

# Metagenomics Reveals Seasonality of Human Pathogenic Bacteria from Hand-Dug Well Water in the Cuvelai Etosha Basin of Namibia

Billy McBenedict<sup>1,2\*</sup>, Wilhelmina Ndapunikwa Hauwanga<sup>1,3</sup>, Heike Wanke<sup>4</sup>, Percy Maruwa Chimwamurombe<sup>5</sup> and Bernard Mudenda Hang'ombe<sup>2</sup>

<sup>1</sup>School of Public Health, Department of Environmental and Occupational Health, University of Namibia, P.O. Box 2654, Eliander Mwatale Street, Oshakati West, Namibia. <sup>2</sup>Microbiology Unit, School of Veterinary Medicine, University of Zambia, P.O. Box 32379, Lusaka, Zambia. <sup>3</sup>Department of Natural Resource Management, Namibia University of Science and Technology, P.O. Bag 13388, 13 Storch Street, Windhoek, Namibia. <sup>4</sup>Geology Department, University of Namibia, 340 Mandume Ndemufayo Avenue, Windhoek, Namibia. <sup>5</sup>Department of Natural and Applied Sciences, Namibia University of Science and Technology, P.O. Bag 13388, 13 Storch Street, Windhoek, Namibia.

\*Correspondence: [billymcbenedict@yahoo.com](mailto:billymcbenedict@yahoo.com); +264813667923

(Received: 15 April 2019; accepted: 10 June 2019)

**Abbreviations:** CEB: Cuvelai Etosha Basin; SASSCAL: Southern African Science Service Centre for Climate Change and Adaptive Land Management; WHO: World Health Organisation; pH: potential of hydrogen; OTU: Operational taxonomic units; H': Shannon-Wiener diversity; D: Simpson diversity; R: Richness; E: Evenness; NMS: Nonmetric Multidimensional Scaling; UTI: urinary tract infections; PCR: Polymerase Chain Reaction.

**Citation:** Billy McBenedict, Wilhelmina Ndapunikwa Hauwanga, Heike Wanke, Percy Maruwa Chimwamurombe and Bernard Mudenda Hang'ombe, Metagenomics Reveals Seasonality of Human Pathogenic Bacteria from Hand-Dug Well Water in the Cuvelai Etosha Basin of Namibia, *J Pure Appl Microbiol.*, 2019; **13**(2): 865-877. doi: 10.22207/JPAM.13.2.22

© The Author(s) 2019. **Open Access.** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License which permits unrestricted use, sharing, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Abstract

Namibia is a large country with a population slightly above 2 million of which 73% live in rural areas. It is the driest country in sub-Saharan Africa with unpredictable rainfall patterns which occur between October and May. The Cuvelai Etosha Basin is a rural setting in which most people depend on groundwater to circumvent water scarcity by the construction of hand-dug wells. Communities consume hand-dug well water throughout the year without; knowledge of its safety, and treatment prior to use. Hence, this study aimed at evaluating the seasonal safety and quality of hand-dug well water for human consumption. A 16S rRNA metagenomics study was conducted to explore the influence of season on the human pathogenic bacterial species richness, diversity, evenness and abundance in the hand-dug wells. Bacterial pathogens were identified and analysis revealed no significant difference in species richness and diversity, and a significant difference in species evenness and abundance between the wet and dry seasons. Hence, seasonality had an influence on human bacterial pathogens evenness and abundance but had no influence on species richness and diversity. In addition, the water in hand-dug wells of the Cuvelai Etosha Basin is not safe for consumption and domestic use unless sanitized. It is recommended that efforts should be made to treat the water prior to use, and hand-dug wells should be properly constructed to reduce contamination.

**Keywords:** Cuvelai Etosha Basin, Diversity, Evenness, Hand-dug well, Human bacterial pathogens, Metagenomics, Richness, Water.

## INTRODUCTION

Namibia is a large country found in southern Africa with a size of 823, 680 km<sup>2</sup>.<sup>1</sup> The country has a population slightly above 2 million of which 73% live in rural areas and the rest in urban settings. Namibia is one of the driest countries in southern Africa with unpredictable rainfall patterns which uniquely occur between October and May<sup>2</sup>. Due to high temperatures, Namibia experiences high evaporation rates, leaving a small quantity (2%) that can be used as surface water.<sup>2</sup> Amakali and Swatuk<sup>3</sup> disclosed that about 83% of the rainfall vaporizes almost immediately while 14% is lost by evapo-transpiration due to high temperatures, and groundwater is recharged by only 1%. Mendelson *et al.*<sup>1</sup> revealed that Namibia has been characterized by an arid climate for millions of years leading to water scarcity in most parts of the country. In addition, rivers in Namibia are dry most of the time and only flow after heavy rains. However, the Namibian transboundary rivers namely Kunene, Zambezi, Kavango and Orange are perennial although far from main centres of demand. This demands that communities rely on groundwater as a viable source for domestic water.

The Cuvelai Etosha Basin (CEB) is a densely populated area located in central northern Namibia, and consists of four regions namely Oshana, Oshikoto, Ohangwena and Omusati<sup>4</sup>.

This Basin harbours about half of the Namibian population amounting to 1 million<sup>4</sup>. Given the water scarcity problem in the country, this Basin is not excused from the hardships of inadequate water supply. Thus, the CEB rural communities mostly depend on the use of groundwater to circumvent water scarcity by the construction of hand-dug wells. However, ground-water presents another problem in that it is saline in most parts of the basin and the situation is worsened by lack of perennial rivers within the regions<sup>5,6</sup>. The CEB has three distinct hand-dug well forms (Fig.; 1, 2 and 3) that differ according to structure. These structural differences determine whether or not animals have access. There is only one report by McBenedict *et al.*<sup>7</sup> that described these hand-dug wells not to be safe for human consumption. However, McBenedict *et al.*<sup>7</sup> used culture dependent methods which cannot reveal the total bacterial species in these hand-dug wells as opposed to the Metagenomics technique used in this study as a follow up.

Hence, the use of metagenomics approaches circumvents the limitation of culturing media which does not appropriately represent in situ conditions since only a certain fraction (<1%) of bacteria can be cultured leading to a poor understanding of the bacterial communities. Metagenomics is a technique that employs the

sequencing and analysis of the entire microbial community genomes to define and understand the genetic content of the environment in question. Metagenomics is also termed environmental genomics owing to its ability to capture and analyse the total DNA of an environment<sup>8</sup>. Metagenomics is a sequence-based and functional analysis of the entire microbial genomes from a microbial habitat and can reveal an inclusive measure of genetic diversity, species composition, evolution, and ecological functions of respective species in microbial communities<sup>9</sup>.

