Development of Real Time RT-PCR Assays for Neuraminidase Subtyping of Avian Influenza Virus

Yanyan Huang[1], Mazhar Khan[1], and Ion Măndoiu

[1] Department of Pathobiology & Veterinary Science, University of Connecticut
61 North Eagleville Road, Storrs, CT 06269-3089, USA {mazhar.khan}@uconn.edu
[2] Computer Science & Engineering Department, University of Connecticut
371 Fairfield Way, Storrs, CT 06269, USA
ion@engr.uconn.edu

1 Introduction

Influenza A viruses (IAVs) are eight-segmented, negative-sense RNA viruses belonging to the family Orthomyxoviridae. IAVs are subtyped according to the difference of the two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). To date, there are total 16 HA and 9 NA subtypes of IAVs, and all of the subtypes have been detected from Avian Influenza Viruses (AIVs). In recent years AIVs are arising more and more human concerns because of the increasing cases of human infection of AIVs directly from birds. Epidemiological surveillance of AIVs in different avian species are reinforced to track the evolution of the viruses and to prepare ourselves for a possible Avian Flu Pandemic. Thus, sensitive, accurate and time-saving assays for AIV diagnosis and subtying, are essential for prompt response to epidemic and routine surveillance of AIV evolution.

The Polymerase Chain Reaction (PCR) has become the method of choice for virus subtype identification, largely replacing traditional immunological assays due to its high sensitivity and specificity, fast response time, and affordable cost [6]. However, designing subtype specific PCR primer pairs is a very challenging task [4]: on one hand, selected primer pairs must result in robust amplification in the presence of a significant degree of sequence heterogeneity within subtypes, on the other, they must discriminate between the subtype of interest and closely related subtypes. In recent work [2], we have developed an open source software tool, called PrimerHunter, that can be used to select highly sensitive and specific primers for virus subtyping. PrimerHunter ensures the desired amplification properties by using accurate estimates of melting temperature with mismatches, computed based on the nearest-neighbor model via an efficient fractional programming algorithm. In [2] we have confirmed the sensitivity and specificity of PrimerHunter's primers for hemagglutinin (HA) subtyping of Avian Influenza Virus (AIV) using individual real time PCR (R-PCR) reactions testing the presence of each subtype.

To increase sensitivity of detection when the quantity of viral RNA is limited, we propose using PCR reactions with pools of subtype-specific primer pairs

designed so that each subtype yields a unique amplification pattern. In this abstract we describe an integer program that can be used in conjunction with freely available solvers to design the minimum number of uniquely decodable pools subject to primer non-dimerization constraints. We applied this method to design four primer pools that can be used to differentiate the nine neuramidinase (NA) subtypes of AIV. Experiments using real time RT-PCR (RRT-PCR) show that the pool-based assay is sensitive and can detect in vitro transcribed RNA of different NA subtypes ranging from 176 to 4000 copies (2-30fg) per reaction.

2 Methods

2.1 Primer pair design

NA subtype specific primers were designed as described in [2]. A total of 588 complete Avian influenza NA sequences from North America were downloaded from the NCBI flu database [1]. PrimerHunter was run once for each subtype; when designing primers for subtype N_i we used all available NA sequences classified as N_i as targets, and all NA sequences labeled with different subtypes as non-targets. Between 7 and 9,665 pairs of primers were selected by PrimerHunter for each subtype (complete primer lists are publicly available at http://dna.engr.uconn.edu/software/PrimerHunter/). We selected for each NA subtype one pair of primers. Detection limits were determined using RRT-PCR on serially diluted RNA standards (see Table 1). The RRT-PCR assays could detect NA RNA ranging from 176 to 4000 copies per reaction, or total RNA concentration ranging from 2 to 30fg per reaction.

2.2 Primer pool design

To reduce the number of reactions needed to identify the subtype of a sample we will perform (RRT-PCR) reactions with pools of primer pairs. Clearly, the pools must be designed so that the result of these reactions uniquely identifies the subtype present in the sample. After the selection of a detection threshold, the result of each reaction can be viewed as being either positive or negative (in practice, additional information is provided by dissociation curves). Thus, pools must be designed so that each subtype results in a unique pattern of positive and negative signals, which can be viewed as the "barcode" for the subtype. As noted by [5], ensuring unambiguous identification is equivalent to ensuring that for every pair of subtypes there is a pool that results in a positive signal for one but not for the other. We additionally require that each subtype result in positive amplification signal from at least one pool.

