



Predicting Neoantigens for Cancer Using Next-Generation IEDB & CEDAR Tools

PEACE KOUDJIWAN¹, BRADY ANDERSON²,
BETH VITALIS³, MATTHEW KATELE¹, EH HSER¹,
GABRIELLE JACKSON¹, JAZMIN MALAGON¹,
JOANNA SANCHEZ⁴, JAN ZAKARIYA¹, ANDRES ROJAS JR.¹,
CIDNEY SMITH¹, NINA BLAZESKA⁵,
ZEYNEP KOŞALOĞLU-YALÇIN⁵, SHEELA VEMU^{1*}

¹Department of Liberal Arts & Sciences, Waubonsee Community College, Sugar Grove, IL 60554, USA

²Cleveland Clinic; Cleveland, OH 44122, USA

³South Puget Sound Community College, Olympia, WA 98512, USA

⁴School of Molecular and Cellular Biology, University of Illinois- Urbana-Champaign, Urbana, IL 61801, USA

⁵La Jolla Institute for Allergy and Immunology, La Jolla, CA 92037, USA

*svemu@waubonsee.edu

Abstract: Cancer is a group of diseases characterized by uncontrolled growth and spread of abnormal cells. The underlying cause of cancer relates to the cell cycle, during which DNA is replicated. Cancer cells accumulate DNA mutations that help them acquire cancerous features, such as evading cell death and indefinite growth [1]. If these DNA mutations are in coding regions, they are translated to mutated proteins. The epitopes that contain these mutations are called neoantigens. Neoantigens are highly tumor-specific and can be targeted with immunotherapies [2]. During cell division, tumor suppressor genes play a role in the case of DNA damage or replication errors. The p53 protein is a tumor suppressor gene product that prevents tumor formation by activating processes that block cell division when DNA damage has occurred [3]. Mutant p53 does not effectively bind DNA or activate the production of proteins necessary for the stop signal. This project explored a hypothesis that a set of distinct p53 protein mutations can be selected to serve as potential targets for cancer immunotherapy and vaccines by using immunoinformatics predictive analysis tools. By comparing these potential targets with experimental results, we can predict epitopes that may serve as neoantigen targets for immunotherapy. We identified candidate immunogenic epitopes using the NCI's TP53 Database (NCI DB - tp53.isb-cgc.org), Cancer Epitope Database and Analysis Resource (CEDAR - cedar.iedb.org), and a powerful new bioinformatics tool (nextgen-tools.iedb.org/) [4] hosted by Immune Epitope Database (IEDB - iedb.org) and CEDAR. Comparing predicted epitopes to highly mutable regions of p53 in tumor variants from NCI DB revealed areas of overlap that may be priority candidate epitopes for immunotherapy. Experimental data from CEDAR tested the immunogenicity of normal and mutated protein versions to help avoid harmful cross-reactions. These results help predict cancer epitope amino acid sequences relevant to understanding the immune system's role in cancer progression, prevention, and treatment. These studies also set the stage for important subsequent undergraduate research projects to further characterize predicted cancer neoantigens.

Keywords: p53, mutations, neoantigens, cancer, CEDAR, bioinformatics, immunotherapeutic, Open Access Resources

© 2024 under the terms of the J ATE Open Access Publishing Agreement

Introduction

The underlying cause of cancer relates to the process by which most human cells grow and repair, known as the cell cycle. During the cell cycle, DNA is replicated so that dividing cells contain their matching set of chromosomes. DNA can be damaged by toxins, radiation, or other sources, leading to mutations that will be passed onto offspring cells if the cell cycle is not stopped. Cancer arises from the accumulation of mutations, which results in uncontrolled cell division and growth. Genes involved in regulating the cell cycle, including



proto-oncogenes and tumor suppressor genes, are often found to be mutated in tumors [5]. Cancer therapies that target oncogenes, such as trastuzumab (Herceptin®) and imatinib (Gleevec®, STI571), both work by targeting kinases that are part of the cell signaling pathway [6].

These drugs have revolutionized cancer treatment by inhibiting oncogenic proteins and blocking accelerated growth. However, reactivating a mutated, inactive protein such as tumor suppressor genes is challenging. While p53 is the most well-characterized contributor to tumors, no targeted drugs are available. Cellular mechanisms involving tumor suppressor genes, such as TP53, have naturally evolved to stop the cell cycle in the case of DNA damage or replication error. The p53 protein prevents uncontrolled growth and tumor formation by responding to stress-induced DNA damage via entering the nucleus and binding DNA to prompt the production of another mediator protein [3]. Mutations in the p53 protein affect DNA binding, so the mediator protein is not made. As a result, the cell loses the stop signal that would prevent uncontrolled growth. In addition, p53 binds DNA as a tetramer of four molecules of functioning p53; thus, if one allele of TP53 is mutated, it will negate the function of the unmutated p53 gene product [7].

