Bioinformatic challenge on prostate cancer and urinary microbiome

Reto bioinformático sobre el cáncer de próstata y el microbioma urinario

Elmira Davasaz Tabrizi[1]* Mushteba Sevil2 Omar Sajer Naser[3] Ercan Arican[4]
[1]*The Istanbul University, Faculty of Science, Molecular Biology and Genetics Department, 34134, Vezneciler-Istanbul/ TÜRKİYE; ORCID:0000-0002-1676-5919. Email: elmira.dt@gmail.com,
[2]The Private Capa hospital, Department of surgery and urology, 34134, Istanbul, Turkey Email: info@mustebasevil.com https://orcid.org/0009-0006-5461-2151
[3]The Istanbul University, Institute of Science, Molecular Biology and Genetics Section, 34134, Vezneciler-Istanbul/ TÜRKİYE; ORCİD: 0009-0003-3762-9304 Email: omarsajer89@gmail.com
[4]Department of molecular biology and genetic, Istanbul University, Istanbul, Turkey. Email: earican@istanbul.edu.tr, ORCID: 0000-0002-1676-5919.
*Corresponding author: Elmira Davasaz Tabrizi
Received: 05/20/2022 Accepted: 08/19/2023 Published: 09/25/2023 DOI: http://doi.org/10.5281/zenodo.10029976

Abstract

The Bioinformatics Resource Center (BRC) program was developed by the National Institute of Allergy and Infectious Diseases (NIAID) to assist researchers in analyzing the increasing amount of genomic sequences and other omics-related data. In this work, whole-genome sequences of prostate cancer and genitourinary diseases (WGS) were examined for genes utilizing the BV-BRC Bioinformatics Resource Center. using the Usegalaxy program to combine the plasma and gut microbiome sequences from prostate cancer patients. Following that, chromosomes, plasmids, and unclassified sequences were subjected to ARG analysis. As a result of comprehensive genomic analysis of all samples, the S2 sequences were of good quality compared to the other sequences. As for virulence factors, intracellular survival is one of the important virulence factors, a common gene of Salmonella, which was represented in the prostate cancer samples but not in the urine microbiome samples.

Keywords: Bioinformatic, Virulence factors, Antibiotic resistance gene, Prostate cancer.

Resumen

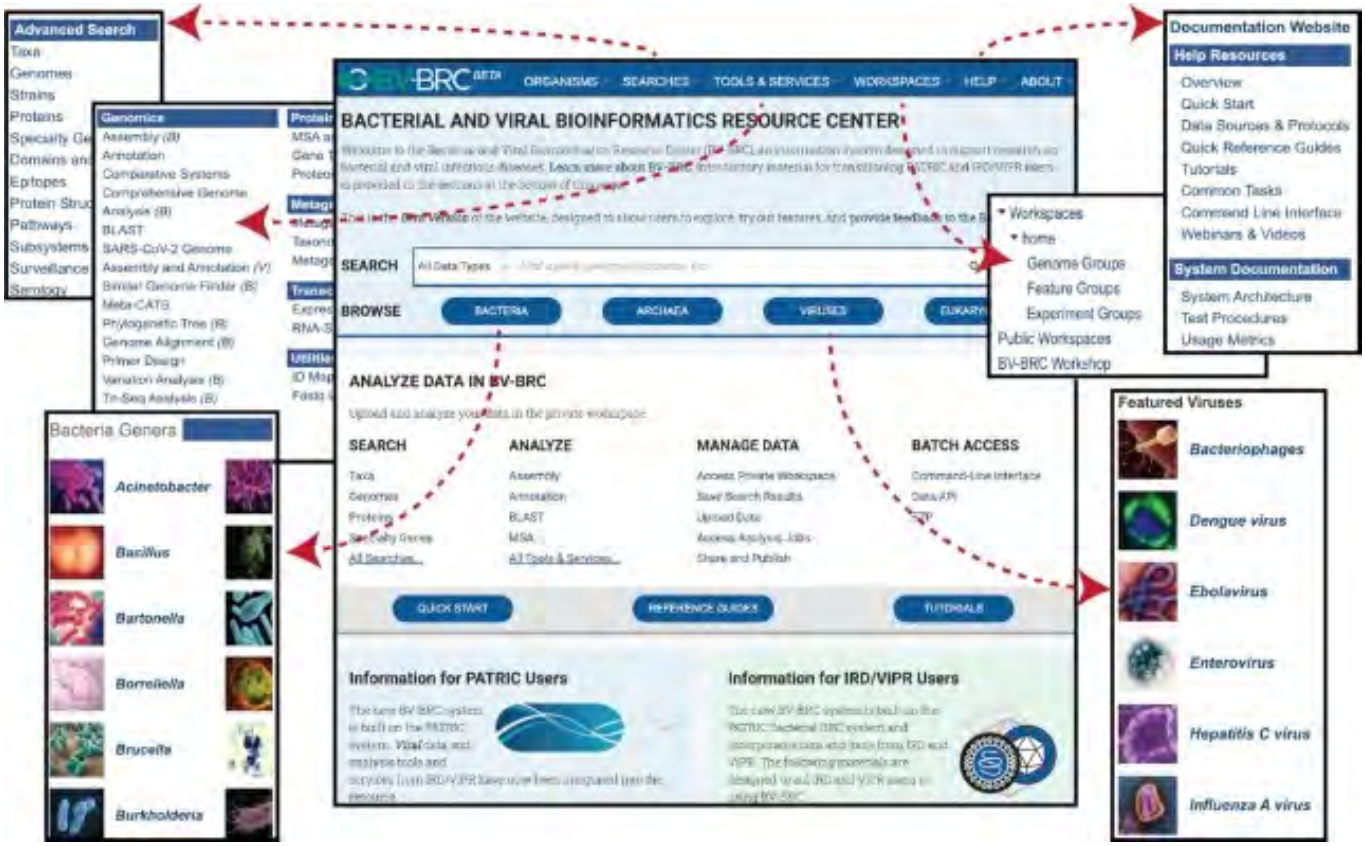
El programa del Centro de Recursos Bioinformáticos (BRC) fue desarrollado por el Instituto Nacional de Alergias y Enfermedades Infecciosas (NIAID) para ayudar a los investigadores a analizar la creciente cantidad de secuencias genómicas y otros datos relacionados con la ómica. En este trabajo, se examinaron secuencias del genoma completo del cáncer de próstata y enfermedades genitourinarias (WGS) en busca de genes utilizando el Centro de recursos de bioinformática BV-BRC. utilizando el programa Usegalaxy para combinar secuencias de microbioma plasmático y intestinal de pacientes con cáncer de próstata. A continuación, los cromosomas, plásmidos y secuencias no clasificadas se sometieron a análisis ARG. Como resultado de un análisis genómico exhaustivo de todas las muestras, las secuencias S2 eran de buena calidad en comparación con las otras secuencias. En cuanto a los factores de virulencia, la supervivencia intracelular es uno de los factores de virulencia importantes, un gen común de Salmonella, que estuvo representado en las muestras de cáncer de próstata, pero no en las muestras de microbioma de orina.

