Accepted Manuscript

Title: A bacterium-based contact assay for evaluating the quality of solid samples – results from an international ring test

Authors: Catarina R. Marques, Najoi El-Azhari, Fabrice Martin-Laurent, Pascal Pandard, Camille Meline, Alice L. Petre, Suzanne Eckert, Jürgen Zipperle, Martin Váňa, Stanislav Maly, Lucie Šindelářová, Anna Slavíková Amemori, Jakub Hofman, Anu Kumar, Hai Doan, Mike McLaughlin, Elizabeth Richter, Jörg Römbke

PII: S0304-3894(18)30174-2
DOI: https://doi.org/10.1016/j.jhazmat.2018.03.022
Reference: HAZMAT 19242

To appear in: *Journal of Hazardous Materials*

Received date: 1-10-2017
Revised date: 5-3-2018
Accepted date: 13-3-2018

Please cite this article as: Marques CR, El-Azhari N, Martin-Laurent F, Pandard P, Meline C, Petre AL, Eckert S, Zipperle J, Váňa M, Maly S, Šindelářová L, Amemori AS, Hofman J, Kumar A, Doan H, McLaughlin M, Richter E, Römbke J, A bacterium-based contact assay for evaluating the quality of solid samples – results from an international ring test, *Journal of Hazardous Materials* (2010), https://doi.org/10.1016/j.jhazmat.2018.03.022

A bacterium-based contact assay for evaluating the quality of solid samples – results from an international ring test

Catarina R. Marques[1,*], Najoi El-Azhari[2], Fabrice Martin-Laurent[3], Pascal Pandard[4], Camille Meline[4], Alice L. Petre[5], Suzanne Eckert[6], Jürgen Zipperle[6], Martin Váňa[7], Stanislav Maly[7], Lucie Šindelářová[8], Anna Slavíková Amemori[8], Jakub Hofman[8], Anu Kumar[9], Hai Doan[9], Mike McLaughlin[9], Elizabeth Richter[10], Jörg Römbke[11]

[1]Departamento de Biologia & Centro de Estudos do Ambiente e do Mar (CESAM), Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

[2]SATT Grand Est, Welience Agro-Environnement, F-21000 Dijon, France

[3]Agroécologie, AgroSup Dijon, INRA, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France

[4]INERIS, Parc Technologique Alata, Verneuil-en-Halatte, France

[5]Analytical Chemistry, Physical Chemistry and Chemical Engineering Department, Universidad de Alcalá, 28871, Madrid, Spain

[6]LUBW Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg, Griesbachstr. 1 76185 Karlsruhe, Germany

[7]Central Institute for Supervising and Testing in Agriculture, Department of Microbiology and Biochemistry, Hroznová 2, Brno, CZ-656 06, Czech Republic

[8]Research centre for toxic compounds in the environment (RECETOX), Faculty of Science, Masaryk University, Kamenice 753/5, Brno, 625 00 Czech Republic

[9]CSIRO Land and Water Reception, Gate 4, Waite Road, SA 5064 Urrbrae, Australia

[10]Goethe University Frankfurt am Main, Institute for Ecology, Evolution and Diversity, Department Aquatic Ecotoxicology, Frankfurt am Main, Germany.

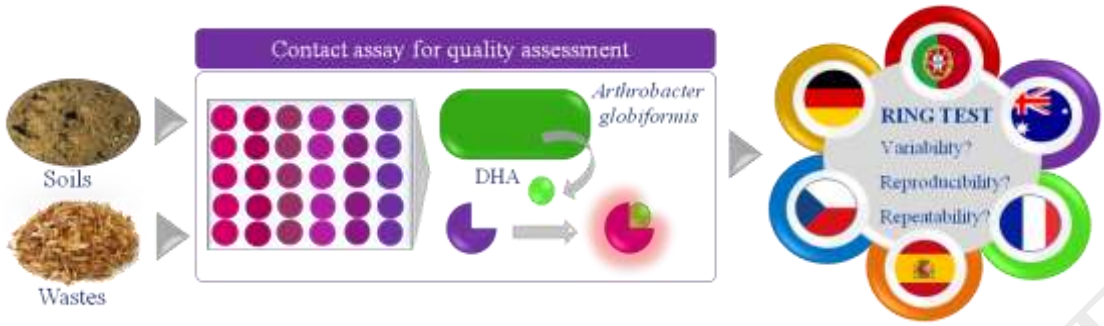
[11]ECT Oekotoxikologie GmbH, Böttgerstraße 2–14, D-65439 Flörsheim a.M., Germany

*Corresponding author: crmarques@ua.pt

T: +351 234 370 773

Fax: +351 234 372 587

Graphical abstarct



HIGHLIGHTS

- *A. globiformis* DHA is a sensitive endpoint for assessing the quality of solid samples
- *A. globiformis* solid contact assay proved to be reproducible and repeatable
- The assay is reliable for routine use towards soil and waste quality evaluation

Abstract

The contact assay measuring the inhibition of *Arthrobacter globiformis* dehydrogenase activity as an endpoint to evaluate the toxicity of solid samples was tested in an international ring-test to validate its performance for ISO standardization (ISO/CD 18187). This work reports the results of the ring-test involving 9 laboratories from six countries. At least 8 valid data sets were obtained for each sample and more than three quarters of the participants attained the validity criteria defined in the standard. The coefficient of variation within (CV_r) and between (CV_R) laboratories was generally on average <15% and <30% for negative and positive controls, respectively. Regarding solid samples, the laboratories provided a similar ranking of the samples based on their toxicity, despite some variation in the LOEC values. The logarithmic within-lab standard deviation <0.50 for soils and <0.25 for wastes evidenced a good repeatability. The between-lab variability assessed by a CV_R <30%, minimum-maximum factor <4 and a reproducibility standard deviation (SD_R) <0.13 for a great part of the solid samples, confirmed the test reproducibility. Overall, this assay proved to be robust, sensitive and feasible for routine use towards the quality assessment of soils and wastes.

Key-words: *Arthrobacter globiformis*; dehydrogenase activity; soils; wastes; ISO/CD 18187

Introduction

The quality assessment of soils and wastes has been recently receiving more attention, given the relevance of soil ecosystem services and their protection for different uses, as well as the need for a sustainable management of the increasingly produced solid wastes. Under this context, the evaluation of solid materials, which might be contaminated with a mixture of unknown compounds, has been broadening beyond the regular physical and chemical characterization as to include their ecotoxicological characterization [1]. This requires the development of new, easy-handled, cost-effective, rapid and sensitive methods to cover ecologically relevant terrestrial organisms that are yet to be considered for assessing the quality of solid samples. Soil microorganisms, for instance, are a key group in the terrestrial trophic chains, since they are responsible for a wide range of ecosystem functions, such as organic matter degradation, assimilation and dissimilation of N and plant-growth promotion [2,3,4]. Therefore, soil microorganisms have been often viewed as possible bioindicators of soil functioning and health [5,6]. Until now, most microbial parameters monitored in soils using ISO standards are at the community level [7], like microbial diversity [8,9], biomass [10,11], metabolic activity (*e.g.*, enzymatic activities, respiration curves [12], ammonium oxidation [13], mineralization and nitrification [14]), and abundance of microbial genes [15].

