5

Ammonium Acetate Precipitation Protocol

- Mix 100µL DNA extract, 400µL dH_20
- Add and mix 250µL 8M Ammonium Acetate
- Add and mix 750µL isopropanol
- Place mixture on ice for 30 minutes
- Centrifuge at 30 000 rpm, for 15 minutes
- Gently pour off supernatant, and wash pellet in 70% ethanol
- Pour off ethanol and air-dry pellet
- Resuspend pellet in 100µL dH_2O

Appendix 6

Sequence Alignments

ETS Sequence Alignment

```
                                                                      60
Venezue2   TTCGAA-GCG TGTCGAGCAG TCATTTAAGG AAGCATAAAC ACTATGCAGG GACCTAACAG
Brazil_5   TTCGAATGCG TGTCGAGCAG TCATTTAAGG AAGCATAAAC ACTATGCAGG GACCTAACAG
Jamaica8   TTCGAACGCG TGTCGAGCAG TCATTTAAGG AAGCATAAAC ACTATGCAGG GACATAACAG
Mexico10   AGCAAATGCG TGTCGAGCAG TCATTTAAGG GAGCATAAAC ACTATGCAGG GACATAACAG
USA_Fl12   TTCAAATGCG TGTCGAGCAG TCATTTAAGG AAGCATAAAC ACTATGCAGG GACCTAACAG

                                                                     120
Venezue2   GGATCCCATG AGACCCATGC CCACATCAGG TACTACATCA AAGAGACCAA CCATGGCTCA
Brazil_5   GGATCCCATG AGACCCATGC CCACATCAGG TACTACATCA AAGAGACCAA CCATGGCTCA
Jamaica8   GGATCCCATG AGACCCATGC CCACATCAGG TACTACATCA AAGAGACCAA CCATGGCTCA
Mexico10   GGATCCCGTG AGACCCATGC CCACATCAGG TACTACATCA AAGAGACCAA CCATGGCTCA
USA_Fl12   GGATCCCATG AGACCCATGC CCACATCAGG TACTACATCA AAGAGACCAA CCATGGCTCA

                                                                     180
Venezue2   ATCTACCACT ACATCAACAA AAGCATTATT GATGTGGTTC AAAGAGACAG GTTTAGGGTT
Brazil_5   ATCTACCACT ACATCAACAA AAGCATTATT GATGTGGTTC AAAGAGACAG GTTTAGGGTT
Jamaica8   ATGTACCACT ACATCAACAA AAGCATTATT GATGTGGTTC AAAGAGACAG GTTTAGGGTT
Mexico10   ATGCACCACT ACATCAACAA AAGCATTATT GATGTGGTTC AAAGAGACAG GTTTAGGGTT
USA_Fl12   ATCTACCACT ACATCAACAA AAGCATTATT GATGTGGTTC AAAGAGACAG GTTTAGGGTT

                                                                     240
Venezue2   CGTCGGTCGT CATAAACATT GACGAAAGAC GAAAGATATT TAGATTAGCC AGATCAATAC
Brazil_5   CGTCGGTCGT CATAAACATT GACGAAAGAC GAAAGATATT TAGATTAGCC AGATCAATAC
Jamaica8   CGTCGGTCGT CATAAACATT GACGATAGAC GAAAGATATT TAGATTAGCC AGATCAATAC
Mexico10   CGTCGGTCGT CATAAACATT GACGAAAGAC AAAAGAAATT TAGATTAGCC AGATCAATAC
USA_Fl12   CGTCGGTCGT CATAAACATT GACGATAGAC GAAAGATATT TAGATTAGCC AGATCAATAC

                                                                     300
Venezue2   CGTATGTTAG GTACGCAACA CATGAAACCC ACAATTTGTT CAGGCATTTT CAGCCCTTAC
Brazil_5   CGTATGTTAG GTACGCAACA CATGAAACCC ACAATTTGTT CAGGCATTTT CAGCCCTTAC
Jamaica8   CGTATGTTAG GTACGCAACA CATGAAACCC ACAATTTGTT CAGGCATTTT CAGCCCTTAC
Mexico10   CTTATGTTAG GTACGCAACA CATGAAACCC ACAATTTGTT CAGGCATTTT CAGCCCTTAC
USA_Fl12   CGTATGTTAG GTACGCAACA CATGAAACCC ACAATTTGTT CAGGCATTTT CAGCCCTTAC

                                                                     360
Venezue2   CAGACAAACA ACCAGGCAAG TTAGGTGGAA GTTGTTGCAC AAGCAAAGCG AGCCGACCAC
Brazil_5   CAGACAAACA ACCAGGCAAG TTAGGTGGAA GTTGTTGCAC AAGCAAAGCG AGCCGACCAC
Jamaica8   CAGACAAACA ACCAGGCAAG TTAGGTGGAA GTTGTTGCAC AAGCAAAGCG AGCCGACCAC
Mexico10   CGGACAAACA ACCAGGCAAG TTAGGTGGAA GTTGTTGCAC AAGCAAAGCG AGCCGACCAC
USA_Fl12   CAGACAAACA ACCAGGCAAG TTAGGTGGAA GTTGTTGCAC AAGCAAAGCG AGCCGACCAC

                                                                     420
Venezue2   CGGTAACAAA CCAAAGACCA CTCACGCACC TTTACGGTAT GCATTCCCAG AAGCCAGAAA
Brazil_5   CGGTAACAAA CCAAAGACCA CTCACGCACC TTTACGGTAT GCATTCCCAG AAGCCAGAAA
Jamaica8   CGGTAACAAA CCAAAGACCA CTCACGCACC TTTACGGTAT GCATTCCCAG AAGCCAGAAA
Mexico10   CGGTAACAAA CCAAAGACCA CTCACGCACC TTTACGGTAT GCATTCCCAG AAGCCAGAAA
USA_Fl12   CGGTAACAAA CCAAAGACCA CTCACGCACC TTTACGGTAT GCATTCCCAG AAGCCAGAAA

                                                                     480
Venezue2   ACCATTAGCC ACCCCCTAAG CAGTTAAGCA AAGTGGGAAA CAAACAGCCA TTAAGACCCC
Brazil_5   ACCATTAGCC ACCCCCTAAG CAGTTAAGCA AAGTGGGAAA CAAACAGCCA TTAAGACCCC
Jamaica8   ACCATTAGCC ACCCCCTAAG CATTTAAGCA AAGTGGGAAA CAAACAGCCA TTAAGACCCC
Mexico10   ACCATTCGCC ACCCCCTAAG CATTTAAGCA ACGTGGGAAA CAAACAGCCA TTAAGACCCC
USA_Fl12   ACCATTAGCC ACCCCCTAAG CAGTTAAGCA AAGTGGGAAA CAAACAGCCA TTAAGACCCC
```

```
                                                                      540
Venezue2   TAGAATGCCC CGAACAATGC GACCTGGACG TGCCTCAGCA TCCCGTTATG TGGCCCCGTA
Brazil_5   TAGAATGCCC CGAACAATGC GACCTGGACG TGCCTCAGCA TCCCGTTATG TGGCCCCGTA
Jamaica8   TAGAATGCCC CGAACAATGC GACCTGGACG TGCCTCAGCA TCCCGTTATG TGGCCCCGTA
Mexico10   TAGAATGCCC CGAACAATGC GACCTGAACG TGCCTCAGCA TCCCGTTATG TGGCCCCGTA
USA_Fl12   TAGAATGCCC CGAACAATGC GACCTGGACG TGCCTCAGCA TCCCGTTATG TGGCCCCGTA

                                                                      600
Venezue2   CTGGTTTAGC AAAGACCATT TGGACGTCGA GCACATAAAC TCACCAATCA AGATGCAATA
Brazil_5   CTGGTTTAGC AAAGACCATT TGGACGTCGA GCACATAAAC TTACCAATCA AGATGCAATG
Jamaica8   CTGGTTTAGC AAAGACCATT TGGACGTCTA GCACATAAAC TCACCAATCA AGATGCAATG
Mexico10   CTGGTTTAGC AAAGACCATT TGGACGTCTA GCACATAAAC TCACCAATCA AGATGCAAT?
USA_Fl12   CTGGTTTAGC AAAGACCATT TGGACGTCGA GCACATAAAC TCACCAATCA AGATGCAAT-


Venezue2   AGAGTGCCAT ATACCGCAAC TAACCAACCA CACTTTCAAA GGCATCC
Brazil_5   AGAGTGCCAT ATACCGCAAC TAACCAACC- CA???????? ???????
Jamaica8   AGAGTGCCAT ATACCGCAAC TAACCCACCG CACTTTCAAA GGCATCC
Mexico10   ?????????? ?????????? ?????????? ?????????? ???????
USA_Fl12   AGAGTGCCAT ATACCGCAAC TAACCAACCG CACTTTCAAA GGCATCC
```

