Stroke Paper

by Vijayakumar R

Submission date: 10-Nov-2022 01:38PM (UTC+0300)
Submission ID: 1950087182
File name: Manuscript_-Plagiarism.docx (597.38K)
Word count: 3299
Character count: 20025

Prediction of protein aggregation on key proteins involved in Ischemic stroke

Abstract

Stroke is a genetic condition comprising multiple subtypes and arising from both genic and other multi factors. Genetic basis of stroke is well established through several studies. Advances in integrating sequencing methods and Genome-wide association studies have shown that genetics of stroke is manifested in several genic disorders. Many of the neurodegenerative disorders show aggravated protein aggregation through amyloid formation. Through the protein aggregation prediction, we observed a higher protein disorder in 46 stroke-associated proteins. Also, we observed a large number of aggregation residues distributed as a pattern in multiple regions of these candidate proteins. Overall, we present a study showing that there is a possible interrelationship between protein aggregation and stroke.

Keywords: Genetic disorder, Ischemic stroke, Protein aggregation, Gene ontology, KEGG pathway

1. Introduction

Stroke is a complex heterogenous condition that is one of the major causes of ailment and death in the world. Characterizing stroke involves classification into subtypes. Several classification systems have been proposed to differentiate subtypes of stroke and distinguish between ischemic and hemorrhagic stroke, subarachnoid hemorrhage, cerebral venous thrombosis, and spinal cord stroke (Amarenco et al., 2009). Studies such as the association of monozygotic twins to stroke provide the evidence of implication of genetic factors in stroke pathophysiology. Next generation sequencing (NGS), genome-wide association studies (GWAS) provided evidence that stroke is both a monogenic and a polygenic disorder. Both above classification systems and genetic studies are vital in grouping patients for therapeutic purposes.

Whole-genome sequencing studies have resulted identifying novel variants associated with stroke subtypes. A genome-wide association study through Trans-Omics for Precision Medicine (TOPMed) Program identified 5 novel loci associated with subtypes of stroke in a multi-ancestry population (Y. Hu et al., 2022). Besides this, MEGASTROKE consortium have performed genotyping and GWAS studies that resulted in identification of stroke risk variants such as *NKX2-5*, *ANK2*, *LRCH1*, *REEP3*, *JAZF1*(de Vries et al., 2019; Malik et al., 2018).

Although multiple stroke specific risk loci are detected, functional characterization of these loci are challenging as these variants fall mostly on non-coding regions of the genome. Hence recent approaches use data from gene expression, DNA methylation to establish a causal relationship to stroke susceptible genes. Analysis of candidate genes involved in ischemic stroke resulted in identification of at least five susceptibility genes such as factor V Leiden Gln506, ACE I/D, MTHFR C677T, prothrombin G20210A, PAI-1 5G (Bentley et al., 2010)

Emerging evidence shows that protein aggregates formed in Ischemic stroke. Misfolded proteins tend to form fibers of aggregates (B. R. Hu et al., 2001). In particular ischemic stroke (Luo et al., 2013); stroke and aggregation (Zhang et al., 2020). Protein aggregates are found to be form deposits in degenerate cells and are involved in cellular toxicity (Tutar et al., 2013). Several neurological disorders such as Parkinson's Disease (PD), Huntington's disease (HD), prion diseases, Amyotrophic lateral sclerosis (ALS) are associated with the formation of protein aggregate (Pedersen & Heegaard, 2013). Aβ-peptide (1-40/1-42) forms amyloid plaque in regions such as cortex, hippocampus and forebrain. Proteins such as Tau, α-synuclein, Ataxins, superoxide dismutase (SOD1) and RNA binding proteins TDP43, FUS, TAF15 are found to form lewy bodies, intranuclear inclusion, axonal spheroids and cytoplasmic aggregates (Kumar et al., 2016).

Efforts on analysis of protein aggregation and characterization shown significant improvement in the development of protein aggregation prediction methods. More than 20 different computational algorithms are available now for the prediction of protein aggregation based on amino acid sequences (Santos et al., 2020). Tools such as TANGO, PASTA2.0, AGGRESCAN uses either protein features or experimental data to predict protein aggregation (Conchillo-Solé et al., 2007; de Groot et al., 2005; Walsh et al., 2014). Taking advantage of the available methods of protein aggregation prediction and sequences available on the stroke dataset, here we explored the connection or common theme between the above stroke and amyloid formation.

2. Methods

2.1 Dataset of candidate genes associated with disorders related to stroke:

We performed text mining and sequence database search through pubmed to obtain a base dataset of genes associated with Ischemic stroke. As stroke is linked to both monogenic and

polygenic disorders, we obtained list of genes reported earlier (Ekkert et al., 2022). Each of these genes in the dataset have certain impact on stroke pathogenesis. Mutiple literature and database search resulted in a dataset of 46 genes linked to risk of stroke.

