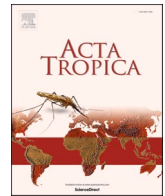




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Peptides of *H. sapiens* and *P. falciparum* that are predicted to bind strongly to HLA-A*24:02 and homologous to a SARS-CoV-2 peptide

Yekbun Adiguzel

Department of Biophysics, School of Medicine, Altinbas University, Kartaltepe Mah. Incirli Cad. No11, 34147 Bakirkoy, Istanbul, Turkey

ARTICLE INFO

Keywords:

SARS-CoV-2
Plasmodium
human leukocyte antigen
molecular mimicry
autoimmunity
disease susceptibility

ABSTRACT

Aim: This study is looking for a common pathogenicity between SARS-CoV-2 and Plasmodium species, in individuals with certain HLA serotypes.

Methods: 1. Tblastx searches of SARS-CoV-2 are performed by limiting searches to five Plasmodium species that infect humans. 2. Aligned sequences in the respective organisms' proteomes are searched with blastp. 3. Binding predictions of the identified SARS-CoV-2 peptide to HLA supertype representatives are performed. 4. Blastp searches of predicted epitopes that bind strongly to the identified HLA allele are performed by limiting searches to *H. sapiens* and Plasmodium species, separately. 5. Peptides with minimum 60% identity to the predicted epitopes are found in results. 6. Peptides among those, which bind strongly to the same HLA allele, are predicted. 7. Step-4 is repeated by limiting searches to *H. sapiens*, followed by the remaining steps until step-7, for peptides sourced by Plasmodium species after step-6.

Results: SARS-CoV-2 peptide with single letter amino acid code CFLGYFCTCYFGLFC has the highest identity to *P. vivax*. Its YFCTCYFGLF part is predicted to bind strongly to HLA-A*24:02. Peptides in the human proteome both homologous to YFCTCYFGLF and with a strong binding affinity to HLA-A*24:02 are YYCARRFGLF, YYCHCPFGVVF, and YYCQQYFFLF. Such peptides in the Plasmodium species' proteomes are FFYTFYFELF, YFVACLFILF, and YFPTTFHLF. The first one belonging to *P. falciparum* has a homologous peptide (YFYLFSLLELF) in the human proteome, which also has a strong binding affinity to the same HLA allele.

Conclusion: Immune responses to the identified-peptides with similar sequences and strong binding affinities to HLA-A*24:02 can be related to autoimmune response risk in individuals with HLA-A*24:02 serotypes, upon getting infected with SARS-CoV-2 or *P. falciparum*.

1. Introduction

COVID-19 pandemic is an ongoing global crisis with devastating effects. Chloroquine derivatives and hydroxychloroquine were among the treatment options (Perricone et al., 2020). However, they were not deemed significantly effective in decreasing death rates or mechanical ventilation risk (Magagnoli et al., 2020), in addition to serious side effects (Touret and De Lamballerie, 2020). Chloroquine and hydroxychloroquine are registered compounds for the treatment and prevention of malaria and for treatment of autoimmune diseases (Perricone et al., 2020). Accordingly, this study is initiated by searching for a relationship between SARS-CoV-2 and five Plasmodium species that cause malaria in humans (Blanquart and Gascuel, 2011; Déchamps et al., 2010), followed by investigating potential of autoimmune reactions through molecular mimicry. In relation, SARS-CoV-2 and *Plasmodium falciparum* common immunodominant regions were suggested to be explaining low

COVID-19 incidence in the "malaria-endemic belt" (Iesa et al., 2020). That study was initiated mainly due to the small number of confirmed cases of COVID-19 at the African region with high malaria-prevalence (references 4–6 of (Iesa et al., 2020)). This lower incidence was highlighted in an earlier study as well (references 1,2 of (Mehta et al., 2020)). The presence of zoonotic viral genome in *P. vivax* (Charon et al., 2019), which was pointed also by others (De Souza, 2020), was mentioned as a basis of related work (Mehta et al., 2020). Also, one case study reported co-infection with COVID-19 and *P. vivax* in a 10-year-old boy (Kishore et al., 2020). The boy received incomplete primaquine therapy six months before, for the *P. vivax* infection. Authors concluded that it was a relapse of earlier infection, triggered by SARS-CoV-2. Yet, re-infection with *P. vivax* simultaneously with COVID-19 or a relapse in its natural course was not ruled out. However, incidences of malaria and SARS-CoV-2 were found to have a negative correlation (Mehta et al., 2020). The genomic similarity between various strains of SARS-CoV-2

E-mail address: yekbun.adiguzel@altinbas.edu.tr.

<https://doi.org/10.1016/j.actatropica.2021.106013>

Received 23 February 2021; Received in revised form 27 May 2021; Accepted 10 June 2021

Available online 17 June 2021

0001-706X/© 2021 Elsevier B.V. All rights reserved.

and the matryoshka RNA virus 1 (MaRNAV-1), associated with *P. vivax*, was also investigated (Charon et al., 2019). Accordingly, 15-base pair (bp) alignment with 93.75% identity (e-value 3.1) between SARS-CoV-2 (NC_045512.2) 22487–22502 bp region and 2796–2781 bp region of MaRNAV-1 was observed. The aligned region, which is encoding surface glycoprotein, was suggested to be small to draw any conclusion. Yet, authors stated that a shared immunogenicity cannot be ruled out. Further, their observation was valid in all different strains' data that was used. Cross-species' peptide sharing and high number of matches of amino acids (aa) are expected. Ultimate claim of this work is that such similarities (homologies) carry the potential of leading to inherently related autoimmune pathologies.

The immune response to pathogen proteins is a defence mechanism of the host. However, it can cause adverse effects when pathogen protein has similar sequences to that of humans. That is molecular mimicry, which can lead to autoimmune response. Researchers Kerkar and Vergani (2018) supported this notion by findings of *de novo* autoimmune hepatitis associated with certain viral infections. Earlier, related work on HIV supported the concept of immune network manipulation and autoimmunity induced by HIV, through mechanisms involving molecular mimicry (Metlas et al., 1994; Metlas et al., 1999a; Metlas et al., 1999b). Associated work pointed at related problems in HIV-1 gp120-based vaccines (Veljkovic et al., 1997). Similarities and pathway analysis of homologous peptides implied viral infection associated immune response pathologies (Carter, 2011). Kanduc and Shoefeld contributed to the field with several related works (Kanduc and Shoefeld, 2016; Kanduc and Shoefeld, 2018; Kanduc and Shoefeld, 2019) and they considered risks inherent in associated vaccines, investigated vaccines for papilloma virus L1 (HPV L1) proteins of 4 strains, and surface antigen of hepatitis B virus (Kanduc and Shoefeld, 2016). Cross-reactions of immune responses against HPV L1 proteins with 20 human proteins were deemed feasible, revealing a probable connection with pathophysiological conditions, implying that peptide-based vaccines need to be pathogen-specific. Regarding SARS-CoV-2, researchers found homologous sequences between spike glycoprotein of the virus, and human surfactant and related proteins, indicative of associated pathologies (Kanduc and Shoefeld, 2020a). In relation, autoimmune reaction risk of vaccines based on whole SARS-CoV-2 antigens was told to have less likelihood of being revealed through tests on primates. This expectation is due to lower peptide sharing between primates and human pathogens, including SARS-CoV-2 (Kanduc and Shoefeld, 2020b). Researchers suggested "aged mice" as suitable animal model for testing SARS-CoV-2 spike glycoproteins-based vaccines (Kanduc and Shoefeld, 2020c).

The relationship of COVID-19 with immunity is multifaceted (Ateyo et al., 2020; Kaneko et al., 2020; Kuri-Cervantes et al., 2020; Laing et al., 2020; Lucas et al., 2020; Mathew et al., 2020). COVID-19 patients experiencing serious health threats displayed hallmarks of extra-follicular B cell activation like in autoimmune settings (Woodruff et al., 2020). Autoinflammatory and autoimmune conditions in COVID-19 are also reviewed (Rodríguez et al., 2020). COVID-19 was suggested to be linked to autoimmunity through mechanisms such as molecular mimicry and bystander activation. SARS-CoV-2 was also hypothesised to be triggering stress-induced autoimmunity through molecular mimicry (Cappello et al., 2020). Also, studies on cerebrospinal fluids of COVID-19 patients were indicative of autoimmunity Lucchese (2020). Peptide sharing between SARS-CoV-2 proteome and human brainstem respiratory pacemaker proteins (Lucchese and Flöel, 2020a); human heat shock proteins 90 and 60 (Lucchese and Flöel, 2020b); Odorant Receptor 7D4, Poly ADP-Ribose Polymerase Family Member 9, and Solute Carrier Family 12 Member 6 proteins of human (Angileri et al., 2020) were reported. Strong immune cross-reactions with SARS-CoV-2 spike protein antibody (Vojdani and Kharratian, 2020) can infer autoimmunity risk in susceptible individuals, through molecular mimicry (Rodríguez et al., 2020). More than one-third of the immunogenic peptides of SARS-CoV-2 have homology with proteins that are of

importance for adaptive immune system (Lyons-Weiler, 2020). Also, various disorders associated with possible autoimmunity against human peptides homologous to immunogenic SARS-CoV-2 peptides were pointed at, and examination of patients' sera for autoantibodies against those human peptides were suggested Kanduc (2020).

Certain major histocompatibility complex (MHC) allele types can cause risk in certain diseases. For instance, patients expressing human leukocyte antigen (HLA)-DRB1*04:01 are suffering from severe form of rheumatoid arthritis more, and in relation, HLA-DRB1*04:01 has a 5mer, which is shared with *E. coli* heat shock protein (Rojas et al., 2018). This similarity can bear risk in rheumatoid arthritis patients with HLA-DRB1*04:01, when exposed to Enterobacteriaceae. Besides, pathogen and human proteins' peptide similarities are present in systemic lupus erythematosus (James and Harley, 1992), systemic sclerosis (Lunardi et al., 2000), and primary biliary cholangitis (Fussey et al., 1990). Regarding SARS-CoV-2, conserved viral proteome regions with high MHC allele-binding affinities are not found under positive or negative selective pressure (Nguyen et al., 2020). Nevertheless, different studies (e.g., (Cappello, 2020a; Cappello, 2020b; Sedaghat and Karimi, 2020), in addition to those mentioned) suggested related autoimmune reaction risk. Also, association of several alleles through exclusive routes can prevent sorting the risk posed by individual allele types in population-based studies. Additionally, HLA-B*46:01 (Lin et al., 2003; Alicia, 2020) and HLA-B*07:03 (Ng et al., 2004; Alicia, 2020) were found to be leading to susceptibility to the associated disease, SARS.