ITS Sequence Alignment

```
                                                                                                                    60
AAdenoph    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAATCCTG CGTAGCAGAA CAACTTGTGA
CBorinqu    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
CSqualid    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
PClemati    ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
CCollina    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAAAA CAACCCGTGA
CMacroce    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CAAAGCAAAA CAACCCGTGA
ARiparia    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAATCCTG CGTAGCAGAA CAACTTGTGA
SA-PSh28    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
AusITS1A    ?????????? ?????????? ?????????? TCGAACCCTG CATGGCAGAA CAACCCGTGA
AusITS1B    ?????????? ?????????? ?????????? TCGAACCCTG CATGGCAGAA CAACCCGTGA
Austra84    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Austra81    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Brazil-5    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Brazil51    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
CostRic6    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Guatema7    TCCGYAGGTG AACCTRCSGA AGKATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Guatem52    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
India-16    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic58    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic59    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic62    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic63    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic69    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic71    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaica9    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic57    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic54    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaic56    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Jamaica8    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Mexico10    ?????????? ?????????? ?????????? ?????????? ????????AA CAACCCGTGA
Maurit15    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
NSumat38    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
PRico100    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
PRico102    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-PSJ50    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-PSh32    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-PSJ45    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-PEd23    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Hlu42    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Dbn13    ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-Pal92    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Pal88    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Tza95    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Tza96    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Esh26    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Tza99    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-PSJ48    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Tza18    ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-Hlu44    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Tza97    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-Mtu27    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
SA-PSh36    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Trinid55    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Trinid11    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Thaila11    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
USA-Fl12    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Venezue2    ?????????? ?????????? ?????????? ?????????? ?????????? ???????TGA
Venezue1    ?????????? ?????????? ?????????? ?????????? ???GGCAGAA CAACCCGTGA
Venezue3    TCCGTAGGTG AACCTGCGGA AGGATC-TTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
Venezue4    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
WAfric17    TCCGTAGGTG AACCTGCGGA AGGATCATTG TCGAACCCTG CATGGCAGAA CAACCCGTGA
```

```
                                                                                                                    120
AAdenoph   ACATGTAACA AC--AAGATG TCTTTGCGGT GGATTGATGC CTATTGTCTT CAAACCTCGT
CBorinqu   ACGTGTATCA ACAAAAGATG GCTTGGCGGG CAATCGAAGC TTTCTGT-TT CTAGCCTCGT
CSqualid   ACGTGTATCA AC--AAGACG GCTTGGCGGG CTGTCGAAGC TTTCTGT-TT CGAGCCTCGT
PClemati   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
CCollina   ACGTGTATCA AC-AAAGACG GCTTGGCGGG YWTT-GAYGC TTTATGT-TT CAAGCCTCGT
CMacroce   ACGTGTACCA AC--AAGATA GCCTTATGGG TATTTGATTC TTTTGCT-TT CAAACCCTGT
ARiparia   ACATGTAACA AC--AARATG TCTTTGCGGT GGATTGATGC CTATTGTYTT CAAACCTCGT
SA-PSh28   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
AusITS1A   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
AusITS1B   ACGTGTATCA AC--AAGACG GCTTGGCGGT CCATCGAAGC TATGTGT-TT CATGCCTCGT
Austra84   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Austra81   ACGTGTATCA AC--AAGACG GCTTGGCGGT CCATCGAAGC TATGTGT-TT CATGCCTCGT
Brazil-5   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Brazil51   ACGTGTATCA AC--AAGACG GCTTGGCGGT CCATCGAAGC TATGTGT-TT CATGCCTCGT
CostRic6   ACGTGTATCA AC--AAGAAG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Guatema7   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Guatem52   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
India-16   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic58   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic59   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic62   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic63   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic69   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic71   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaica9   ACGTGTATCA AC--AAGAMG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic57   ACGTGTATCA AC--AAGAMG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic54   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaic56   ACGTGTATCA AC--AAGAMG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Jamaica8   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Mexico10   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Maurit15   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
NSumat38   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
PRico100   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
PRico102   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-PSJ50   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-PSh32   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-PSJ45   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-PEd23   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Hlu42   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Dbn13   ?????????? ?????AGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Pal92   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Pal88   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Tza95   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Tza96   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Esh26   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Tza99   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-PSJ48   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Tza18   ???????TCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Hlu44   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Tza97   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-Mtu27   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
SA-PSh36   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Trinid55   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Trinid11   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Thaila11   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
USA-Fl12   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Venezue2   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Venezue1   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Venezue3   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
Venezue4   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
WAfric17   ACGTGTATCA AC--AAGACG GCTTGGCGGG CCATCGAAGC TATGTGT-TT CATGCCTCGT
```

```
                                                                      180
AAdenoph   GAAGCCATGT TTACGTGTGT TTTTGGTTTC TTGTCTTGGT CACTCATGAA CATCACGTTG
CBorinqu   TAAGC-CCGT CGACGTGCGT CCGGGGTGCC TCCTTTTGG- CACCTCCGG- CGTCACGTTG
CSqualid   TAAGC-CTGT CGACGTGSGT CCGGGGTGCC TCCTTTTGG- CGCCTCCGG- CGTCACGTTG
PClemati   ?????????? ?????????? ?????????? ?????????? ?????????? ?????????G
CCollina   GAAGC-CTGT CGACGTGTGT TTGYGGAGTC TCTTTTTGG- CACTATTGG- CATCACGTTG
CMacroce   GAGGC-CTGT TGACGTGTGT TTGTGGTGTC TCTTTTTG-T CACYG-TAGG CATCACGTTG
ARiparia   GAAGCCATGT TTACGTGTGT TTTTGGTTTC TTGTTTTGGT CACTCATGAA CATCACGTTG
SA-PSh28   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
AusITS1A   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
AusITS1B   TAAGC-TTGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Austra84   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Austra81   TAAGC-TTGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Brazil-5   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Brazil51   TAAGC-TKGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCAYGTTG
CostRic6   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCAYGTTG
Guatema7   TAAGC-ATGT CGATGTGTGT STGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Guatem52   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
India-16   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic58   TAAGC-ATGT CGAAGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic59   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic62   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic63   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic69   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic71   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaica9   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic57   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic54   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaic56   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Jamaica8   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Mexico10   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Maurit15   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
NSumat38   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
PRico100   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
PRico102   TAAGC-ATGT CGAAGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-PSJ50   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-PSh32   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-PSJ45   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-PEd23   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Hlu42   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Dbn13   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Pal92   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Pal88   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Tza95   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Tza96   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Esh26   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Tza99   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-PSJ48   TAAGC-AAGA CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Tza18   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Hlu44   TAAGC-ATGT CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Tza97   TAAGC-AAGA CGATGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-Mtu27   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
SA-PSh36   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Trinid55   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Trinid11   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Thaila11   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
USA-Fl12   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Venezue2   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Venezue1   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Venezue3   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
Venezue4   TAAGC-ATGT CGAYGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
WAfric17   TAAGC-ATGT CGACGTGTGT CTGGGGTGCT TCTTTTTGG- CACTCCTTT- CGTCACGTTG
```

						240
AAdenoph	ACGCAACAAC	AAC-CCC-GG	CACGACACGT	GCCAGGGAAA	ACTAAACTTA	AAAGGGGGTG
CBorinqu	ACGCAATAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCGAACAKA	AGAGTGCCCG
CSqualid	ACGCAATAAC	AACCCCC-GG	CACGGAACGT	GCCAAGGAAA	AGCGAACATA	AGAGTGCCCT
PClemati	ACCCATTAAC	AACCCCC-GG	CACGGAACGT	GCCAAGGAAA	ACCGAACATA	AGAGCGCCCT
CCollina	ACGCAACAAC	AACCCCC-GG	CACAGYACGT	GCCAAGGAAA	ACCAAACTTA	AGAGGGCCCG
CMacroce	ACGCAACAA-	--CCCCC-GG	CACAGCACGT	GCCAAGGAAA	ACCAAASTTT	AGAGTTCCAT
ARiparia	ACGCAACAAC	AACCCC--GG	CACAACACGT	GCCAAGGAAA	AYTAAACTTA	AAAGGGGGTG
SA-PSh28	ACCCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
AusITS1A	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
AusITS1B	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Austra84	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Austra81	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Brazil-5	ACCCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Brazil51	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
CostRic6	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Guatema7	ACCCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Guatem52	ACCCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
India-16	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic58	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic59	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic62	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic63	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic69	ACSCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic71	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaica9	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic57	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic54	ACSCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaic56	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Jamaica8	ACCCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Mexico10	ACCCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Maurit15	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
NSumat38	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
PRico100	ACCCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
PRico102	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-PSJ50	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-PSh32	ACSCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-PSJ45	ACSCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-PEd23	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Hlu42	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Dbn13	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGYA	AGAGTGCCCG
SA-Pal92	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Pal88	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Tza95	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Tza96	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Esh26	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Tza99	ACSCAACAAC	AACCCC--GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-PSJ48	ACCCAACAAC	AACCCCC-GG	CACGGCATGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Tza18	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Hlu44	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Tza97	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-Mtu27	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
SA-PSh36	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Trinid55	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Trinid11	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Thaila11	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
USA-Fl12	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Venezue2	ACGCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Venezue1	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Venezue3	ACSCAACAAC	AACCCCCSGG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
Venezue4	ACSCAACAAC	AACCCCC-GG	CACGGCAYGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG
WAfric17	ACGCAACAAC	AACCCCC-GG	CACGGCACGT	GCCAAGGAAA	ACCAAACGTA	AGAGTGCCCG

