Assessing soil bacterial community and dynamics by integrated high-throughput absolute abundance quantification

Jun Lou [1] , **Li Yang** [1] , **Haizhen Wang** Corresp., [1] , **Laosheng Wu** [1, 2] , **Jianming Xu** Corresp. [1]

[1] Institute of Soil and Water Resources and Environmental Science, College of Environmental and Resource Sciences, Zhejiang Provincial Key Laboratory of Agricultural Resources and Environment, Zhejiang University, Hangzhou, China

[2] Department of Environmental Sciences, University of California, Riverside, Riverside, United States

Corresponding Authors: Haizhen Wang, Jianming Xu
Email address: wanghz@zju.edu.cn, jmxu@zju.edu.cn

Microbial ecological studies have been remarkably promoted by the high-throughput sequencing approach with explosive information of taxonomy and relative abundance. However, relative abundance does not reflect the quantity of the microbial community and the inter-sample differences among taxa. In this study, we refined and applied an integrated high-throughput absolute abundance quantification (iHAAQ) method to better characterize soil quantitative bacterial community through combining the relative abundance (by high-throughput sequencing) and total bacterial quantities (by quantitative PCR). The proposed iHAAQ method was validated by an internal reference strain EDL933 and a laboratory strain WG5. Application of the iHAAQ method to a soil phenanthrene biodegradation study showed that for some bacterial taxa, the changes of relative and absolute abundances were coincident, while for others the changes were opposite. With the addition of a microbial activity inhibitor (NaN$_3$), the absolute abundances of soil bacterial taxa, including several dominant genera of *Bacillus*, *Flavobacterium*, and *Paenibacillus*, decreased significantly, but their relative abundances increased after 28 days of incubation. We conclude that the iHAAQ method can offer more comprehensive information to reflect the dynamics of soil bacterial community with both relative and absolute abundances than the relative abundance from high-throughput sequencing alone.

1   **Assessing soil bacterial community and dynamics by integrated**

2   **high-throughput absolute abundance quantification**

3   Jun Lou[1, +], Li Yang[1, +], Haizhen Wang[1, *], Laosheng Wu[1, 2] and Jianming Xu[1, *]

5   [1] Institute of Soil and Water Resources and Environmental Science, College of

6   Environmental and Resource Sciences, Zhejiang Provincial Key Laboratory of

7   Agricultural Resources and Environment, Zhejiang University, Hangzhou, China

8   [2] Department of Environmental Sciences, University of California, Riverside, CA,

9   USA

11   [+]These authors contributed equally to this work.

12   [*]Correspondence:

13   Haizhen Wang and Jianming Xu,

15   Email address: wanghz@zju.edu.cn and jmxu@zju.edu.cn.

17 **ABSTRACT**

18 Microbial ecological studies have been remarkably promoted by the high-throughput

19 sequencing approach with explosive information of taxonomy and relative abundance.

20 However, relative abundance does not reflect the quantity of the microbial community

21 and the inter-sample differences among taxa. In this study, we refined and applied an

22 integrated high-throughput absolute abundance quantification (iHAAQ) method to

23 better characterize soil quantitative bacterial community through combining the

24 relative abundance (by high-throughput sequencing) and total bacterial quantities (by

25 quantitative PCR). The proposed iHAAQ method was validated by an internal

26 reference strain EDL933 and a laboratory strain WG5. Application of the iHAAQ

27 method to a soil phenanthrene biodegradation study showed that for some bacterial

28 taxa, the changes of relative and absolute abundances were coincident, while for

29 others the changes were opposite. With the addition of a microbial activity inhibitor

30 (NaN$_3$), the absolute abundances of soil bacterial taxa, including several dominant

31 genera of *Bacillus*, *Flavobacterium*, and *Paenibacillus*, decreased significantly, but

32 their relative abundances increased after 28 days of incubation. We conclude that the

33 iHAAQ method can offer more comprehensive information to reflect the dynamics of

34 soil bacterial community with both relative and absolute abundances than the relative

35 abundance from high-throughput sequencing alone.

36 **Introduction**

37 Bacteria are the essential components of soil ecosystem and play a vital role in the

38 material cycle and energy flow in the Earth ecosystem (Whitman et al. 1998). It was

39 estimated that there are $4 - 6 \times 10^{30}$ cells of prokaryotes on Earth and 2.6×10^{30} in soil,

40 which is the most diverse and abundant cellular life form on Earth (Whitman et al.

41 1998; Curtis & Sloan 2005). Soil bacteria have been studied intensively for more than

42 a hundred years from individuals and its communities, but there are still more than

43 85-99% of bacteria that cannot be isolated and grown in culture mediums (Lok 2015).

44 Therefore, the question remains how many bacteria are there (absolute and relative

45 abundances) and who are they (kinds of species) in a soil, which consists a

46 community.

47 Over the past few decades, scientists have attempted to answer these questions.

48 The classical plate-counting method via counting the different colonies on plate was

49 the simplest way to simultaneously obtain the information of absolute and relative

50 abundances of soil bacteria, but the method has the limitation because of the low

51 percentange of cultivable bacteria in soil (Davey 2011; Blagodatskaya & Kuzyakov

52 2013). Many other methods have been applied to explore the absolute or relative

53 abundances of soil bacterial community. For example, Brookes et al. (1985) and

54 Vance et al. (1987) tried to estimate the total microbial abundance by measuring the

55 microbial biomass-N (MBN) and -C (MBC) in soil. With the development of

56 molecular biology and the more verified specific target genes, the quantitative PCR

57   (qPCR) also becomes a powerful and accurate technique to measure the total or

58   specific microbial absolute abundance in soil (Philippot et al. 2009).

59   Various biomarker and molecule methods were also found to be useful for

60   exploring the community sturctures of soil bacteria. These methods include the

61   phospholipid fatty acids (PLFAs), polymerase chain reaction-denaturing gradient gel

62   electrophoresis (PCR-DGGE), clone library, catalysed reporter deposition

63   fluorescence in situ hybridization (CARD-FISH), microfluidic qPCR, etc. (Li et al.

64   2006; Ding et al. 2011; Cai et al. 2016; Kleyer et al. 2017). Especially, establishment

65   of the high-throughput sequencing method opened a new era for microbiology.

66   Confirmed by the microbial mock community and multiple studies, this method can

67   reveal the relative abundance as well as kinds of species of soil bacterial community

68   with unprecedented amount of information (Wang et al. 2007; Caporaso et al. 2012;

69   Cai et al. 2016; Singer et al. 2016; Tourlousse et al. 2017).

70   Since the high-throughput sequencing technique can acquire rich information of

71   relative abundance (Caporaso et al. 2012; Kozich et al. 2013; Dannemiller et al. 2014;

72   Prest et al. 2014; Stokell et al. 2016; Props et al. 2017; Zhang et al. 2017), it was

73   combined with the absolute quantification methods of total microorganism to

74   characterize the more informative changes of microbial communities in several

75   studies (Dannemiller et al. 2014; Prest et al. 2014; Props et al. 2017; Zhang et al.

76   2017). In the combined methods, the flow cytometric (FCM), adenosine triphosphate

77   (ATP), heterotrophic plate count (HPC), qPCR, PLFAs, and MBC, were used to

78   quantify the absolute abundance of total microorganism. For example, Prest et al.

79   (2014) combined the FCM and 16S rRNA gene pyrosequencing and successfully

80   revealed the change of bacterial community in drinking water, which was

81   undetectable by the ATP and HPC measurements. Dannemiller et al. (2014)

82   quantitatively compared the fungal aerosol populations by combining the qPCR and

83   high-throughput sequencing methods, which were verified by the four selected

84   specific taxa, *Alternaria alternata*, *Cladosporium cladosporioides*, *Epicoccum nigrum*,

85   and *Penicillium/Aspergillus* spp.

86   Recently, Zhang et al. (2017) investigated soil bacterial communities at two

87   different locations by coupling the results of ATP, qPCR, PLFAs, and the relative

88   abundances of high-throughput sequencing technique. Nevertheless, the methods were

89   not verified using specific strains or internal strains as it was done in the previous

90   study by Dannemiller et al. (2014). So far, no research has proved the feasibility of

91   combining the absolute quantification methods of total microorganism and

92   high-throughput sequencing data to characterize soil bacterial communities, which is

93   much more complex than in other environments (Whitman et al. 1998; Smets et al.

94   2016; Tourlousse et al. 2017). As mentioned above, the qPCR method is an accurate

95   method for the absolute abundance of genes, while the high-throughput sequencing is

96   the primary method for the relative abundance of bacteria community. Therefore, we

97   proposed an integrated high-throughput absolute abundance quantification (iHAAQ)

98   method for soil quantitative bacterial ecology through combining the qPCR and

99   high-throughput sequencing techniques in this study.

100   Meanwhile, the accuracy of the iHAAQ method was tested by an internal

101   reference strain, *Escherichia coli* O157:H7 strain EDL933, which can be easily

102   classified and distinguished from soil indigenous bacteria. Also, the internal reference

103   strain can be absolutely quantified by the qPCR method because it contains the

104   flagellar biosynthesis protein FliC gene (*fliC*) (Latif et al. 2014). Furthermore, the

105   iHAAQ method was applied to characterize the bacterial community changes in soil

106   phenanthrene biodegradation.

