

# Hepatitis C Virus Antigenic Convergence

David S. Campo<sup>\*1</sup>, Zoya Dimitrova<sup>1</sup>, Jonny Yokosawa<sup>1,2</sup>, Duc Hoang<sup>1,3</sup>, Nestor O. Perez<sup>1,4</sup>, Sumathi Ramachandran<sup>1</sup> and Yury Khudiyakov<sup>1</sup>

<sup>1</sup>Molecular Epidemiology & Bioinformatics Laboratory, Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, USA, 30329. \*fyv6@cdc.gov

<sup>2</sup>Laboratório de Virologia, Instituto de Ciências Biomédicas, Universidade Federal de Uberlândia, Uberlândia, Brazil.

<sup>3</sup>National Institute of Hygiene and Epidemiology, Hanoi, Vietnam.

<sup>4</sup>Probiomed S.A., Tenancingo, Mexico.

\*Correspondence and requests for materials should be addressed to D.S.C. (fyv6@cdc.gov)

## Supporting information

Table S1. Sequence of the peptides used in the cross-immunoreactivity experiment (Gen= Genotype; Immu = immunogen; Anti = Antigen; Self = reaction with itself; 1 = yes; 0 = not; k = number of positive reactions).

ID	Genbank	Gen	aa seq peptide	Self	Immu	k_immu	Anti	k_anti
P001	M62381	1a	ETHVTGGSAGRTTAGLVGLLTPGAKQNIQLI	1	1	9	1	27
P002	M74807	1b	HTRVTGGVQGHVSTLTSLFRPGASQKIQLV	1	1	15	1	10
P003	D85516	1b	TTHVTGGVQAHGAYGLASLFNVGPHQKIQLV	0	0	0	1	16
P004	D13406	1b	HTLTGGHAARLTSGFAGLFTPGPSQRIQLI	0	0	0	1	28
P005	E03766	1b	HTHVTGGRVASSTQSLVSWLSQGSPQKIQLV	0	0	0	1	16
P006	D49374	3b	TTHTTGGSAQAATAGFTSFFTRGPSQNLQLV	0	0	0	1	20
P007	D28917	3a	STHVTAGQAARNAYGITSLSFVGAKQNLQLI	0	0	0	1	22
P008	D30613	1b	GTHVTGGKVAYTTQGFTSFFSRGPSQKIQLV	0	0	0	1	20
P009	D00689	1b	HTRVTGGQVAFRTHGLVSLFTQGSPQKIQLV	0	0	0	1	22
P010	D00690	1b	ETRVTGAVQGHGALGLASLFTPGPSQKIQLI	0	0	0	1	21
P011	E04259	1a	ETIVSGGQAARAMSGLVSLFTPGAKQNIQLI	0	0	0	1	18
P012	E04260	1b	ETYTSGGAASHTTSTLASLFSPGASQRIQLV	1	1	43	1	18
P013	D10074	1b	TTHVTGGATGHTTSGIASLFLPGASQKIQLI	0	0	0	1	30
P014	D10075	2a	NTRTVAGSAAATTRGFTSMFSSGSKQNLQLI	0	0	0	1	20
P015	D10077	2b	STQVTGGQAAHTVRGVASIFSPGSRQDISLI	1	1	12	1	28
P016	D10934	1b	NTYVTGGAAARGASGITSLSRFGPSQKIQLV	0	0	0	1	22
P017	D10988	2b	TTYSSGQEAGRTVAGFAGLFTTGAKQNLILI	1	1	122	1	15
P018	D11355	1b	VTYTTGGSQARHTQGVASFPTGPAQKIQLI	1	1	16	1	41
P019	D11443	3b	NTYTTAGSMAQSIYRLTDIFSTGPSQKLQLV	1	1	10	1	9

P020	D12956	1b	STQVMGGQQGRAAYSLASLLSPGASQKIQLV	0	0	0	1	12
P021	D12972	1b	QTHTVGGQMGGHGVRLTNLFSAGSAQNIQLI	1	1	12	1	23
P022	D13970	2a	ETHSVGGSAAHTTSRFTSLFSPGPQQNIQLI	0	0	0	1	28
P023	D14853	1c	ETRVGTGGAAGHTAFGFASFLAPGAKQKIQLI	0	0	0	1	15
P024	D16189	1c	DTQVTGGSAAAYDARGLASLFTPGPKQNIQLI	0	0	0	1	40
P025	D16612	3a	TTHVTGSSPAKAAQSFSSLSLGPQKTIQLV	0	0	0	1	54
P026	D16614	3b	DTCTTGGAARNARGFASLFSSGARQQQLV	0	0	0	1	6
P027	D16616	3b	ETHVTGGTVASGASILTRMFASGPKQNLQLV	1	1	44	1	10
P028	D16618	3a	HTNVIGSNAGSTLFSIQRIFSPGAAQRIQLI	1	1	12	1	7
P029	D16620	3a	RTHTTGGVAAHTTSSLTTLFTRGPSQNLQLI	0	0	0	1	24
P030	D31971	6d	NTHVTGAVAGKTVSRFTGLFTAGSQNLQLV	0	1	1	1	6
P031	D31972	2c	NTYASGGAVGHQTASFVRLAPGPQQNIQLI	0	0	0	1	8
P032	D37840	3b	NTRTVGGAAANTVSGFAGLFTRGPSQNLQLV	1	1	207	1	7
P033	D43680	4a	QTHVTGGAARGAAGLASLLSGPEQNLQLI	1	1	58	1	24
P034	D49746	2	GHTTVGHAQGH TLHGVLFPAGPKQNIQLI	0	0	0	1	11
P035	D49747	3k	STRISGGSAAHNTYGLSSLFSSGPKQNIQLI	1	1	42	1	38
P036	D49752	3k	STTISGGSAAARSVWGITSLSFSPGSNQNLQLV	0	0	0	1	2
P037	D49753	3k	HTTTTGGVAARTTSGLTALFTTGPKQHIQLI	0	0	0	1	29
P038	D49754	2c	TTRTVGGTAASATSSFTKFFDLGPRQNIQLI	1	1	156	1	29
P039	D49757	2i	AAGAVGGSVARNTAGVAGLFSPGAKQNVQLI	0	0	0	1	22
P040	D50480	1b	NTYTTGAAQGRTVSRLTSIFTPGASQNIQLI	1	1	80	1	23
P041	D50481	1b	NTYVSGGAKSHTTQGLVSLFAPGAQQKIQLV	1	1	105	1	25
P042	D50484	1b	STRVSGGTAYNTRGLTSLFSSGAQQKIQLV	0	0	0	1	25
P043	D88467	6d	TTRTTGAQAARATLGFTGLFQTGAKQNIHLI	0	0	0	1	9
P044	D88474	6d	ETYIIGAATGRTTAGLTSLSFSSGSQQNLQLV	0	0	0	1	3
P045	D88477	6d	TTHAVGSIAGQQARGFTSLLSPGPQQNLQLV	1	1	30	1	12
P046	D88478	6d	STYTTGSLASQNTAGFVSLFSSGAKQNLQLV	0	1	6	1	5
P047	I08294	1a	ETHVTGGSAGHTVSGFVSLAPGAKQNVQLI	0	0	0	1	22
P048	L02836	1b	DTYASGGAQGRSTLGFTSLFTPGASQKIQLI	0	0	0	1	24
P049	L16889	1b	ETNTIGGQASHIISRTSFFSLGPRQRIQLI	0	0	0	1	10
P050	L16891	1a	RTHTTGGSAAFTTAGLAGIFSPGARQNIQLV	0	0	0	1	28
P051	L16892	1a	GTYYTGGSAAYNAQGFAGFLSQGPRQNIQLI	1	1	4	1	24
P052	L19373	1a	ETHVSGGSLARGAAGLVGLFDSGPKQNIKLI	1	1	54	1	18
P053	L19374	1a	STHTIGRSAARTTSGLAGLFPNGATQNVQLI	0	1	1	1	25
P054	L19375	1a	GHTHTIGGSAARSTSTIASLFPNGASQNVQLI	0	0	0	1	27
P055	L19378	1a	ETQVSGGTAAARAHLTNLFTQGAKQNIQLI	0	0	0	1	15
P056	L19379	1a	ETHVTGGSAAARSVQGFTSFFQKGAKQNVQLI	1	1	118	1	20
P057	L19380	1a	ETYTSGGSAAHTTSGFVSFFSPGAKQNIQLI	0	0	0	1	25
P058	L19381	1a	ETYVTGGSAGRTVSGFTGIFAPGARQNIQLI	1	1	40	1	23
P059	L29578	5a	TTQISGGSSAQTTYGIASFITRGAQQKLQLI	0	0	0	1	16
P060	E08315	1b	DTHVTGGAQAKTTNRLVSMFASGPSQKIQLI	1	1	69	1	24
P061	M74804	1a	ETHVTGGNAGRAAAGIAGLFTLGAKQNVQLI	0	0	0	1	22
P062	M74805	1a	DTYATGGSVASIMAGIARFFSPGARQDIQLI	0	0	0	1	16
P063	M74806	1b	GTHVTGGAARNNAHSLTSLAPGASQKIQLI	0	0	0	1	29
P064	M74808	1a	ETHVTGGSAGHTVTGIASLFTSGAKQNIQLI	0	0	0	1	23

