BRIEF REPORT

REVISED Expected immune recognition of COVID-19 virus by memory from earlier infections with common coronaviruses in a large part of the world population [version 2; peer review: 2 approved]

Johannes M. Dijkstra, Keiichiro Hashimoto

Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Aichi, 470-1192, Japan

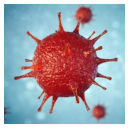
v2 **First published:** 23 Apr 2020, **9**:285
https://doi.org/10.12688/f1000research.23458.1
Latest published: 17 Jul 2020, **9**:285
https://doi.org/10.12688/f1000research.23458.2

Abstract

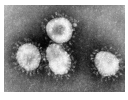
SARS-CoV-2 is the coronavirus agent of the COVID-19 pandemic causing high mortalities. In contrast, the widely spread human coronaviruses OC43, HKU1, 229E, and NL63 tend to cause only mild symptoms. The present study shows, by *in silico* analysis, that these common human viruses are expected to induce immune memory against SARS-CoV-2 by sharing protein fragments (antigen epitopes) for presentation to the immune system by MHC class I. A list of such epitopes is provided. The number of these epitopes and the prevalence of the common coronaviruses suggest that a large part of the world population has some degree of specific immunity against SARS-CoV-2 already, even without having been infected by that virus. For inducing protection, booster vaccinations enhancing existing immunity are less demanding than primary vaccinations against new antigens. Therefore, for the discussion on vaccination strategies against COVID-19, the available immune memory against related viruses should be part of the consideration.

Keywords

Coronavirus, COVID-19, SARS-CoV-2, OC43, HKU1, 229E, NL63, MHC class I, Immunology, Vaccination



This article is included in the Disease Outbreaks gateway.



This article is included in the Coronavirus collection.

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 2 (revision) 17 Jul 2020		
version 1 23 Apr 2020	✓ report	✓ report

1 **Anna Gil**, University of Massachusetts Medical School, Worcester, USA
Liisa K. Selin, University of Massachusetts Medical School, Worcester, USA

2 **Andrea Sant**, University of Rochester Medical Center, Rochester, USA
David H. Smith Center for Vaccine Biology and Immunology, Rochester, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Johannes M. Dijkstra (Dijkstra@fujita-hu.ac.jp)

Author roles: Dijkstra JM: Conceptualization, Investigation, Validation, Visualization, Writing – Original Draft Preparation; **Hashimoto K**: Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

How to cite this article: Dijkstra JM and Hashimoto K. **Expected immune recognition of COVID-19 virus by memory from earlier infections with common coronaviruses in a large part of the world population [version 2; peer review: 2 approved]** F1000Research 2020, **9**:285 https://doi.org/10.12688/f1000research.23458.2

First published: 23 Apr 2020, **9**:285 https://doi.org/10.12688/f1000research.23458.1

REVISED **Amendments from Version 1**

Based on requests of the reviewers, we now have added more background information and references to the article. In Table 1, we added software predictions for HLA-C binding. In the Notifications section, in line with the comments by the reviewers, we additionally included some new information that appeared after the initial submission of our article.

Any further responses from the reviewers can be found at the end of the article

Introduction

SARS-CoV-2 and other human coronaviruses

From the end of 2019, the world experienced the coronavirus disease 2019 (COVID-19) pandemic caused by the emerging severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2; aka 2019 novel coronavirus or 2019-nCoV). SARS-CoV-2 shares ~80% nucleotide identity with SARS-CoV-1 (aka SARS-CoV), the causative agent of the SARS epidemy from 2002, and is even more similar to some coronaviruses in bats (Andersen *et al.*, 2020; Ceraolo & Giorgi, 2020; Wu *et al.*, 2020; Zhou *et al.*, 2020). Coronaviruses are membrane-enveloped positive-strand RNA viruses with, for an RNA virus, a large genome of ~30 kb. That genome encodes several structural components of the virion including the nucleocapsid protein N and the membrane proteins S (spike), M, and E, plus also a number of nonstructural proteins involved in RNA replication and other—partly unknown—functions (Weiss & Navas-Martin, 2005). The coronaviruses infecting humans belong to the serological/phylogenetic clades group I (alphacoronaviruses) and group II (betacoronaviruses); group I includes HCoV-229E (human coronavirus 229E) and HCoV-NL63, while group II includes SARS-CoV-1, SARS-CoV-2, Middle East respiratory syndrome coronavirus (MERS-CoV), HCoV-OC43, and HCoV-HKU1. The viruses SARS-CoV-1 and MERS-CoV, on average, cause the most severe symptoms, and their outbreaks were successfully monitored and halted. At the other end of the spectrum, the viruses HCoV-229E, HCoV-NL-63, HCoV-OC43, and HCoV-HKU1 tend to cause only mild symptoms and are very common.

Prevalence and associated disease of the common human coronaviruses 229E, NL63, OC43, and HKU1

The Centers for Disease Control and Prevention (CDC; https://www.cdc.gov/coronavirus/general-information.html) states: "Common human coronaviruses, including types 229E, NL63, OC43, and HKU1, usually cause mild to moderate upper-respiratory tract illnesses, like the common cold. Most people get infected with one or more of these viruses at some point in their lives." The same agency lists the common symptoms caused by these viruses as runny nose, sore throat, headache, fever, cough, and general feeling of being unwell, but also explains that they occasionally cause lower-respiratory tract illnesses, such as pneumonia or bronchitis. The viruses 229E and OC43 have been known since the 1960s (reviewed in Kahn & McIntosh, 2005), but NL63 (van der Hoek *et al.*, 2004) and HKU1 (Woo *et al.*, 2005) were only (conclusively) identified following the rise in interest in coronaviruses in the wake of the SARS epidemy. These common coronaviruses are believed to be the second most common cause of the common cold (Mäkelä *et al.*, 1998). In the U.S.A., a 3-year RT-PCR surveillance of respiratory samples of patients revealed that the four viruses 229E, NL63, OC43, and HKU1 were present at levels varying by season and region, with all individual viruses peaking at >3% prevalence in each investigated region (Midwest, Northeast, South, West); co-infection with other coronaviruses was found in only ~2% of infected cases, but co-infection with another respiratory virus was found in a substantial ~30% of infected cases (Killerby *et al.*, 2018). This pattern was reminiscent of findings in the United Kingdom (Gaunt *et al.*, 2010) and Japan (Matoba *et al.*, 2015). Serological investigations in countries as diverse as the U.S.A. (Bradburne & Somerset, 1972; Dijkman *et al.*, 2012), China (Zhou *et al.*, 2013), and Qatar (Al Kahlout *et al.*, 2019), found that most healthy blood donors had antibodies against coronaviruses, supporting that these viruses are widespread indeed.

Since immune memory protection can be induced by related pathogens, as exemplified by the eradication of human smallpox virus (Variola) by immunization with a related "cowpox" virus (Vaccinia) (Plotkin & Plotkin, 2018), it is interesting to consider whether common human coronavirus infections may have induced some level of protection against SARS-CoV-2.

The possibility of matching linear epitopes between SARS-CoV-2 and the common human coronaviruses that may stimulate the immune system through MHC class I presentation

The major arms of immune memory concern antibody secretion by B cells, killing of infected cells by $CD8^+$ T cells, and helper/regulatory immune activities (e.g. cytokine secretion) by $CD4^+$ T cells. For a murine coronavirus infection in mouse, both antibody responses and cell-mediated cytotoxicity were needed to efficiently control the virus (reviewed by Weiss & Navas-Martin, 2005). In SARS-CoV-1-infected patients, B cell as well as T cell responses were observed (Li *et al.*, 2008), and, in animal models of SARS, B cell responses (Bisht *et al.*, 2005) as well as $CD4^+$ and $CD8^+$ T cell responses (Channappanavar *et al.*, 2014; Liu *et al.*, 2017; Zhao *et al.*, 2010; Zhao *et al.*, 2016) were shown to have protective value. Notably, in most individuals who had recovered from SARS, SARS-CoV-1-specific memory $CD8^+$ T cells persisted for up to 6 years after SARS-CoV-1 infection whereas memory B cells and antivirus antibodies generally became undetectable (Tang *et al.*, 2011).

Based on theoretical considerations alone, it is difficult to predict effective B cell memory across different virus species (Qiu *et al.*, 2020), which makes it a poor topic for our present study which is based on sequence comparisons. We just note that a recent study concluded that sera from people that likely had been infected with the common human coronaviruses 229E, NL63, OC43, and/or HKU1, possessed no or negligible cross-reactivity with SARS-CoV-2 virus S protein (Amanat *et al.*, 2020) and thus probably possess no neutralizing antibodies.

There may be some recognition of SARS-CoV-2 epitopes by $CD4^+$ T cell memory derived from previous infections with common human coronaviruses. However, as discussed in the

Results and Discussion section, the very limited lengths of identical sequence stretches between the viruses make theoretical predictions of such epitopes difficult, and therefore the current study only concentrates on potential CD8+ T cell memory.

