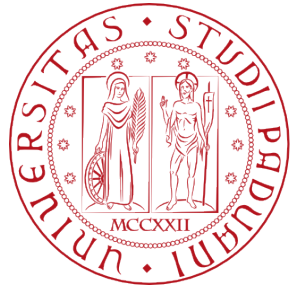
System biology of Activity-regulated cytoskeleton-associated protein (Arc/Arg 3.1)

Nguyen Xuan Tung
July 2022



Università degli Studi di Padova

Overview

Summary

ARC protein is released from neurons in extracellular vesicles that mediate the transfer of ARC mRNA into new target cells. However, the mechanism of Arc development remains unknown. In this project, our work is to investigate the role of protein–protein interactions (PPIs) and pathway interactions network in Arc

Introduction:
Idea of the project

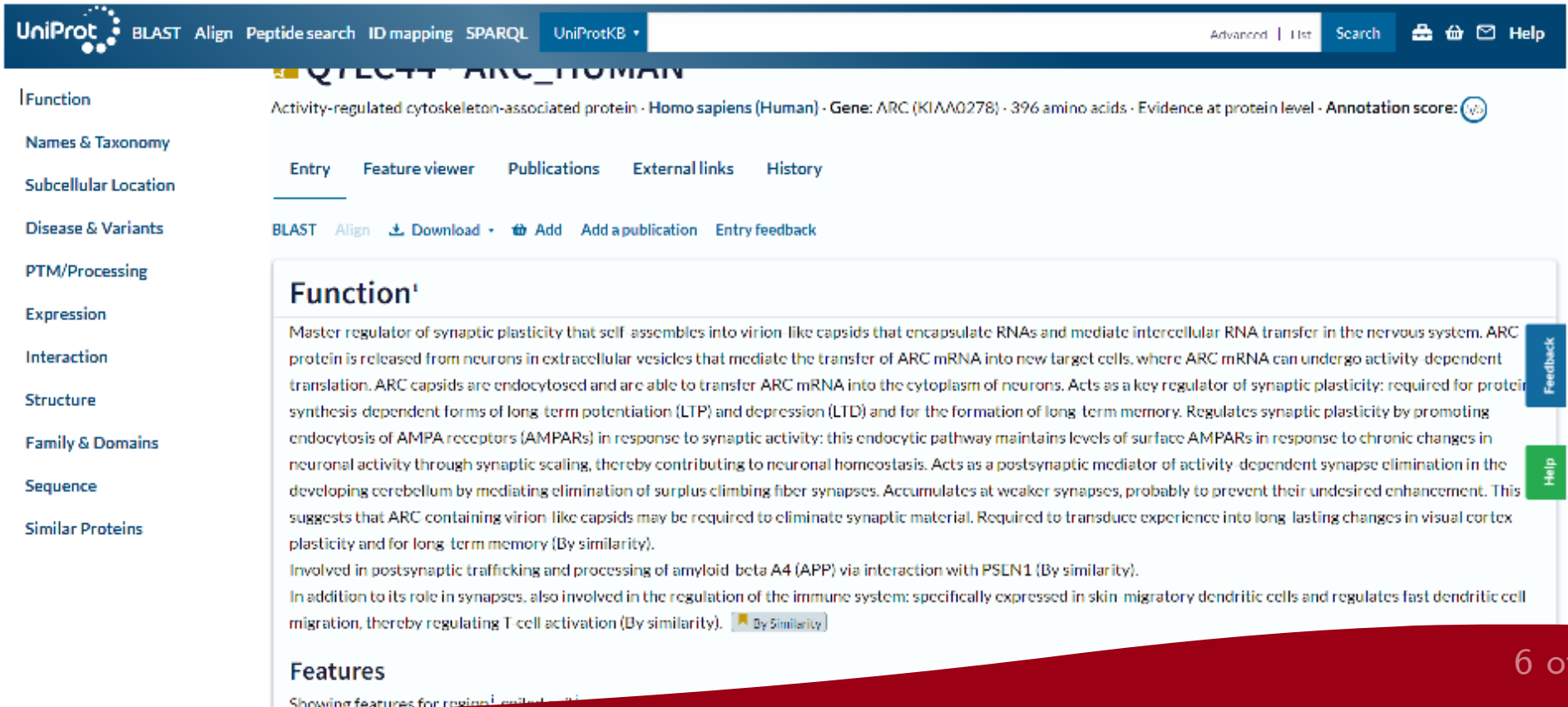
The interaction network between proteins is one of the significant way to reveal the function of proteins the relationship between them. To identify significant proteins associated with disease, a statistical approach is necessary and is an efficient solution against the experimental constraint.

Introduction:
Idea of the project

- The project is divided into 2 parts:
- $\longrightarrow$ Invest the protein-protein interaction (PPI) network and find the hub proteins which contribute the most to the network.
- $\longrightarrow$ With the data of PPI network, we create the pathway network and find the hub nodes which contribute the most to the network.

