Chapter 1

General Introduction

Pollution is a worldwide problem and threatens human health as well as the environment. The European Union (EU) identifies pollution as one of the major threats to proper functioning of soil (Rodrigues et al., 2009). Examples of an impaired ability of soil to function properly are: failure of supporting vegetation or biomass production, disturbance of ecosystems and biological cycling of nutrients, and loss of filtering and buffering capacity affecting groundwater or aquatic ecosystems (Rodrigues et al., 2009; Scullion, 2006; van Straalen, 2002). In the EU metals (~37% of cases) and mineral oil (~34%) have been recognized as the largest groups of soil contaminants followed by polycyclic aromatic hydrocarbons (PAH; ~13%) (Rodrigues et al., 2009). Identification of contaminated sites occurs via soil assessment, which is an independent evaluation of soil’s health using chemical as well as biological methods. One important discipline in soil assessment is ecotoxicology. The science of ecotoxicology investigates the ecological effects of pollutants in the environment. Next to pollutants in the form of chemical substances, heat and noise may be considered as pollution as well. Ecotoxicological methods (especially bioassays) can be applied in ecological risk assessment in two different ways, namely using a so called ‘inverse’ or a ‘forward’ approach (van Straalen, 2002). The (more traditional) inverse approach is used to determine hazardous concentration limits of one (or a group of) toxic compound(s). In bioassays, an ecologically relevant species is exposed to soil containing the compound of interest in gradually increasing concentrations. Endpoints (e.g., reproduction, survival, and growth) are then evaluated and used to define dose-response relationships. Values such as 10% and 50% effect concentrations (EC10 and EC50 respectively) are then estimated and used to predict the maximal acceptable concentration limits of the investigated compound in the field (i.e., the concentration thresholds that are still safe for the environment). Chemical analyses at potentially contaminated sites are then used to answer the question whether the site is contaminated or not. A drawback of the inverse approach however, is that the soils at the field sites generally have properties different from the soils used in the bioassays. Usually, artificial or standardized field soils are applied in bioassays (e.g., OECD or LUFA 2.2, respectively). Soil properties such as pH, structure, clay content, and organic matter content play an important role in bioavailability of toxic compounds and will thus confound the
analysis of ecotoxicological endpoints. Therefore, it is not trivial to translate the concentration limits, as defined by bioassays, to thresholds upon which ecological risks can be reliably assessed in field situations. The forward approach is a bioassay that evaluates the actual soil samples taken from potentially contaminated sites in comparison with clean reference soil samples. A drawback with this approach is to find appropriate clean reference soil, which preferably has the same soil properties but without the contaminant.

Bioassays for soil ecological risk assessment are preferentially performed with ecologically relevant species, such as earthworms and springtails (Collembola). Springtails are one of the most abundant arthropods on earth, and due to their detritivorous role in terrestrial ecosystems, they contribute to soil fertility and cycling of nutrients (Fountain and Hopkin, 2005). The springtail *Folsomia candida* (Willem, 1902; Figure 1) is a species with a cosmopolitan distribution and is often identified as the most sensitive taxon to toxic contaminants in soil (Fountain and Hopkin, 2005). Furthermore, *F. candida*’s reproduction is parthenogenetic and fast which makes them easy to rear in laboratory. Therefore, *F. candida* is frequently used in toxicity tests, and the International Organization for Standardization (ISO) published a protocol in which they standardized the use of this springtail for soil bioassays (ISO, 1999). Very recently, this test has also been validated and adopted by the Organization for Economic Co-operation and Development (OECD, 2009), which emphasizes the high relevance of this assay. The principle of this test is to expose *F. candida* for 28 days to soil containing (gradually increasing concentrations of) a toxic compound, and measure the decrease of reproduction (number of offspring). Subsequently, the above mentioned EC10 and EC50 values can be estimated. Besides ecotoxicological research, *F. candida* has also been a model organism studying the physiology of drought acclimation and cold tolerance (Bayley et al., 2001), predator-prey population dynamics (Read et al., 2006), and phylogenetic analysis (Timmermans et al., 2008). Collembola, which are hexapods, appear take a unique position within the pancrustacean group, between crustaceans such as *Daphnia magna* and other hexapods such as insects (Timmermans et al., 2008). A part of *F. candida*’s transcriptome has recently been elucidated and sequence information is now available for approximately 6,000 unique transcripts (Timmermans et al., 2007), which opens new possibilities for genomic research.
Genomics, a discipline within the field of molecular biology, is often defined as the science that studies the genome (all genes in an organism), at all possible levels (DNA, transcript, protein, and metabolite). However, occasionally the term “-omics” is used, to keep the term genomics reserved for analysis of genomes at DNA level only. Genomic research at the transcript level is more specifically referred to as transcriptomics. The transcriptome is defined as the complete set of transcripts in a certain tissue or organism, at a particular time/condition. Currently, the DNA microarray technology is the most popular tool used in transcriptomics (however, new methods such as next-generation sequencing are emerging; Metzker, 2005). The microarray technology has evolved from Southern blotting, which is a technique in molecular biology used to verify that a specific DNA fragment (target), visualized with electrophoresis, is truly the DNA fragment of interest, using a complementary DNA probe (Southern, 1975). In Southern blotting, first the target is transferred to and linked on a nitrocellulose or positively charged nylon membrane, which then is subjected to a (sequence specific) radio-labeled complementary probe. In case the target and the probe are complementary, and hence hybridize, radioactivity can be measured on the membrane which then confirms that the target DNA fragment has a sequence similar to the probe. A variation of Southern blotting is northern blotting, which enables you to confirm specific RNA (instead
of DNA) fragments. The microarrays used in transcriptomics are to be seen as multiplex northern blotting experiments. However, rather than linking the target to a surface, the probes are covalently linked (immobilized) to a surface (a membrane, or more often to a glass slide in case of microarrays). On a glass slide thousands of microscopic spots (features) can be printed, each containing a probe with unique sequence, e.g., complementary to different genes or transcripts. Furthermore with the microarray technology, target is labeled instead of probe, usually with fluorescent dyes (e.g., Cy3 and Cy5). Therefore with microarrays, transcript abundance of thousands of genes in parallel can be measured on just one glass slide (sometimes referred to as DNA chip). By comparing different transcriptomes, e.g., one derived from treated cells, and one derived from untreated (control) cells, genes can be identified that are altered (up- or downregulated) in response to the treatment. Since whole transcriptomes are evaluated, this technique is very helpful in elucidating unknown regulatory mechanisms or molecular pathways.

