Analyzing HLA Sequences to Predict Organ Rejection and Find Targets for Precise Immunosuppression

Grant Proposal

Samhitha Bodangi

Massachusetts Academy of Math and Science

Worcester, MA

Author Note

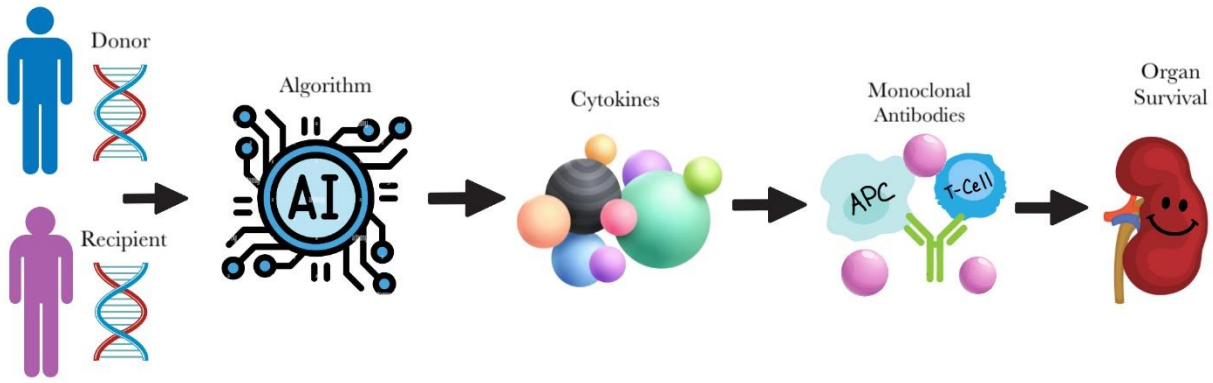
I sincerely thank Dr. David Harlan of UMass Diabetes Center of Excellence, Demetri Maxim of Nephrogen, Dr. Floyd E. Brownewell of Worcester Polytechnic Institute, and Dr. Lawrence Stern of UMass Chan Medical School for their support and insight during the development of my project. Finally, I extend my greatest appreciation to Dr. Kevin Crowthers for his support and feedback during the development of my project and this proposal.

Abstract

Organ transplants are among the greatest advances in modern medicine, saving many lives every year. However, many medical complications may occur after the transplant, such as organ rejection. Currently, all transplant patients are prescribed life-long immunosuppressors to decrease organ rejection. While these medications prevent organ rejection to an extent, about 10-20% of patients will still experience at least one episode of rejection. Additionally, they can also severely weaken the immune system, increasing the risk of cancer, infections, and other diseases. Rejection is primarily caused because of the Human Leukocyte Antigen (HLA) mismatches between the donor and the recipient. HLA genes are very polymorphic and classifying entire HLA mismatches does not account for the allele differences that can start rejection. The main objective is to understand how organ rejection can be decreased with selective T-cell inhibition by analyzing donor and recipient HLA sequences and predicting MHC-peptide complexes. Additionally, understanding the specific MHC-peptide complexes that will initiate rejection can provide greater insight into specific immunosuppressive targets. The model should accurately predict rejection and provide specific targets for precise immunosuppression and can be constructed within an open-source web application.

Keywords: Organ rejection, immune system, T cells, MHC-peptide complex, machine learning

Graphical Abstract



Analyzing HLA Sequences to Predict Organ Rejection and Find Targets for Selective T-Cell Inhibition

Organ transplants are among the greatest advances in modern medicine, saving tens of thousands of lives every year. By increasing life expectancies and improving the quality of life, they remain the best therapy for terminal and irreversible organ failure (Grinyó, 2013). However, there is currently a major problem in the organ transplant industry: the demand is vastly greater than the supply. Due to a lack of organ donations, about seventeen people die each day while waiting for an organ transplant (*Organ, Eye and Tissue Donation Statistics*, n.d.). The immense demand emphasizes that every donated organ has the potential to change lives, and it is crucial to maintain the long-term health of each organ, for the sake of the patient *and* the organ.

Overview of Organ Rejection

Even if a patient is successful in receiving an organ transplant, many medical complications may occur after the transplant, the most common being organ rejection. The immune system is a body system that destroys foreign cells to protect the body from harm. In the case of organ rejection, the immune system recognizes the transplanted organ as foreign and attempts to attack it by producing cells or antibodies that invade the organ (*Understanding Transplant Rejection | Stony Brook Medicine*, n.d.). Currently, all transplant patients are prescribed immunosuppressors to decrease the risk of organ rejection. However, recipients must take immunosuppressive drugs for their entire lives for their bodies to accept a donated organ. While these medications prevent organ rejection to an extent, about 10-20% of patients will still experience at least one episode of rejection within the first three months to one year after a transplant (*Organ Rejection after Renal Transplant | Columbia Surgery*, n.d.). Additionally, they can also severely weaken the immune system, increasing the risk of cancer, infections, and other diseases (Kelly, 2022). New treatments are necessary to prevent organ rejection without using broad immunosuppressors that weaken the entire immune system.

Chronic Rejection

Depending on the mechanisms and timeframe of the rejection episode, rejection can be categorized into many different types. Acute and chronic rejection are categorized based on the time rejection occurred after the transplant. Acute rejection occurs within the first three months to a year after the transplant, while chronic rejection can occur after the first year of the transplant. Chronic rejection is often irreversible and can lead to graft failure or death (Hunt & Saab, 2012). Immunosuppressors are effective in decreasing the risk of acute rejection, but not against chronic rejection. By five years post-transplant, chronic rejection affects up to 50% of kidney transplants (Gautreaux, 2017). Since chronic rejection is often asymptomatic and occurs over an extended period, there is currently no medicine to date that can treat chronic rejection symptoms (*Understanding Transplant Rejection | Stony Brook Medicine*, n.d.). The common treatment method is to increase the dosage of immunosuppressive drugs, which can exacerbate the dangerous side effects. Therefore, it is imperative to understand and target the mechanisms that are involved in chronic rejection to maintain long-term allograft health.

