Genome Sequencer System

RNA Virus Sequencing Using the Genome Sequencer FLX System

454 *LIFE SCIENCES*

RNA Virus Sequencing Using the Genome Sequencer FLX System

Corresponding author: Jan F. Simons and Stephen K. Hutchison, 454 Life Sciences Corporation, Branford, CT, USA.
Email: shutchison@454.com

Introduction

The global spread of the deadly H5N1 avian influenza virus among bird populations and its transmission to humans in endemic areas has emphasized the importance of fast and accurate detection methods for influenza virus strain and variant identification. Standardized assays, typically based on immunological methods, are available for strain classification and are routinely used in epidemiological studies. Other more recently developed methods rely on locus specific PCR combined with assays that interrogate specific genomic regions such as those defining clades, strain or drug resistance (Habib-Bein *et al.*, 2003, Poddar *et al.*, 2002, Pourmand *et al.*, 2006), but are ill-suited for genome-wide analysis.

Immunological assays are widely used but they provide limited information for the understanding of *e.g.*, strain evolution and like the site-specific PCR-based methods might not be useful in the case of novel viral variants. Two recent studies emphasize the value of whole genome influenza virus surveillance sequencing as a way to overcome these shortcomings and further our understanding of influenza virus. In total several hundreds of influenza virus genomes from field isolates collected over several years were sequenced and analyzed for segment reassortment (antigenic shifts), antigenic drift, point mutations and deletions for the understanding of viral evolution and epidemiological dynamics (Holmes *et al.*, 2005, Ghedin *et al.*, 2005). These studies showed that the diversity within the viral reservoir among humans is larger than previously anticipated and that reassortment events between different clades and serologically distinct viruses contribute to the rise to novel lineages and can explain lack of vaccine efficacy. It has also been noted that whereas antigenic and genetic evolution are correlated, the latter is more gradual and might at an earlier stage reveal subtle genetic changes leading to altered tropism, antigenicity or drug response (Smith DJ, *et al.*, 2004).

The significance of monitoring changes on the nucleotide level was further emphasized by a study of the reconstructed H1N1 strain which caused the Spanish flu epidemic in 1918. In this case, minor amino acid changes to the hemagglutinin (HA) protein were shown to drastically change the tropism of the virus by increasing its ability to bind 2,6 sialic acid, abundant in the human respiratory tract (Tumpey *et al.*, 2007). Likewise, a few mutations localized to the M2 protein (Belshe *et al.*, 1988) form the basis for a recent upsurge in adamantane resistance currently observed in several geographic locations (Bright *et al.*, 2005). Importantly, whereas these mutations have dramatic and predictable consequences for the survival and spread of certain strains, the homologous mutations in another strain might not have the same impact (Tumpey *et al.*, 2007). Large-scale whole genome sequencing efforts provide a means to detected genetic changes associated with such dramatic phenotypic changes in time to anticipate and prevent the spread of particularly pathogenic strains with the potential to cause pandemics.

The influenza virus genome consists of 8 unique segments, comprising approximately 13,500 bases and encoding a total of 11 proteins. Until recent large-scale whole genome sequencing efforts (Holmes *et al.*, 2005, Ghedin *et al.*, 2005), partially under the umbrella of the Influenza Genome Sequencing Project (Fauci, 2005), sequencing information related to field isolates has been nearly exclusively limited to HA and neuraminidase (NA), the major antigenic epitopes. Template generation has relied on PCR amplification of specific loci - in the case of whole genome sequencing, up to 96 overlapping amplicons – and have been limited to analysis of variants similar in sequence to the consensus sequence used for primer design.

Here we demonstrate for the first time whole genome sequencing of influenza virus on the Genome Sequencer FLX using a unique cDNA library protocol that is independent of locus specific amplification primers and which allows RNA input amounts as low as ten nanograms. The protocol, which is also applicable to mRNA, takes advantage of temperature-induced RNA fragmentation, effectively eliminating the need to synthesized full-length cDNAs, and generates a library of cDNA fragments flanked by adaptors enabling emPCR amplification and sequencing.