It is sufficing to state that the WHO documented list of water borne pathogens is not comprehensive due to lack of widespread research on water pathogens using highly specific and effective techniques such as metagenomics. Hence, there is need to conduct more microbial research in water in order to reveal the vast microbial life forms and their interactions. This study comprehensively disclosed the human pathogenic bacteria inhabiting the hand-dug wells in the CEB and adds on to the known pathogens for which water is a mode of transmission, and is the first to report the use of metagenomics in assessing the microbial quality of water in Namibia. Therefore, the main aim of this study was to investigate the influence of season on human pathogenic bacterial species richness, diversity and distribution in hand-dug wells of the CEB of Namibia.

## MATERIALS AND METHODS

### Study sites and sample collections

The study targeted the three different structural types of hand-dug wells (Fig. 1 – 3). Although the CEB in Namibia is formed by Oshikoto, Omusati, Ohangwena and Oshana regions, water samples were collected from Omusati and Ohangwena regions since they possess the three different hand-dug well types. Convenience sampling was employed in this study although it targeted the areas in which hand-dug wells were monitored for chemical water quality by Southern African Science Service Centre for Climate Change and Adaptive Land Management (SASSCAL) task 007: *Improving knowledge and understanding of groundwater flow, water quality and quantity variations, improve methodology of groundwater availability study in the Cuvelai* since 2014. Water

samples were collected following standard water sample collection guidelines stipulated by WHO, University of Namibia and the University of Zambia. A total of 40 water samples were collected in sterile 200 ml bottles from the hand-dug wells, half the total number being from the wet and 20 from the dry season respectively. Water samples were collected from the same hand-dug wells in both seasons. The bottles were lowered into the hand-dug wells for water collection using a rope which was tied to the sterile bottles. The samples were then transported on ice to the University of Namibia for analysis. Prior to transportation, the physical parameters; temperature, electrical conductivity, redox potential and potential of hydrogen (pH) of the water were measured.

### DNA extraction and 16S rRNA gene amplification

Each water sample containing a volume of 200 ml was centrifuged at a speed of 100148 *xg* for 30 minutes in order to concentrate the microorganisms. Each volume was then reduced to 40 ml after centrifugation by discarding a portion of the supernatant. The SEEPREP 12™ kit (Seegene, Rockville, USA) was used to extract DNA after which amplification of the 16S rRNA gene by PCR was performed using universal primer sets 27F (5' AGAGTTTGATCMTGGCTCAG 3') and 1492R (5' TACGGYTACCTTGTTACGACTT 3'). The Bio-Rad thermo-cycler (Hercules, CA) was used with reaction conditions described by McBenedict *et al.*<sup>7</sup> The amplicons were then prepared for next generation sequencing diversity assay using Illumina 16S sequencing at Mr. DNA Next Generation Sequencing provider in Texas, United States of America.

### PCR product preparation and sequencing

The amplified amplicons from above were prepared for sequencing as described by the Illumina TruSeq DNA library preparation protocol. A MiSeq was used for Illumina 16s (20k 2x300bp) sequencing performed at MR. DNA ([www.mrdnalab.com](http://www.mrdnalab.com), Shallowater, TX, USA) according to the manufacturer's guidelines, and sequence data were processed using an exclusive analysis pipeline from MR. DNA.

### 16S rRNA metagenomics data collection and analysis

The metagenomics sequence data obtained from Mr. DNA Next Generation Sequencing provider (Texas, United States of

America) were processed and edited using a proprietary analysis pipeline (www.mrdnalab.com, MR. DNA, Shallowater, TX). The Q25 sequence data derived from the sequencing process were depleted of barcodes and primers, and short sequences less than 150 bp were removed. In addition, sequences with ambiguous base calls, and homopolymer runs exceeding 6 bp were removed. The sequences were then denoised and chimeras also removed. Operational taxonomic units (OTUs) were defined after removal of singleton sequences, clustering at 3% divergence (97% similarity) according to other works<sup>10</sup>. OTUs were then taxonomically classified by performing a BLASTn against a curated GreenGenes, RDP11 (<http://rdp.cme.msu.edu>) and NCBI (www.ncbi.nlm.nih.gov) databases and compiled into each taxonomic level<sup>11</sup>. The data was arranged into counts and percentages based on the number of sequences in each sample.

#### **Construction of a phylogenetic tree**

From the total identified OTU's, selection of cleaned bacterial sequences of those bacteria known to cause disease in humans was performed. A phylogenetic tree was constructed from cleaned sequences using the Maximum Likelihood Method<sup>12</sup>. Only the trees with the highest log likelihood were chosen and percentages of trees in which the associated taxa clustered together were shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. Bootstrap was performed and the consensus trees inferred from 1000 replicates were taken to represent the evolutionary history of the taxa analysed<sup>13</sup>. Branches corresponding to partitions reproduced in less than 70% bootstrap replicates were collapsed and the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) were also indicated next to the branches. The evolutionary analyses were conducted in MEGA 7<sup>13</sup>. A *V. cholerae* strain with accession number *KJ725364.1* was retrieved from the NCBI website and used as the outgroup to root the phylogenetic tree.

#### **Influence of season on the abundance of human bacterial pathogens**

The human bacterial pathogens with their respective counts were identified and subjected to analysis. The distribution of these pathogens between the wet and dry seasons was evaluated by entering the data into SPSS version 24 to; generate the Shapiro-Wilk test and Kolmogorov-Smirnov test values, visually inspect histograms, Normal Q-Q Plots, and calculation of Z scores from Skewness and Kurtosis. An Independent Samples Mann-Whitney U test was used to investigate the influence of season on the abundance of the human bacterial pathogens.

#### **Bacterial species diversity, evenness and richness**

Species diversity is the number of species and abundance of each species present in a particular area while species richness is the number of species present in a particular area<sup>14</sup>. Species evenness is the measure of the relative abundance of different species in a particular area<sup>14</sup>. Species diversity, evenness and richness are fundamental in determining ecosystem health, and in the present study it gave an indication of contamination levels. In each hand-dug well, bacterial species richness was counted, and Shannon-Wiener diversity indices, Simpson's diversity indices and species evenness were calculated using the formulas described in literature<sup>14,15</sup>.

Possible differences in Shannon-Wiener diversity indices and species richness of human pathogenic bacteria between the wet and dry seasons were tested using a Paired sample t-test while differences in Simpson diversity indices and species evenness of human pathogenic bacteria between the wet and dry seasons were tested using the Mann-Whitney U test.

#### **Physical parameters analysis**

The physical parameters (pH, Ec, temperature, redox potential) were entered into PC-ORD version 7 in order to investigate the influence of these parameters on the abundance of human pathogenic bacterial species using Nonmetric multidimensional scaling (NMS) multivariate analysis. NMS ordination procedure was done using Sørensen (Bray-Curtis) distance measure. Number of runs used with real data was 200.

