Research Article

Analysis of Multiple Toxicants of Biological Events in *Folsomia* Candida and *Caenorhabditis Elegans*

Cai W[1#], Cui H[2,3,4#], Kong Q[2,3*] and Yang X[2,3*]

[1]Department of Laboratory Medicine, Wen Zhou Traditional Chinese Medicine Hospital, China

[2]School of Laboratory Medicine and Life Sciences, Wenzhou Medical University, Wenzhou, Zhejiang 7 Province China

[3]Key Laboratory of Laboratory Medicine, Ministry of Education, Wenzhou Medical University, Wenzhou, 9 Zhejiang Province China

[4]College of Life Science, Northeast Forestry University, Harbin, P.R. China

[5]College of Life Science, Northeast Agricultural University, Harbin, P.R. China

*Corresponding author: Xu Yang, Key Laboratory of Laboratory Medicine, Ministry of Education, Wenzhou Medical University, Wenzhou, 9 Zhejiang Province China

Abstract

Ecotoxic effects of soil pollutions are attracting attentions owing to the potential toxicity. 41 Generally, the soil toxicants mainly include metals, polycyclic aromatic hydrocarbons, natural toxins and 42 drug residues. Transcriptomic analysis has been demonstrated to detect the toxicity induced by single 43 compounds. However, changes in biological events responding to multiple toxicants still remain unclear. In 44 this study, we performed bioinformatic analysis of gene expression in *Folsomia candida* exposed to 45 cadmium, phenanthrene, 2-phenylethyl isothiocyanate and diclofenac, and compared cadmium-induced 46 toxicological effects on *Folsomia candida* and *Caenorhabditis elegans*. The results indicate cadmium can 47 induce the most Differentially Expressed Genes (DEGs) and variations of biological events in *Folsomia* 48 *candida* and shown the Gene Ontology (GO) terms significantly enriched in metabolic pathways 49 responding to multiple toxicants. And more DEGs and GO terms induced by cadmium were observed in 50 *Folsomia candida*, rather than *Caenorhabditis elegans*. Overall, our study suggests the expression profile 51 of *Folsomia candida* is more suitable to evaluate toxicological effects of chemical toxicants, and the 52 selected DEGs, GO terms and metabolic pathways may be used as early biomarkers for monitoring of soil 53 toxicants.

Keywords: *Folsomia candida*; *Caenorhabditis elegans*; Bioinformatic analysis; Transcriptomics; Biomarkers

Introduction

Environmental pollution is a worldwide issue with increasing concerns, and the adverse effects of soil toxicants could be immense to the terrestrial ecosystem [1]. Diverse anthropogenic means can result in the accumulation of inorganic and organic toxicants in soil [2]. Metals have been identified as the largest 63 class of soil toxicants [3], wherein cadmium belongs to one of the ubiquitous toxic compounds. Polycyclic Aromatic Hydrocarbons (PAHs), derived from anthropogenic activities in the microenvironment [1], are extremely hazardous to the soil ecosystem, and some of them induce severe carcinogenicity [5]. Phenanthrene, belonging to PAHs, is mostly released into the environment as a by-product of organic material combustion [6]. The most poisonous and widespread phytotoxin Isothiocyanate (ITC) comes from the possible products of glucosinolates hydrolysed by myrosinase in damaged tissues [7]. In recent decades, non-steroidal anti-inflammatory drugs have increased to a higher level in soil owing to the global consumption of pharmaceuticals and personal care products, in which diclofenac seems to have the highest acute toxicity (FENT et al., 2006) and commonly be detected in many waste treatment biosolids [8]. It has been shown that soil toxicants may interfere with ecological functions by decreasing the reproduction of organisms. Currently, bioassays based on indicator species have been widely used to identify the ecotoxicological consequences [9].

Arthropods of the order Collembola (springtails) are among the most abundant soil-dwelling invertebrates with a significant contribution to decomposition and nutrient mineralization in the soil ecosystem [10]. Moreover, the parthenogenetic *Folsomia candida* (*F. candida*) is feasible to conduct the standardized detections of soil toxicity, and internationally accepted guidelines have been developed for assessing the ecotoxicity of chemicals in this species. Additionally, nematodes are widely distributed and species-rich, occupying important positions in soil detritus food web [11]. Given the characteristic of sensitivity to toxicants, *Caenorhabditis elegans* (*C. elegans*) is suited as a bioassay organism for both acute and chronic toxicity tests [12]. With the advent of genomic techniques, microarrays are now integrated into the toxicological analysis of toxicants [13,14]. Previous studies in toxicogenomics have shown environmental quality could indeed be diagnosed by gene expression profiling of indicator species [15,16]. Over the last ten years, the molecular toxicity of single toxicant in spiked or naturally polluted soil samples, had been explored by analysing the transcriptional events of *F. candida* [17,18] or *C. elegan* [19,20]. However, in fact, various complex toxicants exist in soils, and mechanistic studies on mixture toxicity are helpful to understand how chemicals interact. And, multivariate analysis, in different receptor organisms, may supply more informative end points to characterize the nature hazard of adverse disturbances [21]. To date, little is known about the functional molecular terms mediated by multiple toxicants in soil, and the extrapolation of ecological risks from one non-target species to another phylogenetically distinct species still remains unclear. In this study, we took advantage of the transcriptional raw datasets of *F. candida* exposed to cadmium [17,22], phenanthrene

[17], 2-phenylethyl ITC [23] and diclofenac [24] -spiked soil, to analyze the toxic effects of four toxicants on biological processes and molecular functions, as well as the metabolic pathways. Further, we compared the transcriptomic responses of *F. candida* and *C. elegans* under the cadmium exposure [5]. Overall, we shown the common and specific Differentially Expressed Genes (DEGs) and the biological events induced by the selected toxicants, and provided insights into the toxicity pathways affected significantly in the indicator species tested.