107

108   **Materials and methods**

109   **Bacteria strains, culture medium and chemicals**

110   The *Escherichia coli* O157:H7 strain EDL933 (ATCC 43895), SMAC (Sorbitol

111   MacConkey)-BCIG (5-bromo-4-chloro-3-indoxyl-β-D-glucuronide) agar (LabM,

112   Lancashire, UK) and Luria-Bertani (LB) medium (Becton, Dickinson and Company,

113   Spark, USA) used herein were described previously (Wang et al. 2014). After

114   incubating at 37 °C on a water orbital shaker (250 rpm), the cultured strain EDL933

115   was harvested by centrifugation at 4 °C (6,000 × g for 10 min) and washed three times

116   with sterile deionized water (Wang et al. 2014). The cell pellets were re-suspended in

117   sterile deionized water to achieve the cell concentration of near 2.1×10^{11} CFU mL^{-1}

118   (CFU, colony-forming units). Diluted with sterile deionized water, the gradient

119   concentrations of strain EDL933 in the order of 2.1×10^{11} to 2.1×10^{6} CFU mL^{-1}

120   were prepared and added to soil samples.

121   The *Massilia* sp. strain WG5 (CCTCC AB 2015362) was isolated from a PAHs

122   contaminated   soil   from   Jiangsu   Province,   China,   which   has   a   high

123   phenanthrene-degrading ability (Lou et al. 2016; Wang et al. 2016). After pre-cultured

124   in LB medium, the strain WG5 was harvested by centrifugation at 4 °C (6,000 × g for

125   10 min) and washed three times with sterile deionized water. The cell pellets were

126   then re-suspended to D_{600nm} 1.0 (approximately 1.5×10^{8} CFU mL^{-1}) for later use.

127   Phenanthrene (PHE; Sigma-Aldrich, Shanghai, China), rifampicin (Fisher

128   Scientific, Fair Lawn, USA), nalidixic acid (Sigma-Aldrich, Saint Louis, USA) and

129   other inorganic chemicals used in this study were of the HPLC or analytical grade.

130   The extraction and detection of PHE in soil were followed by the method of Gu et al.

131   (2017).

132   **Experiment set-up and DNA extractions**

133   The soil was collected from the surface horizon (0-20 cm) near a steel mill in Fujian

134   Province, China. Three composite soil samples were taken to the laboratory using

135   coolers with ice bags. Then, the soil was sieved (2 mm) and stored at 4 °C for use.

136   The soil chemical and physical properties are the same as reported by Gu et al.

137   (2017).

138   The experiment included 10 treatments: control (Treatment Cont, soil only), soil

139   added with strain EDL933 at 6 concentrations (near 10^{9}, 10^{8}, 10^{7}, 10^{6}, 10^{5}, and 10^{4}

140 CFU (g dry wt soil)$^{-1}$, corresponding to Treatments E9, E8, E7, E6, E5 and E4), soil

141 added with PHE (100 μg (g dry wt soil)$^{-1}$, Treatment P), soil added with PHE (100 μg

142 (g dry wt soil)$^{-1}$) and strain WG5 (approximately 1.00×10^{7} CFU (g dry wt soil)$^{-1}$,

143 Treatment W), and soil added with PHE (100 μg (g dry wt soil)$^{-1}$) and sodium azide

144 (NaN_3, 0.1%, w/w, Treatment S). The purpose of adding NaN_3 was to inhibit the

145 native microbial activity (Wei & Pan 2010).

146   In Treatments E9, E8, E7, E6, E5 and E4, 1.0-mL aliquot of strain EDL933 with

147 the gradient concentrations in the order of 2.1×10^{11} to 2.1×10^{6} CFU mL^{-1},

148 respectively, was added to each soil sample (30 g frozen-dry basis). Treatments E9,

149 E8, E7, E6, E5, E4, and Cont were used to develop the iHAAQ method, and

150 Treatments P, W, and S were used to reveal soil bacterial community dynamics

151 during PHE biodegradation.

152   Samples of each treatment were prepared in triplicate by putting 30 g (frozen-dry

153 basis) of soil into a 100-mL flask. All the treated soils were thoroughly mixed (two

154 hours) and soil moisture was adjusted to 100% of the field capacity (soil water content

155 at -33 kPa) by adding sterile deionized water. Ten grams (frozen-dry basis) of soil

156 were sampled immediately from each flask. In addition, all flasks of Treatments P, W,

157 and S were incubated in the dark at 28 ± 1 °C and aerated weekly on the Clean Bench

158 for 30 min. Sterile deionized water was added to compensate the soil moisture loss

159 during incubation and to maintain quasi-constant soil moisture content for the samples.

160   After 28 days of incubation, ten grams (frozen-dry basis) of soil were sampled from

161   each flask of Treatments P, W, and S for DNA extraction.

162   All the soil samples collected from the flasks were immediately frozen and stored

163   in a freezer (-80 °C). Aliquot sample of 0.21 g (frozen-dry basis) per collected sample

164   was used for DNA extraction using the MoBio PowerSoil DNA Isolation kit (MoBio

165   Laboratories, Carlsbad, CA, USA). The extracted DNA was stored in the same freezer

166   (-80 °C) before qPCR analysis.

167   **Quantitative PCR analyses**

168   The qPCR analyses were performed to quantify the specific genes of the extracted

169   DNA with three replicates using a StepOnePlus^{TM} Real-Time PCR System instrument

170   (Applied Biosystems, Foster City, CA, USA). The two pairs of primers, 341F (5'-

171   CCTACGGGAGGCAGCAG -3') and 534R (5'- ATTACCGCGGCTGCTGG -3'),

172   520F   (5'-   AYTGGGYDTAAAGNG   -3')   and   802R   (5'-

173   TACNVGGGTATCTAATCC -3'), were chosen to detect the copy numbers of 16S

174   rRNA gene of total bacteria in V3 and V4 variable regions, respectively (Reddy et al.

175   2012; Guinane et al. 2013). To measure the abundance of strain EDL933, the *fliC*

176   gene was selected from its chromosome, which had only one copy in the genome

177   (Latif et al. 2014). The forward and reverse primers used were 5'-

178   GCGAAGTTAAACACCACGAC -3' and 5'- ACCCGTACCAGCAGTAGATT -3',

179   respectively (Jun et al. 2011).

180   The standard curves were generated using 10-fold serial dilution of a plasmid

181   containing the targeted DNA PCR fragments of 16S rRNA and the *fliC* genes. The 25

182   μL qPCR reactions contained 12.5 μL 2 × SYBR Green qPCR SuperMix-UDG with

183   Rox (Invitrogen, Carlsbad, CA, USA), 0.5 μL of each 10-μM forward and reverse

184   primers, and 10.5 μL sterile and DNA-free water. The known amount and soil DNA

185   samples were added at 1.0 μL per reaction. The reaction was executed under the

186   following thermocycler conditions: 50 °C for 2 min × 1 cycle, 95 °C for 5 min × 1

187   cycle, then 95 °C for 15 s, and 60 °C for 30 s × 40 cycles; and finally followed by the

188   dissociation stage for the melting curve analysis. All the gene copy numbers were

189   calculated from the standard curves of 16S rRNA and the *fliC* genes by using the ΔC_t

190   (cycle threshold) method. All qPCR reactions were conducted in triplicate.

191   **High-throughput sequencing and data processing**

192   According to the method reported by (Ma et al. 2016), each composite DNA sample

193   for sequencing was mixed by combining equal amounts of DNA extraction of the

194   three replicates. The bacterial communities of the Cont, E9 to E4, P, W, and S

195   treatments   were   amplified   with   the   V4   region   of   primers   (520F,   5'-

196   AYTGGGYDTAAAGNG -3' and 802R, 5'- TACNVGGGTATCTAATCC -3'), and

197   then sequenced with Illumina Miseq platform following standard protocols. The

198   quality control of raw sequencing reads was performed using QIIME software

199   (version 1.7.0) (Caporaso et al. 2010). The sequences were removed if the average

200   quality in a moving 5 bps window lower than Q20, length less than 150 bps and

201   having the ambiguous barcode. The paired-end sequences were spliced by the same

202   index with overlap no less than 10 bps and no mismatch using FLASH (Magoc &

203   Salzberg 2011). The chimera sequences were *de novo* identified and removed using

204   UCHIME (Edgar et al. 2011).

205   After quality filtering, the sequences clustering was performed using UCLUST

206   with a 97% similarity and sorted out as operational taxonomic units (OTUs) (Edgar

207   2010). Then, the OTUs were classified using RDP-classifier to match to SILVA

208   database release 119 (Wang et al. 2007; Quast et al. 2013). The raw data obtained in

209   this   research   were   deposited   to   NCBI   SRA   (Sequence   Read   Archive;

210   http://www.ncbi.nlm.nih.gov/sra/)   under   accession   numbers   SRP097773   and

211   SRP105351.

212   **Absolute quantification of soil bacteria and its community by the iHAAQ**

213   **method**

214   The absolute abundance of each phylum or genus in the soil bacterial community was

215   calculated by multiplying total bacterial quantities (the copy number of 16S rRNA

216   gene of total bacteria from qPCR) and the corresponding relative abundance from

217   high-throughput sequencing (Fig. 1). The calculated absolute abundances of

218   *Escherichia-Shigella*, Esch-V3 and Esch-V4, were based on its relative abundance

219   and the total bacterial quantities measured by V3 and V4 regions of 16S rRNA gene,

220   respectively. Since the 16S rRNA gene varied in copy numbers among the different

221   bacteria genomes (Latif et al. 2014; Sun et al. 2013), 7 copies of 16S rRNA gene and

222   1 copy of *fliC* gene in one genome of strain EDL993 (NZ_CP008957, NCBI:

223   http://www.ncbi.nlm.nih.gov) were considered in this study. The absolute abundances

224   of *Escherichia-Shigella* (Esch-V3 and Esch-V4) and $7 \times fliC$ gene copies were

225   logarithmically transformed before performing linear regression analysis.

