16S rRNA Metagenomics of Seminal Fluids from Medical Microbiology Laboratory in a Tertiary Hospital, Southern Nigeria.

Casmir Ndiokwere[1] , Nkechi Augustina Olise[2], Victoria Nmewurum[3], Richard Omoregie[1], Nneka Regina Agbakoba[3], Kingsley C Anukam*[3,4]

1.Medical Microbiology Unit, Medical Laboratory Services, University of Benin Teaching Hospital, Benin City, Nigeria. 2. Department of Medical Laboratory Science, School of Basic Medical Sciences, College of Medical Sciences, University of Benin, Benin City, Nigeria. 3. Department of Medical Laboratory Sciences, Faculty of Health Sciences, Nnamdi Azikiwe University, Nnewi Campus, Anambra State, Nigeria. 4. Uzobiogene Genomics, London, Ontario, Canada.

ABSTRACT

Background: Previous studies have relied on culture-dependent methods to determine microbial communities that may be found in the seminal fluids of men seeking reproductive health care. However, understanding the microbiome composition present in seminal fluids with the state-of-art next-generation sequencing technology is more germane than ever before, instead of culture methods which fails to identify over 99% of bacterial organisms present in biological samples. **Methods**: Forty semen samples were collected and after bacterial DNA extraction, 22 samples that passed quality check were used for amplification of the V4 region of the 16S rRNA using custom bar-coded primers prior to sequencing with Illumina NextSeq 500 platform. Sequencing was performed in a pair-end modality rendering 2 x 150 base-pair sequences. Sequence reads were imported into Illumina BaseSpace Metagenomics pipeline for 16S rRNA recognition. Distribution of taxonomic categories at different levels of resolution was done using Greengenes database.

Results: The taxonomic categories from the dataset produced phyla that ranges from 6 to 25; Class (9-49), Order (16-99), Family (42-214), Genus (55-555) and Species (56-1156). The taxonomic profiles represented 25 phyla, showing 39.5% of the total sequence reads were categorized to *Proteobacteria* as the most abundant. This was followed by *Firmicutes* (33.54%), *Actinobacteria* (20.77%) and *Bacteroidetes* (4.77%), *Fusobacteria* (0.613%), *Tenericutes* (0.31%) and *Verrucomicrobia* (0.12. At the species taxonomic level 1841 species were identified among the seminal fluid samples. *Serratia marcescens* (23.61% sequence reads) was the most abundant species found in 9/22 of the samples followed by *Lactobacillus iners* (18.22%) 13/22, *Serratia entomophila* (5.54%) 17/22, *Haemophilus parainfluenzae* (3.64%) 10/22, *Corynebacterium tuberculostearicum* (3.29%) 21/22, *Gardnerella vaginalis* 2.39%) 12/22, *Lactobacillus taiwanensis* (2.08%) 10/22, *Enterobacter amnigenus* (1.63%) 15/22, *Corynebacterium genitalium* (1.29%) 12/22, *Neisseria lactamica* (1.18%) 8/22, *Finegoldia magna* (1.17%) 16/22, *Prevotella bivia* (1.14%) 10/22, *Corynebacterium imitans* (1.12%) 15/22, *Corynebacterium jeikeium* (1.02%) 17/22 and *Lactobacillus acidophilus* (1.01%) found in 6/22 of the samples. **Conclusions**: We investigated the microbiome compositions from seminal fluids and showed that there are varying bacterial diversities that are unique in each sample in contrast to culture-dependent methods.

Key words: Seminal fluid, microbiome, microbiota, metagenomics, sequencing, 16S rRNA.
***Corresponding authors:** Email: kanukam@gmail.com ; ORCID: 0000-0002-2346-3345

Author's contributions: This work was carried out and approved in collaboration between all the authors. KCA, NRA and CN designed the study; KCA sourced for funding, wrote the protocol and did literature search; CN, NAO and RO collected the samples; VN did taxonomic organization; KCA did bioinformatic analysis, drafted the manuscript, supervised the study and wrote the final manuscript; NRA and KCA proofread the manuscript.
Received: September/15, 2019; **Accepted**: November/30, 2019; **Published**: December 28, 2019.

Citation: Ndiokwere C , Olise NA, Nmewurum V, Omoregie R, Agbakoba NR, Anukam KC. 16S rRNA Metagenomics of Seminal Fluids from Medical Microbiology Laboratory in a Tertiary Hospital, Southern Nigeria**. J Med Lab Sci, 2019; 29 (3): 86-109**

INTRODUCTION

Routine culture-dependent methods are usually employed to isolate and determine microbial compositions of seminal fluids submitted to the microbiology laboratory of teaching hospitals. These seminal fluids originate from men who are seeking reproductive health care as a result of either primary or secondary infertility, and from other useful clinical pathologies. Male infertility has been attributed to several factors, including but not limited to what is referred to as "male factors" comprising of infections affecting the genito-urinary tracts[1], hormonal imbalance, age factor, stress, environmental pollution, some metabolic disorders[2, 3]

Several microbiologists including laboratory scientists at the teaching hospital have at many times tried to understand microbial organisms that are present in seminal fluids. Seminal fluid samples are normally cultured on blood agar, chocolate agar, MacConkey agar, nutrient agar, and sabouraud dextrose agar slants and incubated aerobically and in 5% CO2 at 37°C for 24 to 48 hours. It has been established that culture-dependent methods are now considered obsolete in the 21st century microbial identification as it fails to account for over 99.9% of the microbial compositions present in biological and environmental samples[4].

Some faculty research members have attempted to isolate bacterial organisms present in semen. Onemu and Ibeh[5] previously reported by culture methods that *Staphylococcus aureus* constituted 43.7%, followed by *Klebsiella species* (28.2%), *Escherichia coli* (11.5%), and *Candida albicans* (7.7%). Parallel reports were published subsequently by Momoh *et al*[6], Ibadin *et al*[7], Ekhaise and Richard[8]. Other authors from nearby institutions have presented similar findings whereby *Staphylococcus aureus* (53%), *Staphylococcus saprophyticus* (10%), *Escherichia coli* (11.4%), *Klebsiella spp* (7.1%), *Streptococcus pneumoniae* (4.4%), *Pseudomonas aeruginosa* (7.1%) and *Candida spp* (7.1%) were mainly incriminated[9, 10]. Enwaru *et al* [11] reported 49.4% Gram positive and 21% Gram negative were isolated and *Staphylococcus aureus* (29.6%) and *Escherichia coli* (10.5%) had the highest occurrence for each group respectively.

The objectives of this study are based out of academic/professional curiosity by the laboratory scientists that work on these seminal fluid samples and the opportunity of using culture-independent next-generation high throughput sequencing technology to determine microbial compositions of seminal fluid samples submitted for culture in the laboratory. We compared bacterial organisms isolated by culture methods with the 16S rRNA metagenomics sequencing results of the same samples.

MATERIALS AND METHODS

Ethics approval: Ethics approval was sought and approved at Nnamdi Azikiwe University Teaching hospital, Nnewi as similar study involving large scale metagenomics sequencing from couples (Semen and vaginal fluid) seeking reproductive health care is ongoing.

Study samples:

Seminal fluid samples that were left over in the laboratory after culturing for bacterial isolation were used. Forty semen samples were collected and 100 microlitre from each sample was inoculated into a tube containing lysis and stabilization buffer that preserves the DNA for transport at ambient temperature. The tubes were sent to uBiome Inc. in California, United States America for DNA extraction and sequencing. The

sequencing results were analyzed with bioinformatic tools at Uzobiogene Genomics, London, Ontario, Canada.

Extraction of bacterial DNA from Seminal fluid samples and Sequencing of the amplified 16S rRNA region:
Bacterial DNA was extracted from the seminal fluids using an in-house protocol developed by uBiome Inc. Briefly, samples were lysed using bead-beating, and DNA was extracted in a class 1000 clean room by a guanidine thiocyanate silica column-based purification method using a liquid-handling robot. PCR amplification of the 16S rRNA genes was performed with primers containing universal primers amplifying the V4 region (515F: GTGCCAGCMGCCGCGGTAA and 806R: GGACTACHVGGGTWTCTAAT). In addition, the primers contained Illumina tags and barcodes. DNA samples were barcoded with a unique combination of forward and reverse indexes allowing for simultaneous processing of multiple samples. PCR products were pooled, column-purified, and size-selected through microfluidic DNA fractionation. Consolidated libraries were quantified by quantitative real-time PCR using the Kapa Bio-Rad iCycler qPCR kit on a BioRad MyiQ before loading into the sequencer. Sequencing was performed in a pair-end modality on the Illumina NextSeq 500 platform rendering 2 x 150 bp pair-end sequences. The sequencer has a flow cell with four lanes. This means that each sample was read in four different lanes (L001 to L004), and each produced forward (R1) and reverse (R2) reads.

Metagenomics sequence analysis: Raw sequence reads were demultiplexed using Illumina's BCL2FASTQ algorithm. Reads were filtered using an average Q-score > 30. The 8 paired-end sequence FASTQ reads for each sample were imported into MG-RAST pipeline for quality check (QC). Quantitative Insights into Microbial Ecology (QIIME) pipeline was used for 16S rRNA recognition. Sequences were pre-screened using QIIME-UCLUST algorithms for at least 97% identity to ribosomal sequences from the RNA databases. Reads passing all above filters were aligned to the database of 16S rRNA gene sequences. Microbial taxonomy to species level was generated using the Illumina BaseSpace Greengenes database.

RESULTS

We present the 16S rRNA metagenomics datasets from the twenty-two seminal fluid samples that passed quality check. The eight sequence reads for each sample produced an average 18,673,194 base pair count per read containing 67,395 sequences ranging from 32 bp to 151 bp and averaging 150bp in length (std. deviation from average length 4.720). All of the sequence reads have unique identities. Sequence reads that passed quality check showed that the seminal fluid samples had sequence taxonomic recognition as shown in **Figure 1**. The phyla categories ranges from 6 to 25; Class (9-49), Order (16-99), Family (42-214), Genus (55-555) and Species (56-1156).