Results

DEGs and GO Terms in *F. Candida* Induced By Single Toxicant

Statistical analysis revealed that the differential expression of 964 genes (496 up-regulated and 106 468 down-regulated) to cadmium exposure, 251 genes (122 up-regulated and 129 down-regulated) to 107 phenanthrene exposure, 108 genes (64 up-regulated and 44 down-regulated) to 2-phenylethyl ITC exposure, 108 and 98 genes (79 up-regulated and 19 down-regulated) to diclofenac exposure, respectively (Figure 1a). Moreover, all DEGs of each toxicant were performed GO analysis. We found the biological processes significantly affected by cadmium were "cellular amide metabolic process" (GO: 0043603), "response to hypoxia" (GO: 0001666) and "response to abiotic stimulus" (GO: 0009628), and the significant molecular functions included "N-(5-amino-5-carboxypentanoyl)-L-cysteinyl-D-valine synthase activity" (GO: 0050564) and "metallopeptidase activity" (GO: 0008237). In response to phenanthrene exposure, the biological processes affected significantly included "heterocycle catabolic process" (GO: 0046700), "organic hydroxy compound metabolic process" (GO: 1901615) and "multicellular organismal homeostasis" (GO: 0048871), and the significant molecular functions included "monooxygenase activity" (GO: 0004497) and "retinol dehydrogenase activity" (GO: 0004745). The biological processes significantly responding to 2-phenylethyl ITC contained "DNA modification" (GO: 0006304), "regulation of protein complex disassembly" (GO: 0043244) and "positive regulation of hemopoiesis" (GO: 1903708), and the significant molecular functions included "chitinase activity" (GO: 0008061) and "hexosaminidase activity" (GO: 0015929). As for the effects of diclofenac exposure, the significant biological processes included "embryonic skeletal system morphogenesis" (GO: 0048704), "localization" (GO: 0051179) and "cell-cell signaling" (GO: 0007267), and the significant molecular functions were "interleukin-1 binding" (GO: 0019966) and "substrate-specific channel activity" (GO: 0022838). Overall, in terms of the number of DEGs and GO terms, the most serious effect on *F. candida* is caused by cadmium exposure.

Common DEGs and GO Terms in *F. Candida* Induced By Multiple Toxicants

The numbers of DEGs commonly induced by multiple toxicants are indicated in the overlapping regions of Venn diagram (Figure 1b). Two DEGs, Fcc00057 and Fcc06073, were significantly up regulated in response to the four toxicants (Table 1). Fcc00057 was related to three important biological processes, namely "antibiotic metabolic process" (GO: 0016999), "antibiotic biosynthetic process" (GO: 0017000) and "drug metabolic process" (GO: 0017144). This gene was also involved in two important molecular functions, i.e., "isopenicillin-N synthase activity" (GO: 0016216) and "vitamin

binding" (GO: 0019842). A total of 22 DEGs were responded to three toxicants, and 17 ones exhibited the changed expressions with the same direction (up- or down-regulated). In the treatments of cadmium, phenanthrene or 2-phenylethyl 136 ITC, 9 common genes were differentially expressed significantly, and they were involved in "oxidoreductase activity" (GO: 0046992), "glutathione peroxidase activity" (GO: 0004602) and "carboxylic acid binding" (GO: 0031406) (S1 Table). In response to cadmium, phenanthrene or diclofenac, 14 common genes were differentially expressed significantly. They were relative to the significant biological processes "regulation of defense response" (GO: 0031347), "regulation of dopamine secretion" (GO: 0014059) and "neuronal action potential" (GO: 0019228), as well as the significant molecular functions, like "adenylate cyclase binding" (GO: 0008179), "toxic substance binding" (GO: 0015643), etc.

The biological events behind the DEGs common to two toxicants were revealed by corresponding GO terms. We found 112 DEGs were commonly responded to cadmium and phenanthrene, and some of them were involved in "carbohydrate catabolic process" (GO: 0016052), "oxidation-reduction process" (GO: 0055114) and "catecholamine metabolic process" (GO: 0006584). Additionally, their significant molecular functions included "L-ascorbic acid binding" (GO: 0031418) and "monosaccharide binding" (GO: 0048029). A total of 37 DEGs were shared by cadmium and 2-phenylethyl ITC, and many DEGs were related to the significant biological processes "disaccharide metabolic process" (GO: 0005984), "allantoin catabolic process" (GO: 0000256) and "respiratory system process" (GO: 0003016), together with the significant molecular functions "hydrolase activity" (GO: 0004553) and "allantoinase activity" (GO: 0004038). In the 47 DEGs responded to both cadmium and diclofenac, many of them were involved in the significant biological processes "synaptic transmission" (GO: 0007268), "protein homooligomerization" (GO: 0051260) and "generation of ovulation cycle rhythm" (GO: 0060112), and the significant molecular functions "extracellular ligand-gated ion channel activity" (GO: 0005230) and "cytokine binding" (GO: 0019955). In terms of phenanthrene and 2-phenylethyl ITC, we observed 17 common genes were differentially expressed significantly, and they were relative to several significant biological processes (e.g. "establishment of sister chromatid cohesion" (GO: 0034085), "DNA alkylation" (GO: 0006305), "regulation of growth" (GO: 0040008)) and molecular functions (e.g. "testosterone 6-beta-hydroxylase activity" (GO: 0050649), "antioxidant activity" (GO: 0016209)). As for the 20 DEGs common to phenanthrene and diclofenac, and we found they were related to some biological processes and molecular functions, such as "regulation of response to stress" (GO: 0080134), "hemicellulose metabolic process" (GO: 0010410), "negative regulation of defense response" (GO: 0031348), "channel regulator activity" (GO: 0016247) and "acetylcholine binding" (GO: 0042166). It was worth noting that the GO terms of DEGs shared by two toxicants are also associated with reproduction, development, metabolism, homeostasis etc. Furthermore, the immune-related biological processes and molecular functions are easily affected by all the four toxicants and may be used as promising early biomarkers for detecting soil quality.

