CAAS

The bioinformatic analysis of Argonaute family in Vertebrate

Speaker: Qin Qianshan

BY S2G15

(Gao zeqian、Qin Qianshan、Shui zongxia、Liu Xing)

2013-06-22

Background

Argonaute proteins are the catalytic components of the RNA-induced silencing complex (RISC), the protein complex responsible for the gene silencing phenomenon known as RNA interference (RNAi). Argonaute proteins bind different classes of small non-coding RNAs, including microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). Small RNAs guide Argonaute proteins to their specific targets through sequence complementarity, which typically leads to silencing of the target. Some of the Argonaute proteins have endonuclease activity directed against messenger RNA (mRNA) strands that display extensive complementarity to their bound small RNA, and this is known as Slicer activity .These proteins are also partially responsible for selection of the guide strand and destruction of the passenger strand of the siRNA substrate.

Acquisition of sequences

Query sequence: AGO1_Human

                  NP_036331

Method: Blastp search

Database: Swiss-Prot

Organism: Vertebrate

Filter: E<1e-5

Table1:Argonaute protein Sequences

Argonaute family member	Spceies	Protein Accession Number	Gene Accession Number
AGO1	Homo sapiens	NP_036331	NM_012199
AGO2	Bos taurus	NP_991363	NM_205794
	Homo sapiens	NP_036286	NM_012154
	Mus musculus	NP_694818	NM_153178
	Oryctolagus cuniculus	NP_001076179	NM_001082710
	Rattus norvegicus	NP_067608	NM_021597
	Xenopus (Silurana) tropicalis	NP_001004877	NM_001004877
	Xenopus laevis	NP_001086988	NM_001093519
AGO3	Bos taurus	NP_001001133	NM_001001133
	Danio rerio	NP_001153500	NM_001160028
	Gallus gallus	NP_001026071	NM_001030900
	Homo sapiens	NP_079128	NM_024852
	Mus musculus	NP_700451	NM_153402
AGO4	Gallus gallus	NP_001034365	NM_001039276
	Homo sapiens	NP_060099	NM_017629
	Mus musculus	NP_694817	NM_153177
	Xenopus laevis	NP_001089574	NM_001096105

Table1:Argonaute protein Sequences

Argonaute family member	species	Protein Accession Number	Gene Accession Number
PIWI1	Danio rerio	NP_899181	NM_183338
	Gallus gallus	NP_001092322	NM_001098852
	Homo sapiens	NP_004755	NM_004764
	Mus musculus	NP_067286	NM_021311
PIWI2	Danio rerio	NP_001073668	NM_001080199
	Homo sapiens	NP_001129193	NM_001135721
	Mus musculus	NP_067283	NM_021308
	Oncorhynchus mykiss	NP_001117714	NM_001124242
	Xenopus (Silurana) tropicalis	NP_001106470	NM_001112999
PIWI3	Homo sapiens	NP_001008496	NM_001008496
PIWI4	Homo sapiens	NP_689644	NM_152431
	Mus musculus	NP_808573	NM_177905
	Rattus norvegicus	XP_002727047	XM_002727001

# *Sequence alignment and phylogenetic analysis*

The sequence data set included a total of 30 sequences from 10 species.

The sequence of Argonaute family proteins were aligned by ClustalW and the resulting alignment was manually optimized by Jalview.

Tree reconstructions were done by Neighbor-joining(NJ)method with p-distance model(500 replicates).

The phylogenetic trees were displayed and edited using MEGA5.1.

