

Multifunctional CRISPR/Cas9 with engineered immunosilenced human T cell epitopes

Shayesteh R. Ferdosi^{1,2,a}, Radwa Ewaisha^{1,3,a}, Farzaneh Moghadam^{4,b}, Sri Krishna^{1,4,b}, Jin G. Park¹, Mo R. Ebrahimkhani^{4,5}, Samira Kiani^{4,c}, Karen S. Anderson^{1,3,c}

¹Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, Tempe, AZ 85287.

²School of Molecular Sciences, Arizona State University, Tempe, AZ 85287.

³School of Life Sciences, Arizona State University, Tempe, AZ 85287

⁴School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ 85287.

⁵Division of Gastroenterology and Hepatology, Mayo Clinic, Phoenix, AZ.

^aThe first two authors contributed equally.

^bThe second two authors contributed equally.

^cCorrespondence should be addressed to K.S.A. (Karen.Anderson.1@asu.edu) and S.K. (samira.kiani@asu.edu).

Abstract

The application of Cas9 for genetic and epigenetic therapies in humans raises concerns over immunogenicity of this foreign protein. We report pre-existing human CD8+ T cell immunity to *Streptococcus pyogenes* Cas9 in the majority of healthy individuals screened. In a proof-of-principle study, we demonstrate that Cas9 protein can be modified to eliminate immunodominant epitopes through targeted mutation while preserving its function and specificity.

Introduction

The Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)/Cas9 technology has raised hopes for developing personalized gene therapies for complex diseases such as cancer as well as genetic disorders, and is currently entering clinical trials^{1,2}. The history of gene therapy has included both impressive success stories and serious immunologic adverse events³⁻⁸. The expression of *Streptococcus pyogenes* Cas9 protein (SpCas9) in mice has evoked both cellular and humoral immune responses^{9,10}, which raises concerns regarding its safety and efficacy as a gene or epi-gene therapy in humans. These pre-clinical models and host immune reactions to other exogenous gene delivery systems¹¹⁻¹³ suggest that the pathogenic “non-self” origin of Cas9 may be immunogenic in humans.

Both B cell and T cell host responses specific to either the transgene or the viral components of adenoviral^{14,15} and adeno-associated viral (AAV)^{11,12} vectors have been detected, despite relatively low immunogenicity of AAV vectors. In the case of AAV, specific neutralizing antibodies (Abs) and T cells are frequently detected in healthy donors¹⁶⁻¹⁹ and specific CD8+ T cells have been shown to expand following gene delivery¹⁸. There has been recent progress in developing strategies to overcome this problem, such as capsid engineering and transient immunosuppression²⁰⁻²². The potential consequences of immune responses to expressed proteins from viral vectors or transgenes include neutralization of the gene product; destruction of the cells expressing it, leading to loss of therapeutic activity or tissue destruction; induction of immune memory that prevents re-administration; and fulminant innate inflammatory responses^{23,24}. More potent immune responses to gene therapies have been observed in humans and non-human primate models compared to mice^{8,25}.

Of the Cas9 orthologs derived from bacterial species, the SpCas9 is the best characterized. *S. pyogenes* is a ubiquitous pathogen, with an annual incidence of 700 million worldwide²⁶, but immunity to SpCas9 in humans has not been reported. Here, we sought to characterize the pre-existing immune response to SpCas9 in healthy individuals and to identify the immunodominant T cell epitopes with the aim of developing SpCas9 proteins that have diminished capacity to invoke human adaptive response.

CRISPR application for human therapies will span its use both for gene editing (through DNA double-strand breaks) or epigenetic therapies (without DNA double-strand breaks). In fact, recent reports shed light on CRISPR's ability to activate or repress gene expression in mice²⁷⁻²⁹, which opens the door to a variety of new therapeutic applications such as activating silent genes, compensating for disrupted genes, cell fate reprogramming, or silencing disrupted genes, without the concern over permanent change in DNA sequence. However, unlike the use of Cas9 for gene editing, which may only require Cas9 presence in cells for a few hours, current techniques for CRISPR-based epigenetic therapies require longer term expression of Cas9 *in vivo*, possibly for weeks and months^{28,29}, which poses the challenge of combating pre-existing immune response towards Cas9. This challenge will need to be addressed before CRISPR application for human therapies, especially for epigenetic therapies, can be fully implemented. Delivery of CRISPR *in vivo* by incorporating its expression cassette in adeno-associated virus (AAV), will most likely shape many of the initial clinical trials as AAV-based gene delivery is one of the safest and most prevalent forms of gene therapies in human. AAV will enable longer term expression of Cas9, desirable for epigenetic therapies. Therefore, unlike Cas9 delivery in the form of ribonucleoprotein complexes (which are short term), it is highly likely that CRISPR delivery through AAV and its expression within target cells will engage CD8+ T cell immunity.

Results

We first determined whether healthy individuals have detectable IgG Abs to SpCas9. Of 143 healthy control sera screened, 70 (49.0%) had detectable Abs against *S. pyogenes* lysate using ELISA (**Fig. 1A**). This positive subset along with sera that were borderline negative for Abs to *S. pyogenes* lysate were screened for Abs against recombinant SpCas9, of which 36.6% were positive. At least 21.0% (n=30) of healthy individuals in this study had Cas9-specific Abs (**Fig. 1A**).

Whether Cas9-specific antibodies impact the efficacy or safety of CRISPR application in human remains to be seen. However, cellular immunity is expected to have a more significant impact given the current CRISPR/Cas9 delivery approaches. The gene encoding Cas9 is usually delivered by a vector such as viral vectors to target cells and for intracellular expression, which could evoke a cellular immune response. We thus focused on investigating the T cell immune response against SpCas9. We predicted HLA-A*02:01-restricted T cell epitopes derived from SpCas9 using a model that uses both MHC binding affinity and biochemical properties of immunogenicity³⁰ (**Supplementary Table 1**; the top 5 are shown in **Fig. 1B**). This model incorporates T cell receptor contact residue hydrophobicity and HLA binding prediction, which enhances the efficiency of epitope identification, as we previously reported³⁰. We plotted the calculated normalized binding (S_b) and immunogenicity (S_i) scores for each peptide (**Fig. 1C**) to predict the more immunogenic epitopes, which are expected to have both high HLA binding (low S_b) and more hydrophobicity (high S_i). We chose HLA-A*02:01 because it is the most common HLA type in European/North American Caucasians. We also predicted MHC class II binding epitopes for the SpCas9 protein to HLA-DRB1 (10 alleles), HLA-DQ (5 alleles), and HLA-DP (8 alleles) using the IEDB analysis tool (**Supplementary Table 2**).

We then investigated whether peripheral blood mononuclear cells (PBMCs) derived from healthy individuals had measurable T cell reactivity against the predicted SpCas9 MHC class I

epitopes. We synthesized 38 peptides (**Supplementary Table 1**) and grouped them into 10 pools of 3-4 peptides each. We measured peptide-specific T cell immunity using IFN- γ secretion ELISpot assays with PBMCs derived from 12 healthy individuals (HLA-A*02:01, n=10; non-HLA-A*02:01, n=2) and identified immunoreactive epitopes within pools 3 or 5 in 83.0% of the donors tested (90% of the HLA-A*02:01 donors; **Fig. 1D**). The seven individual peptides from pools 3 and 5 were evaluated by IFN- γ ELISpot and the dominant immunogenic epitopes were SpCas9_240-248 and SpCas9_615-623, designated peptides α and β , from pools 5 and 3, respectively. The subdominant epitopes were found to be γ and δ from pools 3 and 5, respectively. Both peptides α and β are located in the REC lobe of the Cas9 protein (**Fig. 1E**) that binds the sgRNA and the target DNA heteroduplex³¹. The individual peptides within pools that were positive for any donor were evaluated for this donor by IFN- γ ELISpot. The immunoreactivity and position of the 38 predicted peptides (a few of which are overlapping) within the Cas9 protein are shown in **Fig. 1E**.

Peptides α and β are shown as red dots on the epitope prediction plot (**Fig. 1C**) and their sequences and predicted ranking are shown in **Fig. 1B** and **Supplementary Table 1**. As predicted, these peptides had low S_b and high S_i values. Both the immunodominant (α and β) and subdominant (γ and δ) T cell epitopes identified by IFN- γ ELISpot were within the top 5 most immunogenic epitopes predicted by our immunogenicity model³⁰. Their ranking as predicted by the consensus method hosted on the IEDB server using default settings was 14, 5, 18, and 4, respectively. For MHC class II, epitope α is predicted to be a top binder to HLA-DRB1*01:02 and epitope β a top binder to HLA-DPA1*01:03 and DPB1*02:01 (**Supplementary Table 2**). Sequence similarity of peptides α and β to amino acid sequences in known proteins was investigated using Protein BLAST and the IEDB epitope database³². This was done to investigate whether there is any chance that the T cell immune response that we are detecting in healthy individuals could be due to previous exposure to another protein of similar sequence. A peptide was considered 'similar' to α or β if no more than 2 of 9 amino acid residues (that are not the second or ninth) were not matching (78% similarity). None of these two peptides resembled known epitopes in the IEDB database, but similarity to other Cas9 orthologs and other bacterial proteins was detected (**Supplementary Tables 3 and 4**). Epitope β has sequence similarity to a peptide derived from the *Neisseria meningitidis* peptide chain release factor 2 protein (ILEDIVLTL versus ILEGIVLTL). Antigen-specific T cells were expanded for 18 days *in vitro* by coculturing healthy donor PBMCs with peptide β -pulsed autologous antigen presenting cells (APCs). Cas9-specific CD8+ T cell responses were assessed by flow cytometry. CD8+ T cells specific for the HLA-A*0201/ β pentamer were detected after stimulation (3.09%; **Fig. 2A**).

We next hypothesized that mutation of the MHC-binding anchor residues of the identified immunogenic epitopes would abolish specific T cell recognition (**Fig. 2A**). The epitope anchor residues (2nd and 9th) are not only necessary for peptide binding to the MHC groove, but are also crucial for recognition by the T cell receptor³⁰. The percentage of CD8+ β pentamer+ T cells decreased to 0.3% when APCs were pulsed with the mutated peptide (β 2; **Fig. 2B**) compared with 3.09% with the wild type peptide (β ; **Fig. 2A**). We then examined the reactivity of healthy donor T cells to modified peptides α or β with mutations in residues 2, 9, or both (sequences are shown in **Fig. 2C**) using IFN- γ ELISpot assay. The epitope-specific T cell reactivity was markedly reduced with the mutant peptides (**Fig. 2D, Supplementary Fig. 1**). The average reduction for the responsive HLA-A*02:01 donors was 25-fold from α to α 29 (n=7, p<0.03) and 30-fold from β to β 29 (n=8; p<0.03; **Supplementary Fig. 1**). The predicted binding affinity to MHC class II was also decreased for α 2 and β 2 epitopes, although the experimental significance of this alteration is unknown.

We then generated modified Cas9 constructs by mutating the second residue of peptide α (L241G; Cas9- α 2), peptide β (L616G; Cas9- β 2), or both (Cas9- α 2 β 2). To measure the effect of mutating the anchor residue of the immunogenic epitopes on T cell recognition of the Cas9 protein, we transiently transfected healthy donor B cell APCs with mRNA encoding wild type Cas9 (WT-Cas9), Cas9- α 2, Cas9- β 2, or Cas9- α 2 β 2. Protein expression was confirmed by Western blot and the levels were comparable for all four constructs (data not shown). The T cell response measured by IFN- γ ELISpot after coculturing of transfected APCs with autologous PBMCs was significantly decreased for the modified Cas9 proteins (**Fig. 2E**). Introduction of the β 2 mutation was the most effective in reducing T cell immunogenicity (5.5-fold, $p < 0.0001$). This mutation in the REC1 domain (**Fig. 1E and 3A**) is not located in any of the two regions that are absolutely essential for DNA cleavage, the repeat-interacting (97–150) and the anti-repeat-interacting (312–409) regions³¹. These results demonstrate that mutating the anchor amino acid of a highly immunogenic epitope can influence the overall immunogenicity of Cas9. Thus, engineering Cas9 variants with reduced immunogenicity potential can be used in conjunction with other strategies for safer CRISPR therapies and even possibly reduce the dosage of systemic immunosuppression needed for patients.

We then tested the function of Cas9- β 2 in comparison with WT-Cas9 in the context of DNA cleavage and transcriptional modulation. To examine the nuclease activity of Cas9- β 2 and compare with WT-Cas9, we targeted Cas9- β 2 or WT-Cas9 to an endogenous locus (*EMX-1*) and measured percent indel formation (**Fig. 3B, C**). Our data demonstrate that Cas9- β 2 retains nuclease capacity in the locus we studied as well as on a synthetic promoter (**Fig. 3C, Supplementary Fig. 2A**).

Next, we determined whether Cas9- β 2 can successfully recognize and bind its target DNA leading to transcriptional modulation. We first tested this in the context of enhanced transgene expression from a synthetic CRISPR responsive promoter in HEK293 cells using 14nt gRNAs and aptamer-mediated recruitment of transcriptional modulators similar to what we had shown before (**Supplementary Fig. 2B**). Having shown successful transgene activation, we then investigated whether this variant retains such capacity within the chromosomal contexts of endogenous genes. We transfected the cells with plasmids encoding Cas9- β 2 or WT-Cas9 and 14nt gRNAs against two different endogenous genes (*TTN* and *MIAT*). qRT-PCR analysis showed that this variant successfully led to target gene expression (**Fig. 3D-F**). To further characterize Cas9- β 2 specificity, we performed genome-wide RNA sequencing after targeting Cas9- β 2 or WT-Cas9 to the *MIAT* locus for transcriptional activation. The results demonstrated no significant increase in undesired off-target activity by Cas9- β 2 as compared to WT-Cas9 (**Fig. 3G**).

