"Investigating the Zinc Homeostasis Network in *Suillus luteus*: Zinc Transporters"

Gene discovery, transcriptome annotation and gene expression profiling

Laura Coninx (0826685)
Senior Practical Training
12/11/2012 - 26/06/2013
Biomedical Sciences - Environmental Health Sciences
Institutional supervisors: prof. dr. Jan Colpaert
prof. dr. Tom Artois
External supervisor: Emmanuelle Morin
Department: Sciences
Centre for Environmental Sciences (CMK)

Table of Contents

1. List of abbreviations

ANOVA	Analysis of variance
BLAST	Basic local alignment search tool
BLASTp	Protein BLAST
C	Cytosine
Cd	Cadmium
CDF	Cation diffusion facilitator
cDNA	Complement DNA
CMK	Centre for Environmental Sciences
Cu	Copper
DNA	Deoxyribonucleic acid
e.g.	Exempli gratia (Latin: for example)
EC50	Effect concentration 50
ECM	Ectomycorrhiza
ER	Endoplasmatic Reticulum
EST	Expressed sequence tag
Fe	Iron
G	Guanine
GI	Genbank ID
GOI	Gene of interest
ID	Identity document
INRA	Institut National de la Recherche Agronomique
JGI	Joint Genome Institute
K	Potassium
KEGG	Kyoto Encyclopedia of Genes and Genomes
kg	Kilogram
KOG	EuKaryotic Orthologous Groups
M	Average expression stability (GeNorm)
Me	Metal
mg	Milligram
Mn	Manganese
Mo	Molybdenum
N	Nitrogen
NCBI	National Centre for Biotechnology Information
NF	Normalisation factor
Nramp	Natural resistance associated macrophage protein
O	Oxygen
PAM	Point accepted mutation
PCR	Polymerase chain reaction
pH	Power of hydrogen
pm	Picometer
ppm	Parts-per-million
qRT-PCR	Quantitative reverse transcriptase PCR
RNA	Ribonucleic acid
ROS	Reactive oxygen species
S	Sulphur
TC	Transporter Classification
TCDB	Transporter Classification DataBase
TE	Tris-EDTA
TMD	Transmembrane domain
V	Pairwise variation (GeNorm)
yr	Year
ZIP	ZRT/IRT-like proteins
Zn	Zinc

2. Acknowledgements

It would not have been possible to write this master's thesis without the support and help of the kind people around me. It is only possible to mention some of those people in this section.

First of all, I want to extend my gratitude to Professor Jan Colpaert for giving me the opportunity to participate in the *Suillus luteus* project and for enabling me to go abroad. Special thanks to Professor Tom Artois for agreeing to be my second supervisor.

Furthermore, I would like to thank Michiel for not only helping me to get started with the project, but also for his patience and support throughout my thesis. He was always there for me, reassuring me, answering my thousand questions and listening to me mumbling on about my thesis when something was not going as expected. ☺ Thank you, Michiel!

For the work done in France, I would like to thank Professor Francis Martin for allowing me to do an internship at INRA. I also would like to thank Emmanuelle, for getting me settled in and for helping me with the bio-informatics part of my thesis. This must not have been easy given my non-informatical background ☺. Special thanks to Annegret, Elena, Aude, Nicolas, Thibaut, and Nicole for helping me with the manual curation, python, demonstrating lab work, interpretation of the results and administrative documents. Furthermore, I would to thank Sapna and Sanjay for their good advice, support and friendship, which has been invaluable on both academic and personal level. Moreover, I am extremely grateful for the effort made by everyone at INRA in promoting a stimulating and welcoming academic and social environment for me to work in. Laura, Stéphan H., Stéphan U., Sébastien, Yohan, Vincent, Aurélie, Alice, Patrick and Clément, thanks for all the pleasant coffee breaks, sometimes I really needed them!

When returning to Belgium, Joske was the one who took me last minute on as a student. Her experience on metal transport in *Suillus luteus* and advice on all sorts of things have been indispensable. Thank you for helping me (sometimes last minute) and being an extra promoter for me! I think I was the only student privileged with 2 internal promoters. ☺

Special thanks to all the other people in the lab that helped me with a variety of things and that made it a nice and friendly environment to work in: François, Hoai, Tony, Heidi, An, Els, Sarah and everybody else in the lab.

I would also like to thank my family and friends for their love and support throughout the entire course of my study. Thank you! Wesley, Mum, Grandma, Grandpa, Ivan, Sari, Dad, Stefanie, Zoë, Pieter and Joep!

Finally, I would like to acknowledge the financial, academic and technical support of the Hasselt University, the Centre for Environmental Sciences (CMK) and the Institut National de la Recherche Agronomique, (INRA) Nancy.

3. Abstract

INTRODUCTION A lot of research on metal toxicity and homeostasis has already been done, however only a few studies have investigated this thoroughly in mycorrhizas. This study investigates the zinc (Zn) homeostasis network of *Suillus luteus,* such a mycorrhizal fungus. As a pioneer species the fungus is an excellent species to investigate evolutionary adaptation and population dynamics. These processes cannot be easily studied in other fungi. For this project, the emphasis will lie specifically on Zn transporters, as these proteins play a key role in the Zn homeostasis. Zn tolerant isolates of *S. luteus* have already been isolated from the field.

METHODS A bio-informatics approach was used to make an inventory of all the important Zn transporter proteins in the *S. luteus* genome (already available online). Afterwards qRT-PCR was performed to investigate the expression of several of these transporter genes in tolerant and sensitive isolates under normal and excess Zn conditions.

RESULTS Seven putative Cation Diffusion Facilitator (CDF) proteins, four putative ZRT/IRT-like proteins (ZIP) and two putative Natural resistance associated macrophage (Nramp) transporters were identified in the *S. luteus* genome. These transporters were further verified by constructing phylogenetic trees that included characterized reference proteins. Eight of these genes (four CDFs and four ZIPs) were selected for further analysis of the gene expression. One gene was found to be significantly downregulated in all the isolates, when exposed to a high Zn concentration ($p<0.01$). This gene was predicted to be a homolog of the *Saccharomyces cerevisiae* ZIP (Zrt/Irt-like proteins) Zrt1. Two other genes were found to have a different expression among the isolates under normal Zn conditions. These two genes were also predicted to be members of the ZIP protein family: an ATX2 and an YKE4 *S. cerevisiae* homolog. For both genes the sensitive strains were significantly less expressed under normal conditions than the tolerant strains. When exposed to excess Zn, both the genes were upregulated in the sensitive isolates, while they remained stable in the tolerant strains.

CONCLUSIONS This project provided useful insights into the Zn homeostasis network of *S. luteus*, since an overview of the most relevant Zn transporters is now available. An indication of the function of these proteins was also obtained for nine of the twelve transporters via their alignment with characterized reference proteins. Three genes were found to be of particular interest: First of all the *S. luteus* predicted Zrt1 homolog in the context of Zn homeostasis as it was completely downregulated in all the isolates and was predicted to encode an important plasma membrane Zn importer. Secondly, the predicted *S. luteus* ATX2 and YKE4 homologs in terms of Zn tolerance, because these genes showed a difference in expression between tolerant and sensitive isolates. They were predicted to be involved in metal transport into cellular organelles, ATX2 into the Golgi vesicle and YKE4 into the endoplasmatic reticulum (ER).

4. Samenvatting (Abstract in Dutch)

INTRODUCTIE Hoewel er reeds veel onderzoek naar metaaltoxiciteit en homeostase is gedaan, zijn er maar enkele studies die dit grondig hebben onderzocht in mycorrhizas. Deze studie onderzoekt het zink (Zn) homeostase netwerk van *Suillus luteus,* zo een mycorrhiza schimmel. Deze schimmel is een ideale pionierssoort om evolutionaire adaptatie en de dynamiek van een populatie in te bestuderen. Deze processen kunnen slechts moeizaam bestudeerd worden in andere schimmelsoorten. De focus van dit project zal specifiek op Zn-transporters liggen, omdat deze proteïnen een belangrijke rol spelen in de Zn-homeostase.

Zn-tolerante isolaten van *S. luteus* zijn al geïsoleerd uit het veld.

METHODEN Bio-informatica technieken werden aangewend om een inventaris te maken van alle belangrijke Zn-transporters in het *S. luteus* genoom, (dat reeds online beschikbaar was). Een qRT-PCR werd vervolgens gedaan om de expressie van enkele van deze gevonden genen te onderzoeken in tolerante en sensitieve isolaten onder normale (20 µM Zn) en overmatige Zn (1000 µM Zn) condities.

RESULTATEN Zeven vermeende "Cation Diffusion Facilitator" (CDF) proteïnen, vier vermeende "ZRT/IRT-like" proteïnen (ZIP) en twee vermeende "Natural resistance associated macrophage" (Nramp) transporters werden geïdentificeerd in het *S. luteus* genoom. Deze transporters werden verder geanalyseerd door het construeren van fylogenetische bomen die gekarakteriseerde referentieproteïnen bevatten. Acht van deze genen (vier CDFs and vierZIPs) werden geselecteerd voor verdere analyse van de genexpressie. Er werd één gen gevonden dat volledig werd downgereguleerd in alle isolaten onder invloed van een hoge Zn-blootstelling ($p<0.01$). Dit gen was volgens de fylogenetische analyse een homoloog van het *Saccharomyces cerevisiae* ZIP proteïne Zrt1. Voor twee andere genen kon worden aangetoond dat zij een verschillende expressie hadden in sensitieve en tolerante isolaten onder normale omstandigheden. Deze twee genen waren volgens de proteïne sequentie-alignering ook homologen van de ZIP familie, meer bepaald een *S. cerevisiae* ATX2 en YKE4 homoloog. In de sensitieve isolaten hadden beide genen een significant lager expressielevel bij een blootstelling aan 20 µM Zn. Wanneer de isolaten werden blootgesteld aan 1000 µM Zn, werden de genexpressies upgereguleerd. De genexpressie van het *S. luteus* ATX2 en YKE4 homoloog bleef stabiel in de tolerante isolaten tussen de 2 Zn-condities.

CONCLUSIE Dit project leverde nuttige inzichten in het Zn-homeostase netwerk van *S. luteus* op. Er is een overzicht van de meest relevante Zn-transporters bekomen en potentiële functies van sommige van deze proteïnen zijn bepaald (voor negen van de twaalf transporters) via een alignering met gekarakteriseerde referentieproteïnen. Drie bijzonder interessante genen werden gevonden: Ten eerste het *S. luteus* voorspelde Zrt1 homoloog, in de context van Zn-homeostasis. Dit is een zeer belangrijke vondst, omdat het gen compleet werd downgereguleerd in alle isolaten bij een blootstelling aan 1000 µM Zn. Ten tweede de *S. luteus* voorspelde ATX2 en YKE4 homologen, want zij kunnen heel belangrijk zijn voor het mechanisme van Zn-tolerantie omdat deze genen een andere expressie hadden in de tolerante en sensitieve isolaten. ATX2 en YKE4 zijn beide betrokken bij metaaltransport naar de organellen, ATX2 naar het Golgi-apparaat en YKE4 naar het endoplasmatisch reticulum.

5. Introduction

Pollution is a major environmental issue faced by the world today. Among the different types of soil pollution, the contamination of soils by heavy metals is widespread and common. Heavy metal pollution poses serious risks to the general public health and is especially stressful for sessile organisms like fungi, bacteria and plants.[1] Since these organisms don't have the ability to relocate themselves, their survival depends on their ability to adapt. This makes the contaminated areas themselves a great source for finding organisms adapted to heavy metal exposure (Figure 1). In this research project, the zinc (Zn) homeostasis network of such an organism adapted to heavy metal pollution will be investigated. This organism, the ectomycorrhizal (ECM) Basidiomycete *Suillus luteus*, has been found in different contaminated areas throughout Limburg, Flanders (Belgium).[2–4]

Figure 1: A heavy metal-polluted site recolonized by pioneer vegetation[87]

5.1 Heavy metal pollution

Terrestrial environments polluted with heavy metals are often contaminated with a mixture of metals. This multiple metal contamination can originate from a number of sources: **(1)** geochemical origin, **(2)** burning of fossil fuels, **(3)** municipal wastes, **(4)** fertilizers and pesticides, **(5)** mining and processing of ores containing metal elements (e.g. melting) and **(6)** other industrial applications (e.g. batteries, sewage sludge amendments, etc.).[1,5]

The toxicity of these metal cocktails to specific soil biota can show a considerable variation depending on a number of parameters that influence the bioavailability of the metals: **(1)** soil type & characteristics (e.g. metal mobility in the soil, pH, etc.), **(2)** the different metals that are present and their relative ratio's and **(3)** the chemical form in which the metals are present.[5–8] For example, different zinc/cadmium ratios have a great effect on the level of cadmium toxicity in plants, fungi, bacteria and other organisms, because both metals compete for the same binding sites in proteins.[5,9]

Within the framework of this research project particularly the metals Zn, iron (Fe) and cadmium (Cd) are of interest, because **(1)** the Zn homeostasis network of *S. luteus* will be investigated and **(2)** Cd and Fe are chemically very similar to Zn and therefore their exclusion and regulation mechanisms can be important for the regulation of Zn as well (More detailed information on Fe and Cd is given in Annex 11.1).

5.1.1 Zinc: an essential micronutrient and transition metal

Zn is a transition metal that belongs to a group of nutrients classified as micronutrients.[10–13] According to the present knowledge, only elements of the first series transition metals (groups 5-12) and molybdenum (Mo) are essential transition metals, although there are higher order elements with similar characteristics, like Cd.[7,11,13] These essential elements, or micronutrients, are nutrients that are essential for organisms in relatively low concentrations. In the case of most

transition metals this even applies to all life forms, because these metals possess special chemical properties.[10] Zn has several special chemical properties that make it indispensable for life. First of all the Zn ion is characterized by a highly concentrated charge, because of its small atomic radius. Furthermore, Zn is a strong Lewis acid with a high affinity to ligands with sulphur- (S), nitrogen- (N), and oxygen- (O) containing functional groups.[10,11] Zn is able to interact with these ligands in a more flexible geometry than other transition metals, since it has a full d-subshell.[13,14] These properties, combined with its lack of redox activity, make Zn not only a valuable structural element in many proteins (e.g. Zn finger proteins), but also a catalyst for enzymes that have a relatively low substrate specificity and that attack small molecules (e.g. many hydrolytic enzymes). In the soil, most of the Zn is not bioavailable, as it is strongly bound by clay particles and organic matter.[8,15] Baize (1997) found the following median $[Zn]_{total}$ for different soil categories for soils in France: sandy soils 17 mg kg^{-1}, silty soils (< 20% clay) 40 mg kg^{-1}, loams (20-30% clay) 63.5 mg kg^{-1}, clayey soils (30-50% clay) 98 mg kg^{-1}, and very clayey soils (> 50% clay) 132 mg kg^{-1}.[16]

5.2 Heavy metal toxicity

Metal toxicity can be caused by both non-essential (E.g. Cd, Pb) and essential metal elements (E.g. Zn, Fe). Non-essential metals are toxic at very low concentrations, whereas essential metals will exert toxicity only when present in excess. For these essential metal ions the concentration range between adequate and excess concentrations is usually very narrow.[17]

Heavy metals can cause toxicity via primary and secondary effects. This next paragraph will discuss the most important causes of direct toxicity and associated secondary effects [10,11,18–21]:
(1) First of all, metals can cause toxicity by **directly binding to proteins**. Some transition metals have been shown to have a high affinity towards proteins with functional groups that contain S-, O- or N- ligands. When binding to these proteins, metal ions can inhibit them and also cause oxidative stress.[17,22]
(2) Secondly, also the induction of **oxidative stress** is one of the ways for metals to cause toxic effects. Upon exposure to metals, Reactive Oxygen Species (ROS) can be generated: **directly** via Haber-Weiss reactions or overproduction of ROS or **indirectly** by interacting with the antioxidant system and disrupting the electron transport chain, binding to proteins, etc.[11,17,22]
(3) Thirdly, another way for metals to exert toxicity **indirectly** is by **impairing the plasma membrane**. Metals can disturb the normal membrane functionality by stimulating lipid peroxidation and by creating an ion imbalance in the cytoplasm (e.g. inducing a potassium (K^+) efflux). This imbalance can lead to changes in the membrane potential, which in turn can implicate a number of problems (e.g. water loss).[21,23,24]
(4) And finally, metals are also known to have **genotoxic effects** and most metals are therefore classified as (weak) mutagens. Metals will exert their genotoxic effects by interfering with DNA repair processes. They do **not** damage the DNA **directly**.[20,25,26]

5.3 Metal homeostasis

Plants and fungi can depend on a wide variety of mechanisms in dealing with an excess concentration of metals in their environment.[12,22] These mechanisms can include (Figure2): **(1)**

restriction of the metal movement to the roots (e.g. by mycorrhizas), **(2)** extracellular chelation by excreted ligands, **(3)** binding the metal ions to the cell wall and or root exudates, **(4)** reduced influx into the cell, **(5)** active efflux, **(6)** transport into the vacuole or other internal compartments, **(7)** chelation in the cytosol (e.g. metallothionein (MT), glutathione (GSH)) and **(8)** upregulation of the defence mechanisms against oxidative stress (e.g. upregulation of anti-oxidant production, upregulation of repair processes, etc.).[11,12,18,22]