Palabras clave: Bioinformática, Factores de virulencia, Gen de resistencia a antibióticos, Cáncer de próstata.

Introduction

In 2004, the National Institute of Allergy and Infectious Diseases (NIAID) introduced the Bioinformatic Resource Center (BRC) program to facilitate the integration of genomic data with other biological data in pathogen research. The program aimed to enhance basic and applied research through the provision of appropriate data environments, bioinformatic tools, and workflows. Following the evolution of the program, two BRCs now support research on eukaryotic diseases and their vectors in the invertebrate domain (VEuPathDB), and a program (BV-BRC) encourages the study of bacterial and viral diseases. In 2019, the Bacterial and Viral Bioinformatics Research Center (BV-BRC) was created by merging the Viruses Pathogens Database and Analysis Resource (ViPR), the PAThosystems Resource Integration Center (PATRIC), and the Influenza Research Database (IRD). PATRIC, one of the initial BRCs, aimed to advance the bioinformatics of bacterial illnesses. Following the integration of the National Microbial Pathogen Database Resource (NMPDR) and the PATRIC BRC in 2012, the SEED and RAST resources for annotation were recognized. The BV-BRC type 9 has a singular, current resource that is aided by a group of researchers from the University of Chicago, the J. Craig Venter Institute, the University of Virginia, and the Fellowship for the interpretation of genomes, as well as other colleagues from nearby universities. Every bacterial and archaeal genome in the BV-BRC is meticulously annotated by the BV-BRC Annotation Service using the RASTtk tool[10] for more precise gene annotations. Included in these annotations are antibiotic resistance genes, virulence factors, and essential genes such as Mobile Genetic Elements and Metal Resistance Genes, which are crucial for conducting research on infectious diseases. This study involved four distinct WGS samples obtained from NCBI for the purpose of comparative genomic analysis. Figure 1 illustrates the homepage of the BV-BRC. There are a multitude of methods for studying bacteria and viruses.

Fig 1. The home page of the BV-BRC organization. Several methods of obtaining bacterial and viral information are demonstrated.



Adopted from https://doi.org/10.1093/nar/gkac1003.

Trimming

Taxonomy Assignment

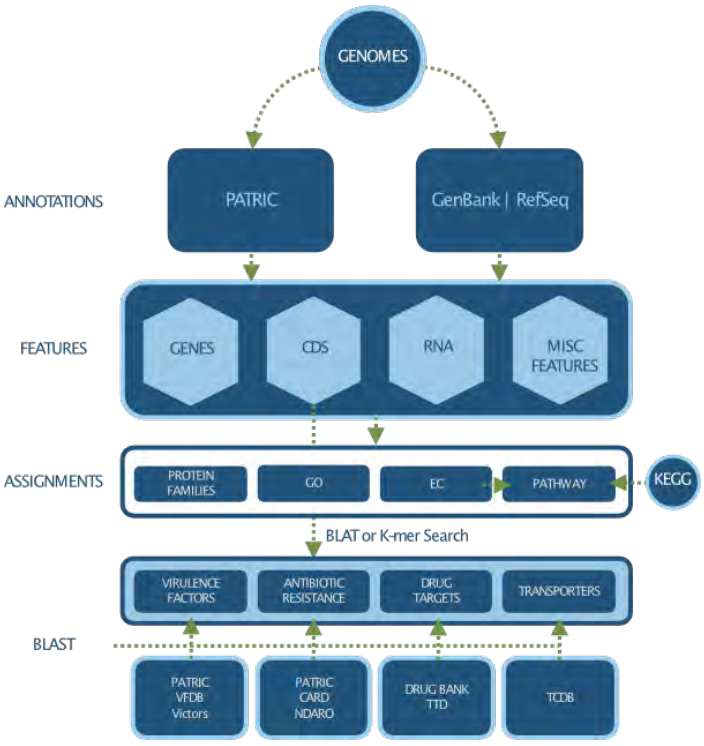
Metagenomic Assembly and the prediction of Plasmid/ chromosomal DNA sequences

