# Genomac Innovative Multi-Omics Research on Infectious Disease in Nigeria (GIMRID)

# Topic: Comprehensive Multi-Omics Research on *Vibrio cholerae* in Nigeria: Integrating Functional Genomics of Infectious Disease, Comparative Genomics, Variant Analysis, and Novel Peptide Therapeutic R&D

## 1.1 Introduction

### 1.1.1 Background

*Vibrio cholerae* is a major cause of severe dehydrating diarrheal disease and it remains a major public health problem in low- and middle-income countries (LMICs) where water, sanitation, and hygiene (WaSH) are inadequate (Sharifi-Mood *et al.,* 2014). Cholera has been historically endemic in the Asian subcontinent, Africa, and the Caribbean. The World Health Organization (WHO) approximates 1.4 million to 4.0 million cases and 21,000 to 143,000 deaths globally due to the disease per annum (World Health Organization, 2017). *V. cholerae* is a Gram-negative, comma-shaped bacterium that is classified serologically into over 200 serogroups. Of them, *V. cholerae* O1 and O139 have caused recent cholera epidemics (Waldor *et al.,* 2014).

Since the first report of a cholera outbreak in Nigeria in 1970 (Babatimehin *et al.,* 2017), the country has remained endemic for the disease with several epidemics. Moreover, cholera in Nigeria has been increasingly linked with environmental and climatic changes, rapid urbanization increasing population growth, and inadequate emergency or public health responses (Legros, 2018). Notably, inadequate access to potable water and poor sanitary conditions remains the principal determinant of cholera transmission in Nigeria, in line with global epidemiology (Impouma *et al.,* 2012).

In 2024, a new strain of *Vibrio cholerae* has been identified in Nigeria. There is concern that the new strain might be more resistant to antibiotics or more virulent, potentially leading to more severe outbreaks. Studying the new cholera strain's spread, resistance, and genetic variations is

essential for adapting treatment strategies, enhancing disease prevention, and improving public health responses.

### 1.1.2 Emergence of New Strains

The evolution of virulence and the spread of drug and antibiotic resistances are now linked to the extensive horizontal gene transfer and recombination events among bacteria and viruses many of which may have occurred in recent years. Genes for antibiotic resistances have been transferred horizontally and have recombined with one another to generate multiple antibiotic resistances throughout the bacterial populations (Davies, 1994).

Natural vectors for horizontal gene transfer comprise replicating, often mobile units of genetic material: plasmids and transposons. Plasmids generally replicate in the cytoplasm, while transposons are integrated into the chromosome, but can move from one site to another in the same or different chromosome, or from chromosome to plasmid and vice versa. Plasmids and transposons typically carry virulence genes and genes for antibiotic resistance. There are three main routes for horizontal gene transfer: infection with viruses (transduction), through pieces of genetic material taken up into cell from the environment (transformation); or by conjugation which is the unusual mating taking place between unrelated species (Mae-Wan Ho *et al.,* 1998).

In the past, only Vibrio cholerae strains of the type O1 are known to have caused epidemics, while non-O1 strains are associated with sporadic cases. However, the recent epidemic in Asia was caused by a non-O1 strain, Vibrio cholerae O139. It turns out that the new strain is identical to an earlier pandemic strain, O1 EL Tor, except for proteins involved in the capsule and the Oantigen synthesis (which is used in typing the strains). This was due to the acquisition of DNA inserted into, and replacing part of the O antigen gene cluster. This suggests that O139 arose by horizontal gene transfer from a non-O1 strain into a type O1 strain (Prager *et al.,* 1995).

In a recent study (Murase *et al.,* 2022) explored the genetic relatedness of all 210 reported serogroups and identified critical distinctions in structural biosynthesis gene clusters on both chromosomes. While only the O1 and O139 serogroups have been known to cause pandemic cholera, members of the remaining serogroups have had significant impacts as environmental strains. While not typically life-threatening, non-O1/non-O139 serogroups can cause sporadic

cases of non-cholera diarrhea, sometimes closely resembling cholera, and some have been shown to act as evolutionary intermediaries in virulence gene acquisition via homologous recombination and horizontal gene transfer (Li *et al.,* 2014, 2019a).

Understanding the genetic characteristics of these strains is essential for effective management and prevention of cholera outbreaks. Continued genetic sequencing of *Vibrio cholerae* would help in understanding the genetic basis of new strains, including mutations that affect virulence or resistance. Research into how genetic changes affect the bacterium's behavior, toxin production, and interaction with the host is crucial for developing effective treatments and antibiotics.

### 1.2 Specific Research Problem

This research aims to address the need to understand the following problem in the new strain of cholera and also come up with possible therapeutic interventions:

Antimicrobial Resistance (AMR): is a major threat to public health globally (Mandal *et al.,* 2011). Antimicrobial resistance (AMR) is influenced by multiple factors, including overuse of antibiotics, insufficient or incomplete dosing, inadequate or insufficient access to clean water, lack of quality medications, high costs of treatment, and low public awareness (WHO, 2022). Cholera outbreaks have been ongoing in Nigeria for the past 4 decades, but the specific strains responsible and their AMR pattern are not well elucidated and their distribution pattern in Nigeria is dynamic. This study aims to proffer some solutions and a better understanding of AMR in cholera by critically studying the mobile gene element and virulence factor which is critical to ensuring that public health responses are effective and that the spread of resistant strains is controlled.

