Oxidative stress after pollutant exposure depends strongly on experimental design and pollutant properties: a meta-analysis

Max V.R. Döring[1], Heike Feldhaar[1,2], Ana L. Antonio Vital[3], Magdalena M. Mair[2,3,*]

[1] Animal Population Ecology, Animal Ecology I, University of Bayreuth, Bayreuth, Germany

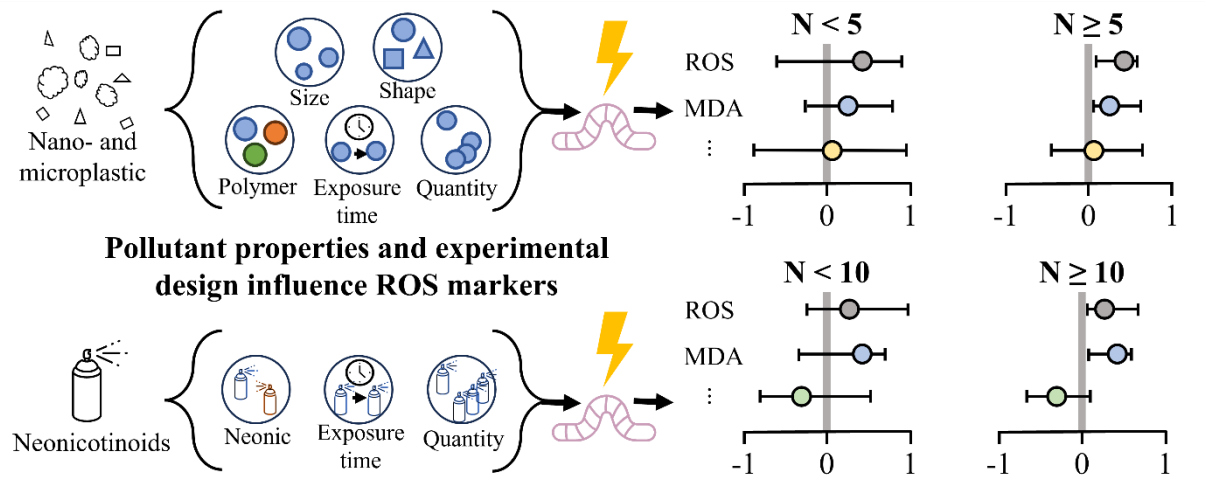
[2] Bayreuth Centre of Ecology and Environmental Research (BayCEER), Bayreuth, Germany

[3] Statistical Ecotoxicology, University of Bayreuth, Bayreuth, Germany

[*] Corresponding author

Correspondence to: magdalena.mair@uni-bayreuth.de

Graphical Abstract



Abstract

Measurements of reactive oxygen species (ROS) are often performed to assess a species' general sublethal stress response to a pollutant. However, ROS bioassays often produce seemingly ambiguous results, and the drivers that lead to these differences are largely unknown. To approach this gap, we conducted a meta-analysis on ROS generation, ROS-associated damage products, enzyme activities, and gene expression levels in response to exposures to two groups of pollutants, nano- and microplastic particles (NMP) and neonicotinoid insecticides (neonics). Based on 2294 ROS-related measurements extracted from 45 studies, we show that measured effects vary substantially with a strong overlap of measured effects with zero. As likely drivers of this variance, we identified multiple parameters of experimental design and pollutant properties. Finally, we performed data simulations and power analyses to investigate how well single experiments are able to detect ROS-related effects. We show that 21 out of 27 ROS markers achieve sufficient power (80%) to demonstrate effects with sample sizes N < 20. Given the pollutant-dependent variability in ROS related responses and the low power of some markers, conclusions derived from single studies with low sample sizes (N < 5) are however at risk of being less informative than previously assumed.

Keywords

Oxidative distress, Contamination, Ecotoxicology, Annelids, Physico-chemical properties, Statistical power

Highlights

- Meta-analysis on effects of pollutants on reactive oxygen species markers.
- Effects are moderated by experimental design and pollutant properties.
- Studies with commonly used sample sizes cannot reliably detect these effects.
- More emphasis should be put on reporting pollutant properties and increasing sample size.

1. Introduction

Reactive oxygen species (ROS) are a class of molecules derived from oxygen (O_2) that is more reactive than O_2 itself [1]. ROS are naturally produced within cells by metabolic processes, including the mitochondrial respiratory chain [2,3] and NADPH oxidases [4] and are important as a redox signalling agent [5,6]. ROS such as hydrogen peroxide have the potential to cause growth arrest and cell death in high concentrations [5], so their levels inside organisms are controlled tightly. This is achieved by regulating the rate of endogenous ROS generation [7] and/or by increasing ROS scavenging via upregulation of antioxidant production (e.g., glutathione [8] and ROS scavenging enzyme expression and activity [9]). However, intracellular ROS levels can additionally be increased by external stressors, overwhelming the ROS scavenging capabilities, and leading to oxidative stress and subsequent damage [10,11]. Consequently, to assess the impact of various external stressors such as pollutants on organisms, it is common practice to monitor various ROS markers such as ROS levels, ROS-associated responses (e.g., ROS scavenging enzyme activities), and cellular damage products. Increased levels in these markers are interpreted as increased oxidative stress and consequently higher toxicity of the tested external stressor.

When reading ecotoxicological literature, we observed that the same ROS markers are often reported to be affected in different directions in different studies (e.g., one study showing an increased enzyme activity while another study showing a decreased activity of the same enzyme; see examples in **Tab. *1***). This observed variability raises the question of its underlying causal factors. For instance, specific ROS markers might be up- or downregulated depending on factors like pollutant type, pollutant properties, and experimental conditions (e.g., exposure time). Given the high variance in observed effects, it is additionally unclear whether currently used experimental designs have sufficient statistical power (i.e., the rate at which true effects can indeed be detected [12]) to reliably detect true effects.

To answer these questions, we used a meta-analytic approach to investigate effects of two different pollutant groups, nano- and microplastic particles (NMP) and neonicotinoid insecticides (neonics), on ROS markers in annelid worms as an example organism. Annelid worms were chosen since they are established model organisms in ecotoxicology and very sensitive to changes in their environment [13]. We chose NMP and neonics as representative examples for pollutants since they exhibit completely different modes of actions. NMP are omnipresent particulate pollutants that have received increasing attention in the last 20 years while their number in the environment is steadily increasing [14]. They elicit mostly sublethal effects, but the mechanisms of how NMP lead to oxidative stress are not sufficiently well understood and likely include both, particle-induced effects (e.g., physical damage and subsequent activation of the immune system [15]) and chemical-induced effects (e.g., through plastic-associated chemicals [16]). Neonics have been one of the most used classes of agricultural insecticides worldwide [17] with well-documented negative effects on various non-target organisms [18]. They seem to induce increased ROS levels by multiple pathways [19] such as disruption of Ca^{2+}-homeostasis [20,21] and potentially by altering key ROS regulatory genes [21]. The extent to which different neonics activate these pathways remains unclear, however.

In total, we extracted 2294 ROS-related measurements from 45 studies (see section 2.1). We examined (1) the directions and strengths of average effects of NMP and neonics on different ROS markers and (2) the explanatory power of pollutant properties and experimental design choices to the observed variance. Based on the estimated average effect sizes and variances, we (3) performed extensive data simulations and power analyses to estimate the statistical power of typical test designs, and the sample sizes required to detect true effects at sufficiently high rates (i.e., with a statistical power of 0.8). To this end, we first assumed a general ROS-related stress response and, second, inspired by the empirical results, allowed for differences in the true stress response due to differences in experimental design parameters (e.g., concentration, exposure duration, sample size) and pollutant properties (e.g., polymer type for NMP or type of neonic).

