CHAPTER VI

General Discussion
1. General discussion

In a world with a growing population and increasing demand for energy resources, a healthy soil must be considered a very important foundation to contribute with a large variety of resources for future generations sustainability. However, inherent to a growing population, soil production activities are also expanding and there is a consequent increasing input of contaminants into the environment, with pesticides being a major contamination input in agricultural soils (Carvalho, 2017; Eugenio et al., 2018). Such increasing pollution calls for improved assessment techniques and instruments, which must be sensitive, specific and provide early detections, aiming for a better and faster evaluation of the risks of contamination, monitor soil quality, and keep it sustainable. Also, the development of such tools should be driven towards involving relevant biological species, as they may be considered indicators of important ecological soil functions. Soil invertebrates play important functions in the soil environment, being among the most important contamination indicators in soil ecotoxicology. Earthworms and springtails are the most widely used species in ecotoxicological studies, for which there are standardized protocols available.

Here, by selecting a representative soil arthropod, *Folsomia candida*, it was primarily intended to evaluate the suitability of omics tools (transcriptomics and proteomics) to assist current methodologies implemented for soil ecotoxicological assessments. These molecular tools proved to be specific for the tested chemicals. Extensive relevant information could be compiled, concerning the cascade of events triggered by the exposures during time, from the sub-cellular to the population level, increasing the knowledge on the interaction between pollutants and non-target organisms (mechanisms of toxic action). Two pesticides (an herbicide and a fungicide) with distinct activities and widely used for agricultural purposes were studied here as contaminating subjects and were characterized in chapter II. To bring more realism and ecological relevance to this work, an agricultural natural soil was used for all evaluations, laboratory and then in the field, under natural environmental conditions. The different characteristics in soil (degradation rates) and effects of both pollutants in the survival and reproduction of the non-target organism *F. candida* were addressed and discussed in chapter II. Here a particular observation on the herbicide (glyphosate) formulation stood out, in the sense that the effects in *F. candida* were not caused by the active substance, and probably must be attributed to other formulating substances present in the formulation instead.

*Chapters III and IV* constituted a major focus of this thesis, where gene expressions and correspondent protein levels were correlated and integrated with effects at a higher level of
biological organization. Both gene and protein expression demonstrated to be good early warning indicators for the tested pesticides with, however, distinct pathways of toxic action, presenting also a different cascade of events that led to an impaired development and reproduction of \textit{F. candida}. For the glyphosate formulation, the majority of the molecular events associated with exposure were more consistently related to effects described in the literature for the used adjuvant POEA, despite the relative scarce toxicological information about this surfactant, as demonstrated in chapter III. In the case of the formulated fungicide (chlorothalonil), the cascade of molecular and cellular events that led to the observed effects on reproduction, development, and survival of \textit{F. candida} agreed with previous toxicological reports for the fungicide, being also concordant with its reported mode of action, as discussed in chapter IV. It must also be acknowledged that different responses to the formulations could be observed over time, at both gene and protein levels, thus highlighting the importance of following different time-points for a more accurate assessment of toxicity mechanisms.

Interestingly, a time-shift correlation between gene expression and correspondent protein levels was verified for the fungicide formulation but no such evidence was observed for the herbicide formulation, where better correlations were observed for datasets from the same time-points. Despite such observations, it is important to clarify that the majority of data for glyphosate are related to 10 days after exposure, where major responses at gene and protein levels were observed. There are known regulatory biological mechanisms (mentioned in chapter IV) that may decouple protein from mRNA levels and need to be considered, as it was observed by Fortelny et al. (2017) in a critical work about protein level predictions from gene levels, but also argued in other studies (Edfors et al., 2016; Liu et al., 2016). The present thesis demonstrated how such mechanisms may be differentially affected, leading to different gene-to-protein synchronizations through time. Including time as a variable in a gene/protein correlation analysis was firstly equated by Gedeon and Bokes (2012), although in a different perspective (apart from a toxicological context) and using minimal time scales and a prokaryotic model system. By considering post translational modification and elongation time in protein synthesis, the authors found better correlations between gene and protein levels. This important variable was considered in the present thesis and for a better clarification of results obtained here, an illustration of two distinct theoretical models explaining gene and protein correlations in organisms exposed to contaminants over time are here proposed and illustrated in Figure 6.1. The gene/protein effects caused by exposure of \textit{F. candida} to glyphosate formulation may be represented by a non-delayed relation model (Figure 6.1A) and the effects
in correlations between gene and protein levels caused by chlorothalonil formulation were demonstrated to be more related to a delayed expression model (Figure 6.1B).