In Plasmodium infections, pathogen-infected cell-detection by cytotoxic CD8(+) T cells of adaptive immune system was thought to be protective through MHC class I molecules in the liver-stage malaria, but not in the blood stage malaria because erythrocytes do not express MHC class I molecules (Rivera-Correa and Rodriguez, 2020). Finding that *P. vivax*-infected reticulocytes keep expressing MHC class I molecules changed this concept (Junqueira et al., 2018). Regarding molecular mimicry, that between histamine-releasing factor of human and translationally controlled tumour protein of *P. falciparum* (MacDonald et al., 2001); vitronectin and erythrocyte membrane protein 1 of *P. falciparum* (Ludin et al., 2011); and ankyrin, spectrin, and actin proteins of erythrocytes and *P. vivax* proteins (Mourão et al., 2018), were mentioned (Mourão et al., 2020).

2. Methods

2.1. Blast search

Tblastx (Altschul et al., 1997) at NCBI (NCBI Resource Coordinators, 2017) is used to compare SARS-CoV-2 (NC_045512.2) and five Plasmodium species that infect humans (*P. falciparum*, taxid:5833; *P. malariae*, taxid:5858; *P. vivax*, taxid:5855; *P. ovale*, taxid:36330; *P. knowlesi*, taxid:5850). Separate searches are performed by limiting searches to individual Plasmodium species. Default algorithm parameters are used, except for the number of maximum target sequences. Accordingly, the highest number of allowed target sequences was used. It was 20000 in the first searches and diminished to 5000 later, by the server, for all blast searches.

Aligned sequences with the highest identity in the tblastx outputs are used as inputs of blastp search at NCBI, performed to check the presence of those peptides in the respective organisms' proteomes. Blastp searches are also performed using default parameters, except for increasing the number of maximum target sequences, as mentioned. Sequences are displayed as single letter amino acid (aa) codes.

2.2. MHC class I-binding

SARS-CoV-2 peptide identified in the previous step is used as input for predicting its binding affinities to MHC class I (*MHC class I genes are HLA-A, -B, and -C genes*) (Nguyen et al., 2020) proteins. NetMHC 4.0 (Andreatta and Nielsen, 2016; Nielsen et al., 2003), NetMHCpan 4.1

(Reynisson et al., 2020), PickPocket 1.1 (Zhang et al., 2009), and NetMHCcons 1.1 (Karosiene et al., 2012) are the tools used for predictions. Predictions for binding affinities to 12 HLA supertype representatives (HLA-A*01:01 [A1], HLA-A*02:01 [A2], HLA-A*03:01 [A3], HLA-A*24:02 [A24], HLA-A*26:01 [A26], HLA-B*07:02 [B7], HLA-B*08:01 [B8], HLA-B*27:05 [B27], HLA-B*39:01 [B39], HLA-B*40:01 [B44], HLA-B*58:01 [B58], HLA-B*15:01 [B62]) are performed. Fasta format is used as the input format, to be able to choose different peptide lengths, into which input sequence is chopped. Then, all peptide lengths are selected, and default parameters are used for the rest. In predictions with NetMHCpan, BA option is also used for binding affinities to be included in the outputs. Only strong binders (SBs) are considered in the analysis of outputs. In PickPocket, 1-log50k(aff) is evaluated manually, by considering those ≥ 0.5 as strong binder (SB). 1-log50k(aff) is calculated by taking the difference of logarithm (with base 50000) of the affinity (in nM) from 1. Possible difference in the outcome due to using a peptide as input is checked using protein sequence containing the peptide sequence, as input.

NetCTLpan 1.1 is used to predict cytotoxic T lymphocyte (CTL) epitopes (Larsen et al., 2007; Stranzl et al., 2020). It is used similarly to predict epitopes that bind with 12 HLA supertypes. Default parameters are used as before, but by performing predictions for different peptide lengths separately, as required by the tool. However, change in the outcome due to using a short sequence as input is checked again, by using protein sequence containing the peptide sequence, as input. The aim there is to adjust C-term cleavage weight in the epitope prediction, in case that a significant difference is observed.

Peptides identified as SBs are considered based on the following criteria: If predictions by at least three of the tools estimate such a binding affinity to the same HLA allele, in addition to epitope prediction by NetCTLpan. Those peptides are termed as SB-peptides. Finally, information at the websites of predictors is suggested for further details on the tools.

2.3. Blast search of SB-peptide and specific HLA allele-binding of SB-peptide homologues

2.3.1. Blast search of SB-peptide and identifying SB-peptide homologues

Peptide sequence used as input in Section 2.2 is searched with blastp, as described in Section 2.1, by limiting search to *H. sapiens* (taxid:9606). SB-peptides of the specific HLA allele are searched in blastp results, as follows: Sequences within the results of blastp search are deemed as homologous if they align with query sequence with more than 60% match. Simultaneously, they are expected to contain whole SB-peptide sequence within the query, while there would be no gap in the aligned subject sequence. Analysis is done with the Excel Program of MS Office.

2.3.2. Specific HLA allele-binding predictions of SB-peptide homologues

After the identification of homologous sequences of SB-peptides, their binding to the respective HLA serotype is predicted as described in Section 2.2, but by choosing the same length as that of SB-peptide, and the specific HLA allele of SB-peptide. The remaining is the same as described in Section 2.2. The prediction with the protein sequence containing SB-peptide homologue is not performed here.

Homologous peptides of SB-peptide are searched in the Plasmodium species' proteomes as well. For that, blastp searches are performed as described in Section 2.1, by limiting searches to Plasmodium species. Homologous sequences of SB-peptides within the alignment results are identified if they had more than 60% match, while there is no gap in the aligned subject sequence. Following that, the binding affinities of resulting peptide sequences to the HLA allele of SB-peptide are predicted, as described above. The resulting peptide with strong binding affinity to the same HLA allele as SB-peptide is termed SB'-peptide. Blastp search of SB'-peptide is performed eventually, by limiting search to *H. sapiens*. This step is followed by identifying homologous sequences of SB'-peptide that have more than 60% match, while there is no gap in

the aligned subject sequence. Finally, the binding affinity of the resulting peptide sequence to the same HLA allele as SB'-peptide (and as SB-peptide) is predicted again, like in Section 2.3.2. The outline of the process until here is schematically displayed in Fig. 1.

2.4. Visualisation

TCRpMHCmodels (Jensen et al., 2019) is used to predict the three-dimensional structure of T cell receptor (TCR) alpha and beta chains that are complexed with SB-peptide and MHC class I. Sequences are retrieved from the Protein Data Bank (Berman et al., 2000). TCR alpha and beta chains' sequences are retrieved from the data of H27-14 TCR that is specific for HLA-A24-Nef134-10 (PDBid: 3VXQ; (Shimizu et al., 2013)). MHC's sequence information is retrieved from the data titled "Crystal structure of peptide-HLA-A24 bound to S19-2 V-del-ta/V-beta TCR" (PDBid: 5XOV; (Shi et al., 2017)).

Seq2Logo 2.0 is a sequence logo generator (Thomsen and Nielsen, 2012). It is used for generating sequence profiles of peptides that are SB to the same HLA allele as SB-peptide and have at least 60% similar sequences with SB-peptide. For logo generation, default parameters are used, other than that for pseudo-count correction (weight on prior) for low counts (see <https://services.healthtech.dtu.dk/service.php?Seq2Logo-2.0> for further explanation). Weight on prior option among the adjustable parameters is set to 1 instead of setting to 0, despite expectedly low number of observations. This choice ends up in a logo that appears comparable to the sequence logo of the HLA allele received from MHCmotifViewer (Rapin et al., 2008).

3. Results

3.1. Blast search

In tblastx (Altschul et al., 1997), query and subject nucleotide sequences are translated in six reading frames, followed by protein blastp comparisons (Wheeler and Bhagwat, 2007). Tblastx retrieved no results for *P. falciparum*. Whereas 1 alignment with the highest 70% identity in case of *P. malariae* (supplementary file (s1)), 12 alignments with the highest 73% identity in case of *P. vivax* (s2), 1 alignment with the highest 75% identity in case of *P. ovale* (s3), and 3 alignments with the highest 67% identity in case of *P. knowlesi* (s4), are observed within alignment results of the others. Aligned sequences' lengths are 10, 15, 8, and 12 aa, respectively. Sequences of those alignments are displayed in Table 1. It can be seen there that *P. knowlesi* sequences VVLWLLCCLCWLC and VVLWLCCCWLC both aligned with the LVLWLLCNCFLC sequence of SARS-CoV-2, with the same maximum identities. Palindromic parts are also present in these aligned query and subject sequences (Table 1, e.g., underlined regions in CNFFNCHYFR, CFLGYFCTCYFGLFC, CFVGYFCTCFVGYFC, FCFCSCFC, LVLWLLCNCFLC, VVLWLLCCLCWLC and VVLWLCCCWLC).

Repetitive regions are present in the sequences displayed in Table 1 (e.g., CHYFR of seq.1', CFVGYFC of seq.2', FC and CFC of seq.3', and WLC of seq.4' and seq.4''). Detection of repetitive regions in comparisons of distant genomes is expected. However, low complexity regions were filtered during tblastx search, as indicated for default algorithm parameters. After analysing results of blastp searches of sequences in Table 1, only CFLGYFCTCYFGLFC sequence is found to be present in translation products of SARS-CoV-2 (s6 of s5–13). Not observing the other sequences in the respective proteomes can be because tblastx compares query and subject after translating nucleotide sequences of each of them in six possible different reading frames. Yet, it can be due to some annotation amendment requirements as well (Böhme et al., 2019).

CFLGYFCTCYFGLFC sequence is not just the query sequence with the highest identity to *P. vivax*. That sequence is present in the alignment results of all respective tblastx searches for SARS-CoV-2 (s14–17). Furthermore, it is the most frequently observed query sequence in the alignments (Fig. 2, Table 2, s18). When *P. malariae* is selected as the

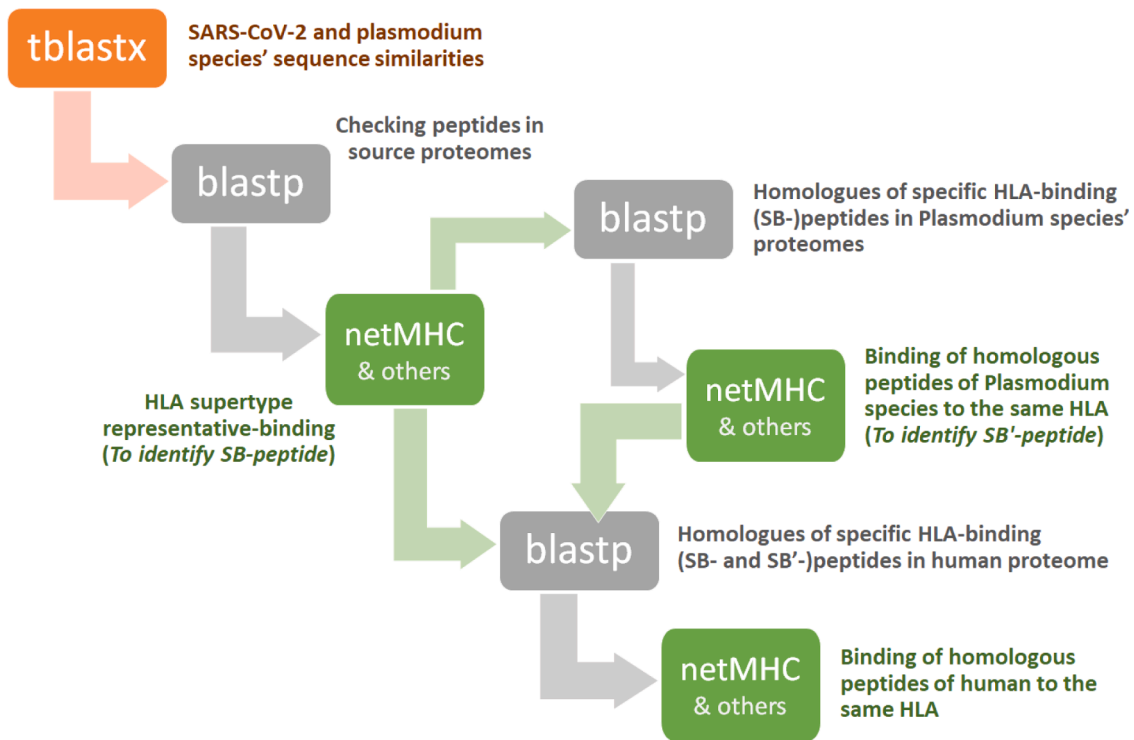


Fig. 1. The outline of the process that is followed.