2.2 Functional annotation of genes linked to stroke disorders

Curated canonical protein sequences are obtained from UniProtKB/Swiss-Prot protein sequence database (The UniProt Consortium, 2021). Only full-length protein sequences are used for each of these genes. Alternative sequences for each ids are avoided to remove redundancy and only unique sequences are further analyzed. For functional annotation of gene list, DAVID knowledgebase which is a webserver for bioinformatics resource providing functional enrichment analysis is utilized (Sherman et al., 2022).

2.3 Prediction of protein aggregation in stroke disorder associated protein sequences

To evaluate the tendency of proteins associated with stroke to form protein aggregate, we performed analysis whether these sequences forms β-sheet enriched secondary structure conformation. Using a pairwise energy potential, intrinsic disorder and secondary structure, protein aggregation calculation for the candidate protein sequences were performed in PASTA2.0 webserver (Walsh et al., 2014).

2.4 Amyloidogenic region in the protein sequences

Smaller fragments of regions in the protein sequences responsible for the amylodogenesis (Ivanova et al., 2004). These regions are composed of aminoacids which are unique and distinct from other non-aggregating regions or peptides. Using expected contact of the residues in the protein sequences, amyloidogenic regions are predicted (Garbuzynskiy et al., 2010).

3. Results

We have shown here that the protein aggregation might occur among the candidate proteins involved in stroke associated disorders. The pathophysiology between the neurodegenerative disorders and protein aggregation are shown to be shared. Our approach has shown that there could be a significant overlap between the pathophysiology of amyloid formation and ischemic stroke.

3.1 Genes involved in monogenic and polygenic disorders associated with stroke

Around 46 genes related to stroke associated disorders are retrieved from different databases. These genes are found to have around 687 splice variants in the UniprotKB. We used full length canonical protein sequences for further analysis. Annotation of genes retrieved through DAVID shows that most genes are related to signaling function including receptors (Table1). Multiple candidate genes functions have stroke phenotypic manifestation

3.2 Prediction of protein aggregation

Formation of amyloid aggregates is implicated in several neurodegenerative disorders. We use protein disorder as a scale for predicting protein aggregation. Propensity of aggregation remains relatively similar across multiple methods for candidate stroke related genes (Figure1). We further used PASTA2.0 to determine the percentage of protein disorder, number of amyloid within the protein sequence, percentage of α-helix, percentage of β-strand etc.Percentage disorder of proteins vary ranging from 1 to upto 80 for stroke associated genes. For a random dataset of non-stroke related genes this range from 1.5 to 63. Median value differs significantly between these two groups of proteins (Figure2). Statistical test (*t*-test) using R between the above two groups of proteins was performed. This test reveals a significantly lower *p*-value (*p*-value = 0.007744). This is highly significant and percentage disorder is higher for stroke associated genes.

3.3 Residue based prediction of aggregation specific protein region

Each protein sequences are predicted to have atleast 20 poteintial amyloid forming short amino acid sequence pattern (Table 2). Consensus pattern derived from multiple amyloid predicting tools such as TANGO, AGGRESCAN, WALTZ through FoldAmyloid shows that these patterns are detected in all methods (Figure1). For example, in protein WNTB-2B, these aggregation forming residues are distributed in 14 different sites (Figure3). Most of these patterns are 4-14 aminoacid length (Figure4).

4. Discussion

Analysis of protein aggregation from 46 protein sequences implicated in stroke manifestation shows that large number of proteins form amyloids with varying degree of protein disorder. The role of protein aggregation and amyloid formation in the neurodegenerative diseases are well established (Pedersen & Heegaard, 2013; Tutar et al., 2013). Also, earlier studies have shown that there is higher levels or induction of protein aggregation after cerebral Ischemia (Wu & Du, 2021). Compared to the random datasets of proteins from UniprotKB, these proteins are found to have higher percentage of protein disorder. This could arise from the β-strand composition of these sequences. Furthermore, structural variation including SNP and small indels in these genes could possibly contribute to the amyloid formation. Our work contributes to the evidence that protein aggregation could be implicated in stroke disorder. Further research following our findings would improve our knowledge on the molecular level overlap between these two related process and disorder. More experimental evidences are needed for implicating the above listed genes (their products) in the protein aggregation. Establishing mouse models for stroke and investigating the

aggregation through electron microscopy, laser-scanning confocal microscopy, and Western blotting could help further.

5. Conclusion

In the present study, we analyzed the aggregation properties of stroke related proteins. We selected a set of stroke-associated candidate proteins and a set of random control dataset. Overall, we observed that most proteins associated with stroke have higher protein disorder compared to a random dataset of protein sequences. These amyloids forming aggregating residues are distributed anywhere between the N-terminal and C-terminal part of the sequence of these candidate proteins. We found the contact frequency profile value of multiple residues are higher than average expected value and is part of disordered region related to protein conformation. Our study suggests that there is an overlap in pathophysiology of protein aggregation, neurological disorders and stroke related disorders.

6. Acknowledgment

The authors extend their appreciations to the deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number (IFP-2020-38).