There are several additional constraints that we must take into account when designing pools of primer pairs. First, pool size cannot be too large, since this results in decreased amplification efficiency due to the reduced primer concentration. We will denote by $2m$ the maximum number of primers allowed

Primer	Sequence	Prod. length (bp)	Prod. T_M ($^\circ C$)	Detection limit (RNA copies)
N1-fwd	5'-TAGACTGCATGAGGCCTTGCTTCTG-3'	137	78.2-79.2	2,080
N1-rev	5'-CACCGTCTGGCCAAGACCAACCTAC-3'			
N2-fwd	5'-ATGTTATCAATTTGCACTTGGGCAG-3'	149	77.1-77.8	1,440
N2-rev	5'-CATGCTATGCACACTTGTTTGGTTC-3'			
N3-fwd	5'-ATGATGTCTCTTGGACAAGCAATAG-3'	104	74.8-76.2	4,000
N3-rev	5'-TGGGCATAAACCCAATGTTGGAACC-3'			
N4-fwd	5'-AAATCATAACCATCGGTAGTGCGAG-3'	194	76.8-78.6	176
N4-rev	5'-TATAGTTGTTCTGCACATTGGTGAC-3'			
N5-fwd	5'-CATTTGTGGCATGTGGTCCCACGGA-3'	147	76.6-77.2	2,160
N5-rev	5'-AGGCATTGGGTGAAGATCCTAATGG-3'			
N6-fwd	5'-GCAAATAGACCAGTAATCACTAT-3'	153	77.9-78.9	2,080
N6-rev	5'-CCAGGATCTGGGTTTCCTCCTGTTA-3'			
N7-fwd	5'-AGCCAAGTATGTTTGGTGGACGAGC-3'	111	79.2-80.3	2,048
N7-rev	5'-TTACGAAAAGTATTGGATTTGTGCC-3'			
N8-fwd	5'-TAATGAGTGTAGAAATAGGGCAATC-3'	127	78.9-79.6	195
N8-rev	5'-GGAATCAGGGCCCGTTACTCCAA-3'			
N9-fwd	5'-ATCGTATTAAACACTGACTGGAGTG-3'	171	78.2-78.9	1,874
N9-rev	5'-ATTCTGTGCTGGAACACATTGATAC-3'			

Table 1. Sequences of NA specific primers. Expected PCR product lengths were calculated according to reference sequences of corresponding NA subtypes in the NCBI database. Detection limits were determined using RRT-PCR on serially diluted RNA standards of each NA subtype, with copy numbers ranging from 1 to 10^9.

in a primer pool. Second, primer-dimers must be avoided since even a single pair of dimerizing primers can lead to complete loss of amplification signal. Although PrimerHunter ensures that forward and reverse primers specific to a subtype do not dimerize, dimerization can occur between primers corresponding to different subtypes. We tested all primers included in Table 1 using the online autodimer tool at the National Institute of Standards and Technology (http://yellow.nist.gov:8444/dnaAnalysis/primerToolsPage.do). With a total score threshold of 4, autodimer reported 10 potentially dimerizing primer pairs. In general, we will denote the set of pairs of subtypes whose primers are predicted to form dimers by $\mathcal{D}$.

Subject to these constraints, we would like to minimize the number of pools, i.e., the number of PCR reactions. This can be modeled as an integer linear program similar to that proposed for the DNA barcoding problem by Rash and Gusfield [5]. Let N denote the number of subtypes, identified for simplicity by the integers from 1 to N ($N = 9$ in the case of NA subtyping). We first construct the set $\mathcal{P}$ of *candidate pools*, i.e., subsets of size at most m of $\{1, \ldots, N\}$ that do not contain any pair in $\mathcal{D}$. We then introduce a 0/1 variable x_p for every candidate pool $p \in \mathcal{P}$, where x_p is set to 1 if pool p is selected in our subtyping assay and to 0 otherwise. The pool design problem can then be written as follows:

Primer pool	N1	N2	N3	N4	N5	N6	N7	N8	N9
A (2,6,7)	-	+	-	-	-	+	+	-	-
B (4,5,7,8)	-	-	-	+	+	-	+	+	-
C (3,5,9)	-	-	+	-	+	-	-	-	+
D (1,4,6,9)	+	-	-	+	-	+	-	-	+

