Hepatitis C Virus Genotyping Using Next-Generation Sequencing: An Efficient Alternative to Sanger Sequencing

To the Editor.—The global prevalence of hepatitis C virus (HCV) is estimated to be approximately 146 to 219 million people.[1] Of those infected, it has been shown that 4 out of 5 will develop a chronic infection. In addition, 33% of chronically infected patients progress to cirrhosis during a period of 20 to 30 years. The patients with cirrhosis may develop hepatocellular carcinoma at a rate of 2% to 5% per year.[2] HCV has a positive-sense, single-stranded RNA genome, and it has 7 major genotypes with more than 70 subtypes.[3]

Hepatitis C virus genotypes vary in geographical distribution and treatment response varies according to genotype.[4] In clinical diagnostic laboratories, HCV genotyping has traditionally been carried out using polymerase chain reaction or first-generation (Sanger) sequencing–based technologies. Issues with the current mainstream genotyping methods include their limited accuracy in subtyping calls and their inability to detect coinfection, which is estimated to be present in 2% to 7% of patients.[5]

We evaluated the performance characteristics of next-generation sequencing (NGS) in HCV genotyping to increase speed, reduce costs, and generate more accurate coinfection profiles. We accomplished this by generating complementary DNA (cDNA) libraries of plasma samples, which we then sequenced with an Illumina MiSeq (Illumina, San Diego, California). We also developed an in-house bioinformatics workflow, FH-HCV-GT, to analyze and genotype our data and compared the results with commercial software, ABL SA-DeepChek-HCV (Advanced Biological Laboratories, Luxembourg, Luxembourg), available for HCV genotyping.

Sequence data of more than 2 000 000 reads, with a mean length of 200 bases per read, were generated from 15 samples. All samples had 10 000 reads or more. There was 100% concordance of genotypes with the Sanger method when using either ABL SA-DeepChek-HCV or FH-HCV-GT programs for genotyping calls (see Table 1). We were able to detect coinfection of 2 different HCV subtypes with an equal mixture of subtypes 1a and 2b and our NGS technique (see Table 1). Of note, the sequencing data by NGS do not reflect the exact percentage of the intended equal mixture. This is likely due to an amplification bias of the consensus primers used. Further optimization of the polymerase chain reaction conditions is needed to achieve the best efficacy of detecting coinfections.

We found that sample preparation time was 16.25 min/sample using the Sanger method and 12.1 min/sample using the NGS technique. Table 2 shows a comparison of the sample preparation time for each method. Reagent cost per sample was $32 to process 30 samples at a time using our NGS technique, with a potential cost of $17 to process 96 samples at a time (see Table 3). In comparison, reagent cost per sample using the Sanger method was $23, but the throughput was much smaller, with a maximum of 16 samples/run (Table 4). The cost-efficient breakpoint for NGS was 52 samples/run; after that, it becomes more economical (in terms of reagent cost, not even considering the saved technologist time using NGS) to use NGS than Sanger.

Our data indicate that HCV genotyping by NGS technology is as accurate as Sanger sequencing, the current gold standard. Our results indicate that NGS has the potential to be more cost effective with a quicker turnaround and simultaneously offering greater throughput. One of the potential important clinical differences between Sanger and NGS technologies is the ability of NGS to detect coinfections by different strains of HCV.

Table 2. A Comparison of Sample Prep Time Using Both Sanger and Next-Generation Sequencing (NGS)

Sequencing Process	Hands-on Time, min		
	Sanger, n = 16 Samples	NGS, n = 30 Samples	NGS, n = 96 Samples
Sample accessioning	24	45	144
DNA extraction	24	45	144
PCR amplification setup	30	52.5	110
Cleanup/washes	30	45	30
QC check	15	30	96
Preparation for sequence run	40	67.5	90
Denature	2	3	2
Program sequencer	5	7.5	5
Analysis	80	60	180
Review and manual entry	10	7.5	25
Total time	**260**	**363**	**826**
Time/sample	16.25	12.1	8.6

Abbreviations: PCR, polymerase chain reaction; Sanger, Sanger sequencing.

