Leveraging Illumina iSeq100 for Whole Genome Sequencing of *Salmonella Typhi*: a practical approach for resource-limited setting

Nishan Katuwal[1,2,3], Shiva Ram Naga[2,3], Aastha Shrestha[1,2,3], Sabin Bikram Shahi[2,3], Dipesh Tamrakar[2,3,4], Rajeev Shrestha[1,2,3,5*]

1 Molecular and Genome Sequencing Research Lab, Dhulikhel Hospital, Kathmandu University Hospital, Dhulikhel, Nepal

2 Center for Infectious Disease Research and Surveillance, Dhulikhel Hospital, Kathmandu University Hospital, Dhulikhel, Nepal

3 Research and Development Division, Dhulikhel Hospital, Kathmandu University Hospital, Dhulikhel, Nepal

4 Department of Community Medicine, Kathmandu University School of Medical Sciences, Dhulikhel, Nepal

5 Department of Pharmacology, Kathmandu University School of Medical Sciences, Dhulikhel, Nepal

***Corresponding Author**

Prof. Dr. Rajeev Shrestha

email: rajeev.shrestha@kusms.edu.np

Word Count (Manuscript): 2491

Word Count (Abstract): 285

Number of Figures: 2

Number of Tables: 2

ABSTRACT:

Bacterial whole genome sequencing helps to improve our understanding of epidemiology and pathogenesis of bacterial infections and allows comprehensive investigation on virulence, evolution and resistance mechanisms. Nepal, in recent times, has seen some increase in sequencing capabilities but faces numerous hurdles for optimum utilization. However, these hurdles can be alleviated with use of Illumina iSeq100. Therefore, this study aimed at performing whole genome sequencing of bacteria isolated utilizing the iSeq100 platform.

For this study, 6 banked *Salmonella enterica* serovar Typhi bacterial isolates were selected. These isolates were extracted for DNA, confirmed by qPCR and then, their libraries were prepared. The libaries were checked and loaded in Illumina iSeq100 at loading concentration of 200pM. The consensus was generated from the raw genomic data by reference-based assembly, mapping onto S. Typhi CT18. These consensus genomes and coverage parameters were compared to data from using HiSeq. The raw reads were also evaluated using pathogenwatch (v22.3.8) to observe for genotype, mutations and resistance genes.

The coverage parameters (coverage width and depth) of the genomes from this study were compared to same genomes sequenced using Illumina HiSeq. The average coverage width (96.81%) and depth (63.75x) of genomes sequenced in iSeq100 were comparable to that of HiSeq (width: 98.72% and depth: 69.87x). When the genomes sequenced from Illumina iSeq100 and HiSeq were compared, the genotypes detected, number of SNPs and genetic determinants of AMR genes were identical.

The data from bacterial whole genome sequencing using the Illumina iSeq100 is equally informative when compared to other high-end sequencers. Therefore, the primary goal of this study is to advocate for optimum utlisation of iSeq100 while still ensuring a high standard of quality. This optimum utilization will create capacity to fill critical surveillance gaps.

Keywords: iSeq100, Nepal, salmonella, whole genome sequencing

BACKGROUND:

As part of most infection and outbreak management protocols, several phenotypic and molecular methods for pathogen characterization are conventionally utilized to monitor and control the spread of infections. [1] However, these conventional control approaches fail to distinguish between closely related strain, resistance and virulence factors, mostly due to limited genomic resolution. [2] For instance, investigation of antimicrobial resistant (AMR) pathogens via genotypic method can only detect the AMR genes and would require another assay for virulence gene. [3] However, to obtain comprehensive information on phylogenetic information and improve outbreak investigations, full genomic infections are needed. [2]

Bacterial whole genome sequencing (WGS) is the one of the most promising approaches of this development, which helps to improve our understanding of epidemiology and pathogenesis of bacterial infections.[4] The genomic information from WGS allows comprehensive investigation on virulence, evolution and resistance mechanisms. For instance, identification of single nucleotide polymorphisms (SNPs) can differentiate the evolution, with low frequency of SNPs indicating bacteria are genetically similar and recently originated from the same source. [2] Such approaches have known to be used in public health surveillance and control for several bacterial infections due to *Escherichia coli, Campylobacter jejuni, Legionella pneumophila and Mycobacterium tuberculosis*, outbreaks and monitor the source of healthcare associated infections. [5-9]

Furthermore, continuous development in high throughput sequencing have pushed current clinical microbiology field due to its vast potential in identification of infectious agents, detection of pathogenicity, antimicrobial resistance and epidemiological surveillance. [5,10-12] Although, WGS approach can comprehend all genomic information, it is only being utlilised in niches, because clinicians and researchers have shown reluctance, due higher costs, data interpretation and burdensome process of early sequencing technologies. [1,13,14] Further, some sequencing instruments have high operational costs, require large multiplexing for effective and cost continuous sequencing. [15]

Nevertheless, the encouraging prospects are recent advancements in sequencing technologies (for instance, Sequencing by Synthesis from Illumina) and investigation tools have made the platform to have high throughput, increase output, decrease analysis time and reduced cost. [16,17]

Nepal, recently, has also seen some increase in sequencing capabilities, with several sites having some type of next generation sequencer. [18] However, not all these sequencers are at the optimum use due to limitation in capacity, as majority of instruments are either Illumina iSeq100 or MiSeq. However, the previously mentioned barriers can be alleviated through the introduction of Illumina iSeq 100, which is an inexpensive benchtop next-generation sequencer that minimizes the up-front instrument costs while maximizing simplicity of use and capability in obtaining bacterial whole genome. [19,20]

Therefore, this study aimed at performing Whole Genome Sequencing (WGS) of bacteria isolated from Dhulikhel Hospital Kathmandu University Hospital, utilizing the Illumina iSeq100 platform.