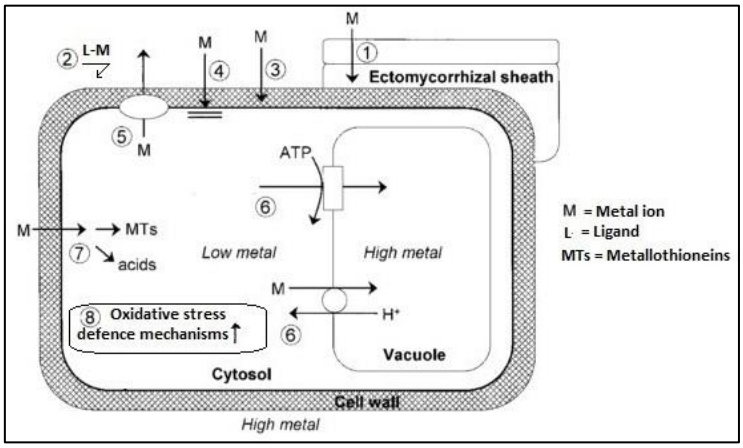


Figure 2: Overview of the general defence mechanisms against metal toxicity. The numbers are corresponding to the numbers used in the text. Modified from Hall, 2002[22]

These mechanisms are all related to pathways involved in the uptake, partitioning, sequestration and accumulation of toxic metals or to pathways involved in the oxidative stress response.[11] These mechanisms can be largely divided into two groups based on their mode of action: **(1)** Metal-specific mechanisms that respond to metals themselves (e.g. specific transporters, chelators, etc.) and **(2)** Secondary mechanisms that do not deal with the metals but with their effects (e.g. repair mechanisms, oxidative stress mechanisms).[12,18,22]

In this study on the Zn homeostasis network of *S. luteus* the focus will lie on the primary defence mechanisms and more specifically on Zn-transporters, because these transporters are very important in the response of the organism to Zn and because they play a key role in the Zn homeostasis. Since Zn tolerant ecotypes of *S. luteus* have been isolated from the field, it will be interesting to investigate whether there are differences between the regulation of these transporters among the different isolates. The next step is to place these transporters in the general picture of basal Zn **resistance** and the adaptive Zn **tolerance** that is found in some subpopulations of *S. luteus*. It is important not to confuse these 2 terms that will inevitably be used when discussing metal toxicity in a species. **Basal metal resistance** is the metal resistance that can be attributed to general defence mechanisms present in the entire population.[11] The level of resistance to a metal can vary greatly among species depending on the efficiency of the defence mechanisms that can be addressed and the level in which the organisms are capable of up- or down-regulating the involved genes and pathways. However, beyond the basal level of metal resistance, which is an inherent trait present in all individuals of a certain species, **adaptive tolerance** to one or several metals has been found in ecotypes of a number of species. [27–29] These ecotypes are subpopulations of a species that have succeeded in further adapting themselves to their environment via a genetic modification (e.g. mutation), so they form a genetically distinct variety within the species.[30] Studies on metal tolerance mechanisms suggest that these mechanisms are in most cases metal-specific, since the evidence for co-tolerance to two or more metals is not strong.[5,22,31,32]

5.4 Zinc homeostasis and transporters

As mentioned earlier, Zn is an essential element with a lot of important cellular functions, but when present in excess it becomes toxic. This explains why in general **(1)** the intracellular concentration of free Zn molecules is extremely low, **(2)** most of the Zn inside the cell is bound by ligands and/or stored inside the vacuole (because it is needed in the cell, so exporting all the Zn outside the cell is no option) and **(3)** the cell has a lot of machinery to quickly deal with Zn shortages and excesses.[33,34] This need to tightly control the concentrations and localization of Zn inside the cell has led to the evolution of a complex network of Zn-transporters, low-molecular weight ligands, metallochaperone proteins, etc. These proteins make it possible to regulate the acquisition, redistribution, storage and sequestration of Zn in response to fluctuating environmental conditions and locally varying internal demands throughout the life cycle.[33,35] The **most important way** for cells to control internal levels of Zn is by regulating **transport processes that move Zn across membranes**.[36] Organisms acquire Zn mostly from their diet or environment by specific membrane transport proteins and chelator proteins. Since these external inputs of Zn through the diet and environment can change very rapidly, the cell requires the permanent presence of these proteins acting in Zn detoxification, even in Zn deficient cells.[37] The cell is not only at risk for Zn toxicity when exposed to high external Zn concentrations, also a deficiency in other essential metal element can cause an enhanced uptake and accumulation of toxic amounts of Zn. This is because most divalent metal cation transporters often exhibit broad substrate specificity, so when a transporter is upregulated with the aim of acquiring more ions of one specific element, Zn is sometimes imported into the cell aspecifically as well. This aspecific transport can occur via several other metal transporters, since Zn does not only show a high degree of similarity with Cd and Fe, but also with other divalent cations such as Mn, Cu, Ni etc. Consequently the cell also needs specific transporters to export excess Zn from the cytoplasm. In the next sections 4.4.1, 4.4.2 and 4.4.3 three important protein families of Zn transporters will be discussed (A lot of information on these protein families is also available on http://www.tcdb.org). Although these transporter families are not the only Zn transporters in the cell, they are by far the most important in the cell.

5.4.1 The Cation Diffusion Facilitator superfamily (CDF)

The CDF transporters are an ubiquitous protein family. Members are found in Archaea, Bacteria and Eukaryota.[38–40] The family is characterized by an unusual degree of sequence divergence and size variation (300-750 amino acids). Most members have six putative transmembrane domains (TMDs), but CDFs with five and twelve TMDs have also been found.[28,39,40] CDFs transport a wide range of divalent metal cations (including Zn^{2+}, Co^{2+}, Fe^{2+}, Cd^{2+}, Ni^{2+} and Mn^{2+}).[28,33,38,39] They are localized in different membranes in the cell and normally catalyse the efflux of metals from the cytoplasm to the outside of the cell or into subcellular compartments.[34,41–43] Montanini and colleagues (2007) investigated the CDFs thoroughly and established a general signature sequence for them [40]: [SC]-X-[ASG]-[LIVM]-[LIVMTAF]-[SATG]-[DAELSTY]-[SGALFMTV]-{DKNPQR}-[HDNEL]-X_3-[DH]-X_2-[ASGLNTI]-X_{20-25}-G-X_2-[KNQRSY]-X-[DEGLNPQRST]. (With X = any amino acid, [ ] = one of the amino acid between brackets is possible and { } = any amino acid is possible except those between brackets)

5.4.2 The ZRT/IRT-like protein family (ZIP)

Another family known to transport Zn is the ZIP protein family. Like the CDF family, the ZIPs are ubiquitous: members are found in the Archaea, Bacteria and Eukaryota.[41,44,45] Most ZIP protein sequences are very divergent and vary in size between 220 and 650 amino acids. In general ZIPs have eight putative TMDs and they are known to primarily transport Zn and Fe (although other metals have also been found as substrates). ZIPs have been localized in different cellular membranes and combined with the CDF family, they account for most of the Zn transport in the cell.[46] Eng and colleagues (1998) proposed the following ZIP signature sequence : [LIVFA]-[GAS]-[LIVMD]-[LIVSCG]-[LIVFAS]-H-[SAN]-[LIVFA]-[LIFMAT]-[LIVDE]-G-[LIVF]-[SAN]-[LIVF]-[GS].[45,47]

5.4.3 The Natural resistance associated macrophage protein family (Nramp)

Besides the two major Zn-transporter families, the CDFs and the ZIPs, another protein family involved in Zn transport and homeostasis is the Nramp family. This family is also found in both Eukaryota and Prokaryota.[48,49] However, unlike the CDFs and the ZIPs, the Nramp family is a really conservative protein family with an average length around the 300 amino acids. The Nramps are known to transport the following metals: Fe^{2+}, Mn^{2+}, Zn^{2+}, Co^{2+}, Ca^{2+}, Cu^{2+} , Ni^{2+} and Pb^{2+} and they have in general eight to twelve TMDs.[50–52] These transporters have been found in both the tonoplast and plasma membranes.[53] Cellier and colleagues (1995) found the following Nramp signature sequence: D-P-G-N.[49,54]

5.5 ***Suillus luteus***

5.5.1 *Suillus luteus*: an ECM Basidiomycete

As stated previously, this study will investigate the Zn homeostasis network in *S. luteus* (Figure 3). *S. luteus* and other fungal species like *Suilllus bovinus*, *Paxillus involutus*, *Rhizopogon luteolus,* etc. are often found on contaminated areas. [3,27] However, *S. luteus* distinguishes itself from these other colonist species by being specifically involved in the primary recolonization of polluted areas. Since the contaminated areas investigated in this project are still in the stage of primary recolonization, it was opted to study *S. luteus*. Later on, this might not be possible anymore, since it can be that other species like *S. bovinus* and *R. luteolus* take over when the secondary colonisation process starts and that they outcompete *S.luteus*.

Figure 3: *S. luteus* basidiocarps; replica of the Botanical Museum in Greifswald, Germany.

S. luteus (L.Fr.) Roussel, commonly referred to as "Slippery Jack", is a Basidiomycete fungus therefore basidiocarps will be formed as sexual reproductive structures (Figure 3). These large basidiocarps or fruiting bodies are often referred to as mushrooms and are typical for most members of the Basidiomycete phylum, although there are exceptions to this feature (e.g. yeasts, rusts). *S. luteus* is a typical root symbiont that associates with the roots of young pine trees, also colonizing the polluted sites.[55,56] The fungus is regarded as a pioneer species that migrates to new areas via its spores that are easily dispersed by wind or animals.[57] A relatively short generation time (3-5 yr) and a regular sexual reproduction with billions of spores, enable a rapid selection of suitable genotypes in these new environments.[27,55] This is demonstrated in our study case as well, where an inhospitable environment polluted with Zn, Cd, Pb and other metals has contributed to the rapid evolution of tolerant ecotypes of *S. luteus.*[27] Previous studies on *S. luteus* have illustrated the importance of this fungal ECM symbiont in reducing the effects of environmental metal pollution in its host. This can be explained by the fact that the fungus is able to **(1)** sustain the host's mineral nutrition and **(2)** reduce metal transfer to the host [5,30,58].

5.5.2 *Suillus luteus:* sampling locations

S. luteus isolations have been made in previous studies at several sampling sites in the northern part of Limburg, Flanders (Belgium) [58–60]: **(1)** One part of the isolations was made at several sampling sites (Lommel-Maatheide, Lommel-Sahara and Neerpelt) in a **polluted region** in Limburg. In this area metal smelters have processed Zn-Pb ores for about one century.[5,27] Many of these Zn-Pb ores, like sphalerite and galena, also contained substantial amounts of Cd. Consequently Zn, Pb and Cd have accumulated in the soils around the factories in this area.[5,27] The bioavailable fraction of these metals was high enough to cause severe toxicity and thus selection pressure on the organisms inhabiting the area. Subsequently, during the decades, the metal contamination resulted in the establishment of very homogenous plant and microbial communities in the surroundings of these smelters. [5,27] **(2)** The second part of the isolations was conducted at several **non-polluted, control sampling sites** (Meeuwen-Gruitrode, Paal and Maasmechelen). All of these sites were located roughly 20 km south of the polluted sampling sites. (See Annex 11.2 for more information)

The phytogeographic district in which the sampling sites are situated is called the Campine region. This area is characterized by base-poor, sandy soils of low fertility with pines (*Pinus sylvestris* and *Pinus nigra)* as the dominant tree species. [27] These pine forests are normally the primary plantation and are in general younger than 30 years.[27]

5.5.3 Adaptive zinc tolerance in *Suillus luteus*

Adaptive tolerance towards heavy metals is a well-documented example of microevolution and in the case of *S. luteus* this evolution has been observed to occur very rapidly.[61,62]

In general, an adaptive tolerance to metals seems to depend on the occurrence of tolerant individuals in the general population that get the chance to colonize and monopolize new, polluted areas.[4,61,62] This colonization is associated with high selective pressures (for tolerant individuals) and apparent founder effects, but despite these 2 features, bottleneck effects are not commonly reported.[4,61,62] Explanations for this absence in bottleneck effects can be: **(1)** a high frequency of

tolerant individuals in the general (source) population, **(2)** successive colonization events (from the primary source population), **(3)** contribution to the gene pool from other surrounding populations (not the primary source population), **(4)** environmental heterogeneity, **(5)** short generation time of the species and **(6)** human disturbance.[4,61,62] In most cases the tolerant subpopulations tend to be at least as variable as the source population and other populations.[4,61,62] This level of genetic variability between the general and tolerant population has also been investigated for ***S. luteus***.[4,55] It was found that the heavy metal pollution has limited effects on the genetic structure of the tolerant *S. luteus* population.[4,55] No evidence for a reduced genetic diversity or historical bottleneck effects could be found. As mentioned earlier this absence of bottleneck effects can be explained by several reasons, but in the case of *S. luteus*, other important factors can be: **(1)** the high frequency of sexual reproduction and the short generation time, **(2)** the extensive gene flow, **(3)** the fact that multiple genotypes may have been introduced when the populations were initially established and **(4)** the fact that on the polluted site sexual reproduction may have led to a rapid evolution of the tolerance trait.[4,55]

In vitro growth experiments with *S. luteus* isolates have shown that there is a high correlation between Zn tolerance and the habitat where the samples are originating from.[3,4] Furthermore were the relative levels of tolerance of the different isolates assessed by means of EC50 (Effect Concentration, definition[1]) values. These values demonstrated that there is a very wide range in EC50 values: there are hyper-sensitive isolates (EC50 values= ± 100 ppm Zn), hyper-tolerant isolates (EC50 values= ± 1100 ppm Zn) and also a number of isolates with EC50 values in between these 2 extremes. This variability in Zn tolerance was further investigated with cross-breeding experiments (Michiel Op de Beeck, unpublished results). These experiments indicate that there is a **genetic basis,** a heredity, for the tolerance trait.

Ruytinx and colleagues (2013) investigated the Zn tolerance mechanism of a closely related *S. luteus* sister species, *S. bovinus*. Zn tolerant isolates of this species were also found on the Zn polluted areas where *S. luteus* was found (4.2.2). It has been observed that the sensitive *S. bovinus* isolates had significantly more Zn accumulation inside the cell than tolerant isolates. Zn efflux experiments further revealed a considerably faster Zn export in a Zn tolerant *S. bovinus* isolate. It was concluded that the adaptive Zn tolerance in *S. bovinus* was achieved by a Zn export out of the cell, instead of the standard transfer of Zn into the tonoplast. Furthermore, this **Zn exclusion mechanism** in Zn tolerant Suilloid ecotypes was not only able to ensure the survival of the fungus but also contributed to the protection of the fungus' host plant.[63] Because this protection of the host plant was also observed for Zn tolerant *S. luteus* isolates together with many other similar Zn exposure-linked effects between the 2 closely related sister species[30,58], *S. luteus* can possibly rely on a similar mechanism to cope with Zn stress.

[1] EC50 values were determined by assessing the concentration of Zn in the medium at which 50% of the samples weren't able to grow anymore.

5.6 Objectives

Although the Zn tolerant *S. luteus* isolates have already been investigated in previous studies [3–5,27,30,58] a lot of important questions still remain. One such question concerns the Zn homeostasis network in *S. luteus*: How is it maintained both **(1)** under normal conditions (no excess Zn in the environment) and **(2)** under conditions where the organism is challenged by higher Zn concentrations in its surroundings. Another question relates to whether Zn will activate these mechanisms related to Zn homeostasis or if these mechanisms already have a high standard expression under normal conditions. This latter can be the case if a strong promoter region is located near one of the Zn-homeostasis related genes. In this way, the cell would be able to immediately respond to Zn stress and it doesn't have to upregulate the related genes first. This study will address these questions by investigating a first selection of proteins crucial for maintaining the Zn homeostasis inside the cell: The CDF, ZIP and Nramp transporters. Although these transporters are not the only Zn transporters in the cell, they are among the most important and will provide us with a general overview of the Zn homeostasis network in *S. luteus*.

Since the genome and transcriptome of one of the *S. luteus* isolates has already been sequenced and made available online (http://www.jgi.doe.gov/) these data will be used to search for CDF-, ZIP- and Nramp-like sequences (both on genomic and transcriptomic level). Afterwards the selected *S. luteus* candidate sequences will be incorporated in phylogenetic trees with other fungal CDF, ZIP and Nramp sequences that are already characterized. *S. luteus* candidate sequences that are real CDF, ZIP and Nramp orthologs will have a good alignment with the other proteins in the three, especially with the conserved regions of these proteins. Afterwards the well-aligned *S. luteus* sequences will be manually curated on http://www.jgi.doe.gov/. This **first part of the study** will be done via a bio-informatics approach using online databases, biotools, etc. and it will be carried out at the "Institut National de la Recherche Agronomique, INRA" in Nancy, France.

The **second part of the project** will consist of investigating the expression of the selected *S. luteus* protein sequences via a quantative real-time PCR (Polymerase Chain Reaction) experiment. Sensitive and tolerant strains - cultured under normal standard conditions (20µM Zn) and under conditions with excess Zn in the medium (1000 µM Zn) - will be investigated to determine if there is a possible up- or down- regulation of these genes by Zn and if there is a difference in expression between tolerant and sensitive individuals.

