

This is a post-peer-review, pre-copyedit version of an article published in *Science China Life Sciences* (Ed. Springer).

The final authenticated version is available online at:
<http://dx.doi.org/10.1007/s11427-018-9364-7>

Under a “All rights reserved” license.

1 **QMEC: A tool for high-throughput quantitative assessment of microbial
2 functional potential in C, N, P, and S biogeochemical cycling**

3 Bang-Xiao Zheng^{1,2,3,4}, Yong-Guan Zhu^{1,5}, Jordi Sardans^{3,4}, Josep Peñuelas^{3,4*}, Jian-
4 Qiang Su^{1**}

5 ¹ *Key Laboratory of Urban Environment and Health, Institute of Urban Environment,
6 Chinese Academy of Sciences, Xiamen 361021, PR China;*

7 ² *University of Chinese Academy of Sciences, Beijing 100049, PR China;*

8 ³ *Consejo Superior de Investigaciones Científicas (CSIC), Global Ecology Unit,
9 Centre for Ecological Research and Forestry Applications (CREAF)–CSIC–
10 Universitat Autònoma de Barcelona (UAB), Bellaterra, 08193 Barcelona, Catalonia,
11 Spain;
12 ⁴ *CREAF, Cerdanyola del Vallès, 08193 Barcelona, Catalonia, Spain;
13 ⁵ *State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-
14 Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China.***

15 * Corresponding Author

16 Phone: (+34) 935812199; Fax: (+34) 935814151; E-mail: josep.penuelas@uab.cat

17 ** Corresponding Author

18 Phone: (+86) 592 6190997; Fax: (+86) 592 6190791; E-mail: jqsu@iue.ac.cn

19

20

21 **Abstract**

22 Microorganisms are major drivers of elemental cycling in the biosphere. Determining
23 the abundance of microbial functional traits involved in the transformation of nutrients,
24 including carbon (C), nitrogen (N), phosphorus (P) and sulfur (S), is critical for
25 assessing microbial functionality in elemental cycling. We developed a high-throughput
26 quantitative-PCR-based chip, QMEC (Quantitative Microbial Element Cycling), for
27 assessing and quantifying the genetic potential of microbiota to mineralize soil organic
28 matter and to release C, N, P and S. QMEC contains 72 primer pairs targeting 64
29 microbial functional genes for C, N, P, S and methane metabolism. These primer pairs
30 were characterized by high coverage (average of 18-20 phyla covered per gene) and
31 sufficient specificity (>70% match rate) with a relatively low detection limit (7-102
32 copies per run). QMEC was successfully applied to soil and sediment samples,
33 identifying significantly different structures, abundances and diversities of the
34 functional genes ($P < 0.05$). QMEC was also able to determine absolute gene abundance.
35 QMEC enabled the simultaneous qualitative and quantitative determination of 72 genes
36 from 72 samples in one run, which is promising for comprehensively investigating
37 microbially mediated ecological processes and biogeochemical cycles in various
38 environmental contexts including those of the current global change.

39 **Key words:**

40 microbial genes, functional potential, high-throughput qPCR, elemental cycling,
41 biogeochemical cycle, ecological process

42 **Introduction**

43 Microorganisms are the major drivers of biogeochemical cycles on Earth (van der
44 Heijden et al., 2008), substantially affecting carbon (C) and sulfur (S) metabolism,
45 organic-matter degradation, nitrogen (N) efflux and phosphorus (P) mobilization in
46 the environment (Zarraonaindia et al., 2013; Vanwonterghem et al., 2014). Those
47 processes may result in CO₂ elevation, greenhouse gases release, nutrient loading and
48 water consumption, which coupled with changes of interacting spheres of the earth
49 (Chapin et al., 2000; Bardgett and van der Putten, 2014). Comprehensive
50 investigation of microbial taxonomic composition and functional-gene diversity and
51 abundance are key for a better understanding of microbially mediated biogeochemical
52 processes and their current global changes (Stevenson and Cole, 1999; Penuelas et al.,
53 2013; Graham et al., 2016).

54 Culture-independent molecular technologies have been widely adopted to
55 investigate microbial phylogenetic and functional diversity and to evaluate their
56 responses to environmental changes (Zhou et al., 2015; Deng et al., 2016; Feng et al.,
57 2017) . The array-based PhyloChip (a high-density 16S gene oligonucleotide
58 microarray) and high-throughput sequencing of 16S rRNA gene fragments are two
59 commonly reported methods for determining the structure of microbial-community
60 composition (Schmidt et al., 1991; Hazen et al., 2010). Metasequencing, including
61 shotgun metagenomics, metatranscriptomics or their combination, enables the
62 functional characterization of novel genes, phylotypes, regulators and metabolic
63 pathways (Weinstock, 2012). The microarray-based GeoChip is more specific than

64 sequencing-based technologies for detecting functional genes of interest, especially
65 genes involved in elemental biogeochemistry (Tu et al., 2014).

66 Sequencing-based metagenomics and hybridization-based microarray
67 technologies have been successfully applied in studies of microbial ecology with high
68 gene coverage and resolution and can generally determine the relative abundance of
69 microbial taxa and functional genes in microbial communities, enabling comparison
70 among environmental samples. Relative abundances are the proportions of specific
71 taxa or genes in microbial communities, enabling the detection of increases or
72 decreases; however, relative abundance cannot well evaluate the impact of microbial-
73 community size on the abundance of taxa and genes. Quantification of absolute
74 abundance, i.e. copy numbers of functional genes or their transcripts, is vital for
75 evaluating the functional capacities and potentials of microbial communities. The
76 prediction of N-cycling processes can be improved more using information of
77 functional-gene abundance than microbial diversity (Graham et al., 2016).

78 Quantitative PCR (qPCR) is the most adopted method to measure the copy number of
79 functional genes. For example, an analysis of oceanic nitrification identified the
80 dominant role of *amoA* (archaeal ammonia mono-oxygenase alpha subunit for aerobic
81 ammonia oxidation) in an archaeal community by correlating its gene abundance with
82 ammonium concentration (Wuchter et al., 2006). The abundances of *amoA*, *nirK/S*
83 and *nosZ* were used to assess the changes in nitrification and denitrification potential
84 across a vegetation gradient (Petersen et al., 2012).

85 Biogeochemical nutrient cycling is a complex process consisting of numerous
86 steps, each mediated by various functional genes. For example, N cycling is
87 composed of several processes, including N fixation, nitrification, denitrification,
88 ammonification, anaerobic ammonium oxidation, organic N mineralization and
89 assimilatory and dissimilatory N reduction, with over 20 key microbial functional
90 genes involved, including different forms of *nifH*, *amoA/B*, *napA*, *narG*, *nirS/K*,
91 *nosZ*, *hzo* and *hzsA/B* (Kuypers et al., 2018). The comprehensive evaluation of
92 microbial functional potential in CNPS biogeochemical cycling requires obtaining
93 quantitative data for all these genes, which is extremely laborious when using
94 conventional qPCR to process many environmental samples. Besides, current
95 knowledge of functional genes involved in P cycling is more limited than for C, N and
96 S cycling. As far as we know, *ppx* (exopolyphosphatase), *ppk* (polyphosphate kinase)
97 and phytase genes are the most reported P-cycling genes (He et al., 2010). Primers for
98 genes responsible for inorganic-P solubilization, alkaline phosphatase (hydrolysis of
99 phosphoric monoesters) and C-P lyase, are not available, which hinders the
100 quantitative evaluation of microbially mediated P cycling. To address these
101 limitations, this study 1) designed a set of primer pairs targeting functional genes
102 involved in P cycling and 2) developed a high-throughput qPCR-based functional-
103 gene chip detection method, QMEC, for the simultaneous quantification of CNPS-
104 cycling genes and further assessment of microbial potentials in CNPS biogeochemical
105 dynamics and microbial responses to environmental changes. QMEC contains 36
106 reported and 36 novel primer pairs involved in C, N, P and S cycles. The coverage,

107 specificity and efficiency of the designed primers were validated, and the performance
108 of QMEC was evaluated, by analyzing the functional-gene abundance and diversity of
109 soil and sediment samples.