Materials and Methods

Procedure

Library Preparation

Viral RNA of the A/PR/8/34 (HlN1) strain isolated from virus grown in specific pathogen free (SPF) embryonated eggs was purchased from Charles River Laboratories (Franklin, CT). To reduce the substantial amounts of contaminating nucleic acids commonly observed in viral RNA preparations an initial and optional procedure to enrich for influenza virus RNA was introduced. This step involved capturing and enriching the viral RNA molecules through the hybridization of a biotin-labeled oligonucleotide directed to the conserved 5'-end of all 8 segments of the influenza A viral genome. The procedure was done as follows: 200 ng of isolated viral RNA (quantified by the Quant-iT Ribogreen assay kit; Invitrogen Corporation, Carlsbad, CA) in 15 µl of 10 mM Tris-HCl (pH 7.5) was incubated in the presence of 50 µl of 6× SSPE buffer (0.9 M NaCl, 60 mM NaH_2PO_4, 7.5 mM EDTA) containing 0.1 units/µl of SUPERase-In RNase inhibitor (Ambion, Austin, TX) and 0.5 µM of the 5'-Capture Oligo (5'-CCT TGT TTC TAC T-biotin-3'; IDT, Coralville, IA) at 70°C for 5 minutes followed by 15 minutes at 39°C. Subsequently, 50 µl of 1× Binding and Wash Buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl) containing 0.05% Sera-Mag 30 beads (Seradyn Inc, Indianapolis, IN) was added and allowed to incubate on a bench top rotator at room temperature for 30 minutes. With the aid of a magnetic particle capture unit the beads were washed once with 0.5× Binding and Wash Buffer and twice with Bead Wash Buffer (10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 30 mM NaCl, 0.1% Tween-20). To recover the isolated viral RNA the beads were placed in 10 µl of 10 mM Tris-HCl (pH 7.5) and heated to 65°C for 5 minutes. The liquid containing the released viral RNA was immediately recovered leaving the empty beads behind.

The viral RNA, once enriched, was converted into single-stranded cDNA (sscDNA). First, the RNA was fragmented into a size range compatible in length with sequencing on the Genome Sequencer FLX. To this end, 2.5 µl of 5× Fragmentation Buffer (200 mM Tris-Acetate, 500 mM potassium acetate, 157.5 mM magnesium acetate [pH 8.1]) was added to the viral RNA (in 10 µl of 10 mM Tris-HCl [pH 7.5]) and was after thorough mixing incubated for 2 min at 82°C and immediately thereafter transferred to ice to stop the fragmentation reaction. The reaction volume was increased to 50 µl with 10 mM Tris-HCl (pH 7.5) and the sample was purified with 80 µl (or 1.6× the volume of the sample) of RNA-Clean (Agencourt, Beverly, MA) for 10 min at room temperature. The mixture was then washed as per the manufacturer's instructions and the fragmented material was eluted from the beads in a volume of 9.5 µl of 10 mM Tris-HCl (pH 7.5).

The fragmented RNA was placed in the presence of 2 µl of 100 µM reverse transcription primer (5'-phosphate-$TNNNTN_6$-3'; IDT) and heated to 70°C for 10 min followed by snap cooling on ice. Next, the RNA was reverse transcribed in a total volume of 20 µl and a final composition of 1× First Strand Synthesis buffer (Invitrogen, Corporation, Carlsbad, CA), 500 µM dNTPs, 10 mM DTT, 10 units of SUPERase-In RNase inhibitor (Ambion, Austin, TX) and 200 U of Superscript II enzyme (Invitrogen Corporation, Carlsbad, CA). The reaction was incubated at 25°C for 5 min and then at 37°C for 60 min. The RNA was removed by hydrolysis by the addition of 20 µl of Denaturation Solution (0.5 M NaOH, 0.25 M EDTA) and incubation at 65°C for 20 minutes. To neutralize the mixture a solution of 0.5 M HCl in 0.5 M Tris-HCl (pH 8.0) was added to bring the pH to between 7.0 and 8.5 (the volume of 0.5 M HCl needed to achieve the

required pH was determined through titration before hand). The resultant sscDNA was recovered with RNAClean solution (1.6× the volume of the sample) as per the manufacturer's instructions and was eluted from the beads with 12 μl of 10 mM Tris-HCl (pH 7.5).