After completing this second part of the study, an overview of the most important Zn transporters in *S. luteus* and hence a better understanding of the Zn homeostasis network in *S. luteus* will be available. This will enable us to better understand the Zn tolerance mechanism and to make better study designs for investigating this tolerance mechanism in future projects. Understanding this tolerance mechanism is important since a better understanding of metal tolerance mechanisms in symbiotic plant-fungus associations might lead to the development of specific metal-adapted plant-fungus combinations suited for bioremediation purposes, like phytostabilisation. Phytostabilisation of metal-polluted soils is urgently needed to reduce the dispersion of metals to other areas. Furthermore is phytostabilisation a low-cost option that also reduces the erosion of the soil and promotes the production of O_2 at the cost of CO_2.[64]

6. Materials & Methods

6.1 Gene discovery and inventory: The CDF, ZIP and Nramp family

3 different search strategies were combined in order to maximize the chance to find all relevant CDF, ZIP and Nramp homologs in the *S. luteus* UH-Slu-Lm8-n1 genome draft v1.0 (http://genome.jgi.doe.gov/Suilu1/Suilu1.home.html). Genes that were found in each of these 3 different ways of searching (5.1.1, 5.1.2 and 5.1.3) were attributed more confidence of being a real homolog.

6.1.1 BLASTp against *Suillus luteus* genome/transcriptome with reference proteins

Fungal CDF, ZIP and Nramp sequences were obtained from TCDB (http://www.tcdb.org/), TransportDB (http://www.membranetransport.org/) and several articles.[40,43] **TCDB** and **TransportDB** are both freely accessible, interactive databases that contain sequence, structural, classification, evolutionary and functional information about transport systems from a variety of organisms. Both databases use the transporter classification (TC) system, an International Union of Biochemistry and Molecular Biology approved system of nomenclature for the classification of transport proteins. TCDB however is a curated database, whereas TransportDB is not.[65,66]
Consequently, the CDF, ZIP and Nramp reference sequences selected from these databases are already identified and characterized proteins that were used to look for orthologous sequences in the filtered model dataset of *S. luteus* UH-Slu-Lm8-n1 v1.0 on the JGI website (Joint Genome Insitute).This search for orthologous sequences was done via a protein BLAST (BLASTp), run with the standard program settings (a maximum E-value of 1E-5 and the BLOSUM62 scoring matrix).
The **JGI** website that was used to access the filtered model dataset of *S. luteus* contains a section 'Fungal genetics program' where a number of fungal genomes and transcriptome data are made available. JGI makes use of a number of statistical prediction models that automatically combine genome and transcriptome data. However, since not all of the matches between the transcriptome and genome data are found or made correctly by the prediction models: **(1)** some parts of the *S. luteus* genome have no corresponding transcriptome data (ESTs, Expressed Sequence Tags) and **(2)** some ESTs are not mapped to the genome. This explains why it was opted to search both the genome (5.1.1 and 5.1.3) and the transcriptome (5.1.2). Furthermore, this also implies that, running a BLAST on JGI (as was done in this section) will involve 2 types of gene models: both **(1)** the models that have ESTs mapped to them (more reliable) and **(2)** the remaining gene predicted models without any EST evidence (less reliable).

6.1.2 Keyword search in the Expressed Sequence Tag (EST) dataset

A second search for CDF, ZIP and Nramp orthologs was executed via a keyword search directly in the annotated EST assembly of *S. luteus* UH-Slu-Lm8-n1 v1.0. This dataset was accessible at the INRA, Mycorweb website (http://mycor.nancy.inra.fr/) together with the **(1)** Swissprot, **(2)** Pfam, **(3)** KOG (EuKaryotic Orthologous Groups) and **(4)** KEGG (Kyoto Encyclopedia of Genes and Genomes) ESTs best hit descriptions.

These 4 databases are widely used bioinformatics data resources.

(1) Swissprot is a curated protein sequence database with a high level of annotation, extensive cross-references, literature citations and computational analyses.

(2) PFAM is a database that encompasses a large collection of multiple-sequence alignments and hidden-Markov models of common protein domains and families.

(3) The **KEGG** database consists of pathway maps that also include diagrams representing the information pathways of interacting molecules or genes.

(4) KOG is a eukaryote-specific database developed for the identification of ortholog and paralog proteins.

Since Swissprot and Pfam are both curated databases, whereas KEGG and KOG are not, information gained from Swissprot and Pfam was regarded with a higher level of confidence. This information together with its level of confidence was also taken into account at the manual curation that happened later on.

To find relevant ESTs the search option 'Hit description contains' was selected, so that every EST with a Swissprot, Pfam, KEGG or KOG hit description containing the entered keyword was found.
The following keywords were used to look for relevant ESTs: "CDF", "cation diffusion facilitator", "cadmium", "ZIP", "zinc transport", "ZRT", "IRT", "zinc iron permease", "Nramp", "iron transport", "natural resistance-associated macrophage protein", "zinc transport", "zinc efflux", "zinc influx", "zinc permease", "iron efflux", "iron influx" and "iron permease".

The hits (ESTs or EST-contig sequences) were blasted against the *S. luteus* UH-Slu-Lm8-n1 v1.0 proteome and additional genes were included in the selection of interesting *S. luteus* genes to be investigated further. Since these additional ESTs were not found in 5.1.1, chances of their predicted gene model being inaccurate were high, therefore the predicted gene models of all these additional genes were immediately checked.

6.1.3 BLASTp in *S. luteus* UH-Slu-Lm8-n1 v1.0 with curated *Laccaria bicolor* transporters

Since many of the fungal reference CDF, ZIP and Nramp proteins (5.1.1) were phylogenetically still relatively distant from *S. luteus*, manually curated *Laccaria bicolor* proteins (Table 1) were used to look for additional orthologous sequences in the filtered model dataset of *S. luteus* UH-Slu-Lm8-n1 v1.0 (http://genome.jgi.doe.gov/Suilu1/Suilu1.home.html). This was also done via a BLASTp, run with the standard program settings.

Table 1: *Laccaria bicolor* orthologues sequences		
Protein family	**JGI protein ID number**	**Manually curated by**
CDF	305317	Dr. Damien Blaudez
CDF	307944	Dr. Damien Blaudez
CDF	625478	Dr. Damien Blaudez
CDF	191080	Dr. Damien Blaudez
CDF	234505	Dr. Damien Blaudez
CDF	681521	Prof. Dr.Ursula Kües
CDF	256277	Dr. Damien Blaudez
CDF	244654	Dr. Damien Blaudez
ZIP	189929	Dr. Damien Blaudez
ZIP	309134	Dr. Damien Blaudez
ZIP	309863	Dr. Damien Blaudez
ZIP	180140	Dr. Damien Blaudez
ZIP	305445	Dr. Damien Blaudez
Nramp	186854	Dr. Montanini Barbara

6.2 Phylogenetic trees

6.2.1 Homologous sequences

With the *S. luteus* UH-Slu-Lm8-n1 v1.0 candidate genes identified in 5.1 a BLASTp search was done in the genomes/transcriptomes of other Fungi to obtain the necessary reference sequences needed to make phylogenetic trees. The most recent version of the following genomes was blasted with the *S. luteus* hits (http://www.jgi.doe.gov/): *Suillus brevipes*, *Heterobasidion irregulare, Schizophyllum commune, Pleurotus ostreatus, Hebeloma cylindrosporum, Laccaria bicolor, Serpula lacrymans* and *Candida tenuis.* These species were selected because they were: **(1)** closely related to *S. luteus*, **(2)** had a manually annotated genome or **(3)** had a resequenced genome (v.2.0 and hence more reliable reference sequences).

Furthermore, if there were well-characterized CDF, ZIP and Nramp sequences available for other Basidiomycetes or Ascomycetes these sequences were also included:

(1) For the CDF family the following proteins were also included:

a. *H. cylindrosporum* ZnT1 (Genbank Accession number (GI) AFB74685)
b. *Paxillus involutus* CDF Manganese transporter (GI AFB74686)
c. *Saccharomyces cerevisiae* COT1 (GI CAA99636), ZRC1 (GI CAA88653), MSC2 (GI DAA12046), ZRG17 (GI DAA10581), MMT1 (GI DAA10074) and MMT2 (GI DAA11212)

(2) For the ZIP family the following proteins were also included:

a. *S. cerevisiae* ZRT1 (JGI 2308), ZRT2 (JGI 4151), ZRT3 (JGI 3650), YKE4 (JGI 1756) and ATX2 (JGI 5732)

(3) For the Nramp family the following proteins were also included:

a. *S. cerevisiae* SMF1 (GI NP_014519), SMF2 (GI EIW10202) and SMF3 (GI EIW09113)

Because the Nramp family is very conserved also the following non-fungal reference sequences were included:

b. Homo sapiens NRAMP1 (GI NP_000569) and NRAMP2 (GI BAA24933)
c. Mus musculus NRAMP1 (GI NP_038640) and NRAMP2 (GI AAC24496)

6.2.2 Control for false positives

A BLASTp can have a number of false positive hits. Therefore certain characteristics of the *S. luteus* sequences and selected homologs (5.2.1) were checked in order to exclude these false positives:

(1) Does the sequence have a conserved domain hit on JGI or on the website of the National Centre for Biotechnology Information (NCBI)?
(2) Does the sequence have a sequence length that is within the normal average range of the protein family?
(3) Does the sequence have the normal number of predicted transmembrane domains (TMDs)? (the predicted number of TMDs is available on the website of JGI)
(4) Does the sequence possess one of the signature sequences that are described for the protein families? [40,46,54,67]
(5) If a sequence didn't fit the criteria above, its automatically selected predicted gene model was examined. Was this a good model or was there editing needed? A better model was selected and it was checked if the criteria were better met.

The first 4 characteristics were used to calculate a criteria score. A positive answer to one of the questions gave a +1 in that category. The final score was calculated by the following formula:

$$\begin{aligned}\boldsymbol{Final\ criteria\ score} = \quad & 2 * conserved\ domain\ (present = 1, absent = 0) \\ & + 1 * correct\ length\ (yes = 1, no = 0) \\ & + 2 * correct\ number\ of\ TMDs\ (yes = 1, no = 0) \\ & + 2 * signature\ sequence\ (present = 1, absent = 0)\end{aligned}$$

If the criteria score was smaller than 3, the sequence was excluded. Sequences of which the predicted gene model needed editing were adjusted or discarded if no good alternative gene prediction model could be found.

6.2.3 Alignment algorithms

Since not every protein family is conserved in the same regions, in the same number of amino acids, in the length of the variable regions, etc., it was opted to test 6 different, often used sequence alignment algorithms: **(1)** ClustalW2, **(2)** Clustal Omega, **(3)** Kalign, **(4)** T-Coffee, **(5)** MAFF-T and **(6)** MUSCLE.

For each algorithm it was checked if the conserved domains and/or signature sequences of the proteins were correctly aligned. The best algorithm model was selected by looking at: **(1)** in the alignment, were the conserved domains well-aligned and **(2)** in the cladogram, was there a subdivision into clusters that was consistent with literature.

6.2.4 MEGA 4.0.2

After obtaining the *S. luteus* homologs (5.1), searching for reference homologous sequences (5.2.1), removing false positives (5.2.2) and selecting the best alignment algorithm (5.2.3), phylogenetic trees were constructed using the MEGA 4.0.2 software. In order to take amino acid substitutions into account the Dayhoff model[2] was selected to predict substitutions and substitution rates. For each phylogenetic tree the neighbour-joining method was used and a bootstrap test of phylogeny with 1000 replicates was done.

Because the ZIP family is really divergent and consequently did not have a good alignment, a phylogenetic tree was made with only reference proteins containing eight predicted TMDs or manually curated ones.

6.3 Manual curation

Based on the data from **(1)** the phylogenetic trees (5.2.4) and **(2)** the control criteria for the removal of false positive hits (5.2.2), the selected *S. luteus* sequences could be more confidently annotated on JGI using the manual annotation tool.
The proteins were named: **(1)** a putative CDF/ZIP/Nramp transporter, when the results strongly indicated the sequence being a CDF/ZIP/Nramp transporter or **(2)** a hypothetical CDF/ZIP/Nramp transporter, when the results merely suggested the sequence being a CDF/ZIP/Nramp transporter.

6.4 Validation: qRT-PCR (quantitative Reverse Transcriptase Polymerase Chain Reaction)

S. luteus sequences predicted to be CDF and ZIP Zn transporters by comparative sequence analysis and phylogenetic analysis were selected for further investigation. The 2 *S. luteus* CDF ZRC1 homologs that were found were excluded, because they had already been intensively studied in previous research (unpublished, Ruytinx). A total of 8 *S. luteus* hits, 4 CDFs and 4 ZIPs, were finally selected for further analysis with qRT-PCR.

6.4.1 Primer design

For the amplification of the reference genes and the genes of interest (GOI), primers were developed with the use of Primer3 software version 3.0.0.[68] Preferably, primers with the following characteristics were chosen: **(1)** an average length of 20 nucleotides, **(2)** an annealing temperature of 60°C, **(3)** an average GC (Guanine, Cytosine) content of 50%, **(4)** a product size between 80 and 160 nucleotides, **(5)** a maximum self-complementarity score of 3, **(6)** a maximum 3'-complementarity score of 0 and **(7)** spanning an exon-exon border.
The nucleotide sequences and other efficiencies of the primers are given in Table 2.
Undissolved primers were ordered at Biolegio (Nijmegen, the Netherlands) and diluted with 1/10 TE buffer (Tris – EDTA buffer) to bring the final primer concentrations to 100 µM. These 100 µM stock primers were diluted with RNase free water to a concentration of 10 µM, so that the primers could be added directly into the mastermix.

[2] Also called the PAM (Point Accepted Mutation) substitution matrix.

Primer efficiencies (Table 2) were calculated by making a standard curve for each gene by plotting the average values of a dilution series. Dilution series ware made with a pooled cDNA sample and 3 technical replicates per dilution (x1, x1/4, x1/16, x1/64, x1/256).

The efficiency was calculated with the following formula:

$$Efficiency = \left(10^{\frac{1}{slope\ of\ the\ standard\ curve}}\right) - 1$$

Table 2: Primer sequences and characteristics

Gene[A]	Primer sequences (F = forward, R = reverse)[B]	Efficiency (%)
Gene 1 - ZIP (229544)	F: 3'-CCACATGCTCGACCTCTGTT-5' R: 5'-CATACCCTCCTCGCGTTCAA-3'	85.58
Gene 2 - ZIP (720881)	F: 3'-TTGAGGCTTTGAGGAACCCA-5' R: 5'-TCGTTGCAGTTATCGGGGTA-3'	82.75
Gene 3 - CDF (810602)	F: 3'- CGGTGACTTTGTTGTTCTGTT-5' R: 5'-TACGAATGGAAGCCGATG-3'	100.64
Gene 4 - ZIP (811220)	F: 3'-GCGTGGTTGTTCCCATTT -5' R: 5'-ATCGTATTGAGCGTGTCTGG-3'	100.09
Gene 5 - ZIP (22926)	F: 3'-GCCAAACGGACAAACTGG-5' R: 5'- GACAGGCACGGAGATGAAAG-3'	96.07
Gene 6 - CDF (72605)	F: 3'-CGTCTTCGCACTCTCTCATCT-5' R: 5'-TCATCCTTGCCTTCTTCACC-3'	97.32
Gene 7 - CDF (72657)	F: 3'-ACCATCCCCACGAACATGAC-5' R: 5'-ATGGGAAAGCGAAGACTGCA-3'	85.73
Gene 8 - CDF (798077)	F: 3'-CCCCGAATAACAATCCACAA-5' R: 5'-GAGGAGCCGAATGAAAAAGA-3'	94.63

[A] (JGI protein number)

[B] A =adenine, C = cytosine, G = guanine and T = thymine

6.4.2 Reference genes

Stability was tested and approved for 3 reference genes using the GeNorm software version 3.5. Reference genes were selected from a set of reference genes used in previous experiments (unpublished, Ruytinx). The following genes were tested: Zn103 (GI AM085296), Zn109 (GI AM085168), TUB1 (GI AY112730) and TDF1 (GI GR975621).

6.4.3 cDNA

Residual cDNA of previous experiments (unpublished, Ruytinx) was used to analyse the gene expression of the eight selected transporter genes and the 3 reference genes. The cDNA was collected from fungal mycelia 48h exposed to 20µM or 1000µM Zn in modified liquid Fries medium.[27,63] cDNA of 4 different isolates was used with five biological replicates per isolate. These isolates were: P4 and Mg4, two Zn sensitive isolates and Ls1 and OF3, two Zn tolerant isolates.

6.4.4 qRT-PCR

Master mixes were prepared with the following reagentia: Each PCR reaction contained Fast 2x SYBR Green dye (Applied Biosystems, CA, USA), 0.3 µM forward primer, 0.3 µM reverse primer and 2.5 µl cDNA. RNase free water was used to bring the final reaction volume to 10 µl.

Samples with their appropriate mastermix were loaded onto 96 well plates and placed in the 7500 Fast Real-Time PCR System (Applied Biosystems, Foster city, USA). After completing fast cycling conditions program (20 s at 95°C, 40 cycles of 3 s at 95°C and 30 s at 60°C), dissociation and

amplification curves were analysed with the 7500 Fast Real-Time PCR Software version 2.0.6 (Applied Biosystems, Foster city, USA). Ct-values were exported and further analysed with the comparative ΔCt method.[69]

For each gene also a "No template control" (NTC) was performed by adding 2.5 µl RNase free water instead of cDNA to the mastermix.

