Diversity and predictive metabolic pathways of the prokaryotic microbial community along a groundwater salinity gradient of the Pearl River Delta, China

Shilei Sang1, Xiaoying Zhang1, Heng Dai2, Bill X. Hu2, Hao Ou2 & Liwei Sun1

Almost half of the groundwater in the Pearl River Delta (PRD) contains salt water originally derived from paleo-seawater due to the Holocene transgression, which then generates intense physicochemical gradients in the mixing zone between freshwater and saltwater. Although some studies have been conducted on the hydrological and geochemical characteristics of groundwater in the PRD to monitor the intrusion of seawater, little attention has been paid to the microbial community of this particular region. In this study, we implemented a high-throughput sequencing analysis to characterize the microbial communities along a salinity gradient in the PRD aquifer, China. Our results indicated that the microbial community composition varied significantly depending on the salinity of the aquifer. The presence of abundant anaerobic microorganisms of the genera Desulfovibrio and Methanococcus in certain saltwater samples may be responsible for the gas generation of H2S and CH4 in the stratum.

In saline water samples (TDS > 10 g/L), the linear discriminant analysis effect size (LEfSe) analysis found two biomarkers that usually live in marine environments, and the aquifers of the PRD still contained large quantity of saltwater, indicating that the impact of the paleo-seawater has lasted to this day. The predictive metagenomic analysis revealed that the metabolic pathways present in the groundwater samples studied, included the degradation of pesticides and refractory organics (dichlorodiphenyltrichloroethane (DDT), atrazine and polycyclic aromatic hydrocarbons), matter cycling (methane, nitrogen and sulfur), and inorganic ion and mineral metabolites. This study can help enhance our understanding of the composition of the microbial assemblages and its implications as an environmental indicator in an aquifer affected by saltwater intrusion.

In China, groundwater is the most valuable water resource, and supplies water for nearly 70% of the total population and 40% of agricultural irrigation1. However, the rapid socio-economic development in the last few decades has increased the emission of various pollutants and created a shortage of fresh groundwater2, caused by the continuous increase in freshwater demand, the over-exploitation of groundwater and the discharge of domestic and industrial wastewater. Microorganisms are almost the only inhabitants in the groundwater environment3, and they have an important role in cycling nutrients and constituents through their metabolic activity4. Recent research suggests that an aquifer is an ecological system affected by the activity of microbial communities5,6, which is closely relate to nutrient transport, geochemical cycles and the degradation of pollutants7,8.

In general, microorganisms in aquifer ecosystems are sensitive to environmental changes and thus can be a useful indicator to biomonitor pollutants4. It is generally known that groundwater is more stable and cleaner compared to surface water, which occurs because of the filtration process through the aquifer material (i.e., gravel, silt, sand, clay) that leads to natural attenuation processes in the groundwater, such as dissolution and adsorption, oxidation-reduction reactions, and biochemical reactions9. However, the distribution and activity of microbial communities are actively influenced by various aquifer conditions10,11, such as geological setting, water quality,
aquifer materials, and the type of aquifer. Smith et al. compared the composition of microbial assemblages inhabiting unconfined and confined aquifer ecosystem and found distinct variations in composition and metabolism driven by different nutrient inputs and aquifer types. Flynn et al. compared the community of microbes in different aquifer materials and concluded that the aquifer types. Thus, determining the characteristics of the groundwater microbial community, the microbial diversity and their potential functions will help to expand our understanding of a complex groundwater system and its related biogeochemical processes in another dimension.

An increasing number of studies have shown that the microbiome is an important part of the underground dynamic ecosystem and is an indispensable parameter to evaluate subsurface water quality. Unlike surface water, such as the Mississippi River, that can contain as many as 3,107 to 5,498 species, an aquifer is mostly oligotrophic, containing a small number of microorganisms. The presence of E. coli is typically used as an environmental indicator of fecal contamination. Recently, other microbial parameters, including the microbial communities, diversity and interaction with physicochemical characterization, have been used as indicators to evaluate water quality. For instance, to find biological parameters to monitor groundwater quality, Unno et al. analyzed the relationship between taxonomic groups and hydrological chemistry and found that microbiome analysis was an effective tool to monitor groundwater health. As expected, many different contaminants, which can be degraded by diverse microbes, have been found in aquifers. Ye et al. compared different microbial fingerprints with respect to sampling location in submarine groundwater, and selected some potential bacterial groups for bioremediation, such as Comamonas spp. To effectively understand and protect fresh groundwater, it is essential to investigate the composition and function of the microbial community in aquifers.

The plain of the Pearl River Delta (PRD) in China is responsible for the water supply of 50 million people, but the groundwater is contaminated due to high salinity and pollution, as well as continuous seawater intrusion into the Pearl River Estuary. The groundwater salinity is relatively stable in most areas of the PRD, since it primarily originated from paleo-seawater and due to the Holocene transgression. In addition, groundwater in many areas of the southern part of the PRD has been found to contain more than 10 g/L of total dissolved solids (TDS). To improve the groundwater quality and management, numerous research focusing on the hydrological and geochemical characteristics of the PRD aquifers has been conducted. However, few studies have investigated the microbial communities of the groundwater in the PRD. To investigate the diversity and taxonomic composition of the microbial biocenosis in groundwater along a salinity gradient of the PRD, we used next generation sequencing (NGS) technology to establish the relationship between microbial communities and environmental variables from a hydrogeological point of view. Through the investigation of the spatial variation of the groundwater quality, microbial communities and functions in the PRD, the response of the underground microbial communities to the salinity gradient was obtained. Our study can help to monitor seawater intrusion and underground water quality from a biological perspective.