For inducing CD8+ T cell memory, the core requirement is merely that an identical short peptide is presented by major histocompatibility complex (MHC) class I (MHC-I) molecules. MHC-I molecules present peptide fragments from intracellular proteins, thus also from viral proteins, at the cell surface for screening by CD8+ cytotoxic T cells (Neefjes *et al.*, 2011). CD8+ T cells recognize the combination of MHC-I molecule with peptide by T cell receptors (TCR) that are unique per T cell clone, and if stimulated these clones can proliferate, kill the presenting (virus-infected) cell, and produce memory cells. MHC-I molecules are polymorphic in that they are represented by many diverse allelic forms that differ between human populations and individuals (Robinson *et al.*, 2020), and mostly bind peptides of 9 amino acids (aa) length in their binding groove which is closed at either end (Bjorkman *et al.*, 1987; Rammensee *et al.*, 1995; Schellens *et al.*, 2015).

In the present study, we analyzed whether there are linear 9 aa epitopes that are identical between proteins encoded by SARS-CoV-2 and one or more of the common human coronaviruses. We found many of such epitopes indeed, and, by using prediction software, found that some are expected to bind well to certain MHC-I alleles. We therefore expect that common human coronaviruses can induce some level of CD8+ T cell-mediated immune memory recognizing SARS-CoV-2, and consider the possibility of enhancing that immune memory by vaccination.

Methods

Proteins encoded by a reported genomic sequence for SARS-CoV-2 (GenBank MN908947; Wu *et al.*, 2020) were compared with those for HCoV-OC43 (NC_005147; Vijgen *et al.*, 2005), HCoV-HKU1 (NC_006577; Woo *et al.*, 2005), HCoV-229E (NC_002645; Thiel *et al.*, 2001), and HCoV-NL63 (NC_005831; van der Hoek *et al.*, 2004) by performing BLAST homology searches at the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and by making multiple sequence alignments using CLUSTALW software (https://www.genome.jp/tools-bin/clustalw); continuous stretches of 9 aa acids identical between SARS-CoV-2 and one of the other viruses were identified manually. All these shared 9 aa epitopes were screened by ANN 4.0 software at IEDB Analysis Resource (http://tools.immuneepitope.org/mhci/) for prediction of their affinity to a set of representative human MHC-I alleles.

Results and discussion

Table 1 lists the 9 aa epitopes that are identical between proteins encoded by SARS-CoV-2 and one or more of the common human coronaviruses. Many identical >9 aa stretches were found with ORF1ab encoded polyprotein, one such identical stretch (of 12 aa) was found with the N protein of the other two type II coronaviruses HCoV-OC43 and HCoV-HKU1, and no such stretches were found when comparing with any of the other gene products; ORF1ab-derived mature proteins with such stretches, expected from cleavage of the polyprotein precursor (Wu *et al.*, 2020), were the transmembrane protein nonstructural protein 4 (NSP4), 3C-like cysteine protease NSP5, RNA binding protein NSP9, RNA dependent RNA polymerase NSP12, helicase NSP13, 3'-to-5' exonuclease NSP14, nidoviral endoribonuclease specific for U NSP15, and S-adenosylmethionine-dependent ribose 2'-O-methyltransferase NSP16 (Table 1). Sequence alignment figures of the ORF1ab and N proteins are shown in *Extended data* (Dijkstra, 2020a) with highlighting of the interesting epitopes. It is of note that the S protein, which is the prime candidate for inducing neutralizing antibodies (Cohen, 2020), is poorly suitable for inducing an MHC-I-restricted immune memory across the investigated viral species as between S protein of SARS-CoV-2 and S proteins of the common human coronaviruses there are no 9 aa matches, and, among the virus isolates compared in this study, only a single 8 aa match (DRLITGRL with HCoV-NL63 and -229E) (not shown).

In Table 1 (for Excel format see *Extended data*) it is shown that there are >200 linear epitopes of 9 aa that are identical between SARS-CoV-2 and at least one of the common human coronaviruses, most of them with OC43 and HKU1 which, like SARS-CoV-2, belong to the group II coronaviruses. In a simplified model, if people would have been exposed to many of these epitopes through common HCoV infections, this kind of equals immunization by a small intracellular protein under natural viral infection conditions. Whereas live virus is commonly considered the gold standard in regard to inducing strong immunity, unless the virus has some tricks up its sleeve to manipulate the immune system, which for common human coronaviruses is not well investigated, a research grant proposal suggesting this as a vaccination strategy would probably fail. Reviewers of such proposal would righteously point out that the strategy would not induce neutralizing antibodies, which for combating some viral infections can be very important, and that for inducing MHC-I-restricted cell-mediated cytotoxicity memory, ideally, a much larger protein or more proteins should be taken. Those reviewers would conclude that for such small intracellular protein to induce strong immune memory it would be too dependent on the MHC alleles of the immunized person and would need too much luck in regard to immunogenicity. Nevertheless, those reviewers would probably also agree that in most persons thus vaccinated some (small) level of immune memory protection would be established, even with such small non-surface protein (e.g. Polakos *et al.*, 2001; Wasmoen *et al.*, 1995; Zhao *et al.*, 2005). Regardless of that this obviously is not the ideal way to induce a population-wide strong protective immunity (see the spread of COVID-19), together with other factors such as health and the number of encountered viruses (the strength of the viral challenge), the induced immune memory could make a difference for whether a person gets sick; at the population scale, it so may somewhat reduce the virus reproduction number. Importantly, by stimulating this HCoV-derived MHC-I restricted immune memory by vaccination (see below), it may become a more significant helper in fighting COVID-19.

Table 1. Stretches of 9 consecutive amino acids that are identical between SARS-CoV-2 and at least one of the common human coronaviruses. Compared sequences were derived from the following GenBank accessions: for ORF1ab SARS-CoV-2, QHD43415; OC43, NP_937947; HKU1, YP_173236; 229E, NP_073549; NL63, YP_003766; and for N SARS-CoV-2, QHD43423; OC43, NP_937954; HKU1, YP_173242; 229E, NP_073556; NL63, YP_003771 Source positions indicate the N-terminal position of the depicted 9 aa sequence in the SARS-CoV-2 ORF1ab protein or N protein (see Supplementary file 1). The ORF1ab protein is only a precursor polyprotein, and the column Mature protein indicates the probable mature protein that possesses the epitope: 3CLpro, 3C-like cysteine protease; RdRp, RNA dependent RNA polymerase; Hel, helicase; ExoN, 3'-to-5' exonuclease; NendoU, nidoviral endoribonuclease specific for U; O-MT, S-adenosylmethionine-dependent ribose 2'-O-methyltransferase. Yellow blocks indicate the presence of identical sequences in the respective common human coronavirus, and gray blocks indicate the absence of such matches. Orange blocks highlight those peptides with predicted IC50 values of <500 nM for one of the twelf investigated MHC-I alleles. No predicted IC50 values of <500 nM were found for HLA-C*0401

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)													
	Mature		Group II		Group I		HLA-A					HLA-B							HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 3004	NSP4	WVLNNDYYR																		
ORF1ab 3005	NSP4	VLNNDYYRS																		
ORF1ab 3006	NSP4	LNNDYYRSL																		
ORF1ab 3007	NSP4	NNDYYRSLP																		
ORF1ab 3008	NSP4	NDYYRSLPG																		
ORF1ab 3289	NSP5 3CLpro	TLNGLWLDD																		
ORF1ab 3299	NSP5 3CLpro	VYCPRHVIC																		
ORF1ab 4208	NSP9	ELEPPCRFV																		
ORF1ab 4551	NSP12 RdRp	KKDWYDFVE																		
ORF1ab 4552	NSP12 RdRp	KDWYDFVEN																		
ORF1ab 4553	NSP12 RdRp	DWYDFVENP																		
ORF1ab 4554	NSP12 RdRp	WYDFVENPD																		
ORF1ab 4555	NSP12 RdRp	YDFVENPDI																		
ORF1ab 4594	NSP12 RdRp	VGVLTLDNQ																		
ORF1ab 4595	NSP12 RdRp	GVLTLDNQD																		
ORF1ab 4596	NSP12 RdRp	VLTLDNQDL																		
ORF1ab 4597	NSP12 RdRp	LTLDNQDLN																		
ORF1ab 4598	NSP12 RdRp	TLDNQDLNG																		
ORF1ab 4599	NSP12 RdRp	LDNQDLNGN																		
ORF1ab 4608	NSP12 RdRp	WYDFGDFIQ																		
ORF1ab 4609	NSP12 RdRp	YDFGDFIQT																		
ORF1ab 4661	NSP12 RdRp	DLLKYDFTE																		
ORF1ab 4725	NSP12 RdRp	IFVDGVPFV						158 nM												
ORF1ab 4726	NSP12 RdRp	FVDGVPFVV						10 nM								143 nM				
ORF1ab 4727	NSP12 RdRp	VDGVPFVVS																		
ORF1ab 4799	NSP12 RdRp	FQTVKPGNF																		
ORF1ab 4800	NSP12 RdRp	QTVKPGNFN																		
ORF1ab 4931	NSP12 RdRp	TQMNLKYAI						278 nM								21 nM				