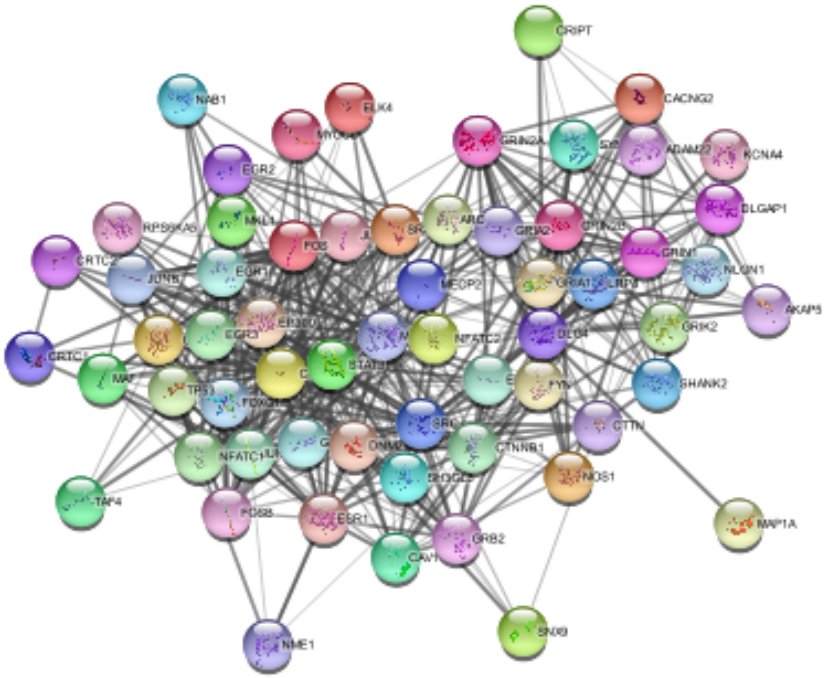
Method: Data collection

- The PPI data is collected in "STRING": https://string-db.org/
- The general data of Arc proten is collected in "Uniprot": https://www.uniprot.org/uniprotkb/Q7LC44/entry



Method: Data analysis

- We use the Search Tool for the Retrieval of Interacting Genes/Proteins database (STRING) to construct the PPI network associated with Arc, the global PPI dataset containing 505 interactions among 61 unique human proteins was obtained with medium confidence score 0.4.
- After get the PPI, we can construct it with Cytoscape app.



Method: PPI analysis

- Using the analysis function in Cytoscape, we adopted several topological measures of the network.
- degree (k), between centrality (BC), eccentricity, closeness centrality (CC), and clustering coefficient(ClusteringCoeff).

	A	B	C	D	E	F	G	H
1	name	BetweennessCentrality	ClosenessCentrality	ClusteringCoefficient	Degree	Eccentricity	NumberOfUndirectedEdges	
2	CREB1	0.103257344	0.75	0.397435897	40	2	40	
3	FOS	0.07227287	0.731707317	0.425320057	38	2	38	
4	DLG4	0.111118507	0.681818182	0.39516129	32	2	32	
5	CTNNB1	0.069697248	0.681818182	0.441532258	32	2	32	
6	JUN	0.039215388	0.674157303	0.483870968	31	2	31	
7	EGR1	0.032046119	0.666666667	0.52183908	30	2	30	
8	EP300	0.037056835	0.625	0.477832512	29	3	29	
9	GRIN2B	0.03768341	0.64516129	0.481481481	27	2	27	
10	SRC	0.032520763	0.64516129	0.524216524	27	2	27	
11	MAPK1	0.022132544	0.64516129	0.54985755	27	2	27	
12	SRF	0.024213181	0.606060606	0.523333333	25	3	25	
13	CREBBP	0.014544175	0.588235294	0.576666667	25	3	25	
14	TP53	0.013934721	0.576923077	0.557971014	24	3	24	
15	STAT3	0.011724405	0.582524272	0.59057971	24	3	24	
16	GRIN1	0.01857389	0.618556701	0.553359684	23	2	23	
17	ARC	0.036289615	0.606060606	0.533333333	21	2	21	
18	DNM2	0.025908058	0.606060606	0.471428571	21	2	21	
19	GRIN2A	0.017301359	0.594059406	0.566666667	21	3	21	
20	GRIA1	0.013837507	0.594059406	0.585714286	21	3	21	
21	ESR1	0.008385064	0.560747664	0.60952381	21			
22	JUND	0.005353123	0.588235294					

Method: Pathway network analysis

- In order to find the pathway of the protein interaction network of Arc, we use Cytoscape plugin named StringApp to generate the enrichment map.
- Using the analysis function in Cytoscape, we adopted several topological measures of the network.
- In this project, we only select KEGG pathway to analysis.