## RESULTS

### Phylogenetic tree

A total of 181 human pathogens were identified and used to generate a phylogenetic tree showing the relationships between the detected human pathogens (Fig. 4). It was disclosed that *L. garvieae*, *L. lactis*, *Lactococcus* sp., *S. lutetiensis*, *S. gordonii*, *S. sanguinis*, *L. pseudomesenteroides*, *L. iners*, *L. paraplantarum*, *L. fermentum*, *L. plantarum*, *E. faecalis*, *B. parabrevis*, *Brevibacillus* sp., *B. coagulans*, *Sporosarcina* spp., *E. aurantiacum*, *Exiguobacterium* sp., *L. massiliensis* and *G. sanguinis* formed a cluster at 77% bootstrap. *C. intestinale*, *C. subterminale*, *C. limosum* and *C. sporogenes* formed a cluster at 81% bootstrap while *A. vaginae*, *O. uli* and *Eggerthella* sp. formed a cluster at 77% bootstrap. *C. falsenii*, *C. mucifaciens*, *C. jeikeium*, *C. tuberculostearicum*, *C. thomssenii*, *C. amycolatum*, *D. papillomatosis*, *W. muralis*, *G. terrae*, *M. septicum*, *M. parascrofulaceum* and *M. ulcerans* formed a cluster at 83% bootstrap, and *Peptostreptococcus* spp. and *P. stomatis* formed a cluster at 99% bootstrap. *D. capnocytophagoides*, *D. gadei*, *Bacteroides* spp., *Dysgonomonas* spp., *Chitinophaga* spp., *Sphingobacterium* spp., *Flavobacterium* spp., *A. shahii*, *Alistipes* spp., *A. finegoldii*, *Empedobacter* sp., *Bergeyella* sp., *E. brevis* and *W. falsenii* formed a cluster at 90% bootstrap. *R. mucosa* and *Roseomonas* spp. formed a cluster at 91% bootstrap while *Brevundimonas* spp., *C. vibrioides*, *B. vesicularis*, *B. diminuta* and *Caulobacter* spp. formed a cluster at 76% bootstrap, and *M. iners* and *M. thiocyanatum* formed a cluster at 72% bootstrap. *P. mendocina*, *P. stutzeri*, *P. putida* and *P. oryzihabitans* formed a cluster at 83% bootstrap while *Dokdonella* sp. and *D. ginsengisoli* formed a cluster at 86% bootstrap, and *L. pneumophila*, *L. lytica* and *L. sainthelensi* formed a cluster at 99% bootstrap. *Acinetobacter* spp., *A. junii*, *A. schindleri*, *A. johnsonii*, *A. septicus*, *A. radioresistens*, *A. calcoaceticus* and *A. lwoffii* formed a cluster at 82% bootstrap, and *J. lividum* and *M. timonae* formed a cluster at 82% bootstrap. *Achromo-bacter* spp., *A. xylooxidans* and *B. petrii* formed a cluster at 83% bootstrap, and *Comamonas* sp., *D. tsuruhatensis*, *A. delafieldii*, *A. temperans*, *C. kerstersii*, *A. facilis*, *Acidovorax* spp. and *C. testosteroni* formed a cluster at 86% bootstrap. The rest of the detected species (Fig.

4) did not form any clusters. The *V. cholerae* strain with accession number *KJ725364.1* was retrieved from the NCBI website and used as the outgroup to root the human pathogen's tree.

### Influence of season on the abundance of human bacterial pathogens

From the pathogens detected, the notable pathogens of public health concern were; *H. parainfluenzae* ( $P < 0.05$ ), *L. lytica* ( $P < 0.05$ ), *L. sainthelensi* ( $P < 0.05$ ), *P. mendocina* ( $P < 0.05$ ), *P. oryzihabitans* ( $P < 0.05$ ), *P. putida* ( $P < 0.05$ ), *P. stutzeri* ( $P < 0.05$ ) and *S. sonnei* ( $P < 0.05$ ) which showed a significant difference in abundance between the wet and dry seasons (Table 1). *L. sainthelensi* ( $P < 0.05$ ), *P. oryzihabitans* ( $P < 0.05$ ), *P. putida* ( $P < 0.05$ ), *P. stutzeri* ( $P < 0.05$ ) and *S. sonnei* ( $P < 0.05$ ) were more abundant in wet season compared to the dry season. *H. parainfluenzae* ( $P < 0.05$ ), *L. lytica* ( $P < 0.05$ ) and *P. mendocina* ( $P < 0.05$ ) were more abundant in the dry season compared to the wet season. However, there was no significant difference in the abundance of *Citrobacter* spp. ( $P > 0.05$ ), *L. jordanis* ( $P > 0.05$ ), *L. pneumophila* ( $P > 0.05$ ) and *V. cholera* ( $P > 0.05$ ) between the wet and dry seasons (Table 1).

### Bacterial species diversity, evenness and richness

The Kolmogorov-Smirnov test revealed that Shannon-Wiener diversity indices and species richness data were normally distributed ( $P > 0.05$ ). A Paired sample t-test revealed that there was no significant difference between Shannon-Wiener diversity ( $H'$ ) indices, and species richness of human pathogenic bacteria in the wet and dry seasons (Table 2). Simpson diversity indices and species evenness data were not normally distributed ( $P < 0.05$ ). According to the Wilcoxon rank test, there was no significant difference in Simpson diversity ( $D$ ) data across the wet and dry seasons. However, species evenness data differed significantly between the wet and dry seasons.

### Physical parameters

The pH values ranged from 7.18 to 8.31 in the wet season and 5.68 to 8.34 in the dry season, and temperature values ranged from 13.2° C to 26.3° C in the wet season and 20.5° C to 34.6° C in the dry season. The electrical conductivity values ranged from 54.5  $\mu\text{S}/\text{cm}$  to 7240  $\mu\text{S}/\text{cm}$  in the wet season and 62.4  $\mu\text{S}/\text{cm}$  to 14420  $\mu\text{S}/\text{cm}$  in the dry season, and the redox potential ranged from -380.1 Mv to 160.9 Mv in the wet season

**Table 1.** Wilcoxon rank test performed to determine the influence of season on the abundance of the detected human bacterial pathogens