P065	M74809	1b	HNHVTGGTSARNTFGITTFLTQGPSQKLQLV	0	0	0	1	15
P066	M74810	1b	STYVSGGAQARAAQGITSLSFRGSSQKIQLV	1	1	18	1	16
P067	M74811	1a	STHVTGGTAAHTVAGFSSLFTVGPKNQIQLI	1	1	69	1	26
P068	M74812	1a	GTHVTGGAAARDAFRFSSLFTRGPSQNIQLI	0	0	0	1	25
P069	M74813	1b	HTYTTGGTVARSTQGLVGLFSPGPSQNIQLI	1	1	3	1	29
P070	M74814	1b	NTYVSGGTAGHTGHGLTALFSPGASQNIQLI	1	1	25	1	30
P071	M74815	1b	STRTAGGAQAFNTYGVASIFSPGPSQRIQLV	0	1	11	1	25
P072	M74886	1b	DTYASGGAAGRATYGITSLFAPGASQNIQLI	1	1	10	1	21
P073	M84754	1b	STIVSGGTVARTTHTSLASLFTQGASQKIQLI	0	0	0	1	16
P074	M86769	1b	GTSLTGGARARAASGLAGLFSSGPSQRIQLI	0	0	0	1	31
P075	U16362	1b	TTHVTGGAQGRAASSLTSFSPGPVQHLQLI	0	0	0	1	34
P076	D17763	3a	HTYTTGGTASRHTQAFAGLFDIGPQQKLQLV	1	1	1	1	15
P077	E06695	1b	TTYVTGGTTGRTTSSFASLFTLGSHQKVQLI	0	0	0	1	21
P078	S55630	1b	NTYVTGGAAAKGTSTFAGLFAPGSRQNIQPI	0	0	0	1	28
P079	S55848	1a	TTYVTGGAASHTVSGLNGLFTSGARQNIQLI	1	1	145	1	22
P080	S55849	1a	STHVSAGSAAHAFTLSLLRLGPKQNIQLV	0	0	0	1	9
P081	S62395	1a	ETHVTGGSAASTTSTLTCLFMPGASQNIQLI	0	0	0	1	29
P082	S71394	1b	STYVTGGASGRSTHGLVSLFNLGAQQKVQLI	0	0	0	1	18
P083	S71864	1b	ATNMTGGAPARTTYKLTTLFSYGASQKIQLI	0	1	8	1	11
P084	S71957	1b	STHVTGAVQGHRSIRGLTSLFTSGPAQKIQLV	1	1	77	1	19
P085	S72725	1a	ETYTTGGSVARATSRLTNLFSQGAKQDVQLV	0	0	0	1	5
P086	U14226	1b	ATYTTGGEAARTTRGIASLFTSGAKQKIQLI	1	1	27	1	26
P087	HCU14228	2a	QTYTVGSSAGHSTRGLASLFTPGARQKLQLI	0	0	0	1	36
P088	HCU14229	2a	QTYTVGSSAGHSTRGLASLFTNGARQELQLI	0	1	9	1	18
P089	U14230	1b	TTHTTGGQAGRSASGLASLFTTGPSQRIQLV	1	1	90	1	30
P090	U14234	1b	DTRVTGGTAGRTARGLTSLFTLGPSQNIQLI	0	1	1	1	25
P091	HCU14236	3a	NTYVTGGSAARDTYGFAGLFNIGANQKLQLV	0	1	0	1	7
P092	HCU14237	3a	TTYVTGGSAARNAYGFSSLFNVGANQKLQLV	0	0	0	1	11
P093	U14238	1b	NTRTTGGTAAYTTRGLTNLFTPGPAQKIQLV	0	0	0	1	31
P094	U28032	1a	TTHVSGGAAGHTTLGFASLFTSGPSQKIQLV	0	0	0	1	19
P095	U45463	1b	QHTTIGKAAHTTNSLTSIFSPGPAQKVQLI	1	1	59	1	33
P096	U45476	1b	TTTVTGGSQARTVYELTSLYTRGPSQRIQLV	1	1	1	1	17
P097	X53131	1a	ETYTSNGNAGHTMTGIVRFFAPGPKQNVHLI	1	1	77	1	19
P098	X53132	1a	ETHRTGGSAARSTAGVASLFTPGARQNIQLI	0	0	0	1	31
P099	X53133	1a	TTYTTGGNAARTTQALTSFFSPGAKQDIQLI	0	0	0	1	23
P100	X53134	1a	ETTVTGGSAAHGALGASLFTNGARQNIQLI	1	1	5	1	21
P101	X60556	1b	NTRVTGGVQGRDTSGLVSLFSLGPSQKIQLV	0	0	0	1	38
P102	X60557	1b	ETHVTGGASARTTQRFTSFFDLGPSQKIQLV	0	0	0	1	21
P103	X60558	1b	VRTTGEVAARTANTFASLFTTGPSQNIQII	0	0	0	1	12
P104	X60559	1b	GTYYTVGGASAFRTSRLTSFALGPSQRIQLV	0	0	0	1	14
P105	X60561	1b	TTYMTGGANARTTQGFVSLFTPGPAQKIQLV	1	1	97	1	28
P106	X60562	1b	ETHVTGGTSARTTQGFVSLFSAGASQKIQLV	0	0	0	1	15
P107	X60564	1a	TTRVSGGTAAHTTAGLTSFSPGPRQNIHLV	1	1	159	1	42
P108	X60565	1b	TTHVSGGTAGRTTASLTSFFAPGASQRIQLV	0	0	0	1	21
P109	X60566	1b	STSVVGGRQASATFRFTSFFSRGPTQEIKLI	1	1	10	1	19

