



Stone, D., Ritz, K., Griffiths, B.S., Orgiazzi, A. and Creamer, R.E. (2016) Selection of biological indicators appropriate for European soil monitoring. *Applied Soil Ecology*, 97, pp.12-22. ISSN 0929-1393.

Copyright © 2015 Elsevier B.V. All rights reserved.

This manuscript version is made available after the end of the 12 month embargo period under the CC-BY-NC-ND 4.0 license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

<http://hdl.handle.net/11262/10854>

<http://dx.doi.org/10.1016/j.apsoil.2015.08.005>

1 **Selection of biological indicators appropriate for European soil monitoring**

2

3 D. Stone<sup>1,2</sup>, K.Ritz<sup>3</sup>, B.G. Griffiths<sup>4</sup>, A. Orgiazzi<sup>5</sup>, R.E. Creamer<sup>2\*</sup>

4

5 <sup>1</sup>*Faculty of Biological Sciences, Leeds University, Leeds, UK,* <sup>2</sup>*Teagasc, Johnstown*  
6 *Castle Research Centre, Ireland,* <sup>3</sup>*Division of Agricultural & Environmental Sciences,*  
7 *University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, UK*  
8 <sup>4</sup>*Crop & Soils Systems Research Group, SRUC, Edinburgh, EH9 3JG, UK,* <sup>5</sup>  
9 *European Commission, Joint Research Centre (JRC), Institute for Environment and*  
10 *Sustainability (IES), Land Resource Management Unit, Via E. Fermi, 2749 – 21027*  
11 *Ispira (VA), ITALY.*

12

13

14 \*Corresponding author ([rachel.creamer@teagasc.ie](mailto:rachel.creamer@teagasc.ie))

15 **Abstract**

16 The selection of biological indicators for monitoring progress towards policy goals for  
17 soil quality should be without bias and in line with individual scenarios of need. Here  
18 we describe the prescription of a suite of appropriate indicators for potential  
19 application in such monitoring schemes across Europe. We applied a structured  
20 framework of assessment and ranking (*viz.* a ‘logical sieve’), building upon published  
21 data and a new survey taken from a wide section of the global soil biodiversity  
22 research and policy community.

23 The top ten indicators included four indicators of biodiversity (three microbial and  
24 one meso-faunal) by various methods of measurement, and three indicators of  
25 ecological function (Multiple enzyme assay, Multiple substrate-induced respiration  
26 profiling, and ‘Functional genes by molecular biological means’). Within the  
27 techniques assessed, seven out of the top ten indicators made use of molecular  
28 methods.

29

30 **Keywords (5)**

31 Logical sieve, indicators, monitoring, soil biodiversity, soil ecosystem function

32

33 **1. Introduction**

34 National and international policy development regarding soil quality requires data on  
35 the current situation to create a baseline, or normal operating range, to which new  
36 policies can be applied. Enforcement of policy in turn requires monitoring, using pre-  
37 described indicators, to identify improvement or degradation (Turbe *et al.*, 2010). In  
38 2001 the OECD laid down a set of criteria for agri-biodiversity indicators. These  
39 need to be measurable, based in sound science, able to be interpreted by policy  
40 makers, and policy relevant. They also need to allow monitoring at multiple spatial  
41 and temporal scales (OECD, 2001).

42

43 The majority of soil processes are intrinsically linked to soil biota, although specific  
44 relationships are complex and as yet not fully described in all cases (Ritz *et al.* 2009).  
45 However it is under these conditions of uncertainty that biological indicators are key in  
46 monitoring soil quality (here defined as the ability to deliver key soil processes) in ways  
47 that other indicators are not. By measuring organisms - which by being inherently  
48 adaptive - integrate multi-dimensional phenomena including time, it is possible at least in  
49 principle to ascertain the full potential of a soil to deliver key soil processes (Ritz *et al.*  
50 2009).

51

52 However large-scale soil monitoring campaigns are expensive both in terms of labour  
53 and capital needed to acquire enough data for accurate conclusions. Selecting the best  
54 group of indicators is therefore important to reduce costs and provide data which is fit

55 for purpose, and can enable policy making and policy implementation in relation to  
56 soil quality into the future. This is however not as easy as it sounds. Previous  
57 attempts, for example the ENVASSO project working towards an operational set of  
58 EU-wide criteria and indicators to provide monitoring in Europe, found it difficult to  
59 arrive at a small set of indicators due to the complexity of soil biota and functions  
60 (Bispo *et al.* 2009). Faber *et al.* (2013), investigating the level of consensus in choice  
61 of indicator found that across Europe the use of different indicators varied from  
62 country to country, and where an indicator was used frequently (for example  
63 nematodes as a faunal indicator) this was driven by innate inclusion in a large scale  
64 project, rather than intrinsic potential of using an indicator alone.

65

66 Working to solve this problem, Ritz *et al.* (2009) developed a ‘logical sieve’  
67 procedure to select indicators in a structured manner and without bias, and applied it  
68 to the context of soil monitoring at the specific scale of the UK. The logical sieve  
69 method formally considers the relation of a candidate indicator to a range of criteria  
70 based upon applicability, discrimination ability, and involvement in more than one  
71 function, plus a range of technical attributes associated with a method. It then invokes  
72 a series of sieving processes to exclude those not fit for purpose, and ranks the  
73 remaining indicators on a numerical basis. This expansion of selection criteria away  
74 from the more common technically-focused assessment (i.e. how easy or cheap is a  
75 specific method), gives the results of a logical sieve greater robustness but also  
76 provides an audit trail to the ranking that allows flexibility in adjusting parameters as  
77 needs or knowledge changes.

78

79 Of course, a single monitoring scheme cannot be a perfect fit to all scenarios. The  
80 logical sieve approach is advantageous, since the method is run on a specific scenario  
81 which has been laid out in advance, and the scenario can be changed and the process  
82 re-run without much additional effort. This provides different suites of indicators for  
83 different scenarios making the process flexible and more useful. Ritz *et al.* (2009)  
84 deployed the logical sieve on a scenario of a monitoring scheme for soil quality within  
85 the UK, incorporating the then contemporary state of methods developed for  
86 measuring soil biological indicators. Expanding on this paper, and working with the  
87 concepts and equations laid down for the logical sieve, we investigated whether a  
88 suite of indicators could be found which would enable a soil quality monitoring  
89 scheme to function over the whole of the European land area. This represents a more  
90 complex scenario than the one laid out in Ritz *et al.* (2009) by encompassing a wider  
91 range of climatic zones. The sieve itself was also updated to take into account  
92 developments in methods for measuring biological indicators of soil quality.

93

94 In the last decade with the rise of molecular techniques in particular, the suite of  
95 available biological indicators has changed, whilst at the same time there have been  
96 advances in knowledge of the functional roles of below-ground biodiversity. (Bardgett  
97 & van der Putten, 2014). This was foreshadowed by Ritz *et al.* (2009), but the pace of  
98 development has accelerated in recent years with massive advances in technology and  
99 the reduction in the costs of using molecular tools, which continue apace (Orgiazzi *et*  
100 *al.*, 2015). For example, with regard to next generation sequencing (NGS) three new  
101 platforms were released in 2011 alone: Ion Torrent's PGM Pacific Biosciences' RS  
102 and the Illumina MiSeq (Quail *et al.*, 2012). Alongside these advances in technology  
103 are developments of bio-informatics tools that manage large data flows compare data

104 with specialised databases and extract relevant information thus opening up new  
105 perspectives for investigating the soil microbiome (Uroz *et al.*, 2013).

106

107 This work expands on in the list of indicators put forward by Faber *et al.* (2013),  
108 incorporates updates on logical sieve selection procedures and is then applied to a  
109 global-range survey of the soil biodiversity community in order to select indicators for  
110 future use in monitoring schemes.

