Computer Simulations of the Humoral Immune System Reveal How Imprinting Can Affect Responses to Influenza HA Stalk with Implications for the Design of Universal Vaccines

Christopher S. Anderson[1], Mark Y. Sangster[1], Hongmei Yang[2], Sidhartha Chaudhury[3], David J. Topham[1*]

[1]New York Influenza Center of Excellence at David Smith Center for Immunology and Vaccine Biology, Department of Microbiology and Immunology, University of Rochester School of Medicine and Dentistry, Rochester, NY. [2]Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, NY. [3]BHSAI/U.S. Army MRMC, Frederick, MD.

*Corresponding author. E-mail: David_Topham@URMC.rochester.edu

Abstract

Antigenic drift of the H1N1 virus results in significant reduction in vaccine efficacy and often necessitates the production of new vaccines that more closely antigenically match the circulating strains. Efforts to develop a vaccine resistant to antigenic drift are ongoing and the HA stalk region of the influenza H1N1 virus has emerged as a potential target for vaccines due to its conservation across antigenically drifted strains. Studies of the 2009 pandemic H1N1 vaccine as well as candidate pandemic avian influenza vaccines have demonstrated that it is possible to boost antibody towards the stalk region, but for reason that are unclear, only in individuals who had not been exposed to antigenically similar viruses. Here we use stochastic simulations of a humoral immune system model to provide theoretical insights into how repeated exposure to influenza vaccines increases stalk-specific antibodies. We found that pre-existing memory B cells are the greatest contributor to stalk-specific antibody boosting and that pre-existing antibody negatively interferes with this boosting. Additionally, we found that increases in cross-reactivity after heterologous boosting occur in both head and stalk specific antibody populations. Moreover, pre-existing memory B cells focus antibody responses

29 towards the stalk region in a manner dependent on the antigenic dissimilarities between
30 other antigenic sites, even when these dissimilarities are minimal. Finally we show stalk-
31 specific antibody can be boosted by repeat exposure to homologous antigen, but this
32 boosting is limited. These finding provide needed insights into universal vaccine
33 regimens, especially those aimed at boosting stalk-specific antibody responses using
34 prime and boost strategies.

35 ## Introduction

36 The influenza glycoprotein hemagglutinin (HA) is a key target antigen for
37 protective antibody responses because it is expressed on the surface of the virus (and
38 infected cells) and is responsible for attachment to cellular receptors. The HA protein
39 contains a head region and a stalk region located proximal and distal to the viral
40 membrane, respectively. The head contains the receptor binding domains (where the
41 influenza viruses attaches itself to the host cell), while the stalk attaches the HA to the
42 viral envelope and mediates virus entry into the host cell. Although variation exists in
43 both the head and stalk region of HA, the stalk region is relatively more conserved and
44 more resistant to antigenic drift[1].The majority of protective antibodies in influenza
45 vaccines are directed to antigenic sites (epitopes) on the head of the HA because the
46 mechanism of action centers on the inhibition of cellular attachment[2]. Antibody
47 responses to the stalk domain are generally lower, at least after seasonal immunization,
48 and it is more difficult to measure correlates of protection because these antibodies do
49 not always affect cellular attachment or neutralization[3]. Antibodies to the stalk domain
50 can, nevertheless, be protective[4],[5]. Early experimental studies of H1N1 HA antigen
51 demonstrated that five, non-overlapping, antigenic sites on the HA exist[6]. More recent

studies have shown that, in addition to the head, the stalk region contains at least a single antigenic site[6,7].

In 2009, a zoonotic influenza H1N1 virus antigenically distinct to currently circulating viruses caused a world-wide pandemic. To combat the virus, a new vaccine was developed from the pandemic strain. Although immune responses to influenza vaccine are usually strain specific, studies of the immune responses to the pandemic vaccine demonstrated an increase in cross-reactive antibody to other antigenically distinct strains. Moreover, these antibody responses were associated with an increase in reactivity to the more conserved stalk region of HA. Close evaluation of the immunoglobulin genes of responding antibody secreting cells demonstrated high levels of somatic mutations suggesting a role of pre-existing memory B cells in the response to H1N1 pandemic virus vaccine[8]. Moreover, avian influenza candidate vaccine trials demonstrated similar findings suggesting this is a general phenomenon of exposure to novel influenza antigens and not specific to H1N1 viruses[9].

Exposure of naïve individuals to influenza virus by vaccination or natural infection leads to differentiation of naïve B cells into antibody secreting cells and memory B cells that make antibodies capable of neutralizing the virus and can exist in the body for decades[10]. In addition to longevity, memory B cells have a lower threshold of B cell receptor activation, higher proliferation rates, and contain somatic mutations providing greater affinity for its cognate antigen[11]. Additionally, since naïve B cells are continually replenished from the bone marrow, differentiation of naïve B cell into memory B cells increases the overall number B cells capable of recognizing influenza virus in circulation. In this way, exposure to influenza vitus by vaccination or infection

leads to an increased ability to mount an effective immune response to the virus upon secondary exposure. Not all HA antigenic sites elicit the same B cell response and formation of memory B cells, with B cells specific to head antigenic sites dominating the response compared to stalk antigenic sites[12], although memory B cells to the stalk due form [13]. The current thinking in the field is stalk-specific memory B cells are responsible for the difference in antibody specificities seen between seasonal and pandemic HA vaccines. HA head antigenic sites are somewhat conserved between seasonal vaccines but highly divergent in pandemic HA vaccines compared to seasonal strains[14]. It is thought that this leads to a lack of pre-existing memory B cells cross-reactive to the HA head antigenic sites of pandemic vaccines leading to decreased competition between stalk-specific and head-specific memory B cells. This decrease in competition leads to an increased stimulation of HA stalk specific memory B cells skewing the immune response (and antibodies) towards the HA stalk.

Elucidating the combined effect of differences in HA antigenic site conservation, pre-existing immunity, epitope dominance, and B cell and antibody specificities is a daunting task for the experimentalist. However, computational models allow explicit experimentation of biological parameters that are not possible to manipulate with typical animal and human models. Perelson et al. hypothesized that B cell receptor repertoires (paratopes) exist in an immunological shape space and antigen binding differences between them are represented as distance in shape space[15]. Smith et al. subsequently derived the parameters of such an immunological shape space for influenza viruses[16]. Moreover, Smith et al. developed a computational model of the humoral immune system and demonstrated that such a model can be used to

understand secondary immune responses to influenza[17]. Recently, Chaudhury et al. developed a stochastic simulation model using the parameters developed by Smith et al. and expanded the model to include multiple antigenic sites of different conservation[18].

We recently developed a method of estimating the antigenic relationships between the five canonical H1N1 HA head antigenic sites[19]. Here we use these estimates to expand the model developed by Chaudhury et al.[18] to include 6 epitopes representing the 5 canonical head antigenic sites and a conserved stalk antigenic site. Additionally, we expand the model to include long-lived plasma cells as was previously included in the Smith et al. model[17]. To gain theoretical insights into the role of memory B cells and pre-existing antibodies in the stalk-specific boosting of the antibody responses, we simulated humoral immune responses to the 2009 H1N1 pandemic vaccine HA antigen in systems previously exposed to antigenically similar (A/South Carolina/1/1918) or distinct (A/Brisbane/59/2007) HA antigen.

Materials and methods

Modeling HA antigen

The influenza virus HA antigen was chosen to model since antibody responses to this antigen is the primary target of vaccination. The following criteria was used to model the HA antigen: Each HA used in the simulation contains 5 distinct, equally dominant, antigenic sites representing the 5-canonical head antigenic sites of H1N1[20]. In addition, each HA contains a single, subdominant, stalk antigenic site[12].

Immunoglobulin interacts with antigens through the shape complementarity between the antigen-binding immunoglobulin paratope and the antigen epitope. To model antigen/immunoglobulin interactions we used the principals of Immunological Shape Space originally theorized by Perelson et al.[15]. Optimal Shape Space Theory parameters for influenza HA antigen were subsequently determined by Smith et al [16]. Following the example set by Smith, the shape of antigenic sites are represented by strings of symbols that symbolically represent their shape. The shape is represented by a 20-character string made up of 4 unique characters. The length and number of unique symbols at each locations gives the following properties: a potential immunoglobulin repertoire of 10^{12} B cells, a 1 in 10^{5} chance of B cells responding to a particular antigen, and an expressed repertoire of 10^{7} B cells [21-25].

In order to model the antigenic differences between strains and across antigenic sites, we used a sequence-based antigenic distance approach[19]. This approach uses HA protein sequence data to estimate the antigenic distance between antigenic sites (epitopic distances). For each head antigenic site (Sa, Sb, Ca1, Ca2, Cb), protein sequences were truncated to include only the amino acids in that site. Once an antigenic site amino acid sequence was obtained, the Hamming distance (i.e. number of amino acids differing between the sequences) was calculated. To normalize for differences in the number of amino acids in each epitope, this value was divided by the total number of amino acids in the site resulting in the percent difference (range 0 to 1) between the HAs. To convert to an individual string representing the epitopes, the percent difference was then multiplied by 20 (the number of symbols in each string) to create antigenic site-specific antigenic distances in the range of 0-20 as derived by

Smith et al.[16]. These antigenic site-specific distances are realized in the model by randomly changing symbols across the 20-symbol string until the Hamming distance between the virtual HA antigenic sites matched the epitopic distance calculated from the protein sequence. This was done for each antigenic site and then all antigenic sites were combined into a single virtual HA antigen in the model.

Unlike head antigenic sites, the exact number and location of the HA stalk region all possible antigenic sites are still largely unknown. Studies have demonstrated there are at least 1-2 epitopes in this region[12], with antibodies directed to the fusion domain having the ability to affect infectivity of the virus. We therefore chose to model a single HA stalk antigenic site (Stk). It is generally accepted that the stalk region of HA is highly conserved amongst H1N1 viruses therefore in our model the stalk antigenic site was completely conserved between HA strains, although it is likely that multiple antigenic sites exist in the stalk region and vary in conservation[1,26].

Modeling the Humoral Immune System

A first principle approach to modeling was used to create a computational model of the immune system. First principal approaches to modeling immunology attempt to use established parameters in immunological theory in order to simulate complex systems with the fewest assumptions or fitted parameters. These methods allow estimation of the true state of the system and can be used to understand complex interactions or reactions that arise from basic biological principles. Borrowing from work by Smith and Chaudhury, we chose to represent the immune response using a simplified version of the immune system representing only the B cell arm.

The model immune system consists of 7 agents: Naïve B cells, stimulated B cells, germinal center B cells, short-lived plasma cells, long-lived plasma cells, memory B cells, and antibody (Fig 1). In the model, Naïve B cells bind antigen and become stimulated B cells. Stimulated B cells then form germinal centers, becoming germinal center B cells capable of stochastically differentiating into plasma cells or memory B cells. Plasma cells secrete antibody capable of binding and removing antigen. Once primed, memory B cells can also be stimulated by antigen leading to stimulated B cells capable of forming germinal centers. Follicular helper T cells and antigen presenting cells are therefore modeled implicitly.

