D O C T O R A A T S P R O E F S C H R I F T

2005

Faculteit Wetenschappen

Ecological genetics of heavy metal tolerant populations of *Suillus luteus*

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen, richting Biologie,
te verdedigen door

Ludo A. H. MULLER

Promotor : Prof. dr. J. Colpaert
Co-promotor : Prof. dr. J. Vangronsveld

LIMBURGS UNIVERSITAIR CENTRUM
PARTNER IN DE UNIVERSITEIT LIMBURG

2005

Faculteit Wetenschappen

Ecological genetics of heavy metal tolerant populations of *Suillus luteus*

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen, richting Biologie,
te verdedigen door

Ludo A. H. MULLER

Promotor : Prof. dr. J. Colpaert
Co-promotor : Prof. Dr. J. Vangronsveld

LIMBURGS UNIVERSITAIR CENTRUM
PARTNER IN DE UNIVERSITEIT LIMBURG

According to the guidelines of the Limburgs Universitair Centrum, a copy of this publication has been filed in the Royal Library Albert I, Brussels,as publication D/2005/2451/33

Samenvatting

De voorgestelde studie beoogde een analyse van de adaptatie aan verhoogde concentraties van zware metalen in populaties van de ectomycorrhizavormende basidiomyceet *Suillus luteus*. Uit voorgaand onderzoek was gebleken dat een verhoogde tolerantie voor zware metalen aanwezig is in populaties die voorkomen in verontreinigde gebieden, alsook dat deze populaties mogelijk een verminderde genetische diversiteit bezitten ten gevolge van stichterseffecten en flessenhalsfenomenen.

Vruchtlichamen van *S. luteus* werden verzameld op negen verschillende locaties in de provincie Limburg (België). Op zes staalnameplaatsen zijn de bodems ernstig verontreinigd met hoge concentraties aan zware metalen door vroegere activiteiten van verschillende zinkfabrieken in de directe omgeving. Een fenotypische karakterisatie van het tolerantiekenmerk aan de hand van *in vitro* dosis-respons experimenten toonde aan dat isolaten uit verontreinigde gebieden een duidelijk hogere tolerantie bezitten voor verhoogde externe concentraties aan zware metalen. Er werden echter geen aanwijzingen gevonden voor het optreden van natuurlijke selectie op genoom-niveau. Analyse aan de hand van AFLP- en microsatellietmerkers onthulde grote genetische diversiteit binnen populaties, zowel in verontreinigde als in niet-verontreinigde gebieden, en een beperkte differentiatie tussen de populaties. Daarenboven werd er geen bewijs gevonden voor huidige of vroegere flessenhalsfenomenen en bleek er geen effect te zijn van de verontreiniging op de genetische differentiatie van de verschillende populaties.

Overeenkomstig de grote genomische diversiteit werd door middel van cDNA-AFLP variatie aangetoond op het niveau van het transcriptoom in tolerante en niet-tolerante *S. luteus*-isolaten. Een deel van deze variatie bleek te wijten aan sequentiedifferentiatie tussen de geteste isolaten, maar verificatie aan de hand van RT-PCR toonde aan dat de meerderheid van de verschillen in expressiepatroon te wijten waren aan verschillen in het niveau van de genexpressie. Hoewel transcriptoom-vingerafdruktechnieken geen uitsluitsel geven over de relevantie van verschillen in genexpressie en een deel van de verschillen te wijten zijn aan "genetisch liften" (genetic hitchhiking) alsook aan variatie in expressieniveaus van nature aanwezig in populaties, konden

sommige genen die een verschillend expressiepatroon vertoonden toch in verband worden gebracht met het verschil in tolerantieniveau. Zo werden genen geïdentificeerd die coderen voor een hitte-geïnduceerd (heat shock) proteïne (HSP60), een metaaltransporter en enkele eiwitten die tussenkomen bij de ubiquitine-gemediëerde proteolyse, en die betrokken zouden kunnen zijn bij de detoxificatie van en cellulaire verdediging tegen zware metalen in het tolerante isolaat. Daarnaast werd een hydrophobine-homoloog geïdentificeerd, een waterafstotend eiwit uit de celwand waarvan aangenomen wordt dat het bescherming levert tegen aspecifiek binnendringen van zware metalen in de cel, dat enkel tot expressie kwam in het tolerante isolaat en dat mogelijks mede verantwoordelijk is voor de hogere externe zinkconcentratie die vereist is om modulatie van transcriptie te veroorzaken in het tolerante isolaat.

Uit de resultaten van de verschillende experimenten in deze studie volgt dat verontreiniging door zware metalen weinig invloed heeft op de genetische structuur van *S. luteus*-populaties, een conclusie die *a priori* onwaarschijnlijk leek maar die reeds werd beschreven in andere organismen. In het geval van *S. luteus* zouden flessenhalseffecten tijdens de kolonisatie van verontreinigde gebieden beperkt kunnen zijn gebleven door de introductie van verschillende tolerante genotypes, gevolgd door regelmatige sexuele reproductie en een snelle evolutie van het tolerantiekenmerk. Daarenboven kan herhaalde migratie van tolerante genotypes, ontstaan door kruising met niet-tolerante genotypes in niet-verontreinigde gebieden, bijgedragen hebben aan de grote diversiteit van populaties in verontreinigde gebieden.

Abstract

Genomic and transcriptomic aspects of the adaptation to increased environmental heavy metal concentrations were studied in relation to phenotypic observations of heavy metal tolerance in populations of the ectomycorrhizal basidiomycete *Suillus luteus*. *S. luteus* sporocarps were collected at nine different locations in the Campine phytogeographic district (province of Limburg, Belgium). At six of these locations the soil is severely polluted with heavy metals due to past activities of several zinc smelters and phenotypic characterization, by means of *in vitro* growth experiments, revealed an obvious increase of the heavy metal tolerance in the populations at these sites. However, evidence for natural selection was not found at the genomic level. AFLP and microsatellite marker analyses both indicated high levels of genetic diversity within the total sample and within geographic subpopulations, but limited population differentiation. No current or past reduction of the genetic diversity of populations inhabiting polluted environments was discovered, nor was population differentiation found to be related to the environmental heavy metal pollution.

In accordance with the high level of genomic diversity revealed by the more-or-less neutral AFLP and microsatellite markers, considerable diversity was found at the level of the transcriptome by cDNA-AFLP analysis of a heavy metal tolerant and a nontolerant isolate. Although transcript profiling should be regarded as a tool for exploratory data analysis and some of the differences in expression level will be due to genetic hitchhiking and natural regulatory variation, some information was obtained about possible mechanisms that may explain the difference in tolerance level. Expression of a hydrophobin homologue, a water-repellent protein in the cell wall which is believed to protect against penetration of heavy metal containing solutions, may be partly responsible for the higher external zinc concentrations required to induce modulation of transcription in the tolerant isolate and candidate genes that may play a role in the detoxification of and cellular defense against heavy metals were identified as encoding a heat shock protein (HSP60), a putative metal transporter and several proteins involved in ubiquitin-dependent proteolysis.

The results of these different experiments show that heavy metal pollution has limited influence on the genetic diversity of *S. luteus* populations, a result

that seems unlikely *a priori* but that has been described in several other organisms. In case of *S. luteus* bottleneck events associated with the foundation of populations in polluted environments may have been prevented by the introduction of different tolerant genotypes, followed by sexual reproduction on the site and rapid evolution of the tolerance trait. Additionally, recurrent migration of tolerant genotypes that originate due to admixture in nonpolluted areas, which is very likely due to the high migration rates, may have attributed to the high levels of genetic diversity of populations in polluted habitats.

To Joke

Acknowledgements

This doctoral work would not have been possible without the help of many people for which I am very grateful. I would especially like to thank Prof. Jan Colpaert and Prof. Jaco Vangronsveld for giving me the opportunity to start this study and for their continuous support during the past four years, Prof. André Van Laere for being so generous to allow me to use his lab and equipment, Prof. Godelieve Gheysen and Dr. Bart Van Droogenbroeck for helping me with the AFLP experiment, Prof. Nathalie Verbruggen and Dr. Adrian Craciun for their help and support with the cDNA-AFLP study, Ilse Bastians, Catherina Coun and Evy Vanderheyden for their high quality-sequencing of many, many sequences, Marc Lambaerts for his help and support in the lab, Dr. Travis Glenn and Dr. Chunlan Lian for providing me with detailed protocols for the isolation of microsatellites, Carine Put and Ann Wijgaerts for their accurate ordering of all the products, and Beryl Muller for her proofreading of this manuscript.

Contents

List of Figures

List of Tables

Chapter 1

General introduction

1.1 Mycorrhizal symbiosis

1.1.1 Types of associations

Under natural conditions, over 90% of all terrestrial plants form symbiotic associations with root-colonizing mycorrhizal fungi (Cairney, 2000). In this type of symbiosis, the fungus facilitates nutrient uptake from the soil by absorbing, assimilating and transporting nutrients to the plant, which in turn may provide the fungus with photosynthates. This bidirectional transfer of nutrients forms the basis for the mutualistic nature of most mycorrhizal symbioses. However, mycorrhizal association may also provide other benefits to the host plant, including improved water relations and resistance to pathogenic microorganisms or environmental pollution (Smith and Read, 1997).

Mycorrhizae can be categorized into seven main types according to their morphology and to the fungal and plant taxa involved. Based on the fungal partner in the symbiotic association, the main types of mycorrhizae are divided into those involving obligately symbiotic aseptate endophytes in the phylum Glomeromycota and those formed by septate fungi in the phyla Ascomycota and Basidiomycota. The associations formed by the members of the Glomeromycota are collectively called vesicular-arbuscular mycorrhizae, which refers to characteristic structures, the arbuscules and vesicles, formed within or between the cortical cells of the host plant. Because vesicles may be absent in mycorrhizae of this type, the term "arbuscular mycorrhizae" has been proposed (Berch, 1987). The plant hosts involved in this type of association include mosses and liverworts (Bryophyta), ferns (Pteridophyta), conifers (Gymnospermae) and flowering plants (Angiospermae). Vesicular-arbuscular mycorrhizae are the most common underground symbiosis and the fossil record suggests that they were important for the colonization of the land by vascular plants some $450 - 500$ million years ago (Smith and Read, 1997; Cairney,

2000).

The remaining types of mycorrhizae are associations formed between septate fungi and autotrophic or heterotrophic trees, shrubs and herbs. In ectomycorrhizae the fungal partner forms a mantle or sheath that encloses the rootlets and from which hyphae or rhizomorphs radiate outwards into the soil. Hyphae also penetrate the root and form a complex intercellular system called the Hartig net. However, little or no intracellular penetration occurs. In ectendomycorrhizae the fungal mantle is reduced or absent, but the Hartig net is generally well-developed and the hyphae do penetrate into the cells of the host plant. The plant hosts involved in the formation of ectomycorrhizae and ectendomycorrhizae are conifers and flowering plants.

Arbutoid mycorrhizae are formed on the roots of various members of the Ericales and have a fungal mantle, external hyphae and a well developed Hartig net, but the hyphae characteristically form extensive coils within the cells of the plant. Orchid and monotropoid mycorrhizae are associated with plant hosts that are achlorophyllous for at least a part of their lives. The fungi forming associations with members of the Orchidaceae can be saprophytes, parasites of other plants or may form mycorrhizae with other, autotrophic, plants and transfer organic carbon and mineral nutrients to the orchids. In the case of monotropoid mycorrhizae, the fungal partner also forms ectomycorrhizae with the roots of neighbouring autotrophic plants and it is believed that organic carbon is transferred to the achlorophyllous plant host.

Finally, Ericoid mycorrhizae are the result of associations between ascomycetes and autotrophic members of the Ericaceae. Many of the ericaceous host plants inhabit soils where most of the nutrients are present in organic form and the associated fungi may have a considerable role in mobilizing these nutrients and making them available to the plant (Smith and Read, 1997).

From this overview it is clear that the term "mycorrhiza" is used to describe a variety of associations found between soil fungi and plant roots. However, both functional and structural unifying features, as outlined above, allow a distinction between the mycorrhizal condition and other associations between soil fungi and plants.

1.1.2 Ectomycorrhizae

Ectomycorrhizae are probably the most diverse type of mycorrhiza, primarily due to the high number of fungal species involved in this type of relationship. Molina et al. (1992) estimated that at least 5000 fungal species are forming ectomycorrhizae and most of these are in the basidiomycetes. Although the majority of ectomycorrhizal fungi have a broad host range, specificity, at least at the level of the plant genus, has been observed in a diverse range of fungi, plant hosts and habitats (Smith and Read, 1997).

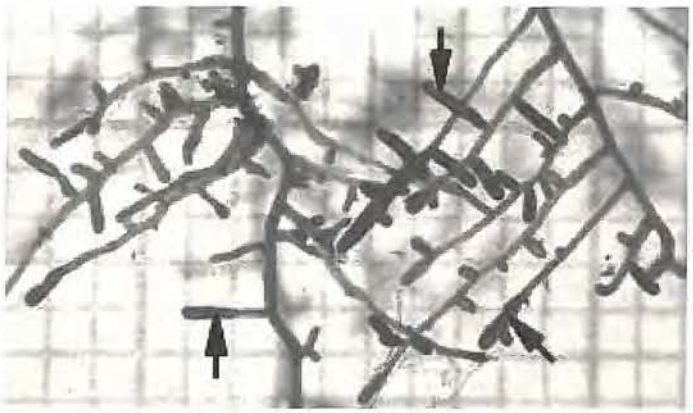

Figure 1.1: Example of ectomycorrhizal short roots (arrows) of birch (*Betula alleghaniensis*). (Copyright ©by Mark Brundrett.)

Conversely, a relatively small number of seed plants, around 3% (Meyer, 1973), are ectomycorrhizal and the vast majority of these are woody perennials. However, because of their disproportionate occupancy of land surface and their economic importance as producers of timber, their global importance is increased. Families of predominantly ectomycorrhizal species include the Pinaceae, which form the major component of the boreal forests of the northern hemisphere, the Fagaceae, which dominate the northern temperate forests, and the Myrtaceae, dominating temperate and subtropical forests of the southern hemisphere.

Most plants with ectomycorrhizae have roots with a modified lateral root branching pattern. This pattern, which is called heterorhizy, consists of short mycorrhizal lateral roots supported by a network of long roots (see figure 1.1). Mycorrhizal lateral roots are more abundant in topsoil layers containing humus than in underlying layers of mineral soil and tend to be colonized by a single fungal species (Harvey et al., 1976). Nonetheless, a single tree may be host to dozens of fungal partners on different parts of its root system. The ectomycorrhizal roots are characterized by the presence of a sheath or mantle made up of fungal tissue enclosing the root, a Hartig net formed by hyphae growing inward between the epidermal and cortical cells and a system of hyphal elements growing outward and forming connections with the soil and with the fruiting bodies of the fungi that form the mycorrhizae (see figure 1.2). Although considerable variation exists in the extent to which the Hartig net, mantle and extraradical mycelium are developed, these characteristic features are well conserved and confer particular functional attributes that further distinguish ectomycorrhizae from other mycorrhizal types.

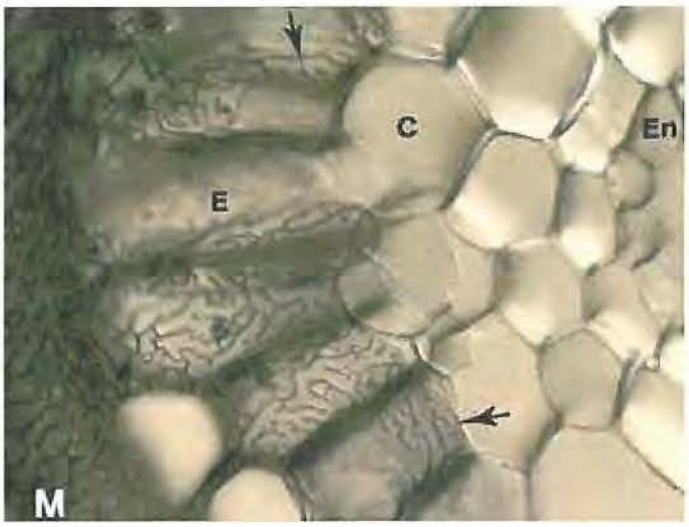

Figure 1.2: Ectomycorrhizal root cross section of poplar (*Populus tremuloides*) showing Hartig net hyphae (arrows) around epidermal cells; C, cortex cell; E, epidermal cell; M, mantle; En, endodermis. (Copyright ©by Mark Brundrett.)

The functional basis of the ectomycorrhizal relationship is similar to that of vesicular-arbuscular mycorrhizae in that the plant host exchanges carbohydrates for mineral nutrients. However, ectomycorrhizae are more complex than arbuscular mycorrhizae. Many ectomycorrhizal fungi have enzyme systems unknown in plants, which allow them to break down organic matter and extract phosporus, micronutrients and, in some cases, nitrogen. These enzyme systems allow the fungi to withdraw nutrients from organic substrates such as the soil litter layer. In addition, the hyphal mantle protects the underlying root from other microorganisms and environmental toxins (Smith and Read, 1997).

1.1.3 The genus *Suillus* and *Suillus luteus*

The ectomycorrhizal fungal genus *Suillus* consists of species that produce generally epigeous mushrooms with tubular hymenophores, which are common in conifer forest ecosystems. Taxonomically, the genus is grouped within the order Boletales (Hawksworth et al., 1995), where the majority of species are ectomycorrhizal. Within Boletales, the genus *Suillus* is grouped within the monogeneric family Suillaceae, which belongs to the suborder Suillineae together with the closely related families Rhizopogonaceae and Gomphidiaceae (Besl and Bresinsky, 1997). The genus *Suillus* is characterised by medium to large fleshy terrestrial boletes with generally a viscid to glutinous cap and a

Figure 1.3: Ectomycorrhizal association between *Pinus radiata* and *Suillus brevipes* showing dichotomously branched mycorrhizal short roots (arrows). (Copyright ©by Nick Malajczuk, Randy Molina and Jim Trappe.)

stipe with or without a ring. The hymenium has small or larger, sometimes radial pores and the spores are elongate, smooth and brownish. *Suillus* mycorrhizas are characterised by a dichotomous to tuberculate form, conspicuous rhizomorphs and abundant mycelia (see figure 1.3).

Suillus species form ectomycorrhizal associations exclusively with members of the Pinaceae and hence, their distribution coincides with the natural distribution of pinaceous conifers. They are predominantly found in boreal and temperate zones of the northern hemisphere, although some species may appear in the paleo- and neotropics. The most important host genera for *Suillus* are *Pinus*, *Larix* and *Pseudotsuga*. The genus *Pinus* comprises 94 species which dominate the natural vegetation in many regions of the northern hemisphere and the genus *Larix* comprises ten species, of which two are widespread in Eurasia and one in North America. Most *Suillus* species associate only with a single host genus and many have a limited geographical range. The highest species richness of *Suillus*, almost 70 species, is found in North America, coinciding with the highest species richness of conifers. Europe and Asia have respectively 31 and 36 species of *Suillus* (Dahlberg and Finlay, 1999).

Suillus luteus (L.:Fr.) Roussel (see figure 1.4) is a typical pioneer ectomycorrhizal fungus that associates with the roots of young pine trees colonising disturbed sites. It has a relatively short generation time (3 to 5 years) and frequent sexual reproduction. Like most basidiomycetes, *S. luteus* has a heterothallic life cycle. In this life cycle, nuclear fusion, meiosis and spore formation occur in the fruit bodies produced by dikaryotic mycelia. After spore germination, monokaryotic mycelia develop and fusion between sexually compatible monokaryons, corresponding to different mating types, leads to the formation of a dikaryon. This dikaryon has a prolonged vegetative stage and is the dominant form in which the ectomycorrhizal fungi exist in their natural habitats. Sexual compatibility between different haploid monokaryons is determined by the mating-type genes. Different mating types correspond to

Figure 1.4: Fruit bodies of *Suillus luteus*.

mycelia that are morphologically identical but self-incompatible. In contrast with the ascomycetes, which have only two mating types controlled by alternative alleles at a single genetic locus, the hymenomycetes have a complex mating-type control, with possibly thousands of mating types determined by one to four multiallelic loci (Debaud et al., 1999). Two classes of heterothallic basidiomycetes are recognized: the bipolar and the tetrapolar species. In the bipolar species (25% of the heterothallic species), sexual compatibility between monokaryons is controlled by a single factor, while in the tetrapolar species (75% of the heterothallic species) two unlinked factors control compatibility. Two monokaryons are sexually compatible if, in case of a bipolar species, the alleles at the mating-type locus differ or, in case of a tetrapolar species, the alleles at both mating-type loci are different. Thus, heterothallism enforces outbreeding. The mating systems of only a limited number of ectomycorrhizal species have been resolved. However, Fries and Neumann (1990) found three *Suillus* species, among which *S. luteus*, to be heterothallic and to possess a bipolar compatibility system.

1.2 Heavy metal pollution and tolerance

1.2.1 Environmental heavy metal pollution

In scientific literature, the term "heavy metal" was introduced by Bjerrum in 1936. He defined heavy metals as those metals having elemental densities above 7 g/cm^3 (Bjerrum, 1936). However, no authoritative definition seems to exist for the term "heavy metal" and since it was introduced the original

definition has been modified by different authors without consistency. These definitions have been based on various characteristics such as density, atomic weight, atomic number and toxicity (Duffus, 2002). In terms of toxicity, following definitions can be found:

> Elements that are commonly used in industry and which are generally toxic to animals and to aerobic and anaerobic processes; however, not all are dense nor entirely metallic. Includes As, Cd, Cr, Cu, Pb, Hg, Ni, Se and Zn. (Scott and Smith, 1981)
>
> Outdated generic term referring to lead, cadmium, mercury and some other elements which generally are relatively toxic in nature; recently, the term toxic element has been used. The term also refers to compounds containing these elements. (Hodgson et al., 1988)

Here, the term "heavy metals" is used to refer to the eight elements which are considered to be of major environmental concern according to the third North Sea Conference held in The Hague in 1990: arsenic (As), cadmium (Cd), chrome (Cr), mercury (Hg), lead (Pb), copper (Cu), nickel (Ni) and zinc (Zn).

Soils contaminated with heavy metals can either occur naturally or as a result of anthropogenic activities such as mining, agriculture, traffic and industrial processes. Atmospheric deposition and several management practices, such as the use of fertiliser, manure, pesticides, sewage sludge and bio waste, cause heavy metal input into soils and despite adsorption processes, which inhibit the migration of heavy metals in the soil, leaching leads to a further contamination of the ground water.

Depending on their bio-availability, heavy metals are taken up by plants and animals and because they are not metabolized, some tend to accumulate (e.g. Cd and Pb). As trace elements, some heavy metals (e.g. Mo, Cu, Mn and Zn) are essential for the survival of all forms of life. They show a variety of functions in the cell, ranging from regulators of biological processes to important structural and functional components in proteins. However, essential or not, most heavy metals become potentially toxic at higher concentrations. This toxicity may be the result of metal ions binding to sulphydryl groups in proteins, which leads to an inhibition of activity or disruption of structure, or of the displacement of essential metal ions. In addition, heavy metal ions may stimulate the formation of free radicals and reactive oxygen species, resulting in oxidative stress (Ochiai, 1987; Hall, 2002).

1.2.2 Evolutionary aspects of heavy metal tolerance

Although soils contaminated with high levels of heavy metals can be extremely toxic, they are often colonized by a variety of plant, fungus and animal species.

Since the first observation of heavy metal tolerance in plant populations originating from polluted sites (Prat, 1934), it has been suggested that these have adaptively evolved as tolerant ecotypes and subsequently, a genetic basis for heavy metal tolerance has been presumed (Bradshaw, 1952; Macnair, 1993). It has been shown that such adaptive heavy metal tolerance is controlled by a small number of major genes, with perhaps some contribution from more minor modifier genes and, in addition, it is believed that specific mechanisms are involved for each heavy metal present at a toxic concentration (Macnair, 1993; Macnair et al., 2000; Schat et al., 2000).

At the cellular level, plants possess a variety of potential mechanisms which may be involved in the detoxification of and tolerance to heavy metals. It appears that these mechanisms are primarily involved in preventing the build-up of toxic concentrations of free heavy metal ions at sensitive sites within the cell, rather than developing proteins tolerant to elevated concentrations of heavy metals. The strategies for avoiding the build-up of toxic concentrations of heavy metals include roles for mycorrhizas and for cell wall and extracellular exudates and may involve a reduction of the uptake of heavy metals or an enhanced efflux pumping of heavy metal ions. Within the protoplast, chelation of heavy metals by organic acids, amino acids or peptides may prevent toxic effects and heat shock proteins or metallothioneins may be involved in the repair of stress-damaged proteins. Additionally, compartmentation away from metabolic processes by transport into the vacuole may also contribute to the detoxification of heavy metals (Hall, 2002).

Among the adapted rhizosphere microorganisms that may assist plants in colonizing polluted sites, the ectomycorrhizal fungi occupy a predominant position. Although the mechanisms involved in conferring an increased tolerance have not been resolved, they seem to be quite diverse and show considerable species and metal specificity. However, most mechanisms that have been proposed involve various exclusion processes that restrict heavy metal movement to the host roots, e.g. absorption of heavy metals by the fungal sheath, chelation by fungal exudates and adsorption onto the external mycelium (Jentschke and Godbold, 2000). The mechanisms employed by the fungi at the cellular level to tolerate increased concentrations of heavy metals are believed to be similar to some of the strategies employed by higher plants, namely binding to extracellular materials or compartmentation by transport into the vacuoles (Hall, 2002).

The evolution of heavy metal tolerance has been shown to be extremely fast in several cases (Wu et al., 1975; Macnair, 1981) and it seems to depend on the presence of tolerant individuals in the surrounding founder populations (Macnair, 1987; Al-Hiyaly et al., 1993). The probability that tolerant individuals occur in natural populations depends on several parameters:

The mutation rate (μ). Little is known about the mutation rate to heavy metal tolerance, but an important factor is the number of genes involved in the tolerance trait. If several genes are involved, it is unlikely that tolerance will evolve, unless it is a truly polygenic trait which is governed by genes that are polymorphic within natural populations.

The selection pressure against tolerance (s). Heavy metal tolerance is believed to be a disadvantageous trait under natural conditions and this may explain the common observation that ecotypes from noncontaminated sites show better growth on noncontaminated substrates than ecotypes from contaminated sites (Antonovics et al., 1971). However, the disadvantage may be small as shown in earlier studies for *Mimulus guttatus* and *Agrostis tenuis* (Macnair, 1987).

The population size (N). Population size is the parameter that is likely to be the most important in many cases. If population size is infinite, then the equilibrium phenotype frequency of a deleterious character is of the order μ/s (Gillespie, 1998). However, stochastic processes will cause considerable variance around the expected value of μ/s in case of small population sizes and many populations will not possess the tolerance genes at all. Thus, common species with large population sizes are more likely to be able to evolve tolerance as compared to rare species, inbreeding species or species with small local population sizes (Macnair, 1987).

Thus the colonization of polluted soils requires genetic variance for tolerance which in turn depends on the mutation rate, selective disadvantage and population size. Also, a certain degree of preadaption to the edaphic conditions of the polluted site may be necessary. Since the occurence of tolerance in natural populations is low, the initial colonization of polluted soils is likely to be by a small number of individuals and, once established, the new population will expand rapidly throughout the new environment where interspecific competition is limited. Founder effects are expected to cause genetic differences between the initial population and the founder population and this differentiation will only slowly decrease because of gene flow between the two populations.

Natural selection will lead to a rapid fixation of the alleles responsible for the tolerance trait and many alleles, which are linked to the tolerance genes, will be taken to fixation by genetic hitchhiking (i.e. the process by which alleles increase in frequency if linked to an allele whose frequency is increasing because of natural selection; Maynard Smith and Haigh (1974)). This will be followed by the selection of modifiers that improve the tolerance and hitchhiking may again increase the frequency of alleles linked to the modifier genes. Other

characters that are unrelated to tolerance but that enhance the fitness in the new environment will also be subjected to natural selection. For example, mine soils are often nutritionally poor and have a lower water retention capacity due to a lack of organic material and a much coarser texture than other soils. Adaptation to these conditions is to be expected (Macnair, 1987).

Genetic differentiation may be accompanied by reproductive isolation and thus can eventually lead to speciation. Both prezygotic and postzygotic reproductive isolation mechanisms have been found in different plant populations inhabiting mining sites. Differences in flowering time have been found in both *Agrostis capillaris* and *Anthoxanthum odoratum* (McNeilly and Antonovics, 1968) and a gene, closely associated with the tolerance gene, causing postmating isolation has been discovered in *M. guttatus* (Macnair and Christie, 1983).

By now, it should be clear that the phenomenon of heavy metal tolerance provides a valuable model system for the study of microevolution. Heavy metal contaminated areas may be considered as ecological "islands", which are surrounded by noncontaminated "mainland", where a predominant environmental factor is imposing severe selection pressure on the residing populations. These islands are of various sizes (a few to several hundreds of hectares) and can be of recent origin (several decades), thus providing the opportunity to investigate the initial steps in the establishment of differentiated populations under severe selection pressure (Lefèbvre and Vernet, 1989).

