



Article

The Proteomic Analysis of Cancer-Related Alterations in the Human Unfoldome

Victor Paromov ^{1,†}, Vladimir N. Uversky ^{2,†} , Ayorinde Cooley ³, Lincoln E. Liburd II ⁴ , Shyamali Mukherjee ⁴, Insung Na ², Guy W. Dayhoff II ⁵ and Siddharth Pratap ^{1,3,*}

- ¹ Meharry Proteomics Core, RCMI Research Capacity Core, School of Medicine, Meharry Medical College, Nashville, TN 37208, USA; vparomov@mmc.edu
- ² Department of Molecular Medicine, USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33613, USA; vuversky@usf.edu (V.N.U.); nais8303@gmail.com (I.N.)
- ³ Meharry Bioinformatics Core, Department of Microbiology, Immunology and Physiology, School of Medicine, Meharry Medical College, Nashville, TN 37208, USA; acooley@mmc.edu
- ⁴ Department of Biochemistry, Cancer Biology, Neuroscience & Pharmacology, School of Medicine, Meharry Medical College, Nashville, TN 37208, USA; lliburd23@mmc.edu (L.E.L.II); smukherjee@mmc.edu (S.M.)
- ⁵ Department of Chemistry, College of Art and Sciences, University of South Florida, Tampa, FL 33613, USA; gdayhoff@usf.edu
- * Correspondence: spratap@mmc.edu
- † These authors contributed equally to this work.

Abstract: Many proteins lack stable 3D structures. These intrinsically disordered proteins (IDPs) or hybrid proteins containing ordered domains with intrinsically disordered protein regions (IDPRs) often carry out regulatory functions related to molecular recognition and signal transduction. IDPs/IDPRs constitute a substantial portion of the human proteome and are termed “the unfoldome”. Herein, we probe the human breast cancer unfoldome and investigate relations between IDPs and key disease genes and pathways. We utilized bottom-up proteomics, MudPIT (Multidimensional Protein Identification Technology), to profile differentially expressed IDPs in human normal (MCF-10A) and breast cancer (BT-549) cell lines. Overall, we identified 2271 protein groups in the unfoldome of normal and cancer proteomes, with 148 IDPs found to be significantly differentially expressed in cancer cells. Further analysis produced annotations of 140 IDPs, which were then classified to GO (Gene Ontology) categories and pathways. In total, 65% (91 of 140) IDPs were related to various diseases, and 20% (28 of 140) mapped to cancer terms. A substantial portion of the differentially expressed IDPs contained disordered regions, confirmed by in silico characterization. Overall, our analyses suggest high levels of interactivity in the human cancer unfoldome and a prevalence of moderately and highly disordered proteins in the network.

Keywords: unfoldome; intrinsically disordered proteins (IDPs); intrinsically disordered protein regions (IDPRs)



Citation: Paromov, V.; Uversky, V.N.; Cooley, A.; Liburd, L.E., II; Mukherjee, S.; Na, I.; Dayhoff, G.W., II; Pratap, S. The Proteomic Analysis of Cancer-Related Alterations in the Human Unfoldome. *Int. J. Mol. Sci.* **2024**, *25*, 1552. <https://doi.org/10.3390/ijms25031552>

Academic Editor: Antonio Lucacchini

Received: 1 December 2023

Revised: 11 January 2024

Accepted: 11 January 2024

Published: 26 January 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

Many functional proteins lack stable 3D structures as a whole or in their substantial parts due to specific and discoverable features at the primary amino acid sequence level. These intrinsically disordered proteins (IDPs) and intrinsically disordered protein regions (IDPRs) often carry out regulatory functions and are related to molecular recognition, regulation, and signal transduction, with disorder-based functionality being complementary to the common catalysis and transport activities of proteins with well-defined, stable three-dimensional structures [1–26]. Therefore, protein functionality may originate from both order and disorder, and the classic structure-function paradigm, which emphasizes that ordered 3D structures represent the indispensable prerequisite to effective protein functioning, should be redefined to include IDPs/IDPRs [1,3,4,6,8,9,12,13,26,27].

Structurally, IDPs range from completely unstructured polypeptides and extended partially structured forms to compact disordered ensembles containing substantial levels of secondary structures [3,4,6]. Furthermore, many proteins are order-disorder hybrids, containing a mix of ordered domains and IDPRs. These hybrid proteins contain regions with different degrees and flavors of disorder and in general, are characterized by a remarkable degree of structural heterogeneity. This very complex and heterogeneous spatiotemporal structural organization of a protein molecule includes foldons (independent foldable units of a protein), inducible foldons (disordered regions that can fold at least in part due to interactions with binding partners), morphing inducible foldons (disordered regions that can fold into different structures upon interaction with different partners), non-foldons (non-foldable protein regions), semi-foldons (regions that are always in a semi-folded form), and unfoldons (ordered regions that have to undergo an order-to-disorder transition to become functional) [28]. In its turn, this complex structural organization defines the unique functionality of IDPs/IDPRs, where differently (dis)ordered structural elements might have well-defined and specific functions, thereby defining the possibility of a protein molecule being multifunctional [28]. This also defines the relationships of protein structure and function in the form of a structure-function continuum concept, where instead of the classical “one gene–one protein–one structure–one function” view, the actual protein structure-function relationship is described by the more convoluted “one-gene–many-proteins–many-functions” model [28].

Since IDPs/IDPRs constitute a substantial portion of any known proteome [29–32], have amazing structural variability, and possess a very wide variety of functions, the concepts of unfoldome and unfoldomics have been introduced [13]. Here, unfoldomics represents a sub-field of protein science that studies the unfoldome, which includes a set of intrinsically disordered proteins (also known as natively unfolded proteins, therefore unfoldome), their functions, structures, interactions, evolution, etc. [13]. Furthermore, intrinsic disorder is tightly connected to the pathogenesis of various human diseases, since the majority of human cancer-associated proteins [7,33–40]—as well as many proteins associated with neurodegeneration [41–47], diabetes [48–51], major psychiatric disorders [52], and cardiovascular diseases [48,53–57]—are either intrinsically disordered or contain long IDPRs, giving rise to the D^2 (disorder in disorders) concept [58]. Curiously, based on the computational analysis of the diseasome network organized by human genetic diseases and related genes [59], it was concluded that this diseasome represents a human-genetic-disease-associated unfoldome, where intrinsic disorder and disorder-based interaction sites are commonly found in proteins associated with human genetic diseases and where proteins from different classes of genetic disease possess different levels of intrinsic disorder [60]. It was also emphasized that disease-associated mutations are commonly found within IDPRs [54].

Based on the systematic comparison of the peculiarities of the amino acid sequences of ordered proteins/domains and IDPs/IDPRs, it was found that IDPs/IDPRs differ from ordered globular proteins and domains at many levels. Among the most notable differences are dissimilarities in amino acid compositions, sequence complexity, hydrophobicity, charge, and flexibility [3,12,61]. At a more detailed level, IDPs/IDPRs were shown to be systematically depleted in order-promoting residues (Cys, Trp, Ile, Tyr, Phe, Leu, His, and Val), being noticeably enriched in disorder-promoting amino acids (Pro, Ala, Glu, Ser, Gln, Lys, Gly, Asp, and Arg) [3,12,61].

Furthermore, the lack of compact structures in highly disordered IDPs/IDPRs (i.e., IDPs with so-called extended disorders, such as coil-like and pre-molten globule-like) was attributed to the characteristic combination of low mean hydrophobicity and relatively high net charge [2]. These peculiarities in the amino acid sequences of extended IDPs define their “turned out” response to heat, where increases in temperature cause the formation of more ordered secondary structures [62–64]. These structure-promoting effects of elevated temperatures were attributed to the increased strength of hydrophobic interactions at higher temperatures, leading to stronger hydrophobic attractions, which is the major driving force

for folding [64]. Analogously, extended IDPs show a “turned out” response to changes in pH [64–67], where partial folding was shown to be induced in extended IDPs by a decrease (or increase) in pH of the media. This pH-induced partial folding of extended IDPs/IDPRs was attributed to the minimization of their large net charge at a neutral pH, leading to a decrease in intramolecular electrostatic repulsion and thereby permitting the hydrophobicity-driven collapse of the partially folded conformation [28,64].

Importantly, it was emphasized that the aforementioned structural features of extended IDPs and their specific conformational behavior can be utilized in the large-scale identification of these important members of the protein kingdom, where the resistance of such proteins to acid- or heat-treatment can be utilized for their targeted identification [68–70]. In fact, one such method is based on the finding that many proteins that fail to precipitate during perchloric acid or trichloroacetic acid treatment are IDPs [68], whereas another method utilizes the fact that IDPs possess high resistance toward heat-induced aggregation [68–70].

Therefore, since IDPs/IDPRs are often associated with various diseases, including cancer, we suggest that a focused unfoldome characterization would represent a valuable addition to cancer discovery research. To this end, here, we probe a human breast cancer unfoldome generated by the treatment of human normal (MCF-10A) and breast cancer (BT-549) cells with 1.5% trichloroacetic acid, followed by IDP precipitation and an LC-MS/MS analysis, and investigated the relations of these IDPs to key disease genes and pathways in silico using bioinformatics approaches.

2. Results and Discussion

2.1. Proteomics Analysis

Utilizing a bottom-up proteomics approach and Multidimensional Protein Identification Technology (MudPIT) methods [71–81], we compared the unfoldomics portions of normal (MCF-10A cell lysates) and cancerous (BT-549 cell lysates) proteomes. In general, IDPs show altered physical properties in solution under physiological conditions—e.g., higher resistance to temperature, light, or pH-induced precipitation [64,68]. To remove folded proteins, we treated the cell lysates with 1.5% trichloroacetic acid and sedimented protein pellets by centrifugation. Next, we sedimented all the proteins using 20% trichloroacetic acid, and thus, we obtained a fraction of a proteome highly enriched in IDPs. Our MudPIT procedure identified 2271 protein groups in these unfoldomic fractions (Figures 1 and 2). The unfoldomic enrichment revealed substantial differences between normal and cancer unfoldomes: 148 IDPs were significantly up or downregulated in cancer cells (≥ 2 -fold; $p \leq 0.05$). The further analysis of these differentially present proteins (DPPs) allowed annotation of 140 proteins using the Database for Annotation, Visualization, and Integrated Discovery knowledgebase (DAVID) (v6.8, accessed on 12/2023) [82–84]. Notably, 65% (91 of 140) IDPs were related to various diseases, and 20% (28 of 140) of IDPs were related to cancer (GAD_DISEASE database). These data suggest the substantial alteration of the human unfoldome in cancer.

2.2. In Silico Protein Disorder Analysis

Figure 3 provides a representative characterization of the 140 DPPs identified in this study. Although the proteins assembled in the human cancer unfoldome range in length from 68 (Copper transport protein ATOX1, UniProt ID: O00244) to 5871 residues (nesprin-2 isoform X12, RefSeq ID: XP_011534885.1), Figure 3A that shows the length distribution of these proteins demonstrates that most identified proteins are rather short (less than 400 residues). Since amino acid compositions of ordered and intrinsically disordered proteins are known to be noticeably different, with IDPs possessing lower levels of order-promoting (mostly hydrophobic) residues and higher levels of polar and charged residues [3,12,61], we utilized the Composition Profiler, a tool for the detection of enrichment or depletion patterns of individual amino acids within query proteins. The results (summarized in Figure 3B) clearly show that proteins in the human cancer unfoldome

are depleted in most order-promoting residues and enriched in most disorder-promoting residues, highly suggesting that these proteins contain significant levels of intrinsic disorder.

