microbial biotechnology

Microbial Biotechnology (2012) **5**(1), 135–141 doi:10.1111/j.1751-7915.2011.00307.x

Brief report

Molecular biomass and MetaTaxogenomic assessment of soil microbial communities as influenced by soil DNA extraction procedure

Sébastien Terrat,[1,2] Richard Christen,[3,7] Samuel Dequiedt,[2] Mélanie Lelièvre,[2] Virginie Nowak,[1] Tiffanie Regnier,[2] Dipankar Bachar,[3,7] Pierre Plassart,[2] Patrick Wincker,[4] Claudy Jolivet,[5] Antonio Bispo,[6] Philippe Lemanceau,[1] Pierre-Alain Maron,[1,2] Christophe Mougel[1,2] and Lionel Ranjard[1,2]*

[1]*INRA-Université de Bourgogne, UMR Microbiologie du Sol et de l'Environnement, CMSE, 17, rue Sully, B.V. 86510, 21065 Dijon Cedex, France.*

[2]*INRA-Université de Bourgogne, Plateforme GenoSol, CMSE, 17, rue Sully, B.V. 86510, 21065 Dijon Cedex, France.*

[3]*CNRS UMR 6543, Laboratoire de Biologie Virtuelle, Centre de Biochimie, Parc Valose. F 06108 Nice, France.*

[4]*Commissariat à l'Energie Antomique (CEA), Institut de Génomique (IG), Genoscope, Evry, France.*

[5]*INRA Orléans – US 1106, Unité INFOSOL, Avenue de la Pomme de Pin – BP 20619 Ardon, 45166 Olivet cedex, France.*

[6]*ADEME, Service Agriculture et Forêt – 20, Avenue du Grésillé – BP 90406-49004 Angers Cedex 01, France.*

[7]*Université de Nice, UMR 6543, Laboratoire de Biologie Virtuelle, Centre de Biochimie, Parc Valose. F 06108 Nice, France.*

Summary

Three soil DNA extraction procedures (homemade protocols and commercial kit) varying in their practicability were applied to contrasting soils to evaluate their efficiency in recovering: (i) soil DNA and (ii) bacterial diversity estimated by 16S rDNA pyrosequencing. Significant differences in DNA yield were systematically observed between tested procedures. For certain soils, 10 times more DNA was recovered with one protocol than with the others. About 15 000 sequences of 16S rDNA were obtained for each sample which were clustered to draw rarefaction curves. These curves, as well as the PCA ordination of community composition based on OTU clustering, did not reveal any significant difference between procedures. Nevertheless, significant differences between procedures were highlighted by the taxonomic identification of sequences obtained at the *phylum* to *genus* levels. Depending on the soil, differences in the number of *genera* detected ranged from 1% to 26% between the most and least efficient procedures, mainly due to a poorer capacity to recover populations belonging to *Actinobacteria*, *Firmicutes* or *Crenarchaeota*. This study enabled us to rank the relative efficiencies of protocols for their recovery of soil molecular microbial biomass and bacterial diversity and to help choosing an appropriate soil DNA extraction procedure adapted to novel sequencing technologies.

Received 22 August, 2011; accepted 26 August, 2011. *For correspondence. E-mail ranjard@dijon.inra.fr; Tel. (+33) 3 80 69 30 88; Fax (+33) 3 80 69 32 24.

Introduction

During the last two decades, novel molecular methods have been developed that are well-suited to the characterization of soil microbial communities as they provide access to previously hidden genetic resources (for review see Torsvik and Øvreås, 2002). These methods are based essentially on characterizing soil DNA and significant efforts have been devoted to optimize the soil DNA extraction procedure in order to obtain suitable representative extracts for quantitative and qualitative characterization of the microbial communities (Zhou *et al*., 1996; Kuske *et al*., 1998; Delmont *et al*., 2010). These efforts led to the development of various homemade DNA extraction protocols, as well as commercial kits, which have been used in more than 1000 articles published yearly. Microbial abundance and community fingerprinting have already been assessed in various soils to compare the technical biases of these procedures (Martin-Laurent *et al*., 2001; Ranjard *et al*., 2003). However, the recent development of high-throughput sequencing technology (such as 454 or Illumina) allows the scientific community to assess 'MetaTaxogenomic' studies by getting hundreds of thousands of ribosomal rDNA gene sequences from a single metagenomic DNA (Roesch *et al*., 2007; Will *et al*., 2010; Maron *et al*., 2011). The impressive power of these emerging tools calls for a more thorough examination of

Table 1. Physico-chemical characteristics and land-use of studied soils.

