

# Lower proportion than expected for transversions and higher for transitions in synonymous SNPs evaluated in dbSNP

Fernanda Stussi<sup>1</sup>, Tetsu Sakamoto<sup>2</sup>, José Miguel Ortega<sup>2</sup>

*1 UFMG*

*2 UFMG, LABORATÓRIO DE BIODADOS*

## Abstract

We have previously investigated the proportions of SNPs involving transversions (A/C, A/T, G/C, G/T) and transitions (A/G and C/T) for human entries in dbSNP. Here we reanalyzed human SNPs together with mouse, rat, pig and cow deposits. For these organisms, SNPs were classified after the region where the SNP is located: Intron, 5'-UTR, 3'-UTR, CDS-missense and CDS-synonymous. We noticed that the deposit for cow SNPs suffered somehow a bias that changed the relative proportions between these categories, suggesting that sampling is not random in the cow project. Therefore we set up to investigate the proportions of transitions and transversions for the other four organisms. The frequency of transitions varied around 33% for all categories but CDS-synonymous, where it was over 40%, and this measurement was attained for man, mouse, rat and pig. Accordingly, all four transversions were lower for CDS-synonymous. C/G Transversions were also lower for Intron and 3'-UTR and especially higher in CDS-missense and 5'-UTR. Thus, these studies support the occurrence of equal amounts in non-coding regions of genome for transitions and all transversions but C/G. And furthermore suggest that using SNP frequency in non-coding transcript regions in man, rat, mouse and pig would be safe for building mutation models based on SNP frequencies.

Funding: CAPES/Biocomp