110 **Results**

111 *Specificity and coverage of the designed primers*

112 We designed 36 primer pairs that could potentially amplify genes involved in C, N, P
113 and S cycling (Table S1). These primers annealed to conserved regions of the target
114 genes and produced amplicons averaging 332 bp (ranging from 240 bp for *gdhA* to
115 464 bp for *xylA*).

116 Matched sequences of each gene (434 335 sequences) were phylogenetically
117 analyzed. A total of 42 phyla were represented, with an average of 18 ± 12 phyla and
118 a range of two to 30 phyla per gene. The dominant matched taxa were
119 Alphaproteobacteria (18.4%), Gammaproteobacteria (16.5%) and Actinobacteria
120 (11.4%).

121 *Assessment of QMEC*

122 The HT-qPCR QMEC results were validated by the quantification of the 18 selected
123 genes using conventional real-time qPCR under optimal PCR conditions (Figure 1c).
124 The accuracy for all 18 genes averaged $101.25 \pm 8.27\%$. The abundances of *apsA*,
125 *dsrB*, *hzsB*, *mcrA*, *nirK1*, *nirS1*, *phnK*, *pqqC* and *ureC* quantified by HT-qPCR were
126 very similar (nearly 100% accuracy) to those quantified by conventional qPCR. The
127 average SD and CV of C_T from the replicate samples were 0.22 ± 0.17 and $1.06 \pm$
128 0.74%, respectively (Table S2). SD was largest for *gcd* in Q1 at 0.99%, with a CV of

129 4.90%. Average LOD estimated from the 18 genes (Table S3) was 78.43 ± 27.08
130 copies per well.

131 *Application of QMEC to environmental samples*

132 QMEC was then applied to the soil and sediment samples to illustrate the patterns of
133 microbial functional-gene structures. Nearly all of the 72 genes were detected; only
134 *hzO* and *hzsA* in soil and *ipu* in sediment were not detected (Table S2). The NMDS
135 analysis identified significantly different functional-gene structures between the soil
136 and sediment, where replicates of each sample clustered together and the soil and
137 sediment samples were well separated along the first axis (Figure 2a). Three replicates
138 of each gene were gathered and many genes were well separated from each other
139 (Figure 2b).

140 Bacterial populations (16S rRNA gene) and the absolute quantities of the
141 functional genes were significantly larger in all three soil samples than the sediment
142 samples, but the relative abundances of all functional genes were significantly higher
143 in the sediment samples (Figure S1, $P < 0.05$). The clustering analysis of relative gene
144 abundance found that the soils and sediments were well separated into two clusters
145 (Figure 2c, $P < 0.01$). The abundance of functional genes is summarized in Table 1
146 based on their functions. Nearly all functional genes were significantly more abundant
147 in soil than sediment except for the genes involved in lignin hydrolysis, anaerobic
148 ammonium oxidation and S oxidation (Figure S2, $P < 0.05$), and *mcrA* abundance
149 was also higher in sediment.

150 The relative abundances of the detected genes were analyzed to determine the
151 differences between the soil and sediment samples. Gene abundance was generally
152 lower in the wheat (H1) than the maize (H2) and soybean (H3) soils, and the
153 abundances of most of the functional genes in H2 and H3 soil differed significantly
154 from those in the sediments ($P < 0.05$). The potentials for starch (*amyX*) and pectin
155 (*pgu*) hydrolysis were significantly higher in H2 than H1 and H3 ($P < 0.05$, Figure
156 3a), and the abundances of genes for hemicellulose (*abfA*) and cellulose (*cex*)
157 hydrolysis were highest in H3. The abundance of the gene for C fixation (*accA*) was
158 significantly higher in soil than sediment ($P < 0.05$, Figure 3b). The abundances of
159 smtA, *frdA*, *mct*, *rbcL* and *acsA* were significantly higher in H2 and H3 than
160 H1 ($P < 0.05$). The genes involved in N fixation were 20-fold more abundant in H3
161 than H1 and H2 (Figure 3c). The *ureC* and *napA* abundances were significantly higher
162 in soil than sediment, and the genes involved in anaerobic ammonia oxidation (*hzO*,
163 *hzsA* and *hzsB*) were significantly more abundant in sediment ($P < 0.05$). The
164 potentials of organic N mineralization (*gdhA*) and nitrification (*amoA* and *amoB*)
165 were significantly higher in H2 and H3 than H1 ($P < 0.05$). Gene abundance (*bpp*,
166 *cphy*, *phoD* and *phoX*) for organic-P mineralization was significantly higher in soil
167 than sediment, and genes for solubilizing inorganic P (*gcd* and *pqqC*) were
168 significantly more abundant in H2 and H3 than H1 ($P < 0.05$, Figure 3d). The
169 abundances of some genes for S and methane cycling (*dsrA*, *pmoA* and *pqq-mdh*)
170 were highest in H1, but others (*soxY*, *apsA* and *mxAF*) were significantly higher in H2
171 and H3 ($P < 0.05$ Figure 3e).

172 **Discussion**

173 *Design and assessment of QMEC*

174 This study developed QMEC based on HT-qPCR for comprehensively profiling
175 functional genes involved in C, N, P, S and methane cycling. Many genes are critical
176 to CNPS cycling, but the lack of appropriate primers hinders the quantification of
177 these genes and the further assessment of microbial potential in CNPS cycling. We
178 successfully designed and introduced 36 new primer pairs targeting these genes to
179 supplement the missing genetic tools for analyzing microbially mediated
180 biogeochemical processes. These genes are involved in C hydrolysis, C fixation,
181 methane metabolism and N, P and S cycling. Previous studies of P cycling have
182 focused on a limited set of functional genes, e.g. only three genes were targeted by
183 GeoChip 4.0 (*ppx*, *ppk* and phytase) (Tu et al., 2014), which may be inadequate for
184 the comprehensive evaluation of microbial potentials of organic- or inorganic-P use.
185 We designed and introduced seven new primer pairs to amplify genes involved in P
186 cycling: including two acid phosphatase genes (*bpp*, β-propeller phytase, which is the
187 dominant phytase in water and soil (Lim et al., 2007), and *cphy*, ruminal cysteine
188 phytase (Sebastian and Ammerman, 2009; Ragot et al., 2017)), two alkaline
189 phosphatase genes for phosphate use (*phoD*, which has been identified in 13 bacterial
190 phyla and 71 families in soil (Ragot et al., 2015; Ragot et al., 2017), and *phoX*, which
191 is widely distributed in aquatic systems) and *phnK*, which hydrolyzes
192 organophosphorus compounds (C-P bonds). The organic acid 2-keto-D-gluconic acid

193 has a high ability to solubilize inorganic P (Hwangbo et al., 2003), which requires a
194 pyrroloquinoline quinone (PQQ) co-factor. We thus designed primer pairs for a PQQ-
195 dependent glucose dehydrogenase gene (*gcd*) and its cofactor gene *pqqC* for assessing
196 the potential for inorganic-P solubilization.