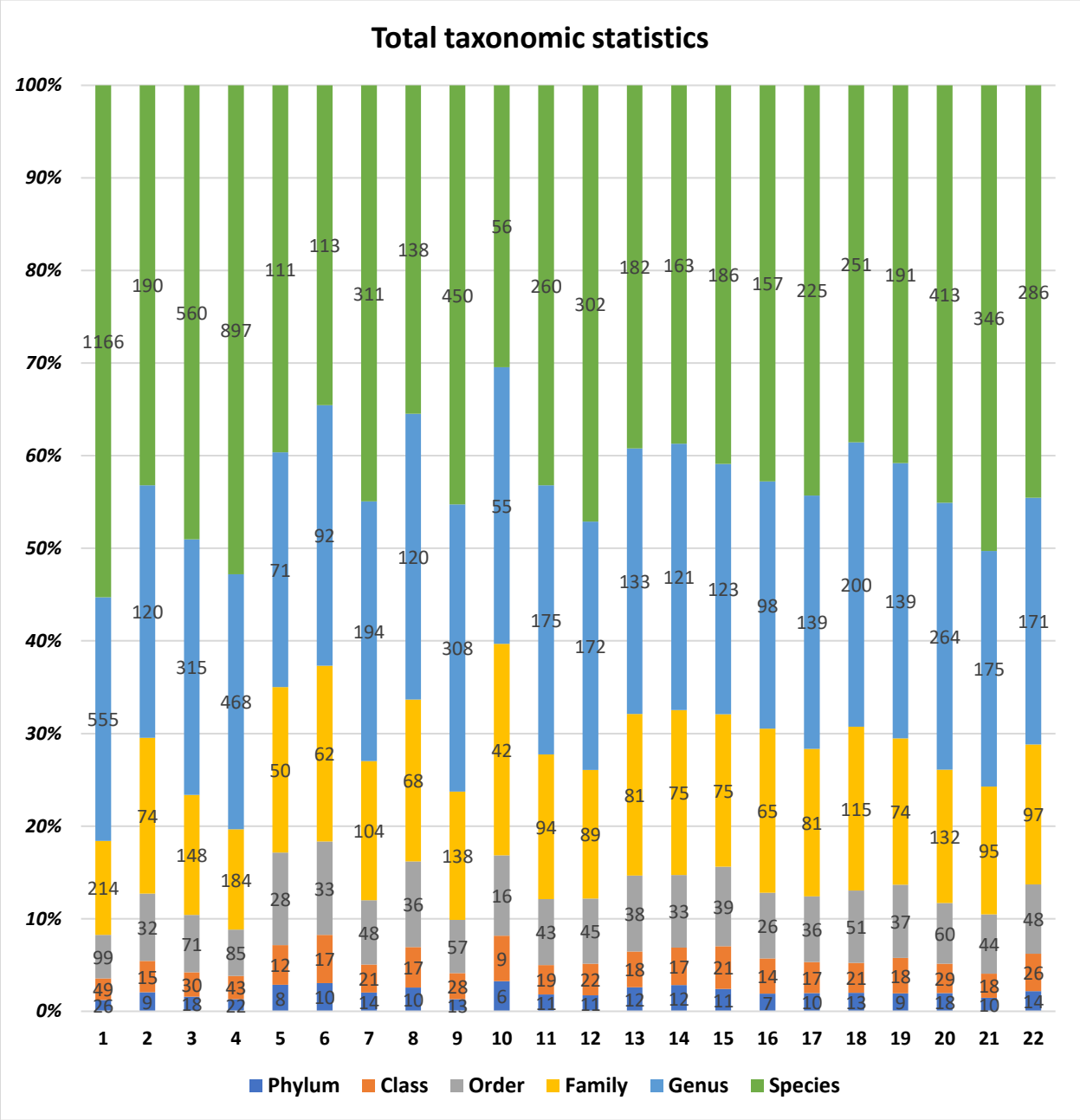

Figure 1: Showing the total taxonomic statistics from sample 1 to sample 22.

The taxonomic profile represented 25 phyla, showing 39.5% of the sequence reads were categorized to *Proteobacteria* as the most abundant **(Figure 2**). This was followed by *Firmicutes* (33.54%), *Actinobacteria* (20.77%) and *Bacteroidetes* (4.77%), *Fusobacteria* (0.613%), *Tenericutes* (0.31%), *Verrucomicrobia* (0.12%) and others.

At the family taxonomic categories, *Enterobacteriaceae* (29.9%) was the most relative abundant, followed by *Lactobacillaceae* (18.17%), *Corynebacteriaceae* (11.64%), *Pasteurellaceae* (4.46%), *Staphylococcaceae* (4.16%), *Prevotellaceae* (3.08%), *Neisseriaceae* (2.66%), *Clostridiaceae* (2.59%), *Streptococcaceae* (1.95%), *Bifidobacteriaceae* (1.89%), *Micrococcaceae* (1.62%), Lachnospiraceae (1.52%), *Veillonellaceae* (1.39%), *Brevibacteriaceae* (1.05%) and others as presented in **Table 1**

.

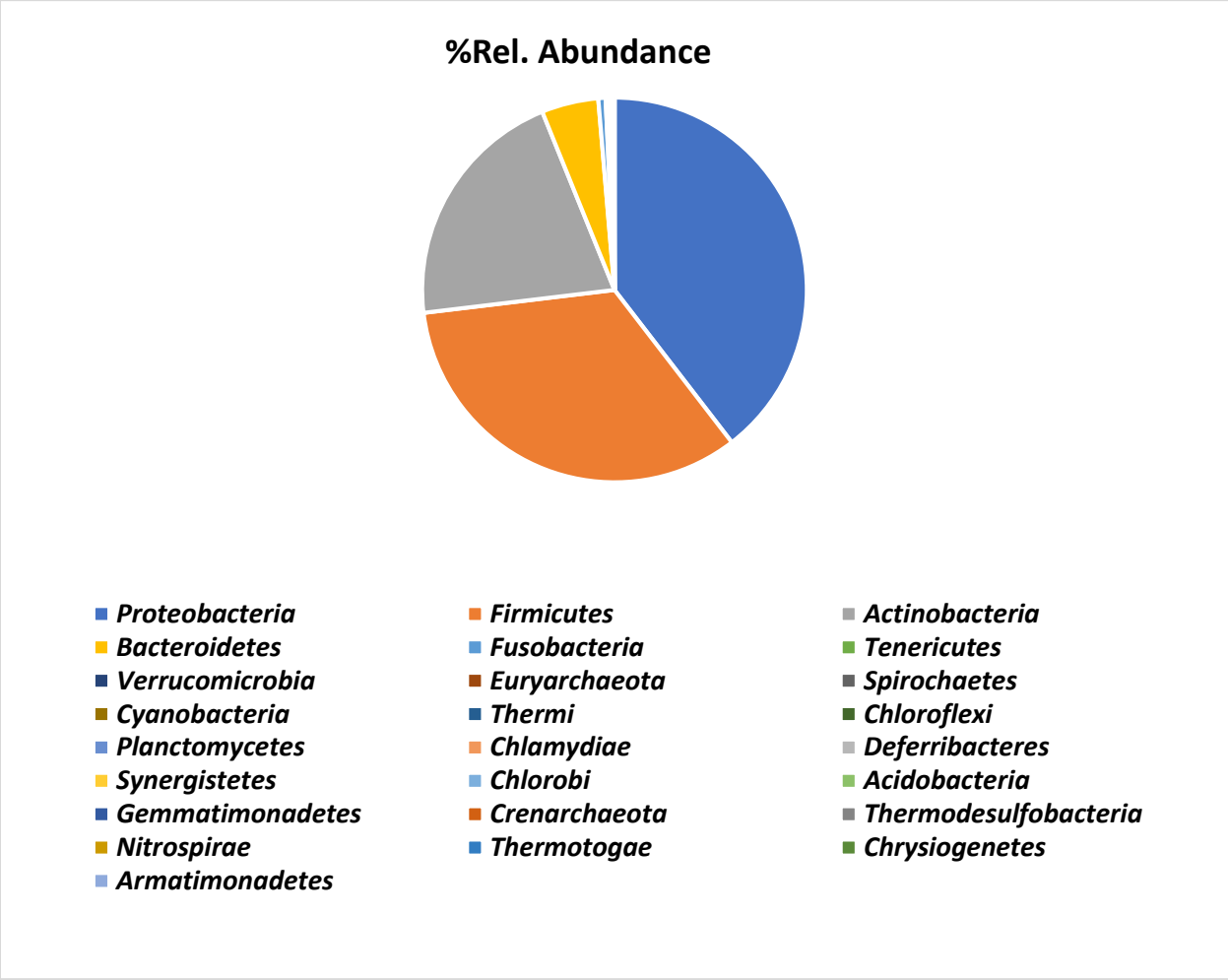

Figure 2: Showing the % relative abundance on bar chart for the 25 phyla identified in the seminal fluid samples.

Table 1: Showing all the 249 family taxa identified in semen samples in % relative abundance.

Taxa/Family	% Rel abundance	Family	% rel abundance	Taxa Family	% rel abundance
Enterobacteriaceae	29.91568462	Nocardioidaceae	0.092441332	Nostocaceae	0.007658768
Lactobacillaceae	18.17280172	Xanthomonadaceae	0.092211569	Pseudanabaenaceae	0.007505593
Corynebacteriaceae	11.64125104	Carnobacteriaceae	0.091675455	Deferribacteraceae	0.006969479
Pasteurellaceae	4.464372559	Rhodobacteraceae	0.088382185	Exiguobacteraceae	0.006739716
Staphylococcaceae	4.161851216	Methylobacteriaceae	0.077736497	Kineosporiaceae	0.006586541
Prevotellaceae	3.085487936	Paenibacillaceae	0.076051568	Bradyrhizobiaceae	0.006356778
Neisseriaceae	2.662111231	Gemellaceae	0.069082089	Coprobacillaceae	0.00628019
Clostridiaceae	2.598620043	Erysipelotrichaceae	0.067243985	Phyllobacteriaceae	0.006203602
Streptococcaceae	1.956049393	Spirochaetaceae	0.066018582	Waddliaceae	0.006203602
Bifidobacteriaceae	1.891103039	Desulfonatronumacea	0.065941994	Oceanospirillaceae	0.006127015
Micrococcaceae	1.626339423	Leuconostocaceae	0.062955074	Legionellaceae	0.005820664
Lachnospiraceae	1.528996479	Chromatiaceae	0.060887207	Rhodothermaceae	0.005744076
Veillonellaceae	1.391445003	Coriobacteriaceae	0.058436401	Yaniellaceae	0.005667488
Brevibacteriaceae	1.057829061	Oxalobacteraceae	0.053228439	Chlorobiaceae	0.005514313
Halomonadaceae	0.818875494	Flexibacteraceae	0.04526332	Ectothiorhodospiraceae	0.005361138
Propionibacteriaceae	0.718392455	Caulobacteraceae	0.044574031	Synergistaceae	0.00528455
Moraxellaceae	0.694037572	Planococcaceae	0.038447016	Beijerinckiaceae	0.005131375
Aerococcaceae	0.660185817	Streptomycetaceae	0.036762087	Streptosporangiaceae	0.005054787
Bacteroidaceae	0.646476622	Acetobacteraceae	0.03140095	Thermoactinomycetaceae	0.004748436
Enterococcaceae	0.616224488	Micromonosporaceae	0.028337442	Isosphaeraceae	0.004518673
Dermabacteraceae	0.59830297	Dietziaceae	0.02213384	Sinobacteraceae	0.004442086
Flavobacteriaceae	0.487021069	Thermoanaerobactera	0.021750902	Thiotrichaceae	0.004442086
Nocardiopsaceae	0.456845522	Dermacoccaceae	0.021597726	Borreliaceae	0.003982559
Cellulomonadaceae	0.446276422	Deinococcaceae	0.020602086	Thermogemmatisporaceae	0.003905972
Leptotrichiaceae	0.439843057	Rivulariaceae	0.020372323	Listeriaceae	0.003676209
Porphyromonadaceae	0.430193009	Syntrophobacteraceae	0.020219148	Thermomonosporaceae	0.003676209
Actinomycetaceae	0.381023717	Rhizobiaceae	0.018763982	Piscirickettsiaceae	0.003523033
Rubrobacteraceae	0.316766652	Symbiobacteriaceae	0.018763982	Caldilineaceae	0.003446446
Mycoplasmataceae	0.307805893	Brucellaceae	0.018534219	Entomoplasmataceae	0.00329327
Bacillaceae	0.304359447	Rhodospirillaceae	0.017921518	Aurantimonadaceae	0.003216683
Pseudonocardiaceae	0.300759826	Erythrobacteraceae	0.016542939	Gemmatimonadaceae	0.003140095
Peptostreptococcaceae	0.295092338	Tsukamurellaceae	0.016160001	Solirubrobacteraceae	0.003140095
Peptococcaceae	0.278855749	Acholeplasmataceae	0.016006825	Vibrionaceae	0.003140095
Alcaligenaceae	0.25901954	Nocardiaceae	0.015394124	Glycomycetaceae	0.00298692
Pseudomonadaceae	0.220266173	Geodermatophilaceae	0.014934598	Eubacteriaceae	0.002833744
Intrasporangiaceae	0.206020864	Halanaerobiaceae	0.014015546	Alcanivoracaceae	0.002603981
Fusobacteriaceae	0.198055745	Salinisphaeraceae	0.013173081	Desulfuromonadaceae	0.002603981
Ruminococcaceae	0.180670341	Anaerobrancaceae	0.012024266	Cardiobacteriaceae	0.002527393
Campylobacteraceae	0.167573848	Rhodocyclaceae	0.011947678	Bdellovibrionaceae	0.002450806
Comamonadaceae	0.150111856	Hyphomicrobiaceae	0.011794503	Euzebyaceae	0.002450806
Sphingomonadaceae	0.146741998	Actinosynnemataceae	0.01156474	Sporolactobacillaceae	0.002450806
Microbacteriaceae	0.143984842	Williamsiaceae	0.010645688	Desulfohalobiaceae	0.002374218
Verrucomicrobiaceae	0.126369675	Idiomarinaceae	0.009956399	Gemmataceae	0.002374218
Gordoniaceae	0.119017258	Desulfovibrionaceae	0.009190522	Saccharospirillaceae	0.00229763
Sphingobacteriaceae	0.115723987	Actinopolysporaceae	0.008730996	Solibacteraceae	0.00229763
Aeromonadaceae	0.11266048	Alteromonadaceae	0.008654408	Dermatophilaceae	0.002221043
Shewanellaceae	0.111128726	Amoebophilaceae	0.008501233	Turicibacteraceae	0.002221043
Halobacteriaceae	0.10967356	Chitinophagaceae	0.008348057	Caldicellulosiruptoraceae	0.002144455
Burkholderiaceae	0.106686641	Methylophilaceae	0.00827147	Hyphomonadaceae	0.002144455
Mycobacteriaceae	0.105154887	Bogoriellaceae	0.008194882	Conexibacteraceae	0.002067867
Paraprevotellaceae	0.100329863	Thermicanaceae	0.008194882	Methylocystaceae	0.002067867

