Building Statistical Profile for SITECON

Task Name: sitecon-build

Builds a statistical profile for SITECON. It can be later used to search for TFBS.

Parameters:

in — semicolon-separated list of input DNA multiple sequence alignment files. An input file must not contain gaps. [String, Required]

out — output file. If several input files have been supplied, then a sitecon profile is built for each input file, i.e. several output files (with different indexes) are generated. [String, Required]

wsize — window size. The window is a region of the alignment used to build the profile. It is picked up from the center of the alignment and occupies the specified length. The edges of the alignment beyond the window are not taken into account. The recommended length is a bit less than the alignment length, but not more than 50 bp. [Number, Optional, Default: 40]

clength — length of a random synthetic sequence used to calibrate the profile. [Number, Optional, Default: 1000000]

rseed — random seed used to calibrate the profile, e.g. to generate the random synthetic sequence. Use the same value to get the same calibration results twice on the same data. By default, new random seed is generated each time a calibration occurs. [Number, Optional, Default: 0]

walg — specifies to use the *Algorithm 2* weight algorithm. In most cases it is not required, but in some cases it can increase the recognition quality. [Boolean, Optional, Default: false]

Example:

```
ugene sitecon-build --in=COI.aln --out=result.sitecon
```