

Comparison of SNP distribution in Human and Mouse

BY:

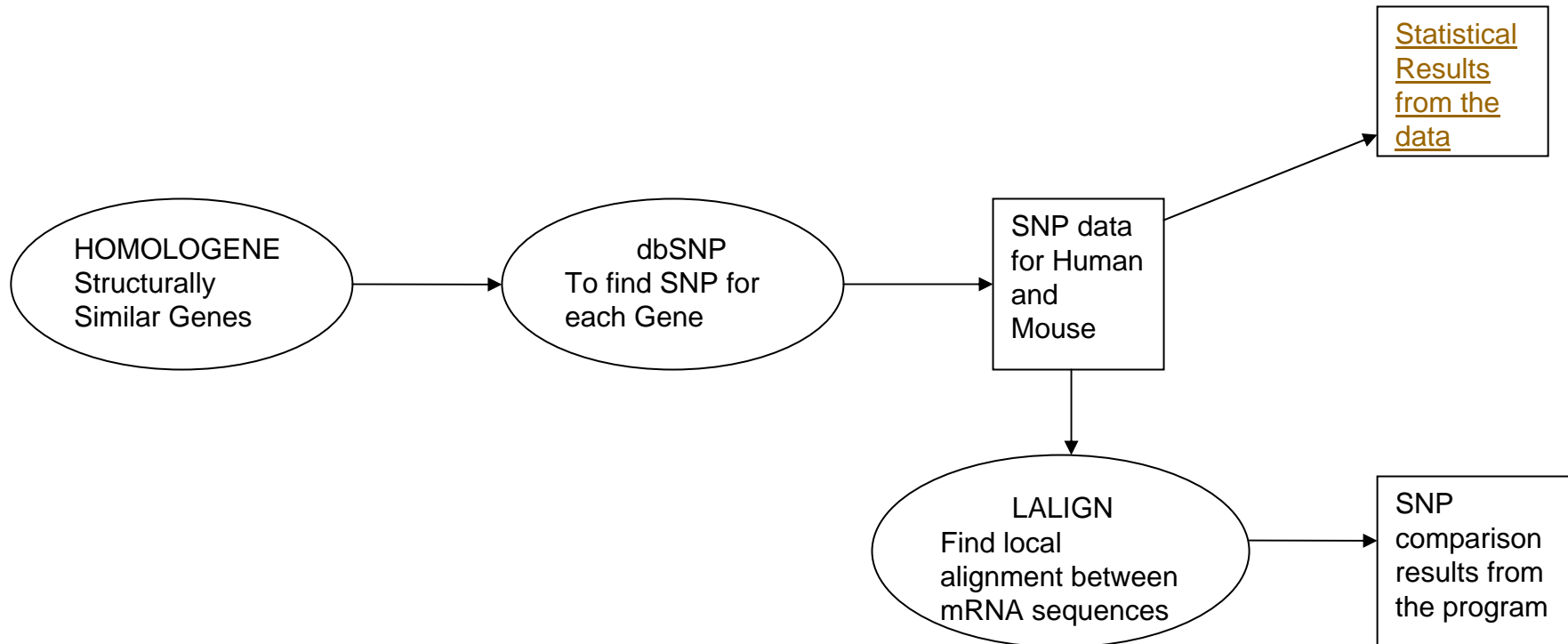
Amira El-Srougy

Shweta Bhargava

Jhelum Naik

Rahul Patil

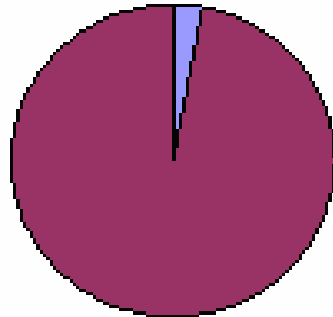
Methodology



Intron/Exon Distribution

Human

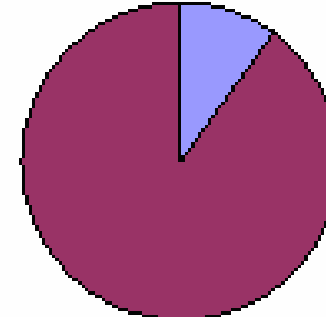
Percentage of SNPs in
Exon
3%



Percentage of SNPs in
Intron
97%

Mouse

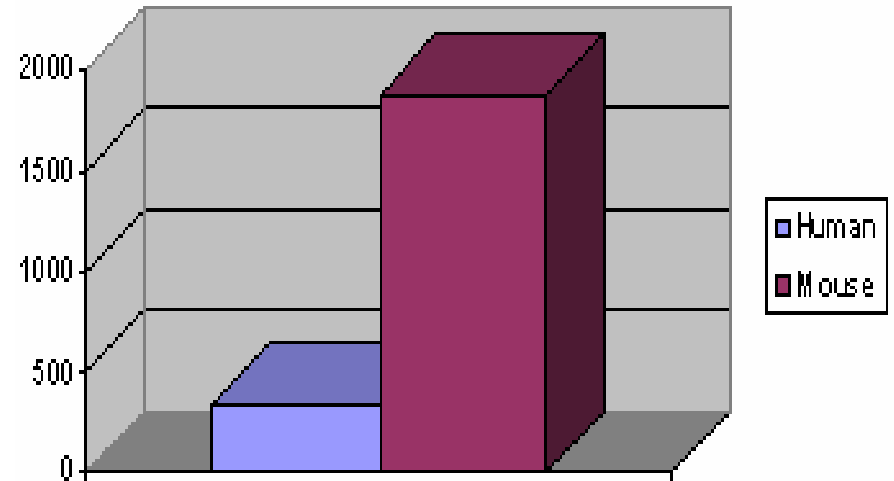
Percentage of SNPs in
Exon
10%



Percentage of SNPs in
Intron
90%

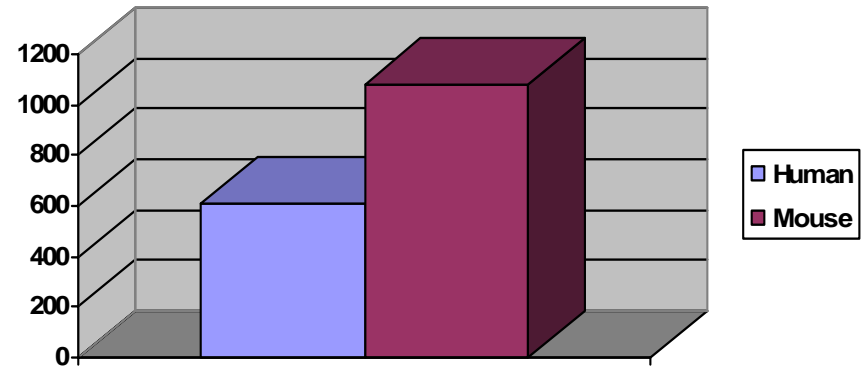
SNP density

- Human : 335
- Mouse : 1877



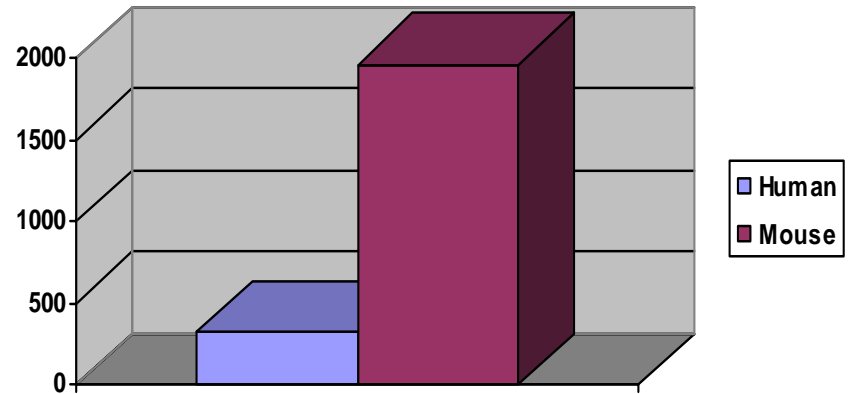
SNP density in Exon

- Human : 612
- Mouse : 1078

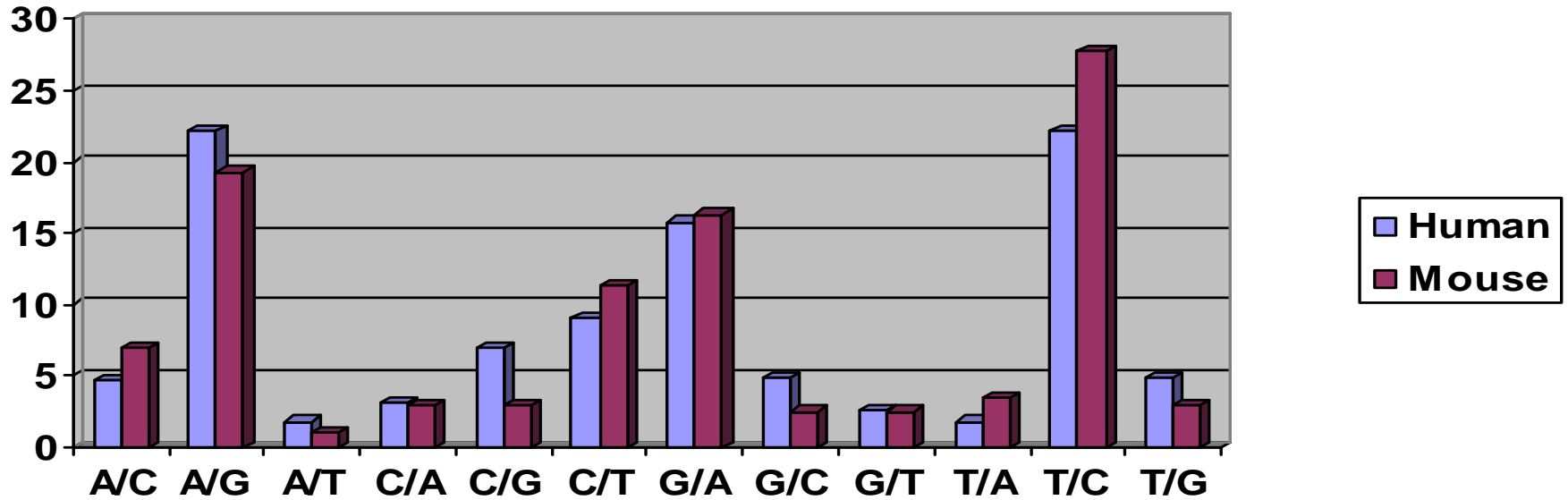


SNP density in Intron

- Human : 328
- Mouse : 1962



Alleles distribution



How to Compare SNPs?

- To test for SNP within a species is easier..
- Sequences are almost 99% similar

AAGAC **G** ATAGTACG

AAGAC **?** ATAGTACG

SNP Comparison.. Two Models

