



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.



## Possible role of HLA class-I genotype in SARS-CoV-2 infection and progression: A pilot study in a cohort of Covid-19 Spanish patients

Ignacio Iturrieta-Zuazo<sup>a,1</sup>, Claudia Geraldine Rita<sup>a,1</sup>, Ana García-Soidán<sup>a</sup>, Ana de Malet Pintos-Fonseca<sup>b</sup>, Nieves Alonso-Alarcón<sup>a</sup>, Roberto Pariente-Rodríguez<sup>a</sup>, Amalia Tejeda-Velarde<sup>a</sup>, Sergio Serrano-Villar<sup>c</sup>, José Luis Castañer-Alabau<sup>a</sup>, Israel Nieto-Gañán<sup>a,\*,1</sup>

<sup>a</sup> Immunology Department. Hospital Universitario Ramón y Cajal. Madrid. Spain

<sup>b</sup> Microbiology Department. Hospital Universitario Marqués de Valdecilla. Santander. Spain

<sup>c</sup> Infectious Disease Department. Hospital Universitario Ramón y Cajal. Madrid. Spain

### ARTICLE INFO

#### Keywords:

SARS-CoV-2  
HLA class I  
Evolution disease  
Viral peptide-MHC class I binding affinity

### ABSTRACT

Human Leukocyte Antigen (HLA) includes a large set of genes with important actions in immune response against viral infection. Numerous studies have revealed the existence of significant associations between certain HLA alleles and the susceptibility and prognosis of different infectious diseases. In this pilot study we analyse the binding affinity between 66 class I HLA alleles and SARS-CoV-2 viral peptides, and its association with the severity of the disease.

A total of 45 Spanish patients with mild, moderate and severe SARS-CoV-2 infection were typed for HLA class I; after that, we analysed if an in silico model of HLA I-viral peptide binding affinity and classical HLA supertypes could be correlated to the severity of the disease. Our results suggest that patients with mild disease present Class I HLA molecules with a higher theoretical capacity for binding SARS-CoV-2 peptides and showed greater heterozygosity when comparing them with moderate and severe groups. In this regard, identifying HLA-SARS-CoV-2 peptides binding differences between individuals would help to clarify the heterogeneity of clinical responses to the disease and will also be useful to guide a personalized treatment according to its particular risk.

### 1. Introduction

At the end of 2019, a novel coronavirus infection was reported in Wuhan, China, being identified as SARS-CoV-2 [1]. Subsequently, the virus has spread rapidly globally, practically affecting all countries in the world with serious health and economic consequences. Approximately 5.5 million infections and 585.000 deaths have been reported worldwide in July 2020. Updated information can be consulted in <https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6> (Coronavirus COVID-19 global cases by the centre for systems science and engineering at Johns Hopkins University) [2].

Over the last twenty years, three novel Betacoronaviruses, Severe Acute Respiratory Syndrome (SARS-CoV), Middle East Respiratory Syndrome (MERS) and SARS-CoV2, have caused a significant high case-fatality rates in humans [3]. Nowadays, researchers are putting all their

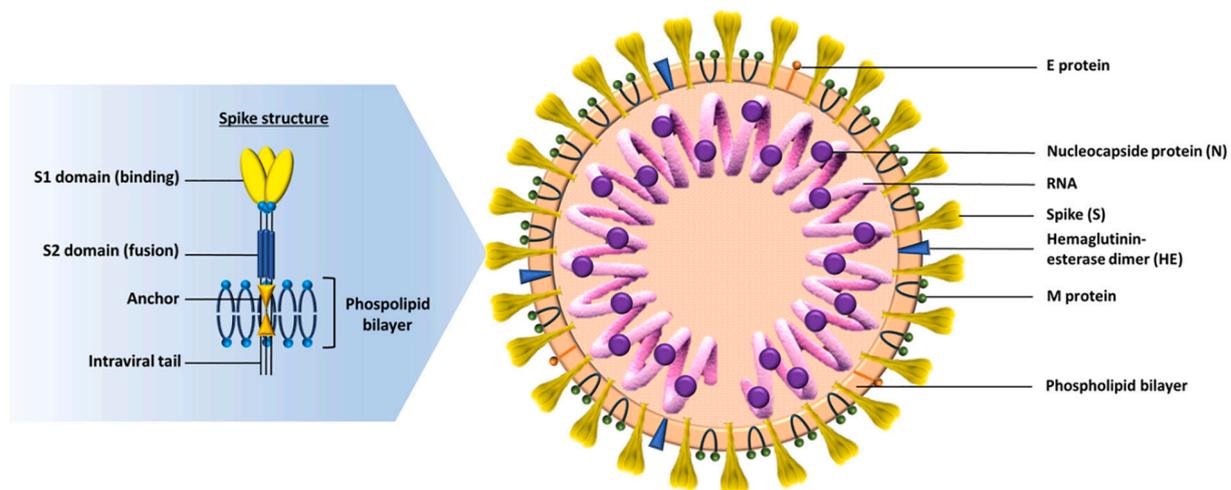
efforts into understanding more about transmissibility, severity and other features associated with novel SARS-CoV-2 infection [4]. Although age and the existence of comorbidities are already known to be risk factors for developing a severe form of the disease, the importance of genetic differences between individuals in the heterogeneity of observed clinical manifestations remains unclear [5].

Some authors suggest that disease severity in patients is due not only to the viral infection but also to the host response [6]. Based on the immunological response to SARS-CoV-2 infection, it has been reported that innate immunity gives the first try to contain the virus through its pathogen-associated molecular patterns (PAMPs) receptors and it is followed by an adaptive immune response that helps to decrease the viral load through the T cell response and the release of antibodies [7]. Therefore, one of the questions that has not been fully answered yet is why in some cases a dysfunctional immune response occurs leading to a severe course of disease [6].

\* Corresponding author.

E-mail address: [israel.ganan@salud.madrid.org](mailto:israel.ganan@salud.madrid.org) (I. Nieto-Gañán).

<sup>1</sup> These three authors contributed equally.