Table 1
Aligned SARS-CoV-2 query sequences and Plasmodium subject sequences with the highest identities.

Query (COVID-19) sequence	Subject (Plasmodium) sequence	Subject name
-	-	<i>P. falciparum</i>
seq.1 CNFFNCHYFR	seq.1' CHYFRCHYFR	<i>P. malariae</i>
seq.2 CFLGYFCTCYFGLFC	seq.2' CFVGYFCTCFVGYFC	<i>P. vivax</i>
seq.3 FCFHKCFC	seq.3' FCFCSFC	<i>P. ovale</i>
seq.4 LVLWLLCNCFLC	seq.4' VVLWLCCLCWLC	<i>P. knowlesi</i>
	seq.4' VVLWLCCCWLC	

organism in tblastx search of SARS-CoV-2, other alignment results than that with the highest identity contain whole or part of CFLGYFCTCYFGLFC in several queries (*s14*, *s18*). YFCT repeat is abundant in the alignments among alignment-hits with *P. malariae* genome assembly chromosome 10 (sequence ID LT594498.1) (*s14*). So, subject sequences in the alignments contain repeats, i.e., low complexity regions. It was mentioned that there are 12 alignments with the highest (73%) identity to CFLGYFCTCYFGLFC, within alignment results of tblastx search of SARS-CoV-2, limited to *P. vivax* (*s15*). That tblastx search result has several other alignments with the same sequence, as part or whole of aligned query (i.e., 73 alignments of a total number of 143 alignments, see Table 2) (*s15*, *s18*). In case of *P. ovale* and *P. knowlesi*, situation is like

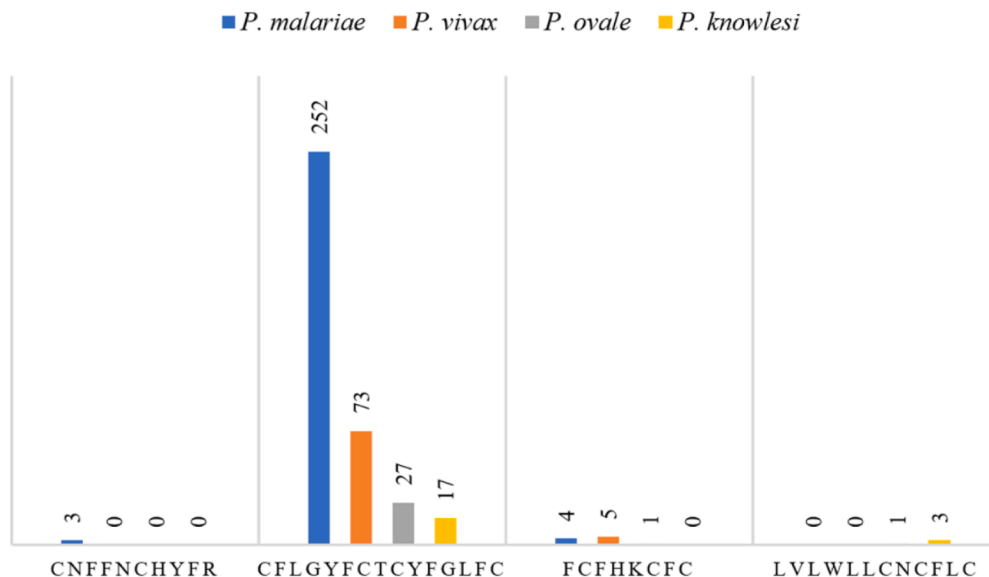


Fig. 2. The number of alignments of whole or part of query sequences with the highest percent identities, within alignment results of tblastx searches of SARS-CoV-2, limited to Plasmodium species.

Table 2

Alignment frequencies of whole or part of query sequences with the highest percent identities, within alignment results of tblastx searches of SARS-CoV-2, limited to *Plasmodium* species.*

	<i>P. malariae</i>	<i>P. vivax</i>	<i>P. ovale</i>	<i>P. knowlesi</i>
CNFFNCHYFR	3 / 693	0 / 143	0 / 145	0 / 124
CFLGYFCTCYFGLFC	252 / 693	73 / 143	27 / 145	17 / 124
FCFHKFCFC	4 / 693	5 / 143	1 / 145	0 / 124
LVLWLLCNCFLC	0 / 693	0 / 143	1 / 145	3 / 124

* Alignment frequency is presented as a ratio of the number of alignments that contain whole or part of the sequence with the highest percent identity to the total number of alignments.

that of *P. malariae* (s16, s17), but the number of alignment-hits with whole or part of CFLGYFCTCYFGLFC is less (Table 2).

Blastp search of CFLGYFCTCYFGLFC sequence reveals that it is part of non-structural protein 6 (nsp6) of replicase polyprotein 1a (ORF1a polyprotein) (s6 of s5–13). Nsp6 is cleaved from ORF1a polyprotein. It is indicated at the UniProtKB (The UniProt Consortium, 2019) (PODTC1) that the subcellular location of nsp6 is host membrane and it is a multi-pass membrane protein, as inferred from sequence similarity to ORF1ab (UniProtKB: P0C6 × 7). The position of CFLGYFCTCYFGLFC sequence is 3784–3798 region of ORF1a polyprotein, corresponding to 215–229 region of nsp6. It is a transmembrane region with helical topology.

3.2. MHC class I-binding

NetMHC 4.0, PickPocket 1.1, and NetMHCpan 4.1 tools are used to predict the binding of CFLGYFCTCYFGLFC to MHC class I proteins. It is indicated (Lundegaard et al., 2008) for an earlier version of NetMHC (v3.0) that the method was used to predict MHC-binding peptides in viral proteomes, including influenza, SARS, and HIV, with an average of 75–80% confirmed binders of MHC.

Peptides of CFLGYFCTCYFGLFC have strong binding affinities for different HLA alleles, as predicted by distinct tools (Table 3, s19, s20). NetMHC predicted 7 SBs and by PickPocket predicted 4 SBs. In case of NetMHC, one 10mer is SB to HLA-B*15:01, and the rest are SBs to HLA-A*24:02. In case of PickPocket, one 9mer is SB to both HLA-A*02:01 and HLA-A*24:02, one 12mer is SB to HLA-A*02:01, and the remaining two are SBs to HLA-A*24:02. NetMHCpan did not predict any SBs (s21). In case of NetMHCcons (Table 3, s22), there are 5 SBs, in which one 12mer is SB to HLA-A*02:01, and the rest are SBs to HLA-A*24:02. As a result, 10mer YFCTCYFGLF and 11mer GYFCTCYFGLF are predicted as SBs to HLA-A*24:02, by NetMHC, PickPocket, and NetMHCcons. Accordingly, they are termed as SB-peptides in the rest of this work.

Prediction results for CFLGYFCTCYFGLFC are like that of nsp6 protein (s23) containing that 15mer (s24–27). SB-peptides are predicted as

Table 3

Results of MHC class I binding predictions for peptides of CFLGYFCTCYFGLFC.* Those written bold are termed as SB-peptides in the rest of this work.

MHC class I type	Length	NetMHC (SB)	PickPocket (≥0.5)	NetMHCcons (SB)	NetCTLpan (E)
HLA-A*24:02	8mer				GYFCTCYF
HLA-A*01:01	9mer				FLGYFCTCY
HLA-A*02:01 HLA-A*24:02			YFCTCYFGL		
HLA-B*15:01	10mer	FLGYFCTCYF			
HLA-A*24:02					GYFCTCYFGL
HLA-A*24:02		YFCTCYFGLF	YFCTCYFGLF	YFCTCYFGLF	YFCTCYFGLF
HLA-A*24:02	11mer			CFLGYFCTCYF	CFLGYFCTCYF
HLA-A*24:02		GYFCTCYFGLF	GYFCTCYFGLF	GYFCTCYFGLF	GYFCTCYFGLF
HLA-A*02:01	12mer		FLGYFCTCYFGL	FLGYFCTCYFGL	
HLA-A*24:02		LGYFCTCYFGLF			
HLA-A*24:02		GYFCTCYFGLFC			
HLA-A*24:02	13mer	FLGYFCTCYFGLF			
HLA-A*24:02	14mer	CFLGYFCTCYFGLF			
				CFLGYFCTCYFGLF	

* NetMHCpan results are not displayed since there was no SB prediction there. (SB: strong binder; ≥0.5 stands for 1-log50k(aff)≥0.5; E: epitope.)

SB to HLA-A*24:02, also when the prediction of the 15mer is performed using the nsp6 protein.

All epitopes, except for one, are predicted to be binding to the HLA-A*24:02 allele (NetCTLpan results in Table 3, s28). One 9mer epitope is predicted to be binding to the HLA-A*01:01 allele. YFCTCYFGLF and GYFCTCYFGLF are predicted as epitopes of HLA-A*24:02, in accordance with the other predictions (Table 3). Those two epitopes are still predicted, and both are binding to HLA-A*24:02 when the 15mer target sequence is predicted as part of the nsp6 protein (s29, note that the sequence is at the 214–228 position of the file). C-term cleavage values are varying, as expected. Accordingly, SB-peptides YFCTCYFGLF and GYFCTCYFGLF of the 15mer are used for further analysis. These two SB-peptides are 91% overlapping, in other words, the 10mer is fully overlapped by the 11mer. Three-dimensional structure model of SB-peptide YFCTCYFGLF in complex with TCR alpha and beta chains, and MHC, is shown in Fig. 3.