197 Primer pairs targeting genes with the same function but covering extended taxa
198 were also designed and introduced in QMEC. Conventional *nirK* and *nirS* primers
199 (*nirK1* and *nirS1* in this study) typically cover denitrifiers from Alpha-, Beta- and
200 Gammaproteobacteria (Katsuyama et al., 2008; Yoshida et al., 2012). The recently
201 reported *nirK* and *nirS* primers (*nirK2*, *nirK3*, *nirS2* and *nirS3* in this study) with a
202 greater diversity of targets, including Actinobacteria, Bacteroidetes, Chloroflexi and
203 Euryarchaeota, were also introduced for comprehensively estimating denitrifying
204 potential (Wei et al., 2015).

205 The specificity and phylogenetic coverage of the primers designed in this study
206 were assessed by BLAST searches against the NCBI database and the analysis of
207 amplicon sequences. These primers covered an average of 18 ± 5 phyla. All newly
208 designed primers had $>70\%$ specificity, and 20 primers had $>80\%$ specificity (Figure
209 1a), suggesting that the designed primers were applicable in functional-gene
210 detection. The 20-30% mismatch rate may have been due to the unified HT-qPCR
211 protocol to ensure the simultaneous detection of multiple functional genes rather than
212 to suboptimal amplifying conditions of primer pairs. The phylogenetic analysis
213 identified distinct taxonomic compositions of functional genes between the soil- and

214 sediment-derived sequences, (Figure 1b), suggesting that the coverage of the designed
215 primers was sufficient to identify various taxa under different environment contexts.

216 Annealing temperature is critical for accurate amplification since diverse
217 annealing temperatures may alter primer-binding kinetics and result in quantification
218 bias, especially in using primers with degenerate positions (Lueders and Friedrich,
219 2003; Gaby and Buckley, 2017). One benefit of HT-qPCR based on SmartChip Reat-
220 time PCR system was one-time amplifying of numerous genes with one PCR
221 protocol, which has been extensively proved to be efficient in antibiotic resistance
222 gene amplification (Chen et al., 2016; Zhu et al., 2017). For further confirm, the
223 accuracy and precision of QMEC were tested by comparing the results for 18
224 randomly selected genes to those using conventional qPCR. The amplification
225 accuracies of QMEC were similar to those for conventional qPCR (Figure 1c, nearly
226 100%), suggesting that the QMEC protocol could simultaneously detect multiple
227 genes. The CV was low ($1.06 \pm 0.74\%$, Table S2), indicating that QMEC
228 amplification was stable and precise. This finding was in accordance with a previous
229 report that WaferGen SmartChip were capable of C_T standard deviations <0.2 or CVs
230 $<3\%$ (Saunders, 2013).

231 *Application of QMEC*

232 QMEC was further applied for profiling functional genes in soils and estuary
233 sediments. Functional-gene structure, gene abundance and gene diversity (Figures 2c,
234 3, S3 & S4) differed significantly between soil and sediment. Interestingly, the

235 relative and absolute abundances of most of the functional genes were reversed
236 between the soil and sediment samples (Figure 2c). For example, the absolute
237 abundance of *acsA* was highest in H2, but its relative abundance was highest in Q3.
238 Absolute gene abundance in pooled replicate samples from one site were well
239 separated from the abundance in samples from different sites (Figure 2a), and
240 absolute gene abundance differed significantly ($P < 0.05$) between soil and sediment
241 and between different soils (Figure 3), suggesting that QMEC could successfully
242 differentiate between functional-gene profiles of different environments. QMEC can
243 quantify absolute gene abundance (copy number), unlike metagenomic sequencing
244 and microarrays (Figure S2). Gene copy number of functional gene reflects the
245 absolute quantities of functional genes in one environmental sample based on qPCR
246 technology. This technology is commonly applied in ecological studies and believed
247 to be precise, high-sensitive, reproducible and easy-to-interpret. Previous studies have
248 found that key biochemical processes were strongly associated with absolute
249 functional-gene copy numbers. For example, the absolute abundances of various N-
250 cycling genes, including *nifH*, *amoA*, *nirS*, *nirK* and *nosZ*, were sensitive to long-
251 term N enrichment in a steppe ecosystem (Zhang et al., 2013). The absolute
252 abundances of *nirS*, *nirK* and other N functional genes were able to account for
253 differences in denitrifying rate, ammonia availability and rate of nitrate transformation
254 in different wastewater-treatment systems (Wang et al., 2015; Wang et al., 2016).
255 Although currently GeoChip and metagenomics are used to acquire microbial
256 functional genes and their structures and to relate them to biogeochemical processes,

257 the gene copy number obtained by qPCR is still of vital importance. For example,
258 GeoChip was applied in a deep-sea hydrothermal vent to determine the differences of
259 metabolic function between samples; however, the qPCR method was also used to
260 quantify the gene copy number of 16S rRNA, mcrA, cbbL and cbbM gene,
261 interpreting the abundance difference between bacterial and archaea community,
262 uncovering the potentials of predominant biogeochemical process (methane
263 metabolism and CO₂ fixation) and making relations with functional community
264 (Wang et al., 2009). A report in Antarctic area used GeoChip to detect the variation of
265 functional genes with different chemical and biogeochemical properties; however, the
266 most highly detected N- and C-cycles genes were also precisely quantified by qPCR
267 to evaluate the functional redundancy (ammonia oxidation, C-fixation, methane
268 oxidation and generation, etc.) among the dominant microbial community members
269 (Yergeau et al., 2007). Similar examples could be found in previous studies (Zhou et
270 al., 2008; Trivedi et al., 2012).

271 The RNA-level measurements could provide information about microbial
272 community dynamics because the RNA could directly relate to the specific function of
273 protein synthesis (Blazewicz et al., 2013). Elser, *et al.* also indicated the RNA,
274 especially ribosomal RNA (rRNA) as rapid protein synthesis, directly or indirectly
275 related with evolutionary processes and consequently ecological dynamics (Elser et
276 al., 2000). However, the RNA abundance may not always be a greater biogeochemical
277 indicator than DNA. For instance, a survey of planktonic *Crenarchaea* in the Pacific
278 Ocean indicated that the gene abundance (copy number) of *amoA*, which expressed

279 ammonia monooxygenase subunit A for aerobic oxidation of ammonia, strongly
280 correlated with ocean depth while the transcript of *amoA* gene (RNA level) showed
281 non-significant relevance (Church et al., 2010). The meta-transcriptome analysis or
282 high-throughput sequencing of RNA genes may comprehensively give a functional
283 profiles of gene expression. For example, a pyrosequencing analysis of microbial
284 community RNA in ocean surface waters, which produces large amounts of cDNA
285 fragments, proved that the genes of key metabolic pathway could be obtained and the
286 abundance of key genes favorably compared to independent qPCR assessments of
287 individual gene expression (Frias-Lopez et al., 2008). However, the pyrosequencing-
288 based technology, no matter in DNA or RNA level, is prone to artifacts where single
289 DNA fragments are duplicately sequenced, which limited its application to relate with
290 biogeochemical potentials. As Gifford, *et al.* have indicated that qPCR approaches can
291 provide absolute numbers with greater sensitivity, the actual limitation is
292 simultaneous detection of a handful of functional genes (Gifford et al., 2011).
293 The idea of QMEC provides an effective solution to both accurate quantification and
294 simultaneous detection. However, as all the degenerate primer designer may face,
295 there existed a non-target amplification of our designed and cited primers. The best
296 way to avoid this problem is to massively expand testing samples and their types,
297 which is far more than easy to achieve in this study. The best usage of QMEC is to
298 make choices of really needed primers rather than using all of them.

