The Function and Oncological Significance of PD-L1 and its Antibodies

Shreya Nagunuri[1], Ananya Garg[2] and Jagath Reddy[#]

[1]Granite Bay High School, Granite Bay CA, USA
[2]American High School, Fremont, CA, USA
[#]Advisor

ABSTRACT

The PD-L1 protein, also known as programmed death-ligand 1 is a protein encoded in the CD274 gene. Specifically, PD-L1 belongs to the immunoglobulin superfamily of proteins, and it is a transmembrane protein that allows nutrients across the cell membrane 11. PD-L1 works in close connection with T cells (thymus cells) and B cells (bone marrow or bursa-derived cells)1. PD-L1 binds to the PD-L receptor on T cells to regulate and sometimes inhibit (in the case of cancer) the activated T cells, B cells, and myeloid cells. Once their activation is inhibited by PD-L1, the T-cells are unable to fight foreign substances in the body like infections, diseases, and cancers, thus allowing cancerous tumors to grow without check3. Elevated levels of PD-L1 have been found in a variety of cancers, including melanoma, non-small cell lung cancer (NSCLC), Hodgkin's lymphoma, bladder cancer, renal cell carcinoma (RCC), head and neck squamous cell carcinoma (HNSCC), breast cancer, Merkel cell carcinoma, hepatocellular carcinoma (HCC) and gastric cancer (GC)4. Among other information, this study also examines the protein sequence of PD-L1, alignments of the sequence, the structure, functional domains, gene expression, copy number, and mutation profiles.

Introduction

The PD-L1 protein (also known as Programmed death-ligand 1) is encoded in the CD274 gene and is a transmembrane protein. It functions to act as a gateway to allow certain materials and nutrients across the membrane and it has receptors to signal other genes and proteins. PD-L1 belongs to the immunoglobulin superfamily of proteins which is known for its ability to recognize and adhere to other molecules. The immunoglobulin superfamily is a large protein superfamily of cell surface and soluble proteins that are involved in the recognition, binding, or adhesion processes of cells4. Molecules are categorized as members of this superfamily based on shared structural features with immunoglobulins (also known as antibodies); they all possess a domain known as an immunoglobulin domain or fold. It was first recognized by the Mayo Clinic as an immuno- regulatory molecule and it was first named B7-H1. Later researchers determined that it was a ligand of PD-L so they renamed it PD-L1. Generally, PDL1 works in close connection with T cells and B cells. T cells (thymus cells) and B cells (bone marrow- or bursaderived cells) are the major cellular components of the adaptive immune response. T cells are involved in cell mediated immunity, whereas B cells are primarily responsible for humoral immunity (relating to antibodies)3. The PD-L1 binds to the PD-L to regulate and sometimes inhibit the activated T cells, B cells, and myeloid cells. T and B cells are activated when they recognize small components of antigens, called epitopes. An antigen is a macromolecule that reacts with components of the immune system3. A given antigen may contain several motifs that are recognized by immune cells4. The expression of PD-L1 can be induced with viruses; one notable example is the 2009 Pandemic Influenza A(H1N1) 3 that led to impaired T-cell responses3. Also, Epstein-Barr Virus (EBV) which overproduced the latent membrane protein-1 (LMP1) induced more PD-L1 responses3. As a result, having a history with the EBV infection increases a person's risk of getting nasopharyngeal cancer (cancer of the area in the back of the nose), certain types of fast-growing lymphomas such as Burkitt lymphoma, Hodgkin lymphoma, and some cases of stomach cancer3. Additionally, PD-L1

plays a crucial role at certain times. For instance, during events such as pregnancy, tissue allografts, autoimmune disease, and disease states such as hepatitis, PD-L1 is thought to have a larger role to play and is often seen in larger quantities in the body.

PD-L1 can be found in many different types of cancers however its antibodies have only been effective in melanoma and non-small cancers. For some background information, T-cells are cells that are part of the adaptive immune system that work to kill infected host cells, activate other immune cells, creating cytokines, and regulating immune responses. PD-L1 is responsible for acting as a "stop sign" and prevents T-cells from getting activated. In addition to this, cells that have extra amounts of PD L1 can hide from an immune attack. This is because PD-L1 is also an immune checkpoint so in a tumor cell the immune system would identify the cell as "foreign" and therefore attack it.

