FRONT MATTER

Title: Polyvalent immunization elicits a synergistic broadly neutralizing immune response to hypervariable region 1 variants of hepatitis C virus

Authors

Alexander I. Mosa[1]*† & David S. Campo[2]†**, Yury Khudyakov[2**], Mounir G. AbouHaidar[3], Adam J. Gehring[4], Atif Zahoor[1], Jonathan K. Ball[5], Richard A. Urbanowicz[6], and Jordan J. Feld[1]

Affiliations

[1]Toronto General Hospital, University Health Network; Toronto, Canada
[2]Molecular Epidemiology & Bioinformatics, CDC; Atlanta, Georgia
[3]Department of Cell & Systems Biology, University of Toronto; Toronto, Canada
[4]Department of Immunology, University of Toronto; Toronto, Canada
[5]Wolfson Centre for Global Virus Infections, University of Nottingham; Nottingham, UK
[6]Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool; Liverpool, UK
*Corresponding author email: alexander.mosa@uhnresearch.ca
† These authors contributed equally to this work.

Abstract

A hepatitis C virus (HCV) vaccine is urgently needed. Vaccine development has been hindered by HCV's genetic diversity, particularly within the immunodominant hypervariable region 1 (HVR1). Here, we developed a new strategy to elicit broadly neutralizing antibodies to HVR1, which had previously been considered infeasible. We first applied a novel information theory-based measure of genetic distance to evaluate phenotypic relatedness between HVR1 variants. These distances were used to model the structure of HVR1's sequence space, which was found to have five major clusters. Variants from each cluster were used to immunize mice individually, and as a pentavalent mixture. Sera obtained following immunization neutralized every variant in a diverse HCVpp panel (n=10), including those resistant to monovalent immunization, and at higher mean titers ($1/ID_{50}$=435) than a glycoprotein E2 ($1/ID_{50}$=205) vaccine. This synergistic immune response offers a novel approach to overcoming antigenic variability and may be applicable to other highly mutable viruses.

Significance

HCV infects 58 million people worldwide, with an estimated 1.5 million new infections annually. Despite remarkable advances in treatment, new infections outpace cures, and a prophylactic vaccine is needed to achieve HCV elimination. Vaccine development has been impeded, however, by the extreme genetic variability of HCV. Our previous work indicates that even in highly variable epitopes HCV is limited by phenotypic constraints that can be exploited by rationale vaccine design. Here we applied a novel measure of genetic distance to model these constraints into a network describing the global HCV sequence space. By combining variants from across this space to immunize mice, we elicited broadly neutralizing antibodies with greater neutralization breadth and potency than a classical vaccine candidate.

MAIN TEXT

Introduction

Hepatitis C is a leading cause of morbidity and mortality from liver disease worldwide[1]. The introduction of curative direct-acting antivirals spurred hopes for global HCV elimination[2]. However, with an estimated 1.5 million new infections and 300,000 deaths annually, it may be challenging to achieve the World Health Organization's 2030 elimination targets with treatment alone[3]. Availability of an effective HCV vaccine would significantly aid in these efforts[4].

Vaccine development has been impeded, however, by the extreme genetic variability of HCV, which renders immune responses produced against one variant ineffective against others[5,6]. Though classified at the full genomic level into eight genotypes differing at 30-35% of nucleotide positions, HCV's heterogeneity is not distributed uniformly along the genome[7]. The most heterogeneous region, Hypervariable Region 1 (HVR1), encodes the N-terminal 27 amino acid (aa) portion of the envelope protein E2[8]. Though HVR1 contains an immunodominant neutralizing epitope, mediates interactions with the HCV co-receptor Scavenger Receptor class B type 1 (SRB1), and is strongly positively selected in natural infection, its application to vaccine development has been limited due to its extraordinary genetic variability[9–12]. Thus, despite the capacity of anti-HVR1 antibodies to prevent homologous infection, and the favorable accessibility of this epitope to neutralizing antibodies, vaccine efforts have been focused on eliciting antibodies to conserved regions outside of HVR1[13,14]. However, even conserved regions seem to be affected by HVR1, which physically shields conserved neutralizing epitopes, modulates envelope conformation, and elicits strain-specific, dominant "decoy" immune responses, thus suppressing recognition of the conserved subdominant epitopes[15–17]. Simply removing HVR1 from E2 did not improve responses following vaccination, but instead was inferior to native E2 in terms of neutralization, possibly related to conformational changes in E2 caused by the HVR1 excision or by disruption of discontinuous antigenic epitopes involving HVR1[17–19].

The role of HVR1 in HCV neutralization, both as a dominant epitope and as a modifier of the response to conserved epitopes, must therefore be considered in the design of any HCV vaccine. Here, we describe a novel strategy to overcome the challenge of HVR1 heterogeneity. Using a novel information theory-based distance, we modelled HVR1 genetic variability and observed discrete, genotype-independent clusters. We selected 5 central sequences from these clusters to synthesize peptides for vaccination. The mixture of HVR1 variants resulted in an antibody response that was more broadly neutralizing than each individual variant or pooled sera, indicating a synergistic interaction among immune responses to related, but distinct, HVR1variants. These findings open a new path for the development of an HCV vaccine using sequence-complementary variants of genetically divergent HVR1 antigenic epitopes.

Results

Selection of Genetic Distance Relevant to Cross-Immunoreactivity

To identify HVR1 variants for immunization experiments, we modelled HVR1's genetic space, with the hypothesis that the space structure could inform variant selection and thus improve coverage. First, we explored how different measures of genetic distance were associated with a previously published cross-immunoreactivity dataset of 26,883 pairwise reactions among 262 HVR1 variants[20] (Fig 1A). We compared the mean distance observed in pairs that did not cross-react, with the mean distance observed in pairs that did cross-react. If the ratio is 1, then the distance is not helping us to differentiate the two types of pairs, but the greater the ratio, the greater the relevance of the distance to cross-immunoreactivity. The ratio calculated using distances based on individual or joint physiochemical properties[21], Hamming distances (number of mismatches; ratio =1.19; t-test, p = 1.3668E-279) or the BLOSUM62 scores (ratio =1.17; t-test, p =5.5753E-270) showed very similar results, all indicating low association with cross-immunoreactivity (Fig 1B).

Considering importance of coordinated substitutions in HCV evolution[22], we devised a novel information-theory-based distance called MIH (Mutual Information and entropy, H)[23]. The MIH distance considers the variability of each position (measured by entropy) and the existence of coordinated substitutions between position (measured by mutual information among position pairs) (Fig 1C). The mean MIH of non-cross-reactive pairs was 1.89 times higher than the mean of cross-reactive pairs (t-test, p = 8.88E-56), a ratio 58.7% greater than the second best, obtained with the Hamming distance. These results indicate that the MIH distance has a higher association with cross-immunoreactivity (Fig. 1B) and thus HVR1 variants with lower average MIH distance to other variants, are also more likely to be broadly cross-immunoreactive.

HVR1 Sequence Space

There are 969 unique C-terminal 8aa sequences in the extended global dataset of 12,245 HVR1 sequences. Out of all HVR1 sequences published with known HCV genotype (n=1,305), 8.6% were found in two or more genotypes, with one being found in all six of them. This result indicates that this small region is genetically convergent at the genotype level.

We then proceeded to measure the MIH distance among every pair of 8aa sequences. This matrix of distances was used to build a k-step network (Fig 2), which is equivalent to the union of all Minimum Spanning Trees and allows one to visualize the distances among all variants present. Thus, the network constitutes our model of the HVR1 sequence space, which we used to find modules and measure the centrality of each variant.

Given that early acute-phase variants (also referred to as Transmitted-Founder variants), are plausible targets for vaccine development, as they are the first variants encountered by the immune system[24], we studied their location in the HVR1 network. First, the eigenvector centrality of each HVR1 variant in the k-step network was established. Second, the average centrality of variants collected during acute (n=119) or chronic (n=251) infection was calculated. The acute HVR1 variants had an average centrality (9.50E-04) that was 9.73 times higher (t-test, p = 0.0077) than the average of chronic variants (9.77E-05). These results indicate that average MIH distance between acute and other variants in the network is significantly reduced relative to chronic variants and thus suggests that acute variants are more likely than chronic to be cross-immunoreactive. In addition, the acute variants were not locally confined but were found globally distributed across the network and independent of HCV genotype. This indicates acute HVR1 variants, owing to their broad spread in the HVR1 genetic space, may possess complementary cross-immunoreactivities, which if combined, may provide broad cross-immunoreactivity leading to broad neutralization.

