IMMUNOME RESEARCH

RESEARCH Open Access

Immunoinformatics analysis of H5N1 proteome for designing an epitope-derived vaccine and predicting the prevalence of pre-existing cellular-mediated immunity toward bird flu virus in Indonesian population

Marsia Gustiananda

Abstract

Background

The persistence of influenza A virus H5N1 among poultry in Indonesia, along with increasing numbers of human infections by this virus, point to risk of a reassortment event with other prevalent influenza A subtypes in Indonesia. In the absence of cross-protective antibody to the emerging strain, the presence of cellular immunity might reduce the influenza illness to subclinical level and could dampen the pandemic. This study evaluated the degree of likely cross-protective T-cell mediated immunity against such hypothetical emerging influenza A strains. All 11 protein sequences from Indonesian H5N1 isolates were evaluated for the presence of T-cell epitopes restricted by Indonesian HLA types.

Results

Among 4433 possible nonamer peptides, 225 were predicted as good CTL epitopes (score < 1%) by NetCTLpan and had strong binding affinity (IC50 < 50 nM) toward HLA molecules by IEDBann and netMHCann. Epitope conservation analysis revealed that at least 60% of the H5N1 Indonesian isolates contained the identical sequences of this 225 peptide set. Sixty-nine peptides were specific for H5N1 and 156 were cross-reactive with other subtypes. Significant numbers (184) of peptides bound to more than one HLA allele (promiscuous). Blast analysis showed that the majority (213) of peptides did not have similarity with the human self peptides likely to induce autoimmunity if used for vaccination. Certain peptides can act as a core for HLA Class II molecules, and some superimposed with the putative linear B-cell epitopes. Eighteen peptides emerged as likely vaccine components optimized to an Indonesian population but likely also to be effective among most other ethnic groups.

Conclusions

Routine exposure to seasonal influenza A likely ensures some level of cross-protective immunity to H5N1 and most reassortment mutants. Identifying of such likely cross-reactive T-cell epitopes may aid in gauging the threat of potential outbreaks and composing the vaccines that could contain them.

BACKGROUND

Influenza is preventable by vaccination. However, the continuous antigenic drift and shift threatens the efficacy of trivalent inactivated influenza vaccines. Pre-existing virus-specific T-cell memory resulting from prior natural exposure to human influenza A or vaccination may provide partial protection against novel strains of the virus. H5N1 influenza A remains a significant human and animal health problem in Indonesia today despite its relatively low transmissibility between humans. The severe morbidity and high risk of mortality associated with H5N1, along with concerns about possible reassortment of this strain in a manner that retains its virulence and ramps up its human-to-human transmissibility, emphasizes the importance of vaccination strategies. T-cells, in particular cytotoxic T-cells, directed against such mutants may be especially vital given their demonstrated importance in diminishing disease severity and promoting early clearance of virus [1]. These cells are activated by peptides derived from the internal protein of the virus and are presented by the major histocompatibility complex (MHC) of infected cells. These viral proteins tend to be relatively conserved among strains and subtypes and may thus represent ideal targets for vaccine formulations.

The major obstacle in the development of such broadly effective vaccines is the high polymorphism of MHC molecules (Human Leukocyte Antigen or HLA in humans). Each individual has their own unique set of HLA class I and II alleles. The type and frequency of alleles are very specific for each population. The HLA polymorphism that exists between ethnic groups will cause the individuals of different ethnicities to focus the response toward different T-cell epitopes. Therefore, probing the repertoire of T-cell specificities responses directed against influenza A viruses in the affected human population is needed.

The role of CTL in clearing up flu infection was first demonstrated in the study involving human volunteers inoculated with live virus of H1N1 subtype [1]. The study demonstrated influenza-specific CTL in the

Eijkman Institute for Molecular Biology, Jl. Diponegoro 69, Jakarta 10430, Indonesia
Email address: marsia@eijkman.go.id

absence of cross-protective antibody rapidly cleared infection and reduced disease to subclinical levels. The Cleveland family study showed that during the H2N2 pandemic of 1957, individuals that had experienced H1N1 influenza A were less likely to develop severe disease or succumb to infection with H2N2 [2].

The outbreak of avian influenza A virus H5N1 in Hongkong in 1997 and again in southeast-Asia and China in 2003 raised concern about the level of immunity that the human population might have toward this virus. A study by Jameson et al. [3] showed that people living in the U.S., who are highly unlikely to have encountered H5N1, have memory T lymphocytes that recognize epitopes on H5N1 viruses Hongkong isolates. Several other studies showed that the adult population living in H5N1-free countries, already had memory T-cells cross-react with H5N1 which probably stemming from previous exposure to seasonal human influenza either by vaccination [4] or infection [5, 6]. One study comparing Viet Nam and UK populations [7] found that both populations had H5N1 cross-reactive $CD8^+$ and $CD4^+$ memory T-cells presumably primed by seasonal influenza A infection.

The presence of cross-reactive T-cells against the pandemic swine origin influenza A virus (S-OIV) H1N1 2009 was also observed in the majority of populations evaluated [8, 9]. PBMCs archived the year before that pandemic showed T-cells recognizing epitopes from S-OIV H1N1 [10]. Immunoinformatics comparison of T-cell epitopes contained in S-OIV H1N1 with epitopes in 2008 - 2009 conventional influenza vaccine also revealed the high conservancy of $CD8^+$ and $CD4^+$ epitopes in HA and NA [11]. Again, the studies support the proposition that in the absence of cross-reactive antibody response to the outbreak strain, cross-reactive T-cells induced by vaccination or natural exposure may reduce disease duration and severity, as shown by the mild nature of the S-OIV H1N1 2009 pandemic.

The long recruitment time for cytotoxic T lymphocyte responses might hinder the effectiveness of the pre-existing cell-mediated immunity toward a highly virulence virus such as H5N1. However, there is evidence in human H5N1 disease that antiviral treatment-induced viral clearance seems to improve clinical outcome even when initiated late in the course of infection, suggesting there is a considerable time window during which enhanced viral clearance may still affect clinical outcome [12, 13]. The study in animal model revealed that pre-exposure to seasonal influenza A strain protected mice against lethal H5N1 challenge [14] and boosting the pre-existing cell-mediated immunity by vaccination will provide protection from H5N1 infection [15]. Thus, a possible clinical benefit of T-cell-targeted approaches cannot be excluded and the vaccine that induced and strengthened the pre-existing cell-mediated immunity deserves further exploration.

T-cell memory, in particular $CD8^+$ CTLs, recognizes peptides derived from virus peptides bound to HLA class I molecules on the exterior surface of infected cells. Recognition of this peptide-MHC complex prompts the CTLs that recognize it to kill the presenting cell. Mapping such epitopes in influenza A virus provides insight on probabilities of cross-protective T-cell responses to novel virus strains or subtypes, and may aid in development of more effective vaccine and diagnostics. Bioinformatics tools for epitope mapping (reviewed in [16]) in combination with the advancement of sensitive T-cell assays such as ELISpot [17], tetramer staining [18] and cytokine staining [19] and the concept of HLA supertypes [20, 21] paved the way to the development of 'immunoinformatics' for mining genome data for T-cell epitopes most likely to be useful in vaccine development [22].

Systematic bioinformatics approach has been used to analyze protein sequences to find conserved region containing T-cell epitopes that bind to HLA class I and II molecules [23]. This approach has been used to predict T-cell epitopes from the conserved protein sequences of both avian and human influenza A viruses and then matched the prediction results with the influenza epitope sequences curated in the Immune Epitope Database (www.immuneepitope.org) [24]. Others have performed in silico genome-, pathogen-, and HLA-wide searches followed by experimental in-vitro confirmation for new CTL epitopes against influenza A viruses, including H5N1 [25]. Some studies focus on identifying epitopes from the H5N1 proteome only and verified the predicted peptides binding to MHC alleles using molecular modeling [26]. Immunoinformatics has also been used to identify H5N1 epitopes from HA and NA proteins both from human and poultry isolates [27].

