Vol 3 | Issue 1 | Pages 91-101

Journal of Soil and Water Science
ISSN: 2643-5799

Research Article

DOI: 10.36959/624/434

Groundwater Evolutionary Processes and Quality Characterization: A Case of Olbanita Aquifer System, Lower Baringo Basin, Kenya Rift

Benjamin Sosi[1], Albert Getabu[1], Samson Maobe[1] and Justus Barongo[2]*

[1]*Department of Natural Resources, Kisii University, Kenya*

[2]*Department of Geology, University of Nairobi, Kenya*

Abstract

A hydro-geochemical relation has been hypothesized through the analyses of physico-chemical data of a fractured volcanic rock aquifer located in the Lower Baringo Basin, Kenyan Rift. Datasets included 15 individual metrics determined in forty-two dry and wet season water samples obtained from six boreholes in the area. Aquifer evolutionary theory was postulated using sequential Principal Component Analysis and Hierarchical Cluster Analysis. In order to eliminate effects of scale dimensionality, PCA decomposed the variable data into four factors namely electrical conductivity, salinity, alkalinity and carbonate equilibrium with external pH control for the dry season, and salinity, carbonate equilibrium with external pH control, alkalinity and electrical conductivity, for the wet season. The main result depicted a major shift in variability factor from electrolytic conductivity (34.8%) in the dry season to salinity (23.5%) in the wet season. Ward's linkage cluster analysis partitioned the aquifer into two spatially discrete associations; the western and the eastern entities respectively, in spite of their shared recharge area. These agglomerative scheduling validated in an integrative approach (with groundwater flow predictions using a calibrated petro-physical groundwater model for the area), linked the four factors to aquifer processes and three pathways: Fault permeability, weathering processes, and water-rock interaction. Statistical approaches are, therefore, useful in conceptualization of pollutant sources and their attenuation for effective groundwater quality management.

Keywords

Groundwater evolutionary processes, Water quality characterization, Olbanita, Kenya rift

Introduction

The Olbanita aquifer system is located in the lower Baringo basin of the Kenyan Rift. It's characterized by proximity to the equator and by its elevation which ranges from 1750 meters to 1880 meters a.s.l. Despite its semi-arid climate and open-faulted drainage, the occurrence of groundwater in it makes it indispensable for human subsistence. As noted by [1], water availability problems in the greater Nakuru and Baringo basins, had risen to an echelon whereby they negatively impacted on resident communities and the regional GDP. The hydrogeology of the area is comprised of fractured and weathered volcanic rocks, and lacustrine sediments. The weathered tuffs sandwiched between the Samburu basalts and the Wasagess flows (phonolites and trachytes) of the Rumuruti group form the best aquifers in Olbanita area. The near N-S trending fault systems provide groundwater high porosity pathways which interrupt the aquifer system at five sites.

The city of Nakuru, situated in the upper Lake Nakuru Basin in the south relies heavily on the eight wells drilled in the adjacent Lower Baringo basin for potable water supply. Despite the significance of groundwater in the region, a wide gap in knowledge exists on the aquifer processes that control the physical and chemical variability of groundwater quality [2] observed that sustainable development and management of the quality of groundwater resources in ASALs is achievable through improved understanding of geochemical evolution and groundwater processes.

Determination of groundwater evolutionally processes in relation to groundwater quality in fractured confined aquifers is complex [3], noted that interpretation of associations on the basis of graphical techniques is fraught

***Corresponding author:** Benjamin Sosi, Department of Natural Resources, Kisii University, Kenya

Accepted: September 28, 2019

Published online: September 30, 2019

Citation: Sosi B, Getabu A, Maobe S, et al. (2019) Groundwater Evolutionary Processes and Quality Characterization: A Case of Olbanita Aquifer System, Lower Baringo Basin, Kenya Rift. J Soil Water Sci 3(1):91-101

with difficulties owing to the simultaneous nature of aquifer processes and their effects. Recent studies by [2,4,5] established the value of multivariate techniques such as Principal Component Analysis (PCA), and Hierarchical Clustering Analysis (HCA) in the provision of quantitative measures of correlation between water quality parameters and fundamental aquifer processes. PCA is widely utilized to characterize groundwater pollution sources [2], while HCA is used to deduce spatial variability amongst sampling sites [6]. An integrative application of PCA and HCA was employed by [5] to interpret processes affecting groundwater hydrochemistry and by [7] to evaluate evolutionary trends of groundwater dynamics and by [8] to assess spatial patterns of pollutants in water. A number of investigators [9,10] utilized z-score or log conversions to normalize physical and chemical data acquired from groundwater sources. However, [11] noted that such transformations could eliminate legitimate data values from the analysis. It is therefore suggested that Wards' Linkage agglomeration procedure could automatically re-scale metric data into a range and yield good recovery of clusters.

In this study, we postulate an evolutionary trend of the fractured volcanic aquifer based on multivariate statistics of groundwater quality data of the area. By way of these evaluations, we determine relationships amongst deep groundwater wells and major sources of field-scale groundwater quality variations.

Materials and Methods

The sampling design in this study involved 6 out of 8 sampling sites constituted of boreholes in the Olbanita aquifer system located in the Kenya Rift (Figure 1). The sampling sites were located using a Garmin GPSmap60CSx model. At each sampling site, 4 replicate samples were taken monthly in 250 ml high density plastic bottles which had already been pre-cleaned by using concentrated nitric acid and drenched in deionized water. The samples were screened through Millipore membrane filters of pore size 0.2 micrometers to remove suspended solids. Thereafter the water samples for sulphate analysis were precipitated using 0.2M $Zn(CH_3COO)_2$ while samples for metal analysis were acidified in concentrated HCl or HNO_3^- 1 mL per 100 mL sample. Insitu parameters (pH, Electrical conductivity and Total Dissolved Solids) were measured using Hanna Instruments multi-parameter meter model HI98194. Out of the 4 samples taken in plastic bottles, two replicate samples were stored in ice and taken to the laboratory for determination of geochemical parameters namely; Na^+, K^+, Ca^{2+}, Cl^-, F^-, SO^{2-}_4, H_2S, Total CO_2, CO^{-2}_3, HCO^-_3, H_2CO_3 and NH^+_4. The concentrations of sulphate and chloride components in this study were quantified using UV/VIS spectro-photometer model UV-1800 while the ions such as Na, F, Ca, K and NH_3 were quantified using an Ion Selective Electrode (ISE) model ELIT 9801 as detailed in [12]. The concentrations of non-ionic components namely Total carbonate carbonate (TCC) and H_2S were quantified using