111

## 112 **2. Method**

113 The logical sieve technique (Ritz *et al.*, 2009) was applied to the list of indicators  
114 originally described in Faber *et al.* (2013) and included here as Table 1. These  
115 indicators were shortlisted from the suite used in Ritz *et al.* (2009), using analysis of  
116 meta-data of soil biodiversity indicators used across Europe (Faber *et al.*, 2013).  
117 Where indicators could be measured or assessed by more than one method, these were  
118 included as separate indicator/method combinations (Table 1). Given that genetic-  
119 based molecular methods are in a huge state of flux in terms of development (i.e.  
120 resolution, throughput, interpretation), indicators based on these approaches were sub-  
121 categorised in relation to the underlying objective of the approach. Fingerprinting  
122 methods (terminal restriction fragment length polymorphism (TRFLP) and automated  
123 ribosomal intergenic spacer analysis (ARISA) were included as an overall means of  
124 measuring microbial community structure. Pyrosequencing was included as a broad  
125 scale approach to soil biodiversity in general. While a molecular method for  
126 measuring the biodiversity of bacteria and archaea (which could include TRFLP,  
127 ARISA and pyrosequencing) was included as a specific determination on these  
128 microbial groups (Table 1).

129 A scenario was established to outline the purpose of the monitoring for which the  
130 indicators were being assessed. The logical sieve methodology was applied to  
131 establish key indicators for the assessment and potential monitoring of ecosystem  
132 functions specifically: (1) habitat for soil biodiversity; (2) C cycling and storage; (3)  
133 cycling of nitrogen (N) and phosphorus (P) across the whole productive non-urban  
134 land area of European member states for use in future potential monitoring.

135

136 Following on from the work of Faber *et al.* (2013) where indicators were scored  
137 against criteria taken from OECD (2002); UNEP (2007), Ritz *et al.* (2009); and Turbé  
138 *et al.* (2010) an online questionnaire was developed. The original questions used in  
139 Faber *et al.* (2013) were based on responses to a pilot run of the survey and  
140 communication with a small panel of experts in the field. The questionnaire used in  
141 this study expanded this initial study using additional criteria adapted from the logical  
142 sieve of Ritz *et al.* (2009) in order to provide a complete assessment of each indicator.

143

144 To facilitate the assessment of indicators by those completing the survey a set of  
145 choices were developed for answering each question. Each question had three  
146 possible answers consisting of the ‘best possible’ option a ‘medium’ option that was  
147 neither ideal nor problematic for the measurement of the indicator and a ‘worst-case’  
148 option where the measurement of the indicator was affected. An additional answer of  
149 ‘essentially no knowledge’ was included for all questions allowing respondents to  
150 indicate that they considered that they did not know enough about a certain indicator  
151 to answer the question with sufficient confidence.

152

153 The final questionnaire contained 15 questions with associated answer choices (Table  
154 2) and was html coded as a web survey by the Joint Research Centre (JRC) of the  
155 European Commission. The questionnaire was made available online for three  
156 months (April-June 2014) by JRC and was open to all who wished to participate. It  
157 was advertised through the Global Soil Biodiversity Institute network  
158 (<http://www.globalsoilbiodiversity.org>), EcoFINDERS project partners, and European  
159 research scientists who work in the field of soil biodiversity and ecosystem function,  
160 and related fields. JRC collated the results of the web survey and returned the raw  
161 data whereupon it was amalgamated, checked, sorted and finally passed through the  
162 equations of the logical sieve as per Ritz *et al.* (2009). These equations are  
163 summarised in Figure 1.

164

165 Questions relating to skill requirements measurement equipment requirements and  
166 costs were used to calculate the technical factor ( $F_T$ ). The weighting factors ( $WC_i$ )  
167 used in the calculation of  $F_T$  were taken from Ritz *et al.*, (2009) which were the result  
168 of stakeholder consultation with the scientific community and likely end-user public-  
169 bodies (reproduced here in Table 2).

170

171 Questions 12 and 14 (Table 3) were incorporated into the discrimination factor [ $S_D$ ]  
172 and taken forward into the applicability/discrimination factor [ $F_{AD}$ ]. A soil function  
173 factor for each indicator was calculated by the amalgamation of scores allocated by a  
174 small panel of experts for each indicator. A final aggregated factor [ $F_A$ ] was  
175 calculated for each indicator.

176

177 **3. Results**

178 There were 61 respondents to the web survey. Three of the responses were rejected  
179 because the participants supplied non-usable answers (i.e. no knowledge). The  
180 responses of the remaining 58 were taken forward for analysis in the logical sieve.

181

182 3.1 Response description

183 The majority of indicator/question combinations received between 25 and 45  
184 respondents out of a possible 58 (range 43-78%). Many participants used the ‘no  
185 knowledge’ option for one or more of the indicator/question combinations. No  
186 question was answered for all indicators by all participants.

187 The lowest response for an indicator/question combination was 14 respondents for a  
188 question on standardisation [Q5] using the indicator of ‘Mites (by molecular  
189 techniques)’. The highest response was 51 respondents for a question regarding  
190 equipment needed [Q2] to measure the indicator of ‘Respiration (using methods to  
191 measure basal respiration)’. Q2 and Q15 (understandability) were answered by the  
192 greatest proportion of respondents.

193

194 Molecular indicators generally had a poor response, overall with the lowest proportion  
195 with regard to questions on standardisation (Q5, 14, 15, 17, 19: all under 20  
196 respondents).

197 Only two indicators returned 50 or more responses (‘Respiration (all basal methods)’  
198 as above and ‘Capital start up needed [Q6] for litter bags’).

199

200 Indicators with a median response value under 30 included mainly molecular methods  
201 ('Enchytraeids', 'Mites', 'Collembola', and 'Protozoa') and 'ChipTechnology'.  
202 These indicators were also ranked lowly in the technical factor.

203

204 An average response was calculated per indicator/question combination. 'No  
205 knowledge' answers were removed from the dataset before averages were calculated.

206

### 207 3.2 Technical rankings

208 Technical Factor Scores [ $F_T$ ] were calculated for each indicator using the sum of  
209 weighted responses to questions (Table 4). Measurements of soil respiration scored  
210 highest with molecular methods scoring poorly over the weighted responses to  
211 technical questions.

212

213 The top three indicators after ranking by  $F_T$  scores ('Basal respiration', 'Substrate  
214 induced respiration' and 'Nitrification potential') all scored highly on four particular  
215 questions. These four questions addressed labour in the lab (Q7: 'Does not require  
216 senior technical staff or an experienced researcher') analysis costs (Q8: 'Low cost per  
217 sample in terms of consumables') ease of sampling (Q11: 'A short one off sampling  
218 occasion in the field with a low volume of soil needed') and reproducibility of results  
219 (Q13: 'Reproducible results across different laboratories'). Though the top indicators  
220 scored more than 2 on a number of technical questions they all scored an average of  
221 more than 2.5 out of a possible 3 in these four specific areas.

222

223 There was no bias of either cost questions or method of measurement questions within  
224 the technical factor scores. Weighted technical scores were grouped by

225 'Measurement' (6 questions) and by 'Cost' (6 questions) and a single index value  
226 produced (i.e. one measurement value as a % of the possible total weighted score and  
227 one cost value as a % of the possible total weighted score). These two index values  
228 were added together to produce a  $F_T$  score that had an equal contribution from  
229 measurement and cost questions. The top ten indicators then remained essentially the  
230 same. The top five indicators did not change in any way; indicators ranked 6 and 7  
231 changed places ('Respiration measured by MSIR' and 'Litter bags' respectively);  
232 indicators ranked 8 and 9 also switched places ('Collembola identified by  
233 morphological methods' and 'Molecular microbial biomass' respectively).

234

### 235 3.3 Applicability and discrimination

236 Applicability/Discrimination Factor Scores [ $S_A$ ] and [ $S_D$ ] were calculated for each  
237 indicator using scores for applicability taken from Ritz *et al.* (2009) and responses to  
238 questions about discrimination ability of indicators (Table 5).

