

Genetic alteration of human *MYH6* is mimicked by SARS-CoV-2 polyprotein: mapping viral variants of cardiac interest

Praveen Anand¹⁺, Patrick J. Lenehan²⁺, Michiel Niesen², Unice Yoo², Dhruvi Patwardhan¹, Marcelo Montorzi^{2,3}, AJ Venkatakrisnan^{2*}, Venky Soundararajan^{1,2*}

¹ nference Labs, Bengaluru, Karnataka 560017, India

² nference, Cambridge, Massachusetts 02139, USA

³ Southcoast Health, Fairhaven, Massachusetts 02719, USA

*These authors contributed equally

*Correspondence to: Venky Soundararajan (venky@nference.net) and AJ Venkatakrisnan (aj@nference.net)

Abstract

Acute cardiac injury has been observed in a subset of COVID-19 patients, but the molecular basis for this clinical phenotype is unknown. It has been hypothesized that molecular mimicry may play a role in triggering an autoimmune inflammatory reaction in some individuals after SARS-CoV-2 infection. Here we investigate if linear peptides contained in proteins that are primarily expressed in the heart also occur in the SARS-CoV-2 proteome. Specifically, we compared the library of 136,704 8-mer peptides from 144 human proteins (including splicing variants) to 9,926 8-mers from all 17 viral proteins in the reference SARS-CoV-2 proteome. No 8-mers were exactly identical between the reference human proteome and the reference SARS-CoV-2 proteome. However, there were 45 8-mers that differed by only one amino acid when compared to the reference SARS-CoV-2 proteome. Interestingly, analysis of protein-coding mutations from 141,456 individuals showed that one of these 8-mers from the SARS-CoV-2 Replicase polyprotein 1a/1ab (KIALKGGK) is identical to a *MYH6* peptide encoded by the c.5410C>A (Q1804K) genetic variation, which has been observed at low prevalence in Africans/African Americans (0.08%), East Asians (0.3%), South Asians (0.06%) and Latino/Admixed Americans (0.003%). Furthermore, analysis of 4.85 million SARS-CoV-2 genomes from over 200 countries shows that viral evolution has already resulted in 20 additional 8-mer peptides that are identical to human heart-enriched proteins encoded by reference sequences or genetic variants. Whether such mimicry contributes to cardiac inflammation during or after COVID-19 illness warrants further experimental evaluation. We suggest that SARS-CoV-2 variants harboring peptides identical to human cardiac proteins should be investigated as ‘viral variants of cardiac interest’.

Introduction

Cardiac injury is a prevalent complication associated with COVID-19.¹ In a study of 100 recently recovered COVID-19 patients, cardiovascular magnetic resonance imaging revealed cardiac involvement or ongoing myocardial inflammation in 78 and 60 patients, respectively.² In another study of 39 consecutive autopsies from patients who died of COVID-19, viral RNA was

detectable in the heart of 24 (62%) patients. A large nationwide study from Israel reported that SARS-CoV-2 infection is associated with increased rates of myocarditis, arrhythmia, myocardial infarction, and pericarditis.³ Myocarditis has also been reported in a small fraction of individuals after receiving an mRNA COVID-19 vaccine.⁴⁻⁶

Despite these phenotypic associations, the mechanisms underlying myocardial inflammation in the setting of COVID-19 infection and vaccination remain unclear. A prevalent hypothesis, known as molecular mimicry, posits that T lymphocytes and/or antibodies that recognize SARS-CoV-2 antigens and mediate virus neutralization may also cross-react against host cardiac proteins and trigger an autoimmune response against cardiomyocytes.⁷ This mechanism has also been suggested to contribute to other inflammatory conditions seen in the context of COVID-19 infection.⁸⁻¹⁰ Indeed, autoimmune sequelae of other infectious diseases have been attributed to mimicry between host and microbial antigens.¹¹⁻¹⁶

Advances in next-generation sequencing technologies have facilitated the rapid development of large-scale multi-omic datasets and genomic epidemiology resources to better understand the COVID-19 pandemic. Bulk and single cell RNA-sequencing datasets have elucidated the transcriptional signatures of most healthy human tissues and cell types.¹⁷⁻²⁰ Amino acid sequences of human proteins, including genetic variants and immunologic epitopes, are available in UniProt²¹, gnomAD²², and Immune Epitope Database (IEDB)²³. The GISAID database currently hosts 4.85 million SARS-CoV-2 genomes from more than 200 countries.²⁴ The availability of such genome-scale data enables us to investigate the potential for molecular mimicry between SARS-CoV-2 and human cardiac proteins.

Here we present a systematic comparison of peptides from human cardiac proteins and SARS-CoV-2 proteins. We show that no 8-mer peptides are identical between the reference sequences of these two groups of proteins. However, when including human and viral genetic variants in this comparison, we found 21 8-mer peptides to be identical between human cardiac proteins and SARS-CoV-2 proteins. Among these, a human genetic variant of *MYH6* (c.5410C>A; Q1804K) is identical to a peptide of the reference SARS-CoV-2 replicase polyprotein. Finally, we propose that the SARS-CoV-2 variants that have peptides identical to human cardiac proteins should be studied as potential 'viral variants of cardiac interest'.

Results

Identification of genes that are overexpressed in cardiac tissue

To identify heart-enriched proteins, we compared the expression of all human protein-coding genes in heart samples ($n = 861$) versus all non-striated muscle samples ($n = 15,718$) from the Genotype Tissue Expression (GTEx) project. There were 137 genes expressed at least 5-fold higher in the heart with a Cohen's D value greater than or equal to 0.5 (**Figure 1, Supplementary Table 1**). Similarly, we compared the expression of all genes between cardiomyocytes ($n = 8.9K$ cells) and non-cardiomyocytes ($n = \sim 2.5$ million cells) across a database of 52 single-cell RNA-sequencing studies covering 62 tissues.¹⁸ There were 46 genes overexpressed in cardiomyocytes based on the same criteria outlined above (**Figure 1,**

Supplementary Table 1). Combining the lists of genes identified from bulk and single-cell RNA-sequencing analyses, we identified a total of 144 candidate cardiac proteins.

MYH6 variant is mimicked by an epitope of SARS-CoV-2 Replicase polyprotein 1a/1ab

We computed peptide 8-mers from the reference sequences of the 144 cardiac proteins (including isoforms) and all 17 proteins from the reference SARS-CoV-2 sequence. We then systematically compared the pairwise sequence identity (using Hamming Distance; see **Methods**) for the 136,704 cardiac protein 8-mers with the 9,926 8-mers derived from the SARS-CoV-2 proteins. No peptides were fully identical between these two groups. However, 45 8-mers were nearly identical, with only a single mismatched amino acid (**Table 1**). To determine whether human genetic variation results in any 8-mers which exactly match the reference SARS-CoV-2 proteome, we then analyzed amino acid mutations from 141,456 individuals using the gnomAD database (including 83,623 mutations in the 144 cardiac proteins).²²

Interestingly, one specific 8-mer from the SARS-CoV-2 Replicase polyprotein 1a/1ab (KIALKGGK) is identical to a mutant peptide encoded by the c.5410C>A(Gln1804Lys) variation in human *MYH6* (KIALKGGK), which is a subunit of a cardiac motor protein. This genetic variation has been identified in Africans/African Americans (0.08% prevalence), East Asians (0.3% prevalence), South Asians (0.06% prevalence) and Latino/Admixed Americans (0.003% prevalence).²² Analysis of peptides from IEDB shows that the non-mutated 7-mer in this peptide (IALKGGK) also overlaps with a known B-cell epitope from SARS-CoV-2 (IALKGGKIVNNWLKQ)²⁵. Whether the mimicry between SARS-CoV-2 Replicase polyprotein 1a/1ab and wild-type or mutant *MYH6* contributes to cardiac inflammation in the setting of COVID-19 warrants further investigation.