Selection of HVR1 variants for immunization

To discover the combination of variants most likely to possess complementary cross-immunoreactivities, we evaluated if the HVR1 network contained modules or clusters, with the hypothesis that each cluster would correspond to distinct HVR1 sub-phenotypes. The distribution of all pairwise MIH distances showed a bimodal distribution, suggesting the existence of modules (Fig. 3A). In contrast, distribution of the Hamming or physicochemical distances was unimodal. The bimodal distribution of MIH distances is an indication of clusters in the HVR1 space modeled, which leads to two types of distances: (i) smaller within-cluster distances and (ii) larger between-cluster distances. The modular organization of the MIH-based network suggests that a combination of HVR1 variants selected from each module may be capable of inducing immune responses covering the entire space. Thus, we created modularity-maximizing partitions between 2 to 40 modules. We identified the five-module solution as the best one, given that it showed the highest difference between average within-module distances and the distance obtained by random partitioning of the same size (Fig 3B). All HCV genotypes were scattered across clusters: genotypes 1 (n=1082), 2 (n=40), 4 (n=27) and 6 (n=59) were present in all 5 clusters; whereas genotypes 3 (n=88) and 5 (n= 9) were present in 4 clusters. Within genotype 1, those belonging to subtype 1a (n=774) were present in all 5 clusters, with those belonging to 1b (n=302) also present in all 5 clusters (Fig 2) . Finally, we identified the most central acute-phase variant in each of the five modules and selected them as immunogens for synthesis **(**Fig 3C**).**

Immunogens Elicit Cross-Reactive Antibodies

To evaluate if our candidate peptides were immunogenic, six groups of Balb/c mice (n=3 per group) were immunized with each of the peptides individually (monovalent) or combined (pentavalent), and terminally bled to characterize humoral responses (Fig 4A**).** Both monovalent and pentavalent formulations elicited high-titer (1:25,000) peptide-specific antibodies following immunization, with higher reactivity observed with the pentavalent sera at the lowest dilution tested (t-test, p=0.003; Fig 4B). Sera from mock immunized mice (adjuvant + PBS) were not reactive at any dilution tested (Fig 4B). A concern in multivalent formulations is diminished reactivity to each of the individual constituent immunogens. We therefore evaluated monovalent immunogenicity, based on self-reactivity, in comparison to the reactivity of the pentavalent immunized sera. Though we observed intrinsic differences in the antigenicity and

immunogenicity of the monovalent immunogens, self-reactivity following pentavalent immunization was not inferior (Fig 4C). Next, using competitive ELISA, we evaluated if antibodies elicited by pentavalent immunization targeted the C-terminal neutralizing epitope of HVR1. We observed significant binding inhibition when sera were pre-incubated with a peptide fragment comprising the C-terminal eight amino acids, suggesting antibodies elicited by the pentavalent formulation predominantly, though not exclusively, target the C-terminus of HVR1 **(**Fig 4D**).**

Next, we sought to characterize heterologous cross-reactivity using a panel of HVR1 peptides (Fig 5) representing global genetic diversity. This was based on prior work to develop a standardized panel of HCV variants representing all major global genotypes, 1a intra-genotypic diversity, and the spectrum of neutralization resistance[25,26]. The sub-panel we selected was enriched for highly neutralization-resistant variants maximally differing in genetic distance from our vaccine immunogens (50-87.5% sequence divergence) (Fig 5A**)**. By ELISA, we observed universal cross-reactivity of pentavalent sera with the panel of HVR1 peptides **(**Fig 5B**).** No correlations between cross-reactivity and either HVR1 genotype or genetic distance to the pentavalent immunogens were observed. These findings indicate that pentavalent immunization elicited broadly cross-reactive antibodies targeting the neutralizing epitope containing HVR1 C-terminus.

Pentavalent Immunogen Elicits Broadly Neutralizing Antibodies

Our previous experiments demonstrated cross-reactivity to genetically diverse HVR1 peptides. Cross-reactivity is necessary but not sufficient for viral neutralization. We therefore sought to characterize the protective breadth of the antibodies elicited by pentavalent immunization using HCV pseudoparticles (HCVpp). Briefly, for each HCV variant in our panel, HCVpp were generated, and residual infectivity in the presence of serial dilutions of mouse sera were used to calculate proportion neutralization and $1/ID_{50}$. We observed potent, universal neutralization across the HCVpp panel (Fig 6A). Even highly neutralization-resistant variants, such as UKNP3.1.2, which are almost completely resistant to neutralization by patient-derived sera[26], were potently neutralized by pentavalent sera ($1/ID_{50}$=1,280). Further, compared to a derivative of a gpE1/E2 vaccine entering clinical trials, neutralization potency against UKNP3.1.2 was more than 10-fold higher, with average heterologous neutralization across the entire panel 2.32-fold higher (t-test, p=0.021; Fig 6B). We found no relationship between sequence divergence from the pentavalent immunogens and neutralization resistance, with HCVpp UKNP1.17.1, which has the greatest Hamming distance from any immunogen in the formulation, potently neutralized ($1/ID_{50}$=817; Fig 6C**).** Collectively, these findings suggest that the antibodies elicited by the pentavalent formulation can potently neutralize even extremely genetically distant variants, with no escape detected for any HCVpp in this antigenically diverse panel.

Pentavalent Neutralization Breadth Exceeds Monovalent Constituents

Next, we evaluated if pentavalent immunization elicited antibodies that could neutralize variants resistant to monovalent immunization. Interestingly, not only was pentavalent neutralization potency against the panel greater than average monovalent potency, but variants completely resistant to neutralization by every monovalent preparation were potently neutralized by pentavalent sera (UKNP1.7.1 and UKNP2.4.1). Across the panel, average pentavalent potency was 3.93-fold greater (t-test, p=0.009) than monovalent potency ($1/ID_{50}$=111), and for eight of the ten variants, was significantly greater than the most potent monovalent against each variant (Fig 7A**).** We also compared the neutralization capacity of sera obtained following pentavalent immunization to sera obtained by sequentially immunizing mice with the same monovalent immunogens. Neutralization was not improved by sequentially administering the monovalent immunogens (mean $1/ID_{50}$=99), and was inferior to simultaneous (pentavalent) immunization (t-test, p=0.004; Fig 7B), indicating that not only the valency, but also the method of immunization influences the humoral response and the potential for synergy. These findings suggest that a qualitatively distinct humoral response, rather than a summation of monovalent polyclonal responses, is operative in the broad neutralization observed following pentavalent immunization.

Discussion

Vaccines are one of the most efficient public health tools to control infectious disease in human populations[27]. However, development of vaccines to highly mutable viruses such as HIV, influenza virus, and HCV is greatly impeded by the genetic variability of dominant epitopes, immune responses against which are largely strain-specific, lacking the breadth of cross-immunoreactivity required for protection against a vast swarm of viral variants[28]. HCV's HVR1 is a well-characterized example of a variable region eliciting only narrowly neutralizing antibodies following natural infection or vaccination[29]. Here we present data suggesting a new strategy, based on a novel model of the HVR1 genetic space, for designing complementary formulations of HVR1 antigens capable of directing the immune response to conserved epitopes within a variable sequence region. We show that immunization of mice with a mixture of HVR1 variants selected from each of the five genetic modules of the space produces antibodies demonstrating broad, potent, and superior neutralization activity. This strategy is distinct from past vaccine approaches to variable viruses, which have attempted to direct immune responses to conserved epitopes[13,14,18]. Though targeting conserved epitopes is a rational approach to addressing antigenic variability, the limitations of this approach are evident in the natural history of HCV infection. Not only can conserved epitopes directly evolve to evade immune pressure, but diversifying selection on HVR1 persists even in the presence of conserved epitope targeting antibodies[6,30,31]. This suggests that HVR1 can evolve to attenuate the neutralizing potency of not only HVR1-specific antibodies, but antibodies targeting other epitopes on the virion, which is mechanistically consistent with findings that HVR1 modulates the accessibility of conserved regions[32]. That the pentavalent candidate reported here neutralized a panel of highly neutralization-resistant, highly diverse HCV variants, suggests that a reappraisal of the role of variable epitopes in vaccine design is warranted, especially when their genetic space indicates the presence of functional constraints bounding variability.

Considering the proximity of HVR1 to the E2 receptor binding sites, the major function constraining the HVR1 genetic space is likely related to transmission and receptor binding. Indeed, HVR1 was shown to affect HCV infectivity by contributing to the optimal composition of virions and membrane fusion[15]. It is a critical region for interaction between E2 and Scavenger Receptor class B type I (SR-BI)[33–35]. HVR1 sequence diversity also has a role in protecting conserved epitopes against antibody binding[36], and thus affecting global HCV neutralization sensitivity. These results add credence to the approach presented here based on generating antibody with a broad HVR1 specificity to protect against HCV infection rather than targeting conserved epitopes exposed only in the neutralization-sensitive conformational state of the HCV envelope. If the HVR1 genetic space is largely shaped by balancing a single important function like transmissibility, with the diversifying selection of host immune pressure, there should be common structural features maintained by patterns of coordinated substitutions that permit immune evasion without compromising infectivity. Conservation of HVR1 size, physiochemical invariance, and extensive epistasis (ie coordinated substitutions) within HVR1 and between HVR1 and other positions in E2, support the existence of fitness-constrained structural features[22,37]. It is reasonable to expect that such conserved structural features, if properly presented to B-cells as antigenic epitopes, would elicit broadly neutralizing antibodies despite marked sequence divergence.

It is not clear what determines the differential presentation of these conserved epitopes among HVR1 variants. It is also unknown what determines cross-reactivity between any two HVR1 variants. The HCV HVR1 genetic space is vast. It is conceivable, though, that sufficiently broad immune responses covering this space may be produced by a combination of HVR1 variants, each eliciting antibody with complementary immune reactivity. Criteria for rational selection of such complementary HVR1 variants are not clear because there are no data associating breadth of immunoreactivity with HVR1 sequences in humans. Among many strategies, we considered a random selection and identification of the most genetically distant variants. However, these criteria do not inform how many such variants need to be selected or how to achieve adequate, or ideally, complete coverage. Another potential criterion is representation of all HCV genotypes. Although this seems to be clearer for identifying the number of potential variants and the space to be covered, antigenic convergence among HVR1 variants from different genotypes and subtypes[20] complicate the selection of the initial HVR1 variants for the study because there are no distinct genotype-specific clusters of HVR1 sequences. Here, we developed a different selection strategy for heuristic identification of complimentary HVR1 variants. It is based on using the MIH distance, which we recently showed to better approximate phenotypic distances in both *in silico* and *in*

vitro datasets[23]. We found that while simple sequence similarity (Hamming distance) could moderately discriminate between cross-reactive pairs, the novel MIH distance was markedly superior. This result is particularly important as it indicates that the distance captures the well-known fact that not all substitutions are equivalent[21], and that the more radical the substitution, measured by capacity to increase MIH distance, the more likely it will abrogate cross-immunoreactivity. When we explored the structure of the HVR1 sequence space using MIH, the network was found to be pentamodular, indicating that the structural features defining breadth of immunoreactivity, and mutual reactivity between any two variants, are distributed across 5 major HVR1 clusters. Acute-phase variants were identified in all 5 clusters and were found to occupy positions of centrality within each module suggesting that founder viruses have a greater breadth of cross-immunoreactivity within each cluster than chronic phase variants. This finding is in concert with the observation that early acute-phase variants, referred to as Transmitted-Founder variants, possess distinct, transmissibility enhancing phenotypes, and occupy central positions within the sequence space, affording greater mutational robustness from which to diversify once infection is established[38,39]. It is important that the acute HVR1 variants are not locally confined but are distributed across the k-step network, entirely independent of HCV genotype, as this indicates the existence of multiple Transmitted-Founder phenotypes, which must all be neutralized by a putative HCV vaccine.

