Supplementary Information

Common gene expression patterns in environmental model organisms exposed to engineered nanomaterials: a meta-analysis

Michael Burkard[1], Alexander Betz[1], Kristin Schirmer[1,2,3] #, Anze Zupanic[1] #

[1] Eawag, Swiss Federal Institute of Technology, Dübendorf, Switzerland

[2] ETH Zürich, Institute of Biogeochemistry and Pollutant Dynamics, Zürich, Switzerland

[3] EPF Lausanne, School of Architecture, Civil and Environmental Engineering, Lausanne, Switzerland

Corresponding author:

Eawag, Swiss Federal Institute of Aquatic Science and Technology

Überlandstrasse 133

CH-8600

Email: Kristin.Schirmer@eawag.ch

Phone: +41 (0)58 765 5266

Table of content

Table S1: Nano- and environmental related studies which were found in public data repositories; the metadata information indicates where the dataset can be found (GSE=GEO; E-MTAB= ArrayExpress) and the criteria if the dataset was selected for analysis.

ID Number Data base	Nanomaterial Type	Organism	Assay Type	Manu-facturer	Concentration & duration of exposure	Pipeline Analysis	Deselection criteria
GSE53989	Cadmium QD	A. thaliana	Microarray	Affymetrix	0, 40, 80 mg/L; 21d	No	<3 replicates
GSE24847	Carbon (MWCNTs)	C. elegans	Microarray	Affymetrix	1mg/L; 4&24h	No	<3 replicates
GSE70842	Carbon (MWCNTs)	S. lycopersicum	Microarray	Affymetrix	25 ug/ml; 1&11d	No	cDNA not available
E-MTAB-3331	Carbon (NTs)	A. thaliana	Microarray	Agilent	25 µg/mL; 48h	No	No probe sequences
GSE22803	Carbon (NTs)	S. lycopersicum	Microarray	Affymetrix	50, 100, 200 ug/ml; 10d	No	<3 replicates
GSE41486	Carbon (SWCNTs)	C. elegans	Microarray	Affymetrix	500 µg/mL;48h	No	<3 replicates
GSE57541	Carbon (SWCNTs)	S. lycopersicum	Microarray	Affymetrix	25 ug/ml; 1&11d	No	cDNA not available
E-MTAB-2454	Cerium	C. reinhardtii	Microarray	Agilent	0.14 ug/l, 10g/l	No	format not readable
GSE80461	Cerium	A. thaliana	Microarray	Affymetrix	500 mg/L; 12, 29 days	Yes	
GSE26331	Copper	E. albidus	Microarray	EBT_UA	400, 600, 800, 1000 mg/kg; 48h	No	no gene ontology
GSE54738	Copper	D. magna	Microarray	Agilent	10% immobilisation; 96 h	No	no gene ontology
GSE46958	Gold	A. thaliana	Microarray	Affymetrix	0.125 mM; 24h	Yes	
GSE32521	Gold	C. elegans	Microarray	Affymetrix	LC10 (5.9 mg/L); 24 h	Yes	
GSE30275	Gold	D. rerio	Microarray	NimbleGen	10,50 ug/mL (TMAT, MES); 24;48	No	No probe sequences
GSE41333	PAMAM	D. rerio	Microarray	Agilent	LC50 (1.8 and 2.3mg/L); 48 h	Yes	
GSE77755	Polystyrene	D. rerio	RNA-seq	Illumina	5 mg/ml; 1;3 days	No	No raw data
GSE73427	Silica	D. rerio	Microarray	customised	3 mg/mL; 24h	Yes	
GSE93182	Silica	A. thaliana	Microarray	Agilent	7,5 mg/L (C), 92,5 mg/L (S); 24h	No	<3 replicates
GSE47064	Silica	D. magna	Microarray	Agilent	1/10 LC50; 24 h	No	no gene ontology
E-MTAB-3331	Silver	A. thaliana	Microarray	Agilent	0.2 µg/mL;48h	No	No probe sequences
GSE14932	Silver	C. elegans	Microarray	Affymetrix	0.1mg/L; 24h	No	<3 replicates
GSE48677	Silver	C. reinhardtii	Microarray	customised	10, 100, 200 nM; 0.15, 1, 5, 16 h	No	no ENMs (only ions)
GSE38125	Silver	D. rerio	RNAseq	Illumina	5 µg/L; 24;48 h	No	<3 replicates
GSE70509	Silver	C. elegans	Microarray	Affymetrix	EC30 (reproduct.); 48 h	Yes	
GSE50718	Silver	D. rerio	Microarray	Agilent	20 ppB; 96 h	Yes	
GSE33488	Silver	O. latipes	Microarray	customised	0.05 mg/L; 48 h	No	no gene ontology
GSE64259	Silver	P. promelas	Microarray	Agilent	6 µg/L; 96 h	No	no gene ontology
GSE35150	Silver	D. magna	Microarray	Agilent	0.43,1, 3.5,8 ug/l; 24 h	No	no gene ontology
GSE42431	Silver	E. albidus	Microarray	Agilent	100; 200 mg/kg; 2d	No	no gene ontology
NERC unpubl.	Silver	C. elegans	Microarray	Affymetrix	EC50; 96 h	Yes	
GSE61186	Silver	D. rerio	Microarray	Agilent	0.01 mg/L,15d	Yes	
MTAB-4041	Silver	D. magna	Microarray	Agilent	1.18 µg/l, 4.70 µg/l;24h	No	no gene ontology
PRJEB7919	Silver	E. fetida	RNA-seq	HiSeq	EC50; 28 d	No	no gene ontology
GSE46520	Silver	P. promelas	Microarray	Agilent	70 µg/L; 96 h	No	no gene ontology
E-MEXP-3950	Silver	A. thaliana	Microarray	Afymetrix	5 mg/l; 10d	No	No probe sequences
E-MTAB-3331	Titanium	A. thaliana	Microarray	Agilent	20 µg/mL;48h	No	Duplicates in data
GSE80461	Titanium	A. thaliana	Microarray	Affymetrix	500 mg/L; 12;29 d	Yes	
GSE59519	Titanium	C. elegans	Microarray	Agilent	0.01, 0.1, 1 mg/L; 24h	No	No raw data
GSE59470	Titanium	C. elegans	Microarray	Agilent	0.01, 0.1, 1 mg/L; 24h	No	No raw data
GSE95603	Titanium	C. elegans	Microarray	Affymetrix	500µg/ml; n.a.	No	<3 replicates
GSE69867	Titanium	M. galloprovincialis	Microarray	customised	n.a.	No	no gene ontology
GSE77148	Zinc	D. rerio	Microarray	Agilent	LC25; 96h	Yes	
GSE54736	Zinc	D. magna	Microarray	Agilent	10% (immob.); 96 h	No	no gene ontology
GSE22051	Zinc	D. magna	Microarray	customised	1/10 LC50, LC25;24 h	No	<3 replicates
GSE49136	Zinc	H. azteca	Microarray	customised	1/10 LC50, LC25; 96 h	No	no gene ontology
GSE46947	Zinc	P. aeruginosa	Microarray	Affymetrix	1 mM; 5h	No	<3 replicates