Adaptor Ligation

For clonal amplification and sequencing on the Genome Sequencer FLX the sscDNA requires the addition of adaptors to each terminus. The adaptors have been designed to enforce directional ligation to the sscDNA, such that one will be uniquely ligated to the 5'-end (sscDNA Adaptor A) and the other to the 3'-end (sscDNA Adaptor B) of the sscDNA. Each adaptor is comprised of two complimentary oligonucleotides which are annealed together as described below. The 3'-end adaptor consists of "sscDNA Oligo B" (5'-biotin-GCCTTGCCAGCCCGCTCAGNNNNNN-phosphate-3') and "sscDNA Oligo B-prime" (5'-phosphate-CTGAGCGGGCTGGCAAGG-dideoxyC-3'), which after annealing result in "sscDNA Adaptor B" with a 3'-random overhang of six nucleotides. Similarly, the 5'-end adaptor consists of "sscDNA Oligo A-prime" (5'-NANNACTGATGGCGCGAGGAGG-dideoxyC-3') and "sscDNA Oligo A" (5'-GCCTCC CTCGCGCCATCAG-3') which form "sscDNA Adaptor A" with a five nucleotide 5'-end overhang. Annealing of the adaptors was done as follows: 1) for "sscDNA Adaptor A", Oligo A and Oligo A-prime were diluted to a final concentration of 50 μM and 60 μM, respectively, in a total volume of 100 μl 10 mM Tris-HCl (pH 7.5); 2) correspondingly for "sscDNA Adaptor B", Oligo B and Oligo B-prime were diluted to a final concentration of 240 μM and 200 μM, respectively, in a final volume of 100 μl 10 mM Tris-HCl (pH 7.5). For maximum annealing of the adaptors the following thermocycler conditions for were used: 5 min at 80°C, 7 min at 65°C, 7 min at 60°C, 7 min at 55°C, 7 min at 50°C, 7 min at 45°C, 7 min at 40°C, 7 min at 35°C, 7 min at 30°C, 7 min at 25°C, hold at 4°C.

The adaptor ligation reaction was done by placing the sscDNA in a total volume of 30 μl 1× Quick Ligase Buffer (New England Biolabs, Ipswich, MA) containing 1.67 μM of the Adaptor A, 6.67 μM of the Adaptor B and 2000 units of T4 DNA Ligase (New England Biolabs, Ipswich, MA) and incubating at 37°C for 2 hours. The reaction was terminated with the addition of 70 μl of 1× TE (pH 8.0). To recover the ligated sscDNA 100 μl of Binding and Wash Buffer containing 0.05 % Sera-Mag 30 beads (Seradyn Inc, Indianapolis, IN) was added and allowed to mix for 10 minutes at room temperature. Unbound material was washed away with 200 μl 0.5× Binding and Wash Buffer followed by two 200 μl washes of Bead Wash Buffer. The ligated products were eluted from the beads with 90 μl of Bead Elusion Buffer (25 mM NaOH, 1 mM EDTA (pH 8.0), 0.1 % Tween-20). The final sscDNA library was subjected to two rounds of purification with RNAClean as per the manufacturer's directions except the amount of beads was reduced to 1.6× the volume of the sample. After the second RNAClean step the final adapted and purified sscDNA library was eluted from the beads in 12 μl of 10 mM Tris-HCl (pH 7.5).

A small aliquot (1 μl) of the final sscDNA library was run on an RNA 6000 Pico chip on a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) to confirm proper size distribution. The material was then quantified with the Quant-iT Ribogreen RNA Assay Kit (Invitrogen Corporation, Carlsbad, CA) on a Synergy HT (Bio-Tek Instruments Inc, Winooski, VT) instrument following the manufacturer's instructions. Based on this data an aliquot of the library was typically diluted to a working concentration of 2×10^5 per μl for immediate use in the emPCR step.

emPCR

When starting with RNA samples of less than 200 ng or if there is substantial amounts of non-viral RNA in the unenriched starting material the final adapted sscDNA will typically require an amplification step to enable accurate quantification of final library. The PCR reaction is done in 50 μl containing 3 μl of the sscDNA (25% of total available), 2 μM each of Primer A (5'-GCC TCC CTC GCG CCA-3') and Primer B (5'-GCC TTG CCA GCC CGC-3'), 400 μM dNTPs, 1× Advantage 2 buffer and 1 μl of Advantage 2 polymerase mix (Clontech, Mountain View, CA). The amplification reaction is performed under the following conditions: 96°C for 4 min. 94°C for 30 sec, 64°C for 30 sec, repeating steps 2 and 3 for a total of 20 cycles, followed by 68°C for 3 minutes and finally 14°C until the samples are removed from the PCR instrument. After the amplification the samples are subjected to two rounds of AMPure bead treatment as per the manufacturer's instructions (Agencourt).