The identification of 5 clusters in the HVR1 MIH genetic space was used to guide selection of 5 HVR1 sequences for peptide synthesis and immunization. We hypothesized that peptides from these clusters could elicit cluster-neutralizing immune responses, the summation of which should provide responses covering the majority of HCV strains. Immunization with a mixture of all five peptides was used to show that the same breadth of neutralization could be maintained after immunization with all peptides simultaneously. We found, however, that the combination of antibody against individual peptides and a mixture of sera from mice immunized with these peptides did not neutralize 2 HCVpp variants, whereas the pentavalent immunization did neutralize all the HCVpp tested. This synergy of pentavalency was not expected but is an important observation.

This suggests that to achieve a universal broad neutralization, the divergent HVR1 immunogens may need to be simultaneously presented. Sequential exposure to each cluster may instead successively direct maturation to module-specific features, limiting breadth of reactivity. This may explain why neutralization breadth and potency observed following sequential immunization with the five HVR1 peptides was inferior, and why chronic infection, with its repeated targeting of immunodominant epitopes followed by escape, does not produce the breadth of neutralization observed following pentavalent immunization[30,31]. The importance of simultaneous presentation is also supported by the finding that antibodies elicited by pentavalent immunization neutralized variants resistant to monovalent immunization. This synergistic interaction indicates that although the HVR1 variants selected for immunization were genetically distant, and occupied distinct modules, they shared the neutralizing epitope.

Polyvalent vaccines are a common strategy to broaden immune responses, and have been in application for decades[40]. The rationale is the expectation that each antigenic variant will be immunologically complementary, and thereby incrementally broaden coverage. However, simply combining antigenic variants is not sufficient for eliciting synergy. For example, despite application of polyvalent vaccines for decades, concerns of reduced protection[41] and antibody-dependent enhancement of disease[42], rather than synergy, predominate. The 50-valent inactivated rhinovirus vaccine is immunogenic only against one-third of circulating human rhinoviral types[43], suggesting that a mere blending of serologically distinct antigenic variants does not result in apparent synergy with certainty. Another example of the complex nature of synergy was reported by Lange et al[44], who showed that immunization with a mixture of 4 HVR1 variants increased strength of immune response compared to immunization with the individual variants, but without a detectable increase in breadth of neutralization. All four HVR1 sequences in that study belonged to HCV genotype 1b. Interestingly, we found that 3 of the variants they evaluated come from a single MIH cluster. This observation indicates that either a highly disproportionate presentation of clusters resulted in a limited synergistic effect focused on the dominant cluster, or presentation of variants from only two clusters is insufficient for increasing breadth of neutralization.

We did not evaluate multi-valency other than pentavalency but believe that all 5 antigens are important. Notably, Fig.4 shows that epitopes of M1 and M2 presented in peptides adsorbed on the solid phase are poorly recognized by antibodies to all HVR1 variants, including M1 and M2-vaccinated mice, but epitopes of M3, M4 and M5 react with all sera. This observation indicates that antibody was elicited to all HVR1 variants and that epitopes of M1 and M2 are highly conformation-dependent, which is in keeping with previous data showing that lack of self-immunoreactivity is a sign of conformational dependence of the epitopes[20]. The M1 and M2 peptides seemingly assume a conformation on the surface of microtiter plates that prevents interaction with the HVR1 antibody. However, as shown in Fig. 7, antibody to all HVR1 variants, including M1 and M2, are neutralizing, which indicates that they all most likely contribute to synergy.

There are some important limitations to the data presented. We recognize that a limited number of HVR1 variants were evaluated in the neutralization experiments. Although we selected known neutralization resistant and diverse HCVpp for the neutralization panel[25,26], the tested set is only an approximation of the entire HCV genetic space. However, the successful neutralization of all HCVpp, including highly neutralization-resistant variants, supports the strength of the approach. We did not compare our use of MIH cluster-based immunogen selection to other possible selection strategies and therefore note that while the results are promising, this may not be the only or the optimal method for immunogen selection. Our data demonstrate the synergistic effect for a mixture of five HVR1 variants, however, the precise mechanism of the immunological synergy remains unclear. Whether the number of epitopes can be reduced while achieving a similar synergistic effect was not evaluated, but given the high valency of approved (pneumococcal 23-valent)[45] and proposed (inactivated rhinovirus 50-valent)[43] vaccines, the current pentavalent strategy is well within the practical limits of vaccine technology. Additionally, all experiments conducted here were based on using HCVpp. It is known that these particles do not incorporate human lipoproteins, which are important for escape from antibody binding. Although there is evidence indicating no major difference in the neutralizing breadth of mAbs measured using HCVpp or HCVcc[46], testing of synergistic neutralization in various experimental conditions is warranted. Future studies will address these open questions and compare antibodies produced against individual HVR1 variants and the polyvalent mixture to understand the mechanism of synergistic immunization for vaccine design. This will hopefully allow translation of the *in vitro* neutralization data to real protection against HCV infection *in vivo*.

In conclusion, synergistic immune responses to HVR1 variants selected using a sequence space model accounting for the heterogeneity of each position and the interactions among amino acid positions, offer a novel approach to overcoming HCV genetic heterogeneity and the dominance of strain-specific immunity by directing the immune response to cross-immunoreactive neutralizing epitopes within HVR1. Application of this approach opens a new avenue for the development of a universal HCV vaccine. This new approach may be generalizable to other highly mutable viruses.

Materials and Methods

HVR1 sequences

All the HVR1 nucleotide sequences covering the Hypervariable region (81bp) were obtained from the Virus Pathogen Database and Analysis Resource (ViPR)[47]. In addition, the following sequences were added from previous studies: 119 sequences obtained from patients with recent HCV infection, 256 sequences from chronic HCV infection, and 262 sequences from our previously published cross-reactivity experiment[20,48].

This set of 12,245 sequences belongs to all known HCV genotypes. All sequences were clipped and translated to the C-terminal HVR1 eight amino acid sites and cleaned in the following manner: (i) only one sequence per patient was allowed, (ii) only sequences without insertions or deletions were allowed, (iii) sequences with Ns or non-coding regions were removed. Finally, there were 969 distinct variants, which were used in all analyses conducted here.

Distance between HVR1 variants

Genetic distances based on physical-chemical properties[21] were calculated as described in[22]. The MIH distance between every pair of variants was recently developed[23]. The MIH is a distance inspired by the Mahalanobis distance that can be applied to any type of categorical data like nucleotide or amino acid sequences. The Mahalanobis distance accounts for the fact that the variance of each variable is different and that there may be covariance between variables. This distance is reduced to the Euclidean distance for uncorrelated variables with unit variance.

The MIH distance considers the variability of each position as measured by entropy and the existence of coordinated substitutions as measured by mutual information. The MIH distance between two sequences x and y is given by the following formula:

$$MIH(x, y) = xy^T . infMat . xy$$

Where xy is the mismatch vector (with 1 where the symbols are different and 0 where they are the same) and xy^T is its transposed form; InfMat is the information matrix, with 1/entropy in the diagonals and 1/mutual information between position pairs in all other entries. Effectively, if the difference between two sequences occurs at a variable position, this difference receives a low weight. In the same manner, if the difference occurs at positions that are highly associated, this difference also receives a low weight. Thus, the MIH distance is reduced to the Hamming distance when the positions have maximum entropy, and every pair of positions has mutual information equal to zero. A modified version of the MIH distance showed the best performance separating known grouping in a biological validation dataset[23].

K-Step Network and clustering

For the set of HVR1 variants we visualized the matrix of MIH distances by means of a k-step network as previously described[49–51]. The k-step network is equivalent to the union of all possible Minimum Spanning Trees and allows for efficient visualization of the distances among all variants present in a sample. This network was then split into clusters using the Girvan-Newman method as implemented in GEPHI, which was also used to draw the networks[52] . The number of clusters was chosen by using the gap statistic: for each desired number of clusters (from 2 to 40), we measured the average distance within clusters in the k-step network and compared it with the distance in 10000 random partitions of the same size[53].

Immunizations

Peptides for immunization experiments were synthesized using Fmoc chemistry, conjugated to keyhole limpet hemocyanin (KLH) via maleimide linkage, and combined in a 1:1 emulsion with Freund's complete (primary) or incomplete (booster) adjuvant as previously described[54]. For immunizations, female Balb/c mice (4-6 weeks old) were ordered through the UHN animal care facility, acclimatized for one week, pre-bled at ~7 weeks of age (5-8), then subcutaneously injected (25 μg peptide + 25 μL adjuvant) at days 0, 28, and 38, with terminal bleed via cardiac puncture at day 48 [3 mice per group – protocol approved by University Health Network (UHN) Animal Care Committee (ACC)]. Pentavalent immunizations were performed by combining 5 μg of each monovalent antigen into a single formulation with adjuvant (25 μL). Sequential immunizations were performed by administering M1/M2 (12.5 μg each peptide + 25 μL adjuvant) at day 0, M3/4 (12.5 μg each peptide + 25 μL adjuvant) at day 28, and the pentavalent formulation (5 μg each peptide + 25 μL adjuvant) at day 38. Mock immunizations were performed with adjuvant and sterile PBS. Both pre-bleed and mock-immunized sera served as negative controls in subsequent assays. gpE1/E2 sera served as positive control and was obtained by immunizing 5-7 week old female CB6F1 mice with H77 E1/E2 (2 μg of purified antigen) in a 1:1 ratio with alum (75 μg) and monophosphoryl lipid A (MPLA). To obtain sera in all groups, blood samples were processed by centrifugation, heat-inactivated, and stored at −80 °C until analysis was performed.