Bacterial species	Z value	P – Value	Bacterial species	Z value	P – Value
<i>Achromobacter spp.</i>	-1.94	0.052	<i>Candidatus</i>	-0.365	0.715
<i>Achromobacter xylooxidans</i>	-3.92	0.000	<i>neoehrlichia mikurensis</i>		
<i>Acidaminococcus spp.</i>	-1.60	0.109	<i>Catabacter</i>	-1.00	0.317
<i>Acidovorax delafieldii</i>	-1.01	0.313	<i>hongkongensis</i>		
<i>Acidovorax facilis</i>	-3.83	0.000	<i>Caulobacter spp.</i>	-2.67	0.008
<i>Acidovorax spp.</i>	-0.112	0.911	<i>Caulobacter vibrioides</i>	-2.53	0.011
<i>Acidovorax temperans</i>	-3.30	0.001	<i>Cellulomonas spp.</i>	-3.42	0.001
<i>Acinetobacter</i>	-0.019	0.985	<i>Chitinophaga spp.</i>	-1.16	0.245
<i>calcoaceticus</i>			<i>Citrobacter spp.</i>	-1.00	0.317
<i>Acinetobacter johnsonii</i>	-3.62	0.000	<i>Clostridium ghonii</i>	-3.54	0.000
<i>Acinetobacter junii</i>	-3.92	0.000	<i>Clostridium intestinale</i>	-1.63	0.102
<i>Acinetobacter lwoffii</i>	-3.88	0.000	<i>Clostridium limosum</i>	-1.00	0.317
<i>Acinetobacter</i>	-3.92	0.000	<i>Clostridium sporogenes</i>	-1.84	0.066
<i>radioresistens</i>			<i>Clostridium subterminale</i>	-1.60	0.109
<i>Acinetobacter schindleri</i>	-3.85	0.000	<i>Comamonas kerstersii</i>	-3.06	0.002
<i>Acinetobacter septicus</i>	-3.73	0.000	<i>Comamonas sp.</i>	-3.62	0.000
<i>Acinetobacter spp.</i>	-1.05	0.296	<i>Comamonas testosterone</i>	-0.242	0.809
<i>Actinomadura spp.</i>	-2.21	0.027	<i>Corynebacterium</i>	-1.84	0.066
<i>Actinomadura vinacea</i>	-2.68	0.007	<i>amycolatum</i>		
<i>Aeromonas spp.</i>	-3.07	0.002	<i>Corynebacterium falsenii</i>	-3.83	0.000
<i>Agromyces sp.</i>	-3.52	0.000	<i>Corynebacterium jeikeium</i>	-3.74	0.000
<i>Alistipes finegoldii</i>	-3.74	0.000	<i>Corynebacterium</i>	-3.44	0.001
<i>Alistipes shahii</i>	-1.00	0.317	<i>mucifaciens</i>		
<i>Alistipes spp.</i>	-2.11	0.035	<i>Corynebacterium</i>	-3.92	0.000
<i>Alteromonas sp.</i>	-1.00	0.317	<i>thomssenii</i>		
<i>Anaerococcus sp.</i>	-2.00	0.046	<i>Corynebacterium</i>	-3.92	0.000
<i>Anaerovorax spp.</i>	-3.54	0.000	<i>tuberculostearicum</i>		
<i>Arthrobacter oxydans</i>	-2.02	0.043	<i>Coxiella burnetii</i>	-1.34	0.180
<i>Arthrobacter spp.</i>	-0.65	0.514	<i>Coxiella spp.</i>	-3.07	0.002
<i>Atopobium vaginae</i>	-1.34	0.180	<i>Cupriavidus spp.</i>	-1.71	0.087
<i>Aurantimonas sp.</i>	-2.03	0.042	<i>Delftia tsuruhatensis</i>	-3.92	0.000
<i>Azospirillum brasilense</i>	-1.00	0.317	<i>Desulfomicrobium spp.</i>	-2.03	0.042
<i>Bacillus coagulans</i>	-3.37	0.001	<i>Desulfovibrio desulfuricans</i>	-1.34	0.180
<i>Bacteroides spp.</i>	-0.69	0.490	<i>Desulfovibrio spp.</i>	0.00	1.000
<i>Bacteroides vulgatus</i>	-1.00	0.317	<i>Dietzia papillomatosis</i>	-2.21	0.027
<i>Bergeyella sp.</i>	-2.54	0.011	<i>Dokdonella spp.</i>	-0.464	0.642
<i>Bordetella petrii</i>	-2.82	0.005	<i>Dyella ginsengisoli</i>	-1.83	0.068
<i>Bosea sp.</i>	-2.83	0.005	<i>Dysgonomonas</i>	-3.22	0.001
<i>Brevibacillus parabrevis</i>	-1.00	0.317	<i>capnocytophagoides</i>		
<i>Brevibacillus sp.</i>	-1.89	0.059	<i>Dysgonomonas gadei</i>	-3.92	0.000
<i>Brevibacterium sp.</i>	-2.99	0.003	<i>Dysgonomonas spp.</i>	-3.09	0.002
<i>Brevundimonas diminuta</i>	-3.74	0.000	<i>Eggerthella sp.</i>	-1.00	0.317
<i>Brevundimonas spp.</i>	-2.99	0.003	<i>Empedobacter brevis</i>	-2.23	0.026
<i>Brevundimonas vesicularis</i>	-1.34	0.180	<i>Empedobacter sp.</i>	-1.84	0.066
<i>Burkholderia spp.</i>	-3.00	0.003	<i>Enterobacter hormaechei</i>	-3.25	0.001
<i>Burkholderia tropica</i>	-1.34	0.180	<i>Enterococcus faecalis</i>	-3.63	0.000
<i>Burkholderia ubonensis</i>	-3.92	0.000	<i>Escherichia hermannii</i>	-2.39	0.017
<i>Butyrivibrio sp.</i>	-1.00	0.317	<i>Eubacterium spp.</i>	-1.25	0.212
<i>Caenispirillum sp.</i>	-1.00	0.317	<i>Exiguobacterium</i>	-2.94	0.003
<i>Campylobacter lari</i>	-1.00	0.317	<i>aurantiacum</i>		