The final double-stranded library is eluted in 12 μl of 10 mM Tris-HCl (pH 7.5) and analyzed on a DNA 7500 LabChip on a 2100 Bioanalyzer to ensure significant amplification. The resultant dscDNA products are quantified with the Quant-iT Picogreen dsDNA Assay Kit (Invitrogen Corporation) on a Synergy HT (Bio-Tek Instruments Inc, Winooski, VT) instrument following the manufacturer's instructions. Once the amplified material has been quantitated a portion can be diluted to a working concentration, as described above, for the emPCR step.

The library was taken through emPCR following the standard GS emPCR Kit User's Manual. Both GS emPCR Kit II and Kit III were used to generate beads allowing for sequencing of the sscDNA molecules in both directions and thereby achieving more even 5'-to-3'-coverage of each genomic segment. Two reactions containing 600,000 beads and 2 copies of template sscDNA per bead (approximately 1.2 million molecules) were made for each

GS emPCR kit. DNA containing beads were enriched, counted and an equal number (15,000 beads each) of Kit II and Kit III beads were pooled and deposited onto a 16 lane (2 × 53 mm lanes) PicoTiterPlate following the standard protocol for bead deposition. The samples were sequenced on a GS FLX instrument under standard operating conditions for 100 cycles of nucleotide flows (Genome Sequencer FLX Operator's Manual).

The sequencing data was analyzed using software version 1.1.01 and the quality filtered sequences (totalPassedFiltering) were retained. Before further analysis the first 10 bases immediately following the A adaptor were removed as these originated from the semi-randomized TNNTN6 primer used for reverse transcription. The resulting sequences were aligned using BLAST 2.0 against the Puerto Rico A/PR/8/34 (HlN1) viral sequence obtained from the Influenza Sequence Database (ISD) **(http://www.flu.lanl.gov/)**.

Results

Library Preparation

In this application note we describe a library preparation protocol that allows for the sequencing of the influenza A viral genome with the Genome Sequencer FLX Instrument. The protocol differs from traditional cDNA or amplicon approaches currently used to sequence influenza virus in that no target specific primers are required and problems associated with full-length cDNA synthesis are evaded. The initial step in this protocol involves enrichment of the viral RNA from contaminating RNAs (host mRNA, mtRNA, rRNA or small RNA molecules) and thus ensures that a maximum number of sequencing reads will correspond to the viral material. This is accomplished by annealing a biotinylated capture-oligonucleotide mix to the conserved 5'-ends of each genomic segment and