Specific DEGs and GO Terms in *F. Candida* Induced by Single Toxicant

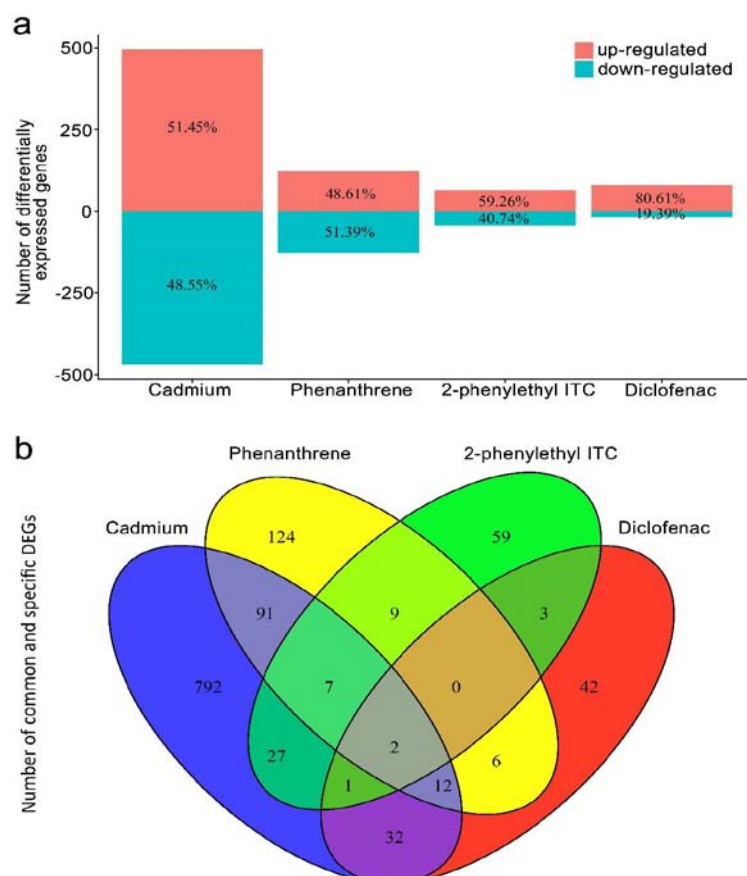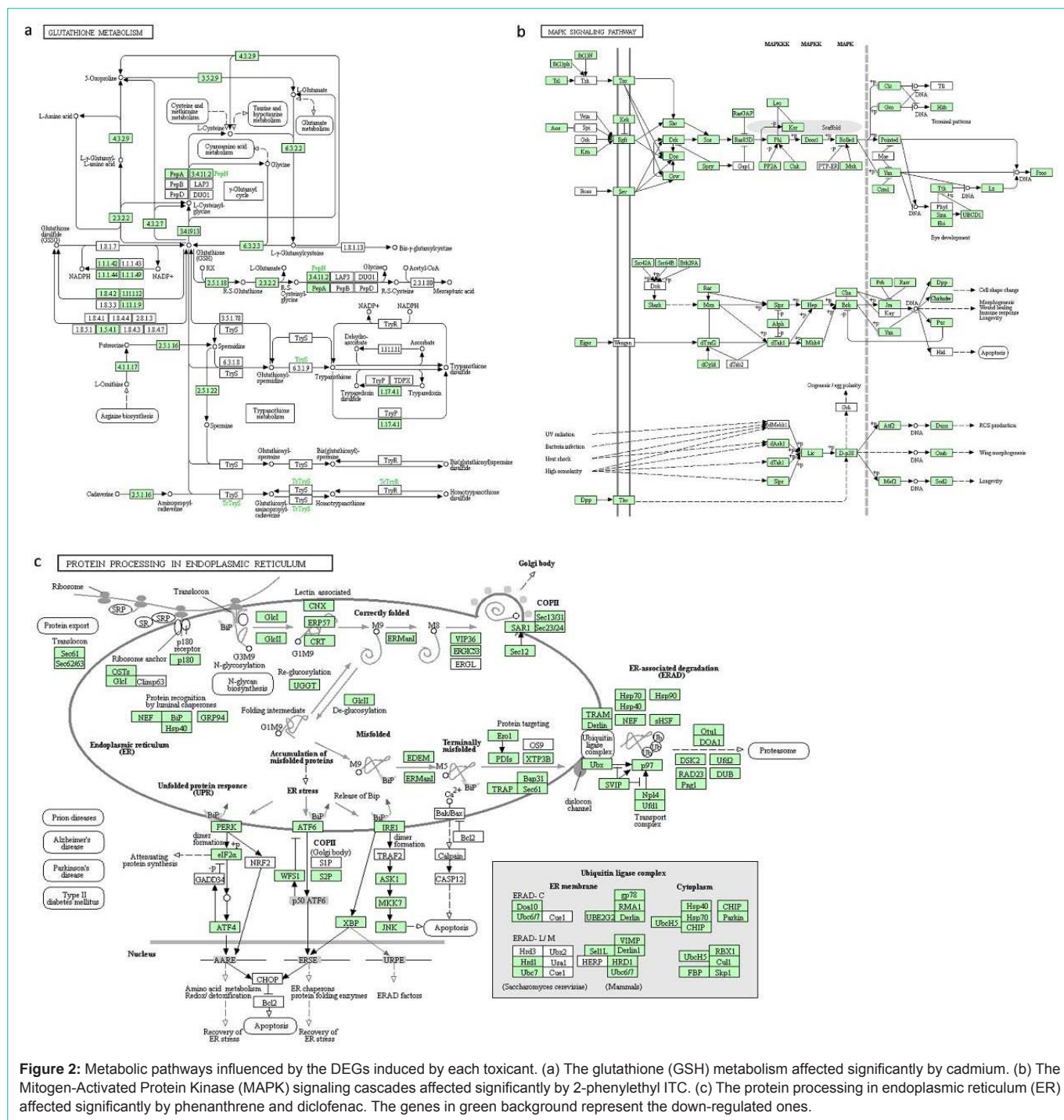As shown in the Venn diagram (Figure 1b), the specific DEGs

Figure 1: All DEGs of *Folsomia candida* induced by cadmium, phenanthrene, 2-phenylethyl ITC, and diclofenac treatment. (a) The numbers of DEGs relative to the control ($p < 0.05$) are shown as a result of the toxicological effect under single toxicant exposure. Their percentages of up- and down-regulated genes are indicated on the bars. (b) Venn diagram showing the number of genes that were significantly different from the other treatments ($p < 0.05$). The numbers in the overlapping regions represent those genes commonly changed to multiple toxicants, and the numbers in non-overlapping regions represent those genes unique to the indicated treatment.

were observed specifically by cadmium, phenanthrene, 2-phenylethyl ITC, or diclofenac. 792 DEGs were specifically responded to cadmium, and the specific biological processes included "arachidonic acid secretion" (GO: 0050482), "lipid localization" (GO: 0010876) and "response to transforming growth factor beta" (GO: 0071559). In addition, the specific molecular functions of 792 DEGs contained "ACP phosphopantetheine attachment site binding" (GO: 0044620) and "exopeptidase activity" (GO: 0008238). 124 DEGs were specifically induced by phenanthrene, and the specific biological processes included "response to alkaloid" (GO: 0043279), "DNA duplex unwinding" (GO: 0032508) and "vitellogenesis" (GO: 0007296). Additionally, the specific molecular functions of the 124 DEGs contained "pyridoxamine-phosphate oxidase activity" (GO: 0004733) and "monooxygenase activity" (GO: 0004497) (S2 Table). 59 DEGs were specifically responded to 2-phenylethyl ITC, and the specific biological processes included "regulation of cell killing" (GO: 0031341), "GTP metabolic process" (GO: 0046039) and "response to pH" (GO: 0009268). And the specific molecular functions of the 59 DEGs contained "acetyl-CoA carboxylase activity" (GO: 0003989) and "purine-nucleoside phosphorylase activity" (GO: 0004731) (S3 Table). 42 DEGs were specifically induced by diclofenac, and the specific biological processes included "oxidative phosphorylation"