6.4.5 Statistical analysis

The statistical analysis was performed using "R" software (Vienna, Austria) version 2.15.1 The following tests were used:

To test if one gene was significantly up- or down- regulated within a specific isolate:

(1) The data were tested for a normal distribution using the **Shapiro-Wilk test.**

(2) The data were tested for homoscedasticity:

a. A parametrical test was done if the data were normally distributed: the **Bartlett Test** of Homogeneity of Variances.
b. A non-parametrical test was done if the data were not normally distributed: the **Fligner-Killeen Test** of Homogeneity of Variances.

(3) To test the hypothesis if the expression of a gene was not different between the two conditions in one isolate:

a. A parametric **1-way analysis of variance** (ANOVA) was done, if all the data were normally distributed and homoscedastic.
b. Else, a non-parametrical **Kruskal-Wallis test** was done.

To test if there were differences between the different isolates (e.g. in gene expression under the 2 Zn conditions, in manner of reacting, etc.):

(1) Step 1 of the previous paragraph was repeated for the combined data of all the isolates.

(2) Step 2 of the previous paragraph was repeated for the combined data of all the isolates.

(3) To test the hypothesis that there were no differences between the different isolates:

a. Parametrical testing was done if the data were normally distributed and homoscedastic: A **2-way ANOVA** followed by a **Tukey's HSD (Honestly Significant Difference) test**.
b. Else non-parametrical testing was done: A 2-way **Kruskal-Wallis test** followed by a **Pairwise Wilcoxon Rank-Sum test**.

7. Results

7.1 Phylogenetic trees

In order to make the phylogenetic trees, a lot of different Fungi were used including *S. brevipes* a sister species of *S. luteus*. **For each of the selected *S. luteus* CDF, ZIP and Nramp candidate sequences a *S. brevipes* homolog was found**. These *S. brevipes* homologous sequences were also included in the trees, except for one homolog that was excluded. This was the CDF ZRG17 homolog, as it wasn't able to make the defined selection criteria and consequently also didn't have a good alignment with all the other selected CDF sequences.

7.1.1 CDF

The search for genes encoding CDF proteins resulted in a large set of sequences that cluster in a number of distinct groups (Figure 4 and 5). Four large subdivisions were observed in the general structure of the CDF cladogram. To avoid confusion these primary clusters will be hereafter referred to as "Subfamilies" and the following secondary subdivisions will be referred to as "Clusters". All of the **seven *S. luteus* CDF homologs** that were found are given in the cladograms (Figure 4 and 5).

The subfamilies and clusters that were found are consistent with the ones seen in literature. Three major groups with a different primary substrate (Zn, Fe/Zn and Mn) and one rest group with unknown specificity were found. [40,70] Subfamily 1 and 4 (Figure 4) are probably transporting Zn, given their alignment with known, characterized Zn transporters. These two subfamilies make up the first part of the phylogenetic tree: the Zn transporters. **Subfamily 1** subdivides into two clusters: the **ZRC1/COT1 cluster,** that has two *S. luteus* homologs and the **MSC2 cluster** that only has one *S. luteus* homolog**.** Both clusters are named after their characterized *S. cerevisiae* protein members. For **Subfamily 4** no underlying clusters are discussed, because the only bifurcation that could be detected was the segregation between the Ascomycetes and Basidiomycetes, which is normal and does not necessarily imply a functional difference between the two groups. This subfamily was also named after its *S. cerevisiae* protein member, the ZRG17 Zn transporter. It includes one *S. luteus* protein sequence.

A second part of the cladogram (Figure 5) consists of two other subfamilies and one undefined rest group. All these groups have one *S. luteus* homolog. Subfamily 2 and 3 are probably transporting Mn and Fe/Zn respectively, based on their alignment with known Mn and Fe transporters. **Subfamily 2** has no underlying clusters and is simply named CDF Mn transporters, like the characterized *P. involutus* CDF Mn transporter. For this family no characterized *S. cerevisiae* homologs could be found. **Subfamily 3** has only a bifurcation between the Ascomycetes and Basidiomycetes and therefore no underlying clusters are defined. This group has two *S. cerevisiae* homologs MMT1 and MMT2 and is consequently named the MMT1/MMT2-like subfamily. The last group that remains is the undefined **rest group** with unknown specificity. This group branches off right before Subfamily 2, the Mn transporters and consists of only three members: the *S. luteus*, *S. brevipes* and *H. cylindrosporum* CDF homologs.

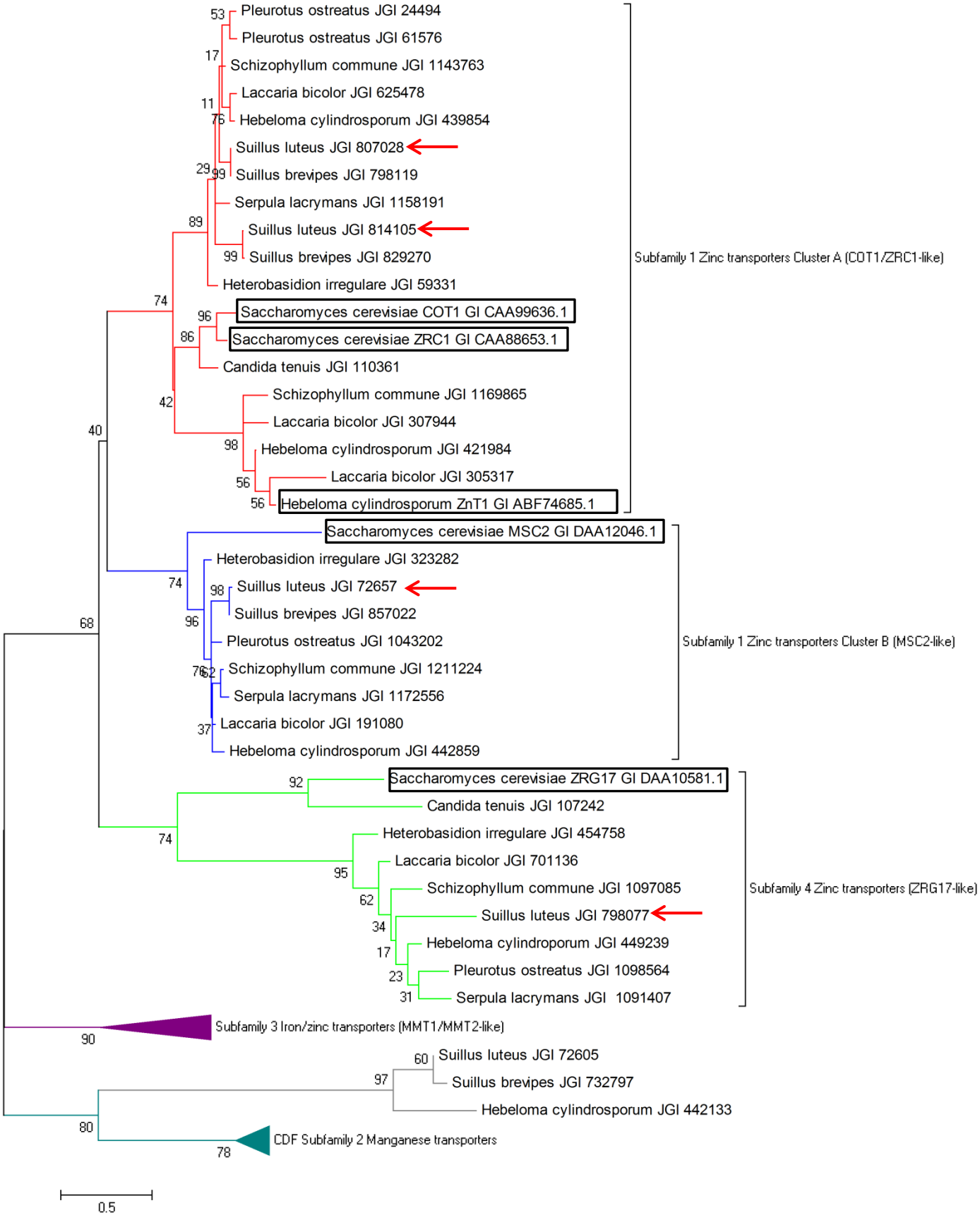


Figure 4: The CDF family (part A) Zn transporters. The JGI or Genbank (GI) identification number of each protein is given. *S. luteus* homologs are designated with an arrow and well-characterized reference proteins are framed. This bootstrap consensus tree was generated using the MEGA 4.0.2 software combined with the MAFFT alingment algorithm v6.850. Bootstrap values (1000 replicates) are also given in the figure.

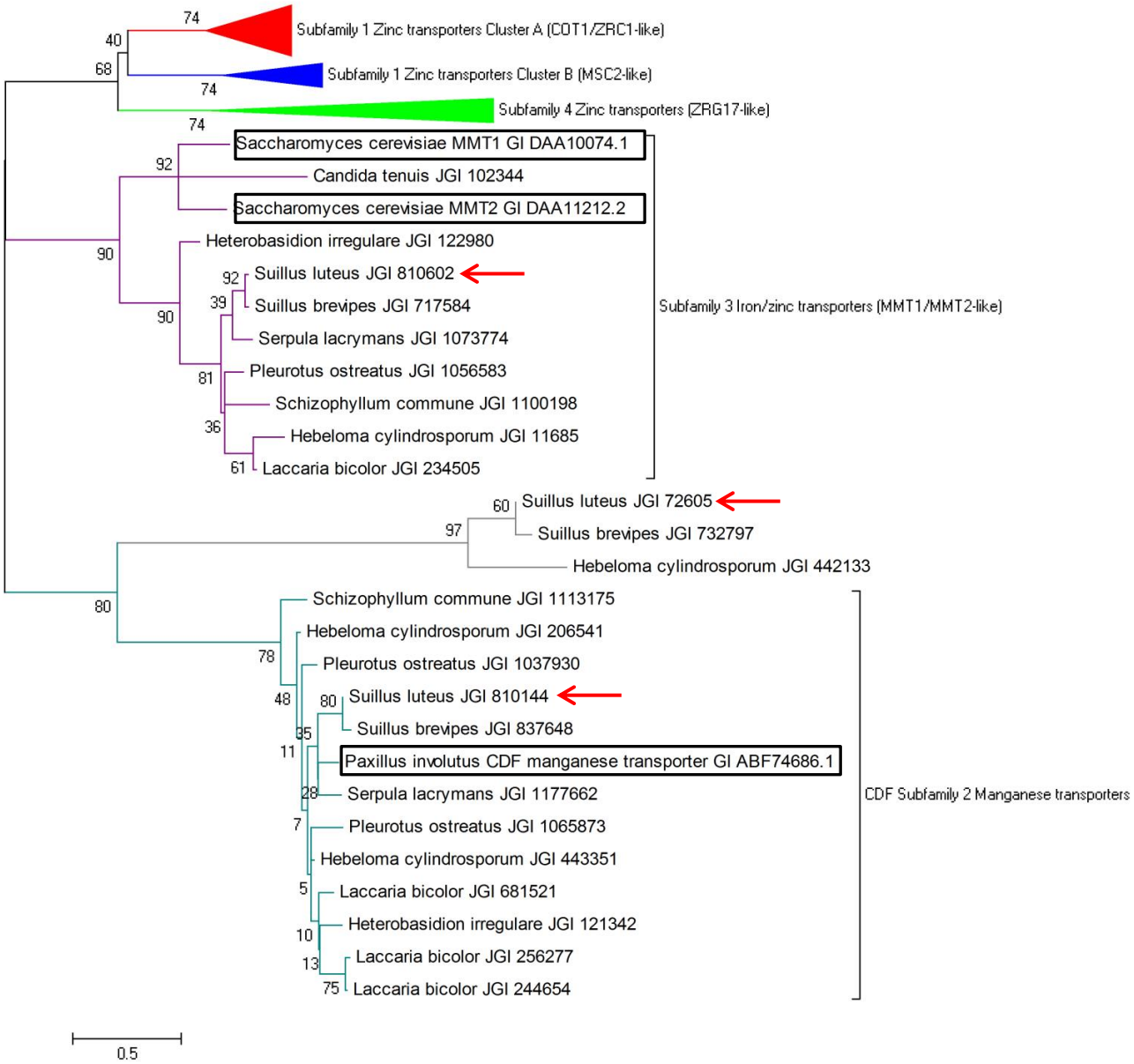


Figure 5: The CDF family (part B) Fe and Mn transporters. The JGI or Genbank (GI) identification number of each protein is given. *S. luteus* homologs are designated with an arrow and well-characterized reference proteins are framed. This bootstrap consensus tree was generated using the MEGA 4.0.2 software combined with the MAFFT alingment algorithm v6.850. Bootstrap values (1000 replicates) are also given in the figure.

7.1.2 ZIP

The general structure of the ZIP phylogenetic tree has three large groups (Figure 6). For one of these groups, **Subfamily 3 the ZRT3-like proteins**, only two protein sequences were found: *S. cerevisiae* ZRT3 and its *C. tenuis* homolog. This is consistent with literature as no homologs have ever been found in the Basidiomycetes.

The other two subfamilies that remain both subdivide into two clusters and each of these clusters contains one of the **4 *S. luteus* ZIP homologs**. **Subfamily 1** has two clusters. One of them is the **ZRT1/ZRT2-like cluster**, named after its characterized *S. cerevisiae* protein members ZRT1 and ZRT2. These are both Zn transporters.[71] The other cluster has no characterized proteins aligned to it and is therefore simply called **cluster B**. No information regarding classification or function of this cluster could be found. Finally the last group, **Subfamily 2**, consist of the **YKE4-like-cluster** and the **ATX2-like cluster**. Both clusters are named after their *S. cerevisiae* members: YKE4, a bidirectional Zn transporter and ATX2, a Mn transporter.[72,73]

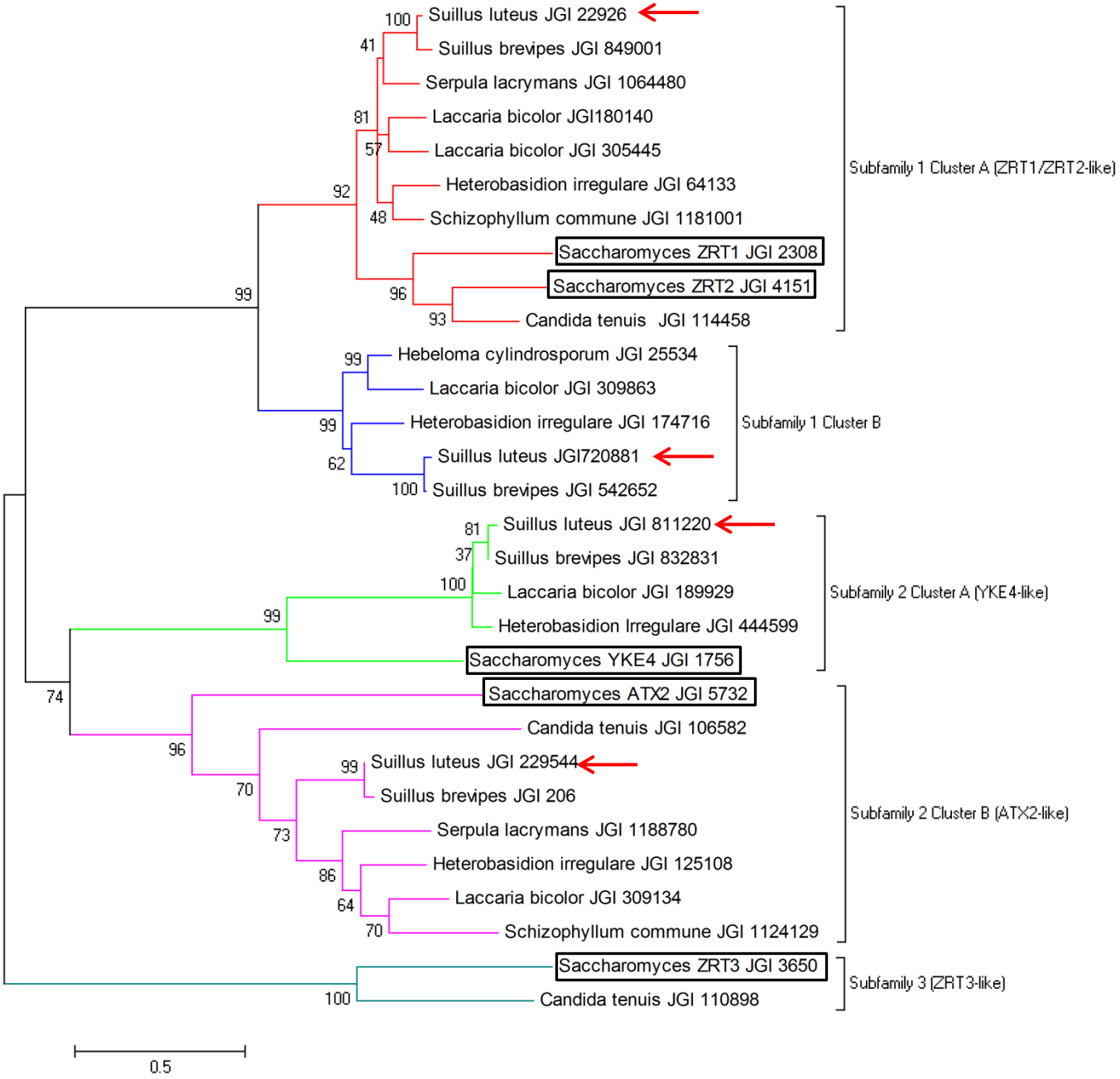


Figure 6: The ZIP family. The JGI or GI ID numbers of each protein is given. *S. luteus* homologs are designated with an arrow and well-characterized reference proteins are framed. This tree was generated using the MEGA 4.0.2 software combined with the MUSCLE alingment algorithm v0.1 . Bootstrap values (1000 replicates) are also given in the figure.

