Effects of 2X Mammal Sequences on Constraint Elements in Pseudogenes
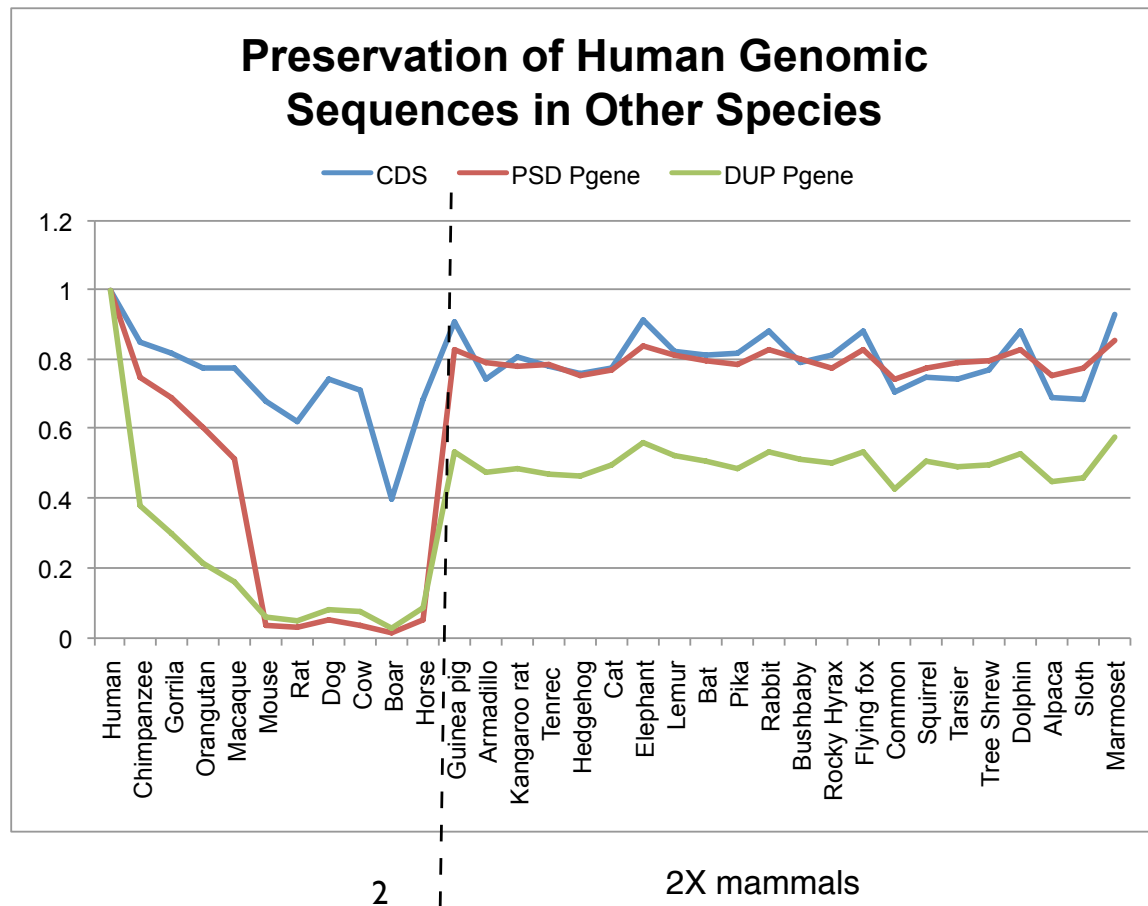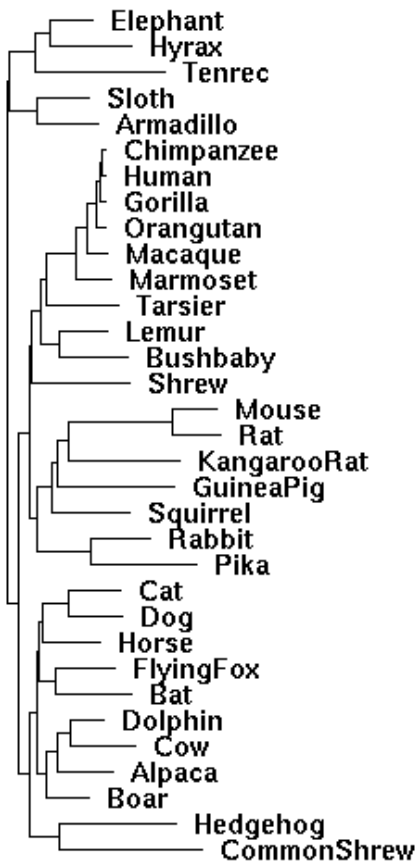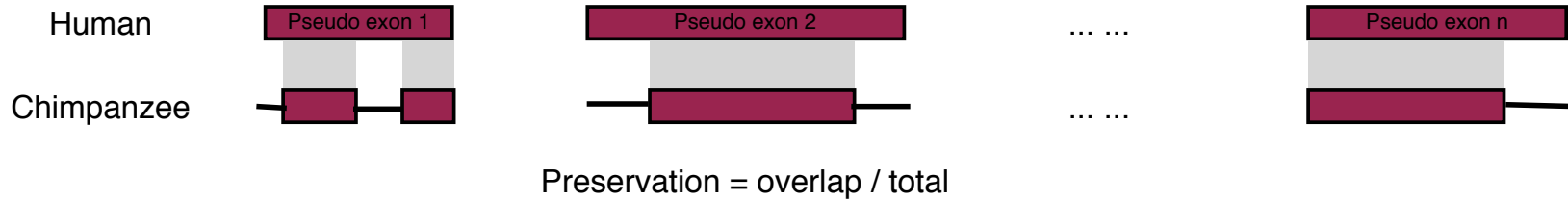
Annotation Meeting
Baikang Pei
06/22/2011

