



Article

# Survival Distinctions for Cases Representing Immunologically Cold Tumors via Intrinsic Disorder Assessments for Blood-Sourced TRB Variable Regions

Arpan Sahoo<sup>1</sup> , Etienne C. Gozlan<sup>1</sup>, Joanna J. Song<sup>1</sup>, George Angelakakis<sup>1</sup>, Michelle Yeagley<sup>2</sup>, Boris I. Chobrutskiy<sup>3</sup>, Taha I. Huda<sup>1</sup> and George Blanck<sup>1,4,\*</sup>

<sup>1</sup> Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA; asahoo@usf.edu (A.S.); thuda@usf.edu (T.I.H.)

<sup>2</sup> University of Pittsburgh Medical Center, Pittsburgh, PA 15261, USA

<sup>3</sup> Department of Internal Medicine, Oregon Health and Science University Hospital, Portland, OR 97239, USA

<sup>4</sup> Department of Immunology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL 33612, USA

\* Correspondence: gblanck@usf.edu

**Abstract:** T cell receptor beta (TRB) sequences were recovered from the Cancer Genome Atlas Uveal Melanoma blood exome files. Intrinsic disorder scores for amino acid (AA) sequences of the entire TRB variable region were obtained and evaluated as potentially representative of overall survival (OS) distinctions, i.e., for cases representing the upper and lower 50th percentiles for intrinsic disorder scores. Analyses using four intrinsic disorder assessment tools indicated that a lower intrinsic disorder of the blood-sourced TRB variable regions, including continuous AA sequences of the V-gene segment, the complementarity-determining region-3, and the J-gene segment, was associated with a better OS probability (with log-rank *p*-values ranging from 0.002 to 0.014). We further determined that intrinsic disorder assessments could be used for OS stratification for a second, immunologically cold cancer: *MYCN* amplified neuroblastoma. Thus, intrinsic disorder assessments of blood-sourced, full TRB variable regions may provide a novel patient stratification approach for patients with immunologically cold cancers.

**Keywords:** TRB variable regions; uveal melanoma; intrinsic disorder; survival probability distinctions; neuroblastoma



**Citation:** Sahoo, A.; Gozlan, E.C.; Song, J.J.; Angelakakis, G.; Yeagley, M.; Chobrutskiy, B.I.; Huda, T.I.; Blanck, G. Survival Distinctions for Cases Representing Immunologically Cold Tumors via Intrinsic Disorder Assessments for Blood-Sourced TRB Variable Regions. *Int. J. Mol. Sci.* **2024**, *25*, 11691. <https://doi.org/10.3390/ijms252111691>

Academic Editor: Apostolos Zaravinos

Received: 24 September 2024

Revised: 21 October 2024

Accepted: 23 October 2024

Published: 30 October 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Uveal melanoma (UVM), while relatively rare, is the most common primary intraocular tumor in adults. The tumor originates from melanocytes in the uvea and often resides in the choroid, a highly vascularized tissue that provides a robust blood source to the retina [1]. Importantly, the choroid is one of the most vascularized tissues in the human body [2], supporting tumor growth and providing a route for metastasis, specifically to the liver in the case of UVM [3].

Primary UVM has classically been treated by enucleation, tumor resection, or radiation therapy, while metastatic UVM has required resection or ablation of metastatic sites in addition to delivery of medication through the hepatic artery (e.g., hepatic artery infusion chemotherapy, immune-embolization with granulocyte-macrophage colony-stimulating factor). Systemic chemotherapy and immune checkpoint inhibitors (e.g., anti-CTLA4 and anti-PD1 agents) have had limited efficacy in treating metastatic UVM [4–6], in contrast to cutaneous melanoma, where immune checkpoint inhibitors (ICIs) are an important option for treatment. The lack of effectiveness of ICIs is likely due to the fact that UVM is, at least comparatively, an immunologically cold tumor, as the eye and the primary site of metastasis, the liver, are considered to be immunologically privileged sites [6].

However, increasing evidence suggests that there may be a minimal or underappreciated ocular immune response that warrants further investigation. For example, a single-cell

RNA-seq analysis revealed clonally expanded T cells in UVM samples and identified LAG3, rather than PD-1 or CTLA4, as a potential target for immune checkpoint inhibition [7]. A recently published phase 2 trial of tebentafusp, a soluble T cell receptor bispecific, i.e., specific for a peptide-HLA complex, and the T cell expressed CD3, showed a marked improvement in overall survival (OS) rate, from 37% to 62% for metastatic UVM [8]. Also, a phase 2 trial of relatlimab, an anti-LAG3 antibody, in combination with nivolumab, an anti-PD1 antibody, showed a 47.7% progression-free survival rate at 1 year compared to 36.0% in patients treated with nivolumab alone [9]. These findings suggest that an adaptive immune response can be mounted against UVM and have supported the goal of better understanding the correlation of T cell receptor chemical features with UVM OS.