SARS-CoV-2 variants harbor epitopes that are identical to peptides of cardiac proteins

We next examined whether SARS-CoV-2 evolution has given rise to variants that harbor peptides identical to human cardiac proteins. To this end, we analyzed 4.85 million SARS-CoV-2 genomes from over 200 countries obtained from the GISAID database.²⁴ We identified 21 8-mer peptides from SARS-CoV-2 variants that are identical to cardiac proteins encoded by reference sequences or genetic variants (**Tables 2 and 3**). Of these 21 peptides, the one present in the highest number of viral sequences was STAVLGVL, which mapped to the NSP3 protein (part of Replicase polyprotein sequence) in 4,501 (0.09%) SARS-CoV-2 genomes. This sequence is present in a transmembrane helix of all three described isoforms of the cardiac protein SLC4A3, an anion exchange protein that is associated with short QT syndrome (**Tables 2 and 3**).^{26,27} Although no epitopes containing this full sequence were present in IEDB, a BLAST search at 90% similarity identified a similar epitope (HEAQAVLGVLL) that is reported to bind to both HLA-B*40:1 and HLA-B*58:01.^{25,28} We additionally analyzed the temporal and geographical emergence of the five variant SARS-CoV-2 8-mers that are identical to cardiac protein 8-mers and that occur in over 100 reported SARS-CoV-2 genomes (**Supplementary Figure S1**). Variants with exactly matching 8-mers have occurred sporadically in different geographic locations, but they have not persisted over time (**Supplementary Figure S1**).

Discussion

Understanding the mechanistic basis of acute cardiac injury in COVID-19 patients is important to develop countermeasures. Viral infections have previously been proposed to trigger autoimmune reactions, and it has been hypothesized that molecular mimicry plays a role in mediating autoimmune reactions.^{7–9,11–13,29} Here, a systematic analysis of gene-expression from cardiac tissues, based on bulk RNAseq data and single cell RNAseq data, combined with analysis of human genetic variation and SARS-CoV-2 genomes has led to the identification of candidate proteins and peptide regions therein that might be involved in immune cross reactivity (**Tables 2 and 3**). These newly identified identical peptides expand the set of known shared peptides between the human proteome and SARS-CoV-2, such as the furin cleavage site 'RRARSVAS' present in both human ENaC- α and the Spike glycoprotein.^{10,30,31} Further research is warranted to ascertain whether these mimicked peptides contribute to autoinflammatory pathology in the context of COVID-19 infection. Additionally, given the occurrence of myocarditis in some individuals shortly after receiving an mRNA COVID-19 vaccine, the potential for molecular mimicry between the antigen encoded by these vaccines (i.e., the pre-fusion stabilized Spike glycoprotein) and human cardiac proteins should be evaluated.^{4–6,32,33}

There are a few limitations to this study. First, the human proteins that we surveyed were shortlisted based on their overexpression in cardiac tissue. There could be mimicked proteins that are shared between cardiac tissues and other tissues that are not accounted for in the current analysis. Second, there are other mechanisms that could contribute to autoimmunity after viral infection such as bystander activation, epitope spreading, and viral persistence³⁴. Third, the presence of identical peptides in cardiac proteins and the SARS-CoV-2 proteome could occur due to chance. Comparing all SARS-CoV-2 8-mers to a set of brain enriched proteins and a set of skin enriched proteins shows a similar probability distribution of Hamming distance (**Supplementary Figure S2**), suggesting that the observed similarity with SARS-CoV-2 peptides is not specific to human cardiac proteins. Fourth, it is possible that peptides with lower degrees of similarity could contribute to immunologic mimicry, as T cells can be highly cross reactive against different major histocompatibility complex (MHC)-presented peptides.^{35–39}

Taken together, by studying the intersection of human genetic variation in cardiac proteins and SARS-CoV-2 evolution, we have identified candidates of molecular mimicry that have potential to contribute to cardiac inflammation in the context of COVID-19. It will be important to perform follow-up functional studies evaluating the potential of SARS-CoV-2 reactive T cells and antibodies (e.g., from active or recovering COVID-19 patients) to cross-react with these peptides. Thus, we propose that SARS-CoV-2 variants harboring peptides identical to host heart-enriched proteins should be studied as 'viral variants of cardiac interest'. We highlight that a similar strategy can be applied to identify and categorize plausible mimicry candidates from any human tissues that are targeted by other autoimmune responses in COVID-19 patients.

Methods

Identification of proteins enriched in cardiac tissue

Bulk RNA-sequencing (RNA-seq) data was accessed from the Genotype Tissue Expression (GTEx) project V8.¹⁷ For each sample, FASTQ files were processed using Salmon (in mapping-based mode) to quantify gene expression in transcripts per million (TPM). Specifically, the expression of each transcript isoform was first determined by passing FASTQ files to Salmon *quant* with the following parameters passed: *validateMappings*, *rangeFactorizationBins* 4, *gcBias*, *biasSpeedSamp* 10. All isoforms are then summed via a transcript-to-gene map, generating a gene-level expression value. GRCh38 was used as the reference, including cDNA and non-coding RNA.

For single cell RNA-seq studies, processed count matrices were accessed from Gene Expression Omnibus or other publicly available data repositories. There were two datasets analyzing heart tissues which captured cardiomyocytes, our main cell type of interest for this report.^{20,40} Other datasets captured a wide variety of immune, stromal, and parenchymal cell types from tissues including the respiratory tract, gastrointestinal tract, genitourinary tract, hepatobiliary system, skeletal muscle, brain, skin, eyes, and endocrine organs. Each dataset was processed using Scrublet and Seurat v3.0 as described previously.^{18,41–43} Cell type annotations were obtained from associated metadata files if available; otherwise, annotation was performed manually, guided by the cell types reported in the associated publication.

To identify genes that are overexpressed in cardiac tissue, we calculated fold change and Cohen's D values between defined sample cohorts. For bulk RNA-seq data, Cohort A was defined as all GTEx heart samples ($n = 861$), and Cohort B was defined as all remaining GTEx samples except for those derived from skeletal muscle ($n = 15,718$). For single cell RNA-seq data, Cohort A was defined as all cells annotated as cardiomyocytes ($n \sim 8900$ cells), and Cohort B was defined as all other cells from all processed studies ($n \sim 2.5$ million cells). Fold change and Cohen's D were calculated as follows:

$$\mathbf{Fold\ Change} = \frac{TPM_{Cohort\ A} + 1}{TPM_{Cohort\ B} + 1}$$

$$\mathbf{Cohen's\ D} = \frac{TPM_{Cohort\ A} - TPM_{Cohort\ B}}{SD_{pooled}}, \text{ where the pooled standard deviation } SD_{pooled} \text{ is defined as:}$$

$$SD_{pooled} = \sqrt{\frac{(N_{Cohort\ A} - 1) \times SD_{Cohort\ A}^2 + (N_{Cohort\ B} - 1) \times SD_{Cohort\ B}^2}{(N_{Cohort\ A} + N_{Cohort\ B} - 2)}}, \text{ where } N_{Cohort\ A} \text{ and } N_{Cohort\ B} \text{ are the number of samples in Cohorts A and B, respectively, and } SD_{Cohort\ A} \text{ and } SD_{Cohort\ B} \text{ are the standard deviation of TPM values for the given gene in Cohorts A and B, respectively.}$$

Genes with Fold Change ≥ 5 and Cohen's D ≥ 0.5 from either the bulk or single cell RNA-seq analysis were considered to be enriched in cardiac tissue. In the volcano plots used to visualize these analyses, we filtered to genes with a TPM or CP10K value ≥ 1 in either Cohort A or Cohort B, and genes meeting the criteria for overexpression in heart or cardiomyocytes are colored in red.