PDL1 is a protein that helps keep immune cells from attacking non-harmful cells in the body however when a patient is PD-L1 positive, the immune checkpoint fails, and the T-cell does not attack the tumor cell. Small-cell lung cancer, squamous cell carcinoma (which is located in the oral cavity), cervical cancer, ovarian cancer, breast cancer, melanoma, bladder cancer, head and neck cancer, soft tissue sarcoma, and prostate cancer all have an increased copy number of chromosome 9p24 which is where CD274 is located. The antibodies that are used to inhibit PD-L1 in cancer are also known as immunotherapy since they are increasing the efficiency of the immune system[4]. There are various ways that PD-L1 signaling pathways can be activated in many kinds of cancers. One way is genetically such as a PD-L1 gene application, 3'-UTR disruption[2], or via an unregulated transcription. Other things that can activate the PD-L1 signaling pathways are the constitutive oncogenic signaling activation, extrinsic factors, and epigenetic mechanisms (up regulation of oncogenic microRNAs, down-regulation of tumor suppressor miRNAs, aberrant DNA methylation, and histone modifications). The signaling pathway for PD-L1 positive cancer is when tumor progression is based on T-cell inhibitory activities and binding to PD-1. So when PD-1 works with PD-L1, this stops the TCR signal, and co-stimulation by using the SHP-1 and SHP-2 phosphatase to the intracellular domain of PD-1[83,84] and up-regulating the expression of CBL E3 ubiquitin ligases to create TCR down modulation.[5] A study in 2004 discovered that PD-L1 positive cancer cells created a protective molecular shield to become more resistant to T-cell cytotoxicity. PD-L1/PD-1-blocking antibodies can break this shield in cell cultures which restore T-cell cytotoxicity. Just like all drugs, therapies for anti-PD-L1 have developed resistance in patients over time. This happened genetically because of the selection of cancer cell variants combined with somatic mutations that inactivate JAK1, or JAK2 or that abrogate β2- microglobulin expression.[6]

Methods

In this research, various bioinformatic tools and methods were used. For instance, cBioPortal was one of the bioinformatic tools that was used. Additionally, UniProt was another bioinformatic tool that was used. NCBI was a bioinformatic tool used in the study. The UCSF Xena Browser is the final tool that was used in the study. cBioPortal is a bioinformatic tool specifically dedicated to cancer genomics. Within the interface, there are various features available. It is used for visualizing and analyzing genomics data. It is useful in viewing different types of cancer and which mutations are present in the genomic DNA for that type of cancer. The figure below illustrates its interface. Additionally, The Universal Protein Resource (UniProt) was used and can be especially useful for protein sequence and annotation data. It can also be used to align various sequences for different species and map the genome for a particular gene. The figure below illustrates the interface used. NCBI (National Center for Biotechnology Information) was a final tool used for this study. It is part of the United States National Library of Medicine. It creates automated systems for storing and analyzing knowledge about molecular biology, biochemistry, and genetics and also facilitates the use of such databases and software by the research and medical community. In other words, it is a very useful tool for the efforts to collect biotechnology information. Also, the GTEX portal and DepMap was used in this study to graph information. Lastly, The UCSF Xena Browser was used.

Results

Gene / Protein Sequence

The gene and protein sequence of PD-L1 was analyzed in this study. To do this, the website UniProt was used. Figure 1 and Figure 2 show the sequence of PD-L1 in two different species: mice and homo sapiens.


```
>sp|Q9NZQ7|PD1L1_HUMAN Programmed cell death 1 ligand 1 OS=Homo sapiens OX=9606 GN=CD274 PE=1 SV=1
MRIFAVFIFMTYWHLLNAFTVTVPKDLYVVEYGSNMTIECKFPVEKQLDLAALIVYWEME
DKNIIQFVHGEEDLKVQHSSYRQRARLLKDQLSLGNAALQITDVKLQDAGVYRCMISYGG
ADYKRITVKVNAPYNKINQRILVVDPVTSEHELTCQAEGYPKAEVIWTSSDHQVLSGKTT
TTNSKREEKLFNVTSTLRINTTTNEIFYCTFRRLDPEENHTAELVIPELPLAHPPNERTH
LVILGAILLCLGVALTFIFRLRKGRMMDVKKCGIQDTNSKKQSDTHLEET
```

