Journal of Global Antimicrobial Resistance 31 (2022) 167–174

Contents lists available at ScienceDirect

Journal of Global Antimicrobial Resistance

journal homepage: www.elsevier.com/locate/jgar

Comparative analysis of two next-generation sequencing platforms for analysis of antimicrobial resistance genes

Twinkle Soni[a], Ramesh Pandit[a], Damer Blake[b], Chaitanya Joshi[a], Madhvi Joshi[a,*]

[a] Gujarat Biotechnology Research Centre (GBRC), Department of Science and Technology, Government of Gujarat, Gandhinagar, India.
[b] The Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Hertfordshire, AL9 7TA, United Kingdom

ARTICLE INFO

Article history:
Received 27 June 2022
Revised 10 August 2022
Accepted 23 August 2022
Available online 30 August 2022

Editor: Stefania Stefani

Keywords:
Antimicrobial Resistance
Bioinformatics pipeline
Illumina
Ion torrent
Next-generation sequencing

ABSTRACT

Objectives: The use of antibiotics in human medicine and livestock production has contributed to the widespread occurrence of Antimicrobial Resistance (AMR). Recognizing the relevance of AMR to human and livestock health, it is important to assess the occurrence of genetic determinants of resistance in medical, veterinary, and public health settings in order to understand risks of transmission and treatment failure. Advances in next-generation sequencing technologies have had a significant impact on research in microbial genetics and microbiome analyses. The aim of the present study was to compare the Illumina MiSeq and Ion Torrent S5 Plus sequencing platforms for the analysis of AMR genes in a veterinary/public health setting.

Methods: All samples were processed in parallel for the two sequencing technologies, subsequently following a common bioinformatics workflow to define the occurrence and abundance of AMR gene sequences. The Comprehensive Antibiotic Resistance Database (CARD), QIAGEN Microbial Insight - Antimicrobial Resistance, Antimicrobial resistance database, and Comprehensive Antibiotic Resistance Database developed by CLC bio (CARD-CLC) databases were compared for analysis, with the most genes identified using CARD.

Results: Drawing on these results, we described an end-to-end workflow for the analysis of AMR genes a using advances in next-generation sequencing. No statistically significant differences were observed among any other genes except the tet-(40) gene between two sequencing platforms, which may be due to the short amplicon length.

Conclusions: Irrespective of sequencing chemistry and platform used, comparative analysis of AMR genes and candidate host organism suggest that the Illumina MiSeq and Ion Torrent platforms performed almost equally. Regardless of sequencing platform, the results were closely comparable with minor differences.

1. Introduction

Antimicrobial resistance (AMR) is a growing challenge to the efficient control of diseases caused by bacteria, parasites, viruses, and fungi, prompting the World Health Organization (WHO) to rank it in the top ten public health hazards worldwide. The consequences of AMR include reduced treatment efficacy and increased pathogen persistence, enhancing the likelihood of disease and transmission to others. Multiple-drug-resistant bacteria may already be responsible for 700 000 or more human deaths each year [1]. The UK's Review on Antimicrobial Resistance states that "Advances in genetics, genomics and computer science will likely change the way that infections and new types of resistance are diagnosed, detected and reported worldwide, so that we can fight back faster when bacteria evolve to resist drugs" [2]. One key advance is the use of next-generation sequencing (NGS) to detect and analyze the presence of genes and organisms responsible for the transmission of AMR [3].

As sequencing platforms and data analysis pipelines evolve, it is important to regularly review their performance for specific applications. An increasingly wide range of NGS platforms are now available, including pyrosequencing, semiconductor-based sequencing, sequencing by synthesis, sequencing by ligation, and third generation sequencers [4], each based upon a distinct sequencing chemistry. For example, the Ion Torrent technique detects hydro-

* Corresponding Author:
E-mail address: madhvimicrobio@gmail.com (M. Joshi).

https://doi.org/10.1016/j.jgar.2022.08.017

gen ions released during the integration of nucleotides into the expanding DNA template [5,6], while Illumina works on sequencing by synthesis chemistry [7]. Platforms such as Ion Torrent and Illumina also have their own specifications for library preparation method, read length, data quality, and total data output per run. Consequentially, it can be challenging for researchers to select the optimal sequencing platform and specifications. The choice of data analysis pipeline to process output data adds additional variables, influenced by the nature of the data and the purpose of the study. Several tools are available to detect AMR genes in NGS data, including multiple pipelines, and different thresholds and databases, hindering comparison between studies. The core objectives of the work presented in this paper were to understand how the use of Illumina or Ion Torrent sequencing platforms impact data generation, analysis, and final outcome for AMR gene detection in biological samples with relevance to public health. Further we also optimize a data analysis pipeline and use different databases for the analysis of AM genes.

2. Materials and methods

2.1. Ethical approval

The work described here was carried out using welfare standards consistent with those established under the Animals (Scientific Procedures) Act 1986, an Act of Parliament of the United Kingdom. All protocols were approved by the Animal Ethics Committee of the Anand Agricultural University (AAU, Gujarat, India) and the Animal Welfare and Ethical Review Body (AWERB) of the Royal Veterinary College, London, UK.