Soils	Clay	Silt	Sand	Corg	C:N	pH	Calcium carbonate	Land use
399	258	394	348	1.8	9.1	8.2	804	Crop system in rotation with grassland
676	318	571	111	0.34	4.6	7.9	30.1	Crop system in rotation without grassland
2034	477	449	74	10.8	15.3	7.9	39.1	Deciduous forest
1850	174	128	698	1.7	14.4	7.8	2.2	Deciduous forest
1868	38	32	930	1.8	27.7	4	0	Coniferous forest
1901	707	275	18	4.9	10.7	5.4	0	Grassland
2116	322	405	273	0.6	7.2	8.3	134	Vineyards

Values are given in g kg^{-1} soil (dw), except for C:N and pH.

the relative efficiencies of different DNA extraction techniques to recover microbial diversity and discriminate soil bacterial composition as well as to detect the rare species.

Delmont and colleagues (2010) recently recommended the combination of several soil sampling and DNA extraction strategies to access the whole metagenome in terms of species richness. However, such a strategy is time- and cost-consuming and inappropriate studying a large number of samples in wide extent studies. In this context, we aimed at comparing classical but different soil DNA extraction procedures for their recovery of soil DNA and ability to estimate bacterial diversity in seven different soils to deduce the most efficient and useful procedures for soil 'MetaTaxogenomic' studies. The seven soils were originated from the soil library of the French Soil Quality Monitoring Network (Dequiedt *et al*., 2009; 2011) and were chosen for their contrasting physico-chemical and land-use characteristics (Table 1), known to strongly influence microbial abundance and diversity (Ranjard *et al*., 2010). Three DNA extraction procedures were chosen for their recurrent use in the literature (Ultraclean soil DNA kit, MOBIO Laboratories) and for their strong optimization for soil bacterial diversity retrieval (GnS-GII and Sy3). The GnS-GII procedure was developed by the platform GenoSol to extract soil DNA in large-scale soil surveys (Ranjard *et al*., 2009). It is based on a combination of mechanical (FastPrep® bead-beating, strongly optimized by V. Nowak and M. Lelievre, pers. comm.) and chemical (SDS) lyses of indigenous cells. The Sy3 procedure has been optimized for low bacterial density samples and gives a good DNA recovery of different bacterial *phyla* (like α-, β-, γ-*Proteobacteria*, *Firmicutes* and *Actinobacteria*) (Maron *et al*., 2005; 2007). It is based on a combination of mechanical (bead-beating with Mikrodismembrator, BBraun Biotech), chemical (SDS), thermal (freeze-thaw cycle) and enzymatic (proteinase K) lyses of indigenous cells. The Ultraclean soil DNA kit (MOBIO) is a commercial kit based on mechanical (bead-beating) and chemical lyses (see '*Experimental procedures*' in the Supporting Information file for a precise description of each procedure). For each protocol, DNA was extracted from one gram of soil (dry weight) according to the prerequisite for soil sampling size (Ranjard *et al*., 2003). Regarding the detailed procedure of the lyses step, an order of practicability (including cost, fastness and technical levels) could be deduced as follows: MOBIO > GnS-GII > Sy3.

Results and discussion

The first step in the evaluation of the various DNA extraction procedures was based on the comparison of soil DNA yields which were quantified as previously described (Bernard *et al*., 2007). A significantly ($P < 0.05$) higher amount of soil DNA (in some cases, almost 10 times more) was recovered from all soils with GnS-GII, as compared with the Sy3 and MOBIO procedures (Table 2). It is important to note that the improvement of DNA yield is not combined with an alteration of size integrity (as visually observed on agarose gel, data not shown). Based on our results the ranking of the protocol according to their efficiency for soil DNA recovery was: GnS-GII >> Sy3 > MOBIO. This suggested that the most stringent cell lysis procedure in Sy3 (including mechanical, chemical, thermal and enzymatic steps) did not provide the highest DNA yield, maybe due to its inadequacy in the soil matrix and/or a possible alteration of DNA integrity. On the other hand, the highest efficiency of the GnS-GII procedure might be explained by the mechanical lysis step, which was strongly optimized in terms of type and size of the glass-beads as well as in terms of the strength and duration of grinding using FastPrep® -24 (V. Nowak and M. Lelièvre, pers. comm., see '*Experimental procedures*' in the Supporting Information file). FastPrep® -24 grinding was carried out in three dimensions and improved cell lysis efficiency. Indeed, a significant increase in DNA recovery was obtained (data not shown) when the bead-beating procedures in the MOBIO and Sy3 protocols were replaced by those of GnS-GII, thereby confirming the importance of this step in the DNA extraction procedure.