Tabel 1 continued

Bacterial species	Z value	P – Value	Bacterial species	Z value	P – Value
<i>Exiguobacterium sp.</i>	-3.62	0.000	<i>parascrofulaceum</i>		
<i>Fastidiosipila sanguinis</i>	-1.00	0.317	<i>Mycobacterium septicum</i>	-3.52	0.000
<i>Finegoldia magna</i>	-3.83	0.000	<i>Mycobacterium ulcerans</i>	-2.97	0.003
<i>Finegoldia spp.</i>	-1.00	0.317	<i>Mycoplasma hominis</i>	-1.00	0.317
<i>Flavobacterium spp.</i>	-3.29	0.001	<i>Mycoplasma salivarium</i>	-1.00	0.317
<i>Francisella spp.</i>	-0.530	0.596	<i>Neisseria subflava</i>	-2.68	0.007
<i>Gemella sanguinis</i>	-3.76	0.000	<i>Nocardiopsis spp.</i>	-1.17	0.242
<i>Gluconobacter spp.</i>	-2.96	0.003	<i>Ochrobactrum intermedium</i>	-1.84	0.066
<i>Gordonia terrae</i>	-2.12	0.034	<i>Ochrobactrum spp.</i>	-2.02	0.043
<i>Haemophilus parainfluenzae</i>	-2.12	0.034	<i>Olsenella uli</i>	-2.97	0.003
<i>Halomonas venusta</i>	-2.55	0.011	<i>Paracoccus spp.</i>	-3.36	0.001
<i>Herbaspirillum rhizosphaerae</i>	-0.46	0.647	<i>Parvimonas spp.</i>	-2.06	0.039
<i>Herbaspirillum spp.</i>	-0.040	0.968	<i>Peptoniphilus asaccharolyticus</i>	-1.00	0.317
<i>Inquilinus spp.</i>	-2.06	0.039	<i>Peptostreptococcus spp.</i>	-1.00	0.317
<i>Janthinobacterium lividum</i>	-3.92	0.000	<i>Peptostreptococcus stomatis</i>	-1.34	0.180
<i>Kocuria rosea</i>	-3.92	0.000	<i>Pseudoclavibacter zimmermannella bifida</i>	-1.00	0.317
<i>Lachnospirillum clostridium symbiosum</i>	-1.34	0.180	<i>Pseudomonas mendocina</i>	-3.54	0.000
<i>Lactobacillus fermentum</i>	-1.60	0.109	<i>Pseudomonas oryzihabitans</i>	-3.84	0.000
<i>Lactobacillus iners</i>	-3.83	0.000	<i>Pseudomonas putida</i>	-3.92	0.000
<i>Lactobacillus paraplantarum</i>	-1.00	0.317	<i>Pseudomonas stutzeri</i>	-3.92	0.000
<i>Lactobacillus plantarum</i>	-2.26	0.024	<i>Ralstonia spp.</i>	-3.31	0.001
<i>Lactococcus garvieae</i>	-2.54	0.011	<i>Rhizobium spp.</i>	-3.92	0.000
<i>Lactococcus lactis</i>	-3.55	0.000	<i>Rhodoplanes spp.</i>	-1.91	0.056
<i>Lactococcus sp.</i>	-1.00	0.317	<i>Robinsoniella peoriensis</i>	-2.02	0.043
<i>Legionella jordanis</i>	-1.00	0.317	<i>Roseomonas mucosa</i>	-1.34	0.180
<i>Legionella lytica</i>	-2.81	0.005	<i>Roseomonas spp.</i>	-0.150	0.881
<i>Legionella pneumophila</i>	-1.84	0.066	<i>Rothia mucilaginoso</i>	-1.34	0.180
<i>Legionella sainthelensi</i>	-3.93	0.000	<i>Ruminococcus flavefaciens</i>	-0.518	0.605
<i>Leifsonia spp.</i>	-1.00	0.317	<i>Ruminococcus spp.</i>	-3.13	0.002
<i>Leuconostoc pseudomesenteroides</i>	-2.23	0.026	<i>Selenomonas spp.</i>	-1.84	0.066
<i>Lysinibacillus massiliensis</i>	-3.83	0.000	<i>Shewanella putrefaciens</i>	-3.92	0.000
<i>Massilia spp.</i>	-1.34	0.180	<i>Shigella sonnei</i>	-3.92	0.000
<i>Massilia timonae</i>	-2.19	0.029	<i>Simkania negevensis</i>	-1.34	0.180
<i>Mesorhizobium spp.</i>	-3.31	0.001	<i>Sphingobacterium spp.</i>	-0.967	0.334
<i>Methylobacterium iners</i>	-2.23	0.026	<i>Spiroplasma sp.</i>	-1.60	0.109
<i>Methylobacterium spp.</i>	-3.22	0.001	<i>Sporosarcina spp.</i>	-1.92	0.054
<i>Methylobacterium tardum</i>	-3.02	0.003	<i>Streptococcus gordonii</i>	-3.92	0.000
<i>Methylobacterium thiocyanatum</i>	-0.55	0.583	<i>Streptococcus lutetiensis</i>	-1.84	0.066
<i>Micrococcus luteus</i>	-0.626	0.538	<i>Streptococcus sanguinis</i>	-2.53	0.011
<i>Micrococcus sp.</i>	-686	0.493	<i>Streptomyces spp.</i>	-0.944	0.345
<i>Microvirgula aerodenitrificans</i>	-1.63	0.102	<i>Synergistes spp.</i>	-1.60	0.109
<i>Mogibacterium timidum</i>	-1.84	0.066	<i>Varibaculum cambriense</i>	-3.93	0.000
<i>Mycobacterium parascrofulaceum</i>	-2.33	0.020	<i>Veillonella parvula</i>	-3.74	0.000
			<i>Vibrio cholera</i>	-1.23	0.219
			<i>Wautersiella falsenii</i>	-1.34	0.180
			<i>Williamsia muralis</i>	-2.46	0.014
			<i>Wolbachia pipientis</i>	-1.34	0.180
			<i>Wolbachia spp.</i>	-3.10	0.002
			<i>Xanthomonas spp.</i>	-1.04	0.299

**Table 2.** Tests performed to determine the influence of season on human pathogenic bacterial species richness (R), evenness (E) and diversity [Shannon (H') and Simpson (D)]

Parameters	Wet season				Dry season			
	H'	D	R	E	H'	D	R	E
Mean	2.05	0.798	60.80	0.10	2.30	0.750	60.25	0.89
St.Dev	0.51	0.104	5.38	0.02	0.74	0.202	12.43	0.22
Test statistics	T =	Z =	T =	Z =	T =	Z =	T =	Z =
value	-1.22	-0.299	0.16	-3.920	-1.22	-0.299	0.16	-3.920
df	19.00	19.00	19.00	19.00	19.00	19.00	19.00	19.00
P-value	P > 0.05	P > 0.05	P > 0.05	P < 0.05	P > 0.05	P > 0.05	P > 0.05	P < 0.05



**Fig. 1.** A picture of type 1 shallow hand-dug wells found in the Ohangwena and Omusati regions of the CEB in which animals can walk into the water.



**Fig. 3.** A picture of type 3 deep hand-dug wells found in Ohangwena and Omusati regions of the CEB in which animals cannot walk into the water.

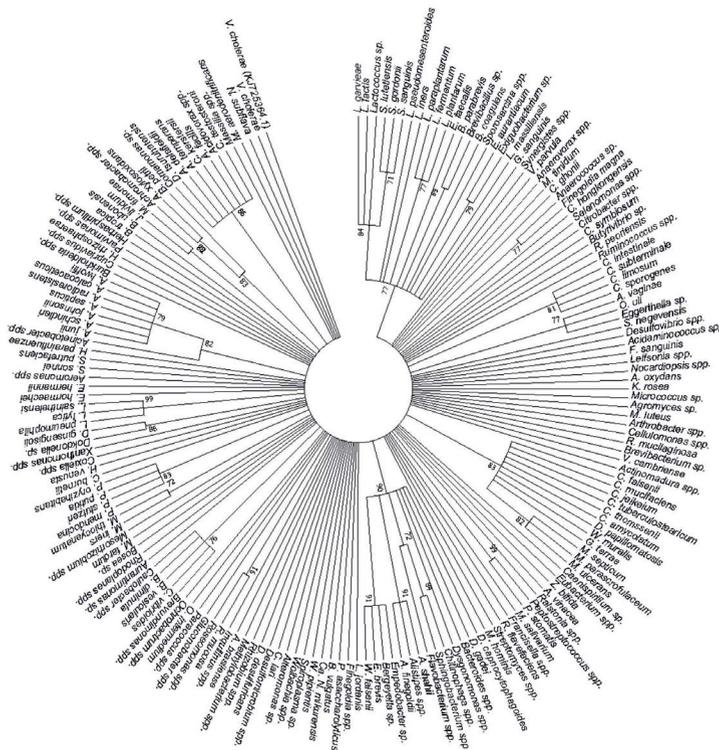


**Fig. 2.** A picture of type 2 shallow hand-dug wells found in Ohangwena and Omusati regions of the CEB in which animals cannot walk into the water.