Comparative Genomics: Previous study shows that V. cholerae has evolved to possess complex signal transduction and gene regulatory systems to survive and grow under various environmental conditions (Bhadra et al., 2008). Comparative Genomics study of Cholera aims to have a deep insight into the new strain of cholera by studying its Functional Analysis or Subsystem Categorization, Phylogenomic, Pan genomics, Orthologus Analysis, Average Nucleotide Identity (ANI), Average Amino Acid Identity (AAI), DNA hybridization, Geographical comparative analysis (Geomap, phenotype), Whole Genome Sequence (WGS) Analysis, Multilocus Sequence

Typing (MLST), Core-genomics. These studies are necessary to track the evolution and transmission of V.cholerae and the ability to modify its genetic content (Zeb *et al.*, 2020).

**Variant Analysis**: Recent studies have revealed the genetic mechanisms in these bacteria by which new variants of *V. cholerae* are generated from type-specific strains; these mechanisms suggest that certain strains are selected by environmental or human factors over time. By understanding the mechanisms and driving forces of historical and current changes in the *V. cholerae* population, it would be possible to predict the direction of such changes and the evolution of new variants (Eun *et al.,* 2015). The aim of carrying out variant analysis by investigating variation in Mobile Genetic element, Antimicrobial Resistance, Whole Genome analysis and Virulence Factor is essential for adapting public health measures, ensuring effective treatment, and guiding antibiotic development, all of which are crucial for controlling and eventually eradicating cholera.

**Therapeutic Research and Development**: For centuries, cholera has wreaked havoc on developing countries with poor infrastructure, sanitation, and access to clean drinking water (Rabaan, 2019). Because cholera outbreaks often are linked to poor infrastructure, lack of access to clean water, or societal disruptions, our aim is that a Multi-pronged, Flexible strategy is needed to combat these infections, and each of these treatment strategies can meet a specific need to reduce the burden of cholera. Also, building a model to identify possible therapeutic peptide we be carried out to help combat future outbreaks alongside providing a immediate solution

### 1.3 Objective/Aim of Your Study

The primary purpose of this study is to advance our understanding of *Vibrio cholerae* (the bacterium responsible for cholera outbreaks) by employing an integrative multi-omics approach. This research is crucial for several reasons: firstly, cholera remains a significant public health challenge in Nigeria, where periodic outbreaks have substantial impacts on communities (Adetoro *et al*., 2022), exacerbated by seasonal rains and flooding, which contaminate water sources with fecal matter, spreading rapidly in areas with inadequate sanitation and poor access to clean water (Jaber *et al*., 2023). These outbreaks lead to widespread health crises, overwhelming local healthcare systems and causing high morbidity and mortality rates (Louw *et al*., 2019). This study aims to investigate the outbreak dynamics and spread of a new strain of *Vibrio cholerae*, providing valuable insights into its epidemiology and transmission patterns.

Secondly, considering that AMR is a growing concern in infectious diseases, focusing on AMR mechanisms in *Vibrio cholerae* will help elucidate how this pathogen acquires resistance to commonly used antibiotics, which is essential for developing effective treatment strategies and combating resistance. Lastly, identifying and understanding the virulence factors and mobile genetic elements involved in the pathogenicity of *Vibrio cholerae* is critical for developing new therapeutic approaches and interventions.

The primary objective of this study is to conduct a comprehensive investigation into the new cholera strain by focusing on the following areas:

1. **Outbreak, Spread, and Resistance Mechanisms:** The study aims to explore the outbreak patterns, spread, and resistance mechanisms of the new *Vibrio cholerae* strain. This includes a detailed examination of AMR profiles, virulence factors, and mobile genetic elements that contribute to the strain's pathogenicity and resistance.

2. **Comparative Genomics Study:** The research will involve a thorough comparative genomics study, including:

- Functional Analysis: Investigating the functional roles of genes and their contributions to the strain's ability to cause disease.
- Subsystem Categorization: Classifying genes into functional categories to understand their roles in the bacterium's physiology and pathogenicity.
- Phylogenomic Analysis: Analyzing the phylogenetic relationships between different strains of *Vibrio cholerae* to uncover evolutionary patterns and strain diversity.
- Pan-Genomics: Exploring the pan-genome to identify core and accessory genes across different *Vibrio cholerae* strains.

3. **Variant Study:** Conducting a detailed variant analysis to explore:

- AMR Gene Variations: Variations in AMR genes that contribute to resistance.
- Mobile Genetic Elements: The presence and variation of mobile genetic elements that facilitate gene transfer and resistance.
- Virulence Factors: Variations in virulence factors that enhance pathogenicity.