2.2. Sample Collection and Processing

Twelve apparently healthy broiler chickens (Cobb 400) were collected from the Central Poultry Research Station of Anand Agricultural University, Anand, Gujarat, India. All 12 chickens were euthanized by cervical dislocation at 37 days of age. The chickens were reared in a deep litter system using rice husk as substrate, in common with local practices. All chickens were fed a standard maize and soybean-based commercial diet, which included bacitracin methylene disalicylate and maduramycin (10%) for routine prophylaxis as described in our previous study [8]. Similarly, samples were collected in RNAprotect Bacteria Reagent (QIAGEN, Germany) as described in our previous study [8] and transported to the laboratory at 4 °C. Upon receipt, total genomic DNA was extracted from each sample immediately using a QIAamp® DNA Stool Mini Kit (QIAGEN, Germany) as described previously [8]. Extracted DNA was stored at -20 °C prior to further processing.

2.3. AMR Gene Sequencing

While comparing two different sequencing platforms, there should be no difference in the workflow including library preparation. Therefore, we used an Ion AmpliSeq™ AMR Research Panel (Thermo Fisher Scientific, MA, USA) for library preparation. This AMR panel consisted of two primer pools targeting 408 and 407 amplicons in each pool. The library preparation flow was also standardized for both the platforms with the exception that Ion-specific adapters and barcodes were ligated for the Ion Torrent platform, while Illumina-specific adapters and indices were used for the Illumina library. For library preparation, we used 10 ng total DNA per primer pool.

2.3.1. Ion Torrent Platform

Amplicon libraries were prepared using an Ion AmpliSeq™ Library Kit Plus (Cat. No. A35907; Thermo Fisher Scientific, MA, USA). Library quality was assessed using a 2100 Bioanalyzer with a DNA high sensitivity assay kit (Agilent CA, USA). Libraries were quantified using the Ion Library TaqMan™ Quantitation kit (Cat. No. 4468802; Thermo Fisher Scientific, MA, USA). Sequencing was performed on an Ion S5 Plus system using 530 chip and 400bp chemistry.

2.3.2. Illumina Platform

Amplicon libraries for the Illumina platform were prepared using AmpliSeq Library PLUS for Illumina kit (Cat. No. 20019102) and checked for quality as described above. Sequencing was carried out using an Illumina MiSeq system with a MiSeq reagent kit v2 500 cycles (2 × 250) paired-end chemistry.

2.4. Data Analysis

Data obtained from Ion torrent and Illumina MiSeq platforms was analyzed using the same bioinformatics pipeline. The initial difference in the paired-end read from Illumina and single end reads from Ion torrent was nullified by merging the paired-end reads of Illumina using PANDAseq v 2.8.1 [9]. Here also, different overlapping parameters i.e. 5 bp, 10 bp, 15 bp overlapping and default, were first assessed for the best results.

The quality of raw data was assessed using FastQC v. 0.11.5 [10]. An average Phred quality score threshold of the reads to retain a read was set ≥ 30 for Illumina and ≥ 20 for Ion Torrent data because of the inherent differences in the base calling accuracy due to differences in the sequencing chemistry of these two platforms [11,12]. Read trimming based on length was not performed as the smallest amplicon targeted in the panel was 72 bp. CARD database version 3.0.7 [13] was used for analysis. Local Basic Local Alignment Search Tools (BLASTn and BLASTx) were performed with the following parameters: no. of alignments retrieved 1, minimum percent identity 95%, and E-value $10e^{-5}$. The downstream statistical analysis was done using Excel and STAMP v2.1.3 [14].*2.5 Comparison of different databases*

Four different databases namely, Comprehensive Antibiotic Resistance Database (CARD) [13], QIAGEN microbial Insight–Antimicrobial Resistance (QMI-AR) [15], Antimicrobial Resistance (AR), and CARD-CLC [16] were used for comparison. These databases were compared using stringent parameters including number of alignments per read as 1, minimum alignment length as 95%, E-value as 10 e-5 and percent identity for BLAST 95%. The downstream analysis was performed using STAMP and Venny 2.1.0 [17].

2.6. Microbiome analysis

The online web-based tool Microbiome Analyst [18,19] was used to perform linear discriminant analysis effect size (LEfSe), principal coordinates analysis with permutational multivariate analysis of variance (PERMANOVA) statistics, and random forest analysis to support statistical comparison. In LEfSe, Log linear discriminant analysis (LDA) cutoff was set as 3.0 with $P \leq 0.05$.

3. Results

3.1. Sequencing results

In total ~15 M reads were obtained using the Ion Torrent S5 Plus platform, correspond to approximately 1 M reads per sample with an average length of 200 bp. In parallel, 4.18 M reads were produced for the same samples using Illumina MiSeq, correspond to 0.2 M reads per sample with an average length of 185 bp.

