

Article

Therapeutic Target Identification and Inhibitor Screening against Riboflavin Synthase of Colorectal Cancer Associated *Fusobacterium nucleatum*

Norah A. Alturki ¹, Mutaib M. Mashraqi ², Khurshid Jalal ³, Kanwal Khan ⁴, Zarrin Basharat ^{5,*}
and Ahmad Alzamami ^{6,*}

- ¹ Clinical Laboratory Science Department, College of Applied Medical Sciences, King Saud University, Riyadh 11433, Saudi Arabia
 - ² Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Najran University, Najran 61441, Saudi Arabia
 - ³ HEJ Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan
 - ⁴ Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan
 - ⁵ Jamil-ur-Rahman Center for Genome Research, Dr. Panjwani Center for Molecular Medicine and Drug Research, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan
 - ⁶ Clinical Laboratory Science Department, College of Applied Medical Sciences, Shaqra University, Al-Quwayyah 11961, Saudi Arabia
- * Correspondence: zarrin.iui@gmail.com (Z.B.); aalzamami@su.edu.sa (A.A.)



Citation: Alturki, N.A.; Mashraqi, M.M.; Jalal, K.; Khan, K.; Basharat, Z.; Alzamami, A. Therapeutic Target Identification and Inhibitor Screening against Riboflavin Synthase of Colorectal Cancer Associated *Fusobacterium nucleatum*. *Cancers* **2022**, *14*, 6260. <https://doi.org/10.3390/cancers14246260>

Academic Editor: Ikuko Kato

Received: 4 November 2022

Accepted: 14 December 2022

Published: 19 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: More and more studies are suggesting the role of microbes in several diseases in addition to the germline and environmental factors. *F. nucleatum* is recently being associated with colorectal cancer and here, we aimed to identify important drug targets from the core genome of colorectal cancer associated *F. nucleatum* through bioinformatics approach. We used one drug target for further analysis and obtained natural product inhibitors against it. Finally, we validated inhibition stability through dynamics simulation approach. We are hopeful that this study could benefit researchers working on colorectal cancer, its microbiome and cure.

Abstract: Colorectal cancer (CRC) ranks third among all cancers in terms of prevalence. There is growing evidence that gut microbiota has a role in the development of colorectal cancer. *Fusobacterium nucleatum* is overrepresented in the gastrointestinal tract and tumor microenvironment of patients with CRC. This suggests the role of *F. nucleatum* as a potential risk factor in the development of CRC. Hence, we aimed to explore whole genomes of *F. nucleatum* strains related to CRC to predict potential therapeutic markers through a pan-genome integrated subtractive genomics approach. In the current study, we identified 538 proteins as essential for *F. nucleatum* survival, 209 non-homologous to a human host, and 12 as drug targets. Eventually, riboflavin synthase (RiS) was selected as a therapeutic target for further processing. Three different inhibitor libraries of lead-like natural products, i.e., cyanobactins ($n = 237$), streptomycins ($n = 607$), and marine bacterial secondary metabolites ($n = 1226$) were screened against it. After the structure-based study, three compounds, i.e., CMNPD3609 (-7.63) > Malyngamide V (-7.03) > ZINC06804365 (-7.01) were prioritized as potential inhibitors of *F. nucleatum*. Additionally, the stability and flexibility of these compounds bound to RiS were determined via a molecular dynamics simulation of 50 ns. Results revealed the stability of these compounds within the binding pocket, after 5 ns. ADMET profiling showed compounds as drug-like, non-permeable to the blood brain barrier, non-toxic, and HIA permeable. Pan-genomics mediated drug target identification and the virtual screening of inhibitors is the preliminary step towards inhibition of this pathogenic oncobacterium and we suggest mouse model experiments to validate our findings.

Keywords: pan-genome; *F. nucleatum*; colorectal cancer; riboflavin synthase; virtual screening; computer aided drug design; natural products

1. Introduction

Colorectal cancer (CRC) is the world's third most common malignant neoplasm and the fourth leading cause of cancer mortality, with a five-year survival rate of about 65% [1]. Statistically, the mortality rate of CRC has decreased in places with high level of healthcare resources, while it has increased in areas with low levels of medical resources. Several variables contribute to the development of colorectal cancer, including hereditary and environmental factors such as diet and lifestyle [2–4]. The progression from precancerous adenomatous polyps to adenocarcinoma generally takes about 10 years [5], and is considered to be the consequence of host mutations that accumulate over time. Improved methods for detecting microorganisms and new insights into the human microbiome have led to a new perspective on several diseases, including CRC [6–8]. According to recent research, this cancer is linked with altered gut microbiota [9] and aided by the anaerobic gram-negative bacteria *F. nucleatum* [4,10–12]. It is often linked to advanced disease, chemoresistance, distant metastases, and a poor prognosis [13–15]. Additional studies have demonstrated that *F. nucleatum* promotes carcinogenesis, changes lymphocyte infiltration of tumor-infiltrating lymphocytes, suppresses NK cell and tumor-infiltrating T cell death, and enhances resistance to chemotherapy in colon cancer [7,15,16]. *Fusobacteria* linked to CRC develops in the oral microbiome and reaches the colon through hematogenous rather than gastrointestinal routes. It has been found in approximately 10–90% of CRC tissue samples, with a higher prevalence in the distal colon than the proximal one. Baik et al. [17] have identified IgA antibodies against the amyloid adhesin factor FadA of this bacterium in the colorectal neoplasia. Cavallucci and colleagues [18] reported that *F. nucleatum* binds Gal-GalNac disaccharide and CECAM-1 on the CRC tumor stem cells, leading to the activation of a MAP kinase. MAP kinase has an implied role in the CRC, therefore, this gives *F. nucleatum* a strong linkage with CRC progression [19,20]. Gao et al. [21] have shown that *F. nucleatum* asserts a positive impact on PD-L1 (an immune frontier point) by modifying an oncogene IFIT1. This causes immunosuppression and tumor progression in the CRC [21]. Lacourse et al. [22] have also shown that the chemotherapeutic agent 5-fluorouracil is an inhibitor of this oncomicrobe but if *E. coli* modifies this compound, then the growth of tumor and this bacterium remains upward. This demonstrates the importance of targeting bacteria alongside the tumor in CRC. Chen et al. [23] have explored CRC therapy options by targeting microbiota, particularly *F. nucleatum*, as infected tumors have shown a significant upsurge in the nitroreductase activity of this bacteria [23]. They instructed an enzyme assembly to dissolve and dismantle tumors through this approach in a mouse model and observed successful results.

In Saudi Arabia (SA), the age-standardized incidence rates (ASIR) and age-standardization mortality rate (ASMR) for all cancers are 88.7 and 43.3, respectively, per 100,000 people. After breast cancer (ASIR-27.3, ASMR-7.5), colorectal cancer (ASIR-13.1, ASMR-6.3) is the second most frequent cancer in Saudi Arabia. According to the GLOBOCON 2018 report, the highest ASMR and ASIR were discovered in Qatar and Oman, respectively, while the greatest ASIR was recorded in the UAE, Saudi Arabia, and Kuwait, according to ICAR, WHO, and GLOBOCON 2018 [14,24,25]. To cope with this disease burden, new therapeutic biomarkers and drug targets are required. New targets may have useful mechanisms [26] that can be harnessed to render the bacteria inactive when bound by drug. Next-generation sequencing (NGS) technology has made it simpler to acquire genetic data of microbial species [8]. Previously, only a small number of genome sequences were available, but now the complete genetic repertoire of organisms may be explored [8,27]. One of the best strategies in the drug development pipeline is to identify therapeutic targets in pathogens using a bioinformatics approach. This is a swift method. Aside from being expensive and time-consuming, conventional drug development in wet laboratories suffers from other drawbacks. An organism's whole genetic repertoire cannot be discerned by a single genomic sequence, therefore the advancement of big data and data mining-based

informatics approaches has a huge advantage [25,28,29]. Pan-genomic studies, which account for the whole set of genes across all strains within a genus, have become possible because of recent algorithmic and computational power developments [30]. Understanding the evolution of a strain requires knowledge of its whole genome sequence, and information of conserved fractions will also enable more specific targets for therapy and inhibitor development. Recently, this has been implemented against gram-negative bacteria such as *Brachypodium hybridum* [31], *Klebsiella pneumoniae* [32], and *Shigella* sp. [28,33]. Subtractive genome analysis is an effective method of finding the gene or group of genes responsible for a certain characteristic in an organism [34]. It helps discover the systemic variety in a given organism by analyzing its full genomic array.

Therefore, to address the challenge of processing data from the wide range of *F. nucleatum* genomes associated with CRC, we did a pan-genomic analysis and examined a large number of *F. nucleatum* pathovar isolates linked with CRC ($n = 14$) to characterize the core and accessory-genome subsets and prioritize therapeutic targets. To our knowledge, this is the first pan-genome mediated study on numerous *F. nucleatum* species linked with CRC, as well as the use of core gene data to investigate therapeutic targets in them. Furthermore, three different libraries of natural products were used to identify potent inhibitors that may be helpful to combat CRC-linked *Fusobacterium* sp.

2. Methodology

2.1. Data Retrieval

An extensive literature survey was performed to retrieve CRC-related *F. nucleatum* genomes for pan-genome analysis. Fourteen *F. nucleatum* genomes were used in this study (Table 1), derived either from the colon cancer site, tissue, or gut of the ailing person. Sub-species included *animalis*, *polymorphum*, and *vincentii*. The Universal Protein Resource (UniProt) database was used to download the human proteome. The significance of therapeutic targets was examined using the Database of Essential Genes (DEG) [35] and the Cluster of Essential Genes (CEG) [36]. The most recent 2022–3 edition of the DrugBank database [37] was used to assess the potential for the druggability of the proteins. Additionally, for the structure-based virtual screening, the cyanobactins ($n = 237$), streptomycins ($n = 607$), and marine bacterial secondary metabolites ($n = 1226$) were utilized.

Table 1. *F. nucleatum* genomes used in this study, either derived from colon cancer site, tissue, or gut of the ailing person.

