Importance of metal biotransformation in cell response to metallic nanoparticles: a transcriptomic meta-analysis study

Alice Balfourier[†,‡], *Anne-Pia Marty*[†], *Florence Gazeau*[†,*]

[†] Université Paris Cité, Laboratoire Matière et Systèmes Complexes (MSC), CNRS, 45 rue des Saints Pères, 75006 Paris, France

[‡] Current address: Laboratoire de BioMolécules (LBM), Département de Chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 24 rue Lhomond, 75005, Paris, France

* Corresponding author: florence.gazeau@u-paris.fr

Supplementary Information

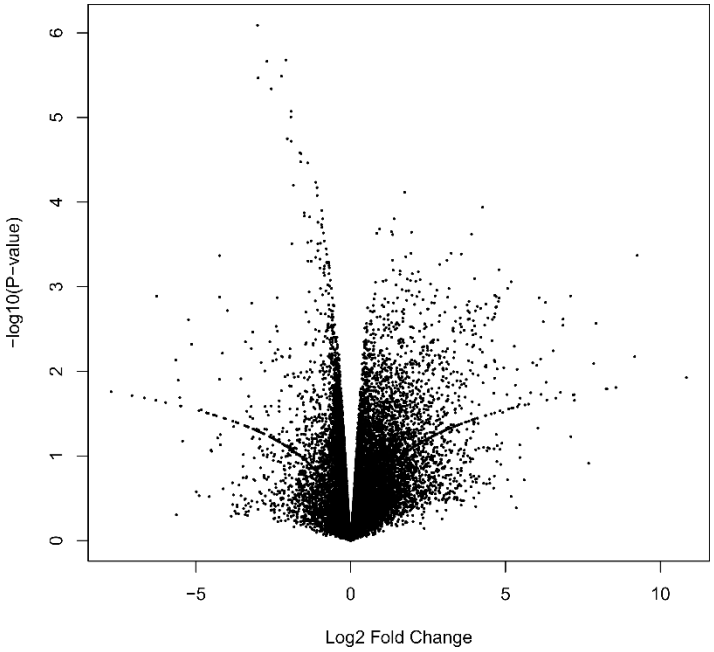

Fig. S1. Example of a volcano plot that does not match quality standard because of unexpected alignment between points. The corresponding original data set was excluded from the analysis.