1.3 Research objectives

Research in the field of evolutionary and ecological functional genomics seeks to understand the functional significance of genomic variation for organisms inhabiting their natural environments (Feder and Mitchell-Olds, 2003). In order to accomplish this goal, it is required to elucidate the biotic and abiotic challenges to function in natural environments, the molecular and cellular mechanisms required to meet these challenges and the evolutionary processes that manipulate these mechanisms through time. Therefore, evolutionary and ecological functional genomics is a multidisciplinary research field that integrates molecular, cellular, organismal, population and ecological approaches.

Ideal model species in evolutionary and ecological functional genomics have to meet a wide range of criteria. Amongst others, relatively undisturbed habitats in the native range of the species have to exist and genetic variation, present in natural populations, must allow local adaptation to biotic or abiotic environments. Each genetic variant should be characterized phenotypically under natural conditions and the biotic or abiotic environmental factors have to correlate with each segregating haplotype. The genetic structure and

dynamics of natural populations should be known and the evolutionary forces underlying the genetic variation have to be inferred.

In this study, the fungal species *Suillus luteus* is proposed as a model species for evolutionary and ecological functional genomics. Although many criteria are not met, this species is potentially a valuable model for studying adaptation to local environments. Prior research has already shown that *S. luteus* populations in heavy metal contaminated areas are dominated by heavy metal tolerant individuals and that a genetic basis for this tolerance trait is very likely. Moreover, preliminary screening indicated possible founder effects occuring at polluted sites and resulting in a reduction of the genetic variation (Colpaert et al., 2000, 2004).

The objective of this study was to investigate diverse aspects of the evolutionary adaptation to increased levels of environmental heavy metal concentrations in *S. luteus* populations. Therefore, several populations, inhabiting both heavy metal contaminated and noncontaminated environments, were sampled and collected isolates were characterised phenotypically. Furthermore, transcript profiling by cDNA-AFLP was used to describe patterns of gene expression in heavy metal tolerant and nontolerant individuals and to identify candidate genes that might influence the tolerance trait. Finally, AFLP and microsatellite analyses were used to infer the genetic structure of the sampled populations and to determine the evolutionary forces involved in the differentiation of these populations. Besides providing insight into microevolutionary aspects of the local adaptation to heavy metal polluted environments, the results obtained by this study could enhance the value of *S. luteus* as a model species in evolutionary and ecological functional genomics.

Chapter 2

Heavy metal tolerance in *Suillus luteus*

2.1 Phenotypic characterization of heavy metal tolerance in *Suillus luteus*

2.1.1 Introduction

Natural selection, the only evolutionary force that can produce adaptation, is active at the level of the phenotype. Genetic variation, in combination with environmental components, determines the phenotypic variation among individuals for the collection of traits expressed by an organism. The amount and nature of the phenotypic variation in turn determine the extent of variation in fitness, i.e. the relative ability of a genotype to pass on its alleles to future generations in a particular environment. Thus, evolutionary adaptation is the result of the selection of the phenotype showing the highest fitness in a particular environment and is manifested at the level of the genotype as a change in allele frequencies.

Recognizing adaptive phenotypic change is often difficult because phenotypic change is usually gradual, occuring only over many generations, and the determination of many phenotypic traits is often complicated and may be influenced by several genes and complex environmental factors. However, rapid evolutionary change is possible in response to strong ecological or environmental forces and human-caused environmental changes, such as the evolution of pesticide resistance and heavy metal tolerance. Adaptive heavy metal tolerance has been observed in various groups of organisms, including plants and fungi, inhabiting heavy metal-enriched environments. In the case of *S. luteus*, it has been shown by *in vitro* tolerance screening that populations in heavy metal contaminated soils have been established by adaptive evolution

Figure 2.1: Geographic locations of the *Suillus luteus* subpopulations sampled for this study: Lommel (L: Lm, Ls, Lc, Lk and Lr), Paal (P), Eksel (E), Hechtel (Hh), Neerpelt (N), Overpelt (Of); black circles, sampled subpopulations; black squares, industrial sources of heavy metal pollution.

(Colpaert et al., 2000).

In order to phenotypically characterize the different *S. luteus* populations that were sampled for assessing the effects of heavy metal pollution on their genetic structure and to reveal their adaptive evolution, *in vitro* zinc tolerance screening was performed. Although this method does not monitor the fitness of tolerant and nontolerant isolates directly, the *in vitro* growth rate is assumed to be a valid indirect estimator of fitness which can be used to demonstrate adaptive evolution.

2.1.2 Material and methods

Sampling sites

Nine different sampling sites were selected along a heavy metal pollution gradient that developed in the previous century due to the activities of several factories that produced zinc by pyrometallurgical processing of zinc ores. From

Site (abbreviation)	Zn pore water (μg/l)	Zn pine needles (μg/g d.wt)
Lommel Maatheide (Lm)	7.7 (3.1)	272 (44)
Neerpelt (N)	3.2 (1.4)	177 (58)
Lommel kanaal (Lk)	4.8 (1.1)	143 (55)
Overpelt fabriek (Of)	4.9 (1.5)	117 (41)
Lommel sahara (Ls)	3.7 (1.2)	110 (63)
Lommel containerpark (Lc)	4.4 (0.4)	101 (29)
Eksel (E)	0.54 (0.04)	43 (15)
Hechtelse heide (Hh)	0.54 (0.1)	31 (16)
Paal (P)	0.12 (0.03)	26 (8)

Table 2.1: Pollution levels of the different sampling sites, measured as the zinc concentration in soil pore water and pine needles (SD in parentheses; Colpaert et al. (2004)).

the end of the nineteenth century until the mid-seventies, these factories emitted large quantities of zinc and several other heavy metals into the atmosphere.

All the study sites are located in the Campine phytogeographic district (province of Limburg, Belgium; see figure 2.1), which is characterised by base-poor, sandy soils of low fertility. Most forests at these sites are pioneer forests or primary plantations of predominantly pine trees (*Pinus sylvestris* and *P. nigra*), sometimes mixed with birches (*Betula* sp.), and all were younger than 30 yr. Six of the sampling sites are located within a radius of 2 km from the zinc smelters of Lommel (Lm, Lc, Lk and Ls) or Overpelt (N and Of) and the soils on these sites are seriously contaminated through atmospheric deposition. The remaining three sampling sites (E, Hh and P) show low levels of pollution and are situated, respectively, 7.6, 14.8 and 15.6 km from a zinc smelter. The zinc pollution at these sites has been assessed by analysing the zinc content of soil pore water and pine needles (Colpaert et al., 2004). In soils with high levels of pollution, the zinc concentration in the soil pore water varied between 7.7 ± 3.1 μg/l and 4.8 ± 1.1 μg/l, whereas in noncontaminated soils the concentration varied between 0.54 ± 0.04 μg/l and 0.12 ± 0.03 μg/l. In the pine needles, the zinc concentration varied between 272 ± 44 μg/g d.wt and 101 ± 29 μg/g d.wt in contaminated areas and between 43 ± 15 μg/g d.wt and 26 ± 8 μg/g d.wt in noncontaminated areas. An overview of the characteristics of the different sampling sites is presented in tables 2.1 and 2.2.

Site (abbreviation)	Geographical coordinates (latitude/longitude)	Forest type	Age of trees (yr)
Lommel Maatheide (Lm)	51°14'12"-N/05°15'28"-E	Industrial area, mostly planted trees	27
Neerpelt (N)	51°14'01"-N/05°24'28"-E	Industrial area, spontaneous colonisation	1 – 15
Overpelt fabriek (Of)	51°13'39"-N/05°24'09"-E	Sand dunes, spontaneous colonisation	1 – 25
Lommel sahara (Ls)	51°14'45"-N/05°16'23"-E	Industrial area, first rotation forest	10
Lommel containerpark (Lc)	51°14'08"-N/05°16'26"-E	First rotation forest, previously grassland	8
Lommel kanaal (Lk)	51°14'49"-N/05°15'42"-E	First rotation forest, disturbed soil	8
Eksel (E)	51°08'55"-N/05°20'49"-E	Road side, spontaneous colonisation	1 – 15
Hechtelse heide (Hh)	51°06'33"-N/05°22'11"-E	Sand dunes, spontaneous colonisation	1 – 25
Paal (P)	51°03'33"-N/05°10'33"-E	Industrial area, first rotation forest and spontaneous colonisation	1 – 26

Table 2.2: Characteristics of the different sampling sites: geographical coordinates, characteristics of the forest stands and age of the pine trees.

Fungal material

Fungal cultures were established from sporocarps collected from the different sampling sites selected for this study on a single day in 2000, 2001 or 2002. During sampling, a minimal distance of approximately 10 m was respected between collected fruit bodies in order to avoid multiple sampling of the same genet. The majority of the cultures were kept in a collection on Fries medium without elevated zinc concentration.

Metal tolerance analysis

In order to avoid possible carry-over effects, which may influence the tolerance analysis, the mycelia were subcultured at least three times on Fries medium without elevated zinc concentration before being included in the screening tests. Zinc tolerance was tested on solid modified Fries medium (Colpaert et al., 2000), which contained 28 mM glucose, 5.4 mM ammonium tartrate, 1.5 mM KH_2PO_4, 0.4 mM $MgSO_4 \cdot 7H_2O$, 0.3 mM NaCl, 0.2 mM $CaCl_2 \cdot 2H_2O$, 4 μM $FeCl_3 \cdot 6H_2O$, 6 μM $MnSO_4 \cdot H_2O$, 0.8 μM $CuSO_4 \cdot 5H_2O$, 56 μM *myo*-inositol, 0.1 μM biotin, 0.5 μM pyridoxine, 0.3 μM riboflavin, 0.8 μM nicotinamide, 0.7 μM *p*-aminobenzoic acid, 0.3 μM thiamine, 0.2 μM Ca-pantothenate and 0.8% agar. $ZnSO_4 \cdot 7H_2O$ was added to this basic nutrient medium in order to establish ten different zinc treatments: Zn^{2+} was added at concentrations of 0.15, 3, 6, 9, ..., 27 mM. The pH of the final media was adjusted to 4.5.

One week old cultures were used to prepare inocula (mycelial plugs of approximately 0.5 cm^2) and in order for these to be uniform, a large number of plugs were preincubated on cellophane-covered basic medium for 2 or 3 days. Single inocula were then transferred to the cellophane-covered test media and incubated at 23°C in the dark. Three replicates per treatment were prepared and five inocula of each isolate were immediately harvested in order to determine the mean dry weight of the mycelia at the start of the treatment. Mycelia were harvested after 10 days of incubation, frozen at −80°C and freeze-dried before weighing. The dry weight increment was determined and a tolerance index was calculated for each isolate as the percentage of biomass retained on the heavy metal-enriched media compared with growth on the control Fries medium. The EC_{50} concentration (zinc concentration which inhibits growth by 50%) was determined for each isolate.

2.1.3 Results and discussion

In total, 121 samples of *S. luteus* were collected from the nine different sampling sites and all were tested for zinc tolerance. The average tolerance indices

for each *S. luteus* population are listed in table 2.3. Increased levels of zinc tolerance are apparent in *S. luteus* populations inhabiting contaminated environments (Lm, Lk, Lc, Ls, N and Of). These populations are strongly dominated by individuals showing a high tolerance to elevated zinc concentrations (EC_{50} > 6 mM), whereas populations in noncontaminated environments consist of strictly nontolerant individuals (P) or a mixture of tolerant and nontolerant individuals (E: 62% tolerant individuals; Hh: 32% tolerant individuals). Although the EC_{50} values obtained by this *in vitro* screening are probably not representative for a field situation, studies with heavy metal tolerant and nontolerant *S. bovinus* isolates showed that the differential response to elevated concentrations of heavy metal ions is maintained in symbiosis (Adriaensen et al., 2004).

Toxic effects of zinc become apparent in biosensor bacteria at a concentration of 40 μM in soil pore water (Chaudri et al., 1999) and this concentration also reduces root elongation in nontolerant grass ecotypes (Al-Hiyaly et al., 1988). Therefore, the zinc concentration at the contaminated sites included in this screening is high enough to cause toxic effects in plants and microorganisms and strong selection pressure for the tolerance trait can be foreseen. The observed differential response to elevated zinc concentrations of the different *S. luteus* populations is clearly related to the pollution level of their respective habitats and this correlation may be the result of a causal relationship. Moreover, it is believed that the zinc tolerance trait is genetically based as frequent subculturing of tolerant isolates on basic medium did not alter their response to elevated zinc concentrations, thus rendering physiological acclimation unlikely as a cause of the zinc tolerance (Colpaert et al., 2004). Hence, it might be expected that the high frequency of tolerant individuals in populations from contaminated areas is the result of adaptive evolution.

The colonisation of contaminated soils might be initiated by tolerant individuals, present at low frequencies in surrounding populations inhabiting noncontaminated sites. A similar mechanism has been proposed for grass species that show rapid evolution of heavy metal tolerant populations and where tolerant variants have been found in nonexposed populations at frequencies between 0.1 and 0.5% (Bradshaw and McNeilly, 1981; Al-Hiyaly et al., 1993; Schat and Verkleij, 1998). Although the short generation time and frequent sexual reproduction of *S. luteus* would allow rapid selection of adapted genotypes, the hypothesis of nonexposed populations with low frequencies of tolerant individuals cannot be supported with the results of this screening as no tolerance was observed among the isolates collected at the noncontaminated site P. However, if heavy metal tolerant mutants of *S. luteus* are present in nonexposed populations at comparable frequencies as in grass populations, samples of much larger size should be collected from populations at noncontaminated sites in order to increase the chance of obtaining tolerant isolates.

		Zinc concentration (mM)									
Site	#	0.15	3	6	9	12	15	18	21	24	27
Lm	12	100	100 (3)	103 (4)	85 (7)	64 (12)	52 (12)	48 (15)	39 (12)	29 (11)	25 (10)
N	14	100	100 (3)	92 (5)	79 (9)	60 (11)	52 (12)	27 (10)	19 (8)	8 (3)	2 (1)
Lk	3	100	105 (5)	104 (7)	88 (20)	74 (30)	54 (24)	39 (26)	31 (26)	31 (31)	28 (28)
Of	7	100	112 (7)	100 (8)	75 (15)	56 (14)	40 (15)	36 (17)	31 (16)	33 (18)	30 (15)
Ls	20	100	101 (3)	92 (4)	69 (7)	60 (8)	51 (9)	47 (10)	42 (10)	40 (11)	31 (10)
Lc	4	100	96 (2)	97 (3)	71 (20)	68 (23)	69 (24)	63 (24)	58 (28)	51 (30)	36 (20)
E	18	100	86 (7)	71 (9)	44 (9)	25 (8)	18 (7)	4 (3)	1 (1)	0	0
Hh	22	100	52 (8)	33 (8)	21 (7)	14 (5)	10 (5)	n.d.	n.d.	n.d.	n.d.
P	19	100	15 (4)	2 (1)	2 (1)	1 (1)	1 (0)	n.d.	n.d.	n.d.	n.d.

Table 2.3: Average zinc tolerance indices of *Suillus luteus* isolates, sampled from nine different populations in heavy metal contaminated (Lm, Lk, Ls, Lc, N and Of) and noncontaminated (P, E and Hh) regions; the sites are ordered from a high to a low level of pollution; #, number of isolates included in the tolerance test; n.d., not determined; SEM is given in parentheses.

The relatively high frequencies of zinc tolerant isolates collected at the sites E and Hh, despite the low level of pollution of the soils at these sites, probably result from considerable gene flow between *S. luteus* populations. Either the populations at these sites have been under the influence of the zinc smelters and nontolerant individuals migrated into populations that originally consisted mainly of tolerant individuals due to a temporary selection pressure or tolerant individuals migrated into nonexposed populations of predominantly nontolerant individuals. The fact that tolerant isolates are not rapidly outcompeted by nontolerant ecotypes or by other pioneer fungi at nonpolluted sites may suggest a relatively small disadvantage of the tolerance trait, although this disadvantage may be obscured by high mutation rates.

2.2 Transcriptomic characterization of tolerance to heavy metals in *Suillus luteus*

2.2.1 Introduction

In order to maintain cytoplasmic metal concentrations at a physiological level, all organisms possess a variety of homeostatic mechanisms which control the uptake, transport, storage and incorporation of metal ions and which provides a basic tolerance to heavy metal toxicity.

Metal ions are actively imported in the cell by specific transport proteins in the cytoplasmic membrane or enter the cell by passive diffusion driven by a chemiosmotic gradient across the cytoplasmic membrane. In the cytosol, they are bound to chelators (e.g. organic acids, amino acids, phytochelatins and metallothioneins) or metallochaperones (e.g. Cu-chaperones CCS, ATX1 and COX17), which prevents them from exhibiting their toxic effects, and transported to specific cellular compartments or proteins for storage and incorporation. Detoxification occurs through sequestration in the vacuoles or plastids and through export by specific transport proteins in the cytoplasmic membrane.

In general, both posttranslational and transcriptional regulatory mechanisms can function together to maintain optimal levels of metal ions in the cell. For example, expression of the high-affinity zinc uptake gene *ZRT1* in *Saccharomyces cerevisiae* increases in response to zinc limitation, whereas under zinc-replete conditions, Zrt1 undergoes zinc-induced endocytosis and is degraged in the vacuole (Eide, 2003). Two zinc-responsive transcription factors that have been characterized in detail, are Zap1 from *S. cerevisiae* and mammalian MTF-1. In response to zinc deficiency, Zap1 increases the expression of three zinc uptake systems encoded by the *ZRT1*, *ZRT2* and *FET4* genes. Zap1 also stimulates the release of zinc from the vacuole by activating

the expression of the *ZRT3* gene, encoding a vacuolar efflux system (Waters and Eide, 2002; Zhao and Eide, 1996a,b). MTF-1 plays an important role in protecting mammalian cells against zinc toxicity by increasing the expression of *MT-1* and *MT-2*, two genes that encode zinc-binding metallothioneins, and by regulating the expression of a zinc efflux system encoded by *ZnT-1* (Heuchel et al., 1994; Langmade et al., 2000).

There is compelling evidence that sensory mechanisms controlling this homeostasis are provided, at least partly, by transcription factors. Metal ions can bind directly to specific domains of these factors and alter their activity accordingly (Rutherford and Bird, 2004). Both Zap1 and MTF-1 use regulatory zinc finger domains to sense zinc and for MTF-1 it has been shown *in vitro* that full DNA-binding activity is only achieved upon full metallation of these regulatory zinc fingers (Heuchel et al., 1994; Westin and Schaffner, 1988). Other mechanisms, such as signal transduction pathways that involve protein kinases, are possibly of equal importance (Rutherford and Bird, 2004).

Currently, limited knowledge is available on the molecular mechanisms responsible for increased tolerance to heavy metals. However, it is assumed that adaptations of the existing homeostasis mechanisms lead to the evolution of increased tolerance. For example, in *Silene vulgaris* a higher expression level of genes encoding metallothioneins as well as an increase of their copy number in the genome was found to be correlated with an increased tolerance to copper (van Hoof et al., 2001), and in *Thlaspi goesingense* the hypertolerance to nickel was found to be due to a higher uptake by the vacuole (Krämer et al., 2000).

The identification of genes that potentially influence any trait of interest can be approached by an analysis of complementary DNA (cDNA), obtained by reverse transcription of messenger RNA (mRNA). Although mRNA is not the ultimate product of a gene, transcription is the first step in gene expression, and information about the transcript level is valuable for understanding gene function and regulation. Transcriptomics or genome-wide expression profiling is one of the tools that is used to determine the function of genes as well as their spatial and temporal expression patterns and the genetic networks in which they participate, the idea being that genes showing similarity in expression pattern may be functionally related and under the same genetic control mechanism. Common techniques used for genome-wide analysis of gene expression are cDNA microarrays, oligonucleotide microarrays, cDNA-AFLP and serial analysis of gene expression (SAGE).

In this study, cDNA-AFLP was used to reveal differences in gene expression at the mRNA level between heavy metal tolerant and nontolerant strains of *S. luteus* and to identify genes that are potentially involved in the tolerance trait. cDNA-AFLP is a robust and reproducible technique that allows genome-wide expression analysis independent of sequence information, thus rendering it a valuable alternative for microarrays when genomic resources are lacking.

It makes use of the AFLP-technique on a cDNA template in order to generate transcript-derived fragments (TDFs) which reflect the expression pattern of the original mRNAs (Bachem et al., 1996). In short, reverse-transcribed messenger RNA is digested using a frequent and a rare cutting restriction enzyme. Subsequently, enzyme specific oligonucleotide adaptors are ligated to the TDFs and expression patterns are obtained by selective amplification of subsets of the pool of TDFs, followed by polyacrylamide gel electrophoresis and visualization. For a more detailed introduction to the AFLP technique, see section 3.1.1 on page 45.

2.2.2 Material and methods

Fungal material

One tolerant and one nontolerant fungal isolate, sampled from the sites Lm and P respectively (see section 2.1.2 on page 14), were cultured on solid modified Fries medium (Colpaert et al., 2000) at three different concentrations of zinc. The basic solution contained 28 mM glucose, 5.4 mM ammonium tartrate, 1.5 mM KH_2PO_4, 0.4 mM $MgSO_4 \cdot 7H_2O$, 0.3 mM NaCl, 0.2 mM $CaCl_2 \cdot 2H_20$, 4 μM $FeCl_3 \cdot 6H_2O$, 6 μM $MnSO_4 \cdot H_2O$, 0.8 μM $CuSO_4 \cdot 5H_2O$, 56 μM *myo*-inositol, 0.1μM biotin, 0.5μM pyridoxine, 0.3μM riboflavin, 0.8μM nicotinamide, 0.7μM *p*-aminobenzoic acid, 0.3μM thiamine, 0.2μM Ca-pantothenate and 0.8% agar. Zinc was added to the nutrient medium under the form of $ZnSO_4 \cdot 7H_2O$ at concentrations of 0.02 mM, 1.5 mM and 3 mM Zn^{2+} for the different treatments. The pH of the final medium was adjusted to 4.5.

Uniform inocula (0.5 cm^2 plugs of fungal mycelium) were prepared from one week old cultures of both isolates, by preincubation of various plugs of mycelium on cellophane-covered agar plates with basic medium for 2 or 3 days. Single plugs were transferred to the test plates covered with cellophane and were allowed to grow in the dark for 7 days at 23°C. Three replicates were made for each treatment.

RNA extraction and cDNA synthesis

Total RNA was extracted from mycelial tissue using the RNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France) and poly(A)$^+$ RNA was extracted from 250 μg samples of total RNA using Oligotex columns (Qiagen, Courtaboeuf, France), both according to the manufacturer's instructions. Double-stranded cDNA was synthesized starting from 0.5 – 1 μg of poly(A)$^+$ RNA using a SMART^{TM} cDNA Library Construction Kit (Clontech, Palo Alto, CA, USA) following the manufacturer's instructions.

cDNA-AFLP procedure

After purification of the double-stranded cDNA using a QIAquick PCR Purification Kit (Qiagen, Courtaboeuf, France), cDNA-AFLP was performed according to Vos et al. (1995) and Bachem et al. (1996), using the restriction enzymes *Bst*YI and *Mse*I for template production (see figure 2.2).

Restriction digestion and adaptor ligation. Around 0.5 μg of double-stranded cDNA was incubated for 2 h at 60°C with 30 U *Bst*YI in 120 μl 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/μl BSA. Next, 30 U *Mse*I, 3 U T4 DNA ligase, 150 pMol *Mse*I-adaptor and 15 pMol *Bst*YI-adaptor in 45 μl 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 1 mM ATP, 50 ng/μl BSA was added to the mixture and the incubation was continued for 3 h at 37°C. Adaptors were prepared by allowing equimolar amounts of both strands to slowly cool to room temperature after incubation at 95°C for 5 min; adaptors were not phosphorylated (see figure 2.2 for adaptor structure). After ligation, the reaction mixture was stored at −20°C until further use.

cDNA-AFLP reactions. Preamplification of the adapted restriction fragments was performed using *Bst*YI+C and *Mse*I+0 primers (see figure 2.2 for the primer sequences). PCRs were performed in 20 μl reactions containing 30 ng *Bst*YI-primer, 30 ng *Mse*I-primer, 5 μl template-DNA, 0.4 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl and 0.2 mM of all four dNTPs. Preamplification reactions were performed for 25 cycles with the following cycle profile: a 30 s DNA denaturation step at 94°C, a 30 s annealing step at 55°C and a 1 min extension step at 72°C. The preamplification product was diluted 10-fold with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0 for subsequent use. Selective amplification was performed in 20 μl reactions containing 5 ng labeled *Bst*YI-primer, 30 ng *Mse*I-primer, 5 μl of the diluted preamplification product, 0.4 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl and 0.2 mM of all four dNTPs. The labelling reactions were performed in 50 μl 25 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 0.5 mM spermidine-3HCl using 500 ng oligonucleotide primer, 100 μCi [γ-^{33}P]ATP and 10 U T4 polynucleotide kinase. The *Bst*YI- and *Mse*I-primers used for selective amplification each were extended with two selective nucleotides (*Bst*YI+CN/*Mse*I+NN; see figure 2.2) and a touchdown-PCR temperature profile was used: a 30 s DNA denaturation step at 94°C, a 30 s annealing step and a 1 min extension step at 72°C. The annealing temperature in the first cycle was 65°C and this temperature was subsequently reduced each cycle by 0.7°C for the next 12 cycles down to 56°C, which was used as annealing temperature for the remaining 23 cycles.

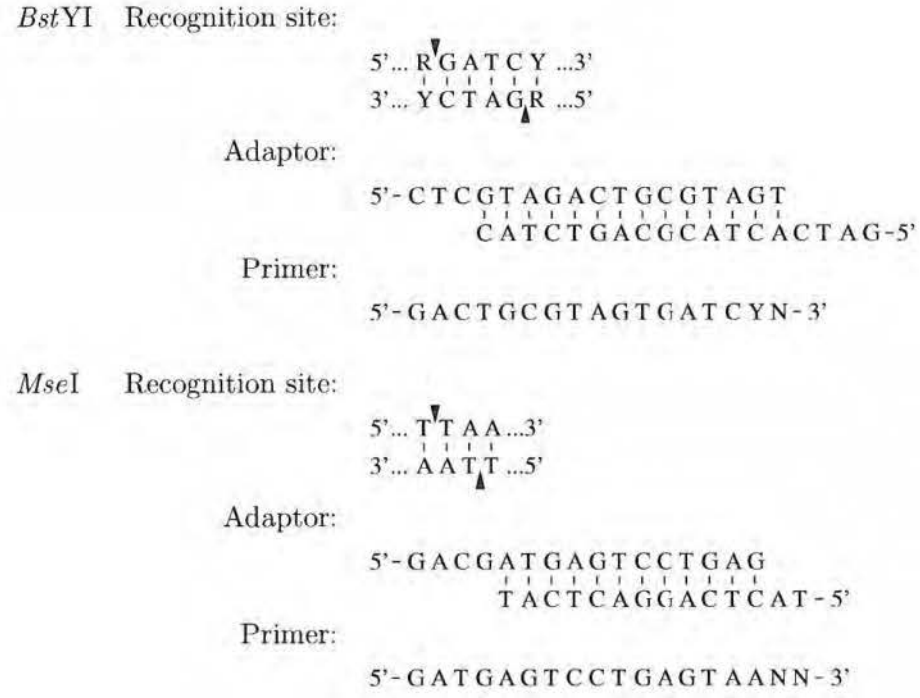

Figure 2.2: Recognition sites, adaptor structures and primer sequences for the restriction enzymes *Bst*YI and *Mse*I used in cDNA-AFLP analysis of heavy metal tolerant and nontolerant *Suillus luteus* isolates; R = G or A; Y = C or T; N = A, C, G or T; arrows indicate cleavage site.

Gel analysis. Selective amplification products were mixed with an equal volume of formamide loading dye (98% formamide, 10 mM EDTA pH 8.0 and bromophenol blue and xylene cyanol as tracking dyes). The resulting mixtures were kept at −20°C overnight, heated for 5 min at 95°C and quickly cooled on ice before loading. Each sample was loaded on a 5% denaturing polyacrylamide gel, prepared using 5% acrylamide, 0.25% methylene bisacryl, 7.5 M urea in 50 mM Tris/50 mM Boric acid/1 mM EDTA. To 100 ml of gel solution, 500 μl of 10% APS and 100 μl TEMED was added and gels were cast using a SequiGen 38 × 50 cm gel apparatus (BioRad Laboratories Inc., Hercules, CA, USA). 100 mM Tris/100 mM Boric acid/2 mM EDTA was used as running buffer. Electrophoresis was performed at constant power, 100 W, for approximately 2.5 h. After electrophoresis, gels were dried under vacuum at 80°C, exposed to RX-U radiography film (Fuji Photo Film, Tokyo, Japan) for approximately 48 h and scanned with a PhosphoImager (Amersham Biosciences, Little Chalfont, UK).