Table S2: Studies considered for the meta-analysis.

Nanomaterial Type	Organism	Concentration & duration	Assay Type	ID Number Data base	Platform ID
Cerium	*A. thaliana*	500 mg/L (non-coated); 12, 29 d	MA;Affymetrix	GSE80461	GPL198
Titanium	*A. thaliana*	500 mg/L (non-coated); 12, 29 d	MA;Affymetrix	GSE80461	GPL198
Gold	*A. thaliana*	0.125 mM (AuO); 24h	MA;Affymetrix	GSE46958	GPL198
Gold	*C. elegans*	LC10 (5.9 mg/L)(citrate-Au); 24 h	MA;Affymetrix	GSE32521	GPL200
Silver	*C. elegans*	EC30 (reproduction); 48 h	MA;Affymetrix	GSE70509	GPL200
Silver	*C. elegans*	EC15 (Unf, PVP, PVP, MUA, AUT); 48h	MA; Agilent	NERC unpub.	GPL10094
Silica	*D. rerio*	3 mg/mL; 24h	MA;custom.	GSE73427	GPL16933
Silver	*D. rerio*	0.01 mg/L; 15 d	MA;Agilent	GSE61186	GPL14664
Zinc	*D. rerio*	LC25 (ZnO, ZnSO4); 96h	MA;Agilent	GSE77148	GPL21361
Silver	*D. rerio*	20 ppB; AgNPs (50, and 150nm); 96 h	MA;Agilent	GSE50718	GPL14664
PAMAM	*D. rerio*	LC50 (1.8 and 2.3mg/L); 48 h	MA;Agilent	GSE41333	GPL14688

Table S3: Identified GO-terms classified into main categories.

