Selection for immune evasion in SARS-CoV-2 revealed by high-resolution epitope mapping combined with viral genome sequence analysis



20th April 2022 Jörg H. Fritz & Ciriaco A. Piccirillo McGill University – Dept. of Microbiology & Immunology

SARS-CoV 2 Structure



Adaptive Immunity to SARS-CoV-2 (T and B cells)





Viral evolution of SARS-CoV-2

(How to generate an assessment tool for public health risk of newly arising SARS-CoV-2 variants)





Amino acid position

High-Density Peptide Array (HDPA) (PEPperPRINT)

SARS-CoV-2 proteome-wide IgG and IgA epitope mapping



Example of raw data









SARS-CoV2-specific Peptides

| | Spike | Nucleocapsid | Envelope | Membrane | ORF1ab | TOTAL |
|---------------------|-------|--------------|----------|----------|--------|-------|
| SARS-CoV-2 negative | 119 | 41 | 1 | 47 | 294 | 502 |
| SARS-CoV-2 positive | 195 | 69 | 6 | 29 | 549 | 848 |
| Overlap | 90 | 35 | 6 | 17 | 353 | 501 |
| Total | 404 | 145 | 13 | 93 | 1196 | 1851 |

Comparison HDPA (PEPperPRINT) with prediction tools (Bepipred, DiscoTope)



Spike protein



R

R

S

С

S

W

S

С

G

S

D

G

Score

1

D

Q

S

D

esidues)

361

421

Α













ORF1A

MESLVPGFNENTHVQLSLPVLQVR S V E E V L S E A R O H L K D G T Q Y G R S G E T L G V L V P C N Q M C L S T L W K C D H C G E T S W Q T A C H N S E Y G P E H S L A C Y H N E S G L DWSYSGQSTOLG. K V F T T V D N I N S H T Q V V D M S M T V G Y Y H T T D P S F L G B Y M S A L H H T H A L Q D A Y Y R A R A G E A A N F G A L CGGGGTTLKG STH PVETSNS RFU Ratio LTIKXTNELSBY CYNYMPYPPTEL 2101 5.10 STAALOV >30 TKSSLPINVIVFE VIGOSCNNYM ESLRPDTRYVLMDDSIIGFPN 4201 G T G T G T J T E L E P P C R F V T D T P K G 4201 G T G T G T J T E L E P P C R F V T D T P K G 4261 F G G A S C G L Y C M C M T J M P N P K G F O I P T T C A N O P V O F T L K N T V C T V C O M W K O V C 4381 CSCDOLREPMLOSADAGSFLNRS 10 20