**Results and Discussion**

**Hydrogeochemical and environmental parameters.** Hydrogeochemical parameters of the groundwater samples analyzed are listed in Table 1. According to the Schukaliev classification method, the primary chemical type of saltwater was Cl-Na, while the freshwater contained HCO3-Cl-Ca-Na, HCO3-Cl-Mg-Na or HCO3-Mg-Ca, which indicated that the groundwater underwent salinization derived from the paleo-seawater. The parameter of the TDS was usually used to quantify groundwater salinity based on the following criteria: non-saline/freshwater (F) (0 < TDS < 1 g/L), brackish water (B) (1 < TDS < 10 g/L) and saline water (S) (10 < TDS < 100 g/L). Based on this classification, the 12 monitoring wells were divided into four saline samples (S), four brackish samples (B) and four freshwater samples (F).

**Alpha and beta diversity of the microbial community.** In this study, a total of 448,645 sequences (294 bp average length) were obtained from twelve groundwater samples (including four saline monitoring wells (S) (TDS < 10 g/L), four brackish monitoring wells (B) (1 < TDS < 10 g/L) and four freshwater samples (F) (TDS < 1 g/L)) generated in a single run on a HiSeq 2500 high-throughput sequencing system. For all 12 groundwater samples, the raw sequence reads ranged from 29,355 to 44,283. The lowest sequence number 29,355 was used to subsample all of the samples at the same level. After clustering and alignment, a total of 984 microbial OTUs were obtained based on a 97% threshold with a range of 364–634 OTUs (Table S2), indicating a remarkable variation of the microbial OTU number among the sampling sites. The lowest OTU number was found in the brackish sample Q140 (B), and the highest OTU number was also in the saltwater sample Q141 (S). Among all the 984 OTUs, some OTUs were only found in either the saltwater samples (n = 295, 29.98%) or the freshwater samples (n = 60, 6.10%), while most OTUs (n = 629, 63.92%) were shared by salt and fresh groundwater samples (Supplementary Fig. S1), indicating a potential exchange between the saltwater originating from paleo-seawater and the freshwater recharged from the surface water and precipitation.

The microbial Chao richness estimator varied between 496 and 785 and the Heip's evenness estimator was between 0.02 and 0.08 (Supplementary Table S1). The alpha richness (H′) and evenness, reflected by the Chao and Heip indices, varied along the salinity gradient but not significantly (P > 0.05) (Supplementary Table S2). As demonstrated by the Shannon index, the diversity of the microbial communities did not vary significantly along the salinity gradient of the aquifer in the PRD area, which was consistent with the study in the relatively stable estuary of the Baltic Sea, where no significant change in the Shannon diversity index was found along the salinity gradient. However, an increase in salinity will result in a substantial change of the microbial diversity in a Bay estuary or aquifer where the environment is unstable. In the research which was conducted by Zhou et al., the bacterial Shannon diversity markedly changed from H′ = 3.22 ± 0.28 (autumn and winter) to H′ = 1.31 ± 0.35
Hydrogeochemical and environmental parameters of the 12 groundwater samples.

| Sample       | Q149 | Q141 | Q137 | Q146 | Q138 | Q140 | Q144 | Q143 | Q130 | Q132 | Q124 | Q135 |
|--------------|------|------|------|------|------|------|------|------|------|------|------|------|
| T (°C)       | 26.0 | 25.2 | 24.0 | 25.1 | 24.9 | 23.0 | 25.1 | 25.0 | 25.8 | 24.5 | 25.0 | 22.7 |
| pH           | 7.68 | 6.46 | 6.6  | 7.04 | 6.53 | 5.82 | 7.09 | 6.97 | 6.88 | 7.26 | 7.37 | 7.82 |
| ORP          | -220.1 | 9.3 | -84.4 | 100.7 | -29.6 | -104.1 | 32.5 | 29.1 | 71.2 | -30.6 | 45.5 | -98.2 |
| DO (%)       | 1.4  | 41.2 | 27.1 | 28.8 | 26.4 | 13.3 | 35.2 | 49  | 28.5 | 60.8 | 48.2 | 25  |
| salinity     | 19.48 | 18.41 | 17.9 | 13.27 | 5.97 | 5.5 | 4.1 | 4.0 | 0.81 | 0.46 | 0.27 | 0.15 |
| TN (mg/L)    | 87.27 | 3.09 | 1    | 1.27 | 4.88 | 2.77 | 3.78 | 1.37 | 47.59 | 32.42 | 0.93 | 25.69 |
| EC (ms/cm)   | 31.22 | 29.6 | 28.92 | 22.02 | 10.48 | 9.741 | 7.35 | 7.19 | 1.42 | 0.84 | 0.76 | 0.19 |
| TDS (g/L)    | 15.3 | 14.52 | 14.17 | 10.79 | 5.136 | 4.77 | 3.61 | 3.52 | 0.74 | 0.41 | 0.22 | 0.10 |
| TP (umol/L)  | 0.8  | 0.92 | 0.92 | 1.04 | 0.56 | 0.56 | 0.68 | 0.56 | 2.58 | 0.5 | 0.32 | 0.32 |
| TP (mg/L)    | 121.59 | 227.2 | 58.67 | 65.27 | 23.43 | 25.62 | 36.34 | 48.36 | 8.01 | 15.6 | 5.43 | 0.95 |
| NO⁻ (μg/L)   | 2.33 | 3.21 | 3.87 | 7.25 | 8.28 | 7.62 | 6.44 | 6.74 | 28.43 | 27.69 | 6.74 | 1.37 |
| NO₃⁻ (μg/L)  | 539.06 | 553.59 | 452.42 | 382.72 | 265.67 | 238.07 | 127.48 | 128.48 | 37.29 | 32.65 | 8.56 | 13.29 |
| Ca²⁺ (mg/L)  | 849.48 | 542.7 | 899.82 | 1754.1 | 1311.1 | 1905.3 | 831.53 | 286.51 | 158.94 | 47.68 | 60.62 | 13.97 |
| Mg²⁺ (mg/L)  | 539.06 | 553.59 | 452.42 | 382.72 | 265.67 | 238.07 | 127.48 | 128.48 | 37.29 | 32.65 | 8.56 | 13.29 |
| Na⁺ (mg/L)   | 2456.1 | 2415.7 | 2321.3 | 2060.7 | 1237.8 | 894.49 | 538.25 | 1252.3 | 158.35 | 104.87 | 57.7 | 5.17 |
| Cl⁻ (mg/L)   | 13.44 | 11.37 | 12.36 | 9.45 | 4.46 | 3.55 | 2.49 | 2.09 | 0.23 | 0.09 | 0.02 | 0.006 |
| Hydrogeochemical type | Cl-Na | Cl-Na | Cl-Na | Cl-Ca-Na | Cl-Ca-Na | Cl-Ca-Na | Cl-Ca-Na | Cl-Na | HCO₃-Cl-Ca-Na-Mg | HCO₃-Cl-Na-Mg | HCO₃-Mg-Ca |