FIG1:Phylogenetic relationships of protein sequence with the Argonaute family in Vertebrate
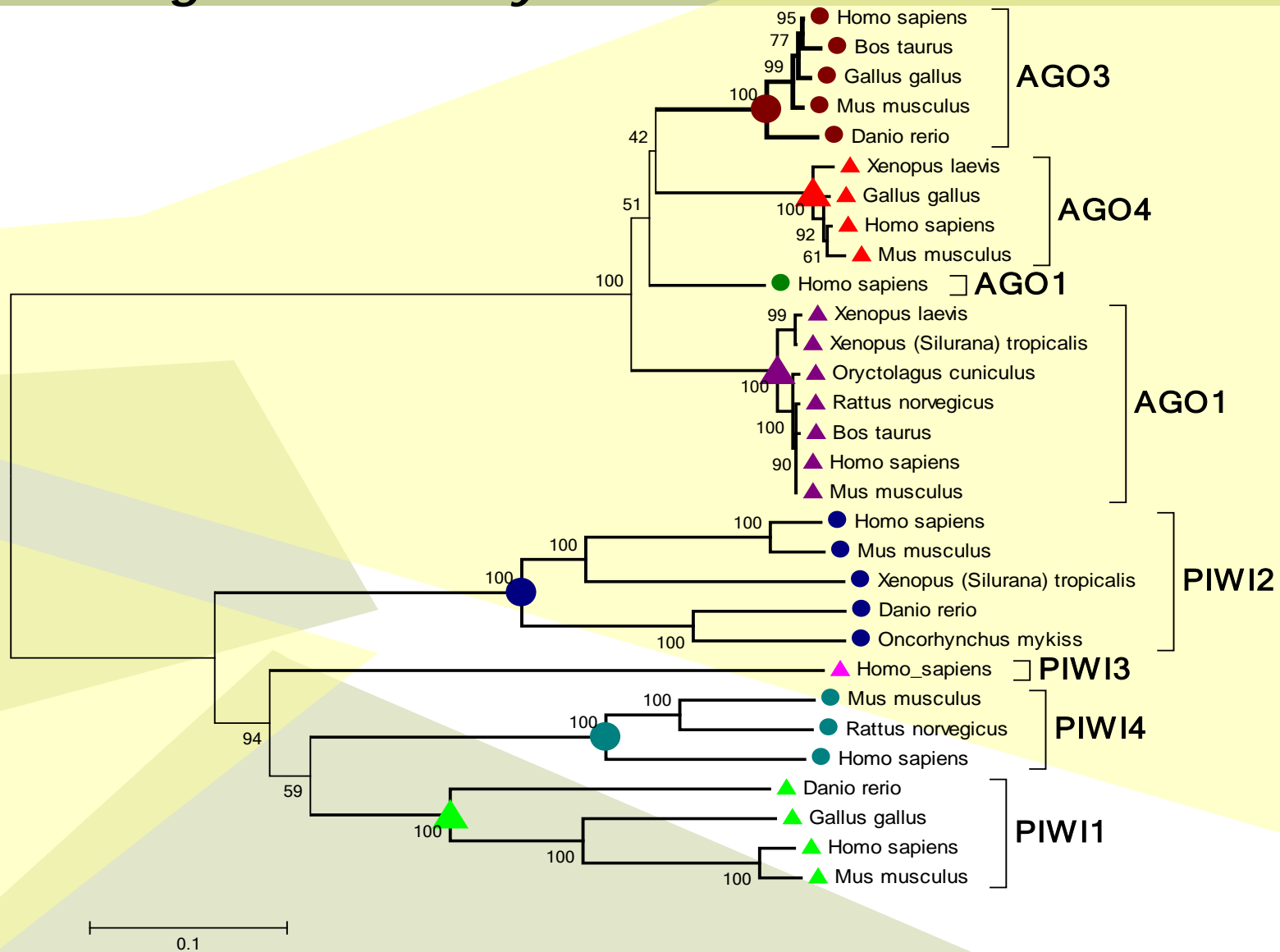
Synonymous codon usage analyses

The Argonaute genes coding sequence were retrieved at the NCBI.

Codon usage bias was estimated by the Effective Number of Codons (ENC) and the proportion of G and C in the third codon position (G/C 3rd).For ENC, lower values indicate stronger synonymous codon usage bias. These measures were calculated for all genes using the CodonW program.

Table2:Mean values of GC%,G/C3rd% and ENC of the Argonaute genes

Gene	GC%	G/C 3rd%	ENC
AGO1	0.547	0.654	49.14
AGO2	0.537	0.630	51.67
AGO3	0.483	0.455	54.39
AGO4	0.483	0.459	54.37
PIWI1	0.481	0.499	54.40
PIWI2	0.504	0.533	53.19
PIWI3	0.437	0.406	54.91
PIWI4	0.493	0.554	53.36