(GO: 0006119), "chromatin silencing at centromere" (GO: 0030702) and "negative regulation of leukocyte proliferation" (GO: 0070664). The specific molecular functions of the 42 DEGs contained "NADH dehydrogenase (ubiquinone) activity" (GO: 0008137) and "morphogen activity" (GO: 0016015) (S4 Table). Altogether, our results show the four toxicants could induce their own specific biological events, which may be used to indicate the specific toxicant in bioassay.

Significantly Enriched Metabolic Pathways by Single Toxicant Induced-DEGs in *F. Candida*

All DEGs of *F. candida* induced by the single toxicant were blasted with KEGG database, and the metabolic pathways enriched significantly were obtained respectively. Our results shown that cadmium toxicant negatively regulated the glutathione (GSH) metabolism (Figure 2a), and the down-regulated genes were mainly associated with the gamma-glutamyl cycle, the transformation between GSH and Glutathione Disulfide (GSSG), and the conjugation reactions with exogenous chemicals. For example, the genes encoding glutamate-cysteine ligase catalytic subunit [EC:6.3.2.2], glutathione synthase [EC:6.3.2.3] and glutathione hydrolase [EC:3.4.19.13] were inhibited in the gamma-Glutamyl cycle. During the 6 transformation from GSH to GSSG, the genes encoding glutathione peroxidase

Figure 2: Metabolic pathways influenced by the DEGs induced by each toxicant. (a) The glutathione (GSH) metabolism affected significantly by cadmium. (b) The Mitogen-Activated Protein Kinase (MAPK) signaling cascades affected significantly by 2-phenylethyl ITC. (c) The protein processing in endoplasmic reticulum (ER) affected significantly by phenanthrene and diclofenac. The genes in green background represent the down-regulated ones.

[EC:1.11.1.9] and protein-disulfide reductase (GSH) [EC:1.8.4.2] were down-regulated. In the conjugation reactions with exogenous chemicals, the genes encoding glutathione S-transferase [EC:2.5.1.18] and glutamyl transpeptidase [EC:2.3.2.2] were also down-regulated. KEGG pathway enrichment analysis of all DEGs induced by 2-phenylethyl ITC shown, this toxicant significantly affected the mitogen-activated protein kinase (MAPK) signaling cascades (Figure 2b). Many down-regulated genes were mainly involved in the extracellular signal-regulated kinase (ERK) 1/2 signaling (e.g.

