Location-Specific Gut Bacterial Biomarkers in Delta State, Nigeria: Lefse Identifies Beneficial Commensals Distinguishing Three Communities

Osinachi M. Nwachukwu[1]*; O. S. Egbule[1]; BC Iweriebor[1]; Nqobile Monate Mkolo[2]; Donald Ibe Ofili[2]; Kingsley Chukwuka Amaihunwa[2]

[1*]Department of Medical Laboratory Science,
Faculty of Allied Health Sciences,
Delta State University, Abraka, Delta State, Nigeria
[1]Department of Microbiology, Faculty of Science,
Delta State University, Abraka, Delta State, Nigeria
[1]Department of Biology and Environmental Science,
Sefako Makgatho Health Sciences
University, South Africa.
[2]Department of Biology and Environmental Science,
Sefako Makgatho Health Sciences
University, South Africa
[2]Medical Laboratory Science Council of Nigeria. (MLSCN).
[2]Department of Medical Laboratory Science,
Faculty of Allied Health Sciences,
Delta State University, Abraka, Delta State, Nigeria

https://doi.org/10.56201/rjpst.vol.9.no8.2026.pg37.42

Abstract

Introduction: *The composition of the gut microbiota is shaped by diet, geography, and environment, and distinct populations can harbour distinguishable communities; yet African and Nigerian populations remain under-represented in microbiome research. This study set out to identify the gut bacterial taxa that are differentially abundant among residents of three communities in Delta State, Nigeria — Abraka, Asaba, and Patani.*

Methods: *In a cross-sectional design, faecal samples were collected from 40 participants (22 adults aged 18–35 years and 18 children aged 2–5 years) across the three communities. The full-length 16S rRNA gene was sequenced on the PacBio SMRT platform; community profiling was performed in MicrobiomeAnalyst, and differentially abundant sequences were identified by Linear Discriminant Analysis Effect Size (LEfSe), with significance set at a raw $p < 0.05$ and an LDA score above 2.0, alongside false-discovery-rate (FDR) correction.*

Results: *LEfSe identified five sequences that discriminated among the communities. Three were exclusive to the Abraka residents; Faecalibacterium prausnitzii (LDA = 4.64), Romboutsia spp. (LDA = 4.44), and Bifidobacterium longum (LDA = 4.23) and two were associated with the Patani residents, Bifidobacterium spp. (LDA = 4.09) and Bifidobacterium pseudocatenulatum (LDA = 3.76); none was significantly enriched in the Asaba residents. All five sequences carried strong effect sizes at raw $p < 0.05$, although their FDR-corrected values (≈ 0.089–0.094) fell marginally above 0.05. Overall alpha diversity did not distinguish the communities. Notably, all five discriminating taxa are recognised beneficial commensals.*

Conclusion: *Among these three Delta State communities, the differences in gut microbiota were carried by beneficial commensals, chiefly the butyrate producer F. prausnitzii and several*

bifidobacteria — rather than by potential pathogens, reflecting a shift in the identity of the dominant beneficial genera. These robust LEfSe-selected trends provide the first location-resolved microbial biomarkers for these communities and warrant confirmation in larger cohorts.

Keywords: *gut microbiota; LEfSe; microbial biomarkers; Faecalibacterium prausnitzii; Bifidobacterium; Delta State, Nigeria*

Introduction

The human gut microbiota is a dense and metabolically active community that contributes to nutrition, immune development, and defence against pathogens, and whose composition is shaped by age, diet, geography, and environment (Rinninella *et al*., 2019). Distinct populations can harbour distinguishable gut communities: ethnicity and geography together structure the microbiota, often through the relative abundance of particular taxa rather than through gross diversity (Deschasaux *et al*., 2018; Brooks *et al*., 2018; Gaulke & Sharpton, 2018), a pattern also reported within Nigeria between the nomadic Fulani and their neighbours (Afolayan *et al*., 2019). African, and specifically Nigerian, populations nonetheless remain markedly under-represented in gut-microbiome research, with the great majority of public data drawn from Europe and North America (Abdill *et al*., 2022). The few Nigerian studies available have focused on rural–urban and lifestyle contrasts (Ayeni *et al*., 2018; Doumatey *et al*., 2020), and no study has characterised or compared the gut microbiota of communities in Delta State.
Differential-abundance methods such as Linear Discriminant Analysis Effect Size (LEfSe) can identify the specific taxa that best discriminate between groups and serve as candidate community biomarkers (Segata *et al*., 2011). The present study applied this approach to residents of three Delta State communities — Abraka (Delta Central), Asaba (Delta North), and Patani (Delta South), which differ in geographic setting and habitual diet — in order to identify the gut bacterial taxa that distinguish them.

Methodology

Study design and population

A cross-sectional, comparative study was conducted among 40 apparently healthy participants; 22 adults (18–35 years) and 18 children (2–5 years) resident in three communities of Delta State, Nigeria: Abraka, Asaba, and Patani. Ethical approval was obtained from the Ethics Committee of Delta State University, Abraka, and written informed consent was provided by all participants, with parents consenting on behalf of children. Individuals with persistent or chronic diarrhoea of unknown cause were excluded.