4406 V S A A R L T P C G T G T S T D V V Y R A F D I Y N D K V A G F A K F L K T N C C R F Q E K D E D D N L I D S Y F 4466 RHT FSNYQHEETIYNLLKDCPAVAKHDFFKFRTDGDMVPHISRQRLTKYTMAD 4526 FDEGNCDTLKEILVTYNCCDDDYFNKKDWYDFVENPDIL<mark>RVYA</mark>NLGERVROALLKTVOFC 4586 DAMRNAGIVGVLTLDNQDLNGNWYDFGDFIQTTPGSGVPVVDSYYSLLMPILTLTRA 4646 ESHVDTDLTKPYIKWDLLKYDFTEERLKLFDRYFKYWDQTYHPNCVNCLDD 4706 N V L F S T V F P P T S F G P L V R K I F V D G V P F V V S T G Y H F R E L G V V H N Q D V N L H S S R L 4766 YAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFOTVKPGNFNKDFYDFAV 4826 SVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLLFVVEVVDKYFDCYDGG 4886 I V N N L D K S A G F P F N K W G K A R L Y Y D S M S Y F D O D A L F A Y T K R N V I P T I T O M N L K Y A 4946 A R T V A G V S I C S T M T N R Q F H Q K L L K S I A A T R G A T V V I G T S K F Y G G W H N M L K 5006 LMGWDYPKCDRAMPNMLRIMASLYLARKHTTCCSLSHRFYRLANECAQVLS Y V K P G G T S S G D A T T A Y A N S V F N I C Q A V T A N V N A L L S T D G N K I A D K Y V R N L Q 5066 5126 N R D V D T D F V N E F Y A Y L R K H F S M M I L S D D A V V C F N S T Y A S Q G L V A S I K N F K S V L Y 5186 N S E A K C W T E T D L T K G P H E F C S Q H T M L V K Q G D D Y V Y L P Y P D P S R I L G A G C F LMIERFVSLAIDAYPLTKHPNOEYADVFHLYLQYIRKLHDELTGHMLDMY 5246 5306 R Y W E P E F Y E A M Y T P H T V L Q A V G A C V L C N S Q T S L R C G A C I R 5366 V L S V N P Y V C N A P G C D V T D V T O L Y L G G M S Y Y C K S H K PISFPLCANGOVFGL GSD 5486 ELHLSWEVGKPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGT 5426 NVTDFNAIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVRE 5546 G D Y F Y L T S H T Y M P L S A P T L Y P Q E H Y Y R I T G L Y P T L N I S D E F **RFU Ratic** 5606 GPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEKALKYLPI<mark>DK</mark>CSRIIPAR RVE e 5666 C F D K F K V N S T L E Q Y V F C T V N A L P E T T A D I V V F D E I S M A T N Y D L S V V N A R L R A K H 2-5 5726 PAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTV L 5726 5786 5786 5-10 YKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVF >30 5846 ASKILGLPTOTVDSSOGSEVDYVIFTOTTETAHSCNVNRFNVAITRAKV NA 5906 Y D K L Q F T S L E I P R R N V A T L Q A E N V T G L F K D C S K V I T G L H P T Q A P T H L S V D \$ 5966 DIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHYRAWIGF LOLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAK PPPGDQFKHLIP 6026 6086 I V O M L S D T L K N L S D R V V F V L W A H G F E L T S M K Y F V K I G P E R T C C L C D R R A T C 6146 S D T Y A C W H H S I G F D Y V Y N P F M I D V Q Q W G F T G N L Q S N H D L Y C Q V H G N A H A V H E C F V K R V D W T I E Y P I I G D E L K I N A A C R K V Q H M V V K A A L L A D K F P V L H D 6206 6266 Y P Q A D V E W K F Y D A Q P C S D K A Y K I E E L F Y S V A T H S D K F T D G V 6326 F D T R V L S N L N L P G C D G G S L Y V N K H A F H T P A F D K S A F V N L K Q L P F F Y Y S D S P C S D I D Y V P L K S A T C I T R C N L G G A V C R H H A N E Y R L Y L D A Y N M M I S A G F S L W V 6386 6446 N T F T R L Q S L E N V A F N V V N K G H F D G Q Q G E V P V S I I N N T V Y T K V D G V D V E L F 6506 A F E L W A K R N I K P V P E V K I L N N L G V D I A A N T V I W D Y K R D A P A H I S T I G V C S TICAPLTVFFDGRVDGOVDLFRNARNGVLITEGSVKGLQPSVGPKQASL 6566 6626 T Q F N Y Y K K V D G V V Q Q L P E T Y F T Q S R N L Q E F K P R S Q M E I D F L E L A M D E F I 6686 HIVYGDFSHSQLGGLHLLIGLAKRFKESPFELEDFIPMDSTVKNYF 6746 VIDLLLDDFVEIIKSQDLSVVSKVVKVTIDYTEISFMLWCKDGH 6806 G V A M P N L Y K M Q R M L L E K C D L Q N Y G D S A T L P K G I M M N V A K Y T Q L C Q Y L N T L T L A V P I H F G A G S D K G V A P G T A V L R Q W L P T G T L L V D S D L N D F V S D A D S T L I G D C A T V 6866 6926 ISDMYDPKTKNVTKENDSKEGFFTYICGFIQQKLALGGSVAIKITEHSWNADLYKLMGHF 6986 A WWTAFVTNVNASSSEAFLIGCNYLGKPREQIDGYVMHANYIFWRNTNPIOLSSYSLFDM 7046 SKFPLKLRGTAVMSLKEG<mark>OT</mark>NDMILSLLSKGRLII<mark>RENNRVVISSDVLVN</mark>N 10 30 40 50 60 20

ORF1B



The Case for Pre-existing Immunity (asymptomatic infections)

Cite as: R. Li *et al.*, *Science* 10.1126/science.abb3221 (2020).

Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)

Ruiyun Li^{1*}, Sen Pei^{2*+}, Bin Chen^{3*}, Yimeng Song⁴, Tao Zhang⁵, Wan Yang⁶, Jeffrey Shaman²⁺

¹MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Faculty of Medicine, Imperial College London, London W2 1PG, UK. ²Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA. ³Department of Land, Air and Water Resources, University of California, Davis, Davis, CA 95616, USA. ⁴Department of Urban Planning and Design, The University of Hong Kong, Hong Kong. ⁵Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 10084, P. R. China. ⁶Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA.

*These authors contributed equally to this work.

+Corresponding author. Email: sp3449@cumc.columbia.edu (S.P.); jls106@cumc.columbia.edu (J.S.)

Estimation of the prevalence and contagiousness of undocumented novel coronavirus (SARS-CoV2) infections is critical for understanding the overall prevalence and pandemic potential of this disease. Here we use observations of reported infection within China, in conjunction with mobility data, a networked dynamic metapopulation model and Bayesian inference, to infer critical epidemiological characteristics associated with SARS-CoV2, including the fraction of undocumented infections and their contagiousness. We estimate 86% of all infections were undocumented (95% CI: [82%–90%]) prior to 23 January 2020 travel restrictions. Per person, the transmission rate of undocumented infections was 55% of documented infections ([46%–62%]), yet, due to their greater numbers, undocumented infections were the infection source for 79% of documented cases. These findings explain the rapid geographic spread of SARS-CoV2 and indicate containment of this virus will be particularly challenging.

Science. 2020 May 1;368(6490):489-493

The Case for Pre-existing Immunity (asymptomatic infections)

21.7 - 85%

Table 2. Antibody Testing

| Study or Report | Tested, n* | Random Sampling* | SARS-CoV-2- Positive, <i>n</i> (%) | Asymptomatic, n (%) |
|---|---------------|---------------------|---------------------------------------|------------------------|
| England residents (55) | 365 104 | Yes | 17 576 (4.8) | 5694 (32.4) |
| Spain residents (56) | 61 075 | Yes | 3053 (5.0) | 1008 (33.0) |
| Detroit, Michigan, hospital staff (57) | 20 614 | No | 1818 (8.8) | 798 (43.9) |
| Wuhan, China, hospital staff (58) | 8553 | No | 424 (5.0) | 148 (34.9) |
| Bavaria, Germany, children aged 1-18 y (59) | 4859 | Yes | 47 (1.0) | 22 (46.8) |
| Louisiana residents (60) | 4778 | Yes | 311 (6.5) | 147 (47.3) |
| Munich, Germany, hospital staff (61) | 4554 | No | 108 (2.4) | 28 (25.9) |
| Cairo, Egypt, hospital staff (62) | 4040 | No | 170 (4.2) | 116 (68.2) |
| Health care personnel at 13 U.S. medical centers (63) | 3248 | No | 194 (6.0) | 56 (28.9) |
| Maranhão, Brazil, residents (64) | 3156 | Yes | 1167 (37.0) | 320 (27.4) |
| Ischgl, Austria, residents (65) | 1473 | No | 622 (42.2) | 529 (85.0) |
| Wuhan dialysis patients (66) | 1027 | No | 99 (9.6) | 50 (50.5) |
| Buenos Aires, Argentina, residents (67) | 873 | No | 466 (53.4) | 396 (85.0) |
| Connecticut residents (68) | 567 | Yes | 23 (4.1) | 5 (21.7) |
| Sweden nursing home staff (69) | 459 | No | 86 (18.7) | 40 (46.5) |
| London, England, dialysis patients (70) | 356 | No | 129 (36.2) | 52 (40.3) |
| Nashville, Tennessee, hospital staff (71) | 249 | No | 19 (7.6) | 8 (42.1) |
| London maternity unit staff (72) | 200 | No | 29 (14.5) | 10 (34.5) |

SARS-CoV-2 = severe acute respiratory syndrome coronavirus 2.

* Boldface indicates details that increase the likelihood of higher-quality evidence.