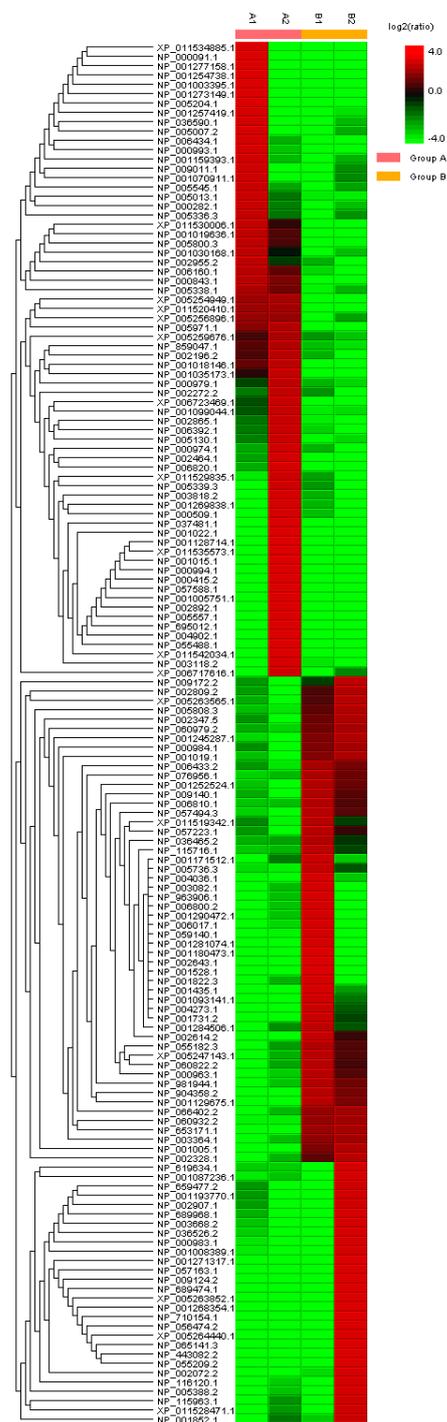


Figure 1. Heat map with hierarchical clustering of the 148 DPPs in the unfoldomics portions of the normal (MCF-10A cell lysates) versus cancerous (BT-549 cell lysates) human proteomes using PEAKS Studio proteomics software (PEAKS Version 9, accessed on 12/2023) [85]. Clustering: rows are centered; unit variance scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage. For the heat map, the legend on the right indicates the color key for intermediate ratios: red color = up-regulated proteins; green color = down-regulated proteins; color gradation according to the \log_2 ratio (as shown).

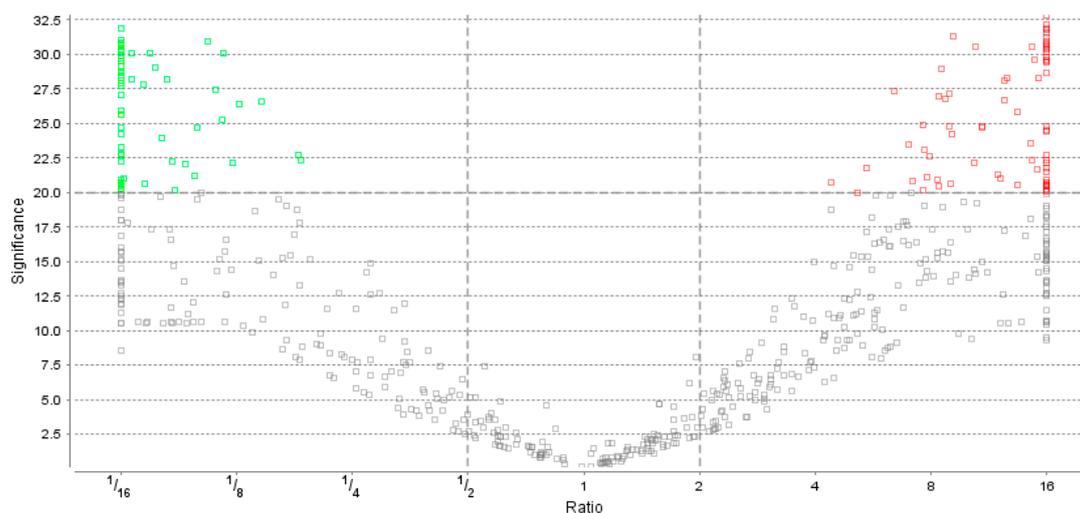


Figure 2. Volcano plot showing distribution of the up-regulated (red) and down-regulated (green) proteins in the unfoldome portions of the normal (MCF-10A cell lysates) and cancerous (BT-549 cell lysates) human proteomes. Filtering parameters: PEAKS significance score ≥ 20 or $-10_{\log P} > 2$ (equivalent to $p \leq 0.01$) and fold change $\geq |2|$. (PEAKS Version 9, accessed on 12/2023).

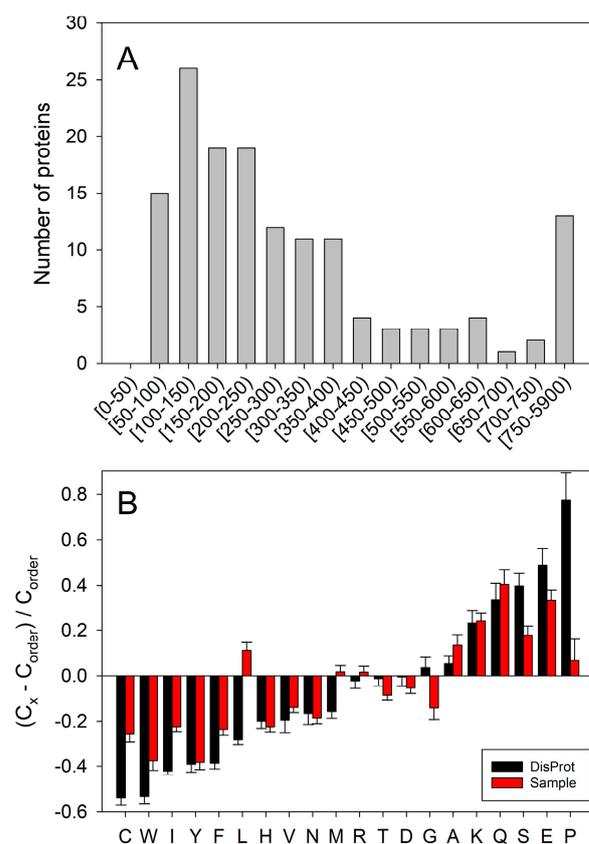


Figure 3. Characteristics of 140 human proteins differently expressed in the cancer unfoldome. (A) Length distribution of proteins. (B) Compositional profiling of proteins conducted using the web-based computation tool, Composition Profiler (<http://profler.cs.ucr.edu>, accessed on 20 December 2023) [86]. This tool generates fractional compositional differences calculated as $(C_x - C_{order}) / C_{order}$, where C is the content of a given amino acid in a query protein (set), and C_{order} is the corresponding value for the set of ordered proteins from PDB Select 25 (a subset of all PDB proteins sharing the set of ordered proteins from

PDB Select 25 (a subset of all PDB proteins sharing $\leq 25\%$ sequence identity) [87]. The magnitudes of the bars indicate fractional differences between amino acid compositions, where positive/negative values indicate enrichment/depletion, respectively. Error bars indicate standard deviations of fractional differences based on Composition Profiler bootstrapping. Black bars correspond to a similar analysis for proteins via DisProt [88].

The validity of this conclusion is supported by Figure 4, which represents the outputs of different intrinsic disorder predictors for a set of query proteins. Since the average disorder score (ADS) of a given protein is not directly related to its percent of predicted disordered residues (PPDR) value, Figure 4A shows the PPDR vs. ADS plots generated for human proteins that are differently expressed in cancer based on the output of Rapid Intrinsic Disorder Analysis Online (RIDAO) [89]. RIDAO yields results for IUPred (short), IUPred (long), PONDR[®] VL3, PONDR[®] VLXT, PONDR[®] VSL2, and PONDR[®] FIT and computes a mean disorder score for each residue based on these predictors. The plots illustrate a reasonable agreement between the outputs of different intrinsic disorder predictors for the analyzed proteins. These data can also be used for the classification of proteins by their overall disorder characteristics. In fact, two arbitrary cutoffs for the levels of intrinsic disorder (PPDR values) are typically used to classify proteins as highly ordered (PPDR < 10%), moderately disordered ($10\% \leq \text{PPDR} < 30\%$), and highly disordered (PPDR $\geq 30\%$) [90]. Similarly, proteins can be classified based on their ADS values, with proteins being considered as highly ordered, moderately disordered, and highly disordered if their ADS values are < 0.25 , between 0.25 and 0.5, or ≥ 0.5 , respectively. Therefore, Figure 4A demonstrates that the majority of analyzed DPPs are either moderately or highly disordered. Furthermore, this analysis allowed us to select the most disordered proteins in the set, which were defined here as proteins possessing ADS ≥ 0.5 or PPDR $\geq 30\%$ (by any of the predictors utilized in this study). We found that 103 (69.6%) and 77 (52.0%) of the DPPs satisfy these criteria, respectively.

Figure 4B compares two more conservative ways of evaluation of intrinsic disorder—PPDRs evaluated by the meta-predictor PONDR[®] FIT and the RIDAO mean disorder score—and shows the presence of reasonable correlation between these approaches.

Next, we compared the outputs of PONDR[®] VLXT, PONDR[®] VSL2, and PONDR[®] FIT in the form of a 3D scatter plot, where the PPDR values generated by these three disorder predictors are plotted against each other for all the DPPs (Figure 4C). Based on this analysis and the aforementioned PPDR thresholds, 21, 51, and 76 members of the human cancer unfoldome are ordered, moderately disordered, and highly disordered proteins, respectively (see Figure 4C). In other words, according to these criteria, 85.8% of these proteins are predicted to possess noticeable levels of intrinsic disorder. Similarly, we created a 3D scatter plot correlating the corresponding ADS values (Figure 4D). Based on this analysis and the aforementioned ADS thresholds, 29, 72, and 47 human proteins differently expressed in cancer are expected to be ordered, moderately disordered, and highly disordered, respectively (see Figure 4D), suggesting that by these criteria, 80.4% proteins contain high levels of disorder.

Important information on the global classification of disorder status of query proteins can be obtained from the analysis of binary disorder predictors (i.e., computational tools that classify proteins as wholly ordered or wholly disordered). Because of the principal differences in the criteria used by two such predictors, the charge-hydrophathy (CH) plot [2,31] and the cumulative distribution function (CDF) plot [31], their combination in the form of a CH-CDF plot gives an opportunity for the unique assessment of intrinsic disorder in several categories, allowing for the predictive classification of proteins into structurally different classes [31]. The results of this analysis are shown in Figure 5A, where based on their positions within the CH-CDF phase space: proteins are classified as ordered (proteins predicted as ordered and compact by both the CDF and CH, which are located within the lower-right quadrant (Q1)), native molten globules, or hybrid proteins containing sizable levels of order and disorder (proteins predicted to be disordered by the CDF but compact by the CH-plot, which are located within the lower-left quadrant (Q2)); proteins with

extended disorder, such as native coils and native pre-molten globules (proteins predicted to be disordered by both methods, which are found within the upper-left quadrant (Q3)); and proteins predicted to be disordered by the CH-plot but ordered by the CDF (located within the upper-right quadrant (Q4)) [31].

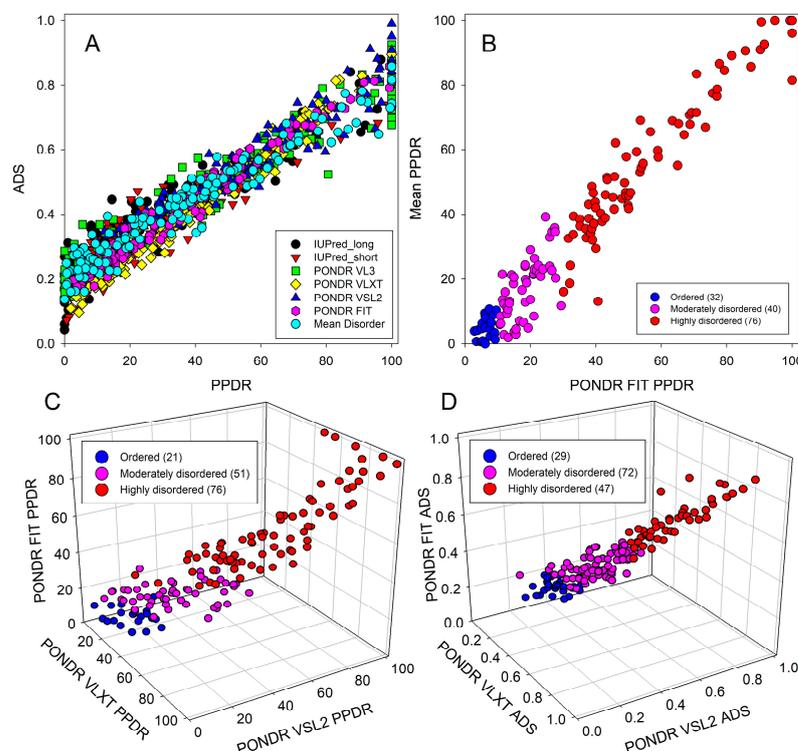


Figure 4. Two- and three-dimensional scatterplots comparing the average disorder scores and average PPDR values generated by three different predictors. (A). Average disorder score (ADS) vs. predicted percent disordered residues (PPDR) for the entire RIDAO suite of predictors and the RIDAO mean disorder score. (B). PPDR based on RIDAO's mean disorder score vs. PPDR based on PONDRA[®] FIT. (C). PPDR evaluated by PONDRA[®] VLXT, PONDRA[®] VSL2, and PONDRA[®] FIT along the x, y, and z axes, respectively. (D). ADS generated by PONDRA[®] VLXT, PONDRA[®] VSL2, and PONDRA[®] FIT along the x, y, and z axes, respectively. Proteins shown in blue, pink, or red are predicted as high ordered, moderately disordered, and highly disordered, respectively, based on their respective PONDRA[®] FIT PPDR values.