Rolled and *Ras85D*), c-Jun N-terminal kinase (JNK) signaling (e.g. *Slpr* and *Bsk*) and p38 kinase (e.g. *Lic* and *Tkv*). We further found that both phenanthrene and diclofenac negatively regulated the protein processing in Endoplasmic Reticulum (ER) (Figure 2c). Many down-regulated genes were mainly involved in the formation of coat protein complex II (COP II), ER-Associated Degradation (ERAD) and Unfolded Protein Response (UPR). And, the important genes encoding Sec61 pore and GTP-binding protein SAR1, associated with newly synthesized peptides entering ER and Golgi apparatus,
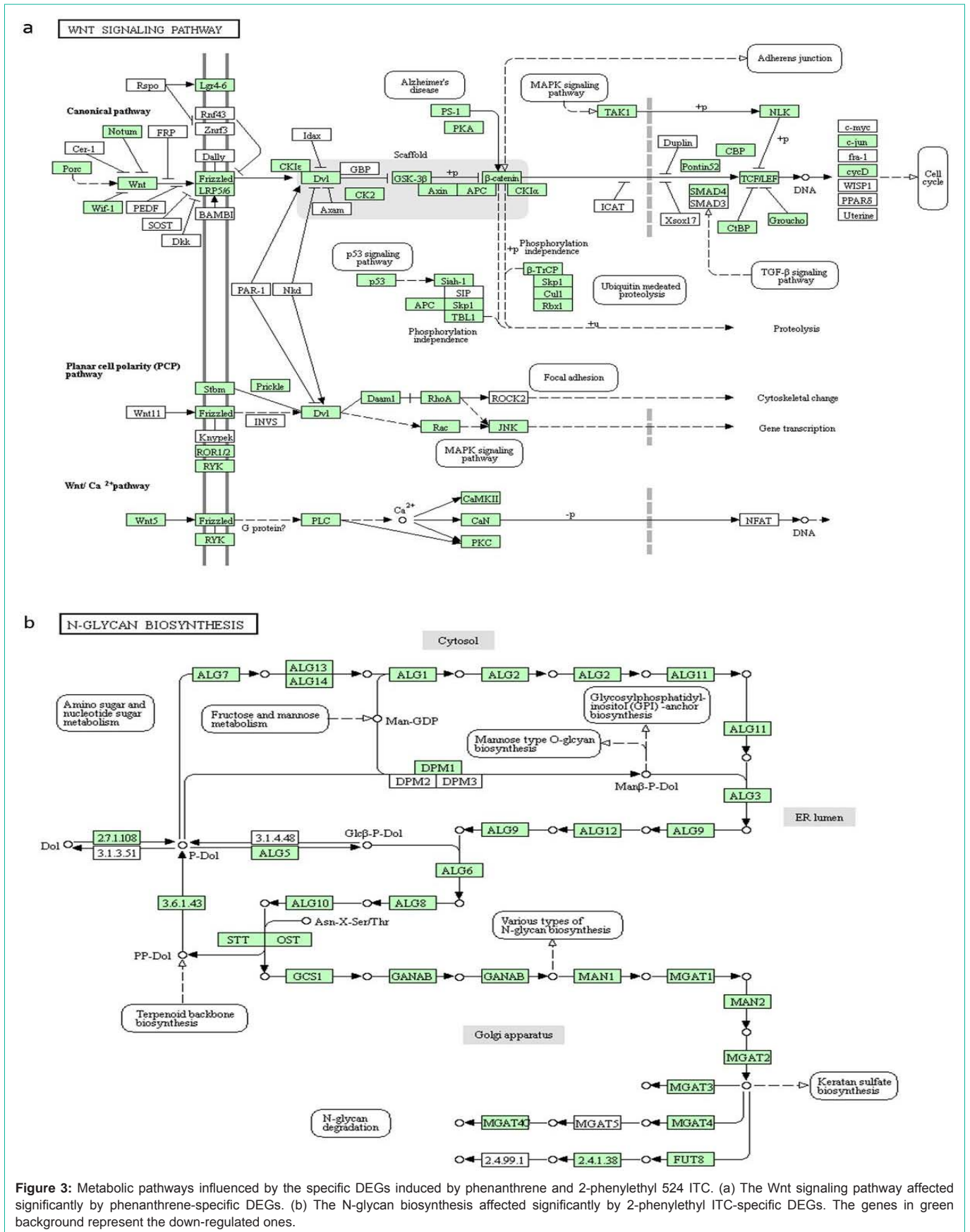
Figure 3: Metabolic pathways influenced by the specific DEGs induced by phenanthrene and 2-phenylethyl 524 ITC. (a) The Wnt signaling pathway affected significantly by phenanthrene-specific DEGs. (b) The N-glycan biosynthesis affected significantly by 2-phenylethyl ITC-specific DEGs. The genes in green background represent the down-regulated ones.
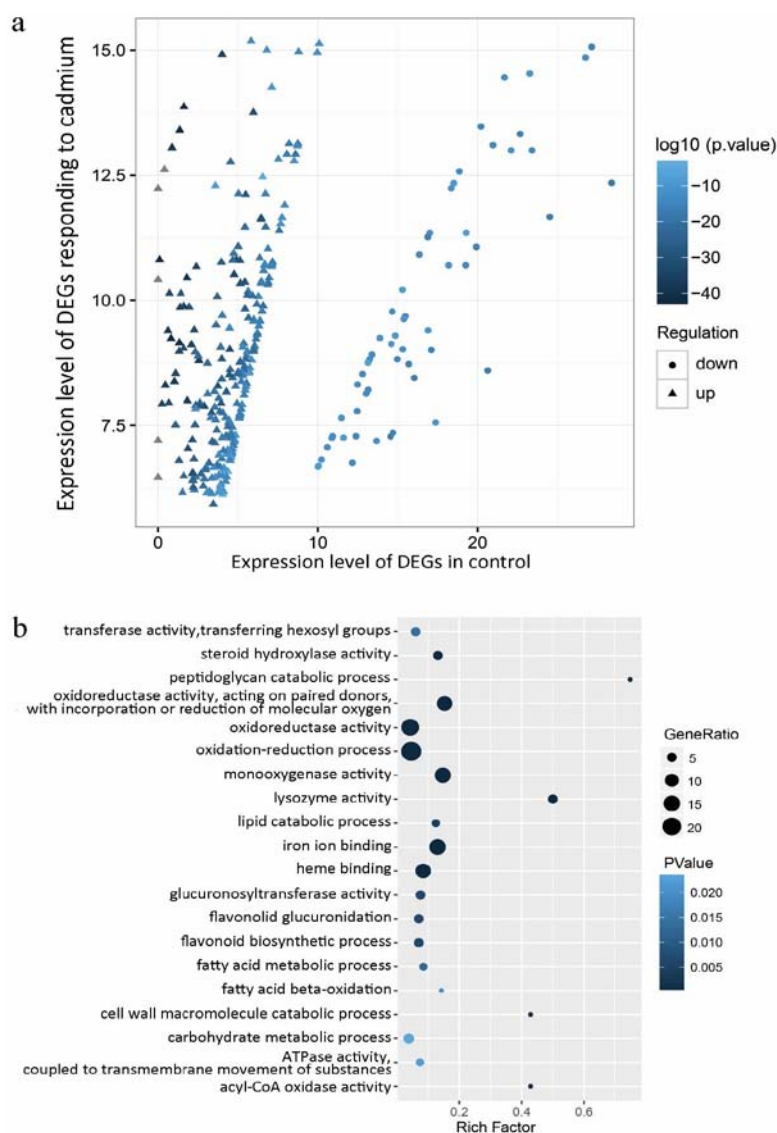
Figure 4: DEGs and GO terms of *C. elegans* induced by cadmium. (a) The expression levels of DEGs are indicated when compared with the control. The distribution of up- and down-regulated genes can be clearly observed. (b) Significant biological processes and molecular functions were displayed (top 20, ranked with p-value). The number of genes affected is represented by the size of circle and the significance of biological events is represented by the color of circle.

were also down-regulated. In addition, the genes related to molecular chaperones and ubiquitin ligase complex, such as *BiP*, *HSP40* and *Doa10*, were inhibited during the process of binding misfolded proteins and ERAD. In ER stress, the expression levels of key genes, like *PERK* and *eIF2-alpha,* were also decreased.

Furthermore, the specific DEGs of *F. candida* caused by single toxicant were also annotated using KEGG database. Results showed that phenanthrene-specific DEGs significantly affected the Wnt signaling pathway (Figure 3a), and many down-regulated genes were associated with the canonical pathway, Planar Cell Polarity (PCP) pathway and Wnt/Ca2+ pathway. In the canonical pathway, the genes encoding glycogen synthase kinase 3 beta (*GSK-3beta*) [EC:2.7.11.26], beta-catenin and TCF (T-cell factor) family transcription factors were down-regulated. The genes in PCP pathway related to small GTPases RhoA (RAS homologue gene-family member A) and RAC1

(Ras-related C3 botulinum toxin substrate 1) also were suppressed. In Wnt/Ca2+ pathway, the genes encoding phosphatidylinositol phospholipase C (PLC) [EC:3.1.4.11] and calcium/calmodulin-dependent protein kinase (CaM K II) [EC:2.7.11.17] were inhibited. Through KEGG analysis of 2-phenylethyl ITC -specific DEGs, we found the N-glycan biosynthesis was significantly affected (Figure 3b) and the down-regulated genes were mainly involved in the N-glycan precursor biosynthesis and glycosylation. The key gene encoding dolichyl-phosphate beta-glucosyltransferase [EC:2.4.1.117] (i.e. *ALG5*) was negatively regulated in the biosynthesis of N-glycan precursor. During sequential addition of monosaccharides, many genes associated with ALG glycosyltransferases and Oligo Saccharyl Transferase (OST) complex were inhibited. We also found the genes encoding mannosyl-oligosaccharide glucosidase [EC:3.2.1.106] (i.e., *GCS1*) and alpha-mannosidase 231 II [EC:3.2.1.114] (i.e., *MAN2*) were down-regulated during the period of precursor's trimming,

Table 1: Significant DEGs shared by the multiple toxicants.