Cytoskeleton	
GO:0046854	phosphatidylinositol phosphorylation
GO:0007015	actin filament organization
GO:0030239	myofibril assembly
GO:0031109	microtubule polymerization or depolymeri...
GO:0051693	actin filament capping
GO:0031110	regulation of microtubule polymerization...
DNA metabolism	
GO:1901068	guanosine-containing compound metabolic ...
GO:0009263	deoxyribonucleotide biosynthetic process
GO:0072528	pyrimidine-containing compound biosynthe...
GO:0001522	pseudouridine synthesis
GO:0031497	chromatin assembly
GO:0046112	nucleobase biosynthetic process
GO:0009206	purine ribonucleoside triphosphate biosy...
GO:0009116	nucleoside metabolic process
GO:0006144	purine nucleobase metabolic process
GO:0019359	nicotinamide nucleotide biosynthetic pro...
DNA repair	
GO:0006281	DNA repair
GO:0006298	mismatch repair
GO:0000724	double-strand break repair via homologou...
AA metabolism	
GO:0006575	cellular modified amino acid metabolic p...

GO:0015807	L-amino acid transport
GO:0009069	serine family amino acid metabolic proce...
GO:0006568	tryptophan metabolic process
Catabolism	
GO:1901606	alpha-amino acid catabolic process
GO:0043171	peptide catabolic process
GO:0016042	lipid catabolic process
GO:0044273	sulfur compound catabolic process
GO:0044248	cellular catabolic process
GO:0044242	cellular lipid catabolic process
GO:0046434	organophosphate catabolic process
GO:0019439	aromatic compound catabolic process
GO:0046700	heterocycle catabolic process
GO:0009063	cellular amino acid catabolic process
GO:0044270	cellular nitrogen compound catabolic pro...
GO:1901361	organic cyclic compound catabolic proces...
GO:0006032	chitin catabolic process
Ion membrane transport	
GO:0071577	zinc II ion transmembrane transport
GO:0071805	potassium ion transmembrane transport
GO:0034765	regulation of ion transmembrane transpor...
GO:0070588	calcium ion transmembrane transport
GO:0010959	regulation of metal ion transport
GO:0070838	divalent metal ion transport
Energy conversion	
GO:0042775	mitochondrial ATP synthesis coupled elec...
GO:0006099	tricarboxylic acid cycle
GO:0006730	one-carbon metabolic process
GO:0033108	mitochondrial respiratory chain complex ...
GO:0072330	monocarboxylic acid biosynthetic process
GO:0017004	cytochrome complex assembly
GO:0019674	NAD metabolic process
GO:0006739	NADP metabolic process
GO:0043648	dicarboxylic acid metabolic process
GO:0043462	regulation of ATPase activity
GO:0015986	ATP synthesis coupled proton transport
GO:0046034	ATP metabolic process
GO:0099132	ATP hydrolysis coupled cation transmembr...
GO:0006084	acetyl-CoA metabolic process

GO:0006094	gluconeogenesis
GO:0006096	glycolytic process
GO:0006006	glucose metabolic process
GO:0043547	positive regulation of GTPase activity
Regulation of energy metabolism	
GO:0071901	negative regulation of protein serine/th...
Cell growth inhibition	
GO:0030308	negative regulation of cell growth
GO:0001708	cell fate specification
Translation	
GO:0000398	mRNA splicing, via spliceosome
GO:0000027	ribosomal large subunit assembly
GO:0006412	translation
GO:0032543	mitochondrial translation
GO:0042254	ribosome biogenesis
GO:0006418	tRNA aminoacylation for protein translat...
GO:0006413	translational initiation
GO:0002181	cytoplasmic translation
GO:0042273	ribosomal large subunit biogenesis
Protein Folding	
GO:0006457	protein folding
GO:0061077	chaperone-mediated protein folding
GO:0006458	'de novo' protein folding
Response to misfolded proteins	
GO:0035966	response to topologically incorrect prot...
GO:0006986	response to unfolded protein
GO:0030968	endoplasmic reticulum unfolded protein r...
Transcription	
GO:0006383	transcription from RNA polymerase III pr...
GO:0045944	positive regulation of transcription fro...
GO:0006352	DNA-templated transcription, initiation
GO:0006355	regulation of transcription, DNA-templat...
Signalling general	
GO:0030111	regulation of Wnt signaling pathway
GO:0007186	G-protein coupled receptor signaling pat...

