biotools

Tetra

Calculates different variations of tetra-nucleotide frequencies of different input data

Functions

`countTetraNucleotide(sequence: string):CountTable[string]`
Counts 4-mer amount of a given string. If a 4-mer does not not occur in the string, it will not occur in Table, to get a Table in which not occuring 4-mers have a value of 0 in the table use `countTetraNucleotideAll`
Parameter 1: String of which to count the mount of 4-mers.
Return: Table in which keys are the 4-mers and the corresponding values are the amounts of it.

`countTetraNucleotideAll(sequence: string):CountTable[string]`
Counts 4-mer amount of a given string. If a 4-mer does not not occur in the string, it will still occur in the table with a value of zero.
Parameter 1: String of which to count the mount of 4-mers.
Return: Table in which keys are the 4-mers and the corresponding values are the amounts of it. Not occuring 4-mers have a value of zero.

`countTetraNucleotideReversed*(sequence: string):OrderedTable[string, int]`
Counts 4-mer amount of a given string. it further collapses 4-mers and their reverse complement into single dictionary entries i.e. if a sequence contains `AAAA` and `TTTT` one time only `AAAA` will appear with a value of 2.
Parameter 1: String of which to count the mount of 4-mers.
Return: Table in which keys are the 4-mers and the corresponding values are the amounts of it. Not occuring 4-mers have a value of zero.

`countAllFreqs(sequences: seq[Sequence], normalized:bool=false): OrderedTable[string, string]`
Counts the amount of 4-mers of a given array of sequences, this funtction is normally used with a FASTA file and will output the tetra-nucleotide amount for each contig.
Parameter 1: String of which to count the mount of 4-mers.
Parameter 2: If normalization is enabled the tetranucleotide-frequency will be normalized by the contig length.
Return: Table in which keys are the 4-mers and the corresponding values are the amounts of it. Not occuring 4-mers have a value of zero.