Gene Name	EC50 cadmium		EC50 phenanthrene		EC50 2-phenylethyl ITC		EC50 diclofenac	
	Log2 Fold change	adj.P.val	Log2 Fold change	adj.P.val	Log2 Fold change	adj.P.val	Log2 Fold change	adj.P.val
Fcc00057	5.827	0	1.471	0.004	0.856	0.05	1.039	0.012
Fcc06073	3.742	0	1.26	0.015	0.717	0.044	1.679	0.016
Fcc00494	1.44	0	0.593	0.009	0.83	0.01		
Fcc00908	-1.84	0.002	-1.235	0.029	-0.852	0.022		
Fcc01630	2.112	0	1.025	0.013	0.738	0.045		
Fcc01912	-2.155	0.003	-1.344	0	0.818	0.01		
Fcc02341	-1.291	0.018	-1.303	0.007	-1.062	0.033		
Fcc02980	2.821	0	0.614	0.01	0.721	0.007		
Fcc03384	-0.604	0.048	0.75	0.002	-1.381	0.026		
Fcc02580	-1.685	0			-0.498	0.05	-0.63	0.045
Fcc00272	3.866	0	0.985	0.001			-2.096	0.025
Fcc00343	2.309	0	2.309	0			2.337	0.027
Fcc01429	1.892	0.036	1.311	0.039			4.058	0
Fcc01688	-5.042	0	-1.582	0.015			-1.327	0.016
Fcc02703	-1.204	0.005	-0.992	0.012			-0.933	0.03
Fcc04207	5.347	0	2.134	0.001			2.586	0.001
Fcc04297	1.046	0.016	-1.219	0.017			1.046	0.016
Fcc04350	-4.777	0	-0.969	0.017			-1.223	0.024
Fcc05034	0.484	0.048	0.732	0.013			1.349	0
Fcc05311	1.607	0.001	1.676	0.002			2.313	0.001
Fcc05719	2.787	0	0.748	0.018			1.525	0.001
Fcc05739	1.306	0.001	-0.84	0.029			1.813	0.001

extension and modification. Altogether, we suggest the specifically affected metabolic pathways should also contribute to the effective detection of specific toxicant based on the indicator species.

Common and Specific Biological Events Caused By Cadmium between *F. Candida* and *C. Elegans*

To further explore the toxic effects of cadmium on cross-species, we compared the biological processes and molecular functions affected by cadmium in *F. candida* and *C. elegans*. We firstly analyzed the DEGs and GO terms of *C. elegans* induced by cadmium. As for the 309 significant DEGs, 249 DEGs (80.6%) were up-regulated and 60 DEGs (19.4%) were down-regulated (Figure 4a). Regarding the GO 240 terms of DEGs in *C. elegans*, 11 biological processes and 14 molecular functions affected significantly by cadmium were identified (S5 Table; S6 Table). The top 20 representatives (ranked with *p*-value) were shown in (Figure 4b). The most significant biological process and molecular function were "oxidation-reduction process" (GO: 0055114) and "iron ion binding" (GO: 0005506), respectively. Furthermore, we found the biological process "carbohydrate metabolic process" (GO: 0005975) and the molecular function "ATPase activity coupled to transmembrane movement of substances" (GO: 0042626) were commonly influenced in *F. candida* and *C. elegans*. Monooxygenase activity was also influenced by cadmium in both *F. candida* (GO: 0004500) and *C. elegans* (GO: 0004497). Moreover, lipid metabolism, indicated by "lipid localization" (GO: 0010876) and "lipid transport" (GO: 0006869) in *F. candida* and "lipid catabolic

process" (GO: 0016042) in *C. elegans,* might be also co-affected. On the other hand, many biological processes and molecular functions are exclusive to each species. Here, some representatives of them in *F. candida* induced by cadmium were described, such as "acid secretion" (GO: 0046717), "ion transport" (GO: 0006811), "circulatory system process" (GO: 0003013), "metallopeptidase activity" (GO: 0008237), "peptide binding" (GO: 0042277) and "glucosidase activity" (GO: 0015926). And the specific biological processes in *C. elegans* responding to cadmium contained "flavonoid biosynthetic process" (GO: 0009813) and "flavonoid glucuronidation" (GO: 0052696), and its specific molecular functions included "lysozyme activity" (GO: 0003796), "steroid hydroxylase activity" (GO: 0008395), "structural constituent of cuticle" (GO: 0042302), etc. In fact, more specific DEGs and biological events induced by cadmium were observed in *F. candida,* indicating that it is a better indicator species to evaluate cadmium toxicity.

Discussion

In the study, we analyzed the DEGs, GO terms and metabolic pathways of *F. candida*, which was respectively exposed to cadmium, phenanthrene, 2-phenylethyl ITC or diclofenac for two days, with the same Effect Concentration (EC50) in LUFA 2.2 soil, and evaluated the toxic effects of cadmium on *F. candida* and *C. elegans* in similar experimental conditions. Some common and specific DEGs, the corresponding GO terms and significant enriched metabolic pathways in response to the toxicants were identified in *F. Candida*.

In addition, more DEGs and GO terms induced by cadmium were observed in *F. candida*, rather than *C. elegans*. Our results indicate *F. candida* is a more sensitive indicator species to monitor soil toxicants, and the screening genes and significantly affected biological events will be helpful to evaluate the molecular toxicity of chemicals in soil.