P110	X60567	1b	ETRV TGQVGR TTQSLTSLFTPGPSQTIQLI	0	0	0	1	33
P111	X60568	1a	NTYVTGGSAGRAVAGFAGLLQPGAKQNVQLI	0	0	0	1	20
P112	X60569	1a	GTHVTGGSAGRATAGIAGLLTPGARQNIQLV	1	1	14	1	25
P113	X60570	1b	HTRVTGGKAGRETMGFVSFFSAGPAQKIQLI	1	1	68	1	30
P114	X60571	1b	GTHVTGGQAARTTQSFTSLFSPGPQQKIQLI	1	1	76	1	42
P115	X60578	1b	RTHVTGGKAAHTTKGFASLFTPGPSQNIQLI	0	0	0	1	36
P116	X60580	1b	ETRRIGEQAQOTTGSLANLFLSGPTQKISLV	1	1	9	1	7
P117	X60582	1b	STYVTGGEASRTTRGFASLFTLGSSQKIQLI	0	0	0	1	21
P118	X60585	1b	QTYVTGGKAAQTVSGFTGLFSSGPSQKIQLI	0	0	0	1	20
P119	X60587	1b	ATYTTGGSVSRITQSLTSLFNPGPIQKIQLI	0	0	0	1	15
P120	X60588	1b	SSSTIGRAQGRQAYGLASIFSPGASQKIQLI	1	1	21	1	44
P121	X60589	1b	ETHVAGGAAGYTTRGLTSLFTLGASQKIQLI	0	0	0	1	12
P122	X60590	1b	KTHVTGMVAGKNAHTLSSIFTSGPSQNIQLI	0	0	0	1	26
P123	X60591	1b	ATHVTGGTEARTTRGLVSLFTPGPSQQLQLV	1	1	102	1	32
P124	X61596	1b	TTYVSVGHASQTRRVASFFSPGSAQKIQLV	0	0	0	1	13
P125	X61593	1b	RTQVTGAQAGHTTSGLASLFTPGPSQKIQLV	1	1	106	1	33
P126	X72979	1a	DTHVTGGTAAHSAGTLVSMFALGPSQKLQLV	0	0	0	1	22
P127	X76412	1b	STYTSGGSAAHTTNRFTSLFTLGASQKIQLV	0	0	0	1	10
P128	X76413	1b	ETSVMGASGHTTRTVTNLFSPGATQNIQLI	1	1	2	1	27
P129	X76415	2i	NTRTVGGNTAQGGSLAGLFSAGSQQHRLRV	0	0	0	1	11
P130	D50483	1b	DTYTTGGAQGH TTSRVASLFTSGASQNIQLI	0	0	0	1	21
P131	D10750	1b	ATYTSGGVAGRTTSGFTSLFSSGASQKIQLV	1	1	37	1	17
P132	E03795	1b	QTRVSGGTAGRNTGHFVSLFRFGASQNIQLI	1	1	35	1	21
P133	D50485	1b	RTTVTGGAQGH TQRLASLFTFGAQQRIQLV	0	0	0	1	11
P134	A32200	1b	EPYTTGGTHGRAAHGLTSLFTPGPAQKIQLV	0	0	0	1	27
P135	E07544	1b	GTHVSGGSAAQT TSGLASLFTSGSAQNIQLV	0	0	0	1	27
P136	D50482	1b	STHVSGGTAYNTRSFTSLFSSGAQQKIQLV	0	0	0	1	23
P137	A22778	1a	ETYTTGGSTARTTQGLVSLFSRGAKQDIQLI	1	1	15	1	27
P138	S55631	1b	KTYVTGGASACHTWASVLSFSPGPSQKIQLI	0	0	0	1	26
P139	X60586	1b	TTLTTGGSASRAVQRYTSIFSAGPSQNIQLI	0	0	0	1	24
P140	M62382	1a	ETHVTGGSAGRSVLGIASFLTRGPKQNIQLI	0	0	0	1	20
P141	X60574	1b	KTNRVGG LQAKHTREVVSIFAPGARQNIQLI	1	1	9	1	31
P142	X60555	1b	QTYVSGGSSGR TTSGLVSI FSPGASQNLQLI	0	0	0	1	23
P143	X60576	1b	KKTFVGQVPAQGAQTLSSFFT LGPQQKIQLI	0	0	0	1	25
P144	X60575	1b	KKKKQGRAVGQT TGGTSLFSLG SQRIQLI	1	1	23	1	46
P145	X60584	1b	KTHVTGGQA AQRASGFVGLFTPGPSQKIQLI	1	1	72	1	38
P146	D12969	1b	QTHTVGGQM GHGVRGLTSLFSAGSAQNIQLI	0	0	0	1	27
P147	E04262	2a	QTHTVGGSTAHNARTLTGMFSLGARQKIQLI	0	1	2	1	17
P148	D12965	1b	QTRTVGGQVGH SVRGLTSLFSAGSAQNIQLI	1	1	12	1	22
P149	D12952	1b	STRVSGGQQGRAAHSLTSLFTLGASQNIQLV	0	0	0	1	7
P150	D12960	1b	QTRTVGGQVGH SVRGFTSLFSAGSAQNIQLI	0	0	0	1	39
P151	D12967	1b	QTRTVGGQM GHGVRGLTSLFSAGSARNIQLI	0	0	0	1	28
P152	X60554	1b	ATYATGAAQGHATNSFVSLFRSGASQNLKLV	1	1	6	1	9
P153	X60572	1b	KKKYVGAPPARAANTLTSMFNP GARQNIQLI	0	0	0	1	37
P154	D12949	1b	NTRVTGGRQGRAAHSLTSLFSPGASQNIQLV	0	0	0	1	31