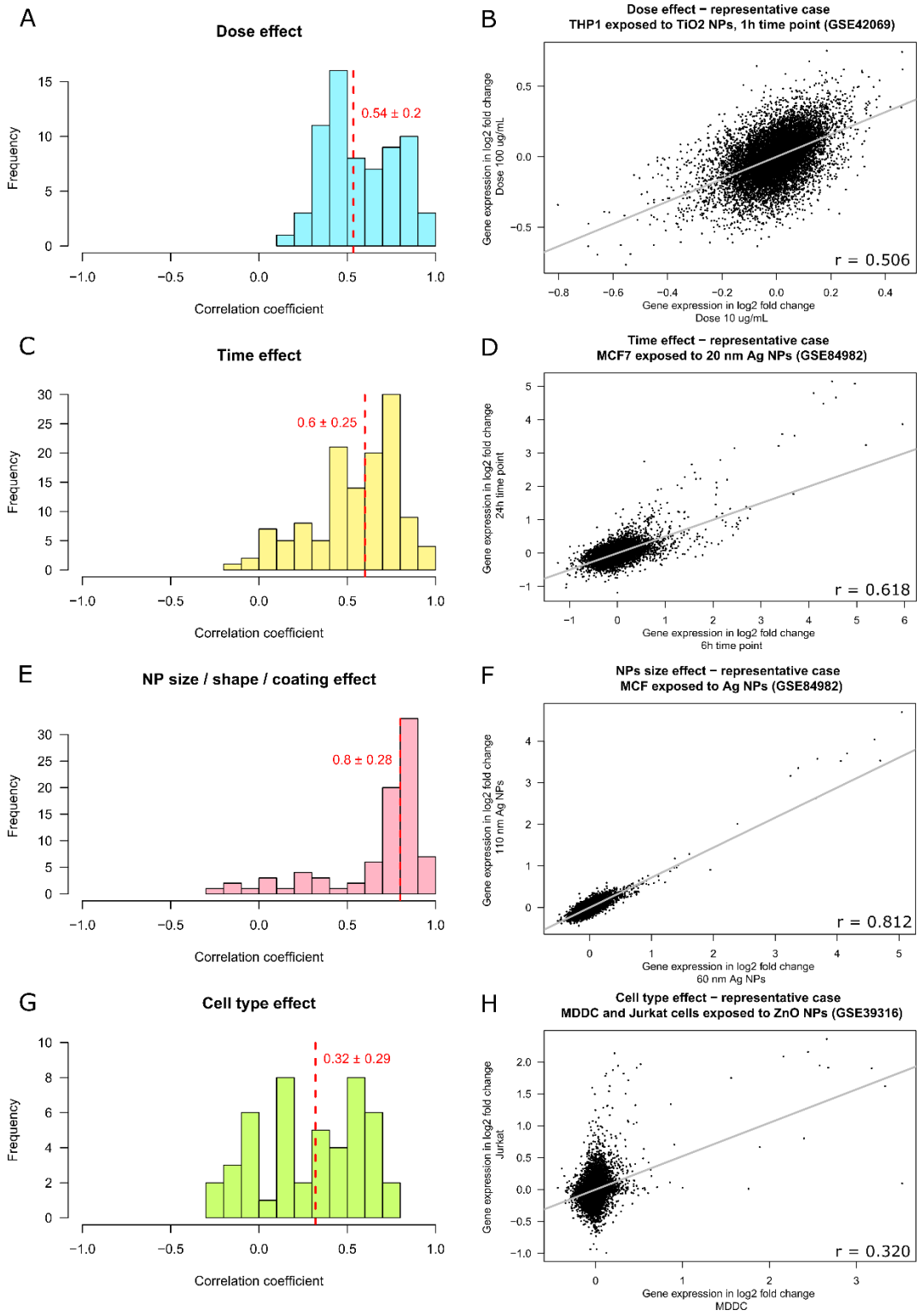

Fig. S2. Distribution of correlation coefficients between log_2 fold change gene expression calculated between the different experimental parameters tested in each dataset (A, C, E, G), and representative examples of these correlation coefficients (B, D, F, H). Experimental parameters displayed are dose (A,B), time (C,D), NP size, shape or coating (E, F) and cell type (G, H). Red dotted lines and numbers indicate the median of the distribution and the standard deviation.

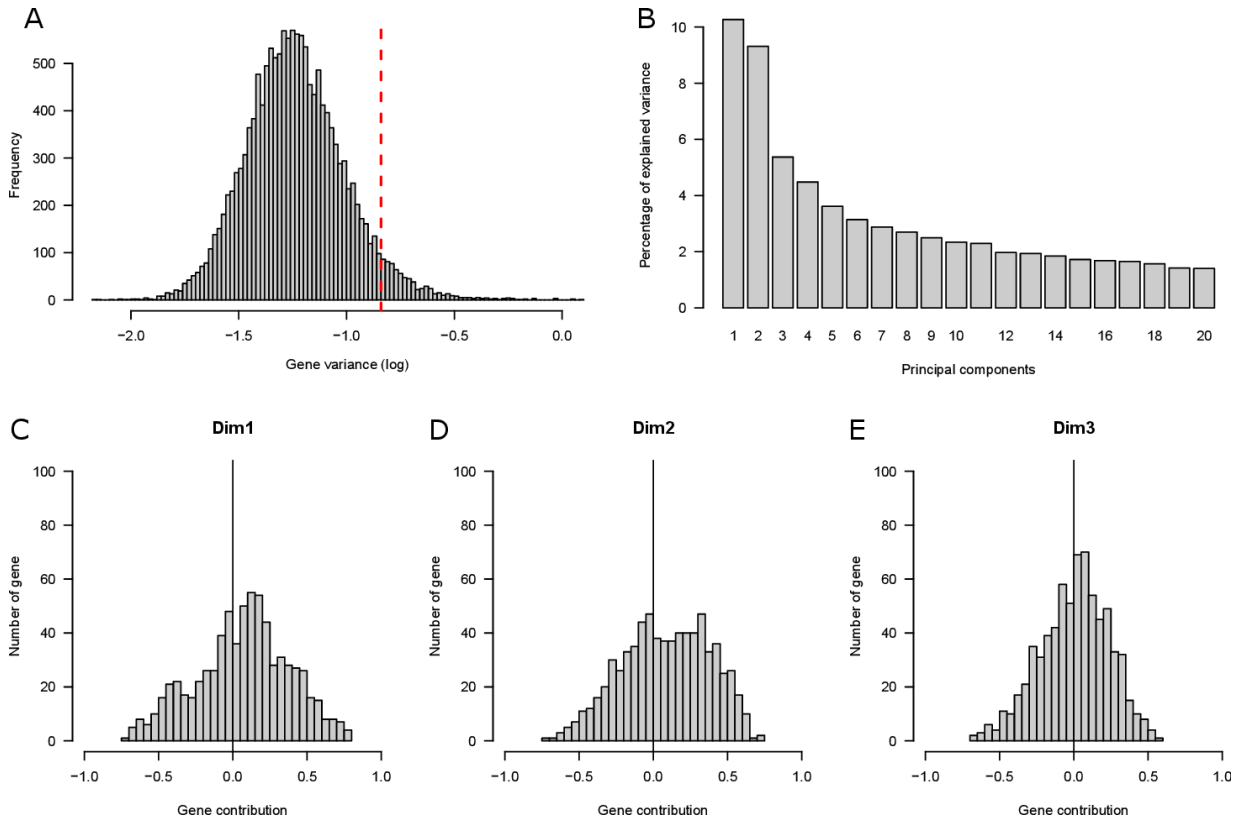

Fig. S3. Principal Component Analysis (PCA) associated graphs and quality controls. (A) Histogram of the variance in logarithmic value for all considered genes. The red dotted line indicates the threshold above which the genes were selected as input for the PCA. (B) Percentage of explained variance for the 20 principal components of the PCA. (C, D, E) Histogram of the gene contribution to the three first principal components of the PCA.