7.1.3 Nramp

As mentioned in the introduction, the Nramp family is a very conservative family. Therefore also homologs from more distinct species were included in the tree (Nramp 1 and Nramp2 from *H. sapiens* en *M. musculus*) (Figure 7). These homologs formed a separate group in the cladogram and were named "**Group 3 Mammalia**".

To discuss the Nramp family the term "Group" is used instead of "Subfamily", because the cladogram clearly consists of clusters that coincide with the different taxonomical groups of the reference species. Subsequently also the remaining groups, **Group 1** and **Group 2**, were named after their corresponding phylum or class: the Basidiomycota and the Ascomycota, respectively. Group 1 subdivides into two clusters: **Cluster A** and **Cluster B.** They each contain one of the **two *S. luteus* homologs** that were found. No information regarding the classification or function of these proteins of both clusters could be found.

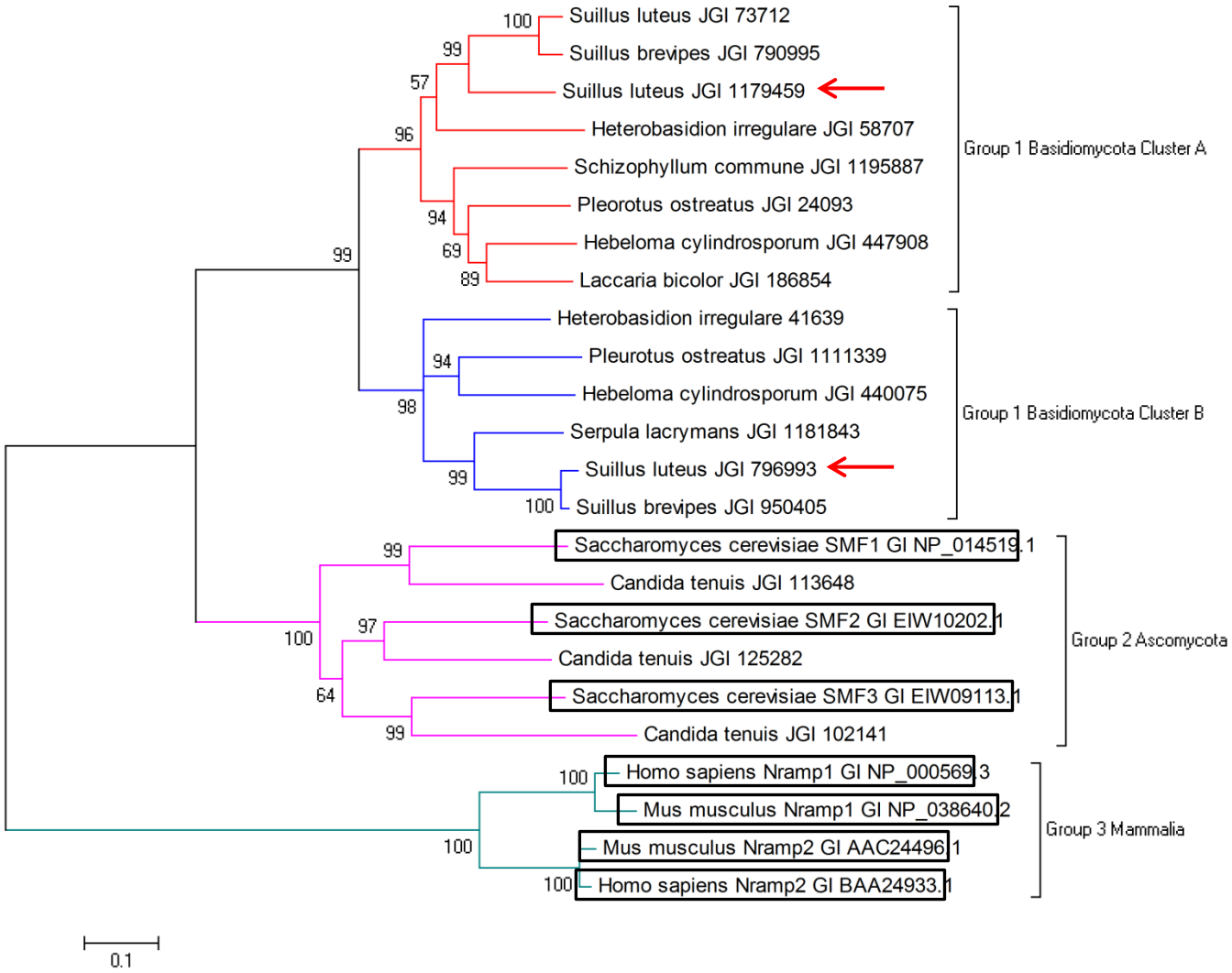


Figure 7: The Nramp family. The JGI or Genbank (GI) identification number of each protein is given. *S. luteus* homologs are designated with an arrow and well-characterized reference proteins are framed. This bootstrap consensus tree was generated using the MEGA 4.0.2 software combined with the MAFFT alingment algorithm v6.850 . Bootstrap values (1000 replicates) are also given in the figure.

7.2 Manual curation

All of the *S. luteus* CDF, ZIP and Nramp homologs that were found were annotated on JGI.

The annotation was based on **(1)** the proteins' criteria scores that reflect how well the proteins correspond to the general characteristics of their protein families, **(2)** their alignment in the phylogenetic trees and **(3)** the occurrence of close homologs (preferably characterized).

A protein with a very high probability of being a member of a certain protein family X, was named "a **putative** X transporter". In some cases, when a potential substrate could be appointed to a transporter, this was also mentioned in the protein name. A protein that had less convincing results of being a particular transporter was named "a **hypothetical** X transporter". All of the annotated proteins are listed in Table 3.

Table 3: *S. luteus* CDF, ZIP and Nramp homologs.				
Protein family	**Protein ID (JGI)**	**Name**	**Description**[A]	**Defline**
CDF	807028	Putative CDF (Zn-) transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Putative CDF (Zn-) transporter, ZRC1-like
CDF	814105	Putative CDF (Zn-) transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Putative CDF (Zn-) transporter, ZRC1-like
CDF	72657	Putative CDF (Zn-) transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Putative CDF (Zn-) transporter, MSC2-like
CDF	798077	Putative CDF (Zn-) transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Putative CDF (Zn-) transporter, ZRG17-like
CDF	810602	Putative CDF (Fe-) transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Putative CDF (Fe-) transporter, Mmt1p-like
CDF	72605	Hypothetical CDF transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Hypothetical CDF transporter with unknown specificity, no similarity with other known CDFs, high level of uncertainty, but homologs in other Basidiomycete fungi have been found
CDF	810144	Putative CDF (Mn-) transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.4) The Cation Diffusion Facilitator (CDF) family	Putative CDF (Mn-) transporter
ZIP	22926	Putative ZIP transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.5) The ZIP (ZRT-IRT-like protein) family	Putative transporter of the ZIP family
ZIP	720881	Putative ZIP transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.5) The ZIP (ZRT-IRT-like protein) family	Putative transporter of the ZIP family
ZIP	811220	Putative ZIP transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.5) The ZIP (ZRT-IRT-like protein) family	Putative transporter of the ZIP family
ZIP	229544	Putative ZIP transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.5) The ZIP (ZRT-IRT-like protein) family	Putative transporter of the ZIP family
Nramp	73712	Putative Nramp transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.55) The Metal Ion (Mn/Fe) Transporter (Nramp) Family	Putative transporter of the Nramp family
Nramp	796993	Putative Nramp transporter *Suillus luteus* UH-Slu-Lm8-n1	(TCDB 2.A.55) The Metal Ion (Mn/Fe) Transporter (Nramp) Family	Putative transporter of the Nramp family

A The TCDB classification number and official protein family name are given

7.3 qRT-PCR

7.3.1 Reference genes

The optimal number of reference genes was determined using the GeNorm version 3.5 software. This algorithm ranks candidate reference genes according to their average expression stability value M. The M-values of the 4 candidate reference genes that were selected for testing (based on previous experiments, Ruytinckx 2013, unpublished) are given in Figure 8. All of them have an M-value below the cut-off value of M<1.5 proposed by Vandesompele et al. 2002. [74] **The most stable genes were TDF1 and Zn109, followed by Zn103 and then TUB1** (Figure 8).

Based on the M-values of these genes a normalization factor (NF) was calculated for the 2 (TDF1 and Zn109), 3 (TDF1, Zn109 and Zn103) and 4 (TDF1, Zn109, Zn103 and TUB1) most stable genes. The optimal number of reference genes was then determined by calculating the pairwise variation (V) between the two sequential NFs of **(1)** the 2 and 3 most stable reference genes (V2/3) and **(2)** the 3 and 4 most stable reference genes (V3/4). These results are given in Figure 9. Since a large V means that the added gene has a noteworthy effect on the normalisation, it needs to be included in order to have a reliable normalization factor. Vandesompele et al. 2002 also proposed a threshold value for V<0.15. However, when working with *S. luteus* it is very difficult to meet this proposed threshold value, as the species is characterized by an unusual degree of genetic variability.[4,59] Therefore, it was opted to accept a higher V value and work with the **3 most stable reference genes: Zn109, TDF1 and Zn103**. Figure 9 shows that the inclusion of an additional 4th reference gene will only increase the pairwise variation and is consequently not beneficial.

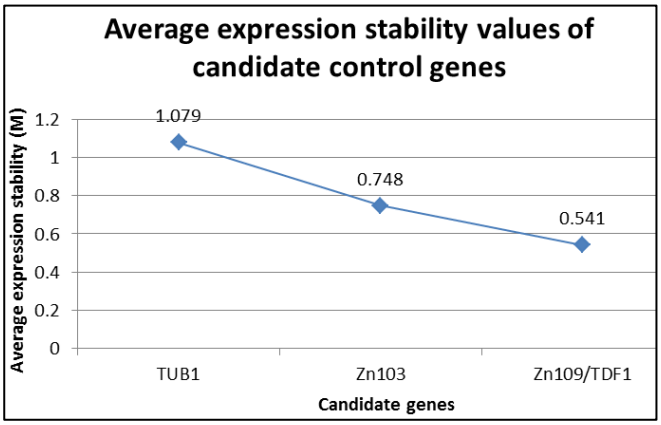


Figure 9: Average expression stability values of candidate control genes. (M-value = the average pairwise variation of a gene with all the other candidate reference genes. M is first calculated for all the genes and then one by one the least stable genes are excluded and new M values are calculated. Consequently, genes are arranged from least stable to most stable from left to right).

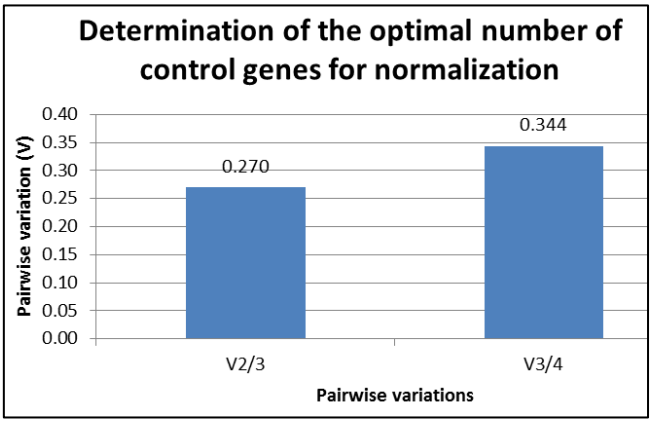


Figure 8: Determination of the optimal number of control genes for normalization (V2/3 = the pairwise variation between the sequential NFs of 2 and 3 control genes, V3/4 = the pairwise variation between the sequential NFs of 3 and 4 control genes)

7.3.2 Gene expression

Results from the qRT-PCR experiment were only regarded as statistically significant if the p-value was lower than 0.05. To describe data with a p-value lower than 0.1 the term "borderline significant" will be used. Furthermore will the term "treatment" be used to discuss the 2 different Zn concentrations that were tested (20µM Zn and 1000µM Zn).

(1) Gene 1: ZIP (229544)

This was the *S. luteus* predicted ZIP - ATX2 homolog (Section 6.1.2). **When exposed for 48h to 1000µM Zn the sensitive isolates were upregulated** (Figure 10.A). P4 was significantly upregulated (p = 0.0133) and Mg4 borderline significantly (p = 0.0715). Meanwhile, the expression of the tolerant isolates Ls1 and OF3 remained stable despite the Zn treatment.

Further analysis, comparing the different isolates with each other, proved that **the isolates were significantly different** from each other (p = 0.00282) and that there was a borderline significant interaction effect between the treatments and the isolates (p = 0.0574). The tolerant isolate, **Ls1, was significantly different from** the sensitive isolates **Mg4** (p = 0.0309) and **P4** (p = 0.00367). Furthermore was the sensitive P4 isolate also borderline significantly different from the other tolerant isolate OF3 (p = 0.0572).

The data that examined "the interaction effect" between all the groups (Figure 11.A) showed that the **P4 sensitive isolate reacted significantly different when exposed to 20µM than the tolerant isolates Ls1** (p = 0.0133) and **OF3** (p = 0.0172) did. For the other sensitive isolate Mg4 the reaction when exposed to 20µM Zn was only borderline significantly different from the reaction of Ls1 (p = 0.0937) and not significantly different from the reaction with OF3 (p = 0.116).

(2) Gene 2: ZIP (720881)

This was the *S. luteus* ZIP protein belonging to an uncharacterized cluster in the phylogenetic tree (Section 6.1.2). Only in isolate OF3 was a difference observed between the 2 Zn treatments. **OF3 was significantly downregulated when exposed to 1000µM Zn** (p = 0.00105) (Figure 10.B).

Both the Zn treatments and the choice of isolate were proven to have a significant effect on the gene expression ($p_{treatment}$= 0.0123 and $p_{isolates}$= 1.55E-11). However, the interaction effect between these two parameters was found to be only borderline significant (p = 0.0698). A**ll the isolates were significantly different from each other** (highest p-value = 0.00150), with the exception of the isolates P4 and OF3. They were only borderline significant different from each other (p = 0.0981).

A **significant different interaction effect** (Figure 11.B) was found for **Ls1 with P4** (p = 7E-6) **and OF3** (p = 0) **and for Mg4 with OF3** (p= 4.2E-6) **and P4** (p = 0.00375), **when** all of these isolates were **exposed to 20µM Zn**. **Ls1 was also** found to have **a significant different interaction effect with Mg4** (p = 0.0248), **OF3** (p= 8.3E-6) and **P4** (p= 0.000426) **under the 1000 µM Zn conditions**.

(3) Gene 3: CDF (810602)

This was the *S. luteus* CDF homolog aligned with the *S. cerevisiae* MMT1 and MMT2 Fe transporters (Section 6.1.1). **None of the isolates showed a significant up- or down- regulation for this gene caused by the Zn treatment** (Figure 10.C).

When a comparison between the different isolates was made, **no significant difference between the isolates and the two Zn conditions** was found. Also **the reaction to the Zn treatment was not significantly different between the isolates** (Figure 11.C).

(4) Gene 4: ZIP (811220) (Figure 10.D)

This was the *S. luteus* predicted ZIP - YKE4 homolog (Section 6.1.2). Figure 10.D shows that **both the sensitive isolates were significantly upregulated when exposed to 1000 µM** (P4 p = 0.0386; Mg4 p = 0.00105). The expression of the tolerant isolates did not change significantly.

The combined data of all the isolates did not have a normal distribution for this gene; therefore isolates were compared with each other via a non-parametrical test. There was a difference among the isolates and their reaction to the Zn-treatments (p = 0.0139).

A **significant difference was found between the gene expression by 20µM Zn of the sensitive isolate Mg4 and the expression of the tolerant isolates LS1** (p = 0.00790) **and OF3** (p = 0.0159) (Figure 11.D). The other sensitive isolate, P4 was borderline significantly different from the tolerant isolates Ls1 (p = 0.0952) and OF3 (p = 0.0952) when exposed to 20µM Zn.

(5) Gene 5: ZIP (22926)

This was the *S. luteus* ZIP protein belonging to the ZRT1/ZRT2-like cluster in the phylogenetic trees (Section 6.1.2). This gene was **significantly downregulated in all isolates** (Figure 10.E). In both tolerant isolates, Ls1 (p = 0.00105) and OF3 (0.00105) and both sensitive isolates, P4 (0.00253) and Mg4 (0.00105) gene 5 was almost completely turned off when exposed to 1000µM Zn.

Further testing confirmed a significant effect of the treatment (p < 2.2E-16). It was found that **all the isolates were significantly differing from one another** (highest p = 0.0357).

Furthermore, it was found that the general interaction effect between the Zn treatments and the isolates was not significant. Individual interaction effects are given in Figure 11.E.

(6) Gene 6: CDF (72605)

This was the *S. luteus* CDF homolog belonging to the uncharacterized rest group in the CDF phylogenetic tree (Section 6.1.1). **Two of the isolates showed a significant downregulation when exposed to 1000 µM**: the sensitive isolate **P4** (p = 0.00254) and the tolerant isolate **Ls1** (p = 0.0163). The other tolerant isolate, OF3 also had a downregulation, but this was only borderline significant (p = 0.0715). For the sensitive isolate Mg4 also a downregulation was observed in the data, but this was not statistical significant (p = 0.122) (Figure 10.F).