239

240 Molecular indicators of biodiversity and indicators of ecosystem function scored  
241 highly in discrimination [ $S_D$ ]. The top three indicators after ranking by  
242 Discrimination Factor Scores [ $S_D$ ] ('Multiple enzyme assay', 'PLFA', and 'Molecular  
243 methods for assessing bacteria and archaea') all scored highly on overall sensitivity  
244 (Q14: indicator is sensitive to more than one of either: land use disturbance or soil  
245 type).

246

247 In terms of Applicability [ $S_A$ ] 'Earthworms' and 'Enchytraeids' were both allocated  
248 low  $S_A$  scores due to the fact that they are not ubiquitous to all European bio-  
249 geographical zones.

250

### 251 3.4 Functionality

252 To assess the usefulness of the indicators in describing changes in functionality of soil  
253 experts were asked to review the influence values taken from Ritz *et al.* (2009) for the  
254 following functions: (1) Food and fibre production [ $S_{FF}$ ]; (2) Environmental  
255 interactions [ $S_{EI}$ ]; (3) Habitat for biodiversity [ $S_{HB}$ ]. The updated influence values  
256 were developed taking into account new methods for analysis and new knowledge of  
257 ecological importance of the indicators involved.

258

259 To test indicator rankings for all functions, scores for each function were multiplied  
260 together to create a Function factor score [ $F_{SF}$ ] (Table 6a). This sieved out indicators  
261 ranked ‘not pertinent’ (0) for any of the three functions in order to assess the overall  
262 potential across functions. In this ranking ‘Potential nitrification’ and ‘Protista  
263 diversity’ scored poorly: ‘Nitrification potential’ due to low influence in the function  
264 of habitat and biodiversity provision, and ‘Protista diversity’ due to low influence in  
265 the function of  $S_{FF}$ .

266

267 The second ranking assumed that functions were equal and indicators did not need to  
268 hold influence over *all* functions in order to score highly. In this ranking  $F_{SF}$  was  
269 calculated from the average of the three functions (Table 6b). ‘Nitrification potential’  
270 scored better in this method of assessing functionality.

271

272 ‘Earthworms’, ‘Fungi’, and ‘Bacteria biodiversity’ indicators scored highly in both.

273

### 274 3.5 Aggregated Factor Scores

275  $F_{AD}$  was calculated from  $S_A$ ,  $S_D$  and  $F_{SF}$  scores (using both multiplied and averaged  
276 [ $F_{SF}$ ] values). The Aggregated factor [ $F_A$ ] was then calculated for each indicator by  
277 multiplying  $F_{AD}$  by  $F_T$  (Figures 2 and 3 for multiplied and averaged [ $F_{SF}$ ] values  
278 respectively).

279

280 As the Aggregated factor scores [ $F_A$ ] were calculated from previous factors scores  
281 those indicators that performed poorly in previous rankings were found at the bottom  
282 of the  $F_A$  rankings. These included ‘Earthworms’ (ranked lowly in  $S_A$  – applicability)  
283 and ‘Nitrification potential’ and ‘Protozoa biodiversity’ (ranked lowly in  $F_{SF}$  –  
284 function).

285

286 The top ten indicators included four indicators of biodiversity (three microbial and  
287 one meso faunal) by various methods of measurement and three indicators of  
288 ecological function (multiple enzyme assay, multiple substrate induced respiration  
289 and the exploration of functional genes by molecular means). Within the techniques  
290 assessed, seven out of the top ten indicators made use of molecular methods.

291

292 Other soil fauna biodiversity indicators (‘Nematodes’ and ‘Micro-arthropods’) were  
293 found in the middle of the rankings, above the remaining indicators for ecological  
294 functions such as ‘Biolog’ and ‘Basal respiration’ (not ranked highly in  $F_{SF}$ ) or ‘Litter  
295 bags’ and ‘Bait lamina’ (ranked lowly in  $S_D$  – discrimination).

## 296 **4 Discussion**

### 297 4.1 Response description

298 Any voluntary survey such as this contains the potential for self-selection bias and we  
299 acknowledge that respondents to the survey potentially had an interest in the outcome.

300 In part, the survey was designed to be self-selecting in order to access the knowledge  
301 and expertise of those specialising in the monitoring of soil biodiversity and soil  
302 ecosystem functioning. However, bearing this in mind, various measures were  
303 incorporated into the survey to reduce the effect of any potential bias.

304 Method assessment on a technical level was broken down into small discrete areas  
305 with clear and measurable criteria for assessment (cost of laboratory equipment,  
306 labour time in the field etc.). This reduced the number of qualitative responses and  
307 the potential for participants to over mark their favourite method.

308 Questions relating to sensitivity or reproducibility of methods will always ultimately  
309 be subjective. To address this, the design of the survey restricted participants to three  
310 categories of response with clear criterion guides for each category (plainly presented  
311 on the screen alongside the question), in order to prevent subtle over- or under-  
312 marking. To address blatant such marking, the survey was advertised to as broad a  
313 group of potential participants as possible in order to balance the bias between areas  
314 of research interest. The authors themselves had no bias to any particular group and  
315 advertised the survey amongst general scientific communities through mailing lists  
316 and fora in order to get a good range of technical backgrounds. In addition, though  
317 some indicators drew a less numerous response, by creating an average value per  
318 question per indicator, the potential for the number of respondents assessing an  
319 indicator to affect the overall value for that indicator was reduced.

320 It was clear from the answers collected by the web survey that scientists though  
321 knowledgeable in their field cannot be expected to have the information to answer  
322 questions about every selected indicator of soil biodiversity and ecosystem function.  
323 The use of the 'no knowledge' response at some point by all participants in some  
324 cases for all indicators except one confirmed that the incorporation of this option was

325 necessary and its general use is reassuring in the sense that such a return would be  
326 anticipated if respondents were being duly considerate in their returns. Without the  
327 option to not answer the question the survey would likely have been biased by  
328 guessed answers and not performed the function it was designed for.

329 It is however, noted that the potential for self-selection bias cannot be completely  
330 denied when acting on the conclusions drawn by this paper.

331

#### 332 4.2 Technical rankings

333 Technical factor scores [ $F_T$ ] give an indication of the practicality associated with the  
334 measurement of a particular indicator. The questions used for  $F_T$  scores are those  
335 most commonly assessed when choosing an indicator i.e. cost and difficulty of the  
336 measurement method. Indicators considered to have a high  $F_T$  are generally  
337 incorporated into monitoring schemes without further assessment. As this paper  
338 shows the indicators in the top 10  $F_T$  scores were not always those which scored  
339 highest overall. This is an advantage of the logical sieve method. The top ten  
340 indicators after ranking by  $F_T$  scores are dominated by measurements of ecosystem  
341 function. Methods used in the measurement of ecosystem function are often simple  
342 and cost effective especially when compared to the high cost of molecular techniques  
343 and the high labour demand associated with soil fauna identification (Wu *et al.* 2009).

344

#### 345 4.3 Applicability and discrimination

346 Applicability [ $S_A$ ] tests the ubiquitous nature of each indicator and thus forms an  
347 intrinsic part of the scenario tested. In the scenario tested in this paper covering the  
348 whole productive non-urban land area of European member states, the restricted range  
349 of earthworms was a major factor, resulting in a low ranking overall despite high  $F_T$

350 scores. Having evidence of why an indicator cannot be recommended for a  
351 monitoring scheme (in this case due to non-presence in some monitoring sites) gives  
352 the potential for the scheme to be adapted before implementation. This highlights  
353 another benefit of the logical sieve in that indicators are ranked and rejected for  
354 practical reasons and the sieve could always be re-run with changed parameters to  
355 adjust the indicators recommended in accord with the context (particularly spatial  
356 scale and range of biomes involved).

357

358 Discrimination [ $S_D$ ] tests the sensitivity of each indicator to a wide range of  
359 environmental conditions. Five of the top ten indicators within this category were for  
360 ecological function (two for soil respiration which is an indicator for C cycling and  
361 storage). Different soils types and management schemes produce soils with different  
362 abilities to perform ecosystem functions. Indicators which measure ecosystem  
363 function are therefore intrinsically able to discriminate between these different soil  
364 conditions and are often recommended for monitoring schemes (Faber *et al.*, 2013).