3.3. Blastp search for SB-peptide and HLA-A*24:02 binding prediction of SB-peptide homologues

3.3.1. Blastp search and identifying SB-peptides' homologues

Blastp search of CFLGYFCTCYFGLFC is performed by limiting search to *H. sapiens*. Top results in description table belong to immunoglobulin heavy chain junction region (s30). Searching for gapless alignments of two SB-peptides YFCTCYFGLF and GYFCTCYFGLF within results revealed no outcome for the latter. In case of the former, 5 sequences (YFCARNFGPF, YFCASSFGSF, YYCARRFGLF, YYCHCPFGVF, and YYCQYFFLF) are identified (Table 4). The last one (YYCQYFFLF) is not present when blastp search is set to return maximum 5000 target sequences instead of 20000.

Homologous sequences of SB-peptides are identified also in results of CFLGYFCTCYFGLFC blastp searches that are performed separately by limiting searches to individual *Plasmodium* species (*P. falciparum*, *P. malariae*, *P. vivax*, *P. ovale*, and *P. knowlesi*) (s31–35). Accordingly, there are 12 homologous sequences of the first SB-peptide, YFCTCYFGLF (Table 5, until the last row, s36–40). Search for homologous sequences of the second SB-peptide, GYFCTCYFGLF, yielded HYFPITTFHLF sequence of *P. ovale*,

3.3.2. HLA-A*24:02 binding predictions of SB-peptides' homologues

After the identification of homologous sequences of SB-peptides, their binding to HLA-A*24:02 is predicted. Accordingly, all the tools predicted YYCARRFGLF, YYCHCPFGVF, and YYCQYFFLF sequences, to be binding to HLA-A*24:02 (Table 4, s41–45).

The sequences displayed in Table 4 belong to the following proteins in human proteome:

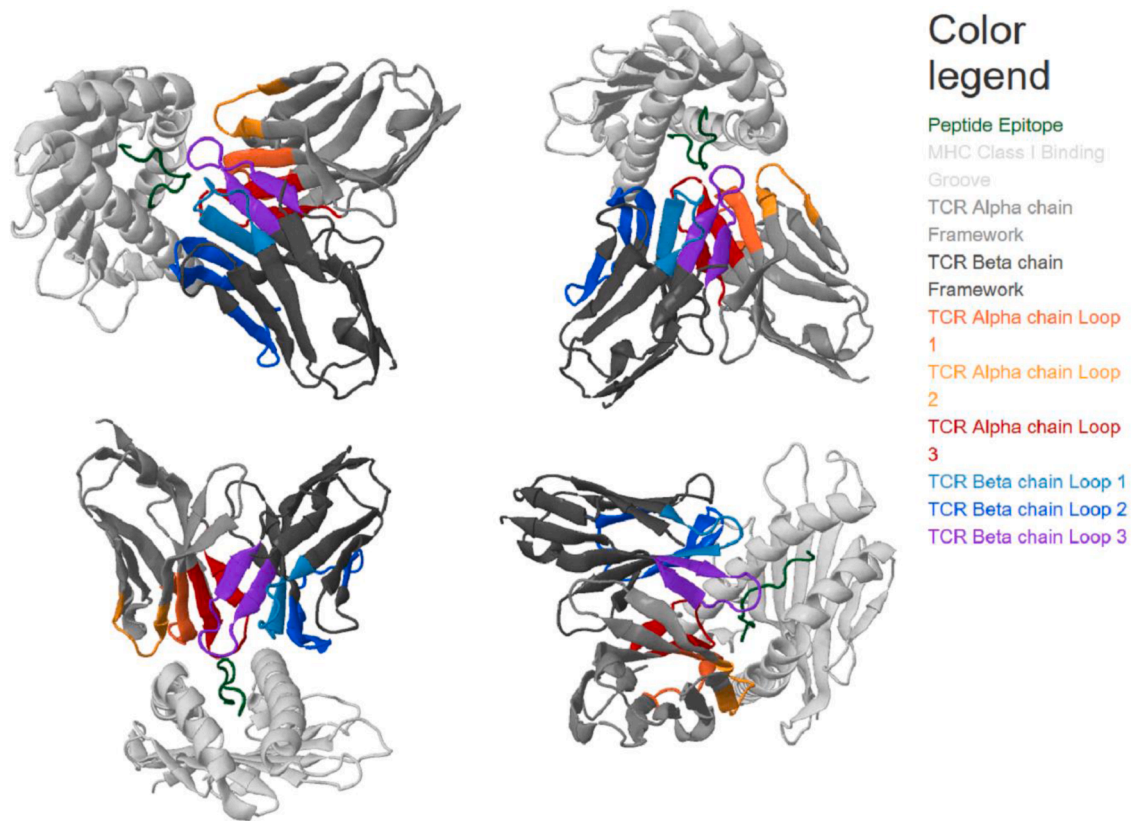


Fig. 3. Three-dimensional structure model of TCR alpha, TCR beta, YFCTCYFGLF, and MHC complex, from different perspectives. The model is obtained through TCRpMHCmodels (Jensen et al., 2019).

Table 4

HLA-A*24:02 binding predictions for human 10mers homologous to the first SB-peptide, YFCTCYFGLF.

Species	10mer sequence	NetMHC (SB / WB / none)*	PickPocket (≥ 0.5 / < 0.5)*	NetMHCpan (SB / WB / none)*	NetMHCcons (SB / WB / none)*	NetCTLpan (E)*
<i>H. sapiens</i>	YFCARNFGPF	SB	< 0.5 (0.436)	WB	SB	E
	YFCASSFGSF	SB	< 0.5 (0.456)	WB	SB	E
	YYCARRFGLF	SB	≥ 0.5 (0.536)	SB	SB	E
	YYCHCPFGVF	SB	≥ 0.5 (0.554)	SB	SB	E
	YYCQYFFLF	SB	≥ 0.5 (0.677)	SB	SB	E

* WB: weak binder; SB: strong binder; ≥ 0.5 and < 0.5 stands for $1-\log_{50k}(\text{aff}) \geq 0.5$ and $1-\log_{50k}(\text{aff}) < 0.5$, respectively; and E: epitope.

Table 5

HLA-A*24:02 binding prediction outcomes of peptides in Plasmodium species' proteomes that are homologous to the first SB-peptide YFCTCYFGLF (until the last row) and of the peptide in human proteome (the last row), which is homologous to the SB'-peptide FFYTFYFELF.*

Species	10mer sequence	NetMHC (SB / WB / none)	PickPocket (≥ 0.5 / < 0.5)	NetMHCpan (SB / WB / none)	NetMHCcons (SB / WB / none)	NetCTLpan (E / none)
<i>P. falciparum</i>	FFCTPFFILF	WB	< 0.5 (0.446)	WB	WB	E
	YICIFYFILF	WB	< 0.5 (0.345)	none	none	none
	YIITCLSGLF	WB	< 0.5 (0.288)	none	none	none
*	FFYTFYFELF	SB	≥ 0.5 (0.500)	WB	SB	E
	YICTGMFSLF	WB	< 0.5 (0.316)	none	WB	none
<i>P. malariae</i>	FFCTKYFAHF	WB	< 0.5 (0.442)	WB	WB	E
	YFVACLFILF	SB	≥ 0.5 (0.503)	WB	SB	E
<i>P. vivax</i>	YFATFYFTLY	SB	< 0.5 (0.334)	none	WB	E
<i>P. ovale</i>	YFPTTFHFLF	SB	≥ 0.5 (0.589)	SB	SB	E
<i>P. knowlesi</i>	YFGTFYFMLY	WB	< 0.5 (0.347)	none	none	E
	FFCACLFLLF	SB	< 0.5 (0.472)	none	SB	E
	YFATFYFTLY	SB	< 0.5 (0.334)	none	WB	E
* <i>H. sapiens</i>	YFYLSLELF	SB	< 0.5 (0.455)	SB	SB	E

* HLA-A*24:02 binding prediction results until the last row present predictions for Plasmodium species' sequences homologous to SB-peptide YFCTCYFGLF. Among those, the ones that are identified as SBs are written bold and termed as SB'-peptide. Results in the last row are the human peptide homologous to FFYTFYFELF, the only peptide identified at the end of the search for homologous peptides of SB'-peptides in human proteome. (WB: weak binder; SB: strong binder; ≥ 0.5 and < 0.5 stands for $1-\log_{50k}(\text{aff}) \geq 0.5$ and $1-\log_{50k}(\text{aff}) < 0.5$, respectively; and E: epitope.)

- YFCARNFGPF: Ig heavy chain variable region (Sequence ID (seq. ID) AAQ05358.1).
- YFCASSFGSF: T cell receptor beta chain variable region (seq.ID ANO56516.1).
- YYCARRFGLF: Immunoglobulin heavy chain variable region (seq. ID QGT38216.1).
- YYCHCPFGVF: Protocadherin FAT4 (seq.ID AHN13824.1), unnamed protein product (seq.IDs BAF84150.1, BAG53346.1), protocadherin Fat 4 isoform X2 (seq.ID XP_011530539.1), Fat tumor suppressor homolog 1 (seq.ID EAX05203.1), and protocadherin Fat 4 isoforms 3, 2, and 1 (seq.IDs NP_078858.4 and Q6V0I7.2, NP_001278214.1, and NP_001278232.1, respectively).
- YYCQYFFLF: Anti-HIV immunoglobulin kappa chain variable region (seq.IDs AVQ94610.1 and AVQ94611.1).

As mentioned, there are 12 homologous sequences of YFCTCYFGLF, within *Plasmodium* species' proteomes (Table 5, until the last row, s36–40). Among those, 3 sequences are predicted as SBs to HLA-A*24:02 (Table 5, bold sequences until the last row, s36–40). Those are termed SB-peptides. To determine if any of those SB-peptides are homologous to human proteins while binding to HLA-A*24:02 with high affinity: First, blastp searches of SB-peptides are performed separately, by limiting searches to *H. sapiens* (s46–48). One homologous sequence (YFYLFSLELF) is identified through analysis of blastp search results (Table 5, the last row). That is homologous to SB-peptide FFYTFYFELF. It belongs to LOC441426 protein in human proteome (seq.IDs AAI36787.1 and AAI36791.1), and is also SB to HLA-A*24:02 (Table 5, the last row, and s49–53).

Peptide in *Plasmodium* species' proteomes homologous to the

second SB-peptide GYFCTCYFGLF is the HYFPTITFHLF sequence of *P. ovale*. It is also predicted as SB to HLA-A*24:02 (s54). Analysis of blastp search results of that sequence in human proteome (s55) did not reveal homologous sequences.

3.4. Visualization

Fig. 4 displays sequence profiles of human- and *Plasmodium*-sourced 10mers that are both SBs to HLA-A*24:02 and have 6 identical residues with YFCTCYFGLF, compared to the 9mer motif of HLA-A*24:02. Three-dimensional structure model of TCR alpha, TCR beta, YFCTCYFGLF, and MHC complex is already presented (Fig. 3).

