ion torrent
Sequencing for all.™

INFLUENZA A WHOLE-GENOME SEQUENCING

Benefits of the PathAmp™ FluA Reagents workflow

- Streamlined workflow—complete influenza A genome sequencing, typically in about a day; much faster than traditional Sanger sequencing methods
- Multiplexing capability—simultaneously sequence up to 10 samples on a single Ion 314™ chip for lower cost per sample
- More accurate sequence data—Ion PGM™ next-generation sequencer offers higher sequencing coverage for more accurate sequence data

Influenza A is responsible for seasonal epidemics that cause approximately three to five million cases of severe illness and 250,000 to 300,000 deaths, according to the World Health Organization (WHO). Reassortant influenza A strains can arise when influenza virus variants are coincident in an animal population—for example in avian and swine reservoirs—and genes from one variant become incorporated in the other. Zoonotic transmission of avian or swine influenza directly to humans as well as transmission of reassortant virus has caused notable human pandemics over the last decade.

Because of the genetic instability in this pathogen, continual surveillance of circulating influenza A subtypes by public health institutions is critical for outbreak prevention, outbreak monitoring, and annual vaccine production. Genetic surveillance of influenza A virus has historically been carried out by Sanger sequencing of multiple overlapping fragments to build a complete sequence for all 8 viral genome segments, a time-consuming and relatively expensive process.

In this publication, we outline a next-generation sequencing strategy for influenza A sequencing using the Ion PGM™ System and the PathAmp™ FluA Reagents, which offers deep sequence coverage, multiplexed sample capability, and a faster, more streamlined workflow (Figure 1). Using this methodology, researchers may in the future collect high-quality sequence data more quickly and cost-effectively than with their current protocols—information that is key for public health labs and vaccine production facilities that are responsible for monitoring the emergence of seasonal flu subtypes and creating therapeutics to combat them.

We report the results of both internal performance testing of the PathAmp™ FluA Reagents as well as results generated by a global network of researchers from leading government and public health and veterinary agencies and research institutes. Members of this network have been sharing tools, experience, and data using the Ion PGM™ System for influenza A whole-genome sequencing.