Fig 1. Schematic Representation of Humoral Immune System Model.

Simulating biological reactions

The agent based simulation method, developed by Chaudhury *et al*. 2013[18], uses a stochastic chemical-kinetics based approach[27,28] to simulate the progression of an immune response at a repertoire-scale, where components such as B cells, antigens, or antibodies are modeled as chemical species, and processes such as antigen-binding, somatic mutations, or B cell replication are modeled as chemical reactions; all parameters in the model were set as previously described[18] with of addition of long-lived plasma cell parameters(Table 1).

Table 1. Model Parameters

Parameter	Symbol	Value
No. of naïve B cells		5×10^7 cells

Ag dose		360 units
GC carrying capacity	k	5000 cells
Ag parameters		
Epitopes		6
Epitopes1		
Immunogenicity	y	0.8
Clearance	p	1.0 x 10^{-4} cells
Epitopic distance		0
Epitope 2-6		
Immunogenicity	y	1.2
Clearance	p	1.0 x 10^{-4} cells
Epitopic distance		Varied
Intrinsic decay	gAg	$(12\ h)^{-1}$
B cell parameters		
B cell enhancement factor	εB	10
Ab enhancement factor	εAb	2.5
Naïve B cell formation rate	kN	(4.6 x 105 $h)^{-1}$
Naïve B cell stimulation rate	σN	$(1\ d)^{-1}$
GC B cell sitmualte rate (base)	σbase	$(8\ h)^{-1}$
GC B cell stimulation rate (maximum)	σmax	$(15\ min)^{-1}$
GC B cell replication rate	r	$(8\ h)^{-1}$
Mutation probability	μ	0.1
Differntiation probability	δ	0.1
Memory cell stimulation	σMax	$(1\ d)^{-1}$
Ab production rate	kAb	1
Naïve B cell decay rate	gB	$(4.5\ d)^{-1}$
GC b cell decay rate (base)	gB	$(4.5\ d)^{-1}$
Plasma cell decay rate	gP	$(3\ d)^{-1}$
Ab decay rate	gAb	$(10\ d)^{-1}$

The simulated immune response to antigen was modeled by using a simplified version of the humoral immune system. The model attempts to describe the process that governs how naïve B cells become stimulated after encounter with antigen and

191 subsequently enter into germinal center reactions leading to production of memory B
192 cells and antibody secreting cells. Additionally, affinity maturation is described in a way
193 that models affinity increasing over time due to competition between B cells and limiting
194 antigen. Antigen is immediately removed from the system if bound by antibody. This
195 simplified approach allows the model to reflect how activation of the adaptive B cell
196 response leads to the eventual clearance of antigen from the host.

197 Simulations were carried out by defining a set of rate equations that describe the
198 underlying biological reactions and then applying the Gillespie algorithm, a dynamic
199 Monte Carlo method [18,29]. The main parameters that were set in order to model B cell
200 responses to HA antigen in the simulation were the number of antigenic sites for each
201 antigen, the immunogenicity of each site, and the epitopic distance for each antigenic
202 site between HA strains. Immunogenicity parameters for HA antigenic sites were
203 chosen from experimental mouse studies[30,31]. These studies suggest that stalk-
204 specific antibody responses make up about 20% of the total response[30,32]. Therefore
205 an immunogenicity parameter value for each antigenic site was used to account for
206 these differences (Table 1). An immunogenicity parameter of one represents equal
207 immunogenicity. Because the HA head is immunodominant compared to the stalk,
208 immunogenicity parameters were adjusted to 0.8 for the stalk antigenic site and 1.2 for
209 all head antigenic sites which resulted in a primary antibody responses in the simulation
210 that was five-fold lower to the stalk antigenic site compared to a single head antigenic
211 site, comparable to experimentally determined responses[30,32].

212 In the simulation, the binding affinity between the paratope and antigenic sites is
213 proportional to the number of symbols that are complementary between them (their

Hamming distance). Two antigenic sites cease to be cross-reactive when a 35% change or more in amino acid sequence in the amino acids that make up each HA antigenic site[33,34]. Therefore, there are eight degrees of cross-reactivity between epitopes corresponding to epitopic distances 0 through 7. Although epitopic distance ranged from 0 to 7, affinity ranged for each site can vary from 4 to 7 reflecting the degeneracy in paratope sequences that can achieve maximum binding affinity to their epitope. This range represents a 10^4-fold difference in binding affinity of the immunoglobulin between naïve and fully matured B cells.

Rate Equations

The binding affinity (Qij) between paratope i and the epitope j is a function of their Hamming distances, d(i,j) (eq. 1). Parameter ε is an enhancement factor that reflects the fold increase in apparent binding of B cells compared to antibodies (10 and 2.5, respectively). The increased avidity of B cells in comparison to a single antibody in the model reflects the many immunoglobulin molecules found on the surface of B cells resulting in many possible interactions between B cells and antigen[18]. Parameters, α and λ, represent the minimum and maximum Hamming distance (4 and 7, respectively).

((1))
$$Q_{ij} = \begin{cases} 1, & d(i,j) < \alpha \\ \varepsilon^{\alpha - d(i,j)}, & \alpha \leq d(i,j) \leq \lambda \\ 0, & d(i,j) > \lambda \end{cases}$$

The model simulates an animal size B cell repertoire of 10^7-10^8 B cells [24,25]. The life expectancy of unstimulated (naïve) B cells is 4.5 days. Naive B cells with randomly generated immunoglobulin are created such that there is steady population of 5 x 10^7 naive B cells specific to each antigen and the numbers of naïve B cells are

236 balanced so equal numbers of B cells are specific to each antigenic site of the antigen.
237 For all antigens in the population (P_{Ag}), the rate ($P_{Ag}gN$) of naive B cell (N) decay was
238 set to $(4.5\ d)^{-1}$ (2a). The formation rate of naïve B cells (kN) was modeled as a first-
239 order reaction where the rate was dependent on the naïve B cell population size ($5x10^7$)
240 per antigen in the population (P_{Ag})) and set to $(4.6x105\ h)^{-1}$ (2b).

241 ((2a)) $$N_j \xrightarrow[P_{Ag}gN]{} 0$$

242
243 ((2b)) $$0 \underset{k_N}{\rightarrow} N_j \quad k_N = 5 \times 10^7 \cdot P_{Ag}\,gN$$

244
245 Naïve B cells stimulation was modeled as a second-order reaction between the
246 antigen and naïve B cells. The rate of this reaction was determined by the base
247 stimulation rate, immunogenicity of the antigenic site, and the binding affinity between
248 the paratope and antigenic site. The stimulation rate was set to 3 days for naïve and
249 memory B cells. The stimulation rate for germinal center B cells was set to 8hrs with a
250 maximum stimulation rate of 15 minutes reflecting the rapid stimulation of germinal
251 center B cells. The rate that naïve B cells (N) form germinal centers (B) was modeled as
252 a second-order rate equation dependent on the affinity of the B cell for an antigen i for
253 epitope j (Q_{ij}), antigen (Ag) epitope immunogenicity (□), and a stimulation rate multiplier
254 σN (3a). The rate of germinal center B cell (B) stimulation was modeled as a second-
255 order rate equation dependent on the affinity of the B cell for an antigen epitope (Q_{ij}),
256 antigen (Ag) epitope immunogenicity (□), and a stimulation rate multiplier σ_B (eq. 3b).

257 ((3a)) $$N_j + Ag_i \xrightarrow[\sigma_N \gamma_i Q_{ij}]{} B_j + Ag_i$$

258
259 ((3b)) $$B_j + Ag_i \xrightarrow[\sigma_B \gamma_i Q_{ij}]{} B_j^* + Ag_i$$

260

Germinal center B cell proliferation was modeled as a first-order reaction. The product of proliferation is a single daughter B cell containing at most a single mutation from the parent genotype. The replication rate (*r*) was set to a doubling time of 8hrs. Although in reality antigen is consumed during this process, for simplicity, antigen was not consumed during B cell activation. A constant rate of differentiation (δ) for germinal center B cells was used with a probability of differentiation set to 0.1. Germinal center B cells have equal probability of differentiating into antibody secreting cells or memory B cells. Antibody secreting cells had a 75% chance of having a half-life of 3 days (short-lived antibody secreting cells), and a 25% chance of having a half-life of 200 days (long-lived antibody secreting cells). Affinity maturation occurs in the germinal center under high apoptotic pressure that drives the selection of higher-affinity immunoglobulin receptors. A carrying capacity for the germinal center was set to 5000 B cells. As the germinal center B cell population expands so does the rate of germinal center B cell decay. When the germinal center reaches the carrying capacity, the germinal center B cell decay rate reaches the replication rate halting further expansion of the germinal center. Germinal center B cell (*B**) proliferation was modeled as a first-order rate equation dependent on the B cell replication rate (*r*) and the probability of mutation from genotype j to genotype k (R_{jk}) and is defined as $R_{jk} = (1 - \mu)^{19}(\mu/3)$ where μ is the mutation rate (eq. 4a). A first-order rate equation was used to model memory B cells (M) (eq. 4b) antibody secreting cells (*P*) dependent on the differentiation rate of *δ* (eq. 4c). Apoptosis of germinal center B cells was modeled using a second-order equation dependent on the apoptosis rate η, which was a function of the B cell replication rate (r) and the total GC B cell population (*B*) relative to the GC carrying capacity *k* (eq. 4d).

((4a)) $$B_j^* \xrightarrow[rR_{jk}]{} B_j + B_k$$

((4b)) $$B_j^* \xrightarrow[\delta]{} M_j$$

((4c)) $$B_j^* \xrightarrow[\delta]{} P_j$$

((4d)) $$B_j^* \xrightarrow[\max(\eta,\ gB)]{} 0 \qquad \eta = r\,(B/\kappa)$$

Antibody is produced from antibody secreting cells with a decay rate based on a half-life of 3 days. Each antibody in the simulation represents a large number of real antibodies. Antibody production was dependent on presence of antibody secreting cells (P_s, P_l), which contain different decay rates. Antibody is production was modeled based on a production rate, k_{Ab} (eq. 5a). Short-lived antibody secreting cell (*sP*) decay was modeled as a first-order reaction with a decay rate of gP_s (eq 5b). Long-lived antibody secreting cell decay was modeled as a first-order reaction with a decay rate of gP_l (eq. 5c).

