DOI:10.1111/j.1750-2659.2010.00161.x
www.influenzajournal.com

Original Article

Highly conserved cross-reactive CD4+ T-cell HA-epitopes of seasonal and the 2009 pandemic influenza viruses

Venkata R. S. K. Duvvuri,[a,b,c] Seyed M. Moghadas,[d,e] Hongbin Guo,[f,g] Bhargavi Duvvuri,[h] Jane M. Heffernan,[a,f] David N. Fisman,[c,i] Gillian E. Wu,[b,h] Jianhong Wu[a,f]

[a]MITACS Centre for Disease Modeling, York Institute of Health Research, Toronto, Ontario, Canada. [b]Department of Biology, York University, Toronto, Ontario, Canada. [c]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada. [d]Institute for Biodiagnostics, National Research Council Canada. [e]Department of Mathematics and Statistics, University of Winnipeg, Winnipeg, Manitoba, Canada. [f]Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada. [g]Modelling & Projection Section, Surveillance and Risk Assessment Division, Centre for Communicable Diseases and Infection Control, Infectious Diseases and Emergency Preparedness Branch, Public Health Agency of Canada, Ottawa, Ontario, Canada. [h]School of Kinesiology and Health Sciences, York University, Toronto, Ontario, Canada. [i]Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada.
Correspondence: Seyed M. Moghadas, Institute for Biodiagnostics, National Research Council Canada. E-mail: Seyed.Moghadas@nrc-cnrc.gc.ca

Accepted 8 July 2010. Published online 18 August 2010.

Background The relatively mild nature of the 2009 influenza pandemic (nH1N1) highlights the overriding importance of pre-existing immune memory. The absence of cross-reactive antibodies to nH1N1 in most individuals suggests that such attenuation may be attributed to pre-existing cellular immune responses to epitopes shared between nH1N1 virus and previously circulating strains of inter-pandemic influenza A viruses.

Results We sought to identify potential CD4+ T cell epitopes and predict the level of cross-reactivity of responding T cells. By performing large-scale major histocompatibility complex II analyses on Hemagglutinin (HA) proteins, we investigated the degree of T-cell cross-reactivity between seasonal influenza A (sH1N1, H3N2) from 1968 to 2009 and nH1N1 strains. Each epitope was examined against all the protein sequences that correspond to sH1N1, H3N2, and nH1N1. T-cell cross-reactivity was estimated to be 52%, and maximum conservancy was found between sH1N1 and nH1N1 with a significant correlation ($P < 0{\cdot}05$).

Conclusions Given the importance of cellular responses in kinetics of influenza infection in humans, our findings underscore the role of T-cell assays for understanding the inter-pandemic variability in severity and for planning treatment methods for emerging influenza viruses.

Keywords Influenza A/H1N1, MHC II, pandemic dynamics, pre-existing immunity, T-cell cross-reactivity.

Please cite this paper as: Duvvuri *et al.* (2010) Highly conserved cross-reactive CD4+ T-cell HA-epitopes of seasonal and the 2009 pandemic influenza viruses. Influenza and Other Respiratory Viruses 4(5), 249–258.

Introduction

In April 2009, a novel triple reassortant influenza H1N1 virus (nH1N1), with a unique genomic profile combining seasonal H3N2 and H1N1 subtypes and a swine-origin H1N1 subtype, was identified in association with human respiratory illness in Mexico and California.[1,2] The rapid global spread of nH1N1 led the World Health Organization to declare an influenza pandemic on June 11, 2009.[3] Despite its novelty and widespread, the 2009 nH1N1 pandemic is characterized by relatively mild clinical outcomes in a vast majority of individuals.[4,5] Furthermore, the incidence of severe cases caused by nH1N1 appeared to be significantly lower than that caused by the human seasonal influenza viruses.[6] While pre-existing immunity because of prior exposure to similar viral strains may provide an explanation for the reduced severity of nH1N1 infection,[7–11] the extent to which different components of the host immune system (i.e., cellular and humoral responses) affect disease outcomes remains controversial.[12]

As a major surface protein of influenza viruses, hemagglutinin (HA) plays a pivotal role in viral infection by binding to surface receptors on respiratory epithelial cells.[13,14] The HA gene consists of two subunits HA1 and HA2, produced by enzymatic cleavage of a precursor HA molecule (HA0).[15–17] The HA1 subunit contains both highly conserved and variable regions. The HA1 gene being a major target for neutralizing antibodies, it is not surprising that it accumulates mutations in response to this strong immunological pressure.[18,19] The HA2 gene, in contrast to HA1, is highly conserved[20,21] and usually does not appear to be target for neutralizing antibodies mainly owing to limitations of its exposure on the viral surface.[21–24] Nevertheless, it has been shown to induce strong CD4+ T-cell responses.[25]

Previous studies have demonstrated a critical role of the CD4+ T-cells in the host's defense against influenza virus infection, in part by regulating the production of neutralizing antibodies from B-lymphocytes. Moreover, many T-cell epitopes are conserved across influenza virus strains, making available memory T-cells that could contribute to protective immunity.[26] In the context of influenza nH1N1 infection, recent serological studies have shown the presence of some level of cross-reactive antibody titers in groups of individuals older than 60 years of age, but no protection for children and younger adults.[27,28] Moreover, the presence of cross-reactive T-cells among influenza strains has been shown, even in the absence of cross-reactive antibodies.[8,11,29,30]

Understanding the nature of pre-existing immune responses in populations when a novel influenza virus strain emerges is critical for the formulation of effective, efficient public health responses to epidemics.[31,32] However, for reasons of practicality and cost, most large-scale efforts to rapidly evaluate population immune responses to emerging infectious diseases emphasize humoral rather than cellular immune responses. Given the importance of CD4+ T-cells in regulating B-cell and cytotoxic T-lymphocyte (CTL) responses,[33] we sought to investigate the degree of conserved CD4+ T-cell epitopes in the HA gene to determine whether prior exposure to seasonal influenza A strains would be expected to provide benefits against nH1N1 strains.