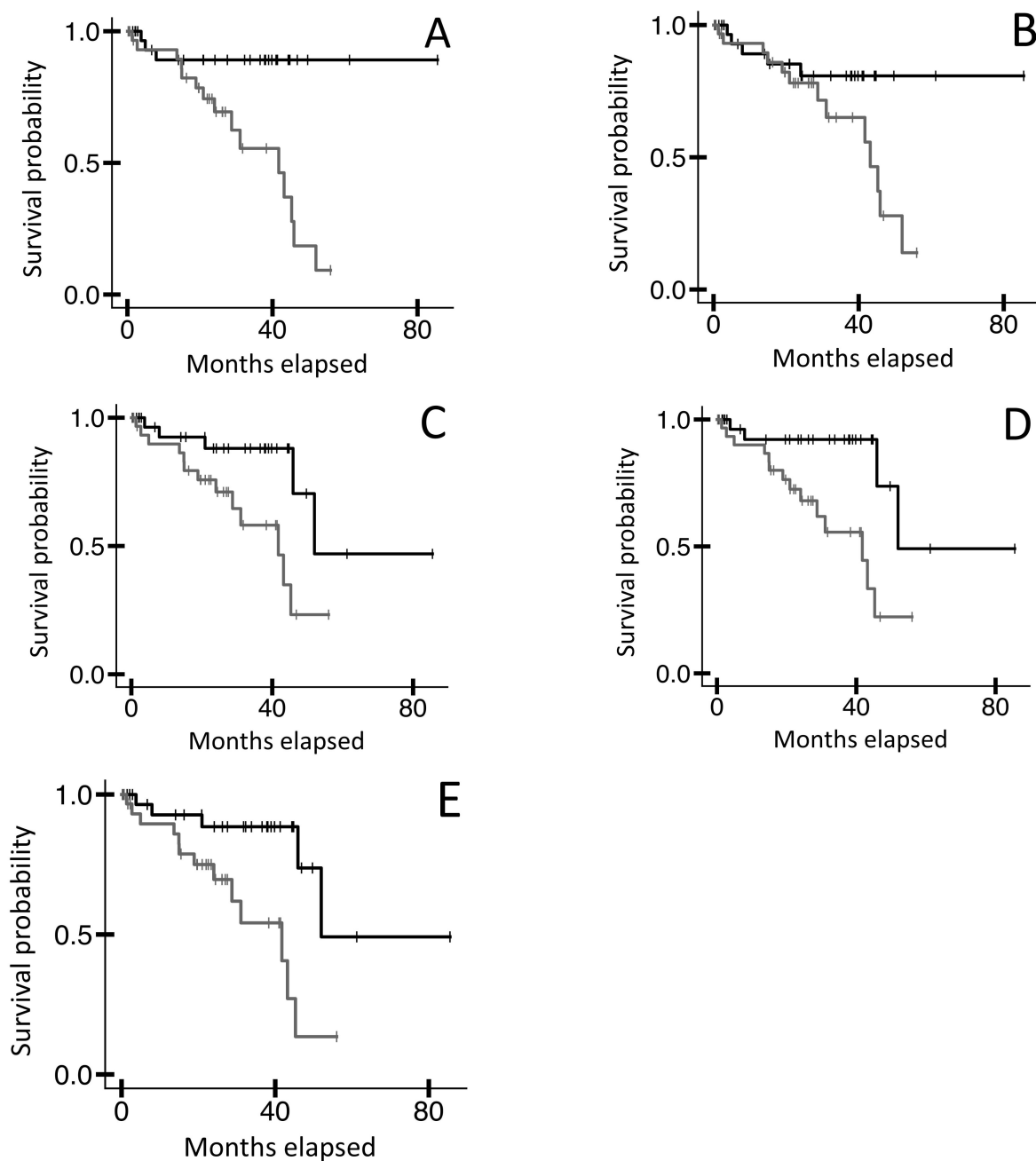
To gain further insight into the chemical features related to antigen recognition and binding, we investigated the intrinsic disorder of the TCR beta chain (TRB) variable region. Intriguingly, unstable conformations, as would be revealed by intrinsic disorder assessments, serve to provide an advantage related to the ability of CLIP to bind multiple MHC class II variations [10]. Furthermore, many intrinsically disordered proteins undergo a disorder-to-order transition upon interaction with a globular molecular binding partner. The potential biological value of having a disorder-to-order transition includes conformational flexibility as a way to bind multiple different targets, the ability to resist factors like extreme pH that would denature ordered proteins, and a reduced half-life to facilitate rapid turnover for regulatory molecules [11,12].

In this report, we evaluated full TRB variable region amino acid (AA) sequences recovered from UVM patients' blood specimens, with a focus on assessing the association between intrinsic disorder of these sequences and OS probability. We also report survival distinctions based on intrinsic disorder of blood-sourced TRB sequences for *MYCN* amplified neuroblastoma, another tumor considered to be immunologically cold [13].

## 2. Results

### 2.1. OS Probability Distinctions for UVM, Based on Assessments of Intrinsic Disorder for Blood-Sourced TRB V-CDR3-J AA Sequences

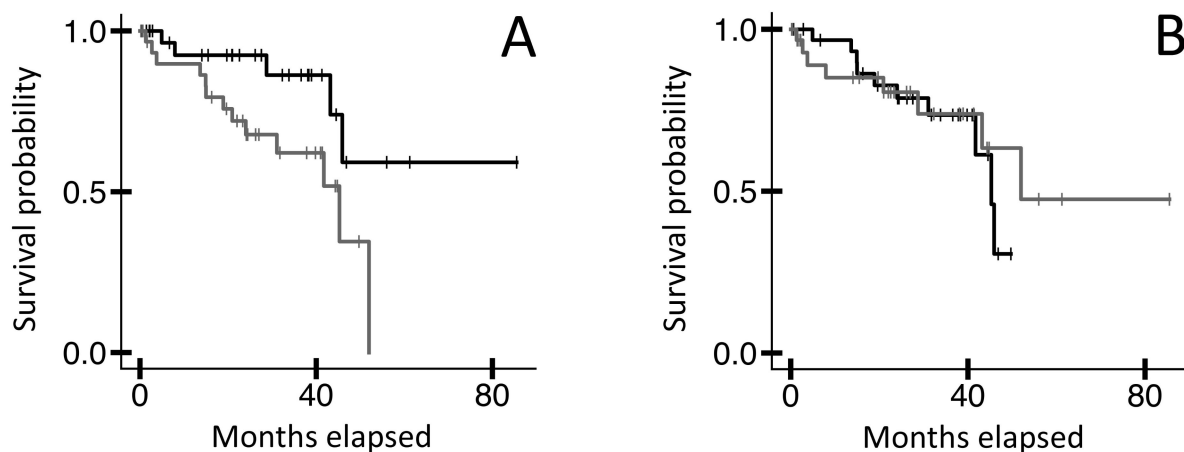
To assess intrinsic disorder for TRB V-CDR3-J AA sequences from blood specimens of UVM cases, we first mined TRB recombination reads from TCGA-UVM blood WXS files (Section 4; Table S1). The recombination reads were each defined by a validated V and J sequence, as previously described [14], and required an intervening, productive CDR3, i.e., a CDR3 AA sequence with no frameshifts or stop codons. For each recovered TRB recombination read, once the V, CDR3, and J were identified, the appropriate V and J AA sequences were integrated into the CDR3 AA sequence to produce a full-length V-CDR3-J AA sequence representing the full TRB variable region. (See Section 4 for further details.) We then employed multiple computational models, termed VSL2, VL3, IUPred2 short, IUPred2 long, and ANCHOR2, to identify regions of intrinsic disorder in the V-CDR3-J AA sequences (Section 4). Finally, we conducted an OS analysis based on each set of scores resulting from those five intrinsic disorder assessment tools. Results indicated that UVM cases represented by VSL2 scores in the lower 50th percentile (corresponding to a lower likelihood of intrinsic disorder) had a higher OS probability than cases represented by the upper 50th percentile (Figure 1A,  $n = 64$ , log-rank  $p$ -value = 0.002). This distinction was also observed with lower VL3, IUPred2 short, IUPred2 long, and ANCHOR2 scores, though the distinction with the VL3 approach represented only a trend (Figure 1B–E, log-rank  $p$ -values = 0.066, 0.014, 0.004, 0.003, respectively).