Serial No.	Name	Biosample Accession	Genome Size (Mbp)	Isolation Source	Coding DNA Sequences	Reference
1.	<i>F. nucleatum</i> Fn146CP	SAMN20819806	2.082	Tissue	1949	[38]
2.	<i>F. nucleatum</i> Fn10-CTX3	SAMN20819805	2.101	Tissue	2026	[38]
3.	<i>F. nucleatum</i> Fn3760T	SAMN20819808	2.299	Tissue	2234	[38]
4.	<i>F. nucleatum</i> Fn173CP	SAMN20819807	2.121	Tissue	2041	[38]
5.	<i>F. nucleatum</i> FnS043-1	SAMN20819807	2.288	Tissue	2245	[38]
6.	<i>F. nucleatum</i> subsp. <i>animalis</i> strain THCT5A4	SAMN20819807	2.491	Gut	2360	[39]
7.	<i>F. nucleatum</i> subsp. <i>polymorphum</i> strain THCT15E1	SAMN18042967	2.526	Gut	2405	[39]
8.	<i>F. nucleatum</i> subsp. <i>animalis</i> strain THCT7A2	SAMN18042965	2.515	Gut	2339	[39,40]

Table 1. Cont.

Serial No.	Name	Biosample Accession	Genome Size (Mbp)	Isolation Source	Coding DNA Sequences	Reference
9.	<i>F. nucleatum</i> subsp. polymorphum strain THCT7E2	SAMN18042966	2.547	Gut	2420	[39]
10.	<i>F. nucleatum</i> subsp. vincentii strain THCT14A3	SAMN18042968	2.053	Gut	1903	[39]
11.	<i>F. nucleatum</i> subsp. animalis strain THCT6B3	SAMN18042964	2.269	Gut	2116	[39]
12.	<i>F. nucleatum</i> subsp. animalis strain P2_CP	SAMN07448031	2.351	Colorectal primary tumor	2346	[41]
13.	<i>F. nucleatum</i> subsp. animalis strain P2_LM	SAMN07448032	2.346	Liver metastasis	2353	-
14.	<i>F. nucleatum</i> CC53	SAMN02469329	2.070	Colon adenocarcinoma	1879	[42]

2.2. Pan-Genome and Core Genome Analysis

The retrieved 14 CRC-related *F. nucleatum* strains were subjected to pan-genome profiling to analyze the variation in the genome content (core genome and dispensable genome) using the BGPA software version 1.2, with parameters described before [43]. A USEARCH clustering algorithm was utilized to cluster only homologous genes detected in the pan-genome, with a 70% cut-off value. The genomes were aligned using the MUSCLE program [44] with default settings, and the phylogenetic analysis and tree-building processes were carried out using UPGMA. Following the classification of gene homologous sets, functional annotation was carried out using the Cluster of Ortholog Groups (COG) [45].

2.3. Potential Drug Targets Identification

Therapeutic targets were discovered by subtractive genomic analysis, using the predicted core genome of *F. nucleatum*. CD-HIT with a sequence identity cutoff of 0.8 (i.e., 80%) was performed to identify paralogous or duplicate protein sequences. The non-paralogous sequences of the *F. nucleatum* genome were removed from the core genome fraction of *F. nucleatum*.

These sequences were then subjected to BLASTP with an E-value cutoff of 10^{-4} [46] against the whole human proteome. The proteins with an E-value less than 10^{-3} were examined as potential targets. A gap penalty of 11 and a gap extension penalty of 1 were both considered standard. To determine the uniqueness of our targets, 83 different species of human microbial flora were compared to the usual gut flora. Based on extensive research, an E-value cut-off of 10^{-2} was selected to differentiate non-homologous proteins [47]. Antibiotic-resistant drug molecules were a primary target since they were unable to interact with human proteins. The non-homologous proteins to the human host and normal intestinal flora proteins were utilized for further assessment.

Further functional characterization of the non-homologous core genome was carried out. Essential genes were considered as those that have a substantial impact on the pathogen's ability to survive. Therefore, the obtained non-homologous proteins were checked for their essentiality in the Database of Essential Gene (DEG) [35]. A BLAST search with an E-value of 10^{-5} was performed on the non-homolog genes. Moreover, an essential gene database (CEG) was used to look into the genes that had high similarities with DEG. By validating essential genes based on both alignment and function, the CEG reduces the likelihood of receiving false-positive findings from essentiality predictions based on alignment. Overlapping the CEG and DEG essential gene lists allowed us to pick genes with similar reports in both lists for further study.

Using BLAST with an E-value of 10^{-5} against the whole DrugBank dataset [37] of prokaryotes (which include drug targets), the obtained essential genes were analyzed to determine the potential drug targets. For *F. nucleatum*, a list of essential, drug-like proteins with less than 30% identity and more than 50% query coverage against the Drugbank database was identified.

The virulent proteins among the druggable targets were also studied to identify the genes responsible for producing the virulent components using a database called VFDB [48]. Bacterial virulence proteins facilitate bacterial colonization and penetration of host immune cells, hence aiding in the destruction of the host immune system. A threshold of 10^{-3} was used to compare the *F. nucleatum* proteins with the VFDB.

2.4. Structural Retrieval and Virtual Screening

A method of in silico structure prediction known as homology modeling may be used to predict the three-dimensional structure of a selected protein, based on its primary amino acid sequence. Drug design requires an understanding of proteins' 3D structure and function [49]. A selected protein 3-D structure was retrieved from the Alpha Fold server using UniProt sequence ID. We uploaded its sequence to the STRING database online server and found interactions with other proteins, (represented by FN0707).

Virtual screening was performed using lead-like cyanobactins ($n = 237$), streptomycins ($n = 607$), and marine bacterial secondary metabolites ($n = 1226$) against the shortlisted drug target riboflavin kinase. Protonation and energy reduction were used to get the compounds ready for screening. Protein structure, on the other hand, was used as a receptor, while compounds served as ligands. The screening was carried out using MOE version 2019.0102, as described previously [28]. To do the screening, the triangle-matching method was used, whereas London dG, affinity dG, and forcefield parameters were used for rescoring. S-score was used to sort docked molecules. The compound with the lowest binding energy was chosen as the top conformation and used for further investigation.

2.5. ADMET Profiling of Shortlisted Drug Candidates

The development of safe drugs is a major problem but the drug development time and expenses may be halved if toxicological side effects can be detected early on. Screened drug-like substances were examined using the SwissADME software [50] for pharmacokinetic properties such as Absorption, Distribution, Metabolism, and Excretion (ADME). pkCSM [51] was used to analyze the compounds' toxicity profile, minimal human side effects, immunotoxicity, mutagenicity, teratogenicity, neurotoxicity, increased penetration, and carcinogenicity. As a result, the most effective method of administration, the impact on diverse species, and the elimination of drug properties in the human body were evaluated.

2.6. Dynamic Simulation Analysis

The complex interaction and stability of the ligand-protein complex were subsequently studied in detail using molecular dynamic (MD) simulation of these molecules. The GROMACS server v2020 [52] was the MD simulation program used, with GROMOS54A71 forcefield to rectify the topologies. The steepest energy minimization algorithm was used to minimize the energy of the ligand topology files. To construct a simulation system, the following parameters were used: size of the box with a 1.0 nm margin from the protein atom border; dodecahedron box as a boundary condition; SPC216 as an accurate water solvation model; Na⁺ ions were used to neutralize complexes and then simulated for 50 ns at a pressure of 1 atm and a temperature of 300 K. NPT and NVT were considered as an ensemble class, containing 50,000 steps of NPT and NVT [53]. A post-simulation examination of interactions was carried out.

3. Results

3.1. Pan-Genome Analysis

To prioritize the potent therapeutic target against *F. nucleatum*, the pan-genome analysis was applied to 14 strains. *F. nucleatum* strains consisted of >2230 DNA coding sequences. Among these, 1193 genes were shared by all strains and classified as a core genome fraction (i.e., 56% of pan-genome). The comparative study showed that the *Fusobacterium nucleatum* subsp. *animalis* strain P2_CP consisted of the maximum number of accessory genes ($n = 1017$), whereas the *F. nucleatum* subsp. *polymorphum* strain THCT15E1 strain encompassed the maximum number of unique genes ($n = 211$). Only one unique gene was present in the *F. nucleatum* subsp. *animalis* strain P2_CP while the minimum number of accessory genes were in the strain *F. nucleatum* Fn10-CTX3 ($n = 567$). Additionally, 11 strains were lacking several genes that were present in other strains (File S1). The pan-genome curve represented the $B_{\text{pan}} = 0.320$ (i.e., <1) resulting in the open nature of the pathogen (Figure 1). However, it is hypothesized that the pan-genome will continue to be accessible to new bacterial species as long as evolution and horizontal gene transfers are there [34].

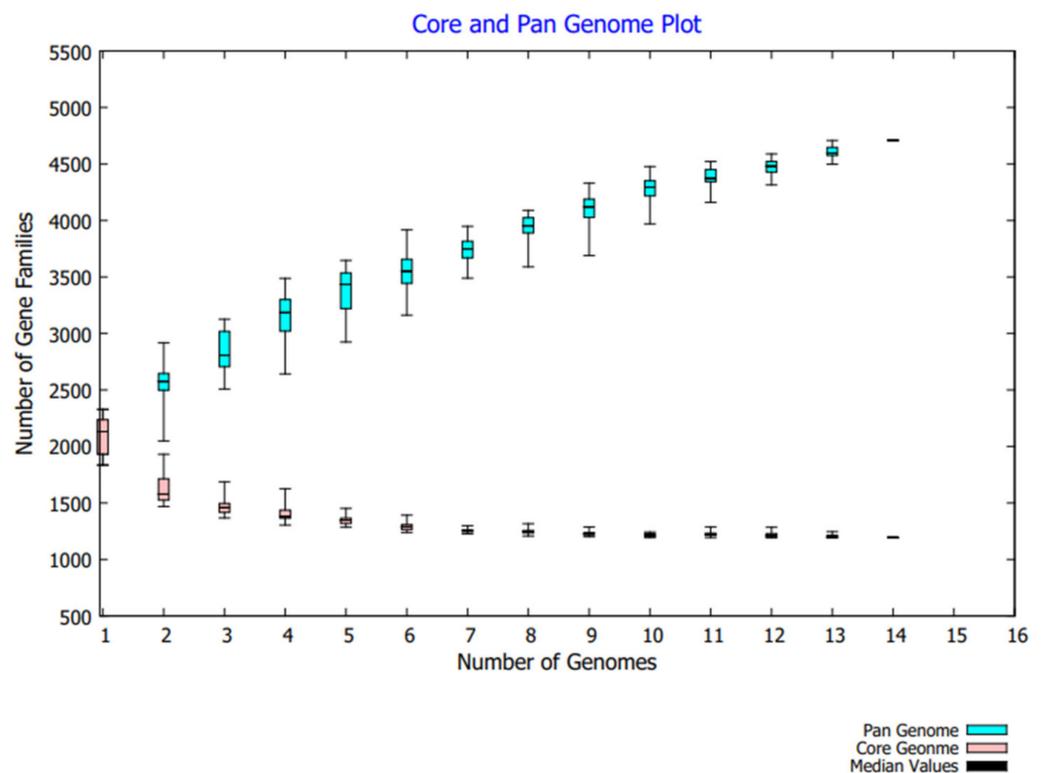


Figure 1. Pan-genome analysis of 14 *F. nucleatum* strains related to CRC.