Results

Analysis of MHC class II predicted conserved epitopes

Amino acid sequences of HA1 and HA2 regions of sH1N1 and vH1N1 strains (as described in Materials and Methods section) shared, respectively, 72% and 91·8% identity with that of nH1N1. We focused our analysis on identifying epitopes recognized by CD4+ T-cells in the context of major histocompatibility complex (MHC) II. Epitope analysis identified 15 amino acids length peptide sequences binding to various MHC II alleles specific to human leukocyte antigen (HLA)-DRB1 (supplementary information, Tables S2–S6). A total of 147 strong binders were predicted with the MHC II allele (DRB1*0101) for vH1N1 and the 2008 sH1N1, whereas 124 strong binders were identified for nH1N1 strains with the same MHC II allele. Next-predicted MHC allele was DRB1*0701 with 54 strong binders in vH1N1 and the sH1N1 2008, and 51 strong binders in nH1N1. There are no binders identified with the following alleles: DRB1*0301, DRB1*0801, DRB1*1101, and DRB1*1301. We considered conserved regions at the level of the 9-mer sequence 'frame' that fits into the MHC binding groove and found a total of 119 CD4+ MHC II epitopes to be 100% conserved among sH1N1, vH1N1, and nH1N1 strains. Of these 119 conserved epitopes, 21 were from the HA1 region and 98 from the HA2 region. CD4+ MHC II epitopes found in HA1 and HA2 regions were predicted to bind with multiple MHC alleles (i.e., promiscuous MHC II epitopes; Table 1). Consistent with previous observations,[7] some of the HA1- and HA2-predicted MHC II binders were also identified to be MHC class I epitopes. A number of the MHC II epitopes from nH1N1 strain binding to MHC class I alleles are A*0101, A*0301 (LSSVSSFER); A*0201 (WTYNAELLV, YNAELLVLL, VTVTHSVNL, YQILAIYST); A*2402 (IYSTVASSL, FWMCSNGSL); and B*0702 (IPSIQSRGL). Notably, all of these eight epitopes were 100% conserved with the vaccine strain A/Brisbane/59/2007 H1N1 used in the 2008–2009 season, and also recommended for the 2009–2010 season.[34] Out of 18 MHC class II epitopes predicted in our analysis, eight were shown to be 100% conserved with MHC class I epitopes as reported by De Groot *et al.*[10] (Table 2). Some of these epitopes were

Table 1. 88·8% (8/9)–100% (9/9) conserved peptides of sH1N1 and nH1N1 with reference to Major Histocompatibility Complex class II alleles

HA region	DRB1*0101	DRB1*0401	DRB1*0404	DRB1*0701	DRB1*1501
HA1 (1–327)	YHANNSTDT LREQLSSVS LRNIPSIQS IPSIQSRGL FGAIAGFIE	LRNIPSIQS	LRNIPSIQS	VTVTHSVNL LSSVSSFER	IPSIQSRGL
HA2 (328–535)	IEKMNTQFT WTYNAELLV YNAELLVLL LVLLENERT YQILAIYST YSTVASSLV FWMCSNGSL YEKVKSQLK	LVLLENERT	LVLLENERT	IYSTVASSL YSTVASSLV	LVLLENERT LAIYSTVAS LVLLVSLGA

also found in the list of B-cell, T-cell class I and T-cell class II epitopes that are deposited in the Immune Epitope Database (IEDB).[11] Cross-reactivity between sH1N1 and nH1N1 was estimated to be 52%, by dividing the conserved epitopes by the total common binders as described in Table 3.

Correlation in epitope conservancy

To ensure that analyzed epitope data in this study correspond to those of the 2009 pandemic nH1N1 strain, we used the conserved epitopes (CD4+ T-cell) predicted from recent sH1N1 (2008) and vH1N1 (2007) strains of influenza. Mapping these conserved epitopes on the sequences submitted during the period of 1968–2009 for H3N2, 1977–2009 for sH1N1 and 2009 for nH1N1 strains, we found that the conservancy among all the individual epitopes ranges from 33·3% to 100% (Table 4). The number of predicted sH1N1-conserved epitopes in nH1N1 remained the same. Figure 1A–C represents mean conser-

Table 2. CD4+ and CD8+ T-cell predicted conserved epitopes of nH1N1, which are, respectively, 88·8% (8/9) and 100% (9/9) conserved with the seasonal influenza type H1N1. Mutations in epitopes are highlighted in red

HA region	Epitope category		Source
	T-cell ($CD4^+$) MHC-II	T-cell ($CD8^+$) MHC-I	
HA1 conserved epitopes (1–327)	LSSVSSFER	LSSVSSFER	This study[10]
	VTVTHSVNL	VTVTHSVNL	This study[10]
	IPSIQSRGL	IPSIQSRGL	This study[10]
	LRNIPSIQS		This study
	LREQLSSVS		This study
	FGAIAGFIE		This study
	YHANNSTDT		This study
HA2 conserved epitopes (328–535)	IEKMNTQFT		This study
	WTYNAELLV	WTYNAELLV	This study[10]
	YNAELLVLL	YNAELLVLL	This study[10]
	LVLLENERT		This study
	YEKVKSQLK		This study
	YQILAIYST	YQILAIYST	This study[10]
	LAIYSTVAS		This study
	IYSTVASSL	IYSTVASSL	This study[10]
	YSTVASSLV		This study
	LVLLVSLGA		This study
	FWMCSNGSL	FWMCSNGSL	This study[10]

MHC, Major Histocompatibility Complex.

