Genomics Guided Target Identification and Virtual Screening of Natural Compound Library to Propose Inhibitors Against *Shigella* Spp.

Thesis submitted for the Degree of

Doctor of Philosophy (Science)

In Biotechnology

By

Sarmishta Mukhopadhyay

Postgraduate and Research Department of Biotechnology

St. Xavier's College (Autonomous), Kolkata

Registration No. Ph.D./21/BMBT/01

Affiliated to the University of Calcutta

2024

Contents

Acknowledgements

Foremost, I would like to express my deep and sincere gratitude to my research mentor, Dr. Sayak Ganguli for giving me this opportunity to work under his supervision and for his constant guidance, motivation and patience with me during this research. His expertise, encouragement, unwavering support and constructive feedback was instrumental in shaping the direction and outcome of this research.

I would like to thank my Co-supervisor Dr. Santanu Chakrabarti for his continuous guidance, patience and mentorship. His constructive criticism, invaluable suggestions and innovative ideas have not only honed my skills but also significantly enriched the quality of the thesis.

The guidance from both my mentors have helped me in all the time of research and writing this thesis.

I am immensely grateful to the members of my PhD committee, Dr. Rajat Banerjee, Associate Professor, Dept. of Biotechnology, University of Calcutta; Dr. Pratiti Ghosh, Professor, Dept. of Physiology, West Bengal State University for their insightful comments, suggestions, and contributions. Their expertise and critical evaluation have significantly improved the quality of this thesis.

I extend my sincere thanks to Rev. Dr. Dominic Savio, S.J. (Principal) for providing a conducive academic environment and access to resources necessary for conducting my research. I would like to express my gratitude to Dr. Samrat Roy, Ph. D. Coordinator for seamlessly coordinating all PhD related official activities during my PhD tenure.

I would like to convey my deepest gratitude to Dr. Gautam Sethi, Associate Professor, Dept. of Pharmacology, NUS Yong Loo Lin School of Medicine, for offering me the internship opportunity in his group, and leading me to work on diverse exciting research avenues.

My gratitude also goes to Dr. Arup Kumar Mitra (Controller of Examination) and Dr. Mahashweta Mitra Ghosh for love, encouragement, constructive criticism and motivation throughout this research work.

I owe my deep gratitude to all the dignified faculty members of Department of Biotechnology, St. Xavier's College, Kolkata, viz., Dr. Jhimli Dasgupta (HOD), Dr. Uma Siddhanta, Dr. Sudipa Saha, Dr. Aniruddha Banerji, Dr. Ronita Nag Chaudhuri, Dr. Priyanka De, Dr. Souvik Roy, Dr. Arindam Bakshi and Dr. Ditipriya Hazra for their insightful comments and help throughout my PhD tenure.

I thank my fellow lab mates: Souradip Basu, Gaurab Aditya Dhar, Rupsha Karmakar, Wrick Chakraborty, Debava Chaudhuri for their camaraderie, stimulating discussions, and willingness to collaborate. Their support and friendship have made this journey more enjoyable and rewarding.

I am grateful to West Bengal Department of Science and Technology and Biotechnology for the financial assistance provided in the form of fellowships and grants, which has helped me to complete the work on time.

I also acknowledge the support of the intramural fund received by my principal investigator from St. Xavier's College (Autonomous), Kolkata, which kept us going and made this research into a reality.

My thanks also go to the departmental support staffs, Mr. Abhijit Ghoshal, Mr. Rajkumar Dolui and Mr. Bidesh Makhal for their cooperation and assistance in handling laboratory equipment.

Lastly, I would like to acknowledge my parents, my husband and my friends, whose belief in my abilities and encouragement have provided the emotional sustenance required to complete this challenging task. Their constant support and love have been the driving force behind my accomplishments.

This thesis would not have been possible without the collective contributions and support of all these individuals and institutions. I am truly grateful for their impact on my academic journey and the completion of this research.

Abbreviations

ADMET: Absorption, Distribution, Metabolism, Excretion, And Toxicity

AMR: Antimicrobial Resistance

ARGs: Antibiotic Resistance Genes

BLAST: Basic Local Alignment Search Tool

BPGA: Bacterial Pan Genome Analysis

CADD: Computer Aided Drug Discovery

CARD: Comprehensive Antibiotic Resistance Database

CDC: Center For Disease Control and Prevention

CLSI: Clinical Laboratory Standards Institute

COGs: Clusters Of Orthologous Genes

CPE: Carbapenemase-Producing Enterobacteriaceae

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

DALY: Disability-Adjusted Life Year

DDD: Drug Discovery and Development

DEGs: Differentially Expressed Genes

DNA: Deoxyribonucleic Acid

EIEC: Enteroinvasive *Escherichia Coli*

ESBL: Extended-Spectrum B-Lactamase

GC-HRMS: Gas Chromatography-High Resolution Mass Spectrometry

GCTM: Global Centre for Traditional Medicines

GDP: Gross Domestic Product

GEMS: Global Enteric Multicenter Study

GWAS: Genome-Wide Association Studies

HPS: Hypothetical Proteins

HTS: High Throughput Screening

IL-18: Interleukin- 18

IL-1β: Interleukin-1β

IpaA-D: Invasion Plasmid Antigens A-D

KEGG: Kyoto Encyclopedia Of Genes and Genomes

LBDD: Ligand Based Drug Discovery

LMICs: Low- And Middle-Income Countries

M Cells: Microfold Cells

MDR: Multidrug Resistant

MGEs: Mobile Genetic Elements

MSD: Multistate Disease

MSM: Men Who Have Sex with Men

NCBI: National Center for Biotechnology Information

NF-Kb: Nuclear Factor Kappa B

NGS: Next Generation Sequencing

NICED: National Institute of Cholera and Enteric Disease

OD: Optical Density

OECD: Organisation For Economic Co-Operation and Development

OMPs: Outer-Membrane Proteins

ORS: Oral Rehydration Salts

PI: Propidium Iodide

PK: Pharmacokinetics

PMN: Polymorphonuclear Neutrophil

PMQR: Plasmid Mediated Quinolone Resistance

QMEAN: Qualitative Model Energy Analysis

qPCR: Quantitative Polymerase Chain Reaction

QRDR: Quinolone Resistance-Determining Region

QSAR: Quantitative Structure Activity Relationship

RGI: Resistance Gene Identifier

SBDD: Structure-Based Drug Discovery

SEM: Scanning Electron Microscope

SMILES: Simplified Molecular Input Line Entry System

SVM: Support Vector Machine

T3SS: Type III Secretion System

TCM: Traditional Chinese Medicine

TKM: Traditional Korean Medicine

WGS: Whole-Genome Sequencing

WHO: World Health Organization

YLL: Year Of Life Lost

BGCs: Biosynthetic Gene Clusters

RNA: Ribonucleic acid

MR: Methyl Red

VP: Voges-Proskauer

List of Figures

List of Tables

Abstract

Since the inception of civilization, diarrheal illnesses have detrimentally affected human well-being, and even today in the twenty-first century, they remain the second-most prevalent reason of morbidity and death among youngsters. The most commonly recognized etiologic agent for diarrheal mortality in children is the Gram-negative bacterium *Shigella*, which is accountable for 13.2% of diarrheal episodes across the globe. While *Shigella* was first identified in 1898, the manifestations of shigellosis remain evident today, a century after the disease was discovered. Within the Enterobacteriaceae family, *Shigella* is a facultative anaerobe that is gram-negative, non-flagellated, non-spore-forming, and non-lactose-fermenting. The disease is mostly spread via the faecal channel, which involves food, faeces, fingers, flies, and fomites as the main vectors, routes, and factors that trigger disease outbreak. Because *Shigella* only requires a minimal infectious load (about 10–100 organisms), the infection spreads effortlessly. The primary hallmark of shigellosis is acute ulcerative damage to the large intestine; with signs of fever, vomiting, cramping in the abdomen, diarrhea, and tenesmus. Symptoms emerge between 2 and 5 days after exposure with the bacteria. Sulfonamides had originally made fighting shigellosis easier, but to our disappointment, the efficacy of the remedy did not endure long and more than 80% of the isolates acquired sulfonamide resistance. This was followed by a quick emergence of *Shigella* isolates that were resilient to a broad range of medications, including tetracycline, fluoroquinolones, and nalidixic acid, and hence unresponsive to therapy. Currently, shigellosis is treated with third-generation cephalosporins (ceftriaxone) with varying degrees of success. The pressing need for new therapeutic alternatives to control the threat of shigellosis has been made explicit by the present trend in antibiotic resistance, the rise in non-typeable strains, and the uneven global

distribution of different *Shigella* serotypes. In this regard, this work undertakes a comprehensive genomic analysis of *Shigella flexneri* with the goal to identify promising and broad spectrum novel therapeutic targets in this bacterium and propose potential small molecule inhibitors of herbal origin against them. In the upstream part, the study involved the isolation and identification of a multidrug resistant *Shigella flexneri* strain from urban environmental sample. This was followed by a pangenome analysis to determine the core proteins which were eventually filtered by an efficient subtractive genomics approach to identify a set of *Shigella* specific proteins that do not exist in the corresponding host. After acquiring the likely targets, the next stage was to find inhibitors from natural herbal resources. Since the dawn of time, medicinal plants have been utilized in conventional therapies to heal illnesses in humans and animals. Nevertheless, substantial scientific investigation must be conducted due to the dearth of concrete evidence confirming the curative effect of these drugs against certain infectious diseases. Consequently, the latter part of this work focussed on evaluating the antimicrobial qualities of selected medicinal plants against *Shigella flexneri*. To rapidly and reliably screen natural compounds derived from plants for antibacterial agents, virtual screening, molecular docking, and molecular simulation were executed. 10 unique and widespread pharmacological targets were identified in the investigation, and 3 compounds were deemed as viable therapeutic candidates against the anticipated targets based on their affinity for binding and analysis of the ADMET attributes. The findings of this study corroborate the long-standing assertion of the ethnomedicinal literatures for the perpetual application of these plants for soothing infectious maladies. The prospective targets together with the small molecules proposed through this work, thus opens up the

avenue to replace existing medications with more efficient natural therapeutics upon further in-vitro and in-vivo investigations, to effectively manage *Shigella* superbugs.

Keywords: *Shigella*, MDR, comparative genomics, drug targets, natural products.

1. Introduction

1.1 Infectious Diarrhoea: A Global Health Threat

The epidemiological shift in prevalence from transmissible to non-communicable ailments have adjusted the priority of health care scientists in due course in the past few years. However, the mortality, morbidity, and disability linked to communicable diseases—also referred to as "primitive diseases"—remain noteworthy (Wolde *et al*., 2022). Even in the most recent years, diarrhoea remains one of the top 10 disorders contributing to worldwide DALY and persists as one of the primary factors for hospitalizations, medical consultations, and the worldwide year of lifespan forfeited (YLL) in individuals of every age (GBD 2016 Causes of Death Collaborators (2017)). Diarrhoea is described by the World Health Organization (WHO) as "the passage of three or more loose or liquid stools per day, or more frequently than is normal for the individual". When it comes to etiology, diarrhoea can be classified as non-infectious (such as nutritional diarrhoea, allergy diarrhoea, and symptomatic diarrhoea) or infectious (such as cholera, bacillary dysentery, amoebic dysentery, and rotavirus). However, the main diarrheal burden stems from infections caused by rotavirus, *Cryptosporidium*, enterotoxigenic *Escherichia coli*, and *Shigella* species (Nemeth and Pfleghaar, 2022). Diarrhoea is evaluated differently based on its length, intensity, and the existence of specific coexisting symptoms. The abrupt onset of no fewer than three watery or loose stool episodes on a single day over a period of not more than fourteen days is referred to as acute diarrhoea. When a bout of diarrhoea persists longer than 14 days, it is categorized as chronic or persistent (Nemeth and Pfleghaar, 2022). Often, an infection results in acute diarrhoea. As diarrhoea becomes chronic, non-infectious etiologies become increasingly prevalent. The onset of diarrhoea is the consequence of an aggressive external factor (the virulence or toxicity of a causative

agent, the dose of the inoculant) confronting the body's defences (Bellido-Blasco and Arnedo-Pena, 2011). Secretory, osmotic, and inflammatory diarrhoea are among the pathophysiological mechanisms that might cause diarrhoea as a secondary symptom. The most common causes of diarrhoea in its secretory form are viral and bacterial diseases (Bhutta and Syed, 2016). When the amount of intestinal fluid produced significantly increases beyond the gastrointestinal epithelium's capacity to absorb it, secretory diarrhoea ensues. In this case, damage to the intestinal epithelium is the cause of the watery stool. The digestive tract is lined with epithelial cells, which aid in the absorption of solutes such as water and electrolytes. Increased intestinal permeability results from epithelial cell destruction caused by infectious etiologies. Loose stool results from the injured epithelium cells' inability to absorb water from the intestinal lumen (Nemeth and Pfleghaar, 2022). Diarrhoea can have a wide range of aftereffects, from mild, transient conditions that clear up on their own to more severe cases where electrolyte and water loss may play a significant role. The ensuing hypovolemic shock can be fatal to the most vulnerable patients and induce renal failure and other organ diseases (Bellido-Blasco and Arnedo-Pena, 2011). There are several variables that affect the likelihood of diarrhoea, and each one has a different relative impact depending on how socioeconomic, environmental, and behavioural factors interact (Boithias *et al*., 2016; Graf *et al*., 2008). Young children in low and middle-income countries are disproportionately affected by diarrhoea, which poses a serious risk to the populace and the economy of resource poor nations (Bhutta and Syed, 2016). Poverty is a significant factor determining the prevalence of diarrheal disease in these nations, which appears to be rising over time. In impoverished nations, a child under five experiences diarrhoea three times a year on average (Anand *et al*., 1994). Together, the five nations of China, India, Nigeria, the Congo, Pakistan,

and Nigeria accounted for half of the 4.249 million minor diarrheal deaths in 2008 (Black *et al*., 2010). One-fifth of children worldwide and one-third of children in underdeveloped nations are affected by the high morbidity of up to 43% of stunted growth caused by the ongoing global burden of enteric infections (Bhutta and Syed, 2016). By the time they are 7–9 years old, children who experienced diarrhoea throughout the first two years of their lives may have experienced an average growth deficit of 8 cm and a 10 IQ point decline (Bhutta and Syed, 2016). In 2017, there were over 0.5 million diarrheal fatalities worldwide among children under the age of five, with South Asia and sub-Saharan Africa accounting for 88% of these deaths (Wang *et al*., 2021). Even though public health awareness and sanitation standards have advanced significantly, diarrhoea continues to be one of the leading sources of epidemics globally. Diarrhoea ranked as the eighth most prevalent cause of death worldwide in 2016, with over 4.4 billion cases of illnesses and 1.6 million fatalities linked to it (GBD 2016 Diarrhoeal Disease Collaborators, 2018). An estimated 300,000 children in India die from diarrhoea each year, making it the third most common cause of death for children under the age of five and accounts for 13% of deaths in this age range (Bassani *et al*., 2010). Enhanced comprehension of the aetiology, epidemiology, and weather fluctuations of acute diarrhoea would be beneficial in developing and implementing focused preventative interventions in addition to antimicrobial therapy. To mitigate the commonality of diarrhoea among children, it is imperative to concentrate on a comprehensive strategy for controlling the disease through optimized handling of cases, improving social indicators of wellness like public hygiene and safe water for consumption, promoting consumer consciousness regarding proactive measures like breastfeeding, and pursuing studies on affordable therapies.

1.2 Shigellosis: A Predominant Contributor to The Global Diarrheal Disease

Shigellosis is a clinical illness produced by *Shigella* species invading the epithelium that lines the terminal ileum, colon, and rectal cavity. Although infections occur worldwide and in persons of all ages, recurrent infections among kids under the age of four, living in low- and middle-income countries account for the majority of the illness burden. *Shigella* is relatively resistant to stomach acid, requiring only a few organisms to induce the sickness. Clinical signs normally appear between 12 hours and 3 days after consumption of the organism, with an average incubation period of 3 days. These extremely contagious organisms exhibit a wide variety of symptoms, from acute watery diarrhoea to fulminant dysentery, which is characterized by regular sparse bloody stools, fever, prostration, and stomach cramps. A wide range of unusual but often severe intestinal and extraintestinal problems might arise. Despite significant reductions in mortality during the last three decades, there remain around 164,000 annual deaths owing to shigellosis (Kotloff *et al*., 2018).

1.2.1 How common is *Shigella*-attributable diarrhoea?

Shigella was recognized as the second most prevalent cause of diarrhoeal mortality in 2016 across all ages, accounting for 212,438 casualties and approximately 13.2% of all diarrheal deaths. *Shigella* caused 63,713 deaths among children under 5 years old and was linked to diarrhoea in adults of all ages, with an increase in the elderly and a wide geographical distribution (Khalil *et al*., 2018). Shigellosis is estimated to affect one hundred eighty-eight million people each year, resulting in around one million fatalities. In wealthy countries, there are approximately 1.5 million instances every year. In the United States, it causes approximately 450,000 cases per year (Kotloff *et al*., 2018). *Shigella* is the leading cause of diarrhoea in children under the age of five in sub-Saharan Africa and South Asia (Liu *et al.*, 2016). There is no gender

or racial preference for shigellosis. Shigellosis is widespread in daycare centres and residential facilities. The new use of quantitative PCR for *Shigella* detection indicated a more than five times higher burden of shigellosis among children in low-resource settings than previously detected using culture-based diagnostics (Platts-Mills *et al*., 2018). Importantly, the greatest share of the *Shigella* burden was associated with watery diarrhoea rather than dysentery. A contemporary statistical analysis found that the proportion of *Shigella* infections that manifest as dysentery appears to be declining. Shigellosis overall had a greater link with death than *Shigella*-associated cases of dysentery (Tickell *et al*., 2017). Furthermore, shigellosis has been linked to reduce linear growth even when there are no diarrheal symptoms (Rogawski *et al*., 2018).

1.2.2 Shigellosis: Infection

Shigellosis, a global health concern in both industrialized and developing countries, is a widespread infectious disease caused by different strains and serotypes of *Shigella* species. Shigellosis is mostly transmitted through human faeces, which can be ingested through infected hands or items. Infection can also be transmitted by contaminated food and drink. The disease primarily affects youngsters in low-income areas, with travellers and MSM being the highest risk groups in high-income areas (ECDC, 2022). Outbreaks can spread quickly, often starting with an edible or water-based source and propagating through personal contact. Houseflies play an underestimated role as a mechanical vector for transmission in areas with insufficient human waste management (Farag *et al*., 2013). Shigellosis symptoms typically begin with discomfort, high temperature, migraine, anorexia, and nausea followed by watery diarrhoea. Most mild diseases in individuals who are otherwise in good health, typically resolve within a few days of symptoms. In certain cases, the condition

progresses to frank dysentery, characterized by frequent small stools with blood and mucus, lower abdominal pains, and tenesmus (Kotloff *et al*., 2018). Severe infections might result in over 20 dysenteric stools in a single day. Shigellosis can cause serious intestinal problems such as rectal prolapse, blockage, toxic megacolon, and perforation, albeit these are uncommon (Avital *et al*., 1982). Seizures are the most common extraintestinal manifestation of shigellosis, occurring in 5-30% of hospitalized children, particularly in young children with fever or metabolic derangements that can lead to seizures (Khan *et al*., 1999). Ekiri syndrome is an uncommon but deadly encephalopathy linked to shigellosis (Pourakbari *et al*., 2012). Shigellosis can cause haematological symptoms such as leukaemia and haemolytic uraemia (Khan *et al*., 2013). *Shigella* vaginitis is commonly seen in prepubertal children (Parisot *et al*., 2016). Long-term consequences of shigellosis include persistent diarrhoea and malnutrition, particularly in kids from lower-middle-class and economically deprived nations. Intestinal protein loss is a typical occurrence that can exacerbate these outcomes (Black *et al*., 1982). *Shigella* infection in adults in high-income settings has been related to post-infection irritable bowel syndrome, which is characterized by changed stool frequency, shape, and passage, as well as abdominal pain (Jung *et al*., 2009). Shigellosis may cause irritable bowel syndrome by disrupting tight junctions with commensal bacterial translocation during severe inflammation, leading to dysregulation and long-term immune, proprioceptive, and motility defects (Verdu and Riddle, 2012). Shigellosis may also trigger invasive infections, including meningitis, osteomyelitis, arthritis, and splenic abscess.

1.2.3 Shigellosis: Causes and Transmission

The sole reservoir for *Shigella* is humans; non-human primates can get dysentery if exposed, but they are not involved in the spread of the infection within communities.

Asymptomatic illness over an extended period is uncommon. *Shigella* spp are not naturally durable, thereby making fomites less effective. The faecal-oral route is the predominant mode of transfer, and *Shigella* spp. have a relatively low infectious dose of around 100 organisms (Bennish and Ahmed, 2020). This virus spreads through contaminated water, food, faeces, fingers, flies, and fomites. Poor personal hygiene among food handlers has frequently been linked to the spread of *Shigella*. Furthermore, because *Shigella* spp. are not associated with any particular dietary product, a rise in the prevalence of food-related ailments can be attributed to a range of environmental variables, including poor food handling, storage, and preparation. What exacerbates these scenarios is *Shigella*'s ability to travel easily and quickly from person to person, potentially infecting a larger number of people. *Shigella*-related outbreaks involving food do not follow any specific trend. The disease can afflict an isolated local community or thousands of people across many states or countries. Fresh vegetables and dishes handled by unwell preparers are common causes of contamination. A depressing reminder of the potential for rapid, international transmission of dysentery is an epidemic connected to the eating of meals supplied by a Minnesota-based airline (Hedberg *et al*., 1992). The most probable cause was contamination of meals during preparation by one or more personnel handling food, who had contracted *Shigella sonnei* in the community. Natural disasters, such as earthquakes and flooding, can have a significant health impact on devastated communities (Lampel *et al*., 2018). In these cases, public health is jeopardized if sections of a society's framework, particularly water supply sources along with delivery platforms, as well as sewage disposal facilities, are compromised. Men who have intercourse with men and those with HIV are more likely to contract *Shigella*, leading to severe illness (McCrickard et al., 2018). While this has altered the age-

related demographics of shigellosis in wealthier nations, it has had little impact on the epidemiology of the disease in poorer countries.

1.2.4 Shigellosis: Clinical Features and Diagnosis

The symptoms of shigellosis are generated upon the infiltration of the mucosa of the small bowel and colon by the microorganism, leading to inflammation. Tissue breakdown damages the epithelial cell layer and produces inflammatory exudates in the tissues and stool (Sansonetti *et al*., 1999). Clinical signs might range from watery diarrhoea without blood or mucus to dysentery with tenesmus. In extreme circumstances, gastrointestinal symptoms can last for a long time, resulting in a chronic relapsing sickness with phases of acute active diarrhoea and low-grade symptoms (Ebrahim, 1991). Shigellosis often leads to a bunch of systemic, non-intestinal problems, which are more prevalent and severe among malnourished children in impoverished nations (Khan *et al*., 2013). *Shigella* infections typically cause high (40°C) and long-lasting fevers and seizures in young children impacting about 5% of hospitalized kids (Khan *et al*., 1999). Severe shigellosis can cause lethargy, confusion, obtundation, and even coma. Shigellosis can produce mental abnormalities due to metabolic issues or neurologic consequences, which are not fully understood (Khan *et al*., 1999). Hypoproteinaemia can result from a combination of intestinal protein loss, poor nutrition during sickness, and systemic inflammation that reduces protein synthesis (Bennish *et al*., 1990). *Shigella* can infect epithelial areas beyond the gut, such as the eye and the genital region. It can also invade the circulation and cause sepsis (Operario *et al*., 2017). The most devastating clinical consequence of shigellosis is haemolytic uremic syndrome (HUS). In addition to microangiopathic anaemia and renal failure, leukemoid response (>50,000 mm3

peripheral blood polymorphonuclear cells) and thrombocytopenia may occur in conjunction with HUS (Khan *et al*., 1999).

Even though the final diagnosis of shigellosis requires the isolation of *Shigella* from stool, antimicrobial treatment is mostly based upon speculative clinical diagnosis due to the length of time needed for isolation (48 hours), the lack of diagnostic microbiologic facilities at the majority of health facilities in developing countries, and the insensitivity of current culture methods (Bennish and Ahmed, 2020). Additional symptoms that corroborate the shigellosis diagnosis include rectal prolapse, fever, and abdominal pain. The laboratory test that people in developing nations are most likely to have access to is microscopic inspection of the excrement using a moist mounting medium or Gram stain, both of which exhibits a high polymorphonuclear cell and erythrocyte density. A peripheral blood count with leucocytosis is one of the additional supporting laboratory tests. While some rapid methods for identifying *Shigella* infections have been developed, none are being used widely at the moment (Bennish and Ahmed, 2020). *Shigella* is isolated from the faeces to provide a conclusive diagnosis when microbiologic laboratory facilities are available. *Shigella* do not endure well and therefore necessitates rapid plating of a stool sample either at the patient's bedside or by an immediate transfer to the lab. Stool samples are placed on media that is specifically designed to support gram-negative bacteria. Since *Shigella* cannot ferment lactose, it looks white or colourless on particular medium that have indicator dyes that are sensitive to acid. The identity of non-lactose-fermenting colonies is then confirmed beyond a reasonable doubt using biochemical analyses.

1.2.5 Shigellosis: Medication and Therapy

The preservation of electrolyte balance and hydration is the bedrock for shigellosis treatment. Oral rehydration salts (ORS) or parenteral rehydration therapy are the main methods used in clinical management of shigellosis to restore lost fluids and prevent or treat dehydration (Victora *et al*., 2000). Though useful in minimizing fatalities, ORS fails to eliminate the pathogen and remove signs of infection, which lowers caregiver compliance and treatment effectiveness (Harrell and Cheng, 2018). Antibiotics are indicated to be effective in treating *Shigella* associated dysentery. Antibiotics shorten fever and diarrhoea by one to two days and stop pathogen shedding, which lowers the danger of transmission from person to person (Kotloff *et al*., 2018). For both adults and children with shigellosis, oral azithromycin and ciprofloxacin are typically recommended as initial-line treatments. Individuals with severe diseases or compromised immune functions are prescribed parenteral ceftriaxone. Widespread resistance has rendered previous therapeutic mainstays like ampicillin, trimethoprim-sulfamethoxazole, nalidixic acid, and tetracycline ineffective (Kotloff *et al*., 2018). Because resistant strains pose a concern, treatment decisions should be wise and informed by susceptibility results wherever feasible in order to achieve the greatest benefit and minimize risk. A community's resistance prevalence also plays a crucial role in choosing an empirical treatment plan. Choosing antimicrobials to treat shigellosis is based on a number of principles. Notably, significant faecal concentrations are not necessary for the medication to reach therapeutic quantities in the serum. This has been observed in two apparently non-absorbable antibiotics, viz. rifaximin and gentamicin, which although exhibited in-vitro efficacy against *Shigella*, but failed to produce similar outcomes in vivo (Taylor *et al*., 2008). Secondly, not every antibiotic that surpasses the minimal inhibitory

concentrations in serum works. Amoxicillin, for instance, appears to be less effective than ampicillin despite being better absorbed. In the more than 20 years since azithromycin was originally investigated for the treatment of shigellosis in 1997, no new medications for the disease have been released, and none are expected anytime soon (Khan *et al*., 1997). This puts the possibility of having a widespread, maybe fatal infection for which there is no efficient antibacterial treatment in place in question.

Early feeding is equally crucial. One major complication of shigellosis is the development of malnutrition due to anorexia, increased metabolism, and loss of protein in the gut. In addition to preventing hypoglycaemia, aggressive feeding combined with a high-protein diet can help improve the nutritional decline that occurs after dysentery with *Shigella (*Kabir *et al*., 1992). It is recommended that infants be breastfed for the duration of their sickness. It has also been observed that taking extra zinc shortens the duration and intensity of the acute illness and decreases the frequency and intensity of diarrhoea during the three months following infection (Chao, 2023). Warm magnesium sulphate compresses help relieve oedema and are the best treatment for rectal prolapse. For the foreseeable future, control of shigellosis thus calls for improvements in nutrition and cleanliness, adequate diagnosis and treatment of sick individuals, and betterment in socioeconomic situations.

1.2.6 *Shigella*- The elusive digestive invader

Shigella, the primary bacteria, responsible for persistent endemic diarrhoea has a long and intriguing history. The diagnosis of shigellosis, or bacillary dysentery, dates back to the Bible. The dysentery bacillus is named after the Japanese physician Kiyoshi Shiga, who is widely acknowledged with discovering the organism in Japan in 1898 (Lampel *et al*., 2018). The disease was dubbed "the most dreaded disease of children

from its fulminating course and high mortality" by Shiga himself. In the years since Shiga's discovery, over a hundred distinct of the varieties of the dysentery bacillus were described. Strong evidence suggests that *E. coli* has several origins, from which *Shigella* species and EIEC are derived, forming a single pathovar. It is thought that EIEC strains were the ancestors of "full-blown" *Shigella*, which sprang from *E. coli (*Peng *et al*., 2009). *Shigella* is thought to have evolved from *E. coli* by acquiring the plasmid that caused invasion and by gaining many chromosomal DNA clusters known as pathogenicity islands (The *et al*., 2016). Within the Enterobacteriaceae family, *Shigella* is a genus of bacteria with rod-like forms. Microbiologically, *Shigella* are classified as gram-negative, nonmotile, nonspore forming bacteria. Their cells have dimensions of 0.4 to 0.6 μm in diameter and 1 to 3 μm in length and can be encountered solitary, in pairs, or in strings. *Shigella* species are classified as Voges Proskauer negative, lysine decarboxylase negative, ornithine decarboxylase negative, oxidase negative, urease negative, catalase positive, and methyl red positive (van der Ploeg *et al*., 2010). When cultured in vitro, *Shigella* clusters have a width of around 2 mm and are spherical, convex, smooth, and translucent. All of them are colourless or pale on MacConkey's agar medium, with the exception of *Shigella sonnei*, which is a late lactose fermenter (Payne, 2019). *Shigella* establishes an infection by penetrating cells and producing severe inflammation in the epithelium of the colon and rectal area. A temperature of 37ºC is ideal for their growth. Genes required for cell invasion, such as those pertaining to a type III secretion system (TTSS) that allows penetration into epithelial cells, are carried by a large virulence plasmid found in *Shigella* species (Lampel et al., 2018). With the discovery of *Shigella* DNA in up to one-third of culture-negative specimens, a recent multicenter analysis of the epidemiology and microbiology of shigellosis in Asia suggested that the disease's

incidence may even surpass earlier estimates, thus making it a devastating health problem, even after 100 years of its discovery (Schroeder and Hilbi, 2008).

1.2.7 Classification of *Shigella*

Shigella is an antigenically diversified bacterium that is serologically classified into four recognized species, viz. *Shigella dysenteriae*, *Shigella flexneri*, *Shigella boydii* and *Shigella sonnei* (Kotloff *et al*., 2018). Each of the four forms of *Shigella* share phenotypic similarities, notably invasive pathogenicity, and are composed of many serogroups. Serogroups A through D are represented by each species, which is made up of several (15–20) serotypes. With the exception of *Shigella sonnei*, which only has one serotype, every species has a variety of distinctive serotypes: *Shigella flexneri* has 14, *Shigella boydii* has 19, and *Shigella dysenteriae* has 15 (Kotloff *et al*., 2018). The wide variation in the makeup, alteration, and quantity of repeats of its lipopolysaccharide (LPS) O antigen—a significant *Shigella* antigenic target of the host's humoral response—is reflected in the multitude of serotypes (Schnupf and Sansonetti, 2019). *Shigella* serogroups and types are distributed differently across time and in different geographic regions.

Shigella flexneri: Of the four species, *Shigella flexneri* is the cause of 60% of infections that occur worldwide and has been well investigated, improving our understanding of *Shigella* pathophysiology and the covert host-pathogen interaction (Schnupf and Sansonetti, 2019). Serotyping data provides us with a great deal of epidemiological information on *Shigella flexneri*. Different *Shigella flexneri* serotypes have different O antigens, and experimental data suggests that the conformation of the O antigen is critical for invasion and innate immunity evasion (West *et al*., 2005). Serotype conversion, on the other hand, is a well-documented

phenomenon in *Shigella flexneri*, wherein plasmids and bacterial phages encoding genes, generate alterations in the composition of O antigen and foster serotype transition in a homogeneous population. SfI, SfII, Sf6, SfIV, SfV, and SfX represent a few of the O antigen-modifying bacteriophages found thus far. These bacteriophages change *Shigella flexneri* Y into serotypes 1a, 2a, 3b, 4a, 5a, and X, respectively (The *et al*., 2016). Moreover, in the last 10 years, a number of new *Shigella flexneri* serotypes have been identified, which has complicated the epidemiology and possible protective effectiveness of any future O antigen-based vaccinations. Numerous reports have noted the appearance of new *Shigella flexneri* serotypes. Numerous of these new serotypes carry multiple O antigen-modifying operons, which leads to further alterations in the already significantly altered tetrasaccharide. For instance, *S. flexneri* 1a is transformed into the novel serotype 1c when gtr1C is introduced. This results in the addition of a glucosyl group on the glucosyl-linked N-acetylglucosamine (Stagg *et al*., 2009). There have also been findings regarding serotype flipping accomplished through plasmids in *Shigella flexneri* Xv, 4s, and Yv. It was discovered that the transfer of phosphoethanolamine (PEtN) to the second rhamnose (RhaII) and RhaIII of the O antigen in *S. flexneri* Xv and *Shigella flexneri* Yv, respectively, required the plasmid-borne O antigen phosphoethanolamine transferase (opt) gene (Sun *et al*., 2013; Zhang *et al*., 2014). In Asian nations, 2a, 3a, and 1a are the three most common serotypes of *Shigella flexneri* (Muthuirulandi Sethuvel et al., 2017). Frequent serotype switching as well as the flourishing of certain descendants, emphasize the requirement for more precise tracking and surveillance of *Shigella flexneri*.

Shigella sonnei: As the most prevalent infectious species of shigellosis in affluent nations and the second most common in low- and middle-income countries (LMICs),

Shigella sonnei is an emerging pathogen on a global scale (Shad and Shad, 2021). The O-antigen of *Shigella sonnei* has been observed to suppress uptake and escape from vacuole, and inhibits activation of inflammasome. This unique mechanism helps distinguish *Shigella sonnei* from other Gram-negative enteropathogens and offers it an edge over species that are closely related (Watson *et al*., 2019). *Shigella sonnei* has been divided into three primary lineages, I, II, and III, which have a most recent common ancestor that lived less than 400 years ago (Holt *et al*., 2012). The overwhelming majority of isolates in South America, Asia, and Africa originated from the most newly flourished lineage III. It's significant to note that resistance to several antimicrobials was correlated with worldwide distribution. The acquisition of class II integrons and mutations in DNA gyrase subunit A (gyrA), which encodes the target protein for the fluoroquinolone (a class of broad-spectrum antibiotic) family, served as the mediating mechanisms for resistance (The *et al*., 2016). It appears that the multiple processes and unique mechanisms thus discovered in *Shigella sonnei* collectively contribute to the organism's increased predominance.

Shigella boydii: Since it is seldom isolated elsewhere, *Shigella boydii* appears to be confined to the Indian subcontinent, where it was initially discovered to exist (The *et al*., 2016). *Shigella boydii* was identified as 5.4% (61/1130) of the 1130 *Shigella* isolates that were collected over the 36-month Global Enteric Multicenter Study (Livio *et al*., 2014). Despite its relatively moderate contribution to the total observed cases of multistate disease (MSD) in contrast to the remaining three *Shigella* serovars, *Shigella boydii* remains a key contributor to the entire load of shigellosis. With an isolation rate of almost 50% of all *Shigella boydii* isolates, serotype 2 is the most common and clinically significant serotype in developing nations (Livio *et al*., 2014). The utilisation of genomics to investigate such neglected species like *S. boydii* are

considered to facilitate further advancements in the domains of pathogen identification, phylogenetic classification, and possible functional analysis of the revealed genomic areas specific to clades and species.

Shigella dysenteriae: With a high attack rate, fatality rate, and a number of sequelae, *Shigella dysenteriae* garners the most attention due to its potential to cause an epidemic and its link with the majority of serious cases of dysentery (Dutta *et al*., 2003). The hypervirulence of *Shigella dysenteriae* serotype 1, one of the best studied serotypes, can be somewhat clarified by the presence of the shu cluster, which is upregulated in response to the host body temperature and uses haem as an iron source, improving host adaptation (Kouse *et al*., 2013). It can also be possibly explained by the secretion of Stx, that restricts protein synthesis. During the two world wars, inadequate sanitation and high human mobility allowed two lineages of *Shigella dysenteriae* 1 to spread quickly across continents. It appears that selection during epidemics led to the independent acquisition of antimicrobial-resistance genes by these lineages, which also explains why, despite rare isolation rates, it can nevertheless trigger catastrophic epidemics in populations who are already at risk (Rohmer *et al*., 2014).

1.2.8 Epidemiology of *Shigella*

In both endemic and epidemic forms, shigellosis is a global illness. In most cases, the victims are younger than five years old. An estimated 164.7 million cases of shigellosis occur each year worldwide; children under the age of five account for 69% of all cases and 61% of all deaths linked to the disease (Taneja and Mewara, 2016). Numerous observations have led to the conclusion that the degree of economic development in a given country determines the geographical stratification of *Shigella*

species (Anderson et al., 2016). While *Shigella sonnei* rates rise with economic growth, *Shigella flexneri* ranks the most prevalent species in impoverished nations. According to a 5-year investigation conducted in the southern part of India to ascertain the changing trends of *Shigella* serovars in occasions of dysentery, the most prevalent type of *Shigella* spp. between 2001 and 2003 was *Shigella flexneri* (45%), followed by *Shigella dysenteriae* (29%), *Shigella boydii* (15%), and *Shigella sonnei* (10%). The same study found that there was a significant shift in the *Shigella* serogroups between 2003 and 2006. While *Shigella flexneri* (45%) remained the most common isolate, it was followed by *Shigella sonnei* (31%), *Shigella boydii* (15%), and finally *Shigella dysenteriae* (8%) (Muthuirulandi Sethuvel et al., 2017). It turned out that serotype 2a was the most frequently encountered *Shigella flexneri* strain, followed by 3a, 6 and 2b. However, new data suggests that *Shigella sonnei* has emerged in areas undergoing economic transformation, essentially dispensing of *Shigella flexneri* as the primary cause of shigellosis (Thompson *et al*., 2015). Numerous Asian nations, including Vietnam, Thailand, and Bangladesh, have regularly reported on this species replacement occurrence. Shigellosis was found to be endemic and the primary cause of acute paediatric diarrhoea in hospital-based bacteriological surveillance conducted in the Andaman and Nicobar Islands between 2000 and 2011 (Bhattacharya *et al*., 2014). It was shown that the most widespread and ubiquitous serogroup detected in the islands was *Shigella flexneri* 2a, with *Shigella sonnei* rarely competing, as reported in prior research from India. Apart from the significant species shift that has been noted in developing nations, there has also been an alteration in the epidemiology of *Shigella flexneri* in specific populations in advanced economies. In communities of men who have sex with men (MSM) in Wales, England, and Canada, the isolation rate of *Shigella flexneri* 3a was found to

rise consistently (Kotloff *et al*., 2018). Since the increased isolation rate cannot be attributed to an import from one or more low-income nations, it is possible that this variant's ecology has evolved to allow for transmission within MSM communities. Periodic changes in the most common serogroups of *Shigella* were noted in India, at a tertiary care referral center serving patients from eight neighbouring states (Chandigarh, Punjab, Haryana, Himachal Pradesh, Jammu and Kashmir, western portions of Uttar Pradesh, Uttaranchal, and portions of Rajasthan), covering a wide geographic area (Taneja and Mewara, 2016). From 1994 to 2002, *Shigella flexneri* was the most prevalent serogroup. A decade later, in 2003, *Shigella dysenteriae* type 1 became the leading serogroup, only to be replaced by *Shigella flexneri* again in 2004. The National Institute of Cholera and Enteric Disease (NICED), Kolkata, in eastern India, has also documented similar cyclical variations, where *Shigella dysenteriae* epidemics recur around every ten years (Pazhani *et al*., 2005). Travelers from the developed world to developing nations are susceptible to contracting diarrheal illness from *Shigella*. About 15-20 million tourists who visit developing nations each year are estimated to get diarrhoea. Out of the 64,039 enteric diseases that FoodNet received reports of, 8270 (13%) were linked to travel. Of the bacterial agents, *Shigella* (13%) was the most common, coming in third place after *Campylobacter* (42%), and non-typhoidal *Salmonella* (32%) (Kendall *et al*., 2012). Furthermore, *Shigella* might constitute the primary cause of mild-to-serious intestinal infections provided accurate molecular diagnostic approaches are implemented and adjustments are introduced for those with no symptoms.

1.2.9 Pathogenesis of *Shigella*

There are several steps involved in contracting the invasive pathogen *Shigella* spp (Schroeder and Hilbi, 2008). Following oral consumption, *Shigella* makes it past the

stomach's acidic environment and the gut's competitive microbiota to enter the terminal ileum, colon, and rectum, where it breaks through the mucous layer (Kotloff *et al*., 2018). Proton consumption systems, resistance to locally produced antimicrobial peptides, and mucinase synthesis are some of the mechanisms that allow *Shigella* to get past these obstacles (Anderson *et al*., 2016). The intestinal epithelium, which developed as a structural and functional barrier to guard the body against the invasion of commensal and pathogenic bacteria, is crossed by *Shigella flexneri* in order to reach the intestinal mucosa (Sansonetti, 2004). In the early stages of infection, *Shigella flexneri* appears to induce uptake into microfold cells (M cells) and transcytose across the epithelial layer rather than breaching the epithelial barrier from the apical side (Anderson *et al*., 2016). The low levels of mucus and antimicrobial peptides make M cells easily vulnerable and lucrative targets, which *Shigella* exploits to cross the intestinal barrier and enter polarized colonocytes from the basolateral side (Man *et al*., 2004). Following transcytosis, *Shigella flexneri* escapes into an intraepithelial pocket where it comes in contact with local macrophages that devour and destroy foreign material. *Shigella* prefers to swiftly elude macrophages and infiltrate colonocytes in order to gain access to the cytoplasm of epithelial cells, which serves as their replicative niche (Anderson *et al*., 2016). *Shigella* causes macrophage pyroptosis, which allows the bacteria to leave the cell. The proinflammatory cytokines IL-18 and Interleukin-1β (IL-1β) are released in conjunction with macrophage cell death. The acute and widespread inflammatory response that *Shigella* elicits is significantly mediated by both cytokines. IL-18 has a role in the development of a potent antibacterial response, whereas IL-1β signalling causes the severe intestinal inflammation that is specific to shigellosis (Schnupf and Sansonetti, 2019). Once in the cytosol, *Shigella* penetrates to the neighbouring

epithelial cells, by carefully manipulating actin polymerization, while bypassing external immune system components. However, this technique is a dual-edged weapon because the penetration of epithelial cells triggers an intense inflammatory cascade. The Nod1-mediated intracellular surveillance system detects bacterial peptidoglycan fragments produced by *Shigella* and activates NF-κB, leading to the up-regulation and production of the chemokine IL-8 (Schnupf and Sansonetti, 2019). IL-8 stimulates a huge flux of polymorphonuclear neutrophil leukocytes (PMN) to the region of inflammation. Invading PMN jeopardize the unity of the epithelial barrier, allowing more bacteria from the lumen to break into the submucosa regardless of the assistance of M cells. Furthermore, the pathogen weakens the binding of the epithelial cell junctions that are tightly connected by manipulating the protein content of tight-junction. Thus, macrophage death, epithelial layer degradation, and a huge influx of PMNs all contribute to the bacterial infection and tissue injury (Sakaguchi *et al*., 2002). Both the onset of diarrhoea and the distinctive pathophysiology of shigellosis depend on these processes. The extensive breakdown of tissues inflicted by *Shigella* spp. culminates in reduced adsorption of electrolytes and water, which might reflect in the form of watery diarrhoea and blood-stained stools, that are indicative of infection with *Shigella*. Although severe inflammation favours invasion and infection of *Shigella* at first, new research shows that bacteria release effectors that actively suppress proinflammatory signals, possibly to keep the intensity of the inflammation at a level that is advantageous to the bacteria (Anderson *et al*., 2016). *Shigella* replicates and spreads throughout the intestinal epithelium by using an assortment of protein effectors, like IpaA- IpaD, and IcsA/VirG. The effectors and a needle-like type III secretion mechanism that injects them into the host cell cytoplasm, are all encoded by a sizeable virulence plasmid shared by all *Shigella*

strains (Killackey *et al*., 2016). Even in the absence of traditional adhesion proteins, *Shigella* first clings to the host cell before delivering effectors. Recent research has shown that, following the first activation of the T3SS, the *Shigella* surface protein IcsA undergoes a bile salt dependent activation and promotes contact with host cells by acting as an adhesin (Brotcke Zumsteg et al., 2014). Additionally, bile salts induce the expulsion of OspE1 and OspE2, which persist on the exterior membrane of bacteria and improves adherence to polarized cells. The ultimate assemblage of the T3SS in an instantaneously activatable configuration is facilitated by bile salts, specifically deoxycholate. The tip nexus, which is made up of IpaB, IpaC, and IpaD, is located at the distal extremity of the T3SS. IpaD helps IpaB and IpaC assemble onto the needle while IpaB and IpaC are hydrophobic proteins that intrudes into eukaryotic membranes and create a pore that permits effector delivery (Killackey *et al*., 2016). In the host membrane IpaB communicates with cholesterol, promoting membrane insertion and T3SS activity. IpaB acts as a molecular stopper to prevent secretion through the T3SS prior to implantation, that is eliminated upon implantation into the host membrane. Intriguingly, in a study by Marteyn et al. *Shigella* was observed to inhibit the secretion via T3SS in anaerobic environments by suppressing the transcription of spa32 and spa33 through fumarate and nitrate reductase (Marteyn et al., 2010). When taken in common our findings imply that *Shigella* has developed the ability to detect, with acute precision, when it is in the right gut environment to prompt elevated T3SS activity and adherence.

1.3 The Evolving Epidemic of Antimicrobial Resistance

The healthcare community believed that with the discovery of antibiotics, the struggle against infectious diseases had been won. But now that so many bacteria are resistant to various antimicrobial drugs, it appears like the tide of battle is shifting in the

bacterium's favour (Reygaert, 2018). The 'Faceless Pandemic' that has the entire planet captivated is Antibiotic Resistance. Antimicrobial resistance (AMR) is now recognized as one of the most pressing public health issues of the twenty-first century, which poses a threat to the efficient prevention and treatment of an expanding array of infections brought on by bacteria, parasites, viruses, and fungi that are resistant to the common medications used to combat them (Prestinaci *et al*., 2015). The "golden era" of antibiotics lasted from the 1930s to the 1960s and gave rise to numerous antibiotics (Aslam *et al.*, 2018). However, this period came to an end when researchers were unable to keep up with the fast pace of antibiotic discovery in the face of expanding resistance infections (Aslam et al., 2018). The contributing variables linked to the expansion of antimicrobial resistance involve the continual inability to develop new antibiotics and negligent consumption of prescription drugs. For the first time in the late 1950s and early 1960s, enteric bacteria such as *Salmonella, Shigella,* and *Escherichia coli* showed signs of developing resistance to multiple antimicrobial agents (Aslam *et al*., 2018). These resistant strains primarily affected developing nations, causing enormous clinical, financial, and human casualties. Nonetheless, in the advanced nations it was seen as a minor health issue limited to intestinal microorganisms. This false belief was dispelled in the 1970s with the discovery of multidrug-resistant strains of both *Haemophilus influenzae* and *Neisseria gonorrhoeae*, that showed resistance to ampicillin, tetracycline and chloramphenicol (Aslam et al., 2018). Microbes' extraordinary genetic capacities have benefited from man's overuse of antibiotics, allowing them to exploit every source of resistance genes and every mode of horizontal gene transmission to develop multiple resistance mechanisms for every antibiotic introduced into clinical, agricultural, or other practice (Davies and Davies, 2010). Although uncommon and

difficult to foresee, evolutionary events that result in the creation of novel resistance factors in pathogens may have far-reaching effects (Larsson and Flach, 2022). Conversely, propagation episodes of pre-existing widely distributed resistant variants are frequent, quantitative, and more predictable. To comprehend and mitigate the drug resistance dilemma overall, it is essential to quantify the pathways and pinpoint the factors that contribute to environmental evolution and the bottlenecks that prevent antibiotic resistance from spreading.

1.3.1 The multifaceted phenomenon

The overwhelming threat of antimicrobial resistance is especially significant in the domain of bacteria (Dadgostar, 2019). The impact of AMR on patients and their families is largely invisible, yet it manifests itself in protracted bacterial infections that lengthen hospital stays and result in unnecessary deaths (Laxminarayan *et al*., 2013). Although an inherent feature of microorganisms, the prevalence of clinically relevant antimicrobial resistance (AMR) has been worsened by the improper consumption of antimicrobial drugs, which is fueled by a number of causes.

Abuse of Antibiotics: Ever since the discovery of antibiotics in the 1940s, Sir Alexander Fleming has cautioned the public about the potential for antibiotic misuse due to anticipated high demand (Zaman *et al*., 2017). Numerous global polls reveal that a significant number of patients strongly feel antibacterial drugs could aid in the treatment of viral infections such as the flu or common cold (Dadgostar, 2019). The overprescription of antibiotics, the availability of over-the-counter medications, and the application of antibiotics when they are not genuinely necessary for the therapy, all foster antibiotic resistance.

Utilizing Antibiotics in Agriculture: Another significant variable driving clinical antibiotic resistance is the routine application of antibiotics in agriculture. Worldwide, cattle agriculture used 63,200 tons of antibiotics in 2010- a substantial excess above what was used by humans. Antibiotics are used not just to cure sick animals but also to mostly avoid animal illness, raise herds at subtherapeutic levels, and increase feed efficiency. For example, in the United States alone, around 80% of all antibiotic sales are used on food intended for animal consumption (Bartlett *et al*., 2013).

Economic Growth: Antibiotic usage has been found to be positively connected with rises in GDP and living standards in low- and middle-income countries (LMICs) (Dadgostar, 2019). Klein et al for example, have reported that 65% more antibiotics were used worldwide between 2000 and 2015. While United States, France, Spain, New Zealand, and Hong Kong had the highest rates of antibiotic consumption in 2000, low- and middle-income nations—Turkey, Tunisia, Algeria, and Romania—had the greatest rates of antibiotic consumption in 2015 (Klein *et al*., 2018).

Frequent Travels: According to studies, the ease with which people, animals, and consumables may now travel around the world has greatly aided in the proliferation of antimicrobial resistance. Human visitors possess an elevated risk of returning home infested and diseased due to exposure to resistant diseases. For instance, research conducted by Ruppe et al. found that following return from their trip to India, European visitors who had no interaction with the country's healthcare system, still tested positive for carbapenemase-producing Enterobacteriaceae (CPE) (Ruppé *et al*., 2014).

Environment-driven evolution of resistance: The environment influences the advancement of antimicrobial resistance as well as serves as a conduit for its spread.

Antibiotic resistance can result from foreign DNA absorption as well as alterations in the bacterium's pre-existing genome. Water, soil, and other habitats with highly varying biological niches offer an unparalleled gene pool with a variety that considerably surpasses that of the human and domestic animal microbiota when it comes to the uptake of novel resistance factors (Larsson and Flach, 2022). Antibiotic resistance has been observed in at least some of the pathogens targeted by all recognized classes of antibiotics to date, including natural, semi-synthetic, and synthetic drugs. This implies that, external surroundings currently include resistance elements for any antibiotics that will ever be developed (Larsson and Flach, 2022).

