A BIO-INFORMATIC INVESTIGATION OF CATS' SUSCEPTIBILITY TO CORONAVIRUSES-DERIVING EPITOPES

M. Buonocore*,†, D. De Biase†, D. Sorrentino‡, A.M. D'Ursi†, S. Papparella*, A. Giordano§ and O. Paciello*

**Department of Veterinary Medicine and animal production, University of Naples "Federico II", Naples, IT, †Department of Pharmacy, University of Salerno, Fisciano, IT, ‡Department of Pharmacy, University of Salerno, IT and §Sbarro Institute for Cancer Research and Molecular Medicine, Temple University, Philadelphia, US*

INTRODUCTION

The mechanisms underlying the susceptibility to SARS-CoV-2 infection in humans and animals are still elusive, but a recent study demonstrated that different Class I/II human leukocyte antigen (HLA) alleles might define an individual susceptibility to SARS-CoV-2. Similarly, several studies suggested that the Feline Leukocyte Antigen (FLA) plays a pivotal role in the transmission of viruses to cats.

With this study, we explored a novel bio-informatic approach in order to predict which FLA-I allele might be correlated with transmissibility potential and enhanced immunogenic response in domestic cats infected with FCoVs and SARS-CoV-2.

MATERIALS AND METHODS

We performed epitope mapping of aminoacids deriving from SARS-CoV-2, FeCV and FIPV glycoproteins using the online server NetMHCPan 4.1 and predicted their affinities for different alleles of the three main loci in class I FLAs. The predicted complexes with the most promising affinities were then subjected to molecular docking and molecular dynamics simulations with HPepDock.

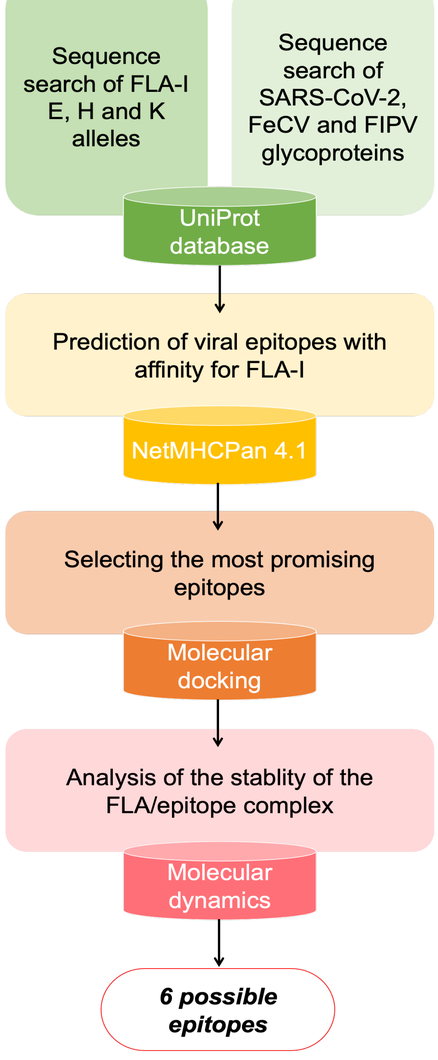


Figure 1. Workflow of investigation.

RESULTS

We performed an exploratory epitope mapping using the online server NetMHCPan 4.1. Among the entries in UniProt database marked as "reviewed", we selected **12 variants for FLA-I E, 11 for FLA-I H and 9 for FLA-I K**. Concerning the restrained peptide sequences, **we considered viral proteins from the three coronaviruses in exam: SARS-CoV-2, FIPV, and FeCV.** The amino acid sequences were retrieved from UniProt and only the entries marked as "reviewed" were selected. **Outcomes were filtered according to their Mass Spectrometry Eluted Ligands (EL) score.**