((5a)) $$P_j \xrightarrow[\kappa_{Ab}]{} P_j + Ab_j$$

((5b)) $$sP_j \xrightarrow[gP_s]{} 0$$

((5c)) $$lP_j \xrightarrow[gP_l]{} 0$$

Memory B cells do not decay in the simulation. In the model, memory B cells can also give rise to germinal center B cells and antigens. Memory B cells have an

311 increased rate of simulation compared to naïve B cells giving them a competitive
312 advantage independent of their genotype. Additionally, given that memory B cells can
313 arise from mutated germinal center B cells with increased affinity, memory B cells can
314 also have greater affinity for the antigen compared to naive B cells, giving them an
315 additional rate advantage over naïve B cells. The rate of memory B cells was dependent
316 on the stimulation rate set to $(1d)^{-1}$ and the immunogenicity parameter (γ).
317
318 ((6)) $$M_j + Ag_i \xrightarrow[\sigma_M \gamma_i Q_{ij}]{} B_j + Ag_i$$
319
320 Antibodies bind and remove antigen using a second-order reaction with a
321 reaction rate that is the function of the binding affinity between the antibody paratope
322 and antigen epitope, as well as the clearance and neutralization parameter (these
323 values were constant between all epitopes) (eq. 7a). Intrinsic antibody (Ab) decay was
324 based on a half-life of 10 days and modeled using a first-order rate equation dependent
325 on the decay rate g_{Ab} (eq. 7b). Intrinsic antigen decay was modeled based on a half-life
326 of 12hrs and modeled using a first-order reaction dependent the antigen decay rate g_{Ag}
327 (eq. 8).
328
329 ((7a)) $$Ab_j + Ag_i \xrightarrow[\rho_i Q_{ij}]{} Ab_j$$
330
331 ((7b)) $$Ab_j \xrightarrow[g_{Ab}]{} 0$$
332
333 ((8)) $$Ag_j \xrightarrow[g_{Ag}]{} 0$$
334

Historical Strain Protein Sequences

Influenza HA protein sequences used in the model were obtained from Genbank: A/California/07/2009 [NC_026433], A/Brisbane/59/2007 [KP458398], A/South Carolina/01/1918 [AF117241], A/Beijing/262/1995[AAP34323], A/Brazil/11/1978 [A4GBX7], A/Chile/1/1983 [A4GCH5], A/New Caledonia/20/99 [AY289929], A/Singapore/6/1986 [ABO38395], A/Solomon Islands/3/2006 [ABU99109], A/USSR/90/1977 [P03453], A/New Jersey/11/1976 [ACU80014].

2009 H1N1 Vaccine Clinical Trial Human Serum

As a means to test specific predictions of the simulations in a real world situation, healthy adults and children were enrolled in age cohorts as previously described[13]. Results of this clinical trial have been published previously[13]. Subjects received a single intramuscular (i.m.) injection of inactivated influenza A/California/07/2009 (H1N1) monovalent subunit vaccine (Novartis). Each 0.5-ml dose contained 15µg of HA antigen. Administration of the vaccine (study day 0) took place from January 2010 to March 2010. The study was conducted under a protocol approved by the University of Rochester Research Subjects Review Board. Informed written consent was obtained from each participant. ClinicalTrials.gov identifier NCT01055184.

Enzyme-linked Immunosorbent Assay

Recombinant HA proteins were obtained from Influenza Reagent Resource (Cat#: FR-67, FR-692, FR-65, FR-180, FR-699) and BEI Resources (Cat# NR-19240, NR-48873). Chimera proteins were a gift from Dr. Florian Krammer.

Enzyme-linked immunosorbent assays were performed using recombinant HA proteins coated on MaxiSorb 96-well plates (ThermoSci; 439454) overnight at 4°C. Plates were blocked with 3% bovine serum albumin (BSA) in phosphate buffered saline

(PBS) for 1hr at room temperature. Serum was diluted 1:1000 in PBS/0.5% BSA/0.05% Tween-20. Plates were washed and incubated with alkaline phosphatase (AP)-conjugated secondary antibody for 2 hrs at room temperature. Plates were washed and developed using AP substrate (ThermoSci 34064).

Antigen Clearance Kinetics

A separate exponential decay model was fit to the data for each group to assess the difference in exponential decay rate. The model is specified as:

$$y_{ij} = (V_1 + b_{1i}) * \exp\{-(\beta_1 + b_{2i})t_j\} + \epsilon_{ij} \text{ for group SC18,}$$

And

$$y_{ij} = (V_2 + b_{1i}) * \exp\{-(\beta_2 + b_{2i})t_j\} + \epsilon_{ij} \text{ for BR07}$$

where y_{ij} is the viral load for the i^{th} subject at the j^{th} time point, and t_j represents the time points in hours. Random effects, b_{1i} and b_{2i}, are included to account for between-individual variability. The exponential decay rate is represented by β_2 and β_2 for group 1 and 2, respectively. To test for differences in the decay rates, the model was fit using PROC NLMIXED within SAS v9.4.

Statistics

All group comparisons were done using Student's two-tailed t-test. A p-value of 0.05 or less was considered statistically significant.

Results

To better understand how repeated exposure to vaccines can focus the immune response to the stalk region of the influenza HA protein, we expanded the stochastic model developed by Chaudhury et al.[18] (Fig 1). The model represents a simplified humoral immune system where a B cell is represented as a character string (e.g. "AAAAABBBBBCCCCCDDDDD") which are randomly generated by a random number

generator, reflecting the random nature of B cell receptors development *in vivo*. Naive B cells are continually generated and naturally decay unless stimulated by antigen, where they differentiate into memory B cells and plasma cells. During the simulation, B cells had a probability of being stimulated by their cognate antigen strings, replicating, differentiating into plasma or memory B cells, and producing antibody. The six HA antigenic sites (5 head, 1 stalk) are also represented as character strings in the model(e.g. "BBAAABBBBBCCCCCDDDDD"), but these strings are derived from virus sequence data using a sequence-based antigenic distance approach[19]. During the simulation antibody can then bind the antigen and remove it from the simulation. In this way, the model captures the antigenic differences between strains and reflects the ability of the immune system to adapt to inoculum producing antigenic site-specific antibodies from randomly generated B cell receptors.

Antigenic Distance Determination

We first determined antigenic distances (AD) using protein sequence data for 11 HA proteins using the sequence-based antigenic distance approach previously described[19]. Vaccine and prototypical influenza virus strains were chosen to represent antigenically distinct strains that have circulated since 1918 (see methods section). Given that each HA in the model contains 6 antigenic sites, and each antigenic site in the model contains 20 positions, the maximum epitopic distance (antigenic-site-specific antigenic distance; ED) is 20 and the maximum AD for each antigen is 120. 0verall, SC18 and CA09 had the greatest similarity with an AD of 21 (Table 2) with BR07 and CA09 having the greatest difference with 53 AD. Antibody cross-reactivity in the model occurs when an epitopic distance is seven or less in the model[17,18]. SC18 and CA09

had four of the five head epitopes with an ED of less than or equal to seven, with the Sa antigenic site having the least distance (Table 3). Alternatively, SC18 and BR07 had only one antigenic site with an ED of less than seven. Thus, in the model SC18 was antigenically more similar to CA09 while BR07 was largely antigenically distinct (Table 2). Using these distances, strings virtually representing the antigenic sites of HA were constructed.

Table 2 Antigenic Distances

	SC18	PR34	NJ76	US77	BR78	SI86	CH83	BE95	NC99	SI06	BR07	CA09
SC18	0											
PR34	43	0										
NJ76	16	43	0									
US77	50	46	50	0								
BR78	50	46	50	0	0							
SI86	45	48	53	18	18	0						
CH83	50	48	53	4	4	16	0					
BE95	53	48	53	22	22	21	23	0				
NC99	48	45	55	31	31	18	29	10	0			
SI06	47	50	54	35	35	23	33	18	10	0		
BR07	49	50	56	34	34	22	32	17	7	8	0	
CA09	21	48	24	55	55	50	55	58	53	51	53	0

Table 3 Epitopic Distances

Epitope	ED to CA09
SC18-Sa	2
BR07-Sa	8
SC18-Sb	3
BR07-Sb	15
SC18-Ca1	5
BR07-Ca1	7
SC18-Ca2	8

BR07-Ca2	10
SC18-Cb	3
BR07-Cb	13

Simulating the 2009 Pandemic

Simulations were carried out representing the real-life scenarios that occurred during the 2009 influenza pandemic. Two scenarios were simulated with 50 simulations carried out for each scenario. In scenario one, the model was “immunized” (primed) with the 1918 pandemic virus HA, A/South Carolina/1/1918 (SC18), then the immune response was allowed to resolve for 365 days, during which antibody level to returned close to baseline. Scenario two was identical to scenario one except the model was primed with the 2008 vaccine strain HA, A/Brisbane/59/2007 (BR07). Both groups were then re-immunized (boosted) with the 2009 pandemic virus vaccine HA (CA09). In this way, the “SC18 primed group” represented individuals that in 2009 had been previously exposed to 1918-like viruses and the “BR07 primed group” represented individuals primed by more recent seasonal influenza strains. B cell and antibody counts, genotype, and antigen specificities were tracked during the simulation allowing quantification of antigenic-site-specific B cells and antibodies during the simulation.

Antigenic Site Specific Antibody Responses

Given that the model represents a naïve immune system, immune responses specific to priming antigen (SC18 or BR07) should be identical between groups. To determine that the model was unbiased towards which priming antigen was used, antigenic site specific antibodies and memory B cells specific to the priming Ag were measured throughout the simulation, while in humans they are typically measured at

440 day 28-30 post immunization. Counts of antibody and memory B cells in the simulation
441 reactive to the priming antigen were similar across head epitopes and between groups
442 (Fig 2A, S1 Fig A-D) and these similarities remained up until boosting (S1 Fig E-F).
443 Stalk-specific antibody and memory B cell counts were significantly less than head
444 antigenic sites making up about 10% of the total response and were similar for both
445 groups (S1 Fig E and F). Overall, antibody and memory B cell counts and specificities to
446 their priming antigen were similar for both priming groups demonstrating that the model
447 is unbiased towards the priming antigen.
448
449 **Fig 2. Immune Responses After Prime and Boost.**
450 (A) Antigenic-site-specific antibody titers to each HA epitope included in the model for
451 the SC18 primed group (right) and BR07 primed group (left). Curves represent average
452 titers for 50 simulations and range is the standard deviation. (B) Average antibody titers
453 cross-reactive to CA09 pre-boost (Day 365) and (C) memory B cells. (D) Affinity
454 (antigenic distances 1-7) of memory B cells to CA09 HA antigen. (E) Percent of stalk
455 antigenic site specific antibody for each priming group.
456
457 Although immune responses to their priming antigen were similar between
458 groups, we sought to determine if the cross-reactivity to CA09 was different between the
459 priming groups prior to boosting as expected by the closer antigenic distance between
460 SC18 and CA09 compared to BR07 an CA09. Unlike reactivity to priming antigen,
461 antibody and memory B cells cross-reactive to CA09 after priming was markedly
462 differently between groups with significantly higher CA09 cross-reactive antibodies in

the SC18 primed group compared to the BR07 primed group with a greater than 2-fold difference in cross-reactive antibodies and memory B cells to CA09 just prior to boosting (Fig 2B-C). The affinity of the cross-reactive memory B cells to CA09 was also different between the groups, with the SC18 group having higher affinity compared to the BR07 group (Fig 2D). Therefore, although immune responses specific to the priming antigen were similar between groups, cross-reactive antibody and memory B cells were significantly different.