Table 1. Hydrogeochemical and environmental parameters of the 12 groundwater samples.
Figure 1. The calculated α-diversity indices of freshwater (F), brackish water (B) and saline water (S): (A) the rarefaction curves and the number of the observed OTUs. (B) Shannon's diversity. (C) Chao richness. (D) Heip's evenness. (E) Sequencing coverage. The numbers in parentheses are the mean and standard deviation. For more specific data of the α-diversity indices, please refer to Supplementary Table S1.

Figure 2. Beta diversity of saltwater and freshwater samples: (A) Hierarchical clustering tree on the OTU level based on the unweighted-unifrac distance. (B) Principal co-ordinates analysis (PCoA) on the OTU level based on the unweighted-unifrac distance.
in a carbonate aquifer subject to seawater intrusion and found that the phylum with the highest abundance was Proteobacteria with a percentage of 58.8–93%. Consistent with our results, a majority of the classified sequences were assigned to Proteobacteria with a percentage of 50.34–86.51%. Other dominant phyla were also similar, including Firmicutes, Bacteroidetes, and Actinobacteria. It is notable that the community structure of the archaea was dominated by Thaumarchaeota, Euryarchaeota, and Woesearchaeota, which were significantly more abundant in saltwater than freshwater. Except for the eight major phyla, another 36 phyla with a relative abundance that was no more than 2% in any sample were classified as “other phyla” and included Acidobacteria, Verrucomicrobia, and Actinobacteria. In addition, the percentage of “other phyla” was between 0.78 and 8.89% with an average of 3.21%.

The composition of the microbial communities at the class level is compared further, as shown in Fig. 3B (Supplementary Table S4). Among the 12 representative classes with a relative abundance >2%, at least one sample, and as many as eight classes, were affiliated with Proteobacteria, Firmicutes, Thaumarchaeota, and Euryarchaeota. In Proteobacteria, theγ-proteobacteria was dominant in most samples with an abundance of 11.01–52.87%, followed by β-, ε-, δ-, and α-proteobacteria. It was obvious that β-proteobacteria was more abundant in freshwater than saltwater, except for the Q140 sample, which coincides with previous reports that β-proteobacteria was dominant in fresh water on a global scale40,41. Simultaneously, the classes Marine Group I and Methanococci, which were affiliated with Thaumarchaeota and Euryarchaeota, respectively, were primarily found in saltwater. In Firmicutes, the dominant class was Bacilli with a relatively higher percentage in saltwater.
than in freshwater. Except for the classes belonging to the phyla Proteobacteria, Firmicutes, Thaumarchaeota and Euryarchaeota, other dominant classes were affiliated with Nitrospira, Woesearchaeota and Bacteroidetes.

The hierarchical heatmap at the genus level (top 50) of the groundwater samples analyzed is presented in Fig. 3C (Supplementary Table S5). The freshwater samples had Acidovorax, Aquabacterium, Denitratisoma and Comamonas with a relatively high abundance, while the saltwater contained some genera related to carbon, nitrogen and sulfur cycles, such as Methanococcus, Candidatus_Nitrosoarchaeum and Desulfovibrio. The presence of abundant heterotrophic anaerobic microorganisms of the genera Desulfovibrio and Methanococcus validated the previous assumption that there was microbial SO$_4^{2-}$ reduction and methanogenesis happening in the PRD aquifer. When hydrogeological surveys were conducted by drilling boreholes, the odor of rotten eggs was commonly smelled, indicating the emission of H$_2$S gas. The presence of methanogens in the aquifer is confirmed by the existence of CH$_4$ gas in the quaternary aquifer identified by the Guangdong Geological Survey. The results indicate that alternative bacterial metabolic processes, such as SO$_4^{2-}$ reduction and methanogenesis, may be responsible for the gas generation of H$_2$S and CH$_4$ in the stratum. In addition, the suitable primers to amplify sulfate-reducing bacteria (SRB) and methanogenic archaea (MA) were also designed and synthesized (Supplementary Fig. S2). Their amplification with the primers used was satisfactory, which verified the results of the high-throughput sequencings and demonstrated that SRB and MA were present in the microbial community. In addition, some types of clones, which correlated with aerobic microbial groups, such as Exiguobacterium, Pseudomonas and Acinetobacter, were revealed in all the groundwater samples with a relatively high yield. In addition, 40 of the top 50 genera were present in all the groundwater samples, which may reflect the ecological coherence of municipal groundwater ecosystems.