7. References

Amarenco, P., Bogousslavsky, J., Caplan, L. R., Donnan, G. A., & Hennerici, M. G. (2009). Classification of stroke subtypes. *Cerebrovascular Diseases (Basel, Switzerland)*, *27*(5), 493–501. https://doi.org/10.1159/000210432

Bentley, P., Peck, G., Smeeth, L., Whittaker, J., & Sharma, P. (2010). Causal Relationship of Susceptibility Genes to Ischemic Stroke: Comparison to Ischemic Heart Disease and Biochemical Determinants. *PLOS ONE*, *5*(2), e9136. https://doi.org/10.1371/journal.pone.0009136

Conchillo-Solé, O., de Groot, N. S., Avilés, F. X., Vendrell, J., Daura, X., & Ventura, S. (2007). AGGRESCAN: A server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *BMC Bioinformatics*, *8*(1), 65. https://doi.org/10.1186/1471-2105-8-65

De Groot, N. S., Pallarés, I., Avilés, F. X., Vendrell, J., & Ventura, S. (2005). Prediction of "hot spots" of aggregation in disease-linked polypeptides. *BMC Structural Biology*, *5*(1), 18. https://doi.org/10.1186/1472-6807-5-18

De Vries, P. S., Sabater-Lleal, M., Huffman, J. E., Marten, J., Song, C., Pankratz, N., Bartz, T. M., de Haan, H. G., Delgado, G. E., Eicher, J. D., Martinez-Perez, A., Ward-Caviness, C. K., Brody, J. A., Chen, M.-H., de Maat, M. P. M., Frånberg, M., Gill, D., Kleber, M. E., Rivadeneira, F., … Smith, N. L. (2019). A genome-wide association study identifies new loci for factor VII and implicates factor VII in ischemic stroke etiology. *Blood*, *133*(9), 967–977. https://doi.org/10.1182/blood-2018-05-849240

Ekkert, A., Šliachtenko, A., Grigaitė, J., Burnytė, B., Utkus, A., & Jatužis, D. (2022). Ischemic Stroke Genetics: What Is New and How to Apply It in Clinical Practice? *Genes*, *13*(1), 48. https://doi.org/10.3390/genes13010048

Garbuzynskiy, S. O., Lobanov, M. Yu., & Galzitskaya, O. V. (2010). FoldAmyloid: A method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics*, *26*(3), 326–332. https://doi.org/10.1093/bioinformatics/btp691

Hu, B. R., Janelidze, S., Ginsberg, M. D., Busto, R., Perez-Pinzon, M., Sick, T. J., Siesjö, B. K., & Liu, C. L. (2001). Protein aggregation after focal brain ischemia and reperfusion. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *21*(7), 865–875. https://doi.org/10.1097/00004647-200107000-00012

Hu, Y., Haessler, J. W., Manansala, R., Wiggins, K. L., Moscati, A., Beiser, A., Heard-Costa, N. L., Sarnowski, C., Raffield, L. M., Chung, J., Marini, S., Anderson, C. D., Rosand, J., Xu, H., Sun, X., Kelly, T. N., Wong, Q., Lange, L. A., Rotter, J. I., … null, null. (2022). Whole-Genome Sequencing Association Analyses of Stroke and Its Subtypes in Ancestrally Diverse Populations From Trans-Omics for Precision Medicine Project. *Stroke*, *53*(3), 875–885. https://doi.org/10.1161/STROKEAHA.120.031792

Ivanova, M. I., Sawaya, M. R., Gingery, M., Attinger, A., & Eisenberg, D. (2004). An amyloid-forming segment of β2-microglobulin suggests a molecular model for the fibril. *Proceedings of the National Academy of Sciences*, *101*(29), 10584–10589. https://doi.org/10.1073/pnas.0403756101

Kumar, V., Sami, N., Kashav, T., Islam, A., Ahmad, F., & Hassan, Md. I. (2016). Protein aggregation and neurodegenerative diseases: From theory to therapy. *European Journal of Medicinal Chemistry*, *124*, 1105–1120. https://doi.org/10.1016/j.ejmech.2016.07.054

Luo, T., Park, Y., Sun, X., Liu, C., & Hu, B. (2013). Protein misfolding, aggregation, and autophagy after brain ischemia. *Translational Stroke Research*, *4*(6), 581–588. https://doi.org/10.1007/s12975-013-0299-5

Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.-K., van der Laan, S. W., Gretarsdottir, S., Anderson, C. D., Chong, M.,

Adams, H. H. H., Ago, T., Almgren, P., Amouyel, P., Ay, H., Bartz, T. M., Benavente, O. R., … Dichgans, M. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nature Genetics*, *50*(4), 524–537. https://doi.org/10.1038/s41588-018-0058-3

Pedersen, J. T., & Heegaard, N. H. H. (2013). Analysis of Protein Aggregation in Neurodegenerative Disease. *Analytical Chemistry*, *85*(9), 4215–4227. https://doi.org/10.1021/ac400023c