All 4 WGS metagenomic samples consist of S1: 4.6 GB, S2: 981.8 MB, S3: 2.9, and S4: 1.3 GB, respectively. All paired and individual sample reads were submitted to the BV-BRC Bioinformatics Reference. The genome assembly pipeline (version 2.9.1) performs paired (S1, S3, S4) and single-end (S2) on the illumine, NextSeq, and HighSeq tool platforms. In order to improve genome assembly designs, bases were changed, misassemblies were fixed, and gaps were filled using the SPAdes v3.12.0 assembly tool. The components' quality was then evaluated with QUAST (version 5.0.2) after two rounds of polishing with Pilon 1.23. Furthermore, based on neural network models trained on whole genome and plasmid sequences, Plasflow (usegalaxy v 21.09) default settings for all metagenomic assembled contigs combined predicted plasmid and chromosomal sequences with over 96% accuracy.

Functional Annotation of Metagenome-Assembled Genomes (MAGs)

For prediction of specific gene and functional classifications, annotations were performed using the BV-BRC v3.30.19a (PATRIC) annotation pipeline and the RAST toolkit (RASTtk). Each sample-assembled contig was submitted to the pipeline along with a unique genome identifier, the genetic code used for translation by 11 Achaea and most bacteria for calling functional traits. Each annotated feature was added to the KEGG database to forecast the dataset's functional assignment. Following annotation, BV-BRC evaluates all genes, including those that are known to exist in the genome, against specialist gene information databases like ARDB, CARD, NDARO, and PATRIC. Management of AMR offers insights on pathogenicity, antibiotic resistance, and k-mers. This makes it easier to find antimicrobial resistance genes and gene counts by utilizing BLASTP and k-mer based approaches to explore metagenomic data for resistance profiles based on high sequence similarity. The frequencies of the various classes of resistance mechanisms were calculated after computing the frequency and percentage distributions of ARGs in the dataset. These frequencies were then combined with the frequencies of the ARG subtypes connected to each class of resistance mechanisms. Calculate the percentage of various antibiotic resistance gene types present on plasmids, chromosomes, and unclassified sequences by combining all metagenomically assembled contigs.

Fig 2. BV-BRC (PATRIC) Functional Annotation Workflow



Genes for Antimicrobial Resistance

Utilizing a carefully selected set of AMR gene variations, PATRIC's Genome Annotation Service employs a k-mer-based technique to find AMR genes[14]. Each functional annotation of an AMR gene is given a drug class, a specific resistant antibiotic, and a mechanism of antibiotic resistance. Despite this, having AMR genes in a genome does not automatically mean that a certain phenotype—antibiotic resistance—exists. Certain AMR pathways must be regarded as crucial, especially those with SNP mutations that give resistance. Fig. 2 displays the AMR genes and related processes found in the genome.

Statistical Analysis

The output data was initially evaluated prior to statistical analysis. For the Plasflow result, a bar chart was used to quantify the ratios of various ARG types identified on plasmid, chromosomal, and unclassified sequences, as well as their densities and relative dispersion across all datasets included. Utilizing the usegalaxy tool, Circos and the Heatmap presentation were created.

Results

Assessments of WGS Metagenomic Assembly

Genome Assembly

The Comprehensive Genome Analysis service received bacterial S1 (ERR3840084), S2 (SRR15244457), S3 (SRR5535749), and S4 (SRR13740644) assembled genomes. Table 1 shows the total number of contigs, the estimated length of the genome, the average G+C content, the N50 length, or the shortest sequence length at 50% of the genome, and the L50 count, or the least number of contigs whose length sum equals the N50.

Table1: Comparison of genome assembly

Samples	Genome Length	Contigs	GC content	contigL50.	Contig N50
S1	4,300,058bp	176	38.22%	16	92,433
S2	44,651,884bp	2,309	41.13%	130	84,738
S3	23,411,304bp	1,263	41.08%	77	75,259
S4	16,671,310bp	1,689	43.96%	137	26,211

Evaluation of Functional Genome Annotations

In the initial stage, all annotated data must be compressed. Following this, each intended gene will be thoroughly evaluated and documented. The RASTtk (RAST tool kit) system generates several features for all MAGs, including CDS, tRNAs, rRNAs, proteins with functional annotations, CRISPR elements, proteins with Enzyme Commission (EC) numbers, proteins mapped to KEGG pathways, proteins with Gene Ontology (GO) assignments, as well as two types of protein families (PLFams and PGFams) identified from the PATRIC database. These characteristics are used to obtain a comprehensive view of the targeted genes.

ARG types are many and diverse, as demonstrated by the metagenomic assembly sample.

The frequency of antibiotic resistance genes in each sample is shown in Fig. 2A. The information came from the BV-BRC rescores reference. The types of ARGs that were abundantly present in all of the samples were identified through metagenomic analysis; the four most prevalent types were aminoglycoside, peptide antibiotic, cycloserine, and elfamycin resistance genes, followed by fluoroquinolone, fosfomycin, rifamycin, fusidic acid, and mupirocidal resistance genes. When compared to peptide antibiotics, aminoglycosides are shown to be more abundant in the prostate cancer sample (SRR1524457) and slightly less abundant in the urinary tract infection sample (SRR5535749). The ARG types of tetracycline and rifamycin were present in roughly equal amounts in the SRR15244457, SRR5535749, and SRR13740644 samples. The sample with the lowest abundance of all ARGs is ERR3840084, however. Triclosan and sulfonamides have the lowest abundances of ARG types. With aminoglycosides being the most prevalent kind, this observation showed an increase and decrease in the abundance of classes of antibiotic-resistant genes between comparing prostate cancer and urinary tract infection microbiome samples.