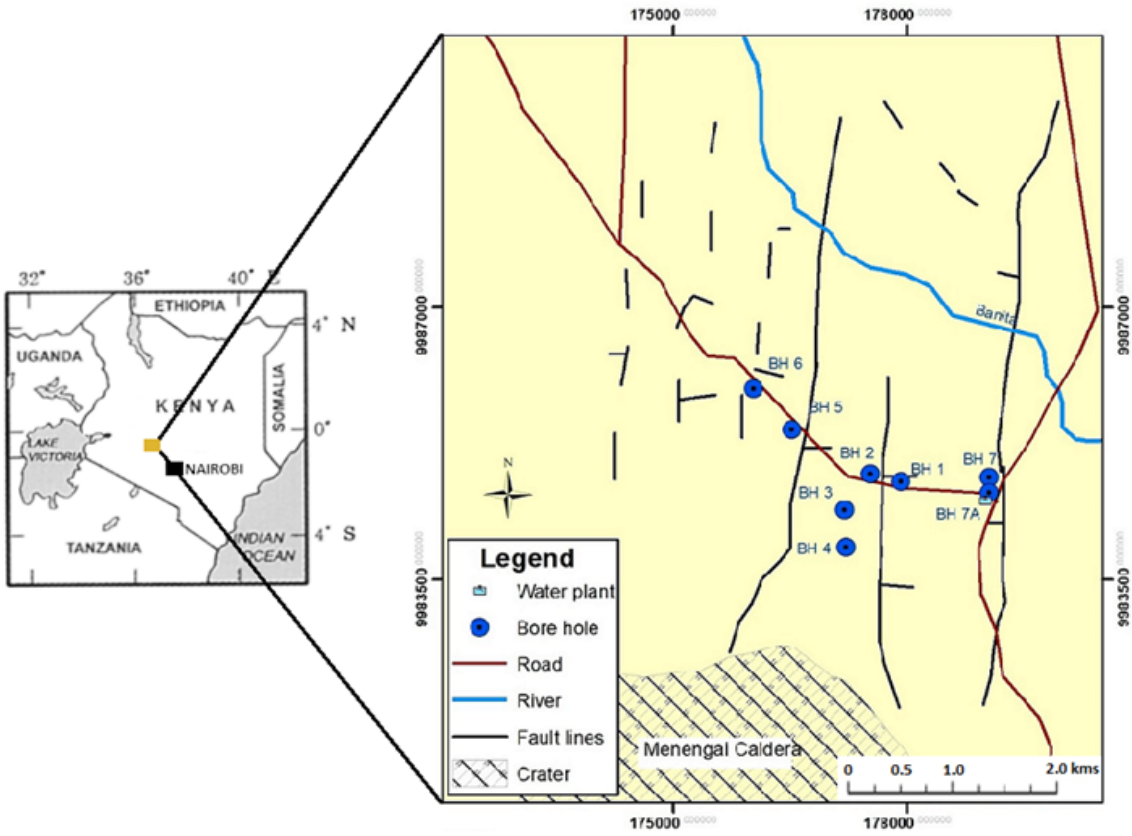

Figure 1: Location of Olbanita aquifer in Kenya showing sampled boreholes (blue circled dots).

Titrimetric methods as also described in [12]. Prior to ISE analysis, the acidified samples were re-digested using a strong alkali solution. TCC was further speciated into CO_3^{2-}, CO_2, $H2CO^-_3$, and HCO_3 components.

To evaluate the appropriateness of geochemical data for factor analysis, Kaiser-Mayer-Olkin (KMO) measure of sampling adequacy and Barlett's test of sphericity were performed [13,14]. The KMO index was found to be 0.6 indicating a moderate degree of correlation amongst the variables could be appropriately carried out. Additionally, the Barlett's test of sphericity was significant (Bartlett χ^2 = 0.570, 66 df, $p < 0.001$); confirming the suitability of factor analysis on the data collected in this study. The results of the assessment depict a lower number of variables (factors) which may be used to elucidate the variability in the hydrochemical data. The correlation matrix was further inspected for correlation coefficients greater than 0.3 [15]. Statistically significant correlations in physical chemical data sets for each season were identified through application of the Spearman's rho for non-normally distributed data. Most of the correlation coefficients are over 0.3 indicating that factor analysis may be utilized to provide significant reductions in data dimensionality [16,17] warn about the dangers of such simplified substitutions for non-detectable concentrations (by using numerical surrogates such as one-half the detection limit). Data were, therefore, automatically re-scaled for HCA (because each predictor variable adopted a different scale of measurement), though exempted datasets for PCA. The laboratory geochemical datasets were utilized for PCA considering that the bottom-line correlation matrix (based on the Spearman rank) has the effect of standardizing the variable data [18]. Therefore, it is a more robust estimation technique which is less responsive to outliers compared to the widely used Pearson correlation matrix [18].

Groundwater quality data were further subjected to PCA and HCA. All statistical computations were executed using of MS Excel spreadsheet and SPSS software's version 20.0. For the PCA matrix, orthogonalization of factors was based on rotated varimax method (with significant eigen value loadings > ± 0.5) and a derived scree plot (with the criterion of eigen values > 1) were inspected for purposes of extracting varifactors [19]. For HCA, the software's algorithm utilized Euclidean distances and 'sum of squared errors' to minimize the criterion function.