ELISA assessment of HVR1 binding

As previously described, ELISA was performed to measure HVR1-specific antibody responses in mouse sera[55]. Briefly, 96-well plates (MaxiSorp, Thermo Fisher Scientific), were coated overnight with 2 μg/mL of HVR1 peptides at 4 °C. The next morning, plates were washed 5x with PBS containing 0.05% Tween

20 (PBST) and incubated with group-pooled, serially diluted mouse (PBST) sera for 1 hour at room temperature. Post-incubation, plates were washed 5x with PBST, and incubated for 1 hour with a 1:10,000 dilution of HRP-conjugated anti-mouse IgG secondary antibody. After a final 5 washes, 3,3',5,5'-tetramethylbenzidine (TMB) substrate was added to each well, dark-incubated for 15 min, then the reaction was terminated with Stop-Solution (0.16 M sulfuric acid). Absorbance was read at 450nm, in triplicate, with measurements corresponding to visual colour change in each well. For competitive ELISA, the same protocol was followed, except for the additional incubation of inhibiting peptides (C-terminal 8 AA of HVR1, full-length (FL) HVR1, or FL-HVR1 conjugated to KLH) with diluted sera for 1 hour prior to plate application. ELISA cut-off was calculated by multiplying (2x) the mean of negative controls (adjuvant immunized sera). Statistical analysis was done by unpaired *t*-test using Prism8 software[56].

Neutralization assays

HCVpp neutralization assays were performed as previously described[55]. Briefly, HCVpp were generated by co-transfecting HEK 293T cells with the pNL4-3.lucR^{-}E^{-} packaging plasmid and expression plasmids encoding patient-derived E1E2. To test sera for neutralizing activity, Huh7 cells were plated in 96-well plates (15,000 per well), and incubated overnight. The following day, HCVpp were incubated with heat-inactivated, group-pooled, serially diluted mouse serum for 1 hour at 37°C, and then added in triplicate to Huh7 plated wells. Plates were then incubated in a CO_2 incubator at 37°C for 4 hours before media was replaced. 72 hours later, media was removed and cells were lysed using cell lysis buffer (Promega, Southampton, UK) and placed on a rocker for 15 min. Luciferase activity was then measured in relative light units (RLUs) using a FLUOstar Omega plate reader (BMG Labtech, Aylesbury, UK) with MARS software. Each sample was tested in triplicate. The $1/ID_{50}$ was calculated as the serum dilution that caused a 50% reduction in relative light units compared to pseudoparticles incubated with pre-bleed serum. Values were calculated using a variable slope dose-response curve fit with nonlinear regression, and ordinary one-way ANOVA was used to compare difference between vaccine groups using Prism 9.3.1 (GraphPad Software, San Diego, CA, USA).

References

1. Perz JF, Armstrong GL, Farrington LA, Hutin YJF, Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol*. 2006;45(4):529-538. doi:10.1016/j.jhep.2006.05.013

2. Dore GJ, Bajis S. Hepatitis C virus elimination: laying the foundation for achieving 2030 targets. *Nat Rev Gastroenterol Hepatol*. 2021;18(2):91-92. doi:10.1038/s41575-020-00392-3

3. Page K, Melia MT, Veenhuis RT, et al. Randomized Trial of a Vaccine Regimen to Prevent Chronic HCV Infection. *N Engl J Med*. 2021;384(6):541-549. doi:10.1056/NEJMoa2023345

4. Liang TJ, Feld JJ, Cox AL, Rice CM. Controlled Human Infection Model — Fast Track to HCV Vaccine? *N Engl J Med*. 2021;385(13):1235-1240. doi:10.1056/NEJMsb2109093

5. Pierce BG, Keck ZY, Foung SK. Viral evasion and challenges of hepatitis C virus vaccine development. *Curr Opin Virol*. 2016;20:55-63. doi:10.1016/j.coviro.2016.09.004

6. von Hahn T, Yoon JC, Alter H, et al. Hepatitis C Virus Continuously Escapes From Neutralizing Antibody and T-Cell Responses During Chronic Infection In Vivo. *Gastroenterology*. 2007;132(2):667-678. doi:10.1053/j.gastro.2006.12.008

7. Martinez MA, Franco S. Therapy Implications of Hepatitis C Virus Genetic Diversity. *Viruses*. 2020;13(1):41. doi:10.3390/v13010041

8. Prentoe J, Bukh J. Hypervariable Region 1 in Envelope Protein 2 of Hepatitis C Virus: A Linchpin in Neutralizing Antibody Evasion and Viral Entry. *Front Immunol*. 2018;9:2146. doi:10.3389/fimmu.2018.02146

9. Farci P, Shimoda A, Wong D, et al. Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. *Proc Natl Acad Sci*. 1996;93(26):15394-15399. doi:10.1073/pnas.93.26.15394

10. Shimizu YK, Igarashi H, Kiyohara T, et al. A Hyperimmune Serum against a Synthetic Peptide Corresponding to the Hypervariable Region 1 of Hepatitis C Virus Can Prevent Viral Infection in Cell Cultures. *Virology*. 1996;223(2):409-412. doi:10.1006/viro.1996.0497

11. Bartosch B, Bukh J, Meunier JC, et al. *In vitro* assay for neutralizing antibody to hepatitis C virus: Evidence for broadly conserved neutralization epitopes. *Proc Natl Acad Sci*. 2003;100(24):14199-14204. doi:10.1073/pnas.2335981100

12. Drummer HE. Challenges to the development of vaccines to hepatitis C virus that elicit neutralizing antibodies. *Front Microbiol*. 2014;5. doi:10.3389/fmicb.2014.00329

13. Pierce BG, Boucher EN, Piepenbrink KH, et al. Structure-Based Design of Hepatitis C Virus Vaccines That Elicit Neutralizing Antibody Responses to a Conserved Epitope. Diamond MS, ed. *J Virol*. 2017;91(20):e01032-17. doi:10.1128/JVI.01032-17

14. Tzarum N, Giang E, Kadam RU, et al. An alternate conformation of HCV E2 neutralizing face as an additional vaccine target. *Sci Adv*. 2020;6(30):eabb5642. doi:10.1126/sciadv.abb5642

15. Bankwitz D, Steinmann E, Bitzegeio J, et al. Hepatitis C Virus Hypervariable Region 1 Modulates Receptor Interactions, Conceals the CD81 Binding Site, and Protects Conserved Neutralizing Epitopes. *J Virol*. 2010;84(11):5751-5763. doi:10.1128/JVI.02200-09

16. Prentoe J, Velázquez-Moctezuma R, Foung SKH, Law M, Bukh J. Hypervariable region 1 shielding of hepatitis C virus is a main contributor to genotypic differences in neutralization sensitivity. *Hepatology*. 2016;64(6):1881-1892. doi:10.1002/hep.28705

17. Prentoe J, Velázquez-Moctezuma R, Augestad EH, et al. Hypervariable region 1 and N-linked glycans of hepatitis C regulate virion neutralization by modulating envelope conformations. *Proc Natl Acad Sci*. 2019;116(20):10039-10047. doi:10.1073/pnas.1822002116

18. Law JLM, Logan M, Wong J, et al. Role of the E2 Hypervariable Region (HVR1) in the Immunogenicity of a Recombinant Hepatitis C Virus Vaccine. Ou JHJ, ed. *J Virol*. 2018;92(11):e02141-17. doi:10.1128/JVI.02141-17

19. Law M, Maruyama T, Lewis J, et al. Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge. *Nat Med*. 2008;14(1):25-27. doi:10.1038/nm1698

20. Campo DS, Dimitrova Z, Yokosawa J, et al. Hepatitis C Virus Antigenic Convergence. *Sci Rep*. 2012;2(1):267. doi:10.1038/srep00267

21. Atchley WR, Zhao J, Fernandes AD, Drüke T. Solving the protein sequence metric problem. *Proc Natl Acad Sci*. 2005;102(18):6395-6400. doi:10.1073/pnas.0408677102

22. Campo DS, Dimitrova Z, Mitchell RJ, Lara J, Khudyakov Y. Coordinated evolution of the hepatitis C virus. *Proc Natl Acad Sci*. 2008;105(28):9685-9690. doi:10.1073/pnas.0801774105

23. Campo DS, Mosa A, Khudyakov Y. A Novel Information-Theory-Based Genetic Distance That Approximates Phenotypic Differences. *J Comput Biol*. Published online January 3, 2023:cmb.2022.0395. doi:10.1089/cmb.2022.0395

24. Li H, Zahid MN, Wang S, Shaw GM. Molecular Identification of Transmitted/Founder Hepatitis C Viruses and Their Progeny by Single Genome Sequencing. In: Law M, ed. *Hepatitis C Virus Protocols*. Vol 1911. Methods in Molecular Biology. Springer New York; 2019:139-155. doi:10.1007/978-1-4939-8976-8_9

25. Urbanowicz RA, McClure CP, Brown RJP, et al. A Diverse Panel of Hepatitis C Virus Glycoproteins for Use in Vaccine Research Reveals Extremes of Monoclonal Antibody Neutralization Resistance. Diamond MS, ed. *J Virol*. 2016;90(7):3288-3301. doi:10.1128/JVI.02700-15

26. Salas JH, Urbanowicz RA, Guest JD, et al. An Antigenically Diverse, Representative Panel of Envelope Glycoproteins for Hepatitis C Virus Vaccine Development. *Gastroenterology*. 2022;162(2):562-574. doi:10.1053/j.gastro.2021.10.005

27. Pollard AJ, Bijker EM. A guide to vaccinology: from basic principles to new developments. *Nat Rev Immunol*. 2021;21(2):83-100. doi:10.1038/s41577-020-00479-7