**Fig. 1.** On the right the SARS-CoV-2 structure particle: The virion has a nucleocapsid composed of genomic RNA and phosphorylated nucleocapsid (N) protein, which is buried inside phospholipid bilayers and covered by the spike glycoprotein trimer (S). The membrane (M) protein (a type III transmembrane glycoprotein) and the envelope (E) protein are located among the S proteins in the virus envelope. On the left a detailed structure of the spike: it consists of 3 domains, the ectodomain located in the virus crown and divided into the S1 receptor (binding subunit) and the S2 receptor (fusion subunit), the anchor and the intraviral tail (self-made figure).

In this context, special importance should be given to the human leukocyte antigen (HLA) system, that governs our adaptive immunity, by presenting pathogen-derived peptides to T cells [8]. HLA alleles have been shown to influence susceptibility or resistance to several infectious diseases, such as malaria, tuberculosis, leprosy, HIV and hepatitis virus persistence [9]. Some studies suggest that certain HLA alleles are associated not only with susceptibility or resistance to previous coronaviruses but also with an increased severity of the disease [8,10–12]. Nguyen et al. [13] performed an *in silico* analysis of viral peptide-MHC class I binding affinity between 145 HLA-I alleles and all possible 8- to 12-mers from the SARS-CoV-2 proteome using two degrees of affinity: loosely (threshold < 500 nM) and tightly (threshold < 50 nM). These peptides belong to various structures of the virus, which consists of genomic RNA and a phosphorylated nucleocapsid protein (N) inside of a phospholipid bilayer, covered by the spike glycoprotein trimer (S). Some coronaviruses have an additional protein spike called hemagglutinin-esterase (HE). The membrane (M) protein (a type III transmembrane glycoprotein) and the envelope (E) protein are located among the S proteins in the virus envelope (Fig. 1) [14,15]. Protein S is the main antigenic component against which the host's immune response occurs. It is composed of 3 domains (Fig. 1): the ectodomain located in the virus crown and divided into the S1 receptor (binding subunit) and the S2 receptor (fusion subunit), the anchor and the intraviral tail [14].

Nguyen et al. found that HLA-A\*02:02, HLA-B\*15:03, and HLA-C\*12:03 were presenters of a higher number of peptides while A\*25:01, B\*46:01, C\*01:02 were the worst presenting SARS-CoV-2 peptides [13].

This pilot study was conducted to analyse the association between HLA in SARS-CoV-2 infection and the severity of the disease. Following the protein-presenting alleles of the virus model proposed by Nguyen et al. [13], we wanted to check if having a HLA Class I genotype that recognizes a greater number of SARS-CoV-2 peptides can be associated with a better evolution of the disease.

## 2. Material and methods

### 2.1. Patients

A total of 45 confirmed cases of COVID-19 diagnosed at Ramon and Cajal University Hospital between March and May 2020 were included in this pilot study. All patients had at least one positive test result for

SARS-CoV-2 by RT-PCR from nasopharyngeal swabs or bronchoalveolar lavage. The sample was divided into three main groups according to the severity of the disease: mild group ( $n = 5$ ), including outpatients whom presented with mild clinical features or even asymptomatic, moderate group ( $n = 20$ ), which included hospitalized patients but not requiring admission at ICU and the severe group ( $n = 20$ ), that required ICU admission and supportive care. Patients with pathologies that lead to greater morbidity or who had additional immunosuppression were excluded (HIV, active cancer in treatment with chemotherapy, immunodeficiency, autoimmune diseases with immunosuppressants, transplant patients, chronic kidney disease on dialysis, etc.). Patients dismissed for ICU due to their age or medical history were also excluded from the study. All individuals included in the study belong to the northeast area of Madrid and are native of Spain. Epidemiologic, demographic and outcome data were collected and listed in Table 1.

### 2.2. HLA typing

HLA Class I typing (locus A, B and C) of all patients was performed using the Sequence Specific Oligonucleotide reverse (SSO) technique through Luminex xMAP® technology. We used the kits RSSOW1A, RSSOW1B and RSSOW1C (One lambda inc) were used, which allows resolution at the level of Common and Well-Documented (CWD) alleles in a Flexmap 3D Luminex system.

### 2.3. Viral peptide-MHC class I binding affinity data

The Nguyen et al. [13] study data are available in the web <https://github.com/pdxgx/covid19>. To calculate the number of SARS-CoV-2 peptides bound per patient we list all the peptides binding by each HLA allele by locus and class I haplotype of each patient and excluded the repeated peptides (recognized by more than one allele), then we count the total number of different peptides recognized tightly and loosely. Taking this into account, in this pilot study we wanted to check whether having a HLA Class I genotype that recognizes a greater number of SARS-CoV-2 peptides can be associated with a better evolution of the disease.

### 2.4. HLA supertype data (Fig. 2)

To group the alleles analysed by Nguyen et al. into superotypes, we used various references available in this regard for loci A and B

**Table 1**  
Demographic and clinical data of mild, moderate and severe patients.

	Mild (n = 5)	Moderate (n = 20)	Severe (n = 20)
Age, y (mean ± SD)	48 ± 7,02	66 ± 16,04	65 ± 9,20
<b>Gender</b>			
Female	2 (40%)	10 (50%)	6 (30%)
Male	3 (60%)	10 (50%)	14 (70%)
<b>Comorbidities</b>			
None	4 (80%)	5 (25%)	2 (10%)
1	–	2 (10%)	3 (15%)
> 1	1 (20%)	13 (65%)	18 (75%)
<b>Cardiovascular disease</b>			
Hypertension	1 (20%)	12 (60%)	12 (60%)
Coronary artery disease	–	1 (5%)	2 (10%)
Stroke	–	1 (5%)	1 (5%)
<b>Metabolic disease</b>			
Obesity	1 (20%)	2 (10%)	3 (15%)
Dyslipidemia	1 (20%)	4 (20%)	7 (35%)
Hyperuricemia	–	–	2 (10%)
Diabetes	–	4 (20%)	5 (25%)
<b>Chronic respiratory disease</b>			
Asthma	–	–	1 (5%)
Chronic obstructive pulmonary disease	–	1 (5%)	–
Obstructive sleep apnea	–	1 (5%)	2 (10%)
<b>Neoplasm<sup>a</sup></b>	–	2 (10%)	3 (15%)
<b>Others</b>	–	2 (8%) <sup>b</sup>	1 (4%) <sup>c</sup>
<b>Outcome</b>			
Death	0	0	15 (75%)
Discharged alive	0	20 (100%)	3 (12%)
Still in hospital	0	0	2 (10%)
Outpatient	5 (100%)	0	0

<sup>a</sup> without active treatment, cured, remission; <sup>b</sup>SLE, thalassemia minor; <sup>c</sup>hemochromatosis.