**4. Therapeutic R&D:** The focus of this will be on the discovery and design of peptides that can effectively inhibit these targets, thereby neutralizing the bacteria's ability to cause disease. This will be done through:

- Identifying Novel Peptide Candidates: Using data from genomics, proteomics, and transcriptomics to discover peptides with potential antimicrobial properties against *V. cholerae*.
- Optimization of Peptide Efficacy: Refining the identified peptides for enhanced stability, specificity, and efficacy in combating *V. cholerae* without harming beneficial microbiota.
- Preclinical Testing: Conducting rigorous in vitro and in vivo studies to assess the safety and therapeutic potential of the developed peptides.
- Addressing Resistance: Investigating the potential for resistance development against the novel peptides and designing strategies to mitigate this risk.

## 1.4. Research Contribution

This research suggests a significant and multidisciplinary research effort focusing on understanding and addressing cholera caused by *Vibrio cholerae* in Nigeria through advanced genomic and therapeutic approaches.

### 1.4.1 Significance of the Research

Toxigenic V. cholerae produces multiple virulence factors that play a key role in its pathogenesis. While the exact mechanisms of pathogenicity are not yet fully understood, it is generally recognized that they involve a combination of these virulence factors and the bacterium's capacity to colonize and persist in the small intestine. (Montero *et al.,* 2023).

This omics approach can help identify how *V. cholerae* strains regulate gene expression in response to environmental conditions, including antibiotic exposure. Understanding the expression patterns of resistance genes and stress response pathways can shed light on how *V. cholerae* adapts at the transcriptional level, contributing to AMR (Domman *et al.,* 2017). By investigating DNA methylation and other epigenetic modifications, insights into non-genetic factors that influence the bacterium's ability to survive antibiotics can be provided, offering another layer of understanding regarding AMR development.

Understanding the evolution of *V. cholerae* using multi-omics data can enhance surveillance programs by providing detailed information on the prevalence and spread of AMR genes in *V. cholerae* populations in Nigeria. This can inform public health strategies to control the spread of resistant strains, aid in the development of more effective therapeutic drugs, potentially targeting specific pathways or proteins involved in resistance.

### 1.4.2 Broader Implication of Research

In recent decades, V. cholera has emerged as a notorious multidrug-resistant enteric pathogen. V. cholera is a severe problem in Asia and Africa. Resistance patterns vary across different geographical regions. Novobiocin (0%) and ofloxacin (0%) are most effective in Africa, gatifloxacin (0%) and levofloxacin (0%) in Asia, and ciprofloxacin (0%) in North America. However, resistance to furazolidone, nalidixic acid, nitrofurantoin, and cephalothin has increased over time. Continuous monitoring of antibiotic resistance and careful selection of appropriate antibiotics are crucial for controlling resistance (Rostami *et al.*, 2022)

The distinctive genetic traits and impressive adaptability of *Vibrio cholerae* are key factors that enable this cholera pathogen to quickly adjust to harsh environmental conditions and withstand the effects of antimicrobial agents. Over the past few decades, *V. cholerae*, which causes acute watery diarrheal disease, has become a notorious multidrug-resistant (MDR) enteric pathogen. While chromosomal mutations contribute to antimicrobial resistance (AMR), the frequent acquisition of extrachromosomal mobile genetic elements (MGEs) from both closely and distantly related bacterial species plays a significant role in *V. cholerae*'s drug resistance (Das *et al.*, 2020).

Our findings can be compared with data from other regions, contributing to a global understanding of cholera and AMR. This can help in developing international strategies for disease management and control. Carrying out this multi-omics research on cholera in Nigeria will not only deepen the understanding of the pathogen but also have far-reaching implications for public health strategies, antimicrobial resistance management, global health security, and policy development.

## 1.5. Methodology

The overview of this research is to use a multi-omics methodological research approach to achieve our desired result. that would revolve around using molecular biology techniques and sequencing techniques such as whole genome sequencing (WGS) which has been championed as the obvious

and inevitable future of diagnostics in multiple reviews and opinion pieces dating back to 2012 (Didelot et al., 2012; Köser *et al.*, 2012; Fricke *et al.,* 2014). WGS is the process of determining the complete nucleotide sequence of an organism's genome. This is generally achieved by 'shotgun' sequencing of short reads that are either assembled de novo or mapped onto a high-quality reference genome (Balloux *et al.*, 2018) This method is essential for identifying single nucleotide polymorphisms (SNPs) and other genomic features crucial for strain characterization (Liu *et al.*, 2019).

### 1.5.1 Sample Collection and Processing

Multiple sites across Nigeria where cholera outbreaks are prevalent will be identified and selected for sample collection. These sites would include both rural and urban areas and regions with varying environmental conditions. Environmental samples (water, soil) and clinical samples (stool, vomit) from cholera patients will be collected for analysis, ensuring that all samples are collected following strict ethical guidelines and with the necessary permissions from relevant health authorities. The samples will be appropriately packaged, labeled, and transported to the laboratory for analysis.

