Decoding Neural Networks: Discovery of Anti-Tumor B Cell Receptor Motifs Using a Novel Sequence-Based Computational Framework

Cynthia Chen
Society for Science and the Public, United States of America

Abstract: Deep learning models have been successfully employed for various challenging biological tasks; however, the complexity and depth of neural networks render them black boxes. To address this problem and reveal the important features learned by deep learning models, we developed a novel computational pipeline for decoding neural network models trained on protein sequence data. Our pipeline consists of several stages: generating random input sequences, running the model to rank sequences, clustering top sequences to characterize motifs, and visualizing motif clusters with sequence logos. Using our pipeline, we deciphered the binding motifs learned by a deep learning model trained on a pan-cancer dataset containing more than 30 million B cell receptor (BCR) protein sequences from ~5,000 patients. We discovered 65 BCR binding motifs among 13 cancer types and validated the robustness of the motifs through extensive correlation analyses. Our study is the first to reveal and validate anti-tumor BCR binding motifs that target specific tumor antigens, a discovery that is critical to the future synthesis of new antibody drugs for cancer treatments. Furthermore, we demonstrated the versatility of our computational pipeline by using it to decode a second deep learning model, showing that our methods are applicable to a variety of neural networks.

1 Introduction

9.6 million people worldwide die from cancer each year [1]. To combat cancer, it is critical to understand the adaptive immune system, which possesses the ability to recognize a diverse set of tumor antigens and mutations. In particular, B cells are an integral part of the human immune response, as they recognize and bind to disease-causing tumor antigens and signal the release of antibodies to attack tumor cells [2]. Through these mechanisms, B cells can effectively control tumor growth in humans.

In this study, we focus on the B cell receptor (BCR), which is situated on the surface of the B cell. The BCR is the driving force behind the activation of a B cell: when a BCR recognizes and binds to an antigen presented by a tumor cell, it triggers the secretion of disease-preventing antibodies into the tissue and blood. B cells are extremely effective in targeting specific antigens because of the complementarity determining region 3 (CDR3). Since tumor cells belonging to the same cancer type have a set of shared antigens, such as neoantigens, the B cell CDR3 also has similar binding specificities among different cancer types to recognize the tumor antigens [3].

The DeepBCR model is a multi-layer deep neural network that was developed prior to this study [4]. Based on input BCR repertoires, DeepBCR predicts the corresponding cancer-type association scores to model the binding affinity between the BCR and antigen. DeepBCR was trained on a dataset that contains BCR sequences for $\sim$5,000 patient samples across the 13 cancer types.

After being trained on the preprocessed dataset, the DeepBCR neural network achieved an average prediction accuracy of 39.3% and area under the curve (AUC) of 0.80 across all 13 cancer types, compared to a benchmark random predictor (accuracy 7.7% and AUC 0.5). The predictive ability of DeepBCR to associate BCR sequences to tumor-types motivated us to determine the hidden information the model learned from the TCGA dataset that helped it make accurate predictions.

Many neural networks such as DeepBCR have been trained on large sequence datasets and have achieved high prediction accuracy, yet we are unable to reveal the motif patterns they have learned. In this study, we aimed to address the lack of computational methods for decoding sequence-based neural networks by interpreting and characterizing the tumor-specific BCR motifs learned by the DeepBCR model from the TCGA dataset. Therefore, we established two primary objectives: 1) develop a pipeline to decode neural networks trained on protein-sequence data and 2) apply the pipeline to DeepBCR to discover anti-tumor BCR binding motifs.