Characterization of cDNA-AFLP fragments. Bands corresponding to differentially expressed transcripts were cut out from the gel and the DNA was eluted by incubation in 200 μl TE buffer for 2 h at room temperature. The eluted DNA was purified using a Purelink^{TM} 96 PCR Purification Kit (Invitrogen Corporation, Carlsbad, CA, USA) and reamplified under the same conditions used for selective amplification. The amplified fragments were ligated into pCR®2.1-TOPO®vectors (Invitrogen Corporation, Carlsbad, CA, USA) according to the manufacturer's instructions and the plasmids were transformed into *Escherichia coli*, TOP10 strain (Invitrogen Corporation, Carlsbad, CA, USA). Plasmid inserts from four positive clones were PCR amplified in 20 μl 0.5 μM M13f-primer, 0.5 μM M13r-primer, 1.5 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 2 mM of all four dNTPs. After an initial cell lysis at 95°C for 5 min, PCR was performed with the following temperature profile: 35 cycles of 95°C for 20 s, 50°C for 20 s and 72°C for 1.5 min. The PCR products were diluted 20 times and 1 μl was used in a 20 μl sequencing reaction with the M13f-primer using ABI BigDye®Terminator reaction mix (Applied Biosystems, Foster City, CA, USA), following the manufacturer's instructions. The sequences were run on an ABI PRISM®3100-Avant Genetic Analyzer. The local BLAST function in the BioEdit software package (Hall, 1999) was used to identify redundant TDFs and database searches were performed using the BLAST Network Service (NCBI, National Center for Biotechnology Service; http://www.ncbi.nlm.nih.gov/BLAST). The functions of function-known genes by BLASTX and TBLASTX sequence alignments (Altschul et al., 1997) with the nonredundant (nr) databases were classified according to their puta-

*Bst*YI-primer +	*Mse*I-primer +
CA	AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG, GT, TA, TC, TG, TT
CC	AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG, GT, TA, TC, TG, TT
CG	AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG, GT, TA, TC, TG, TT
CT	AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG, GT, TA, TC, TG, TT

Table 2.4: Overview of the cDNA-AFLP primer combinations used for selective amplification of the TDFs.

tive function as described in the *Saccharomyces cerevisiae* functional catalogue (http://mips.gsf.de/genre/proj/yeast/).

RT-PCR analysis

In order to verify cDNA-AFLP expression patterns, RT-PCR analysis of several differentially regulated genes and a nonregulated control gene was performed using cDNA from the original *S. luteus* isolates as template. Based on the sequence of the recovered TDFs, gene-specific primers were designed using PRIMER3 software (Rozen and Skaletsky, 1996) and PCR was performed with equal amounts of template cDNA (approximately 25 ng) in 20 μl reactions containing 0.5 μM of each primer, 1 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 200 μM of all four dNTPs. The temperature profile of the PCR was the following: initial denaturation at 95°C for 2 min; 30 cycles of 95°C for 30 s, primer-specific annealing temperature for 30 s and 72°C for 1 min, followed by a final extension at 72°C for 10 min. Amplified PCR products (10 μl) were electrophoresed on a 2% (w/v) agarose gel and visualized by ethidium bromide staining.

2.2.3 Results

cDNA was synthesized from RNA extracted from a heavy metal tolerant and a nontolerant *S. luteus* isolate, grown on solid medium with three different concentrations of zinc (0.02 mM, 1.5 mM and 3 mM Zn^{2+}). A clear differential response to elevated zinc concentrations by both isolates was observed, as growth rate of the nontolerant isolate was significantly more restricted at zinc

concentrations of 1.5 mM and 3 mM when compared to growth rate of the tolerant isolate. cDNA expression patterns were identified by selective amplification using 64 of the 128 possible primer combinations (see table 2.4) and 59 revealed transcript derived fragments (TDFs) with differential expression patterns (see figure 2.3). Fourty to seventy TDFs were obtained per primer combination and 213 differentially expressed fragments were identified. One hundred and fourty-three TDFs were exclusively expressed either in the tolerant isolate (72 TDFs, 70 of which constitutively and 2 repressed at elevated zinc concentration) or in the nontolerant isolate (71 TDFs, 67 of which constitutively and 4 induced at elevated zinc concentration). The remaining TDFs were expressed in the tolerant and the nontolerant isolate, either constitutively in both but at different expression levels (38 TDFs) or regulated in response to an elevated zinc concentration (32 TDFs). In the tolerant isolate, four TDFs were induced at elevated zinc concentrations and ten repressed, whereas in the nontolerant isolate, seven were induced and three repressed at increased zinc concentrations.

One hundred and ninety-seven TDFs were succesfully recovered from the gel matrix and PCR-amplified. The recovered fragments, ranging in length between 50 and 450 bp, were cloned before sequencing in order to prevent problems associated with direct sequencing of PCR-products (Ditt et al., 2001; Durrant et al., 2000) and correct sequences were obtained for 156 TDFs. Due to the presence of comigrating fragments, the correct sequence of 41 TDFs could not be identified unambiguously. Based on pairwise comparisons among the 156 sequences of the recovered fragments, 144 TDFs were found to represent potentially different genes. Redundancy was due to the presence of highly homologous sequences, resulting from mispriming during PCR amplification or representing either different alleles from the same gene or different instances of multicopy genes. Bachem et al. (1996) reported the occurence of TDFs with the same mobility and expression pattern in fingerprints obtained with primers having similar sequence extensions, e.g. one base exchange at either end, and showed this to be due to mismatch primed PCR of highly abundant transcripts. This was the case for TDFs 109, 117, 118, 124 and 140 which resulted from mismatch primed PCR of respectively TDFs 47, 57, 58, 41 and 71 (see figure 2.4).

A second, highly homologous TDF with a different expression and/or mobility pattern was found for each of seven different fragments. These TDFs may represent either different alleles of a gene or different instances of a multicopy gene. Highly homologous fragments with different mobility patterns were found for the TDFs 18, 38, 57, 67, 90, 104 and 132 and sequence analysis revealed that length differences of the homologous TDFs were either due to internal deletion events (TDF pairs 38/42 and 57/58; see figure 2.5) or to mutations introducing new restriction sites (TDF pairs 18/138, 67/68, 71/76 and

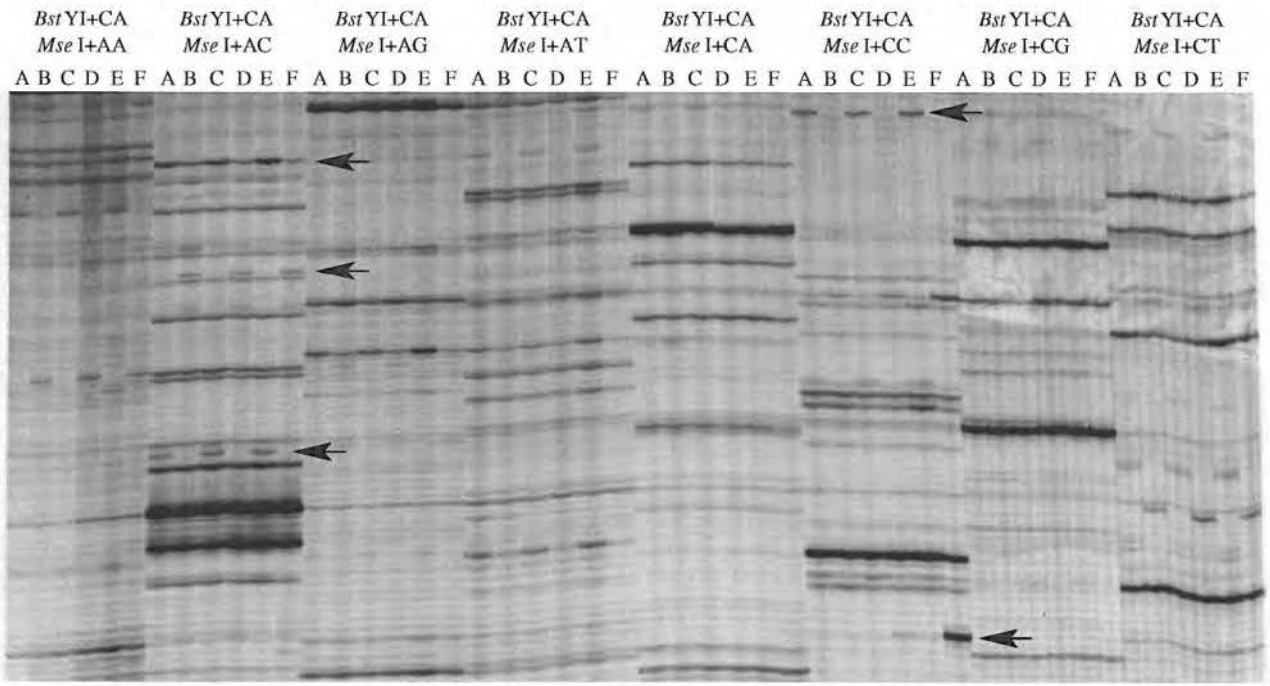

Figure 2.3: Example of an autoradiograph obtained by cDNA-AFLP analysis of a heavy metal tolerant and a nontolerant *Suillus luteus* isolate. The primer combinations that were used to generate the transcriptome fingerprints are indicated above the lanes; A, tolerant isolate grown on agar medium with 0.02 mM Zn^{2+}; B, nontolerant isolate grown on agar medium with 0.02 mM Zn^{2+}; C, tolerant isolate grown on agar medium with 1.5 mM Zn^{2+}; D, nontolerant isolate grown on agar medium with 1.5 mM Zn^{2+}; E, tolerant isolate grown on agar medium with 3 mM Zn^{2+}; F, nontolerant isolate grown on agar medium with 3 mM Zn^{2+}; arrows indicate TDFs with differential expression patterns.

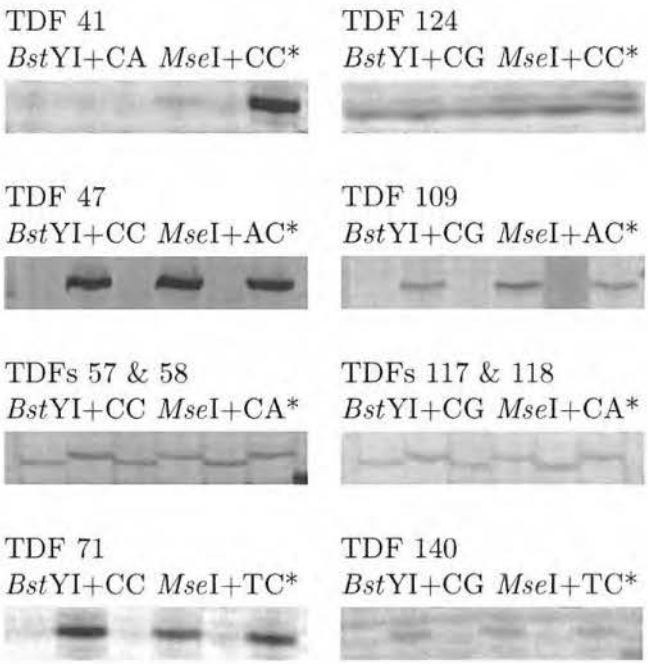

Figure 2.4: Redundant TDFs obtained by cDNA-AFLP analysis of heavy metal tolerant and nontolerant isolates of *Suillus luteus*; redundancy is due to mismatch primed PCR of abundant transcripts; (*), the combination of primers used to PCR amplify the TDF.

90/91; see figure 2.5). Two homologous fragments, TDFs 123 and 139, were of equal length but differed in the nucleotides of the selective extension (see figure 2.6) and since their expression patterns correspond to differently expressed genes (TDF 123 is constitutively expressed only in the tolerant isolate and TDF 139 is expressed in both isolates, constitutively in the nontolerant isolate and repressed at 3 mM Zn^{2+} in the tolerant isolate; see figure 2.6), they are expected to represent differently expressed instances of a multicopy gene.

In order to identify sequences in the nonredundant databases showing significant homology, a threshold value of 1E-05 for the *E*-value is often used. However, the *E*-value also considers the length of the fragments (e.g. short fragments (< 100 bp) showing high levels of homology often have a higher *E*-value than longer fragments showing less homology) and therefore homology was evaluated taking into account the percentages of identities and positives returned by the BLASTX and TBLASTX search algorithms.

Fourty-seven of the 144 nonredundant TDFs showed homology with genes of known function, 22 showed homology with function-unknown genes and 75 showed no homology by BLAST searches (see table 2.6). Some TDFs showed homology with different parts of the same gene and are either derived from the same transcript or represent different alleles or different instances of a multicopy gene: TDFs 27 and 47, homologous with glutathione S-transferase, TDFs 28 and 133, homologous with ARF small monomeric GTPase, TDFs 63 and 156, homologous with NAD+ dependent aldehyde dehydrogenase, and TDFs 81 and 100, homologous with arginine-tRNA ligase. Of the 43 nonredundant function-known genes, twenty-two (51%) are involved in cellular metabolism and energy production, three (7%) in transcription and six (14%) in protein synthesis. Other genes have functions in cellular transport (7%), signal transduction (7%), cell rescue and defense (12%) and cellular organisation (2%).

In order to validate the cDNA-AFLP expression patterns, 8 differentially expressed TDFs were selected for RT-PCR analysis using cDNA from the original isolates as template (see table 2.5 for the primer combinations). The comparisons showed that 6 of the 8 cDNAs examined had the same expression profiles as revealed by cDNA-AFLP (see figure 2.7). A different expression pattern was obtained for TDFs 89 and 134. In case of TDF 134, which was absent in the cDNA-AFLP profile of the nontolerant isolate while RT-PCR showed the corresponding gene to be expressed, the difference could be due to sequence divergence of the corresponding genes in both isolates. However, in general cDNA-AFLP is shown to be a reliable method for identifying differentially expressed genes in *S. luteus* exposed to heavy metal stress.

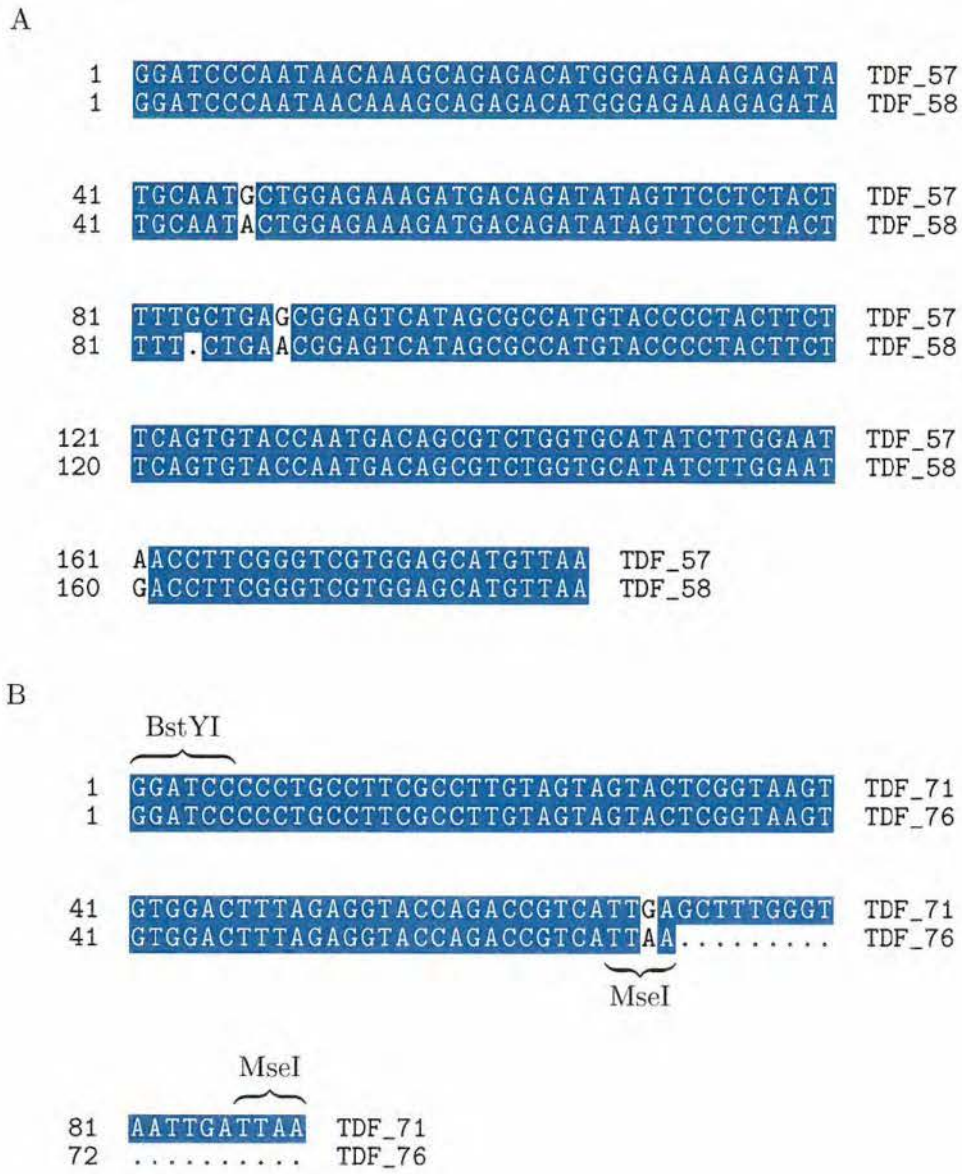

Figure 2.5: Example of homologous TDFs with different mobility patterns due to internal deletion events (A) or mutations introducing new restriction sites (B); the TDFs were obtained by cDNA-AFLP analysis of a heavy metal tolerant and a nontolerant isolate of *Suillus luteus*.

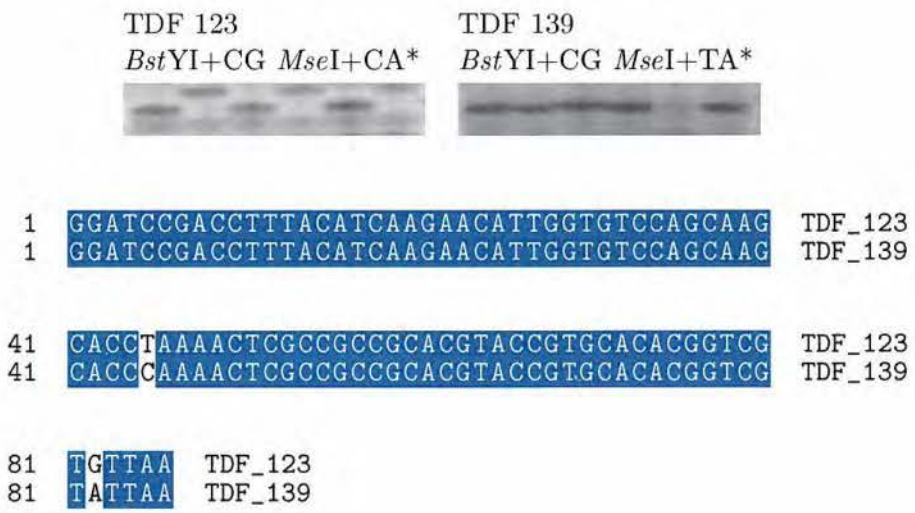

Figure 2.6: Two homologous TDFs with expression patterns that correspond to differently expressed instances of a multicopy gene; the TDFs were obtained by cDNA-AFLP analysis of a heavy metal tolerant and a nontolerant isolate of *Suillus luteus*; (*), the combination of primers used to PCR amplify the TDF.

TDF	Primer sequence (5' → 3')	T_a (°C)	Product size (bp)
C	F: TGAAGGCAAAGTCACGAATG	56	247
	R: TTCCCGAGCATTGATGTCTAC		
16	F: CCCGTCTCCTCAGCATTATC	56	250
	R: GGAACACCCAAATCATCGAC		
27	F: GCAGTGGTATCAACGCAGAG	55	245
	R: TGCTCGACGAACATCCATAC		
61	F: ATCCTGGAAGACGGCTACAC	56	222
	R: GCACAATCACAAACGTGGTC		
66	F: CAACCCAAGATCATTCATGC	56	214
	R: GCCTGGACTGAAACCATGTC		
89	F: ACTTGCAGGCGCCTTTACAC	57	213
	R: TTCCCAAGTCCCTGTTTCTG		
128	F: CGCACTGTATTCGTAGTTGA	53	139
	R: GCGTGGTTATGCGATGTTTA		
134	F: ACTCCTGAACGACCCAACC	57	223
	R: AAGGCTGGCCTTGAGAAGC		
147	F: ATCCGATCGTCACGACTGC	55	100
	R: TTAAGGTACTGGAGGTAGAAAGCA		

Table 2.5: Primer combinations used for RT-PCR verification of differentially expressed cDNA-AFLP fragments in heavy metal tolerant and nontolerant *Suillus luteus* isolates; T_a, annealing temperature.

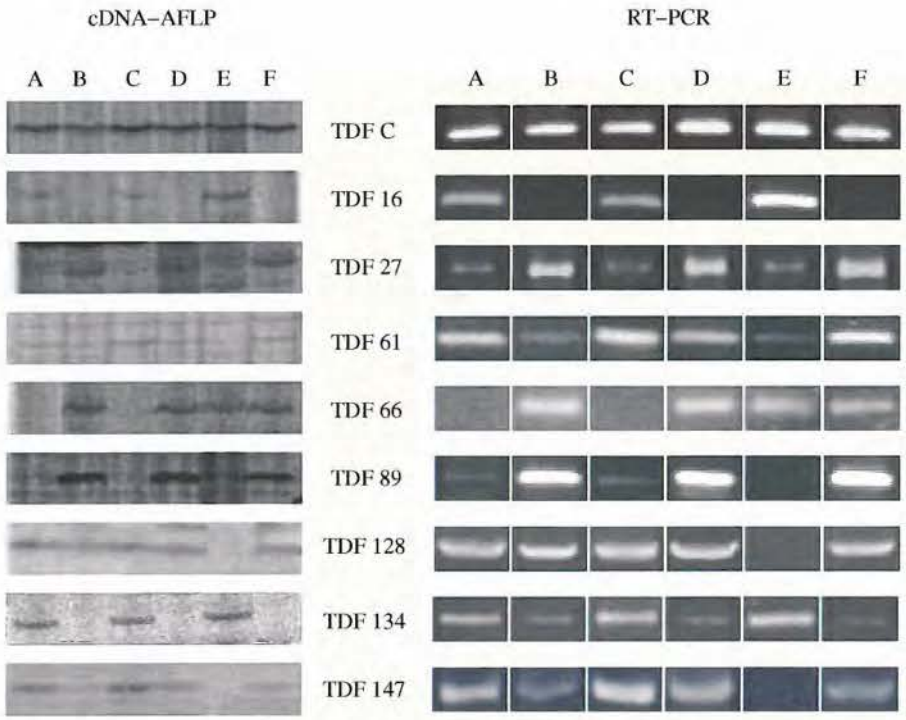

Figure 2.7: Comparison of expression patterns of 8 differentially expressed genes and one constitutively expressed control gene (TDF C) by both cDNA-AFLP and RT-PCR; A, tolerant isolate grown on agar medium with 0.02 mM Zn^{2+}; B, nontolerant isolate grown on agar medium with 0.02 mM Zn^{2+}; C, tolerant isolate grown on agar medium with 1.5 mM Zn^{2+}; D, nontolerant isolate grown on agar medium with 1.5 mM Zn^{2+}; E, tolerant isolate grown on agar medium with 3 mM Zn^{2+}; F, nontolerant isolate grown on agar medium with 3 mM Zn^{2+}.

TDF	Size (bp)	Expression: Tolerant A	Tolerant B	Tolerant C	Nontolerant A	Nontolerant B	Nontolerant C	Homology	Function category	Id. (%)	Pos. (%)	*E*-value
1	195	■	■	■	□	□	□	Membrane protein	Unknown	45	64	3.20E+00
2	160	■	■	■	▪	▪	▪	Aconitate hydratase	Metabolism and energy	64	71	1.00E−09
6	286	■	■	■	■	▪	□	Predicted protein (*Neurospora crassa*)	Unknown	35	52	5.50E+00
9	240	▪	▪	▪	■	■	■	Taka-amylase	Metabolism and energy	51	62	1.00E−13
10	106	■	■	■	▪	▪	▪	Farnesyl diphosphate synthetase	Metabolism and energy	77	90	2.00E−06
11	105	■	■	■	□	□	□	6,7-dimethyl-8-ribityllumazine synthase	Metabolism and energy	60	82	5.00E−01
14	169	■	■	■	□	□	□	Ubiquitin carboxyl-terminal hydrolase	Metabolism and energy	43	68	1.20E−02
16	371	▪	▪	■	□	□	□	Putative copper ion transporter	Transport	39	50	6.00E−13
20	183	□	□	□	■	■	■	FUSCA protein FUS6	Signal transduction	37	62	2.50E+00
21	135	■	■	■	□	□	□	Predicted protein (*Aspergillus nidulans*)	Unknown	41	65	1.50E−02
27	183	▪	▪	▪	■	■	■	Glutathione S-transferase	Cell rescue and defense	57	68	2.00E−12
28	159	■	■	■	□	□	□	ARF small monomeric GTPase	Transport	84	96	6.00E−20
29	131	□	□	□	■	■	■	Cytochrome P450	Metabolism and energy	47	67	2.60E−02
30	226	■	■	■	▪	▪	▪	Proteasome subunit beta type 2	Metabolism and energy	63	81	1.00E−20
31	139	■	■	■	□	□	□	G protein subunit alpha	Signal transduction	60	84	2.00E−12
34	211	□	□	□	■	■	■	Expressed protein (*Cryptococcus neoformans*)	Unknown	42	59	6.30E−01
36	69	□	□	□	▪	▪	■	ATP synthase β	Metabolism and energy	82	95	1.00E−03
39	279	■	■	■	▪	▪	▪	Predicted protein (*Magnaporthe grisea*)	Unknown	35	59	1.00E−03
43	89	■	■	■	□	□	□	Predicted protein (*Ustilago maydis*)	Unknown	72	82	2.00E−03
44	193	■	■	■	□	□	□	Predicted protein (*Cryptococcus neoformans*)	Unknown	80	85	1.00E−12
46	94	■	■	■	□	□	□	Endoplasmic reticulum protein	Unknown	67	83	2.00E−05
47	280	□	□	□	■	■	■	Glutathione S-transferase	Cell rescue and defense	71	79	8.00E−31
50	153	■	■	■	□	□	□	Predicted protein (*Ustilago maydis*)	Unknown	54	70	3.00E−07
51	123	■	■	■	□	□	□	Predicted protein (*Cryptococcus neoformans*)	Unknown	77	92	3.00E−11
57	177	□	□	□	■	■	■	Fumarate reductase	Metabolism and energy	66	85	8.20E−01
59	159	■	■	■	□	□	□	Glyceraldehyde-3-phosphate dehydrogenase	Metabolism and energy	92	98	3.00E−20
60	315	□	□	□	■	■	■	Hydrolase	Metabolism and energy	46	63	2.00E−12

Continued on next page

Table 2.6: Function analysis of differentially expressed TDFs and their expression patterns obtained by cDNA-AFLP of a heavy metal tolerant and a nontolerant *S. luteus* isolate grown at different zinc concentrations; A, 0.02 mM Zn^{2+} treatment; B, 1.5 mM Zn^{2+} treatment; C, 3 mM Zn^{2+} treatment; Expression levels, relative to the highest level observed for each TDF independently, are indicated by ■ (highest level) > ▪ (reduced expression) > □ (no expression); Id., BLAST identities; Pos., BLAST positives.