365

366 Three indicators of biodiversity (two indicators of microbial diversity plus ‘Nematode  
367 diversity’) were placed in the top 10 rankings. A great deal of species data is  
368 collected when molecular methods to determine biodiversity are employed. This  
369 species data can be used to create a more detailed picture of the community present  
370 and with the massive niche diversity in soil habitats species diversity varies from  
371 scenario to scenario leading to high levels of discrimination. In addition if trait  
372 information is available species data can be linked to soils’ ability to perform certain  
373 ecological functions leading to greater discrimination between them.

374

#### 375 4.4 Functionality

376 Functionality [ $F_{SF}$ ] changes the recommended indicators dependant on the function or  
377 functions wanted, removing certain indicators from the logical sieve that might  
378 otherwise be recommended. The example of the indicator ‘Potential nitrification’  
379 shows how vital it is to be clear about which functions should be included in the  
380 scenario used for each logical sieve process. When the scenario was that *all* functions  
381 needed to be monitored by the indicators selected, potential nitrification scored poorly  
382 despite being reasonably ranked in all other factors. However, when the scenario was  
383 changed to a looser monitoring of three functions but each indicator was not required  
384 to be able to monitor all three, ‘Potential nitrification’ scored higher in the ranking.

385

386 If the criteria for selection requires indicators necessarily that are ubiquitous in terms  
387 of function then it is the molecular methods for determining the biodiversity of the  
388 microbial community that score most highly. If the guidelines are a little looser,  
389 allowing the inclusion of indicators that are relevant to less than all of the functions  
390 required then methods of measuring ecological function rise up the rankings.

391

#### 392 4.5 Aggregated factor scores

393 The top ten indicators after the full suite of sieves was applied ( $[F_A]$  scores) were a  
394 mix of measures of soil biodiversity and soil ecosystem function. When re-examining  
395 the sieving processes it becomes clear that this is due to the discrimination potential  
396 and applicability to a range of functions for biodiversity indicators, and a mix of  
397 technical factors and discrimination potential for indicators of soil ecosystem function  
398 provision. The top ten rankings were dominated by molecular methods of measuring  
399 both biodiversity and ecological function (seven indicator/method combinations).

400 This is interesting, bearing in mind that in terms of responses to the questionnaire,  
401 molecular indicators generally had a poor response overall when compared to the  
402 number of respondents who answered questions on other indicators. The results of  
403 this suggest that despite rapid advances in molecular techniques over the last decade,  
404 the potential of molecular methods has not yet percolated out to the majority of the  
405 scientific community surveyed in this questionnaire.

406

407 Within the top indicators for biodiversity, both ‘Bacteria and Archaea diversity’ and  
408 ‘Fungi diversity’ were only in the middle rankings for Technical factors [F<sub>T</sub>] but  
409 scored in the top five for Discrimination potential [S<sub>D</sub>] and high for relevance to  
410 function [F<sub>SF</sub>] (‘Bacteria and Archaea’ were number one). ‘Mite biodiversity’ scored  
411 poorly in [F<sub>T</sub>] and was only in the middle of the table for [S<sub>D</sub>], but scored in the top  
412 ten for relevance to function [F<sub>SF</sub>]. Though methods of determining biodiversity are  
413 often time consuming and/or costly (Lawton *et al.*, 1998, Andre *et al.*, 2001 with  
414 regard to ‘scientist hours’ needed for morphological identification), it can be seen  
415 from the results of this exercise that the potential of these indicators in terms of  
416 discrimination and relevance to more than one function results in them being  
417 recommended for use as indicators for monitoring schemes.

418

419 Both ‘Respiration (Multiple Substrate Induced Respiration)’ and ‘Molecular microbial  
420 biomass’ were in the top ten for Technical rankings [F<sub>T</sub>] and the top ten for  
421 Discrimination potential [S<sub>D</sub>]. ‘Functional Genes’ and ‘Multiple enzyme assay’ were  
422 mid-range in Technical rankings [F<sub>T</sub>] but in the top five for Discrimination potential  
423 [S<sub>D</sub>]. ‘Functional genes’ also scored within the top five for Function factor scores  
424 [F<sub>SF</sub>]. With regard to indicators for ecological function, methods such as ‘MSIR’ or

425 'Molecular microbial biomass', which scored well in ease of use and discrimination  
426 potential, are restricted in the number and variety of functions to which they are  
427 relevant and are therefore not ranked in the top five. Whereas 'Functional genes' and  
428 'Multiple enzyme assays', though not ranked as highly in Technical rankings [F<sub>T</sub>],  
429 scored highly both in discrimination potential and in relevance to more than one  
430 function. This resulted in these two indicators being ranked higher in the overall  
431 standings.

432

433 When not using the logical sieve, it is often indicators that score highly in Technical  
434 factor [F<sub>T</sub>] criteria (those considered the easiest to measure) that are selected for  
435 monitoring schemes. However, as seen in this paper, when comparing indicators with  
436 high F<sub>T</sub> scores as opposed to those that were still ranked highly at the end of the  
437 sieving process, this can produce a decision making process that fails to take into  
438 account low discrimination ability of an indicator or does not recognise that an  
439 indicator may not be present across the full range of environments found within the  
440 monitoring scheme.

441

442 The functions selected within the scenario, in this case soil biodiversity habitat, C  
443 cycling and nutrient cycling, have an intrinsic effect on the highest ranking indicators.  
444 This allows the logical sieve method to be precisely calibrated for each monitoring  
445 scheme to provide scheme specific indicators.

446

## 447 **5 Conclusion**

448 For the scenario provided, i.e. the assessment and potential monitoring of ecosystem  
449 functions specifically: (1) habitat for soil biodiversity; (2) C cycling and storage; (3)

450 cycling of nitrogen (N) and phosphorus (P) across the whole productive non-urban  
451 land area of European member states, the top ranking biological indicator was  
452 'Bacteria and archaea diversity measured by molecular methods'.

453

454 The top ten indicators included three related to biodiversity ('Bacteria and archaea',  
455 'Fungi' and 'Mites') and four related to ecological function ('Functional genes',  
456 'Multiple enzyme assay', 'Respiration (multiple substrate induced respiration)' and  
457 'Molecular microbial biomass'). The list was dominated by molecular methods to  
458 measure both biodiversity and ecological function (7 indicator/method combinations).

459

460 These selected indicators are recommended only under advisement. As shown in this  
461 paper, it is only when applying the full suite of logical sieves that the indicators  
462 selected can be termed fit for monitoring the (precise) scenario intended. Careful  
463 prescription of the scenario is also important as the scale of the monitoring  
464 programme, the level of discrimination needed, and the need for pertinence to more  
465 than one function, all have the potential to change the ranking of indicators.

466

#### 467 **Acknowledgements**

468 This work was supported by the European Union within the projects EcoFINDERS  
469 (FP7-264465), the Joint Research Council (JRC), the Scottish Government Rural and  
470 Environment Science and Analytical Services Division Theme 3 Land Use. We  
471 sincerely thank the individual respondents to the survey questionnaires, and members  
472 of the Joint Research Centre, and the Global Soil Biodiversity Initiative, for their  
473 assistance and engagement.

474

475 **References**

476 Andre, H.M., Ducarme, X., Anderson, J.M., Crossley, D.A., Koehler, H.H., Paoletti,  
477 M.G., Walter, D.E., Lebrun, P., 2001. Skilled eyes are needed to go on studying the  
478 richness of the soil. *Nature*. 409, 761.

