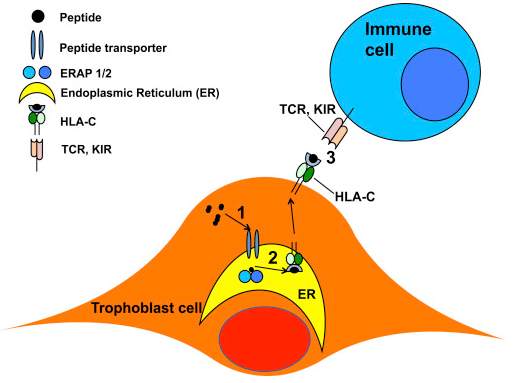
***Genetically Modifications on ERAP2 and its Effect on Lymphocyte Activation***

**I. Introduction**

According to Mayo Clinic, a spontaneous miscarriage is a loss of a fetal life before the 20th week. It is the most common type of pregnancy loss, with more than 3 million cases per year in the US alone. One of the reasons why this devastating event happens to so many women can be because of immune system imbalances.9 The immune system is a complex unit with many parts that work together constantly to keep the human body healthy. Lymphocytes (T cells, natural killer cells and helper T cells) are cells which produce antibodies. Antibodies are tiny proteins that attach on to antigens, which are proteins on the surface of cells. These antibodies serve as flags for the lymphocytes that that signal them become activated.9 Lymphocyte activation happens when an antibody on a lymphocyte is activates by the antigen on another cell.9 These antigens are presented on the outside of the cell by a major histocompatibility complex (MHC). These complexes are found on every cell in the body.9 Epitopes are the peptides that are presented on the MHC and recognized by the antibodies and lymphocytes. They have to be a specific peptide sequence in order for the activation to happen.9

NKCs are the most abundant lymphocytes (70%) during early gestation and decrease greatly after pregnancy.10 The balance of these lymphocytes and the fetal cells needs to be very precise in order to keep the baby healthy9. The lymphocytes secrete antibodies that could be detrimental to the fetal health.9 An excess of the antibodies CD16 and CD56 (which denote NKCs) in the peripheral cells of the mother’s body can lead to activation by lymphocytes and release tumor neurosis factors (TNF) that can destroy the placenta and therefore hurting the fetus’ chances of survival.9 Women with natural killer cells (NKC) with CD16+ and CD56+ in excess of 20% are at risk of miscarriage because of the TNF9. Activation of different types of lymphocytes can be determined by antibodies. For example, CD16 and CD56 are used by NKCs and become activated. CD8 and CD4 are used by T cells to become active. And CD69 are used by both for early activation, which is needed for them to be active9.

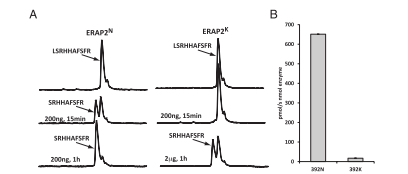
Endoplasmic Reticulum Aminopeptidase 2 (ERAP2) is an enzyme that cleaves peptides in the cell and determines the epitopes are that are presented on the MHC class-1.2 ERAP2 receives pre-cleaved peptides that are random amino acid sequences from protostomes (enzymes that cleave proteins) in the cytosol from the transporter associated with antigen processing (TAP), an ATP-driven transporter that transports cleaved peptides, from the proteasome to the ER.5 ERAP1 and ERAP2 both work together to cleave the random peptides into 8-10 amino acid sequences, which get presented to the MHC. 2 This pathway is illustrated in **figure 1**.