Despite the relevance of these endpoints to estimate possible impacts of stressors in the composition and normal functioning of the soil microbial community, they may provide misleading sensitivity due to community shifts, namely associated with the disappearance of sensitive populations and the dominance of tolerant ones [16]. In addition, the methods used are expensive, time-consuming and effort-demanding. Thereby, the development of standard methods to quickly assess the quality of solid samples on specific microbial populations can be an excellent and reliable alternative. Indeed, such methods are not constrained by the functional redundancy occurring at microbial community level, besides being usually rather simple, cost-effective and easily standardized. As such, they are valuable tools to support risk assessment and hazard characterization schemes applied to soils [*e.g.*,17,18] and wastes [19], as often required by risk managers, industry and academics.

In particular, *Arthrobacter globiformis* is a ubiquitous and non-pathogenic aerobic soil bacterium that synthesizes an extracellular enzyme during different metabolic processes [20], which activity was suggested to be a potential indicator of the effect of

contaminants on solid samples [21]. As such, a solid contact test based on the measurement of *A. globiformis* dehydrogenase activity (DHA) was preliminary purposed [21,22]. The principle of the assay relies on the reduction of resazurin into resorufin that is fluorimetrically detected and used as a proxy of DHA. Whenever a solid sample inhibits the *A. globiformis* DHA, the level of resorufin production and, hence, the emitted fluorescence is reduced, indicating that the sample is toxic or presents reduced quality. Several studies confirmed the bacterium sensitivity to different pollutants [*e.g.*,22,23,24], which was often more pronounced comparatively to other test methods and terrestrial test organisms [25,26]. Therefore, due to the short life-cycle, fast response, sensitivity and easy maintenance of *A. globiformis*, as well as the high surface-to-volume ratio and the requirement of small amounts of test sample, this solid contact assay was proposed for standardization to the International Organization for Standardization (ISO).

A relevant step in the standardization process is the validation of the test procedures through an international ring-test (IRT) at the Committee stage (CD) to reach the Enquiry stage (DIS) [7]. Therefore, this work aims to evaluate the within- and between-laboratories variability of the method 'Solid contact test using the dehydrogenase activity of *Arthrobacter globiformis,* ISO/CD 18187' [27], as requested by the ISO/TC 190/SC4 'Soil quality – Biological methods'. The IRT joined 9 laboratories from 6 countries to evaluate and validate the test method in what regards (i) its understandability and practicability, (ii) achievement of the validity criteria, (iii) suitability of the reference substance, (iv) sensitivity and responsiveness of the assay to different soil and waste samples, (v) repeatability of the assay, (vi) reproducibility and applicability of the assay for routine use in different laboratories and countries.

2. Material and methods

2.1. Organization and participants

The organization, scientific coordination, development of practical procedures and statistical analyses were carried out at the Department of Biology of the University of Aveiro (DBio-UA) (Aveiro, Portugal), in collaboration with ECT Oekotoxikologie GmbH (ECT). A total of 9 participants from 6 different countries: France (2 labs), Germany (2 labs), Czech Republic (2 labs), Australia (1 lab), Spain (1 lab) and Portugal (1 lab) joined the IRT held in 2014. Each participant was provided with an IRT kit constituted by vials with lyophilized *A. globiformis*, the reference substance, the control

substrates and the solid samples for testing (*i.e.*, soils and wastes). Upon reception of the IRT kit, the participants had to prepare the appropriate dilutions of the solid samples (*cf.* section 2.3) and perform the contact test as described in the draft ISO 18187. The final results were sent to DBio-UA for further treatment and analysis.

2.2. Test organism

A. globiformis (Conn 1982) Conn and Dimmick 1947 was obtained from Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSMZ) GmbH and cultured under 30±1ºC and 150 rpm into liquid media as described in the draft ISO 18187 [27]. Fresh bacterial cultures were then lyophilized into individual vials, ready to use into contact test trials. The viability of the bacterial lots prepared in Portugal was first confirmed before being sent to the participants (see section 2.4).

2.3. Solid samples, characterization and dilutions

Different soil and waste samples were used to perform the IRT (Table 1). Soil samples were collected from: a low-contaminated site (S1), a construction site (S2), an abandoned uranium mine in Portugal (S3), and a phosphogypsum deploying site in Tunisia (S4). The soils were collected, handled and stored according to the procedures outlined in ISO 10381-6 [28] and in the draft ISO 18187 [27]. The pH, conductivity [29], water holding capacity [30], organic matter [31] and silt/clay [32] contents were determined according to the respective standard procedures. Pseudo-total metal concentrations were quantified by ICP-MS (inductively coupled plasma mass spectrometry; Thermo S-quadrupole ICP-MS apparatus) after extraction in *aqua regia* [24].

The four representative waste types used were: wood waste (WOO) treated with wood preservatives containing copper (W1; [33,34]), dredge material from a harbor (W2; waste code 17 05 06); fluidized bed ash from a coal plant (W3; waste code 10 01 17); crushed glass material (W4; waste code 19 12 05). The waste samples were previously characterized, stored and prepared according to EN 14735 [35], being their main physical and chemical properties and the content of contaminants either determined by ICP-MS (following the method mentioned above), though some values were retrieved from Becker et al. [33] and the ABANDA database (organized by the German state of North-Rhine-Westphalia) [36].

The 2-mm particle size fraction of both soil and waste samples was obtained for metal analysis and/or to conduct the contact tests, being stored at 4ºC in the dark for no longer than 2 weeks, until being sent to the participants for performing the *A. globiformis* contact assay.

Since the soil and waste samples presented different levels of toxicity, a range of dilutions were prepared by each participant, by adding the appropriate mass of the respective control substrates [*i.e.*, Lufa 2.2 natural soil (Speyer, Germany) and quartz sand (QS), respectively]. For soils, the geometric dilutions series considered were G1 (100%), G2 (50%), G4 (25%), G8 (12.5%) and G16 (6.3%), except for soil S1, which was only tested at G1 since it presented low contamination/toxicity. For wastes, the evaluated dilutions in terms of sample dry mass were G2 (50%), G4 (25%), G8 (12.5%), G16 (6.3%), G32 (3.1%), although one laboratory (L4) also tested two lower dilutions [G64 (1.6%) and G128 (0.8 %)].

Table 1

2.4. Chemicals

The reference substance benzyldimethylhexadecylammonium chloride (C16-BAC; CAS #122-18-9) belonging to the family of quaternary ammonium compounds was purchased from Sigma-Aldrich, Germany. The C16-BAC solution was prepared in distilled water and used to spike Lufa 2.2 soil at 600 mg C16-BAC kg^{-1}, in order to obtain the positive control substrate. At this concentration, C16-BAC induces between 30 and 80% inhibition of *A. globiformis* DHA. All other chemicals used for preparing culture media and phosphate buffer were of high quality (microbiology-grade; 80 to >99% purity) and purchased from Sigma-Aldrich and Merck.

2.5. The contact assay

The test procedures followed those outlined in the draft ISO 18187. In brief, the test was conducted in 24-well microplates, being considered 4 replicates (*i.e.*, wells) per sample dilution, negative (Lufa 2.2 and QS) and positive control (C16-BAC-spiked Lufa 2.2) substrates. In order to validate the test for the three negative control substrates suggested in the draft ISO 18187 [27] (*i.e.*, Lufa 2.2, QS, OECD artificial soil hereinafter referred as OECD soil), additional 4 replicates for the OECD soil [37] were

also prepared. Each replicate/well contained 600 mg of pre-moistened sample/substrate. The potential DHA of native microbial community in solid samples is deactivated by keeping the microplates at 85ºC for 10 min.