![Figure 6.1 Illustration of two distinct theoretical models by which contaminants can affect gene/protein interactions. A - non-delayed relation model; B - delayed relation model. x, y, z - consecutive sampling time-points; t - time.](image)

Both models here presented assume better correlations with higher differential expression levels and consider chronic effects caused by contaminants. The first theoretical model (6.1A) is characterized by less pronounced effects. At an initial stage of contamination there are few differentially expressed genes or proteins to correlate (x) but with increasing exposure time, effects become more evident, triggering response mechanisms. In this case, however, organisms are still able to keep the same level of synchronization between transcription and translation mechanisms, which are observed by significant correlations between datasets (y). In such model, it is expected that RNA changes will translate faster to the differential protein levels observed due to lower post-translational regulatory events. At a final stage, organisms lower their differential expression patterns (z), but possibly keeping expression of important molecules, at a post-exposure new adaptive steady-state level. This was the case observed by Ravaschiere et al. (2017) for a particular heat shock protein (Hsp70), where basal levels of the protein in the mussel *Geukensia demissa* were demonstrated to be adapted to a contaminated environment. Heat shock proteins are good candidate molecules to adapt and maintain their expression levels, since they are a group of highly conserved proteins with major physiological roles in protein homeostasis (Nguyen et al., 2016). Model A may be adopted in situations where there is a cellular stimulation caused by the contaminant (e.g immune response), but no extensive oxidative stress damage in cells is observed, therefore maintaining a stable modulation in gene-to-protein metabolism (Jovanovic et al., 2015).
The second theoretical model (6.1B) is defined by a pronounced response from the beginning of exposure, firstly observed at transcriptional level \((x)\) and later at protein levels. There is a delay between gene to protein expressions, caused by exposure to the pollutant. In addition, a time-shift window along exposure may be considered, where higher levels of differential expression are observed and may present better biological correlations \((y)\). This shift may however adopt different amplitudes along the exposure period, with a trend to increase with time due to increasing adverse effects caused. For this model, a recovery to basal levels of expression is less likely to occur, at least for the protein levels \((z)\). These estimations are based on evidence showed in a simpler but elucidative biological system, where the authors also reported a delayed response in protein levels compared to differential gene expression, after inducing misfolding stress (Cheng et al., 2016).

Therefore, the exposure to glyphosate formulation fits better the first model (A), with low initial expression levels for both datasets, which increased only after 10 days of exposure. There are indications that organisms possibly recovered overall gene expression to a basal level, since the big majority of previous differentially expressed genes \((y)\) were down-regulated, and at the end of the exposure there were no effects on survival of \(F.\) candida. As for the exposure to the fungicide formulation, a delayed relation model (B) is better fitted. Severe effects were observed since the beginning of the exposure, by early triggering extensive stress and detoxification responses at the transcription level \((x)\). The exposure caused a delay in translation which was verified along the entire exposure period. A differential expression could be observed for both datasets, despite not correlated within the same time-point due to the delay resulting from the effects caused by the contaminant \((y)\). Good correlations could be observed between consecutive time-points during the exposure to the fungicide. At a final stage, an eventual return to a basal stage was less likely to exist \((z)\). This observation is also mentioned by Cheng et al. (2016), where many protein ratios increased without reverting back to their levels along time.