P155	D12942	1b	STRVTGGQQGRAVHGIASLFSLGASQKIQLV	0	1	1	1	9
P156	D12963	1b	QTRTVGGQVGHSVRGFTSLLSAGSAQNIQLI	0	0	0	1	19
P157	D12948	1b	STRVTGGQQGHAAHSLTSLFRLGASQNIQLV	1	1	5	1	12
P158	X60577	1b	KKGLVGEVSGQGARGLTSLFTHGSNQRIQLI	1	1	166	1	29
P159	D12957	1b	RTRTVGGQVGHSVRGFTSLFSAGSAQNIQLI	0	0	0	1	23
P160	D12943	1b	STRVTGGQQGRAVQGFASLFRLGASQEIQLV	0	0	0	1	6
P161	D12951	1b	GTRVTGGRQGRAAHSLTSLFSPGASENIRLV	0	0	0	1	22
P162	E07468	2b	TTYSTGQQAGRTAEGIASLFTNGAKQNLHLI	0	0	0	1	31
P163	D12953	1b	STQVMGGQQGRAAYSLASLLSPGANQKIQLV	0	0	0	1	11
P164	E04263	2b	STQVTGGQAAHTVRGFASIFSPGSRQDISLI	0	0	0	1	27
P165	E08446	1c	ETQVSGGSAAQTTYGLTALFRTGPNQKIQLI	1	1	195	1	12
P166	D30722	2b	GTYASGGTAGHTVHGVTSLFAPGPRQKIDLI	0	0	0	1	39
P167	D30725	2b	KTYVSGGQAARTAYRLAGLFTVGPQKIDLI	0	0	0	1	19
P168	D30723	2b	GTYASGGTAGHTVYGVTSLFAPGPRQKIDLI	0	0	0	1	29
P169	D30727	2b	GTYVTGGRAGHTVRGLTSLFAPGPRQKIDLI	1	1	63	1	40
P170	D30728	2b	GTHTTGGTAGHTVRGIASLFPSPGPKQKIDLI	0	0	0	1	34
P171	X60563	1a	TTRVSGGAAGYVTSGLVGLFTRGPKQNIQLV	1	1	53	1	8
P172	D30726	2b	ETYVTGGRAGRTVSSFAGLFTSGPKQKIDLI	0	0	0	1	18
P173	L25447	1a	ETHVTGGSAAHGASTFTSFFSRGASQNIQLI	0	0	0	1	22
P174	E06696	1b	KTTVTGGSAAFQSRKLVSSFSPGPKQNIQLD	0	1	4	0	0
P175	S83317	1b	STHVSGAASARGAHLASIFTPGARQNIQLI	1	1	52	1	34
P176	E06370	1b	GTHVTGGKVAYTTQSFTSFFSRGPSQRIQLV	1	1	29	1	18
P177	E06173	1b	RTHVTGGKVAYTTQRFTSFFSRGPSQKIQLI	0	0	0	1	33
P178	L16893	1a	ITHTTGSAGFTTVGLAGIFSQGAKQNIQLI	0	1	3	1	20
P179	I40242	1a	ETHVTGGNVARTAAARFAGLFTPGAQQNVQLI	1	1	34	1	18
P180	M74885	1b	DTYVSGGAAARSISGFTSLFTPASQKIQLI	0	0	0	1	23
P181	M74887	1b	RTYASGGAAGRTHHGFTSLFSTGARQNIQLI	0	0	0	1	29
P182	U51795	1a	STYTTGGSAAQAAAGIASFFRSGPKQDIQLI	0	0	0	1	19
P183	U51794	1a	HTITTGGSAAHYHTRAFTNLNFRGAQQNIQLV	0	0	0	1	10
P184	U51791	1a	ETYVTGGAARTMTGLTGLFSPGAKQNIQLI	0	0	0	1	24
P185	U51796	2b	STYTIGSHAGQTVSKLSSLFTPGPQQRISLV	0	0	0	1	6
P186	Z68710	3a	TTYTTAGSAARHVSFGTGLFDQGPQNQLV	0	0	0	1	24
P187	Z68725	3a	TTYTTAGSAARHVSRTGLFAQGPQNQLV	0	0	0	1	14
P188	Z68709	3a	TTYTTAGSAARHVSFTGLFAQGPQNQLV	1	1	23	1	11
P189	Z68728	3a	NTYTTGGSAAARSVSVLAGIFSPGPQQLELI	0	0	0	1	17
P190	U28036	2a	TTYTTGGATARTAAGLTRLFSHGPTQNQLI	0	0	0	1	21
P191	Z68712	3a	TTYTTAGSAARHVSFTGLFDAGPKQNQLV	0	1	5	1	15
P192	Z68723	3a	TTYTTAGSAARHVNVTGLFGQGPQNQLV	0	0	0	1	10
P193	Z68711	3a	TTYTTAGSAARHVSFTGLFDVGPQNQLV	0	0	0	1	10
P194	Z68717	3a	TTYTTAGSAARSVSTFTGLFDLGPQNQLV	0	1	0	1	10
P195	Z68722	3a	TTYTIAGSAARHVNVTGLFGSGPKQNQLV	0	1	11	1	9
P196	Z68720	3a	TTYTTAGSAARSVSVFTGLFDAGPKQNQLV	0	0	0	1	15
P197	Z68732	3a	VTYTTGGSAAYSTRAFTGLFSRGAQQNQLV	1	1	15	1	15
P198	Z68727	3a	STYTTAGSAAHSTSMVASLFTPGAKQNQLI	0	0	0	1	28
P199	U51792	1a	GTYVSGGAAGYTTAGFAGLSPGPKQNVHLI	1	1	37	1	28

P200	Z68715	3a	TTYTTAGSAARSVSAFTGLFDLGPQNQLV	0	0	0	1	12
P201	Z68734	3a	GTYYVSGGSVARDANSLVKLFDVGPQNQLV	0	0	0	1	16
P202	Z68714	3a	TTYTTAGSAARSVSAFTAVFDLGPQNQLV	0	0	0	1	13
P203	Z68708	3a	TTYTTAGSVARRVSTFTGPFLVGPKNPQLV	1	1	4	1	9
P204	D43679	4a	QTHVTGGAAARGAAGLASLLSGPKQNQLI	0	0	0	1	23
P205	S74442	1a	HTTVSGGAAAHAVSGFVGFFAPGPKQNVQLI	0	0	0	1	19
P206	U24609	1b	HTYTTGGAAGRTASGLTSLFTPGAQKIQLI	1	1	79	1	32
P207	U24612	1a	ETYTTGGSAAARTTSALTSLFMQGPQNIQLV	1	1	10	1	9
P208	U24611	1b	DTYVSGGTQGRVLHGFTSFLTSGPSQKIQLI	1	1	10	1	21
P209	U24616	1a	ATYTTGGSAAKTAHRLASFFTVGPKQDIQLI	1	1	15	1	16
P210	X76414	1b	DTYATGAATSRPRTALLTFFTPGSRQNQLI	1	1	24	1	23
P211	S79536	1b	TTHVSGGSAAHTTSGFASLFTQGASQKIQLI	1	1	53	1	31
P212	U24619	1b	STYSMGGAAAHNARGLTSLFSSGASQRIQLI	1	1	223	1	34
P213	U56520	1a	ETTVTGGSAAARDAAGLVGLFSSGAKQNQLQ	0	0	0	1	9
P214	S72417	1b	DTYTTGAAQGRTVSRLTSIFTPGASQNIQLI	0	0	0	1	21
P215	S74282	1b	RTYTTGGQAGFTVRGFTSFFTAGPQQKIQLI	0	0	0	1	1
P216	U24597	1b	GTHVLGGQAARGTQGIASLFSFGASQKIQLV	0	0	0	1	17
P217	U24602	1a	ETHTTGGSAAARATFGIANFFTPGAKQNIQLI	0	0	0	1	4
P218	S74294	1b	YTYTSGGRAGETTYSLTSLFTPGASQNIQLI	0	1	91	1	14
P219	U24614	1a	DTHVTGGTAARATSTLVSLSPGPKQDIQLI	0	0	0	1	9
P220	U24600	1a	ETYTTGGSAAARATSVLTSLFTPGAKQDIQLI	0	0	0	1	4
P221	U24259	1a	DTYTTGGSVAQATSSFVRLFAPGPKQNIQLI	0	0	0	1	5
P222	X92451	1a	TTYTTGGSAAKAILGITGLFSPGPRQNIQLM	0	0	0	1	4
P223	S74281	1b	GTYYTTGGQAGSTVRGLTSFFAVGPQQKIQLI	1	1	130	1	8
P224	S74280	1b	VTYVTGGQAGSTIKGLTSIFASGPQQKIQLI	0	0	0	1	3
P225	U24606	1a	ETHTTGGSVARTAYGLAGFFSQGAKQDIQLI	0	0	0	1	1
P226	S74295	1b	HTRTTGAQAARTVSTITSFFSPGASQNIQLI	0	0	0	1	7
P227	X92452	2c	TTYTTGGSAAKAILGITGLFAPGPRQNIQLM	0	0	0	1	6
P228	U24613	1a	ATYTTGGSAAARTASTLAGLFSFGAKQDIQLI	1	1	29	1	14
P229	S74288	1b	TTYTTGGQAGKTVNALTSLFAPGASQNIQLI	0	1	4	1	8
P230	S74291	1b	HTFTTGGQAGRTVSTITSFAPGASQNIQLI	0	1	16	1	8
P231	U24615	1a	HTHTTGGVAARTTKAFTGFFSQGAKQDIQLI	0	1	13	1	4
P232	U24604	1a	ATHTTGGTAAHTTSLTSLFSPGAKQNVQLI	0	0	0	1	13
P233	U24622	1b	ATRVSGGSPARGAAGLASLFSFGASQKIQLV	0	0	0	1	2
P234	HCU24601	2a	TTYTTGGSAAARAASSTIANLFTPGAKQDIQLI	0	0	0	1	8
P235	S72416	1b	HTRTTGAQAGRTVSTITSFFAPGASQNIQLI	1	1	12	1	11
P236	U24258	1a	DTYTTGGSVAHATSVLVKLFAPGPKQNIQLI	0	0	0	1	5
P237	S74287	1b	ETYVTGGQAGRTAQGFTSLFAPGASQDIQLV	0	0	0	1	18
P238	U24603	1a	VTHVTGGSAAQAAGFTSLFSPGPKQDIQLV	0	0	0	1	12
P239	S74290	1b	HTHTTGAQAGRTVSTITSFFAPGASQNIQLI	0	0	0	1	9
P240	U24599	1a	GTHVTGGSAAARDAFQFTSLFSQGAKQDIQLI	1	1	255	1	11
P241	S74283	1b	ETHTIGGQAASSTVRAFTSLFTAGPQQKIQLI	0	0	0	1	7
P242	S74289	1b	RTYTTGGRAGGTAYSFTSLFTPGASQNIQLI	0	0	0	1	11
P243	S74286	1b	DTYTTGGQAGKTVSTFTSLFAPGASQNIQLI	0	0	0	1	8
P244	U24610	1b	GTTVVGGQQAAYTTNRLASLFTLGSHQKIQLI	1	1	166	1	3