Ann Intern Med. 2020 Sep 1;173(5):362-367

The Case for Pre-existing Immunity (seasonal human Coronaviruses)



Fig. 2. Phylogenetic relationships in the subfamily *Coronavirinae*. Bayesian phylogeny of an 816-nucleotide *RNA-dependent RNA polymerase* fragment, as described previously (Drexler et al., 2010) of the subfamily *Coronavirinae* using MrBayes V3.1 (Ronquist and Huelsenbeck, 2003) under assumption of a GTR + G + I substitution model, using 2,000,000 trees sampled every 100 steps, annotated with a burn-in of 25% using TreeAnnotator V1.7.4 and visualized using FigTree V1.4 from the BEAST package (Drummond et al., 2012). Cavally virus (Zirkel et al., 2011) was used as an outgroup. Values at deep nodes indicate statistical support from Bayesian posterior probabilities, scale bar genetic distance.



Fig. 1 Summary diagram of the animal groups representing natural hosts and the putative intermediate hosts for the six CoVs found in humans.

Adv Virus Res. 2018;100:163-188

The Case for Pre-existing Immunity (seasonal human Coronaviruses) NTD RBD S1/S2 HR2 а FP HR1 SARS-CoV-2 SARS-CoV **MERS-CoV** similarity to OC43-CoV HKU1-CoV % NTD RBD S1/S2 HR2 FP HR1 b % identity to SARS-CoV-2 SARS-CoV **MERS-CoV** OC43-CoV HKU1-CoV > 75 > 50 > 25 < < 25 = 50 amino acids

High-Density Peptide Array (HDPA) (PEPperPRINT)

Seasonal coronavirus (hCoV) proteome-wide IgG and IgA epitope mapping (OC43, NL63, 229E, HKU1)



В

How conserved are identified epitopes? (seasonal human Coronaviruses)

OC43-specific Peptides

| | Spike | Nucleocapsid | Envelope | Membrane | ORF1ab | TOTAL |
|---------------------|-------|--------------|----------|----------|--------|-------|
| SARS-CoV-2 negative | 126 | 37 | 6 | 4 | 293 | 466 |
| SARS-CoV-2 positive | 209 | 70 | 10 | 27 | 508 | 824 |
| Overlap | 104 | 35 | 2 | 8 | 280 | 429 |
| Total | 439 | 142 | 18 | 39 | 1081 | 1719 |

HKU1-specific Peptides

| | Spike | Nucleocapsid | Envelope | Membrane | ORF1ab | TOTAL |
|---------------------|-------|--------------|----------|----------|--------|-------|
| SARS-CoV-2 negative | 104 | 43 | 7 | 17 | 293 | 464 |
| SARS-CoV-2 positive | 220 | 105 | 7 | 23 | 503 | 858 |
| Overlap | 90 | 35 | 2 | 10 | 254 | 391 |
| Total | 414 | 183 | 16 | 50 | 1050 | 1713 |

NL63-specific Peptides

| | Spike | Nucleocapsid | Envelope | Membrane | ORF1ab | TOTAL |
|---------------------|-------|--------------|----------|----------|--------|-------|
| SARS-CoV-2 negative | 139 | 54 | 12 | 17 | 296 | 518 |
| SARS-CoV-2 positive | 183 | 77 | 8 | 24 | 571 | 863 |
| Overlap | 70 | 56 | 3 | 14 | 269 | 412 |
| Total | 392 | 187 | 23 | 55 | 1136 | 1793 |

229E-specific Peptides

| | Spike | Nucleocapsid | Envelope | Membrane | ORF1ab | TOTAL |
|---------------------|-------|--------------|----------|----------|--------|-------|
| SARS-CoV-2 negative | 116 | 43 | 3 | 15 | 306 | 483 |
| SARS-CoV-2 positive | 158 | 99 | 7 | 38 | 592 | 894 |
| Overlap | 72 | 46 | 2 | 12 | 325 | 457 |
| Total | 346 | 188 | 12 | 65 | 1223 | 1834 |

How conserved are identified epitopes? (seasonal human Coronaviruses)

| HCoV-229E HCoV-NL63 HCoV-NL63 HCoV-HKU1 HCoV-OC43 Sars-Cov-2 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCoV-DC43 HCOV- |
|--|
| the burght have the for the balance of the former |
| athe is a det a che to the second and the second at a |
| <u>ellerete</u> <u>state de lettele</u> |

To evaluate conservation of epitopes:

- Aligned protein sequences of viral strains and calculated conservation score based on physico-chemical properties

- Defined cross-reactivity per amino acid sites within 15-mer peptides

Sites with conservation score > 6 for which Ab response
to SARS-CoV-2 and at least one shCoV were considered cross-reactive

- ~ 27% of the pool of detected epitope sites are cross-reactive
- Local alignment of HDPA response for S protein of all 5 viruses shown on the left

Are cross-reactive epitope sites that particularly important for the humoral immune response after exposure to SARS-CoV-2?