28. Kennedy RB, Ovsyannikova IG, Palese P, Poland GA. Current Challenges in Vaccinology. *Front Immunol*. 2020;11:1181. doi:10.3389/fimmu.2020.01181

29. Vieyres G, Dubuisson J, Patel AH. Characterization of antibody-mediated neutralization directed against the hypervariable region 1 of hepatitis C virus E2 glycoprotein. *J Gen Virol*. 2011;92(3):494-506. doi:10.1099/vir.0.028092-0

30. Duan H, Struble E, Zhong L, et al. Hepatitis C virus with a naturally occurring single amino-acid substitution in the E2 envelope protein escapes neutralization by naturally-induced and vaccine-induced antibodies. *Vaccine*. 2010;28(25):4138-4144. doi:10.1016/j.vaccine.2010.04.024

31. Bailey JR, Laskey S, Wasilewski LN, et al. Constraints on Viral Evolution during Chronic Hepatitis C Virus Infection Arising from a Common-Source Exposure. *J Virol*. 2012;86(23):12582-12590. doi:10.1128/JVI.01440-12

32. Stejskal L, Lees WD, Moss DS, et al. Flexibility and intrinsic disorder are conserved features of hepatitis C virus E2 glycoprotein. Wei G, ed. *PLOS Comput Biol*. 2020;16(2):e1007710. doi:10.1371/journal.pcbi.1007710

33. Scarselli E, Ansuini H, Cerino R, et al. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J*. 2002;21(19):5017-5025. doi:10.1093/emboj/cdf529

34. Bartosch B, Verney G, Dreux M, et al. An Interplay between Hypervariable Region 1 of the Hepatitis C Virus E2 Glycoprotein, the Scavenger Receptor BI, and High-Density Lipoprotein Promotes both Enhancement of Infection and Protection against Neutralizing Antibodies. *J Virol*. 2005;79(13):8217-8229. doi:10.1128/JVI.79.13.8217-8229.2005

35. Johnson J, Freedman H, Logan M, et al. A Recombinant Hepatitis C Virus Genotype 1a E1/E2 Envelope Glycoprotein Vaccine Elicits Antibodies That Differentially Neutralize Closely Related 2a Strains through Interactions of the N-Terminal Hypervariable Region 1 of E2 with Scavenger Receptor B1. James Ou JH, ed. *J Virol*. 2019;93(22):e00810-19. doi:10.1128/JVI.00810-19

36. Augestad EH, Castelli M, Clementi N, et al. Global and local envelope protein dynamics of hepatitis C virus determine broad antibody sensitivity. *Sci Adv*. 2020;6(35):eabb5938. doi:10.1126/sciadv.abb5938

37. Penin F, Combet C, Germanidis G, Frainais PO, Deléage G, Pawlotsky JM. Conservation of the Conformation and Positive Charges of Hepatitis C Virus E2 Envelope Glycoprotein Hypervariable Region 1 Points to a Role in Cell Attachment. *J Virol*. 2001;75(12):5703-5710. doi:10.1128/JVI.75.12.5703-5710.2001

38. Campo DS, Zhang J, Ramachandran S, Khudyakov Y. Transmissibility of intra-host hepatitis C virus variants. *BMC Genomics*. 2017;18(S10):881. doi:10.1186/s12864-017-4267-4

39. Astrakhantseva IV, Campo DS, Araujo A, Teo CG, Khudyakov Y, Kamili S. Variation in physicochemical properties of the hypervariable region 1 during acute and chronic stages of hepatitis C virus infection. In: *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. IEEE; 2011:72-78. doi:10.1109/BIBMW.2011.6112357

40. Schlingmann B, Castiglia KR, Stobart CC, Moore ML. Polyvalent vaccines: High-maintenance heroes. Spindler KR, ed. *PLOS Pathog*. 2018;14(4):e1006904. doi:10.1371/journal.ppat.1006904

41. Chakradhar S. Updated, augmented vaccines compete with original antigenic sin. *Nat Med*. 2015;21(6):540-541. doi:10.1038/nm0615-540

42. Normile D. Safety concerns derail dengue vaccination program. *Science*. 2017;358(6370):1514-1515. doi:10.1126/science.358.6370.1514

43. Lee S, Nguyen MT, Currier MG, et al. A polyvalent inactivated rhinovirus vaccine is broadly immunogenic in rhesus macaques. *Nat Commun*. 2016;7(1):12838. doi:10.1038/ncomms12838

44. Lange M, Fiedler M, Bankwitz D, et al. Hepatitis C Virus Hypervariable Region 1 Variants Presented on Hepatitis B Virus Capsid-Like Particles Induce Cross-Neutralizing Antibodies. Ray R, ed. *PLoS ONE*. 2014;9(7):e102235. doi:10.1371/journal.pone.0102235

45. Cannon K, Elder C, Young M, et al. A trial to evaluate the safety and immunogenicity of a 20-valent pneumococcal conjugate vaccine in populations of adults ≥65 years of age with different prior pneumococcal vaccination. *Vaccine*. 2021;39(51):7494-7502. doi:10.1016/j.vaccine.2021.10.032

46. Kinchen VJ, Bailey JR. Defining Breadth of Hepatitis C Virus Neutralization. *Front Immunol*. 2018;9:1703. doi:10.3389/fimmu.2018.01703

47. Pickett B, Greer D, Zhang Y, et al. Virus Pathogen Database and Analysis Resource (ViPR): A Comprehensive Bioinformatics Database and Analysis Resource for the Coronavirus Research Community. *Viruses*. 2012;4(11):3209-3226. doi:10.3390/v4113209

48. Icer Baykal PB, Lara J, Khudyakov Y, Zelikovsky A, Skums P. Quantitative differences between intra-host HCV populations from persons with recently established and persistent infections. *Virus Evol*. 2021;7(1):veaa103. doi:10.1093/ve/veaa103

49. Quirin A, Cordón O, Guerrero-Bote VP, Vargas-Quesada B, Moya-Anegón F. A quick MST-based algorithm to obtain Pathfinder networks (∞, $n-1$). *J Am Soc Inf Sci Technol*. 2008;59(12):1912-1924. doi:10.1002/asi.20904

50. Campo DS, Dimitrova Z, Yamasaki L, et al. Next-generation sequencing reveals large connected networks of intra-host HCV variants. *BMC Genomics*. 2014;15(S5):S4. doi:10.1186/1471-2164-15-S5-S4

51. Campo DS, Xia GL, Dimitrova Z, et al. Accurate Genetic Detection of Hepatitis C Virus Transmissions in Outbreak Settings. *J Infect Dis*. 2016;213(6):957-965. doi:10.1093/infdis/jiv542

52. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proc Int AAAI Conf Web Soc Media*. 2009;3(1):361-362. doi:10.1609/icwsm.v3i1.13937

53. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a data set via the gap statistic. *J R Stat Soc Ser B Stat Methodol*. 2001;63(2):411-423. doi:10.1111/1467-9868.00293

54. Mosa AI, Urbanowicz RA, AbouHaidar MG, Tavis JE, Ball JK, Feld JJ. A bivalent HCV peptide vaccine elicits pan-genotypic neutralizing antibodies in mice. *Vaccine*. 2020;38(44):6864-6867. doi:10.1016/j.vaccine.2020.08.066

55. Mosa AI, AbouHaidar MG, Urbanowicz RA, Tavis JE, Ball JK, Feld JJ. Role of HVR1 sequence similarity in the cross-genotypic neutralization of HCV. *Virol J*. 2020;17(1):140. doi:10.1186/s12985-020-01408-9

56. Yu X, Gilbert PB, Hioe CE, Zolla-Pazner S, Self SG. Statistical Approaches to Analyzing HIV-1 Neutralizing Antibody Assay Data. *Stat Biopharm Res*. 2012;4(1):1-13. doi:10.1080/19466315.2011.633860

Acknowledgments

The authors gratefully acknowledge John Law and Michael Houghton for providing sera from mice immunized with gpE1/E2. The authors also acknowledge the input and guidance of John Tavis in designing the study and providing recommendations to improve the manuscript. AIM also acknowledges the Toronto Centre for Liver Disease for financially supporting the project.

Funding: Not applicable.

Author contributions:

Conceptualization: AIM, DSC, YK, RAU, JJF
Methodology: AIM, DSC, RAU
Investigation: AIM, DSC, RAU, YK, JJF
Visualization: AIM, DSC
Writing – original draft: AIM, DSC, YK, RAU, JJF
Writing – review & editing: AIM, DSC, YK, RAU, JJF, MGA, AGG, AZ, JKB

Competing interests:

The University Health Network has filed a patent related to the approach to immunogen design described herein, of which AIM, DSC, RAU, MGA, and JJF are listed as inventors. The other authors declare no competing interests. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention (CDC).

Data and materials availability:

All data associated with this study are present in the paper or publicly available[47]. Physical clones and HVR1 peptides for L47, H77, 1a23, 1.4.1, 1.17.1, 1.4.1, JFH1, 2.4.1, 3.1.2, 4.1.1, 5.1.1, and 6.1.1. can be shared upon reasonable request.

Figure Headings

Figure 1. Association of Genetic Distance and Cross-Reactivity.
A) Overview of the cross-reactivity experiment (For more details see Campo et al (20)), which generated a total of 26,833 HVR1 pairwise cross-immunoreactive assays. B). Ratio between mean distance of non-cross reactive pairs and cross-reactive pairs using different types of distances: Hamming, BLOSUM62 scores, MIH and Euclidean distance of 5 physicochemical factors (F1, polarity; F2, secondary structure; F3, molecular size; F4, codon diversity ; F5, charge) (21). C). HVR1 Information matrix using the entire sequence dataset. The diagonal shows the Shannon entropy of each position and the other entries of the matrix show the Mutual information among all pairs of positions.