3.2. Optimization of overlapping parameter for Illumina

Based on our previous experience with PANDAseq in merging 16S Illumina amplicon data and the toll's citations (> 1600), we selected this toll for this study. Merging of Illumina forward and reverse reads required attention [20]. Initially, we merged the reads using PANDAseq's default parameters; later, the overlapping length was optimized. Forward-reverse read overlaps of 5, 10, and 15 bp were analyzed in addition to the default parameters. The 10 base pairs overlap was found to be optimal due to its appropriate representation of merged reads (Supplementary Fig. S1). These results showed that overlapping parameters for merging forward-reverse amplicon reads may incur important differences in apparent gene abundance, as an appropriate overlapping parameter leads to false positive and negative results in Illumina sequencing platforms while analyzing AMR data. Overlap read length means the minimum number of exact nucleotide matches required between forward and reverse reads in order to create a singale consensus sequence. The less the overlap, the higher will be the chances of false positive results, whereas higher overlap length may result in false negative results (i.e. due to sequencing defects, there may be a mismatch at some position, which if it occurs in the overlap length, will result in false negative results).

3.3. Optimization of BLAST parameters

The BLAST algorithm is used widely, but output is influenced by the parameters applied. Therefore, in this study, various BLAST parameters were optimized along with the overlapping length used in PANDAseq. Three conditions were set: default BLAST and default PANDAseq overlap length: 10 bp overlap and default BLAST and 10 bp overlap in PANDAseq; and BLAST query hsp percentage 90 (Supplementary Fig. S2). The default overlap with default BLAST could not be used for analysis due to nonspecific read merges. Specifically, the PANDAseq default merge length is 1 bp, indicating that any two reads possessing a common base at the 5´end will be merged. The 10 bp overlap and BLAST qcov hsp percentage 90 was also not efficient, as it hampered estimation of occurrence for genes such as *ErmB*. The 10 bp overlap with default BLAST was found to be most accurate as it avoided all these issues and was applied for all subsequent analyses.

3.4. Comparison of Ion Torrent and Illumina MiSeq for AMR gene detection

More AMR genes were detected using the Ion Torrent Platform compared to Illumina MiSeq (average number of genes detected 369 ± 58 compared to 206 ± 38, respectively, from all 12 samples). In total, the Ion Torrent platform detected the presence of 31.9% more AMR genes compared to Illumina MiSeq, although the percentage abundance of these genes was very low (i.e. less than 0.004%). Additionally, 6% of genes detected using Illumina MiSeq were missing from the Ion Torrent results, but again the abundance of these genes was very low (i.e. less than 0.004%). Overall, 62.1% of genes were common across both the platforms, but, when genes with abundance ≥1% were considered, the results from both sequencing platforms were similar (Table 1, Fig. 1). The *APH (3')-IIIa* gene was found to be most abundant in both the platforms followed by *tetW* and *tetQ*. The occurrence of only nine genes was found to be significantly different between both sequencing platforms (Supplementary Fig. S3). Out of these nine genes, *tet-(40)* was found to be most variable with almost 4% difference between two platforms. Sample-specific comparison highlighted similar platform-associated variation for the occurrence of *tetO* and aminoglycoside phosphotransferase genes (Supplementary Fig. S4, Supplementary Fig. S5). Direct sample-specific comparison revealed comparable gene detection profiles using Illumina MiSeq and Ion Torrent S5 Plus for genes with greater than 1% read abundance (Fig. 2).

Table 1
Comparative analysis of the presence or absence of AMR genes represented by ≥ 1% sequence abundance within Illumina MiSeq or Ion Torrent amplicon sequencing datasets.

	Number of AMR genes detected			
Sample	Total	Illumina AND Ion Torrent	Illumina	Ion Torrent
S1	19	18	1	0
S2	16	15	1	0
S3	13	13	0	0
S4	14	14	0	0
S5	13	13	0	0
S6	12	12	0	0
S7	13	13	0	0
S8	16	15	1	0
S9	18	14	1	3
S10	16	11	0	5
S11	16	12	3	1
S12	16	14	2	0

Table 2
Comparative analysis of the presence or absence of organisms predicted to host AMR genes detected within Illumina MiSeq or Ion Torrent amplicon sequencing datasets. Organisms representing AMR genes with ≥ 1% abundance are shown.

	Number organisms identified by CARD			
Sample	Total	Illumina AND Ion Torrent	Illumina	Ion Torrent
S1	15	14	0	1
S2	13	13	0	0
S3	12	12	0	0
S4	13	13	0	0
S5	12	12	0	0
S6	13	12	1	0
S7	13	13	0	0
S8	14	14	0	0
S9	15	14	0	1
S10	14	10	0	4
S11	14	13	1	0
S12	14	14	0	0

3.5. Tet-(40) and Lnu C comparison

The abundance of *tet-(40)* was found to be higher in Illumina MiSeq data (6.21 ± 1.26%) when compared with Ion Torrent (2.5 ± 1.0%). Annotation using the CARD database indicated *tet-(40)* carriage by a group of uncultured bacteria. Thus, a comparable trend was also observed when samples were compared on taxonomy (Illumina 6.2 ± 1.6%; Ion Torrent 2.5 ± 1.0%). In contrast, *lnuC* was abundant in the Ion Torrent dataset (9.79 ± 5.15%) compared with Illumina (7.9±4.1%) (Table 3). The *lnuC* gene was predicted to be carried by *Streptococcus agalactiae* and hence, the same trend in the percentage of *S. agalactiae* was observed (Fig. 1, Fig. 3).