Santos, J., Pujols, J., Pallarès, I., Iglesias, V., & Ventura, S. (2020). Computational prediction of protein aggregation: Advances in proteomics, conformation-specific algorithms and biotechnological applications. *Computational and Structural Biotechnology Journal*, *18*, 1403–1413. https://doi.org/10.1016/j.csbj.2020.05.026

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, gkac194. https://doi.org/10.1093/nar/gkac194

The UniProt Consortium. (2021). UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480–D489. https://doi.org/10.1093/nar/gkaa1100

Tutar, Y., Özgür, A., & Tutar, L. (2013). Role of Protein Aggregation in Neurodegenerative Diseases. In *Neurodegenerative Diseases*. IntechOpen. https://doi.org/10.5772/54487

Walsh, I., Seno, F., Tosatto, S. C. E., & Trovato, A. (2014). PASTA 2.0: An improved server for protein aggregation prediction. *Nucleic Acids Research*, *42*(Web Server issue), W301–W307. https://doi.org/10.1093/nar/gku399

Wu, S., & Du, L. (2021). Protein Aggregation in the Pathogenesis of Ischemic Stroke. *Cellular and Molecular Neurobiology*, *41*(6), 1183–1194. https://doi.org/10.1007/s10571-020-00899-y

Zhang, X., Wesén, E., Kumar, R., Bernson, D., Gallud, A., Paul, A., Wittung-Stafshede, P., & Esbjörner, E. K. (2020). Correlation between Cellular Uptake and Cytotoxicity of Fragmented α-Synuclein Amyloid Fibrils Suggests Intracellular Basis for Toxicity. *ACS Chemical Neuroscience*, *11*(3), 233–241. https://doi.org/10.1021/acschemneuro.9b00562