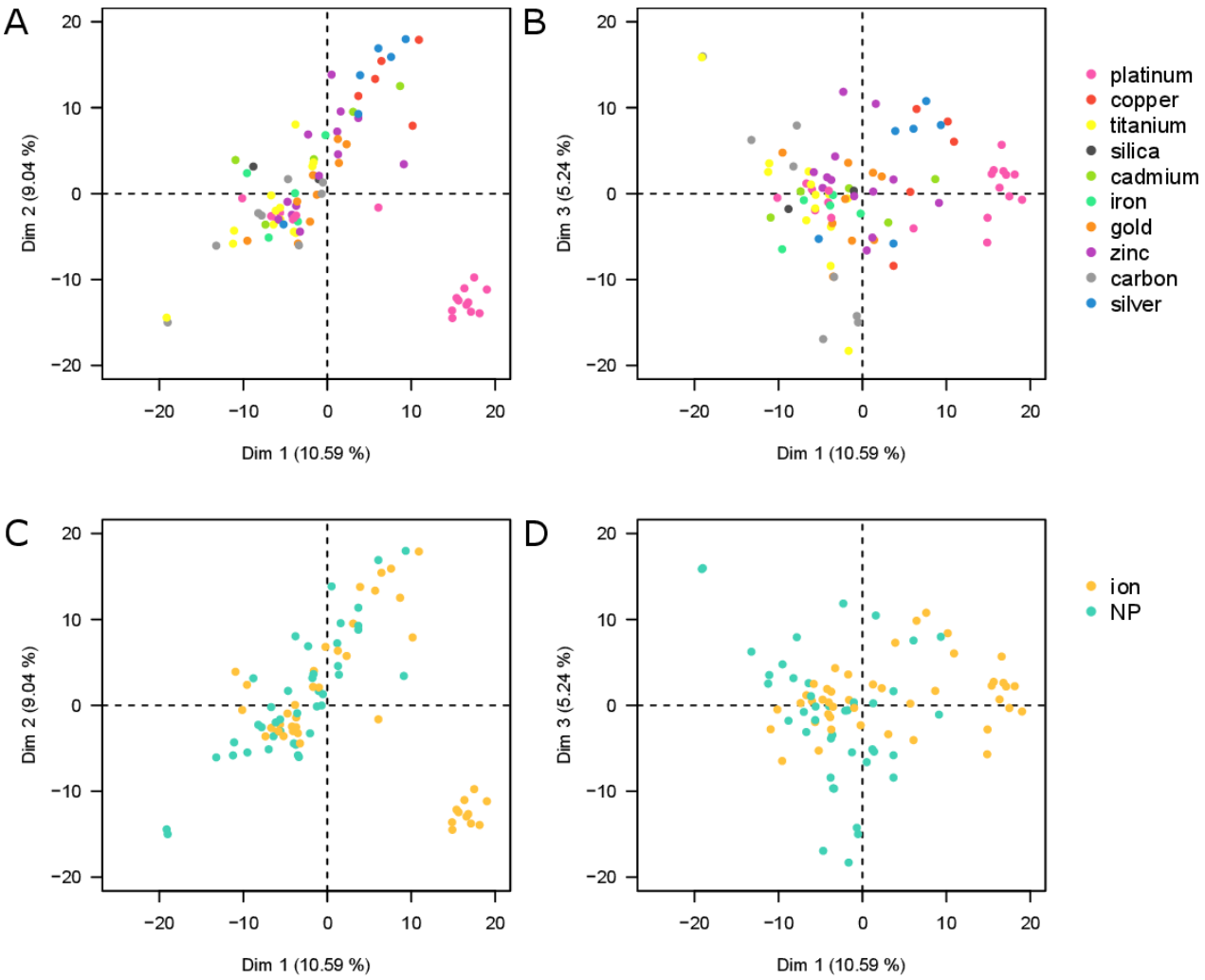

Fig. S4. Principal component analysis performed on the 85 studied conditions and the genes with the highest variance. Representation of the two first principal components (A and C) or the first and third component (B and D) with visualization of the metal type (A and B) or the formulation of the metal (C and D).