A. globiformis inoculum was obtained after reconstituting the lyophilized bacterial vials. The inoculum was added to each well and the plates were incubated at 30ºC for 2 h (the contact reaction). Afterwards, the resazurine dye was added per well, being fluorimetrically followed the kinetics of its transformation into resorufin through DHA during 1 h. At the end, the test was validated if:

(i) the absolute value of the average relative fluorescence of the negative control increased by a factor ≥5,

(ii) the percentage inhibition of DHA in the positive control was between 30 and 80%,

(iii) the coefficient of variation of the relative fluorescence in the control soil was <15% (ISO 18187).

2.6. Statistical analyses

The datasets generated by each laboratory were carefully checked for their validity and were only accepted for further statistical analysis if at least one validity criterion was fulfilled, and if the procedures followed by the laboratories were within those outlined in the draft ISO 18187 [27].

The results were expressed both as slope of the relative fluorescence readings and percentage of *A. globiformis* DHA inhibition relatively to the negative control [24]. These data were then used to assess the variability of the assay in terms of its repeatability (*i.e.,* within laboratory variation) and reproducibility (*i.e.,* between laboratory variation). Such approach was conducted for the negative and positive controls, as well as for the soil and waste samples.

2.6.1. Assessing assay variability in negative and positive controls

The within-laboratory variation of *A. globiformis* response to the negative control substrates (QS, and Lufa 2.2 and OECD soils) was evaluated by the coefficient of variation (CV_r) of the slope values obtained from DHA kinetics for each laboratory and substrate.

For the positive control, the assay variability was analyzed by the CV calculated from the average % inhibition of DHA, either within (CV_r) and between (CV_R) laboratories.

2.6.2. Assessing assay variability in soil and waste samples

A one-way ANOVA followed by the Dunnett's test was used to determine the LID (lowest ineffective dilution) and LOEC (lowest observed effect concentration) of *A. globiformis* DHA for each test substrate ($p < 0.05$). Whenever the one-way ANOVA assumptions were not fulfilled after data transformation, it was applied the non-parametric Dunn's test to detect significantly different substrates from the control ($p < 0.05$). The minimum detected difference (MDD) was determined for the standard deviation of the average response corresponding to the LOEC value, and expressed as a percentage of the respective negative control. The EC_{20} and EC_{50} values and respective 95 % confidence limits were estimated by fitting the least-squares regression model to the data. The tests in which the EC_x or LID/LOEC could not be calculated were disregarded from the downstream statistical analyses.

The presence of outliers, as well as the repeatability and reproducibility of the *A. globiformis* contact test were evaluated according to ISO 5725-2 [38] procedures, though some amendments were done. The assay repeatability was evaluated by the logarithmic within-laboratory standard deviation (WLSD), using the log-transformed 95%-confidence limits of the EC*x*. Reproducibility was assessed by different statistical approaches based on the EC*x* values calculated: (i) standard deviation (SD_R) of the average EC*x* values computed for all the laboratories *per* test substrate, (ii) CV_R, (iii) warning limits approach, and (iv) min-max factor.

3. Results & Discussion

The sensitivity of *A. globiformis* to discriminate contaminated soils [*e.g.*, 24] and waste [*e.g.*,19] samples, together with the test system simplicity, cost-effectiveness and short time required for attaining reliable results, were major driving forces for its approval for standardization. Within this process, an IRT was organized and carried out for completing the validation of the assay, being estimated its repeatability and reproducibility for routine use towards quality assessment of solid samples with distinct properties and/or contamination patterns.

3.1 Validity criteria and IRT data consistency

More than 78% of the participant laboratories were able to fulfill at least two of the validity criteria set in the draft ISO 18187 [27] (Table 2), being each criterion attained in more than 80% of the tests performed. Such outcome highlights the practicability and understandability of the standard procedures to the participants, who did not mention particular difficulties or constraints to perform the test.

The scrutiny for data consistency (*H* and *K* statistics) and presence of outliers (Cochran's and Grubb's tests) [38] showed that some test values were out of the range reached by most laboratories, being thereby discarded from the statistical analysis (Table 2). L6 presented out-of-range data variability between replicates in negative (CV_r's of 92-98%) and positive (CV_r's of ~99%) controls, besides obtaining discrepant responses for *A. globiformis* DHA relatively to the other laboratories. For this reason, it was dismissed from some analyses as well. No plausible explanation was found for this discrepancy, since no inappropriate procedure or bacterial damage was noticed by the participant. Notwithstanding, at least 8 (over a total of 9) valid datasets (tests) for soils or wastes were selected to evaluate the variability of IRT results concerning the testing of solid samples (Table 2). Overall, only 7% outliers were identified (*i.e.*, 5 in 72 tests), suggesting a generally high data consistency.

Table 2

3.2 Variability of the test method - controls

3.2.1 Negative controls

Negative controls are required to estimate the percentage of inhibition of *A. globiformis* DHA under contaminated matrices [39]. In compliance with the draft ISO 18187 [27], two standard soils (Lufa 2.2 and OECD soil) and one substrate (quartz sand) were used as negative controls to analyze the inherent variability of the test system. Such analysis will also serve to confirm the CV established in the third validity criterion (*cf.* section 2.5,), and to define an internal quality control for each laboratory.

The average CV_r values obtained by participant laboratories were below 15%, irrespectively of the negative control (Figure 1). These results showed that for all laboratories the within-variability in *A. globiformis* DHA was low for the three substrates. They further confirmed the suitability of the selected substrates as negative controls and the good repeatability of the contact-test procedure.

Nevertheless, it is noticeable a higher variability on *A. globiformis* response under the OECD soil (Figure 1C), being recorded for 4 laboratories (L2, L4, L5 and L8) a CV_r higher than 15%. Such within-lab variability can be explained by the heterogeneous OECD soil matrix due to the high peat content (10%) and presence of peat particles with varied size (up to 2 mm).

Figure 1

Consequently, small discrepancies during soil weighing and distribution into the microplate wells might induce an observable variance in the results. This is further constrained by the fact that certain types of organic compounds in peat may promote fluorescent light quenching effects, therefore affecting the measured enzymatic activities [22]. As such, it is recommended to use a lower amount of peat (*e.g.*, 5%) as to reduce variability in the test outcome.

3.2.2 Positive control – suitability of C16-BAC

Regarding the test performance under the reference substance C16-BAC at 600 mg Kg^{-1}, it was clearly demonstrated that all the laboratories achieved DHA inhibitions between 30 and 80% (Figure 2), which is within the range required to fulfill the second validity criterion of the contact assay [27].

Figure 2

Hence, the confirmed sensitivity of *A. globiformis*, either within or between laboratories, reinforces the adequacy of C16-BAC to be used as a reference substance for evaluating the responsiveness of the test method and of the bacterium, being in agreement with previous works [*e.g.*,22]. Moreover, a good repeatability and reproducibility of the contact test was proved, given the CV_r values broadly below 30% (except for L3) and CV_R = 26%, respectively (Figure 2).

3.3 Sensitivity and variability of the contact assay for solid samples

Two major approaches were followed for evaluating the effects of soils (S1/S2 to S4) and wastes (W1 to W4) on *A. globiformis* response: (i) the use of a threshold of 30% for DHA inhibition, above which the microbial metabolic function might be compromised

[27], and (ii) the application of statistical analysis (ANOVA) for determining the LID and LOEC values. The toxicity outcomes provided by both approaches were quite coherent for most laboratories, reinforcing the adequacy of the threshold for DHA inhibition defined in the standard (Figures 3 and 4, Table 3).