An alternate strategy for cancer treatment focuses on leveraging the adaptive immune system to identify and destroy cells displaying foreign antigens presented on their cell surface (cell-mediated branch). Neoantigens are a class of peptides carrying somatic mutations resulting in a “new” antigen that can be identified as foreign and marked for destruction [8]. Targeting these neoantigens provides a mechanism for tumor-specific immunotherapies carried out on behalf of the adaptive immune system. Since tumor-associated mutations in p53 can produce neoantigens, cancerous cells that produce them can be distinguished from normal p53 epitopes and cleared by immune effector cells.

In the case of alterations to the TP53 gene, the mutant p53 proteins are degraded into short peptides and transported into the endoplasmic reticulum. In the lumen of the endoplasmic reticulum, the peptide fragments may bind with major histocompatibility complex (MHC) Class I proteins present in all nucleated cells to mediate antigen presentation. This epitope-MHC complex is displayed on the cell’s surface, where it may be found by a T-cell with a complementary receptor (TCR), forming a tight MHC-TCR complex. T-cell receptors are specific to foreign antigens and bind only epitope-MHC ‘peptides in a bun’ shaped complexes. The resulting immune response makes neoantigens promising immunotherapeutics, especially for highly immunogenic epitopes on the surface of tumor cells [8]. Current bioinformatics techniques such as sequence analysis, machine learning-aided binding, and immunogenicity predictions [9-10] (see Figure 1) help identify tumor-specific neoantigen epitopes that may be effective immunotherapeutic targets and cancer vaccines. Maximizing computational predictions helps minimize expensive and laborious experimental approaches.

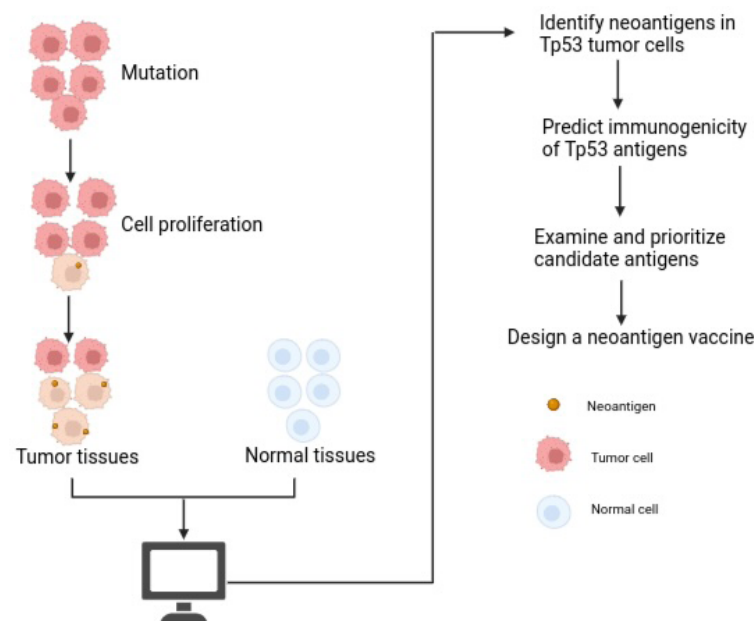


Fig. 1. [10-11] Bioinformatics tools enable prediction of candidate immunotherapy targets. (Figure adapted from "Neoantigen vaccine: an emerging tumor immunotherapy" and made with BioRender. (<https://bit.ly/3TLZV4T>))



The undergraduate research described here is at the forefront of cancer immunotherapy; leveraging the cancer-specific NCI TP53 database, CEDAR, and IEDB tools to predict immunogenic p53 tumor antigen peptides computationally. By utilizing a next-generation pipeline tool to predict the processing of intracellular events, the results are then compared to experimental p53 epitope data to identify epitopes most likely to elicit an immune response to a large set of tumors while minimizing cross-reactivity to normal tissue. This project combines a current understanding of p53 role in cancer with available databases and bioinformatics tools to identify and characterize priority peptide epitopes that may serve as powerful neoantigens for targeted immunotherapies.