All 14 strains were grouped together into a phylogenetic tree using data from the pan and core genomes. There was a difference in distance between the pan and core genome-based tree branches (Figure 2). Phylogenetic variation was also shown by cluster analysis using both genomic fractions. In comparison with the pan-genome, the distance in the core gene tree branches was much reduced.

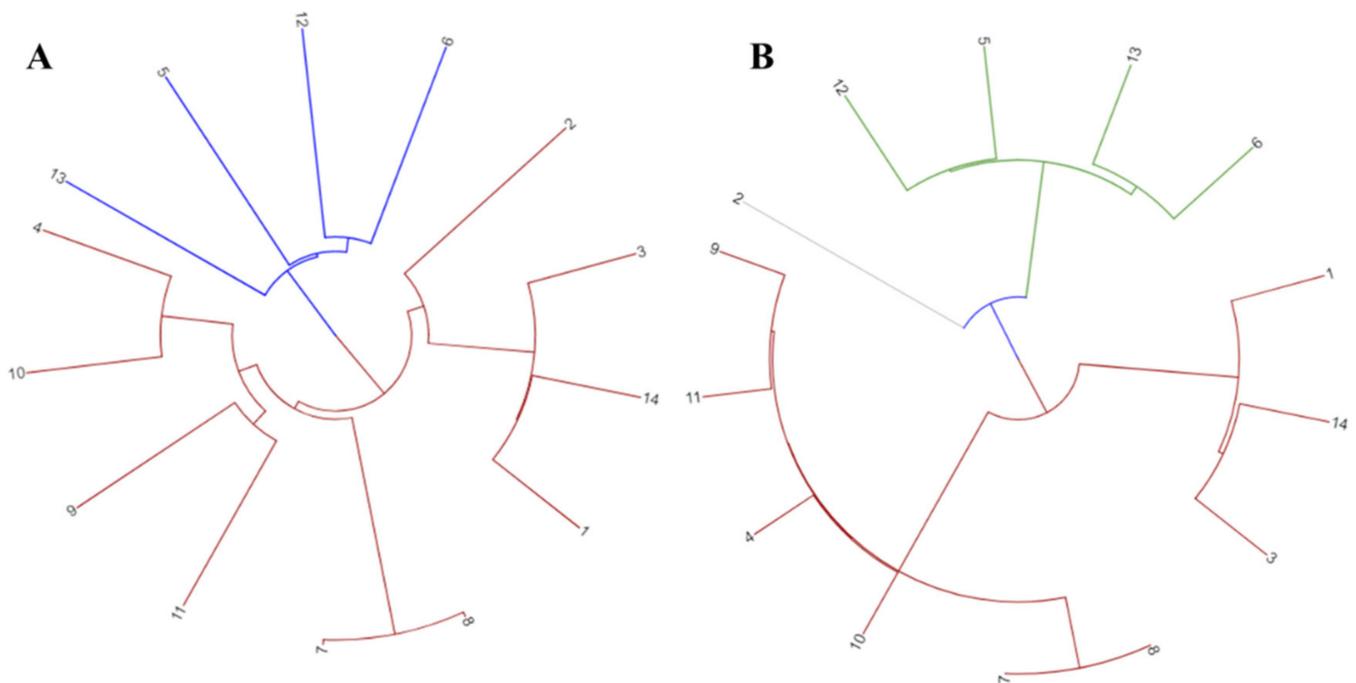


Figure 2. Phylogenetic tree depicting (A) Pan-genome phylogeny of 14 CRC related *F. nucleatum* strains used in this study (B) Core genome phylogeny of studied *F. nucleatum* strains. 1 = *F. nucleatum* CC53; 2 = *F. nucleatum* Fn10-CTX3; 3 = *F. nucleatum* Fn146CP; 4 = *F. nucleatum* Fn173CP; 5 = *F. nucleatum* Fn3760T; 6 = *F. nucleatum* FnS0431; 7 = *F. nucleatum* subsp. animalis strain P2_CP; 8 = *F. nucleatum* subsp. animalis strain P2_LM; 9 = *F. nucleatum* subsp. animalis strain THCT5A4; 10 = *F. nucleatum* subsp. animalis strain THCT6B3; 11 = *F. nucleatum* subsp. animalis strain THCT7A2; 12 = *F. nucleatum* subsp. polymorphum strain THCT7E2; 13 = *F. nucleatum* subsp. polymorphum strain THCT15E1; 14 = *F. nucleatum* subsp. vincentii strain THCT14A3.

3.2. Functional Annotation Studies

The COG functional analysis of the pan-genome was performed to look into the conserved proteins related to their specific metabolic pathways. The COG functional annotation for pan-genome analysis revealed that the core genome was found to be enriched in metabolic pathways, i.e., information storage, processing, and metabolic related pathways mainly enriched in amino acid transport metabolism, ribosomal, translational/biogenesis pathways, and inorganic ions transport metabolism. The accessory genome was identified to be highly involved in information storage, processing, and metabolic related pathways, mainly enriched in amino acid transport metabolism, and having genes linked with general function prediction. However, unique genes were mainly involved in information storage and processing pathways and poorly characterized (functionally unknown) pathways, i.e., general function prediction, cell wall, cell membrane, and envelop biogenesis (Figure 3).

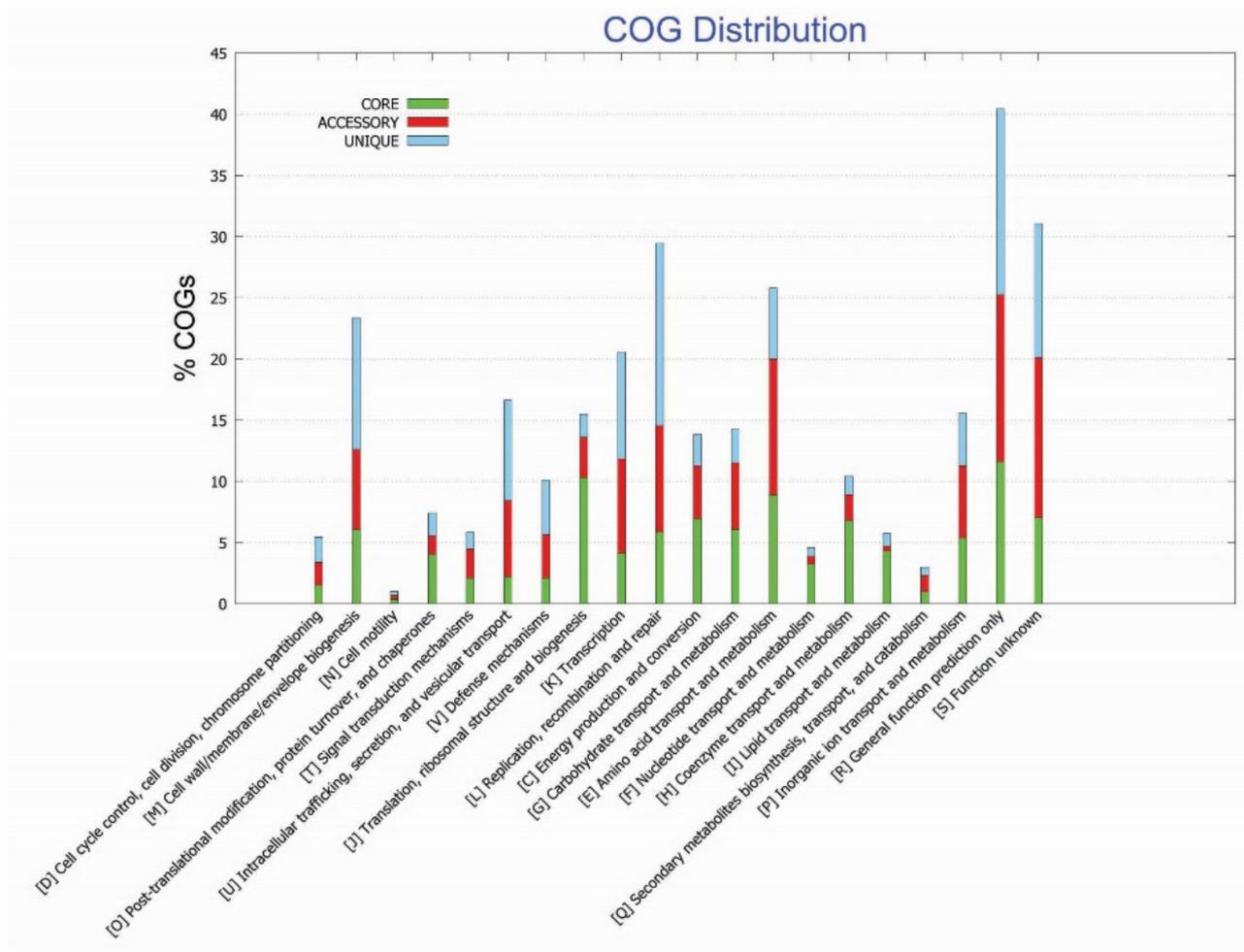


Figure 3. Functional enrichment analysis of identified genome of *F. nucleatum* strains.

3.3. Potential Drug Target Identification

For the purpose of predicting genes necessary for metabolic pathways and the overall survival of the *F. nucleatum*, a differential analysis of 1193 core protein-coding sequences was performed. The non-paralogous proteins from the core genome were retrieved via the CD-HIT tool. It resulted in the identification of 1190 proteins as non-paralogous while three proteins were excluded as paralogous to *F. nucleatum*. These identified 1190 proteins were further used for drug target identification.

Potentially useful and novel therapeutic treatment targets are those that have been highly conserved throughout evolution and are found in almost all pathogen strains of a genus. This explains why the study of such genes has long been a priority for researchers. RNA transcript inhibition and gene knock-out techniques, which are similar to mutant insertion for the loss of function in the gene, are usually used to verify the necessity of genes [54]. In recent decades, it has been possible to integrate this kind of traditional laboratory-based data into an online database, making the process more efficient, uniform, and repeatable. In this study, we evaluated the *F. nucleatum* essential genes using two databases, DEG and CEG. DEG predicts essential genes using a homology-based search algorithm with ~30,000 genes collected from ~70 different species [35], while CEG makes use of a forecasting technique using pre-determined homology-dependent clusters, which allows for the depiction of preservation and the specificity of gene functions. The DEG findings showed that 613 genes were crucial to the pathogen's survival, whereas the CEG study yielded 550 crucial genes; consequently, the predicted number of essential

proteins shared by both sets of genes was 538. These crucial genes were retained for future examination.