The ranking of the different soils according to the average DHA inhibition of *A. globiformis* was generally consistent between laboratories: S1 (only the threshold-based approach was used since no concentration range was tested, hence preventing the calculation of LID/LOEC values) and S2 induced lower inhibitions comparatively to S3 and, especially, to S4 (Figure 3). Similarly, the EC_{20} and EC_{50} mean values calculated for all laboratories were also ranking soils S4>S3>S2=S1 into a decreasing order of DHA inhibition. In fact, S3 and S4 are originally from two areas of mine tailings deposition, being characterized by an acidic pH and a high metal content (Table 1). The considerably high organic matter content in soil S4 might have also affected the fluorescence readings due to quenching effects (see above), therefore overestimating the real level of fluorescent light emitted by the reduced resorufin produced by DHA [24].

The response of *A. globiformis* to waste samples was more variable between laboratories, though based on the 30% threshold all laboratories identified W1 and W3 as being the most inhibitory of DHA. However, based on statistical methods (*i.e.*, on LOEC and EC_x values), W1, W2 or W3 (Figure 4, Table 3) induced higher depletions of DHA. W4, in turn, was generally the least inhibitory waste for *A. globiformis* DHA.

Figure 3

The metal levels in W1 (Cu, Cr) and W3 (Cu, Pb, As, Zn) were above the soil benchmark values defined for soil organisms or microbial processes in USA [40], which together with the respective acidic and alkaline pH of W1 and W3 might have significantly depleted DHA. In W2, however, the Cr and PHA contents could have been responsible for the small number of recorded inhibitions [40]. Moser et al. [34] also reported the extremely high toxicity of W1 (WOO) for this bacterium, which proved to be more sensitive than other terrestrial organisms and test methods/endpoints, such as earthworm avoidance and reproduction, and enchytraeidae and collembolans reproduction. Similarly, Römbke et al. [25] observed that *A. globiformis* DHA (LID<16) was more sensitive than the avoidance behavior of *Eisenia fetida* under W3

(LID = 16) and equally sensitive as *Brassica napus* growth (LID<16). Notwithstanding, the authors confirmed that the three endpoints/test species were not that responsive to W2 or W4 (LID = 4), what is in agreement with the lower inhibitions herein observed.

Although a wide range of LOEC/LID values could be determined among laboratories either for soils or wastes, it was noticeable that ≥50% of laboratories were achieving the same or only two different LOEC values. The greater variability in S4, W3 and W4 could be attributed to the contamination pattern or the type of solid sample (*i.e.*, metal contamination, ash waste and crushed glass, respectively), which may induce some variation in *A. globiformis* and/or fluorescence readings. Even though, for hazard assessment purposes regarding the HP14 property of wastes, a LID of 8 was suggested as the threshold above which wastes can be considered toxic [cf. 19]. Hence, the range of LOEC values herein obtained for wastes would not affect the decision about the ecotoxicological *status* of the waste samples.

EC_{20} and EC_{50} values, in turn, resulted from the fitting of the least-squares regression model to the experimental data. Hence, a more accurate toxicity value can be estimated, which is not constrained by the concentration range under testing, as for LOEC/LID values. Therefore, and particularly the average EC_{50} values obtained for all laboratories were seemingly less variable (*cf.* Figures 3 and 4, Table 3) than the LOECs/LIDs. Notwithstanding, both point estimates should still be applied, for example for HP14-waste-property classification purposes, as to get more robust and feasible decisions and /or quality evaluations.

Table 3

The variability of the contact assay for testing the quality of solid samples was evaluated through the determination of repeatability (MDD, WLSD) and reproducibility (*i.e.*, upper and lower warning limits, CV_R and SD_R). MDD provides an estimation of the power of the hypothesis test to discriminate a variation in the endpoint provoked by a sample relatively to the control [41,42].

For the four evaluated soils, the average MDD values varied between 1.36 and 9.98% (Table 2), whilst for wastes it was between 0.12 and 24.41%. The maximum MDD

value was observed by one laboratory (L6) for W4, although >90% of the waste tests resulted in MDD < 7%.

Figure 4

Notwithstanding, the wider range of MDD values for wastes might also be associated with the heterogeneity of some sample matrices (*e.g.*, W1, W2 and W4) (*cf.* Table 1), either in terms of density (W1 - wood material), contamination with oils (W2 - dredged material from a harbor) or particle size (W4 - glass-based waste), therefore constraining the consistency on the fluorescence readings. In spite of this, most MDD values were in conformity with the ones determined by Höss et al. [43] (7.9%) and Marques et al. [24] (2.5-4.7%) for *A. globiformis* exposed to different sediment and soil samples, respectively. The consistency in MDD indirectly reinforced the sensitivity of the *A. globiformis* contact assay to indicate changes in the bacterium response upon exposure to the test samples.

The logarithmic within-lab standard deviations (WLSD) were below 0.50 for soils and 0.25 for wastes, being most of them (76% of the WLSD's *per* laboratory) below 0.07, thereby showing a generally high accuracy and a good repeatability of the IRT results for solid samples.

Concerning the between-lab variability of the assay, all EC_{20} and EC_{50} values computed for soils and wastes were within the respective lower and upper warning limits, as well as the min-max factors were generally <4.0, which is the maximum threshold acceptable for the variation of the EC_x values between laboratories (Table 3) [44]. In a general view, the EC_{50}-based min-max factors were lower for soils, whilst for wastes they were lower for EC_{20}-based parameters. From the average EC_x's for soils and wastes, it was verified that five tests presented a $CV_R \leq 30\%$ (Table 3), three were between 30 and 50%, and two were >50%. The CV_R-based variability trend was confirmed by the reproducibility standard deviation. Once CV and SD values should be up to 30% and 0.132 (EC [45]), the outcome achieved for soil and wastes indicated a satisfactory reproducibility of the test.

4. Conclusion

The analysis of the IRT comprising at least 8 datasets generated by 9 laboratories led to the conclusion that the *A. globiformis* solid contact assay is valid to assess the quality of soils and wastes with good repeatability and reproducibility. Overall, no constraints or difficulties in assay practicability were referred by the participants, which generally attained the assay validity criteria, as well as reproducible and comparable results, either for the reference substance, control substrates and solid samples tested. Therefore, the *A. globiformis* contact test proved to be sensitive, robust and feasible for routine use towards the quality assessment of solid samples. Once the IRT was completed and succeeded to demonstrate the valuable applicability of the assay, the standard was hence published into a final document (ISO 18187:2016 [46]). At the light of the above, we recommend its inclusion in test batteries for the ecotoxicological characterization of soils and wastes.

5. Acknowledgments

This study was funded by ECT Oekotoxikologie, by the Portuguese Foundation for Science and Technology (FCT) and by the European Regional Development Fund (ERDF) through the Portugal 2020 Partnership Agreement between Portugal and the European Union, the Competitiveness and Internationalization Operational Programme (COMPETE 2020) and the Regional Operational Programme Lisboa (POR-Lisboa), under the project NANOBINDERS (PTDC/AAG-REC/3004/2014). Catarina R. Marques is granted as an Invited Scientist within the project NANOBINDERS. One soil sample was obtained within the scope of a project funded by NATO SfPS (ref. SFP 983311). Thanks are further due for the financial support to CESAM (UID/AMB/50017), FCT/MEC through national funds, and the co-funding by the FEDER, within the PT2020 Partnership Agreement and COMPETE 2020. The funding institutions had no direct involvement in the development of the study. The authors declare that there are no conflicts of interest.