Table 2. Pools obtained by solving (2) using CPLEX. Notice that each column corresponds to a unique amplification pattern, allowing unambiguous identification of the subtype present in the sample.

$$\begin{aligned} \text{minimize} \quad & \sum_{p \in \mathcal{P}} x_p && (1) \\ \text{s.t.} \quad & \sum_{p \in \mathcal{P}: |\{i,j\} \cap p| = 1} x_p \geq 1, \ 1 \leq i < j \leq N \\ & \sum_{p \in \mathcal{P}: i \in p} x_p \geq 1, \ 1 \leq i \leq N \\ & x_p \in \{0,1\}, \ p \in \mathcal{P} \end{aligned}$$

To break ties in favor of solutions that use pools with fewer primers, objective (1) can be replaced by

$$\text{minimize} \quad \sum_{p \in \mathcal{P}} (M + |p|) x_p \qquad (2)$$

where M is a constant larger than $|\mathcal{P}|N$.

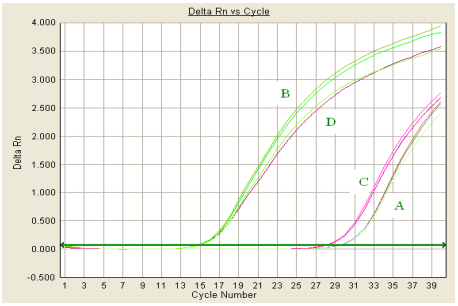
For NA subtyping this results in a formulation with 85 variables (pool candidates) and 45 constraints that was solved to optimality in a fraction of a second using CPLEX. The resulting assay consisting of four pools is shown along with expected amplification signatures in Table 2.

3 Results

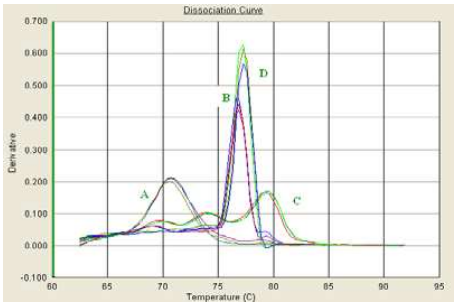
One step RRT-PCR with in vitro-transcribed RNA of each NA subtype showed linear amplification in a wide RNA concentration range (from 10 to 10^9 copies per reaction). Primer Pool tests of one step RRT-PCR demonstrated that RRT-PCR can detect and differentiate NA RNA of all nine subtypes extracted from AIV-infected allantoid fluids. Figure 1 shows the amplification and dissociation properties of N4 RNA in primer pool tests; full results will be presented in [3].

Acknowledgments

This work was supported in part by NSF awards IIS-0546457 and DBI-0543365.



(a)



(b)

Fig. 1. Amplification (a) and dissociation curves (b) for RRT-PCR reactions performed using the four primer pools on RNA extracted from H8N4 AIV-infected allantoid fluids.

References

1. Y. Bao, P. Bolotov, D. Dernovoy, B. Kiryutin, L. Zaslavsky, T. Tatusova, J. Ostell, and D. Lipman. The influenza virus resource at the national center for biotechnology information. *J. Virol.,* 82(2), pp. 596–601, 2008.
2. J. Duitama, D.M. Kumar, E. Hemphill, M. Kahn, I.I. Mandoiu, and C.E. Nelson. PrimerHunter: A Primer Design Tool for PCR-Based Virus Subtype Identification, *Nucleic Acids Research* 37, pp. 2483-2492, 2009.
3. Y. Huang, M. Kahn, and I.I. Mandoiu. Manuscript in preparation.
4. S. Gardner, T. Kuczmarski, E. Vitalis, and T. Slezak. Limitations of TaqMan PCR for detecting divergent viral pathogens illustrated by hepatitis A, B, C, and E viruses and Human Immunodeficiency Virus. *J. Clin. Microbiol.,* 41, pp. 2417–2427, 2003.
5. S. Rash and D. Gusfield. 'String barcoding: Uncovering optimal virus signatures', *Proc. 6th Annual International Conference on Computational Biology*, pp. 254–261, 2002.
6. D. Suarez, A. Das, and E. Ellis. Review of rapid molecular diagnostic tools for avian influenza. *Avian Diseases,* 51, pp. 201–208, 2007.