Argonaute family gene structure

N-terminal   C-terminal

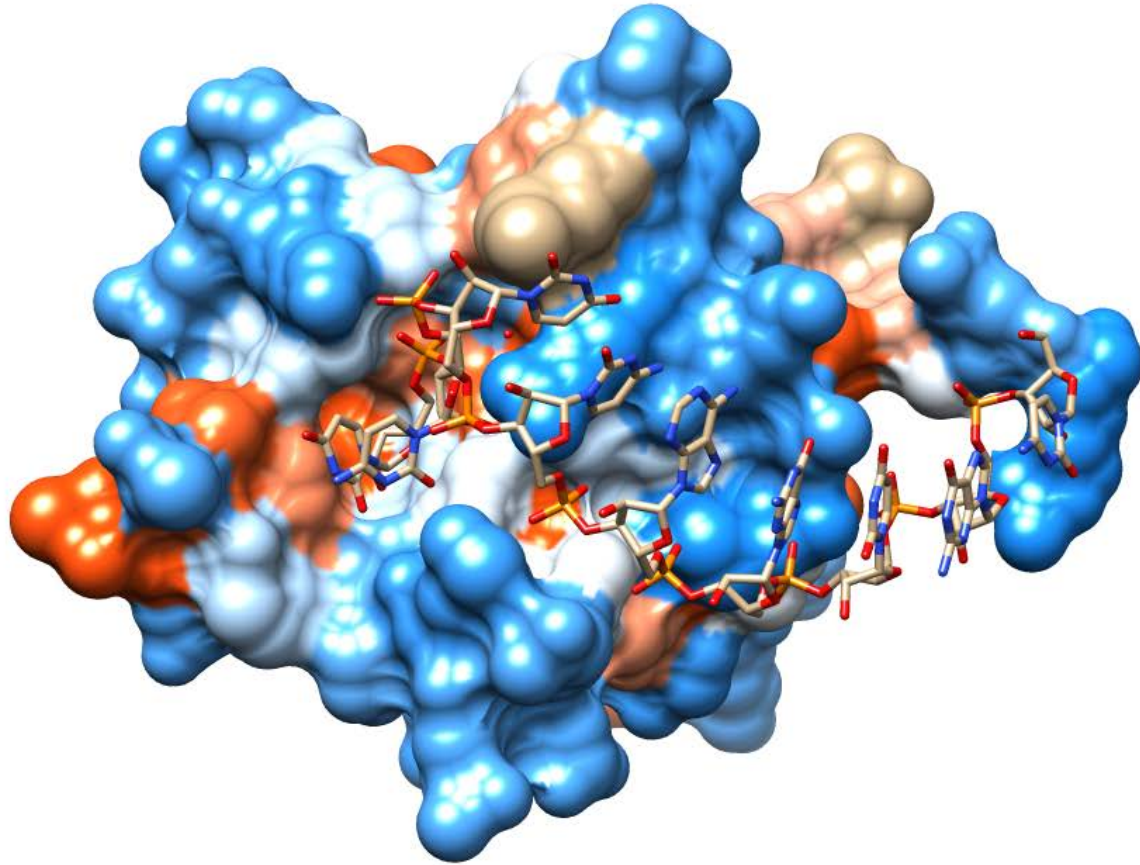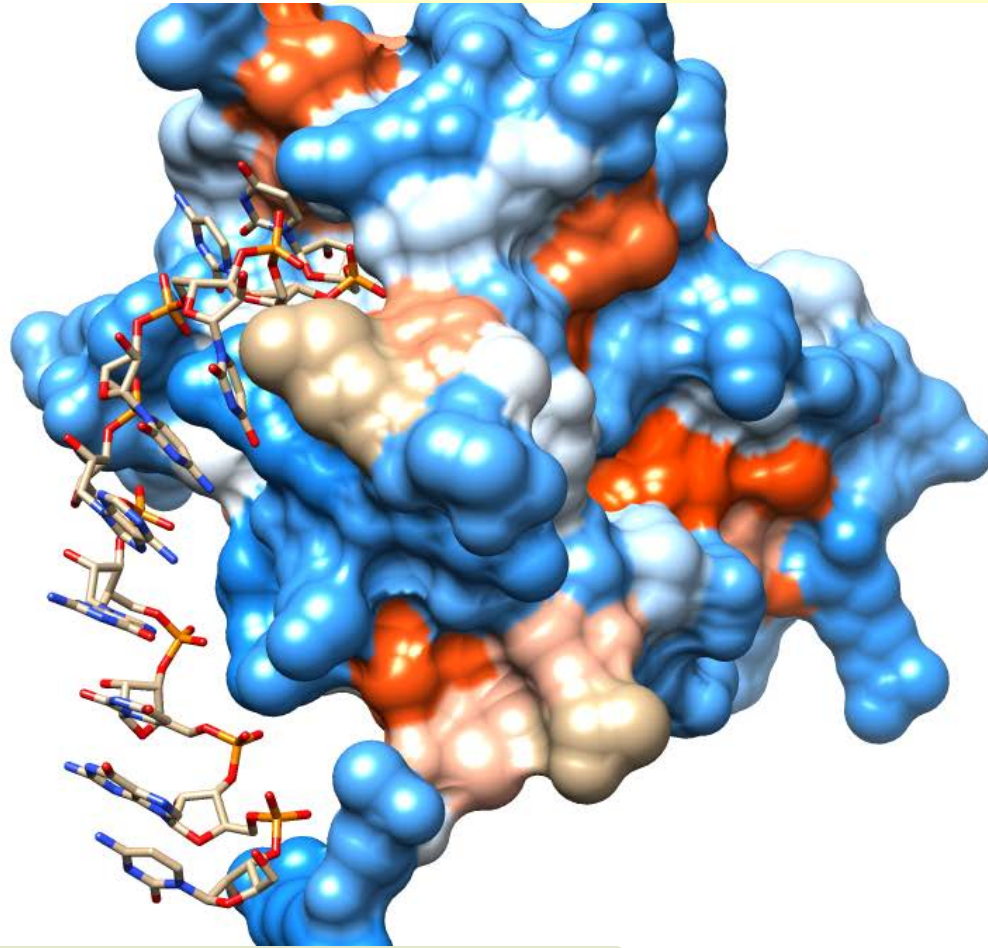| PAZ Domain | | PIWI   Domain | |