Results and Discussions

General consideration of data sets

Based on the test of skewness which uses the arithmetic mean and standard deviation, the ions Cl^-, Na^+, F^-, Ca^{+2}, SO_4, $H_2CO^-_3$, CO_3^{2-} and H_2S are not considered to follow a normal frequency distribution across the study area and across the sampling seasons. On the other hand, the physical - chemical parameters such as TDS, pH, EC, as well as CO_2 and NH_4 depicted a Gaussian distribution (Table 1). The binomial nature of the former is an indication of natural substandard waters characterized by high mineralization. The super-saturation of Olbanita groundwater may be linked to processes like dissolution of halide, ion exchange and weathering of sodium-rich plagioclases (usually giving rise to clay mineralogy). The high

Table 1: Descriptive statistical data for all the parameters for the sampling period (n = 6, N = 42).

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic						Statistic	Std. Error
pH	42	1.670	7.400	9.070	8.16452	0.441495	0.105	0.365
TDS	42	156.000	242.000	398.000	320.16667	45.152272	-0.231	0.365
EC	42	312.000	482.000	794.000	641.80952	92.302324	-0.162	0.365
Cl	42	10.640	10.180	20.820	14.69536	2.686202	0.948	0.365
Na	42	127.420	134.320	261.740	173.22667	35.178236	1.232	0.365
SO_4	42	19.890	0.900	20.790	7.85314	6.578745	0.705	0.365
F	42	9.580	3.230	12.810	6.36310	3.310061	0.804	0.365
Ca	42	15.190	2.15	17.340	6.40024	4.675027	1.659	0.365
K	42	11.500	1.970	13.470	6.56071	3.822435	0.284	0.365
CO_2	42	168.206	134.404	302.610	230.4072	52.722064	-0.151	0.365
H_2CO_3	42	28.816	0.758	29.574	8.04064	7.570068	1.358	0.365
HCO_3	42	220.384	181.518	401.903	306.08954	69.488132	-0.118	0.365
CO_3	42	33.031	0.490	33.522	5.42143	6.540314	2.469	0.365
H_2S	42	0.720	0.000	0.720	0.03374	0.109123	6.36	0.365
NH_4	42	0.530	0.360	0.890	0.56310	0.159381	0.441	0.365
Valid N (listwise)	42							
Units: ppm (except for pH and EC. EC in microSiemen/cm).								

Table 2: Spearman'rank correlation coefficients of the physicochemical parameters of groundwater for the dry season (November, 2017-February, 2018). Strong correlation between variables are specified by coefficients in bold fonts.

	pH	**TDS**	**EC**	**Cl^-**	**Na^+**	**SO^{-2}_4**	**F^-**	**Ca^{2+}**	**K^+**	**CO_2**	**H_2CO_3**	**HCO^-_3**	**$CO^=_3$**	**H_2S**	**NH^-_4**
pH	1.000	-0.009	0.067	-0.156	0.533**	-0.003	**0.763****	-0.525**	-0.216	-0.109	-0.590**	0.018	0.611**	0.161	-0.683**
TDS		1.000	**0.985****	-0.463*	0.702**	-0.446*	0.374	0.278	-0.017	**0.813****	0.226	**0.756****	-0.015	-0.373	0.147
EC			1.000	-0.459*	0.735**	-0.404	0.430*	0.247	-0.031	0.779**	0.182	**0.756****	0.034	-0.341	0.098
Cl^-				1.000	-0.373	**0.842****	-0.200	0.288	0.526**	-0.676**	0.230	-0.692**	-0.371	0.383	0.037
Na^+					1.000	-0.400	0.679**	-0.065	-0.018	0.499*	-0.151	0.549**	0.298	-0.221	-0.233
SO^{-2}_4						1.000	0.010	0.134	0.297	-0.714**	0.093	-0.687**	-0.225	0.437*	-0.223
F^-							1.000	-0.546**	-0.425*	0.106	-0.337	0.202	0.416*	-0.004	**-0.752****
Ca^{2+}								1.000	0.820**	0.203	0.472*	0.126	-0.424*	-0.072	**0.774****
K^+									1.000	-0.209	0.294	-0.235	-0.314	0.082	0.513*
CO_2										1.000	0.089	**0.929****	0.12	-0.486*	0.331
H_2CO_3											1.000	0.063	**-0.966****	-0.358	0.545**
HCO^-_3												1.000	0.149	-0.593**	0.273
$CO=_3$													1.000	0.239	-0.492*
H_2S														1.000	-0.235
NH^-_4															1.000

**Correlation is significant at α = 0.01 level (2-tailed); *Correlation is significant at α = 0.05 level (2-tailed).

Table 3: Spearman'rank correlation coefficients for physicochemical parameters of groundwater for the wet season (March-May, 2018). Strong correlation between variables are specified by coefficients in bold fonts.

	pH	**TDS**	**EC**	**Cl^-**	**Na^+**	**SO^{-2}_4**	**F^-**	**Ca^{2+}**	**K^+**	**CO_2**	**H_2CO_3**	**HCO^-_3**	**$CO^=_3$**	**H_2S**	**NH^-_4**
pH	1.000	0.010	-0.007	-0.239	0.621**	0.225	**0.873****	-0.557*	-0.161	-0.317	**-0.980****	-0.24	**0.978****	-0.041	-0.666**
TDS		1.000	**0.991****	-0.636**	0.704**	-0.235	0.258	0.003	-0.019	0.231	0.045	0.287	0.072	-0.350	0.134
EC			1.000	-0.645**	0.686**	-0.226	0.224	0.015	-0.011	0.249	0.067	0.298	0.053	-0.319	0.124
Cl^-				1.000	-0.600**	0.538*	-0.265	0.112	0.629**	-0.406	0.166	-0.459	-0.259	0.575*	0.272
Na^+					1.000	-0.261	**0.788****	-0.370	-0.150	0.119	-0.542*	0.193	0.697**	-0.433	-0.280
SO^{-2}_4						1.000	0.082	-0.022	0.534*	-0.686**	-0.335	-0.699**	0.108	**0.760****	-0.081
F^-							1.000	-0.705**	-0.208	-0.381	**-0.853****	-0.309	**0.867****	-0.169	-0.581*
Ca^{2+}								1.000	0.461	0.598**	0.558*	0.556*	-0.490*	0.123	0.656**
K^+									1.000	0.012	0.159	0.007	-0.106	0.511*	0.337
CO_2										1.000	0.418	**0.979****	-0.189	-0.385	0.225
H_2CO_3											1.000	0.362	**-0.930****	-0.068	0.637**
HCO^-_3												1.000	-0.098	-0.454	0.181
$CO^=_3$													1.000	-0.143	-0.579*
H_2S														1.000	0.010
NH^-_4															1.000

**Correlation is significant at the 0.01 level (2-tailed); *Correlation is significant at the 0.05 level (2-tailed).

concentration of H_2S is an indication of very deep circulation attaining anoxic conditions accompanied with abundant bacterial activity.