Most studies exploring T-cell epitopes by screening PBMC have employed predominantly European or North American ethnic groups. The study by Lee at al [7] is exceptional in this regard. Although in some studies [25] the prediction was made for T-cell epitopes restricted by HLA supertypes, the relevance to broader human population of the identified epitopes is uncertain, especially where the major HLA allele may not compose part of the HLA supertype.

H5N1 presents an ideal model for evaluation of potentially cross-protective T-cell epitopes, and to identify novel epitopes of this relatively very new human virus. Several studies mapped the antibody epitopes of H5N1 among the few survivors of this infection [28, 29, 30], but only one study explored T-cell epitopes [31]. The report identified CTL epitopes in H5 HA restricted by HLA A*02:01 and demonstrated immune responses in H5N1 survivors. That epitope (KLYQNPTTYI) derived from the poorly conserved HA protein of the virus. The epitope was not recognized by a healthy control having HLA A*02:01, thus demonstrating the unique character of the epitope. CTL epitopes may thus be derived not only from internal protein of the virus but also from surface proteins such as HA.

The present study applied an immunoinformatic analysis of the proteomes of H5N1 isolates from humans resident

in Indonesia in order to identify candidate T-cell epitopes (nonamer peptides) occurring on conserved internal virus proteins and likely to induce memory immune responses in an Indonesian population represented by Javanese and Sundanese-Javanese ethnic groups. The HLA allele types and frequencies of these two populations have been well characterized [32].

METHODS

Virus and proteins used in this study

The A/Indonesia/7261/2008 (H5N1) [GISAID: EPIISL23815] virus genome sequence housed in the GISAID database (http://platform.gisaid.org/) is the most recent Indonesian H5N1 isolate available in the public database. All 11 proteins encoded by the virus were used in the current analysis to identify T-cell epitopes: HA [GISAID: EPI162940], M1 [GISAID: EPI163111], M2 [GISAID: EPI163111], NA [GISAID: EPI163112], NEP [GISAID: EPI163115], NP [GISAID: EPI163113], NS1 [GISAID: EPI163115], PA [GISAID: EPI163116], PB1 [GISAID: EPI163117], PB1-F2 [GISAID: EPI163117], and PB2 [GISAID: EPI163118]. All other available H5N1 Indonesian human isolates were obtained from GISAID to be used in conservancy analysis. For conservancy analysis against other subtypes, protein sequences of the seasonal influenza A viruses H3N2, H1N1, and S-OIV H1N1 regardless of the origin of the isolates were extracted from the influenza virus resource at the NCBI database (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html).

Prediction of T-cell epitopes using immunoinformatic tools

The proteins listed above were analyzed for the cytotoxic T lymphocyte epitopes using several algorithms: netCTLpan, netMHCann, and IEDBann. All methods are based on artificial neural network (ann) algorithm. NetCTLpan (http://www.cbs.dtu.dk/services/NetCTLpan/) is a state of the art prediction tool which integrates predictions of proteasomal cleavage, transporter associated with antigen processing (TAP) transport efficiency, and MHC class I binding affinity into a MHC class I pathway for identification of the immunogenic epitopes. This tool, in particular, was chosen because it can predict peptide binding to HLA Class I molecules where experimental data is absent as long as HLA sequence is known. Therefore it can provide guidance for identification of epitopes likely to bind to HLA alleles in Javanese and Sundanese-Javanese population including those that has not been characterized experimentally. The nonamers peptides predicted by netCTLpan which score highest (< 1%) was selected. The other two algorithms used in the analysis were IEDBann (http://tools.immuneepitope.org/analyze/html/mhc_binding.html) and NetMHCann (http://www.cbs.dtu.dk/services/NetMHC/) which predicted peptides binding to 57 and 43 HLA class I alleles, respectively. The nonamers peptides with IC50 < 50 nM (strong binding) were selected. Herein, we applied a high stringency of choosing only the peptides that bound strongly to the HLA molecules, because the stronger the binding, the more likely the peptide to become T-cell epitopes. The peptides which showed binding by all the three analytical procedures (netCTLpan, IEDBann, netMHCann), were selected and subjected to further analysis.

Identification of predicted nonamers which were conserved across H5N1 Indonesian isolates. The sequences of H5N1 Indonesian isolates were downloaded from GISAID database (September 2010). The conservancy analysis was done using epitope analysis tools housed in the IEDB server (http://tools.immuneepitope.org/tools/conservancy/iedb_input). To avoid the bias, we eliminated the duplicate sequences from the analysis. The nonamer peptides where the identical sequences were found in more than 60% of the isolates were annotated and named Pepset1.

Identification of the predicted nonamers that were specific for H5N1 and those which were cross-reactive with other influenza A subtypes. The sequences of seasonal H3N2, H1N1, and S-OIV H1N1 were downloaded from NCBI (September 2010). Using the same epitope conservancy analysis tool as above, we analyzed whether the predicted nonamers sequences were also found in other influenza A subtypes. The nonamers where the identical sequences were found in more than 60% of any of the other subtype isolates were annotated and named Pepset2. The nonamers in which sequences were found only in H5N1 subtypes were annotated and named Pepset3.

Analysis for the probability that the predicted nonamers can act as a core peptide for HLA class II. The analysis was conducted using netMHCIIpan prediction tool (http://www.cbs.dtu.dk/services/NetMHCIIpan/) which can take nonamer peptides as the input and predict peptides binding to all HLA class II including those that are not well characterized. All HLA Class II alleles that are found in Javanese and Sundanese-Javanese population [32] were analyzed. The nonamer peptides predicted as core for HLA class II molecules (IC50 < 500 nM) were annotated.

Analysis for the sequence similarity with the human self peptides. The peptides having similarity with the self peptides can potentially induce autoimmunity if it is used for vaccine. Therefore nonamer peptides derived from H5N1 were blasted against the non redundant protein sequences of human [taxid: 9606] using NCBI Blastp suite program (http://blast.ncbi.nlm.nih.gov/Blast.cgi). NCBI Blast parameters were: word size 2; expectation value 30.000; matrix used PAM30; low complexity filter disabled; composition-based statistics was set to no adjustment. The Blast results were parsed using excel program. Nonamers which shared 7 out of 9 (7/9) identical amino acid sequences with the human peptides with no gap and no mismatches residue were triaged and not included in the vaccine construct.

Analysis for the presence of experimentally proven epitopes. Using epitope search tools in the Influenza Resources Database server (http://www.fludb.org), we

Proteins	Length	Possible number of nonamer peptides	Predicted epitopes	
			Number	Percentage
PB2	754	746	27	3.6%
PB1	757	749	62	8.3%
PB1-F2	90	82	1	1.2%
PA	716	708	40	5.6%
HA	568	560	23	4.1%
NP	498	490	27	5.5%
NA	444	436	12	2.8%
M1	252	244	13	5.3%
M2	96	88	5	5.7%
NS1	225	217	10	4.6%
NEP	121	113	5	4.4%
Total		4433	225	5.1%

Table 1 – The number of predicted peptides across H5N1 proteome.
The number of nonamer peptides identified as T-cell epitopes restricted by HLA alleles predominant in Javanese and Sundanese-Javanese population which was predicted by consensus method (netCTLpan score < 1%, IEDBann and netMHCann IC50 < 50 nM) and conserved in 60% of the H5N1 Indonesian isolates. The set of 225 peptides was named Pepset1 throughout the manuscript.

checked (September 2010) if the predicted nonamers had been confirmed experimentally either by T-cell assay or HLA binding assay. The nonamers that had been already reported in the database and those that were newly identified (in this study) were annotated.

Analysis for the presence of putative B-cell epitopes. Following Parida et al. [26], the B-cell epitope prediction was done using ABCpred server (http://www.imtech.res.in/raghava/abcpred/). The nonamer peptides which significantly superimposed (> 7 amino acid overlap) on putative B-cell epitope (decamer peptides) was annotated. According to the relay hypothesis [33], superimposition of CTL epitopes with B-cell epitope and T helper epitope would ensure good T-cell response with specific T-cell memory.

Selection of nonamers for vaccine candidates. The peptides which are selected for vaccine formulation should meet several criteria such as promiscuous binding to HLA Class I and Class II, conserve in H5N1 Indonesian isolates and other subtypes, do not induce autoimmunity, and the peptides are present at the putative linear B-cell epitopes of the proteins.