Methods

Database and Resource Tools Used for Bioinformatics Analyses

This study used a select set of open access resources. In 2021, The Cancer Epitope and Analysis Resource (CEDAR), (<https://cedar.iedb.org/>) funded by the National Cancer Institute NCI, was developed as a companion to the Immune Epitope Database Analysis Resource (IEDB) (iedb.org) created by National Institute of Allergy and Infectious Diseases in 2003 with ongoing updates [12]. CEDAR serves as a repository of cancer-specific experimental peptide and epitope data as it catalogs experimental data on antibodies and T cell epitopes studied primarily in humans regarding cancer disease. CEDAR and IEDB collectively host next-generation tools that assist in predicting and analyzing epitopes (nextgen-tools.iedb.org/). For given protein sequences, the tools predict each step in the antigen processing and display process, including proteasomal cleavage, transporter associated with antigen processing (TAP), MHC Class I binding, cell-surface display, and T-Cell recognition (see Figure 2). The underlying machine learning algorithms have been trained on extensive empirical data sets to predict how each candidate epitope will behave at each step, thus avoiding costly empirical testing for large sets of new candidate epitopes. The sequence processing workflow follows the biological process by which peptides are internally processed and externally displayed for interaction with T-cell receptors. This project leveraged the newer CEDAR database to predict and analyze immunogenic p53 cancer epitopes.

The National Cancer Institute's (NCI) TP53 Database (<https://dceg.cancer.gov/tools/public-data/tp53-database>) has nearly 28,000 mutations of TP53 tumor variants characterized and available to the public. This extensive TP53 mutation variant dataset was used to map mutation frequency across the p53 protein and align it with computation predictions. The computational predictions from the generation tools in CEDAR were compared to experimental data in NCI TP53 DB and the CEDAR database, as described below.

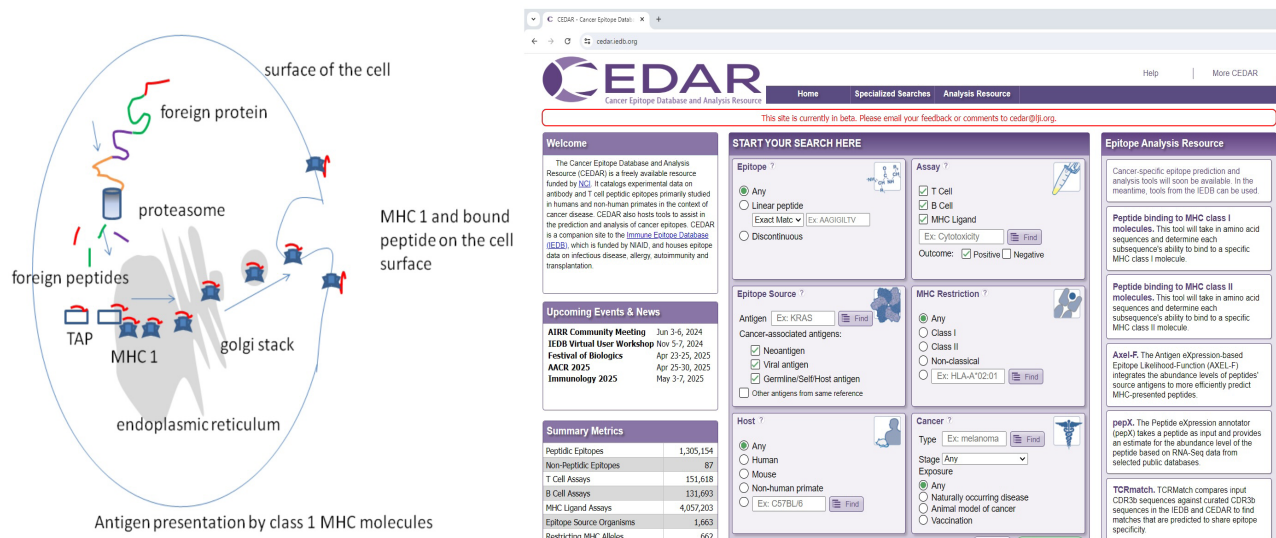


Fig. 2. [12-13] The CEDAR database and Next-Generation tools enable computational prediction of all steps in the antigen processing and presentation pathway (Figure adapted from Colm and Koşaloğlu-Yalçın et al.)



MHC Class I Presenting Peptide Predictions Using Next-Generation Tools

The Next-Generation Epitope Prediction Tools platform (<https://nextgen-tools.iedb.org/>) was used to predict a set of peptide epitopes that MHC Class 1 proteins may present on the cell surface. The tool links predictions of intracellular events of antigen processing into one workflow. The computational pipeline used included the following predictions: proteasomal cleavage, predictions of selective specificity of peptides that are transported into the cytosol of the endoplasmic reticulum lumen, and MHC1 binding. Our pipeline and its parameters (<https://nextgen-tools.iedb.org/pipeline>) using p53 (UniProt: P04637).

p53 Tumor Variant Mutation Frequency Distribution

TP53 mutation variant data (n=27,847) from the NCI p53 database was used to identify protein regions showing high mutation frequency. The NCI codon distribution tool (https://portal.gdc.cancer.gov/analysis_page?app=ProteinPaintApp) was applied to the variant data to build a tumor variant distribution chart displaying the mutation frequency of amino acid segments along linear p53.