For a control analysis, we also identified genes overexpressed in the brain or skin by bulk RNA-sequencing from the GTEx project. We used the same approach as described above, except that Cohort A was defined as either all brain samples (n = 2,351) or all skin samples (n = 1,305), and Cohort B was defined as all other samples.

Comparison of 8-mers from reference sequences of cardiac proteins and SARS-CoV-2 proteins

The translated proteome from reference the SARS-CoV-2 genome (NC_045512.2) was downloaded from UniProt.^{44,45} A sliding window approach was used to enumerate all 8-mers from the 17 proteins in this viral proteome. Similarly, we used a sliding window approach to generate all 8-mers from the reference amino acid sequences of the previously defined 144 cardiac proteins, including the canonical isoforms and all described isoforms indicated in UniProt. We then performed a pairwise comparison of all 8-mers in these two groups by calculating the Hamming distance using the *stringdist* function from the *stringdist* package (version 0.9.8) in R (version 4.0.3). In a control analysis, we used the same approach to calculate the Hamming distance between all SARS-CoV-2 8-mers and the control sets of 369 human proteins enriched in the brain or 198 human proteins enriched in the skin (described above).

Assessing the impact of human and SARS-CoV-2 variants on cardiac peptide matches

To assess the impact of human genetic variation on potential molecular mimicry, we retrieved all missense variants from the gnomAD database for the previously identified cardiac proteins that had at least one 8-mer similar to peptide in the SARS-CoV-2 reference proteome (Hamming distance = 1).²² We used the gnomad-api (<https://gnomad.broadinstitute.org/api>) to fetch the variant calls from gnomad_r2_1 version from the Human GRCh37 genome assembly. The variants in this gnomad version (GRCh37/hg19) are derived from 125,748 exome sequences and 15,708 whole-genome sequences from unrelated individuals sequenced as part of various disease-specific and population genetic studies. For any variants that alter the amino acid sequence of a potentially mimicked peptide, we determined whether the mutation resulted in an exact match (Hamming distance = 0) to the corresponding 8-mer from the SARS-CoV-2 reference proteome.

To assess the impact of viral evolution on potential molecular mimicry, we queried the cardiac 8-mers with Hamming distance of 1 (including any alterations of these 8-mers arising from human genetic variation as described above) against all protein variants encoded in 4,854,709 SARS-CoV-2 genomes deposited in the GISAID database (last accessed 11/5/2021).²⁴ Here, we determined whether any mutations in viral genomes (relative to the reference sequence) resulted in 8-mers which exactly match one or more cardiac peptides.

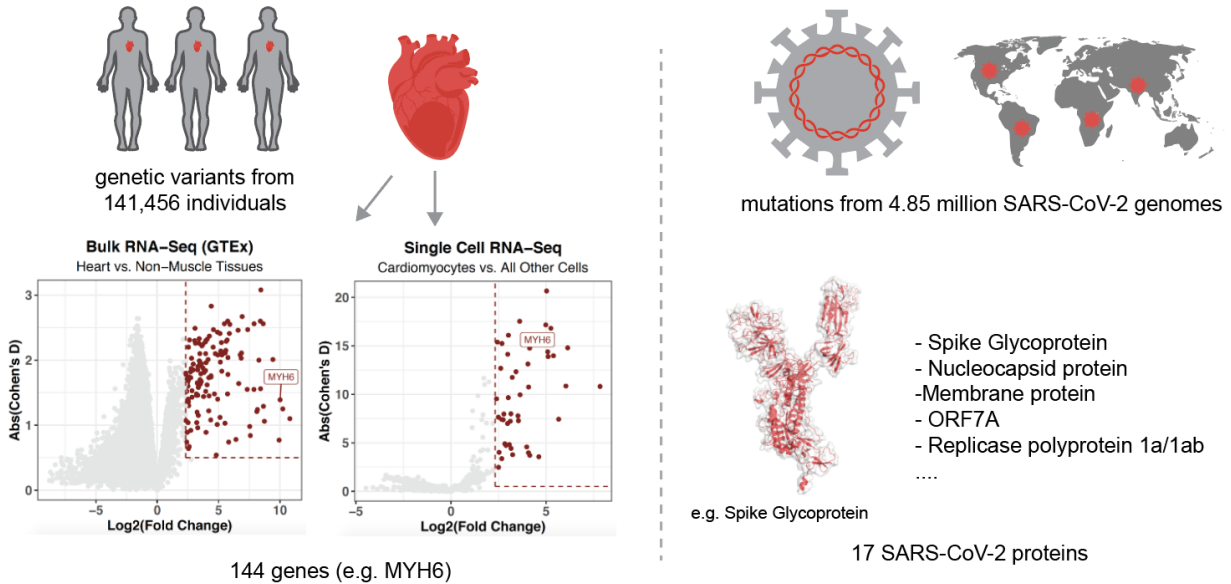
Evaluation of mimicked peptides for inclusion in immune epitopes

For any 8-mers which showed an exact match between a cardiac peptide (reference or variant sequences) and a SARS-CoV-2 peptide (reference or variant sequences), we queried the 8-mer using the Immune Epitope Database (IEDB; www.iedb.org) and Analysis Resource.²³

We searched for any linear peptide epitope with a Blast similarity of at least 90% from any human host that had positive experimental evidence in any assay (T cell, B cell, or MHC Ligand). No MHC class restrictions or disease filters were applied.

Figures

a. Identification of cardiac specific proteins and SARS-CoV-2 proteins



b. Comparison of peptide libraries of human cardiac proteins and SARS-CoV-2 proteins

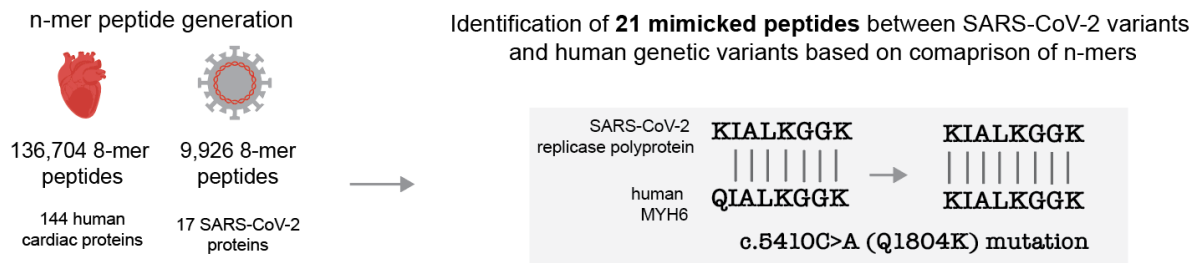


Figure 1. Identification of mimicked peptides between SARS-CoV-2 and human proteins

a. Identification of cardiac specific proteins based on analysis of bulk RNAseq and single-cell RNAseq data and identification of SARS-CoV-2 proteins. **b.** Comparison of peptide libraries of human cardiac proteins and SARS-CoV-2 proteins.