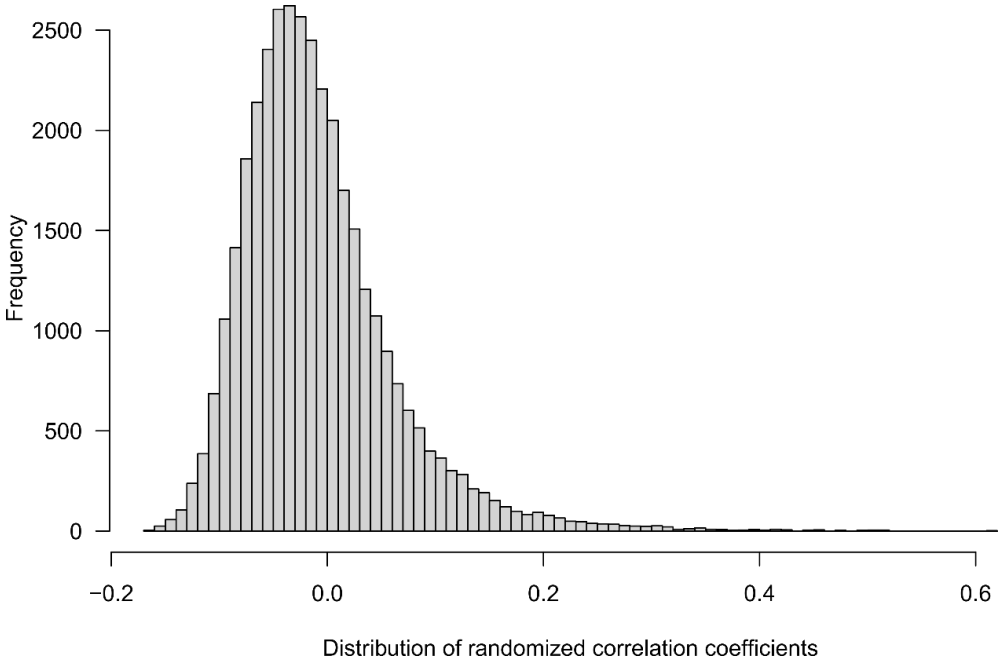

Fig. S5. Distribution of weighted correlation coefficients from randomized data of Figure 2. All combination of vectors are here equally represented.

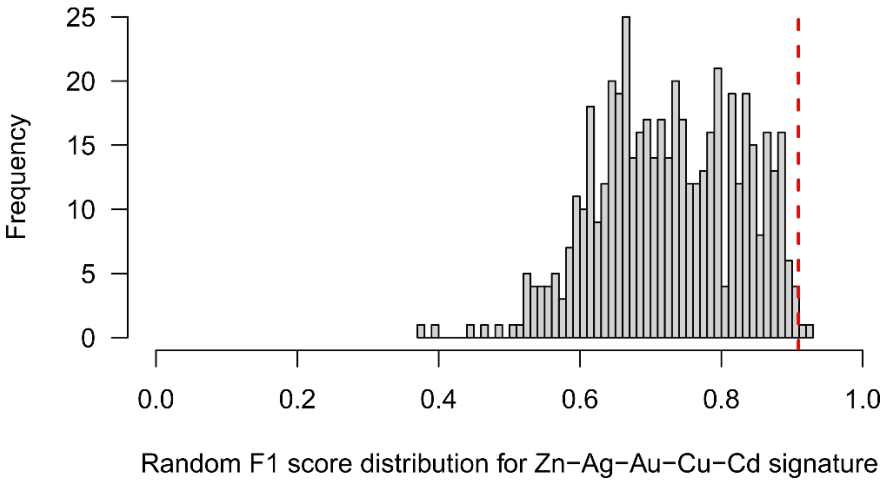

Fig. S6. Distribution of F_1 scores obtained for random set of genes from the 50 top genes corresponding to Zn, Ag, Au, Cu and Cd signature. The number of randomly selected gene is chosen according to the size of the optimized selection. Red dotted lines indicate the optimized F1 score obtained, which is here 0.9090. The calculated p-value is 0.0040.