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)														
	Mature		Group II		Group I		HLA-A					HLA-B								HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 4932	NSP12 RdRp	QMNLKYAIS																		
ORF1ab 4933	NSP12 RdRp	MNLKYAISA																		
ORF1ab 4934	NSP12 RdRp	NLKYAISAK							198 nM											
ORF1ab 4935	NSP12 RdRp	LKYAISAKN																		
ORF1ab 4936	NSP12 RdRp	KYAISAKNR																		
ORF1ab 4937	NSP12 RdRp	YAISAKNRA																		
ORF1ab 4938	NSP12 RdRp	AISAKNRAR																		
ORF1ab 4939	NSP12 RdRp	ISAKNRART																		
ORF1ab 4940	NSP12 RdRp	SAKNRARTV											58 nM							
ORF1ab 4941	NSP12 RdRp	AKNRARTVA																		
ORF1ab 4942	NSP12 RdRp	KNRARTVAG																		
ORF1ab 4943	NSP12 RdRp	NRARTVAGV													468 nM					
ORF1ab 4944	NSP12 RdRp	RARTVAGVS																		
ORF1ab 4945	NSP12 RdRp	ARTVAGVSI													219 nM					
ORF1ab 5006	NSP12 RdRp	LMGWDYPKC																		
ORF1ab 5007	NSP12 RdRp	MGWDYPKCD																		
ORF1ab 5008	NSP12 RdRp	GWDYPKCDR																		
ORF1ab 5009	NSP12 RdRp	WDYPKCDRA																		
ORF1ab 5010	NSP12 RdRp	DYPKCDRAM																		
ORF1ab 5011	NSP12 RdRp	YPKCDRAMP																		
ORF1ab 5012	NSP12 RdRp	PKCDRAMPN																		
ORF1ab 5043	NSP12 RdRp	RFYRLANEC																		
ORF1ab 5044	NSP12 RdRp	FYRLANECA																		
ORF1ab 5045	NSP12 RdRp	YRLANECAQ																		
ORF1ab 5046	NSP12 RdRp	RLANECAQV						53 nM												
ORF1ab 5047	NSP12 RdRp	LANECAQVL																	26 nM	
ORF1ab 5048	NSP12 RdRp	ANECAQVLS																		
ORF1ab 5049	NSP12 RdRp	NECAQVLSE																		
ORF1ab 5066	NSP12 RdRp	YVKPGGTSS																		
ORF1ab 5067	NSP12 RdRp	VKPGGTSSG																		
ORF1ab 5068	NSP12 RdRp	KPGGTSSGD																		
ORF1ab 5069	NSP12 RdRp	PGGTSSGDA																		
ORF1ab 5070	NSP12 RdRp	GGTSSGDAT																		

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)														
	Mature		Group II		Group I		HLA-A					HLA-B								HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 5071	NSP12 RdRp	GTSSGDATT																		
ORF1ab 5072	NSP12 RdRp	TSSGDATTA																		
ORF1ab 5074	NSP12 RdRp	SGDATTAYA																		
ORF1ab 5075	NSP12 RdRp	GDATTAYAN																		
ORF1ab 5076	NSP12 RdRp	DATTAYANS																		
ORF1ab 5077	NSP12 RdRp	ATTAYANSV																		
ORF1ab 5078	NSP12 RdRp	TTAYANSVF									245 nM			29 nM						
ORF1ab 5079	NSP12 RdRp	TAYANSVFN																	466 nM	
ORF1ab 5080	NSP12 RdRp	AYANSVFNI								56 nM										
ORF1ab 5082	NSP12 RdRp	ANSVFNICQ																		
ORF1ab 5083	NSP12 RdRp	NSVFNICQA																		
ORF1ab 5084	NSP12 RdRp	SVFNICQAV						85 nM												
ORF1ab 5085	NSP12 RdRp	VFNICQAVT																		
ORF1ab 5086	NSP12 RdRp	FNICQAVTA																		
ORF1ab 5087	NSP12 RdRp	NICQAVTAN																		
ORF1ab 5088	NSP12 RdRp	ICQAVTANV																		
ORF1ab 5140	NSP12 RdRp	YLRKHFSMM									128 nM	134 nM	4 nM	47 nM						
ORF1ab 5141	NSP12 RdRp	LRKHFSMMI													313 nM					
ORF1ab 5142	NSP12 RdRp	RKHFSMMIL														310 nM				
ORF1ab 5143	NSP12 RdRp	KHFSMMILS																		
ORF1ab 5144	NSP12 RdRp	HFSMMILSD																		
ORF1ab 5145	NSP12 RdRp	FSMMILSDD																		
ORF1ab 5177	NSP12 RdRp	VLYYQNNVF												25 nM						
ORF1ab 5178	NSP12 RdRp	LYYQNNVFM																		
ORF1ab 5179	NSP12 RdRp	YYQNNVFMS																		
ORF1ab 5180	NSP12 RdRp	YQNNVFMSE																		
ORF1ab 5196	NSP12 RdRp	DLTKGPHEF																		
ORF1ab 5197	NSP12 RdRp	LTKGPHEFC																		
ORF1ab 5198	NSP12 RdRp	TKGPHEFCS																		
ORF1ab 5199	NSP12 RdRp	KGPHEFCSQ																		
ORF1ab 5200	NSP12 RdRp	GPHEFCSQH																		
ORF1ab 5201	NSP12 RdRp	PHEFCSQHT																		

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)													
	Mature		Group II		Group I		HLA-A					HLA-B							HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 5202	NSP12 RdRp	HEFCSQHTM														54 nM	10 nM			
ORF1ab 5203	NSP12 RdRp	EFCSQHTML																		
ORF1ab 5204	NSP12 RdRp	FCSQHTMLV																		
ORF1ab 5205	NSP12 RdRp	CSQHTMLVK							227 nM											
ORF1ab 5217	NSP12 RdRp	DYVYLPYPD																		
ORF1ab 5218	NSP12 RdRp	YVYLPYPDP																		
ORF1ab 5219	NSP12 RdRp	VYLPYPDPS																		
ORF1ab 5220	NSP12 RdRp	YLPYPDPSR																		
ORF1ab 5221	NSP12 RdRp	LPYPDPSRI																		
ORF1ab 5222	NSP12 RdRp	PYPDPSRIL																		
ORF1ab 5223	NSP12 RdRp	YPDPSRILG																		
ORF1ab 5224	NSP12 RdRp	PDPSRILGA																		
ORF1ab 5225	NSP12 RdRp	DPSRILGAG																		
ORF1ab 5226	NSP12 RdRp	PSRILGAGC																		
ORF1ab 5227	NSP12 RdRp	SRILGAGCF																		
ORF1ab 5228	NSP12 RdRp	RILGAGCFV						153 nM												
ORF1ab 5229	NSP12 RdRp	ILGAGCFVD																		
ORF1ab 5230	NSP12 RdRp	LGAGCFVDD																		
ORF1ab 5248	NSP12 RdRp	IERFVSLAI															221 nM			
ORF1ab 5249	NSP12 RdRp	ERFVSLAID																		
ORF1ab 5250	NSP12 RdRp	RFVSLAIDA																		
ORF1ab 5251	NSP12 RdRp	FVSLAIDAY									477 nM			223 nM						
ORF1ab 5252	NSP12 RdRp	VSLAIDAYP																		
ORF1ab 5253	NSP12 RdRp	SLAIDAYPL						21 nM								487 nM				
ORF1ab 5349	NSP13 Hel	LCCKCCYDH																		
ORF1ab 5350	NSP13 Hel	CCKCCYDHV																		
ORF1ab 5371	NSP13 Hel	PYVCNAPGC																		
ORF1ab 5372	NSP13 Hel	YVCNAPGCD																		
ORF1ab 5373	NSP13 Hel	VCNAPGCDV																		
ORF1ab 5387	NSP13 Hel	LYLGGMSYY																		
ORF1ab 5388	NSP13 Hel	YLGGMSYYC						89 nM												

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)													
	Mature		Group II		Group I		HLA-A					HLA-B							HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 5450	NSP13 Hel	CTERLKLFA					303 nM													
ORF1ab 5451	NSP13 Hel	TERLKLFAA																		
ORF1ab 5452	NSP13 Hel	ERLKLFAAE																		
ORF1ab 5453	NSP13 Hel	RLKLFAAET																		
ORF1ab 5559	NSP13 Hel	LSAPTLVPQ																		
ORF1ab 5560	NSP13 Hel	SAPTLVPQE																		
ORF1ab 5605	NSP13 Hel	QGPPGTGKS																		
ORF1ab 5606	NSP13 Hel	GPPGTGKSH																		
ORF1ab 5634	NSP13 Hel	SHAAVDALC																		
ORF1ab 5635	NSP13 Hel	HAAVDALCE																		
ORF1ab 5636	NSP13 Hel	AAVDALCEK																		
ORF1ab 5637	NSP13 Hel	AVDALCEKA																		
ORF1ab 5656	NSP13 Hel	RIIPARARV																		
ORF1ab 5657	NSP13 Hel	IIPARARVE																		
ORF1ab 5658	NSP13 Hel	IPARARVEC										213 nM								
ORF1ab 5716	NSP13 Hel	RAKHYVYIG																		
ORF1ab 5717	NSP13 Hel	AKHYVYIGD																		
ORF1ab 5718	NSP13 Hel	KHYVYIGDP																		
ORF1ab 5719	NSP13 Hel	HYVYIGDPA																		
ORF1ab 5720	NSP13 Hel	YVYIGDPAQ																	312 nM	
ORF1ab 5721	NSP13 Hel	VYIGDPAQL								206 nM										
ORF1ab 5722	NSP13 Hel	YIGDPAQLP																		
ORF1ab 5723	NSP13 Hel	IGDPAQLPA																		
ORF1ab 5724	NSP13 Hel	GDPAQLPAP																		
ORF1ab 5725	NSP13 Hel	DPAQLPAPR																		
ORF1ab 5771	NSP13 Hel	EIVDTVSAL									16 nM					298 nM			463 nM	
ORF1ab 5772	NSP13 Hel	IVDTVSALV						114 nM												
ORF1ab 5773	NSP13 Hel	VDTVSALVY																		
ORF1ab 5775	NSP13 Hel	TVSALVYDN																		
ORF1ab 5776	NSP13 Hel	VSALVYDNK																		
ORF1ab 5777	NSP13 Hel	SALVYDNKL																		
ORF1ab 5778	NSP13 Hel	ALVYDNKLK																		
ORF1ab 5779	NSP13 Hel	LVYDNKLKA																		