P245	U24617	1a	ETYTTGGAVASTARGLANLFMPGPKQNIQLV	0	0	0	1	6
P246	U24256	1b	TTRVSGGEAARNAHGLASLFSSGASQKIQLV	1	1	21	1	26
P247	U24618	1a	ETHVTGGSVARATNGLAGLFAPGAKQNIQLI	1	1	163	1	21
P248	S74279	1b	ETRVTTGGQAGSTIKGLTSIFASGPQQKIQLI	0	0	0	1	7
P249	X92459	3a	GTYTGGTAARTVSGFTSLFSAGANQNLQLI	0	0	0	1	8
P250	D88473	6a	WTTTVGSASVSSNVRGLTSLFTPGPRQNLQLI	0	0	0	1	16
P251	S74271	1b	DTPVWGGTAGGQVSGVVSFFSPGPSQQIQLV	1	1	8	1	12
P252	X92453	1a	ETYTTGGSVAQAASLTSLFSPGAKQDIQLV	0	0	0	1	11
P253	X92457	3a	GTYTGGTAARTVSGFTSLFSVGANQNLQLI	0	0	0	1	9
P254	S78727	2a	TTHVTGAQAGRTTSALATLFSHGSKQNIQLI	0	0	0	1	9
P255	X92461	3a	GTYTGGTAHTVSGTLGLFSIGAKQNLQLI	0	0	0	1	40
P256	D16568	1b	STRTTGQVAGHGASSLVSLFTPGASQKIQLV	0	0	0	1	15
P257	X92458	3a	GTYTGGTAAHTVSGFTSLFSVGAKQNLQLI	1	1	216	1	15
P258	D16565	1b	GTRTTGQVAGHGASSLVSLFTPGASQKIQLV	1	1	24	1	9
P259	S74275	1b	HTRVTGGTVGGQVAGFVSLLSAGPSQKIQLV	0	0	0	1	9
P260	S74272	1b	HSRVTGGTVGGQVSGFVSLLSAGPSQKIQLV	0	0	0	1	7
P261	D49756	2i	ARTLVANTVAQGAAGLSSLFKAGARQNIQLI	0	0	0	1	6
P262	N/A	1a	ETHVTGGSVGRTVSGLASLFTPDAKQNIQLI	0	0	0	1	23

Table S2. Mean diameter of samples. The right hand of the table shows the p value of the pairwise t-test of equality of means.

	<b>Total sample s</b>	<b>Unique aa sequences</b>	<b>Mean</b>	<b>S.D.</b>	<b>Acute</b>	<b>Chronic</b>	<b>Follow-up</b>
<b>Acute</b>	23	90	9.8532	12.4542			
<b>Chronic</b>	89	799	17.2699	21.9667	0.0375		
<b>Follow-up</b>	27	968	18.8686	17.6907	0.0460	0.7305	
<b>Genotype</b>	8	4615	107.698	11.073	0.0001	0.0001	0.0001

Table S3. Sequences of the unrelated peptides used in the cross-immunoreactivity experiment.

Peptide ID	Sequence	Source
1057	TRKTIKKLEEENPWLGNIVG	Hepatitis Delta Virus
1058	EEENPWLGNIVGIIRKGKDG	Hepatitis Delta Virus
1059	NPWLGNIVGIIRKGKDGEGA	Hepatitis Delta Virus
1060	GNIVGIIRKGKDGEGAPPAK	Hepatitis Delta Virus
1151	GGHRIYSSGG	Hepatitis Delta Virus
1157	GGLPRVVDGG	Hepatitis Delta Virus
1163	GEARDSHGIY	Hepatitis Delta Virus
1166	HGIYSSQPTL	Hepatitis Delta Virus
1167	IYSSQPTLPPGG	Hepatitis Delta Virus
1170	TLPLPRVVD	Hepatitis Delta Virus
1173	GEARDSHRIY	Hepatitis Delta Virus
1177	ARDSHGIYSSQPTLPLPRV	Hepatitis Delta Virus
1182	ERPGAASSPACKESRSPPLL	Hepatitis Delta Virus
2340	YAMTRIRDTLHLVECPTPAI	Hepatitis G virus
2343	GAGLTGGFYEPLVRRCSKLM	Hepatitis G virus
2344	RRCskLMGSRNPVCPGFAWL	Hepatitis G virus
2367	VPTDAKTTTEPPVPAKGVF	Hepatitis G virus
2371	SVVICDECHSHDSTVLLGIG	Hepatitis G virus
2381	LRTVPASAEslMQRRGRtGR	Hepatitis G virus
2385	MCRETlSPGSPDDPQWAGLK	Hepatitis G virus
2386	GPSDDPQWAGLGPNPVPLL	Hepatitis G virus



2387	PVPLLLRWGNDLPSKVAGHH	Hepatitis G virus
2390	TDWDVKGGGAPLYRHGDQAT	Hepatitis G virus
2391	GGAPLYRHGDQATPQPVVQV	Hepatitis G virus
2392	LYRHGDQATPQPVVQVPPVD	Hepatitis G virus
2393	TPQPVVQVPPVDHRPGGESA	Hepatitis G virus
2398	AYGASRSPPLAAAASYLMGL	Hepatitis G virus
2401	RRPLARMSSEYSDPMASAIG	Hepatitis G virus
2402	LTCAFRGGGTPSDPVWCQVH	Hepatitis G virus
2403	VHGNYYKFPLDKLPNIIVAL	Hepatitis G virus
2404	LRVTADTTKTKEAGKVLSD	Hepatitis G virus
2405	KLPGLAVHRKKAGALRTRML	Hepatitis G virus
2407	PGLRLPPPEIAGIPGGFPLS	Hepatitis G virus