TDF	Size (bp)	Expression: Tolerant A	Tolerant B	Tolerant C	Nontolerant A	Nontolerant B	Nontolerant C	Homology	Function category	Id. (%)	Pos. (%)	*E*-value
61	261	■	■	■	■	■	■	Ketol-acid reductoisomerase	Metabolism and energy	89	96	3.00E−38
62	174	□	□	□	■	■	■	Transcription factor btf3 homologue	Transcription	58	79	4.00E−09
63	102	□	□	□	■	■	■	Aldehyde dehydrogenase (NAD+)	Metabolism and energy	60	72	2.00E−04
66	223	□	□	■	■	■	■	Predicted protein (*Cryptococcus neoformans*)	Unknown	68	89	1.00E−08
73	159	■	■	■	□	□	□	Predicted protein (*Cryptococcus neoformans*)	Unknown	69	82	4.00E−12
81	266	■	■	■	□	□	□	Cytoplasmic arginyl-tRNA synthetase	Protein synthesis	54	80	9.00E−22
83	162	■	■	■	□	□	□	Nitrilase NIT3	Metabolism and energy	71	85	1.30E−01
84	112	□	□	□	■	■	■	Predicted protein (*Aspergillus nidulans*)	Unknown	72	88	3.20E+00
85	362	■	■	■	□	□	□	Hydrophobin-3 precursor	Cellular organisation	38	54	2.00E−12
87	148	■	■	■	□	□	□	60s ribosomal protein	Protein synthesis	68	78	4.00E−10
89	259	■	■	■	■	■	■	Metallothionein	Cell rescue and defense	55	73	4.00E−06
98	219	■	■	■	□	□	□	Predicted protein (*Ustilago maydis*)	Unknown	40	61	5.00E−07
100	254	□	□	□	■	■	■	Arginine-tRNA ligase	Protein synthesis	49	72	1.00E−13
104	99	■	■	■	■	■	■	Enolase	Metabolism and energy	75	84	2.00E−07
106	110	■	■	■	■	■	■	CTP synthase	Metabolism and energy	88	97	2.00E−12
107	118	■	■	■	□	□	□	Dihydroxy-acid dehydratase	Metabolism and energy	82	89	6.00E−10
109	146	■	■	■	■	■	■	Myc1	Transcription	64	75	5.00E−08
111	217	□	□	□	■	■	■	Threonine synthase	Metabolism and energy	43	59	1.00E−06
112	113	■	■	■	■	■	■	expressed protein (*Yarrowia lipolytica*)	Unknown	66	74	2.00E−02
113	366	■	■	■	□	□	□	Predicted protein (*Neurospora crassa*)	Unknown	32	47	3.00E−04
115	99	■	■	■	■	■	■	60s ribosomal protein l11	Protein synthesis	84	96	1.00E−09
116	295	■	■	■	□	□	□	Postsynaptic protein CRIPT	Signal transduction	50	64	6.00E−12
120	158	■	■	■	■	□	■	Predicted protein (*Neurospora crassa*)	Unknown	43	67	8.40E−01
121	114	■	■	■	■	■	■	Predicted protein (*Aspergillus nidulans*)	Unknown	39	58	7.10E+00
123	77	■	■	■	□	□	□	60s ribosomal protein l17	Protein synthesis	88	92	1.00E−04
126	164	□	□	□	■	■	■	Predicted protein (*Giberella zeae*)	Unknown	61	75	2.00E−05
128	142	■	■	□	■	■	■	Thioredoxin peroxidase	Cell rescue and defense	87	95	5.00E−17
129	120	■	■	■	■	■	■	Proteasome component pts1	Metabolism and energy	89	97	3.00E−11
131	143	□	□	□	■	■	■	Nicotinamidase	Metabolism and energy	44	76	1.00E−01

Continued on next page

Table 2.6: Function analysis of differentially expressed TDFs and their expression patterns obtained by cDNA-AFLP of a heavy metal tolerant and a nontolerant *S. luteus* isolate grown at different zinc concentrations; A, 0.02 mM Zn^{2+} treatment; B, 1.5 mM Zn^{2+} treatment; C, 3 mM Zn^{2+} treatment; Expression levels, relative to the highest level observed for each TDF independently, are indicated by ■ (highest level) > ■ (reduced expression) > □ (no expression); Id., BLAST identities; Pos., BLAST positives.

		Expression									
	Size	Tolerant			Nontolerant				Id.	Pos.	
TDF	(bp)	A	B	C	A	B	C	Homology	Function category	(%)	(%)	*E*-value
132	313	□	□	□	■	■	■	Mitochondrial translocase subunit tim8	Transport	51	80	5.00E−20
133	276	■	■	■	■	■	■	ARF small monomeric GTPase	Transport	95	98	1.00E−44
134	232	■	■	■	□	□	□	Heat shock protein HSP60	Cell rescue and defense	87	94	1.00E−30
135	85	■	■	■	□	□	□	Peptidyl-prolyl cis-trans isomerase	Transcription	68	77	5.40E+00
136	237	■	■	□	□	■	■	Translation elongation factor EF1-alpha	Protein synthesis	98	100	2.00E−37
137	212	■	■	■	□	□	□	40S ribosomal protein S12	Protein synthesis	68	79	4.00E−13
142	288	■	■	■	■	■	■	Predicted protein (*Gibberella zeae*)	Unknown	30	48	5.00E−04
144	133	■	■	■	■	■	■	Predicted protein (*Magnaporthe grisea*)	Unknown	33	60	1.70E−01
147	102	■	■	□	■	■	■	Manganese-superoxide dismutase	Cell rescue and defense	90	93	6.00E−10
150	285	■	■	■	■	■	■	Predicted protein (*Cryptococcus neoformans*)	Unknown	36	54	1.60E−02
152	283	■	■	■	■	■	■	Retinal short-chain dehydrogenase/reductase	Metabolism and energy	40	60	1.00E−12
155	213	□	□	□	■	■	■	Cyanate lyase	Metabolism and energy	51	73	6.00E−12
156	221	□	□	□	■	■	■	Aldehyde dehydrogenase (NAD+)	Metabolism and energy	56	69	2,00E−16

Table 2.6: Function analysis of differentially expressed TDFs and their expression patterns obtained by cDNA-AFLP of a heavy metal tolerant and a nontolerant *S. luteus* isolate grown at different zinc concentrations; A, 0.02 mM Zn^{2+} treatment; B, 1.5 mM Zn^{2+} treatment; C, 3 mM Zn^{2+} treatment; Expression levels, relative to the highest level observed for each TDF independently, are indicated by ■ (highest level) > ■ (reduced expression) > □ (no expression); Id., BLAST identities; Pos., BLAST positives.

2.2.4 Discussion

Estimations suggest that the genome of filamentous fungi, including ectomycorrhizal fungi, usually varies between 20 and 40 Mb, with a complement of about 8000 genes (Kupfer et al., 1997; Le Quéré et al., 2002). Using this estimation for the number of genes in *S. luteus*, the present set of 140 nonredundant TDFs corresponds to approximately 2% of the total expected complement of genes. A total of 65 TDFs (46%) corresponded to genes homologous to GenBank entries, including genes of known function as well as hypothetical proteins. The remaining TDFs (54%) did not have significant matches within the GenBank databases and may represent novel fungal genes, although this could also be partly due to the sequences being derived from either the 3' or the 5' untranslated mRNA regions. This relatively high proportion of TDFs showing no homology to GenBank entries falls within the range of $50 - 65\%$ reported for ESTs in other filamentous fungi (Ospina-Giraldo et al., 2000; Skinner et al., 2001; Lee et al., 2002) and might reflect the low number of database entries of fungal origin.

Considerable variation in transcript profiles was revealed between the tolerant and the nontolerant *S. luteus* isolate, both at physiological zinc concentration and at increased zinc concentrations. Although in several cases the differential expression pattern was shown to be caused by sequence divergence of the corresponding transcripts, RT-PCR analysis verified the differential transcript profiles to be mainly due to differential gene expression in both isolates. Adaptive evolution by natural selection will most likely explain this difference in gene expression to a certain extent, but other causes, including genetic hitchhiking and natural regulatory variation, may be equally important. In genomic regions with a low recombination rate, positive selection of alleles responsible for the tolerance trait and modifiers that enhance the tolerance may be associated with a reduction of the genetic variation around the selected sites (Maynard Smith and Haigh, 1974) and this signature of selection may be important at the genome-wide level and more common than usually accepted (Quesada et al., 2003). Besides the variation due to this genetic hitchhiking effect, part of the observed differential expression in both *S. luteus* isolates will be due to variation in gene expression level present in natural populations. Although natural regulatory variation has been shown to be extensive in *Saccharomyces cerevisiae* (Townsend et al., 2003), transcript profiles of more isolates should be analyzed in order to assess its importance in *S. luteus*.

Among the differentially expressed genes, five showed significant similarity with genes encoding proteins involved in cell rescue and defense: glutathione S-transferase (TDFs 27/47), metallothionein (TDF 89), thioredoxin peroxidase (TDF 128), heat shock protein HSP60 (TDF 134) and manganese

superoxide-dismutase (TDF 147). Three of these proteins, namely glutathione S-transferase, thioredoxin peroxidase and manganese superoxide-dismutase, are known to be responsive to oxidative stress. Reactive oxygen species, such as hydroxyl radical (·OH), superoxide radical ($\cdot O_2^-$) and hydrogen peroxide (H_2O_2), are generated during the successive reduction of O_2 to H_2O and exposure to heavy metal ions may shift the balance of free radical metabolism towards the production of reactive oxygen species. Enhanced generation of reactive oxygen species may result in oxidative stress and induce damage to lipids, proteins and DNA, thus causing cellular dysfunctioning. Aerobic organisms have evolved defense mechanisms against the deleterious effects of reactive oxygen species, consisting of enzymes (e.g. superoxide dismutase, catalase, peroxidase and glutathione S-transferase), as well as nonenzymatic reductants and free radical scavengers (e.g. glutathione and phenolics). Thioredoxin peroxidase reduces organic hydroperoxides and H_2O_2 to harmless products, using thioredoxin as the electron donor, and is induced in response to oxidative stress (Godon et al., 1998; Lee et al., 1999). Glutathione S-transferase is a general detoxification enzyme that covalently binds toxic compounds to glutathione, forming a less toxic glutathione S-conjugate that can be transported into and sequestered by the vacuole, and that converts H_2O_2 to H_2O by oxidizing reduced glutathione (peroxidase activity). Toxic products of lipid peroxidation and DNA damage, caused by reactive oxygen species, are metabolized by glutathione S-transferase (Marrs, 1996) and overexpression of glutathione S-transferase in transgenic tobacco seedlings has been shown to increase the tolerance to oxidative stress (Roxas et al., 2000). Manganese superoxide-dismutase protects cells against oxidative stress by catalyzing the conversion of superoxide radicals to hydrogen peroxide and oxygen. However, expression analysis in *Paxillus involutus* revealed no change in manganese superoxide-dismutase transcript levels following cadmium treatment and post-translational mechanisms are believed to be involved in activating a pool of pre-existing superoxide-dismutase in response to cadmium stress (Jacob et al., 2001). Heat shock proteins are induced under a variety of stress factors which cause protein unfolding, misfolding or aggregation (e.g. heat, UV radiation and heavy metal ions), and are capable of re-establishing the balance between protein synthesis, assembly and degradation. Many members of the HSP60 family are constitutively expressed and are essential mitochondrial chaperones that promote the folding of proteins imported in the mitochondrial matrix and that direct several proteins to the intermembrane space. HSP60 is believed to protect enzymes possessing Fe/S clusters, such as aconitase and succinate dehydrogenase, from oxidative inactivation and release of their iron, thus preventing additional oxidative stress (Cabiscol et al., 2002).

Although induction of several of these genes by oxidative stress or heavy metal ions has been shown in various organisms, no increase in transcript

levels was observed in *S. luteus* following zinc treatment. On the contrary, transcriptional repression was observed in the tolerant isolate for manganese superoxide-dismutase and thioredoxin peroxidase. Vallino et al. (2004) reported a similar observation in the ericoid mycorrhizal fungus *Oidiodendron maius*, where EST analysis revealed no differential response of stress-related genes following zinc treatment, and proposed this to be due to the applied zinc concentration not being high enough to induce severe toxic effects. However, growth of the nontolerant *S. luteus* isolate was clearly inhibited by both zinc treatments, rendering their nontoxicity unlikely. Possibly, posttranslational mechanisms control these defense systems, as proposed for manganese superoxide-dismutase, or other systems dominate the stress response.

Relatively few genes were found to be responsive to heavy metal induced stress at the transcript level, but a clear difference in transcriptional response to elevated zinc concentrations was apparent between the tolerant and the nontolerant isolate. In the tolerant isolate, a transcriptional response following zinc treatment was in general only induced at the highest zinc concentration, whereas in the nontolerant isolate zinc responsive modulation of gene expression was already observed at a lower zinc concentration. Both isolates also differed in the set of genes responsive to increased zinc concentrations and in the way that transcription was affected, repression being the main effect in the tolerant isolate and induction in the nontolerant isolate. In the nontolerant isolate, elevated zinc concentrations induced the expression of homologues of ATP synthase β, ketol-acid reductoisomerase and translation elongation factor 1, respectively involved in energy production, amino acid biosynthesis and protein synthesis. Increased transcription levels have been shown for ketol-acid reductoisomerase following heat shock and for mitochondrial ATP synthase and translation elongation factor 1 in response to heavy metal induced stress, but the functional significance of this increase remains unclear (Rosen et al., 2002; Hamilton et al., 2001; Joseph et al., 2002). Among the genes of which the transcription is downregulated in the tolerant isolate, two encode homologues of thioredoxin peroxidase and manganese superoxide-dismutase. The repression of these oxidative stress-related genes coincides with an increased transcription of a putative metal transporter, homologous to the low-affinity copper transporter Ctr2 in the vacuolar membrane of *S. cerevisiae* (Rees et al., 2004). The increased transcription level of this putative metal transporter may indicate a role in the detoxification of zinc and the transcriptional downregulation of manganese superoxide-dismutase and thioredoxin peroxidase could follow from the associated reduction of the level of oxidative stress. Expression levels of homologues of ketol-acid reductoisomerase and translation elongation factor 1, which were upregulated in the nontolerant isolate, were downregulated in the tolerant isolate in response to increased zinc concentrations. Other downregulated genes in the tolerant isolate encoded for homologues of Myc1

and a 60S ribosomal protein, respectively involved in transcription and protein synthesis.

Besides absence of transcription of the putative metal transporter in the nontolerant isolate, other differences in gene expression levels may partly account for the difference in heavy metal tolerance between both isolates. Several genes that were constitutively expressed at higher levels in the tolerant isolate encode for homologues of proteins that may provide additional protection against heavy metal induced stress and associated oxidative stress. Among the stress related genes, expression of a HSP60 homologue was higher in the tolerant isolate, whereas the expression levels of homologues of thioredoxin peroxidase, glutathione S-transferase and metallothionein were equal or reduced compared to their expression levels in the nontolerant isolate. Because expression of manganese superoxide-dismutase was downregulated following the 3 mM zinc treatment, its importance for the different tolerance level is believed to be limited, although it may play a role at lower zinc concentrations. Several genes encoding homologues of proteins involved in the ubiquitin-dependent degradation of proteins, which provides protection against potentially cytotoxic damaged or aberrant proteins, were constitutively expressed at higher levels in the tolerant isolate. The principal mechanism of ubiquitin-dependent proteolysis involves the covalent ligation of ubiquitin to other proteins, which can then be degraded by the 26S protease, an ATP-dependent multicatalytic protease. The genes that were differentially expressed in both isolates encode homologues of ubiquitin carboxyl-terminal hydrolase, which is involved in the regeneration of ubiquitin after protein degradation, and components of the proteasome complex. It has been shown that elevated expression of genes encoding for proteasome subunits induces enhanced tolerance to oxidative stress (Chondrogianni et al., 2005). Furthermore, the hydrophobin homologue, which was only expressed in the tolerant isolate, may be related to the different response to elevated zinc concentrations in the tolerant isolate. Hydrophobins are water-repellent proteins in the cell wall, which may protect against indiscriminate intracellular penetration of heavy metal-containing solutions (Pawlik-Skowrońska et al., 2002). Thus, the presence of hydrophobins in the cell wall of the tolerant isolate may be partly responsible for the higher external zinc concentration needed for a heavy metal induced modulation of transcription in this isolate.

Chapter 3

Genetic structure of *Suillus luteus* populations in heavy metal contaminated and noncontaminated areas

3.1 Introduction

The gene pool of nearly all natural populations shows genetic variation and this variation forms the basis for the adaptation in populations to changing environmental conditions. Both the gene pool allele frequencies and genotype frequencies define the genetic structure of a population: allele frequencies reflect the amount of genetic variation, whereas the genotype frequencies reflect the distribution of the variation among the members of the population. Microevolution concerns the change of genetic structure of a population over time, which results from the interaction of several microevolutionary forces: mutation, gene flow or migration, genetic drift, nonrandom or assortative mating and natural selection.

The force of mutation is the ultimate source of new genetic variation within a population and although most mutations are selectively neutral or detrimental, some mutations enhance the fitness of the carrier and these variants may be selected for gradualistic adaptive evolution. Within finite populations the mutational variation is affected by genetic drift, assortative mating, gene flow and natural selection. Genetic drift comprises the random changes in allele frequencies due to the stochastic nature of sexual reproduction (Mendel's law of segregation and demographic stochasticity). It is a dispersive force that removes genetic variation from populations at a rate that is inversely proportional to the population size. Although genetic drift is a weak dispersive force

in most natural populations, its effects are important in case of genetic bottlenecks and founder events. Assortative mating occurs when mating between individuals having the same or different genotypes occurs more often than expected by chance alone and this nonrandom mating changes the genotype frequencies, whereas the allele frequencies remain unaffected. In case of inbreeding, mating between close relatives occurs more frequently than expected by chance and an increase of the homozygosity follows. The opposite case, outbreeding, results in a reduction of the homozygosity. Migration between divergent populations acts as a homogenizing agent that erodes population differentiation. Within populations, migration causes a change of allele and genotype frequencies and immigration may introduce new alleles. Natural selection, the differential reproductive succes of alleles, is the only mechanism of adaptive evolution and may change allele frequencies or conserve them over longer periods of time.

Inferences about the occurence of microevolution and the forces that drive this evolution can be made by comparing the observed genetic structure of populations with the structure that is expected under Hardy-Weinberg equilibrium. The Hardy-Weinberg equilibrium describes the gene pool of infinitely large populations that consist of randomly mating individuals and that is free of other microevolutionary forces (Hardy, 1908; Weinberg, 1908). If these assumptions are met, the allele and genotype frequencies remain constant over time (generations) and a simple relationship exists between both frequencies. Thus, in case the observed genetic structure of a population differs from the expected equilibrium structure, it can be concluded that evolution occurs and that one or more of the assumptions of the Hardy-Weinberg equilibrium are not valid in this population.

The genetic structure of a population can be described using genetic markers, measurable characters that allow the detection of variation in either protein or DNA sequences. Biochemical markers or protein markers (seed storage proteins, isozymes) were among the first group of markers employed for genetic diversity assessment. They are generated through electrophoresis, taking advantage of the migrational properties of proteins and enzymes and are visualized by histochemical stains specific to the enzymes being assayed. Although variation revealed by protein markers is due to differences in the amino acid sequence and hence may have a genetic basis, they are influenced by the environment and by changes in developmental stages, which limits their applicability. DNA markers detect polymorphisms in nuclear and organellar DNA and are considered to be objective measures of genetic variation. Many different types of DNA markers exist, all differing with respect to important features such as genomic abundance, level of polymorphism, locus specificity, codominance, reproducibility, technical requirements and financial investment. None of the available techniques is superior for a wide range of applications

and the choice of the most appropriate marker will depend on the type of application, the presumed level of polymorphism, the presence of sufficient technical facilities, time constraints and financial limitations.

In order to describe the genetic structure of *S. luteus* populations and to interpret the results in the context of heavy metal pollution, amplified fragment length polymorphisms (AFLPs) and microsatellites were the markers of choice in this study. The AFLP technique does not require prior DNA sequence information and generates dominant markers randomly distributed over the entire genome, with a high multiplex ratio. Microsatellites, on the other hand, have a low multiplex ratio, are locus specific and codominant. Because of their potentially high levels of polymorphism, AFLP and microsatellite markers can be applied for population genetic analysis and the combination of both marker types should allow an accurate description of the population genetic structure.

3.1.1 Amplified fragment length polymorphism (AFLP)

Amplified fragment length polymorphism or AFLP is a robust DNA fingerprinting technique that does not require prior DNA sequence information (Vos et al., 1995). The technique is based on the selective amplification of subsets of the restriction fragments from a total digest of genomic DNA and it involves the following five steps (see figure 3.1 for a graphical representation):

Restriction digest of genomic DNA. Restriction fragments for amplification are generated by two different restriction enzymes, one frequent and one rare cutting enzyme. The frequent cutter will generate small DNA fragments that will amplify well and that are in the optimal size range for separation on denaturing polyacrylamide gels, while the rare cutter will reduce the number of fragments to be amplified, thus limiting the number of selective nucleotides necessary in subsequent steps.

Ligation of oligonucleotide adaptors. Double-stranded adaptors are ligated to the restriction fragments. The sequence of the adaptors and the adjacent restriction sites serve as primer binding sites for subsequent amplification of the restriction fragments. The adaptors are constructed in such a way that after ligation, the original restriction site is altered. This allows for the combination of the restriction digest and the adaptor ligation in one single step.

Preselective amplification. A subset of the pool of restriction fragments is PCR amplified using oligonucleotide primers complementary to the adaptor and restriction site sequences, but extended with one or two selective nucleotides at their 3'-end. This step reduces the amount of mismatch amplification during the selective PCR, which enhances the

reproducibility of the AFLP technique, especially in the case of complex genomes. Although the vast majority of the fragments end with restriction sites for the frequent cutter at both sides, the amplification of these fragments is inefficient due to the presence of inverted repeats at the ends which form stem–loop structures during primer annealing.

Selective amplification with labelled primers. For the selective amplification of restriction fragments the oligonucleotide primer complementary to the adaptor ligated at the rare restriction site is either fluorescently or radioactively labelled. Both primers are now extended with up to four selective nucleotides. Typically, 50 – 200 different AFLP-fragments are amplified and different primer combinations allow for the amplification of different subsets of restriction fragments.

Gel-based analysis of the AFLP-fragments. Polyacrylamide gel electrophoresis under denaturing conditions is used to separate the amplified fragments and subsequently the fragments are visualized through detection of the fluorescent or radioactive signal. DNA polymorphisms are observed as presence or absence of specific markers and are either due to sequence changes in the restriction sites or to insertions/deletions within the restriction fragments.

The major advantages of the AFLP technique are the high level of polymorphism and the high multiplex ratio. The large number of revealed DNA polymorphisms allows for the differentiation of accessions across a wide range of scales, from individual/clone discrimination up to species level differences, as well as for the estimation of variation across the entire genome rather than from a limited number of segments. Moreover, the technique does not rely on prior DNA sequence information and since it is PCR based, it allows high throughput analysis.

An important problem of the AFLP technique is the presumed homology of comigrating AFLP fragments, which is not always the case (Vekemans et al., 2002). Comigrating bands may originate from different chromosomal regions and the absence of a band may be due to different mutation events. Although size homoplasy introduces bias into the data, the extent to which parameter estimation is influenced depends on the species under study. As was suggested for RAPD markers, homology of AFLP markers may be a function of the taxonomic distance: the more closely related the accessions are, the greater the probability of homology of shared comigrating fragments (Rieseberg, 1996). Also the mutation rate of the species under study is of importance, as high levels of variation due to high mutation rates may be accompanied by higher levels of size homoplasy. Thus, care should be taken when applying AFLP markers for the analysis of more distantly related accessions.

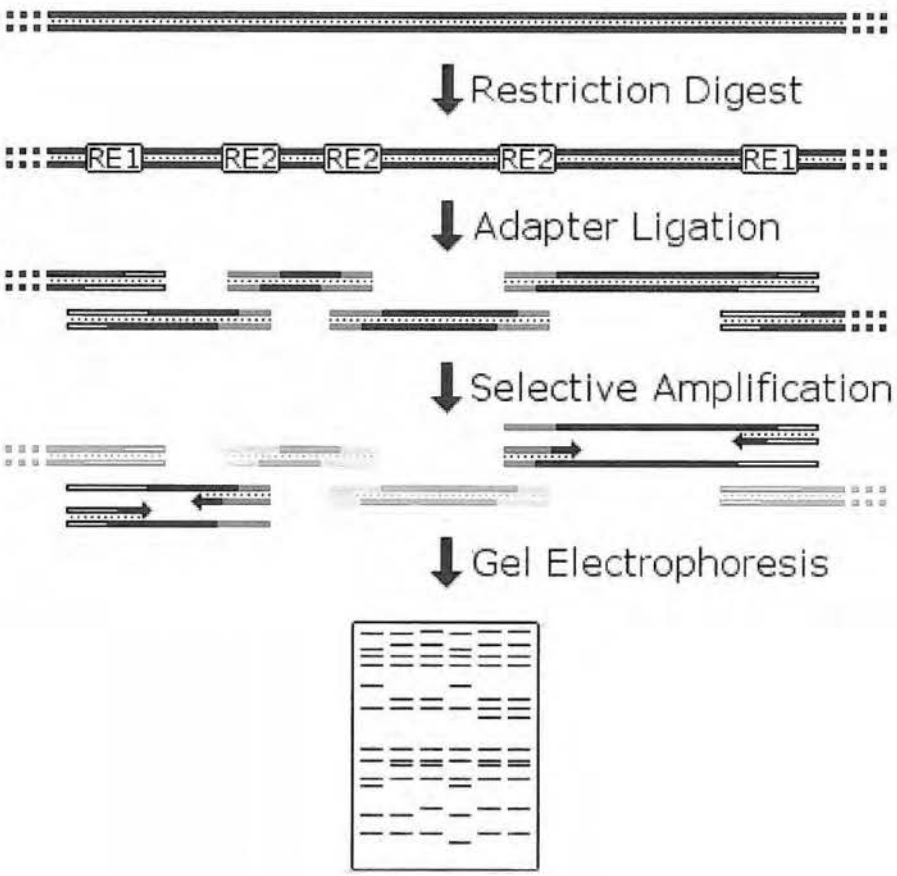

Figure 3.1: Schematic representation of the AFLP technique. The different steps include: restriction digest of genomic DNA using a rare cutter (RE1) and a frequent cutter (RE2), ligation of double-stranded adaptors to the sticky ends of the restriction fragments, selective amplification of the restriction fragments using oligonucleotide primers complementary to the adaptor sequences and extended with selective nucleotides, and separation of the AFLP-fragments by polyacrylamide gel electrophoresis.

Although mixed dominance/codominance may occur, AFLP markers are considered to be dominant markers in most applications: the heterozygote cannot be distinguished directly from the dominant homozygote phenotype, unless breeding or pedigree studies are carried out to determine the inheritance pattern of each band. Because of this lower information content, dominant markers are less efficient for population genetic studies when compared to codominant markers and require a larger number of individuals to be sampled (Lynch and Milligan, 1994).

Despite these drawbacks, the AFLP technique remains a valuable DNA fingerprinting method that is frequently applied in studies concerning genetic mapping, forensic genotyping, paternity analysis and population genetics.

3.1.2 Microsatellites or simple sequence repeats

Microsatellites or simple sequence repeats (SSRs) are a special class of tandemly repeated DNA in which a specific motif of $1-6$ bp is repeated up to ~ 100 times. SSRs have been detected in the genomes of all organisms analyzed so far and their genomic abundance, as well as their various functions and effects, are associated with their mutation rate, which is very high (typically $10^{-6}-10^{-2}$ mutation events per locus per generation) compared to the rate of point mutations at coding sequences (Li et al., 2002). Although the mutation process varies with species, repeat type, locus, allele, age and sex, the instability of microsatellite loci manifests itself mainly as changes in the number of repeat motifs (Schlötterer, 2000).

In order to explain the high mutation rates of microsatellite loci, two mutational mechanisms are being considered: replication slippage and recombination. Many changes of repeat numbers at SSR loci are caused by slipped-strand mispairing errors during DNA replication. In its simplest form, slipped-strand mispairing involves local denaturation and displacement of the strands of a DNA duplex, followed by mispairing of complementary bases at the site of an existing SSR (see figure 3.2). The consequences of this mispairing, when followed by replication or repair, can lead to insertions or deletions of one or more short repeat units. Some of these errors are corrected by exonucleolytic proofreading and mismatch DNA repair, but many escape repair and become mutations (Eisen, 1999; Levinson and Gutman, 1987). Recombination can potentially change the SSR length by unequal crossover between sister chromatids or by gene conversion (nonreciprocal recombination). Although gene conversion and crossover are often associated, it is mostly gene conversion without crossover that destabilizes microsatellites (Richard and Pâques, 2000). The SDSA model (synthesis-dependent strand annealing) accounts for non-crossover gene conversions and is associated with double-stranded break repair. In this model, the two newly synthesized strands unwind from their

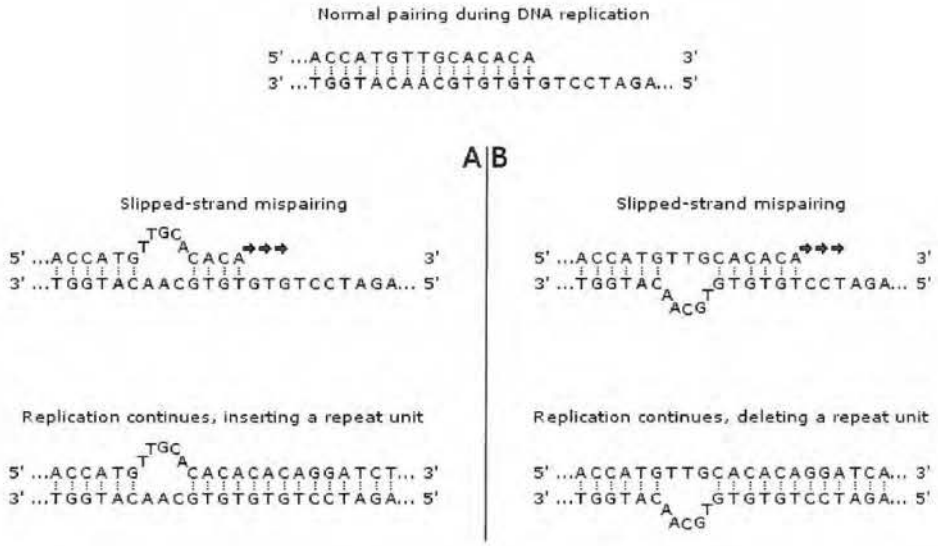

Figure 3.2: Slipped-strand mispairing as a cause of SSR instability. Two-base slippage in a CA-repeat during replication of a DNA duplex, followed by continued chain elongation; slippage in the $3' \rightarrow 5'$ direction (A) results in insertion of one CA unit, whereas slippage in the other direction (B) results in deletion of one CA unit; small arrows indicate the direction and starting point of DNA synthesis.

template and anneal with each other. Out-of-frame reannealing would then lead to tandem repeat rearrangements (see figure 3.3).