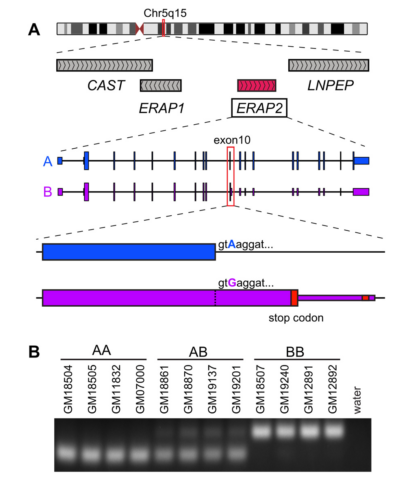
**Figure 1:**10 **The role of ERAP2 in the immune system.** 10

ERAP2 has two alleles which code different proteins: ERAP2K and ERAP2N.2 The difference between them is a difference in one amino acid at the 392nd position.2 ERAPK has Lysine and ERAP2N has asparagine.2 Lysine is a basic, polar amino acid where as asparagine is neutral and polar.

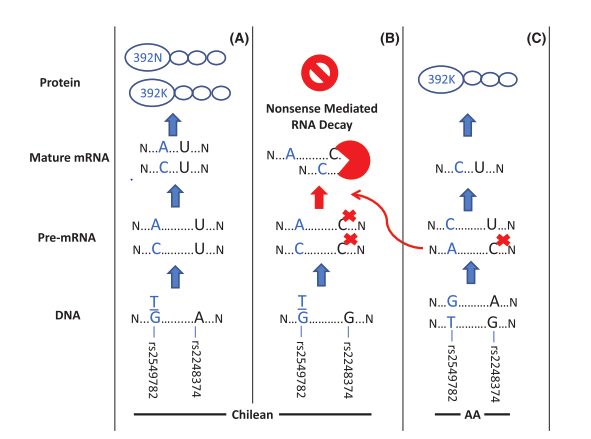
Although both of the alleles cleave at positively charged N terminal amino acids, the change from lysine and asparagine at the 392 position results in 165-fold greater activity for cleaving.2 The working assumption for why this hyper cleaving happens is that the asparagine results in results in a specificity switch with the peptides that it is trimming.2 This phenomenon is described as the largest functional changes described for a common coding polymorphism.2 This major change in cleaving activity could have a direct effect in the immune response, effecting the rate at which the immune system detects the epitopes. This can have detrimental effects to the fragile balance of the immune system during pregnancy.

**figure 4:2 RP-HPLC chromatograms of products of the enzymatic digestion of peptide LSRHHAFSFR (substrate) by the two ERAP2 variants.**

Although an allele, ERAP2N is not normally expressed in nature.4 The genetic difference between these two alleles is determined by single nucleotide polymorphism (SNP), which is a difference in one nucleotide in the DNA sequence.4 There are two SNPs that are in linkage disequilibrium (LD), which means that one SNP effects the other. 4 They are rs-2248374 and rs2549782. SNP rs2549782 has a switch from A to T. Changing it from A to T changes the codon of mRNA from AA[G] (which codes for lysine) to AA[U] (which codes for asparagine).4,2 If SNP rs2549782 is A, it is the natural state and the other SNP does not affect the exon at all (haplotype A in figure 2).4 Haplotype A’s exon is the correct size and can go on to translation. When rs2549782 is T, however, SNP rs2248374 elongates the exon by 56 nucleotides because splice site at position 69 is skipped.6 This elongation makes it possible for two stop codons to be transcribed downstream (as shown in figure 3). The ERAP2 mRNA from haplotype A produces a full length ERAP2K protein consisting of 960 amino acids.6 However, the mRNA from haplotype B is predicted to produce a truncated protein of 534 amino acids because of the premature stop codons.6 This is shown part C on figure 4. The stop codon on haplotype B is shown in figure 3.

**Figure 36: A) the two haplotypes, A: which expresses ERAP2K and haplotype B: which is supposed to express ERAP2N, but is instead a nonsense mutation due to SNP rs-2248374. And B) the lack of ERAP2 protein in haplotype B cells.**

It is proven that the alleles are in LD in most of the world.6 But there is one Chilean population that do not have the alleles linked.4 Theoretically, they should express ERAP2N, but they do not. The reason is unknown. One hypothesis of why ERAP2N is not seen in either fetus or mother’s cells is because the hyper trimming capability of the N allele presents epitopes to the MHC at a faster rate, which then induces more of a response from the mother’s immune system, (i.e. activating more lymphocytes) and the uterus becomes a hostile, fatal environment, resulting in termination of trophoblast cell or fetal survival.10 This would also explain why there is an LD in most of the world’s population, because it is an evolutionary adaptation that came about because of the effects of ERAP2N on fetal health.4



