

SOM Table 1. General features of the *Kuenenia stuttgartiensis* genome and comparison with *Rhodopirellula baltica*²³

	<i>K. stuttgartiensis</i>	<i>R. baltica</i>
Total size, bases	4,218,325	7,145,576
G+C content, %	41	55.4
Coding sequences	4,664	7,325
Coding density, %	85.9	95
Average gene length, bases	776	939
Genes with similarities in databases*	3,279 (70.3%)	4,613 (63%)
Genes with functional assignments**	1,385 (29.7%)	2,712 (37%)
rRNAs	1x 16S-23S-5S	1x (16S) and (23S-5S)
tRNAs	44	70
<i>Archaea</i> -like coding sequences***	6.9%	9%
<i>Eukarya</i> -like coding sequences***	3.7%	8%

* Threshold for BLASTP E-value < 1×10^{-3} (includes hits to hypothetical proteins). Value was recalculated for *R. baltica* by taking into account genomes published since the *R. baltica* genome was published.

** Value was recalculated for *R. baltica* by taking into account genomes published since the *R. baltica* genome was published.

*** Coding sequences with a first BLASTP hit with an E-value < 1×10^{-3} with proteins from *Archaea* and *Eukarya*, respectively.