To prevent unwanted side effects, we need to ensure that the drug is selectively toxic against *F. nucleatum* but not against human and microbial gut flora. This will prevent the drug from binding to the active sites of the homologous proteins in the host [55]. All 538 of the obtained proteins resultant from the intersection of the DEG and CEG analysis were subjected to BLASTP against the human and microbial gut flora proteome to identify the non-homologous proteins. From a total of 538 proteins, BLAST determined 209 to be dissimilar to the human proteome, whereas 48 were dissimilar to the proteome of the human digestive tract microorganisms.

Furthermore, the druggability of these 48 essential, non-homologous proteins was determined through the BLAST against the DrugBank database. It revealed that only 15 proteins had drug target likeability and can act as potential therapeutics against *F. nucleatum*. Moreover, the virulent factors released from pathogenic proteins are mainly involved in the infliction of infections. Around 12 virulent proteins responsible for the pathogenic condition were identified.

3.4. Significant and Novel Drug Target Prediction

It has been reported that cytoplasmic proteins are an excellent therapeutic target and may be readily targeted with drugs [56]. Additionally, ~70 percent FDA approved medications are claimed to target enzymatic proteins because of their substantial participation in numerous pathways. Finally, among 12 candidate therapeutic targets, one protein was identified as an essential, non-homologous, druggable target against *F. nucleatum*, i.e., riboflavin synthase (RiS). Based on their cytoplasmic subcellular distribution, length > 100 amino acids, enzymatic nature, and role in important metabolic pathways, this discovered protein was used for structure-based study.

Among these 12 proteins, the enzyme RiS was chosen as the intended target (Figure 4A). Due to its lack of presence in humans, RiS is thought to be a promising therapeutic target since microorganisms are reliant on this metabolic process. The enzyme (EC:2.5.1.9) is essential in the bacterial secondary metabolite biosynthesis (Flavin Biosynthesis I). It catalyzes the final step in the biosynthesis of vitamin B2, i.e., the conversion of two molecules of 6,7-dimethyl-8-ribityllumazine to riboflavin and 5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione.

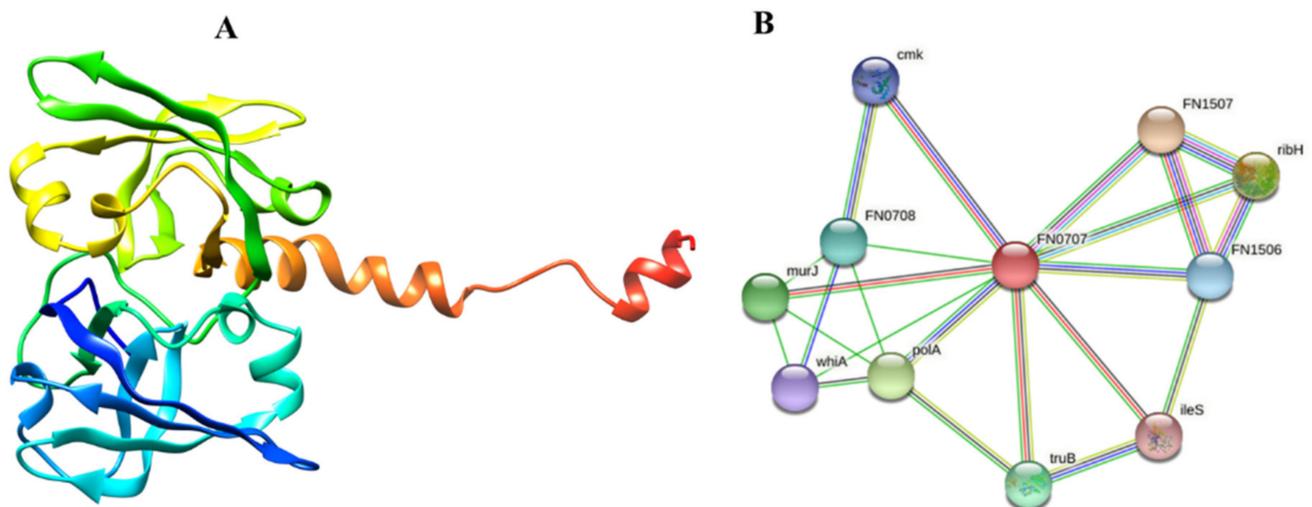


Figure 4. Representation of the selected drug target as a solo protein and part of its network (A) 3D structure of RiS retrieved from AlphaFold (B) PPI analysis of RiS.

Tutino et al. [57] show that CRC causes significant changes in the transcription and translation of RFVTs. Additionally, L-methionine and riboflavin metabolisms were the most prominent pathways affected in CRC cells as the cancer advances [58]. Moreover, this is not only in colon cancer as riboflavin is reportedly involved in the onset of gastric cancer [59]. Nevertheless, it is also involved in numerous infectious diseases and is reported as a potential marker for *Candida albicans* [59], *Brucella* spp. [60], and *Aspergillus fumigatus* [61]. While *F. nucleatum* has been linked to colon cancer before, it has never been investigated for therapeutic targets.

Protein-protein interactions and their functional annotation form the backbone of cellular machinery, which is responsible for regulating a plethora of biological activities [62]. To fully comprehend the PPI and its functional significance in the cell, it is necessary to recognize the several interactions that take place and regulate their outcomes [63]. The STRING analysis indicated that our prioritized protein likely interacted with many adjacent proteins to carry out essential functions, suggesting that it is a hub protein. Since proteins often work in groups, inhibiting this one may also impact the activity of its interactors. RiS mediates interactions with other proteins nearby, such as FN1507 (probability score: 0.964), ribH (probability score: 0.960), murJ (probability score: 0.918), FN0708 (probability score: 0.880), FN1506 (probability score: 0.851), cmk (probability score: 0.833), ileS (probability score: 0.819), polA (probability score: 0.935), truB (probability score: 0.891), and whiA (probability score: 0.827). The PPI results showed that isocitrate lyase had a total number of 11 nodes, an average node number of 4.18, an average local clustering coefficient 0.79, a PPI enrichment *p*-value 0.0266, and 15 expected numbers of edges, as shown in (Figure 4B). Many essential processes rely on the presence of these proteins. If RiS is inhibited, the other interactor proteins may eventually become dysfunctional. As a result, RiS may be presented as a viable pharmacological target without apprehension about the repercussions.

3.5. Virtual Screening Studies

The 3D structure of RiS was retrieved from the Alphafold server with the identifier ID: AF-Q8RF19-F1, with 319 amino acids. The retrieved structure was further used for the structure-based docking to prioritize potential inhibitors against it. Molecular docking is an excellent method for determining how studied complexes interact with a biological target, which is crucial in treatment. To comprehend the compounds-RiS interactions and analyze the likely binding mechanism and energy, the produced complexes were studied through the MOE tool. The docking analysis was performed for Ribityl (9-D-ribityl-1,3,7-trihydropurine-2,6,8-trione) as a reference control where the RiS was utilized as a protein receptor. The result showed the binding of the reference compound with the protein with different conformations and orientations. We selected conformation 1 of the ligand based on its binding affinity energy, i.e., -6.40 kcal/mol.

Three natural product libraries were taken and only lead-like compounds were retained for screening, such as (i) cyanobactins ($n = 237$), (ii) streptomycins ($n = 607$), and (iii) marine bacterial secondary metabolites ($n = 1226$), using rigorous docking to the active site of RiS having residues as Gly4, Leu5, Val6, Glu7, Glu8, Lys30, Arg137, Ala143, Ser144, Leu145, Thr146, Val157, Ser158, Leu159, Ile160, His162, Thr163, Lys166, and Ile167. It gave several docked configurations of compounds that were defined by docking scores. Lower binding affinity compounds (i.e., those with a binding affinity less than or comparable to -6 kcal/mol) were excluded from assessment as potential hit candidates. Due to their lower binding affinity than the Ribityl inhibitor, only three compounds (one from each library) were chosen for further study in this work because of their significant inhibitory effect against RiS, i.e., CMNPD3609 (Marine), Malyngamide V (Cyanobactin), and ZINC06804365 (Streptomycin metabolite) to inhibit *F. nucleatum*.

3.6. Interaction Analysis of Shortlisted Compounds

Shortlisted compounds were examined using post-molecular docking interaction analysis to better understand RiS pharmacological activity and its binding mechanism. Each

ligand had several interactions with the receptor in this molecular docking research. CMNPD3609 (-7.63) > Malyngamide V (-7.03) > ZINC06804365 (-7.01) is the docking rank order based on docking score. The hydroxyl group of the CMNPD3609 site chain mediates two hydrogen bonds with Lys30, and Val6, each with a bond distance of 3.18 \AA and 2.94 \AA along with an energy of -0.8 and -1.8 kcal/mol , correspondingly (Figure 5A). Malyngamide V, through its aromatic ring-OH, mediates two hydrogens bind with Val6 having a bond distance of 3.07 \AA and an energy of -1.2 – 2.8 kcal/mol (Figure 5B). ZINC06804365 showed two pi-hydrogen interactions between Leu5, Thr163, and the aromatic ring mediates, with a bond distance of 4.28 – 4.44 \AA and an energy range of -1.6 to -0.5 kcal/mol (Figure 5C). The control compound docked with RiS at -6.40 kcal/mol formed six hydrogen bonds with Ser158, Thr163, Val6, Leu159, Ile160, and Thr146 whereas its aromatic ring mediates a single pi-hydrogen bond with the Leu145 residue (Figure 5D). A description of binding interactions formed inside the RiS active cavity by the selected compounds is shown in Table 2.