There are two major factors that may possibly link the pesticides tested to the models proposed in Figure 6.1, which need further clarifications: 1) exposure concentration effects; and 2) mode of action of the pesticides.

1) The correlation between transcriptional and translational effects caused by the glyphosate formulation may possibly be attributed to the lower overall effects of the herbicide, as it only caused an effective inhibition in reproduction of about 25% (chapter III), while the fungicide caused a reproductive reduction of around 75%, accompanied by 26% survival effects.
(chapter IV). Moreover, for the herbicide formulation, major effects were observed only after 10 days of exposure, while for the fungicide formulation effects were registered after 2 days. Since the effects observed at a higher level (reproduction and survival), caused by the pesticides, deviated from the initial EC50 estimations (50% inhibition in reproduction), the deviation observed for the fungicide formulation (higher concentration and stronger effects) might therefore be associated to the delayed translation observed.

2) The biological pathways being affected by the two pesticides are clearly distinct, due to their reported modes of toxic action. The way RNA and protein regulation determines homeostasis at a post-exposure stage depends on the nature of treatment and its implications on the fitness of the cells (Cheng et al., 2016). Most likely, this could help to explain the differences observed for the gene/protein correlations for the two pesticide formulations. It has been demonstrated that overall relationships between mRNA and protein abundance may vary, depending on gene categories (Morimoto and Yahara, 2018). For instance, the authors reported highly similar time series correlations between stress related genes and proteins, although using single-celled yeast. Similarly, in the present work, the herbicide formulation affected the organisms through changes in normal cellular respiration and lipid metabolism, with stress-related molecules presenting the higher correlations. This is indicative of faster protein synthesis mechanisms, when triggered by a stress response. Moreover, since effects were less pronounced in *F. candida* for this formulation, probably the organisms could still be able to focus energetic expenses on triggering stress response mechanisms more rapidly, inhibiting energy-expensive processes such as moulting, as observed. For the fungicide (non-systemic) formulation, it was an entirely different scenario, since it caused much more conspicuous and generalized effects in *F. candida*, firmly related to the mechanistic ways of toxic action described for the active substance, chlorothalonil (Elskus, 2012). With crucial energetic pathways such as glycolysis and citric acid cycles, protein metabolism and therefore general energy homeostasis being affected, possibly the organisms canalized energy to cope with the adverse events caused by the fungicide, loosing therefore efficiency in overall cellular transcription, translation, and post-transcriptional mechanisms, resulting in a gap between general gene and correspondent protein levels. It is also possible that these mechanisms could be directly targeted by the reactive oxygen species (ROS) metabolites produced upon exposure to chlorothalonil, generated during an initial detoxification process, as response to the fungicide. Although research on RNA oxidative damage has only gained attention more recently and needs further investigation, ROS metabolites have been demonstrated to be implicated in ribosome activity (Willi et al., 2018).
These authors found that a folded structure and protein interactions do not protect ribosomal function from these reactive species. They ultimately cause ribosomal RNA damage by oxidative stress (tested in vivo), which conclusively lead to impaired protein biosynthesis, as demonstrated by the authors (tested in vitro). In the present research there is enough evidence that the severe damages inflicted by the fungicide probably relate to the temporal gap observed between gene expression changes and the corresponding changes in protein levels. The early observed molecular evidences of inhibition in reproduction and developmental cycles are very corroborative of the premature injurious effects in *F. candida*, that later translated in an overall reduced reproduction and survival (chapter IV).

In *chapter V*, the validation of the laboratory findings, now under real environmental conditions, exposing *F. candida* in the field was performed, only with the fungicide formulation. The observed effects for the herbicide formulation under laboratory conditions could not be attributed to the active substance glyphosate, which caused no effects on survival or reproduction in *F. candida*, even at the higher concentrations tested. Moreover, the exact composition of the surfactant POEA in the formulation could not be determined, as it is discussed in *chapter III*. It is also mentioned in *chapter III* that after being demonstrated that POEA was more toxic to non-plant organisms than glyphosate itself, the surfactant was banned from glyphosate-based herbicide formulations in 2016 and POEA-containing herbicides were no longer available in the market. By considering the above-mentioned factors, a decision was made to exclude the herbicide formulation from the field exposure validation.