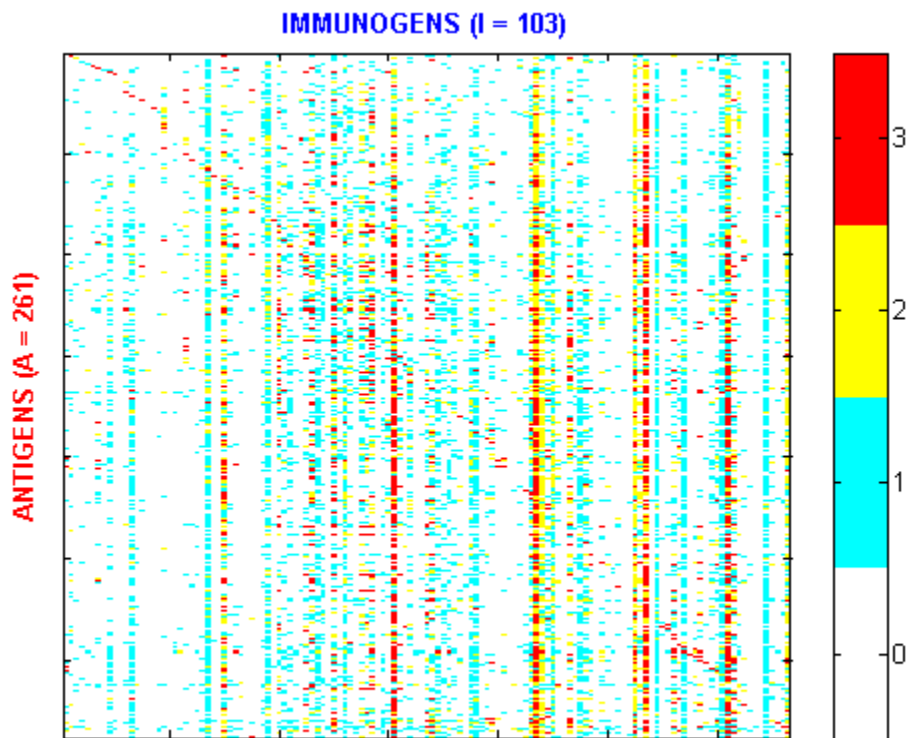


Fig. S1. **Experimental cross-immunoreactivity.** The color indicates the number of mice with a positive reaction between each tested immunogen and antigen. The percentage of cross-immunoreactive reactions with one or more mice is 18.74%, with two or more mice

is 7.09% and with three or more mice is 2.43%. The average agreement between mice immunized with the same BSA-conjugated peptide was 88.78 (S.D. 14.98).

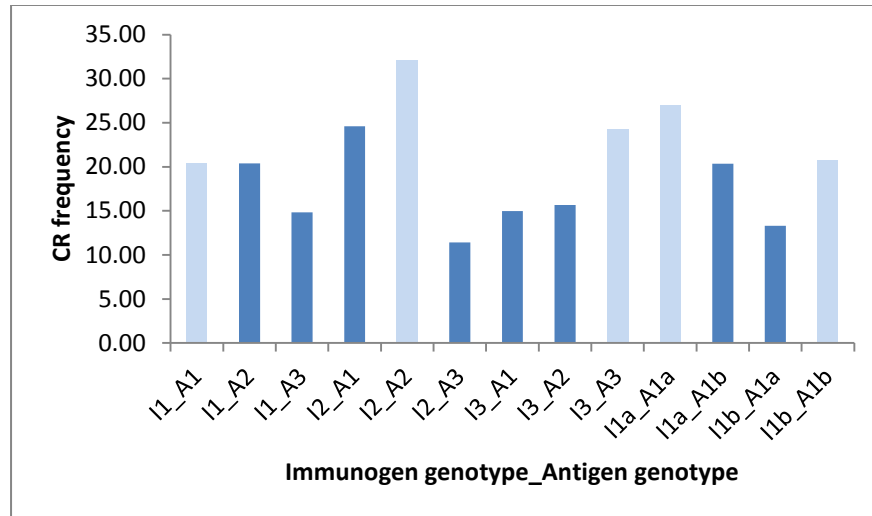


Fig. S2. **Cross-immunoreactivity frequency among genotypes.** The cross-immunoreactivity frequency is equal to the number of positive reactions divided by the total number of pairwise comparisons among tested peptides, expressed as a percentage. Light blue bars correspond to within-genotype comparisons, dark blue bars correspond to between-genotype comparisons. Data are shown only for genotypes 1 (subtypes 1a and 1b), 2 and 3, which are the most represented among the tested HVR1 variants.

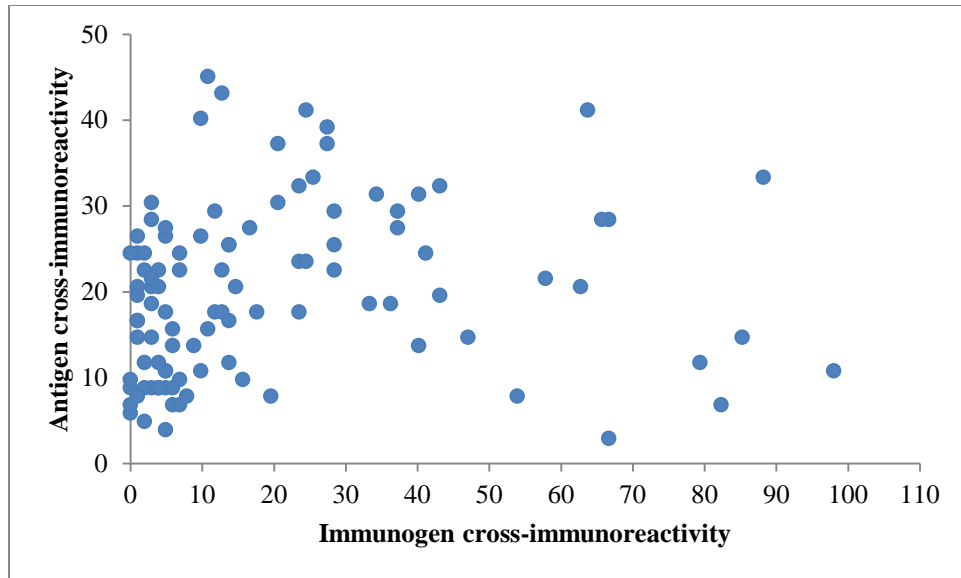


Fig. S3. **Scatterplot of the cross-immunoreactivity among HVR1 immunogens and antigens.** The correlation between the two variables is 0.1468 ( $p = 0.1409$ ).