4. Discussion

CFLGYFCTCYFGLFC sequence is identified here as query sequence with the highest identity to a *P. vivax* sequence. This 15mer sequence is prevalently observed within tblastx alignment result and is common within respective results of the other *Plasmodium* species as well (Fig. 2, Table 2). It is suggested to be a transmembrane region (UniProtKB: P0DTC1). In relation, it is mentioned in other studies (Bianchi et al., 2017) that MHC class I molecules present transmembrane helices. YFCTCYFGLF part of this 15mer peptide is predicted as SB to HLA-A*24:02 of HLA supertype representatives. Yet, the other alleles can also be SB to parts of this 15mer and they are eliminated here because we looked for agreement of prediction results of different tools. YYCARRFGLF, YYCHCPFGVF, and YYCQYFFLF sequences in human proteome, which are among homologous sequences of YFCTCYFGLF, are predicted as SB to the same HLA allele (Table 4, bold sequences). Upon

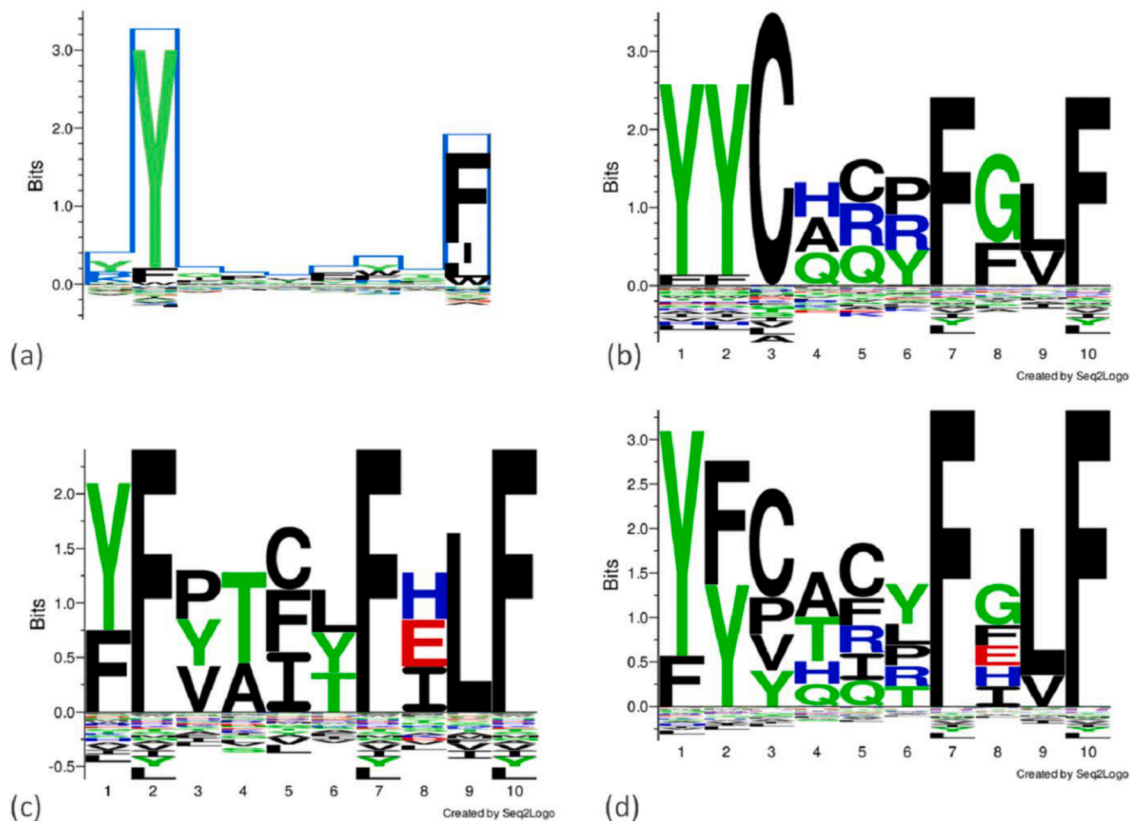


Fig. 4. Sequence logos of 10mers that are both SBs to HLA-A*24:02 and have 6 identical residues with SB-peptide YFCTCYFGLF, compared to HLA-A*24:02 motif. (a) is the 9mer motif of HLA-A*24:02. (b) is the logo generated from YYCARRFGLF, YYCHCPFGVF, and YYCQYFFLF sequences (HLA-A*24:02 binding peptides in human proteome, homologous to YFCTCYFGLF). (c) is the logo generated from FFYTFYFELF, YFVACLFILF, and YFPTITFHLF sequences (HLA-A*24:02 binding peptides in *Plasmodium* species' proteomes, homologous to YFCTCYFGLF). (d) is the logo generated from sequences in (b) and (c). Dimensions of the original file of data in (a) is adjusted to comply with figure dimensions. Website of file for 9mer motif of HLA-A*24:02 in jpg format: https://services.healthtech.dtu.dk/services/MHCmotifViewer/HLA-A_files/Media/HLA-A2429/HLA-A2429.jpg?disposition=download.

being infected with SARS-CoV-2, individuals with HLA-A*24:02 can develop autoimmune responses, accordingly. Further, among sequences in Plasmodium species' proteomes homologous to YFCTCYFLGF, 3 sequences are predicted as SBs to the same HLA allele (Table 5, bold sequences until the last row). Those pathogen sequences can trigger the immune response in HLA-A*24:02 serotypes, upon getting infected with respective Plasmodium species. Yet, the question here is whether it can lead to an autoimmune response as well, through homology to SARS-CoV-2 peptide. Therefore, we looked if any of those are homologous to human proteins that are also SB to the same HLA allele. YFYFLSLELF is identified accordingly (Table 5, the last row).

Aligned SARS-CoV-2 query sequences with gaps are eliminated in this study. Additionally, aligned subject sequences with gaps in regions other than the positions that align with the investigated sequence regions were also eliminated. This process diminished the number of sequences that were obtained. In support, relevant work Adiguzel, 2021a reported a higher number of results by presenting alignments of the same 15mer with human proteome, unrestricted to alignments' gaps. Yet, results outlined here are obtained although restrictions implemented.

Other than the discussions above, this study is conducted using reference genomes, meaning that individual variations in indicated sequences can influence outcomes, but they were not evaluated. Further genetic, physiological, and environmental variations can also contribute. All can influence possible immune responses and autoimmune reactions, but they cannot be evaluated in studies with single, representative reference genomes.

Peptides with homologous regions and strong binding affinities to the same HLA allele possess autoimmune reaction risk (Zabriskie and Freimer, 1966; Kohm et al., 2003; Yuki, 2007; Lule et al., 2017; Negi et al., 2017; Trost et al., 2010; Vellozzi et al., 2014). Intense rise in the number of shared heptapeptides among bacteria and human proteomes is reported (Trost et al., 2010). Yet, antibody generating pathogen proteins are not homologous to human proteins, and vice versa (Amela et al., 2007). These indicate that mere resemblance is generally not the only reason for developing autoimmunity, which should involve more parameters, including environmental factors (Rojas et al., 2018). Distinguishing self from non-self is a probable contributor (Matzinger, 2002). To account for molecular mimicry leading to an autoimmune disease, in addition to homologous host and pathogen proteins, the presence of cross-reacting T-cells or antibodies, epidemiological link and autoimmune disease generation, and reproducibility in an animal model, are required. In accordance, autoimmune reactions were observed in animal models (e.g., (Fujinami et al., 1983; Fujinami and Oldstone, 1985)). However, the latter two criteria are still challenging and prone to concerns (Rojas et al., 2018). In any case, findings of *in silico* studies as this one need to be supported by further evidence, in terms of autoimmunity susceptibility risk. Considering these criticisms, this work presents a condition of autoimmunity susceptibility risk of certain HLA serotypes, upon being infected with SARS-CoV-2 and *P. falciparum*, through homology to the same SARS-CoV-2 peptide.

Implications of the current study also highlight the importance of accounting for genetic variations in vaccine development. Far more genetic variations than HLA serotypes exist. Cross-reactivity between adjuvant-vaccine and patient proteome can possess autoimmune reaction risk in genetically susceptible individuals Kanduc (2012). Therefore, homologies between adjuvant vaccines and human proteome need to be investigated (Kanduc and Shoenfeld, 2016; Kanduc and Shoenfeld, 2018; Kanduc and Shoenfeld, 2019; Kanduc and Shoenfeld, 2020a; Kanduc and Shoenfeld, 2020b; Kanduc and Shoenfeld, 2020c) by taking genetic variations into account. Individuals to be included in the respective clinical trials should be representative of allelic variation of HLA types. As suggested (Nguyen et al., 2020), the implementation of HLA typing into ongoing and upcoming clinical trials and COVID-19 tests would be extremely useful. Such considerations are of importance for tests on animals as well since autoimmune responses would not

necessarily be indicative of the same situation in humans, and vice versa (Kanduc and Shoenfeld, 2020b).

In case of COVID-19, HLA class I and II alleles' prediction from transcriptome sequencing data of bronchoalveolar lavage fluid samples of 5 COVID-19 patients from China also led to the identification of HLA-A*24:02, and it was deemed significant Warren and Birol (2020), compared to the frequency of the allele in South Han Chinese population (Middleton et al., 2003). Authors mentioned that HLA-A*24:02 (A24 allele group) was unknown as a risk factor for SARS, but it is associated with diabetes (Adamashvili et al., 1997; Noble et al., 2002; Nakanishi and Inoko, 2006; Kronenberg et al., 2012), which is known as a risk factor in COVID-19 (Guan et al., 2020). HLA-A*24 is reported as "an independent predictor of 5-year progression to diabetes in autoantibody-positive first-degree relatives" of patients with Type I diabetes (Mbunwe et al., 2013). Malaria is also associated with diabetes (Danquah et al., 2010; Acquah, 2019; Carrillo-Larco et al., 2019). In the mentioned COVID-19 study (Warren and Birol, 2020), HLA class II alleles DPA1*02:02 and DPB1*05:01 were found to be frequent as well. Those alleles are common in Han Chinese and associated with autoimmune diseases Graves' (Chu et al., 2018) and narcolepsy (Ollila et al., 2015). The latter one, DPB1*05:01, is associated with chronic hepatitis B in Asians, in addition to being a risk factor effecting viral infection clearance ability ((Kamatani et al., 2009; Ollila et al., 2015) cited in (Middleton et al., 2003)). Accordingly, class II allele-binding predictions of SARS-CoV-2 15mer identified here is planned as a future work.