299 **Materials and Methods**

300 QMEC is a qPCR-based chip containing 71 microbial CNPS primers and 1
301 bacterial taxa primer, which could parallel quantify 72 DNA samples or 24 samples
302 with 3 replicates in one time.

303 *QMEC primers*

304 QMEC contained a total of 72 primer pairs: 36 designed pairs, 35 published pairs and
305 one pair targeting the bacterial 16S rRNA gene as the reference gene (Table S1). Most
306 published primer pairs originated from previous studies, including those targeting a
307 new functional gene (*pqqC*) from a recent study (Zheng et al., 2017) or genes with
308 extended phylogenies such as *nirK2*, *nirK3*, *nirS2* and *nirS3*. We designed 36 novel
309 primer pairs, in which primers specific for *acsA*, *korA*, *lig*, *mmoX*, *phnK*, *pqq-mdh*,
310 *ppx*, *soxY* and *yedZ* genes were designed based on the conserved regions of amino
311 acid sequences aligned using ClustalW2 (Larkin et al., 2007). Degenerate primers for
312 *abfA*, *accA*, *acsE*, *amyA*, *amyX*, *apu*, *cdaR*, *cdh*, *cex*, *chiA*, *exo-chi*, *frdA*, *gcd*, *gdhA*,
313 *glx*, *ipu*, *manB*, *mct*, *mnp*, *naglu*, *pccA*, *pgu*, *pox*, *ppk*, *sga*, *smtA* and *xylA* were
314 designed using Primer Premier 5.0 (Lalitha, 2000).

315 *Soil and sediment sampling*

316 Samples of surface soils (0-15 cm) from wheat (*Triticum aestivum* L.) (H1), maize
317 (*Zea mays* L.) (H2) and soybean (*Glycine max* L.) (H3) fields were collected after
318 harvest in June 2014 from a long-term cropped site in Hailun, Heilongjiang, China

319 (47°26'N, 126°38'E). The samples were lyophilized, sieved (2.0 mm) and stored at -
320 20 °C for further analysis. Sediment samples (top 15 cm) (Q1, Q2 and Q3) were
321 collected using a grab sampler from the estuary of the Qiantang River in Hangzhou,
322 Zhejiang, China (30°39'N, 120°52'E) during summer 2013 (Figure S3) (Zhu et al.,
323 2017). All samples were transferred to the laboratory on dry ice and stored at -20 °C
324 before analysis.

325 *DNA extraction and quantification*

326 DNA was extracted from the soil and sediment samples using the FastDNA Spin Kit
327 for Soil (MP Biomedicals, Santa Ana, USA) following the manufacturer's
328 instructions. DNA quality was checked by ultraviolet absorbance (ND1000,
329 NanoDrop, Thermo Fisher Scientific, Waltham, USA). DNA concentration was
330 determined using the QuantiFluor dsDNA kit (Promega, Fitchburg, USA). DNA
331 extracts were diluted to 50 ng μL^{-1} with sterilized water and stored at -20 °C before
332 use.

333 *Validation of primers*

334 The specificity and taxonomic coverage of the 36 designed primer pairs were assessed
335 by sequence analysis of the corresponding amplicons from the environmental
336 samples. The DNA extracts from the soil (H1-H3) and sediment (Q1-Q3) samples
337 were equally mixed as soil- and sediment-derived DNA templates (HD and QD,
338 respectively). Each 50- μL PCR reaction contained 25 μL of *Premix Ex Taq*
339 (TAKARA, Dalian, China), 0.2 μM each primer, 1 ng μL^{-1} DNA template and 0.1 mg

340 mL⁻¹ bovine serum albumin. The samples were amplified with an initial denaturation
341 at 95 °C for 5 min and 35 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C
342 for 30 s and extension at 72 °C for 30 s, followed by a final extension at 72 °C for 5
343 min. The PCR products were purified (Universal DNA purification kit, TIANGEN,
344 Beijing, China), quantified (QuantiFluor dsDNA kit, Promega), pooled at equal molar
345 concentrations and sequenced using an Illumina Hiseq2500 platform (Novogen,
346 Tianjin, China). The raw reads were filtered and aligned with bacterial sequences in
347 the Reference Sequence (RefSeq) database (<ftp://ftp.ncbi.nlm.nih.gov/refseq/release>)
348 using Local Blast 2.2.27+ (<ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.2.27/>).
349 The e-value of the alignments was set at 10⁻⁵, and the highest score was accepted. The
350 aligned result with hypothetical protein was excluded. All sequences were submitted
351 to the National Center for Biotechnology Information Sequence Read Archive with
352 the accession numbers SRP107153 and SRP107154.

353 *HT-qPCR*

354 The H1-H3 soil and Q1-Q3 sediment samples were used as examples for QMEC
355 detection quantified by HT-qPCR (SmartChip Real-time PCR system, WaferGen
356 Biosystems, Fremont, USA) using the 16S rRNA gene (F525/R907) as the reference
357 gene (Su et al., 2015). The chip reaction systems were prepared following manual
358 instructions. The qPCR protocol was an initial denaturation at 95 °C for 10 min with
359 40 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s and extension
360 at 72 °C for 30 s. The melting curve was automatically generated by the WaferGen

software. Three replicates of each sample were amplified to analyze the reproducibility of QMEC. Results with multiple melting peaks or amplification efficiencies <80% and >120% were excluded by the SmartChip qPCR software. The results with a threshold cycle (C_T) <31 were used for further analysis. Relative copy number was calculated as described by (Looft et al., 2012). Relative gene abundance was defined as the proportion of the abundance of a functional gene to the abundance of the 16S rRNA gene (equation 1). Absolute gene abundance was calculated based on the absolute 16S rRNA gene copy number quantified by conventional qPCR where *Fun* and *16S* indicate the functional and 16S rRNA genes, respectively (equation 2) (Zhu et al., 2017).

$$371 \text{ Gene relative copy number } GR = (31 - C_T)/(10/3) \quad (1)$$

$$372 \text{ Gene absolute copy number } GA_{Fun} = \frac{GA_{16S} \cdot GR_{Fun}}{GR_{16S}} \quad (2)$$