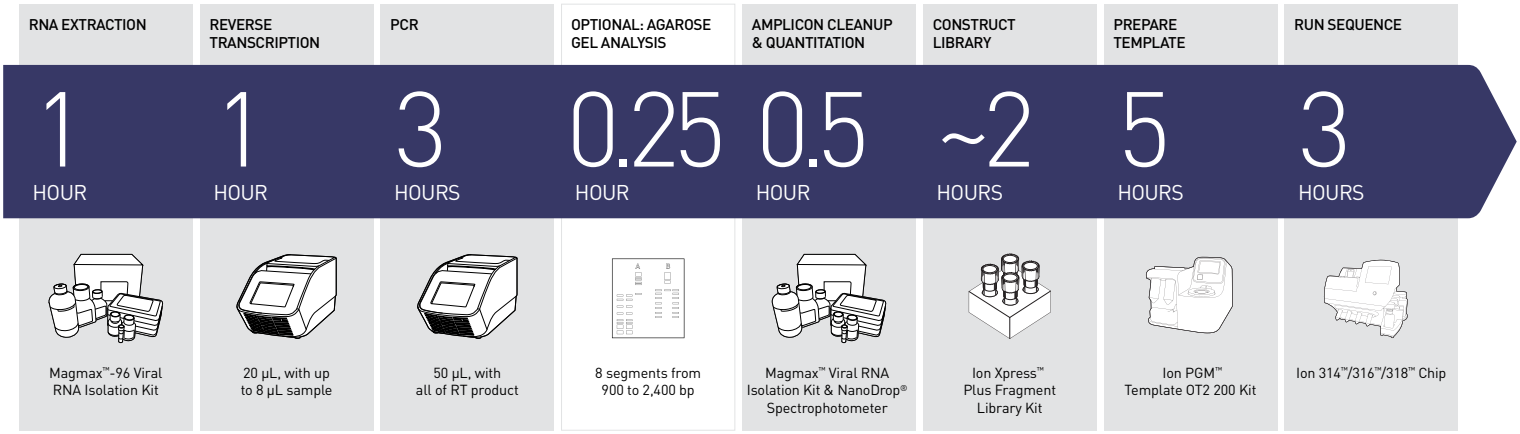

Figure 1. The PathAmp™ FluA workflow. RNA extraction was performed using the MagMAX™-96 Viral RNA Isolation Kit (Cat. No. AM1836) without carrier RNA or Xeno™ RNA (internal positive control) added. Reverse transcription and PCR amplification was performed using the PathAmpTM FluA Reagents (Cat. No. 4485019). Amplicons were purified and quantitated. The library was prepared with Ion Xpress™ Plus Fragment Library Kit (Cat. No. 4471269). Enzymatic fragmentation was used (15 minute incubation) for the 200 bp read protocol. Samples were barcoded with Ion Xpress™ Barcode Adapters 1-16 Kit (Cat. No. 4471250). Template generation was performed with the automated template preparation using the Ion OneTouch™ 2 System (Ion PGM™ Template OT2 200 Kit–Cat. No. 4480974). Sequencing was performed with the Ion PGM™ Sequencing 200 Kit v2 (Cat. No. 4482006) using either an Ion 314™ Chip Kit v2 (Cat. No. 4482261), Ion 316™ Chip Kit v2 (Cat. No. 4483188), or Ion 318™ Chip Kit v2 (Cat. No. 4484354).

Multiplexed sequencing of all 8 genomic segments

The influenza viral genome is ~13.5 kb and comprises eight segments that encode up to 11 proteins (Figure 2A). The terminal segment-specific nucleotide regions are unique to each segment and are conserved across influenza A viruses, which makes it possible for researchers to design primers for reverse transcription and PCR amplification (Figure 2B) and then to sequence the entire viral genome [1,2].

The PathAmp™ FluA Reagents are a set of highly specific, universal influenza primers along with a high-fidelity master mix for the amplification of all eight influenza A genomic segments in a single tube. The reaction produces cDNA fragments that range in size from 900 bp to 2.4 kb, which can then be used as an input into Ion PGM™ library preparation workflow.

Researchers have already published data demonstrating the power of the Ion PGM™ System for influenza surveillance. In a study of pigs at an Ohio agricultural fair, Bowman et al. [3], in conjunction with National Veterinary Services Laboratories (NVSL), used primers to amplify all eight segments of the influenza A genome and then sequence the products on the Ion PGM™ instrument using an Ion 314™ chip. Follow-up molecular phylogenetic analysis of the human and swine viruses indicated a common source of dissemination. In addition, whole-genome sequencing of influenza A virus on the Ion PGM™ System can be used for more effective vaccine development. To this end, Hause et al. identified a mutation that reversed antibody-mediated neutralization of swine influenza virus [4].

Table 1. Average coverage of all eight influenza A segments using the Ion 314™ and Ion 316™ chips, respectively.
On each chip, 10 samples (H1N1, H1 2009 pandemic, and H3) were run simultaneously. (Data provided by S. Glavas.)

RNA Segment	Length (bp)	Average reads	Average coverage (fold)
Ion 314™ chip			
PB2	2,310	1,776	123
PB1	2,274	3,259	237
PA	2,167	669	52
HA	1,701	1,543	149
NP	1,537	2,501	256
NA	1,410	2,380	275
M	982	4,151	710
NS	844	1261	201
Ion 316™ chip			
PB2	2,310	25,398	1,865
PB1	2,274	15,592	1,176
PA	2,167	24,073	1,893
HA	1,701	20,709	2,130
NP	1,537	24,138	2,634
NA	1,410	25,747	3,179
M	982	30,583	5,260
NS	844	9,989	1,650