Sample collection, DNA extraction, and sequencing

Stool samples were collected in sterile containers, transferred into nucleic-acid preservation tubes, and stored until processing. Total genomic DNA was extracted, and its concentration and integrity were assessed spectrophotometrically. Sequencing libraries were prepared as SMRTbell templates, and the full-length 16S rRNA gene was sequenced on the PacBio SMRT platform.

Bioinformatics and differential abundance analysis

Raw sequence quality was assessed with FastQC (Andrews, 2015). Sequence variants were inferred and taxonomically classified with DADA2 against the GreenGenes reference database (Callahan *et al*., 2016), and diversity metrics were computed in QIIME (Caporaso *et al*., 2010). Community composition and diversity were profiled in MicrobiomeAnalyst (Dhariwal *et al*., 2017; Chong *et al*., 2020). Differentially abundant sequences among the three communities

were identified using LEfSe (Segata *et al*., 2011), which applies the non-parametric Kruskal–Wallis rank-sum test followed by Linear Discriminant Analysis (LDA) for effect-size estimation; sequences with a raw $p < 0.05$ and an LDA score above 2.0 were regarded as discriminating, with false-discovery-rate (FDR) correction reported alongside.

Result

LEfSe identified five sequences that significantly discriminated among the three communities (raw $p < 0.05$; LDA > 2.0) (Table 1). Three of these were recorded exclusively in the Abraka residents — seq26 (86,537 reads; LDA = 4.64), seq11 (54,972 reads; LDA = 4.44), and seq27 (34,007 reads; LDA = 4.23) — each being wholly absent from the Asaba and Patani samples. The remaining two were associated with the Patani residents: seq40 (24,810 reads in Patani, with a trace of 131 reads in Asaba; LDA = 4.09) and seq99 (11,632 reads; LDA = 3.76). No sequence was significantly enriched in the Asaba residents.

Table 1. *Sequences significantly differentially abundant among the three communities (LEfSe; raw $p < 0.05$, LDA > 2.0), with read counts per community.*

Sequence	p-value	FDR	Asaba	Patani	Abraka	LDA
seq26	0.010	0.089	0	0	86,537	4.64
seq11	0.010	0.089	0	0	54,972	4.44
seq27	0.010	0.089	0	0	34,007	4.23
seq40	0.018	0.094	131	24,810	0	4.09
seq99	0.018	0.094	0	11,632	0	3.76

FDR, false-discovery rate; LDA, linear discriminant analysis. Read counts are the summed reads recorded for each community.

The taxonomic identities of the five discriminating sequences are given in Table 2. The three Abraka-exclusive sequences corresponded to *Faecalibacterium prausnitzii* (seq26), *Romboutsia* spp. (seq11), and *Bifidobacterium longum* (seq27), while the two Patani-associated sequences were *Bifidobacterium* spp. (seq40) and *Bifidobacterium pseudocatenulatum* (seq99).

Table 2. *Taxonomic identity of the five discriminating sequences.*

Sequence	Bacterium	Phylum	Community
seq26	*Faecalibacterium prausnitzii*	Bacillota	Abraka
seq11	*Romboutsia* spp.	Bacillota	Abraka
seq27	*Bifidobacterium longum*	Actinomycetota	Abraka
seq40	*Bifidobacterium* spp.	Actinomycetota	Asaba and Patani
seq99	*Bifidobacterium pseudocatenulatum*	Actinomycetota	Patani

Bacillota = Firmicutes; Actinomycetota = Actinobacteria.

Although all five sequences were significant at the raw $p < 0.05$ level and carried strong LDA effect sizes, their FDR-corrected values (≈ 0.089–0.094) fell marginally above the conventional 0.05 threshold; they are therefore interpreted as robust LEfSe-selected trends warranting confirmation in a larger cohort. Overall alpha diversity (Shannon and Chao1) did not clearly distinguish the three communities, indicating that the between-community signal was carried by specific taxa rather than by community-wide diversity. Notably, all five discriminating taxa are recognised beneficial commensals.

Discussion

The principal finding of this study is that the gut microbiota of these three Delta State communities is distinguished not by pathogens but by a shift in the identity of its dominant beneficial commensals. The Abraka residents were characterised by three exclusive taxa of recognised health value: *Faecalibacterium prausnitzii*, a major butyrate producer with well-documented anti-inflammatory properties (Sokol *et al*., 2008); *Romboutsia* spp.; and *Bifidobacterium longum*. The Patani residents were characterised by bifidobacteria — *Bifidobacterium* spp. and *Bifidobacterium pseudocatenulatum* — a genus widely regarded as beneficial and central to carbohydrate and fibre fermentation in the gut (O'Callaghan & van Sinderen, 2016). That the discriminating taxa are beneficial genera suggests that the communities differ in which favourable organisms predominate rather than in the presence of harmful ones. This dissociation — overall diversity failing to separate the communities while specific taxa did — mirrors the wider literature, in which ethnicity and geography structure the gut microbiota chiefly through particular taxa rather than through gross diversity (Deschasaux *et al*., 2018; Brooks *et al*., 2018; Gaulke & Sharpton, 2018), including within Nigeria (Afolayan *et al*., 2019). The most plausible driver of the observed community differences is habitual diet, which is among the strongest determinants of gut microbial composition (De Filippo *et al*., 2010; Rothschild *et al*., 2018); the three communities occupy distinct geographic settings — upland, central, and riverine — with correspondingly different food environments, although this cannot be established directly from the present cross-sectional data. Several limitations temper these findings. The FDR-corrected significance values fell marginally above 0.05, so the five sequences are best regarded as candidate biomarkers requiring confirmation rather than definitively established associations. The sample was modest (approximately 13 participants per community), the taxa were resolved at the sequence level, and community membership was inevitably confounded with diet and other geographic factors.