**Figure 3:4 C): The linkage seen in most populations vs. the lack of linkage in Chilean population (A, B). This is a prediction of what would theoretically happen since there is no linkage, but in the same study they found out that ERAP2N still is not found. Leaving only K at the protein stage in scenario (A).**

Mutations happen all the time in cells, and nothing stops them from happening on SNPs. If we can change the SNP rs-2248374 to A so that ERAP2N is expressed in the cell, it would show that it is possible to override the LD on SNP-rs2248374 from a point mutation (possible in nature). Also, if we can see the correlation between ERAP2N, and the number of activated lymphocytes that it produces, we can see the effect that the aminopeptidase has on the immune system and ultimately if ERAP2 can be a factor in immune induced miscarriages. This experiment is going to test the expression of ERAP2N in JEG-3 cells through genetic alterations and the effect it has on the amount of activated NKCs and T cells. I am going to achieve this by preforming Crispr Cas9 mediated mutagenesis, western blot, and a lymphocyte activation assay.

**II. Experiment**

The aim of this experiment is to genetically modify SNP rs-2248374 to induce expression of ERAP2N in JEG-3 cells and the to determine the number of activated lymphocytes. My hypothesis is that ERAP2N can be induced by genetic alteration, and transcribing it to a working enzyme would present more epitopes, leading to more activated lymphocytes. I am going to achieve this by preforming Crispr Cas9 mediated mutagenesis, western blot, and a lymphocyte activation assay.

*Cell Line*

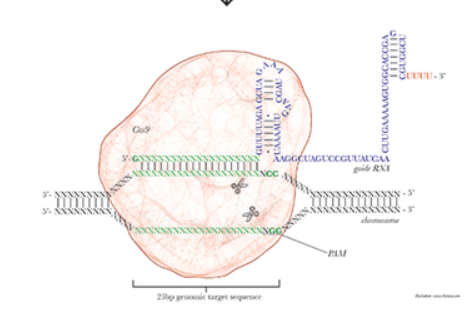
Jeg-3 cell lines are going to be used in this experiment. It is a type of choriocarcinoma cell that are trophoblast cells which have both TAP (transporter protein) and MHC-Class one.13 They have the genotype for ERAP2N (SNP rs2549782 homozygous TT).11 And the SNP that suppresses the transcription of ERAP2N, (SNP rs-2248374 homozygous GG)11. Because they have both homogenous SNPs, the heightened genotype for the N allele causes the SNP rs-2248374 to induce premature mRNA decay. So natural JEG-3 cells do not express ERAP2N, but have the genes for it.11 This makes them an ideal cell line to do gene editing in the SNP rs-2248374.

IIA. Gene editing

*IIAa. Crispr Gene Editing*

The CRISPR Cas 9 editing is going to heighten transcription of ERAP2N by using single nucleotide replacement to change SNP rs-2248374 from G to A, which should induce the transcription of ERAP2N because the SNP would not be able to cause the premature mRNA decay. CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats20. It is a tool which is used by prokaryotes to protect themselves from viral DNA.20 When a virus penetrates a bacterium and inserts its own DNA, Cas9, an endonuclease, is able to cleave the foreign DNA.20 Cas9 can be induced into eukaryotic cells by transfection in order to make precise double stranded breaks in the DNA.

Parts of the Cas9 enzyme include a sequence of guide RNA, and crRNA. The guide RNA is the same for all Cas9 enzymes.20 The crRNA, however, is a 20nt RNA sequence that is specific to the target region.20 It should be directly complimentary to the target DNA. The DNA sequence was found in BLAST and the sequence is in **Table 2.** The crRNA covalently binds to the guide RNA to form the sgRNA.20 The crRNA directs the Cas9 to cleave the targeted DNA sequence if they are adjacent to proto- spacer adjacent motifs (PAMs) towards the 5’ end (colored green in **Table 2**).20 This PAM sequence is NGG, the N is able to be replaced by any other nucleotide. 20 As shown in the nucleotide sequence of ERAP2 rs2248374, “TGG” is right after the SNP site.