Tables

Table 1: List of peptide pairs from SARS-CoV-2 proteins and human cardiac proteins that have a Hamming distance less than or equal to 1.

| SARS-CoV-2 Proteins | | | Human Cardiac-Enriched Proteins | | |
|------------------------------|----------------------|---------------------|---------------------------------|--|---------------------|
| Protein Name | Amino Acid Positions | Amino Acid Sequence | Protein Name | Isoforms Containing Sequence | Amino Acid Sequence |
| Spike glycoprotein WT/2P | 491-498 | PLQSYGFQ | KLHL41 | 1, 2 | PLQSYFFQ |
| Spike glycoprotein WT/2P | 856-863 | NGLTVLPP | FHOD3 | 1, 2, 3, 4 | IGLTVLPP |
| Spike glycoprotein WT/2P | 857-864 | GLTVLPLL | FHOD3 | 1, 2, 3, 4 | GLTVLPPP |
| Spike glycoprotein WT/2P | 1087-1094 | AHFPREGV | CMYA5 | 1 | AHFPAEGV |
| Replicase polyprotein 1ab | 4937-4944 | KYAISAKN | TTN | 1, 3, 2, 5, 12, 13, 7, 8, 4, 9, 10, 11 | KYIISAKN |
| Replicase polyprotein 1ab | 5604-5611 | LQGPPGTG | MYLK3 | 1, 4 | KQGPPGTG |
| Replicase polyprotein 1ab | 5605-5612 | QGPPGTGK | MYLK3 | 1, 4 | QGPPGTGR |
| Replicase polyprotein 1ab | 5813-5820 | NRPQIGVV | CASQ2 | 1, 2 | FRPQIGVV |
| Replicase polyprotein 1ab | 5814-5821 | RPQIGVVR | CASQ2 | 1, 2 | RPQIGVWN |
| Replicase polyprotein 1ab | 5955-5962 | DTKFKTEG | TTN | 1, 3, 2, 5, 12, 13, 7, 8, 4, 9, 10, 11 | DTKFKTTG |
| Replicase polyprotein 1ab | 5955-5963 | DTKFKTEGL | TTN | 1, 3, 2, 5, 12, 13, 7, 8, 4, 9, 10, 11 | DTKFKTTGL |
| Replicase polyprotein 1ab | 5956-5963 | TKFKTEGL | TTN | 1, 3, 2, 5, 12, 13, 7, 8, 4, 9, 10, 11 | TKFKTTGL |
| Replicase polyprotein 1ab | 6516-6523 | KPVPEVKI | CMYA5 | 1 | KPSPEVKI |
| Replicase polyprotein 1ab | 6516-6523 | KPVPEVKI | TTN | 1, 2, 5, 12, 13, 7, 8, 4, 11 | KPVPEEKI |
| Replicase polyprotein 1a/1ab | 207-214 | RAGKASCT | TTN | 1, 2, 5, 12, 13, 7, 8, 4, 11 | EAGKASCT |
| Replicase polyprotein 1a/1ab | 208-215 | AGKASCTL | TTN | 1, 2, 5, 12, 13, 7, 8, 4, 11 | AGKASCTT |
| Replicase polyprotein 1a/1ab | 345-352 | GTEMLTKE | TNNI3K | 1, 3, 4 | GTESLTKE |
| Replicase polyprotein 1a/1ab | 459-466 | VNINIVGD | ENO3 | 1, 3, 2 | VNIQIVGD |
| Replicase polyprotein 1a/1ab | 492-499 | KGLDYKAF | CASQ2 | 2 | KKLDYKAF |
| Replicase polyprotein 1a/1ab | 512-519 | TKGKAKKG | MYH7 | 1 | GKGKAKKG |

| | | | | | |
|------------------------------|-----------|----------|---------|------------------------------|----------|
| Replicase polyprotein 1a/1ab | 513-520 | KGKAKKGA | MYH7 | 1 | KGKAKKGS |
| Replicase polyprotein 1a/1ab | 879-886 | VIKTLQPV | LMOD3 | 1 | VIKTLKPV |
| Replicase polyprotein 1a/1ab | 963-970 | GATSAALQ | ANKRD2 | 1, 2 | GAQSAALQ |
| Replicase polyprotein 1a/1ab | 1143-1150 | VLLAPLLS | HJV | 1, 2, 3 | TLLAPLLS |
| Replicase polyprotein 1a/1ab | 1144-1151 | LLAPLLSA | HJV | 1, 2, 3 | LLAPLLSG |
| Replicase polyprotein 1a/1ab | 1197-1204 | KQVEQKIA | GOT1 | 1, 2 | KKVEQKIA |
| Replicase polyprotein 1a/1ab | 2246-2253 | STAALGVL | SLC4A3 | 1, 2, 3 | STAVLGVL |
| Replicase polyprotein 1a/1ab | 2533-2540 | KGSLPINV | TTN | 1, 2, 5, 12, 13, 7, 8, 4, 11 | KGSLPITV |
| Replicase polyprotein 1a/1ab | 2550-2557 | EESAKSA | MYH7B | 1 | EESAKSA |
| Replicase polyprotein 1a/1ab | 2630-2637 | LSTFISAA | TENM2 | 1, 2 | LSTFFSAA |
| Replicase polyprotein 1a/1ab | 2757-2764 | KIALKGGK | MYH6 | 1 | QIALKGGK |
| Replicase polyprotein 1a/1ab | 2757-2764 | KIALKGGK | MYH7 | 1 | QIALKGGK |
| Replicase polyprotein 1a/1ab | 2758-2765 | IALKGGKI | MYH6 | 1 | IALKGGKK |
| Replicase polyprotein 1a/1ab | 2758-2765 | IALKGGKI | MYH7 | 1 | IALKGGKK |
| Replicase polyprotein 1a/1ab | 3908-3915 | FEKMSVLL | MYH6 | 1 | EEKMSVLL |
| Replicase polyprotein 1a/1ab | 3908-3915 | FEKMSVLL | MYH7 | 1 | EEKMSVLL |
| Replicase polyprotein 1a/1ab | 3909-3916 | EKMVSLLS | MYH6 | 1 | EKMVSLLQ |
| Replicase polyprotein 1a/1ab | 3909-3916 | EKMVSLLS | MYH7 | 1 | EKMVSLLQ |
| Replicase polyprotein 1a/1ab | 4137-4144 | VKLQNNEL | TBX20 | 1 | VKLQNNEL |
| Putative ORF9c protein | 47-54 | AAVGELLL | ASB10 | 1, 2, 3 | AAVVELLL |
| ORF7b protein | 14-21 | LAFLFLV | TMEM182 | 2, 1, 3 | LAGLLFLV |
| ORF7a protein | 43-50 | NSPFHPLA | FLNC | 1, 2 | NSPFHVLA |
| Nucleoprotein | 192-199 | NSSRNSTP | CMYA5 | 1 | NSSRSSTP |
| Nucleoprotein | 374-381 | KKADETQA | MYPN | 1 | EKADETQA |
| Nucleoprotein | 375-382 | KADETQAL | MYPN | 1 | KADETQAR |

Table 2: List of identical cardiac peptides found in SARS-CoV-2 variants in GISAID. The NSP proteins are cleaved products of the replicase polyprotein.