Tab. 1 Reactive oxygen species (ROS) concentrations, damage products, and ROS responses are highly variable within and among studies. 11 example studies from our data set that measured ROS concentration, malondialdehyde (MDA) concentration, catalase (CAT) activity, glutathione-S-transferase (GST) activity, superoxide dismutase (SOD) activity, and/or peroxidase (POD) activity in annelids after nano- and microplastic (NMP) exposure. As examples, we chose studies from our dataset that measured at least three of the six ROS markers. Upwards pointing arrows with blue background symbolize reported significant increases in enzyme activity or molecule concentration compared to the control, downwards pointing arrows with yellow background symbolize a significant decrease. Hyphens with grey background represent no significant difference between NMP treatment and control. Enzyme activities or concentrations that were not determined are shown as NA. Most enzyme activities or molecule concentrations show mixed results within the same ROS marker and study when multiple exposure times, concentrations, shapes, sizes, or plastic ages were tested.

Authors	Date	DOI	ROS	MDA	CAT	GST	SOD	POD
Baihetiyaer et al.	2023	10.1016/j.envpol.2023.121285	↑	↑↓	↓	NA	–↓	–↑↓
Cheng et al.	2020	10.1016/j.scitotenv.2020.14128	↑	–↑	↓	↓	–↓	NA
He et al	2023	10.1016/j.scitotenv.2023.16285	–↑	–↑	–↓	–↑↓	–↑↓	NA
Holzinger et al.	2022	10.1016/j.scitotenv.2022.15638	NA	–↓	–↑	–	NA	NA
Li et al.	2021	10.1016/j.scitotenv.2021.14700	NA	–↑↓	–↓	–↓	–↓	NA
Li et al.	2021	10.1016/j.scitotenv.2020.14403	NA	↑	↑	NA	↑	↑
Li et al.	2023	10.1016/j.chemosphere.2022.13	↑	–↑	NA	NA	↑	NA
Liu et al.	2022	10.1016/j.envint.2022.107158	–↑	–↑	–↑↓	–↑↓	–↑	–↑↓
Shang et al.	2023	10.1016/j.scitotenv.2023.16695	↑↓	–↓	–↑↓	NA	–↑↓	–↑↓
Wang et al.	2019	10.1016/j.envpol.2019.03.102	NA	–↑	–↑	–↓	–↓	↑
Zhao et al.	2023	10.1016/j.scitotenv.2022.16009	NA	↑↓	–↑↓	–↑↓	↑↓	–↑↓

We extracted ROS-related measurements (e.g., enzyme activity measurements), experimental design parameters (e.g., number of replicates) and reported statistical outcomes from 21 and 24 studies examining the effects of NMP and neonics on annelids, respectively. The log-transformed ratio of means (logROM) for differences between treatments (numerator) and controls (denominator) were calculated as effect sizes. Overall average effects of different pollutants on measured ROS markers (i.e., one effect size estimate per ROS marker for NMP and neonics separately) were derived from mixed meta regression models [23] (see section 2.2). The following ROS markers were evaluated: ROS formation (overall ROS and hydroxyl radical (•OH) concentration), ROS responses (enzyme activities of catalase (CAT), superoxide dismutase (SOD), glutathione-S-transferase (GST), peroxidase (POD), glutathione reductase (GR), carboxylesterase (CarE), mRNA expression levels of CAT, SOD, and GST, total antioxidant capacity (TAC), glutathione (GSH) concentration), and ROS-associated damage products (malondialdehyde (MDA), 8-hydroxy-2-deoxyguanosine (8-OHdG), and protein carbonyl group (PC) concentrations).

2. Methods

2.1 Data collection

We conducted a systematic search for all peer-reviewed publications that were published before December 2023 and that examined the effects of NMP or neonics on ROS generation, ROS scavenging enzyme activities, and ROS-associated damage products in terrestrial annelids. We started the search for suitable articles by choosing 10 already published, representative articles for each NMP and neonics. Afterwards, we optimised our search string such that these 10 previously selected articles appeared in the 50 most relevant (i.e., the first 50) papers on Web of Science (*https://www.webofscience.com*). For search term optimization, different combinations of keywords were tested including terms related to oxidative stress responses (e.g., ros, oxidative stress, reactive oxygen species, detoxification and names of specific ROS markers), terms associated with the focal taxon of annelid worms (e.g., annelid, Eisenia, earthworm) and terms describing NMP (e.g., microplastic*, nanoplastic*) and neonicotinoids as focal pollutant types (e.g., neonic* and different specific neonicotinoid names). Finally, on the 14th and 15th of December 2023, we searched on Web of Science with the following search string for NMP (49 hits): "(annelid* OR eisenia) AND (reactive oxygen species OR oxidative stress) AND (microplastic* OR nanoplastic*)" and neonics-associated studies (51 hits), respectively: "(annelid* OR eisenia OR earthworm*) AND (oxidative stress OR detoxification OR inhibition) AND (neonic* OR imidaclo* OR thiaclo*)".

The titles and abstracts of all publications analysing the effects of NMP or neonics were screened to fulfil the following criteria: The studies investigated the oxidative stress of terrestrial annelids following an *in vivo* pollutant exposure (i.e., no single cell analyses) and included a negative control without the respective pollutant. We excluded review articles and meta-analyses. Studies using tire wear particles (TWPs) or in which the worms were co-exposed to other pollutants in addition to NMP or neonics were excluded to avoid confounding by additional substances (e.g., softener and vulcanization agents in TWPs [24]). We also excluded studies which did not report

essential data for statistical analysis (e.g., missing number of biological replicates) and where the authors have not responded to a request for raw data. Lastly, for an improved comparability within our NMP data, we excluded data generated after a recovery period (i.e., period after the pollutant exposure without pollutant; 84 data points) and data obtained through a neonic contact test (i.e., annelids are exposed to the neonic on a filter paper in petri dishes) from our main analysis (174 data points). For the latter, we included a separate analysis in the supplement (**Supplementary Table S1, Supplementary Fig. S1Fehler! Verweisquelle konnte nicht gefunden werden.**). This left us with 21 NMP studies (898 data points) and 24 neonic studies (1396 data points). For a visual representation (PRISMA (Preferred Reporting Items for Systematic reviews and Meta-analyses) Flow diagram) of the article selection process see **Supplementary Fig. S2**.

From each of the remaining articles, we extracted data regarding the species and life stage of annelids at the start of exposure, the exposure route (whether the pollutants were mixed in the soil or food or if a contact test was performed), type of soil (artificial or collected from the environment), soil parameters (pH and mean temperature), exposure and recovery duration, nominal and measured pollutant concentration, added chemicals, screening for chemical and NMP contamination of the soil, uptake validation (whether pollutant uptake was verified) and validation method, food type and dose, measured ROS marker and respective unit, measurement method, the number of biological and technical replicates, and the pollutant manufacturer. The extracted ROS markers were: measured ROS levels, oxidative stress associated enzyme activities (CAT, SOD, GST, POD, GR, CarE) and some of their mRNA expression levels, GSH concentrations, the TAC, and concentrations of ROS-induced cell damage products (MDA, Protein carbonyl, 8-OHdG). For NMP, we additionally extracted whether NMP were cleaned prior to exposure and the solvent used for cleaning, the polymer type, shape, nominal and measured mean particle size, and if applied, method of aging (including artificial UV-weathering and exposure to environment prior to their use in exposure bioassays). For neonics, we determined the specific neonicotinoid identity (i.e., the name of the chemical) that was used. ROS marker measurements were

167 preferentially extracted from the text or calculated from the raw data. If this was not possible,
168 measurements were extracted from figures using the *metaDigitise* package (version 1.0.1) [25] in
169 R (version 4.3.1) [26]. All extracted data were double-checked and validated by a second person
170 and discrepancies were discussed until a consensus was reached.