MHC-Peptide Presentation and T-Cell Activation

Early chronic organ rejection is primarily caused by T-cell-mediated rejection (Chong, 2020). T-cells are a type of immune cell that play a crucial role in identifying and eliminating foreign cells. When T-cells misinterpret donated organ cells as foreign, it can lead to T-cell activation and an attack on the transplanted organ. MHC peptide presentation plays a vital role in T-cell activation and can lead to developing strategies to prevent transplant rejection. The major histocompatibility complex (MHC) is a group of genes that code for MHC molecules found on the surface of cells. These proteins play a vital role in the immune system's ability to distinguish between "self" and "non-self" (*General, Non-Specific Defenses Against Infection*, n.d.). There are two main types of MHC molecules: MHC class I and MHC class II molecules. As MHC class I molecules are present on all nucleated body cells and directly interact

with T-cells, this project will focus on MHC class I peptide presentation (Hewitt, 2003). In the case of organ transplantation, intracellular proteins are broken down into smaller peptides inside the organ cells. These peptides are transported into the endoplasmic reticulum and bind to a groove in the MHC class I molecule, forming a peptide-MHC complex. This complex then travels to the cell surface and is displayed for T-cells to recognize. If T-cell receptors (TCRs) recognize a peptide from the transplanted organ on an MHC molecule, it activates, starting the immune response against the transplanted organ. Once T-cells recognize foreign antigens displayed on antigen-presenting cells (APC), proteins called cytokines synthesize and allow for the proliferation and differentiation of T-cells that attack the organ (Ingulli, 2010). Multiple cytokines or pathways may be active during rejection and may be different from person to person (Chen et al., 2019).

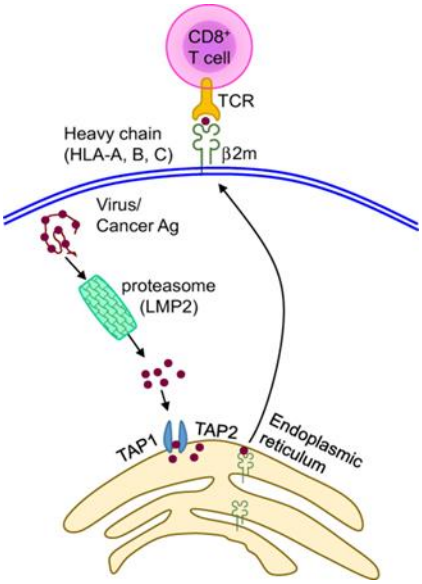


Figure 1: Intracellular antigenic peptides are presented to CD8+ T cells by MHC class I. Even though this figure depicts antigens of virus or cancer origins, the process is synonymous with organ cells. The antigens are processed by the immunoproteasome into peptides that are transported into the endoplasmic reticulum through TAP1 and TAP2 transporter proteins. These peptides are then loaded onto the MHC class I molecule (β2m). The MHC-peptide complex is presented on the cell surface to CD8+ T-cells (Vijayan et al., 2019).

Tissue Typing and Immune Profiling

This document proposes selective T-cell inhibition with a personalized medicine approach to induce donor-specific tolerance. When looking for organ matches, doctors perform Human Leukocyte Antigen[1] (HLA) typing to understand the similarity in antigens between the donor and the recipient. The HLA is a group of genes that provide instructions to make antigens present on the surface of cells (Manski et al., 2019). Six specific HLAs are looked for, and a higher similarity results in a likely chance of an organ match (*Matching and Compatibility | Transplant Center | UC Davis Health*, n.d.). However, HLA

[1] The human leukocyte antigen (HLA) complex is synonymous with the human MHC. MHC is a general term describing MHCs found in all vertebrates, while HLAs describe MHCs found in humans (Viatte, 2023).

genes are the most polymorphic genes in the human genome. This means that HLAs have many different allele combinations, and their variant alleles have high degrees of sequence similarity. The similarity can be difficult to establish with current serological and low-resolution tests (Dasgupta, 2016). Therefore, understanding the exact differences in HLAs between the donor and recipient can result in a better treatment method that is personalized and accurate for the recipient.

Benefits of Machine Learning

Machine learning is a subset of artificial intelligence that uses statistical techniques that allow computer systems to automatically learn and develop from experience without being explicitly programmed (Costa, 2019). Previous studies have employed machine learning techniques to sift through massive datasets of gene expression data. Machine learning algorithms can analyze data to identify patterns and establish relationships from complex datasets. For this project, machine learning would allow for HLA sequence data analysis and make a prediction model. By training the model on datasets of HLA sequences and peptide binding affinities, the algorithm can predict these complexes with high accuracy, paving the way for personalized and targeted immunosuppression. There have been many studies that employ machine learning to predict organ rejection. However, those models focus on "whole" HLA mismatches, which do not account for HLA polymorphism or the peptide sequences. Therefore, by focusing on HLA sequences and peptides, a more accurate and robust model can be created to prevent organ rejection. This way, we can protect the patient *and* the organ from harm.