Aligning Predicted Peptides to Full-Length p53 Protein and Tumor Variant Data

The resulting 23 peptides were modeled on linearized p53 (Uniprot: P04637) (1-393 aa) and juxtaposed along the NCI p53 tumor variant mutation distribution chart described above to identify regions of interest for neoantigen targets.

Comparison of Predicted and Empirical Results in CEDAR to Obtain Experimental Data for NGP Peptides

To describe the immunogenicity of the 23 NGP Peptides as non-mutated (self-antigen) and mutated (neoantigen) epitopes, human T cell assays of TP53 (UniProt: P04637, E7EQX7, J3KP33) epitopes were exported from CEDAR's database and stratified. Self-antigen assays (n=25) were collected by filtering self-antigens with negative assay results. Neoantigen assays (n=76) were collected by filtering neoantigens with positive assay results. Using BLAST, sequences from CEDAR and NGP Peptides were matched, and assays for each NGP Peptide were counted.

Results and Discussion

Next-Generation Pipeline Predicted Peptides

The next-generation pipeline (NGP) feature of the CEDAR and IEDB resources was used to compute a set of candidate neoantigen targets. The NGP predicts products of intracellular steps of antigen processing to display for immune system surveillance and is a relatively new resource available to the public. (<http://workshop.iedb.org/>)

The results of the next-generation pipeline applied to the p53 protein sequence included a set of 23 peptides, listed in Table 1. Amino acid location on the p53 protein shows a broad distribution with a few clusters. Epitopes derived from sequences with the highest incidence of mutations must be a higher priority, as the resulting therapy will be effective against a broader set of tumors across diverse populations.



Table 1

Cluster.Sub-Cluster Number	Peptide Number	Alignment	AA Position
1.1	Consensus	FEMFRELNEALELK	338-351
2.1	Consensus	RMPEAAPPVAPAP	65-77
3.1	Consensus	EYFTLQIRGRERF	326-338
4.1	Consensus	YQGSYGFR LGFLH	103-115
5.1	Consensus	GTRVRAMAIYK	154-164
6.1	Consensus	APAPAAPTPAA	74-84
7.1	Consensus	LSQETFSDLWKL	14-25
8.1	Consensus	VEYLDDRNTFR	203-213
9.1	Consensus	NLLGRNSFEVR	263-273
10.1	Consensus	MLSPDDIEQWF	44-54
11.1	Consensus	EVRVCACPGRDRR	271-283
12.1	Consensus	DSTPPPGTRVR	148-158
13.1	Consensus	RGRERFEMFREL	333-344
14.1	Consensus	QSQHMTEVVRR	165-175
15.1	Consensus	VVVPYEPPEV	216-225
16.1	Consensus	APAPAPSWPL	84-93
17.1	Consensus	VGSDCTTIHY	225-234
18.1	Consensus	HLIRVEGNLR	193-202
19.1	Consensus	WKLLPENNVL	23-32
20.1	Consensus	RNSFEVRVCA	267-276
21.1	Consensus	RNTFRHSVVV	209-218
22.1	Consensus	RRPILTIITL	248-257
23.1	Consensus	RVEGNLRVEY	196-205

The 23 next-generation pipeline (NGP) predicted peptides and their position on the p53 protein sequence after processing through IEDB's Next-generation Pipeline tools. Each NGP Peptide is the consensus alignment of its respective clustering results.

Positioning of NGP Predicted Peptides on p53 Mutations

To determine the prevalence of each mutation and thus the real-world relevance of the predicted epitopes, the NCI's TP53 database was accessed and analyzed in the context of the linear p53 protein. As seen in Figure 3, the frequency distribution of TP53 mutation variants along full-length p53 was visualized by building a codon distribution chart, with each codon representing an amino acid. The peaks and valleys show the frequency of mutations around a specific section of the linear protein. The codon chart showed a high incidence of mutations in specific regions along p53.