Calculation of the population coverage. The selected nonamers and their restricted HLA alleles were uploaded into the epitope analysis tools housed in IEDB server (http://tools.immuneepitope.org/tools/population/iedb_input). The server was used to calculate the population coverage of the selected peptide for Javanese and Sundanese-Javanese as well as for the other 11 major ethnic groups in the world.

RESULTS AND DISCUSSION

Prediction of T-cell epitopes by consensus method and analysis of epitopes conservancy across H5N1 Indonesian isolates

The A/Indonesia/7261/2008 (H5N1) virus genome sequence housed in the GISAID database (ID: EPIISL23815) was the most recent Indonesian H5N1 isolate available in the public database at the time the author started the analysis. In principle, we can choose any strain, but since we would like to asses T-cell responses to H5N1 in Indonesian population, then the locally circulating H5N1strain was used. This strain was not included in the earlier studies [24, 26, 27] of T-cell epitope prediction of influenza A viruses before. One study only included strain from Thailand [26], while another study [24] analyzed all influenza A viruses sequences published in the database as of September 2006. Some of the epitopes predicted in this study match the epitopes reported in earlier works [26, 27] and mapped to the conserved region described in one paper [24].

All 11 proteins of H5N1 were used as a starting dataset for prediction. Nine of the 11 proteins exist in the virion (HA, M1, M2, NA, NP, NEP, PA, PB1, and PB2) and 2 are expressed in the host-cells (PB1-F2 and NS1). All overlapping nonamers peptides were generated from this data set and were screened for potential T-cell antigens using netCTLpan, IEDBann, and netMHCann algorithms. IEDBann and netMHCann are among the best prediction server available online that have been evaluated [34], while NetCTLpan is a cutting edge methods to predicts not only peptide binding to all HLA class I alleles with known protein sequence [35] but also peptide processing inside the cells (proteasomal cleavage and TAP transport efficiency).

In order to cover prediction for all HLA alleles, we used NetCTLpan instead of using the server that predicts peptide binding to HLA supertype. NetCTLpan server can predict peptide binding to all HLA class I molecules including those that has not been characterized previously. One of the HLA allele that is important for Javanese and Sundanese-Javanese population is HLA A*24:07 (allele frequency of 22%). The peptide binding specificity of this allele has not been characterized. Futhermore, the allele is not classified into HLA A*24 supertype (represented by HLA-A*24:02). Both A*24:07 and A*24:02 have high homology with only one amino acid difference (residue no 70 is histidine in A*24:02 and

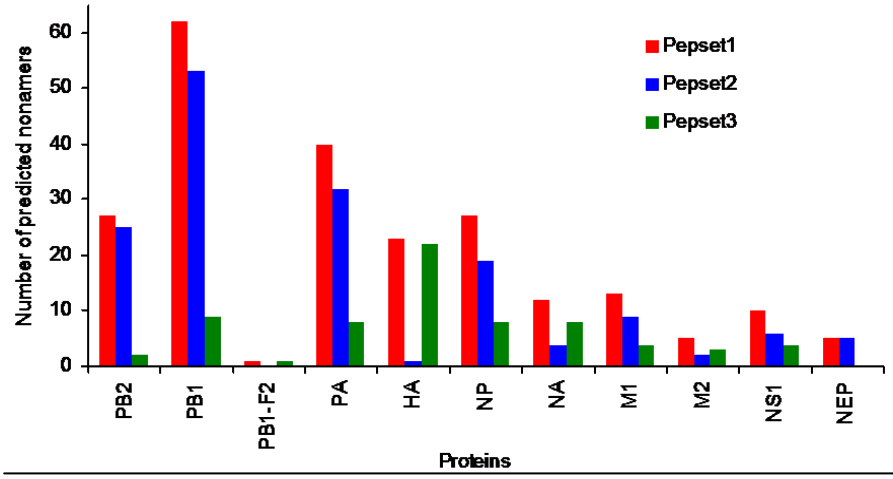

Figure 1 – The number of predicted peptides across the H5N1 proteome.

Proportion of the Pepset1 (predicted nonamers conserved in H5N1 Indonesian isolates, total 225 peptides), Pepset2 (predicted nonamers inducing T-cell cross-reactivity with other subtypes: H3N2, H1N1, and S-OIV H1N1, total 156 peptides), and Pepset3 (predicted nonamers specific for H5N1, total 69 peptides) across all 11 proteins of influenza A.

glutamine in A*24:07). However, this one amino acid is a contact residue which is part of the B pocket of the peptide binding sites. The difference could have some impact on binding specificity that makes A*24:07 differs from the rest of the member of A*24 supertype.

The consensus method in choosing the peptides was applied here, following the paper which stated that better prediction performance can be achieved by consensus approach (selecting peptides that are predicted as binders by several prediction tools) [36]. In total, there are 4433 possible nonamer peptides that can be generated from the proteome of A/Indonesia/7261/2008 (H5N1). Consensus prediction and conservancy analysis reduced this number to 225 peptides (5.1% of 4433) (Table 1). This set of 225 peptides were among the top highest score (< 1%) of the netCTLpan results and have high binding affinity (IC50 < 50 nM) toward HLA class I alleles in both IEDBann and netMHCann algorithms.

We set a high stringency in selecting the peptides (IC50 < 50 nM, and netCTLpan score < 1%) to ensure that the false positive is minimized. The peptides which have higher affinity for HLA molecules are expected to be more likely recognized by T-cells (the true T-cell epitopes). The peptides which score highest by netCTLpan are more likely to be generated in vivo, because this tool integrates the score for proteasomal cleavage, TAP transport efficiency, and HLA binding affinity. The high stringency in choosing the peptides has some advantage because false positive is minimized. However, the false negatives were increased.

We calculated the conservancy level of these 225 nonamer peptides. The conservancy level is defined as the number of the strains containing the epitope with exact sequences identity level divided by the total number of strains. Our analysis revealed that the identical sequences of each of these 225 peptides were found in more than 60% of the isolates. Out of 225 predicted peptides, 84 peptides have the sequences with 100% match identity level in all H5N1 Indonesian isolates. 19 peptides have 60-80% conservancy level and 206 peptides have more than 80% conservancy level. The significance of this information is that the epitopes predicted in this study could cover almost all strains of Indonesian H5N1 and might still present in the next pandemic H5N1 strain.

The number of predicted epitopes across H5N1 proteome

Our analysis revealed that all 11 proteins were capable of contributing nonamers for T-cell epitopes (Table 1). This is in agreement with other published papers [37] which experimentally showed that T-cells recognize the epitopes generated from all influenza proteins. PB1 contributes most of the predicted epitopes (60 nonamers), followed by PA, and then PB2 and NP. The proteins that contribute the least are PB1-F2, M2, and NEP, which are relatively small proteins. However, the number of the predicted peptides does not always correlate with the length of the proteins. Bigger protein does not always give larger number of predicted binding peptides such as in the case of PB2 (27 predicted nonamers out of 746 possibility). Also noteworthy is that nonamer peptides potential to become $CD8^+$ T-cell epitopes come not only from the internal protein, but also from the transmembrane surface protein HA, NA, and M2. One study reported however, that HA is not a major target for $CD8^+$ T-cell response [37].