GO:0010817	regulation of hormone levels
GO:0008593	regulation of Notch signaling pathway
GO:0007210	serotonin receptor signaling pathway
GO:0009581	detection of external stimulus
GO:0009582	detection of abiotic stimulus
GO:0050906	detection of stimulus involved in sensor...
GO:0050911	detection of chemical stimulus involved ...
GO:0043405	regulation of MAP kinase activity
GO:0043401	steroid hormone mediated signaling pathw...
O:0010469	regulation of receptor activity
Methylation	
GO:0001510	RNA methylation
GO:0032259	methylation
Defense	
GO:0050832	defense response to fungus
GO:0010038	response to metal ion
Oxidative stress	
GO:0070482	response to oxygen levels
GO:0055114	oxidation-reduction process
GO:0072593	reactive oxygen species metabolic proces...
GO:0036293	response to decreased oxygen levels
GO:0045454	cell redox homeostasis
Neuronal activity	
GO:0016079	synaptic vesicle exocytosis
GO:0007608	sensory perception of smell
GO:0007413	axonal fasciculation
GO:0007268	chemical synaptic transmission
GO:0060079	excitatory postsynaptic potential
GO:0050808	synapse organization
GO:0048666	neuron development
GO:0031175	neuron projection development
GO:0008045	motor neuron axon guidance
GO:0022008	neurogenesis
GO:0048699	generation of neurons
GO:0045664	regulation of neuron differentiation
GO:0050767	regulation of neurogenesis

Table S4: Motivation and main finding of selected studies.

Study	Main objective	Finding	Toxicological effect
GSE80461	Investigation of the mechanisms associated with plant growth and development.	Nano-ceria had more differentially expressed genes in roots and induced expression of genes related to activating transcription factors (ethylene responsive element binding protein family).	No effects visible.
GSE80461	Investigation of the mechanisms associated with plant growth and development.	Nano-titania induced more differentially expressed genes in rosette leaves and induced genes associated with photosynthesis.	Increase of the number of rosette leaves.
GSE46958	Monitorring of candidate genes involved in metal uptake and transport.	A discreet number of genes known to be involved in the transport of copper, cadmium, nickel and iron were down-regulated.	Effect on seedling root length.
GSE32521	Assessment of particle specific effects and associated biological pathways.	Several genes which are indicative for endoplasmic reticulum stress, protective mechanisms, apoptosis and necrosis were regulated.	Effect on mortality (EC10).
GSE70509	Assessment of the transcriptomic response of pristine and artifically aged particles.	No description available.	No information available.
NERC unpub.	Investigation of the modes of action of different Ag-ENPs with different physicochemical properties.	Distinctive response patterns were found for different NPs, whereby neurotoxic effects were of concern for all types. Commercially produced NPs affected other functions.	Effect on reproduction (EC15).
GSE73427	Investigation of the toxic effect and mechanisms.	Several pathways were found to be regulated including vascular smooth muscle contraction, metabolic pathways, apoptosis, MAPK signaling pathway, calcium signaling pathway and JAK-STAT signaling pathway.	No Observed Adverse Effect Level.
GSE61186	Transcriptional profiling.	No description available.	No information available.
GSE77148	Identification of expression profiles and modes of action.	Go terms related to (GPCR)-signaling pathways, cell-surface receptor-linked signal transduction and immune response were enriched.	Effect on tail malformation, pericardial edema, and yolk-sac edema (LC25).
GSE50718	Transcription of different sized AgNPs to differentiate their modes of toxicity.	Immune response, antigen processing and presentation, response to estradiol stimulus and regulation of RNA metabolic process were found as significant Go terms.	No information available.
GSE41333	Expression profiling between two different polyamidoamine dendrimers.	5% of genes were affected. Genes were involved in oxidation-reduction processes and in the nervous system development.	Effect on survival (EC10).