Fig.2. A. The abundance of ARGs in all samples. Fig. 2.B.

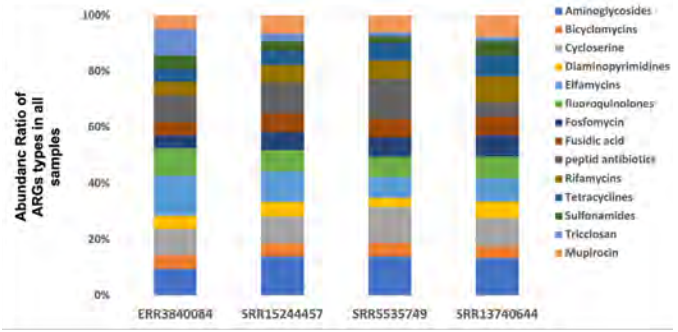
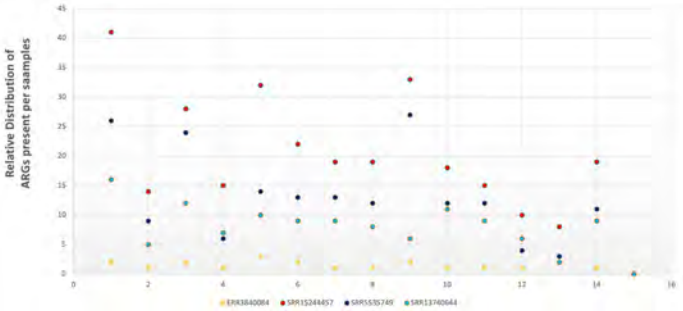


Fig. 2. The ratio of ARGs in each sample. (A) ARG type abundance ratio (B) Sum total abundance normalized relative distribution of the ARG types in each sample.



As it is clear in the presentation of the number of ARGs in each sample, S1 (ERR3840084), the plasma sample of prostate cancer patients includes all kinds of ARGs in the lowest account. However, the other sample of prostate cancer patients from gut microbiome contains the highest number of ARG types.

Fig. 3. A. The heatmap presentation of ARGs for all samples

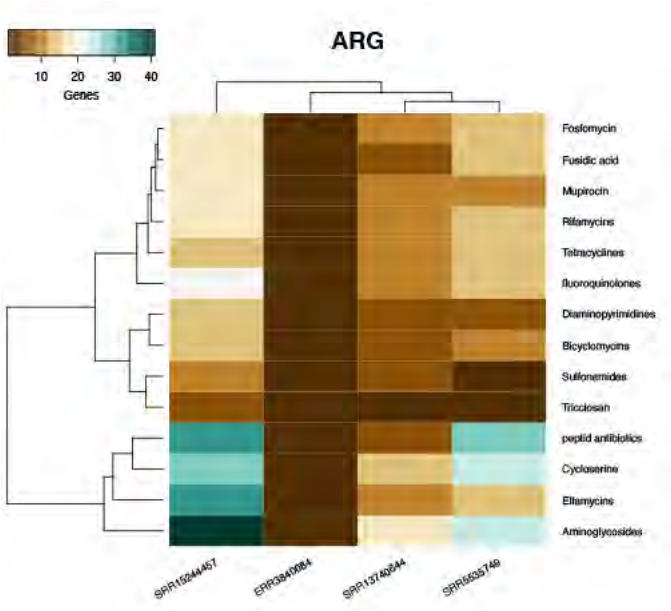
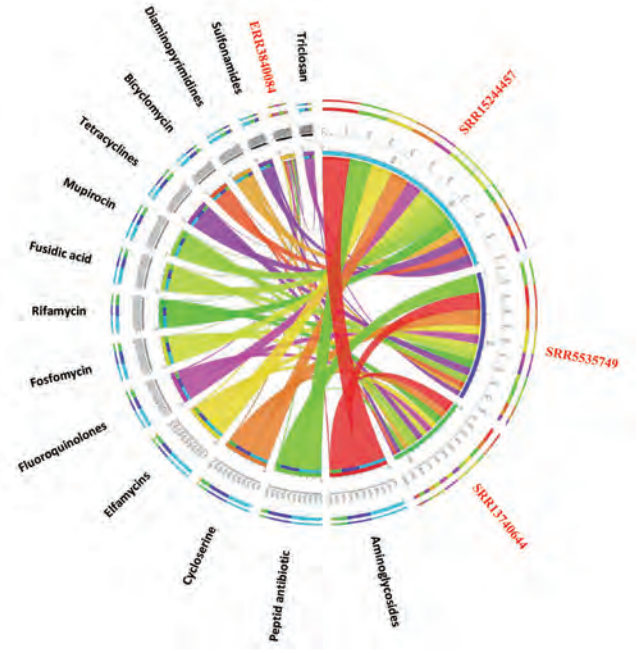


Fig. 3. The heatmap presentation of ARGs for all samples created by the usegalaxy.eu tool. Clustering showed three distinct groupings of antibiotic resistance in the prostate cancer and urinary disease samples. As it is shown UTI and genitourinary patients microbiome samples are close to each other in terms of ARGs in comparison with the prostate cancer samples. Summarly, Aminoglycosides, Elfamycins, Cycloserin, and peptid agents are almost the highest number of ARGs types that S2 and S3 samples include.