						300
AAdenoph	TG-CCATGAC	ACCC--TCTT	A---------	-------GTG	GA----TTTT	TTATAA-T-C
CBorinqu	TGTGGTGAKG	CCCC-GTAKM	TGGTGGCMTC	GTTGCATTCG	GMTTGCTTT-	TG-TAAAAKC
CSqualid	TGTGGTGATG	CCCC-GTATT	TGGTGGCCTC	GTTGCATGCG	GC-TGCTTTT	TT-TAAAT-C
PClemati	AGTGGCGATG	CCCC-GTATT	TGGTGGCGTC	GTTGSGTGCG	GC-CGCTTTT	--ATAAAT-C
CCollina	TGCCATGATG	CCCC-GTATT	AGGTGTGTTC	ATTGTATGTG	GC-TTCTTT-	-G-TAA-T-C
CMacroce	GTTCCATGAS	TATCCGTT-T	AGGTGTGTTC	ATTGTACGTG	GC-TTCTCTT	-G-TAATG--
ARiparia	TG-CCATGAC	MCCC--TT-T	A---------	-------GTG	GA----TTTT	TTATAA-T-C
SA-PSh28	TGTCGTGATG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
AusITS1A	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
AusITS1B	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GT-TGCTTT-	-G-TAAAT-C
Austra84	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Austra81	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GT-TGCTTT-	-G-TAAAT-C
Brazil-5	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Brazil51	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GT-TGCTTT-	-G-TAAAT-C
CostRic6	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Guatema7	TGTCGTGATG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAA-C
Guatem52	TGTCGTGATG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAA-C
India-16	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic58	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic59	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic62	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic63	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic69	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic71	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaica9	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic57	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic54	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaic56	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Jamaica8	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Mexico10	TGTCGTGATG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAA-C
Maurit15	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
NSumat38	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
PRico100	TGTCGTGATG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
PRico102	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-PSJ50	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-PSh32	TGTCGTGATG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGCG	GC-TGCTTT-	-G-TAAAT-C
SA-PSJ45	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-PEd23	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Hlu42	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Dbn13	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-WAAAW-C
SA-Pal92	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Pal88	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Tza95	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Tza96	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Esh26	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Tza99	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-PSJ48	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Tza18	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Hlu44	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Tza97	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-Mtu27	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
SA-PSh36	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCAKGTG	GC-TGCTTT-	-G-TAAAT-C
Trinid55	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Trinid11	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Thaila11	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
USA-Fl12	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Venezue2	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Venezue1	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Venezue3	TGTCGTGAWG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
Venezue4	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C
WAfric17	TGTCGTGAAG	CCCC-GTATT	TGGTGGCCTC	ATTGCATGTG	GC-TGCTTT-	-G-TAAAT-C

```
                                                                                                                  360
AAdenoph   TTAAACGACT CTCGGCAACG GATATCTTGG CTCATGCATC GATGAAGAAC GTAGCAAA-T
CBorinqu   AKAAACGACT CTCGGCAACG GATA-CTCGG ATCACGCATC GATGAAGAAC GTAGCAAAAT
CSqualid   ATAAACGACT CTCGGCAACG GATATCTGGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
PClemati   ATAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
CCollina   ATAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
CMacroce   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
ARiparia   TTWAACGACT CTCGGCAACG GATATCTTGG CTCATGCATC GATGAAGAAC GTAGCAAAAT
SA-PSh28   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
AusITS1A   T????????? ?????????? ?????????? ?????????? ?????????? ??????????
AusITS1B   T????????? ?????????? ?????????? ?????????? ?????????? ??????????
Austra84   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Austra81   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Brazil-5   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Brazil51   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
CostRic6   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Guatema7   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Guatem52   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
India-16   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic58   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic59   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic62   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic63   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic69   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC -ATGAAGAAC GTAGCAAAAT
Jamaic71   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaica9   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic57   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC -ATGAAGAAC GTAGCAAAAT
Jamaic54   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaic56   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Jamaica8   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Mexico10   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Maurit15   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
NSumat38   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
PRico100   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
PRico102   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GAWGAAGAAC GTAGCAAAAT
SA-PSJ50   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGC--AAT
SA-PSh32   TTAAWCGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGA-C GTAGCAAAAT
SA-PSJ45   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-PEd23   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Hlu42   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Dbn13   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Pal92   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Pal88   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Tza95   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Tza96   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Esh26   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Tza99   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-PSJ48   TTAAACGACT CTCGGCAACG GATATCTCGG -TCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Tza18   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-Hlu44   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGC-AAAT
SA-Tza97   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GAAGAAGAAC GTAGCAAAAT
SA-Mtu27   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
SA-PSh36   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Trinid55   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Trinid11   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Thaila11   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
USA-Fl12   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Venezue2   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Venezue1   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Venezue3   TTAAACGACT ATCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
Venezue4   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
WAfric17   TTAAACGACT CTCGGCAACG GATATCTCGG CTCACGCATC GATGAAGAAC GTAGCAAAAT
```

```
                                                                                                                  420
AAdenoph   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTT-G AACGCAAGTT
CBorinqu   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
CSqualid   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
PClemati   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
CCollina   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
CMacroce   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
ARiparia   GCGATACTTG GT-GTGAATT GCAGAATCCC --GAACCAT- CGAGTTTTTG AACGCAAGTT
SA-PSh28   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
AusITS1A   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
AusITS1B   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Austra84   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Austra81   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Brazil-5   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Brazil51   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
CostRic6   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Guatema7   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Guatem52   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCATT -GAGTTTTTG AACGCAAGTT
India-16   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic58   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic59   GCGATACTTG GTAGTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic62   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic63   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic69   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic71   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaica9   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic57   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic54   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaic56   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Jamaica8   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Mexico10   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Maurit15   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
NSumat38   GCGATACTTG GT-GTGAATT GCASAATCCC GTGAACCAT- CKAGTTTTTG AACGCAAGTT
PRico100   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
PRico102   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-PSJ50   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-PSh32   RCGATACTTG KT-GTGAATT GCASAATCCC GTGAACCAT- CKAGTTTTTG AACGCAAGTT
SA-PSJ45   GCGATACTTG KT-GTGAATT GCAGAATCCM GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-PEd23   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Hlu42   GCGATMCTTG KT-GTGAATT GCASAATCCC GTGAACCAT- CKAGTTTTTG AACGCAAGTT
SA-Dbn13   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Pal92   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Pal88   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Tza95   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Tza96   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Esh26   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Tza99   GCAACCCT-G GTAGTGAAWT GCAGAAWCCC GAGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-PSJ48   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCATT CGAGTTTTTG AACGCAAGTT
SA-Tza18   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Hlu44   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Tza97   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-Mtu27   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
SA-PSh36   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Trinid55   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Trinid11   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Thaila11   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
USA-Fl12   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Venezue2   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Venezue1   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Venezue3   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
Venezue4   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
WAfric17   GCGATACTTG GT-GTGAATT GCAGAATCCC GTGAACCAT- CGAGTTTTTG AACGCAAGTT
```

```
                                                                   480
AAdenoph   GCGCCT-AAG CCACTAGGT- TGAGGGCACG -TCTGCCTGG GCGTCACACA TCATGTTGCC
CBorinqu   GCGCCTGAAG CCTCCCGGCC TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
CSqualid   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
PClemati   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
CCollina   GCGCCCGAAG CCACCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
CMacroce   GCGCCTGAAG CCGCTCGGGT TGAAGGCACG ATCTGCCTGG GCGTCACGCG TAATGTCGCC
ARiparia   GCGCCTGAAG CCACTAGGA- TGAGGGAACG -TCTGCCTGG GCGTCACACA TCATGTTGCC
SA-PSh28   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
AusITS1A   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
AusITS1B   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Austra84   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
Austra81   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
Brazil-5   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Brazil51   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
CostRic6   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Guatema7   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Guatem52   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
India-16   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
Jamaic58   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Jamaic59   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Jamaic62   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Jamaic63   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Jamaic69   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Jamaic71   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Jamaica9   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Jamaic57   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Jamaic54   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Jamaic56   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Jamaica8   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Mexico10   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Maurit15   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
NSumat38   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
PRico100   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
PRico102   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-PSJ50   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-PSh32   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTMACC
SA-PSJ45   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TSTGCCTGG GCGTCACGCA TCACGTMACC
SA-PEd23   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Hlu42   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TSTGCCTGG GCGTCACGCA TCACGTMACC
SA-Dbn13   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Pal92   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Pal88   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Tza95   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Tza96   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Esh26   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Tza99   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCYGCCTGG GCGTCACGCA TCACGTCRCC
SA-PSJ48   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
SA-Tza18   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Hlu44   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Tza97   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-Mtu27   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
SA-PSh36   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Trinid55   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Trinid11   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
Thaila11   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTMACC
USA-Fl12   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
Venezue2   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Venezue1   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
Venezue3   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCGCC
Venezue4   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCRCC
WAfric17   GCGCCTGAAG CCTCCCGGC- TGAGGGCACG -TCTGCCTGG GCGTCACGCA TCACGTCACC
```