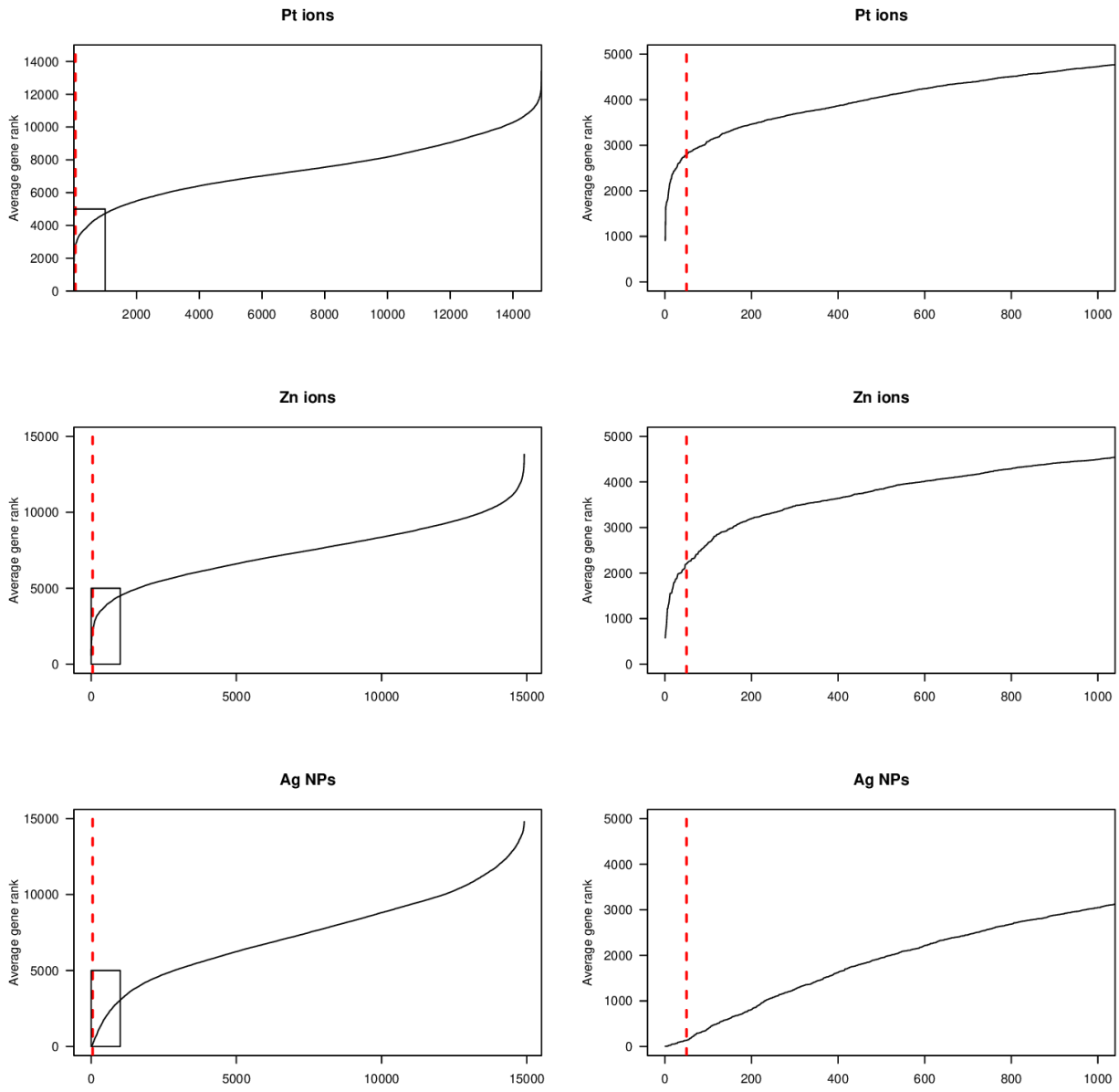

Fig. S7. Representative average rank plots for three materials/ionic forms that average 19 conditions (Pt ions), 7 conditions (Zn ions) and 3 conditions (Ag NPs). The dark square areas are highlighted by the graphics on the right side. The red dotted line delineates the fifty first genes that have been selected for further study.

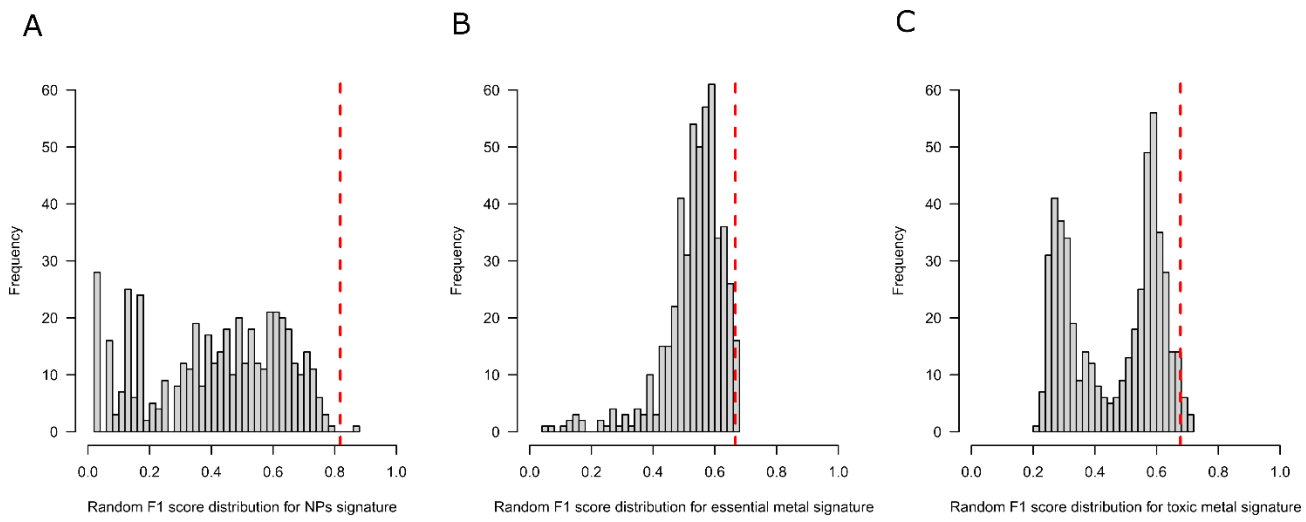

Fig. S8. Distribution of F_1 scores obtained for random set of genes of the same size of the optimized selection. (A) F_1 score distribution obtained from the 50 top genes corresponding to persistent nanoparticles. (B) F_1 score distribution obtained from the 50 top genes corresponding to essentials metals. (C) F_1 score distribution obtained from the 50 top genes corresponding to highly toxic metals. Red dotted lines indicate the optimized F_1 score obtained in each case. Calculated p-values are respectively of 0.0021 (A), 0.020 (B) and 0.022 (C).