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)													
	Mature		Group II		Group I		HLA-A					HLA-B							HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 5832	NSP13 Hel	KAVFISPYN																		
ORF1ab 5833	NSP13 Hel	AVFISPYNS																		
ORF1ab 5834	NSP13 Hel	VFISPYNSQ																		
ORF1ab 5835	NSP13 Hel	FISPYNSQN																		
ORF1ab 5855	NSP13 Hel	QTVDSSQGS																		
ORF1ab 5856	NSP13 Hel	TVDSSQGSE																		
ORF1ab 5857	NSP13 Hel	VDSSQGSEY																		
ORF1ab 5858	NSP13 Hel	DSSQGSEYD																		
ORF1ab 5859	NSP13 Hel	SSQGSEYDY																		
ORF1ab 5860	NSP13 Hel	SQGSEYDYV																		
ORF1ab 5861	NSP13 Hel	QGSEYDYVI																		
ORF1ab 5862	NSP13 Hel	GSEYDYVIF																		
ORF1ab 5880	NSP13 Hel	CNVNRFNVA																		
ORF1ab 5881	NSP13 Hel	NVNRFNVAI																		
ORF1ab 5882	NSP13 Hel	VNRFNVAIT																		
ORF1ab 5883	NSP13 Hel	NRFNVAITR													65 nM					
ORF1ab 5884	NSP13 Hel	RFNVAITRA																		
ORF1ab 5885	NSP13 Hel	FNVAITRAK																		
ORF1ab 6031	NSP14 ExoN	PLQLGFSTG																		
ORF1ab 6198	NSP14 ExoN	DAIMTRCLA																		
ORF1ab 6199	NSP14 ExoN	AIMTRCLAV						29 nM					24 nM							
ORF1ab 6307	NSP14 ExoN	CLFWNCNVD																		
ORF1ab 6320	NSP14 ExoN	NSIVCRFDT																		
ORF1ab 6321	NSP14 ExoN	SIVCRFDTR																		
ORF1ab 6322	NSP14 ExoN	IVCRFDTRV																		
ORF1ab 6323	NSP14 ExoN	VCRFDTRVL																		
ORF1ab 6341	NSP14 ExoN	GGSLYVNKH																		
ORF1ab 6342	NSP14 ExoN	GSLYVNKHA																		
ORF1ab 6343	NSP14 ExoN	SLYVNKHAF											279 nM	154 nM						
ORF1ab 6344	NSP14 ExoN	LYVNKHAFH																		
ORF1ab 6345	NSP14 ExoN	YVNKHAFHT																		
ORF1ab 6346	NSP14 ExoN	VNKHAFHTP																		
ORF1ab 6347	NSP14 ExoN	NKHAFHTPA																		
ORF1ab 6389	NSP14 ExoN	DYVPLKSAT																		

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)													
	Mature		Group II		Group I		HLA-A					HLA-B							HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 6390	NSP14 ExoN	YVPLKSATC																		
ORF1ab 6391	NSP14 ExoN	VPLKSATCI																		
ORF1ab 6392	NSP14 ExoN	PLKSATCIT																		
ORF1ab 6393	NSP14 ExoN	LKSATCITR																		
ORF1ab 6394	NSP14 ExoN	KSATCITRC																331 nM		
ORF1ab 6395	NSP14 ExoN	SATCITRCN																		
ORF1ab 6396	NSP14 ExoN	ATCITRCNL																		
ORF1ab 6397	NSP14 ExoN	TCITRCNLG																		
ORF1ab 6398	NSP14 ExoN	CITRCNLGG																		
ORF1ab 6399	NSP14 ExoN	ITRCNLGGA																		
ORF1ab 6400	NSP14 ExoN	TRCNLGGAV																		
ORF1ab 6401	NSP14 ExoN	RCNLGGAVC																		
ORF1ab 6682	NSP15 NendoU	YAFEHIVYG																	177 nM	
ORF1ab 6698	NSP15 NendoU	GGLHLLIGL																		
ORF1ab 6746	NSP15 NendoU	VIDLLLDDF																		
ORF1ab 6747	NSP15 NendoU	IDLLLDDFV																		
ORF1ab 6839	NSP16 O-MT	MMNVAKYTQ																		
ORF1ab 6840	NSP16 O-MT	MNVAKYTQL											259 nM							
ORF1ab 6841	NSP16 O-MT	NVAKYTQLC																		
ORF1ab 6842	NSP16 O-MT	VAKYTQLCQ																		
ORF1ab 6843	NSP16 O-MT	AKYTQLCQY																		
ORF1ab 6844	NSP16 O-MT	KYTQLCQYL								139 nM										
ORF1ab 6845	NSP16 O-MT	YTQLCQYLN																		
ORF1ab 6846	NSP16 O-MT	TQLCQYLNT																		
ORF1ab 6869	NSP16 O-MT	GAGSDKGVA																		
ORF1ab 6870	NSP16 O-MT	AGSDKGVAP																		
ORF1ab 6871	NSP16 O-MT	GSDKGVAPG																		
ORF1ab 6872	NSP16 O-MT	SDKGVAPGT																		
ORF1ab 6922	NSP16 O-MT	WDLIISDMY																		
ORF1ab 6923	NSP16 O-MT	DLIISDMYD																		
ORF1ab 6924	NSP16 O-MT	LIISDMYDP																		
ORF1ab 6943	NSP16 O-MT	SKEGFFTYI																		
ORF1ab 6958	NSP16 O-MT	KLALGGSVA																		
ORF1ab 6959	NSP16 O-MT	LALGGSVAI																	11 nM	

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

			common human coronaviruses				IC50 prediction by ANN 4.0 program of IEDB software (only IC50 valuesof <500 nM are shown)														
	Mature		Group II		Group I		HLA-A					HLA-B								HLA-C	
SARS-CoV-2 source	protein	Sequence (9aa)	OC43	HKU1	229E	NL63	*0101	*0201	*0301	*2402	*2601	*0702	*0801	*1501	*2705	*3901	*4001	*5801	*0303	*0401
ORF1ab 6960	NSP16 O-MT	ALGGSVAIK							110 nM											
ORF1ab 6961	NSP16 O-MT	LGGSVAIKI																		
ORF1ab 6962	NSP16 O-MT	GGSVAIKIT																		
ORF1ab 6963	NSP16 O-MT	GSVAIKITE																		
ORF1ab 6973	NSP16 O-MT	SWNADLYKL																		
ORF1ab 6974	NSP16 O-MT	WNADLYKLM																		
ORF1ab 6993	NSP16 O-MT	TNVNASSSE																		
ORF1ab 6998	NSP16 O-MT	SSSEAFLIG																		
ORF1ab 7024	NSP16 O-MT	HANYIFWRN																		
N 106	Nucleocapsid	PRWYFYYLG																		
N 107	Nucleocapsid	RWYFYYLGT																		
N 108	Nucleocapsid	WYFYYLGTG																		
N 109	Nucleocapsid	YFYYLGTGP																		

F1000Research 2020, 9:285 Last updated: 17 JUL 2020

Software predictions of MHC-I-binding epitopes

Based on combinations of experimental results and computer learning, various software has been created that with some degree of reliability can predict how efficiently peptides can bind to the grooves of various MHC-I alleles. In the present study, we used the artificial neural network (ANN) function (Lundegaard *et al.*, 2008) of the IEDB Analysis Resource (http://tools.immuneepitope.org/mhci/) (Dhanda *et al.*, 2019) which may achieve >75% reliability for predicting binding (Lundegaard *et al.*, 2008). The software designers state that IC50 values of <50 nM and <500 nM are considered high and intermediate affinity, respectively, and are found for most epitopes known to stimulate cytotoxic T cells. Therefore, Table 1 only indicates the predicted IC50 values if lower than 500 nM. Table 1 shows these expected affinities for fourteen MHC-I alleles that are rather representative for sets of MHC-I alleles with similar binding properties (supertypes) and so represent a large part of the human MHC-I binding repertoire (Doytchinova *et al.*, 2004; Lund *et al.*, 2004): HLA-A*0101 (supertype A1), HLA-A*0201 (A2), HLA-A*0301 (A3), HLA-A*2402 (A24), HLA-A*2601(A26), HLA-B*0702 (B7), HLA-B*0801 (B8), HLA-B*1501 (B62), HLA-B*2705 (B27), HLA-B*3901 (B39), HLA-B*4001 (B44), HLA-B*5801 (B58), HLA-C*0303 (C1), and HLA-C*0401 (C4). It is of note that Li *et al.* (2008) found that a SARS-CoV-1 15 aa peptide sequence (their "Replicase 4701-4715" peptide) encompassing the SARS-CoV-2/HCoV-shared ORF1ab4725 and ORF1ab4726 epitopes that are predicted to bind well to the MHC-I alleles HLA-A*0201 and HLA-B*3901 (see our Table 1) was associated with a $CD8^+$ T cell response against SARS-CoV-1 in humans. However, Li *et al.* (2008) also found such $CD8^+$ T cell response associated with a SARS-CoV-1 15 aa peptide (their "Nucleocapsid 106-120" peptide) encompassing the SARS-CoV-2/HCoV-shared N 106, N 107, N 108, and N 109 epitopes for which our analyses did not predict MHC-I binding (see our Table 1).