To offset the effects on mineral dissolution caused by dilution by meteoric waters from the analyses, seasonal data sets were analyzed independently. The initial step in the analysis was to account for the extent of mutual variability between individual pairs of water quality variables during the separate seasons. The Inter-Item Correlation Matrix of the measured parameters during the dry season

Table 4: Extracted Factor loadings of the measured parameters during the dry season which suited the provisions of Orthogonal Varimax rotation.

	Component				**Communalities**
	1	**2**	**3**	**4**	**Extraction**
pH	0.037	-0.230	0.885*	0.087	0.844
TDS	0.968*	-0.112	0.091	-0.003	0.957
EC	0.959*	-0.083	0.156	0.012	0.952
Cl^-	-0.583*	0.763*	-0.126	-0.132	0.955
Na^+	0.725*	0.094	0.613*	-0.104	0.921
SO^{-2}_4	-0.548*	0.775*	0.086	-0.111	0.92
F^-	0.373	-0.053	0.897*	-0.056	0.95
Ca^{2+}	-0.205	0.889*	-0.299	-0.022	0.923
K^+	0.134	0.870*	-0.234	-0.183	0.862
CO_2	0.952*	-0.174	-0.111	0.111	0.961
H_2CO_3	0.219	0.204	-0.555*	-0.543	0.693
HCO^-_3	0.943*	-0.184	-0.077	0.056	0.932
$CO^=_3$	0.139	-0.168	0.254	0.904*	0.93
H_2S	0.063	-0.045	-0.185	0.902*	0.854
NH^-_4	0.370	0.440	-0.728*	-0.094	0.869
Probable Process	Electrical conductivity	Water salinity	Water alkalinity	Carbonate equilibrium with exogenic pH control	

Table 5: The Factor loadings of the measured parameters during the wet season which suited the provisions of Orthogonal Varimax rotation.

	Component				**Commonalities**
	1	**2**	**3**	**4**	**Extraction**
pH	-0.166	-0.169	0.948*	0.157	0.980
TDS	-0.148	0.292	0.060	0.934*	0.983
EC	-0.194	0.280	0.010	0.921*	0.964
Cl^-	0.830*	-0.292	-0.155	-0.379	0.942
Na^+	-0.071	-0.203	0.285	0.925*	0.983
SO^{-2}_4	0.694*	-0.702*	0.043	-0.059	0.980
F^-	-0.071	-0.246	0.735*	0.596*	0.961
Ca^{2+}	0.865*	0.002	-0.191	-0.307	0.879
K^+	0.929*	-0.040	-0.198	0.148	0.926
CO_2	-0.088	0.967*	-0.147	0.092	0.974
H_2CO_3	0.102	0.485	-0.762*	-0.023	0.827
HCO^-_3	-0.102	0.968*	-0.122	0.100	0.973
$CO^=_3$	-0.118	0.237	0.944*	0.061	0.966
H_2S	0.572*	-0.503*	0.081	-0.237	0.643
NH^-_4	0.519*	0.367	-0.529*	-0.038	0.685
	Aquifer salinity	Carbonate equilibrium with external pH controls	Alkalinity of water	Electrolytic conduction	

and wet season is provided in Table 2 and Table 3, respectively.

Generally, the rotated component matrix was found to contain both positive and negative loadings (Table 4 and Table 5). The work of [20] as enumerated in [21] observed that Eigen value loadings near ± 1 designate a strong association between a variable and a Principal Component (PC); Eigen values exceeding ± 0.75 represent strong correlation, eigen values between ± 0.5 and ± 0.74 represent moderate correlation and those approaching 0 depict weak correlations. Each PC was attributed to a process owing to which the corresponding variables are probably linearly linked. The underlying processes occurring within the aquifer as construed from the consequent Eigen value loadings are presented in Table 4 and Table 5.

Dry season water quality parameters

For the dry season data, the pairs; pH-F, TDS-EC, TDS-CO_2, TDS-HCO_3, EC-CO_2, EC-HCO_3, Cl-SO_4, Ca-K, Ca-NH_4, F-NH_4, CO_2-HCO_3, as well as HCO_3-CO_3 showed strongly significant relationships. The pairs; pH-Na, pH-Ca, pH-H_2CO_3, pH-NH_4, TDS-Na, EC-Na, K-Cl, Cl-CO_2, Cl-HCO_3, Na-F, Na-HCO_3, SO_4-CO_2, SO_4-HCO_3, Ca-F, NH_4-H_2CO_3 and HCO_3-H_2S depicted moderate correlations (Table 2).

Principal components extracted: On the basis of the eigen values > 1 criterion, four principal components explained variability in groundwater quality at the site (Figure 2).

Based on the Cumulative variance of the rotation sums of squared loadings of the dry season, the retained latent constructs account for 90.1% of the variance in the dataset (Table 6). PC1 with the largest Eigen value accounted for maximum of the total variability (34.8%). PC2 accounted for the total variation of 21%. The third PC explained 20.8% of the total variance. The fourth and final PC explained 13.6% of the remaining variation in the data. The observed Eigen value decomposition corresponds to earlier observations by [22] that after the first PC, the second PC explains the greatest of the residual variance and so forth.