### 1.5.2 Wet Lab

Isolation and Characterization of *Vibrio cholerae* Strains: The samples will be cultured on selective media to isolate *Vibrio cholerae* strains. Biochemical testing, sensitivity test, and serotyping will be done to confirm the presence of *Vibrio cholerae,* the resistant strains, and the identity of the specific serogroups, after which high-quality genomic DNA will be extracted from the confirmed resistant *Vibrio cholerae* isolates using a standardized protocol like phenol-chloroform extraction, ensuring DNA purity and concentration are suitable for downstream applications (Lucena-Aguilar *et al*., 2016).

### 1.5.3 Insilco Multi-omics Approach

The Insilco Multi-Omics Approach refers to the usage of computational methods and tools to combine and examine data from several "omics" layers—genomics, transcriptomics, proteomics, metabolomics, and epigenomics. the Insilco Multi-omics approach will be conducted in this research so that the molecular mechanisms driving the disease can be revealed and identify

potential biomarkers for diagnosis or treatment. Alongside identifying new drug targets while understanding complex biological data.

### 1.5.4 Genomics of Infectious Diseases

- **Data Acquisition and Preprocessing:** Collection of WGS (NGS) Genomic Data: Gather whole genome sequencing data of multi-drug-resistant pathogenic bacteria strains from relevant sources and databases
- **Comprehensive Genome Analysis**: Functional Annotation: Gene prediction, Protein features, Specialty features, Chromosomal properties, and Circus-view, among others
- **Genomic Insights into Drug Resistance and Mechanisms of Adaptation:** Resistome Profile Study: Understanding the mechanisms underlying antibiotic resistance and its relevance to the spread and evolution of multi-drug-resistant pathogenic bacteria.
- **Genomics Insights into the Mechanisms of Bacteria's Pathogenicity:** Virulome Profile Study: Understanding the mechanisms through which the strains evade host defenses and cause infections. Plasmid/Prophages prediction: Mobilome Analysis:
- **Identification of Genetic Markers: Variant Annotation:** Annotate SNPs to identify variants with potential functional significance, such as drug resistance-associated mutations, etc. Pathogenicity Assessment: Predict the impact of variants on protein function and assess their potential contributions to pathogenicity
- **Comparative Genome Analysis (Insights into Spread and Transmission):** Phylogeographic Analysis: • Phylogenetic Tree Construction: Build a phylogenetic tree using SNPs to reveal the evolutionary relationship among the strains. • Phylogeographic
- Geo-mapping: Overlay phylogenetic data onto geographical maps to visualize patterns of spread and transmission routes.
- Virulence Factors: The virulence genes will be identified using databases like VFDB (Virulence Factors of Pathogenic Bacteria), using tools like BLAST to compare these genes across different strains. Detailed variant analysis will be performed on virulence factors using GATK or FreeBayes, focusing on variants that may alter the function or expression of these genes. Using in silico tools (e.g., SIFT, PolyPhen) the impact of variants on virulence factor function will be predicted and the findings correlated with experimental

data on strain virulence. The genomic data will be complemented with transcriptomic analysis to assess the expression levels of virulence genes under different conditions.

**1.5.5. Comparative Genomics:**

- Assembly and Annotation: The raw sequence reads will be assembled into complete genomes using bioinformatics tools, which will be annotated using databases like BV-BRC. The coding sequences, regulatory elements, and other genomic features will be identified.
- Subsystem Categorization: Using tools like SEED or the Subsystems Technology integrated within RAST, the genes will be categorized into subsystems that represent distinct functional units, such as cell wall synthesis, motility, and toxin production. The distribution and variations of subsystems across multiple strains will be analyzed to determine their contributions to the bacterium's physiology and pathogenicity, focusing on identifying unique subsystems in pathogenic versus non-pathogenic strains. The subsystems that are overrepresented in highly virulent strains will be identified and their contribution to the pathogen's ability to invade, survive, and cause disease will be explored.
- Phylogenomic Analysis: A diverse set of *Vibrio cholerae* strains will be selected, including both clinical and environmental isolates from different geographic regions in Nigeria which will integrate functional genomics of Infectious Disease, Comparative Genomics, Variant Analysis, and Novel Peptide Therapeutic R&D
- . Core genome sequences will be aligned and phylogenetic trees will be constructed using tools like MAFFT, MEGA, RAxML, or IQ-TREE before inferring their evolutionary relationships. The genetic distances and divergence times between strains will be calculated to understand the evolutionary history of *Vibrio cholerae*. The clusters of closely related strains will be identified and correlated these with specific epidemiological or geographical data.
- Pan/Core-Genomics: Using tools like BRIG software for pan genomics and genius software to identify unique genes and Roary or Panaroo, core and accessory genomes across different *Vibrio cholerae* strains will be identified. The core genome consists of genes present in all strains, while the accessory genome includes genes present in only some strains. A presence/absence matrix will be

generated to visualize the distribution of genes across the strains, which will be analyzed to identify genes that are unique to pathogenic strains. The functional roles of accessory genes will be investigated, particularly those involved in virulence, antibiotic resistance, and environmental adaptation.