6. References

[1] P. Pandard, J. Devillers, A.-M. Charissou, V. Poulsen, M.-J. Jourdain, J.-F. Férard, C. Grand, A. Bispo A., Selecting a battery of bioassays for ecotoxicological characterization of wastes, Sci. Total Environ. 363 (2006) 114-125.

[2] L. Torstensson, M. Pell, B. Stenberg, Need of a strategy for evaluation of arable soil quality, Ambio 27 (1998) 4-8.

[3] P.G. Falkowski, T. Fenchel, E.F. Delong, The microbial engines that drive Earth's biogeochemical cycles, Science 320 (2008) 1034–1039.

[4] M.G.A. van der Heijden, R.D. Bardgett, N.M. van Straalen, The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems, Ecol. Lett. 11 (2008) 296–310.

[5] G.P. Sparling, Soil microbial biomass activity and nutrient cycling as indicators of soil health, in: C.E. Pankhurst, B.M. Doube, V.V.S.R. Gupta (Eds.), Biological Indicators of Soil Health, CAB International, Wallingford, 1997, pp. 97-119.

[6] P. Pereira, J.P. Sousa, R. Ribeiro, F. Gonçalves, Microbial indicators in mine soils (S. Domingos mine, Portugal), Soil Sediment. Contam. 15 (2006) 147–67.

[7] L. Philippot, K. Ritz, P. Pandard, S. Hallin, M. Martin-Laurent, Standardisation of methods in soil microbiology: progress and challenges, FEMS Microbiol. Ecol. 82 (2012) 1-10.

[8] ISO/TS 29843-1, Soil quality - Determination of soil microbial diversity - Part 1: Method by phospholipid fatty acid analysis (PLFA) and phospholipid ether lipids (PLEL) analysis, International Standardization Organization (ISO), Geneve, Switzerland, 2010.

[9] ISO/TS 29843-2, Soil quality - Determination of soil microbial diversity - Part 2: Method by phospholipid fatty acid analysis (PLFA) using the simple PLFA extraction method, International Standardization Organization (ISO), Geneve, Switzerland, 2011.

[10] ISO 14240-1, Soil quality - Determination of soil microbial biomass - Part 1: Substrate-induced respiration method, International Standardization Organization (ISO), Geneve, Switzerland, 1997.

[11] ISO 14240-2, Soil quality - Determination of soil microbial biomass - Part 2: Fumigation-extraction method, International Standardization Organization (ISO), Geneve, Switzerland, 1997.

[12] ISO 17155, Soil quality - Determination of abundance and activity of soil microflora using respiration curves, International Standardization Organization (ISO), Geneve, Switzerland, 2012.

[13] ISO 15685, Soil quality - Determination of potential nitrification and inhibition of nitrification - Rapid test by ammonium oxidation, International Standardization Organization (ISO), Geneve, Switzerland, 2004.

[14] ISO 14238, Soil quality - Biological methods - Determination of nitrogen mineralization and nitrification in soils and the influence of chemicals on these processes, International Standardization Organization (ISO), Geneve, Switzerland, 2012.

[15] ISO 17601, Soil quality - Estimation of abundance of selected microbial gene sequences by quantitative PCR from DNA directly extracted from soil, International Standardization Organization (ISO), Geneve, Switzerland, 2016.

[16] J. Ahtiainen, Microbiological tests and measurements in the assessment of harmful substances and pollution, in: Monographs of the Boreal Environment Research 22, 2002, pp. 7-51.

[17] J. Jensen, M. Mesman, Ecological Risk Assessment of Contaminated Land. Decision Support for site specific investigations, RIVM report number 711701047, National Institute for Public Health and the Environment, The Netherlands, 2006, pp. 136.

[18] M. Wagelmans, J.G.M. Derksen, S.A.E. Kools, J. Faber, J. van der Pol, M. Mesman, D. Lud, Evaluation of the implementation of Triad in assessing ecological risks, SKB (The Centre for Soil Quality Management and Knowledge Transfer) project report, The Netherlands, 2009, p. 46.

[19] P. Pandard, J. Römbke, Proposal for a "Harmonized" strategy for the assessment of the HP 14 property, Integr. Environ. Assess. Manag. 9 (2013) 665-72

[20] K. Thomson, D. Liu, K.L.E. Kaiser, A direct resazurin test for measuring chemical toxicity, Toxic. Assess. 1 (1986) 407-418.

[21] K. Rönnpagel, E. Janßen, W. Ahlf, Asking for the indicator function of bioassays evaluating soil contamination: are bioassay results reasonable surrogates of effects on soil microflora?, Chemosphere 36 (1998) 1291-1304

[22] H. Neumann-Hensel, K. Melbye, Optimization of the solid-contact test with *Arthrobacter globiformis*, J. Soils Sediment. 6 (2006) 207.

[23] H. Moser, J. Römbke, Ecotoxicological Characterization of Waste - Results and Experiences of an International Ring Test, Springer, NY, USA, 2009, p. 308.

[24] C.R. Marques, A. Caetano, A. Haller, R. Pereira, F. Gonçalves, J. Römbke, Toxicity screening of soils from different mine areas - A contribution to track the

sensitivity and variability of *Arthrobacter globiformis* assay, J. Hazard. Mat. 274 (2014) 331-341.

[25] J. Römbke, R. Ketelhut, B. Förster, E. Richter, A. Scheffczyk, Hazard classification of 24 waste materials using three terrestrial standard methods: the bacterial contact test, the plant growth test and the earthworm avoidance test, (unpublished data).

[26] J. Römbke, Th. Moser, H. Moser, Ecotoxicological characterization of 12 incineration ashes using 6 laboratory tests, Waste Manage. 29 (2009) 2475-2482.

[27] ISO 18187, Soil Quality – Quality of Solid Samples - Solid contact test using the dehydrogenase activity of *Arthrobacter globiformis*, ISO/CD 18187, International Organization for Standardization, 2012.

[28] ISO 10381-6, Soil quality - Sampling - Part 6: Guidance on the collection, handling and storage of soil under aerobic conditions for the assessment of microbiological processes, biomass and diversity in the laboratory, International Standardization Organization (ISO), Geneve, Switzerland, 2009.

[29] ISO 10390, Soil quality - Determination of pH, International Standardization Organization (ISO), Geneve, Switzerland, 2005.

[30] ISO 11268-2, Soil quality - Effects of pollutants on earthworms - Part 2: Determination of effects on reproduction of *Eisenia fetida/Eisenia Andrei*, International Standardization Organization (ISO), Geneve, Switzerland, 2012.

[31] ISO 10694, Soil quality - Determination of organic and total carbon after dry combustion (elementary analysis), International Standardization Organization (ISO), Geneve, Switzerland, 2005.

[32] ISO 11277, Soil quality - Determination of particle size distribution in mineral soil material - Method by sieving and sedimentation, International Standardization Organization (ISO), Geneve, Switzerland, 2009.

[33] R. Becker, U. Kalbe, A. Buchholz, Selection and Characterization of Test Samples and Eluates, in: H. Moser, J. Römbke (Eds), Ecotoxicological Characterization of Waste. Springer, New York, NY, 2008.

[34] H. Moser, J. Römbke, G. Donnevert, R. Becker, Evaluation of biological methods for a future methodological implementation of the Hazard criterion H14 'ecotoxic' in the European waste list (2000/532/EC), Waste Manage. Res. 29 (2011) 180-187.

[35] EN 14735, Characterization of waste - preparation of waste samples for ecotoxicity tests. CEN – EN (European Committee for Standardization), 2005.