Taxa Family	% rel abundance	Taxa Family	% rel abundance
Bartonellaceae	0.00199128	Balneolaceae	0.000536114
Phormidiaceae	0.00199128	Desulfobacteraceae	0.000536114
Anaplasmataceae	0.001914692	Hahellaceae	0.000459526
Xanthobacteraceae	0.001914692	Nannocystaceae	0.000459526
Rickettsiaceae	0.001838104	Rhodobiaceae	0.000459526
Sulfobacillaceae	0.001838104	Anaeroplasmataceae	0.000382938
Helicobacteraceae	0.001761517	Brachyspiraceae	0.000382938
Kiloniellaceae	0.001761517	Myxococcaceae	0.000382938
Patulibacteraceae	0.001684929	Pedosphaeraceae	0.000382938
Alicyclobacillaceae	0.001608341	Pelagicoccaceae	0.000382938
Geobacteraceae	0.001608341	Psychromonadaceae	0.000382938
Oscillochloridaceae	0.001608341	Chroococcaceae	0.000306351
Acidimicrobiaceae	0.001531754	Dehalobacteriaceae	0.000306351
Flammeovirgaceae	0.001531754	Halobacteroidaceae	0.000306351
Puniceicoccaceae	0.001455166	Methanobacteriaceae	0.000306351
Scytonemataceae	0.001455166	Nitrospiraceae	0.000306351
Pirellulaceae	0.001378578	Thermotogaceae	0.000306351
Sanguibacteraceae	0.001378578	Chrysiogenaceae	0.000229763
Carboxydocellaceae	0.001301991	Dethiosulfovibrionaceae	0.000229763
Heliobacteriaceae	0.001301991	Gomphosphaeriaceae	0.000229763
Promicromonosporaceae	0.001225403	Halothiobacillaceae	0.000229763
Acaryochloridaceae	0.001148815	Holophagaceae	0.000229763
Coxiellaceae	0.001148815	Hydrogenophilaceae	0.000229763
Brocadiaceae	0.001072228	Microcystaceae	0.000229763
Haliangiaceae	0.001072228	Rikenellaceae	0.000229763
Planctomycetaceae	0.001072228	Sphaerochaetaceae	0.000229763
Cyanobacteriaceae	0.00099564	Succinivibrionaceae	0.000229763
Gallionellaceae	0.00099564	Aminiphilaceae	0.000153175
Syntrophaceae	0.00099564	Catenulisporaceae	0.000153175
Anaerolinaceae	0.000919052	Chthoniobacteraceae	0.000153175
Cystobacteraceae	0.000919052	Francisellaceae	0.000153175
Odoribacteraceae	0.000919052	Koribacteraceae	0.000153175
Pseudoalteromonadaceae	0.000919052	Microviridae	0.000153175
Acidobacteriaceae	0.000842464	Chloroherpetaceae	7.65877E-05
Desulfomicrobiaceae	0.000842464	Chthonomonadaceae	7.65877E-05
Leptospiraceae	0.000842464	Cohaesibacteraceae	7.65877E-05
Bacteriovoracaceae	0.000765877	Desulfurellaceae	7.65877E-05
Dehalococcoidaceae	0.000765877	Ferrimonadaceae	7.65877E-05
Desulfobulbaceae	0.000765877	Frankiaceae	7.65877E-05
Opitutaceae	0.000765877	Iamiaceae	7.65877E-05
Polyangiaceae	0.000765877	Nitrosomonadaceae	7.65877E-05
Saprospiraceae	0.000765877	Oscillatoriaceae	7.65877E-05
Syntrophomonadaceae	0.000765877	Rarobacteraceae	7.65877E-05
Synechococcaceae	0.000689289	Thermodesulfovibrionaceae	7.65877E-05
Thermobaculaceae	0.000689289	Thiobacteraceae	7.65877E-05
Methylococcaceae	0.000612701		
Nitrososphaeraceae	0.000612701		
Rhabdochlamydiaceae	0.000612701		
Thermaceae	0.000612701		
Thermodesulfobacteriaceae	0.000612701		
Thiohalorhabdaceae	0.000612701		

At the genera taxonomic categories, 725 genera were identified in this cohort of men. Interestingly *Serratia* was identified in 17/22 (77.27%) of the seminal fluid samples as the most abundant representing 24.12% of the total sequence reads. This was followed by *Lactobacillus* found in 20/22 (90.90%) of the samples representing 18.89% of the total sequence reads (**Figure 3**). *Corynebacterium* (12.26%), and *Staphylococcus* (3.98%), *Propionibacterium* (0.677%), *Acinetobacter* (0.668%) were identified in all the samples.

Bacterial vaginosis associated genera such as *Prevotella* (3.98%) was found in 17/22 samples, *Gardnerella* (1.95%) found in 12/22 samples; *Haemophilus* (3.11%):12/22; *Neisseria* (2.78%):15/22; *Streptococcus* (2.05%):21/22; *Veillonella* (1.26%):16/22; *Finegoldia* (1.20%):17/22; *Peptoniphilus* (0.88%):20/22; *Escherichia* (0.65%):19/22; *Anaerococcus* (0.40%):19/22; and *Sneathia* (0.32%):8/22.

.

Figure 3: Showing the most relative abundant genera from 0.1% and above.

At the species taxonomic level 1841 species were identified among the seminal fluid samples. *Serratia marcescens* (23.61% sequence reads) was the most abundant species found in 9/22 of the samples **(Figure 4**) followed by *Lactobacillus iners* (18.22%) 13/22, *Serratia entomophila* (5.54%) 17/22, *Haemophilus parainfluenzae* (3.64%) 10/22, *Corynebacterium tuberculostearicum* (3.29%) 21/22, *Gardnerella vaginalis*

2.39%) 12/22, *Lactobacillus taiwanensis* (2.08%) 10/22, *Enterobacter amnigenus* (1.63%) 15/22, *Corynebacterium genitalium* (1.29%) 12/22, *Neisseria lactamica* (1.18%) 8/22, *Finegoldia magna* (1.17%) 16/22, *Prevotella bivia* (1.14%) 10/22, *Corynebacterium imitans* (1.12%) 15/22, *Corynebacterium jeikeium* (1.02%) 17/22 and *Lactobacillus acidophilus* (1.01%) found in 6/22 of the samples

.

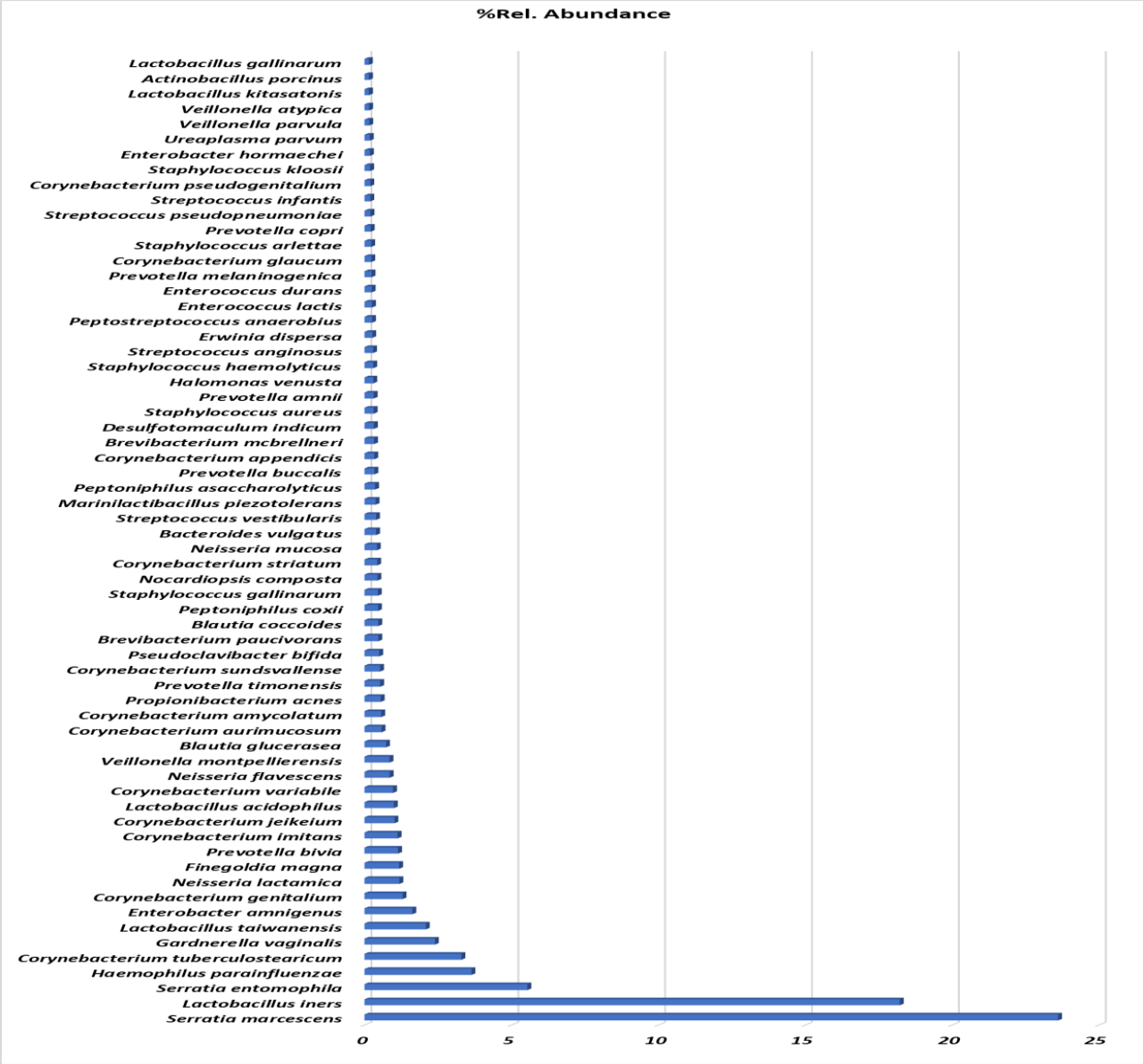

Figure 4: Showing the most % relative abundant species from 0.15% and above.