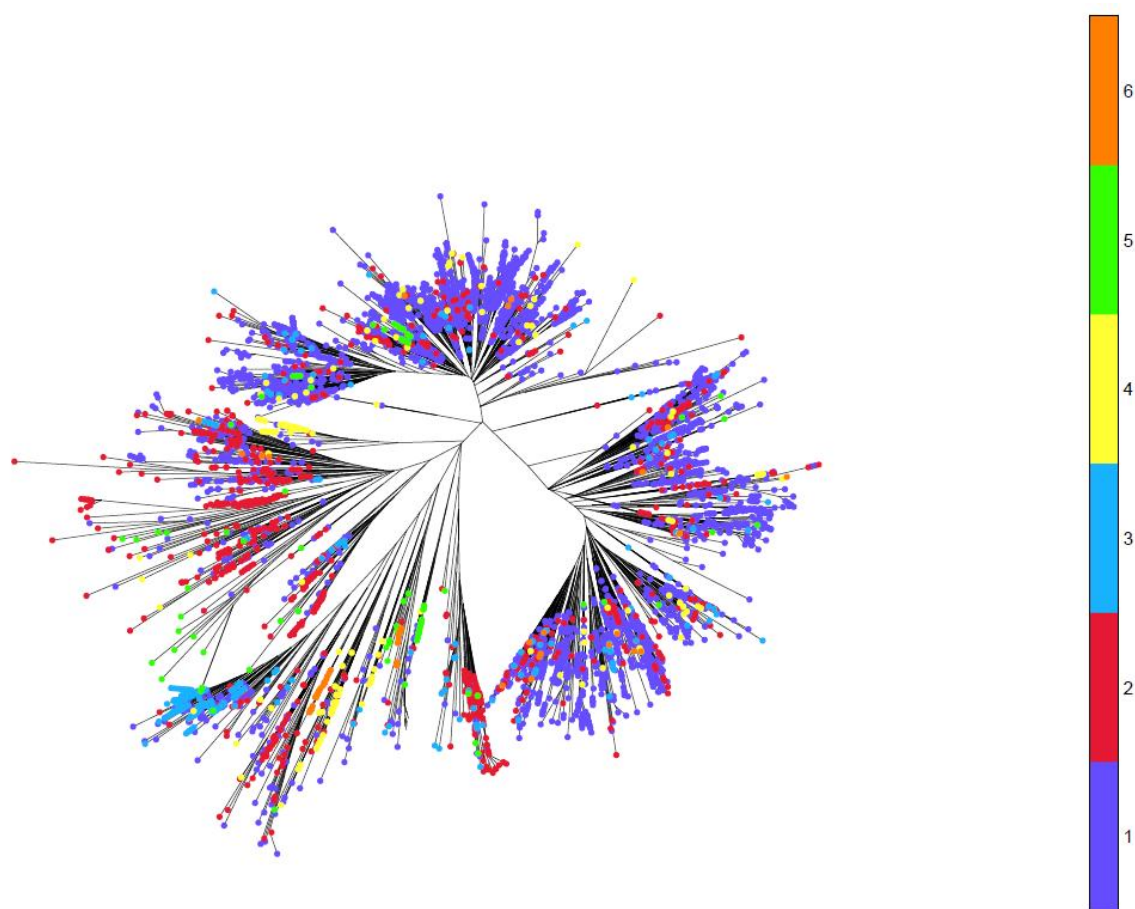
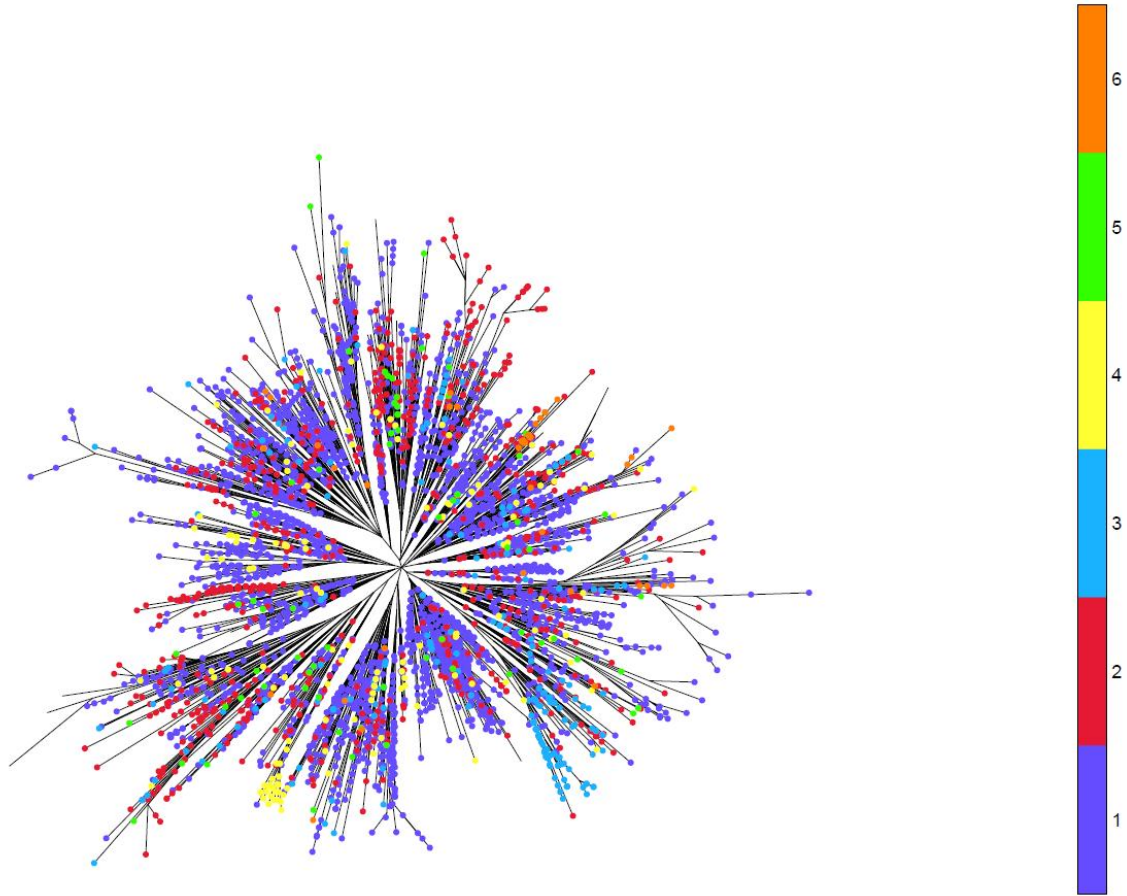


Fig. S4. **Maximum likelihood tree of 6018 HVR1 nucleotide sequences.** The tree was built using HyPhy<sup>1</sup> with the GTR model, global variation of branches and observed nucleotide frequencies. Each genotype is shown in a different color.



**Fig. S5. Maximum likelihood tree of 4757 HVR1 aa sequences.** The tree was built using HyPhy<sup>1</sup> with the JTT model<sup>2</sup>, global variation of branches and observed aa frequencies. Each genotype is shown in a different color.

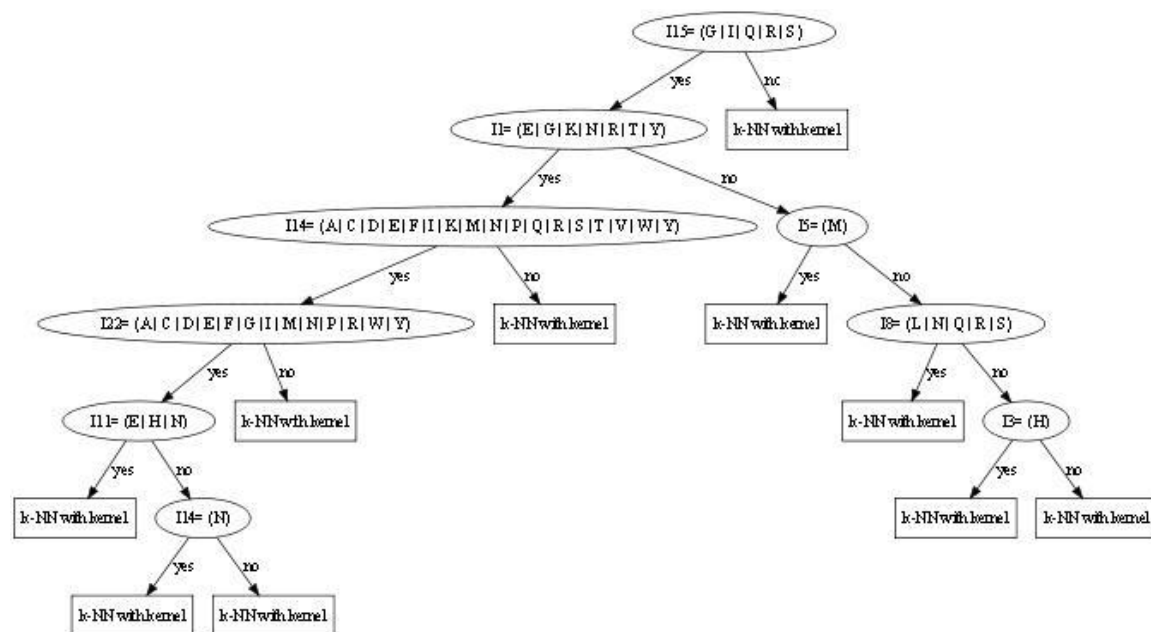


Fig. S6. Decision tree for the prediction of cross-immunoreactivity.