Significant variations in DNA recovery between soils were observed with the GnS-GII protocol. The greatest amount of soil DNA (8.21 ± 1.21 μg DNA g^{-1} soil) was

Table 2. Numbers of filtered sequences produced by pyrosequencing the V4 region of the 16S rDNA gene; richness was estimated with numbers of taxonomic groups found at different taxonomic levels within soil samples, according to the DNA extraction procedure used; Evenness index was determined with taxonomy assigned at the *genus* level. Soil DNA recovery for each sample is also indicated.

Sample	Soil DNA recovery μg g^{-1} soil (dry weight) (± standard deviation)	Sequences after equalization and filtering steps	Richness: number of taxonomic groups found (Evenness index)				
			Phyla	*Class*	*Order*	*Family*	*Genus*
399 GnS-GII	*5.61* (± 0.98)	13 420	**29**	57	126	195	**292** (0.99)
399 MoBio	0.43 (± 0.04)	13 750	24	56	122	183	275 (0.98)
399 Sy3	1.77 (± 0.43)	13 268	**29**	**60**	**129**	**196**	285 (1.00)
676 GnS-GII	*5.1* (± 1.16)	13 428	29	58	129	196	284 (0.93)
676 MoBIO	1.01 (± 0.28)	13 772	**33**	63	128	188	278 (0.97)
676 Sy3	1.11 (± 0.48)	13 475	29	**65**	**138**	**211**	**302** (0.98)
2034 GnS-GII	*8.21* (± 1.21)	13 415	24	49	**114**	**170**	**271** (0.97)
2034 MoBIO	0.85 (± 0.12)	13 232	22	49	105	162	252 (0.95)
2034 Sy3	1.42 (± 0.59)	13 299	**27**	**52**	110	168	258 (0.89)
1850 GnS-GII	*4.36* (± 0.45)	13 476	24	**53**	**120**	**177**	**269** (0.95)
1850 MoBIO	0.44 (± 0.07)	13 740	21	55	115	167	243 (0.89)
1850 Sy3	1.35 (± 0.21)	7 806	**25**	51	101	153	212 (0.90)
1868 GnS-GII	*1.15* (± 0.11)	14 391	**20**	31	52	83	**127** (0.70)
1868 MoBIO	0.26 (± 0.05)	14 472	18	28	48	68	106 (0.72)
1868 Sy3	0.61 (± 0.08)	10 347	**20**	**34**	**57**	**86**	125 (0.76)
1901 GnS-GII	*1.66* (± 0.13)	14 217	19	43	87	122	174 (0.74)
1901 MoBIO	0.35 (± 0.15)	14 265	15	33	66	94	129 (0.50)
1901 Sy3	0.99 (± 0.24)	14 149	**21**	**48**	**91**	**130**	**182** (0.70)
2116 GnS-GII	*1.20* (± 0.39)	13 596	24	54	122	187	**277** (0.96)
2116 MoBIO	0.05 (± 0.001)	0	NA	NA	NA	NA	NA
2116 Sy3	0.58 (± 0.03)	13 805	**27**	**59**	**123**	**189**	275 (0.96)

Numbers in italic indicate the significant differences between soil DNA recovery and soils for a given procedure. Bold numbers indicated the samples with the highest numbers of detected taxonomic groups for each studied soil between protocols for each taxonomic level.
NA, not applicable.

obtained from an alkaline fine-textured soil with high organic carbon content and under deciduous forest (soil 2034), whereas the smallest amounts were detected in an acid sandy soil type with high C:N ratio and under coniferous forest (1.15 ± 0.11 μg DNA g^{-1} soil, soil 1868), and in an alkaline silty-clay soil under vineyards (1.20 ± 0.39 μg DNA g^{-1} soil, soil 2116). As already demonstrated in a large-scale study, soil DNA recovery is positively correlated with soil characteristics such as texture (silt content $r = 0.56$, $P < 0.011$), organic carbon content ($r = 0.549$, $P < 0.010$) and pH ($r = 0.564$, $P < 0.008$) (Marstorp *et al.*, 2000; Dequiedt *et al.*, 2011). In addition, our results confirm the significant impact of land use, the amount of soil DNA being much lower under coniferous forest and vineyards, independently of soil characteristics (Dequiedt *et al.*, 2011). Similar trends between soils were also observed for the MOBIO and Sy3 protocols but the differences were less significant. Consequently, in a context where the amount of soil DNA might be used as a robust indicator of soil microbial biomass (Marstorp *et al.*, 2000; Dequiedt *et al.*, 2011), our study underlined the need to choose a reliable DNA extraction procedure suitable for sensitive and early detection of quantitative changes due to soil characteristics and management.