4.1. Evolutionary perspective

Peptide sharing events or high number of matches in peptides potentially carry autoimmune response risk in individuals with genetic susceptibilities. Observation of matches in peptides of different species that possess such risks is expected. The probability of a hexapeptide to occur is $1/64000000$ Kanduc (2020), which is calculated through equal likelihood of occurrence of 20 aa at any of the 6 positions in a hexapeptide. Proteome size of the organism, to which blastp search is limited, is the site to realise that possibility: $1/64000000$. An unbiased proteome with 64 million aa would be sufficient for observing any 6mer for once, considering that such a proteome would be comprised of 64 million hexamers (ignoring the dissection of proteome into proteins). 122962 proteins with the total number of 80769298 aa are present in human proteome (retrieved from the protein-list at NCBI, genome assembly id 582967). This size of a proteome can let every possible 6mer to be observed once, at least. In this study, we looked for a minimum of 6 aa-matches in a 10mer and did not observe more than 6 aa-matches. The expected occurrence of 6 aa-matches in a 10mer is 210 and that of observing 6mers is 5. 210 is the number of 6-membered subsets in a 10-membered set, calculated by $10!/(6! \times 4!)$ (each member is the aa at each position of the 10mer). However, we observed only few sequences with 6 aa-matches (Table 4), namely, few of the 6-membered subsets. So, the observation is less than expected. This lessening is because the calculated number is expected number of observations for unbiased conditions, but human proteome is biased, and most proteins have several isomers. These results somehow support views about involvement of evolutionary processes between human hosts and SARS-CoV-2 (Kanduc, 2019; Kanduc, 2020), in relation to a different part of the concept. The evolution of viral proteome can be providing a selection advantage based on infectivity, or elimination by the host. For instance, proteome of *P. falciparum* strain 3D7 reference genome has 5387 proteins with total number of 4103467 aa (data retrieved from the protein-list at NCBI, genome assembly id 895506). It is about 20 times smaller than that of the human proteome. Yet, that proteome also revealed 5 peptides with 6 aa-matches to the 10mer query peptide, the first SB-peptide (Table 5). This amount can be interpreted as less bias in *P. falciparum* proteome than in human proteome. *P. falciparum* proteins do not have several isomers like in case of human proteins, which supports this interpretation. Still, this observation should be validated with

a larger and varied query set. Currently, it is interesting that only one of the three SB-peptides of *Plasmodium* species in Table 5 has a 6 aa-matching peptide in human proteome. This observation can be related to host-pathogen evolution or co-evolution of pathogens, or both. Also, 32 *Plasmodium* sequences are present among 17657 sequences in descriptions table of blastn search of SARS-CoV-2 with somewhat similar sequences (s56), which contributed to this view. In relation, 10mer part of SARS-CoV-2 peptide investigated here have homologous peptides with 6 aa-matches both in human proteome and in *Plasmodium* species' proteomes. Yet only one of those *Plasmodium* species' peptides has one such peptide in human proteome. In relation, shared immunodominant regions between *P. falciparum* and SARS-CoV-2 were suggested to be explaining low incidence of COVID-19 within the malaria-endemic belt (Iesa et al., 2020). Here, autoimmune reaction risk is studied through such a similarity. Accordingly, low incidence in malaria-endemic belt can be due to the higher number of homologous peptides of *Plasmodium* species and SARS-CoV-2. It can cause more protection by triggering immune response than an autoimmune reaction. This claim is supported here by that only one of the three SB-peptides of *Plasmodium* species have a homologous peptide in human proteome that is SB to the same allele (Table 5). Accordingly, probability of a protective effect due to shared peptides between pathogens is higher than their risk of causing autoimmune response. This difference is because only a subset, i.e., a smaller number, of those shared peptides are SB to the same HLA allele and have a homologous peptide in human proteome, which is also SB to the same HLA allele. It can be interpreted also as follows: Host-pathogen evolution and pathogen co-evolution at the "malaria-endemic belt" worked towards diminished autoimmune responses that involve molecular mimicry. It can be protective against severe cases of COVID-19 with autoimmune reactions involving associated mechanisms. Yet, this scenario does not rule out the potential of inherently related autoimmune pathologies through peptide homologies. Besides, such similarities are expected, although physiological outcomes are complex. Still, results of *in silico* studies like this one require experimental validation or clinical support, or both.

Eventually, studies like this one can pave grounds to certain HLA serotypes' recognition as risk groups and identification of novel alleles (Marsh et al., 2010; Cheranev et al., 2021; Robinson et al., 2020). It also highlights the importance of relevant considerations in vaccine studies and can contribute to studies involving commonalities between pathogens of human host, driven by host-pathogen evolution. As outlined nicely (Córdoba-Aguilar et al., 2021), what is as much important is the possibility of lessons that will be learned from the current pandemic to trigger an understanding leading to the development of disease prevention and environmental protection strategies for the well-being of not only humanity but also nature itself.

5. Conclusion

In this study, 10mer peptide at positions 5–14 of CFLGYFCTCYFGLFC sequence of SARS-CoV-2 is identified as a possible relationship between SARS-CoV-2 and *P. falciparum*. Homologous 10mers of that peptide in human proteome are predicted to have strong binding affinities to the HLA-A*24:02 allele. *Plasmodium* species other than *P. falciparum* have homologous 10mers of that peptide with high affinities to the same HLA allele. However, only the 10mer of *P. falciparum* has a homologous 10mer in human proteome, with strong affinity to that HLA allele. These results imply autoimmunity susceptibility of HLA-A*24:02 serotypes through molecular mimicry in SARS-CoV-2 and *P. falciparum* infections, due to homology to the same SARS-CoV-2 peptide. This result is yet to be supported by clinical and experimental findings.

Supporting information

Supplementary files available at shorturl.at/bmzP6 and deposited at

Mendeley Data under the title Dataset1_ADIGUZEL (Adiguzel, 2021b)

Availability of data and materials

Data as supplementary files deposited at Database: Mendeley Data (Dataset1_ADIGUZEL) (Adiguzel, 2021b) and available as supplementary files under the shortened URL (URL shortened at <https://www.shorurl.at/>) of the Google Drive link.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

CRedit authorship contribution statement

Yekbun Adiguzel: Conceptualization, Funding acquisition, Formal analysis, Data curation, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Ecology and Evolutionary Biology Society of Turkey is acknowledged.

References

- Acquah, S., 2019. Linking malaria to type 2 diabetes mellitus: a review. *J. Ghana Sci. Ass.* 18, 56–70.
- Adamashvili, I., McVie, R., Gelder, F., Gautreaux, M., Jaramillo, J., Roggero, T., McDonald, J., 1997. Soluble HLA class I antigens in patients with type 1 diabetes and their family members. *Human Immunol.* 55, 176–183.
- Adiguzel, Y., 2021a. Molecular mimicry between SARS-CoV-2 and human proteins. *Autoimmun. Rev.* 20, 102791.
- Adiguzel, Y., 2021b. Dataset1_ADIGUZEL. Mendeley Data, v1. <https://doi.org/10.17632/ncrs9cy9r.1>.
- Alicia, S.M., 2020. HLA studies in the context of coronavirus outbreaks. *Swiss Med. Wkly.* 150, w20248.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25, 3389–3402.
- Amela, I., Cedano, J., Querol, E., 2007. Pathogen proteins eliciting antibodies do not share epitopes with host proteins: a bioinformatics approach. *PLoS One* 2, e512.
- Andreata, M., Nielsen, M., 2016. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* 32, 511–517.
- Angileri, F., Legare, S., Gammazza, A.M., de Macario, E.C., Macario, A.J., Cappello, F., 2020. Molecular mimicry may explain multi-organ damage in COVID-19. *Autoimmun. Rev.* 19, 102591.
- Atyeo, C., Fischinger, S., Zohar, T., Slein, M.D., Burke, J., Loos, C., McCulloch, D.J., Newman, K.L., Wolf, C., Yu, J., Shuey, K., Feldman, J., Hauser, B.M., Caradonna, T., Schmidt, A.G., Suscovich, T.J., Linde, C., Cai, Y., Barouch, D., Ryan, E.T., et al., 2020. Distinct early serological signatures track with SARS-CoV-2 survival. *Immunity* 53, 524–532.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, E., 2000. The protein data bank. *Nucleic Acids Res* 28, 235–242.
- Bianchi, F., Textor, J., van den Bogaart, G., 2017. Transmembrane helices are an overlooked source of Major Histocompatibility Complex Class I epitopes. *Front. Immunol.* 8, 1118.