Although microsatellites are often considered to be selectively neutral sequences that are randomly distributed over the euchromatic genome, several studies showed structural patterns with allele size constraints and functional significance, as well as a nonrandom distribution over the genome (Li et al., 2003; Kashi et al., 1997; Schlötterer and Wiehe, 1999). Except for triplet SSRs, differences in the distribution pattern of tandem repeats in coding and noncoding regions appear to be common. In many species, the majority of SSRs ($48 - 67\%$) are dinucleotide repeats (Schug et al., 1998) and Dokholyan et al. (2000) reported a significant difference in the distribution of dimeric tandem repeats in coding and noncoding DNA: for exons, the length distributions of dimeric tandem repeats were found to be exponential, whereas for introns and intergenic regions, the length distributions deviated strongly from an exponential function and long chains of up to 100 repeats were found. The absence of lengthy nontriplet repeats in coding regions is thought to be due to negative selection against frameshift mutations resulting from length changes in the nontriplet repeats and leading to a loss of protein functionality (as e.g. in Huntington's disease (Wells, 1996)). Nevertheless, 14% of all proteins contain repeated sequences, with a three times higher abundance of repeats in eukaryotes as compared to prokaryotes (Marcotte et al., 1999).

The functional significance of several microsatellites has been proven for various biological phenomena such as chromatin organization, regulation of DNA metabolic processes and regulation of gene activity (Li et al., 2002). More specifically, some of the putative functions of microsatellites are:

DNA structure. Microsatellite sequences are capable of forming a wide variety of unusual DNA structures with simple and complex folding patterns. These structures may have important regulatory effects on gene expression, provide specific protein recognition motifs and may be involved in the unwinding of DNA during transcription (Catasti et al., 1999; Fabregat et al., 2001).

Centromere structure. In many species, the centromeric region of chromosomes is composed of numerous tandem repeats, which affect the centromeric organization and which may be involved in sister chromatid cohesion and kinetochore formation or function (Murphy and Karpen, 1995).

Recombination. Microsatellites have shown to be hotspots for recombination events and this is thought to be due to their high affinity for recombination enzymes or to their influence on DNA structure (Biet et al., 1999).

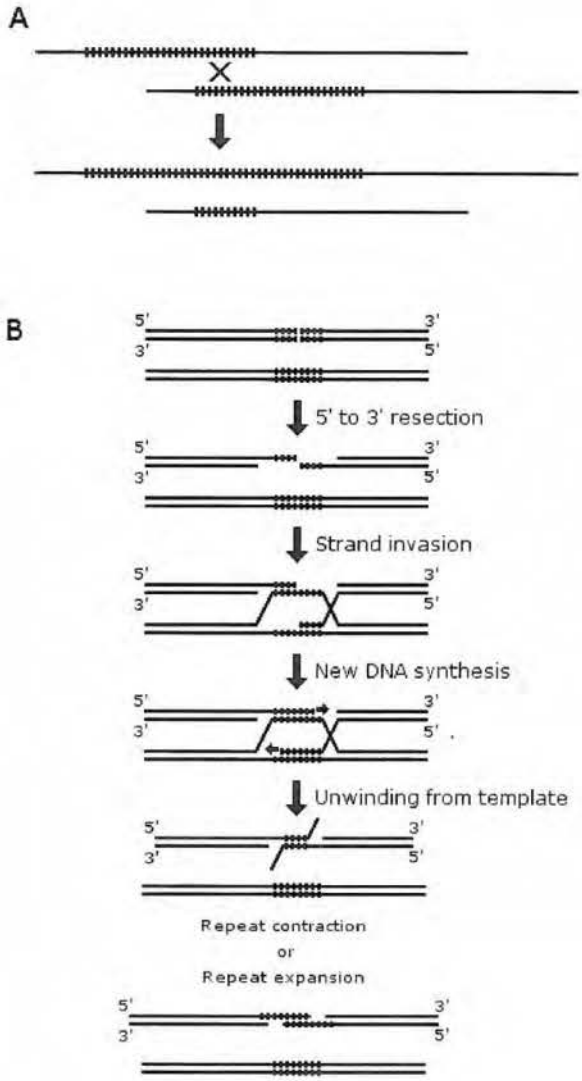

Figure 3.3: Tandem repeat rearrangements by unequal crossover and gene conversion. (A), tandem repeat rearrangement by unequal crossover between sister chromatids; reciprocal exchange leads to simultaneous contraction of one of the repeats and expansion of the other. (B), SDSA model of tandem repeat rearrangements; the newly synthesized strands both contain repeats and their out-of-frame reannealing can lead to contraction or expansion of the repeat. Small arrows indicate the direction and starting point of DNA synthesis.

Transcription. SSRs located in promotor and intronic regions can affect transcriptional activity, as shown by the marked changes in gene transcription following the deletion of various di-, tri- and tetra-SSR tracts (Li et al., 2002).

Although the mechanism of microsatellite evolution is still not fully understood and their selective neutrality as well as their randomness throughout the genome is debatable, SSRs are being widely employed as genetic markers, for which their high variability renders them powerful. Microsatellites are being used extensively in genome mapping, but their applications span over different areas, such as forensic DNA studies, population genetics and conservation/management of biological resources (Schuler et al., 1996; Jarne and Lagoda, 1996). As with AFLP markers, the great advantage of microsatellites is the large number of polymorphisms that are revealed at a wide range of scales. However, microsatellites are codominant markers and thus more powerful for estimating population genetic parameters when compared to AFLP markers.

One of the problems when using microsatellites is their need to be isolated *de novo* from most species being studied for the first time. This is due to their preference for noncoding regions where the nucleotide substitution rate is higher than in coding regions. Consequently, the design of "universal" primers matching the conserved sequences surrounding the repeats becomes problematic for most microsatellite loci.

The isolation of microsatellites can be a very involving task in terms of effort and time because it traditionally involves screening genomic libraries with appropriate oligonucleotide probes (Rassmann et al., 1991). The proportion of positive clones, which contain microsatellites, that can be obtained by means of this method usually ranges from 12% to less than 0.04%. In order to increase the efficiency of microsatellite isolation, several alternative strategies have been developed. The construction of genomic libraries enriched for microsatellites allows for higher frequencies of positive clones and in order to avoid library construction and screening altogether, PCR-based methods have been introduced. Two different methods that were used to isolate microsatellites from *S. luteus*, one that involves library screening (Glenn and Schable, 2005) and one PCR-based method (Lian and Hogetsu, 2002), are respectively detailed in sections 3.2.1 and 3.2.2.

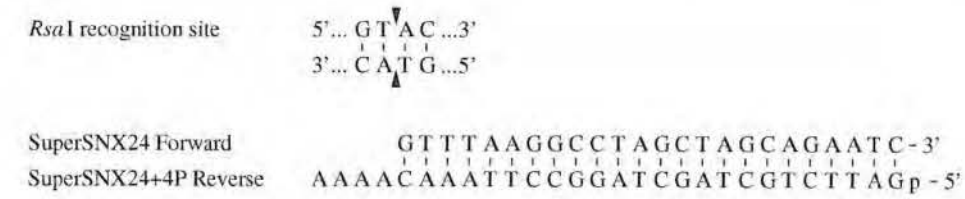

Figure 3.4: *Rsa*I restriction site and structure of the double-stranded SuperSNX adaptor used for isolation of microsatellite loci from *Suillus luteus*. Note the phosphorylation of the 5'-end of the reverse primer, which allows ligation; arrows indicate cleavage site.

3.2 Microsatellite isolation in *Suillus luteus*

3.2.1 Microsatellite isolation by enriched library screening

Material and methods

DNA extraction, restriction digest and adaptor ligation. Approximately 5 μg of genomic DNA, extracted from a single sporocarp with the DNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France), was incubated for 1 h at 37°C with 10 U *Rsa*I in 40 μl 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 10 mM dithiothreitol and 1 μg BSA. Subsequently, double-stranded SuperSNX adaptors were ligated to the restriction fragments (see figure 3.4 for adaptor structure; Schable et al. (2002)). Double-stranded adaptors were prepared by incubation of equimolar amounts of SuperSNX24 and SuperSNX24+4p primers in 100 mM NaCl at 95°C for 5 min after which the mixture was allowed to slowly cool to room temperature. In order to ligate the adaptors to the restriction fragments, 10 μl of a solution containing 800 U T4 DNA ligase and 20 U *Xmn*I in 5 μM SuperSNX adaptor, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 10 mM dithiothreitol and 1 μg BSA was added to the restricted DNA and this mixture was incubated overnight at 16°C.

Enrichment for microsatellite containing DNA fragments. The adapted DNA fragments were PCR amplified in a 50 μl reaction containing 2 μl of adapted restriction fragments, 0.5 μM SuperSNX24, 1.5 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 2 mM of all four dNTPs. After an initial denaturation step at 95°C for 2 min, PCR was performed for 25 cycles with the following temperature profile: 95°C for 20 s, 60°C for 20 s and 72°C for 1.5 min. The PCR product was diluted twofold prior to the enrichment. Enrichment for microsatellite containing DNA fragments was performed in four separate reactions, each with a different mixture

of biotinylated oligonucleotide probes (see table 3.1). Seventy μl of 6× SSC (0.9 M NaCl, 0.1 M NaCitrate), 0.1% SDS and 5 μl of a biotinylated oligonucleotide mixture (3 or 4 oligonucleotides, 16.7 μM each) were added to 25 μl of adapted restriction fragments and this mixture was incubated at 95°C for 5 min, after which it was allowed to slowly cool to room temperature. Twenty-five μl Vectrex®Avidin D beads (Vector Laboratories, Burlingame, CA, USA) were washed three times with 400 μl TBST (0.1 M Tris pH 7.5, 150 mM NaCl, 0.1% Tween 20), resuspended in 300 μl TBST and added to the mixture of adapted restriction fragments and biotinylated oligonucleotide probes. Avidin-biotin binding was allowed by incubating the mixture at 37°C for 1 h and the avidin beads were then subsequently washed three times with 400 μl TBST and three times with 400 μl 0.2× SSC, 0.1% SDS. The avidin beads were then resuspended in 100 μl TLE, incubated at 95°C for 5 min and pelleted by centrifugation at 12000×g for 5 min. In order to increase the amount of product and to recover double-stranded DNA, a PCR was performed with 2 μl of eluted DNA as template in a 25 μl reaction containing 0.5 μM SuperSNX24, 1.5 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 2 mM of all four dNTPs. The temperature profile was: 95°C for 2 min; 25 cycles of 95°C for 20 s, 60°C for 20 s and 72°C for 1.5 min; 72°C for 10 min. The enrichment procedure was repeated with a two-fold dilution of the recovered double-stranded DNA in order to further increase the proportion of repeat sequences in the enriched library.

Identification of microsatellite containing DNA fragments. The enriched and recovered DNA fragments were subcloned using the TOPO TA Cloning®Kit with pCR®2.1-TOPO®vector and TOP 10 cells according to the manufacturer's instructions (Invitrogen Corporation, Carlsbad, CA, USA). Succesfully transformed bacteria were selected and plated out in raster form (11 × 11 colonies per plate), incubated overnight at 37°C and blotted onto Hybond^{TM} N+ nylon membranes (Amersham Biosciences, Buckinghamshire, UK). The cells were subsequently lysed using 10% SDS and the DNA denatured with 0.5 M NaOH, 1.5 M NaCl. The pH was neutralised with 0.5 M Tris-HCl pH 6.8, 1.5 M NaCl and the membranes were washed in 400 ml 2× SSC. After drying the membranes, the DNA was bound by baking at 80°C for 2 h. Oligonucleotide colony hybridisation was performed in order to identify DNA fragments containing tandem repeat sequences using the ECL Direct^{TM} Nucleic Acid Labelling and Detection System (Amersham Biosciences, Buckinghamshire, UK) according to the manufacturer's instructions with the same mixtures of oligonucleotide probes as used for the enrichment procedure (see table 3.1).

Mix	Oligonucleotide primers
1	5'-AACAACAACAACAACAACAACAACAACAAC-3'
	5'-ACACACACACACACACACACACACACACAC-3'
	5'-AGAGAGAGAGAGAGAGAGAGAGAGAGAG-3'
	5'-CAGCAGCAGCAGCAGCAGCAGCAGCAG-3'
2	5'-AATAATAATAATAATAATAATAATAAT-3'
	5'-ACCACCACCACCACCACCACCACCACC-3'
	5'-ACGACGACGACGACGACGACGACGACG-3'
	5'-ATCATCATCATCATCATCATCATCATC-3'
3	5'-AAAGAAAGAAAGAAAGAAAGAAAGAAAG-3'
	5'-ACAGACAGACAGACAGACAGACAGACAG-3'
	5'-ACTCACTCACTCACTCACTCACTCACTC-3'
4	5'-ACATACATACATACATACATACATACAT-3'
	5'-AGGCAGGCAGGCAGGCAGGCAGGCAGGC-3'
	5'-ATGCATGCATGCATGCATGCATGCATGC-3'

Table 3.1: Overview of the different mixtures of biotinylated oligonucleotide probes used for the enrichment of microsatellite containing DNA fragments in *Suillus luteus*.

Sequencing of microsatellite containing DNA fragments. Positive colonies were manually picked and transferred to 96-wells PCR plates containing 20 μl 0.5 μM M13f-primer, 0.5 μM M13r-primer, 1.5 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 2 mM of all four dNTPs. After an initial cell lysis at 95°C for 5 min, PCR was performed with the following temperature profile: 35 cycles of 95°C for 20 s, 50°C for 20 s and 72°C for 1.5 min. The PCR products were diluted 20 times and 1 μl was used in a 20 μl sequencing reaction with the M13f-primer using ABI BigDye®Terminator reaction mix (Applied Biosystems, Foster City, CA, USA), following the manufacturer's instructions. The sequences were run on an ABI PRISM®3100-Avant Genetic Analyzer.

Primer design and testing. Primer pairs were designed to sequences flanking the microsatellite region using PRIMER3 software (Rozen and Skaletsky, 1996). Microsatellite analysis PCR reactions were set up in total volumes of 20 μl containing 0.5 μM of each primer, 1 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl, 200 μM of all four dNTPs and approximately 2 ng of template DNA. The following PCR temperature profile was used: initial denaturation at 95°C for 2 min followed by 30 cycles of 95°C for 20 s, locus specific annealing temperature (see table 3.2) for 20 s and 72°C

for 1 min. The primer pairs were tested and the PCR conditions optimized using genomic DNA from ten randomly chosen individuals in the sample used for this study (see section 3.3.1 on page 62). The primers that allowed amplification in these individuals were then used for a screening in a larger subset of the total sample (40 individuals) in orther to be characterised. For each locus, the observed number of alleles, the observed heterozygosity (H_O) and the expected heterozygosity (H_E) under the hypothesis of Hardy-Weinberg equilibrium (Nei, 1978) were calculated using GENETIX (Version 4.04; Belkhir et al. (2002)). Linkage disequilibrium between the microsatellite loci was tested with the web-based version of GENEPOP (Version 3.4: dememorization 1000, batches 100, iterations 1000; Raymond and Rousset (1995)).

Results and discussion

In total, 360 bacterial colonies were screened by oligonucleotide colony hybridisation for tandem repeat sequences and 24 nonredundant microsatellite loci were identified, a proportion that reflects the scarceness of microsatellites in ectomycorrhizal genomes (Lian et al., 2003). The most abundant microsatellite sequences were those with (AC) as repeat motif (7 out of 24), followed by the sequences that consist of (ACC)-repeats (5 out of 24). However, only five of these loci allowed proper amplification in the test sample of ten randomly selected individuals and three of these allowed amplification of one or two bands, thus being single-locused and codominant (see table 3.2). PCR amplification with primers specific for the microsatellite loci Sulu01 and Sulu02 resulted in patterns with up to six bands, indicating genomic duplications of the corresponding DNA segments. The microsatellite loci Sulu03, Sulu04 and Sulu05 revealed relatively high allelic variation, having a total number of alleles per locus ranging from 8 to 22. Expected heterozygosities varied between 0.786 and 0.889, and a heterozygote deficit was apparent for locus Sulu03. No linkage disequilibrium was found and high allele frequencies were not biased to a single allele at these loci, suggesting that they are appropriate for population genetic analysis.

Most of the problems were encountered during the design of appropriate primers and in setting up proper PCR conditions. More specifically, PCR amplification of the majority of the isolated loci resulted in high numbers of null alleles or amplification of multiple loci, either aspecific loci or duplicated microsatellite loci. An interesting observation is the frequent asymmetry of flanking regions between different isolated microsatellite loci: high similarities (more than 95%) between long stretches of the flanking sequences (up to several hundreds of basepairs) on only one side of different microsatellites were observed for several loci (see figure 3.5). These chimeric sequences are not uncommon in microsatellite isolation procedures that are based on enriched

Locus	Repeat	Primer sequence (5' → 3')	T_a (°C)	Size range (bp)	No. of alleles	H_O	H_E
Sulu01*	$(ATCC)_9$	F: CCCTTTAATCGTCCCAATCC R: ACATCTTGTCGCGTCTTGTG	55	172 – 264	18	–	–
Sulu02*	$(CCA)_5TCA(CCA)_5$	F: TGACGGTTCACAGATGAAGG R: AGCAACAGCAGGAGTGGTG	55	212 – 311	21	–	–
Sulu03	$(GAT)_3GAC(GAT)_{13}$	F: CCCAACACCGTGAATTGAG R: TTCCATGAATTCGGCTTCTC	53	209 – 296	22	0.321	0.889
Sulu04	$(CTT)_8(CCT)_2$	F: GCATCATGATATAGTCGGGTCAG R: AGATTTAGACGGACAGGAATGC	55	181 – 263	15	0.792	0.853
Sulu05	$(AGCAGG)_6$	F: TAGCGCTATATCAAGCCTGACG R: CTATGTTTGCTCTTCGCTCTTCC	55	193 – 247	8	0.729	0.786

Table 3.2: Overview of the characteristics of the five microsatellite loci, isolated according to Glenn and Schable (2005), that allowed proper PCR amplification and appeared to be polymorphic in *Suillus luteus*. T_a, annealing temperature used for PCR amplification; H_O, observed heterozygosity; H_E, expected heterozygosity under Hardy-Weinberg equilibrium (Nei, 1978); *, indicates amplification of multiple loci.

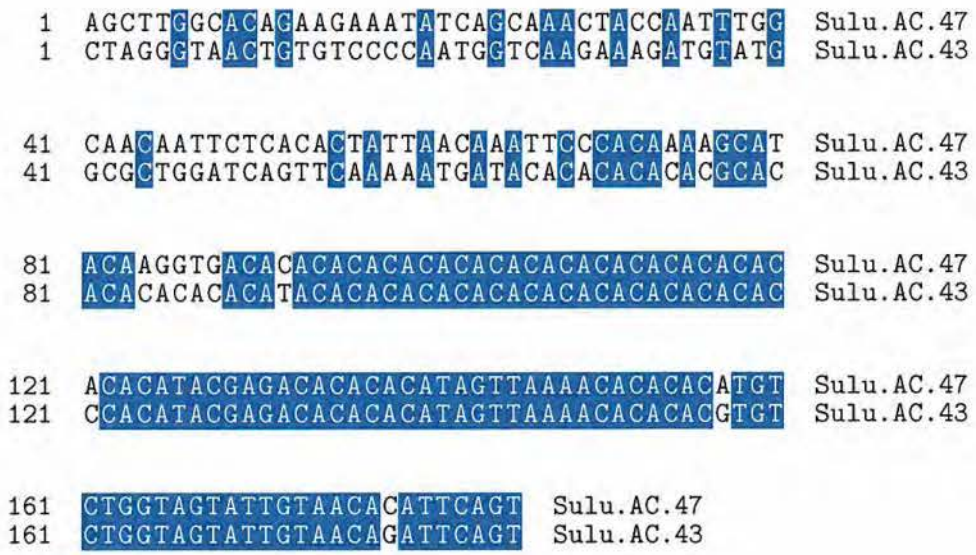

Figure 3.5: Example of asymmetry found in the flanking sequences of different microsatellite loci isolated from *Suillus luteus*.

libraries and most often, they are not suited for the development of primer pairs (Squirrell et al., 2003). It has been suggested that these sequences are the consequence of crossing over between nonhomologous microsatellite loci, leading to the exchange between the flanking regions of different microsatellites (Meglecz et al., 2004). Alternatively, PCR-mediated recombination during the enrichment procedure can also generate asymmetry of the flanking regions, but therefore PCR elongation has to be interrupted at the microsatellite motif which is believed to be less probable. In this case, the microsatellite that is being elongated can reanneal with a nonrelated microsatellite, leading to the formation of a chimeric fragment with two unrelated flanking sequences. A nearly complete identity, diminished only by the errors incorporated by the DNA polymerase employed, would then be expected between the flanking sequences at one side of the microsatellite. However, because no chimeric fragments were found using the PCR-based isolation method developed by Lian and Hogetsu (2002), they are believed to be artifacts of the enrichment procedure rather than the result of nonhomologous crossover events.

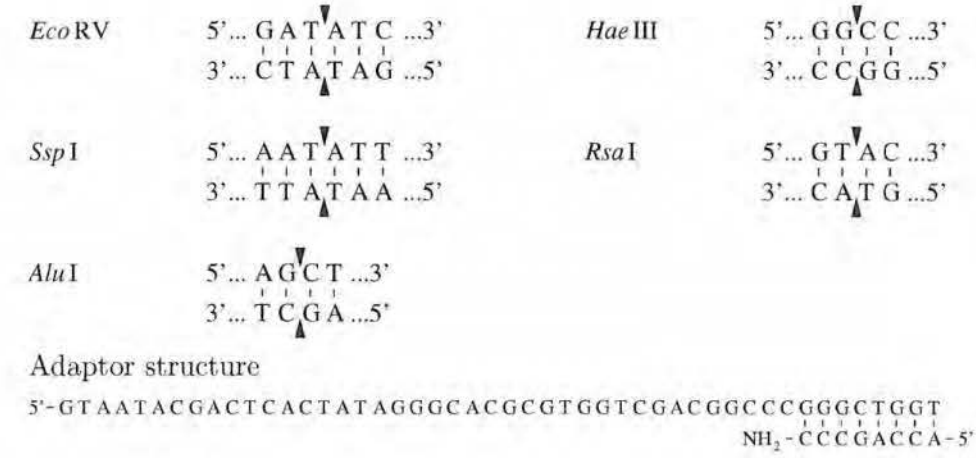

Figure 3.6: Restriction enzyme recognition sites and structure of the double-stranded adaptor used for the isolation of microsatellite loci from *Suillus luteus* following Lian and Hogetsu (2002). Note the amine residue at the 3'-end of the reverse strand, which prevents PCR-elongation; arrows indicate cleavage site.

3.2.2 Microsatellite isolation by a dual-suppression-PCR technique

Material and methods

DNA extraction, restriction digest and adaptor ligation. Adaptor-ligated, restricted DNA libraries were constructed according to Siebert et al. (1995). Four μg of genomic DNA extracted from a single *S. luteus* sporocarp with the DNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France), was separately incubated for 4 h at 37°C with 40 U of *Eco*RV, *Ssp*I, *Alu*I, *Rsa*I and *Hae*III restriction enzymes in 100 μl 50 mM NaCl, 10 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$ and 1 mM dithiothreitol. The reaction was terminated by purifying the DNA using the MinElute Reaction Cleanup Kit (Qiagen, Courtaboeuf, France). Double-stranded adaptors (see figure 3.6) were prepared by incubation of equimolar amounts of both forward and reverse strands in 100 mM NaCl at 95°C for 5 min after which the mixture was allowed to slowly cool to room temperature. Five μl of restricted DNA was ligated to an excess of adaptor overnight at 16°C under the following conditions: 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP, 10 mM dithiothreitol, 5 μM double-stranded adaptor and 100 U T4 DNA ligase in a total volume of 20 μl. Finally, the adapted and restricted DNA was diluted five-fold with TE buffer and stored at -20°C until further use.

PCR amplification of potential microsatellite loci. As a first step for isolating microsatellite loci, fragments flanked by $(AC)_n$, $(AG)_n$, $(ACC)_n$ or $(ACG)_n$ sequences at one end were amplified from the constructed *Eco*RV DNA library using respectively $(AC)_{10}$, $(AG)_{10}$, $(ACC)_8$ and $(ACG)_8$ primers together with the adaptor primer AP2 (5'-CTATAGGGCACGCGTGGT-3') designed from the longer strand of the adaptor. The presence of an amine residue at the 3'-end of the short adaptor strand prevents DNA polymerase-catalyzed extension of this strand. There is no sequence complimentary to the adaptor primers AP1 and AP2, unless one of the microsatellite primers is extended beyond the longer adaptor strand. Therefore, these primers facilitate only the amplification of the fragment containing the specific microsatellite primer. PCR amplification was performed separately for each microsatellite primer in 20 μl reactions containing 0.5 μl of the *Eco*RV DNA library, 0.5 μM of each primer, 1 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 200 μM of all four dNTPs. The temperature profile of the PCR was the following: initial denaturation at 95°C for 2 min; 35 cycles of 95°C for 20 s, primer specific annealing temperature ($(AC)_{10}$ and $(AG)_{10}$: 60°C, $(ACC)_8$ and $(ACG)_8$: 62°C) for 20 s, 72°C for 1.5 min, followed by a final extension at 72°C for 10 min. The amplified fragments were ligated into pCR®2.1-TOPO®vectors (Invitrogen Corporation, Carlsbad, CA, USA) according to the manufacturer's instructions and the plasmids were transformed into *Escherichia coli*, TOP10 strain (Invitrogen Corporation, Carlsbad, CA, USA). Plasmid inserts from positive clones were PCR amplified in 20 μl 0.5 μM M13f-primer, 0.5 μM M13r-primer, 1.5 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 2 mM of all four dNTPs. After an initial cell lysis at 95°C for 5 min, PCR was performed with the following temperature profile: 35 cycles of 95°C for 20 s, 50°C for 20 s and 72°C for 1.5 min. The PCR products were diluted 20 times and 1 μl was used in a 20 μl sequencing reaction with the M13f-primer using ABI BigDye®Terminator reaction mix (Applied Biosystems, Foster City, CA, USA), following the manufacturer's instructions. The sequences were run on an ABI PRISM®3100-Avant Genetic Analyzer.

The second step of the isolation procedure was to determine the sequence of the other flanking region of each potential microsatellite. Two primers, IP1 and IP2, were designed using PRIMER3 software (Rozen and Skaletsky, 1996) from the newly sequenced region flanking the repeat sequence, IP2 being designed from the region between IP1 and the repeat sequence. Nested PCR was performed using the primers IP1 and IP2 together with the adaptor primers AP1 (5'-CCATCGTAATACGACTCACTATAGGGC-3') and AP2. The primary PCR was conducted with each of the constructed libraries using primers IP1 and AP1. The secondary PCR was conducted with a 100-fold dilution of the primary PCR products using primers IP2 and AP2. Both PCR amplifica-

tions were performed in 20 μl volumes containing 0.5 μl of the DNA library, 0.5 μM of each primer, 1 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl and 200 μM of all four dNTPs. The temperature profile of the PCR was as follows: initial denaturation at 95°C for 2 min; 30 cycles of 95°C for 20 s, 60°C for 20 s and 72°C for 1.5 min, followed by a final extension at 72°C for 10 min. Usually, the reaction product for at least one of the DNA libraries consisted of a single fragment and these fragments were cloned and sequenced as described above. A second specific primer was then designed from the newly defined flanking region for amplification of the entire region containing the microsatellite.

Characterisation of the microsatellite loci. In order to verify the repeat sequence, PCR amplification products obtained with the microsatellite specific primers were subsequently cloned and sequenced. PCR reactions were set up in total volumes of 20 μl containing 0.5 μM of each primer, 1 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl, 200 μM of all four dNTPs and approximately 2 ng of template DNA. The following PCR temperature profile was used: initial denaturation at 95°C for 2 min, 30 cycles of 95°C for 20 s, locus specific annealing temperature (see table 3.3) for 20 s, 72°C for 1 min, followed by a final extension at 72°C for 10 min. The primer pairs were tested and the PCR conditions optimized using genomic DNA from ten randomly chosen individuals in the sample used for this study (see section 3.3.1 on page 62). Cloning and sequencing were performed as described above. The primers that allowed amplification of true microsatellite loci in these individuals were then used for a screening in a larger subset of the total sample (40 individuals) in order to be characterised. For each locus, the observed number of alleles, the observed heterozygosity (H_O) and the expected heterozygosity (H_E) under the hypothesis of Hardy-Weinberg equilibrium (Nei, 1978) were calculated using GENETIX (Version 4.04; Belkhir et al. (2002)). Linkage disequilibrium between the microsatellite loci was tested with the web-based version of GENEPOP (Version 3.4; dememorization 1000, batches 100, iterations 1000; Raymond and Rousset (1995)).