Section II: Specific Aims

The objective is to make a machine learning model that can predict rejection and provide specific targets that will cause rejection, given donor and recipient HLA sequences. The model will work by predicting the MHC-peptide complex on the donor organ by focusing on the specific HLA allele mismatches. Ideally, this model will use mismatches to provide information on targets for personalized immunosuppression. In a clinical setting, we imagine this model being used before an organ transplant

to analyze donor and recipient blood samples or peripheral blood mononuclear cells (PBMCs)[2]. Traditional HLA typing techniques can be used such as serotyping or molecular typing to identify HLA alleles for the donor and the recipient (Casey, 2023). This HLA typing data can then be inputted into the machine learning model which can identify HLA mismatches and predict the MHC-peptide complex, giving us more information on the potential immunosuppressive targets. The evidence that arises from this research can greatly impact the field of organ transplantation and immunology.

Our long-term goal is to decrease the risk of organ rejection without the need for non-discriminate immunosuppressors. The central aim of this proposal is to use HLA mismatches between the donor and the recipient to predict the donor peptides that are likely to bind to the recipient's HLA molecules, forming MHC-peptide complexes capable of triggering T-cell responses. Identified complexes could guide tailored immunosuppressive regimens or tolerance induction approaches. Previous models have attempted to predict organ rejection by using HLA mismatches. However, these models focus on antibody-mediated rejection (AMR) and do not provide specific immunosuppressive targets (Duquesnoy & Askar, 2007). While this research focuses on T-cell mediated rejection (TCMR), the results of previous studies can be used as a rationale for the objectives described in this proposal, which will be elaborated upon in the upcoming sections.

The work we propose here will use three specific aims as stages in developing the final model. The expected outcome of this work is to produce a machine-learning model that can effectively predict organ rejection using HLA typing data with at least 80% accuracy.

Specific Aim 1: Analyze and identify HLA mismatches in donor-recipient immunology profiles.

Specific Aim 2: Use the donor-recipient HLA profile to predict the MHC-peptide complex.

Specific Aim 3: Predict the risk of rejection and determine precise immunosuppressive targets

[2] PBMCs are isolated from blood samples and identified as any blood cell with a round nucleus (i.e. lymphocytes). All nucleated cells in an individual harbor the same genetic material, including the HLA genes (Viatte, 2023).

Section III: Project Goals and Methodology

Relevance/Significance

The supply and demand issue in the organ transplant industry is killing many patients every year. Along with the lack of organ donations, the risk of organ rejection negatively affects patients after receiving a transplant. In all transplant cases, it becomes a weighing of risks and benefits. On one side, the patient is experiencing severe organ failure and needs a new organ to live. On the other side, gaining a new organ has its own complications. Doctors use biopsies to assess the health of the transplanted organ. If there is evidence of rejection, the most common treatment is increasing the dosage of immunosuppressors (*Transplant Rejection*, n.d.). However, these immunosuppressors can increase the patient's susceptibility to other diseases and infections. Increasing the dosages of these harmful medications will only worsen the side effects. Antibody induction therapy has had promising results in decreasing organ rejection and is currently used in about 80% of kidney transplants to decrease rejection and improve graft function (Zaza et al., 2014). Additionally, it has proven to be more effective in reducing the frequency of acute rejection episodes compared to conventional immunosuppression. However, different mAbs have different effects on people depending on their genotype (Mahmud et al., 2010). Therefore, predicting rejection can help doctors administer the optimal mAb against rejection based on the donor and recipient's HLA sequences. Predicting rejection before the transplant occurs could also enhance compatibility matching. Accurately predicting donor-recipient compatibility based on the MHC-peptide complex could significantly improve transplant success rates and reduce the need for immunosuppressive drugs altogether.

Innovation

Previous research has found that HLA mismatches contribute to a greater chance of organ rejection. As such, transplants happen to minimize the number of HLA mismatches to promote allograft

tolerance. However, comparing "whole" HLAs does not account for the allele differences or the amino acid differences that can initiate rejection. By classifying mismatches down to the allele or amino acid level, it can give insight into the smaller differences that can initiate rejection. Those differences can be used to create specific, precise immunosuppressive treatments that would not have the dangerous side effects as broad immunosuppressors. Additionally, predicting the MHC-peptide complex that recipient T-cells will recognize can also provide a greater understanding of the rejection outcome. The PIRCHE-II model is an algorithm that predicts MHC class II peptides for donor and recipient HLA sequences. This successful model has been linked to better transplant outcomes (Unterrainer et al., 2021). However, the model focuses on peptide mismatches rather than HLA mismatches and specifically focuses on MHC class II. MHC class I plays a significant role in the immune processes, which is why this model will include MHC class I peptide presentation. While many servers aid in predicting peptide presentation such as HLAMatchmaker and NetMHCpan, they all have different prediction methods and are yet to be used in an organ transplant setting. Overall, more research is needed to develop better treatments against organ rejection to improve the health of organ transplant patients.

Methodology

To create the proposed model, the necessary data and servers must be collected. HLA allele sequence data must be collected that contains accurate data from donor-recipient HLA typing results and the rejection outcome. Preferably, the data should have similar HLA allele resolutions[3], to resemble the current serological HLA typing methods. The data will be stored and used for model testing and validation after the model is created. Additionally, allele sequence data will be taken from datasets such as IPD-IMGT/HLA from the European Bioinformatics Institute (EBI), which provides HLA allele-specific

[3] Allelic resolution is a measure of the level of detail in HLA typing results. It defines a single allele as a unique DNA sequence for the HLA gene (Nunes et al., 2011). For example, a 2-field resolution for the HLA-A gene might be A02:01, while the 4-field resolution might be A02:01:01:01.

information such as HLA sequences and known peptide-binding motifs. The data will be used to find specific amino acid mismatches between HLA allele sequences between the donor and the recipient data, and only solvent-accessible amino acid mismatches will be stored. Solvent-accessible amino acids are exposed to the surrounding solvent. These exposed mismatches have a higher chance of being recognized by the recipient's immune system as "foreign", potentially triggering an immune response (Kramer et al.). Therefore, considering these mismatches can narrow down the number of features the model considers, making it more efficient. After, public bioinformatic servers such as NetMHCpan and HLAMatchmaker will be used to predict the MHC-peptide complex. These servers will be used as a guideline to create the machine-learning model, as NetMHCpan uses Artificial Neural Networks, and HLAMatchmaker uses experimental data. Models will be created using different machine learning architectures, such as Support Vector Machine, K-Nearest Neighbor, Random Forest, and other classification frameworks. The models will be tested using the collected HLA sequence data to assess the accuracy of the model algorithm. If necessary, the model will be modified to increase its accuracy. In the end, an open-source web application can be made that predicts the degree of rejection using MHC-peptide prediction and gives insight into the specific targets that could lead to personalized immunosuppression.