Argonaute family genes can be defined by four sections:an N-terminal,PAZ domain,PIWI domain and C-terminal,in which PAZ ,PIWI belong to main functional domains.
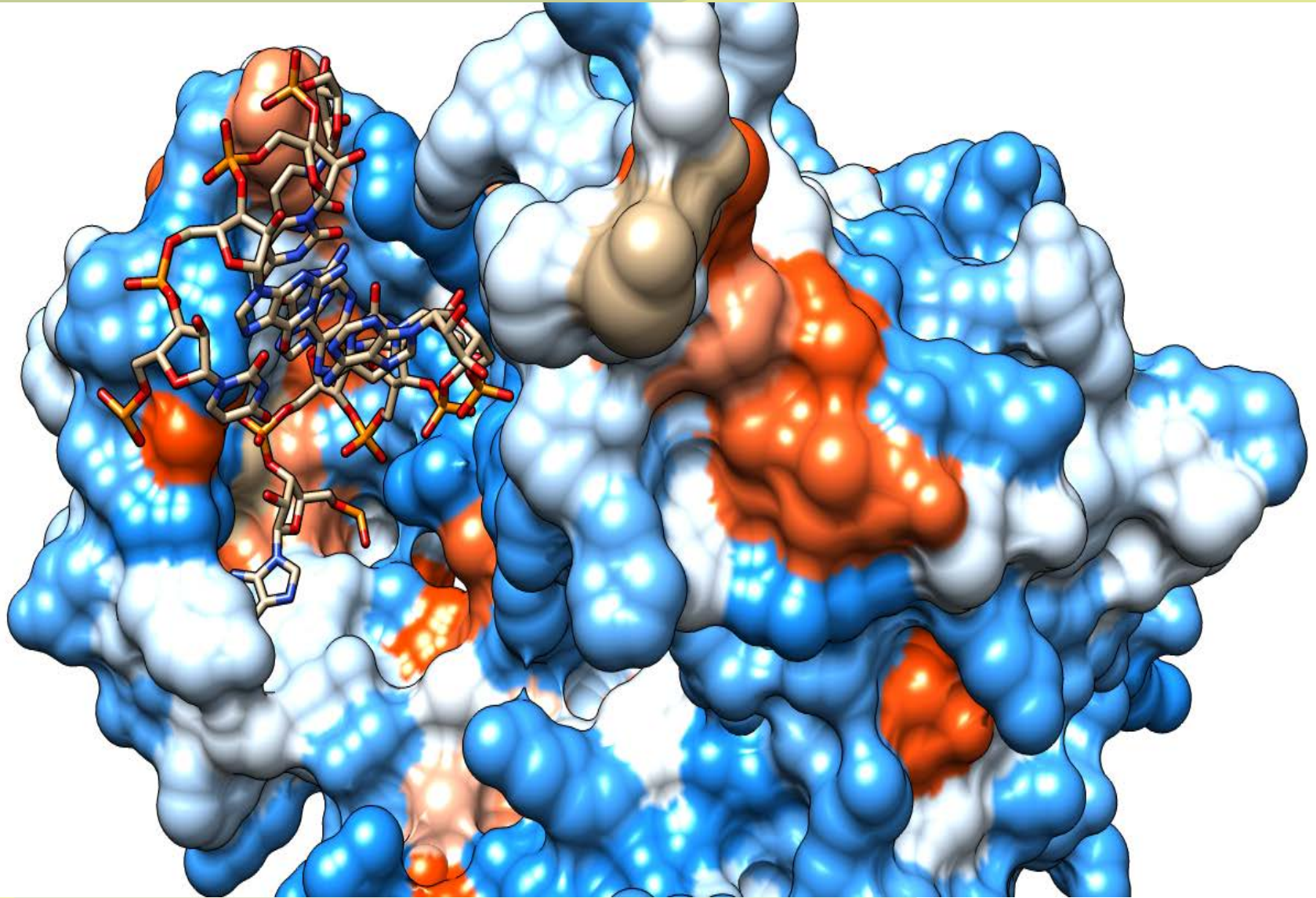
The representative of Argonaute family--Human AGO1

Number of aminoacids:834

Molecular weight:94438.5

**Theoretical PI:9.18**

```
Amino acid composition:
Ala (A)    35        4.2%
Arg (R)    44        5.3%
Asn (N)    42        5.0%
Asp (D)    38        4.6%
Cys (C)    17        2.0%
Gln (Q)    45        5.4%
Glu (E)    35        4.2%
Gly (G)    53        6.4%
His (H)    19        2.3%
Ile (I)    53        6.4%
Leu (L)    70        8.4%
Lys (K)    52        6.2%
Met (M)    23        2.8%
Phe (F)    39        4.7%
Pro (P)    46        5.5%
Ser (S)    76        9.1%
Thr (T)    38        4.6%
Trp (W)     9        1.1%
Tyr (Y)    37        4.4%
Val (V)    63        7.6%
Pyl (O)     0        0.0%
Sec (U)     0        0.0%
```