171 2.2 Data analysis

172 Units of all measured concentrations, enzyme activities, and expression levels were converted
173 into nine defined units (fluorescence intensity, fluorescence intensity/mg protein, nmol/mg
174 protein, nmol/ml homogenate, U/mg protein, ng/g fresh weight, ng/l, ng/mg DNA, and fold
175 change). In case results were reported as a fold change, we set the mean of the control per
176 definition to 1 when data were obtained via the *metaDigitise* package. We did not do this for 48
177 data points of one study, where the measurements were reported as fold change, but the mean
178 value of the control was clearly different from 1. Two articles [27,28] quantified the thiobarbituric
179 acid reactive substances concentration as a proxy for lipid peroxidation. These ROS markers were
180 treated as malondialdehyde concentrations [29]. All pollutant concentrations were log
181 transformed and converted to weight percentages (% w/w). If the nominal as well as the measured
182 concentration or size were reported, we used the measured concentration in our analysis. The
183 NMP concentration in one article (2 data points) could not be converted as it was only reported in
184 particles per kg soil. These two data points were neglected in all models considering NMP
185 concentration. Most articles that measured enzyme activity and expression levels did not report
186 the exact enzyme or gene they investigated. Therefore, we summarized the genes by their function
187 (CAT, SOD, GST, POD, CarE, GR) rather than their specific names (e.g., glutathione peroxidase 1).

188 All statistical analyses were done in R (version 4.3.1) [26]. The data were filtered using the package
189 *dplyr* (version 1.1.3) [30]. For each treatment-control pair, we calculated the *log transformed ratio*
190 *of means* (logROM) as an effect size and corresponding sampling variance (vi) using the *metafor*
191 package (version 4.4-0) [23]. We excluded data points where both control and treated worms had
192 measured values of zero (9 cases in total) which made it impossible to calculate the logROM. We

also excluded data points with a sampling variance of zero (21 cases in total) since the model cannot calculate the proper weight of these data points and a true variance of zero seems highly unlikely. To assess whether the exclusion of these data points infcluenced our results, we repeated all analyses on the complete dataset and replacing the zero variances with 10^{-7}. This did not change our general results (i.e., no changes in significances; see **Supplementary Fig. S3** and **Supplementary Table S2**).

We multiplied the logROM with minus one, so that positive logROM indicate higher measurements of the treatment group compared to the control. Mixed meta regression models without intercept were fit to the data to evaluate the effects of NMP and neonics. We included the ROS markers as a fixed effect and a nested random intercept (~1|DOI/individual level), where the article's Digital Object Identifier (DOI) accounted for the heterogeneity between studies and the individual level random effect accounted for the heterogeneity between samples within studies. We additionally included a quadratic term and intercept in models that examined the effects of neonics and NMP over time, with different sizes, or at different concentrations. These models were plotted with *metafor* and rated as significant, if at least one of the terms had a *P* value below 0.05. Data and models were illustrated using the packages *orchaRd* (version 2.0) [31] and *ggplot2* (version 3.4.4) [32]. Since the GST mRNA expression was only measured at two time points after NMP exposure (14 and 28 days), we only assumed a linear relationship in these cases and excluded the quadratic terms. Additionally, the effects of different NMP sizes on two ROS markers (TAC and GST mRNA expression) could not be evaluated because only one NMP size was used.

Since some pollutant properties determined the effects they have on certain ROS markers, we wanted to figure out if including the properties as fixed factors in a model would significantly improve its fit compared to models without them. Using likelihood ratio tests (LRT), we compared a full model containing all properties (exposure time, concentration, shape, polymer, and size for the NMP dataset, and exposure time, concentration, and neonic identity for the neonic dataset) with models, where one property at a time was excluded (i.e., reduced models). Additionally, we

wanted to figure out if the effects of experimental design and pollutant properties are different among ROS markers. We therefore created a full model with all the above-mentioned properties plus their interactions with ROS marker and compared it with reduced models via LRT, where each of the interaction terms was excluded one by one. Finally, we calculated the proportion of the variance in the data that was explained by the pollutant properties and experimental design parameters. To this end, we fitted a model containing only ROS marker as a fixed effect and calculated the sum of both σ^2 variance components ($\sigma^2_{endpoint}$). We repeated the procedure with a full model containing all the above-mentioned properties plus their interactions with ROS marker ($\sigma^2_{predictors}$). Finally, we calculated the residual heterogeneity (i.e., variance in the true effects) by subtracting $\sigma^2_{predictors}$ from $\sigma^2_{endpoint}$, divided by $\sigma^2_{endpoint}$. Since the proportion of the sampling variance to the total unaccounted variance was extremely low (< 0.6 %), we used the residual variance as an approximation for the total unaccounted variance.

2.3 Power analysis

We conducted multiple power analyses to determine, at which rates commonly performed experiments can detect true effects. To this end, we calculated the means and standard deviations (sd) of each control and treatment group for each ROS marker (NMP: 12 ROS markers times two groups = 24 separate values; neonics: 15 ROS markers times two groups = 30 separate values). Means and sd were then used to draw random samples from normal distributions, both for the control and the treatment group, thus simulating experiments in the lab. This procedure was repeated 100 times for each ROS marker-pollutant-combination. Each simulated control-treatment pair was then analysed via a t-test. The statistical power was finally calculated as the proportion of t-tests with *P* values equal or below 0.05. The simulation process was repeated with increasing sample sizes (i.e., the number of random samples drawn) ranging from three to 100 (incremented by one). Finally, we determined the lowest sample size at which the statistical power was equal or higher than 0.8 (which would translate into significant results in 80 % of

experiments) and the statistical power at sample sizes of three and eight, which were the most frequently used and highest sample sizes in our dataset, respectively.

Considering differential stress responses dependent on experimental design parameters and pollutant properties, we performed a second set of power analyses that took these differences into account. To this end, we calculated the means and sd separately for each combination of pollutant properties and experimental design parameters (NMP: 813 combinations, neonic: 1241combinations). For each combination, we performed a power analysis as described above with a maximum sample size of 50 instead of 100 (47 sample sizes x 100 t-tests = 4700 t-tests per combination). The statistical power for different property-experimental design combinations was averaged for each sample size and ROS marker. Finally, the lowest sample size at which the statistical power was equal to or greater than 0.8 was extracted, along with the statistical power at sample sizes of three and eight.