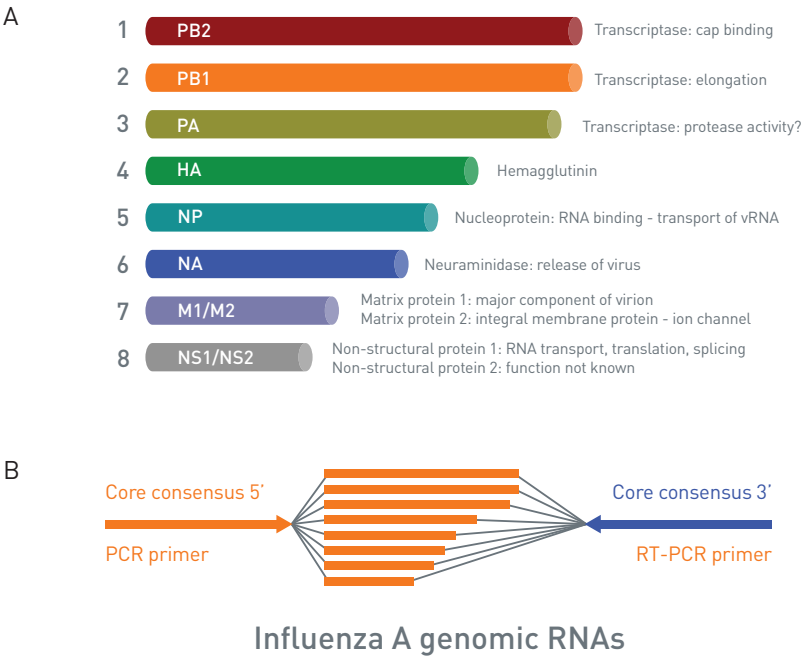

Figure 2. Influenza A genome and sequencing primer design strategy. (A) Influenza A genome with 8 RNA fragments. The hemagglutin protein (HA) and neuraminidase enzyme (NA) that are present on the surface of the viral envelope are used to categorize influenza A into subtypes. In addition to HA and NA segments, the other six segments encode: the polymerase basic protein 1 (PB1) and PB1-F2 protein; polymerase basic protein 2 (PB2); polymerase acidic protein (PA); nucleoprotein (NP), matrix proteins (M) M1 and M2, and non-structural proteins (NS) NS1 and NEP. (B) Primer design overview for the amplification of all eight influenza A genomic segments. Conserved regions on the end of each segment of the virus are targeted with primers that allow for reverse transcription and PCR amplification, prior to DNA library preparation, of influenza A viruses from human and non-human hosts.

Deep sequencing coverage from the PathAmp™ FluA Reagents workflow

Steve Glavas, head of the NGS Platform at the Swedish Institute for Communicable Disease Control (SMI), tested the performance of the PathAmp™ FluA Reagents on the Ion PGM™ System using both Ion 314™ and Ion 316™ chips, running 10 samples simultaneously on each chip (20 isolates from H1N1, H1 2009 pandemic, and H3 cases). Results were 100% in concordance with data generated by complete Sanger-based sequencing of HA, NA, and M fragments as well as sections of the PB1, PB2, PA, NS1/NS2, and NP genes.

Average coverage of all eight segments ranged from 52x (PA) to 712x (M) on an Ion 314™ chip and from 1865x (PB2) to 5260x (M) on an Ion 316™ chip when running 10 samples simultaneously (Figure 3 and Table 1). Note that coverage depth varies inversely with segment length (shorter segments tending to have more coverage). The longer segments such as PB1 (~2.3 kb), PB2 (~2.3 kb) and PA (~2.2 kb) are more deeply covered at the ends, whereas the shorter segments such as NS (~0.9 kb) and MP (~1.0 kb) are more deeply covered in the center (Figure 3).

Glavas and co-workers at SMI established a custom data analysis pipeline for the fast and automated analysis of influenza A whole-genome data. Reads from Torrent Suite™ Software are mapped against reference sequences and the average coverage depth for each segment and on a per-base basis is reported. The consensus sequence, which is available in FASTA format, can be used for further analysis. The plug-in for this is called "PathogenAnalyzer" and is currently available in the Ion Community Torrent Browser Plug-in Store.

Partnership develops faster, more reliable flu subtyping methods

Each year, public health agencies around the world collect samples from infected individuals and share data about flu subtypes circulating in their regions. The pooled data are used by the WHO to determine the strains used to design a vaccine that will be effective against that year's epidemic. Costs of sequencing, however, have limited the data set to about 20% of the patient samples collected.

In an effort to increase the speed and efficiency of influenza monitoring and vaccine development, Life Technologies recently announced the establishment of the Global Influenza Network. This partnership includes Steve Glavas, head of the NGS Platform, Mattias Mild, head of the Genomics Platform, and Mia Brytting, PhD, head of the Microbial Typing Unit at the Swedish Institute for Communicable Disease Control (SMI); Gabriele Vaccari, PhD, researcher at the Istituto Superiore di Sanità in Rome; Mary Lea Killian, microbiologist at the U.S. National Veterinary Service Laboratories in Ames, IA; and David Wentworth at the J. Craig Venter Institute in Rockville, MD.

