DNA Statistics

The *DNA Statistics* plugin provides exportable statistic reports.

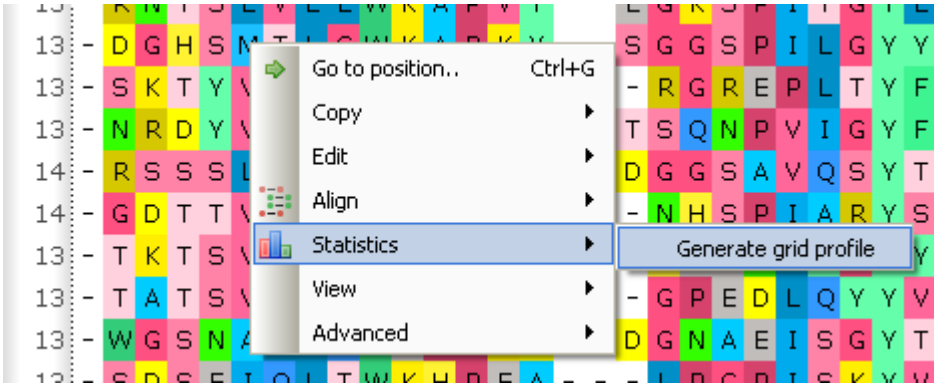
In the current UGENE version the *DNA Statistics* plugin provides only *Alignment Grid Profile* report. The *Alignment Grid Profile* shows positional amino acid or nucleotide counts highlighted according to the frequency of symbols in a row.

The original idea of the MSA Grid Profile is described in the following paper:

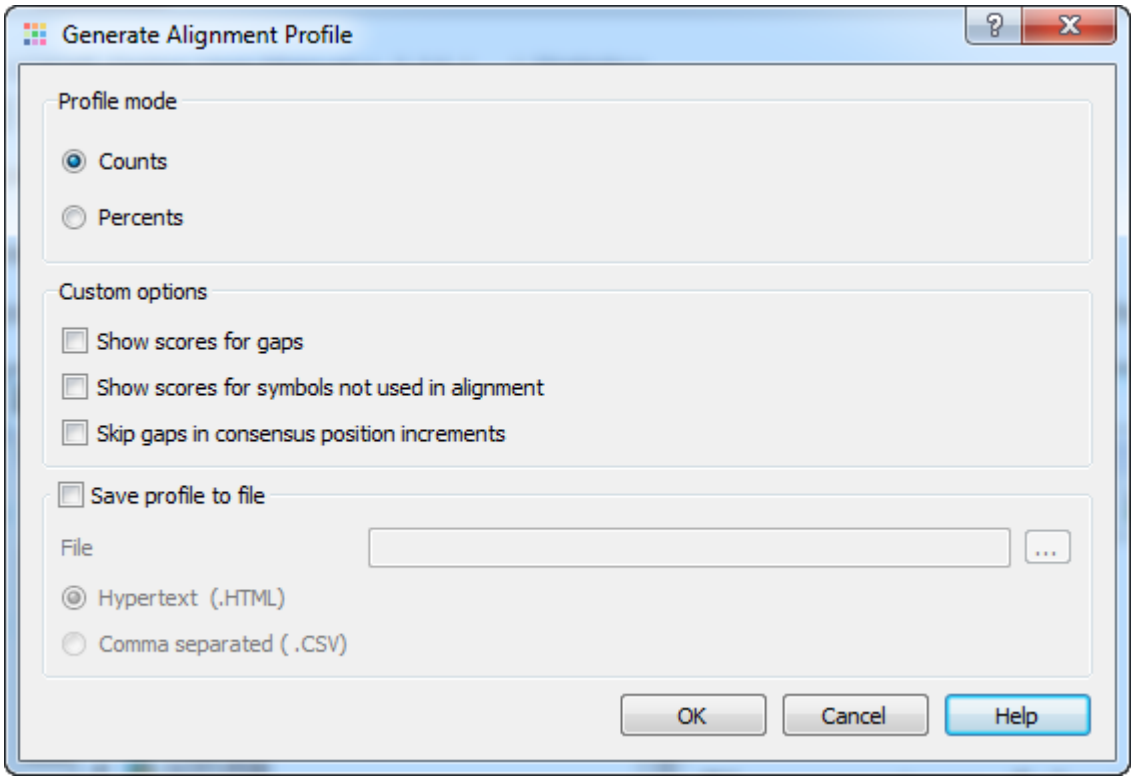
“Alberto Roca, Albert Almada and Aaron C Abajian: ProfileGrids as a new visual representation of large multiple sequence alignments: a case study of the RecA protein family, BMC Bioinformatics 2008, 9:554”

Usage example:

Open a sequence alignment in the *Alignment Editor* and use the *Statistics Generate grid profile* context menu item.



The dialog will appear:



Here is a brief description of the options that can be set in the dialog:

Profile mode: Counts/Percents — select the *Percents* to have scores shown as percents in the report.

