






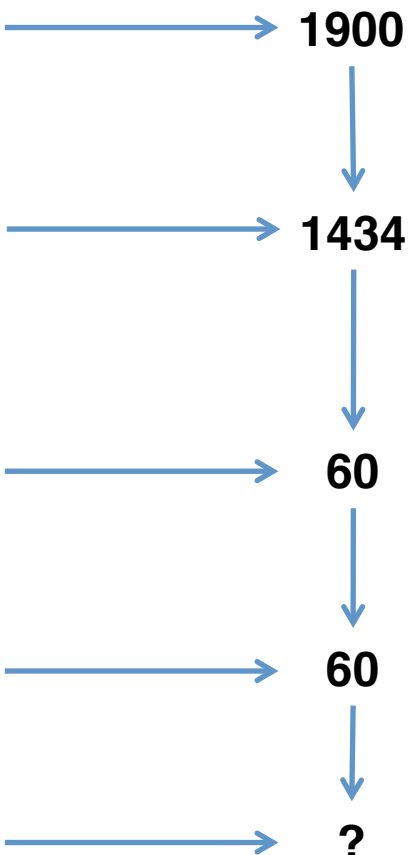
Human and Mouse Pseudogene Conservation

- Human Pseudogenes
 - GENCODE 10 HAVANA pseudogenes
- Mouse Pseudogenes
 - PseudoPipe output
 - Using ENSEMBL genome release 65
 - 15,887 pseudogenes

Pipeline

- Extract mouse sequences aligned to human pseudogenes, using UCSC multiz alignment
- Keep the mouse alignments satisfying the following criteria:
 - Aligned to > 70% human pseudogene sequence
 - Sequence identity > 40%
- Keep the mouse alignments overlapping with mouse pseudogenes from pseudopipe
 - Overlapping criterion: > 1 bp
- Filter out conserved mouse pseudogenes overlapping with coding regions
- Check conserved human and mouse pseudogenes for canonical splicing junctions