Table 3. Estimation of cross-reactivity based on the conserved binders versus common binders. The probability of cross-immunity is the ratio of conserved strong binders to the total common strong binders (represented in percentage)

	MHC	Conserved strong binders	Total common strong binders	CD4+ T-cell cross-reactivity
nH1N1 (2009) versus vH1N1 (1999, 2006, 2007) and sH1N1(1977–2008)	DRB1*0101	58	119	48.3%
	DRB1*0301	0	0	0
	DRB1*0401	10	16	62%
	DRB1*0404	8	19	42%
	DRB1*0701	16	37	43%
	DRB1*0801	0	0	0
	DRB1*1101	0	9	0
	DRB1*1301	0	0	0
	DRB1*1302	0	0	0
	DRB1*1501	27	27	100
Cumulative		119	227	52%

Table 4. Conservancy ratios of predicted epitopes in sH1N1, H3N2, and nH1N1

Epitopes		Immune response	Within sH1N1 strains (1977–2009)	Within nH1N1 strains (2009)	Within H3N2 strains (1968–2009)
Position	Motif				
7–15	YHANNSTDT	CD4+	100%	100%	33·3–44·4%
24–32	VTVTHSVNL	CD4+ &CD8+	100%	88·8–100%	33·3–44·4%
101–109	LREQLSSVS	CD4+	88·8–100%	100%	33·3–44·4%
105–113	LSSVSSFER	CD4+ & CD8+	100%	88·8–100%	44·4%
317–325	LRNIPSIQS	CD4+	88·8–100%	88·8–100%	44·4–66·6%
320–328	IPSIQSRGL	CD4+ & CD8+	88·8–100%	88·8–100%	44·4–66·6%
329–337	FGAIAGFIE	CD4+	100%	100%	88·8–100%
382–390	IEKMNTQFT	CD4+	100%	100%	44·4–55·5%
418–426	WTYNAELLV	CD4+ & CD8+	100%	100%	66·6–88·8%
420–428	YNAELLVLL	CD4+ & CD8+	100%	88·8–100%	66·6–88·8%
423–433	LVLLENERT	CD4+	100%	100%	55·5–66·6%
445–452	YEKVKSQLK	CD4+	100%	88·8–100%	33·3–44·4%
510–518	YQILAIYST	CD4+ & CD8+	100%	100%	33·3–55·5%
514–522	LAIYSTVAS	CD4+	100%	100%	44·4%
515–523	IYSTVASSL	CD4+ & CD8+	88·8–100%	100%	44·4–55·5%
516–524	YSTVASSLV	CD4+	100%	100%	44·4–55·5%
523–531	LVLLVSLGA	CD4+	100%	88·8–100%	44·4–55·5%
534–542	FWMCSNGSL	CD4+ & CD8+	100%	100%	33·3–44·4%

vancy for each epitope over years between strain groups (nH1N1 versus sH1N1, nH1N1 versus H3N2, and sH1N1 versus H3N2). Mutated epitopes within years were also considered. This figure shows highly conserved epitopes dispersed in the HA region since the emergence of strains. Statistically significant correlation was found between sH1N1 versus nH1N1 epitope conservancy ($r = 0{\cdot}51$, P value = 0·03), whereas other groups H3N2 versus nH1N1 and sH1N1 versus H3N2 showed no significant correlation. This analysis supports our estimation of 52% cross-reactivity based on the conservancy.

Figure 2 highlights CD4+ and CD8+ specific as well as overlapping epitopes in the nH1N1-HA protein sequence. The predicted epitopes – YHANNSTDT (7–15), VTVTHSVNL (24–32), LREQLSSVS (101–109), and LSSVSSFER (105–113) – were found to be highly conserved (percentage of conservancy, 88·8–100%) in the sH1N1 (1977–2009) and nH1N1 (2009) strains. These epitopes overlap with experimentally verified neutralising antibody-binding sites,[35–38] as represented in Figure 2.

Discussion

The 2009 H1N1 pandemic appears (with reference to case–fatality rates) to have been the mildest influenza pandemic on record,[4] although severity has varied markedly across geographies and communities.[39] This lack of severity has resulted in a degree of retrospective criticism of the response to the 2009 pandemic as overly aggressive.[40,41] Older adults at highest risk of complications of severe influenza appear to have had a high degree of immunity to infection, although given that this has been a feature noted in prior pandemics, it is unlikely that this observation is sufficient to explain between-pandemic variability in severity.[42–45] We suggest that an important feature of the 2009 influenza A (H1N1) pandemic, which may have contributed substantially to the diminished severity of this pandemic, is the circulation of sH1N1 as a sometime-dominant seasonal influenza strain for some 23 years prior to the emergence of nH1N1. Although extensive cross-protection against nH1N1 in younger individuals would not have been expected based on documented sero-epidemiological profiles,[27,28] we demonstrate that atypical features of this pandemic are compatible with a major (and under-appreciated) role for pre-existing T-cell immunity against influenza nH1N1 infection.[10,11]

To elucidate whether there could be some level of cross-reactivity from CD4+ T-cells between sH1N1 strains and nH1N1, we conducted an epitope prediction-based analysis. Our results show the existence of a high level of CD4+ T-cell cross-reactivity that could influence disease outcomes. We failed to identify T-cell cross-reactivity between H3N2 and nH1N1 subtypes, which may be because of distinctive surface antigens.[46] Consistent with previous work, our analysis indicates that minimal CD4+ T-cell epitopes (i.e. core region) from nH1N1 HA1 and HA2 exhibit overlapping epitopes with CD8+ T-cell.[47,48] The generation of classical CD8+ effector CTL