373 *Assessment of QMEC*

374 We randomly selected 18 genes (the 16S rRNA gene, *apsA*, *cdaR*, *chiA*, *dsrB*, *frdA*,
375 *hzsB*, *manB*, *mcrA*, *nifH*, *nirK1*, *nirS1*, *phnK*, *phoD*, *pqqC*, *rbcL*, *smtA* and *ureC*) for
376 quantification by conventional real-time qPCR (LightCycler 480, Hoffmann-La
377 Roche, Basel, Switzerland) for comparison with the results from the HT-qPCR. The
378 16S rRNA gene was used as the reference gene (Zheng et al., 2017). DNA from the
379 HD and QD samples was prepared (n=3) and used for quantification in the
380 LightCycler 480 System. The reaction mixture consisted of 0.2 μM each primer, 10 ng

381 of template DNA and 1× SYBR premix Ex *Taq*. The protocols for the various genes
382 are listed in Table S4. Standard plasmids of each genes were prepared by amplifying
383 the 18 genes in 50- μ L volumes containing 1 μ L of *Premix Ex Taq* (TAKARA), 0.2 μ M
384 each primer and 1 ng μ L⁻¹ DNA template. The amplicons were then inserted into pMD
385 19-T vectors (TAKARA). The concentration of plasmid DNA was measured using a
386 QuantiFluor dsDNA kit (Promega). qPCR Standard curves were generated using 10-
387 fold serially diluted plasmid DNA with 90-110% amplification efficiency. The copy
388 numbers of the target genes were calculated based on the standard curve and were
389 compared with those generated from the QMEC results.

390 The limit of detection (LOD) of QMEC was determined using 10-fold serial
391 dilutions of the standard plasmids of the 18 genes from 10⁶ to 10¹² copies μ L⁻¹, with at
392 least three replicates for quantification. The reaction system and protocol were the
393 same as for the HT-qPCR described above. Replicates with C_T standard deviations <1
394 were included in the calculation of LOD.

395 *Statistical Analysis*

396 Mean, standard deviation (SD) and coefficient of variation (CV, the ratio of SD to
397 mean) were calculated using Office 365. Correlation and variance (ANOVA) analyses
398 used SPSS Statistics 21.0. Bar plotting and cluster figure generation, nonmetric
399 multidimensional scaling (NMDS), an analysis of similarity (ANOSIM) and a
400 heatmap analysis were conducted using the ggplot2 (Ito and Murphy, 2013), plotly

401 (Sievert et al., 2016), vegan (Oksanen et al., 2013) and pheatmap (Kolde and Kolde,
402 2015) packages of RStudio version 3.4.2, respectively.

403

404 **Compliance and ethics**

405 None declared.

406 **Acknowledgements**

407 This work was financially supported by the Strategic Priority Research Program of the
408 Chinese Academy of Sciences (XDB15020402, XDB15020302) and the Natural
409 Science Foundation of China (41571130063, 41430858). J. P. and J. S. acknowledge
410 the financial support from the European Research Council Synergy grant ERC-SyG-
411 2013-610028 IMBALANCE-P.

412

413 **Figure Legends**

414 **Figure 1. Assessment of QMEC.** (a) Specificity of the 36 designed primer pairs. The
415 specificities are based on alignment with the Reference Sequence (RefSeq) database.
416 (b) Relative abundances of the dominant bacterial phyla in the soil and sediment
417 samples using the 36 pairs. (c) Accuracy of QMEC for the 18 randomly selected
418 genes. The accuracy indicates the ratio of HT-qPCR copy number to conventional
419 qPCR copy number. Relative abundances are based on the proportions of DNA
420 sequences classified at the phylum level.

421 **Figure 2.** Nonmetric multidimensional scaling (NMDS) analyses of (a) all replicates
422 of the six soil and sediment samples and (b) functional genes with three replicates
423 based on their abundance and diversity. (c) Heatmap analysis of absolute functional-
424 gene abundances (left) and relative abundances (right). The plotted values were
425 natural-logarithm transformed.

426 **Figure 3. Analysis of functional-gene differences between soil and sediment.**
427 Three soils with different plant hosts (H1-H3) were compared to sediment samples
428 (Q). (a) Differences in abundance of C-hydrolysis genes. The genes are arranged by
429 the biodegradability of their target substrates, from labile to recalcitrant. Differences
430 in the abundances of genes involved in (b) C fixation, (c) N processes and (d) P
431 cycling. (e) Differences in the abundances of genes involved in methane metabolism
432 and S cycling. Error bars represent standard errors. Different letters indicate

433 significant differences between soil samples at $P < 0.05$. ** indicates significant
434 differences between soil and sediment samples at $P < 0.05$.

435 **References**

- 436 Baker, M., 2011. qPCR: quicker and easier but don't be sloppy. Nat Methods 8, 207-
437 212.
- 438 Bardgett, R.D., van der Putten, W.H., 2014. Belowground biodiversity and ecosystem
439 functioning. Nature 515, 505-511.
- 440 Blazewicz, S.J., Barnard, R.L., Daly, R.A., Firestone, M.K., 2013. Evaluating rRNA
441 as an indicator of microbial activity in environmental communities: limitations
442 and uses. ISME J 7, 2061-2068.
- 443 Chapin, F.S., 3rd, Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M., Reynolds,
444 H.L., Hooper, D.U., Lavorel, S., Sala, O.E., Hobbie, S.E., Mack, M.C., Diaz, S.,
445 2000. Consequences of changing biodiversity. Nature 405, 234-242.
- 446 Chen, Q.L., An, X.L., Li, H., Su, J.Q., Ma, Y., Zhu, Y.G., 2016. Long-term field
447 application of sewage sludge increases the abundance of antibiotic resistance
448 genes in soil. Environ Int 92-93, 1-10.
- 449 Chen, Y., Gelfond, J.A., McManus, L.M., Shireman, P.K., 2009. Reproducibility of
450 quantitative RT-PCR array in miRNA expression profiling and comparison with
451 microarray analysis. BMC Genomics 10, 407.
- 452 Church, M.J., Wai, B., Karl, D.M., DeLong, E.F., 2010. Abundances of crenarchaeal
453 *amoA* genes and transcripts in the Pacific Ocean. Environ Microbiol 12, 679-
454 688.

- 455 De Wilde, B., Lefever, S., Dong, W., Dunne, J., Husain, S., Derveaux, S., Hellemans,
456 J., Vandesompele, J., 2014. Target enrichment using parallel nanoliter
457 quantitative PCR amplification. BMC Genomics 15.
- 458 Deng, Y., He, Z.L., Xiong, J.B., Yu, H., Xu, M.Y., Hobbie, S.E., Reich, P.B., Schadt,
459 C.W., Kent, A., Pendall, E., Wallenstein, M., Zhou, J.Z., 2016. Elevated carbon
460 dioxide accelerates the spatial turnover of soil microbial communities. Glob
461 Chang Biol 22, 957-964.
- 462 Elser, J.J., Sterner, R.W., Gorokhova, E., Fagan, W.F., Markow, T.A., Cotner, J.B.,
463 Harrison, J.F., Hobbie, S.E., Odell, G.M., Weider, L.J., 2000. Biological
464 stoichiometry from genes to ecosystems. Ecol Lett 3, 540-550.
- 465 Feng, W., Liang, J., Hale, L.E., Jung, C.G., Chen, J., Zhou, J., Xu, M., Yuan, M., Wu,
466 L., Bracho, R., Pegoraro, E., Schuur, E.A.G., Luo, Y., 2017. Enhanced
467 decomposition of stable soil organic carbon and microbial catabolic potentials by
468 long-term field warming. Glob Chang Biol 23, 4765-4776.
- 469 Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W.,
470 DeLong, E.F., 2008. Microbial community gene expression in ocean surface
471 waters. Proc Nat Acad Sci USA 105, 3805-3810.
- 472 Gaby, J.C., Buckley, D.H., 2017. The use of degenerate primers in qPCR analysis of
473 functional genes can cause dramatic quantification bias as revealed by
474 investigation of *nifH* primer performance. Microb Ecol 74, 701-708.