3. Results and Discussion

3.1 Effects of pollutants of ROS markers

In summary, we observed that on average both pollutant groups increased ROS concentrations and all damage products and significantly altered 8 of the 19 ROS responses. We found that on average, exposure to NMP induced a significant increase in the ROS concentration (mean logROM ± sd: 0.173 ± 0.246, Z = 3.265, $P = 0.001$), CAT mRNA expression level (0.354 ± 0.350, Z = 4.444, $P < 0.001$), GSH concentration (0.184 ± 0.591, Z = 2.151, $P = 0.031$), MDA concentration (0.156 ± 0.385, Z = 3.427, $P < 0.001$), and 8-OHdG concentration (0.199 ± 0.105, Z = 3.290, $P = 0.001$) (**Fig. 1**, **Supplementary Table S3**). In contrast, the GST mRNA expression level (-0.178 ± 0.165, Z = -1.996, $P = 0.046$) decreased on average. Neonic exposure increased ROS concentration (mean logROM ± sd: 0.183 ± 0.130, Z = 4.462, $P < 0.001$), $^{\bullet}OH$ concentration (0.177 ± 0.157, Z = 2.244, $P = 0.025$), CAT activity (0.096 ± 0.348, Z = 2.548, $P = 0.011$), SOD activity (0.148± 0.353, Z = 3.837, $P < 0.001$), SOD mRNA expression (0.399 ± 0.594, Z = 5.790, $P < 0.001$), CAT mRNA expression

(0.182 ± 0.595, Z = 2.648, *P* = 0.008), MDA concentration (0.207 ± 0.250, Z = 5.241, *P* < 0.001), 8-OHdG concentration (0.247 ± 0.190, Z = 3.432, *P* = 0.001), and PC concentration (0.166± 0.079, Z = 2.994, *P* = 0.003) (**Supplementary Table S4**). Only the CarE activity (-0.241± 0.466, Z = -3.173, *P* = 0.002) was significantly decreased under exposure to neonics. So on average, both pollutant groups (NMP and neonics) increased ROS levels and damage product concentrations and altered the gene expression levels of many of the analysed enzymes and enzyme activities, which is consistent with results from previous meta-analyses examining the effects of NMP [33] and neonics [18] on soil biota.

3.2 Sample sizes assuming a general stress response

We observed large variances and frequent overlap of observed data ranges with zero (**Fig. 1**) as well as small sample sizes in the compiled literature (maximum sample size: N = 8, median: N = 3). Small effect size, high variance and low sample size both decrease the statistical power. Therefore, our findings raise the question of whether sample sizes of experiments commonly conducted in individual studies are large enough to provide sufficient statistical power.

To investigate statistical power in more detail, we conducted simulation experiments and power analyses based on the means and standard deviations of measured effects for the control (without pollutant) and treatment group (with pollutant) (see section 2.3). To this end, we assumed the difference between control and treatment means as true effects for each ROS marker and pollutant group. For the sample sizes reported in the compiled literature (median of all 45 articles: N = 3), our simulations showed very low statistical power (**Supplementary Table S5)** with an average power of 0.085 ± 0.102 (mean ± sd) and 0.065 ± 0.069 over all ROS markers for NMP and neonics, respectively. This means that effects that are truly there would be detected in less than ten percent of the experiments. At a sample size of N = 8 (maximum sample size reported in the compiled literature), the average power for NMP and neonic ROS markers was 0.220 ± 0.300 and 0.163 ± 0.229, respectively (**Supplementary Fig. S4, Supplementary Fig. S5**). Usually, a power of 0.8 is recommended for experimental design [34,35]. We thus used additional

simulations to investigate, whether sufficient power could be achieved by increasing sample sizes up to a maximum of 100. Most ROS markers (17 out of 27) did not reach a power of 80% with a sample size of 100 (**Supplementary Table S5)**. This indicates that, assuming a generalized ROS-related stress response (similar true effects across all NMP and neonics, respectively), the effects of exposure to NMP and neonics on ROS markers are hardly detectable by single experimental studies.

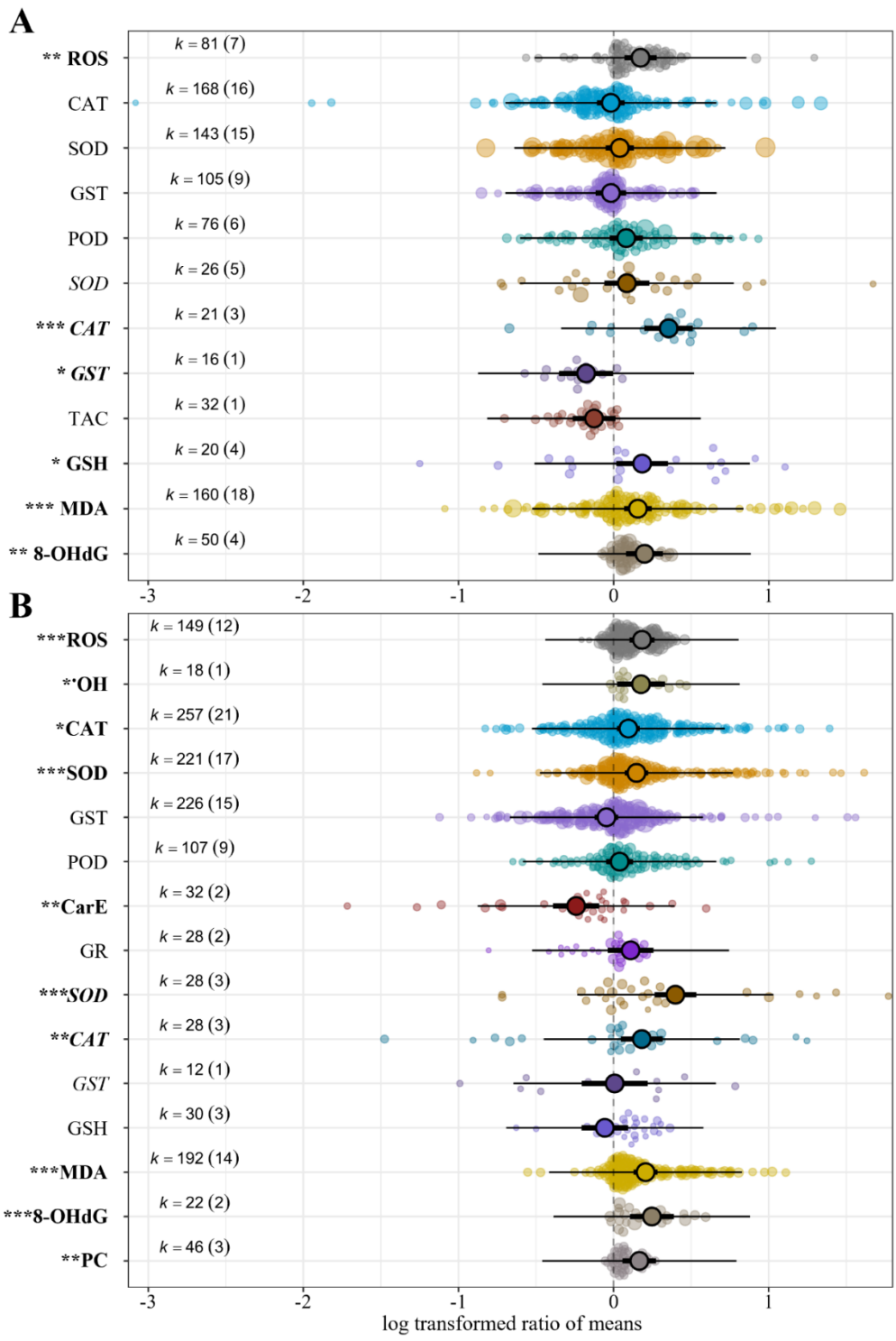


Fig. 1 Nano- and microplastic (NMP) and neonicotinoids (neonics) induce the formation of reactive oxygen species (ROS) and alter ROS responses on average, but with large variances. Overall effects of NMP (**A**) and neonics (**B**) on the generation of ROS, ROS-scavenging enzyme activities and expression, and ROS-associated damage products. K: number of data points (number of studies in brackets). The black circle displays the mean, thicker black lines show 95% confidence intervals, and narrow lines show prediction intervals. These estimates were calculated using mixed meta regression models without intercept for each pollutant separately. Point sizes correlate with inverse standard errors. ROS markers in bold significantly differ from zero (no effect); *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. Italic ROS markers represent mRNA expression levels of respective enzymes. ROS: reactive oxygen species concentration; •OH: hydroxyl concentration; CAT: catalase activity; SOD: superoxide dismutase activity; GST: glutathione-S-transferase activity; POD: peroxidase activity; CarE: carboxylesterase activity; GR: glutathione reductase activity; TAC: total antioxidant capacity; GSH: glutathione concentration; MDA: malondialdehyde concentration; 8-OHdG: 8-hydroxy-2-deoxyguanosine concentration; PC: protein carbonyl group concentration.