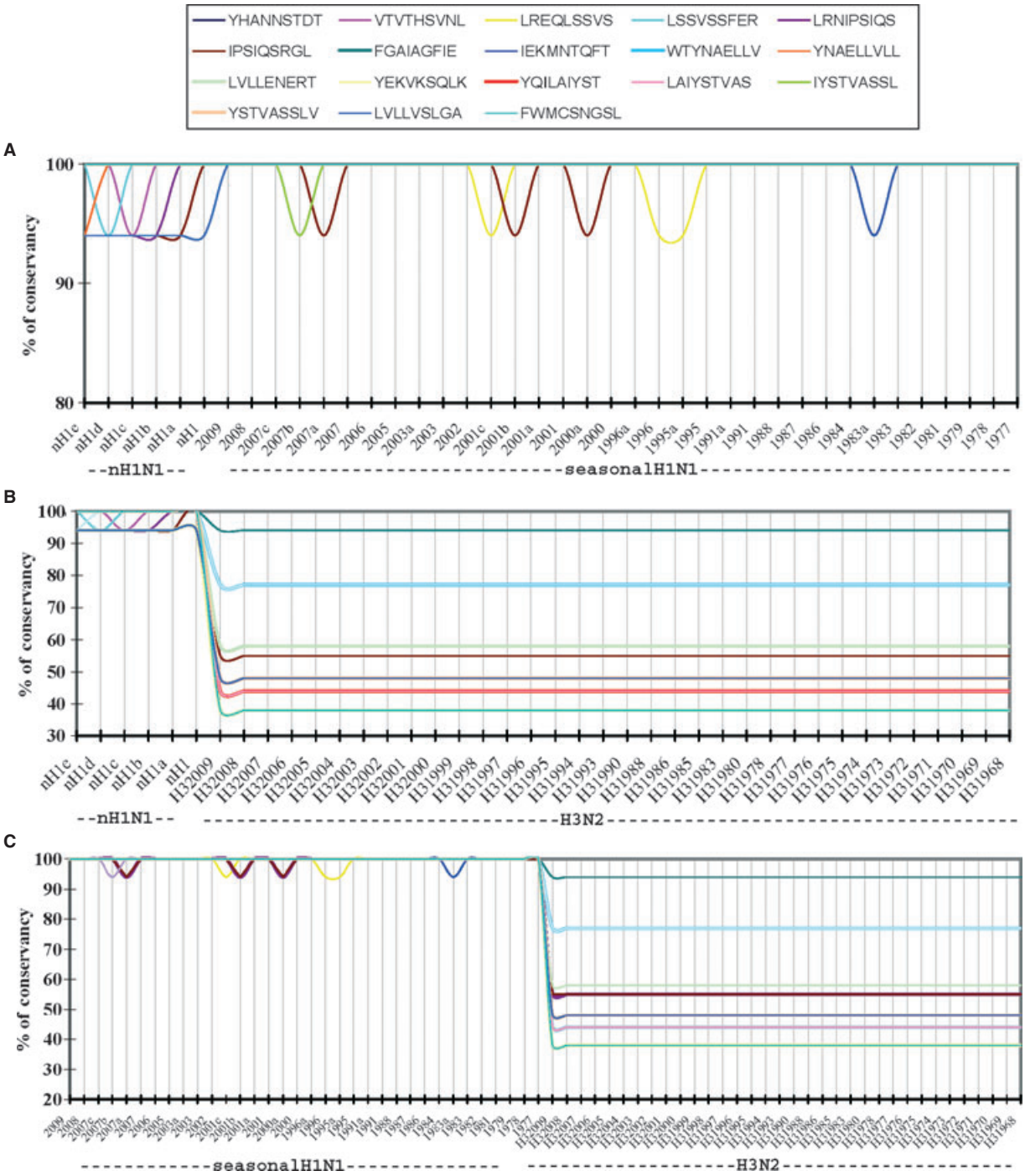

Figure 1. Comparative epitope conservancy: (A) nH1N1 and sH1N1. (B) nH1N1 and H3N2, and (C) sH1N1 and H3N2. Note: sH1N1 (1985, 1989, 1990, 1992, 1993, 1994, 1997, 1998, 1999, and 2004) sequences and H3N2 (1979, 1981, 1982, 1984, 1987, 1989, 1991, and 1992) sequences are not available in National Center for Biotechnology Information Influenza database. Hence, they are not represented in these figures. Mutated epitopes within the same year are represented as year followed by a, b, c. Accession numbers are given in Table S7 of the supplementary information.

responses generally require *in vivo* priming, either through natural infection or vaccination,[47] involving licensing of antigen-presenting cells (APC) because of APC and CD4+ T helper cell interaction in the context of MHC II. Such APC licensing is crucial for efficient induction of CTL responses.[49,50] Our study identifies epi-

gs	glycosylation site		CD4+ epitopes
as	antigenic site		CD4+ & CD8+ epitopes
rbc	receptor-binding site	eA	epitope A
tm	transmembrane site	eB	epitope B
		eC	epitope C
		eD	epitope D
		eE	epitope E

										gs	gs	gs									gs	gs	gs			eC	eC	eC	eC	eC	
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
D	T	I	C	I	G	Y	H	A	N	N	S	T	D	T	V	D	T	V	L	E	K	N	V	T	V	T	H	S	V	N	L	L
eC	eC	as	eC	eC			eE		eE			eE	eE	as			eE		as							eE			eE		eE	eE	eE
34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66
E	R	T	H	N	G	K	L	C	K	L	R	G	V	A	P	L	H	L	G	K	C	N	I	A	G	W	I	L	G	N	P	E
eE			eE	eE	eE			eE	eE		eE		eD			eE			eD	eD				rbs		eE						
67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
C	E	S	L	S	T	A	S	S	W	S	Y	I	V	E	T	P	S	S	D	N	G	T	C	Y	P	G	D	F	I	D	Y	E
	eD				eD	eA		eA		eA		eB	eB	eA	eA/as	eA	eA		eA		eA	eA		eA/as		eA/as	eA	eA/as	eA/rbs	eA		rbs
100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132
E	L	R	E	Q	L	S	S	V	S	S	F	E	R	F	E	I	F	P	K	T	S	S	W	P	N	H	D	S	N	K	G	V
	eA		eA/as		as	eB	eB/as	eB/as	eB	eB	eB			eB		eB/rbs		eD	eA/as	as	eD/gs	eD/gs	eD/gs	eD	eD	eD	eD	eD		eD		gs
133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165
A	A	C	P	H	A	G	A	K	S	F	Y	K	N	L	I	W	L	V	K	K	G	N	S	Y	P	K	L	S	K	S	Y	I
eD/gs	gs			eB	eB	eB	eB	eB	rbs	eB	eB	eB	rbs	eB	eB	eB/as			eD/as	rbs	eD		as/rbs	rbs	eD	eD	eD			eD	eD	eD
166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198
N	D	K	G	K	E	V	L	V	L	W	G	I	H	H	P	S	T	S	A	D	Q	Q	S	I	Y	Q	N	A	D	T	Y	V
eD	eD	eD	eD	eD							eD	eD	eD	eD	eD				rbs			rbs	eD/rbs		eD/rbs		eD		eD		eD	eD
199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231
F	V	G	S	S	R	Y	S	K	K	F	K	P	E	I	A	I	R	P	K	V	R	D	Q	E	G	R	M	N	Y	Y	W	T
eD											eE	eE	eE				eE								eC		eC	eC		eC	eC	eC
232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264
L	V	E	P	G	D	K	I	T	F	E	A	T	G	N	L	V	V	P	R	Y	A	F	A	M	E	R	N	A	G	S	G	I
													eC			eC		eC	eC		GS	GS	eC/gs	eC		eC	eC	eC	eC	eC	eC	
265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297
I	I	S	D	T	P	V	H	D	C	N	T	T	C	Q	T	P	K	G	A	I	N	T	S	L	P	F	Q	N	I	H	P	I
298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330
T	I	G	K	C	P	K	Y	V	K	S	T	K	L	R	L	A	T	G	L	R	N	I	P	S	I	Q	S	R	G	L	F	G
331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363
A	I	A	G	F	I	E	G	G	W	T	G	M	V	D	G	W	Y	G	Y	H	H	Q	N	E	Q	G	S	G	Y	A	A	D
364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396
L	K	S	TM	Q	N	A	I	D	E	I	T	N	K	V	N	S	V	I	E	K	M	N	T	Q	F	T	A	V	G	K	E	F
															as																	
397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429
N	H	L	E	K	R	I	E	N	L	N	K	K	V	D	D	G	F	L	D	I	W	T	Y	N	A	E	L	L	V	L	L	E
																as																
430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462
N	E	R	T	L	D	Y	H	D	S	N	V	K	N	L	Y	E	K	V	R	S	Q	L	K	N	N	A	K	E	I	G	N	G
																					gs	gs	gs									
463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495
C	F	E	F	Y	H	K	C	D	N	T	C	M	E	S	V	K	N	G	T	Y	D	Y	P	K	Y	S	E	E	A	K	L	N
																	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm	tm
496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528
R	E	E	I	D	G	V	K	L	E	S	T	R	I	Y	Q	I	L	A	I	Y	S	T	V	A	S	S	L	V	L	V	V	S
tm	tm																															
529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548													
L	G	A	I	S	F	W	M	C	S	N	G	S	L	Q	C	R	I	C	I													