Table S5: Comparison of probe annotation update efficiency between gene id mapping and probe re-alignment. The percent of probes that could not be matched to up-to-date Ensembl gene annotations are shown. GPL16933 has a newer array design that uses a varying number of probes per target (mean=1.44).

| Series ID	Platform ID	Organism	Microarray Manufacturer	ID Mapping	Probe Re-alignment	Number of Probes	Probes per target
GSE80461	GPL198	*A. thaliana*	Affymetrix	5.7%	6.3%	250'206	11
GSE46958	GPL198	*A. thaliana*	Affymetrix	5.7%	6.3%	250'206	11
GSE32521	GPL200	*C. elegans*	Affymetrix	6.9%	5.9%	248'028	11
GSE70509	GPL200	*C. elegans*	Affymetrix	6.9%	5.9%	248'028	11
GSE73427	GPL16933	*D. rerio*	Affymetrix	30.6%	31.0%	1'258'175	1-190
GSE61186	GPL14664	*D. rerio*	Agilent	25.6%	19.0%	43'663	1
GSE77148	GPL21361	*D. rerio*	Agilent	26.0%	19.1%	43'659	1
GSE50718	GPL14664	*D. rerio*	Agilent	25.6%	19.0%	43'663	1
GSE41333	GPL14688	*D. rerio*	Agilent	28.0%	19.9%	45'220	1
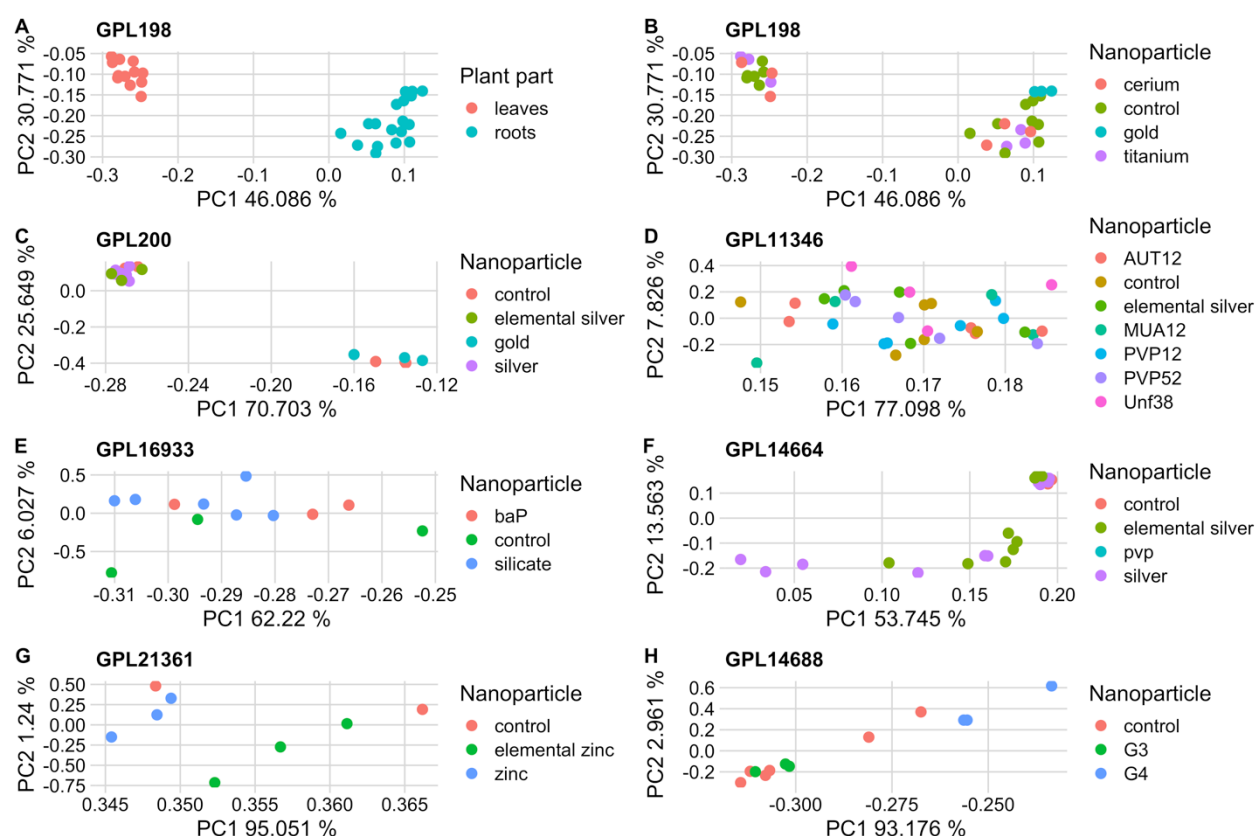
Figure S1: Principal Component Analysis (PCA)