[36] ABANDA database, [http://www.abfallbewertung.org/; accessed on 15.07.2017]

[37] OECD 207, Earthworm, Acute Toxicity Test, Organisation for Economic Co-operation and Development (OECD), 1984.

[38] ISO 5725-2, Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method, 1994.

[39] H. Løkke, C.A.M. van Gestel, Handbook of Soil Invertebrate Toxicity Tests. Wiley & Sons, Chichester, 1998.

[40] R.A. Efroymsom, M.E. Will, G.W. Suter, Toxicological benchmarks for contaminants of potential concern for effects on soil and litter invertebrates and heterotrophic process: 1997 Revision, ES/ER/TM-126/R2, U.S. Department of Energy, Office of Environmental Management, Oak Ridge, U.A.A., 1997.

[41] J. Römbke, S. Jänsch, T. Junker, B. Pohl, A. Scheffczyk, H.J. Schallnass, An improvement of the applicability of ecotoxicological tests with earthworms, springtails, and plants for the risk assessment of metals in natural soils, Environ. Toxicol. Chem. 25 (2006) 776-787.

[42] T.C.M. Brock, M. Hammers-Wirtz, U. Hommen, T.G. Preuss, H.T. Ratte, I. Roessink, T. Strauss, P.J. Van den Brink, The minimum detectable difference (MDD) and the interpretation of treatment-related effects of pesticides in experimental ecosystems, Environ. Sci. Pollut. Res. 22 (2015) 1160-1174.

[43] S. Höss, W. Ahlf, C. Fahnenstich, D. Gilberg, H. Hollert, K. Melbye, M. Meller, M. Hammers-Wirtz, P. Heininger, H. Neumann-Hensel, R. Ottermanns, H.T. Ratte , T.B. Seiler, D. Spira, J. Weber, U. Feiler, Variability of sediment-contact tests in freshwater sediments with low-level anthropogenic contamination - determination of toxicity thresholds, Environ. Pollut. 158 (2010) 2999-3010.

[44] P.M. Chapman, Ecotoxicology and pollution – key issues, Marine Poll. Bull. 31 (1995) 167-177.

[45] EC (Environment Canada), Guidance document on control of toxicity test precision using reference toxicants, Environmental Protection Series, Ottawa, Ont., Rept. EPS 1/RM/12, 1990.

[46] ISO 18187:2016, Soil quality - Contact test for solid samples using the dehydrogenase activity of *Arthrobacter globiformis*. International Organization for Standardization, (2016).

FIGURE CAPTIONS

Figure 1. Coefficient of variation within (CV_r) laboratories (L) for the slope values obtained for the dehydrogenase activity (DHA) of *A. globiformis* exposed to quartz sand (A; dark blue circles), Lufa 2.2 (B; green circles) and OECD soil (C; purple circles). The average of CV_r values for 8 laboratories (L) is shown by the continuous lines, whilst the dashed lines highlight the upper and lower limits of the respective standard deviation. The number of different tests performed by each laboratory in which the calculated CV_r values were based on is shown on the top of the respective symbols.

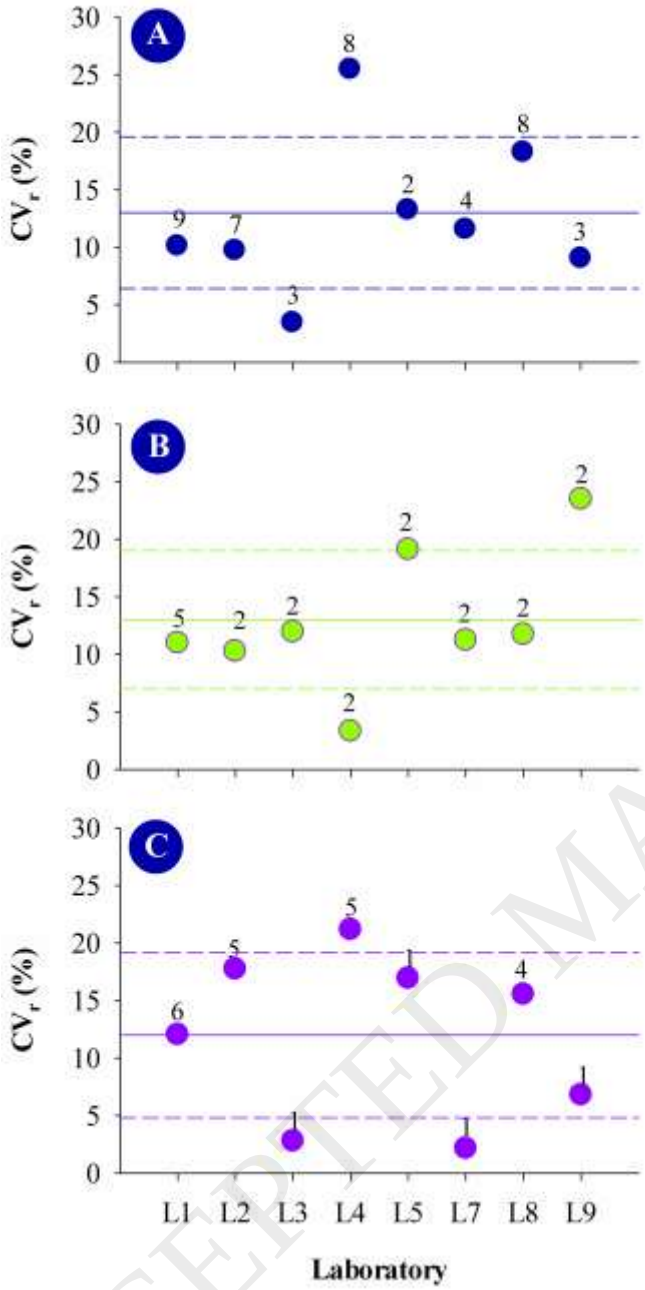


Figure 2 – Average inhibition of the dehydrogenase activity (DHA) of *A. globiformis* attained by 8 participant laboratories (L) for the reference substance at 600 mg C16-BAC Kg^{-1}. The error bars indicate standard deviation (n=4) and the numbers on the top represent the coefficient of variation within (CV_r; %) laboratories based on the following number of tests: 5 for L1, 1 for L2, 3 for L3, and 2 for L4 to L9. CV_R is the coefficient of variation between laboratories.