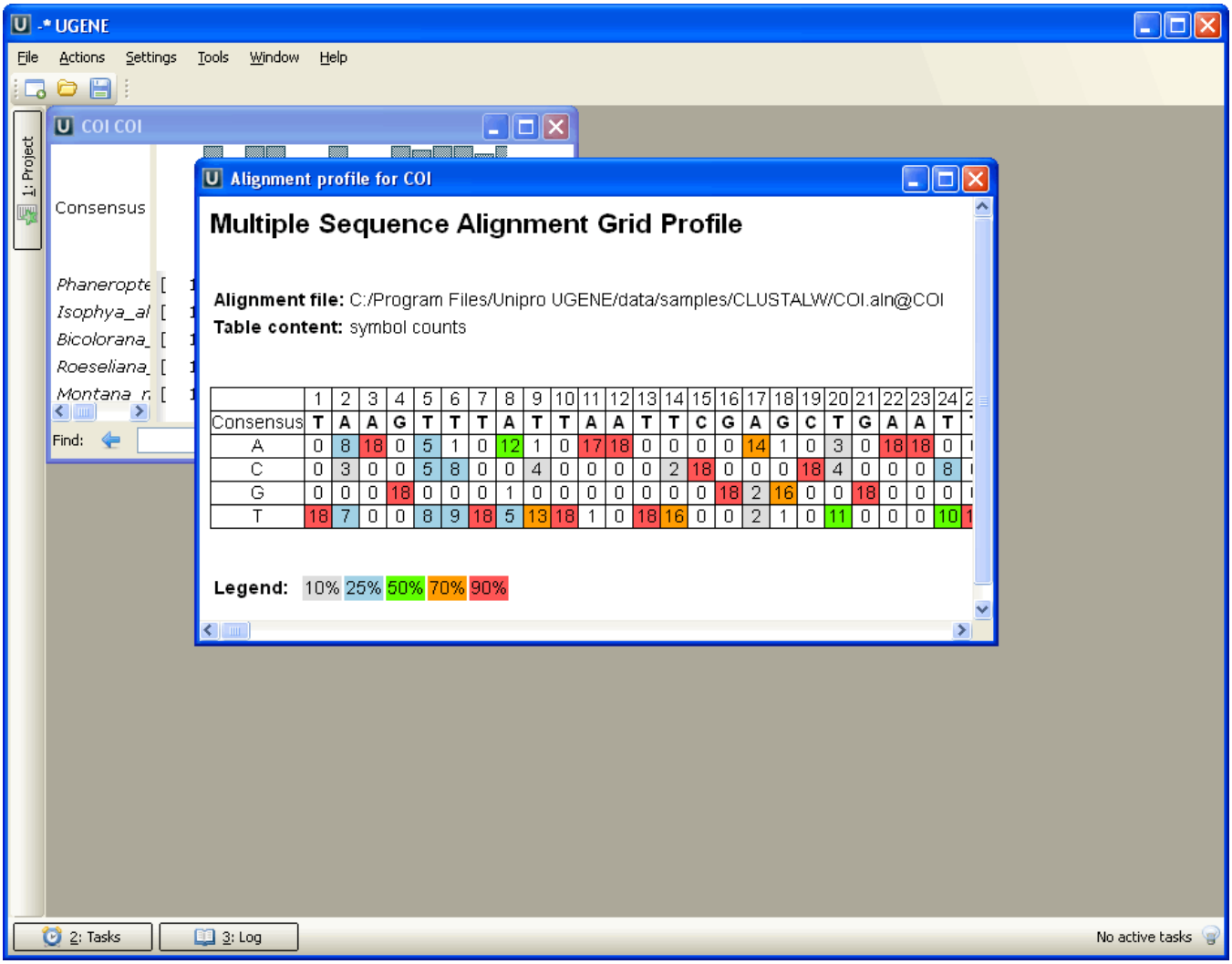
Show scores for gaps — check this item if you want gap characters (‘—’) statistics to be shown in the report.

Show scores for symbols not used in alignment — if a symbol is not used in the alignment at all it won't be shown in the report. Check this item to make all symbols of alignment alphabet reported.

Skip gaps in consensus position increments — consensus ruler configuration. If checked the gaps in consensus will not lead to ruler increments.

Save profile to file — allows to save profile to a file in the HTML or CSV format. The CSV format is convenient for further processing in worksheets editors like Excel.

The result profile in the HTML mode:



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Consensus	T	A	A	G	T	T	T	A	T	T	A	A	T	T	C	G	A	G	C	T	G	A	A	T
A	0	8	18	0	5	1	0	12	1	0	17	18	0	0	0	0	14	1	0	3	0	18	18	0
C	0	3	0	0	5	8	0	0	4	0	0	0	0	2	18	0	0	0	18	4	0	0	0	8
G	0	0	0	18	0	0	0	1	0	0	0	0	0	0	0	18	2	16	0	0	18	0	0	0
T	18	7	0	0	8	9	18	5	13	18	1	0	18	16	0	0	2	1	0	11	0	0	0	10