enriching the bound viral RNA onto streptavidin beads. Figures 1A and 1B illustrate the RNA size profile before and after applying the enrichment procedure to a sucrose-gradient purified viral preparation . As shown, even with a rather clean viral preparation a substantial amount of smaller sized contaminating material is removed by the enrichment procedure. A small loss of viral material is noted, however the major products in the final RNA sample correspond to the viral RNA. The importance of this step is more evident when working with unpurified viral RNA samples harvested from cell cultures and associated with proportionally larger amounts of host contaminating RNA. In such cases the enrichment method described here has achieved up to 100-fold enrichment of the viral RNA (data not shown).
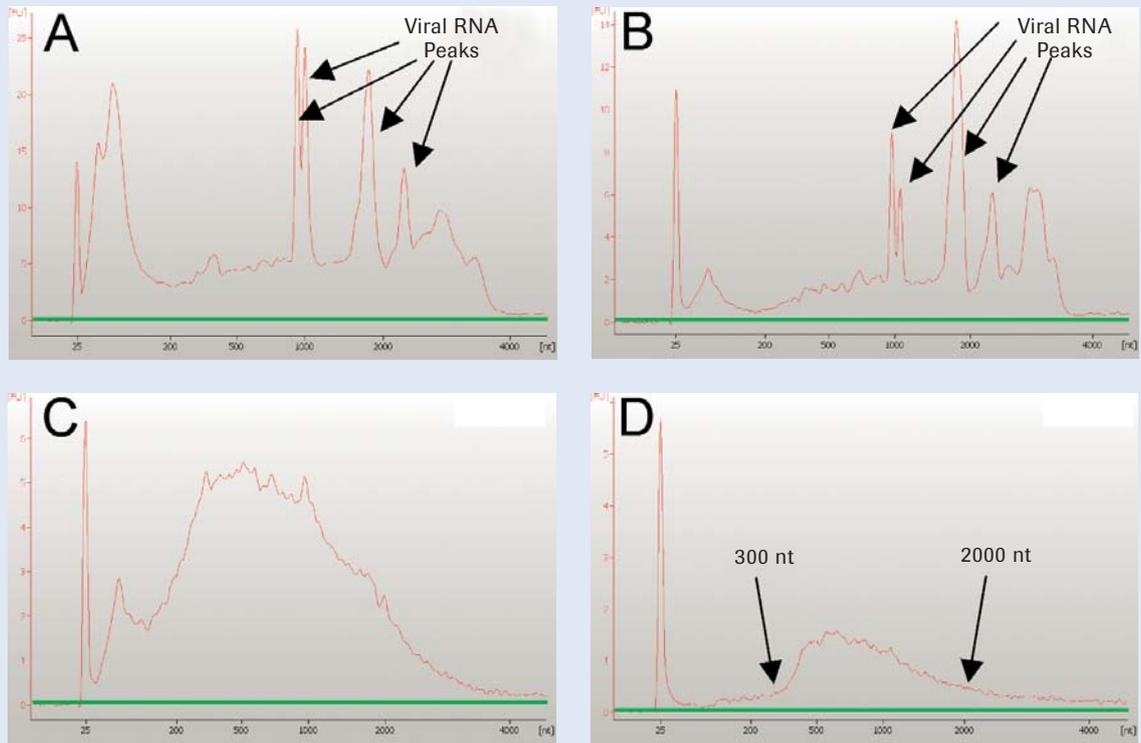
Figure 1: 2100 Bioanalyzer size distribution profiles of the RNA/cDNA products generated at the different steps of the sample preparation protocol. A) Stock sample of the A/PR/8/34 (HIN1) influenza viral RNA. Peaks corresponding to the segments of influenza virus are indicated. **B)** Influenza virus sample after the viral RNA enrichment step. **C)** Viral RNA post-fragmentation. **D)** Final adapted sscDNA library.

A key feature of the protocol is the heat-induced fragmentation of the RNA prior to the RT-mediated cDNA synthesis step. This reduces coverage bias of distal parts of the RNA typical of protocols relying on full-length cDNA synthesis. In fact, the random fragmentation method augments recovery and sequencing of the very 5'-end of template RNAs. The conditions presented here for RNA fragmentation generated a fragment size distribution from slightly less than 100 nucleotides to just above 2000 nucleotides with a majority of the material ranging from 300 to 900 nucleotides (Figure 1C). The majority of the small sized material is removed by the RNAClean bead purification in the subsequent steps.