This analysis revealed that Q1 includes 71 DPPs, of which 21 proteins are predicted to be highly ordered, 40 proteins classified as moderately disordered, and 10 proteins classified as highly disordered by PONDRA[®] FIT. Q2 includes 30 DPPs (3 moderately and 27 highly disordered proteins according to the PONDRA[®] FIT analysis), whereas in Q3, one can find a total of 38 proteins, of which 2 and 36 are predicted as moderately and highly disordered (respectively) by PONDRA[®] FIT. Finally, Q4 contains 8 DPPs, of which 4 and 4 were classified as moderately and highly disordered (respectively) by PONDRA[®] FIT. Figure 5B further illustrates these observations in a 3D scatter plot, where in addition to the CH-CDF plane, the PONDRA[®] FIT-generated PPDR values are used as a third dimension. Overall, these analyses revealed that more than half of the proteins altered in the human cancer unfoldome are globally disordered.

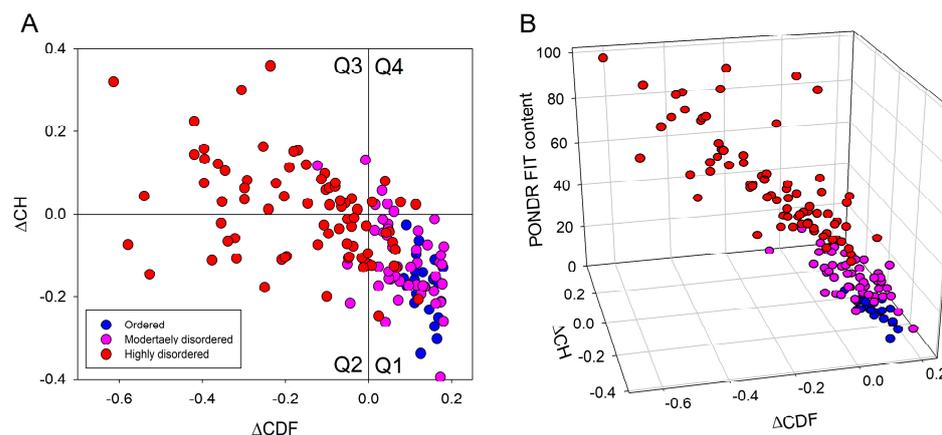


Figure 5. Evaluation of the overall disorder status of 148 human proteins differently expressed in cancer. **(A).** CH-CDF plot analysis of 166 human TPR proteins. Quadrants are numbered in a clockwise direction starting with Q1 in the lower right and ending with Q4 in the upper right. The following characterize proteins within each quadrant: Q1 (lower-right quadrant)—proteins predicted to be ordered by both predictors; Q2 (lower-left quadrant)—proteins predicted to be ordered by CH but disordered by CDF; Q3 (upper-left quadrant)—proteins predicted to be disordered by both predictors; and Q4 (upper-right quadrant)—proteins predicted to be disordered by CH and ordered by CDF. **(B).** Comparison of the CH-CDF plane with the output of the per-residue disorder predictor PONDR® FIT.

2.3. Biological Network and Pathway Analyses

These DPPs (140 protein groups) were classified using major Gene Ontology (GO) categories, and a pathway analysis was performed using an interacting genes database (DAVID) [82–84]. For insight into the potential function of the comorbid signature proteins, we performed a protein–protein interaction (PPI) analysis of the identified DPPs in each group comparison using GeneMANIA [91] and with previously established bioinformatics workflows from our lab [92].

First, we analyzed the inter-set interactivity of these DPPs using the Search Tool for the Retrieval of Interacting Genes (STRING (<http://string-db.org/>, accessed on 20 December 2023)) [93–95]. Figure 6 represents the results of this analysis for 146 DPPs (no STRING information was available for XP_005264440.1 and XP_011528471.1) and shows that most of these proteins are involved in the formation of a rather dense PPI network. In fact, only two of the proteins (TMA16 and TMEM263 with UniProt IDs: Q96EY4 and Q8WUH6, respectively) are not included in this network. As a result, this network contains 145 nodes (human proteins from the cancer unfoldome) connected by 1603 edges (PPIs). In this network, the average node degree is 22.1, and its average local clustering coefficient (which defines how close its neighbors are to being a complete clique; the local clustering coefficient is equal to 1 if every neighbor connected to a given node N_i is also connected to every other node within the neighborhood, and it is equal to 0 if no node that is connected to a given node N_i connects to any other node that is connected to N_i) is 0.422. Since the expected number of interactions among proteins in a comparable size set of proteins randomly selected from the human proteome is equal to 769, this internal PPI network has significantly more interactions than expected, being characterized by a PPI enrichment p -value $< 10^{-16}$. Therefore, these data indicate that the majority of human proteins differently expressed in cancer can interact with each other (although the confidence of this inter-family PPI network is low).

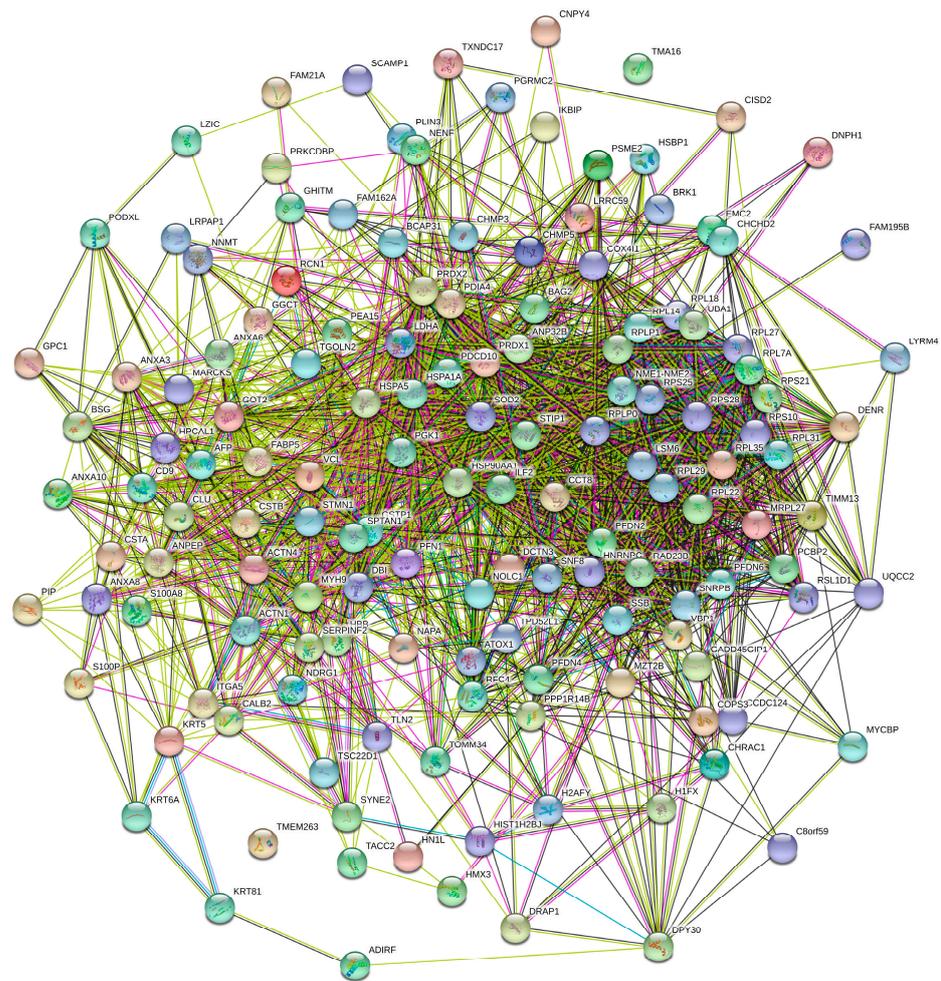


Figure 6. STRING-based analysis of the inter-set interactivity of 140 human proteins differently expressed in cancer using the low confidence level of 0.15. This confidence level was selected to ensure maximal inclusion of these proteins in the resulting PPI. STRING generates a network of predicted associations based on predicted and experimentally validated information on the interaction partners of a protein of interest [93]. In the corresponding network, the nodes correspond to proteins, whereas the edges show predicted/known functional associations. Seven types of evidence are used to build the corresponding network and are indicated by the differently colored lines: a green line represents neighborhood evidence; a red line—the presence of fusion evidence; a purple line—experimental evidence; a blue line—co-occurrence evidence; a light blue line—database evidence; a yellow line—text mining evidence; and a black line—co-expression evidence [93].

Next, we compared the interactivity of highly ordered (PPDR < 10%), moderately disordered (10% ≤ PPDR < 30%), and highly disordered (PPDR ≥ 30%) DPPs. Figure 7 shows the interaction networks for the 125 most-related proteins out of 140 annotated DPPs (excluding 8 unknown proteins), which are grouped together in clusters.

Figures 8 and 9 show the same comparison of interactivity for the DPPs related to cancer and diseases other than cancer (respectively) according to the DAVID functional annotation. The cancer PPI network consists of 21 (+12 singletons) proteins, where 13 (+9 singletons) proteins are moderately or highly disordered, and 7 (+3 singletons) proteins are ordered.