PCA was applied to investigate the variability in gene expression between the different treatments of the model organisms *A.thaliana, C.elegans* and *D.rerio*. The 500 genes with the highest variance across all samples were used for the PCA. The distance between points (treatment samples) emphasizes the variation of the gene expression profile. *A.thaliana* samples were analysed with the platform GPL198. Points are colored with respect to the different sample sources (A) and nanoparticles (B-H). For *C.elegans*, the platforms GPL200(C) and GPL11346(D) are shown and for *D.rerio* the platforms GPL16933(E), GPL14664(F), GPL21361(G) and GPL14688(H) are shown. The axis labels show the percent variance explained.
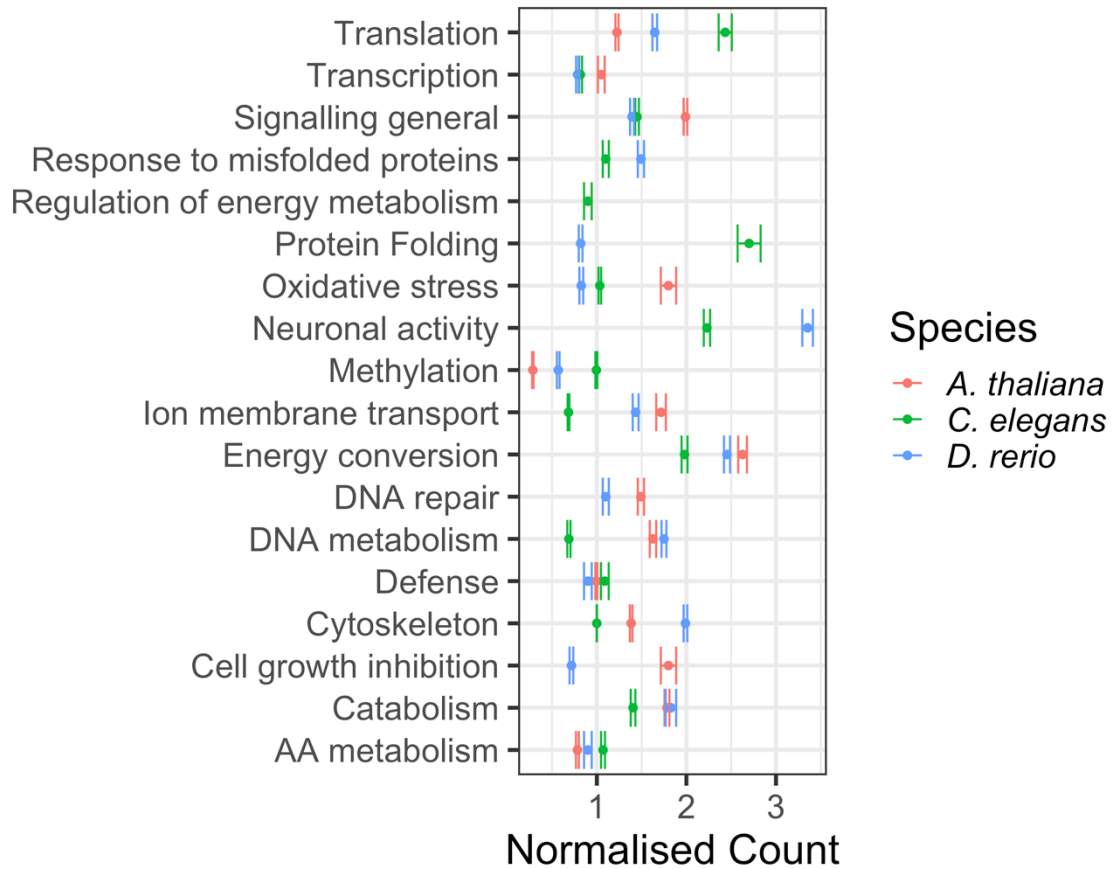
Figure S2: Uncertainty estimation with respect to species

