WGS Analysis of *Salmonella enterica* from VT Backyard Chickens Using MinION: Advantages and Pitfalls

Calleigh S. Herren[1] and Andrea J. Etter[2]

[1]Department of Microbiology and Molecular Genetics, University of Vermont, Burlington VT
[2]Department of Nutrition and Food Sciences, University of Vermont, Burlington VT

Introduction

In 2016, there were over 46,000 cases of culture-confirmed *Salmonella* infections reported to the Laboratory-based Enteric Disease Surveillance system (Basler et al., 2016), and this is a small fraction of the 1.3 million estimated cases per year (Scallan et al., 2011). *Salmonella* is the foremost foodborne pathogen in terms of cost of illness and mortality – $3.3 billion and 378 deaths (Batz et al., 2011). Backyard poultry ownership has been increasing across Vermont and the United States, which also increases contact with zoonotic pathogens such as *Salmonella* (Vermont Department of Health, 2018, McDonagh et al., 2019). In 2020, the number of reported illnesses linked to backyard flock outbreaks was higher than any year previously investigated (CDC, 2020). Next-generation sequencing techniques, such as the Oxford Nanopore MinION platform, allow for investigation of smaller and more dispersed foodborne outbreaks such as salmonellosis in backyard poultry flocks. We chose the ONT MinION platform for our sequencing due to apparent lower initial cost, accessibility, and the *de novo* assembly possibilities stemming from its long reads.

Purpose: The purpose of this study is to determine the serovars and genomic traits, such as antimicrobial resistance, that are found in *Salmonella enterica* isolates from backyard poultry flocks across Vermont.

Materials & Methods

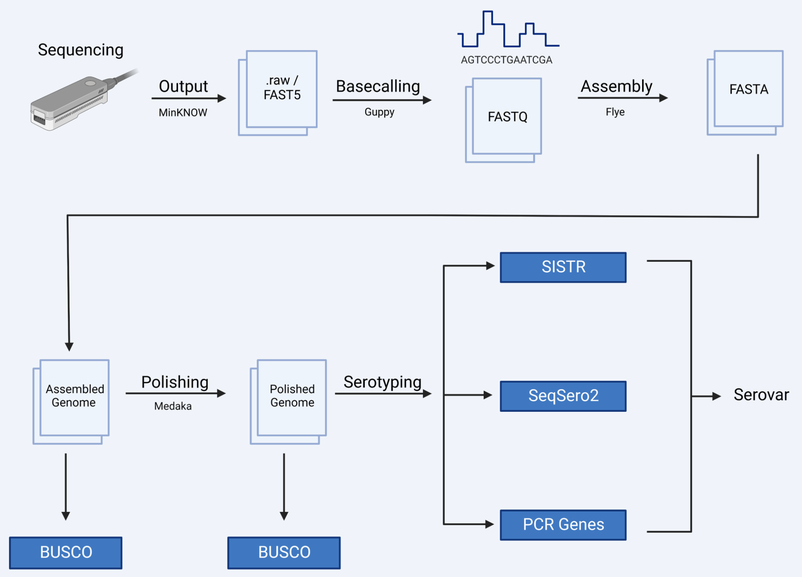

Figure 1: Post-sequencing workflow

Created with BioRender.com

Chick shipping pads were tested for *S. enterica* with the standard USDA *Salmonella* enrichment protocol and confirmed with PCR for *hilA*. Pure cultures were streaked onto TSA plates and inoculated in TSB; 1 mL of overnight culture was used an input for standard phenol:chloroform DNA extraction. DNA quality, concentration, and purity were assessed with the QuBIT 4 Fluorometer and Nanodrop spectrophotometer. Library preparation was performed with Oxford Nanopore's Rapid PCR Barcoding Kit (RBP-004) and three to five samples were barcoded and pooled prior to sequencing. Sequencing was done on the MinION using R.9.4 flow cells with a 6-8 hour run time.

Genomes were assembled and analyzed on the Vermont Advanced Computing Core (VACC) supercomputers using Flye (v. 2.6), and polished with Medaka (v. 1.6.0). Genome quality was assessed with BUSCO (v. 5.3.2) before and after polishing. Assembled and polished genomes were then used as the input for SISTR (v. 1.1.1) and SeqSero2 to determine serovar, ResFinder (v. 4.1) to determine the presence of antimicrobial resistance, and PlasmidFinder (v. 2.1) to determine the presence of plasmids. Assembled genomes were also run against gene markers specific to serovars using BLAST (Yang et al., 2021).