Specific Aim #1

Analyze and identify HLA mismatches in donor-recipient immunology profiles.

Justification and Feasibility. A study by Manski et al., 2019 created an algorithm that predicted graft survival at the one-year and five-year survival marks. However, their data only included data for three out of six HLAs and included the Kidney Donor Profile Index (KDPI) as a rejection factor. Table 1 contains the predictions for

Table 1: Manski et al., 2019 created a partial algorithm that compared three HLA sequences. Below is a table that predicts the one-year survival rate with different mismatched and KDPI scores.

	KDPI = 25	KDPI = 50	KDPI = 75
DR mm = 0	0.97	0.96	0.89
	[0.91, 1]	[0.85, 1]	[0.80, 0.96]
DR mm = 1	0.97	1.00	1.00
	[0.90, 1]	[1, 1]	[1, 1]
DR mm = 2	0.92	0.96	0.94
	[0.84, 0.98]	[0.87, 1]	[0.90, 0.98]

different KDPI scores and a different number of HLA-DR mismatches (mm) with the confidence interval in brackets. Nonetheless, the decrease in survival with a greater number of HLA-DR allele mismatches shows a clear correlation between kidney survival and HLA matching.

Summary of Preliminary Data. Sample HLA alleles for the donor and the recipient were used to compare amino acid sequences from the IPD-IMGT/HLA database. As seen in Table 2, the sample alleles in the HLA-B locus are displayed in the FASTA format. NetSurfP is a server that can provide solvent-accessible predictions based on the FASTA format of HLA alleles. The HLA Epitope Mismatch Algorithm is a model that compares donor and recipient amino acid sequences to find mismatches. After verifying the results in HLA-EMMA, it was evident that the proposed method was accurate and showed evidence that these HLA-B allele combinations have a low immunogenicity and a low risk for rejection (Kramer et al.).

Table 2: Modeled After the HLA Epitope Mismatch Algorithm (HLA-EMMA) Format, Table Comparing Sample Recipient and Donor Alleles. Displays Amino Acid Mismatches and Solvent Accessible Mismatches in FASTA Format.

Info	Allele	33	91	93	94	95	138	176	180	202
Recipient	B*08:01	D	F	T	N	T	N	V	D	T
Recipient	B*40:02	H	S	T	N	T	N	V	L	T
Donor	B*07:02	Y	Y	A	Q	A	D	E	R	K
Total AA MM	9	Y	Y	A	Q	A	D	E	R	K
Solvent Accessible MM	3	---	---	---	Q	---	---	E	---	K

Expected Outcomes. The overall outcome of this aim is to develop a successful model that can analyze the allele mismatches between the donor and the recipient. Afterward, the mismatch information can be used to assess the intensity of rejection, where a greater number of mismatches indicates a greater risk for rejection.

Potential Pitfalls and Alternative Strategies. It is possible that the donor and recipient HLA allele sequence data are not all based on the same resolution. However, based on the dataset that is

acquired, multiple trials can be run to assess the model's accuracy at different resolutions. In this case, the dataset will be split into smaller subsets with each one having the same HLA typing data resolution.

Specific Aim #2

Use the donor-recipient HLA profile to predict the MHC-peptide complex.

Justification and Feasibility. The Predicted Indirectly Recognizable HLA Epitopes (PIRCHE-II) model was developed to predict T-cell-related immune responses against donor HLA-derived peptides. The PIRCHE-II algorithm calculates the number of theoretical HLA-epitopes that can cause an indirect alloreactive response that involves CD4+ T-cell recognition of HLA class-II presented donor HLA-peptides (Spitznagel et al., 2022). While the PIRCHE-II model counts the number of epitope mismatches rather than HLA mismatches, Figure 2 shows a direct relationship between high PIRHCE-II scores and a greater number of mismatches. As HLA mismatches are considered the main cause of rejection, the PIRHCE-II model provides a rationale to focus on MHC-peptide complexes to predict rejection.

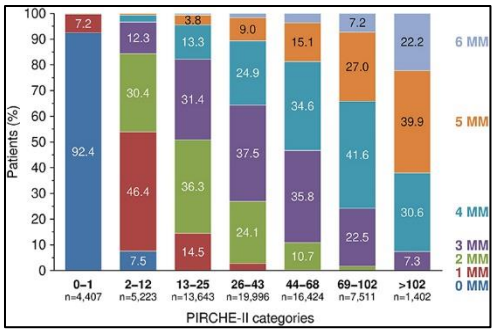


Figure 2: The relationship between PIRCHE-II score ranges and the proportion of patients with HLA mismatches (MM). A greater PIRCHE-II score indicates a greater number of HLA mismatches (Unterrainer et al., 2021).