113 **METHODS:**

114 *Isolate Selection*

115 For this study, 6 banked *Salmonella enterica* serovar Typhi bacterial isolates were selected
116 randomly from Surveillance for Enteric Fever in Asia Project (SEAP). These glycerol stock
117 isolates were subcultured on MacConkey Agar (Oxoid, Cat: CM0007) and reconfirmed
118 biochemically and serologically. An average of 30 isolated colonies were picked for each
119 isolate.

120 *Processing, DNA Extraction and qPCR*

121 The loop full of colonies were resuspended in 800ul of sterile normal saline, vortexed and
122 centrifuged for 10 minutes at 8000rpm, to obtain cell pellets. The supernatant was discarded
123 and then the pellets were resuspended in 200ul sterile 1XPBS. The pellets were subsequently
124 revotexed and 100ul of the resuspension was taken for extraction.

125 The extraction was done as per manufacturer manual, using Qiagen DNeasy Blood and Tissue
126 Kit. (Qiagen, Cat: 69504), with modification in elution. The final elution was done in 50ul,
127 which was used for re-elution to increase the yield.

128 The S. Typhi isolated were re-confirmed by qPCR targeting Ty21a gene, using forward primer:
129 5'-CGCGAAGTCAGAGTCGACATAG-3', reverse primer 5'-
130 AAGACCTCAACGCCGATCAC-3' and probe [6-FAM]
131 CATTTGTTCTGGAGCAGGCTGACGG [BHQ1a-Q]. The assays were run in BioRad CFX
132 96 Dx qPCR machine with reaction mix that included: 10ul of 2X master mix (Quantabio
133 Perfecta, Cat: 95113-012), 0.8ul of forward primer (400nM), 0.8 ul of reverse primer (400nM),
134 0.4ul of probe (200nM), 4ul of nuclease free water and 4ul of DNA template. The samples will
135 Ct values <20.00 were selected for library preparation.

136 *Library Preparation and Whole Genome Sequencing*

137 The concentration of the samples was checked by Qubit Fluorometer 4 (Invitrogen, Cat:
138 Q33226) using dsDNA HS Kit (Invitrogen, Cat: Q32854). All the samples were diluted to have
139 input amount of 75ng. The libraries were subsequently prepared using NEBNext Ultra II FS
140 DNA Library Preparation Kit (New England Biolabs, Cat: 7805L). The libraries were quality
141 checked for library size and concentration using Agilent Tapestation 4150 using D5000 HS
142 Assay Kit (Agilent Technologies, Cat: 5067-5593). The average library size was 222.5bp. The
143 isolates were loaded in two batches of two and four isolates at loading concentration of 200pM.
144 The sequencing was done using pair ended barcoding primers at 2x146bp.

145 *Assembly, Consensus and Bioinformatical Investigation*

146 The raw genomic data were trimmed for adapters (fastp), assembled (bowtie2, samtools),
147 consensus generated, and coverage was calculated (bamCoverage, samtools depth, awk) and
148 viewed in Integrated Genome Browser. The consensus was built using reference-based
149 assembly and mapping on reference S. Typhi CT18. (Figure 1)

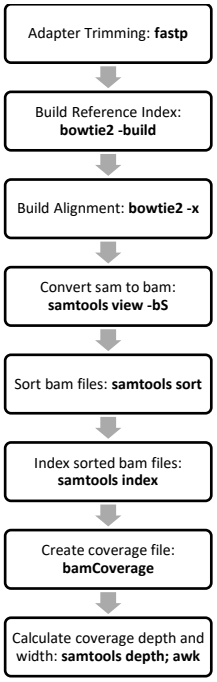


Figure 1: Bioinformatics workflow for generating and evaluating consensus genome

These consensus genomes and coverage parameters were compared to genomic data of same isolates sequenced using HiSeq (Illumina Inc.), which was initially used by SEAP Project. These sequenced genomic data were downloaded from European Nucleotide Archive (ENA).

The raw reads were evaluated using Pathogenwatch (v22.3.8) to observe for genotype detected, number of Single Nucleotide Polymorphisms (SNPs), antimicrobial resistance determinants. Pathogenwatch is a platform which facilitates rapid identification of genomic markers of antimicrobial resistance (AMR) and includes latest analytics for typing along with epidemiological contextualization with public genomic data. [21]

Data Availability

The genomic data, from Illumina iSeq100, can be obtained from National Center for Biotechnology Information's BioProject PRJNA1110744 (Accession: SRX24530188 to SRX24530193).