Cross-reactive peptides and H5N1-specific peptides

We conducted an epitope analysis to elucidate whether there could be some level of cross-reactivity from $CD8^+$ T-cells between H5N1 and the other influenza A subtypes (seasonal H3N2 and/or seasonal H1N1 and/or S-OIV H1N1). The high level of cross reactivity between H5N1 and seasonal influenza A virus could influence the diseases outcome and provide partial protection from bird flu infection. We identified that out of 225 predicted peptides in Pepset1, 156 of them (named Pepset2) also found in other influenza A subtypes, while 69 of them (Pepset3) were specific for H5N1. As shown in Table 2, there are 76 predicted peptides that are conserved across all the influenza subtypes analyzed (H5N1, H3N2, H1N1, and S-OIV H1N1) with conservancy level at least 60%. The majority of the nonamers from the internal influenza proteins (PB1, PB2, PA, NP, M1, and NS1) were cross-reactive with other subtypes, while majority of those from the surface proteins (HA, NA, and M2) were H5N1-specific (Figure 1). Pepset3 peptides that were specific from H5N1 were generated from HA (22 nonamers), M1

The number of Pepset2 peptides that cross react	H1N1 (109)	H3N2 (108)	S-OIV H1N1 (135)	Note
76	+	+	+	76 peptides cross react with all 3 subtypes
5	+	+	-	5 peptides cross react with H1N1 and H3N2
16	+	-	+	16 peptides cross react with H1N1 and S-OIV H1N1
26	-	+	+	26 peptides cross react with H3N2 and S-OIV H1N1
10	+	-	-	10 peptides cross react with H1N1 only
1	-	+	-	1 peptide cross reacts with H3N2 only
18	-	-	+	18 peptides cross react with S-OIV H1N1 only

Table 2 – The number of predicted nonamers and cross-reactivity with other influenza A subtypes.

The number of peptides from Pepset2 that cross-reacted with H1N1 were 109, with H3N2 were 108, and with S-OIV H1N1 were 135. The presence of H5N1 peptides in other subtypes were indicated with '+' symbol and cells were highlighted in grey.

(4 nonamers), M2 (3 nonamers), NA (8 nonamers), NP (8 nonamers), NS1 (4 nonamers), PA (8 nonamers), PB1-F2 (1 nonamers), PB1 (9 nonamers), and PB2 (2 nonamers).

The fact that some H5N1 specific nonamers could be generated from the internal proteins reflect that the sequence variations among influenza A subtypes exists event in the region that is considered 'conserved' such as the internal proteins. These sequence variation and subtype specificity of the internal protein suggests that despite considered as 'conserved; the proteins undergo evolutionary change [38].

We identified that more epitopes were shared between H5N1 and S-OIV H1N1 than with the seasonal H3N2 or H1N1 (Table 2). In total, out of 156 epitopes in Pepset2, 135 were cross reactive with S-OIV H1N1, 109 with seasonal H1N1, and 108 with seasonal H3N2. 18 epitopes were cross-reactive with S-OIV H1N1 only, 1 with seasonal H3N2 only, and 10 with seasonal H1N1 only. The fact that S-OIV H1N1 shared larger number of epitopes with H5N1 is probably because one third of the S-OIV genome comes from avian virus [39, 40] and that S-OIV and H5N1 have high genetic compatibility [41]. It is noteworthy that in our analysis we found one nonamer (TYNAELLVL) from H5 HA which will induce T-cell cross reactivity with H1 HA and S-OIV HA but not with H3 HA. This epitope was found in the HA2 segment of HA protein which is known to be a stalk region [42] and relatively conserve across subtypes. As shown in Sup1, this nonamer was restricted by HLA A*24:07, HLA A*24:02, and HLA A*24:10. Two of these alleles (A*24:07 and A*24:02) are very common in Javanese and Sundanese-Javanese population. The study needs to be done to determine whether T-cells recognizing this epitope will be identified in the patients of Javanese and Sundanese-Javanese ethnic origin.

The conservancy of the epitopes across influenza A subtypes also reflects that the region of the protein in which these epitopes originated is of functional importance. It was proposed that the region between amino acids 506 and 659 of the PB1 protein is involved in the interaction with the PB2 subunit [43]. We found 9 nonamer peptides in this region, in which 8 of them were cross-reactive with other subtypes while one of them was H5N1 specific. Another region with functional importance was PB2 between amino acid 1 and 269 and between amino acid 580 and 683 which involved in the interaction with NP [44]. In this region we found 11 nonamer peptides in which all of them were conserved across subtypes.

Binding specificity of peptides to HLA class I alleles present in Javanese and Sundanese-Javanese population

Immunoinformatic analysis carried out in this study revealed that A*74:01 was predicted to bind the highest number of peptides in the Pepset1 (44 out of 225) (Figure 2). However, this allele was not relevant for the target population since it presents at a very low frequency in Sundanese-Javanese (0.25%) and absent in Javanese population. The four most frequent HLA A alleles in Javanese and Sundanese-Javanese population are HLA A*24:07 (21.52%), A*11:01 (16.03%), A*33:03 (15.61%), and A*24:02 (14.35%) which accounted for 67.8% of the total HLA-A alleles [32]. Immunoinformatic analysis carried out in this study revealed that A*24:07 bound 14 peptides, A*11:01 bound 31 peptides, A*33:03 bound 39 peptides, and A*24:02 bound 15 peptides (Figure 2). The two most common HLA B alleles are HLA B*15:02 (11.6%) and HLA B*15:13 (11.18%) which bound 15 and 18 peptides, respectively (Figure 2). It will be worthwhile to investigate whether these predicted nonamers are recognized by the bird flu survivors (some of them are Javanese and Sundanese-Javanese ethnic origin), and to see the antiviral activity of CTL recognizing these epitopes or to investigate its role in immunopathology of bird flu infection.

HLA A*02:01, which is the allele that is mostly predominant among Caucasian population and covers around 7% of the Javanese and Sundanese-Javanese population, bound 30 peptides. Majority of these peptides were promiscuous as they bound also to other allelic variants that belong to HLA A*02 supertypes, as well as to HLA B and HLA DRB1 alleles (Sup1). In Javanese and Sundanese-Javanese population, HLA A*02 is represented by HLA A*02:01 (6.96%), A*02:03 (4.22%), A*02:06 (3.38%), and A*02:11 (0.21%). In total, the frequency of the HLA A*02 supertypes is 14.77%, which implied that the immunodominant epitope M1 (58-66) GILGFVFTL which is restricted by HLA A*02:01 might be recognized by 15% of the Javanese and Sundanese-Javanese population, as the allele that belong to the same supertype will bind to the same nonamers. However, our analysis revealed that GILGFVFTL epitope bound to HLA A*02:01, A*02:06, A*02:11, and A*32:01 only and not