The information provided in *chapter V* confirms that most of the tested genes are specific for the effects associated with the fungicide’s mode of action and were not modulated by the extra environmental factors. A real field scenario presents much more variability compared to laboratorial environment, where conditions are controlled. In the field, subjects may experience several biotic and abiotic constraints that may reflect in biased results (Calisi and Bentley, 2009), especially when dealing with sensitive endpoints such as gene expression. As an example, by exposing honey bees to sub-lethal doses of imidacloprid (neonicotinoid insecticide) under laboratory and field conditions, De Smet et al. (2017) found discrepant inductions in a set of targeted genes related to immunity and detoxification. Therefore, given that in *chapter V* the same differential expression responses were obtained in the field for six of the relevant tested genes, with a trend to increase response with increased fungicide concentrations, this makes them optimal early warning indicators to be included in a fast detection tool-kit to be used in
the field under chlorothalonil-based exposures, helping to predict hazardous effects of the fungicide in the environment.

1.1 A basis for Adverse Outcome Pathway integration

The importance of the Adverse Outcome Pathway (AOP) conceptual framework was previously mentioned in chapter I. It consists in a structured representation of measurable mechanistic biological events, from molecular to higher biological levels (population/community), usually following some key points. It is elaborated with a starting Molecular Initiating Event (MIE) that describes an action on a specific biomolecule, through a series of causally-linked Key Events (KE) representing downstream effects at the levels of molecular, cellular, tissue, organ, individual, and population-level response, leading to some Adverse Outcome (AO) (Carusi et al., 2018). It thus serves as a knowledge assembly tool to facilitate the transparent translation of mechanistic data into outcomes which are meaningful to regulatory chemical safety assessment (Carusi et al., 2018).