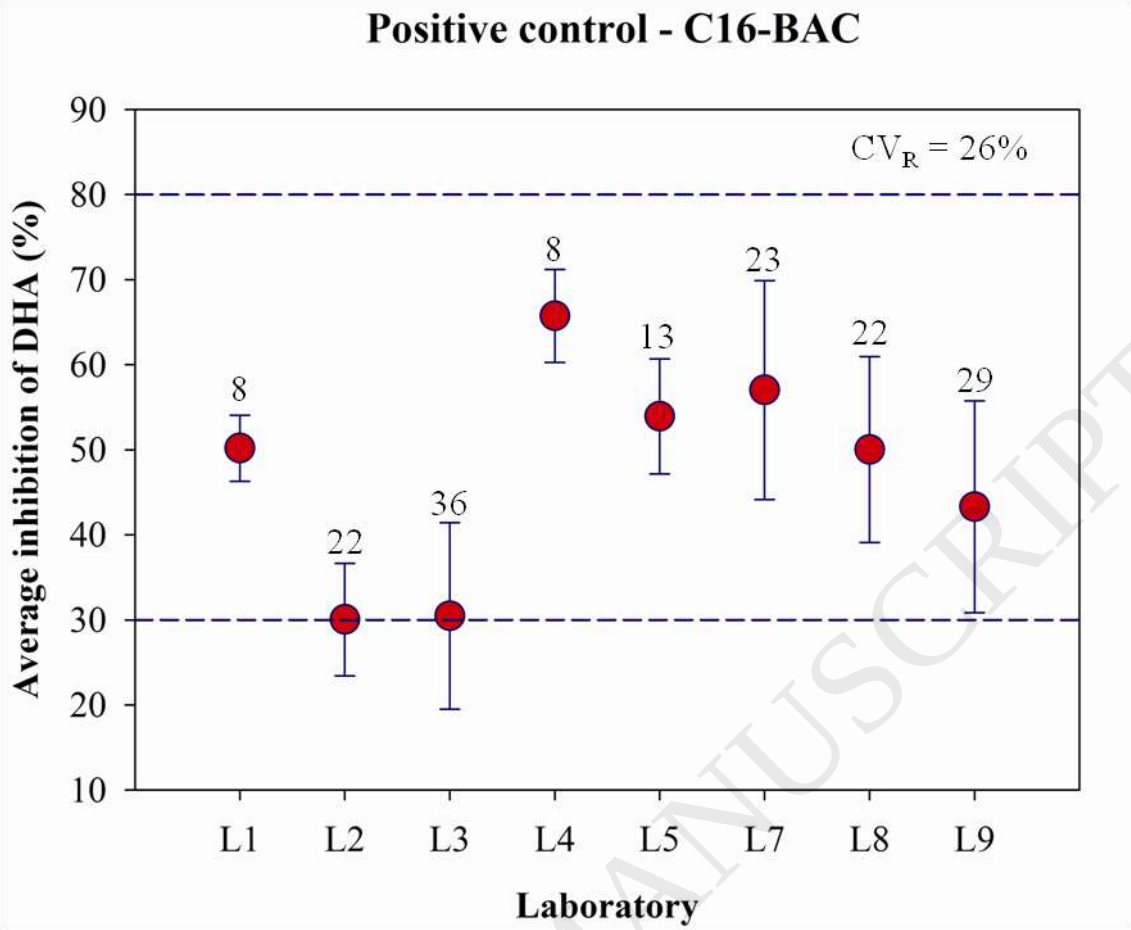


Figure 3 – Average inhibition of dehydrogenase activity (DHA) of *A. globiformis* recorded by the participant laboratories (L) in response to different dilutions of the soil samples S1 to S4. The dashed line indicates the 30% inhibition threshold (ISO 18187). The dilutions on the top of the bars highlight the LOEC value ($p < 0.05$). The error bars represent standard deviation values. The orange circles represent the EC_{50} values and upper confidence limit calculated for DHA for each laboratory [EC_{50} <6.3% (S4: L3) or >100% (S2: all labs; S3: L3 and L4) are not represented].

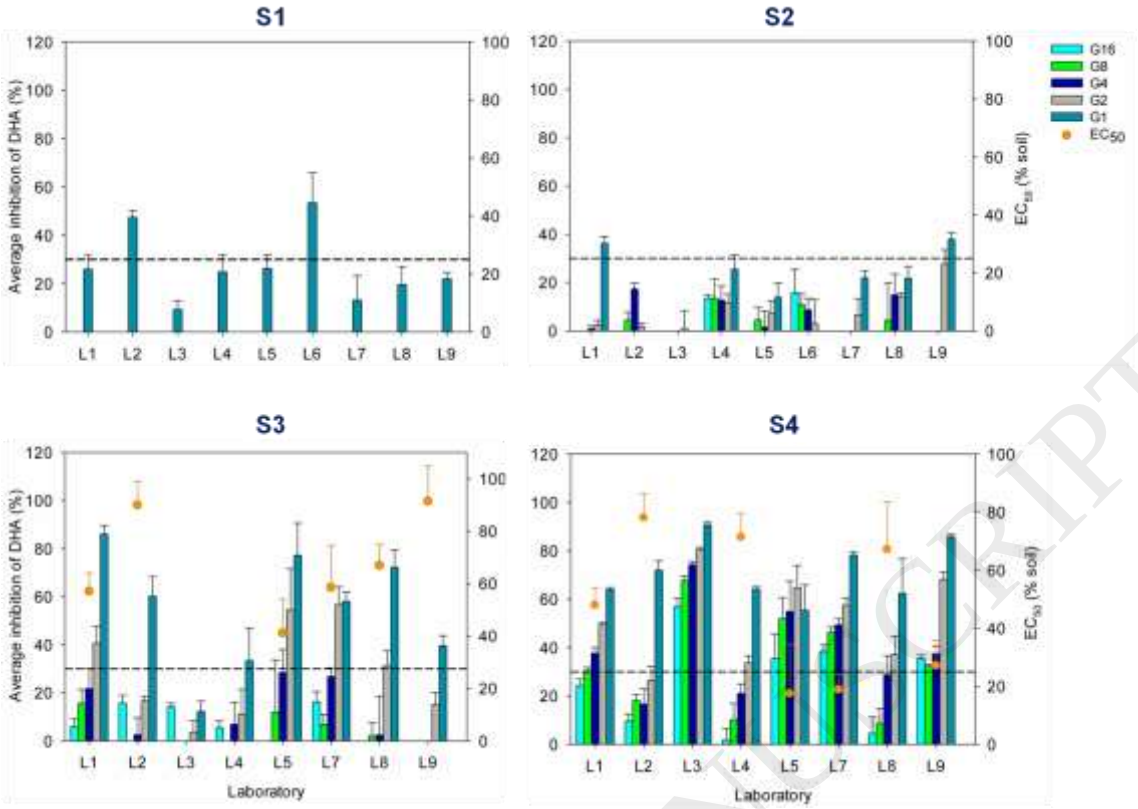
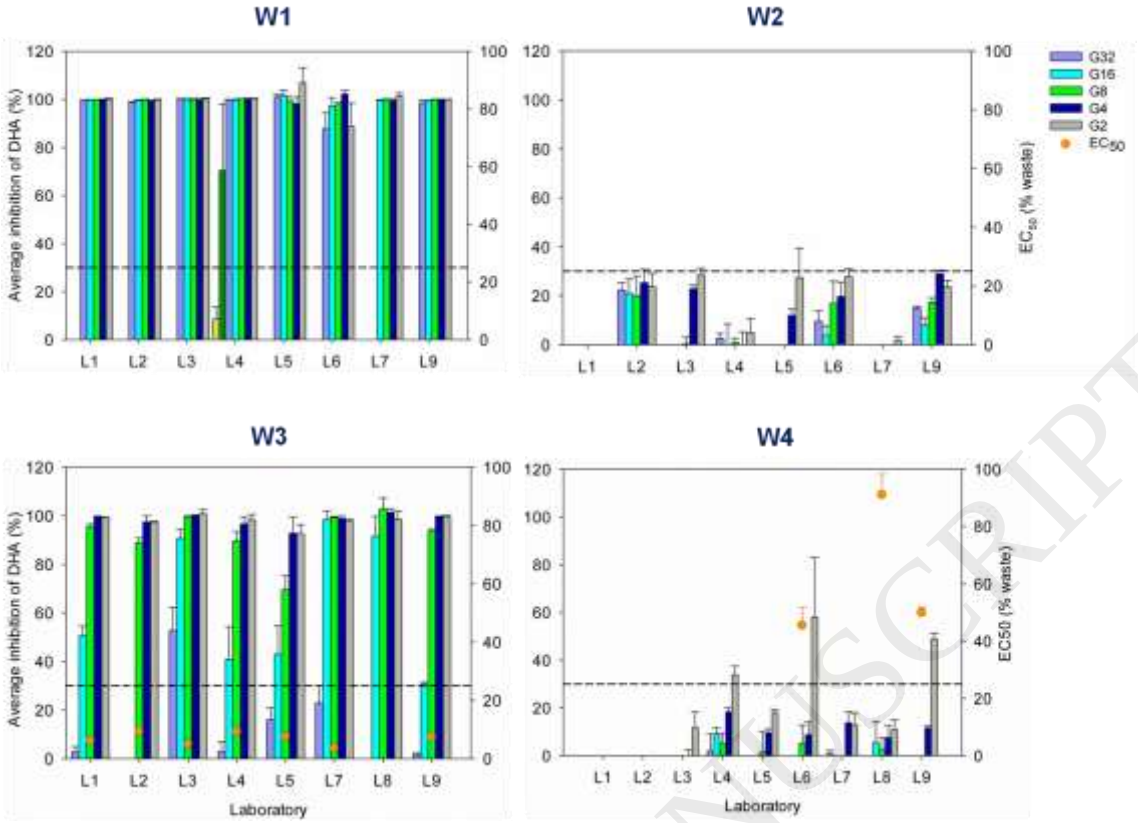