Fig 3. B. The Circos presentation of ARGs types for all samples.

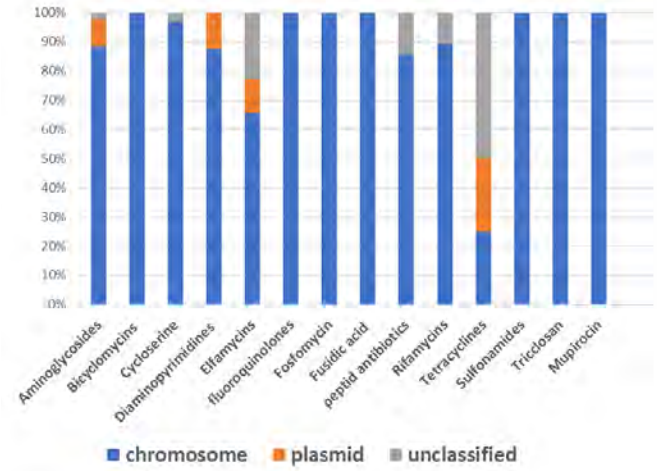


The Circos presentation of ARGs for all samples reveals that all samples include the lowest number of genes for triclosan agent. The S1 sample totally contains the lowest number of ARGs, while the gut microbiome sample of prostate cancer S2 shows the highest number of ARGs. Moreover, it is possible to name the aminoglycosides antibiotics classification, which almost all samples contain a high number of resistance genes against them.

Proportion of ARGs located in Plasmid and Chromosomal Sequence.

All samples received their respective ARG types for chromosomes, plasmids, and unclassified sequences. Chromosome contigs possessed a much higher percentage of ARG types than plasmid and unclassified contigs, according to contig assembly. The usegalaxy.eu tool joined the contig sequences of the prostate cancer samples together. The merged collection containing chromosomal, plasmid, and unclassified sequences was put together for the thorough genome analysis. Tetracycline encoded in plasmid and cycloserine is revealed to be the most dominant ARGs subtypes encoded in chromosomal contigs of prostate cancer samples sequences as a result of thorough analysis of merge contigs of prostate cancer patients. It has been demonstrated that the ARGs for the drugs bicyclomycin, fluoroquinolones, fosfomycin, fusidic acid, sulfonamides, triclosan, and mupirocin are solely encoded on chromosomal DNA.

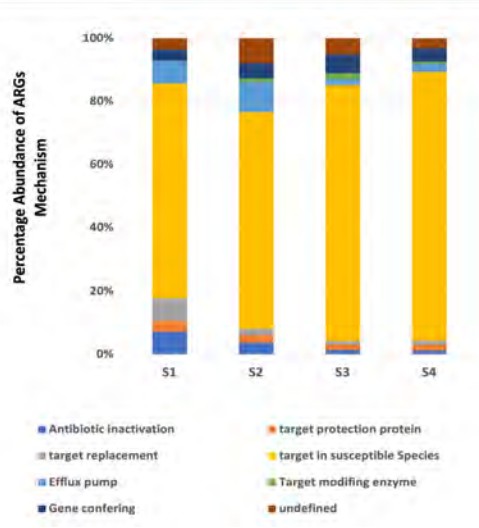
Fig 4. Prostate cancer samples' proportions of various ARG types found on plasmid, chromosomal, and unclassified sequences.



These results demonstrated that different chromosomal and plasmid subtypes carry ARGs in varied ways, and some ARG types appeared to be located on unclassified regions of bacterial DNA.

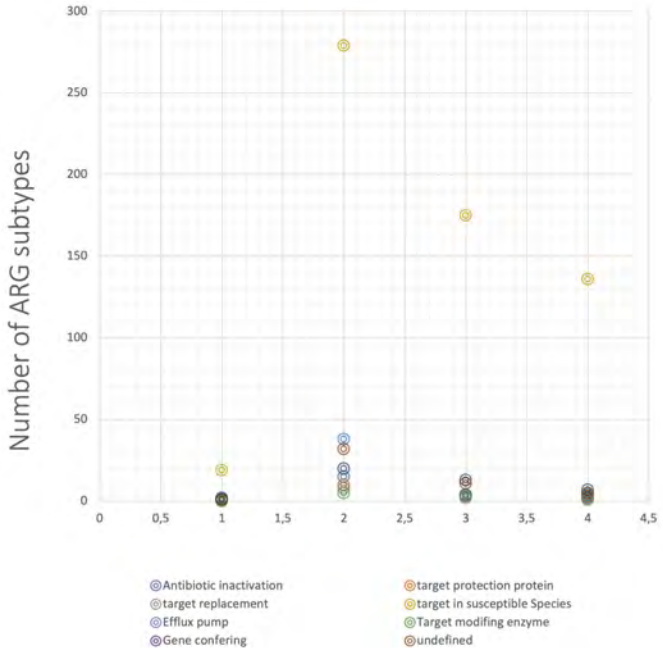
All ARGs Mechanism in this study are divided into seven major categories based on the Blasts results from the CARD, NDARO, and PATRIC AMR-related curation databases: efflux pump, antibiotic inactivation, antibiotic target protection, antibiotic target replacement, antibiotic target in susceptible species, target modifying enzyme, and gene conferring. According to Fig. 5, "undefined" refers to the mechanism's unclassified categories.