[15–18]. Since no supertypes have been described for the C locus, we used instead the classic C1 and C2 groups established as ligands for the KIR receptors of NK cells based on the presence of Ser77 and Asn80 (group C1) or Asn77 and Lys80 (group C2) [19]. Only alleles analysed by Nguyen et al. [13] are showed in Fig. 2. To assess the clinical relevance of HLA supertypes, we firstly analysed theoretically whether the alleles with the highest binding to virus peptides clustered in some of the supertypes; and then we apply the analysis of supertypes in our patients and count the number of alleles of each supertype found in all patients of each group (mild, moderate or severe).

### 2.5. Frequency HLA alleles

Allelic and haplotypic frequencies in Spain were obtained from <http://www.allelefreqencies.net> [20]. In this web, frequencies are aggregated by country but reflect the distribution of HLA data available in this Net database. In order to calculate the Spanish allelic frequency, we used the average of the frequencies of each allele in all Spanish studies uploaded on this website.

### 2.6. Statistical analysis

Data were analysed using GraphPad Prism v8 for Windows software. Shapiro Wilks test confirmed the normal distribution of data. For this reason, the Student test for two group was performed to compare allele frequency between groups. To compare the different HLA supertypes with peptides binding affinity for loci A and B, an ANOVA for multiple comparisons (Tukey's test) was performed while for locus C we used a Student's *t*-test. To assess the association of HLA alleles with the severity we compared all three Class I locus alleles separately and the HLA-I genotype with the number of proteins recognized in the three groups of patients using the ANOVA analysis for multiple comparisons.

## 3. Results

Forty-five patients with confirmed SARS-Cov-2 infection were included in this pilot study. Although those patients with severe comorbidities were excluded from the study, 82% of them had at least one underlying disease. In the mild group only one patient had history of previous pathologies. The median age of the study population was 60 years-old and over 60% were male. More than 75% of the severe patients died during admission, while none of the patients of the mild and moderate group deceased. Table 1 includes demographic and clinical data of each group of severity.

The mild group showed a significantly statistical affinity for a greater number of viral peptides when comparing it to moderate and severe groups by locus and by complete HLA Class I genotype (Tukey's test,  $p < 0.0001$ ). However, no significant difference was obtained when analysing the locus and genotype HLA-number of recognized proteins in severe patients versus those with moderate evolution (Fig. 3, Table 2).

Table 3 contains the summary of the alleles found in our study. Fig. 4 shows each allele with the number of binding SARS-CoV-2 peptides according to loose (< 500 nM) or tight affinity (< 50 nM).

The degree of HLA homozygosity is associated with poorer disease resistance [21]. In our study, it was found that the percentage of homozygous individuals for loci A and C was higher in the group of patients with severe evolution compared to those with moderate evolution (20% vs. 10% in locus A, 15% vs. 5% in locus C). However, no statistically significant differences were obtained. For locus B the degree of homozygosity was 5% in both groups. It should be noticed that although we only had 5 mild patients, none of them was homozygous for any of the loci.

Analysis between HLA supertypes and SARS-CoV-2 peptide affinity (Table 4) indicates that A2 supertypes theoretically recognize significantly more peptides than A1 and A3 supertypes ( $p < 0.0001$ ), for both loose and tight affinity. Also, C1 supertype shows a theoretical higher viral peptide recognition for loose ( $p = 0.031$ ) and tight ( $p = 0.016$ ) binding affinity. No differences were found for B supertypes.

However, when trying to apply this hypothesis to our cohort of patients, although distribution of supertypes according to the HLA alleles found in the moderate and severe evolution groups showed correlation it did not reach statistical significance ( $t = 0.1756$ ;  $p = 0.86$ ). This could be because of the sample size. Mild group was not included in this analysis for its small size. Results appear collected in Table 4.

## 4. Discussion

Numerous studies have revealed the existence of significant associations between certain HLA alleles and the susceptibility and prognosis of different infectious diseases [22,23]. HLA includes a large set of genes with important actions in regulation and activity of the immune system against viral infection. In this regard, the association between SARS-Cov-2 infection and the HLA genotype has not been well established, with a few studies in this context focused on finding specific alleles of susceptibility.

To the best of our knowledge, this is one of the first works that study the role played by HLA molecules on individual responses to SARS-CoV-2 infection in a cohort of patients from Madrid, the Spanish region with the highest incidence of confirmed cases of COVID-19. Our pilot study examined the viral peptide-HLA binding predictions of an in-silico model in a clinical context. The bioinformatic approach of Nguyen et al. [13] focused on making a prediction of the number of total virus-derived peptides that a series of HLA class I allelic variants were capable of presenting, that would act as CD8+ T cell epitopes, under the premise that the more peptides they can present, the better the immune response is. Although these theoretical bioinformatic peptide-binding predictions would require their experimental validation and