Next, we sought to determine the effect of priming on secondary immune responses to CA09 and assess differences in B cell/antibody totals and antigenic site specificities. After boosting with CA09, total antibody levels reactive to CA09 in the SC18 group were higher compared to the BR07 group, although this difference did not reach significance (S2 Fig A). The SC18 primed group produced a Sa-antigenic-site dominant response to CA09 with Sb and Cb antigenic-site specific antibodies also boosted (Fig 2A). Stalk antigenic site antibody was also boosted but to a lesser extent than the head epitopes making up about 15% of the antibody response (Fig 2E). In contrast, stalk-specific antibody responses for the BR07 primed group dominated (Fig 2A) comprising 35% of the total antibody response (Fig 2E) and showed a more moderate increase in other antigenic site specific antibodies (Fig 2A). These antigenic site-specific differences between groups generally corresponded to differences in epitopic distances between the primary and secondary antigen (Table 3) except for the stalk antigenic site, which is conserved between priming and boost antigens in both groups.

Pre-Exposure Affects Cross-Reactivity of Secondary Responses

Given differences in antigenic-site-specificities between groups and the dissimilarities in conservation of those sites between strains, we set out to determine antibody cross-reactivity after boosting with CA09 to a panel of antigenically distinct strains that have circulated since 1918. Cross-reactive antibody responses were measured day 30 post-boosting with CA09. Both groups had strong responses to the antigens to which they had been exposed, but differed largely in responses to other strains (Fig 3A). Generally, the SC18 group was cross-reactive to strains antigenically similar to CA09, while the BR07 group's antibody response was cross-reactive to more antigenically distinct strains. Cross-reactive titers in the SC18 primed group correlated well with the antigenic distance from CA09, while the BR07 primed group antibody cross-reactivity showed no linear correlation with antigenic distance (pval = 0.0001, pval = 0.4983; respectively). Therefore, although antigenic distance was a good predictor of cross-reactivity during the primary response of the simulation, for secondary immune responses, epitopic distance alone is not sufficient to predict cross-reactive immune responses.

Fig 3. Crossreacitity After Boosting with CA09.

(A) Heatmap of all simulations for the SC18 primed group and the BR07 primed group measured at day 30 post-boost (day 395) to 12 historical HA antigens. Values were log transformed. (B) Pie-chart representing the percent of simulations that contain cross-reactive antibodies to 1-11 strains for each group. Asterisk represents statistically significant difference (pValue < 0.05) between SC18 and BR07 groups.

Overall, the BR07 primed group had a statistically significant increase in cross-reactive responses to all antigens compared to the SC18 group with 20% of the antibodies cross-reactive to all 11 strains (Fig 3B). Additionally, cross-reactivity to 5-7 strains was also boosted indicating that the increase in cross-reactivity in the BR07 group was not only due to boosting of stalk-specific antibodies but also increased cross-reactivity of antibodies specific to the HA head antigenic sites. Interestingly, antigen was cleared more quickly in the SC18 group compared to the BR07 group ($p = 0.0001$; S2 Fig B) suggesting that pre-existing cross-reactive immunity affects antigen load, and may limit the duration of antigen stimulation. Overall, the BR07 primed group produced a greater cross-reactive antibody response compared to the SC18 primed group due to both an increase in cross-reactive stalk and head antigenic site specific antibodies.

Contribution of Memory B cells and Antibody on Cross-Reactivity

Since cross-reactive antibodies and memory B cells to CA09 existed prior boosting with CA09, we sought to address the contribution of preexisting antibody and memory B cells on the increase in stalk-specific antibodies and cross-reactivity after exposure to CA09. To this end, two perturbed models were created. For one model ("No Clearance") the antibody clearance was removed from the simulations in such that only basal decay of the antigen occurred. In this way, the effect of antibody-mediated removal of antigen on the secondary immune response was assessed. For the other model ("No Memory") the memory B cell activation was removed from the simulation such that only naïve B cells contributed to the germinal center reactions. In this way, the effect of memory B cell germinal center seeding on the cross-reactivity of the secondary immune response could be assessed. By comparing these two models, the relative

532 contribution of antibodies and memory B cells on cross-reactivity of the secondary
533 immune response in the simulation could be assessed.

534 Both perturbations affected the cross-reactive and stalk specific response after
535 boosting with CA09 for both priming groups, although in surprisingly different ways. For
536 the SC18 group, removal of the memory B cell germinal center contribution relatively
537 increased the cross-reactivity to historical antigens compared to the unperturbed
538 ("Normal") model (Fig 4A) although stalk-specific antibody was decreased compared to
539 "Normal" model (Fig 4C). Removal of antibody clearance for the SC18 group also
540 increased antibody cross-reactivity, but to a lesser extent compared to the "No Memory"
541 model (Fig 4A). For the BR07 group, removal of antibody clearance also increased the
542 cross-reactive response, but unlike the SC18 group, removal of memory B cells from
543 the germinal centers drastically decreased the cross-reactive response (Fig 4B). Stalk-
544 specific antibody was also significantly increased in the "No Clearance" model, but
545 significantly decreased in the "No Memory" model. Therefore, both antibodies and
546 memory B cells affect the antigenic sites targeted during the secondary immune
547 response, but how memory B cell affects the immune response depends on the
548 antigenic relationship between the priming strain and secondary strain.

549

550 **Fig. 4 Antibody Titers After Perturbation of the Model**

551 (A) Antibody titers measured at day 30 post-boost with CA09 for the SC18 primed group
552 (B) or BR07 primed group. Data was generated from the Normal model, No Ab
553 Clearance model, and No Memory Stimulation model. Values were log transformed. (C)
554 Stalk-specific antibodies 30 days post boost with CA09.

555

556 ## Boosting HA Stalk Responses

557 In order to better understand stalk-specific antibody responses in the model,
558 additional simulations were performed where a single parameter was changed and stalk
559 antigenic site-specific antibody levels were measured. Four parameters were tested:
560 stalk antigenic site immunogenicity, the number of antigen exposures, the number of
561 head antigenic sites, and head antigenic site epitopic distances. 50 simulations were
562 performed for each type of simulation and the count of stalk reactive antibodies was
563 tracked. Data is presented as the average of the 50 simulations.

564 The antigenic site immunogenicity parameter simulates changes in the minimum
565 B cell receptor affinity required to stimulate a B cell. Low immunogenicity antigenic sites
566 require higher affinity B cells compared to higher immunogenicity antigenic sites. For
567 simplicity, simulations were prime and boosted with a two-antigenic-site antigen (i.e.
568 head and stalk) with equal epitopic distance (homologous). The antigenic site
569 immunogenicity was varied over a two-fold range (0.6-1.2). As the stalk antigenic site
570 immunogenicity was increased the stalk-specific antibodies steadily increased (Fig 5A).
571 A two-fold increase in the immunogenicity parameter (0.6 to 1.2) led to a 20% increase
572 in stalk epitope-specific antibodies on average (5045 to 6044). Therefore,
573 immunogenicity of an antigenic site does modestly impact the level antibodies against
574 that antigenic site.

575

576 **Fig. 5 Model Parameters on Stalk Binding Antibody Counts**

577 Counts of antibodies specific to stalk antigenic site were determined at day 30 post-
578 boost of the second antigen challenge. (A) Stalk specific antibody counts in models with

579 different stalk epitope immunogenicity parameters. (B) Stalk specific antibody counts in
580 model with different numbers of homologous antigen exposures (C) Stalk specific
581 antibody counts in models with antigens that contain different numbers of head
582 antigenic sites D) Stalk specific antibody counts in models where the head epitopic
583 distance is increased.
584
585 Recently, Nachbagaer et al demonstrated that both stalk and head-specific
586 antibodies are increased upon repeated exposure to influenza [35]. In our original
587 simulations, each group was exposed to antigen twice (prime and boost) and both
588 groups showed an increase in stalk specific antibodies (see Fig 2A), but it is not clear to
589 what extent this increase was due to repeat exposure to antigen and how much was
590 due to antigenic properties of the boosting antigen. In order to separate heterologous
591 affects and repeat exposure affects, the number of exposures to homologous antigen
592 was varied (1-5 exposures). Stalk-antigenic-site-specific antibodies increase from prime
593 to boost by 32% and rose only slightly after with additional exposures (Fig 5B).
594 Therefore, stalk-specific antibody responses can be boosted by repeat exposure to
595 homologous antigen, but there is a limit and the levels quickly plateau.
596 Next, the number of antigenic sites defined in the model was varied. Although 5
597 canonical antigenic sites have been described, others have reported additional
598 antigenic regions[36]. All parameters were kept constant except the number of head
599 antigenic sites, which was varied from 1-6 sites. Stalk-antigenic-site specific antibodies
600 decreased as the number of head epitopes increased (Fig 5C). Increasing the number
601 of head epitopes from one to six led to a 68% decrease in the number stalk-antigenic-

site specific antibodies. Therefore, the number of head epitopes used is not arbitrary, and the choice does affect the level of antibodies specific to the stalk. Although we chose to explicitly model subdominance as an intrinsic property of the stalk antigenic site as has been reported [8], it is likely that this subdominance also occurs as a result of the ratio of stalk to head antigenic sites.

Lastly, although the change in epitopic distance of head epitopes is thought to be the cause of the increase in stalk-antigenic site specific antibodies seen after boosting with CA09 in the BR07 group in our simulations, the extent that antigenic change in the head increases antibody responses to the stalk was not directly tested. Therefore, to evaluate the effect of epitopic distance of head antigenic sites on the stalk-specific antibody response, a two-antigenic-site antigen (head and stalk) was used. All parameters were kept constant except the epitopic distance of the head antigenic site, which was increased from 0 (fully conserved) to 10 (highly variable). Stalk antigenic site-specific antibody increased linearly as epitopic distance was increased from 0 to 5 (over 200% increase) and plateaued when epitopic distance was increased beyond 5 (Fig 5D). Therefore, the epitopic distance between head antigenic sites greatly affects antibody responses to the stalk.

Taken together, epitopic distance increases of the head epitope had the largest effect on stalk antigenic site specific antibody levels after boosting. Although all parameters demonstrated some effect on the stalk-antigenic site specific antibodies, these were modest when compared to the effect of epitopic distance. The decrease in stalk antigenic site specific antibodies when the number of head antigenic sites was increased may lend itself to the still unanswered question in the field of how difference

in the ratio of head to stalk epitopes of HA affects the subdominance of the stalk antigenic site. If indeed the head contains more antigenic sites than the stalk, the model predicts that stalk-antigenic site response will be decreased. It is important to note that this analysis demonstrates stalk-antigenic site-specific antibody truly decreases with the addition of head antigenic sites, and it is not only that stalk-specific antibodies remain constant and only the relative amount compared to the head is changed. It also suggests that the immunologic subdominance of the stalk does not necessarily mean it is inherently less immunogenic, having implications for targeting this domain in universal vaccination.