**Figure 1.** Kaplan–Meier (KM) overall survival (OS) analyses of TCGA-UVM case IDs based on intrinsic disorder assessments as applied to TRB V-CDR3-J AA sequences from blood sample WXS files. **(A)** Case IDs representing the bottom 50% of VSL2 scores (black line,  $n = 32$ , median OS N/A) versus the top 50% (grey line,  $n = 32$ , median OS 41.7 months). Log-rank  $p$ -value = 0.002. **(B)** Case IDs representing the bottom 50% of VL3 scores (black line,  $n = 32$ , median OS N/A) versus the top 50% (grey line,  $n = 32$ , median OS 43.2 months). Log-rank  $p$ -value = 0.066. **(C)** Case IDs representing the bottom 50% of IUPred2 short scores (black line,  $n = 32$ , median OS 52.0 months) versus the top 50% (grey line,  $n = 32$ , median OS 41.7 months). Log-rank  $p$ -value = 0.014. **(D)** Case IDs representing the bottom 50% of IUPred2 long scores (black line,  $n = 32$ , median OS 52.0 months) versus the top 50% (grey line,  $n = 32$ , median OS 41.7 months). Log-rank  $p$ -value = 0.004. **(E)** Case IDs representing the bottom 50% of ANCHOR2 scores (black line,  $n = 32$ , median OS 52.0 months) versus the top 50% (grey line,  $n = 32$ , median OS 41.7 months). Log-rank  $p$ -value = 0.003.

## 2.2. OS Probability Distinctions for UVM, Based on Physico-Chemical Parameters of Blood-Sourced TRB CDR3 AA Sequences

Because the CDR3, a subset of the variable region, typically represents the most important part of the adaptive IR for antigen binding, we made intrinsic disorder assessments for the CDR3 using parameters effective for the smaller CDR3 AA length. Note, the intrinsic disorder assessment tools used above are not functional for the shorter CDR3 AA lengths. Thus, we instead assessed the following four parameters for the TRB CDR3 AA sequences represented by TRB recombination reads from the blood WXS files in the TCGA-UVM dataset: proportion of (a) disorder-promoting AA residues; (b)  $\alpha$ -helix promoting residues; (c)  $\beta$ -turn promoting residues; and (d)  $\beta$ -sheet promoting residues (Section 4). We then assessed potential OS probability distinctions represented by these four parameters. Results indicated that UVM cases represented by the lower 50th percentile of (a) disorder-promoting AA residues in the TRB CDR3s had a higher OS probability than cases represented by the upper 50th percentile (Figure 2A,  $n = 64$ , log-rank  $p$ -value = 0.022). However, the parameters (b)  $\alpha$ -helix promoting residues, (c)  $\beta$ -turn promoting residues, and (d)  $\beta$ -sheet promoting residues, which represent structural stability indicators as opposed to disorder indicators, did not reflect any OS probability distinctions (for example, Figure 2B, log-rank  $p$ -value = 0.631).

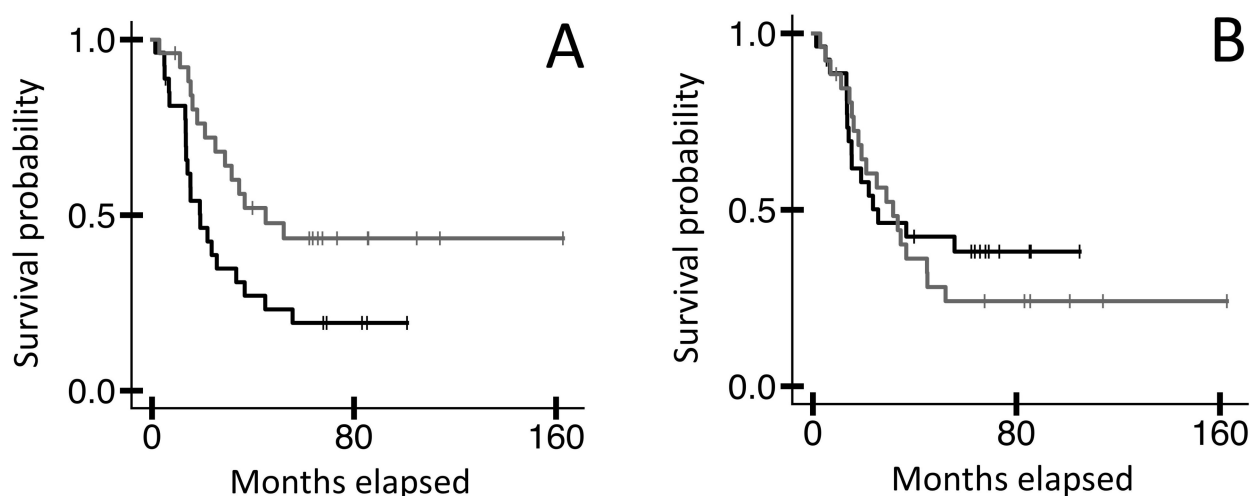


**Figure 2.** KM OS analyses of TCGA-UVM case IDs based on physico-chemical parameters of TRB CDR3 AA sequences (Table S5) from blood sample WXS files. **(A)** Case IDs representing the bottom 50% of the proportion of disorder-promoting residues (black line,  $n = 32$ , median OS N/A months) versus the top 50% (grey line,  $n = 32$ , median OS 45.3 months). Log-rank  $p$ -value = 0.022. **(B)** Case IDs representing the bottom 50% of the proportion of helix-promoting residues (black line,  $n = 32$ , median OS 45.3 months) versus the top 50% (grey line,  $n = 32$ , median OS 52.0 months). Log-rank  $p$ -value = 0.631. Similarly to the proportion of helix-promoting residues, two other physico-chemical parameters, namely the proportion of sheet-promoting residues and turn-promoting residues, did not reflect significant OS distinctions.