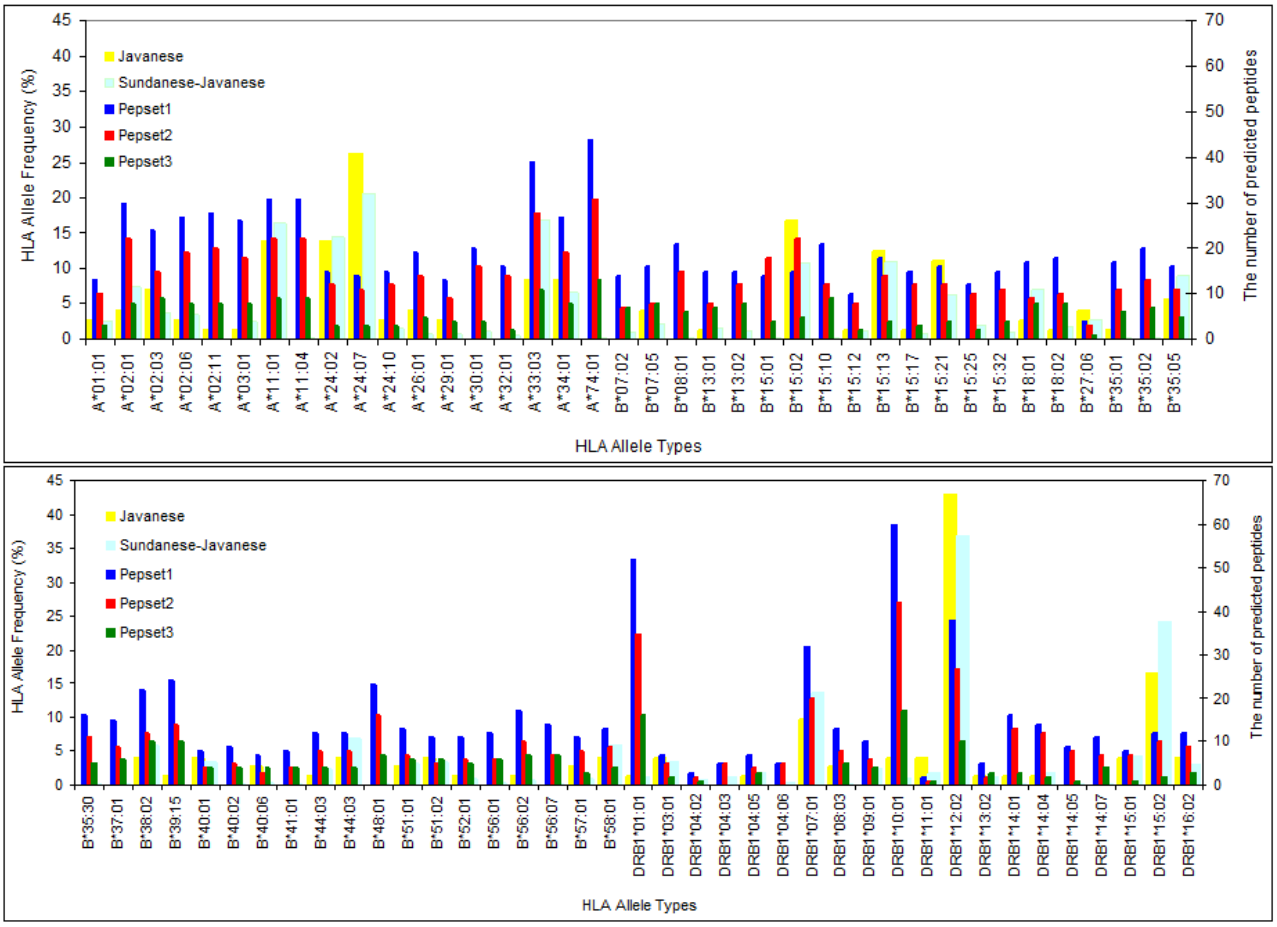

Figure 2 – The number of predicted peptides across HLA A and B alleles.
Proportion of the Pepset1 (predicted nonamers conserved in H5N1 Indonesian isolates, total 225 peptides), Pepset2 (predicted nonamers that will induce T-cell cross-reactivity with other subtypes: H3N2, H1N1, and S-OIV H1N1, total 156 peptides), and Pepset3 (predicted nonamers specific for H5N1, total 69 peptides) across all HLA Class I and II alleles found in Javanese and Sundanese-Javanese population.

to A*02:03 (Sup1). This was because in the analysis we selected only the nonamers which belong to the highest 1% of the netCTLpan score, while HLA A*02:03 - GILGFVFTL pair had a score that belong to the top 4% (data not shown). This indicated that the high stringency in choosing the peptides would cause higher false negative rates.

HLA B*27 supertype has been associated with strong efficiency in antiviral CTL responses to influenza A virus and several other viruses such as HIV, EBV, mump, and measles [45]. Relative to other class I molecules, HLA-B27 is a dominant restricting element in antiviral responses [45]. However, in Javanese and Sundanese-Javanese population, HLA B*27 is not predominant. The member of HLA B*27 is represented by HLA B*27:06 (2.95%) which predicted to bind 4 peptides, B*38:02 (5.7%) bind 22 peptides, B*15:10 (0.21%) bind 21 peptides, B*39:15 (0.21%) bind 24 peptides, and B*48:01 (0.21%) bind 23 peptides (Figure 2). The reported influenza T-cell epitopes restricted by HLA B*27 (SRYWAIRTR, NP 383-391) [46] was not included in the list of the predicted peptides, because of the high stringency in the selection criteria that had been applied in choosing the peptides, which was the highest 1% rank of the epitopes score. The peptide SRYWAIRTR was ranked as the 3% highest of the T-cell epitopes predicted by netCTLpan (data not shown). Again, this data implied that false negative might occur if one applied high stringency in the peptides selection step.

Binding of peptides to HLA Class II

Binding of the H5N1 derived peptides to all HLA class II alleles present in the Javanese and Sundanese-Javanese population had been carried out by using netMHCIIpan prediction tool. This tool is the best individual predictor for peptide binding to HLA Class II molecules [47] and can be used to predict peptide binding to all HLA Class II alleles including those where the experimental binding data is lacking. It was found that 75 out of 225 predicted nonamers capable of binding to the HLA Class II alleles (Sup1). This was in line with previous finding which suggested that nonamer peptides restricted by HLA class I was capable of becoming the core for HLA class II [48]. The peptides binding to HLA Class II molecules will be presented to the $CD4^+$ T-cell, which is important in regulating B-cell and CTL responses. The three most frequent alleles found in Javanese and Sundanese-Javanese populations are DRB1*12:02, DRB1*15:02, and DRB1*07:01 with frequencies of 37.8%, 23%, and 13.1%, respectively and bound to 38, 12, and 32 peptides, respectively (Figure 2). The two alleles that bound the largest number of peptides were HLA DRB1*01:01

and HLA DRB1*10:01 which bound 52 peptides and 60 peptides, respectively. These two alleles, however, present at a low frequency in Javanese and Sundanese-Javanese population (Figure 2).

Promiscuity of the predicted epitopes

Out of 225 predicted peptides, 41 peptides bind to only one HLA allele, while the rest (184) were promiscuous peptides. For the purpose of the vaccine development, it is important that the chosen peptides cover as many HLA alleles as possible to maximize the population coverage. The peptide capable of binding the greatest number of alleles was MQIRGFVYF, which could bind 29 HLA class I alleles (Sup1. The peptide capable of binding the greatest number of HLA class I and II alleles was LTITYSSSM (bound 26 alleles).

Cross-reactivity with the human peptides

The predicted peptides resulted in this study were compared with human proteome sequences to investigate the similarity to human self-antigens that could trigger an autoimmune response. The 225 nonamer peptides were blasted using the NCBI Blastp program (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against non redundant protein sequences of human (taxid: 9606). In particular we would like to know the degree of overlap between the predicted peptides and the human self peptides. Using the Microsoft excell program, we parsed the Blast results to identify peptides which have 100% identity with the peptides from the human proteome, with no gap or mismatches residues. All peptides from Pepset1 had similarity of 6 or less amino acid residues with human peptides. Following Tan et al. [49], we disregarded search results containing predicted sequences and human peptides with fewer than six contiguous identical residues (≤ 6/9 similarity) as the probability of matching five or less residues were high and non-significant. 8-mer peptides were the shortest that can bind to HLA molecules [50], therefore, the peptides which shared 8 contiguous amino acids with the human peptides were likely to induce autoimmune response. However, none of the predicted peptides in Pepset1 had 8/9 similarity (89%) or higher. The longest influenza A sequence with an identical human counterpart was 7 amino acid long (Table 3). They showed similarity to peptides of several human proteins such as tetraspanin 33, microtubule-associated protein 2, cystatin -11, dystrophin, myosin X, etc.

Selection of peptides for vaccine formulation and the calculation of the population coverage

The predicted epitopes were rationally chosen to be incorporated into a multi-epitope vaccine. Out of 225 predicted peptides, we selected 18 peptides to be included in the design of the influenza vaccine construct (Table 4). We selected the peptides which posses certain criteria such as: conserve in more than 60% of the H5N1 Indonesian isolates, cross-reactive with other subtypes, bound to HLA class I as well as class II, do not have similarity of more than 6 contiguous amino acids with the peptides derived from the human proteome, and superimpose with the putative linear B-cell epitopes. Here, we selected nonamer peptides which significantly superimposed (> 7 amino acid overlap) on putative linear B-cell epitope, because according to the relay hypothesis [33], superimposition of CTL epitopes with B-cell epitope and T-helper epitope would ensure good T-cell response with specific T-cell memory and will be beneficial for the formulation of the vaccine. The high conservancy of the peptides epitopes over time and across subtypes may reduce the incidence of the generation of escape mutant and is useful for the universal vaccine formulation.

Of note is that peptides restricted by HLA A*24:07 did not superimpose with the putative linear B-cell epitopes, so they were not selected. Other major HLA alleles for Javanese and Sundanese-Javanese population such as HLA A*24:02, A*11:01, A*33:03, B*15:02, B*15:13, DRB1*12:02, DRB1*07:01, and DRB1*15:02 bound to the peptides in the selected set (Table 4). It might be beneficial to include one or two peptides that were conserved and restricted by HLA A*24:07 even though the peptides did not fulfil the other selection criteria so as to increase the population coverage of the peptides set.