Figure 2. Predicted epitopes and functional related sites represented on the novel H1N1-HA protein sequence.

topes that are conserved among different influenza strains and also represents overlapping CD4+ and CD8+ T-cell epitopes, which represent attractive novel candidates for the development of T-cell-based vaccines.

Human leukocyte antigen (HLA) is an important genetic regulator of adaptive immunity, especially for T-cell immune responses. In the current study, all the predicted CD4+ T-cell HA-epitopes are restricted to HLA-DRB1*0101, and some of these epitopes are promiscuous with other sub-alleles of DRB1, *0401, *0404, *0701, and *1501. The promiscuity between these epitopes suggests the possibility of acquired cross-immune responses to novel influenza infections from earlier exposures. Understanding the association between the immune responses to natural infection and HLA polymorphic genes is therefore crucial for the development of universal influenza vaccines based on the highly conserved and strain cross-reactive epitopes. Earlier work on seasonal influenza A viruses has identified

the importance of class II HLA-DR alleles and shown HLA-DR3 and DR4 to be associated with reduced elicitation of vaccine-induced immunity in patients with type I diabetes.[51] An increased frequency of DRB1*0701 has been shown among non-responders to trivalent subunit vaccines; however, these individuals were found to recognize identical CD4+ T-cell HA-epitopes of influenza viruses.[33] These observations warrant further investigation into the role of HLA polymorphisms and immune responses to infection, vaccination, and autoimmune diseases.

Our conclusions drawn from a bioinformatics study on HA protein corroborate a recent experimental analysis of cross-reactive CD4 T-cell memory response against nH1N1 conferred by prior exposure to sH1N1 viruses.[52] The immunodominant HA-epitopes, HA_{316}, (TGLRNIPSIQSRGLFGAIA), HA_{381} (SVIEKMNTQFTAVGK), and HA_{424} (ELLVLLENERTLDYH) (see Table S2), are shown to be highly conserved between sH1N1 and nH1N1.[52] In structural perspective, these conserved epitopes are found in the HA2 segment of HA protein, which is known to be a stalk region.[53,54] In line with previous work,[55] we have shown that there may be potential CD4 T-cell help for the B cells targeting the HA2 region, where the majority of conserved epitopes (66·6%) are unveiled and seems to be in the stalk of HA structure found in our analysis (Table 4).

Several factors may influence the degree of immunological cross-reactivity, including immunological history and frequency of exposure to variants of a specific viral strain,[56–58] and therefore conservancy of epitopes does not necessarily correspond to cross-reactivity. Using a highly efficient epitope prediction tool (NETMHCIIPan[66]) and considering all the HA protein sequences of H3N2, sH1N1, and nH1N1 strains available in the National Center for Biotechnology Information (NCBI) since their emergence, we have revised prior estimates of 41% CD4+ T-cell cross-reactivity[11] upwards by a large margin to 52%. Our analysis included nH1N1 HA protein sequences from April to August 2009 submissions from Influenza Virus Sequence Database and is limited to some degree by the lack of entries for more recent strains in the database.

A further practical implication of these findings relates to the urgent public health response to newly emerged influenza strains with epidemic or pandemic potential in humans: in this context population immunity is generally assayed via sero-epidemiological studies.[25,56] However, as our analysis demonstrates, estimating the prevalence of neutralizing antibody may not be sufficient for characterizing the epidemiology of the disease or accurately projecting the future course of epidemics. Although evaluation of the prevalence of pre-existing cellular immune responses to a novel influenza virus is likely to be more time-consuming, complex, and expensive than traditional sero-epidemiological studies, the widespread availability of commercial cytokine elaboration assays suggests that this may not be an unattainable goal in relatively resource-rich settings, and gathering information on cellular responses may allow for a more nuanced and efficient response to future epidemics and pandemics. It is notable that the behavior of several other respiratory pathogens (including mumps and pertussis) has proven difficult to predict using models based on sero-epidemiological data alone, suggesting a future role for the integration of information on cellular immune responses into such models.[57]

Our study has several limitations. Most notably, because of the tendency of influenza A viruses toward antigenic drift via high mutation rates for surface proteins,[59] further analysis of the sequences past August 2009 would be required to determine the conservancy of epitopes. To calculate the conservancy ratio of available sequences, we considered only epitopes that are 100% conserved. While some of the predicted epitopes in our analysis were in agreement with previous experimental studies,[11,25] new cross-reactive T-cell epitopes were identified – in particular CD4+ and CD8+ T-cell overlapping epitopes – and therefore further investigation should be conducted for the quantification of T-cell responses. Considering the role of MHC class II in HA epitope selection,[25] we analyzed all the HA protein sequences of seasonal H1N1, H3N2, and the novel H1N1 available at the NCBI influenza genome databank (Figure 1). While the analysis aimed at determining T-cell cross-reactivity, we note that the MHC presentation is not necessarily reflective of T-cell response, and therefore proliferation assays are needed to confirm the T-cell response to these predicted epitopes. Nevertheless, overlap of the predicted epitopes with experimental work[25] provides a good degree of validation and confidence to derive the implications of the cross-reactivity, which highlights the importance of pre-existing memory T-cell responses against an emerging influenza virus.