To estimate the uncertainty of our analysis, we deleted 10% of all samples from our dataset and re-ran the analysis. We repeated this process for all combinations of samples and from this distribution we calculated 95% confidence intervals of the result metrics. The sum of shared GO-terms per category divided by the number of comparisons considered for the respective species is shown along with the corresponding confidence intervals.
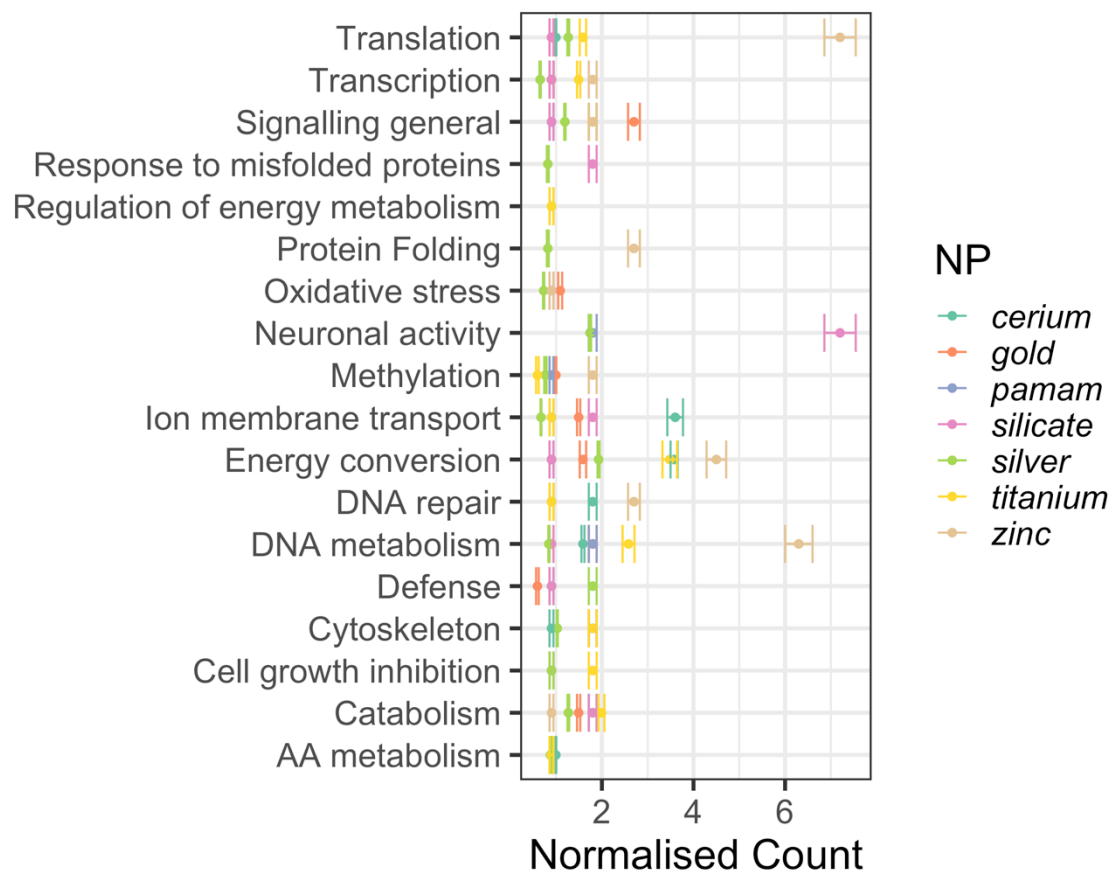
Figure S3: Uncertainty estimation with respect to nanoparticle type