The current collaboration is a pilot program to evaluate influenza virus typing by semiconductor sequencing on the Life Technologies Ion Torrent™ platform: the Ion PGM™ sequencer. After implementing Life Technologies' protocol for whole-genome sequencing of Influenza A virus, the network partners determined it to be: (1) accurate, (2) highly sensitive, and (3) economical (less than $100 per isolate) because it enables scientists to multiplex at least 10 samples in a single run on the Ion PGM™ sequencer. Global Influenza Network partners have agreed to share their data and experiences in order to refine the initial protocol, if needed. The results of the collaborative study will be submitted to a peer-reviewed research journal for publication.

"Using next-generation sequencing technology makes whole influenza genome sequencing much easier, and much less expensive than older sequencing techniques, when used appropriately."

– **Steve Glavas**
Head of NGS Platform, SMI, Sweden

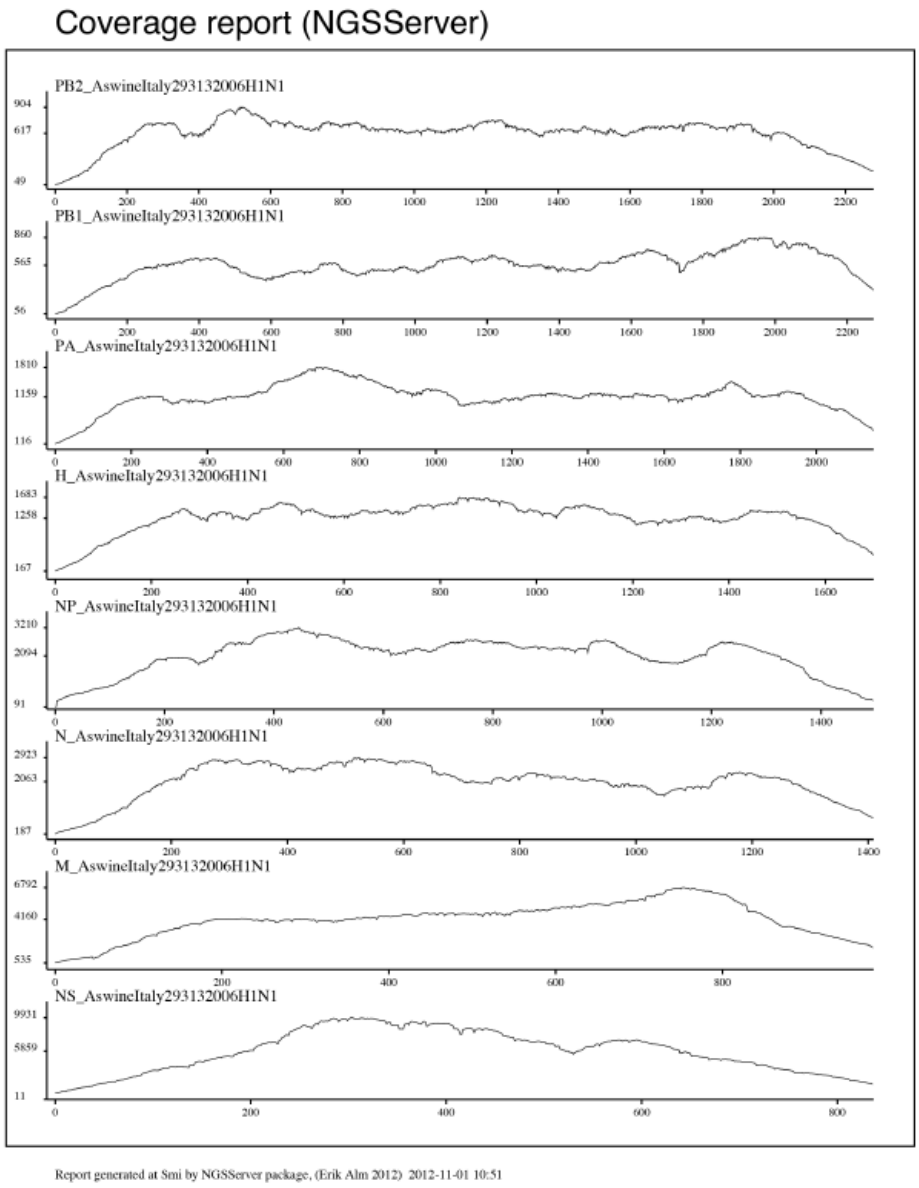

Figure 3. Coverage report for one sample (H1N1) that was run in 10-plex on a Ion 314™ Chip using the PathAmp™ Flu A Reagents. There is sufficient coverage for all segments and the coverage within each segment is very even, without any gaps. [Data provided by G. Vaccari.]