To show the extensibility of our approach, we tested the function of Cas9- α 2, that has a mutation located in the REC2 domain (**Fig. 1E, 3A**). Cas9- α 2 also demonstrated DNA cleavage and transcriptional modulation functionality comparable with WT-Cas9 (**Supplementary Fig. 3A-E**). This is consistent with a previous study which showed that Cas9 with a deleted REC2 domain retains its nuclease activity³¹. When T cells were stimulated with APCs spiked with peptide α 2, the percentage of CD8⁺ CD137⁺ T cells (a marker of T cell activation³³) was decreased by 2.3-fold as compared to WT peptide α stimulation (**Supplementary Fig. 3F**).

Discussion

The detection of pre-existing B cell and T cell immunity to the most widely used nuclease ortholog of the CRISPR/Cas9 tool in a significant proportion of healthy humans confirms previous studies in mice^{9,10} and sheds light on the need for more studies of the immunological

risks of this system. The CD8+ T cell immunity we observed is likely memory responses, as they are observed without ex vivo stimulation. However, following 18 days of T cell stimulation by peptides α or β , expansion of naïve T cells is not precluded. This suggests that the expression of Cas9 in naïve individuals may trigger a T cell response that could prevent subsequent administration. This could be avoided by switching to Cas9 orthologs from other bacterial species, but attention needs to be given to individual and distinct immune repertoires. This can be difficult given the epitope conservation across Cas9 proteins from multiple *Streptococcus* species and resemblance to sequences from other bacterial proteins such as the common pathogen *N. meningitidis*, that asymptotically colonizes the nasopharynx in 10% of the population³⁴. Therefore, selective deimmunization (immunosilencing) of Cas9 can represent an attractive alternative, particularly in patients where high-dose systemic immunosuppression is contraindicated, such as in patients with chronic infectious diseases. This strategy can be important in particular when longer term expression of Cas9 will be desired.

Conventional methods of deimmunizing non-human therapeutic proteins rely on trial-and-error mutagenesis or machine learning and often includes deletion of whole regions of the protein³⁵⁻³⁹. Here, as a general principle, we show that alteration of one of the anchor residues of an immunodominant epitope abolished specific T cell recognition. However, HLA allotype diversity and the existence of numerous epitopes in the large Cas9 protein complicate the process of complete deimmunization. The overall impact of removal of select immunodominant epitopes remains to be seen; both reduction⁴⁰ and enhancement⁴¹ of the immunogenicity of subdominant epitopes have been reported with similar approaches for other proteins.

The top binding T cell epitopes within Cas9 that are most promiscuous for common HLA class I and class II alleles have been recently predicted *in silico* using IEDB⁴². However, this is the first study that experimentally validates predicted immunodominant epitopes. None of the epitopes we report overlap with the peptides previously predicted⁴². This is not unsurprising since we restricted our analysis to one HLA haplotype. Additionally, improved algorithms are needed to predict epitopes that hold up in experimental validation, as we show here. The use of CRISPR/Cas9 in humans may eventually necessitate creating HLA type-specific Cas9 variants, particularly for applications that require long-term Cas9 expression.

Non-specific localized immune suppressive approaches, such as those used by tumor cells and some viruses may complement these strategies for complete deimmunization. One attractive strategy is the transient and inducible co-expression of programmed death-ligand 1 (PD-L1) or indoleamine 2,3-dioxygenase 1 (IDO1) activating gRNAs inside cells that express Cas9 to protect them against attack by T cells. Alternatively, antigen presentation can be blocked by viral proteins interfering with antigen presentation (VIPRs), such as the adenoviral E319K or US2 and US11 from the human cytomegalovirus⁴³ or molecules that inhibit proteasomal antigen processing such as the Epstein-Barr virus Gly-Ala repeat⁴⁴. Deimmunized Cas9 may be useful in reduction of the dosage of other immunomodulatory measures needed to be co-administered in patients, thus facilitating therapeutic CRISPR applications as we develop better understanding of the immunological consequences of this system.

Acknowledgements

The authors thank Dr. Erich Sturgis for providing the serum samples and Dr. Diego Chowell for support with the computational prediction model. We thank all the other members of the Anderson, Kiani, and Ebrahimkhani labs for their assistance and insightful discussions. This work was supported by ASU institutional funds and the startup fund by the School of Biological

and Health Systems Engineering of ASU. We thank the Center for Computational and Integrative Biology (CCIB) DNA core facility at Massachusetts General Hospital and Genomics and Bioinformatics (TCGB) core at UCLA for the DNA and RNA sequencing services.

Author Contributions

S.R.F. and R.E. designed experiments, performed experiments, and analyzed data. F.M. generated Cas9 variant constructs, performed Cas9 functional analysis experiments, and analyzed data. S. Krishna generated predicted Cas9 epitopes and designed and assisted with the T cell experiments. J.G.P. analyzed RNA seq data. M.R.E. helped with the design of experiments and interpretation of data. S.K. and K.S.A. supervised this study. S.R.F., R.E., S.K., and K.S.A. wrote the manuscript with the support of all the other authors.

Competing Financial Interests

Arizona State University has submitted a patent application, with authors SRF, RE, FM, SK, MRE, SK and KSA. JGP declares no competing financial interest.

References

- 1 Cyranoski, D. Chinese scientists to pioneer first human CRISPR trial. *Nature News* **535**, 476 (2016).
- 2 Reardon, S. First CRISPR clinical trial gets green light from US panel. *Nature* **531** (2016).
- 3 Cavazzana-Calvo, M. *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669-672 (2000).
- 4 Gaspar, H. B. *et al.* Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* **364**, 2181-2187, doi:10.1016/S0140-6736(04)17590-9 (2004).
- 5 Howe, S. J. *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* **118**, 3143-3150, doi:10.1172/JCI35798 (2008).
- 6 Marshall, E. Gene Therapy Death Prompts Review of Adenovirus Vector. *Science* **286**, 2244-2245, doi:10.1126/science.286.5448.2244 (1999).
- 7 Hacein-Bey-Abina, S. *et al.* Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N Engl J Med* **346**, 1185-1193, doi:10.1056/NEJMoa012616 (2002).
- 8 Manno, C. S. *et al.* Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med* **12**, 342-347, doi:10.1038/nm1358 (2006).
- 9 Chew, W. L. *et al.* A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat Methods* **13**, 868-874, doi:10.1038/nmeth.3993 (2016).
- 10 Wang, D. *et al.* Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR/Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. *Hum Gene Ther* **26**, 432-442, doi:10.1089/hum.2015.087 (2015).
- 11 Mays, L. E. & Wilson, J. M. The complex and evolving story of T cell activation to AAV vector-encoded transgene products. *Mol Ther* **19**, 16-27, doi:10.1038/mt.2010.250 (2011).
- 12 Mingozi, F. & High, K. A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **122**, 23-36, doi:10.1182/blood-2013-01-306647 (2013).
- 13 Yin, H. *et al.* Non-viral vectors for gene-based therapy. *Nat Rev Genet* **15**, 541-555, doi:10.1038/nrg3763 (2014).
- 14 Ahi, Y. S., Bangari, D. S. & Mittal, S. K. Adenoviral Vector Immunity: Its Implications and circumvention strategies. *Current gene therapy* **11**, 307-320 (2011).
- 15 Aldhamen, Y. A. & Amalfitano, A. in *Adenoviral Vectors for Gene Therapy (Second Edition)* 391-422 (Academic Press, 2016).
- 16 Thwaite, R., Pages, G., Chillon, M. & Bosch, A. AAVrh.10 immunogenicity in mice and humans. Relevance of antibody cross-reactivity in human gene therapy. *Gene Ther* **22**, 196-201, doi:10.1038/gt.2014.103 (2015).
- 17 Boutin, S. *et al.* Prevalence of serum IgG and neutralizing factors against adeno-associated virus (AAV) types 1, 2, 5, 6, 8, and 9 in the healthy population: implications for gene therapy using AAV vectors. *Hum Gene Ther* **21**, 704-712, doi:10.1089/hum.2009.182 (2010).
- 18 Mingozi, F. *et al.* CD8(+) T-cell responses to adeno-associated virus capsid in humans. *Nat Med* **13**, 419-422, doi:10.1038/nm1549 (2007).
- 19 Scallan, C. D. *et al.* Human immunoglobulin inhibits liver transduction by AAV vectors at low AAV2 neutralizing titers in SCID mice. *Blood* **107**, 1810-1817, doi:10.1182/blood-2005-08-3229 (2006).

- 20 Bartel, M., Schaffer, D. & Buning, H. Enhancing the Clinical Potential of AAV Vectors by Capsid Engineering to Evade Pre-Existing Immunity. *Front Microbiol* **2**, 204, doi:10.3389/fmicb.2011.00204 (2011).
- 21 Martino, A. T. *et al.* Engineered AAV vector minimizes in vivo targeting of transduced hepatocytes by capsid-specific CD8+ T cells. *Blood* **121**, 2224-2233, doi:10.1182/blood-2012-10-460733 (2013).
- 22 Mingozzi, F. *et al.* Overcoming preexisting humoral immunity to AAV using capsid decoys. *Sci Transl Med* **5**, 194ra192, doi:10.1126/scitranslmed.3005795 (2013).
- 23 Jiang, H. *et al.* Effects of transient immunosuppression on adenoassociated, virus-mediated, liver-directed gene transfer in rhesus macaques and implications for human gene therapy. *Blood* **108**, 3321-3328, doi:10.1182/blood-2006-04-017913 (2006).
- 24 Kay, M. A. State-of-the-art gene-based therapies: the road ahead. *Nat Rev Genet* **12**, 316-328, doi:10.1038/nrg2971 (2011).
- 25 Gao, G. *et al.* Adeno-associated virus-mediated gene transfer to nonhuman primate liver can elicit destructive transgene-specific T cell responses. *Hum Gene Ther* **20**, 930-942, doi:10.1089/hum.2009.060 (2009).
- 26 Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. The global burden of group A streptococcal diseases. *Lancet Infect Dis* **5**, 685-694, doi:10.1016/S1473-3099(05)70267-X (2005).
- 27 Ibrahim, S. H. & Robertson, K. D. Use of the CRISPR/Cas9-Based Epigenetic Gene Activation System in Vivo: a New Potential Therapeutic Modality. *Hepatology*, doi:10.1002/hep.29860 (2018).
- 28 Liao, H. K. *et al.* In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell* **171**, 1495-1507 e1415, doi:10.1016/j.cell.2017.10.025 (2017).
- 29 Zheng, Y. *et al.* CRISPR interference-based specific and efficient gene inactivation in the brain. *Nat Neurosci* **21**, 447-454, doi:10.1038/s41593-018-0077-5 (2018).
- 30 Chowell, D. *et al.* TCR contact residue hydrophobicity is a hallmark of immunogenic CD8+ T cell epitopes. *Proc Natl Acad Sci U S A* **112**, E1754-1762, doi:10.1073/pnas.1500973112 (2015).
- 31 Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935-949, doi:10.1016/j.cell.2014.02.001 (2014).
- 32 Vita, R. *et al.* The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* **43**, D405-412, doi:10.1093/nar/gku938 (2015).
- 33 Wolfl, M. *et al.* Activation-induced expression of CD137 permits detection, isolation, and expansion of the full repertoire of CD8+ T cells responding to antigen without requiring knowledge of epitope specificities. *Blood* **110**, 201-210, doi:10.1182/blood-2006-11-056168 (2007).
- 34 Pollard, A. J. & Maiden, M. C. J. *Meningococcal Vaccines*. (Humana Press, 2001).
- 35 King, C. *et al.* Removing T-cell epitopes with computational protein design. *Proc Natl Acad Sci U S A* **111**, 8577-8582, doi:10.1073/pnas.1321126111 (2014).
- 36 Mazor, R. *et al.* Rational design of low immunogenic anti CD25 recombinant immunotoxin for T cell malignancies by elimination of T cell epitopes in PE38. *Cell Immunol* **313**, 59-66, doi:10.1016/j.cellimm.2017.01.003 (2017).
- 37 Salvat, R. S. *et al.* Computationally optimized deimmunization libraries yield highly mutated enzymes with low immunogenicity and enhanced activity. *Proc Natl Acad Sci U S A* **114**, E5085-E5093, doi:10.1073/pnas.1621233114 (2017).
- 38 Cantor, J. R. *et al.* Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift. *Proc Natl Acad Sci U S A* **108**, 1272-1277, doi:10.1073/pnas.1014739108 (2011).