In addition, redundancy analysis (RDA) was performed on the microbial communities and the main environmental factors (Fig. 3D–F). At the phylum level, acute angles emerged among the DO, TDS and ORP, indicating a synergistic effect on the microbial community, which suggested an opposite effect of the pH, temperature and total nitrogen (TN). However, at the class and genus level, the DO and ORP both exhibited an opposite effect on the TDS, indicating that the relationship between the environmental variables and microbial communities in groundwater ecosystem is complex. We can therefore conclude that the structure of the microbial communities was determined by the combined effects of several factors. Among all the environmental parameters, the pH had a significant influence on the microbial communities at the phylum, class and genus level ($P < 0.05$), which indicated that the pH would have a more powerful influence on highly abundant species.

**Significant differences in microbial communities.** Biomarker analysis was performed using LEfSe to determine the significant differences in microbial abundance between saltwater and freshwater. As depicted in Supplementary Fig. S3A, a total of 14 microbial clades had significant differences with an LDA threshold value of 4.0 (Supplementary Fig. S3B). Most of the microbes were significantly more abundant in freshwater samples, while only two clades were enriched in saltwater samples. Notably, Alteromonadales (order) and Marinobacter (genus) were enriched in the saline water samples (TDS > 10 g/L) (Supplementary Fig. S3C, $P < 0.05$), known as halophilic or halotolerant microorganisms, which usually live in marine environments. The aquifer in the PRD has been confirmed to have a high salinity, and a previous $^{13}$C analysis concluded that the saltwater in the confined PRD aquifer most likely originated from seawater during the Holocene transgression period. Many studies have demonstrated that the PRD underwent at least two large-scale transgressions during the Holocene period, leading to a long period of interaction between the paleo-seawater and groundwater. Although some researchers concluded that the groundwater of the PRD has been undergoing freshening during deltaic evolution of the PRD, we found that the aquifers of the PRD still contained saline water on a large scale and contained some microbes that live in the marine environment, such as Alteromonadales (order) and Marinobacter (genus). Thus, it can be concluded that the groundwater environment of high salinity originated from seawater, and this impact of the seawater intrusion has lasted to this day. In contrast, the freshwater samples in our study were primarily dominated by Bacteroidetes (phylum) and β-proteobacteria (class). In general, β-proteobacteria is a freshwater environmental indicator species, and many studies have documented that β-proteobacteria predominate in low salinity environments. Simultaneously, it was again confirmed by LEfSe analysis that the abundance of β-proteobacteria was larger in the freshwater samples, which is consistent with the results of the microbial community composition analysis (Fig. 3B). In addition, there were three orders, three families and four genera that were enriched in the freshwater samples, indicating that most of the microbes were more likely to survive in a low salinity environment.

**Microbial functional predictive analysis.** Based on the KEGG orthologous groups, the functional profiles of the microbial communities were predicted but not measured for the 12 groundwater samples. At the metabolic pathway level, the primary functions included the degradation of pesticides and refractory organics (dichlorodiphenyltrichloroethane (DDT), atrazine and polycyclic aromatic hydrocarbons), matter cycling (methane, nitrogen and sulfur), as well as inorganic ion and mineral metabolism (Fig. 4). These underground microorganisms would have some level of capacity for natural pollutants degradation. In addition, this study also revealed that the microbes participated widely in various metabolic pathways, such as carbon, nitrogen and sulfur circulation, which might be attributed to the genera Methanococcus, Candidatus_Nitrosoarchaeum and Desulfovibrio and had an important role in cycling constituents and the maintenance of the environmental balance. In addition, the phylum Proteobacteria, which was involved in nitrogen cycling, was dominant in all the samples. Sulfate-reducing bacteria (SRB), which play an important role in the biodegradation of organic matter, have been detected in groundwater (Fig. 3C). Although microorganisms have some natural purification capacity, naturally occurring ammonium nitrogen has been estimated to be up to 8600 × 10$^{5}$ kg in the aquitard of the PRD. A substantial amount of ammonium was expected to be released from the aquitard. The Guangdong Hydrogeology Team found that the groundwater in PRD contains ammonium at concentrations as...
high as 560 mg/L. The ammonium originated in the overlying organic-rich Holocene-Pleistocene aquitard and entered the aquifer through groundwater transport and diffusion. Due to the low-permeability of the aquitard and the low rate of recharge, the large amount of ammonium may be gradually migrating into the river water and coastal seawater, which may disrupt the ecological balance and cause substantial harm to the environment.