Information Gaps: Since data and specifics regarding the use of various antibiotics in both the healthcare sector and in animal agriculture are not consistently collected globally, there are many gaps in our understanding of antibiotic resistance, which contributes to the ongoing trends of antimicrobial resistance (AMR) (Dadgostar, 2019). For example, only 42 nations worldwide at present maintain statistics on the use of antibiotics in livestock in a structured way. Moreover, there is a dearth of public acknowledgement about the rational utilization of antibiotics and their possible hazards. For example, national survey findings from several nations, both developed and developing, including Japan, the United States, Sri Lanka, and others, show that the majority of individuals have little to no understanding of the proper use of antibiotics (Aslam et al., 2018).

1.3.2 The costs and implications

Antibiotic resistance typically manifesting itself in the form of our inability to effectively treat infections, have implications not only for health care spending but also for a number of other economic areas, with potentially disastrous outcomes (Ait Ouakrim *et al*., 2020). Determining the precise economic setback for treating resistant

bacterial infections remains an immense issue worldwide. The site and severity of infections, a species' innate resistance, and the availability of substitute antimicrobials that may be used to treat infections brought on by resistant strains are just a few of the many variables that affect a pathogenic microorganism's burden of resistance to a specific antibiotic (Ait Ouakrim et al., 2020). The fact that every patient has unique symptoms, complications, and treatment outcomes adds complexity to the effects of AMR (Ait Ouakrim et al., 2020). In a perfect world, assessments for every unique microbe and every distinct antibiotic in various hosts would be included in the burden of resistance estimate, along with consideration for multidrug resistance. Because of the magnitude of the undertaking, estimates of AMR burden are currently based on specific infection locations, bacteria, and antimicrobial combinations. The primary difficulty in estimating the economic impact of antimicrobial resistance (AMR) stems from the fact that some of its costs are concealed, as neither the antibiotic's supplier nor its direct customer must pay for improper use (Ait Ouakrim et al., 2020). A further challenge is that antimicrobial resistance (AMR) undermines the efficacy of several therapeutic measures, such as immunosuppressive treatments, chemotherapy, and surgical procedures, that rely on the efficient management and avoidance of infections. According to estimates, treating a resistant illness can cost anywhere from $10,000 to $40,000 more than treating a susceptible condition (Tansarli *et al*., 2013). The extra health care costs associated with antimicrobial resistance (AMR) were estimated to be over $3.5 billion annually by the Organisation for Economic Co-operation and Development (OECD) based on a recent modelling study involving 33 EU and OECD nations (OECD, 2018). AMR is predicted to cost those nations' healthcare systems a total of $134 billion by 2050 (OECD, 2018). Though significant, AMR's influence on healthcare expenditures pales in comparison to the potential

monetary and human consequences of resistance to civilization. AMR's impact on the social landscape is mostly dictated by the increased rates of illness and mortality that it causes. The magnitude of the workforce and employee productivity are both affected by these two elements. More precisely, AMR is linked to costs to society that occur from missed wages as a result of extended absences from work, medical expenses, and, ultimately, mortality. The influence of AMR on global GDP and on particular global economic components between 2017 and 2050 was evaluated by the World Bank using a general equilibrium model (Adeyi *et al*., 2017). Two different AMR effect scenarios were emulated as disruptions to the supply of human resources. Based on the representations, the global GDP would probably decrease by 1.1% year by 2050 in the optimistic scenario of minimal AMR consequences, as opposed to a base-case scenario in which there would be no AMR effects. Following 2030, the associated GDP gap would be greater than $1 trillion yearly. The world will lose 3.8% of its yearly GDP by 2050 in the high AMR effect scenario, resulting in an annual deficit of $3.4 trillion by 2030 (Adeyi *et al*., 2017). On a broader note, it has been indicated that AMR's effects could mirror those of epidemic outbreaks turning into pandemics.

1.4 Ethnomedicine: A Boon

Natural products with an abundance of structures and bioactivities, has its own way of defending humanity against the horrors of disease and illness. These products are not only essential for maintaining good health but also for fending off various infectious diseases, brought on by viruses, bacteria, fungi, protozoa, and other parasites (Mahapatra *et al*., 2019). For the past century, humanity has attempted to understand nature's riddles, only to be amazed at its revelations till now. "Ethnomedicine" refers to the study of natural resources that have historically been

utilized in many ethnic cultures to treat or cure illnesses. The oldest medical specialty in the world, ethnomedicine is used to treat and prevent both physical and mental diseases (Yuan *et al*., 2016). History has witnessed the development of numerous practical therapeutic techniques by different civilizations to battle a wide range of ailments. Traditional Chinese medicine (TCM), Traditional Korean medicine (TKM), Ayurveda, Unani, each employ elements of nature and have been cultivated for hundreds, if not thousands, of years, blooming into orderly-regulated health care practices. Fossil data suggests, human usage of plants as remedies may be traced back at least 60,000 years (Yuan *et al*., 2016). However, traditional medicine practitioners adhered to their traditions, views, and beliefs while remaining oblivious of current therapeutic theories. As a result, traditional medicines were long overlooked by doctors and biological practitioners for a variety of reasons, including concerns over purity, safety, and potency (Mahapatra *et al*., 2019). The unknown composition and quantum of source components and lack of standardization of the raw materials using contemporary quality control measures, raised concerns about their purity. However, there seems to be a limit to the amount of progress that can be made in the development of novel pharmaceuticals using only contemporary technologies. The limited half-life of most synthetic agents and drugs, combined with the issues of drug resistance, latency, and reactivation in emerging diseases, have prompted scientists to search nature's treasure trove for information on plant metabolites for the treatment of various ailments (Aprilio and Wilar, 2021). Because traditional medicines are more widely available and are perceived as safer, they become more enticing in light of contemporary drugs' limited availability and access. A thorough and continuous investigation into the various facets of ethnomedicine has been made possible by the newly established WHO Global Centre for Traditional Medicines (GCTM) in

Jamnagar, Gujarat, which serves as a knowledge centre for indigenous medicines (Priya, 2022). To maximize the contribution of ethnomedicines to global health and sustainable development, the WHO has included research and education, analytics and data mining, sustainability and equity, as well as innovation and technology, as part of its overall traditional medicines' strategy, while respecting the rights, resources, and cultural legacy of the area as a fundamental value (Priya, 2022). Therefore, addressing the many functions of traditional medicine can help with the development of regionally and nationally suitable health systems. Transdisciplinary research can result in more comprehensive, contemporary, and person-empowering conceptions of health, wellness and disease treatment.

1.5 *In-Silico* Pharmacology for Drug Discovery

Drug discovery and development (DDD), which typically takes an average of 15 years and over US$2 billion to produce a small-molecule medicine, remains sluggish and costly despite remarkable advancements in the fundamental life sciences and biotechnology (Austin and Hayford, 2021). The likelihood of a successful trial for all medications (marketed and under development) was only 13.8%, according to a study that examined 406,038 trials from January 2000 to October 2015 (Wong *et al*., 2019). Optimization of the absorption, distribution, metabolism, excretion, and toxicology (ADMET) and pharmacokinetic (PK) descriptors of larger accumulates of superior chemical entities, hits, and leads during the inception phase of DDD was therefore required to improve preclinical and clinical study outcomes and pave the way for the development of safer, more readily available, and more effective pharmaceuticals. The notion of computer-aided drug discovery, first introduced in the 1970s and made popular by Fortune magazine in 1981, has seen multiple phases of exhilaration and disappointment (Sadybekov and Katritch, 2023). The process of finding and creating

potential medication candidates using computer tools is referred to as *in-silico* pharmaceutical research. By aiding in the screening, design, and therapeutic potential prediction of novel pharmaceuticals, they expedite the process of identifying promising drug candidates (Miah and Mohd Aluwi, 2024). In addition, by assisting with toxicity prediction, *in-silico* techniques help research teams save both money and time by identifying potentially detrimental impacts early in the development process (Miah and Mohd Aluwi, 2024). Computing techniques for example molecular modeling, docking-based computational screening, and digital screening contribute to discovering new and intriguing compounds by predicting potential binding modes of possible therapeutic candidates and assessing their themes of communications. Furthermore, by leveraging existing information to support future protocols, *in-silico* research boosts the flexibility and moral acceptability of early drug development (Miah and Mohd Aluwi, 2024). With some campaigns particularly claiming target-to-lead timelines as low as 1-2 months or target-to-clinic time under 1 year, computationally driven discovery efforts are yielding an increasing number of clinical candidates, albeit it is still too early to expect approved medications from them (Sadybekov and Katritch, 2023). Taking everything into account, the benefits of *in-silico* drug discovery have led to the formation of a $8.3 billion sector annually in 2022, and this momentum is projected to continue (Miah and Mohd Aluwi, 2024). Notwithstanding the difficulties and limitations, it is compelling to promote for the switch from computer-assisted to largely computer-guided DDD ecosystem due to the growing capacity of computational approaches to efficiently access the vast diversity and abundance of lead like chemical moieties at the fundamental transitions from target, through hit to lead till clinical trials.

2. Review of Literature

2.1 The Expanding Spectrum of Antimicrobial Resistance in *Shigella*

Since the early 1960s, *Shigella* has developed and disseminated antibiotic resistance, posing a serious threat to public health worldwide (Muthuirulandi Sethuvel *et al*., 2017). Multiple processes, including drug ejection through operative efflux pumps, a reduction in membrane permeation, amplified levels of enzymes that alter and inactivate medicines, and target alteration through mutagenesis, can lead to drug resistance in *Shigella* species. The most significant contributing factor to the emergence of multidrug-resistant strains of these bacteria has been their capacity to incorporate and spread exogenous genes via mobile genetic elements such plasmids, transposons, insertion sequences, and genomic islands (Muthuirulandi Sethuvel *et al*., 2017). The earliest antibiotics administered for relieving infections with *Shigella* were sulphonamides; tetracycline and chloramphenicol came next. *Shigella a*cquired resistant to each of them, subsequently ampicillin and co-trimoxazole were used as a replacement in treatment (Puzari *et al*., 2018). But as *Shigella* became resistant to the earlier medications, the suggested course of treatment was once more modified to nalidixic acid. Eventually, the ability to withstand nalidixic acid emerged, and fluoroquinolones were subsequently introduced (Puzari *et al*., 2018). Yet, strains resistant to fluoroquinolones have recently been identified from a number of sources. The World Health Organization advises using ceftriaxone, pivmecillinam, and azithromycin to treat *Shigella* infections that are resistant to fluoroquinolones. However, several locations have also reported isolates that are resistant to azithromycin and ceftriaxone (Puzari *et al*., 2018). This has now become a menace to humanity, provoking great fear. *Shigella* species that exhibit resistance to nearly all antibiotic classes are becoming more common and assuming a dominant position

worldwide (Baker and Scott, 2023). The urgency of the situation is underscored by a pattern that other intestinal bacterial pathogens are also following. To address a possible public health emergency, new approaches to the prevention and treatment of this illnesses is necessary.

2.1.1 The Interplay between Genetic Diversity and Antimicrobial Resistance in *Shigella*

The effectiveness of *Shigella* spp. as pathogens and the susceptibility of various human groups to infection with these species may be attributed to their genomic flexibility and capacity to accept genes conferring resistance to antibiotics (The *et al*., 2016). *Shigella* spp. are not an exception to the "arms race" that many bacterial organisms have experienced with antimicrobials, since their widespread usage. Single base pair mutations in important genes that code for antimicrobial targets is inherited by *Shigella* species, rendering the occurrence of multidrug-resistant and extremely drug-resistant forms unavoidable (Baker and Scott, 2023). *Shigella* spp. resistance to trimethoprim was linked to the existence of the dhfrIa or dhfrIIIc genes. While *Shigella sonnei* isolates possessed the dhfrIIIc gene, the majority of *Shigella flexneri* strains were discovered to have the dhfrIa gene, which is thought to be the most prevalent di-hydrofolate reductase gene in the species (Muthuirulandi Sethuvel *et al*., 2017). Sulfonamide tolerance among Gram-negative microbes is likewise typically the result of acquiring any one of the following genes, sulI, sulII, or sulIII. While *Shigella* species generating extended-spectrum β-lactamases (ESBLs) have been getting more commonplace worldwide in recent decades, Gram-negative microorganisms are mostly known for carrying TEM-, SHV-, OXA-, and CTX-M-type β-lactamases. The bla-OXA gene was primarily responsible for the ampicillin resistance of *Shigella flexneri* (Muthuirulandi Sethuvel *et al*., 2017). *Shigella* spp. are

resilient to cephalosporins at a rate of 2·0–5·2%, according to studies from Southeast Asia. Thus far, *Shigella* species have been identified to harbour the genes bla-CTX-M-3, bla-CTX-M-15, bla-CMY-2 and bla-CTX-M-14. Quinolone resistance in *Shigella* is triggered by multiple methods. A key strategy involves mutations in the topoisomerase IV (parC and parE) and DNA gyrase (gyrA and gyrB). An alternative approach relies on the PMQR determinants qnrA, qnrB, qnrS, aac(6')-Ib-cr, and qepA, while *Shigella* spp. are more likely to have qnrS (Muthuirulandi Sethuvel *et al*., 2017). Additionally, *Shigella* species have been shown to exhibit efflux-modulated resistance through mdfA, acrA, and acrB. The overexpression of the acrA gene was found to be the cause of multiple antibiotic resistance (Mar) phenotypes in China based on an analysis of clinical isolates of *Shigella flexneri* obtained from rural hospitals (Puzari *et al*., 2018). *Shigella* species' resistance to several antimicrobial agents, such as streptomycin, chloramphenicol and tetracycline, is attributed to the presence of strA, catA1 and tetB genes. An important contributing element to the spread of AMR is the spread of MDR plasmids among Gram-negative bacteria. *Shigella* spp. frequently harbour numerous AMR genes on plasmids, primarily IncFII, IncI1, IncI2, and IncB/O/K/Z plasmids (Yang *et al*., 2023). In the last three months of 2021, an enormous flare-up of *Shigella sonnei* contamination was reported, affecting multiple European nations. A plasmid called IncFII was recovered from this strain of *Shigella*, which was found to host several resistance genes, including mphA, in addition to the blaCTX-M-27 gene (Yang *et al*., 2023). Analysis reveals that such vectors not just encourage the growth of multidrug resistant *Shigella* strains yet additionally have the ability to transmit between *Shigella flexneri* and *Shigella sonnei*, suggesting an elevated risk of circulation.

2.1.2 Mechanisms of Antimicrobial Resistance in *Shigella*

Shigella's antibiotic resistance strategies can be put into four groups: **acquired resistance**, where the bacteria can obtain new resistance genes and DNA from other resistant bacteria; **intrinsic resistance**, where the bacteria can alter their structures or constituent parts. Next is the **genetic alterations in DNA** that can modify the way proteins are produced, resulting in different components and receptors that the antibiotic cannot recognize. Lastly, DNA can be transferred between bacteria horizontally through transformation, transduction, or conjugation (Breijyeh *et al*., 2020). Naturally acquired resistance to antimicrobial medications can occur through a variety of processes that either reduce affinity with the drug's target or prevent the drug itself from being absorbed. Gram-negative bacteria like *Shigella* spp. may become less susceptible to certain antimicrobial medications if they have a mutation in or lack the ~39 kDa porin in their membrane (Ranjbar and Farahani, 2019). According to a study, permeability of outer-membrane proteins (OMPs) was linked to increased imipenem resistance in mutants of *Shigella dysenteriae*. This was proven by the observation that three imipenem-resistant mutants of the bacteria derived from India had reduced concentrations of significant OMPs (~43 and ~38 kDa). The occurrence of antimicrobial resistant phenotypes in *Shigella* coupled with the removal of harmful substances from their cells can be largely attributed to active efflux pumps. The AcrAB–TolC tripartite complex, a member of the resistance–nodulation–division family of efflux pumps, has been linked to quinolone efflux and is a major contributor to the emergence of resistance in *Shigella* isolates (Yang *et al*., 2008). Certain strains of *Shigella flexneri* have been identified as having AcrAB-associated bile-salt resistance, wherein their manifestation was demonstrated to rise upon encounters with bile salts, thus enabling the bacteria to withstand the killing action of bile. In another

study, Kim and colleagues showed that *Shigella* spp. with elevated expression of the MdfA (a member of the Major Facilitator Superfamily) efflux pump, exhibited resistance to fluoroquinolones (Kim *et al*., 2008). Tet genes, which belong to the MFS antibiotic-efflux system, have been linked to tetracycline efflux and resistance in Gram-negative bacteria like *Shigella* spp, where TetA and tetB were found to mediate tetracycline resistance in *Shigella sonnei* and *Shigella flexneri*, respectively. *Shigella* also expresses a handful of antibiotic-inactivating enzymes as one of its resistance strategies to antimicrobials. *Shigella* isolates have been shown to possess extended-spectrum β-lactamases (ESBLs), that provide resistivity against cephalosporins of third-generation. Despite the fact that the majority of ESBLs are variants within the SHV β-lactamase and TEM families, *Shigella* spp. can also express the CTX-M family, of which CTX-M-15 is one of the most pertinent findings associated with the current epidemiology of ESBLs, essentially recognized among naturally occurring and pathogenic *Shigella* isolates globally (Bialvaei *et al*., 2017). The majority of resistance against fluoroquinolones is brought about by recombination in the chromosomal target site. Quinolone resistance-determining region (QRDR), a short region close to the beginning of the gyrA gene, is where the majority of mutations have been discovered (Ranjbar and Farahani, 2019). The occurrence of plasmid genes designated as plasmid-mediated quinolone-resistance regions (PMQRs), or qnr (qnrA, qnrB, qnrC, qnrD, qnrS, qep, aac[6′]-lb-cr) genes, are typically connected to quinolone resistance in *Shigella* isolates (Ranjbar and Farahani, 2019). The *Shigella flexneri* strains recovered from Chinese patients were the first to have fosfomycin-modifying enzymes identified in them (Liu *et al*., 2017). Several studies have proposed that the rapid diffusion of IncN and IncI plasmids may be accountable for the rising incidence of novel transferable fosfomycin-resistance

determinants, such as fosA3 (Liu *et al*., 2017). Indeed, transconjugant plasmids from several incompatibility groups were found to cocarry the ESBL and fosA3 genes, and expressed resistance to cefotaxime, ceftriaxone, and fosfomycin (Liu *et al*., 2017). Ribosomal modification, active efflux pumps, and enzymatic inactivation were linked to aminoglycoside resistance; while the most prevalent of these mechanisms was aminoglycoside-modifying enzymes (Ranjbar and Farahani, 2019). *Shigella* isolates are frequently found to host aadA gene cassettes, which confer resistance to spectinomycin and streptomycin. Therefore, *Shigella* spp. display multifaceted resistance to antimicrobial agents, which can arise from innate, acquired, or adaptive pathways, making infections resulting from such resistance incredibly challenging to treat.

2.2 Environmental Basins as Reservoirs for Antimicrobial Resistance

The environment has been acknowledged in recent years as an important reservoir and conduit for the propagation of resistance, however our knowledge of how it operates is still fairly limited (Bengtsson-Palme *et al*., 2018). Mitigation of emergence and spread of mobile resistance elements is difficult owing to the lack of comprehension as to how and when the environment fosters resistance development. The term "antibiotic resistome" refers to the group composed of every antibiotic resistance gene (ARGs) including cryptic, acquired, and protoresistance DNA fragments, as well as the products they encode. Comprising intrinsic mechanisms and other resistance factors found in environmental bacteria, the environmental resistome is a subset of the antibiotic resistome (Bobate *et al*., 2023). New elements that contribute to antibiotic resistance may appear at any time or place. As a result, new resistance factors probably emerge on a daily basis, even though the great majority of these occurrences go unnoticed. However, there are a number of explanations for why

new resistance genes do not always end up in generating overwhelming pathogens (Bengtsson-Palme *et al*., 2018). It is imperative to ensure that there be a little or insignificant fitness cost and a substantial selection pressure to preserve the appearance and mobilization of novel resistance elements (Bengtsson-Palme *et al*., 2018). The next question concerns where selection forces are strong enough to encourage the dispersal of chromosomal resistance determinants and the maintenance for the already mobiles ones. A reservoir of resistance genes found in the environment is mostly preserved by natural selection processes (Bengtsson-Palme *et al*., 2018). Thus, environmental hotspots that support a resistance gene's continued survival at whatever cost are crucial for the mobilization of resistance determinants. Water, soil, and other habitats with highly varying biological niches offer an unparalleled gene pool with a variety that considerably surpasses that of the human and domestic animal microbiota when it comes to the uptake of novel resistance factors (Larsson and Flach, 2022). Faecal contamination of the environment creates opportunities for gene exchange between resident ambient bacteria and bacteria that have been evolved to live in the digestive tracts of humans or domestic animals, by enhancing the chances of physical contact between the two types of bacteria (Larsson and Flach, 2022). Furthermore, ARGs found in faecal bacteria that are released into the environment may move horizontally, one or more steps at a time, to pathogens that may eventually infect humans, thus contributing to the evolution of resistance that is clinically important (Larsson and Flach, 2022). Antibiotics infiltrate the environment through anthropogenic sources such as hospital effluents, animal waste, industrial effluents, and aquaculture operations (Bobate *et al*., 2023). The presence of antibiotic residues in the surrounding environment imposes selective pressure on the community of microbes, so that only cells possessing resistance mechanisms—which can be

acquired by lateral gene transfer or spontaneous mutations—survive (Bobate *et al*., 2023). The overabundance of antibiotics in the environment may promote the development of resistance through a number of processes, such as increased mutation frequencies, DNA recombination, and the transmission of MGEs across bacteria (Bobate *et al*., 2023). The World Bank Group prioritized the environmental spread of antimicrobial resistance in 2018. Environmental settings act as "hotspots of resistance," where lateral gene transfer events cause ARGs to intensify and produce new resistant strains. Consequently, in conceptual terms, the most effective mitigation strategies for the emergence of environmental resistance include (i) avoiding the establishment of conditions that favour, mobilize, and permit persistence of resistance genes in bacterial communities; (ii) limiting the pathways via which resistant bacteria can spread throughout the human microbiome; and (iii) lowering the selection pressure for resistant pathogens (i.e., judicious use of antibiotics for humans and animals) (Bengtsson-Palme *et al*., 2018). While the environment plays a role in the development of antibiotic resistance and in its transmission, it can also offer strategies for controlling it (Larsson and Flach, 2022). With the goal to progress drug investigations, environmental microorganisms may be explored as a source for numerous novel candidate antibiotic compounds. Assessing the regional resistance situation might be made feasible by examining the pattern and abundance of resistance in the environmental microbiota (Larsson and Flach, 2022). Further studies may be conducted to determine the degree to which these more common antibiotic pollutants contribute to the spread of antibiotic resistance in the environmental resistome (Flores-Vargas *et al*., 2021).

2.3 Genomics in Drug Discovery

Genomic research has been utilized to speed up the discovery of novel medications for more than 20 years (Spreafico *et al*., 2020). Massive amounts data on gene expression is now capable of being measured and analysed because of advancements in high-throughput sequencing technology and the rise of gene expression profile databases over the past few years (Fang *et al*., 2023). Research on ailments and the development of novel pharmaceuticals depends greatly upon the efficient utilization of genetic data. In-depth analysis of the data in gene expression profile databases using genomics-based tools and technologies can highlight the associations and interplay between medications, illnesses, and genes (Fang *et al*., 2023). Target identification, target prioritizing and tractability, and the forecasting of effects from therapeutic manipulations are aided by a variety of conceptual methods and methodologies in pharmacogenomics (Spreafico *et al*., 2020). The integration of genomic innovations, such as statistical genealogy, genome sequencing, and evaluation of gene expression, to prescription medicines that are being developed and marketed is commonly referred to as pharmacogenomics. This strategy leverages genomics to locate medication response markers at the target, metabolism, and disease pathway levels (Emilien *et al*., 2000). In accordance with the desire to expedite the discovery of innovative therapeutics while also emphasizing the importance for individualized care, genomics technologies are now progressively becoming a crucial part of the drug development process.

2.3.1 System Genetics to Uncover Therapeutic Targets

The discovery novel biological targets, rather than the synthesis of new chemical compounds, is what confines the present paradigm of target-based drug design

(Penrod *et al*., 2011). By shedding light on the underlying genetics of disease, the integration and analysis of genomic data may help identify new targets for pharmaceutical development. Given the enormous amount of data that is routinely collected, the usefulness of such genome-wide metrics mostly hinges on the computational interventions performed to convert the data into an interpretable format (Penrod *et al*., 2011). Computational approaches that address the impact of disease-associated genetic variables in the wider setting of the complete genetic governing network are crucial for finding fresh targets for therapy.

Genotyping and Genome-wide association studies (GWAS):

GWASs uncover connections between SNP alleles with the presence or absence of a medical condition, which allows for finding variants that are linked to certain diseases (Penrod *et al*., 2011). Knowledge of the biological underpinnings of common ailments like diabetes, cardiovascular disease, infectious diseases, inflammatory and autoimmune disorders, and more has advanced through the identification of over 1200 genomic areas by GWASs (Spreafico *et al*., 2020). With little consideration for previous biological information, GWASs are intended to detect disease-associated variations based on statistical correlations, which offers researchers a starting point to formulate testable hypotheses addressing the genetic basis of disease (Penrod *et al*., 2011).

Functional genomics:

Nowadays, an increasing number of functional assays harness technological advancements and available DNA sequence data to gather functional data at the genomic level. The emergence of novel research domains such as transcriptomics (which includes assessing gene expression profiles), proteomics (which involves

analysing protein expression profiles), cistromics (which entails investigating transcription factor binding profiles), and epigenomics (which requires DNA methylation and chromatin remodelling) has been facilitated by such assays (Penrod *et al*., 2011).

Network pharmacology:

Network analysis offers a means of modelling the intricate relationships that exist both inside and among the various tiers of genetic information and govern biological activity in context to a potential drug candidate (Chandran *et al*., 2017). The framework is implemented as a symbolic illustration of genomic information, wherein nodes stand for a number of genomic units (e.g., mRNA, gene, protein), and edges denote some sort of connection or collaboration, such as transcriptional control, concurrent expression, or tangible contact within the node coordinates they link.

The era of genomics has thus ushered a technological boom that altered data collection methods, which has resulted in concurrent unbiased screens for therapeutic targets in the drug discovery process and unbiased screens for genes or genetic variations causing diseases (Penrod *et al*., 2011).

2.3.2 Identification of Drug Targets using Subtractive Genomics

Drug discovery has benefited greatly from the breakthroughs in computational biology and bioinformatics techniques that employ omics data, such as genomics, proteomics and metabolomics, thus bringing down the expense and length of in vivo and experimental testing for drug development. Subtractive genomics is one such approach for determining the sequence of a gene, protein, or generic area that resides in a wider genomic landscape (Asalone *et al*., 2019). By comparing the proteomes of the pathogen and the host, subtractive genomics enables identification of non-host

proteins with distinct metabolic pathways that are essential to the pathogens' survival (Asalone *et al*., 2019). In several investigations, this methodology has been applied to ascertain putative therapeutic targets and candidates for vaccines directed against various pathogenic bacteria. The technique starts off with extensive genomic data and applies Basic Local Alignment Search Tool (BLAST) against a reference sequence or sequences to eliminate matching alleles, solely retaining the target. The process essentially integrates multiple computing stages into an iterative cycle that sequentially excludes multiple reference sequences thus optimizing the quest for the target (Asalone *et al*., 2019). One of the primary advantages of subtractive genomics is that it is fair and independent given that it does not require prior positive identification of the target. Reductive genomics in this instance guarantees that the medications generated would not have any unintended consequences on the host. Although subtractive genomics is a robust technique, it is not a one-size-fits-all method, requiring customizations at key steps and precise selection of reference sequences (Barh *et al*., 2010). The choice of an acceptable e-value constitutes a crucial stage in the BLAST process as it dictates the stringency or relaxation of the subtraction. Since non-matching sequences are not subtracted, the approach inverts the idea of a more rigorous match to reference being a less severe subtraction (Asalone *et al*., 2019). Subtractive genomics has been used in numerous research recently to identify and recognize novel therapeutic targets that are particular to a given species. Using this method, recent research has predicted two distinct metabolic pathways of *Mycoplasma genitalium* as possible targets for drugs: the phosphotransferase system (PTS) and the bacterial secretion system (Fatoba *et al*., 2021). The sophisticated *in-silico* genome subtraction method was used by Ashraf et al. in a different investigation to find 38 putative, pathogen-specific therapeutic

targets for *Moraxella catarrhalis (*Ashraf *et al*., 2022). The ability of *in-silico* subtractive genomics to quickly and affordably screen targets at the genome level accounts for its significance in the identification of therapeutic targets (Barh *et al*., 2010). Finding presumably essential genes in pathogens, which may be later on verified by mutagenesis studies, is another significant benefit.

2.3.3 Comparative Genomics in *Shigella*

The chances of a molecule or medication being therapeutically effective are increased when it interacts with a logical target or a particular set of targets (Zhang *et al*., 2022). By coupling massive amounts of data with computational methodologies, computer-aided target identification greatly decreases the number of experimental targets, minimizes the drug discovery and development cycle, and lowers experimental costs (Zhang *et al*., 2022). Comparative genomics has emerged as one of the most frequently implemented *in-silico* methodologies for finding possible drug targets. In most biological inquiries, it is now an accepted procedure to examine and compare sequence properties within and between species (Zhang *et al*., 2022). By aligning and evaluating the genes and genomes of extant or extinct creatures related by differing degrees of evolutionary divergence from a common ancestor, comparative genomics studies the processes of evolution. The foundation of comparative genomics is the idea that potential targets play a crucial role in pathogen survival and are an essential part of their metabolic processes (Zhang *et al*., 2022). Comparative genomics addresses two key features of therapeutic targets. One, the success of the newly developed medication is ensured by the specified targets' substantial effects on a few key physiological activities of the pathogen. Two, any adverse impacts on the human body when the drug interacts with the target can be prevented by comparing the protein sequences of possible therapeutic targets and the host to assess whether there

is homology (Zhang *et al*., 2022). This optimizes the safety of the physiological effects of novel pharmaceuticals. In a similar vein, comparative genomics has also been used to pinpoint treatment targets from bacterial pathogens. For instance, a comparative genomic analysis carried out by Wang et al. with 15 clinically obtained strains of *Shigella flexneri* demonstrated the coevolution of antibiotic resistance and virulence factors in *Shigella* (Wang *et al*., 2019). Similar genomic analyses of 12 *Shigella sonnei* strains were conducted in a different study by Zhu et al. The findings demonstrated a twofold increase in the number of type 2 secretion system virulence factors amongst resistant bacteria and a positive correlation between higher levels of antibiotic resistance and an abundance of virulence markers (Zhu *et al*., 2021). An *in-silico* comparative genomics approach was implemented in a study by Shrestha et al. to foresee pathogenic trehalase genes in the Antarctic *Shigella* sp. PAMC28760 genome. The study provided insights into how bacteria use readily available carbohydrates like glucose, which are produced by the trehalose degradation pathway, and it identified two different types of trehalase genes in several *Shigella* species, thereby emphasizing the significance of periplasmic trehalase in bacterial virulence (Shrestha *et al*., 2022). In another research project by Quereshi et al., two novel therapeutic targets and four drug-like compounds with likely inhibitory actions against 27 strains of *Shigella dysenteriae* were identified through the use of an integrated approach combining comparative genomics and subtractive genomics (Qureshi *et al*., 2024). Comparative genomics thus integrates data gathered from genome information records and algorithmic programs to expose pathogens' catastrophic flaws which impact how they thrive and reproduce within the human body, involving traits required by the pathogen to endure, develop, and

perform important functions, making this approach the gold standard for identifying therapeutic targets.

2.3.4 Pangenome Analysis in Core Target Identification

As sequencing technologies have advanced, the realm of genomic analysis has evolved, switching from single strain to species (or even higher taxa)-wide genomic resolution in an effort to document the "totality" of biological variety (Vernikos, 2020). The term "pangenome" was first used over 20 years ago by Tettelin et al. (2005) in an effort to characterize and simulate the entire genetic makeup of a particular taxon (Tettelin *et al*., 2005). The pangenome is a compilation of the entire DNA pool of a species which include sequences common to all individuals (core genome), as well as details about sequences distinctive for every individual (unique genome) (Gong *et al*., 2023). Pangenomes have progressively emerged as the new reference points for genomics research owing to the abundance of available genomic data and the shortcomings in earlier methods of describing the genetic variation across species. Pangenome analysis at all possible phylogenetic precision scale, is carried out utilizing an array of modelling frameworks, presumptions, and fundamental homology algorithms (Vernikos, 2020). The implementation of pan-genome analysis has become increasingly recognized as a valuable method to comprehend a clade of pathogenic bacteria, as it aids in the development of therapeutic strategies that are operational against the entire clan as well as towards a specific species or strain within a clade, based on biological similarities and differences across the group (Kim *et al*., 2020). A comprehensive pan-genome investigation entails the acquisition of genomic sequences, orthology grouping based on sequence alignment of multiple allele copies and screening of central and supplementary genomes. Phylogenetic distances, the presence or absence pattern of

genes within a target class, and the functional distribution of proteins are a few instances of relevant scientific information that can be derived from the pan-genome analysis (Kim *et al*., 2020). Finding the genes that code for necessary functionalities can be accomplished by pangenome analysis. Genes associated to fundamental housekeeping tasks, like metabolic activities, cell membrane proteins, transshipment proteins, regulatory factors, and binders make up for the majority of the genes that are conserved across strains in an organism's core genome. As a result, a pool of possible therapeutic targets is provided by the core genome discovered through pangenome analysis (Anani *et al*., 2020). Despite the fact that pangenome analysis has not yet produced any antimicrobial medications, a number of published studies have demonstrated the pangenome's potential to identify logically acceptable therapeutic targets in pathogens like *Helicobacter pylori*, and *Clostridium botulinum* and *Mycobacterium tuberculosis* (Anani *et al*., 2020). Pangenome analysis techniques were utilized in an alternative study by Fatoba et al. to determine potential therapeutic targets from *Clostridium difficile* using *in-silico* methods (Fatoba *et al*., 2022). Pangenome screening yielded a total of 2556 core genes, of which 5 cytoplasmic proteins were ultimately projected to be unique prospective targets for therapeutic intervention. The potential value of two component systems (TCSs) as possible therapeutic targets against ESKAPEE (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter* spp., and *Escherichia coli*) pathogens was established through pangenome analysis in another study by Rajput et al. In addition to revealing a variety of tactics used by the ESKAPEE pathogens to demonstrate pathogenicity and antibiotic resistance through two-component systems, this pangenome assessment also raised the possibility of TCSs as a desirable set of prospective targets for

addressing antibiotic resistance considering their conserved characteristics (Rajput *et al*., 2021). With an unprecedented rise in accessible genome sequences, pangenome investigations, which allow an unparalleled window into the varied genomes of human pathogens, holds the ability to become an indispensable instrument in medical microbiology in the years to come.

2.4 Phytochemicals as An Arsenal against Antibiotic Resistance

The negligent application of antibiotics in medical care has resulted in the rapid emergence, persistence, and dissemination of antimicrobial resistance (Khare *et al*., 2021). Coupled with the prolonged illness and rising costs, the futility of current antibiotics comes as a consequence of expanding antibiotic-resistance mechanisms in microorganisms (Khare *et al*., 2021). The World Health Organization has released the first-ever compilation of priority pathogens resistant to antibiotics, featuring a list of twelve bacterial families that pose the biggest threats to human wellness (World Health Organization, 2017). Thus, there exists a dire requirement for innovative and unique therapies as well as novel potent medications to treat MDR infections. Aside from the various natural and synthetic medications being explored, plants could possibly own the code to offering a vast spectrum of therapeutics in the form of their secondary metabolites with the potential for combating microbial infections. This dates back to the paper published in 1803 detailing the extraction of morphine from *Papaver somniferum* (Schmitz, 1985). The members of the coumerins, flavonoids, alkaloids, quinones, and several other groups are examples of these secondary metabolites, or phytochemicals. For instance, Dahiya and Purkayastha, examined ten MDR clinical isolates (containing Gram-positive and Gram-negative bacteria) using several solvent-based extracts of medicinal plants, such as Lemongrass, Neem, Aloevera, Oregano, Rosemary, Thyme, Tulsi, and Bryophyllum (Dahiya and

Purkayastha, 2012). The antibacterial activity against methicillin-resistant *Staphylococcus aureus* were shown as being linked with both tannins and flavonoids found in a methanolic preparation. In addition to the extracts demonstrating a complex combination of several phytochemicals in one formulation, several of the individual proactive components in their pure state were additionally evaluated for their ability to inhibit extensively drug resistant organisms and the mechanism they are intended to affect. For example, the isoquinoline alkaloid berberine, which is present in the roots and stem of the Berberis species, has been observed to demonstrate antibacterial activity (Yi *et al*., 2007). Phytochemicals have shown the capacity to forestall crucial factors that contribute to resistance-gaining, involving penetration of cells, pumps for effusion, replicating apparatus and further steps essential to the bacteria's sustenance and resilience, thus presenting themselves as a viable substitute to the dwindling supply of traditional antibiotics.

2.4.1 New Horizons in Shigellosis Treatment

A successful treatment for shigellosis typically manifests itself by limiting the spread of the bacteria and minimizing the infection's course (Ranjbar and Farahani, 2019). Conventional antibiotic therapies are incredibly ineffective against shigellosis due to the diverse range of antibiotic-resistance mechanisms, which has led to the rise in extensively drug resistant *Shigella* strains. This makes the creation of novel therapeutics and complementary approaches necessary for the avoidance and management of *Shigella* infections.

Natural and organic ingredients-based therapy options:

It has been documented that natural product, consisting of diverse chemicals originating from microorganisms, plants, and animals, have benefits in treating

Shigella infections. The synthesis of unified organic acids, bacteriocin and organic acid derivatives, have all been proposed as ways to induce antibacterial activity of bacteria against enteric pathogens. Zhang and colleagues, for instance, curated 91 *Lactobacilli* for evaluation of antimicrobial activity against *Shigella* isolates; of which, 16 *Lactobacilli* illustrated significant antibacterial activity against strains of *Shigella sonnei* (Zhang *et al*., 2011).

Shigella treatment with nanoparticles:

An increasing proportion of investigations have centered around nanoparticles in light of their bactericidal properties, that have displayed wide-ranging antimicrobial properties against harmful microbes. In addition to producing free oxygen radicals, nanoparticles typically damage bacterial targets, inflicting damage to membranes and disrupting their integrity (Ranjbar and Farahani, 2019). Nanoparticles made of copper oxide have just been discovered as an antibacterial for handling shigellosis. Both Fe2O3 and Ag–Fe2O3 nanoparticles were found to exhibit antimicrobial effects against *Shigella dysenteriae* strains in a study that assessed their antibacterial qualities (Kareem *et al*., 2018). Treatment for drug-resistant *Shigella* also involves the use of formulations containing nano antibiotics. The synthesis of a nanosized form of tetracycline by loading it in calcium phosphate nanoparticles was reported by Mukherjee et al. (Mukherjee *et al*., 2019) Their findings revealed that this treatment significantly reduced the likelihood of microbial colonization, loss of weight, mushy-stool discharge, and colon-length reduction in the digestive tracts of *Shigella*-infected mice.

Phage Intervention:

Phage therapy has gained recognition for its many benefits, which include but are not limited to less adverse reactions compared to other medications, high specificity for selecting bacteria with minimal impact on the body's regular microbiome, site specific replication, and its ability to selectively kill drug resistant pathogens. Research conducted by Jamal et al. have looked into the use of phages as a potential substitute to antibiotics for treating MDR *Shigella dysenteriae* isolated from wastewater (Jamal *et al*., 2015).

Vaccination Methods:

A number of prospective vaccines have lately been designed to forestall infection by *Shigella* spp., the bulk of these are presently getting scrutinized to ensure security and antigenicity. Vaccine developers have deployed various strategies, that can be put into two categories: parenteral subunit strategies, which primarily involve glycoconjugates and seek to focus on the immune response to crucial *Shigella* antigens, and oral whole cell strategies, which chiefly involve live attenuated as well as inactivated vaccines (MacLennan *et al*., 2022). Recently, a *Shigella sonnei* O-antigen glycoconjugate vaccine designed by John Robbins and associates at the National Institutes of Health (NIH), USA, was shown to be fruitful in immunizing Israeli military recruits (MacLennan *et al*., 2022). However, presently there exists no licensed vaccine for safeguarding against shigellosis, even after more than a century of research into *Shigella* vaccine development and dozens of clinical trials. With a better knowledge of the foundation for *Shigella* protection and numerous candidate vaccines moving on to late-stage clinical trials, this situation is expected to alter in the years to come.

2.4.2 The Treasure of Ethnomedicine

The scientific study of "traditional medicine"—the wisdom and customs that indigenous peoples have passed down orally over generations and refined all along centuries of mankind's existence—is widely regarded as "ethnomedicine" (Mahapatra *et al*., 2019). In broad terms, ethnomedicine relates to the conventional medical procedures that focus on how different communities perceive wellness, disease, and illness with the mission to deliver care and improve recovery. This art of healing has been prevalent in many historically significant civilizations, ranging from African Muti to Mughal Indian Unani medicine, and from Indian Ayurveda to Traditional Chinese Medicine (TCM) China (Mahapatra *et al*., 2019). As a testament to their convictions in "healing from within", the main thrust of their therapy was to first relieve the patient's pain prior to finding the precise cause of it. Between 3000 and 2730 BC, the ancient Chinese utilized species of *Hydnocarpus* to treat leprosy symptoms. The discovery of castor beans and opium poppies in Egyptian tombs demonstrated that phytomedicine was used in Africa as early as 1500 BC. Over 750 plants, including *Shorea*, *Aconitum*, *Clitoria*, *Cosinium*, and many more, are used in Ayurveda, the earliest-surviving healthcare system in India, which dates back to approximately 5000 BC (Mahapatra et al., 2019). One of the primary tenets of Ayurveda is the application of plant-based remedies utilizing various herbal components, such as roots, leaves, fruits, bark, or even seeds, in conjunction with the therapeutic resources used by ethnic communities (Mahapatra *et al*., 2019). Of the estimated 645 indigenous tribes who call India home, most have not yet had their rich understanding of the nation's wildlife and plants spelled out. This expertise has emerged as invaluable for dealing with and curing diseases. The last couple of years have witnessed an emphasis on endorsement of such unexplored knowledge in

Indian research policy (Mahapatra *et al*., 2019). The antiviral activity of a useful herb, *Ophiorrhiza nicobarica*, collected from the Galathia River basin of Great Nicobar Island, and widely used for skin ailments by the local and tribal communities, is one of the many effective traditional medicines of the Bay Islands identified and validated by Chattopadhyay et al (Chattopadhyay *et al*., 2003). Likewise, it has been discovered that the stem bark of *Odina wodier*, a traditional Jangalmahal remedy used to treat ulcers, cardiac disorders, and skin diseases, can protect HSV-1-infected animals by preventing viral replication via host immune regulation (Ojha *et al*., 2013). Creative strategies that will expose and help bring the entire spectrum of molecular plurality of such naturally occurring elements into the course of developing drugs will be sought as new routes for pharmaceutical research arise.

2.4.3 Traditional Medicines against Diarrheal Diseases

Millions of people worldwide rely on alternative health care systems that include traditional medications as a key component. The treatment of diarrhoea and related gastrointestinal diseases also relies on the consumption of many different indigenous medicinal herbs. This evidenced by the use of opium (*Papaver somniferum*) combined with minerals as a herbo-mineral formula against diarrhoea, documented in the ancient texts of Yoga Ratnakara (1700–1800 CE), and the use of opium, camphor (*Cinnamomum camphora*) and marijuana (*Cannabis indica*) for diarrhoea in the Bhaisajya Ratnavali book (Mahapatra *et al*., 2019). A multitude of mechanisms, including increased intestinal absorption, anti-motility or anti-peristaltic impact, antibacterial activity, and anti-spasmodic action, have been identified as contributing to the anti-diarrheal function of most herbal extracts (Rawat *et al*., 2017). In a randomized, double blind, placebo-controlled clinical trial, the effectiveness of *Mentha piperita*, popularly known as peppermint, against diarrhoea was examined.

43 (83%) of the 52 individuals receiving peppermint therapy had less stools and exhibited improvement in other symptoms (Tyagi and Malik, 2011). Subsequent research revealed the anti-microbial action of essential oil derived from *M. piperita* against a variety of microorganisms, like *Escherichia coli*, *Pseudomonas aeruginosa*, *Bacillus subtilis*, and *Staphylococcus aureus*. To assess the reliability and efficacy of TXNG, a Chinese polyherbal formulation made up of four herbs—*Paeonia lactiflora* (root), *Atractylodes macrocephala* (rhizome), *Citrus reticulate* (green, unripe exocarp), and *Allium macrostemon* (bulb)—a randomized, placebo-controlled, double-blind clinical trial was carried out on a total of 60 subjects with diarrhoea over a 10-month period (Rawat *et al*., 2017). When comparing the treatment group to the placebo group, the TXNG group's diarrhoea relief time was noted to be 47% shorter, along with a noticeable improvement in the form and appearance of the stools as well as a reduction in their frequency. A number of phytochemicals that were extracted from the plants were also evaluated for their ability to shield against diarrhoea. For example, a castor oil-induced diarrhoea model was used to study 1,8-cineole, a constituent found in varying amounts in several biologically active aromatic herbs and spice oils, including *Mentha longifolia*, *Artemisia dracunculus*, *Coriandrum sativum*, *Origanum vulgare*, *Rosmarinus officinalis*, *Thymus vulgaris*, and *Zingiber officina* (Jalilzadeh-Amin and Maham, 2014). The study showed notable anti-spasmodic and anti-secretory efficacy, with a 37% delay in the commencement of diarrhoea in the treatment group compared to the control group. Friedelin, which was extracted from *Azima tetracantha* (Salvadoraceae) leaves, was tested for its anti-diarrheal properties using models of diarrhoea caused by magnesium sulfate and castor oil (Antonisamy *et al*., 2015). At a dosage of 20 mg/kg, the ingredient produced 89% inhibition in diarrhoea generated by castor oil, while 83% inhibition

was observed in diarrhea induced by magnesium sulfate. Although there are several traditional anti-diarrheal plants that are readily accessible and have been shown to be profitable, there hasn't been much focus on using these phytopreparations for ameliorating diarrhea and related conditions (Rawat *et al*., 2017). These promising herbal extracts thus deserve to be deployed as a source for novel alternative medicine development, tapped to identify active ingredients, and subjected to rigorous pre-clinical and clinical studies to ensure their security and effectiveness.

2.4.4 Revival of interest in Pharmacologically Active Plant Ingredients

The early 1980s have witnessed a decline in the use of herbal products, primarily due to technological obstacles to high-throughput studies for natural product profiling against target candidates and worries about the challenges involved in repeatedly isolating already recognized chemicals (Jantan *et al*., 2015). The poor rate of triumph in finding therapeutics derived from flora together with innovations in synthetic organic chemistry and biotechnology that offered greater chances for building novel medicines in the lab contributed to the declining interest in embracing plants as a source of drugs (Jantan *et al*., 2015). Nevertheless, in the latter half of the 1980s, ethnobotany-driven drug development saw a resurgence, primarily because of developments in chromatographic and spectroscopic techniques, which greatly aided in the extraction and structural characterization of the constituents of medicinal plants as well as the formulation of a number of quick, simple, quantitative bioassays that allowed for the detection of bioactive molecules even at low levels of concentration (Jantan *et al*., 2015). A further factor driving the current renewal of attention in natural product research is the unappealing yield of powerful chemical discoveries from combinatorial chemistry and high throughput screening (HTS) (Harvey *et al*., 2015). The conventional method of isolating natural products using bioassay guidance has

been altered to capitalize on technological developments, adapt to the state-of-the-art clinical chemistry knowledge, and investigate chemical space that is physiologically important. The bottleneck in structure-elucidation and isolation has been overcome through the use of sensitive NMR techniques in conjunction with simpler fractions generated by chromatographic methods. The magnitude of screening can thus be lowered and the pace increased by preparing fractions of lower complexity (Harvey *et al*., 2015). Natural-product fraction libraries can be subjected to the pre-filtering method used in present pharmaceutical research, which involves screening libraries for lead or drug resemblance. An investigation has demonstrated that a screening library of highly varied plant-derived ingredients could be created by analysing the Dictionary of Natural Products, which would facilitate the pre-selection of compounds based on the similitude of their physicochemical features to established drugs (Jantan *et al*., 2015). On the other hand, the use of chemoinformatic techniques in the process of discovering new drugs from natural products can be equally beneficial. These technologies enable compounds to be evaluated for attributes related to absorption, distribution, metabolism, excretion, and toxicity (ADMET) prior to enrolment in drug development programs (Harvey *et al*., 2015). Metabolomics technologies have the potential to be proven advantageous in natural product discovery processes on several fronts. The arena of metabolomics emerged from metabolic profiling and aims to analyse all of an organism's metabolites both quantitatively and qualitatively at a given time and under a given set of circumstances (Harvey et al., 2015). The three main analytical platforms include NMR spectroscopy, liquid chromatography (LC) or capillary electrophoresis (CE), and mass spectrometry with gas chromatography (MS-GC) (Jantan *et al*., 2015). To preferentially synthesize biologically active secondary metabolites, a biosynthetic pathway can be optimized

using an integrated approach of genetics and metabolomics (Jantan *et al*., 2015). By mining metabolomics data, biosynthetic precursors can be identified, which upon manipulation using engineering pathways, can boost the yield of the operational natural product (Harvey et al., 2015). For instance, *Nicotiana benthamiana*, the tobacco plant, has been investigated as a potential source for precursors of artemisinin, which is typically isolated from *Artemisia annua* to fix malaria (van Herpen *et al*., 2010). Through the agro-infiltration of *N. benthamiana*, the genes encoding the three biosynthetic enzymes required to synthesize artemisinic acid were expressed. Approximately twice as much artemisinic acid 12-beta-diglucoside was produced by this heterologous route in tobacco plants as compared to *A. annua* (van Herpen *et al*., 2010). Therefore, the holistic strategy that combines traditional and modern techniques—like botanical, phytochemical and pharmaceutical chemistry, HTS, biological, computational, molecular and multi-target approach, metabolomics, proteomics, and genomic techniques—has rendered the applied problems of natural product drug discovery simpler to address (Jantan *et al*., 2015). It is highly probable that the great majority of plant species have not been thoroughly examined in efforts to find new drugs product (Harvey et al., 2015). In the long run, innovations in technology might be enough to spur the worth of using natural ingredients as foundations for drug discovery exploration.

2.5 *In-Silico* Approaches for Therapeutic Innovation

In-silico methods are crucial to determining both targets and lead compounds, as a means of expediting the process of finding novel drugs and reducing the upfront costs associated with bringing a product to market. Similar to the Latin terms in vivo, in vitro, and in situ, the term "*in-silico*" was first used in 1989. In simple terms, *in-silico* drug design and drug discovery refer to the logical design of drugs by computational

techniques that aid in the discovery of targets and drug-like substances through the use of bioinformatics tools (Jabalia *et al*., 2021). Implementing *in-silico* approaches comes with multiple ramifications for the drug discovery process: (1) comprehending target elements with active sites for drug binding; (2) gauging drug-like characteristics; (3) scrutinizing drug molecules' affinity for association with targets; (4) comparing the highest-scoring molecules; and (5) optimizing them more extensively (Jabalia *et al*., 2021). *In-silico* drug discovery has led to the discovery of numerous potential medicinal candidates thus far. For example, raltegravir, an antiviral drug, was discovered using *in-silico* techniques (Miah and Mohd Aluwi, 2024). Another instance is the cancer drug vemurafenib, which came into existence through virtual screening. Moreover, employing *in-silico* methods, promising COVID-19 therapy candidates have been established, including the drug molnupiravir (Miah and Mohd Aluwi, 2024). The possibility of a novel drug coming to fruition escalates when *in-silico* techniques receive integration into the interdisciplinary drug discovery workflow. As a result, *in-silico* drug discovery methods give researchers a huge edge over random screening of billions of molecules. Industry projections reveal that leveraging an *in-silico* approach could preserve roughly 30 percent of the time and expenses associated with the drug discovery process (Jabalia *et al*., 2021).