- 39 Tangri, S. *et al.* Rationally engineered therapeutic proteins with reduced immunogenicity. *J Immunol* **174**, 3187-3196 (2005).
- 40 Yeung, V. P. *et al.* Elimination of an Immunodominant CD4⁺ T Cell Epitope in Human IFN- β Does Not Result in an In Vivo Response Directed at the Subdominant Epitope. *The Journal of Immunology* **172**, 6658-6665, doi:10.4049/jimmunol.172.11.6658 (2004).
- 41 Mok, H., Lee, S., Wright, D. W. & Crowe, J. E. Enhancement of the CD8⁺ T cell response to a subdominant epitope of respiratory syncytial virus by deletion of an immunodominant epitope. *Vaccine* **26**, 4775-4782, doi:<https://doi.org/10.1016/j.vaccine.2008.07.012> (2008).
- 42 Chew, W. L. Immunity to CRISPR Cas9 and Cas12a therapeutics. *Wiley Interdiscip Rev Syst Biol Med* **10**, doi:10.1002/wsbm.1408 (2018).
- 43 Yewdell, J. W. & Hill, A. B. Viral interference with antigen presentation. *Nature Immunology* **3**, 1019, doi:10.1038/ni1102-1019 (2002).
- 44 Levitskaya, J. *et al.* Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1. *Nature* **375**, 685-688, doi:10.1038/375685a0 (1995).
- 45 Anderson, K. S. *et al.* HPV16 antibodies as risk factors for oropharyngeal cancer and their association with tumor HPV and smoking status. *Oral oncology* **51**, 662-667, doi:10.1016/j.oraloncology.2015.04.011 (2015).
- 46 Wang, J. *et al.* A versatile protein microarray platform enabling antibody profiling against denatured proteins. *Proteomics Clin Appl* **7**, 378-383, doi:10.1002/prca.201200062 (2013).
- 47 Krishna, S. & Anderson, K. S. T-Cell Epitope Discovery for Therapeutic Cancer Vaccines. *Vaccine Design: Methods and Protocols: Volume 1: Vaccines for Human Diseases*, 779-796 (2016).
- 48 Moutaftsi, M. *et al.* A consensus epitope prediction approach identifies the breadth of murine TCD8⁺-cell responses to vaccinia virus. *Nature biotechnology* **24**, 817 (2006).
- 49 Hoof, I. *et al.* NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* **61**, 1 (2009).
- 50 Rammensee, H.-G., Bachmann, J., Emmerich, N. P. N., Bachor, O. A. & Stevanović, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**, 213-219 (1999).
- 51 Tenzer, S. *et al.* Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. *Cellular and Molecular Life Sciences* **62**, 1025-1037 (2005).
- 52 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
- 53 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562-578, doi:10.1038/nprot.2012.016 (2012).
- 54 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

Figure Legends

Figure 1. Detection of pre-existing B cell and T cell immune responses to SpCas9 in healthy donors and identification of two immunodominant T cell epitopes. **A.** Specific serum Abs were detected against *S. pyogenes* lysate in 49.0% (above the dotted line) of 143 healthy controls (left). The subset shown in black circles was screened for Abs against recombinant Cas9 protein (right), of which 36.6% (21.0% of total samples screened) were positive (above the dashed line). **B.** The top 5 predicted SpCas9 T cell epitopes and their predicted S_b and S_i scores and ranking (based on the $S_b \cdot S_i$ value)³⁰. These top 5 peptides include the identified immunodominant (α and β ; gray) and subdominant (γ and δ) epitopes that were shown to be immunogenic by IFN- γ ELISpot. **C.** Plot of S_b and S_i of predicted HLA-A*02:01 epitopes for the SpCas9 protein. Red dots represent the immunodominant and subdominant epitopes. **D.** IFN- γ ELISpot assay of T cell reactivity of 12 healthy donors (the two non- HLA-A*02:01 are shown as open circles) to 38 predicted epitopes grouped in 10 pools, CEF (positive control), and DMSO (negative control). Peptides α and δ were in pool 5 while β and γ were in pool 3. **E.** IFN- γ ELISpot reactivity of healthy donor T cells (n=12) to epitopes across the different domains of the Cas9 protein. Donors 1-10 were HLA-A*02:01, while 11 and 12 were not. Peptides α and δ overlap in 5 amino acid residues. S_b , normalized binding score; S_i , normalized immunogenicity score.

Figure 2. SpCas9 immunodominant epitope-specific CD8+ T cell recognition is abolished after anchor residue mutation. **A.** Epitope β -specific CD8+ T cell response detected using β -specific pentamer in PBMCs stimulated with peptide β -pulsed antigen presenting cells. **B.** The percentage of CD8+ pentamer β + T cells was reduced to 0.3% when healthy donor B cell APCs were pulsed with the mutated peptide β . **C.** Positions, sequences and IEDB HLA binding percentile rank of epitopes α and β before and after mutation of the anchor (2nd and/or 9th) residues. S_b , normalized binding score; S_i , normalized immunogenicity score. **D.** IFN- γ ELISpot assay in triplicate wells comparing T cell reactivity to wild type or mutated epitopes α and β . These results are representative of 12 donors and two independent replicates. **E.** IFN- γ ELISpot comparing T cell reactivity to APCs expressing WT or modified Cas9 proteins. APCs expressing FluM1 were used as a positive control. APCs expressing GAPDH or spiked with peptide α 2 were used as negative controls. Data represent mean +/- SEM of 5 replicates (right).

Figure 3. Mutated SpCas9 protein (Cas9- β 2) retains its function and specificity. **A.** 3D structure of the SpCas9 protein, showing the location of the identified immunodominant epitopes α and β . **B.** Schematic of the experiment assessing mutagenesis capacity of Cas9- β 2. Cells were transfected with either WT-Cas9, Cas9- β 2, or an empty plasmid as well as 20nt gRNA targeting *EMX-1* locus. 72 hrs after transfection, percent cleavage was assessed by DNA extraction and illumina sequencing. **C.** Percentage of indel formation in *EMX-1* locus. Data represent mean +/- SD of three individual transfections. **D.** Schematic of the experiment assessing gRNA binding, DNA targeting and transcriptional modulation with Cas9- β 2. Cells were transfected with either WT-Cas9, Cas9- β 2, or an empty plasmid as well as 14nt gRNA targeting *TTN* or *MIAT* in the presence of MS2-P65-HSF1 (transcriptional modulation). 72 hrs after transfection, mRNA was assessed by qRT-PCR. **E,F.** Shown is the mRNA level relative to an untransfected control experiment. Each individual dot represents an individual transfection. **G.** Mean expression levels of 24,078 protein-coding and non-coding RNA genes for WT-Cas9 and Cas9- β 2 (each in duplicate) are shown. For visualization purposes, the values were transformed to a $\log_2(\text{CPM}+1)$

scale. *MIAT*, the gRNA target gene, is highlighted in red, and R denotes Pearson correlation coefficient between two groups.

Methods

Detection of Cas9-Specific Serum Antibodies in Healthy Controls

Healthy control sera ($n = 183$) used in this study, and previously described⁴⁵, are a subset of a molecular epidemiology study of head and neck cancer at the MD Anderson Cancer Center, collected between January 2006 and September 2008. Samples were collected using a standardized sample collection protocol and stored at -80°C until use. Written informed consent was obtained from all participants under institutional review board approval. *S. pyogenes* lysate was prepared by sonication of bacterial pellets from overnight cultures of *S. pyogenes* ATCC 19615 in the presence of 1 pill of cOmplete Protease Inhibitor (Sigma-Aldrich) after 3 cycles of freezing and thawing. Serum antibody detection was performed using ELISA. 96-well plates were coated with $20\ \mu\text{g}/\text{mL}$ of recombinant *S. pyogenes* Cas9 nuclease (New England Biolabs, Ipswich, MA) or *S. pyogenes* lysate. Sera were diluted 1:50 in 10% *E. coli* lysate prepared in 5% milk-PBST (0.2% tween)⁴⁶, incubated with shaking for 2 hrs at room temperature, and added to the specified wells in duplicate. Horseradish peroxidase (HRP) anti-human IgG Abs (Jackson ImmunoResearch Laboratories, West Grove, PA) were added at 1:10,000, and detected using Supersignal ELISA Femto Chemiluminescent substrate (Thermo Fisher Scientific, Waltham, MA). Luminescence was detected as relative light units (RLU) on a Glomax 96 Microplate Luminometer (Promega, Madison, WI) at 425 nm. To establish cut-off values, a RLU ratio $>$ (the mean + 2 standard deviations) of all samples with signal below the mean RLU (horizontal black lines) was designated positive (**Fig. 1A**, dotted and dashed lines for bacterial lysate and Cas9 protein, respectively).

Cas9 candidate T cell epitope prediction

We used our previously described prediction strategies^{30,47} to predict candidate Cas9 T cell epitopes. Briefly, we predicted MHC class I restricted 9-mer and 10-mer candidate epitopes derived from the Cas9 protein (Uniprot - Q99ZW2) for HLA A*02:01. The protein reference sequence was entered into 5 different prediction algorithms; 3 MHC-binding: IEDB-consensus binding⁴⁸, NetMHCpan binding⁴⁹, Syfpeithi⁵⁰ and 2 antigen-processing algorithms: IEDB-consensus processing, ANN processing⁵¹. The individual scores from each of the prediction algorithms were then normalized within the pool of predicted peptides after exclusion of poor binders as previously detailed^{30,47}, and the average normalized binding scores were used to re-rank the candidate peptides. The top 38 candidate peptides (**Supplementary Table 1**) were selected for experimental testing. The IEDB consensus MHC-binding prediction algorithm (<http://www.iedb.org/>) was applied to obtain a list of high binding Cas9 peptides, each of which was assigned a normalized binding score (S_b). The immunogenicity score (S_i) was calculated for each peptide based on its amino acid hydrophobicity (ANN-Hydro)³⁰.

Ex vivo stimulation and epitope mapping of Cas9 by ELISpot

All peripheral blood mononuclear cells (PBMCs) were obtained from healthy individuals with written informed consent under ASU's Institutional Review Board. PBMCs were isolated from fresh heparinized blood by Ficoll–Hypaque (GE Healthcare, UK) density gradient centrifugation and stimulated as previously described⁴⁷. Briefly, predicted Cas9 peptides with $S_b < 0.148$ ($N=38$) were synthesized ($> 80\%$ purity) by Proimmune, UK. Each peptide was reconstituted at $1\text{mg}/\text{mL}$ in sterile PBS and pools were created by mixing 3-4 candidate peptides. Sterile multiscreen ELISpot plates (Merck Millipore, Billerica, MA, USA) were coated overnight with $5\mu\text{g}/\text{well}$ of anti-IFN- γ capture antibody (clone D1K, Mabtech, USA) diluted in sterile PBS. Frozen PBMCs were thawed rapidly and recombinant human IL-2 ($20\text{U}/\text{mL}$, R&D Systems) was added. They were then stimulated in triplicates with $10\mu\text{g}/\text{mL}$ Cas9 peptide pools (or individual peptides), pre-mixed CEF pool as a positive control (Proimmune, UK), or DMSO as a negative

control in the anti-IFN- γ -coated ELISpot plates, (Merck Millipore, Billerica, MA, USA) and incubated in a 37°C, 5% CO₂ incubator for 48 hrs. Plates were washed three times for 5 min each with ELISpot buffer (PBS + 0.5% FBS) and incubated with 1 μ g/mL anti-IFN- γ secondary detection antibody (clone 7-B6-1, Mabtech, USA) for 2 hrs at room temperature, washed and incubated with 1 μ g/mL Streptavidin ALP conjugate for 1 hr at room temperature. The wells were washed again with ELISpot buffer and spots were developed by incubating for 8-10 min with detection buffer (33 μ L NBT, 16.5 μ L BCIP, in 100mM Tris-HCl pH 9, 1mM MgCl₂, 150mM NaCl). Plates were left to dry for 2 days and spots were read using the AID ELISpot reader (Autoimmun Diagnostika GmbH, Germany). The average number of spot forming units for each triplicate was calculated for each test peptide or peptide pool and subtracted from the background signal.

Autologous APC generation from healthy individual PBMCs

Autologous CD40L-activated B cell APCs were generated from healthy donors by incubating whole PBMCs with irradiated (32 Gy) K562-cell line expressing human CD40L (KCD40L) at a ratio of 4:1 (800,000 PBMCs to 200,000 irradiated KCD40Ls) in each well. The cells were maintained in B cell media (BCM) consisting of IMDM (Gibco, USA), 10% heat-inactivated human serum (Gemini Bio Products, CA, USA), and Antibiotic-Antimycotic (Anti-Anti, Gibco, USA). BCM was supplemented with 10ng/mL recombinant human IL-4 (R&D Systems, MN, USA), 2 μ g/mL Cyclosporin A (Sigma-Aldrich, CA, USA), and insulin transferrin supplement (ITES, Lonza, MD, USA). APCs were re-stimulated with fresh irradiated KCD40Ls on days 5 and 10, after washing with PBS and expanding into a whole 24-well plate. After two weeks, APC purity was assessed by CD19⁺ CD86⁺ expressing cells using flow cytometry, and were used for T cell stimulation after >90% purity. APCs were either restimulated up to 4 weeks or cryopreserved for re-expansion as necessary.

T cell stimulation by autologous APCs

Antigen-specific T cells were detected by stimulating healthy donor B cell APCs by either peptide pulsing of specific Cas9 epitopes, or by transfecting with mRNA encoding the whole WT or modified Cas9 proteins. Peptide pulsing of APCs was done under BCM 5% human serum, with recombinant IL-4. Transfection of APCs was done with primary P3 buffer in a Lonza 4D Nucleofector and program EO117 (Lonza, MD, USA) and incubated in BCM-10% human serum and IL-4. Twenty-four hrs later, on day 1, APCs were washed and incubated with thawed whole PBMCs at a ratio of 1:2 (200,000 APCs : 400,000 PBMCs) in a 24-well plate in BCM supplemented with 20U/mL recombinant human IL-2 (R&D Systems, MN, USA) and 5ng/mL IL-7 (R&D Systems, MN, USA). On day 5, partial media exchange was performed by replacing half the well with fresh BCM and IL-2. On day 10, fresh APCs were peptide pulsed in a new 24-well plate. On day 11, expanded T cells were restimulated with peptide-pulsed APCs similar to day 1. T cells were used for T cell assays or immunophenotyped after day 18.