Performance characteristics of PathAmp™ FluA Reagents workflow

Internally, Life Technologies tested the performance characteristics of the PathAmp™ FluA Reagents on a diverse panel of 15 influenza viral strains, seven avian and eight swine. The PathAmp™ FluA results were compared to data obtained using a collaborator's lab–developed influenza A preamplification workflow.

The PathAmp™ FluA workflow accuracy was verified internally and externally on the diverse panel of 15 influenza subtypes, with subtyping calls consistent, except for one discrepancy discussed below, with a published lab–developed influenza A preamplification workflow. Further, the subtyping lineage calls were 100% concordant with real-time PCR subtypes and Sanger sequencing results.

A single segment mapped to different lineages based on preamplification reagents used. The NP segment of avian influenza A virus subtype H7N9 was determined to be E lineage using the collaborator's preamplification primers, whereas this segment was called as H lineage using the PathAmp™ kit. However, the NP segment allele had 100% alignment of reads with the PathAmp™ workflow, while with the collaborator preamplification method, it had 98%. Further, the alignment of sequencing reads to lineage E contained a ~200 bp gap in the coverage, possibly indicative of an incorrect call. Sanger sequencing confirmation identified the correct allele as the H lineage, as predicted by the PathAmp™ kit.

PathAmp™ FluA Reagents have high specificity for influenza A over, host genome, with detection down to 200 viral copies

When sequencing viral genomes, background host nucleic acids may be co-processed during library preparation, resulting in a sequencing reaction in which a majority of reads arise from the host genome. One solution is to run a preamplification reverse transcription (RT)-PCR reaction on the extracted nucleic acid with primers designed to amplify only the viral genome prior to library preparation.

Using the PathAmp™ FluA Reagents, the percentage of sequencing reads aligning to influenza A reference was improved over a collaborator preamplification workflow (Table 2). Additionally, a reduction in the percentage of reads aligning to a host reference was greatly improved using the PathAmp™ approach. Taken together, the results indicate that PathAmp™ FluA Reagents are highly specific to influenza A, with extremely low background mapping of reads to the host cell genome.

When the concentration of the influenza A virus is relatively high, amplification of the virus predominates over non-specific host amplification. However, in the presence of low viral concentrations, the nonspecific amplification of the host DNA can account for the vast majority of the library reads. This can be more pronounced when the host DNA is more concentrated, such as is the case for tonsil samples.

Table 2. The 15 influenza A subtypes tested comprising seven avian (AIV) and eight swine (SIV) strains.
All 15 cultured samples were barcoded and run in a single sequencing reaction.

Subtype	Sample	Collaborator preamplification workflow		PathAmp™ FluA Reagents	
		Percent reads mapping to influenza*	Percent reads mapping to reference*‡	Percent reads mapping to influenza*	Percent reads mapping to reference*‡
H1N2	SIV-1	48	45.90	98	0.25
H3N2	SIV-2	75	21.50	99	0.09
H3N2	SIV-3	43	51.50	98	0.23
H1N1	SIV-4	33	61.20	98	0.17
H3N2	SIV-5	43	50.70	98	0.22
H1N2	SIV-6	15	77.40	98	0.61
H1N1	SIV-7	47	46.60	97	0.19
H3N2	SIV-8	47	47.00	98	0.14
H7N9	AIV-1	86	9.60	98	0.31
H9N2	AIV-2	90	5	97	1.13
H5N2	AIV-3	93	0.20	97	0.04
H5N2	AIV-4	94	1.10	98	0.06
H5N9	AIV-5	95	1.20	98	0.06
H3N6	AIV-6	94	2.30	98	0.04
H8N4	AIV-7	95	0.60	99	0.14

*Percents do not sum to 100 due to reads not mapping to either reference.
‡SIV reads against dog genome; AIV against chicken.

To mimic field samples, a swine influenza A virus isolate was serially diluted in nucleic acid extracted from influenza A–negative porcine nasal swabs and tonsil tissue. The dilution range spanned 1x to 1:10M; quantification with qPCR using the VetMAX®-Gold SIV Detection kit and a eukaryotic 18S rRNA reference assay determined the approximate viral copy number to range from 10 to 40,000,000. The dilution samples were processed through the standard PathAmp™ preamplification protocol and DNA library preparation. Accurate genomic sequence and lineage calls were correct down to approximately 200 viral copies (Table 3) in both sample collection site backgrounds. Further, >99% of all swine influenza A segments were covered by reads down to 1:100,000x for the tonsil dilutions and 1:1,000,000x for the nasal swab dilutions.