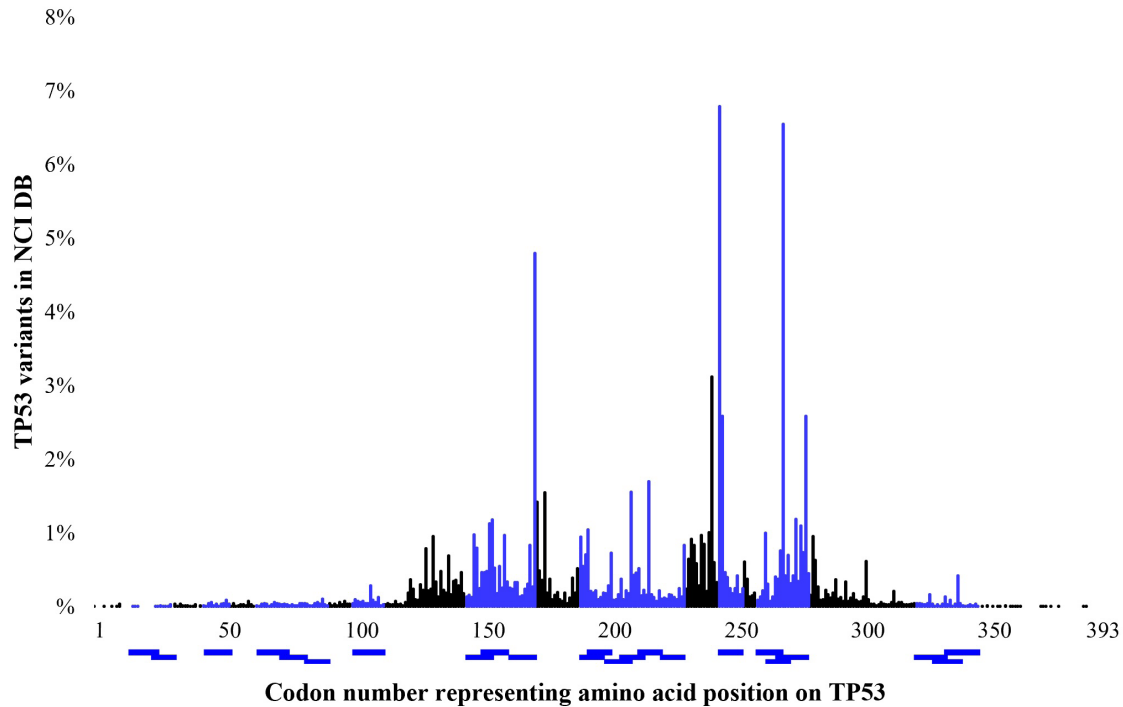


Fig. 3. The frequency distribution of known human p53 variants from the NCI's TP53 Database (n=27,847) is mapped onto the full-length p53 codons (1-393 aa), representing amino acids. In blue, the 23 NGP Peptides are both spatially aligned along TP53 and highlighted within the distribution. The association between NGP Peptides and variant data represents the mutability within each predicted peptide.

Juxtaposing the 23 NGP Peptides against the mutation distribution of computationally predicted epitopes on this frequency chart helped zone in on neoantigen epitope targets that will be effective across the broadest range of tumors and populations. Multiple spikes in prevalence are observed between the 150th and 300th codon. This area acts on the cell cycle, inhibiting the moderating actions that monitor the cell cycle, resulting in the formation of cancer cells, and it makes sense that mutations will lead to tumor development [14]. The peaks in blue show overlap between regions of high mutagenicity and predicted immunogenicity, which will be priority amino acid regions for immunotherapy.

The Summation of TP53 Point Mutation Frequencies within Each NGP Eptide

To further characterize the incidence of real-life mutations in these predicted epitope peptides, the percentage of p53 variants found in the NCI's TP53 database represented by each of the amino acids within each of the 23 peptides was combined and graphed in Figure 4.