3.3 Influence of experimental design and pollution properties

A potential explanation for the high data spread in the measured ROS markers is a differential stress response with effects being moderated by differences in the pollutants the organisms are exposed to or by variations resulting from differences in the experimental designs. An indication for such a differential stress response is the high residual heterogeneity we found in the meta-analytical models (Test of residual heterogeneity: NMP: QE = 969050.145, $P < 0.001$; neonics: QE = 153342.470, $P < 0.001$). As a next step, we thus examined whether differences in effects are associated with different exposure designs, particle characteristics, and different neonicotinoid identities. Effects on life-history traits of several soil-dwelling organisms have been shown to be dependent on NMP properties [36], while different neonics have different effects on the reproduction and survival of various soil animals at similar concentrations [37]. As moderators of NMP effects, we considered exposure time, concentration, particle shape, particle size, and polymer type in our analysis. As moderators of neonic effects, we included exposure time, concentration, and the neonic identity.

We used two approaches to examine whether each property contributed to the observed variance. First, we investigated differences among different pollutant properties and experimental parameters across all pooled ROS markers. To this end, we fitted a full model containing all moderators as fixed effects and used likelihood ratio tests (LRT) to compare this model to different reduced models, where one moderator at a time was excluded. For this analysis, we assumed that effects on different ROS markers are similar. This assumption could be too simplistic however, since a pollutant could have different effects on different ROS markers. For example, higher NMP concentrations can increase damage product concentrations such as MDA but decrease the activity of multiple ROS scavenging enzymes e.g., CAT, SOD, and GST [38]. So, in a second approach, we additionally considered differences in effects among different ROS markers. To this end, we included interaction terms for each moderator with ROS markers to the model structure. Again, we compared the full model with all interactions to reduced models, where one moderator-ROS marker interaction at a time was excluded. In addition, we used the

variance components σ^2 of the models to calculate the amount of the variance in the data that can be explained by the pollutants' properties and experimental design parameters.

For models without interactions, the experimental exposure time (LR = 16.276, $P < 0.001$) and the NMP properties polymer type (LR = 38.465, $P < 0.001$), shape (LR = 12.634, $P = 0.027$), and size (LR = 6.312, $P = 0.010$) explained a considerable amount of the variability, while concentration (LR = 2.023, $P = 0.154$) did not. In contrast, for neonics only the neonic concentration (LR = 26.409, $P < 0.001$) explained a significant amount of variability, while exposure time (LR = 0.090, $P = 0.764$) and the neonic identity (LR = 21.461, $P = 0.161$) did not.

When including interaction terms, we found that different NMP concentrations (LR = 34.730, $P < 0.001$), exposure times (LR = 32.457, $P < 0.001$), and polymers (LR = 30.043, $P = 0.026$) had different effects among the ROS markers, but different shapes (LR = 21.950, $P = 0.234$) and sizes (LR = 3.564, $P = 0.468$) did not. Similarly, different neonic concentrations (LR = 59.767, $P < 0.001$), exposure times (LR = 35.283, $P = 0.001$), and neonic identities (LR = 345.732, $P < 0.001$) differed in the patterns observed for different ROS markers. These results indicate that pollutant properties and experimental design parameters explain at least part of the observed variance in our data (Test of moderators of full model with interactions: NMP: $QM_{91} = 427.725$, $P < 0.001$; neonics: $QM_{127} = 738.941$, $P < 0.001$). In total, pollutant properties and experimental design parameters accounted for 46.7 % (NMP dataset) and 21.2 % (neonic dataset) of the variance in the data. An in-depth analysis of how experimental designs and pollutant properties affect individual ROS markers are reported in **Supplementary Fig. S6 - S13**, and **Supplementary Tab. S6 - S13**.

Our results indicate that ROS responses are strongly dependent on the pollutant properties, their chemical identity, and experimental design parameters. This complicates generalizations and predictions of effects of pollutants on ROS markers. Interestingly, only neonic concentration, but not NMP concentration, explained a significant amount of variance in our data. This is surprising since other meta-analyses found concentration dependent effects of NMP on survival, growth

rate, and reproduction of soil biota [33,39]. However, our analyses revealed significant interactions of both pollutant properties (e.g., NMP concentration) and experimental design parameters with ROS markers indicating that effects of these parameters act stronger on certain ROS markers than on others (see **Supplementary Fig. S6 - S13**). In addition, pollutant parameters may not only interact with ROS markers but also with each other, potentially leading to additional confounding effects. A more in-depth mechanistic understanding is thus needed to explain these differences and allow to extrapolate from tested pollutant-ROS marker combinations to untested ones.

How different pollutant properties and experimental designs affect ROS formation, subsequent damage product formation, and ROS responses is still unclear. Neonics seem to induce ROS generation by disrupting intracellular Ca^{2+}-homeostasis [21], but why different neonics differ in their effects on ROS markers remains understudied. For NMP, few hypotheses have been formulated on how these particles can induce ROS and our knowledge about mediating effects remains fragmentary to date. For instance, Qiao and colleagues suggested a general activation of the immune system, which in turn may lead to increased ROS responses [15], while Hu and Palić hypothesized that ROS formed during the environmental degradation of NMP might be taken up by organisms together with the particles and thus lead to higher ROS levels in these organisms [40]. In addition, surface modifications of NMP including surface morphology and surface charge were shown to significantly affect particle-immune cell interactions [41,42], which might give first insights into moderating effects of these particle properties on ROS related responses. Further research investigating mechanistic pathways leading to and moderating ROS responses will be valuable to get a better understanding of the observed patterns.

3.4 Sample sizes depending on properties

Since the pollutant-induced effects on ROS generation, ROS responses, and damage product formation seem to depend on experimental design and partially on the pollutants' properties, our first power analysis (which assumed a generalized ROS response) is likely too simplistic and may

have overestimated sample variance and consequently underestimated the statistical power of the test system. We thus performed more detailed data simulations and power analyses that consider differences in effects based on pollutant properties and experimental designs. To this end, we calculated the means and standard deviations of the measured ROS markers for each combination of moderators (NMP: exposure duration, concentration, polymer, size, shape (813 unique combinations); neonic: exposure duration, concentration, neonic identity (1241 unique combinations)). These values were used as a basis for extensive data simulations and the calculation of statistical power dependent on sample size for each combination separately. As expected, for most ROS markers, statistical power at fixed sample sizes was notably higher when all moderators were considered. For 21 out of the 27 ROS markers, sample sizes smaller than 20 were sufficient to reach a power of at least 0.8, and 8 ROS markers required sample sizes smaller than ten (**Tab. *2***). For example, based on our simulations, sample sizes of N = 3 and N = 5 are sufficient to achieve significant results in over 80 % of experiments when measuring CAT or SOD mRNA expression after NMP exposure (**Supplementary Fig. S14**). For neonics, only the CarE activity, GR activity, and GSH concentration showed rather low power, not reaching the 0.8 threshold with a sample size of N ≤ 50 (**Supplementary Fig. S15**).