Physical and chemical parameters of AGO1 Protein sequence was predicted by ProtParam
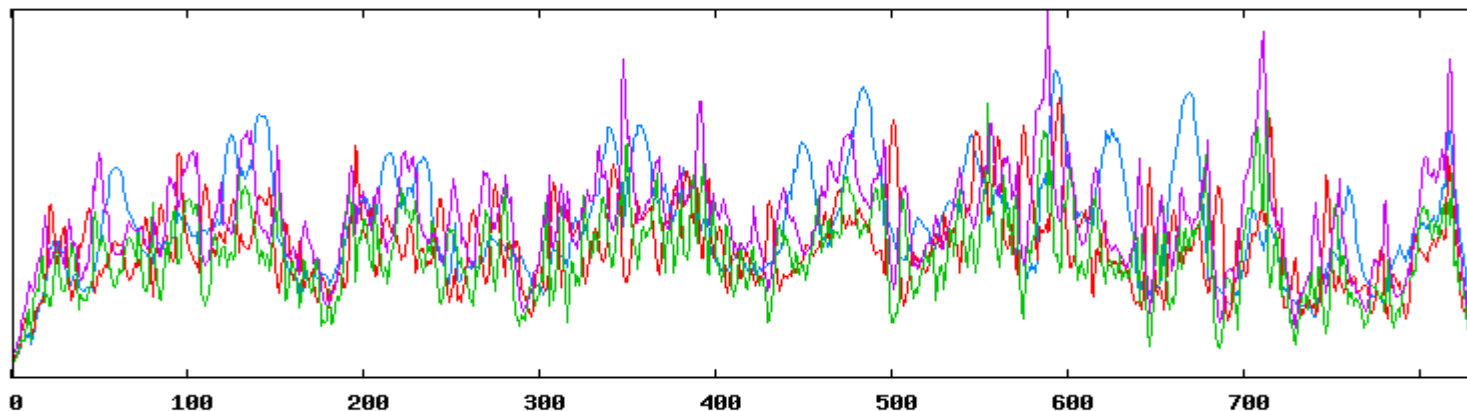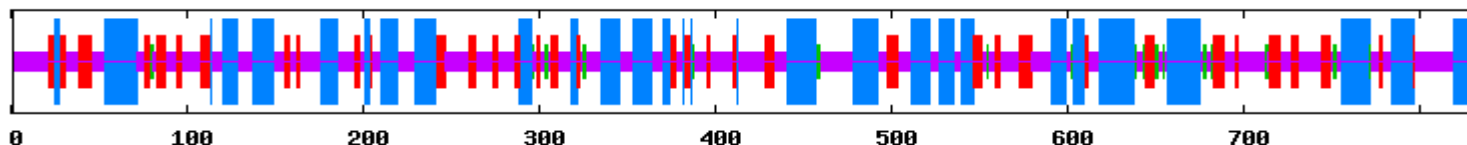
The representative of Argonaute family--Human AGO1
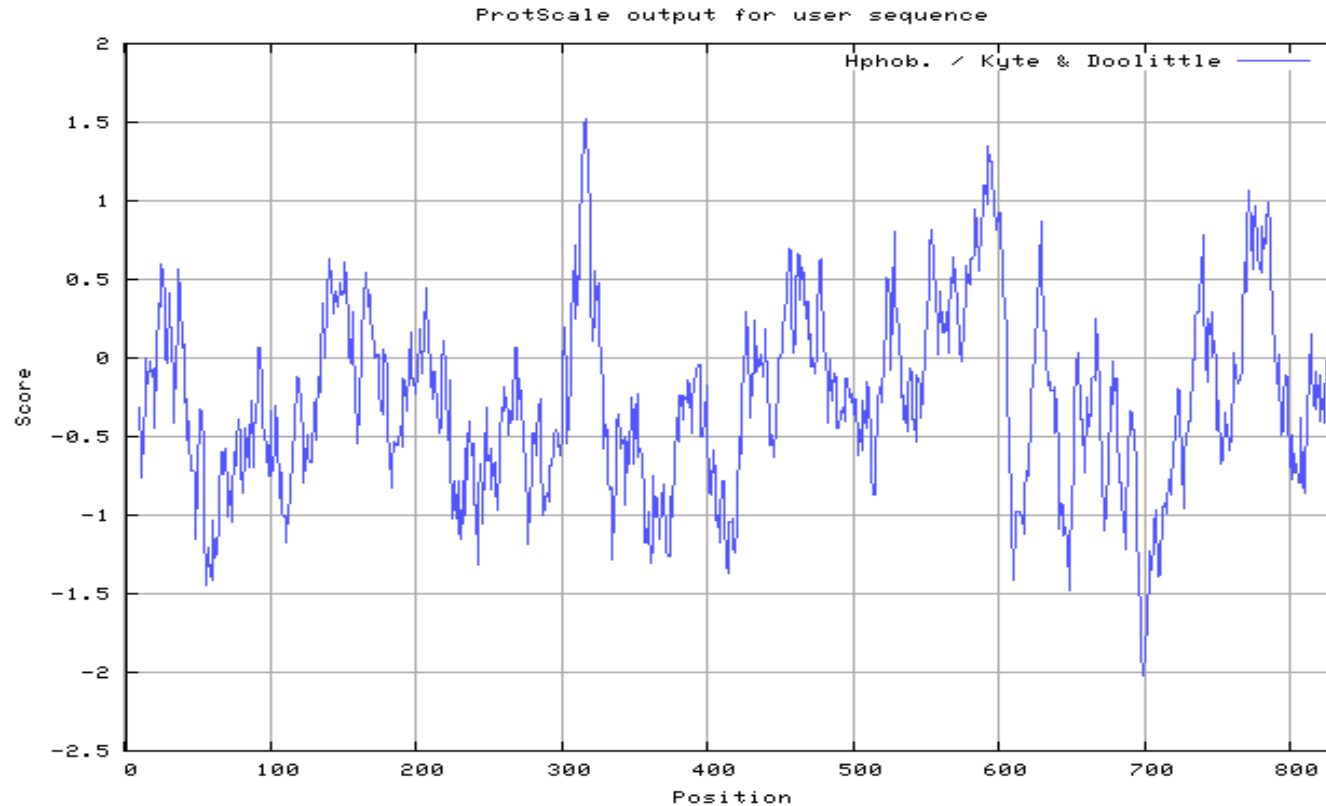
The Hydropathicity of AGO1 was predicted by ProtScale