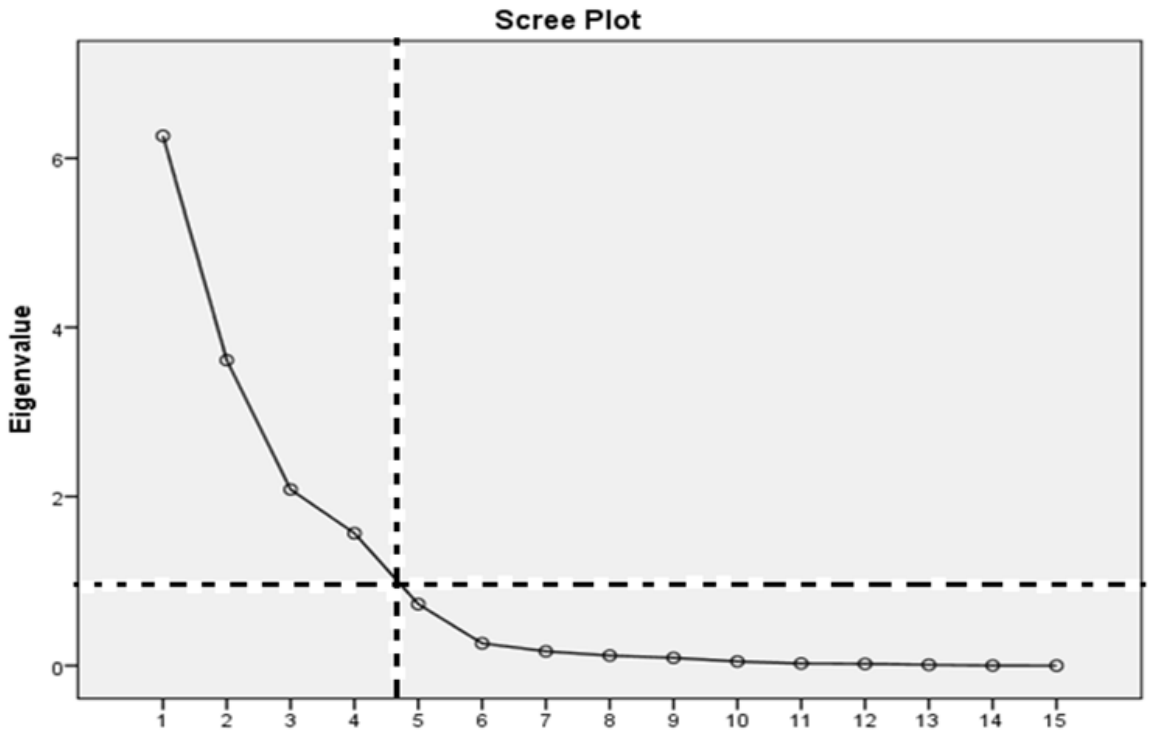

Figure 2: Cattel Scree plot with Eigenvalues > 1 criteria.

Table 6: Total variance explained for the dry season data.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.265	41.763	41.763	6.265	41.763	41.763	5.226	34.842	34.842
2	3.611	24.077	65.840	3.611	24.077	65.840	3.143	20.955	55.797
3	2.082	13.882	79.722	2.082	13.882	79.722	3.117	20.783	76.58
4	1.565	10.434	90.156	1.565	10.434	90.156	2.036	13.576	90.156

Extraction Method: Principal Component Analysis.

The first principal component, PC1 (the conductivity component) is associated with significantly high concentrations of electrolytic ions indicated by TDS, EC, Na^+, Cl^-, SO^{-2}_4, CO_2 and HCO^-_3. TDS depicted a strong correlation with EC (+0.99, α = 0.01) due to the extensive range in solubility of/and mineral diversity within the aquifer system. Under sluggish flow during the dry season, ground water can attain chemical saturation with respect to TDS. Statistical analyses associate Na^+ with the dissolution or chemical weathering of sodium-rich plagioclases (to produce clay minerals) or the dissolution of halide [5]. The moderate correlation observed between Na and F (+0.7) and Ca-F (-0.5) coupled with a relatively weak association between the pairs; Na-Cl (-0.4) supports partial derivation of electrolytic ions from weathering of Na-rich feldspathic rocks, dissolution of accessory mineral apatite as well as carbonate materials as opposed to dissolution of either halide rocks. The insoluble products of rock weathering such Cl^-, and SO^{2-}_4 show a strong positive correlation (+0.84, α = 0.01) but cumulatively tend to inhibit electrical conductivity of groundwater. Within the hydro-geological framework, the pattern and therefore, provenance of weathering can be accounted for by the roughly N-S fracture-fissure zones (Figure 1).

The second principal component, PC2 (the salinity component) is depicted mainly by Cl^-, SO^{-2}_4, Ca^{+2} and K^+ ions in water. The anomalous distribution of Cl^-, SO^{-2}_4, Ca^{+2} is attributable to ion exchange mechanisms in saturated aquifer zones. The relatively strong association between the pairs Cl-SO_4 (+0.8), and the moderate association between the pairs Ca-F, K-Cl, SO_4-CO_2, Cl-HCO_3, and SO_4-HCO_3, indicate that aquifer water salinity is chiefly attributed to geologic derivation. Additionally, agriculture is equally a major pollution source owing to the skewed distribution of Ca^{2+} which depicts strong correlations with K^+ ions (+0.8) and NH^-_4 (+0.8) at α = 0.01.

The third principal component, PC3 is the alkalinity component as indicated by pH, and Na^+, F^-, H_2CO_3 and NH^-_3. The pH, Na and F have positive loadings whereas H_2CO_3 and NH_3 have moderate negative loading on this PC. H_2CO_3 and NH_3 are slightly broken down (that is, at about 25 °C) to release H^+, HCO^{3-}, $CO^=_3$ and NH^-_3 thereby reducing the pH. Reduced pH significantly increases the rates of weathering introducing more Na and F from geologic sources.

The fourth PC represents Carbonate equilibrium with exogenic pH control. The external factor controlling pH is H_2S. Sulphide in borehole waters is probably due to inorganic and bacterial changes in deep aquifer under low dissolved oxygen, optimum growth range in pH (between 5.5 and 8.5) and optimum temperature (between 24 °C and 42 °C). Deeper depths within the bedrock which are readily flushed by low mobility, oxygen-deficient ground water; effects which are conditioned by low permeability at depth. The sulfate reducing bacteria, *Desulfovibrio desulfuricans* obtain energy via the inter-conversion between sulfates and sulphides within the larger sulphur cycle in the aquifer system [23,24]. Under these conditions, sulphates may be reduced to sulphides producing metallic sulphide which is again changed to H_2S under the action of H_2CO_3. The dissociation of H_2CO_3 yields the $CO^=_3$ and donate H^+ which consequently reduce sulphates to sulphides.