3.6. Microbial diversity comparison

Prediction of bacterial community associated with AMR genes was found to be comparable in both the platforms (Table 2, Fig. 3). *Campylobacter coli* CVM N29710 was the most abundant organism identified, followed by *Bacteroides fragillis*. Only *Staphylococcus epidermidis* was found to be significantly different between the platforms (q-value (corrected) = 0.001) (Abundance < 0.0014) (Supplementary Fig. S6). Comparison of bacterial representation in individual samples was also undertaken, illustrating the comparable taxonomic classification between the two sequencing platforms (Fig. 4, Supplementary Fig. S7, Supplementary Fig. S8).

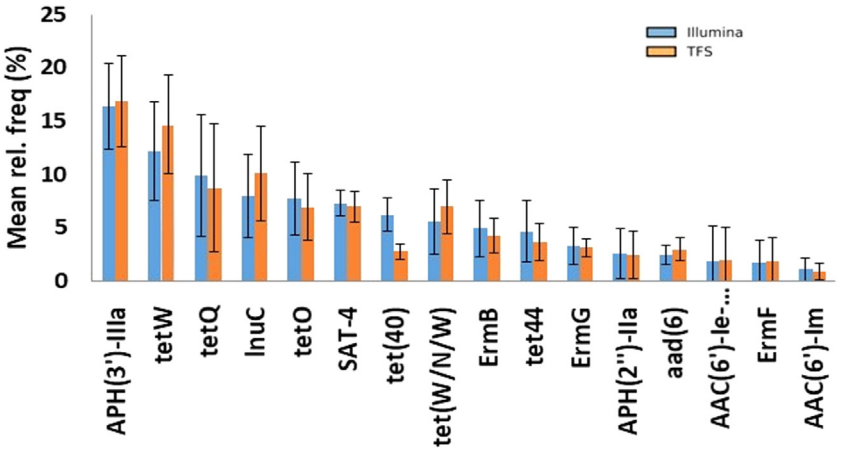

Fig. 1. The relative sequencing read abundance of genes with ≥ 1% abundance within the Illumina and Ion Torrent platform datasets.

Fig. 2. Heatmap depicting the abundance of top 25 AMR genes in all 12 samples from Illumina and Ion Torrent, plotted using statistical analysis of metagenomic profiles (STAMP) Abundance of AMR genes in all samples is graphically represented where individual values in a sample are marked by color gradient. The color coding representing the abundances is indicated by the color scale.

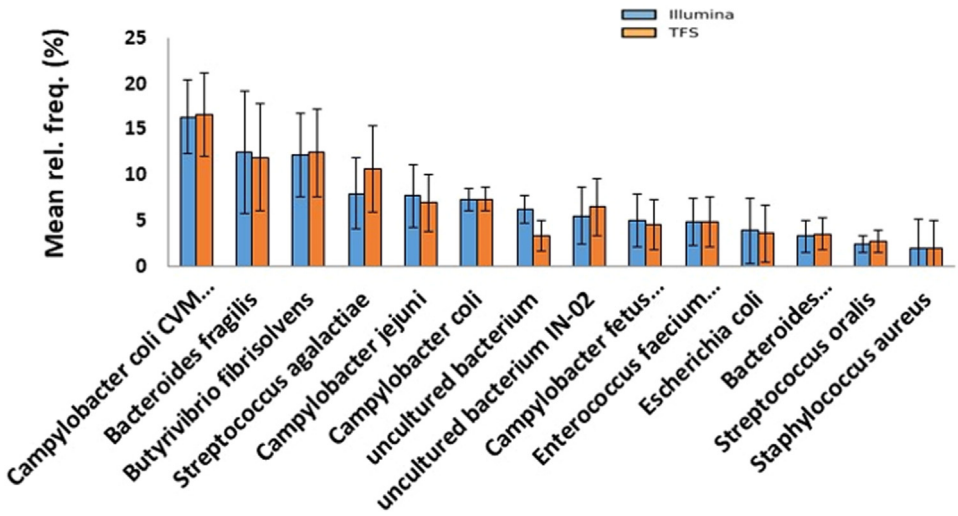

Fig. 3. The relative sequence abundance with ≥ 1% organisms corresponding to the AMR (on the basis of CARD database) between the Illumina and Ion Torrent platforms.

Table 3
Variation in the relative abundance of *tet(40)* and *lnuC* gene amplicons detected in chicken caecal bacterial populations using Illumina MiSeq or Ion Torrent amplicon sequencing. Likely host organism (as predicted by CARD) and amplicon length is shown.

GENE	ILLUMINA	ION TORRENT	ORGANISM	AMPLICON LENGTH
tet(40)	High (6.2%±1.3 %)	Low (2.5%±1.0%)	Uncultured bacteria	80
lnuC	Low (7.9%±4.1%)	High (9.8%±5.1%)	*S. agalactiae*	224

3.7. Database comparison

Several databases are available for the analysis of AMR genes. Comparison of the CARD, QMI-DB, AR and CARD-CLC databases with stringent parameters produced varied results with limited correlation or similarity (Fig. 5). In the absence of clear complementarity, the CARD database was chosen for downstream analysis because it is easily available and hosts the largest number of genes and organisms among the four databases. CARD is a curated database prepared after varyfying the gene with the laboratory experiment and supplemented with the Antibiotic Resistance Ontology (ARO) for each gene. The microbiological analysis module in CLC genomic workbench (version 21.1) was utilized to compare results. The investigation also made use of the CARD database in CLC genomic workbench.