Conclusion

In this study, we integrated methodologies from hydrogeology, ecology and microbiology and applied them to study the physicochemical water parameters and microbial changes induced by saltwater intrusion in groundwater. Most previous studies on microbial ecology across salinity gradients or in hypersaline environments focused on salt lakes, wetlands and marine environments. Only a few studies addressed the diversity of the structure, the distribution of the microbial species, and their relationship based on different salt concentrations with a large spatial scale in groundwater environment. Our study demonstrated that salinity was the primary driving force of the microbial community composition in the groundwater, but the alpha diversity did not completely follow the salinity gradient, and there were many other co-varying factors that could also influence the formations of bacterial and archaeal communities, such as the pH and TDS, as well as the TN. The most abundant phylum in the groundwater of the PRD area was **Proteobacteria**, followed by **Firmicutes** and **Bacteroidetes**, and the community structure of the archaea was dominated by **Thaumarchaeota** and **Euryarchaeota** which were more enriched in some saline water.

Our results indicated that the microbial community composition varied significantly along the aquifer salinity. The genera **Acidovorax**, **Aquabacterium**, **Denitratisoma** and **Comamonas** had a relatively high abundance in freshwater samples, while the saltwater samples contained more genera related to carbon, nitrogen and sulfur cycles, such as **Methanococcus**, **Candidatus_Nitrosoarchaeum** and **Desulfovibrio**. The results of the high-throughput sequencing and the functional genes (**mcrA** and **dsrA**) illustrated that sulfate reducers and methanogens were present in the groundwater, which may be responsible for the gas generation of H$_2$S and CH$_4$ in the stratum. Although some researchers concluded that the groundwater of the PRD has been undergoing freshening during deltaic evolution of the PRD, we found that the aquifers of the PRD still contained large amounts of saltwater and contained some microbes that live in marine environments, such as **Alteromonadales** (order) and **Marinobacter** (genus), indicating that the impact of the seawater intrusion has lasted to this day.

Additionally, the predictive metagenomic analysis showed that the metabolic pathway included the degradation of pesticides and refractory organics (DDT, atrazine and polycyclic aromatic hydrocarbons), matter cycling (methane, nitrogen and sulfur), as well as inorganic ion and mineral metabolism that were present in the PRD aquifer. To our knowledge, this may be the first report about the microbial communities in groundwater along a salinity gradient in the PRD area using a high-throughput sequencing approach. Therefore, this study provides a baseline measurement of the prokaryotic microbial community in groundwater affected by seawater intrusion, laying a foundation for further study on the ecological characteristics of microorganism along a salinity gradient in the aquifer.

Methods

**Site description, sample collection and physicochemical analysis.** Twelve sampling sites were selected to collect samples on May 10, 2017 in the central and southern regions of the PRD area. The twelve sampling sites include four saline monitoring wells (S) with TDS > 10 g/L, four brackish monitoring wells (B) with 1 g/L < TDS < 10 g/L, and four freshwater monitoring wells (F) with TDS < 1 g/L as shown in Fig. 5. The PRD has a subtropical monsoon climate, which is warm and humid all year round with an annual average temperature of 22°C. The annual average precipitation ranges from 1,600 to 2,000 mm, primarily from April to October. According to the hydrogeologic survey of the PRD provided by the Guangdong Geological Survey, there is a deposition consist of very fine-grained silt and clay overlying the terrestrial aquifer. The thickness of the deposition units was between 5 and 20 m. All the collected groundwater samples were below 10 m overlying a thickness of deposition units. Hence, the water type of groundwater samples collected were confined aquifer. The
investigated lithostratigraphic groups all belonged to the quaternary. The detailed information for each well is shown in Supplementary Table S6.

The wells were cleaned by pumping groundwater (three times well volumes) with an electric submersible pump to remove stagnant water before sampling. The physicochemical parameters, such as oxygen content (DO), oxidation-reduction potential (ORP), TDS, electrical conductivity (EC), pH and temperature (T) were measured using the freshly collected groundwater samples using a portable meter (Hanna Instrument, Milan, Italy). All the groundwater samples were collected in sterile 10-L plastic containers for filtration to collect the microbial samples, and an additional 500 mL for each sample was collected in triplicate for physicochemical analyses. All samples were kept at 4 °C during transportation and were refrigerated until they were used. Water samples used for the physicochemical analyses were filtered through a sterile 0.45-μm nitrocellulose membrane filter (Millipore, Sigma., Burlington, MA, USA) using a vacuum system. Physicochemical analysis was performed according to international standards. Cations (K⁺, Ca²⁺, Na⁺ and Mg²⁺) were measured using Inductively Coupled Plasma Optical Emission Spectrometry (ICAP 7600, ICP-OES; Thermo Fisher Scientific, Waltham, MA, USA). Anions (NO₃⁻, Cl⁻ and SO₄²⁻) were measured using ion chromatography (Swiss Wantong type 883 chromatograph; Metrohm Schweiz AG, Zofingen, Switzerland). HCO₃⁻ was measured using acid-based titration analysis (DZ/T 0064.49–93). Total nitrogen (TN) was detected using the alkaline potassium persulfate digestion and UV spectrophotometric method; total phosphorus (TP) was detected using the persulfate digestion and spectrophotometric method; total organic carbon (TOC) was measured using a total carbon analyzer (Elementar, Liquid TOC II; Elementar Analysensysteme GmbH, Langenselbold, Germany).