226   Linear regression was performed using Origin 2016 (OriginLab, Northampton,

227   USA), and analysis of variance (ANOVA) used with SPSS 20 (IBM, Armonk, USA).

228   The homogeneity of variances was tested by Levene's test before ANOVA. Tukey's

229   honestly significant difference test was used when the variances were homogeneous.

230   Otherwise Tamhane's T2 test was used. Heatmaps and cluster analyses were

231   performed using R software version 3.3.1 with pheatmap package (RCore 2013;

232   Kolde 2015).

233   **Results**

234   **Total bacterial quantity and strain EDL933 measurement by qPCR**

235   The background total bacterial quantities (the copy numbers of 16S rRNA gene) were

236   $3.53 \times 10^9 \pm 1.23 \times 10^8$ and $5.82 \times 10^9 \pm 3.29 \times 10^8$ copies (g dry wt soil)$^{-1}$,

237   respectively, for V3 and V4 regions. The total bacterial quantities in Treatments from

238   E9 to E4 ranged from 7.00×10^{10} to 3.53×10^9 and from 7.23×10^{10} to 5.82×10^9

239   copies (g dry wt soil)$^{-1}$ for V3 and V4 regions, respectively. The total bacterial

240   quantities of V3 and V4 regions in Treatments E9 and E8 were significantly ($p < 0.05$)

241   higher than those in other treatments (Fig. S1, Supporting Information). Compared

242   with the background total bacterial quantity ($\sim 10^9$ copies (g dry wt soil)$^{-1}$), the added

243   strain EDL933 in the order of 10^4 to 10^7 copies (g dry wt soil)$^{-1}$ (Table 1) only had a

244   slight influence on the total bacterial quantities in Treatments E4 to E7 (Fig. S1,

245   Supporting Information).

246   A standard curve of the *fliC* gene in the range of 3.60×10^1 to 3.60×10^9 copies

247   μL^{-1} was constructed for qPCR. The C_t value of the standard sample containing $3.60 \times$

248   10^1 copies μL^{-1} was 32.62 ± 0.76, and its standard deviation (SD) was greater than 0.5.

249   However, the C_t value of Cont (33.05 ± 0.98) was even greater than the C_t value of

250   the standard sample containing 3.60×10^1 copies μL^{-1}. This qPCR result indicated

251   that the detected gene in the control soil was not the *fliC* gene (Fig. S2, Supporting

252   Information).

253   In Treatments from E9 to E4, the C_t values (from 13.73 ± 0.05 to 29.66 ± 0.06) of

254   the *fliC* gene were all within the range of the standard curve, and all SD values were

255   lower than 0.5. The *fliC* gene copies and added concentrations of strain EDL933 (data

256   converted into Log_{10}) fitted well to a linear model ($R^2 = 0.999$) (Fig. S3, Supporting

257   Information). With or without intercept, the slopes were close to 1, indicating that the

258   quantified *fliC* gene in the extracted soil DNA can be used to predict the

259   concentrations of strain EDL993 in soil with high accuracy.

260   **Establishment and validation of the iHAAQ method**

261   After quality filtering, there was a total of 465,220 high-quality sequences (66460

262   sequences per sample), and the sequences were classified into 35 phyla. At the genus

263   level, strain EDL933 was identified and classified as genus *Escherichia-Shigella* in

264 the SILVA database. The relative abundances of genus *Escherichia-Shigella* in

265 Treatments from E4 to E9 were much higher than that in Cont, and they increased

266 when more strain EDL933 was added to the soil (Fig. 2). The percentage of genus

267 *Escherichia-Shigella* in phylum *Proteobacteria* was 89.76% in Treatment E9,

268 compared with 0.07% in Cont.

269 After the *fliC* gene copies were multiplied by 7 times (marked as $7 \times fliC$), they

270 were in the same order of magnitude of the absolute abundance of Esch-V3 and

271 Esch-V4, except for Treatment E4 (Table 1). The ANOVA results showed that there

272 was no significant difference ($p > 0.05$) between the $7 \times fliC$ and Esch-V4 in all the

273 treatments, except for Treatment E4 (Table 1). In Treatment E4, both the *fliC* gene

274 copy number and $7 \times fliC$ gene copy number were significant lower than the absolute

275 abundances of Esch-V3 and Esch-V4.

276 Furthermore, the $7 \times fliC$ gene copies were linearly related to the calculated

277 absolute copies of Esch-V3 and Esch-V4 (Table 1, Fig. 3). All the four linear

278 equations fitted to the experimental data very well (R^2, 0.998 - 0.999), indicating that

279 the $7 \times fliC$ gene copies of the strain EDL933 can be reliably predicted from the

280 absolute abundances of Esch-V3 and Esch-V4 (Fig. 3). Thus, it is apparent that the

281 iHAAQ method based on the integration of qPCR and high-throughput sequencing

282 can be used to calculate the absolute abundance (Fig. 1, Fig. 4).

283 **Examining dynamics of soil bacterial communities by the iHAAQ method.**

284   From 7 days to 28 days incubation, the residual PHE was declined rapidly from 22.01

285   to 8.44 μg (g dry wt soil)$^{-1}$ in Treatment P and from 7.53 to 4.38 μg (g dry wt soil)$^{-1}$ in

286   Treatment W. Compared with Treatments P and W, Treatment S contained the highest

287   residual PHE and had the slightest change during the 28-d incubation (47.79 - 60.45

288   μg (g dry wt soil)$^{-1}$). Following the procedures as shown in Fig. 1, the dynamics of

289   soil bacterial community in Treatments P, W, and S were assessed by the iHAAQ

290   method. The total bacterial quantities increased with the incubation time in

291   Treatments P and W, but decreased with the incubation time in Treatment S (Fig. S4,

292   Supporting Information).

293   There were mainly 17 phyla in which the relative abundances were greater than

294   0.1% in each of the samples (Table S1, Fig. S5, Supporting Information). The

295   dynamics of relative abundances reflected the soil bacterial composition changes

296   during incubation (Fig. S5, Supporting Information). Overall, the relative abundances

297   decreased with the increase of incubation time in more than half of the phyla. The

298   phyla   of   *Firmicutes*,   *Proteobacteria*,   *Planctomycetes*,   *Bacteroidetes*,   and

299   *Actinobacteria* decreased by 5.58%, 1.93%, 1.59%, 1.28%, and 1.24%, respectively,

300   in Treatment P; the phyla of *Proteobacteria*, *Firmicutes*, and *Verrucomicrobia*

301   decreased by 9.24%, 3.41%, and 0.93%, respectively, in Treatments W; and the phyla

302   of *Actinobacteria*, *Chloroflexi*, *Gemmatimonadetes*, and *Proteobacteria* decreased by

303   6.57%, 4.24%, 2.77%, and 1.40%, respectively, in Treatment S. However, the phylum

304   *Firmicutes* in Treatment S increased substantially from 8.53% to 25.87%, and

305   eventually it became the dominated bacteria in soil.

306   Nevertheless, relative abundance does not provide information about change of

307   total bacterial quantities in soil. Based on the absolute abundances calculated by our

308   proposed iHAAQ method, there were only 5 phyla showing decrease in both the

309   relative and absolute abundances in Treatment P, while the relative and absolute

310   abundances of other phyla showed opposite trend (Fig. 1, Fig. 5, Fig. 6, Fig. 7; Table

311   S1, Fig. S5, Supporting Information). As an example, the difference in relative and

312   absolute abundances was especially evident for phylum *Proteobacteria* in Treatment

313   P, in which the relative abundance decreased by 6.67%, while the absolute

314   abundances increased by 42.22% and 21.02% in V3 and V4 regions, respectively.

315   Such inconsistent result was also observed in Treatments W and S. In Treatment W,

316   the relative abundance of 4 and 6 phyla decreased, while the absolute abundances of

317   all phyla increased. In Treatment S, none of the absolute abundances of phylum

318   increased, but the relative abundance of phylum *Firmicutes* increased by 17.54%.

319   Discrepancy between relative and absolute abundances also occurred at the genus

320   level (Fig. 7, Table S1, Supporting Information). In Treatment P, 135 and 57 opposite

321   results in relative and absolute abundances were observed in V3 and V4 regions,

322   respectively. While in Treatment W, 77 and 160 opposite relative and absolute

323   abundances were found in V3 and V4 regions, respectively. Even more discrepancies

324   (258 and 245 genera in V3 and V4 regions, respectively) were observed in Treatment

325   S. While it was a general trend that the relative abundances of genera decreased while

326   their absolute abundances increased in Treatments P and W (Table S1, Supporting

327   Information). The opposite trend was observed in Treatment S, in which the relative

328   abundance increased while the absolute abundance decreased. For instance, the

329   relative abundance of the dominant genus *Bacillus*, which belonged to the phylum

330   *Firmicutes*, increased from 3.67% to 13.11% in Treatment S (Fig. 7). However, its

331   absolute abundance decreased by more than 20% (from 8.88×10^7 to 5.26×10^7

332   copies (g dry wt soil)$^{-1}$ in V3 region, and from 1.10×10^8 to 8.69×10^7 copies (g dry

333   wt soil)$^{-1}$ in V4 region).