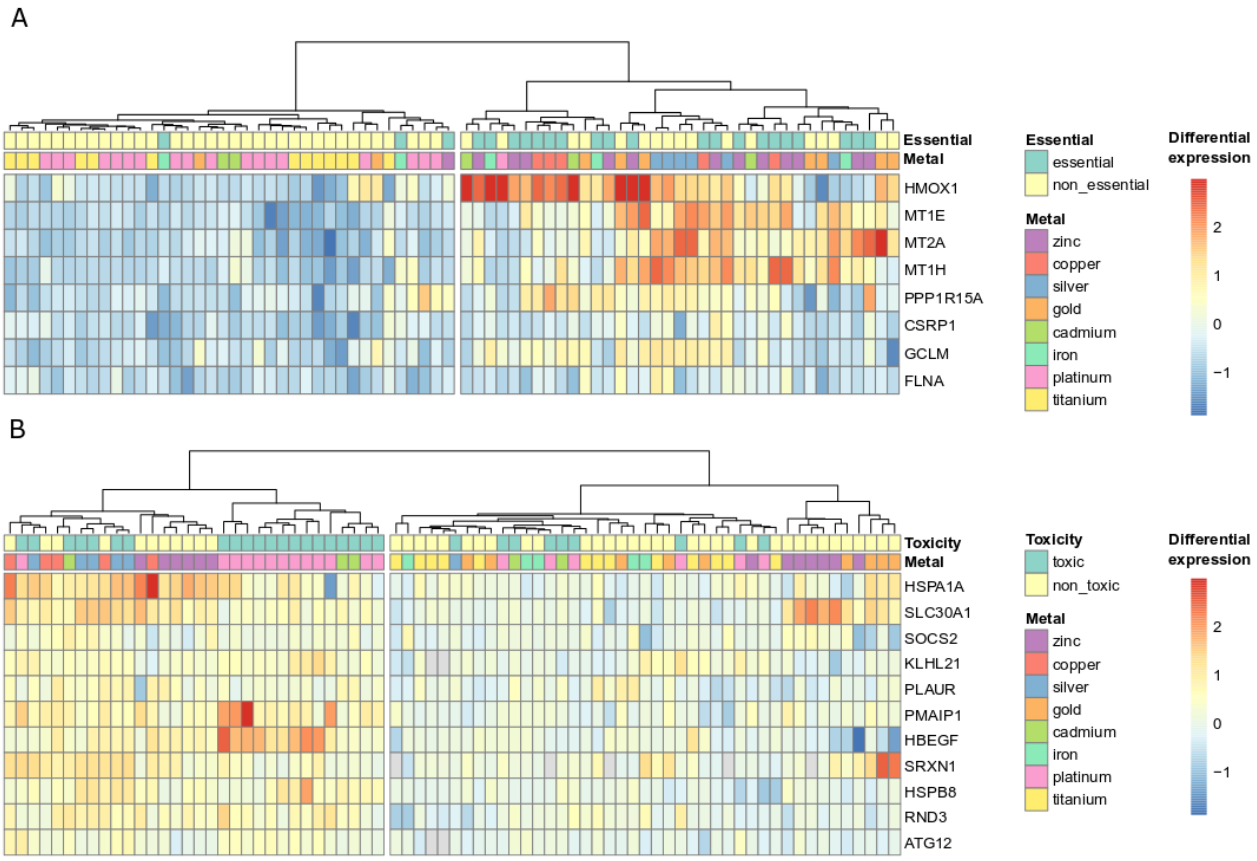

Fig. S9. Heatmap of gene expression for the genes related to bioessential (A) or highly-toxic metals-treated cells (B) across data sets and conditions. Bioessential metals are iron, copper and zinc, while highly-toxic metals are silver, cadmium and platinum. Clustering was performed by ascending hierarchical classification using Ward method on squared distances.

