1b: There is quite a lot of variance in entropy for the values between positions 200 and 300, which indicates that EVO2's confidence in the nucleotide for the position varies dramatically depending on the position.

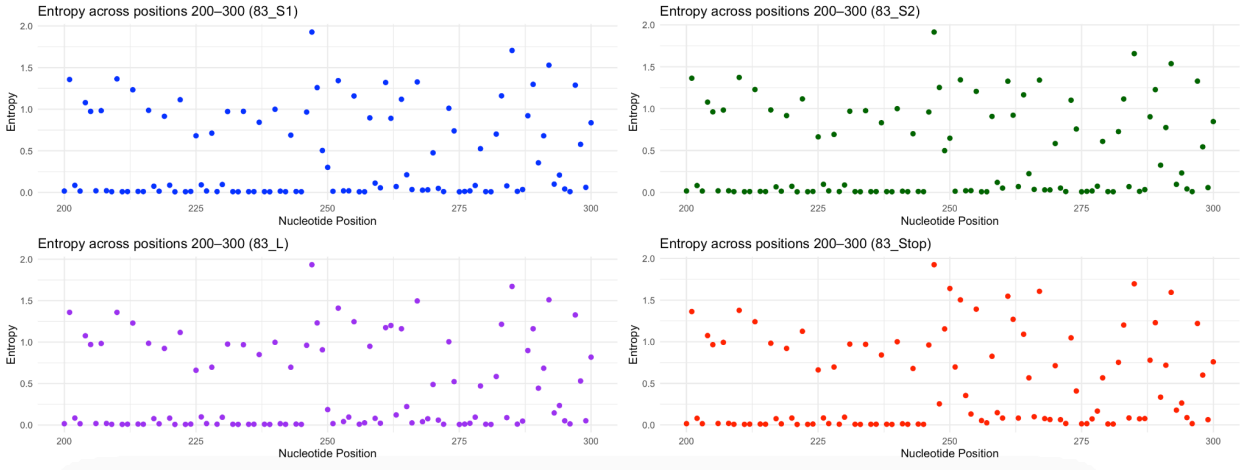
1d: The probability of EVO2 choosing the correct base is 0.877 for 83_S1. This indicates a pretty high likelihood that for a given position, EVO2 will choose the correct nucleotide (as the highest probability).

(Photo from Elena)

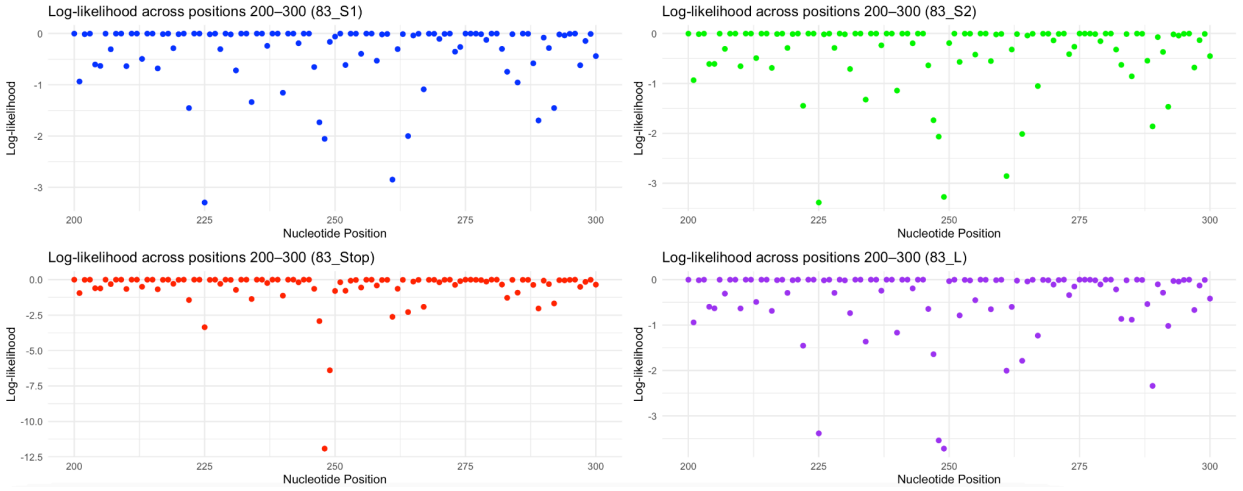
2a: 83_S1 -> AGC (valid serine at 83)
Stop Codon 83_S -> TGA (valid stop codon at 83)
83_S2 -> `AGT (`synonymous substitution at 83 (T))

83_L -> ATC (isoleucine not leucine)???