Figure 1. This image shows the sequence of PD-L1 in homos sapiens.


```
>sp|Q0EP73|PD1L1_MOUSE Programmed cell death 1 ligand 1 OS=Mus musculus OX=10090 GN=Cd274 PE=1 SV=1
MRIFAGIIFTACCHLLRAFTITAPKDLYVVEYGSNVTMECRFPVERELDLLALVVYWEKE
DEQVIQFVAGEEDLKPQHSNFRGRASLPKDQLLKGNAALQITDVKLQDAGVYCCIISYGG
ADYKRITLKVNAPYRKINQRISVDPATSEHELICQAEGYPEAEVIWTNSDHQPVSGKRSV
TTSRTEGMLLNVTSSLRVNATANDVFYCTFWRSQPGQNHTAELIIPELPATHPPQNRTHW
VLLGSILLFLIVVSTVLLFLRKQVRMLDVEKCGVEDTSSKNRNDTQFEET
```

Figure 2. This image shows the sequence of PD-L1 in mice.

Alignment

When comparing and aligning all three of the species (homo sapiens, mice, and canine dog), many interesting results were found. For one, it was seen through the individual sequences in each animal, there is a lot of overlap and similarities in the sequences. The figure 3 illustrates the alignment of the 3 species using UniProt.

Figure 3. This image shows three different species and how the sequence for PD-L1 differs for all of them. While the three species are very different the expressions of PD-L1 seem overall very similar.

According to UniProt, when comparing the human, mouse, and canine, there were 69 similar positions. Although this may not seem like much, when comparing the three for identical positions, it was found that there were 173 identical positions. Overall, within the comparison of the three, there was a 59.45 percent match.

Additionally, the comparison of just the human and the mouse was also done. In this alignment, very profound results were found. It was found that there were 200 identical positions. Additionally, it was found that there were 57 similar positions. There was a 68.72% identity that was similar.

Domain

As can be seen in Figure 4, V set and CD-set 2 are the two functional domains. Further analysis of these two domains shows profound results in terms of the function of these two domains. For the V set, it was found that this domain is found in antibodies as well as neural protein P0 and CTL4 amongst others. Additionally, it was found that this domain is involved in cell-cell recognition, cell-surface receptors, muscle structure, and the immune system.

Figure 4. The image illustrated the two major domains in the PD-L1 protein.

For CD-set 2, it was found that these domains belong to the immunoglobulin superfamily. As for their function, it was found that most C2-domains bind Ca2+ and mediate interactions of their proteins with phospholipids. While the majority of CD-set 2 domains bind Ca2+, it was found that some CD-set 2 domains actually do not bind Ca2+.Additionally, the comparison of just the human and the mouse was also done. In this alignment, very profound results were found. It was found that there were 200 identical positions. Additionally, it was found that there were 57 similar positions. There was a 68.72% identity that was similar.

Gene Expression

While using GETx Portal we were able to see that PD-L1 was mostly expressed on cells, lungs, and the spleen. The expression for males and females were about the same except for the kidneys, the reason for this is unknown. The figure below represents the highest median TPM from lowest to highest. PD-L1's median is highest for the cells EBV-transformed lymphocytes since the PD-L1 is considered as an immune checkpoint so it would be more common on lymphocytes compared to the brain specifically the hippocampus. The gene expressions for all of these tissues went significantly up when seen in the cancerous types of tissues(Figure 5). Because of increases of the IFNγ stimulation, the PD-L1 protein is over-produced preventing T cells from attacking cancerous cells. Before the medians had a range of 0.5 to 2 TPM but the gene expressions for cancerous cells had a range of .5 to 8 TPM.
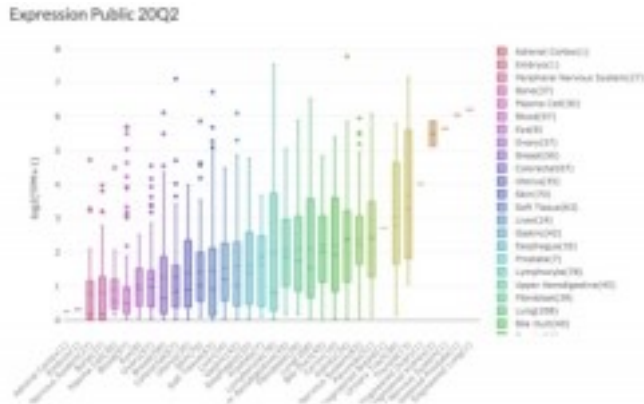
Expression Public 20Q2

Figure 5: This image represents the gene expression of PD-L1 in various types of cancerous tissues. As seen here, PD-L1 is vital for the progression of certain cancers but for others it is not necessary.