334   The added strain WG5 in the W0 sample was also detected by high-throughput

335   sequencing and classified as genus *Massilia*. The relative abundance of *Massilia* was

336   higher than those in the P0 and S0 samples (< 0.10%). The change in relative and

337   absolute abundances of *Massilia* were also different in Treatment W after the soil was

338   incubated at 28 °C for 28 days. Its relative abundances in the samples were 10.71%

339   and 2.71%, respectively, in Day 0 (W0) and Day 28 (W28) (Fig. 7), dropped almost

340   three quarters (74.70%). Whereas its absolute abundances in W0 and W28 samples

341   decreased from 3.86×10^8 to 1.40×10^8 copies (g dry wt soil)$^{-1}$, and from 5.13×10^8

342   to 2.47×10^8 copies (g dry wt soil)$^{-1}$, representing less than two thirds (63.69%) and

343   one half (51.84%) of decrease, respectively, in V3 and V4 regions.

344   Cluster analysis also showed the discrepancy between the relative and absolute

345   abundances of bacteria in the treatments (Fig. 6, Fig. S6, Supporting Information).

346   Based on the relative abundances, the samples of P0, W0 and P28, S0 were clustered

347   into two close groups, respectively, while the samples of W28 and S28 were far away

348   from those groups (Fig. 6A). However, the cluster based on the absolute abundances

349   showed totally different results. The samples of P28, W28 in V3 and V4 regions were

350   clustered into the same group (Fig. 6B, Fig. S6, Supporting Information). Meanwhile,

351   the samples of P0, W0 and S0, S28 in V4 region were clustered into two different

352   groups (Fig. 6B), which is more rational than the result in V3 region (Fig. S6,

353   Supporting Information).

354   **Discussion**

355   The V3 and V4 regions of 16S rRNA gene were universally used to detect the total

356   bacterial quantity in soil via qPCR (Reddy et al. 2012; Guinane et al. 2013). Youssef

357   et al. (2009) observed that there exists difference in soil bacterial community in V3

358   and V4 regions. Except for Treatment E9, all other treatments or samples showed that

359   the total bacteria quantities in V4 region were significantly higher than those in V3

360   region. Comparing with the nearly complete 16S rRNA gene, the results of V3 region

361   generally underestimated the total quantity of soil bacterial community, while the

362   results of V4 region were usually comparable to it (Youssef et al. 2009). Meanwhile,

363   Wang et al. (2007) also reported that V4 region has the highest reliability in

364   classification among the 16S rRNA variable regions. Thus, it was suggested that V4

365   region was the best choice of 16S rRNA gene variable regions for bacterial

366   community characterization. Cluster analysis also proved that the results from the

367   same region (V4) of qPCR and high-throughput sequencing were better than those

368   from the different regions of them. Therefore, in the following discussion, we focused

369   our discussion only on the results of 16S rRNA gene in V4 region.

370   Comparing with the selected taxa reported by Dannemiller et al. (2014), it is very

371   difficult to select a strain from soil to verify the iHAAQ method. As the previous

372   research reported, the *fliC* gene of strain EDL933 can be precisely measured by qPCR

373   (Jun et al. 2011; Li & Chen 2012). Therefore, to test the accuracy of absolute

374   abundance calculated by our method, an internal reference strain EDL933 was

375   selected and introduced into the soil for validation. The high R^2 values of the linear

376   relationships indicate strong association between the *fliC* gene and the actual added

377   strain. The similar result was also found by Tourlousse et al. (2017) that the slopes of

378   the spike-in read counts and input amounts were close to 1. Thus, we concluded the

379   iHAAQ method based on qPCR and high-throughput sequencing in the same V4

380   region can be used to reliably predict the absolute abundance of *Escherichia-Shigella*.

381   The validation of the iHAAQ method was further verified by the added *Massilia*

382   sp. strain WG5 in W0 sample. The absolute abundance of genus *Massilia* calculated

383   by the iHAAQ method (5.13×10^8 copies (g dry wt soil)$^{-1}$) was one order of

384   magnitude higher than the added strain WG5 concentration (1.00×10^7 CFU (g dry wt

385   soil)$^{-1}$). After considering that one chromosome of strain WG5 has 9 copies of 16S

386   rRNA gene (Lou et al. 2016), the calculated absolute abundance of *Massilia* was close

387   to the added concentration of strain WG5.

388   When the iHAAQ method is applied to quantify soil bacterial community

389   dynamics, both similar and opposite trends of relative and absolute abundances were

390   observed, which agreed with early investigations (Dannemiller et al. 2014; Prest et al.

391   2014; Smets et al. 2016; Stokell et al. 2016; Zhang et al. 2017). We observed that both

392   the relative and absolute abundances of genus *Massilia* decreased in Treatments W

393   after 28 days of incubation. The low residual PHE and its slow decrease in Treatment

394   W indicated its low bioavailability, which might lead to the slight decrease of the

395   genus *Massilia*.

396   In Treatment S, the addition of NaN_3 significantly decreased the absolute

397   abundances of soil bacterial taxa, which agreed with the report of Rouwane et al.

398   (2016). The genus *Bacillus* became the dominant composition with the increased

399   relative abundance but the decreased absolute abundance in the bacteria community

400   of Treatment S. Klein et al. (1994) stated that some species in genus *Bacillus* were

401   strongly resistant to NaN_3. It might result in the less decrease of genus *Bacillus* than

402   other bacteria in Treatment S.

403   The relative abundance can reflect the composition changes of the microbial

404   communities, while the absolute abundance can uncover the dynamics of the total

405   quantities (Reddy et al. 2012; Guinane et al. 2013; Dannemiller et al. 2014; Prest et al.

406   2014; Props et al. 2017; Zhang et al. 2017). The relative abundance from

407   high-throughput sequencing can only partially characterized the bacterial community

408   (Smets et al. 2016; Props et al. 2017; Tourlousse et al. 2017). The actual microbial

409   community dynamics can be more objectively described by both the relative and

410   absolute abundances from the iHAAQ method than by the relative abundance alone.

411   Furthermore, even if the relative abundance of samples had been obtained from

412   the high-throughput sequencing, the absolute abundance of each taxon in the samples

413   will be achieved with the supplementary qPCR analysis. It was recommended to use

414   the same variable region of 16S rRNA gene for qPCR and high-throughput

415   sequencing. Hence, the iHAAQ method is suitable to determine the absolute

416   abundance of bacterial community and can be applied to studying the quantitative

417   bacterial ecology in soil.

418   It is worthy to mention that the range of the valid absolute abundance of each

419   taxon is limited by the high-throughput sequencing technique. Based on

420   pyrosequencing data of a fungal aerosol population, the exact measure of absolute

421   abundance of the fungal species can be achieved at the sufficiently high relative

422   abundance (above 4.4×10^{-4}) (Dannemiller et al. 2014). In our study, the least

423   sequence number of indigenous bacteria in soil were 5 reads (i.e. the least relative

424   abundance was about 8.0×10^{-5}). The qPCR results showed that the total bacteria

425   quantities in our samples were from 10^8 to 10^{10} copies (g dry wt soil)$^{-1}$. Therefore, the

426   indigenous bacteria in the soil samples can be detected by the iHAAQ method when

427   their absolute abundances were above 10^3 copies (g dry wt soil)$^{-1}$. With the increase

428   of high-throughput sequencing depth, more rare species with lower abundance can be

429 detected (Agogue et al. 2011; Polka et al. 2015), which may help to better

430 characterize microbial communities.

431 **Conclusion**

432 Our results indicate that the proposed iHAAQ method allows to obtain the absolute

433 abundance of each taxon in soil bacteria communities by combining qPCR (total

434 bacterial quantity) and high-throughput sequencing (relative abundance of each taxon),

435 which could not be achieved separately by either method. The absolute abundance

436 reflects the dynamics of a bacterial community more accurately than the relative

437 abundance data. In our study, the iHAAQ method was successfully applied to

438 characterize the soil bacterial communities with both relative and absolute abundances,

439 which could provide more information than with the relative abundance alone by

440 high-throughput sequencing. The proposed iHAAQ method could help to study soil

441 quantitative bacterial ecology in future application.

442 **REFERENCES**

443 Agogue H, Lamy D, Neal PR, Sogin ML, and Herndl GJ. 2011. Water mass-specificity of

444   bacterial communities in the North Atlantic revealed by massively parallel

445   sequencing. *Mol Ecol* 20:258-274. 10.1111/j.1365-294X.2010.04932.x

446 Blagodatskaya E, and Kuzyakov Y. 2013. Active microorganisms in soil: Critical review of

447   estimation criteria and approaches. *Soil Biol Biochem* 67:192-211.

448   10.1016/j.soilbio.2013.08.024

449 Brookes PC, Landman A, Pruden G, and Jenkinson DS. 1985. Chloroform fumigation and the

450   release of soil nitrogen: a rapid direct extraction method to measure microbial

451   biomass nitrogen in soil. *Soil Biol Biochem* 17:837-842.

452   10.1016/0038-0717(85)90144-0

453 Cai Y, Zheng Y, Bodelier PL, Conrad R, and Jia Z. 2016. Conventional methanotrophs are

454   responsible for atmospheric methane oxidation in paddy soils. *Nat Commun* 7:11728.

455   10.1038/ncomms11728

456 Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N,

457   Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE,

458   Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR,

459   Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, and Knight R.