2.5.1 Computer-aided Drug Discovery

Computer-aided drug design (CADD) approaches have grown to become an important and fundamental element in the drug research and development procedure (Sadybekov and Katritch, 2023). CADD capitalizes on two different approaches: structure-based drug discovery (SBDD) and ligand-based drug discovery (LBDD) (Shaker *et al*., 2021). The availability of structural information about the target protein

influences the choice of an appropriate CADD technique. For success with the SBDD technique, structural data about the target protein is required, which can frequently be collected experimentally using nuclear magnetic resonance or X-ray crystallography (Shaker *et al*., 2021). The 3D structure of the target protein can be predicted using *in-silico* prediction techniques like homology modelling or ab initio modelling in absence of experimentally derived coordinates. Upon the availability of the structure, molecular docking and structure-based virtual screening become feasible. In instances when *in-silico* approaches are unable to anticipate a high-quality structure or when the structure is unobtainable, the LBDD methodology is frequently used as a replacement strategy (Shaker *et al*., 2021). Algorithmic target identification, virtual screening of vast chemical archives for promising therapeutic candidates, subsequent refinement of prospective compounds, and *in-silico* evaluation of their potential toxicity are all included in CADD. Homology modelling, virtual screening, molecular docking, quantitative structure activity relationship (QSAR), and pharmacophore modelling are the most noteworthy and routinely employed *in-silico* approaches in the course of drug development (Jabalia *et al*., 2021). Protein structures are generated via homology modelling based on data on related proteins. Finding ligands that bind most favourably within receptor binding sites and figuring out the receptor's most energetically preferred binding orientations are the main objectives of molecular docking (Baig *et al*., 2018). The scoring function rates the generated poses according to their binding affinities and determines the most advantageous receptor/ligand binding modes while the search algorithm is in charge of looking for various ligand poses and conformations within a specific target protein. The virtual screening method is an effective strategy for seeking possible ligands, primarily because it leverages the information base that already exists (Jabalia *et al*., 2021). The virtual

screening method is a productive way to identify low molecular weight, readily accessible chemicals that have the ability to bind firmly with target molecules as ligands (Jabalia *et al*., 2021). The foundation of QSAR techniques is the idea that biological activities are directly correlated with molecular frameworks, and that biological activities are subsequently modified by molecular or structural alterations (Baig *et al*., 2018). The process of creating mathematical or computational models utilizing chemometric techniques in order to find meaningful connections between a number of structures and functions is known as QSAR. The application of CADD, therefore, lends to the most notable responsibility in the current drug discovery program and delivers computational tools and algorithms that restore money, time, and minimize the chance of yielding non-viable developmental leads (Baig *et al*., 2018).

2.5.2 Virtual Screening

A high-throughput computerized approach called virtual screening is deployed in drug discovery initiatives to find the hit compounds by scanning a pool of ligands against a biological target (Rizzuti and Grande, 2020). Analysing a compound's ability to bind to a target is an aspect of the screening process. Virtual screening is the broad term for the practice of rating and evaluating compounds in huge chemical libraries based on how likely they are to possess affinity for a particular target (Rizzuti and Grande, 2020). The phrase itself first came into use in the late 1990s, as computer-based procedures became adequately sophisticated to deliver an option to conventional high-throughput screening (HTS) methods, which were regrettably performing worse and costing more than expected. An information-driven method, virtual screening necessitates structural data about the target itself (target-based virtual screening) or about bioactive compounds for the target of interest (ligand-

based virtual screening) (Lavecchia and Di Giovanni, 2013). So, without significantly affecting performance in advance, the final decision about which approach to employ will primarily rely on the kind and volume of data that is accessible. Virtual screening techniques built around ligands come in a wide variety. Chemical analogy estimates are the foundation of ligand-based virtual screening since all of these techniques are hinged on the core similarity-property concept, which holds that molecules with comparable features should display similar properties (Rizzuti and Grande, 2020). As a result, each molecule in a given database can be ordered according to decreasing likelihood of being active by first scoring its resemblance to one or more bioactive ligands (Ekins *et al*., 2007). Topological signatures implying the existence of substructural pieces in molecules have been undoubtedly the most popular structure interpretations for ligand-based virtual screening. Techniques based on geometrical illustrations of molecular structures can be employed as an alternative to topological approaches. Among them, a tried-and-true virtual screening strategy involves the flexible superposition of molecules onto one or more conformations of a reference bioactive ligand (Ekins *et al*., 2007). Target-based virtual screening techniques rely on the truthfulness of the target's structural details, which can be obtained computationally using homology modelling methods or established experimentally. Such approaches seek to present a realistic assumption of the binding affinity of the ligand as well as a decent estimate of the anticipated form and alignment of the ligand into the protein pocket (Ekins *et al*., 2007). Virtual screening recruits an array of docking techniques. Docking algorithms like as Autodock4, Autodock Vina, DOCK6, GOLD, and GLIDE are frequently employed for virtual screening of small compounds. A slew of novel approaches for evaluating the pharmacological profile of compounds on numerous targets have recently emerged, adding to the biological

component of virtual screening. These have the power to have an immediate effect on drug discovery by realizing the possible adverse reactions of compounds caused by off-target affinities during the optimization phase (Ekins *et al*., 2007). The introduction of documented chemical libraries, that bring literature-based pharmacological data into standard chemical archives, has been instrumental in the advancement of ligand-based affinity profiling tools. The establishment of target-based affinity monitoring approaches has made profit of the truly exponential expansion in the sheer number of experimentally characterized protein structures, thereby offering functional elucidation of protein families (Rizzuti and Grande, 2020). Furthermore, a great deal of attention has recently been given to the implementation of docking methods to targets for which empirically established structures are not yet available. This is especially true for the numerous clinically significant targets that are members of the superfamily of G-protein-coupled receptors (GPCRs) (Ekins *et al*., 2007). As opposed to random selection or other conventional methods, these approaches show an evolving ability to provide enrichment in figuring out active compounds for the target of interest, which is a step forward in drug innovation programs (Lavecchia and Di Giovanni, 2013). To sum up, virtual screening is a game-changing tool that sits at the nexus of computational power and biological complexity in contemporary drug discovery.

2.5.3 Lead Discovery and Optimization

There are certainly multiple phases involved in the discovery of a pharmaceutical product, including Hit to Lead and Lead Optimization (Arya and Coumar, 2021). The Hit to Lead phase involves assessing the activity of small molecules and how they interact with the target to produce lead compounds. In the realm of pharmaceutical research, a chemical entity with biological or pharmacological activity is termed a

lead compound (Arya and Coumar, 2021). Lead optimization, the last phase of drug discovery, aims to minimize structural flaws while simultaneously preserving or enhancing the desired characteristics of the compounds that have been chosen and thus represents the process of producing and refining a pre-selected lead compound (Arya and Coumar, 2021). It involves chemical alterations to the compound and modifies several of its properties to satisfy all necessary stereoelectronic, physicochemical, pharmacokinetic, and toxicologic standards for its therapeutic use. Computational techniques serve as a viable approach to investigate lead generation and optimization. This phase entails the use of computerized methods like QSAR, molecular docking, pharmacophore investigations, and molecular dynamics. Lead compound generation can be achieved in a variety of ways, broadly classified into two categories: database searching and de novo design (Xiang *et al*., 2012). De novo synthesis produces novel chemicals with intended pharmacological attributes in the binding region of a target protein, whereas database searching finds pre-existing molecules from chemical compound catalogues. When a high-resolution framework for the target protein is accessible, de novo design can theoretically be achieved by algorithmically assembling molecular fragments to create an entity with the requisite structural and electrostatic properties compatible with the target protein (Xiang *et al*., 2012). Database searching appears as an option to de novo design. Target proteins containing recognized binding sites and well depicted three-dimensional structures can be used to sift databases of known chemicals to uncover lead compounds (Xiang *et al*., 2012). Another particular approach used to identify lead compounds in paucity of receptor structure information is pharmacophore-based database mining. The three-dimensional organization of the key elements that lend a medication its biological impact is called a pharmacophore. Undoubtedly, evaluating the quality of

lead compounds that have been identified remains a crucial task, which is frequently manifested in terms of binding affinity with the intended target receptor (Xiang *et al*., 2012). It is widely accepted that a compound's affinity for a macromolecular receptor closely relates to its biological action. The key elements in determining the affinity are factors like the electrostatic interactions between the ligand and the receptor, the solvation and desolvation investments, the geometric complementarity of both binding partners, the enthalpic and entropic expenditures resulting from shifts in the number of degrees of freedom, and the structural modifications of the ligand and receptor encountered over complex formation (Verma and Pathak, 2022). Scoring functions are frequently employed during the lead generation phase to accurately measure binding affinity. Three major categories can be used to classify scoring functions: knowledge-based, force-field-based, and empirical-based (Verma and Pathak, 2022). The force field scoring functions generally tally up the individual contributions from non-bond interactions such van de Waals and electrostatic interactions to predict the binding affinity of a protein-ligand complex. Empirical scoring functions make the assumption that a number of distinct, independent terms, such as hydrogen bonding, hydrophobic contacts, rotor terms, desolvation, etc., can be added together to determine binding affinity (Verma and Pathak, 2022). By fitting a regression model to a test set of protein-ligand complexes with known binding affinities, the weighting of components is roughly estimated. Large collections of protein-ligand complex structures undergo statistical evaluation to build the foundation for knowledge-based scoring systems (Verma and Pathak, 2022). An alternative to the conventional scoring functions is molecular dynamic simulations (MD). Using MD, the impact of specific solvent molecules on protein structure could be thoroughly investigated (Verma and Pathak, 2022). Additional thermodynamic

parameters, such as interaction energies and entropies, are also achievable using MD. Following lead discovery comes the step of lead optimization. At this point, efforts are focused on enhancing a lead compound's drug-like qualities by making minor structural modifications in accordance with the proposition that every minor alteration in a chemical structure result in small differences in bioactivity, either positive or negative (Barcelos *et al*., 2022). Lead optimization typically involves a methodical procedure that includes the construction of several analogues and testing of their biological activities in addition to computational analysis using structure-activity relationship (SAR) data, which offer recommendations for tweaking specific sites of lead molecules (Barcelos *et al*., 2022). Therefore, the two primary objectives of developing pharmaceuticals —lead discovery and lead optimization—have been substantially backed up by computerized drug development (CADD), which has shown remarkable success in the development and deployment of novel techniques.

3. Research Questions and Objectives of The Study

There is an immediate need to characterize, annotate and identify targets in the *Shigella* genomes so that we are able to find answers to the following questions:

• Q1: Is there any such set of genes, that are ubiquitously present across all *Shigella* genomes and constitute an integral part of the intricate pathogenesis network?

• Q2: Of the set of core genes, are there any such genes which code for uncharacterized proteins and which when annotated, can help us to reconcile the missing links in the pathogenesis cascade and thereby serve as valuable drug target candidates?

• Q3: Does our ethnomedicinal literature and practices offer us with insights on natural products of plant origin which can be effective against the potential targets?

• Q4: What are the alterations in expression of the genome that are brought about by the administration of existing drugs and proposed natural formulations?

In this study, we undertook the following objectives with the intent to find answers to the aims stated above.

• Answer to Q1: Critical analysis of the *Shigella* genome to identify core genome and pangenome from different strains of *Shigella* spp. followed by structural and functional characterization of disease modifier proteins encoded by the core genome to prioritize therapeutic targets.

Answer to Q2: Curation and functional characterization of Hypothetical Proteins (HPs) from *Shigella* to analyse their suitability as drug targets using comparative modelling, and structural analyses.

• Answer to Q3: Virtual screening of a library of plant based small molecules identified from ethnomedicinal literature and identified using standard Mass Spectrometry techniques for optimization of lead compounds.

• Answer to Q4: A comparative study of the gene expression profile to specifically investigate the genes that are differentially regulated upon exposure of the pathogen to the herbal formulations.

4. Materials and Methods

4.1 Overall workflow

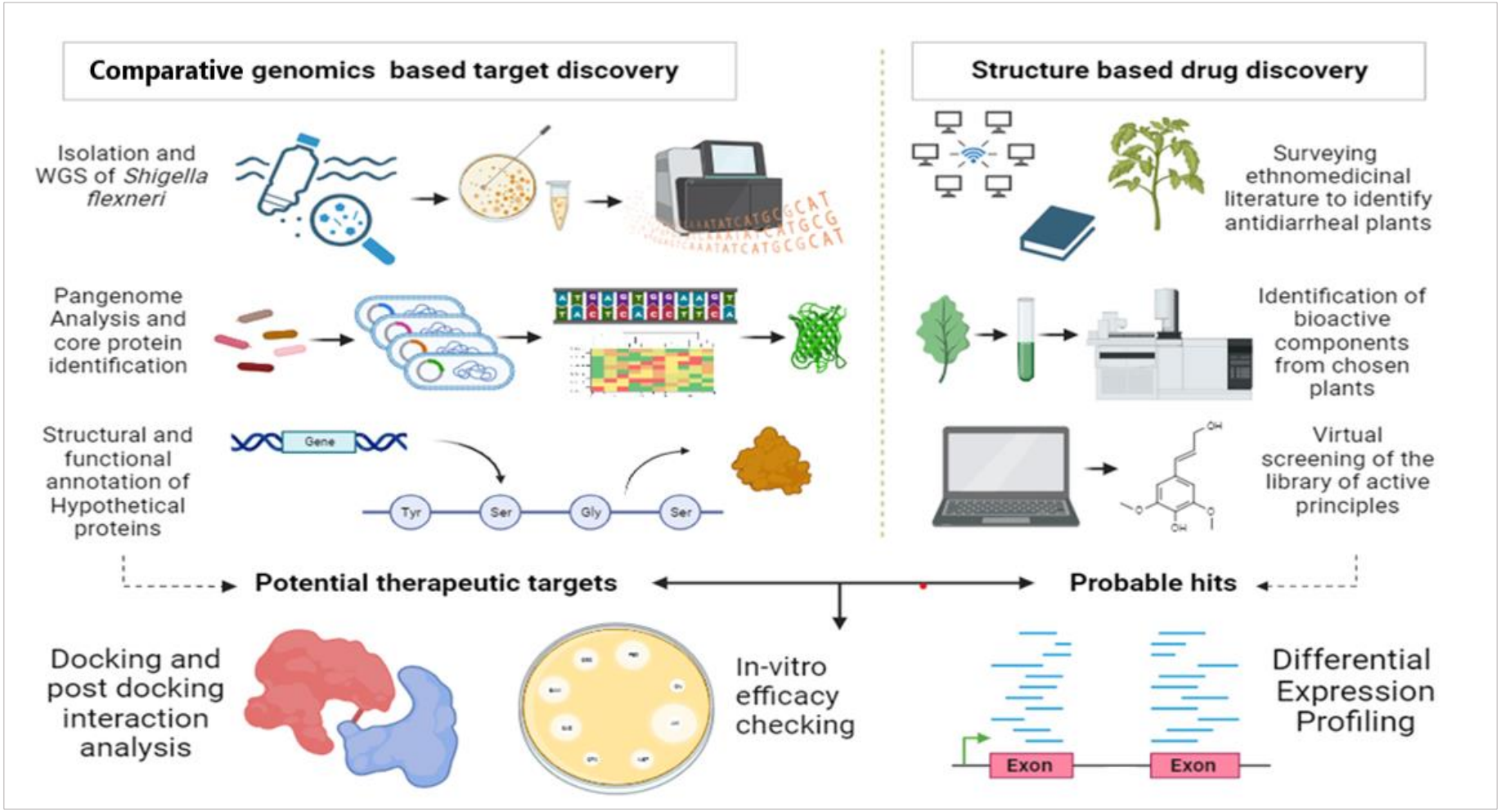


Figure 4.1.1 Scientific methodology.

The work can be broadly summarized under two heads, viz. Comparative genomics-based target discovery and Structure based lead discovery and optimization.

4.2 Comparative genomics-based target discovery

4.2.1 Selection of drug targets

The available *Shigella* genomes were analysed using a two-pronged approach where the first was to identify the core and accessory genomes. The core genomes and the corresponding pathogenesis proteins were then checked for their suitability as drug targets based on their functions and impact in the pathogenesis cascade. The second strategy was to annotate the potentially uncharacterized (hypothetical) proteins of the core genome, that had been exposed by the public sequencing projects of different *Shigella* strains and constitute a significant fraction of the available genomes.

4.2.1.1 Isolation and characterization of *Shigella flexneri*

4.2.1.1.1 Sample Collection

Purulia has been chosen as the location of this research. It is positioned between latitudes 22°42′35′′ and 23°42′00′′ North and between longitudes 85°49′25′′ and 86°54′37′′ East. The Chotanagpur plateau region contributes to the warm, humid temperature in this region, which can be explained by a dry and arid climate. The following climatic conditions were recorded at the time of collection: 42% relative humidity and 42°C temperature, Time of pickup: 2:00 p.m. A sample of liquid sludge was taken from the central sewerage ditch in the suburban Purulia region of West Bengal, India. This system picks up unprocessed hospital discharges, trash from household, and farmland runoffs. Following stringent aseptic procedures, the sample was drawn up in clean jars, moved to the research facilities in ten hours, and preserved at -20°C.

4.2.1.1.2 Pure Colony Isolation

After the collected sample was cleaned of soil particles and other granular materials using membrane filtering, it was serially diluted ten times (10^{-1}-10^{-6}). After removing and diluting 1ml of the filtered stock solution combined with 9ml of water obtained from distillation, the steps were reiterated up until a 10-6 gradation was reached. Following this, 100**µ**l of the diluted versions were spread on nutrient agar (NA) and allowed to incubate for a maximum of 48 hours at 37°C. Morphologically diverse colonies were identified by closely observing a plate with a fathomable count of distinguished independent colonies. The colonies were then picked up and plated fresh.

4.2.1.1.3 Biochemical Characterization

We subsequently conducted a series of biochemistry tests, that have been documented to deliver typical results for *Shigella* as a primary step to pinpoint the bacteria of preference, from the cultures of pure bacteria segregated thus far. The tests executed for the purpose (catalase test, citrate utilization test, oxidase test, indole test, lactose utilization test, MR test and VP test) were performed following the protocols in Bergey's Manual.

Catalase test: The catalase test's usefulness in the presumed distinction of specific Enterobacteriaceae has been documented in a number of investigations (Brenner *et al*., 2005). The hydrogen peroxide's bactericidal effects are countered by the catalase enzyme. Hydrogen peroxide (H_2O_2) is broken down faster by catalase into oxygen and water. The quick production of bubbles is indicative of this response.

$$2H_2O_2 + \text{Catalase} \rightarrow 2H_2O + O_2$$

Citrate utilization test: The citrate test is frequently a component of a panel of tests meant to detect environmental isolates and gram-negative bacteria. A bacterial isolate is tested for its capacity to use citrate as a source of carbon and energy using the citrate test (Brenner *et*

al., 2005). After being absorbed by the cell, citrate is broken down into acetate and oxaloacetate by citrate lyase. Following metabolization, the oxaloacetate yields CO_2 and pyruvate. Following its release, the carbon dioxide combines with the water and sodium ions present in the medium to form sodium carbonate, an alkaline substance that elevates pH. A change in pH indicator hue and the obvious presence of growth on the medium indicate that an organism is able to import citrate and use it as its exclusive source of carbon and energy.

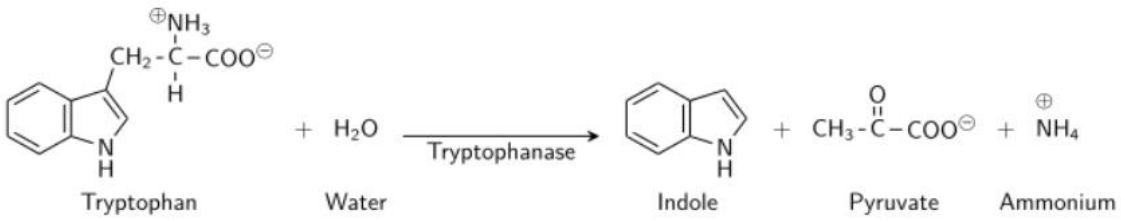
$$\text{citrate} = \text{oxaloacetate} + \text{acetate}$$

$$\text{oxalacetate} = \text{pyruvate} + CO_2$$

Oxidase test: An enzyme sometimes referred to as indophenol oxidase, cytochrome oxidase, is detected by the oxidase test, a biochemical process (Brenner *et al*., 2005). The utilization of the enzyme cytochrome oxidase, which catalyses the oxidation of cytochrome c while reducing oxygen to generate water, is the last step in the respiratory chain of bacteria. Tetra-methyl-p-phenylenediamine dihydrochloride (Kovács oxidase reagent) is a reagent that is frequently used in the oxidase test to act as a synthetic electron donor for cytochrome c. The reduced colourless reagent oxidizes and transforms into the dark blue or purple chemical indophenol blue when an organism containing the cytochrome oxidase enzyme is present.

$$2 \text{ reduced Cytochrome} + 2H^+ + 1/2\ O_2 \rightarrow 2 \text{ oxidized cytochrome} + H_2O$$

Indole test: The purpose of the indole test is to determine whether an organism can break down the amino acid tryptophan and create indole. It is a component of the IMViC



procedures, which are a series of examinations intended to differentiate between Enterobacteriaceae species (Brenner *et al*., 2005). When a microorganism is cultivated in a tryptophan-rich media, the presence of indole indicates that the organism has the ability to break down tryptophan. As a byproduct of tryptophan metabolism, indole can be detected by its chemical interaction with p-dimethylaminobenzaldehyde (DMAB) in an acidic environment, which yields the red dye rosindole.

Lactose utilization test: A lactose utilization test determines whether microbes are capable of fermenting lactose, a particular type of carbohydrate (Brenner *et al*., 2005). Enterobacteriaceae members can be distinguished from one another by their fermentation patterns. The final byproduct of lactose fermentation is either lactic acid or acid with gas. When acid products are produced, a pH indicator's colour changes, signalling a fermentation process. The pH only increases when lactose fermentation produces excess acid since bacteria can also consume peptones in the medium to make alkaline byproducts.

Methyl Red (MR) and Voges-Proskauer (VP) test: The Voges-Proskauer and methyl red tests are among the biochemical tests performed in clinical labs that are referred to as IMViC. Members of the Enterobacteriaceae family can be distinguished from one another using the paired MR-VP tests (Brenner *et al*., 2005). *Escherichia coli* and other low-ratio organisms use the mixed acid route to ferment carbohydrates, producing gas with a low CO_2 to H_2 ratio as a byproduct of fermentation. For every mol of fermented glucose, the mixed acid route yields 4 mol of acidic products (mostly lactic and acetic acid), 1 mol of neutral fermentation products (ethanol), 1 mol of CO_2, and 1 mol of H_2. On the other hand, only one mol of acid is produced for every mol of glucose when sugars are fermented via the butanediol fermentation route by high-ratio organisms (those that create a high ratio of CO_2 to H_2). The culture media becomes less acidic as a result of this route. The culture has a positive result for the MR test when the culture media turns red upon the addition of

methyl red, owing to a pH at or below 4.4 from the fermentation of glucose. Through the butanediol pathway, bacteria ferment carbohydrates to create acetoin, an intermediate that can then be further reduced to 2,3-butanediol. Alpha-naphthol appears as a catalyst for the oxidation of acetoin to diacetyl. Diacetyl subsequently forms a pinkish-red effect, which is identified as a positive VP reaction, by reacting with the guanidine group linked to molecules added by peptone in the medium. The results of the MR-VP tests thus have an inverse connection.

4.2.1.1.4 Molecular Characterization

Following Chen and Kuo's 1993 protocol, genomic DNA extraction was applied to the bacterial strain that demonstrated the characteristics of *Shigella (*Chen and Kuo, 1993). An individual colony of the bacterial culture was incubated overnight at 37°C in sterile nutritional broth. The bacterial culture that had grown over night was used to extract genomic DNA, and utilizing this extracted DNA, universal primers were employed to perform PCR. The procedures listed below were used to isolate genomic DNA:

i. A 1.5 ml culture was centrifuged for 10 minutes at 10,000 rpm.
ii. The pellet was suspended in 500µL of pH 8.0 TE buffer, which also contained 1 mM EDTA and 10 mM TrisHCl.
iii. Lysozyme was added to the mixture at a concentration of 1 mg/mL-1, and it was incubated for 1 hour at 37°C. RNaseH was added to the mixture at a concentration of 20µg mL-1 in order to eliminate the RNA contamination.
iv. The mixture was mixed with 1% SDS, and it was incubated for 15 minutes at 55°C.
v. After adding an equal volume of Tris saturated phenol, the mixture was thoroughly mixed by inverting the tube.
vi. The mixture was centrifuged for ten minutes at 10,000 rpm in order to perform phase separation.

vii. After gathering the aqueous phase in a brand-new centrifuge tube, an equal volume of a 1:1 mixture of Tris-saturated phenol and chloroform was added.

viii. After the aqueous phase was collected and separated as previously, an equivalent volume of chloroform was added to eliminate any phenol that may have been present in trace amounts.

ix. The aqueous phase was extracted once more and gathered in a brand-new tube.

x. To aid in the nucleic acid precipitation process, 2.5 volume of cold absolute ethanol and 0.3M sodium acetate were added.

xi. Following an hour of incubation at -20°C, centrifugation was carried out for 20 minutes at 15,000 rpm.

xii. After twice washing the pellet in 70% ethanol, it was allowed to air dry.

xiii. Ultimately, 30µL of 2.5mM Tris HCl (pH 8.0) was used to breakdown the genomic DNA.

The isolated microbial DNA was selectively amplified for the16SrRNA gene, using the two universal primers27f [5'AGAGTTTGATCCTGGCTCAG3'] and 1492r [5'TACGGTTACCTTGTTACGACTT3']. To create the master mix, the following estimations were performed.

Table 4.2.1.1 Composition of the master mix used in PCR amplification.

Ingredients	**Volume**
10X PCR reaction buffer (MgCl2 present)	2.5 µl
DNA	1 µl
dNTP (2.5 Mm)	2 µl
27f Primer	1 µl
1492r Primer	1 µl
Taq DNA polymerase (5U)	0.5 µl
Sterile water	X µl

The figure below outlines the necessary temperature parameters for PCR amplification.

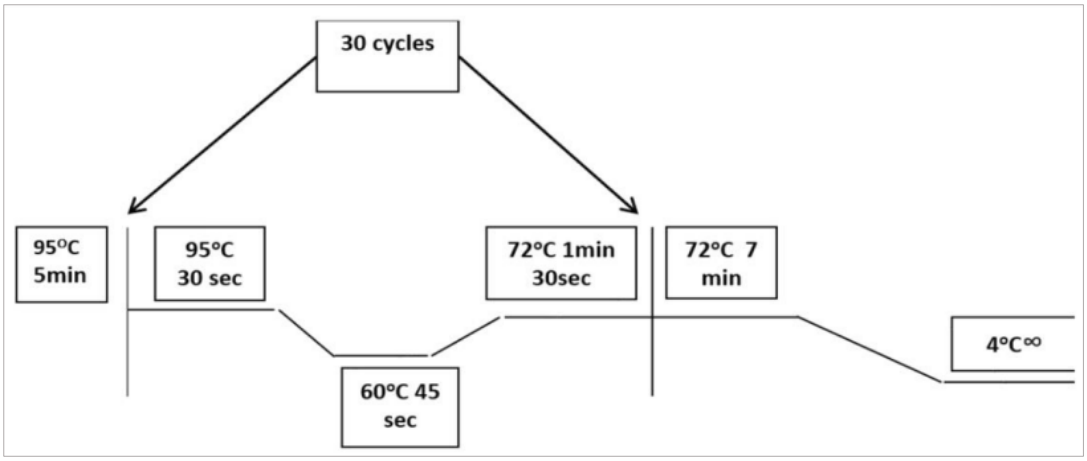


Figure 4.2.1.1 Cycle of PCR amplification.

The following protocol was executed to purify the PCR product using agarose gel extraction.

i. 1% agarose gel was casted with the PCR output in it.

ii. The gel was examined with a UV transilluminator. Utilizing a knife, the bands were cut and transferred into sterile eppendorf tubes.

iii. The weight was determined and the extraction kit QIAGEN Gel method was subsequently followed.

iv. Three volumes of buffer QG were added to the gel and incubated for 10 mins at 50°C.

v. Next, 1 volume of isopropanol was added and mixed thoroughly.

vi. The material was centrifuged right after inserting it into the spin column. The flow through was discarded.

vii. Next, 500 µl of buffer QG was added and centrifuged.

viii. Finally, 750 µl of buffer PE was added to wash and centrifuge, following which it was airdried for a minute.

ix. 20 µl of deionized water was added for elution, and centrifuged further.

The following steps were performed for sequencing of the purified PCR products.

i) Two microliters of 125 millimolar EDTA (pH 8.0) was added to the purified PCR products.

ii) For twenty minutes at room temperature, the resulting suspension was precipitated using 2 µl of 3 M sodium acetate (pH 4.6) and 50 µl of absolute ethanol.

iii) Centrifugation was performed to recover DNA, which was then cleaned with 70% alcohol, dried, and resuspended in 12µl of Hi-Di formamide (Applied Biosystems, USA).

iv) The DNA was denatured for five minutes at 96°C following a 20-minute dark incubation period.

v) Samples were stored at 4°C before being loaded.

vi) The Applied Biosystems 3130XL Genetic Analyzer was used for the sequencing process.

The extracted genomic DNA was shipped to Molsys Pvt. Ltd. (Bengaluru, India) to perform whole-genome sequencing (WGS). The reads thus obtained upon sequencing were quality verified and refined through the FASTQC software, and the cleaned reads were compiled utilising the unicycler algorithm available on the KBase data analysis medium, aided by the template genome NC-004337. Accession CP123365 from the National Center for Biotechnology Information (NCBI) allows the viewing of the entire genome sequence information of the isolated *Shigella flexneri* (SFMMGSG_23).

4.2.1.1.5 Whole Genome Analysis

Prokaryotic Genome Annotation Pipeline (PGAP) from the NCBI and the Prokka tool V.1.1.0 were utilized to annotate the whole genome (Seemann, 2014; Tatusova *et al*., 2016). By using the Resistance Gene Identifier (RGI) tool to search the Comprehensive Antibiotic Resistance Database (CARD), it was possible to identify the likely antibiotic resistance

genes contained in the isolated strain (Alcock *et al.*, 2020). The CRISPRCasFinder tool (Couvin *et al.*, 2018) was adopted to identify the potential Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) arrays and the Cas proteins that are linked to them. The Proksee-Genome Annotation system was deployed to create a premium engaging visualization of the entire cyclic genome. (Grant *et al*., 2023). Using the Proksee server's ORF function, the open reading frames that extend over various stretches of the genome, and coding segments were anticipated. The web-based program Alien-Hunter was used to infer the possibility of horizontal gene transfer activities inside the microbial genome (da Silva Filho *et al*., 2018). By taking into account the abundance and probable arrangement of varying order motifs ranging between one to eight pairs of bases, this method identifies unusual genomic areas based on various compositional metrics (codon/amino acid prejudices, dinucleotide regularity, GC content, etc.) (Vernikos and Parkhill, 2006). It was possible to figure out the existence of transferable elements within the microbial genome through the use of the web-based application of mobileOG database (Brown *et al*., 2022). All of the data has been obtained by supplying the complete genome record in genbank layout and leaving all of the factors at their standard configurations. After the isolated pathogen's draft genome was fed to antiSMASH, the quantity and kinds of biosynthetic gene clusters (BGCs) encoding secondary metabolites were determined (Blin et al., 2023).

4.2.1.2 Pangenome analysis

The isolated strain and 15 other *Shigella flexneri* strains—which typically originated from geographic regions that exhibit marked susceptibility to shigellosis—were contrasted genetically through a pangenome analysis to perceive the adaptive capability and genetic diversification of *Shigella flexneri*. The selection of *Shigella* was motivated by the pathogens' unchecked spread and rising incidence (700,000 cases worldwide each year), with low- and middle-income countries accounting for 99% of cases (Puzari *et al*., 2018).

Bacterial Pan Genome Analysis (BPGA), a perl-based program, was used to accomplish pangenome analysis contingent upon sorting of orthologous protein sequences (Chaudhari *et al*., 2016).

4.2.1.2.1 Strain Selection

By mining the NCBI Assembly database, 15 distinct *Shigella flexneri* variants were selected. These variations were apparently collected from seven exclusive Southeast Asian nations: Singapore, South Korea, China, Hong Kong, Bangladesh, Taiwan and India (Kitts *et al*., 2016). The whole genome data of the chosen isolates were acquired from the Genbank archive. Pangenome analysis was performed on a total of 16 strains—1 isolated and 15 chosen (Table 4.2.1.2). Typically, the strains were selected from regions that showed an elevated level of shigellosis vulnerability.

Table 4.2.1.2 Selected *Shigella flexneri* Genomes from Asian Nations for Pangenome Analysis.

Serial no	Assembly Id	Date of submission	Collection date	Region of isolation	Species name	Strain Name
1	ASM692v2	8/3/2011	1984	China	*Shigella flexneri 2a*	str. 301
2	ASM2235470v1	2/22/2022	2018	Hong Kong	*Shigella flexneri*	Strain: SWHIN_101
3	ASM719759v1	7/20/2019	2019	South Korea	*Shigella flexneri*	Strain: C32
4	ASM217469v2	3/26/2019	2015	India	*Shigella flexneri*	Strain: FC906
5	ASM2286984v1	4/11/2022	2000	Taiwan	*Shigella flexneri*	Strain: 3160_NCHU22
6	ASM417118v1	2/12/2019	1984	Bangladesh	*Shigella flexneri*	Strain: SFL1520
7	ASM2473228v1	8/22/2022	2022	Singapore	*Shigella flexneri*	Strain: E
8	ASM158017v1	3/3/2016	1998	China	*Shigella flexneri 2a*	Strain: 981
9	ASM157812v1	3/1/2016	2002	China	*Shigella flexneri 1a*	Strain: 0228
10	ASM1358v1	7/21/2006		China	*Shigella flexneri 5*	Strain: 8401

11	ASM157996v1	3/3/2016	2012	China	*Shigella flexneri 4c*	Strain: 1205
12	ASM2249435v1	3/7/2022	2018	Hong Kong	*Shigella flexneri*	Strain: STLIN_17
13	ASM2473230v1	8/22/2022	4/13/2022	Singapore	*Shigella flexneri*	Strain: D
14	ASM1979357v1	8/29/2021	10/4/2020	China	*Shigella flexneri*	Strain: WW1
15	ASM244299v2	8/20/2019	1988	Bangladesh	*Shigella flexneri 1c*	Strain: Y394

4.2.1.2.2 Clustering of Orthologous Protein Sequences

All protein sequences were clustered by implementing the USEARCH algorithm, obeying the 50% sequence identity cut off (default setting). To eliminate bias, BPGA allows pangenome calculation by sequentially adding genomes in 30 different combinations, and it calculates pan and core genome dimensions based on the average measures for every instance (Chaudhari *et al*., 2016).

4.2.1.2.3 Functional Classification of Pan/Core Genes

The Pangenome Functional Analysis Application of the BPGA software were utilised to annotate all genes in order to comprehend the connection between the regularity of each gene and its corresponding function. Based on protein BLAST against reference COG and KEGG databases, COG and KEGG identifiers are allocated to all sampled protein sequences belonging to every orthologous protein cluster. Next, for central genes, auxiliary genes, and exclusive genes, the relative proportions of COG and KEGG classifications are computed, which are illustrated as histograms using gnuplot.

4.2.1.2.4 Phylogenetic Deduction using Pan/Core Gene Alignments

The same BPGA platform was adopted for building a phylogenetic tree without the support of any reference genome, built on 20 arbitrarily picked groups of common genes amongst the 16 variants. The evolutionary relationship between the 16 *Shigella flexneri* strains was

ascertained using a thorough phylogenetic reconstruction dependent on concatenated core gene agreements and the absence or presence of pan genes.

4.2.1.2.5 Pathway Enrichment Analysis of Core Genes

Using BLAST, the bi-directional best hit technique, and the KEGG Automated Annotation Server (KAAS) with an established cutoff bit-score limit of 60, the collection of genes that compose the core genome was mapped to conventional biological routes (Moriya *et al*., 2007). KAAS allocates KO identifiers to single protein sequences and permits pathway restoration by coupling a group of orthologous genes to an element in the KEGG networks and BRITE functional hierarchy.

4.2.1.3 Annotation of hypothetical proteins

A comprehensive evaluation of the genomes of *Shigella flexneri* revealed that roughly 18% of the encoded proteins had unidentified biological and cellular roles (Jin *et al*., 2002). It was envisaged that functional annotation of these hypothetical proteins (HPs) would provide pertinent details on how they operate in immunological responses, other metabolic pathways, and pathogenicity (Naveed *et al*., 2018). In this study, 432 hypothetical proteins from *Shigella flexneri* were investigated and their functions classified using a computational pipeline that involved several bioinformatics databases and computational tools, which allowed us to gain a better understanding of the functional characteristics of these unidentified proteins, some of which could be novel therapeutic targets, thereby laying the groundwork for the development of new treatments and appropriate medication formulations.

4.2.1.3.1 Curation of Hypothetical Proteins

A text-based search was undertaken to extract complete sequences of possible *Shigella flexneri* proteins from the NCBI Protein database, excluding putative and partial sequences

(NCBI Resource Coordinators, 2016). After the search yielded 717 hits, the proteins with 200 amino acids or more were further filtered out. Less than 200 amino acid proteins were eliminated since they have not been demonstrated to be stable therapeutic targets (Bakheet and Doig, 2010; Oany *et al*., 2018). The other 432 HPs' sequences were downloaded in FASTA format from the NCBI data warehouse's Protein database, along with the corresponding accession number (Yousafi *et al*., 2019).

4.2.1.3.2 Functional Analysis of Hypothetical Proteins

Understanding a protein's subcellular compartmentalization is essential for effectively characterizing its activity. CELLO and PSORTb, two different programs, were used to predict the sub-cellular localization of these unidentified bacterial proteins (Yu *et al*., 2006; Yu *et al*., 2010). To accurately infer protein localizations, CELLO employs a two-level support vector machine (SVM) algorithm. The first level contains a small number of SVM classifiers, each of which pertains to a distinct feature taken from the protein sequence itself, such as amino acid composition, di-peptide composition, sequence profiling based on the physico-chemical properties of the amino acids, etc. As a juror, the next level SVM classifier projects the probability value for each localization category by averaging the scores from various first level classifiers (Yu *et al*., 2006). Domains are 3-dimensional functional elements that exist independently and are typically conserved throughout protein families. Finding the conserved domains that are integrated into these unidentified proteins can provide us important hints about how they function. This was accomplished utilizing computer programs such as BLASTp and InterPro, which work on the tenet that proteins exhibiting homology in several species are functionally related (Johnson *et al*., 2008; Mitchell *et al*., 2015; Finn *et al*., 2017). Searches for proteins with known activities and significant sequence similarity with the query protein were conducted using a non-redundant protein sequence database. Next, information about domains and families was

determined based on the best hits. Determining the function of these query proteins in biological processes requires a thorough examination of their protein-protein interactions. The STRING database was utilized to anticipate the potential interaction partners of the HPs. This database integrates data from a wide range of organisms to reveal both direct and indirect contacts of the HPs (Franceschini *et al*., 2013).

4.2.2 Construction and Validation of 3-Dimensional Structures

The putative target sequences with functional relevance from both core genome as well as hypothetical proteins thus proposed were subjected to comparative modelling techniques to construct 3-Dimensional protein structures. A combination of template free modelling together with fold recognition-based threading techniques was employed to generate the 3-D models of the selected proteins under study using the online server IntFold (McGuffin *et al*., 2019). The server presents five top scoring models in the descending order of their global model quality score and confidence level, for each input sequence. Global model quality scores are in the range of 0-1, with higher scores indicating more complete and confident models. The models thus generated were next checked for their stereochemical stability by looking at the Ramachandran distribution and verification of Z-score using QMEAN scoring function (Williams *et al*., 2018; Benkert *et al*., 2011). The conformationally stable proteins were retained for subsequent analysis.

4.3 Structure Based Drug Discovery

Natural products, especially plant derived metabolites represent a prolific fount of bioactive ingredients, having thousand years long history of clinical applications in different ethnomedicinal approaches. The bountiful scaffold diversity and evolutionary optimized biological effects of herbal products, makes them a lucrative choice from conventional synthetic compounds (Erb and Kliebenstein, 2020). Although plants remain an

unquestionable reservoir of bioactive principles, the molecular validation and marketing of herbal drugs remain at a distance, owing to the resurgent challenges of accessing, screening and optimization (Harvey *et al*., 2015). Through this work, we intend to isolate, characterize and optimize such active pharmaceutical ingredients from herbal extracts using advanced analytical tools, with the aim to test their efficacy against the diarrheal disease, shigellosis, which will reinvigorate the international interest in phytochemicals as alternative medicines and promote their industrialization. Here, we have utilized high throughput virtual screening in conjunction with the analytical chemistry platform of mass spectrometry (MS) hyphenated with gas chromatography (GC) to efficiently identify and capture the therapeutically promising small molecules found in herbal extracts that have been reported to have antidiarrheal properties in ethnomedicinal literature.

4.3.1 Curation and identification of antidiarrheal plants

Plants are distinctive and competitive in their own habitats because they produce a vast array of organic molecules, which are frequently categorized into overlapping groups of main and secondary metabolites depending on their roles (Erb and Kliebenstein, 2020). While a number of antiquated medical texts highlight the historical importance of plant-based natural products in the treatment of various illnesses, the majority of the information comes from countless human trial and error experiments conducted over hundreds of years with little to no understanding of the underlying bioactive constituents (Dzobo, 2022). Plants used in traditional ethnomedicinal approaches against diarrheal diseases were identified through text-based search of IMPPAT (https://cb.imsc.res.in/imppat/) and TKDL (http://www.tkdl.res.in/tkdl/langdefault/common/Home.asp?GL=Eng) database, which houses a list of medicinal plants mentioned in Ayurveda, Unani, Siddha and by mining the texts of Ayurvedic Pharmacopoeia of India (API) The selected plants were then collected from Medicinal Plant Gardens maintained by the Government of West Bengal.

4.3.2 Preparation of herbal extracts and spectrometric analysis

The plant leaves were gathered in sterile zip-lock bags with the appropriate labelling, and the bags were then transported to the lab where the leaves were air dried and cleaned to get rid of any remaining dust and debris. The dry leaves were then ground in a mechanical blender until they were powdered, after which the powder was smashed with methanol in a mortar and pestle. Whatman filter paper was used to filter the resultant crude extract, and the filtrate was used for all subsequent analysis. Drug development is accelerated by the use of metabolomics in conjunction with bioinformatic evaluation, making it an invaluable device for qualitative study of the diverse range of metabolites found in plant extracts (Salem *et al*., 2020). In order to build a virtual library of natural products, GC-HRMS based metabolomic profiling of the chosen herbal extracts were performed at the Sophisticated Analytical Instrument Facility (SAIF) at the Indian Institute of Technology, Bombay, using a Jeol AccuTOF GCv mass spectrometer with a direct capillary column HP5 Capillary Column (30m length × 0.25 microfilm thickness × 0.25mm).

4.3.3 Identification of Hits and Leads

The library of small molecules thus assembled were screened using drug likeness and ADMET filters to identify suitable hits (Lipinski *et al*., 2001; Guan *et al*., 2018). Structure based virtual screening was undertaken utilizing molecular docking to rationalize the activity of the promising hits towards the set of specific target candidates identified thus far. Molecular docking is a widely recognized virtual screening technique, especially when the three-dimensional structure of the target protein has been determined (Lavecchia and Di Giovanni, 2013). It makes it possible to anticipate the structure of the ensuing protein-ligand complex as well as the binding affinity between a ligand and a protein, which serves as an insightful data when optimizing leads during drug discovery. Understanding the ways

in which drugs interact with bacterial targets becomes easier with the use of docking, which also makes molecular structures visually appealing and analytically practicable. Utilized for more than thirty years in pharmaceutical research, molecular docking has been invaluable in our efforts to find new target-ligand combinations that are effective against *Shigella* species.

4.3.3.1 Virtual Screening of Small Molecules

Following a metabolomic investigation, chemicals from the chosen plants were evaluated for their drug-likeness according to structural or physicochemical parameters that have a major bearing on the molecules' pharmacokinetic and pharmacodynamic attributes (Guan *et al*., 2018). Databases such as PubChem and ZINC were used to get the three-dimensional structure of the recognized chemical entities or specification in the form of a line notation (SMILES) expressing their structure using short ASCII characters. The structural coordinates of a large variety of compounds that can be utilized for virtual screening are easily accessible through these sources available online. Pharmacological screening was accomplished using the two web-based servers, SwissADME (http://www.swissadme.ch/) and Molinspiration (https://molinspiration.com/) respectively (Daina *et al*., 2017; Khan *et al*., 2019). The canonical SMILES of the small compounds were used as inputs.

4.3.3.2 Docking Interaction Analysis

The structure of a protein target-small molecule interaction was predicted using molecular docking. Docking programs are essentially a hybrid of a scoring function and a search algorithm. The goal of the search technique was to locate the exact 3D geometry, or poses, of a ligand within a certain targeted protein. Predicting the binding affinity was the goal of the scoring function, which would evaluate how effectively ligands bind to proteins. Using the SeamDock website, an *in-silico* evaluation of protein receptor interactions with small

molecules was successfully done (Murail *et al*., 2021). In order to facilitate simple interaction site identification, the SeamDock web service unifies many docking technologies (AutoDock, AutoDock Vina, Smina, or Qvina) into a single framework that supports both global and local ligand docking as well as a structured approach of combining the two (Murail *et al*., 2021). The prospective targets were thereafter screened against the identified hits for their optimization to leads and subsequent in vitro confirmatory assays.

4.3.4 In-vitro assays for efficacy testing of herbal extracts/leads against *Shigella*

In order to draw a comprehensive inference on the effectiveness of the selected potential leads and their corresponding target candidates, the application of computational methods was followed by an array of enriched in-vitro assays. To address microbial resistance, there has been a boom in enthusiasm for the research and creation of natural antibacterial agents derived from plant sources in recent years. As a result, techniques for screening and assessing antimicrobial activity have received greater emphasis. In addition to the well-known and frequently used conventional bioassays like disk-diffusion, well diffusion, and broth or agar dilution protocols, other techniques like flow cytofluorometric and bioluminescent methods can yield quick results pertaining to the effects of the antimicrobial agent and a deeper comprehension of their influence on the viability and cell damage inflicted upon the microorganism under investigation (Balouiri *et al*., 2016). In this work, a combination of disc diffusion assay and bacterial live/dead assay using Flow cytofluorometric method was deployed confirm the antishigellosis effect of the projected leads. Furthermore, morphological changes in the isolated organism and their colonies exposed to treatment with specific leads were also observed utilizing a scanning electron microscope (SEM).

4.3.4.1 Antimicrobial Susceptibility Testing Using Disc Diffusion Assay

To assess the isolated strain's susceptibility and resistance to various antibiotic classes, herbal extracts and active ingredients (leads), the Kirby-Bauer disc diffusion test was carried out in accordance with CLSI standards (33rd edition) (CLSI Performance Standards for Antimicrobial resistance Testing 2023). Nutrient Broth (NB) was used to cultivate the bacterial isolate for an entire night. 200**µ**l of the enriched culture were then put into a nephelometric flask with 20ml of nutrient broth and incubated at 37°C with shaking after 18 hours. On an hourly basis, the optical density (O.D.) of the nephelometric flask was recorded at 600 nm. Fresh nutrient agar (NA) dishes were seeded with 200 **µ**l of the microbial suspension following plate spreading tactics. Antibiotic discs were promptly placed, one into each quadrant, right after when the O.D. value attained 0.5 units. After a 24-hour incubation period at 37°C, zone diameter readings were taken, along with accompanying pictures. Initially the Kirby Bauer disc diffusion assays were performed for standardizing the activity of the crude extracts with the commercially available antibiotics kept as positive control. Based on their effectiveness in restricting bacterial proliferation, the top two performing crude extracts were identified and corelated with the data obtained from virtual screening and docking interaction to finalize upon the active pharmaceutical ingredient, the lead. The isolated bacterium was tested for its vulnerability to antimicrobial agents from various categories and genera., viz. ampicillin- 25mcg (i), tetracycline- 30mcg (ii), norfloxacin- 10mcg (iii), amoxycillin- 30mcg (iv), gentamicin- 10mcg (v), amikacin- 10mcg (vi), cefoxitin- 30mcg (vii), imipenem- 10mcg (viii), cefuroxime-30mcg (ix), trimeth-sulfa- 25mcg (x), ciprofloxacin-5mcg (xi).

4.3.4.2 Bacterial Cell Viability Assay Using Flow Cytometry

The findings of the Kirby-Bauer disc diffusion assay, were further justified by performing bacterial cell viability assay to comment on the bactericidal effect of the crude extracts and chosen chemical leads. The merits of flow cytometry for assessing the susceptibility/resistance phenotype of microorganisms were proposed a number of years ago. As a result, many authors adopted this methodology to study the antibacterial and antifungal properties of numerous medications (Davey and Guyot, 2020). This method's quick identification of injured cells is dependent on the legitimate staining of proper dyes. Propidium iodide (PI), a fluorescent and intercalating chemical, was hence frequently employed as a DNA stain. However, using a single stain makes it a challenging task to accurately discriminate dead cells from other cellular debris. Therefore, to circumvent this problem, a dual staining live dead assay was performed in this work to precisely record the antibacterial activity of the crude extracts and purified chemical leads. The principal stain, propidium iodide, was employed since it only seeps into cells with damaged membranes. As a counterstain, thiazole orange was employed; it penetrates all cells, living and dead, to differing degrees. As a result, combining these two dyes offered a quick and accurate way to distinguish between living and dead bacteria. The bacterial cells were stained using 0.5 ug/ml of PI. Following a 15-minute incubation period, the cells were subjected to a 473-nm laser run on a BD Accuri C6 plus flow cytometer, and the PI signal was captured using a detector that had a 588/15-nm filter.

4.3.4.3 Tracking Morphological Changes using Scanning Electron Microscopy

The surface structure of *Shigella* samples, in both treated and untreated cultures, was examined using scanning electron microscopy (SEM) (Czerwińska-Główka and Krukiewicz, 2021). Glass slides were utilized as the sample preparation platform in order

to minimize interactions with bacterial metabolites. The bacterial pellets were separated from the suspensions by centrifuging them for 10 minutes at 2000 rpm. Following the removal of the supernatant, the pellets were centrifuged again under the same settings after being washed in 1 ml of distilled water. 5 to 10 µl of the condensed bacterial suspension was placed on the surface of a clean glass slide. The sample was fixed, dehydrated, and dried in the order specified by the standard SEM sample preparation protocol. Fixation was typically done using an aldehyde buffer, and osmium tetroxide (OsO_4) was used as a post-fixation step. Specimens were then typically sprayed with a conductive wrapping to create a conducive investigation surface and avoid electron charging effect-induced image distortion. The prepared sample was then observed under a scanning electron microscope. The resulting images were used to examine the morphology of bacterial cells, their adhesion to the surface, as well as their number and distribution in the area under observation.

4.3.5 Differential Expression Profiling

A crucial component for comprehending phenotypic variation is accurately identifying differentially expressed genes (DEGs) under particular circumstances. Differential expression profiling using standard NGS based transcriptome analysis were performed to evaluate the upregulated and downregulated genes under the effect of the crude extracts in culture (Jiang *et al*., 2015). The profiles were compared to the identified multidrug resistant strain (MDR). The up/downregulated genes were henceforth corelated with the protein targets recognized from core genome to pinpoint the differentially regulated genes bearing strong resemblance to the identified targets and impacting the same biological pathway. The observed up/downregulation of the hand-picked genes were eventually substantiated by performing quantitative PCR assays (Rocha *et al*., 2020).