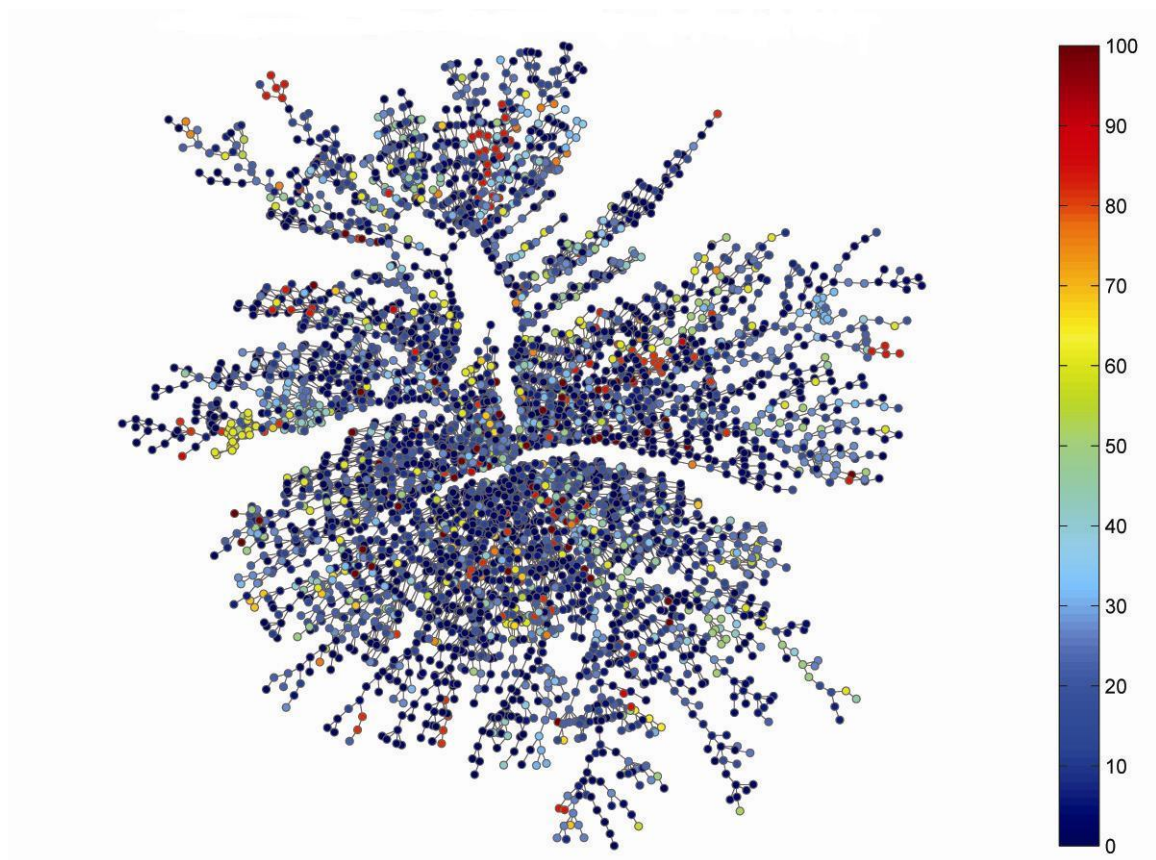


Fig. S7. **Predicted immunogen cross-immunoreactivity among 4,757 HVR1 variants.**  
The color bar corresponds to the percent immunoreactivity.



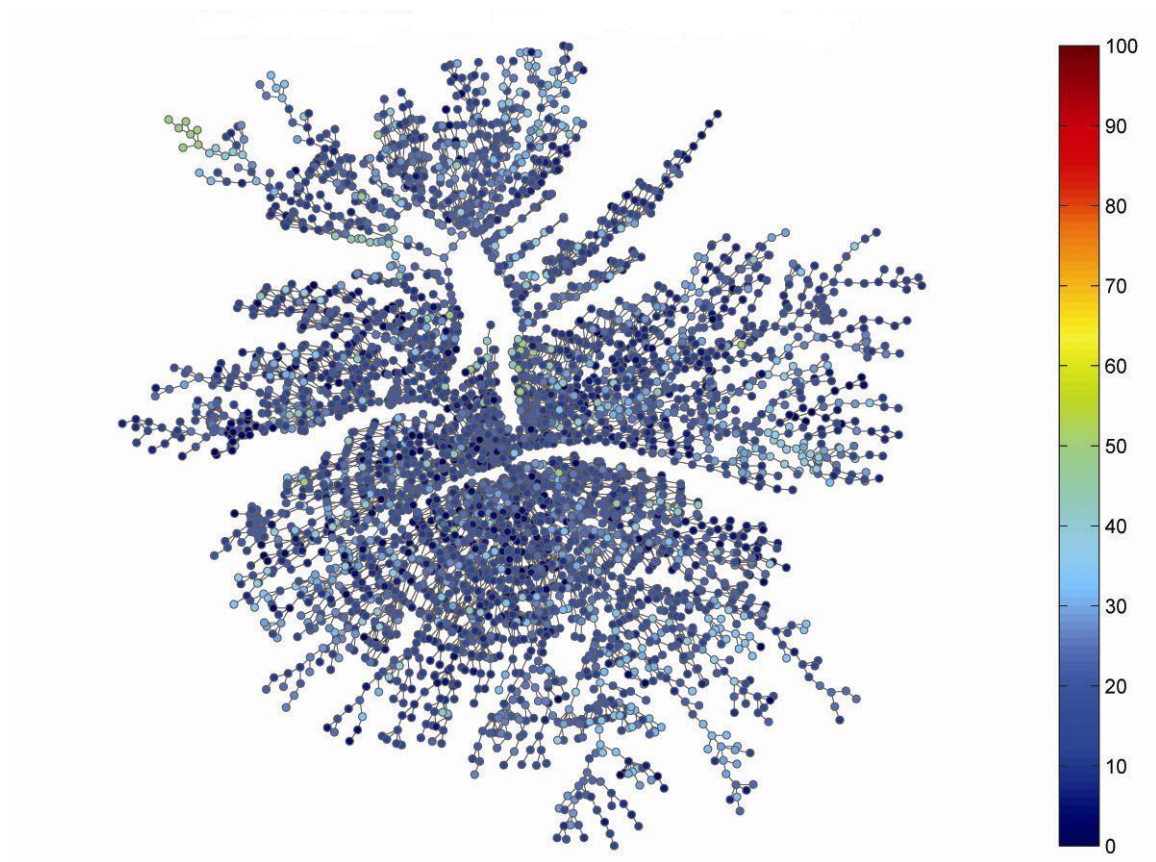


Fig. S8. **Predicted antigen cross-immunoreactivity level among 4,757 HVR1 variants.** The color bar corresponds to the percent immunoreactivity.

## References

- 1 Kosakovsky, S., Frost, S. & Muse, S. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
- 2 Jones, D., Taylor, W. & Thornton, J. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275-282 (1992).