						540
AAdenoph	TACA-CACA-	CATCTTTGCT	TGAAATGTGC	GGTATGTGGG	CGGAGCCTGG	TCTCCTGTGC
CBorinqu	CGCATCAAA-	CGTCCCTGCT	TGGATTGTGG	TGTATGCGGG	CGGAGACTGG	TCTCCTGTGC
CSqualid	CGCATCAAAA	CGTCCTTGTT	TGGATTGTGA	TGTATGCGGG	CGGAGACTGG	TCTCCTGTGC
PClemati	TGCTACAAA-	CGTCCTTGCT	TGGATTGTGA	TGTATGCGGG	CGGAGACTGG	TCTCCCGTGC
CCollina	CACATCAAA-	CATCCTTGTT	TGGATTGTGT	TGAATGTGTG	CGGAGACTGG	TCTCCTGTGC
CMacroce	CACATCAAA-	CATCCTCGTT	TGGATCGTGT	TGCTTGTGGG	-GGAGACTGG	TCTCCCGTGC
ARiparia	TACA-CAACA	CATCTTCGCT	TGAAATGTGC	ACCATGTGGG	CGGAGCCTGG	TCTCCTGTGC
SA-PSh28	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
AusITS1A	??????????	??????????	??????????	??????????	??????????	??????????
AusITS1B	??????????	??????????	??????????	??????????	??????????	??????????
Austra84	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Austra81	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Brazil-5	TACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
Brazil51	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
CostRic6	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Guatema7	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
Guatem52	CACATCAAA-	CGTCCTTGTT	TGGATTGCGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
India-16	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Jamaic58	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaic59	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
Jamaic62	YACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaic63	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaic69	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
Jamaic71	YACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaica9	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaic57	YACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaic54	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaic56	YACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Jamaica8	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
Mexico10	CACATCAAA-	CGTCCTTGTT	TGGATTGCGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
Maurit15	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
NSumat38	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
PRico100	TACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
PRico102	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGKGGG	YGGAGACTGG	TCTCCTGTGC
SA-PSJ50	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGKGGG	YGGAGACTGG	TCTCCTGTGC
SA-PSh32	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
SA-PSJ45	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
SA-PEd23	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
SA-Hlu42	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Dbn13	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Pal92	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Pal88	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Tza95	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Tza96	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGAATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Esh26	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Tza99	CACATCAAA-	CGGCCTTGTT	TGGATTGTGA	GGAAAGGGGG	CGGAGACTGG	TCTCCCGGGC
SA-PSJ48	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	TCTCCTGTGC
SA-Tza18	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-Hlu44	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGMATGTGG?	??????????	??????????
SA-Tza97	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	GGAATGTGGG	CGGAGACTGG	YCTCCTGTGC
SA-Mtu27	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
SA-PSh36	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Trinid55	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGAATGTGGG	YGGAGACTGG	TCTCCTGTGC
Trinid11	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Thaila11	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
USA-Fl12	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Venezue2	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Venezue1	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC
Venezue3	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
Venezue4	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	KGWATGTGGG	YGGAGACTGG	TCTCCTGTGC
WAfric17	CACATCAAA-	CGTCCTTGTT	TGGATTGTGA	TGTATGTGGG	TGGAGACTGG	TCTCCTGTGC

```
                                                                          600
AAdenoph   CTAGGGTACG GTTGGCCTAA AAA-GAAGTC TTGTT-AAGA GCGATAGGGA CGCATGACTG
CBorinqu   CCATGGTGCG GTTGGCCTAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
CSqualid   CCATGGTGCG GTTGGCCTAA ATACG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
PClemati   CCATGGCGCG GCTGGCCTAA ATACG-AGTC CGGTT-AGGA G------TGA CGCACGACTT
CCollina   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA A------AGA CACACGACTG
CMacroce   CTATGGTGCG GTTGGCCCAA ATACG-AGCT CCCGGTAAGA G------TGA CGCACGACTA
ARiparia   CTAGGGTACG GTTGGCCCAA AAAT--AGTC TGGGT-AAGA GGGATAGGGA CGCATGACTG
SA-PSh28   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
AusITS1A   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
AusITS1B   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Austra84   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
Austra81   CCATGGTGCG GCAGGCCCAA ATTTG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Brazil-5   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Brazil51   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
CostRic6   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Guatema7   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Guatem52   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
India-16   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
Jamaic58   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Jamaic59   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Jamaic62   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Jamaic63   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Jamaic69   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Jamaic71   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Jamaica9   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Jamaic57   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Jamaic54   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Jamaic56   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Jamaica8   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Mexico10   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
Maurit15   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
NSumat38   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
PRico100   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
PRico102   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-PSJ50   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-PSh32   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
SA-PSJ45   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
SA-PEd23   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Hlu42   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Dbn13   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Pal92   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Pal88   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Tza95   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
SA-Tza96   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Esh26   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Tza99   CCATGGTGCG GTTGGCCCAA ATATG-AGCC CG???????? ?????????? ??????????
SA-PSJ48   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
SA-Tza18   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-Hlu44   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-Tza97   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-AAGA G------TGA CGCACGACTG
SA-Mtu27   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
SA-PSh36   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Trinid55   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Trinid11   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
Thaila11   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
USA-Fl12   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
Venezue2   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Venezue1   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Venezue3   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
Venezue4   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-WAGA G------TGA CGCACGACTG
WAfric17   CCATGGTGCG GTTGGCCCAA ATATG-AGTC CGCTT-TAGA G------TGA CGCACGACTG
```

```
                                                                                                                    660
AAdenoph   GTGGTGGTTG ATTGGACGGT CGTCCGGTGT TGTGTGCTTT AATTCTTGAT -GGTAAAGAC
CBorinqu   GTGGTGGTTG ACTACACGGT CGTCTCGTGT CGTGTGTTAC GATTCTTAAA -GGGAAAAGC
CSqualid   GTGGTGGTTG ACTACCAGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGGAAACGC
PClemati   CTGGTGGTTG ACTACGCGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA CGG-AAACTC
CCollina   GTGGTGGTTG ATTACACAGT CGTCTCGTGT CGTGTGTTTC GATTCTTGAA CGGAAAACTC
CMacroce   GTGGTGGTTG ATTACACAGT CGTCTCGTGT CGTGTGTAAT GACTCTTAAA -GG--AAAGC
ARiparia   GTGGTGGTTG ATTGGACGGT CGTCCGGTGT T-TGTGCTTT AATTCTTGAT -GGTAAAGAC
SA-PSh28   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
AusITS1A   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
AusITS1B   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Austra84   GTGGTGGTTG ACTTAACGGW CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Austra81   GTGGTGGWTG ACTTAACGGA CGTCTCGTGT CGTGKGWTTC GATTCTYAAA -GGYGAAAAC
Brazil-5   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Brazil51   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
CostRic6   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Guatema7   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Guatem52   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
India-16   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic58   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic59   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic62   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic63   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic69   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic71   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaica9   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic57   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic54   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaic56   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Jamaica8   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Mexico10   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Maurit15   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
NSumat38   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
PRico100   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
PRico102   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-PSJ50   GTGGTGGKTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-PSh32   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-PSJ45   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-PEd23   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Hlu42   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Dbn13   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Pal92   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Pal88   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Tza95   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Tza96   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Esh26   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Tza99   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-PSJ48   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Tza18   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Hlu44   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-Tza97   GTGGTGGGTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-Mtu27   GTGGTGGTTG ACTAAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
SA-PSh36   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Trinid55   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Trinid11   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Thaila11   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
USA-Fl12   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Venezue2   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Venezue1   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Venezue3   GTGGTAGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
Venezue4   GTGGTGGTTG ACTWAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
WAfric17   GTGGTGGTTG ACTTAACGGT CGTCTCGTGT CGTGTGTTTC GATTCTTAAA -GGTGAAAAC
```

```
                                                                                                                  720
AAdenoph   TCTTCCAATA CCCTGATGTG T-TGTCTTAT GATAACTCTT TGATGGCGAC CCCAGGTCAG
CBorinqu   TCTTGAAGTA CCCTGACGCG C-CGTCTTGT GACGGCCCTT CGATCGCGAC CCCAGGTCAG
CSqualid   TCTTGAACTA CCCTGATGCG C-CGACTTGC GACGGCCTTT CGATCGCGAC CCCAGGTCAG
PClemati   TCTTRAACTA CCCTGATGCG C-CGACTTGK RACGGCCCTT CRATCGCGAC CCCAGGTCAG
CCollina   T-TAAAAGTA CCCYGTTGTG --GGCCTTGA GATAGCCCTT CAATCGCGAC CCCAGGTCAG
CMacroce   TCTTAAAGTA CCCTCATGTG CCTGCCTTTT GGTAGCTCTT CGATCGCGAC CCCAGGTCAG
ARiparia   TCTTACAATA CCCTAATGTG T-TGTCTTAC GATAACTCTT TGATGGCGAC CCCAGGTCAG
SA-PSh28   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
AusITS1A   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
AusITS1B   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
Austra84   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Austra81   TCTTCAAGTA CCCTGATGCG C-TGTCTTGA AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Brazil-5   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Brazil51   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
CostRic6   TCTTCASGTA CCCTGATGCG C-TGTCTTGT AACAGCTCTT CKATCGCGAC CCCAGGTCAG
Guatema7   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Guatem52   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
India-16   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic58   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic59   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic62   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic63   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic69   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic71   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaica9   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic57   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACRGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic54   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaic56   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Jamaica8   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Mexico10   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Maurit15   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
NSumat38   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
PRico100   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
PRico102   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-PSJ50   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-PSh32   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-PSJ45   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-PEd23   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Hlu42   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Dbn13   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGG????? ?????????? ??????????
SA-Pal92   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Pal88   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Tza95   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Tza96   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Esh26   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Tza99   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-PSJ48   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Tza18   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Hlu44   ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
SA-Tza97   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-Mtu27   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
SA-PSh36   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Trinid55   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Trinid11   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Thaila11   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
USA-Fl12   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Venezue2   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
Venezue1   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACRGCTCTT CGATCGCGAC CCCAGGTCAG
Venezue3   TCTTCA-GTA CCCTGATGCG C-TGTCTTGT AACRGCTCTT CGATCGCGAC CCCAGGTCAG
Venezue4   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CGATCGCGAC CCCAGGTCAG
WAfric17   TCTTCAAGTA CCCTGATGCG C-TGTCTTGT AACGGCTCTT CAATCGCGAC CCCAGGTCAG
```

```
AAdenoph   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
CBorinqu   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
CSqualid   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
PClemati   --GCGGGACT ACCCGCTGAG TTWAAGCATA TCAATAA-GC GGAGG
CCollina   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
CMacroce   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
ARiparia   --GCGGGACT ACCCGCTGAG CTTAAGCATA TCAATAA-GC GGAGG
SA-PSh28   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
AusITS1A   ?????????? ?????????? ?????????? ?????????? ?????
AusITS1B   ?????????? ?????????? ?????????? ?????????? ?????
Austra84   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Austra81   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Brazil-5   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAACGC GGAGG
Brazil51   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
CostRic6   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Guatema7   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Guatem52   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
India-16   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATA???? ?????
Jamaic58   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic59   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic62   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic63   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic69   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic71   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaica9   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic57   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic54   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaic56   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Jamaica8   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Mexico10   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATA???? ?????
Maurit15   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
NSumat38   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
PRico100   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
PRico102   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-PSJ50   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-PSh32   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-PSJ45   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-PEd23   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Hlu42   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Dbn13   ?????????? ?????????? ?????????? ?????????? ?????
SA-Pal92   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Pal88   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Tza95   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Tza96   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Esh26   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Tza99   ?????????? ?????????? ?????????? ?????????? ?????
SA-PSJ48   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Tza18   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Hlu44   ?????????? ?????????? ?????????? ?????????? ?????
SA-Tza97   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-Mtu27   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
SA-PSh36   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Trinid55   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Trinid11   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATA???? ?????
Thaila11   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
USA-Fl12   CTGCGGGACT ACCCGCTGAG CTTAAGCATA TGAATA???? ?????
Venezue2   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATA???? ?????
Venezue1   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Venezue3   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
Venezue4   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
WAfric17   --GCGGGACT ACCCGCTGAG TTTAAGCATA TCAATAA-GC GGAGG
```