Wet season water quality parameters

During the wet sampling season the pairs; pH-F, pH-H_2CO_3, pH-CO_3, TDS-EC, Na-F, F-H_2CO_3, F-CO_3, CO_2-H_2CO_3, and H_2CO_3-CO_3 showed strong positive correlations at α = 0.01 (Table 3). The pairs; pH-Na, pH-NH_3, TDS-Cl, TDS-Na, TDS-F, EC-Cl, EC-Na, Cl-Na, K-Cl, Na-CO_3, SO_4-CO_2, SO_4-HCO_3, SO_4-H_2S, Ca-F, Ca-CO_2, and Ca-NH_4, depicted moderate correlations at α = 0.01 (Table 3).

Principal components extracted: Based on the Cumulative variance of the rotation sums of squared loadings of the wet season, the retained latent constructs account for 91.1% of the variability of the data set (Table 7). PC1 with the largest Eigen value accounted for maximum of the total variability (23.5%). PC2 accounted for the total variation of 22.8% and corresponds in concept to the first PC. The third and fourth PCs explained 22.8% and 22% of the total variance, respectively. The Cattel's Scree test plot is presented in Figure 2.

The first principal component, PC1 (the salinity component) is depicted mainly by K^+, Ca^{+2}, Cl^-, SO^{-2}_4, H_2S, and NH^-_4 ions in water. Since the component explains the largest variance in the data, it can be inferred that the groundwater in the study area is mainly saline. Abnormal distribution Ca^{+2}, SO^{-2}_4, and Cl^- may be associated with the ion exchange mechanisms in saturated aquifer zones. At the wet season pH for instance, the divalent Ca^{+2} is depleted from groundwater as it substitutes for monovalent Na^+ on exchangeable micro-pore surface-water interfaces. The ions SO^{-2}_4 and H_2S depict a strong positive correlation (+0.8 at α = 0.01) indicating their geologic provenance; probably due to deep circulation of oxygen saturated waters causing aerobic conditions. The moderate correlations within the pairs; NH^-_4Ca (+0.7) and NH_4-H_2CO_3 (+0.6) imply anthropogenic contributions to the pH controls on aquifer water salinity. During the wet season, decreases

Table 7: Cumulative variance explained for the wet season data.

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.691	37.938	37.938	3.519	23.460	23.460
2	4.423	29.489	67.427	3.427	22.844	46.304
3	2.273	15.154	82.581	3.420	22.799	69.103
4	1.278	8.521	91.102	3.300	21.999	91.102

Extraction Method: Principal Component Analysis.

in H^+ causes a reduction on the water concentrations of K^+, Ca^{+2}, and Cl^-. Aquifer salinity, therefore, is chiefly a construct of water-rock interactions and to a lesser extent anthropogenic inputs.

The second principal component, PC2 depicts carbonate equilibrium with external pH control. The dissociation of H_2CO_3 yields the ions $CO_3^=$ and H^+ which cause the reduction of sulphates to sulphides. The external factors controlling pH are SO^{-2}_4 and H_2S, which are strongly correlated (+0.8, α = 0.01). Sulphate and sulphide transformations in borehole waters are invariably mediated by bacterial changes under anaerobic conditions. Due to high mobility of oxygen-saturated water in the wet season, reduced dissociation of H_2CO_3 yielding less $CO_3^=$ and H^+ which consequently causes oxidation of sulphides to sulphates. *Desulfovibrio desulfuricans* may also produce H_2S under conditions of the measured pH range (optimum growth range in pH is between 5.5 and 8.5) and optimum temperature between 24 °C and 42 °C. Under these conditions, sulphides in the form of metal sulphide undergo oxidation to sulphates which is again converted to H_2S under the action of H_2CO_3.

The third principal component, PC3 is the alkalinity component as indicated by pH, $CO_3^=$, $H_2CO_3^-$ and F^- and NH^-_4. The pH, $CO_3^=$ and F have positive loading whereas $H_2CO_3^-$ and NH^-_4 have negative loading on this PC. The PC explains the dissolution of fluoride through microbial activity releasing ammonia and thereby decreasing the pH. Additionally, weak acids such as H_2CO_3 ionize sequentially releasing $CO_3^=$, which raises the pH of water.

The fourth PC is the electrolytic conductivity component which is associated with significantly high concentrations of electrolytic ions indicated by TDS, EC, Na^+ and F^-. Statistical analyses (e.g. [2] associate F^- with weathering of the fluoro-apatite and silicate mineralogy, whereas [5] linked Na^+ with the dissolution or weathering of sodium-rich plagioclases (clay mineralogy) or the dissolution of halide. The strong correlation observed between Na and F (+0.9) and the weak association between Na and Cl (-0.6) supports weathering plagioclase feldspars as the chief source as opposed to dissolution of halide. Na and F ions are the intrinsic constructs responsible for electrolytic conduction, as supported by their strong correlations with TDS. Within the hydro-geological framework, the pattern and therefore, provenance of weathering can be accounted for by the roughly N-S fracture-fissure zones. The component accounts for the lowest variability because of reduced residence time of groundwater during the wet season exerting substantial reduction in TDS.