Fig.5. A.



As the chart reveal antibiotic target in susceptible species mechanism is the common ARGs mechanism of all prostate and urinary microbiome samples. Hence, sample S2 has the highest percentage with efflux pump, antibiotic inactivation and the unclassified mechanisms. Comprised of the antibiotic target in susceptible species, efflux pump and gene conferring are the other two most abundant in the S2 sample, while gene conferring ARG mechanisms were more prevalent in S3 and S4 too. In order to control their internal environment by eliminating hazardous compounds, bacteria have transport proteins called efflux pumps[15], which are typically linked to multiple drug resistance[16].

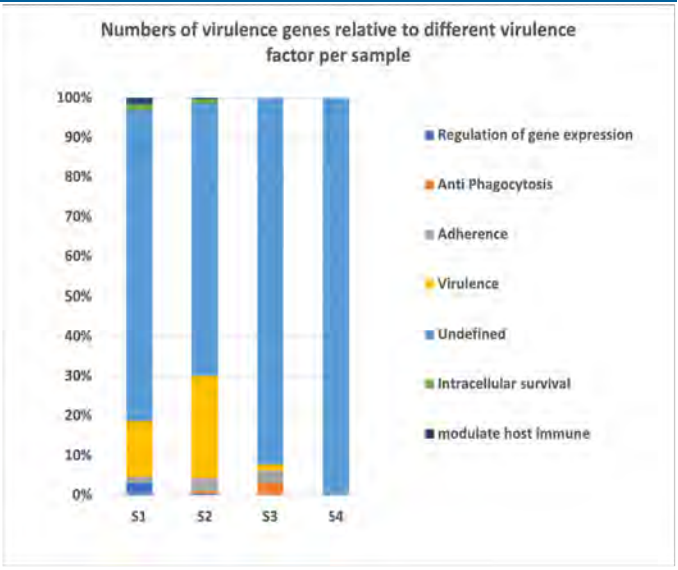
Fig 5. B. ARG resistance mechanism categories: (A) the proportion of each category of resistance mechanisms; (B) the number of subtypes included in each category of resistance mechanisms.



Analysis of Virulence Genes Related to Different Factors in the Samples.

The total number of virulence genes was assessed in this study after comparison with the VFDB, Vectors, and PATRIC_VF databases using the high sequence similarity method with BLAT. Six virulence factors and the undefined factors were the focus of the investigation. According to Fig. 6, Sample S2 possesses the greatest number of virulence genes compared to the other samples. Analysis reveals that Sample S4 contains the largest percentage of undescribed genes. The whole set of virulence genes were uniformly distributed throughout all study samples in the urine, gut, and plasma microbiomes of patients with prostate cancer and other genitourinary diseases. Bacterial adhesion factors are frequently trained cell surface structures that engage with surfaces like MSCRAMMs or Pili[17]. The anti-phagocytosis VF gene was predominately detected in Bacillus bacteria in the S2 and S3 samples. Some pathogenic bacteria are naturally able to resist the bactericidal components of host tissues because of their structural makeup. For instance, the poly-D-glutamate capsule of Bacillus anthracis protects the organisms from cationic proteins (defensins) produced by phagocytes. The virulence (regulation of gene expression) genes for S1 and S2 prostate cancer samples that were isolated from plasma and gut were predominately detected in Escherichia coli and Salmonella enterica. Well-adapted gastrointestinal pathogens like Escherichia coli or Salmonella have the capacity to modulate the expression of a number of genes as they transition from one host flora to another during their passage through the gut, including passing through the gastric barrier and surviving in macrophages or intestinal epithelial cells[18]. The Salmonella intracellular survival gene is the other gene that is shared by the S1 and S2 samples. Salmonella enterica serovar Typhimurium, a common facultative intracellular pathogen, causes millions of cases of food-borne gastroenteritis each year. In vivo, the bacterium can be found in a variety of phagocytic and non-phagocytic cells, and intracellular survival and reproduction are important factors of virulence[19].

Fig. 6. Numbers of virulence genes relative to different virulence factor per sample.

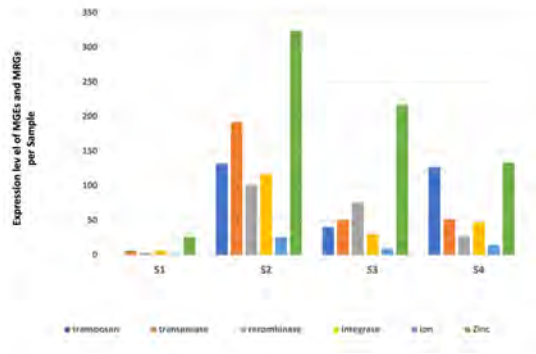


Expression level of Mobile Gene Elements (MGE) and Metal Resistance Genes (MRG)

From the full WGS assembly samples, four major categories of mobile gene elements and two metal resistance genes have been identified, and their expression rates per sample were assessed. Sample S2 is shown to have the most dominantly expressed MGE, with transposase being the most expressed, followed by transposon within the same sample compared to others. Sample S4, shown to follow S2 in expression of the transposon, is the most expressed MGE compared to others. MRGs were found in all samples, with zinc predominating in samples S2, S3, S4, and S1, respectively. Ion appears to be the MGE and MRG gene with

the lowest percentage of expression in all samples. In the S2 sample, ion is seen to be highly expressed. In terms of MGE and MRG, S2 also showed the highest expiry. Presentation of MGE and MRG for all samples is shown in Fig. 7.

