Prediction of highly expressed genes in microbes based on chromatin accessibility

BMC Molecular Biology 2007, 8:11 Willenbrock & Ussery

> Overview by Age Tats

Journal Club Feb 26th, 2007

Background

- Gene expression is dependent on chromatin structure in eukaryotes.
- The 'position preference'(PP) measure is a DNA structural measure, which reflects the preference of a given di-/trinucleotide for being found in a region where the DNA minor groove faces either towards or away from the nucleosome histone core.
- High absolute PP reflects a high preference for nucleosomes, while low absolute PP reflects di-/trinucleotides which tend to exclude nucleosomes.
- The PP measure also describes a more general structural property of DNA that is, how easily can it be wrapped around chromatin proteins.

Nucleosome structure in eukaryotes



Prokaryotes do not have nucleosomes but they also have chromatin, and the DNA is compacted to similar levels (i.e., more than 1000x) in both prokaryotes and eukaryotes.



Willenbrock & Ussery 2004

Productive binding of DNase I requires DNA to be bent.

DNase I interacts with a 6 bp contact surface of the minor groove and bends the DNA molecul away from the enzyme, towards the major groove. Therefore, base sequences that are flexible or inherently bent towards the major groove should be more accessible to DNase I cleavage.

Trinucleotide step	No. of occurrences in dataset	DNase I-derived trinucleotide parameter (ln p)
AAT/ATT	89	-0.280
AAA/TTT	278	-0.274
CCA/TGG	45	-0.246
AAC/GTT	81	-0.205 ^b
ACT/AGT	77	-0.183 ^b
CCG/CGG	73	-0.136
ATC/GAT	112	-0.110
AAG/CTT	110	-0.081
CGC/GCG	84	-0.077
AGG/CCT	101	-0.057
GAA/TTC	117	-0.037
ACG/CGT	84	-0.033
ACC/GGT	87	-0.032
GAC/GTC	81	-0.013
CCC/GGG	141	-0.012
ACA/TGT	52	-0.006 ^b
CGA/TCG	84	-0.003
GGA/TCC	71	0.013
CAA/TTG	74	0.015 ^b
AGC/GCT	35	0.017
GTA/TAC	83	0.025
AGA/TCT	127	0.027
CTC/GAG	102	0.031 ^b
CAC/GTG	55	0.040
TAA/TTA	99	0.068 ^b
GCA/TGC	34	0.076
CTA/TAG	64	0.090
GCC/GGC	57	0.107
ATG/CAT	71	0.134 ^b
CAG/CTG	61	0.175
ATA/TAT	80	0.182
TCA/TGA	127	0.194

Table I. DNA bending and/or bendability parameters as revealed by

DNase I: parameters of trinucleotide steps^a

Brukner et al 1995



- Codon adaptation index is highly correlated with the expression level in fast growing bacteria, but
 - it cannot consider tRNAs, ribosomal RNAs and other noncoding RNAs;
 - it is less effective predictor in slow growing organisms;
 - it requires the identification of a representative subset of highly expressed genes in an organism (problematic for more distant microbes).
- Alternative using the PP for the prediction of highly expressed genes in microbial genomes?

- The average position preference for ribosomal protein encoding genes is lower than for other protein encoding genes (Wilcoxon P-value 4e-11).
- rRNAs, tRNAs, and miscellaneous RNAs have significantly lower position preference values than translated genes (P-value = 6e-34).



Figure 2. Gene density plots. (A) Density plot of position preference differences for 328 microbial genomes. Differences between mean position preference of translated coding sequences (CDSs) and ribosomal proteins (red) or between mean position preference of CDSs and ribosomal RNA (rRNA) (blue). Most microbial genomes CDSs have a higher mean position preference values than ribosomal proteins and rRNA (mean above 0). (B) Position preference densities for the 10% most highly expressed genes, non-highly expressed genes, rRNAs, tRNAs and ribosomal proteins in *E. coli*.