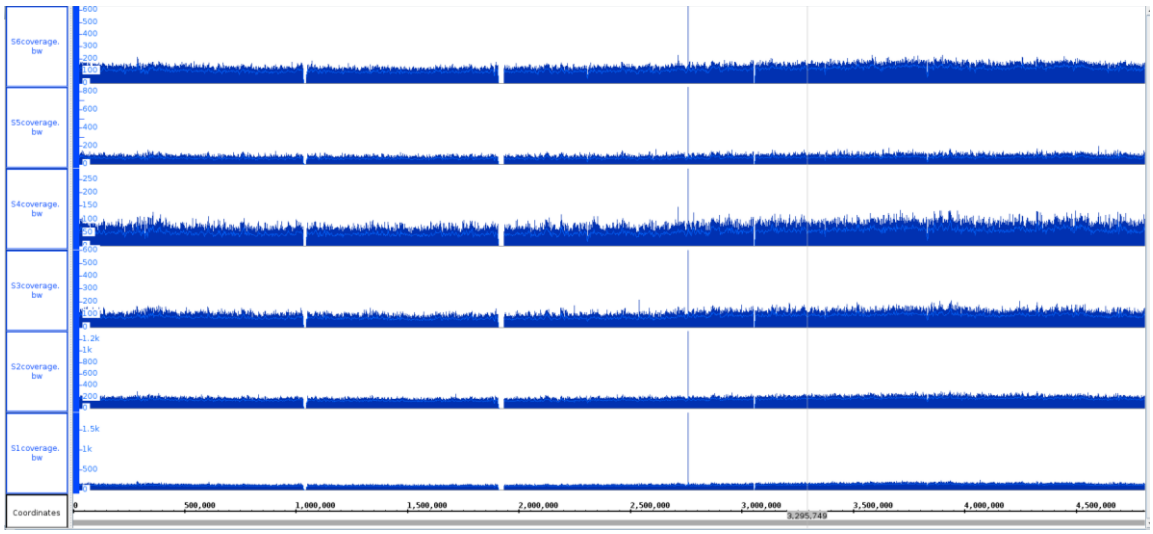
173 **RESULTS**:

174 *Extraction and qPCR*

175 The S. Typhi genomic DNA when evaluated for Ty21a gene, had average Ct value of 15.14,
176 while the concentrations were variable, ranging from 1.58ng/ul to 20.6ng/ul.

177 *Next Generation Sequencing*

178 After the reference-based assembly, the Salmonella genomes were mapped over the S. Typhi
179 CT18 strain in Integrated Genome Browser, with gaps observed in from ~1,033,000bp to
180 1,047,000bp and ~1,908,000bp to ~1,933,000bp and ~3,053,000bp to ~3,059000bp. (Figure 2)

181 

182 **Figure 2:** Coverage Visualization of S. Typhi genomes (S1 to S6) against S. Typhi CT18 strain
183 reference genome, from Integrated Genome Browser

184 *Coverage Parameters*

185 The coverage parameters (coverage width and depth) of the genomes from this study were
186 compared to same genomes sequenced using Illumina HiSeq. The average coverage width
187 (96.81%) and depth (63.75x) of genomes sequenced in iSeq100 were comparable to that of
188 HiSeq (width: 98.72% and depth: 69.87x).

189 **Table 1**: Comparison of coverage depth and width between Illumina HiSeq and Illumina
190 iSeq100

	Illumina iSeq100		**Illumina HiSeq**	
Sample	**Average Coverage Depth**	**Coverage Width (%) at 20x**	**Average Coverage Depth**	**Coverage Width (%) at 20x**
S1	72.23x	98.91	50.43x	98.90
S2	101.91x	98.93	92.77x	99.97
S3	52.40x	98.41	65.47x	98.92

S4	34.23x	89.04	88.49x	98.94
S5	46.67x	97.00	46.81x	97.05
S6	75.02x	98.58	75.29x	98.58

Comparison of Genotype, AMR genes and SNPs

When the genomes sequenced from Illumina iSeq100 and HiSeq were compared, the genotypes detected, number of SNPs and genetic determinants of AMR genes were identical.

Table 2: Comparison of genetic determinants, genotype and SNPs among genomes, evaluated by Pathogenwatch scheme (v22.3.8) for *Salmonella enterica* serovar Typhi (*S*. Typhi)

	Illumina iSeq100			**Illumina HiSeq**		
Sample	**Genetic Determinants of AMR Genes**	**Genotype**	**SNPs**	**Genetic Determinants of AMR Genes**	**Genotype**	**SNPs**
S1	gyrA_S83F	4.3.1.2	87	gyrA_S83F	4.3.1.2	87
S2	gyrA_S83F	4.3.1.2	87	gyrA_S83F	4.3.1.2	87
S3	blaTEM-1D, catA1, gyrA_D87N, sul1; sul2, dfrA7	4.3.1.1	87	blaTEM-1D, catA1, gyrA_D87N, sul1; sul2, dfrA7	4.3.1.1	87
S4	blaTEM-1D, catA1, gyrA_S83F, sul1; sul2, dfrA7	4.3.1.1	87	blaTEM-1D, catA1, gyrA_S83F, sul1; sul2, dfrA7	4.3.1.1	87
S5	blaTEM-1D, catA1, gyrA_S83F, sul1; sul2, dfrA7	4.3.1.1	87	blaTEM-1D, catA1, gyrA_S83F, sul1; sul2, dfrA7	4.3.1.1	87
S6	blaTEM-1D, catA1, gyrA_S83F, sul1; sul2, dfrA7	4.3.1.1	87	blaTEM-1D, catA1, gyrA_S83F, sul1; sul2, dfrA7	4.3.1.1	87