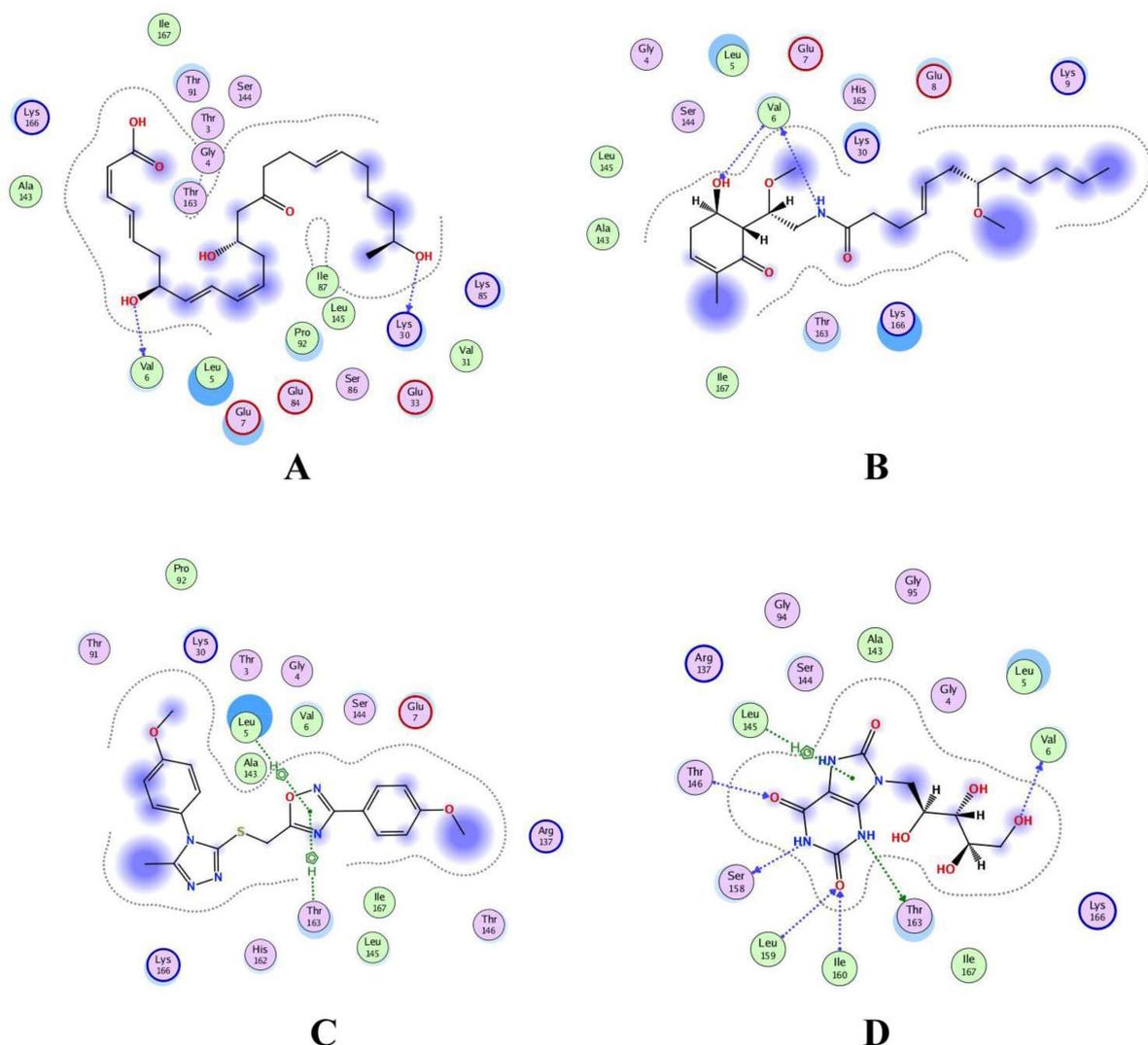


Figure 5. Molecular docking analysis of (A) RiS-CMNPD3609, (B) RiS-Malyngamide V, (C) RiS-ZINC06804365, and (D) RiS-control.

Table 2. Docking scores and identified bond types predicted through MOE tool for shortlisted compounds within the binding cavity of RiS.

S. No.	Ligand	Receptor	Interaction	Distance	E (kcal/mol)	S Value (kcal/mol)
Ribityl control						
1	N 2	O SER 158	H-donor	2.87	−5.9	−6.40
	N 4	OG1 THR 163	H-donor	3.06	−2.0	
	O 21	O VAL 6	H-donor	2.91	−1.4	
	O 9	CA LEU 159	H-acceptor	3.43	−0.5	
	O 9	N ILE 160	H-acceptor	3.22	−2.8	
	O 10	N THR 146	H-acceptor	2.95	−2.5	
	5-ring	CA LEU 145	pi-H	4.49	−1.0	
CMNPD3609						
2	O 27	O LYS 30	H-donor	3.18	−0.8	−7.63
	O 56	O VAL 6	H-donor	2.94	−1.8	
Malyngamide V						
3	N 13	O VAL 6	H-donor	3.08	−2.8	−7.03
	O 7	N VAL 6	H-acceptor	3.07	−1.2	
ZINC06804365						
4	5-ring	CA LEU 5	pi-H	4.44	−1.6	−7.01
	5-ring	CG2 THR 163	pi-H	4.28	−0.5	

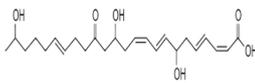
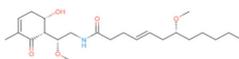
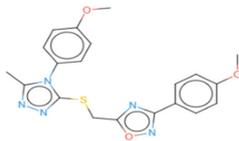
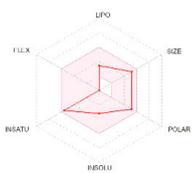
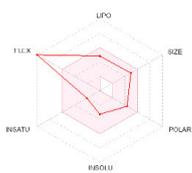
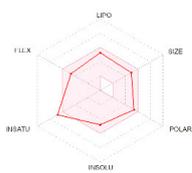
3.7. ADMET Profiling of Shortlisted Drug Candidates

To avoid drug adverse reactions and toxicity, the computational evaluation of ADMET properties is one of the crucial steps to investigate the drug safety assessment at the initial step of a drug candidate. The shortlisted four compounds (three natural products and one standard) were found to be non-inhibitors of CYP2C19 and CYP2D6, with significant permeability to skin and Caco2. Only CMNPD3609 was identified as a non-inhibitor while the other two compounds were identified as P-glycoprotein inhibitors, whereas all three prioritized compounds were impermeable to the BBB.

The ADMET characterization is a significant part of the drug discovery process since it reduces costs and development times in clinical trials [64]. When categorizing compounds for analogies to drugs, the Lipinski rule of five was used as a preliminary step. Ideally, drug-like molecules would have lipophilicity values (LogP) below five, hydrogen bond acceptor sites below ten, and a molecular weight of 500 amu or less. This is the ideal molecular weight range for orally administered drugs and chemicals, as suggested by RO5. This resulted in the finding that all three compounds precisely adhered to the RO5 guidelines as shown in Table 3.

Additionally, the Ames mutagenesis study used the pkcsm tool to identify the chemicals and predict their toxicity, Max. tolerated dose (human), Minnow toxicity, *T. Pyriformis* toxicity, Oral Rat Acute Toxicity (LD50), Hepatotoxic, and Skin Sensitization evaluation. Consequently, all compounds showed a negative Ames test except CMNPD3609. This means that these molecules do not cause mutagenicity. Only ZINC6804365 showed hepatotoxicity while all compounds were observed to have no skin sensitization. *T. pyriformis* showed the maximum tolerance to Malyngamide-V (0.575 log ug/L), while less tolerance was seen for the remaining compounds. The detailed information is shown in Table 3.

Table 3. Three shortlisted compounds as possible drug candidates against RiS along with their ADMET properties.

Compounds ID	CMNPD3609	Malyngamide-V	ZINC6804365
Compound Name	Isomacrolactic Acid	Malyngamide-V	34M (PDB ID)
Structure			
GI absorption	100	92.819	98.36
Caco2	1.317	1.076	1.242
Water solubility	−0.102	−4.386	−3.651
Skin permeability	−2.282	−2.827	−2.735
BBB permeant		No	
Lipinski		Yes	
Binding Affinity (kcal/mol)	−7.63	−7.03	−5.9
Radar			
Ames	Yes	No	No
Max. tolerated dose (human)	0.959	0.21	0.805
Hepatotoxicity	No	No	Yes
Skin sensation	No	No	No
<i>T.Pyriformis</i> toxicity	−0.964	0.575	0.286
Minnow toxicity	4.806	0.929	−2.353

3.8. MD Simulation of Protein-Ligand Complex

The MD simulation was performed for the shortlisted inhibitors to validate the complex interactions and flexibility. The GROMACS server was used to find the movements of molecules and atoms of protein complexes at 50 ns. All the compounds showed mild to moderate fluctuations within the range of 1.5–2 Å while all compounds were observed to be stable after 5 ns (Figure 6A). The RMSF and Rg for all four compounds highlights the identical patterns of protein and its binding pocket stability (Figure 6B,C). Importantly, the binding pattern of these compounds showed that the control mediated five hydrogen bonds, CMNPD3609 made four hydrogen bonds, while Malyngamide V and ZINC06804365 mediated three to four hydrogen bonds throughout the 50 ns simulations (Figure 6D). CMNPD3609 was observed to have high to moderate fluctuation in terms of RMSD, and a radius of gyration, while Malyngamide V and ZINC06804365 formed a stable complex throughout the simulation.