- Average Nucleotide Identity (ANI): High-quality and annotated genomic sequences of different *Vibrio cholerae* strains will be collected and processed for alignment to calculate the ANI using tools like MUMmer or BLAST. Established thresholds will be used to categorize the degree of similarity as ANI values above 95-96% often indicate strains or species within the same species. The ANI results will be correlated with functional data to understand how genetic similarities and differences might impact the pathogen behavior, such as virulence or resistance.
- DNA - DNA Hybridization: Well-labeled and unlabeled DNA samples of different Cholera strains are heated to separate the double-stranded DNA into single strands at 80-95°C for a few minutes. The denatured DNA samples are mixed with a hybridization buffer that facilitates the binding of complementary DNA strands and incubated around 60-65°C to allow complementary strands to form hybrid DNA duplexes for about 16-24 hours and washed thereafter ensuring that only closely related sequences remain hybridized. The percentage of DNA-DNA hybridization is calculated between the reference and the test samples as this value represents the degree of genetic relatedness and the results can also be used to refine or confirm the taxonomic classification of pathogens, including species and strain identification.
- Whole Genome Alignment: Whole genome alignment will be performed by collecting complete genome sequences of the Cholera strains of interest and normalized to account for differences in sequencing depth and other factors. Tools like MAUVE, Mugsy, or progressive Mauve will be used for comparing multiple genomes simultaneously and aligning multiple genomes by identifying conserved genomic regions and rearrangements. The alignments will be refined to address any gaps due to insertions or deletions. Conserved gene orders and arrangements (synteny) across genomes will be identified to understand evolutionary relationships and genome organization. The alignments will be further analyzed to

detect genetic variations such as single nucleotide polymorphisms (SNPs), insertions, and deletions that may be associated with pathogenic traits or drug resistance.

- Multilocus Sequence Typing (MLST): A set of housekeeping genes that are conserved and universally present across the strains of the Cholera species under study will be chosen and amplified using Polymerase Chain Reaction (PCR) ensuring that the primers are designed specifically to the target genes. The PCR products are sequenced either by Sanger sequencing or next-generation sequencing (NGS) methods to obtain the DNA sequences of the housekeeping genes and the obtained sequences are aligned to identify nucleotide variations. The allelic profiles of the different genes are combined to create a unique sequence type (ST) for each strain. Each ST is a specific combination of allele numbers across the selected genes and can be used to assess genetic diversity, track the evolution of strains, and identify patterns of genetic variation.
- The role of horizontal gene transfer in the acquisition of accessory genes will be investigated using tools like ICEberg or PHASTER to identify mobile genetic elements like plasmids, transposons, and integrons.

### 1.5.6. Variant Study

- Variant Calling and prediction: Variant calling is the process of identifying variants from sequence data by comparing the mapped reads (from the previous step) to the reference genome. The data is inputted as a BAM file containing aligned reads and output in VCF (Variant call Format) which lists the variants detected, along with details like quality scores, allele frequencies, and annotations. Identification of genetic variants such as single nucleotide polymorphisms (SNPs), insertions and deletions (indels), and structural variants will be carried out. Tools like GATK, FreeBayes, or SAMtools are used to call SNPs, indels, and other variants.

- Whole Genome Variation: Whole genome variation analysis involves the comprehensive examination of an organism's entire genome to identify genetic variations. These variations can include single nucleotide polymorphisms (SNPs),

insertions, deletions, copy number variations (CNVs), and structural variants (SVs).

- AMR Gene Variations: The AMR genes will be identified using specialized databases such as CARD or ResFinder, and compared across strains to detect variations. Variant calling will be performed, focusing on single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) in AMR genes. The functional impact of detected variants on the AMR genes will be predicted, including any changes in protein structure or function. Genetic variants will be correlated with phenotypic resistance profiles (e.g., MIC testing) to determine their impact on resistance levels.
- Mobile Genetic Elements (MGEs): MGEs will be detected using tools like PlasmidFinder, ISfinder, or mobile genetic element databases; focusing on identifying plasmids, transposons, and integrons. The presence and sequence of MGEs will be compared across different *Vibrio cholerae* strains to understand their role in the spread of resistance and virulence genes.

**1.5.7. Therapeutic R&D**

- Machine Learning: Develop a deep understanding of applying machine learning algorithms to analyze complex biological datasets, particularly within the realm of multi-omics research. This includes harnessing these advanced computational techniques to identify patterns, correlations, and predictive models that can significantly contribute to the rational design and optimization of therapeutic agents, such as anti-cholera peptides. By integrating machine learning with multi-omics data, we aim to advance therapeutic research and development, driving innovation in the identification of novel drug targets and the improvement of treatment strategies.Identifying Novel Peptide Candidates: Data from genomics, proteomics, and transcriptomics will be integrated to identify potential peptide targets, antimicrobial peptides (AMPs) will be predicted based on the protein sequences of *Vibrio cholerae*. Using bioinformatics tools, such as galaxy the peptides will be screened for potential antimicrobial activity against *Vibrio cholerae* considering factors like hydrophobicity, charge, and secondary structure.