Taxonomic categories that occurred from 0.1476% to 0.0615% relative abundance are represented in **Figure 5**. At the individual genera, 5 species of *Serratia* were identified which include *Serratia entomophila*, *Serratia marcescens*, *Serratia nematodiphila*, *Serratia symbiotica*, and *Serratia ureilytica*.

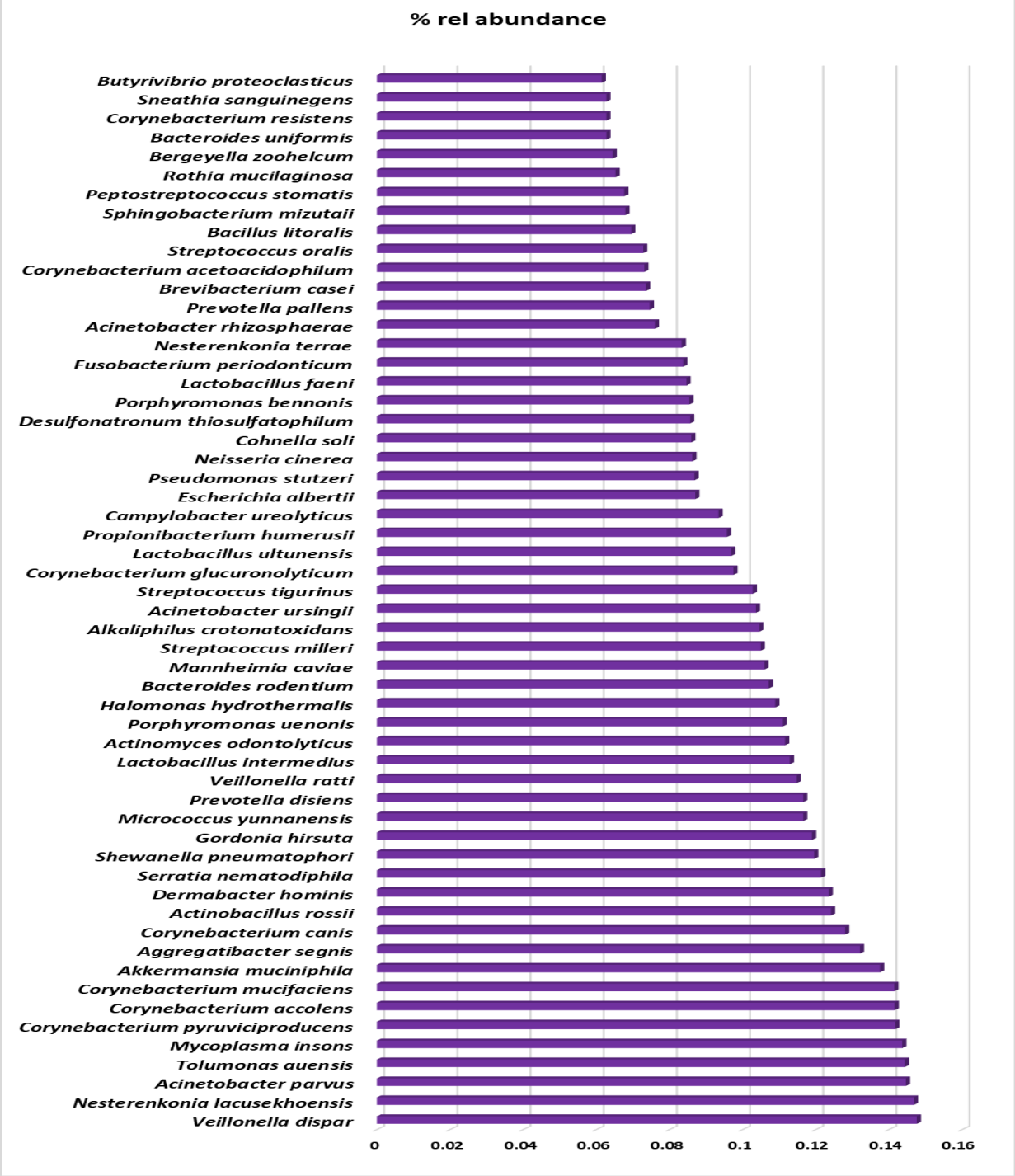

Figure 5: Showing the species that occurred from 0.1476% to 0.0615% relative abundance

Lactobacillus had 45 species among these samples shown in **Figure 6** with *Lactobacillus iners* predominating.

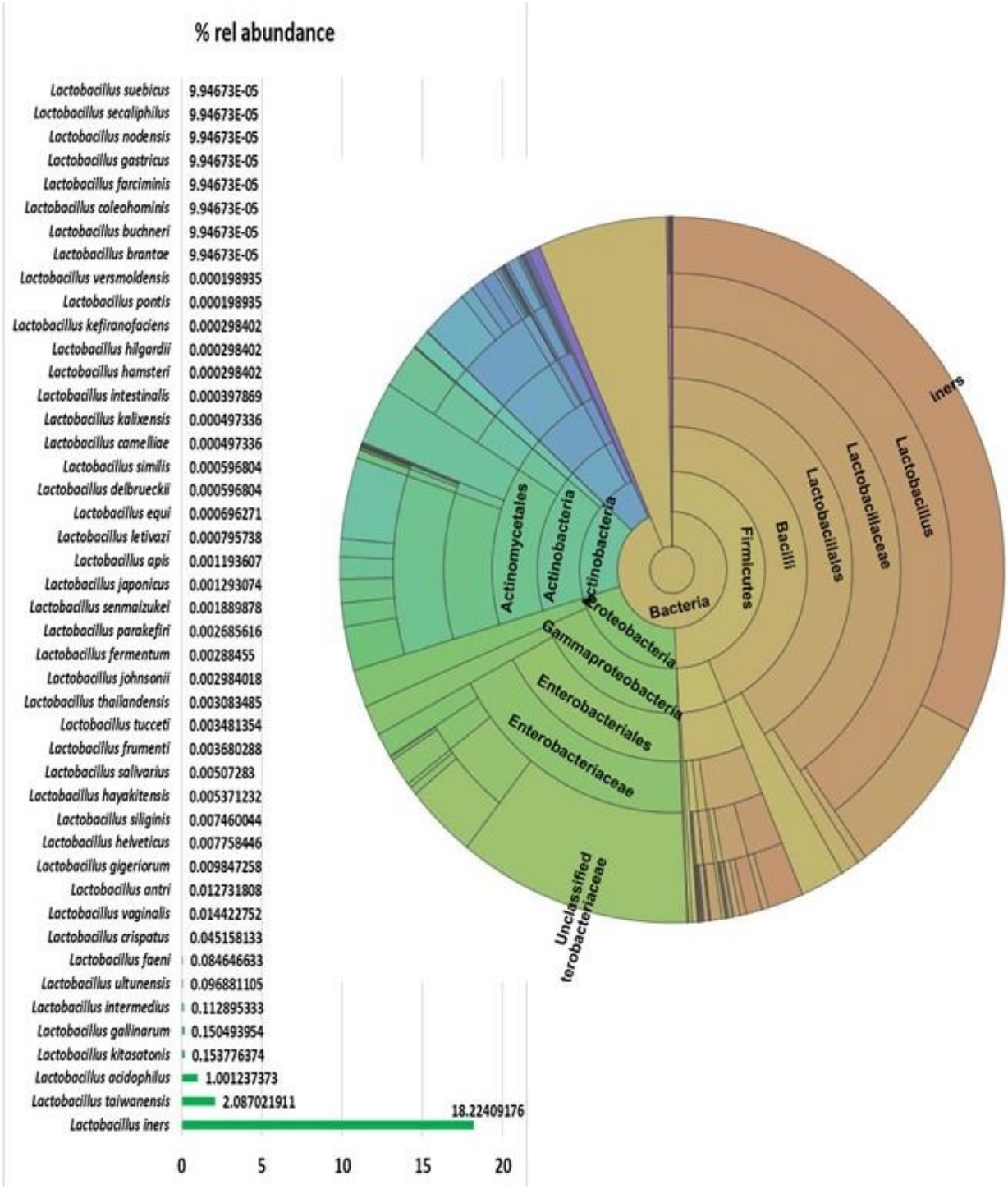

Figure 6: Showing the relative abundance of all the Lactobacillus species identified and sunburst chart within each taxonomic level.

Corynebacterium genera had 61 species identified in the samples **(Figure 7**).