3.8. Statistical comparison of AMR gene occurrence detected by Illumina MiSeq and Ion Torrent sequencing

The random forest method generates decision trees from data samples, generating multiple predictions before identifying the best solution. Random forest is an ensemble method that is superior to a single decision tree because it averages results to reduce over-fitting [21]. Here, random forest analysis was performed in order to identify any outliers in each dataset. Comparison of Illumina MiSeq and Ion Torrent datasets revealed the absence of outliers, supporting the comparison of both platforms (Supplementary Fig. S9). Similarly, principal coordinates analysis was used to confirm that, irrespective of the sequencing platform, all the samples were tightly clustered (Fig. 6). For AMR genes and respective organism comparisons, there were no significant differences (PERMANOVA: AMR genes; F value 1.3421, r^2 value 0.057498, $P < 0.219$; and organisms; F value 0.82178, r^2 value 0.036009, $P < 0.514$).

3.9. LEfSe analysis

LEfSe was performed for both genes and organisms with minimum LDA score 3.0 and $P \leq 0.05$. Only 4 out of 300 organisms were found to be significantly different between both sequencing platforms. *Enterococcus faecalis*, Plasmid_pGT633, and *Bacteroides coprosuis* were more abundant in the Ion Torrent dataset, while uncultured bacteria were more common using the Illumina platform. However, the abundance of all four organisms was low, i.e. less than 0.07% and 0.02% in the Ion Torrent and Illumina datasets, respectively, and both below the 1% threshold set earlier (Supplementary Fig. S10). LEfSe analysis of the AMR genes identified five genes that were significantly different between platforms (Supplementary Fig. S11). The genes *tet32, ErmT, tetS,* and *Erm35* were found to be more abundant in Ion Torrent dataset, while *tet-(40)* was more common in the Illumina data. Again, the percent abundance of these gene-specific reads was less than 0.04% in the Ion Torrent sequencing. Only detection of the gene *tet-(40)* was found to be significantly different, with 6.2 ± 1.3 % abundance, presenting almost 2.5 fold higher abundance in the Illumina MiSeq data.

4. Discussion

To support national and global priority setting, public health initiatives, and treatment decisions, a credible base of knowledge that appropriately captures and characterizes the worldwide burden and transmission of AMR is required. The study was planned to answer the very basic question associated with the use of NGS sequencing platforms for AMR analysis. Therefore, in this study we compared two sequencing platforms, Ion Torrent and Illumina MiSeq for the analysis of AMR genes and set a bioinformatics data analysis pipeline after consideration of all the differences between two platforms. All experimental variables were fixed with the exception of sequencing platform and a minor difference in Illumina data analysis, i.e. merging forward and reverse reads. Li-

Fig. 4. Heatmap depicting the abundance of top 25 organisms as per CARD database in all 12 samples from Illumina and Ion Torrent, plotted using Statistical analysis of Metagenomic Profiles (STAMP). Abundance of organisms in all samples is graphically represented where individual values in a sample are marked by color gradient. The color coding representing the abundances are indicated by the color scale.

brary preparation kit, data analysis pipeline, and database stringency were all kept the same to maintain uniformity. The difference in the sequencing chemistry between platforms may result in internal biases which have already been reported by various researchers [22–24].

This study was performed with 12 chicken cecum samples to estimate the abundance of AMR genes and corresponding organisms. The bioinformatics pipeline generated for data analysis kept constant for both platforms, although the initial parameters, such as quality score threshold and merging of forward and revers reads, varied a bit between platforms. Due to higher confidence at quality score greater than 30 in Illumina and greater than 20 at Ion Torrent, we set different initial quality cutoffs for the data. In addition to this, Ion Torrent generates single-end sequencing reads while paired-end reads are generated by Illumina platforms. In order to merge the forward and reverse reads of Illumina, an extra step of read merges was performed. These two changes bring data from both platforms onto same page. Later, the local BLAST parameters and statistical analysis parameters were kept stringent and constant for both data sets.