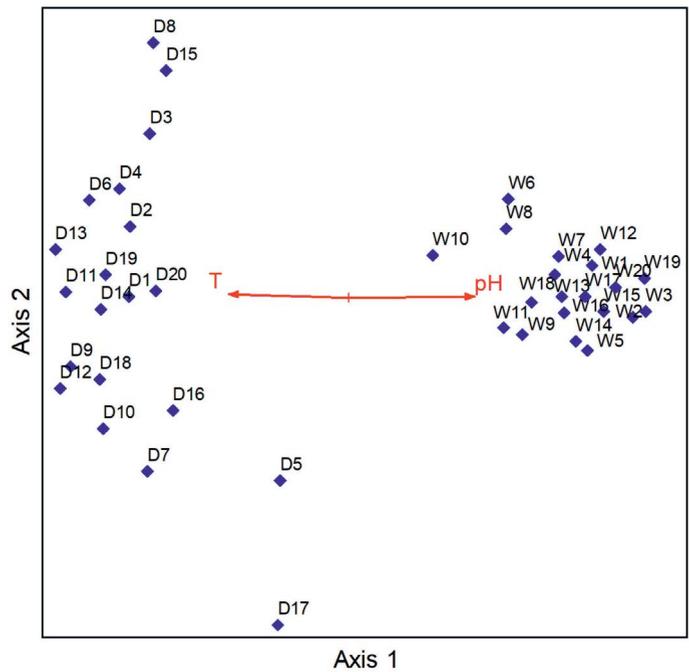
and -403.6 Mv to 141.2 Mv in the dry season. The NMS analysis performed had a total of 500 iterations upon which a final solution was reached with a stress of 11.807 and a final instability of 0.000001. NMS results revealed that the main physical factors that influenced the distribution of human pathogenic bacteria in hand-dug wells were temperature  $r=-0.628$   $\tau=-0.365$  and pH  $\tau=0.645$   $\tau=0.401$ . pH was positively correlated with the wet season while temperature was correlated with the dry season (Fig. 5).

## DISCUSSION

Most clusters of the phylogenetic tree were formed by species belonging to the same genus indicating their close relation. The detection of multiple species in each genus and their close relation confirmed intra-genus versatility. Human pathogenic bacterial species diversity and richness did not show a significant difference between the wet and dry season indicating that hand-dug well bacterial diversity and richness is independent of season. The sustained species diversity and



**Fig. 4.** The phylogenetic tree depicts the detected human bacterial pathogens by using the Maximum Likelihood method based on the Tamura-Nei model on 182 nucleotide sequences and 70% bootstrap.



**Fig. 5.** NMS result of human pathogenic bacteria illustrating a Biplot overlaid with physical factors (D=Dry season samples, W=Wet season samples).

richness was caused by the poor structure, lack of protection from animal access and lack of a covered top throughout the year allowing bacteria to be deposited into these wells. The easy access of livestock, other domestic and wild animals or birds to the water exposes these hand-dug wells to different human pathogenic bacteria regardless of season.

The diversity and richness of human pathogenic bacteria was similar in the hand-dug wells in both season due to continuous contact between soil and water in the hand-dug wells. Since various bacteria belonging to different phyla have mostly been reported to be found in the soil,<sup>16</sup> it is plausible that most of the bacteria detected originated from the soil, and were either active or dormant which is an inherent limitation of Metagenomics. Moreover, it is known that water availability and nutrients are among the main limiting factors influencing bacterial growth and survival.<sup>17</sup> The water-soil bacteria interface allows these bacteria to survive thereby maintaining the diversity within the hand-dug wells and this agrees with Bull<sup>18</sup> who reported that bacteria can grow and survive at various ranges of physicochemical parameters, and growth occurs at a slow rate because the environment is not well optimized. The findings of this study confirmed Sun *et al.*<sup>19</sup>'s results which showed a high diversity and richness of bacteria in a river in both the wet and dry seasons with no significant difference between the two seasons ( $P > 0.05$ ). Sun *et al.*<sup>19</sup> recorded Shannon indices values of 9.53 and 9.42 respectively for dry and wet seasons. However, the present study recorded lower values of Shannon indices than reported by Sun *et al.*<sup>19</sup>, but still displayed high diversity and richness in both seasons. Furthermore, the high diversity of human pathogenic bacteria in both seasons confirmed that bacteria have pronounced ecophysiological plasticity that permits them to adapt to various freshwater ecosystems and dynamic seasonal changes<sup>19,20</sup>.

Hand-dug wells had animal droppings and visible floating debris which provided a carbon rich environment propitious for microbial growth. This carbon rich environment coupled with the soil-water interface can enhance the ability of the microorganisms to cope with various or fluctuating environmental conditions by the transfer and

exchange of genes owed to microbial interactions. This is evident because *Escherichia* species were once described to be unable to survive lengthy periods outside the intestines of warm blooded animals but other works<sup>21,22</sup> have indicated that *E. coli* strains survive in soil and water that's not known to be faecally contaminated. The current status quo demands the development of a more suitable indicator of recent faecal contamination and further research to explore these emerging patterns.

The present study indicated that there was a significant difference in abundance and evenness of human pathogenic bacterial species between the dry and wet seasons. The wet season had higher abundances of human pathogenic bacterial species than the dry season, and human bacteria pathogen species in the dry season were more even than in the wet season. The variation in abundance was due to surface runoff and the downward transportation of bacteria by water through the permeable soil layers in the wet season. Surface runoff transports various bacteria into the hand-dug wells thereby elevating the abundance of some human pathogenic bacteria and leading to none even distribution of bacterial abundance in the wet season. Most of the detected pathogens were soil and gut bacteria and this agrees with Thomas<sup>23</sup> who disclosed that Namibia is amongst the highest rated southern African countries with open defecation. The poor architecture of these hand-dug wells and the livestock and human faeces that are found at sites near the hand-dug wells promote the entry of faecal matter especially in the wet season. However, it's worth noting that surface runoff mainly affects the abundance and evenness than diversity and richness especially that the water is mostly in contact with soil. It is not to rule out that new species are introduced into these hand-dug wells but to emphasize that the amount is either negligible with respect to diversity or it is sustained throughout the year.