- Blanquart, S., Gascuel, O., 2011. Mitochondrial genes support a common origin of rodent malaria parasites and *Plasmodium falciparum*'s relatives infecting great apes. *BMC Evol. Biol.* 11, 70.
- Böhme, U., Otto, T.D., Sanders, M., Newbold, C.I., Berriman, M., 2019. Progression of the canonical reference malaria parasite genome from 2002–2019. *Wellcome Open Res* 4, 58.
- Cappello, F., 2020a. Is COVID-19 a proteiform disease inducing also molecular mimicry phenomena. *Cell Stress Chaperones* 25, 381–382.
- Cappello, F., 2020b. COVID-19 and molecular mimicry: the Columbus' egg. *J. Clin. Neurosci.* 77, 246.
- Cappello, F., Gammazza, A.M., Dieli, F., de Macario, E.C., Macario, A.J., 2020. Does SARS-CoV-2 trigger stress-induced autoimmunity by molecular mimicry? A hypothesis. *J. Clin. Med.* 9, 2038.
- Carrillo-Larco, R., Altez-Fernandez, C., Ugarte-Gil, C., 2019. Is diabetes associated with malaria and malaria severity? A systematic review of observational studies. *Wellcome Open Res* 4, 136.
- Carter, C.J., 2011. Extensive viral mimicry of 22 AIDS-related autoantigens by HIV-1 proteins and pathway analysis of 561 viral/human homologues suggest an initial treatable autoimmune component of AIDS. *FEMS Immunol. Med. Microbiol.* 63, 254–268.
- Charon, J., Grigg, M.J., Eden, J.S., Piera, K.A., Rana, H., William, T., Rose, K., Davenport, M.P., Anstey, N.M., Holmes, E.C., 2019. Novel RNA viruses associated with *Plasmodium vivax* in human malaria and Leucocytozoon parasites in avian disease. *PLoS Pathog* 15, e1008216.
- Cheranev, V., Loginova, M., Jankevici, T., Kutuyavina, S., Rebrikov, D., 2021. A novel allele, HLA-C*15:227, identified when typing COVID-19 patients. *HLA* 97, 377–378.
- Chu, X., Yang, M., Song, Z.J., Dong, Y., Li, C., Shen, M., Zhu, Y.Q., Song, H.D., Chen, S.J., Chen, Z., Huang, W., 2018. Fine mapping MHC associations in Graves' disease and its clinical subtypes in Han Chinese. *J. Med. Genet.* 55, 685–692.
- Córdoba-Aguilar, A., Ibarra-Cerdena, C.N., Castro-Arellano, I., Suzan, G., 2021. Tackling zoonoses in a crowded world: lessons to be learned from the COVID-19 pandemic. *Acta Trop.* 214, 105780.
- Danquah, I., Bedu-Addo, G., Mockenhaupt, F.P., 2010. Diabetes Mellitus and increased risk for malaria infection. *Emerging Infect. Dis.* 16, 1601–1604.
- De Souza, W., 2020. Covid-19 and parasitology. *Parasitol. Res.* 119, 2369–2370.
- Déchamps, S., Maynadier, M., Wein, S., Gannoun-Zaki, L., Maréchal, E., Vial, H.J., 2010. Rodent and nonrodent malaria parasites differ in their phospholipid metabolic pathways. *J. Lipid. Res.* 51, 81–96.
- Fujinami, R.S., Oldstone, M.B., 1985. Amino acid homology between the encephalitogenic site of myelin basic protein and virus: mechanism for autoimmunity. *Science* 230, 1043–1045.
- Fujinami, R.S., Oldstone, M.B., Wroblewska, Z., Frankel, M.E., Koprowski, H., 1983. Molecular mimicry in virus infection: crossreaction of measles virus phosphoprotein or of herpes simplex virus protein with human intermediate filaments. *Proc. Natl. Acad. Sci. Unit. States Am.* 80, 2346–2350.
- Fussey, S.P., Ali, S.T., Guest, J.R., James, O.F., Bassendine, M.F., Yeaman, S.J., 1990. Reactivity of primary biliary cirrhosis sera with *Escherichia coli* dihydrolipoamide acetyltransferase (E2p): characterization of the main immunogenic region. *Proc. Natl. Acad. Sci. Unit. States Am.* 87, 3987–3991.
- Guan, W.J., Liang, W.H., Zhao, Y., Liang, H.R., Chen, Z.S., Li, Y.M., Liu, X.Q., Chen, R.C., Tang, C.L., Wang, T., Ou, C.Q., Li, L., Chen, Y., Sang, L., Wang, W., Li, J.F., Li, C.C., Ou, L.M., Cheng, B., Xiong, S., et al., 2020. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur. Respir. J.* 55, 2000547.
- Iesa, M.A., Osman, M.E., Hassan, M.A., Dirar, A.L., Abuzeid, N., Mancuso, J.J., Pandey, R., Mohammed, A.A., Borad, M.J., Babiker, H.M., Konozy, E.H., 2020. SARS-CoV-2 and *Plasmodium falciparum* common immunodominant regions may explain low COVID-19 incidence in the malaria-endemic belt. *New Microbe., New Infect* 38, 100817.
- James, J.A., Harley, J.B., 1992. Linear epitope mapping of an Sm B/B' polypeptide. *J. Immunol.* 148, 2074–2079. Available: Pubmed ID 1372022.
- Jensen, K.K., Rantos, V., Jappe, E.C., Olsen, T.H., Jespersen, M.C., Jurtz, V., Jessen, L.E., Lanzarotti, E., Lanzarotti, S., Peters, B., Nielsen, M., 2019. TCRpMHCmodels: structural modelling of TCR-pMHC class I complexes. *Sci. Rep.* 9, 14530.
- Junqueira, C., Barbosa, C.R., Costa, A., Teixeira-Carvalho, A., Castro, G., Sen Santana, S., Barbosa, R.P., Dotiwala, F., Pereira, D.B., Antonelli, L.R., Lieberman, J., Gazzinelli, R.T., 2018. Cytotoxic CD8(+) T cells recognize and kill *Plasmodium vivax*-infected reticulocytes. *Nat. Med.* 24, 1330–1336.
- Kamatani, Y., Wattanapokayakit, S., Ochi, H., Kawaguchi, T., Takahashi, A., Hosono, N., Kubo, M., Tsunoda, T., Kamatani, N., Kumada, H., Puseenam, A., Sura, T., Daigo, Y., Chayama, K., Chantrata, W., Nakamura, Y., Matsuda, K., 2009. A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat. Genet.* 41, 591–595.
- Kanduc, D., 2012. Peptide cross-reactivity: the original sin of vaccines. *Front. Biosci.* 4, 1393–1401.
- Kanduc, D., 2019. The comparative biochemistry of viruses and humans: an evolutionary path towards autoimmunity. *Biol. Chem.* 400, 629–638.
- Kanduc, D., 2020. From anti-SARS-CoV-2 immune responses to COVID-19 via molecular mimicry. *Antibodies* 9, 33.
- Kanduc, D., Shoenfeld, Y., 2016. From HBV to HPV: Designing vaccines for extensive and intensive vaccination campaigns worldwide. *Autoimmun. Rev.* 15, 1054–1061.
- Kanduc, D., Shoenfeld, Y., 2018. Inter-pathogen peptide sharing and the original antigenic sin: solving a paradox. *Open Immuno. J.* 8, 16–27.
- Kanduc, D., Shoenfeld, Y., 2019. Human Papillomavirus epitope mimicry and autoimmunity: the molecular truth of peptide sharing. *Pathobiology* 86, 285–295.
- Kanduc, D., Shoenfeld, Y., 2020a. On the molecular determinants of the SARS-CoV-2 attack. *Clin. Immunol.* 215, 108426.
- Kanduc, D., Shoenfeld, Y., 2020b. Medical, genomic, and evolutionary aspects of the peptide sharing between pathogens, primates, and humans. *Global Med. Genet.* 7, 64–67.
- Kanduc, D., Shoenfeld, Y., 2020c. Molecular mimicry between SARS-CoV-2 spike glycoprotein and mammalian proteomes: implications for the vaccine. *Immunol. Res.* 68, 310–313.
- Kaneko, N., Kuo, H.H., Boucay, J., Farmer, J.R., Allard-Chamard, H., Mahajan, V.S., Piechocka-Trocha, A., Lefteri, K., Osborn, M., Bals, J., Bartsch, Y.C., Bonheur, N., Caradonna, T.M., Chevalier, J., Chowdhury, F., Diefenbach, T.J., Einkauf, K., Fallon, J., Feldman, J., Finn, K.K., et al., 2020. Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. *Cell* 183, 1–15.
- Karosiene, E., Lundegaard, C., Lund, O., Nielsen, M., 2012. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* 64, 177–186.
- Kerker, N., Vergani, D., 2018. De novo autoimmune hepatitis –is this different in adults compared to children. *J. Autoimmun.* 95, 26–33.
- Kishore, R., Dhakad, S., Arif, N., Dar, L., Mirdha, B.R., Aggarwal, R., Kabra, S.K., 2020. COVID-19: possible cause of induction of relapse of plasmodium vivax infection. *Indian J. Pediatr.* 3, 1–2.
- Kohm, A.K., Fuller, K.G., Miller, S.D., 2003. Mimicking the way to autoimmunity: an evolving theory of sequence and structural homology. *Trends Microbiol.* 11, 101–105.
- Kronenberg, D., Knight, R.R., Estorninho, M., Ellis, R.J., Kester, M.G., de Ru, A., Eichmann, M., Huang, G.C., Powrie, J., Dayan, C.M., Skowera, A., van Veelen, A., Peakman, M., 2012. Circulating preproinsulin signal peptide-specific CD8 T cells restricted by the susceptibility molecule HLA-A*24 are expanded at onset of type 1 diabetes and kill β -cells. *Diabetes* 61, 1752–1759.
- Kuri-Cervantes, L., Pampena, M.B., Meng, W., Rosenfeld, A.M., Ittner, C.A., Weisman, A. R., Agyekum, R.S., Mathew, D., Baxter, A.E., Vella, L.A., Kuthuru, O., Apostolidis, S. A., Bershaw, L., Dougherty, J., Greenplate, A.R., Pattekar, A., Kim, J., Han, N., Gouma, S., Weirick, M.E., et al., 2020. Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci. Immunol.* 5, eabd7114.
- Vantourout Laing, A.G., Lorenc, A., del Molino del Barrio, I., Das, A., Fish, M., Monin, L., Muñoz-Ruiz, M., McKenzie, D.R., Hayday, T.S., Francos-Quijorna, I., Kamdar, S., Joseph, M., Davies, D., Davis, R., Jennings, A., Zlatareva, I., Wu, Y., Sofra, V., Cano, F., et al., 2020. A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* 26, 1623–1635.
- Larsen, M.V., Lundegaard, C., Lamberth, K., Buus, S., Lund, O., Nielsen, M., 2007. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics* 8, 424.
- Lin, M., Tseng, H.K., Trejaut, J.A., Lee, H.L., Loo, J.H., Chu, C.C., Chen, J., Su, Y.W., Lim, K.H., Tsai, Z.U., Lin, R.Y., Lin, R.S., Huang, C.H., 2003. Association of HLA class I with severe acute respiratory syndrome coronavirus infection. *BMC Med. Genet.* 4, 9.
- Wong-Lu Lucas, C., Klein, J., Castro, T.B., Silva, J., Sundaram, M., Ellingson, M.K., Mao, T., Oh, J.E., Israelow, B., Takahashi, T., Tokuyama, M., Venkataraman, A., Park, A., Mohanty, S., Wang, H., Wylie, A.L., Vogels, C.B., Earnest, R., et al., 2020. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* 584, 463–469.
- Lucchese, G., 2020. Cerebrospinal fluid findings in COVID-19 indicate autoimmunity. *Lancet Microbe* 1, e242.
- Lucchese, G., Flöel, A., 2020a. Molecular mimicry between SARS-CoV-2 and respiratory pacemaker neurons. *Autoimmun. Rev.* 19, 102556.
- Lucchese, G., Flöel, A., 2020b. SARS-CoV-2 and Guillain-Barré syndrome: molecular mimicry with human heat shock proteins as potential pathogenic mechanism. *Cell Stress Chaperones* 25, 731–735.
- Ludin, Nilsson, D., 2011. Genome-wide identification of molecular mimicry candidates in parasites. *PLoS One* 6, e17546.
- Lule, S., Colpak, A.I., Balci-Peynircioglu, B., Gursoy-Ozdemir, Y., Peker, S., Kalyoncu, U., Can, A., Tekin, N., Demiralp, D., Dalkara, T., 2017. Behcet disease serum is immunoreactive to neurofilament medium which share common epitopes to bacterial HSP-65, a putative trigger. *J. Autoimmun.* 84, 87–96.
- Lunardi, C., Bason, C., Navone, R., Millo, E., Damonte, G., Corrocher, R., Puccetti, A., 2000. Systemic sclerosis immunoglobulin G autoantibodies bind the human cytomegalovirus late protein UL94 and induce apoptosis in human endothelial cells. *Nat. Med.* 6, 1183–1186.
- Lundegaard, C., Lamberth, K., Harndahl, M., Buus, S., Lund, O., Nielsen, M., 2008. 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* 36, W509–W512.
- Lyons-Weiler, J., 2020. Pathogenic priming likely contributes to serious and critical illness and mortality in COVID-19 via autoimmunity. *J. Transl. Autoimmun.* 3, 100051.
- MacDonald, S.M., Bhisutthibhan, J., Shapiro, T.A., Rogerson, S.J., Taylor, T.E., Tembo, M., Langdon, J.M., Meshnick, S.R., 2001. Immune mimicry in malaria: *Plasmodium falciparum* secretes a functional histamine-releasing factor homolog in vitro and in vivo. *Proc. Natl. Acad. Sci. Unit. States Am.* 98, 10829–10832.
- Magagnoli, J., Narendran, S., Pereira, F., Cummings, T.H., Hardin, J.W., Sutton, S.S., Ambati, J., 2020. Outcomes of hydroxychloroquine usage in United States veterans hospitalized with COVID-19. *Med* 1, 1–14.
- Parham Marsh, S.G., Albert, E.D., Bodmer, W.F., Bontrop, R.E., Dupont, B., Erlich, H.A., Fernandez-Vina, M., Geraghty, D.E., Holdsworth, R., Hurley, C.K., Lau, M., Lee, K. W., Mach, B., Maiers, M., Mayr, W.R., Müller, C.R., Petersdorf, E.W., Sasazuki, T., Strominger, J.L., et al., 2010. Nomenclature for factors of the HLA system. *Tissue Antigens* 75, 291–455.