- Optimization of Peptide Efficacy: Using molecular modeling software, peptide structures will be refined for enhanced stability, specificity, and binding affinity to target molecules. Using docking simulations, the interaction between peptides and bacterial targets will be predicted and optimized, selecting peptides with high binding affinity and specificity for further testing.
- Peptide Docking: Expertise in In-silico analyses of designed peptides will be developed, focusing on predicting their phytochemical properties, structural characteristics, and functional roles. The 3D structures of target proteins can be obtained from databases like the Protein Data Bank (PDB). If the structure is not available, use homology modeling tools (e.g., SWISS-MODEL, I-TASSER) to predict the structure based on related proteins.

  The 3D structures of the peptides will be gotten and modeled using Auto-dock tools to assign appropriate charges. The docking parameters will then be set suchas the search space around the active site of the target protein, the flexibility of the peptide, and scoring functions to evaluate binding affinity. The docking simulations will then run to predict the binding mode of the peptides to the target protein. This step may involve generating multiple conformations and poses to explore different potential binding interactions.

- Molecular Docking: Docking studies will be carried out to predict the binding affinity of phytochemicals performed to target proteins or enzymes critical for cholera's pathogenesis. This helps in narrowing down potential therapeutic candidates. Skills in molecular docking techniques are also enhanced by utilizing various software tools to identify potential drug targets, select effective ligands, and rigorously evaluate results, all within the context of multi-omics research on cholera.

- Preclinical Testing: The antimicrobial activity of the identified peptides will be evaluated using standard in vitro assays, such as the minimum inhibitory concentration (MIC) assay and bactericidal kinetics assay. The safety of the peptides will be assessed by conducting cytotoxicity assays (e.g., MTT assay) on mammalian cell lines to ensure selective toxicity against *Vibrio cholerae*.

- Addressing Resistance: Using in silico evolutionary models, the potential for resistance development against the novel peptides will be predicted, which can involve simulating peptide exposure over multiple bacterial generations. The potential for combining the peptides with other antimicrobials will be investigated to reduce the likelihood of resistance development. Test the effectiveness of combination therapy in vitro and in vivo.

### 1.5.8. Expected Outcome

At the end of the study, we expect to critically cover the following areas:

1. **Infectious Disease**

The study is anticipated to provide a deeper understanding of the pathogenic mechanisms and epidemiology of *V. cholerae* in Nigeria. By integrating multi-omics data, the research will reveal critical insights into the infection patterns, virulence factors, and environmental triggers that drive cholera outbreaks. This knowledge could enhance public health strategies, leading to more effective prevention and control measures against cholera in Nigeria.

2. **Comparative Genomics**

Through comparative genomics, this research will identify genetic variations and evolutionary relationships between *V. cholerae* strains from Nigeria and those from other regions. This will help to elucidate the unique genomic features associated with the virulence, antibiotic resistance, and adaptability of strains isolated in Nigeria. The findings will contribute to the global understanding of *V. cholerae* diversity and might reveal region-specific targets for intervention.

3. **Variant Study**

The variant analysis is expected to uncover novel genetic mutations and polymorphisms in *V. cholerae* strains that are linked to increased virulence, antibiotic resistance, or environmental resilience. This data will aid in understanding how different strains evolve and adapt to local conditions in Nigeria. The results could also inform the development of diagnostic tools to quickly identify high-risk strains in clinical settings.

4. **Therapeutic R&D**

The research will likely lead to the identification of novel peptide therapeutics that are effective against *V. cholerae*, particularly targeting the strains prevalent in Nigeria. By integrating functional genomics with peptide-based drug discovery, the study could pave the way for new treatment options that are more effective and less prone to resistance. The outcomes might also =include potential candidates for antibiotic development, providing long-term solutions to cholera prevention in endemic regions.

**1.6. Budget**

| Category | Description | Qty | Subtotal excluding VAT (?) | VAT | Subtotal including VAT (?) |
|---|---|---|---|---|---|
| Sample Sequencing | Genomic DNA extraction, QC, SBB Microbial library prep and Sequencing per sample, NGS Data QC | 15 | 4,776,374.70 | 0.23 | 5,134,602.81 |
| Publication Costs | Submission fees, open-access charges, and related expenses | - | 2,000,000.00 | 0 | 2,000,000.00 |
| Internet & IT Infrastructure | High-speed internet, cloud storage, and IT support | - | 2,000,000.00 | 0 | 2,000,000.00 |
| Microbiology Supplies | Necessary microbiology supplies and lab work | - | 2,500,000.00 | 0 | 2,500,000.00 |
| Capacity building and insilico bioinformatics analysis with tools | All are covered by the Genomac institute | 50 | 2,000,000.00 | 0 | 100,000,000.00 |
| Interns budget | Interns to be trained | 10 | 2,400,000 | 0 | 2,400,000.00 |
| Miscellaneous Expenses | Additional unforeseen expenses, travel, equipment, etc. | - | TBD | TBD | TBD |
| | | | | TOTAL | 114,034,602.80 |

