

Application of Next Generation Sequencing in a Local Public Health Laboratory – The Benefit of Innovative Partnering

Sanjib Bhattacharyya*, Manjeet Khubbar, Jose Navidad, and Steve Gradus: City of Milwaukee Health Department Laboratory

Background

Next generation sequencing (NGS) is opening opportunities for improved microbial detection and advanced genomic analysis, thus revealing fresh insight into how microbes impact humans, animal populations and the environment. Public health laboratories (PHLs) are exploring and integrating NGS to accelerate infectious disease surveillance, real-time outbreak investigations and strain typing, as well as improve definitive microbial ID. In the process, PHLs can uncover and describe the virulence factors and mechanism of antimicrobial resistance of microbial pathogens in complex clinical and environmental matrices at a genomic level. The Milwaukee Health Department Laboratory has shown the benefit of innovative partnering with academic and private sector researchers with NGS capabilities to conduct pilot projects on a variety of applications to address the unique problems and concerns of public health in our jurisdiction and beyond. Several studies were conducted at the MHD laboratory and other partner sites, which highlight the success of NGS applications in a local PHL and the benefits of this new technology for both individual patients and population health outcomes.

1. Metagenomic analysis of human fecal microbiome to understand microbial diversity and pathogen association with diarrheal disease outbreaks¹:

In an effort to delineate differences in human fecal microbiomes in symptomatic and non-symptomatic subjects from multiple geographic locations, 16S rRNA-based metagenomic approach was used. Sequencing was performed on Ion Torrent platform (Thermo Fischer Scientific/Life Technologies, CA) by our academic partner at the College of Biomedical Sciences Larkin Health Sciences Institute (South Miami, FL), followed by bioinformatics and statistical analysis using SPSS version 21.0 (IBM, IL). Screening of gastrointestinal (GI) pathogens followed by bacterial microbiome analysis shed light on clinical and epidemiology connections and understanding toward genomic diversity among gut microbiota (Figures 1-3).

Figure 1.

Rarefaction curve analysis shows that 203 different species were identified in 59 human stool samples associated with gastrointestinal illness collected from 5 geographic sites. Additional bacterial species associated with health issues and human diseases were also identified in the study samples.

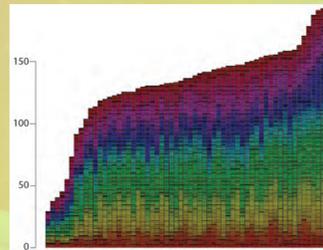


Figure 2.

Distribution of bacterial species (order) by site. The ANOVA analysis demonstrated a significant (p < 0.01) interaction. Follow-up tests revealed that the *Verrucimicrobiales* species was higher (P < 0.05) in WI state samples and those from Milwaukee (SE WI) (~25%) specimens compared to the other sites.

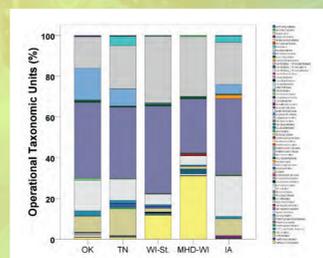
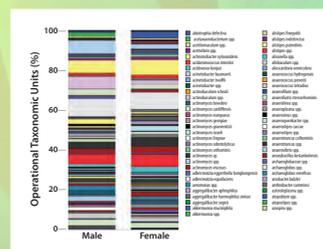


Figure 3.

Distribution of bacterial species (genus) by gender. *Butyrivomax* and *Lactatifermentans* bacterial species have been found to be associated with anti-inflammation properties and dairy consumption, respectively.



Methods and Results

2. Characterization of multi-drug resistant *Mycobacterium tuberculosis* complex from immigrants residing in the USA using full-gene sequencing^{2,3}:

NGS analysis was performed by private sector partner Longhorn Vaccines & Diagnostics (San Antonio, TX) on MHD TB isolates and compared to results from traditional drug sensitivity testing (DST). Ion Torrent sequencing was performed according to the manufacturer's instructions (Thermo Fischer Scientific/Life Technologies, CA) using 314 V2 chips with 200-bp sequencing chemistry. Sequencing runs produced average read lengths of 107 bp, greater than 30,000 high-quality (AQ20) reads, and a 300x average depth of coverage. Sequence assembly, alignments, and protein translations were performed using SeqMan NGen (V4) and LaserGene (V10) Core Suite (DNASTar, USA). Gene sequences for rpoB, katG, inhA, pncA, rpsL, gyrA, and rrs for all cases were deposited in GenBank (Accession numbers KC692347-KC692364). Genetic characterization of seven, full-length resistance-associated genes revealed two MDRs and one highly resistant strain with important drug-resistant mutations. NGS analysis revealed resistance genes not detected by conventional methods. Such detection of new resistance mutations has potentially major public health/epidemiologic implications for disease control of MDR TB at the local level and globally (Tables 1-2).