Upon the completion of analysis, we found that for the analysis of AMR genes, both sequencing platforms almost performed equally, and data were highly comparable. Similar results were obtained by Lahens et al. [25], upon comparative analysis of differential expression of genes among Ion torrent and Illumine [25]. A similar study was performed by Allali et al. [26], in which they analyzed the chicken gut microbiome over different sequencing platform for the detection of *Salmonella* infection [26]. They came to the same conclusion as presented in this paper in that despite differences in the sequencing platform, chemistry used, and bioinformatics pipeline, the same conclusion can be drawn. The difference in other insignificant hits may arise from sequencing errors and poor quality. The *tet*-(40) gene was found to be significantly different between both platforms. Upon detailed analysis of *tet*-(40) abundance, it was found that the amplicon length of *tet*-(40) gene is only 80 bp. It is among the shortest amplicon present in the AMR panel. This short amplicon length may result in the phenomenon of competitive binding on Illumina flow cells during cluster generation. Competitive binding means a shorter amplicon tends to bind to flow cells more as compared to a larger one. The bacterium which corresponds to this *tet-(40)* is uncultured bacteria. Hence, the same trend was observed in the abundance of uncultured bacteria. Inverse to this, during emulsion polymerase chain reaction in the Ion Torrent sequencing, shorter fragments tended to form polyclonal antibodies, and therefore, the reads tended to be discarded. Therefore, we expected that this could be one of the possible reasons for the *tet-(40)* gene's lesser abundance in the Ion Torrent dataset and higher abundance in the Illumina dataset. Similarly, the lincosamide resistance gene is one of the largest amplicons (224 bp) in the AMR panel used in this study. The phenomenon opposite that of *tet*-(40) may work here, i.e. lower abundance of *lnuC* in Illumina data as compared with Ion

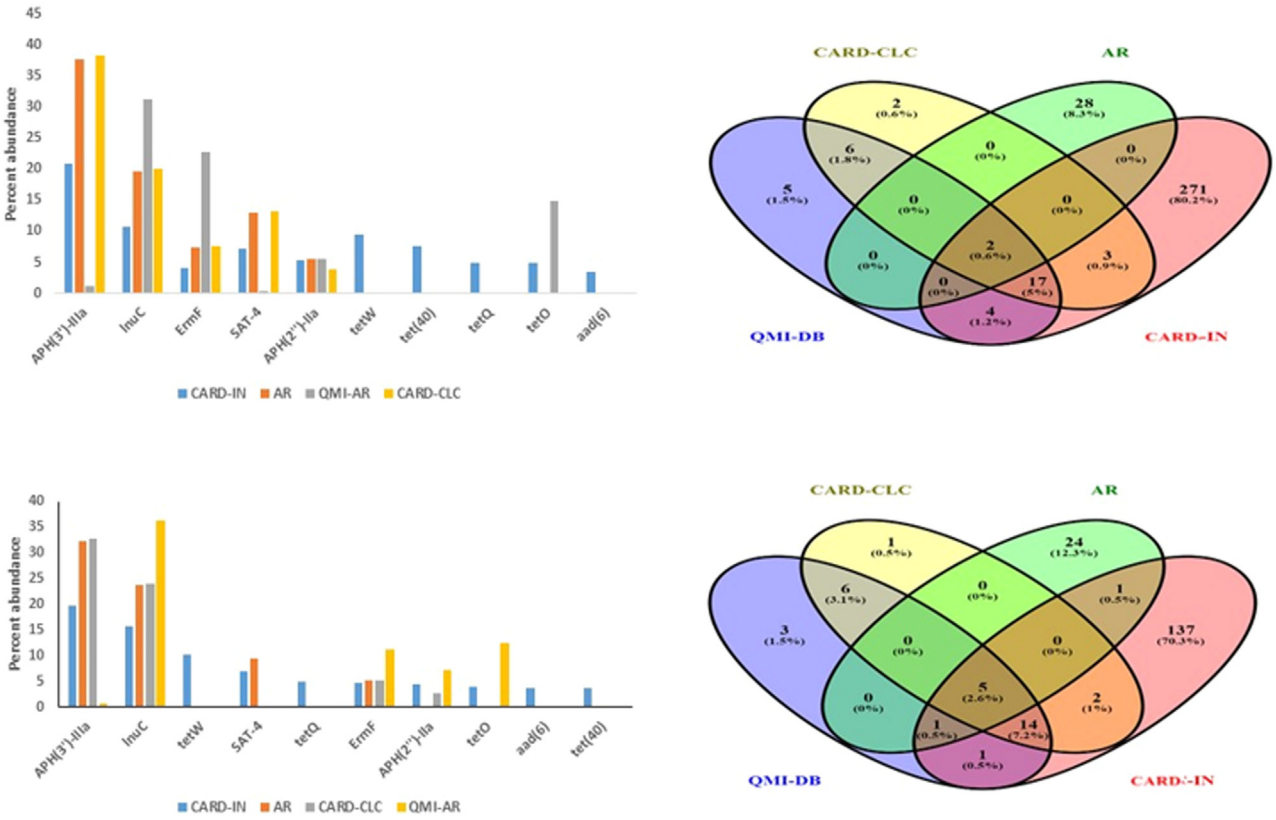

Fig. 5. Database comparison for same sample in Illumina and Ion Torrent (CARD-CLC- CARD database present in CLC genomic workbench microbial genomic module, AR – Antibiotic resistance database, QMI-DB- QIAGEN microbial Insight – AR, CARD-IN – CARD database downloaded from CARD site and run locally).