Soil bacterial diversity was assessed by pyrosequencing the V4 region of 16S rDNA using the genome sequencer FLX system, with long-read GS FLX Titanium chemistry (Roche Applied Science, Indianapolis, Indiana) at the Genoscope center (Evry, France) (see '*Experimental procedures*' in the Supporting Information). The raw sequence libraries were first equalized and filtered to remove reads originating from sequencing errors or putative chimeric sequences (see '*Experimental procedures*' in the Supporting Information file for detailed bioinformatic procedure). As a result, the number of high quality sequences ranged from 13 232 to 14 472 (average read length: 256 bp), except for three soils which gave 10 000 or fewer raw sequences, or no PCR amplification (Table 2). The final step included a clustering of rare reads to abundant reads, with k = six differences, but not counting differences in homopolymer lengths, which are a major source of errors with the 454 technology. Rarefaction curves were drawn with the obtained clusters (at $k = 6$ differences, Fig. 1) and showed that: (i) saturation was reached for all soils and for each DNA extraction procedure, thereby indicating that about 15 000 sequences were sufficient to obtain a representative covering of soil bacterial diversity; and (ii) significant differences between the DNA extraction procedures were recorded in the slope and level of the plateau of the curve for three soils (1850, 1901 and 1868, Fig. 1). However, it was difficult to deduce a general ranking of the procedures according to their efficiency to recover OTUs. Indeed, GnS-GII and MOBIO were more efficient for soil 1850, Sy3 and GnS-GII for soil 1901 and Sy3 for soil 1868.

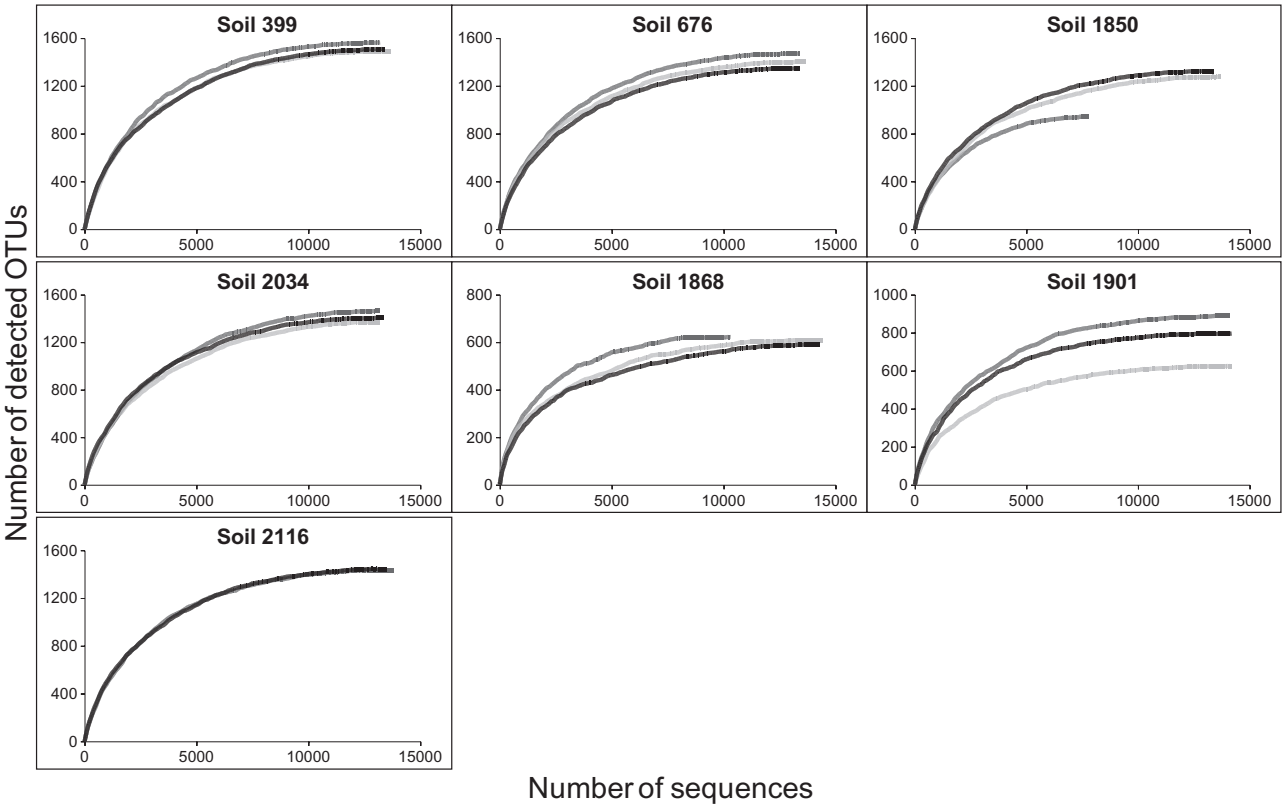

Fig. 1. Rarefaction curves determined by pyrosequencing of V4 region of the 16S rDNA gene obtained for each soil with different soil DNA extraction procedures, based on randomly selected and filtered sequences. Rarefaction curves were determined using clustering at $k = 6$ differences between sequences (see Supporting Information for more information). Black line: GnS-GII; light grey line: MOBIO; dark grey line: Sy3.