- There was no significant correlations between CAI triplet weights and PP scores.
- The correlation between CAI weights and PP triplet values did not increase for fast replicating bacteria, indicating that PP may be a useful supplement for predicting highly expressed non-translated genes even in slow-growing microbes.
- The PP measure could be useful for identifying rRNA, tRNA and other non-coding RNA genes in pre-annotated DNA sequences, because those genes tend to have lower PP than the genomic average.



Table 1. Predicted highly expressed non-translated E. coli genes by the position

preference measure.

snl bi977 FRM 122 0.1076 leup bi356 FRNA 122 0.1076 selC bi3558 FRNA 617 0.1296 snU bi989 FRNA 121 0.1076 selC bi3558 FRNA 617 0.1296 snU bi989 FRNA 121 0.1076 ser/U bi3973 FRNA 139 0.1310 alW bi9666 FRNA 121 0.1076 ser/U bi3973 FRNA 149 0.1310 alW bi9666 FRNA 124 0.1172 rrsA bi3973 FRNA 172 0.1316 alY b0746 FRNA 129 0.1212 rrsG b33751 FRNA 172 0.1332 alY b0206 FRNA 129 0.1212 rrsG b33753 FRNA 172 0.1332 alY b0206 FRNA 129 0.1212 rrsG b33753 FRNA 172 0.1338<	Gene	bnumber	Type	Gene expression rank	ЪР	Gene	bnumber	Type	Gene expression rank	dд
Sinv b1984 tRNA 121 0.1076 selC b3558 tRNA 617 0.1205 sinv b1986 tRNA 101 0.1076 thrb1986 tRNA 131 0.1300 rv b1986 tRNA 127 0.1076 thrb1987 tRNA 143 0.1300 rv b1666 tRNA 1544 0.1172 trsH b0201 RIbosomal 15 0.1318 alf b0746 tRNA 134 0.1212 trsH b0201 RIbosomal 15 0.1312 alf b2403 tRNA 124 0.1212 trsG b3571 RIbosomal 17 0.1322 alf b2403 tRNA 129 0.1212 trsG b3775 RIbosomal 17 0.1322 alf b2307 tRNA 129 0.1215 trrf b3303 RIbosomal 17 0.1322 alf b2307 tRNA 233 0.1215 <td>InT</td> <td>b1977</td> <td>tRNA</td> <td>122</td> <td>0.1076</td> <td>leuP</td> <td>b4369</td> <td>tRNA</td> <td>404</td> <td>0.1294</td>	InT	b1977	tRNA	122	0.1076	leuP	b4369	tRNA	404	0.1294
snU b1986 FNA 101 0.1076 thrT b3979 FRNA 143 0.1300 nV b1939 FRNA 101 0.1076 serW b0833 FRNA 143 0.1310 nV b1636 FRNA 1544 0.1154 rrsA b3851 FIbosomal 15 0.1310 alf b0744 FRNA 1544 0.1154 rrsA b3851 FIbosomal 172 0.1318 alf b0744 FRNA 114 0.1212 rrsA b3851 FIbosomal 172 0.1318 alf b0746 FRNA 124 0.1212 rrsG b2591 FIbosomal 172 0.1318 alf b0206 FRNA 129 0.1212 rrsG b3775 FIBosomal 172 0.1328 alf b0206 FRNA 129 0.1215 rrfF b3275 FIBosomal 124 0.1338 sp10 b0206 FRNA	Wusi	b1984	tRNA	121	0.1076	selC	b3658	tRNA	617	0.1295