The representative of Argonaute family--Human AGO1

```
SOPMA :
    Alpha helix       (Hh) :     280  is   33.57%
    3_10  helix       (Gg) :       0  is    0.00%
    Pi helix          (Ii) :       0  is    0.00%
    Beta bridge       (Bb) :       0  is    0.00%
    Extended strand   (Ee) :     148  is   17.75%
    Beta turn         (Tt) :      28  is    3.36%
    Bend region       (Ss) :       0  is    0.00%
    Random coil       (Cc) :     378  is   45.32%
    Ambigous states   (?)  :       0  is    0.00%
    Other states           :       0  is    0.00%
```


Secondary structure prediction of AGO1 with the SOPMA program

The representative of Argonaute family--Human AGO1

SignalP-4.1 prediction (euk networks): Sequence
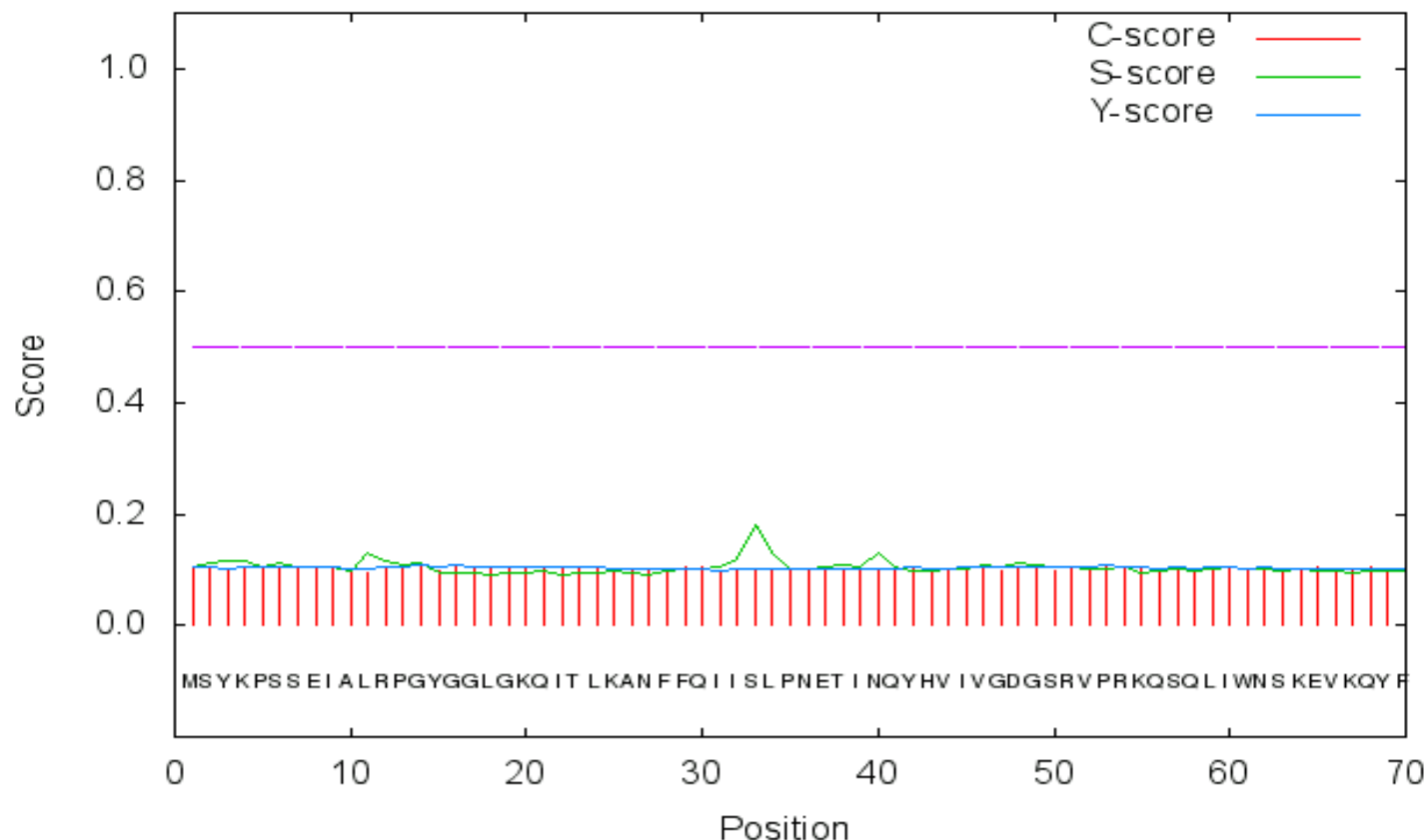
Signal Peptide prediction of AGO1 using SignalP4.1 program