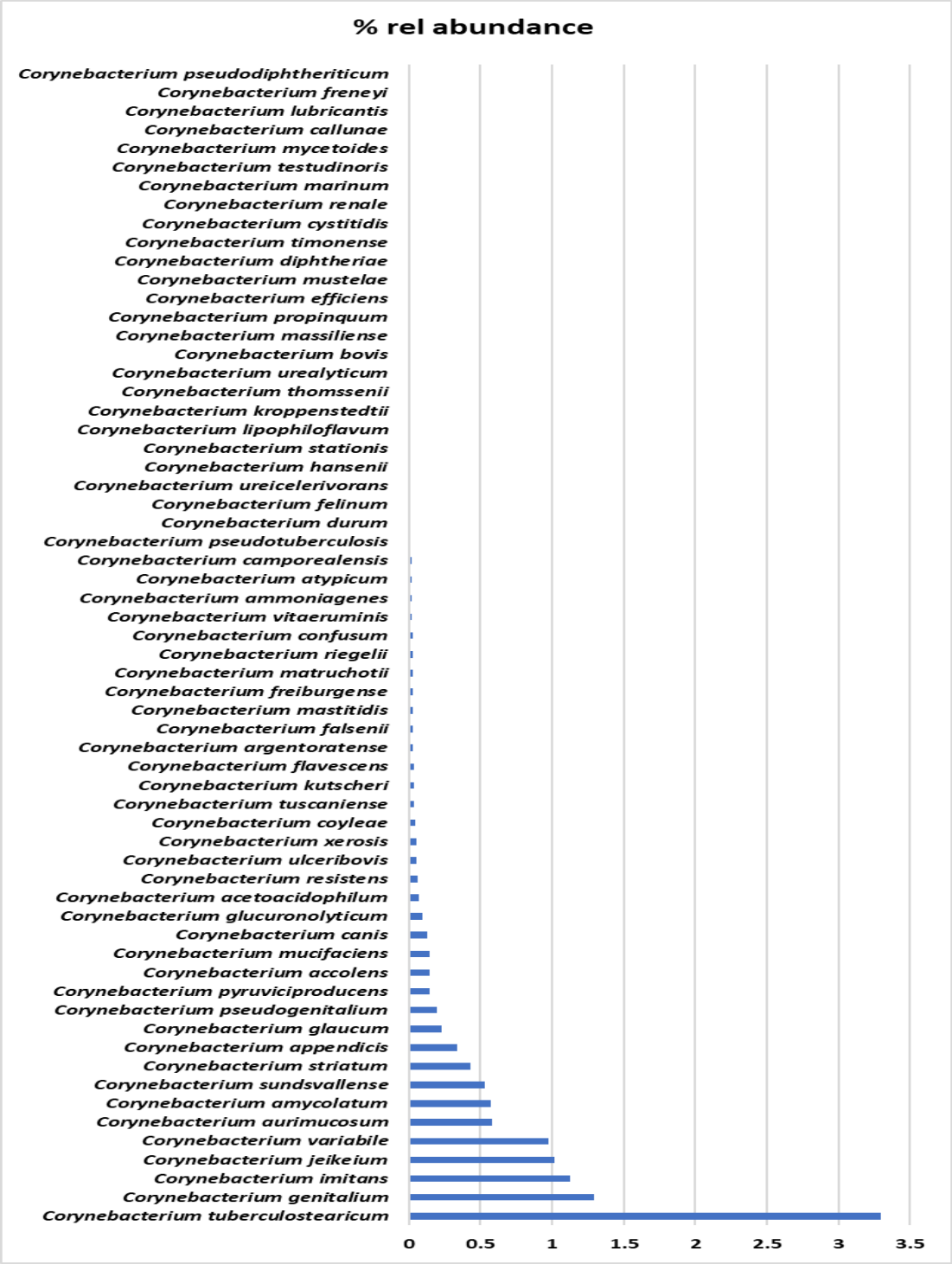

Figure 7: Showing % relative abundance of all the 61 *Corynebacterium* species in the semen samples.

Staphylococcus had 32 species with *Staphylococcus gallinarum* (0.45%) 19/22 predominating followed by *Staphylococcus aureus* (0.31%) 19/22, *Staphylococcus haemolyticus* (0.29%) 19/22, *Staphylococcus arlettae* (0.22%) 18/22, *Staphylococcus kloosii* (0.19%) 15/22 and others (**Figure 8).**

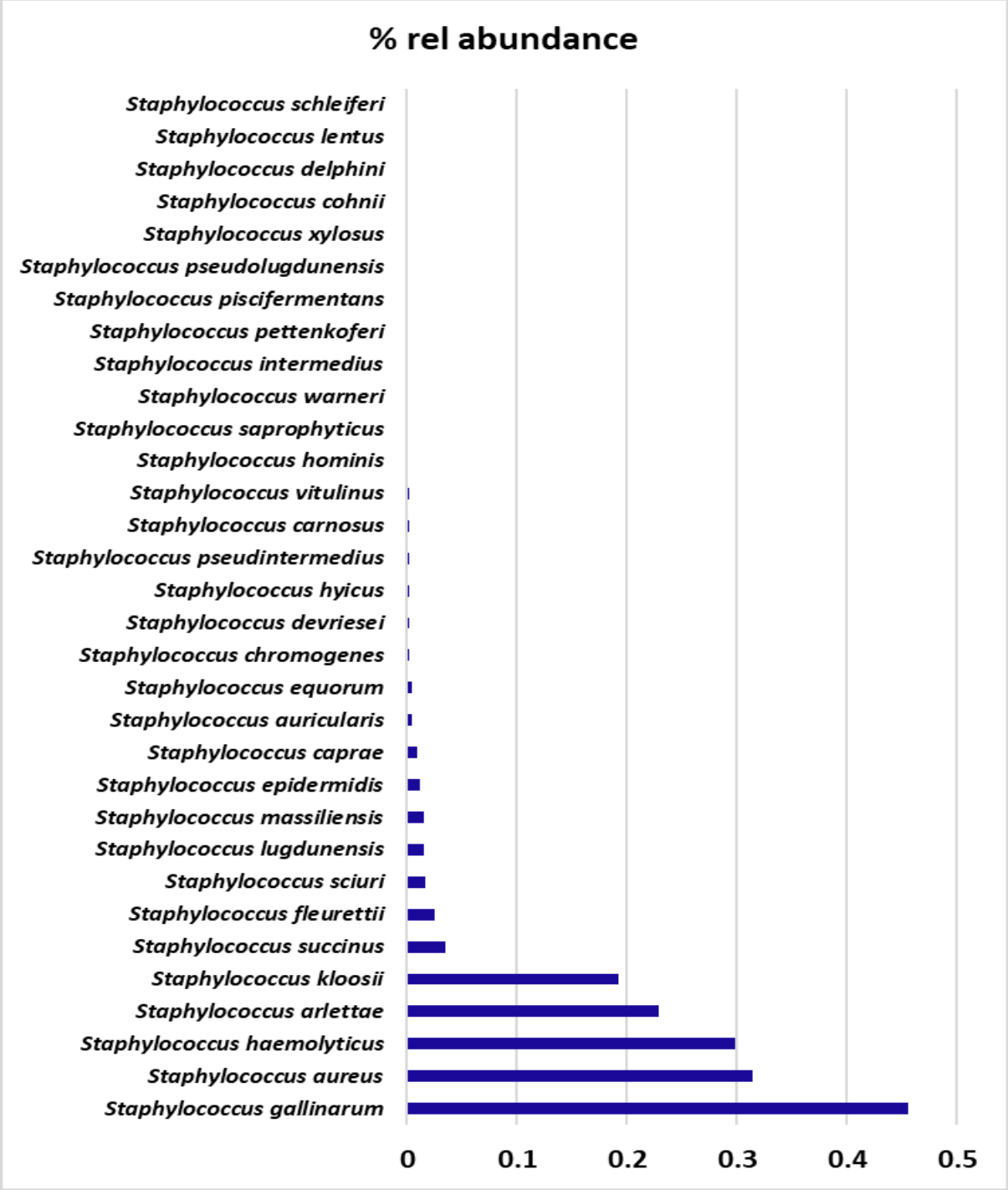

Figure 8: Showing the % relative abundance of all the 32 identified *Staphylococcus* species

Prevotella genera had 37 species with *Prevotella bivia* (1.14%) 10/22 as the most abundant, followed by *Prevotella timonensis* (0.53%) 12/22, *Prevotella buccalis* (0.33%) 8/22, *Prevotella amnii* (0.31%) 9/22, *Prevotella melaninogenica* (0.23%) 7/22,

Prevotella copri (0.21%) 5/22 and others (**Figure 9**).

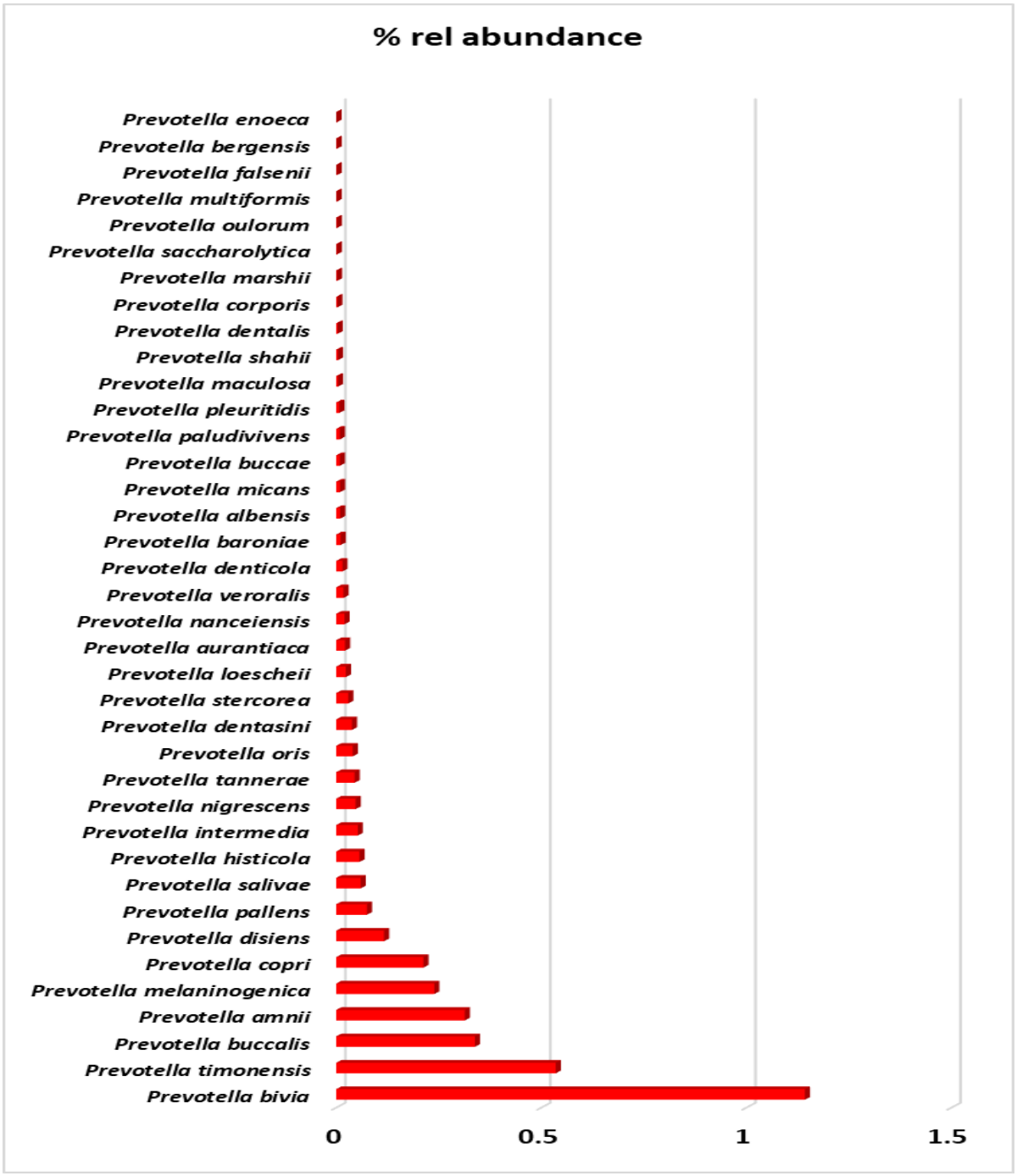

Figure 9: Showing the % relative abundance of all the 37 *Prevotella* species

Likewise, *Streptococcus* had 37 species with *Streptococcus vestibularis* (0.394%) 6/22, followed by *Streptococcus anginosus* (0.291%) 6/22, *Streptococcus pseudopneumoniae* (0.204%) 4/22, *Streptococcus infantis* (0.199%) 12/22, *Streptococcus milleri* (0.104%) 11/22 and others as shown in **Figure 10**