Summary of Preliminary Data. The same sample donor and recipient HLA-B alleles were used in Table 2 to generate peptides using donor sequence. The sequence was inputted into NetMHCpan, which generates peptides and predicts the binding affinity to the recipient HLA alleles. The threshold for a peptide to have a strong binding to 0.5. As seen in Figure 3, there are very few donor-specific peptides that have a strong

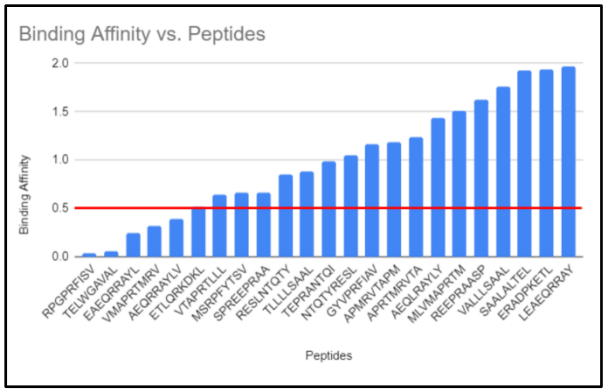


Figure 3: Histogram of Donor-Generated Peptides with Significant Binding Affinities. Peptides with Binding Affinity Below Red Line Have a String Binding Affinity. Peptides Above the Red Line Have Weak Binding Affinity. Peptides With Binding Affinity Greater Than 2 Were Not Considered as Significant Peptides.

binding to recipient HLA molecules. This means that there is a low chance of donor peptides binding to recipient HLA molecules, which demonstrates a low immunogenicity. Once again, this procedure confirms the results of the previous aim and can be used to analyze other HLA alleles as well.

Expected Outcomes. The expected outcome of this aim is to use HLA mismatches between the donor and the recipient to predict the MHC-peptide complex that could initiate rejection. Understanding the specific peptide can provide insight into better medications or rejection therapies.

Potential Pitfalls and Alternative Strategies. One concern is to make sure the model is not overfitting to training data. If the model is overly complex and captures noise in the training data, it may overfit and perform poorly on new, unseen data. To decrease the possibility of overfitting, multiple servers and a substantial amount of peptide prediction data will be used.

Specific Aim #3

Predict the risk of rejection and determine precise immunosuppressive targets.

Justification and Feasibility. In a recent study, Spitznagel et al., 2022 analyzed how PIRHCE-II scores relate to the immunological risk and histopathological changes of immune-related injury in kidney transplants. After doing multiple biopsies, they found that kidney transplant recipients with higher PIRCHE-II scores were more likely to develop TCMR in the follow-up biopsy. Additionally, these differences were significant for both HLA-class I and HLA-class II-derived PIRCHE-II scores. Figure 3 further shows lower PIRCHE-II scores being associated with no rejection or borderline rejection, while the higher scores were associated with graft rejection. The study confirms the rationale for using MHC-peptide complexes to assess the risk of rejection.

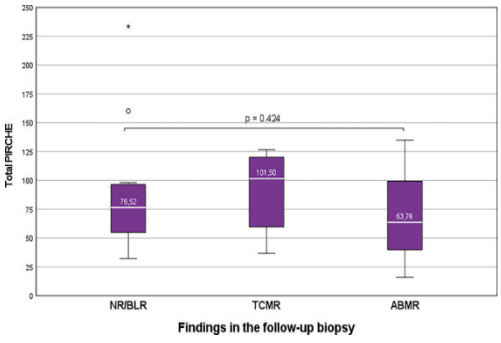


Figure 4: Boxplots show higher PIRCHE-II scores among kidney transplant recipients who show T-cell mediated rejection (TCMR), no rejection/borderline rejection (NR/BLR), and antibody-mediated rejection (AMR) (Spitznagel et al., 2022).

Summary of Preliminary Data. A rudimentary K-Nearest-Neighbor (KNN) algorithm was developed for T-cell mediated rejection (TCMR) versus no rejection samples from a Gene Expression Omnibus (GEO) data set. The KNN algorithm is a classification algorithm that makes predictions by finding pattern groups in a dataset. A flowchart of this model is provided in Appendix 2. With of accuracy of 80% (Figure 4), this initial algorithm can be expanded upon to find specific rejection targets for each sample. The most significant features of the KNN algorithm were multiple cytokine genes such as CCL2 and CCL15. Cytokines are active participants in T-cell-mediated rejection, giving a rationale to look at the rejection processes for potential targets for immunosuppression.

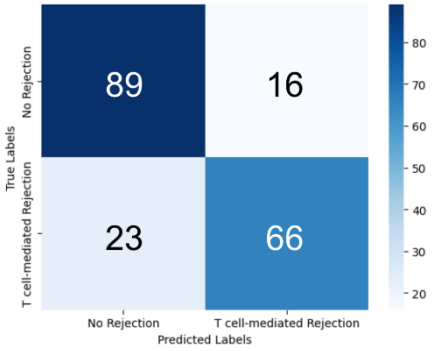


Figure 5: Confusion Matrix of K-Nearest Neighbor Algorithm Using Gene Expression Omnibus (GEO) Dataset GSE212160.

Expected Outcomes. The overall outcome of this aim is to assess the risk of rejection for a given donor and recipient HLA combination and provide specific targets for immunosuppression based on the predicted MHC-peptide complex. In the end, the model can be used to enhance donor-recipient matching and decrease the need for broad immunosuppressors.

Potential Pitfalls and Alternative Strategies. We expect the most difficult aspect of this aim to be finding data. Using accurate data is crucial to testing the accuracy of the model, which is why the data must be as accurate and detailed as possible. If proper data is not found through literature or database searches, we will likely contact medical professionals who have access to recipient records and create a model based on those records.