DISCUSSION:

Whole Genome Next Generation Sequencing

In clinical microbiology, conventional pathogen identification methods are slow, labor and resource intensive despite having good efficacy. [22,23] These approaches cannot act as standalone due to limited genomic resolution. Since late 90s, DNA fingerprinting was being implemented for identification for serotypes of Mycobacterium Tuberculosis.[24] This was followed by shift to genomic approaches such as multi-locus sequence typing (MLST) and the variable number of tandem repeats (VNTR). [25,26]

The recent advancement in high throughput genomic approaches have transformed clinical microbiology, offering unparalleled capabilities in pathogen detection and characterization. They have known to provide faster and comprehensive characterization of pathogens. [27,28] In this regards, whole genome sequencing, in particular, gained traction for comprehensive investigation of bacteria. [29] This approach due to its comprehensiveness on pathogen biology, drug resistance, transmission, evolution, community profiling, clinical metagenomics and pathogen discovery are already transforming the research landscape in microbiology.[30] This is evident by the large number of whole genomes stored in public repositories. [31,32] Further, in the context of outbreak management, traditional methods often fall short in distinguishing closely related outbreak strains while WGS can facilitate rapid and accurate identification of virulence factors of the pathogen and can also identify the transmission pathway within a population and provide information on the probable source. [33] This has been more evident by its use in COVID-19 pandemic. [34,35]

However, despite its potential, the widespread adoption of WGS has been hindered by various challenges, including high costs, time-consuming protocols, and data interpretation complexities. [36]

This study, therefore, addresses the limitations by using a low-cost next generation sequencer, Illumina iSeq100 for WGS of *Salmonella Typhi.* The iSeq100 represents a significant advancement in sequencing technology, offering a cost-effective and user-friendly solution that overcomes many barriers associated with traditional NGS platforms. [20] This platform has been known to be more useful when used with amplicon-based sequencing wet-lab approach. [37-39]

Coverage Parameters

The coverage parameters of our approach were high (63.75x) and comparable to results from Illumina HiSeq (69.87x), which we agree could depend on the pooling. Though the threshold on the coverage depth depends upon the target pathogen, studies have estimated that that 40X are required for high-quality detection of virulence genes, while 21 bp reads at 4X coverage are sufficient for target gene detection and coverage width should preferably be between 50x – 100x. [40-42] In addition, in another study, for antibiotic resistance gene, at the read level, is considered present with 20x-30x coverage depth. [43,44] A study that utilized iSeq100 for WGS of bacterial isolates observed coverage depth of 17x - 149x with average of 59x, which is slightly lower than our observation. [19]

Similarly, high coverage width of at least 90% is recommended, which is well below the observed average coverage width of 96.81% of this study. [45] This high coverage width is

required for optimum quality data during bioinformatics analysis, as this threshold helps achieve correct gene identification. [44]

The genome gaps were observed in the regions of ~1,033,000bp to 1,047,000bp and ~1,908,000bp to ~1,933,000bp and ~3,053,000bp to ~3,059,000bp, which correspond to STY_RS04890 gene (DUF2213 domain-containing protein), WP_001197078.1 (baseplate J/gp47 family protein), STY_RS09550 (phage minor head protein), STY_RS09735 (tyrosine-type recombinase/integrase), STY_RS15100 (DUF1460 domain-containing protein) and STY_RS15130 (site-specific integrase). [46] These gaps could be a result of genomic rearrangements, horizontal gene transfer as well as sequencing and alignment errors and presence of low complexity regions. [47]

Pooling and cost calculation

The pooling of isolates with varying genome size results in unequal sample coverage even with equimolar pooling, while some protocols do not mandatorily require equimolar pooling. [19, 48] Further, pooling also depends on the complexity of the genome. For instance, though Escheria coli and Salmonella enterica have similar genome size, E. coli is more complex and has variable accessory genome. Therefore, comparatively larger number of S. enterica genomes can be pooled. [48] However, the cost per run also depends on pooling. The number of samples pooled would be inversely proportional to cartridge/flow cell requirement, while the amount of library preparation reagents remains the same. The average cost of run per gb is $60.65 for HiSeq (using HiSeq SBS Kit V4 kit - 250 cycles PE 2x125bp, which has now been discontinued), while for iSeq would be ~$130 per gb when 4 samples are pooled and ~$87 when 6 samples are pooled. [20, 49] Unfortunately, we are not sure what number of samples were pooled for sequencing in HiSeq, to make more informed comparisons. During the discussion of low cost NGS, Oxford Nanopore provides great argument. The flongle, which gives output upto 2Gb, costs ~$90, while iSeq100 cartridge/flowcell costs ~$630 with 1.2Gb output. [48] However, the shorter shelf life of flow cell and lower accuracy of individuals reads are counter intuitive.