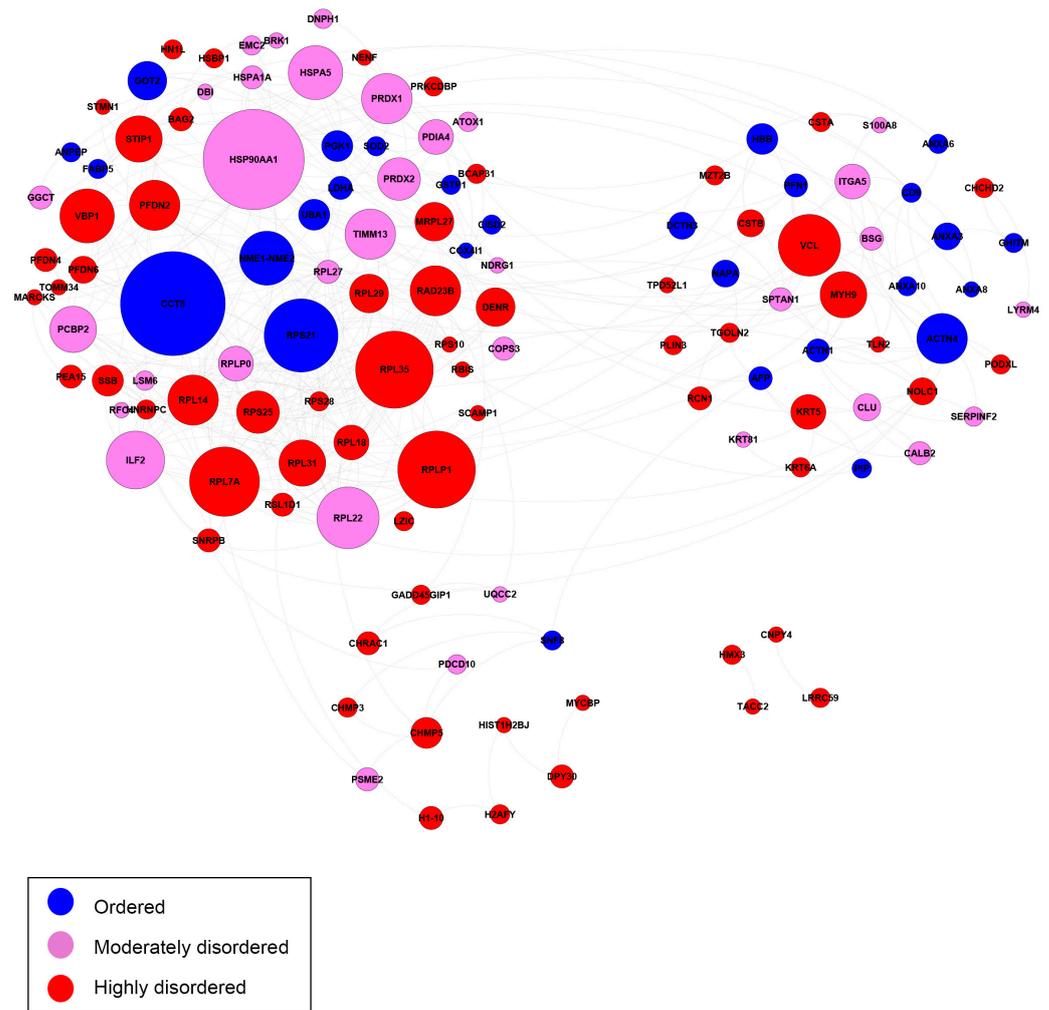


Figure 7. Biological interaction networks of the differentially present proteins (DPPs). In total, 125 most-related proteins (a subgroup of 148 DPPs excluding 8 unknown proteins) are shown and grouped together in clusters. Node size is proportional to the number of connections to other nodes. Nodes are colored according to the levels of intrinsic disorder (PPDR values): highly ordered, $PPDR < 10\%$; moderately disordered, $10\% \leq PPDR < 30\%$; and highly disordered, $PPDR \geq 30\%$ (see the explanation in Section 3).

The moderately disordered heat shock protein 90 alpha family class A (HSP90AA1) interacted with the most proteins in the network. HSP90AA1 has roles in tumor formation and activates oncogenic proteins to enhance tumor growth and invasiveness [96,97]. The highly disordered protein with the most PPIs was stress induced phosphoprotein 1 (STIP1). STIP1 overexpression is associated with high metastatic potential in diverse types of cancer [98–100]. Heat shock protein family A, member 1A (HSPA1A), was also highly connected. In colon cancer patients, elevated levels of HSPA1A are associated with a poor prognosis [101]. Among the other disordered proteins, Keratin 5 (KRT5) is a biomarker in basal-like breast cancers associated with a poor prognosis [102].

Several ordered proteins were highly connected. Heat shock 70 kDa protein 5 (HSPA5) has been shown to promote tumor growth by regulating cell proliferation in various cancers [103,104]. Phosphoglycerate kinase 1 (PGK1) has roles in angiogenesis and DNA repair. PGK1 is overexpressed in many cancers, and its overexpression typically correlates with poor prognoses in specific cancer types. However, PGK1 can act as a tumor suppressor by suppressing angiogenesis and tumor growth under certain conditions [105]. Different post-translational modifications of PGK1 are related to the positive or negative regulation of tumor metabolism and growth [106].

Superoxide dismutase 2 is an antioxidant that regulates oxidative stress in the mitochondria through the dismutation of superoxide radicals. The upregulation of SOD2 contributes to cancer cell proliferation, angiogenesis, and invasion [107,108] or tumor suppression [107]. The stability of the SOD2 protein observed in this study is consistent with findings indicating the SOD2 gene is rarely mutated in cancer. It acts as a context-dependent regulator of tumorigenesis or tumor suppression [107].

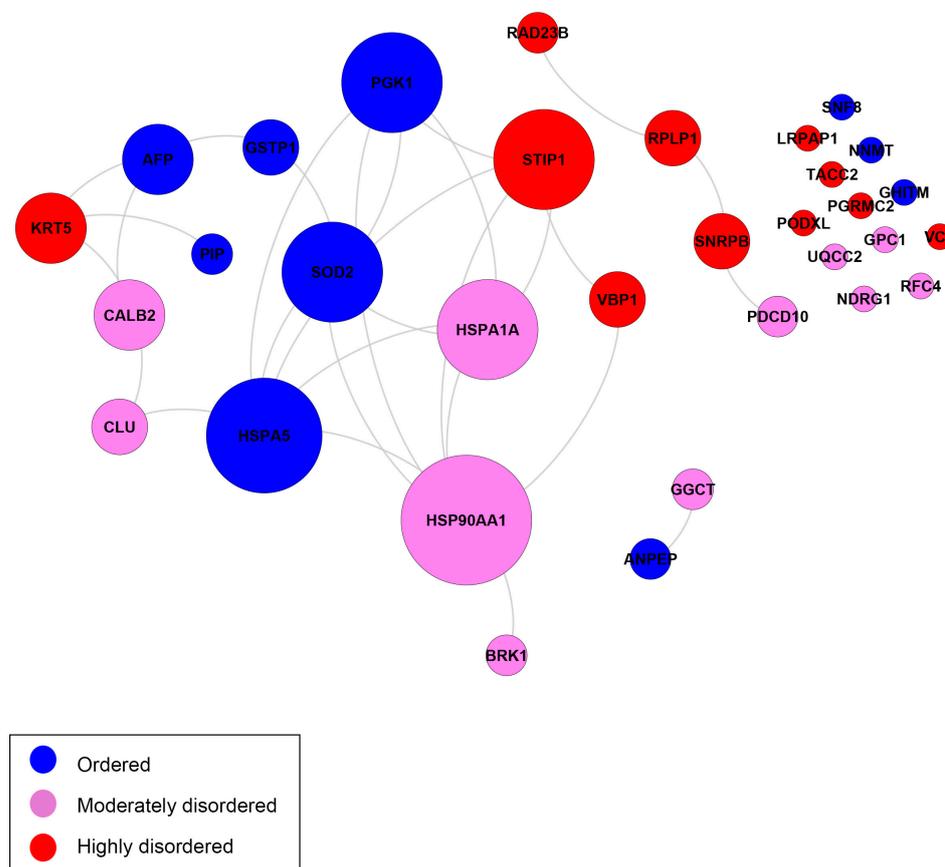


Figure 8. Biological interaction networks of the DPPs related to cancer (a subgroup of 148 DPPs) according to the DAVID functional annotation. In total, 20 most-related proteins out of 28 cancer related proteins are shown and grouped together in clusters. Node size is proportional to the number of connections to other nodes. Nodes are colored according to the levels of intrinsic disorder (PPDR values): highly ordered, $PPDR < 10\%$; moderately disordered, $10\% \leq PPDR < 30\%$; and highly disordered, $PPDR \geq 30\%$ (see the explanation in Section 3).

The network of annotations for other diseases consisted of 48 proteins. In total, 34 proteins were moderately or highly disordered, and 14 proteins were ordered. The highly disordered protein Vinculin (VCL) had the most PPIs. VCL is responsible for the progression of different cancers through tumor cell invasion and proliferation [109,110]. In prostate cancer patients, the down-regulation of the VCL promoter leads to the repression of tumor invasion, movement, and migration [110]. The human ribosomal large subunit protein L7a (RPL7A) contributes to the risk of breast cancer with adequate alcohol consumption. Specifically, the protein functions as an ethanol-response factor in breast cancer cells and promotes tumor growth and metastasis [111]. NME1-NME2 is capable of suppressing metastasis in several types of cancers [112,113]. Actinin alpha 4 (ACTN4) is extensively associated with cancer development. Both in vivo and in vitro studies have implicated this protein in tumor invasion and metastasis [114,115]. Depending on the cellular context, ACTN4 may also function as a tumor suppressor [116].

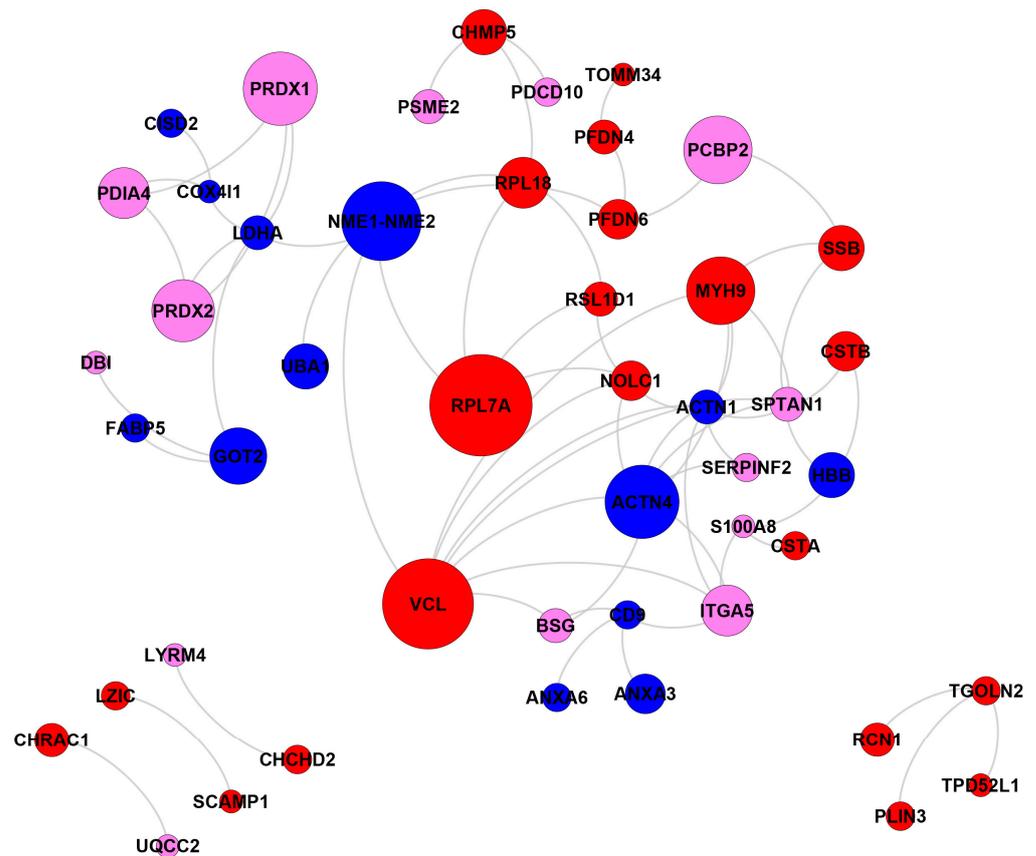


Figure 9. Biological interaction networks of the DPPs related to diseases other than cancer (a subgroup of 148 DPPs) according to DAVID functional annotation (GAD_DISEASE). In total, 49 most-related proteins out of 63 disease related proteins (other than cancer) are shown and grouped together in clusters. Node size is proportional to degree (the number of connections to other nodes). Nodes are colored according to the levels of intrinsic disorder (PPDR values): highly ordered, $PPDR < 10\%$; moderately disordered, $10\% \leq PPDR < 30\%$; and highly disordered, $PPDR \geq 30\%$ (see the explanation in Section 3).

Finally, we analyzed the pathway enrichment for the 140 annotated DPP using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and associated proteins database (<https://www.genome.jp/kegg/>, accessed on 20 December 2023) [117–121]. We observed significant enrichment of protein processing in the endoplasmic reticulum pathway (04141, $adj_p = 0.0014$). Cancer cells frequently undergo ER stress, triggering the unfolded protein response (UPR) to promote cell survival [122]. HSP90AA1, a component of this pathway, may serve as a target of interest due to its extensive interactions with ordered and disordered proteins. We observed the enrichment of numerous pathways corresponding to cell adhesion, junctions, and extracellular matrix function. Functional defects in these structures are typically associated with increased metastatic potential by the promotion of epithelial-mesenchymal transition [123,124]. Alterations in spliceosome pathway (03040, $adj_p = 0.0315$) components have been associated with aggressive forms of cancer, promoting cell invasion, angiogenesis, and drug resistance [125].