479

480 Bardgett, R.D., & van der Putten, W.H., 2014. Belowground biodiversity and  
481 ecosystem functioning. *Nature*. 515, 505–511

482

483 Bispo, A., Cluzeau, D., Creamer, R., Dombos, M., Graefe, U., Krogh, P.H., Sousa,  
484 J.P., Peres, G., Rutgers, M., Winding, A., and Rombke, J., 2009. Indicators for  
485 monitoring soil biodiversity. *International Environmental Assessment Policy and*  
486 *Management*. 5(4), 717-719

487

488 Faber, J.H., Creamer, R.E., Mulder, C., Römcke, J., Rutgers, M., Sousa, J.S., Stone,  
489 D., Griffiths, B.S., 2013. The Practicalities and Pitfalls of Establishing a  
490 Policy-Relevant and Cost-Effective Soil Biological Monitoring Scheme. *Integrated*  
491 *Environmental Assessment and Management*. 9(2), 276–284

492

493 Lawton, J.H., Bignell, D.E., Bolton, B., Bloemers, G.F., Eggleton, P., Hammond,  
494 P.M., Hodda, M., Holt, R.D., Larsen, T.B., Mawdsley, N.A., Stork, N.E., Srivastava,  
495 D.S., Watt, A.D., 1998. Biodiversity inventories, indicator taxa and effects of habitat  
496 modification in tropical forest. *Nature*. 391, 72-76.

497

498 Morvan, X., Saby, N.P.A., Arrouays, D., Le Bas, C., Jones, R.J.A., Verheijen, F.G.A.,  
499 Bellamy, P.H., Stephens, M., Kibblewhite, M.G., 2008. Soil monitoring in Europe: A

500 review of existing systems and requirements for harmonisation. *Science of The Total*  
501 *Environment*. 391(1), 1-12

502

503 [OECD] Organization for Economic Co-Operation and Development, 2002. Report on  
504 the OECD expert meeting on agri-biodiversity indicators. November 2001: Summary  
505 and recommendations. OECD Joint Working Party on Agriculture and Environment  
506 COM/AGR/CA/ENV/EPOC (2002) 35. Paris: OECD.

507

508 Orgiazzi, A., Dunbar, M.B., Panagos, P., de Groot, G.A., Lemanceau, P., 2015. Soil  
509 biodiversity and DNA barcodes: opportunities and challenges. *Soil Biology and*  
510 *Biochemistry*. 80, 244-250

511

512 Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni,  
513 A., Swerdlow, H.P., and Gu, Y., 2012. A tale of three next generation sequencing  
514 platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq  
515 sequencers. *BMC Genomics*. 13, 341

516

517 Ritz, K., Black, H.I.J., Campbell, C.D., Harris, J.A., Wood, C., 2009. Selecting  
518 biological indicators for monitoring soils: A framework for balancing scientific and  
519 technical opinion to assist policy development. *Ecological Indicators*. 9, 1212-1221

520

521 Turbé A., De Toni, A., Benito, P., Lavelle, P., Ruiz, N., Van der Putten, W.H.,  
522 Labouze, E., Mudgal, S., 2010. Soil biodiversity: Functions, threats and tools for  
523 policy makers. Bio Intelligence Service, IRD, and NIOO, Report for European  
524 Commission (DG Environment).

525

526 UNEP/CBD (2007). In-depth review of the programme of work on agricultural  
527 biodiversity – Note by the Executive Secretary., UNEP.

528

529 Uroz, S., Ioannidis , P., Lengelle, J., Cébron , A., Morin, E., Buée, M., Martin, F.,  
530 2013. Functional Assays and Metagenomic Analyses Reveals Differences between the  
531 Microbial Communities Inhabiting the Soil Horizons of a Norway Spruce Plantation.

532 *PLOS* *ONE*.

533 <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0055929>

534

535 Wu, T., Ayres, E., Li, G., Bardgett, R.D., Wall, D.H., Garey, J.R., 2009. Molecular  
536 profiling of soil animal diversity in natural ecosystems: Incongruence of molecular  
537 and morphological results. *Soil Biology & Biochemistry*. 41, 849–857

538 Table 1. List of indicators processed by logical sieve and their importance to either  
 539 Biodiversity or Ecosystem Function.

<b>Indicator</b>	<b>Biodiversity</b>	<b>Function</b>
<b>Macro and Meso faunal diversity</b>		
Earthworms (Morphological identification or Molecular methods)	X	X
Enchytraeids (Morphological identification or Molecular methods)	X	X
Mites (Morphological identification or Molecular methods)	X	X
Collembola (Morphological identification or Molecular methods)	X	X
Nematodes (Morphological identification or Molecular methods)	X	X
Protista diversity (Morphological identification or Molecular methods)	X	X
<b>Microfaunal diversity</b>		
Bacteria and Archaea species by molecular methods	X	
Fungi species (Morphological identification or Molecular methods),	X	
Bacteria & Fungi diversity through fingerprint methods (TRFLP, ARISA)*	X	
Pyrosequencing, of soil DNA**	X	X
PLFA,	X	
<b>Ecosystem Function performed by soil biology</b>		
Functional Genes (targeting antibiotic producers, nitrifiers, denitrifiers)		X
ChipTechnology (up regulation or down regulation of specific genes related to Ecosystem Function)		X
Molecular microbial biomass		X
Respiration (All basal methods),		X
Respiration (SIR-Glucose),		X
Respiration (Multiple Substrate Induced Respiration),		X
Respiration (Biolog)		X
Nitrification potential		X
Multiple Enzyme Assay,		X
Bait Lamina,		X
Litter Bags		X

540 \*As a method of assessing microbial community structure

541 \*\*Broad-spectrum pyrosequencing of whole soil DNA

542

543 Table 2. The weighting factors (WC<sub>i</sub>) used in the calculation of F<sub>T</sub>, taken from Ritz *et*  
 544 *al.*, (2009) using the question: “What *weighting* would you assign to the criterion  
 545 when considering a trans-UK (cross-habitat) measuring and monitoring programme?  
 546 Weight from 0 (i.e. dismiss entirely) to 4 (maximum relative weight).”

	CATEGORY:	Weight	Question in sieve
1	THROUGHPUT: <i>How important is it to be able to have a high level of throughput (i.e. 100's per week) e.g. 0 = not important - dismiss 1 = relevant but not essential 2 = valuable but not essential 3 = valuable and preferred 4 = vital!</i>	3	Q9
2	STORAGE: <i>How important is it to be able to store samples until they can be analysed, for up to 2 weeks post sampling?</i>	3	Not used
3	ARCHIVABILITY: <i>How important do you consider archiving of samples (or analytical products) e.g. for future monitoring comparisons or for currently unknown analyses to answer new questions?</i>	2	Q3
4	SAMPLE COLLECTION: <i>Does it matter that the site would need to be visited more than once for a particular method to get the data?</i>	3	Q11
5	HOW MUCH SOIL: <i>Smaller soil samples cost less, easier to sample and handle etc; is a smaller sample preferred?</i>	2	Q11
6	COST – HARDWARE: <i>Does it matter how much it costs, in terms of hardware, to analyse the soil?</i>	1	Q6
7	COST- LABOUR: <i>Does it matter how much it costs, in terms of people, to analyse the soil?</i>	3	Q7, Q9
8	EASE OF USE: <i>Is it important that the method is relatively easy to carry out?</i>	2	Not used
9	POTENTIAL REFERENCE MATERIAL: <i>How important is quality control (QC) via reference material?</i>	2	Q4
10	REPRODUCIBILITY OF RESULTS: <i>How much do you care about being able to reproduce the same results time after time?</i>	4	Q1, Q5, Q13
11	READY-TO-USE DEPLOYMENT STATUS: <i>Is it important that the method is well established and has standard operating procedures?</i>	0	Q5
12	INTERNATIONAL COMPARISONS: <i>If the method is used in soil monitoring schemes elsewhere, is this important for UK soil monitoring?</i>	2	Not used
13	UK INFRASTRUCTURE: <i>Is it important that we have the capacity at present to deliver this method?</i>	3	Not used

547

548

Table 3. Questionnaire questions and answer choices.

