

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Evolutionary Analysis of Oct4 Orthologs

Aprajita

Stem Cell Engineering, Barkatullah University, Bhopal (M.P), India.

ABSTRACT: Oct4 is a transcription factor which plays an important role in maintaining Stem Cell pluripotency. It is also one of the four reprogramming factors used in production of induced pluripotent stem cells. As stem cell science is developing, the scientists have found many organisms expressing Oct4 gene along with human beings. The present study shows the evolutionary aspects of Oct4 gene found in different organism and what percentage of similarity these orthologs share. Humans share great Oct4 gene similarity with non-primates as compared to fishes and birds. By defining variations in Oct4 gene, we will be able to study the POU5F1 gene thoroughly and will be to correlate the regenerative capacity of different organism with humans.

KEYWORDS: homolog, genealogy pluripotent, alternative splicing tetrapods, differentiation.

I. INTRODUCTION

Oct4 or octamer binding transcription factor is a highly conserved gene. Usually Oct4 is state as protein that in humans is encoded by POU5f gene. **Oct4** is a gatekeeper for ES cell pluripotency. Pluripotent embryonic stem cell identity is governed by a core of transcription factors involving **Oct4** that acts as a key regulator of pluripotent cells across mammalian species(h) Pluripotency, is defined as the ability of a cell to generate all of the cell types of an organism. A group of transcription factors is essential for the establishment and maintenance of the pluripotent state (Shi G et.al).

Oct4 being the member of the POU transcription factor family,. Oct4 becomes silenced upon differentiation. Interestingly, the precise expression level of Oct4 determines the fate of embryonic stem cells.Oct-4 expression must be closely regulated; too much or too little will cause differentiation of the cells.

The expressed Oct-4 is a member of the octamer transcription factor family, which binds the octameric (8-unit) DNA nucleotide sequence "ATTTGCAT".

II. ORTHOLOGS

Orthologs are genes in different species that evolved from a common ancestral gene by speciation. They retain the same function in the course of evolution . There are five major orthologs of Oct4 gene . Identification of orthologs accomplishes two goals: delineating the genealogy of genes to investigate the forces and mechanisms of evolutionary process, and creating groups of genes with the same biological functions. (Gang Fang et.al). To analyze these evolutionary forces in this paper we will study about these five orthologs of different class of Chordata phylum.

Species	Gene ID	Gene name
<i>Mus musculus</i> (mouse)	18999	Pou5f1 POU domain, class 5, transcription factor 1
<i>Homo sapiens</i> (human)	5460	POU5F1 POU class 5 homeobox 1
<i>Rattus norvegicus</i> (rat)	294562	Pou5f1 POU class 5 homeobox 1
<i>Danio rerio</i> (zebrafish)	303333	pou5f3 POU domain, class 5, transcription factor 3
<i>Gallus gallus</i> (chicken)	427781	Pou5f3 POU domain class 5 transcription factor 3 *updated data on 20 June 2018

Figure: Above table lists Gene ID and Name of orthologs (* data from NCBI Database)

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

III. METHODOLOGY

In order to find out the phylogenetic relationship between Orthologs of Oct4 . I performed multiple sequence alignment for Oct4 gene Orthologs . Multiple Sequence alignment characterizes gene families and identifies the shared regions of homology. These orthologous sequences have possibility to be derived from some common ancestral sequence indicating that the sequence could have arisen by convergent evolution towards common structure or function.

After the discovery of Oct4 ortholog in birds , the evolutionary aspect of Oct4 expression has become complex. In order to analyze the evolutionary variation between these Orthologs , I aligned all the genetic sequence using alignment software “ MUSCLE” via web service of European Bioinformatics Institute , although the genetic data was extracted from National Center for Biotechnology Information Search Database. MUSCLE is a Multiple Sequence Comparison by Log Expectation and is significantly a faster software than Clustal which is a widely used software for protein sequences alignment .

After performing MSA , as a result I got Percentage Identity Matrix , this matrix will help me in analyzing the percentage of genetic similarity between all these Orthologs which will in turn put on some light on the evolutionary relationship between them.

The Orthologs sequences I used to study are those widely accepted by scientist as Orthologs and have data in NCBI database, for *Gallus gallus*, although it is a recent research stating that it has PouV gene that encodes for Oct4 but I still mentioned it in order to have a sharp analysis.