Bacterial OTU matrices were subjected to principal component analysis (PCA) to obtain an ordination on a factorial map and compare the soil DNA procedures based on OTU composition (Fig. 2). The discrimination between soils was similar irrespective of the soil DNA extraction protocol used, with soil 1868 being discriminated from soil 1901 and from the other soils. This result suggests that although various soil DNA extraction procedures could lead to different soil DNA yields for a given soil, the procedures led to a similar rough discrimination between soils, based on the global composition of the bacterial community. This observation is in accordance with the results of previous studies involving DNA fingerprinting approach which concluded that all procedures were biased and provided a robust global discrimination of bacterial community structure (Martin-Laurent *et al.*, 2001). Although only a limited number of soils were studied, it is interesting to note that the soil bacterial community discrimination revealed a strong influence of soil pH since soils 1868 and 1901 were significantly discriminated and more acidic than the other soils (Table 1). Therefore, our results corroborate the high structuring effect of soil pH on soil bacterial community composition (Fierer and Jackson, 2006; Griffiths *et al.*, 2011).

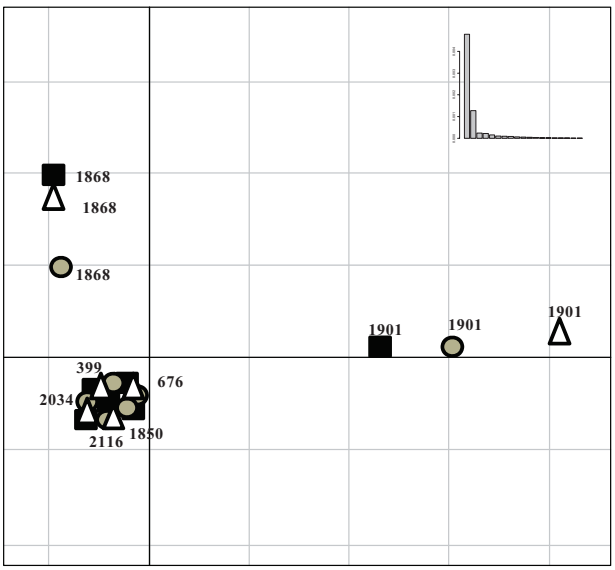

Fig. 2. Principal component (PC1xPC2) plots generated from the clustering of 16S rDNA OTU sequences at $k = 6$ obtained from the different soils using the various soil DNA extraction procedures (black square: GnS-GII, grey circle: Sy3, white triangle: MOBIO).