2 Methods

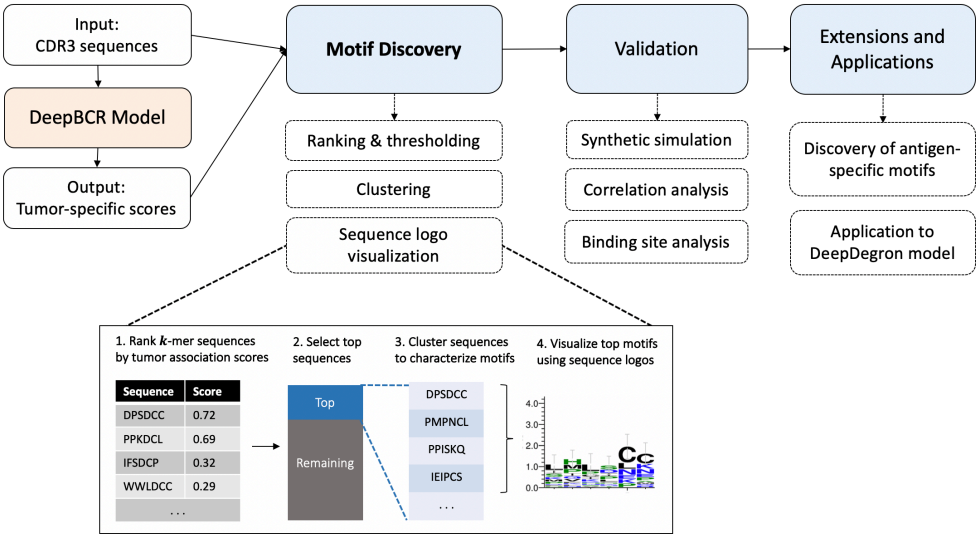
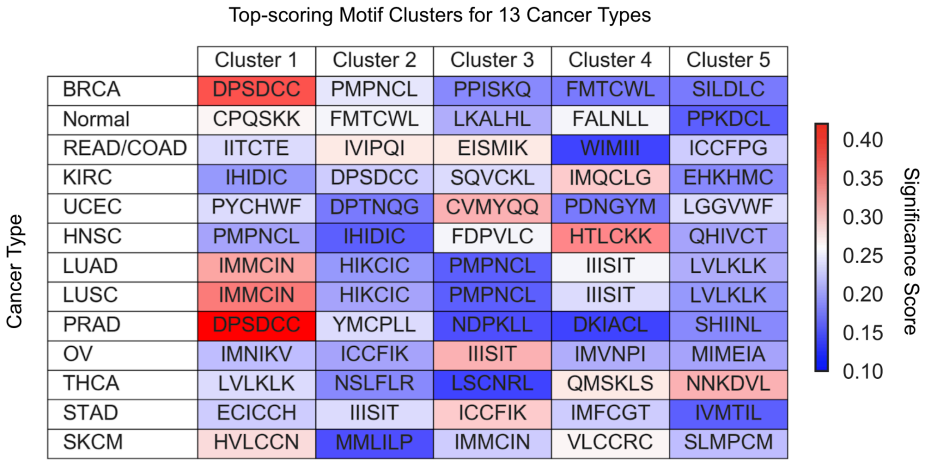


Figure 1: An overview of our multi-step computational motif discovery pipeline. In addition to DeepBCR, the pipeline framework can be adapated to decode the motifs learned by other deep neural networks trained on protein sequence data.

In this study, we developed a novel computational pipeline to characterize the protein sequence motifs that DeepBCR learned from the TCGA dataset. As outlined in Figure 1, our computational pipeline, created using Python and R, consists of several stages. First, a set of random k-mer sequences is generated and fed into a trained DeepBCR model, which predicts tumor-association scores for each input sequence. Next, we perform motif discovery, the main component of our pipeline. To extract the motifs learned by DeepBCR, we rank all the sequences based on the tumor-specific scores, determine a selection threshold for each cancer type, and use this threshold to select the top-scoring sequences for each cancer type. Then, greedy clustering is performed on these sequences to identify motif clusters, which are visualized using sequence logos. In the validation stage, we verify the robustness of our pipeline by running a synthetic data simulation, comparing correlations between models with different training parameters and iterations, and conducting binding site analyses on full-length CDR3s.

To demonstrate the generalizability of our methods, we applied our computational pipeline to decode a second deep neural network, the DeepDegron model. DeepDegron predicts degron regulatory potential (DRP), the likelihood of an input protein sequence to contain a degron [5]. By interpreting the motifs learned by this model, we hoped to identify degrons, short linear amino acid sequences that regulate the protein degradation process [6] [7].