By comparing the DEGs and GO terms induced by cadmium, phenanthrene, 2-phenylethyl ITC 272 or diclofenac, we found cadmium can cause the most DEGs and most significant changes in biological events in *F. candida*. For example, we found the molecular function "metallopeptidase activity" (GO: 0008237), which plays an important role on the digestion and absorption of food proteins in the digestive tract of insect [25], was badly influenced. So we considered the toxic effects of cadmium on soil ecosystem should be more severe. By analyzing the DEGs and GO terms commonly affected by the four toxicants, we found the gene Fcc00057 was expressed differentially, and it was associated with the biological process "antibiotic biosynthetic process" (GO: 0017000) and the molecular function "isopenicillin-N synthase activity" (GO: 0016216). That is consistent with a previous study which demonstrated that Fcc00057 was involved in immune responses [22]. Thus, this gene may be used as one universal predictor of future toxin exposures. The specific DEGs and GO terms induced by one specific toxicant were also identified. Cadmium could specifically affect "arachidonic acid secretion" (GO: 0050482), which functions in oxidative stress [26]. Phenanthrene could specifically influence "monooxygenase activity" (GO: 0004497), which is involved in Cytochrome P450 enzymes and phase I of the biotransformation and detoxification [27]. In consistent with the previous findings, we found the "acetyl-CoA carboxylase activity" (GO: 0003989) was specifically affected by 2-phenylethyl ITC, suggesting it is able to induce metabolic processes [23]. The "NADH dehydrogenase (ubiquinone) activity" (GO: 0008137), critical for catalytic 288 mechanism of cell respiration [28], was specifically influenced by diclofenac. Overall, our results show the stress-specific response profiles of *F. candida* may be useful to effectively detect soil toxicants and evaluate their ecotoxicity on the molecular level.

The enriched metabolic pathways affected by the toxicant tested were also analyzed in F. candida. We found that the biosynthesis and transformation of GSH, which functions during the detoxification and antioxidant, were negatively regulated by cadmium [29]. This observation indicates that cadmium should cause an adverse effect on the antioxidant and immunoregulatory mechanisms. MAPK signaling cascades, which participate in the regulation of inflammatory reactions, were down-regulated by 2-phenylethyl ITC [30], suggesting that some inflammatory factors are induced to cope with the bad environmental stress. The DEGs induced by both phenanthrene and diclofenac were mainly involved in the protein processing in ER. It is likely that phenanthrene and diclofenac could disturb the survival of *F. candida* by interfering with the germination of vesicles [31]. We also found the phenanthrene-specific DEGs were associated with Wnt signaling pathway, indicating the remodelling of cytoskeleton and cell adhesion may be influenced [32]. In addition, the biosynthesis of N-glycan, which belongs to one of major glycoproteins in eukaryotes [33], was negatively regulated by 2-phenylethyl ITC -specific DEGs. This observation shows 2-phenylethyl ITC could decrease the process of glucose metabolism, and further affect the production and distribution of energy in *F. candida*. Altogether, our results provide insights into the metabolic pathways in response to each toxicant. We believe the specifically affected metabolic pathways will contribute to detect the specific toxicant in the future bioassay.

The indicator species play an extremely important role in the bioassay of environmental pollutants. *F. candida* has been frequently applied in the standardized detections of soil toxicity, while *C. elegans* also was reported as a model of toxin effects to detect chemical toxicants. In this study, the effects of cadmium on *F. candida* and *C. elegans* were compared. We observed the biological events affected commonly in *F. candida* and *C. elegans*, namely "carbohydrate metabolic process", "ATPase activity coupled to transmembrane movement of substances" and "monooxygenase activity", suggesting that detoxification and energy consuming processes could be invoked by cadmium in both species, in order to survive under the stressful conditions [23]. We also identified many DEGs and GO terms with specific changes in *F. candida* or *C. elegans*. Compared with *C. elegans*, more specific DEGs and GO terms were revealed in *F. candida*, suggesting it is a more sensitive indicator species to test chemical toxicants in soil.

In this study, we show the common and specific DEGs of *F. candida*, the GO terms and the enriched metabolic pathways in response to the selected toxicants. We also provide the same and different biological events caused by cadmium in *F. candida* and *C. elegans.* In conclusion, the results demonstrate that *F. candida* is a more sensitive indicator species to detect soil toxicants, and the screening genes commonly and specifically responded to the chemical toxicants, together with the significantly affected biological events, should enable to facilitate the diagnosis of their toxicological impacts on the terrestrial ecosystem.

Methods

Analysis of DEGs

Microarray datasets of gene expression from accession numbers for *F. Candida* exposure to cadmium (GSE11122), phenanthrene (GSE14207), 2-phenylethyl ITC (GSE29239), diclofenac (GSE59589), and *C. elegans* exposure to cadmium (GSE7535), were accessed from the National Center for Biotechnology Information. Under the exposure condition of each toxicant, the degree of gene differential expression was measured with mean $\log_2$ expression ratios and the corresponding *p*-values (BH corrected *p < 0.05).*

Gene Ontology (GO) Analysis of DEGs

To understand the molecular biological events associated with DEGs, GO terms of gene clusters in *F. candida* were annotated through the top GO (0.9.6) package in the R (3.1.0) software environment [34]. GO analysis of gene sets in *C. elegans* was performed with the functional classification tool Database for Annotation, Visualization, and Integrated Discovered (DAVID) [35] using default parameters. Each list of the significant GO terms was summarized applying REVIGO with default parameters [36]. The biological processes and molecular functions affected significantly by the single or multiple toxicants, were categorized to observe the same and different features. We also compared the common and specific biological events induced by cadmium in *F. candida* and *C. elegans*. In detail, all parameters of GO analysis in each defined condition were separately counted in different Microsoft Excels.

Kyoto Encyclopedia of Genes and Genome (KEGG) Analysis

In order to unveil the metabolic pathways under toxicant treatment in *F. candida*, all DEGs and the specific DEGs induced by four individual toxicant were performed KEGG analysis [37]. To identify 349 the significantly enriched metabolic pathways, we focused on the expression levels of key regulatory genes and their associations.

Declaration of Conflicting Interests

We have no conflict of interest to declare.

Funding

References

1. Kowalska JB, Mazurek R, Gąsiorek M, Zaleski T. Pollution indices as useful tools for the comprehensive evaluation of the degree of soil contamination–A review. Environmental Geochemistry and Health. 2018; 40: 2395-2420.

2. Biruk LN, Moretton J, Iorio AFD, Weigandt C, Etcheverry J, Filippetto J, et al. Toxicity and genotoxicity assessment in sediments from the Matanza-Riachuelo river basin (Argentina) under the influence of heavy metals and organic contaminants. Ecotoxicology and environmental safety. 2017; 135: 302-311.

3. Rodrigues SM, Pereira ME, Silva EFD, Hursthouse AS, Duarte AC. A review of regulatory decisions for environmental protection: part I - challenges in the implementation of national soil policies. Environment international. 2009; 35: 202-213.

4. Olawoyin R. Application of back propagation artificial neural network prediction model for the PAH bioremediation of polluted soil. Chemosphere. 2016; 161: 145-50.