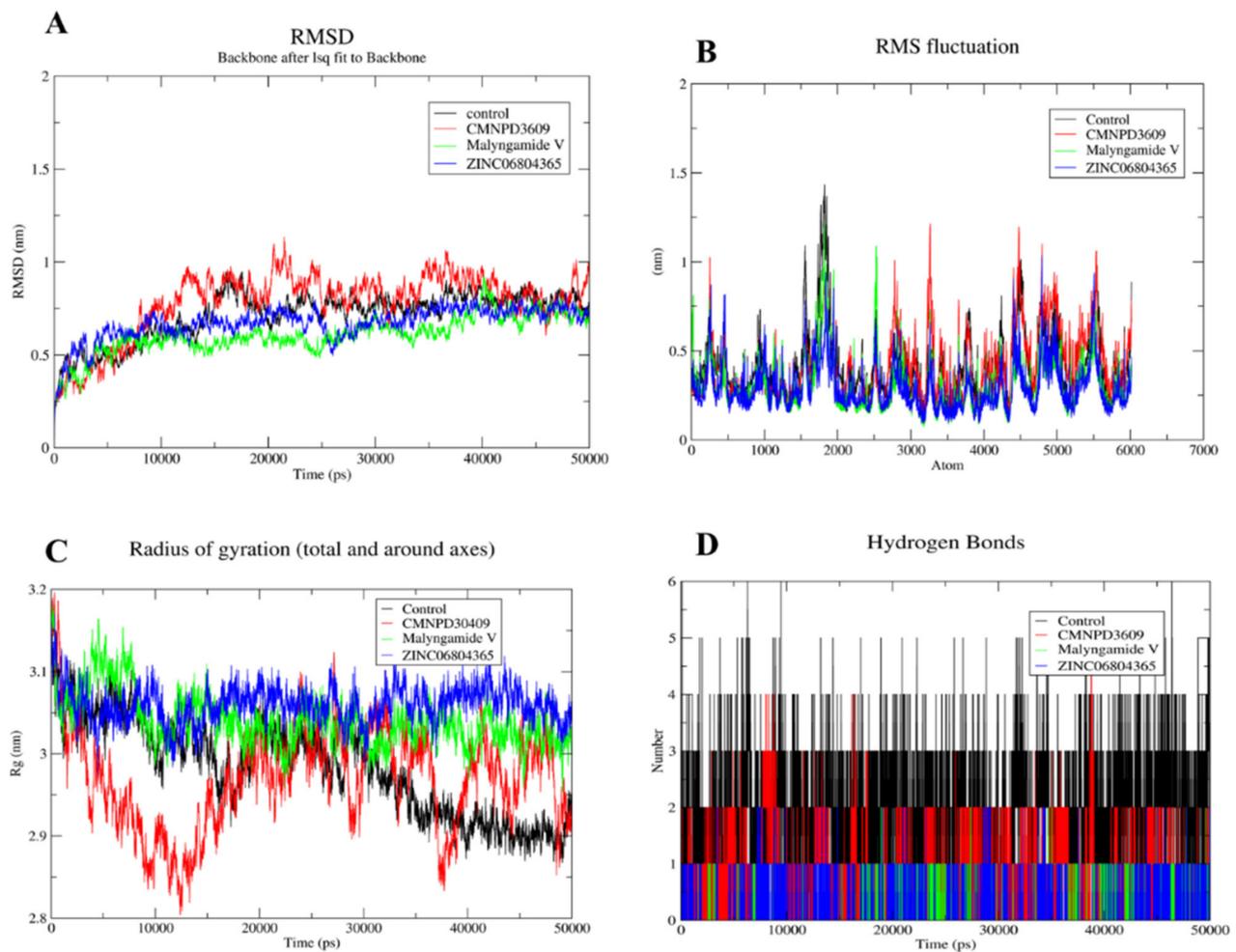


Figure 6. MD simulation for shortlisted compounds, depicting (A) RMSD (B) RMSF (C) Radius of gyration and (D) Hydrogen bonds.

4. Discussion

CRC is a pressing health issue worldwide [65], with a proportion of 9.2% among various cancer diagnoses and the second highest fatality rate of any cancer type [66]. The underlying mechanism of CRC malignancy is still to be determined but inflammation has been recognized as an essential risk factor for colorectal cancer, which is produced by a complex interaction of environmental, dietary, behavioral, and hereditary variables. The anaerobic, gram-negative bacillus *F. nucleatum* lives in species-specific reservoirs throughout the human body, including the mouth, gastrointestinal system, and other sites. Analysis of the 16S ribosomal ribonucleic acid (rRNA) gene sequence and the use of metagenomic sequencing techniques revealed a strong association between *F. nucleatum* and CRC [67]. Compared with the control group, it was significantly higher in patients with CRC [68]. Additionally, *F. nucleatum* was associated with a worse prognosis for CRC patients and likely aided in the development of chemoresistance [69].

Since pan-genomics has evolved as a standard for understanding the molecular evolution of bacterial populations in respect to the ever-increasing diversity of bacterial genomes [70], we analyzed pan-genomes to study species specific variations in *F. nucleatum*. The analysis of the 14 *F. nucleatum* genomes isolated from the CRC patients revealed the presence of various numbers of unique genes and accessory genes, with a shared core genome amounting to 1193 genes. Among the identified 12 potential therapeutic markers of *F. nucleatum* related to CRC, the RiS was selected as a novel and promiscuous drug/therapeutic prognostic target/marker involved in the biosynthesis of the secondary

metabolite Vitamin B12. Its reported role in CRC makes it a highly relatable therapeutic marker and drug target for CRC control. Tutino et al. [57] have shown that CRC causes significant changes in the transcription and translation of riboflavin transporters [57]. Additionally, L-methionine and riboflavin metabolisms were the most prominent pathways affected in CRC cells as the cancer advances [58]. Riboflavin synthase has also been previously earmarked as a drug target in several pathogenic bacteria [71–73], with an effect on the host's immune response. Since *F. nucleatum* is generally linked with immune suppression in CRC, inhibiting it with RiS binding of drugs can escalate the host cell pathogen clearance. Therefore, RiS inhibition can be a two-pronged approach, where both pathogen and host immunity can be targeted through simultaneous bactericidal activity and mitigation of host cell inflammation. Therefore, virtual screenings of lead-like natural products from various sources were performed to prioritize a new potent drug candidate against RiS, using Ribityl (9-D-ribityl-1,3,7-trihydropurine-2,6,8-trione) as a reference control. It resulted in the shortlisting of three drug candidates as potential binders, i.e., CMNPD3609 (Marine), Malyngamide V (Cyanobactin), and ZINC06804365 (Streptomycin metabolite) to inhibit *F. nucleatum*. The ZINC06804365 has been tested and reported as a potential inhibitor for many signaling pathways such as Wnt-3a and also as having inhibitory activity against the Tankyrase-2 enzyme which reportedly has anti-cancer activities [74]. Malyngamide V has been widely studied for its anti-inflammatory and antinociceptive role [75] and CMNPD3609 possesses several antibacterial activities [76]. The diverse role of these compounds in different cancer and antimicrobial activities makes them a possible drug candidate to curb CRC caused by *F. nucleatum*. However, the discovery of these compounds as drug candidates for CRC or *F. nucleatum* infections needs further experimental studies to validate their function. Furthermore, MD simulation was performed for these shortlisted compounds for 50 ns, to identify whether the binding and inhibition would be long lasting or transitory. It showed stable binding of all complexes (fluctuation within the range of 1.5–2 Å) after the initial phase of ~5 ns.

Eventually, the current computational pipeline will help in the identification of natural therapeutic products against the targets prioritized by the high throughput screening of the *F. nucleatum* core genome. Assuming the hefty cost, coupled with the small success rate in drug discovery and development, repurposing available compounds or the screening of natural product libraries against new targets and diseases is an important venture. It has lower development budgets and shorter screening periods, saving cost and time. Since natural products have been used in traditional medicine for hundreds of years, they are ideal for drug repositioning due to their medicinal value. However, this approach is best for prioritization although limitations exist. Therefore, we suggest extensive lab analysis in model organisms followed by clinical trials.

5. Conclusions

Cancer rate is increasing and complications associated with it, including role of microbes is a complex science. We need to maneuver the intricacies of the diseases like CRC, through swift approaches that are based on computation and artificial intelligence, both by understanding the mechanisms and finding cure for it. Bioinformatics based subtractive genomic approach is a state of the art method for identifying drug targets from genome sequences. Herein, we utilized this approach for finding out the therapeutic proteins from the CRC related oncomicrobe *F. nucleatum* genomes. MOE was then used for screening out inhibitors from natural product libraries against the selected target riboflavin synthase. Although, we got quick results but other parallel approaches may also be used for screening inhibitors and compare to our findings. The procedure could be further improved to Kdi or MIC evaluation of these inhibitors using artificial intelligence algorithms. Shortlisted natural product inhibitors may further be tested in the laboratory.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14246260/s1>, File S1.

Author Contributions: Z.B., A.A. and N.A.A. conceived and designed the study; N.A.A., K.J. and K.K. performed the experiments; K.K., N.A.A. and K.J. wrote the original draft; M.M.M., A.A. and Z.B. wrote, reviewed, and edited; N.A.A., M.M.M. and Z.B. provided software, resource support, and validated findings; A.A. and Z.B. supervised the study; Z.B. and A.A. operated funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deanship of Scientific Research at Shaqra University Research Support Program under the code SU-ANN-202211.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the genomic data used in this manuscript is publicly available from databases like NCBI and BV-BRC server, with accession numbers of genomes mentioned at appropriate places in the manuscript.

Acknowledgments: The authors extend their appreciation to the Deanship of Scientific Research at Shaqra University for funding this research work through the project number SU-ANN-202211.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2016. *CA Cancer J. Clin.* **2016**, *66*, 7–30. [[CrossRef](#)]
2. Abed, J.; Emgård, J.E.; Zamir, G.; Faroja, M.; Almogy, G.; Grenov, A.; Sol, A.; Naor, R.; Pikarsky, E.; Atlan, K.A. Fap2 mediates *Fusobacterium nucleatum* colorectal adenocarcinoma enrichment by binding to tumor-expressed Gal-GalNAc. *Cell Host Microbe* **2016**, *20*, 215–225. [[CrossRef](#)] [[PubMed](#)]
3. Bellerba, F.; Serrano, D.; Harriet, J.; Pozzi, C.; Segata, N.; NabiNejad, A.; Piperni, E.; Gnagnarella, P.; Macis, D.; Aristarco, V.; et al. Colorectal cancer, Vitamin D and microbiota: A double-blind Phase II randomized trial (ColoViD) in colorectal cancer patients. *Neoplasia* **2022**, *34*, 100842. [[CrossRef](#)] [[PubMed](#)]
4. Udayasuryan, B.; Ahmad, R.N.; Nguyen, T.T.; Umaña, A.; Monét Roberts, L.; Sobol, P.; Jones, S.D.; Munson, J.M.; Slade, D.J.; Verbridge, S.S. *Fusobacterium nucleatum* induces proliferation and migration in pancreatic cancer cells through host autocrine and paracrine signaling. *Sci. Signal.* **2022**, *15*, eabn4948. [[CrossRef](#)] [[PubMed](#)]
5. Vogelstein, B.; Kinzler, K.W. The multistep nature of cancer. *Trends Genet.* **1993**, *9*, 138–141. [[CrossRef](#)]
6. Abreu, M.T.; Peek, R.M., Jr. Gastrointestinal malignancy and the microbiome. *Gastroenterology* **2014**, *146*, 1534–1546.e3. [[CrossRef](#)] [[PubMed](#)]
7. Zundler, S.; Mardin, C.; Bertz, S.; Vitali, F.; Strauß, R.; Fürst, J.; Neurath, M.F.; Strobel, D. Rectal Cancer Presenting with Absceding Infection Due to *Fusobacterium nucleatum*. *Pathogens* **2022**, *11*, 1113. [[CrossRef](#)]
8. Ahn, H.; Min, K.; Lee, E.; Kim, H.; Kim, S.; Kim, Y.; Kim, G.; Cho, B.; Jeong, C.; Kim, Y.; et al. Whole-Transcriptome Sequencing Reveals Characteristics of Cancer Microbiome in Korean Patients with GI Tract Cancer: *Fusobacterium nucleatum* as a Therapeutic Target. *Microorganisms* **2022**, *10*, 1896. [[CrossRef](#)]
9. Huh, J.W.; Kim, M.J.; Kim, J.; Lee, H.G.; Ryoo, S.B.; Ku, J.L.; Jeong, S.Y.; Park, K.J.; Kim, D.; Kim, J.F.; et al. Enterotypical *Prevotella* and three novel bacterial biomarkers in preoperative stool predict the clinical outcome of colorectal cancer. *Microbiome* **2022**, *10*, 203. [[CrossRef](#)]
10. Keku, T.O.; McCoy, A.N.; Azcarate-Peril, A.M. *Fusobacterium* spp. and colorectal cancer: Cause or consequence? *Trends Microbiol.* **2013**, *21*, 506–508. [[CrossRef](#)]
11. Kostic, A.D.; Chun, E.; Robertson, L.; Glickman, J.N.; Gallini, C.A.; Michaud, M.; Clancy, T.E.; Chung, D.C.; Lochhead, P.; Hold, G.L. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **2013**, *14*, 207–215. [[CrossRef](#)] [[PubMed](#)]
12. Gur, C.; Ibrahim, Y.; Isaacson, B.; Yamin, R.; Abed, J.; Gamliel, M.; Enk, J.; Bar-On, Y.; Stanietsky-Kaynan, N.; Copenhagen-Glazer, S. Binding of the Fap2 protein of *Fusobacterium nucleatum* to human inhibitory receptor TIGIT protects tumors from immune cell attack. *Immunity* **2015**, *42*, 344–355. [[CrossRef](#)] [[PubMed](#)]
13. Flanagan, L.; Schmid, J.; Ebert, M.; Soucek, P.; Kunicka, T.; Liska, V.; Bruha, J.; Neary, P.; Dezeeuw, N.; Tommasino, M. *Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur. J. Clin. Microbiol. Infect. Dis.* **2014**, *33*, 1381–1390. [[CrossRef](#)] [[PubMed](#)]
14. Wu, Z.; Ma, Q.; Guo, Y.; You, F. The Role of *Fusobacterium nucleatum* in Colorectal Cancer Cell Proliferation and Migration. *Cancers* **2022**, *14*, 5350. [[CrossRef](#)]
15. Lo, C.-H.; Wu, D.-C.; Jao, S.-W.; Wu, C.-C.; Lin, C.-Y.; Chuang, C.-H.; Lin, Y.-B.; Chen, C.-H.; Chen, Y.-T.; Chen, J.-H.; et al. Enrichment of *Prevotella intermedia* in human colorectal cancer and its additive effects with *Fusobacterium nucleatum* on the malignant transformation of colorectal adenomas. *J. Biomed. Sci.* **2022**, *29*, 88. [[CrossRef](#)]