Flow cytometry staining for T cells

Cells were washed once in MACS buffer (containing PBS, 1% BSA, 0.5mM EDTA), centrifuged at 550g for 5 min and re-suspended in 200 μ L MACS buffer. Cells were stained in 100 μ L of staining buffer containing anti-CD137, conjugated with phycoerythrin (PE, clone 4B4-1; BD Biosciences, USA), anti-CD8-PC5 (clone B9.11; Beckman Coulter 1:100), anti-CD4 (clone SK3; BioLegend, 1:200), anti-CD14 (clone 63D3; BioLegend, 1:200), and anti-CD19 (clone HIB19; BioLegend, 1:200), all conjugated to Fluorescein isothiocyanate (FITC) for exclusion gates, for 30 min on ice. Samples were covered and incubated for 30 min on ice, washed twice in PBS, and resuspended in 1mL PBS prior to analysis.

Pentamer staining for T cell immunophenotyping

The following HLA-A*02:01 PE-conjugated Cas9 pentamers were obtained from ProImmune: F2A-D-CUS-A*02:01-ILEDIVLTL-Pentamer, 007-Influenza A MP 58-66-GILGFVFTL-Pentamer. T cells were washed twice in MACS buffer with 5% human serum and centrifuged at 550g for 5 min each time. They were then re-suspended in 100 μ L staining buffer (MACS buffer, with 5% human serum and 1mM Dasatanib (ThermoFisher Scientific, MA, USA). Each of the pentamers was added to resuspended T cells, stimulated with the respective peptide or APCs at a concentration of 1:100. Samples were incubated at room temperature for 30 min in the dark, then washed twice in MACS buffer. Cells were stained in 100 μ L MACS buffer with anti-CD8-PC5, anti-CD4-FITC, anti-CD14-FITC, and anti-CD19-FITC for exclusion gates. Samples were then washed twice with PBS and analyzed by flow cytometry. For flow cytometric analysis, all samples were acquired with Attune flow cytometer (ThermoFisher Scientific, MA, USA) and analyzed using the Attune software. Gates for expression of different markers and pentamers were determined based on flow minus one (FMO) samples for each color after doublet discrimination (**Supplementary Fig. 4**). Percentages from each of the gated populations were used for the analysis.

Vector Design and Construction

Modified Cas9 plasmids - Human codon-optimized *Streptococcus pyogenes* Cas9 sequence was amplified from pSpCas9 (pX330; Addgene plasmid ID: 42230), using forward and reverse primers and inserted within gateway entry vectors using golden gate reaction. Desired mutations were designed within gBlocks (Integrated DNA Technologies). The gblocks and amplicons were then cloned into entry vectors using golden gate reaction. All the primers and gblocks sequences are listed in supplementary notes. Next, the Cas9 vectors and CAG promoter cassettes were cloned into an appropriate gateway destination vector via LR reaction (Invitrogen).

U6-sgRNA-MS2 plasmids - These plasmids were constructed by inserting either 14bp or 20bp spacers of gRNAs (supplementary note) into sgRNA (MS2) cloning backbone (Addgene plasmid ID: 61424) at BbsI site. All the gRNA sequences are listed in supplementary note.

Cell culture for endogenous target mutation and activation

HEK293FT cell line was purchased from ATCC and maintained in Dulbecco's modified Eagle's medium (DMEM - Life Technologies) containing 10% fetal bovine serum (FBS - Life Technologies), 2mM glutamine, 1mM sodium pyruvate (Life Technologies) and 1% penicillin-streptomycin (Life Technologies) in incubators at 37°C and 5% CO₂. Polyethylenimine (PEI) was used to transfect HEK293FT cells seeded into 24-well plates. Transfection complexes were prepared according to manufacturer's instructions.

Fluorescent Reporter Assay for Quantifying Cas9 Function

HEK293FT cells were co-transfected with 10ng gRNA, 200ng Cas9 constructs, 100ng reporter plasmid and 25ng EBFP2 expressing plasmid as the transfection control. Fluorescent reporter experiments were performed 48 hrs after transfection. Flow cytometry data was analyzed using FlowJo. Cells were gated for positive EBFP expression to remove the un-transfected cells from the analysis (**Supplementary Fig. 4**). Untransfected controls were included in each experiment.

Quantitative RT-PCR Analysis

HEK293FT cells were co-transfected with 10ng gRNA, 200ng Cas9 constructs, 100ng MS2-P65-HSF1 (Addgene plasmid ID: 61423) and 25ng transfection control. Cells were lysed, and RNA was extracted using RNeasy Plus mini kit (Qiagen) 72 hrs post transfection, followed by cDNA synthesis using the High-Capacity RNA-to-cDNA Kit (Thermo fisher). qRT-PCR was performed using SYBR Green PCR Master Mix (Thermo fisher). All analyses were normalized to 18S rRNA (ΔCt) and fold changes were calculated against un-transfected controls ($2^{-\Delta\Delta Ct}$). Primer sequences for qPCR are listed in supplementary notes.

Endogenous Indel Analysis

HEK293FT cells were co-transfected with 200ng of Cas9 plasmids, 10ng of gRNA coding cassette and 25ng transfection control. 72 hrs later, transfected cells were dissociated and spun down at 200 g for 5 min at room temperature. Genomic DNA was extracted using 50 μ l of QuickExtract DNA extraction solution (Epicentre) according to the manufacturer's instructions. Genomic DNA was amplified by PCR using primers flanking the targeted region. Illumina Tru-Seq library was created by ligating partial adaptors and a unique barcode to the DNA samples. Next, a small number of PCR cycles was performed to complete the partial adaptors. Equal amounts of each sample were then pooled and sequenced on Illumina Tru-Seq platform with 2x150 run parameters, which yielded approximately 80,000 reads per sample. Sequencing was performed using a 2x150 paired-end (PE) configuration by CCIB DNA Core Facility at Massachusetts General Hospital (Cambridge, MA, USA). The reads were aligned to the target gene reference in *Mus musculus* genome using Geneious software, 9-1-5. To detect the indels (insertions and deletions of nucleic acid sequence at the site of double-strand break), each mutation was evaluated carefully in order to exclude the ones that are caused by sequencing error or any off-target mutation. The variant frequencies (percentage to total) assigned to each read containing indels were summed up. i.e. indel percentage = total number of indel containing reads/ total number of reads. The minimum number of analyzed reads per sample was 70,000.

RNA Sequencing for Quantifying Activator Specificity

HEK293FT cells were co-transfected with 10ng gRNA for MIAT locus, 200ng Cas9 constructs, 100ng MS2-P65-HSF1 (Addgene plasmid ID: 61423) and 25ng transfection control. Total RNA was extracted 72 hrs post transfection using RNeasy Plus mini kit (Qiagen) and sent to UCLA TCGB core on dry ice. Ribosomal RNA depletion, and single read library preparation were performed at UCLA core followed by RNA sequencing using NextSeq500. Coverage was 14 million reads per sample. FASTQ files with single-ended 75bp reads were then aligned to the human GRCh38 reference genome sequence (Ensembl release 90) with STAR⁵², and uniquely-mapped read counts (an average of 14.8 million reads per sample) were obtained with Cufflink⁵³. The read counts for each sample were then normalized for the library size to CPM (counts per million reads) with edgeR⁵⁴. Custom R scripts were then used to generate plots.

Figure 1

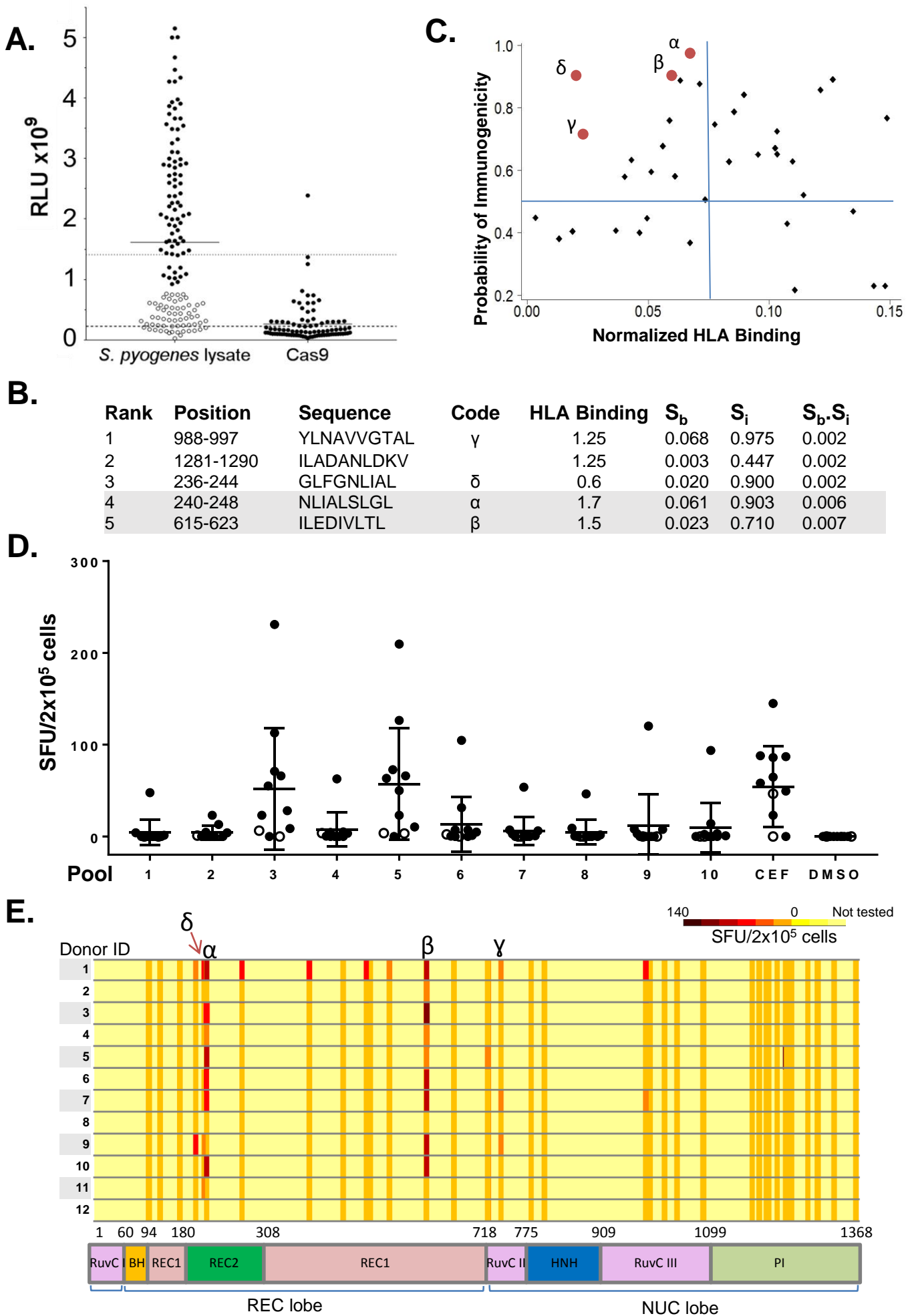
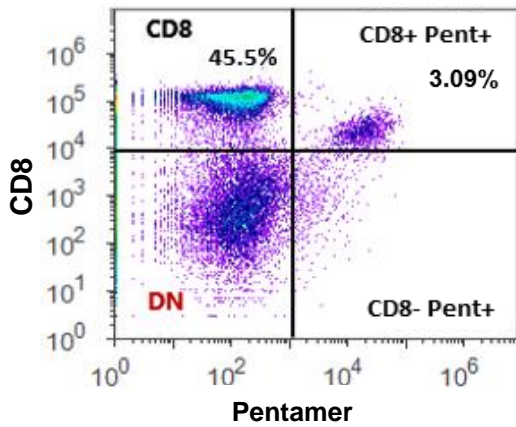


Figure 2

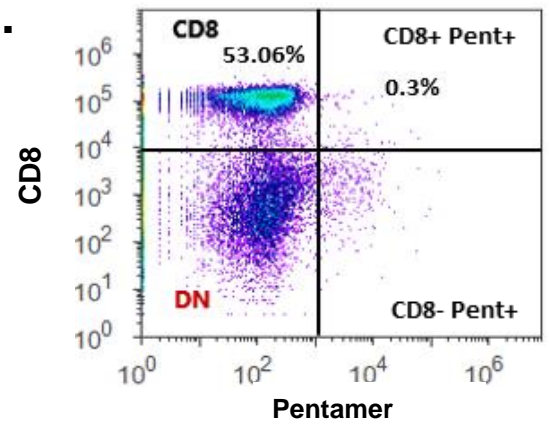
WT (β)

Mut ($\beta 2$)

A.