Although pre-existing immunity is a self-protection mechanism, its effects often extend well beyond the individuals, by influencing the transmission dynamics of the pathogen in the population as a whole. These effects may appear as a prolonged incubation period,[5] reduced severity of the disease,[4,26] and reduced infectiousness.[5,7] Further evaluation of cross-reactive T-cell immunity and its implications for epidemic dynamics at the population level remains an important task for modeling and simulations of disease spread and control. In this context, previous studies involving within-host models of viral-immune dynamics have demonstrated that pre-existing cellular immunity can also interfere with the evolutionary responses of influenza viruses and prevent *in vivo* emergence of drug resistance and its spread between individuals.[60,61] We hope that this

study, combined with the ongoing research on real-time monitoring of the 2009 influenza pandemic, will guide future research and help foster the design of scientific frameworks that strengthen links between viral-immune dynamics at the individual level and disease transmission and control at the population level. Such an integrative and trans-disciplinary approach will foster understanding of influenza virology, immunology, and epidemiology and consequently will improve the ability to respond to and control this still-deadly disease.

Materials and methods

Sequence analysis of influenza A viruses

A total of 217, 532, and 56 HA protein sequences of sH1N1, seasonal H3N2, and nH1N1 strains, respectively, from years 1968 to 2009 were obtained from the Influenza Virus Resource at the NCBI.[62] Identical sequences were deleted by using the option 'collapse identical sequences' (Sequence accession numbers were included in supplementary information S7). We employed BioEdit, a biological sequence alignment editor,[63] for the analysis of HA sequences. Multiple sequence alignments were performed using ClustalW[64] with default parameters offered by BioEdit as an external program. To predict the strongly conserved common epitopes against MHC II alleles, the following sequences, A/New Jersey/AF09/2008 for sH1N1; A/Brisbane/59/2007 for seasonal H1N1 vaccine (vH1N1); and A/California/04/2009 for nH1N1 were used. These epitope numbers were used to calculate the fraction of CD4+ T-cell cross-reactivity.

Major Histocompatibility Complex (MHC) class II allele's selection

For the selection of MHC alleles, we considered a previous investigation into the human CD4+ T-cell repertoire response toward influenza A virus HA gene following natural infection.[25] The subjects utilized were 12 unrelated healthy adult donors (ages 21–55 years) with a history of influenza (A/Beijing/32/92) infection without vaccination, and the control group consisted of 6 healthy individuals (ages 28–42 years) with no history of influenza-like illness during the preceding 4 years (Table 1 in[25]). Donors who expressed HLA-DRB1 and HLA-DQB1 MHC alleles with a recent history of influenza infection induced strong cell-mediated responses to the peptide pools derived from HA.

Predicting binding affinities of HLA-DR alleles and epitopes of HA gene

Selection of peptide binding prediction tool

Major Histocompatibility Complex-II peptide binding prediction servers have been evaluated[65] by measuring the prediction accuracy in terms of the area under the receiver operating characteristic curve (A_{ROC}), NETMHCIIPAN[66] has been identified as best predictor ($A_{ROC} > 0{\cdot}9$), and closely followed by PROPRED,[67] IEDB,[68] and MULTIPRED.[69] Hence, NETMHCIIPAN was chosen to calculate the binding affinities of peptide-HLA-DR alleles and to identify the optimal peptides.

HA gene comparative analysis of sH1N1 and nH1N1 against HLA-DR alleles

FASTA format of HA protein sequences of sH1N1, vH1N1, and nH1N1 were individually analyzed for binding affinities against the selected alleles DRB1*0101, DRB1*0301, DRB1*0401, DRB1*0701, DRB1*0801, DRB1*1101, DRB1*1301, and DRB1*1501 in the NETMHCIIPAN online server. The following information was provided by the server for each peptide: position, core region, log-transformed binding value (1−log50k), and binding affinity (nm) of strong (threshold 50·000) and weak (threshold 500·00) peptides.

Epitope conservancy analysis

For the initial analysis to predict the conserved epitopes for HLA-DR alleles, only recent strains from sH1N1 (2008), vH1N1 (2007), and nH1N1 (2009) were used. Predicted epitopes were mapped among the list of sequences (supplementary information S7) to elucidate the epitope conservancy over the years 1968–2009. Epitope Conservancy Tool[70] was employed to investigate the conservancy across the seasonal and novel HA proteins. Statistical analysis was performed using GraphPad Prism 5.0 to determine the correlation between conservancy of epitopes within the groups: sH1N1 versus nH1N1, H3N2 versus nH1N1, and sH1N1 versus H3N2.

Funding

This work was supported by the Canadian Institutes of Health Research (CIHR) – (Grant no. 101785), Canadian Federal Network of Centers of Excellence: Mathematics of Information Technology and Complex Systems (MITACS) – (Grant no. 561915), and GEOmatics for Informed DEcision (GEOIDE) – (Grant no. 563025), The Arthritis Society (Grant no. 560102), Natural Sciences and Engineering Research Council of Canada (NSERC) – (Grant no. 500700, 509502), and Early Researcher Award (ERA) – (Grant no. 541121). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing of interests

The authors declare that they have no competing interests.

Acknowledgements

SM, DF, and JW are members of Pan-InfORM and acknowledge the support of CIHR Catalyst Grant for Pandemic Preparedness (PI: SM). JW acknowledges the support of PHAC, GW acknowledges the support from The Arthritis Society and JH acknowledges the support of NSERC and ERA. VRSKD is supported by Canadian Consortium for Pandemic Preparedness Modelling (CanPan).