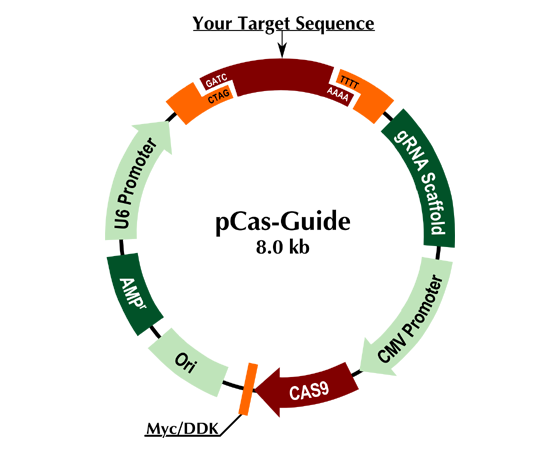


**Figure 424: green= crRNA, blue= tracrRNA. The PAM sequence is pointed out also**

*IIAb. Plasmid Construction*

A plasmid is a circular piece of DNA hat can be used to introduce new proteins and genes into a cell. A pCas-Guide template is going to be ordered from ORIGENE®25. The plasmid contains a U6 promoter for the target sequence, the gRNA scaffold, CMV promoter for the Cas9 enzyme, the Ori (point of transcription). It also has AMP, which is the gene for ampicillin, which makes the bacteria that we will be transforming in resistant to ampicillin for selection purposes. All of this is in **Figure 5.**

**Figure 5:25 the ORIgene® plasmid kit.**



After the plasmid is made, it needs to be amplified in bacteria before it gets transfected into the cell line. In order to transfer the plasmid into the cell, the FuGENE® method will be used to open up the membrane. And bacteria will be inserted into the solution.

*IIB: Transfection*

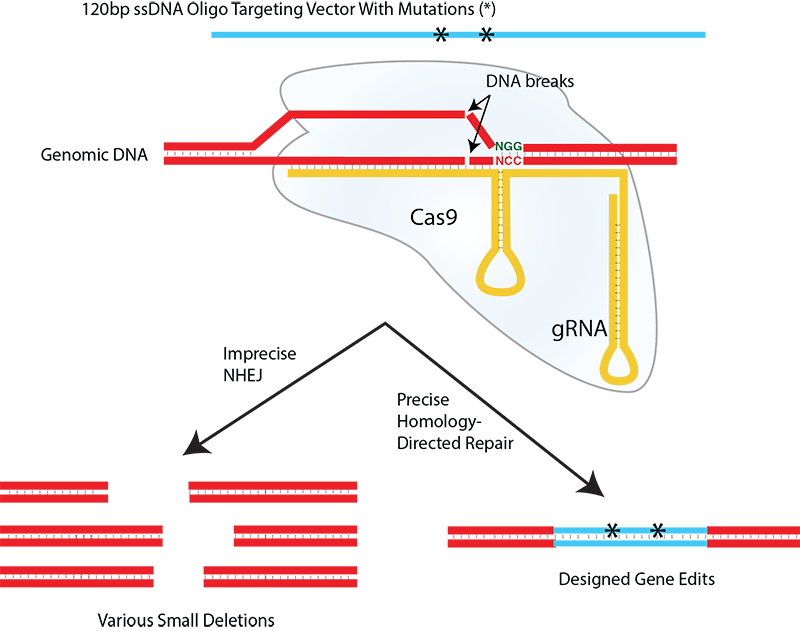
Transfection is the method of introducing nucleic acids into eukaryotic cells. In this experiment, the FuGENE® HD Transfection Reagent will be used. This is so the membrane of the JEG-3 cells can open up for the plasmid to enter through, it will then be incubated for 48 hours for the transfection to occur. During transfection in the JEG-3 cells, the primers should recruit its own RNA polymerase and ribosomes so it transcribes make Cas9, the sgRNA, and the ampicillin resistant gene.