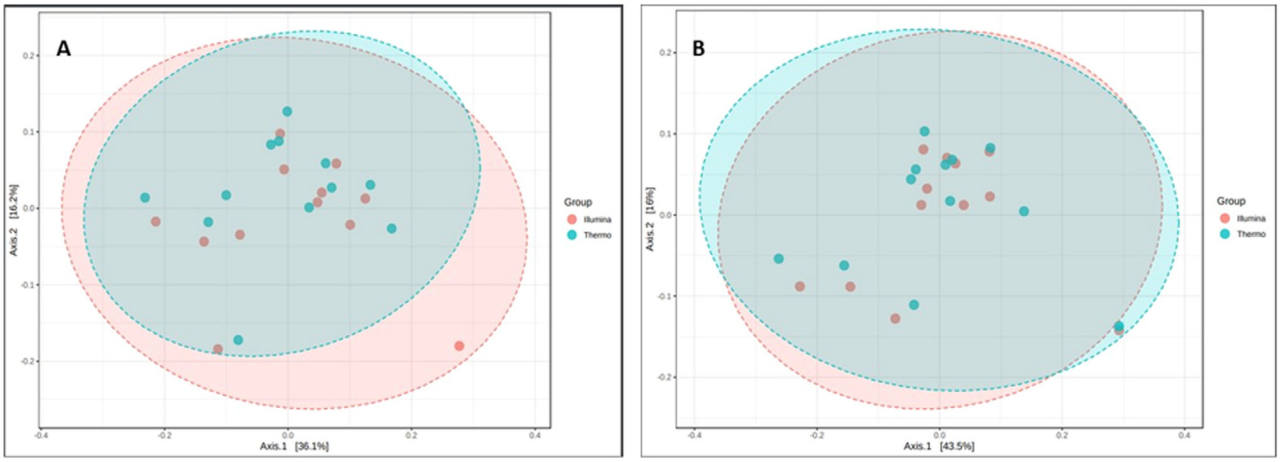

Fig. 6. PERMANOVA analysis of AMR gene and organism. (A) AMR gene PERMANOVA with $F = 1.3421$, $r^2 = 0.057498$ and $^*P < 0.219$ (B) Organism PERMANOVA analysis with $F = 0.82178$, $r^2 = 0.036009$ and $^*P < 0.514$.

torrent. *LunC* gene is mostly contributed from the *S. agalactiae,* and therefore, a similar trend also is observed in this case. The only statistically significant difference was found for one organism, i.e. *Staphylococcus epidermidis.* The variation in the abundance of *S. epidermidis* is almost negligible, as its abundance is very less.

Two different bioinformatics platforms were used to identify any database correlation. One of these platforms was the CARD local database and another was the CLC workbench with QIAGEN microbial insight module providing a different database for the AMR search (QMI-AR, AR, CARD). The CARD local database was preferred due to its having higher number of genes as compare to the other. Other researchers have also compared the CARD database for AMR gene analysis with the other available databases and validated the effectiveness and accuracy of the CARD database [27–29]. The main disadvantage of the CLC workbench is that it is not freely available. Both the CLC workbench license and the microbiological insight module have separate costs.

In summary, the present study has effectively demonstrated that both sequencing platforms, i.e. Illumina MiSeq and Ion Torrent, produce comparable results for the analysis of AMR genes from poultry samples. For analyzing the AMR genes in the sample, one shound selcet the either of the sequencer i.e. Illumina or Ion Torrent based on the availability of the fund and instrument in the Institute. The only limitation of the present study is that we did

not perform the same exercise on a mock community, as a mock community for the analysis of AMR genes is not available.

4. Conclusion

Irrespective of sequencing chemistry and platform used, comparative analysis among AMR genes and candidate host organisms suggest that the Illumina MiSeq and Ion Torrent platforms performed almost equally. According to the findings, the authors suggest that using any platform (Ion Torrent or Illumina) or sequencing chemistry has little effect on the outcome of the AMR data analysis. Comparative analysis of the organisms identified in each sample rarely varied significantly. A statistical significance difference among the *tet*-(40) gene was observed, which may arise with short length amplicons. Furthermore, in order to correctly assess AMR in biological samples, standard methods and a pipeline for sample analysis must be established. Database selection and parameters for analysis can change the outcome considerably.

Table 3

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

Department of Science and Technology (DST), Government of Gujarat, Gandhinagar, Gujarat, India and One Health Poulty Hub

Ethical approval

All protocols were approved by the Animal Ethics Committee of the Anand Agricultural University (AAU, Gujarat, India) and the Animal Welfare and Ethical Review Body (AWERB) of the Royal Veterinary College, London, UK.

Acknowledgements

The authors would like to acknowledge global challenge research fund (GCRF), UK research and innovation (UKRI) and Dr. Prakash Koringa and his team from Anand agricultural University (AAU), Anand for the support.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jgar.2022.08.017.