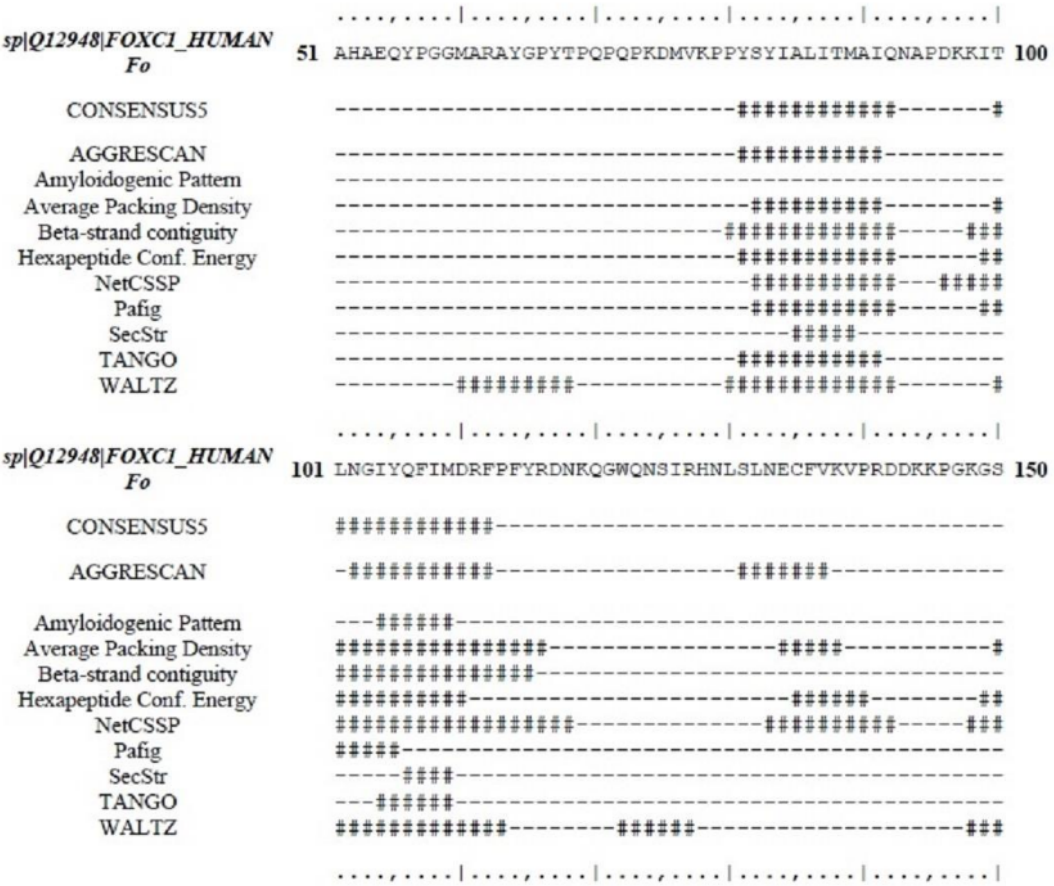
LEGENDS

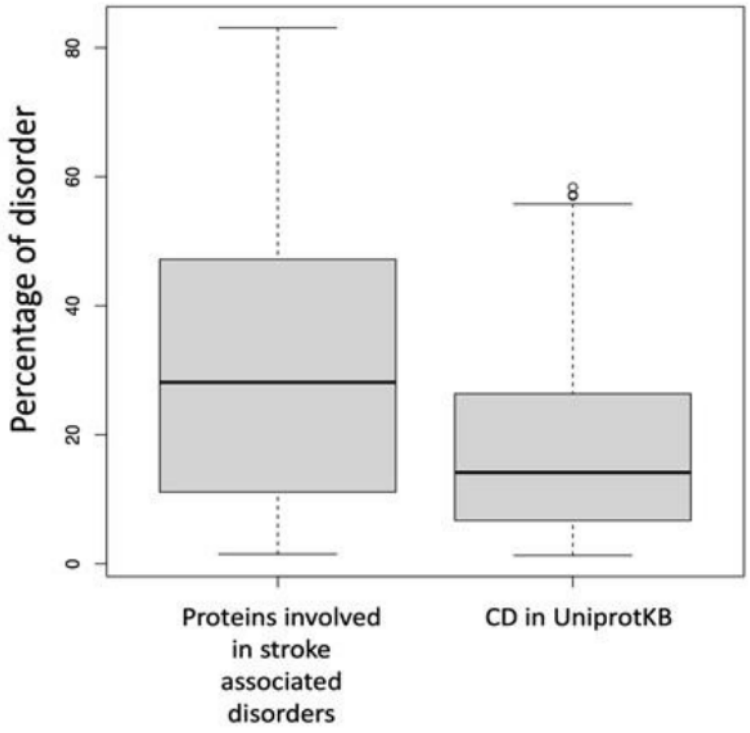
Figure1: Consensus methods predicting same amyloid regions

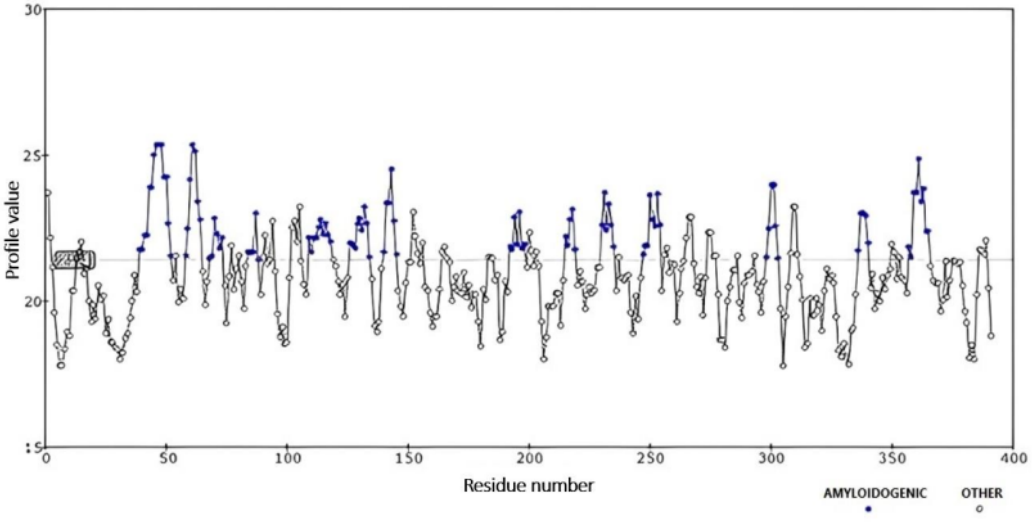
Figure2: Boxplot showing the differences in the disorder among two groups of proteins

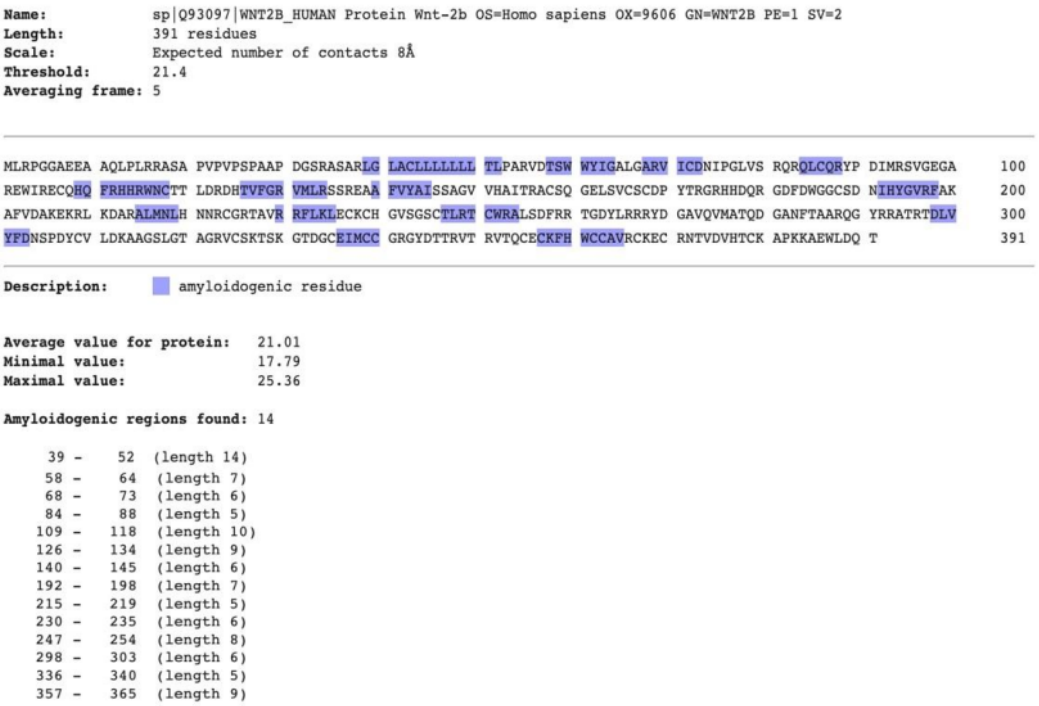
Figure3: Contact frequency predicted for different residues across the protein sequence is shown here

