Supplementary IVd: Prediction of DNA from context using neural networks. Fourier, zebrafish genome, GC/AT content, high frequencies.

Christian Grønbæk, Yuhu Liang, Desmond Elliott, and Anders Krogh

April 17, 2022

This file contains the second of two sets of plots based on the Fourier analysis showing the L2-norm of the Fourier coefficients in a running window. This first set covers the frequency range from 200 to 45000 using a window length of 1000 (and a step size of 100); this second file/set covers the frequencies from 40000 to 140000 and used a window length of 5000 (and step size of 100).

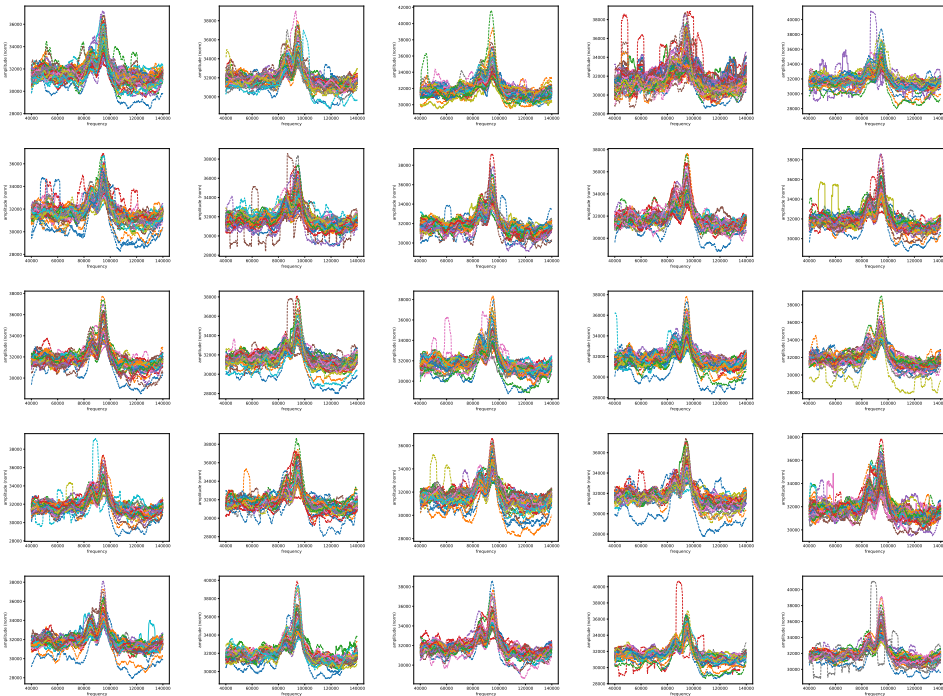# Fourier plots, GC/AT content, frequency range 40000 to 140000

Figure 1: Zebrafish GC/AT content. Fouriers on GC/AT content arrays, higher frequency range. Each plot covers one chromosome, listed in increasing order (chr1 to chr25). The genome string is divided in adjacent segments of 1Mb (per chromosome); each plot shows the results for all segments in the chromosome (with ratio of qualified positions > 0.9).