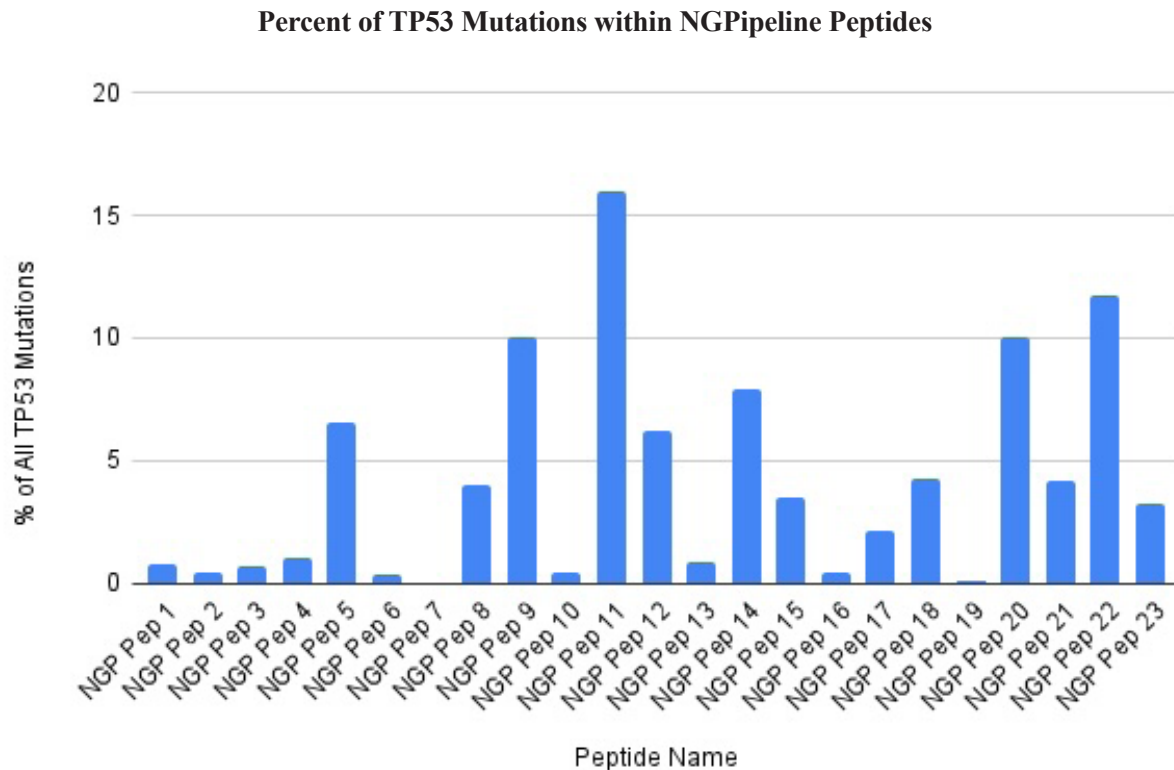


Fig. 4. This graph summarizes the frequency of tumor variants from the NCI's TP53 database, represented in each of the next-generation pipeline (NGP) predicted peptides (Peps). Calculations were performed by combining the mutational frequencies for each amino acid within the amino acid range of the NGP Peps.

Certain NGP Peptides contained higher frequencies of mutations than other peptides. Those accumulating more than 5% of known mutations were NGP Peptides 5, 9, 11, 12, 14, 20, and 22. Peptides containing 3% to 5% of the mutations were NGP Peptides 8, 15, 18, 21, and 23. Peptides with fewer than 3% of the mutations were NGP Peptides 1, 2, 3, 4, 6, 7, 10, 13, 16, 17, and 19. These data give confidence to our predictive model, where predicted epitopes can be found in the literature. The next step was to use the *in vitro* data within the CEDAR database to assess real-world immunogenicity as measured by T-cell assays.

Comparing *In Vitro* Immunogenicity of Self-Antigen and Neoantigen NGP Sequences

Non-mutated self-antigens are non-immunogenic or would otherwise be autoimmune. Depending on the mutation, variable levels of immunogenicity are possible as the neoantigen is dissimilar from the self-antigen [9]. Figure 5 shows *in vitro* human T-cell assays exported from CEDAR where the experimental epitopes matched NGP Peptides. Self-antigen assays were the accumulation of self-antigen stimulation assays that did not elicit an immunogenic response, and neoantigen assays were those from neoantigen peptides that did elicit an immunogenic response. CEDAR was vital because it collected these epitopes and their assay data to analyze post hoc. Without these open-access resources, this project would be challenging to process using currently available undergraduate research resources.