B	C	D	E	F	G	H	I
nrichmen	EnrichmentMap::GS_DESCR	EnrichmentMap::gs_size	Degree	AverageShortestPathLength	BetweennessCentrality	ClosenessCentrality	ClusteringCoefficient
TNNB1 | N	Kaposi sarcoma-associated herpesvirus infection	12	18	2.232758621	0.327358505	0.447876448	0.33333333
TNNB1 | N	Prostate cancer	7	28	2.25	0.060361433	0.444444444	0.36243386
TNNB1 | N	Human papillomavirus infection	7	28	2.25	0.060361433	0.444444444	0.36243386
MAPK1 | EG	Hepatitis B	14	20	2.327586207	0.036978039	0.42962963	0.36842105
MAPK1 | FG	Yersinia infection	6	23	2.353448276	0.035066003	0.424908425	0.57707509
MAPK1 | FG	Choline metabolism in cancer	4	31	2.362068966	0.040694005	0.423357664	0.41505376
TNNB1 | N	Colorectal cancer	6	29	2.396551724	0.035901462	0.417266187	0.41625615
MAPK1 | FG	Human immunodeficiency virus 1 infection	5	21	2.396551724	0.046531098	0.417266187	0.54761904
TNNB1 | N	Pathways in cancer	11	17	2.396551724	0.028825965	0.417266187	0.44117647
TNNB1 | N	Breast cancer	7	28	2.405172414	0.030370587	0.415770609	0.42328042
MAPK1 | TP	Glioma	3	27	2.413793103	0.017465801	0.414285714	0.4900284
MAPK1 | TP	Chronic myeloid leukemia	3	27	2.413793103	0.017465801	0.414285714	0.4900284
MAPK1 | ST	PD-L1 expression and PD-1 checkpoint pathway in cancer	6	20	2.431034483	0.027625714	0.411347518	0.57894736
MAPK1 | TP	Apoptosis	4	28	2.431034483	0.051111066	0.411347518	0.39417989
MAPK1 | ST	Th17 cell differentiation	6	20	2.431034483	0.027625714	0.411347518	0.57894736
MAPK1 | TP	Endocrine resistance	7	29	2.431034483	0.0377619	0.411347518	0.41379310
MAPK1 | CA	Oxytocin signaling pathway	7	19	2.448275862	0.020619355	0.408450704	0.6549707
MAPK1 | CF	Growth hormone synthesis, secretion and action	8	15	2.465517241	0.029072202	0.405594406	0.47619047
TNNB1 | N	Human cytomegalovirus infection	9	16	2.5	0.050056755	0.4	0.40833333
MAPK1 | TP	Neurotrophin signaling pathway	5	20	2.50862069	0.027802188	0.39862543	0.4

Results: Protein interaction networks

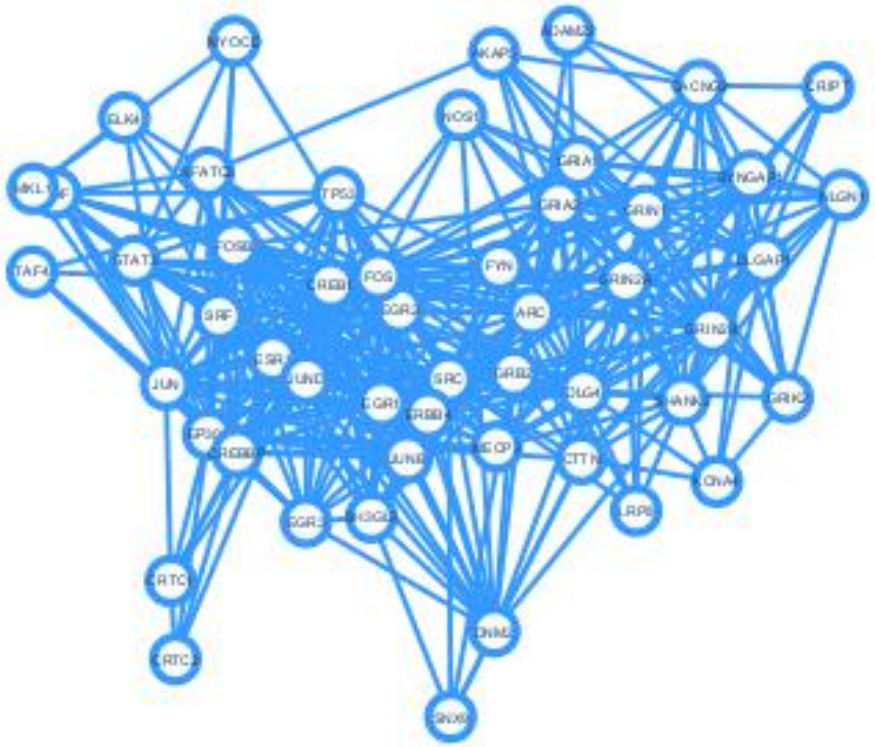


Figure: The Arc PPI network. The network contains 61 nodes with 505 edges. Data was downloaded from String database in TSV file and generated by Cytoscape.