Results and discussion

In the first step of the isolation of microsatellite loci according to Lian and Hogetsu (2002), 24 different plasmid inserts from each of the four PCR amplifications were selected to be sequenced. Most of these loci allowed the design of two different primers for nested PCR amplification in the second step of the isolation procedure and in most cases, this resulted in single amplification products with at least one of the constructed DNA libraries. Based on the complete sequences of these loci, 15 were found to be microsatellite loci. In

total, 10 microsatellite loci allowed proper PCR amplification of a single locus in the test sample and 8 appeared to be polymorphic and codominant (see table 3.3). The number of alleles detected at these loci ranged from 5 to 16 per locus, with an average of 8.13. The expected heterozygosity under the hypothesis of Hardy-Weinberg genotype frequencies (Nei, 1978) varied between 0.532 and 0.811, and no linkage disequilibrium was observed. Except for loci Sulu06 and Sulu08, high allele frequencies were not biased to a single allele, suggesting that these microsatellite loci are appropriate for the detection of genotype polymorphism and population genetic analysis.

3.3 Genetic structure of *Suillus luteus* populations in heavy metal polluted and nonpolluted areas, revealed by AFLP and microsatellite analysis

3.3.1 Material and methods

Fungal material and DNA extraction

In order to sample populations in heavy metal contaminated and noncontaminated areas, fruit bodies of *S. luteus* were collected in 2002 at the nine different sampling sites introduced in section 2.1.2 on page 14. At each site, *S. luteus* sporocarps approximately 10 m apart were collected, frozen in liquid nitrogen, and stored at −80°C until later use. The sampling design was organized to specifically address small scale population genetic patterns and therefore the number of sampled individuals per site, rather than the number of sampling sites, was maximized. In total 164 samples were collected and although sample sizes of approximately 30 individuals per site were aimed at, the sample sizes for several locations were reduced because the number of available sporocarps was insufficient (see table 3.4 for the number of samples per geographic subpopulation). Total DNA was extracted from fungal tissue taken sterile from the centre of the collected sporocarps using the DNeasy Plant Mini Kit (Qiagen, Courtaboeuf, France) and stored at −20°C.

AFLP analysis

Restriction of genomic DNA and adaptor ligation. Approximately 0.5 μg genomic DNA was incubated for 1 h at 37°C with 5 U *Eco*RI and 5 U *Mse*I in 40 μl 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/μl BSA. Next, 10 μl of a solution containing 5 pMol *Eco*RI-adaptors, 50 pMol *Mse*I-adaptors, 1 U T4 DNA-ligase, 1 mM ATP in 10 mM Tris-HAc pH 7.5, 10 mM MgAc, 50 mM KAc, 5 mM DTT, 50 ng/μl BSA was added, and the incubation was continued for 3 h at 37°C. Adaptors were prepared

Locus	Repeat	Primer sequence (5' → 3')	T_a (°C)	Size range (bp)	No. of alleles	H_O	H_E
Sulu06	$(GA)_8$	F: CAATGTCTGAACTAGCACAAGG R: GTCACCACCCTCACACAGG	55	210 – 244	5	0.453	0.532
Sulu07	$(GCA)_7$	F: GGAGTGCGAAAAACATTTGC R: TCTTGATTCCAGCAATGAGC	55	204 – 252	7	0.583	0.693
Sulu08	$(GAC)_8$	F: GATAGCTTTCATGCCAATCG R: GACTAGGCGTGTTGGAGACG	55	196 – 238	5	0.512	0.587
Sulu09	$(AGC)_9$	F: GAGTCCCTTTGCAAATCTCG R: GGTGCAAAGAATGAACAACG	55	220 – 262	6	0.357	0.628
Sulu10	$(GAT)_{23}$	F: CGAGCTCCAGCAGTTACACG R: AAAACGCTTCTTCTGGTTGG	55	189 – 264	16	0.852	0.811
Sulu11	$(GAC)_8$	F: AACACTGAACGGATTGCTAGG R: GTGGAACAATGAGGGTTTGC	55	204 – 219	6	0.536	0.666
Sulu12	$(GAC)_9$	F: TGATTGTGAAAGCCCAGTAGC R: GATGAGTCGCAAGCAGACC	55	210 – 279	8	0.524	0.754
Sulu13	$(GAA)_7(GAT)_{12}\ldots(GAC)_7$	F: GGCAATGCCAATAACAACG R: CCTCCCCTACCACCTCTCC	55	330 – 411	12	0.837	0.793

Table 3.3: Overview of the characteristics of the eight microsatellite loci, isolated according to Lian and Hogetsu (2002), that allowed proper PCR amplification and appeared to be polymorphic in *Suillus luteus*. T_a, annealing temperature used for PCR amplification; H_O, observed heterozygosity; H_E, expected heterozygosity under Hardy-Weinberg equilibrium (Nei, 1978).

Contaminated		Noncontaminated	
Subpopulation	# samples	Subpopulation	# samples
Lm	10	Hh	9
Ls	36	E	12
Lc	11	P	33
Of	6		
N	27		
Lk	20		
Total	110	Total	54

Table 3.4: Overview of the number of *Suillus luteus* samples collected at nine different sampling sites in the province of Limburg (Belgium).

by allowing equimolar amounts of both strands to slowly cool to room temperature after incubation at 95°C for 5 min; adaptors were not phosphorylated (see figure 3.7 for adaptor structure). After ligation, the reaction mixture was diluted to 500 μl with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0 and stored at −20°C.

AFLP reactions. Preamplification of the adapted restriction fragments was performed using *Eco*RI and *Mse*I primers, respectively extended with the selective nucleotides A and C. PCRs were performed in 20 μl reactions containing 30 ng *Eco*RI-primer, 30 ng *Mse*I-primer, 5 μl template-DNA, 0.4 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl and 0.2 mM of all four dNTPs. Preamplification reactions were performed for 36 cycles with the following touch-down cycle profile: a 30 s DNA denaturation step at 94°C, a 30 s annealing step and a 1 min extension step at 72°C. The annealing temperature in the first cycle was 65°C and this temperature was subsequently reduced each cycle by 0.7°C for the next 12 cycles down to 56°C, which was used as annealing temperature for the remaining 23 cycles. The preamplication product was diluted 10-fold with 10 mM Tris-HCl, 0.1 mM EDTA pH 8.0 for subsequent use. Selective amplification was performed in 20 μl reactions containing 5 ng labeled *Eco*RI-primer, 30 ng *Mse*I-primer, 5 μl of the diluted preamplification product, 0.4 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl and 0.2 mM of all four dNTPs. The labelling reactions were performed in 50 μl 25 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 5 mM DTT, 0.5 mM spermidine-3HCl using 500 ng oligonucleotide primer, 100 μCi [γ-^{33}P]ATP and 10 U T4 polynucleotide kinase. The *Eco*RI- and *Mse*I-primers used for selective amplification each were extended

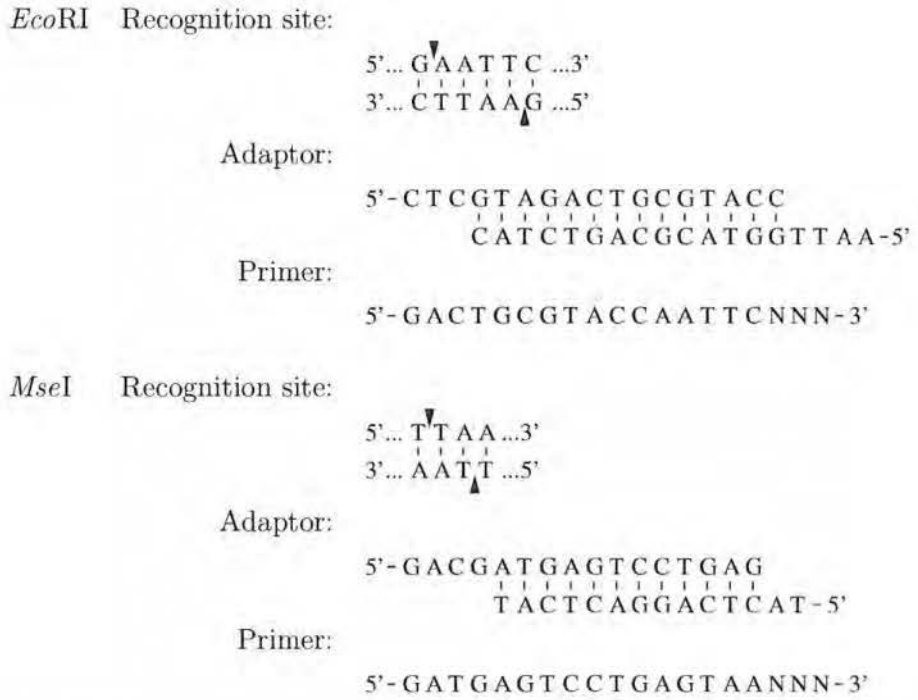

Figure 3.7: Recognition sites, adaptor structures and primer sequences for the restriction enzymes *Eco*RI and *Mse*I used in AFLP analysis of *Suillus luteus* populations; arrows indicate cleavage site.

with three selective nucleotides (see table 3.5 on page 71) and the same PCR temperature profile was used as described above for the preamplification step.

Gel analysis. Selective amplification products were mixed with an equal volume of formamide loading dye (98% formamide, 10 mM EDTA pH 8.0 and bromophenol blue and xylene cyanol as tracking dyes). The resulting mixtures were kept at −20°C overnight, heated for 5 min at 95°C and quickly cooled on ice before loading. Each sample was loaded on a 5% denaturing polyacrylamide gel, prepared using 5% acrylamide, 0.25% methylene bisacryl, 7.5 M urea in 50 mM Tris/50 mM Boric acid/1 mM EDTA. To 100 ml of gel solution, 500 μl of 10% APS and 100 μl TEMED was added and gels were cast using a SequiGen 38 × 50 cm gel apparatus (BioRad Laboratories Inc., Hercules, CA, USA). 100 mM Tris/100 mM Boric acid/2 mM EDTA was used as running buffer. Electrophoresis was performed at constant power, 100 W, for approximately 2.5 h. After electrophoresis, gels were dried under vacuum at 80°C and exposed to Kodak BioMax MR film for approximately 24 h.

Microsatellite analysis

Microsatellite analysis was performed using four primer pairs of microsatellite loci (Sulu03, Sulu05, Sulu10 and Sulu11; see table 3.2 on page 57 and table 3.3 on page 63). Although thirteen polymorphic and codominant microsatellite loci were isolated in *S. luteus*, nine loci were not considered useful in this study because of null alleles or genomic duplications occuring at relatively high frequencies in the total sample. PCR reactions were set up in total volumes of 20 μl containing 0.5 μM of each primer, 1 U *Taq* DNA polymerase, 10 mM Tris-HCl pH 9, 15 mM $MgCl_2$, 50 mM KCl, 200 μM of all four dNTPs and approximately 2 ng of template DNA. Forward primers of each primer pair were labelled with a fluorescent near-infrared dye (IRDye^{TM} 700 or IRDye^{TM} 800) and the following PCR temperature profile was used: initial denaturation at 95°C for 3 min, 35 cycles of 95°C for 20 s, locus specific annealing temperature (see tables 3.2 and 3.3) for 20 s, 72°C for 1 min, followed by a final extension at 72°C for 60 min. PCR products were analyzed using the Global IR^2 system (LI-COR, Lincoln, NE, USA) and the microsatellites were scored with SAGA^{GT} software (LI-COR, Lincoln, NE, USA).

Statistical analysis of AFLP data

The AFLP fingerprints were manually scored for the presence or absence of individual fragments. In order to analyze the fingerprints, bands of equal fragment size were assumed to be homologous and the relative intensity of the

bands was not considered to be informational. Only loci with clearly amplified bands were retained for further data analysis.

The population structure of the sample was inferred using the program STRUCTURE (Version 2.1; Pritchard et al. (2000)). This program implements a model-based clustering method for the inference of population structure using genotype data consisting of unlinked markers. In brief, a model is assumed in which there are K populations (where K may be unknown), each of which is characterized by a set of allele frequencies at each locus. Individuals in the sample are assigned probabilistically to populations, or jointly to two or more populations if their genotypes indicate admixture. It is assumed that within populations, the loci are at Hardy-Weinberg and linkage equilibrium, and individuals are assigned to populations in such a way as to achieve this. Because it is not possible to distinguish all the genotypes with AFLP data, each genotype class was treated as being, effectively, a haploid allele and the model of no admixture was applied. This model states that each individual originates purely from one of the sampled populations and does not implement mixed ancestry of the individual genotypes. Furthermore, allele frequencies were assumed to be correlated among populations. In order to estimate K, the number of populations most appropriate for interpreting the data, a series of independent runs for each value of K between 1 and 9 was conducted. A burn-in period of 50000 iterations was chosen and data were collected for 250000 iterations. Three independent runs were done for each value of K and all produced highly consistent results.

The genetic structure of the geographic subpopulations was analyzed using the Bayesian method suggested by Holsinger et al. (2002), which allows for the estimation of the fixation index F_{ST} (Wright, 1978) from dominant markers without prior knowledge of the degree of inbreeding within populations and without assuming Hardy-Weinberg equilibrium within populations. Some information about the level of inbreeding within populations is also obtained by using this method. The program HICKORY (Version 1.0) was used with a full model and a model which assumes no inbreeding within populations. Several runs, using noninformative priors for the coefficient of inbreeding (F_{IS}) and the fixation index, were performed and all gave consistent results (burn-in = 50000, sample = 250000, thin = 50). Pairwise genetic distances between the geographic subpopulations were calculated as pairwise F_{ST} estimates in ARLEQUIN (Version 2.001; Schneider et al. (2000)) and the significance of the obtained estimates was assessed with a permutation test (10000 permutations).

Linkage disequilibrium between individual AFLPs was assessed according to Miyashita et al. (1999) by using a χ^2 test implemented in the program POPGENE (Version 1.32; Yeh et al. (1995)). Both Bayesian methods, implemented in STRUCTURE and HICKORY, assume that the markers are unlinked within populations. However, there seems to be more linkage disequilibrium

than expected by chance alone within the sampled subpopulations (see section 3.3.2). Therefore, the robustness of the different fitted models was tested by running the respective programs using random subsets of the data.

Analysis of molecular variance (AMOVA; Excoffier et al. (1992)), as implemented in the software program ARLEQUIN (Version 2.001; Schneider et al. (2000)), was applied to the data in order to partition the genotypic variance into components attributable to different hierarchical levels. Two population structure models were examined. In the first model no grouping of the geographic subpopulations was defined and AMOVA partitioned the total variance into components due to differences between subpopulations and differences between individuals within subpopulations. In the second model the structure suggested by the Bayesian cluster analysis was subjected to AMOVA. AMOVA was based on the pairwise squared Euclidian distances among the AFLP phenotypes and the significance of the variance components at the different hierarchical levels was assessed with a permutation procedure (15000 permutations).

In order to assess the effect of the heavy metal contamination on the pattern of genetic variation, the genetic diversity of subpopulations inhabiting polluted soils was compared to the diversity of subpopulations inhabiting non-polluted soils and the genetic differentiation between these two groups of subpopulations was examined using nested analysis of molecular variance. Genetic diversity of the geographic subpopulations was calculated as the percentage of polymorphic loci and Nei's unbiased heterozygosity (1978), averaged over loci, using the AFLP-SURV program (Version 1.0; Vekemans et al. (2002)). Because AFLP loci segregate as dominant markers, Hardy-Weinberg equilibrium was assumed in order to estimate allele frequencies and subpopulation heterozygosities. The frequency of the recessive allele was estimated at each locus using a Bayesian method with a non-uniform prior distribution of the allele frequencies, resulting in less biased estimates of the genetic diversity (Zhivotovsky, 1999). Alternatively, heterozygosities were calculated using the average value of the F_{IS} estimates that were obtained with the microsatellite data.

Statistical analysis of microsatellite data

For each geographic subpopulation of *S. luteus*, genetic diversity was estimated by standard genetic parameters obtained with GENETIX (Version 4.04 (Belkhir et al., 2002)): mean number of alleles per locus (A_O), observed heterozygosity (H_O) and expected heterozygosity (H_E) under the hypothesis of Hardy-Weinberg genotypic proportions (Nei, 1978). Estimates of diversity obtained with microsatellite markers were compared with the estimates obtained with AFLPs by computing Spearman's rank coefficient correlation (Sokal and Rohlf, 1995). A significant correlation would indicate congruence

between both types of markers and the impact of marker nonspecific factors on the level of diversity. Conversely, a lack of correlation is likely to reveal that marker specific factors are sufficiently different between populations to generate different levels of diversity. Departures from Hardy-Weinberg equilibrium and linkage equilibrium were calculated and tested with the web-based version of GENEPOP (Version 3.4; Raymond and Rousset (1995)). In both cases, an unbiased estimate of the exact P-value of the test was obtained by using a Markov-chain algorithm with 10000 dememorization steps, 1000 batches and 10000 iterations per batch.

A model-based clustering method was used to infer the genetic structure of the populations and to define the number of clusters (gene pools) in the dataset using the software STRUCTURE (Version 2.1; Pritchard et al. (2000)). A model that allows admixture (mixed ancestry of individuals) and that assumes correlation of allele frequencies among populations (due to migration or shared ancestry) was applied. In order to estimate K, the number of populations most appropriate for interpreting the data, a series of three independent runs for each value of K between 1 and 9 was performed. A burn-in period of 100000 iterations was chosen and data were collected for 350000 iterations. Because locus Sulu03 showed significant deviation from Hardy-Weinberg equilibrium, which is one of the assumptions of the model, data for locus Sulu03 were excluded from the analysis.

Analysis of molecular variance (AMOVA; Excoffier et al. (1992)), implemented in the software program ARLEQUIN (Version 2.001; Schneider et al. (2000)), was applied to the multilocus SSR data in order to partition the genotypic variance into components attributable to different hierarchical levels. Three population structure models were examined. In the first model no grouping of the geographic subpopulations was defined and AMOVA partitioned the total variance into components due to differences between subpopulations and differences between individuals within subpopulations. In the second model the structure suggested by the Bayesian cluster analysis based on AFLP data was subjected to AMOVA. In the third model nested analysis of molecular variance was applied to estimate the genetic differentiation between *S. luteus* subpopulations inhabiting polluted soils and the subpopulations inhabiting nonpolluted soils. In all cases pairwise distances were defined as the sum of squared differences in allele size (R_{ST}) and the significance of the variance components at the different hierarchical levels was assessed with a permutation procedure (15000 permutations).

Pairwise genetic differentiation between the geographic subpopulations of *S. luteus* was quantified by estimation of R_{ST} (Slatkin, 1995) using the software RSTCALC (Version 2.2; Goodman (1997)). R_{ST}, estimated as *Rho* values which account for differences in sample sizes, was calculated from the data after standardization to correct for differing variance in allele size between

loci (Goodman, 1997). The significance of the R_{ST} values was tested using permutation tests (Hudson et al., 1992) and bootstrapping (Efron, 1979) implemented in RSTCALC.

Allele frequency data was tested for evidence of recent genetic bottleneck events using the software BOTTLENECK (Piry et al., 1999). Populations that have experienced a recent reduction of their effective population size (between $2N_e$ and $4N_e$ generations according to Piry et al. (1999)) exhibit a correlative reduction of the allele numbers (A_O) and the genetic diversity (H_E) at polymorphic loci. The number of alleles is thereby reduced faster than the genetic diversity and therefore, the observed genetic diversity in a recently bottlenecked population is higher than the expected genetic diversity (H_{eq}), calculated from the observed number of alleles, under the assumption of mutation-drift equilibrium (Cornuet and Luikart, 1996). Calculations were performed using three possible microsatellite mutation models: the infinite allele model (IAM), the stepwise mutation model (SMM) and the two-phase model of mutation (TPM), which is intermediate to the IAM and SMM (95% of SMM and 5% of multiple-step mutation and a variance of 12 among multiple steps, as recommended by Piry et al. (1999)). A Wilcoxon sign rank test was conducted in order to determine whether there is a significant heterozygote excess, which may be indicative of a recent bottleneck. Because the Wilcoxon sign rank test is particularly sensitive to deviations from Hardy-Weinberg equilibrium, data collected at locus Sulu03 was excluded from the analysis.

The critical significance level applied in all statistical tests was 0.05. In all simultaneous statistical tests the critical significance levels were corrected using the sequential Bonferroni test (Rice, 1989) to enable overall significance to be examined.

3.3.2 Results

AFLP and microsatellite markers

AFLP and microsatellite analyses were performed with DNA extracted from 164 different individuals sampled across nine different geographic subpopulations of *S. luteus*, six of which inhabit heavy metal polluted soils (Lm, Ls, Lc, Lk, N and Of; 110 individuals) and three which inhabit nonpolluted soils (P, E and Hh; 54 individuals). AFLP data were generated with thirteen different primer combinations (see table 3.5) and these revealed a total of 361 usable markers, 347 (96%) of which were polymorphic in at least one subpopulation (figure 3.8 presents an example of an AFLP autoradiograph). AFLP analysis did not identify private alleles, nor were any of the alleles restricted to subpopulations inhabiting heavy metal polluted soils.

Primer combination	*Eco*RI-primer +	*Mse*I-primer +	# Loci scored
1	AAG	CCA	40
2	AAG	CTA	22
3	AAG	CTG	16
4	AAG	CTT	46
5	ACA	CAT	30
6	ACA	CCT	32
7	ACA	CTC	30
8	ACA	CTT	39
9	AGA	CAT	16
10	AGA	CCA	17
11	AGA	CCT	26
12	AGA	CTC	30
13	AGA	CTG	17
Total			361

Table 3.5: Overview of the AFLP primer combinations used for selective amplification in the analysis of *suillus luteus* populations and the number of loci scored for each primer combination.

Only four of the thirteen microsatellite loci that were isolated in *S. luteus* were applied in this study. Nine of the microsatellite loci showed relatively high frequencies of null alleles or amplification at more than one locus in the total sample and this is believed to be due to the high genetic diversity of *S. luteus* populations. The amplification of more than two alleles with a single microsatellite primer pair, observed in several sampled individuals, might be due to genomic duplications or to the presence of supernumerary chromosomes which have been reported to cause genomic variability within fungal species (Le Quéré et al., 2002). All four microsatellite loci were polymorphic in all geographic subpopulations and sixty-eight different alleles were detected in total (see figure 3.9 for an example of a microsatellite fingerprint gel image). The most variable locus was Sulu03 with thirty alleles and the least variable locus was Sulu11 with six alleles. A total of 24 private alleles was observed at these four loci and all subpopulations had at least one private allele. An overview of the allele size distributions at the four microsatellite loci is presented in figure 3.10.

Based on these AFLP and microsatellite markers, 163 different genotypes could be distinguished, as two individual samples from the Hh population shared the same AFLP and microsatellite fingerprints. Because genetic identity among fungal isolates has been suggested as the basis for the concept of individuality (Rayner, 1991), the data of only one of these two individual samples was retained for further analysis.

Within-population diversity

AFLP markers. The genetic diversity in the total sample, estimated as Nei's unbiased heterozygosity (Nei, 1978), was 0.326. The proportion of polymorphic loci within subpopulations varied between 44% (Of) and 89% (Ls), whereas heterozygosity ranged from 0.265 (Lk) to 0.352 (Lm). The average diversity was 0.306 ± 0.027 for subpopulations inhabiting polluted soils and 0.298 ± 0.016 for subpopulations inhabiting nonpolluted soils. Using the average fixation index F_{IS} obtained from the microsatellite data, heterozygosity ranged from 0.257 (Lk) to 0.341 (Lm) and the average diversity was 0.297 ± 0.028 for subpopulations in polluted habitats and 0.287 ± 0.018 for subpopulations in nonpolluted habitats (see table 3.6).

Microsatellite markers. Allelic richness, measured as the mean observed number of alleles per locus (A_O), varied between 4.50 in Of and 9.25 in Ls. The genetic diversity in terms of the observed heterozygosity (H_O) ranged from 0.588 ± 0.175 in Lk to 0.778 ± 0.203 in Hh. The average diversity (H_O) was 0.685 ± 0.0484 for subpopulations inhabiting polluted soils and 0.673 ± 0.0650 for subpopulations inhabiting nonpolluted soils. Values of H_O in all

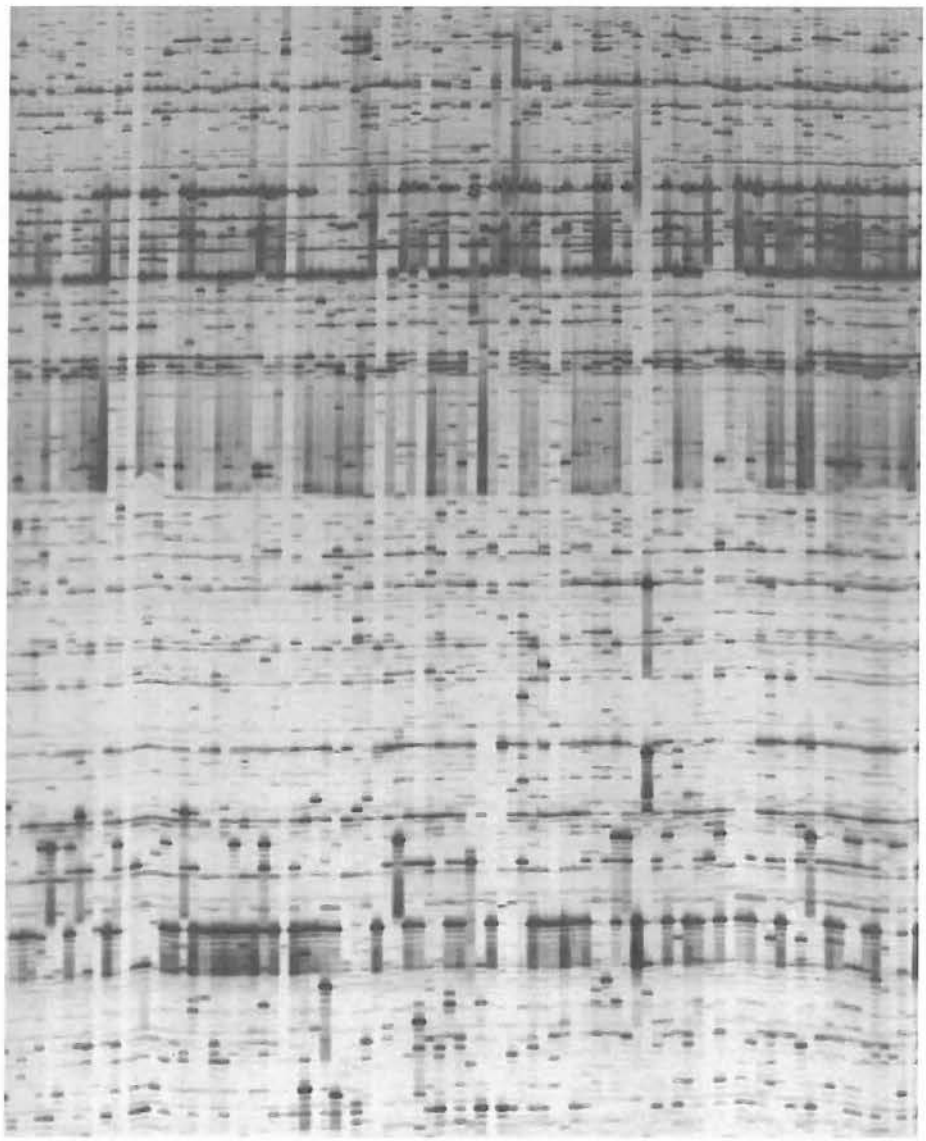

Figure 3.8: Example of an autoradiograph obtained by AFLP analysis of samples from several *Suillus luteus* populations, using the primer combination *Eco*RI+ACA/*Mse*I+CAT. Vertical lanes correspond to the fingerprints of different individuals; fragments of equal length, which are considered homologous, migrated the same distance from the top of the gel and form horizontal lanes.