Section IV: Resources/Equipment

The algorithm of this project will be done *in silico*, so the necessary equipment would require a computer, data, and the necessary software. The main programming language that will be used for the machine-learning model will be Python. With its comprehensive library ecosystem, libraries such as

Matplotlib, numpy, and more can be used to visualize and analyze data. Google Collaboratory will be used to write and execute Python scripts within the browser. R will be the second programming language that is used for data analysis and visualizing data. The BioConductor and GEOquery packages make it easy to analyze gene expression data. Lastly, all datasets and data files will be processed and converted into a table-like format that can be readable as an Excel file. As the United Network for Organ Sharing (UNOS) requires a Statistical Analysis System (SAS) to read its dataset, the SAS software will be used solely to read and analyze the UNOS patient data. Databases such as Gene Expression Omnibus (GEO) and the Immune Epitope Database (IEDB) will be searched for HLA sequence data. Servers such as IPD-IMGT/HLA, HLAMatchmaker, NetSurfP and NetMHCpan will be used as resources to build the MHC-peptide prediction model.

Section V: Ethical Considerations

As this project will mainly use computational and programming software to fulfill the main objectives, there are no major safety concerns. Regarding the machine learning model itself, one important consideration is the potential for bias. Bias refers to the systematic errors or unfairness built into a model, leading to inaccurate predictions for certain groups or situations (Buhl, 2023). Bias will be mitigated by using multiple data analysis techniques such as external validation and using multiple different types of models. Concerning the data, all the data used in the model will come from publicly available databases such as Gene Expression Omnibus (GEO) and NetMHCpan or the United Network for Organ Sharing (UNOS). UNOS contains STAR File datasets that have transplant data since 10/1/1987. These files maintain the anonymity of patients by removing personal and identifiable information. When using the UNOS dataset, the project will comply with the terms of the signed Organ Procurement and Transplantation Network (OPTN) Data Use Agreement contract. All datasets used in this project will not include any details that can be traced back to specific individuals, and all datasets will be properly cited.

Section VI: Timeline

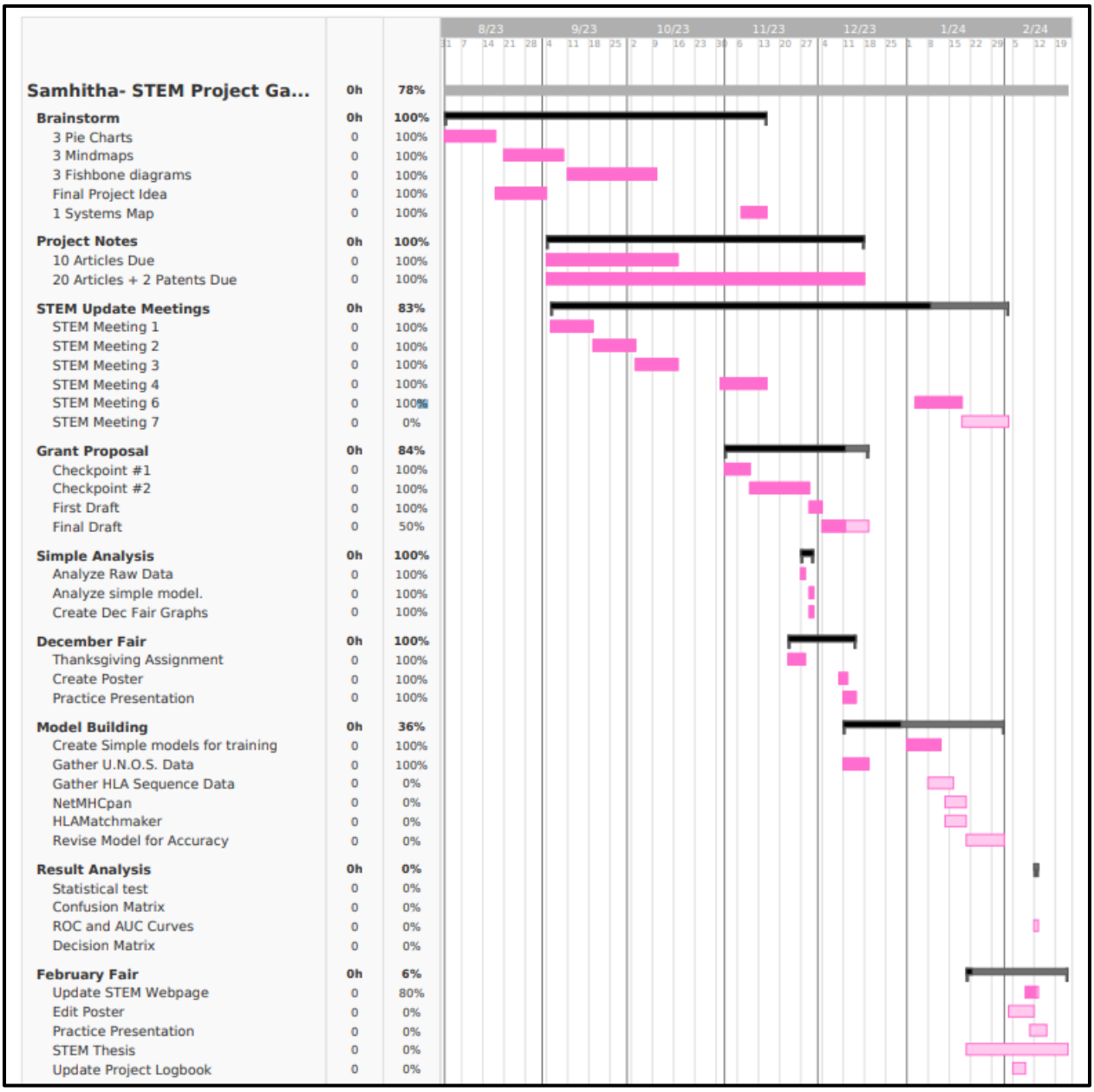


Figure 6: The Gantt Chart is a rough timeline of tasks for the project until the February Fair which is on February 15, 2023.