The influence of host DNA and sample site collection on nonspecific background amplification effects on the PathAmp™ reagents was also examined by serial dilution experiments (Table 3). In the tonsil dilution samples, most of the reads aligned to influenza A or host sequences. At 1:10k (copy number 9,191) dilution, the percentage of reads aligning to influenza predominates. At a 1:100k (copy number 1,182) dilution, the percentage of influenza A reads drops precipitously to 7.12%, with non-specific amplification of the host DNA accounting for 78.27%. In the nasal tissue dilutions, the marked decline occurs at 1:100k (copy number 1,486), with 8.29% of reads aligning to influenza A and only 17.08% mapping to the host. It appears that when the influenza A concentration is low, most reads are coming from something else in nasal tissue. Possibly the non-specific amplification is coming from nucleic acid from other organisms such as bacteria in the nasal passages.

PathAmp™ FluA Reagents allowed detection of additional mutations

Collaborators at the WuXi Genome Center in China used the PathAmp™ reagents to identify genomic mutations in oseltamivir (TAMIFLU®)-resistant influenza A H1N1 strains. They sequenced 10 strains of oseltamivir-resistant influenza A H1N1, a single wild-type and 9 mutants, using the PathAmp™ FluA workflow and were able to identify multiple mutations (Table 4). A subset of samples was Sanger sequenced, confirming the mutations identified on the Ion PGM™ Sequencer.

Conclusions

PathAmp™ FluA Reagents provide a robust solution for influenza A virus sequencing. External and internal performance testing demonstrated high accuracy when results were compared to the collaborator's preamplification workflow. When using 10-sample multiplexing on Ion 316™ chip, collaborators observed sufficient coverage of all eight segments of the influenza A genomes with higher multiplexing possible.

The limit of detection of the PathAmp™ sequencing workflow was demonstrated to provide accurate lineage calls down to approximately 200 viral copies in the presence of high amounts of contaminating host nucleic acid. The workflow was shown to significantly decrease the amount of contaminating host nucleic acid reads to improve influenza coverage during the sequencing reaction.

In conclusion, the PathAmp™ FluA Reagents provide a rapid, accurate, and sensitive solution for influenza A viral sequencing on the Ion PGM™ Sequencer—and the resulting data can be utilized for future outbreak monitoring as well as routine surveillance.

Table 3. Limit of detection and the influence of host DNA background and sample collection site on the PathAmp™ reagents. A swine influenza virus sample was serially diluted in the presence of porcine nasal swab and tonsil tissue. Lineage calls for both sample matrices were correct down to 200 viral copies. More than 99% of all SIV segments were covered by reads down to 1:100,000x for the tonsil dilutions and 1:1,000,000x for the nasal swab dilutions. This corresponds to approximately 1,000 copies and 200 copies, respectively (determined by qRT-PCR). Average read depth for correct calls was >50x. N/A = not applicable; ND = not determined.

Sample	Dilution	Barcode	~ Copy number	SIV [Ct]	Lineage call	Total number of reads	Average read depth	Percent fraction covered	Percent match, influenza A	Percent match, host
SIV diluted in porcine tonsil tissue	1x	1	40,000,000	13.97	100% (8/8)	65,902	810	100	98.8	0.3
	1:100x	2	662,126	20.64	100% (8/8)	155,996	1803	100	97.6	1.6
	1:1,000x	3	78,440	23.99	100% (8/8)	180,118	1921	100	92.0	6.4
	1:10k	4	9,191	27.37	100% (8/8)	82,189	563	100	60.5	33.3
	1:100k	5	1,182	30.60	100% (8/8)	162,057	145	100	7.1	78.3
	1:1M	6	222	33.23	100% (8/8)	174,372	3	86	0.3	81.6
	1:10M	7	10	38.14	87.5% (7/8)	97,842	0	35	0.2	84.2
	Negative	8	N/A	ND	37.5% (3/8)	162,430	0	19	0.2	84.0
SIV diluted in porcine nasal swabs	1x	9	40,000,000	13.97	100% (8/8)	160,450	1964	100	99.1	0.2
	1:100x	10	791,955	20.35	100% (8/8)	223,193	2536	100	98.7	0.1
	1:1,000x	11	89,426	23.79	100% (8/8)	172,326	2030	100	98.6	0.1
	1:10k	12	10,083	27.22	100% (8/8)	181,147	2011	100	95.2	0.7
	1:100k	13	1,486	30.24	100% (8/8)	184,680	1702	100	76.4	2.9
	1:1M	14	201	33.39	100% (8/8)	68,766	63	100	8.3	17.1
	1:10M	15	9	38.29	87.5% (7/8)	159,746	25	77	1.8	12.6
	Negative	16	N/A	ND	75% (6/8)	84,073	3	32	0.3	14.5