Tab. 2 Most reactive oxygen species (ROS) measurements have high enough power at sample sizes lower than 20. Smallest sample size (N) needed to achieve statistical power of at least 0.8 and statistical power of ROS markers at N = 3 (most used sample size) and N = 8 (highest used sample size) after nano- and microplastic (NMP) or neonicotinoid (neonic) exposure in experiments investigating pollutant effects on ROS markers in annelid worms. Means and standard deviations for each property and experimental design combination were used as a basis for data and power simulations. NA: no data on this ROS marker available. ROS: reactive oxygen species concentration; •OH: hydroxyl concentration; CAT: catalase activity; SOD: superoxide dismutase activity; GST: glutathione-S-transferase activity; POD: peroxidase activity; CarE: carboxylesterase activity; GR: glutathione reductase activity; TAC: total antioxidant capacity; GSH: glutathione concentration; MDA: malondialdehyde concentration; 8-OHdG: 8-hydroxy-2-deoxyguanosine concentration; PC: protein carbonyl group concentration.

	NMP			neonics		
ROS marker	smallest N with power ≥ 0.8	power at N = 3	power at N = 8	smallest N with power ≥ 0.8	power at N = 3	power at N = 8
ROS concentration	12	0.57	0.76	11	0.63	0.78
•OH concentration	NA	NA	NA	10	0.44	0.61
CAT activity	16	0.56	0.73	14	0.51	0.74
SOD activity	8	0.63	0.8	12	0.51	0.75
GST activity	9	0.59	0.79	24	0.53	0.7
POD activity	12	0.55	0.76	23	0.42	0.66
CarE activity	NA	NA	NA	> 50	0.48	0.62
GR activity	NA	NA	NA	> 50	0.26	0.45
SOD mRNA expression	5	0.69	0.89	10	0.59	0.79
CAT mRNA expression	3	0.88	0.91	17	0.63	0.73
GST mRNA expression	6	0.61	0.85	13	0.45	0.71
TAC	15	0.54	0.77	NA	NA	NA
GSH concentration	9	0.63	0.78	> 50	0.19	0.41
MDA concentration	15	0.52	0.75	13	0.56	0.76
8-OHdG concentration	5	0.69	0.87	8	0.62	0.82
PC concentration	NA	NA	NA	25	0.38	0.61

Statistical power depends on sample size, true effect size, type I error rate (i.e., chosen alpha level) and the variance of the data [12]. While the true effect size is fixed by the studied system and the type I error rate is usually fixed by the research community, sample size and variance can be at least partially adjusted. Considering differential ROS responses dependent on differences in NMP properties and neonic identity, the effects of pollutants on ROS generation, ROS responses, and damage products can likely be measured reliably with reasonably large sample sizes (at least N = 5 and N = 10 for most ROS markers in NMP and neonics, respectively). To further increase power, the variance in the data could be decreased by reducing measurement error. For example, Murphy and colleagues argued that measurement techniques frequently used for measuring ROS markers are often not optimal [1]. Using enzyme-linked immunosorbent assay (ELISA) to measure 8-OHdG concentrations [43] or MDA to examine lipid peroxidation [44] lack specificity compared to more sophisticated methods like UPLC-MS/MS. Our data indicate that

442 expression measurements via qPCR require smaller sample sizes compared to other ROS
443 makers, especially after NMP exposure. Improved methodology could thus help to strengthen
444 statistical power where adequate sample sizes are difficult to achieve.

445 We investigated how NMP and neonics affect ROS markers in terrestrial annelids. However,
446 caution should be taken when applying our insights to other organisms in other ecosystems. Since
447 the underlying mechanisms of how these pollutants induce ROS are not well understood,
448 especially for NMP, it is difficult to estimate how well our results can be generalized to other
449 organisms. For example, if ROS are primarily produced by immune cells in response to pollutant
450 exposure, the effects of these pollutants could differ greatly between organisms with different
451 immune systems, such as annelids and vertebrates. On the other hand, property dependent
452 effects were observed in other organisms as well, e.g., polymer dependent effects on bacterial
453 and fungal diversity [39]. Therefore, the influence of pollutant properties and experimental design
454 on ROS markers probably applies to other taxa as well. While we have mainly analysed the effects
455 of pristine NMP, we have not considered the presence of additives or the numerous changes to
456 NMP due to environmental factors. Plastic materials contain on average about 20 additives
457 including among others antioxidants, plasticizers, and flame retardants [45] which can have
458 additional adverse effects [16]. Moreover, abiotic factors such as photooxidation via sunlight can
459 lead to changes in the particles' surface structure and increase the number of functional groups
460 on the surface [46], which could explain the greater toxicity of aged compared to pristine NMP in
461 some studies [47]. Finally, the scarce mechanistic understanding of how pollutants influence
462 individual ROS markers limits the ability to test for antagonistic and synergistic effects among
463 ROS markers. For example, it is uncertain whether the increase in CAT and SOD activity ultimately
464 led to lower amounts of MDA, 8-OHdG, and PC after neonic exposure, since these enzymes
465 potentially decomposed ROS before they reacted and produced new damage products.

466 **4. Conclusion and future prospects**

Our results show that NMP and neonic exposure impact multiple ROS markers, such as increasing ROS and damage product concentrations, in terrestrial annelids. The effects are however heavily influenced by multiple parameters of experimental design and pollutant properties and often strongly overlap with zero. Singe experiments with low sample size will therefore potentially find it difficult to obtain sufficient statistical power to detect these effects.

Based on these results, we identified three major improvements for future studies when analysing the impacts of pollutants on ROS markers. First, authors should report their methods (e.g., the specific ROS marker, experimental design choices, and pollutant properties) with as much detail as possible [48]. If possible, materials, methods, and results should additionally be reported in a machine-readable format and uploaded to an online data repository which follows the FAIR principles (e.g., Zenodo, Open Science Framework) and additionally published as online supplemental material.

Second, to understand how ROS responses are moderated by pollutant properties, more detailed knowledge of molecular mechanisms leading to these differential responses and the further investigation of adverse outcome pathways are necessary. Dedicated experiments are needed to clarify the mechanisms by which pollutants trigger effects.

Third, caution should be exercised when interpreting studies that use small sample sizes and suboptimal measurement techniques since they most likely possess insufficient statistical power. Instead, more emphasis should be placed on increasing the statistical power of individual experiments by increasing the sample size, using more precise measurement techniques such as qPCR or LCMS, or both. Based on the data we compiled, we believe that a minimum of five replicates for NMP and ten replicates for neonic studies seem to be necessary to achieve sufficient power. Ideally, a priori power analyses based on lab-specific experimental conditions and variance estimates derived from previous experiments should be conducted to ensure that sample sizes meet statistical power requirements (i.e., a statistical power of at least

492 80%). This way, true effects can be detected at higher rates, and true effect sizes and potential
493 risks to environmental health can be estimated more accurately.

494 **Acknowledgements**

495 We want to thank Otto Schmölder for helping with the data extraction, Wolfgang Viechtbauer for
496 giving advice on parts of the data analysis, and Matthias Schott, Lisa Hülsmann, and Holger Kress
497 for commenting on earlier versions of the manuscript.

498 **Author's contributions**

499 **Max Döring:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology,
500 Software, Visualization, Writing – original draft, Writing – review and editing. **Heike Feldhaar:**
501 Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing.
502 **Ana Antonio Vital**: Data curation, Validation, Writing – review and editing. **Magdalena Mair:**
503 Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Supervision,
504 Writing – review and editing.