Results

Isolate Name	cgMLST Type	Assembly Length (bp)
FML-BYC-001	Cerro*[a], 102[b]	3,184,259
FML-BYC-002	Cerro*, 6	522,908
FML-BYC-003	Enteritidis*, 9	391,693
FML-BYC-004	Enteritidis*, 19	768,945
FML-BYC-005	Enteritidis*, 2	250,221
FML-BYC-006	Cerro, 248	4,734,944
FML-BYC-008	-[c]	-
FML-BYC-010	Senftenberg*, 0	1,743,005
FML-BYC-012	-	-
FML-BYC-013	Braenderup, 227	4,849,109
FML-BYC-014	Enteritidis*, 7	737,624
FML-BYC-015	Enteritidis*, 72	4,240,004
FML-BYC-016	Enteritidis*, 28	1,916,357
FML-BYC-017	-	-
FML-BYC-018	Infantis*, 58	2,620,469
FML-BYC-019	Mbandaka, 192	4,989,014
FML-BYC-020	Enteritidis, 237	4,902,965
FML-BYC-021	Kentucky, 226	4,931,704
FML-BYC-022	Kentucky*, 4	246,279
FML-BYC-023	Kentucky*, 97	4,622,441
FML-BYC-024	Senftenberg*, 18	685,746
FML-BYC-025	Enteritidis*, 67	3,160,671
FML-BYC-026	Cerro*, 114	4,770,731
FML-BYC-027	Enteritidis, 247	4,786,787
FML-BYC-028	Enteritidis, 231	4,795,745
FML-BYC-029	Enteritidis, 234	4,812,499

Table 1: Isolates sequenced using MinION, with corresponding cgMLST serovar predictions by SISTR (Yoshida et al., 2016) and assembly lengths after Flye (Kolmogorov et al., 2019). The average size of a *Salmonella* genome is 4.8 Mbp (Wu et al., 2021)

[a] An * indicates the assembly failed the SISTR quality control check

[b] The number of matched cgMLST loci, SISTR recommends 297 for confidence in serovar prediction