16. Parhi, L.; Alon-Maimon, T.; Sol, A.; Nejman, D.; Shhadeh, A.; Fainsod-Levi, T.; Yajuk, O.; Isaacson, B.; Abed, J.; Maalouf, N. Breast cancer colonization by *Fusobacterium nucleatum* accelerates tumor growth and metastatic progression. *Nat. Commun.* **2020**, *11*, 1–12. [[CrossRef](#)]
17. Baik, J.E.; Li, L.; Shah, A.; Freedberg, D.E.; Jin, Z.; Wang, T.C.; Han, Y.W. Circulating IgA antibodies against *Fusobacterium nucleatum* amyloid adhesin FadA are a potential biomarker for colorectal neoplasia. *Cancer Res. Commun.* **2022**, *2*, 1497–1503. [[CrossRef](#)]
18. Cavallucci, V.; Palucci, I.; Fidaleo, M.; Mercuri, A.; Masi, L.; Emoli, V.; Bianchetti, G.; Fiori, M.E.; Bachrach, G.; Scaldaferrri, F.; et al. Proinflammatory and Cancer-Promoting Pathobiont *Fusobacterium nucleatum* Directly Targets Colorectal Cancer Stem Cells. *Biomolecules* **2022**, *12*, 1256. [[CrossRef](#)]
19. Koveitypour, Z.; Panahi, F.; Vakilian, M.; Peymani, M.; Seyed Forootan, F.; Nasr Esfahani, M.H.; Ghaedi, K. Signaling pathways involved in colorectal cancer progression. *Cell Biosci.* **2019**, *9*, 97. [[CrossRef](#)]
20. Slattery, M.L.; Lundgreen, A.; Wolff, R.K. MAP kinase genes and colon and rectal cancer. *Carcinogenesis* **2012**, *33*, 2398–2408. [[CrossRef](#)]
21. Gao, Y.; Zou, T.; Xu, P.; Wang, Y.; Jiang, Y.; Chen, Y.X.; Chen, H.; Hong, J.; Fang, J.Y. *Fusobacterium nucleatum* stimulates cell proliferation and promotes PD-L1 expression via IFIT1-related signal in colorectal cancer. *Neoplasia* **2022**, *35*, 100850. [[CrossRef](#)] [[PubMed](#)]
22. LaCourse, K.D.; Zepeda-Rivera, M.; Kempchinsky, A.G.; Baryiamas, A.; Minot, S.S.; Johnston, C.D.; Bullman, S. The cancer chemotherapeutic 5-fluorouracil is a potent *Fusobacterium nucleatum* inhibitor and its activity is modified by intratumoral microbiota. *Cell Rep.* **2022**, *41*, 111625. [[CrossRef](#)] [[PubMed](#)]
23. Chen, J.; Zhang, P.; Zhao, Y.; Zhao, J.; Wu, X.; Zhang, R.; Cha, R.; Yao, Q.; Gao, Y. Nitroreductase-instructed supramolecular assemblies for microbiome regulation to enhance colorectal cancer treatments. *Sci. Adv.* **2022**, *8*, eadd2789. [[CrossRef](#)] [[PubMed](#)]
24. Almatroudi, A. The incidence rate of colorectal cancer in Saudi Arabia: An observational descriptive epidemiological analysis. *Int. J. Gen. Med.* **2020**, *13*, 977. [[CrossRef](#)] [[PubMed](#)]
25. Tran, H.N.H.; Thu, T.N.H.; Nguyen, P.H.; Vo, C.N.; Doan, K.V.; Nguyen Ngoc Minh, C.; Nguyen, N.T.; Ta, V.N.D.; Vu, K.A.; Hua, T.D.; et al. Tumour microbiomes and *Fusobacterium* genomics in Vietnamese colorectal cancer patients. *NPJ Biofilms Microbiomes* **2022**, *8*, 87. [[CrossRef](#)]
26. Tondi, D. Novel Targets and Mechanisms in Antimicrobial Drug Discovery. *Antibiotics* **2021**, *10*, 141. [[CrossRef](#)]
27. Costa, D.; Iraola, G. Pathogenomics of emerging *Campylobacter* species. *Clin. Microbiol. Rev.* **2019**, *32*, e00072-18. [[CrossRef](#)]
28. Basharat, Z.; Khan, K.; Jalal, K.; Ahmad, D.; Hayat, A.; Alotaibi, G.; Al Mouslem, A.; Alkhayl, F.F.A.; Almatroudi, A. An in silico hierarchal approach for drug candidate mining and validation of natural product inhibitors against pyrimidine biosynthesis enzyme in the antibiotic-resistant *Shigella flexneri*. *Infect. Genet. Evol.* **2022**, *98*, 105233. [[CrossRef](#)]
29. Vasaikar, S.; Huang, C.; Wang, X.; Petyuk, V.A.; Savage, S.R.; Wen, B.; Dou, Y.; Zhang, Y.; Shi, Z.; Arshad, O.A. Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell* **2019**, *177*, 1035–1049.e19. [[CrossRef](#)]
30. Caputo, A.; Fournier, P.-E.; Raoult, D. Genome and pan-genome analysis to classify emerging bacteria. *Biol. Direct* **2019**, *14*, 5. [[CrossRef](#)]
31. Gordon, S.P.; Contreras-Moreira, B.; Levy, J.J.; Djamei, A.; Czedik-Eysenberg, A.; Tartaglio, V.S.; Session, A.; Martin, J.; Cartwright, A.; Katz, A. Gradual polyploid genome evolution revealed by pan-genomic analysis of *Brachypodium hybridum* and its diploid progenitors. *Nat. Commun.* **2020**, *11*, 1–16. [[CrossRef](#)] [[PubMed](#)]
32. Allemailem, K.S. A Comprehensive Computer Aided Vaccine Design Approach to Propose a Multi-Epitopes Subunit Vaccine against Genus *Klebsiella* Using Pan-Genomics, Reverse Vaccinology, and Biophysical Techniques. *Vaccines* **2021**, *9*, 1087. [[CrossRef](#)] [[PubMed](#)]
33. Basharat, Z.; Jahanzaib, M.; Rahman, N. Therapeutic target identification via differential genome analysis of antibiotic resistant *Shigella sonnei* and inhibitor evaluation against a selected drug target. *Infect. Genet. Evol.* **2021**, *94*, 105004. [[CrossRef](#)] [[PubMed](#)]
34. Basharat, Z.; Akhtar, U.; Khan, K.; Alotaibi, G.; Jalal, K.; Abbas, M.N.; Hayat, A.; Ahmad, D.; Hassan, S.S. Differential analysis of *Orientia tsutsugamushi* genomes for therapeutic target identification and possible intervention through natural product inhibitor screening. *Comput. Biol. Med.* **2022**, *141*, 105165. [[CrossRef](#)] [[PubMed](#)]
35. Luo, H.; Lin, Y.; Liu, T.; Lai, F.-L.; Zhang, C.-T.; Gao, F.; Zhang, R. DEG 15, an update of the Database of Essential Genes that includes built-in analysis tools. *Nucleic Acids Res.* **2021**, *49*, D677–D686. [[CrossRef](#)]
36. Liu, S.; Wang, S.-X.; Liu, W.; Wang, C.; Zhang, F.-Z.; Ye, Y.-N.; Wu, C.-S.; Zheng, W.-X.; Rao, N.; Guo, F.-B. CEG 2.0: An updated database of clusters of essential genes including eukaryotic organisms. *Database* **2020**, *2020*, baaa112. [[CrossRef](#)]
37. Wishart, D.S.; Feunang, Y.D.; Guo, A.C.; Lo, E.J.; Marcu, A.; Grant, J.R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018**, *46*, D1074–D1082. [[CrossRef](#)]
38. Queen, J.; Domingue, J.C.; White, J.R.; Stevens, C.; Udayasuryan, B.; Nguyen, T.T.D.; Wu, S.; Ding, H.; Fan, H.; McMann, M.; et al. Comparative Analysis of Colon Cancer-Derived *Fusobacterium nucleatum* Subspecies: Inflammation and Colon Tumorigenesis in Murine Models. *mBio* **2022**, *13*, e0299121. [[CrossRef](#)]
39. Bi, D.; Zhu, Y.; Gao, Y.; Li, H.; Zhu, X.; Wei, R.; Xie, R.; Cai, C.; Wei, Q.; Qin, H. Profiling *Fusobacterium* infection at high taxonomic resolution reveals lineage-specific correlations in colorectal cancer. *Nat. Commun.* **2022**, *13*, 3336. [[CrossRef](#)]
40. Bi, D.; Zhu, Y.; Gao, Y.; Li, H.; Zhu, X.; Wei, R.; Xie, R.; Wei, Q.; Qin, H. A newly developed PCR-based method revealed distinct *Fusobacterium nucleatum* subspecies infection patterns in colorectal cancer. *Microb. Biotechnol.* **2021**, *14*, 2176–2186. [[CrossRef](#)]