Section VII: Appendix

Appendix 1: Description of Grant Application

Description of RFA-AI-10—19, Genomics of Transplantation Cooperative Research Program (U01, U19) Grant from the National Institute of Allergy and Infectious Diseases (NIAID), a program of the National Institute of Health (NIH).

Purpose:

The National Institute of Allergy and Infectious Diseases (NIAID) invites new or renewal applications from institutions or consortia of institutions to participate in a cooperative interdisciplinary research program for large-scale, broad-scope genomic studies in clinical transplantation of solid organs, tissues, and cells. The goals of this program are to identify and characterize gene polymorphisms and gene expression patterns that: (1) correlate with and/or predict transplantation outcomes; (2) define immune responses relating to the onset and severity of acute and chronic graft rejection; (3) predict responses to immunosuppressive intervention to allow tailoring of therapy; and (4) elucidate the genetic basis of variation in graft survival between individuals and/or populations. The long-term goal of the program is to understand the genetic basis of immune-mediated graft rejection and differences in transplant outcomes, and thereby provide a rational basis for the development of more effective treatments, and prevention strategies to improve long-term graft survival and provide a better quality-of-life for transplant recipients.

This Funding Opportunity Announcement (FOA) will utilize the single-project (U01), and multi-project (U19) Cooperative Agreement award mechanism.

Appendix 2: Model Flow Chart

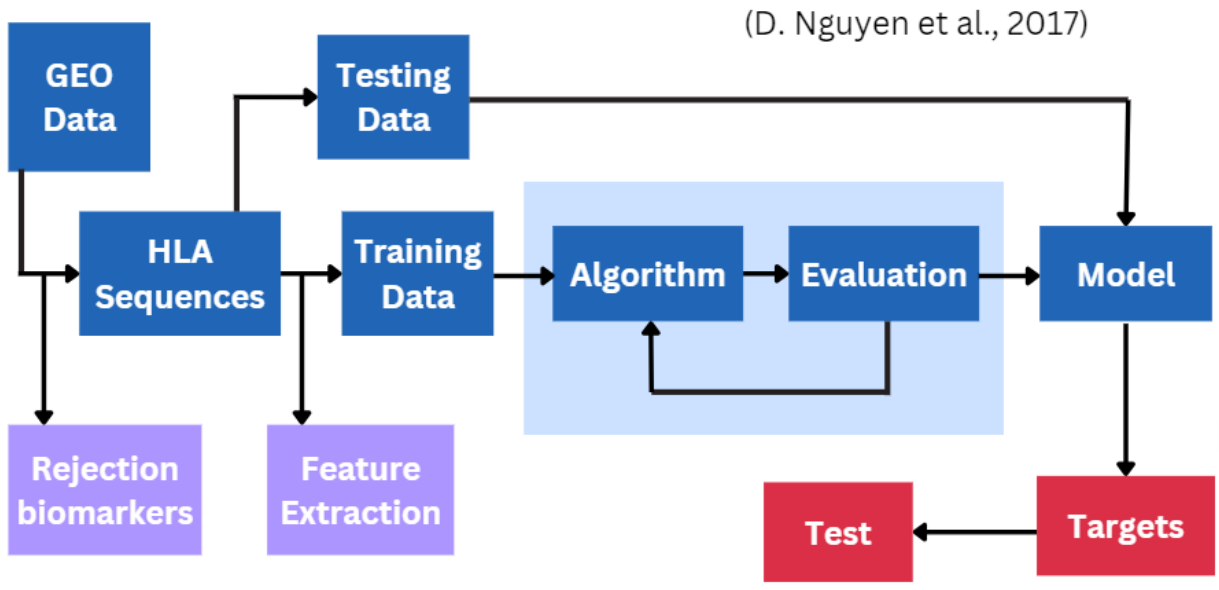


Figure 7: Flowchart of machine learning model that takes GEO data to predict rejection, inspired by D. Nguyen et al., 2017

Section VIII: References

Casey, R. (2023, May 9). A Quick Guide to Human Leukocyte Antigen (HLA) Typing Techniques. The Sequencing Center. https://thesequencingcenter.com/a-quick-guide-to-human-leukocyte-antigen-hla-typing-techniques/

Chen, H., Yang, J., Zhang, S., Qin, X., Jin, W., Sun, L., Li, F., & Cheng, Y. (2019). Serological cytokine profiles of cardiac rejection and lung infection after heart transplantation in rats. *Journal of Cardiothoracic Surgery*, *14*(1), 26. https://doi.org/10.1186/s13019-019-0839-5

Chong, A. S. (2020). B cells as antigen-presenting cells in transplantation rejection and tolerance. *Cellular Immunology*, *349*, 104061. https://doi.org/10.1016/j.cellimm.2020.104061

Costa, C. D. (2019, August 26). What Is Machine Learning & Deep Learning? Medium. https://medium.com/@clairedigitalogy/what-is-machine-learning-deep-learning-7788604004da

Dasgupta, A. (2016). Chapter 2 - Limitations of immunoassays used for therapeutic drug monitoring of immunosuppressants. In M. Oellerich & A. Dasgupta (Eds.), *Personalized Immunosuppression in Transplantation* (pp. 29–56). Elsevier. https://doi.org/10.1016/B978-0-12-800885-0.00002-3

Duquesnoy, R. J., & Askar, M. (2007). HLAMatchmaker: A Molecularly Based Algorithm for Histocompatibility Determination. V. Eplet Matching for HLA-DR, HLA-DQ, and HLA-DP. *Human Immunology*, *68*(1), 12–25. https://doi.org/10.1016/j.humimm.2006.10.003