Question	Best Option	Middle Option	Worst Option
<p><b>Q1. Skills</b> What is the amenability of the method to ready application via a standard operating procedure when presented to a competent technician? Does it include a training element? (Ritz <i>et al.</i>, 2009)</p> <p><b>Q2. Equipment</b> Are you likely to find the required equipment in all labs (i.e. pH meter) or very few (confocal microscope).</p>	<p><b>Straightforward</b> No more than 1 day of training is required to carry out the analysis</p> <p><b>All laboratories</b> All laboratories can perform this method. Also applies to methods carried out in the field.</p>	<p><b>Moderate</b> A PhD student or skilled technician can carry out the analysis after no more than 1 week of training.</p> <p><b>Standard medium level laboratories</b> The method can be successfully performed in a mid-range laboratory</p>	<p><b>Specialised</b> A fully trained specialist is required full time to carry out the analysis</p> <p><b>Very few, specialised laboratories</b> The method can only be carried out in a specialised laboratory designed for that purpose</p>
<p><b>Q3. Archive</b> What is the potential for archiving samples (i.e. over decades) in order to accurately re-determine these properties? (Ritz <i>et al.</i>, 2009)</p>	<p><b>Air dried soil/ Fauna stored in Alcohol/Formalin</b> Samples can be archived as air dried soil or, if fauna, in alcohol/formalin.</p>	<p><b>Freeze dried/-80 freezer</b> Samples are only able to be archived for any length of time either freeze-dried, in a -20 degree freezer or in a -80 degree freezer.</p>	<p><b>Not able to be archived</b> Samples cannot be archived for the length of time described.</p>
<p><b>Q4. Potential reference material</b> Is the method amenable to the prescription and provision of reference material (Ritz <i>et al.</i>, 2009)</p>	<p><b>Readily available materials exist</b> Reference materials are available and affordable for most mid-range laboratories OR "in house" QCs are possible to use.</p>	<p><b>Yes, but materials are expensive/difficult to obtain</b> Materials are expensive (cost more than 500 Euros)</p>	<p><b>No</b> The method does not allow the use of a reference material.</p>
<p><b>Q5. Method Standardised</b> Is the method standardised (i.e. is there an ISO/SOP available?)</p>	<p><b>ISO available</b> An ISO is available and used by all.</p>	<p><b>Peer reviewed SOP in the literature</b> A clear and detailed SOP is available in the literature and generally followed by peers</p>	<p><b>Local lab SOP</b> The method follows an SOP that is only described within the local laboratory</p>
<p><b>Q6. Capital Start-up</b> What level of finance is required to equip a laboratory to perform this analysis? (Instrument cost only, not salary or</p>	<p><b>Low</b> Less than 2000 Euros</p>	<p><b>Moderate</b> Up to 50,000 Euros</p>	<p><b>Very Expensive</b> More than 50,000 Euro</p>

<p>training)*  <i>Additional notes:  In DNA sequencing methods, if you do your sequencing "in-house" then the cost of sequencing equipment should be included when answering this question. If sequencing is contracted out, this will be covered in the later question "cost per sample" (Q8).</i></p>	<p>Q7. Labour requirement  What are the human resource costs to realise the method and initial interpretation (including consideration of skill level required and associated salary). (Ritz <i>et al.</i>, 2009)</p>	<p><u>Low cost</u>  Technical staff at a salary of 25,000 Euros per year</p>	<p><u>Moderately expensive</u>  Senior Technical staff or Experienced Researcher at a salary of 35,000 Euros per year</p>	<p><u>Very Expensive</u>  Senior Researcher or Specialist at a salary of 50,000 Euros per year</p>
<p>Q8. Cost per sample (consumables)  What are the consumable resource costs to realise the method and initial interpretation. (Ritz <i>et al.</i>, 2009)  Cost per average sample. If cost varies with number of individuals per sample, then the cost of a mid-range sample should be recorded.  <i>Additional notes:  In DNA sequencing methods, if you do not do your sequencing "in-house" then the cost of contracting out sequencing should be included when answering this question.  Cost per sample in terms of labour is addressed in the next question (Q9).</i></p>	<p><u>Low cost</u>  Less than 2 Euros per sample</p>	<p><u>Moderately expensive</u>  Less than 20 Euros per sample</p>	<p><u>Very Expensive</u>  More than 20 Euros per sample</p>	<p><u>Very Expensive</u>  More than 20 Euros per sample</p>
<p>Q9. Labour intensity (laboratory)  How long does it take to carry out the method from start to finish, up to the point of production of initial data results?  Does not include pre-processing of sample or post processing of data.</p>	<p><u>Less than 1 day</u>  The method requires less than the full time labour of one person for one day.</p>	<p><u>1 day to 1 week</u>  The method requires either the full time labour of one person for more than one day OR takes more than one full day due to the length of time of some steps</p>	<p><u>More than 1 week</u>  The method requires either the full time labour of one person for over one week OR takes more than one week due to the length of time of some steps.</p>	<p><u>More than 1 week</u>  The method requires either the full time labour of one person for over one week OR takes more than one week due to the length of time of some steps.</p>
<p>Q10. Labour intensity (pre-processing)  How much pre-processing does the sample require before analysis (from field sample)?</p>	<p><u>&lt;1 hour</u>  Sample processing requires less than the full</p>	<p><u>1 day</u>  Sample processing requires either the full time</p>	<p><u>1 week</u>  Sample processing requires either the full time labour of</p>	<p><u>1 week</u>  Sample processing requires either the full time labour of</p>

	time labour of one person for one hour.	labour of one person for one day OR takes one full day due to the length of time of some steps.	one person for one week OR takes one week due to the length of time of some steps.
<p><b>Q11. Labour intensity (field)</b> How long does field sampling take? Consider: Is one stop sampling in the field tenable? What mass of soil is needed for sampling and determination?</p>	<p><u>Low</u> One- off sampling required AND less than 500 g needed from the field site AND sampling time less than one hour.</p>	<p><u>Medium</u> Either one-off sampling required, but this sampling is a composite of a large area OR &lt; 2.5 kg of sample is required OR sampling will take more than three hours.</p>	<p><u>High</u> Either timelapse sampling is required, therefore multiple visits over short time frame (&lt; 1 month) OR large amounts of sample are required (&gt;5 kg), OR sampling time in total (if multiple visits are required) will take more than six hours.</p>
<p><b>Q12. Field Variability</b> What is the inherent spatial variability of the parameter measured under average field conditions?</p>	<p><u>Low</u> A single sample per site is adequate. This category includes the taking of a single composite sample per site.</p>	<p><u>Moderate</u> Three to five individual field replicates are required within a single site to account for field variability. This does not apply to composite samples.</p>	<p><u>High</u> More than five individual field replicates are required within a single site to account for field variability. This does not apply to composite samples.</p>
<p><b>Q13. Reproducibility of results</b> What is the inherent ability for the method to generate reproducible results, given that full quality-control protocols are available and applied, including (assumed) availability of reference material.</p>	<p><u>High</u> Reproducible results are expected across different laboratories when following an SOP.</p>	<p><u>Moderate</u> Validation through an inter laboratory comparison is required for reproducibility of results.</p>	<p><u>Inherently poor</u> The method is so sensitive to operator controlled conditions that it cannot be reproduced across laboratories</p>
<p><b>Q14. Overall sensitivity</b> Is this indicator sensitive enough to be of use in monitoring programmes?</p>	<p><u>High sensitivity</u> The indicator is sensitive to more than one of either: land use, disturbance or soil type.</p>	<p><u>Moderate sensitivity</u> The indicator is sensitive to one of either: land use, disturbance or soil type.</p>	<p><u>Poor sensitivity</u> The indicator is not sensitive to either: land use, disturbance or soil type OR the indicator is too sensitive to one or more of these parameters such that it</p>