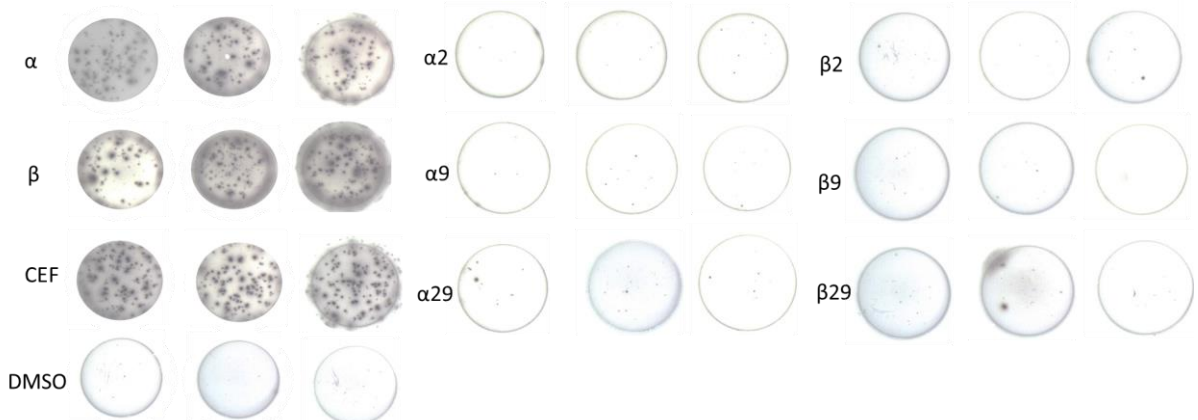
B.



C.

Peptide Code (Position)	Peptide Sequence	HLA binding (percentile rank)	Peptide Code (Position)	Peptide Sequence	HLA binding (percentile rank)
α (240-248)	NLIALSLGL	1.7	β (615-623)	ILEDIVLTL	1.5
$\alpha 2$	NGIALSLGL	26	$\beta 2$	IGEDIVLTL	23
$\alpha 9$	NLIALSLGG	14	$\beta 9$	ILEDIVLTG	12
$\alpha 29$	NGIALSLGG	62	$\beta 29$	IGEDIVLTG	49

D.



E.

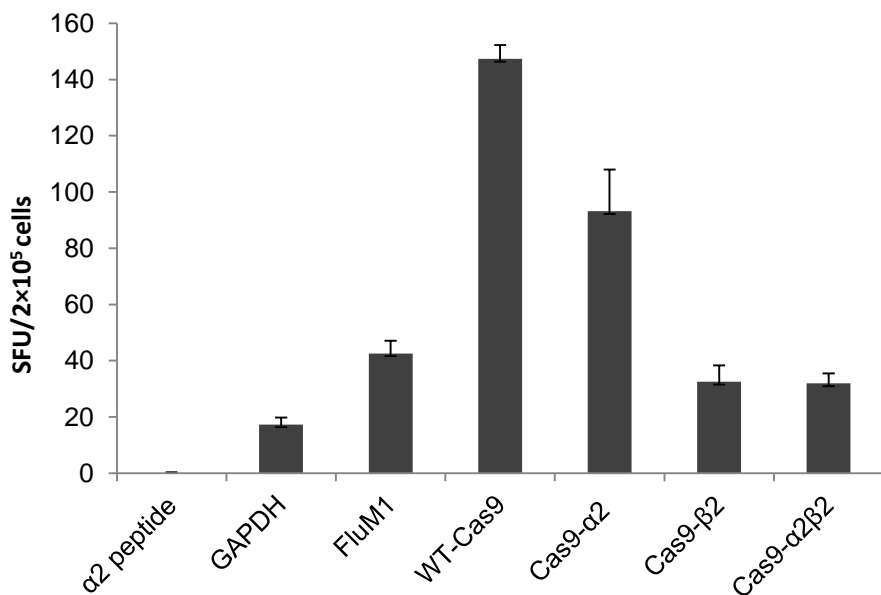
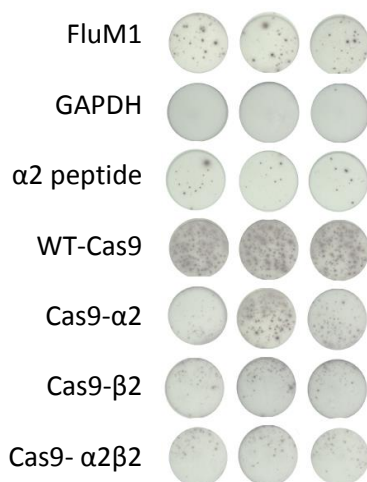
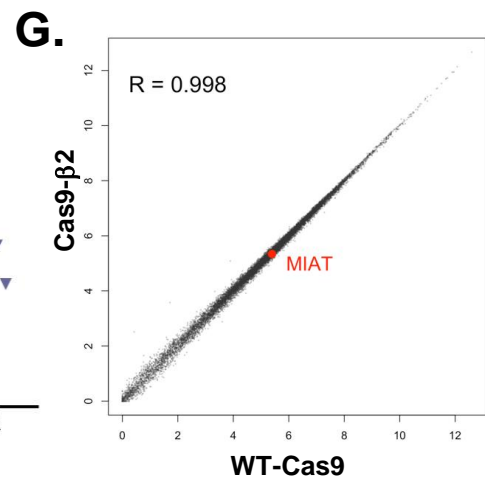
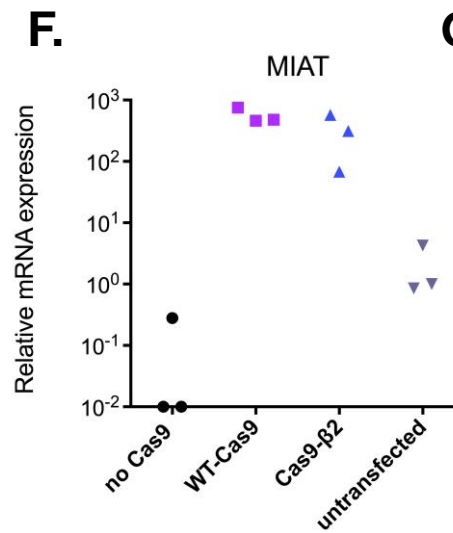
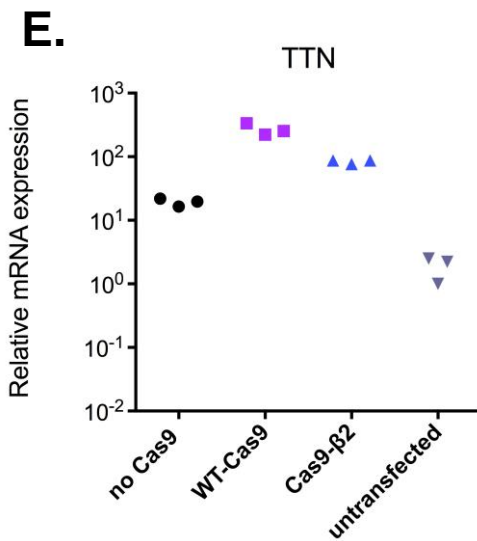
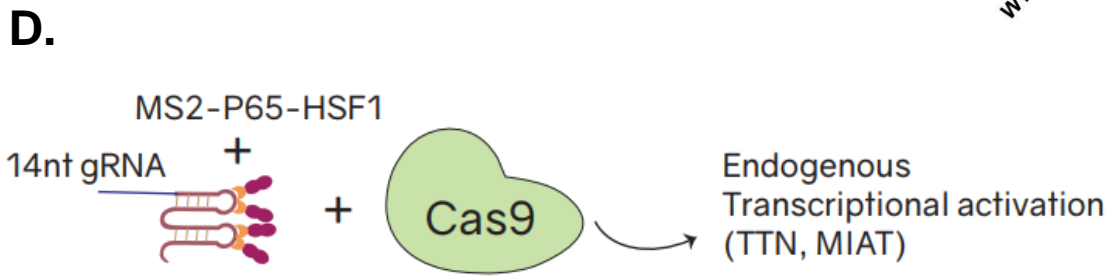
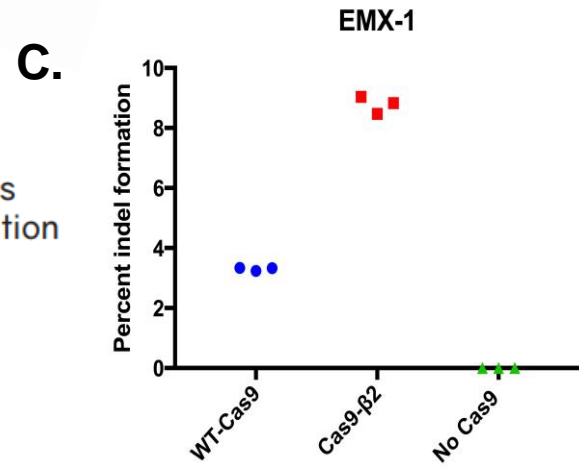
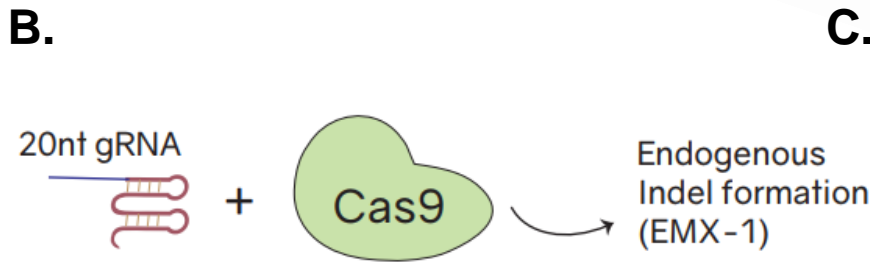
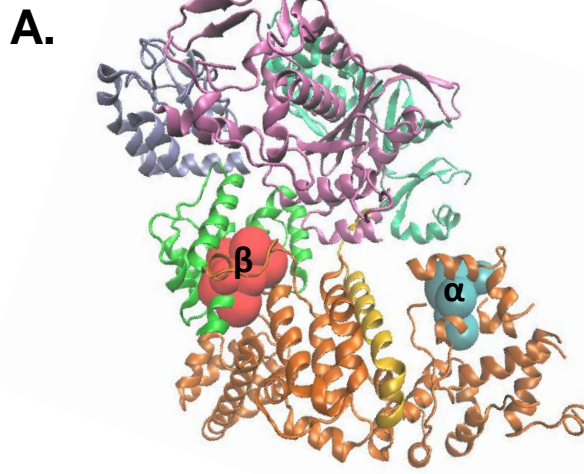


Figure 3



Supplementary Table 1. Predicted Cas9 immunogenic T cell epitopes

Rank	Position	Sequence	Code	Binding			Protein Processing			S_b	S_i	$S_b \cdot S_i$
				IEDB	NetMHC	Syfypeithi	IEDB	ANN				
1	988-997	YLNAVVG TAL	γ	1.25	21.5	24	0.27	0.02	0.068	0.975	0.002	
2	1281-1290	ILADANLDKV		1.25	11.37	31	-0.06	-0.49	0.003	0.447	0.002	
3	236-244	GLFGNLI AL	δ	0.6	10.12	29	1.15	1.04	0.020	0.900	0.002	
4	240-248	NLI ALSGL	α	1.7	61.18	25	0.15	0.22	0.061	0.903	0.006	
5	615-623	ILEDIVL TL	β	1.5	53.29	29	0.28	0.56	0.023	0.710	0.007	
6	614-623	DILEDIVL TL		4.6	3105.79	28	-1.53	-1.02	0.063	0.888	0.007	
7	719-727	SLHEHIANL		1.4	9.14	30	0.93	0.82	0.013	0.380	0.008	
8	415-423	HLGELHAIL		4.4	276.73	25	-0.75	-0.81	0.071	0.876	0.009	
9	300-308	ILLSDILRV		0.3	6.51	29	0.67	0.7	0.019	0.404	0.011	
10	1086-1095	VLSMPQVNIV		3.65	178.87	26	-1.05	-1.43	0.059	0.758	0.014	
11	719-728	SLHEHIANLA		4.7	60.17	19	-0.98	-1.74	0.126	0.890	0.014	
12	1194-1203	LIKLPKYSL		8.5	966.31	25	-0.97	-1.04	0.090	0.841	0.014	
13	1346-1355	TLIHQSITGL		1.95	57.8	27	0.12	-0.06	0.043	0.632	0.016	
14	1197-1207	KLPKYSLFEL		1.2	10.93	27	0.9	0.5	0.040	0.579	0.017	
15	1041-1050	NIMNFFKTEI		2.65	314.8	19	-1.03	-0.9	0.121	0.857	0.017	
16	512-520	SLLYEYFTV		0.4	4.56	25	0.67	0.55	0.056	0.678	0.018	
17	1309-1318	IIHLFTLTNL		4.25	1083.6	24	-1.04	-0.78	0.085	0.787	0.018	
18	661-670	RLSRKLINGI		3.5	278.03	24	-0.82	-1.05	0.078	0.746	0.020	
19	1227-1236	ALPSKYVNFL		4.3	111.14	27	0.05	-0.26	0.051	0.594	0.021	
20	996-1004	ALIKKYPKL		2.6	154.09	28	-0.27	0	0.037	0.407	0.022	
21	221-229	RLENLIAQL		4.2	242.87	26	-0.46	-0.46	0.061	0.581	0.026	
22	1237-1245	YLASHYEKL		1.2	10.3	26	0.9	0.84	0.050	0.446	0.027	
23	1265-1273	YLDEIIEQI		0.3	4.8	26	0.62	0.6	0.046	0.399	0.028	
24	1042-1050	IMNFFKTEI		3.2	131.4	21	-0.69	-0.87	0.103	0.724	0.028	
25	815-824	YLQNGRDMYV		0.25	13.01	22	-0.18	-0.07	0.083	0.627	0.031	
26	1212-1220	RMLASAGEL		3.2	333.2	22	-0.64	-0.51	0.095	0.650	0.033	
27	1020-1029	KMIKSEQEI		3.1	64.01	21	-0.36	-0.9	0.103	0.671	0.034	
28	793-801	SQILKEHPV		2.8	191.23	16	-1.4	-1.36	0.149	0.766	0.035	
29	742-750	KVDELVKV		2.8	44.75	24	-0.06	-0.26	0.074	0.505	0.036	
30	1181-1190	FLEAKGYKEV		3.25	105.27	21	-1.08	-1.42	0.103	0.651	0.036	
31	160-169	HMIKFRGHFL		4.75	324.13	21	-0.59	-0.73	0.110	0.628	0.041	
32	551-559	LLFKTNRKV		3	381.3	25	-1.52	-1.25	0.067	0.368	0.043	
33	141-149	KLVDSTDKA		3.4	274.05	20	-1.48	-1.17	0.114	0.520	0.055	
34	472-481	TITPWNFEEV		4.45	124.55	21	-0.84	-1.21	0.107	0.429	0.061	
35	194-203	QLFEENPINA		1.65	67.94	17	-0.71	-0.79	0.135	0.469	0.072	
36	518-527	FTVYNELTKV		2.55	169.93	20	-1.12	-1.15	0.111	0.216	0.087	
37	473-481	ITPWNFEEV		6.4	351.14	18	-1.25	-1.65	0.143	0.229	0.110	
38	970-978	FQFYKVREI		2.7	135.61	16	-0.66	-0.39	0.148	0.229	0.114	

The table shows Cas9 HLA-A*02:01 epitopes predicted using an integrative prediction model and ranked according to their $S_b \cdot S_i$ score (the lower the more immunogenic). The immunodominant and subdominant epitopes as confirmed by ELISpot are highlighted in dark gray and light gray, respectively. S_b , binding score; S_i , immunogenicity score.