Appendix 7

Jukes-Cantor Distance Matrix

		1	2	3	4	5	6	7	8
1	USA Fl12	-							
2	Brazil 5	0.01670	-						
3	Guatema7	0.01673	0.00411	-					
4	Venezue4	0.00276	0.00138	0.00276	-				
5	CostRic6	0.01109	0.01377	0.01378	0.00410	-			
6	Jamaica9	0.00276	0.00138	0.00139	0.00000	0.00273	-		
7	SA Esh26	0.00416	0.00138	0.00138	0.00000	0.00548	0.00000	-	
8	Trinid11	0.00276	0.01388	0.01392	0.00000	0.00831	0.00000	0.00140	-
9	SA PSh28	0.01669	0.00136	0.00274	0.00136	0.01376	0.00000	0.00000	0.01389
10	Jamaica8	0.00831	0.00138	0.00137	0.00000	0.00961	0.00000	0.00000	0.00553
11	Maurit15	0.00274	0.01375	0.01376	0.00000	0.00823	0.00000	0.00138	0.00000
12	India 16	0.00276	0.01388	0.01392	0.00000	0.00831	0.00000	0.00140	0.00000
13	WAfric17	0.00412	0.01513	0.01515	0.00137	0.00961	0.00137	0.00275	0.00138
14	SA Mtu27	0.00416	0.00138	0.00138	0.00000	0.00548	0.00000	0.00000	0.00140
15	SA PEd23	0.00836	0.00138	0.00137	0.00000	0.00962	0.00000	0.00000	0.00556
16	SA Tza18	0.00305	0.00148	0.00157	0.00000	0.00445	0.00000	0.00000	0.00000
17	SA Dbn13	0.00000	0.00163	0.00000	0.00000	0.00493	0.00000	0.00000	0.00000
18	Mexico10	0.01935	0.00578	0.00144	0.00443	0.01616	0.00293	0.00293	0.01629
19	Venezue2	0.00302	0.00442	0.00598	0.00000	0.00589	0.00000	0.00000	0.00000
20	Venezue1	0.00292	0.01024	0.01026	0.00000	0.00282	0.00000	0.00147	0.00000
21	SA PSh36	0.00276	0.00138	0.00139	0.00000	0.00410	0.00000	0.00000	0.00000
22	Thaila11	0.00274	0.01238	0.01240	0.00000	0.00823	0.00000	0.00000	0.00000
23	SA PSh32	0.01531	0.00410	0.00550	0.00275	0.01520	0.00137	0.00137	0.01248
24	SA PSJ45	0.01108	0.00273	0.00411	0.00000	0.01099	0.00000	0.00000	0.00828
25	NSumat38	0.00274	0.01375	0.01377	0.00000	0.00823	0.00000	0.00138	0.00000
26	SA Hlu44	0.00335	0.00411	0.00405	0.00000	0.00332	0.00000	0.00000	0.00335
27	SA Hlu42	0.00277	0.00274	0.00276	0.00000	0.00548	0.00000	0.00000	0.00000
28	Jamaic54	0.00549	0.00136	0.00277	0.00000	0.00547	0.00000	0.00000	0.00273
29	Jamaic57	0.00697	0.00000	0.00138	0.00000	0.00136	0.00000	0.00000	0.00418
30	Venezue3	0.00696	0.00412	0.00414	0.00273	0.00411	0.00273	0.00273	0.00419
31	Jamaic56	0.00836	0.00000	0.00137	0.00000	0.00411	0.00000	0.00000	0.00557
32	Jamaic58	0.00414	0.00274	0.00276	0.00136	0.00548	0.00136	0.00136	0.00138
33	Jamaic59	0.01255	0.00137	0.00136	0.00000	0.00962	0.00000	0.00000	0.00975
34	Jamaic62	0.00836	0.00000	0.00137	0.00000	0.00548	0.00000	0.00000	0.00557
35	Jamaic63	0.00413	0.00137	0.00139	0.00000	0.00410	0.00000	0.00000	0.00137
36	CBorinqu	0.07401	0.07176	0.06734	0.06469	0.07179	0.06329	0.06327	0.07093
37	SA Tza97	0.01679	0.00686	0.00687	0.00555	0.01517	0.00557	0.00554	0.01394
38	PRico100	0.01809	0.00000	0.00411	0.00274	0.01515	0.00138	0.00138	0.01528
39	PRico102	0.00414	0.00275	0.00414	0.00136	0.00549	0.00136	0.00136	0.00138
40	SA PSJ50	0.00836	0.00137	0.00139	0.00000	0.00828	0.00000	0.00000	0.00558
41	Brazil51	0.00965	0.01516	0.01518	0.00414	0.00960	0.00414	0.00552	0.00688
42	Trinid55	0.00416	0.00138	0.00138	0.00000	0.00548	0.00000	0.00000	0.00140
43	Jamaic69	0.01108	0.00136	0.00274	0.00000	0.00961	0.00000	0.00000	0.00828
44	Jamaic71	0.00417	0.00000	0.00138	0.00000	0.00548	0.00000	0.00000	0.00140
45	CSqualid	0.07981	0.07735	0.07604	0.07030	0.07897	0.06883	0.06876	0.07674
46	PClemati	0.09497	0.08178	0.08368	0.08040	0.08946	0.07856	0.07844	0.09079
47	CCollina	0.09896	0.09336	0.09188	0.08762	0.09494	0.08616	0.08604	0.09579
48	Guatem52	0.01813	0.00545	0.00136	0.00414	0.01515	0.00277	0.00276	0.01530
49	SA Pal92	0.00416	0.00138	0.00138	0.00000	0.00548	0.00000	0.00000	0.00140
50	SA Pal88	0.00554	0.00137	0.00138	0.00000	0.00548	0.00000	0.00000	0.00277
51	SA Tza95	0.00416	0.00138	0.00138	0.00000	0.00410	0.00000	0.00000	0.00140
52	SA Tza96	0.00693	0.00137	0.00137	0.00000	0.00684	0.00000	0.00000	0.00416
53	Austra81	0.01671	0.02224	0.02223	0.01113	0.01660	0.01116	0.01253	0.01390
54	SA PSJ48	0.01815	0.00410	0.00551	0.00278	0.01515	0.00279	0.00277	0.01532
55	CMacroce	0.16511	0.16360	0.16208	0.15554	0.15855	0.15557	0.15554	0.16158
56	Austra84	0.00274	0.01375	0.01377	0.00000	0.00823	0.00000	0.00138	0.00000
57	ARiparia	0.19105	0.18715	0.19109	0.18369	0.18716	0.18373	0.18388	0.19096
58	SA Tza99	0.02882	0.02085	0.02282	0.01920	0.02862	0.01922	0.01917	0.02887
59	AAdenoph	0.18405	0.17678	0.18061	0.17176	0.17673	0.17181	0.17194	0.18034
60	AusITS1A	0.00000	0.01130	0.01539	0.00000	0.00780	0.00000	0.00000	0.00000
61	AusITS1B	0.01535	0.01948	0.02357	0.01185	0.01566	0.01194	0.01191	0.01533