The antigen processing and presentation pathway (04612, $adj_p = 0.0087$) plays a significant role in both tumor evasion and the development of cancer immunotherapies. IDPs may have a role in the functions of antigen-presenting machinery in tumor cells, possibly interfering with the antitumor response. Proteasome activator subunit 2 (PSME2, PA28) is constitutively expressed in antigen-presenting cells and is involved in MHC class I antigen presentation. The down-regulation of this protein has been observed to impair the

antigen presentation of endogenous tumor antigens [126,127]. Disorder in PSME2 protein may also impair its activity and inhibit the antigen-presenting response.

3. Materials and Methods

3.1. Cell Culture

Human normal (MCF-10A) and breast cancer (BT-549) cell lines were cultured in Dulbecco's Modified Eagle's medium (DMEM) with 2 mmol/L L-glutamine and 10% fetal bovine serum (VWR). Upon confluence, cells were trypsinized, sedimented at $1000\times g$ for 6 min, and stored at $-80\text{ }^{\circ}\text{C}$.

3.2. Sample Preparation

Prior to the LC-MS analysis, cell pellets ($\sim 10^7$ cells per sample) were lysed on ice with 0.1% Nonidet P-40 in PBS in the presence of the Protease Inhibitor Cocktail (Sigma-Aldrich) for 30 min. Cell debris were removed by centrifugation at $13,000\times g$ for 30 min. Then, fully folded proteins were precipitated at $13,000\times g$ for 30 min after incubation with 1.5% trichloroacetic acid (TCA). The pellets were removed and discarded. Then, IDPs were precipitated at $13,000\times g$ for 60 min after incubation with 20% TCA. The IDP containing pellets (50 μg total protein per sample) were denatured in 8M urea and 50 mM Tris-HCl, pH 8.0, reduced with 10 mM TCEP for 60 min, alkylated with 2 mM iodoacetamide for 60 min in the dark, and then diluted to 2M urea with 50 mM Tris-HCl, pH 8.0, at RT. Two micrograms of Trypsin Gold (Promega) were added for overnight digestion (18 h, $37\text{ }^{\circ}\text{C}$), and then the tryptic peptides were immediately desalted using Pierce C18 spin columns (Thermo Fischer Scientific, Norcross, GA USA) at RT. Peptides were eluted with 70% acetonitrile and 0.1% formic acid (FA) and dried completely on a SpeedVac Concentrator.

3.3. Nano-LC-MS/MS Analysis

Peptides were resuspended in 5 μL of 0.5% FA and loaded onto a three-phase MudPIT column (150 $\mu\text{m} \times 2\text{ cm}$ C18 resin, 150 $\mu\text{m} \times 4\text{ cm}$ strong cation exchange SCX resin, filter union, and 150 $\mu\text{m} \times 20\text{ cm}$ C18 resin) as described previously [128]. A 10-step MudPIT (0 mM, 25 mM, 50mM, 100 mM, 150 mM, 200 mM, 300 mM, 500 mM, 750 mM, and 1000 mM ammonium acetate, each salt pulse followed by a 120-min acetonitrile gradient 5–50% B [Buffer A: 0.1% FA; Buffer B: 0.1% FA in acetonitrile]) was executed for the LC-MS analysis using an Eksigent™ AS-1 autosampler and Eksigent™ nano-LC Ultra 2D pump online with an Orbitrap LTQ XL Linear Ion Trap Mass Spectrometer (Thermo Finnigan, Norcross, GA USA) with a nanospray source. MS data acquisition was undertaken with a data-dependent six-event method (a survey FTMS scan [res. 30,000] followed by five data-dependent IT scans for the five consequent most abundant ions). The general mass spectrometric settings were as follows: spray voltage, 2.4 kV; no sheath and no auxiliary gas flow; ion transfer tube temperature, 200 C; CID fragmentation (for MS/MS), 35% normalized collision energy; activation $q = 0.25$; activation time, 30 ms. The minimal threshold for the dependent scans was set to 1000 counts, and a dynamic exclusion list was used with the following settings: repeat count of 1, repeat duration of 30 s, exclusion list size of 500, and exclusion duration of 90 s.

3.4. Proteins Identification and Quantification

Database searches were undertaken using PEAKS v9 software (Bioinformatics Solutions Inc., Waterloo, ON, Canada) against the forward and reverse human trypsin sequences (UniProtKB/Swiss-Prot) [85,129–131]. The parameters for the database search were as follows: full tryptic digestion; up to three missed cleavage sites; 20 ppm peptide mass tolerance; 0.5 Da fragment mass tolerance; cysteine carbamidomethylation (+57 Da) and methionine oxidation (+16 Da) as variable modifications. The relative label-free quantification (LFQ) of the identified proteins was performed with the Q module of the PEAKS software based on the extracted ion currents of the identified unique peptides' parent ions

or a spectral counting approach, and statistically significant changes have been confirmed with Fisher's exact test ($p \leq 0.005$; Benjamini-Hochberg FDR < 0.05).

3.5. Protein Annotation, Biological Network and Pathway Analyses

Annotation in DAVID (<https://david.ncifcrf.gov/>, accessed on 20 December 2023) [82–84] was conducted to identify proteins associated with cancer and other non-cancer diseases. Biological networks were constructed with the Cytoscape software platform (<https://cytoscape.org/>, accessed on 20 December 2023) [132] and visualized with GEPHI (<https://gephi.org/>, accessed on 20 December 2023) [133]. The pathway enrichment analysis was performed with WEBGESTALT web analysis software (<http://bioinfo.vanderbilt.edu/wg2/>, accessed on 20 December 2023) [134,135] by mapping proteins corresponding to KEGG pathways and conducting a hypergeometric test for significant enrichment. The significance for pathway level enrichment was defined as having an enrichment score False Discovery Rate (FDR) adjusted p -value ≤ 0.05 ($adj_p \leq 0.05$).

3.6. Compositional Profiling of the Members of Human Cancer-Related Unfoldome

A comparative analysis of the amino acid composition of the members of the human cancer-related unfoldome with respect to DisProt 3.4 (<http://www.disprot.org>, accessed on 20 December 2023) [88] and PDB Select 25 [87] reference protein sets was conducted using the Composition Profiler online service (<http://www.cprofiler.org>, accessed on 20 December 2023) [86]. The set DisProt 3.4 comprises consensus sequences of experimentally determined disordered regions, while PDB Select 25 contains PDB structures with less than 25% sequence identity and is biased towards the composition of proteins amenable to crystallization studies. Amino acids are arranged in the order of increase in their disorder propensity, according to the scale by Radivojac et al. [12].

3.7. Per-Residue Intrinsic Disorder Analysis

The per-residue intrinsic disorder predisposition analysis of proteins included in the unfoldome of human cancer and proteins included in the human cancer-related unfoldome was conducted using the RIDAO platform [89], which integrates a suite of six established disorder predictors: PONDR[®] VLXT [136], PONDR[®] VL3 [137], PONDR[®] VLS2B [138], PONDR[®] FIT [139], IUPred (short), and IUPred (long) [140,141]. RIDAO enables the rapid generation of disorder profile plots for individual polypeptides as well as arrays of polypeptides. We note that the use of multiple computational tools for prediction of intrinsic disorder in proteins is an accepted practice in the field. This is because different computational tools use different attributes (such as amino acid composition, hydrophathy, sequence complexity, etc.) and models to calculate a disorder predisposition score for every amino acid residue in a query protein. Therefore, it is not uncommon for different tools to generate rather different outputs. Furthermore, there is not an accepted consensus regarding which disorder predictor is the best for evaluating disorder predisposition of a query protein. In fact, since different computational tools are sensitive to different disorder-related aspects of amino acid sequences, all of them contain some useful information. The per-residue disorder propensities that these tools outputs are real numbers between 1 (ideal prediction of disorder) and 0 (ideal prediction of order). To identify the disordered residues and regions in query proteins, a threshold of ≥ 0.5 is used. For each query protein in this study, the predicted percentage of intrinsic disorder (PPID) was calculated based on the outputs of each predictor in the RIDAO suite, such that PPID in a query protein represents the percent of residues with disorder scores satisfying the aforementioned threshold.

3.8. Binary and CH-CDF Intrinsic Disorder Analysis

Next, RIDAO was used to perform a CH-CDF analysis [142]. A CH-CDF analysis takes into consideration the outputs of two binary predictors—i.e., the charge-hydrophathy (CH) plot [2,31] and the cumulative distribution function (CDF) plot [31]. RIDAO yields CH values based on the Kyte-Doolittle hydrophathy scale [143] and CDF values based on

PONDR[®] VLXT [136]. Conducting a CH-CDF analysis allows for the classification of proteins based on their position within the CH-CDF phase space as: (i) ordered (proteins predicted to be ordered by both binary predictors), (ii) putative native “molten globules” or hybrid proteins (proteins determined to be ordered/compact by the CH but disordered by the CDF), (iii) putative native coils and native pre-molten globules (proteins predicted to be disordered by both methods), and (iv) proteins predicted to be disordered by the CH-plot but ordered by the CDF.

3.9. Evaluation of the Interactability within the Human Cancer-Related Unfoldome

Information on the interactability within the human cancer-related unfoldome was retrieved using the Search Tool for the Retrieval of Interacting Genes; STRING (<http://string-db.org/>, accessed on 20 December 2023) [93–95]. STRING generates a network of protein-protein interactions based on predicted and experimentally validated information on the interaction partners of a protein of interest [93–95]. In the corresponding network, the nodes correspond to proteins, whereas the edges show predicted or known functional associations. Seven types of evidence are used to build the corresponding network, where they are indicated by differently colored lines: a green line represents neighborhood evidence; a red line—the presence of fusion evidence; a purple line—experimental evidence; a blue line—co-occurrence evidence; a light blue line—database evidence; a yellow line—text mining evidence; and a black line—co-expression evidence [93].

In this study, STRING was utilized to generate the inter-unfoldome network of protein-protein interactions (PPI), where the protein interaction network was obtained from STRING using a custom confidence level (minimum required interaction score) of 0.15. We also analyzed the individual PPI networks of the 10 most disordered proteins from the analyzed set using a custom value of 500 maximum first-shell interactions at medium, high, and highest confidence levels (minimum required interaction score of 0.4, 0.7, and 0.9 respectively). The resulting PPI networks were further analyzed using STRING-embedded routines in order to retrieve network-related statistics, such as the following: the number of nodes (proteins); the number of edges (interactions); average node degree (average number of interactions per protein); average local clustering coefficient (which defines how close the neighbors are to being a complete clique—if a local clustering coefficient is equal to 1, then every neighbor connected to a given node N_i is also connected to every other node within the neighborhood, and if it is equal to 0, then no node that is connected to a given node N_i connects to any other node that is connected to N_i); expected number of edges (which is the number of interactions among the proteins in a random set of proteins of similar size); and a PPI enrichment p -value (which is a reflection of the fact that query proteins in the analyzed PPI network have more interactions among themselves than what would be expected for a random set of proteins of similar size drawn from the genome; it has been pointed out that such an enrichment indicates that the proteins are at least partially biologically connected as a group).

4. Conclusions

The results reported in this study confirmed and further characterized the role of IDPs in the molecular mechanisms of breast cancer. The involvement of IDPs in several fundamental cellular processes underscores the need to capture their distinct contributions to cancer. Understanding the structural dynamics of these IDPs and the mechanics of their binding interactions with other proteins can provide new insights into the role of protein dysregulation in cancer development. As approaches for characterizing IDP structures and their interactors develop, the unique characteristics of specific cancer phenotypes can be elucidated.

Author Contributions: Conceptualization, V.P. and V.N.U.; methodology, V.P., V.N.U., S.M. and S.P.; software, I.N., G.W.D.II and V.N.U.; validation, V.P., V.N.U., A.C. and S.M.; formal analysis, V.P., V.N.U., A.C., S.M., I.N. and G.W.D.II; investigation, V.P., V.N.U., I.N. and G.W.D.II; resources, S.M. and S.P.; data curation, L.E.L.II and S.P.; writing, V.P. and V.N.U.; writing—review and editing,

V.P., V.N.U., A.C., S.M., I.N., G.W.D.II and S.P.; visualization, V.P., V.N.U., A.C., L.E.L.II and S.P.; supervision, V.P., V.N.U. and S.P.; project administration, V.P., V.N.U. and S.P.; funding acquisition, V.P. and S.P. All authors have read and agreed to the published version of the manuscript.