Comparing genotypes and genomic determinants of AMR genes

In this study, both genotypes (4.3.1.1 and 4.3.1.2) belonging to lineage I and II, respectively, of H58 strain, were detected. Both strains were first detected in a pediatric study conducted in Kathmandu.[50] Furthermore, the point mutations (S83F and D87N) were observed in quinolone resistance determining region of gyrA gene, which is associated with fluoroquinolone resistance. [51-53] Similarly, mutations were observed in genes conferring resistance to ampicillin (blaTEM), chloramphenicol (catA1) and resistance to co-trimoxazole (sul) and resistance to trimethoprim (dfrA7). [54-57] In this study, the genotypes, genomic determinants and number of SNPs among the S. Typhi genomes were identical, as expected, in genomes sequenced from iSeq100 and HiSeq. This adds to the fact that the genomes sequenced have similar quality.

Our findings have demonstrated that iSeq platform is able to generate high-quality, accurate data for bacterial WGS, and it is a cost-effective addition to the AMR surveillance armamentarium. The genomes generated from iSeq have comparable coverage parameters to those obtained from more established commercial sequencing platforms like the Illumina HiSeq, which are not available, nor viable (due to its large number of sample requirements) in

Nepal. The findings of this study support premise that iSeq100 is well suited to small laboratories with low sample throughput, which is mostly the case in Nepal as it Nepal has small molecular and genomic research ecosystem.

CONCLUSION

The data from bacterial whole genome sequencing using the Illumina iSeq100 is equally informative when compared to other high-end sequencers. Therefore, the primary goal of this study is to advocate for optimum utlisation of iSeq100 while still ensuring a high standard of quality. This optimum utilization will create capacity to fill critical surveillance gaps. We believe that Illumina iSeq, through primarily designed for microbial metagenomics, is able to perform whole genome sequencing of bacteria when combined with a modified wet-lab protocol. This approach will also provide high-quality genomes for other pathogens of human health importance, with some limitations on throughput.

Looking ahead, the findings of this study pave the way for broader applications of the Illumina iSeq100 platform in clinical microbiology, outbreak surveillance, and antimicrobial resistance monitoring. Future research should focus on optimizing sample preparation protocols, expanding sequencing capacity, and integrating bioinformatics tools to further enhance the utility and accessibility of WGS technologies in pathogen investigation.

DECLARATIONS

Ethics approval and consent to participate

This study investigated banked bacterial isolates and did not contact human subjects or obtain data associated with them. The ethical approval was obtained from Institutional Review Committee at Kathmandu University School of Medical Sciences.

Availability of Data and Materials

The raw reads are available as an NCBI BioProject PRJNA1110744 (Accession: SRX24530188 to SRX24530193).

Competing interests

The authors declare no competing interests.

Funding

This study was independently funded and supported by Center for Infectious Disease Research and Surveillance, Dhulikhel Hospital Kathmandu University Hospital.

Acknowledgements

We gratefully thank Jason R. Andrews and the SEAP team for providing accessions numbers (European Nucleotide Archive) for the *S. Typhi* isolates. We are also thankful to all the research staffs at Center for Infectious Disease Research and Surveillance, Dhulikhel Hospital Kathmandu University Hospital.

REFERENCES

1. Vourli, S., Kontos, F., Pournaras, S. (2021). WGS for Bacterial Identification and Susceptibility Testing in the Clinical Lab. In: Moran-Gilad, J., Yagel, Y. (eds) Application and Integration of Omics-powered Diagnostics in Clinical and Public Health Microbiology. Springer, Cham. https://doi.org/10.1007/978-3-030-62155-1_3
2. Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, Wertheim HFL. Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clin Microbiol Rev. 2017 Oct;30(4):1015-1063. doi: 10.1128/CMR.00016-17. Erratum in: Clin Microbiol Rev. 2017 Nov 1;31(1):e00082-17. doi: 10.1128/CMR.00082-17.
3. Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A. 2014. Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. J Clin Microbiol **52**:2365–2370. doi: 10.1128/JCM.00262-14.
4. Garaizar J, Laorden L. Bacterial Genomics and Epidemiology. Microorganisms. 2023 May 29;11(6):1428. doi: 10.3390/microorganisms11061428. PMID: 37374930; PMCID: PMC10303121.
5. Gautam SS, Kc R, Leong KW, Mac Aogáin M, O'Toole RF. A step-by-step beginner's protocol for whole genome sequencing of human bacterial pathogens. J Biol Methods. 2019 Mar 15;6(1):e110. doi: 10.14440/jbm.2019.276. PMID: 31453259; PMCID: PMC6706130.
6. Grad YH, Lipsitch M, Feldgarden M, Arachchi HM, Cerqueira GC, et al. (2012) Genomic epidemiology of the Escherichia coli O104:H4 outbreaks in Europe, 2011. Proc Natl Acad Sci USA 109: 3065-3070. doi: 10.1073/pnas.1121491109.
7. Llarena A, Taboada E, Rossi M. (2017) Whole-Genome Sequencing in Epidemiology of Campylobacter jejuni Infections. J Clin Microbiol 55: 1269-1275. doi: 10.1128/JCM.00017-17.
8. Reuter S, Harrison TG, Köser CU, Ellington MJ, Smith GP, et al. (2013) A pilot study of rapid whole-genome sequencing for the investigation of a Legionella outbreak. BMJ Open 3: e002175. doi: 10.1136/bmjopen-2012-002175.
9. Gautam SS, Mac Aogáin M, Cooley LA, Haug G, Fyfe JA, et al. (2018) Molecular epidemiology of tuberculosis in Tasmania and genomic characterisation of its first known multi-drug resistant case. PLoS One 13: e0192351. doi: 10.1371/journal.pone.0192351.
10. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, et al. (2015) Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. Lancet Infect Dis 15: 1193-1202. doi: 10.1016/S1473-3099(15)00062-6.
11. Takiff HE, Feo O. (2015) Clinical value of whole-genome sequencing of Mycobacterium tuberculosis. Lancet Infect Dis 15: 1077-1090. doi: 10.1016/S1473-3099(15)00071-7.
12. Harris SR, Cartwright EJP, Török ME, Holden MTG, Brown NM, et al. (2012) Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. Lancet Infect Dis 13: 130-136. doi: 10.1016/S1473-3099(12)70268-2.)
13. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman

KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, et al.. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380.

14. Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A, Johnson SM. 2008. A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. Genome Res 18:1051–1063. doi: 10.1101/gr.076463.108.
15. Christensen KD, Dukhovny D, Siebert U, Green RC. Assessing the Costs and Cost-Effectiveness of Genomic Sequencing. J Pers Med. 2015 Dec 10;5(4):470-86. doi: 10.3390/jpm5040470. PMID: 26690481; PMCID: PMC4695866.
16. Harris SR, Cartwright EJ, Török ME, Holden MT, Brown NM, Ogilvy-Stuart AL, Ellington MJ, Quail MA, Bentley SD, Parkhill J, Peacock SJ. 2013. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant Staphylococcus aureus: a descriptive study. Lancet Infect Dis 13:130–136. doi: 10.1016/S1473-3099(12)70268-2.
17. Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, Holden MTG, Dougan G, Bentley SD, Parkhill J, Peacock SJ. 2012. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathog 8:e1002824. doi: 10.1371/journal.ppat.1002824.
18. WHO International. Building In-country Capacity for Pathogen Genetic Sequencing in Nepal. Available at : https://www.who.int/nepal/news/detail/13-12-2021-building-in-country-capacity-for-pathogen-genetic-sequencing-in-nepal. Accessed on: 6 September 2024.
19. Bruzek S, Vestal G, Lasher A, Lima A, Silbert S. Bacterial Whole Genome Sequencing on the Illumina iSeq 100 for Clinical and Public Health Laboratories. J Mol Diagn. 2020 Dec;22(12):1419-1429. doi: 10.1016/j.jmoldx.2020.09.003. Epub 2020 Sep 19. PMID: 32961318.
20. Illumina iSeq100 instrument manual and product description. Available at: https://www.illumina.com/systems/sequencing-platforms/iseq.html. Accessed on: 7 September 2024
21. Argimón, S., Yeats, C.A., Goater, R.J. et al. A global resource for genomic predictions of antimicrobial resistance and surveillance of Salmonella Typhi at pathogenwatch. Nat Commun **12**, 2879 (2021). https://doi.org/10.1038/s41467-021-23091-2
22. Allen, J.D.; Ross, T.M. H3N2 influenza viruses in humans: Viral mechanisms, evolution, and evaluation. Hum. Vaccines Immunother. **2018**, 14, 1840–1847
23. Vashisht V, Vashisht A, Mondal AK, Farmaha J, Alptekin A, Singh H, Ahluwalia P, Srinivas A, Kolhe R. Genomics for Emerging Pathogen Identification and Monitoring: Prospects and Obstacles. BioMedInformatics. 2023; 3(4):1145-1177. https://doi.org/10.3390/biomedinformatics3040069
24. Guthrie, J.L.; Gardy, J.L. A brief primer on genomic epidemiology: Lessons learned from Mycobacterium tuberculosis. Ann. N. Y. Acad. Sci. **2017**, 1388, 59–77
25. Maiden MC, Van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol. 2013;11(10):728–36.
26. Ramisse V, Houssu P, Hernandez E, Denoeud F, Hilaire V, Lisanti O, Ramisse F, Cavallo JD, Vergnaud G. Variable number of tandem repeats in Salmonella enterica subsp. enterica for typing purposes. J Clin Microbiol. 2004 Dec;42(12):5722–30.

27. Gardy, J.L.; Johnston, J.C.; Sui, S.J.H.; Cook, V.J.; Shah, L.; Brodkin, E.; Rempel, S.; Moore, R.; Zhao, Y.; Holt, R.; et al. Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. N. J. Med. 2011, 364, 730–739.
28. Ko, K.K.; Chng, K.R.; Nagarajan, N. Metagenomics-enabled microbial surveillance. Nat. Microbiol. 2022, 7, 486–496.

