

LOW COMPLEXITY REGIONS IN PROTEINS AND DNA ARE POORLY CORRELATED

JM ENRIGHT, ZW DICKSON, AND GB GOLDING

October 6, 2022

¹ Department of Biology, McMaster University, Hamilton, ON, Canada

* Author for correspondence: G. Brian Golding, Department of Biology, Life Science Building, McMaster University, Hamilton, ON, Canada, L8S 4K1. Email: golding@mcmaster.ca.

Abstract

Key Words: Low complexity, entropy, amino acid sequence, DNA sequence

Introduction

Materials and Methods

All custom scripts and commands used in this analysis can be found on GitHub at <https://github.com/JohannaEnright/LCREntropyProject/>.

For a detailed protocol, see Supplementary files on GitHub at <https://github.com/JohannaEnright/LCREntropyProject/>.

Sequence Data

LCR Simulation

Confidence Intervals for Correlation Coefficient

Results

Entropy of LCRs in Protein and DNA correlate Poorly with Corresponding Sequence Entropies

Entropy Correlation is Higher in Simulated Sequences

Discussion