The representative of Argonaute family--Human AGO1

```
### targetp v1.1 prediction results ################################
Number of query sequences:  1
Cleavage site predictions not included.
Using NON-PLANT networks.


Name                      Len          mTP     SP   other  Loc   RC
----------------------------------------------------------------------
Sequence                  834          0.109  0.089  0.851   _     2
----------------------------------------------------------------------
cutoff                                 0.000  0.000  0.000
```

Cellular location prediction of AGO1 using TargetP1.1 program

The representative of Argonaute family--Human AGO1
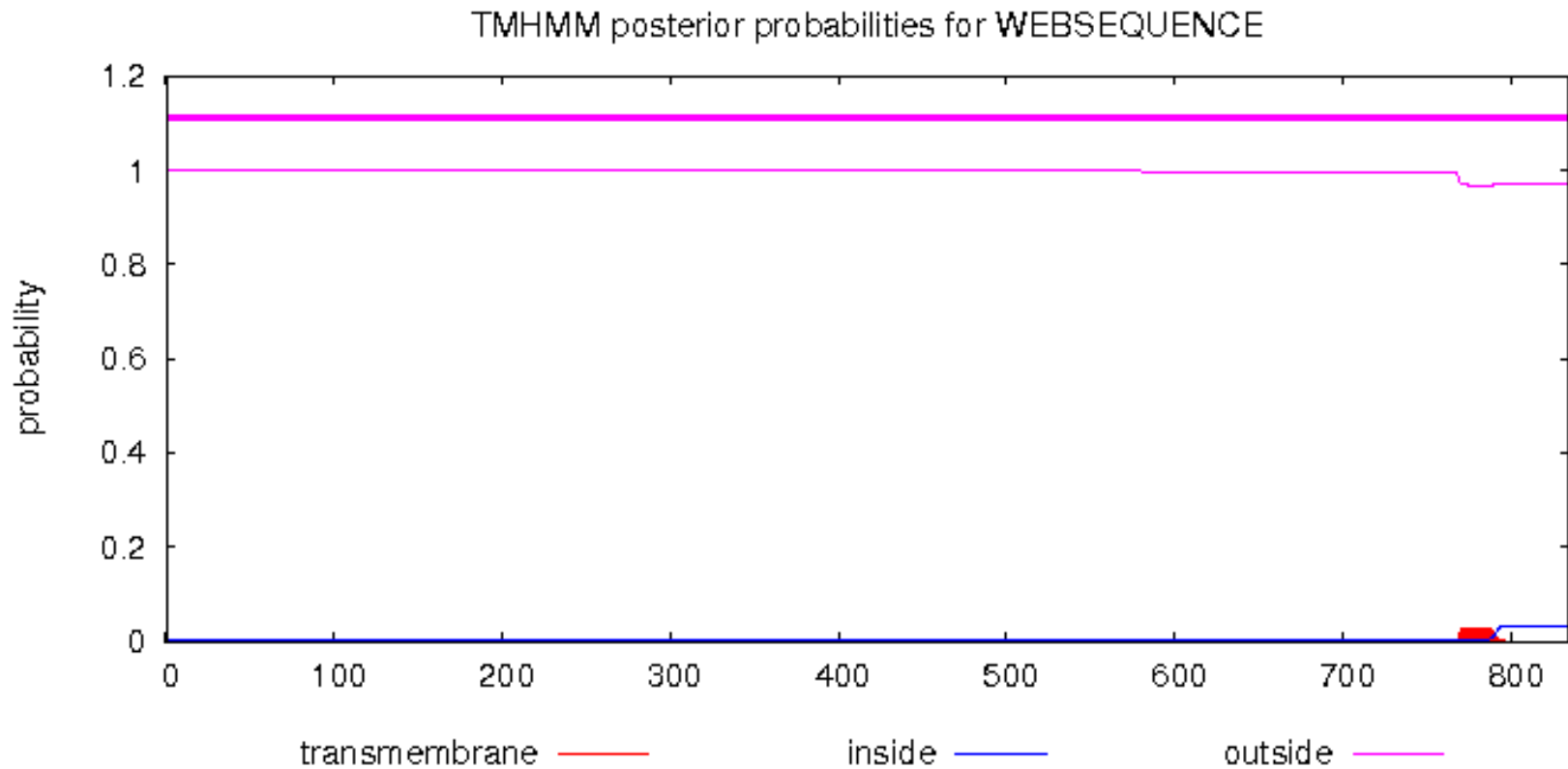
Transmembrane helix prediction of AGO1 using TMHMM program