			<p>becomes a problem when undertaking monitoring programmes OR there is not enough data available in the literature to make a conclusion.</p>
<p>Q15. Understandability How understandable is this indicator in the area of soil monitoring for soil health and biodiversity?</p>	<p><b>Public</b> This indicator can be easily used to explain issues of soil health and biodiversity to the general public.</p>	<p><b>Policy makers/Land managers</b> This indicator can be used to explain monitoring of soil health and biodiversity to policy makers. And can be used by them in future soil legislation and policy documents.</p>	<p><b>Scientists</b> This indicator can be used to explain monitoring of soil health and biodiversity only to other scientists within the general fields of ecology, soil science, biology, and environmental chemistry.</p>

550 Table 4. Ranking of indicators by Technical Factor Scores [ $F_T$ ]. (Scores are out of a  
 551 potential total of 93)

<b>Indicator</b>	<b><math>F_T</math></b>
Respiration (All basal methods),	73.5
Respiration (SIR-Glucose),	73.2
Nitrification,	72.8
Earthworms-Morphology,	72.7
Biolog,	71.1
Litter Bags	70.3
Respiration (Multiple Substrate Induced Respiration),	69.9
Collembola-Morphology,	69.8
Molecular microbial biomass,	69.6
Bait Lamina,	69.2
Enchytraeids-Morphology,	68.9
Multiple Enzyme Assay,	68.4
Nematodes-Morphology,	68.1
Mites-Morphology,	67.7
PLFA,	67.2
Bacteria and Archaea-Molecular,	67.2
Fungi-Morphology,	65.8
Fungi-Molecular,	65.8
Bacteria & Fungi-fingerprints (TRFLP, ARISA,...),	65.3
Nematodes-Molecular,	64.3
Functional Genes (targetting antibiotic producers, nitrifiers, denitrifiers)	64.3
Collembola-Molecular,	63.8
Mites-Molecular,	63.5
Protozoa-Molecular, Bacteria and Archaea-Molecular,	63.4
Enchytraeids-Molecular,	63.4
Protozoa-Morphology,	62.4
Earthworms-Molecular,	60.7
ChipTechnology,	57.6
Pyrosequencing,	57.6

552

553 Table 5. Applicability/Discrimination Factor Scores [ $S_A$ ] and [ $S_D$ ]

Indicator	$S_D$
Multiple Enzyme Assay,	5.0
PLFA,	5.0
Bacteria and Archaea-Molecular,	4.7
Nitrification,	4.7
Fungi-Molecular,	4.7
Respiration (Multiple Substrate Induced Respiration),	4.6
Molecular microbial biomass,	4.6
Functional Genes (targetting antibiotic producers, nitrifiers, denitrifiers)	4.6
Nematodes-Molecular,	4.6
Respiration (SIR-Glucose),	4.5
Biolog,	4.5
Bacteria & Fungi-fingerprints (TRFLP, ARISA,...),	4.5
Mites-Molecular,	4.5
Pyrosequencing,	4.4
Fungi-Morphology,	4.3
Protozoa-Molecular,	4.3
ChipTechnology,	4.3
Nematodes-Morphology,	4.3
Respiration (All basal methods),	4.3
Collembola-Molecular,	4.2
Enchytraeids-Molecular,	4.0
Mites-Morphology,	4.0
Earthworms-Molecular,	3.8
Protozoa-Morphology,	3.8
Collembola-Morphology,	3.8
Earthworms-Morphology,	3.7
Enchytraeids-Morphology,	3.7
Litter Bags	3.7
Bait Lamina,	3.5
Indicator	$S_A$ †
Earthworms-Morphology, Earthworms-Molecular, Enchytraeids-Morphology, Enchytraeids-Molecular,	0.1
<b>Macro and Meso faunal diversity</b>	
Mites-Morphology, Mites-Molecular, Collembola-Morphology, Collembola-Molecular Nematodes-Morphology, Nematodes-Molecular Protozoa-Morphology, Protozoa-Molecular	1
<b>Microfaunal diversity</b>	
Bacteria and Archaea-Molecular Fungi-Morphology, Fungi-Molecular Functional Genes (targetting antibiotic producers, nitrifiers, denitrifiers) Bacteria & Fungi-fingerprints (TRFLP, ARISA,...) Pyrosequencing, ChipTechnology PLFA	1
<b>Ecosystem Function performed by soil biology</b>	
Molecular microbial biomass Respiration (All basal methods), Respiration (SIR-Glucose) Respiration (Multiple Substrate Induced Respiration), Biolog, Nitrification Multiple Enzyme Assay Bait Lamina, Litter Bags	1

554 †Value of 0.1 used for non-ubiquitous species rather than 0 as in Ritz *et al.* (2009) in

555 order not to lose these indicators from the sieve whilst indicating their problems.

556 Table 6a. Function Factor Scores [ $F_{SF}$ ] (Food and Fibre x Environmental interactions  
 557 x Habitat and Biodiversity provision) Total:  $2 \times 2 \times 2 = 8$

<b>Indicator</b>	<b><math>F_{SF}</math></b>
Bacteria and Archaea-Molecular,	8.00
ChipTechnology,	8.00
Earthworms-Molecular,	8.00
Earthworms-Morphology,	8.00
Functional Genes (targetting antibiotic producers, nitrifiers, denitrifiers)	8.00
Fungi-Molecular,	8.00
Fungi-Morphology,	8.00
Mites-Molecular,	8.00
Pyrosequencing,	8.00
Collembola-Molecular,	4.00
Collembola-Morphology,	4.00
Mites-Morphology,	4.00
Molecular microbial biomass,	4.00
Multiple Enzyme Assay,	4.00
Nematodes-Molecular,	4.00
Nematodes-Morphology,	4.00
Respiration (Multiple Substrate Induced Respiration),	4.00
Bait Lamina,	2.00
Biolog,	2.00
Enchytraeids-Molecular,	2.00
Enchytraeids-Morphology,	2.00
Litter Bags	2.00
Respiration (All basal methods),	2.00
Respiration (SIR-Glucose),	2.00
Bacteria & Fungi-fingerprints (TRFLP, ARISA,...),	1.00
PLFA,	1.00
Nitrification,	0.00
Protozoa-Molecular,	0.00
Protozoa-Morphology,	0.00

558 Table 6b. Function Factor Scores [ $F_{SF}$ ] (Average of Food and Fibre, Environmental  
 559 interactions, Habitat and Biodiversity provision) Total:  $(2 + 2 + 2)/3 = 2$

Indicator	$F_{SF}$
Bacteria and Archaea-Molecular,	2.00
ChipTechnology,	2.00
Earthworms-Molecular,	2.00
Earthworms-Morphology,	2.00
Functional Genes (targetting antibiotic producers, nitrifiers, denitrifiers)	2.00
Fungi-Molecular,	2.00
Fungi-Morphology,	2.00
Mites-Molecular,	2.00
Pyrosequencing,	2.00
Collembola-Molecular,	1.67
Collembola-Morphology,	1.67
Mites-Morphology,	1.67
Molecular microbial biomass,	1.67
Multiple Enzyme Assay,	1.67
Nematodes-Molecular,	1.67
Nematodes-Morphology,	1.67
Respiration (Multiple Substrate Induced Respiration),	1.67
Nitrification,	1.33
Bait Lamina,	1.33
Biolog,	1.33
Enchytraeids-Molecular,	1.33
Enchytraeids-Morphology,	1.33
Litter Bags	1.33
Respiration (All basal methods),	1.33
Respiration (SIR-Glucose),	1.33
Bacteria & Fungi-fingerprints (TRFLP, ARISA,...),	1.00
PLFA,	1.00
Protozoa-Molecular,	0.67
Protozoa-Morphology,	0.67

560

561

562 Figure 1. Equations of the logical sieve as per Ritz *et al.* (2009).

563

564 Figure 2. Aggregated Factor Scores  $[F_A]$  for the list of indicators assessed.  $[F_A]$  is  
565 calculated using *multiplied* Function Factor Score  $[F_{SF}]$  values and incorporates the  
566 ability of each indicator to be relevant to all the functions intended to be monitored.

