

Abstract Submitted  
for the MAR09 Meeting of  
The American Physical Society

**Blasting and Zipping: Sequence Alignment and Mutual Information** ORION PENNER, Complexity Science Group, Department of Physics and Astronomy, University of Calgary, PETER GRASSBERGER, Complexity Science Group, Department of Physics and Astronomy, University of Calgary and Institute for Biocomplexity and Informatics., MAYA PACZUSKI, Complexity Science Group, Department of Physics and Astronomy, University of Calgary — Alignment of biological sequences such as DNA, RNA or proteins is one of the most widely used tools in computational bioscience. While the accomplishments of sequence alignment algorithms are undeniable the fact remains that these algorithms are based upon heuristic scoring schemes. Therefore, these algorithms do not provide model independent and objective measures for how similar two (or more) sequences actually are. Although information theory provides such a similarity measure - the mutual information (MI) - numerous previous attempts to connect sequence alignment and information have not produced realistic estimates for the MI from a given alignment. We report on a simple and flexible approach to get robust estimates of MI from global alignments. The presented results may help establish MI as a reliable tool for evaluating the quality of global alignments, judging the relative merits of different alignment algorithms, and estimating the significance of specific alignments.

Orion Penner  
Complexity Science Group,  
Dept. of Physics and Astronomy, University of Calgary

Date submitted: 22 Nov 2008

Electronic form version 1.4