Conclusion

Using LEfSe, this study identified five gut bacterial sequences that distinguished residents of three Delta State communities — three (including the butyrate producer *Faecalibacterium prausnitzii*) exclusive to Abraka and two bifidobacteria associated with Patani, with none enriched in Asaba. The community differences were carried by beneficial commensals rather than potential pathogens, reflecting a shift in the identity of the dominant beneficial genera. To our knowledge, these are the first location-resolved gut microbial biomarkers reported for communities in Delta State; they represent robust, biologically plausible trends that should be confirmed in larger cohorts using cold-chain sample preservation and deeper sequencing.

Acknowledgement

The research was funded by TETFUND and SAMRC.

References

Abdill, R. J., Adamowicz, E. M., & Blekhman, R. (2022). Public human microbiome data are dominated by highly developed countries. *PLoS Biology, 20*(2), e3001536. https://doi.org/10.1371/journal.pbio.3001536

Afolayan, A. O., Ayeni, F. A., Moissl-Eichinger, C., Gorkiewicz, G., Halwachs, B., & Högenauer, C. (2019). Impact of a nomadic pastoral lifestyle on the gut microbiome in the Fulani living in Nigeria. *Frontiers in Microbiology, 10*, 2138. https://doi.org/10.3389/fmicb.2019.02138

Andrews, S. (2015). *FastQC: A quality control tool for high throughput sequence data* [Computer software]. Babraham Bioinformatics. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Ayeni, F. A., Biagi, E., Rampelli, S., Fiori, J., Soverini, M., Audu, H. J., … Turroni, S. (2018). Infant and adult gut microbiome and metabolome in rural Bassa and urban settlers from Nigeria. *Cell Reports, 23*(10), 3056–3067. https://doi.org/10.1016/j.celrep.2018.05.018

Brooks, A. W., Priya, S., Blekhman, R., & Bordenstein, S. R. (2018). Gut microbiota diversity across ethnicities in the United States. *PLoS Biology, 16*(12), e2006842. https://doi.org/10.1371/journal.pbio.2006842

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods, 13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods, 7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Chong, J., Liu, P., Zhou, G., & Xia, J. (2020). Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nature Protocols, 15*(3), 799–821. https://doi.org/10.1038/s41596-019-0264-1

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., … Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences, 107*(33), 14691–14696. https://doi.org/10.1073/pnas.1005963107

Deschasaux, M., Bouter, K. E., Prodan, A., Levin, E., Groen, A. K., Herrema, H., … Nieuwdorp, M. (2018). Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nature Medicine, 24*(10), 1526–1531. https://doi.org/10.1038/s41591-018-0160-1

Dhariwal, A., Chong, J., Habib, S., King, I. L., Agellon, L. B., & Xia, J. (2017). MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Research, 45*(W1), W180–W188. https://doi.org/10.1093/nar/gkx295

Doumatey, A. P., Adeyemo, A., Zhou, J., Lei, L., Adebamowo, S. N., Adebamowo, C., & Rotimi, C. N. (2020). Gut microbiome profiles are associated with type 2 diabetes in urban Africans. *Frontiers in Cellular and Infection Microbiology, 10*, 63. https://doi.org/10.3389/fcimb.2020.00063

Gaulke, C. A., & Sharpton, T. J. (2018). The influence of ethnicity and geography on human gut microbiome composition. *Nature Medicine, 24*(10), 1495–1496. https://doi.org/10.1038/s41591-018-0210-8

O'Callaghan, A., & van Sinderen, D. (2016). Bifidobacteria and their role as members of the human gut microbiota. *Frontiers in Microbiology, 7*, 925. https://doi.org/10.3389/fmicb.2016.00925

Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G. A. D., Gasbarrini, A., & Mele, M. C. (2019). What is the healthy gut microbiota composition? A changing

ecosystem across age, environment, diet, and diseases. *Microorganisms, 7*(1), 14. https://doi.org/10.3390/microorganisms7010014

Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., … Segal, E. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature, 555*(7695), 210–215. https://doi.org/10.1038/nature25973

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology, 12*(6), R60. https://doi.org/10.1186/gb-2011-12-6-r60

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L. G., Gratadoux, J. J., … Langella, P. (2008). Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences, 105*(43), 16731–16736. https://doi.org/10.1073/pnas.0804812105