The MHC-I binding affinity is considered the most selective in determining which peptides are presented, but also steps in the peptide processing and loading pathways may play selective roles which are difficult to capture in prediction software (Nielsen *et al.*, 2005). We argue that, if such steps would be selective for presentation, in most cases they would probably not differentiate between the 9 aa epitope in the SARS-CoV-2 context versus the respective HCoV context, since most of those epitopes are within stretches that also show many similarities in the neighboring residues (*Extended data*).

Not all stable complexes of MHC-I with non-self peptides elicit a strong immune response, but "immunogenicity" features are hard to predict with meaningful reliability by *in silico* analysis (Calis *et al.*, 2013), and in the present study we refrain from such predictions. Table 1 should, foremost, be understood as evidence of principle and a list of promising peptides, whereas only future experiments can prove MHC-I-mediated immune memory involving these or other peptides.

In regard to SARS-CoV-2 recognition, the common human coronaviruses may also induce some MHC-II-mediated immune memory by $CD4^+$ helper T cells (as an example for shared epitope use by different coronaviruses see Zhao *et al.*, 2016). $CD4^+$ helper T cells can help stimulate cells involved in antibody or cell-mediated cytotoxic immune responses (Neefjes *et al.*, 2011). However, for this topic, in the present article, we have refrained from detailed (software) predictions because comparison of MHC-II epitopes across different viruses is harder than for MHC-I epitopes. Namely, although the core of MHC-II bound peptides is also only 9 aa, the surrounding amino acids are also part of the bound peptide that tends to be 12–25 aa (Brown *et al.*, 1993; Rammensee *et al.*, 1995; Stern & Wiley, 1994) and can affect how the peptide interacts with the receptors on the $CD4^+$ helper T cells (Arnold *et al.*, 2002).

Vaccination potential

Immune memory means that a secondary immune response, upon renewed encounter with the same pathogen, is faster and stronger than the primary immune response during the first encounter with the pathogen. This is based on expansion of specific B and T cell clones, which specifically recognize pathogen(-derived) epitopes, with some of those cells becoming memory cells (Paul, 2013). This principle also causes that for a booster vaccination/immunization the requirements for efficiently inducing an immune response are lower than for a first vaccination/immunization (e.g. Du *et al.*, 2008; Goding, 1996; Schulze *et al.*, 2008). Especially in elderly people, who have a decreased ability to mount adaptive immune responses against new antigens, vaccination that stimulates an immune memory response may be beneficial (Kaml *et al.*, 2006; Reber *et al.*, 2012; Wagner & Weinberger, 2020). As discussed above, people's past infections with common coronaviruses probably did not induce a B cell memory for making antibodies that can neutralize SARS-CoV-2. However, as the current study shows by analysis of linear 9 aa epitopes, these common human coronaviruses are expected to induce $CD8^+$ T cells that may potentially kill SARS-CoV-2-infected cells and so can help eradicate the virus. There are several possible ways to exploit this probable immune memory. For example, if using RNA for immunization (Cohen, 2020), it may be best to also include SARS-CoV-2 genes that encode MHC-I epitopes that match those of the common coronaviruses. Alternatively, delivery of these epitopes to the MHC-I presentation system may be tried by peptide or protein based vaccines (e.g. Kohyama *et al.*, 2009; Slingluff, 2011; van Montfoort *et al.*, 2014; Yadav *et al.*, 2014), possibly in combination with some of the strategies that are currently being explored for non-specific stimulation of the immune system against COVID-19 (Kupferschmidt & Cohen, 2020). Protein (-coding) vaccines, for example encompassing a large part of the SARS-CoV-2 ORF1ab product, would have an advantage over peptide-vaccines by including multiple possible MHC-I and also MHC-II epitopes, and be less dependent on MHC-allele matching and the quality of software predictions. Naturally, as for any new vaccine strategy, it should be carefully assessed whether the benefits of the induced type of immunity outweigh the potential deleterious health effects caused by, for example, an increased inflammation response (Cohen, 2020; Weingartl *et al.*, 2004). Another fundamental concern is the maximum level of protection that can be generated by vaccination against coronavirus infections in humans,

considering that infection of volunteers with HCoV-229E live virus gave only partial protection upon infection with the same virus one year later (Callow *et al.*, 1990). Additional questions specifically related to the contents of our study are whether the history of previous—especially recent—infections with common coronaviruses, or people's MHC alleles, affect people's resistance to SARS-CoV-2. Most definitely, if discussing possible strategies for vaccination against SARS-CoV-2, pre-existing MHC-I-based immunity derived from previous infections with common coronaviruses should be part of the consideration.

Notifications

Although we were not aware of this at the time of writing, a recent paper appeared with overlapping contents (Nguyen *et al.*, 2020). The Nguyen *et al.* study was more complete on SARS-CoV-2 MHC epitope predictions and made an association with global MHC allele distributions. The advantage of our study is a more concentrated focus on the MHC-I mediated memory expected from previous coronavirus infections, and the vaccination potential deriving from that memory.

After we had submitted our study, two studies reported *in vitro* responses of T cells against SARS-CoV-2 peptides, which might represent memory from previous infections with common coronaviruses (Braun *et al.*, 2020; Grifoni *et al.*, 2020). However, both studies only used peptide mixes without identifying the responsible peptide, and at least several of the observed responses necessitated the allowance of peptide ligand sequence mismatches for T cell receptor to MHC/peptide binding (T cell cross-reactivity). Negative control donors, who with certainty had never been infected with common coronaviruses, were not available for the experiments, and conclusions that the observed responses were from T cell memory from previous coronavirus infections, and have *in vivo* relevance, should be considered only cautiously. Discussion of this topic is important because the two studies concluded a potential of the common coronavirus S proteins to induce $CD4^+$ T cell memory (Braun *et al.*, 2020; Grifoni *et al.*, 2020) and $CD8^+$ T cell memory (Grifoni *et al.*, 2020), whereas these proteins do not share 9 aa identical stretches with SARS-CoV-2 (see our article and Supplementary Fig. 1 in Braun *et al.*, 2020), and would arguably necessitate the allowance of peptide sequence mismatches (T cell cross-reactivity) for inducing an efficient MHC-mediated T cell response. As we pointed out in our article, although SARS-CoV-2 S protein is the prime vaccine component candidate for inducing neutralizing antibodies, for a more realistic chance to efficiently boost existing T cell memory it probably would be better to additionally include other SARS-CoV-2 proteins that do share identical MHC epitopes with common oronaviruses.

Regarding the potential of existing $CD8^+$ T cell memory cells to help fight COVID-19 disease, a recent observation by Liao *et al.*, (2020) might be interesting. Their study suggests that in COVID-19 patients with pneumonia, $ZNF683^+$ $CD8^+$ T cell clonal expansion may protect the patient from more severe disease.

Data availability

Underlying data

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome, Accession number MN908947: https://www.ncbi.nlm.nih.gov/nuccore/MN908947

Human coronavirus OC43, complete genome, Accession number NC_005147.1: https://www.ncbi.nlm.nih.gov/nuccore/NC_005147.1?report=genbank

Human coronavirus HKU1, complete genome, Accession number NC_006577: https://www.ncbi.nlm.nih.gov/nuccore/NC_006577

Human coronavirus 229E, complete genome, Accession number NC_002645: https://www.ncbi.nlm.nih.gov/nuccore/NC_002645

Human Coronavirus NL63, complete genome, Accession number NC_005831: https://www.ncbi.nlm.nih.gov/nuccore/NC_005831

Extended data

Harvard Dataverse: Extended data. Sequence alignments of SARS-CoV-2 ORF1ab and N proteins with their counterparts in the common human coronaviruses, https://doi.org/10.7910/DVN/CNPUPA (Dijkstra, 2020a).

Harvard Dataverse: Excel format version of Table 1. https://doi.org/10.7910/DVN/LOBKLV (Dijkstra, 2020b).