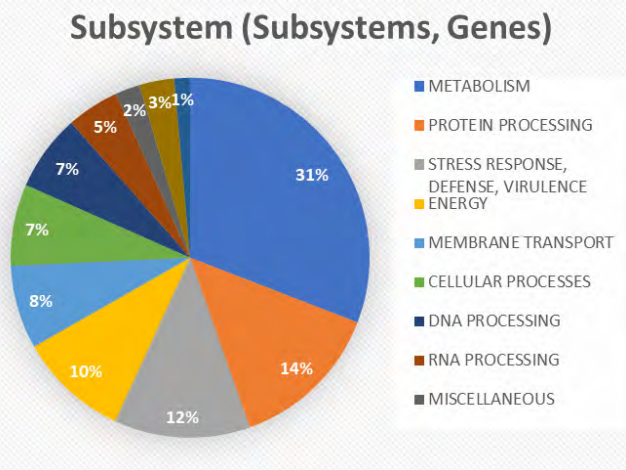
Fig. 7. MGE and MRG Expression Levels in Each Assembly Sample.



Microbial Subsystems and Gene Classifications for all samples.

A subsystem is a group of proteins that work together to carry out a particular biological process or structural complex[20]; the BV-BRC annotation also examines the subsystems that are peculiar to each genome. In the Fig. 8 presentation above, a summary of the subsystems for each sample is presented.

Fig 8. Subsystems, Genes Classifications for all samples.



Discussion

One of the newest subfields in molecular microbiology involves research on microbial communities at the ecosystem level (such as the microbiome). NGS-based panbacterial metagenomic data analysis requires carefully curated, top-notch reference datasets in order to produce pertinent results. The BV-BCR database is dedicated to provide the scientific community with an extensive knowledge base. The S2 sequence was of good quality in comparison to other sequences as a result of thorough genome analyses for all samples. While urine microbiome samples do not contain it, prostate cancer samples show intracellular survival, one of the significant virulence factors that is a common gene of Salmonella.

Against prokaryotic rivals and eukaryotic hosts, pathogenic bacteria have developed a number of strategies to invade, pierce, and disrupt the host immune response. Empirical studies have demonstrated that many bacterial VFs are complex proteins with a variety of functions. For instance, Listeria monocytogenes, a facultative intracellular pathogen, has a major virulence component called Listeriolysin O (LLO) that plays a role in numerous stages of the intracellular lifecycle and has unique characteristics[21]. The production of pili in uropathogenic bacteria is, however, suppressed by pilicides from the class of bicyclic 2-pyridones, which are effective against type I pili generated by Enterobacteriaceae[22]. A different class of anti-colonization drugs includes amyloid aggregate inhibitors[23]. The study's findings indicate that all samples showed significant numbers of resistance genes for aminoglycoside antibiotics, whereas none of the samples showed high numbers of resistance genes for triclosan antibiotic families. Aminoglycosides are bactericidal, broad-spectrum antibiotics that are routinely administered to children, primarily to treat conditions caused by Gram-negative bacteria. Some of the aminoglycosides include gentamicin, amikacin, tobramycin, neomycin, and streptomycin.

Conclusions

The presence of highly resistant genes in the urinary microbiome suggests that aminoglycoside antibiotics may not be the best option for treating genitourinary infections, based on our findings. Interestingly, our analysis of the prostate cancer microbiome samples (S1, S2) revealed triclosan antibiotics to possess the lowest number of resistance genes. Further

studies could investigate whether triclosan could prove effective in treating prostate cancer. Looking ahead, the BV-BRC database will likely expand its knowledge on bacterial pathogens to better serve global users.alue.

References

1- Greene, J. M., Collins, F., Lefkowitz, E. J., Roos, D., Scheuermann, R. H., Sobral, B., ... & Di Francesco, V. (2007). National Institute of Allergy and Infectious Diseases bioinformatics resource centers: new assets for pathogen informatics. Infection and immunity, 75(7), 3212-3219.

2- Amos B., Aurrecoechea C., Barba M., Barreto A., Basenko E.Y., Belnap R., Blevins A.S., Böhme U., Brestelli J., Brunk B.P. VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center. Nucleic Acids Res. 2022; 50:D898–D911. [PMC free article] [PubMed] [Google Scholar]

3- Zhang Y., Aevermann B.D., Anderson T.K., Burke D.F., Dauphin G., Gu Z., He S., Kumar S., Larsen C.N., Lee A.J. Influenza Research Database: An integrated bioinformatics resource for influenza virus research. Nucleic Acids Res. 2017; 45: D466–D474. [PMC free article] [PubMed] [Google Scholar]

4- Pickett B.E., Sadat E.L., Zhang Y., Noronha J.M., Squires R.B., Hunt V., Liu M., Kumar S., Zaremba S., Gu Z. ViPR: an open bioinformatics database and analysis resource for virology research. Nucleic Acids Res. 2012; 40:D593–D598. [PMC free article] [PubMed] [Google Scholar]

5- Davis J.J., Wattam A.R., Aziz R.K., Brettin T., Butler R., Butler R.M., Chlenski P., Conrad N., Dickerman A., Dietrich E.M. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. Nucleic Acids Res. 2020; 48:D606–D612. [PMC free article] [PubMed] [Google Scholar].