[c] A – indicates that Flye was unable to create an assembly

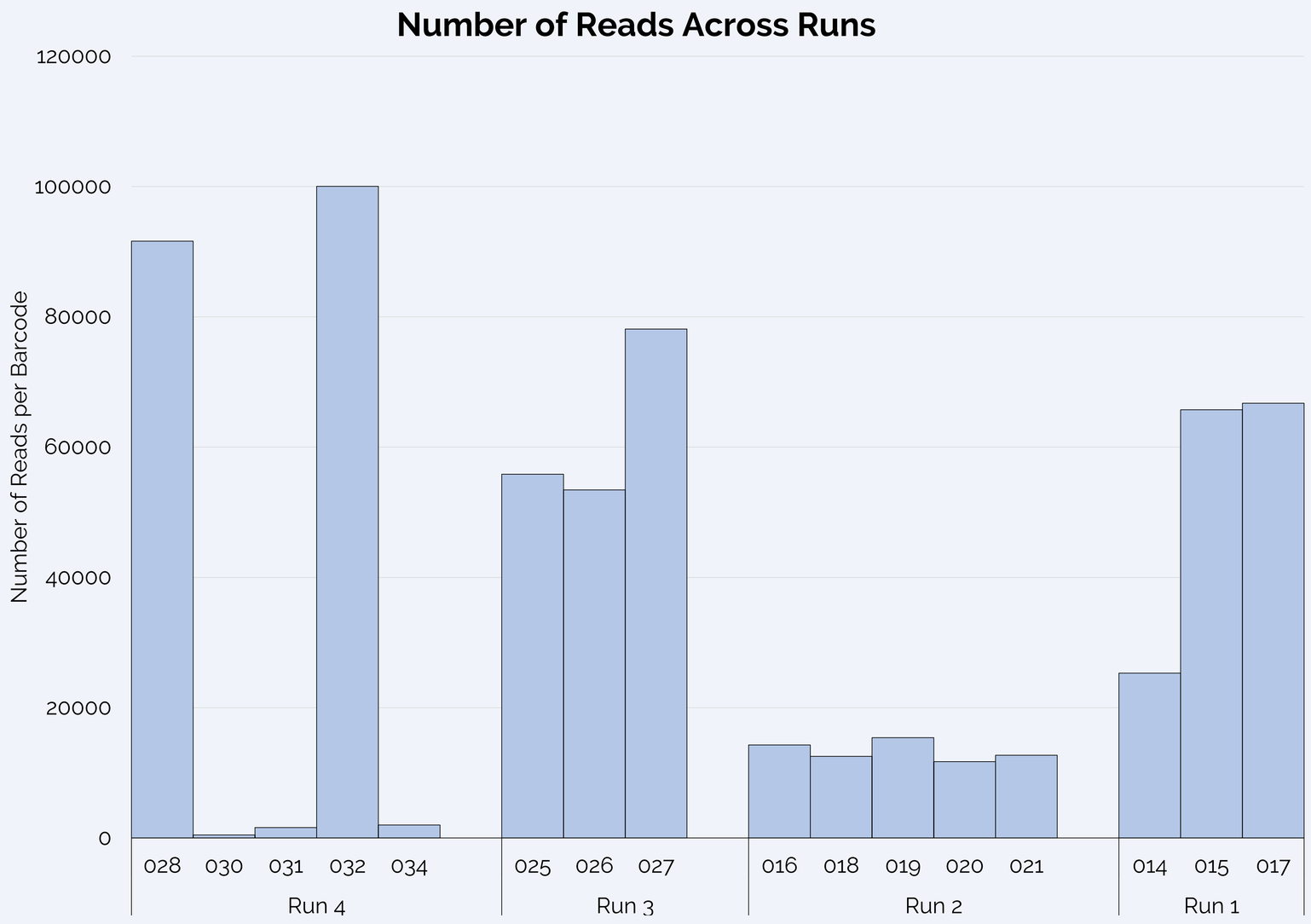

Figure 2: Four of the sequencing runs we ran, each with three to five isolates (identified by experimental identification number) and the number of reads generated for each isolate. We experienced difficulty with reads varying widely within and between runs, even when pooling ratios were even

Results & Significance

We sequenced 26 isolates from our collection of 152 *S. enterica* from chick shipping boxes. Serovars identified after MinION sequencing included Cerro (n = 3), Braenderup (n = 1), Enteritidis (n = 8), Infantis (n = 1), Mbandaka (n = 1), and Kentucky (n = 2). Plasmids found in the samples so far include IncFIB(S), IncFII(S), IncX1, and Col(pHAD28); 42.31% of isolates had plasmids.

Cerro and Kentucky are normal chicken flora and rarely cause disease in humans, however serovars Braenderup, Infantis, Enteritidis, and Mbandaka have all been implicated in outbreaks of *Salmonella* from backyard poultry (Foley et al., 2013). Results from this project will inform risk assessments and best practices guidelines for backyard chicken owners, stores selling baby chicks, and the hatcheries that supply them.

Although there are many DNA extraction kits on the market, after much trial and error we discovered only a few protocols that would allow us to reliably extract sequencing-grade DNA from our isolates. From there, the library preparation kits offered by ONT are expensive and can be finicky. While the Rapid PCR Barcoding Kit (RPB-004) is more user friendly and time effective, it picks up fewer short fragments and thus misses portions of the genome. Other kits, such as the Ligation Sequencing Kit, can capture a wider range of fragment sizes but are more expensive and significantly more time-intensive. Once libraries are prepared, determining the optimal pooling ratios and run specifications can be difficult and require many attempts. Additionally, getting even and consistent coverage across barcodes is a challenge during a sequencing run (Figure 2). Once sequencing is performed, the data is often not reliable enough to make accurate serovar predictions or elucidate other genomic information. Out of 26 isolates, 3 were unable to be assembled by Flye, and 15 additional isolates failed the quality control check in SISTR (Table 1).

It should be noted that Oxford Nanopore does offer some kits specializing in high accuracy. However, at the present this is not a viable option for this project due to the inability of the high accuracy kits to handle multiplexing, as well as being significantly more expensive per sample.