Multiple Sequence Alignments

Human | Pseudo exon 1 | Pseudo exon 2 | ... ... | Pseudo exon n

Chimpanzee

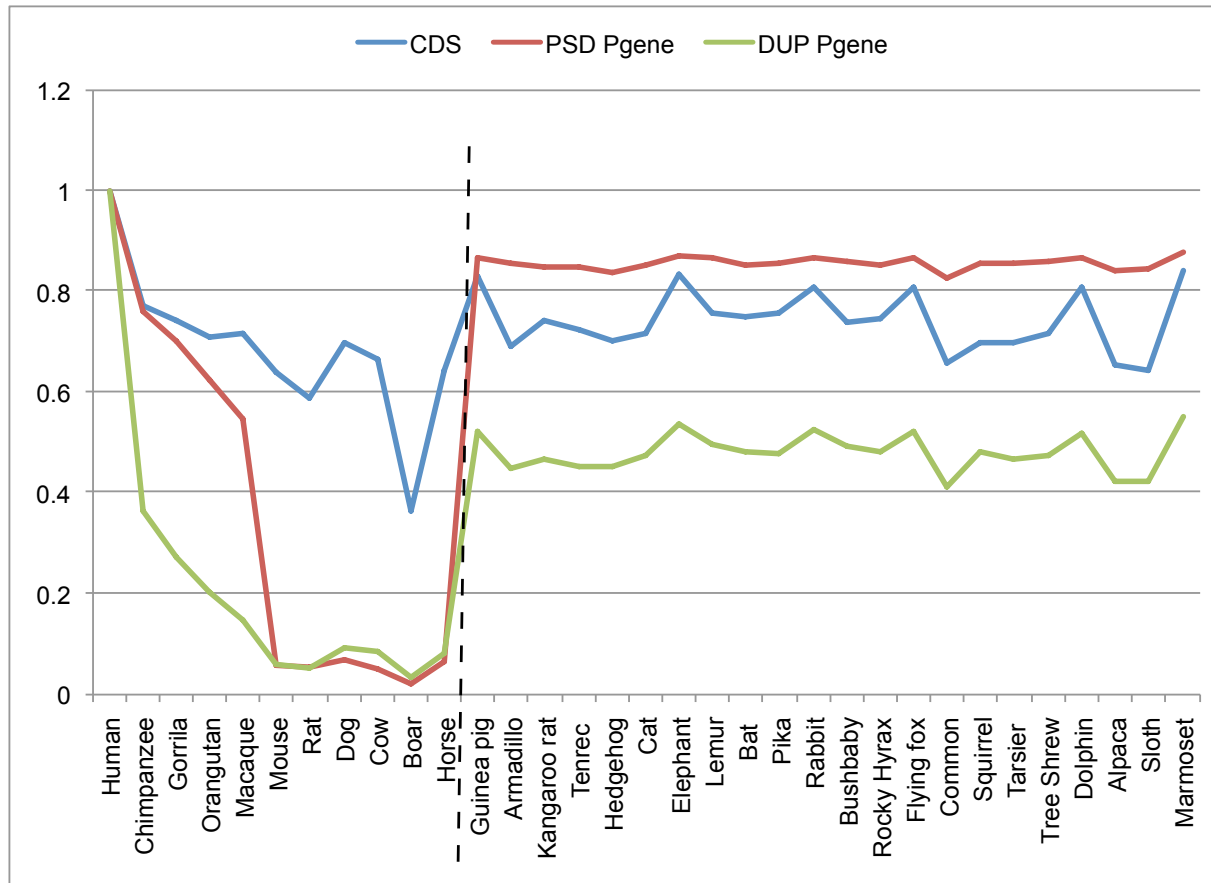Preservation = overlap / total

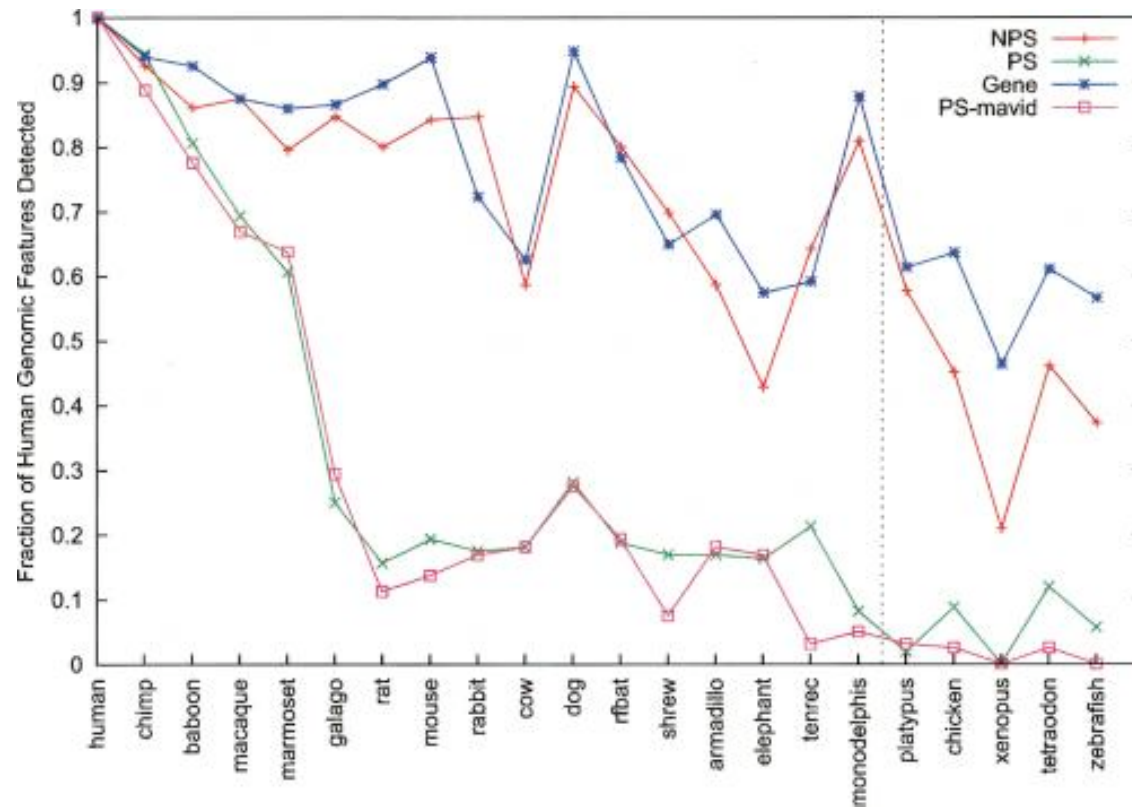# Other Preservation Criterion -- 50 nt