6- Snyder E., Kampanya N., Lu J., Nordberg E.K., Karur H., Shukla M., Soneja J., Tian Y., Xue T., Yoo H.. PATRIC: the VBI pathosystems resource integration center. Nucleic Acids Res. 2007; 35:D401–D406. [PMC free article] [PubMed] [Google Scholar].

7- McNeil L.K., Reich C., Aziz R.K., Bartels D., Cohoon M., Disz T., Edwards R.A., Gerdes S., Hwang K., Kubal M.. The National Microbial Pathogen Database Resource (NMPDR): a genomics platform based on subsystem annotation. Nucleic Acids Res. 2007; 35:D347–D353. [PMC free article] [PubMed] [Google Scholar]

8- Overbeek R., Olson R., Pusch G.D., Olsen G.J., Davis J.J., Disz T., Edwards R.A., Gerdes S., Parrello B., Shukla M. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014; 42:D206–D214. [PMC free article] [PubMed] [Google Scholar]

9- Olson RD, Assaf R, Brettin T, Conrad N, Cucinell C, Davis JJ, Dempsey DM, Dickerman A, Dietrich EM, Kenyon RW, Kuscuoglu M, Lefkowitz EJ, Lu J, Machi D, Macken C, Mao C, Niewiadomska A, Nguyen M, Olsen GJ, Overbeek JC, Parrello B, Parrello V, Porter JS, Pusch GD, Shukla M, Singh I, Stewart L, Tan G, Thomas C, VanOeffelen M, Vonstein V, Wallace ZS, Warren AS, Wattam AR, Xia F, Yoo H, Zhang Y, Zmasek CM, Scheuermann RH, Stevens RL. Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res. 2023 Jan 6;51(D1):D678-D689. doi: 10.1093/nar/gkac1003. PMID: 36350631; PMCID: PMC9825582.

10- Brettin T., Davis J.J., Disz T., Edwards R.A., Gerdes S., Olsen G.J., Olson R., Overbeek R., Parrello B., Pusch G.D. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci. Rep. 2015; 5:8365. [PMC free article] [PubMed] [Google Scholar]

11- Overbeek R., Begley T., Butler R.M., Choudhuri J.V., Chuang H.-Y., Cohoon M., de Crécy-Lagard V., Diaz N., Disz T., Edwards R. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 2005; 33:5691–5702. [PMC free article] [PubMed] [Google Scholar].

12- Alcock B.P., Raphenya A.R., Lau T.T., Tsang K.K., Bouchard M., Edalatmand A., Huynh W., Nguyen A.-L.V., Cheng A.A., Liu S. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 2020; 48:D517–D525. [PMC free article] [PubMed] [Google Scholar]

13- Liu B., Zheng D., Zhou S., Chen L., Yang J. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 2022; 50:D912–D917. [PMC free article] [PubMed] [Google Scholar]

14- Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, et al. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45:D535-D542.].

15- Webber, M. A., & Piddock, L. J. V. (2003). The importance of efflux pumps in bacterial antibiotic resistance. Journal of antimicrobial chemotherapy, 51(1), 9-11.

16- Sun, J. R., Perng, C. L., Lin, J. C., Yang, Y. S., Chan, M. C., Chang, T. Y., ... & Chiueh, T. S. (2014). AdeRS combination codes differentiate the response to efflux pump inhibitors in tigecycline-resistant isolates of extensively drug-resistant Acinetobacter baumannii. European journal of clinical microbiology & infectious diseases, 33, 2141-2147.

17- Jeffery, C. (2018). Intracellular proteins moonlighting as bacterial adhesion factors. AIMS microbiology, 4(2), 362.

18- Thomas MS, Wigneshweraraj S. Regulation of virulence gene expression. Virulence. 2014;5(8):832-4. doi: 10.1080/21505594.2014.995573. PMID: 25603428; PMCID: PMC4601333.

19- Ibarra, J. A., & Steele-Mortimer, O. (2009). Salmonella–the ultimate insider. Salmonella virulence factors that modulate intracellular survival. Cellular microbiology, 11(11), 1579-1586.

20- Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang H-Y, Cohoon M, de Crécy-Lagard V, Diaz N, Disz T, Edwards R. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33:5691-5702.]

21- Nguyen B.N., Peterson B.N., Portnoy D.A. Listeriolysin O: a phagosome-specific cytolysin revisited. Cell. Microbiol. 2019; 21:e12988. [PMC free article] [PubMed] [Google Scholar]

22- Pinkner, J. S., Remaut, H., Buelens, F., Miller, E., Åberg, V., Pemberton, N., ... & Almqvist, F. (2006). Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. Proceedings of the National Academy of Sciences, 103(47), 17897-17902.

23- Cegelski, L., Pinkner, J. S., Hammer, N. D., Cusumano, C. K., Hung, C. S., Chorell, E., ... & Hultgren, S. J. (2009). Small-molecule inhibitors target Escherichia coli amyloid biogenesis and biofilm formation. Nature chemical biology, 5(12), 913-919.

24- Sadowski MC, Pouwer RH, Gunter JH, Lubik AA, Quinn RJ, Nelson CC. The fatty acid synthase inhibitor triclosan: repurposing an anti-microbial agent for targeting prostate cancer. Oncotarget. 2014 Oct 15;5(19):9362-81. doi: 10.18632/oncotarget.2433. PMID: 25313139; PMCID: PMC4253440