29. Parkhill J, Dougan G, James K, Thomson N, Pickard D, Wain J, Churcher C, Mungall K, Bentley S, Holden M. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature. 2001;413(6858):848–52.)
30. Pallen MJ. Et al. High-throughput sequencing and clinical microbiology: progress, opportunities and challenges. Current Opinion in Microbiology. 2010. 13(5). Doi: https://doi.org/10.1016/j.mib.2010.08.003
31. Supratim Mukherjee, Dimitri Stamatis, Cindy Tianqing Li, Galina Ovchinnikova, Jon Bertsch, Jagadish Chandrabose Sundaramurthi, Mahathi Kandimalla, Paul A. Nicolopoulos, Alessandro Favognano, I-Min A. Chen , Nikos C. Kyrpides and T.B.K. Reddy. Twenty-five years of Genomes OnLine Database (GOLD): data updates and new features in v.9. Nucl. Acids Res. (2022) doi: doi.org/10.1093/nar/gkac974
32. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, Connor R, Funk K, Kelly C, Kim S, Madej T, Marchler-Bauer A, Lanczycki C, Lathrop S, Lu Z, Thibaud-Nissen F, Murphy T, Phan L, Skripchenko Y, Tse T, Wang J, Williams R, Trawick BW, Pruitt KD, Sherry ST. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-D26. doi: 10.1093/nar/gkab1112. PMID: 34850941; PMCID: PMC8728269.
33. Gilchrist CA, Turner SD, Riley MF, Petri WA Jr, Hewlett EL. Whole-genome sequencing in outbreak analysis. Clin Microbiol Rev. 2015 Jul;28(3):541-63. doi: 10.1128/CMR.00075-13. PMID: 25876885; PMCID: PMC4399107.
34. Berggreen H, Løvestad AH, Helmersen K, Jørgensen SB, Aamot HV. Lessons learned: use of WGS in real-time investigation of suspected intrahospital SARS-CoV-2 outbreaks. J Hosp Infect. 2023 Jan;131:81-88. doi: 10.1016/j.jhin.2022.10.003. Epub 2022 Oct 29. PMID: 36404573; PMCID: PMC9617632.
35. Francis RV, Billam H, Clarke M, Yates C, Tsoleridis T, Berry L, Mahida N, Irving WL, Moore C, Holmes N, Ball JK, Loose M, McClure CP; COVID-19 Genomics UK (COG-UK) consortium. The Impact of Real-Time Whole-Genome Sequencing in Controlling Healthcare-Associated SARS-CoV-2 Outbreaks. J Infect Dis. 2022 Jan 5;225(1):10-18. doi: 10.1093/infdis/jiab483. PMID: 34555152; PMCID: PMC8522425.
36. Lucassen A, Houlston RS. The challenges of genome analysis in the health care setting. Genes (Basel). 2014 Jul 22;5(3):576-85. doi: 10.3390/genes5030576. PMID: 25055201; PMCID: PMC4198918.
37. Hourdel V, Kwasiborski A, Balière C, Matheus S, Batéjat CF, Manuguerra JC, Vanhomwegen J, Caro V. Rapid Genomic Characterization of SARS-CoV-2 by Direct Amplicon-Based Sequencing Through Comparison of MinION and Illumina iSeq100™ System. Front Microbiol. 2020 Sep 25;11:571328. doi: 10.3389/fmicb.2020.571328. PMID: 33101244; PMCID: PMC7546329
38. Tamrakar D, Katuwal N, Adhikari N, Madhup SK, Dhimal M, Shrestha R. Genomic analysis of SARS-CoV-2 Circulating during Second and Third Wave of COVID-19 in Nepal. J Nepal Health Res Counc. 2023 Sep 8;21(1):57-62. doi: 10.33314/jnhrc.v21i1.4489. PMID: 37742150.
39. Shrestha R , Katuwal N , Adhikari N , Vanaerschot M , Tamrakar D , Dhimal M , Gyanwali P , Bhattarai S , Madhup SK , Devkota B . Whole Genome Sequence Analysis

to Identify SARS-CoV-2 Variant in Nepal. Kathmandu Univ Med J (KUMJ). 2021 Apr-Jun;19(74):137-142. PMID: 34819443.

40. Lindsey RL, Pouseele H, Chen JC, Strockbine NA, Carleton HA. Implementation of whole genome sequencing (WGS) for identification and characterization of Shiga toxin-producing Escherichia coli (STEC) in the United States. Front Microbiol. 2016;7:1–9. doi: 10.3389/fmicb.2016.00766.