GEO dataset reference	Composition	Form(s)	Form details	Cell type
GSE103115	Platinum	Ionic	Cisplatin	Breast cancer cells (BT549, HCC38, MDA-MB-157, MDA-MB-231)
GSE109375	Copper	Ionic	Copper sulfate	Alveolar epithelium cells (BEAS-2B), colon cancer cells (HCT116)
GSE110410	Titanium	NP	Titanium oxide	Epithelial colorectal adenocarcinoma cells (Caco-2)
GSE110425	Platinum	Ionic	Oxaliplatin	Colon cancer cells (HCT116)
GSE117056	Titanium	NP	Titanium oxide, D = 22 to 214 nm	Liver cancer cells (HepG2)
GSE127773	Copper	NP	Copper oxide, D < 400 nm, nude or -COOH coating	Epithelium cells
GSE127962	Silicon	NP	Silica, spheres, D = 40 to 200 nm	Alveolar epithelium cells (BEAS-2B)
GSE13525	Platinum	Ionic	Carboplatin	Ovarian cancer cells (36M2)
GSE136595	Cadmium	Ionic		Breast cancer cells (MCF7)
GSE16247	Copper, platinum	Ionic	Thioxotriazole copper (II), cisplatin	Fibrosarcoma cells (HT1080)
GSE16349	Iron	Ionic	Ferrous sulfate	Dermal fibroblasts
GSE16425	Titanium	NP	Titanium oxide, D = 7 to 200 nm	Keratinocyte (HaCat)
GSE17676	Titanium	NP	Titanium oxide, nanotubes, L = 1 μm, porous	Endothelial cells, vascular smooth muscle cells
GSE20320	Cadmium	Ionic	Cadmium chloride	Spleen lymphoblasts (TK6)
GSE20677	Gold	NP	Sphere, D = 10 nm, functionalized with an oligonucleotide	Peripheral blood mononuclear cells
GSE2111	Zinc	Ionic	Zinc sulfate	Bronchial epithelial cells
GSE25167	Titanium, zinc	NP	Titanium oxide, D = 5 nm; zinc oxide, D = 9 nm	Keratinocyte (HaCat), melanoma cells (SK-MEL-28)
GSE27204	Cadmium	Ionic	Cadmium chloride	Kidney cells (HK-2)
GSE2964	Zinc	Ionic	Zinc acetate	Lymphoma B cells (Ramos)
GSE30774	Iron	Ionic	Ferric ammonium citrate	Bone osteosarcoma cells (MG-63)
GSE31286	Cadmium	Ionic	Cadmium chloride	Liver cancer cells (HepG2)
GSE3573	Iron	Ionic	Ferric ammonium citrate	Epithelial colorectal adenocarcinoma cells (Caco-2)
GSE38545	Platinum	Ionic	Cisplatin	Ovarian cancer cells (NIH:OVCAR-3, SK-OV-3, TOV-21G)
GSE39316	Titanium, zinc	NP	Titanium oxide, spheres, D = 30 nm; zinc oxide, spheres, D = 15 nm	T lymphocytes (Jurkat), monocytes derived macrophages (hMDM), monocytes derived dendritic cells (mdDC)
GSE42069	Titanium, carbon	NP	Titanium oxide, nanobelt, L =7 μm, D = 10 nm; nanotubes, L =7 μm, D = 25nm	Monocytes derives macrophages (THP1), epithelial colorectal adenocarcinoma cells (Caco-2), small airway epithelial cells
GSE43515	Carbon	NP	Multiwalled carbon nanotubes, carbon black nanoparticles	Lung Fibroblast
GSE45322	Zinc	NP	Zinc oxide, D < 100nm, coated and uncoated	Olfactory neurosphere derived cells
GSE47980	Platinum	Ionic	Cisplatin	Melanoma cells (MM200, SK-MEL-28, igR3, Me4405, Mel-RM), melanocyte

GSE53700	Silicon	NP	Silica, spheres, D = 9 and 18 nm,	Adenocarcinomic alveolar epithelial cells (A549)
GSE54962	Iron	Ionic	Ferric ammonium citrate	Adenocarcinomic alveolar epithelial cells (A549)
GSE55349	Gold	NP	Spheres, D = 5 nm, citrate coated	Epithelial colorectal adenocarcinoma cells (Caco-2)
GSE56432	Gold	NP	Spheres, D = 20 nm, citrate coated	Dermal fibroblasts, prostate cancer cells (PC3)
GSE60159	Zinc	NP	Zinc oxide, D < 100nm, uncoated	Primary human hepatic stellate cells
GSE60408	Gold, zinc	Ionic	Auranofin, zinc pyrithione	Lymphoma B cells (U2932, OCI-LY3, OCI-LY7)
GSE62253	Silver	Ionic and NP	Silver nitrate; spheres, D = 15 nm, polymer coated	Epithelial colorectal adenocarcinoma cells (Caco-2)
GSE6818	Gold	NP	D = 2 to 200 nm	Skin fibroblasts
GSE69644	Platinum	Ionic	Cisplatin	Kidney cells (HK-2)
GSE7035	Platinum	Ionic	Carboplatin	Adenocarcinomic alveolar epithelial cells (A549)
GSE73302	Platinum	Ionic	Cisplatin	Adenocarcinomic alveolar epithelial cells (A549)
GSE80695	Iron	NP	Iron oxide (maghemite), D = 7 nm, polymer coated	Bone derived mesenchymal stem cells
GSE84982	Silver	Ionic and NP	Silver nitrate; spheres, D = 20 to 110 nm	Epithelial colorectal adenocarcinoma cells (Caco-2), breast cancer cells (MCF7)
GSE92901	Carbon	NP	Multiwalled carbon nanotubes, L = 1 to 30 μm, D = 10 to 50 nm; graphite fiber, L = 10 μm, D = 140 nm; fullerene, D = 100 nm	Monocytes derives macrophages (THP1)
GSE92987	Gold	NP	Spheres, D = 14 nm	Retinoblastoma cells (Y79)
GSE9539	Copper	Ionic	Copper sulfate	Liver cancer cells (HepG2)
GSE9951	Cadmium	Ionic	Cadmium chloride	Prostate epithelial cells