Table 1.

Comparison of genotypic and phenotypic drug resistance* Ion Torrent sequencing* and drug sensitivity testing

	Case 1		Case 2		Case 3	
	Ion Torrent	Bactec MGIT 960	Ion Torrent	Bactec MGIT 960	Ion Torrent	Bactec MGIT 960
Rifampin	Resistant	Resistant	Sensitive	Sensitive	Resistant	Resistant
Isoniazid	Resistant	Resistant	Resistant	Resistant	Resistant	Resistant
Pyrazinamide	Resistant	Resistant	Sensitive	Sensitive	Sensitive	Sensitive
Fluoroquinolone	Sensitive	Sensitive	Sensitive	Sensitive	Sensitive	Sensitive
Streptomycin	Resistant	Resistant	Resistant	Partial Resistance ^d	Sensitive	Sensitive
Resistance type	Multidrug resistant		Drug resistant		Multidrug Resistant	

* Critical concentrations for indicated antibiotics are: rifampin (2-9µg/ml); isoniazid (0-1 µg/ml and 0-4µg/ml); pyrazinamide (100µg/ml); fluoroquinolone (2 µg/ml); streptomycin (2-4 µg/ml and 6-0 µg/ml).
¹ Previously confirmed amino-acid substitutions giving rise to antibiotic resistance.
² Resistant at 2 µg/ml but sensitive at 6 µg/ml.

Table 2.

Amino-acid substitutions in drug resistance genes from *M. Tuberculosis* samples isolated from Burmese, Hmong, and Indian immigrants residing in the USA using Ion Torrent sequencing*

Isolate no.	rpoB (rifampin)	katG (isoniazid)	inhA (pyrazinamide)	pncA (fluoroquinolones)	gyrA (streptomycin)	rpsL (aminoglycosides/streptomycin)	rrs (16S)
Case 1	S531L	R463L	121T	K96R	S95T, E21Q, G688D	K43R	WT
Case 2	WT	S315T, R463L	WT	WT	WT	K43R	WT
Case 3	S531L	S315T, R463L	WT	WT	WT	WT	WT

WT refers to the 'wild type' H37Rv reference strain. * Amino-acid changes in bold indicate a known resistance mutation.

* Compare to previously confirmed amino-acid substitutions known to confer antibiotic resistance.
¹ Compared to H37Rv wild-type reference strain.
² Gene sequences: Case 1 (TB11-91), Case 2 (TB11-111), Case 3 (TB11-153) were deposited in Genbank (accession numbers KC692347-KC692364).

3. Whole genome sequence for Methicillin-resistant *Staphylococcus aureus*:

NGS library was prepared using the Nextera XT sample prep kit (Illumina, San Diego, CA, USA). MiSeq NGS was performed by private sector partner Longhorn Vaccines & Diagnostics (San Antonio, TX) according to the manufacturer's instructions (Illumina) using the MiSeq reagent kit (version 3) with 600 cycles. Following fragmentation, end preparation and sample tagging, the sequencer produced 6.2 million and 6.3 million paired reads for reference strains yielding appreciable coverage of ~200. The average read quality ranged between 34.8 and 35.3, indicating high-quality sequencing data (Figure 4). Assembly was performed using LaserGene DNASTar N-Gen version 12.1 (LaserGene, Inc., Madison, WI, USA) de novo assembly (algorithm: Velvet) and the control sequences were analyzed using BioNumerics v7.5. WGS of MRSA USA300 strain, along with mutation search for blue-light irradiated strains, allowed investigation of possible genetic effects underlying the antimicrobial effects of 470-nm blue light on MRSA and any genetic differences between irradiated and control MRSA strains, which likely shed light to future blue-light treatment on combating MRSA infections (Tables 3, 4a-b; Figure 5).

Figure 4.

The plot shows average sequence read quality between 34.8 and 35.3, assuring high quality sequencing data for treated as well as control MRSA strains.

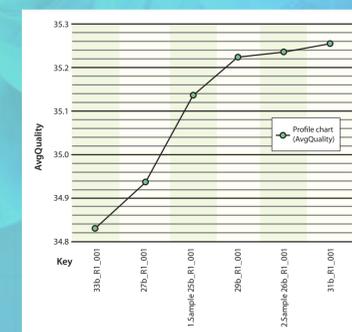


Table 3.