Figure 2. K-step network of global HVR1 sequence space
All non-redundant HVR1 8aa sequences (n=969) pooled across datasets were used to construct a k-step network with node radius corresponding to the square root of the haplotype frequency. Nodes in blue correspond to those sequences present only in 1a, yellow to those sequences only present in 1b and green to those sequences present in both subtypes.

Figure 3. Clusters in the HVR1 sequence space.
A) Histogram of distances among all pairs of sequences. Three types of distances are considered: Hamming, MIH and Euclidean distances between physiochemical profiles. Each distance type is normalized by dividing by its maximum value. B) Scatterplot of the goodness of each clustering (gap Z score) according to the number of clusters. C) k-step network of all HVR1 sequences. Nodes are colored by membership to each cluster and the big nodes correspond to the most central one in each cluster.

Figure 4. Self and cross-reactivity of HVR1 Antigens
Mice were immunized with monovalent or pentavalent immunogens conjugated to KLH and formulated with either complete (CFA) or incomplete (IFA) Freunds adjuvant and terminally bled at day 48 (A) to evaluate anti-immunogen (HVR1-KLH) titers (B). Sera from each group were evaluated for self and cross-reactivity to each of the five antigens used for immunizations (M1-5) and a genotype 1 patient derived isolate (L47) heterologous to all monovalent sera. Homologous monovalent sera are shown in red, pentavalent in blue, and the heterologous monovalents in black (C). Pentavalent sera were incubated with peptides containing either the immunogen (FL+KLH), full-length HVR1 alone

(FL), or the c-terminal eight AA of HVR1 (C8) to measure binding inhibition to immunogen-coated ELISA plates (D). Error bars indicate mean with standard deviation. *P<0. 05.

Figure 5. Pentavalent Sera Broadly Cross-React with Antigenically Diverse Panel of HVR1 Peptides

HCV variants with the greatest pairwise divergence in their eight C-terminal aa from each peptide used in the pentavalent formulation (A) were synthesized and used to evaluate pentavalent cross-immunoreactivity (blue circles) compared to adjuvant control (black diamonds) (B). Error bars indicate the mean with standard deviation. Dotted line indicates two times the SD of adjuvant control. *, P<0. 05.

Figure 6. Pentavalent Sera Neutralize Panel of Antigenically Diverse HCVpp in Excess of gpE1/E2 Vaccine

Neutralizing activity of pentavalent sera against a multi-genotype panel of HCVpp was evaluated in serial dilutions starting at 1:50, with the exception of 4.1.1 which was additionally tested at 1:20 (A). Neutralizing potencies (1/ID50s) were compared between pentavalent sera and sera obtained from mice immunized with a gpE1/E2 vaccine candidate (B). The 1/ID50 of pentavalent sera was evaluated as a function of the minimal Hamming distance between each HCVpp HVR1 (C-terminal eight aa) and the pentavalent peptides (C). Error bars indicate standard deviation. *,P<0.05.

Figure 7. Pentavalent sera neutralize variants resistant to neutralization by its monovalent constituents

Neutralizing potencies (1/ID50s) were compared across monovalent (orange) and pentavalent (blue) groups (A). Neutralizing potencies (1/ID50s) were compared between pentavalent sera and sera obtained from mice immunized with the same immunogens sequentially (B). Error bars indicate mean with standard deviation. *, P<0. 05.

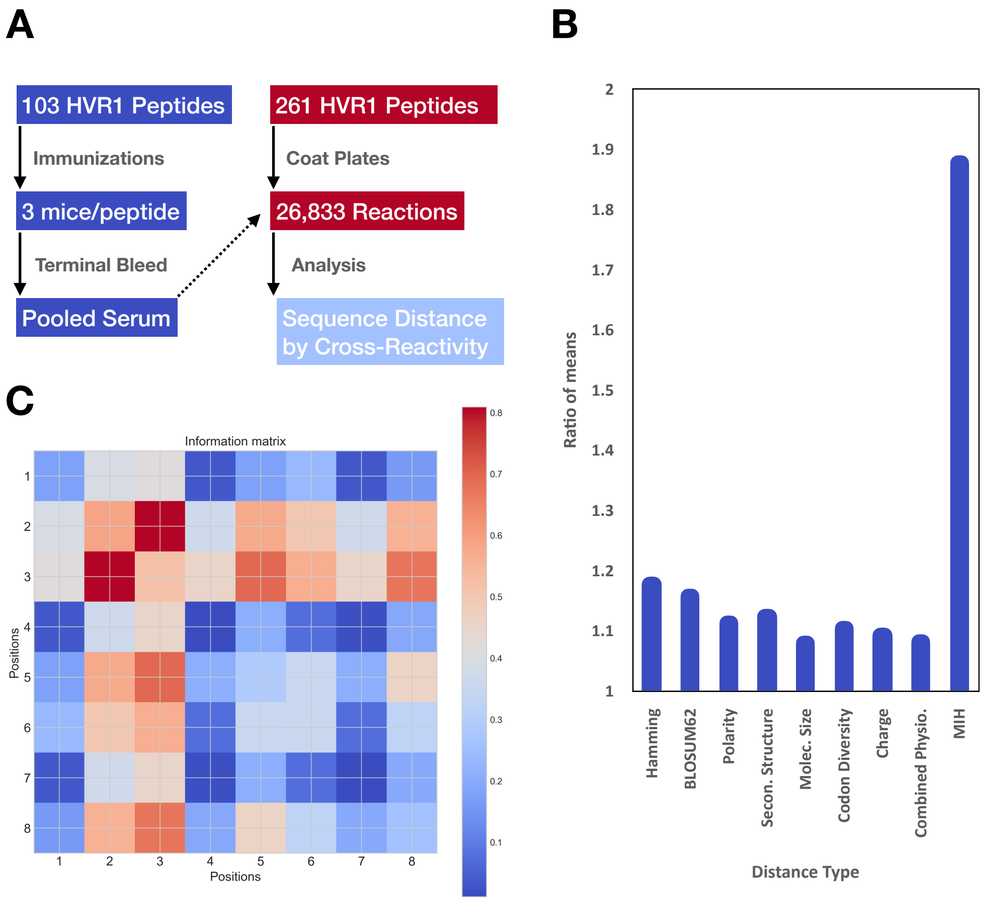
A
103 HVR1 Peptides
Immunizations
3 mice/peptide
Terminal Bleed
Pooled Serum
261 HVR1 Peptides
Coat Plates
26,833 Reactions
Analysis
Sequence Distance by Cross-Reactivity
B
Ratio of means
2
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
Hamming
BLOSUM62
Polarity
Secon. Structure
Molec. Size
Codon Diversity
Charge
Combined Physio.
MIH
Distance Type
C
Information matrix
Positions
1
2
3
4
5
6
7
8
Positions
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

