

**piRNA DIVERSITY ACROSS MODEL SPECIES AND HUMANS REVEALED THROUGH KOHONEN MAPS****VINEET VISHAL<sup>1\*</sup>, PANKAJ K. SINGH<sup>1</sup>, ABHISEK R. BERA<sup>1</sup>,  
PROTIP BASU<sup>1</sup>, SAYAK GANGULI<sup>1</sup> AND ABHIJIT DATTA<sup>2</sup>**<sup>1</sup>*DBT Centre for Bioinformatics, Presidency University, Kolkata*<sup>2</sup>*Department of Botany, Jhargram Raj College, West Bengal*

**Transposons (Transposable Elements or TEs) are a constant menace for human genome integrity as various diseases like SCID, DMD, Haemophilias, have been attributed to retrotransposon activity etc. Piwi-interacting RNA (piRNA) plays a pivotal role in RNA interference and their mechanism involved silencing selfish genetic material (Transposonable elements) in the germline and there by preserving genome integrity between generation. This study attempts to gain insights on the conserved features of piRNA using an artificial intelligence technique – Self Organizing Map (Kohonen Map).**

The Piwi Interacting RNAs (piRNAs) are the largest class of small non coding RNA with a diverse function across invertebrate and vertebrate, but still has not been reported in plants [1]. piRNAs were first reported in mouse testes by four groups (Aravin *et. al.* 2006; Girard *et. al.* 2006; Grivna *et. al.* 2006; Watanabe *et. al.* 2006;) derived from genomic clusters [2]. Owing to 24-31 nt long expressed only in eukaryotic germline but recent studies reveals that it was also found in somatic cell, human cancer cell and control gene expression more broadly [3]. The piRNA interact with PIWI protein to form protein-RNA complex which are involved in gene silencing by both epigenetic and post transcriptional pathway, a mechanism called RNA Interference pathway [4].

Among other small non coding RNA, piRNA was considered as enigmatic in nature due to their importance in- protecting genome integrity from invasion by genomic parasite-SELFISH GENETIC MATERIAL such as retrotransposons and repetitive sequence as well as viruses by silencing them [5]. It establishes and maintains heterochromatin structure and regulates endogenous gene expression [6].

The piRNA also plays significant function in epigenetic, required for germ cell and stem cell development and spermatogenesis [7]. The animals which lack piRNA function show defect in gametogenesis and exhibit sterility [8].

The Kohonen map commonly known as SELF ORGANIZING MAP (SOMs) is a clustering technique based on artificial neural network, which uses unsupervised learning [9-10]. The SOM algorithms are broadly used to cluster huge amount of data depending upon similarity

---

\*Address for correspondence : Dr. Vineet Vishal, E-mail: vineet.vishal73@gmail.com

[11-12] .This study attempts to gain insights on the conserved features of piRNA using an artificial intelligence technique–Self Organizing Map (Kohonen Map).

## MATERIALS AND METHODS

### Data Mining and Curation

piRNA sequences of *Homo sapiens*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Mus musculus* were retrieved and were curated to remove redundancies. The test set was a curated dataset of ~ 38000 sequences that were obtained from available resources. Finally the test set was used to generate the map.

### Generation of Self organizing Maps

The training of the SOM was initially done with an assemblage of piRNA sequences of diverse origin map and the SOM's were generated as described in Ganguli and Datta 2014.

## RESULTS AND DISCUSSION

The self organizing map of different organisms were generated a cluster based on similarity is displayed by using different colour pattern. In case of *C.elegans* self organizing map generated 1 clusters contain 14.7 % of gene with coverage area of 46.1% in the map. *D.melanogaster* SOM 4 clusters contain 14.4 % of gene with coverage area of 31.6% in the map. While in *M .musculus* SOM 4 clusters contain 8.9 % of gene with coverage area of 7.0% in the map. The following diagram (Fig.1) summarizes the results obtained. Most significantly no clusters were formed in case human piRNA which shows the lack of similar sequence entities in the entire dataset. But when we add sequence of all organisms the self organizing map 4 clusters contain 6.9 % of gene with coverage area of 6.2% in the map (Fig1).