*IIBa. Transfection of the Donor DNA*

In order for the Cas9 to switch the A to a T in the rs22 after the cleaving has taken place, it must go through homologous directed repair (HDR).23 This is when the double stranded break uses a template DNA in order to fill in the hole for the broken DNA.23 The donor DNA has to be introduced with the transfection of the Cas9 plasmid. The plasmid that I will use will consist of a primer sequence, **(table 1),** and the donor DNA sequence **(table 2)** this sequence will have the nucleotide T instead of A. The primer was checked using an online primer map.29 An empty vector will be ordered and EcoRI will be used to cleave it. And the DNA strand and the T7 promotors will be inserted. Using a DNA ligation protocol.27 This will also be transformed and amplified and then introduced into the JEG-3 cell line. **Figure 5** shows the different effects the non-homologous end joining and HDR have on the resulting DNA strand.

The different types of cells that I am aiming to obtain is JEG-3 with ERAP2N (by gene editing) and without (natural JEG-3 cells). The efficiency of the gene editing will be checked by western blot. SDS page will be used to look at the weight of the peptide. ERAP2 should be 110kDa4. Beta actin internal control will be also checked in order to keep the amount of protein the same.

**Figure 5: 26 insertions of nucleotides by HDR**



IIB. Lymphocyte Activation Assay

*IIBa.* *Isolation of lymphocytes*

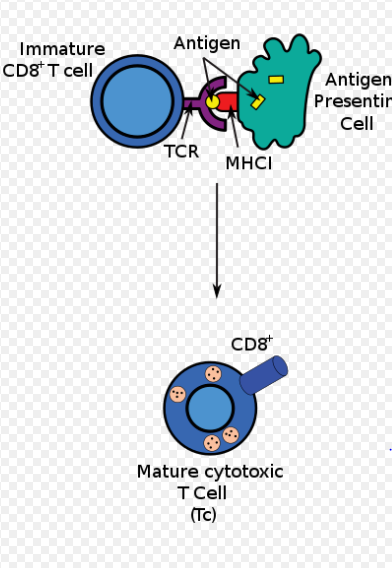
Isolation of lymphocytes needs to be done from freshly donated blood. The PBMC isolation protocol14 will be used to isolate lymphocytes based on density gradient centrifugation. The fresh blood with be centrifuges at 800gs for thirty-minute increments with addition of saline in between each centrifugation. The lymphocytes will be in the white fluffy layer after. NK cells will be used because they are found in abundance in pregnant women, and T cells to see if it induces any other immune response.

*IIBb. Insertion of antibodies*

The antibodies will be commercially bought with fluoresce already tagged on them. The antibodies that will be inserted are anti-CD56, CD16, CD69, CD8, and CD4. CD69 needs to be positive in order for the results to be valid because without eraly activation, the antibodies will have no effect.17

*IIBc. Quantification of Activated Lymphocytes*

After 48 hours of keeping the lymphocytes, antibodies, and JEG-3 cells together in incubation, the lymphocytes will be rinsed so to reduce the number of floating antibodies. The only antibodies that will be present after they are washed are the ones that are on the lymphocytes, activating them. Flow cytometry will be used to quantify activated lymphocytes. Flow cytometry works by releasing cells one by one on to a laser. The laser can measure the detentions of the cell we know which type of cell it is, and it can also measure how many fluorescence (antibodies) are tagged on to the lymphocytes. This will give us a quantitative result as to what type of lymphocytes were activated and how strong the activation was and with which antibody.

The activated lymphocytes (NKC and T Cells) will have CD\* protein on the surface of the cell like in **figure 6**. For NKCs, instead of TCR (T cell receptor) it is KIR (Killer-cell Immunoglobulin-like Receptors). Also shown in **Figure 1**.