FIG 1

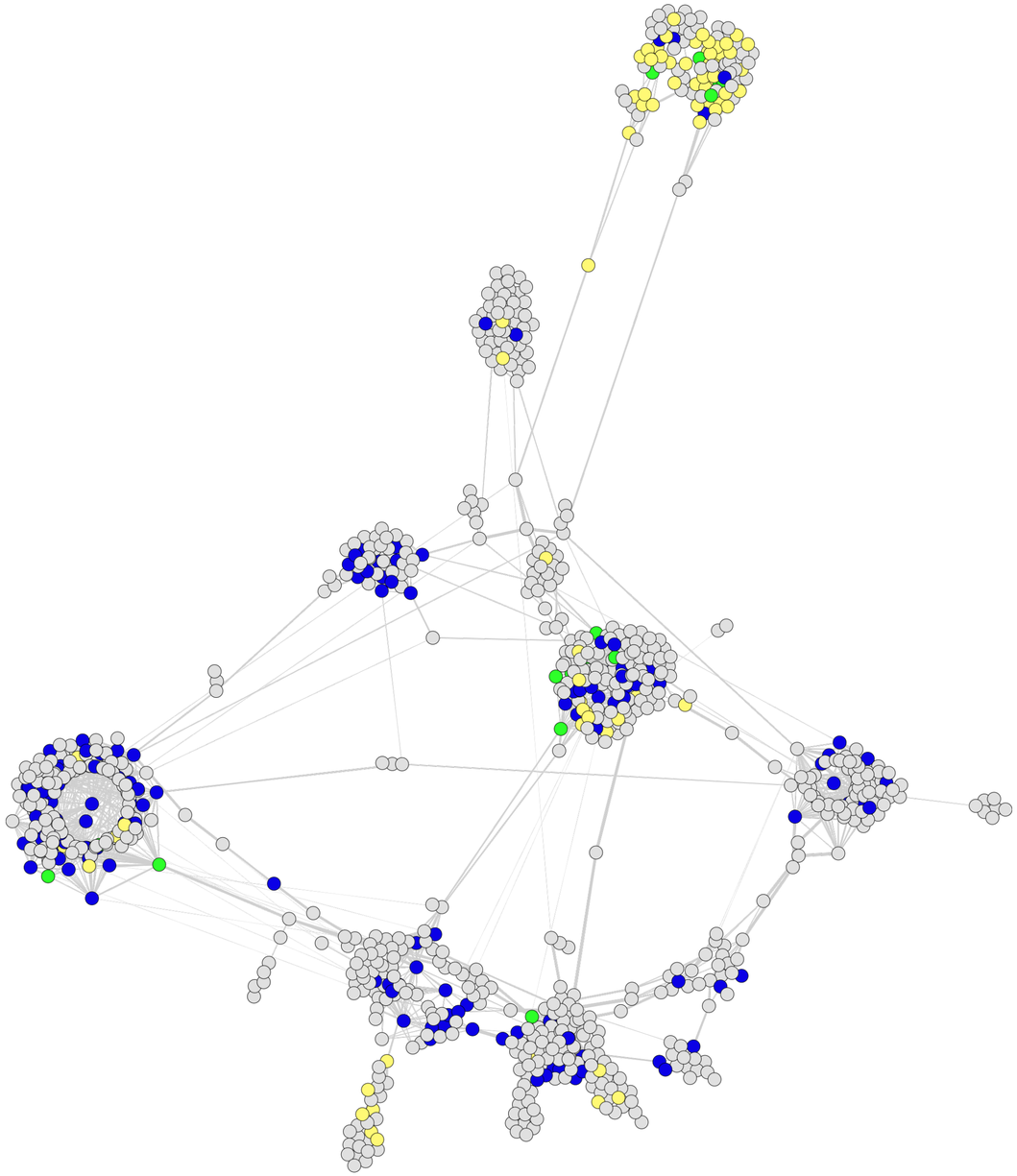

FIG 2

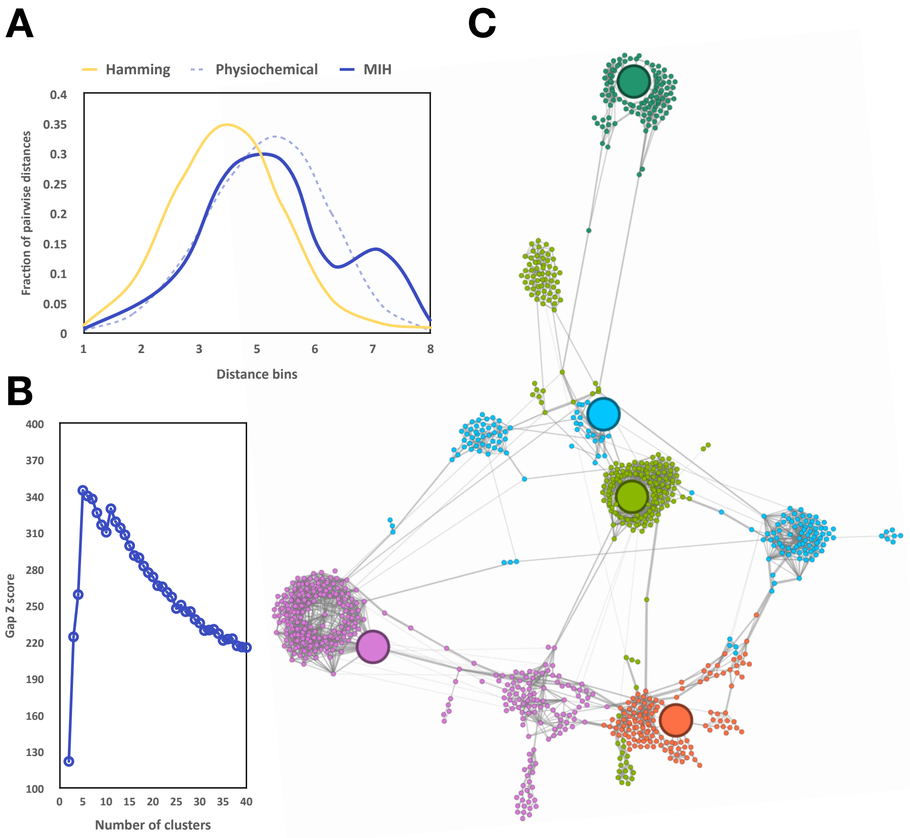
A
Hamming
Physiochemical
MIH
Fraction of pairwise distances
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1
2
3
5
6
7
8
Distance bins
B
Gap Z score
400
370
340
310
280
250
220
190
160
130
100
0
5
10
15
20
25
30
35
40
Number of clusters
C

FIG 3

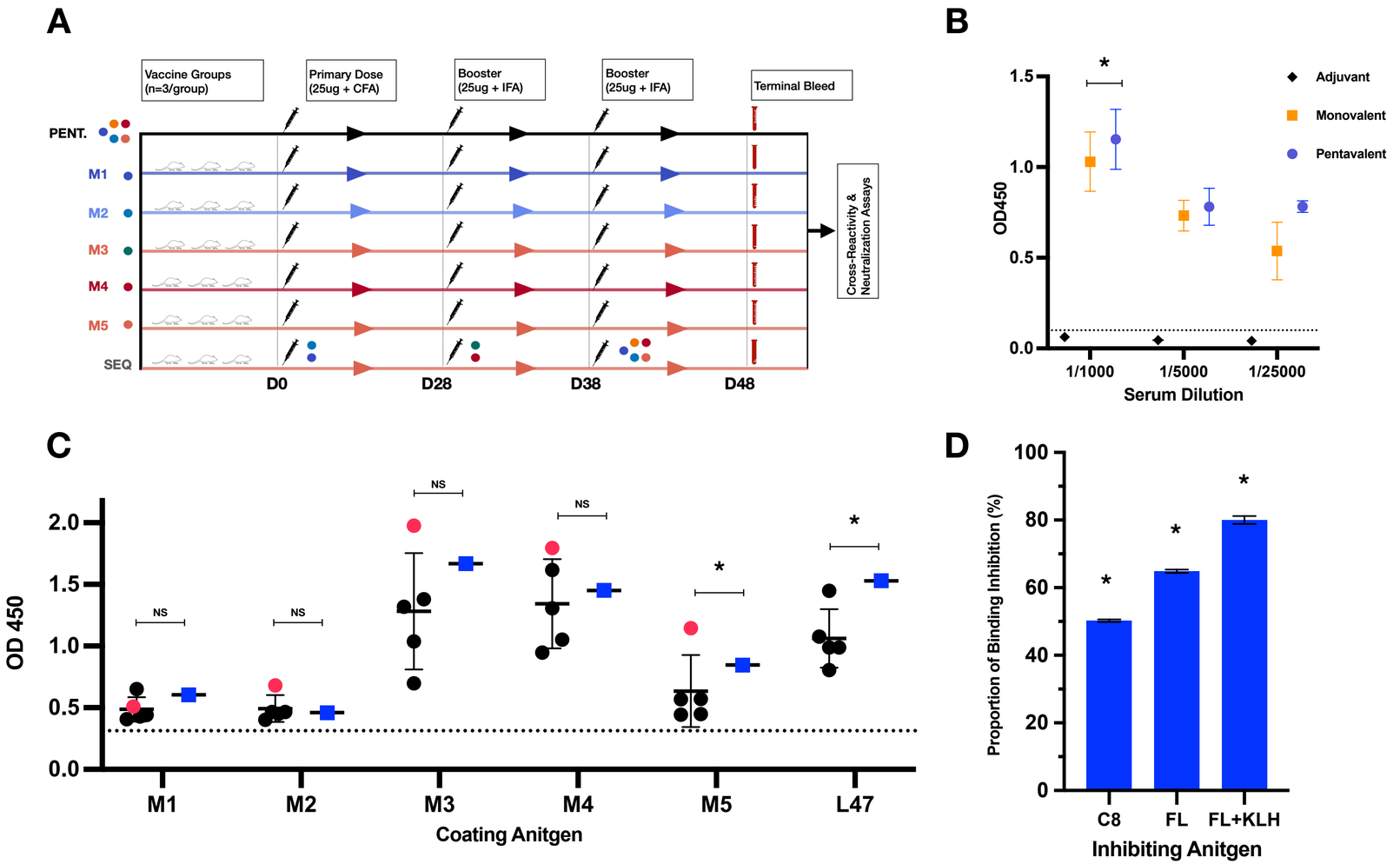
A
Vaccine Groups (n=3/group)
Primary Dose (25ug + CFA)
Booster (25ug + IFA)
Booster (25ug + IFA)
Terminal Bleed
PENT.
M1
M2
M3
M4
M5
SEQ
Cross-Reactivity & Neutralization Assays
D0
D28
D38
D48
B
OD450
Adjuvant
Monovalent
Pentavalent
1/1000
1/5000
1/25000
Serum Dilution
C
OD 450
NS
*
M1
M2
M3
M4
M5
L47
Coating Anitgen
D
Proportion of Binding Inhibition (%)
C8
FL
FL+KLH
Inhibiting Anitgen