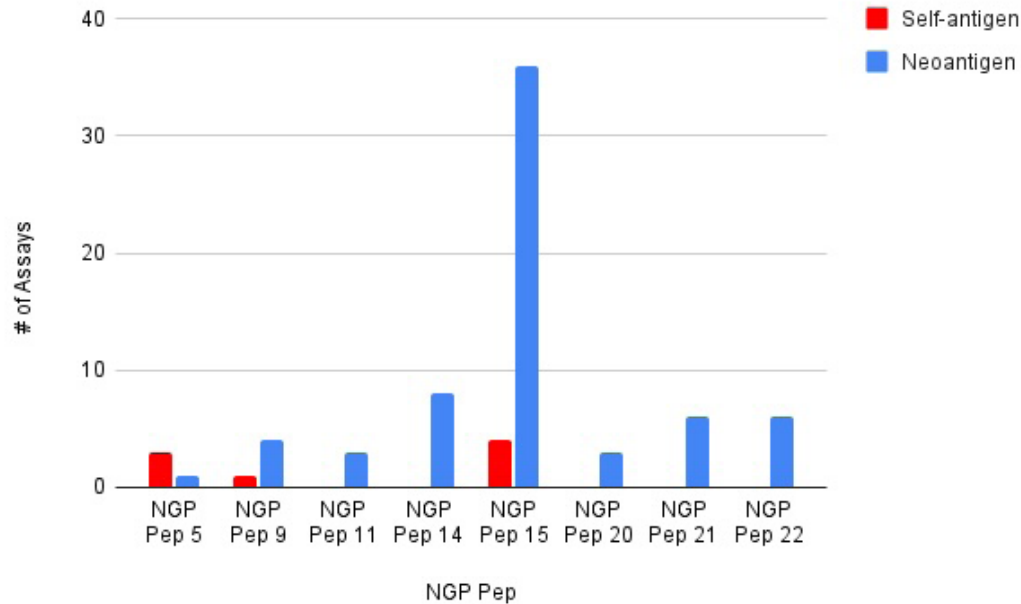


Fig. 5. The number of *in vitro* T cell assays exported from CEDAR with experimental epitopes matching any of the 23 NGP Peptides. Self-antigen assays were the accumulation of self-antigen stimulation assays that did not elicit an immunogenic response ($n=8$ of 25), and neoantigen assays were those from neoantigen peptides that did elicit an immunogenic response ($n=67$ of 76). NGP peptides 5, 9, and 15 are high priority as the non-mutated sequence did not produce T cell reactivity, and the mutated version did.

Conclusion

To summarize, this work identified three p53 epitope sequences representing a significant set of real-world p53 mutations found in tumors. In addition, *in vitro* data supports that the neoantigens are immunogenic, while the non-mutated sequences are not. These results suggest that the results of the CEDAR prediction tool can be used to predict real-world data. These results also help recommend further *in vitro* and *in silico* testing of epitopes to increase our confidence in whether our other NGP Peptides are suitable candidates.

These findings help support the value of computational prediction in identifying high-priority immunotherapy and vaccine targets. This is important because challenges remain in cancer immunotherapy, especially in solid tumors. p53 is an attractive target since it is a critical tumor suppressor [15]. Mutations in the p53 gene have been found in 50% of cancers, and failures in the p53 pathway contribute to almost all cancers [16]. Furthermore, prior research suggests it has a dominant negative phenotype. To aid computational predictions, an extensive set of p53 tumor antigen variants have been identified and are available in NCI TP53 DB, with supporting experimental data in CEDAR. We evaluated these resources and leveraged components that helped meet the project objective to determine whether computational predictions can successfully identify p53 immunogenic neoantigens that cover the spectrum of clinical mutations. Requirements of a successful immunotherapy target include effective antigen processing and T cell reactivity, a non-immunogenic wild-type, and tumor antigen variants that are clinically prevalent across diverse populations. Comparing predicted results with available experimental data best enables effective immunotherapy target epitope identification.

Computational prediction is critical to identifying high-priority immunotherapy targets given that the immune system sees only a tiny fraction of tumor antigens, so data alone do not give a complete picture. To predict immunogenicity, neoantigen prediction tools must cover all steps, from mutant protein production to T cell activation. The development of CEDAR and the next-generation tools platform has enabled students to



embark on medically essential and timely research to help develop broad immunotherapy targets for cancer diagnoses and vaccines. Follow-on student projects will use CEDAR and its next-generation tools to further refine the priority immunotherapy targets by assessing protein expression and the critical T-cell recognition of the candidate epitopes. As CEDAR expands, we expect future projects to utilize their growing repertoire of tools and conduct deeper analyses [12]. This research also sets the stage for future student projects that could explore features of the high-priority epitopes (overlap of computed and empirical epitopes), such as the effect of mutations on protein structure and how this may impact function and immunogenicity. Critical subsequent research will also explore how well the predicted epitopes represent diverse populations [17-18]. It will suggest ways that data can be accessed and utilized differently so that the resultant immunotherapy would benefit all populations.

In conclusion, concepts of the role of p53 in cancer were applied with the novel, open-access databases and bioinformatics tools to identify and characterize priority peptide epitopes that may serve as powerful neoantigen targets. The work sets the stage for follow-up undergraduate projects that use current bioinformatics capabilities to address and help solve immunotherapy and vaccine challenges.

Acknowledgments. The students would like to thank their faculty advisor, Sheela Vemu, Ph.D., Associate Professor of Biology at Waubensee Community College. They would like to express their appreciation for the insight and guidance from the Cancer Epitope Database and Analysis Resource (CEDAR) Bioinformatics scientist, Dr. Zeynep Koşaloğlu-Yalçın, and Senior Project Manager Nina Blazeska, La Jolla Institute for Allergy and Immunology, La Jolla, CA. The students want to acknowledge the significant contribution of a Waubensee Alumnus, Brady Anderson, for his research mindset that propelled the project's direction and the editing of the manuscript, and Beth Vitalis for her dedication and expertise in improving and finalizing the manuscript. The initial ideation of this project started with the Antibody Engineering Hackathons, which were supported by the National Science Foundation (NSF) under the award DUE ATE # 2055036. This work was in part supported by the National Science Foundation (NSF) under the award DUE ATE #2325500.

Disclosures. The authors declare no conflicts of interest.