567

568 Figure 3. Aggregated Factor Scores  $[F_A]$  for the list of indicators assessed.  $[F_A]$  is  
569 calculated using *averaged* Function Factor Score  $[F_{SF}]$  values and assess indicators in  
570 a scenario where they do not need to be relevant to all the functions intended to be  
571 monitored.

572

573

574 Figure 1.

575

576

$$F_A = F_{AD} \times F_T \quad (1)$$

577 Where

$$F_T = \sum_{i=1}^n (S_{Ci} \times W_{Ci}) \dots + \dots (S_{Cn} \times W_{Cn})$$

578

(2)

579 Where  $F_T$  is the technical factor;  $S_{Ci}$  is the individual score for the indicator/question

580 combination  $i$ ;  $W_{Ci}$  is the weighting value for the individual indicator/question

581 combination  $i$ ; and  $n$  = number of indicator/question combinations.

582

583 And

584

$$F_{AD} = S_A \times S_D \times F_{SF} \quad (3)$$

585 Where  $F_{AD}$  is the applicability/discrimination factor;  $S_A$  is the score for applicability,

586 taken from Ritz *et al.* (2009);  $S_D$  is the score for discrimination taken from the

587 averaged answer for questions 12 and 14; and  $F_{SF}$  is the soil function factor,

588 amalgamated from scorings given by a small panel of EcoFINDERS experts for each

589 indicator.

590

591 And

592

$$F_{SF} = S_{FF} \times S_{EI} \times S_{HB} \quad (4)$$

593 Where  $F_{SF}$  is the soil function factor;  $S_{FF}$  is the indicator score for the ecosystem

594 function of food and fibre production;  $S_{EI}$  is the indicator score for the ecosystem

595 function of environmental interaction; and  $S_{HB}$  is the indicator score for the ecosystem

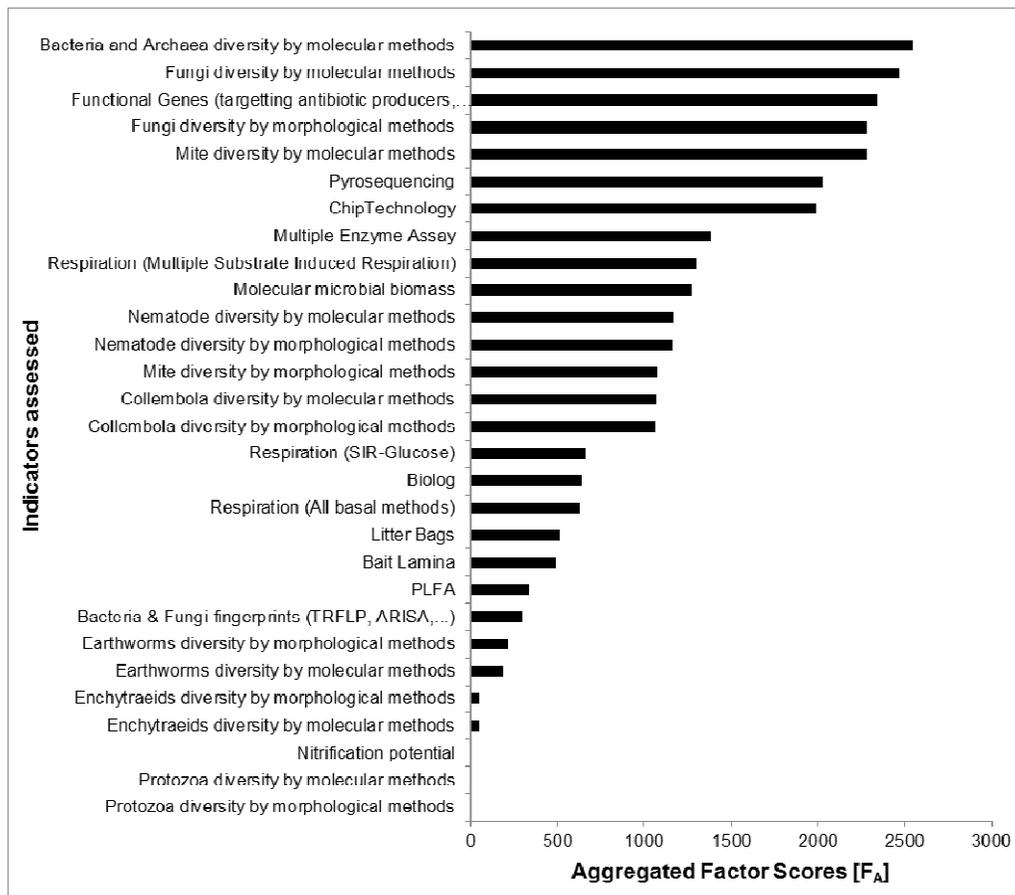
596 function of promotion of habitat and biodiversity. The scores for each indicator for

597 these soil functions were amalgamated from scorings given by a small panel of

598 EcoFINDERS experts for each indicator.

599

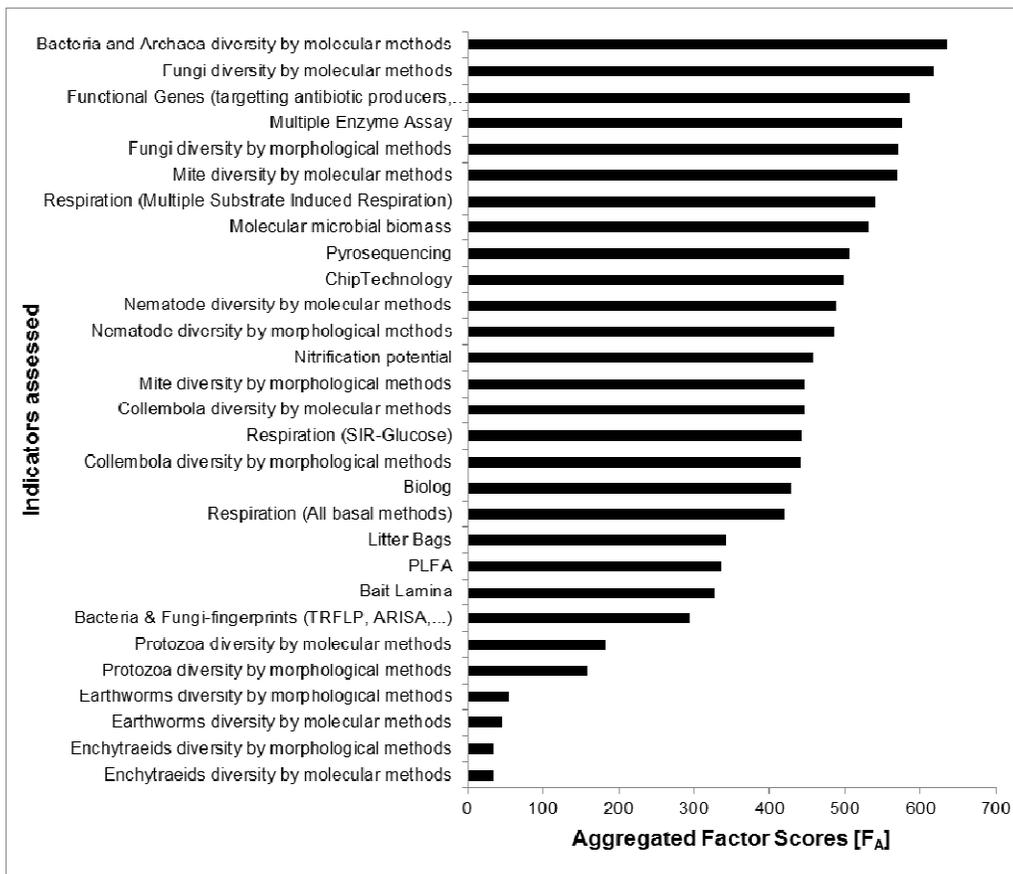
600 Figure 2



601

602

603 Figure 3



604