FIG 4

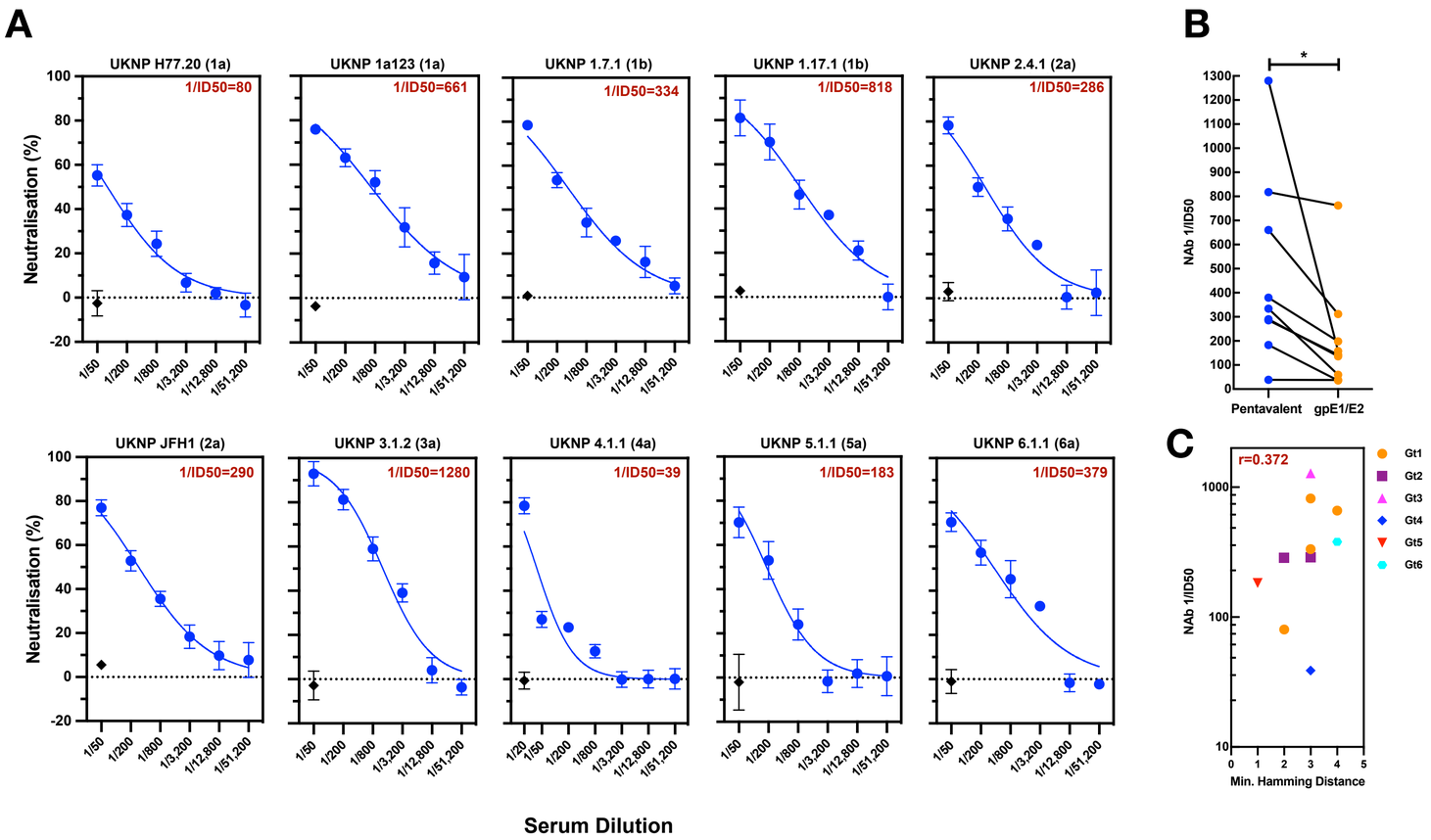
A
UKNP H77.20 (1a)
1/ID50=80
UKNP 1a123 (1a)
1/ID50=661
UKNP 1.7.1 (1b)
1/ID50=334
UKNP 1.17.1 (1b)
1/ID50=818
UKNP 2.4.1 (2a)
1/ID50=286
UKNP JFH1 (2a)
1/ID50=290
UKNP 3.1.2 (3a)
1/ID50=1280
UKNP 4.1.1 (4a)
1/ID50=39
UKNP 5.1.1 (5a)
1/ID50=183
UKNP 6.1.1 (6a)
1/ID50=379
Neutralisation (%)
Serum Dilution
B
NAb 1/ID50
Pentavalent
gpE1/E2
C
r=0.372
NAb 1/ID50
Min. Hamming Distance
Gt1
Gt2
Gt3
Gt4
Gt5
Gt6

FIG 6

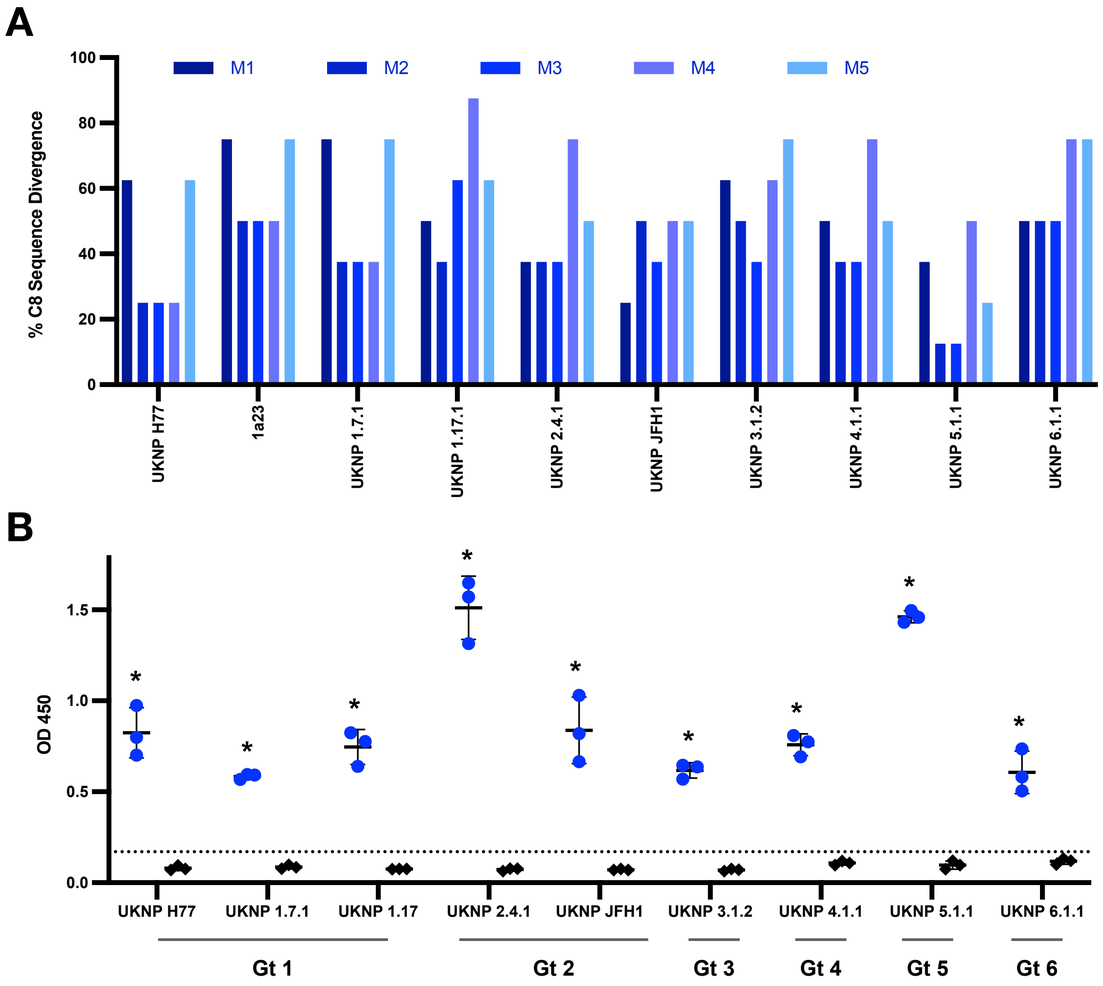
A
M1
M2
M3
M4
M5
% C8 Sequence Divergence
100
80
60
40
20
0
UKNP H77
1a23
UKNP 1.7.1
UKNP 1.17.1
UKNP 2.4.1
UKNP JFH1
UKNP 3.1.2
UKNP 4.1.1
UKNP 5.1.1
UKNP 6.1.1
B
OD 450
1.5
1.0
0.5
0.0
UKNP H77
UKNP 1.7.1
UKNP 1.17
UKNP 2.4.1
UKNP JFH1
UKNP 3.1.2
UKNP 4.1.1
UKNP 5.1.1
UKNP 6.1.1
Gt 1
Gt 2
Gt 3
Gt 4
Gt 5
Gt 6

FIG 5

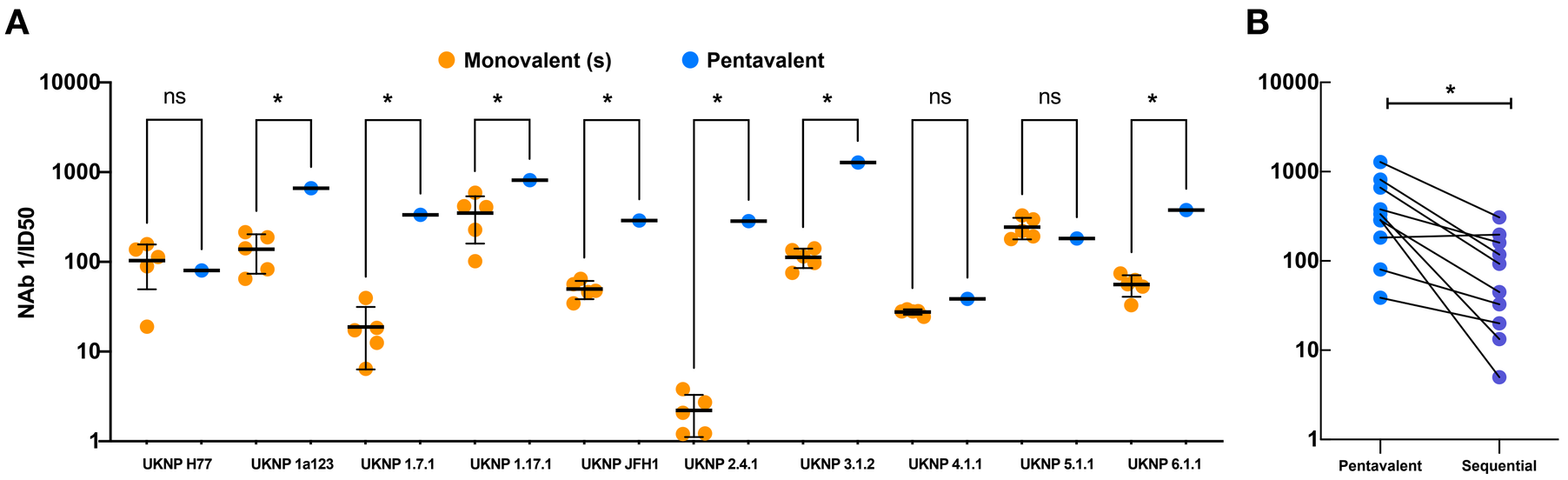
A
Monovalent (s)
Pentavalent
NAb 1/ID50
10000
1000
100
10
1
ns
*
UKNP H77
UKNP 1a123
UKNP 1.7.1
UKNP 1.17.1
UKNP JFH1
UKNP 2.4.1
UKNP 3.1.2
UKNP 4.1.1
UKNP 5.1.1
UKNP 6.1.1
B
Pentavalent
Sequential

FIG 7