

# Study of Domains of Plant Argonaute Protein Sequences

**Protip Basu**

Assistant Professor  
Department of Botany  
Siliguri College

## **Abstract:**

Argonaute proteins are found in all types of living organisms, existing across all domains and kingdoms of life. Though they fall under the category of uncharacterized proteins, they ubiquitously act as slicers of various non-coding RNAs and this process leads to the silencing of mRNAs, which become function-less. This paper deals with the detection of analysis of various domains found in the plant argonaute sequences of some plants. Using this work plan, work can be undertaken for other domains of life, may help in furthering our knowledge about the Argonautes.

## **1. INTRODUCTION**

Argonautes are proteins found in across all domains of the living world. These proteins are uncharacterized proteins, which act as slicers of various non-coding RNAs. This process destroys cognate mRNAs, leading to 'silencing' viz. rendering them function-less. The Argonautes play a part in the formation of the RNA induced silencing Complex (RISC). Argonautes of different groups may vary in number, but have the same role(s). The work mentioned in this paper includes detection of domains present in these proteins, using the CD search server of the NCBI website.

## **2. MATERIALS AND METHODS**

### **2.1 Data Mining and Data Curation.**

The data-set comprised of Argonautes (AGO 1 – AGO 10) from *Arabidopsis thaliana*, the model plant. Protein sequences retrieved from the NCBI Protein database, which comprised of sequences from sources including protein databases and other non-protein databases as well, pertaining to annotated coding regions.

### **2.2 BLASTp Analysis.**

The target database Phytozome v10 and the eleven target species – *Arabidopsis thaliana*, *Brassica rapa*, *Manihot esculenta*, *Glycine max*, *Phaseolus vulgaris*, *Gossypium raimondii*, *Solanum tuberosum*, *Solanum lycopersicum*, *Oryza sativa*, *Sorghum bicolor*, and *Zea mays*, were exposed to subsequent BLASTp and protein Sequences, the same as was used in the CPF pathway [2].

### **2.3 Prediction and Analysis of Domains.**

Prediction and analysis of domains was done using the CD search server of the NCBI website. [1]

## **3. RESULTS AND DISCUSSION**

### **Primary Structural analysis of the Argonaute proteins: Analysis of Domains (Tables 1 to 10):**

The Conserved Domain (CD) search server of NCBI was utilized to identify the conserved domains present in Argonaute Proteins.

**Specific Conserved Domain Hits (SCDH):**

All Argonautes contain Piwi\_ago-like (cd04657), PAZ\_argonaute\_like (cd02846) and DUF1785 (pfam08699) conserved domains among the Specific Conserved Domain Hits (SCDH) other than Argonaute 1 which also contains the additional Gly-rich\_Ago1 (pfam12764) conserved domain as a SCDH.

**Non-specific Conserved Domain Hits (NCDH):**

All Argonautes contain Piwi (pfam02171), Piwi (smart00950), Piwi\_piwi-like\_Euk (cd04658), and Piwi-like (cd02826) conserved domains among the Non-specific Conserved Domain Hits (NCDH). Apart from these, there were other frequently found conserved domains like PAZ (pfam02170) – absent in Argonaute 4, PAZ (smart00949) – absent in Argonaute 8, and PAZ (cd02825) – absent in both Argonaute 2 & Argonaute 8. However, Piwi\_piwi-like\_ProArk (cd04659) was found in only four types of Argonautes, viz. AGO04, AGO06, AGO07 & AGO09 whereas PAZ\_piwi\_like (cd02845) was found in AGO03 & AGO05. UBA\_e1\_C (pfam09358) was found only in Argonaute 10.

Hence, it is proposed here that all these argonautes should interact with the corresponding consensus ncRNA, as all of them contain the two SCDH and four NCDH, all of which are required for binding to ncRNAs and are hallmark structural features of all plant Argonautes.

**Tables**

**Table 1: Conserved domains of Argonaute 1**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	pfam02171	Piwi	cl00628
non-specific	smart00950	Piwi	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	pfam02170	PAZ	cl00301
specific	pfam12764	Gly-rich_Ago1	cl15078
superfamily	cl15078	Gly-rich_Ago1 superfamily	-
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	cd02825	PAZ	cl00301
non-specific	smart00949	PAZ	cl00301
multi-dom	PLN03202	PLN03202	-

**Table 2: Conserved domains of Argonaute 2**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	smart00950	Piwi	cl00628
non-specific	pfam02171	Piwi	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628

non-specific	pfam02170	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	smart00949	PAZ	cl00301
multi-dom	PLN03202	PLN03202	-

**Table 3: Conserved domains of Argonaute 3**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	smart00950	Piwi	cl00628
non-specific	pfam02171	Piwi	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	pfam02170	PAZ	cl00301
non-specific	smart00949	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	cd02825	PAZ	cl00301
non-specific	cd02845	PAZ_piwi_like	cl00301
multi-dom	PLN03202	PLN03202	-