Predicting Antibody Responses

Although not the primary aim of this work, the fact that our simulations stem from real life virus strains allows us to explore the possibility of using such an algorithm to predict immune responses to real life vaccines. Perfect validation would require specimens from age-matched subjects after vaccination with monovalent CA09 vaccine with documented exposure histories or accurately measured antibody and memory B cell repertoires, but this is not currently possible. Therefore, we attempted to determine if the simulations can be used to accurately predict the increase in stalk-specific antibody and increased cross reactivity seen in the BR07 exposed groups by using an age-stratified cohort under the assumption that those born prior to 1947 were originally exposed to 1918-like strains and those born after 1977 were exposed to the more BR07-like recent strains. Specifically, serum was collected from an age-stratified cohort (ages 18 - 32 or 60+) vaccinated during the 2009-2010 flu season with the monovalent 2009 H1N1 pandemic vaccine (A/California/07/2009) before and 28 days after vaccination. Antibody levels were measured against recombinant HA proteins derived

649 from historical antigens via ELISA. We report the relative change in antibody (d28/d0) in
650 order to account for age-specific differences in basal antibody cross-reactivity.
651 The similarity and differences in the responses of each group was assessed first.
652 Although the sample size for the two groups was limited (n = 8 and n = 9), the 18-32
653 group clustered separately from the 60+ group by hierarchical clustering although this
654 grouping was not exact (Fig 6A). Consistent with our model's findings, cross-reactivity
655 was generally increased in the BR07 representative group except for FM47, NC99, and
656 BR07 in which both groups had similar levels (Fig 6B). Stalk specific antibody
657 responses were measured using chimeric HA proteins that contained an "exotic" HA
658 head but retained the conserved stalk region (cH9.1 and cH6.1, Fig 6B). The BR07
659 group had an increased response to the stalk region compared to the SC18 group for
660 both chimeras, although this was more pronounced in the cH9.1 assay. These findings
661 were consistent with the increase in stalk-specific antibody in the BR07 primed group
662 compared to the SC18 primed group in the model. Although our validation cohort was
663 underpowered, and differences did not reach statistical significance (with the exception
664 of NC99, t-test p=0.049), we found that the qualitative trends of the data match closely
665 with that of the model. This suggests that the model can at least qualitatively predict
666 differences in the cross-reactivity and relative stalk-specific antibody of secondary
667 immune responses.
668
669 **Fig 6 Serum Antibody Levels for Age-Stratified Cohort**
670 (A) Hierarchical clustering of antibody binding measured against recombinant HA
671 proteins using ELISA. (B) Heatmap of ELISA antibody binding data. Data represents the

672 relative change in binding. Data was log transformed and standardized, values represet
673 column z-scores.

674 ## Discussion

675 In the current study, we aimed to understanding theoretically how prior exposure
676 to influenza virus antigens affects the antigenic site specificity of the antibodies elicited
677 by vaccination. This work was an extension of the work originally performed by Smith et
678 al.[37] and the theory of Shape Space originally developed by Perelson et al.[15].
679 Consist with Smith et al. findings, we found that the antigenic relationship between the
680 first and secondary exposure antigens largely affect the specificity of the antibody
681 response. Moreover, during secondary immune responses in the model, antigen was
682 removed from the system more quickly in the group previously exposed to an
683 antigenically similar strain during the primary exposure, consistent with the notion of
684 antibody mediated negative interference[17]. Additionally, the increased antibody
685 response to the CA09 strain in the SC18 exposed group after boosting supports the
686 notion of positive interference, in which antibody responses from preexisting memory B
687 cells are increased. Taken together, our findings support the Antigenic Distance
688 Hypothesis described by Smith et al.[17] .

689 The expansion of the Shape Space based model to include multiple antigenic
690 sites by Chaudhury et al. was a major advancement in use of the model to understand
691 B cell specificity across complex antigens[38]. By incorporating multiple antigenic sites,
692 the model creates competition for antigen between B cells complementary to different
693 antigenic sites on the same antigen. Although Chaudhury et al. modeled a multivalent
694 vaccine, our findings are consistent with their finding that antibody responses to a
695 normally subdominant antigenic site will dominate when the antigenic distance between

head antigens are large. Additionally, the large increase in stalk-specific antibodies in the BR07 group is consistent with reports on universal vaccine development that apply a similar strategy to boost stalk specific antibodies[39,40].

One of the most significant findings of the 2009 pandemic, was the ability of 2009 pandemic vaccine to induce antibodies able to bind antigenically distinct viruses[7,12]. Our model agrees with these findings demonstrating that BR07 primed individuals will have an increased antibody reactivity to 1918-like viruses (CA09, SC18, NJ76) as well as seasonal H1N1 viruses, while SC18 primed individuals will only cross-react to viruses antigenically close to CA09. These findings are consistent with the reports suggesting that original virus exposures, not age, affected the vaccine response to the 2009 vaccine[41,42]. Furthermore, although only slightly different, SC18 antibody titers were higher than CA09 titers after boosting with CA09 in the SC18 group but CA09 titers were higher than BR07 in the BR07 primed group, consistent with the phenomenon known as original antigenic sin[43,44].

Other reports of the immune responses to the 2009 pandemic vaccine can be used to further validate our model. Pre-boost titers of the SC18 primed group were almost 3-fold greater for CA09 than those primed for BR07, similar to what has been reported[45]. Additionally, the fold-change in antibody response to the stalk is consistent with published reports[13]. The Sa antigenic site dominance is the SC18 group is consistent with experimental data showing that antibody responses from the 60+ year old individuals had antibody responses focused on the Sa site of CA09[46]. Furthermore fold change titers (pre-boost/post-boost) were decreased in the SC18 primed group

718 suggesting it is important to take into account priming history of the elderly when trying
719 to assess immunosenescence or predict responses in different age groups[17,47-49].
720 Boosting the cross-reactivity of the antibody response (i.e. the number of strains
721 an immune system has antibodies against) is crucial to the design of universal vaccines.
722 Here we demonstrate that cross-reactivity of secondary immune responses is
723 dependent on priming antigen and therefore different strategies may be required for
724 individuals with different exposure histories. This was clearly demonstrated in the
725 context of pandemic vaccines where more highly cross-reactive antibodies were
726 observed in subjects primed with an A/Hong Kong/97 H5 vaccine and later boosted with
727 an A/Vietnam/04 vaccine, who then subsequently mounted antibody responses
728 recognizing both vaccine strains, as well as a third H5 strain (A/Indonesia/05) not
729 included in either vaccination [50]. This suggests strategies to broaden cross-reactive
730 immunity may be possible with existing vaccine technologies. Although the model does
731 not directly examine susceptibility to infection, it does demonstrate how antigenic
732 distance between heterologous antigenic sites can shift responses to particular
733 conserved antigenic sites leading to increases in cross-reactivity and thus immunity to a
734 greater number of variant influenza stains. Hence, incorporation of a model, such as the
735 one presented here, into the vaccination selection process may allow targeting of
736 vaccine strains specific to the individual in order to produce broadly reactive responses
737 in individuals with different exposure histories.
738 Lastly, the work described here demonstrates the limitations with the current
739 vaccine selection process that relies only on antigenic and phylogenetic distances
740 between strains. Here, the shorter antigenic distance between SC18 and CA09

741 compared to BR07 and CA09 led to two different immune system states. For instance,
742 the SC18 primed group had low titers to US77 after boost with CA09, while the BR07
743 primed group had greater titers. Therefore, although the antigenic distance between
744 CA09 and US77 is fixed, and reflects expected responses from naïve individuals,
745 previously exposed individuals produce antibody responses inconsistent with antigenic
746 distance estimates. Therefore, this suggests that serum samples are not 'impartial
747 observers' of antigenic similarity and they are highly biased by their own immune
748 histories. This is an inherent challenge with the current vaccine approach and highlights
749 the need to take into account prior exposure histories when trying to predict antibody
750 specificities after vaccination.
751
752 Acknowledgments

753 We thank the Center for Integrated Research Computing and the Health
754 Sciences Center for Computational Innovation for computational assistance and
755 resources. Thank you Carrie A. Anderson and Elaine Smolock for help with the
756 manuscript. Thank you Alan Perelson, Martin Zand, Juilee, Thakar, John Treanor, Jim
757 Miller, and Paige Lawrence for supervision of this work. Thank you Derek Smith for
758 helpful comments early on in this project. Thank you Anthony DiPiazza, Matt Brewer,
759 and Patrick McCall for discussions on the project. Funding for this work was supported
760 by the New York Influenza Center of Excellence NIH/NIAID/DMID,
761 HHSN272201400005C.

762 REFERENCES

763 1. Anderson CS, Ortega S, Chaves FA, Clark AM, Yang H, Topham DJ, et al.
764 Natural and directed antigenic drift of the H1 influenza virus hemagglutinin stalk
765 domain. Sci Rep. Nature Publishing Group; 2017;7: 14614. doi:10.1038/s41598-

017-14931-7

2. WILEY DC, SKEHEL JJ. The Structure and Function of the Hemagglutinin Membrane Glycoprotein of Influenza-Virus. Annual Review of Biochemistry. 1987;56: 365–394. doi:10.1146/annurev.biochem.56.1.365

3. Henry Dunand CJ, Leon PE, Kaur K, Tan GS, Zheng N-Y, Andrews S, et al. Preexisting human antibodies neutralize recently emerged H7N9 influenza strains. J Clin Invest. 2015;125: 1255–1268. doi:10.1172/JCI74374

4. Jegaskanda S, Laurie KL, Amarasena TH, Winnall WR, Kramski M, De Rose R, et al. Age-associated cross-reactive antibody-dependent cellular cytotoxicity toward 2009 pandemic influenza A virus subtype H1N1. J Infect Dis. Oxford University Press; 2013;208: 1051–1061. doi:10.1093/infdis/jit294

5. Henry Dunand CJ, Leon PE, Huang M, Choi A, Chromikova V, Ho IY, et al. Both Neutralizing and Non-Neutralizing Human H7N9 Influenza Vaccine-Induced Monoclonal Antibodies Confer Protection. Cell Host and Microbe. 2016;19: 800–813. doi:10.1016/j.chom.2016.05.014

6. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). CELL. 1982;31: 417–427. doi:10.1016/0092-8674(82)90135-0

7. Li G-M, Chiu C, Wrammert J, McCausland M, Andrews SF, Zheng N-Y, et al. Pandemic H1N1 influenza vaccine induces a recall response in humans that favors broadly cross-reactive memory B cells. Proc Natl Acad Sci USA. National Acad Sciences; 2012;109: 9047–9052. doi:10.1073/pnas.1118979109

8. Wrammert J, Koutsonanos D, Li G-M, Edupuganti S, Sui J, Morrissey M, et al. Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. J Exp Med. 2011;208: 181–193. doi:10.1084/jem.20101352

9. Joyce MG, Wheatley AK, Thomas PV, Chuang G-Y, Soto C, Bailer RT, et al. Vaccine-Induced Antibodies that Neutralize Group 1 and Group 2 Influenza A Viruses. CELL. 2016;166: 609–623. doi:10.1016/j.cell.2016.06.043

10. Crotty S, Felgner P, Davies H, Glidewell J, Villarreal L, Ahmed R. Cutting Edge: Long-Term B Cell Memory in Humans after Smallpox Vaccination. The Journal of Immunology. 2003;171: 4969–4973. doi:10.4049/jimmunol.171.10.4969

11. Kurosaki T, Kometani K, Ise W. Memory B cells. Nature Reviews Immunology. 2015;15: 149–159. doi:10.1038/nri3802

12. Wrammert J, Koutsonanos D, Li GM, Edupuganti S, Sui J, Morrissey M, et al. Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. Journal of Experimental Medicine.