Figure 4 – Average inhibition of dehydrogenase activity (DHA) of *A. globiformis* recorded by the participant laboratories (L) in response to different dilutions of the waste samples W1 to W4. The dashed line indicates the 30% inhibition threshold. The dilutions on the top of the bars highlight the LOEC value ($p < 0.05$). The error bars represent standard deviation. The orange circles represent the EC_{50} values and upper confidence limit calculated for DHA for each laboratory [EC_{50} <3.1/6.3% (W1: all labs; W3: L8) or >50.0% (W2: all labs; W4: L1-L5 and L7) are not represented].



TABLES

Table 1 - Physical and chemical characterization and metal content of the soil and waste samples. *Data retrieved from ABANDA database [37].

	Soils				Wastes			
Test Items	**S1**	**S2**	**S3**	**S4**	**W1**	**W2**	**W3**	**W4**
Waste code	-	-	-	-	WOO	17 05 06	10 01 17	19 12 05
Description	Control soil	Construction site	Uranium mine	Phosphogypsum deploying site	Cu-treated Wood	Dredged harbor material	Fluidized bed ash	Crushed glass material
pH-value	6.23	6.93	4.25	4.14	5.18	6.89 *	9.96 *	9.06 *
Conductivity (µS cm^{-1})	271	159	293	2723	n.d.	n.d.	n.d.	n.d.
OM content (%)	2.93	4.20	2.36	5.39	46.4	3.91	6.54	1.04
Texture	Silty clay	Sandy loam	Loamy sand	Sandy loam				
Sand (%)	9.9	55.2	85.9	69.1	n.d.	n.d.	n.d.	n.d.
Silt/clay (%)	99.1	44.8	14.1	24.9				
Metals (mg Kg^{-1})								
Cu	3.06	24.32	22.83	11.46	1057.31	92.7	254.1*	5.69*
Pb	1.36	35.47	32.01	20.88	14.30	162.2	2187*	148.2*
Cd	0.38	0.34	21.81	0.11	0.27	2.1	8*	0.76*
Cr	8.67	14.71	33.36	9.86	298.16	171.8	n.d.	3.37*
U	<dl	<dl	169.2	1.63	n.d.	n.d.	n.d.	n.d.
Zn	50.72	155.15	118.83	89.51	43.50	582.2	4049*	1222.09*
Fe	742.37	1631.71	2413.54	10870.4	n.d.	n.d.	n.d.	n.d.
Ni	1.88	6.36	32.96	17	n.d.	n.d.	n.d.	n.d.
As	n.d.	n.d.	n.d.	n.d.	n.d.	52.5	99.8*	n.d.

PAH (µg Kg^{-1})	n.d.	8.5	n.d.	n.d.	n.d.	6.1	n.d.	n.d.

dl – detection limit; n.d. – not determined.

Table 2 - Summary of the ring-test results and interval of minimum detected difference (MDD) values determined for *A. globiformis* response under soil and waste samples (data also presented in [45]).

Solid sample	Control substrate	≥ 2 validity criteria met (%)	Total # tests	# tests with outliers	# accepted tests	# tests for MDD calculation	Mean MDD (%)	SD
S1	Lufa	89	9	0	9	0	nd	nd
S2	Lufa	89	9	0	9	5	3.8	1.78
S3	Lufa	78	9	1	8	6	5.5	2.86
S4	Lufa	89	9	1	8	4	7.1	1.23
W1	QS	89	9	1	8	8	1.2	2.18
W2	QS	89	9	1	8	6	6.0	3.81
W3	QS	78	9	1	8	5	2.9	2.08
W4	QS	78	9	0	9	6	6.8	8.88

S – soil; W – waste; # – number; Lufa – Lufa 2.2 natural soil; QS – quartz sand; n.d. – not determined.

Table 3 - Outcome of the ECx-based repeatability (given by WLSD) and reproducibility (given by LWL/UWL, CV_R, SD_R, Min-Max factor) analysis of *A. globiformis* contact assay for soil and waste samples (data also presented in [45]).

		Waste number	No. of tests	Mean ECx (%)	WLSD (min-max)	LWL	UWL	CV_R (%)	S_R	Min-Max factor
Soils	EC20	S1	0	nd	nd	nd	nd	nd	nd	nd
		S2	5	69.5	0.03-0.27	38.08	119.4	26	0.12	2.1
		S3	6	40.5	0.04-0.14	10.98	114.16	52	0.25	4.6
		S4	3	32.5	0.04-0.11	12.37	74.16	47	0.19	2.4
	EC50	S1	0	nd	nd	nd	nd	nd	nd	nd
		S2	0	>100.0	nd	nd	nd	nd	nd	nd
		S3	6	67.4	0.02-0.07	35.41	119.17	29	0.13	2.2
		S4	7	46.9	0.02-0.50	11.01	144.59	55	0.28	4.5
Wastes	EC20	W1	0	<1.6/3.1/6.3[1]	nd	nd	nd	nd	nd	nd
		W2	4	24.8	0.06-0.25	15.92	37.32	19	0.09	1.6
		W3	6	5.6	0.002-0.05	2.68	10.15	33	0.14	2.2
		W4	5	37.9	0.02-0.06	20.85	64.54	30	0.12	1.9
	EC50	W1	0	< 1.6/3.1/6.3[1]	nd	nd	nd	nd	nd	nd
		W2	0	>50.0	nd	nd	nd	nd	nd	nd
		W3	7	6.9	0.002-0.03	3.37	12.85	30	0.14	2.5
		W4	3	62.1	0.008-0.03	27.80	125.69	40	0.16	2.0

WLSD – logarithmic within-lab standard deviation; LWL/UWL – lower and upper warning limits; CV_R – reproducibility coefficient of variation; SD_R, – reproducibility standard deviation determined from the average of the logarithm of EC_x values; Min-Max factor – factor determined from the minimum and maximum values of EC_x's; n.d. – not determined; [1] the minimum dilution tested by some laboratories was G64 (= 1.6% waste), G32 (= 3.1% waste) or G16 (= 6.3% waste).