| No. of GISAID Entries with Identical Match | SARS-CoV-2 Gene | Mimicked Peptide | Approx. Star-End in SCOVID Protein | Cardiac Protein | Cardiac Protein Start-End | Exact IEDB epitope match | Identical IEDB Epitope Seq (HLA/antigen infor) [$\geq 90\%$ identity] |
|--|-----------------|------------------|------------------------------------|-----------------|---------------------------|--------------------------|--|
| 4501 | NSP3 | STAVLGVL | 1429-1436 | SLC4A3 | 746-753 | No | HEAQA VLGVL (HLA-B*40:01); HEAQA VLGVL (HLA-B*40:01) |
| 1322 | NSP2 | GTESLTKE | 166-173 | TNNI3K | 294-301 | No | NA |
| 580 | N | NSSRSSTP | 193-200 | CMYA5 | 88-95 | No | TINSSRSSQESY (B-cell epitopes, MHC ligands) |
| 205 | NS9c | AAVVELL | 48-55 | ASB10 | 307-314 | No | MVDPQLDGPQLAALA AVVELGSFDA () |
| 118 | NSP2 | KGKAKKGS | 334-341 | MYH7 | 635-642 | Yes | AGADAPIEKKGKAKKGS (MHC ligand) |
| 33 | NSP3 | VIKTLKPV | 60-67 | LMOD3 | 514-521 | No | KVAIKTLKPGTMS (HLA-A*02:01); TKVAIKTLKPGTMSPE (HLA-A*02:01) |
| 17 | NSP3 | KKVEQKIA | 380-387 | GOT1 | 55-62 | | GEKVEQKIEGKWNKKAQEDKLQ (MHC Class I, II, Bcell epitope) |
| 14 | NS7a | NSPFHVLA | 44-51 | FLNC | 1727-1734 | No | NA |
| 11 | N | EKADETQA | 375-382 | MYPN | 89-96 | No | ADETQALPQRQKKQQ (HLA Class II) |
| 9 | NSP3 | TLLAPLLS | 304-311 | HJV | 409-416 | | NFNQHEVLLAPLLS (Bcell epitope and MHC ligand) |
| 6 | NSP3 | LSTFFSAA | 1811-1818 | TENM2 | 1036-1043 | No | AGTLSTFFGVPLVLT (HLA class II MHC restriction) |
| 6 | NSP3 | LLAPLLSG | 327-334 | HJV | 410-417 | No | ENFNQHEVLLAPLLS (Bcell and MHC) |
| 6 | Spike | IGLTVLPP | 854-861 | FHOD3 | 971-978 | No | GFIKQYGDCLGDIAA RDLICAQKFNGLTVL PPLLTDEMIAQYT (Tcell, Bcell and MHC ligands) |
| 2 | NSP13 | QGPPGTGR | 282-289 | MYLK3 | 373-380 | No | ILYGPPGTGK (HLA-A*03:01) |
| 1 | NSP15 | KPVPEEKI | 65-72 | TTN | 10277-10284 | No | EAPLYVVDKPVPEESE (HLA-DRB1*04:01) |
| 1 | NSP3 | KGSLPITV | 1716-1722 | TTN | 5437-5444 | No | SLPITVYYAV (Tcell, Bcell and MHC ligands) |

Table 3. List of mutated cardiac peptide nmers from human genetic variants identical to SARS-CoV-2 variants. The NSP proteins are cleaved products of the replicase polyprotein.

| Human Cardiac Gene | rsID | Mutation consequence | Mutated cardiac Peptide | Wild-type cardiac peptide | No. of GISAID genomes | SARS-CoV-2 gene | IEDB epitope Exact match | IEDB epitope info (>=90 % Seq Identity) |
|--------------------|-------------|----------------------|-------------------------|---------------------------|-----------------------|-----------------|--------------------------|---|
| FLNC | rs374848954 | p.Val1732Leu | NSPFHLLA | NSPFHVLA | 1661 | NS7a | Yes | NSPFH (HLA-A*01:01) |
| LMOD3 | rs370869958 | p.Lys519Arg | VIKTLRPV | VIKTLKPV | 85 | NSP3 | No | NA |
| TMEM182 | rs774398171 | p.Gly215Val | LAVLLFLV | LAGLLFLV | 21 | NS7b | No | NA |
| MYLK3 | rs771870674 | p.Arg380Cys | QGPPGTGC | QGPPGTGR | 1 | NSP13 | No | GPPGTGKSHFAIGLA (B cell, Tcell, MHC ligand) |

References

1. Chung, M. K. *et al.* COVID-19 and Cardiovascular Disease: From Bench to Bedside. *Circ. Res.* **128**, 1214–1236 (2021).
2. Puntmann, V. O. *et al.* Outcomes of Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From Coronavirus Disease 2019 (COVID-19). *JAMA Cardiol* **5**, 1265–1273 (2020).
3. Barda, N. *et al.* Safety of the BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Setting. *N. Engl. J. Med.* **385**, 1078–1090 (2021).
4. CDC. Myocarditis and Pericarditis After mRNA COVID-19 Vaccination. <https://www.cdc.gov/coronavirus/2019-ncov/vaccines/safety/myocarditis.html> (2021).
5. Witberg, G. *et al.* Myocarditis after Covid-19 Vaccination in a Large Health Care Organization. *New England Journal of Medicine* (2021) doi:10.1056/nejmoa2110737.
6. Simone, A. *et al.* Acute Myocarditis Following COVID-19 mRNA Vaccination in Adults Aged 18 Years or Older. *JAMA Intern. Med.* (2021) doi:10.1001/jamainternmed.2021.5511.
7. Bozkurt, B., Kamat, I. & Hotez, P. J. Myocarditis With COVID-19 mRNA Vaccines.

- Circulation* **144**, 471–484 (2021).
8. Proal, A. D. & VanElzakker, M. B. Long COVID or Post-acute Sequelae of COVID-19 (PASC): An Overview of Biological Factors That May Contribute to Persistent Symptoms. *Front. Microbiol.* **12**, 698169 (2021).
 9. Galeotti, C. & Bayry, J. Autoimmune and inflammatory diseases following COVID-19. *Nature reviews. Rheumatology* vol. 16 413–414 (2020).
 10. Kanduc, D. From Anti-SARS-CoV-2 Immune Responses to COVID-19 via Molecular Mimicry. *Antibodies (Basel)* **9**, (2020).
 11. Gowthaman, U. & Eswarakumar, V. P. Molecular mimicry: good artists copy, great artists steal. *Virulence* **4**, 433–434 (2013).
 12. Cusick, M. F., Libbey, J. E. & Fujinami, R. S. Molecular mimicry as a mechanism of autoimmune disease. *Clin. Rev. Allergy Immunol.* **42**, 102–111 (2012).
 13. Oldstone, M. B. A. Molecular Mimicry, Microbial Infection, and Autoimmune Disease: Evolution of the Concept. in *Molecular Mimicry: Infection-Inducing Autoimmune Disease* (ed. Oldstone, M. B. A.) 1–17 (Springer Berlin Heidelberg, 2005).
 14. Adderson, E. E., Shikhman, A. R., Ward, K. E. & Cunningham, M. W. Molecular analysis of polyreactive monoclonal antibodies from rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin antibody V region genes. *J. Immunol.* **161**, 2020–2031 (1998).
 15. Cunningham, M. W., Antone, S. M., Smart, M., Liu, R. & Kosanke, S. Molecular analysis of human cardiac myosin-cross-reactive B- and T-cell epitopes of the group A streptococcal M5 protein. *Infect. Immun.* **65**, 3913–3923 (1997).
 16. Cunningham, M. W. Autoimmunity and molecular mimicry in the pathogenesis of post-streptococcal heart disease. *Front. Biosci.* **8**, s533–43 (2003).
 17. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).