41. Bullman, S.; Pedomallu, C.S.; Sicinska, E.; Clancy, T.E.; Zhang, X.; Cai, D.; Neuberger, D.; Huang, K.; Guevara, F.; Nelson, T.; et al. Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer. *Science* **2017**, *358*, 1443–1448. [[CrossRef](#)]
42. Castellarin, M.; Warren, R.L.; Freeman, J.D.; Dreolini, L.; Krzywinski, M.; Strauss, J.; Barnes, R.; Watson, P.; Allen-Vercoe, E.; Moore, R.A.; et al. Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Res.* **2012**, *22*, 299–306. [[CrossRef](#)] [[PubMed](#)]
43. Basharat, Z.; Jahanzaib, M.; Yasmin, A.; Khan, I.A. Pan-genomics, drug candidate mining and ADMET profiling of natural product inhibitors screened against Yersinia pseudotuberculosis. *Genomics* **2021**, *113*, 238–244. [[CrossRef](#)]
44. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]
45. Galperin, M.Y.; Wolf, Y.I.; Makarova, K.S.; Vera Alvarez, R.; Landsman, D.; Koonin, E.V. COG database update: Focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res.* **2021**, *49*, D274–D281. [[CrossRef](#)] [[PubMed](#)]
46. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [[CrossRef](#)] [[PubMed](#)]
47. Sarangi, A.N.; Lohani, M.; Aggarwal, R. Proteome mining for drug target identification in Listeria monocytogenes strain EGD-e and structure-based virtual screening of a candidate drug target penicillin binding protein 4. *J. Microbiol. Methods* **2015**, *111*, 9–18. [[CrossRef](#)]
48. Chen, L.; Xiong, Z.; Sun, L.; Yang, J.; Jin, Q. VFDB 2012 update: Toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res.* **2012**, *40*, D641–D645. [[CrossRef](#)]
49. Kiefer, F.; Arnold, K.; Künzli, M.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res.* **2009**, *37*, D387–D392. [[CrossRef](#)]
50. Daina, A.; Michielin, O.; Zoete, V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* **2017**, *7*, 42717. [[CrossRef](#)]
51. Pires, D.E.; Blundell, T.L.; Ascher, D.B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J. Med. Chem.* **2015**, *58*, 4066–4072. [[CrossRef](#)] [[PubMed](#)]
52. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; Van Der Spoel, D. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854. [[CrossRef](#)] [[PubMed](#)]
53. Jalal, K.; Khan, K.; Haleem, D.J.; Uddin, R. In silico study to identify new monoamine oxidase type a (MAO-A) selective inhibitors from natural source by virtual screening and molecular dynamics simulation. *J. Mol. Struct.* **2022**, *1254*, 132244. [[CrossRef](#)]
54. Luo, J.; Wu, J. A new algorithm for essential proteins identification based on the integration of protein complex co-expression information and edge clustering coefficient. *Int. J. Data Min. Bioinform.* **2015**, *12*, 257–274. [[CrossRef](#)]
55. Shanmugham, B.; Pan, A. Identification and characterization of potential therapeutic candidates in emerging human pathogen Mycobacterium abscessus: A novel hierarchical in silico approach. *PLoS ONE* **2013**, *8*, e59126. [[CrossRef](#)] [[PubMed](#)]
56. Khan, K.; Jalal, K.; Khan, A.; Al-Harrasi, A.; Uddin, R. Comparative Metabolic Pathways Analysis and Subtractive Genomics Profiling to Prioritize Potential Drug Targets Against Streptococcus pneumoniae. *Front. Microbiol.* **2021**, *12*, 796363. [[CrossRef](#)]
57. Tutino, V.; Defrancesco, M.L.; Tolomeo, M.; De Nunzio, V.; Lorusso, D.; Paleni, D.; Caruso, M.G.; Notarnicola, M.; Barile, M. The expression of riboflavin transporters in human colorectal cancer. *Anticancer Res.* **2018**, *38*, 2659–2667.
58. Mohd Yusof, H. *Characterisation of Metabolic Changes in Colorectal Cancer Cells of Different Stages*; Universiti Teknologi MARA: Shah Alam, Malaysia, 2020.
59. Lu, Y.-T.; Gunathilake, M.; Lee, J.; Choi, I.J.; Kim, Y.-I.; Kim, J. Riboflavin intake, MTRR genetic polymorphism (rs1532268) and gastric cancer risk in a Korean population: A case-control study. *Br. J. Nutr.* **2022**, *127*, 1026–1033. [[CrossRef](#)]
60. Serer, M.I.; Carrica, M.d.C.; Trappe, J.; López Romero, S.; Bonomi, H.R.; Klinke, S.; Cerutti, M.L.; Goldbaum, F.A. A high-throughput screening for inhibitors of riboflavin synthase identifies novel antimicrobial compounds to treat brucellosis. *FEBS J.* **2019**, *286*, 2522–2535. [[CrossRef](#)]
61. Dietl, A.-M.; Meir, Z.; Shadkchan, Y.; Oshero, N.; Haas, H. Riboflavin and pantothenic acid biosynthesis are crucial for iron homeostasis and virulence in the pathogenic mold Aspergillus fumigatus. *Virulence* **2018**, *9*, 1036–1049. [[CrossRef](#)]
62. Braun, P.; Gingras, A.C. History of protein–protein interactions: From egg-white to complex networks. *Proteomics* **2012**, *12*, 1478–1498. [[CrossRef](#)] [[PubMed](#)]
63. Zhang, A. *Protein Interaction Networks: Computational Analysis*; Cambridge University Press: Cambridge, MA, USA, 2009.
64. Shahinozzaman, M.; Taira, N.; Ishii, T.; Halim, M.A.; Hossain, M.A.; Tawata, S. Anti-inflammatory, anti-diabetic, and anti-Alzheimer's effects of prenylated flavonoids from Okinawa propolis: An investigation by experimental and computational studies. *Molecules* **2018**, *23*, 2479. [[CrossRef](#)] [[PubMed](#)]
65. Tam, S.Y.; Wu, V.W. A review on the special radiotherapy techniques of colorectal cancer. *Front. Oncol.* **2019**, *9*, 208. [[CrossRef](#)] [[PubMed](#)]
66. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
67. Sun, C.-H.; Li, B.-B.; Wang, B.; Zhao, J.; Zhang, X.-Y.; Li, T.-T.; Li, W.-B.; Tang, D.; Qiu, M.-J.; Wang, X.-C. The role of Fusobacterium nucleatum in colorectal cancer: From carcinogenesis to clinical management. *Chronic Dis. Transl. Med.* **2019**, *5*, 178–187. [[CrossRef](#)] [[PubMed](#)]

68. Fukugaiti, M.H.; Ignacio, A.; Fernandes, M.R.; Ribeiro Júnior, U.; Nakano, V.; Avila-Campos, M.J. High occurrence of *Fusobacterium nucleatum* and *Clostridium difficile* in the intestinal microbiota of colorectal carcinoma patients. *Braz. J. Microbiol.* **2015**, *46*, 1135–1140. [[CrossRef](#)]
69. Yu, T.; Guo, F.; Yu, Y.; Sun, T.; Ma, D.; Han, J.; Qian, Y.; Kryczek, I.; Sun, D.; Nagarsheth, N. *Fusobacterium nucleatum* promotes chemoresistance to colorectal cancer by modulating autophagy. *Cell* **2017**, *170*, 548–563.e16. [[CrossRef](#)]
70. 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. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome”. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13950–13955. [[CrossRef](#)]
71. Jalal, K.; Khan, K.; Hayat, A.; Ahmad, D.; Alotaibi, G.; Uddin, R.; Mashraqi, M.M.; Alzamami, A.; Aurongzeb, M.; Basharat, Z. Mining therapeutic targets from the antibiotic-resistant *Campylobacter coli* and virtual screening of natural product inhibitors against its riboflavin synthase. *Mol Divers* **2022**. [[CrossRef](#)]
72. Long, Q.; Ji, L.; Wang, H.; Xie, J. Riboflavin biosynthetic and regulatory factors as potential novel anti-infective drug targets. *Chem. Biol. Drug Des.* **2010**, *75*, 339–347. [[CrossRef](#)]
73. Aris, S.N.A.M.; Rahman, R.N.Z.R.A.; Ali, M.S.; Jonet, M.A.; Motomura, T.; Noor, N.D.M.; Shariff, F.M.; Hsu, K.; Leow, T.C. Unraveling the crystal structure of *Leptospira kmetyi* riboflavin synthase and computational analyses for potential development of new antibacterials. *J. Mol. Struct.* **2022**, *165*, 133420. [[CrossRef](#)]
74. Lien, V.T.; Kristiansen, M.K.; Pettersen, S.; Haugen, M.H.; Olberg, D.E.; Waaler, J.; Klaveness, J. Towards dual inhibitors of the MET kinase and WNT signaling pathway; design, synthesis and biological evaluation. *RSC Adv.* **2019**, *9*, 37092–37100. [[CrossRef](#)] [[PubMed](#)]
75. Li, Z.; Zhang, L.; Zhao, Z. Malyngamide F Possesses Anti-Inflammatory and Antinociceptive Activity in Rat Models of Inflammation. *Pain Res. Manag.* **2021**, *2021*, 4919391. [[CrossRef](#)]
76. Cheng, Z.; Lou, L.; Liu, D.; Li, X.; Proksch, P.; Yin, S.; Lin, W. Versiquinazolines A–K, Fumiquinazoline-Type Alkaloids from the Gorgonian-Derived Fungus *Aspergillus versicolor* LZD-14-1. *J. Nat. Prod.* **2016**, *79*, 2941–2952. [[CrossRef](#)] [[PubMed](#)]