CORONAVIRUS

Viral epitope profiling of COVID-19 patients reveals cross-reactivity and correlates of severity

Ellen Shrock*, Eric Fujimura*, Tomasz Kula†, Richard T. Timms†, I-Hsiu Lee, Yumei Leng, Matthew L. Robinson, Brandon M. Sie, Mamie Z. Li, Yuezhou Chen, Jennifer Logue, Adam Zuiani, Denise McCulloch, Felipe J. N. Lelis, Stephanie Henson, Daniel R. Monaco, Meghan Travers, Shaghayegh Habibi, William A. Clarke, Patrizio Caturegli, Oliver Laeyendecker, Alicja Piechocka-Trocha, Jonathan Z. Li, Ashok Khatri, Helen Y. Chu, MGH COVID-19 Collection & Processing Team, Alexandra-Chloé Villani, Kyle Kays, Marcia B. Goldberg, Nir Hacohen, Michael R. Filbin, Xu G. Yu, Bruce D. Walker, Duane R. Wesemann, H. Benjamin Larman, James A. Lederer, Stephen J. Elledge‡

- analyzed if the humoral immune response to SARS-CoV-2 epitopes correlated with the number of cross-reactive epitopes identified.
- to what extent is the response to SARS-CoV-2 predictable based on cross-reactivity to other endemic hCoVs?
- We defined cross-reactive epitopes as peptide sequences with at least five cross-reactive epitope sites
 - 16 epitopes being cross-reactive

Viral Genome Sequencing



Metric value is lower in epitope sites (FDR-adjusted Wilcoxon test p < 0.05)
Metric value is higher in epitope sites (FDR-adjusted Wilcoxon test p < 0.05)

- We tracked the evolution of identified SARS-CoV-2 B cell epitopes using single nucleotide variants (SNVs) identified in 38,685 SARS-CoV-2 genome sequences from the NCBI sequence read archive (Wave 1: 01-07/2020; Wave 2: 08-12/2020) sequenced using Illumina paired-end amplicons with a minimum average depth of coverage of 200x and fewer than 10,000 sites with a depth of coverage lower than 100x. Combined with additional filters to remove sequencing errors

es? - Such deep coverage allowed us to identify SNVs that are polymorphic within patients, reflecting withinpatient evolution, as well as those that are shared between the consensus sequences of different patients.

- Mutations in epitope sites or non-epitope sites for within host and between host genomic viral sequences

Viral Genome Sequencing



Metric value is lower in epitope sites (FDR-adjusted Wilcoxon test p < 0.05)
Metric value is higher in epitope sites (FDR-adjusted Wilcoxon test p < 0.05)

These observations indicate that nonsynonymous substitutions in S and N epitope sites accumulate most rapidly upon transmission, rather than within patients.

Taken together these results support the notion that most of the selective pressure for immune evasion of SARS-CoV-2 occurs upon transmission between hosts, consistent with the asynchrony model

Assessing Immune Evasion Potential of SARS-CoV-2 Variants



Assessing Immune Evasion Potential of SARS-CoV-2 Variants



Acknowledgments

Carsten Haber and Volker Stadler, PEPperPRINT GmbH, Heidelberg, Germany

Arnaud N'Guessan and Jesse Shapiro, Department of Microbiology and Immunology, McGill University and McGill Genome Centre

Senthil Kumar Duraikannu Kailasam and Guillaume Bourque, Canadian Center for Computational Genomics, and Department of Human Genetics, McGill University

Raphael Poujol, Jean-Christophe Grenier, Fatima Mostefai, Julie Hussin, Research Centre, Montreal Heart Institute, and Département de Médecine, Université de Montréal

Paola Contini and Raffaele De Palma, Department of Internal Medicine, University of Genoa and IRCCS IST-Ospedale San Martino, Genoa, Italy

Jörg H. Fritz and Ciriaco A. Piccirillo, Department of Microbiology and Immunology, McGill University and McGill University Research Center on Complex Traits (MRCCT), McGill University

McGill Interdisciplinary Initiative in Infection and Immunity (MI4) for financial support for our study