Moreover, it was found that there was general **significant effect of the treatment** (p = 0.000296) and that there were **no similar expressed isolates for this gene** (highest p = 0.0209 for the difference between Mg4 and Ls1).

No significant difference was found among the different isolates **in the manner of reacting** to the treatment (general interaction effect between isolate and treatment p = 0.733) (Figure 11.F).

(7) Gene 7: CDF (72657)

This was the *S. luteus* CDF homolog aligned with the *S. cerevisiae* MSC2 Zn transporter (Section 6.1.1). **None of the isolates had a significant up- or down- regulation when exposed to 1000µM Zn.** (Figure 10.G)

When comparing the different isolates with each other, it was found that only the sensitive isolate **Mg4 was significantly different from** the other sensitive isolate **P4** (p = 0.0485).

For **none of the isolates a significant interaction effect** was found between the choice of isolate and the Zn treatment (Figure 11.G).

(8) Gene 8: CDF (798077)

This was the *S. luteus* CDF homolog aligned with the *S. cerevisiae* ZRG17 (Section 6.1.1). **None of the isolates showed a significant up- or down- regulation caused by the Zn treatment** (Figure 10.H).

It was found that the sensitive isolate **Mg4 was significantly different from** the other sensitive isolate **P4** (p = 0.00234) **and** the tolerant isolate **Ls1** (p = 0.0331). (Also a borderline significantly difference was found between Mg4 and the tolerant isolate OF3 (p = 0.0786)).

Figure 11.H demonstrates that **only a borderline significant** difference was found between the **interaction effect of Mg4 and P4** when exposed to 20 µM Zn (p = 0.0873).

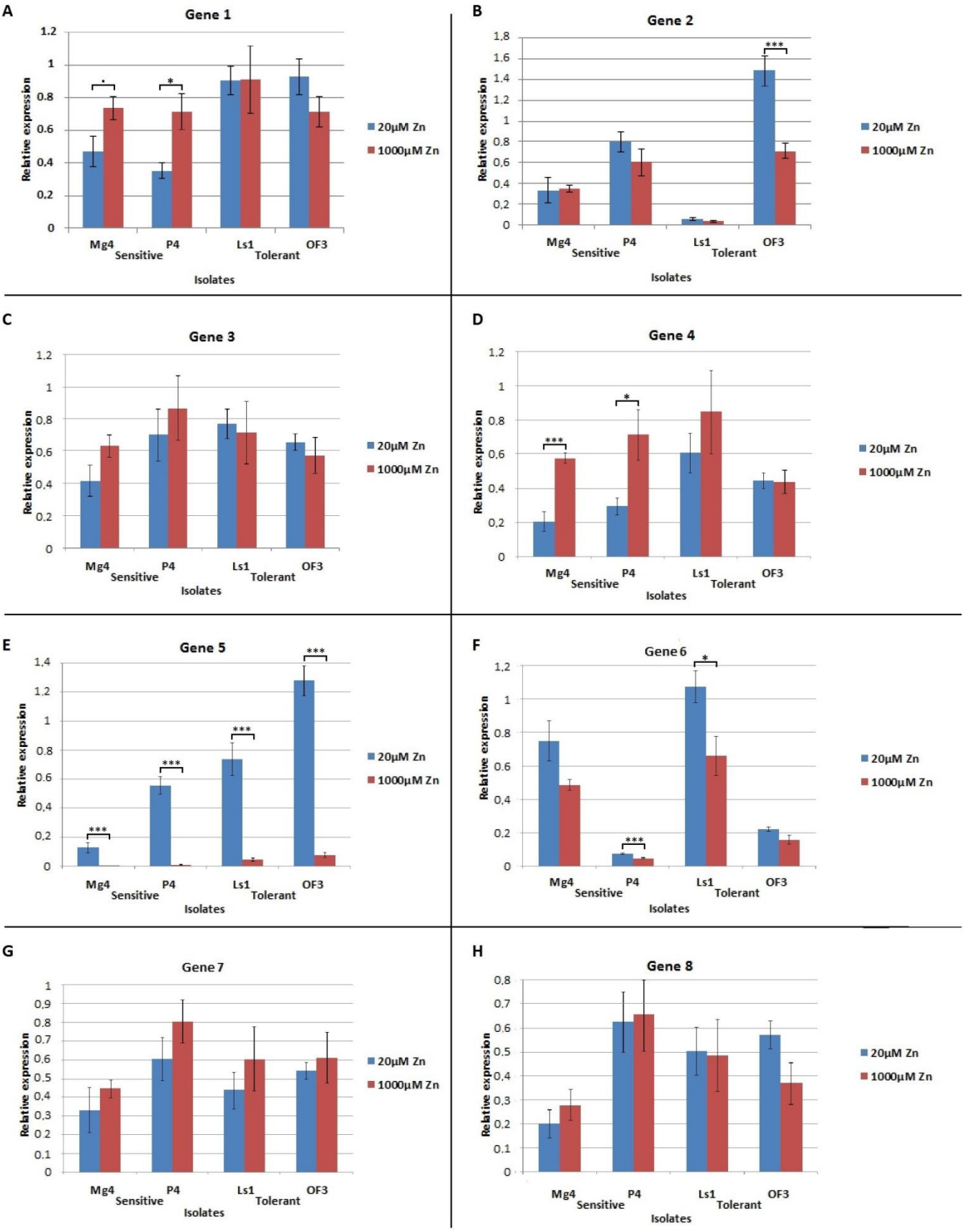


Figure 10: 1-way ANOVA; Gene up- or down-regulation within one specific isolate.
Data shown are means (of 5 biological replicates) and standard errors. The following symbols are used to describe the level of significance: "*" for p < 0.005, "**" for p < 0.01, "*" for p < 0.05 and "." for p < 0.1.**

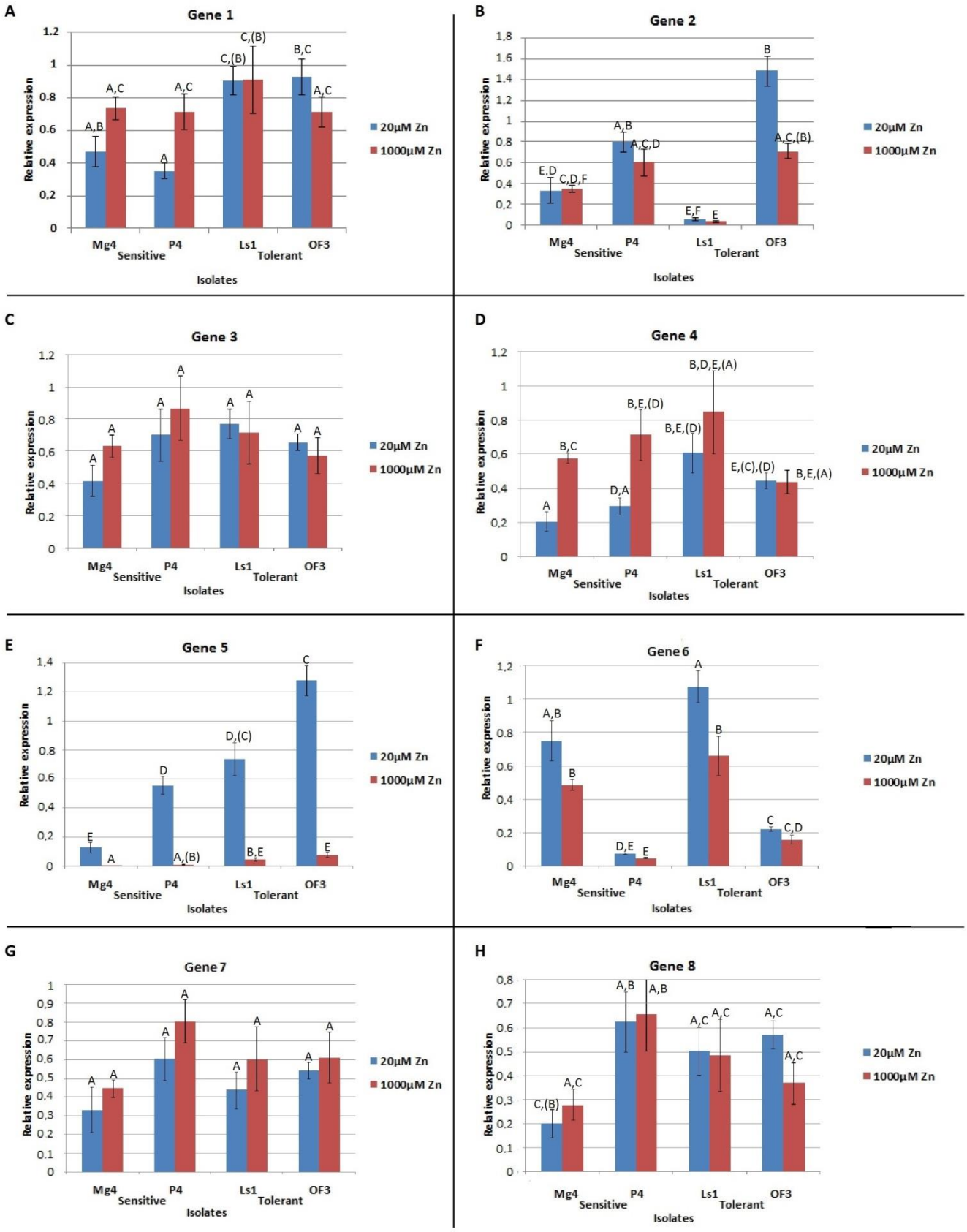


Figure 11: Differences among the isolates; 2-way ANOVA and Tukey HSD test: Interaction effects between Zn treatment and isolates. Data shown are means (of 5 biological replicates) and standard errors. The level of significance is indicated by the letters on top of each column: columns with the same letter are not significantly different from each other (p >0.1), columns that have a letter between brackets are only borderline significantly different from each other (0.1 > p > 0.05).
Gene 4 was not done using the Tukey HSD test as the data were not normally distributed, the data were obtained via a Kruskal-Wallis test followed by a Pairwise Wilcoxon Rank-Sum test.

8. Discussion

8.1 General discussion

8.1.1 The phylogenetic trees

The phylogenetic trees clearly illustrated that that all of the *S. luteus* proteins had a high chance of being real CDF, ZIP and Nramp homologs, because **(1)** all of the homologs clustered within a larger, encompassing group that had a high (for the Nramp family) or relatively high (for the CDF and ZIP family) bootstrap value. "Relatively high", because it needs to be kept in mind that the CDF and ZIP are really divergent protein families and therefore will have lower bootstrap values than average, more conservative protein families (e.g. the Nramp family). **(2)** All of the *S. luteus* homologs were also found in a closely related sister species *S. brevipes*. So it can be assumed that the *S. luteus* proteins are really present in the genome and not the result of a sequencing error or a contamination in the samples.

All of the homologs between *S. luteus* and *S. brevipes* were really similar, except for the *S. brevipes* CDF ZRG17 homolog. This homolog was not able to make the selection criteria and did not have a good alignment with the other ZRG17 proteins.
On the other hand, the *S. luteus* ZRG17 protein did have a good alignment with the other reference proteins and did meet the selection criteria. Therefore, it is assumed that the data for the *S. luteus* ZRG17 protein were reliable.
As the other 12 homologs of the 2 Suilloid fungi were all very similar to each other (Figure 4, 5, 6 and 7), it might be that the sequencing for the ZRG17 gene in *S. brevipes* was not very accurate. Another possibility is that maybe the 2 species are just for this gene very divergent.

Concerning the functional information that was obtained by including more characterized protein sequences in the tree it is important to realize that this is only a hypothesized function for the proteins. In order to have a reliable functional characterization of the proteins, their function needs to be validated in a real-life experiment (e.g. via a yeast complementation experiment).

Most of the characterized homologs that were included in the tree as reference proteins were *S. cerevisiae* sequences. Other characterized proteins that were more closely related to *S. luteus* could not be found. This makes the phylogenetic trees difficult to interpret, as the species *S. luteus* and *S. cerevisiae* are not closely related and have a very different lifestyle (ECM versus yeast).

Another factor that complicates the interpretation of the phylogenetic trees is the nature of the proteins themselves. These three protein families have all been found to transport more than one substrate and furthermore, it is often difficult to assess which the primary transport substrate of these transporters is.[41,51] Also the location and direction (import or export) of these transporters is sometimes difficult to predict as a lot of these proteins have been found in several membranes in the cell and have been shown to import a metal in one organism and export it in another.[41,52,53] This inconsistency can be explained by the fact that **(1)** A modification of the signal-peptide or of a certain residue in a transporter can lead to the conversion of an importer into an exporter (and vice

versa) or to alter the transporter's localization in the cell[75,76] and **(2)** some of these transporters have shown that they are sometimes able to transport bidirectionally, changing their transport direction depending on fluctuations in proton and/or divalent cation concentrations.[53,77]

Consequently, the functional information assigned to the different clusters in the phylogenetic trees needs to be interpreted very carefully. A good alignment with a characterized transporter will not automatically guarantee that all the proteins in that cluster have the same function, although the chances of them having the same function are relatively high. Therefore the functional information provided by the phylogenetic trees can be used to predict the function of the *S. luteus* homologs and to interpret the data of the qRT-PCR experiment, but they cannot be used as a tool for to determine accurate protein functions.

A few clusters in the phylogenetic trees (Figure 4, 5, 6 and 7) had no alignment with an already characterized reference protein; therefore no functional information could be assigned to these homologs. This lack of characterized homologs is partly explained by the fact that these protein families (especially the ZIP and the Nramp family) have been less studied than other protein families and particularly in Fungi. Another reason for this absence can be that maybe these uncharacterized groups represent a specific evolutionary adaptation in the Basidiomycota, hereby forming a specific group of transporters typical for this taxonomical group. Figure 6, Subfamily 3 the ZRT3-like proteins, represents such an already characterized group. Homologs of this group have only been found in the Ascomycota.

In the CDF phylogenetic tree (Figure 4 and 5) only the function of the rest group was not indicated by an already characterized reference protein. This rest group can represent a new evolutionary branch that is present in some members of the Agaricales (an order of Basidiomycetes) as this cluster (Figure 5), only has three members and they are all Agaricales: *S. luteus, S. brevipes* and *H. cylindrosporum*. This latter species was also included in a sub-cluster in the ZRC1/COT1 cluster, but no *S. luteus* homolog was assigned to this group.
Figure 4 illustrates that *S. luteus* has in general an average number of Zn CDFs: two homologs in the ZRC1/COT1-like cluster, one in the MSC2-like cluster and one in the ZRG17-like cluster. *L. bicolor* showed a remarkable paralog abundance in Subfamily 1 and 4.
For the second part of the CDF phylogenetic tree, *S. luteus* also has an average number of homologs (Figure 5): one in Subfamily 3 and in Subfamily 2, (although also a number of species have 2 paralogous sequences for this latter subfamily.)
In general, it can be concluded that for the CDF family no apparent members were missing from the inventory and that a possible extra CDF homolog was found. ("Possible", because its functionality needs to be verified, confirmation that the homolog is not a pseudogene and "extra", because this bifurcation has never been described before.)

For the ZIP family (Figure 6) no functional information could be obtained for the group that had a bifurcation with the ZRT1/ZRT2 cluster. This group consisted of 5 members: *H. cylindrosporum, L. bicolor, H. irregulare, S. brevipes and S. luteus* and except from *H. irregulare*, these species are all Agaricales. Therefore, this cluster and its lack of a characterized reference protein can be explained

by the same hypothesis as the one for the rest group of the CDF family: namely, the cluster represents a new evolutionary branch.
For the other clusters that all contained one *S. luteus* ZIP homolog everything seems "standard".
In general, it can be concluded that also for the ZIP family no apparent members are missing from the inventory and possibly an additional ZIP is found. However, this phylogenetic tree only consisted of a small number of sequences, due the necessity make the selection criteria more strict. This was necessary as the ZIP sequences were even more divergent than the CDFs.

Since the Nramp family tree (Figure 7) had a clustering according to the taxonomical groups of the fungi, no functional information could be appointed to these homologs. The Nramp phylogenetic tree had distinct clusters for the Basidiomycota and the Ascomycota. As only *S. cerevisiae* characterized Nramp proteins could be used as fungal reference proteins, none of the 2 *S. luteus* homologs could be aligned with an already the characterized Nramp protein. In general most of the Basidiomycete fungi also had 2 Nramp homologs, except for *L. bicolor* that was in this case a lesser represented species.

To summarize, this project identified 7 CDF, 4 ZIP and 2 Nramp homologs in *S. luteus*. Neighbour-joining phylogenetic trees were constructed in order to verify the identity of the predicted genes and to appoint a potential function to them. The trees also provided a global picture of the conservation or reduction of the protein families during evolution in the investigated species.

8.1.2 Gene expression data

Eight *S. luteus* proteins were selected for the qRT-PCR experiment: four CDF and four ZIP sequences. The Nramp family was excluded from further investigation, **(1)** because the CDF and ZIP family are more important for Zn transport and homeostasis[37], **(2)** because they primarily transport Mn and Fe[51] and **(3)** because no functional information could be obtained for these homologs. Furthermore were also three CDF homologs excluded: the two CDF ZRC1/COT1 homologs, as they were already intensively investigated in previous experiments (unpublished, Ruytinx) and the predicted CDF Mn transporter, as this homolog probably has Mn as primary substrate.