475 Gifford, S.M., Sharma, S., Rinta-Kanto, J.M., Moran, M.A., 2011. Quantitative
476 analysis of a deeply sequenced marine microbial metatranscriptome. ISME J 5,
477 461-472.

478 Graham, E.B., Knelman, J.E., Schindlbacher, A., Siciliano, S., Breulmann, M.,
479 Yannarell, A., Beman, J., Abell, G., Philippot, L., Prosser, J., 2016. Microbes as
480 engines of ecosystem function: when does community structure enhance
481 predictions of ecosystem processes? Front Microbiol 7.

482 Hazen, T.C., Dubinsky, E.A., DeSantis, T.Z., Andersen, G.L., Piceno, Y.M., Singh, N.,
483 Jansson, J.K., Probst, A., Borglin, S.E., Fortney, J.L., Stringfellow, W.T., Bill,
484 M., Conrad, M.E., Tom, L.M., Chavarria, K.L., Alusi, T.R., Lamendella, R.,
485 Joyner, D.C., Spier, C., Baelum, J., Auer, M., Zemla, M.L., Chakraborty, R.,
486 Sonnenthal, E.L., D'Haeseleer, P., Holman, H.Y., Osman, S., Lu, Z., Van
487 Nostrand, J.D., Deng, Y., Zhou, J., Mason, O.U., 2010. Deep-sea oil plume
488 enriches indigenous oil-degrading bacteria. Science 330, 204-208.

489 He, Z., Deng, Y., Van Nostrand, J.D., Tu, Q., Xu, M., Hemme, C.L., Li, X., Wu, L.,
490 Gentry, T.J., Yin, Y., Liebich, J., Hazen, T.C., Zhou, J., 2010. GeoChip 3.0 as a
491 high-throughput tool for analyzing microbial community composition, structure
492 and functional activity. ISME J 4, 1167-1179.

493 He, Z., Deng, Y., Zhou, J., 2012. Development of functional gene microarrays for
494 microbial community analysis. Curr Opin Biotech 23, 49-55.

- 495 Hwangbo, H., Park, R.D., Kim, Y.W., Rim, Y.S., Park, K.H., Kim, T.H., Suh, J.S.,
- 496 Kim, K.Y., 2003. 2-Ketogluconic acid production and phosphate solubilization
- 497 by Enterobacter intermedium. *Curr Microbiol* 47, 0087-0092.
- 498 Ito, K., Murphy, D., 2013. Application of ggplot2 to Pharmacometric Graphics. *CPT*
- 499 *Pharmacometrics Syst Pharmacol* 2, e79.
- 500 Katsuyama, C., Kondo, N., Suwa, Y., Yamagishi, T., Itoh, M., Ohte, N., Kimura, H.,
- 501 Nagaosa, K., Kato, K., 2008. Denitrification activity and relevant bacteria
- 502 revealed by nitrite reductase gene fragments in soil of temperate mixed forest.
- 503 *Microb Environ* 23, 337-345.
- 504 Kolde, R., Kolde, M.R., 2015. Package ‘pheatmap’.
- 505 Kuypers, M.M.M., Marchant, H.K., Kartal, B., 2018. The microbial nitrogen-cycling
- 506 network. *Nat Rev Microbiol* 16, 263-276.
- 507 Lalitha, S., 2000. Primer premier 5. Biotech Software & Internet Report: The
- 508 Computer Software Journal for Scient 1, 270-272.
- 509 Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A.,
- 510 Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G.,
- 511 Huttenhower, C., 2013. Predictive functional profiling of microbial communities
- 512 using 16S rRNA marker gene sequences. *Nat Biotechnol* 31, 814-821.
- 513 Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGgettigan, P.A.,
- 514 McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson,
- 515 J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and clustal X version 2.0.
- 516 *Bioinformatics* 23, 2947-2948.

- 517 Liang, Y., Van Nostrand, J.D., Deng, Y., He, Z., Wu, L., Zhang, X., Li, G., Zhou, J.,
- 518 2011. Functional gene diversity of soil microbial communities from five oil-
- 519 contaminated fields in China. *ISME J* 5, 403-413.
- 520 Lim, B.L., Yeung, P., Cheng, C., Hill, J.E., 2007. Distribution and diversity of
- 521 phytate-mineralizing bacteria. *ISME J* 1, 321-330.
- 522 Looft, T., Johnson, T.A., Allen, H.K., Bayles, D.O., Alt, D.P., Stedtfeld, R.D., Sul,
- 523 W.J., Stedtfeld, T.M., Chai, B., Cole, J.R., Hashsham, S.A., Tiedje, J.M., Stanton,
- 524 T.B., 2012. In-feed antibiotic effects on the swine intestinal microbiome. *Proc*
- 525 *Natl Acad Sci USA* 109, 1691-1696.
- 526 Lu, Z., Deng, Y., Van Nostrand, J.D., He, Z., Voordeckers, J., Zhou, A., Lee, Y.J.,
- 527 Mason, O.U., Dubinsky, E.A., Chavarria, K.L., Tom, L.M., Fortney, J.L.,
- 528 Lamendella, R., Jansson, J.K., D'Haeseleer, P., Hazen, T.C., Zhou, J., 2012.
- 529 Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume.
- 530 *ISME J* 6, 451-460.
- 531 Lueders, T., Friedrich, M.W., 2003. Evaluation of PCR amplification bias by terminal
- 532 restriction fragment length polymorphism analysis of small-subunit rRNA and
- 533 mcrA genes by using defined template mixtures of methanogenic pure cultures
- 534 and soil DNA extracts. *Appl Environ Microbiol* 69, 320-326.
- 535 Mamanova, L., Coffey, A.J., Scott, C.E., Kozarewa, I., Turner, E.H., Kumar, A.,
- 536 Howard, E., Shendure, J., Turner, D.J., 2010. Target-enrichment strategies for
- 537 next-generation sequencing. *Nat Methods* 7, 111-118.

538 Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.,
539 Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2013. Package
540 'vegan'. Community ecology package, version 2.

541 Penuelas, J., Poulter, B., Sardans, J., Ciais, P., van der Velde, M., Bopp, L., Boucher,
542 O., Godderis, Y., Hinsinger, P., Llusia, J., Nardin, E., Vicca, S., Obersteiner, M.,
543 Janssens, I.A., 2013. Human-induced nitrogen-phosphorus imbalances alter
544 natural and managed ecosystems across the globe. Nat Commun 4.

545 Petersen, D.G., Blazewicz, S.J., Firestone, M., Herman, D.J., Turetsky, M., Waldrop,
546 M., 2012. Abundance of microbial genes associated with nitrogen cycling as
547 indices of biogeochemical process rates across a vegetation gradient in Alaska.
548 Environ Microbiol 14, 993-1008.