References

Al Kahlout RA, Nasrallah GK, Farag EA, *et al.*: **Comparative Serological Study for the Prevalence of Anti-MERS Coronavirus Antibodies in High- and Low-Risk Groups in Qatar.** *J Immunol Res.* 2019; **2019**: 1386740.
PubMed Abstract | Publisher Full Text | Free Full Text

Amanat F, Nguyen T, Chromikova V, *et al.*: **A serological assay to detect SARS-CoV-2 seroconversion in humans.** *medRxiv.* 2020.
Publisher Full Text

Andersen KG, Rambaut A, Lipkin WI, *et al.*: **The proximal origin of SARS-CoV-2.** *Nat Med.* 2020; **26**(4): 450–452.
PubMed Abstract | Publisher Full Text

Arnold PY, La Gruta NL, Miller T, *et al.*: **The majority of immunogenic epitopes generate $CD4^+$ T cells that are dependent on MHC class II-bound peptide-flanking residues.** *J Immunol.* 2002; **169**(2): 739–49.
PubMed Abstract | Publisher Full Text

Bisht H, Roberts A, Vogel L, *et al.*: **Neutralizing antibody and protective immunity to SARS coronavirus infection of mice induced by a soluble recombinant polypeptide containing an N-terminal segment of the spike glycoprotein.** *Virology.* 2005; **334**(2): 160–5.
PubMed Abstract | Publisher Full Text | Free Full Text

Bjorkman PJ, Saper MA, Samraoui B, *et al.*: **Structure of the human class I histocompatibility antigen, HLA-A2.** *Nature.* 1987; **329**(6139): 506–12.
PubMed Abstract | Publisher Full Text

Bradburne AF, Somerset BA: **Coronative antibody tires in sera of healthy adults and experimentally infected volunteers.** *J Hyg (Lond).* 1972; **70**(2): 235–44.
PubMed Abstract | Publisher Full Text | Free Full Text

Braun J, Loyal L, Frentsch M, *et al.*: **Presence of SARS-CoV-2 reactive T cells in COVID-19 patients and healthy donors.** *MedRxiv.* 2020.
Publisher Full Text

Brown JH, Jardetzky TS, Gorga JC, *et al.*: **Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1.** *Nature.* 1993; **364**(6432): 33–9.
PubMed Abstract | Publisher Full Text

Calis JJ, Maybeno M, Greenbaum JA, *et al.*: **Properties of MHC class I presented peptides that enhance immunogenicity.** *PLoS Comput Biol.* 2013; **9**(10): e1003266.
PubMed Abstract | Publisher Full Text | Free Full Text

Callow KA, Parry HF, Sergeant M, *et al.*: **The time course of the immune response to experimental coronavirus infection of man.** *Epidemiol Infect.* 1990; **105**(2): 435–46.
PubMed Abstract | Publisher Full Text | Free Full Text

Ceraolo C, Giorgi FM: **Genomic variance of the 2019-nCoV coronavirus.** *J Med Virol.* 2020; **92**(5): 522–528.
PubMed Abstract | Publisher Full Text

Channappanavar R, Fett C, Zhao J, *et al.*: **Virus-specific memory CD8 T cells provide substantial protection from lethal severe acute respiratory syndrome coronavirus infection.** *J Virol.* 2014; **88**(19): 11034–44.
PubMed Abstract | Publisher Full Text | Free Full Text

Cohen J: **Vaccine designers take first shots at COVID-19.** *Science.* 2020; **368**(6486): 14–16.
PubMed Abstract | Publisher Full Text

Dhanda SK, Mahajan S, Paul S, *et al.*: **IEDB-AR: immune epitope database-analysis resource in 2019.** *Nucleic Acids Res.* 2019; **47**(W1): W502–W506.
PubMed Abstract | Publisher Full Text | Free Full Text

Dijkman R, Jebbink MF, Gaunt E, *et al.*: **The dominance of human coronavirus OC43 and NL63 infections in infants.** *J Clin Virol.* 2012; **53**(2): 135–9.
PubMed Abstract | Publisher Full Text | Free Full Text

Dijkstra H: **Extended data. Sequence alignments of SARS-CoV-2 ORF1ab and N proteins with their counterparts in the common human coronaviruses.** *Harvard Dataverse.* V1. 2020a.
http://www.doi.org/10.7910/DVN/CNPUPA

Dijkstra H: Excel format version of Table 1. *Harvard Dataverse.* V1. 2020b.
http://www.doi.org/10.7910/DVN/LOBKLV

Doytchinova IA, Guan P, Flower DR: **Identifiying human MHC supertypes using bioinformatic methods.** *J Immunol.* 2004; **172**(7): 4314–23.
PubMed Abstract | Publisher Full Text

Du L, Zhao G, Lin Y, *et al.*: **Intranasal vaccination of recombinant adeno-associated virus encoding receptor-binding domain of severe acute respiratory syndrome coronavirus (SARS-CoV) spike protein induces strong mucosal immune responses and provides long-term protection against SARS-CoV infection.** *J Immunol.* 2008; **180**(2): 948–56.
PubMed Abstract | Publisher Full Text | Free Full Text

Gaunt ER, Hardie A, Claas EC, *et al.*: **Epidemiology and clinical presentations of the four human coronaviruses 229E, HKU1, NL63, and OC43 detected over 3 years using a novel multiplex real-time PCR method.** *J Clin Microbiol.* 2010; **48**(8): 2940–7.
PubMed Abstract | Publisher Full Text | Free Full Text

Goding JW: **Production of Monoclonal Antibodies.** *Monoclonal Antibodies* (Third Edition) Ed.: Goding JW. Academic Press Limited, London, 1996; 141–191.
Reference Source

Grifoni A, Weiskopf D, Ramirez SI, *et al.*: **Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals.** *Cell.* 2020; **181**(7): 1489–1501.
PubMed Abstract | Publisher Full Text | Free Full Text

Kahn JS, McIntosh K: **History and recent advances in coronavirus discovery.** *Pediatr Infect Dis J.* 2005; **24**(11 suppl): S223–7, discussion S226.
PubMed Abstract | Publisher Full Text

Kaml M, Weiskirchner I, Keller M, *et al.*: **Booster vaccination in the elderly: their success depends on the vaccine type applied earlier in life as well as on pre-vaccination antibody titers.** *Vaccine.* 2006; **24**(47–48): 6808–11.
PubMed Abstract | Publisher Full Text

Killerby ME, Biggs HM, Haynes A, *et al.*: **Human coronavirus circulation in the United States 2014-2017.** *J Clin Virol.* 2018; **101**: 52–6.
PubMed Abstract | Publisher Full Text | Free Full Text

Kohyama S, Ohno S, Suda T, *et al.*: **Efficient induction of cytotoxic T lymphocytes specific for severe acute respiratory syndrome (SARS)-associated coronavirus by immunization with surface-linked liposomal peptides derived from a non-structural polyprotein 1a.** *Antiviral Res.* 2009; **84**(2): 168–77.
PubMed Abstract | Publisher Full Text | Free Full Text

Kupferschmidt K, Cohen J: **Race to find COVID-19 treatments accelerates.** *Science.* 2020; **367**(6485): 1412–3.
PubMed Abstract | Publisher Full Text

Li CK, Wu H, Yan H, *et al.*: **T cell responses to whole SARS coronavirus in humans.** *J Immunol.* 2008; **181**(8): 5490–500.
PubMed Abstract | Publisher Full Text | Free Full Text

Liao M, Liu Y, Yuan J, *et al.*: **Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19.** *Nat Med.* 2020; **26**(6): 842–844.
PubMed Abstract | Publisher Full Text

Liu WJ, Zhao M, Liu K, *et al.*: **T-cell immunity of SARS-CoV: Implications for vaccine development against MERS-CoV.** *Antiviral Res.* 2017; **137**: 82–92.
PubMed Abstract | Publisher Full Text | Free Full Text

Lund O, Nielsen M, Kesmir C, *et al.*: **Definition of supertypes for HLA molecules using clustering of specificity matrices.** *Immunogenetics.* 2004; **55**(12): 797–810.
PubMed Abstract | Publisher Full Text

Lundegaard C, Lamberth K, Harndahl M, *et al.*: **NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11.** *Nucleic Acids Res.* 2008; **36**(Web Server issue): W509–12.
PubMed Abstract | Publisher Full Text | Free Full Text

Mäkelä MJ, Puhakka T, Ruuskanen O, *et al.*: **Viruses and bacteria in the etiology of the common cold.** *J Clin Microbiol.* 1998; **36**(2): 539–42.
PubMed Abstract | Publisher Full Text | Free Full Text

Matoba Y, Abiko C, Ikeda T, *et al.*: **Detection of the human coronavirus 229E, HKU1, NL63, and OC43 between 2010 and 2013 in Yamagata, Japan.** *Jpn J Infect Dis.* 2015; **68**(2): 138–41.
PubMed Abstract | Publisher Full Text

Neefjes J, Jongsma ML, Paul P, *et al.*: **Towards a systems understanding of MHC class I and MHC class II antigen presentation.** *Nat Rev Immunol.* 2011; **11**(12): 823–36.
PubMed Abstract | Publisher Full Text

Nguyen A, David JK, Maden SK, *et al.*: **Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2.** *J Virol.* 2020; **94**(13): e00510–20.
PubMed Abstract | Publisher Full Text

Nielsen M, Lundegaard C, Lund O, *et al.*: **The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage.** *Immunogenetics.* 2005; **57**(1–2): 33–41.
PubMed Abstract | Publisher Full Text

Paul WE: **The immune system**. In *Fundamental Immunology* (seventh edition). Ed.: Paul WE. Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2013; 1–21.
Reference Source

Plotkin SL, Plotkin SA: **A Short History of Vaccination**. In *Plotkin's Vaccines* (7th Edition). Eds.: Plotkin SA, Orenstein W, Offit PA, Edwards KM. Elsevier, Philadelphia, 2018; 1–15.
Publisher Full Text