Epitope mapping

FLA-I E locus

Allele	Score	Total number of viral epitopes
FLA-I E*00501	EL > 0.5	89
	EL > 0.8	1
FLA-I E*00101	EL > 0.5	332
	EL > 0.8	41
FLA-I E*00701	EL > 0.5	93
	EL > 0.8	4
FLA-I E*01301	EL > 0.5	159
	EL > 0.8	1
FLA-I E*01401	EL > 0.5	31
	EL > 0.8	0
FLA-I E*01101	EL > 0.5	303
	EL > 0.8	11
FLA-I E*00303	EL > 0.5	31
	EL > 0.8	0
FLA-I E*01001	EL > 0.5	206
	EL > 0.8	3
FLA-I E*00302	EL > 0.5	31
	EL > 0.8	0
FLA-I E*00902	EL > 0.5	129
	EL > 0.8	3
FLA-I E*00601	EL > 0.5	12
	EL > 0.8	0
FLA-I E*01201	EL > 0.5	81
	EL > 0.8	8

FLA-I H locus

Allele	Score	Total number of viral epitopes
FLA-I H*00102	EL > 0.5	285
	EL > 0.8	12
FLA-I H*00601	EL > 0.5	205
	EL > 0.8	3
FLA-I H*003011	EL > 0.5	202
	EL > 0.8	3
FLA-I H*00701	EL > 0.5	332
	EL > 0.8	37
FLA-I H*00501	EL > 0.5	289
	EL > 0.8	7
FLA-I H*00401	EL > 0.5	402
	EL > 0.8	46
FLA-I H*008012	EL > 0.5	494
	EL > 0.8	46
FLA-I H*003012	EL > 0.5	202
	EL > 0.8	3
FLA-I H*008011	EL > 0.5	494
	EL > 0.8	46
FLA-I H*00201	EL > 0.5	266
	EL > 0.8	5
FLA-I H*00101	EL > 0.5	285
	EL > 0.8	12

FLA-I K locus

Allele	Score	Total number of viral epitopes
FLA-I K*00401	EL > 0.5	19
	EL > 0.8	0
FLA-I K*00302	EL > 0.5	33
	EL > 0.8	0
FLA-I K*00801	EL > 0.5	310
	EL > 0.8	49
FLA-I K*00201	EL > 0.5	0
	EL > 0.8	0
FLA-I K*00701	EL > 0.5	155
	EL > 0.8	2
FLA-I K*00101	EL > 0.5	0
	EL > 0.8	0
FLA-I K*00501	EL > 0.5	57
	EL > 0.8	0
FLA-I K*00303	EL > 0.5	33
	EL > 0.8	0
FLA-I K*017	EL > 0.5	245
	EL > 0.8	7

Total number of epitopes > 250

Table 1. Number of epitopes deriving from viral glycoproteins predicted by NetMHCPan with EL score > 0.5 and > 0.8 for each allele in FLA-I loci. **For all 3 loci, the peptides with the highest EL scores were almost exclusively derived from the Spike (S) protein and Replicase 1ab (R1ab) protein of SARS-CoV-2 and FIPV.**