The abundance is easily increased since water serves as a transporter of these contaminants especially in rain season when the water penetrates the permeable soil layers reaching the aquifers below that are shared by the hand-dug wells within the same vicinity,<sup>24</sup> eliminating potential differences in contamination

levels that would otherwise exist due to different types hand-dug wells. Further-more, it was noted that temperature was positively correlated with the dry season and negatively correlated with the wet season while pH was positively correlated with the wet season and negatively correlated with the dry season. It is established that pH and temperature have an influence on the growth and survival of bacteria. Optimal bacterial growth occurs at various levels of pH ranging from low, moderate, and high in different species. Similarly, various bacterial species have different temperature ranges for optimal growth.<sup>25</sup> This agrees with literature<sup>16,18</sup> disclosing that pH, salinity, water content, temperature, pressure and radiation are the major physicochemical parameters that regulate bacterial growth and survival.

The notable human pathogens of public health concern were; *Citrobacter spp.* known to cause urinary tract infections (UTI), meningitis, bacteraemia and haemolytic–uraemic syndrome, *H. parainfluenzae* known to cause sinusitis, otitis media, pneumonia, abscesses, endocarditis, and biliary tract infections<sup>26</sup>, *Legionella* species (*L. jordanis*, *L. lytica*, *L. pneumophila* and *L. sainthelensi*) known to cause Legionnaires' disease, respiratory tract infections and Pontiac fever, *Pseudomonas* species (*P. mendocina*, *P. oryzihabitans*, *P. putida*, *P. stutzeri*) known to cause bacteraemia, UTI, wound infection, abscesses, septic arthritis, conjunctivitis, endocarditis, meningitis, otitis, sepsis, peritonitis, pneumonia and urinary tract infections<sup>27</sup>, *S. sonnei* known to cause enteric infections<sup>28,29</sup>, and *V. cholerae* known to cause cholera<sup>30</sup>.

*H. parainfluenzae*, *L. lytica*, *L. pneumophila*, *L. sainthelensi*, *P. mendocina*, *P. oryzihabitans*, *P. putida*, *P. stutzeri* and *S. sonnei* showed a significant difference in abundance between the wet and dry seasons. *L. sainthelensi*, *P. oryzihabitans*, *P. putida*, *P. stutzeri* and *S. sonnei* were more abundant in wet season compared to the dry season, demonstrating that the communities of the CEB are exposed to these pathogens to a higher extent in the wet season than the dry season. The presence of *Shigella* species in these hand-dug wells is of grave concern because it is highly infectious. Bacterial counts of 0-100 are adequate to induce

shigellosis in humans<sup>31</sup>. *H. parainfluenzae*, *L. lytica*, *L. pneumophila* and *P. mendocina* were more abundant in the dry season compared to the wet season, indicating that diseases caused by these species are expected to surge in the dry season. However, there was no significant difference in the abundance of *Citrobacter spp.*, *L. jordanis* and *V. cholerae* between the wet and dry seasons, indicating that the CEB communities are exposed to these pathogens continuously. This explains the non-seasonal sporadic cholera outbreaks that occur in these communities and highlights the necessity of adhering to hygiene practices and implementing routine hand-dug well water bacteriological analysis. The rest of the detected human pathogens are reported to mostly cause endocarditis, meningitis and bacteraemia.

Metagenomics provided vast information regarding the microbial communities and safety of the hand-dug wells in the CEB for household consumption by bypassing limitations of culturing based methods that lead to the inability to quantify the total natural diversity within a given habitat<sup>32</sup>. Since Meta-genomics is a PCR-based analyses of microbial diversity, it is entrenched with some biases that are inherent to PCR applications that have been described by other studies<sup>33,34,35,36</sup>. However, it should be noted that since Meta-genomics as opposed to Metatranscriptomics is a DNA based technique, the microbial communities detected potentially included DNA from dead bacteria thereby displaying an over representation of bacterial communities or omitted some bacteria due to DNA extraction difficulties as described by Filippidou et al<sup>34</sup>. This might have led to a low coverage of less abundant taxa known as “depth bias” and under representation of certain taxa. The findings of this study led to the conclusions that seasonality has an influence on evenness and overall abundance levels but none on species diversity and richness of human bacterial pathogens in hand-dug wells. Generally, this study showed that the water in hand-dug wells of the CEB is not safe for consumption and domestic use unless sanitized.

#### AVAILABILITY OF MATERIALS

The 16S rRNA sequences data that were generated in the current study are available on GenBank, with the following accession numbers:

MH050446 - MH050607. The raw sequence files deposited into the Sequence Read Archive (SRA) are also available on GenBank under BioProject ID; PRJNA478691 and BioSample accessions; SAMN 09521643, SAMN 09521644, SAMN 09521645, and SAMN 09521646.

#### ACKNOWLEDGEMENTS

We thank Dr. Wanke through SASSCAL for funding this research and Josephine Hamutoko for the combined sampling with Billy McBenedict, and recording of pH, temperature and electrical conductivity.

#### CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

#### FUNDING

This work was funded by Dr. Heike Wanke through SASSCAL through a Doctorate research scholarship. The funding body did not have a role in the analysis and interpretation of the data, and preparation or writing of this manuscript.

#### AUTHOR CONTRIBUTIONS

BM collected the water samples and performed the research. WH and BM analysed microbial diversity, richness, evenness and the statistical analysis. BM, WH, HW, PC and BH wrote the manuscript. BM, PC and BH reviewed the manuscript.

#### FUNDING

This work was funded by Dr. Heike Wanke through SASSCAL through a Doctorate research scholarship. The funding body did not have a role in the analysis and interpretation of the data, and preparation or writing of this manuscript.

#### DATA AVAILABILITY

The 16S rRNA sequences data that were generated in the current study are available on GenBank, with the following accession numbers: MH050446 - MH050607. The raw sequence files deposited into the Sequence Read Archive (SRA) are also available on GenBank under BioProject ID; PRJNA478691 and BioSample accessions; SAMN 09521643, SAMN 09521644, SAMN 09521645, and SAMN 09521646.