The representative of Argonaute family--Human AGO1

S-2	PKC	0.84	S-203	cdc2	0.52	Y-582	INSR	0.50
T-23	PKC	0.73	S-210	PKA	0.76	S-592	PKA	0.50
S-33	PKA	0.74	S-222	cdc2	0.52	S-599	PKC	0.72
Y-42	INSR	0.53	S-236	PKC	0.74	Y-607	EGFR	0.57
S-50	cdc2	0.53	S-236	PKA	0.67	S-613	DNAPK	0.62
S-57	RSK	0.54	S-267	PKC	0.74	T-639	cdc2	0.52
S-57	ATM	0.57	S-267	cdc2	0.54	S-660	CKII	0.51
S-57	PKA	0.79	S-270	cdc2	0.56	S-667	CKII	0.51
S-57	PKG	0.57	S-288	CKII	0.52	S-667	DNAPK	0.54
S-72	PKC	0.63	S-288	PKA	0.60	S-667	cdc2	0.51
T-95	PKC	0.86	S-288	PKG	0.52	T-688	PKC	0.78
S-104	CKII	0.59	T-331	PKC	0.84	T-689	PKC	0.85
S-116	PKC	0.74	Y-371	INSR	0.54	T-715	PKC	0.61
S-117	PKC	0.62	T-373	DNAPK	0.57	T-715	cdc2	0.50
S-117	PKA	0.64	T-435	PKC	0.76	S-741	cdc2	0.52
S-117	cdc2	0.51	S-436	PKC	0.55	T-775	PKC	0.71
T-123	PKC	0.65	T-437	PKC	0.86	S-776	RSK	0.56
S-125	DNAPK	0.55	T-451	PKC	0.73	S-776	CaM-II	0.51
S-125	ATM	0.58	T-451	cdc2	0.51	S-776	PKA	0.63
S-130	cdc2	0.51	S-478	CKII	0.51	T-804	CKII	0.57
S-140	PKC	0.52	S-478	RSK	0.55	T-808	CKII	0.60
S-140	cdc2	0.54	S-478	PKA	0.73	S-809	CKII	0.50
S-141	cdc2	0.54	T-522	PKC	0.50	S-809	CKI	0.56
S-154	PKA	0.75	S-528	DNAPK	0.57	S-812	CKII	0.53
S-172	CKI	0.60	S-528	ATM	0.59	S-824	PKC	0.68
Y-184	EGFR	0.50	S-539	PKC	0.91	S-825	PKC	0.86
S-186	PKC	0.84	S-539	PKA	0.58	T-829	PKC	0.76

Kinase-specific phosphorylation site prediction of AGO1 using NetPhosK program

Rerference

Cenik ES, Zamore PD. Argonaute proteins [J]. Current Biology,2011, 21 (12): R446–9.

Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe [J]. Nat. Rev. Genet ,2009,10 (2): 94–108. ^

Tolia, N. H. Slicer and the argonautes [J]. Nature Chem. Biol,2007,3, 36–43 (2007).

Rand TA, Petersen S, Du F, Wang X et al. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation [J]. Cell,2005,123 (4): 621–9.

Wei K F, Wu L J, Chen J et al. Structual evolution and functional diversification analysis of Argonaute protein [J].Journal of Cellular Biochenistry,2012,113:2576-2585.

Thank you

LOGO