Table 4. Mutation detection of drug-resistant influenza A subtype H1N1. A subset of mutations, highlighted in green, was confirmed by Sanger sequencing.

Ion PGM™ sequencing											
Segment	**Position**	**Wild-type**	**Mutant 1**	**Mutant 2**	**Mutant 3**	**Mutant 4**	**Mutant 5**	**Mutant 6**	**Mutant 7**	**Mutant 8**	**Mutant 9**
2	1,387	G	A	A	A	A	A	A	A	A	A
4	1,200	A	G	G	G	G	G	.	.	.	.
4	1,660	T	G	G	G	G	G	G	G	G	.
5	237	C	T	T	T	T	T	T	T	T	T
5	881	A	T	T	T	.	.	.	T	T	T
6	164	G	A	.	.	A	A	A	.	.	.
6	832	C	T	T	T	T	T	T	T	T	.
7	637	A	.	.	T	T	T	T	T	T	.
7	640	C	A	.	A	A	A	A	A	A	.
7	648	A	.	.	G	G/C	G/C	G/C	.	G/C	.
7	651	G	A	A	A	.	A	A	A	.	A
7	659	G	A	A	A	A	A	A	A	A	A
7	679	G	A	A	A	A	A	A	A	A	A
7	708	G	A	A	A	A	A	A	A	A	A
7	764	A	G	G	G	G	G	G	G	G	G
7	784	G	A	A	A	A	A	A	A	A	A
7	858	T	C	C	C	C	C	C	C	C	C
7	871	C	T	T	T	T	T	T	T	T	T
7	877	C	T	T	T	T	T	T	T	T	T
7	927	A	G	G	G	G	G	G	G	G	G
7	932	G	A	A	A	A	A	A	A	A	A
7	980	A	.	.	G	G	G	G	G	.	.
7	987	G	.	.	A	A	A	A	A	.	.
Sanger sequencing											
4	1,200	A	G		G	ND	ND	ND			
6	164	G	A		.	A	A	A			
6	832	C	T		T	T	T	T			

Acknowledgments

The authors greatly appreciate the opportunity to work with Steve Glavas, Mattias Mild, and Dr. Mia Brytting from the Swedish Institute for Communicable Disease Control (SMI) and Dr. Gabriele Vaccari, Dr. Livia Di Trani, Dr. Guendalina Zaccaria, and Dr. Adrianna Boni from the Department of Veterinary Public Health and Food Safety at the Istituto Superiore di Sanità in Rome.

We would also like to acknowledge Dr. Amy Vincent and Dr. Mary Lea Killian from the USDA for providing samples for us to test and to thank Dr. Lele Sun from the WuXi Genome Center for use of her sequencing data. This collaboration was led by Angela Burrell, Life Technologies, USA and Astrid Ferlinz, FALCON team, Life Technologies, Europe.

References

1. Hoffmann E, Stech J, Guan Y, et al. (2001). Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol 146:2275–2289.
2. Zhou B, Donnelly ME, Scholes DT, et al. (2009). Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and swine origin human influenza a viruses. J Virol 83:10309–10313.
3. Bowman AS, Sreevatsan S, Killian ML, et al. (2012). Molecular evidence for interspecies transmission of H3N2pM/H3N2v influenza A viruses at an Ohio agricultural fair, July 2012. Emerg Microbes Infect 1, e33; doi:10.1038/emi.2012.33.
4. Hause BM, Stine DL, Sheng, et al. (2012). Migration of the swine influenza virus δ-cluster hemagglutinin N-linked glycosylation site from N142 to N144 results in loss of antibody cross-reactivity. Clin Vaccine Immunol 19:1457–1464.