inqaba biotec ™

VAT No: 17949735-0001

# Quotation

| Prepared for: | Quotation Number: NG2024/60458 | Delivery Address: |
|---|---|---|
| Genomac Hub<br>Ogbomoso 00<br>Nigeria<br>Phone: 08077191794 | Quotation Date: 30 August 2024<br>Your Reference: SA2024/179772<br>Sales Contact: Olabode Omotoso | Genomac Hub<br>Ogbomoso 00<br>Nigeria<br>Phone: 08077191794 |

| Code | Description | Supplier | In Stock | Qty | Price Per Unit | % Disc | VAT | Subtotal excl. VAT |
|---|---|---|---|---|---|---|---|---|
| IB DNA EX | Genomic DNA extraction | inqaba biotec | N/A | 15 | ₦ 12 523.50 | 0.00 | 7.5 % | ₦ 187 852.50 |
| IB PACB_SBB_MM | QC, SBB Microbial library prep and Sequencing per sample | inqaba biotec | N/A | 15 | ₦ 297 556.88 | 0.00 | 7.5 % | ₦ 4 463 353.20 |
| IB NGSANA | NGS Data QC | inqaba biotec | N/A | 2 | ₦ 62 584.50 | 0.00 | 7.5 % | ₦ 125 169.00 |

| | |
|---|---|
| **Total Excluding VAT** | ₦ 4 776 374.70 |
| VAT 7.5 % | ₦ 358 228.11 |
| **Total Including VAT** | ₦ 5 134 602.81 |

Bank Details;
Account Name: Inqaba Biotec West Africa Limited
Bank Name: Stanbic IBTC Bank
Account number: 0014120597
Kindly send proof of payment by email to olabode.omotoso@inqababiotec.africa and info@inqababiotec.africa

All quotes are valid for 7 days unless otherwise noted.

All purchase order documents submitted to Inqaba Biotec West Africa Ltd must state the relevant quote number.

| Product Category | ETA |
|---|---|
| Instruments | 2-12 Weeks |
| Reagents | 2-6 Weeks |
| Consumables | 2-6 Weeks |
| Siemens | 2-6 Weeks |
| Chemicals | 2-8 Weeks |
| Oligonucleotides/Probes | 2-3 weeks |

ETA's are subject to change based on availability and shipping times

**Figure 1: A total breakdown of the cost for sample sequencing from Inqaba biotech**

**References**


Adetoro, O. I. O., Salami, A., Awotoye, O. O., & Oroboade, J. (2022). Nexus Between Flooding and Cholera in Nigeria: A Review of Occurrence, Consequences, and Mitigation. *Disaster Risk Reduction for Resilience: Disaster Risk Management Strategies*, 361-377.

Babatimehin O, Uyeh J, Onukogu A. Analysis of the re-emergence and occurrence of cholera in Lagos state, Nigeria. In: Chodkowska-Miszczuk J, Szymańska D, editors. Bull Geogr. Toruń: Nicolaus Copernicus University; 2017. p. 21–32.

Balloux, F., Brynildsrud, O. B., Van Dorp, L., Shaw, L. P., Chen, H., Harris, K. A., ... & Eldholm, V. (2018). From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends in microbiology*, *26*(12), 1035-1048.

Das, B., Verma, J., Kumar, P., Ghosh, A., & Ramamurthy, T. (2020). Antibiotic resistance in Vibrio cholerae: Understanding the ecology of resistance genes and mechanisms. *Vaccine*, *38 Suppl 1*, A83–A92. https://doi.org/10.1016/j.vaccine.2019.06.031)

Davies HG, Bowman C, Luby SP. 2017. Cholera—management and prevention. *J Infect* 74:S66–S73.

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E., & Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nature Reviews Genetics*, *13*(9), 601-612.

Domman, D., et al. (2017). *Integrated genomic and functional analyses of Vibrio cholerae during toxigenic outbreaks in Bangladesh.* Nature Microbiology, 2, 1381-1390.

Eun Jin Kim, Chan Hee Lee, G. Balakrish Nair, Dong Wook Kim, trends in microbiology, Volume 23, Issue 8, August 2015, Pages 479-489

Fricke, W. F., & Rasko, D. A. (2014). Bacterial genome sequencing in the clinic: bioinformatic challenges and solutions. *Nature Reviews Genetics*, *15*(1), 49-55.

Goel AK, Jiang SC. 2010. Genetic determinants of virulence, antibiogram and altered biotype among the Vibrio cholerae O1 isolates from different cholera outbreaks in India. *Infect Genet Evol* 10:814–818.