The sum of shared GO-terms per category divided by the number of comparisons considered for the respective nanoparticle is shown along with the corresponding confidence intervals.
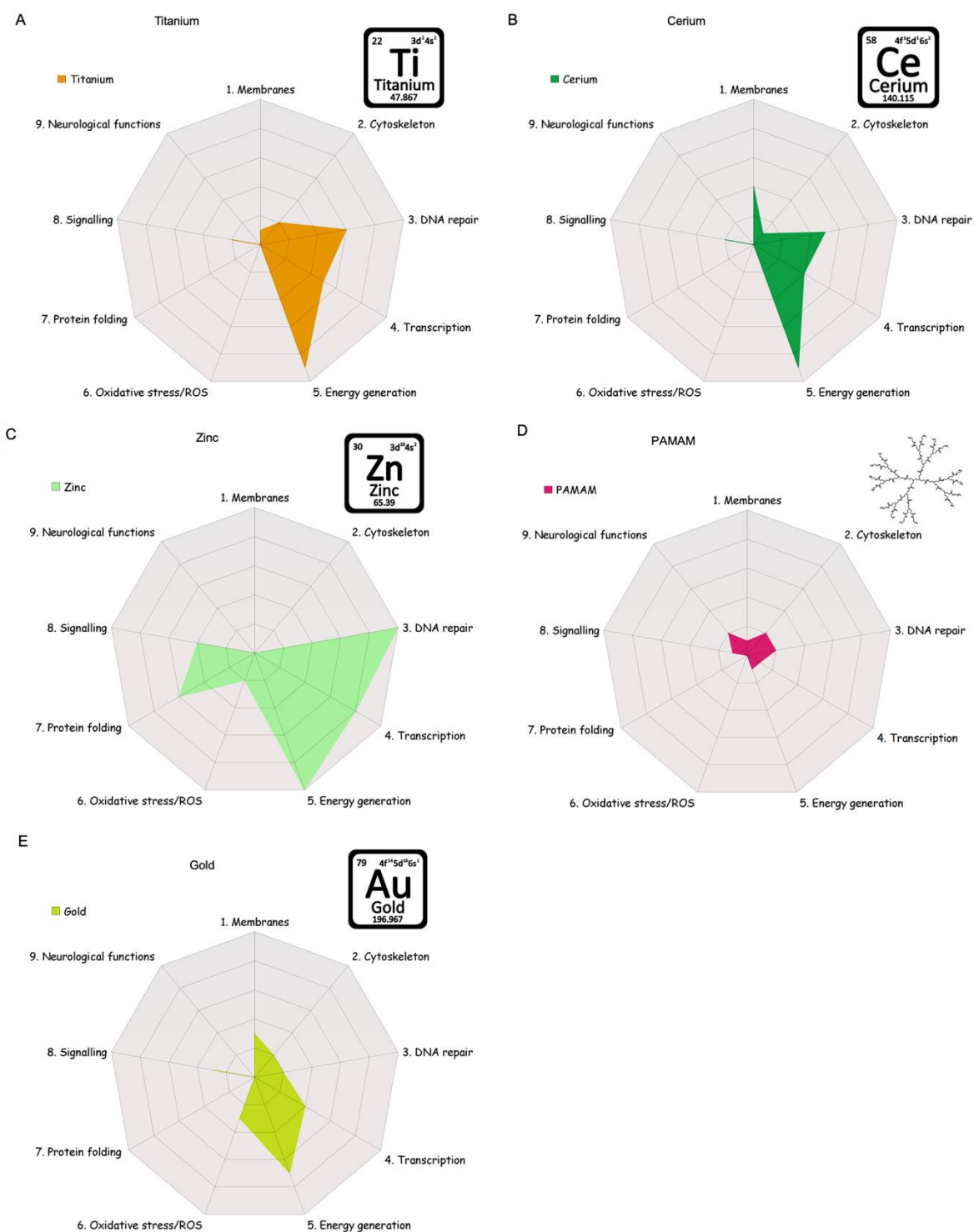
Figure S4: Toxicity profiles for ENMs

The response profiles were determined by assessment of the number of enriched functional annotations (GO-terms) associated with respective common toxicity mechanisms illustrated as general profile for titanium (A), cerium (B), zinc (C), PAMAM (D), gold (E). The score values indicate the number of functional annotations which were found for each category and is normalised to the number of tested contrasts for each organism.