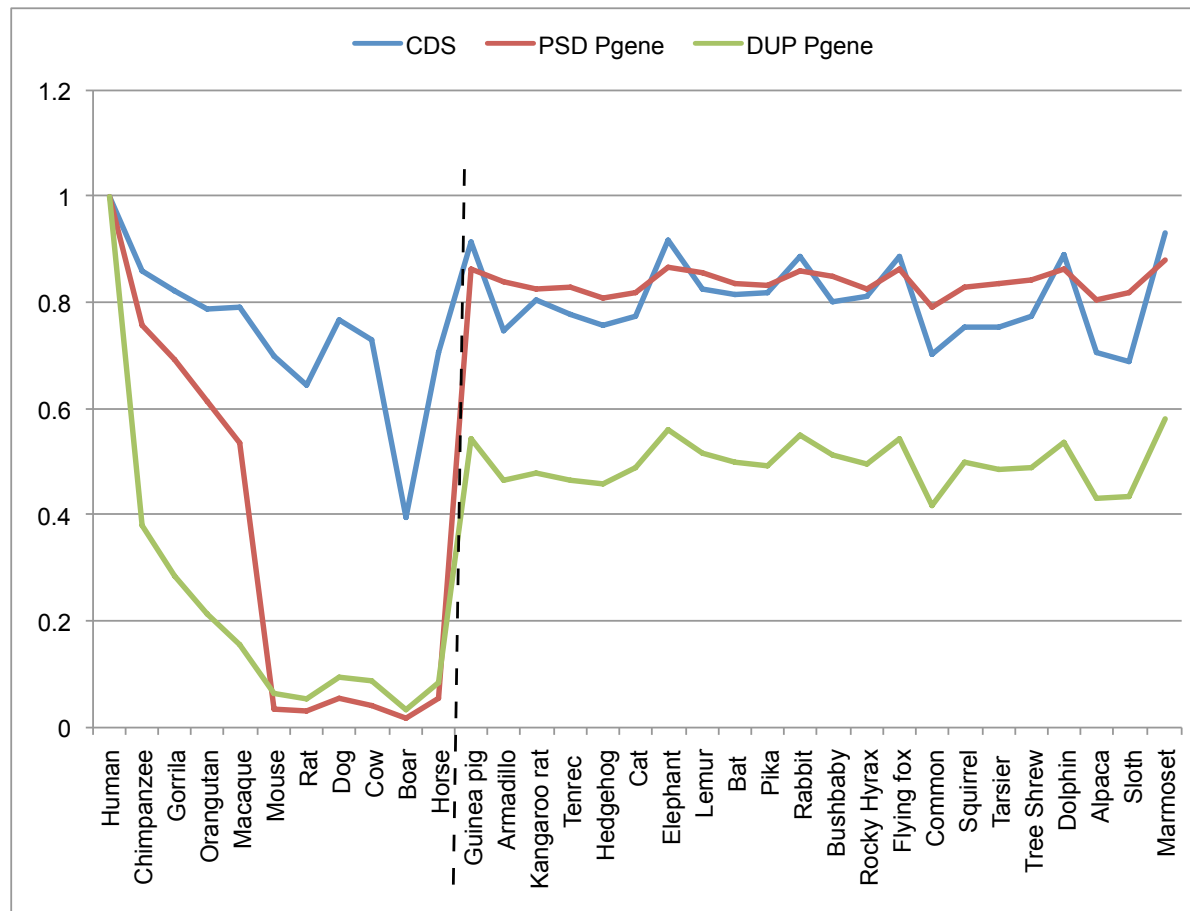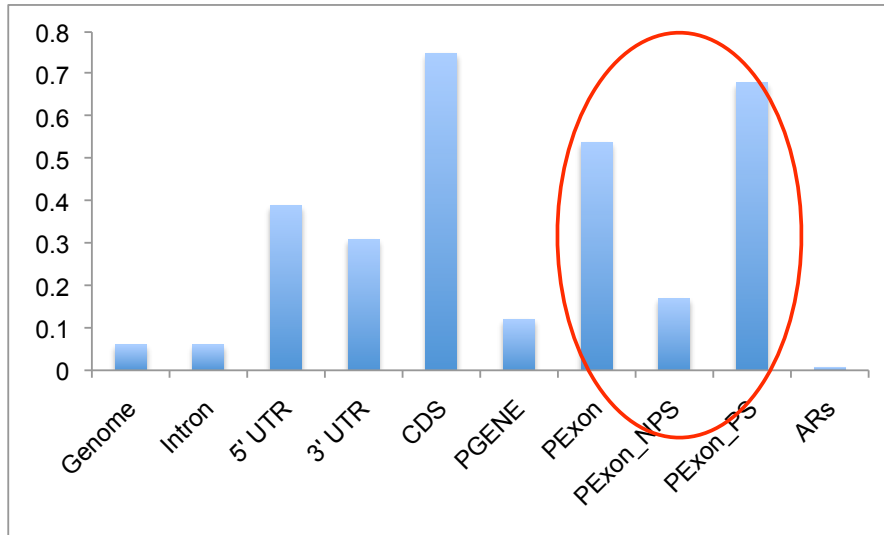# Figure in Pilot Companion Paper