snv b1989 tRNA 127 0.1076 serW b0883 tRNA 196 0.1318 iV b3273 tRNA 1544 0.1172 trsH b0201 Rbosomal 15 0.1318 alf b0746 tRNA 114 0.1172 trsH b3377 tRNa 172 0.1318 alf b0746 tRNA 114 0.1212 trsH b3377 tRNa 172 0.1318 alf b0746 tRNA 124 0.1212 trsG b3377 tRNa 172 0.1318 alf b0206 tRNA 124 0.1212 trsG b3377 tRNa 172 0.1338 sp1 b0206 tRNA 126 0.1212 trsG b3377 tRNA 172 0.1338 sp1 b0206 tRNA 126 0.1221 trsG b3377 tRNA 172 0.1338 sp1 b0206 tRNA 1280	Unsi	b1986	tRNA	101	0.1076	thrT	b3979	tRNA	143	0.1300
IV b3273 FINA B98 0.1154 trsH b0201 Ribosomal 15 0.1318 alf b1666 FRNA 1844 0.1172 trsH b0201 Ribosomal 15 0.1318 alf b0746 FRNA 1344 0.1172 trsH b3377 Ribosomal 15 0.1318 alf b0206 FRNA 134 0.1212 trsG b33756 Ribosomal 15 0.1312 alf b2205 FRNA 124 0.1212 trsG b3756 Ribosomal 15 0.1328 alf b1975 FRNA 129 0.1215 trrH b0204 Ribosomal 14 0.1328 spU b0206 FRNA 233 0.1221 trrH b0204 Ribosomal 14 0.1338 spT b1975 FRNA 233 0.1229 trrH b0204 Ribosomal 15 0.1338 alv b1975 FRNA	Vusi	b1989	tRNA	127	0.1076	serW	b0883	tRNA	196	0.1310
all block fRNA 1544 0.1172 rrsA b3851 Ribosomal 1 0.1318 all b0744 RNA 114 0.1212 rrsG b3756 Ribosomal 172 0.1338 all b0746 fRNA 124 0.1212 rrsG b3756 Ribosomal 172 0.1338 all b2006 fRNA 129 0.1212 rrsG b3756 Ribosomal 172 0.1338 spU b0206 fRNA 129 0.1212 rrsG b3756 Ribosomal 173 0.1328 spU b0206 fRNA 129 0.1215 rrfF b3272 Ribosomal 14 0.1338 spV b0216 fRNA 1420 0.1215 rrfF b3272 Ribosomal 14 0.1338 spV b1655 fRNA 1420 0.1216 rrfF b3272 Ribosomal 14 0.1338 stV b1665 fRNA 219 0.1215 rrfF b3272 Ribosomal 14 0.1338 stV b1665 fRNA 2202 0.1233 rrfA b0666 fRNA 229 0.1351 alV b1665 fRNA 274 0.1229 mefU b0666 fRNA 229 0.1351 alV b1665 fRNA 274 0.1239 rrfC b3758 Ribosomal 12 0.1351 alV b1665 fRNA 274 0.1239 rrfC b3758 Ribosomal 18 0.1352 alV b1665 fRNA 274 0.1239 rrfC b3758 Ribosomal 18 0.1355 rrfC b3758 Ribosomal 18 0.1357 rrfV b2694 fRNA 236 0.1241 rrfD b32589 Ribosomal 18 0.1355 rrfV b2694 fRNA 130 0.1286 rrfG b3758 Ribosomal 18 0.1355 rrfV b2694 fRNA 236 0.1241 rrfD b3758 Ribosomal 18 0.1355 rrfV b2694 fRNA 100 0.1241 rrfD b3758 Ribosomal 18 0.1355 rrfV b2694 fRNA 101 0.1241 rrfD b3758 Ribosomal 18 0.1357 rrfV b2694 fRNA 101 0.1286 rrfG b3758 Ribosomal 18 0.1357 rrfV b2694 fRNA 101 0.1286 rrfG b3758 Ribosomal 18 0.1357 rrfV b2694 fRNA 1995 0.1286 rrfG b3758 Ribosomal 18 0.1357 rrfV b2694 fRNA 1995 0.1286 rrfG b3758 Ribosomal 10 0.1357 rrfV b2694 fRNA 1995 0.1286 rrfG b3758 Ribosomal 10 0.1357 rrfV b3798 rrfN 1995 0.1286 rrfC b304 0.1367 rrfN 136 0.1357 rrfN 130 0.1357 rf	hrV	b3273	tRNA	868	0.1154	rrsH	b0201	Ribosomal	15	0.1318