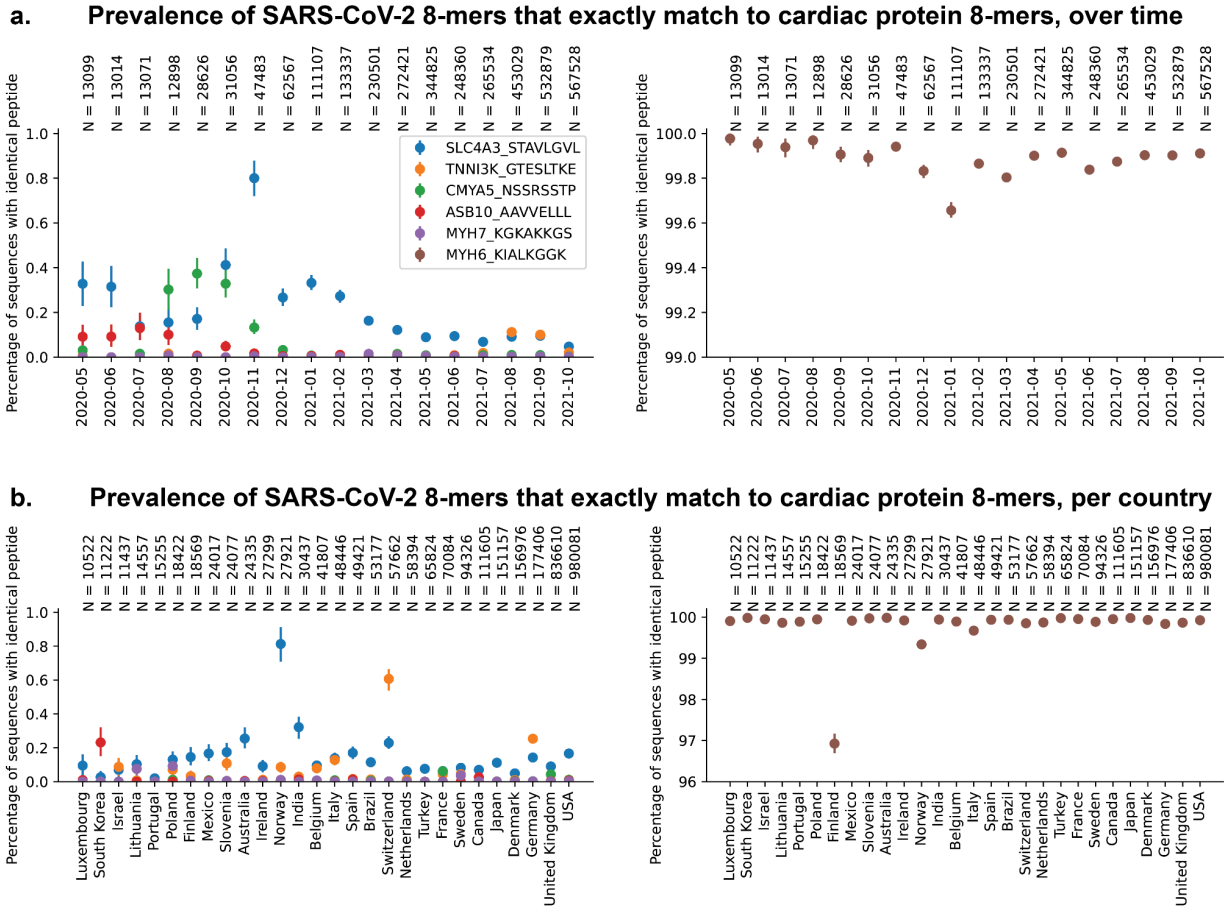
18. Venkatakrisnan, A. J. *et al.* Knowledge synthesis of 100 million biomedical documents augments the deep expression profiling of coronavirus receptors. *Elife* **9**, (2020).
19. Rozenblatt-Rosen, O., Stubbington, M. J. T., Regev, A. & Teichmann, S. A. The Human Cell Atlas: from vision to reality. *Nature* **550**, 451–453 (2017).
20. Han, X. *et al.* Construction of a human cell landscape at single-cell level. *Nature* **581**, 303–309 (2020).
21. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
22. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
23. Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339–D343 (2019).
24. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro Surveill.* **22**, (2017).
25. IEDB.org: Free epitope database and prediction resource.
https://www.iedb.org/home_v3.php.
26. Thorsen, K. *et al.* Loss-of-activity-mutation in the cardiac chloride-bicarbonate exchanger AE3 causes short QT syndrome. *Nat. Commun.* **8**, 1696 (2017).
27. Anion exchange protein 3. <https://www.uniprot.org/uniprot/P48751>.
28. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* vol. 215 403–410 (1990).
29. Ehrenfeld, M. *et al.* Covid-19 and autoimmunity. *Autoimmunity Reviews* vol. 19 102597 (2020).
30. Anand, P., Puranik, A., Aravamudan, M., Venkatakrisnan, A. J. & Soundararajan, V. SARS-CoV-2 strategically mimics proteolytic activation of human ENaC. *Elife* **9**, (2020).
31. Venkatakrisnan, A. J. *et al.* Benchmarking evolutionary tinkering underlying human-viral

- molecular mimicry shows multiple host pulmonary-arterial peptides mimicked by SARS-CoV-2. *Cell Death Discov* **6**, 96 (2020).
32. Corbett, K. S. *et al.* SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* **586**, 567–571 (2020).
 33. Walsh, E. E. *et al.* Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates. *N. Engl. J. Med.* **383**, 2439–2450 (2020).
 34. Fujinami, R. S., von Herrath, M. G., Christen, U. & Whitton, J. L. Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. *Clin. Microbiol. Rev.* **19**, 80–94 (2006).
 35. Sewell, A. K. Why must T cells be cross-reactive? *Nat. Rev. Immunol.* **12**, 669–677 (2012).
 36. Borbulevych, O. Y., Santhanagopalan, S. M., Hossain, M. & Baker, B. M. TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. *J. Immunol.* **187**, 2453–2463 (2011).
 37. Scott, D. R., Borbulevych, O. Y., Piepenbrink, K. H., Corcelli, S. A. & Baker, B. M. Disparate degrees of hypervariable loop flexibility control T-cell receptor cross-reactivity, specificity, and binding mechanism. *J. Mol. Biol.* **414**, 385–400 (2011).
 38. Garcia, K. C. *et al.* Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* **279**, 1166–1172 (1998).
 39. Wooldridge, L. *et al.* A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177 (2012).
 40. Wang, L. *et al.* Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat. Cell Biol.* **22**, 108–119 (2020).
 41. Doddahonnaiah, D. *et al.* A Literature-Derived Knowledge Graph Augments the Interpretation of Single Cell RNA-seq Datasets. *Genes* **12**, (2021).
 42. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21

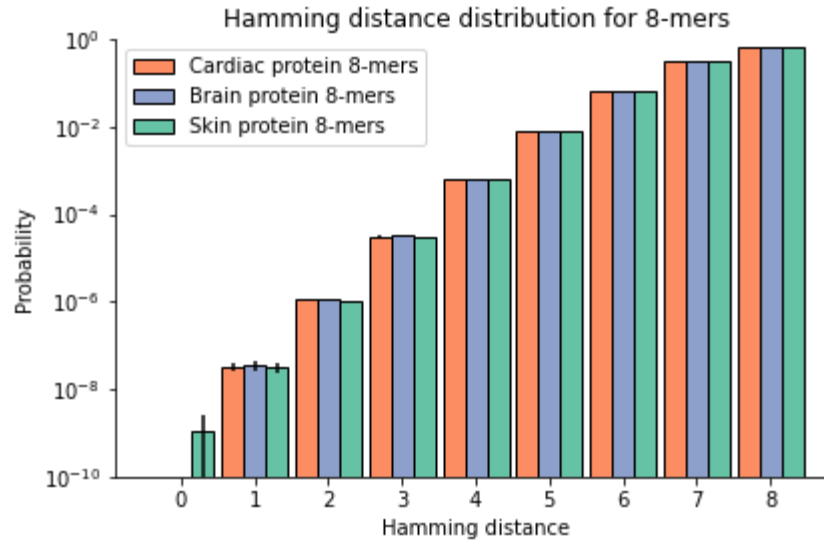
(2019).

43. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019).
44. UniProt.
https://covid-19.uniprot.org/uniprotkb?query=*%26amp%26facets=other_organism:Severe%20acute%20respiratory%20syndrome%20coronavirus%202
45. Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, co - Nucleotide - NCBI. https://www.ncbi.nlm.nih.gov/nucore/NC_045512.2.

Supplementary Material



Supplementary Figure S1: Prevalence of SARS-CoV-2 variants with exact 8-mer matches to human cardiac proteins. Prevalence is reported as the percentage of SARS-CoV-2 sequences with exact 8-mer matches during 1-month time intervals (a) or across all sequences reported by a country (b). Prevalence of the wild-type SARS-CoV-2 8-mer that exactly matches with the MYH6 peptide, described in the main text, is shown separately on the right (brown). Data is only shown for months and countries in which more than 10,000 SARS-CoV-2 genomes were reported to GISAID.



Supplementary Figure S2. Observed probability distribution for the Hamming distance between all 8-mers in the SARS-CoV-2 proteome and all 8-mers in reference sets of human proteins. Shown is the distribution of Hamming distance between all SARS-CoV-2 linear 8-mer peptides (n=9,926 peptides) and all linear 8-mer peptides from canonical isoforms of human cardiac proteins (orange, n=129,415 peptides), brain proteins (blue, n=241,046 peptides), and skin proteins (green, n=88,628 peptides). Error bars represent 95% confidence intervals.

Supplementary Table S1. List of human cardiac proteins identified based on Bulk RNA-seq and Single cell RNA-seq. Rows with human genes from Tables 2 and 3 are highlighted.

| Gene | Bulk RNAseq Cohen'sD | Bulk RNAseq Fold Change | Single cell RNAseq Cohen'sD | Single cell RNAseq Fold Change |
|---------------|----------------------|-------------------------|-----------------------------|--------------------------------|
| <i>NPPA</i> | 1.1 | 1796.96 | 10.81 | 228.63 |
| <i>MYL7</i> | 1.25 | 1206.31 | 13.98 | 41.82 |
| <i>MYH6</i> | 1.39 | 1035.51 | 14.39 | 33.95 |
| <i>NPPB</i> | 0.77 | 978.15 | 7.39 | 3.08 |
| <i>TNNI3</i> | 2.01 | 688.88 | 13.9 | 34.44 |
| <i>MYL2</i> | 1.06 | 426.75 | 14.8 | 70.24 |
| <i>MYH7</i> | 1.29 | 403.97 | 20.66 | 32.61 |
| <i>MYBPC3</i> | 2.56 | 398.67 | 14.13 | 7.9 |
| <i>TNNT2</i> | 3.08 | 349.01 | 15.97 | 40.17 |
| <i>CSRP3</i> | 1.56 | 342.89 | 17.53 | 12.22 |
| <i>MB</i> | 2.6 | 339.57 | 10.75 | 16.91 |
| <i>NMRK2</i> | 2 | 312.46 | 11.29 | 3.98 |
| <i>MYL3</i> | 1.24 | 287.28 | 12.33 | 10.69 |
| <i>MYL4</i> | 1.21 | 278.12 | 4.43 | 8.92 |
| <i>COX6A2</i> | 2.56 | 238.14 | 13.13 | 16.49 |
| <i>ANKRD1</i> | 1.55 | 236.39 | 16.81 | 38.11 |
| <i>BMP10</i> | 0.84 | 192.77 | 0.02 | 1 |
| <i>XIRP1</i> | 0.98 | 186.09 | 11.67 | 3.3 |
| <i>TCAP</i> | 2.47 | 165.66 | 11.75 | 9.32 |
| <i>LMOD2</i> | 1.55 | 161.06 | 15.38 | 5.41 |
| <i>CKM</i> | 2.67 | 121.39 | 14.76 | 17.56 |
| <i>TNNC1</i> | 2.11 | 116.52 | 9.95 | 12.41 |
| <i>ACTC1</i> | 2.03 | 114.83 | 17.17 | 31.65 |
| <i>TECRL</i> | 2.08 | 108.63 | 15.24 | 6.41 |
| <i>NRAP</i> | 1.64 | 100.1 | 9.86 | 2.67 |
| <i>MYOZ2</i> | 2.46 | 96.37 | 16.1 | 8.11 |

| | | | | |
|----------------|------|-------|-------|-------|
| <i>ACTA1</i> | 0.78 | 85.4 | 7.79 | 11.48 |
| <i>NKX2-5</i> | 2.55 | 60.66 | 2.68 | 1.24 |
| <i>SMPX</i> | 2.41 | 58.97 | 11.02 | 3.39 |
| <i>ANKRD2</i> | 0.74 | 56.65 | 3.14 | 1.48 |
| <i>MYOM2</i> | 2.04 | 54.35 | 3.28 | 2.9 |
| <i>SYNPO2L</i> | 1.78 | 54.28 | 7.88 | 3.21 |
| <i>FABP3</i> | 2.6 | 53.66 | 4.9 | 7.11 |
| <i>DHRS7C</i> | 1.48 | 53.13 | 3.91 | 1.26 |
| <i>HRC</i> | 2.31 | 52.1 | 9.54 | 5.25 |
| <i>ACTN2</i> | 2.29 | 49.24 | 7.42 | 6.34 |
| <i>CKMT2</i> | 2.15 | 48.13 | 5.57 | 3.19 |
| <i>TRIM63</i> | 2.11 | 46.3 | 3.35 | 1.75 |
| <i>SLN</i> | 1.03 | 45.97 | 4.25 | 3.84 |
| <i>HSPB3</i> | 2.56 | 44.41 | 6.57 | 2.16 |
| <i>FBXO40</i> | 2.08 | 38.61 | 8.3 | 2.2 |
| <i>APOBEC2</i> | 1.83 | 38.37 | 2.02 | 1.33 |
| <i>SRL</i> | 2.27 | 37.54 | 5.02 | 1.85 |
| <i>MYOM3</i> | 2.14 | 36.29 | 3.86 | 1.64 |
| <i>UNC45B</i> | 1.97 | 35.35 | 4.96 | 1.59 |
| <i>TTN</i> | 2.41 | 34.78 | 10.85 | 65.81 |
| <i>SBK2</i> | 1.12 | 34.73 | 0.08 | 1 |
| <i>MYLK3</i> | 2.08 | 30.66 | 3.38 | 1.59 |
| <i>RPL3L</i> | 1.4 | 30.36 | 3.51 | 1.44 |
| <i>TRIM54</i> | 2.43 | 30.03 | 1.55 | 1.25 |
| <i>MYBPHL</i> | 1.01 | 29.58 | 3.8 | 1.64 |
| <i>XIRP2</i> | 0.54 | 28.03 | 12.68 | 6.12 |
| <i>CMYA5</i> | 2.1 | 26.54 | 7.28 | 11.68 |
| <i>SMYD1</i> | 1.48 | 25.33 | 6.66 | 1.58 |
| <i>FITM1</i> | 2.47 | 23.91 | 4.26 | 1.96 |