Funding: NIH grant U54MD007586 (NIMHD) to the Research Centers in Minority Institutions (RCMI) program at Meharry Medical College. NIH grant UL1TR002243 (NCATS) to the Vanderbilt Institute for Clinical Translational Research at Vanderbilt University Medical Center.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD047284 and 10.6019/PXD047284 (<https://www.proteomexchange.org>, accessed on 1 January 2024).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331. [[CrossRef](#)]
2. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415–427. [[CrossRef](#)] [[PubMed](#)]
3. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W.; et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [[CrossRef](#)]
4. Dunker, A.K.; Obradovic, Z. The protein trinity—linking function and disorder. *Nat. Biotechnol.* **2001**, *19*, 805–806. [[CrossRef](#)] [[PubMed](#)]
5. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533. [[CrossRef](#)]
6. Uversky, V.N. Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* **2002**, *11*, 739–756. [[CrossRef](#)]
7. Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **2002**, *323*, 573–584. [[CrossRef](#)]
8. Dunker, A.K.; Brown, C.J.; Lawson, J.D.; Iakoucheva, L.M.; Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **2002**, *41*, 6573–6582. [[CrossRef](#)]
9. Dunker, A.K.; Brown, C.J.; Obradovic, Z. Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* **2002**, *62*, 25–49.
10. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208. [[CrossRef](#)]
11. Tompa, P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* **2005**, *579*, 3346–3354. [[CrossRef](#)] [[PubMed](#)]
12. Radivojac, P.; Iakoucheva, L.M.; Oldfield, C.J.; Obradovic, Z.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder and functional proteomics. *Biophys. J.* **2007**, *92*, 1439–1456. [[CrossRef](#)] [[PubMed](#)]
13. Dunker, A.K.; Oldfield, C.J.; Meng, J.; Romero, P.; Yang, J.Y.; Chen, J.W.; Vacic, V.; Obradovic, Z.; Uversky, V.N. The unfoldomics decade: An update on intrinsically disordered proteins. *BMC Genom.* **2008**, *9* (Suppl. 2), S1–S26. [[CrossRef](#)] [[PubMed](#)]
14. Galea, C.A.; Wang, Y.; Sivakolundu, S.G.; Kriwacki, R.W. Regulation of cell division by intrinsically unstructured proteins: Intrinsic flexibility, modularity, and signaling conduits. *Biochemistry* **2008**, *47*, 7598–7609. [[CrossRef](#)] [[PubMed](#)]
15. Tompa, P. Unstructural biology coming of age. *Curr. Opin. Struct. Biol.* **2011**, *21*, 419–425. [[CrossRef](#)]
16. Hong, S.; Choi, S.; Kim, R.; Koh, J. Mechanisms of Macromolecular Interactions Mediated by Protein Intrinsic Disorder. *Mol. Cells* **2020**, *43*, 899–908. [[CrossRef](#)]
17. Pajkos, M.; Dosztanyi, Z. Functions of intrinsically disordered proteins through evolutionary lenses. *Prog. Mol. Biol. Transl. Sci.* **2021**, *183*, 45–74. [[CrossRef](#)]
18. Malagrino, F.; Pennacchietti, V.; Santorelli, D.; Pagano, L.; Nardella, C.; Diop, A.; Toto, A.; Gianni, S. On the Effects of Disordered Tails, Supertertiary Structure and Quinary Interactions on the Folding and Function of Protein Domains. *Biomolecules* **2022**, *12*, 209. [[CrossRef](#)]
19. Cermakova, K.; Hodges, H.C. Interaction modules that impart specificity to disordered protein. *Trends Biochem. Sci.* **2023**, *48*, 477–490. [[CrossRef](#)]
20. Davey, N.E.; Simonetti, L.; Ivarsson, Y. The next wave of interactomics: Mapping the SLiM-based interactions of the intrinsically disordered proteome. *Curr. Opin. Struct. Biol.* **2023**, *80*, 102593. [[CrossRef](#)]
21. Evans, R.; Ramisetty, S.; Kulkarni, P.; Weninger, K. Illuminating Intrinsically Disordered Proteins with Integrative Structural Biology. *Biomolecules* **2023**, *13*, 124. [[CrossRef](#)]

22. Fenton, M.; Gregory, E.; Daughdrill, G. Protein disorder and autoinhibition: The role of multivalency and effective concentration. *Curr. Opin. Struct. Biol.* **2023**, *83*, 102705. [[CrossRef](#)] [[PubMed](#)]
23. Handa, T.; Kundu, D.; Dubey, V.K. Perspectives on evolutionary and functional importance of intrinsically disordered proteins. *Int. J. Biol. Macromol.* **2023**, *224*, 243–255. [[CrossRef](#)] [[PubMed](#)]
24. Holehouse, A.S.; Kragelund, B.B. The molecular basis for cellular function of intrinsically disordered protein regions. *Nat. Rev. Mol. Cell. Biol.* **2023**. [[CrossRef](#)]
25. Moses, D.; Ginell, G.M.; Holehouse, A.S.; Sukenik, S. Intrinsically disordered regions are poised to act as sensors of cellular chemistry. *Trends Biochem. Sci.* **2023**, *48*, 1019–1034. [[CrossRef](#)] [[PubMed](#)]
26. Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **2014**, *83*, 553–584. [[CrossRef](#)] [[PubMed](#)]
27. Deiana, A.; Forcelloni, S.; Porrello, A.; Giansanti, A. Intrinsically disordered proteins and structured proteins with intrinsically disordered regions have different functional roles in the cell. *PLoS ONE* **2019**, *14*, e0217889. [[CrossRef](#)] [[PubMed](#)]
28. Uversky, V.N. Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta* **2013**, *1834*, 932–951. [[CrossRef](#)]
29. Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* **2000**, *11*, 161–171.
30. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004**, *337*, 635–645. [[CrossRef](#)]
31. Oldfield, C.J.; Cheng, Y.; Cortese, M.S.; Brown, C.J.; Uversky, V.N.; Dunker, A.K. Comparing and combining predictors of mostly disordered proteins. *Biochemistry* **2005**, *44*, 1989–2000. [[CrossRef](#)]
32. Tompa, P.; Dosztanyi, Z.; Simon, I. Prevalent structural disorder in E. coli and S. cerevisiae proteomes. *J. Proteome Res.* **2006**, *5*, 1996–2000. [[CrossRef](#)]
33. Lu, Y.; Yang, Y.; Zhu, G.; Zeng, H.; Fan, Y.; Guo, F.; Xu, D.; Wang, B.; Chen, D.; Ge, G. Emerging Pharmacotherapeutic Strategies to Overcome Undruggable Proteins in Cancer. *Int. J. Biol. Sci.* **2023**, *19*, 3360–3382. [[CrossRef](#)]
34. Santofimia-Castano, P.; Rizzuti, B.; Xia, Y.; Abian, O.; Peng, L.; Velazquez-Campoy, A.; Neira, J.L.; Iovanna, J. Targeting intrinsically disordered proteins involved in cancer. *Cell. Mol. Life Sci.* **2020**, *77*, 1695–1707. [[CrossRef](#)]
35. Kulkarni, V.; Kulkarni, P. Intrinsically disordered proteins and phenotypic switching: Implications in cancer. *Prog. Mol. Biol. Transl. Sci.* **2019**, *166*, 63–84. [[CrossRef](#)]
36. Garg, N.; Kumar, P.; Gadhave, K.; Giri, R. The dark proteome of cancer: Intrinsic disorder and functionality of HIF-1alpha along with its interacting proteins. *Prog. Mol. Biol. Transl. Sci.* **2019**, *166*, 371–403. [[CrossRef](#)]
37. Nussinov, R.; Jang, H.; Tsai, C.J.; Liao, T.J.; Li, S.; Fushman, D.; Zhang, J. Intrinsic protein disorder in oncogenic KRAS signaling. *Cell. Mol. Life Sci.* **2017**, *74*, 3245–3261. [[CrossRef](#)]
38. Kumar, D.; Sharma, N.; Giri, R. Therapeutic Interventions of Cancers Using Intrinsically Disordered Proteins as Drug Targets: c-Myc as Model System. *Cancer Inform.* **2017**, *16*, 1176935117699408. [[CrossRef](#)] [[PubMed](#)]
39. Mooney, S.M.; Jolly, M.K.; Levine, H.; Kulkarni, P. Phenotypic plasticity in prostate cancer: Role of intrinsically disordered proteins. *Asian J. Androl.* **2016**, *18*, 704–710. [[CrossRef](#)] [[PubMed](#)]
40. Russo, A.; Manna, S.L.; Novellino, E.; Malfitano, A.M.; Marasco, D. Molecular signaling involving intrinsically disordered proteins in prostate cancer. *Asian J. Androl.* **2016**, *18*, 673–681. [[CrossRef](#)] [[PubMed](#)]
41. Uversky, V.N. Intrinsic disorder in proteins associated with neurodegenerative diseases. *Front. Biosci. (Landmark Ed.)* **2009**, *14*, 5188–5238. [[CrossRef](#)] [[PubMed](#)]
42. Birol, M.; Melo, A.M. Untangling the Conformational Polymorphism of Disordered Proteins Associated With Neurodegeneration at the Single-Molecule Level. *Front. Mol. Neurosci.* **2019**, *12*, 309. [[CrossRef](#)] [[PubMed](#)]
43. Ayyadevara, S.; Ganne, A.; Balasubramaniam, M.; Shmookler Reis, R.J. Intrinsically disordered proteins identified in the aggregate proteome serve as biomarkers of neurodegeneration. *Metab. Brain Dis.* **2022**, *37*, 147–152. [[CrossRef](#)]
44. Di Liegro, C.M.; Schiera, G.; Schiro, G.; Di Liegro, I. RNA-Binding Proteins as Epigenetic Regulators of Brain Functions and Their Involvement in Neurodegeneration. *Int. J. Mol. Sci.* **2022**, *23*, 14622. [[CrossRef](#)] [[PubMed](#)]
45. Chakraborty, P.; Zweckstetter, M. Role of aberrant phase separation in pathological protein aggregation. *Curr. Opin. Struct. Biol.* **2023**, *82*, 102678. [[CrossRef](#)]
46. Tsoi, P.S.; Quan, M.D.; Ferreon, J.C.; Ferreon, A.C.M. Aggregation of Disordered Proteins Associated with Neurodegeneration. *Int. J. Mol. Sci.* **2023**, *24*, 3380. [[CrossRef](#)]
47. Cieplak, M.; Mioduszewski, L.; Chwastyk, M. Contact-Based Analysis of Aggregation of Intrinsically Disordered Proteins. *Methods Mol. Biol.* **2022**, *2340*, 105–120. [[CrossRef](#)]
48. Mei, Y.; Glover, K.; Su, M.; Sinha, S.C. Conformational flexibility of BECN1: Essential to its key role in autophagy and beyond. *Protein Sci.* **2016**, *25*, 1767–1785. [[CrossRef](#)]
49. Du, Z.; Uversky, V.N. A Comprehensive Survey of the Roles of Highly Disordered Proteins in Type 2 Diabetes. *Int. J. Mol. Sci.* **2017**, *18*, 2010. [[CrossRef](#)]
50. Nguyen, P.H.; Ramamoorthy, A.; Sahoo, B.R.; Zheng, J.; Faller, P.; Straub, J.E.; Dominguez, L.; Shea, J.E.; Dokholyan, N.V.; De Simone, A.; et al. Amyloid Oligomers: A Joint Experimental/Computational Perspective on Alzheimer's Disease, Parkinson's Disease, Type II Diabetes, and Amyotrophic Lateral Sclerosis. *Chem. Rev.* **2021**, *121*, 2545–2647. [[CrossRef](#)]