**Figure 6:16 The activation of T cells from the bonding with MHC-C1**

**III. Discussion**

If this experiment goes well, it will show if an alteration of in SNP rs2248374 will result in the expression of hyper cleaving ERAP2N. This will show that although the genes are linked to not be expressed in nature, an alteration in the SNP (which is possible in nature) could make the expression of ERAP2N. ERAP2N has not been induced by gene alteration before, only by transfection. The possibility of the enzyme ERAP2N being expressed could then be one factor of spontaneous miscarriages that happen with no other warning. And would need to be further researched. If the gene alteration is successful, it will also show how the alleles have an effect on the immune system through activation of T cells and NKCs.

The amount of activation increasing would go for the broader hypothesis the genes that are not linked in Chilean population do not result in people expressing the enzyme (although they theoretically should) because the increase in lymphocytes caused by ERAP2N would result in hyperimmune defenses such as TNFs releasing, being fatal to the fetal survival.

One problem that I could come across is that although JEG-3 cell lines do not express ERAP2, it could have other aminopeptidases that are presenting to the MHC. Another complication is that the antibodies sometimes do not have the same affinity to bond when the lymphocytes are isolated instead of doing the experiment in actual blood. Since I would be isolating lymphocytes, the results might not be as strong. Also, HDR is known to not be as efficient as NHEJ. This experiment can be repeated with single stranded nicking to increase the efficiency. (20) An invitro transcription assay can also be used in order to check if the transfection worked.

In this proposed experiment, JEG-3 cell lines are used to see the effects of the different alleles clearer because they are homogenous for both SNPs. But another cell line that has heterozygous for the two SNPs can be used to see if the difference in immune response is as big. And to gear it towards miscarriages, cell line from fetal cells would be ideal.

If this experiment were to work, it would have many implications and would open op doors for further research. Further research needs to be done on the spontaneous miscarriages and ERAP2N. One experiment that would help the hypothesis is genotyping the mothers and the terminated fetal tissue for ERAP2N and K. This would show the simple association of the two.

Although spontaneous miscarriages are an awful phenomenon to happen to women and their families, advances in not only research, but also technology have lead us closer to the causes of it. We have come far, but there are still more generations of knowledge that is needed to fully assess spontaneous miscarriages.

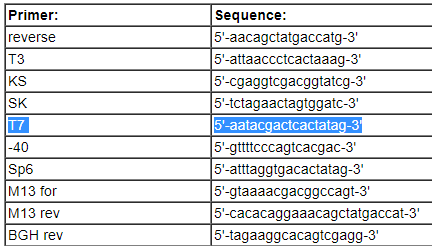
Funcüon 
K N [Asn] 
missense 
K N [Asn] 
missense 
K N [Asn] 
missense 
ncRNA 
K N [Asn] 
missense 
ncRNA 
ncRNA 
SNP to mRNA 
mRNA 
Accession 
NM 001130140.2 
Position 
1450 
1752 
1887 
1887 
1359 
1359 
1807 
Allele change 
Accession 
NP 001123612.1 
NP 001316158.1 
NP 071745.1 
XP 011541846_1 
Protein 
Position 
392 
347 
392 
392 
Residue change 
AA G 
AA T 
NM 00132g22g.1 
NM 0223504 
NR 1376371 
XM 011543544.1 
XR 001742179.1 
XR 948283.2 