Author contributions

Conceived and designed the study: VD, SMM, GEW, JW. Performed sequence and epitope prediction analysis: VD, BD. Wrote the paper: VD, BD, SMM. Contributed reagents/materials/analysis tools: JW, GEW, DNF, JH, HG. All the authors have read the final version of the paper and approved it. The authors thank the reviewers for their insightful comments that have improved the paper.

References

1 Morens MD, Taubenberger JK, Fauci AS. The Persistent Legacy of the 1918 influenza virus. N Engl J Med 2009; 361:225–229.
2 Garten RJ, Davis CT, Russell CA *et al.* Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science 2009; 325:197–201.
3 Chan M. World now at start of 2009 influenza pandemic: statement to the press by WHO director-general, on 11 Jun 2009. http://www.who.int/mediacentre/news/statements/2009/h1n1_pandemic_phase6_20090611/en/. Accessed 01 January 2010.
4 Reed C, Angulo FJ, Swerdlow DL *et al.* Estimates of the prevalence of pandemic (H1N1) 2009, United States, April-July 2009. Emerg Infect Dis 2009; 15(12):2004–2007.
5 Tuite AR, Greer AL, Whelan M *et al.* Estimated epidemiologic parameters and morbidity associated with pandemic H1N1 influenza. CMAJ. 2009; 182(2):131–136. DOI:10.1503/cmaj.091807.
6 Presanis AM, De Angelis D, New York City Swine Flu Investigation Team *et al.* The severity of pandemic H1N1 influenza in the United States, from April to July 2009: a Bayesian analysis. PLoS Med 2009; 6(12):e1000207.
7 McMichael AJ, Gotch FM, Noble GR, Beare PA. Cytotoxic T-cell immunity to influenza. N Engl J Med 1983; 309:13–17.
8 Webby RJ, Andreansky S, Stambas J *et al.* Protection and compensation in the influenza virus-specific CD8+ T cell response. Proc Natl Acad Sci U SA 2003; 100:7235–7240.
9 McElhaney JE, Xie D, Hager WD *et al.* T cell responses are better correlates of vaccine protection in the elderly. J Immunol 2006; 176:6333–6339.
10 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.
11 Greenbaum JA, Kotturi MF, Kim Y *et al.* 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.
12 Ahmed R, Oldstone BA, Palese P. Protective immunity and susceptibility to infectious diseases: lessons from the 1918 influenza pandemic. Nat Immunol 2007; 8:1188–1193.
13 Skehel JJ, Wiley DC. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 2000; 69:531–569.
14 Steinhauer DA, Skehel JJ. Genetics of influenza viruses. Annu Rev Genet 2002; 36:305–332.
15 Wiley DC, Wilson IA, Skehel JJ. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 1981; 289:373–378.
16 Webster RG, Laver WG, Air GM, Schild GC. Molecular mechanisms of variation in influenza viruses. Nature 1982; 296:115–121.
17 Steinhauer AD. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. Virology 1999; 258:1–20.
18 Duvvuri VRSK, Duvvuri B, Cuff WR, Wu GE, Wu J. Role of positive selection pressure on the evolution of H5N1 hemagglutinin. Genomics Proteomics Bioinformatics 2009; 7:47–56.
19 Bok K, Abente EJ, Realpe-Quintero M *et al.* Evolutionary dynamics of GII.4 noroviruses over a 34-year period. J Virol 2009; 83:11890–11901.
20 Krystal M, Elliott RM, Benz EW Jr, Young JF, Palese P. Evolution of influenza A and B viruses: conservation of structural features in the hemagglutinin genes. Proc Natl Acad Sci USA 1982; 79:4800–4804.
21 Kawaoka Y, Yamnikova S, Chambers TM, Lvov DK, Webster RG. Molecular characterization of a new hemagglutinin, subtype H14, of influenza A virus. Virology 1990; 179:759–767.
22 Brand CM, Skehel JJ. Crystalline antigen from the influenza virus envelope. Nat New Biol 1972; 238:145–147.
23 Hedestam GBK, Fouchier RAM, Phogat S, Burton DR, Sodroski J, Wyatt RT. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. Nat Rev Microbiol 2008; 6:143–155.
24 Becht H, Huang RT, Fleischer B, Boschek CB, Rott R. Immunogenic properties of the small chain HA2 of the haemagglutinin of influenza viruses. J Gen Virol 1984; 65:173–183.
25 Gelder CM, Welsh KI, Faith A, Lamb JR, Askonas BA. Human CD4+ T-cell repertoire of responses to influenza A virus hemagglutinin after recent natural infection. J Virol 1995; 69:7497–7506.
26 Lee LY, Ha do LA, Simmons C *et al.* 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(10):3478–3490.
27 Katz J, Hancock K, Veguilla V *et al.* Serum cross-reactive antibody response to a novel influenza A (H1N1) virus after vaccination with seasonal influenza vaccine. MMWR Morb Mortal Wkly Rep 2009; 58:521–524.
28 Hancock K, Veguilla V, Lu X *et al.* Cross-reactive antibody responses to the 2009 pandemic H1N1 influenza virus. N Engl J Med 2009; 361:1945–1952.
29 Grebe KM, Yewdell JW, Benninik JR. Heterotypic immunity to influenza A virus: where do we stand? Microbes Infect 2008; 10:1024–1029.
30 Zimmer MS, Burke DS. Historical perspective — emergence of influenza A (H1N1) viruses. N Engl J Med 2009; 361:279–285.
31 Gojovic MZ, Sander B, Fisman D, Krahn MD, Bauch CT. Modelling mitigation strategies for pandemic (H1N1) 2009. CMAJ 2009; 181(10):673–680.
32 Tuite AR, Fisman DN, Kwong JC, Greer AL. Optimal pandemic influenza vaccine allocation strategies for the Canadian population. PLoS ONE 2010; 5(5):e10520.
33 Gelder C, Davenport M, Barnardo M *et al.* Six unrelated HLA-DR-matched adults recognize identical CD4+ T cell epitopes from influenza A haemagglutinin that are not simply peptides with high HLA-DR binding affinities. Int Immunol 1998; 10:211–222.
34 FDA. News Release: U.S. Food and Drug Administration approves Vaccine for 2009-2010 Seasonal Influenza. 2009. 07/20/2009.