Sequence alignment

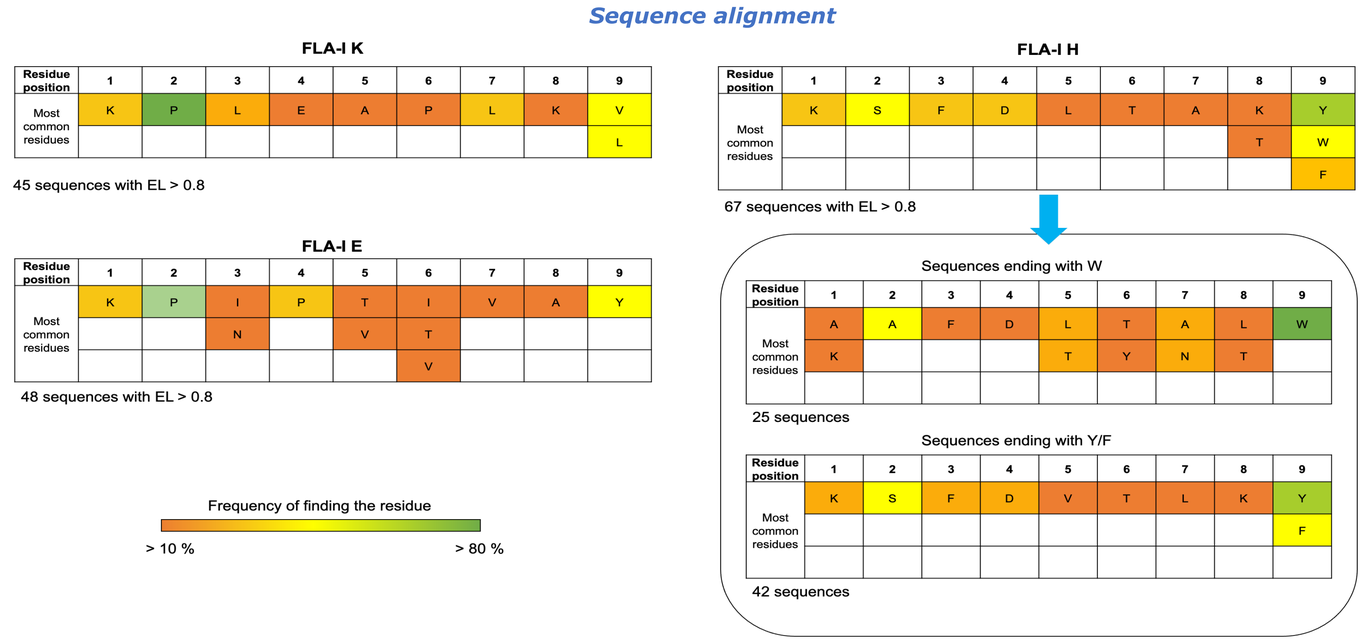
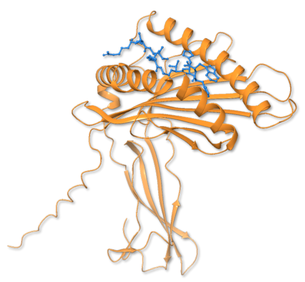


Figure 2. Sequence alignment of the peptides resulted in an EL-score above 0.8 from NetMHCPan analysis. The frequency of finding the amino acid reported in the table in the corresponding position is indicated in a colour scale (dark orange for frequency > 10%, bright green for frequency > 80%). As shown in the figure, proline is very common in position 2 for epitopes showing an affinity for E and K loci.

Figure 3. Molecular docking. Molecular docking of the NetMHCPan outputs with an EL-score above 0.8 using HPepDock for *ab initio* generation of the complex and AutoDock CrankPep (ADCP for the refinement. The peptides were sampled in two conformations, helix and extended. Considering an ADCP docking score lower than -18.0 kcal/mol, the results report that among the epitopes, the SARS-CoV-2 R1a-deriving sequence $_{3574}$RTIKGTHHW$_{3582}$ showed to bind with the highest affinity in both conformations the alleles FLA-I H*00501, *00401, *008012 and FLA-I K*00701.

Molecular docking



Viral glycoprotein	Epitope	FLA-I Allele	Extended docking score (kcal/mol)	Helix docking score (kcal/mol)
SARS-CoV-2 R1ab	$_{3574}$RTIKGTHHW$_{3582}$	FLA-I H*00501	-22.6	-22.5
SARS-CoV-2 R1ab	$_{6437}$KQFDTYNLW$_{6445}$	FLA-I K*00701	-20.8	-19.1
FIPV R1ab	$_{3756}$AANELNITW$_{3764}$	FLA-I H*00501	-20.5	-19.0
FIPV R1ab	$_{4533}$RLYYETLSY$_{4541}$	FLA-I H*00501	-19.8	-20.2
FIPV S	$_{1325}$RPNWTVPEF$_{1333}$	FLA-I E*00701	-16.6	-15.4
SARS-CoV-2 S	$_{625}$HADQLTPTW$_{633}$	FLA-I E*01001	-16.3	-14.5
SARS-CoV-2 S	$_{321}$QPTESIVRF$_{329}$	FLA-I E*00701	-15.6	-16.3
FIPV S	$_{771}$TTTPNFYYY$_{779}$	FLA-I H*00501	-16.5	-16.3
FIPV S	$_{1228}$TAYETVTAW$_{1236}$	FLA-I H*00401	-15.5	-15.5