Other Preservation Criterion -- 50%

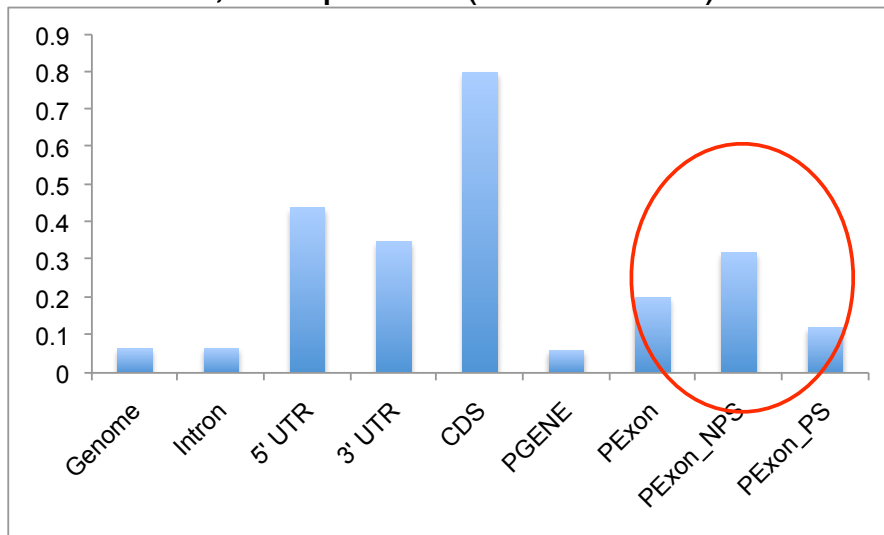# Constraint Regions from Different Alignments

EPO, 33 species (from ENCODE)

TBA, 36 species (from UCSC)

Difference of Species included in Alignment

	ENCODE	UCSC
Boar	x	-
Baboon	-	x
Wallaby	-	x
Opossum	-	x
Platypus	-	x

Difference of 2X mammals included in Alignment

	ENCODE	UCSC
Guinea Pig	x	-
Wallaby	-	x
Opossum	-	x
Platypus	-	x
Gorilla	-	x

There is a filter step in the TBA alignment: those for 2X mammalian genomes were filtered to retain only alignments of best quality in both the target and query ("reciprocal best")

Distribution of Constraint Elements in Pexons

EPO33

TBA36