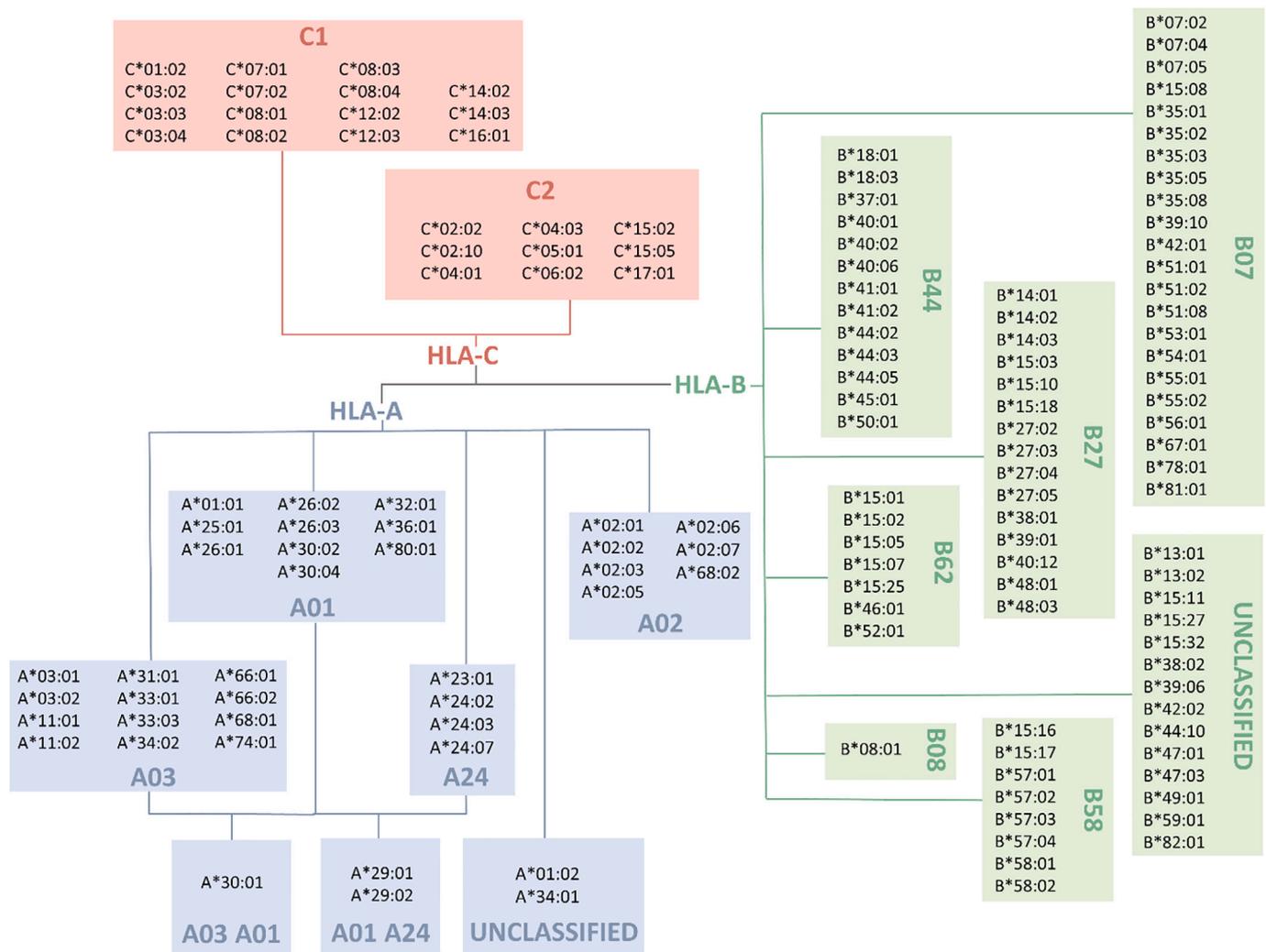


Fig. 2. HLA supertypes according to Sidney et al. [16]. HLA-C alleles are classified by C1 (Ser77 and Asn80) and C2 (Asn77 and Lys80) motifs (self-made figure).

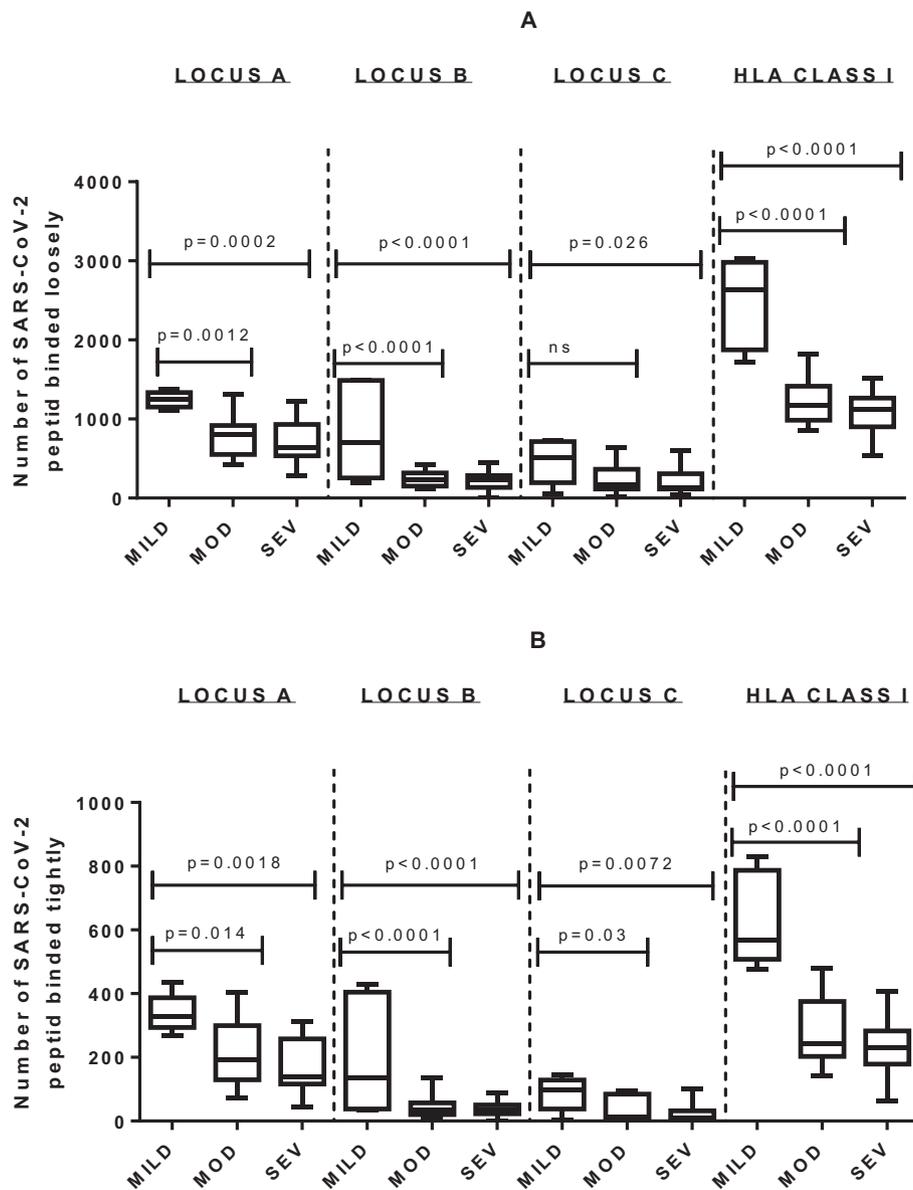
demonstration in the laboratory, we wanted to make an approximation of their representability in the clinical context of our study cohort.