4.3.5.1 Transcriptome Analysis

Comparing different physiological states or strains can yield valuable insights into cellular processes through the analysis of global gene expression, a feature that makes RNA-Sequencing especially beneficial (Utturkar *et al*., 2020). In this work, the RNA-Sequencing technology using the Illumina (NOVASEQ 6000) platform was employed for differential expression analysis under specific treatment conditions by pooling the samples simultaneusly (Costa-Silva *et al*., 2017). Prior to being sequenced using a high throughput technology, the RNA samples were first broken up into tiny complementary DNA sequences (cDNA). The sequencing generated raw fastq reads were assessed for quality using FastQC v.0.11.9. Fastp v.0.20.1 was used to preprocess the raw fastq reads, which were then re-assessed for quality using FastQC and summarized using MultiQC. To filter rRNA reads, the trimmed fastp reads were matched against the silva database with bowtie2. Subsequently, the rRNA-filtered reads were aligned with the reference genome of an untreated multidrug-resistant *Shigella flexneri* strain (GCF_030169125.1) using Bowtie2 v2.4.5. Following that, an estimate of each gene's or isoform's expression levels was determined. The gene counts were compared to *Shigella flexneri'*s unique reference genome and utilized as inputs to edgeR with exactTest to estimate differential expression. The differentially expressed genes (DEGs) were eventually identified by normalizing the mapped data and applying statistical and machine learning techniques. Gene expression was statistically estimated using a log2 fold change of approximately 1.0 and an adjusted p-value threshold of less than 0.05. The up and downregulated genes identified thus far, were finally justified from a biological perspective.

4.3.5.2 quantitative PCR

Quantitative PCR (qPCR) performed in tandem with NGS-based transcriptome analysis offers an achievable solution to any inaccurate interpretations that may have been inadvertently introduced and provides a precise estimation of differential gene expression (Rocha *et al*., 2020). Gene expression analysis using the qPCR was executed by performing the following steps in a sequential manner (Duan *et al*., 1997):

i. The total RNA was extracted from the control and treated samples by performing single step RNA isolation using a monophasic solution of phenol and guanidine isothiocyanate (GITC) following the steps as outlines below:
    a) The bacteria were cultured for 3-4 hours following which the culture was centrifuged at 8000rpm for 10 minutes to obtain the bacterial pellet.
    b) 1ml of lysis buffer containing GITC, sodium acetate, phenol and SDS was added to the pellet and mixed using micropipette.
    c) The microcentrifuge tube was then incubated at room temperature for 30 minutes, following which 300 µl of chloroform was added to the tube and mixed well by inverting the tube for 10 seconds.
    d) Next the tube was centrifuged at 12000 rpm for 20 minutes at 4ºC and the aqueous phase thus obtained was transferred to a fresh microcentrifuge tube.
    e) Equal volume of propanol was added to the aqueous phase and incubated at -20 ºC for an hour.
    f) After incubation, the tube was centrifuged again at 12000 rpm for 15 minutes at 4ºC. The supernatant was discarded following centrifugation.
    g) Next 750 µl of ethanol (75%) was added to the tube and centrifuged further at 10000 rpm for 10 minutes at 4ºC.

h) The pellet thus obtained was airdried and dissolved in 50 µl of nuclease free water for further downstream analysis.

ii. The RNA was next checked for its concentration and purity using Multiscan SkyHigh Microplate Spectrophotometer from Thermo Scientific.

iii. The extracted RNA was next synthesized into cDNA and carried forward for qPCR using the Applied Biosystems QuantStudio® Real Time PCR Systems. For a sensitive and reliable detection of RNA expression, the GoTaq® 2-Step RT-qPCR Kit was used. This kit contains the components of GoScriptTM Reverse Transcriptase and GoTaq® qPCR Master Mix. The primers were designed for the up and down regulated genes identified in transcriptome analysis using the Primer3Plus web interface.

iv. The obtained data was normalized using double delta Ct analysis method. Four Ct values—Gene being Tested Experimental (TE), Gene being Tested Control (TC), Housekeeping Gene Experimental (HE), and Housekeeping Gene Control (HC)—were obtained by averaging the Ct values for the housekeeping gene and the gene under test in the experimental and control conditions. The ΔCt values for the experimental (ΔCTE) and control (ΔCTC) conditions, respectively, were obtained by calculating the differences between the housekeeping gene and the tested gene under both experimental (TE – HE) and control (TC – HC) circumstances. The double delta Ct value (ΔΔCt) was then computed by taking the difference between the ΔCT values for the experimental and control circumstances (ΔCTE − ΔCTC). The value of $2^{-\Delta\Delta Ct}$ was calculated to finally deduce the change in gene expression in terms of fold change.

5. Results

5.1 Isolation and Identification of *Shigella flexneri*

From the obtained liquid sludge sample, the study was able to extract and purify a strain of *Shigella flexneri*. After a 24-hour incubation period, the plates that received inoculation with varying gradients of serially diluted bacterial suspensions were examined for bacterial growth. Based on the common colony morphology traits of colour, opacity, texture, etc., 19 unique microbial colonies were selected from the plate containing a countable number of colonies. The selected 19 pure bacterial cultures were then subjected to a battery of distinct biochemical tests depending on their ability to utilize carbohydrates and produce particular enzymes. The procedures outlined in Bergey's Manual of Systematic Bacteriology were followed while conducting the tests (catalase, citrate utilization, oxidase, indole, lactose utilization, MR, and VP) (Brenner *et al*., 2005). The bacterial isolate that showed consistent *Shigella*-positive results (methyl red positive, lactose negative, oxidase negative, catalase positive, citrate negative, and Voges Proskauer negative) was then put forward for 16S rRNA sequencing (Figure 5.1.1).

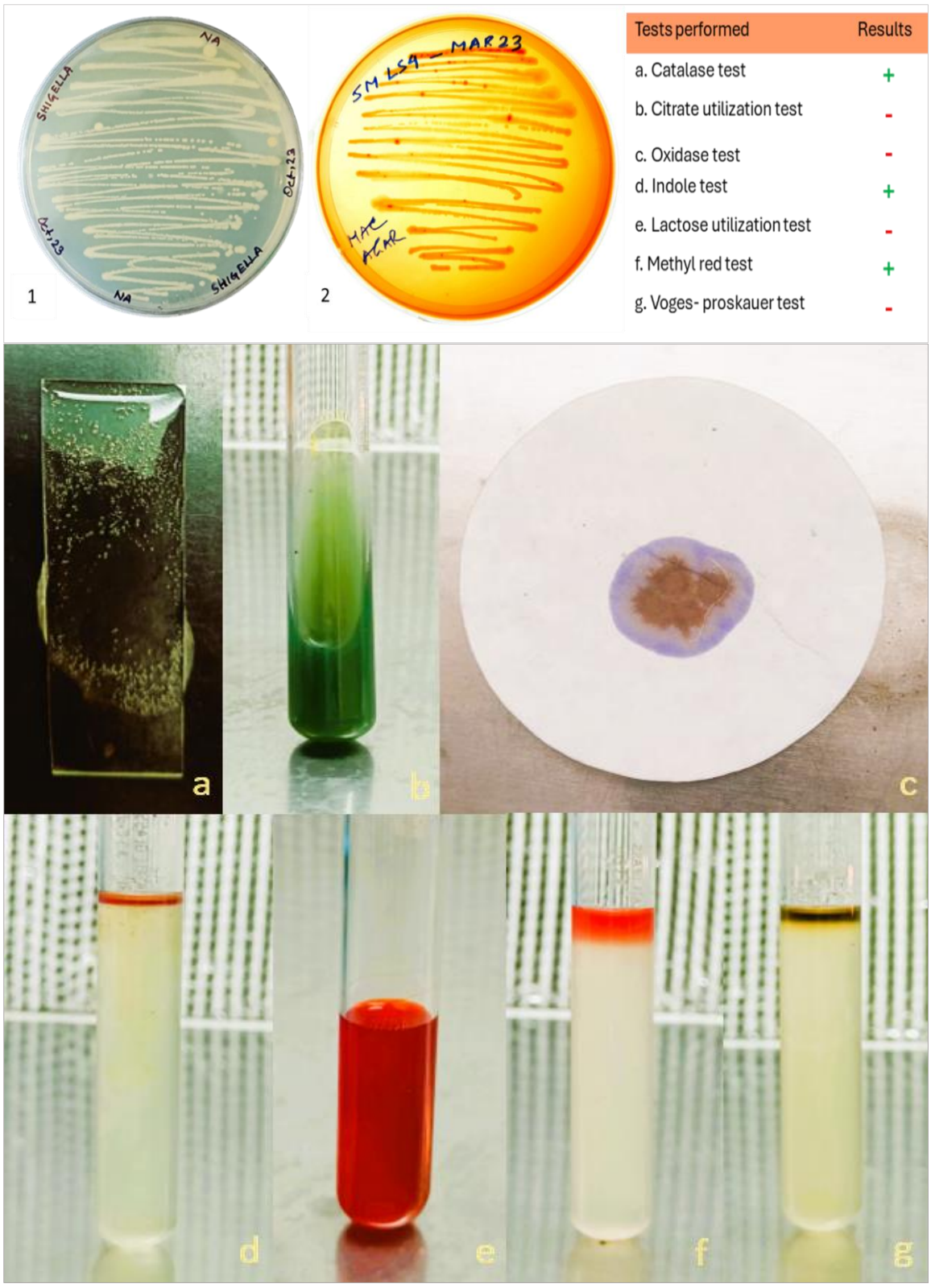


Figure 5.1.1 Colony characteristics of the isolated *Shigella flexneri*, grown on: 1. Nutrient agar 2. McConkey agar plates; Biochemical traits of the naturally isolated *Shigella flexneri*, catalase +ve, citrate utilization -ve, oxidase – ve, indole +ve, lactose utilization -ve, methyl red +ve, Voges Proskauer -ve.

5.2 Molecular Characterization of *Shigella flexneri*

The bacterial isolate was subjected to PCR amplification for the 16SrRNA gene following genomic DNA extraction (Figure 5.2.1; Figure 5.2.2). The amplified PCR products was next sequenced, which further confirmed the microbial isolate as *Shigella flexneri* (Figure 5.2.3).

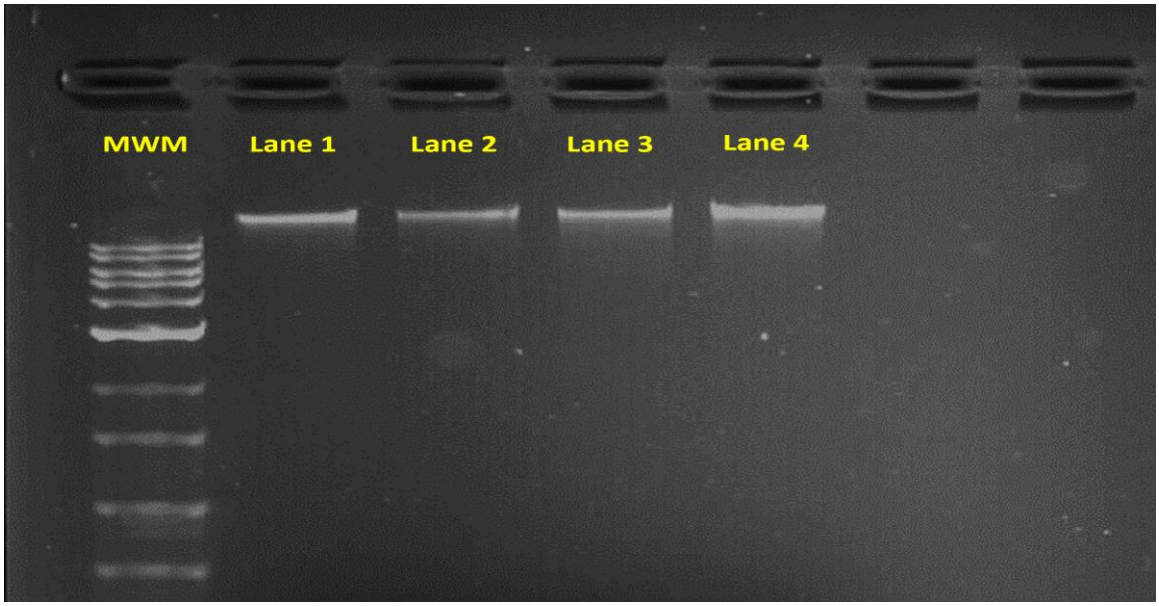


Figure 5.2.1 Agarose gel electrophoresis for visualization of genomic DNA; where MWM indicates molecular weight of marker, and lane 1-4 indicates genomic DNA isolated from *Shigella flexneri*.

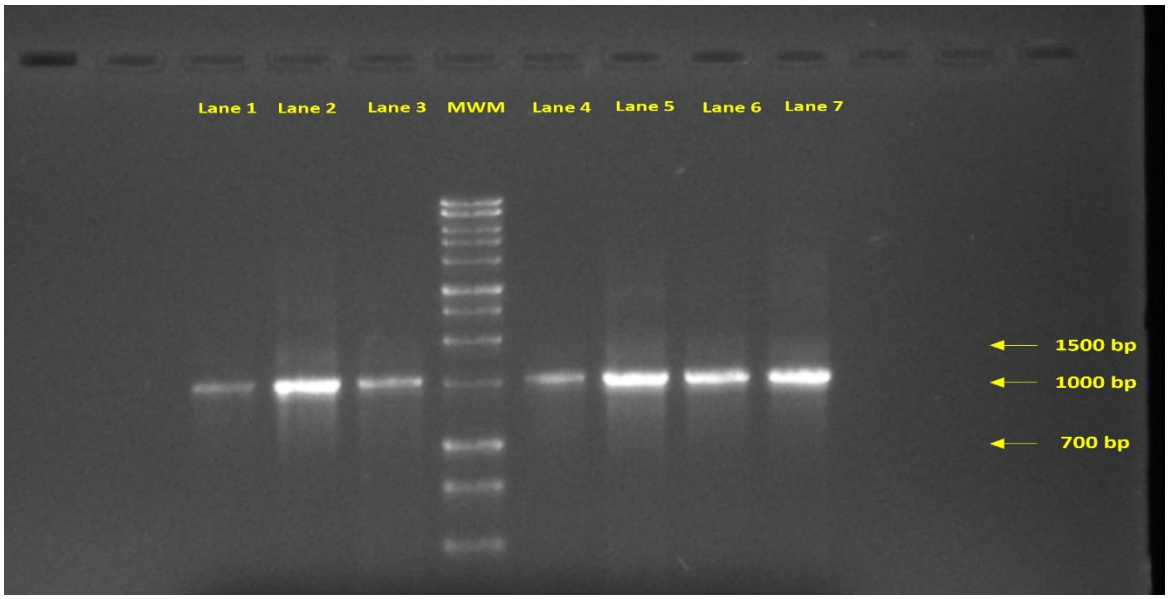


Figure 5.2.2 PCR amplification of 16S rRNA gene; lane 5 indicates amplified 16SrRNA gene for *Shigella flexneri* and MWM indicates molecular weight of marker.

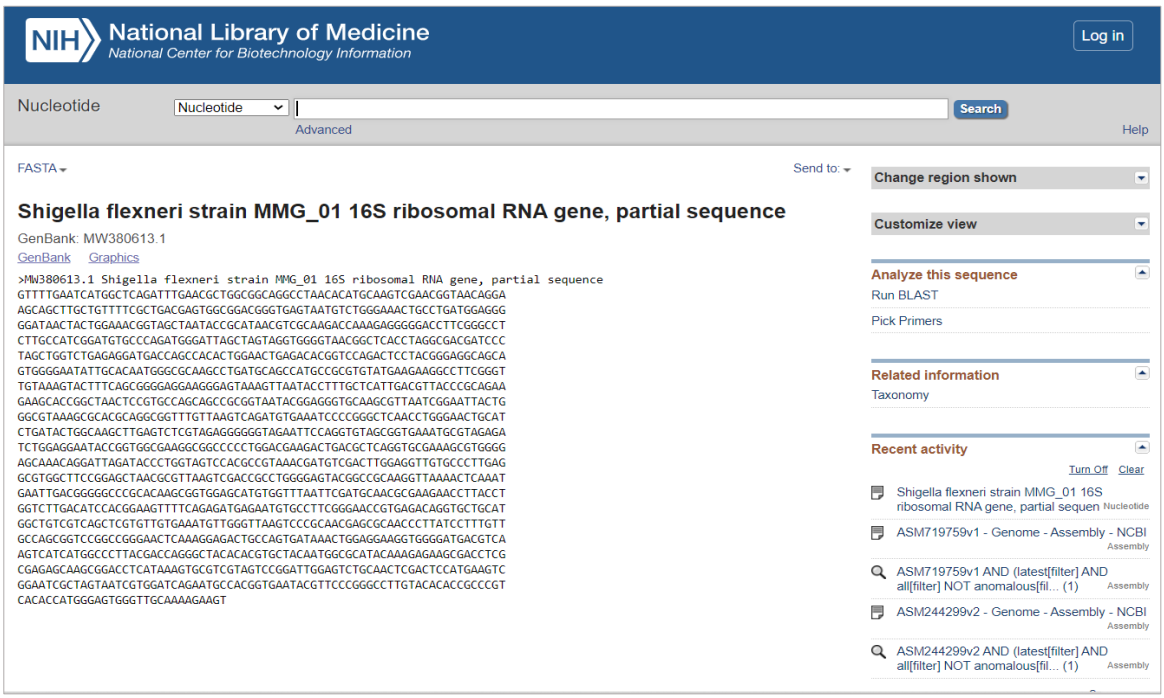


Figure 5.2.3 Published 16SrRNA gene sequence of the isolated *Shigella flexneri* strain in NCBI.

5.3 Whole genome analysis of the isolated *Shigella flexneri*

The entire genome sequence was analyzed using the web-based Proksee service, which uses a variety of algorithms to assess genomic characteristics (Figure 5.3.1). 12000 Mb of unprocessed sequence reads were obtained from whole genome sequencing using the Illumina HiSeq technology. Following quality assurance, 4607202 base pairs of unambiguous reads were put together into a circular genome. The NCBI has incorporated the genome (SFMMGSG_23) under accession number CP123365. The Prokaryotic Genome Annotation Pipeline (PGAP) hosted by NCBI estimated 4640 genes, of which 3845 were found to code for proteins, 127 for specific RNAs, and the balance of 668 for pseudogenes because of their dubious residue composition or propensity to produce truncated peptides.

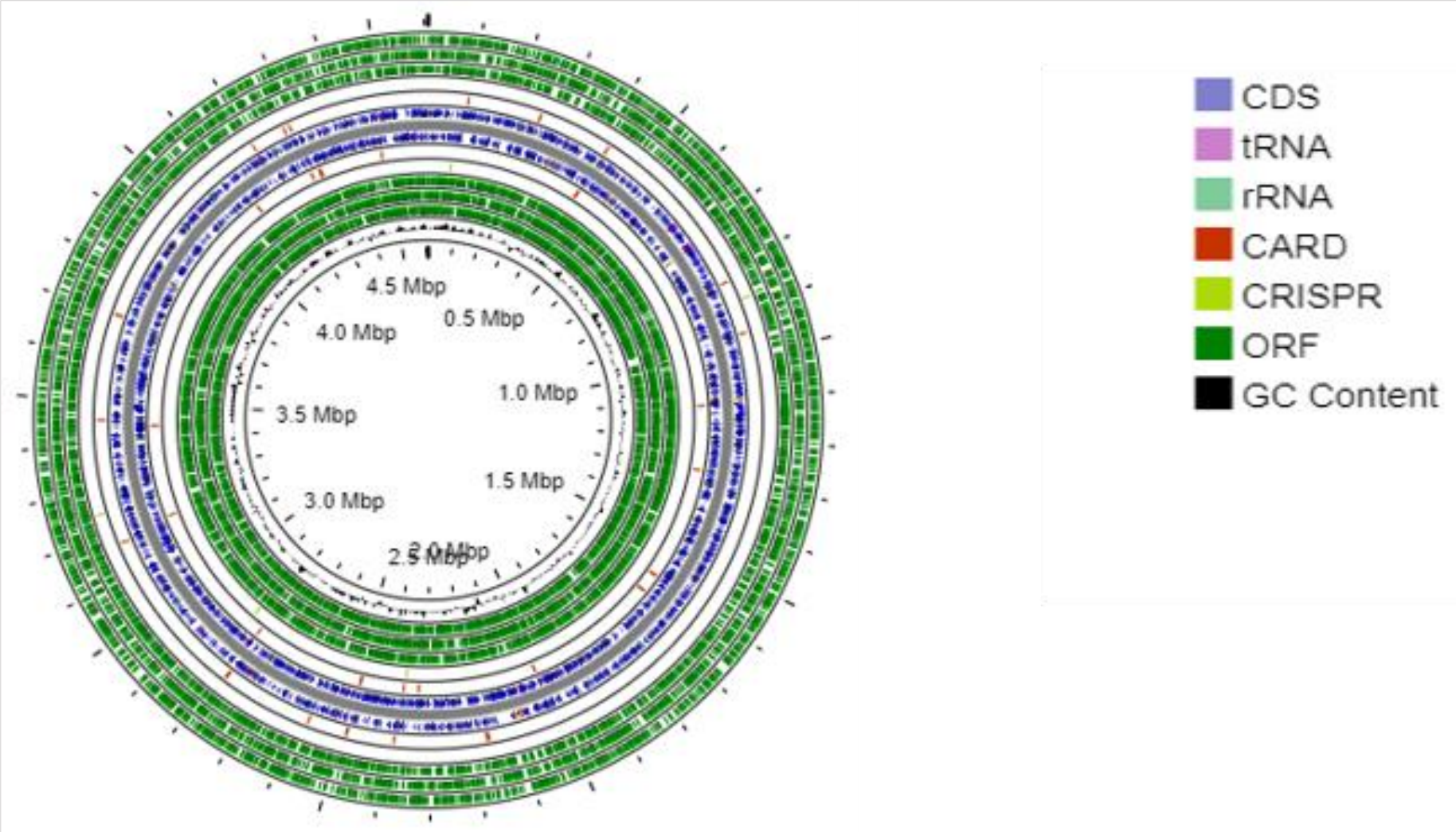


Figure 5.3.1 Circular chromosome map of *Shigella flexneri* strain SFMMGSG_23. A total of 9167 open reading frames (ORFs) are pointed out, with 4834 coding sequences, 57 resistance genes and 5 CRISPR arrays.

A total of 4834 probable coding genes were identified by Prokka, of which 2830 could be classified into different COG categories (Figure 5.3.2). Additionally, 1105 genes were determined to code for hypothetical proteins.

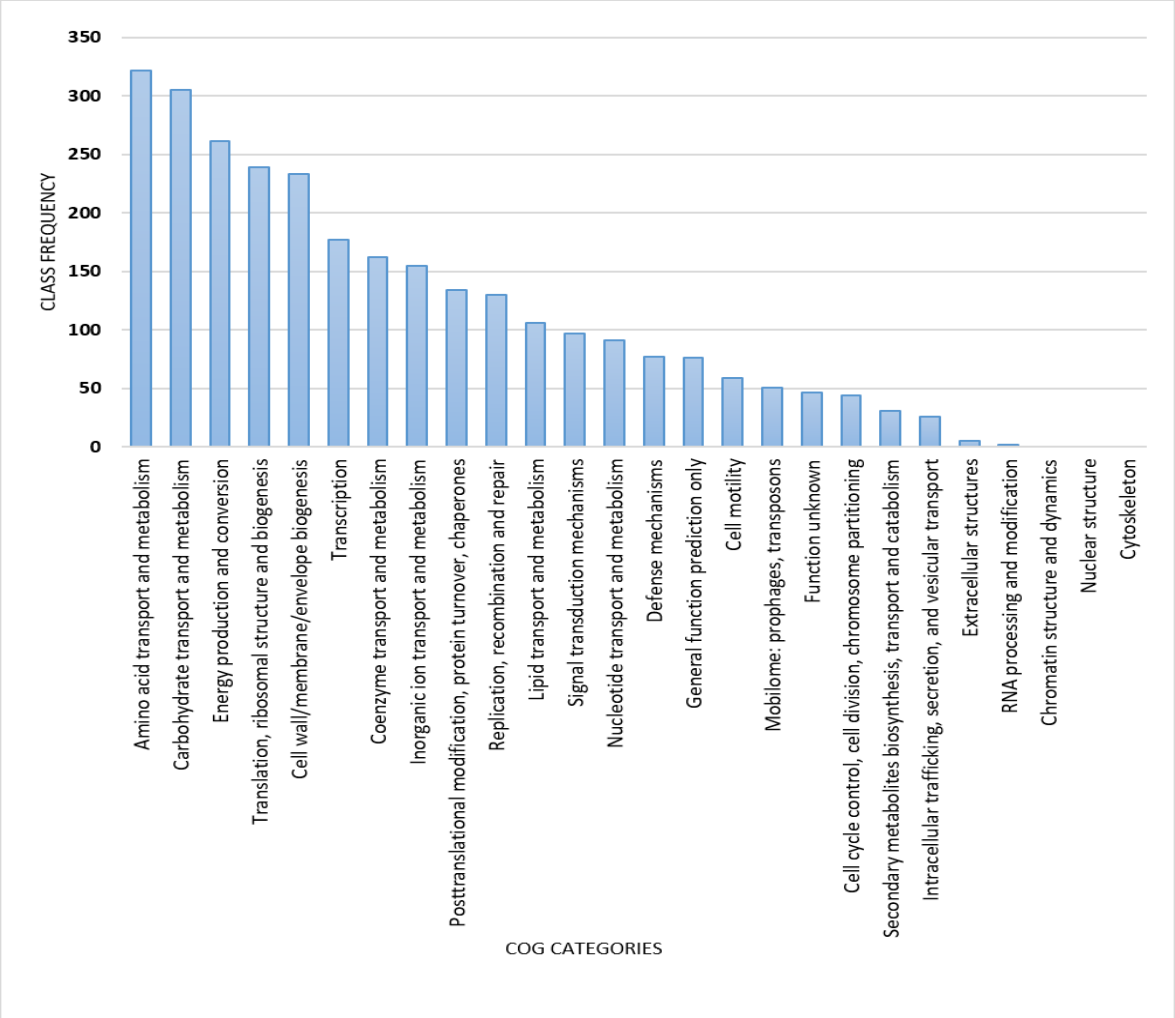


Figure 5.3.2 Distribution of coding genes into different functional COG categories.

Using the RGI tool, a search was conducted against the CARD database which revealed 57 genes associated with antimicrobial resistance that fall into 11 discrete families of genes (Figure 5.3.3). It turned out that over 50% of the genes encoded antibiotic efflux proteins, whereas 15% of the genes coded for proteins that participate in target transformation.

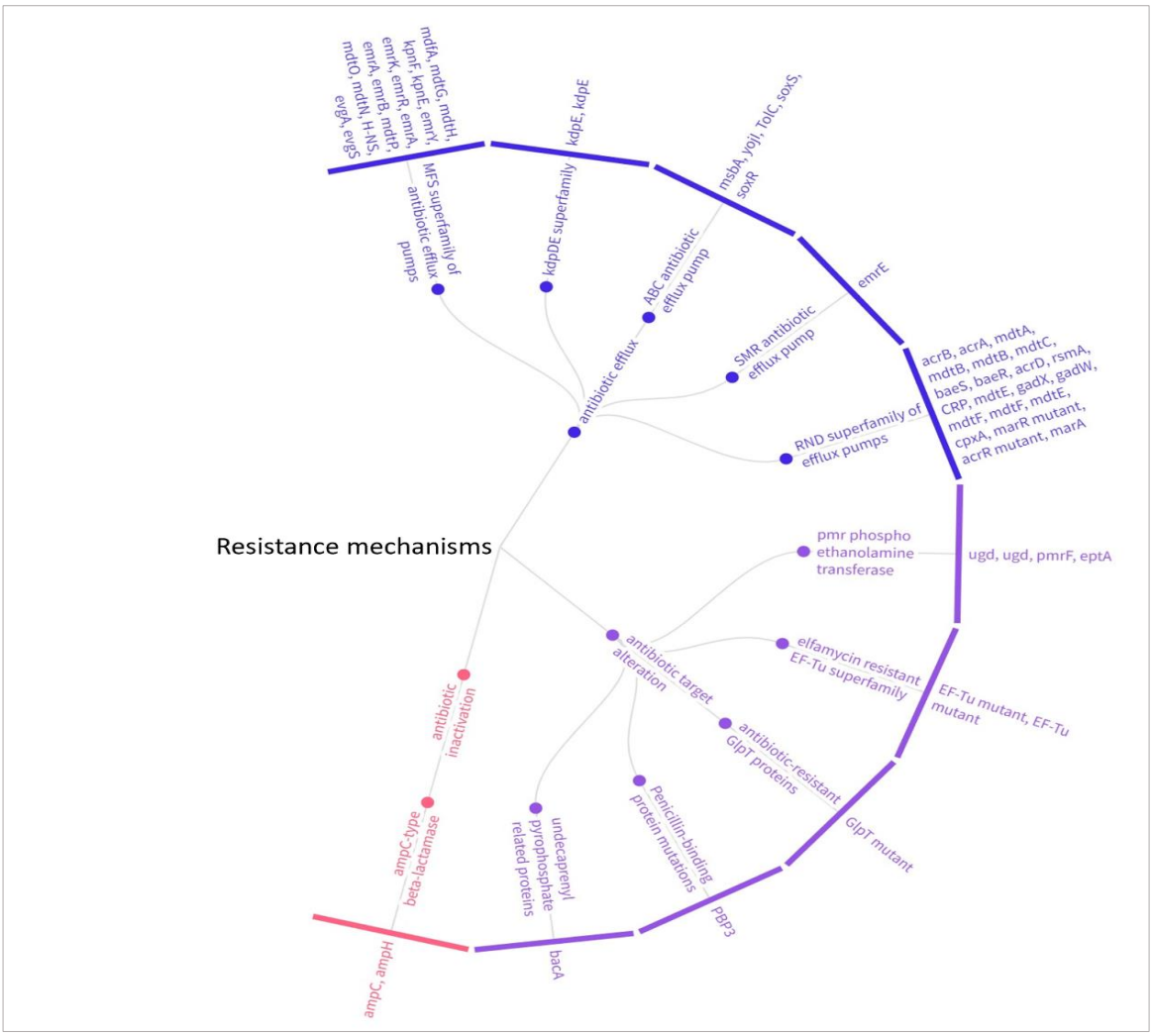


Figure 5.3.3 A hierarchy chart summarizing the AMR genes across different categories based on their mechanism of action. (Mukhopadhyay *et al*., 2024)

The genome yielded five distinct CRISPR arrays, each including one or two spacer sequences and direct repeats spanning 24-53 base pairs on either side (Table 5.3.1). However, it was noticed that not a single CRISPR-associated gene, also known as Cas gene, were detected in the vicinity of any of the CRISPR arrays. The genomic neighbourhood for four of the five isolated CRISPR arrays detected from the isolated pathogen, were found to be in immediate vicinity to the anticipated mobile genomic elements, firmly indicating potential transmission of CRISPR panels via MGEs (Figure 5.3.4).

Table 5.3.1 Characteristics of the computationally identified CRISPR arrays from the SFMMGSG_23 genome.

CRISPR Id	CRISPR Start	CRISPR End	CRISPR Length	Repeat Length	No. of Spacers	Mean size of Spacers	Repeats Conservation (% identity)	Spacers Conservation (% identity)
CP123365_1	64543	64658	115	31	1	54	96.77	100
CP123365_2	877065	877213	148	29	2	31	75.86	32.35
CP123365_3	2363954	2364080	126	39	1	49	100	100
CP123365_4	2832890	2833036	146	53	1	41	100	100
CP123365_5	3246675	3246776	101	24	1	54	100	100

Figure 5.3.4 Correlations between the identified CRISPR arrays and mobile genetic elements in terms of positional proximity.

826 sequences that govern various crucial life cycle processes of portable genetic components (MGEs) were revealed by the mobile-OG-db annotation of the entire genome (Figure 5.3.5).

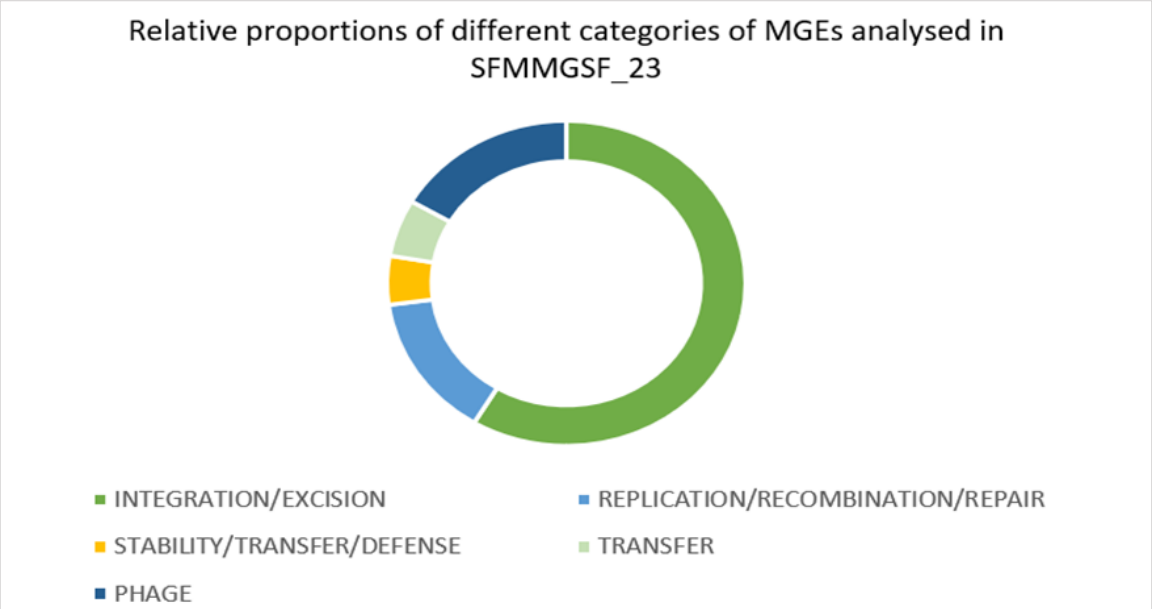


Figure 5.3.5 Distribution of mobile genetic elements derived from phages, insertion sequences, integrative genomic elements (IGEs) and plasmids into different life cycle processes mediated by them (Mukhopadhyay *et al*., 2024).

To reinforce a link between the presence of MGEs and the prevalence of ARGs, the positional proximity between the two were investigated, and 10 ARGs were found to be in close vicinity to the predicted mobile genetic elements, hence compellingly indicating the likely transmission of resistance genes through MGEs (Figure 5.3.6).

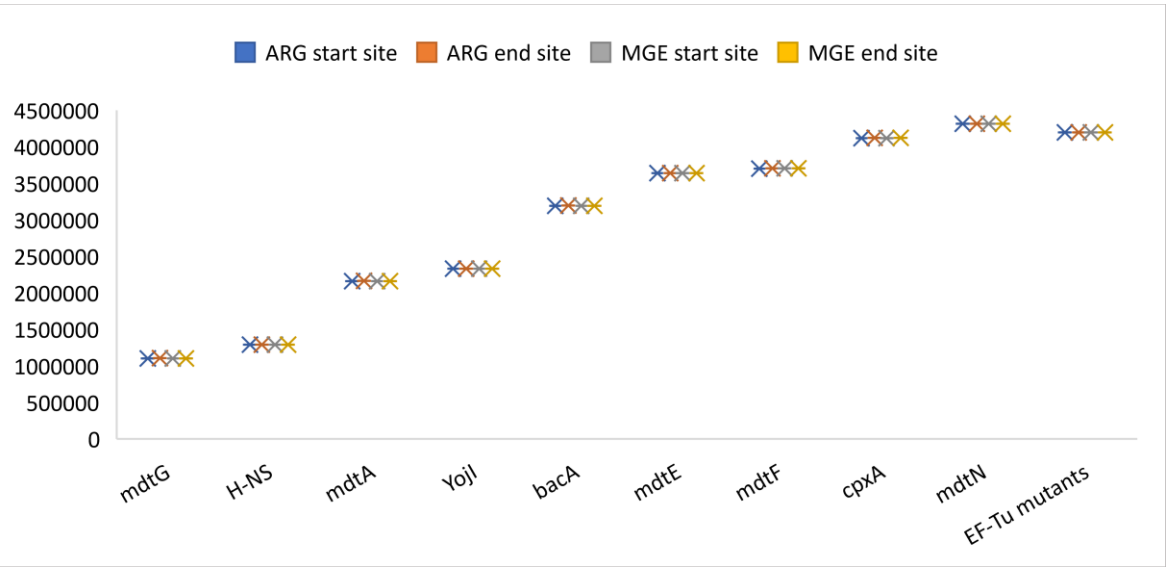


Figure 5.3.6 A box and whisker chart showing the positional proximity between the identified ARGs and MGEs (Mukhopadhyay *et al*., 2024).

AlienHunter was implemented for locating laterally transmitted genomic areas that constitute the bacterial isolate. An inverse association was noticed when the relative incidence of these foreign sections was compared to their sequence length, suggesting that smaller implants might find it less challenging to settle and propagate vertically. (Figure 5.3.7).

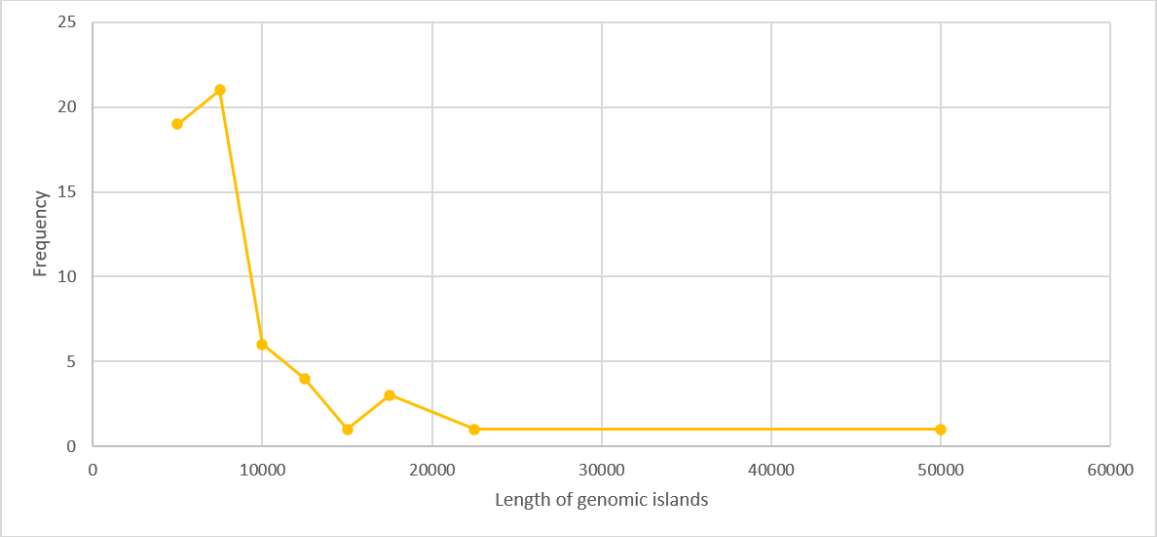


Figure 5.3.7 A scatter chart mapping the association between the length of horizontally acquired genomic regions and their relative frequency.

The independently operated algorithm AntiSMASH was employed to examine the isolated genome for any evidence of biosynthetic gene clusters (Figure 5.3.8).

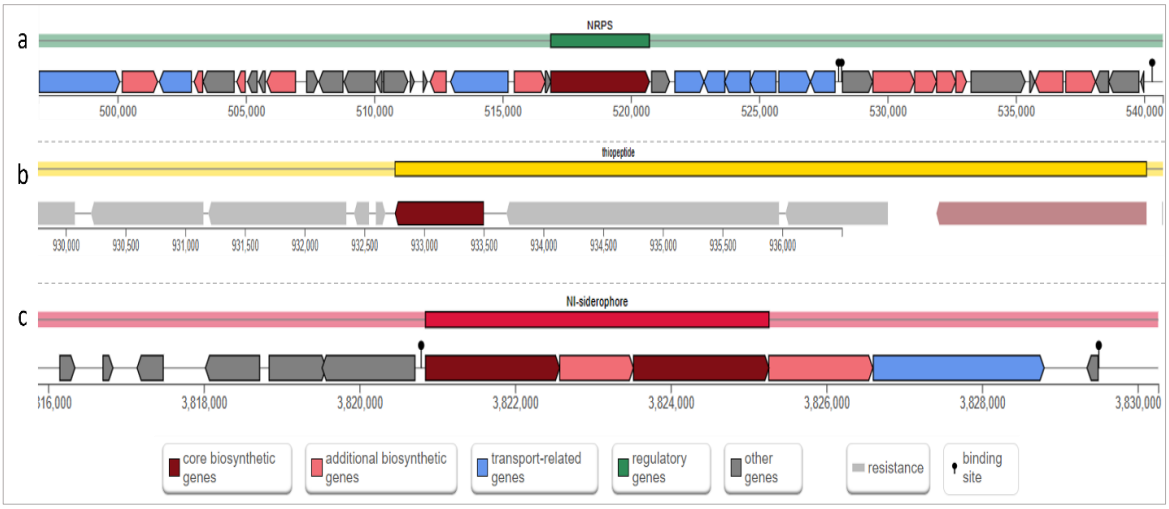


Figure 5.3.8 Schematic representation of the distribution of biosynthetic gene clusters across the assembled genome of SFMMGSG_23; (a) non-ribosomal peptide synthetase

(NRPS), similar to enterobactin biosynthetic gene cluster from *Escherichia coli* str. K-12 substr. MG1655; (b) thiopeptide, similar to O-antigen biosynthetic gene cluster from *Pseudomonas aeruginosa*; (c) NRPS-independent (NI), IucA/IucC-like siderophores, similar to aerobactin biosynthetic gene cluster from *Pantoea ananatis*. (The colour of the BGCs and coding sequences are autogenerated by the server used) (Mukhopadhyay *et al*., 2024).

5.4 Pangenome analysis of *Shigella flexneri*

Using Bacterial Pan Genome Analysis (BPGA), a perl-based tool, a total of 16 strains—1 isolated and 15 chosen—were run through pangenome estimation designed around aggregation of orthologous protein sequences. Clustering revealed a pangenome size of 6907 gene families, of which 2448 families (~35%) made up core genome of *Shigella flexneri*. The unique genome had 1700 gene families, or nearly 25% of the whole gene pool, whereas the accessory genome contained 2759 gene families, or over 40% of the total. (Figure 5.4.1).

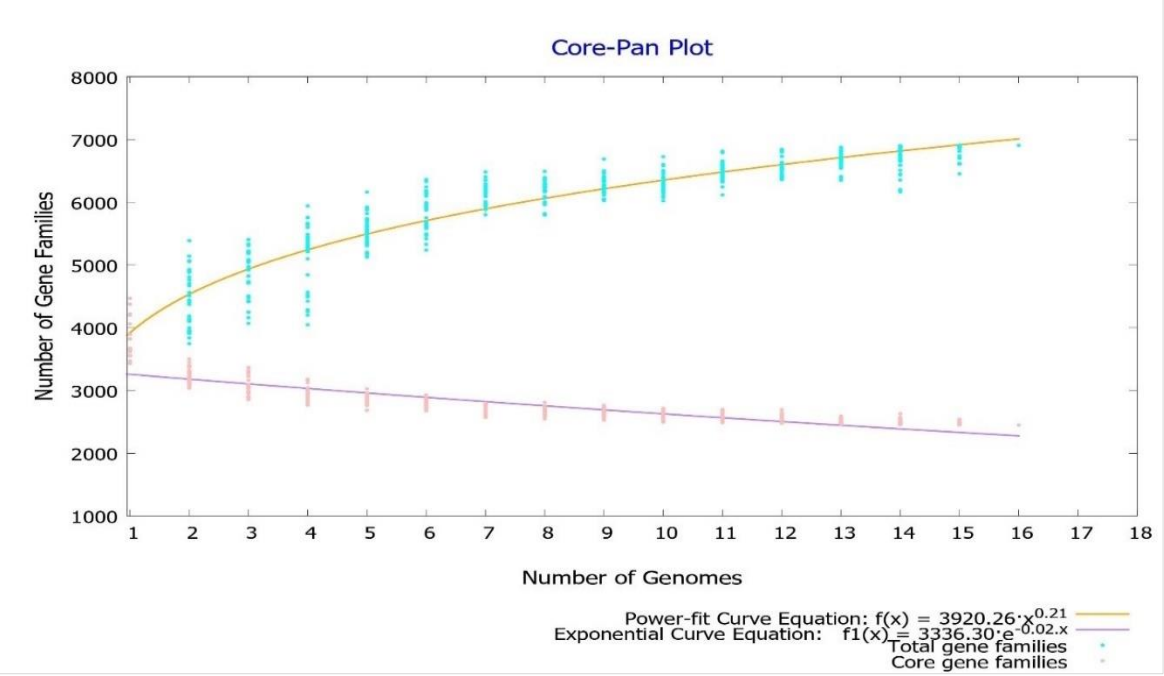


Figure 5.4.1 Graph portraying the effect of dataset size on the pan/core genome dimensions (Mukhopadhyay *et al*., 2024).

All genes were given annotations using the Pangenome Functional Analysis Application within the BPGA program in order to assess the relationship between the prevalence of genes and its related functions (Figure 5.4.2; Figure 5.4.3). A general enrichment of universal genes in molecular events essential to pathogen survival in any circumstance was found by COG analysis of the pangenome compilation, whereas strain-specific exclusive and optional genes were more engaged in niche-specific tasks like membrane biosynthesis, synthesis of virulence determinants, and secretion, which accounted for a large portion of the observed serotype diversity.

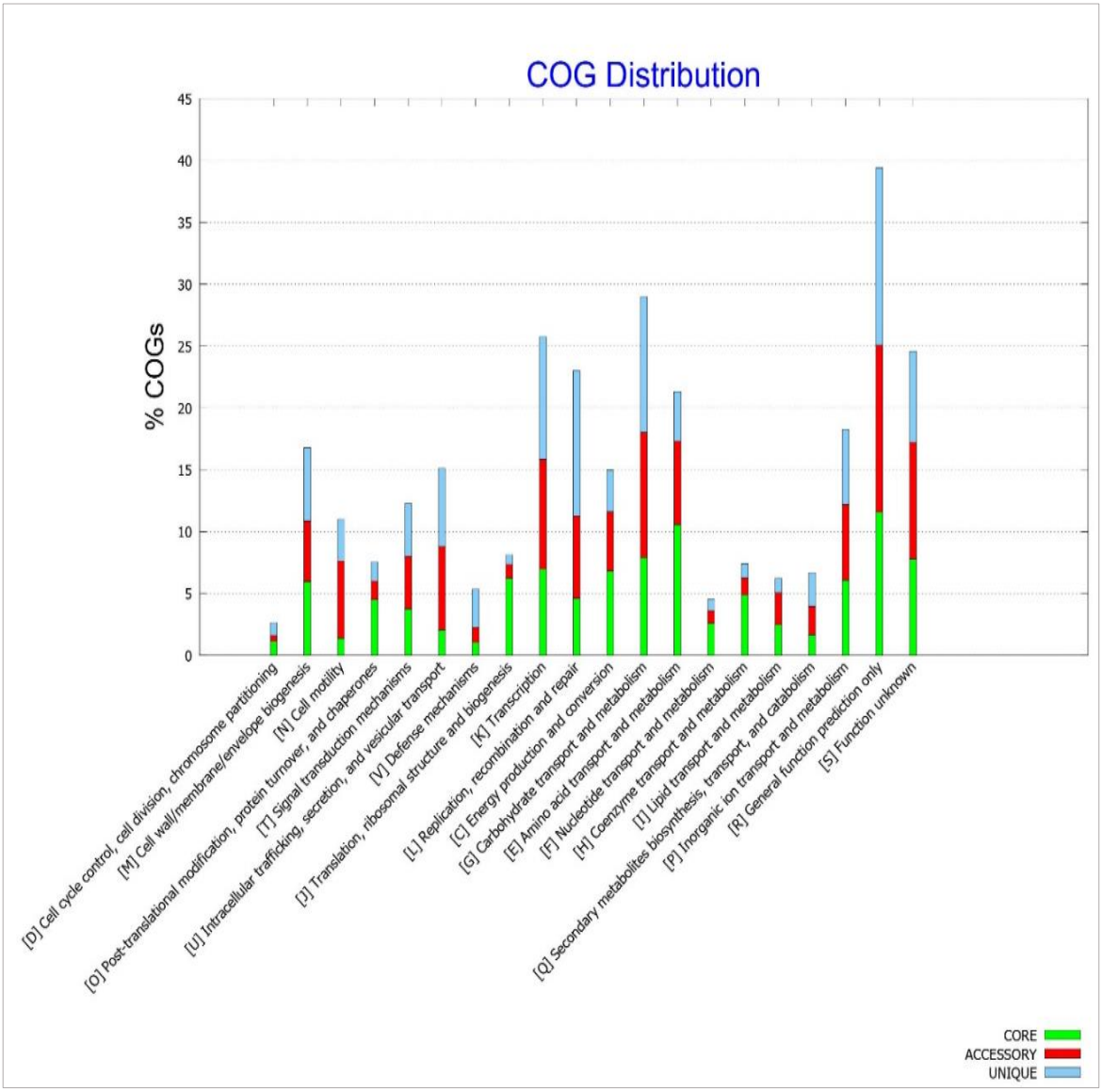


Figure 5.4.2 Functional breakdown of the core, accessory and unique genes into different COG categories.

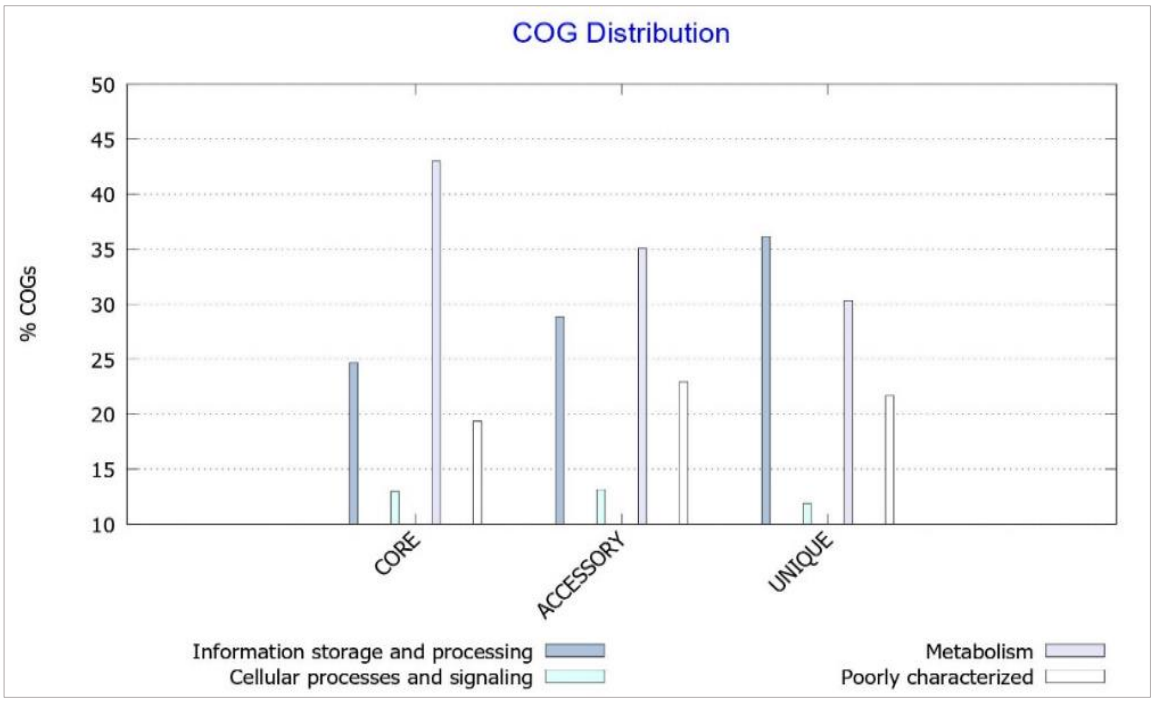


Figure 5.4.3 Bar plots showing the average distribution of COG classes in the core, accessory and unique genome of 16 annotated *Shigella flexneri* strains.