Figure4: Amyloidogenic residues predicted for human Wnt-2b protein




Percentage of disorder
0
20
40
60
80
Proteins involved in stroke associated disorders
CD in UniprotKB


Profile value
15
20
25
30
0
50
100
150
200
250
300
350
400
Residue number
AMYLOIDOGENIC
OTHER



Tables

Table1: Candidate gene ids and their functional annotation of genes involved in stroke associated disorders

Table2: Protein aggregation propensity for the genes involved in stroke associated disorders

Table1:
Candidate gene ids and their functional annotation of genes involved in stroke associated disorders

SI	From	Species	David Gene Name
1	NOTCH3	Homo sapiens	notch receptor 3(NOTCH3)
2	FOXC1	Homo sapiens	forkhead box C1(FOXC1)
3	CASZ1	Homo sapiens	castor zinc finger 1(CASZ1)
4	WNT2B	Homo sapiens	Wnt family member 2B(WNT2B)
5	LINC01492	Homo sapiens	long intergenic non-protein coding RNA 1492(LINC01492)
6	HTRA1	Homo sapiens	HtrA serine peptidase 1(HTRA1)
7	ADCY2	Homo sapiens	adenylate cyclase 2(ADCY2)
8	PRPF8	Homo sapiens	pre-mRNA processing factor 8(PRPF8)
9	HDAC9	Homo sapiens	histone deacetylase 9(HDAC9)
10	ABO	Homo sapiens	ABO, alpha 1-3-N-acetylgalactosaminyltransferase and alpha 1-3-galactosyltransferase(ABO)
11	ZCCHC14	Homo sapiens	zinc finger CCHC-type containing 14(ZCCHC14)
12	EDNRA	Homo sapiens	endothelin receptor type A(EDNRA)
13	SH3PXD2A	Homo sapiens	SH3 and PX domains 2A(SH3PXD2A)
14	CBS	Homo sapiens	cystathionine beta-synthase(CBS)
15	PITX2	Homo sapiens	paired like homeodomain 2(PITX2)
16	ZNF566	Homo sapiens	zinc finger protein 566(ZNF566)
17	NKX2-5	Homo sapiens	NK2 homeobox 5(NKX2-5)
18	SH2B3	Homo sapiens	SH2B adaptor protein 3(SH2B3)
19	HABP2	Homo sapiens	hyaluronan binding protein 2(HABP2)
20	RGS7	Homo sapiens	regulator of G protein signaling 7(RGS7)
21	FGA	Homo sapiens	fibrinogen alpha chain(FGA)
22	ZFHX3	Homo sapiens	zinc finger homeobox 3(ZFHX3)
23	FOXF2	Homo sapiens	forkhead box F2(FOXF2)
24	TREX1	Homo sapiens	three prime repair exonuclease 1(TREX1)
25	ABCC6	Homo sapiens	ATP binding cassette subfamily C member 6(ABCC6)
26	ANK2	Homo sapiens	ankyrin 2(ANK2)
27	PDZK1IP1	Homo sapiens	PDZK1 interacting protein 1(PDZK1IP1)
28	TBX3	Homo sapiens	T-box transcription factor 3(TBX3)
29	MMP12	Homo sapiens	matrix metallopeptidase 12(MMP12)
30	COL3A1	Homo sapiens	collagen type III alpha 1 chain(COL3A1)