Supplementary Table 3. Sequence Homology of epitope α to amino acid sequences from known proteins.

	Sequence	Similarity (%)	Protein	Sequence ID	Source
1	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_014612333.1	<i>Streptococcus dysgalactiae</i>
2	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_054279288.1	<i>Streptococcus phocae</i>
3	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_067062573.1	<i>Streptococcus pantholopis</i>
4	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_048800889.1	<i>Streptococcus constellatus</i>
5	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_002304487.1	<i>Streptococcus mutans</i>
6	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_049516684.1	<i>Streptococcus anginosus</i>
7	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_003079701.1	<i>Streptococcus macacae</i>
8	NLIALSGL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	GAD40915.1	<i>Streptococcus intermedius</i> SK54
9	NLIAFSLGL	8/9 (89%)	Full=RNA polymerase-associated protein RapA; AltName: Full=ATP-dependent helicase HepA	Q6LV34.1	<i>Photobacterium profundum</i> SS9
10	NLI S SLGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_096633625.1	<i>Streptococcus parauberis</i>
11	NLIAL A LGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_075103982.1	<i>Streptococcus cuniculi</i>
12	NLIAL A LGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_058692367.1	<i>Streptococcus gallolyticus</i>
13	NLIAL A LGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_061100419.1	<i>Streptococcus pasteurianus</i>
14	NLIAL A LGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_018363470.1	<i>Streptococcus caballi</i>
15	NLIAL A LGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_099412266.1	<i>Streptococcus macedonicus</i>
16	NLIAL A LGL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_014334983.1	<i>Streptococcus infantarius</i>
17	D LIALYLGL	7/9 (78%)	Full=NADH-quinone oxidoreductase subunit N; AltName: Full=NADH dehydrogenase I subunit N; AltName: Full=NDH-1 subunit N	A8I421.1	<i>Azorhizobium caulinodans</i> ORS 571
18	NL L ALALGL	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_007896501.1	<i>Streptococcus pseudoporcinus</i>
19	NLIG L ALGL	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_061587801.1	<i>Streptococcus oralis</i>
20	NL V ALALGL	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_074862269.1	<i>Streptococcus equinus</i>
21	NL V AL V LGL	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_020917064.1	<i>Streptococcus lutetiensis</i>
22	S LIAFSLGL	7/9 (78%)	ectoine/hydroxyectoine ABC transporter permease subunit EhuD	WP_086160327.1	<i>Streptomyces</i> sp. SCSIO 03032
23	Y LIALALGL	7/9 (78%)	ectoine/hydroxyectoine ABC transporter permease subunit EhuD	WP_026413155.1	<i>Actinomadura oligospora</i>

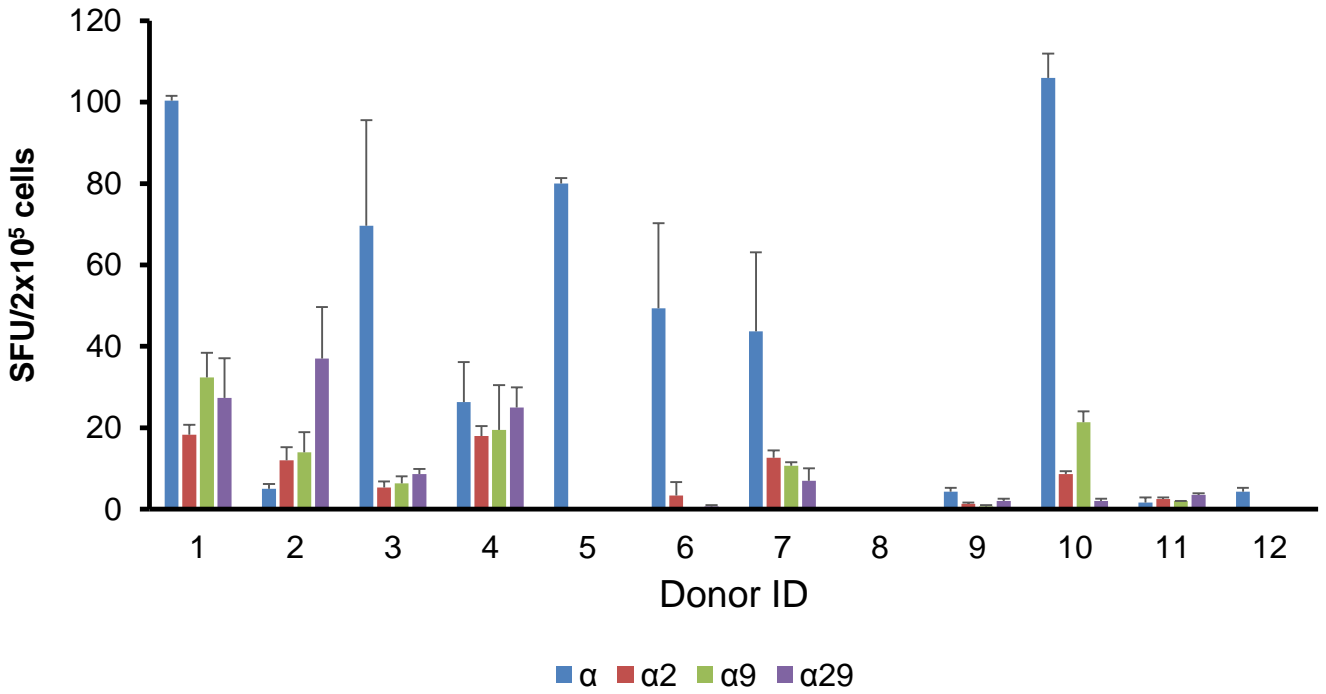
Supplementary Table 4. Sequence Homology of epitope β to amino acid sequences from known proteins.

	Sequence	Similarity (%)	Protein	Sequence ID	Source
1	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_084916602.1	<i>Streptococcus dysgalactiae</i>
2	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_074484960.1	<i>Streptococcus henryi</i>
3	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_003088697.1	<i>Streptococcus rattii</i>
4	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_044681799.1	<i>Streptococcus suis</i>
5	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_024786433.1	<i>Streptococcus mutans</i>
6	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_057491067.1	<i>Streptococcus orisasini</i>
7	ILEDIVLTL	9/9 (100%)	type II CRISPR RNA-guided endonuclease Cas9	WP_082312238.1	<i>Streptococcus intermedius</i>
8	I L E G IVLTL	8/9 (89%)	peptide chain release factor 2	NP_275123.1	<i>Neisseria meningitidis</i> MC58
9	ILEDIV Q TL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	EAO61901.1	<i>Streptococcus agalactiae</i>
10	ILEDIV Q TL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_070454905.1	<i>Streptococcus</i> sp. HMSC063D10
11	V LEDIVLTL	8/9 (89%)	type II CRISPR RNA-guided endonuclease Cas9	WP_075346866.1	<i>Streptococcus</i> sp. 'caviae'
12	V LEDIV L S L	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_093650272.1	<i>Streptococcus varani</i>
13	I L E N I V H T L	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	KYF37509.1	<i>Streptococcus mitis</i>
14	I L E N I V H T L	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_084972088.1	<i>Streptococcus oralis</i>
15	I L E N I V H T L	7/9 (78%)	type II CRISPR RNA-guided endonuclease Cas9	WP_045635197.1	<i>Streptococcus gordonii</i>

Supplementary Figure 1

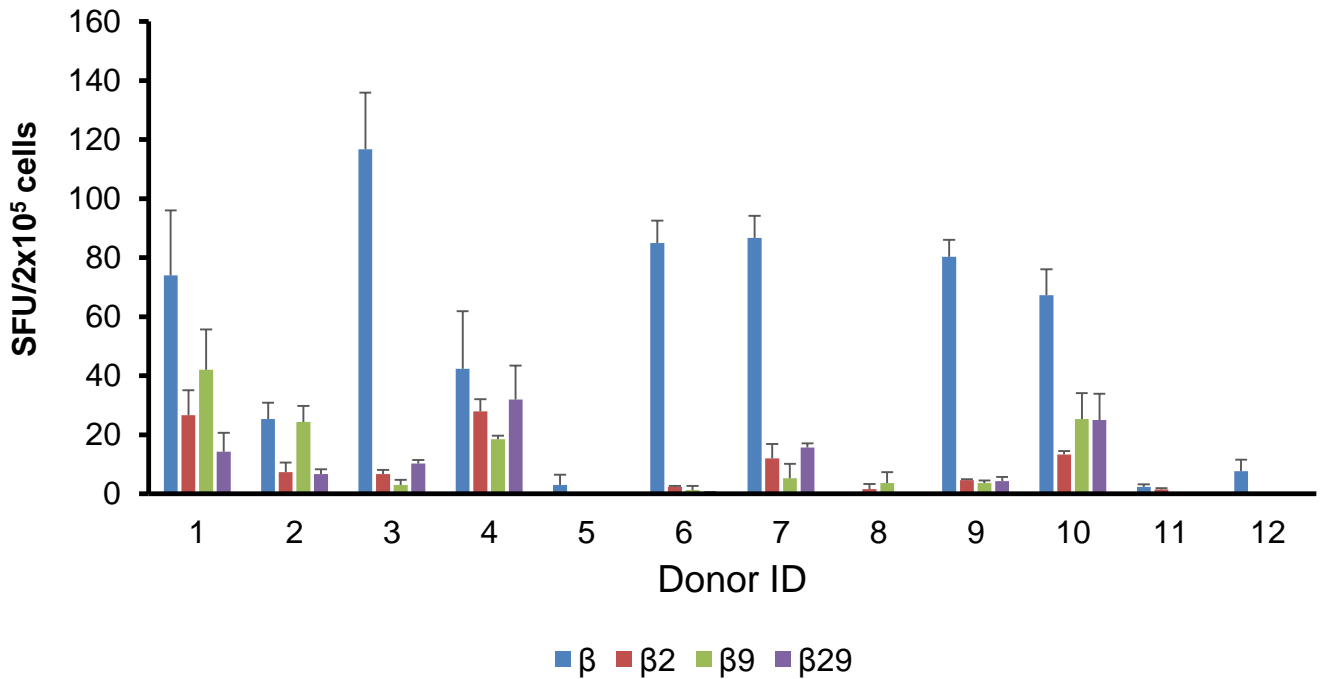
A.