5. Baird WM, Hooven LA, Mahadevan B. Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action. Environmental and Molecular Mutagenesis. 2005; 45: 106-114.

6. Morillo E, Romero AS, Maqueda C, Madrid L, Ajmone-Marsan F, Grcman H, et al. Soil pollution by PAHs in urban soils: a comparison of three European cities. Journal of environmental monitoring : JEM. 2007; 9: 1001.

7. Halkier BA, Gershenzon J. Biology and biochemistry of glucosinolates. Annual review of plant biology. 2006; 57: 303-333.

8. Rosal R, Rodriguez A, Perdigon-Melon JA, Petre A, Garcia-Calvo E, et al. Occurrence of emerging pollutants in urban wastewater and their removal through biological treatment followed by ozonation. Water Research. 2010; 44: 578-88.

9. Rotini A, Gallo A, Parlapiano I, Berducci MT, Boni R, Tosti E, et al. Insights into the CuO nanoparticle ecotoxicity with suitable marine model species. Ecotoxicology and environmental safety. 2018; 147: 852-860.

10. Fountain MT, Hopkin SP. Folsomia candida (Collembola): a "standard" soil arthropod. Annual review of entomology. 2005; 50: 201-222.

11. Wei C, Zheng H, Li Q, Lü X, Yu Q, Zhang H, et al. Nitrogen Addition Regulates Soil Nematode Community Composition through Ammonium Suppression. PLoS ONE. 2012; 7.

12. Graves AL, Boyd WA, Williams PL. Using Transgenic Caenorhabditis elegans in Soil Toxicity Testing. Archives of Environmental Contamination and Toxicology. 2005; 48: 490-494.

13. Hou J, Liu XH, Cui BS, Bai JH, Wang XK. Microarray analysis and real-time PCR assay developed to find biomarkers for mercury-contaminated soil. Toxicology Research. 2016; 5: 1539-1547.

14. Hou J, Liu X, Wang J, Zhao S, Cui B. Microarray-based analysis of gene expression in lycopersicon esculentum seedling roots in response to cadmium, chromium, mercury, and lead. Environmental science & technology. 2015; 49: 1834-1841.

15. Hara-Yamamura H, Nakashima K, Hoque A, Miyoshi T, Kimura K, Watanabe Y, et al. Evaluation of whole wastewater effluent impacts on HepG2 using DNA microarray-based transcriptome analysis. Environmental science & technology. 2013; 47: 5425-5432.

16. Lettieri T. Recent Applications of DNA Microarray Technology to Toxicology and Ecotoxicology. Environmental Health Perspectives. 2006; 114: 4-9.

17. Nota B, Bosse M, Ylstra B, Straalen NMV, Roelofs D. Transcriptomics reveals extensive inducible biotransformation in the soil-dwelling invertebrate Folsomia candida exposed to phenanthrene. BMC Genomics. 2009; 10: 236.

18. Qiao M, Wang G, Zhang C, Roelofs D, Straalen NMV, Zhu Y. Transcriptional profiling of the soil invertebrate Folsomia candida in pentachlorophenol□ contaminated soil. Environmental Toxicology and Chemistry. 2015; 34: 1362-1368.

19. Inokuchi A, Yamamoto R, Morita F, Takumi S, Matsusaki H, Ishibashi H, et al. Effects of lithium on growth, maturation, reproduction and gene expression in the nematode Caenorhabditis elegans. Journal of Applied Toxicology. 2015; 35: 999-1006.

20. Jeong SW, Kim HL, Seo YR. Microarray analysis of global gene expression in Caenorhabditis elegans exposed to potassium dichromate. BioChip Journal. 2011; 5: 151-157.

21. Huguier P, Manier N, Chabot L, Bauda P, Pandard P. Ecotoxicological assessment of organic wastes spread on land: Towards a proposal of a suitable test battery. Ecotoxicology and environmental safety. 2015; 113: 103-111.

22. Nota B, Timmermans MJ, Franken O, Montagne-Wajer K, Marien J, et al. Gene expression analysis of collembola in cadmium containing soil. Environ Sci Technol. 2008; 42: 8152-8157.

23. Kloeke AEEV, van Gestel CAM, Styrishave B, Hansen M, Ellers J, Roelofs D. Molecular and life-411 history effects of a natural toxin on herbivorous and non-target soil arthropods. Ecotoxicology. 2012; 21: 1084-1093.

24. Chen G, Braver MWD, Gestel CAMV, Straalen NMV, Roelofs D. Ecotoxicogenomic assessment of diclofenac toxicity in soil. Environmental pollution. 2015; 199: 253-260.

25. Segundo BS, Martínez MC, Vilanova M, Cuchillo CM, Avilés FX. The severed activation segment of porcine pancreatic procarboxypeptidase A is a powerful inhibitor of the active enzyme. Isolation and characterisation of the activation peptide. Biochimica et biophysica acta. 1982; 707: 74-80.

26. Balboa MA, Balsinde J. Oxidative stress and arachidonic acid mobilization. Biochimica et biophysica acta. 2006; 1761: 385-391.

27. Guengerich FP. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. Chemical research in toxicology. 2001; 14: 611-650.

28. Feng Y, Li W, Li J, Wang J, Ge J, Xu D, et al. Structural insight into the type-II mitochondrial NADH dehydrogenases. Nature. 2012; 491: 478-482.

29. Fournier D, Bride JM, Poirie M, Bergé JB, Plapp FW. Insect glutathione S-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. The Journal of biological chemistry. 1992; 267: 1840-1845.

30. Kim EK, Choi E. Compromised MAPK signaling in human diseases: an update. Archives of Toxicology. 2015; 89: 867-882.

31. Zeuschner D, Geerts WJ, Donselaar EV, Humbel BM, Slot JW, Koster AJ, et al. Immuno-electron tomography of ER exit sites reveals the existence of free COPII-coated transport carriers. Nature Cell Biology. 2006; 8: 377-383.

32. Tzima E. Role of Small GTPases in Endothelial Cytoskeletal Dynamics and the Shear Stress Response. Circulation Research. 2006; 98: 176-185.

33. Nguema-Ona E, Vicré-Gibouin M, Gotté M, Plancot B, Lerouge P, Bardor M, et al. Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. Frontiers in Plant Science. 2014; 5.

34. Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics. 2006; 22: 1600-1607.

35. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols. 2008; 4: 44-57.

36. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLoS ONE. 2011; 6: e21800.

37. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Research. 2016; 44: D457-D462.