Single-stranded cDNA (sscDNA) was made using Superscript II reverse transciptase (Invitrogen) and a 5'-phosphorylated TNNTN6-OH primer. Following RNA hydrolysis and a purification step the sscDNA fragments are ready for direct ligation to the adaptors, which are designed to enforce directional ligation. Figure 2 depicts the ligation of each adaptor to the 5'- or 3'-ends of the sscDNA molecules. Through a biotin incorporated into the non-ligated strand of the 3'-adaptor, the ligated material is bound to magnetic streptavidin beads and cleansed of unligated sscDNA and excess adaptors. Following elution of the sscDNA and cleanup steps a final adaptor ligated library ranging in size from 300 to about 2000 nucleotides was obtained (Figure 1D).
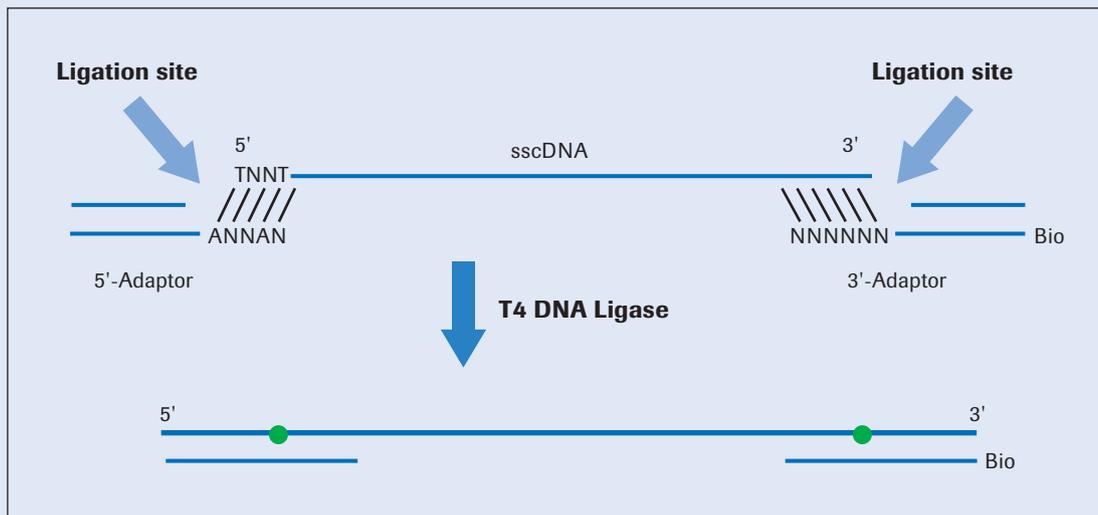
Figure 2: Schematic representing the ligation of the adaptors to their specific termini of the sscDNA via T4 DNA ligase. Diagonal lines between the sscDNA and the adaptor units represent base pairing between the randomized ends of the adaptors and the 5'- and 3'-ends of the sscDNA. Green circles in the final product represents ligation events.

The final yield of single-stranded library was determined by fluorometry to be 6.9 ng, which is typical when starting with 200 ng of RNA. The library was quantified by fluorometry and processed through emulsion PCR at a concentration of 2 molecules of sscDNA to each capture bead with a total of 600,000 beads per reaction. A total of 53,000 enriched beads were recovered using 2 reactions for the GS emPCR Kit II and 2 reactions for 32,900 beads for Kit III, representing enrichment efficiencies of 4.5% and 2.5%, respectively. To maximize sequence coverage, especially at the ends of each viral segment, samples were sequenced from both the A and the B primers simultaneously. To this end, 15,000 beads of each type were pooled along with 1,000 run control beads and loaded into 2 × 53 mm PicoTiterPlate regions (16 lane format). The PicoTiterPlate was run on the Genome Sequencer FLX Instrument under standard sequencing conditions of 100 cycles of nucleotide flows.

Sequencing Data Analysis

Standard analysis of the sequencing run resulted in 14,323 quality filtered reads generating a total of 3,580,065 nucleotides with an average read length of 250 bases. BLAST 2.0 analysis with an e value of 1.0E-6 to the Puerto Rican A/PR/8/34 (HlN1) genomic sequence gave 15,437 hits. All BLAST hits with lengths less than 35 nucleotides or reads that hit more than one region in the reference were discarded leaving 14,025 reads (97.9 % of the high quality reads) for a total of 3,509,701 quality filtered nucleotides. Table 1 presents a break down of the final sequencing results over all 8 segments of the genome. A total of 99.51% of the entire genome and 100% of the coding regions were covered.