Our work indicates that the mild group presents HLA molecules with a theoretical capacity for binding of SARS-Cov-2 peptides higher than the other groups. In addition, they present a greater percentage of heterozygosity. Although these findings are consistent with the bioinformatic predictions and previous works [13,21,24], these data must be interpreted with caution mainly due to the limited sample size. In this sense, global efforts are currently being made to collect samples from different types of patients, including asymptomatic individuals, in order to carry out studies with greater statistical power. For instance, our hospital has begun to carry out seroprevalence studies in addition to implementing strategies for donating hyperimmune serum from asymptomatic seropositive individuals or recovered patients, which will allow for a larger sample.

Although this is a pilot study with a few patients, based on the results obtained, it could be deduced that although the degree of affinity of an individual's HLA molecules for different SARS-CoV-2 structures influences the evolution of the disease, it is evident that many other genetic and epigenetic factors independent of the HLA participate in the immunopathogenesis of the virus. For example, the particular immunogenicity of each of the virus-derived peptides should be taken into account. Although in this regard it has been found that most of the viral peptides presented in the context of HLA class I could be immunogenic [25], not all are expressed in the same concentration or with the same cell dynamics. This is highlighted when finding alleles with a high

affinity for the virus in patients with a serious evolution of the disease without associated comorbidity that could justify this evolution. Therefore, we believe that genetic studies related to different structures of innate immunity such as PAMP receptors and everything related to the inflammasome should be undertaken since one of the main problems that has been observed in SARS-CoV-2 patients has been an exaggerated response of the innate immunity that has its maximum exponent in the cytokine release syndrome [24]. In addition, if possible, it would be advisable to carry out the same approach for HLA Class II molecules, since their presence on the surface of antigen-presenting cells play a fundamental role in the synthesis of antibodies to develop an effective humoral adaptive immunity. In this regard, Ellinghaus et al [26] employed a peptide binding prediction model for HLA-Class II in which they found no significant difference between the control and the patient groups. Unfortunately, bioinformatic models are having difficulty designing such models for HLA Class II [13]. In our case, Nguyen et al. only designed a model for class I, which is why we did not analyse the clinical relevance of Class II.

Regarding the utility of the classification by supertypes of the different HLA alleles, although they are the first simple bioinformatics models that grouped HLA molecules based on the structure of the peptide-binding cleft [16], the development of more complex in silico models, which allows establishing allelic differences in binding against various peptides of the same microorganism, such as the case at hand, makes the latter more useful for searching for possible allelic and genotypic associations of severity and protection against disease. Our



**Fig. 3.** Comparison of severe and moderate groups with a third small cohort of 5 mild patients. Significant *p*-value of Tukey Test for multiple comparison is showed. A) Results for loosely binding affinity analysis. B) Results for tightly binding affinity analysis. MOD: Group of patients with moderate evolution; MILD: Group of patients with mild evolution; SEV: Group of patients with severe evolution.

**Table 2**

Summary of binding affinity analysis for A, B, C locus and HLA Class I genotype between the three groups through an ANOVA for multiple comparison test.

Binding affinity analysis	Affinity	Mild group (mean /SD)	Moderate group (mean /SD)	Severe group (mean /SD)	f-value (ANOVA)	p-value	p-value severe vs moderate	p-value severe vs mild	p-value moderate vs mild
LOCUS A	< 500 nm	1245 ± 100.7	776.9 ± 243	700.4 ± 264.7	10.10	<b>0.0003</b>	0.58	<b>0.0002</b>	<b>0.0012</b>
	< 50 nm	338 ± 60.69	208.7 ± 93.45	175.5 ± 87.16	6.83	<b>0.0027</b>	0.4636	<b>0.0018</b>	<b>0.0144</b>
LOCUS B	< 500 nm	836.8 ± 622.3	238.6 ± 91.51	208.3 ± 112.4	18.12	< <b>0.0001</b>	0.89	< <b>0.0001</b>	< <b>0.0001</b>
	< 50 nm	204 ± 188.1	44.55 ± 34.29	33.55 ± 22.52	14.79	< <b>0.0001</b>	0.85	< <b>0.0001</b>	< <b>0.0001</b>
LOCUS C	< 500 nm	467.6 ± 279.3	240.8 ± 189.7	208.2 ± 169.8	3.719	<b>0.033</b>	0.85	<b>0.026</b>	0.057
	< 50 nm	86.2 ± 53.57	36.75 ± 37.91	26.2 ± 32.6	5.136	<b>0.01</b>	0.65	<b>0.0072</b>	<b>0.03</b>
HLA Class I (ABC)	< 500 nm	2469 ± 573	1212 ± 261.9	1082 ± 248	44.05	< <b>0.0001</b>	0.37	< <b>0.0001</b>	< <b>0.0001</b>
	< 50 nm	631.6 ± 148.7	285 ± 103.6	233.3 ± 87.44	30.86	< <b>0.0001</b>	0.26	< <b>0.0001</b>	< <b>0.0001</b>