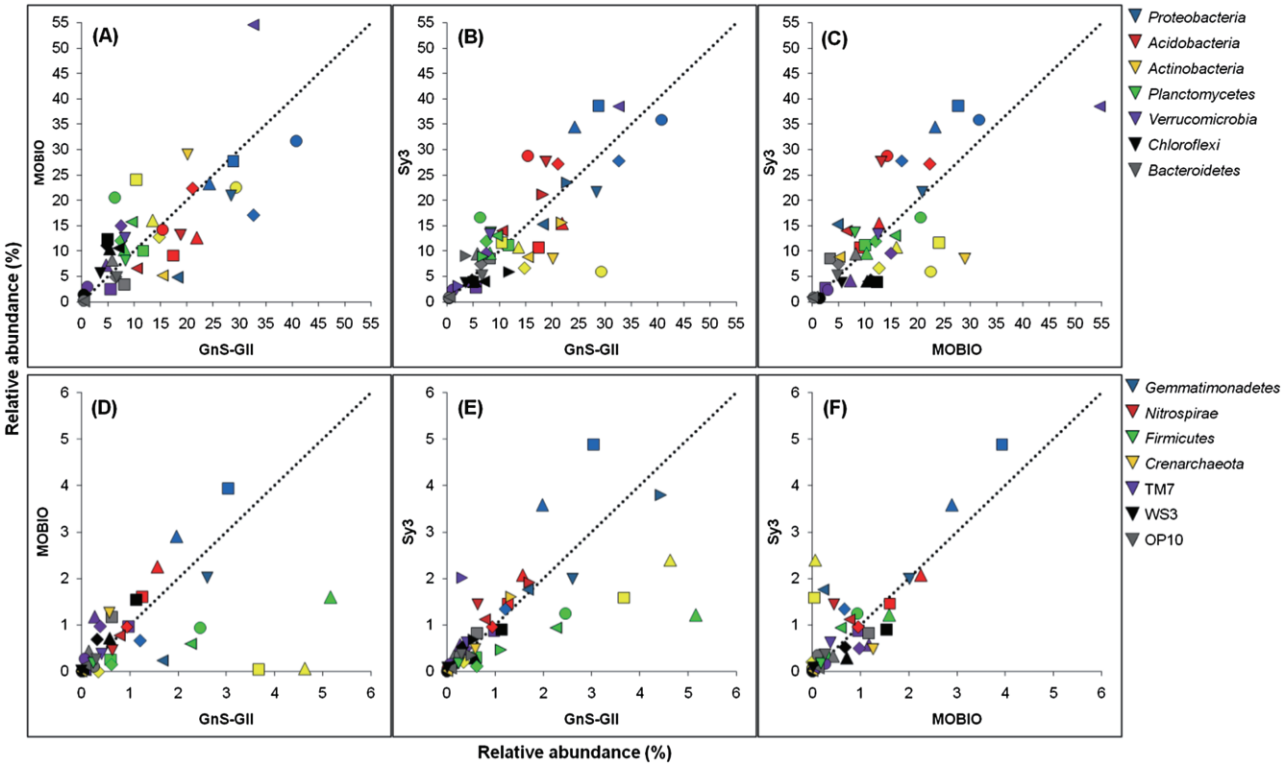

Fig. 3. Scatter plot comparison of relative abundance of each detected phylogenetic group in the studied soils (399: ◆, 676:●, 2034:▼, 1850:▲, 1868: ■, 1901: ◄ and 2116: ►) with each procedure (GnS-GII, MOBIO and Sy3). Only the 14 most abundant *phyla* are plotted for the comparison. (A), (B) and (C) scatter plots show the dominant *phyla* detected; and (D), (E) and (F) scatter plots the minor *phyla*. (A) and (D): GnS-GII versus MOBIO; (B) and (E): GnS-GII versus Sy3; (C) and (F): MOBIO versus Sy3. Dotted lines are unity lines.

The obtained reads were taxonomically identified at *phyla, class, order, family* and *genus* levels using a previously published procedure (Stoeck *et al*., 2010; Behnke *et al*., 2011), that had been slightly modified to incorporate the Silva 104 reference sequences (see Supporting Information for the detailed bioinformatic pipeline). Interestingly, application of the described method allowed the affiliation of as much as about 98% of the obtained sequences to known bacterial *phyla* (Fig. S1). Furthermore, significant differences in bacterial richness were observed between DNA extraction procedures for all soils (Table 2). At the *phylum* level, the GnS-GII and Sy3 procedures always detected a higher number of taxa than MOBIO, excepted for soil 676. At the *genus* level, the GnS-GII protocol was the most efficient in recovering the highest number of taxa for all soils, except for two soils (676 and 1901) where Sy3 and GnS-GII gave similar results. Differences between the most and the least efficient procedures for recovering richness were soil dependent and ranged from 1% (soil 2116) to more than 26% (soil 1901). Thus, the deduced ranking of bacterial richness recovery efficiency was GnS-GII > Sy3 > MOBIO that might partly explain the higher DNA recovery obtained with the GnS-GII procedure. On the other hand, when evenness was calculated at the *genus* level, the results implied that the variations between procedures were less significant than for richness (Table 2). This could suggest that the global pattern of population frequency was only slightly influenced by the DNA extraction procedure as discussed above for PCA analysis (Fig. S1).

Pairwise comparison of the relative abundances of identified taxa between procedures suggested that the higher bacterial richness obtained with GnS-GII could mainly be explained by its greater efficiency to recover particular *phyla*. Figure 3A–C clearly indicated that MOBIO underestimated the dominant *Proteobacteria* and *Acidobacteria* groups and that Sy3 underestimated *Actinobacteria*. Furthermore, the GnS-GII procedure was better to detect some *phyla*, such as *Firmicutes* and *Crenarchaeota* (*Archaeabacteria*) (Fig. 3D–F). This observation was highly significant for several classes of *Firmicutes* (*Bacilli* and *Clostridia*) (Fig. S1 and Table S1). Both *Crenarchaeota* and *Firmicutes* are known to be recalcitrant to detection due to their particular habitat, physiology or their spore forming ability (Fierer *et al*., 2007; Bates *et al*., 2011). Altogether, these data confirm that the lysis stringency developed for the GnS-GII protocol led to a better detection of these taxa and allowed the access to a more genetically diverse proportion of the soil metagenome. From an ecological point of view *Crenarchaeota, Firmicutes* and

Acidobacteria are known to be abundant and ubiquitous in soil environments and to be of major interest in soil functioning (Jones *et al*., 2009; Bates *et al*., 2011). Therefore, the ranking of procedures (GnS-GII > Sy3 > MOBIO) deduced from the efficiency of bacterial richness recovery could facilitate the choice of an adequate procedure which, in turn, might have ecological relevance in a context where the major challenge in soil microbial ecology is to better understand community assembly and to link microbial diversity to ecosystem goods and services (Maron *et al*., 2011).