41. Lambert D, Carrillo CD, Koziol AG, Manninger P, Blais BW. GeneSippr: A rapid whole-genome approach for the identification and characterization of foodborne pathogens such as priority Shiga toxigenic Escherichia coli . PLoS One. 2015;10:1–19
42. Sayeb Maroura. Guidance document for WGS-Benchmarking. 2021. EURL CPS. Available at : https://sitesv2.anses.fr/en/system/files/NGS%20EURL%20WG%20-%20Del5%20-%20Benchmarking.pdf. Accessed on: 7 September 2024
43. Sanabria, A.M., Janice, J., Hjerde, E. et al. Shotgun-metagenomics based prediction of antibiotic resistance and virulence determinants in Staphylococcus aureus from periprosthetic tissue on blood culture bottles. Sci Rep 11, 20848 (2021). https://doi.org/10.1038/s41598-021-00383-7
44. Vieira, A.A., Piccoli, B.C., y Castro, T.R. et al. Pipeline validation for the identification of antimicrobial-resistant genes in carbapenem-resistant Klebsiella pneumoniae. Sci Rep **13**, 15189 (2023). https://doi.org/10.1038/s41598-023-42154-6
45. Doyle, R. M. et al. Discordant bioinformatic predictions of antimicrobial resistance from whole-genome sequencing data of bacterial isolates: An inter-laboratory study. Microbial. Genom. **6**(2), 1–13. https://doi.org/10.1099/mgen.0.000335 (2020)
46. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, Churcher C, Mungall KL, Bentley SD, Holden MT, Sebaihia M, Baker S, Basham D, Brooks K, Chillingworth T, Connerton P, Cronin A, Davis P, Davies RM, Dowd L, White N, Farrar J, Feltwell T, Hamlin N, Haque A, Hien TT, Holroyd S, Jagels K, Krogh A, Larsen TS, Leather S, Moule S, O'Gaora P, Parry C, Quail M, Rutherford K, Simmonds M, Skelton J, Stevens K, Whitehead S, Barrell BG. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature. 2001 Oct 25;413(6858):848-52. doi: 10.1038/35101607. PMID: 11677608.
47. Nagarajan, N., Pop, M. Sequence assembly demystified. Nat Rev Genet **14**, 157–167 (2013). https://doi.org/10.1038/nrg3367 AND Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. PLoS One. 2014 Aug 11;9(8):e104984. doi: 10.1371/journal.pone.0104984. PMID: 25110940; PMCID: PMC4128722.
48. Mitchell PK, Wang L, Stanhope BJ, Cronk BD, Anderson R, Mohan S, Zhou L, Sanchez S, Bartlett P, Maddox C, DeShambo V, Mani R, Hengesbach LM, Gresch S, Wright K, Mor S, Zhang S, Shen Z, Yan L, Mackey R, Franklin-Guild R, Zhang Y, Prarat M, Shiplett K, Ramachandran A, Narayanan S, Sanders J, Hunkapiller AA, Lahmers K, Carbonello AA, Aulik N, Lim A, Cooper J, Jones A, Guag J, Nemser SM, Tyson GH, Timme R, Strain E, Reimschuessel R, Ceric O, Goodman LB. Multi-laboratory evaluation of the Illumina iSeq platform for whole genome sequencing of Salmonella, Escherichia coli and Listeria. Microb Genom. 2022 Feb;8(2):000717. doi: 10.1099/mgen.0.000717. PMID: 35113783; PMCID: PMC8942033.
49. La Trobe University Genomics. Available at : https://www.latrobe.edu.au/__data/assets/pdf_file/0005/724748/LTU-NGS-prices-AH.pdf Accessed on: 7 September 2024

50. Holt, K. E. et al. High-throughput bacterial SNP typing identifies distinct clusters of Salmonella Typhi causing typhoid in Nepalese children. BMC Infect. Dis. 2010. 10, 144.
51. Jacoby GA. Mechanisms of resistance to quinolones. Clin Infect Dis. 2005;41:7.
52. Brown JC, Shanahan PM, Jesudason MV, Thomson CJ, Amyes SG. Mutations responsible for reduced susceptibility to 4 quinolones in clinical isolates of multi resistant Salmonella typhi in India. J Antimicrob Chemother. 1996;37:891–900.
53. Wain J, Hoa NT, Chinh NT, Vinh H, Everett MJ, Diep TS, et al. Quinolone resistant Salmonella typhi from Vietnam; molecular basis of resistance and clinical response to treatment. Clin Infect Dis. 1997;25:1404–10.

54. Tamang MD, Oh JY, Seol SY, Kang HY, Lee JC, Lee YC, Cho DT, Kim J. Emergence of multidrug-resistant Salmonella enterica serovar Typhi associated with a class 1 integron carrying the dfrA7 gene cassette in Nepal. Int J Antimicrob Agents. 2007 Oct;30(4):330-5. doi: 10.1016/j.ijantimicag.2007.05.009. Epub 2007 Jul 12. PMID: 17629465.
55. Kongsoi S, Changkwanyeun R, Yokoyama K, Nakajima C, Changkaew K, Suthienkul O, Suzuki Y. Amino acid substitutions in GyrA affect quinolone susceptibility in Salmonella typhimurium. Drug Test Anal. 2016 Oct;8(10):1065-1070. doi: 10.1002/dta.1910. Epub 2015 Oct 30. PMID: 26514939.
56. Yusof NY, Norazzman NII, Zaidi NFM, Azlan MM, Ghazali B, Najib MA, Malik AHA, Halim MAHA, Sanusi MNSM, Zainal AA, Aziah I. Prevalence of Antimicrobial Resistance Genes in Salmonella Typhi: A Systematic Review and Meta-Analysis. Trop Med Infect Dis. 2022 Sep 27;7(10):271. doi: 10.3390/tropicalmed7100271. PMID: 36288012; PMCID: PMC9611315.
57. da Silva KE, Tanmoy AM, Pragasam AK, Iqbal J, Sajib MSI, Mutreja A, Veeraraghavan B, Tamrakar D, Qamar FN, Dougan G, Bogoch I, Seidman JC, Shakya J, Vaidya K, Carey ME, Shrestha R, Irfan S, Baker S, Luby SP, Cao Y, Dyson ZA, Garrett DO, John J, Kang G, Hooda Y, Saha SK, Saha S, Andrews JR. The international and intercontinental spread and expansion of antimicrobial-resistant Salmonella Typhi: a genomic epidemiology study. Lancet Microbe. 2022 Aug;3(8):e567-e577. doi: 10.1016/S2666-5247(22)00093-3. Epub 2022 Jun 21. PMID: 35750070; PMCID: PMC9329132.