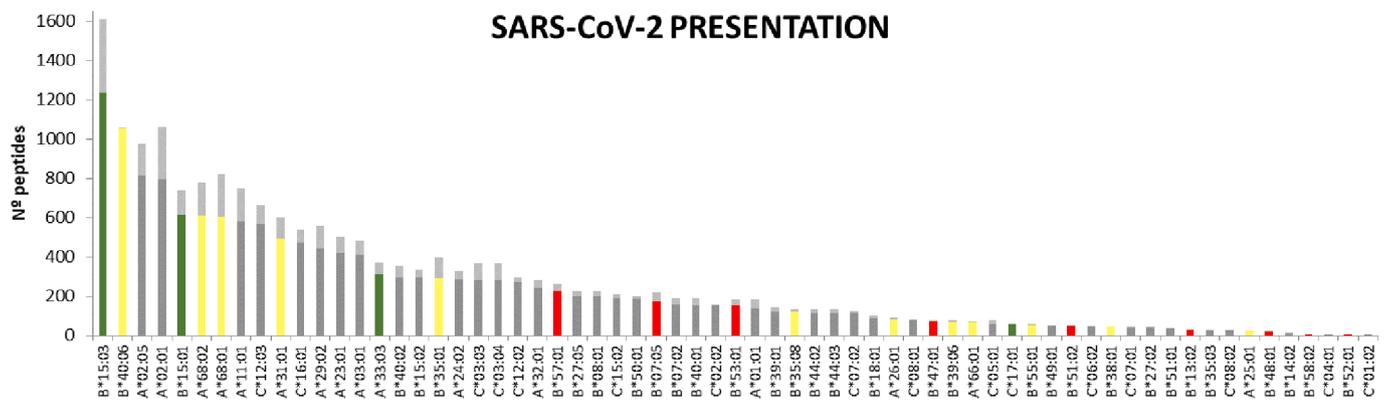
In bold letter significant *p*-values. Calculations are based on the number of binding peptides loosely and tightly by the patients included in each of the groups.

**Table 3**

Summary of the alleles found in the studied patients indicating their frequency in Spain and in each group of patients.

Allele	Hla supertype	Freq Spain	Presence in severe groupP (n = 40 alleles; 20 patients)	Freq. in severe group	Presence in moderate group (n = 40 alleles; 20 patients)	Freq. in moderate group	Presence in mild group (n = 10 alleles; 5 patients)	Freq. in mild group	N° loosely binding peptides (< 500 nm)	N° tightly binding peptides (< 50 nm)
A*01:01	A01	0,1	10	0,25	4	0,1	0	0	139	44
A*02:01	A02	0,25	5	0,125	6	0,15	4	0,4	795	267
A*02:05	A02	0,01	2	0,05	0	0	2	0,2	817	158
A*03:01	A03	0,06	6	0,15	6	0,15	1	0,1	409	73
A*11:01	A03	0,05	2	0,05	6	0,15	1	0,1	582	168
A*23:01	A24	0,03	1	0,025	1	0,025	0	0	418	82
A*24:02	A24	0,07	6	0,15	3	0,075	0	0	286	43
A*25:01	A01	0,01	0	0	2	0,05	0	0	19	1
A*26:01	A01	0,03	0	0	3	0,075	0	0	84	10
A*29:02	A01 A24	0,05	2	0,05	1	0,025	1	0,1	445	112
A*31:01	A03	0,03	0	0	2	0,05	0	0	494	107
A*32:01	A01	0,03	2	0,05	1	0,025	0	0	242	42
A*33:03	A03	0,01	0	0	0	0	1	0,1	312	61
A*66:01	A03	0,02	0	0	1	0,025	0	0	69	2
A*68:01	A03	0,02	0	0	1	0,025	0	0	604	217
A*68:02	A02	0,01	0	0	1	0,025	0	0	612	164
B*07:02	B07	0,05	4	0,1	4	0,1	0	0	158	36
B*07:05	B07	0,01	1	0,025	0	0	0	0	175	46
B*08:01	B08	0,06	5	0,125	3	0,075	0	0	200	25
B*13:02	B62	0,01	1	0,025	0	0	0	0	28	0
B*14:02	B27	0,04	1	0,025	1	0,025	0	0	13	1
B*15:01	B62	0,02	0	0	0	0	2	0,2	616	123
B*15:02	B62	0,01	0	0	2	0,05	0	0	295	40
B*15:03	B27	0,01	0	0	0	0	2	0,2	1236	373
B*18:01	B44	0,06	2	0,05	2	0,05	1	0,1	89	12
B*27:02	B27	0,01	1	0,025	1	0,025	0	0	38	1
B*27:05	B27	0,02	1	0,025	2	0,05	1	0,1	203	22
B*35:01	B07	0,04	0	0	3	0,075	0	0	294	104
B*35:03	B07	0,03	2	0,05	2	0,05	0	0	26	2
B*35:08	B07	0,01	0	0	1	0,025	0	0	119	14
B*38:01	B27	0,02	0	0	2	0,05	0	0	41	1
B*39:01	B27	0,01	1	0,025	2	0,05	0	0	122	23
B*39:06	B27	0,01	2	0,05	0	0	0	0	70	6
B*40:01	B44	0,04	1	0,025	1	0,025	0	0	156	38
B*40:02	B44	0,03	1	0,025	0	0	1	0,1	297	57
B*40:06	B44	0,02	0	0	1	0,025	0	0	1054	6
B*44:02	B44	0,04	2	0,05	2	0,05	1	0,1	116	19
B*44:03	B44	0,06	2	0,05	1	0,025	1	0,1	115	22
B*47:01	B44	0	1	0,025	0	0	0	0	74	1
B*48:01	B27	0	1	0,025	0	0	0	0	18	1
B*49:01	XXX	0,04	1	0,025	2	0,05	0	0	49	4
B*50:01	B44	0,02	1	0,025	2	0,05	1	0,1	188	13
B*51:01	B07	0,06	1	0,025	3	0,075	0	0	33	1
B*51:02	B07	0	1	0,025	0	0	0	0	48	2
B*52:01	B62	0,03	2	0,05	0	0	0	0	4	0
B*53:01	B07	0,01	1	0,025	0	0	0	0	152	32
B*55:01	B07	0,02	0	0	2	0,05	0	0	53	7
B*57:01	B58	0,03	2	0,05	0	0	0	0	225	38
B*58:02	B58	0,01	1	0,025	0	0	0	0	7	0
C*01:02	C1	0,03	1	0,025	4	0,1	1	0,1	4	0
C*02:02	C2	0,04	2	0,05	2	0,05	1	0,1	153	2
C*03:03	C1	0,04	0	0	1	0,025	1	0,1	284	86
C*03:04	C1	0,03	1	0,025	2	0,05	0	0	284	86
C*04:01	C2	0,1	2	0,05	5	0,125	0	0	6	0
C*05:01	C2	0,09	3	0,075	3	0,075	1	0,1	60	16
C*06:02	C2	0,07	3	0,075	1	0,025	1	0,1	44	6
C*07:01	C1	0,13	9	0,225	5	0,125	0	0	40	4
C*07:02	C1	0,06	8	0,2	8	0,2	0	0	115	8
C*08:01	C1	0	1	0,025	1	0,025	0	0	79	4
C*08:02	C1	0,07	1	0,025	1	0,025	0	0	24	2
C*12:02	C1	0,02	2	0,05	1	0,025	0	0	271	26
C*12:03	C1	0,07	1	0,025	2	0,05	1	0,1	569	94
C*15:02	C2	0,04	1	0,025	1	0,025	0	0	193	16
C*15:05	C2	0,01	0	0	0	0	0	0	118	5
C*16:01	C1	0,07	2	0,05	2	0,05	1	0,1	472	66
C*17:01	C2	0,02	0	0	0	0	1	0,1	55	3