alf b0744 tRNA 114 0.1212 tyrU b3977 tRNA 172 0.1318 al2 b0746 tRNA 130 0.1212 rrsG b2591 Ribosomal 6 0.1319 al2 b0206 tRNA 124 0.1212 rrsG b2591 Ribosomal 2 0.1328 spU b0206 tRNA 219 0.1212 rrsG b1909 fRNA 73 0.1328 spV b0216 tRNA 219 0.1215 rrfF b3272 Ribosomal 14 0.1338 spV b1975 tRNA 1420 0.1216 rrfF b3272 Ribosomal 14 0.1338 e1U b1975 tRNA 1240 0.1221 metU b0666 tRNA 309 0.1351 alU b1975 tRNA 20204 RIbosomal 14 0.1358 alU b1975 tRNA 2022 0.1239 rr1A b0204 Ribosomal 14 0.1358 alU b1975 tRNA 2022 0.1239 rr1A b0204 Ribosomal 14 0.1358 alU b1975 tRNA 2022 0.1229 rr1A b0204 Ribosomal 14 0.1358 alU b1975 tRNA 2022 0.1239 rr1A b0204 Ribosomal 14 0.1358 alU b3376 tRNA 2022 0.1239 rr1A b0204 Ribosomal 14 0.1351 alU b3376 tRNA 2022 0.1239 rr1A b0204 Ribosomal 14 0.1351 alU b3376 tRNA 2022 0.1239 rr1A b0204 Ribosomal 12 0.1351 alU b3376 tRNA 2022 0.1239 rr1A b0204 Ribosomal 18 0.1355 rr1C b3758 Ribosomal 18 0.1355 rr1C b3758 Ribosomal 18 0.1355 rr1C b3758 Ribosomal 18 0.1355 rr1A b0375 tRNA 2023 rr1A b0745 tRNA 122 0.1357 rgV b2694 tRNA 765 0.1241 ysT b0745 tRNA 122 0.1357 rgV b2694 tRNA 3047 0.1285 yr1A 0.1256 ysY b0747 tRNA 122 0.1355 rgV b2694 tRNA 3047 0.1285 yr2A b0745 tRNA 123 0.1355 rgV b2694 tRNA 3047 0.1285 yr5A b0745 tRNA 123 0.1355 rgV b3745 tRNA 3047 0.1285 yr5A b0745 tRNA 131 0.1355 rgV b3745 tRNA 1995 0.1286 yr5A b0745 tRNA 131 0.1357 rgV b3545 tRNA 171 0.1287 proK b3545 tRNA 131 0.1367 0.1367	alW	b1666	tRNA	1544	0.1172	rrsA	b3851	Ribosomal	-	0.1318
al b0746 FNN 130 0.1212 rrsG b2591 Ribosomal 6 0.1319 al b2403 FNN 124 0.1212 rrsG b3756 Ribosomal 2 0.1320 al b0206 FNN 2129 0.1215 rrFG b3756 Ribosomal 2 0.1332 spU b0206 FNN 219 0.1215 rrFF b3123 misc RNA 73 0.1332 spU b0206 FNN 233 0.1215 rrFF b3272 Ribosomal 1 4 0.1338 erU b1975 FNN 1420 0.1221 metU b0666 FNN 309 0.1351 al b1975 FNN 233 0.1229 metT b0673 RNA 309 0.1351 al b1975 FNN 233 0.1229 metT b0673 RNA 309 0.1351 al b3376 FNN 2322 0.1223 rrF b0673 RNA 309 0.1351 al b3376 FNN 2322 0.1229 rrFG b33758 Ribosomal 1 4 0.1353 erU b3377 FNN 274 0.1229 rrFG b4009 Ribosomal 12 0.1351 al b3376 FNN 2360 0.1241 lysT b0743 FNN 229 0.1351 al b3376 FNN 2360 0.1241 lysT b0743 FNN 229 0.1351 al b3376 FNN 2360 0.1241 lysT b0743 FNN 123 0.1355 rrG b2691 FNN 236 0.1241 lysT b0743 FNN 123 0.1355 rrG b2691 FNN 3047 0.1256 lysY b0745 FNN 113 0.1357 al b3376 FNN 3047 0.1265 lysY b0745 FNN 113 0.1357 b04 b134 FNN 3047 0.1285 lysY