Depth of coverage plots were generated for the 8 segments and an average of 100 to 200 fold coverage was seen for all segments except segment 3 which is lower (examples in Figure 3). A significantly higher coverage was observed for the last 245 to 260 nucleotides corresponding to the 5'-end of the viral RNA segments. These reads started at the very 5'-end of the original RNA and their length correspond to the read length expected for a 100 cycle sequencing run. A simulated coverage curve for a 100 cycle sequencing run of randomly shared and primed cDNAs was consistent with the observed elevation in 5'-end coverage (data not shown).
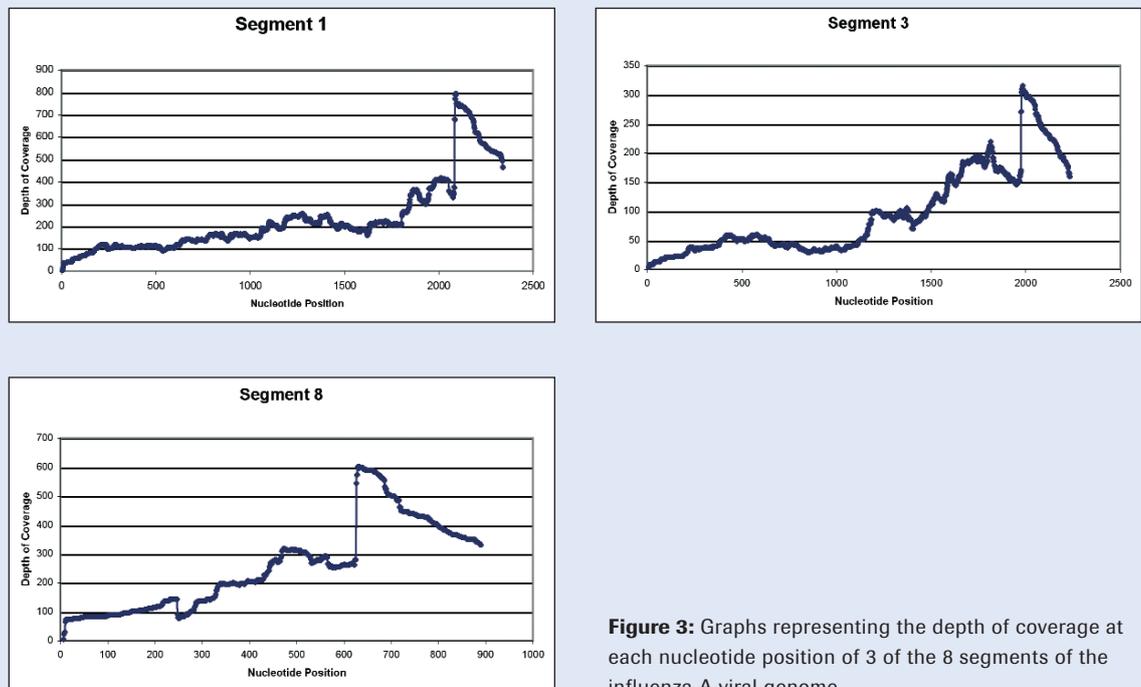

Figure 3: Graphs representing the depth of coverage at each nucleotide position of 3 of the 8 segments of the influenza A viral genome.

The lower coverage of the ultimate 3'-end of the viral RNA may be attributed to a lower frequency of hybridization of the reverse transcription primer to this portion of the RNAs. Alternatively, the 3'-ends of the RNAs may either form an uncharacterized secondary structure rendering them inaccessible or be degraded, either of which is consistent with poor success trying to enrich the viral RNAs through the conserved 3'-ends (data not shown).

Segment	Length (nt)	Size of coding region (nt)	Number of sequencing reads	% of the nucleotides covered	Coding Completeness (%)
PB2	2341	2277	2213	99.53	100
PB1	2341	2271, 261	1858	99.74	100
PA	2233	2148	868	99.42	100
HA	1775	1698	1927	100	100
NP	1565	1494	2002	99.3	100
NA	1413	1362	3370	99.5	100
MP	1027	756, 291	878	99.32	100
NS	890	690, 363	909	98.65	100

Table 1: Distribution of the high quality reads to the 8 segments of the Influenza A Puerto Rico A/PR/8/34 (H1N1) genome.

Assembly of the sequences corresponding to each segment was performed with Contig Assembly Program (CAP, Huang, 1991) with criteria set at 20 base minimum overlap with a minimum match of 85%. Comparison of consensus sequence to the A/PR/8/34 reference sequence downloaded from the Influenza Sequence Database (ISD) **(http://www.flu.lanl.gov/)** revealed 1 sequence variation resulting in an amino acid substitution in segment 1 encoding the PB2 polypeptide. The change (A substituted for G) is observed at nucleotide position 1705 and causes a valine to isoleucine switch. Analysis of the individual reads exposed 221 that were used to generate the final contig spanning this nucleotide. Of these, 3 or 1.36% corresponded to the nucleotide sequence from the database. The observed difference may be a true representation of a mutation resulting from years of *in vitro* culturing of this viral strain. A comparison of several submitted sequences of the A/PR/8/34 strain from the Influenza Sequence Database (ISD) revealed some variations at other genomic positions, however this is the first report of a sequence variant at position 1705 of PB2.