Total results were filtered following an analysis of the exposition of the epitopes on the surfaces of the glycoproteins

Analysis of the epitope exposition on the viral glycoproteins

Molecular dynamics simulations

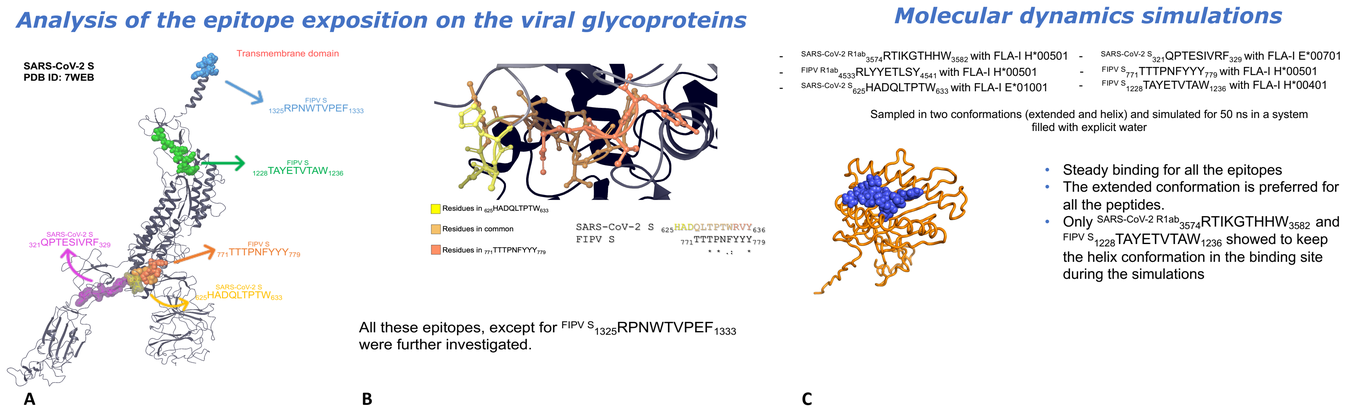


Figure 4. A) Cryo-EM structure of the monomer of SARS-CoV-2 S (PDB ID: 7WEB [26]) in ribbon representation. The possible localization of the FIPV S deriving epitopes is shown in CPK visualization, while the actual localization of SARS-CoV-2 S deriving epitopes is shown as transparent surfaces. The sequence alignment of the FIPV and SARS-CoV-2 S sequences on the left reports a high identity in this domain between the two proteins (asterisk indicates conserved residues, colon indicates amino acids with a strong similarity and dot indicates amino acids with a weak similarity). **B)** Focus on the overlapping residues in the possible location of the FIPV S deriving epitope $_{771}$TTTPNFYYY$_{779}$ with the real location of the SARS-CoV-2 S deriving epitope $_{625}$HADQLTPTW$_{633}$. C**) 50 ns classical MD simulations** in water at the average body temperature of cats (311.65 K).

CONCLUSIONS

Our data provide a basis to further explore the interaction between the feline immune system and peptide vaccines able to prevent untreatable viral diseases. Moreover, our investigation may be used as a valid approach to predict other viral glycoproteins potentially immunogenic in a wide number of species.

REFERENCES: 1) Liang, R. et al. Major Histocompatibility Complex Class I (FLA-E*01801) Molecular Structure in Domestic Cats Demonstrates Species-Specific Characteristics in Presenting Viral Antigen Peptides. J. Virol. 92, (2017). (2) Correale, P. et al. HLA-B*44 and C*01 Prevalence Correlates with Covid19 Spreading across Italy. Int. J. Mol. Sci. 21, 5205 (2020). (3) Buonocore, M. et al. New putative animal reservoirs of SARS-CoV-2 in Italian fauna: a bioinformatic approach. Heliyon 6, e05430 (2020). (4) Holmes, J. C. et al. Polymorphisms and tissue expression of the feline leukocyte antigen class I loci FLAI-E, FLAI-H, and FLAI-K. Immunogenetics 65, 675–689 (2013).