Table S1. GEO references and experimental details of the datasets that have been analyzed to determine the cellular response to metallic nanoparticles and metal ions for eight metals (Ag, Au, Cd, Cu, Fe, Pt, Ti, Zn) and non-metallic nanoparticles (C, Si).

Rank	PC1	PC2	PC3
1	GADD45B	DUSP5	HSD17B7
2	NXF1	SPRED2	SHCBP1
3	SERTAD1	TXNRD1	TSC22D3
4	HIST2H2AA3	MT1A	RAD51AP1
5	TUFT1	DNAJB1	TTBK2
6	GADD45A	ABTB2	HERPUD1
7	JUN	SRGAP1	ZNF12
8	ATF3	SPSB1	ZNF148
9	TOB1	PPP1R15A	NEIL3
10	HBEGF	KLF6	ACAT2
11	MAP1LC3B	MT1E	SERPINI1
12	TUBB2A	IGF2BP2	CHAC1
13	DUSP1	SPRY4	GABARAPL1
14	IER5	MT1G	PCYOX1
15	CDKN1A	TRIB3	SGK3
16	ZNF408	LDLR	INHBE
17	OSGIN1	JARID2	IFRD1
18	TNFSF9	NEDD4L	SLC46A3
19	FOSB	MCL1	MTHFD2
20	GPR3	SERPINE2	PIR
21	PMAIP1	SLC7A11	CLK1
22	PHLDA2	MT1B	LYZ
23	IER2	MT2A	KIF18A
24	PDRG1	MAFF	DNAJB4
25	CLCF1	LPP	ZNF652

Table S2. Table of the 25 top genes that contributes the most to the first three principal components of the principal component analysis.

	Gold (NP)	Zinc (ion)	Cadmium (ion)	Copper (NP)	Copper (ion)	Silver (NP)	Silver (ion)	Gold (ion)	Zinc (NP)	Iron (NP)	Carbon (NP)	Iron (ion)	Silicon (NP)	Platinum (ion)	Titanium (NP)
Gold (NP)	< 0.001	< 0.001	< 0.001	< 0.001	0.0033	< 0.001	< 0.001	< 0.001	< 0.001	0.3400	0.0733	0.0767	0.9733	0.0133	0.0033
Zinc (ion)		< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.3467	0.1933	0.1867	0.7933	0.1000	0.0800
Cadmium (ion)			< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.5700	1.0000	0.0100	< 0.001	0.0133	0.7433
Copper (NP)				< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.1700	0.9733	< 0.001	< 0.001	< 0.001	0.1467
Copper (ion)					< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	0.4367	0.3433	0.0033	< 0.001	< 0.001	0.1400
Silver (NP)						< 0.001	< 0.001	< 0.001	< 0.001	0.5700	0.2267	< 0.001	0.0067	< 0.001	0.0767
Silver (ion)							< 0.001	< 0.001	< 0.001	0.4833	0.1867	< 0.001	0.0200	< 0.001	0.0633
Gold (ion)								< 0.001	< 0.001	0.4367	0.9433	0.0100	0.1233	0.0033	0.0400
Zinc (NP)									< 0.001	0.0833	0.2300	0.0033	0.0067	< 0.001	0.0033
Iron (NP)										< 0.001	0.4167	< 0.001	0.0133	1.0000	0.4233
Carbon (NP)											< 0.001	0.8000	0.0367	0.0133	< 0.001
Iron (ion)												< 0.001	< 0.001	0.0233	0.1267
Silicon (NP)													< 0.001	0.0133	0.0167
Platinum (ion)														< 0.001	0.3300
Titanium (NP)															< 0.001

Table S3. Table of the p-values calculated from randomization method to estimate the significancy of the calculated correlation coefficients displayed in Figure 2

Data file S1. General table of gene expression in log2fold change for the 85 studied conditions and the 14912 selected genes.

Data file S2. Table of the 50 genes with the lowest average rank for each studied nanoparticulate or ionic form.

Data file S3 to S14. Tables of the enriched biological functions for each studied nanoparticulate or ionic form, with the associated FDR adjusted p-value and the gene driving the enrichment.