Conclusion

Here we demonstrate sequencing of the Influenza A genome on the Genome Sequencer FLX Instrument, by employing a unique library preparation method that maximizes coverage across the entire viral genome. The protocol has been shown to work reproducibly with as little as 10 ng input RNA and involves heat fragmentation of the initial RNA sample, generation of single-stranded cDNA (sscDNA) by way of randomly primed reverse transcription and direct ligation of adaptors to the sscDNA. The final library averages 400 and 600 nucleotides in length and allows efficient emPCR and sequencing on the Genome Sequencer FLX Instrument. Unlike PCR-based approaches this method allows the viral genome to be sequenced without any prior knowledge of specific sequence variants or segmental rearrangements. This is of particular importance as the large sequence variability found among viral isolates sometimes makes strategies targeting specific genomic regions constrained in their use to specific viral clades, or isolates from restricted geographic locations (Spackman *et al.*, 2002).

To accommodate the need for processing multiple samples in parallel, the procedure is based on the use of 96 well plates and allows manual preparation of 96 samples at a time, with the possibility of adapting the entire process to a robotic platform to further increase throughput. By using the small multi-lane format GS FLX sequencing format, up to 16 samples can be sequenced simultaneously, while achieving genome coverage in excess of 99%. The depth of oversampling across the genome achieved with this sequencing format is typically 10 to 100-fold (excluding the very 3'-end) allowing for determination of a high-confidence consensus sequence and the possibility to detect lower level sequence variants. The detection of very low level mutations can be accomplished by simply scaling up the number of emulsion PCR reactions and performing the sequencing on a larger PicoTiterPlate format, making investigations into viral diversity and evolution within a single host possible.

Finally, it should be noted that the library preparation approach described here, barring the initial enrichment step, is not restricted to influenza virus but can be applied to any pathogen or organism with an RNA genome. In addition, the protocol has been used without modification and with success on mRNA samples.

References

1. Belshe RB, Smith MH, Hall CB, Betts R, Hay AJ. Genetic basis of resistance to rimantadine emerging during treatment of influenza virus infection. 1988. J. Virol.62:1508-1512.

2. Bright *et al.,* Lancet 2005; 366:1175-1181

3. Fauci AS. 2005. Race against time. Nature 435: 423-424.

4. Ghedin E., Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, Spiro DJ, Sitz J, Koo H, Bolotov P, Dernovoy D, Tatusova T, Bao Y, St George K, Taylor J, Lipman DJ, Fraser CM, Taubenberger JK, Salzberg SL. 2005. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. Nature 437:1162-1166.

5. Habib-Bein NF, Beckwith, WH, Mayo D, Landry ML. 2003. Comparison of SmartCycler real-time reverse transcription-PCR assay in a public health laboratory with direct immnofluorescence and cell culture assays in a medical center for detection of influenza A virus. J. Clin. Microbiol. 41: 3597-3601.

6. Holmes CE, Ghedin E, Miller N, Taylor J, Bao Y, St. George K, Grenfell B., Salzberg SL, Fraser CM, Lipman .J, Taubenberger JK. Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. 2005. PLOS Biology 3(9):1570-1589.

7. Huang, X. 1991. Contig Assembly Program (CAP). http://www.cs.sunysb.edu/~algorith/implement/cap/

8. Poddar SK, Espina R, Schnurr DP. Evaluation of a single-step multiplex RT-PCR for influenza virus type and subtype detection in respiratory samples. 2002. J. Clin. Lab. Anal. 16:163-166.

9. Pourmand N, Diamond L, Garten R, Erickson JP, Kumm J, Donis RO, Davis RW. 2006. Rapid and Highly Informative Diagnostic Assay for H5N1 Influenza Viruses. PLoS ONE 1 (1).

10. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. Mapping the antigenic and genetic evolution of influenza virus. 2004. Science 305:371-6.

11. Spackman *et al.*, J. Clin. Microbiol. 2002; 40:3256-3260.

12. Tumpey TM, Maines TR, Van Hoeven N, Glaser L, Solórzano A, Pappas C, Cox NJ, Swayne DE, Palese P, Katz JM García-Sastre A. 2007. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. Science 315:655-659.

Notes:

For additional references and detailed information, visit
www.genome-sequencing.com

454 LIFE SCIENCES