460   2010. QIIME allows analysis of high-throughput community sequencing data. *Nat*

461   *Methods* 7:335-336. 10.1038/nmeth.f.303

462 Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM,

463   Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, and Knight R. 2012.

464   Ultra-high-throughput microbial community analysis on the Illumina HiSeq and

465   MiSeq platforms. *ISME J* 6:1621-1624. 10.1038/ismej.2012.8

466 Curtis TP, and Sloan WT. 2005. Microbiology. Exploring microbial diversity--a vast below.

467   *Science* 309:1331-1333. 10.1126/science.1118176

468 Dannemiller KC, Lang-Yona N, Yamamoto N, Rudich Y, and Peccia J. 2014. Combining

469   real-time PCR and next-generation DNA sequencing to provide quantitative

470   comparisons of fungal aerosol populations. *Atmospheric Environment* 84:113-121.

471   10.1016/j.atmosenv.2013.11.036

472 Davey HM. 2011. Life, death, and in-between: meanings and methods in microbiology. *Appl*

473   *Environ Microbiol* 77:5571-5576. 10.1128/aem.00744-11

474 Ding N, Hayat T, Wang J, Wang H, Liu X, and Xu J. 2011. Responses of microbial

475   community in rhizosphere soils when ryegrass was subjected to stress from PCBs. *J*

476   *Soils Sediments* 11:1355-1362. 10.1007/s11368-011-0412-x

477 Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST.

478   *Bioinformatics* 26:2460-2461. 10.1093/bioinformatics/btq461

479 Edgar RC, Haas BJ, Clemente JC, Quince C, and Knight R. 2011. UCHIME improves

480   sensitivity and speed of chimera detection. *Bioinformatics* 27:2194-2200.

481   10.1093/bioinformatics/btr381

482   Gu H, Chen Y, Liu X, Wang H, Shen-Tu J, Wu L, Zeng L, and Xu J. 2017. The effective

483   migration of *Massilia* sp. WF1 by *Phanerochaete chrysosporium* and its

484   phenanthrene biodegradation in soil. *Science of the Total Environment*

485   593-594:695-703. 10.1016/j.scitotenv.2017.03.205

486   Guinane CM, Tadrous A, Fouhy F, Ryan CA, Dempsey EM, Murphy B, Andrews E, Cotter

487   PD, Stanton C, and Ross RP. 2013. Microbial composition of human appendices from

488   patients following appendectomy. *MBio* 4:e00366-00312.

489   Jun L, Xie Y, Xuan X, Chen Z, Yang W, Chun-Xia M, Shuai H, Peng H, Li-Gan X, and Xie

490   Y. 2011. Establishment of real-time quantitative PCR for detection of *Escherichia*

491   *coli* O157:H7. *Animal Husbandry & Veterinary Medicine* 43:5-9.

492   Klein M, Hofmann B, Klose M, and Freudl R. 1994. Isolation and characterization of a

493   *Bacillus subtilis secA* mutant allele conferring resistance to sodium azide. *FEMS*

494   *Microbiol Lett* 124:393-397.

495   Kleyer H, Tecon R, and Or D. 2017. Resolving Species Level Changes in a Representative

496   Soil Bacterial Community Using Microfluidic Quantitative PCR. *Front Microbiol*

497   8:2017. 10.3389/fmicb.2017.02017

498   Kolde   R.   2015.   pheatmap: Pretty   Heatmaps.   R   package   version   1.0.8.

499   https://cran.r-project.org/package=pheatmap.

500   Kozich JJ, Westcott SL, Baxter NT, Highlander SK, and Schloss PD. 2013. Development of a

501   dual-index sequencing strategy and curation pipeline for analyzing amplicon

502   sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol*

503   79:5112-5120. 10.1128/AEM.01043-13

504   Latif H, Li HJ, Charusanti P, Palsson BO, and Aziz RK. 2014. A gapless, unambiguous

505   genome sequence of the enterohemorrhagic *Escherichia coli* O157:H7 strain EDL933.

506   *Genome Announc* 2:e00821-00814. 10.1128/genomeA.00821-14

507   Li B, and Chen JQ. 2012. Real-time PCR methodology for selective detection of viable

508   *Escherichia coli* O157:H7 cells by targeting Z3276 as a genetic marker. *Appl Environ*

509   *Microbiol* 78:5297-5304. 10.1128/aem.00794-12

510   Li Z, Xu J, Tang C, Wu J, Muhammad A, and Wang H. 2006. Application of 16S rDNA-PCR

511   amplification and DGGE fingerprinting for detection of shift in microbial community

512   diversity in Cu-, Zn-, and Cd-contaminated paddy soils. *Chemosphere* 62:1374-1380.

513   10.1016/j.chemosphere.2005.07.050

514   Lok C. 2015. Mining the microbial dark matter. *Nature* 522:270-273. 10.1038/522270a

515   Lou J, Gu H, Wang H, An Q, and Xu J. 2016. Complete genome sequence of *Massilia* sp.

516   WG5, an efficient phenanthrene-degrading bacterium from soil. *J Biotechnol*

517   218:49-50. 10.1016/j.jbiotec.2015.11.026

518   Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z, Brookes PC, Xu J, and Gilbert JA. 2016.

519   Geographic patterns of co-occurrence network topological features for soil microbiota

520   at continental scale in eastern China. *ISME J* 10:1891-1901. 10.1038/ismej.2015.261

521   Magoc T, and Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve

522   genome assemblies. *Bioinformatics* 27:2957-2963. 10.1093/bioinformatics/btr507

523   Philippot L, Bru D, Saby NP, Cuhel J, Arrouays D, Simek M, and Hallin S. 2009. Spatial

524   patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S

525   rRNA   bacterial   tree.   *Environ   Microbiol*   11:3096-3104.

526   10.1111/j.1462-2920.2009.02014.x

527   Polka J, Rebecchi A, Pisacane V, Morelli L, and Puglisi E. 2015. Bacterial diversity in typical

528   Italian salami at different ripening stages as revealed by high-throughput sequencing

529   of 16S rRNA amplicons. *Food Microbiol* 46:342-356. 10.1016/j.fm.2014.08.023

530   Prest EI, El-Chakhtoura J, Hammes F, Saikaly PE, van Loosdrecht MC, and Vrouwenvelder

531   JS. 2014. Combining flow cytometry and 16S rRNA gene pyrosequencing: a

532   promising approach for drinking water monitoring and characterization. *Water Res*

533   63:179-189. 10.1016/j.watres.2014.06.020

534   Props R, Kerckhof FM, Rubbens P, De Vrieze J, Hernandez Sanabria E, Waegeman W,

535   Monsieurs P, Hammes F, and Boon N. 2017. Absolute quantification of microbial

536   taxon abundances. *Isme j* 11:584-587. 10.1038/ismej.2016.117

537   Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, and Glockner FO.

538   2013. The SILVA ribosomal RNA gene database project: improved data processing

539   and web-based tools. *Nucleic Acids Res* 41:D590-596. 10.1093/nar/gks1219

540   RCore T. 2013. R: A language and environment for statistical computing. R Foundation for

541   Statistical Computing, Vienna, Austria. URL http://www.r-project.org/.

542   Reddy BV, Kallifidas D, Kim JH, Charlop-Powers Z, Feng Z, and Brady SF. 2012. Natural

543   product biosynthetic gene diversity in geographically distinct soil microbiomes. *Appl*

544   *Environ Microbiol* 78:3744-3752. 10.1128/aem.00102-12

545   Rouwane A, Rabiet M, Bourven I, Grybos M, Mallet L, and Guibaud G. 2016. Role of

546   microbial reducing activity in antimony and arsenic release from an unpolluted

547   wetland soil: a lab scale study using sodium azide as a microbial inhibiting agent.

548   *Environmental Chemistry* 13:945-954. 10.1071/en16029

549   Singer E, Andreopoulos B, Bowers RM, Lee J, Deshpande S, Chiniquy J, Ciobanu D, Klenk

550   HP, Zane M, Daum C, Clum A, Cheng JF, Copeland A, and Woyke T. 2016. Next

551   generation sequencing data of a defined microbial mock community. *Sci Data*

552   3:160081. 10.1038/sdata.2016.81

553   Smets W, Leff JW, Bradford MA, McCulley RL, Lebeer S, and Fierer N. 2016. A method for

554   simultaneous measurement of soil bacterial abundances and community composition

555   via 16S rRNA gene sequencing. *Soil Biology & Biochemistry* 96:145-151.

556   10.1016/j.soilbio.2016.02.003

557   Stokell JR, Hamp TJ, and Steck TR. 2016. Examining changes in bacterial abundance in

558   complex communities using next-generation sequencing is enhanced with

559   quantitative PCR. *Antonie Van Leeuwenhoek International Journal of General and*

560   *Molecular Microbiology* 109:1161-1166. 10.1007/s10482-016-0707-4

561   Tourlousse DM, Yoshiike S, Ohashi A, Matsukura S, Noda N, and Sekiguchi Y. 2017.

562   Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon

563   sequencing. *Nucleic acids research* 45:14. 10.1093/nar/gkw984

564   Vance E, Brookes P, and Jenkinson D. 1987. An extraction method for measuring soil

565   microbial biomass C. *Soil Biol Biochem* 19:703-707.