Significant differences were also observed when a given procedure was used to compare bacterial diversity between different soils (Table 2). For example, the acidic soils 1868 and 1901 harboured the lowest bacterial richness, thereby corroborating the positive correlation between pH and bacterial richness previously depicted by Fierer and Jackson (2006). In addition, these soils exhibited a very different pattern of detected phylogenetic groups (Fig. S1 and Table S1) with several particular dominant *phyla* such as *Verrucomicrobia* for soil 1901. This *phylum* is known to be one of the most common and diverse *phyla* in soil habitats (Kielak *et al*., 2010) and is commonly found in acidic grassland soils (Kuramae *et al*., 2010). Furthermore, the lowest bacterial diversity observed in soil 1868 might also result from the coarse texture, low organic carbon content, high C:N ratio and a coniferous land cover (Table 1). All these parameters are known to be deleterious for soil microbial diversity (Lejon *et al*., 2005; Dequiedt *et al*., 2009).

Altogether, these results indicate that our biological vision of soil microbial abundance and biodiversity might be limited by the procedures used to recover soil metagenomic DNA. This statement has become critical with the recent expansion of 16S rDNA pyrosequencing technologies, which allow a more profound investigation of microbial diversity. Technically, our study confirms the pivotal importance of cell lysis in the soil DNA extraction procedure. It also emphasizes the need for incessant technical surveys to increase species richness per sequencing effort, without increasing the cost and practicality of the procedure, to be useful for wide extent studies. Consequently, we need to revisit our choice of extraction protocol to ensure that the soil DNA recovered is not only of good quality but also sufficiently representative in terms of richness and evenness of the bacterial populations. However, fungal communities constitute a significant portion of soil biodiversity and are also a key component in soil processes. In order to draw meaningful conclusions about the representativeness of the extracted soil DNA, our study could therefore be completed by an evaluation of fungal diversity recovery. This statement is of particular relevance in the current context where soil microbial ecology studies are carried out to better understand community assembly in space and time and its link with soil functioning and more widely with ecosystem services.

Acknowledgements

This study was granted by ADEME (French Environment and Energy Management Agency), the Regional Council of Burgundy and by the Genoscope (Evry, France). RMQS soil sampling and physico-chemical analyses were supported by a French Scientific Group of Interest on soils: the 'GIS Sol', involving the French Ministry for Ecology and Sustainable Development (MEDAD), the French Ministry of Agriculture (MAP), the French Institute for Environment (IFEN), the Environment and Energy Management Agency (ADEME), the French Institute for Research and Development (IRD) and the National Institute for Agronomic Research (INRA). We thank all the soil surveyors and technical assistants involved in sampling the sites.

References

Bates, S.T., Berg-Lyons, D., Caporaso, J.G., Walters, W.A., Knight, R., and Fierer, N. (2011) Examining the global distribution of dominant archaeal populations in soil. *ISME J* **5:** 908–917.

Behnke, A., Engel, M., Christen, R., Nebel, M., Klein, R.R., and Stoeck, T. (2011) Depicting more accurate pictures of protistan community complexity using pyrosequencing of hypervariable SSU rRNA gene regions. *Environ Microbiol* **13:** 340–349.

Bernard, L., Mougel, C., Maron, P.A., Nowak, V., Lévêque, J., Henault, C., *et al*. (2007) Dynamics and identification of soil microbial populations actively assimilating carbon from 13C-labelled wheat residue as estimated by DNA- and RNA-SIP techniques. *Environ Microbiol* **9:** 752–764.

Delmont, T.O., Robe, P., Cecillon, S., Clark, I.M., Constancias, F., Simonet, P., *et al*. (2010) Accessing the soil metagenome for studies of microbial diversity. *Appl Environ Microbiol* **77:** 1315–1324.

Dequiedt, S., Thioulouse, J., Jolivet, C., Saby, N.P.A., Lelievre, M., Maron, P.-A., *et al*. (2009) Biogeographical patterns of soil bacterial communities. *Environ Microbiol Rep* **1:** 251–255.

Dequiedt, S., Saby, N.P.A., Lelievre, M., Jolivet, C., Thioulouse, J., Toutain, B., *et al*. (2011) Biogeographical patterns of soil molecular microbial biomass as influenced by soil characteristics and management. *Glob Ecol Biogeogr* **20:** 641–652.

Fierer, N., and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* **103:** 626–631.

Fierer, N., Morse, J.L., Berthrong, S.T., Bernhardt, E.S., and Jackson, R.B. (2007) Environmental controls on the landscape-scale biogeography of stream bacterial communities. *Ecology* **88:** 2162–2173.