References

[1] Farha MA, Brown ED. Drug repurposing for antimicrobial discovery. Nat Microbiol 2019;4:565–77. doi:10.1038/s41564-019-0357-1.
[2] O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations; 2014. London, UK.
[3] Angers A, Petrillo M, Patak A, Querci M, Van den Eede G. The role and implementation of next-generation sequencing technologies in the coordinated action plan against antimicrobial resistance. EUR 2017;28619.
[4] Tripathi V, Kumar P, Tripathi P, Kishore A, Kamle M. Microbial genomics in sustainable agroecosystems (Vol. 2). Singapore: Springer Nature Singapore; 2019.
[5] Gupta AK, Gupta UD, Verma AS, Singh A. Next generation sequencing and its applications. In: Animal biotechnology: models in discovery and translation. Amsterdam: Elsevier; 2014. p. 395–421.
[6] Rusk N. Torrents of sequence. Nat Methods 2011;8:44. doi:10.1038/nmeth.f.330.
[7] Clark DP, Pazdernik N, McGehee M. Molecular Biology. London: Academic Press; 2019.
[8] Pandit RJ, Hinsu AT, Patel NV, Koringa PG, Jakhesara SJ, Thakkar JR, et al. Microbial diversity and community composition of caecal microbiota in commercial and indigenous Indian chickens determined using 16s rDNA amplicon sequencing. Microbiome 2018;6:115. doi:10.1186/s40168-018-0501-9.
[9] Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 2012;13:1–7. doi:10.1186/1471-2105-13-31.
[10] Andrews S. FastQC Version 0.11. 5. A Quality Control Tool for High Throughput Sequence Data. Babraham Bioinformatics; 2016.
[11] Osborne CA. Terminal restriction fragment length polymorphism (T-RFLP) profiling of bacterial 16s rRNA Genes. Environmental microbiology. methods in molecular biology. Paulsen I, Homes A, editors, Totowa, NJ: Humana Press; 2014. vol. 1096.
[12] Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 2012;13:341. doi:10.1186/1471-2164-13-341.
[13] Wein T, Wang Y, Hülter NF, Hammerschmidt K, Dagan T. Antibiotics interfere with the evolution of plasmid stability. Curr Biol 2020;30:3841–7 e4. doi:10.1016/j.cub.2020.07.019.
[14] Liu X, Cheng YW, Shao L, Sun SH, Wu J, Song QH, et al. Gut microbiota dysbiosis in Chinese children with type 1 diabetes mellitus: an observational study. World J Gastroenterol 2021;27:2394–414. doi:10.3748/wjg.v27.i19.2394.
[15] Roachford OSE, Alleyne AT, Kuelbs C, Torralba MG, Nelson KE. The cervicovaginal microbiome and its resistome in a random selection of Afro-Caribbean women. Hum Microbiome J 2021;20:100079. doi:10.1016/j.humic.2021.100079.
[16] Dubin K, Callahan MK, Ren B, Khanin R, Viale A, Ling L, et al. Intestinal microbiome analyses identify melanoma patients at risk for checkpoint-blockade-induced colitis. Nat Commun 2016;7:1–8. doi:10.1038/ncomms10391.
[17] Song D, Tian J, Hu Y, Wei Y, Lu H, Wang Y, et al. Identification of biomarkers associated with diagnosis and prognosis of gastroesophageal junction adenocarcinoma-a study based on integrated bioinformatics analysis in GEO and TCGA database. Medicine (Baltimore) 2020;99:e23605. doi:10.1097/MD.0000000000023605.
[18] Chong J, Liu P, Zhou G, Xia J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat Protoc 2020;15:799–821. doi:10.1038/s41596-019-0264-1.
[19] Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res 2017;45:W180–8. doi:10.1093/nar/gkx295.
[20] Liu T, Chen CY, Chen-Deng A, Chen YL, Wang JY, Hou YI, et al. Joining Illumina paired-end reads for classifying phylogenetic marker sequences. BMC Bioinformatics 2020;21:1–13. doi:10.1186/s12859-020-3445-6.
[21] Pedregosa Fabian, et al. Scikit-learn: Machine learning in Python. the Journal of machine Learning research 2011;12:2825–30.
[22] Dechesne A, Musovic S, Palomo A, Diwan V, Smets BF. Underestimation of ammonia-oxidizing bacteria abundance by amplification bias in amoA-targeted qPCR. Microb Biotechnol 2016;9:519–24. doi:10.1111/1751-7915.12366.
[23] Jones MB, Highlander SK, Anderson EL, Li W, Dayrit M, Klitgord N, et al. Library preparation methodology can influence genomic and functional predictions in human microbiome research. Proc Natl Acad Sci USA 2015;112:14024–9. doi:10.1073/pnas.1519288112.
[24] Yu G, Fadrosh D, Goedert JJ, Ravel J, Goldstein AM. Nested PCR biases in interpreting microbial community structure in 16S rRNA gene sequence datasets. PLoS ONE 2015;10:1–12. doi:10.1371/journal.pone.0132253.
[25] Lahens NF, Ricciotti E, Smirnova O, Toorens E, Kim EJ, Baruzzo G, et al. A comparison of Illumina and Ion Torrent sequencing platforms in the context of differential gene expression. BMC Genomics 2017;18:602. doi:10.1186/s12864-017-4011-0.
[26] Allali I, Arnold JW, Roach J, Cadenas MB, Butz N, Hassan HM, Koci M, Ballou A, Mendoza M, Ali R, Azcarate-Peril MA, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. BMC Microbiol 2017;17:194. doi:10.1186/s12866-017-1101-8.
[27] Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 2020;48:D517–25. doi:10.1093/nar/gkz935.
[28] Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 2017;45:D566–73. doi:10.1093/nar/gkw1004.
[29] McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 2013;57:3348–57. doi:10.1128/AAC.00419-13.