566   Wang H, Ibekwe AM, Ma J, Wu L, Lou J, Wu Z, Liu R, Xu J, and Yates SR. 2014. A glimpse

567   of *Escherichia coli* O157:H7 survival in soils from eastern China. *Science of the*

568   *Total Environment* 476:49-56. 10.1016/j.scitotenv.2014.01.004

569   Wang H, Lou J, Gu H, Luo X, Yang L, Wu L, Liu Y, Wu J, and Xu J. 2016. Efficient

570   biodegradation of phenanthrene by a novel strain *Massilia* sp. WF1 isolated from a

571   PAH-contaminated soil. *Environ Sci Pollut Res Int* 23:13378-13388.

572   10.1007/s11356-016-6515-6

573   Wang Q, Garrity GM, Tiedje JM, and Cole JR. 2007. Naive Bayesian classifier for rapid

574   assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ*

575   *Microbiol* 73:5261-5267. 10.1128/aem.00062-07

576   Wei S, and Pan S. 2010. Phytoremediation for soils contaminated by phenanthrene and

577   pyrene with multiple plant species. *J Soils Sediments* 10:886-894.

578   10.1007/s11368-010-0216-4

579   Whitman WB, Coleman DC, and Wiebe WJ. 1998. Prokaryotes: the unseen majority. *Proc*

580   *Natl Acad Sci U S A* 95:6578-6583.

581   Youssef N, Sheik CS, Krumholz LR, Najar FZ, Roe BA, and Elshahed MS. 2009.

582   Comparison of species richness estimates obtained using nearly complete fragments

583   and simulated pyrosequencing-generated fragments in 16S rRNA gene-based

584   environmental surveys. *Appl Environ Microbiol* 75:5227-5236.

585   10.1128/aem.00592-09

586   Zhang Z, Qu Y, Li S, Feng K, Wang S, Cai W, Liang Y, Li H, Xu M, Yin H, and Deng Y.

587   2017. Soil bacterial quantification approaches coupling with relative abundances

588   reflecting the changes of taxa. *Sci Rep* 7:4837. 10.1038/s41598-017-05260-w

589

Detection and calculation of the added internal reference strain EDL933 (copies (g dry wt soil)$^{-1}$) in soil by qPCR (*fliC* gene and 7 × *fliC* gene) and iHAAQ method (Esch-V3 and Esch-V4).

Treatments E9 to E4 are corresponding to with 6 different concentrations of 10^9 to 10^4 CFU (g dry wt soil)$^{-1}$. Different letters indicate significant difference ($p < 0.05$) among 7 × *fliC* gene, Esch-V3 and Esch-V4 of each treatment detected by ANOVA test.

1 **Table 1**

Treatment	*fliC*	$7 \times fliC$	Esch-V3	Esch-V4
E9	$4.79 \times 10^9 \pm 1.30 \times 10^9$	$3.35 \times 10^{10} \pm 9.11 \times 10^8$ a	$4.67 \times 10^{10} \pm 9.07 \times 10^8$ a	$4.82 \times 10^{10} \pm 1.09 \times 10^9$ a
E8	$5.96 \times 10^8 \pm 1.20 \times 10^7$	$4.17 \times 10^9 \pm 8.38 \times 10^7$ ab	$3.59 \times 10^9 \pm 2.20 \times 10^8$ b	$4.37 \times 10^9 \pm 3.53 \times 10^8$ a
E7	$5.28 \times 10^7 \pm 1.40 \times 10^6$	$3.70 \times 10^8 \pm 9.80 \times 10^6$ a	$2.94 \times 10^8 \pm 1.17 \times 10^7$ a	$3.39 \times 10^8 \pm 1.13 \times 10^7$ a
E6	$5.70 \times 10^6 \pm 4.62 \times 10^5$	$3.99 \times 10^7 \pm 3.23 \times 10^6$ a	$2.82 \times 10^7 \pm 3.86 \times 10^6$ b	$4.08 \times 10^7 \pm 2.52 \times 10^6$ a
E5	$5.65 \times 10^5 \pm 5.19 \times 10^4$	$3.95 \times 10^6 \pm 3.63 \times 10^5$ ab	$3.48 \times 10^6 \pm 2.83 \times 10^4$ b	$6.47 \times 10^6 \pm 2.07 \times 10^5$ a
E4	$8.11 \times 10^4 \pm 3.15 \times 10^3$	$5.68 \times 10^5 \pm 2.21 \times 10^4$ c	$1.47 \times 10^6 \pm 7.66 \times 10^4$ b	$3.14 \times 10^6 \pm 1.25 \times 10^5$ a

Figure 1

A schematic workflow of the iHAAQ method and its validation procedure.

For method validation, the internal reference strain (*IRS*) was introduced into soil before DNA extraction and measured by iHAAQ and qPCR methods. These two results were combined and analyzed by liner regression to validate the iHAAQ method.
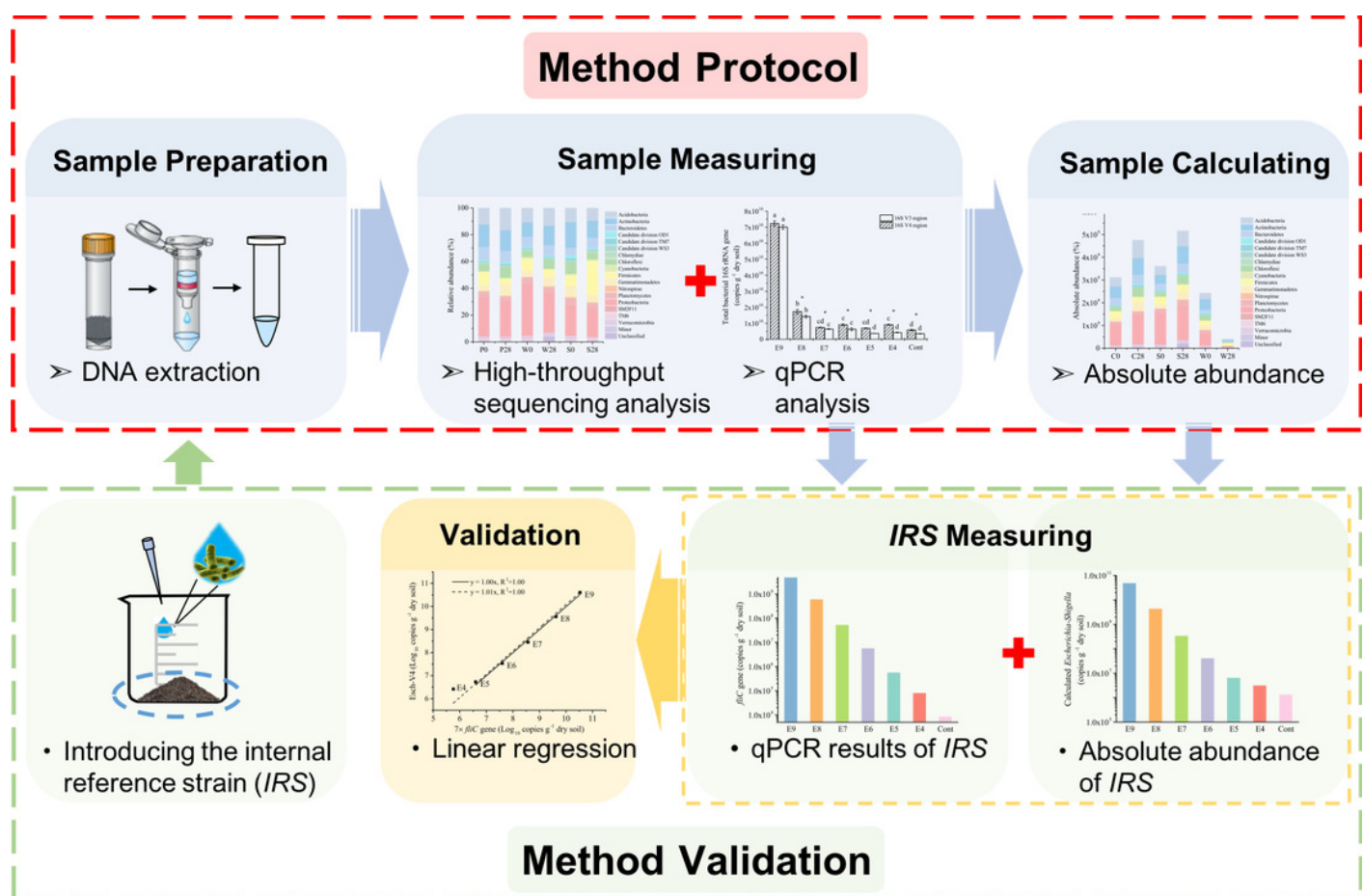
Figure 2

Relative abundances of *Escherichia-Shigella* genus in the different treatments.