b0745 FNN 131 0.1357 b01 b1379 FNN 3047 0.1285 lysY b0745 FNN 131 0.1357 b01 b1379 FNN 33798 FNN 131 0.1357 b01 b1379 FNN 3044 0.1367 0.1365 lysY b0745 FNN 3076 0.1367 b10 b1379 FNN 3076 0.1285 lysY b0745 FNN 3076 0.1367 b10 b1379 FNN 3076 0.1285 lysY b0745 FNN 3076 0.1367 b10 b1379 FNN 3076 0.1285 lysY b0745 FNN 3076 0.1367 b10 b1379 FNN 3076 0.1285 lysY b0745 FNN 3076 0.1367 b10 b1367 lysN b0745 FNN 3076 0.1367 b10 b1367 lysN b0745 FNN 3076 0.1367 0.1367 b10 b1367 lysN b0745 FNN 3076 0.1367 0.13	alT	b0744	tRNA	114	0.1212	tvrU	b3977	tRNA	172	0.1318
alX b2402 tRNA 124 0.1212 rrsC b3756 Ribosomal 2 0.1320 alY b22403 tRNA 109 0.1215 rrsC b3756 Ribosomal 2 0.1332 spU b0206 tRNA 233 0.1215 rrlH b0204 Ribosomal 14 0.1338 srU b1975 tRNA 150 0.1215 rrlH b0204 Ribosomal 14 0.1338 srU b1975 tRNA 150 0.1221 metU b0666 tRNA 309 0.1351 alV b1665 tRNA 233 0.1221 metU b0666 tRNA 309 0.1351 alV b1665 tRNA 232 0.1221 metU b0666 tRNA 309 0.1351 alV b1665 tRNA 236 0.1233 rrlA b3854 Ribosomal 12 0.1351 alV b1665 tRNA 236 0.1233 rrlA b3854 Ribosomal 12 0.1351 alV b1665 tRNA 236 0.1233 rrlA b3854 Ribosomal 12 0.1351 alV b1665 tRNA 236 0.1233 rrlA b3854 Ribosomal 12 0.1351 alV b1665 tRNA 236 0.1233 rrlA b3854 Ribosomal 12 0.1351 alV b1665 tRNA 236 0.1235 rrlA b3854 Ribosomal 12 0.1355 alV b2694 tRNA 765 0.1241 lysT b0743 tRNA 229 0.1357 alV b2694 tRNA 765 0.1241 lysT b0743 tRNA 123 0.1355 al0 b2691 tRNA 1483 0.1265 lysV b0745 tRNA 123 0.1355 al0 b1357 tRNA 1995 0.1285 lysV b0743 tRNA 123 0.1357 al1 b3798 tRNA 1995 0.1285 lysV b0743 tRNA 131 0.1357 beU b1354 tRNA 1995 0.1285 lysQ b0749 tRNA 131 0.1357 beU b1354 tRNA 1995 0.1285 lysQ b0749 tRNA 131 0.1357 beU b1378 tRNA 1995 0.1285 lysQ b0749 tRNA 131 0.1367 beU b1378 tRNA 171 0.1287 proK b3545 tRNA 0.1367 0.1367	alZ	b0746	tRNA	130	0.1212	rrsG	b2591	Ribosomal	9	0.1319
alY b2403 tRNA 109 0.1212 mpB b3123 misc_RNA 73 0.1332 spU b0206 tRNA 219 0.1215 iuZ b1909 tRNA 920 0.1338 spU b0206 tRNA 233 0.1215 rrfF b3272 Ribosomal 14 0.1338 spT b3760 tRNA 1420 0.1221 metU b0666 tRNA 920 0.1338 e1U b1656 tRNA 3814 0.1229 metU b0666 tRNA 3209 0.1351 nU b1656 tRNA 2022 0.1233 rrIA b0204 Ribosomal 14 0.1351 nU b3976 tRNA 3814 0.1229 metT b0573 tRNA 229 0.1351 eU b3277 tRNA 2022 0.1233 rrIC b4009 Ribosomal 18 0.1352 eU b3277 tRNA 274 0.1239 rrIC b3758 Ribosomal 18 0.1355 a10 b2691 tRNA 274 0.1239 rrIC b4009 Ribosomal 18 0.1355 rrIC b2692 tRNA 236 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 