| | | | | |
|----------------|-------------|--------------|-------------|-------------|
| SCN5A | 2.2 | 23.56 | 2.62 | 1.26 |
| TRDN | 2.83 | 21.15 | 7 | 4.6 |
| LMOD3 | 1.74 | 20.88 | 2.32 | 1.78 |
| MYPN | 1.76 | 20.84 | 3.81 | 1.39 |
| TRIM55 | 1.08 | 20.55 | 2.44 | 1.4 |
| LDB3 | 2.37 | 19.06 | 7.97 | 6.87 |
| SMCO1 | 1.45 | 18.71 | 4.26 | 1.25 |
| KLHL31 | 1.73 | 18.44 | 5.15 | 1.73 |
| LRRC14B | 1.22 | 18.39 | 0.07 | 1 |
| PXDNL | 2.01 | 17.72 | 0.72 | 1.13 |
| POPDC2 | 2.32 | 16.72 | 3.57 | 1.89 |
| PPP1R3A | 1.62 | 16.72 | 5.73 | 1.46 |
| ABRA | 0.89 | 15.75 | 5.75 | 1.63 |
| MYO18B | 1.96 | 15.7 | 3.25 | 1.38 |
| HHATL | 2.3 | 15.55 | 0.82 | 1.2 |
| LRRC10 | 1.52 | 15.46 | 3.59 | 1.13 |
| MYH7B | 2.1 | 15.44 | 1.42 | 1.29 |
| MYOM1 | 1.84 | 15.35 | 7.65 | 5.61 |
| SGCG | 2.28 | 15.18 | 2.37 | 1.45 |
| ENO3 | 1.84 | 14.26 | 4.15 | 2.92 |
| RYR2 | 1.81 | 14.09 | 4.02 | 5.77 |
| TBX20 | 1.44 | 13.73 | 0.06 | 1 |
| HSPB7 | 1.88 | 13.18 | 9.66 | 8.07 |
| ADPRHL1 | 2.24 | 13.08 | 1.81 | 1.49 |
| ALPK2 | 2.22 | 12.49 | 2.64 | 1.44 |
| ASB11 | 1.67 | 12.1 | 0.22 | 1.02 |
| CASQ2 | 1.93 | 11.98 | 7.04 | 3.65 |
| TMEM182 | 1.95 | 11.83 | 1.64 | 1.45 |
| PERM1 | 1.79 | 11.75 | 0.77 | 1.06 |

| | | | | |
|-----------------|------|-------|------|-------|
| <i>ASB15</i> | 1.63 | 11.34 | 1.24 | 1.18 |
| <i>TXLNB</i> | 1.7 | 11.25 | 2.56 | 1.74 |
| <i>MYZAP</i> | 1.97 | 11.15 | 3.56 | 2.28 |
| <i>ITGB1BP2</i> | 2.1 | 11.09 | 2.37 | 1.25 |
| <i>RD3L</i> | 1.99 | 10.81 | 1.21 | 1.1 |
| <i>LRRC39</i> | 1.92 | 10.76 | 1.1 | 1.23 |
| <i>SLC25A4</i> | 2.33 | 10.73 | 3.77 | 9.45 |
| <i>NEBL</i> | 2.26 | 10.45 | 3.36 | 6.36 |
| <i>FAM155B</i> | 1.6 | 10.34 | 1.02 | 1.11 |
| <i>HJV</i> | 1.58 | 10.24 | 1.84 | 1.19 |
| <i>KLHL41</i> | 1.15 | 10.21 | 1.67 | 1.27 |
| <i>ASB10</i> | 1.9 | 9.93 | 1.16 | 1.05 |
| <i>TBX5</i> | 1.46 | 9.89 | 3.04 | 1.45 |
| <i>METTL7B</i> | 0.96 | 9.85 | 0.15 | 1.02 |
| <i>RBM24</i> | 1.85 | 9.81 | 2.26 | 1.57 |
| <i>PKP2</i> | 1.65 | 9.56 | 1.59 | 1.89 |
| <i>KLHL38</i> | 1.2 | 9.49 | 0.06 | 1 |
| <i>SBK3</i> | 1.04 | 9.31 | 0.26 | 1.01 |
| <i>CORIN</i> | 1.46 | 9.19 | 1.59 | 1.26 |
| <i>CAVIN4</i> | 1.48 | 9.16 | 2.79 | 1.27 |
| <i>CAV3</i> | 2.26 | 8.63 | 0.9 | 1.07 |
| <i>COX7A1</i> | 2.49 | 8.45 | 7.43 | 50.55 |
| <i>SLC4A3</i> | 2.3 | 8.27 | 2.78 | 1.82 |
| <i>GATA4</i> | 1.94 | 7.67 | 0.6 | 1.14 |
| <i>PEBP4</i> | 1.85 | 7.67 | 0.7 | 1.25 |
| <i>MLIP</i> | 1.43 | 7.14 | 2.02 | 1.65 |
| <i>DOK7</i> | 1.53 | 7.13 | 0.23 | 1.03 |
| <i>RBM20</i> | 2.03 | 6.94 | 2 | 1.48 |
| <i>EEF1A2</i> | 1.87 | 6.78 | 1.28 | 1.35 |

| | | | | |
|------------------|------|------|-------|-------|
| <i>CHRNE</i> | 0.92 | 6.65 | 0.03 | 0.99 |
| <i>ATP2A2</i> | 1.73 | 6.64 | 2.08 | 4.52 |
| <i>PLA2G5</i> | 1.82 | 6.58 | 1.04 | 1.54 |
| <i>PLPP7</i> | 2.31 | 6.2 | -0.05 | 0.99 |
| <i>FSD2</i> | 1.74 | 6.11 | 1.56 | 1.23 |
| <i>FLNC</i> | 0.67 | 6.11 | 7.38 | 6.69 |
| <i>CRYAB</i> | 1.3 | 5.93 | 2.47 | 5.66 |
| <i>S100A1</i> | 2.2 | 5.86 | 0.1 | 0.8 |
| <i>TENM2</i> | 1.36 | 5.86 | 0.4 | 1.08 |
| <i>TNNI1</i> | 0.64 | 5.85 | 1.55 | 1.27 |
| <i>PLN</i> | 1.4 | 5.77 | 7.23 | 8.69 |
| <i>CDH2</i> | 1.68 | 5.73 | 1.74 | 2.01 |
| <i>FHL2</i> | 1.04 | 5.7 | 0.48 | 1.35 |
| <i>CYP2J2</i> | 1.6 | 5.67 | 0.52 | 1.17 |
| <i>FGF12</i> | 1.65 | 5.5 | 1.48 | 1.91 |
| <i>CCDC141</i> | 1.6 | 5.48 | 1.07 | 1.36 |
| <i>TNNI3K</i> | 1.71 | 5.47 | 0.31 | 1.05 |
| <i>SLC5A1</i> | 0.89 | 5.47 | 0.88 | 1.24 |
| <i>TNNT1</i> | 0.74 | 5.31 | 0.63 | 1.4 |
| <i>FHOD3</i> | 1.52 | 5.29 | 0.23 | 1.05 |
| <i>CLIC5</i> | 1.62 | 5.1 | 1.58 | 1.74 |
| <i>TPM1</i> | 1.78 | 5.06 | 3.71 | 17.66 |
| <i>LRRC2</i> | 1.07 | 5.05 | 1.53 | 1.51 |
| <i>GOT1</i> | 1.8 | 5.03 | 1.21 | 1.88 |
| <i>DES</i> | 1.2 | 3.88 | 7.99 | 8.88 |
| <i>NEXN</i> | 1.05 | 2.92 | 4.73 | 7.42 |
| <i>MTRNR2L8</i> | 0.27 | 1.32 | 3.56 | 24.52 |
| <i>MTRNR2L1</i> | 0.31 | 1.2 | 3.98 | 16.82 |
| <i>MTRNR2L10</i> | 0.45 | 1.12 | 4.58 | 15.78 |

| | | | | |
|-----------------|-------|------|------|------|
| <i>MTRNR2L3</i> | -0.08 | 1 | 6.99 | 7.85 |
| <i>MSRB3</i> | -0.35 | 0.65 | 4.82 | 8.38 |