IV. RESULT ANALYSIS

MUSCLE

[Tools \(/Tools/\)](#) > [Multiple Sequence Alignment \(/Tools/msa\)](#) > MUSCLE

Results for job muscle-l20180718-164728-0951-33759007-p1m

```
#
#
# Percent Identity Matrix - created by Clustal2.1
#
#
1: Danio      100.00   44.06   47.31   52.21   51.25
2: Gallus    44.06  100.00   55.38   57.56   57.97
3: Homo      47.31   55.38  100.00   69.17   68.34
4: Rattus    52.21   57.56   69.17  100.00   84.59
5: Mus       51.25   57.97   68.34   84.59  100.00
```

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

Results for job simple_phylogeny-I20180718-164958-0169-95252994-p1m

Phylogram

Branch length: Cladogram Real

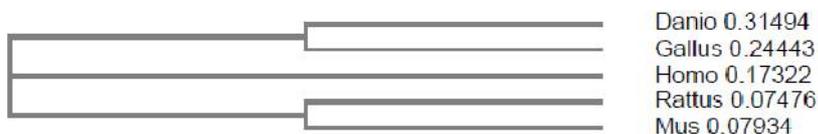


Figure1: Obtained data from Multiple Sequence Alignment

The obtained matrix is symmetric ($m(i,j)=m(i,j)$). Following points can be obtained from resulting matrix.

- *Danio rerio* 's genetic sequence is most identical to *Rattus norvegicus* Oct4 ortholog (52.21%) .
- *Gallus gallus* shows greater similarity with *Rattus norvegicus* (57.56%) and *Mus Musculus* (57.97%) and along with 2% difference with *Homo sapiens* .
- The similarity between *Danio rerio* And *Gallus gallus* is less (44.06%) but in the obtained Cladogram they have been grouped under the same node which is slight contradictory.
- *Homo sapiens* share similarity of *Rattus norvegicus* and *Mus musculus* and less similarity with *Danio rerio*.

V. DISCUSSION

After observing the similarity percentage, we can say that may be there was a POU5F gene (OCT4) gene present in clad Euteleostomi , because we can see that animals belonging to of its largest subgroup are having Oct4 genes as orthologs . All the above mentioned sequences are identified as Orthologs, this clarifies the fact that they have been evolved from a common ancestral gene. The aim of this study was to locate the probable ancestral clad for Orthologs and to state the relationship .

For fishes , fishes of clad Actinopterygii could be the root node for Oct4 gene i,e, this gene could have been present in Actinopterygii clad organisms as mid blastula embryo culture of *Oryzias latipes* also shows the features of mouse embryonic stem cells in in- vitro and *Danio rerio* belongs to the same class, but till now there is no identification of oct4 has been made by scientist in *Oryzias latipes*. *Oryzias latipes* also has Actinopterygii like *Danio rerio*.

For humans and mouse or rats , may be the gene expression made his way from clad of extant members "Eutheria" that got diverged in early Cretaceous or Late Jurassic . The human OCT4 gene encodes three distinct transcripts due to alternative splicing : oct4A , oct4b AND oct4b1, in all these three variants , OCT4A is highly expressed in the nucleus of compacted embryos and blastocyst and functions as a main regulator in maintaining the pluripotency and self renewal capacities of ES Cells and in mice , Oct 4 encodes at least 2 variants' designated Oct4A and Oct4B (Lei Liu et. al).

As for birds , the recent research says only about *Gallus gallus* expressing Oct4 gene that is similar to human gene . After alignment we can say that the identified sequence in gallus is very similar to mammals, 55.38% to humans and 58% to Murinae class.

The Cladogram obtained shows a bit different relationship between aves and fishes , it connects them to a same inter node but according to the evolutionary maps provided by various institutes and databanks, the birds and aves belong to two different clad, although their super clad is same with humans and mice that is " Euteleosomi" which further divides in two groups of bony fishes and tetrapods .(The cladogram depends on the model of evolution used by software.) In several articles the clad of birds is separated from phylum chordata and is known as Ornithurae which is stated ancestral to aves . As the gallus gene shows a noticeable similarity with *rerio* , more research is needed to define the perfect relationship between clades "Archelosauria and Ornithurae".