Note that each patient has two alleles from each locus, so the total number of alleles in each group has been used for calculating frequencies and not the number of patients. Consequently, the groups with 20 patients include 40 alleles, while the group with 5 includes 10 alleles. For each one, the HLA supertype as well as the number of SARS-CoV-2 peptides recognized with loosely (< 500 nm) and tightly (< 50 nm) high and low affinity according to Nguyen et al. (11) is also indicated.



**Fig. 4.** Schematic of the SARS-CoV-2 binding capacity of different HLA alleles. Only those found in some of the patient groups studied are indicated. Alleles is shown as a series of vertical bars, with dark and light shading indicating the number of loosely (< 500 nM) and tightly (< 50 nM) binding peptides respectively. In green, yellow and red, those alleles that were only found in the group of mild patients without hospitalization, patients with moderate evolution and severe group respectively. In grey, those alleles found more than one group. Alleles are sorted in descending order based on the number of peptides they bind. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

results show that when performing the theoretical comparison, a higher degree of recognition of SARS-CoV-2 peptides is observed for the A2 and C1 supertypes. However, no correlation was deduced with the clinical evolution in our patient groups. Although it still has to be verified the usefulness of supertypes with a larger cohort of patients, the development of more complex bioinformatic models [13,24] would be more recommendable for this type of studies.

Concerning the different allele frequencies observed, it should be taken into account that they vary a lot between regions, so it could happen that alleles that have been identified as risk alleles in other association studies did not have weight in other populations, given their low representability on them [20]. An example is the HLA-B \* 46: 01 allele. It has been reported to have a low affinity for the virus [13] and to be statistically significantly associated with severe SARS-CoV-1 disease and COVID-19 progression in studies conducted in various regions of China [12,27]. In these places, its population frequency reaches 0.25 (20), while in other studies in populations where its frequency is lower, this allele has not been found relevant [28].

In relation to our work, the frequency of this allele in the Spanish population is almost nil [20]. In this regard, although the sample size of our study does not allow us to demonstrate it, the fact of having found theoretically protective alleles, such as HLA-B \* 15: 03 in patients with severe evolution who have died, and alleles with low affinity for the viruses such as HLA-A \* 25: 01 in patients with a moderate evolution of the disease, makes us think that rather than looking for protective or harmful alleles, it seems more appropriate to study the affinity of the complete HLA genotype for the virus.

And it is here where the bioinformatics models, such as those of Nguyen et al., play a fundamental role. It is in this context that our study establishes a clinical relationship between a greater number of

recognized peptides and a better clinical evolution when comparing the mild group, who did not require hospital admission, to the other two. Even though there are no statistically significant differences between the groups of patients with severe and moderate evolution, the truth is that when observing the raw data there is a slightly higher number of proteins recognized by the HLA genotype of patients with moderate evolution. Although this is a pilot study and considering the possible limitations of the results interpretation, it can be a basis for conducting similar studies with a larger number of patients from different regions. It can be as well an opportunity to share our data with global databases (e.g. <http://www.hla-covid19.org/>) to perform a collaborative analysis that allows to get more conclusive results, avoiding the biases derived from the different allelic and haplotypic population frequencies, and taking advantage of the valuable bioinformatic models available. We plan to continue performing HLA typing of a greater number of SARS-CoV-2 positive patients and considering a series of clinical variables that will allow to establish more homogeneous groups, and to carry out multivariate analysis to try to elucidate more precisely the importance of the HLA in SARS-CoV-2 infection.

In conclusion, there is potential for further research in the role of HLA in COVID-19. It could be reasonable to consider HLA typing of every COVID-19 patient because it would help to clarify the heterogeneity of responses to the disease and it will also be of great value for the development of specific vaccines and personalized treatments.

**Funding**

“This research received no external funding”.

**Table 4**  
Summary of HLA supertype-SARS-CoV-2 binding affinity analysis.

HLA Supertype analysis	Affinity	F / t value	p value	Tukey's multiple comparisons	p value
HLA-A SUPERTYPES	< 500 nm	6.845	0.0002	A2 vs A1 supertypes	< 0.0001
	< 50 nm	5.57	0.0008	A2 vs A3 supertypes	0.0014
HLA-B SUPERTYPES	< 500 nm	1.775	0.13	A2 vs A1 supertypes	0.0002
	< 50 nm	1.302	0.27	A2 vs A3 supertypes	0.0029
HLA-C SUPERTYPES	< 500 nm	2.292	0.031		
	< 50 nm	2.593	0.016		

In those loci with more than two supertypes in which significant ANOVA values were obtained, a Tukey's multiple comparison was performed in pairs. Only those pairs of comparisons in which statistically significant p-values were obtained are shown. For locus C, a t student test was used to compare the two supertypes (C1 and C2).

## Ethical approval

The study was developed in accordance with the ethical standards of the Helsinki Declaration. The bioethics committee of the University Ramón y Cajal Hospital approved the study (approval number 225–20).

## Declaration of Competing Interest

“The authors declare no conflict of interest.”