## 2.3. An OS Probability Distinction for MCYN Amplified NBL, Based on ANCHOR2 Assessment of Intrinsic Disorder for Blood-Sourced TRB V-CDR3-J AA Sequences

To further explore the value of assessing intrinsic disorder for the entire TRB variable region, we aimed to identify OS distinctions for another (relatively) non-immunogenic cancer type: MYCN amplified neuroblastoma [15,16]. We recovered TRB recombination reads from the blood WXS files in the TARGET-NBL dataset and established the full-length V-CDR3-J AA sequences represented by those TRB recombination reads as described above (Section 4). Then, using the intrinsic disorder assessment tools, we evaluated those V-CDR3-J AA sequences. Results indicated that MCYN amplified NBL cases represented by ANCHOR2 scores in the upper 50<sup>th</sup> percentile had a higher OS probability compared to those in the lower 50<sup>th</sup> percentile (Figure 3A,  $n = 53$ , log-rank  $p$ -value = 0.024). However,

the VSL2, VL3, IUPred2 short, and IUPred2 long approaches and scores did not reflect any OS probability distinctions (for example, Figure 3B, log-rank  $p$ -value = 0.600).



**Figure 3.** KM OS analyses of TARGET-NBL case IDs based on intrinsic disorder assessments as applied to TRB V-CDR3-J AA sequences from blood sample WXS files. **(A)** Case IDs representing the bottom 50% of ANCHOR2 scores (black line,  $n = 27$ , median OS 19.1 months) versus the top 50% (grey line,  $n = 26$ , median OS 45.0 months). Log-rank  $p$ -value = 0.024. **(B)** Case IDs representing the bottom 50% of VSL2 scores (black line,  $n = 27$ , median OS 25.6 months) versus the top 50% (grey line,  $n = 26$ , median OS 31.5 months). Log-rank  $p$ -value = 0.600. Like VSL2, the other intrinsic disorder assessments, namely VL3, IUPred2 short, and IUPred2 long, did not reflect significant OS distinctions.

#### 2.4. Multivariate Analysis

Given the statistically significant OS distinctions between the upper and lower 50th percentiles for the different intrinsic disorder assessments reported above, we aimed to account for potentially confounding clinical variables. For UVM, survival distinctions based on the VSL2, IUPred2 short, IUPred2 long, and ANCHOR2 approaches maintained their statistical significance; additionally, the survival distinction based on the proportion of disorder-promoting residues maintained its statistical significance (Table 1). For MYCN amplified NBL, the survival distinction based on the ANCHOR2 approach maintained its statistical significance (Table 2).

**Table 1.** Multivariate analysis of UVM survival per intrinsic disorder assessment and clinical features.

<i>Analysis 1</i>			
<b>Covariate</b>	<b>Exp(B)</b>	<b>Significance</b>	<b>95% CI for Exp(B)</b>
<b>VSL2</b>	9.812	0.006	1.919–50.175
<i>Age at diagnosis</i>	1.049	0.169	0.980–1.124
<i>AJCC pathologic stage</i>	4.647	0.031	1.155–18.691
<i>AJCC pathologic T</i>	0.492	0.093	0.215–1.126
<i>Received treatment</i>	3.875	0.135	0.655–22.935
<i>Fraction of genome altered</i>	8.068	0.351	0.100–647.679
<i>MSIsensor score</i>	$1.591 \times 10^3$	0.072	$0.511\text{--}4.951 \times 10^6$
<i>Analysis 2</i>			
<b>IUPred2 short</b>	5.910	0.009	1.569–22.259
<i>Age at diagnosis</i>	1.039	0.237	0.975–1.107
<i>AJCC pathologic stage</i>	4.469	0.016	1.316–15.184
<i>AJCC pathologic T</i>	0.585	0.128	0.294–1.167
<i>Received treatment</i>	2.179	0.334	0.449–10.581
<i>Fraction of genome altered</i>	16.602	0.240	$0.153\text{--}1.797 \times 10^3$
<i>MSIsensor score</i>	$2.573 \times 10^3$	0.044	$1.251\text{--}5.292 \times 10^6$

Table 1. Cont.

<i>Analysis 3</i>			
<b>IUPred2 long</b>	9.355	0.002	2.218–39.453
<i>Age at diagnosis</i>	1.041	0.227	0.975–1.111
<i>AJCC pathologic stage</i>	5.052	0.013	1.405–18.167
<i>AJCC pathologic T</i>	0.568	0.122	0.277–1.163
<i>Received treatment</i>	2.073	0.374	0.415–10.354
<i>Fraction of genome altered</i>	21.022	0.212	0.175–2.519 × 10 <sup>3</sup>
<i>MSIsensor score</i>	2.206 × 10 <sup>3</sup>	0.051	0.969–5.024 × 10 <sup>6</sup>
<i>Analysis 4</i>			
<b>ANCHOR2</b>	6.939	0.006	1.732–27.803
<i>Age at diagnosis</i>	1.044	0.184	0.980–1.112
<i>AJCC pathologic stage</i>	4.787	0.015	1.362–16.823
<i>AJCC pathologic T</i>	0.498	0.061	0.240–1.033
<i>Received treatment</i>	3.357	0.132	0.695–16.216
<i>Fraction of genome altered</i>	26.585	0.187	0.204–3.468 × 10 <sup>3</sup>
<i>MSIsensor score</i>	879.865	0.089	0.357–2.171 × 10 <sup>6</sup>
<i>Analysis 5</i>			
<b>Proportion of disorder-promoting residues</b>	4.652	0.029	1.166–18.562
<i>Age at diagnosis</i>	1.045	0.157	0.983–1.110
<i>AJCC pathologic stage</i>	2.268	0.149	0.747–6.889
<i>AJCC pathologic T</i>	0.697	0.265	0.369–1.316
<i>Received treatment</i>	3.179	0.138	0.691–14.623
<i>Fraction of genome altered</i>	17.480	0.239	0.149–2.054 × 10 <sup>3</sup>
<i>MSIsensor score</i>	4.477 × 10 <sup>3</sup>	0.030	2.213–9.059 × 10 <sup>6</sup>

Table 2. Multivariate analysis of NBL survival per intrinsic disorder assessment and clinical features.