Gautreaux, M. D. (2017). Chapter 17 - Histocompatibility Testing in the Transplant Setting. In G. Orlando, G. Remuzzi, & D. F. Williams (Eds.), *Kidney Transplantation, Bioengineering and Regeneration* (pp. 223–234). Academic Press. https://doi.org/10.1016/B978-0-12-801734-0.00017-5

General, Non-specific Defenses Against Infection. (n.d.). Defense Mechanism. Retrieved January 26, 2024, from https://sphweb.bumc.bu.edu/otlt/mph-modules/ph/ph709_defenses/ph709_defenses_print.html

Geneugelijk, K., & Spierings, E. (2020). PIRCHE-II: an algorithm to predict indirectly recognizable HLA epitopes in solid organ transplantation. Immunogenetics, 72(1), 119–129. https://doi.org/10.1007/s00251-019-01140-x

Grinyo, J. M. (2013). Why Is Organ Transplantation Clinically Important? *Cold Spring Harbor Perspectives in Medicine*, *13*(11). https://doi.org/10.1101/cshperspect.a014985

Hewitt, E. W. (2003). The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology*, *110*(2), 163–169. https://doi.org/10.1046/j.1365-2567.2003.01738.x

Hunt, D., & Saab, S. (2012). Post–Liver Transplantation Management - ScienceDirect. In *Zakim and Boyer's Hepatology* (Sixth, pp. 869–882). https://www.sciencedirect.com/science/article/abs/pii/B9781437708813000498

Ingulli, E. (2010). Mechanism of cellular rejection in transplantation | Pediatric Nephrology. *Pediatric Nephrology*, *25*. https://doi.org/10.1007/s00467-008-1020-x

Kelly, J. (2022, April 27). *End of anti-rejection transplant drugs? A clinical trial at Hume-Lee hopes so.* VCU Health. https://www.vcuhealth.org/news/end-of-anti-rejection-transplant-drugs-a-clinical-trial-at-hume-lee-hopes-so

Kramer, Cynthia S. M., et al. "HLA-EMMA: A User-friendly Tool to Analyse HLA Class I and Class II Compatibility on the Amino Acid Level." *Hla*, vol. 96, no. 1, July 2020, pp. 43–51, https://doi.org/10.1111/tan.13883

Manski, C. F., Tambur, A. R., & Gmeiner, M. (2019). Predicting kidney transplant outcomes with partial knowledge of HLA mismatch. *Proceedings of the National Academy of Sciences*, *116*(41), 20339–20345. https://doi.org/10.1073/pnas.1911281116

Matching and Compatibility. (n.d.). UC Davis Health. Retrieved November 8, 2023, from https://health.ucdavis.edu/transplant/livingkidneydonation/matching-and-compatibility.html

Nunes, E., Heslop, H., Fernandez-Vina, M., Taves, C., Wagenknecht, D. R., Eisenbrey, A. B., Fischer, G., Poulton, K., Wacker, K., Hurley, C. K., Noreen, H., & Sacchi, N. (2011). Definitions of histocompatibility typing terms. *Blood*, *118*(23), e180–e183. https://doi.org/10.1182/blood-2011-05-353490

Organ, Eye and Tissue Donation Statistics. (n.d.). Donate Life America. Retrieved November 8, 2023, from https://donatelife.net/donation/statistics/

Organ Rejection after Renal Transplant. (n.d.). Columbia Surgery. Retrieved November 8, 2023, from https://columbiasurgery.org/kidney-transplant/organ-rejection-after-renal-transplant

Reits, E., & Neefjes, J. (2022). HLA molecules in transplantation, autoimmunity and infection control: A comic book adventure. *Hla*, *100*(4), 301–311. https://doi.org/10.1111/tan.14626

Spitznagel, T., Matter, L. S., Kaufmann, Y. L., Nilsson, J., von Moos, S., & Schachtner, T. (2022). PIRCHE-II scores prove useful as a predictive biomarker among kidney transplant recipients with rejection: An analysis of indication and follow-up biopsies. *Frontiers in Immunology*, *13*, 949933. https://doi.org/10.3389/fimmu.2022.949933

Understanding Transplant Rejection. (n.d.). Stony Brook Medicine. Retrieved November 8, 2023, from https://www.stonybrookmedicine.edu/patientcare/transplant/rejection

Unterrainer, C., Döhler, B., Niemann, M., Lachmann, N., & Süsal, C. (2021). Can PIRCHE-II Matching Outmatch Traditional HLA Matching? *Frontiers in Immunology*, *12*. https://www.frontiersin.org/articles/10.3389/fimmu.2021.631246

Viatte, S. (2023, September 8). Human leukocyte antigens (HLA): A roadmap - UpToDate. Uptodate. https://www.uptodate.com/contents/human-leukocyte-antigens-hla-a-roadmap/print

Vijayan, S., Sidiq, T., Yousuf, S., van den Elsen, P. J., & Kobayashi, K. S. (2019). Class I transactivator, NLRC5: a central player in the MHC class I pathway and cancer immune surveillance. Immunogenetics, 71(3), 273–282. https://doi.org/10.1007/s00251-019-01106-z

Why Is Organ Donation Important? (2022, March 29). INTEGRIS Health. https://integrisok.com/resources/on-your-health/2022/march/why-is-organ-donation-important

Zaza, G., Tomei, P., Granata, S., Boschiero, L., & Lupo, A. (2014). Monoclonal Antibody Therapy and Renal Transplantation: Focus on Adverse Effects. *Toxins*, *6*(3). https://doi.org/10.3390/toxins6030869