Thus, while the Oxford Nanopore MinION is an excellent tool for well-trained individuals or those who are perhaps more interested in surface level genomic analysis such as classification, those working on projects lacking experienced personnel or funds should be aware of potential drawbacks and challenges. In order to use the ONT MinION to its highest potential, teams should be willing to invest considerable time and funds into the project or have individuals with significant training and ONT experience on board.

Conclusions

The Oxford Nanopore MinION platform is an incredible tool with a low start-up cost compared to other next generation sequencing platforms. However, while the apparent benefits to the platform are appealing, obtaining reliable genomic data in a research laboratory setting is difficult and expensive, and there is a tradeoff between accuracy and feasibility/cost for sequencing bacterial collections. We are excited for the future of MinION, especially the new 10.4.1 flow cells and associated ligation kits, but at this time the reliability, cost, time investment, and accuracy of genomic information from Nanopore MinION is not suitable for many smaller, less well-resourced academic laboratories.

References

Basler, C., Nguyen, T.-A., Anderson, T. C., Hancock, T., & Behravesh, C. B. (2016). Outbreaks of Human Salmonella Infections Associated with Live Poultry, United States, 1990–2014. *Emerging Infectious Diseases*, *22*(10), 1705–1711. https://doi.org/10.3201/eid2210.150765

Batz, M., Hoffmann, S., & Morris, J. G. (2011). *Ranking the Risks: The 10 Pathogen-Food Combinations With The Greatest Burden on Public Health*. University of Florida. https://www.epi.ufl.edu/media/epiufledu/Ranking-the-Risks-REPORT.pdf

CDC. (2020, December 17). *Outbreak of Salmonella Infections Linked to Backyard Poultry*. https://www.cdc.gov/salmonella/backyardpoultry-05-20/index.html

Foley, S. L., Johnson, T. J., Ricke, S. C., Nayak, R., & Danzeisen, J. (2013). Salmonella Pathogenicity and Host Adaptation in Chicken-Associated Serovars. *Microbiology and Molecular Biology Reviews*, *77*(4), 582–607. https://doi.org/10.1128/MMBR.00015-13

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), Article 5. https://doi.org/10.1038/s41587-019-0072-8

McDonagh, A., Leibler, J. H., Mukherjee, J., Thachil, A., Goodman, L. B., Riekofski, C., Nee, A., Smyth, K., Forrester, J., & Rosenbaum, M. H. (2019). Frequent human-poultry interactions and low prevalence of Salmonella in backyard chicken flocks in Massachusetts. *Zoonoses and Public Health*, *66*(1), 92–100. https://doi.org/10.1111/zph.12538

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., Jones, J. L., & Griffin, P. M. (2011). Foodborne Illness Acquired in the United States—Major Pathogens. *Emerging Infectious Diseases*, *17*(1), 7–15. https://doi.org/10.3201/eid1701.P11101

Vermont Department of Health. (2018, March 27). *Know How to Avoid Salmonella from Live Poultry*. Vermont Department of Health. https://www.healthvermont.gov/media/newsroom/know-how-avoid-salmonella-live-poultry-march-27-2018

Wu, X., Luo, H., Xu, F., Ge, C., Li, S., Deng, X., Wiedmann, M., Baker, R. C., Stevenson, A., Zhang, G., & Tang, S. (2021). Evaluation of Salmonella Serotype Prediction With Multiplex Nanopore Sequencing. *Frontiers in Microbiology*, *12*, 637771. https://doi.org/10.3389/fmicb.2021.637771

Yang, S.-M., Kim, E., Kim, D., Kim, H.-B., Baek, J., Ko, S., Kim, D., Yoon, H., & Kim, H.-Y. (2021). Rapid Real-Time Polymerase Chain Reaction for Salmonella Serotyping Based on Novel Unique Gene Markers by Pangenome Analysis. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.750379

Yoshida, C. E., Kruczkiewicz, P., Laing, C. R., Lingohr, E. J., Gannon, V. P. J., Nash, J. H. E., & Taboada, E. N. (2016). The Salmonella In Silico Typing Resource (SISTR): An Open Web-Accessible Tool for Rapidly Typing and Subtyping Draft Salmonella Genome Assemblies. *PLoS ONE*, *11*(1), e0147101. https://doi.org/10.1371/journal.pone.0147101