<i>Analysis 1</i>			
<b>Covariate</b>	<b>Exp(B)</b>	<b>Significance</b>	<b>95% CI for Exp(B)</b>
<b>ANCHOR2</b>	0.401	0.012	0.197–0.816
<i>Age at diagnosis</i>	0.795	0.196	0.561–1.126
<i>Gender</i>	0.602	0.194	0.280–1.294

### 3. Discussion

Establishing survival distinctions by assessment of the patient's adaptive immune receptors (IRs) has been difficult for UVM, given the disease's low immunogenicity. We hypothesized that the application of intrinsic disorder assessment tools, which integrate an array of features ranging from charge-hydrophobicity ratio to intra-protein interaction energies and which can be applied over relatively long polypeptide lengths, would provide an opportunity for greater specificity and power of assessments. Results reported here indicate a significant association between low intrinsic disorder of the blood-sourced TRB variable region and greater overall survival (OS) for UVM patients. In fact, these findings are consistent across various models of intrinsic disorder; we found that low VSL2, IUPred2 short, IUPred2 long, and ANCHOR2 scores were all associated with a greater OS probability for UVM. These results encouraged us to also exploit blood-sourced TRB V-CDR3-Js for the identification of a survival distinction for MYCN amplified NBL, another immunologically cold tumor.

These full-length TRB variable region assessment approaches are in contrast to related approaches that have relied on the CDR3 alone. As far as the authors are aware, in all previous attempts to establish survival probability distinctions based on TCR chemistries or chemical fingerprints, the assessments were limited to the CDR3 AA sequences, based on the understanding that the CDR3 AA sequences play a major role in antigen binding. However, the relatively short AA sequence of the CDR3 places an upper limit on specificity and gradations of distinctions in comparison to assessing the entire TRB variable region.

Although the CDR3 AA sequence has a relatively unique importance in antigen contacts, it is obvious that other chemical features of the variable region also contribute to the specificity of antigen binding, even if indirectly. Thus, in sum, the results above indicate that a finer and more comprehensive assessment of adaptive IR chemistries (i.e., analysis of the full variable region) may provide opportunities to stratify patients when the immune response is comparatively low. And, again, as far as the authors are aware, this is the first time the survival of UVM cases has been distinguished based on TRB variable region chemical features.

More specifically, in the case of UVM, cases represented by lower TRB intrinsic disorder for the V-CDR3-J region represented better OS probabilities. This is somewhat in contrast to the results obtained with MYCN amplified NBL; the MYCN amplified NBL cases represented by higher ANCHOR2 scores exhibited greater OS. Note, however, that the ANCHOR2 assessment tool specifically identifies disordered binding sites within proteins that undergo a disorder-to-order transition upon interaction with a hypothetical globular binding partner. In other words, higher ANCHOR2 scores indicate the presence of context-dependent disorder, not just strict disorder. This approach differs from the other intrinsic disorder assessment tools, which more generally detect intrinsically disordered regions without regard for such context-dependent transitions. While there are no data that point to specific antigens, and certainly no data that directly connects the status of blood-sourced TRB polypeptides to an immune response against these two cancers, overall the intrinsic disorder approaches that provided the OS distinctions in this report raise the question of whether there is a specific antigen, or a relatively small set of antigens, that is the target of the TRB variable regions associated with better OS. In the case of UVM, the lower intrinsic disorder would imply rigidity and higher specificity in TRB binding to the antigen. In the case of NBL, the ANCHOR results imply the likelihood of a specific binding partner that facilitates a disorder-to-order transition for the TRB before binding the antigen.

The work reported here has several limitations. First, we analyzed only TRB recombination reads, and although TRB is generally considered more important for directly contacting antigens, TRA chemistry would obviously be important for the function of the entire T cell receptor. Second, we were unable to analyze TRB recombination reads from primary tumor specimens due to extremely low sample sizes and low immune infiltrate. Blood-sourced results are limited to a general picture of the patient's adaptive immune status, though they may have value as a non-invasive prognostic biomarker, depending, of course, on follow-up studies and verifications of results reported here. Third, while TRB chemical fingerprints may indeed lead to useful patient stratification biomarkers as well as point in the direction of more focused research involving TRB-antigen contacts, a full understanding of TRB-tumor antigen interaction would require not only *in vitro* studies but also high-resolution computational assessments, as would be provided by molecular dynamics or cryoEM approaches. However, it is important to keep in mind that our results, which are facilitated by large patient datasets and informed by survival outcomes, provide a very credible foundation for further investment into the role of the T cell receptor in regulating two immunologically cold tumors: UVM and MYCN amplified NBL. As noted in the Section 1, this is consistent with recent, specific attempts to revisit the idea of immunotherapy for UVM. And as for NBL, treatment options have recently expanded to include immunotherapy, such as anti-GD2 monoclonal antibodies as well as CAR-T cell therapy [17], the latter of which specifically motivates research to also explore TRB intrinsic disorder in NBL.

#### 4. Materials and Methods

##### 4.1. Recovery of TRB Recombination Reads from Uveal Melanoma and Neuroblastoma Datasets

We used previously described methods to recover TRB V(D)J recombination reads from whole exome sequencing (WXS) files [14,18,19]. Access to uveal melanoma (UVM) and neuroblastoma (NBL) WXS files from The Cancer Genome Atlas (TCGA, phs000178) dataset and Therapeutically Applicable Research to Generate Effective Treatments (TARGET, phs000218)

dataset was via database of genotypes and phenotypes (dbGaP) project approval numbers 6300 and 16,405, respectively. The latest version of the software for mining adaptive immune receptor (IR) recombination reads is publicly available at <https://github.com/arpansahoo/vdj> (accessed on 23 September 2024). Complete collections of the TRB recombination read data, for both UVM and NBL, are in the Supplementary Materials Tables S1 and S2.