Jukes-Cantor distance matrix (continued)

		9	10	11	12	13	14	15	16
9	SA PSh28	-							
10	Jamaica8	0.00000	-						
11	Maurit15	0.01375	0.00548	-					
12	India 16	0.01389	0.00553	0.00000	-				
13	WAfric17	0.01514	0.00687	0.00136	0.00138	-			
14	SA Mtu27	0.00000	0.00000	0.00138	0.00140	0.00275	-		
15	SA PEd23	0.00000	0.00000	0.00548	0.00556	0.00687	0.00000	-	
16	SA Tza18	0.00000	0.00000	0.00000	0.00000	0.00155	0.00000	0.00000	-
17	SA Dbn13	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
18	Mexico10	0.00434	0.00291	0.01620	0.01629	0.01771	0.00293	0.00291	0.00309
19	Venezue2	0.00440	0.00150	0.00000	0.00000	0.00154	0.00000	0.00000	0.00000
20	Venezue1	0.01023	0.00584	0.00000	0.00000	0.00146	0.00147	0.00584	0.00000
21	SA PSh36	0.00000	0.00000	0.00000	0.00000	0.00137	0.00000	0.00000	0.00000
22	Thaila11	0.01238	0.00411	0.00000	0.00000	0.00136	0.00000	0.00411	0.00000
23	SA PSh32	0.00274	0.00136	0.01238	0.01248	0.01378	0.00137	0.00136	0.00151
24	SA PSJ45	0.00136	0.00000	0.00822	0.00828	0.00960	0.00000	0.00000	0.00000
25	NSumat38	0.01376	0.00548	0.00000	0.00000	0.00136	0.00138	0.00549	0.00000
26	SA Hlu44	0.00207	0.00000	0.00332	0.00335	0.00332	0.00000	0.00207	0.00000
27	SA Hlu42	0.00136	0.00000	0.00000	0.00000	0.00136	0.00000	0.00000	0.00000
28	Jamaic54	0.00000	0.00000	0.00273	0.00273	0.00411	0.00000	0.00000	0.00000
29	Jamaic57	0.00000	0.00000	0.00413	0.00418	0.00550	0.00000	0.00000	0.00000
30	Venezue3	0.00273	0.00273	0.00412	0.00419	0.00551	0.00273	0.00273	0.00307
31	Jamaic56	0.00000	0.00000	0.00550	0.00557	0.00687	0.00000	0.00000	0.00000
32	Jamaic58	0.00137	0.00136	0.00136	0.00138	0.00274	0.00136	0.00136	0.00136
33	Jamaic59	0.00000	0.00000	0.00962	0.00975	0.01100	0.00000	0.00000	0.00000
34	Jamaic62	0.00000	0.00000	0.00550	0.00557	0.00687	0.00000	0.00000	0.00000
35	Jamaic63	0.00000	0.00000	0.00137	0.00137	0.00275	0.00000	0.00000	0.00000
36	CBorinqu	0.07021	0.06611	0.07017	0.07093	0.07166	0.06327	0.06613	0.06998
37	SA Tza97	0.00549	0.00551	0.01376	0.01394	0.01516	0.00554	0.00551	0.00624
38	PRico100	0.00136	0.00138	0.01514	0.01528	0.01653	0.00138	0.00138	0.00149
39	PRico102	0.00274	0.00136	0.00136	0.00138	0.00274	0.00136	0.00136	0.00136
40	SA PSJ50	0.00000	0.00000	0.00551	0.00558	0.00691	0.00000	0.00000	0.00000
41	Brazil51	0.01515	0.00964	0.00686	0.00688	0.00823	0.00552	0.00963	0.00446
42	Trinid55	0.00000	0.00000	0.00138	0.00140	0.00275	0.00000	0.00000	0.00000
43	Jamaic69	0.00000	0.00000	0.00822	0.00828	0.00961	0.00000	0.00000	0.00000
44	Jamaic71	0.00000	0.00000	0.00138	0.00140	0.00275	0.00000	0.00000	0.00000
45	CSqualid	0.07584	0.07164	0.07584	0.07674	0.07735	0.06876	0.07167	0.07597
46	PClemati	0.08374	0.08011	0.08964	0.09079	0.08971	0.07844	0.08212	0.07882
47	CCollina	0.09179	0.08744	0.09487	0.09579	0.09332	0.08604	0.08739	0.09365
48	Guatem52	0.00409	0.00274	0.01514	0.01530	0.01653	0.00276	0.00274	0.00308
49	SA Pal92	0.00000	0.00000	0.00138	0.00140	0.00275	0.00000	0.00000	0.00000
50	SA Pal88	0.00000	0.00000	0.00275	0.00277	0.00413	0.00000	0.00000	0.00000
51	SA Tza95	0.00000	0.00000	0.00138	0.00140	0.00275	0.00000	0.00000	0.00000
52	SA Tza96	0.00000	0.00000	0.00412	0.00416	0.00550	0.00000	0.00000	0.00000
53	Austra81	0.02222	0.01668	0.01383	0.01390	0.01522	0.01253	0.01667	0.01229
54	SA PSJ48	0.00274	0.00275	0.01515	0.01532	0.01654	0.00277	0.00275	0.00314
55	CMacroce	0.16347	0.15726	0.16002	0.16158	0.16172	0.15554	0.15715	0.16702
56	Austra84	0.01376	0.00548	0.00000	0.00000	0.00136	0.00138	0.00548	0.00000
57	ARiparia	0.19061	0.18551	0.18886	0.19096	0.19066	0.18388	0.18393	0.19340
58	SA Tza99	0.01899	0.01911	0.02864	0.02887	0.02864	0.01917	0.01908	0.02195
59	AAdenoph	0.18017	0.17525	0.17840	0.18034	0.18017	0.17194	0.17350	0.18025
60	AusITS1A	0.01498	0.00368	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
61	AusITS1B	0.02328	0.01564	0.01536	0.01533	0.01536	0.01191	0.01195	0.01383

Jukes-Cantor distance matrix (continued)