In this research, the results obtained for chlorothalonil provided enough extensive biological evidence, from the initial exposure to the biological effects observed in *F. candida*. This information provides the scientific basis to link this particular compound to an adverse outcome pathway that is not yet available in the public repository AOPWiki (https://aopwiki.org/) but will be highly relevant from a risk assessment perspective. The adverse effects observed for chlorothalonil are presented and discussed in chapter IV. Briefly, this compound exerts its effects by binding to glutathione and other thiol-rich molecules. Data showed that by binding to these molecules, this compound’s mode of action firstly involves activation of detoxification, excretion mechanisms, and normal protein processing. With these processes being insufficient to fully compensate the toxicity, chlorothalonil eventually causes unfolded protein response, ROS accumulation and posterior oxidative stress defence, DNA damage, and abnormal cell proliferation in detoxifying and excretory tissues. Ultimately, it led to an impairment in moulting cycle, and also reproduction effects and reduced survival of *F. candida*. As chlorothalonil is a chlorinated chemical (four chlorine covalent bonds), there are other chlorinated stressors with AOPs available in AOPWiki (Table 6.1), to which chlorothalonil effects may be related. The comparison between the data gathered for chlorothalonil in the present thesis and data collected in Table 6.1 provided the possible basis for a future AOP repository integration for this chemical. The details for the basis of a particular adverse outcome pathway of events for chlorothalonil are also proposed in Table 6.1.
Table 6.1  Adverse outcome pathways developed (or under development) for seven chlorinated chemicals and three fungicides, deposited in the international repository AOPWiki (https://aopwiki.org/). TFHA – Task Force for Hazard Assessment, EAGMST – Extended Advisory Group on Molecular Screening and Toxicogenomics, WNT - Working Group of the National Coordinators of the Test Guideline Programme. The events leading to an AOP inclusion are also suggested for the chemical chlorothalonil.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Molecular Initiating Event (MIE)</th>
<th>Key Events (KE)</th>
<th>Adverse Outcome (AO)</th>
<th>AOP Number</th>
<th>OECD Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chloroform</td>
<td>Event:783 - Increase, Cytotoxicity (tubular epithelial cells)</td>
<td>Event:710 - Increase, Regenerative cell proliferation (tubular epithelial cells)</td>
<td>Event:713 - Increase, Adenomas/carcinomas (renal tubular)</td>
<td>AOP 116</td>
<td>Under development</td>
</tr>
<tr>
<td>Carbon tetrachloride</td>
<td>Event:244 - Protein Alkylation</td>
<td>Event:55 - N/A, Cell injury/death Event:1492 - Tissue resident cell activation Event:68 - Accumulation, Collagen</td>
<td>Event:344 - N/A, Liver fibrosis</td>
<td>AOP 38</td>
<td>TFHA/WNT Endorsed</td>
</tr>
<tr>
<td><strong>Fungicides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maneb Manez Zineb</td>
<td>Event:1462 - Thiol group of chemicals interact with sulfhydryl groups of proteins to form thiol adducts</td>
<td>Event:1464 - Reduction of collagen crosslinking Event:1466 - Notochord distortion or malformations</td>
<td>Event:1467 - Growth, reduction</td>
<td>AOP 242</td>
<td>Under development</td>
</tr>
<tr>
<td>Chlorothalonil</td>
<td>Events:1487; 1462</td>
<td>Events:710; 783; 784; 854; 1088; 1392; 1115; 1512</td>
<td>Events:351; 713; 856; 1467</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The mechanistic ways of toxic action of chlorothalonil resembles reported effects for other chlorinated fungicides (Long and Siegel, 1975). Although the represented pathways in Table 6.1 are mainly concluded from studies with vertebrates, there are striking similarities between the cascade effects verified for these selected chemicals and the observed effects on *F. candida* in the present study. The AOPs described for chloroform and diclofenac sodium seem to be the most closely related to chlorothalonil, with exception to their molecular initiating events (MIE). For chlorothalonil, the initiating event seems to be more related to the other presented chemicals, by binding to thiol/seleno-proteins, and subsequently causing cytotoxic effects. The represented key events in Table 6.1 also reflect relevant similarities to chlorothalonil, which are related to increased detoxifying and excretory tissues injuries, ROS and oxidative stress, unfolded protein response, immune responses or hyperplasia. All these effects are reported in *chapter IV*. The adverse outcomes discussed for exposed *F. candida* can also be observed here. Reduction in growth and survival, and carcinogenic effects are also discussed and suggested in *chapter IV*.

Belonging to the same class of pesticides as chlorothalonil, the MIE observed for the selected fungicides in Table 6.1 (Maneb, Mancozeb and Zineb) may be related to the organochlorine pesticide, despite their distinct chemical composition, with dissimilar reactive groups. It has been documented that chlorothalonil interacts not only with thiol groups, but also with other sulfhydryl groups in molecules, although to a lesser extent (Baier-Anderson and Anderson, 2000; Long and Siegel, 1975). It is therefore possible that chlorothalonil may simulate some of the adverse effects of well-characterized chemicals, as the ones represented in Table 6.1, affecting normal development in *F. candida* afterwards, as documented in this thesis. The information provided in Table 6.1 for chemicals with adverse pathways of effects and physico-chemical properties related to chlorothalonil and its cascade effects observed in the present thesis, present therefore a comparative scientific support to predict the effects of the fungicide and its further integration in an AOP framework.

