1080-37-287 **David Koslicki** and **Daniel J Thompson*** (thompson@math.psu.edu). *Topological pressure and a priori coding sequence density estimation in the human genome.*

Inspired by concepts from ergodic theory, we introduce a new approach to coding sequence (CDS) density estimation for the human genome. Our approach is based on the introduction and study of topological pressure: a quantity assigned to any finite sequence based on an appropriate notion of 'weighted information content'. For human DNA sequences, each codon is assigned a suitable weight, and using a window size of approximately 60,000bp, we obtain a very strong positive correlation between CDS density and topological pressure. The weights are selected by an optimization procedure, and can be interpreted as quantitative data on the relative importance of different codons for the density estimation of coding sequences. We emphasize that topological pressure is a flexible tool and we expect it to be useful for the investigation of many other features of DNA sequences such as interspecies comparison of codon usage bias. We give a first result in this direction, investigating CDS density in the mouse genome and comparing our results with those for the human genome. (Received January 30, 2012)