Peptide α



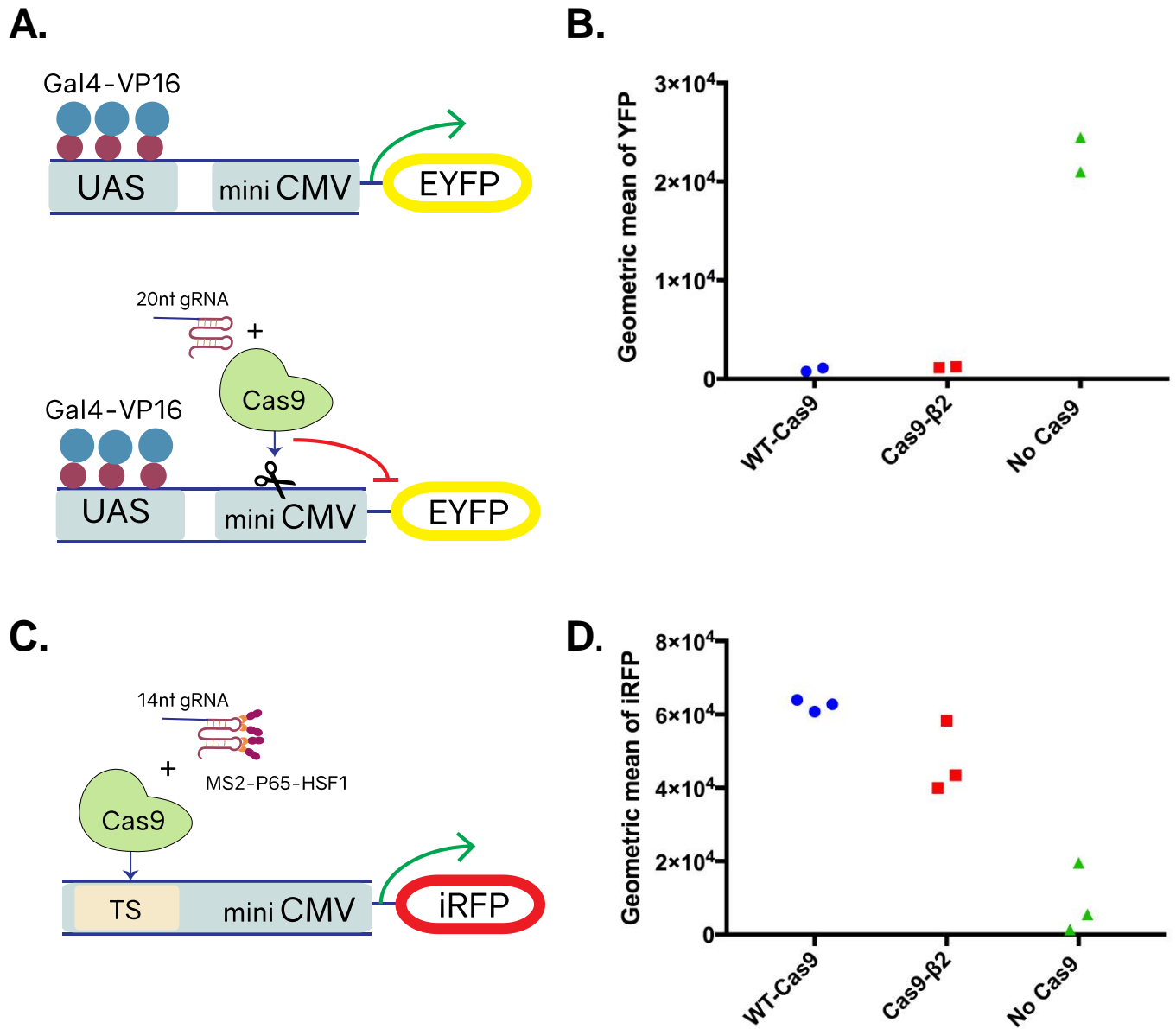
B.

Peptide β



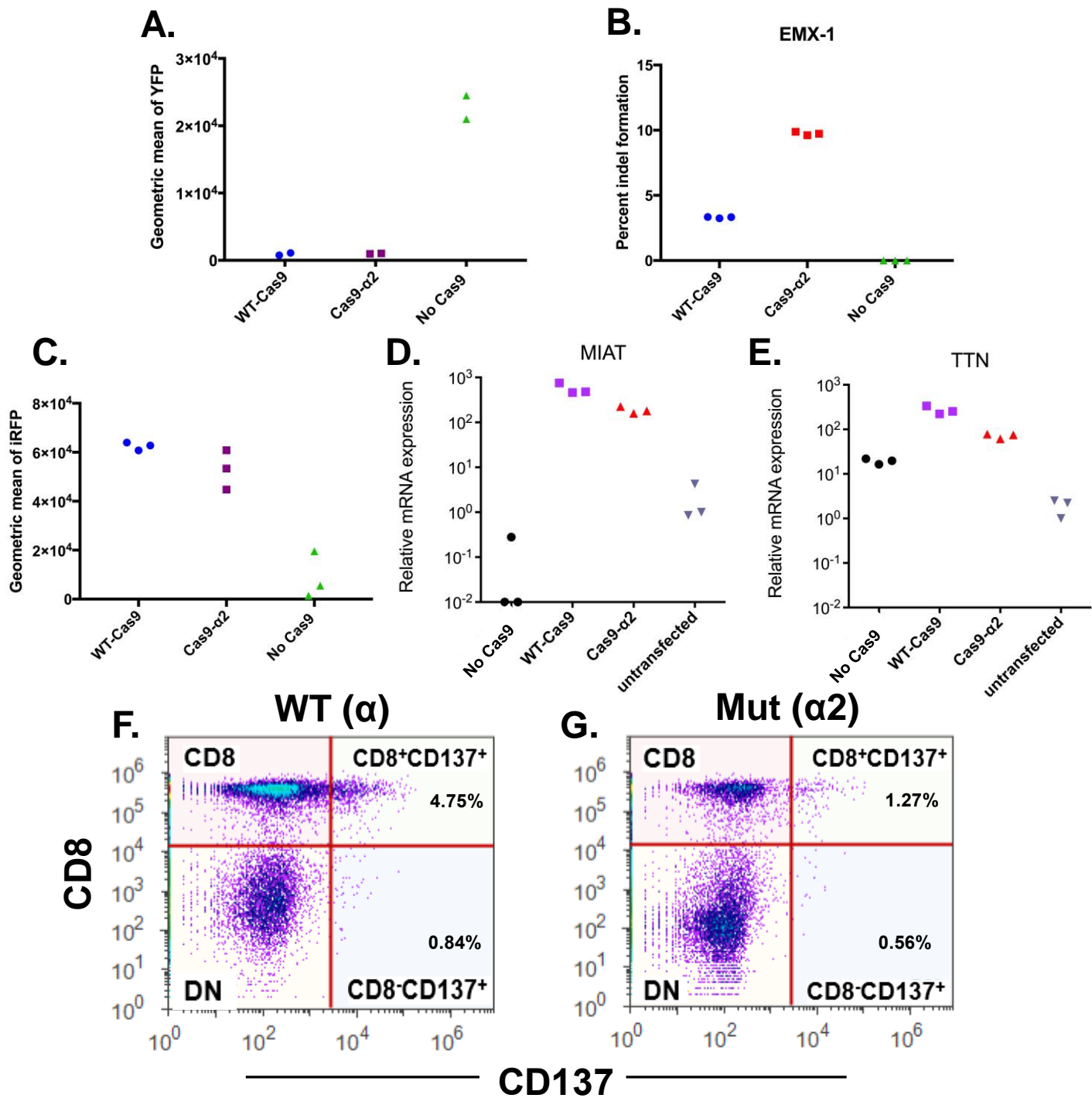
Supplementary Figure 1. Reduced T cell response to epitopes α and β after mutation of the anchor residues. **A and B.** IFN- γ ELISpot for 12 healthy donor PBMCs stimulated with wild type or mutated peptide α (**A**) or β (**B**). Data represent mean \pm SEM.

Supplementary Figure 2



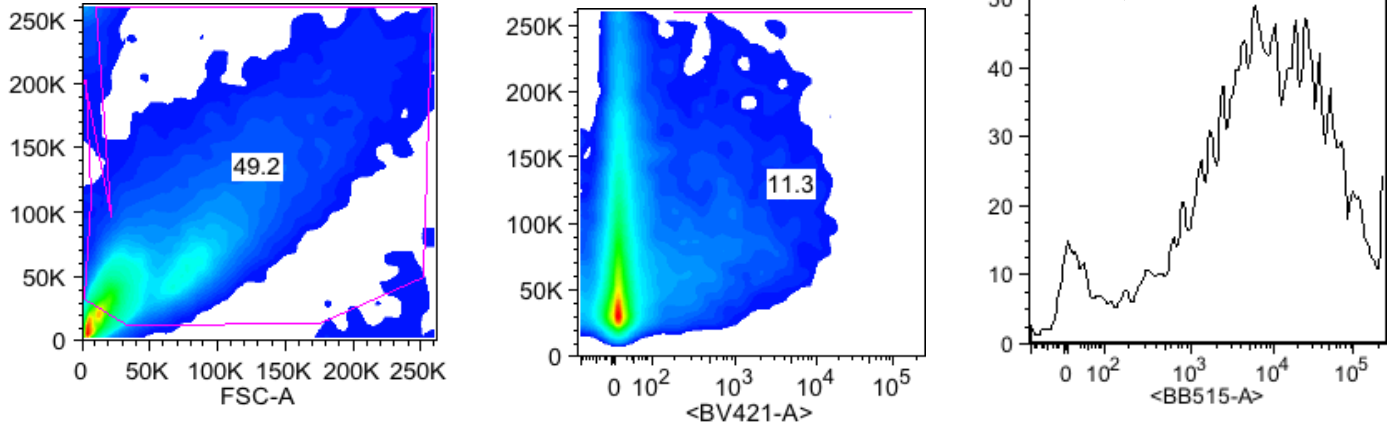
Supplementary Figure 2. A. Schematic of the experiment assessing Cas9-β2 cleavage capacity at a synthetic promoter. Cells were transfected with either WT-Cas9, Cas9-β2 or an empty plasmid as well as 20nt gRNA targeting a synthetic CRISPR promoter that harbors two gRNA target sites flanking a mini-CMV promoter. The targeting and cleavage at the promoter should disrupt the promoter and decrease EYFP expression. **B.** Each individual dot represents EYFP expression 48 hrs after transfection in cells expressing $>2 \times 10^2$ A.U. of a transfection marker measured by flow cytometry ($n=2$ individual transfections represented by individual dots). **C.** Schematic of the experiment assessing Cas9-β2 transcriptional activation capacity at a synthetic promoter. Cells were transfected with either WT-Cas9, Cas9-β2 or an empty plasmid as well as aptamer binding transcriptional activation domains, and a 14nt gRNA targeting a synthetic CRISPR promoter that harbors multiple target sites upstream of a mini-CMV promoter. Targeting at the promoter should enable iRFP expression. **D.** Each individual dot shows iRFP expression 48 hrs after transfection in cells expressing $>2 \times 10^2$ A.U. of a transfection marker measured by flow cytometry ($n=3$ individual transfections represented by individual dots, $n=2$ for the no Cas9 group).

Supplementary Figure 3

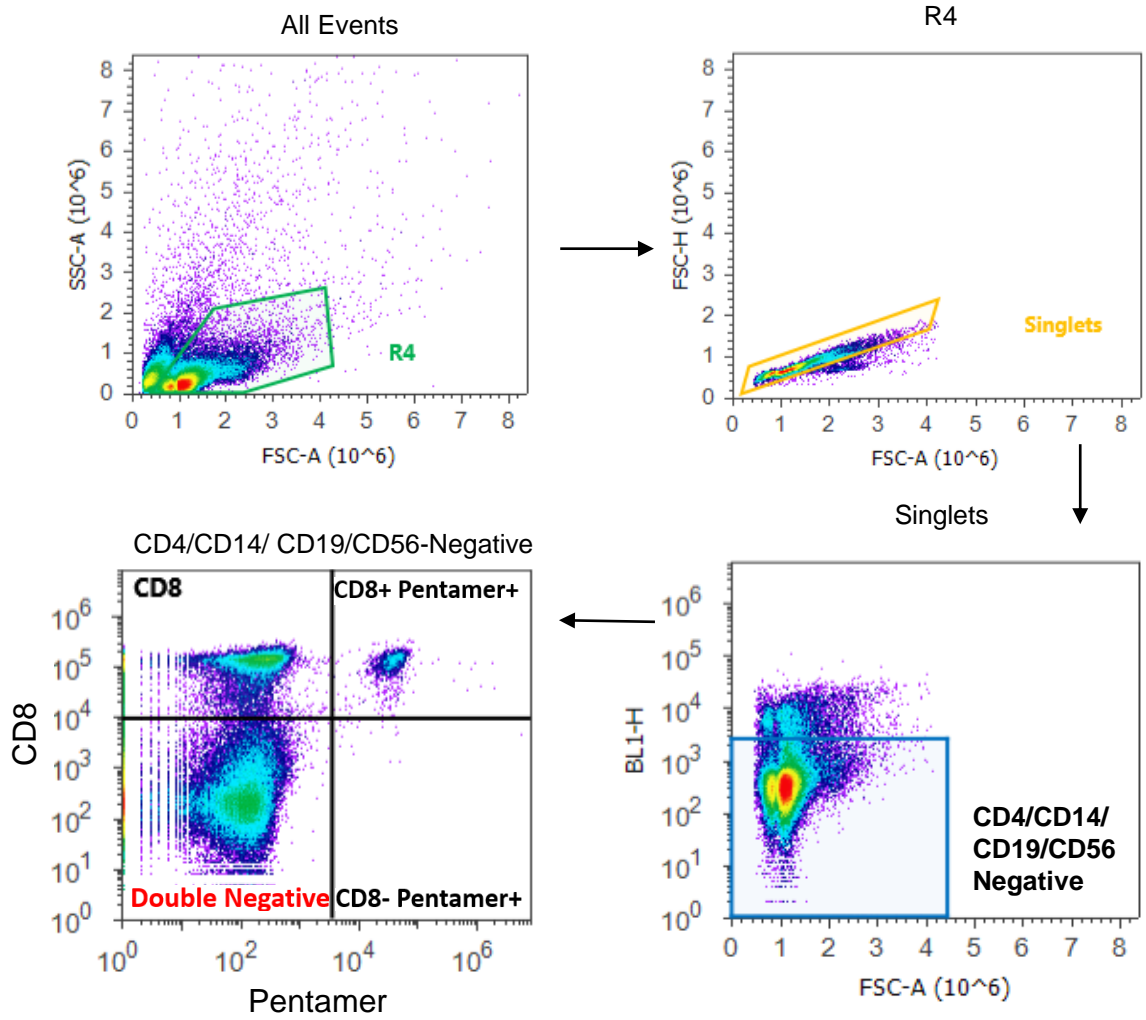


Supplementary Figure 3. **A.** Analysis of the mutagenesis capacity of Cas9- $\alpha 2$ as compared to WT-Cas9 in a synthetic promoter. Each individual dot shows EYFP expression 48 hrs after transfection in cells expressing $>2 \times 10^2$ A.U. of a transfection marker measured by flow cytometry ($n=2$ individual transfections, represented by individual dots). **B.** Mutagenesis in the endogenous *EMX-1* locus. Percentage of indel formation in the *EMX-1* locus. Dots represent three individual transfections. **C.** Transcriptional modulation by Cas9- $\alpha 2$ at a synthetic promoter. Each individual dot shows iRFP expression 48 hrs after transfection in cells expressing $>2 \times 10^2$ A.U. of a transfection marker measured by flow cytometry ($n=3$ individual transfections, $n=2$ for no Cas9 group represented by dots). **D, E.** Shown is the mRNA level relative to an untransfected control experiment. Each individual dot represents an individual transfection. Note for A and C WT-Cas9 and no Cas9 data are also reported in Supplementary Fig.2. For B-E. WT-Cas9 and no Cas9 data are also reported in Fig.3. **F.** Activated CD8+CD137+ T cells detected in PBMCs stimulated with peptide α or peptide $\alpha 2$ (data representative of 3 individual stimulations).