**Table 4: Conserved domains of Argonaute 4**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	pfam02171	Piwi	cl00628
non-specific	smart00950	Piwi	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	pfam02170	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	cd04659	Piwi_piwi-like_ProArk	cl00628
non-specific	cd02825	PAZ	cl00301
non-specific	smart00949	PAZ	cl00301
multi-dom	PLN03202	PLN03202	-

**Table 5: Conserved domains of Argonaute 5**

Hit type	Accession	Short name	Superfamily
Specific	cd04657	Piwi_ago-like	cl00628
Superfamily	cl00628	Piwi-like superfamily	-
Specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-

non-specific	smart00950	Piwi	c100628
non-specific	pfam02171	Piwi	c100628
non-specific	cd04658	Piwi_piwi-like_Euk	c100628
non-specific	cd02826	Piwi-like	c100628
non-specific	pfam02170	PAZ	c100301
Specific	pfam08699	DUF1785	c107356
superfamily	c107356	DUF1785 superfamily	-
non-specific	smart00949	PAZ	c100301
non-specific	cd02825	PAZ	c100301
non-specific	cd02845	PAZ_piwi_like	c100301
multi-dom	PLN03202	PLN03202	-

**Table 6: Conserved domains of Argonaute 6**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	c100628
superfamily	c100628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	c100301
superfamily	c100301	PAZ superfamily	-
non-specific	pfam02171	Piwi	c100628
non-specific	smart00950	Piwi	c100628
non-specific	cd04658	Piwi_piwi-like_Euk	c100628
non-specific	cd02826	Piwi-like	c100628
non-specific	pfam02170	PAZ	c100301
specific	pfam08699	DUF1785	c107356
superfamily	c107356	DUF1785 superfamily	-
non-specific	cd04659	Piwi_piwi-like_ProArk	c100628
non-specific	smart00949	PAZ	c100301
non-specific	cd02825	PAZ	c100301
multi-dom	PLN03202	PLN03202	-

**Table 7: Conserved domains of Argonaute 7**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	c100628
superfamily	c100628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	c100301
superfamily	c100301	PAZ superfamily	-
non-specific	smart00950	Piwi	c100628
non-specific	pfam02171	Piwi	c100628
non-specific	cd02826	Piwi-like	c100628
non-specific	cd04658	Piwi_piwi-like_Euk	c100628

non-specific	pfam02170	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	smart00949	PAZ	cl00301
non-specific	cd04659	Piwi_piwi-like_ProArk	cl00628
non-specific	cd02825	PAZ	cl00301
multi-dom	PLN03202	PLN03202	-

**Table 8: Conserved domains of Argonaute 8**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	smart00950	Piwi	cl00628
non-specific	pfam02171	Piwi	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	pfam02170	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
multi-dom	PLN03202	PLN03202	-

**Table 9: Conserved domains of Argonaute 9**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	smart00950	Piwi	cl00628
non-specific	pfam02171	Piwi	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	pfam02170	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	smart00949	PAZ	cl00301
non-specific	cd04659	Piwi_piwi-like_ProArk	cl00628
non-specific	cd02825	PAZ	cl00301
multi-dom	PLN03202	PLN03202	-

**Table 10: Conserved domains of Argonaute 10**

Hit type	Accession	Short name	Superfamily
specific	cd04657	Piwi_ago-like	cl00628
superfamily	cl00628	Piwi-like superfamily	-
specific	cd02846	PAZ_argonaute_like	cl00301
superfamily	cl00301	PAZ superfamily	-
non-specific	pfam02171	Piwi	cl00628
non-specific	smart00950	Piwi	cl00628
non-specific	cd04658	Piwi_piwi-like_Euk	cl00628
non-specific	cd02826	Piwi-like	cl00628
non-specific	pfam02170	PAZ	cl00301
specific	pfam08699	DUF1785	cl07356
superfamily	cl07356	DUF1785 superfamily	-
non-specific	cd02825	PAZ	cl00301
non-specific	smart00949	PAZ	cl00301
non-specific	pfam09358	UBA_e1_C	cl09615
superfamily	cl09615	UBA_e1_C superfamily	-
multi-dom	PLN03202	PLN03202	-

#### 4. CONCLUSION

In our approach, a bulk of sequence data was found in sequence databases that consists of predicted, and redundant sequences. The same queries as used in the CPF method, were used to search for protein domains. The consequent results may be used for doing structural studies of Argonaute proteins. It can be thus concluded that similar studies, using a similar work plan, can be undertaken for other domains of life, as well, to increase our knowledge about the Argonautes.

#### REFERENCES:

- [1] Marchler-Bauer A & Bryant SH: CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 2004 Jul 1; 32 (Web Server issue): W327-31.
- [2] Basu et al: “Exploring Computational Protein Fishing (CPF) to identify Argonaute Proteins from Sequenced Crop Genomes” *International Letter of Natural Sciences* 6 (2015) 27-36.