Result summary

Aim: Identify the most likely breed of a dog from its sequence

The breed most similar to the mystery DNA file is the ENGLISH\_SPRINGER\_SPANIEL its percent identity is 99.678% the alignment of the mystery dog breed and ENGLISH\_SPRINGER\_SPANIEL is displayed here:

A colorful lines with numbers

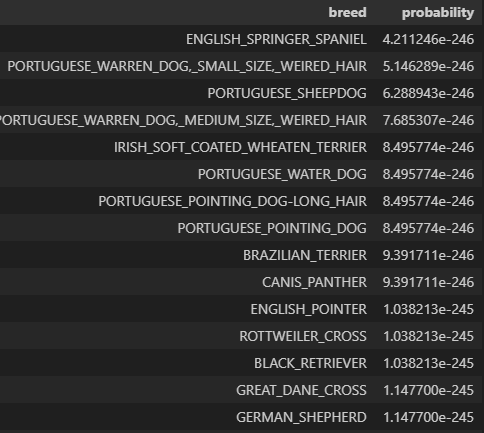
Description automatically generated with medium confidence

Full sequence alignment can be found in the results folder.

The probability that a random sequence alignment has a certain similarity score is displayed in the graph below.

The score of the alignment between the ENGLISH\_SPRINGER\_SPANIEL and the mystery sequence is 16698.0 which has the probability of occurring by chance of 4.211246e-246.A graph of a normal distribution

Description automatically generated



To visualise this a MSA alignment has been performed between all Breed sequences. The branch length of the mystery sequence with ENGLISH\_SPRINGER\_SPANIEL if fairly long especially in comparison to the BLACK\_RETRIVER and GERMAN\_SHEPARD where the branch length is much shorter and nevertheless these are two different breeds. Therefore, it is possible that the mystery sequence is of a breed not included in the database.

A diagram of a structure

Description automatically generated with medium confidence

Using a larger database and comparing the sequence alignment scores of different dogs of the same breed could be useful in determining the true identity of the mystery sequence.