By working with an organism with a relatively short life-cycle and collecting data at different time-points after exposure, a vast amount of data was gathered in the present work to integrate in an AOP framework, from the MIE to the adverse outcomes seen at a higher level of biological organization. Despite most of these frameworks are still being developed and similar work needs to be developed for other species, there is promising evidence here to integrate the stressor chlorothalonil in the AOP system.
1.2 Perspectives

In brief, this thesis evaluated the suitability of gene and protein expression to be used in ecotoxicological assessments, by linking these tools to traditional ecotoxicological analysis supported by international guidelines. The ultimate goal would be to integrate molecular analysis into risk assessment trials, as they are sensitive, specific and early warning tools. It was possible to validate results in the field, providing a set of molecular markers to be used in the field, under similar exposure scenarios. With these observations, the reliability of such tools could also be demonstrated, at least for chlorothalonil homolog formulations. It was also proved here that a time-series approach is very important in this type of experimental designs, since there is often a gap between transcription, translation, and even metabolic regulation, particularly when working with eukaryotic organisms (Yugi and Kuroda, 2018). Most of similar studies do not include a time-frame in their evaluations and this may be the main reason why low relationships between datasets are often presented.

Timing is indeed very important for these sensitive markers and minor differences in timing of the sampling are prone to cause large differences in the responding gene or protein composition. Also, it is often found that a specific set of genes responding to a treatment are also very specific for the conditions to which subjects are exposed. These often neglected factors may result in different differentially expressed genes when an experiment is repeated, leading them to be discarded as good sensitive tools. Despite this, potential biases were in part considered in the present thesis for the field experiment, by theoretically selecting a chlorothalonil-specific set of genes that maintained their differential expression patterns constant (up- or down-regulation) for all the sampling points under laboratory exposure. However, there are additional considerations that may be addressed as next steps to evaluate potential interferences of relevant factors in the field. One way to tackle this, also adding ecological relevance to the work, could be through an evaluation of different soil types. As discussed in chapter II, soils from different regions present different characteristics, such as moisture, pH, organic matter or microbiota that are prone to affect the activity of pollutants (and consequently their effects in *F. candida*) or affect the organisms directly, possibly triggering different responsive mechanisms. Additional environmental factors such as temperature, salinity, or even biotic interactions could be accounted to be tested under laboratory conditions. By knowing the “natural variation” of the omics profiles, we would be able to deploy them more wisely and accurately in the field.
Regarding the relations observed here between gene and protein expressions for the different contaminant formulations tested, and to elucidate some of the possibilities discussed, it would be interesting to assess if these effects are particularly related to the contaminants, their concentrations, or if both may influence gaps between the levels of omics datasets. This could be addressed with, for instance, more targeted analyses, using different sets of contaminant concentrations. Further similar investigations are also recommended for other classes of contaminants to evaluate if they follow the same or different patterns as the proposed here between gene expression and correspondent protein levels for glyphosate and chlorothalonil formulations. Regarding the herbicide formulation containing glyphosate as active substance, it was shown that the observed effects were most probably due to the surfactant POEA and not caused by the active substance (chapter III). Although this chemical was recently banned from most of these herbicide formulations, it is still widely used by the industry and it would be pertinent to perform several exposure tests with different classes of this chemical, evaluating the effects at different biological levels. In chapter V, only two concentrations were used to validate the laboratorial results under realistic field conditions. Although the results proved to be very consistent, there is the need for a more consistent dose-response effect, as similar responses should be addressed at lower doses and a contamination bottom-line should be established for risk assessment purposes. A set of specific markers retrieved from chapter IV could be used to validate such achievements.

Technologies like gene silencing or genetic transformation could also be used in future experiments to corroborate the evidence effects caused by the pollutants for particular doubtful genes. There is a particularly recent study demonstrating this, where the authors resorted to gene silencing of a tomato NADPH oxidase gene (RBOH1) to evaluate its involvement in chlorothalonil metabolism in tomato crops (Hou et al., 2018). Moreover, such technologies have already been implemented in horticulture industry to improve crop production (Choudhary et al., 2018).

As final remark, with the high content that molecular-level datasets are able to reveal, a more comprehensive information on the environmental fate and effects of chemicals can be gathered and integrated into decisive frameworks to better define the term sustainability, concerning the environment and human health. This work plays a significant basal contribution in such direction, by providing encouraging results and also critical considerations to be followed, in the development of innovative environmental “omics” approaches.
References