3 Results



	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
BRCA	DPSDCC	PMPNCL	PPISKQ	FMTCWL	SILDLC
Normal	CPQSKK	FMTCWL	LKALHL	FALNLL	PPKDCL
READ/COAD	IITCTE	IVIPQI	EISMIK	WIMIII	ICCFPG
KIRC	IHIDIC	DPSDCC	SQVCKL	IMQCLG	EHKHMC
UCEC	PYCHWF	DPTNQG	CVMYQQ	PDNGYM	LGGVWF
HNSC	PMPNCL	IHIDIC	FDPVLC	HTLCKK	QHIVCT
LUAD	IMMCIN	HIKCIC	PMPNCL	IIISIT	LVLKLK
LUSC	IMMCIN	HIKCIC	PMPNCL	IIISIT	LVLKLK
PRAD	DPSDCC	YMCPLL	NDPKLL	DKIACL	SHIINL
OV	IMNIKV	ICCFIK	IIISIT	IMVNPI	MIMEIA
THCA	LVLKLK	NSLFLR	LSCNRL	QMSKLS	NNKDVL
STAD	ECICCH	IIISIT	ICCFIK	IMFCGT	IVMTIL
SKCM	HVLCCN	MMLILP	IMMCIN	VLCCRC	SLMPCM

Figure 2: The 65 anti-tumor BCR motif clusters that we discovered among 13 cancer types. Each cluster is represented in the table by its leading sequence, and the background color of each table cell denotes the size of a certain motif cluster.

Our computational pipeline clusters sequences within cancer types to characterize motifs. Figure 2 shows a table of the top five tumor-specific motifs that were discovered by applying our pipeline to analyze the inputs and outputs of DeepBCR. Each cell represents a single cluster of top sequences for a specific tumor-type condensed into a single 6-letter leading sequence. The background color of each cell denotes the significance of the respective motif cluster, which was determined by the cluster size.

We validated our results by finding strong tumor-specific correlations between the motifs from two DeepBCR models with different training parameters. Figure 3A shows the 13x13 correlation heat map, in which the color of square denotes the magnitude of the correlation between the motif clusters for Model 1 and Model 2 at the specific tumor-type combination. The heat map displays a strong correlation along the diagonal, an expected pattern because motifs for the same tumor-type should remain consistent, even if they are determined by different DeepBCR models.

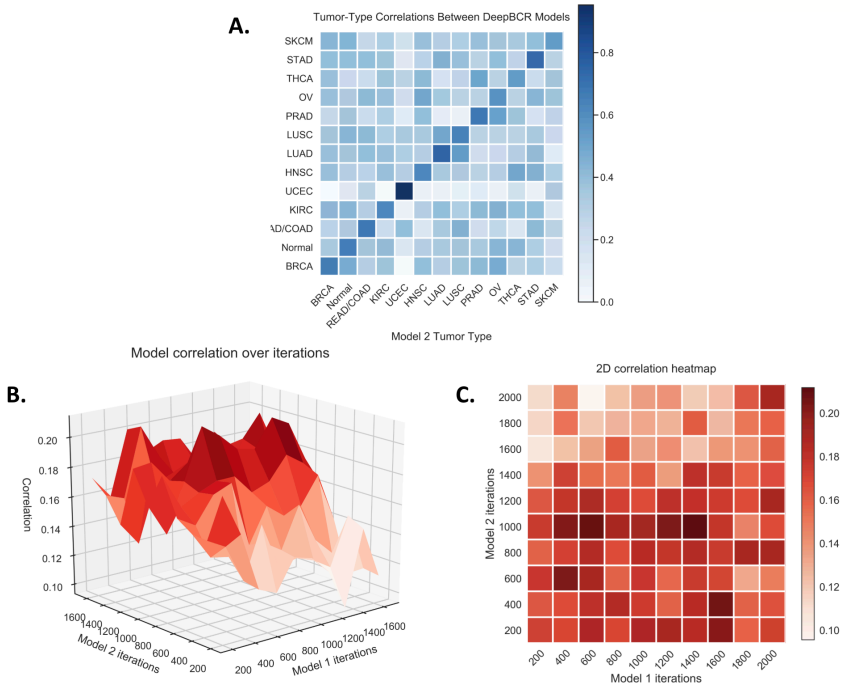


Figure 3: (A) A heat map that displays the tumor-specific correlations between two models with different training parameters. (B) and (C) are 3D and 2D heat maps, respectively, that show overall correlation scores between two DeepBCR models at training iteration combinations from 200 to 2000. The overall correlation score was calculated by finding the difference between average diagonal and non-diagonal correlation values.