Another difference between the two gene expressions was that Engineered Lung Cancer became the highest tissue sample that had TPM instead of the EBV-transformed lymphocytes cells.

Copy Number

The copy number for PD-L1 changes depending on the type of cancer however they all have a range of zero to seven. Only lung cancer has reported cases of having a PD-L1 copy number of zero. The majority of the cancer had a median of two copies of the gene. All the copy number increases were found to be in chromosome number nine. It makes sense that there are no copy numbers that are negative in cancer cells since obviously PD-L1 benefits the cancer and they over-produce it. In addition to this, a higher copy number usually means a higher expression. As seen in the figure below, there is a linear relation between most of these types of cancer; however, an increase in one copy number for all cancer does not mean that the gene expression will have the *same* increase. For some types of cancer, there is a higher slope than others.
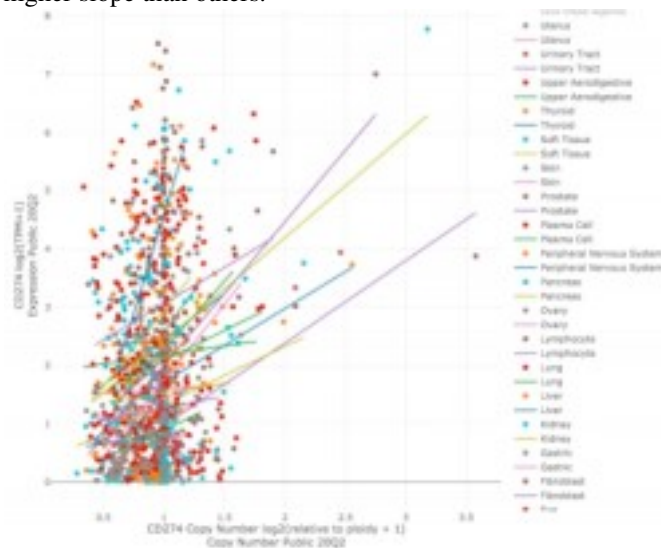

Figure6: This image represents the gene expression of PD-L1 in various types of cancerous tissues. As seen here, PD-L1 is absolutely vital for the progression of certain cancers but for others it is not necessary.

The reason for this is related back to the expression level of PD-L1 in cancer since some types don't utilize PD-L1 as much. As visualized in the graph to the left, using the DepMap database, we were able to compare the copy number

to expression level in TPM. The slope for uterus cancer was 0.864 explaining why PD-L1 inhibitors/ antibodies are not commonly used for uterus cancer however for AML the slope was 1.65 explaining why AML patients usually take those kinds of antibodies. What was confusing was the fact that AML had a high slope yet ALL had a much smaller slope despite their similarities. B-cell Mantle cells had one of the highest slopes (10.53) which makes sense considering that B-cells are immune cells that actually kill cancerous cells around the body.