505 **Funding**

506 This study was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research
507 Foundation) – SFB 1357 Mikroplastik – Project Number 391977956.

508 **Competing interests**

509 The authors declare no competing interests.

510 **Data availability**

511 All data and code pertaining to this manuscript will be made available upon publication via
512 Zenodo and github.

513

514 **References**

[1] M.P. Murphy, H. Bayir, V. Belousov, C.J. Chang, K.J.A. Davies, M.J. Davies, T.P. Dick, T. Finkel, H.J. Forman, Y. Janssen-Heininger, D. Gems, V.E. Kagan, B. Kalyanaraman, N.-G. Larsson, G.L. Milne, T. Nyström, H.E. Poulsen, R. Radi, H. Van Remmen, P.T. Schumacker, P.J. Thornalley, S. Toyokuni, C.C. Winterbourn, H. Yin, B. Halliwell, 2022. Guidelines for measuring reactive oxygen species and oxidative damage in cells and in vivo. Nat Metab 4, 651–662. https://doi.org/10.1038/s42255-022-00591-z.

[2] G. Loschen, A. Azzi, C. Richter, L. Flohé, 1974. Superoxide radicals as precursors of mitochondrial hydrogen peroxide. FEBS Lett 42, 68–72. https://doi.org/10.1016/0014-5793(74)80281-4.

[3] M.P. Murphy, 2009. How mitochondria produce reactive oxygen species. Biochem J 417, 1–13. https://doi.org/10.1042/BJ20081386.

[4] K. Bedard, K.-H. Krause, 2007. The NOX family of ROS-generating NADPH oxidases: physiology and pathophysiology. Physiol Rev 87, 245–313. https://doi.org/10.1152/physrev.00044.2005.

[5] H. Sies, D.P. Jones, 2020. Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. Nat Rev Mol Cell Biol 21, 363–383. https://doi.org/10.1038/s41580-020-0230-3.

[6] A. Tauffenberger, P.J. Magistretti, 2021. Reactive oxygen species: beyond their reactive behavior. Neurochem Res 46, 77–87. https://doi.org/10.1007/s11064-020-03208-7.

[7] Y.S. Bae, H. Oh, S.G. Rhee, Y.D. Yoo, 2011. Regulation of reactive oxygen species generation in cell signaling. Mol Cells 32, 491–509. https://doi.org/10.1007/s10059-011-0276-3.

[8] R.M. Day, Y.J. Suzuki, 2005. Cell proliferation, reactive oxygen and cellular glutathione. Dose-Response 3, 425–442. https://doi.org/10.2203/dose-response.003.03.010.

[9] D. Martins, A.M. English, 2014. Catalase activity is stimulated by H2O2 in rich culture medium and is required for H2O2 resistance and adaptation in yeast. Redox Biol 2, 308–313. https://doi.org/10.1016/j.redox.2013.12.019.

[10] T.L. De Jager, A.E. Cockrell, S.S. Du Plessis, 2017. Ultraviolet light induced generation of reactive oxygen species. In: S.I. Ahmad (Ed.) Ultraviolet Light in Human Health, Diseases and Environment. Springer International Publishing, Cham, pp 15–23. https://doi.org/10.1007/978-3-319-56017-5_2.

[11] Z. Leni, L. Künzi, M. Geiser, 2020. Air pollution causing oxidative stress. Curr Opin Toxicol 20–21, 1–8. https://doi.org/10.1016/j.cotox.2020.02.006.

[12] K.R. Murphy, B. Myors, A. Wolach, 1998. Statistical power analysis: a simple and general model for traditional and modern hypothesis tests. Choice Rev Online 36, 36-1020-36–1020. https://doi.org/10.5860/CHOICE.36-1020.

[13] C.A. Edwards, N.Q. Arancon, 2022. Biology and ecology of earthworms. Springer US, New York, NY. https://doi.org/10.1007/978-0-387-74943-3.

[14] R.C. Thompson, W. Courtene-Jones, J. Boucher, S. Pahl, K. Raubenheimer, A.A. Koelmans, 2024. Twenty years of microplastics pollution research—what have we learned? Science eadl2746. https://doi.org/10.1126/science.adl2746.

[15] R. Qiao, C. Sheng, Y. Lu, Y. Zhang, H. Ren, B. Lemos, 2019. Microplastics induce intestinal inflammation, oxidative stress, and disorders of metabolome and microbiome in zebrafish. Sci Total Environ 662, 246–253. https://doi.org/10.1016/j.scitotenv.2019.01.245.

[16] Y. Xue, X. Gu, X. Wang, C. Sun, X. Xu, J. Sun, B. Zhang, 2009. The hydroxyl radical generation and oxidative stress for the earthworm *Eisenia fetida* exposed to tetrabromobisphenol A. Ecotoxicology 18, 693–699. https://doi.org/10.1007/s10646-009-0333-2.

[17] P. Jeschke, R. Nauen, M. Schindler, A. Elbert, 2011. Overview of the status and global strategy for neonicotinoids. J Agric Food Chem 59, 2897–2908. https://doi.org/10.1021/jf101303g.

[18] Y. Cao, W. Zhao, J. Zhang, D.M. Figueiredo, M. Zhao, S. Ren, H. Mu, Y. Li, H. Lu, H. Shi, X. Li, J. Li, F. Zhao, J. Han, K. Wang, 2024. Effects of neonicotinoid residues on non-target soil animals: a case study of meta-analysis. J Hazard Mater 476, 135022. https://doi.org/10.1016/j.jhazmat.2024.135022.

[19] X. Xu, X. Wang, Y. Yang, I. Ares, M. Martínez, B. Lopez-Torres, M.-R. Martínez-Larrañaga, X. Wang, A. Anadón, M.-A. Martinez, 2022. Neonicotinoids: mechanisms of systemic toxicity based on oxidative stress-mitochondrial damage. Arch Toxicol 96, 1493–1520. https://doi.org/10.1007/s00204-022-03267-5.

[20] S. Li, Y. Cao, Q. Pan, Y. Xiao, Y. Wang, X. Wang, X. Li, Q. Li, X. Tang, B. Ran, 2021. Neonicotinoid insecticides triggers mitochondrial bioenergetic dysfunction via manipulating ROS-calcium influx pathway in the liver. Ecotoxicol Environ Saf 224, 112690. https://doi.org/10.1016/j.ecoenv.2021.112690.

[21] F. Wei, F. Cheng, H. Li, J. You, 2024. Imidacloprid affects human cells through mitochondrial dysfunction and oxidative stress. Sci Total Environ 951, 175422. https://doi.org/10.1016/j.scitotenv.2024.175422.

[22] J. Li, C. Zhu, Y. Xu, H. He, C. Zhao, F. Yan, 2024. Molecular mechanism underlying ROS-mediated AKH resistance to imidacloprid in whitefly. Insects 15, 436. https://doi.org/10.3390/insects15060436.

[23] W. Viechtbauer, 2010. Conducting meta-analyses in R with the metafor package. J Stat Softw 36, 1–48. https://doi.org/10.18637/jss.v036.i03.

[24] F. Sommer, V. Dietze, A. Baum, J. Sauer, S. Gilge, C. Maschowski, R. Gieré, 2018. Tire abrasion as a major source of microplastics in the environment. Aerosol Air Qual Res 18, 2014–2028. https://doi.org/10.4209/aaqr.2018.03.0099.