- Compare the SNP signatures.
- Align the mRNA sequences of the two homologous genes.



SNP signature

- For a mRNA of length 5000 the 7 letter sequence would be unique because..
- $.25^7 * 5000 = .3$
- AGACTGACTGATC**G**TTAAGCTTG
CTGATC**G**
ATC**G**TTA
GTTAAGC

Alignment Method

- Pair wise alignment between mRNA of Homologous genes.
 - Search for the Conserved SNPs in the alignment.
 - Local or Global Alignment ?
-

Lalign Output

```
89.2% identity in 240 nt overlap; score: 946 E(10,000): 1e-71
              70          80          90          100         110
human  CCGA-CGCGGAGCTGCAGATCGAGCGGCGCTTCGTGCCCGACGAGTGCCCGCGCACCGTG
      :: : :::: : : ::::: : : : : : : : : : : : : : : : : : : : : : : : :
mouse  GTGAGCTCGGAACTTCAGATCCAGCAGAGCTTCGTGCCTGATGAGTGTCCGCGCACGGTG
      60          70          80          90          100         110

      120          130          140          150          160         170
human  CGCAGCGCGGACTTCGTGCGCTACCACTACGTGGGGACGTTCCCCGACGGCCAGAAGTTC
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
mouse  CACAGTGCGGACTTCGTGCGCTACCACTACGTGGGGACTTTCCTCGACGGCCAGAAGTTC
      120          130          140          150          160         170
```

Alignment Method

- Read in the start point (nucleotide position) of the alignments from the given input file
 - Read in the SNPs locations for both human and the mouse from the given input file
 - Read in the number of total coding SNPs in human and mouse from the given input file
 - Find the relative distance between the starting point of the alignment and the SNP
 - Compare if the SNPs in humans and mouse occur at the same relative position.
 - If they occur at the same position and their alleles are the same, then the SNP is said to be conserved.
-

Are SNPs Conserved?

- None of the SNPs tested were found to be conserved.

WebSite : www.angelfire.com/sk3/compbio601

Thank You