51. Usher, E.T.; Showalter, S.A. Biophysical insights into glucose-dependent transcriptional regulation by PDX1. *J. Biol. Chem.* **2022**, *298*, 102623. [[CrossRef](#)]
52. Tovo-Rodrigues, L.; Recamonde-Mendoza, M.; Paixao-Cortes, V.R.; Bruxel, E.M.; Schuch, J.B.; Friedrich, D.C.; Rohde, L.A.; Hutz, M.H. The role of protein intrinsic disorder in major psychiatric disorders. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **2016**, *171*, 848–860. [[CrossRef](#)]
53. Cheng, Y.; LeGall, T.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N. Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry* **2006**, *45*, 10448–10460. [[CrossRef](#)]
54. Vacic, V.; Iakoucheva, L.M. Disease mutations in disordered regions—exception to the rule? *Mol. Biosyst.* **2012**, *8*, 27–32. [[CrossRef](#)] [[PubMed](#)]
55. Tarnovskaya, S.; Kiselev, A.; Kostareva, A.; Frishman, D. Structural consequences of mutations associated with idiopathic restrictive cardiomyopathy. *Amino Acids* **2017**, *49*, 1815–1829. [[CrossRef](#)] [[PubMed](#)]
56. Johnston, J.R.; Landim-Vieira, M.; Marques, M.A.; de Oliveira, G.A.P.; Gonzalez-Martinez, D.; Moraes, A.H.; He, H.; Iqbal, A.; Wilnai, Y.; Birk, E.; et al. The intrinsically disordered C terminus of troponin T binds to troponin C to modulate myocardial force generation. *J. Biol. Chem.* **2019**, *294*, 20054–20069. [[CrossRef](#)]
57. Theophall, G.G.; Ramirez, L.M.S.; Premo, A.; Reverdatto, S.; Manigrasso, M.B.; Yepuri, G.; Burz, D.S.; Ramasamy, R.; Schmidt, A.M.; Shekhtman, A. Disruption of the productive encounter complex results in dysregulation of DIAPH1 activity. *J. Biol. Chem.* **2023**, *299*, 105342. [[CrossRef](#)] [[PubMed](#)]
58. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the D² concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246. [[CrossRef](#)]
59. Goh, K.I.; Cusick, M.E.; Valle, D.; Childs, B.; Vidal, M.; Barabasi, A.L. The human disease network. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8685–8690. [[CrossRef](#)]
60. Midic, U.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Protein disorder in the human diseaseome: Unfoldomics of human genetic diseases. *BMC Genom.* **2009**, *10* (Suppl. 1), S12. [[CrossRef](#)]
61. Williams, R.M.; Obradovi, Z.; Mathura, V.; Braun, W.; Garner, E.C.; Young, J.; Takayama, S.; Brown, C.J.; Dunker, A.K. The Protein Non-Folding Problem: Amino Acid Determinants of Intrinsic Order and Disorder. In *Biocomputing 2001—Proceedings of the Pacific Symposium*; World Scientific Publishing: Singapore, 2001; pp. 89–100.
62. Timm, D.E.; Vissavajhala, P.; Ross, A.H.; Neet, K.E. Spectroscopic and chemical studies of the interaction between nerve growth factor (NGF) and the extracellular domain of the low affinity NGF receptor. *Protein Sci.* **1992**, *1*, 1023–1031. [[CrossRef](#)]
63. Kim, T.D.; Ryu, H.J.; Cho, H.I.; Yang, C.H.; Kim, J. Thermal behavior of proteins: Heat-resistant proteins and their heat-induced secondary structural changes. *Biochemistry* **2000**, *39*, 14839–14846. [[CrossRef](#)]
64. Uversky, V.N. Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J.* **2009**, *28*, 305–325. [[CrossRef](#)]
65. Konno, T.; Tanaka, N.; Kataoka, M.; Takano, E.; Maki, M. A circular dichroism study of preferential hydration and alcohol effects on a denatured protein, pig calpastatin domain I. *Biochim. Biophys. Acta* **1997**, *1342*, 73–82. [[CrossRef](#)] [[PubMed](#)]
66. Lynn, A.; Chandra, S.; Malhotra, P.; Chauhan, V.S. Heme binding and polymerization by Plasmodium falciparum histidine rich protein II: Influence of pH on activity and conformation. *FEBS Lett.* **1999**, *459*, 267–271. [[CrossRef](#)]
67. Johansson, J.; Gudmundsson, G.H.; Rottenberg, M.E.; Berndt, K.D.; Agerberth, B. Conformation-dependent antibacterial activity of the naturally occurring human peptide LL-37. *J. Biol. Chem.* **1998**, *273*, 3718–3724. [[CrossRef](#)] [[PubMed](#)]
68. Cortese, M.S.; Baird, J.P.; Uversky, V.N.; Dunker, A.K. Uncovering the unfoldome: Enriching cell extracts for unstructured proteins by Acid treatment. *J. Proteome Res.* **2005**, *4*, 1610–1618. [[CrossRef](#)]
69. Csizmok, V.; Szollosi, E.; Friedrich, P.; Tompa, P. A novel two-dimensional electrophoresis technique for the identification of intrinsically unstructured proteins. *Mol. Cell Proteom.* **2006**, *5*, 265–273. [[CrossRef](#)]
70. Galea, C.A.; Pagala, V.R.; Obenauer, J.C.; Park, C.G.; Slaughter, C.A.; Kriwacki, R.W. Proteomic studies of the intrinsically unstructured mammalian proteome. *J. Proteome Res.* **2006**, *5*, 2839–2848. [[CrossRef](#)]
71. Kaddour, H.; Lyu, Y.; Welch, J.L.; Paromov, V.; Mandape, S.N.; Sakhare, S.S.; Pandhare, J.; Stapleton, J.T.; Pratap, S.; Dash, C.; et al. Proteomics Profiling of Autologous Blood and Semen Exosomes from HIV-infected and Uninfected Individuals Reveals Compositional and Functional Variabilities. *Mol. Cell. Proteom.* **2020**, *19*, 78–100. [[CrossRef](#)]
72. Kaddour, H.; Kopcho, S.; Lyu, Y.; Shouman, N.; Paromov, V.; Pratap, S.; Dash, C.; Kim, E.Y.; Martinson, J.; McKay, H.; et al. HIV-infection and cocaine use regulate semen extracellular vesicles proteome and miRNAome in a manner that mediates strategic monocyte haptotaxis governed by miR-128 network. *Cell. Mol. Life Sci.* **2021**, *79*, 5. [[CrossRef](#)] [[PubMed](#)]
73. Schirmer, E.C.; Yates, J.R., 3rd; Gerace, L. MudPIT: A powerful proteomics tool for discovery. *Discov. Med.* **2003**, *3*, 38–39. [[PubMed](#)]
74. Cagney, G.; Park, S.; Chung, C.; Tong, B.; O'Dushlaine, C.; Shields, D.C.; Emili, A. Human tissue profiling with multidimensional protein identification technology. *J. Proteome Res.* **2005**, *4*, 1757–1767. [[CrossRef](#)] [[PubMed](#)]
75. Kislinger, T.; Gramolini, A.O.; MacLennan, D.H.; Emili, A. Multidimensional protein identification technology (MudPIT): Technical overview of a profiling method optimized for the comprehensive proteomic investigation of normal and diseased heart tissue. *J. Am. Soc. Mass. Spectrom.* **2005**, *16*, 1207–1220. [[CrossRef](#)] [[PubMed](#)]
76. Delahunty, C.M.; Yates, J.R., 3rd. MudPIT: Multidimensional protein identification technology. *Biotechniques* **2007**, *43*, 563, 565, 567 passim. [[PubMed](#)]

77. Link, A.J.; Washburn, M.P. Analysis of protein composition using multidimensional chromatography and mass spectrometry. *Curr. Protoc. Protein Sci.* **2014**, *78*, 23.1.1–23.1.25. [[CrossRef](#)]
78. Hoek, K.L.; Samir, P.; Howard, L.M.; Niu, X.; Prasad, N.; Galassie, A.; Liu, Q.; Allos, T.M.; Floyd, K.A.; Guo, Y.; et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. *PLoS ONE* **2015**, *10*, e0118528. [[CrossRef](#)] [[PubMed](#)]
79. Samir, P.; Rahul; Slaughter, J.C.; Link, A.J. Environmental Interactions and Epistasis Are Revealed in the Proteomic Responses to Complex Stimuli. *PLoS ONE* **2015**, *10*, e0134099. [[CrossRef](#)]
80. Gerbasi, V.R.; Browne, C.M.; Samir, P.; Shen, B.; Sun, M.; Hazelbaker, D.Z.; Galassie, A.C.; Frank, J.; Link, A.J. Critical Role for *Saccharomyces cerevisiae* Asc1p in Translational Initiation at Elevated Temperatures. *Proteomics* **2018**, *18*, e1800208. [[CrossRef](#)]
81. Samir, P.; Browne, C.M.; Rahul; Sun, M.; Shen, B.; Li, W.; Frank, J.; Link, A.J. Identification of Changing Ribosome Protein Compositions using Mass Spectrometry. *Proteomics* **2018**, *18*, e1800217. [[CrossRef](#)]
82. Sherman, B.T.; Hao, M.; Qiu, J.; Jiao, X.; Baseler, M.W.; Lane, H.C.; Imamichi, T.; Chang, W. DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **2022**, *50*, W216–W221. [[CrossRef](#)]
83. Huang da, W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [[CrossRef](#)]
84. Huang, D.W.; Sherman, B.T.; Tan, Q.; Kir, J.; Liu, D.; Bryant, D.; Guo, Y.; Stephens, R.; Baseler, M.W.; Lane, H.C.; et al. DAVID Bioinformatics Resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* **2007**, *35*, W169–W175. [[CrossRef](#)]
85. Xin, L.; Qiao, R.; Chen, X.; Tran, H.; Pan, S.; Rabinoviz, S.; Bian, H.; He, X.; Morse, B.; Shan, B.; et al. A streamlined platform for analyzing tera-scale DDA and DIA mass spectrometry data enables highly sensitive immunopeptidomics. *Nat. Commun.* **2022**, *13*, 3108. [[CrossRef](#)]
86. Vacic, V.; Uversky, V.; Dunker, A.K.; Lonardi, S. Composition Profiler: A tool for discovery and visualization of amino acid composition differences. *BMC Bioinform.* **2007**, *8*, 211. [[CrossRef](#)]
87. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)]
88. Sickmeier, M.; Hamilton, J.A.; LeGall, T.; Vacic, V.; Cortese, M.S.; Tantos, A.; Szabo, B.; Tompa, P.; Chen, J.; Uversky, V.N.; et al. DisProt: The Database of Disordered Proteins. *Nucleic Acids Res.* **2007**, *35*, D786–D793. [[CrossRef](#)] [[PubMed](#)]
89. Dayhoff, G.W., 2nd; Uversky, V.N. Rapid prediction and analysis of protein intrinsic disorder. *Protein Sci.* **2022**, *31*, e4496. [[CrossRef](#)] [[PubMed](#)]
90. Rajagopalan, K.; Mooney, S.M.; Parekh, N.; Getzenberg, R.H.; Kulkarni, P. A majority of the cancer/testis antigens are intrinsically disordered proteins. *J. Cell Biochem.* **2011**, *112*, 3256–3267. [[CrossRef](#)]
91. Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **2010**, *38*, W214–W220. [[CrossRef](#)] [[PubMed](#)]
92. Olokpa, E.; Mandape, S.N.; Pratap, S.; Stewart, M.V. Metformin regulates multiple signaling pathways within castration-resistant human prostate cancer cells. *BMC Cancer* **2022**, *22*, 1025. [[CrossRef](#)]
93. Szklarczyk, D.; Franceschini, A.; Kuhn, M.; Simonovic, M.; Roth, A.; Minguez, P.; Doerks, T.; Stark, M.; Muller, J.; Bork, P.; et al. The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* **2011**, *39*, D561–D568. [[CrossRef](#)] [[PubMed](#)]
94. Szklarczyk, D.; Kirsch, R.; Koutrouli, M.; Nastou, K.; Mehryary, F.; Hachilif, R.; Gable, A.L.; Fang, T.; Doncheva, N.T.; Pyysalo, S.; et al. The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **2023**, *51*, D638–D646. [[CrossRef](#)] [[PubMed](#)]
95. Snel, B.; Lehmann, G.; Bork, P.; Huynen, M.A. STRING: A web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **2000**, *28*, 3442–3444. [[CrossRef](#)]
96. Liu, H.; Zhang, Z.; Huang, Y.; Wei, W.; Ning, S.; Li, J.; Liang, X.; Liu, K.; Zhang, L. Plasma HSP90AA1 Predicts the Risk of Breast Cancer Onset and Distant Metastasis. *Front. Cell Dev. Biol.* **2021**, *9*, 639596. [[CrossRef](#)] [[PubMed](#)]
97. Tian, Y.; Wang, C.; Chen, S.; Liu, J.; Fu, Y.; Luo, Y. Extracellular Hsp90alpha and clusterin synergistically promote breast cancer epithelial-to-mesenchymal transition and metastasis via LRP1. *J. Cell Sci.* **2019**, *132*, jcs.228213. [[CrossRef](#)]
98. Zhang, S.; Shao, J.; Su, F. Prognostic significance of STIP1 expression in human cancer: A meta-analysis. *Clin. Chim. Acta* **2018**, *486*, 168–176. [[CrossRef](#)]
99. Zhang, Z.; Ren, H.; Yang, L.; Zhang, X.; Liang, W.; Wu, H.; Huang, L.; Kang, J.; Xu, J.; Zhai, E.; et al. Aberrant expression of stress-induced phosphoprotein 1 in colorectal cancer and its clinicopathologic significance. *Hum. Pathol.* **2018**, *79*, 135–143. [[CrossRef](#)]
100. Luo, X.; Liu, Y.; Ma, S.; Liu, L.; Xie, R.; Li, M.; Shen, P.; Wang, S. STIP1 is over-expressed in hepatocellular carcinoma and promotes the growth and migration of cancer cells. *Gene* **2018**, *662*, 110–117. [[CrossRef](#)]
101. Guan, Y.; Zhu, X.; Liang, J.; Wei, M.; Huang, S.; Pan, X. Upregulation of HSPA1A/HSPA1B/HSPA7 and Downregulation of HSPA9 Were Related to Poor Survival in Colon Cancer. *Front. Oncol.* **2021**, *11*, 749673. [[CrossRef](#)]