**Figure 7:28 The of allelic and residue change for ERAP2 rs2549782. First row is position 392.**

**Table 1: Primer29**

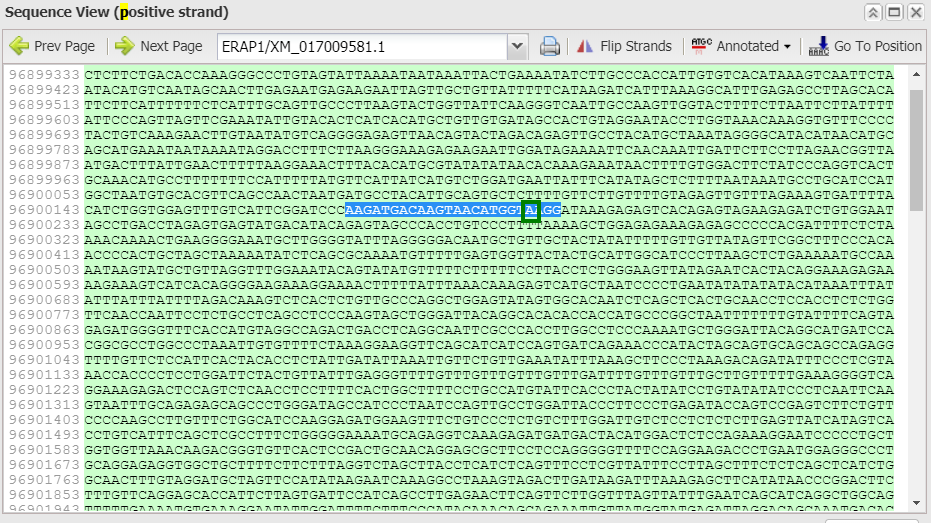
|  |  |
| --- | --- |
| T7 primer | F: aatacgactcactatag |

Other primers that can also be used:



**Table 2: Sequences28**

|  |  |
| --- | --- |
| crRNA G-Rs2248374 | F: AAGATGACAAGTAACATGGTGAGG  R: TCTGTGACTCTCTTTATCCTCACC |
| HDR DNA donor  A-rs2248374 | F:  ATCTGGTGGAGTTTGTCATTCGGATCCCAAGATGACAAGTAACATGGTAAGGATAAAGAGAGTCACAGAGTAGAAGAGATCTGTGGAAT  R: ATTCCACAGATCTCTTCTACTCTGTGACTCTCTTTATCCTTACCATGTTACTTGTCATCTTGGGATCCGAATGACAAACTCCACCAGATGT |