The length of the obtained *de novo* assemblies ranged from 2,866,141 up to 2,882,402 bp. The N50 values ranged between 41,087 bp and 101,466 bp, resulting in *de novo* assemblies between 185 and 94 contigs.

Key	N50	NrContigs	Length
27b_R1_001	73563	105	2881254
29b_R1_001	90967	96	2881121
31b_R1_001	101466	94	2880783
33b_R1_001	41087	185	2866141
Control1.Sample 25b_R1_001	82742	105	2882402
Control2.Sample 26b_R1_001	49178	154	2874512

Table 4a.

Mutation table versus Control 25b

Count	Entry	Label	Position reference	Position sequence	Type	NA change	AA change	Quality
412	33b		2882318	2882318	indel			10000.0000
413	27b		2882340	2882340	indel			0.2132
414	31b		2882346	2882346	indel			0.1890
415	29b		541573	541573	intergenic	t → g		0.0077
416	29b	conserved hypothetical protein	310411	310411	missense	c → a	A → S	0.0077
417	29b	conserved hypothetical protein	531684	531684	missense	a → g	D → G	0.0120
418	27b	conserved hypothetical protein	534323	534323	missense	g → c	D → H	0.0102
419	31b	conserved hypothetical protein	2281089	2281089	missense	c → a	V → L	0.0129
420	27b	methicillin resistance protein	2771639	2771639	missense	c → a	L → I	0.0140
421	29b	sdrD protein, putative	2111639	2111639	silent	a → g		0.0819

Table 4b.

Mutation table versus Control 26b

Count	Entry	Label	Position reference	Position sequence	Type	NA change	AA change	Quality
607	27b		2874434	2882318	indel			10000.0000
608	29b		2874434	2874434	indel			10000.0000
609	31b		2874434	2874434	indel			10000.0000
610	33b		2874434	2874434	indel			10000.0000
611	29b		195227	195227	intergenic	t → g		0.0085
612	29b	conserved hypothetical protein	185338	185338	missense	a → g	D → G	0.0159
613	27b	SA2099	187977	187977	missense	g → c	D → H	0.0123
614	27b	methicillin resistance protein	1599919	1599919	missense	c → a	L → I	0.0140
615	31b	conserved hypothetical protein	2299490	2299490	missense	c → a	V → L	0.0129
616	29b	conserved hypothetical protein	2520885	2520885	missense	g → t	A → S	0.0228

Figure 5.

Circular view of the blue-light irradiated and control MRSA whole genomes



* Presenting Author:

Email: sbhatt@milwaukee.gov

Telephone: 414-286-5702

Presented at AMD Day, CDC, September 2016

Conclusions

- **NGS applications in a local PHL useful for:**
 - Understanding genomic diversity in gut microbiota
 - Efficient and improved detection of drug-resistant *M. tuberculosis* Complex
 - Elucidating mechanism of light-irradiated MRSA treatment
- **Bioinformatics challenges:** Lack of expertise, cost for data storage and multiple different commercial and proprietary software, preference of use
- **Innovative partnerships:** Successful partnership allows bringing in applications of AMD technologies (e.g., real-time PCR, digital droplet PCR, bead-based multiplexing, Sanger and/or NGS) in PHLs
- **Exploring unconventional funding support:** Limited federal funding limits start-up costs for instrumentation, NGS data analysis and bioinformatics – critical for sustained NGS capacity in PHLs

Acknowledgments

We acknowledge the staff at the City of Milwaukee Health Department Laboratory, and all the support from partner sites including:

- College of Biomedical Sciences, Larkin Health Sciences Institute (South Miami, FL)
- Center of Computational Science, University of Miami (Miami, FL)
- Longhorn Vaccines & Diagnostics (San Antonio, TX)
- LeBonheur Hospital (Memphis, TN)
- Oklahoma State Department of Health (Oklahoma City, OK)
- Wisconsin State Laboratory of Hygiene (Madison, WI)
- State Hygienic Laboratory University of Iowa (Ankeny, IA)
- University of Wisconsin-Milwaukee College of Health Sciences (Milwaukee, WI)

Special thanks to Nancy Hills for graphics and Julie Becker for editorial review of the poster.

Funding Sources:

Genome Sequence for Methicillin-resistant *Staphylococcus aureus* study was supported by a UW-Milwaukee College of Health Sciences Stimulus for Enhancing Extramural Development (SEED) grant award, along with a Clinical and Translational Science Award (CTSA) to the UWM College of Health Sciences.