		17	18	19	20	21	22	23	24
17	SA Dbn13	-							
18	Mexico10	0.00175	-						
19	Venezue2	0.00000	0.00757	-					
20	Venezue1	0.00000	0.01179	0.00000	-				
21	SA PSh36	0.00000	0.00294	0.00000	0.00000	-			
22	Thaila11	0.00000	0.01476	0.00000	0.00000	0.00000	-		
23	SA PSh32	0.00173	0.00733	0.00443	0.01026	0.00137	0.01101	-	
24	SA PSJ45	0.00000	0.00583	0.00145	0.00585	0.00000	0.00686	0.00136	-
25	NSumat38	0.00000	0.01621	0.00000	0.00000	0.00000	0.00000	0.01238	0.00822
26	SA Hlu44	0.00000	0.00657	0.00217	0.00136	0.00000	0.00332	0.00405	0.00000
27	SA Hlu42	0.00000	0.00443	0.00000	0.00000	0.00000	0.00000	0.00137	0.00000
28	Jamaic54	0.00000	0.00438	0.00146	0.00145	0.00000	0.00273	0.00137	0.00000
29	Jamaic57	0.00000	0.00293	0.00146	0.00000	0.00000	0.00413	0.00275	0.00136
30	Venezue3	0.00355	0.00592	0.00305	0.00290	0.00273	0.00412	0.00552	0.00412
31	Jamaic56	0.00000	0.00292	0.00146	0.00147	0.00000	0.00413	0.00275	0.00136
32	Jamaic58	0.00136	0.00443	0.00153	0.00136	0.00136	0.00136	0.00276	0.00137
33	Jamaic59	0.00000	0.00289	0.00145	0.00585	0.00000	0.00825	0.00273	0.00136
34	Jamaic62	0.00000	0.00292	0.00146	0.00147	0.00000	0.00413	0.00275	0.00136
35	Jamaic63	0.00000	0.00293	0.00145	0.00000	0.00000	0.00137	0.00137	0.00000
36	CBorinqu	0.07568	0.07446	0.07105	0.07169	0.06333	0.06873	0.06897	0.06888
37	SA Tza97	0.00714	0.00892	0.00767	0.01175	0.00557	0.01242	0.00693	0.00550
38	PRico100	0.00163	0.00579	0.00586	0.01170	0.00138	0.01377	0.00411	0.00273
39	PRico102	0.00136	0.00594	0.00153	0.00136	0.00136	0.00136	0.00416	0.00138
40	SA PSJ50	0.00000	0.00294	0.00145	0.00439	0.00000	0.00414	0.00137	0.00000
41	Brazil51	0.00491	0.01766	0.00584	0.00435	0.00415	0.00687	0.01379	0.00962
42	Trinid55	0.00000	0.00293	0.00000	0.00147	0.00000	0.00138	0.00137	0.00000
43	Jamaic69	0.00000	0.00434	0.00145	0.00585	0.00000	0.00686	0.00136	0.00000
44	Jamaic71	0.00000	0.00293	0.00000	0.00147	0.00000	0.00000	0.00138	0.00000
45	CSqualid	0.08233	0.08395	0.07734	0.07779	0.06881	0.07441	0.07465	0.07450
46	PClemati	0.08793	0.08696	0.08371	0.08609	0.07863	0.08772	0.08411	0.08402
47	CCollina	0.10224	0.09931	0.09454	0.09627	0.08617	0.09346	0.09381	0.09043
48	Guatem52	0.00174	0.00000	0.00751	0.01170	0.00277	0.01378	0.00687	0.00546
49	SA Pal92	0.00000	0.00293	0.00000	0.00147	0.00000	0.00000	0.00137	0.00000
50	SA Pal88	0.00000	0.00292	0.00145	0.00147	0.00000	0.00137	0.00137	0.00000
51	SA Tza95	0.00000	0.00293	0.00000	0.00000	0.00000	0.00138	0.00137	0.00000
52	SA Tza96	0.00000	0.00291	0.00145	0.00292	0.00000	0.00275	0.00137	0.00000
53	Austra81	0.01390	0.02532	0.01358	0.01177	0.01117	0.01384	0.02089	0.01664
54	SA PSJ48	0.00362	0.00737	0.00609	0.01173	0.00279	0.01380	0.00553	0.00410
55	CMacroce	0.18764	0.17220	0.16552	0.16420	0.15556	0.15843	0.16589	0.16049
56	Austra84	0.00000	0.01621	0.00000	0.00000	0.00000	0.00000	0.01239	0.00822
57	ARiparia	0.21209	0.20450	0.19361	0.19343	0.18384	0.18867	0.19332	0.18896
58	SA Tza99	0.02276	0.02712	0.02368	0.02908	0.01925	0.02860	0.02098	0.01905
59	AAdenoph	0.19952	0.19270	0.18050	0.18213	0.17191	0.17823	0.18275	0.17871
60	AusITS1A	0.00000	0.01654	0.00000	0.00000	0.00000	0.00000	0.01521	0.00759
61	AusITS1B	0.01417	0.02521	0.01721	0.01233	0.01192	0.01536	0.02340	0.01577

Jukes-Cantor distance matrix (continued)

	25	26	27	28	29	30	31	32
25 NSumat38	-							
26 SA Hlu44	0.00332	-						
27 SA Hlu42	0.00000	0.00000	-					
28 Jamaic54	0.00273	0.00000	0.00000	-				
29 Jamaic57	0.00413	0.00000	0.00136	0.00000	-			
30 Venezue3	0.00412	0.00197	0.00413	0.00273	0.00273	-		
31 Jamaic56	0.00550	0.00000	0.00136	0.00000	0.00000	0.00273	-	
32 Jamaic58	0.00136	0.00195	0.00136	0.00138	0.00138	0.00410	0.00137	-
33 Jamaic59	0.00962	0.00206	0.00136	0.00000	0.00000	0.00273	0.00000	0.00137
34 Jamaic62	0.00550	0.00000	0.00136	0.00000	0.00000	0.00273	0.00000	0.00137
35 Jamaic63	0.00137	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00138
36 CBorinqu	0.07020	0.07274	0.06487	0.06608	0.06484	0.06640	0.06468	0.06468
37 SA Tza97	0.01378	0.00798	0.00558	0.00557	0.00555	0.00836	0.00551	0.00694
38 PRico100	0.01515	0.00412	0.00274	0.00136	0.00000	0.00412	0.00000	0.00274
39 PRico102	0.00136	0.00197	0.00136	0.00139	0.00139	0.00410	0.00138	0.00000
40 SA PSJ50	0.00551	0.00209	0.00000	0.00000	0.00000	0.00277	0.00000	0.00139
41 Brazil51	0.00686	0.00754	0.00414	0.00551	0.00551	0.00829	0.00688	0.00552
42 Trinid55	0.00138	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00136
43 Jamaic69	0.00822	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00137
44 Jamaic71	0.00138	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00136
45 CSqualid	0.07590	0.06825	0.07037	0.07161	0.07042	0.07189	0.07022	0.07019
46 PClemati	0.08972	0.07380	0.08061	0.08050	0.07671	0.08242	0.07644	0.07853
47 CCollina	0.09492	0.07660	0.08779	0.08914	0.08626	0.08816	0.08753	0.08763
48 Guatem52	0.01515	0.00600	0.00413	0.00413	0.00277	0.00552	0.00276	0.00415
49 SA Pal92	0.00138	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00136
50 SA Pal88	0.00275	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00138
51 SA Tza95	0.00138	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00136
52 SA Tza96	0.00412	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00137
53 Austra81	0.01383	0.00733	0.01116	0.01254	0.01252	0.01539	0.01389	0.01255
54 SA PSJ48	0.01515	0.00600	0.00415	0.00279	0.00277	0.00556	0.00275	0.00416
55 CMacroce	0.16013	0.16998	0.15759	0.15878	0.15575	0.15789	0.15708	0.15719
56 Austra84	0.00000	0.00332	0.00000	0.00273	0.00413	0.00412	0.00550	0.00136
57 ARiparia	0.18906	0.18211	0.18585	0.18727	0.18229	0.18637	0.18378	0.18547
58 SA Tza99	0.02868	0.01650	0.01742	0.01914	0.01921	0.02110	0.01914	0.02106
59 AAdenoph	0.17858	0.18117	0.17381	0.17536	0.17045	0.17436	0.17193	0.17356
60 AusITS1A	0.00000	0.00379	0.00000	0.00766	0.00388	0.00000	0.00387	0.00392
61 AusITS1B	0.01536	0.01196	0.01188	0.01576	0.01191	0.01190	0.01193	0.01582

Jukes-Cantor distance matrix (continued)

	33	34	35	36	37	38	39	40
33 Jamaic59	-							
34 Jamaic62	0.00000	-						
35 Jamaic63	0.00000	0.00000	-					
36 CBorinqu	0.06744	0.06470	0.06470	-				
37 SA Tza97	0.00547	0.00551	0.00557	0.07205	-			
38 PRico100	0.00137	0.00000	0.00137	0.07171	0.00688	-		
39 PRico102	0.00138	0.00138	0.00139	0.06605	0.00555	0.00412	-	
40 SA PSJ50	0.00000	0.00000	0.00000	0.06779	0.00419	0.00137	0.00140	-
41 Brazil51	0.01100	0.00688	0.00414	0.07327	0.01517	0.01654	0.00555	0.00827
42 Trinid55	0.00000	0.00000	0.00000	0.06480	0.00554	0.00138	0.00136	0.00000
43 Jamaic69	0.00000	0.00000	0.00000	0.06744	0.00550	0.00136	0.00138	0.00000
44 Jamaic71	0.00000	0.00000	0.00000	0.06339	0.00554	0.00000	0.00136	0.00000
45 CSqualid	0.07302	0.07018	0.07021	0.03936	0.07920	0.07734	0.07170	0.07333
46 PClemati	0.08203	0.07641	0.07859	0.06335	0.08618	0.08171	0.08000	0.08215
47 CCollina	0.08876	0.08750	0.08764	0.10639	0.09508	0.09334	0.08921	0.08947
48 Guatem52	0.00273	0.00276	0.00277	0.06893	0.00827	0.00546	0.00553	0.00277
49 SA Pal92	0.00000	0.00000	0.00000	0.06327	0.00554	0.00138	0.00136	0.00000
50 SA Pal88	0.00000	0.00000	0.00000	0.06466	0.00554	0.00137	0.00139	0.00000
51 SA Tza95	0.00000	0.00000	0.00000	0.06327	0.00554	0.00138	0.00136	0.00000
52 SA Tza96	0.00000	0.00000	0.00000	0.06615	0.00551	0.00137	0.00138	0.00000
53 Austra81	0.01803	0.01389	0.01115	0.08090	0.02221	0.02364	0.01256	0.01536
54 SA PSJ48	0.00273	0.00275	0.00279	0.07043	0.00275	0.00411	0.00417	0.00279
55 CMacroce	0.15853	0.15713	0.15721	0.17936	0.16579	0.16520	0.15734	0.15954
56 Austra84	0.00962	0.00550	0.00137	0.07019	0.01376	0.01515	0.00136	0.00551
57 ARiparia	0.18534	0.18388	0.18555	0.21868	0.19315	0.18878	0.18557	0.18678
58 SA Tza99	0.01898	0.01913	0.01918	0.09182	0.02463	0.02089	0.01919	0.01724
59 AAdenoph	0.17495	0.17201	0.17363	0.20573	0.18259	0.17838	0.17365	0.17626
60 AusITS1A	0.00382	0.00387	0.00388	0.11210	0.01179	0.01503	0.00389	0.00383
61 AusITS1B	0.01195	0.01189	0.01188	0.12275	0.01993	0.02328	0.01576	0.01192

