

BIOINFORMATICS ANALYSIS OF PROTEIN DOMAINS INVOLVED IN SIGNAL TRANSDUCER ACTIVITY IN HUMAN AND MOUSE

Satabdi Basu and Debaprasad Mukherjee*

**Department of Bioinformatics, Purabi Das School of Information Technology,
Bengal Engineering and Science University, Shibpur, Howrah-711 103,
West Bengal, India.**

E-mail Address: - dm@pdsit.becdu.ac.in

***corresponding author**

The complexity of the entire signaling system makes for challenging research, but once understood it holds the promise for better treatments for cancer and other diseases. Signal transduction systems are therefore important in multicellular organisms, because of the need to coordinate the activities of hundreds to trillions of cells. At the cellular level it refers to the movement of signals from outside the cell to inside. The whole Signal Transduction is conducted with the help of signaling proteins and the domains within the proteins.

Here, we have emphasized our study on the signaling proteins and domains. Domains are basic evolutionary units of proteins. Domains or modules are defined as structurally compact, independently folding parts of protein molecules. Domains' shuffling is mechanism of protein evolution. Protein domain architectures link evolutionarily related proteins and underscore their shared functions. Permutation and combination of protein domain may be responsible of complexity of domain architecture and functions of organisms. The analysis of signaling domains can reveal much information about the evolutionary complexity of signaling process. So in this work we have made a comparative analysis of usage of protein domain and their domain architecture in "Signal Transducer Activity". Here we analyzed the different characteristics based on their process, function, location, composition, parent,children number of protein count and differences in protein count of these signaling domains in mouse and human and observed that ([IPR000014](#)) "PAS" domains showed highest protein number of matches in mouse and human. PAS domains are used as a signal sensor domain. The significant difference is found to be 25 in ([IPR011025](#)) "G protein alpha subunit, helical insertion",([IPR009143](#)) "Wnt-6 protein" and

(IPR004105) “kinase_dim” showed no difference in protein matches. **(IPR009143)** has equal number of protein count in mouse and human.

In order to gain insight into the complexity and reason behind the variations in protein count we have further extended our study on domain architecture (IDA) for all 38 signaling domains. It has been also observed that though **(IPR000014)** “PAS” showed less difference in protein count but it showed high number of domain architectures whereas **(IPR011025)** “G protein alpha subunit, helical insertion” showed significant difference but it has only one domain architecture. We have also observed that both **(IPR009143)** “Wnt-6 protein” and **(IPR004105)** “kinase_dim” showed no differences in protein count but **IPR009143** “Wnt-6 protein” has one domain architecture whereas other one doesn’t have any architecture. Therefore we can conclude that distribution or the arrangement of domain architectures in the domains can give rise to the complexity in the mammals.

This analytical study provided better understanding about domain distribution of these functional domains present in higher eukaryotes. The distribution and arrangement of domain architectures increase the complexity in signaling pathway. Similarity and diversity in the usage of protein domains and their architecture results in the unity and diversity of form and function of living organism.

So far we have analyzed the scenario in humans and mouse but would like to extend the study to other mammals especially those showing close similarity with humans like Chimpanzee etc. These will provide better understanding of the complexity of function in higher mammals.

SL NO.	INTERPRO ID	DESCRIPTION	NO. OF PROTEIN MATCHES	NO. OF PROTEIN MATCHES IN	DIFFERENCE BETWEEN MOUSE AND HUMAN
1	IPR000014	PAS	105	94	11
2	IPR011025	G protein alpha subunit, helical insertion	39	64	25
3	IPR009143	Wnt-6 protein	1	1	0
4	IPR004105	kinase_dim	0	0	0

Table 1 represents the domain **(IPR000014)** having highest protein count and domains**(IPR011025)**, **(IPR008341)**, **(IPR009143)** and **(IPR004105)** showing high, low and no difference in protein count in human and mouse.

SL No.	INTERPRO ID	NO. DOMAIN ARCHITECTURES		INTERPRO DOMAIN ARCHITECTURE (IDA)CODE	PROTEIN COUNT CORRESPONDING TO IDA		DIFFERENCE IN PROTEIN COUNT
		Human	Mouse		Human	Mouse	
1	IPR000014	16	18	IDA1092,13767,13655	28	37	11
				IDA13655	10	20	
				IDA1092,13767	7	7	
				IDA1092,13767,14935,14920,10011	7	6	
				IDA13655,5821,595	6	8	
				IDA1092,13655,14887	4	2	
				IDA13938,1789,13767,2073	4	3	
				IDA1092,13767,13655,10578	3	2	
				IDA13767,719	3	1	
				IDA1092,13767,14935,14920,10011x2	3	2	
				IDA13767,5821,595	3	6	
				IDA1092	2	0	
				IDA13767	1	1	
				IDA13767,13655	1	0	
				IDA5821	1	2	
				IDA5821,595	1	1	
				IDA13938,1789,13767	0	1	
				IDA1789,13656,2073	0	1	
IDA1092,13767,14935	0	1					
IDA1092,13767,13655,14887	0	2					
2	IPR011025	1	1	IDA1019	39	64	25
3	IPR009143	1	1	IDA5817	3	3	0
4	IPR004105	0	0	-	0	0	0

Table 2 represents the domain architectures of the domain showing highest protein count, domains having highest and no differences in protein count.