**Figure 8:28 SNPrs2248374 and 20 nts upstream highlighted**

**Citations**

1. Hill LD, Hilliard DD, York TP, Srinivas S, Kusanovic JP, Gomez R, et al. Fetal ERAP2 variation is associated with preeclampsia in African Americans in a case-control study. BMC Med. Genet. 2011a;12:64.
2. Evnouchidou I, Birtley J, Seregin S, Papakyriakou A, Zervoudi E, Samiotaki M, et al. A common single nucleotide polymorphism in endoplasmic reticulum aminopeptidase 2 induces a specificity switch that leads to altered antigen processing. J. Immunol. 2012;189:2383–2392
3. Nguyen, T. T., S. C. Chang, I. Evnouchidou, I. A. York, C. Zikos, K. L. Rock, A. L. Goldberg, E. Stratikos, and L. J. Stern. 2011. Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1. Nat. Struct. Mol. Biol. 18: 604–613
4. Vanhille DL, Hill LD, Hilliard DD, et al. A novel *ERAP2* haplotype structure in a Chilean population: implications for ERAP2 protein expression and preeclampsia risk. *Molecular Genetics & Genomic Medicine*. 2013;1(2):98-107. doi:10.1002/mgg3.13.
5. Cifaldi L, Romania P, Lorenzi S, Locatelli F, Fruci D. Role of Endoplasmic Reticulum Aminopeptidases in Health and Disease: from Infection to Cancer. *International Journal of Molecular Sciences*. 2012;13(7):8338-8352. doi:10.3390/ijms13078338.
6. Andrés AM, Dennis MY, Kretzschmar WW, et al. Balancing Selection Maintains a Form of *ERAP2* that Undergoes Nonsense-Mediated Decay and Affects Antigen Presentation. Gojobori T, ed. *PLoS Genetics*. 2010;6(10):e1001157. doi:10.1371/journal.pgen.1001157.
7. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*. 2013;154(2):442-451. doi:10.1016/j.cell.2013.06.044.
8. Rothbard, J. B., & Taylor, W. R. (1988). A sequence pattern common to T cell epitopes. *The EMBO Journal*, *7*(1), 93–100. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC454220/>
9. Colucci F, Boulenouar S, Kieckbusch J, Moffett A. How does variability of immune system genes affect placentation? *Placenta*. 2011;32(8):539-545. doi:10.1016/j.placenta.2011.05.001.
10. Lee D, Eun Endoplasmic Reticulum Aminopeptidase 2, a common immunological link to adverse pregnancy outcomes and cancer clearance? *Placenta.* 2017;(56)40-43 <https://doi.org/10.1016/j.placenta.2017.03.012>
11. Lee D, Eun Comparative expression profiling of endoplasmic reticulum aminopeptidase 2 and human leukocyte antigen-C expression in choriocarcinoma cell lines *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health* 2015;5(1)
12. Saric, Tomo Chang, Shih-Chung Hattori, Akira York, Ian A. Markant, Shirley Rock, Kenneth L. Tsujimoto, Masafumi Goldberg, Alfred L. An IFN-γ–induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I–presented peptides *Nature Immunology* 2002;3 <http://dx.doi.org/10.1038/ni859>
13. Easterfield, A.J., B.M. Austen, and O.M. Westwood, Inhibition of antigen transport by expression of infected cell peptide 47 (ICP47) prevents cell surface expression of HLA in choriocarcinoma cell lines. J Reprod Immunol, 2001. 50(1): p. 19-?‐40.
14. Lauren E. Higdon, PhD, Karim Lee, PhD, Qizhi Tang, PhD, and Jonathan S. Maltzman, MD, PhD Virtual Global Transplant Laboratory Standard Operating Procedures for Blood Collection, PBMC Isolation, and Storage *Transplantation Direct* 2016;2: e101; doi: 10.1097/TXD.0000000000000613
15. L. May, D. van Bodegoma, M. Kuningas, J.J. Meij, A.J.M. de Craen, M. Frölich, R.G.J. Westendorp Performance of the whole-blood stimulation assay for assessing innate immune activation under field conditions *Cytokine 45* (2009) 184–189 doi:10.1016/j.cyto.2008.12.010
16. <https://en.wikipedia.org/wiki/Cytotoxic_T_cell#/media/File:Antigen_presentation.svg>
17. Human and Mouse CD Marker Handbook *BD Biosciences* 2010
18. Scott M.SchaefferPaul A.Nakata CRISPR/Cas9-mediated genome editing and gene replacement in plants: Transitioning from lab to field *Plant Science* Volume 240, November 2015, Pages 130-142 <https://doi.org/10.1016/j.plantsci.2015.09.011>
19. Mallmann, P., Mallmann, R. & Krebs, D. Arch Gynecol Obstet (1991) 249: 73. https://doi.org/10.1007/BF02390365
20. Sander J, Joung J, Nature Biotechnology, CRISPR-Cas systems for editing, regulating and targeting genomes vol. 32, issue 4 (2014) pp. 347-350 Published by Nature Publishing Group
21. Price AA, Sampson TR, Ratner HK, Grakoui A, Weiss DS. Cas9-mediated targeting of viral RNA in eukaryotic cells. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(19):6164-6169. doi:10.1073/pnas.1422340112.
22. Transfection, *Protocols for transfection and information on optimization*
23. Song F, Stieger K. Optimizing the DNA Donor Template for Homology-Directed Repair of Double-Strand Breaks. Molecular Therapy Nucleic Acids. 2017;7:53-60. doi:10.1016/j.omtn.2017.02.006.
24. Church Lab CRISPR Plasmids Available from Addgene
25. ORIgene p-Cas9 cloning kit <http://www.origene.com/CRISPR-CAS9/Detail.aspx?sku=GE100001>
26. CRISPR/Cas9 <http://ko.cwru.edu/services/crispr_cas.shtml>
27. Genescript *DNA Ligation Protocol* <https://www.genscript.com/ligation-protocol.html>
28. BLAST ERAP2 <https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2248374>