A thorough phylogenetic inquiry was undertaken to ascertain the ancestral connections among the 16 *Shigella flexneri* strains by means of concatenated core gene alliances and the presence and absence order of pan genes. (Figure 5.4.4).

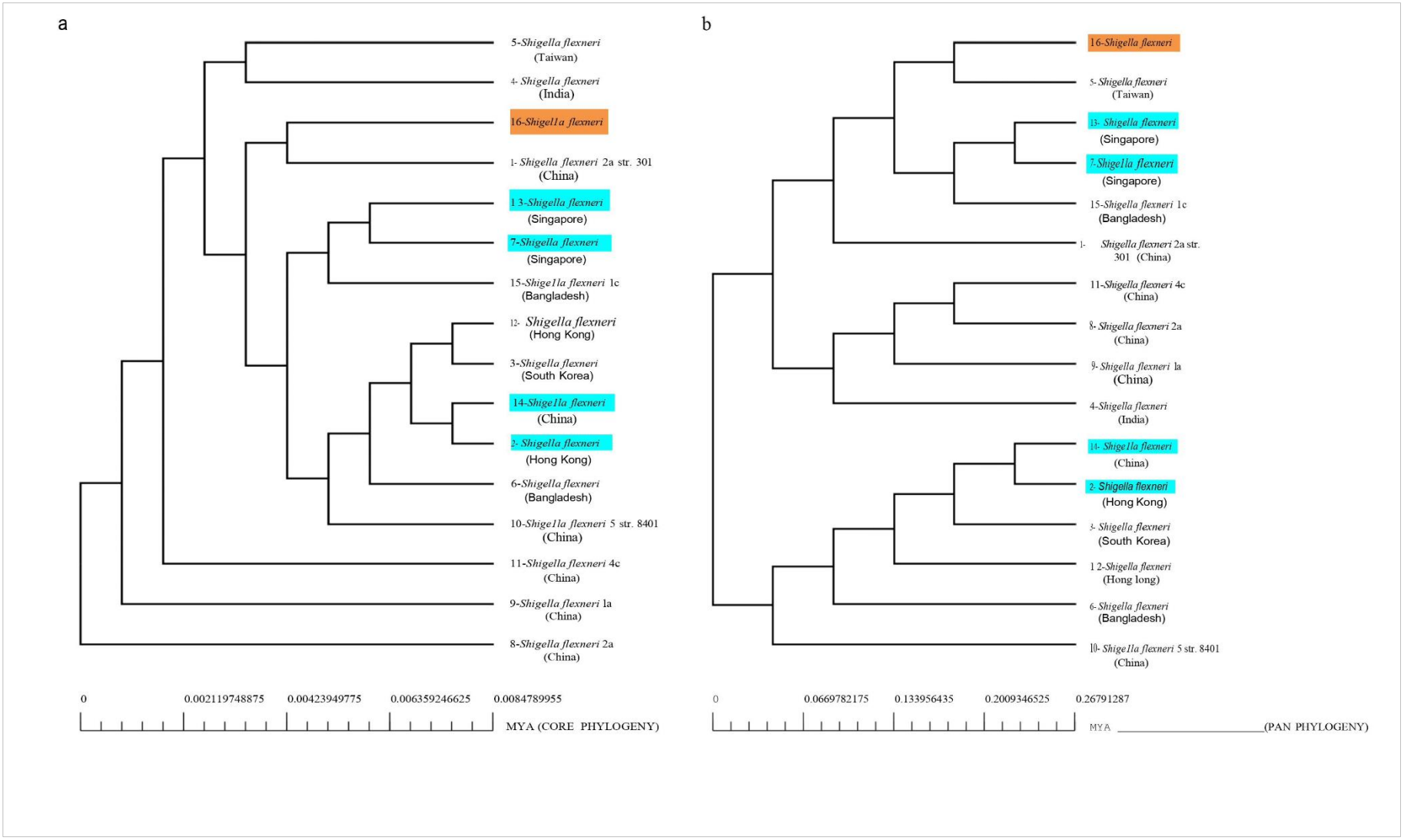


Figure 5.4.4 Evolutionary hierarchy of 16 *Shigella flexneri* strains based on (a) core-genes; (b)pan-matrix. Recurrent clustering patterns are highlighted in blue. The isolated strain is highlighted in orange (Mukhopadhyay *et al*., 2024).

Pathway analysis of the obtained core genomic determinants (2448 core genes) revealed 205 affected pathways (Figure 5.4.5).

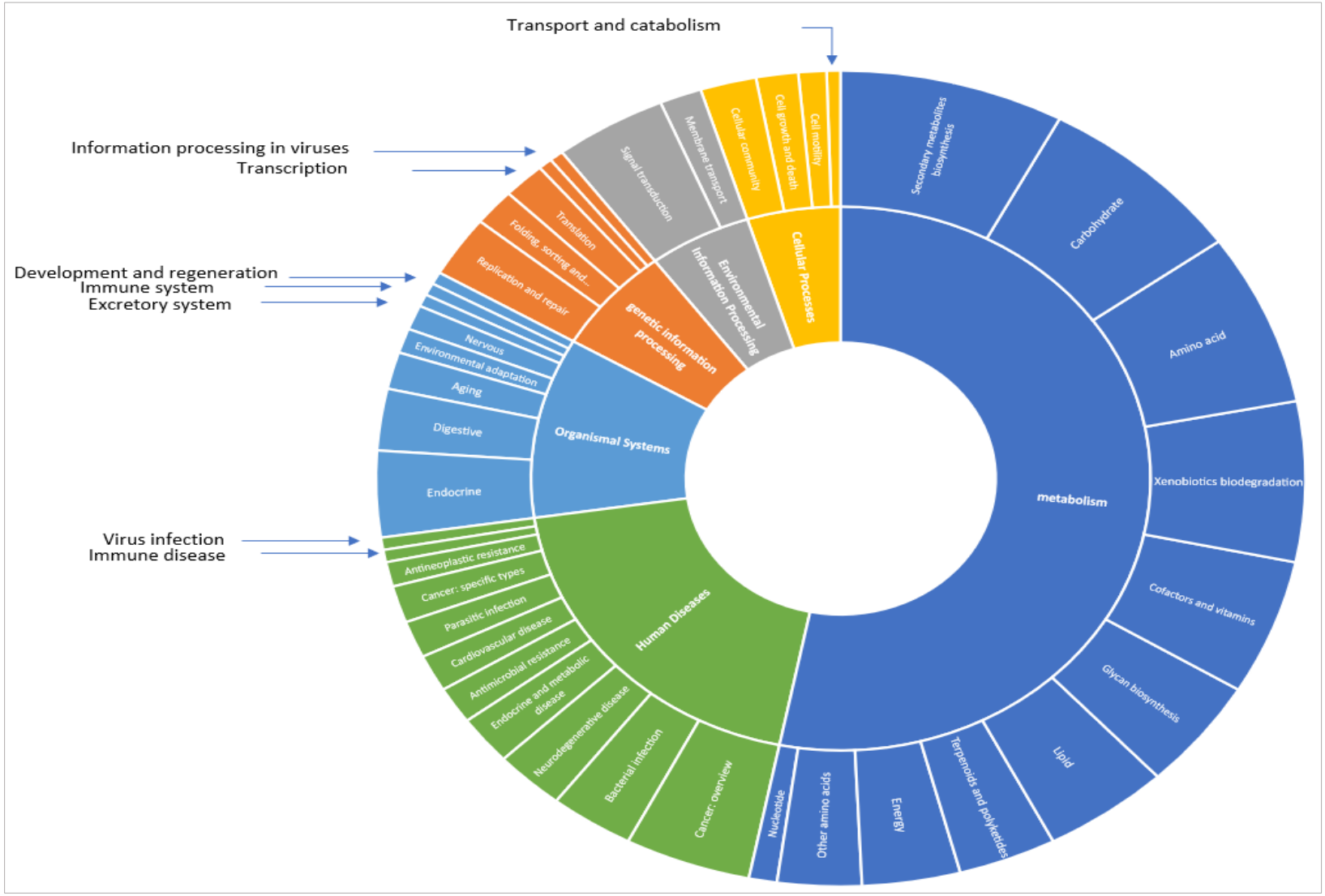


Figure 5.4.5 KEGG pathways influenced by core genes (Mukhopadhyay *et al*., 2024).

Two resistance routes with 22 distinct genes that can either directly or indirectly promote resistance to antibiotics were uncovered by KAAS pathway tracing of 2448 core genes (Figure 5.4.6).

Pathway	Gene	Product
Beta-Lactam resistance	AmpG	MFS transporter, PAT family, beta-lactamase induction signal transducer
	Opp	oligopeptide transport system substrate-binding protein
	OMP	outer membrane protein
	RND	multidrug efflux pump
	MFP	membrane fusion protein, multidrug efflux system
	PBP1a/2	penicillin-binding protein 1A
	PBP2	penicillin-binding protein 2
	FtsI	penicillin-binding protein 3
	MexA/AcrA	multidrug efflux system
	MexB/AcrB	multidrug efflux pump
Cationic antimicrobial peptide (CAMP) resistance	PhoQ	two-component system, OmpR family, sensor histidine kinase PhoQ
	PhoP	two-component system, OmpR family, response regulator PhoP
	PmrB	two-component system, OmpR family, sensor histidine kinase BasS
	PmrA	two-component system, OmpR family, response regulator BasR
	SapB	cationic peptide transport system permease protein
	SapC	cationic peptide transport system permease protein
	SapD	cationic peptide transport system ATP-binding protein
	SapF	cationic peptide transport system ATP-binding protein
	CpxA	two-component system, OmpR family, sensor histidine kinase CpxA
	CpxR	two-component system, OmpR family, response regulator CpxR
	AcrA	membrane fusion protein, multidrug efflux system
	AcrB	multidrug efflux pump

Figure 5.4.6 Ensemble of genes associated with antibiotic resistance pathways (identified by assigning KO numbers to the query genes, followed by gene mapping to the biological pathways using KAAS server) (Mukhopadhyay *et al*., 2024).

Our dataset revealed that the resistance pathways harbored six genes that were identical to the two-component signal transduction systems (TCSs): PmrB, PmrA, PhoQ, PhoP, CpxA, and CpxR (Figure 5.4.7). Additionally, an abundance of genes was discovered which encoded for multidrug efflux pumps (Figure 5.4.8).

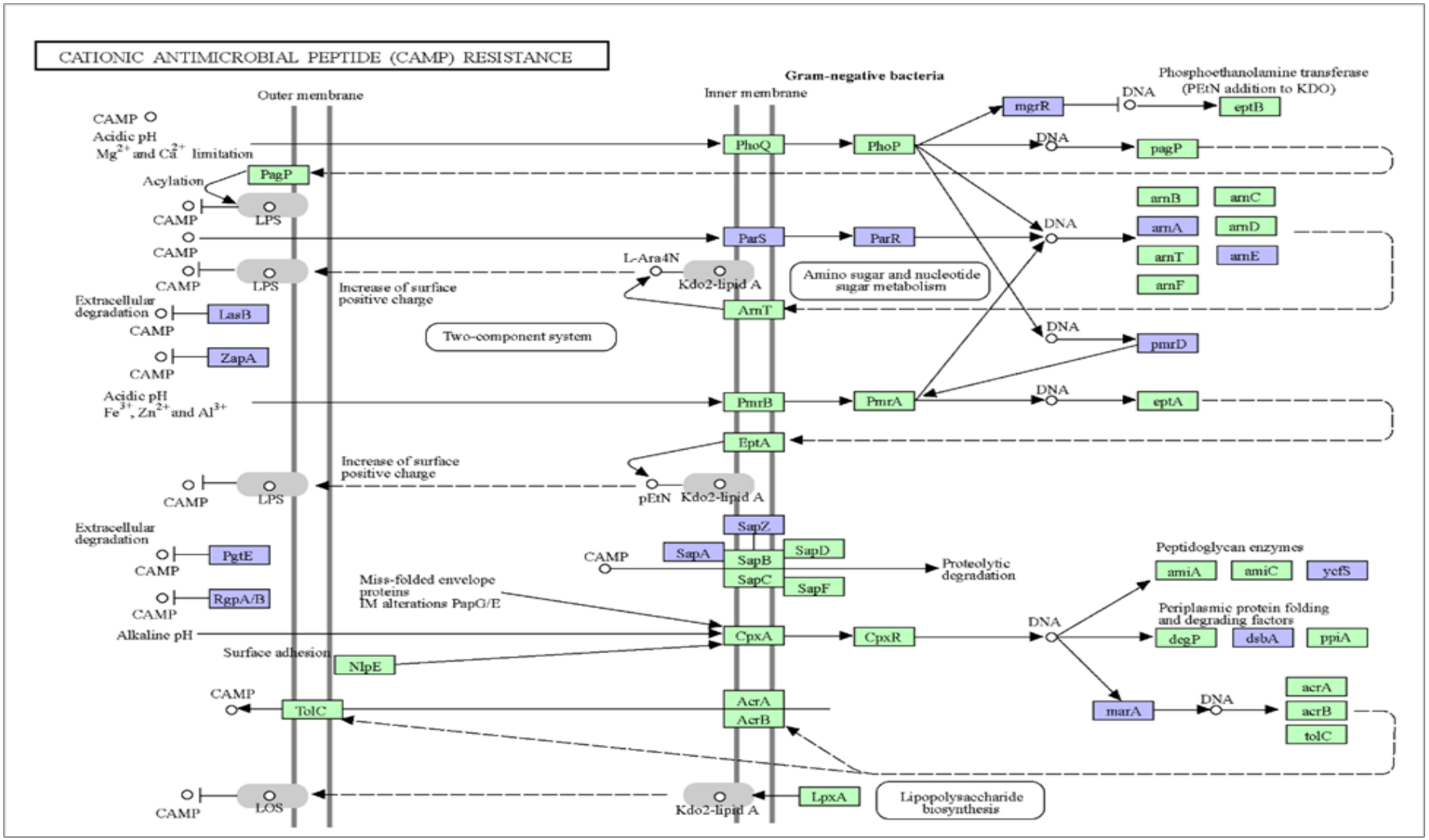


Figure 5.4.7 Cationic antimicrobial peptide resistance pathway in Gram-negative bacteria (the identified resistance genes from the core genome are highlighted in green) (Mukhopadhyay *et al*., 2024).

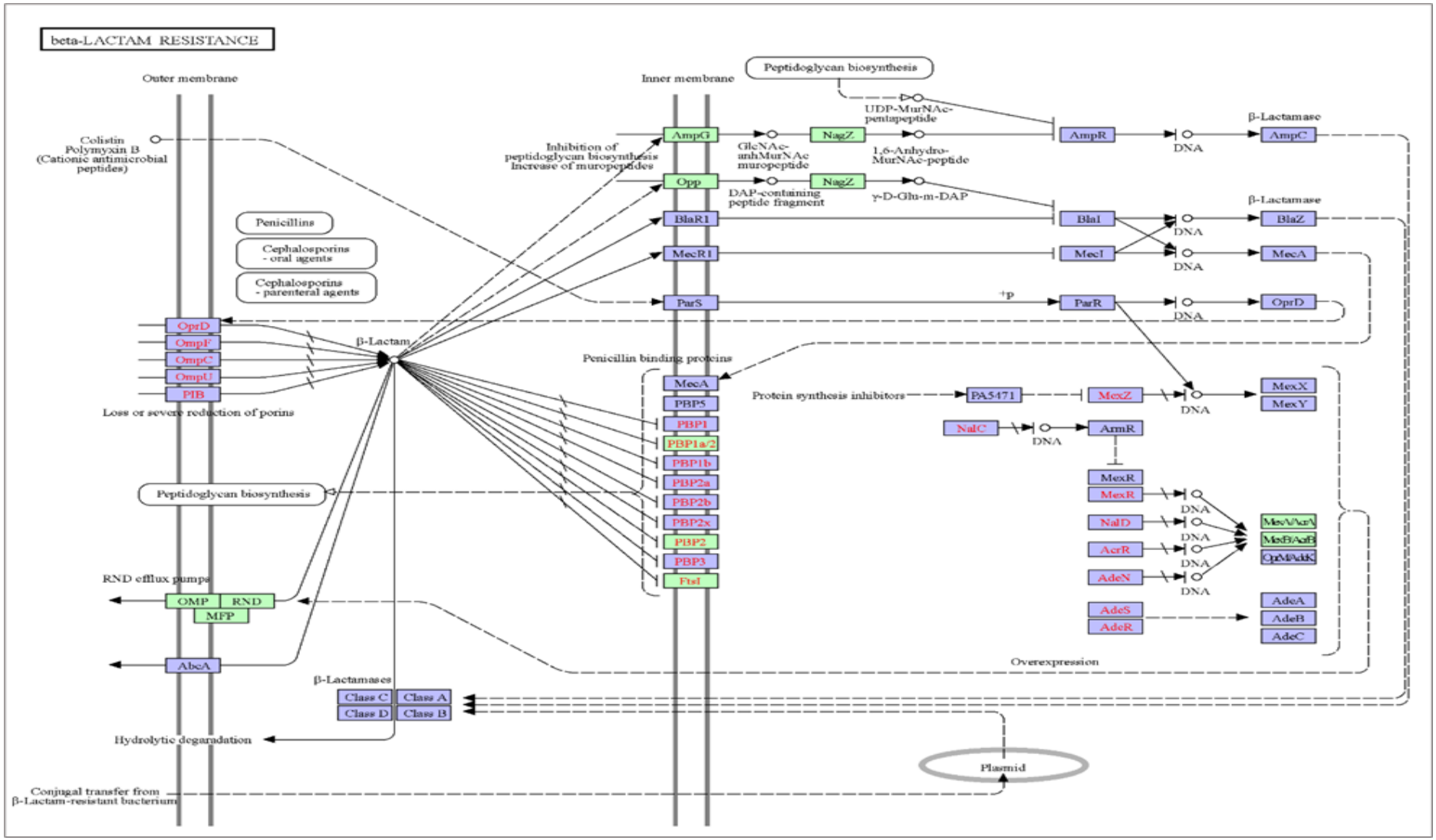


Figure 5.4.8 Beta-lactam resistance pathway in Gram-negative bacteria (the identified resistance genes from the core genome are highlighted in green) (Mukhopadhyay *et al*., 2024).

5.5 Annotation of Hypothetical Proteins

In this work, we have attempted to assign functions to 432 hypothetical proteins from *Shigella flexneri* (Figure 5.5.1). We used a computational pipeline to analyse the gene ontology of the dataset (CC- Cellular Compartment predicted using localization tools, MF- Molecular Function predicted using homology analyses tools, and BP- Biological Pathway predicted using interactomics tools).

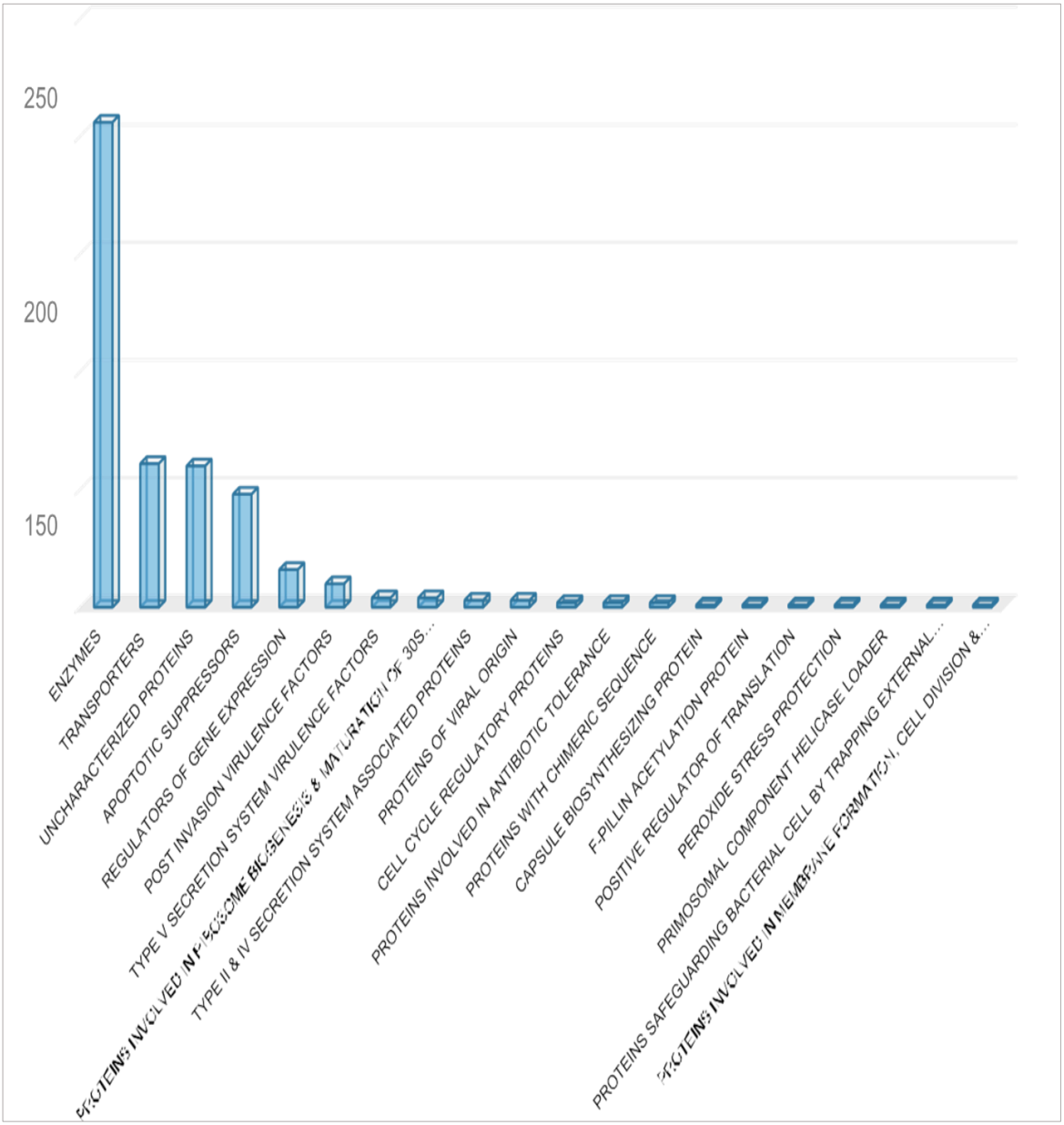


Figure 5.5.1 The cumulative distribution of the pool of 432 hypothetical proteins analysed, into distinct functional categories (Mukhopadhyay *et al*., 2022).

The sequence-based homology method successfully deduced the superfamily of 381 undeciphered proteins (Figure 5.5.2).

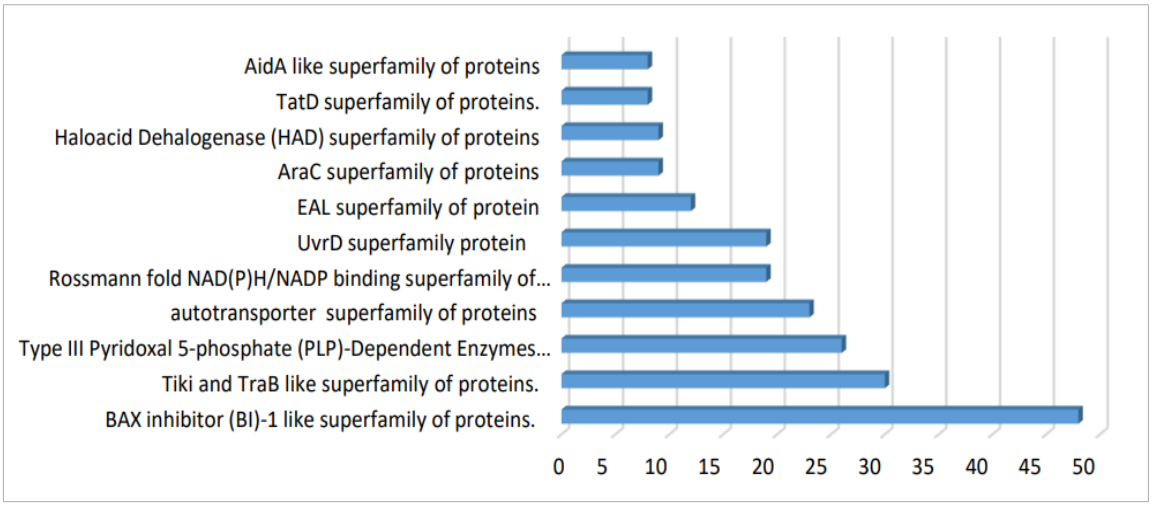


Figure 5.5.2 A graph illustrating the different ubiquitous protein superfamilies detected within the set of unknown proteins examined, along with proportion of proteins belonging to each mentioned head (Mukhopadhyay *et al*., 2022).

The InterPro database was successfully used to identify certain conserved domains of hypothetical proteins. The identification of certain domains such as the CNNM transmembrane domain, CBS domain, AAA+ATPase domain, FAD/NADP binding domain, EAL domain, and others from these proteins indicated their potential participation in important pathogenic pathways of *Shigella*. The placement of HPs within cells is thought to be crucial for their activity. For this goal, two independent computational tools were used: CELLO and PSORTb. Approximately half of the proteins were discovered in the cytoplasm, whereas 25% were located in the membrane. 88% of the potential proteins' sub-cellular addresses were identified with high confidence. 277 proteins were determined to have identical results in both applications, 72 proteins had contradicting results, and 83 proteins had predictions from either tool (Figure 5.5.3).

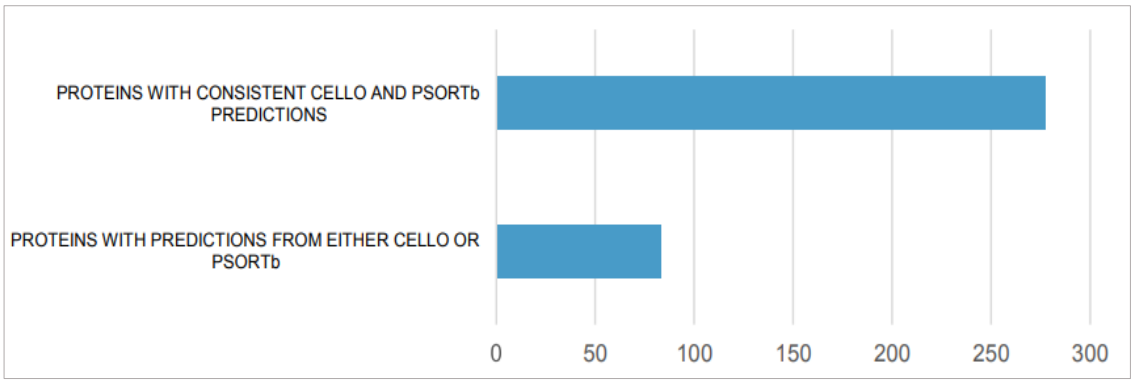


Figure 5.5.3 A bar graph evaluating the accuracy and consistency of the two localisation prediction tools, viz. CELLO and PSORTb, in ascertaining the sub-cellular localisation, for the pool of 432 proteins analysed in the study (Mukhopadhyay *et al*., 2022).

5.6 Prioritization of drug targets from core genome and hypothetical proteins

The core proteins and hypothetical proteins with key functions were searched against the DEG (Database of Essential Genes) database for eukaryotes to perform non-homology analysis with host, followed by an essentiality analysis, which was accomplished by using the DEG database for bacteria. The obligatory non-homologous proteins thus identified were next searched against the core dataset of VFDB (Virulence Factor Database). A combination of template free modelling together with fold recognition-based threading techniques was employed to generate the 3-D models of the potential target proteins that qualified the above check points. The obtained models were assessed for structural stability by using the available benchmarks like Ramachandran plot distribution, Q-mean score and Root Mean Square Fluctuation values. The conformationally sound proteins were pursued for docking interactions with suitable hits to recognize the most target potent candidates (Figure 5.6.1).

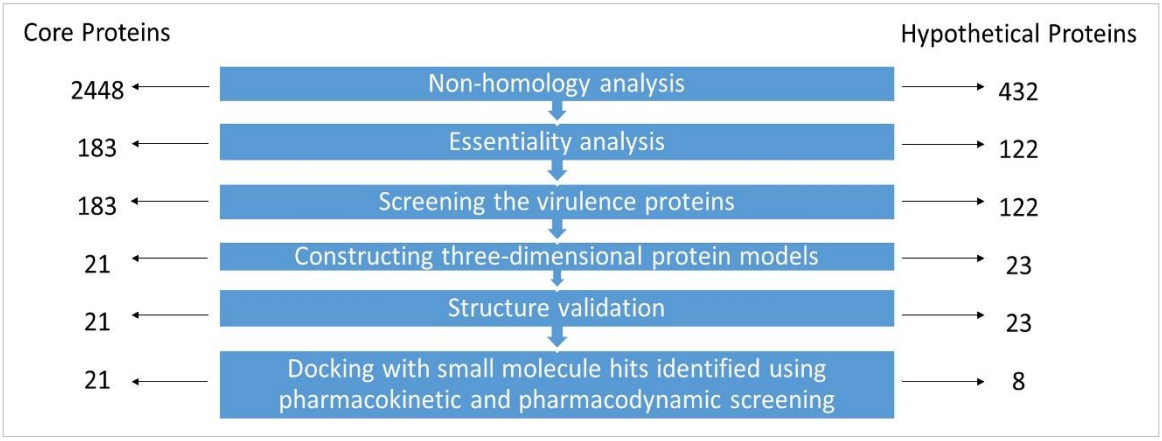


Figure 5.6.1 Screening of drug targets from core genome and hypothetical proteins.

5.7 Curation of Ethnomedicinal Plants and Preparation of Crude Extracts

Several folk medicinal practices have conventionally used plants in treating endemic diarrhoea. We underwent a rigorous text mining to identify such plants and collected them from various medicinal gardens across West Bengal, India (Table 5.7.1). The cleaned and dried leaves were ground in a mechanical blender until they were powdered, and the powder was smashed with methanol in a mortar and pestle. Whatman filter paper was used to filter the resultant crude extract, and the filtrate was used for subsequent analysis.

Table 5.7.1 List of plants collected and analysed.

Sl.no.	Plant names	Family
1	*Glinus oppositifolius* (Gima shag)	Molluginaceae
2	*Aegle marmelos* (Indian Byel)	Rutaceae
3	*Bruguiera gymnorhiza* (Oriental mangrove)	Rhizophoraceae
4	*Avicennia marina* (Grey mangrove)	Acanthaceae
5	*Psidium guajava* (Guava)	Myrtaceae
6	*Paederia foetida* (Gandal lata)	Rubiaceae
7	*Eclipta prostrata* (Bhringraj)	Asteraceae
8	*Mangifera indica* (Mango)	Anacardiaceae
9	*Centella asiatica* (Thankuni)	Apiaceae
10	*Scoparia dulcis* (Bon dhaniya)	Plantaginaceae

5.8 Phenotypic Assays for Antimicrobial Susceptibility Testing

The in vitro susceptibility profile of the selected strain revealed resistance to all the 11 different antibiotics used, thus confirming the isolate as multidrug resistant (MDR) (Figure 5.8.1, Table 5.8.1).

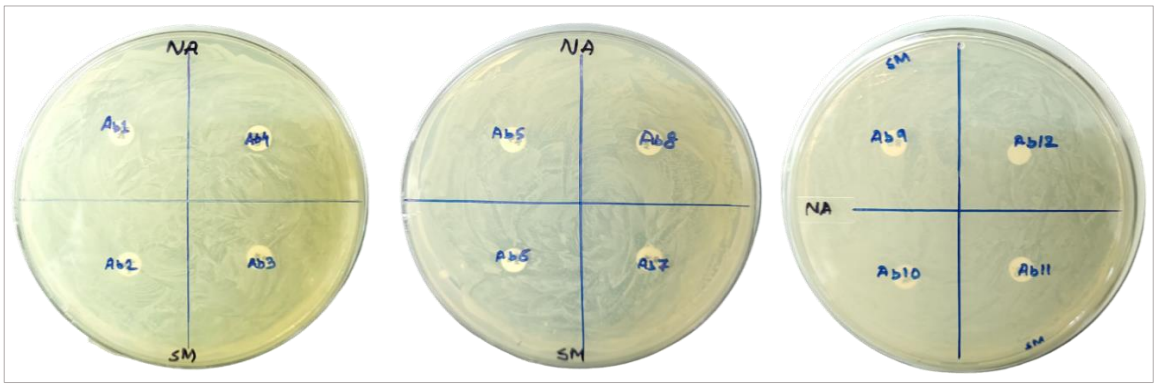


Figure 5.8.1 Resistance profile for the isolated bacteria, obtained on Nutrient agar against; Ab1: Ampicillin, Ab2: Tetracycline, Ab3: Norfloxacin, Ab4: Amoxycillin, Ab6: Gentamicin, Ab7: Amikacin, Ab8: Cefoxitin, Ab9: Imipenem, Ab10: Cefuroxime, Ab11: Trimeth-sulfa, Ab12: Ciprofloxacin.

Table 5.8.1 Antibiotic susceptibility pattern of the isolated bacteria.

Antimicrobial Agent	Disk Content (in µg/ disc)	Observed zone diameter ± SE (in mm)	Interpretive categories and zone diameter breakpoints based on CLSI standards (nearest whole mm)			Susceptibility
			S[1]	I[2]	R[3]	
Ampicillin (Ab1)	10	6.5±0.5	≥17	14-16	≤13	R
Tetracycline (Ab2)	30	6.5±0.5	≥15	12-14	≤11	R
Norfloxacin (Ab3)	10	6±0.6	≥17	13-16	≤12	R
Amoxycillin (Ab4)	30	7±0.5	≥18	14-17	≤13	R
Gentamicin (Ab5)	10	6±0.8	≥18	15-17	≤14	R
Amikacin (Ab6)	30	6.5±0.65	≥20	17-19	≤16	R
Cefoxitin (Ab7)	30	6.5±1.20	≥18	15-17	≤14	R

[1] S: Susceptible
[2] I: Intermediate
[3] R: Resistant

Imipenem (Ab8)	10	7.5±1.30	≥23	20-22	≤19	R
Cefuroxime (Ab9)	30	7.5±0.8	≥18	15-17	≤14	R
Trimeth-sulfa (Ab10)	25	6.5±0.5	≥16	11-15	≤10	R
Ciprofloxacin (Ab11)	5	6.5±1.20	≥31	21-30	≤20	R
Ceftriaxone (Ab12)	30	6.5±0.5	≥ 26 - ≥ 23	23-25 and 20-22	≤22 - ≤ 19	R

The best inhibition was observed with the crude extracts of *Psidium guajava* and *Scoparia dulcis* (Figure 5.8.2, Table 5.8.2).

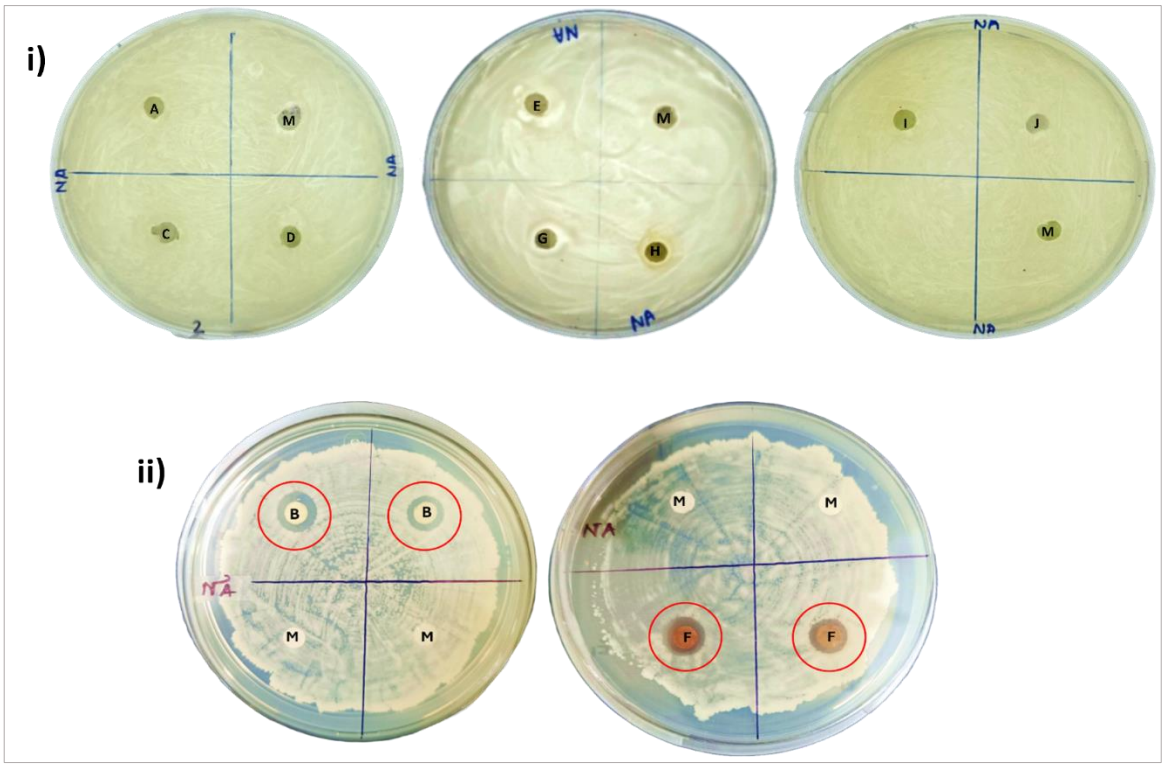


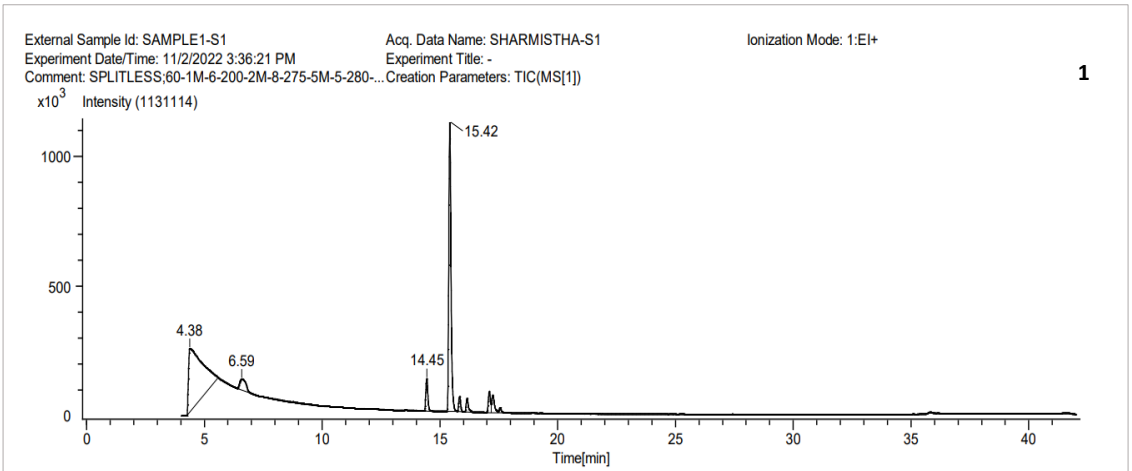
Figure 5.8.2 Susceptibility/resistance profile of the isolated strain of *Shigella flexneri* against the selected herbal extracts; i) No zone of inhibition was observed on nutrient agar for the isolated bacteria against crude extracts of A: *Mangifera indica*; C: *Centella asiatica;* D: *Avicennia marina;* E: *Aegle marmelos;* G: *Bruguiera gymnorhiza;* H: *Glinus oppositifolius;* I: *Eclipta prostrata;* J: *Paederia foetida* ii) Zone of inhibition obtained on nutrient agar for the isolated bacteria against crude extracts of B: *Psidium guajava*; F: *Scoparia dulcis*; M: methanol was used as a control in each assay.

Table 5.8.2 Susceptibility pattern of the isolated bacteria against crude herbal extracts.

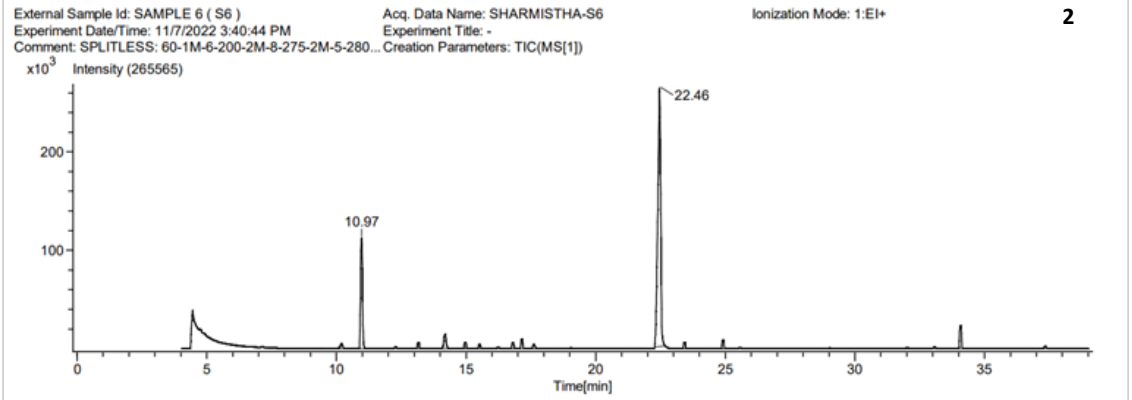
SOURCE PLANT	ZONE OF INHIBITION ± SE (measured in mm)
A. *Mangifera indica* (Mango)	7.5±0.5
B. *Psidium guajava* (Guava)	12.6± 1.20
C. *Centella asiatica* (Thankuni)	6.4±0.5
D. *Avicennia marina* (Grey mangrove)	6±0.5
E. *Aegle marmelos* (Indian Byel)	7.6±0.4
F. *Scoparia dulcis* (Bon dhaniya)	9.8±0.6
G. *Bruguiera gymnorhiza* (Oriental mangrove)	7±0.3
H. *Glinus oppositifolius* (Gima shag)	7.5±0.4
I. *Eclipta prostrata* (Bhringraj)	7.2±0.6
J. *Paederia foetida* (Gandal lata)	6.5±0.5
M. Methanol	6±0.6

5.9 Spectrometric analysis and virtual screening of small molecules

2152 compounds were identified by GC-MS analysis, and these were then evaluated for drug-likeness based on structural or physicochemical characteristics that have a significant impact on the molecules' pharmacokinetics and pharmacodynamics ((Figure 5.9.1, Table 5.9.1). The pipeline made it possible to identify 23 promising molecules, which were then further correlated with the results of in-vitro assays. Finally, three compounds were identified, namely limonene from *Psidium guajava* and 5-hydroxy-1-isopropyl-6,6-dimethyl-5-phenyl-piperidin-2-one and N-(2-hydroxyphenyl)-2,6-dimethoxybenzamide from *Scoparia dulcis*, (Figure 5.9.2, Figure 5.9.3).



Peak Number	Time [min]	Type	Peak Width(FWH [min]	Area [Intens. * sec]	Height	Description	Start Point		End Point	
							Time[min]	Height	Time[min]	Height
1	4.38	BB	1.2607	9405129.35	244457.07		4.24	44	5.58	147789
2	6.59	BB	0.3081	789211.24	42176.85		6.37	106722	6.95	87899
3	14.45	BB	0.0946	751769.51	124958.12		14.25	18906	14.67	18444
4	15.42	BB	0.1004	7345234.60	1113728.02		15.29	17026	15.71	18199
5	15.84	BB	0.0939	337520.76	57231.91		15.74	16962	16.01	15320
6	16.16	BB	0.0980	356740.96	52413.32		16.04	15023	16.41	14148
7	17.10	BV	0.1027	498888.22	81857.13		16.95	12503	17.17	12260
8	17.25	VV	0.1271	550199.84	67768.90		17.17	12260	17.47	11928
9	17.56	VB	0.0972	136311.07	20620.68		17.47	11928	17.84	11524



Peak Number	Time [min]	Type	Peak Width(FWH [min]	Area [Intens. * sec]	Height	Description	Start Point		End Point	
							Time[min]	Height	Time[min]	Height
1	10.20	BB	0.1056	38004.76	5162.59		9.93	143	10.33	53
2	10.97	BB	0.0879	605215.05	112726.42		10.82	45	11.12	202
3	13.15	BB	0.0866	31005.06	5962.62		13.03	46	13.24	54
4	14.19	BB	0.1015	97135.53	14598.06		14.04	42	14.35	181
5	14.97	BB	0.0866	31919.55	6144.83		14.82	39	15.05	54
6	15.51	BB	0.0866	22600.94	4339.33		15.43	53	15.66	55
7	16.79	BB	0.0867	32746.60	6287.59		16.69	52	16.92	43
8	17.14	BB	0.0866	50357.01	9692.74		17.07	43	17.23	50
9	17.62	BB	0.0924	26505.80	4460.27		17.47	48	17.80	101
10	22.46	BB	0.1246	2049616.89	263202.04		22.27	1733	22.67	3057
11	23.43	BB	0.0866	35087.71	6754.33		23.33	49	23.50	53
12	24.90	BB	0.0866	45163.17	8663.46		24.81	49	25.05	49
13	34.07	BB	0.0866	121004.49	23259.37		33.98	47	34.22	117

Figure 5.9.1 Representative images of peaks observed on GC-MS analysis of methanolic extract for 2 selected plants, viz. 1.: *Psidium* and 2.: *Scoparia*.

Table 5.9.1 Table representing some of the active ingredients identified on GC-MS analysis of methanolic extract of 2 chosen plants at different time intervals.

Retention Time (in minutes)	Active principles in MeOH extract of Guava
4.38	1-Fluoroforskolin
6.59	Limonene
14.44	Ylangene
15.42	Isocaryophillene
15.84	Aromadendrene
16.16	α-Caryophyllene
17.1	Docosahexaenoic acid, 1,2,3-propanetriyl ester
17.25	Lycopene
17.56	Zeaxanthin

Retention Time (in minutes)	Active Principles in MeOh Extract of *Scoparia*
10.2	1H-Imidazole, 1-methyl-5-nitro
10.97	4-tert-Butoxystyrene
13.15	2-Methoxy-4-vinylphenol
14.19	Borazine, 2,4,6-trimethyl
14.97	Decane, 2,4-dimethyl
15.51	Styrene
16.79	Malonic acid, eicosyl isobutyl ester
17.14	Methenamine
17.62	Mesalamine

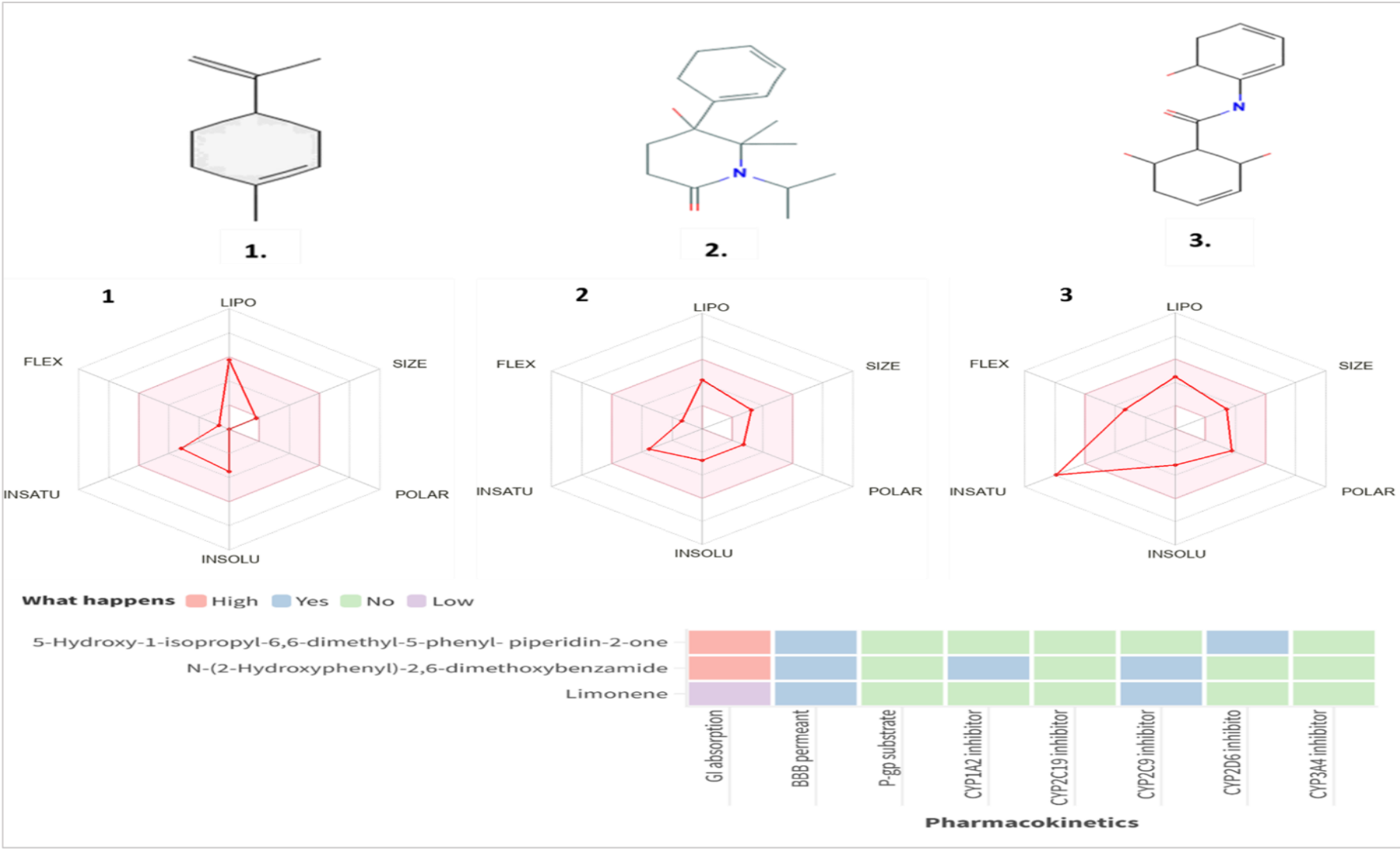


Figure 5.9.2 2-D Structures of the selected natural ingredients along with their drug-likeness and pharmacokinetic properties; 1.: Limonene; 2.: 5-Hydroxy-1-isopropyl-6,6-dimethyl-5-phenyl- piperidin-2-one; 3.: N-(2-Hydroxyphenyl)-2,6-dimethoxybenzamide.