Polakos NK, Drane D, Cox J, *et al.*: **Characterization of hepatitis C virus core-specific immune responses primed in rhesus macaques by a nonclassical ISCOM vaccine.** *J Immunol.* 2001; **166**(5): 3589–98.
PubMed Abstract | Publisher Full Text

Qiu T, Mao T, Wang Y, *et al.*: **Identification of potential cross-protective epitope between a new type of coronavirus (2019-nCoV) and severe acute respiratory syndrome virus.** *J Genet Genomics.* 2020; **47**(2): 115–7.
PubMed Abstract | Publisher Full Text | Free Full Text

Rammensee HG, Friede T, Stevanoviíc S: **MHC ligands and peptide motifs: first listing.** *Immunogenetics.* 1995; **41**(4): 178–228.
PubMed Abstract | Publisher Full Text

Reber AJ, Chirkova T, Kim JH, *et al.*: **Immunosenescence and Challenges of Vaccination against Influenza in the Aging Population.** *Aging Dis.* 2012; **3**(1): 68–90.
PubMed Abstract | Free Full Text

Robinson J, Barker DJ, Georgiou X, *et al.*: **IPD-IMGT/HLA Database.** *Nucleic Acids Res.* 2020; **48**(D1): D948–D955.
PubMed Abstract | Publisher Full Text | Free Full Text

Schellens IM, Hoof I, Meiring HD, *et al.*: **Comprehensive Analysis of the Naturally Processed Peptide Repertoire: Differences between HLA-A and B in the Immunopeptidome.** *PLoS One.* 2015; **10**(9): e0136417.
PubMed Abstract | Publisher Full Text | Free Full Text

Schulze K, Staib C, Schätzl HM, *et al.*: **A prime-boost vaccination protocol optimizes immune responses against the nucleocapsid protein of the SARS coronavirus.** *Vaccine.* 2008; **26**(51): 6678–84.
PubMed Abstract | Publisher Full Text | Free Full Text

Slingluff CL Jr: **The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination?** *Cancer J.* 2011; **17**(5): 343–50.
PubMed Abstract | Publisher Full Text | Free Full Text

Stern LJ, Wiley DC: **Antigenic peptide binding by class I and class II histocompatibility proteins.** *Structure.* 1994; **2**(4): 245–51.
PubMed Abstract | Publisher Full Text

Tang F, Quan Y, Xin ZT, *et al.*: **Lack of peripheral memory B cell responses in recovered patients with severe acute respiratory syndrome: a six-year follow-up study.** *J Immunol.* 2011; **186**(12): 7264–8.
PubMed Abstract | Publisher Full Text

Thiel V, Herold J, Schelle B, *et al.*: **Infectious RNA transcribed *in vitro* from a cDNA copy of the human coronavirus genome cloned in vaccinia virus.** *J Gen Virol.* 2001; **82**(Pt 6): 1273–1281.
PubMed Abstract | Publisher Full Text

van der Hoek L, Pyrc K, Jebbink MF, *et al.*: **Identification of a new human coronavirus.** *Nat Med.* 2004; **10**(4): 368–73.
PubMed Abstract | Publisher Full Text | Free Full Text

van Montfoort N, van der Aa E, Woltman AM: **Understanding MHC class I presentation of viral antigens by human dendritic cells as a basis for rational design of therapeutic vaccines.** *Front Immunol.* 2014; **5**: 182.
PubMed Abstract | Publisher Full Text | Free Full Text

Vijgen L, Keyaerts E, Moës E, *et al.*: **Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event.** *J Virol.* 2005; **79**(3): 1595–604.
PubMed Abstract | Publisher Full Text | Free Full Text

Wagner A, Weinberger B: **Vaccines to Prevent Infectious Diseases in the Older Population: Immunological Challenges and Future Perspectives.** *Front Immunol.* 2020; **11**: 717.
PubMed Abstract | Publisher Full Text | Free Full Text

Wasmoen TL, Kadakia NP, Unfer RC, *et al.*: **Protection of cats from infectious peritonitis by vaccination with a recombinant raccoon poxvirus expressing the nucleocapsid gene of feline infectious peritonitis virus.** *Adv Exp Med Biol.* 1995; **380**: 221–8.
PubMed Abstract | Publisher Full Text

Weingartl H, Czub M, Czub S, *et al.*: **Immunization with modified vaccinia virus Ankara-based recombinant vaccine against severe acute respiratory syndrome is associated with enhanced hepatitis in ferrets.** *J Virol.* 2004; **78**(22): 12672–6.
PubMed Abstract | Publisher Full Text | Free Full Text

Weiss SR, Navas-Martin S: **Coronavirus pathogenesis and the emerging pathogen severe acute respiratory syndrome coronavirus.** *Microbiol Mol Biol Rev.* 2005; **69**(4): 635–64.
PubMed Abstract | Publisher Full Text | Free Full Text

Woo PC, Lau SK, Chu CM, *et al.*: **Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia.** *J Virol.* 2005; **79**(2): 884–95.
PubMed Abstract | Publisher Full Text | Free Full Text

Wu F, Zhao S, Yu B, *et al.*: **A new coronavirus associated with human respiratory disease in China.** *Nature.* 2020; **579**(7798): 265–269.
PubMed Abstract | Publisher Full Text | Free Full Text

Yadav M, Jhunjhunwala S, Phung QT, *et al.*: **Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing.** *Nature.* 2014; **515**(7528): 572–6.
PubMed Abstract | Publisher Full Text

Zhao P, Cao J, Zhao LJ, *et al.*: **Immune responses against SARS-coronavirus nucleocapsid protein induced by DNA vaccine.** *Virology.* 2005; **331**(1): 128–35.
PubMed Abstract | Publisher Full Text | Free Full Text

Zhao J, Zhao J, Perlman S: **T cell responses are required for protection from clinical disease and for virus clearance in severe acute respiratory syndrome coronavirus-infected mice.** *J Virol.* 2010; **84**(18): 9318–25.
PubMed Abstract | Publisher Full Text | Free Full Text

Zhao J, Zhao J, Mangalam AK, *et al.*: **Airway Memory CD4^{+} T Cells Mediate Protective Immunity against Emerging Respiratory Coronaviruses.** *Immunity.* 2016; **44**(6): 1379–91.
PubMed Abstract | Publisher Full Text | Free Full Text

Zhou W, Wang W, Wang H, *et al.*: **First infection by all four non-severe acute respiratory syndrome human coronaviruses takes place during childhood.** *BMC Infect Dis.* 2013; **13**: 433.
PubMed Abstract | Publisher Full Text | Free Full Text

Zhou P, Yang XL, Wang XG, *et al.*: **A pneumonia outbreak associated with a new coronavirus of probable bat origin.** *Nature.* 2020; **579**(7798): 270–273.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 22 June 2020

https://doi.org/10.5256/f1000research.25889.r63246

✔ **Andrea Sant**
[1] Department of Microbiology and Immunology, University of Rochester Medical Center, Rochester, NY, 14642, USA
[2] David H. Smith Center for Vaccine Biology and Immunology, Rochester, NY, USA

These analyses of potential cross reactive CD8 T cell epitopes between the current SARS-CoV-2 and “seasonal” endemic human CoV is useful and timely and the discussion is balanced.

There are several modifications that I believe would improve the clarify and value of the manuscript.

Based on the first sentence of the paragraph entitled “the possibility of matching linear epitopes…”, the authors sate that the two major arms of immune memory…are antibody and CD8 T cells” I believe this is incorrect, as CD4 T cells can directly impact lung pathology and contribute to both protective and pathological immune responses. In fact a recent paper uploaded to BioRxiv suggested that it was populations in the CD4 T cell compartments that correlated with disease severity. The authors should acknowledge that all three subsets of the adaptive response (B cells, CD8 and CD4 T cells) are likely to be important, but this manuscript focusses on CD8 epitopes.

The authors refer to the “software owners” when describing cutoffs. They are perhaps better described as software “designers”.

When discussing “Vaccine Potential”, the authors state that the secondary response is “faster and stronger”. This should be more accurately described, with some references, in a way that points out the higher frequency of responding cells during memory recall, and lower thresholds of TcR engagement needed for T cell activation, both qualities that contribute to a competitive advantage of memory cells.

Because the nature of CD8 memory to the different antigens screened by the authors is not known, the epitopes identified may or may not be targets of cross reactive memory recall. Therefore, the word “expected” should be substituted for “Potential” or some other word that indicates that the epitope list includes candidates but not expected epitopes.

I think the Table could be made quite a lot smaller and thus more valuable to the reader. The source

proteins could be indicated as an abbreviation provided in the legend as could the various seasonal strains. The boxes could then be quite small, and either be positive or negative. In any case, an effort should be made to condense this table.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Immunology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 08 Jul 2020

Johannes M. Dijkstra, Fujita Health University, Toyoake, Japan

Dear Dr. Sant,

Thank you for reviewing our article. We highly appreciate your comment that our article is timely and well balanced. Especially the latter is important, since we did not want to make a general audience too enthusiastic, while we simultaneously wanted to stress that there is a chance that this MHC-I-mediated immunity might (possibly after boostering by vaccines) give real protection.

You are correct that immune memory by $CD4^+$ cells is not only relevant to their control of B cell and CD8 T cell responses, and we have rewritten the sentence and paragraph on the “major arms of immune memory.” As you state, our paper focusses on $CD8^+$ T cell memory indeed, because that is the memory response that can be predicted most reliably by sequence analysis alone. In the notification section we now shortly discuss two papers that appeared after we submitted our paper, and which claim to have found SARS-CoV-2 recognizing $CD4^+$ and $CD8^+$ T cell memory derived from previous common coronavirus infections.