References

- [1] A. K. Basu, "DNA Damage, Mutagenesis and Cancer," *International Journal of Molecular Sciences*, vol. 19, no. 4, Mar. 2018, doi: 10.3390/ijms19040970.
- [2] N. Xie, G. Shen, W. Gao, Z. Huang, C. Huang, and L. Fu, "Neoantigens: promising targets for cancer therapy," *Signal Transduction and Targeted Therapy*, vol. 8, no. 1, Jan. 2023, doi: 10.1038/s41392-022-01270-x.
- [3] B. Vogelstein, D. Lane, and A. J. Levine, "Surfing the p53 network," *Nature*, vol. 408, no. 6810, pp. 307–310, Nov. 2000, doi: 10.1038/35042675.
- [4] Z. Yan et al., "Next-generation IEDB tools: a platform for epitope prediction and analysis," *Nucleic Acids Research*, vol. 52, no. W1, pp. W526–W532, May 2024, doi: 10.1093/nar/gkae407.
- [5] E. N. Kontomanolis et al., "Role of Oncogenes and Tumor-suppressor Genes in Carcinogenesis: a Review," *Anticancer Research*, vol. 40, no. 11, pp. 6009–6015, Nov. 2020, doi: 10.21873/anticancer.14622.
- [6] S. J. Baker and E. P. Reddy, "Targeted Inhibition of Kinases in Cancer Therapy," *The Mount Sinai Journal of Medicine*, New York, vol. 77, no. 6, pp. 573–586, 2010, doi: 10.1002/msj.20220.
- [7] J. Gencel-Augusto and G. Lozano, "p53 tetramerization: at the center of the dominant-negative effect of mutant p53," *Genes & Development*, vol. 34, no. 17–18, pp. 1128–1146, Sep. 2020, doi: 10.1101/gad.340976.120.
- [8] X. Zhao, X. Pan, Y. Wang, and Y. Zhang, "Targeting neoantigens for cancer immunotherapy," *Biomarker Research*, vol. 9, no. 1, Jul. 2021, doi: 10.1186/s40364-021-00315-7.



- [9] Yat-tsai Richie Wan, Zeynep Koşaloğlu-Yalçın, B. Peters, and M. Nielsen, “A large-scale study of peptide features defining immunogenicity of cancer neo-epitopes,” *NAR cancer*, vol. 6, no. 1, Jan. 2024, doi: 10.1093/narcan/zcae002.
- [10] V. Yurina and O. R. Adianingsih, “Predicting epitopes for vaccine development using bioinformatics tools,” *Therapeutic Advances in Vaccines and Immunotherapy*, vol. 10, no. January-December 2022, p. 251513552211002, Jan. 2022, doi: 10.1177/25151355221100218.
- [11] M. Peng et al., “Neoantigen vaccine: an emerging tumor immunotherapy,” *Molecular Cancer*, vol. 18, no. 1, Aug. 2019, doi: 10.1186/s12943-019-1055-6.
- [12] Z. Koşaloğlu-Yalçın et al., “The Cancer Epitope Database and Analysis Resource (CEDAR),” *Nucleic Acids Research*, vol. 51, no. D1, pp. D845–D852, Oct. 2022, doi: 10.1093/nar/gkac902.
- [13] G. Colm, “Antigen Presentation,” Wikipedia. Accessed: May 01, 2024. [Online]. Available: https://commons.wikimedia.org/wiki/File:Antigen_Presentation.png
- [14] L. Bouaoun et al., “TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data,” *Human Mutation*, vol. 37, no. 9, pp. 865–876, Jul. 2016, doi: 10.1002/humu.23035.
- [15] Y. Lu, M. Wu, Y. Xu, and L. Yu, “The Development of p53-Targeted Therapies for Human Cancers,” *Cancers*, vol. 15, no. 14, p. 3560, Jan. 2023, doi: 10.3390/cancers15143560.
- [16] T. Ozaki and A. Nakagawara, “Role of p53 in Cell Death and Human Cancers,” *Cancers*, vol. 3, no. 1, pp. 994–1013, Mar. 2011, doi: 10.3390/cancers3010994.
- [17] A. Madbouly and Y. T. Bolon, “Race, ethnicity, ancestry, and aspects that impact HLA data and matching for transplant,” *Frontiers in genetics*, vol. 15, Mar. 2024, doi: 10.3389/fgene.2024.1375352.
- [18] N. W. Fischer, Y. -H. V. Ma, and J. Gariépy, “Emerging insights into ethnic-specific TP53 germline variants,” *JNCI: Journal of the National Cancer Institute*, vol. 115, no. 10, pp. 1145–1156, Jun. 2023, doi: 10.1093/jnci/djad106.