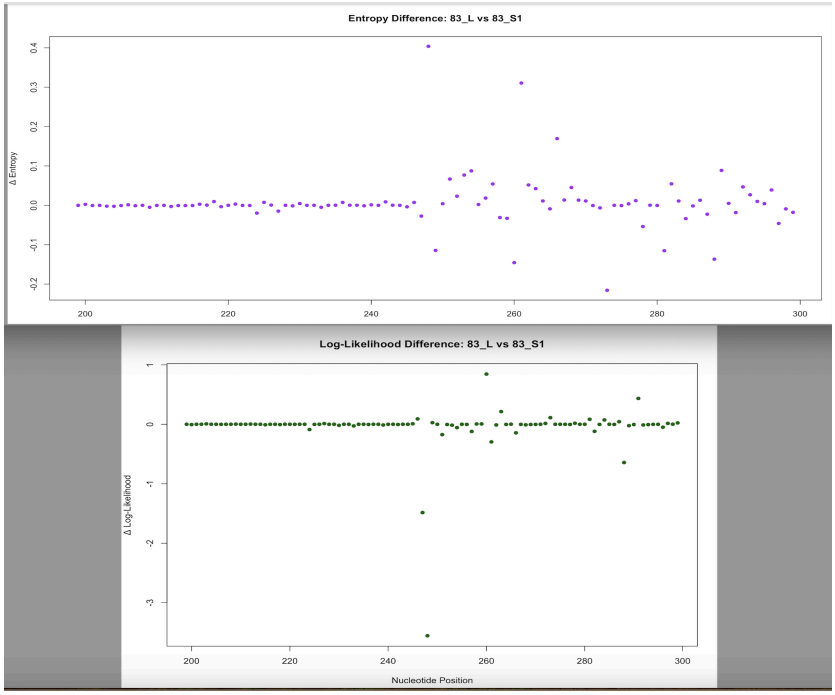
Single entropy plots for all genes



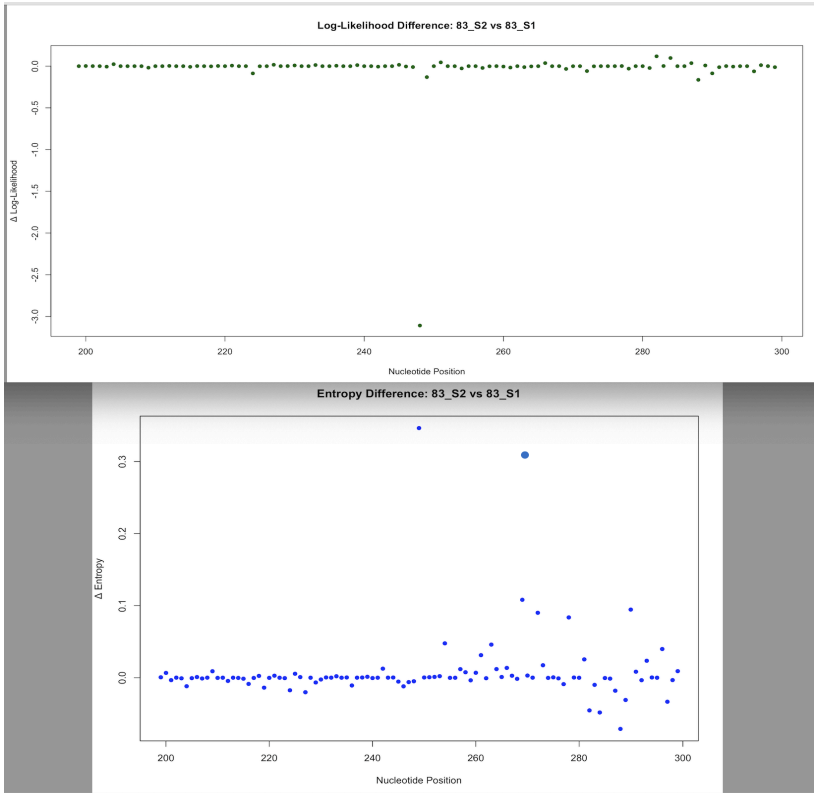
Log-likelihood plots for all genes



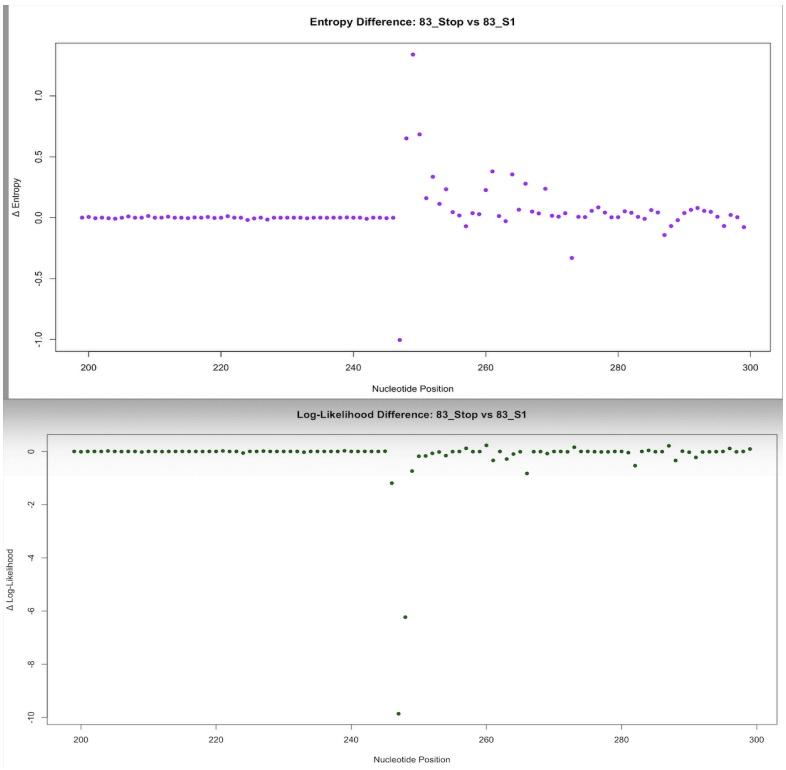
83_S1 vs 83_L entropy and log-likelihood differences



83_S1 vs 83_S2 entropy and log-likelihood differences



83_S1 vs 83_Stop entropy and log-likelihood differences



2d: The difference in entropy and the difference in log-likelihood graphs tended to spike at positions 247-249 for all of the genes. For entropy, this indicates that EVO2 is unsure about what nucleotide goes there and that a mutation is likely to occur (EVO2 can predict mutation). For log-likelihood, this indicates that EVO2 is much more likely to choose the incorrect nucleotide at these positions, which could be indicative of a mutation.

Comparing total log-likelihood