## References

- [1] World Health Organization, Laboratory Testing of Human Suspected Cases of Novel Coronavirus (nCoV) Infection: Interim Guidance, 10 January 2020, (2020).
- [2] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, *Lancet Infect. Dis.* 20 (2020) 533–534.
- [3] S. Felsenstein, J.A. Herbert, P.S. McNamara, C.M. Hedrich, COVID-19: immunology and treatment options, *Clin. Immunol.* 108448 (2020).
- [4] S.P. Adhikari, S. Meng, Y. Wu, Y. Mao, R. Ye, Q. Wang, C. Sun, S. Sylvia, S. Rozelle, H. Raat, Epidemiology, causes, clinical manifestation and diagnosis, prevention and control of coronavirus disease (COVID-19) during the early outbreak period: a scoping review, *Infect. Dis. Poverty* 9 (2020) 1–12.
- [5] R.E. Jordan, P. Adab, K.K. Cheng, Covid-19: risk factors for severe disease and death, *BMJ* 368 (2020) m1198.
- [6] M.Z. Tay, C.M. Poh, L. Rénia, P.A. MacAry, L.F. Ng, The trinity of COVID-19: immunity, inflammation and intervention, *Nat. Rev. Immunol.* (2020) 1–12.
- [7] G. Li, Y. Fan, Y. Lai, T. Han, Z. Li, P. Zhou, P. Pan, W. Wang, D. Hu, X. Liu, Coronavirus infections and immune responses, *J. Med. Virol.* 92 (2020) 424–432.
- [8] A. Sanchez-Mazas, HLA Studies in the Context of Coronavirus Outbreaks, *Swiss Medical Weekly* 150, (2020).
- [9] J.M. Blackwell, S.E. Jamieson, D. Burgner, HLA and infectious diseases, *Clin. Microbiol. Rev.* 22 (2009) 370–85, Table of Contents.
- [10] S. Wang, K. Chen, M. Chen, W. Li, Y. Chen, C. Tsao, M. Yen, J.C. Huang, Y.A. Chen, Human-leukocyte antigen class I Cw 1502 and class II DR 0301 genotypes are associated with resistance to severe acute respiratory syndrome (SARS) infection, *Viral Immunol.* 24 (2011) 421–426.
- [11] A.H. Hajeer, H. Balkhy, S. Johani, M.Z. Yousef, Y. Arabi, Association of human leukocyte antigen class II alleles with severe Middle East respiratory syndrome-coronavirus infection, *Ann Thorac Med* 11 (2016) 211–213.
- [12] M. Lin, H. Tseng, J.A. Trejaut, H. Lee, J. Loo, C. Chu, P. Chen, Y. Su, K.H. Lim, Z. Tsai, Association of HLA class I with severe acute respiratory syndrome coronavirus infection, *BMC Med. Genet.* 4 (2003).
- [13] A. Nguyen, J.K. David, S.K. Maden, M.A. Wood, B.R. Weeder, A. Nellore, R.F. Thompson, Human leukocyte antigen susceptibility map for SARS-CoV-2, *J. Virol.* 94 (13) (2020) e00510–e00520, <https://doi.org/10.1128/JVI.00510-20>.
- [14] G. Li, Y. Fan, Y. Lai, T. Han, Z. Li, P. Zhou, P. Pan, W. Wang, D. Hu, X. Liu, Coronavirus infections and immune responses, *J. Med. Virol.* 92 (2020) 424–432.
- [15] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, *Lancet* 395 (2020) 565–574.
- [16] M. Wang, M.H. Claesson, Classification of human leukocyte antigen (HLA) super-types, *Immunoinformatics*, Springer, 2014, pp. 309–317.
- [17] J. Sidney, B. Peters, N. Frahm, C. Brander, A. Sette, HLA class I supertypes: a revised and updated classification, *BMC Immunol.* 9 (2008) 1.
- [18] O. Lund, M. Nielsen, C. Kesmir, A.G. Petersen, C. Lundegaard, P. Worning, C. Sylvester-Hvid, K. Lamberth, G. Røder, S. Justesen, Definition of supertypes for HLA molecules using clustering of specificity matrices, *Immunogenetics* 55 (2004) 797–810.
- [19] P. Parham, MHC class I molecules and KIRs in human history, health and survival, *Nat. Rev. Immunol.* 5 (2005) 201–214.
- [20] F.F. González-Galarza, L.Y. Takeshita, E.J. Santos, F. Kempson, M.H.T. Maia, Silva, Silva Andrea Luciana Soares da, André Luiz Teles e, G.S. Ghataoraya, A. Alfirevic, A.R. Jones, Allele frequency net 2015 Update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations, *Nucleic Acids Res.* 43 (2015) D784–D788.
- [21] J. Arora, F. Pierini, P.J. McLaren, M. Carrington, J. Fellay, T.L. Lenz, HLA heterozygote advantage against HIV-1 is driven by quantitative and qualitative differences in HLA allele-specific peptide presentation, *Mol. Biol. Evol.* 37 (2020) 639–650.
- [22] A. Sanchez-Mazas, A Review of HLA Allele and SNP Associations with Highly Prevalent Infectious Diseases in Human Populations, *Swiss Medical Weekly* 150, (2020).
- [23] H. Spinola, HLA loci and respiratory infectious diseases, *J. Respirat. Res.* 2 (2016) 56–66.
- [24] N.D. Ellinghaus, F. Degenhardt, L. Bujanda, M. Buti, A. Albillos, P. Invernizzi, J. Fernandez, D. Prati, et al., Genomewide association study of severe Covid-19 with respiratory failure, *N. Engl. J. Med.* (Jun 17 2020), <https://doi.org/10.1056/NEJMoa2020283>.
- [25] N.P. Croft, S.A. Smith, J. Pickering, J. Sidney, B. Peters, P. Faridi, M.J. Witney, P. Sebastian, I.E.A. Flesch, S.L. Heading, A. Sette, N.L. La Gruta, A.W. Purcell, D.C. Tschärke, Most viral peptides displayed by class I MHC on infected cells are immunogenic, *Proc. Natl. Acad. Sci. U. S. A.* 116 (2019) 3112–3117.
- [26] H.K. Siddiqi, M.R. Mehra, COVID-19 illness in native and immunosuppressed states: a clinical-therapeutic staging proposal, *J. Heart Lung Transplant* 39 (5) (2020) 405–407, <https://doi.org/10.1016/j.healun.2020.03.012>.
- [27] R.L. Warren, I. Birol, HLA predictions from the bronchoalveolar lavage fluid samples of five patients at the early stage of the Wuhan seafood market COVID-19 outbreak, arXiv preprint arXiv:2004.07108, (2020).
- [28] M.H. Ng, K. Lau, L. Li, S. Cheng, W.Y. Chan, P.K. Hui, B. Zee, C. Leung, J.J. Sung, 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 (2004) 515–518.