Table 1. Hepatitis C Virus Genotyping and Software Analysis Validation

Subtype by Sanger	Sample, No.	Read Range	Subtyping by NGS; Analyzing With DeepChek[a]	Subtyping by NGS; Analyzing With FH-HCV-GT,[b] Type (%)
1a	7	16 529–119 112	1a	1a
1b	2	55 135–65 659	1b	1b
2a	1	22 240	2a	2a
2b	2	13 717–45 505	2b	2b
3a	3	20 435–104 031	3a	3a
1a and 2b	2	37 527–46 158	NA	1a (4–6), 2b (94–96)

Abbreviations: NA, not applicable; NGS, next-generation sequencing.

[a] Illumina MiSeq (Illumina, San Diego, California).
[b] FH-HCV-GT is an in-house bioinformatics workflow process.

Table 3. Costs of Next-Generation Sequencing (NGS) Components

Illumina Sequencing Components	Quantity, No. Samples	Price,[a] US$
PhiX Control[b]	192	160
Nextera Index Set[b]	96	950
Nano 500 Flowcell[b]	1 Flowcell	639
Cost/sample[c]	30	32
Cost/sample[c]	52	23
Cost/sample[c]	96	17

[a] Prices are list prices from 2016.
[b] Illumina, San Diego, California.
[c] Cost/sample = (160/192) + (950/96) + [639/No. samples(30, 52, or 96)].

Table 4. Costs of Sanger Sequencing Components

Sanger Sequencing Component	Quantity, No. Samples	Price,[a] US$
BigDye Terminator v1.1 Cycle Sequencing Reagents-CG[b]	100	1314
BigDye XTerminator Purification Kit[b]	100	245
POP-6 Polymer for 3130/3130xl Genetic Analyzers[b]	200	206
3500 Dx Series Sequencing Standard, BigDye Terminator v1.1[b]	200	292
3130xl/3100 Genetic Analyzer 16-Capillary Array, 36 cm[b]	200	1074
Cost/sample[c]	16	23

[a] Prices are list prices from 2015.
[b] Thermo Fisher Scientific Life Sciences, Waltham, Massachusetts.
[c] Cost/sample = (1314/100) + (245/100) + (206/200) + (292/200) + (1074/200).

In conclusion, there exists tremendous potential for the application of NGS in the clinical laboratory for cost savings, efficiency, and more-accurate, clinically relevant genotyping information, particularly for laboratories that are already performing NGS for other diseases.

Cameron Wales, BS[1]; George Corpus, MS[2]; ChungChe Chang, MD, PhD[1,2]

[1] College of Medicine, University of Central Florida, Orlando; [2] Molecular Diagnosis Laboratory, Florida Hospital, Orlando

1. Lavanchy D. The global burden of hepatitis C. *Liver Int*. 2009;29(suppl 1):74–81.
2. Myers RP, Krajden M, Bilodeau M, et al. Burden of disease and cost of chronic hepatitis C virus infection in Canada. *Can J Gastroenterol Hepatol*. 2014;28(5):243–250.
3. Smith DB, Bukh J, Kuiken C, et al. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. *Hepatology*. 2014;59(1):318–327.
4. American Association for the Study of Liver Diseases (AASLD); Infectious Disease Society of America (IDSA). HCV guidance: recommendations for testing, managing, and treating hepatitis C. http://www.hcvguidelines.org/full-report-view. Accessed on December 12, 2015.
5. Sy T, Jamal MM. Epidemiology of hepatitis C virus (HCV) infection. *Int J Med Sci*. 2006;3(2):41–46.

The authors have no relevant financial interest in the products or companies described in this article.

doi: 10.5858/arpa.2016-0338-LE