A.



B.



Supplementary Figure 4. A. Representative flow cytometry gating for analysis of Cas9 function on synthetic promoters. Cells are gated based on Forward (FSC) and Side Scatter (SSCs). Transfected population then is gated based on expression of EBFP (BV421-A) more than 2×10^2 . The geometric mean of the output (EYFP) or iRFP is determined in this population. **B.** Flow cytometry gating for analysis of Cas9 pentamer+ CD8+ T lymphocytes. Cells are gated based on FSC and SSC and negatively gated on CD4/CD14/CD19/CD56. The CD8-, CD8+, and Cas9 pentamer+ population are shown (bottom right).

Supplementary notes:

Sequences of Cas9:

Cas9-β2:

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAAAGACGATGACGA
TAAGATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAG
TACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAA
GGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCCTG
ATCGGAGCCCTGCTGTTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCA
GAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGAT
GGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCTTCTGCTGGAAGAGGATAAGA
AGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCC
ACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTA
TCTGGCCCTGGCCACATGATCAAGTTCGGGGCCACTTCTGATCGAGGGCGACCTGAACCCCG
ACAACATCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAA
AACCCCTCAACGCGCGGTGGACGCAAGGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGAG
ACGGCTGGAAAATCTGATCGCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGAACCTTA
TTGCCCTGAGCCTGGGCTGACCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAA
CTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCA
GTACGCCGACCTGTTTCTGGCCGCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAG
AGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACC
ACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCAGCAGCAGCTGCCTGAGAAGTACAAAGAGATT
TTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT
CTACAAGTTCATCAAGCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGA
ACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCTGAAGGACAACCG
GGAAAAGATCGAGAAGATCCTGACCTTCGCATCCCTACTACGTGGGCCCTCTGGCCAGGGGAA
ACAGCAGATTCGCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTTCGAGGA
AGTGGTGGACAAGGGCGCTTCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACC
TGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAG
CTGACCAAAGTGAATAACGTGACCGAGGGAATGAGAAAGCCCGCCTTCTGAGCGGCGAGCAGA
AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGA
GGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAAGATCGGTTCA
ACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCTGGACAAT
GAGGAAAACGAGGACATTGGTGAAGATATCGTGTGACCCTGACACTGTTTGAGGACAGAGAGAT
GATCGAGGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGAAGC
GGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCA
GTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGC
TGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCAGGTGTCCGGCCAGGGC
GATAGCCTGCACGACACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCA
GACAGTGAAGGTGGTGGACGAGCTCGTGAAGTGATGGGCCGGCACAAGCCCCGAGAACATCGTG
ATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATG
AAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAA
ACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTG
GACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTT
TCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGC
GACAACGTGCCCTCCGAAGAGGTCTGTAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACG
CCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGGCGGCCTGAGCGA
ACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTG
GCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGT
GAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGT
GCGCGAGATCAACAACCTACCACCACGCCACGACGCCTACCTGAACGCCGTCTGGGAACCGCCC
TGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTG
CGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAG
CAACATCATGAACTTTTTCAAGACCGAGATTACCTGGCCAACGGCGAGATCCGGAAGCGGCCTC
TGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTG

CGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCT
TCAGCAAAGAGTCTATCTGCCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTG
GGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCA
AAGTGGAAAAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCAT
GGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTG
AAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAG
AATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGA
ACTTCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCGAGGATAATGAGCAGAAA
CAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTTTC
CAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAAGTGTGTCCGCCTACAACAAGCACCAGG
ATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC
CCTGCCGCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACAAAGAGGT
GCTGGACGCCACCCTGATCCACCAGAGCATCACCGCCTGTACGAGACACGGATCGACCTGTCTC
AGCTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAGGCCGGCCAGGCAAAAAGAAAAGT
AA

Cas9- $\alpha 2$:

ATGGACTATAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAAAGACGATGACGA
TAAGATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGTCCCAGCAGCCGACAAGAAG
TACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAA
GGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCCTG
ATCGGAGCCCTGCTGTTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCA
GAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGAT
GGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCTTCTGCTGGAAGAGGATAAGA
AGCAGGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCC
ACCATCTACCACCTGAGAAAGAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTA
TCTGGCCCTGGCCACATGATCAAGTTCGGGGCCACTTCTGATCGAGGGCGACCTGAACCCCG
ACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAA
AACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAG
ACGGCTGGAAAATCTGATCGCCCAGCTGCCCGCGGAGAAGAAGAATGGCCTGTTTCGAAAACGGTA
TTGCCCTGAGCCTGGGCTGACCCCCAATTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAA
CTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCAGATCGGCGACCA
GTACGCCGACCTGTTTCTGGCCGCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAG
AGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACC
ACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCAGGACGAGCTGCCTGAGAAGTACAAAGAGATT
TTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT
CTACAAGTTCATCAAGCCATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGA
ACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCATTCTGAAAGGACAACCG
GGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCTACTACGTGGGCCCTCTGGCCAGGGGAA
ACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTTCGAGGA
AGTGGTGGACAAGGGCGCTTCCGCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACC
TGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAG
CTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCTGAGCGGCGAGCAGA
AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGA
GGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAAGATCGGTTCA
ACGCCTCCCTGGGCACATAACCAGATCTGCTGAAAATTATCAAGGACAAGGACTTCTGGACAAT
GAGGAAAACGAGGACATTCTTGAAGATATCGTGTGACCCTGACACTGTTTGAGGACAGAGAGAT
GATCGAGGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAAGTGTGATGAAGCAGCTGAAG
GGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCA
GTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGC
TGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCAGGTGTCCGGCCAGGGC
GATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCA
GACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCCGAGAACATCGTG
ATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATG
AAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAA

ACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTG
GACCAGGAAGTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTT
TCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGC
GACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCAGGCTGCTGAACG
CCAAGCTGATTACCCAGAGAAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGA
ACTGGATAAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTG
GCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGT
GAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGT
GCGCGAGATCAACAACCTACCACACGCCACGACGCTACCTGAACGCCGTCGTGGGAACCGCCC
TGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTG
CGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAG
CAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTC
TGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAAGGGCCGGGATTTTGCCACCGTG
CGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGA AAAAGACCGAGGTGCAGACAGGCGGCT
TCAGCAAAGAGTCTATCTGCCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTG
GGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCA
AAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCAT
GGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTG
AAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAG
AATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGA
ACTTCTGTACTTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCGAGGATAATGAGCAGAAA
CAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTC
CAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGTGTCCGCCTACAACAAGCACCGGG
ATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC
CCTGCCGCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAGAGGT
GCTGGACGCCACCCTGATCCACCAGAGCATACCGGCCTGTACGAGACACGGATCGACCTGTCTC
AGCTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAGGCCGGCCAGGCAAAAAGAAAAGT
AAG

Cas9- $\alpha 2$ - $\beta 2$

ATGGACTATAAAGGACCACGACGGAGACTACAAGGATCATGATATTGATTACAAAGACGATGACGA
TAAGATGGCCCCAAAGAAGAAGCGGAAGGTTCGGTATCCACGGAGTCCACAGCAGCCGACAAGAAG
TACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAA
GGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTG
ATCGGAGCCCTGCTGTTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCA
GAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGAT
GGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCTTCTGTTGGTGAAGAGGATAAGA
AGCAGGAGCGGCACCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCC
ACCATCTACCACCTGAGAAAAGAACTGGTGGACAGCACCAGCAAGGCCGACCTGCGGCTGATCTA
TCTGGCCCTGGCCACATGATCAAGTTCGGGGGCCACTTCTGATCGAGGGCGACCTGAACCCCG
ACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTTCGAGGAA
AACCCCATCAACGCCAGCGGCTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAG
ACGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTTCGGAACCGTA
TTGCCCTGAGCCTGGGCTGACCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAA
CTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCAGATCGGCGACCA
GTACGCCGACCTGTTTCTGGCCGCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAG
AGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCTCTATGATCAAGAGATACGACGAGCACC
ACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCAGCAGCAGCTGCCTGAGAAGTACAAAGAGATT
TTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT
CTACAAGTTCATCAAGCCATCCTGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGA
ACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCACCAGATCCAC
CTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCATTCTGAAGGACAACCG
GGAAAAGATCGAGAAGATCCTGACCTTCGCATCCCTACTACGTGGGCCCTCTGGCCAGGGGAA
ACAGCAGATTCGCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTCGAGGA
AGTGGTGGACAAGGGCGCTTCGCCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACC

TGCCAACGAGAAGGTGCTGCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAG
 CTGACCAAAGTCAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGA
 AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGA
 GGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCA
 ACGCCTCCCTGGGCACATAACCAGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAAT
 GAGGAAAACGAGGACATTGGTGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGAT
 GATCGAGGAACGGCTGAAAACCTATGCCACCTGTTTCGACGACAAAGTGATGAAGCAGCTGAAGC
 GGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCA
 GTCCGGCAAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGC
 TGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCAGGTGTCCGGCCAGGGC
 GATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCA
 GACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTG
 ATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATG
 AAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAA
 ACACAGCTGCAGAACGAGAAGCTGTACTGTACTACCTGCAGAATGGGCGGGATATGTACGTG
 GACCAGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTT
 TCTGAAGGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCAGGGGCAAGAGC
 GACAACGTGCCCTCCGAAGAGTCCGTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACG
 CCAAGCTGATTACCCAGAGAAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGGCGCCTGAGCGA
 ACTGGATAAAGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTG
 GCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGT
 GAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGT
 GCGCGAGATCAACAACCTACCACCGCCACGACGCTACCTGAACGCCGTCGTGGGAACCGCCC
 TGATCAAAAAGTACCCTAAGCTGGAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTG
 CGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAG
 CAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTC
 TGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAAGGGCCGGGATTTTGCCACCGTG
 CGGAAAGTGCTGAGCATGCCCAAGTGAATATCGTGAAAAGACCGAGGTGCAGACAGGCGGCT
 TCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTG
 GGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCA
 AAGTGGAAAAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCAT
 GGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTG
 AAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAG
 AATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAACGAACTGGCCCTGCCCTCCAAATATGTGA
 ACTTCTGTACTTGGCCAGCCACTATGAGAAGCTGAAGGGCTCCCCGAGGATAATGAGCAGAAA
 CAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTC
 CAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGG
 ATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC
 CCTGCCGCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGT
 GCTGGACGCCACCCTGATCCACCAGAGCATCACCGCCTGTACGAGACACGGATCGACCTGTCTC
 AGCTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAGGCCGGCCAGGCCAAAAAAGAAAAAAGT

Sequences of gRNAs:

MIAT-14bp gRNA	GAGGCTGAGCGCAC
TTN-14bp gRNA	GGAAGTCTCCTTTG
Reporter2-20bp gRNA	GTCCCCTCCACCCACAGTG
CR10-14bp-gRNA	GCATCAGGAACATGT
EMX1- 20bp gRNA	CACC GAGTCCGAGCAGAAGAAGAA

Sequences of primers:

Cas9 fragment1- FW	ttttGGTCTCTAGGTCCACCATGGACTATAAGGACCACGA
Cas9 fragment1- RV	tttgtctcaGAACAGCTGGTTGTAGGTCTGCA

Cas9 fragment2-FW	ttttGGTCTCTACCAACCGGAAAGTGACCGTGAAG
Cas9 fragment2-RV	ttttGGTCTCAAAGCTTACTTTTTCTTTTTTGCC
qPCRMIAT-FW	TGGCTGGGGTTTGAACCTTT
qPCR-MIAT RV	AGGAAGCTGTTCCAGACTGC
qPCRTTN FW	TGTTGCCACTGGTGCTAAAG
qPCR-TTN-RV	ACAGCAGTCTTCTCCGCTTC
PCR-EMX1-FW	CCATCCCCTTCTGTGAATGT
PCR-EMX1-RV	GGAGATTGGAGACACGGAGA