31	LRCH1	Homo sapiens	leucine rich repeats and calponin homology domain containing 1(LRCH1)
32	CDK6	Homo sapiens	cyclin dependent kinase 6(CDK6)
33	GAL	Homo sapiens	galanin and GMAP prepropeptide(GAL)
34	COL4A2	Homo sapiens	collagen type IV alpha 2 chain(COL4A2)
35	COL4A1	Homo sapiens	collagen type IV alpha 1 chain(COL4A1)
36	PDE3A	Homo sapiens	phosphodiesterase 3A(PDE3A)
37	KCNK3	Homo sapiens	potassium two pore domain channel subfamily K member 3(KCNK3)
38	LOC100505841	Homo sapiens	zinc finger protein 474-like(LOC100505841)
39	FBN1	Homo sapiens	fibrillin 1(FBN1)
42	ILF3	Homo sapiens	interleukin enhancer binding factor 3(ILF3)
43	CDKN2A	Homo sapiens	cyclin dependent kinase inhibitor 2A(CDKN2A)
44	ZNF318	Homo sapiens	zinc finger protein 318(ZNF318)
45	FURIN	Homo sapiens	furin, paired basic amino acid cleaving enzyme(FURIN)
46	TM4SF4	Homo sapiens	transmembrane 4 L six family member 4(TM4SF4)
47	PMF1	Homo sapiens	polyamine modulated factor 1(PMF1)
48	SMARCA4	Homo sapiens	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4(SMARCA4)

Table2: Protein aggregation propensity for the genes involved in stroke associated disorders.

Protein name (from fasta header)	length	# amyloids	best energy	% disorder	% α-helix	% β-strand	% coil
spQ08462	1091	20	-39.221871	4.216	64.25	8.07	27.68
spO95255	1503	20	-19.941325	8.715	59.41	6.25	34.33
spQ9Y2L9	728	20	-19.362124	27.33	42.58	10.03	47.39
spP09958	794	20	-18.84564	18.26	15.49	25.06	59.45
spQ13113	114	20	-17.268009	16.66	48.25	8.77	42.98
spO14649	394	20	-17.165792	28.17	65.48	2.79	31.73
spP25101	427	20	-16.917126	7.259	64.64	6.32	29.04
spQ9UM47	2321	20	-15.802264	13.09	17.36	16.59	66.05
spP48230	202	20	-15.752635	7.92	54.46	2.97	42.57
spQ6P2Q9	2335	20	-14.995744	1.498	46.55	13.28	40.17
spQ14432	1141	20	-14.680867	33.74	41.89	7.01	51.1
spP16442	354	20	-14.397722	3.672	33.9	16.95	49.15
spP02671	866	20	-10.907301	35.33	16.17	24.02	59.82
spQ01484	3957	20	-10.281311	42.91	26.38	14.43	59.19
spP35520	551	20	-10.195829	15.06	37.75	14.7	47.55
spQ5TCZ1	1133	20	-9.781373	45.18	11.92	21.27	66.81
spQ15911	3703	20	-9.775551	52.17	32.19	8.37	59.44
spQ93097	391	20	-9.619822	10.48	45.78	14.07	40.15

spQ8N726	132	20	-9.546539	66.66	11.36	14.39	74.24
spQ12906	894	20	-9.535643	47.2	28.75	10.18	61.07
spQ00534	326	20	-9.230796	13.49	43.86	16.87	39.26
spQ92743	480	20	-8.959704	13.33	18.96	30.83	50.21
spQ9UQQ2	575	20	-8.630576	45.04	24.35	14.61	61.04
spQ12947	444	20	-8.290021	52.02	22.52	3.83	73.65
spP51532	1647	20	-8.254159	50.15	46.63	5.22	48.15
spQ9UKV0	1011	20	-8.19617	29.57	40.26	8.51	51.24
spP35555	2871	20	-7.913752	3.065	4.25	31.8	63.95
spP39900	470	20	-7.626874	4.042	23.83	23.83	52.34
spP49802	495	11	-7.622278	11.11	52.93	2.42	44.65
spP02462	1669	20	-7.454205	83.1	2.64	8.63	88.74
spQ14520	560	20	-7.387258	6.607	14.46	26.43	59.11
spQ9NSU2	314	20	-6.854083	27.7	40.45	6.69	52.87
spP08572	1712	20	-6.646477	62.44	2.69	9.35	87.97
spO15119	743	9	-6.556453	32.03	28.4	11.84	59.76
spQ12948	553	6	-6.366574	65.82	22.78	6.69	70.52
spQ5VUA4	2279	11	-6.276438	50.89	27.29	10.79	61.91
spQ969W8	418	1	-5.915617	3.588	22.73	18.66	58.61
spP42771	156	7	-5.761273	28.2	50	0	50
spQ86V15	1759	20	-5.735543	44.11	19.56	17.45	62.99
spQ8N6F7	178	4	-5.629332	28.08	21.91	13.48	64.61
spP22466	123	3	-5.60644	80.48	52.03	0	47.97
spQ99697	317	3	-5.452123	34.7	37.54	3.79	58.68
spQ8WYQ9	949	1	-5.372969	50.36	18.02	16.97	65.02
spP02461	1466	2	-5.361421	77.96	3.96	6.48	89.56
spQ6P1K2	205	1	-5.021246	22.43	76.59	0	23.41
spP52952	324	0	-4.668274	15.12	30.56	11.11	58.33