102. Cheang, M.C.; Voduc, D.; Bajdik, C.; Leung, S.; McKinney, S.; Chia, S.K.; Perou, C.M.; Nielsen, T.O. Basal-like breast cancer defined by five biomarkers has superior prognostic value than triple-negative phenotype. *Clin. Cancer Res.* **2008**, *14*, 1368–1376. [[CrossRef](#)] [[PubMed](#)]
103. Wang, Q.; Ke, S.; Liu, Z.; Shao, H.; He, M.; Guo, J. HSPA5 Promotes the Proliferation, Metastasis and Regulates Ferroptosis of Bladder Cancer. *Int. J. Mol. Sci.* **2023**, *24*, 5144. [[CrossRef](#)] [[PubMed](#)]
104. Cerezo, M.; Rocchi, S. New anti-cancer molecules targeting HSPA5/BIP to induce endoplasmic reticulum stress, autophagy and apoptosis. *Autophagy* **2017**, *13*, 216–217. [[CrossRef](#)] [[PubMed](#)]
105. O'Reilly, M.S.; Holmgren, L.; Shing, Y.; Chen, C.; Rosenthal, R.A.; Moses, M.; Lane, W.S.; Cao, Y.; Sage, E.H.; Folkman, J. Angiostatin: A novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* **1994**, *79*, 315–328. [[CrossRef](#)] [[PubMed](#)]
106. He, Y.; Luo, Y.; Zhang, D.; Wang, X.; Zhang, P.; Li, H.; Ejaz, S.; Liang, S. PGK1-mediated cancer progression and drug resistance. *Am. J. Cancer Res.* **2019**, *9*, 2280–2302. [[PubMed](#)]
107. Kim, Y.S.; Gupta Vallur, P.; Phaeton, R.; Mythreye, K.; Hempel, N. Insights into the Dichotomous Regulation of SOD2 in Cancer. *Antioxidants* **2017**, *6*, 86. [[CrossRef](#)] [[PubMed](#)]
108. Fukai, T.; Ushio-Fukai, M. Superoxide dismutases: Role in redox signaling, vascular function, and diseases. *Antioxid. Redox Signal* **2011**, *15*, 1583–1606. [[CrossRef](#)]
109. Zheng, X.; Xu, H.; Gong, L.; Cao, D.; Jin, T.; Wang, Y.; Pi, J.; Yang, Y.; Yi, X.; Liao, D.; et al. Vinculin orchestrates prostate cancer progression by regulating tumor cell invasion, migration, and proliferation. *Prostate* **2021**, *81*, 347–356. [[CrossRef](#)]
110. Li, H.; Wang, C.; Lan, L.; Behrens, A.; Tomaschko, M.; Ruiz, J.; Su, Q.; Zhao, G.; Yuan, C.; Xiao, X.; et al. High expression of vinculin predicts poor prognosis and distant metastasis and associates with influencing tumor-associated NK cell infiltration and epithelial-mesenchymal transition in gastric cancer. *Aging* **2021**, *13*, 5197–5225. [[CrossRef](#)]
111. Zhu, Y.; Lin, H.; Li, Z.; Wang, M.; Luo, J. Modulation of expression of ribosomal protein L7a (rpL7a) by ethanol in human breast cancer cells. *Breast Cancer Res. Treat.* **2001**, *69*, 29–38. [[CrossRef](#)] [[PubMed](#)]
112. Pamidimukkala, N.; Puts, G.S.; Kathryn Leonard, M.; Snyder, D.; Dabernat, S.; De Fabo, E.C.; Noonan, F.P.; Slominski, A.; Merlino, G.; Kaetzl, D.M. Nme1 and Nme2 genes exert metastasis-suppressor activities in a genetically engineered mouse model of UV-induced melanoma. *Br. J. Cancer* **2021**, *124*, 161–165. [[CrossRef](#)] [[PubMed](#)]
113. Khan, I.; Gril, B.; Steeg, P.S. Metastasis Suppressors NME1 and NME2 Promote Dynamin 2 Oligomerization and Regulate Tumor Cell Endocytosis, Motility, and Metastasis. *Cancer Res.* **2019**, *79*, 4689–4702. [[CrossRef](#)] [[PubMed](#)]
114. Fang, C.; Li, J.J.; Deng, T.; Li, B.H.; Geng, P.L.; Zeng, X.T. Actinin-4 as a Diagnostic Biomarker in Serum of Breast Cancer Patients. *Med. Sci. Monit.* **2019**, *25*, 3298–3302. [[CrossRef](#)] [[PubMed](#)]
115. Hsu, K.S.; Kao, H.Y. Alpha-actinin 4 and tumorigenesis of breast cancer. *Vitam. Horm.* **2013**, *93*, 323–351. [[CrossRef](#)]
116. Nikolopoulos, S.N.; Spengler, B.A.; Kisselbach, K.; Evans, A.E.; Biedler, J.L.; Ross, R.A. The human non-muscle alpha-actinin protein encoded by the ACTN4 gene suppresses tumorigenicity of human neuroblastoma cells. *Oncogene* **2000**, *19*, 380–386. [[CrossRef](#)]
117. Kanehisa, M.; Furumichi, M.; Sato, Y.; Kawashima, M.; Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **2023**, *51*, D587–D592. [[CrossRef](#)]
118. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **2019**, *28*, 1947–1951. [[CrossRef](#)]
119. Kanehisa, M.; Goto, S.; Sato, Y.; Furumichi, M.; Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **2012**, *40*, D109–D114. [[CrossRef](#)]
120. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)]
121. Kanehisa, M. A database for post-genome analysis. *Trends Genet.* **1997**, *13*, 375–376. [[CrossRef](#)]
122. Madden, E.; Logue, S.E.; Healy, S.J.; Manie, S.; Samali, A. The role of the unfolded protein response in cancer progression: From oncogenesis to chemoresistance. *Biol. Cell* **2019**, *111*, 1–17. [[CrossRef](#)]
123. Kyuno, D.; Takasawa, A.; Kikuchi, S.; Takemasa, I.; Osanai, M.; Kojima, T. Role of tight junctions in the epithelial-to-mesenchymal transition of cancer cells. *Biochim. Biophys. Acta Biomembr.* **2021**, *1863*, 183503. [[CrossRef](#)]
124. Le Bras, G.F.; Taubenslag, K.J.; Andl, C.D. The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression. *Cell Adhes. Migr.* **2012**, *6*, 365–373. [[CrossRef](#)]
125. Anczukow, O.; Krainer, A.R. Splicing-factor alterations in cancers. *RNA* **2016**, *22*, 1285–1301. [[CrossRef](#)] [[PubMed](#)]
126. Miyagi, T.; Tatsumi, T.; Takehara, T.; Kanto, T.; Kuzushita, N.; Sugimoto, Y.; Jinushi, M.; Kasahara, A.; Sasaki, Y.; Hori, M.; et al. Impaired expression of proteasome subunits and human leukocyte antigens class I in human colon cancer cells. *J. Gastroenterol. Hepatol.* **2003**, *18*, 32–40. [[CrossRef](#)]
127. Min, L.; Xu, H.; Wang, J.; Qu, L.; Jiang, B.; Zeng, Y.; Meng, L.; Jin, H.; Shou, C. N-alpha-acetyltransferase 10 protein is a negative regulator of 28S proteasome through interaction with PA28beta. *FEBS Lett.* **2013**, *587*, 1630–1637. [[CrossRef](#)]
128. Link, A.J.; Eng, J.; Schieltz, D.M.; Carmack, E.; Mize, G.J.; Morris, D.R.; Garvik, B.M.; Yates, J.R., 3rd. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999**, *17*, 676–682. [[CrossRef](#)] [[PubMed](#)]
129. Zhang, X.; Zhao, S.; He, Y.; Zheng, N.; Yan, X.; Wang, J. Pipeline for Targeted Meta-Proteomic Analyses to Assess the Diversity of Cattle Rumen Microbial Urease. *Front. Microbiol.* **2020**, *11*, 573414. [[CrossRef](#)] [[PubMed](#)]

130. Sakaya, G.R.; Parada, C.A.; Eichler, R.A.; Yamaki, V.N.; Navon, A.; Heimann, A.S.; Figueiredo, E.G.; Ferro, E.S. Peptidomic profiling of cerebrospinal fluid from patients with intracranial saccular aneurysms. *J. Proteom.* **2021**, *240*, 104188. [[CrossRef](#)] [[PubMed](#)]
131. Shahbazy, M.; Ramarathinam, S.H.; Illing, P.T.; Jappe, E.C.; Faridi, P.; Croft, N.P.; Purcell, A.W. Benchmarking Bioinformatics Pipelines in Data-Independent Acquisition Mass Spectrometry for Immunopeptidomics. *Mol. Cell. Proteom.* **2023**, *22*, 100515. [[CrossRef](#)]
132. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)] [[PubMed](#)]
133. Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An Open Source Software for Exploring and Manipulating Networks. In Proceedings of the International Aaai Conference on Web And Social Media, San Jose, CA, USA, 17–20 May 2009; pp. 361–362.
134. Liao, Y.; Wang, J.; Jaehnig, E.J.; Shi, Z.; Zhang, B. WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **2019**, *47*, W199–W205. [[CrossRef](#)] [[PubMed](#)]
135. Zhang, B.; Kirov, S.; Snoddy, J. WebGestalt: An integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* **2005**, *33*, W741–W748. [[CrossRef](#)]
136. Romero, P.; Obradovic, Z.; Li, X.; Garner, E.C.; Brown, C.J.; Dunker, A.K. Sequence complexity of disordered protein. *Proteins* **2001**, *42*, 38–48. [[CrossRef](#)] [[PubMed](#)]
137. 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](#)]
138. Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K.; Obradovic, Z. Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol.* **2005**, *3*, 35–60. [[CrossRef](#)]
139. Xue, B.; Dunbrack, R.L.; Williams, R.W.; Dunker, A.K.; Uversky, V.N. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* **2010**, *1804*, 996–1010. [[CrossRef](#)]
140. Dosztanyi, 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](#)]
141. Dosztanyi, Z.; Csizmok, V.; Tompa, P.; Simon, I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J. Mol. Biol.* **2005**, *347*, 827–839. [[CrossRef](#)]
142. Mohan, A.; Sullivan, W.J., Jr.; Radivojac, P.; Dunker, A.K.; Uversky, V.N. Intrinsic disorder in pathogenic and non-pathogenic microbes: Discovering and analyzing the unfoldomes of early-branching eukaryotes. *Mol. Biosyst.* **2008**, *4*, 328–340. [[CrossRef](#)]
143. Kyte, J.; Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **1982**, *157*, 105–132. [[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.