#### *4.2. Determination of AA Sequences Representing the Full-Length TRB Variable Region from Recovered TRB Recombination Reads*

We identified the V-gene segment, the J-gene segment, and the complementarity-determining region-3 (CDR3) in each TRB recombination read using the software indicated above and following established methods [14]. The CDR3 AA sequences were translated from the recovered TRB recombination reads, whereas the V- and J-gene segment AA sequences were obtained from <https://imgt.org/> [20]. To represent the entire TRB variable region, we subsequently extended the CDR3 AA sequences into full-length V-CDR3-J AA sequences. This was performed programmatically through a string alignment procedure between the CDR3 AA sequence and its associated V and J AA sequences. We first integrated each CDR3 AA sequence with its associated V-gene segment AA sequence (i.e., the V segment as identified by the 5' end of the sequencing read to just past the first few CDR3 nucleotides). To accomplish this, we defined the longest common substring (LCS) between the CDR3 and V AA sequences, i.e., the AAs shared by the indicated V-gene segment and the first few AAs of the CDR3. (Note, the LCS had to start with the conserved cysteine from the V-gene segment). Then, the LCS and any following residues were removed from the V-gene segment AA sequence, and the full CDR3 AA sequence, including the LCS, was concatenated to the end to build the full "V-CDR3" AA sequence. Next, we integrated the "V-CDR3" sequence with its associated J-gene segment AA sequence, also identified by the sequencing read. AA residues in the J sequence preceding the phenylalanine-glycine doublet motif were removed, and the remaining J sequence was concatenated to the end of the "V-CDR3" AA sequence. This process was repeated for all CDR3 sequences to produce full-length V-CDR3-J AA sequences representing the full TRB variable regions from the WXS files. The software for this procedure can be found at the following link: <https://github.com/arpansahoo/vdj>. The full-length TRB V-CDR3-J AA sequences for both UVM and NBL can be found in Supplementary Materials Tables S1 and S2. Note, we excluded from further analyses V-CDR3-J sequences that had a single-residue "C" as their LCS for their associated V and CDR3 sequences. This filtering practice was due to an observed lack of concordance, for cases where the LCS was a "single C", between CDR3 AA sequences identified by the authors' algorithm/script and those identified by the [imgt.org](https://imgt.org) web tool that translates nucleotide sequences into V-CDR3-J AA sequences [20]. This lack of concordance was rarely seen when the LCS for the V and CDR3 obtained from the TRB recombination included at least one or two AAs following the starting C residue of the CDR3. Also, for the analyses of this report, we only considered TRB CDR3 and V-CDR3-J AA sequences derived from blood specimens and excluded such sequences from primary tumor specimens.

#### *4.3. Application of Computational Prediction Models to Assess Intrinsic Disorder of the V-CDR3-J AA Sequences*

We employed several tools to detect regions of intrinsic disorder in the blood-sourced TRB V-CDR3-J AA sequences. Note, the use of full-length V-CDR3-J sequences was necessary, e.g., in contrast to inputting the CDR3 AA sequences alone, because these tools typically require inputs to be relatively long AA sequences. We first used two models from the PONDR family: VSL2 (various, short-long version 2) and VL3 (various, long version 3), which were available at <http://pondr.com> as a web application; we designed and scripted a web crawler (Python/Selenium) to obtain VSL2 and VL3 results in a high-throughput manner. The VSL2 and VL3 models apply machine learning approaches to predict whether a region is disordered. VSL2 is a meta-predictor that combines two linear support vector machines, one trained on relatively short, disordered regions ( $\leq 30$  residues)



and one on long, disordered regions (>30 residues). The VSL2 model's features include charge-to-hydrophobicity ratio, secondary structure probability scores from PSIPRED and PHDsec, and multiple sequence alignment outputs from PSI-BLAST [21]. VL3 is a feed-forward neural network trained solely on long disordered regions. VL3's features include frequencies of certain AAs, average flexibility, and sequence complexity [22]. Additionally, we employed the IUPred2 short, IUPred2 long, and ANCHOR2 models, which are available for download at <https://iupred.elte.hu/>. In contrast to VSL2 and VL3 (the machine learning-based PONDR approaches), IUPred2 uses a biophysics-based model to identify regions of intrinsic disorder. At the core of IUPred2 is an energy estimation approach based on the assumption that disordered proteins result from insufficient stabilization of inter-residue interactions. IUPred2 uses a  $20 \times 20$  energy predictor matrix to sum interaction energies between a residue and its intrachain neighbors, classifying residues with unfavorable summed interaction energies as disordered [23]. Note that while "IUPred2 long" identifies global disorder, IUPred2 short identifies shorter disordered regions [24]. In addition, ANCHOR2 uses an energy estimation approach like that of IUPred2, but it specifically identifies disordered binding regions that transition into an ordered state upon ligand binding. The ANCHOR2 model determines the likelihood of a residue being within a disordered binding site by considering three parameters: (1) the average IUPred2 score across the residue's neighbors (does the residue belong to a disordered region?); (2) the residue's individual IUPred2 score (is the residue unable to form sufficient favorable interactions with its neighbors?); and (3) the increase in favorable energy when the residue interacts with a hypothetical globular protein instead of its intrachain neighbors (is there a disorder-to-order transition?) [23,25]. Note that all of the five above-described approaches (VSL2, VL3, IUPred2 short, IUPred2 long, and ANCHOR2) output, for each residue in an AA sequence, the probability that the residue is within an intrinsically disordered region (or disordered binding site, in the case of ANCHOR2). To obtain a single value from each of the five intrinsic disorder assessments for each V-CDR3-J AA sequence, we averaged the probabilities over AA residues, resulting in five values (mean VSL2, mean VL3, etc.) for each V-CDR3-J AA sequence. Note, there are no composite scores in this report; each algorithm (e.g., VSL2 or VL3) is used independently, and the specific use of each algorithm is indicated and detailed in Section 2. See Supplementary Materials Tables S3 and S4 for a list of the intrinsic disorder values for each V-CDR3-J AA sequence assessed in this report.

#### 4.4. Characterization of Four Physico-Chemical Parameters of CDR3 AA Sequences

We also assessed the blood-sourced TRB CDR3 AA sequences, independently of the V- and J-gene segments. However, the above intrinsic disorder assessment tools are not appropriate for such short sequences, so we instead assessed four other physico-chemical parameters representing disorder or stability: proportion of (a) disorder-promoting AA residues; (b)  $\alpha$ -helix promoting residues; (c)  $\beta$ -turn promoting residues; and (d)  $\beta$ -sheet promoting residues. This was performed with the software at the link indicated above in this Section 4, specifically with functions from Biopython's ProtParam module [26] and the Pappu group's localCIDER package [27]. The proportion of disorder-promoting residues is calculated as the fraction of residues in an AA sequence that tend to promote disorder in a protein's structure; these residues are T, A, G, R, D, H, Q, K, S, E, and P [28]. The other three parameters are defined as the fraction of residues in a sequence that tend to be enriched in an  $\alpha$ -helix,  $\beta$ -turn, or  $\beta$ -sheet. For helix, the residues are M, A, L, E, and K [29]; for turn, the residues are N, P, G, S, and D [30]; and, for sheet, the residues are V, I, T, Y, F, W, and L [31]. See Supplementary Materials Table S5 for a list of the physico-chemical parameter values for each CDR3 AA sequence assessed in this report.