DNA extraction and PCR amplification. Total DNA was extracted from 5 L of water filtered through a sterile 0.2-μm nitrocellulose membrane filter (Millipore, Sigma., Burlington, MA, USA) using a vacuum system. A MOBIO PowerSoil® DNA Isolation Kit (Qiagen/MO BIO Laboratories Inc., Carlsbad, CA, USA) was used to extract the DNA. The V4 region of the prokaryotic microbial 16S rRNA gene was amplified by PCR using the forward primer 515F (5′-GTGCGCACMGCCGCGTAA-3′) and the reverse primer 806R (5′-GGACTACHVGGGTWTCTAAT-3′) [56]. The functional gene mcrA of the methanogens was amplified by PCR using the forward primer ME1-F (5′-GCMATGCAARATHGGGWATGTG-3′) and the reverse primer ME2-R (5′-TCATKCRTAGTDDGGRTA-3′) [57]. The functional gene dsrA of the sulfate reducers was amplified by PCR using the forward primer dsrA 290-F (5′-CCGGGTGTGGCGCATTYACAYCVVT-3′) and the reverse primer dsrA 660-R (5′-GCGGACAGATGGAGHTCRTCTGGRWA-3′) [58]. PCR reactions were conducted on a BioRad S1000 thermocycler (Bio-Rad Laboratories Inc., Hercules, CA, USA) under the following conditions: 94°C for 5 min; 30 cycles: 94°C for 30 s, 52°C for 30 s and 72°C for 30 s; and 72°C for 10 min. Amplicons were extracted from 1.0% agarose gel using a QIAquick gel extraction kit (Qiagen, Hilden, Germany).
agarose gels and purified using an EZNA Gel Extraction Kit (Omega, Bio-Tek, Norcross, GA, USA) according to the manufacturer's instructions. Libraries were prepared using an NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's instructions, and sequencing was performed on an Illumina HiSeq 2500 system at Magi Gene Technology (Guangzhou, China).

**Sequence processing and statistical analysis.** Paired-end raw reads were demultiplexed, quality-filtered by the Trimmomatic software and merged by the Fast Length Adjustment of SHort reads (FLASH) software using the following criteria: (i) The reads were truncated at any site receiving an average quality score <20 over a 50 bp sliding window; (ii) The primers were exactly matched allowing 2 nucleotide mismatches, and reads containing ambiguous bases were removed, and (iii) The sequences that overlap longer than 10 bp were merged according to their overlap sequence. The quality sequences were assigned into operational taxonomic units (OTUs) with a 97% similarity cutoff in the UPARSE platform59, and chimeric microbial sequences were screened using UCHIME60. The taxonomy of each 16S rRNA gene sequence was analyzed by the RDP Classifier algorithm (http://rdp.cme.msu.edu/) against the Silva 16S rRNA database using a confidence threshold of 70%61.

Rarefaction curves were plotted for each sample based on the OTU information62. Alpha-diversity indices, including community diversity indices (Shannon and Simpson), community richness parameters (Chao and ACE), community evenness indices (Heip), as well as a sequencing depth index (Good's coverage), were performed using Mothur software63. In addition, the beta-diversity of the groundwater samples was determined based on the unweighted-unifrac distance including principal co-ordinate analysis (PCoA) and hierarchical clustering analysis using the Quantitative Insights Into Microbial Ecology (Qime 1.7.0) software64. The associations between physicochemical variables and microbial community structure were determined by Spearman's rank correlation analysis using the SPSS software (IBM Corp., Armonk, NY, USA), and the corresponding heatmap was obtained using the plots package in R65. Redundancy analysis (RDA) was performed to reveal microbe-environment relations with the CANOCO 4.5 software (Biometris, Wageningen, The Netherlands). P-values < 0.05 were considered statistically significant.

The potential biomarkers analysis was performed by the linear discriminant analysis (LDA) effect size (LEfSe)66 using the Kruskal-Wallis test to determine the significant differences between the saltwater and freshwater. LDA was performed to assess the difference of each microbial taxon with an LDA threshold value of 4.0. In addition, the OTU sequences were normalized by phylogenetic investigation of the communities by the reconstruction of unobserved states (PICRUSt) and compared to the KEGG databases for functional predictive analysis67.

**Data Access.** All the raw sequence data were deposited in the NCBI Sequence Read Archive (SRA) under the accession number SRP118856.

**Data Availability**

The data analyzed during this study are included in this published article (and its Supplementary Information files).