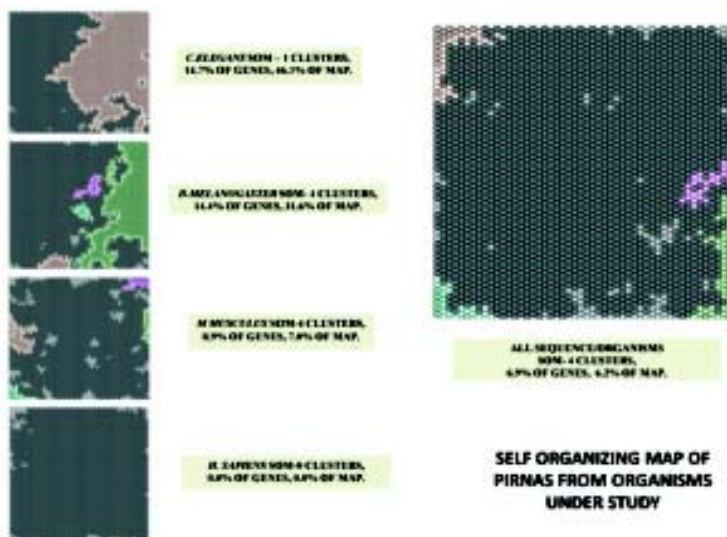
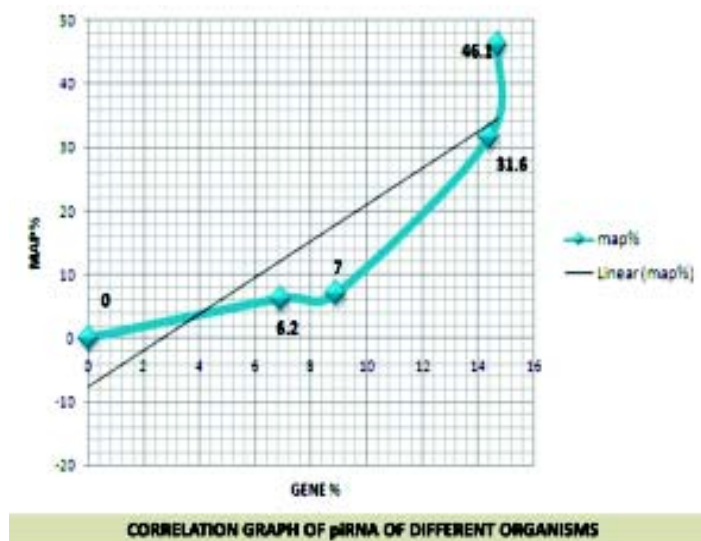


Fig 1: The various Self Organizing Maps obtained for analysis.



**Fig.2 Graphical Representation of the estimation of Correlation and Regression.**

Correlation between percentage of gene and percentage of map was estimated and was found to be **0.8803851** (> the table value of 0.811 at 0.05 probability level) which shows both are correlated

### CONCLUSIONS

piRNAs have been implicated in the host germline pathways specific to development. Thus inherently they are diverse both in their function and occurrence. To add to this their ability to silence transposon activity in the genome, transforms them into a dynamic ensemble, ready to act in host defense and maintain genome integrity. Through evolution organisms have survived adaptive pressures, which has led to the complexities in genome form and function – the SOM clustering obtained above is a clear indication of the diverse nature of piRNAs in various levels of hierarchical genome organization.

### ACKNOWLEDGEMENT

The authors acknowledge the contributions of the DBT RNAiTask Force funded project entitled “Development of RNAiIDB” for the support.

### REFERENCES

- Ganguli S. and Datta A. Phylogenetic Trends of Plant Glutathione peroxidases Revealed by Kohonen Maps (SOM's). (2014): Annual Research & Review in Biology; **4** (24): 3810-3815.
- Halic M., and Moazed D. Transposon Silencing by piRNAs. (2009): Cell; **138**, 1058-1060.
- Hartig V. Julia, Tomari Y. and Förstemann K. piRNAs--the ancient hunters of genome invaders. (2007): Genes Dev; **21**, 1707-1713.

PIRNA DIVERSITY ACROSS MODEL SPECIES AND HUMANS REVEALED THROUGH KOHONEN MAPS

Humberto J F., Holger H., Xavier G. del M., August V., Sara L., Clara M., Alberto V., and Manel S. Epigenetic loss of the PIWI/piRNA machinery in human testicular tumorigenesis. (2014): *Epigenetics*; **9**: 1: 113–118.

Ishizu H, Siomi H, Siomi M.C. Biology of PIWI interacting RNAs: new insights into biogenesis and function inside and outside of germlines. (2012): *Genes Dev*; **26** 2361-2373.

Kim N. V. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. (2006): *Genes Dev*; **20**: 1993-1997.

Kim N.V. Small RNA: Classification, Biogenesis, and Function. (2005): *Mol. Cells*; **19**: 1-15.

Kohonen, T., Oja, E., Simula, O., Visa, A. and Kangas, J. (1996): Engineering applications of the self-organizing map. (1996) *Proc. IEEE*; **84**, 1358–1384.

Kohonen T. Essentials of the self-organizing map. (2012): *Neural Networks*; **37**: 52–65.

Mikiko C. Siomi, Kaoru S., Dubravka P. and Alexei A. Aravin. PIWI-interacting small RNAs: the vanguard of genome defence. (2011): *Nature*; **12**, 246-258.

Paloma M Guzzardo, Felix Muerdter and Gregory J Hannon. (2013): The piRNA pathway in flies: highlights and future directions. *Genes Dev*; **23**, 1–9.

Supek F., and ek Vlahovic K. (2004): INCA: synonymous codon usage analysis and clustering by means of self-organizing map. *Bioinformatics*; **20**, 2329–2330.