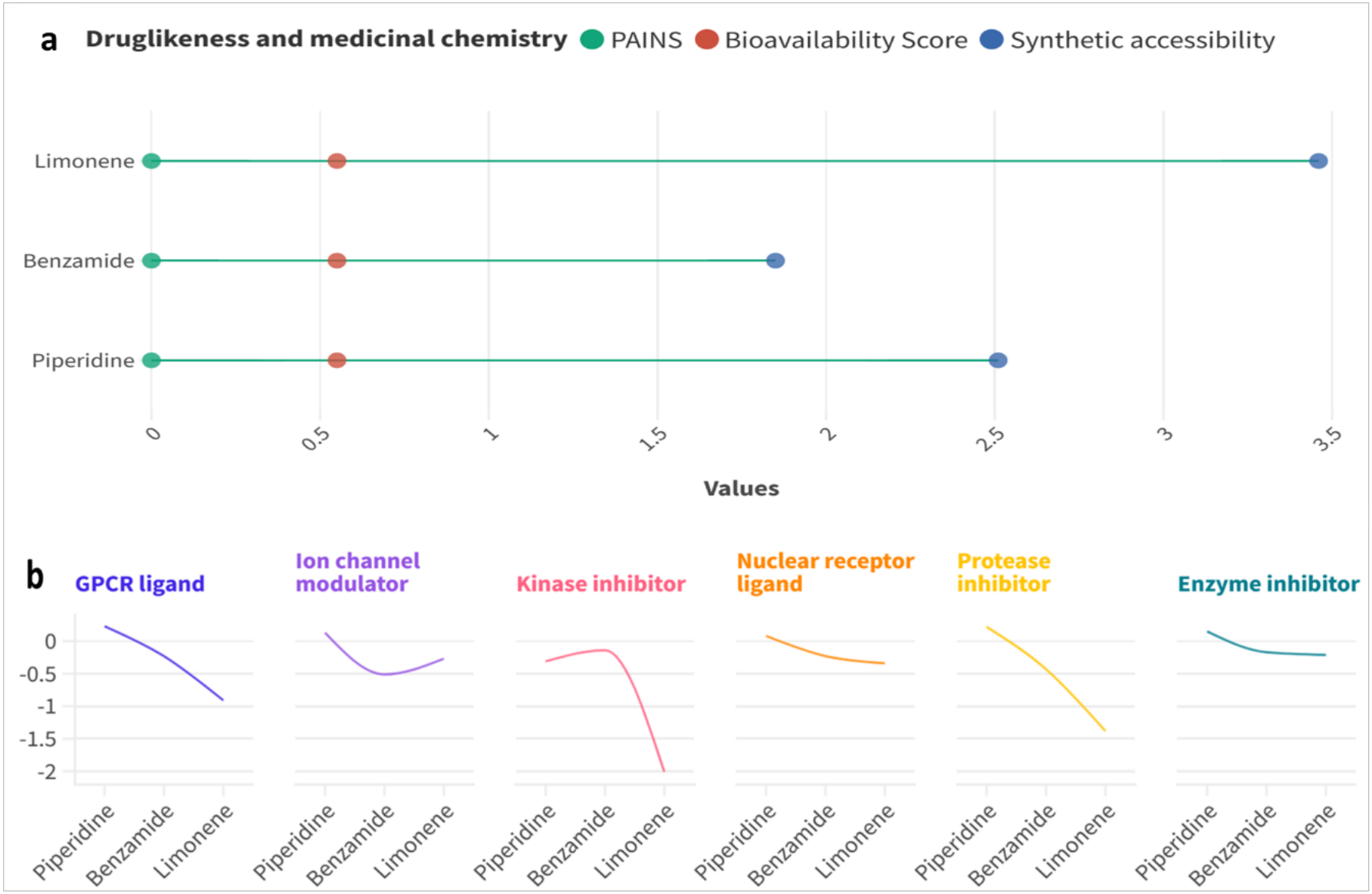


Figure 5.9.3 Pharmacodynamic properties and bioactivity profiling of selected hits.

5.10 Docking interaction analysis and identification of potential targets

The prospective targets were screened against the 3 selected herbal compounds, viz. Limonene, Piperidine and Benzamide separately to verify the binding efficacy of these target candidates using docking and post docking analysis (Figure 5.10.1; Figure 5.10.2; Figure 5.10.3; Figure 5.10.4; Figure 5.10.5; Figure 5.10.6)

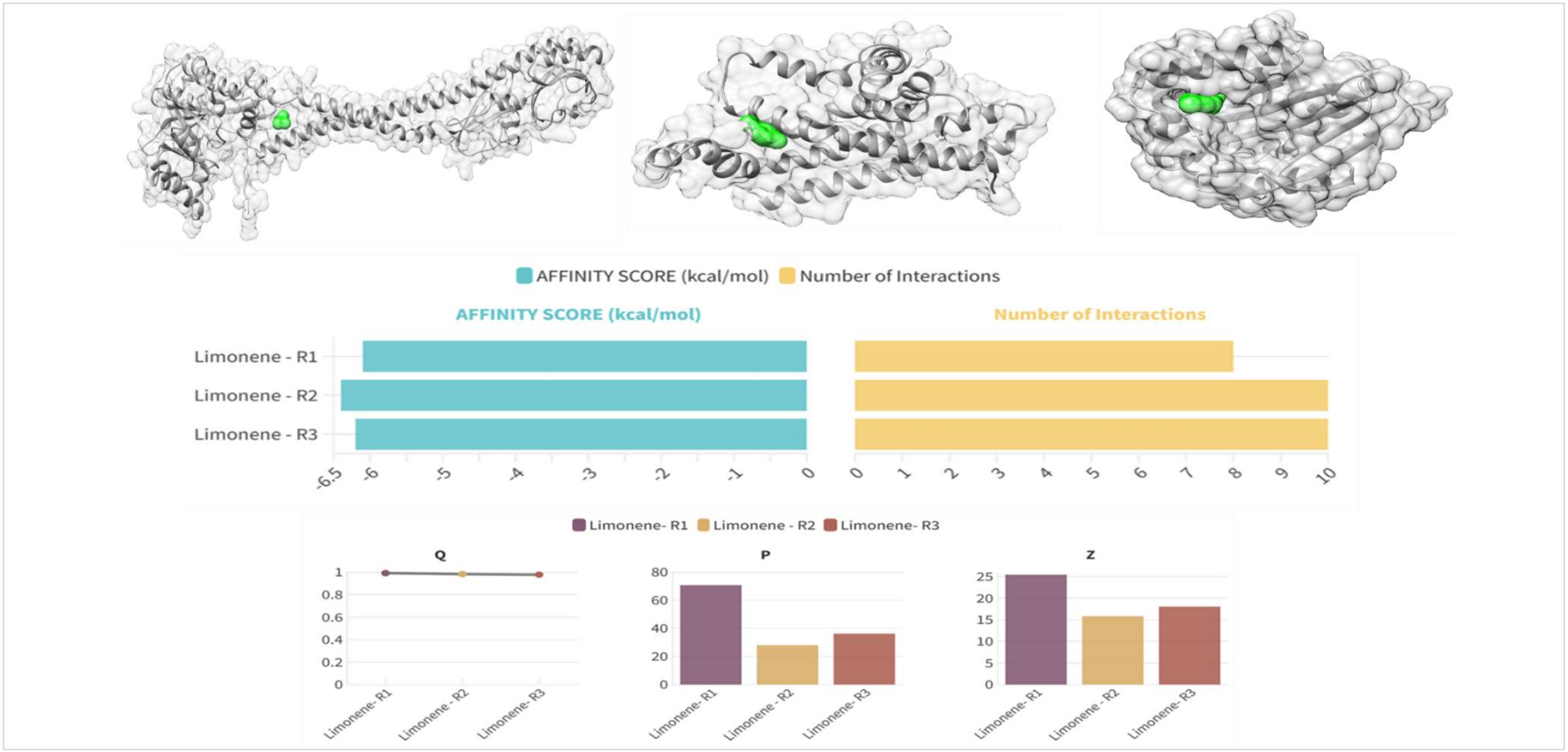


Figure 5.10.1 Docking profile of limonene with top 3 targets from core proteins (R1, R2, R3), based on binding scores and structural alignments; target proteins in grey and ligands in green.

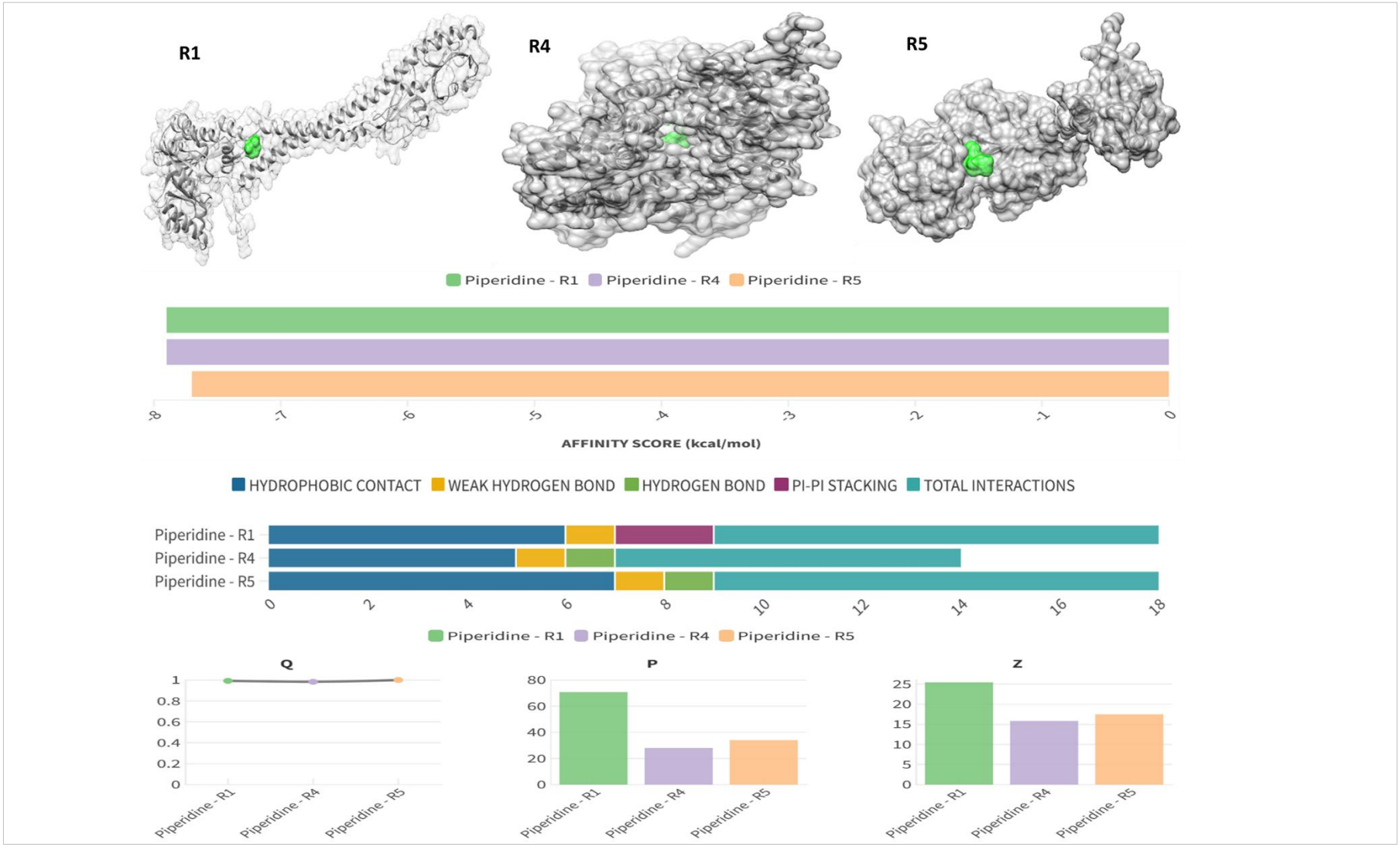


Figure 5.10.2 Docking profile of piperidine with top 3 targets from core proteins (R1, R4, R5), based on binding scores and structural alignments; target proteins in grey and ligands in green.

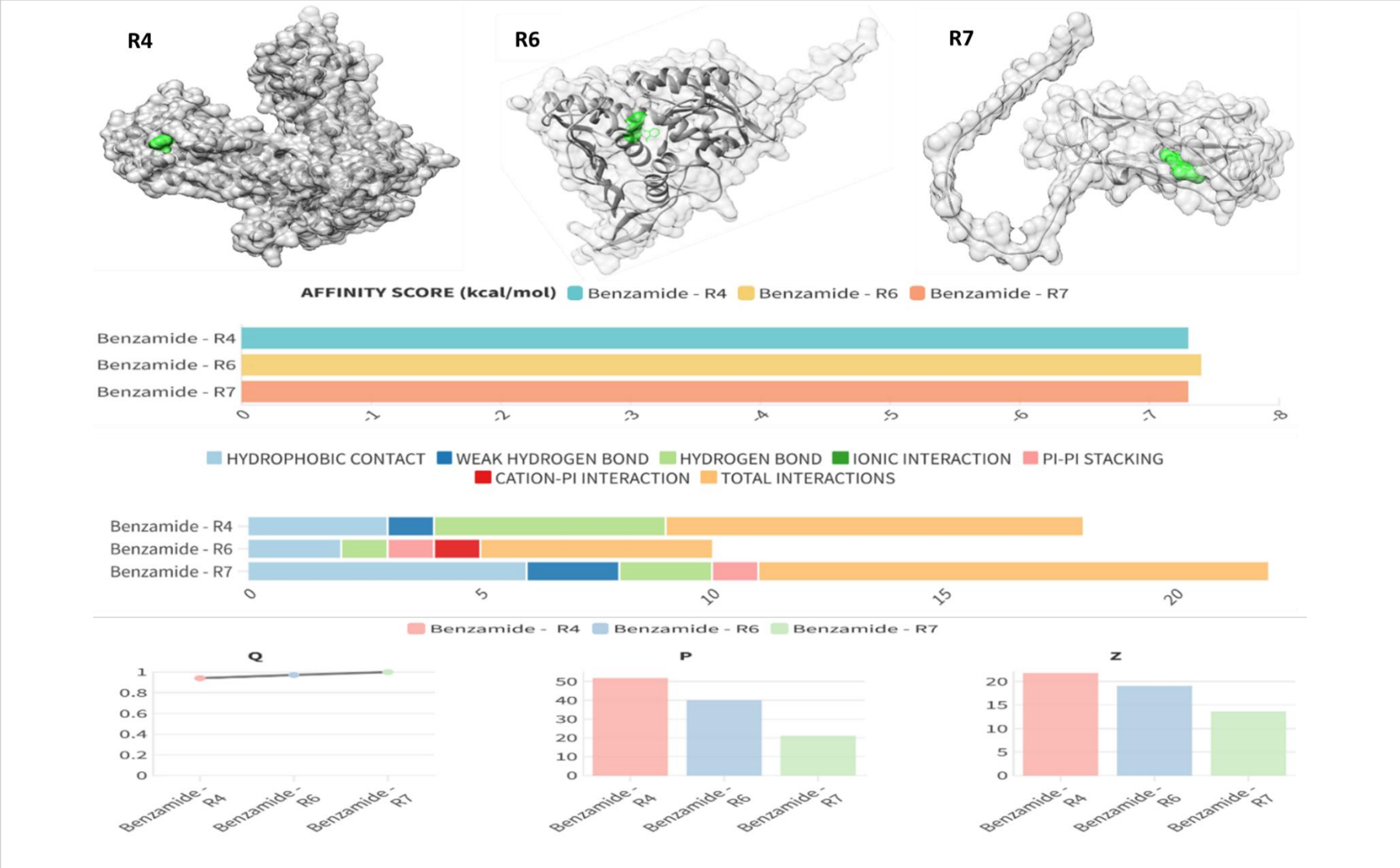


Figure 5.10.3 Docking profile of benzamide with top 3 targets from core proteins (R4, R6, R7), based on binding scores and structural alignments; target proteins in grey and ligands in green.

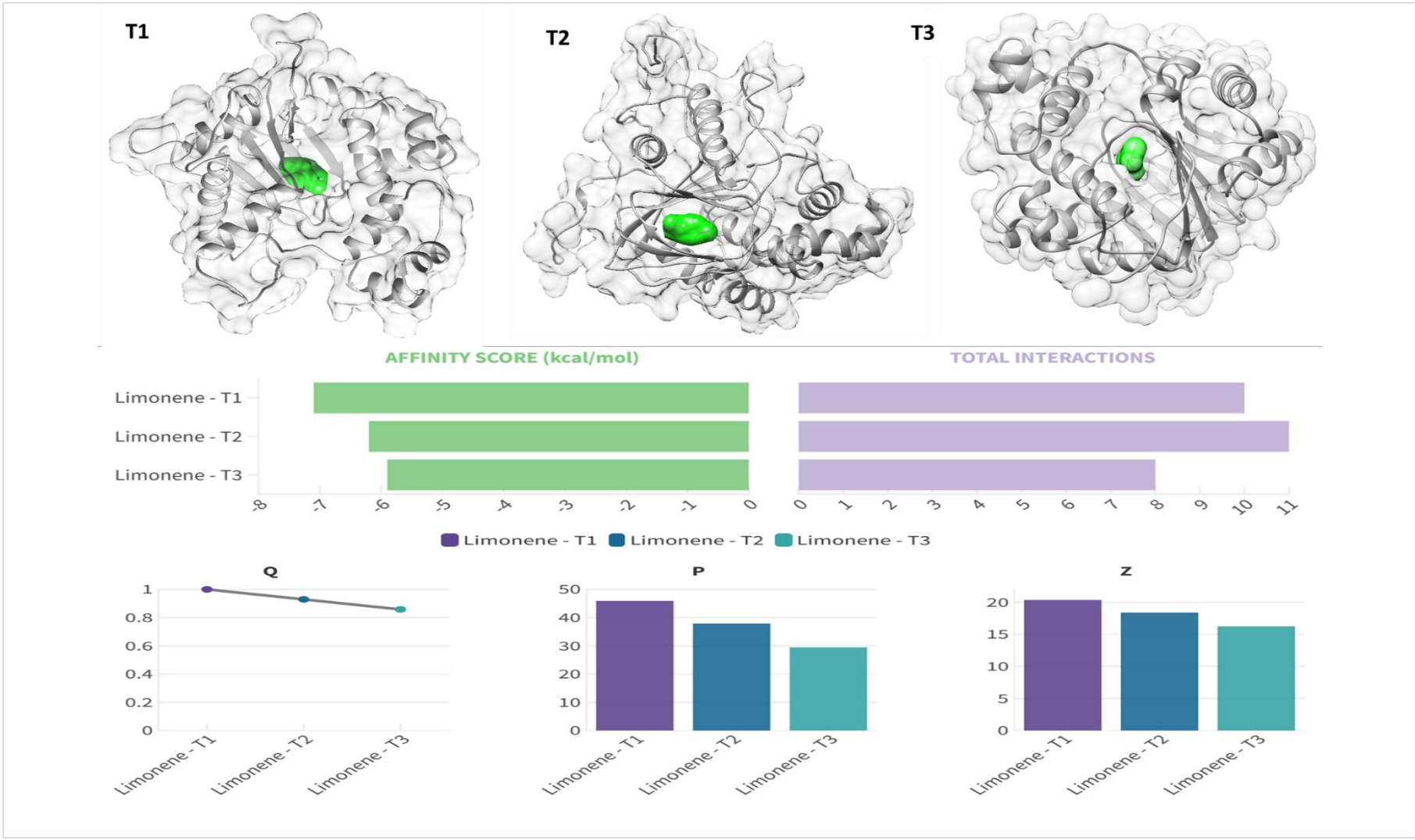


Figure 5.10.4 Docking profile of limonene with top 3 targets from hypothetical proteins (T1, T2, T3), based on binding scores and structural alignments; target proteins in grey and ligands in green.

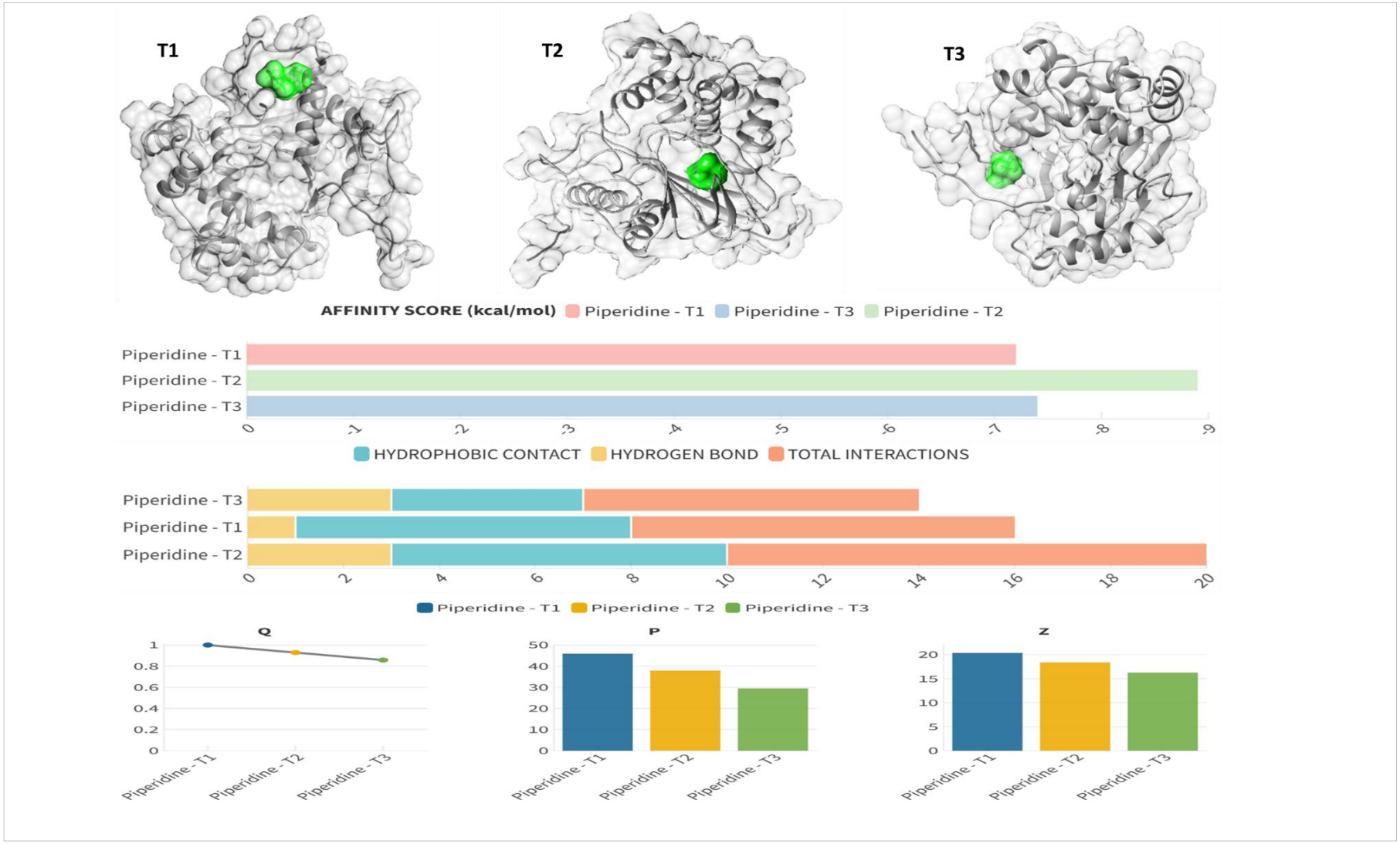


Figure 5.10.5 Docking profile of piperidine with top 3 targets from core proteins (T1, T2, T3), based on binding scores and structural alignments; target proteins in grey and ligands in green.

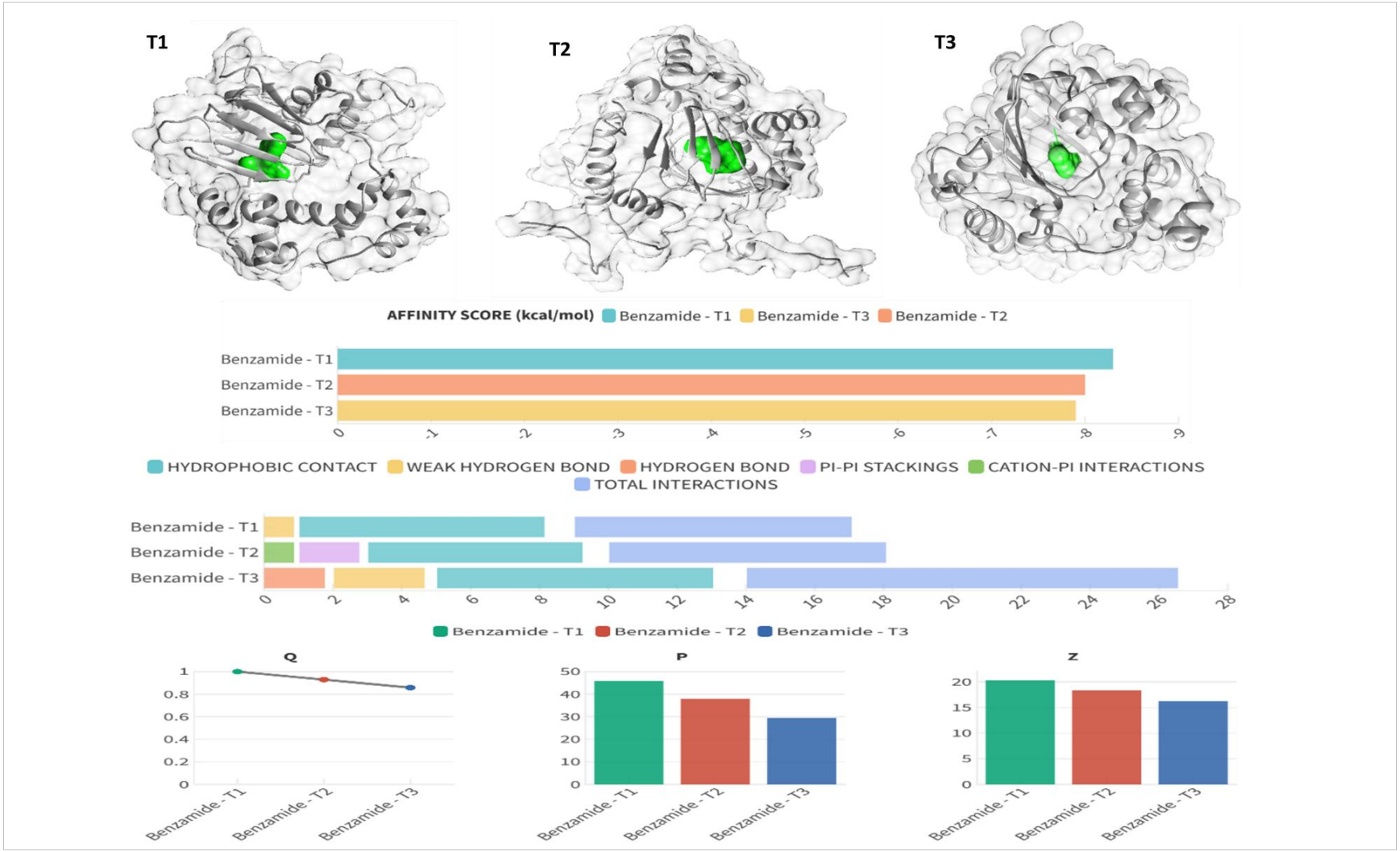


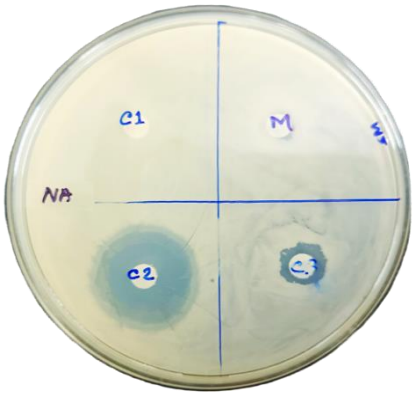
Figure 5.10.6 Docking profile of benzamide with top 3 targets from core proteins (T1, T2, T3), based on binding scores and structural alignments; target proteins in grey and ligands in green.

Table 5.10.1 Prioritized Targets from Core Genome and Hypothetical Proteins along with their Molecular Functions and Biological Pathways.

TARGETS	CODING GENES	MOLECULAR FUNCTIONS	BIOLOGICAL PATHWAYS
R1	adrB, pdeD	c-di-GMP phosphodiesterase	Biofilm formation - *Escherichia coli*
R2	modB	molybdate transport system permease protein	ABC transporters
R3	rimL	ribosomal-protein-serine acetyltransferase	Genetic information processing
R4	thiP	thiamine transport system permease protein	ABC transporters
R5	yhaJ	LysR family of transcriptional regulators	Regulation of DNA-templated transcription
R6	ydgJ	Conserved expressed oxidoreductase	Various metabolic processes
R7	yfcV	fimbrial adhesion protein elfA	Cell adhesion involved in single-species biofilm formation and pilus organization
T1	perA	AraC superfamily of proteins	Regulation of DNA-templated transcription
T2	ydiP	AraC superfamily of proteins	Regulation of DNA-templated transcription
T3	ripA	AraC superfamily of proteins	Regulation of DNA-templated transcription

5.11 In-vitro assays for efficacy testing of leads against the isolated *Shigella* strain

Zones of inhibition of bacterial growth were monitored by performing disc assays with the identified leads following identical protocol as executed earlier (Figure 5.11.1).



ACTIVE INGREDIENTS	ZONE OF INHIBITION ± SE (measured in mm)
C1. Benzamide	7± 0.5
C2. Piperidine	35± 3.4
C3. Limonene	15± 0.3
M. Methanol	6±0.2

Figure 5.11.1 Susceptibility pattern of the isolated *Shigella* strain observed using selected lead compounds.

Of the three chosen leads, piperidine and limonene demonstrated remarkable zones of inhibition in the disc diffusion assay, while benzamide did not exhibit any susceptibility zone. Thus, all subsequent analysis were carried out using these two compounds. To further establish the potential of the leads, bacterial cell viability assays were performed to evaluate the impact of the leads as well as the crude extracts on bacterial survival. In theory, this assay distinguishes between cells with damaged membranes (dead cells) and those with intact membranes (living cells) using two nucleic acid-binding stains. The side scatter (SSC-A) and forward scatter (FSC-A) quadrant plots display the cells that were identified based on physical parameters. Dead cells were identified using propidium iodide marker, whereas thiazole orange was used to label living cells (Figure 5.11.2). The results revealed increased permeability of cells to propidium iodide in cultures treated with crude extracts and identified chemical leads. Whereas in untreated bacterial culture, the cells were less permeable to propidium iodide as evidenced by a steady proportion of live cells (96.12%). The transition from untreated to treated cultures reduced the cell viability to less than 20% upon administration of *Psidium guajava* extract, while the viability was brought down to 31.75%, 56.65% and 71.45% upon application of the *Scoparia dulcis* extract, piperidine and limonene respectively.

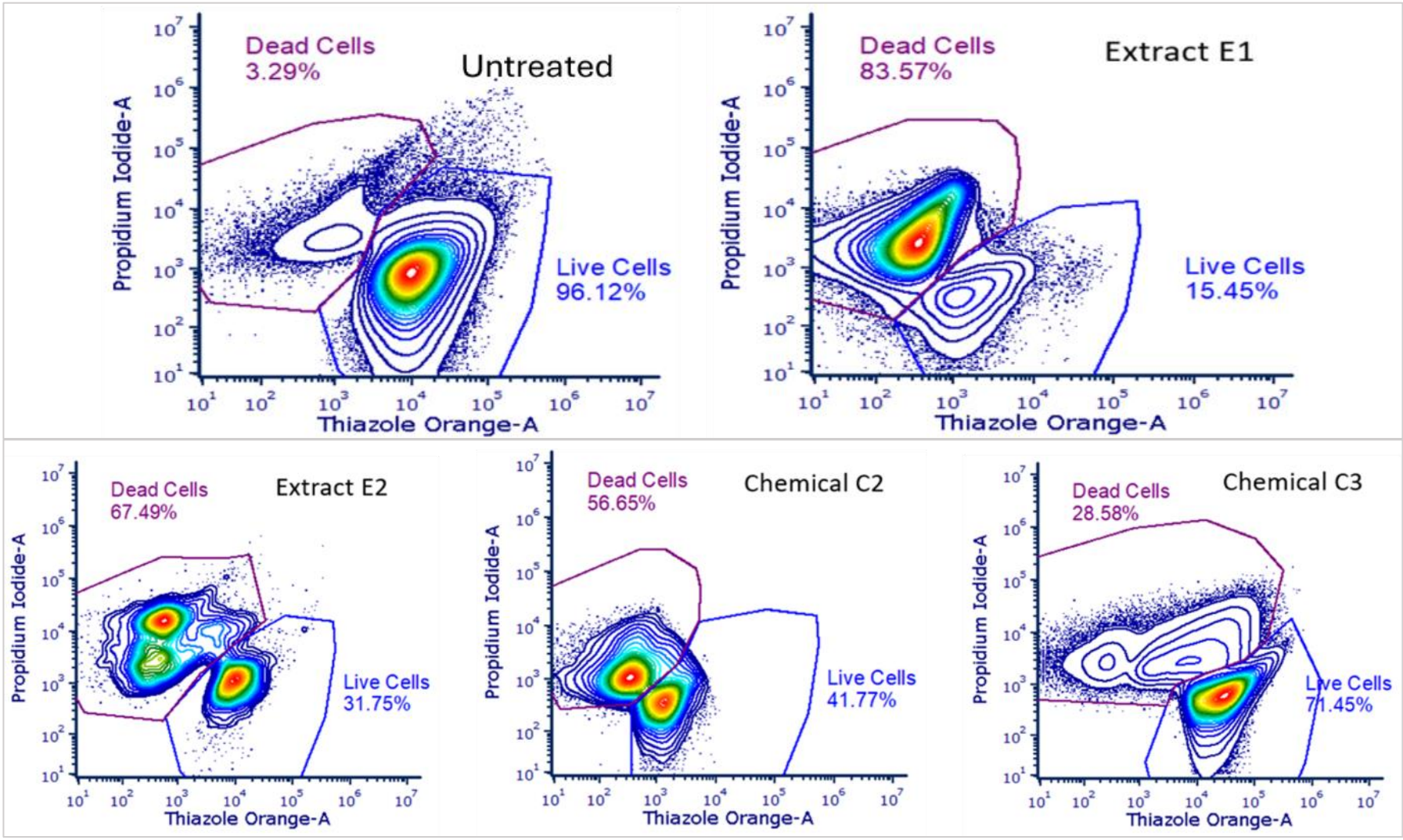
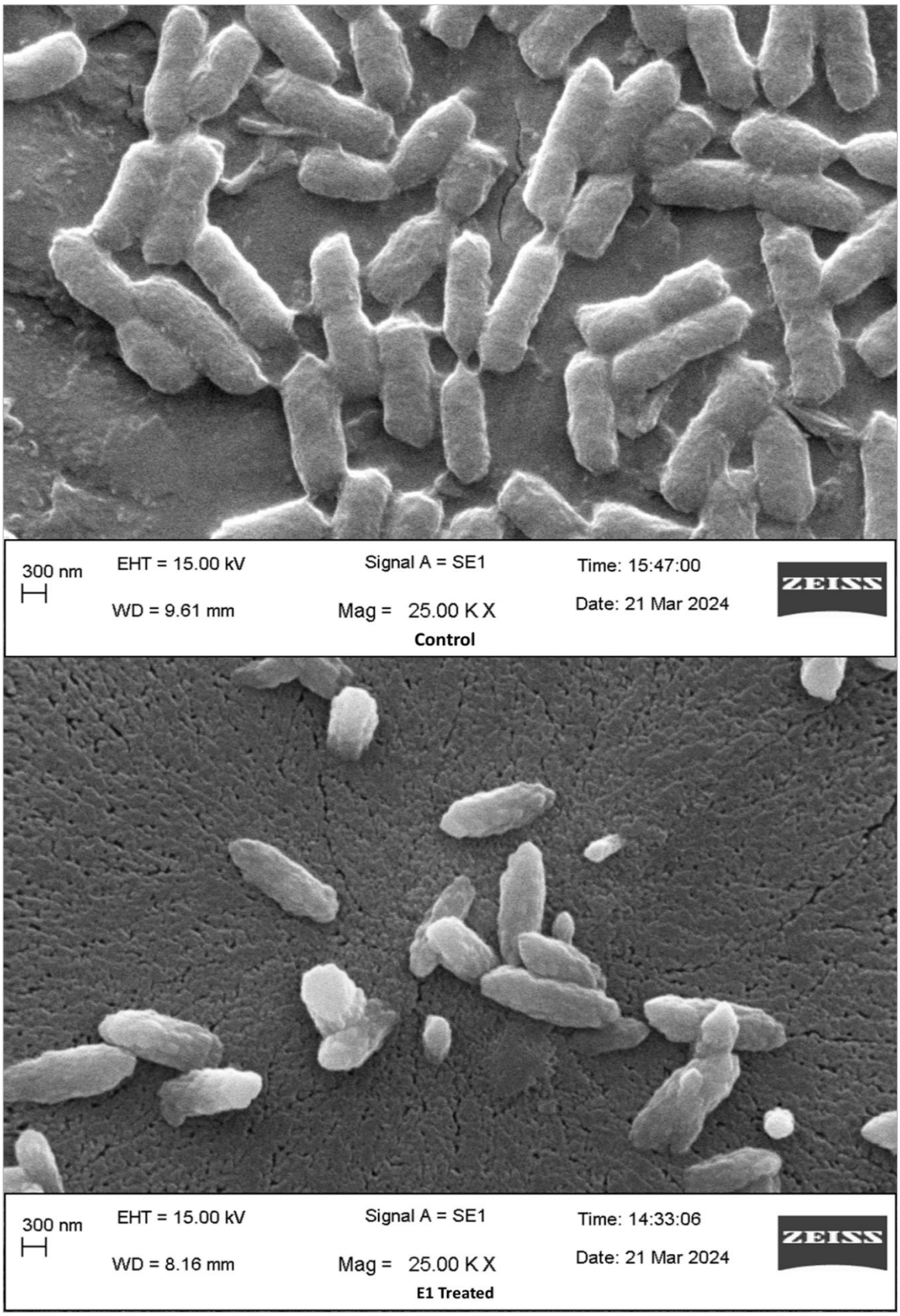
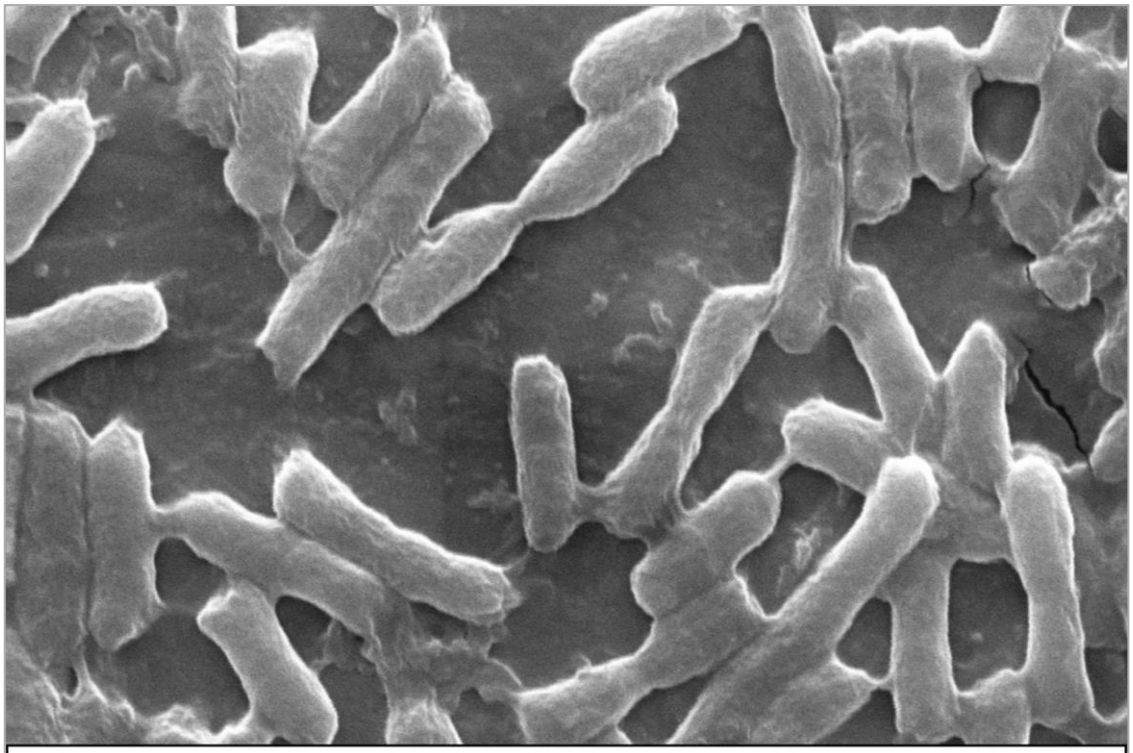


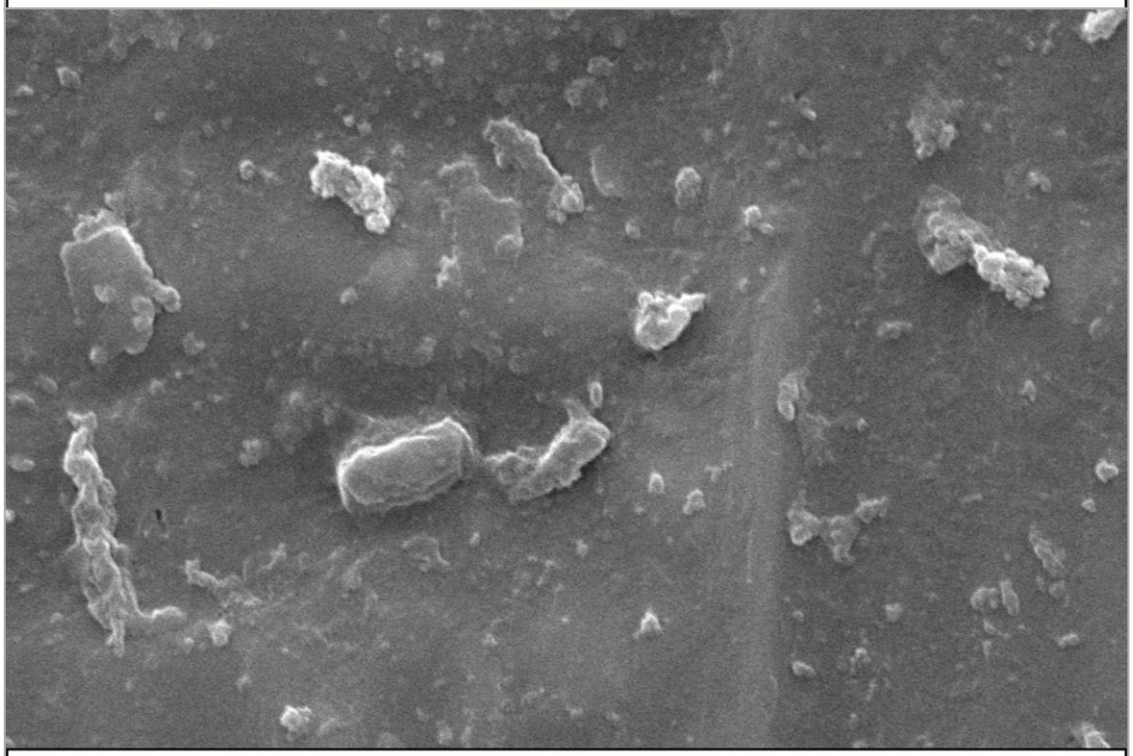
Figure 5.11.2 Bacterial Cells checked for Viability by FACS technique using Propidium Iodide (PI) and Thiazole orange dye; E1: treated with *Psidium* extract, E2: treated with *Scoparia* extract, C2: treated with piperidine, C3: treated with limonene.

The changes in the surface morphology of the bacteria upon treatment with the leads and crude extracts were analysed by using scanning electron microscopy (Figure 5.11.3).





E2 Treated



C2 Treated

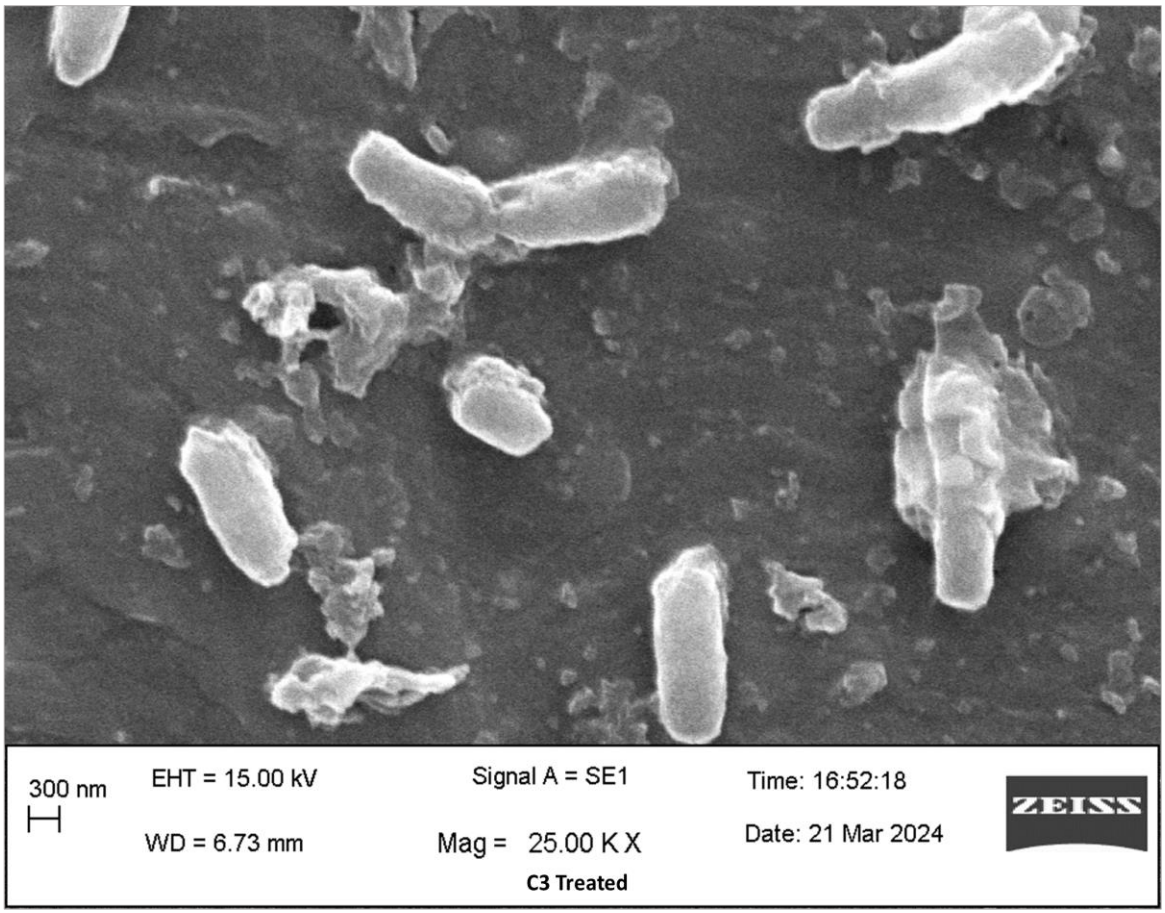


Figure 5.11.3 Scanning electron microscopy (SEM) images of Control: Untreated bacteria appeared as expected, with a smooth surface; E1: Treated with methanolic extract of *Psidium guajava*; E2: Treated with methanolic extract of *Scoparia dulcis*; both the extract treatments had bacteria with a withered and recessed surface; C2: Treated with Piperidine; C3: Treated with Limonene. Both the chemical treatments exhibited greater bacterial damage.

Upon investigation of the alterations in cell morphology and structure using SEM, the untreated bacteria showed up as complete rods with no signs of cell wall rupture or collapse (Figure 5.11.3). Cells exposed to herbal extracts, on the other hand, grew more twisted and coarser. Regular creases were harmed, and the surface developed rifts that resembled grooves. The cells were also observed to have undergone a significant change in shape, becoming shorter and wider (Figure 5.11.3**)**. This was in complete contrast to the surface structure of rod-shaped *Shigella flexneri* in untreated culture, where the bacteria remained

nearly intact exhibiting regular wrinkles at nanoscale resolution. Alternatively, after applying chemical leads, the cells demonstrated severe cell wall damage and leakage of internal components, indicating cell lysis (Figure 5.11.3). The observations recorded using SEM thus amply demonstrated the efficacy of the chosen herbal extracts and small molecule inhibitors in triggering membrane injuries, plasmolysis, and deformation of cells in the MDR *Shigella flexneri* strain.

5.12 Identification of Differentially Regulated Genes

Transcriptome sequencing generated reads were mapped to the reference genome of *Shigella flexneri* (SFMMGSG_23) and the gene counts were used as inputs to edgeR with exactTest for differential expression estimation. An adjusted p-value threshold <= 0.05 and log2 fold change of ∓1.0 was used for statistical estimation of gene expression (Figure 5.12.1).

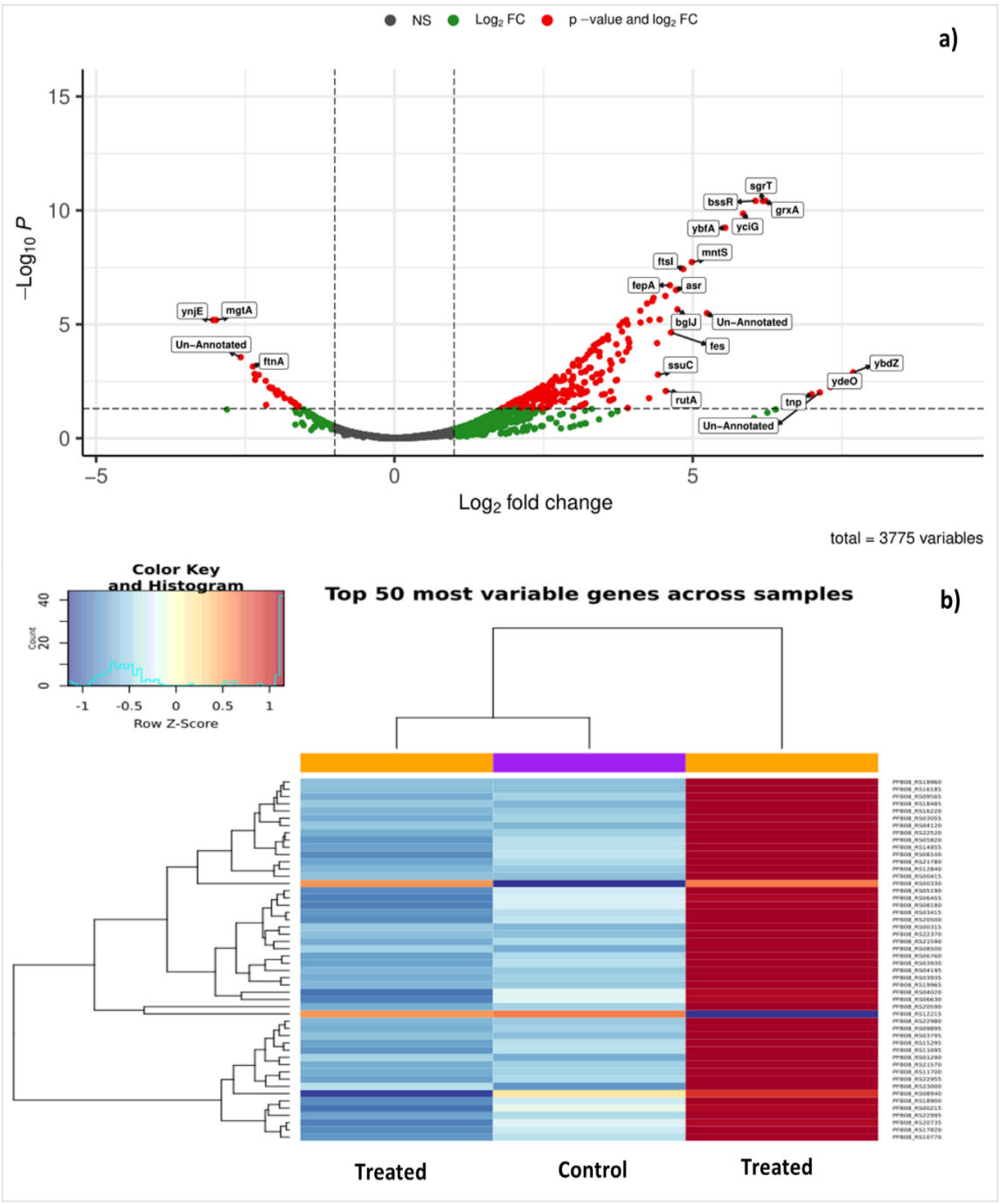


Figure 5.12.1 a) Volcano plot of Treated vs Control; b) Top 50 variable gene heatmap Treated vs Control.

The differentially regulated protein sequences were subjected to the COG classifier (Cluster of Orthologous Genes) v1.0.5 for functional annotation, classification, and analysis of each gene (Figure 5.12.2).