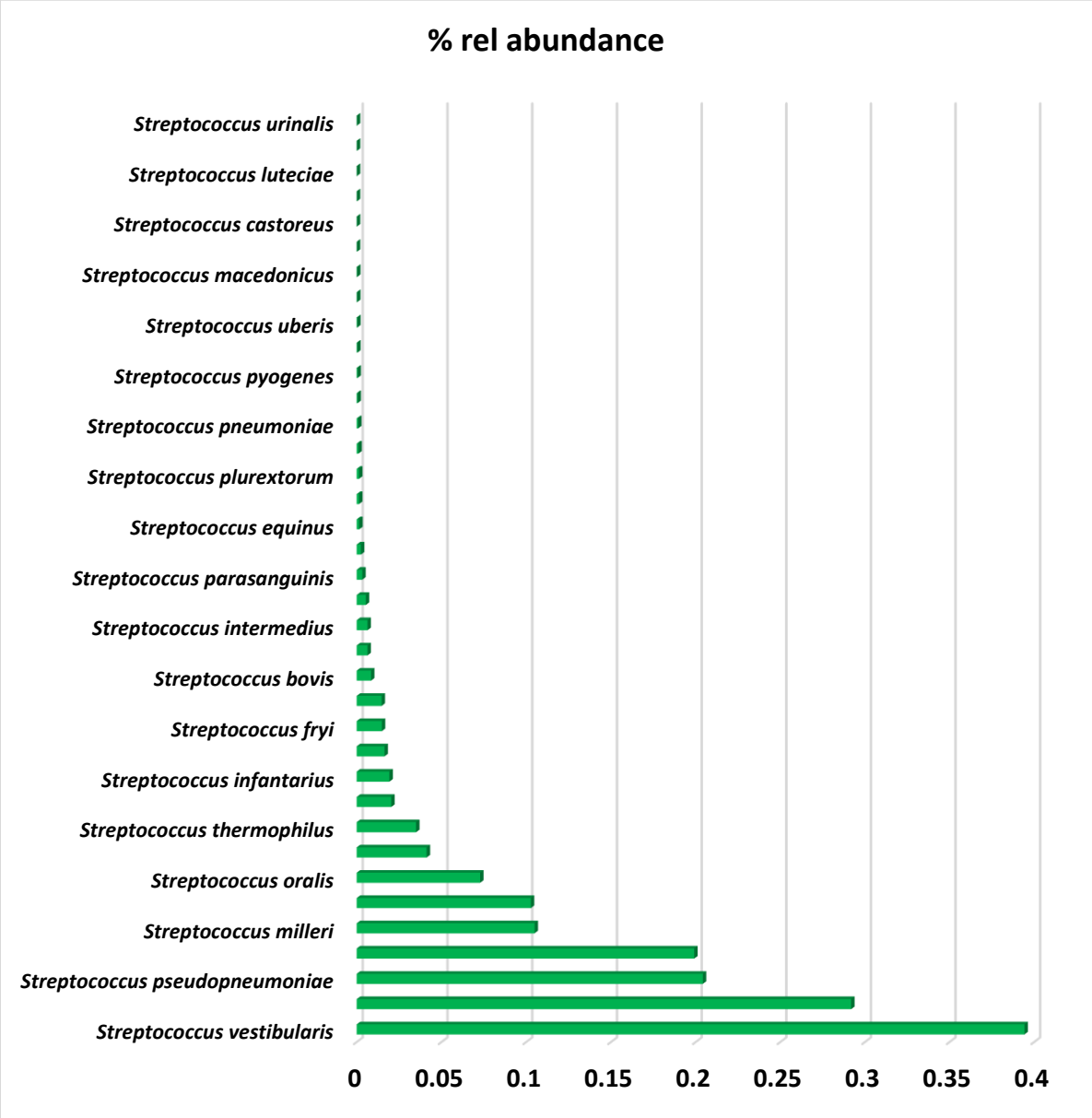

Figure 10: Showing the % relative abundance of all the Streptococcus species identified.

Anaerococcus that are non-motile, Gram-positive cocci that are strictly anaerobic and commonly found in the human vagina and various purulent secretions had five species identified in the semen samples. Notably, *Anaerococcus hydrogenalis* (8/22), *Anaerococcus lactolyticus* (7/22), *Anaerococcus Octavius* (13/22),

Anaerococcus prevotii (8/22), *Anaerococcus tetradius* (11/22) and *Anaerococcus vaginalis* (9/22) were present in the studied samples.
Similarly, 9 species were identified from the *Neisseria* genera (**Figure 11**) showing that *Neisseria lactamica* (1.18%) was the most abundant species occurring in 8/22 of the samples. ***Neisseria gonorrhoeae*** was identified in one sample (1/22) with relative abundance of 0.000298402% of the sequence reads.

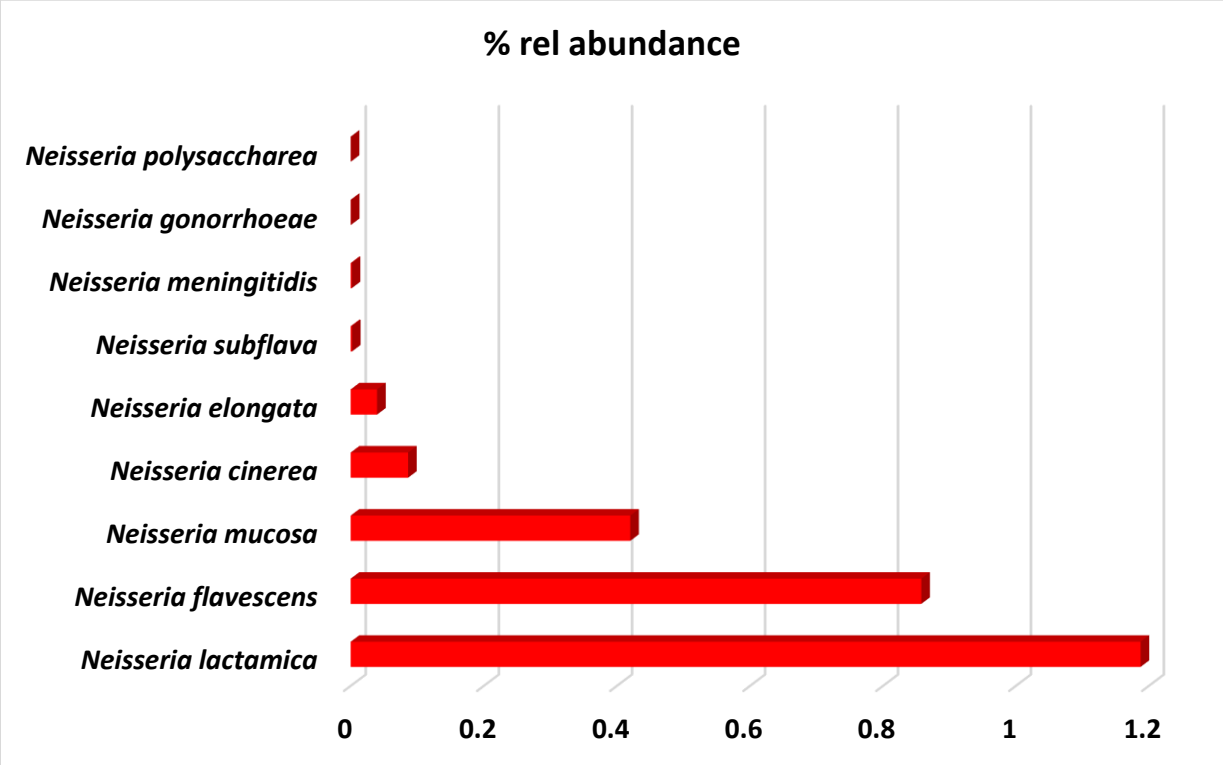

Figure 11: Showing % relative abundance of all the Neisseria species identified in the semen samples.

Among the *Mycoplasma* genus, 9 species were identified from the samples (**Figure 12** *Mycoplasma insons* (0.14%) 8/22, was the most a bundant, followed by *Mycoplasma timone* (0.003%) 5/22, *Mycoplasma iguanae* (0.001%) 4/22 and others.

Ureaplasma genus produced 7 species with *Ureaplasma parvum* (0.18%) as the most abundant species occurring in 9 out of the 22 samples, followed by *Ureaplasma gallorale* (0.007%) 5/22 and others **(Figure 13)**.