Spatial variability between sampling sites

The results of hierarchical clustering procedures were discrete clusters presented graphically in the form of a dendrogram by an averaging algorithm (Figure 3). Based on rescaled Euclidean distances and the 'sums of squared errors', two borehole clusters (within cluster medium depict translational invariance in composition) are conspicuous in the area. The first cluster (forming the left-hand group) consists of the western zone cases (boreholes 2, 3, 4, 5 and 6) whereas the second cluster (forming the right-hand group) consists of the eastern zone borehole cases (boreholes 1, 7 and 7A). The former boreholes were deciphered to be hydraulically connected by a major inferred NW-SE fault which corresponds to a

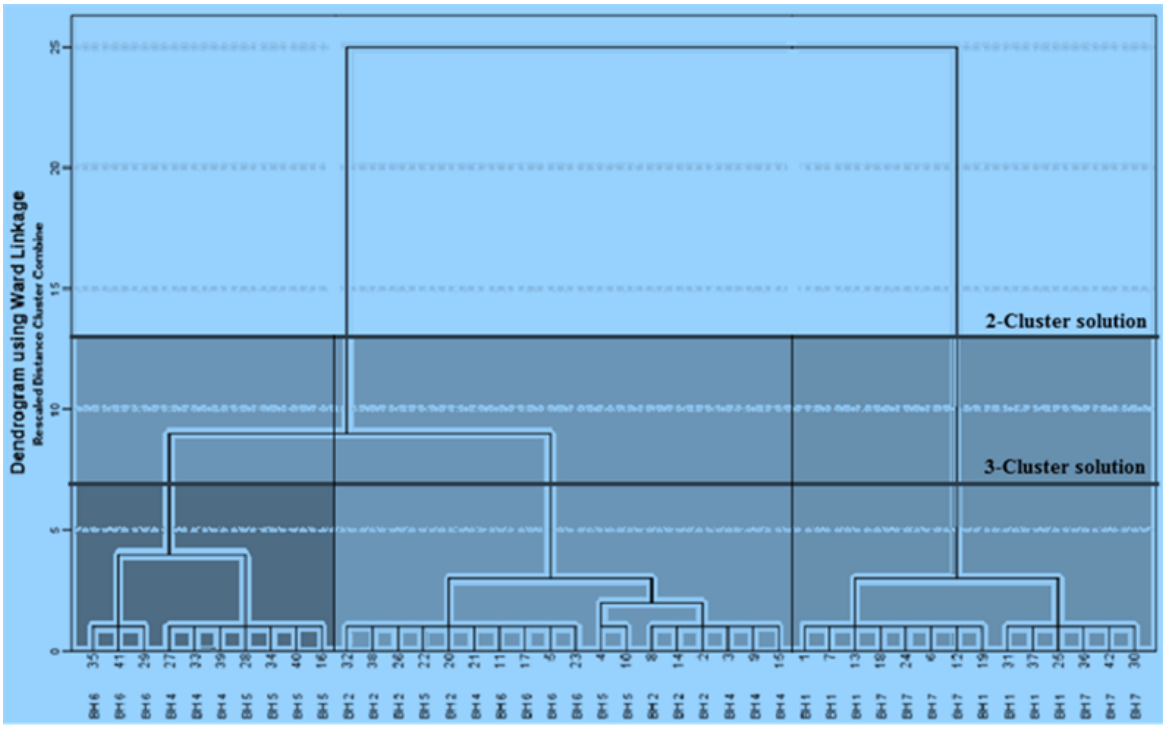

Figure 3: Dendrogram for clustering of groundwater sampling sites using the Euclidean distance metric (Ward's method). Where to 'prune' the tree (e.g. using the continuous bold lines) is a vital factor in interpreting results of the analysis. Alternate shading was introduced to facilitate review.

calibrated transmissivity-formation resistivity model for the aquifer. Intrinsic permeability was empirically elevated along major fracture traces, consequently increasing yields of the affected boreholes. Additionally, some wet season samples obtained from boreholes; 4 (sample Nos. 27 and 33), 5 (sample Nos. 28, 34 and 40) and 6 (sample Nos. 29, 35 and 41) formed a mini-cluster within former main group. The samples represent effects of a high permeability fault/fracture structure which accentuates deep circulation of oxygen saturated waters from recent precipitation events coupled with dilution within the corresponding season. We suggested low residence times of groundwater in the zone.

The latter cluster contains boreholes located in the eastern compartment of the aquifer. Elevated values of Cl^-, SO^{-2}_4, and exceptionally low values of EC, TDS, pH, Na^+, T-CO_2 and HCO^-_3 were recorded in these groundwater boreholes. However, values of EC, TDS, Na^+, SO^{-2}_4, and K^+ depicted an upward trend, whereas those of Ca^{+2}, T-CO_2, $CO^{=}_3$ and HCO^-_3 showed a downward trend from the dry season towards the wet season for these boreholes. By contrast, boreholes in the zone are not hydraulically connected via major fault structures. We therefore suggested that pore-level adsorption/desorption processes in which vast quantities of monovalent ions such as Na^+ and K^+ are added groundwater in exchange for divalent ions control the observed variability in groundwater hydrochemistry in this zone. Lack of aquifer-scale hydraulic networks and presence of clay micropores imply extended groundwater residence time favouring ion exchange reactions in the zone. Borehole 8 drilled in this zone, dried-up after its completion because a clay layer was inadvertently targeted for production.

Conclusions

In this analysis, the joint applicability of two main advanced multivariate statistical methods for aquifer evolutionary structure has been established. PCA technique condensed the two huge data sets of (24 × 15) and (18 × 15) matrices into two matrices, each of (15 × 4), (that is, variables × factors). For the dry season data set, four principal factors were found to explain the highest cumulative variance (90.1%) in water quality. The principal factors were; electrolytic conductivity (34.8%), aquifer salinity (21%), water alkalinity (20.8%), and carbonate equilibrium with external pH (13.6%). On the other hand, the wet season water chemistry explained net variability at 91.1%. The principle processes underlying the variation in water chemistry were identified as aquifer salinity (23.5%), carbonate equilibrium with external pH (22.8%), water alkalinity (22.8%) and electrolytic conductivity (22%).