549 Ragot, S.A., Kertesz, M.A., Bunemann, E.K., 2015. *phoD* alkaline phosphatase gene
550 diversity in soil. Appl Environ Microbiol 81, 7281-7289.

551 Ragot, S.A., Kertesz, M.A., Meszaros, E., Frossard, E., Bunemann, E.K., 2017. Soil
552 *phoD* and *phoX* alkaline phosphatase gene diversity responds to multiple
553 environmental factors. FEMS Microbiol Ecol 93.

554 Saunders, N.A., 2013. Real-time PCR Arrays. Real-time PCR: Advanced
555 Technologies and Applications, 199.

556 Schmidt, T.M., DeLong, E.F., Pace, N.R., 1991. Analysis of a marine picoplankton
557 community by 16S rRNA gene cloning and sequencing. J Bacteriol 173, 4371-
558 4378.

- 559 Sebastian, M., Ammerman, J.W., 2009. The alkaline phosphatase PhoX is more
560 widely distributed in marine bacteria than the classical PhoA. ISME J 3, 563.
- 561 Sievert, C., Parmer, C., Hocking, T., Chamberlain, S., Ram, K., Corvellec, M.,
562 Despouy, P., 2016. plotly: Create interactive web graphics via Plotly's JavaScript
563 graphing library [Software].
- 564 Stevenson, F.J., Cole, M.A., 1999. Cycles of soils: carbon, nitrogen, phosphorus,
565 sulfur, micronutrients. John Wiley & Sons.
- 566 Su, J.Q., Wei, B., Ou-Yang, W.Y., Huang, F.Y., Zhao, Y., Xu, H.J., Zhu, Y.G., 2015.
567 Antibiotic resistome and its association with bacterial communities during
568 sewage sludge composting. Environ Sci Technol 49, 7356-7363.
- 569 Trivedi, P., He, Z., Van Nostrand, J.D., Albrigo, G., Zhou, J., Wang, N., 2012.
570 Huanglongbing alters the structure and functional diversity of microbial
571 communities associated with citrus rhizosphere. ISME J 6, 363-383.
- 572 Tu, Q., Yu, H., He, Z., Deng, Y., Wu, L., Van Nostrand, J.D., Zhou, A., Voordeckers,
573 J., Lee, Y.J., Qin, Y., Hemme, C.L., Shi, Z., Xue, K., Yuan, T., Wang, A., Zhou,
574 J., 2014. GeoChip 4: a functional gene-array-based high-throughput
575 environmental technology for microbial community analysis. Mol Ecol Resour
576 14, 914-928.
- 577 van der Heijden, M.G.A., Bardgett, R.D., van Straalen, N.M., 2008. The unseen
578 majority: soil microbes as drivers of plant diversity and productivity in terrestrial
579 ecosystems. Ecol Lett 11, 296-310.

- 580 Vanwonterghem, I., Jensen, P.D., Ho, D.P., Batstone, D.J., Tyson, G.W., 2014.
- 581 Linking microbial community structure, interactions and function in anaerobic
582 digesters using new molecular techniques. Curr Opin Biotechnol 27, 55-64.
- 583 Wang, F., Zhou, H., Meng, J., Peng, X., Jiang, L., Sun, P., Zhang, C., Van Nostrand,
584 J.D., Deng, Y., He, Z., Wu, L., Zhou, J., Xiao, X., 2009. GeoChip-based analysis
585 of metabolic diversity of microbial communities at the Juan de Fuca Ridge
586 hydrothermal vent. Proc Natl Acad Sci USA 106, 4840-4845.
- 587 Wang, H., Ji, G., Bai, X., He, C., 2015. Assessing nitrogen transformation processes
588 in a trickling filter under hydraulic loading rate constraints using nitrogen
589 functional gene abundances. Bioresource Technol 177, 217-223.
- 590 Wang, L., Zhang, Y., Luo, X., Zhang, J., Zheng, Z., 2016. Effects of earthworms and
591 substrate on diversity and abundance of denitrifying genes (*nirS* and *nirK*) and
592 denitrifying rate during rural domestic wastewater treatment. Bioresource
593 Technol 212, 174-181.
- 594 Wei, W., Isobe, K., Nishizawa, T., Zhu, L., Shiratori, Y., Ohte, N., Koba, K., Otsuka,
595 S., Senoo, K., 2015. Higher diversity and abundance of denitrifying
596 microorganisms in environments than considered previously. ISME J 9, 1954-
597 1965.
- 598 Weinstock, G.M., 2012. Genomic approaches to studying the human microbiota.
599 Nature 489, 250-256.
- 600 Wuchter, C., Abbas, B., Coolen, M.J., Herfort, L., van Bleijswijk, J., Timmers, P.,
601 Strous, M., Teira, E., Herndl, G.J., Middelburg, J.J., Schouten, S., Sinninghe

- 602 Damste, J.S., 2006. Archaeal nitrification in the ocean. Proc Natl Acad Sci USA
- 603 103, 12317-12322.
- 604 Yergeau, E., Kang, S., He, Z., Zhou, J., Kowalchuk, G.A., 2007. Functional
- 605 microarray analysis of nitrogen and carbon cycling genes across an Antarctic
- 606 latitudinal transect. ISME J 1, 163-179.
- 607 Yoshida, M., Ishii, S., Fujii, D., Otsuka, S., Senoo, K., 2012. Identification of active
- 608 denitrifiers in rice paddy soil by DNA-and RNA-based analyses. Microb Environ
- 609 27, 456-461.
- 610 Zarraonaindia, I., Smith, D.P., Gilbert, J.A., 2013. Beyond the genome: community-
- 611 level analysis of the microbial world. Biol & Philos 28, 261-282.
- 612 Zhang, X., Liu, W., Schloter, M., Zhang, G., Chen, Q., Huang, J., Li, L., Elser, J.J.,
- 613 Han, X., 2013. Response of the abundance of key soil microbial nitrogen-cycling
- 614 genes to multi-factorial global changes. PLoS One 8, e76500.
- 615 Zheng, B.X., Hao, X.L., Ding, K., Zhou, G.W., Chen, Q.L., Zhang, J.B., Zhu, Y.G.,
- 616 2017. Long-term nitrogen fertilization decreased the abundance of inorganic
- 617 phosphate solubilizing bacteria in an alkaline soil. Sci Rep-UK 7.
- 618 Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S.G., Alvarez-Cohen, L., 2015. High-
- 619 throughput metagenomic technologies for complex microbial community
- 620 analysis: open and closed formats. MBio 6.
- 621 Zhou, J., Kang, S., Schadt, C.W., Garten, C.T., Jr., 2008. Spatial scaling of functional
- 622 gene diversity across various microbial taxa. Proc Natl Acad Sci USA 105, 7768-
- 623 7773.

- 624 Zhu, Y.G., Zhao, Y., Li, B., Huang, C.L., Zhang, S.Y., Yu, S., Chen, Y.S., Zhang, T.,
- 625 Gillings, M.R., Su, J.Q., 2017. Continental-scale pollution of estuaries with
- 626 antibiotic resistance genes. *Nat Microbiol* 2.

627 **Supporting Information**

628 Table S1. QMEC primer pairs.

629 Table S2. Mean, standard deviation (SD) and coefficient of variation (CV) of cycle
630 threshold (CT number) for each gene in the soil and sediment samples.