**References**

1. Lu, Q., Li, C. & Li, T. Overview of groundwater pollution in China. *Chinese Journal of Industrial hygiene and occupational diseases (In Chinese)* 24, 317–320 (2006).
2. Qi, S. & Qiu, Q. Environmental hazard from saltwater intrusion in the Laizhou Gulf, Shandong Province of China. *Nat Hazards* 56, 563–566 (2011).
3. Goldscheider, N., Hunkeler, D. & Rossi, P. Review: Microbial biocenoses in pristine aquifers and an assessment of investigative methods. *Hydrogeol J* 14, 926–941 (2006).
4. Humphreys, W. F. Hydrogeology and groundwater ecology: Does each inform the other? *Hydrogeol J* 17, 5–21 (2009).
5. Hancock, P. J., Boulton, A. J. & Humphreys, W. F. Aquifers and hyporheic zones: Towards an ecological understanding of groundwater. *Hydrogeol J* 13, 98–111 (2005).
6. Pruss, G. & Schminke, H. K. A global ecosystem - Groundwater is alive! *Chem Unserer Zeit* 38, 340–347 (2004).
7. Griebler, C. & Udert, K. Microbial biodiversity in groundwater ecosystems. *Freshwater Biol* 54, 649–677 (2009).
8. Beltke, C. M., Ding, D., Jin, Q. & Sanford, R. A. Origin of microbiological zoning in groundwater flows. *Geology* 36, 739–742 (2008).
9. Pronk, M., Goldscheider, N. & Zopfi, J. Microbial communities in karst groundwater and their potential use for biomonitoring. *Hydrogeol J* 17, 37–48 (2009).
10. Dash, R. R. *et al.* River bank filtration in Haridwar, India: removal of turbidity, organics and bacteria. *Hydrogeol J* 18, 973–983 (2010).
11. Hahn, H. I. & Fuchs, A. Distribution patterns of groundwater communities across aquifer types in south-western Germany. *Freshwater Biol* 54, 848–860 (2009).
12. Korbel, K. L., Hancock, P. J., Serov, P., Lim, R. P. & Hose, G. C. Groundwater ecosystems vary with land use across a mixed agricultural landscape. *J Environ Qual* 42, 380–390 (2013).
13. Smith, R. J. *et al.* Metagenomic comparison of microbial communities inhabiting confined and unconfined aquifer ecosystems. *Environ Microbiol* 14, 240–253 (2012).
14. Flynn, T. M., Sanford, R. A. & Beltke, C. M. Attached and suspended microbial communities in a pristine confined aquifer. *Water Resour Res* 44 (2008).
15. Korbel, K. L. & Hose, G. C. A tiered framework for assessing groundwater ecosystem health. *Hydrobiologia* 661, 329–349 (2011).
16. Zhou, Y., Kellermann, C. & Griebler, C. Spatio-temporal patterns of microbial communities in a hydrologically dynamic pristine aquifer. *Fems Microbiol Ecol* 81, 230–242 (2012).
17. Staley, C. *et al.* Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *J Appl Microbiol* 115, 1147–1158 (2013).
18. Collins, R. *et al.* Best management practices to mitigate faecal contamination by livestock in New Zealand waters. *New Zealand J Agr Res* 50, 267–278 (2007).
19. Sirisena, K. A., Daughney, C. J., Moreau-Fournier, M., Ryan, P. G. & Chambers, G. K. National survey of molecular bacterial diversity of New Zealand groundwater: relationships between biodiversity, groundwater chemistry and aquifer characteristics. *Fems Microbiol Ecol* 86, 490–504 (2013).
20. Unno, T. *et al.* Influence of seawater intrusion on microbial communities in groundwater. *Sci Total Environ* 532, 337–343 (2015).
21. Ye, Q., Liu, J., Du, J. & Zhang, J. Bacterial Diversity in Submarine Groundwater along the Coasts of the Yellow Sea. *Front Microb* 6, 1519 (2015).
22. Zhi, F. et al. Hydrochemical evidence of salt water formation in the Pearl River Delta. Hydrogeology and Engineering Geology (In Chinese) 42, 133–139 (2015).
23. Liu, K., Jiao, J. J. & Gu, J. Investigation on bacterial community and diversity in the multilayer aquifer-aquitard system of the Pearl River Delta, China. *Ecotoxicology* 23, 2041–2052 (2014).
24. Lu, F., Ni, H., Liu & Zeng, E. Y. Occurrence of nutrients in riverine runoff of the Pearl River Delta, South China. *J Hydrol* 376, 107–115 (2009).
25. Chen, S., Fang, L., Zhang, L. & Huang, W. Remote sensing of turbidity in seawater intrusion reaches of Pearl River Estuary—A case study in Modaomen water way, China. *Estuarine, Coastal and Shelf Science* 82, 119–127 (2009).
26. Wang, Y. & Jiao, J. J. Origin of groundwater salinity and hydrogeochemical processes in the confined Quaternary aquifer of the Pearl River Delta, China. *J Hydrog* 438, 112–124 (2012).
27. Jiao, J. J. et al. Anomally High Ammonium of Natural Origin in a Coastal Aquifer–Aquitard System in the Pearl River Delta, China. *Environ Sci Technol* 44, 7470–7475 (2010).
28. Zong, Y., Huang, G., Switzer, A. D., Yu, F. & Yim, W. An evolutionary model for the Holocene formation of the Pearl River delta, China. *The Holocene* 19, 129–142 (2009).
29. Huang, Z., Li, P., Zhang, Z., Li, K. & Qiao, P. Zhuijiang (Pearl) Delta. *Guangzhou: General Scientific Press* 56, 117 (1982).
30. Wang, Y., Jiao, J. J., Cherry, J. A. & Lee, C. M. Contribution of the aquitard to the regional groundwater hydrochemistry of the underlying confined aquifer in the Pearl River Delta, China. *Sci Total Environ* 461, 663–671 (2013).
31. Farid, I., Treladi, R., Zouari, K., Abid, K. & Ayachi, M. Hydrogeochemical processes affecting groundwater in an irrigated land in Central Tunisia. *Environ Earth Sci* 68, 1215–1231 (2013).
32. Renquan, Z., Xing, L., Menggui, J., Li W. & Qingchun, Y. (Beijing:Geological Publishing House, 2017).
33. Freeze, R. A. & Cherry, J. A. 1979. Groundwater. Prentice—Hall, Inc Englewood cliffs, New Jersey *604*, 21–53 (1974).