Stroke Paper

ORIGINALITY REPORT

19%	18%	14%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ebi.ac.uk Internet Source	2%
2	repository.kaust.edu.sa Internet Source	1%
3	hdl.handle.net Internet Source	1%
4	mafiadoc.com Internet Source	1%
5	repositorio.uam.es Internet Source	1%
6	www.ncbi.nlm.nih.gov Internet Source	1%
7	fp.amegroups.cn Internet Source	1%
8	Submitted to Cardiff University Student Paper	1%
9	www.oalib.com Internet Source	1%

10	Sarah R. Evans, Colista West, Judith Klein-Seetharaman. "Similarity of the non-amyloid-β component and C-terminal tail of monomeric and tetrameric alpha-synuclein with 14-3-3 sigma", Computational and Structural Biotechnology Journal, 2021 Publication	1 %
11	ddd.uab.cat Internet Source	1 %
12	repository.usmf.md Internet Source	1 %
13	Submitted to University of Westminster Student Paper	1 %
14	bdb2.ucsd.edu Internet Source	<1 %
15	www.genscript.com Internet Source	<1 %
16	Maxwell Korang-Yeboah, Stephanie Ketcham, Mack Shih, Ann-Marie Ako-Adounvo et al. "Effect of formulation and peptide folding on the fibrillar aggregation, gelation, and oxidation of a therapeutic peptide", International Journal of Pharmaceutics, 2021 Publication	<1 %
17	E. Andersson. "High-resolution genomic profiling reveals gain of chromosome 14 as a	<1 %

	predictor of poor outcome in ileal carcinoids", Endocrine Related Cancer, 05/20/2009 Publication	
18	www.genes2cognition.org Internet Source	<1 %
19	Choi, Y.W.. "Identification of Differentially Expressed Genes Using Annealing Control Primer-based GeneFishing in Human Squamous Cell Cervical Carcinoma", Clinical Oncology, 200706 Publication	<1 %
20	Submitted to Universiti Teknologi MARA Student Paper	<1 %
21	coek.info Internet Source	<1 %
22	edocs.fu-berlin.de Internet Source	<1 %
23	publikationen.bibliothek.kit.edu Internet Source	<1 %
24	pubmed.ncbi.nlm.nih.gov Internet Source	<1 %
25	www.biorxiv.org Internet Source	<1 %
26	www.medrxiv.org Internet Source	<1 %

27	www.science.gov Internet Source	<1%
28	"Global Virology II - HIV and NeuroAIDS", Springer Science and Business Media LLC, 2017 Publication	<1%
29	Hye Ran Yang, Jae Sung Ko, Jeong Kee Seo. "Analysis of Gene Expression in Helicobacter pylori-associated Nodular Gastritis in Children Using Microarray", Korean Journal of Pediatric Gastroenterology and Nutrition, 2010 Publication	<1%
30	Palanisamy Manikandan, Rajendran Vijayakumar, Bader Alshehri, Subramanian Senthilkumar et al. "Novel approach to unravel the Heat shock proteins (HSPs) with anti-ischemic stroke and human infections", Journal of Infection and Public Health, 2022 Publication	<1%
31	Swagata Das, Uttam Pal, Supriya Das, Khyati Bagga, Anupam Roy, Arpita Mrigwani, Nakul C. Maiti. "Sequence Complexity of Amyloidogenic Regions in Intrinsically Disordered Human Proteins", PLoS ONE, 2014 Publication	<1%
32	Vijay Kumar, Neha Sami, Tara Kashav, Asimul Islam, Faizan Ahmad, Md. Imtaiyaz Hassan.	<1%

	"Protein aggregation and neurodegenerative diseases: From theory to therapy", European Journal of Medicinal Chemistry, 2016 Publication	
33	academic.oup.com Internet Source	<1 %
34	chs-nhlbi.org Internet Source	<1 %
35	worldwidescience.org Internet Source	<1 %
36	Rajendran Vijayakumar, Palanisamy Manikandan, Faiz Alfaiz, Mohammad Saleh Al Aboodi et al. "Mutational signatures on ischemic stroke-associated genes in saudi human reference genome", Journal of King Saud University - Science, 2022 Publication	<1 %
37	Stéphanie Debette, Hugh S. Markus. "Stroke Genetics: Discovery, Insight Into Mechanisms, and Clinical Perspectives", Circulation Research, 2022 Publication	<1 %
38	Helmar C. Lehmann. "Human Schwann cells retain essential phenotype characteristics after immortalization", Stem Cells and Development, 05/17/2011 Publication	<1 %

39 Yasushi TAKAGI, Tomohiro AOKI, Jun C. TAKAHASHI, Kazumichi YOSHIDA et al. "Differential Gene Expression in Relation to the Clinical Characteristics of Human Brain Arteriovenous Malformations", Neurologia medico-chirurgica, 2014

Publication

<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Stroke Paper

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18