Population coverage analysis was done considering all the 18 peptides selected for the vaccine construct. The population coverage of the chosen peptides for Javanese and Sundanese-Javanese as well as for other 11 major ethnic groups in the world (Australia, Europe, North Africa, North America, North-East Asia, Oceania, Other, South America, South-East Asia, South-West Asia, and Sub-Saharan Africa) was calculated using the IEDB epitope analysis tools. The results (Table 5) showed that the set of 18 peptides to be included in the vaccine construct had good population coverage for HLA Class I (87% of the world population), but not so much for HLA Class II (62% of the world population). The coverage for Javanese and Sundanese-Javanese population are more than 99% for both Class I and Class II, since we based our prediction on the peptide binding to HLA alleles found in these populations.

CONCLUSIONS

Indonesia has had 176 (1 April 2011) confirmed human cases of avian influenza A virus (AIV) H5N1, and 145 (82%) of those ended in death. More than half of these cases occurred on the island of Java and thus most affected people of Javanese and Sundanese-Javanese ethnic origin. In this study, we did computational analysis for possible T-cell epitopes likely to be recognized by Javanese and Sundanese-Javanese ethnic groups using state-of-the art T-cell epitope prediction tools. This study is the first effort to identify influenza A virus-specific T-cells epitopes in these populations at high risk infection by H5N1 virus. The epitopes identified in this study were by consensus prediction, i.e., each epitope had to be predicted by at least two immunoinformatic methods. The epitopes thus identified were conserved among influenza A subtypes and this suggests functional significance of these regions. The likelihood or diversity of viable mutants may thus be considered relatively low compared to antibody epitopes among hyper-variable surface proteins.

Influenza peptides	Influenza protein	Human peptides	Human protein
GILGFVFTL	M1	GILGFVF	Tetraspanin 33
NEAVAVLKY	NA	EAVAVLK	Microtubule-associated protein 2
CTELKLSDY	NP	TELKLSD	CDA05
VASGYDFER	NP	VASGYDF	Immunolobulin heavy chain variable region
LQNSQVFSL	NP	QNSQVFS	Star-related lipid transfer protein 9
FMQALQLLL	NEP	QALQLLL	Cystatin-11
YYLEKANKI	PA	YLEKANK	Dystrophin
EITGTMRRL	PA	ITGTMRR	Family with sequence similarity 79, member A, isoform CRA_a
SLPPNFSSL	PA	PPNFSSL	Cadherin-like protein VR8
FLLMDALKL	PA	LLMDALK	Methyltransferase 10 domain containing
DPLASLLEM	PB2	DPLASLL	Contactin associated protein-like 5
RTSGSSVTK	PB2	TSGSSVT	Myosin X

Table 3 – Cross reactivity with the human proteome.
List of the predicted influenza A virus epitopes that shared 7 contiguous amino acid sequences with the human peptides. 7 is the highest number of amino acid residues shared between these influenza peptides and human proteins.

No.	Protein	Peptide Sequences	HLA alleles in which the peptide was predicted to bind
1	M1 (3 – 11)	LLTEVETYV	A*02:01; A*02:03; A*02:06; A*02:11; DRB1*01:01; DRB1*10:01
2	M1 (64 – 72)	FTLTVPSER	A*33:03; A*74:01; DRB1*10:01; DRB1*07:01; DRB1*12:02; DRB1*01:01; DRB1*14:04; DRB1*14:01; DRB1*04:03; DRB1*04:06
3	NP (39 – 47)	FYIQMCTEL	A*24:02; A*24:10; DRB1*10:01
4	NP (55 – 63)	RLIQNSITI	A*32:01; DRB1*13:02
5	NP (57 – 65)	IQNSITIER	A*74:01; DRB1*10:01; DRB1*12:02; DRB1*14:05; DRB1*14:04; DRB1*14:01; DRB1*03:01
6	NP (63 – 71)	IERMVLSAF	B*18:01; B*18:02; B*40:02; B*41:01; B*44:03; DRB1*10:01; DRB1*01:01; DRB1*15:02; DRB1*16:02
7	NS1 (123 – 131)	IILKANFSV	A*02:01; A*02:06; A*02:11; DRB1*15:01; DRB1*15:02; DRB1*01:01; DRB1*16:02; DRB1*12:02; DRB1*07:01; DRB1*10:01
8	PB1 (94 – 102)	FLEESHPGI	A*02:01; A*02:03; A*02:06; A*02:11; DRB1*01:01; DRB1*12:02
9	PB1 (166 – 174)	FLKDVMESM	A*02:01; A*02:03; A*02:06; A*02:11; A*26:01; A*34:01; B*15:01; B*15:02; B*15:21; B*15:32; DRB1*10:01
10	PB1 (340 – 348)	APIMFSNKM	B*07:02; B*07:05; B*35:01; B*35:02; B*35:05; B*35:30; B*56:02; DRB1*10:01
11	PB1 (355 – 363)	YMFESKSMK	A*03:01; A*11:01; A*11:04; A*33:03; A*34:01; A*74:01; DRB1*10:01
12	PB1 (407 – 415)	MMMGMFNML	A*02:01; A*02:03; A*02:06; A*02:11; B*08:01; B*13:02; B*15:10; B*35:02; B*37:01; B*38:02; B*39:15; B*48:01; B*56:01; B*56:02; B*56:07; DRB1*10:01; DRB1*01:01
13	PB1 (413 – 421)	NMLSTVLGV	A*02:01; A*02:06; A*02:11; DRB1*12:02; DRB1*14:01; DRB1*07:01
14	PB1 (441 – 449)	LQSSDDFAL	B*15:10; B*39:15; B*48:01; DRB1*01:01; DRB1*10:01
15	PB2 (49 – 57)	WMMAMKYPI	A*02:01; A*02:06; B*08:01; B*13:02; B*37:01; B*38:02; B*39:15; B*48:01; DRB1*12:02; DRB1*10:01; DRB1*09:01; DRB1*01:01; DRB1*08:03; DRB1*14:04; DRB1*07:01; DRB1*14:01; DRB1*14:07
16	PB2 (88 – 96)	RVMVSPLAV	A*30:01; DRB1*10:01; DRB1*09:01; DRB1*12:02; DRB1*01:01; DRB1*07:01
17	PB2 (90 – 98)	MVSPLAVTW	A*32:01; B*13:01; B*15:13; B*15:17; B*35:01; B*56:07; B*57:01; B*58:01; DRB1*01:01; DRB1*10:01
18	PB2 (201 – 209)	LMVAYMLER	A*74:01; DRB1*12:02; DRB1*03:01; DRB1*14:05; DRB1*14:01; DRB1*10:01; DRB1*14:04; DRB1*04:03; DRB1*04:06

Table 4 – List of candidate peptides for the development of multi-epitope vaccine construct.
Candidate peptides for vaccine should fulfil several criteria: promiscuous (bind to several HLA alleles), conserve across influenza A subtypes, do not share sequence similarity of more than 6 amino acid with the peptides from the human proteome, and occur in the putative linear B-cell epitopes.

Population / Area	Class I coverage	Class II coverage
Javanese	99.43%	99.69%
Sundanese-Javanese	99.70%	99.82%
Australia	99.96%	41.04%
Europe	96.68%	36.65%
North Africa	53.04%	71.40%
North America	94.47%	19.52%
North-East Asia	81.03%	74.74%
Oceania	98.78%	78.99%
Other	90.68%	57.22%
South America	58.62%	58.44%
South-East Asia	94.68%	51.64%
South-West Asia	86.14%	74.80%
Sub-Saharan Africa	76.32%	45.62%
Average (Standard deviation)	86.89% (15.07%)	62.28% (22.96%)

Table 5 – Population coverage calculation result.

The predicted population coverage is the fraction of individuals expected to respond to a given epitope set. In this case, the calculation was based on the set of 18 peptides that had been chosen for vaccine construct. The population coverage was divided for HLA Class I and Class II.