The added internal standard strain EDL933 is classified as *Escherichia-Shigella* genus in the SILVA database. Treatments E9 to E4 are the same as described in Table 1, and Cont is the original control soil.
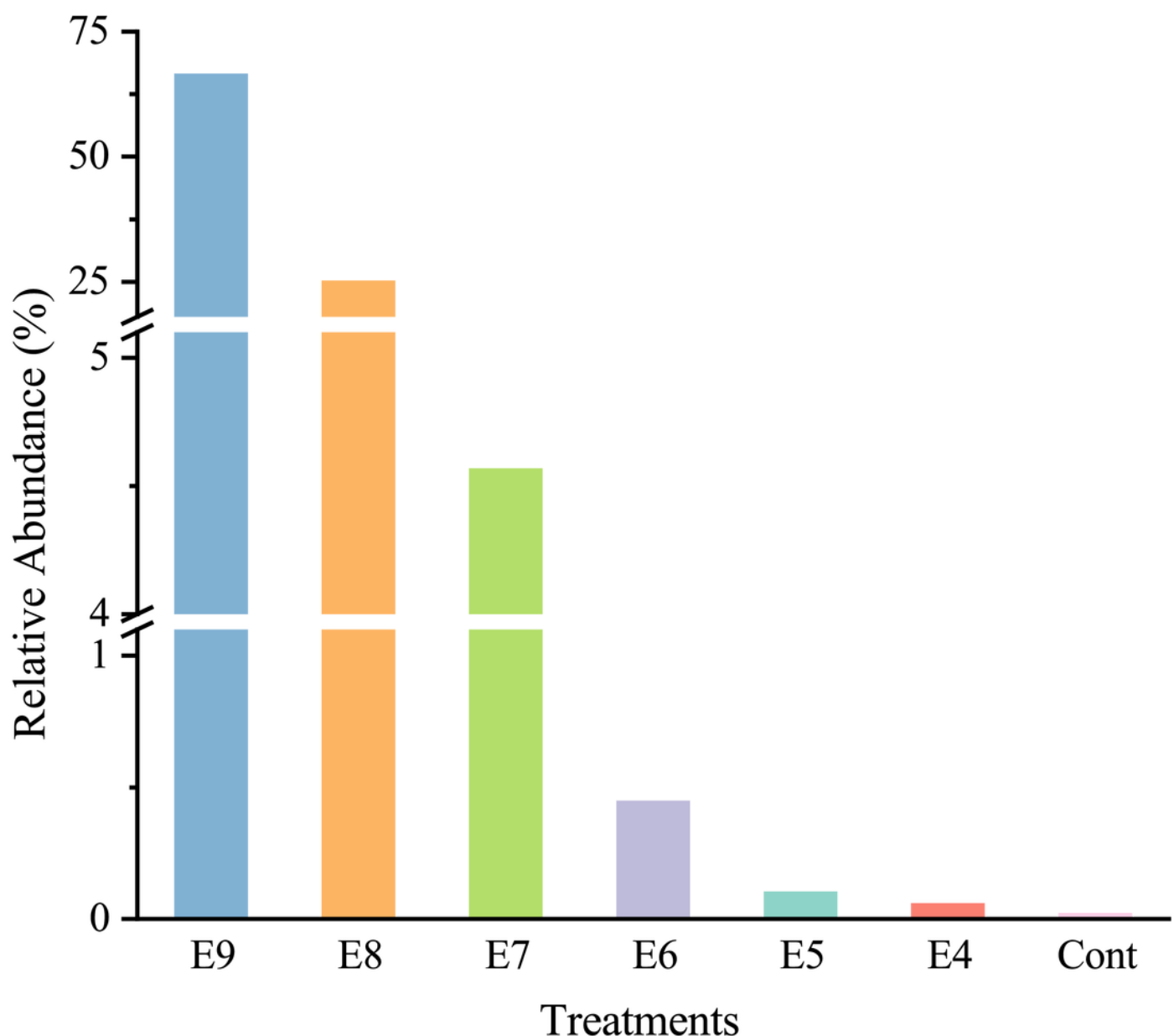
Figure 3

Linear regressions for the quantities of added internal standard strain EDL933 from qPCR and iHAAQ methods in the different treatments.

The 7 × *fliC* gene copies of added internal standard strain EDL933 detected by qPCR. The Esch-V3 (A) and Esch-V4 (B) represent *Escherichia-Shigella* copies calculated by iHAAQ method with its relative abundance multiplying by the total bacterial 16S rRNA gene copies in V3 and V4 variable regions, respectively. All intercepts were fixed to the origin point (0). Treatments E9 to E4 are the same as described in Table 1. The solid and dash lines represent the linear models without and with Treatment E4, respectively.
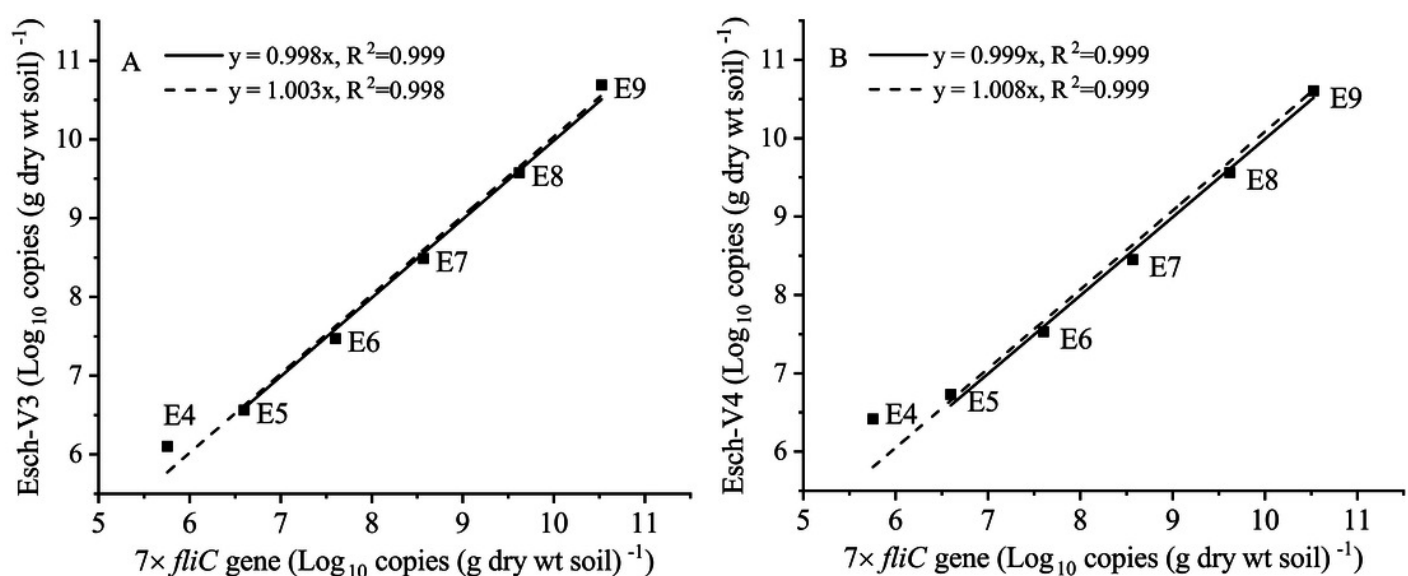
Figure 4

Absolute abundances of the major phyla in different treatments.

All classified phyla with a relative abundance of < 0.1% in a sample were combined and reported as Minor. Treatments E9 to E4 are the same as in Table 1, and the Cont is the original control soil. No shade and shadow histograms represent the phyla number quantified by the V3 and V4 regions of 16S rRNA gene, respectively.
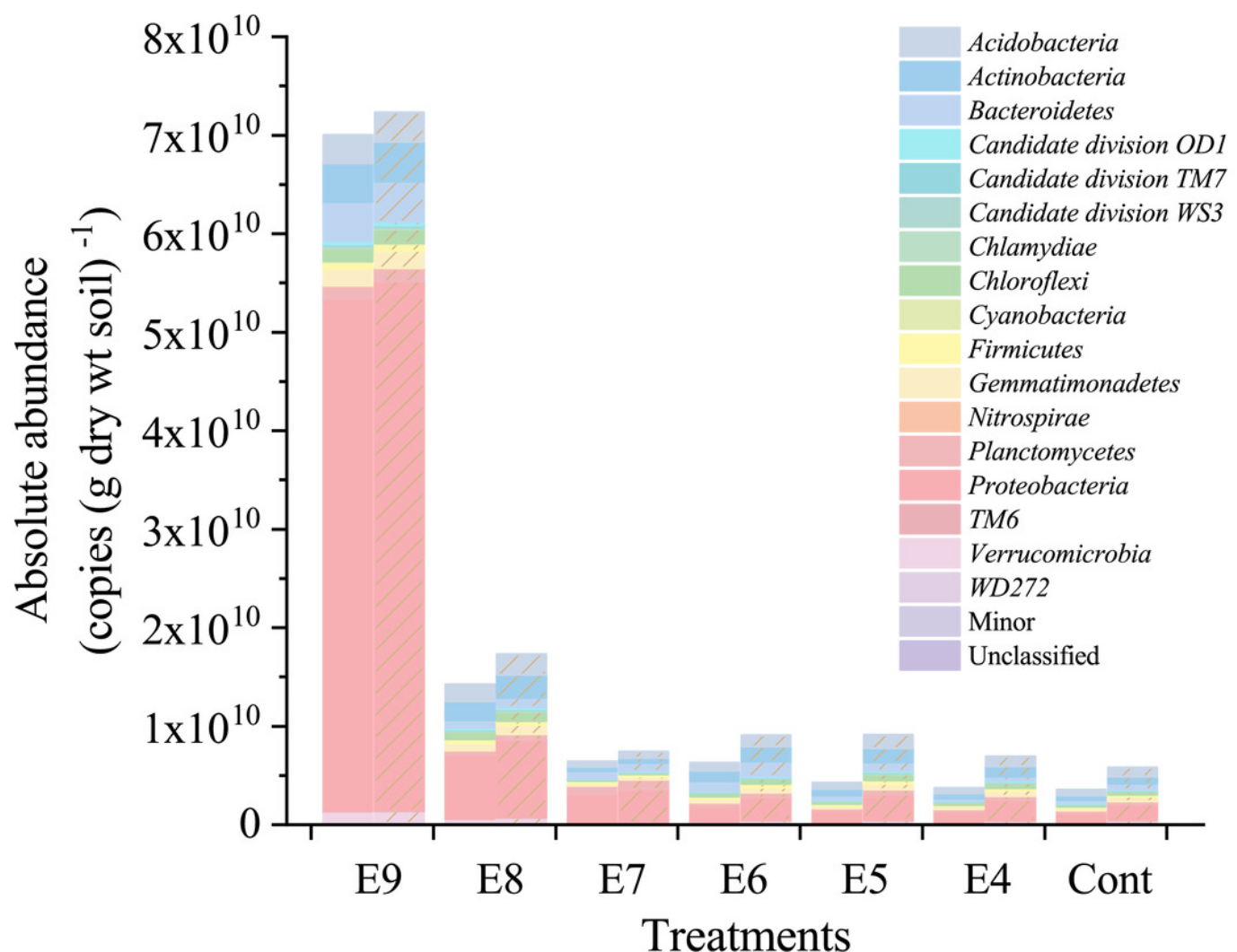
Figure 5

Absolute abundances of the major phyla determined by the iHAAQ method in different soil samples.