#### 4.5. Survival Analysis

Clinical data for the UVM and NBL cases were obtained from [cBioPortal.org](https://www.ncbi.nlm.nih.gov/bioproject/1000000000) [32,33]. We filtered NBL cases to those marked by the presence of MYCN amplification. For each cancer, we analyzed the association of various parameters representing the cases, such

as the average proportion of disorder-promoting AA residues or the average ANCHOR2 score, with OS. Note, each parameter was calculated as the average of the total set of values (e.g., ANCHOR2 values) for the collection of the many blood-sourced TRB CDR3 or V-CDR3-J AA sequences obtained for any individual case. Cases were then split into two groups based on quantification of the case average for the parameter being assessed, i.e., upper vs. lower 50th percentile, and Kaplan–Meier (KM) survival analyses were performed with the Python package “lifelines” [26]. Statistical comparison of OS plots was conducted with the “lifelines” package’s implementation of the log-rank test. See Supplementary Materials Tables S6–S8 for the KM input data.

#### 4.6. Multivariate Analysis

Multivariate Cox regression analyses were performed with the Statistical Product and Service Solutions (formerly, Statistical Package for the Social Sciences; IBM SPSS v29). These multivariate analyses were performed to expand upon the univariate survival analyses detailed above while accounting for clinical covariates [34,35]. All parameters used in the multivariate analyses first underwent a univariate assessment, and no parameter was used in the multivariate analyses without having a univariate result of  $p = 0.1$  or less.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms252111691/s1>.

**Author Contributions:** A.S.: Conceptualization; Methodology; Software; Visualization; Writing—original draft; Writing—review and editing. E.C.G.: Formal analysis; Methodology. J.J.S.: Methodology. G.A.: Methodology. M.Y.: Formal analysis; Methodology. B.I.C.: Methodology; Software. T.I.H.: Methodology; Formal analysis. G.B.: Methodology; Project administration; Resources; Supervision; Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Access to TCGA uveal melanoma (UVM) and the TARGET NBL blood WXS files was via dbGaP project approval numbers 6300 and 16405, respectively, to G.B.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Most of the data upon which these analyses were based are available in the Supplementary Materials; additional data are publicly available at [cbioportal.org](https://www.ncbi.nlm.nih.gov/bioproject/1000000000).

**Acknowledgments:** Authors thank USF research computing; Corinne Walters for expert admin support of data access applications to the NIH; and the taxpayers of the State of Florida.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Kaliki, S.; Shields, C.L. Uveal melanoma: Relatively rare but deadly cancer. *Eye* **2017**, *31*, 241–257. [[CrossRef](#)] [[PubMed](#)]
2. Nickla, D.L.; Wallman, J. The multifunctional choroid. *Prog. Retin. Eye Res.* **2010**, *29*, 144–168. [[CrossRef](#)] [[PubMed](#)]
3. Tarlan, B.; Kiratli, H. Uveal melanoma: Current trends in diagnosis and management. *Turk. J. Ophthalmol.* **2016**, *46*, 123–137. [[CrossRef](#)] [[PubMed](#)]
4. Carvajal, R.D.; Sacco, J.J.; Jager, M.J.; Eschelmann, D.J.; Olofsson Bagge, R.; Harbour, J.W.; Chieng, N.D.; Patel, S.P.; Joshua, A.M.; Piperno-Neumann, S. Advances in the clinical management of uveal melanoma. *Nat. Rev. Clin. Oncol.* **2023**, *20*, 99–115. [[CrossRef](#)] [[PubMed](#)]
5. Gelmi, M.C.; Jager, M.J. Uveal melanoma: Current evidence on prognosis, treatment and potential developments. *Asia Pac. J. Ophthalmol.* **2024**, *13*, 100060. [[CrossRef](#)] [[PubMed](#)]
6. Heptt, M.V.; Steeb, T.; Schlager, J.G.; Rosumeck, S.; Dressler, C.; Ruzicka, T.; Nast, A.; Berking, C. Immune checkpoint blockade for unresectable or metastatic uveal melanoma: A systematic review. *Cancer Treat. Rev.* **2017**, *60*, 44–52. [[CrossRef](#)] [[PubMed](#)]
7. Durante, M.A.; Rodriguez, D.A.; Kurtenbach, S.; Kuznetsov, J.N.; Sanchez, M.I.; Decatur, C.L.; Snyder, H.; Feun, L.G.; Livingstone, A.S.; Harbour, J.W. Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat. Commun.* **2020**, *11*, 496. [[CrossRef](#)] [[PubMed](#)]
8. Carvajal, R.D.; Butler, M.O.; Shoushtari, A.N.; Hassel, J.C.; Ikeguchi, A.; Hernandez-Aya, L.; Nathan, P.; Hamid, O.; Piulats, J.M.; Rieth, M.; et al. Clinical and molecular response to tebentafusp in previously treated patients with metastatic uveal melanoma: A phase 2 trial. *Nat. Med.* **2022**, *28*, 2364–2373. [[CrossRef](#)] [[PubMed](#)]