The search for T-cell epitopes using Influenza Research Database (http://www.fludb.org/) revealed no experimental data available for flu virus specific T-cell epitopes restricted by HLA A*24:07, A*11:01, A*33:03, and A*24:02, which are the four most frequent HLA A alleles in Javanese and Sundanese-Javanese population. Also there are no data for flu T-cell epitopes restricted by HLA B*15:02 and HLA B*15:13, which are the two most common HLA B alleles in these populations. In our immunoinformatic analysis we found a number of epitopes predicted to bind to these alleles. The lack of experimental data on T-cell epitopes restricted by HLA alleles predominant in Indonesian populations emphasizes the necessity of experimentally validating the finding of epitopes predicted to bind to these alleles. While some of the predicted epitopes in our analysis (mostly those binding to HLA A*02:01) were in agreement with the previous experimental studies (the data were curated in fludb.org database). New epitopes (cross-reactive as well as H5N1-specific) were also identified (Sup1).

T-cell immunologic assays involving samples obtained from H5N1 patients and survivors would serve to validate the immunoinformatic findings reported here. At the population level, the prevalence of cellular immunity toward H5N1 may aid in outbreak risk assessments. The prevalence of cellular immunity toward H5N1 has been done for mainly Caucasian populations, but this is the first such survey among a population living with endemic H5N1. The epitopes predicted in this study provide a rational starting point for evaluating cell-mediated immune responses against H5N1 in Indonesian patients.

COMPETING INTERESTS

The author declares there are no competing interests.

AUTHORS' CONTRIBUTIONS

MG designed the study, conducted the analysis, interpreted the data, and wrote the manuscripts.

ACKNOWLEDGEMENTS

MG is a recipient of the Australian Endeavour Award Postdoctoral Fellowship and Small Grant Award from the International Society for Infectious Diseases. MG would like to thank Frances Terry (Epivax Inc) for valuable comments on the manuscript and Kevin Baird (Eijkman Oxford Clinical Research Unit) for proofreading the manuscript.

Additional Files

www.immunome-research.net/journal/homeimmunomeojsfiles/journals/1/articles/22/supp/22-118-2-SP.pdf

REFERENCES

1. McMichael AJ, Gotch FM, Noble GR, Beare PAS: **Cytotoxic T-cell immunity to influenza.** *N. Engl. J. Med* 1983, **309**:13–17.
2. Epstein SL: **Prior H1N1 influenza infection and susceptibility of Cleveland Family Study participants during the H2N2 pandemic of 1957: an experiment of nature.** *J Infect Dis* 2006, **193**:49–53.
3. Jameson J, Cruz J, Terajima M, Ennis FA: **Human CD8+ and CD4+ T Lymphocyte Memory to Influenza A Viruses of Swine and Avian Species**. *The Journal of Immunology* 1999, **162**:7578–7583.
4. Gioia C, Castilletti C, Tempestilli M, Piacentini P, Bordi L, Chiappini R, Agrati C, Squarcione S, Ippolito G, Puro V, Capobianchi MR, Poccia F: **Cross-subtype immunity against avian influenza in persons recently vaccinated for influenza.** *Emerg Infect Dis* 2008, **14**:121-128.
5. Kreijtz JH, de Mutsert G, van Baalen CA, Fouchier RA, Osterhaus AD, Rimmelzwaan GF: **Cross-recognition of avian H5N1 influenza virus by human cytotoxic T-lymphocyte populations directed to human influenza A virus.** *J Virol* 2008, **82**:5161-5166.
6. Roti M, Yang J, Berger D, Huston L, James EA, Kwok WW: **Healthy Human Subjects Have CD4+ T-cells Directed Against H5N1 Influenza Virus**. *The Journal of Immunology* 2008, **180**:1758–1768.
7. Lee LY, Ha Do LA, Simmons C, de Jong MD, Chau NV, Schumacher R, Peng YC, McMichael AJ, Farrar JJ, Smith GL, Townsend AR, Askonas BA, Rowland-Jones S, Dong T: **Memory T-cells established by seasonal human influenza A infection cross-react with avian influenza A (H5N1) in healthy individuals**. *J Clin Invest* 2008, **118**:3478-3490.
8. Tu W, Mao H, Zheng J, Liu Y, Chiu SS, Qin G, Chan PL, Lam KT, Guan J, Zhang L, Guan Y, Yuen KY, Peiris JS, Lau YL: **Cytotoxic T lymphocytes established by seasonal human influenza cross-react against 2009 pandemic H1N1 influenza virus**. *J Virol* 2010, **84**:6527-