631 Table S3. Limit of detection (LoD) for the 18 randomly selected genes.

632 Table S4. Amplification protocols for the 18 randomly selected genes for real-time
633 qPCR.

634 Figure S1. Abundance and relative ratio of 16S rDNA and total functional genes in the
635 soil and sediment samples.

636 Figure S2. Abundance of functional genes in the soil and sediment samples.

637 Figure S3. Sediment sampling sites in Qiantang Estuary, Hangzhou Bay.

638

639

Table 1. Summary of functional-gene abundance.

Gene category	Gene number	Gene names	Gene abundance (copies $\times 10^7$ g $^{-1}$ soil)				
			H1	H2	H3	Q1	Q2
<i>C cycling</i>	35						
<i>Carbon hydrolysis</i>	18						
Starch	5	<i>amyA, amyX, apu, sga, ipu</i>	43.93 \pm 15.15 bc	135.3 \pm 38.43 a	13.01 \pm 4.573 d	20.13 \pm 1.388 cd	31.01 \pm 3.056 cd
Hemicellulose	3	<i>abfA, manB, xylA</i>	27.01 \pm 3.920 cd	88.37 \pm 10.64 a	71.09 \pm 10.30 b	9.228 \pm 0.465 e	14.34 \pm 0.696 de
Cellulose	3	<i>cdh, cex, naglu</i>	7.147 \pm 1.278 cd	21.03 \pm 5.546 a	15.73 \pm 2.155 b	3.218 \pm 0.144 d	4.897 \pm 0.413 d
Chitin	2	<i>chiA, exo-chi</i>	3.319 \pm 1.629 bc	8.546 \pm 1.532 a	7.251 \pm 1.937 a	1.852 \pm 0.216 c	2.497 \pm 1.132 c
Pectin	1	<i>pgu</i>	1.640 \pm 0.236 cd	5.341 \pm 0.356 b	7.023 \pm 0.852 a	0.408 \pm 0.102 e	1.032 \pm 0.183 de
Lignin	4	<i>glx, lig, mnp, pox</i>	26.50 \pm 12.49 b	77.38 \pm 13.47 a	17.33 \pm 3.776 bc	11.67 \pm 1.571 c	24.43 \pm 1.941 bc
<i>C fixation</i>	13	<i>accA, aclB, acsA, acsB, acsE, frdA, cdaR, korA, mcrA, mct, pccA, rbcL, smtA</i>	282.3 \pm 76.16 c	757.6 \pm 136.5 a	769.8 \pm 115.4 a	132.5 \pm 7.144 d	224.0 \pm 18.81 cd
<i>Methane metabolism</i>	4						
Methane production	2	<i>mxaF, pqq-mdh</i>	14.64 \pm 5.961 d	43.15 \pm 5.451 b	71.21 \pm 15.72 a	11.01 \pm 1.067 e	14.94 \pm 1.445 d
Methane oxidation	2	<i>mmoX, pmoA</i>	7.704 \pm 2.294 c	25.66 \pm 6.044 a	27.59 \pm 6.419 a	6.135 \pm 0.358 c	9.343 \pm 0.661 c
<i>N cycling</i>	22						
<i>N fixation</i>	1	<i>nifH</i>	0.791 \pm 0.297 b	2.623 \pm 0.292 b	31.70 \pm 4.262 a	1.394 \pm 0.046 b	1.360 \pm 0.093 b
<i>Nitrification</i>	4	<i>amoA1, amoA2, amoB, hao, nxrA</i>	23.53 \pm 9.141 d	105.7 \pm 16.26 a	66.05 \pm 10.75 b	13.70 \pm 3.023 d	26.14 \pm 2.589 cd
<i>Denitrification</i>	9	<i>narG, nirK1, nirK2, nirK3, nirS1, nirS2, nirS3, nosZ1, nosZ2</i>	97.36 \pm 23.98 bc	304.7 \pm 55.08 a	357.2 \pm 73.10 a	60.71 \pm 10.04 c	104.0 \pm 8.829 bc
<i>Ammonification</i>	1	<i>ureC</i>	34.53 \pm 8.571 b	86.40 \pm 21.53 a	81.10 \pm 10.91 a	3.718 \pm 0.284 c	5.058 \pm 0.236 c
<i>Anaerobic ammonium oxidation</i>	3	<i>hzo, hzsA, hzsB</i>	0.039 \pm 0.027 b	0.087 \pm 0.009 b	0.093 \pm 0.013 b	0.080 \pm 0.022 b	0.139 \pm 0.015 b
<i>Assimilatory N reduction</i>	1	<i>nasA</i>	0.035 \pm 0.007 c	0.267 \pm 0.048 b	0.793 \pm 0.106 a	0.032 \pm 0.004 c	0.091 \pm 0.029 c
<i>Dissimilatory N reduction</i>	1	<i>napA</i>	1.857 \pm 0.397 c	5.317 \pm 0.716 a	2.748 \pm 0.370 b	0.547 \pm 0.069 d	0.637 \pm 0.079 d
<i>Organic N mineralization</i>	1	<i>gdhA</i>	14.31 \pm 3.632 bc	44.77 \pm 6.415 a	52.52 \pm 11.62 a	8.270 \pm 2.268 c	14.12 \pm 2.306 bc
<i>P cycling</i>	9						
<i>Organic P mineralization</i>	5	<i>bpp, cphy, phnK, phoD, phoX</i>	67.47 \pm 29.84 bc	181.9 \pm 35.04 a	120.1 \pm 33.62 b	20.04 \pm 1.751 c	33.51 \pm 2.430 c
<i>Inorganic P solubilization</i>	3	<i>gcd, pqqC</i>	17.87 \pm 3.411 b	70.94 \pm 20.22 a	46.66 \pm 37.61 ab	3.891 \pm 0.332 e	6.484 \pm 0.285 d
<i>Inorganic P biosynthesis</i>	1	<i>ppk</i>	0.185 \pm 0.085 c	0.471 \pm 0.119 b	0.782 \pm 0.105 a	0.178 \pm 0.029 c	0.297 \pm 0.063 c
<i>Inorganic P hydrolysis</i>	1	<i>ppx</i>	104.4 \pm 52.52 cd	286.9 \pm 38.77 b	359.4 \pm 60.22 a	67.81 \pm 11.96 d	93.15 \pm 13.82 cd
<i>S cycling</i>	5						
<i>S reduction</i>	3	<i>apsA, dsrA, dsrB</i>	23.75 \pm 8.191 d	80.17 \pm 12.81 b	109.2 \pm 16.01 a	17.28 \pm 1.275 d	24.97 \pm 4.329 d
<i>S oxidation</i>	2	<i>soxY, yedZ</i>	16.39 \pm 2.048 c	54.00 \pm 11.53 a	46.83 \pm 16.40 a	16.73 \pm 3.162 c	27.87 \pm 2.802 bc
<i>Phylogenetic marker</i>	1	16S rRNA gene	312.7 \pm 68.49 b	519.2 \pm 66.00 a	576.8 \pm 77.56 a	60.66 \pm 3.062 c	71.81 \pm 3.857 c
							86.20 \pm 12.14 c

641 Means \pm standard errors. Different letters within a row indicate significant differences at $P < 0.05$.