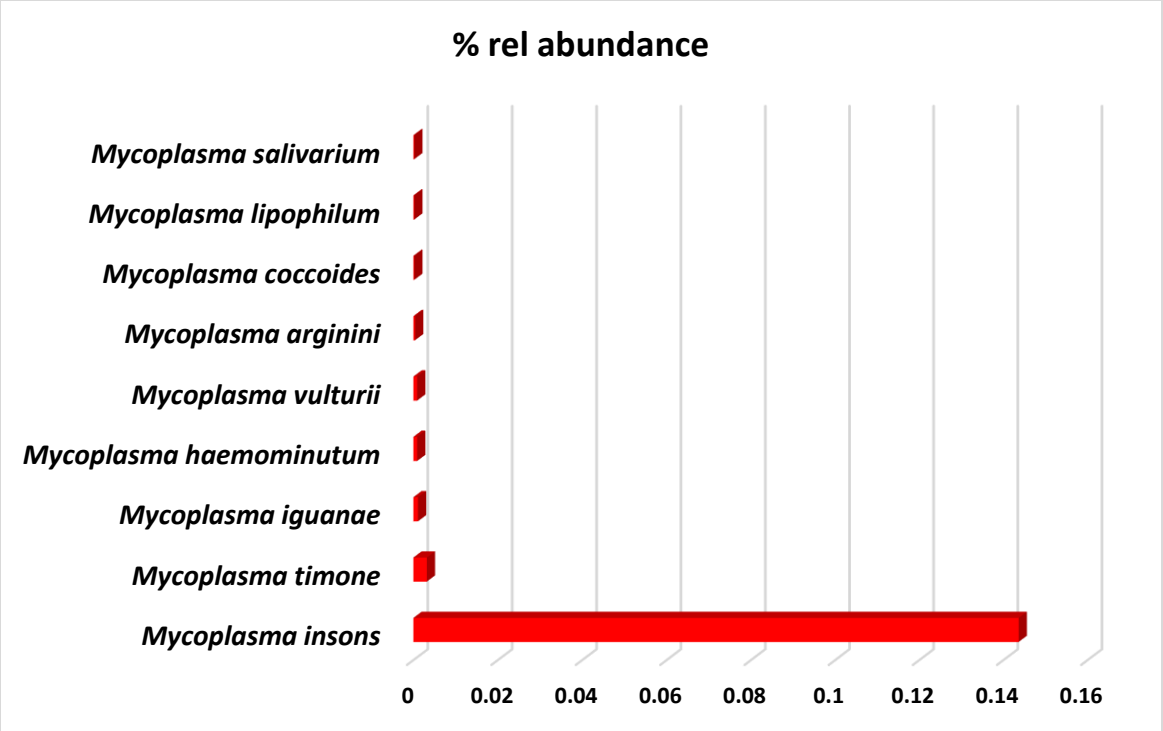

Figure 12: Showing the % relative abundance of all the 9 *Mycoplasma* species

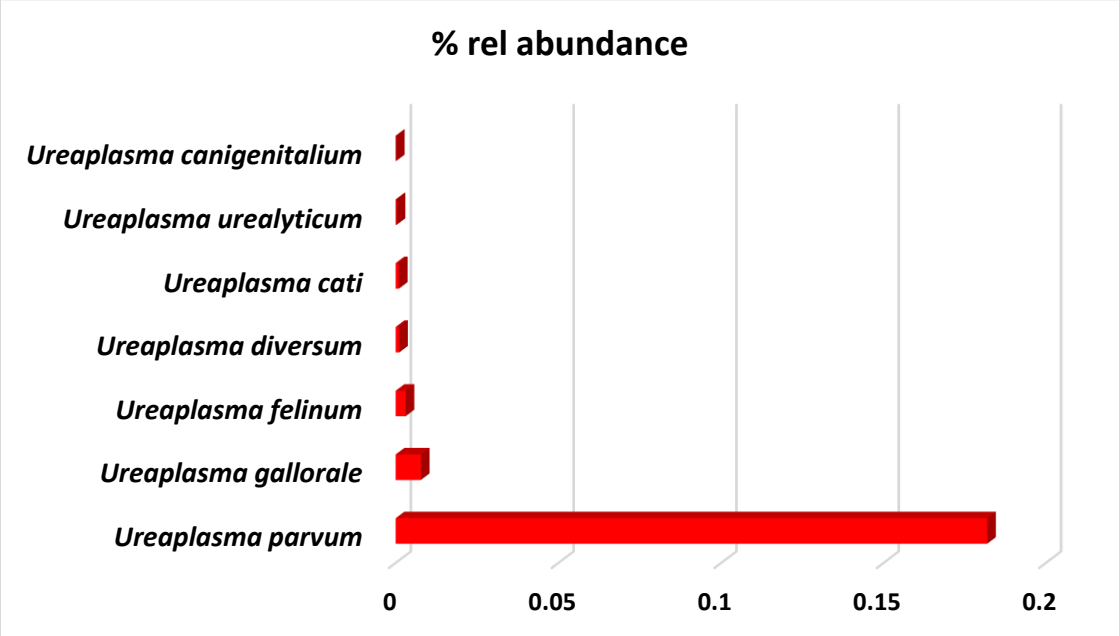

Figure 13: Showing the relative abundance of all the *Ureaplasma* species identified.

Treponema genus had 13 species showing *Treponema amylovorum* (0.032%) 2/22 as the most abundant species, followed by *Treponema medium* (0.018%) 2/22 and others (**Figure 14**).

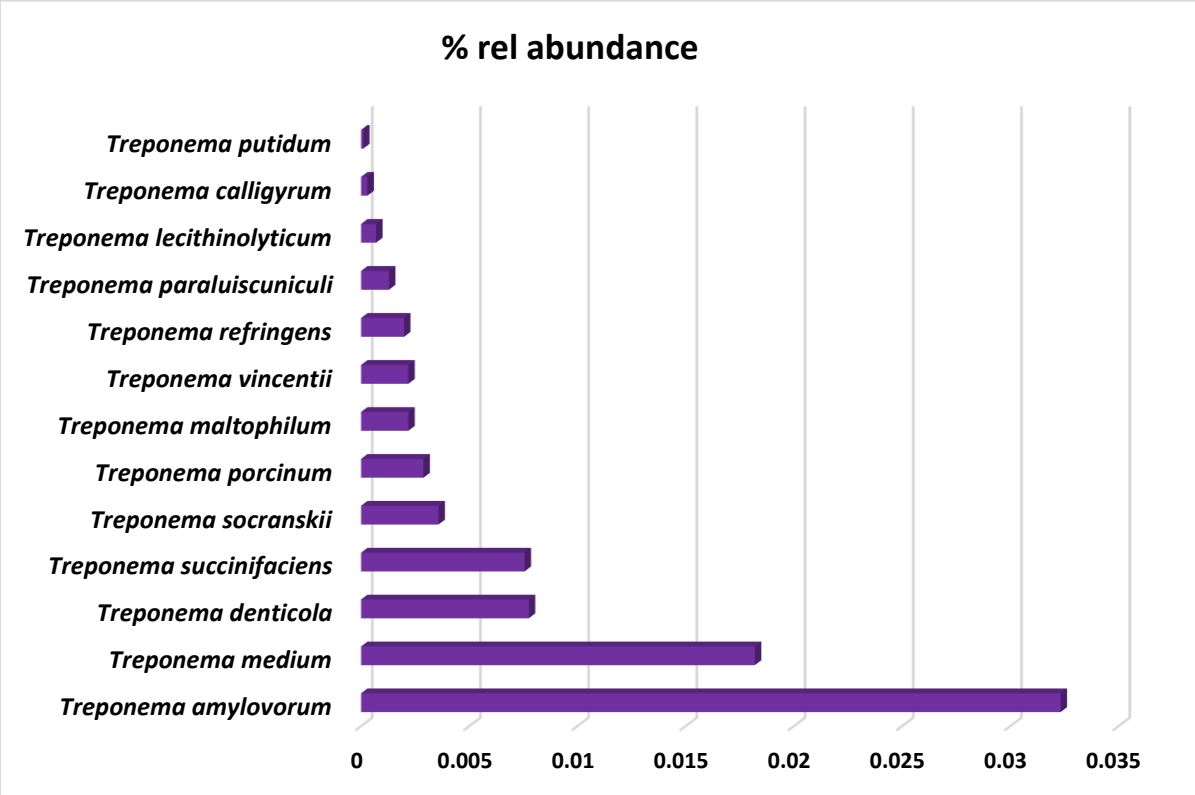

Figure 14: Showing the % relative abundance of all the 13 *Treponema* species found.

Pseudomonas genus constituted 0.22% of the total sequence reads and had 39 species identified with *Pseudomonas stutzeri* found in 15/22 of the samples, followed by *Pseudomonas aeruginosa* (6/22) *Pseudomonas mendocina* (7/22), *Pseudomonas xanthomarina* (9/22), *Pseudomonas oryzihabitans* (6/22) and others (**Figure 15**).

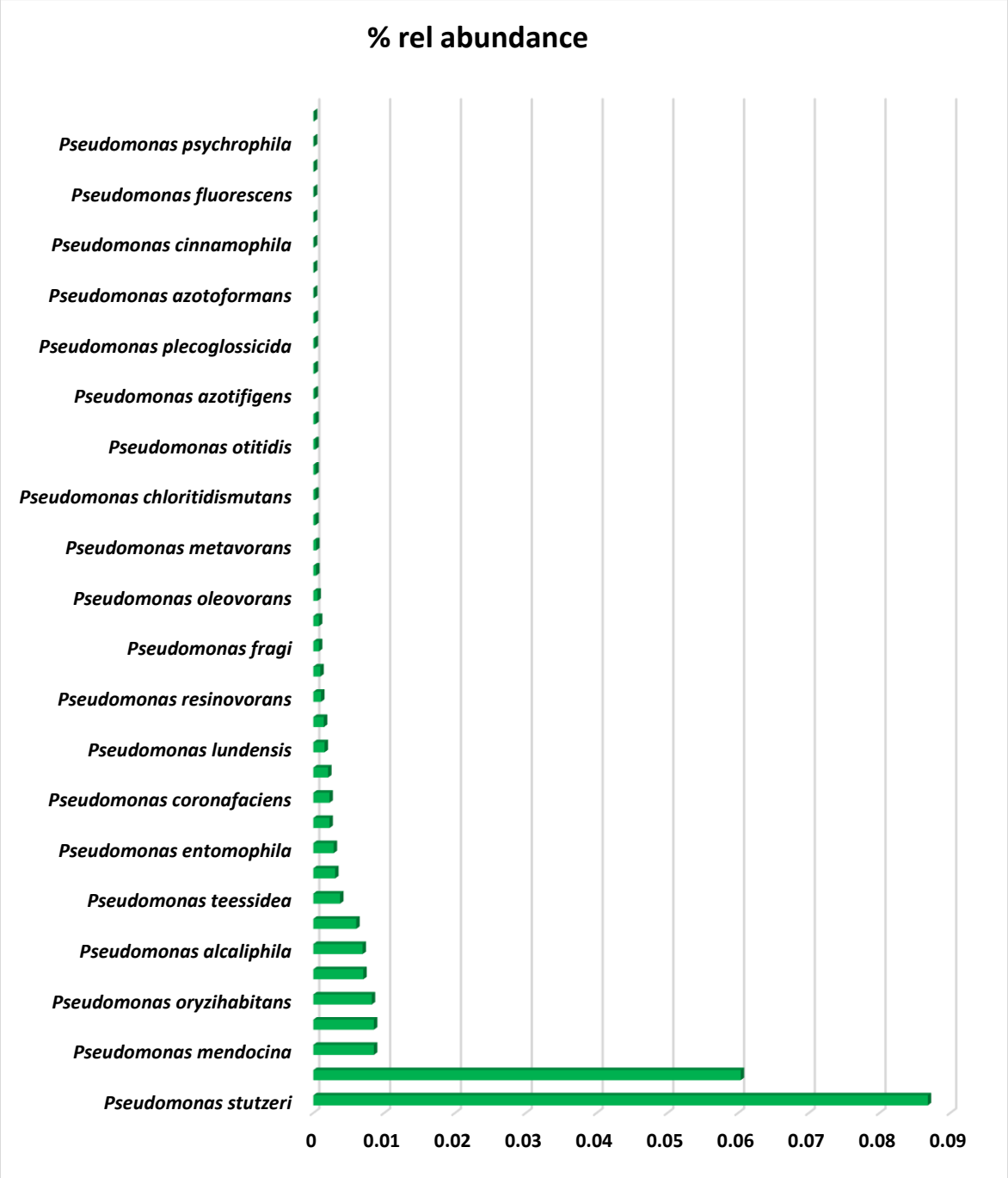

Figure 15: Showing the % relative abundance of all the 39 *Pseudomonas* species identified

It should be noted that other species associated with bacterial vaginosis infection occurred in the samples. For example, *Gardnerella vaginalis* is the only species identified but it occurred in 12/22 of the samples representing 2.39% of the total sequence reads.

In contrast, the results of the culture methods used to process these samples at the teaching hospital reported the isolation of single *Staphylococcus aureus* in three samples, *Haemophilus ducreyi* in one sample, *Enterobacter spp* in one sample and *Streptococcus species* in one sample. All the remaining samples were either reported as no significant growth or no growth after 48 hours of incubation.