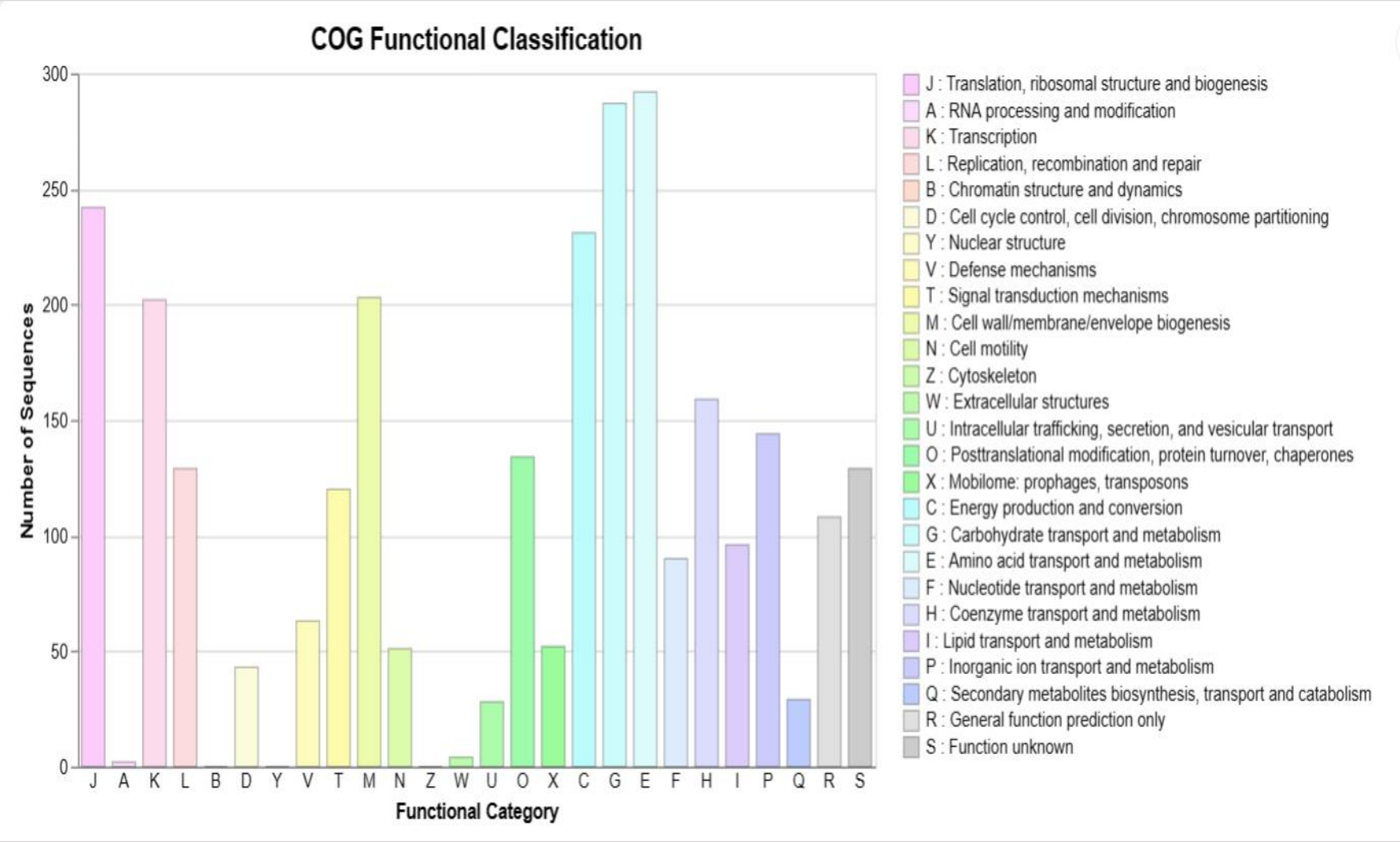


Figure 5.12.2 Functionally Enriched COG categories visualized in bar plot based on up and down regulated genes of Treated vs Control.

5.13 Differentially Regulated Genes and Their Homogeneity to Identified Targets

The differentially regulated genes were screened against prioritized targets based on sequence homogeneity, functional similarity and the presence of overlapping domains, which effectively lead us to 4 such genes (2 upregulated and 2 down regulated) that further established the correlation between the projected targets from core genome and the up/down regulated genes (Table 5.13.1).

Table 5.13.1 Differentially Regulated Genes and Their Homogeneity to Identified Targets.

Genes	Expression Status	Encoded Protein	Molecular Function	Biological Pathway
pdeG	Downregulated	PDEG	c-di-GMP phosphodiesterase	Biofilm Formation
nikC	Downregulated	NIKC	nickel transport system permease protein	ABC transporter
torR	Upregulated	TORR	two-component system OmpR family torCAD operon response regulator TorR	Two-component system
torY	Upregulated	TORY	trimethylamine-N-oxide reductase (cytochrome c) cytochrome	Microbial metabolism in diverse environments
pykF	Unchanged	PYKF	Pyruvate kinase I	Pyruvate metabolism

5.14 Validation of the differential gene expression using quantitative PCR

Quantitative polymerase chain reaction (qPCR) was used to determine the RNA load in the treated cell cultures in order to evaluate the differential expression of the selected genes more extensively (Figure 5.14.1). Nutrient broth was inoculated with enriched bacterial culture (1%) and treated with equal volume of herbal extracts and identified leads, separately. After an incubation period of 3 hours, cell culture supernatants were harvested

and the load of chosen ribonucleic acids (RNA) was quantified by reverse-transcription/real-time quantitative polymerase chain reaction (qPCR) using the Quant Studio™ system (Applied Biosystems) with pyruvate kinase (pykF) as internal control.

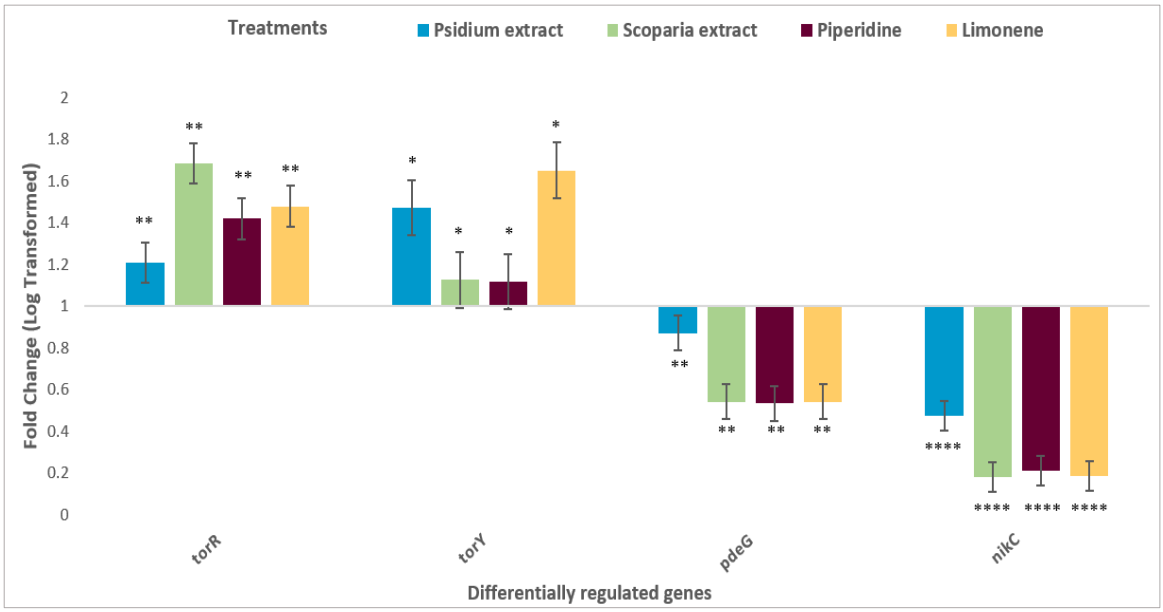


Figure 5.14.1 qPCR verification of differentially regulated genes. Graphs show relative fold change values of selected up regulated and down regulated genes in the naturally isolated *Shigella flexneri*, upon treatment with identified herbal extracts and chemical leads. An untreated bacterial culture of the same strain was used as control with a reference value of 1. Plotted with standard errors are the mean values obtained from trials conducted in triplicate. With respect to the unpaired t-test, the *, **, ***, and **** p-values denote the < 0.05, < 0.01, < 0.001, and <0.0001 levels of significance, in that order.

6. Discussion

Shigella is an important contributor of bacillary dysentery on a global scale, especially in countries that struggle with inadequate hygiene and sanitation. The rise of multidrug-resistant forms compounds the complexity of addressing *Shigella* infections, notably in areas that have restricted opportunities for alternative therapies and healthcare services. Therefore, with the intention to conquer this possible community health emergency, novel approaches to shigellosis prevention and therapy are imperative. Through the use of an integrated *in-silico* technique and experimental in-vitro assays to ensure further endorsement, this investigation attempted to identify and qualitatively characterize prospective therapeutic targets from *Shigella flexneri*. Given the state of affairs, wherein successive generations of antimicrobials have either been or will soon prove inept against *Shigella* spp., it is evident to assume that multiple concomitant initiatives are called for to keep shigellosis in check for a considerable amount of time. In an effort to manage shigellosis affected by antibiotic resistance, the final component of this research focussed in recognizing and validating novel antimicrobials derived from natural herbal products.

In this work, we successfully managed to isolate and identify a strain of *Shigella flexneri* from a natural reservoir (drainage system). The National Center for Biotechnology Information (NCBI) provides access to the entire genome sequence information of the isolated *Shigella flexneri* (SFMMGSG_23) under Accession CP123365. We further reported on the environmental isolate's whole genome sequencing (WGS) analysis, alongside the characterization of its resistance genes and associated antimicrobial phenotype. Approximately 4500 protein-coding genes were identified in the naturally isolated *Shigella flexneri* strain, of which 57 genes conferred antibiotic resistance. 15% of the resistance genes were found to code for proteins involved in target alteration, and around 50% of the resistance genes were identified to code for antibiotic efflux proteins.

Five different CRISPR arrays were obtained from the genome; however, none of the CRISPR arrays had CRISPR-associated genes (Cas genes) located nearby. In a recent study by Shmakov et al., 2020, such isolated CRISPR arrays were found to constitute 25% of the known CRISPR-Cas systems in complete bacterial genomes and were hypothesized to give rise to functional diversification of CRISPR-Cas systems by acting in conjunction with novel Cas proteins (Shmakov *et al*., 2020). Even though there are a number of theories as to how orphan arrays came to be i.e., i. Cas genes were eventually deleted from CRISPR-Cas loci; ii. isolated CRISPR arrays were inserted via mobile genetic elements; iii. de novo CRISPR arrays emerged from off-target spacer integration; the genomic segment for four of the five isolated CRISPR arrays that were discovered in our isolated bacteria was found to lie close to the predicted mobile genomic elements, which certainly supports the feasible dissemination of CRISPR arrays within MGEs. The draft genome's annotation using mobile-OG-db revealed 826 sequences, of which 484 were functional at integrating or cutting off from one genetic locus to another; 137 were linked to bacteriophage-related life cycle processes; 118 were proficient to set forth nucleic acid replication, recombination, and repair; 47 were suited for interbacterial dissemination; and 40 had effects that extend to the durability and protection of such elements. Analysing horizontally acquired genomic spots from whole genome sequences proves essential in deciding the scope of lateral gene transfer events, that constitute crucial variables in shaping up bacterial genome and evolution of diseases. The horizontally transferred genomic sections frequently adopt the shape of genomic islands, which are collections of acquired genes that are often grouped together in bacterial genomes with differing codon use, dinucleotide frequencies, and GC contents than the genes next to them. Nearly 11% of the isolated bacteria's genomic material were found to consist of an immense pool of such horizontally transmitted genomic possibilities. Given the isolated strain's remarkable proportion of genomic islands, the

genome is versatile enough to encode a wide variety of accessory genes for enhanced fitness, pathogenicity, resistance potential, ecological flexibility, etc. therefore rendering it to develop novel properties over a short duration. Such substantial ratio of horizontally acquired genomic areas are also anticipated to contribute to the subsequent acquisition and deletion of supplementary genes inside genomic islands, thus optimizing the genome while supporting adaptations in stressful circumstances. Three BGCs were discovered to be present in the genome, of which two were engaged in the production and internalization of siderophores of the chemical kinds catechols (enterobactin) and hydroxamic acids (aerobactin). These high affinity iron sequestering mechanisms give the pathogen the ability to thrive in iron-deficient settings within the host by allowing it to solubilize and remove ferric iron from insoluble complexes and host iron binding proteins (Sheldon *et al*., 2016)). In addition to aiding in the uptake of iron, siderophores have been linked to increased virulence, detoxification of reactive oxygen species (ROS), and resistance to oxidative stress. It has also been discovered that the genome contains putative thiazolyl-like biosynthetic genes, many of which are thought to function as signalling chemicals that affect bacterial phenotypes in ways that are important to humans and the environment (Bleich *et al*., 2015). Thus, the possibility to investigate if and how unique thiopeptides function as chemical cues in determining the distribution and phenotypes of medically significant infections is made possible by the finding of such likely thiopeptide synthesis genes in *Shigella*.

In an effort to investigate pathogen genomics within this species, we further opted to perform a pangenome analysis. By determining the magnitude of the pangenome, the proportions of the core and dispensable gene pools, the gain/loss of putative virulence factors, and the presence of genomic islands, this analysis allowed us to gain deeper insight into the species' current genomic diversity and pointed out the fundamental drivers that led

to the emergence of rapid escape variants following vaccination. A comparative pan-genomic analysis with 16 different strains of *Shigella flexneri* (1 isolated and 15 collected from public databases), demonstrated a startling ~64% genomic variability, common conservation of core genes in key metabolic activities, and enrichment of unique/accessory genes in defence and virulence mechanisms that help explain a large portion of the observed resistance to antibiotics and the quick dismissal of infection induced immunity against related *Shigella flexneri* serotypes. The pangenome curve shows that the total repertoire of gene families increases steadily with each new genome added. In the case of core genes, opposing tendencies were seen as their numbers steadily decreased as the number of genomes increased. The pangenome and core genome sizes did not reach a stationary phase, despite the fact that we could see a slowdown in the rate of increase of pangenome dimensions. Similar trends were observed in a study conducted by Parajuli et al., where pangenome analysis between 11 different *Shigella flexneri* strains showed a pangenome including 6,056 homologous groups, of which around half were shared by all strains (Parajuli *et al*., 2019). In another investigation, pangenome analysis was carried out for 1,246 methodically collected *Shigella* samples from seven nations in sub-Saharan Africa and South Asia, with the objective of studying genomic diversity in *Shigella* spp. in low- and middle-income countries (LMICs) (Bengtsson *et al*., 2022). Examination of the varying gene content revealed open pangenomes in all four species, suggesting that additional genome inclusions may still allow for the insertion of new genes, as evidenced from our findings. The fluid character of core genes, where inclusion/exclusion of core genes is determined by a complex interplay between the genetic background and environment, is highlighted by the dynamic genome composition, as demonstrated by shifting pan/core genome sizes. There was no clear pattern of correlation between the nation or period of origin of the strains and their phylogenetic distance, according to a thorough phylogenetic

inference based on concatenated core gene alignments and the presence/absence of pan or accessory genes. This implies that similar selection pressures acting under the same environmental conditions may have compelled the evolution of shared traits, which are then influenced by serotype switching and horizontal gene transfer events (Blount *et al*., 2018). 22 genes were identified by pathway analysis of the core genome, mapped to 2 antimicrobial resistance pathways. Multidrug efflux pumps, outer membrane permeability proteins, and two-component systems controlling gene expression in response to antibiotic stress are instances of specific resistance activities encoded by the core genomic elements. Latest research has demonstrated the role of TCSs, PmrA/B, and PhoP/Q, which was discovered as being a component of the core genome, in executing lipid A alterations that impart resistance to polymixins and a variety of cationic antimicrobial peptides. It has been further shown that the CpxA/R genes identified by KAAS annotation substantially influence β-lactam resistance in a variety of clinically significant enteropathogens. The emergence of resistance behaviour in bacterial populations can be attributed to a multitude of factors, however one of the most effective and fast-acting resistance mechanisms found in gram negative pathogens is the presence of chromosomally encoded efflux pumps (Gaurav *et al*., 2023). The existence of genes coding for outer membrane channel (TolC), inner membrane transporter (AcrB), and their periplasmic counterpart, membrane fusion protein (AcrA), that operate as a coordinated structure and comprise a tripartite Resistance Nodulation Division (RND) complex, is particularly significant in clinical terms among our collection of core genomic elements traced to various resistance pathways. A build-up of multi drug resistance in Gram-negative organisms has often been linked to an increase in the efflux activity of RND pumps (Colclough *et al*. 2020;). The enhanced expression of RND efflux pumps can either be a result of a random mutation originating within the transporter genes itself in response to environmental signals, or it could result from an

intricate interaction between transcription factors (Huang *et al*., 2022). New investigations have additionally shown that TCSs modulate the expression of efflux pumps, as illustrated by the contribution of CpxAR in inducing multidrug efflux pumps in a number of Enterobacteriaceae species (Dbeibo *et al*. 2018). Therefore, the coexistence of genes encoding multidrug efflux pumps and TCSs as a fundamental component of the core genome demonstrates an intricate framework of biological responses that support antimicrobial resistance in *Shigella flexneri*. The reality that such efflux pumps and two component systems are omnipresent and evolutionary preserved across various bacterial genera, and lack obvious human homologs also renders them as appealing options for future target discovery.

The *Shigella flexneri* genome expresses an overwhelming number of hypothetical proteins, in addition to core proteins, which makes it tricky to prioritize treatment targets until their functions are understood and logically categorized. An *in-silico* process was used to functionally annotate and classify a set of 432 hypothetical proteins. In addition to expanding the spectrum of potential therapeutic targets, the functional characterization of these unidentified proteins and the identification of their interaction partner facilitated a deeper comprehension of the pathogenic pathways.

Following functional annotation of both core and hypothetical proteins, the work undertook a gradual screening of these proteins following an integrated comparative and subtractive genomics approach, which eventually led us to recognize 10 potential targets (7 from core proteins and 3 from hypothetical proteins), disclosing the metabolic pathways they regulate and therefore expanding the repertoire of druggable genes beyond the ones that are already known. Out of the 10 projected targets, 4 are enzymes (PdeD, RimL, YdgJ, YfcV) catalysing critical steps in fundamental cellular processes like genetic information processing, bacterial motility and metabolism of principal biomacromolecules, 2 proteins

are ion- transporters (ModB, ThiP) with strong implications in maintenance of ion-homeostasis, and 4 are transcriptional regulators (YhaJ, RipA, PerA, ydiP).

PdeD is a GGDEF/EAL domain containing protein, that is documented to generate and degrade the second messenger cyclic di-GMP (c-di-GMP) and is reported to have implications in bacterial adaptation to host specific niche by influencing the adherence mechanism (Sarenko *et al*., 2017).

RimL is a chromosomally encoded acetyltransferase, which besides acetylating ribosomal proteins are also known to acetylate various toxic aminoacyl nucleotides, which are often produced as chemical intermediates after antibiotic administration. This detoxification process executed by RimL significantly contributes to bacterial fitness and might be regarded as a source of antibiotic resistance, thus making this protein an appealing target (Kazakov *et al*., 2014).

The ydgJ gene encodes an oxidoreductase that participates in oxidative function in bacteria. A number of findings suggested that ydgJ may be involved in cell metabolism or ROS detoxification. In 2016, the Bolt lab conducted a genetic screen to identify and characterize this putative oxidoreductase named YdgJ, that conferred resistance to the antibiotic nalidixic acid among *Escherichia coli* isolates (Cusin, 2020).

YfcV genes are evidenced to code for a variety of fimbrial or afimbrial adhesins, which enables attachment to the host epithelial cells and are thus regarded as notable virulence factors in bacterial pathogenesis (Nunes *et al*., 2022).

Molybdenum (Mo) serves as a crucial element in bacteria as it plays a role in the microbial breakdown of carbon, nitrogen, and sulfur and acts as a cofactor for several enzymes that catalyse different oxidation/reduction reactions. In order to produce molybdoenzymes, bacteria have to transport molybdate, convert it into the proper form by activation, and then

integrate it into the organic portion of the molybdenum cofactor. This investigation revealed the ModB protein as a target, which forms the transmembrane component of an ABC-type transport system, where it functions as a permease to facilitate the import of molybdenum into the cell (Xia *et al*., 2018).

Transmembrane carriers allow bacteria to employ exogenous thiamin and thiamin components to augment their own de novo production. In this work, the ThiP protein is suggested as a potential therapeutic target, which functions as a transmembrane thiamin channel, allowing thiamin to enter the cell (Liu *et al*., 2022).

Through concentration-dependent inhibition of the bacteria's type III secretion system, the host metabolite D-serine can significantly affect the course of an infection. In this work, the highly conserved core protein YhaJ is shortlisted as a potential pharmacological target. It regulates the uptake of D-serine from the environment, which in turn affects global gene expression and aids in the perception of D-serine levels. Thus, the protein serves as a layer of protection and is an adaptation of the core genome relevant to infections (Connolly *et al*., 2016).

Several studies have reported on the iron-dependent control of the citric acid cycle by the RipA protein, which is identified as a prospective therapeutic target in this work. In iron deficient circumstances, RipA inhibits the function of aconitase, which catalyses the stereospecific and reversible isomerization of citrate to isocitrate via cis-aconitate (Wennerhold *et al*., 2005). RipA's control of internal iron consumption thereby enhances iron acquisition processes and allows bacteria to survive under harsh settings.

The PerA protein, put forward as one of the therapeutic targets in this study, has been shown to act as an activator of many virulence genes in Enteropathogenic *E. coli* (EPEC) (Lara-Ochoa *et al*., 2021). EPEC typically adheres to the small intestinal epithelial cells,

by producing an attaching and effacing (A/E) lesion. The TTSS encoded within the nexus of enterocyte effacement (LEE) pathogenicity island, acts as the catalyst for the formation of the A/E lesions (Lara-Ochoa *et al*., 2021). The PerA protein, a member of the AraC/XylS family of transcriptional activators produced by the per (perABC) operon, has been confirmed to activate LEE gene expression, thus establishing itself as a key driver of the infectious process induced by normal EPEC strains (Lara-Ochoa *et al*., 2021). Therefore, PerA protein's fundamental importance in closely related *Shigella* homologs improves its suitability as a therapeutic target in *Shigella flexneri*.

While a number of antiquated medical texts, highlight the historical importance of plant-based natural products in the treatment of various illnesses, the majority of the information comes from countless human trial and error experiments conducted over hundreds of years with little to no understanding of the underlying bioactive constituents (Dzobo, 2022). In this investigation, we adopted a direct inoculation approach to test the antibacterial characteristics of 10 different herbal extracts against the isolated multidrug resistant *Shigella flexneri* strain, replicating the conventional disc diffusion experiment. In essence, we made an effort to subject these herbal extracts to the same stringent testing procedures that are usually employed to assess any possible antimicrobial agents obtained from pharmaceuticals. Subsequently a wider zone of inhibition on agar plates indicated that 20% of the screened herbal extracts (methanolic extracts of *Psidium guajava* and *Scoparia dulcis*) had in vitro antibacterial activity against the MDR-*Shigella flexneri* and were thus regarded as the best performing extracts. Our approach of combining the GC-HRMS platform together with high throughput virtual screening, could effectively capture and identify 3 promising small molecules (Piperidine, Limonene and Benzamide) from the 2 selected plants. All the three compounds were found to demonstrate binding energies in the range of -7 kcal/mol to -9 kcal/mol with the proposed therapeutic targets and were found to

obey all the parameters of Lipinski's rule. The favourable pharmacokinetic and pharmacodynamic metrices and spontaneous binding between the identified leads and proposed target candidates demonstrated by *in-silico* findings, were further legitimized by confirmatory invitro assays, viz. live and dead assay in conjunction with flow cytometry, which exhibited significant increase in percentage of dead cells upon treatment with selected crude extracts and lead compounds (Govindarajan *et al*., 2017). Additionally, changes in bacterial cellular morphology following treatment with the best functioning herbal extracts and specific chemical leads were assessed using scanning electron microscopy (SEM) analysis. Scanning electron microscopy images demonstrated that *Shigella flexneri* cells in the untreated control set had uniform and smooth membranes. On the other hand, the bacterial cells in the treatment group underwent major degeneration and intracellular content leakage as evidenced from the uneven and collapsed surfaces often with signs of severe cellular damage as also observed in similar works monitoring morphological changes in *Shigella* (Nickerson *et al*., 2017; Ahamed *et al*., 2023; Wang *et al*., 2022). These findings indicate that the mechanism of action of the plants and the constituents is possibly linked to alterations in bacterial membranes resulting in intracellular content leakage and eventual bacterial death as concluded from similar observations in the work of Li et al (Li *et al*., 2021). The possible mechanisms of action of the herbal extracts in curbing bacterial growth were investigated by RNA-sequencing in combination with bioinformatics analysis. Results showed that treatment with crude herbal extracts of *Psidium guajava* and *Scoparia dulcis* enhanced the activities of two component signal transduction systems that are preliminary sensors of environmental stress thus prompting the organism to cope up with changing conditions. RNA sequencing results also highlighted downregulation of genes related to adherence mechanism and transport systems, thus indicating the synergistic effect of membrane damage and oxidative stress

leading to bacterial growth inhibition and death (Chandrangsu *et al*., 2017; Wan *et al*., 2021). Limonene is a monoterpenoid compound and has been observed to exhibit antimicrobial activity against the gram-positive pathogen *Listeria monocytogenes* possibly working by manipulating membrane permeability (Han *et al*., 2019). Piperidine and its derivatives were also reported to have antimicrobial activity against *Fusarium verticilliodes*, *Candida utilus* and *Penicillium digitatium*, however no investigations have reported on the potential antibacterial activity of pieridine so far (Naicker *et al*., 2015). Thus, the effectiveness of these two phytocompounds are well in conformation with our observations and results, which indicates that these may be effective as important therapeutic interventions against multiple pathogenic microbes. The target candidates and potential leads nominated through our work can be investigated further experimentally, which may eventually open up new therapeutic strategies for effective treatment in the face of emerging antimicrobial resistance. Last but not least it is usually an ideal move when herbs can offer less expensive therapy options in settings with limited resources.

The possibility of the emergence of novel serotypes in the near future is strengthened by the potential evolvability of *Shigella* species, as evidenced by the notable differences in genomic content, the strain-specific essentiality of unique/accessory genes, and the inclusion of potent resistance mechanisms within the core genome (Bengtsson *et al*., 2022). This underscores the significance of tracking down genomic diversity in drug/vaccine design and AMR governance. We believe that the sequence of our isolated genome can be a valuable resource for upcoming research on comparative genomics and devising therapeutic interventions against shigellosis.

7. Future Prospect:

Our findings established the selective potential of *Psidium guajava* and *Scoparia dulcis* extracts, as a source of novel leads for treating infections with multi-drug resistant *Shigella* strains. The novel therapeutic targets and the potential leads proposed in this work, can be taken up by pharmaceutical giants for conducting intensive in-vivo investigations and stipulated clinical trials, passing which the compounds can be marketed as efficient *shigella* therapeutics. The efficacy of the extracts and the identified compounds can be further utilized for investigating anticancer effects using assays directed towards evaluation of invitro toxicity, such as MTT assay (Microculture Tetrazolium Assay); using suitable cell lines, along with investigations which enable identification of alterations in morphology, formation of acidic vesicular organelles (Acridine orange staining followed by fluorescence microscopy and apoptotic cell death (DAPI staining).

8. Summary of Work

Shigellosis is classified as a diarrheal illness that has a significant fatality rate, particularly in paediatric patients, older adults, and individuals with impaired immune systems. In light of the bacteria's expanding resistance patterns, the World Health Organization has recommended innovative therapeutic design to combat shigellosis in more recent times. Thus, finding new therapeutic targets and possible natural inhibitors of *Shigella flexneri* was the goal of this investigation. Using a comprehensive computational method that combined subtractive and comparative genomics, novel therapeutic targets from *Shigella flexneri* were recognized by analysing all the relevant genes and proteins across the entire genome. From the extensive diversity of metabolites discovered through spectrometric analysis of traditional medicinal herbs, promising inhibitors were additionally identified by executing ligand-based virtual screening and ADMET studies. Following multiple iterations of prioritization, 10 distinct proteins (PdeD, ModB, RimL, ThiP, YhaJ, YdgJ, YfcV, PerA, YdiP, RipA) were ultimately suggested as prospective therapeutic targets, accompanied by the 3 small molecule inhibitors (Piperidine, Limonene, Benzamide) that were most effective against them. The proteins and natural drug leads proposed through this investigation, have not been recognized as therapeutic targets or antimicrobials against *Shigella flexneri* thus far. In addition, this study also accurately characterized a novel multidrug resistant strain of *Shigella flexneri* isolated from urban wastewater of West Bengal, India. The identification of the mechanisms granting high resistance to antibiotics in *Shigella flexneri* strains apparently isolated from various places across the globe could potentially be facilitated by the genetic information and comparative genomic profile of the isolated MDR strain reported in this study. To end, we anticipate that our recommended targets for therapy and their natural inhibitors might lead to an exciting advance in the efficient treatment of shigellosis in the near future.

9. Data availability

Bioproject	Accession	Description
-	MW380613	*Shigella flexneri* strain MMG_01 16S ribosomal RNA gene, partial sequence.
PRJNA922074	CP123365	*Shigella flexneri* strain MMGSG_23 chromosome, complete genome.
PRJNA1135748	SRR29832218	RNA-Seq of untreated *Shigella flexneri*
PRJNA1135748	SRR29832216	RNA-Seq of *Shigella flexneri* treated with *Scoparia dulcis* extract
PRJNA1135748	SRR29832217	RNA-Seq of *Shigella flexneri* treated with *Psidium guajava* extract

10.References

1. Adeyi, OO, Baris, E, Jonas, OB et al. (2017). Drug-resistant infections: a threat to our economic future (Vol. 2): final report. Washington, DC: World Bank Group. (http://documents.worldbank.org/curated/en/323311493396993758/final-report

2. Ahamed, S. K. T., Rai, S., Guin, C., Jameela, R. M., Dam, S., Muthuirulandi Sethuvel, D. P., Balaji, V., & Giri, N. (2023). Characterizations of novel broad-spectrum lytic bacteriophages Sfin-2 and Sfin-6 infecting MDR Shigella spp. with their application on raw chicken to reduce the Shigella load. Frontiers in microbiology, 14, 1240570. https://doi.org/10.3389/fmicb.2023.1240570

3. Ait Ouakrim, D., Cassini, A., Cecchini, M., Plachouras, D., & North, J. (2020). The health and economic burden of antimicrobial resistance. In M. Anderson, M. Cecchini, & E. Mossialos (Eds.), Challenges to Tackling Antimicrobial Resistance: Economic and Policy Responses (pp. 23–44). chapter, Cambridge: Cambridge University Press.

4. Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H. K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., Faltyn, M., … McArthur, A. G. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic acids research, 48(D1), D517–D525. https://doi.org/10.1093/nar/gkz935

5. Anand, K., Sundaram, K. R., Lobo, J., & Kapoor, S. K. (1994). Are diarrheal incidence and malnutrition related in under five children? A longitudinal study in an area of poor sanitary conditions. Indian pediatrics, 31(8), 943–948.

6. Anani, H., Zgheib, R., Hasni, I., Raoult, D., & Fournier, P. E. (2020). Interest of bacterial pangenome analyses in clinical microbiology. Microbial pathogenesis, 149, 104275. https://doi.org/10.1016/j.micpath.2020.104275

7. Anderson, M., Sansonetti, P. J., & Marteyn, B. S. (2016). Shigella Diversity and Changing Landscape: Insights for the Twenty-First Century. Frontiers in cellular and infection microbiology, 6, 45. https://doi.org/10.3389/fcimb.2016.00045

8. Antonisamy, P., Duraipandiyan, V., Ignacimuthu, S., & Kim, J. (2015). Anti-Diarrhoeal Activity of Friedelin Isolated from Azima tetracantha Lam. in Wistar Rats. South Indian Journal of Biological Sciences, 1, 34-37.

9. Aprilio, K., & Wilar, G. (2021). Emergence of Ethnomedical COVID-19 Treatment: A Literature Review. Infection and drug resistance, 14, 4277–4289. https://doi.org/10.2147/IDR.S327986

10. Arya, H., Coumar, M.S. (2021). Chapter 4 - Lead identification and optimization, Editor(s): Tarun Kumar Bhatt, Surendra Nimesh. The Design & Development of Novel Drugs and Vaccines. Academic Press, Pages 31-63, ISBN 9780128214718. https://doi.org/10.1016/B978-0-12-821471-8.00004-0.

11. Asalone, K. C., Nelson, M. M., & Bracht, J. R. (2019). Novel Sequence Discovery by Subtractive Genomics. Journal of visualized experiments: JoVE, (143), 10.3791/58877. https://doi.org/10.3791/58877

12. Ashraf, B., Atiq, N., Khan, K., Wadood, A., & Uddin, R. (2022). Subtractive genomics profiling for potential drug targets identification against Moraxella catarrhalis. PloS one, 17(8), e0273252. https://doi.org/10.1371/journal.pone.0273252

13. Aslam, B., Wang, W., Arshad, M. I., Khurshid, M., Muzammil, S., Rasool, M. H., Nisar, M. A., Alvi, R. F., Aslam, M. A., Qamar, M. U., Salamat, M. K. F., & Baloch, Z. (2018). Antibiotic resistance: a rundown of a global crisis. Infection and drug resistance, 11, 1645–1658. https://doi.org/10.2147/IDR.S173867

14. Austin, D. & Hayford, T. (2021). Research and development in the pharmaceutical industry. CBO https://www.cbo.gov/publication/57126.

15. Avital, A., Maayan, C., & Goitein, K. J. (1982). Incidence of convulsions and encephalopathy in childhood Shigella infections. Survey of 117 hospitalized patients. Clinical pediatrics, 21(11), 645–648. https://doi.org/10.1177/000992288202101101

16. Baig, M. H., Ahmad, K., Rabbani, G., Danishuddin, M., & Choi, I. (2018). Computer Aided Drug Design and its Application to the Development of Potential Drugs for Neurodegenerative Disorders. Current neuropharmacology, 16(6), 740–748. https://doi.org/10.2174/1570159X15666171016163510

17. Baker, S., & Scott, T. A. (2023). Antimicrobial-resistant Shigella: where do we go next?. Nature reviews. Microbiology, 21(7), 409–410. https://doi.org/10.1038/s41579-023-00906-1

18. Bakheet, T. M., & Doig, A. J. (2010). Properties and identification of antibiotic drug targets. BMC bioinformatics, 11, 195. https://doi.org/10.1186/1471-2105-11-195

19. Balouiri, M., Sadiki, M., & Ibnsouda, S. K. (2016). Methods for in vitro evaluating antimicrobial activity: A review. Journal of pharmaceutical analysis, 6(2), 71–79. https://doi.org/10.1016/j.jpha.2015.11.005

20. Barcelos, M.P. et al. (2022). Lead Optimization in Drug Discovery. In: Taft, C.A., de Lazaro, S.R. (eds) Research Topics in Bioactivity, Environment and Energy. Engineering Materials. Springer, Cham. https://doi.org/10.1007/978-3-031-07622-0_19

21. Barh, D., Tiwari, S., Jain, N., Ali, A., Santos, A. R., Misra, A. N., … Kumar, A. (2010). In silico subtractive genomics for target identification in human bacterial pathogens. Drug Development Research, 72(2), 162–177. https://doi:10.1002/ddr.20413

22. Bartlett, J. G., Gilbert, D. N., & Spellberg, B. (2013). Seven ways to preserve the miracle of antibiotics. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America, 56(10), 1445–1450. https://doi.org/10.1093/cid/cit070

23. Bellido-Blasco, J. B., & Arnedo-Pena, A. (2011). Epidemiology of Infectious Diarrhea. Encyclopedia of Environmental Health, 659–671. https://doi.org/10.1016/B978-0-444-63951-6.00689-6

24. Bengtsson, R. J., Simpkin, A. J., Pulford, C. V., Low, R., Rasko, D. A., Rigden, D. J., Hall, N., Barry, E. M., Tennant, S. M., & Baker, K. S. (2022). Pathogenomic analyses of Shigella isolates inform factors limiting shigellosis prevention and control across LMICs. Nature microbiology, 7(2), 251–261. https://doi.org/10.1038/s41564-021-01054-z

25. Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2018). Environmental factors influencing the development and spread of antibiotic resistance. FEMS microbiology reviews, 42(1), fux053. https://doi.org/10.1093/femsre/fux053

26. Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics (Oxford, England), 27(3), 343–350. https://doi.org/10.1093/bioinformatics/btq662

27. Bennish, M. L., Azad, A. K., Rahman, O., & Phillips, R. E. (1990). Hypoglycemia during diarrhea in childhood. Prevalence, pathophysiology, and outcome. The New England journal of medicine, 322(19), 1357–1363. https://doi.org/10.1056/NEJM199005103221905

28. Bhattacharya, D., Bhattacharya, H., Thamizhmani, R., Sayi, D. S., Reesu, R., Anwesh, M., Kartick, C., Bharadwaj, A. P., Singhania, M., Sugunan, A. P., & Roy, S. (2014). Shigellosis in Bay of Bengal Islands, India: clinical and seasonal patterns, surveillance of antibiotic susceptibility patterns, and molecular characterization of multidrug-resistant Shigella strains isolated during a 6-year period from 2006 to 2011. European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology, 33(2), 157–170. https://doi.org/10.1007/s10096-013-1937-2

29. Bhutta, Z., & Syed, S. (2016). Diarrheal Diseases. Encyclopedia of Food and Health, 361–372. https://doi.org/10.1016/B978-0-12-384947-2.00223-3

30. Bialvaei, A. Z., Pourlak, T., Aghamali, M., Asgharzadeh, M., Gholizadeh, P., & Kafil, H. S. (2017). The Prevalence of CTX-M-15 Extended-spectrum β-Lactamases Among Salmonella spp. and Shigella spp. Isolated from three Iranian Hospitals. European journal of microbiology & immunology, 7(2), 133–137. https://doi.org/10.1556/1886.2017.00004

31. Black, R. E., Brown, K. H., Becker, S., Alim, A. R., & Huq, I. (1982). Longitudinal studies of infectious diseases and physical growth of children in rural Bangladesh. II. Incidence of diarrhea and association with known pathogens. American journal of epidemiology, 115(3), 315–324. https://doi.org/10.1093/oxfordjournals.aje.a113308

32. Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., Jha, P., Campbell, H., Walker, C. F., Cibulskis, R., Eisele, T., Liu, L., Mathers, C., & Child Health Epidemiology Reference Group of WHO and UNICEF (2010). Global, regional, and national causes of child mortality in 2008: a systematic analysis. Lancet (London, England), 375(9730), 1969–1987. https://doi.org/10.1016/S0140-6736(10)60549-1

33. Bleich, R., Watrous, J. D., Dorrestein, P. C., Bowers, A. A., & Shank, E. A. (2015). Thiopeptide antibiotics stimulate biofilm formation in Bacillus subtilis. Proceedings of the National Academy of Sciences of the United States of America, 112(10), 3086–3091. https://doi.org/10.1073/pnas.1414272112

34. Blin, K., Shaw, S., Augustijn, H. E., Reitz, Z. L., Biermann, F., Alanjary, M., Fetter, A., Terlouw, B. R., Metcalf, W. W., Helfrich, E. J. N., van Wezel, G. P., Medema, M. H., & Weber, T. (2023). antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic acids research, gkad344. Advance online publication. https://doi.org/10.1093/nar/gkad344

35. Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. Science (New York, N.Y.), 362(6415), eaam5979. https://doi.org/10.1126/science.aam5979

36. Bobate, S., Mahalle, S., Dafale, N.A., Bajaj, A. (2023). Emergence of environmental antibiotic resistance: Mechanism, monitoring and management. Environmental Advances, Volume 13(100409), ISSN 2666-7657. https://doi.org/10.1016/j.envadv.2023.100409.

37. Boithias, L., Choisy, M., Souliyaseng, N., Jourdren, M., Quet, F., Buisson, Y., Thammahacksa, C., Silvera, N., Latsachack, K., Sengtaheuanghoung, O., Pierret, A., Rochelle-Newall, E., Becerra, S., & Ribolzi, O. (2016). Hydrological Regime and Water Shortage as Drivers of the Seasonal Incidence of Diarrheal Diseases in a Tropical Montane Environment. PLoS neglected tropical diseases, 10(12), e0005195. https://doi.org/10.1371/journal.pntd.0005195

38. Breijyeh, Z., Jubeh, B., & Karaman, R. (2020). Resistance of Gram-Negative Bacteria to Current Antibacterial Agents and Approaches to Resolve It. Molecules (Basel, Switzerland), 25(6), 1340. https://doi.org/10.3390/molecules25061340

39. Brenner, D. J., Krieg, N. R., Staley, J. T., & Garrity, G. M. (Eds.). (2005). Bergey's Manual® of Systematic Bacteriology. https://doi.org/10.1007/0-387-28021-9

40. Brotcke Zumsteg, A., Goosmann, C., Brinkmann, V., Morona, R., & Zychlinsky, A. (2014). IcsA is a Shigella flexneri adhesin regulated by the type III secretion system and required for pathogenesis. Cell host & microbe, 15(4), 435–445. https://doi.org/10.1016/j.chom.2014.03.001

41. Brown, C. L., Mullet, J., Hindi, F., Stoll, J. E., Gupta, S., Choi, M., Keenum, I., Vikesland, P., Pruden, A., & Zhang, L. (2022). mobileOG-db: a Manually Curated Database of Protein Families Mediating the Life Cycle of Bacterial Mobile Genetic Elements. Applied and environmental microbiology, 88(18), e0099122. https://doi.org/10.1128/aem.00991-22

42. Chandran, U., Mehendale, N., Patil, S., Chaguturu, R., & Patwardhan, B. (2017). Network Pharmacology. Innovative Approaches in Drug Discovery, 127–164. https://doi.org/10.1016/B978-0-12-801814-9.00005-2

43. Chandrangsu, P., Rensing, C., & Helmann, J. D. (2017). Metal homeostasis and resistance in bacteria. Nature reviews. Microbiology, 15(6), 338–350. https://doi.org/10.1038/nrmicro.2017.15

44. Chao H. C. (2023). Zinc Deficiency and Therapeutic Value of Zinc Supplementation in Pediatric Gastrointestinal Diseases. Nutrients, 15(19), 4093. https://doi.org/10.3390/nu15194093

45. Chattopadhyay, D., Arunachalam, G., Mandal, A. B., & Bhattacharya, S. K. (2006). Dose-dependent therapeutic antiinfectives from ethnomedicines of bay islands. Chemotherapy, 52(3), 151–157. https://doi.org/10.1159/000092859

46. Chaudhari, N. M., Gupta, V. K., & Dutta, C. (2016). BPGA- an ultra-fast pan-genome analysis pipeline. Scientific reports, 6, 24373. https://doi.org/10.1038/srep24373

47. Chen, W. P., & Kuo, T. T. (1993). A simple and rapid method for the preparation of gram-negative bacterial genomic DNA. Nucleic acids research, 21(9), 2260. https://doi.org/10.1093/nar/21.9.2260

48. CLSI. Performance Standards for Antimicrobial Susceptibility Testing. 33rd ED. CLSI Supplement M100. Clinical and Laboratory Standards Institute, 2023.

49. Colclough, A. L., Alav, I., Whittle, E. E., Pugh, H. L., Darby, E. M., Legood, S. W., McNeil, H. E., & Blair, J. M. (2020). RND efflux pumps in Gram-negative bacteria; regulation, structure and role in antibiotic resistance. Future microbiology, 15, 143–157. https://doi.org/10.2217/fmb-2019-0235

50. Connolly, J. P., Gabrielsen, M., Goldstone, R. J., Grinter, R., Wang, D., Cogdell, R. J., Walker, D., Smith, D. G., & Roe, A. J. (2016). A Highly Conserved Bacterial D-Serine Uptake System Links Host Metabolism and Virulence. PLoS pathogens, 12(1), e1005359. https://doi.org/10.1371/journal.ppat.1005359

51. Costa-Silva, J., Domingues, D., & Lopes, F. M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. PloS one, 12(12), e0190152. https://doi.org/10.1371/journal.pone.0190152

52. Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B., Rocha, E. P. C., Vergnaud, G., Gautheret, D., & Pourcel, C. (2018). CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic acids research, 46(W1), W246–W251. https://doi.org/10.1093/nar/gky425

53. Cusin, Lola M.L. (2020) Characterisation of a putative oxidoreductase in Escherichia coli: YdgJ. MRes thesis, University of Nottingham.

54. Czerwińska-Główka, D., & Krukiewicz, K. (2021). Guidelines for a Morphometric Analysis of Prokaryotic and Eukaryotic Cells by Scanning Electron Microscopy. Cells, 10(12), 3304. https://doi.org/10.3390/cells10123304

55. da Silva Filho, A. C., Raittz, R. T., Guizelini, D., De Pierri, C. R., Augusto, D. W., Dos Santos-Weiss, I. C. R., & Marchaukoski, J. N. (2018). Comparative Analysis of Genomic Island Prediction Tools. Frontiers in genetics, 9, 619. https://doi.org/10.3389/fgene.2018.00619

56. Dadgostar P. (2019). Antimicrobial Resistance: Implications and Costs. Infection and drug resistance, 12, 3903–3910. https://doi.org/10.2147/IDR.S234610

57. Dahiya, P., & Purkayastha, S. (2012). Phytochemical Screening and Antimicrobial Activity of Some Medicinal Plants Against Multidrug Resistant Bacteria from Clinical Isolates. Indian journal of pharmaceutical sciences, 74(5), 443–450. https://doi.org/10.4103/0250-474X.108420

58. Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific reports, 7, 42717. https://doi.org/10.1038/srep42717

59. Davey, H., & Guyot, S. (2020). Estimation of Microbial Viability Using Flow Cytometry. Current protocols in cytometry, 93(1), e72. https://doi.org/10.1002/cpcy.72

60. Davies, J., & Davies, D. (2010). Origins and evolution of antibiotic resistance. Microbiology and molecular biology reviews: MMBR, 74(3), 417–433. https://doi.org/10.1128/MMBR.000

61. Dbeibo, L., van Rensburg, J. J., Smith, S. N., Fortney, K. R., Gangaiah, D., Gao, H., Marzoa, J., Liu, Y., Mobley, H. L. T., & Spinola, S. M. (2018). Evaluation of CpxRA as a

Therapeutic Target for Uropathogenic Escherichia coli Infections. Infection and immunity, 86(3), e00798-17. https://doi.org/10.1128/IAI.00798-17

62. Duan, J., Cai, X., Zhou, L., & Wang, J. (1997). Single-step method of total RNA isolation by sodium dodecyl sulfate/phenol extraction from cultured cells. Analytical biochemistry, 251(2), 291–292. https://doi.org/10.1006/abio.1997.2275

63. Dutta, S., Dutta, S., Dutta, P., Matsushita, S., Bhattacharya, S. K., & Yoshida, S. (2003). Shigella dysenteriae serotype 1, Kolkata, India. Emerging infectious diseases, 9(11), 1471–1474. https://doi.org/10.3201/eid0911.020652

64. Dzobo K. (2022). The Role of Natural Products as Sources of Therapeutic Agents for Innovative Drug Discovery. Comprehensive Pharmacology, 408–422. https://doi.org/10.1016/B978-0-12-820472-6.00041-4

65. Ebrahim G. J. (1991). Shigellosis. Journal of tropical pediatrics, 37(3), 98–99. https://doi.org/10.1093/tropej/37.3.98

66. Ekins, S., Mestres, J., & Testa, B. (2007). In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. British journal of pharmacology, 152(1), 9–20. https://doi.org/10.1038/sj.bjp.0707305

67. Emilien, G., Ponchon, M., Caldas, C., Isacson, O., & Maloteaux, J. M. (2000). Impact of genomics on drug discovery and clinical medicine. QJM: monthly journal of the Association of Physicians, 93(7), 391–423. https://doi.org/10.1093/qjmed/93.7.391

68. Erb, M., & Kliebenstein, D. J. (2020). Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. Plant physiology, 184(1), 39–52. https://doi.org/10.1104/pp.20.00433

69. European Centre for Disease Prevention and Control. Shigellosis. In: ECDC. Annual epidemiological report for 2020. Stockholm: ECDC; 2022. Farag TH, Faruque AS, Wu Y, et al. Housefly population density correlates with shigellosis among children in Mirzapur, Bangladesh: a time series analysis. PLoS Negl Trop Dis 2013; 7: e2280

70. Fang, J., Chen, Q., Wu, G. (2023). Genomics-based tools for drug discovery and development: From network maps to efficacy prediction. Journal of Holistic Integrative Pharmacy, 4 (3), 199-209. ISSN 2707-3688. https://doi.org/10.1016/j.jhip.2023.11.001.

71. Fatoba, A. J., Okpeku, M., & Adeleke, M. A. (2021). Subtractive Genomics Approach for Identification of Novel Therapeutic Drug Targets in Mycoplasma genitalium. Pathogens (Basel, Switzerland), 10(8), 921. https://doi.org/10.3390/pathogens10080921

72. Fatoba, A.J., Fatoba, D.O. & Babalola, S.O. (2022). Pangenome and subtractive genomic analysis of Clostridioides difficile reveals putative drug targets. J Proteins Proteom, 13, 247–256 https://doi.org/10.1007/s42485-022-00097-y

73. Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., Chang, H. Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G. L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., … Mitchell, A. L. (2017). InterPro in 2017-beyond protein family and domain annotations. Nucleic acids research, 45(D1), D190–D199. https://doi.org/10.1093/nar/gkw1107

74. Flores-Vargas, G., Bergsveinson, J., Lawrence, J. R., & Korber, D. R. (2021). Environmental Biofilms as Reservoirs for Antimicrobial Resistance. Frontiers in microbiology, 12, 766242. https://doi.org/10.3389/fmicb.2021.766242

75. Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C. and Jensen, L.J. (2013) 'STRING v9.1: protein-protein interaction net-works, with increased coverage and integration', Nucleic Acids Res., Vol. 41, Database issue, pp. D808–D815. https://doi.org/10.1093/nar/gks1094

76. Gaurav, A., Bakht, P., Saini, M., Pandey, S., & Pathania, R. (2023). Role of bacterial efflux pumps in antibiotic resistance, virulence, and strategies to discover novel efflux pump inhibitors. Microbiology (Reading, England), 169(5), 001333. https://doi.org/10.1099/mic.0.001333

77. GBD 2016 Causes of Death Collaborators (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet (London, England), 390(10100), 1151–1210. https://doi.org/10.1016/S0140-6736(17)32152-9

78. GBD 2016 Diarrhoeal Disease Collaborators (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. Infectious diseases, 18(11), 1211–1228. https://doi.org/10.1016/S1473-3099(18)30362-1

79. Gong, Y., Li, Y., Liu, X., Ma, Y., & Jiang, L. (2023). A review of the pangenome: how it affects our understanding of genomic variation, selection and breeding in domestic animals?. Journal of animal science and biotechnology, 14(1), 73. https://doi.org/10.1186/s40104-023-00860-1

80. Govindarajan, G., Kamaraj, R., Balakrishnan, K., Santhi, V. S., & Jebakumar, S. R. D. (2017). In-vitro assessment of antimicrobial properties and lymphocytotoxicity assay of benzoisochromanequinones polyketide from Streptomyces sp JRG-04. Microbial pathogenesis, 110, 117–127. https://doi.org/10.1016/j.micpath.2017.06.034

81. Graf, J., Meierhofer, R., Wegelin, M., & Mosler, H. J. (2008). Water disinfection and hygiene behaviour in an urban slum in Kenya: impact on childhood diarrhoea and influence of beliefs. International journal of environmental health research, 18(5), 335–355. https://doi.org/10.1080/09603120801966050

82. Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C. Y., Graham, M., Van Domselaar, G., & Stothard, P. (2023). Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic acids research, gkad326. Advance online publication. https://doi.org/10.1093/nar/gkad326

83. Guan, L., Yang, H., Cai, Y., Sun, L., Di, P., Li, W., Liu, G., & Tang, Y. (2018). ADMET-score - a comprehensive scoring function for evaluation of chemical drug-likeness. MedChemComm, 10(1), 148–157. https://doi.org/10.1039/c8md00472b

84. Han, Y., Sun, Z., & Chen, W. (2019). Antimicrobial Susceptibility and Antibacterial Mechanism of Limonene against Listeria monocytogenes. Molecules (Basel, Switzerland), 25(1), 33. https://doi.org/10.3390/molecules25010033

85. Harrell, J. E., & Cheng, S. X. (2018). Inability to reduce morbidity of diarrhea by ORS: can we design a better therapy?. Pediatric research, 83(3), 559–563. https://doi.org/10.1038/pr.2017.295

86. Harvey, A. L., Edrada-Ebel, R., & Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. Nature reviews. Drug discovery, 14(2), 111–129. https://doi.org/10.1038/nrd4510

87. Hedberg, C. W., Levine, W. C., White, K. E., Carlson, R. H., Winsor, D. K., Cameron, D. N., MacDonald, K. L., & Osterholm, M. T. (1992). An international foodborne outbreak of shigellosis associated with a commercial airline. JAMA, 268(22), 3208–3212.