We also applied our pipeline to decode the enriched motifs learned by the DeepDegron model, a second neural network. By using similar ranking, clustering, and visualizing algorithms, we were able to identify 101 degron motifs across motif lengths of 2, 3, and 4. Figure 4 shows the top motifs for each length, along with their enrichment p-value scores and the corresponding sequence logos.

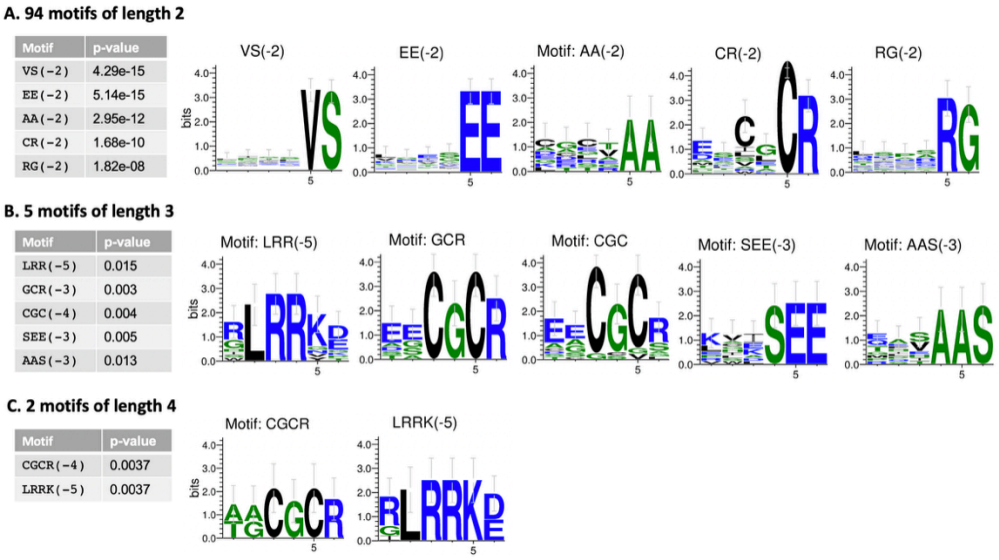


Figure 4: In total, 101 degron motifs were identified by our pipeline. The tables and sequence logos show the enriched degron motifs and their respective p-value enrichment scores for motif lengths 2, 3, and 4.

4 Conclusion

In this study, we developed a novel computational pipeline for decoding deep neural networks trained on protein sequence data. Our pipeline consists of several algorithms and stages: using the model to rank input sequences, running statistical analysis to identify top sequences, clustering sequences to form motifs, and visualizing motifs using sequence logos.

We applied our computational pipeline to decipher the information that DeepBCR neural network learned from the comprehensive TCGA cancer dataset, yielding the discovery of 65 B cell receptor motif signatures that are implicated in the BCR-antigen binding interaction. The reliability and accuracy of these motifs were determined via three independent validation methods, including the synthetic data simulation, tumor-specific correlation matrices, and binding site analyses. Our discovery and validation of these novel anti-tumor motifs can be leveraged to drive the development of targeted cancer immunotherapies and improve the identification of biomarkers used for early detection of cancers.

Finally, through applications to the transfer learning and DeepDegron models, we demonstrated that our computational pipeline can be readily adapted towards neural networks trained on other data types and sequences. The versatility and generalizability of our pipeline will enable it to be reused and employed for decoding a wide range of deep learning models, which will ultimately help lead to more transparent and reliable AI.

References

[1] Cancer Fact Sheets. From World Health Organization website: http://gco.iarc.fr/today/fact-sheets-cancers.

[2] Bruce Alberts et al. “B cells and antibodies”. In: Molecular Biology of the Cell. 4th edition. Garland Science, 2002.

[3] Yong-Chen Lu and Paul F Robbins. “Cancer immunotherapy targeting neoantigens”. In: Seminars in Immunology. Vol. 28. 1. Elsevier. 2016, pp. 22–27.

[4] Sherlock Hu. The DeepBCR model. personal conversation.

[5] Collin Tokheim. The DeepDegron model. personal conversation. June 10, 2019.

[6] Avram Hershko and Aaron Ciechanover. “The ubiquitin system for protein degradation”. In: Annual Review of Biochemistry 61.1 (1992), pp. 761–807.

[7] Balint Meszaros et al. “Degrons in cancer”. In: Science Signaling 10.470 (2017).