34. Herlemann, D. P. R. et al. Distribution and biogeochemical importance of bacterial communities along a salinity gradient in a carbonate aquifer system in Modaomen water way, China. *Geomicrobiol J* 33, 273–291 (2013).
35. Thomas, F., Hehemann, J., Rebuffet, E., Czajek, M. & Michel, G. Environmental and gut Bacteroidetes: the food connection. *Front Microbiol* 2 (2011).
36. Lefort, T. & Gasol, J. M. Global-scale distributions of marine surface bacterioplankton groups along gradients of salinity, temperature, and chlorophyll: a meta-analysis of fluorescence in situ hybridization studies. *Aquat Microb Ecol* 70, 111–113 (2013).
37. Hery, M. et al. Diversity and geochemical structuring of bacterial communities along a salinity gradient in a carbonate aquifer subject to seawater intrusion. *Fems Microbiol Ecol* 90, 922–934 (2014).
38. Filmaid, K. A. & Shen, A. Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Curr Opin Microbiol* 24, 88–95 (2015).
39. Thomas, F., Hehemann, J., Rebuffet, E., Czajek, M. & Michel, G. Environmental and gut Bacteroidetes: the food connection. *Front Microbiol* 2 (2011).
40. Lefort, T. & Gasol, J. M. Global-scale distributions of marine surface bacterioplankton groups along gradients of salinity, temperature, and chlorophyll: a meta-analysis of fluorescence in situ hybridization studies. *Aquat Microb Ecol* 70, 111–113 (2013).
41. Logares, R. et al. Infrequent marine-freshwater transitions in the microbial world. *Trends Microb* 17, 414–422 (2009).
42. Team, G. G. H. Regional Hydrogeological Survey Report (1981).
43. Gregory, S. P., Maurice, L. D., West, J. M. & Goody, D. C. Microbial communities in UK aquifers: current understanding and future research needs. *Q J Eng Geol Hydro* 47, 145–157 (2014).
44. Fuhrman, J. A. et al. A latitudinal diversity gradient in planktonic marine bacteria. *P Natl Acad Sci USA* 105, 7774–7778 (2008).
45. Beyer, A. et al. Aquifer community structure in dependence of lithostratigraphy in groundwater reservoirs. *Environ Sci Pollat R* 22, 19342–19351 (2015).
46. Jiang, H., Deng, S., Huang, Q., Dong, H. & Yu, B. Response of Aerobic Anoxicogenic Phototrophic Bacterial Diversity to Environment Conditions in Saline Lakes and Daotang River on the Tibetan Plateau, NW China. *Geomicrobiol J* 27, 400–408 (2010).
47. Zhang, Y., Jiao, N., Cottrell, M. T. & Kirchman, D. L. Contribution of major bacterial groups to bacterial biomass production along a salinity gradient in the South China Sea. *Aquat Microb Ecol* 43, 233–241 (2006).
48. Team, G. G. H. Regional Hydrogeological Survey Report (1981).
49. Gregory, S. P., Maurice, L. D., West, J. M. & Goody, D. C. Microbial communities in UK aquifers: current understanding and future research needs. *Q J Eng Geol Hydro* 47, 145–157 (2014).
50. Beyer, A. et al. Aquifer community structure in dependence of lithostratigraphy in groundwater reservoirs. *Environ Sci Pollat R* 22, 19342–19351 (2015).
51. Zhou, Q. et al. Quality and seasonal variation of rainwater harvested from concrete, asphalt, ceramic tile and green roofs in Chongqing, China. *J Environ Manage* 132, 178–187 (2014).
52. Tozichi, J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl Environ Microb* 79, 5112–5120 (2013).
53. Zhang, Q. et al. Quality and seasonal variation of rainwater harvested from concrete, asphalt, ceramic tile and green roofs in Chongqing, China. *J Environ Manage* 132, 178–187 (2014).
54. Tozichi, J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl Environ Microb* 79, 5112–5120 (2013).
55. Roesch, C., Mergel, A. & Bothe, H. Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Appl Environ Microb* 68, 3818–3829 (2002).
56. Perrenaut, E. M. et al. Response and recovery of sulfate-reducing biochemical reactions from aerobic stress events. *Proceedings America Society Of Mining And Reclamation*, 1041–1062 (2009).
57. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10, 996 (2013).
58. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200 (2011).
59. Quast, C., et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41, D590–D596 (2013).
60. Crist, T. O. & Veech, J. A. Additive partitioning of rarefaction curves and species-area relationships: unifying α-, β- and γ-diversity with sample size and habitat area. *Ecol Lett* 9, 923–932 (2006).
61. Schloss, P. D. et al. Introducing mothur: platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microb* 75, 7537–7541 (2009).
62. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7, 335 (2010).
63. Team, R. C. R: A language and environment for statistical computing (2013).
64. Segata, N. et al. Metagenomic biomarker discovery and explanation. *Genome Biol* 12, R60 (2011).
Acknowledgements
The first, second and fourth authors were supported by the key research and development project of the Ministry of Science and Technology (Grant No. 2016YFC0402805). The third author was supported by the National Natural Science Foundation of China (Grant No. 41807182). This work was also supported by the Guangzhou water technology project (Grant No. 400117017). We also thank Dr. Jin Zhang, Dr. Yanru Hao, Mr. Guangxiong Qin and Ms. Liang Zhang for their help.

Author Contributions
Wrote the paper: S.L.S., H.D. Conceived and designed the experiments: B.X.H. Sampling and performing the analytical work: S.L.S., H.O., L.W.S., X.Y.Z.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-35350-2.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2018