All classified phyla with a relative abundance of < 0.1% in a sample were combined and reported as Minor. P, S, and W represent the treatments of soil added with PHE (100 µg (g dry wt soil)$^{-1}$), NaN$_3$ (0.1%, w/w), and PHE (100 µg (g dry wt soil)$^{-1}$) and strain WG5 (approximately 1.00×10^7 CFU (g dry wt soil)$^{-1}$), respectively. 0 and 28 represent sampling at 0 and 28 days of incubation. No shade and shadow histograms represent the phyla number quantified by the V3 and V4 regions of 16S rRNA gene, respectively.
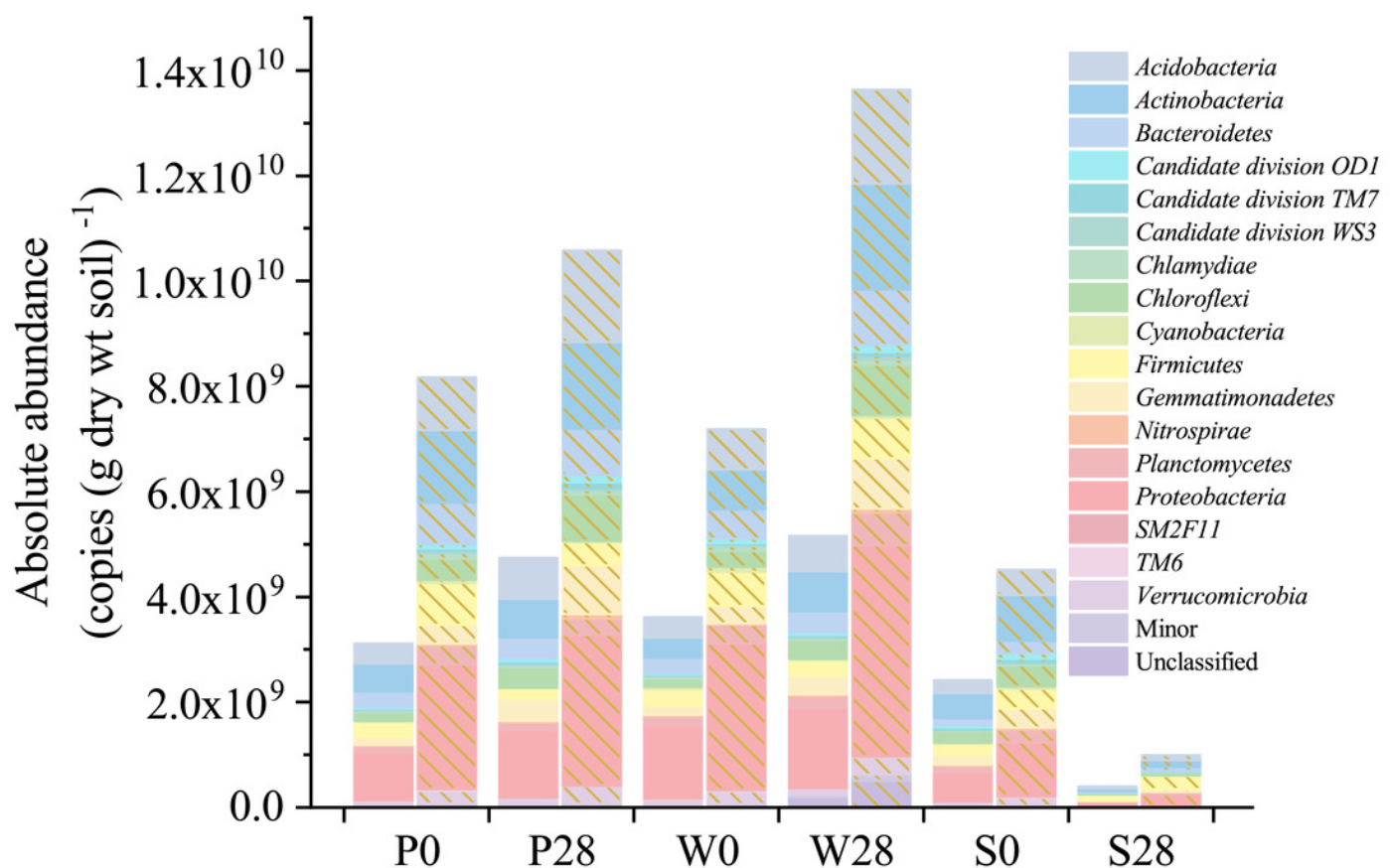
Figure 6

Heatmaps showing the relative abundances (A) and absolute abundances (B) determined by the iHAAQ method of the major phyla quantified by the V4 region of 16S rRNA gene.

All classified phyla with a relative abundance of < 0.1% in a sample were combined and reported as Minor. P0, P28, S0, S28, W0 and W28 are the same as in Fig. 5. The color codes indicate the relative (A) or absolute (B) abundances, ranging from blue (low abundance) to red (high abundance).
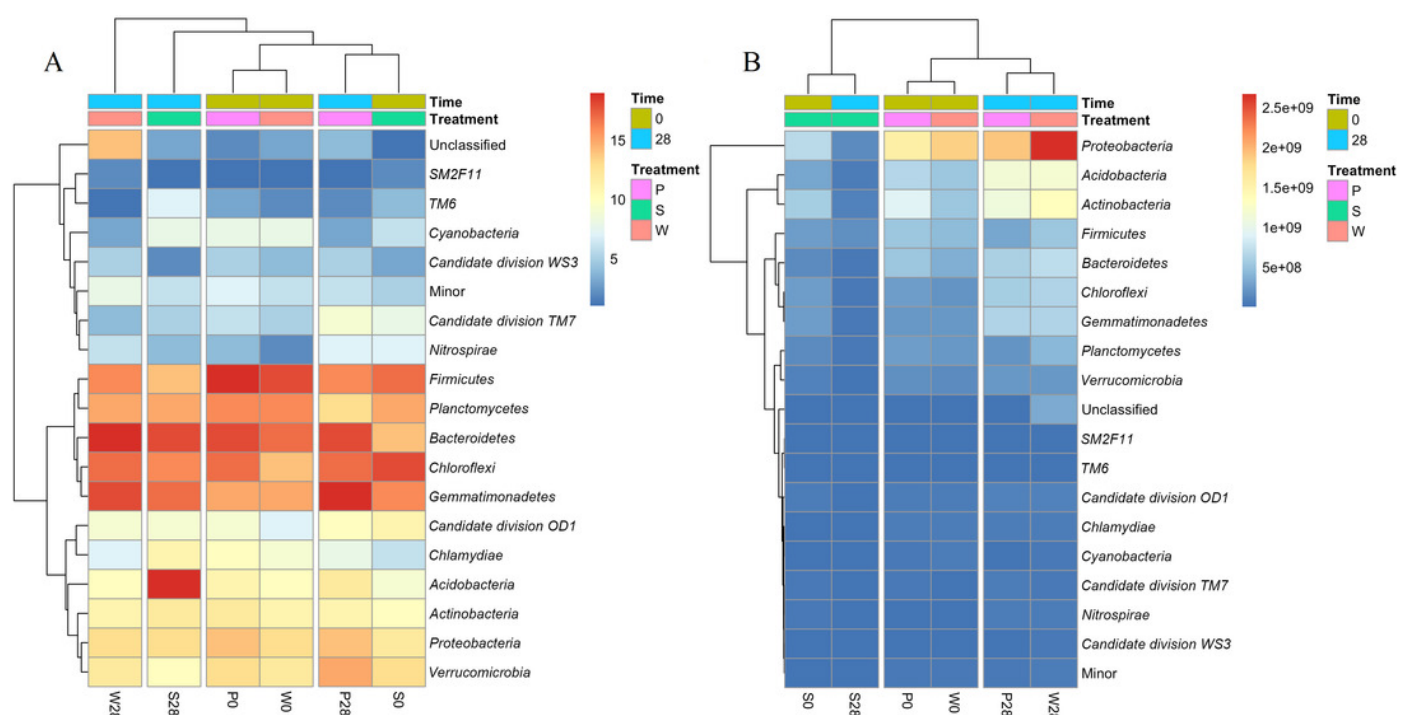
Figure 7

The representative genera of Treatments P, W, and S.

P0, P28, S0, S28, W0 and W28 are the same as in Fig. 5.