The predicted CDF Fe transporter was not excluded, because as mentioned in the introduction Fe and Zn are chemically very similar elements and consequently this transporter could be relevant for Zn as well.

Interpreting the gene expression data is not evident as **(1)** *S. luteus* is known for its high level of genetic variation that is also present in the gene expression [78] and **(2)** only a small number of isolates (two tolerant and two sensitive) were tested. Additionally, this small sample size will only enable big differences in gene expression to be statistical significant.
However, despite these difficulties, there were still several very apparent trends among the tested genes. Therefore this first experiment investigating the gene expression could be regarded as a

success, since the most pronounced effects of the Zn conditions and differences among the isolates are characterized.

In the next paragraphs the investigated genes will be discussed one by one in order of descending stability among the isolates and conditions:

Gene 3, the CDF predicted MMT homolog, **gene 7**, the CDF predicted MSC2 homolog and **gene 8**, the CDF predicted ZRG17 homolog were stably expressed in all the isolates among the 2 Zn conditions.

For **gene 3** this can be explained by the fact that the gene probably encodes a mitochondrial metal importer with Fe^{2+} as primary substrate.[40] However, as this transporter is also know to transport Zn^{2+}, Co^{2+}, Cd^{2+} and Ni^{2+}[40] its gene expression could also be only slightly affected with the minor up- or down- regulation not statistically significant. Figure 11.C shows that the sensitive isolates P4 and Mg4 are slightly upregulated when exposed to 1000 µM Zn (not significantly, only when looking at the mean values) and that the tolerant isolates Ls1 and OF3 are slightly downregulated. This difference between the tolerant and sensitive isolates in reacting to the higher Zn concentration can be expected, keeping in mind that the tolerant isolates are suspected of transporting the Zn outside the cell. While the sensitive isolates would probably try to store the excess Zn in subcellular compartments.
Gene 7 was predicted to be an endoplasmic reticulum/Golgi Zn importer MSC2-like, so an upregulation under the influence of Zn would be likely. Figure 11.G shows that all the genes had a higher mean value of expression when exposed to 1000 µM Zn. This trend was not statistical significant. **Gene 8** had exact the same expression profile as Gene 7 (Figures 10.G, 10.H, 11.G and 11.H): There are no relevant significant differences between the isolates and between the two Zn conditions. When looking at the mean values of the expression levels the 2 genes have the same isolates corresponding to highest and lowest mean expression (for both the 2 Zn conditions): P4 > Ls1 > OF3 > Mg4 (when arranged in descending level of gene expression).
This similarity in expression between gene 7 and 8 can be explained be the fact that they form a hetero-oligomeric complex together that maintains the homeostasis in the ER.[41]

Gene 6 was predicted to be a member of the CDF protein family, but could not be aligned with an already characterized reference protein. Figure 11.F illustrates that for this gene all the mean expression values of the isolates were lower in the 1000 µM Zn condition when compared to the 20 µM Zn condition. Only for the tolerant isolate Ls1 was this downregulation significant.
Besides the general trend of being downregulated, all of the isolates also show a different level of expression.

Gene 2 is another protein that could not be aligned to a characterized reference protein. It was identified as a member of the ZIP protein family and was significantly downregulated in the tolerant isolate OF3 when exposed to 1000 µM Zn (Figure 10.B). In the other isolates Ls1 and P4 the mean expression values also showed a downregulation under the influence of Zn, but this was not statistically significant. The isolate Mg4 did not show a downregulation (Figure 10.B). When

looking at the figure it becomes clear that all of the isolates have a very dissimilar level of expression (Figure 11.B).

This large variability in the expression levels among the different isolates, specifically observed for Gene 2 and Gene 6, can be explained by the high level of genetic variation in *S. luteus*.
A high level of genetic variation is common in most pioneer species and in *S. luteus* this variation is also visible in the gene expression levels.[78]

Gene 4 was predicted to be a ZIP YKE4 homolog, this protein functions as an ER bidirectional Zn transporter in *S. cerevisiae*.[73] Depending on the cytoplasmic Zn concentrations the transporter functions as an importer or exporter. Figure 11.D shows that the sensitive isolates were significantly (Mg4) and borderline significantly (P4) upregulated when exposed to 1000 µM Zn, while the expression of the tolerant isolates remained stable between the two Zn treatments. When comparing the tolerant and sensitive isolates, no difference in the expression levels was detected in the 1000 µM Zn condition. In the 20 µM Zn condition however, the sensitive isolates had a borderline significantly lower expression compared to the tolerant isolates.
These results suggest that this gene might contribute to the difference in level of Zn tolerance between the tolerant and sensitive isolates, as the tolerant isolates already have a high expression of the gene under normal conditions and do not have to upregulate the gene when exposed to higher concentrations of Zn. In contrast, the sensitive isolates need to first upregulate the gene and can therefore not immediately react to the higher Zn concentrations.

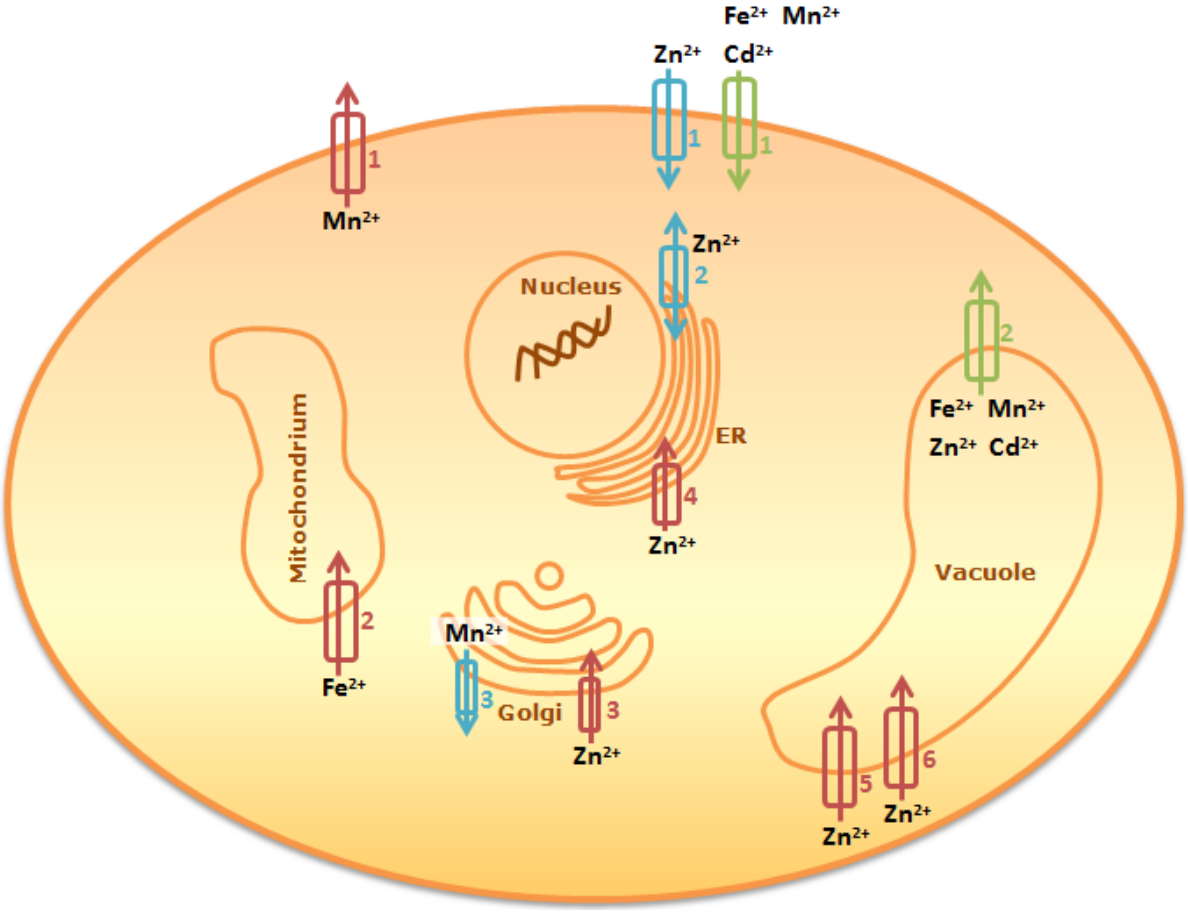
Gene 1 aligned with the *S. cerevisiae* ATX2 Mn transporter localised to the membrane of the Golgi vesicle.[70,72] This gene had a similar expression profile as gene 4, although the difference between the tolerant and sensitive isolates was less pronounced (Figure 11.A). Because of the similar expression pattern, the same hypothesis as for gene 4 can be made: namely, that this gene might contribute to the tolerance difference between the sensitive and tolerant isolates.

To summarize, both Gene 1 and Gene 4 are very interesting results concerning the mechanism of Zn tolerance.

Gene 5, the predicted ZRT1/ZRT2 homolog, was significantly downregulated in all the isolates when exposed to 1000 µM Zn. This can be expected as the ZRT1 gene codes a high-affinity Zn transporter of the plasma membrane, responsible for the majority of Zn uptake in *S. cerevisiae*.[45,47] Based on the expression data (Figure 10.E and 11.E) it can be stated that the *S. luteus* homolog has probably the same function as the *S. cerevisiae* ZRT1 protein.

To summarize, Gene 5 is a very important find in the light of the Zn homeostasis.

Figure 12 gives an overview of all the transporters that were found. Two homologs could not be illustrated in the figure, as no functional information was obtained for them via the phylogenetic trees (one CDF homolog and one ZIP homolog). The illustrated (Figure 12) localization of the Nramp transporters is purely hypothesized, based upon literature.[50,52,79] Furthermore, illustrates Figure 12 clearly the antagonistic action of the CDF and ZIP family.[37] Based on the results of this project the most important transporter for the Zn homeostase is ZIP 1 (gene 5 in the qRT-PCR) and for the Zn tolerance mechanism ZIP 2 (gene 4 in the qRT-PCR) and ZIP 3 (gene 1 in the qRT-PCR).



8.2 Conclusion

As discussed in the previous section 8.1, all the data should be interpreted with caution.
However, the good alignment of the phylogenetic trees attributed more confidence to these results and hence it can be concluded that now a **general overview of the most relevant Zn transporters** is available: **a successful inventory of the CDF, ZIP and Nramp famil**y was made.

Although all the investigated genes had an interesting result, 3 genes are particularly interesting: Gene 1, Gene 4 and Gene 5.
Gene 1 and Gene 4, because of their **potential role in the Zn tolerance mechanism** observed in some ecotypes of *S. luteus* (in this study in the tolerant isolates OF3 and Ls1).
Gene 5, because of its **importance for the general Zn homeostasis** in the cell.

Furthermore these results again confirmed the **high inherent genetic variation that is also visible in the gene expression of *S. luteus***.[78]

9. Future perspectives

For this project all of the bio-informatics data were based on version 1.0 of a sequenced *S. luteus* isolate. Therefore a resequencing could be done to improve the reliability of the data and other isolates could be sequenced as well to enable the direct comparison between isolates via several online bio-tools.

The isolate that was sequenced had an average level of tolerance, so it would be interesting to also have the sequence of a hyper-sensitive and a hyper-tolerant isolate available as well. This would make it easier to study and identify new genes and investigate the differences in expression-related parameters among different isolates. Expression-related parameters like the promotor region and the gene copy number could be compared for all the sequenced strains online with only a minimum amount of time and effort. Furthermore, it would also be easier to make a selection of interesting genes first and then validate the genes in a real-life experiment, hereby lowering the laboratory costs of the study.

As mentioned before, our expression data need further confirmation. It would be interesting to repeat the experiment with: **(1)** a higher sample size, **(2)** more isolates, **(3)** more conditions (Zn concentrations), **(4)** more time points (harvesting of the mycelia) and **(5)** different sample origins (mycelium monokaryons, mycelium dikaryons and *in mycorrhiza*)
(1) Because, a higher sample size will enable smaller differences in gene expression to become more apparent and reliable (statistically significant). **(2)** The use of more isolates will make it easier to see trends between the tolerant and sensitive strains and/or to distinguish groups among the isolates. **(3)** More conditions (several Zn concentrations) will provide insights into potential dose-response relationships between the Zn concentrations and gene expressions and enable the identification of possible threshold values for the up- or down- regulation of a gene. **(4)** As gene expression is dynamic through time, it would be interesting to also measure several time points in order to get a more comprehensive understanding of the response during the whole stress period. Hereby it would be possible to observe a difference between the genes first to be affected, involved in the primary response to Zn and the genes that are affected later on, dealing with the secondary toxicity effects. **(5)** Furthermore, it will also be interesting to compare the expression of genes in mycelia and in mycorrhiza and between mono- and di- karyons. In a dikaryon the gene expression will be influenced by the genomes of both nuclei, and therefore the copy number and promotor sequences of the genes in both genomes need to be taken into account.

In addition, it would also be interesting to study whether the genes are also influenced by other metals like Fe and Cd that are chemically very similar to Zn. Especially, Cd because it was also found as a pollutant on the contaminated sampling areas (See annex 11.1 and 11.2).
Several of the characterized reference proteins have already been implicated in Fe and/or Cd transport.[11,79]

Moreover, while this project was a study that focused on Zn transporters, it is also interesting to investigate other proteins and genes important for the Zn homeostasis and Zn tolerance, like chelating molecules (e.g. MTs, Gluthatione, etc.) and other oxidative stress-linked molecules (e.g. Super Oxide Dismutases, Heat Shock Proteins, etc.).

In the future, the knowledge and insights into Zn homeostasis and tolerance will hopefully contribute to the development of better phytostabilisation techniques. Specific metal-adapted plant-fungus combinations suited for bioremediation purposes, like phytostabilisation could be developed when a clear and comprehensive understanding of the relevant tolerance and homeostasis mechanisms is available.

10. References

1. Leyval C, Turnau K & Haselwandter K. Effect of heavy metal pollution on mycorrhizal colonization and function: physiological, ecological and applied aspects. *Mycorrhiza* **7**, 139–153 (1997).
2. Adriaensen K, Vangronsveld J & Colpaert JV. Zinc-tolerant Suillus bovinus improves growth of Zn-exposed Pinus sylvestris seedlings. *Mycorrhiza* **16**, 553–8 (2006).
3. Colpaert JV, Adriaensen K, Muller LAH, Lambaerts M, Faes C, Carleer R & Vangronsveld J. Element profiles and growth in Zn-sensitive and Zn-resistant Suilloid fungi. *Mycorrhiza* **15**, 628–34 (2005).
4. Muller LAH, Vangronsveld J & Colpaert JV. Genetic structure of Suillus luteus populations in heavy metal polluted and nonpolluted habitats. *Molecular ecology* **16**, 4728–37 (2007).
5. Krznaric E, Wevers JHL, Cloquet C, Vangronsveld J, Vanhaecke F & Colpaert JV. Zn pollution counteracts Cd toxicity in metal-tolerant ectomycorrhizal fungi and their host plant, Pinus sylvestris. *Environmental microbiology* **12**, 2133–41 (2010).
6. Norrstrôm AC. Concentration and chemical species of iron in soils from groundwater / surface water ecotones. *Hydrological Sciences Journal* **40**, 319- 29 (1995).
7. Frausto da Silva JJR & Williams RJP. The biological chemistry of the elements (2nd edition). *Clarendon Press* (2001).
8. Alloway BJ. Zinc in SoilS and Crop Nutrition. *International Zinc Association* **1**, 1-137 (2008).
9. Hart JJ, Welch RM & Norvell WA. Transport interactions between cadmium and zinc in roots of bread and durum wheat seedlings. *Physiol Plant.* **116**, 73–78 (2002).
10. Clemens S. Evolution and function of phytochelatin synthases. *Journal of plant physiology* **163**, 319–32 (2006).
11. Clemens S. Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants. *Biochimie* **88**, 1707–19 (2006).
12. Bellion M, Courbot M, Jacob C, Blaudez D & Chalot M. Extracellular and cellular mechanisms sustaining metal tolerance in ectomycorrhizal fungi. *FEMS microbiology letters* **254**, 173–81 (2006).
13. Petrucci RH, Herring FG, J. D. M. and C. B. General Chemistry: Principles and Modern Applications (9th edition) *Prentice Hall* (2010).
14. Berg J & Shi Y. The galvanization of biology: a growing appreciation for the roles of zinc. *Science* **271**, 1081–1085 (1996).
15. Schulte EE. Soil and Applied Zinc (A2528). *Understanding plant nutrients* **A2528**, 1–2 (2004).
16. Baize D. *Teneurs totales en elements traces metalligues dans les sols (France) INRA Editions, Paris*. 1-409 (1997).
17. Yadav SK. Heavy metals toxicity in plants: An overview on the role of glutathione and phytochelatins in heavy metal stress tolerance of plants. *South African Journal of Botany* **76**, 167–179 (2010).
18. Clemens S. Molecular mechanisms of plant metal tolerance and homeostasis. *Planta* **212**, 475-486 (2001).
19. Li S, Yu J, Zhu M, Zhao F & Luan S. Cadmium impairs ion homeostasis by altering K+ and Ca2+ channel activities in rice root hair cells. *Plant, cell & environment* **35**, 1998–2013 (2012).
20. Abbà S. A PLAC8-containing protein from an endomycorrhizal fungus confers cadmium resistance to yeast cells by interacting with Mlh3p. *Nucleic acids research* **39**, 7548–63 (2011).
21. Kenderesová L, Stanova A, Pavlovkin J, Durisova E, Nadubinska M, Ciamporova M & Ovecka M. Early Zn2+-induced effects on membrane potential account for primary heavy metal susceptibility in tolerant and sensitive Arabidopsis species. *Annals of botany* **110**, 445–59 (2012).
22. Hall JL. Cellular mechanisms for heavy metal detoxification and tolerance. *Journal of experimental botany* **53**, 1–11 (2002).
23. Blaudez D. Differential responses of ectomycorrhizal fungi to heavy metals in vitro. **104**, 1366–1371 (2000).
24. Bearoff FM & Fuller-espie SL. Alteration of mitochondrial membrane potential (Δ Ψm) and phosphatidylserine translocation as early indicators of heavy metal-induced apoptosis in the earthworm Eisenia hortensis. *ISJ* **8**, 98–108 (2011).
25. Beyersmann D & Hartwig A. Carcinogenic metal compounds: recent insight into molecular and cellular mechanisms. *Archives of Toxicology* **82**, 493–512 (2008).
26. Jin YH, Clark AB, Slebos RJ, Al-Refai H, Taylor JA, Kunkel TA & Resnick M. Cadmium is a mutagen that acts by inhibiting mismatch repair. *Nature Genetics* **34**, 326–329 (2003).
27. Colpaert JV, Muller LAH, Lambaerts M, Adriaensen K & Vangronsveld J. Evolutionary adaptation to Zn toxicity in populations of Suilloid fungi. *New Phytologist* **162**, 549–559 (2004).
28. Lan H, Wang Z, Wang Q, Wang M, Bao Y, Huang J & Zhang H. Characterization of a vacuolar zinc transporter OZT1 in rice (Oryza sativa L.). *Molecular biology reports* **40**, 1201–10 (2013).
29. Haydon MJ & Cobbett CS. A novel major facilitator superfamily protein at the tonoplast influences zinc tolerance and accumulation in Arabidopsis. *Plant physiology* **143**, 1705–19 (2007).
30. Colpaert JV, Wevers J, Krznaric E & Adriaensen K. How metal-tolerant ecotypes of ectomycorrhizal fungi protect plants from heavy metal pollution. *Annals of Forest Science* **68**, 17–24 (2011).
31. Schat H, Llugany M & Bernhard R. Metal-specific patterns of tolerance, uptake and transport of heavy metals in hyperaccumulating and non-hyperaccumulating metallophytes. *Phytoremediation of Contaminated Soil and Water* **1**, 171–188 (2000).
32. Verbruggen N, Hermans C & Schat H. Molecular mechanisms of metal hyperaccumulation in plants. *New Phytol* **181**, 759–776 (2009).
33. Sinclair SA & Krämer U. The zinc homeostasis network of land plants. *Biochimica et biophysica acta* **1823**, 1553–67 (2012).