The dry to wet seasonal shift in variability from electrolytic conduction to salinity, respectively, is probably due to differential seasonal rates of weathering, flow and dilution processes in the aquifer. Significant correlations (at α = 0.01) in the pairs; Cl-SO_4, Ca-K, Ca-NH4, F-NH_4, CO_2-HCO_3, HCO_3-CO_3, Na-F and F-H_2CO_3, indicate that the water is alkaline to mildly acidic which are a manifestation of geogenic and to a lesser scale anthropogenic imprints. Key groundwater evolutionary trends suggest that; silicate, carbonate and/or accessory mineral apatite dissolution as well as ion exchange at sorption sites with clay-water interface are the central sources of variability in the groundwater chemistry of the aquifer. Ward's linkage cluster analysis partitioned the aquifer into two discrete spatial associations; the western and the eastern entities respectively, in spite of their indicated shared recharge area. These agglomerative scheduling validated in an integrative approach (with groundwater flow predictions using a calibrated petro-physical groundwater model for the area), linked the aquifer compartments to three pathways: fault permeability and rock dissolution processes in the western zone, and water-rock interaction through pore-level adsorption/desorption processes in the eastern zone. We incontrovertibly associate the western zone with diminutive residence periods and the eastern zone to protracted residence periods. It is convincingly essential, therefore, based on pH shifts per season (basic in dry and acidic in wet) to design a groundwater quality monitoring plan and policy that reduces the number of measured parameters provides an opportunity cost in terms of resources for measurements elsewhere. A sustainable alternative would be to measure EC during the dry season and K^+, Ca^{2+} and Cl^- during the wet seasons. These parameters can be used as surrogates for the presence of the remaining parameters for the respective periods. The additional analyses may be required during extended dry periods accompanied by upward trend in EC measurements.

Data Availability

The datasets used and generated in this study are available upon request to sosibenjamin@yahoo.com.

Author Contributions

"Conceptualization, B.S. and A.G.; methodology, B.S. and A.G.; software, B.S.; validation, B.S., A.G., S.M. and J.B.; formal analysis, B. S. and A.G.; investigation, B.S.; resources, A.G., S.M. and J.B.; data curation, B.S. and A.G.; writing-B.S., writing-review and editing, B.S., A.G., S. M. and J.B.; visualization, B.S. and A.G.; supervision, A.G., S.M. and J.B.; project administration, B.S. and S.M.; funding acquisition, A.G., S.M. and J.B.."

Funding

"This research was funded by Kenya Government and through the National Research Fund (website: http://www.nrf.go.ke), grant number NRF/R/2016/FORM-1A".

Conflicts of Interest

"The authors declare no conflict of interest". "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

References

1. (2004) Rift valley water supply and sanitation project appraisal report.
2. Pazand K (2016) Geochemistry and multivariate statistical anal-

ysis for fluoride occurrence in groundwater in the Kuhbanan basin, Central Iran. Modeling Earth Systems and Environment 2: 72.

3. Güler C, Thyne GD, Tağa H, et al. (2017) Processes governing alkaline groundwater chemistry within a fractured rock (ophiolitic melange) aquifer underlying a seasonally inhabited headwater area in the Aladağlar range (Adana, Turkey). Geofluids.
4. Nwankwoala HO (2013) Interpretation of hydro hydro-geochemical characteristics of deep aquifers in parts of port harcourt, Eastern Niger Delta. Standard Scientific Research and Essays 1: 154-163.
5. Moghimi H (2017) The study of processes affecting groundwater hydrochemistry by multivariate statistical analysis (Case study: Coastal aquifer of Ghaemshahr, NE-Iran). Open Journal of Geology 7: 830.
6. Esmaeili S, Moghaddam AA, Barzegar R, et al. (2018) Multivariate statistics and hydrogeochemical modeling for source identification of major elements and heavy metals in the groundwater of Qareh-Ziaeddin plain, NW Iran. Arabian Journal of Geosciences 11: 5.
7. Yidana SM, Bawoyobie P, Sakyi P, et al. (2018) Evolutionary analysis of groundwater flow: Application of multivariate statistical analysis to hydrochemical data in the Densu Basin, Ghana. Journal of African Earth Sciences 138: 167-176.
8. Majeed S, Rashid S, Qadir A, et al. (2018) Spatial patterns of pollutants in water of metropolitan drain in Lahore, Pakistan, using multivariate statistical techniques. Environ Monit Assess 190: 128.
9. Smeti EM, Golfinopoulos SK (2016) Characterization of the quality of a surface water resource by multivariate statistical analysis. Analytical Letters 49: 1032-1039.
10. Cloutier V, Lefebvre R, Therrien R, et al. (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. Journal of Hydrology 353: 294-313.
11. Everitt BS, Landau S, Leese M, et al. (2011) Hierarchical clustering. Cluster analysis, 71-110.
12. APHA (2009) Guidelines for quality water. Environmental Protection Agency, 2011, 1-6.
13. Voza D, Vuković M (2018) The assessment and prediction of temporal variations in surface water quality-a case study. Environmental Monitoring and Assessment 190: 434.
14. Joshi RR, Mulay P (2018) Deep incremental statistical closeness factor based algorithm (DIS-CFBA) to assess Diabetes Mellitus. International Research Journal of Engineering and Technology 5: 316-320.
15. Tabachnick BG, Fidell LS, Using multivariate statistics. Boston: Pearson Education Inc.
16. Helsel D (2010) Much ado about next to nothing: Incorporating nondetects in science. Ann Occup Hyg 54: 257-262.
17. Rangeti I, Dzwairo B, Barratt GJ, et al. (2015) Validity and errors in water quality data-A review. In Research and Practices in Water Quality.
18. US Department of Energy (2012) Multivariate statistical analysis of water chemistry in evaluating the origin of contamination in many devils wash, Shiprock, New Mexico.
19. Costello AB, Osborne JW (2005) Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical Assessment Research & Evaluation 10: 1-9.
20. Liu CW, Lin KH, Kuo YM (2003) Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Sci Total Environ 313: 77-89.
21. Mohapatra PK, Vijay R, Pujari PR, et al. (2011) Determination of processes affecting groundwater quality in the coastal aquifer beneath Puri city, India: A multivariate statistical approach. Water Sci Technol 64: 809-817.
22. Hossain M, Patras A, Barry-Ryan C, et al. (2011) Analysis of principal component analysis, hierarchical cluster analysis to classify different spices based on in vitro antioxidant activity and individual polyphenolic antioxidant compound. Journal of Functional Food 3: 179-189.
23. Stanley W, Southam G (2018) The effect of gram-positive (Desulfosporosinus orientis) and gram-negative (Desulfovibrio desulfuricans) sulfate-reducing bacteria on iron sulfide mineral precipitation. Can J Microbiol 64: 629-637.
24. Anantharaman K, Brown CT, Hug LA, et al. (2016) Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nature Communications.

DOI: 10.36959/624/434