Mutation Profile

As stated, before all the mutations related to PD-L1 are in Chromosome 9. The type of mutation varies between base substitutions, point mutations, deletions, and insertions; the most common for PD-L1 was a missense mutation10, a point mutation where single nucleotides result in a change of codon that codes for an unassociated amino acid, and least common was a frameshift deletion. By far, mutations are the most likely thing to happen to the PD-L1 gene. However, there are various times where it can be amplified or (although unrelated to the progression of cancer) deleted as well. The most common alterations in CD274 are CD274 Mutation (0.37%), CD274 Amplification (0.32%), CD274 Loss (0.31%), CD274 R260H (0.02%), and CD274 R260C (0.01%). CD274 R260H and CD274 R260C are both mutations of the CD274 gene. Overall, with the PD-L1 gene, it does mutate and many inhibitors target that one specific mutation instead of the regular PD-L1 normal gene.
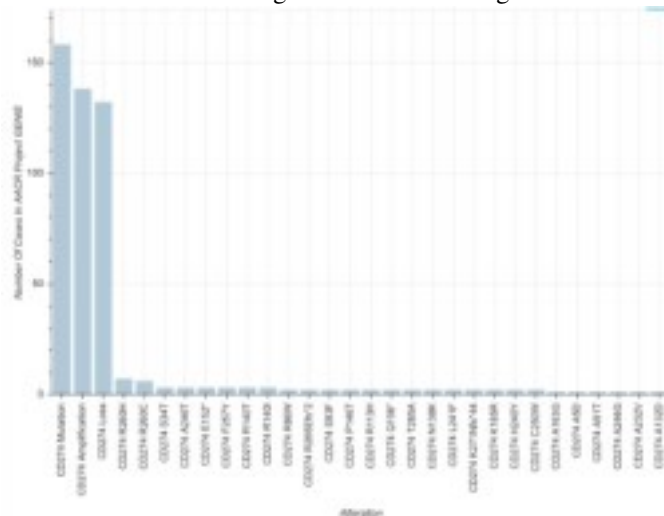

Figure7: The image shows the most common mutations in the PDL-1 gene the least common.

There are many locations in Chromosome 9 where a mutation can occur; one specific spot that is more prone to mutations is a place with an A A Change of R125*. This is a single base substitution which involves a replacement/ substitution of a single nucleotide base with another in DNA or RNA molecule and it has a big effect on the patient's wellbeing. Sometimes there can be two different types of mutations at the same spot. For example, at location 188, there is a single base substitution and small insertion and they had different effects on the patient.

Figure 8: The image shows the exact locations of where the mutations happen.

Conclusion

From our research, we examined that the PD-L1 gene sequences in both humans and in mice. In addition, we aligned them to each other and found how much alike they are. Also, we discovered information related to domains, mutation hotspots, structure, gene expression, copy number, and how the gene expression and the copy number relate to each other.

Limitations

Since this is a research article, our research is limited by the data available on the web, but other than what we can not access, there are few limitations to our research.

Acknowledgements

References

Transmembrane Protein - Chemistry Encyclopedia - structure, water, proteins http:// www.chemistryexplained.com/ Te-Va/Transmembrane-Protein.html (accessed Aug 2, 2020).

Cristino, A. S.; Nourse, J.; West, R. A.; Sabdia, M. B.; Law, S. C.; Gunawardana, J.; Vari, F.; Mujaj, S.; Thillaiyampalam, G.; Snell, C.; et al. EBV MicroRNA-BHRF1-2-5p Targets the 3' of Immune Checkpoint Ligands PD-L1 and PD-L2. *Blood* **2019**, *134* (25), 2261–2270.

Valero-Pacheco, N.; Arriaga-Pizano, L.; Ferat-Osorio, E.; Mora-Velandia, L. M.; Pastelin-Palacios, R.; Villasís-Keever, M. Á.; Alpuche-Aranda, C.; Sánchez-Torres, L. E.; Isibasi, A.; Bonifaz, L.; et al. PD L1 Expression Induced by the 2009 Pandemic Influenza A(H1N1) Virus Impairs the Human T Cell Response. *J. Immunol. Res.* **2013**, *2013*. https://doi.org/10.1155/2013/989673. PD-L1 protein overview: Sequence, structure, function and protein interaction | sino biological https:// www.sinobiological.com/resource/pd-l1/proteins (accessed Aug 2, 2020).

Immune checkpoint inhibitors and their side effects https://www.cancer.org/treatment/treatments-and side- effects/ treatment-types/immunotherapy/immune-checkpoint-inhibitors.html (accessed Aug 6, 2020).

 Messerschmidt, J. L.; Prendergast, G. C.; Messerschmidt, G. L. How Cancers Escape Immune Destruction and Mechanisms of Action for the New Significantly Active Immune Therapies: Helping Nonimmunologists Decipher Recent Advances. Oncologist 2016, 21 (2), 233–243.

Missense Mutation https://www.genome.gov/genetics-glossary/Missense-Mutation (accessed Aug 6, 2020).