9. Tawbi, H.A.; Schadendorf, D.; Lipson, E.J.; Ascierto, P.A.; Matamala, L.; Castillo Gutierrez, E.; Rutkowski, P.; Gogas, H.J.; Lao, C.D.; De Menezes, J.J.; et al. Relatlimab and Nivolumab versus Nivolumab in Untreated Advanced Melanoma. *N. Engl. J. Med.* **2022**, *386*, 24–34. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
10. Uversky, V.N.; Tu, Y.N.; Nwogu, O.; Butler, S.N.; Ramsamoj, M.; Blanck, G. High-level intrinsic disorder explains the universality of CLIP binding to diverse MHC class II variants. *Cell. Mol. Immunol.* **2017**, *15*, 76–78. [[CrossRef](#)] [[PubMed](#)]
11. Uversky, V.N. Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* **2002**, *11*, 739–756. [[CrossRef](#)] [[PubMed](#)]
12. Uversky, V.N. Paradoxes and wonders of intrinsic disorder: Stability of instability. *Intrinsically Disord. Proteins* **2017**, *5*, e1327757. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
13. Wienke, J.; Dierselhuis, M.P.; Tytgat, G.A.M.; Künkele, A.; Nierkens, S.; Molenaar, J.J. The immune landscape of neuroblastoma: Challenges and opportunities for novel therapeutic strategies in pediatric oncology. *Eur. J. Cancer* **2021**, *144*, 123–150. [[CrossRef](#)] [[PubMed](#)]
14. Chobrutskiy, B.I.; Zaman, S.; Tong, W.L.; Diviney, A.; Blanck, G. Recovery of T-cell receptor V(D)J recombination reads from lower grade glioma exome files correlates with reduced survival and advanced cancer grade. *J. Neurooncol.* **2018**, *140*, 697–704. [[CrossRef](#)] [[PubMed](#)]
15. Layer, J.P.; Kronmuller, M.T.; Quast, T.; van den Boorn-Konijnenberg, D.; Effern, M.; Hinze, D.; Althoff, K.; Schramm, A.; Westermann, F.; Peifer, M.; et al. Amplification of N-Myc is associated with a T-cell-poor microenvironment in metastatic neuroblastoma restraining interferon pathway activity and chemokine expression. *Oncoimmunology* **2017**, *6*, e1320626. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
16. Zhang, P.; Wu, X.; Basu, M.; Dong, C.; Zheng, P.; Liu, Y.; Sandler, A.D. MYCN Amplification Is Associated with Repressed Cellular Immunity in Neuroblastoma: An In Silico Immunological Analysis of TARGET Database. *Front. Immunol.* **2017**, *8*, 1473. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
17. Zhou, X.; Wang, X.; Li, N.; Guo, Y.; Yang, X.; Lei, Y. Therapy resistance in neuroblastoma: Mechanisms and reversal strategies. *Front. Pharmacol.* **2023**, *14*, 1114295. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
18. Tong, W.L.; Tu, Y.N.; Samy, M.D.; Sexton, W.J.; Blanck, G. Identification of immunoglobulin V(D)J recombinations in solid tumor specimen exome files: Evidence for high level B-cell infiltrates in breast cancer. *Hum. Vaccines Immunother.* **2017**, *13*, 501–506. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
19. Gill, T.R.; Samy, M.D.; Butler, S.N.; Mauro, J.A.; Sexton, W.J.; Blanck, G. Detection of Productively Rearranged TcR-alpha V-J Sequences in TCGA Exome Files: Implications for Tumor Immunoscoring and Recovery of Antitumor T-cells. *Cancer Inf.* **2016**, *15*, 23–28. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
20. Giudicelli, V.; Duroux, P.; Ginestoux, C.; Folch, G.; Jabado-Michaloud, J.; Chaume, D.; Lefranc, M.P. IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res.* **2006**, *34*, D781–D784. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
21. Peng, K.; Radivojac, P.; Vucetic, S.; Dunker, A.K.; Obradovic, Z. Length-dependent prediction of protein intrinsic disorder. *BMC Bioinform.* **2006**, *7*, 208. [[CrossRef](#)] [[PubMed](#)]
22. Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K. Predicting intrinsic disorder from amino acid sequence. *Proteins Struct. Funct. Bioinform.* **2003**, *53*, 566–572. [[CrossRef](#)] [[PubMed](#)]
23. Mészáros, B.; Erdős, G.; Dosztányi, Z. IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **2018**, *46*, W329–W337. [[CrossRef](#)] [[PubMed](#)]
24. Dosztányi, Z.; Csizmok, V.; Tompa, P.; Simon, I. IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **2005**, *21*, 3433–3434. [[CrossRef](#)] [[PubMed](#)]
25. Mészáros, B.; Simon, I.; Dosztányi, Z. Prediction of protein binding regions in disordered proteins. *PLOS Comput. Biol.* **2009**, *5*, e1000376. [[CrossRef](#)] [[PubMed](#)]
26. Cock, P.J.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
27. Holehouse, A.S.; Das, R.K.; Ahad, J.N.; Richardson, M.O.; Pappu, R.V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys. J.* **2017**, *112*, 16–21. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
28. Campen, A.; Williams, M.R.; Brown, J.C.; Meng, J.; Uversky, N.V.; Dunker, K.A. TOP-IDP-Scale: A new amino acid scale measuring propensity for intrinsic disorder. *Protein Pept. Lett.* **2008**, *15*, 956–963. [[CrossRef](#)] [[PubMed](#)]
29. Haimov, B.; Srebnik, S. A closer look into the  $\alpha$ -helix basin. *Sci. Rep.* **2016**, *6*, 38341. [[CrossRef](#)] [[PubMed](#)]
30. Hutchinson, E.G.; Thornton, J.M. A revised set of potentials for  $\beta$ -turn formation in proteins. *Protein Sci.* **1994**, *3*, 2207–2216. [[CrossRef](#)] [[PubMed](#)]
31. Kim, C.A.; Berg, J.M. Thermodynamic  $\beta$ -sheet propensities measured using a zinc-finger host peptide. *Nature* **1993**, *362*, 267–270. [[CrossRef](#)] [[PubMed](#)]
32. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]

33. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2012**, *2*, 401–404. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
34. Patel, A.R.; Patel, D.N.; Tu, Y.N.; Yeagley, M.; Chobrutskiy, A.; Chobrutskiy, B.I.; Blanck, G. Chemical complementarity between immune receptor CDR3s and candidate cancer antigens correlating with reduced survival: Evidence for outcome mitigation with corticosteroid treatments. *J. Biomol. Struct. Dyn.* **2022**, *41*, 4632–4640. [[CrossRef](#)] [[PubMed](#)]
35. Yeagley, M.; Chobrutskiy, B.I.; Gozlan, E.C.; Medikonda, N.; Patel, D.N.; Falasiri, S.; Callahan, B.M.; Huda, T.; Blanck, G. Electrostatic Complementarity of T-Cell Receptor-Alpha CDR3 Domains and Mutant Amino Acids Is Associated with Better Survival Rates for Sarcomas. *Pediatr. Hematol. Oncol.* **2021**, *38*, 251–264. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.