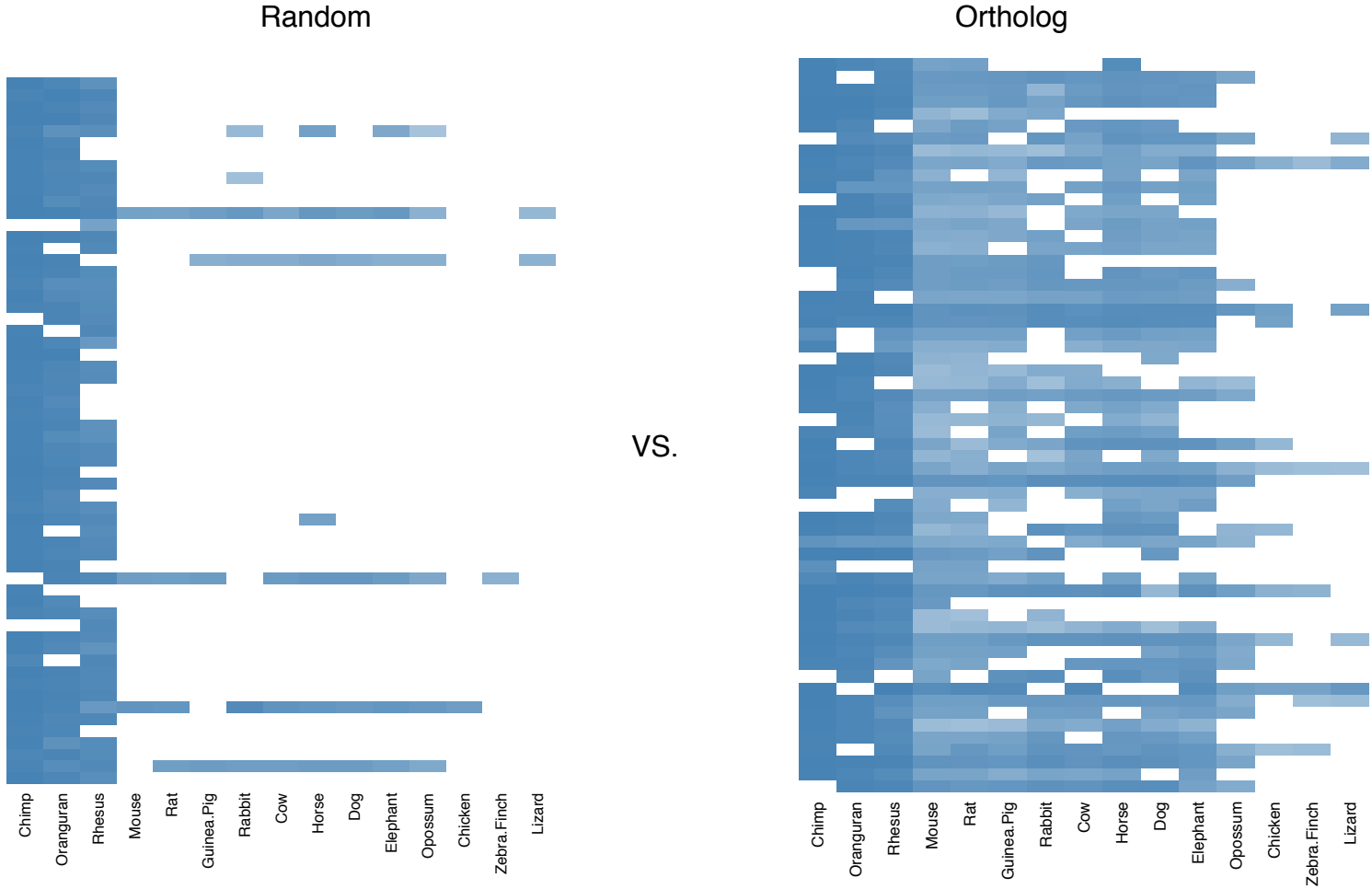
Results

- Extract mouse sequences aligned to human pseudogenes, using UCSC multiz alignment  **1900**
 - Keep the mouse alignments satisfying the following criteria:
 - Aligned to > 70% human pseudogene sequence  **1434**
 - Sequence identity > 40%
 - Keep the mouse alignments overlapping with mouse pseudogenes from pseudopipe  **60**
 - Filter out conserved mouse pseudogenes overlapping with coding regions  **60**
 - Check conserved human and mouse pseudogenes for canonical splicing junctions  **?**
- 

Conserved Human Pseudogenes

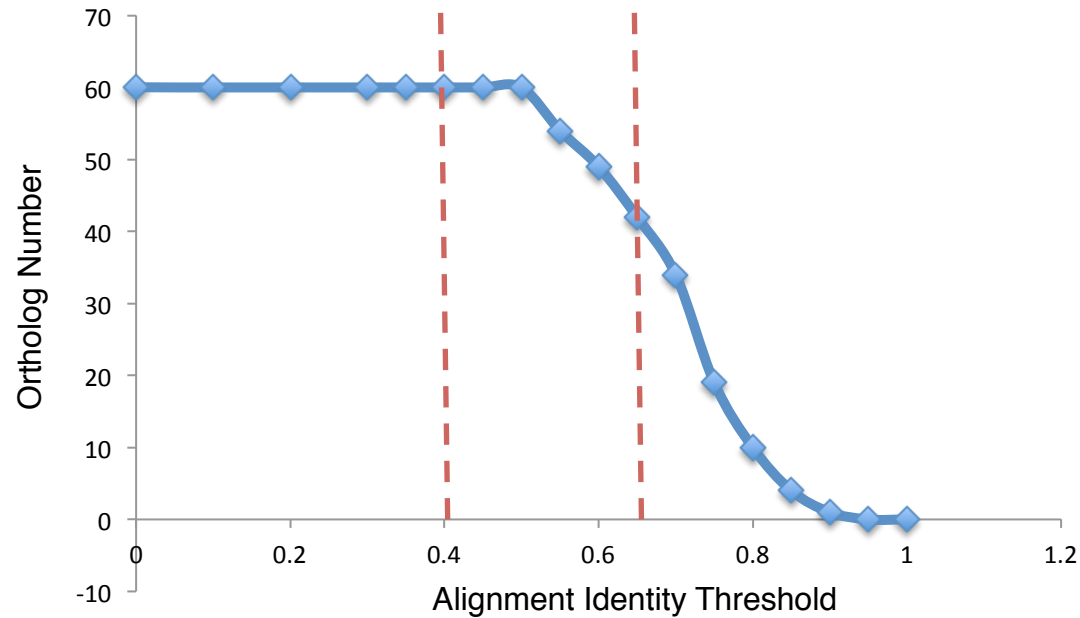
- 60 human pseudogenes are shown to be conserved since the divergence of human and mouse
- Within the 60 pseudogenes:
 - Transcribed: 13; non-transcribed: 47
 - Duplicated: 33; processed: 21; unitary: 3; ambiguous: 3
 -

Conservation in other organisms



Most human pseudogenes conserved since human-mouse speciation are conserved in more ancient time

Tweaking the criteria?



0.4: same threshold used in pseudopipe;

0.65: sequence identity for human-mouse orthologs under no selection