- Mathew, D., Giles, J.R., Baxter, A.E., Oldridge, D.A., Greenplate, A.R., Wu, J.E., Alanio, C., Kuri-Cervantes, L., Pampena, M.B., D'Andrea, K., Manne, S., Chen, Z., Huang, Y.J., Reilly, J.P., Weisman, A.R., Ittner, C.A., Kuthuru, O., Dougherty, J., Nzingha, K., Han, N., et al., 2020. Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* 369, eabc8511.
- Matzinger, 2002. The danger model: a renewed sense of self. *Science* 296, 301–305.
- Mbunwe, E., Van der Auwera, B.J., Vermeulen, I., Demeester, S., Van Dalem, A., Balti, E. V., Van Aken, S., Derdelinckx, L., Dorchy, H., De Schepper, J., van Schravendijk, C., Wenzlau, J.M., Hutton, J.C., Pipeleers, D., Weets, I., Gorus, F.K., Belgian Diabetes Registry, 2013. HLA-A*24 is an independent predictor of 5-year progression to diabetes in autoantibody-positive first-degree relatives of Type 1 diabetic patients. *Diabetes* 62, 1345–1350.
- ParikhKulkarni Mehta, Aggarwal, S., Batra, A., Patel, A., Sahoo, R.K., Gupta, V.G., Bahl, C., Kumar, R., Pabbi, S., Sharma, A.P., 2020. Has India met this enemy before? From an eternal optimist's perspective: SARS-CoV-2. *Indian J. Med. Sci.* 72, 8–12.
- Metlas, R., Skerl, V., Veljkovic, V., Colombatti, A., Pongor, S., 1994. Immunoglobulin-like domain of HIV-1 envelope glycoprotein gp120 encodes putative internal image of some common human proteins. *Viral Immunol* 7, 215–219.
- Metlas, R., Trajkovic, D., Srdic, T., Veljkovic, V., Colombatti, A., 1999a. Human immunodeficiency virus V3 peptide-reactive antibodies are present in normal HIV-negative sera. *AIDS Res. Hum. Retrovir.* 15, 671–677.
- Metlas, R., Trajkovic, D., Srdic, T., Veljkovic, V., Colombatti, A., 1999b. Anti-V3 and anti-IgG antibodies of healthy individuals share complementarity structures. *J. Acquir. Immune. Defic. Syndr.* 21, 266–270.
- Middleton, D., Menchaca, L., Rood, H., Komerofsky, R., 2003. New allele frequency database: <http://www.allele-frequencies.net>. *Tissue Antigens* 61, 403–407.
- Grynberg Mourão, L.C., Baptista, R.P., de Almeida, Z.B., Pucci, M.M., Castro-Gomes, T., Fontes, C.J., Rathore, S., Sharma, Y.D., da Silva-Pereira, R.A., Bemquerer, M.P., Braga, E.M., 2018. Anti-band 3 and anti-spectrin antibodies are increased in *Plasmodium vivax* infection and are associated with anemia. *Sci. Rep.* 8, 8762.
- Mourão, L.C., Cardoso-Oliveira, G.P., Braga, E.M., 2020. Autoantibodies and malaria: where we stand? Insights into pathogenesis and protection. *Front. Cell. Infect. Microbiol.* 10, 262.
- Nakanishi, K., Inoko, H., 2006. Combination of HLA-A24, -DQA1*03, and -DR9 contributes to acute-onset and early complete beta-cell destruction in type 1 diabetes: longitudinal study of residual beta-cell function. *Diabetes* 55, 1862–1868.
- NCBI Resource Coordinators, 2017. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 46, D8–D13.
- Negi, S., Singh, H., Mukhopadhyay, A., 2017. Gut bacterial peptides with autoimmunity potential as environmental trigger for late onset complex diseases: in-silico study. *PLoS One* 12, e0180518.
- Ng, M.H., Lau, K.M., Li, L., Cheng, S.H., Chan, W.Y., Hui, K., Zee, B., Leung, C.B., Sung, J. J., 2004. Association of human-leukocyte-antigen class I (B*0703, and class II (DRB1*0301, genotypes with susceptibility and resistance to the development of severe acute respiratory syndrome. *J. Infect. Dis.* 190, 515–518.
- Nguyen, A., David, J.K., Maden, S.K., Wood, M.A., Weeder, B.R., Nellore, A., Thompson, R.F., 2020. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J. Virol.* 94, e00510–e00520.
- Worning Nielsen, M., Lundegaard, C., Lauemoller, S.L., Lamberth, K., Buus, S., Brunak, S., Lund, O., 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 12, 1007–1017.
- Noble, J.A., Valdes, A.M., Bugawan, T.L., Apple, R.J., Thomson, G., Erlich, H.A., 2002. The HLA class I A locus affects susceptibility to type 1 diabetes. *Human Immunol* 63, 657–664.
- Jennum Ollila, H.M., Ravel, J.M., Han, F., Faraco, J., Lin, L., Zheng, X., Plazzi, G., Dauvilliers, Y., Pizza, F., Hong, S.C., Knudsen, S., Kornum, B.R., Dong, X.S., Yan, H., Hong, H., Coquillard, C., Mahlios, J., Jolanki, O., Einen, M., et al., 2015. HLA-DPBI and HLA class I confer risk of and protection from narcolepsy. *Am. J. Hum. Genet.* 96, 136–146.
- Triggianese Perricone, C., Bartoloni, E., Cafaro, G., Bonifacio, A.F., Bursi, R., Perricone, R., Gerli, R., 2020. The anti-viral facet of anti-rheumatic drugs: lessons from COVID-19. *J. Autoimmun.* 111, 102468.
- Rapin, N., Hoof, I., Lund, O., Nielsen, M., 2008. MHCmotif viewer. *Immunogenetics* 60, 759–765.
- Reynisson, B., Alvarez, B., Paul, S., Peters, B., Nielsen, M., 2020. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* 48, W449–W454.
- Rivera-Correa, J., Rodríguez, A., 2020. Autoimmune anemia in malaria. *Trends Parasitol.* 36, 91–97.
- Flicek Robinson, J., Barker, D.J., Georgiou, X., Cooper, M.A., Marsh, S.G., 2020. IPD-IMGT/HLA database. *Nucleic Acids Res* 48, D948–D955.
- Rodríguez, Y., Novelli, L., Rojas, M., De Santis, M., Acosta-Ampudia, Y., Monsalve, D.M., Ramírez-Santana, C., Costanzo, A., Ridgway, W.M., Ansari, A.A., Gershwin, M.E., Selmi, C., Anaya, J.M., 2020. Autoinflammatory and autoimmune conditions at the crossroad of COVID-19. *J. Autoimmun.* 114, 102506.
- Restrepo-Jiménez Rojas, M., Monsalve, D.M., Pacheco, Y., Acosta-Ampudia, Y., Ramírez-Santana, C., Leung, S., Ansari, A.A., Gershwin, M.E., Anaya, J.M., 2018. Molecular mimicry and autoimmunity. *J. Autoimmun.* 95, 100–123.
- Sedaghat, Z., Karimi, N., 2020. Guillain Barre syndrome associated with COVID-19 infection: a case report. *J. Clin. Neurosci.* 76, 233–235.
- Shi, Y., Kawana-Tachikawa, A., Gao, F., Qi, J., Liu, C., Gao, J., Cheng, H., Ueno, T., Iwamoto, A., Gao, G.F., 2017. Conserved V delta 1 binding geometry in a setting of locus-disparate pHLA recognition by delta /alpha beta T cell receptors (TCRs): insight into recognition of HIV peptides by TCRs. *J. Virol.* 91, e00717–e00725.
- Shimizu, A., Kawana-Tachikawa, A., Yamagata, A., Han, C., Zhu, D., Sato, Y., Nakamura, H., Koibuchi, T., Carlson, J., Martin, E., Brumme, C.J., Shi, Y., Gao, G.F., Brumme, Z.L., Fukai, S., Iwamoto, A., 2013. Structure of TCR and antigen complexes at an immunodominant CTL epitope in HIV-1 infection. *Sci. Rep.* 3, 3097.
- Stranzl, T., Larsen, M.V., Lundegaard, C., Nielsen, M., 2020. NetCTLpan. Pan-specific MHC class I pathway epitope predictions. *Immunogenetics* 62, 357–368.
- The UniProt Consortium, 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47, D506–D515.
- Thomsen, M.F., Nielsen, M., 2012. Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Res* (40), W281–W287.
- Touret, F., De Lamballerie, X., 2020. Of chloroquine and COVID-19. *Antiviral Res* 177, 104762.
- Trost, B., Lucchese, G., Stufano, A., Bickis, M., Kusalik, A., Kanduc, D., 2010. No human protein is exempt from bacterial motifs, not even one. *Self Nonself* 1, 328–334.
- Veljkovic, V., Johnson, E., Metlas, R., 1997. Molecular basis of the inefficacy and possible harmful effects of AIDS vaccine candidates based on HIV-1 envelope glycoprotein gp120. *Vaccine* 15, 473–474.
- Vellozzi, C., Iqbal, S., Broder, K., 2014. Guillain-Barre syndrome, influenza, and influenza vaccination: the epidemiologic evidence. *Clin. Infect. Dis.* 58, 1149–1155.
- Vojdani, A., Kharratian, D., 2020. Potential antigenic cross-reactivity between SARS-CoV-2 and human tissue with a possible link to an increase in autoimmune diseases. *Clin. Immunol.* 217, 108480.
- Warren, R.L., Birol, I., 2020. In: *HLA Predictions from the bronchoalveolar lavage fluid samples of five patients at the early stage of the Wuhan seafood market COVID-19 outbreak* [Preprint] ArXiv. 2004.07108v3.
- Wheeler, D., Bhagwat, M., 2007. BLAST QuickStart. In: Bergman, N.H. (Ed.), *Comparative Genomics*. Humana Press, Totowa (NJ).
- Woodruff, M.C., et al., 2020. In: *Dominant extrafollicular B cell responses in severe COVID-19 disease correlate with robust viral-specific antibody production but poor clinical outcomes*. <https://doi.org/10.1101/2020.04.29.20083717> [Preprint] MedRxiv.
- Yuki, N., 2007. Ganglioside mimicry and peripheral nerve disease. *Muscle Nerve* 35, 691–711.
- Zabriskie, J.B., Freimer, E.H., 1966. An immunological relationship between the group A streptococcus and mammalian muscle. *J. Exp. Med.* 124, 661–678.
- Zhang, H., Lund, O., Nielsen, M., 2009. The PickPocket method for predicting binding specificities for receptors based on receptor pocket similarities: application to MHC-peptide binding. *Bioinformatics* 25, 1293–1299.