Griffiths, R.I., Thompson, B.C., James, P., Bell, T., Bailey, M., and Whitheley, A.S. (2011) The bacterial biogeography of British soils. *Environ Microbiol* **13:** 1642–1654.

Jones, R.T., Robeson, M.S., Lauber, C.L., Hamady, M., Knight, R., and Fierer, N. (2009) A comprehensive survey

of soil acidobacterial diversity using pyrosequencing and clone library analyses. *ISME J* **3:** 442–453.
Kielak, A., Rodrigues, J.L.M., Kuramae, E.E., Chain, P.S.G., Van Veen, J.A., and Kowalchuk, G.A. (2010) Phylogenetic and metagenomic analysis of *Verrucomicrobia* in former agricultural grassland soil. *FEMS Microbiol Ecol* **71:** 23–33.
Kuramae, E.E., Gamper, H.A., Yergeau, E., Piceno, Y.M., Brodie, E.L., DeSantis, T.Z., *et al*. (2010) Microbial secondary succession in a chronosequence of chalk grasslands. *ISME J* **4:** 711–715.
Kuske, C.R., Banton, K.L., Adorada, D.L., Stark, P.C., Hill, K.K., and Jackson, P.J. (1998) Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. *Appl Environ Microbiol* **64:** 2463–2472.
Lejon, D.P.H., Chaussod, R., Ranger, J., and Ranjard, L. (2005) Microbial community structure and density under different tree species in an acid forest soil (Morvan, France). *Microb Ecol* **50:** 614–625.
Maron, P.A., Lejon, D.P.H., Carvalho, E., Bizet, K., Lemanceau, P., Ranjard, L., and Mougel, C. (2005) Assessing genetic structure and diversity of airborne bacterial communities by DNA fingerprinting and 16S rDNA clone library. *Atmos Environ* **39:** 3687–3695.
Maron, P.A., Mougel, C., Siblot, S., Abbas, H., Lemanceau, P., and Ranjard, L. (2007) Protein extraction and fingerprinting optimization of bacterial communities in natural environment. *Microb Ecol* **53:** 426–434.
Maron, P.A., Mougel, C., and Ranjard, L. (2011) Soil microbial diversity: methodological strategy, spatial overview and functional interest. *C R Biol* **334:** 403–411.
Marstorp, H., Guan, X., and Gong, P. (2000) Relationship between dsDNA, chloroform labile C and ergosterol in soils of different organic matter contents and pH. *Soil Biol Biochem* **32:** 879–882.
Martin-Laurent, F., Philippot, L., Hallet, S., Chaussod, R., Germon, J.C., Soulas, G., and Catroux, G. (2001) DNA extraction from soils: old bias for new microbial diversity analysis methods. *Appl Environ Microbiol* **67:** 2354–2359.
Ranjard, L., Lejon, D.P.H., Mougel, C., Schehrer, L., Merdinoglu, D., and Chaussod, R. (2003) Sampling strategy in molecular microbial ecology: influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. *Environ Microbiol* **5:** 1111–1120.
Ranjard, L., Dequiedt, S., Lelievre, M., Maron, P.A., Mougel, C., Morin, F., and Lemanceau, P. (2009) Platform GenoSol: a new tool for conserving and exploring soil microbial diversity. *Environ Microbiol Rep* **1:** 97–99.
Ranjard, L., Dequiedt, S., Jolivet, C., Saby, N.P.A., Thioulouse, J., Harmand, J., *et al*. (2010) Biogeography of soil microbial communities: a review and a description of the ongoing French national initiative. *Agron Sustain Dev* **30:** 359–365.
Roesch, L.F.W., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K.M., Kent, A.D., *et al*. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* **1:** 283–290.
Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.W., and Richards, T.A. (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **19:** 21–31.
Torsvik, V., and Øvreås, L. (2002) Microbial diversity and function in soil: from genes to ecosystems. *Curr Opin Microbiol* **5:** 240–245.
Will, C., Thurmer, A., Wollherr, A., Nacke, H., Herold, N., Schrumpf, M., *et al*. (2010) Horizon-specific bacterial community composition of German Grassland Soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. *Appl Environ Microbiol* **76:** 6751–6759.
Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Appl Environ Microbiol* **62:** 316–322.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Relative abundance of phylogenetic groups found in each soil with different soil DNA extraction procedures. Procedures are indicated below the graph (G: GnS-GII, M: MOBIO and S: Sy3). Shown are the percentages of the classified sequences. Phylogenetic groups accounting for $\leq$ 0.5% of each sample are summarized in the artificial group 'Others'.
Table S1. Relative abundance of phylogenetic groups found in each soil with different soil DNA extraction procedures.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.