2011;208: 181–193. doi:10.1084/jem.20101352

13. Sangster MY, Baer J, Santiago FW, Fitzgerald T, Ilyushina NA, Sundararajan A, et al. B Cell Response and Hemagglutinin Stalk-Reactive Antibody Production in Different Age Cohorts following 2009 H1N1 Influenza Virus Vaccination. Clinical and Vaccine Immunology. 2013;20: 867–876. doi:10.1128/CVI.00735-12

14. Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, et al. Antigenic and Genetic Characteristics of Swine-Origin 2009 A(H1N1) Influenza Viruses Circulating in Humans. Science. American Association for the Advancement of Science; 2009;325: 197–201. doi:10.1126/science.1176225

15. Perelson AS, Oster GF. Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. Journal of Theoretical Biology. 1979;81: 645–670. doi:10.1016/0022-5193(79)90275-3

16. Smith DJ, Forrest S, Hightower RR, Perelson AS. Deriving shape space parameters from immunological data. Journal of Theoretical Biology. 1997;189: 141–150. doi:10.1006/jtbi.1997.0495

17. Smith DJ, Forrest S, Ackley DH, Perelson AS. Variable efficacy of repeated annual influenza vaccination. Proceedings of the National Academy of Sciences. 1999;96: 14001–14006.

18. Chaudhury S, Reifman J, Wallqvist A. Simulation of B cell affinity maturation explains enhanced antibody cross-reactivity induced by the polyvalent malaria vaccine AMA1. J Immunol. American Association of Immunologists; 2014;193: 2073–2086. doi:10.4049/jimmunol.1401054

19. Anderson CS, McCall PR, Stern HA, Yang H, Topham DJ. Antigenic cartography of H1N1 influenza viruses using sequence-based antigenic distance calculation. BMC Bioinformatics. BioMed Central; 2018;19: 51. doi:10.1186/s12859-018-2042-4

20. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). CELL. 1982;31: 417–427. doi:10.1016/0092-8674(82)90135-0

21. Edelman GM. Origins and mechanisms of specificity in clonal selection. Soc Gen Physiol Ser. 1974;29: 1–38.

22. NOSSAL GJV, ADA GL. ANTIGENS AND THE AFFERENT LIMB OF THE IMMUNE RESPONSE. Antigens, Lymphoid Cells and the Immune Response. Elsevier; 1971. pp. 5–23. doi:10.1016/b978-0-12-521950-1.50009-1

23. JERNE NK. Clonal selection in a lymphocyte network. Soc Gen Physiol Ser. 1974;29: 39–48.

24. Kohler G. Frequency of precursor cells against the enzyme beta-galactosidase: an estimate of the BALB/c strain antibody repertoire. Eur J Immunol. WILEY-VCH Verlag GmbH; 1976;6: 340–347. doi:10.1002/eji.1830060507

25. Cunningham AJ. The generation of antibody diversity: a new look. 1976.

26. Clark AM, DeDiego ML, Anderson CS, Wang J, Yang H, Nogales A, et al. Antigenicity of the 2015-2016 seasonal H1N1 human influenza virus HA and NA proteins. PLoS ONE. 2017;12: e0188267. doi:10.1371/journal.pone.0188267

27. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. Journal of Computational Physics. 1976;22: 403–434. doi:10.1016/0021-9991(76)90041-3

28. Gillespie DT. Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry. 1977;81: 2340–2361. doi:10.1021/j100540a008

29. Woo HJ, Reifman J. A quantitative quasispecies theory-based model of virus escape mutation under immune selection. … of the National Academy of Sciences. 2012. doi:10.1073/pnas.1117201109/-/DCSupplemental

30. Eggink D, Goff PH, Palese P. Guiding the Immune Response against Influenza Virus Hemagglutinin toward the Conserved Stalk Domain by Hyperglycosylation of the Globular Head Domain. J Virol. 2013;88: 699–704. doi:10.1128/JVI.02608-13

31. Ellebedy AH, Krammer F, Li GM, Miller MS, Chiu C, Wrammert J, et al. Induction of broadly cross-reactive antibody responses to the influenza HA stem region following H5N1 vaccination in humans. Proceedings of the National Academy of Sciences. 2014;111: 13133–13138. doi:10.1073/pnas.1414070111

32. Ellebedy AH, Krammer F, Li G-M, Miller MS, Chiu C, Wrammert J, et al. Induction of broadly cross-reactive antibody responses to the influenza HA stem region following H5N1 vaccination in humans. Proc Natl Acad Sci USA. 2014;111: 13133–13138. doi:10.1073/pnas.1414070111

33. East IJ, Todd PE, Leach SJ. On topographic antigenic determinants in myoglobins. Mol Immunol. 1980;17: 519–525.

34. Champion AB, Soderberg KL, Wilson AC. Immunological comparison of azurins of known amino acid sequence. Dependence of cross-reactivity upon sequence resemblance. J Mol Evol. 1975;5: 291–305.

35. Nachbagauer R, Choi A, Izikson R, Cox MM, Palese P, Krammer F. Age Dependence and Isotype Specificity of Influenza Virus Hemagglutinin Stalk-Reactive Antibodies in Humans. MBio. American Society for Microbiology; 2016;7: e01996–15. doi:10.1128/mBio.01996-15

875 36. Lee AJ, Das SR, Wang W, Fitzgerald T, Pickett BE, Aevermann BD, et al.
876 Diversifying Selection Analysis Predicts Antigenic Evolution of 2009 Pandemic
877 H1N1 Influenza A Virus in Humans. J Virol. 2015;89: 5427–5440.
878 doi:10.1128/JVI.03636-14

879 37. Smith DJ. Derek J. Smith Thesis Dissertation. 1997. pp. 1–124.

880 38. Chaudhury S, Reifman J, Wallqvist A. Simulation of B Cell Affinity Maturation
881 Explains Enhanced Antibody Cross-Reactivity Induced by the Polyvalent Malaria
882 Vaccine AMA1. The Journal of Immunology. 2014;193: 2073–2086.
883 doi:10.4049/jimmunol.1401054

884 39. Pica N, Palese P. Toward a Universal Influenza Virus Vaccine: Prospects and
885 Challenges. Annu Rev Med. 2013;64: 189–202. doi:10.1146/annurev-med-
886 120611-145115

887 40. Krammer F, García-Sastre A, Palese P. Is It Possible to Develop a “Universal”
888 Influenza Virus Vaccine? Toward a Universal Influenza Virus Vaccine: Potential
889 Target Antigens and Critical Aspects for Vaccine Development. Cold Spring
890 Harbor Perspectives in Biology. 2017. doi:10.1101/cshperspect.a028845

891 41. Li Y, Myers JL, Bostick DL, Sullivan CB, Madara J, Linderman SL, et al. Immune
892 history shapes specificity of pandemic H1N1 influenza antibody responses. J Exp
893 Med. 2013;210: 1493–1500. doi:10.1084/jem.20130212

894 42. Andrews SF, Huang Y, Kaur K, Popova LI, Ho IY, Pauli NT, et al. Immune history
895 profoundly affects broadly protective B cell responses to influenza. Science
896 Translational Medicine. American Association for the Advancement of Science;
897 2015;7: –316ra192. doi:10.1126/scitranslmed.aad0522

898 43. Lessler J, Riley S, Read JM, Wang S, Zhu H, Smith GJD, et al. Evidence for
899 Antigenic Seniority in Influenza A (H3N2) Antibody Responses in Southern China.
900 Basler CF, editor. PLoS Pathog. 2012;8: e1002802–11.
901 doi:10.1371/journal.ppat.1002802

902 44. Francis T. On the doctrine of original antigenic sin. 1960. doi:10.2307/985534

903 45. Hancock K, Veguilla V, Lu X, Zhong W, Butler EN, Sun H, et al. Cross-reactive
904 antibody responses to the 2009 pandemic H1N1 influenza virus. N Engl J Med.
905 2009;361: 1945–1952. doi:10.1056/NEJMoa0906453

906 46. Krause JC, Tumpey TM, Huffman CJ, McGraw PA, Pearce MB, Tsibane T, et al.
907 Naturally occurring human monoclonal antibodies neutralize both 1918 and 2009
908 pandemic influenza A (H1N1) viruses. J Virol. 2010;84: 3127–3130.
909 doi:10.1128/JVI.02184-09