Schena et al. (1995) was the first to publish a gene expression study using microarrays printed on glass slides and used transcripts from the model plant Arabidopsis thaliana. Soon thereafter, microarrays became very popular in the fields of oncology and pharmacology, and eventually found its way into ecology and evolutionary biology. The term ecotoxicogenomics was proposed by Snape et al. (2004) to describe the integration of genomic tools into ecotoxicology. The main goal of ecotoxicogenomics is to elucidate toxic modes of action and mechanisms of environmental pollutants in ecologically relevant species, which can supplement traditional ecotoxicological endpoints. Many aquatic species have been subjected to ecotoxicogenomic research using microarrays, such as fish (Moens et al., 2006; van der Ven et al., 2005), molluscs (Venier et al., 2006), and crustaceans (Poynton et al., 2007). However, for soil-living animals, until now only worms have been used, e.g., Eisenia fetida (Gong et al., 2007), and Lumbricus rubellus (Owen et al., 2008). The use of collembolan species in ecotoxicogenomic research could therefore provide a broader spectrum of ecological targets in the soil ecosystem and elucidate common or species-specific mechanisms.

**Aim of the thesis**

The main purpose of this thesis is to explore the possibilities of the application of the microarray technology in soil ecotoxicology using springtails. The springtail F. candida is a very easy and fast culturable organism for soil quality assessment, and also very sensitive to many pollutants, and therefore the ideal candidate for ecotoxicogenomic research. Gene
expression, almost without exception, is altered in response to toxicant exposure (Nuwaysir et al., 1999). Changes in the transcriptome are one of the first (direct or indirect) responses in organisms to stress, and therefore one of the earliest signs to be detected. Environmental stress (e.g., heat, xenobiotics) usually causes intracellular damage (e.g., denatured proteins, damaged membranes), which elicits intracellular signals that evoke the induction of certain transcripts (Roelofs et al., in press). Furthermore, previous studies (e.g., Feder and Hofmann, 1999; Girardot et al., 2004) provide evidence for a causal link between the early onset of expression of stress response genes and toxic effects of contaminants at the physiological/organismal level. Genomic tools in soil ecotoxicology have been suggested to improve traditional endpoints (e.g., survival, and reproduction) by being: (i) faster, (ii) more sensitive, and (iii) compound specific (van Straalen and Roelofs, 2008). The standardized soil toxicity protocol (ISO, 1999), and the available transcriptome sequence information are used as starting point for the research described in this thesis.

Outline of the thesis

In chapter 2 the microarray technology was used to measure the expression of more than 5,000 transcripts of F. candida, in response to heavy metal containing soil. Different exposure time points (2, 4, and 7 days) were assessed of the non-essential metal cadmium. The concentration of cadmium (57.9 mg kg⁻¹ dry soil) that was used represented the EC50 for reproduction after 28 days of exposure, and therefore the identified altered genes and mechanisms after 2, 4 and 7 days of exposure inform us about the early responses of F. candida to metal toxicity.

Chapter 3 describes F. candida’s transcriptional response to PAHs in soil after 2 days of exposure, using the compound phenanthrene as a model for PAHs. Phenanthrene was assessed at two different concentrations (24.95 and 45.80 mg kg⁻¹ dry soil; EC10 and EC50, respectively), to investigate the effect of different concentrations as well. In this way we are able to study dose-dependent regulation of genes, which may implicate that gene expression can potentially be validated as biomarker for soil quality.

A very different type of stress was investigated in chapter 4, namely heat. With the anticipated climate change, temperature might become a crucial factor for proper soil functioning, and therefore more understanding of the impact of heat on ecologically relevant soil-living species such as springtails is needed. The transcriptome of springtails exposed to heated soil (30°C) was compared with that of springtails exposed to soil under normal
conditions (20°C). The exposure to the elevated temperature was 30 minutes, and becomes lethal after approximately 3 days.

Chapter 5 reports on a diagnostic method of the microarray technology: classifier analysis. In this chapter gene expression profiles are generated from exposures to six different metals in soil (and clean reference soil). Using the uncorrelated shrunken centroid (USC) algorithm (Yeung and Bumgarner, 2003), a gene set (classifier) was selected which was most discriminative between the soils containing one of the six metals (and clean reference soil). This derived classifier was then used to predict the metal content of independent soil samples.

Finally, in chapter 6 the results and findings of the preceding chapters are discussed and suggestions for further research are made.