Jukes-Cantor distance matrix (continued)

	41	42	43	44	45	46	47	48
41 Brazil51	-							
42 Trinid55	0.00552	-						
43 Jamaic69	0.00962	0.00000	-					
44 Jamaic71	0.00552	0.00000	0.00000	-				
45 CSqualid	0.07898	0.07030	0.07306	0.06885	-			
46 PClemati	0.08994	0.08040	0.08214	0.07651	0.04773	-		
47 CCollina	0.09811	0.08604	0.08887	0.08608	0.10587	0.11839	-	
48 Guatem52	0.01654	0.00276	0.00409	0.00277	0.07756	0.08593	0.09347	-
49 SA Pal92	0.00552	0.00000	0.00000	0.00000	0.06876	0.07844	0.08604	0.00276
50 SA Pal88	0.00551	0.00000	0.00000	0.00000	0.07018	0.07850	0.08755	0.00276
51 SA Tza95	0.00414	0.00000	0.00000	0.00000	0.06876	0.07844	0.08604	0.00276
52 SA Tza96	0.00687	0.00000	0.00000	0.00000	0.07168	0.08036	0.08746	0.00275
53 Austra81	0.00691	0.01253	0.01664	0.01254	0.08511	0.09602	0.10298	0.02362
54 SA PSJ48	0.01654	0.00277	0.00273	0.00277	0.07912	0.08403	0.09508	0.00686
55 CMacroce	0.16380	0.15566	0.15879	0.15568	0.18570	0.18435	0.12642	0.16374
56 Austra84	0.00687	0.00138	0.00822	0.00138	0.07586	0.08977	0.09492	0.01515
57 ARiparia	0.18915	0.18386	0.18734	0.18235	0.22252	0.21582	0.17356	0.19261
58 SA Tza99	0.03250	0.01918	0.01904	0.01920	0.08905	0.10989	0.09510	0.02469
59 AAdenoph	0.17871	0.17359	0.17713	0.17047	0.21239	0.19655	0.17419	0.18204
60 AusITS1A	0.01541	0.00000	0.00758	0.00000	0.11891	0.18265	0.12815	0.01532
61 AusITS1B	0.00000	0.01191	0.01577	0.01190	0.13219	0.19453	0.14256	0.02347

Jukes-Cantor distance matrix (continued)

	49	50	51	52	53	54	55	56
49 SA Pal92	-							
50 SA Pal88	0.00000	-						
51 SA Tza95	0.00000	0.00000	-					
52 SA Tza96	0.00000	0.00000	0.00000	-				
53 Austra81	0.01253	0.01253	0.01113	0.01389	-			
54 SA PSJ48	0.00277	0.00277	0.00277	0.00275	0.02366	-		
55 CMacroce	0.15554	0.15713	0.15554	0.15716	0.17282	0.16558	-	
56 Austra84	0.00138	0.00275	0.00138	0.00412	0.01245	0.01515	0.16011	-
57 ARiparia	0.18388	0.18561	0.18369	0.18566	0.19697	0.19270	0.19195	0.18894
58 SA Tza99	0.01917	0.01914	0.01917	0.01911	0.03758	0.02286	0.18949	0.02859
59 AAdenoph	0.17194	0.17369	0.17176	0.17541	0.18828	0.18216	0.19866	0.17846
60 AusITS1A	0.00000	0.00387	0.00000	0.00386	0.01478	0.01940	0.28941	0.00000
61 AusITS1B	0.01191	0.01189	0.01191	0.01191	0.00000	0.02749	0.30884	0.01537

Jukes-Cantor distance matrix (continued)

	57	58	59	60	61
57 ARiparia	-				
58 SA Tza99	0.20621	-			
59 AAdenoph	0.02869	0.20898	-		
60 AusITS1A	0.27216	0.00686	0.28430	-	
61 AusITS1B	0.27875	0.01583	0.29026	0.01567	-

School of Environmental Science
University College Northampton
Park Campus
Boughton Green Road
Northampton
NN2 7AL
UK
6 December 2002

Registrars Division
Rhodes University
Prince Alfred Road
Grahamstown
6139
South Africa

To Whom It May Concern:

Letter of Masters Thesis Submission

Herewith please find my Masters Thesis entitled "An Assessment of the Genetic Diversity and Origin of the Invasive Weed *Chromolaena odorata* (L.) King and Robinson in South Africa". The corrections to this thesis, as suggested by the two external examiners are as below. The corrections requested and made have been commented on in a paragraph-by-paragraph fashion, according to the paragraphs annotated in the attached documents.

Examiner: Dr. McFadyen

Paragraph 1: The number of heading levels in Chapter 1 were decreased, and the section "History of *Chromolaena odorata*'s spread" was moved to a later and more suitable position in the chapter. A few sentences were added, explaining that the locality names in Vos's thesis were unclear, and could not be further defined through further investigation. The spelling and grammatical changes as commented in Dr. McFadyen's copy of the thesis were corrected.

Paragraph 2: Further weed texts were sought and referenced, and all references to "Cruttwell McFadyen" were changed to "McFadyen", however the reference "Cruttwell, 1974" was not changed, as the original paper has no reference to the name "McFadyen" on or in it.

Paragraph 3: Reference was found, and added to the thesis regarding the ability of *Chromolaena odorata* to replace grasses and prevent regeneration of forests.

Paragraph 4: No changes are required in this paragraph.

Paragraph 5: On suggestion from my supervisor, Dr. N. P. Barker, no changes have been made based on the comments in this paragraph. Dr. McFadyen mentioned that high CI values indicate that the data is probably saturated, however Dr. Barker advises that the high CI values indicate that the data does in fact contain high levels of informativeness and a lack of conflicting signal.

Paragraph 6: The analysis with *Chromolaena* borinquensis as the outgroup was done as suggested, but there was almost no resolution in the resulting tree; there were several two-taxon clades, and the remainder of the tree was unresolved. Because of the lack of informativess and poor resolution, it was decided not to include this tree in the dissertation, but mention was made of it. The use of this taxon as outgroup was employed in a neighbor joining analysis presented slightly later in the thesis.

Paragraph 7: The discussion was altered to better reflect the possibility of both apomictic and sexual reproduction.

Examiner: Dr. Crawford

Paragraph 1: The family name (Asteraceae) was added in the title of Chapter 1.

Paragraphs 2 & 3: Additional references were sought and more indepth definitions of polyploidy were incorporated.

Paragraph 4: The full generic name (*Chromolaena*) was spelt out as suggested, and the name of the species studied was spelt out in full in all subheadings.

Paragraph 5: The text referred to was modified to remove any ambiguity in the meaning of the text.

Paragraph 6: Further references were sought and incorporated to give a fuller understanding of the influence of polyploidy on the species.
Paragraph 7: The heading was changed as suggested.
Paragraph 8: The text was altered to avoid any ambiguity in its meaning.
Paragraph 9: The sentence was removed.
Paragraph 10: The text was modified to better explain what is meant by “necessary resolving power”.
Paragraph 11: Text was modified to further emphasise Vos’s (1989) work, and the limitations thereof, on peroxidases.
Paragraph 12: Text was modified to address the questions posed.
Paragraph 13: References were sought, and further text was added regarding the usefulness of micro- and minisatellites.
Paragraph 14: As suggested by my supervisor, Dr N. P. Barker, no changes were made regarding this as it was considered to be a comment rather than instruction to modify or change the thesis.
Paragraph 15: References were sought and their ideas regarding low-copy nuclear genes were incorporated into the text.
Paragraph 16: Text was modified to better explain both the advantages and disadvantages of the 18-26S ribosomal repeat in phylogenetic inferences.
Paragraph 17: References were sought and text modified accordingly to better reflect the value of cpDNA and ITS for detecting parents of hybrids.
Paragraph 18: The sentence referred to was deleted.
Paragraph 19: As suggested by Dr. Barker, the concept and context of geographic isolation resulting in genetic distinctiveness either side of a barrier was expanded upon. The term “intraspecific use” was expanded to remove ambiguity.

All of the changes above have been made after consultation with Dr. Barker, and are to the best of my knowledge accurate.

Yours Faithfully

Inge von Senger