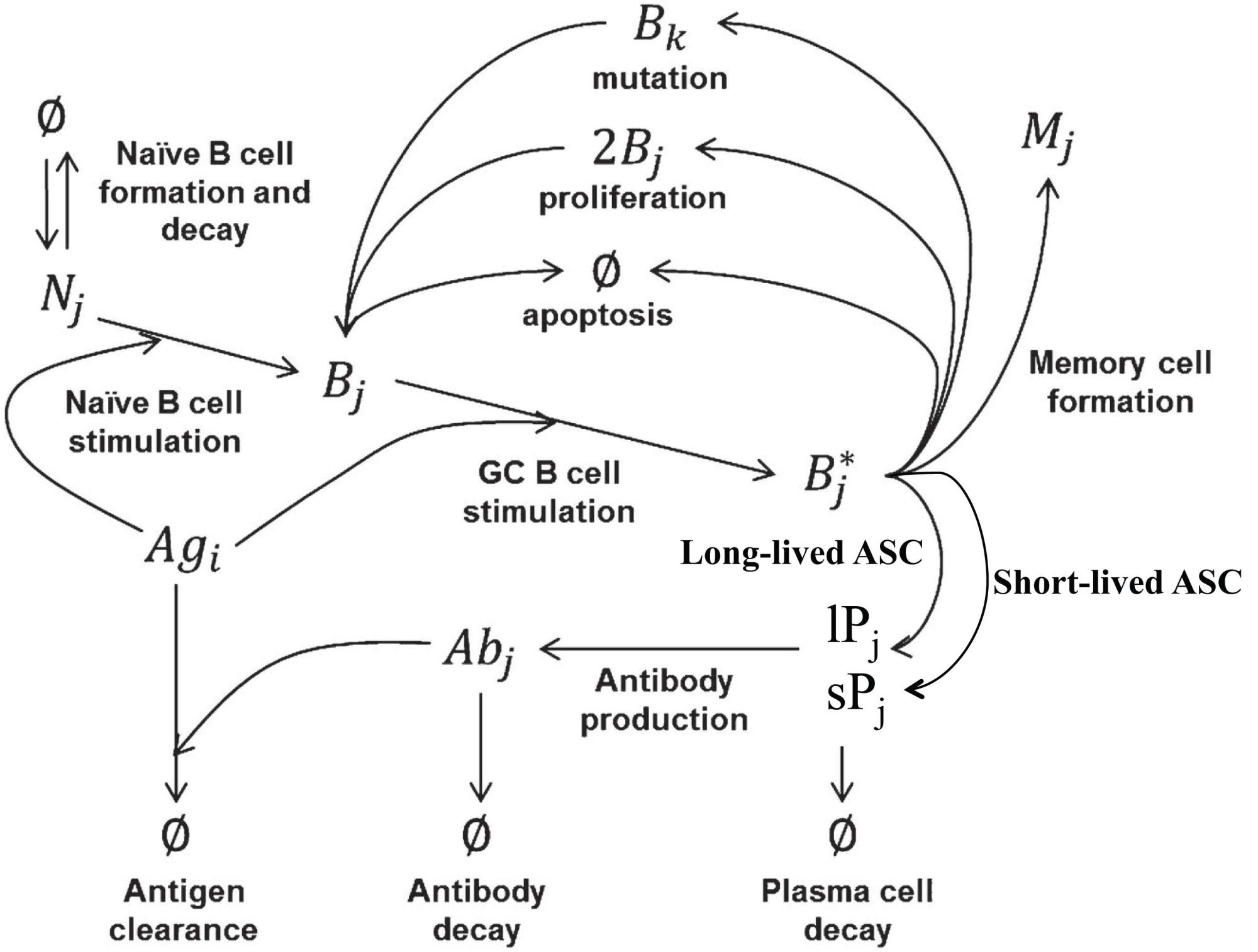
910 47. Thomas Francis J. On the Doctrine of Original Antigenic Sin. Proceedings of the
911 American Philosophical Society. 2012;: 1–8.

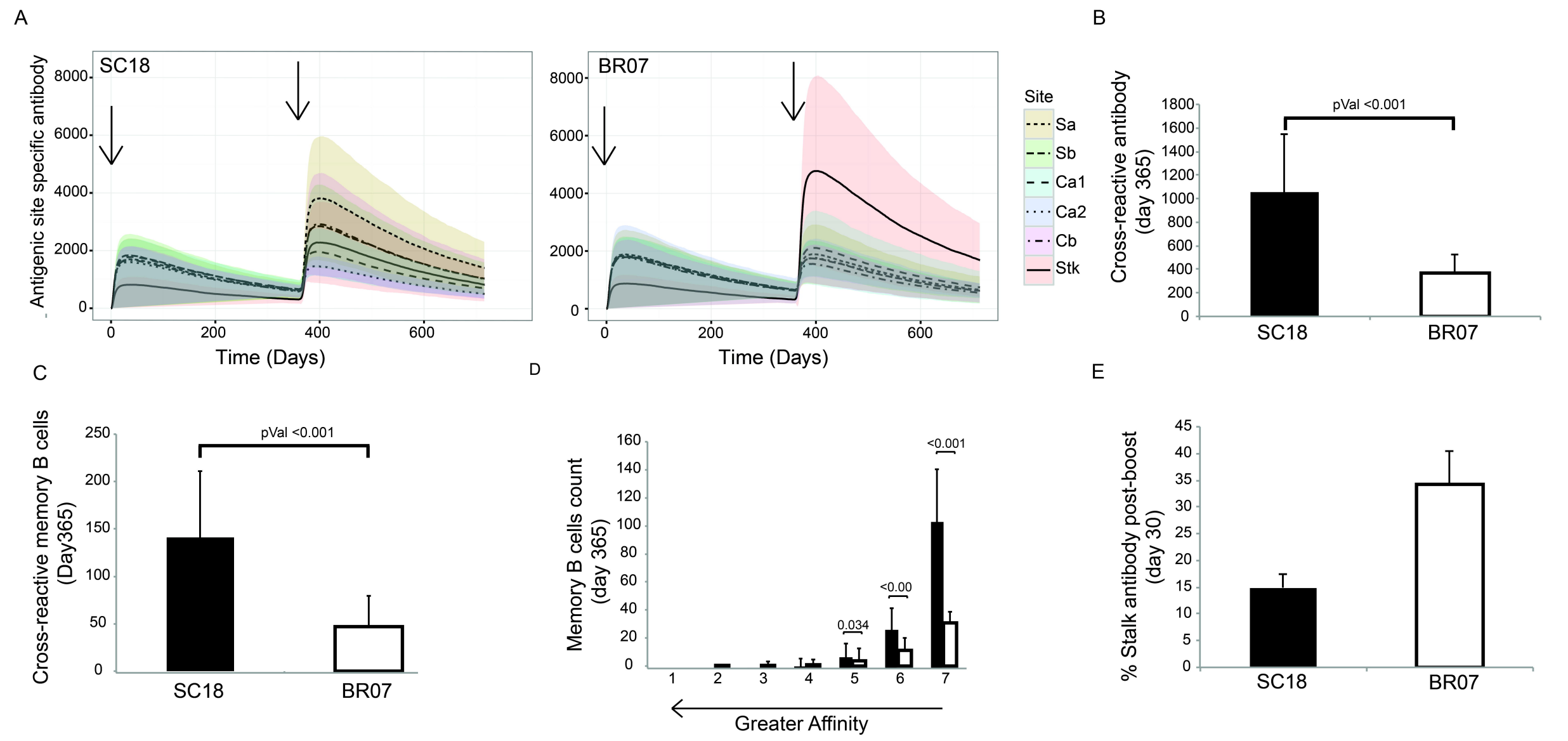
912 48. Fonville JM, Wilks SH, James SL, Fox A, Ventresca M, Aban M, et al. Antibody
913 landscapes after influenza virus infection or vaccination. Science. 2014;346: 996–
914 1000. doi:10.1126/science.1256427

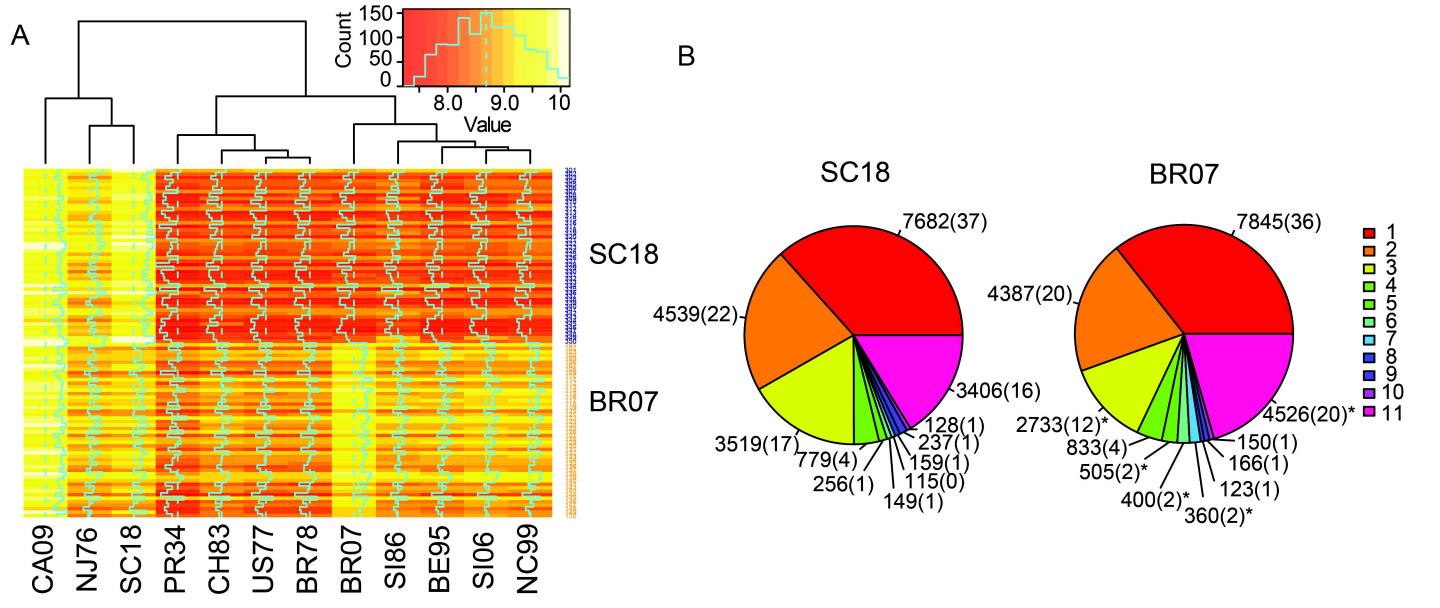
915 49. Höpping AM, McElhaney J, Fonville JM, Powers DC, Beyer WEP, Smith DJ. The
916 confounded effects of age and exposure history in response to influenza
917 vaccination. Vaccine. 2016;34: 540–546. doi:10.1016/j.vaccine.2015.11.058