As requested, we have now changed “software owners” to “software designers.”

As requested, we now have added more references of an enhanced immune reaction after a second (booster) immunization. However, we do not feel that for our type of paper it is necessary to discuss the mechanisms of memory T cells besides just mentioning their involvement.

As for the words “expected” versus “potential.” We feel that we used the word “expected” correctly. Table 1, where all the peptides are listed, carries neither of these two words and has a very neutral title. In the text, some peptides are referred to as “expected” to bind a particular MHC molecule, a term clearly relating to the indicated software and literature. Given the large number of potential MHC-I epitopes shared between the viruses, we “expect” previous common coronavirus infections to have induced some CD8+ T cell immune memory that recognizes SARS-CoV-2; this claim is not about the protective value of this memory, and we feel, therefore, that the word “expect” is within reasonability.

As for the Table format. The format was chosen by the journal editorial team, and we can see that for some uses it has advantages. However, we understand your concern, and now have added an Excel format variant of the Table to the supplement section so that readers can more easily view and interact with the data.

Apart from addressing the reviewer’s comments, we corrected a mistake and now informed the readers that there is a single 8 aa match between compared S proteins.

Apart from addressing the reviewer’s comments, we now also added the information that the study by Callow *et al.* (1990) on HCoV-229E, concluded imperfect immune memory protection even by live virus infection one year before challenge.

We are aware of the time and effort reviewing takes, and are very thankful of your thoughtful contribution. It lifts our spirits that you and the other reviewers consider our article a nice contribution to the COVID-19 studies.

Sincerely,
also on behalf of Keiichiro Hashimoto

Hans (J.M.) Dijkstra

Competing Interests: No competing interests were disclosed.

Reviewer Report 22 June 2020

https://doi.org/10.5256/f1000research.25889.r63655

Anna Gil
University of Massachusetts Medical School, Worcester, MA, USA

Liisa K. Selin
Department of Pathology, University of Massachusetts Medical School, Worcester, MA, USA

This manuscript reports a very useful study that extends our knowledge of peptide-MHC recognition by CD8+ T cells during emerging virus infections such as SARS-CoV-2. Detailed *in silico* analysis showed the presence of potential epitopes shared between new types of betacoronavirus: SARS-CoV-2 and common human alphacoronaviruses: OC43, HKU1, 229E and NL63. Due to the high prevalence of the common coronaviruses authors suggest that the large part of the human population has already some degree of specific memory T cell response before having been infected with the virus.

As authors already mentioned in their manuscript the similar study by Nguyen A. et al (JVI, 2020[1]) demonstrated the HLA binding affinity of all possible 8- to 12-mers from SARS-CoV-2 proteome. This group found that HLA-B15:03 type has the greatest capacity to present highly conserved peptides which are shared among coronaviruses suggesting a cross-protective T cell immune response. In current manuscript using different prediction software authors identified and showed the sequence of epitopes which bind well to similar HLA type, HLA-B15:01. Interestingly, one of the epitopes (YLRKHFSM) can be bound by 4 different HLA types. The obvious strength of this study is the demonstration that certain epitopes, which are identical between SARS-CoV-2 and the common human coronaviruses are being predicted as high affinity binders in multiple HLA-A and B types.

Overall, the work reports important new details about SARS-CoV-2 epitopes theoretically being recognized by human CD8+ T cells. Undeniably, future experiments can prove if generated memory immune responses are specific to the proposed epitopes.

There are some suggestions:

1. The analysis of p/MHCI binding for HLA-C type (if available) would certainly complete the list of presented epitopes

2. The introduction part subtitled: ”The possibility of matching linear epitopes…” has missing information about previously published reports regarding T cell response in individuals infected with coronaviruses, either common or SARS-CoV.

3. In the discussion part readers may wonder why the authors did not discuss their findings with those already published (although they may not have been out at the time of submission) but should be included in the revision.

References

1. Nguyen A, David JK, Maden SK, Wood MA, et al.: Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2.*J Virol*. 2020; **94** (13). PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: T cel viral immunology

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 08 Jul 2020

Johannes M. Dijkstra, Fujita Health University, Toyoake, Japan

Dear Dr. Gil and Dr. Selin,

Thank you for your kindness to review our article. We are glad that you find our study very useful, and that you agree with the conclusions. Coming from experts like you, this is very reassuring.

We have now added a bit more information on the Nguyen *et al*. study in the Notification section. The advantage of the Nguyen *et al.* study was that they were more complete on SARS-CoV-2 MHC epitope predictions and made an association with MHC allele distributions. The advantage of our paper is a more concentrated focus on the memory expected from previous coronavirus infections, and the vaccination potential deriving from that memory. They were first, which we acknowledged, although we wrote our paper independent of their article, and during the submission process of our article theirs was not an indexed publication yet. Presumably, this type of overlap will happen a lot with a topic as intensively studied as coronavirus, and we feel we took a reasonable approach for dealing with their study. From your reviews, we understand that you and the other reviewer, Dr. Sant, find this within acceptability.

Thank you for referring to the interesting YLRKHFSMMIL stretch which is identical between HCoV-NL63 and SARS-CoV-2, as indeed it harbors predicted binding epitopes for several MHC-I supertypes. However, we prefer not to discuss this in text form, because there are many uncertainties (e.g., about recent HCoV-NL63 distributions in the world population) and a textual discussion may not add clarity to the table presentation.

Based on your advice, we now have added HLA-C predictions to Table 1.

Likewise, we now have added a more extensive summary of previous reports on T cell memory after coronavirus infections.

Apart from addressing the reviewer's comments, we corrected a mistake and now informed the

readers that there is a single 8 aa match between compared S proteins.

Apart from addressing the reviewer's comments, we now also added the information that the study by Callow *et al.* (1990) on HCoV-229E, concluded imperfect immune memory protection even by live virus infection one year before challenge.

Again, we like to thank you for the reviewing, as we are aware of the time and effort that it takes. We are very happy that our article is appreciated, since it deals with such an important topic.

Sincerely,
also on behalf of Keiichiro Hashimoto

Hans (J.M.) Dijkstra

Competing Interests: No competing interests were disclosed.

Comments on this article

Version 1

Author Response 02 Jun 2020

Johannes M. Dijkstra, Fujita Health University, Toyoake, Japan

In this comment I would just like to state that we submitted this article to F1000Research on April 15 (Japanese time). F1000Research only lists the publication date (April 23) which was after the journal edited our manuscript.

Also on behalf of Keiichiro Hashimoto,

Hans Dijkstra

Competing Interests: No competing interests were disclosed.

Author Response 28 Apr 2020

Johannes M. Dijkstra, Fujita Health University, Toyoake, Japan

Dear Dr. Bercovier,

Thank you for your comments and interesting links. Especially the data on the protective effects against common coronaviruses by previous infections by homologous virus are relevant.
In the F1000Research publication system, two or three reviewers will comment on-line, after which we as authors (probably) have to modify the manuscript following their comments. If those comments are in line with yours, we can implement some of the information that you provided.
Having said that, we also like to keep the message simple and concentrate on the MHC-I restricted

immune memory against SARS-CoV-2 that can be expected to already exist in many people as induced by common coronaviruses. Especially for the elderly this is interesting in regard to vaccination.

Sincerely,

Hans Dijkstra

Competing Interests: No competing interests were disclosed.

Reader Comment 27 Apr 2020

Herve Bercovier, Hebrew University of Jerusalem, Faculty of Medicine Ein Kerem, Israel

Dear Colleagues,
I am in the middle of writing a similar paper and I would like you to take in consideration the following data that will make my paper unnecessary (I do not need any additional paper for my career) if it was added to your interesting article.
1) data on serological cross reaction among the Coronavirus that last for a while. (For à review https://www.medrxiv.org/content/10.1101/2020.04.14.20065771v1.full.pdf ; and I have more papers including challenged studies in volunteer with old corona viruses, https://academic.oup.com/jid/article/219/12/1913/5307035; https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2271881/pdf/epidinfect00023-0213.pdf) 2)CD4 data in a BioRxiv paper in healthy blood donors in Germany that cross with SARS-COV-2 peptides (https://www.medrxiv.org/content/10.1101/2020.04.17.20061440v1.full.pdf) and cross CD4 peptides among new coronavirses (https://pubmed.ncbi.nlm.nih.gov/27287409/?from_term=Coronavirus+cross+protection&from_pos=6) 3) The hypothesis that these immune cross reactions (CD8, CD4, antibodies) prevent colonization of SARS-COV-2 in children aged 0-9 years who are permanently infected in kinder garden and primary school with these old common cold corona viruses. These children do not seem, as a result, able to contaminate neither siblings nor parents. The remaining immune memory would also explain why children aged 10-19 years who can be infected, do not usually develop any invasive serious disease caused by SARS-COV-2 but are able to infect other children and adults. Likewise, it would explain why the COVID19 clinical presentation is worsening with age.

I hope you will consider my remarks and integrate these data into your paper and it could be then a very comprehensive article on the subject of potential cross protection between the different beta Coronavirus.

sincerely yours,

H. Bercovier.

Competing Interests: I have no competing interests.