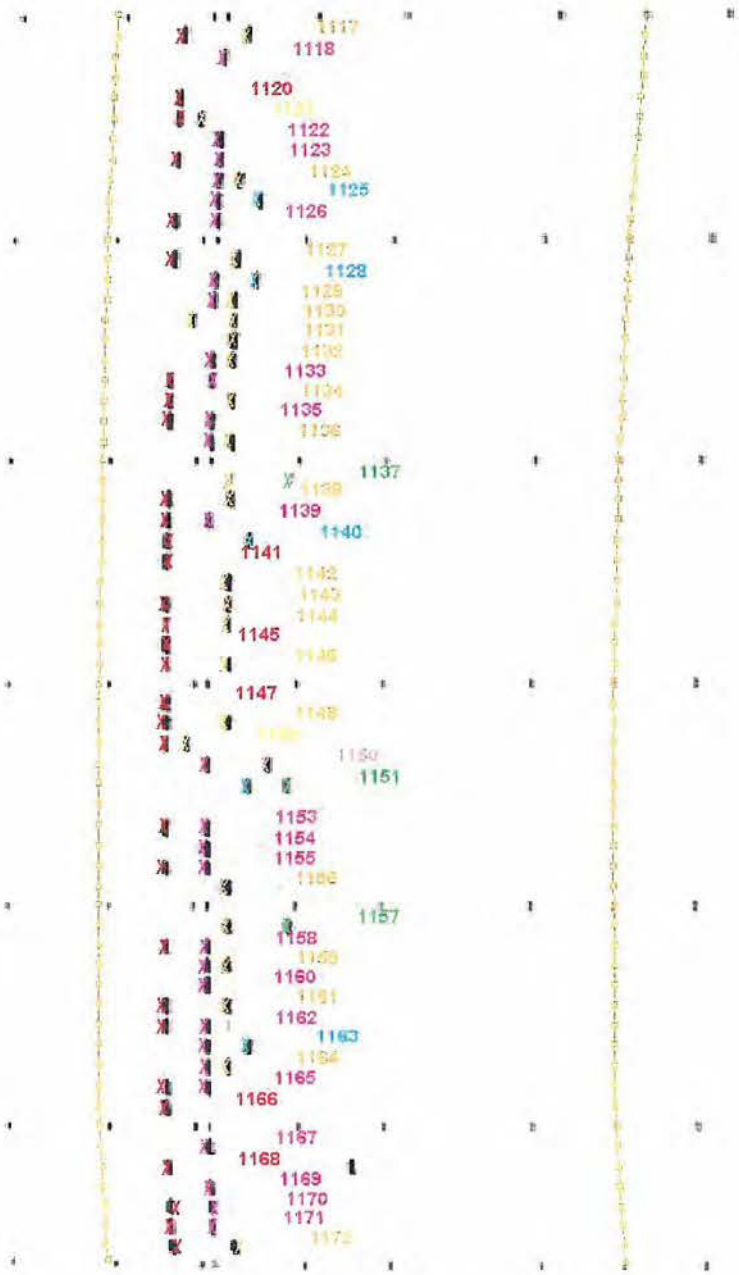

Figure 3.9: Example of a microsatellite fingerprint gel image obtained by PCR of locus Sulu11 in samples from different *Suillus luteus* populations. Vertical lanes correspond to the fingerprints of different individuals; alleles of equal length migrated the same distance from the top of the gel and form horizontal lanes.

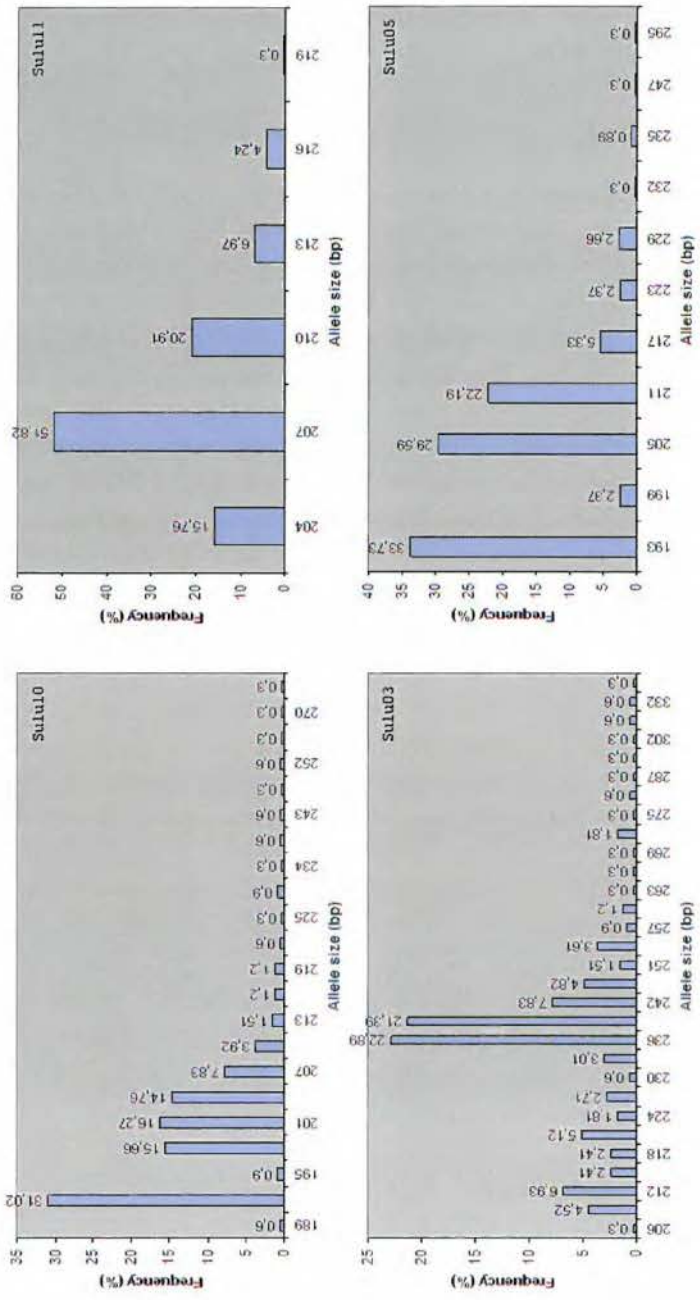

Figure 3.10: Distribution of microsatellite allele sizes at four loci in a sample of 164 *Suillus luteus* individuals.

populations were lower than H_E values, indicating a possible deficiency of heterozygous individuals with respect to Hardy-Weinberg expectations (see table 3.6). Spearman's rank coefficient correlation calculated with the diversity estimates (H and H_O) based on AFLP and microsatellite data was 0.78, which is significant at the 0.05 level.

Deviations from Hardy-Weinberg and linkage equilibria

AFLP markers. Bayesian analysis of the geographic population structure suggested that there was only weak evidence that F_{IS}, the inbreeding coefficient, was significantly different from zero, as the slightly higher value of the Deviance Information Criterion (DIC) for the model with $F_{IS} = 0$ ($DIC = 12013$), in comparison with the value of DIC for the full model ($DIC = 12004$), was entirely due to model dimension (Holsinger and Wallace, 2004). Linkage disequilibrium was examined using a χ^2 test between all the polymorphic loci within the different subpopulations. The proportion of significant linkage disequilibria, using a significance level of 0.05, varied between 9.8% and 17.1%, which is higher than expected by chance alone. In the total sample 22.8% of the pairwise combinations showed significant linkage disequilibrium at the 5% level. However, the linkage disequilibrium seems to be sporadic, as no specific locus pair was in linkage disequilibrium in all subpopulations.

Microsatellite markers. Significant deviation from Hardy-Weinberg equilibrium was found in six geographic subpopulations (Ls, Lc, Lk, N, P and Hh; see table 3.7). However, the significant heterozygosity deficit after Bonferroni corrections was entirely due to the highly significant disequilibrium at locus Sulu03. All other loci did not show significant departure from Hardy-Weinberg equilibrium in most populations (except Sulu11 in Hh; see table 3.7). Linkage disequilibrium across subpopulations was found to be significant between loci Sulu03 and Sulu05 and between Sulu03 and Sulu11. When linkage disequilibrium was assessed within subpopulations, significant disequilibrium between loci Sulu03 and Sulu05 and between Sulu03 and Sulu11 was only observed in Lm. Therefore, it is unlikely that these loci are physically linked.

Genetic population structure

AFLP markers. In the model-based clustering analysis, the estimate of the posterior probability of the value of K (the number of subpopulations) more-or-less plateaus for $K = 2$. The major structure in this data set is thus captured by a model population structure consisting of two differentiated subpopulations. Table 3.8 reports the number of individuals of each sam-

			AFLP markers			SSR markers	
Habitat type	Pop.	N	P (%)	H	H'	A_O	H_O	H_E
Contaminated	Lm	10	73	0.352	0.341	6.75	0.746 (0.0827)	0.764 (0.0957)
	Ls	36	89	0.318	0.309	9.25	0.699 (0.187)	0.780 (0.0480)
	Lc	11	79	0.343	0.338	6.75	0.727 (0.223)	0.804 (0.0753)
	Of	6	44	0.304	0.302	4.50	0.700 (0.258)	0.711 (0.232)
	N	27	74	0.290	0.277	9.00	0.694 (0.108)	0.783 (0.110)
	Lk	20	58	0.265	0.257	7.25	0.588 (0.175)	0.750 (0.0988)
Weighted average				0.306 (0.027)	0.297 (0.028)	8.08 (1.35)	0.685 (0.0484)	0.772 (0.0213)
Noncontaminated	Hh	8	60	0.314	0.305	6.75	0.778 (0.203)	0.817 (0.0835)
	E	12	67	0.322	0.312	6.50	0.739 (0.176)	0.785 (0.0962)
	P	33	75	0.285	0.273	8.75	0.624 (0.243)	0.778 (0.104)
Weighted average				0.298 (0.016)	0.287 (0.018)	7.94 (1.05)	0.673 (0.0650)	0.785 (0.0137)

Table 3.6: Genetic diversity estimates based on 361 AFLP markers and four microsatellite loci in nine geographic subpopulations of *Suillus luteus* inhabiting heavy metal contaminated and noncontaminated areas in Belgium; N, subpopulation sample size; P, percentage of polymorphic loci; H, unbiased heterozygosity (Nei, 1978); H', unbiased heterozygosity calculated with the value of F_{IS} obtained with the microsatellite data; A_O, mean number of alleles per locus; H_O, the mean observed heterozygosity; H_E, the mean unbiased expected heterozygosity (Nei, 1978); SD is given in parentheses.

	Sulu03		Sulu05		Sulu10		Sulu11		Over all loci	
Population	F_{IS}	P-value	F_{IS}	P-value	F_{IS}	P-value	F_{IS}	P-value	F_{IS}	P-value
Lm	0.091	0.23	0.018	0.46	0.12	0.49	−0.15	0.70	0.023	0.58
Ls	0.46	< 0.001	−0.055	0.40	−0.057	0.48	0.056	0.61	0.10	< 0.001
Lc	0.49	< 0.001	0.042	0.23	−0.17	0.72	−0.0060	0.99	0.10	0.0075
Lk	0.58	< 0.001	0.024	0.68	0.12	0.40	0.098	0.22	0.22	0.005
N	0.35	< 0.001	0.027	0.071	0.015	0.55	0.015	0.45	0.11	< 0.001
Of	0.35	0.15	−0.10	0.99	−0.21	0.85	−0.067	0.99	0.018	0.85
P	0.65	< 0.001	−0.080	0.18	−0.051	0.65	0.20	0.051	0.20	< 0.001
E	0.10	0.23	0.036	0.27	−0.15	0.15	0.27	0.19	0.057	0.12
Hh	0.28	0.019	−0.076	0.40	−0.23	0.17	0.22	0.033	0.051	0.010

Table 3.7: Deviation from Hardy-Weinberg equilibrium measured by Wright's fixation index, F_{IS} (Weir and Cockerham, 1984), in populations of *Suillus luteus* at four microsatellite loci; significance was tested using Markov-chain randomization (dememorization 10000, batches 1000, iterations per batch 10000) implemented in GENEPOP (Version 3.4; Raymond and Rousset (1995)).

Population	# Individuals in cluster 1	#Individuals in cluster 2
Lm	6	4
Ls	25	11
Lc	4	7
Of	6	0
N	26	1
Lk	20	0
Hh	7	1
E	7	5
P	31	2

Table 3.8: The numbers of individuals from nine geographic subpopulations of *Suillus luteus* that belong to each of the two clusters (gene pools) derived by a model-based clustering analysis.

pled geographic subpopulation that belongs to each of the derived clusters for $K = 2$. Eighty-one percent of the sampled individuals was assigned to cluster 1, which contained individuals from all geographic subpopulations. Cluster 2 contained 19% of the samples and these originated mainly from the subpopulations Lm, Ls, Lc and E. All individuals were assigned with high probabilities to either one of the two clusters, except for sample N19, which was assigned to each of the derived clusters with equal probability. The highest probabilities for K using subsets of the data were similar to those obtained for the whole data set, which suggests that the result was robust and not due to the linkage disequilibria within the subpopulations. Bayesian analysis of the geographic population structure indicated limited differentiation between the sampled subpopulations. The values of the posterior mean of F_{ST} were similar when using the full model and the model with $F_{IS} = 0$, being 0.036 ± 0.004 and 0.034 ± 0.003, respectively. Again, similar results were obtained when running the analysis with subsets of the entire data set. Pairwise genetic distances (F_{ST}) between the geographic subpopulations varied between 0.00740 ($P > 0.05$) and 0.185 ($P < 0.05$; see table 3.11). In the first model examined, AMOVA indicated most genetic variation among individuals within geographic subpopulations (95%), although variation among subpopulations was highly significant ($P < 0.001$; see table 3.9a). Grouping of the individual samples according to the Baeysian clustering allowed 19% of the total variation ($P < 0.001$) to be accounted for by differences between the two clusters, suggesting considerable genetic differentiation (see table 3.9b). Nested AMOVA showed no significant grouping of the subpopulations inhabiting contaminated

soils versus the subpopulations inhabiting noncontaminated soils (see table 3.9c).

Microsatellite markers. Model-based clustering analysis could not detect population substructuring of the total sample based on the microsatellite data. For all values of K between 1 and 9, the proportion of the total sample assigned to each cluster was roughly symmetric ($\sim 1/K$ in each cluster) and almost all individuals were indicated to be admixed. AMOVA indicated most of the genetic variation to be due to differences among individuals within geographic subpopulations (99%), whereas the differantiation between subpopulations was not found to be significant (1%, $P = 0.38$; see table 3.10a). No significant differentiation was found when AMOVA was performed using the structure inferred by the model-based clustering analysis with AFLP data (see table 3.10b) and nested AMOVA showed no significant grouping of the subpopulations inhabiting contaminated soils versus the subpopulations inhabiting noncontaminated soils (see table 3.10c). Pairwise genetic distances, in terms of R_{ST} estimates, between the geographic subpopulations varied between 0.0010 ($P > 0.05$) and 0.1207 ($P < 0.05$; see table 3.11).

Genetic evidence for population bottlenecks

There is no evidence for significant heterozygosity excess in any of the sampled geographic subpopulations. Under all three mutation models, all subpopulations appear to be in mutation–drift equilibrium in the Wilcoxon sign rank test ($P > 0.05$).

3.3.3 Discussion

Bayesian inference based on AFLP data suggested a model population structure that consists of two differentiated clusters of *S. luteus* samples, one of which comprises 81% of the total sample. Analysis of molecular variance supported this result and indicated considerable genetic differentiation between the two clusters. However, no additional support for this population structure was provided by the microsatellite data. Cornuet et al. (1999) reported that limited statistical power is an important constraint faced when using microsatellite markers in any type of individual-based population assignment procedure when population differentiation is low ($F_{ST} < 0.05$) and in these cases, AFLP markers have been shown to be an excellent alternative to microsatellites in order to enhance the resolution of the assignment procedure (Campbell et al., 2003). Although genetic differentiation was relatively high between the two derived clusters (AMOVA based on AFLP data indicated 19% of the total variance to be due to between cluster differences), no differentia-

Source of variation	D.f.	Sum of squares	Variance component	% of total	P-value
(a) No partitioning of the sampled populations					
Among populations	8	485.005	1.65684	4.93	< 0.001
Within populations	154	4916.155	31.92308	95.07	
Total	162	5401.160	33.57993		
(b) Structure based on Bayesian clustering					
Among clusters	1	395.210	7.25209	18.91	< 0.001
Within clusters	161	5005.949	31.09285	81.09	
Total	162	5401.160	38.34494		
(c) Nested analysis – contaminated vs. noncontaminated soils					
Among groups	1	48.242	-0.41037	-1.23	0.6
Among populations within groups	7	436.763	1.86866	5.60	< 0.001
Within populations	154	4916.155	31.92308	95.63	< 0.001
Total	162	5401.160	33.38137		

Table 3.9: Analysis of molecular variance (AMOVA; Excoffier et al. (1992)) for 163 individuals of *Suillus luteus* sampled from 9 geographic subpopulations in Belgium, employing 361 AFLP markers. (a) AMOVA without further partitioning of the sampled subpopulations; (b) AMOVA contrasting 2 differentiated groups of individuals inferred by Bayesian clustering; (c) Nested analysis contrasting subpopulations colonizing heavy metal contaminated soils (N, Of, Lm, Ls, Lc, Lk) and subpopulations inhabiting noncontaminated soils (E, Hh, P). Data show the degrees of freedom (d.f.), the sum of squared deviations, the variance component estimates, the percentage of total variance contributed by each component and the significance of the variance components (P-value) estimated computing 15000 permutations.

Source of variation	D.f.	Sum of squares	Variance component	% of total	*P*-value
(a) No partitioning of the sampled populations					
Among populations	8	6647.405	6.53778	1.05	0.38
Within populations	303	186011.931	613.90076	98.95	
Total	311	192659.337	620.43854		
(b) Structure inferred by Bayesian clustering with AFLP data					
Among clusters	1	1300.112	7.04502	1.13	0.19
Within clusters	310	191359.064	617.28730	98.87	
Total	311	192659.176	624.33232		
(c) Nested analysis – contaminated vs. noncontaminated soils					
Among groups	1	225.712	-6.04788	-0.98	0.77
Among populations within groups	7	6421.694	9.56493	1.55	0.29
Within populations	303	186011.931	613.90076	99.43	0.38
Total	311	192659.337	617.41781		

Table 3.10: Analysis of molecular variance (AMOVA; Excoffier et al. (1992)) for 163 individuals of *Suillus luteus* sampled from 9 geographic subpopulations in Belgium, employing four microsatellite markers. (a) AMOVA without further partitioning of the sampled subpopulations; (b) AMOVA contrasting 2 differentiated groups of individuals inferred by Bayesian clustering with AFLP data; (c) Nested analysis contrasting subpopulations colonizing heavy metal contaminated soils (N, Of, Lm, Ls, Lc, Lk) and subpopulations inhabiting noncontaminated soils (E, Hh, P). Data show the degrees of freedom (d.f.), the sum of squared deviations, the variance component estimates, the percentage of total variance contributed by each component and the significance of the variance components (*P*-value) estimated computing 15000 permutations.

	Lm	Lc	Lk	Ls	N	Of	E	Hh	P
Lm		0.0163	0.0881*	neg.	0.0576*	0.0515*	0.0243	0.0326	0.0586*
Lc	0.0188		0.185*	0.0458*	0.147*	0.130*	0.0326*	0.120*	0.151*
Lk	0.0518	0.00290		0.0663*	0.0165*	0.0536*	0.0900*	0.0418*	0.0178*
Ls	0.0333	neg.	neg.		0.0403*	0.0546*	0.00740	0.0299*	0.0424*
N	0.0452*	neg.	neg.	neg.		0.0152	0.0492*	0.0251*	0.00880
Of	0.101*	0.0498	neg.	0.0443	neg.		0.0497*	0.0299	0.0232
E	0.0792*	neg.	0.0338	neg.	0.00520	0.0646		0.0562*	0.0528*
Hh	0.0981*	neg.	0.0219	neg.	0.0390	0.121*	0.0340		0.0242*
P	0.0558*	0.00360	neg.	0.00100	neg.	neg.	0.0233	0.0345	

Table 3.11: Pairwise genetic distances between nine geographic subpopulations of *Suillus luteus*, measured by estimates of R_{ST} (Slatkin, 1995) with data from four microsatellite loci (below the diagonal) and by estimates of F_{ST} with data from 361 AFLP markers (above the diagonal); neg., indicates negative estimates obtained due to lack of population differentiation; *, indicates significant values ($P < 0.05$, based on 10000 permutations).

tion was apparent when using the microsatellite data. Because microsatellite markers show higher mutation rates than other marker types, lower levels of population differentiation can be expected and homoplasy due to these high mutation rates may further increase the underestimation of differentiation (Estoup et al., 1995). Thus, the failure of recognizing any structure based on the microsatellite data may be due to a lack of statistical power obtained with these loci. On the other hand, the dominant property of the AFLP markers is also responsible for a bias in the estimation of population differentiation. However, Krutovskii et al. (1999) showed that reasonable estimates of differentiation can be obtained with RAPD markers if sample sizes are larger than 30 individuals and therefore the estimates obtained with the AFLP markers obtained in this study may be reasonably accurate, although an overestimation of differentiation cannot be excluded.

The presence of two differentiated clusters in this sample may indicate the existence of different mating groups in *S. luteus*, as sexual incompatibility is often correlated with genetic divergence (Aanen et al., 2000). However, no intercompatibility groups have been detected in *S. luteus* populations so far. One study described the conspecificity of American and European isolates, but pairwise compatibility between only five isolates was examined (Fries and Neumann, 1990). Alternatively, this result can be explained by considering *S. luteus* to be a very heterogeneous species or a species complex that consists of two or more cryptic species (i.e. species indistinguishable by typical morphological criteria or sexual compatibility). Based on molecular phylogenies, cryptic species have already been suggested for several morphospecies within various fungal genera, including *Fusarium* (Skovgaard et al., 2002), *Tricholoma* (Horton, 2002) and *Cenococcum* (Douhan and Rizzo, 2005).

Both AFLP and microsatellite markers revealed high levels of genetic diversity within *S. luteus* geographic subpopulations, but genetic differentiation between subpopulations was limited. Most of the genetic variation in the total sample was due to differences between individuals within subpopulations and the estimated value of F_{ST} (0.036 based on AFLP data) was low. Moreover, model-based clustering using AFLP data grouped most of the sampled individuals (81%) into one subpopulation, while clustering based on microsatellite data did not reveal any substructure in the total sample. These findings are most likely the result of substantial gene flow between the subpopulations and frequent sexual reproduction, which was also suggested by the high proportion of unique multilocus genotypes. Comparable estimates of genetic diversity and differentiation were reported for populations of *S. grevillei* (Zhou et al., 2001) and *S. pungens* (Bonello et al., 1998). Bayesian analysis of the AFLP data and estimation of F_{IS} with the microsatellite data indicated absence of inbreeding, which was expected *a priori* due to the outbreeding nature of *S. luteus*. However, a significant reduction of the heterozygosity was detected at

one microsatellite locus (Sulu03), both at the level of the total sample and in most of the geographic subpopulations. The homozygosity excess at this locus may be due to the presence of null alleles or may be fitness related (either directly, which is unlikely for microsatellite loci, or due to physical linkage of the marker with a trait locus).

Considerably larger estimates of genetic diversity (more than two-fold) were obtained with the microsatellite data compared to the values obtained with the AFLP data. Diversity assessed at different marker types results from either marker specific factors (e.g. mutation) or marker nonspecific factors (e.g. genetic drift or migration). Hence, it is expected that diversity estimates of a given population will be different for different marker types. However, if two populations went through different evolutionary histories involving drift and migration, then they would differ for their level of diversity in a similar manner for different marker types. In other words, congruent rankings of the populations between markers would be expected. As a result, the comparative analysis of diversity between markers is not based on the comparison of absolute values of diversity levels within populations, but on the comparison of the ranking of different populations by correlation analysis. Estimates of diversity obtained with AFLP and microsatellites were congruent, suggesting that observed subpopulation differences in diversity are more related to marker nonspecific than marker specific causes.

A priori it was expected that heavy metal pollution would have a large effect on the genetic structure of *S. luteus* populations. *In vitro* growth experiments have shown a high correlation between Zn tolerance and habitat pollution in *S. luteus* populations and a genetic basis for the tolerance trait is assumed. Hence, it would be expected that founder effects and a strong selection pressure for heavy metal tolerance force a population to pass through a bottleneck in which size and genetic variation are drastically reduced. However, no evidence was found for a consistent reduction of the genetic variation of *S. luteus* populations in polluted habitats and nested analysis of molecular variance revealed no effect of the heavy metal pollution on the genetic differentiation between subpopulations in contaminated and noncontaminated areas. Although significant linkage disequilibrium was found between several microsatellite loci in the Lm population, which may suggest past founder events or bottlenecks followed by rapid population expansion, no evidence was found for past bottleneck events using specific statistical tests based on microsatellite allele frequency data. Furthermore, none of the other subpopulations in contaminated areas showed significant linkage disequilibrium at these loci. Other possible causes of the observed linkage disequilibrium in the Lm subpopulation are stochastic sampling bias, admixture or natural selection unrelated to the heavy metal pollution.

High levels of morphological, enzymatic or genetic variation within popu-

lations inhabiting heavy metal polluted soils, which were at least as high as for populations inhabiting nonpolluted soils, were reported previously for several plant species such as *Silene vulgaris* (Baker and Dalby, 1980), *Silene paradoxa* (Mengoni et al., 2000), *Agrostis stolonifera* (Wu et al., 1975), *Armeria maritima* (Lefèbvre and Kokes, 1981; Vekemans and Lefèbvre, 1997) and *Arrhenatherum elatius* (Ducousso et al., 1990). Several nonexclusive processes have been proposed to explain this surprisingly high level of variation in tolerant populations. Successive colonization events or a high frequency of tolerance in natural populations could have reversed the effects of an initial genetic bottleneck. Migration from the neighbouring populations could also be a cause of variation as well as environmental heterogeneity (Hedrick et al., 1976) and human disturbance (Gouyon et al., 1983). The results of this study, however, suggest that a bottleneck might never have occurred in populations of *S. luteus* inhabiting polluted soils. Multiple genotypes could have been introduced when the plantations were initially established, followed by sexual reproduction on the site and rapid evolution of the tolerance trait, which would result in genetic diversity levels being higher than expected. In addition, recurrent migration of tolerant genotypes that originate due to admixture in nonpolluted areas may attribute to the high level of genetic diversity of populations inhabiting contaminated soils. Admixture between the tolerant populations and surrounding populations inhabiting nonpolluted soils is very likely, as indicated by the low level of population differentiation, and would explain the high frequency of tolerance observed in some of the populations surrounding polluted areas (see section 2.1.2 on page 14).

Chapter 4

General conclusions and future directions

In the present study, genomic and transcriptomic aspects of the adaptation to increased environmental heavy metal concentrations were analyzed in relation to phenotypic observations of heavy metal tolerance in populations of the ectomycorrhizal basidiomycete *Suillus luteus*. A substantial amount of ecological and physiological information was already available, suggesting that natural selection might have been at work: a clear difference in heavy metal tolerance, a trait which is likely to have a genetic basis, was shown between populations inhabiting polluted and populations inhabiting nonpolluted environments, and indications were found for a reduction of the genetic diversity of populations in polluted habitats (Colpaert et al., 2000).

S. luteus sporocarps were collected at nine different locations in the Campine phytogeographic district (province of Limburg, Belgium). At six of these locations the soil is severely polluted with heavy metals due to past activities of several zinc smelters and phenotypic characterization, by means of *in vitro* growth experiments, revealed an obvious increase of the heavy metal tolerance in the populations at these sites. However, evidence for natural selection was not found at the genomic level. AFLP and microsatellite marker analyses both indicated high levels of genetic diversity within the total sample and within geographic subpopulations, but limited population differentiation. No current or past reduction of the genetic diversity of populations inhabiting polluted environments was discovered, nor was population differentiation found to be related to the environmental heavy metal pollution. Although independent of the heavy metal contamination, inference of the population structure based on the AFLP data resulted in the identification of two differentiated clusters, possibly indicating the presence of different mating groups or cryptic species in *S. luteus*.

In accordance with the high level of genomic diversity revealed by the more-or-less neutral AFLP and microsatellite markers, considerable diversity was found at the level of the transcriptome by cDNA-AFLP analysis of a heavy metal tolerant and a nontolerant isolate. Part of the variation in transcript profiles was shown to be due to sequence divergence between both isolates, but RT-PCR verified differences in expression level to cause most of the observed variation. Although transcript profiling should be regarded as a tool for exploratory data analysis and some of the differences in expression level will be due to genetic hitchhiking and natural regulatory variation, some information was obtained about possible mechanisms that may explain the difference in tolerance level. Expression of a hydrophobin homologue, a water-repellent protein in the cell wall which is believed to protect against penetration of heavy metal containing solutions, may be partly responsible for the higher external zinc concentrations required to induce modulation of transcription in the tolerant isolate and candidate genes that may play a role in heavy metal detoxification were identified as encoding a heat shock protein (HSP60), a putative metal transporter and several proteins involved in ubiquitin-dependent proteolysis.

The results of these different experiments show that heavy metal pollution has limited influence on the genetic diversity of *S. luteus* populations, a result that seems unlikely *a priori* but that has been described in several other organisms. In case of *S. luteus* bottleneck events associated with the foundation of populations in polluted environments may have been prevented by the introduction of different tolerant genotypes, followed by sexual reproduction on the site and rapid evolution of the tolerance trait. Additionally, recurrent migration of tolerant genotypes that originate due to admixture in nonpolluted areas, which is very likely due to the high migration rates, may have attributed to the high levels of genetic diversity of populations in polluted habitats. These findings are possibly correlated with the pioneer characteristics of this species. High genetic diversities in combination with frequent sexual reproduction and high migration rates would allow adaptation to a wide range of environmental conditions and rapid evolution of traits required in disturbed environments.