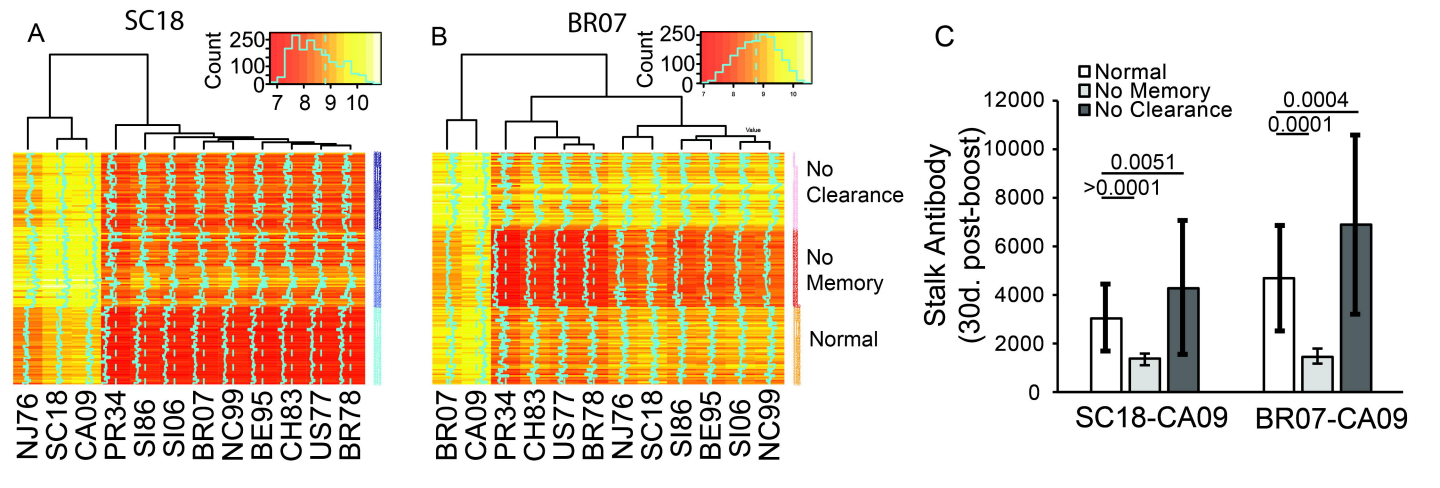
918 50. Goji NA, Nolan C, Hill H, Wolff M, Noah DL, Williams TB, et al. Immune
919 responses of healthy subjects to a single dose of intramuscular inactivated
920 influenza A/Vietnam/1203/2004 (H5N1) vaccine after priming with an antigenic
921 variant. Journal of Infectious Diseases. 2008;198: 635–641. doi:10.1086/590916

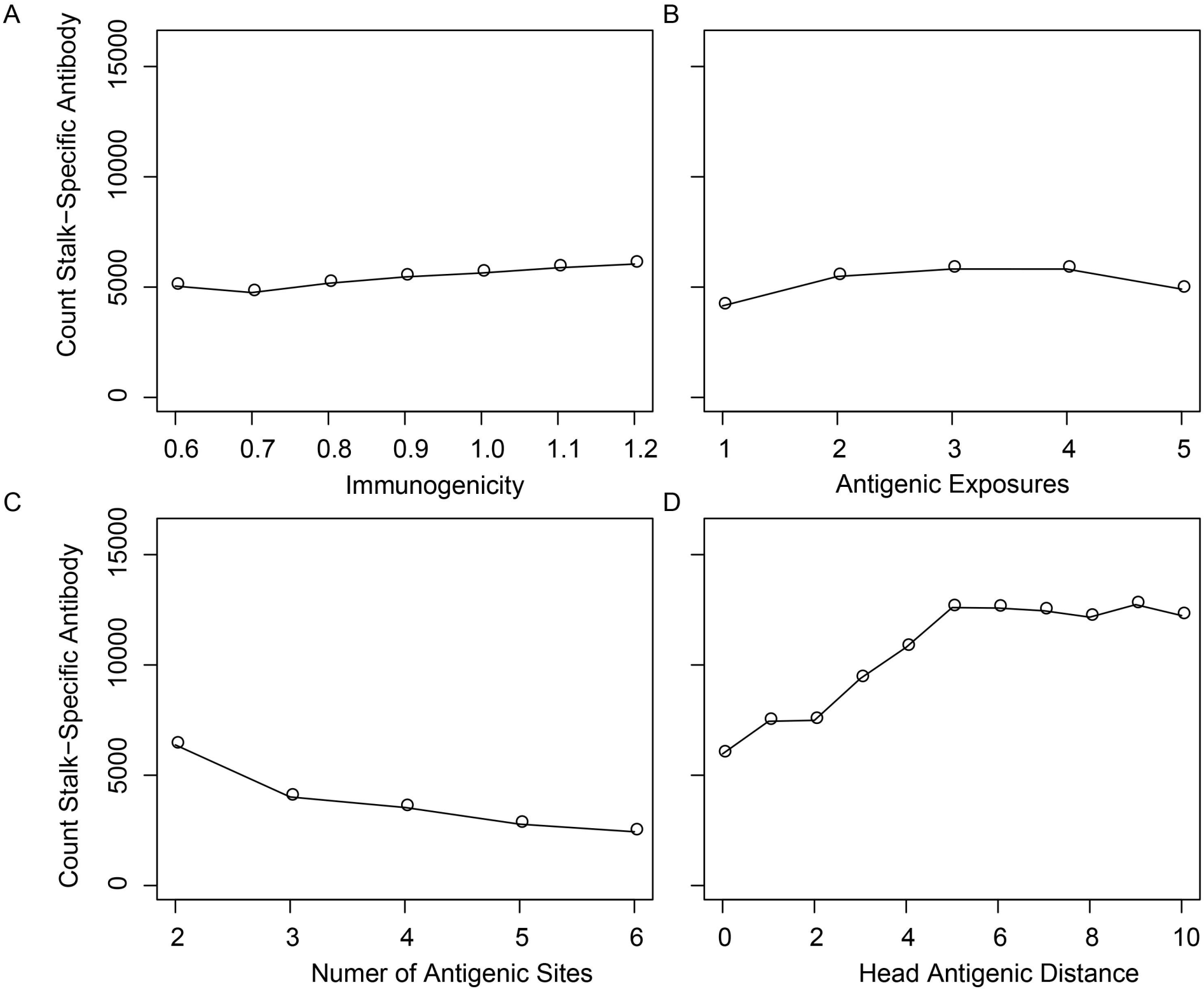
922

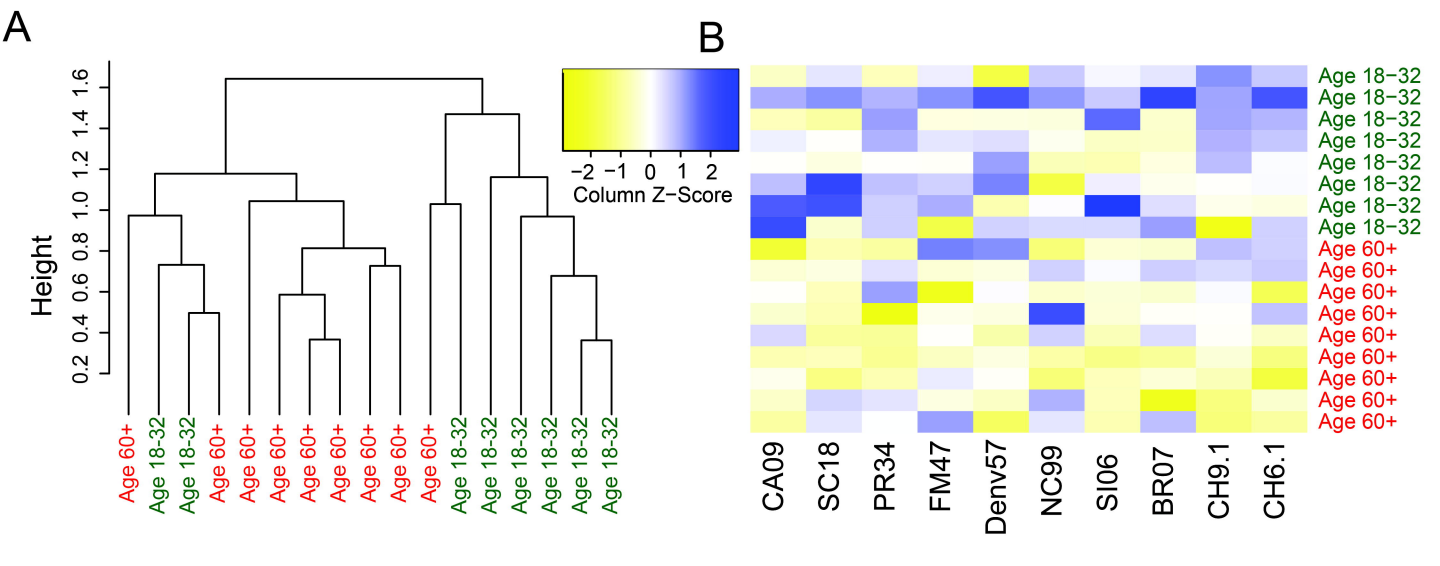

B_k
mutation
$2B_j$
proliferation
M_j
$\emptyset$
Naïve B cell formation and decay
N_j
$\emptyset$
apoptosis
Memory cell formation
B_j
Naïve B cell stimulation
GC B cell stimulation
B_j^*
Ag_i
Long-lived ASC
Short-lived ASC
lP_j
sP_j
Ab_j
Antibody production
$\emptyset$
$\emptyset$
$\emptyset$
Antigen clearance
Antibody decay
Plasma cell decay


A
SC18
BR07
Antigenic site specific antibody
Time (Days)
Site
Sa
Sb
Ca1
Ca2
Cb
Stk
B
Cross-reactive antibody (day 365)
pVal <0.001
SC18
BR07
C
Cross-reactive memory B cells (Day365)
pVal <0.001
SC18
BR07
D
Memory B cells count (day 365)
0.034
<0.00
<0.001
Greater Affinity
E
% Stalk antibody post-boost (day 30)
SC18
BR07


A
Count
150
100
50
0
8.0
9.0
10
Value
SC18
BR07
CA09
NJ76
SC18
PR34
CH83
US77
BR78
BR07
SI86
BE95
SI06
NC99
B
SC18
7682(37)
4539(22)
3406(16)
3519(17)
128(1)
237(1)
779(4)
159(1)
256(1)
115(0)
149(1)
BR07
7845(36)
4387(20)
4526(20)*
2733(12)*
150(1)
833(4)
166(1)
505(2)*
123(1)
400(2)*
360(2)*
1
2
3
4
5
6
7
8
9
10
11


A
SC18
Count
250
100
0
7 8 9 10
NJ76
SC18
CA09
PR34
SI86
SI06
BR07
NC99
BE95
CH83
US77
BR78
B
BR07
Count
250
100
0
Value
No Clearance
No Memory
Normal
BR07
CA09
PR34
CH83
US77
BR78
NJ76
SC18
SI86
BE95
SI06
NC99
C
Normal
No Memory
No Clearance
Stalk Antibody (30d. post-boost)
12000
10000
8000
6000
4000
2000
0
0.0051
>0.0001
0.0004
0.0001
SC18-CA09
BR07-CA09


A
Count Stalk−Specific Antibody
0
5000
10000
15000
0.6
0.7
0.8
0.9
1.0
1.1
1.2
Immunogenicity
B
1
2
3
4
5
Antigenic Exposures
C
Count Stalk−Specific Antibody
0
5000
10000
15000
2
3
4
5
6
Numer of Antigenic Sites
D
0
2
4
6
8
10
Head Antigenic Distance


A
B
Height
0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6
−2 −1 0 1 2
Column Z−Score
Age 60+
Age 18−32
Age 18−32
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 18−32
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
Age 60+
CA09
SC18
PR34
FM47
Denv57
NC99
SI06
BR07
CH9.1
CH6.1