88. Holt, K. E., Baker, S., Weill, F. X., Holmes, E. C., Kitchen, A., Yu, J., Sangal, V., Brown, D. J., Coia, J. E., Kim, D. W., Choi, S. Y., Kim, S. H., da Silveira, W. D., Pickard, D. J., Farrar, J. J., Parkhill, J., Dougan, G., & Thomson, N. R. (2012). Shigella sonnei genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. Nature genetics, 44(9), 1056–1059. https://doi.org/10.1038/ng.2369

89. Huang, L., Wu, C., Gao, H., Xu, C., Dai, M., Huang, L., Hao, H., Wang, X., & Cheng, G. (2022). Bacterial Multidrug Efflux Pumps at the Frontline of Antimicrobial Resistance: An Overview. Antibiotics (Basel, Switzerland), 11(4), 520. https://doi.org/10.3390/antibiotics11040520

90. Jabalia, N., Kumar, A., Kumar, V., Rani, R. (2021). In Silico Approach in Drug Design and Drug Discovery: An Update. In: Singh, S.K. (eds) Innovations and Implementations of Computer Aided Drug Discovery Strategies in Rational Drug Design. Springer, Singapore. https://doi.org/10.1007/978-981-15-8936-2_10

91. Jalilzadeh-Amin, G., & Maham, M. (2014). The application of 1,8-cineole, a terpenoid oxide present in medicinal plants, inhibits castor oil-induced diarrhea in rats. Pharmaceutical Biology, 53(4), 594–599. https://doi.org/10.3109/13880209.2014.935862

92. Jamal, M., Chaudhry, W. N., Hussain, T., Das, C. R., & Andleeb, S. (2015). Characterization of new Myoviridae bacteriophage WZ1 against multi-drug resistant (MDR) Shigella dysenteriae. Journal of basic microbiology, 55(4), 420–431. https://doi.org/10.1002/jobm.201400688

93. Jantan, I., Bukhari, S. N. A., Mohamed, M. A. S., Wai, L. K., & Mesaik, M. A. (2015). The Evolving Role of Natural Products from the Tropical Rainforests as a Replenishable Source of New Drug Leads. InTech. https://doi.org/0.5772/59603

94. Jiang, Z., Zhou, X., Li, R., Michal, J. J., Zhang, S., Dodson, M. V., Zhang, Z., & Harland, R. M. (2015). Whole transcriptome analysis with sequencing: methods, challenges and potential solutions. Cellular and molecular life sciences: CMLS, 72(18), 3425–3439. https://doi.org/10.1007/s00018-015-1934-y

95. Jin, Q., Yuan, Z., Xu, J., Wang, Y., Shen, Y., Lu, W., Wang, J., Liu, H., Yang, J., Yang, F., Zhang, X., Zhang, J., Yang, G., Wu, H., Qu, D., Dong, J., Sun, L., Xue, Y., Zhao, A., Gao, Y., … Yu, J. (2002). Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157. Nucleic acids research, 30(20), 4432–4441. https://doi.org/10.1093/nar/gkf566

96. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., & Madden, T. L. (2008). NCBI BLAST: a better web interface. Nucleic acids research, 36(Web Server issue), W5–W9. https://doi.org/10.1093/nar/gkn201

97. Jung, I. S., Kim, H. S., Park, H., & Lee, S. I. (2009). The clinical course of postinfectious irritable bowel syndrome: a five-year follow-up study. Journal of clinical gastroenterology, 43(6), 534–540. https://doi.org/10.1097/MCG.0b013e31818c87d7

98. Kabir, I., Butler, T., Underwood, L. E., & Rahman, M. M. (1992). Effects of a protein-rich diet during convalescence from shigellosis on catch-up growth, serum proteins, and insulin-like growth factor-I. Pediatric research, 32(6), 689–692. https://doi.org/10.1203/00006450-199212000-00014

99. Kareem, Z.H., Shareef, H.K., & Alkaim, A.F. (2018). Evaluation of antibacterial activity of Fe2O3 nanoparticles against Shigella dysenteriae. 10(8):1980–1982, 10(8):1980–1982.

100. Kazakov, T., Kuznedelov, K., Semenova, E., Mukhamedyarov, D., Datsenko, K. A., Metlitskaya, A., Vondenhoff, G. H., Tikhonov, A., Agarwal, V., Nair, S., Van Aerschot, A., & Severinov, K. (2014). The RimL transacetylase provides resistance to translation inhibitor microcin C. Journal of bacteriology, 196(19), 3377–3385. https://doi.org/10.1128/JB.01584-14

101. Kendall, M. E., Crim, S., Fullerton, K., Han, P. V., Cronquist, A. B., Shiferaw, B., Ingram, L. A., Rounds, J., Mintz, E. D., & Mahon, B. E. (2012). Travel-associated enteric infections diagnosed after return to the United States, Foodborne Diseases Active Surveillance Network (FoodNet), 2004-2009. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America, 54 Suppl 5, S480–S487. https://doi.org/10.1093/cid/cis052

102. Khalil, I. A., Troeger, C., Blacker, B. F., Rao, P. C., Brown, A., Atherly, D. E., Brewer, T. G., Engmann, C. M., Houpt, E. R., Kang, G., Kotloff, K. L., Levine, M. M., Luby, S. P.,

MacLennan, C. A., Pan, W. K., Pavlinac, P. B., Platts-Mills, J. A., Qadri, F., Riddle, M. S., Ryan, E. T., … Reiner, R. C., Jr (2018). Morbidity and mortality due to shigella and enterotoxigenic Escherichia coli diarrhoea: the Global Burden of Disease Study 1990-2016. The Lancet. Infectious diseases, 18(11), 1229–1240. https://doi.org/10.1016/S1473-3099(18)30475-4

103. Khan, T., Lawrence, A. J., Azad, I., Raza, S., Joshi, S., & Khan, A. R. (2019). Computational Drug Designing and Prediction of Important Parameters Using in silico Methods- A Review. Current computer-aided drug design, 15(5), 384–397. https://doi.org/10.2174/1573399815666190326120006

104. Khan, W. A., Dhar, U., Salam, M. A., Griffiths, J. K., Rand, W., & Bennish, M. L. (1999). Central nervous system manifestations of childhood shigellosis: prevalence, risk factors, and outcome. Pediatrics, 103(2), E18. https://doi.org/10.1542/peds.103.2.e18

105. Khan, W. A., Griffiths, J. K., & Bennish, M. L. (2013). Gastrointestinal and extra-intestinal manifestations of childhood shigellosis in a region where all four species of Shigella are endemic. PloS one, 8(5), e64097. https://doi.org/10.1371/journal.pone.0064097

106. Khan, W. A., Seas, C., Dhar, U., Salam, M. A., & Bennish, M. L. (1997). Treatment of shigellosis: V. Comparison of azithromycin and ciprofloxacin. A double-blind, randomized, controlled trial. Annals of internal medicine, 126(9), 697–703. https://doi.org/10.7326/0003-4819-126-9-199705010-00004

107. Khare, T., Anand, U., Dey, A., Assaraf, Y. G., Chen, Z. S., Liu, Z., & Kumar, V. (2021). Exploring Phytochemicals for Combating Antibiotic Resistance in Microbial Pathogens. Frontiers in pharmacology, 12, 720726. https://doi.org/10.3389/fphar.2021.720726

108. Killackey, S. A., Sorbara, M. T., & Girardin, S. E. (2016). Cellular Aspects of Shigella Pathogenesis: Focus on the Manipulation of Host Cell Processes. Frontiers in cellular and infection microbiology, 6, 38. https://doi.org/10.3389/fcimb.2016.00038

109. Kim, J. Y., Kim, S. H., Jeon, S. M., Park, M. S., Rhie, H. G., & Lee, B. K. (2008). Resistance to fluoroquinolones by the combination of target site mutations and enhanced expression of genes for efflux pumps in Shigella flexneri and Shigella sonnei strains isolated in Korea. Clinical microbiology and infection: the official publication of the

European Society of Clinical Microbiology and Infectious Diseases, 14(8), 760–765. https://doi.org/10.1111/j.1469-0691.2008.02033.x

110. Kim, Y., Gu, C., Kim, H. U., & Lee, S. Y. (2020). Current status of pan-genome analysis for pathogenic bacteria. Current opinion in biotechnology, 63, 54–62. https://doi.org/10.1016/j.copbio.2019.12.001

111. Kitts, P. A., Church, D. M., Thibaud-Nissen, F., Choi, J., Hem, V., Sapojnikov, V., Smith, R. G., Tatusova, T., Xiang, C., Zherikov, A., DiCuccio, M., Murphy, T. D., Pruitt, K. D., & Kimchi, A. (2016). Assembly: a resource for assembled genomes at NCBI. Nucleic acids research, 44(D1), D73–D80. https://doi.org/10.1093/nar/gkv1226

112. Klein, E. Y., Van Boeckel, T. P., Martinez, E. M., Pant, S., Gandra, S., Levin, S. A., Goossens, H., & Laxminarayan, R. (2018). Global increase and geographic convergence in antibiotic consumption between 2000 and 2015. Proceedings of the National Academy of Sciences of the United States of America, 115(15), E3463–E3470. https://doi.org/10.1073/pnas.1717295115

113. Kotloff, K. L., Riddle, M. S., Platts-Mills, J. A., Pavlinac, P., & Zaidi, A. K. M. (2018). Shigellosis. Lancet (London, England), 391(10122), 801–812. https://doi.org/10.1016/S0140-6736(17)33296-8

114. Kouse, A. B., Righetti, F., Kortmann, J., Narberhaus, F., & Murphy, E. R. (2013). RNA-mediated thermoregulation of iron-acquisition genes in Shigella dysenteriae and pathogenic Escherichia coli. PloS one, 8(5), e63781. https://doi.org/10.1371/journal.pone.0063781

115. Lampel, K. A., Formal, S. B., & Maurelli, A. T. (2018). A Brief History of Shigella. EcoSal Plus, 8(1), 10.1128/ecosalplus.ESP-0006-2017. https://doi.org/10.1128/ecosalplus.ESP-0006-2017

116. Lara-Ochoa, C., González-Lara, F., Romero-González, L. E., Jaramillo-Rodríguez, J. B., Vázquez-Arellano, S. I., Medrano-López, A., Cedillo-Ramírez, L., Martínez-Laguna, Y., Girón, J. A., Pérez-Rueda, E., Puente, J. L., & Ibarra, J. A. (2021). The transcriptional activator of the bfp operon in EPEC (PerA) interacts with the RNA polymerase alpha subunit. Scientific reports, 11(1), 8541. https://doi.org/10.1038/s41598-021-87586-0

117. Larsson, D. G. J., & Flach, C. F. (2022). Antibiotic resistance in the environment. Nature reviews. Microbiology, 20(5), 257–269. https://doi.org/10.1038/s41579-021-00649-x

118. Lavecchia, A., & Di Giovanni, C. (2013). Virtual screening strategies in drug discovery: a critical review. Current medicinal chemistry, 20(23), 2839–2860. https://doi.org/10.2174/09298673113209990001

119. Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A. K., Wertheim, H. F., Sumpradit, N., Vlieghe, E., Hara, G. L., Gould, I. M., Goossens, H., Greko, C., So, A. D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A. Q., Qamar, F. N., Mir, F., Kariuki, S., … Cars, O. (2013). Antibiotic resistance-the need for global solutions. The Lancet. Infectious diseases, 13(12), 1057–1098. https://doi.org/10.1016/S1473-3099(13)70318-9

120. Li, Y., Liu, F., Zhang, J., Liu, X., Xiao, P., Bai, H., Chen, S., Wang, D., Sung, S. H. P., Kwok, R. T. K., Shen, J., Zhu, K., & Tang, B. Z. (2021). Efficient Killing of Multidrug-Resistant Internalized Bacteria by AIEgens In Vivo. Advanced science (Weinheim, Baden-Wurttemberg, Germany), 8(9), 2001750. https://doi.org/10.1002/advs.202001750

121. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced drug delivery reviews, 46(1-3), 3–26. https://doi.org/10.1016/s0169-409x(00)00129-0

122. Liu, J., Platts-Mills, J. A., Juma, J., Kabir, F., Nkeze, J., Okoi, C., Operario, D. J., Uddin, J., Ahmed, S., Alonso, P. L., Antonio, M., Becker, S. M., Blackwelder, W. C., Breiman, R. F., Faruque, A. S., Fields, B., Gratz, J., Haque, R., Hossain, A., Hossain, M. J., … Houpt, E. R. (2016). Use of quantitative molecular diagnostic methods to identify causes of diarrhoea in children: a reanalysis of the GEMS case-control study. Lancet (London, England), 388(10051), 1291–1301. https://doi.org/10.1016/S0140-6736(16)31529-X

123. Liu, X., Wang, X., Sun, B., & Sun, L. (2022). The Involvement of Thiamine Uptake in the Virulence of Edwardsiella piscicida. Pathogens (Basel, Switzerland), 11(4), 464. https://doi.org/10.3390/pathogens11040464

124. Liu, Y., Cheng, Y., Yang, H., Hu, L., Cheng, J., Ye, Y., & Li, J. (2017). Characterization of Extended-Spectrum β-Lactamase Genes of Shigella flexneri Isolates With Fosfomycin

Resistance From Patients in China. Annals of laboratory medicine, 37(5), 415–419. https://doi.org/10.3343/alm.2017.37.5.415

125. Livio, S., Strockbine, N. A., Panchalingam, S., Tennant, S. M., Barry, E. M., Marohn, M. E., Antonio, M., Hossain, A., Mandomando, I., Ochieng, J. B., Oundo, J. O., Qureshi, S., Ramamurthy, T., Tamboura, B., Adegbola, R. A., Hossain, M. J., Saha, D., Sen, S., Faruque, A. S., Alonso, P. L., … Levine, M. M. (2014). Shigella isolates from the global enteric multicenter study inform vaccine development. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America, 59(7), 933–941. https://doi.org/10.1093/cid/ciu468

126. MacLennan, C. A., Grow, S., Ma, L. F., & Steele, A. D. (2022). The Shigella Vaccines Pipeline. Vaccines, 10(9), 1376. https://doi.org/10.3390/vaccines10091376

127. Mahapatra, A. D., Bhowmik, P., Banerjee, A., Das, A., Ojha, D., & Chattopadhyay, D. (2019). Ethnomedicinal Wisdom: An Approach for Antiviral Drug Development. New Look to Phytomedicine, 35–61. https://doi.org/10.1016/B978-0-12-814619-4.00003-3

128. Man, A. L., Prieto-Garcia, M. E., & Nicoletti, C. (2004). Improving M cell mediated transport across mucosal barriers: do certain bacteria hold the keys?. Immunology, 113(1), 15–22. https://doi.org/10.1111/j.1365-2567.2004.01964.x

129. Marteyn, B., West, N. P., Browning, D. F., Cole, J. A., Shaw, J. G., Palm, F., Mounier, J., Prévost, M. C., Sansonetti, P., & Tang, C. M. (2010). Modulation of Shigella virulence in response to available oxygen in vivo. Nature, 465(7296), 355–358. https://doi.org/10.1038/nature08970

130. McCrickard, L. S., Crim, S. M., Kim, S., & Bowen, A. (2018). Disparities in severe shigellosis among adults - Foodborne diseases active surveillance network, 2002-2014. BMC public health, 18(1), 221. https://doi.org/10.1186/s12889-018-5115-4

131. McGuffin, L. J., Adiyaman, R., Maghrabi, A. H. A., Shuid, A. N., Brackenridge, D. A., Nealon, J. O., & Philomina, L. S. (2019). IntFOLD: an integrated web resource for high performance protein structure and function prediction. Nucleic acids research, 47(W1), W408–W413. https://doi.org/10.1093/nar/gkz322

132. Miah, R., Mohd Aluwi, M. F. F. (2024). The importance of in-silico studies in drug discovery, Intelligent Pharmacy, ISSN 2949-866X. https://doi.org/10.1016/j.ipha.2024.01.010.

133. Michael L. Bennish, Sabeena Ahmed. 48 - Shigellosis, Editor(s): Edward T. Ryan, David R. Hill, Tom Solomon, Naomi E. Aronson, Timothy P. Endy. Hunter's Tropical Medicine and Emerging Infectious Diseases (Tenth Edition), Elsevier, 2020, Pages 492-499, ISBN 9780323555128. https://doi.org/10.1016/B978-0-323-55512-8.00048-X.

134. Million Death Study Collaborators, Bassani, D. G., Kumar, R., Awasthi, S., Morris, S. K., Paul, V. K., Shet, A., Ram, U., Gaffey, M. F., Black, R. E., & Jha, P. (2010). Causes of neonatal and child mortality in India: a nationally representative mortality survey. Lancet (London, England), 376(9755), 1853–1860. https://doi.org/10.1016/S0140-6736(10)61461-4

135. Mitchell, A., Chang, H-Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S. and Sangrador-Vegas, A., Scheremetjew, M., Rato, C., Yong, S-Y., Bateman, A., Punta, M., Attwood, T.K., Sigrist, C.J.A., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P., Letunic, I., Gough, J., Oates, M., Haft, D., Huang, H., Natale, D.A., Wu, C.H., Orengo, C., Sillitoe, I., Mi, H., Thomas, P.D., Finn, R.D. (2015) 'The interPro protein families database: the classification resource after 15 years', Nucleic Acids Res., Vol. 43, Database issue, pp. D213–21. https://doi.org/10.1093/nar/gku1243

136. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic acids research, 35(Web Server issue), W182–W185. https://doi.org/10.1093/nar/gkm321

137. Mukherjee, R., Dutta, D., Patra, M., Chatterjee, B., & Basu, T. (2019). Nanonized tetracycline cures deadly diarrheal disease 'shigellosis' in mice, caused by multidrug-resistant Shigella flexneri 2a bacterial infection. Nanomedicine: nanotechnology, biology, and medicine, 18, 402–413. https://doi.org/10.1016/j.nano.2018.11.004

138. Mukhopadhyay, S., Singh, M., Ghosh, M. M., Chakrabarti, S., & Ganguli, S. (2024). Comparative Genomics and Characterization of Shigella flexneri Isolated from Urban Wastewater. Microbes and environments, 39(2), 10.1264/jsme2.ME23105. https://doi.org/10.1264/jsme2.ME23105

139. Murail, S., de Vries, S. J., Rey, J., Moroy, G., & Tufféry, P. (2021). SeamDock: An Interactive and Collaborative Online Docking Resource to Assist Small Compound Molecular Docking. Frontiers in molecular biosciences, 8, 716466. https://doi.org/10.3389/fmolb.2021.716466

140. Muthuirulandi Sethuvel, D. P., Devanga Ragupathi, N. K., Anandan, S., & Veeraraghavan, B. (2017). Update on: Shigella new serogroups/serotypes and their antimicrobial resistance. Letters in applied microbiology, 64(1), 8–18. https://doi.org/10.1111/lam.12690

141. Naicker, L., Venugopala, K. N., Shode, F. & Odhav*, B. (2015). Antimicrobial and antioxidant activities of piperidine derivatives. African Journal of Pharmacy and Pharmacology, 9(31), 783-792.

142. Naveed, M., Chaudhry, Z., Ali, Z., Amjad, M., Zulfiqar, F. and Numan, A. (2018) Annotation and curation of hypothetical proteins: prioritizing targets for experimental study', Adv. Life Sci., Vol. 5, No. 3, pp.73–87.

143. NCBI Resource Coordinators (2016). Database resources of the National Center for Biotechnology Information. Nucleic acids research, 44(D1), D7–D19. https://doi.org/10.1093/nar/gkv1290

144. Nemeth, V., & Pfleghaar, N. (2022). Diarrhea. In StatPearls. StatPearls Publishing.

145. Nickerson, K. P., Chanin, R. B., Sistrunk, J. R., Rasko, D. A., Fink, P. J., Barry, E. M., Nataro, J. P., & Faherty, C. S. (2017). Analysis of Shigella flexneri Resistance, Biofilm Formation, and Transcriptional Profile in Response to Bile Salts. Infection and immunity, 85(6), e01067-16. https://doi.org/10.1128/IAI.01067-16

146. Nunes, P. H. S., Valiatti, T. B., Santos, A. C. M., Nascimento, J. A. D. S., Santos-Neto, J. F., Rocchetti, T. T., Yu, M. C. Z., Hofling-Lima, A. L., & Gomes, T. A. T. (2022). Evaluation of the Pathogenic Potential of Escherichia coli Strains Isolated from Eye Infections. Microorganisms, 10(6), 1084. https://doi.org/10.3390/microorganisms10061084

147. Oany, A. R., Mia, M., Pervin, T., Hasan, M. N., & Hirashima, A. (2018). Identification of potential drug targets and inhibitor of the pathogenic bacteria Shigella flexneri 2a

through the subtractive genomic approach. In silico pharmacology, 6(1), 11. https://doi.org/10.1007/s40203-018-0048-2

148. OECD (2018), Stemming the Superbug Tide: Just A Few Dollars More, OECD Health Policy Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264307599-en.

149. Ojha, D., Mukherjee, H., Ghosh, S., Bag, P., Mondal, S., Chandra, N. S., Mondal, K. C., Samanta, A., Chakrabarti, S., & Chattopadhyay, D. (2013). Evaluation of anti-infective potential of a tribal folklore Odina wodier Roxb against some selected microbes and herpes simplex virus associated with skin infection. Journal of applied microbiology, 115(6), 1317–1328. https://doi.org/10.1111/jam.12330

150. Operario, D. J., Platts-Mills, J. A., Nadan, S., Page, N., Seheri, M., Mphahlele, J., Praharaj, I., Kang, G., Araujo, I. T., Leite, J. P. G., Cowley, D., Thomas, S., Kirkwood, C. D., Dennis, F., Armah, G., Mwenda, J. M., Wijesinghe, P. R., Rey, G., Grabovac, V., Berejena, C., … Houpt, E. R. (2017). Etiology of Severe Acute Watery Diarrhea in Children in the Global Rotavirus Surveillance Network Using Quantitative Polymerase Chain Reaction. The Journal of infectious diseases, 216(2), 220–227. https://doi.org/10.1093/infdis/jix294

151. Parajuli, P., Deimel, L. P., & Verma, N. K. (2019). Genome Analysis of Shigella flexneri Serotype 3b Strain SFL1520 Reveals Significant Horizontal Gene Acquisitions Including a Multidrug Resistance Cassette. Genome biology and evolution, 11(3), 776–785. https://doi.org/10.1093/gbe/evz026

152. Parisot, M., Jolivet, A., Boukhari, R., & Carles, G. (2016). Shigellosis and Pregnancy in French Guiana: Obstetric and Neonatal Complications. The American journal of tropical medicine and hygiene, 95(1), 26–30. https://doi.org/10.4269/ajtmh.15-0669

153. Payne S. M. (2019). Laboratory Cultivation and Storage of Shigella. Current protocols in microbiology, 55(1), e93. https://doi.org/10.1002/cpmc.93

154. Pazhani, G. P., Ramamurthy, T., Mitra, U., Bhattacharya, S. K., & Niyogi, S. K. (2005). Species diversity and antimicrobial resistance of Shigella spp. isolated between 2001 and 2004 from hospitalized children with diarrhoea in Kolkata (Calcutta), India. Epidemiology and infection, 133(6), 1089–1095. https://doi.org/10.1017/S0950268805004498

155. Peng, J., Yang, J., & Jin, Q. (2009). The molecular evolutionary history of Shigella spp. and enteroinvasive Escherichia coli. Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases, 9(1), 147–152. https://doi.org/10.1016/j.meegid.2008.10.003

156. Penrod, N. M., Cowper-Sal-lari, R., & Moore, J. H. (2011). Systems genetics for drug target discovery. Trends in pharmacological sciences, 32(10), 623–630. https://doi.org/10.1016/j.tips.2011.07.002

157. Piddock L. J. (2006). Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. Clinical microbiology reviews, 19(2), 382–402. https://doi.org/10.1128/CMR.19.2.382-402.2006

158. Platts-Mills, J. A., Liu, J., Rogawski, E. T., Kabir, F., Lertsethtakarn, P., Siguas, M., Khan, S. S., Praharaj, I., Murei, A., Nshama, R., Mujaga, B., Havt, A., Maciel, I. A., McMurry, T. L., Operario, D. J., Taniuchi, M., Gratz, J., Stroup, S. E., Roberts, J. H., Kalam, A., … MAL-ED Network Investigators (2018). Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. The Lancet. Global health, 6(12), e1309–e1318. https://doi.org/10.1016/S2214-109X(18)30349-8

159. Pourakbari, B., Mamishi, S., Kohan, L., Sedighi, L., Mahmoudi, S., Fattahi, F., & Teymuri, M. (2012). Lethal toxic encephalopathy due to childhood shigellosis or Ekiri syndrome. Journal of microbiology, immunology, and infection = Wei mian yu gan ran za zhi, 45(2), 147–150. https://doi.org/10.1016/j.jmii.2011.09.005

160. Prestinaci, F., Pezzotti, P., & Pantosti, A. (2015). Antimicrobial resistance: a global multifaceted phenomenon. Pathogens and global health, 109(7), 309–318. https://doi.org/10.1179/2047773215Y.0000000030

161. Priya R. (2022). The role of traditional medicine in public health. Indian journal of public health, 66(2), 89–90. https://doi.org/10.4103/ijph.ijph_753_22

162. Puzari, M., Sharma, M., & Chetia, P. (2018). Emergence of antibiotic resistant Shigella species: A matter of concern. Journal of infection and public health, 11(4), 451–454. https://doi.org/10.1016/j.jiph.2017.09.025

163. Qureshi, H., Basheer, A., Sajjad, W., Faheem, M., & Babar Jamal, S. (2024). An integrated in-silico approach for drug target identification in human pathogen Shigella dysenteriae. PloS one, 19(5), e0303048. https://doi.org/10.1371/journal.pone.0303048

164. Rajput, A., Seif, Y., Choudhary, K. S., Dalldorf, C., Poudel, S., Monk, J. M., & Palsson, B. O. (2021). Pangenome Analytics Reveal Two-Component Systems as Conserved Targets in ESKAPEE Pathogens. mSystems, 6(1), e00981-20. https://doi.org/10.1128/mSystems.00981-20

165. Ranjbar, R., & Farahani, A. (2019). Shigella: Antibiotic-Resistance Mechanisms And New Horizons For Treatment. Infection and drug resistance, 12, 3137–3167. https://doi.org/10.2147/IDR.S219755

166. Rawat, P., Singh, P. K., & Kumar, V. (2017). Evidence based traditional anti-diarrheal medicinal plants and their phytocompounds. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 96, 1453–1464. https://doi.org/10.1016/j.biopha.2017.11.147

167. Reygaert W. C. (2018). An overview of the antimicrobial resistance mechanisms of bacteria. AIMS microbiology, 4(3), 482–501. https://doi.org/10.3934/microbiol.2018.3.482

168. Rizzuti, B. and Grande, F. (2020). Chapter 14 - Virtual screening in drug discovery: a precious tool for a still-demanding challenge. Editor(s): Angel L. Pey, Protein Homeostasis Diseases, Academic Press, Pages 309-327, ISBN 9780128191323. https://doi.org/10.1016/B978-0-12-819132-3.00014-2.

169. Rocha, D. J. P. G., Castro, T. L. P., Aguiar, E. R. G. R., & Pacheco, L. G. C. (2020). Gene Expression Analysis in Bacteria by RT-qPCR. Methods in molecular biology (Clifton, N.J.), 2065, 119–137. https://doi.org/10.1007/978-1-4939-9833-3_10

170. Rogawski, E. T., Liu, J., Platts-Mills, J. A., Kabir, F., Lertsethtakarn, P., Siguas, M., Khan, S. S., Praharaj, I., Murei, A., Nshama, R., Mujaga, B., Havt, A., Maciel, I. A., Operario, D. J., Taniuchi, M., Gratz, J., Stroup, S. E., Roberts, J. H., Kalam, A., Aziz, F., … MAL-ED Network Investigators (2018). Use of quantitative molecular diagnostic methods to investigate the effect of enteropathogen infections on linear growth in children in low-resource settings: longitudinal analysis of results from the MAL-ED cohort study.

The Lancet. Global health, 6(12), e1319–e1328. https://doi.org/10.1016/S2214-109X(18)30351-6

171. Rohmer, L., Jacobs, M. A., Brittnacher, M. J., Fong, C., Hayden, H. S., Hocquet, D., Weiss, E. J., Radey, M., Germani, Y., Talukder, K. A., Hager, A. J., Kemner, J. M., Sims-Day, E. H., Matamouros, S., Hager, K. R., & Miller, S. I. (2014). Genomic analysis of the emergence of 20th century epidemic dysentery. BMC genomics, 15(1), 355. https://doi.org/10.1186/1471-2164-15-355

172. Ruppé, E., Armand-Lefèvre, L., Estellat, C., El-Mniai, A., Boussadia, Y., Consigny, P. H., Girard, P. M., Vittecoq, D., Bouchaud, O., Pialoux, G., Esposito-Farèse, M., Coignard, B., Lucet, J. C., Andremont, A., & Matheron, S. (2014). Acquisition of carbapenemase-producing Enterobacteriaceae by healthy travellers to India, France, February 2012 to March 2013. Euro surveillance: bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin, 19(14), 20768. https://doi.org/10.2807/1560-7917.es2014.19.14.20768

173. Sadybekov, A. V., & Katritch, V. (2023). Computational approaches streamlining drug discovery. Nature, 616(7958), 673–685. https://doi.org/10.1038/s41586-023-05905-z

174. Sakaguchi, T., Köhler, H., Gu, X., McCormick, B. A., & Reinecker, H. C. (2002). Shigella flexneri regulates tight junction-associated proteins in human intestinal epithelial cells. Cellular microbiology, 4(6), 367–381. https://doi.org/10.1046/j.1462-5822.2002.00197.x

175. Salem, M. A., Perez de Souza, L., Serag, A., Fernie, A. R., Farag, M. A., Ezzat, S. M., & Alseekh, S. (2020). Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. Metabolites, 10(1), 37. https://doi.org/10.3390/metabo10010037

176. Sansonetti P. J. (2004). War and peace at mucosal surfaces. Nature reviews. Immunology, 4(12), 953–964. https://doi.org/10.1038/nri1499

177. Sansonetti, P. J., Tran Van Nhieu, G., & Egile, C. (1999). Rupture of the intestinal epithelial barrier and mucosal invasion by Shigella flexneri. Clinical infectious diseases: an official publication of the Infectious Diseases Society of America, 28(3), 466–475. https://doi.org/10.1086/515150

178. Sarenko, O., Klauck, G., Wilke, F. M., Pfiffer, V., Richter, A. M., Herbst, S., Kaever, V., & Hengge, R. (2017). More than Enzymes That Make or Break Cyclic Di-GMP-Local Signaling in the Interactome of GGDEF/EAL Domain Proteins of Escherichia coli. mBio, 8(5), e01639-17. https://doi.org/10.1128/mBio.01639-17

179. Schmitz R. (1985). Friedrich Wilhelm Sertürner and the discovery of morphine. Pharmacy in history, 27(2), 61–74.

180. Schnupf, P., & Sansonetti, P. J. (2019). Shigella Pathogenesis: New Insights through Advanced Methodologies. Microbiology spectrum, 7(2), 10.1128/microbiolspec.BAI-0023-2019. https://doi.org/10.1128/microbiolspec.BAI-0023-2019

181. Schroeder, G. N., & Hilbi, H. (2008). Molecular pathogenesis of Shigella spp.: controlling host cell signaling, invasion, and death by type III secretion. Clinical microbiology reviews, 21(1), 134–156. https://doi.org/10.1128/CMR.00032-07

182. Seemann T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics (Oxford, England), 30(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

183. Shad, A. A., & Shad, W. A. (2021). Shigella sonnei: virulence and antibiotic resistance. Archives of microbiology, 203(1), 45–58. https://doi.org/10.1007/s00203-020-02034-3

184. Shaker, B., Ahmad, S., Lee, J., Jung, C., & Na, D. (2021). In silico methods and tools for drug discovery. Computers in biology and medicine, 137, 104851. https://doi.org/10.1016/j.compbiomed.2021.104851

185. Sheldon, J. R., Laakso, H. A., & Heinrichs, D. E. (2016). Iron Acquisition Strategies of Bacterial Pathogens. Microbiology spectrum, 4(2), 10.1128/microbiolspec.VMBF-0010-2015. https://doi.org/10.1128/microbiolspec.VMBF-0010-2015

186. Shmakov, S. A., Utkina, I., Wolf, Y. I., Makarova, K. S., Severinov, K. V., & Koonin, E. V. (2020). CRISPR Arrays Away from cas Genes. The CRISPR journal, 3(6), 535–549. https://doi.org/10.1089/crispr.2020.0062

187. Shrestha, P., Karmacharya, J., Han, S. R., Park, H., & Oh, T. J. (2022). In silico analysis and a comparative genomics approach to predict pathogenic trehalase genes in the complete genome of Antarctica Shigella sp. PAMC28760. Virulence, 13(1), 1502–1514. https://doi.org/10.1080/21505594.2022.2117679

188. Spreafico, R., Soriaga, L. B., Grosse, J., Virgin, H. W., & Telenti, A. (2020). Advances in Genomics for Drug Development. Genes, 11(8), 942. https://doi.org/10.3390/genes11080942

189. Stagg, R. M., Tang, S. S., Carlin, N. I., Talukder, K. A., Cam, P. D., & Verma, N. K. (2009). A novel glucosyltransferase involved in O-antigen modification of Shigella flexneri serotype 1c. Journal of bacteriology, 191(21), 6612–6617. https://doi.org/10.1128/JB.00628-09

190. Sun, Q., Lan, R., Wang, J., Xia, S., Wang, Y., Wang, Y., Jin, D., Yu, B., Knirel, Y. A., & Xu, J. (2013). Identification and characterization of a novel Shigella flexneri serotype Yv in China. PloS one, 8(7), e70238. https://doi.org/10.1371/journal.pone.0070238

191. Taneja, N., & Mewara, A. (2016). Shigellosis: Epidemiology in India. The Indian journal of medical research, 143(5), 565–576. https://doi.org/10.4103/0971-5916.187104

192. Tansarli, G. S., Karageorgopoulos, D. E., Kapaskelis, A., & Falagas, M. E. (2013). Impact of antimicrobial multidrug resistance on inpatient care cost: an evaluation of the evidence. Expert review of anti-infective therapy, 11(3), 321–331. https://doi.org/10.1586/eri.13.4

193. Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., & Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. Nucleic acids research, 44(14), 6614–6624. https://doi.org/10.1093/nar/gkw569

194. Taylor, D. N., McKenzie, R., Durbin, A., Carpenter, C., Haake, R., & Bourgeois, A. L. (2008). Systemic pharmacokinetics of rifaximin in volunteers with shigellosis. Antimicrobial agents and chemotherapy, 52(3), 1179–1181. https://doi.org/10.1128/AAC.01108-07

195. Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., Angiuoli, S. V., Crabtree, J., Jones, A. L., Durkin, A. S., Deboy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J. D., Hauser, C. R., Sundaram, J. P., Nelson, W. C., Madupu, R., … Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome". Proceedings of the National Academy of Sciences of the United States of America, 102(39), 13950–13955. https://doi.org/10.1073/pnas.0506758102

196. The, H. C., Thanh, D. P., Holt, K. E., Thomson, N. R., & Baker, S. (2016). The genomic signatures of Shigella evolution, adaptation and geographical spread. Nature reviews. Microbiology, 14(4), 235–250. https://doi.org/10.1038/nrmicro.2016.10

197. Thompson, C. N., Duy, P. T., & Baker, S. (2015). The Rising Dominance of Shigella sonnei: An Intercontinental Shift in the Etiology of Bacillary Dysentery. PLoS neglected tropical diseases, 9(6), e0003708. https://doi.org/10.1371/journal.pntd.0003708

198. Tickell, K. D., Brander, R. L., Atlas, H. E., Pernica, J. M., Walson, J. L., & Pavlinac, P. B. (2017). Identification and management of Shigella infection in children with diarrhoea: a systematic review and meta-analysis. The Lancet. Global health, 5(12), e1235–e1248. https://doi.org/10.1016/S2214-109X(17)30392-3

199. Tyagi, A.K., & Malik, A. (2011). Antimicrobial potential and chemical composition of Mentha piperita oil in liquid and vapour phase against food spoiling microorganisms. Food Control, 22, 1707-1714. https://doi.org/10.1016/j.foodcont.2011.04.002

200. Utturkar, S., Dassanayake, A., Nagaraju, S., & Brown, S. D. (2020). Bacterial Differential Expression Analysis Methods. Methods in molecular biology (Clifton, N.J.), 2096, 89–112. https://doi.org/10.1007/978-1-0716-0195-2_8

201. van der Ploeg CA, Vinas MR, Terragno R, et al. (2010) Laboratory protocol: “Serotyping of Shigella spp”, 1–24

202. van Herpen, T. W., Cankar, K., Nogueira, M., Bosch, D., Bouwmeester, H. J., & Beekwilder, J. (2010). Nicotiana benthamiana as a production platform for artemisinin precursors. PloS one, 5(12), e14222. https://doi.org/10.1371/journal.pone.0014222

203. Verdu, E. F., & Riddle, M. S. (2012). Chronic gastrointestinal consequences of acute infectious diarrhea: evolving concepts in epidemiology and pathogenesis. The American journal of gastroenterology, 107(7), 981–989. https://doi.org/10.1038/ajg.2012.65

204. Verma, S., Pathak, R.K. (2022). Chapter 16 - Discovery and optimization of lead molecules in drug designing, Editor(s): Dev Bukhsh Singh, Rajesh Kumar Pathak. Bioinformatics. Academic Press, Pages 253-267, ISBN 9780323897754. https://doi.org/10.1016/B978-0-323-89775-4.00004-3.

205. Vernikos, G. S. (2020). A Review of Pangenome Tools and Recent Studies. In H. Tettelin (Eds.) et. al., The Pangenome: Diversity, Dynamics and Evolution of Genomes. (pp. 89–112). Springer.

206. Vernikos, G. S., & Parkhill, J. (2006). Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the Salmonella pathogenicity islands. Bioinformatics (Oxford, England), 22(18), 2196–2203. https://doi.org/10.1093/bioinformatics/btl369

207. Victora, C. G., Bryce, J., Fontaine, O., & Monasch, R. (2000). Reducing deaths from diarrhoea through oral rehydration therapy. Bulletin of the World Health Organization, 78(10), 1246–1255.

208. Wan, X., Yang, J., Ahmed, W., Liu, Q., Wang, Y., Wei, L., & Ji, G. (2021). Functional analysis of pde gene and its role in the pathogenesis of Xanthomonas oryzae pv. oryzicola. Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases, 94, 105008. https://doi.org/10.1016/j.meegid.2021.105008

209. Wang, F., Xiao, Y., Lu, Y., Deng, Z. Y., Deng, X. Y., & Lin, L. B. (2022). Bacteriophage Lytic Enzyme P9ly as an Alternative Antibacterial Agent Against Antibiotic-Resistant Shigella dysenteriae and Staphylococcus aureus. Frontiers in microbiology, 13, 821989. https://doi.org/10.3389/fmicb.2022.821989

210. Wang, L. P., Zhou, S. X., Wang, X., Lu, Q. B., Shi, L. S., Ren, X., Zhang, H. Y., Wang, Y. F., Lin, S. H., Zhang, C. H., Geng, M. J., Zhang, X. A., Li, J., Zhao, S. W., Yi, Z. G., Chen, X., Yang, Z. S., Meng, L., Wang, X. H., Liu, Y. L., … Chinese Centers for Disease Control and Prevention (CDC) Etiology of Diarrhea Surveillance Study Team (2021). Etiological, epidemiological, and clinical features of acute diarrhea in China. Nature communications, 12(1), 2464. https://doi.org/10.1038/s41467-021-22551-z

211. Wang, L., Zhu, Z., Qian, H., Li, Y., Chen, Y., Ma, P., & Gu, B. (2019). Comparative genome analysis of 15 clinical Shigellaflexneri strains regarding virulence and antibiotic resistance. AIMS microbiology, 5(3), 205–222. https://doi.org/10.3934/microbiol.2019.3.205

212. Watson, J. L., Sanchez-Garrido, J., Goddard, P. J., Torraca, V., Mostowy, S., Shenoy, A. R., & Clements, A. (2019). Shigella sonnei O-Antigen Inhibits Internalization, Vacuole

Escape, and Inflammasome Activation. mBio, 10(6), e02654-19. https://doi.org/10.1128/mBio.02654-19

213. Wennerhold, J., Krug, A., & Bott, M. (2005). The AraC-type regulator RipA represses aconitase and other iron proteins from Corynebacterium under iron limitation and is itself repressed by DtxR. The Journal of biological chemistry, 280(49), 40500–40508. https://doi.org/10.1074/jbc.M508693200

214. West, N. P., Sansonetti, P., Mounier, J., Exley, R. M., Parsot, C., Guadagnini, S., Prévost, M. C., Prochnicka-Chalufour, A., Delepierre, M., Tanguy, M., & Tang, C. M. (2005). Optimization of virulence functions through glucosylation of Shigella LPS. Science (New York, N.Y.), 307(5713), 1313–1317. https://doi.org/10.1126/science.1108472

215. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., 3rd, Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein science: a publication of the Protein Society, 27(1), 293–315. https://doi.org/10.1002/pro.3330

216. Wolde, D., Tilahun, G. A., Kotiso, K. S., Medhin, G., & Eguale, T. (2022). The Burden of Diarrheal Diseases and Its Associated Factors among Under-Five Children in Welkite Town: A Community Based Cross-Sectional Study. International journal of public health, 67, 1604960. https://doi.org/10.3389/ijph.2022.1604960

217. Wong, C. H., Siah, K. W., & Lo, A. W. (2019). Estimation of clinical trial success rates and related parameters. Biostatistics (Oxford, England), 20(2), 273–286. https://doi.org/10.1093/biostatistics/kxx069

218. World Health Organization (2017). WHO Publishes List of Bacteria for Which New Antibiotics Are Urgently Needed. Available at: https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

219. Xia, Z., Lei, L., Zhang, H. Y., & Wei, H. L. (2018). Characterization of the ModABC Molybdate Transport System of Pseudomonas putida in Nicotine Degradation. Frontiers in microbiology, 9, 3030. https://doi.org/10.3389/fmicb.2018.03030

220. Xiang, M., Cao, Y., Fan, W., Chen, L., & Mo, Y. (2012). Computer-aided drug design: lead discovery and optimization. Combinatorial chemistry & high throughput screening, 15(4), 328–337. https://doi.org/10.2174/138620712799361825

221. Yang, C., Xiang, Y., & Qiu, S. (2023). Resistance in Enteric Shigella and nontyphoidal Salmonella: emerging concepts. Current opinion in infectious diseases, 36(5), 360–365. https://doi.org/10.1097/QCO.0000000000000960

222. Yang, H., Duan, G., Zhu, J., Lv, R., Xi, Y., Zhang, W., Fan, Q., & Zhang, M. (2008). The AcrAB-TolC pump is involved in multidrug resistance in clinical Shigella flexneri isolates. Microbial drug resistance (Larchmont, N.Y.), 14(4), 245–249. https://doi.org/10.1089/mdr.2008.0847

223. Yi, Z. B., Yan Yu, Liang, Y. Z., & Bao Zeng (2007). Evaluation of the antimicrobial mode of berberine by LC/ESI-MS combined with principal component analysis. Journal of pharmaceutical and biomedical analysis, 44(1), 301–304. https://doi.org/10.1016/j.jpba.2007.02.018

224. Yousafi, Q., Kanwal, S., Rashid, H., Khan, M. S., Saleem, S., & Aslam, M. (2019). In silico structural and functional characterization and phylogenetic study of alkaline phosphatase in bacterium, Rhizobium leguminosarum (Frank 1879). Computational biology and chemistry, 83, 107142. https://doi.org/10.1016/j.compbiolchem.2019.107142

225. Yu, C. S., Chen, Y. C., Lu, C. H., & Hwang, J. K. (2006). Prediction of protein subcellular localization. Proteins: Structure, Function and Genetics, 64(3), 643-651. https://doi.org/10.1002/prot.21018

226. Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S. C., Ester, M., Foster, L. J., & Brinkman, F. S. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics (Oxford, England), 26(13), 1608–1615. https://doi.org/10.1093/bioinformatics/btq249

227. Yuan, H., Ma, Q., Ye, L., & Piao, G. (2016). The Traditional Medicine and Modern Medicine from Natural Products. Molecules (Basel, Switzerland), 21(5), 559. https://doi.org/10.3390/molecules21050559

228. Zaman, S. B., Hussain, M. A., Nye, R., Mehta, V., Mamun, K. T., & Hossain, N. (2017). A Review on Antibiotic Resistance: Alarm Bells are Ringing. Cureus, 9(6), e1403. https://doi.org/10.7759/cureus.1403

229. Zhang, N., Lan, R., Sun, Q., Wang, J., Wang, Y., Zhang, J., Yu, D., Hu, W., Hu, S., Dai, H., Du, P., Wang, H., & Xu, J. (2014). Genomic portrait of the evolution and epidemic spread of a recently emerged multidrug-resistant Shigella flexneri clone in China. Journal of clinical microbiology, 52(4), 1119–1126. https://doi.org/10.1128/JCM.02669-13

230. Zhang, X., Wu, F., Yang, N., Zhan, X., Liao, J., Mai, S., & Huang, Z. (2022). In silico Methods for Identification of Potential Therapeutic Targets. Interdisciplinary sciences, computational life sciences, 14(2), 285–310. https://doi.org/10.1007/s12539-021-00491-y

231. Zhang, Y., Zhang, L., Du, M., Yi, H., Guo, C., Tuo, Y., Han, X., Li, J., Zhang, L., & Yang, L. (2011). Antimicrobial activity against Shigella sonnei and probiotic properties of wild lactobacilli from fermented food. Microbiological research, 167(1), 27–31. https://doi.org/10.1016/j.micres.2011.02.006

232. Zhu, Z., Wang, L., Qian, H., Gu, F., Li, Y., Zhang, H., Chen, Y., Shi, J., Ma, P., Bao, C., & Gu, B. (2021). Comparative genome analysis of 12 Shigella sonnei strains: virulence, resistance, and their interactions. International microbiology: the official journal of the Spanish Society for Microbiology, 24(1), 83–91. https://doi.org/10.1007/s10123-020-00145-x

11. List of Publications

Peer Reviewed Publications

- Mukhopadhyay, S., Singh, M., Ghosh, M. M., Chakrabarti, S., & Ganguli, S. (2024). Comparative Genomics and Characterization of *Shigella flexneri* Isolated from Urban Wastewater. Microbes and environments, 39(2), 10.1264/jsme2.ME23105. https://doi.org/10.1264/jsme2.ME23105

- Mukhopadhyay, S., Ganguli, S., & Chakrabarti, S. (2022). Exploring the functions and interactions of undeciphered proteins from *Shigella flexneri*. International Journal of Computational Biology and Drug Design, 15(1), 60-75. https://doi.org/10.1504/IJCBDD.2022.124772

- Mukhopadhyay, S., Ganguli, S., & Chakrabarti, S. (2022). Insights Into The Structure And Dynamics Of *Shigella* Invasion Proteins For Use As Potential Drug Targets. Journal of Environment and Sociobiology, 37-42.

- Mukhopadhyay, S., Ganguli, S., & Chakrabarti, S. (2020). Functional Annotation of Pathogenesis Proteins in *Shigella flexneri* using Comparative Genomics. Journal of Environment and Sociobiology, 65-78.

- Mukhopadhyay, S., Ganguli, S., & Chakrabarti, S. (2020). *Shigella* pathogenesis: molecular and computational insights. AIMS Molecular Science, 7(2), 99-121. https://www.aimspress.com/article/doi/10.3934/molsci.2020007

Book Chapters

- Mukhopadhyay, S., Karmakar, R., Chakrabarti, S., Ghosh, M.M., Ganguli, S. (2024). Evaluating the Impact of Climate Change on Antimicrobial Resistance and Rise in Dysentery Using Next Generation Sequencing Based Approaches. In: Gupta, J., Verma, A. (eds) Microbiology-2.0 Update for a Sustainable Future. Springer, Singapore. https://doi.org/10.1007/978-981-99-9617-9_17

- Dhar GA., Karmakar R., Mukhopadhyay S., Basu S., Ghosh MM., Ganguli S. “Insights into Natural Product based Drug Discovery using a Systems Biology Approach- Communicated in a book entitled: Potential Bioactive ingredients for healthcare & wellness Industry: Recent Advances and Future Perspectives. Publishers: Springer Nature (in press)

Abstract Publications in Seminars/ Symposiums

- Evaluating The Antimicrobial and Cytotoxic Activities of Ethnomedicinal Plants to Propose New Drug Candidates Against Multidrug Resistant *Shigella*. International Symposium on Biotechnology. 12th & 13th October, 2023. Post Graduate and Research Department of Biotechnology. St. Xavier's College, (Autonomous) Kolkata.

- Virtual Screening of Herbal Product Library to Propose New Chemical Entities Effective Against *Shigella* Associated Diarrhoea. Presented At: 1st International Conference on Drug Discovery and Development for Infectious Diseases: Cutting Edge Research and Challenges 3rd and 4th March, 2023 Organized By: Eminent College of Pharmaceutical Technology.

- A Structural Bioinformatic Approach to Evaluate *Shigella* Invasion Proteins as Emerging Therapeutic Targets. International Conference on Climate Change: Global Cooperation. St. Xavier's College, (Autonomous) Kolkata. 26th and 27th August, 2022 (Awarded 3rd prize).

- Pangenome Analysis of *Shigella flexneri* Strains to Explore the Potential of Core Genes as Promising Therapeutic Targets. International Seminar on Recent Trends in Microbiology. Department Of Microbiology, Vijaygarh Jyotish Ray College, Kolkata. 18th May, 2022.

- Insights Into the Structure and Dynamics of *Shigella* Invasion Proteins for Use as Potential Drug Targets. 'When Science Meets Life' - A Symposium of The Society of Biological Chemists (I), Kolkata Chapter. Sister Nivedita University, West Bengal. 9th -10th April, 2022
- Mukhopadhyay S, Ganguli S, Chakrabarti S. Functional Annotation of Pathogenesis Proteins in *Shigella flexneri* using Comparative Genomics. National Seminar on Water Conservation and Harvesting Focusing Biodiversity Issues and Management. Social Environmental and Biological Association (Seba). Jadavpur University, Kolkata. 8th February, 2020.

- Impact Of Global Climate Change on Shigellosis and New Horizons for Treatment. International Symposium on Environment and Climate Crises. Department Of Environmental Studies, St. Xavier's College, (Autonomous) Kolkata. 20th

December, 2019.

- Mukhopadhyay S, Vishal V, Ganguli S, Chakrabarti S. Comparative Genomics to Identify Potential Drug Targets in *Shigella*. National Seminar on “Frontiers in Biological Sciences: Chapter III”. Department of Microbiology, St. Xavier’s College, (Autonomous) Kolkata. 21st and 22nd September, 2019.