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

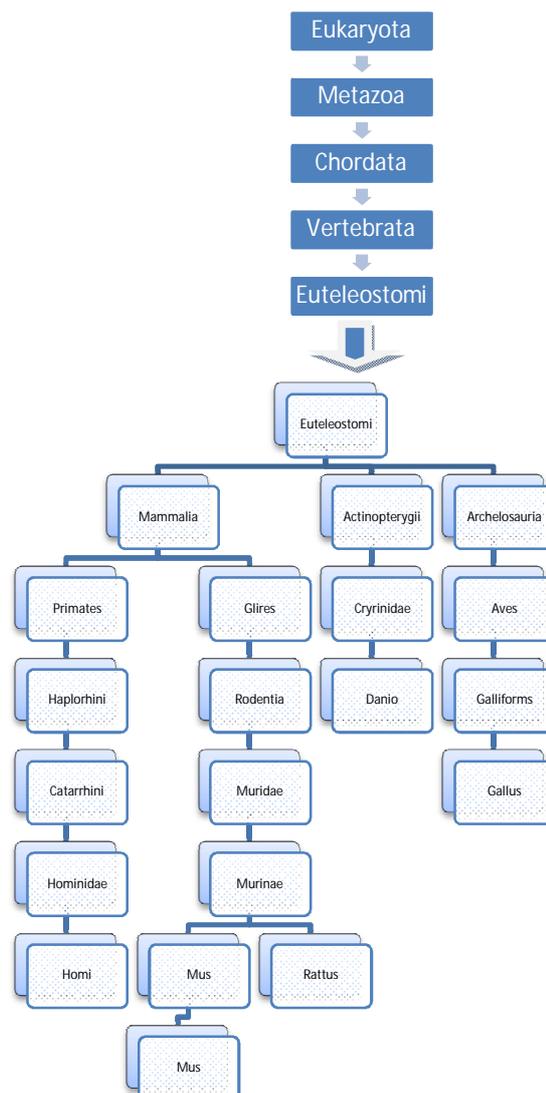
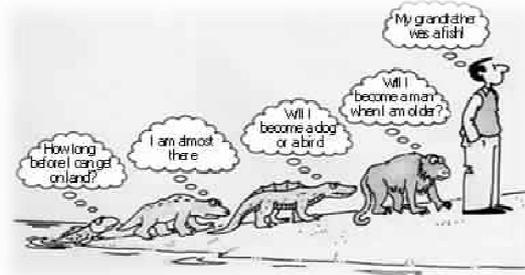


Figure2: Evolutionary Relationship among all five organisms having OCT4 Orthologs

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

- As the extracted tables (below) describe the similar pattern of encoding of Oct 4gene in stated organisms but the percentage variation of sequences highlights the fact that there were some changes in expression mechanism of gene , position of gene o chromosome throughout the evolutionary process.
- And if we correlate Oct4 expression with the regenerative capability of organism, we can give some amount of proof to the statement that through the years of evolution , organisms have developed the mechanism and metabolism of their body according to adaption processes.

Humans	Fish	Bird	Mouse
: encodes a transcription factor containing a POU domain which plays a major role in stem cell pluripotency	Transcriptional regulation in pluripotent stem cell	Studies indicate that oct4 and Nanog genes were involved in characteristics and maintenance of chicken pluripotency	Transcription factor bind to DNA motif and regulate cell differentiation pathway ;, encodes protein ;, helps in stem cell pluripotency

Figure3 : Table comparing gene expression of oct4 orthologs

<u>Organism With Orthologs</u>	<u>Regeneration Capability</u>
Zebra fish	Heart has endogenous regenerative capacity , retina liver and a variety of tissue have this ability.
Mouse	New born mouse hearts have been shown to have regenerative capacity similar to zebrafish.
Human	Broad expression in lung, small intestine and 21 other tissues
Birds	Elbow regeneration is found in aves and feather regeneration capability in <i>Gallus gallus</i> .

Figure4: Above table shows the regenerative capacity of different organism who are known to express Oct4 gene.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 7, July 2018

VI. CONCLUSION

Orthologs are fundamental to comparative genomics. Through this study we can conclude that the evolution of gene Oct4 is highly variable in evolutionary classes compared to mammals, for example only 50% identity percentage between humans and birds Oct4 ortholog. If we can study the genetic variations of identified orthologs at nucleotide level than may be through research and help of some bioinformatics tool, we will be able to predict the specific sequences of gene which are more vulnerable to variations in future. We can also understand the genetic mechanism of stem cells in different organisms comparatively.

REFERENCES

1. "Differential Regenerative Capacity Of Neonatal Mouse Hearts After Cryoinjury" by Ali Darechzershi et.al, Developmental Biology, Volume 399, Issue 1,1 March 2015
2. "Loss And Restoration Of Regeneration Capacity In Tissue And Organ Of Animals" by Lev Vladimirovich Polezhaev.
3. "Role Of OCT4 In Maintaining And Regaining Stem Cell Pluripotency" by Guilai Shi and Ying Jin, Stem Cell Research and Therapy (2010)
4. Stem Cell Pluripotency And Transcription Factor Oct4, by Guang Jin Pan, Zeng Yi Chang et.al., Cell Research, Volume 2, (2002).
5. "Dictionary Of Stem Cells Regenerative Medicine And Translational Medicine" By Frank J. Dye (2014)
6. "Zebrafish As A Model For Understanding The Evolution Of Vertebrate Immune System And Human Primary Deficiency"
7. Inferring Orthologs: Open Questions and Perspectives by Fredj Tekaia, 2016, PMC
8. National Center for Biotechnology Information Search Database.
9. https://www.ncbi.nlm.nih.gov/gene?cmd=Retrieve&dopt=full_report&list_uids=30333,5460,294562,427781,18999
10. "Evolution Of Humans" by Encyclopedia Britannica