#### ETHICS STATEMENT

Not applicable

#### REFERENCES

1. Mendelson J., Jarvis A., Roberts C. and Robertson, T. Atlas of Namibia: A Portrait of the Land and its People. Cape Town. 2009.
2. Christellis G., Struckmeier W. and Baumie R. Groundwater in Namibia: an explanation to the hydrogeological map. 2001
3. Amakali M., Swatuk L.A. Different approaches to local level participation in river basin management in Namibia: A comparison between the Kuiseb and Cuvelai basins. *In Transboundary Water Governance in Southern Africa* 2009 Jan 14 (pp. 111-132). Nomos Verlagsgesellschaft mbH & Co. KG.
4. Zimmermann M. The coexistence of traditional and large-scale water supply systems in central northern Namibia. *Journal of Namibian Studies: History Politics Culture*, 2010; **7**:55-84.
5. Kluge T., Liehr S., Lux A., Moser P., Niemann S., Umlauf N., Urban W. IWRM concept for the Cuvelai Basin in northern Namibia. *Physics and Chemistry of the Earth, Parts A/B/C*, 2008; **33**(1-2): 48-55.
6. Lindenmaier F., Miller R., Fenner J., Christelis G., Dill H.G., Himmelsbach T., Kaufhold S., Lohe C., Quinger M., Schildknecht F., Symons G. Structure and genesis of the Cubango Megafan in northern Namibia: implications for its hydrogeology. *Hydrogeology Journal*, 2014; **22**(6): 1307-28.
7. McBenedict B., Wanke H., Hang'ombe B. M. and Chimwamuroombe P.M. Bacteriological analysis of household water from hand-dug wells in the Cuvelai-Etoshia Basin of Namibia. *International Science and Technology Journal of Namibia*, 2017; **10**: 23-32.
8. Sharon I., Banfield J.F. Genomes from meta-genomics. *Science*, 2013; **342**(6162): 1057-8.
9. Simon C., Daniel R. Metagenomic analyses: past and future trends. *Applied and Environmental Microbiology*, 2011; **77**(4): 1153-61.
10. Capone K.A., Dowd S.E., Stamatas G.N., Nikolovski J. Diversity of the human skin microbiome early in life. *Journal of Investigative Dermatology*, 2011; **131**(10): 2026-32.
11. DeSantis T.Z., Hugenholtz P., Larsen N., Rojas M., Brodie E.L., Keller K., Huber T., Dalevi D., Hu P., Andersen G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, 2006; **72**(7): 5069-72.
12. Tamura K., Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, 1993; **10**(3): 512-26.
13. Kumar S., Stecher G., Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 2016; **33**(7): 1870-4.
14. Pielou E.C. Ecological diversity. *J. Wiley and Sons, New York*. 1975

15. Magurran A.E. Measuring biological diversity. John Wiley & Sons; 2013 Apr 18.
16. Mohammadipanah F., Wink J. Actinobacteria from arid and desert habitats: diversity and biological activity. *Frontiers in Microbiology*. 2016; **28**(6): 1541.
17. Stevenson A., Hallsworth J.E. Water and temperature relations of soil Actinobacteria. *Environmental microbiology reports*, 2014; **6**(6): 744-55.
18. Bull A.T. Actinobacteria of the extremobiosphere. In *Extremophiles handbook 2011* (pp. 1203-1240). Springer Japan.
19. Sun W., Xia C., Xu M., Guo J., Sun G. Seasonality Affects the Diversity and Composition of Bacterioplankton Communities in Dongjiang River, a Drinking Water Source of Hong Kong. *Frontiers in Microbiology*, 2017; **31**(8): 1644.
20. Allgaier M., Breckner S., Jaspers E., Grossart H.P. Intra and inter lake variability of free living and particle associated Actinobacteria communities. *Environmental Microbiology*, 2007; **9**(11): 2728-41.
21. Ishii S., Hansen D.L., Hicks R.E., Sadowsky M.J. Beach sand and sediments are temporal sinks and sources of *Escherichia coli* in Lake Superior. *Environmental Science and Technology*, 2007; **41**(7): 2203-9.
22. NandaKafle G., Seale T., Flint T., Nepal M., Venter S.N., Brzdzel V.S. Distribution of Diverse *Escherichia coli* between Cattle and Pasture. *Microbes. and Environments*, 2017; **32**(3): 226-33.
23. Thomas A. Strengthening post-ODF programming: reviewing lessons from sub-Saharan Africa. *Sustainable Sanitation for All: Experiences, challenges, and innovations*, 2016: 83.
24. Van der Wal A. Understanding Groundwater & Wells in manual drilling. Unicef. <http://www.unicef.org/wash/files/04.pdf>. 2008.
25. Madigan M.T., Martinko J.M., Bender K.S., Buckley D.H. and Stahl D.A. Brock Biology of Micro-organisms. 14<sup>th</sup> edition. Pearson Education. 2015
26. Frankard J., Rodriguez-Villalobos H., Struelens M.J., Jacobs F. *Haemophilus parainfluenzae*: an under diagnosed pathogen of biliary tract infections. *European Journal of Clinical Micro-biology and Infectious Diseases*. 2004; **23**(1): 46-8.
27. Tena D., Fernandez C. *Pseudomonas oryzi-habitans*: an unusual cause of skin and soft tissue infection. *Infectious Diseases*, 2015; **47**(11): 820-4.
28. Bowen A., Eikmeier D., Talley P., Siston A., Smith S., Hurd J., Smith K., Leano F., Bicknese A., Norton J.C., Campbell D. Notes from the Field: Outbreaks of *Shigella sonnei* Infection with Decreased Susceptibility to Azithromycin Among Men Who Have Sex with Men-Chicago and Metropolitan Minneapolis-St. Paul, 2014. *MMWR. Morbidity and mortality weekly report*, 2015; **64**(21): 597-8.
29. Thompson C.N., Duy P.T., Baker S. The rising dominance of *Shigella sonnei*: an intercontinental shift in the etiology of bacillary dysentery. *PLoS Neglected Tropical Diseases*, 2015; **9**(6): e0003708.
30. Bhuiyan T.R., Hoq M.R., Nishat N.S., Al Mahbuba D., Rashu R., Islam K., Hossain L., Dey A., Harris J.B., Ryan E.T., Calderwood S.B. Enumeration of gut-homing  $\beta$ 7-positive, pathogen-specific antibody-secreting cells in whole blood from enterotoxigenic *Escherichia coli*-and *Vibrio cholerae*-infected patients, determined using an enzyme-linked immunosorbent spot assay technique. *Clinical and Vaccine Immunology*, 2016; **23**(1): 27-36.
31. Alamanos Y., Maipa V., Levidiotou S., Gessouli E. A community waterborne outbreak of gastro-enteritis attributed to *Shigella sonnei*. *Epidemiology and Infection*, 2000; **125**(3): 499-503.
32. Lombard N., Prestat E., van Elsas J.D., Simonet P. Soil specific limitations for access and analysis of soil microbial communities by metagenomics. *FEMS Microbiology Ecology*, 2011; **78**(1): 31-49.
33. De Bruijn F.J. Handbook of molecular microbial ecology II: metagenomics in different habitats. *John Wiley & Sons*; 2011 Oct 14.
34. Filippidou S., Junier T., Wunderlin T., Lo C.C., Li P.E., Chain P.S., Junier P. Under-detection of endospore-forming Firmicutes in metagenomic data. *Computational and Structural Bio-technology Journal*, 2015; **13**: 299-306.
35. Pinar D., de Winter A., Sarkis G.J., Gerstein M.B., Tartaro K.R., Plant R.N., Egholm M., Rothberg J.M., Leamon J.H. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. *BMC Genomics*, 2006; **7**(1): 216.
36. Wunderlin T., Junier T., Roussel Delif L., Jeanneret N., Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore forming Firmicutes. *Environmental Microbiology Reports*, 2013; **5**(6): 911-24.