Impouma B, Kasolo F, Yada A, Yoti Z, Yaya S, Woodfill C, et al. Recurring epidemics in the WHO African region: situation analysis, preparedness and response. Afr Health Monit. 2012;(2).

Jaber, T., Fame, T., Agho, O. A., Van de Walle, B., Bartram, J., & Boelee, E. (2023). Environmental, social, and WASH factors affecting the recurrence of cholera outbreaks in displacement camps in Northeast Nigeria: a rapid appraisal. *Journal of Water, Sanitation and Hygiene for Development*, 13(7), 520-526.

Köser, C. U., Ellington, M. J., Cartwright, E. J., Gillespie, S. H., Brown, N. M., Farrington, M., ... & Peacock, S. J. (2012). Routine use of microbial whole genome sequencing in diagnostic and public health microbiology.

Legros D, Partners of the Global Task Force on cholera control. Global cholera epidemiology: opportunities to reduce the burden of cholera by 2030. J infect dis Oxford University Press. 2018;218:S137–40.

Li, F., Du, P., Li, B., Ke, C., Chen, A., Chen, J., et al. (2014). Distribution of virulence-associated genes and genetic relationships in non-O1/O139 *Vibrio cholerae* aquatic isolates from China. *Appl. Environ. Microbiol.* 80, 4987–4992. doi: 10.1128/AEM.01021-14

Li, Z., Lu, X., Wang, D., Liang, W. L., Zhang, J., Li, J., et al. (2019a). Genomic comparison of serogroups O159 and O170 with other *Vibrio cholerae* serogroups. *BMC Genomics* 20:241. doi: 10.1186/s12864-019-5603-7

Liu, B., Zheng, D., Jin, Q., Chen, L., & Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic acids research*, *47*(D1), D687-D692.

Louw, E., Olanrewaju, C. C., Olanrewaju, O. A., & Chitakira, M. (2019). Impacts of flood disasters in Nigeria: A critical evaluation of health implications and management. *Jàmbá: Journal of Disaster Risk Studies*, 11(1), 1-9.

Lucena-Aguilar, G., Sánchez-López, A. M., Barberán-Aceituno, C., Carrillo-Avila, J. A., López-Guerrero, J. A., & Aguilar-Quesada, R. (2016). DNA source selection for downstream applications based on DNA quality indicators analysis. *Biopreservation and biobanking*, *14*(4), 264-270.

Mandal S, Mandal MD, Pal NK. Cholera: a great global concern. Asian Pac J Trop Med. 2011;4(7):573-580. doi:10.1016/S1995-7645(11)60149-1

Montero, D. A., Vidal, R. M., Velasco, J., George, S., Lucero, Y., Gómez, L. A., Carreño, L. J., García-Betancourt, R., & O'Ryan, M. (2023). *Vibrio cholerae*, classification, pathogenesis, immune response, and trends in vaccine development. *Frontiers in medicine*, *10*, 1155751. https://doi.org/10.3389/fmed.2023.1155751)

Murase, K., Arakawa, E., Izumiya, H., Iguchi, A., Takemura, T., Kikuchi, T., et al. (2022). Genomic dissection of the *Vibrio cholerae* O-serogroup global reference strains:

reassessing our view of diversity and plasticity between two chromosomes. *Microb. Genom.* 8:mgen000860. doi: 10.1099/mgen.0.000860

Nelson EJ, Nelson DS, Salam MA, Sack DA. 2011. Antibiotics for both moderate and severe cholera. *N Engl J Med* 364:5–7.

Rabaan AA: Cholera: an overview with reference to the Yemen epidemic. *Front Med.* 2019;13(2):213–228. 10.1007/s11684-018-0631-2

Rostami, A., Zadeh, F. A., Ebrahimzadeh, F., Jafari-Sales, A., & Gholami, S. (2022). Globally Vibrio cholera antibiotics resistance to RNA and DNA effective antibiotics: A systematic review and meta-analysis. *Microbial pathogenesis*, *172*, 105514. https://doi.org/10.1016/j.micpath.2022.105514)

Rupak K. Bhadra *et al.* (2008) Functional analysis of the essential GTP-binding-protein-coding gene cgtA of Vibrio cholera Genetics and Molecular Biology

Samia Zeb, Sardar Muhammad Gulfam, Habib Bokhari , 2020

Sharifi-Mood B, Metanat M. 2014. Diagnosis, clinical management, prevention, and control of cholera; a review study. *Int J Infect* 1:e18303.

Waldor MK, Edward TR. 2014. Vibrio cholerae, p 2471–2479. *In* Bennett JE, Dolin R, Blaser MJ (ed), *Mandell, Douglas, and Bennett's principles and practice of infectious diseases*, 8th ed. Elsevier, Philadelphia, PA.

World Health Organization. 2017. Cholera vaccines: WHO position paper—August 2017. *Wkly Epidemiol Rec* 92:477–500.

World Health Organization. Antimicrobial resistance. October, 2020. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance. Accessed April 2, 2022