**Comparative Genomic Studies of ORFan genes in Mammalian Genomes**

|  |
| --- |
| Purpose |
| Investigate the function/s, if any, of the ORFan gene sequences identified in previous studies from the literature, with a special focus of Clamp *at el* study. |

|  |
| --- |
| Method |
| There are three main phases of this proposed investigation   1. Find ORFan genes from the ORFanFinder online tool 2. Results Validation 3. Determine the function/s, if any, of the ORFan genes   ***Find ORFan genes from the ORFanFinder online tool***  Step 1: Find all genes of the chromosome  Ensembl BioMart[[1]](#footnote-1) online tool will be used to retrieve all the genes. Following parameters will be used in the BioMart search:  *Dataset*: Homo sapiens genes (GRCh38.p7)  *Filters*: Chromosome: 1-22 and X  *Attributes*: Peptide, Ensembl Gene ID, Ensembl Transcript ID  Step 2: Remove non-coding genes  Genes that do not contain a protein sequence will be removed from the output file.  Step 3: Filter orthologues genes to dog and mouse, and pfam domain proteins  By applying filters of the Ensembl Biomart tool, genes that are orthologous to dog and mouse and also proteins that has a pfam ID(belongs to known protein family) can be filtered out separately. Those transcripts are filtered out from the complete gene list to obtain candidate orphan genes to input into the ORFanFinder tool.  Step 4: Find orphan genes  Sequences of all the genes will be input into the ORFanFinder online tool[[2]](#footnote-2) to find orphan genes. Following parameters were used for the analysis:  *NCBI Taxonomy ID*: 9606  *Program*: blastp  *Use Premade Database*: nr  *E-value*: 1e-3  **Results Validation**  The orphan genes reported from the ORFanFinder can be compared against Clamp *at*. el study results. Further, each gene found as an orphan gene can be viewed in UCSC Brower for further validation such as:   * Retroposon/ Pseudogenes (RetroGenes V9 Track) * Pfam Domain (Pfam in UCSC Gene Track) * Repeating Elements (RepeatMasker Track) * Compare with mouse and dog sequence (Conservation Track)   **Determine the function of the ORFan genes**  Gene Ontology[[3]](#footnote-3) and GeneCard[[4]](#footnote-4) tools can be used to find the known functions of these genes |

|  |
| --- |
| Results |
|  |

1. http://asia.ensembl.org/biomart/martview [↑](#footnote-ref-1)
2. http://cys.bios.niu.edu/orfanfinder/index.php [↑](#footnote-ref-2)
3. http://www.geneontology.org [↑](#footnote-ref-3)
4. http://www.genecards.org [↑](#footnote-ref-4)