[25] J.L. Pick, S. Nakagawa, D.W.A. Noble, 2018. Reproducible, flexible and high-throughput data extraction from primary literature: the metaDigitise r package. Methods Ecol Evol 10, 426–431. https://doi.org/10.1111/2041-210X.13118.

[26] R Core Team, 2023. R: A language and environment for statistical computing. https://www.r-project.org/.

[27] A. Rodríguez-Seijo, J.P. Da Costa, T. Rocha-Santos, A.C. Duarte, R. Pereira, 2018. Oxidative stress, energy metabolism and molecular responses of earthworms (*Eisenia fetida*) exposed to low-density polyethylene microplastics. Environ Sci Pollut Res 25, 33599–33610. https://doi.org/10.1007/s11356-018-3317-z.

596 [28] A. Rodríguez-Seijo, B. Santos, E. Ferreira Da Silva, A. Cachada, R. Pereira, 2018. Low-density
597 polyethylene microplastics as a source and carriers of agrochemicals to soil and
598 earthworms. Environ Chem 16, 8. https://doi.org/10.1071/EN18162.

599 [29] J.A. Buege, S.D. Aust, 1978. [30] Microsomal lipid peroxidation. In: S. Fleischer, L. Packer,
600 (Eds.) Methods in Enzymology. Elsevier, pp 302–310. https://doi.org/10.1016/S0076-
601 6879(78)52032-6.

602 [30] H. Wickham, R. François, L. Henry, K. Müller, D. Vaughan, 2023. dplyr: a grammar of data
603 manipulation. https://CRAN.R-project.org/package=dplyr.

604 [31] S. Nakagawa, M. Lagisz, R.E. O'Dea, P. Pottier, J. Rutkowska, A.M. Senior, Y. Yang, D.W.A.
605 Noble, 2023. orchaRd 2.0: an R package for visualising meta-analyses with orchard plots.
606 Methods Ecol Evol 14, 2003–2010. https://doi.org/10.1111/2041-210X.14152.

607 [32] H. Wickham, 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York.
608 https://doi.org/10.1007/978-3-319-24277-4.

609 [33] P. Su, J. Wang, D. Zhang, K. Chu, Y. Yao, Q. Sun, Y. Luo, R. Zhang, X. Su, Z. Wang, N. Bu, Z. Li,
610 2022. Hierarchical and cascading changes in the functional traits of soil animals induced by
611 microplastics: a meta-analysis. J Hazard Mater 440, 129854.
612 https://doi.org/10.1016/j.jhazmat.2022.129854.

613 [34] J. Cohen, 1962. The statistical power of abnormal-social psychological research: a review. J
614 Abnorm Soc Psychol 65, 145–153. https://doi.org/10.1037/h0045186.

615 [35] G.W. Oehlert, P. Whitcomb, 2001. Sizing fixed effects for computing power in experimental
616 designs. Qual Reliab Eng Int 17, 291–306. https://doi.org/10.1002/qre.402.

617 [36] M. Liu, J. Feng, Y. Shen, B. Zhu, 2023. Microplastics effects on soil biota are dependent on
618 their properties: a meta-analysis. Soil Biol Biochem 178, 108940.
619 https://doi.org/10.1016/j.soilbio.2023.108940.

620 [37] M. Renaud, T. Akeju, T. Natal-da-Luz, S. Leston, J. Rosa, F. Ramos, J.P. Sousa, H.M.V.S.
621 Azevedo-Pereira, 2018. Effects of the neonicotinoids acetamiprid and thiacloprid in their
622 commercial formulations on soil fauna. Chemosphere 194, 85–93.
623 https://doi.org/10.1016/j.chemosphere.2017.11.102.

624 [38] F. He, H. Shi, R. Liu, G. Tian, Y. Qi, T. Wang, 2023. Randomly-shaped nanoplastics induced
625 stronger biotoxicity targeted to earthworm *Eisenia fetida* species: differential effects and the
626 underlying mechanisms of realistic and commercial polystyrene nanoplastics. Sci Total
627 Environ 877, 162854. https://doi.org/10.1016/j.scitotenv.2023.162854.

628 [39] L. Wan, H. Cheng, Y. Liu, Y. Shen, G. Liu, X. Su, 2023. Global meta-analysis reveals
629 differential effects of microplastics on soil ecosystem. Sci Total Environ 867, 161403.
630 https://doi.org/10.1016/j.scitotenv.2023.161403.

631 [40] M. Hu, D. Palić, 2020. Micro- and nano-plastics activation of oxidative and inflammatory
632 adverse outcome pathways. Redox Biol 37, 101620.
633 https://doi.org/10.1016/j.redox.2020.101620.

634 [41] A.F.R.M. Ramsperger, J. Jasinski, M. Völkl, T. Witzmann, M. Meinhart, V. Jérôme, W.P.
635 Kretschmer, R. Freitag, J. Senker, A. Fery, H. Kress, T. Scheibel, C. Laforsch, 2022.
636 Supposedly identical microplastic particles substantially differ in their material properties

637 influencing particle-cell interactions and cellular responses. J Hazard Mater 425, 127961.
638 https://doi.org/10.1016/j.jhazmat.2021.127961.

639 [42] S. Wieland, A.F.R.M. Ramsperger, W. Gross, M. Lehmann, T. Witzmann, A. Caspari, M. Obst,
640 S. Gekle, G.K. Auernhammer, A. Fery, C. Laforsch, H. Kress, 2024. Nominally identical
641 microplastic models differ greatly in their particle-cell interactions. Nat Commun 15, 922.
642 https://doi.org/10.1038/s41467-024-45281-4.

643 [43] T. Henriksen, A. Weimann, E.L. Larsen, H.E. Poulsen, 2021. Quantification of 8-oxo-7,8-
644 dihydro-2'-deoxyguanosine and 8-oxo-7,8-dihydro-guanosine concentrations in urine and
645 plasma for estimating 24-h urinary output. Free Radic Biol Med 172, 350–357.
646 https://doi.org/10.1016/j.freeradbiomed.2021.06.014.

647 [44] H. Yin, N.A. Porter, 2003. Specificity of the ferrous oxidation of xylenol orange assay: analysis
648 of autoxidation products of cholesteryl arachidonate. Anal Biochem 313, 319–326.
649 https://doi.org/10.1016/S0003-2697(02)00621-8.

650 [45] L. Van Oers, E. Van Der Voet, V. Grundmann, 2011. Additives in the plastics industry. In: B.
651 Bilitewski, R.M. Darbra, D. Barceló (Eds.) Global Risk-Based Management of Chemical
652 Additives I. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 133–149.

653 [46] N. Meides, T. Menzel, B. Poetzschner, M.G.J. Löder, U. Mansfeld, P. Strohriegl, V. Altstaedt, J.
654 Senker, 2021. Reconstructing the environmental degradation of polystyrene by accelerated
655 weathering. Environ Sci Technol 55, 7930–7938. https://doi.org/10.1021/acs.est.0c07718.

656 [47] X. Jiang, J. Cao, Z. Ye, G. Klobučar, M. Li, 2023. Microplastics - back to reality: impact of
657 pristine and aged microplastics in soil on earthworm *Eisenia fetida* under environmentally
658 relevant conditions. Environ Sci Technol 57, 16788–16799.
659 https://doi.org/10.1021/acs.est.3c04097.

660 [48] V.N. De Ruijter, P.E. Redondo-Hasselerharm, T. Gouin, A.A. Koelmans, 2020. Quality criteria
661 for microplastic effect studies in the context of risk assessment: a critical review. Environ
662 Sci Technol 54, 11692–11705. https://doi.org/10.1021/acs.est.0c03057.

663