6535.
9. Ge X, Tan V, Bollyky PL, Standifer NE, James EA, Kwok WW: **Assessment of seasonal influenza A virus-specific CD4 T-cell responses to 2009 pandemic H1N1 swine-origin influenza A virus**. *J Virol* 2010, **84**:3312-3319.
10. Greenbaum JA, Kotturi MF, Kim Y, Oseroff C, Vaughan K, Salimi N, Vita R, Ponomarenko J, Scheuermann RH, Sette A, Peters B: **Pre-existing immunity against swine-origin H1N1 influenza viruses in the general human population**. *Proc Natl Acad Sci U S A* 2009, **106**:20365-20370.
11. De Groot AS, Ardito M, McClaine EM, Moise L, Martin WD: **Immunoinformatic comparison of T-cell epitopes contained in novel swine-origin influenza A (H1N1) virus with epitopes in 2008-2009 conventional influenza vaccine**. *Vaccine* 2009, **27**:5740-5747.
12. De Jong MD, Tran TT, Truong HK, Vo MH, Smith GJ, Nguyen VC, Bach VC, Phan TQ, Do QH, Guan Y, Peiris JS, Tran TH, Farrar J: **Oseltamivir resistance during treatment of influenza A (H5N1) infection**. *N Engl J Med.* 2005, **353**:2667-2672.
13. Kandun IN, Tresnaningsih E, Purba WH, Lee V, Samaan G, Harun S, Soni E, Septiawati C, Setiawati T, Sariwati E, Wandra T: **Factors associated with case fatality of human H5N1 virus infections in Indonesia: a case series**. *Lancet* 2008, **372**:744-749.
14. Kreijtz JH, Bodewes R, van den Brand JM, de Mutsert G, Baas C, van Amerongen G, Fouchier RA, Osterhaus AD, Rimmelzwaan GF: **Infection of mice with a human influenza A/H3N2 virus induces protective immunity against lethal infection with influenza A/H5N1 virus.** *Vaccine* 2009, **27**:4983-4989.
15. Epstein SL, Tumpey TM, Misplon JA, Lo CY, Cooper LA, Subbarao K, Renshaw M, Sambhara S, Katz JM: **DNA vaccine expressing conserved influenza virus proteins protective against H5N1 challenge infection in mice**. *Emerg Infect Dis* 2002, **8**:796–801.
16. Lafuente EM, Reche PA: **Prediction of MHC-peptide binding: a systematic and comprehensive overview**. *Curr Pharm Des* 2009, **15**:3209-3220.
17. Lalvani A, Brookes R, Wilkinson RJ, Malin AS, Pathan AA, Andersen P, Dockrell H, Pasvol G, Hill AV: **Human cytolytic and interferon gamma-secreting CD8+ T lymphocytes specific for Mycobacterium tuberculosis**. *Proc Natl Acad Sci USA* 1998, **95**: 270-275.
18. Altman JD, Moss PA, Goulder PJ, Barouch DH, McHeyzer-Williams MG, Bell JI, McMichael AJ, Davis MM: **Phenotypic analysis of antigen-specific T lymphocytes**. *Science* 1996, **27**:94-96.
19. Asemissen AM, Nagorsen D, Keilholz U, Letsch A, Schmittel A, Thiel E, Scheibenbogen C: **Flow cytometric determination of intracellular or secreted IFN gamma for the quantification of antigen reactive T-cells**. *J. Immunol. Meth* 2001, **251**:101-108.
20. Sette A, Sidney J: **HLA supertypes and supermotifs: a functional perspective on HLA polymorphism.** *Curr Opin Immunol* 1998, **10**:478-482.
21. Sidney J, Peters B, Frahm N, Brander C, Sette A: **HLA class I supertypes: a revised and updated classification.** *BMC Immunol* 2008, **9**:1.
22. De Groot AS, Sbai H, Aubin CA, McMurry J, Martin W: **Immunoinformatic: Mining genomes for vaccine components.** *Immunology and Cell Biology* 2002, **80**:255–269.
23. Khan AM, Miotto O, Heiny AT, Salmon J, Srinivasan KN, Nascimento E, Marques ET, Brusic V, Tan TW, and August JT. **A systematic bioinformatics approach for selection of epitope-based vaccine targets.** *Cell Immunol* 2006, **244**:141–147.
24. Heiny AT, Miotto O, Srinivasan KN, Khan AM, Zhang GL, Brusic V, Tan TW, August JT: **Evolutionarily conserved protein sequences of influenza A viruses, avian and human, as vaccine targets.** *Plos ONE* 2007, **2**:e1190.
25. Wang M, Lamberth K, Harndahl M, Røder G, Stryhn A, Larsen MV, Nielsen M, Lundegaard C, Tang ST, Dziegel MH, Rosenkvist J, Pedersen AE, Buus A, Claesson MH, Lund O: **CTL epitopes for influenza A including the H5N1 bird flu; genome-, pathogen-, and HLA-wide screening.** *Vaccine* 2007, **25**:2823-2831.
26. Parida R, Shaila MS, Mukherjee S, Chandra NR, Nayak R: **Computational analysis of proteome of H5N1 avian influenza virus to define T-cell epitopes with vaccine potential.** *Vaccine* 2007, **25**:7530-7539.
27. Somvanshi P, Singh V, Seth PK: **Prediction of epitopes in hemagglutinin and neuraminidase proteins of influenza A virus H5N1 strain: a clue for diagnostic and vaccine development**. *OMICS* 2008, **12**:61-69.
28. Khurana S, Suguitan AL Jr, Rivera Y, Simmons CP, Lanzavecchia A, Sallusto F, Manischewitz J, King LR, Subbarao K, Golding H: **Antigenic fingerprinting of H5N1 avian influenza using convalescent sera and monoclonal antibodies reveals potential vaccine and diagnostic targets.** *PLoS Med* 2009, **6**:e1000049.
29. Sun L, Lu X, Li C, Wang M, Liu Q, Li Z, Hu X, Li J, Liu F, Li Q, Belser JA, Hancock K, Shu Y, Katz JM, Liang M, Li D: **Generation, characterization and epitope mapping of two neutralizing and protective human recombinant antibodies against influenza A H5N1 viruses.** *PLoS One* 2009, **4**:e5476.
30. Kashyap AK, Steel J, Oner AF, Dillon MA, Swale RE, Wall KM, Perry KJ, Faynboym A, Ilhan M, Horowitz M, Horowitz L, Palese P, Bhatt RR, Lerner RA: **Combinatorial antibody libraries from survivors of the Turkish H5N1 avian influenza outbreak reveal virus neutralization strategies.** *Proc Natl Acad Sci USA* 2008, **105**:5986-5991.
31. Sun Y, Liu J, Yang M, Gao F, Zhou J, Kitamura Y, Gao B, Tien P, Shu Y, Iwamoto A, Chen Z, Gao GF: **Identification and structural definition of H5-specific CTL epitopes restricted by HLA-A*0201 derived from the H5N1 subtype of influenza A viruses.** *J Gen Virol* 2010, **91**:919-930.
32. Yuliwulandari R, Kashiwase K, Nakajima H, Uddin J, Susmiarsih TP, Sofro AS, Tokunaga K: **Polymorphisms of HLA genes in Western Javanese (Indonesia): close affinities to Southeast Asian populations.** *Tissue Antigens* 2009, **73**:46-53.
33. Lal G, Shaila MS, Nayak R: **Activated mouse T-cells downregulate, process and present their surface TCR to cognate anti-idiotypic CD4+ T-cells.** *Immunol Cell Biol* 2006, **84**:145-153.
34. Lin HH, Ray S, Tongchusak S, Reinherz EL, Brusic V: **Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research**. *BMC Immunol* 2008, **16**:8.
35. Stranzl T, Larsen MV, Lundegaard C, Nielsen M: **NetCTLpan: pan-specific MHC class I pathway epitope predictions.** *Immunogenetics* 2010, **62**:357-368.
36. Moutaftsi M, Peters B, Pasquetto V, Tscharke DC, Sidney J, Bui HH, Grey H, Sette A: **A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus.** *Nat Biotechnol* 2006, **24**:817-819.
37. Assarsson E, Bui HH, Sidney J, Zhang Q, Glenn J, Oseroff C, Mbawuike IN, Alexander J, Newman MJ, Grey H, Sette A: **Immunomic analysis of the repertoire of T-cell specificities for influenza A virus in humans.** J Virol 2008, **82**:12241-12251.
38. Hiromoto Y, Yamazaki Y, Fukushima T, Saito T, Lindstrom SE, Omoe K, Nerome R, Lim W, Sugita S, Nerome K: **Evolutionary characterization of the six internal genes of H5N1 human influenza A virus.** *J Gen Virol* 2000, **81**:1293-1303.
39. Garten RJ., Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, Okomo-Adhiambo M, Gubareva L, Barnes J, Smith CB, Emery SL, Hillman MJ, Rivailler P, Smagala J, de Graaf M, Burke DF, Fouchier RA, Pappas C, Alpuche-Aranda CM, López-Gatell H, Olivera H, López I, Myers CA, Faix D, Blair PJ, Yu C, Keene KM, Dotson PD Jr, Boxrud D, Sambol AR, Abid SH, St George K, Bannerman T, Moore AL, Stringer DJ, Blevins P, Demmler-Harrison GJ, Ginsberg M, Kriner P, Waterman S, Smole S, Guevara HF, Belongia EA, Clark PA, Beatrice ST, Donis R, Katz J, Finelli L, Bridges CB, Shaw M, Jernigan DB, Uyeki TM, Smith DJ, Klimov AI, Cox NJ: **Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans.** *Science* 2009, **325**:197-201.
40. Neumann G, Noda T, Kawaoka Y: **Emergence and pandemic potential of swine-origin H1N1 influenza virus**. *Nature* 2009, **459**:731-739.
41. Octaviani CP, Ozawa M, Yamada S, Goto H, Kawaoka Y: **High genetic compatibility between swine-origin H1N1 and highly pathogenic avian H5N1 influenza viruses.** *J Virol* 2010, **84**:10918-10922.
42. Wang TT, Palese P. **Universal epitopes of influenza virus hemagglutinins?.** *Nat Struct Mol Biol* 2009, **16**:233-234.
43. González S, Zürcher T, Ortín J: **Identification of two separate domains in the influenza virus PB1 protein involved in the interaction with the PB2 and PA subunits: a model for the viral RNA polymerase structure.** *Nucleic Acids Research* 1996, **24**: 4456–4463.
44. Poole E, Elton D, Medcalf L, Digard P: **Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites.** *Virology* 2004, **321**: 120–133.
45. Gomard E, Sitbon M, Toubert A, Begue B, Levy JP: **HLA-B27, a dominant restricting element in antiviral responses?**. *Immunogenetics* 1984, **20**:197–204.
46. López de Castro JA: **HLA-B27: portraying immunodominant viral epitopes.** *Eur J Immunol* 2005, **35**:336-340.
47. Lin HH, Zhang GL, Tongchusak S, Reinherz EL, Brusic V: **Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research.** *BMC Bioinformatic*. 2008, **9 Suppl 12**:S22.
48. Wang M, Larsen MV, Nielsen M, Harndahl M, Justesen S, Dziegiel MH, Buus S, Tang ST, Lund O, Claesson MH: **HLA class I binding 9mer peptides from influenza A virus induce CD4 T-cell responses.** *PLoS One* 2010, **5**:e10533.
49. Tan PT, Heiny AT, Miotto O, Salmon J, Marques ET, Lemonnier F, August JT: **Conservation and diversity of influenza A H1N1 HLA-restricted T cell epitope candidates for epitope-based vaccines.** *PLoS One* 2010, **5**:e8754.
50. Biddison WE, Martin R: **Peptide binding motifs for MHC class I and II molecules.** *Curr Protoc Immunol* 2001, **Appendix 1**: Appendix 1I.