34. Sekler I, Sensi SL, Hershfinkel M & Silverman WF. Mechanism and Regulation of Cellular Zinc Transport. *Mol Med* **3**, 1–6 (2007).
35. Oteiza PI. Zinc and the modulation of redox homeostasis. *Free radical biology & med* **53**, 1748–59 (2012).
36. Lichten LA & Cousins RJ. Mammalian zinc transporters: nutritional and physiologic regulation. *Annual review of nutrition* **29**, 153–76 (2009).
37. Eide DJ. Zinc transporters and the cellular trafficking of zinc. *Biochimica et biophysica acta* **1763**, 711–22 (2006).
38. Nies DH & Silver S. Ion efflux systems involved in bacterial metal resistances. *Journal of industrial microbiology* **14**, 186–99 (1995).
39. Gustin JL, Zanis MJ & Salt DE. Structure and evolution of the plant cation diffusion facilitator family of ion transporters. *BMC evolutionary biology* **11**, 76-86 (2011).
40. Montanini B, Blaudez D, Jeandroz S, Sanders D & Chalot M. Phylogenetic and functional analysis of the Cation Diffusion Facilitator (CDF) family: improved signature and prediction of substrate specificity. *BMC genomics* **8**, 107 (2007).
41. Kambe T, Suzuki T, Nagao M & Yamaguchi-Iwai Y. Sequence similarity and functional relationship among eukaryotic ZIP and CDF transporters. *Genomics, proteomics & bioinformatics* **4**, 1–9 (2006).
42. Delhaize E, Kataoka T, Hebb DM, White RG & Ryan PR. Genes Encoding Proteins of the Cation Diffusion Facilitator Family That Confer Manganese Tolerance. *The Plant Cell* **15**, 1131–1142 (2003).
43. Blaudez D & Chalot M. Characterization of the ER-located zinc transporter ZnT1 and identification of a vesicular zinc storage compartment in Hebeloma cylindrosporum. *Fungal genetics and biology* **48**, 496–503 (2011).
44. Grotz N. Identification of a family of zinc transporter genes from Arabidopsis that respond to zinc deficiency. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 7220–4 (1998).
45. Eng BH, Guerinot ML, Eide MHS. Membrane Biology Sequence Analyses and Phylogenetic Characterization of the ZIP Family of Metal Ion. *J. Membrane Biol.* **166**, 1–7 (1998).
46. Yang X, Huang J, Jiang Y & Zhang HS. Cloning and functional identification of two members of the ZIP (Zrt, Irt-like protein) gene family in rice (Oryza sativa L.). *Molecular biology reports* **36**, 281–7 (2009).
47. Guerinot ML. The ZIP family of metal transporters. *Biochimica et biophysica acta* **1465**, 190–8 (2000).
48. Chen SL, Wang ZJ, Xu MY & Gui JF. Molecular identification and expression analysis of natural resistance associated macrophage protein (Nramp) cDNA from Japanese flounder (Paralichthys olivaceus). *Fish & shellfish immunology* **20**, 365–73 (2006).
49. Cellier M, Prive G, Belouchi A, Kwan T, Rodrigues V, Chia W & Gros P. Nramp defines a family of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 10089–93 (1995).
50. Pinner E, Gruenheid S, Raymond M & Gros P. Functional complementation of the yeast divalent cation transporter family SMF by NRAMP2, a member of the mammalian natural resistance-associated macrophage protein family. *The Journal of biological chemistry* **272**, 28933–8 (1997).
51. Nevo Y & Nelson N. The NRAMP family of metal-ion transporters. *Biochimica et biophysica acta* **1763**, 609–20 (2006).
52. Forbes JR & Gros P. Divalent-metal transport by NRAMP proteins at the interface of host-pathogen interactions. *Trends in microbiology* **9**, 397–403 (2001).
53. Techau ME, Valdez-Taubaz J, Popoff J, Francis R, Seaman M & Blackwell J. Evolution of differences in transport function in Slc11a family members. *The Journal of biological chemistry* **282**, 35646–56 (2007).
54. Haemig HAH & Brooker RJ. Importance of conserved acidic residues in mntH, the Nramp homolog of Escherichia coli. *The Journal of membrane biology* **201**, 97–107 (2004).
55. Muller LAH, Lambaerts M, Vangronsveld J & Colpaert JV. AFLP-based assessment of the effects of environmental heavy metal pollution on the genetic structure of pioneer populations of Suillus luteus. *New Phytologist* **164**, 297–303 (2004).
56. Peterson RL & Massicotte HB. Mycorrhizas : Anatomy and Cell Biology. *NRC Research Press* 1-182 (2004).
57. Bruns TD, Tan J, Bidartando M & Szaro TM. Survival of Suillus pungens and Amanita francheti ectomycorrhizal genets was rare or absent after a stand-replacing fire. *New Phytologist* **155**, 517–523 (2002).
58. Adriaensen K, Van der Lelie D, Van Laere A, Vangronsveld J & Colpaert JV. A zinc-adapted fungus protects pines from zinc stress. *New Phytologist* **161**, 549–555 (2004).
59. Muller LAH, Craciun AR, Ruytinx J, Lambaerts M, Verbruggen N, Vangronsveld J & Colpaert JV. Gene expression profiling of a Zn-tolerant and a Zn-sensitive Suillus luteus isolate exposed to increased external zinc concentrations. *Mycorrhiza* **17**, 571–80 (2007).
60. Muller LAH, Lambaerts M, Vangronsveld J & Colpaert JV. Isolation and characterization of microsatellite loci from the ectomycorrhizal basidiomycete Suillus luteus. *Molecular Ecology Notes* **6**, 165–166 (2006).
61. Macnair MR. The genetic control of copper tolerance in the yellow monkey flower, Mimulus guttatus. *Heredity* **50**, 283–293 (1983).
62. Macnair MR. The genetics of metal tolerance in vascular plants. *New Phytol* **124**, 541–559 (1993).
63. Ruytinx J, Nguyen H, Van Hees M, Op De Beeck M, Vangronsveld J, Carleer R & Adriaensen K. Zinc export results in adaptive zinc tolerance in the ectomycorrhizal basidiomycete Suillus bovinus. *Metallomics : integrated biometal science* **1**, 1-9 (2013).
64. Krznaric E, Verbruggen N, Wevers JHL, Carleer R, Vangronsveld J & Colpaert JV. Cd-tolerant Suillus luteus: a fungal insurance for pines exposed to Cd. *Environmental pollution* **157**, 1581–8 (2009).
65. Saier MH, Tran CV & Barabote RD. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic acids research* **34**, D181–6 (2006).
66. Ren Q, Chen K & Paulsen IT. TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic acids research* **35**, D274–9 (2007).

67. López-Millán AF, Ellis DR & Grusak M. a Identification and characterization of several new members of the ZIP family of metal ion transporters in Medicago truncatula. *Plant molecular biology* **54**, 583–96 (2004).
68. Rozen S & Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology* **132**, 365-387 (1994)
69. Schmittgen TD & Livak KJ. Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* **3**, 1101–1108 (2008).
70. Migeon A, Blaudez D, Wilkins O, Montanini B, Campbell MM, Richaud P, Thomine S & Chalot M. Genome-wide analysis of plant metal transporters, with an emphasis on poplar. *Cellular and molecular life sciences : CMLS* **67**, 3763–84 (2010).
71. Bird AJ. *Metallosensors, the ups and downs of gene regulation. Advances in microbial physiology* **53**, 231–67 (2008).
72. Lin SJ. Suppression of oxidative damage by Saccharomyces cerevisiae ATX2, which encodes a manganesetrafficking protein that localizes to Golgi-like vesicles. *Mol Cell Biol* **16**, 6303–6312 (1996).
73. Kuma´NA, Poruk KE, Osborn KA, Ward DM. YKE4 (YIL023C) encodes a bidirectional zinc transporter in the endoplasmic reticulum of Saccharomyces cerevisiae. *J Biol Chem* **281**, 22566–22574 (2006).
74. Vandesompele J, De Prete K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology* **3**, RESEARCH0034 (2002).
75. Tjalsma H, Bolhuis A, Jongbloed JD, Bron S & Van Dijl JM. Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *Microbiology and molecular biology reviews : MMBR* **64**, 515–47 (2000).
76. Von Heijne G. Signal Peptides. *eLS* **1***,* 1–3 (2005).
77. Goswami T, Bhattacharjee A, Baball P, Searle S, Moore E Ming L & Blackwell JM. Natural resistance associated macrophage protein 1 is an H+/bivalent cation antiporter. *Biochem J* **519**, 511–519 (2001).
78. Ruytinx J, Craciun AR, Verstraelen K, Vangronsveld J, Colpaert JV & Verbruggen N. Transcriptome analysis by cDNA-AFLP of Suillus luteus Cd-tolerant and Cd-sensitive isolates. *Mycorrhiza* **21**, 145–54 (2011).
79. Lin YF & Aarts MGM. The molecular mechanism of zinc and cadmium stress response in plants. *Cellular and molecular life sciences* **69**, 3187–206 (2012).
80. Blaby-Haas CE & Merchant SS. The ins and outs of algal metal transport. *Biochimica et biophysica acta* **1823**, 1531–52 (2012).
81. Aisen P, Enns C & Wessling-Resnick M. Chemistry and biology of eukaryotic iron metabolism. *Int. J. Biochem. Cell Biol.* **33**, 940–959 (2001).
82. Papanikolaou G & Pantopoulos K. Iron metabolism and toxicity. *Toxicology and applied pharmacology* **202**, 199–211 (2005).
83. Vig K. Bioavailability and toxicity of cadmium to microorganisms and their activities in soil: a review. *Advances in Environmental Research* **8**, 121–135 (2003).
84. Helmke PA. Chemistry of cadmium in soil solution. *Cadmium in Soils and Plants.* **1,** 39–64 (1999).
85. Lane T & Morel F. A biological function for cadmium in marine diatoms. *Proc Natl Acad Sci USA* **97**, 4627–4631 (2000).
86. Lane TW, Saito MA, George GN & Pickering IJ. Biochemistry: a cadmium enzyme from a marine diatom. *Nature* **435**, 42 (2005).
87. Odum HT. Heavy Metals in the Environment. *Taylor & Francis Ltd,* 1-344 (2000).

11. Annex

11.1 Heavy Metal pollution

11.1.1 Iron and Cadmium

Fe is, like Zn, also a transition metal and an essential micronutrient. It is the most common redox-active cofactor used in proteins. Fe plays a crucial role in vital biochemical activities (e.g. oxygen sensing and transport, electron transfer, etc.) and is consequently indispensable for life.[80,81] Fe is able to form various coordination complexes with organic ligands in a dynamic and flexible way and to switch between ferrous, Fe(II) and ferric, Fe(III) states.[13,82] However, the bioavailability of this essential element is very limited under normal aerobic conditions, because ferrous Fe(II) is readily oxidized to Fe(III) which is insoluble at pH > 4.[6,82] As a consequence higher bioavailable [Fe] will be found in soils fed by deep reduced groundwater, while lower [Fe] will be found at sites that receive more shallow groundwater.[6] In a study by Norrström (1995) $[Fe]_{total}$ was measured in the soil at 7 different locations in Sweden. It was found that the $[Fe]_{total}$ in these soils varied between 24 and 625 µmol g^{-1} dry mass.[6] All the soils were classified as mollic gleysols and 6 of the 7 sites were located in coniferous forests. The study also confirmed that most of the Fe was bound organically or in Fe-oxides and that the exchangeable Fe-fraction was of minor importance.[6]

Cd, is also a transition metal but a non-essential one[3]. It is chemically very similar to Zn[13] and consequently, substantial amounts of Cd are usually present in most Zn ores, but this type of Cd is not bioavailable.[7,8,83] Bioavailable Cd is the Cd present in the soil solution, where it often forms soluble complexes with chloride-, hydroxyl-, sulfhydryl- and thiol- groups. These complexes largely govern the biological activity of Cd.[83] Soils with no additional Cd-inputs from anthropogenic activities have usually $[Cd]_{soil\ solution}$ that range from 0.3 to 22.5 mg kg^{-1}, depending on the geological nature of the soil.[84]

11.1.2 Iron, Cadmium and Zinc: chemically similar elements

Zn, Fe and Cd are all transition metals and therefore share the common characteristics of this group of elements: **(1)** they are very hard (high melting and boiling points), **(2)** have a high electrical conductivity and malleability (because d-subshell electrons are loosely bound) and **(3)** have low ionization energies.[7,13] **Fe and Zn** are comparable elements, because they have roughly the same atomic and covalent radius (Table 1).[4] Despite the difference in atomic and covalent radius (Table 1), **Cd and Zn** are also very similar[13]. This is because they are both elements belonging to group 12 of the periodic table. Group 12 differs from the other groups of transition metals (groups 3-11) in that they never form compounds in which their oxidation states are higher than +II.[7,13] The explanation for this is that group 12 elements have an analogous shell

[3] There is one documented case of Cd having a biological function in the diatom *Thalassiosira weissflogii,* Cd was found acting as co-factor in a distinct isoform of carboanhydrase. This is normally a Zn^{2+}-dependent enzyme.[85,86]

[4] The Atomic radius of an atom represents it radius when it is not bound to another atom; the covalent radius represents the radius of that same atom when it is covalently bound to another atom.[13]

distribution: Zn = $1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$ $3d^{10}$ $4s^2$ and Cd = $1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$ $3d^{10}$ $4s^2$ $4p^6$ $4d^{10}$ $5s^2$ (both have a full d en s subshell in their outer shells).[13]

Table 1: Atomic and Covalent radii of Fe, Zn and Cd[7,13]

Radius	Fe	Zn	Cd
Atomic (pm)	140	135	155
Covalent (pm)	125	131	148

11.2 Sampling sites

Krznaric and colleagues (2009) assessed the Cd and Zn pollution at the *S. luteus* sampling sites by analysing the soil pore water. Their results are given in Table 2, Cd and Zn concentrations in the pore water.[64]

Table 2: Cd and Zn pollution in Lommel-Maatheide & a control area, Paal[64]

Site	pH	Pore water (µM)		Total (mg kg^{-1})	
		Zn	Cd	Zn	Cd
Meeuwen-Gruitorde	3.7	0.6	<0.1[A]	20	0.7
Paal	4.1	0.8	<0.1[A]	21	0.4
Maasmechelen	4.9	0.9	<0.1[A]	32	0.9
Lommel-Maatheide	5.9	120	6.2	1750	14
Lommel-Sahara	3.7	47	1.0	254	3.1
Neerpelt	4.4	49	1.2	442	3.9

[A] Detection limit