DISCUSSIONS

To our knowledge, this is probably the first metagenomics study that determined the microbiota compositions of seminal fluid in a tertiary hospital in Nigeria using the state-of-the-art sequencing technology. The study has demonstrated the polymicrobial nature of bacterial organisms present in the seminal fluids sampled. In contrast to the results obtained from the conventional culture methods, there are large methodological lacuna in the processing of seminal fluids for bacterial isolation at the teaching hospital. Granted that culture methods lacks merits in bacterial isolation and identification, more efforts need to be made in terms of providing new equipment for both aerobic and anaerobic growth conditions and supply of permanent electricity in all teaching and research institutions in the country.

Previous studies relied on culture methods for information about the bacterial communities found in the seminal fluids of men[12, 13]. However, in the last decade, there has been tremendous reliance on molecular techniques to decipher microbial compositions of seminal fluids. The results showed that seminal fluid of these men seeking reproductive health care are highly colonized by diverse bacterial communities including both aerobic and anaerobic fastidious organisms that have never been reported in our environment with the culture-dependent methods. It remains to be determined if these bacterial genera and species play any clinical or physiological role in the reproductive health of men.

This study observed that the bacterial communities at all the taxonomic categories especially the genera and species taxonomic levels are unique to each individual seminal fluid sample and most importantly, the species richness vary widely. No sample had the same bacterial composition thus suggesting that every seminal fluid is personalized in terms of microbiota colonization.

It is noteworthy that most of the bacterial communities found in the seminal fluid of these cohorts of men seeking reproductive care are closely related to bacterial organisms we identified in the female vagina as shown in our previous studies[14, 15]. Other studies have documented similar findings in the vagina[16] and some studies found that the microbiome of seminal fluid are closely related to those found in urine[17], and in the urethra[18].

Previous study by Onemu and Ibeh[5] and other parallel studies[6,7,8] at the same institution, used similar culture methods and reported that *Staphylococcus aureus* constituted 43.7%, followed by *Klebsiella species* (28.2%), and *Escherichia coli* (11.5%) suggests that culture methods needs improvement as none of the studies identified any anaerobe or fastidious growing bacteria in the semen samples. Identification of *Staphylococcus aureus* and *Escherichia coli*

as the only culprits in most infections in Nigerian health teaching institutions leaves much to be desired. In similar studies by culture detection, the microbiota in semen of healthy men was shown to be characterized by Gram-positive bacteria, notably, *lactobacilli*, coagulase-negative *staphylococci*, *streptococci* and *corynebacteri*a[19].

In addition, the identification of 35 *Lactobacillus* species in the semen samples which constituted over 18% of the total sequence reads in this cohorts and found in 20/22 (90.90%) lends credence to the importance of *Lactobacilli* in playing significant role in health maintenance. It is a travesty that the curriculum in medical schools and school of basic medical sciences in Nigerian Universities teach only pathogens to students. In clinical laboratories, scientists and clinicians look for only pathogens to determine genital health. However, urogenital health in adults can also be measured by looking for and measuring beneficial bacteria. We know that the presence of *Lactobacillus* in the vagina can be a good marker of vaginal health. It is less well known that determining the amount of *Lactobacillus* in semen can be a marker of health in males. A recent study showed that *Lactobacillus crispatus* has a positive association with quality of sperm concentration and Kruger's strict morphology[20]. This study identified *Lactobacillus crispatus* in two seminal fluid samples but we may not be able to ascertain if their presence is associated with semen quality. The predominance of *Lactobacillus iners* in the seminal fluids of these subjects raise more questions on the semen quality as *Lactobacillus iners* was found to be at the crossroads of bacterial vaginosis and healthy vagina[21].

Hou *et al*.[22] found that *Lactobacillus* was one of the most predominant bacteria in the semen of people described as normal. Weng et al [20] showed that the most abundant genera among all semen samples tested were *Lactobacillus* (19.9%), *Pseudomonas* (9.85%), *Prevotella* (8.51%) and *Gardnerella* (4.21%). The proportion of *Lactobacillus* and *Gardnerella* was significantly higher in the normal samples, while that of *Prevotella* was significantly higher in the low-quality samples[20]. However, semen quality was not measured in this study, but *Lactobacillus* genera occurred in higher proportion in over 90% of the samples.

In this study, we characterized the microbial communities based on the amplification and sequencing of only the V4 region of 16S rRNA. There may be limitation based on this method as the primers are not universal[23]. Some taxa may be mixed as other similar studies utilised V1-V2 region of the 16S rRNA[22].

A proposal has been suggested by Weng *et al*.[20] towards classification of semen samples into three different microbiome community types by using four genera of bacteria: *Lactobacillus*, *Prevotella*, *Pseudomonas*, and *Haemophilus* to serve as potential markers for future clinical applications and investigations of male infertility. This study may not fit into this classification as no information was collected on semen quality but the classification could serve as a reference point in subsequent large-scale seminal fluid metagenomics investigations.

The identification of BV associated genera in almost all the samples in this cohort of men raises very critical questions as to which bacteria to look out for from semen samples of patients seeking reproductive health care.

Most of the BV associated organisms found in this study are strict anaerobes that requires well equipped laboratory facilities for isolation. For example it is very uncertain if the management of the teaching hospitals across the country will value to need to equip medical microbiology laboratories with the capability to detect *Gardneralla*, *Atopobium*, *Serratia*, *Ureaplasma*, *Veillonella*, *Prevotella*, *Mycoplasma*, *Treponema*, *Haemphilus*, *Sneathia*, *Finegoldia* and other potential pathogenic organisms, in addition to determining the *Lactobacillus* communities affecting semen quality in a positive way. If this happens, there is likelihood that misdiagnosis and inappropriate treatment will reduce tremendously in tertiary health institutions in Nigeria.

CONCLUSIONS

For the first time we have provided an insight into the bacterial microbiome compositions originating from seminal fluids of men seeking reproductive health care and showed that the culture-dependent method should now be regarded as a relic of the past. Every effort should be made to upgrade on the equipment and other facilities required for optimum microbiological investigations in our research and teaching institutions. There is greater needs to upgrade the facilities for microbiological investigations in order to meet the challenges of the 21st century.

ACKNOWLEDGMENT

The study was funded by uBiome Inc., San Francisco, California, USA, and Uzobiogene Genomics London, Ontario, Canada, under an academic consortium grant-in-kind award to Dr. Kingsley Anukam as part of African Microbiome Initiatives (AMI).

Conflicts of interest:
There are no conflicts of interest to declare.

REFERENCES

1.Diemer T, Huwe P, Ludwig M, Hauck EW, Weidner W. Urogenital infection and sperm motility. Andrologia, 2003; 35: 283–287

2. Pellati D, Mylonakis I, Bertoloni G, Fiore C, Andrisani A, et al. Genital Tract Infections and Infertility. European J Obst Gyneacol Reprod Biol 2008; 140: 3-11.

3.Sanocka-Maciejewsky D, Ciupinska M, Kurpisz M. Bacterial Infection and Semen Quality. J Reprod Immunol 2005; 67:51-56.

4. Al-Awadhi H, Dashti N, Khanafer M, Al-Mailem D, Ali N, Radwan S. Bias problems in culture-independent analysis of environmental bacterial communities: a representative study on hydrocarbonoclastic bacteria. Springerplus. 2013; 2: 369. doi: 10.1186/2193-1801-2-369

5.Onemu SO, Ibeh IN. Studies on the significance of positive bacterial semen cultures in male fertility in nigeria. Int J Fertil Womens Med 2001; 46(4): 210-214.

6. Momoh AR, Odike MA, Samuel SO, Momoh AA, Okolo PO. Infertility in male: risk factors, causes and management-A review. Benin J Post Grad Med 2007; 9(1): 22-27.

7. Ibadin O, Kai IN. Bacteriospermia and sperm quality in infertile male patient at University of Benin Teaching Hospital, Benin City, Nigeria. Malaysian J Microbiol 2008; 4(2): 65-67.

8. Ekhaise, FO, Richard FR. Common bacterial isolates associated with semen of men attending the fertility clinic of the University of Teaching Hospital (U.B.T.H),

Benin City, Nigeria. Afr J microbiol 2001; 5(22): 3805-3809.

9. Nwadioha SI, Odimayo MS, Jombo GTA, Microbiologic Review of Seminal Fluids in a Nigerian Tertiary Health Centre. Arch Clin Microbiol. 2016, 7:4

10. Audu BM, Massa AA, Bukar M, Melah GS, Kudi A. Effect of bacterial isolates on the seminal indices of men investigated for infertility in Gombe. Nigerian J Clin Practice 2010; 13: 16-19.

11. Enwaru CA, Iwalokam VN, Ezechi O, Oluwadun A. The effect of presence of facultative bacteria species on semen and sperm quality of men seeking fertility care. African J Urol, 2016; 22(3): 213-222.

12. Busolo F, Zanchetta R, Lanzone E, Cusinato R. Microbial flora in semen of asymptomatic infertile men. Andrologia. 1984; 16:269–275.

13. Rodin DM, Larone D, Goldstein M. Relationship between semen cultures, leukospermia, and semen analysis in men undergoing fertility evaluation. Fertil Steril. 2003; 79:1555–8

14. Anukam KC, Agbakoba NR, Okoli AC, Oguejiofor CB. Vaginal bacteriome of Nigerian women in health and disease: A study with 16S rRNA metagenomics. Trop J Obstet Gynaecol 2019;36:96-104.

15. Okoli AC, Agbakoba NR, Ezeanya CC, Oguejiofor CB, Anukam KC. Comparative Abundance and Functional Biomarkers of the Vaginal and Gut Microbiome of Nigerian Women with Bacterial Vaginosis: A Study with 16S rRNA Metagenomics. J Med Lab Sci, 2019; 29:1-26

16. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, et al. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A. 2010; 108(Suppl 1):4680-4687

17. Nelson DE, Van Der Pol B, Dong Q, Revanna KV, Fan B, Easwaran S, et al. Characteristic male urine microbiomes associate with asymptomatic sexually transmitted infection. Plos One. 2010; 5:e14116

18. Spaine DM, Mamizuka EM, Cedenho AP, Srougi M. Microbiologic aerobic studies on normal male urethra. Urology. 2005; 56:207–10

19 Ivanov IB, Kuzmin MD, Gritsenko VA. Microflora of the seminal fluid of healthy men and men suffering from chronic prostatitis syndrome. Int J Androl 2009; 32: 462–467.

20. Weng SL, Chiu CM, Lin FM, Huang WC, Liang C, Yang T, et al. Bacterial Communities in Semen from Men of Infertile Couples: Metagenomic Sequencing Reveals Relationships of Seminal Microbiota to Semen Quality. PLoS ONE, 2014, 9(10): e110152. https://doi.org/10.1371/journal.pone.0110152

21. Macklaima JM, Gloor GB, Anukam KC, Cribby S, Reid G. At the crossroads of vaginal health and disease, the genome sequence of *Lactobacillus iners* AB-1. PNAS, 2010;1-8 www.pnas.org/cgi/doi/10.1073/pnas.1000086107

22. Hou D, Zhou X, Zhong X, Settles M, Herring J, Wang L, Abdo Z, Forney LJ, Xu C. Microbiota of the seminal fluid from

healthy and infertile men. Fertil Steril. 2013; 100(5): 1261–1269. doi:10.1016/j.fertnstert.2013.07.1991

23. von Wintzingerode F, Gobel UB, Stackebrandt E. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol Rev. 1997; 21:213–29