Results: Protein interaction networks

- CREB1 was a hub (with the largest degree k=40) and bottleneck (with the second highest BC=0.103257) in the PPI network
- DLG4 was a bottleneck (with the highest BC=0.1111) and also a hub (with higher degree than average)
- Another node can be considered is FOS with the second largest degree k and third largest BC (38 and 0.0722 respectively).

Results: Protein interaction networks

- In 3 sample KEGG pathways: amphetamine, cocain, and nicotine addiction, we find the proteins involved in them.
- CREB1 as one of its center which contribute in amphetamine and cocain addiction pathway
- FOS is also involved in the development of amphetamine, cocain and nicotine addiction.

Pathway interaction network

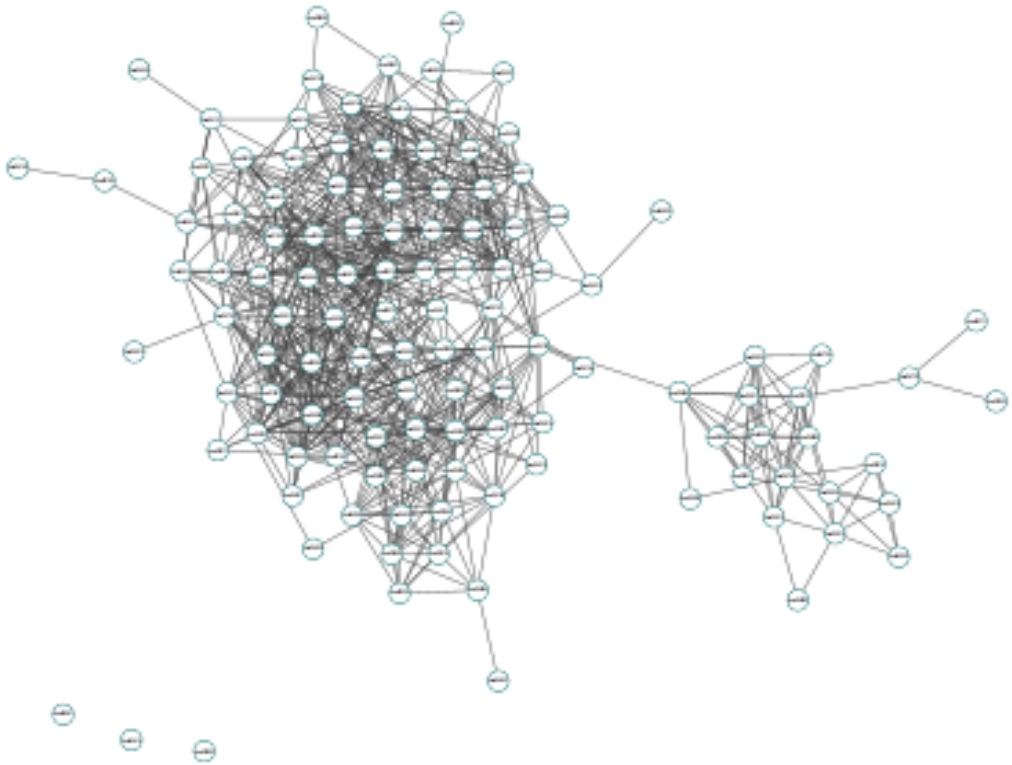


Figure: Pathway interaction network. Nodes were on behalf of pathways and edges stood for the interaction between any two pathways

Pathway interaction network

- we can find that the hub pathway and the center of the network is Choline metabolism in cancer (with highest degree $k = 31$)
- together with second and third largest degree respectively, Colorectal cancer, Endocrine resistance ($k = 29$), and Apoptosis, Prostate cancer ($k = 28$).

Conclusion

- The results show that there is a relationship between proteins interaction in PPI network.
- CREB1 as one of its center which contribute in amphetamine and cocain addiction pathway.
- FOS is also involved in the development of amphetamine, cocain and nicotine addiction.
- It is shown that CREB1 and FOS are the keys protein for studying the formulation of protein interaction network of Arc.

Conclusion

- We find that hsa05231 (Choline metabolism in cancer), hsa05210 (Colorectal cancer), hsa1522 (Endocrine resistance), hsa04210 (Apoptosis), and hsa05215 (Prostate cancer) are the hub nodes of the network.
- The pathway Choline metabolism in cancer was regarded as the seed pathway since it is the center of the network (with lowest eccentricity and sixth largest CC).
- However, the role of them in Arc still remains unclear and requires more studies.


Thank
You