

Appendix 4.

Comparing *Firmicutes* with similarly AT-rich *Proteobacteria*

	<i>Proteobacteria</i>		<i>Firmicutes</i>							
	HI	HP	LI		LM		SPn		TT	
AT-content	61.9%	60.8%	62.6%		62.0%		60.28%		62.4%	
<i>g</i>	1673	1478	2962	HI - HP	2837	HI - HP	1861	HI - HP	2588	HI - HP
<i>sp_{CI}</i>	19	18	36	<i>34-36</i> (*)	33	<i>32-35</i>	23	<i>21-23</i>	33	<i>29-32</i>
<i>sp_{CII}</i>	24	23	45	<i>43-46</i>	41	<i>41-44</i>	31	<i>27-29</i>	42	<i>37-40</i>
<i>upsp_{CI}</i>	1	1	2	<i>2-2</i>	2	<i>2-2</i>	1	<i>1-1</i>	2	<i>2-2</i>
<i>upsp_{CII}</i>	8	7	16	<i>14-14</i>	15	<i>14-13</i>	9	<i>9-9</i>	16	<i>12-12</i>

Table 4.1. Comparison of the numbers of putative strong promoters identified in "average" genomes randomly generated, for two *Proteobacteria* and four *Firmicutes* characterized by close (high) AT-contents (range [60.2%, 62.4%]), under conditions *CI* and *CII*, and with or without UP element required. N.B.: This Table is complementary to Table 1. *g*: total number of genes encoding proteins in the genome considered; HI: *Haemophilus influenza*, HP: *Helicobacter pylori*; LI: *Listeria innocua*, LM: *Listeria monocytogenes*, SPn: *Streptococcus pneumoniae*, TT: *Thermoanaerobacter tengcongensis*. (*) With HI then HP taken as a reference, these columns in italics yield the predicted values *sp_{CI}*, ..., *upsp_{CII}* on the basis of proportionality to total gene number; the two predicted values in italics have to be compared with the value on their left. Example: the reference being HI, predicted *sp_{CI}* for LI is $sp_{CI}(LI) = \frac{sp_{CI}(HI) \times 2962}{1673} = 34$, with $sp_{CI}(HI) = 19$; $sp_{CI}(LI)$ has to be compared with the value of 36 observed for the "average" randomly generated genome relative to LI.