accessed 01 January 2010 http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm172772.htm
35 Skehel JJ, Barnet BC, Burt DS *et al.* Immune recognition of influenza virus haemagglutinin. Phil Trans 1989; B 323:479–485.
36 Matrosovich M, Tuzikov A, Bovin N *et al.* Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. J Virol 2000; 74(18):8502–8512.
37 Lin T, Wang G, Li A *et al.* The hemagglutinin structure of an avian H1N1 influenza A virus. Virology 2009; 1:73–81.
38 Shen J, Ma J, Wang Q. Evolutionary trends of A(H1N1) influenza virus hemagglutinin since 1918. PLoS ONE 2009; 4(11):e7789. doi:10.1371/journal.pone.0007789
39 La Ruche G, Tarantola A, Barboza P, Vaillant L, Gueguen J, Gastellu-Etchegorry M. The 2009 pandemic H1N1 influenza and indigenous populations of the Americas and the Pacific. Euro Surveill 2009; 14(42): 1–6.
40 Allam MF. Influenza A (H1N1) pandemic: true or false alarm. J Epidemiol Community Health 2009; 63(10):862.
41 Donaldson LJ, Rutter PD, Ellis BM *et al.* Mortality from pandemic A/H1N1 2009 influenza in England: public health surveillance study. BMJ 2009; 339(101):b5213.
42 Laver G, Garman E. Pandemic influenza: its origin and control. Microbes Infect 2002; 4:1309–1316.
43 Nicholson KG. Human influenza; in Nicholson KG, Webster RG, Hay AJ (eds): Textbook of Influenza. Oxford: Blackwell Science Ltd., 1998; 219–264.
44 Reid AH, Taubenberger JK. The origin of the 1918 pandemic influenza virus: a continuing enigma. J Gen Virol 2003; 84:2258–2292.
45 Bush RM. Influenza evolution, Chapter 13 in Encyclopedia of Infectious Diseases: Modern Methodologies, Edt. Tibayrenc M, New Jersey, Hoboken: John Wiley & Sons, Inc., 2007.
46 Couch RB, Kasel JA. Immunity to influenza in man. Annu Rev Microbiol 1983; 37:529–549.
47 Ou D, Mitchell LA, Décarie D, Gillam S, Tingle AJ. Characterization of an overlapping CD8+ and CD4+ T-cell epitope on rubella capsid protein. Virology 1997; 235:286–292.
48 Carreno BM, Turner RV, Biddison WE, Coligan JE. Overlapping epitopes that are recognized by $CD8^{++}$. J Immunol 1992; 148:894–899.
49 Celis E, Tsai V, Crimi C *et al.* Induction of anti-tumour cytotoxic T lymphocytes in normal human using primary cultures and synthetic peptide epitopes. Proc Natl Acad Sci USA 1994; 91:2105–2109.
50 Stuhler G, Walden P. Collaboration of helper and cytotoxic T lymphocytes. Eur J Immunol 1993; 23:2279–2286.
51 Ruben FL, Fireman P, LaPorte RE, Drash AL, Uhrin M, Vergona R. Immune responses to killed influenza vaccine in patients with type 1 diabetes: altered responses associated with HLA-DR 3 and DR 4. J Lab Clin Med 1988; 112(5):595–602.
52 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(7):3312–3319.
53 Sui J, Hwang WC, Perez S *et al.* Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 2009; 16(3):265–273.
54 Wang TT, Palese P. Universal epitopes of influenza virus hemagglutinins? Nat Struct Mol Biol 2009; 16(3):233–234.
55 Steel J, Lowen AC, Wang T *et al.* An influenza virus vaccine based on the conserved hemagglutinin stalk domain, mBio doi:10.1128/mBio.00018-10. (downloaded on May, 01, 2010 from http://mbio.asm.org/mBio.00018-10v1.pdf)
56 Chen H, Wang Y, Liu W *et al.* Serologic survey of pandemic (H1N1) 2009 virus, Guangxi Province, China. Emerg Infect Dis 2009; 15(11):1849–1850.
57 Luz PM, Codeco CT, Werneck GL, Struchiner CJ. A modelling analysis of pertussis transmission and vaccination in Rio de Janeiro, Brazil. Epidemiol Infect 2006; 134(4):850–862.
58 Evavold BD, Sloan-Lancaster J, Wilson KJ, Rothbard JB, Allen PM. Specific T cell recognition of minimally homologous peptides: evidence for multiple endogenous ligands. Immunity 1995; 2:655–663.
59 Selin LK, Welsh RM. Plasticity of T cell memory responses to viruses. Immunity 2004; 20:5–16.
60 Handel A, Regoes RR, Antia R. The role of compensatory mutations in the emergence of drug resistance. PLoS Comput. Biol 2006; 2:1262–1270.
61 Moghadas SM. Dynamics of resistance emergence in influenza infection with compensatory mutations. Math Popul Stud (In press).
62 Bao Y, Bolotov P, Dernovoy D *et al.* The Influenza Virus Resource at the National Center for Biotechnology Information. J Virol 2008; 82:596–601.
63 Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 1999; 41:95–98.
64 Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22:4673–4680.
65 Lin HH, Zhang GL, Tongchusak S, Reinherz EL, Brusic V. Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research. BMC Bioinformatics 2008; 9(Suppl 12):S22.
66 Nielsen M, Lundegaard C, Blicher T *et al.* Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. PLoS Comput Biol 2008; 4:e1000107. doi:10.1371/journal.pcbi.1000107
67 Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics 2001; 17:1236–1237.
68 Peters B, Sidney J, Bourne P *et al.* The immune epitope database and analysis resource: from vision to blueprint. PLoS Biol 2005; 3:e91. doi:10.1371/journal.pbio.0030091
69 Zhang GL, Khan AM, Srinivasan KN, August JT, Brusic V. MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides. Nucleic Acids Res 2005; 33 (Web Server issue): W172–W179.
70 Bui HH, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinformatics 2007; 8:361.