Based on the findings of this study future research may focus on genetic aspects of *S. luteus* populations, both related and unrelated to heavy metal pollution. A few of the many questions that still require explanation include:

- Does heavy metal tolerance appear in unexposed populations of *S. luteus*? If so, what is the frequency of this occurence?
- Do the differentiated clusters found with AFLP data correspond to different mating groups or to cryptic species? Do they interbreed?
- What are the characteristics of the genome of *S. luteus* (size, complexity,

ploidy)? How variable is this genome structure?

- What is the extent of natural regulatory variation in *S. luteus* populations?
- Do the genes identified by transcript profiling play a role in heavy metal tolerance? Do they induce increased tolerance in sensitive strains?

Ideally, an integrated approach should be undertaken to answer these questions, combining techniques from a wide range of disciplines. Eventually this may lead to the discovery of the genes and polymorphisms that affect the heavy metal tolerance trait and the mechanisms that underlie these effects. Therefore *S. luteus* should be considered as a valuable model species in evolutionary and ecological functional genomics and it is hoped that the results of this study may help in its acceptance.

Bibliography

D.K. Aanen, T.W. Kuyper, T.H.M. Mes, and R.F. Hoekstra. The evolution of reproductive isolation in the ectomycorrhizal *Hebeloma crustuliniforme* aggregate (basidiomycetes) in northwestern Europe: a phylogenetic approach. *Evolution*, 54:1192–1206, 2000.

K. Adriaensen, D. Van der Lelie, A. Van Laere, J. Vangronsveld, and J.V. Colpaert. A zinc-adapted fungus protects pines from zinc stress. *New Phytologist*, 161:549–555, 2004.

S.A. Al-Hiyaly, T. McNeilly, and A.D. Bradshaw. The effect of zinc contamination from electricity pylons – Evolution in a replicated situation. *New Phytologist*, 110:571–580, 1988.

S.A.K. Al-Hiyaly, T. McNeilly, A.D. Bradshaw, and A.M. Mortimer. The effect of zinc contamination from electricity pylons. Genetic constraints on selection for zinc tolerance. *Heredity*, 70:22–32, 1993.

S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25:3389–3402, 1997.

J. Antonovics, A.D. Bradshaw, and R.G. Turner. Heavy metal tolerance in plants. In J.B. Cragg, editor, *Advances in Ecological Research*, volume 7, pages 1–85. Academic Press, New York, NY, US, 1971.

C.W.B. Bachem, R.S. van der Hoeven, S.M. de Bruijn, D. Vreugdenhil, M. Zabeau, and R.G.F. Visser. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. *The Plant Journal*, 9: 745–753, 1996.

A.J.M. Baker and D.H. Dalby. Morphological variation between some isolated populations of *Silene maritima* With. in the british isles with particular

reference to inland populations on metalliferous soils. *New Phytologist*, 84: 123–138, 1980.

K. Belkhir, P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. GENETIX 4.04, logiciel sous Windows [tm] pour la génétique des populations, 2002. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier, France.

S.M. Berch. Endogonaceae: taxonomy, specificity, fossil record phylogeny. In K. G. Mukerji and V. P. Singh, editors, *Frontiers of Applied Microbiology*, volume 2. Print House, Luknow, India, 1987.

H. Besl and A. Bresinsky. Chemosystematics of Suilaceae and Gomphidiaceae (suborder Suillineae). *Plant Systematics and Evolution*, 206:223–242, 1997.

E. Biet, J. Sun, and M. Dutreix. Conserved sequence preference in DNA binding among recombination proteins: an effect of ssDNA secondary structure. *Nucleic Acids Research*, 27:596–600, 1999.

N. Bjerrum. *Bjerrum's inorganic chemistry*. Heinemann, London, UK, third Danish edition, 1936.

P. Bonello, T.D. Bruns, and M. Gardes. Genetic structure of a natural population of the ectomycorrhizal fungus *Suillus pungens*. *New Phytologist*, 138: 533–542, 1998.

A.D. Bradshaw. Populations of *Agrostis tenuis* resistant to lead and zinc poisoning. *Nature*, 169:1098, 1952.

A.D. Bradshaw and T. McNeilly. *Evolution and pollution*. Edward Arnold, London, UK, 1981.

E. Cabiscol, G. Belli, J. Tamarit, P. Echave, E. Herrero, and J. Ros. Mitochondrial Hsp60, resistance to oxidative stress, and the labile iron pool are closely connected in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 277:44531–44538, 2002.

J.W.G. Cairney. Evolution of mycorrhiza systems. *Naturwissenschaften*, 87: 467–475, 2000.

D. Campbell, P. Duchesne, and L. Bernatchez. AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. *Molecular Ecology*, 12:1979–1991, 2003.

P. Catasti, X. Chen, S.V. Mariappan, E.M. Bradbury, and G. Gupta. DNA repeats in the human genome. *Genetica*, 106:15–36, 1999.

A.M. Chaudri, B.P. Knight, V.L. Barbosa-Jefferson, S. Preston, G.I. Paton, K. Killham, N. Coad, F.A. Nicholson, B.J. Chambers, and S.P. McGrath. Determination of acute Zn toxicity in pore water from soils previously treated with sewage sludge using bioluminescence assays. *Environmental Science and Technology*, 33:1880–1885, 1999.

N. Chondrogianni, C. Tzavelas, A.J. Pemberton, I.P. Nezis, A.J. Rivett, and E.S. Gonos. Overexpression of proteasome β_5 subunit increases the amount of assembled proteasome and confers ameliorated response to oxidative stress and higher survival rates. *Journal of Biological Chemistry*, 280: 11840–11850, 2005.

J.V. Colpaert, P. Vandenkoornhuyse, K. Adriaensen, and J. Vangronsveld. Genetic variation and heavy metal tolerance in the ectomycorrhizal basidiomycete *Suillus luteus*. *New Phytologist*, 147:367–379, 2000.

J.V. Colpaert, L.A.H. Muller, M. Lambaerts, K. Adriaensen, and J. Vangronsveld. Evolutionary adaptation to Zn toxicity in populations of Suilloid fungi. *New Phytologist*, 162:549–560, 2004.

J.M. Cornuet and G. Luikart. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144:2001–2014, 1996.

J.M. Cornuet, S. Piry, G. Luikart, A. Estoup, and M. Solignac. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, 153:1989–2000, 1999.

A. Dahlberg and R.D. Finlay. *Suillus*. In J.W.G. Cairney and S.M. Chambers, editors, *Ectomycorrhizal Fungi. Key Genera in Profile*, pages 33–64. Springer-Verlag, Berlin, DE, 1999.

J.C. Debaud, R. Marmeisse, and G. Gay. Intraspecific genetic variation and populations of ectomycorrhizal fungi. In A.K. Varma and B. Hock, editors, *Mycorrhiza: Structure, Molecular Biology and Function*, pages 75–110. Springer Verlag, Berlin, DE, 1999.

R.F. Ditt, E.W. Nester, and L. Comai. Plant gene expression response to *Agrobacterium tumefaciens*. *Proceedings of the National Academy of Sciences of the United States of America*, 98:10954–10959, 2001.

N.V. Dokholyan, S.V. Buldyrev, S. Havlin, and H.E. Stanley. Distributions of dimeric tandem repeats in coding and noncoding DNA sequences. *Journal of Theoretical Biology*, 202:273–282, 2000.

G.W. Douhan and D.M. Rizzo. Phylogenetic divergence in a local population of the ectomycorrhizal fungus *Cenococcum geophilum*. *New Phytologist*, 166: 263–271, 2005.

A. Ducousso, D. Petit, M. Valero, and P. Vernet. Genetic variation between and within populations of a perennial grass: *Arrhenatherum elatius*. *Heredity*, 65:179–188, 1990.

J.H. Duffus. "Heavy metals" - a meaningless term? *Pure and Applied Chemistry*, 74:793–807, 2002.

W.E. Durrant, O. Rowland, P. Piedras, K.E. Hammond-Kosack, and J.D.G. Jones. cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. *Plant Cell*, 12:963–977, 2000.

B. Efron. Bootstrap methods: another look at the jack-knife. *Annals of Statistics*, 7:1–26, 1979.

D.J. Eide. Multiple regulatory mechanisms maintain zinc homeostasis in *Saccharomyces cerevisiae*. *The journal of Nutrition*, 133:1532–1535, 2003.

J.A. Eisen. Mechanistic basis for microsatellite instability. In D.B. Goldstein and C. Schlötterer, editors, *Microsatellites: Evolution and Applications*, pages 34–48. Oxford University Press, Oxford, UK, 1999.

A. Estoup, C. Tailliez, J.M. Cornuet, and M. Solignac. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular Biology and Evolution*, 12:1074–1084, 1995.

L. Excoffier, P.E. Smouse, and J.M. Quattro. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131:479–491, 1992.

I. Fabregat, K.S. Koch, and T. Aoki. Functional pleiotropy of an intramolecular triplex-forming fragment from the 3'-UTR of the rat *Pigr* gene. *Physiological Genomics*, 5:53–65, 2001.

M.E. Feder and T. Mitchell-Olds. Evolutionary and ecological functional genomics. *Nature Reviews Genetics*, 4:649–655, 2003.

N. Fries and W. Neumann. Sexual incompatibility in *Suillus luteus* and *S. granulatus*. *Mycological research*, 94:64–70, 1990.

J.H. Gillespie. *Population Genetics. A Concise Guide*. The Johns Hopkins University Press, Baltimore, MD, US, 1998.

T.C. Glenn and N.A. Schable. Isolating microsatellite DNA loci. *Methods in Enzymology*, 395:202–222, 2005.

C. Godon, G. Langniel, J. Lee, J.M. Buhler, S. Kieffer, M. Perrot, H. Boucherie, M.B. Tolendano, and J. Labarre. The H_2O_2 stimulon in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 273:22480–22489, 1998.

S.J. Goodman. R_{ST}Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology*, 6:881–885, 1997.

P.H. Gouyon, R. Lumaret, G. Valdeyron, and P.H. Vernet. Reproductive strategies and disturbance by man. In H.A. Mooney and M. Godron, editors, *Disturbance and Ecosystems. Component of Response. Ecological Studies 44*. Springer Verlag, London, UK, 1983.

J.L. Hall. Cellular mechanisms for heavy metal detoxification and tolerance. *Journal of Experimental Botany*, 53:1–11, 2002.

T.A. Hall. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series*, 41:95–98, 1999.

C.A. Hamilton, A.G. Good, and G.J. Taylor. Induction of vacuolar ATPase and mitochondrial ATP synthase by aluminium in an aluminium-resistant cultivar of wheat. *Plant Physiology*, 125:2068–2077, 2001.

G.H. Hardy. Mendelian proportions in a mixed population. *Science*, 37:49–50, 1908.

A.E. Harvey, M.J. Larsen, and M.F. Jurgensen. Distribution of ectomycorrhizae in a mature Douglas-fir/larch soil in western montana. *Forest Science*, 22:393–633, 1976.

D.L. Hawksworth, P.M. Kirk, B.C. Sutton, and D.N. Pegler. *Dictionary of the fungi*. CAB International, Wallingford, UK, 1995.

P.W. Hedrick, M.E. Ginevan, and E.P. Exing. Genetic polymorphism in heterogeneous environment. *Annual Review of Ecology and Systematics*, 7:1–32, 1976.

R. Heuchel, F. Radtke, O. Georgiev, G. Stark, M. Aguet, and W. Schaffner. The transcription factor MTF-1 is essential for basal and heavy-metal induced metallothionein gene expression. *The EMBO Journal*, 13:2870–2875, 1994.

E. Hodgson, R.B. Mailman, and J.E. Chambers, editors. *Macmillan Dictionary of Toxicology*. Macmillan, London, UK, 1988.

K.E. Holsinger and L.E. Wallace. Bayesian approaches for the analysis of population genetic structure: an example from *Platanthera leucophaea* (Orchidaceae). *Molecular Ecology*, 13:887–894, 2004.

K.E. Holsinger, P.O. Lewis, and K.D. Dipak. A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology*, 11:1157–1164, 2002.

T.R. Horton. Molecular approaches to ectomycorrhizal diversity studies: variation in ITS at a local scale. *Plant and Soil*, 244:29–39, 2002.

R.R. Hudson, D.D. Boos, and N.L. Kaplan. A statistical test for detecting geographical subdivision. *Molecular Biology and Evolution*, 9:138–151, 1992.

C. Jacob, M. Courbot, A. Brun, H.M. Steinman, J.P. Jacquot, B. Botton, and M. Chalot. Molecular cloning, characterization and regulation by cadmium of a superoxide dismutase from the ectomycorrhizal fungus *Paxillus involutus*. *European Journal of Biochemistry*, 268:3223–32, 2001.

P. Jarne and P.J.L. Lagoda. Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, 11:424–429, 1996.

G. Jentschke and D.L. Godbold. Metal toxicity and ectomycorrhizas. *Physiologia Plantarum*, 109:107–116, 2000.

P. Joseph, Y.-X. Lei, W.-Z. Whong, and T.-M. Ong. Oncogenic potential of mouse translation elongation factor-1δ, a novel cadmium-responsive proto-oncogene. *Journal of Biological Chemistry*, 277:6131–6136, 2002.

Y. Kashi, D.G. King, and M. Soller. Simple sequence repeats as a source of quantitative genetic variation. *Trends in Genetics*, 13:74–78, 1997.

U. Krämer, I.J. Pickering, R.C. Prince, I. Raskin, and D.E. Salt. Subcellular localization and speciation of nickel in hyperaccumulator and non-accumulator *Thlaspi* species. *Plant Physiology*, 122:1343–1353, 2000.

K.V. Krutovskii, S.Y. Erofeeva, J.E. Aagaard, and S.H. Strauss. Simulation of effects of dominance on estimates of population genetic diversity and differentiation. *Journal of Heredity*, 60:499–502, 1999.

D.M. Kupfer, C.A. Reece, S.W. Clifton, B.A. Roe, and R.A. Prade. Multicellular ascomycetous fungal genomes contain more than 8000 genes. *Fungal Genetics and Biology*, 21:364–372, 1997.

S.J. Langmade, R. Ravindra, P.J. Daniels, and G.K. Andrews. The transcription factor MTF-1 mediates metal regulation of the mouse *ZnT1* gene. *Journal of Biological Chemistry*, 275:34803–34809, 2000.

A. Le Quéré, T. Johansson, and A. Tunlid. Size and complexity of the nuclear genome of the ectomycorrhizal fungus *Paxillus involutus*. *Fungal Genetics and Biology*, 36:234–241, 2002.

J. Lee, C. Godon, G. Lagniel, D. Spector, J. Garin, J. Labarre, and M.B. Toledano. Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast. *Journal of Biological Chemistry*, 274:16040–16046, 1999.

S.H. Lee, B.G. Kim, K.J. Kim, J.S. Lee, D.W. Yun, J.H. Hahn, G.H. Kim, K.H. Lee, D.S. Suh, S.T. Kwon, C.S. Lee, and Y.B. Yoo. Comparative analysis of sequences expressed during the liquid-cultured mycelia and fruit body stages of *Pleurotus ostreatus*. *Fungal Genetics and Biology*, 35:115–134, 2002.

C. Lefèbvre and P. Kokes. Variation électrophorétique des estérases des feuilles d' *Armeria maritima* (Mill.) Willd: quelques aspects taxonomiques et évolutifs. *Bulletin de la Societé Royale de Botanique du Belgique*, 177: 114–131, 1981.

C. Lefèbvre and P. Vernet. Microevolutionary processes on contaminated deposits. In A.J. Shaw, editor, *Heavy Metal Tolerance in Plants: Evolutionary Aspects*, pages 286–299. CRC Press, Boca Raton, FL, US, 1989.

G. Levinson and G.A. Gutman. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*, 4: 203–221, 1987.

Y.C. Li, A.B. Korol, T. Fahima, A. Beiles, and E. Nevo. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, 11:2453–2465, 2002.

Y.C. Li, T. Fahima, M.S. Röder, V.M. Kirzhner, A. Beiles, A.B. Korol, and E. Nevo. Genetic effects on microsatellite diversity in wild emmer wheat (*Triticum dicoccoides*) at the Yehudiyya microsite, Israel. *Heredity*, 90: 150–156, 2003.

C. Lian and T. Hogetsu. Development of microsatellite markers in black locust (*Robinia pseudoacacia*) using a dual-suppression-pcr technique. *Molecular Ecology Notes*, 2:211–213, 2002.

C. Lian, T. Hogetsu, N. Matsushita, A. Guerin-Laguette, K. Suzuki, and A. Yamada. Development of microsatellite markers from an ectomycorrhizal fungus, *Tricholoma matsutake*, by and ISSR-suppression-PCR method. *Mycorrhiza*, 13:27–31, 2003.

M. Lynch and B.G. Milligan. Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, 3:91–99, 1994.

M.R. Macnair. The tolerance of higher plants to toxic materials. In J.A.Bishop and L.M.Cook, editors, *Genetic consequences of man made change*, pages 177–208. Academic Press, London, UK, 1981.

M.R. Macnair. Heavy metal tolerance in plants: a model evolutionary system. *Trends in Ecology and Evolution*, 2:354–359, 1987.

M.R. Macnair. The genetics of metal tolerance in vascular plants. Tansley Review N° 49. *New Phytologist*, 124:541–559, 1993.

M.R. Macnair and P. Christie. Reproductive isolation as a pleiotropic effect of copper tolerance in *Mimulus guttatus*? *Heredity*, 50:295–302, 1983.

M.R. Macnair, G.H. Tilstone, and S.E. Smith. The genetics of metal tolerance and accumulation in higher plants. In N. Terry and G. Banuelos, editors, *Phytoremediation of contaminated soil and water*, pages 235–250. CRC Press LLC, 2000.

E.M. Marcotte, M. Pellegrini, T.O. Yeates, and D. Eisenberg. A census of protein repeats. *Journal of Molecular Biology*, 293:151–160, 1999.

K.A. Marrs. The function and regulation of glutathione S-transferases in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47:127–158, 1996.

J. Maynard Smith and J. Haigh. The hitch-hiking effect of a favourable gene. *Genetical Research*, 23:23–35, 1974.

T. McNeilly and J. Antonovics. Evolution in closely adjacent plant populations. IV. Barriers to gene flow. *Heredity*, 23:205–218, 1968.

E. Meglecz, F. Petenian, E. Danchin, A. Coeur d'Acier, J.Y. Rasplus, and E. Faure. High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia*. *Molecular Ecology*, 13:1693–1700, 2004.

A. Mengoni, C. Gonnelli, F. Galardi, R. Gabbrielli, and M. Bazzicalupo. Genetic diversity and heavy metal tolerance in populations of *Silene paradoxa*

L. (Caryophyllaceae): a random amplified polymorphic DNA analysis. *Molecular Ecology*, 9:1319–1324, 2000.

F.H. Meyer. Distribution of ectomycorrhizae in native and man-made forests. In G.C. Marks and T.T. Kozlowski, editors, *Ectomycorrhizae*, pages 79–105. Academic Press, New York, NY, US, 1973.

N.T. Miyashita, A. Kawabe, and H. Innan. DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism analysis. *Genetics*, 152:1723–1731, 1999.

R. Molina, H. Massicotte, and J.M. Trappe. Specificity phenomena in mycorrhizal symbiosis: community-ecological consequences and practical implications. In M.F. Allen, editor, *Mycorrhizal Functioning*, pages 357–423. Chapman & Hall, London, UK, 1992.

T.D. Murphy and G.H. Karpen. Localisation of centromere function in a *Drosophila* minichromosome. *Cell*, 82:599–609, 1995.

M. Nei. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89:583–590, 1978.

E.I. Ochiai. *General principles of biochemistry of the elements*. Plenum Press, New York, NY, US, 1987.

M.D. Ospina-Giraldo, P.D. Collopy, C.P. Romaine, and D.J. Royse. Classification of sequences expressed during the primordial and basidiome stages of the cultivated mushroom *Agaricus bisporus*. *Fungal Genetics and Biology*, 29:81–94, 2000.

B. Pawlik-Skowrońska, L. Sanità di Toppi, M.A. Favali, F. Fossati, J. Pirszel, and T. Skowroński. Lichens respond to heavy metals by phytochelatin synthesis. *New Phytologist*, 156:95–102, 2002.

S. Piry, G. Luikaert, and J.M. Cornuet. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, 90:502–503, 1999.

S. Prat. Die Erblichkeit der Resistenz gene Kupfer. *Berichte der Deutschen Botanischen Gesellschaft*, 52:65–78, 1934.

J.K. Pritchard, M. Stephans, and P. Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155:945–959, 2000.

H. Quesada, U.E.M. Ramiréz, J. Rozas, and M. Aguadé. Large-scale adaptive hitchhiking upon high recombination in *Drosophila simulans*. *Genetics*, 165: 895–900, 2003.

K. Rassmann, C. Schlötterer, and D. Tautz. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, 12:113–118, 1991.

M. Raymond and F. Rousset. GENEPOP v. 3.1c: Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86:248–249, 1995.

A.D.M. Rayner. The challenge of the individualistic mycelium. *Mycologia*, 83: 48–71, 1991.

E.M. Rees, J. Lee, and D.J. Thiele. Mobilization of intracellular copper stores by the Ctr2 vacuolar copper transporter. *Journal of Biological Chemistry*, 279:54221–54229, 2004.

W.R. Rice. Analyzing tables of statistical tests. *Evolution*, 43:223–225, 1989.

G.F. Richard and F. Pâques. Mini- and microsatellite expansions: the recombination connection. *EMBO Reports*, 11:122–126, 2000.

L.H. Rieseberg. Homology among RAPD fragments in interspecific comparisons. *Molecular Ecology*, 5:99–103, 1996.

R. Rosen, K. Büttner, D. Becher, K. Nakahigashi, T. Yura, M. Hecker, and E.Z. Ron. Heat shock proteome of *Agrobacterium tumefaciens*: evidence for new control systems. *Journal of Bacteriology*, 184:1772–1778, 2002.

V.P. Roxas, S.A. Lodhi, D.K. Garrett, J.R. Mahan, and R.D. Allen. Stress tolerance in transgenic tobacco seedlings that overexpress glutathione S-transferase/glutathione peroxidase. *Plant and Cell Physiology*, 41:1229–1234, 2000.

S. Rozen and J.H. Skaletsky. Primer3. Code available at http://www-genome.wi.mt.edu/genome_software/other.primer3.html, 1996.

J.C. Rutherford and A.J. Bird. Metal-responsive transcription factors that regulate iron, zinc and copper homeostasis in eukaryotic cells. *Eukaryotic Cell*, 3:1–13, 2004.

N.A. Schable, R.U. Fischer, and T.C. Glenn. Tetranucleotide microsatellite DNA loci from the dollar sunfish (*Lepomis marginatus*). *Molecular Ecology Notes*, 2:509–511, 2002.

H. Schat and J.A.C. Verkleij. Biological interactions: The role for non-woody plants in phytorestoration: Possibilities to exploit adaptive heavy metal tolerance. In *Metal-contaminated soils: in situ inactivation and phytorestoration*. Springer Verlag, Berlin, DE, 1998.

H. Schat, M. Llugany, and R. Bernhard. Metal-specific patterns of tolerance, uptake and transport of heavy metals in hyperaccumulating and nonhyperaccumulating metallophytes. In N. Terry and G. Banuelos, editors, *Phytoremediation of contaminated soil and water*, pages 171–188. CRC Press LLC, 2000.

C. Schlötterer. Evolutionary dynamics of microsatellite DNA. *Chromosoma*, 109:365–371, 2000.

C. Schlötterer and T. Wiehe. Microsatellites, a neutral marker to infer selective sweeps. In D.B. Goldstein and C. Schlötterer, editors, *Microsatellites: Evolution and Applications*, pages 238–247. Oxford University Press, Oxford, UK, 1999.

S. Schneider, D. Roessli, and L. Excoffier. ARLEQUIN: A software for population genetic data, 2000. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

M.D. Schug, K.A. Wetterstrand, M.S. Gaudette, R.H. Lim, C.M. Hutter, and C.F. Aquadro. The distribution and frequency of microsatellite loci in *Drosophila melanogaster*. *Molecular Ecology*, 7:57–70, 1998.

G. D. Schuler, M. S. Boguski, E. A. Stewart, L. D. Stein, G. Gyapay, K. Rice, R. E. White, P. Rodriguez-Tomé, A. Aggarwal, E. Bajorek, et al. A gene map of the human genome. *Science*, 274:540–546, 1996.

J.S. Scott and P.G. Smith. *Dictionary of Waste and Water Treatment*. Butterworths, London, UK, 1981.

P.D. Siebert, A. Chenchik, D.E. Kellogg, K.A. Lukyanov, and S.A. Lukyanov. An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Research*, 23:1087–1088, 1995.

W. Skinner, J. Keon, and J. Hargreaves. Gene information for fungal plant pathogens from expressed sequences. *Current Opinion in Microbiology*, 4: 381–386, 2001.

K. Skovgaard, L. Bodker, and S. Rosendahl. Population structure and pathogenicity of members of the *Fusarium oxysporum* complex isolated from soil and root necrosis of pea (*Pisum sativum* l.). *FEMS Microbiology and Ecology*, 42:367–374, 2002.

M. Slatkin. A measure of population subdivision based on microsatellite allele frequency. *Genetics*, 139:457–462, 1995.

S.E. Smith and D.J. Read. *Mycorrhizal Symbiosis*. Academic Press, London, UK, second edition, 1997.

R. Sokal and F.J. Rohlf. *Biometry*. Freeman, New York, NY, USA, 1995.

J. Squirrell, P.M. Hollingsworth, M. Woodhead, J. Russell, A.J. Lowe, M. Gibby, and W. Powell. How much effort is required to isolate nuclear microsatellites from plants? *Molecular Ecology*, 12:1339–1348, 2003.

J.P. Townsend, D. Cavalieri, and D.L. Hartl. Population genetic variation in genome-wide gene expression. *Molecular Biology and Evolution*, 20:955–963, 2003.

M. Vallino, V. Drogo, S. Abba', and S. Perotto. Gene expression of the ericoid mycorrhizal fungus *Oidiodendron maius* in the presence of high zinc concentrations. *Mycorrhiza*, doi:10.1007/s00572-004-0335-0, 2004.

N.A. van Hoof, V.H. Hassinen, H. Hakvoort, K.F. Ballintijn, H. Schat, J.A.C. Verkleij, W.H.O. Ernst, S.O. Kärenlampi, and A.I. Tervahauta. Enhanced copper tolerance in *Silene vulgaris* (Moench) Garcke populations from copper mines is associated with increased transcript levels of a 2b-type metallothionein gene. *Plant Physiology*, 126:1519–1526, 2001.

X. Vekemans and C. Lefèbvre. On the evolution of heavy-metal tolerant populations in *Armeria maritima*: Evidence from allozyme variation and reproductive barriers. *Journal of Evolutionary Biology*, 10:175–191, 1997.

X. Vekemans, T. Beauwens, M. Lemaire, and I. Roldán-Ruiz. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11:139–151, 2002.

P. Vos, R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23:4407–4414, 1995.

B.M. Waters and D.J. Eide. Combinatorial control of yeast *FET4* gene expression by iron, zinc and oxygen. *Journal of Biological Chemistry*, 277: 33749–33757, 2002.

W. Weinberg. Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg*, 64:368–382, 1908.

B. Weir and C. Cockerham. Estimating *F*-statistics for the analysis of population structure. *Evolution*, 38:1358–1370, 1984.

R.D. Wells. Molecular basis of genetic instability of triplet repeats. *Journal of Biological Chemistry*, 271:2875–2878, 1996.

G. Westin and W. Schaffner. A zinc-responsive factor interacts with a metal-regulated enhancer element (MRE) of the mouse metallothionein-1 gene. *The EMBO Journal*, 7:3763–3770, 1988.

S. Wright. Variability within and among natural populations. In *Evolution and the genetics of populations*, volume 4. University of Chicago Press, Chicago, IL, US, 1978.

L. Wu, A.D. Bradshaw, and D.A. Thurman. The potential for evolution of heavy metal tolerance in plants. III. The rapid evolution of copper tolerance in *Agrostis stolonifera*. *Heredity*, 34:165–178, 1975.

F. Yeh, R.C. Yang, and T. Boyle. Popgene, the user-friendly shareware for population genetic analysis, 1995. Molecular Biology and Biotechnology Center, University of Alberta, Edmonton.

H. Zhao and D.J. Eide. The yeast *ZRT1* gene encodes the zinc transporter protein of a high-affinity uptake system induced by zinc limitation. *Proceedings of the National Academy of Sciences of the United States of America*, 93:2454–2458, 1996a.

H. Zhao and D.J. Eide. The *ZRT2* gene encodes the low affinity zinc transporter in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 271: 23203–23210, 1996b.

L.A. Zhivotovsky. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, 8:907–913, 1999.

Z. Zhou, M. Miwa, and T. Hogetsu. Polymorphism of simple sequence repeats reveals gene flow within and between ectomycorrhizal *Suillus grevillei* populations. *New Phytologist*, 149:339–348, 2001.