1483 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 3047 0.1256 lysY b0745 tRNA 131 0.1357 s b3065 tRNA 3047 0.1256 lysY b0745 tRNA 131 0.1357 neU b4134 tRNA 3047 0.1265 lysY b0748 tRNA 131 0.1357 neU b4134 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1357 neU b4134 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neU b4134 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neU b4134 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neU b4134 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neU b4134 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neV b2967 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neV b2967 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neV b2967 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neV b2967 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neV b3545 tRNA 1995 0.1286 lysQ b0749 tRNA 131 0.1367 neV b3545 tRNA 1304 0.1367 0.1367 lysQ b0749 tRNA 131 0.1367 lysG	alX	b2402	tRNA	124	0.1212	rrsC	b3756	Ribosomal	2	0.1320
spU b0206 FRNA 219 0.1215 Int b1909 FRNA 920 0.1338 spV b0216 FRNA 233 0.1215 rrff b3272 Ribosomal 14 0.1338 spV b1975 FRNA 150 0.1215 rrff b3272 Ribosomal 14 0.1338 sit b1975 FRNA 150 0.1215 rrff b3272 Ribosomal 14 0.1338 erU b1975 FRNA 120 0.1215 rrff b3272 Ribosomal 14 0.1338 alV b1665 FRNA 2012 0.1233 rrff b3275 Ribosomal 12 0.1351 alV b1666 FRNA 205 0.1233 rrff b3758 Ribosomal 12 0.1351 alV b3854 FRNA 205 0.1233 rrff b3758 Ribosomal 13 0.1351 alV b38565 FRNA	alY	b2403	tRNA	109	0.1212	rnpB	b3123	misc RNA	73	0.1332
spV b0216 tRNA 233 0.1215 rrff b3272 Ribosomal 14 0.1338 spT b3750 tRNA 150 0.1215 rrff b3272 Ribosomal 14 0.1338 erU b1975 tRNA 150 0.1215 rrff b0204 Ribosomal 4 0.1348 erU b1975 tRNA 1420 0.1223 rrff b0666 tRNA 309 0.1351 aU b3174 tRNA 3814 0.1233 rrff b0673 tRNA 309 0.1351 aU b3377 tRNA 2022 0.1233 rrfd b3354 Ribosomal 12 0.1351 aU b3377 tRNA 203 n12 b3375 Ribosomal 12 0.1355 aU b3352 tRNA 236 0.1241 tyst b0743 tRNA 12 0.1357 aU b3369 p1800 p1241 tyst </td <td>spU</td> <td>b0206</td> <td>tRNA</td> <td>219</td> <td>0.1215</td> <td>IeuZ</td> <td>b1909</td> <td>tRNA</td> <td>920</td> <td>0.1338</td>	spU	b0206	tRNA	219	0.1215	IeuZ	b1909	tRNA	920	0.1338
spT b3760 tRNA 150 0.1215 rrlH b0204 Ribosomal 4 0.1351 erU b1975 tRNA 1420 0.1215 rrlH b0073 tRNA 309 0.1351 aU b1975 tRNA 1420 0.1221 metU b0666 tRNA 309 0.1351 aU b3174 tRNA 2222 0.1233 rrlA b3854 Ribosomal 12 0.1351 aV b1665 tRNA 2022 0.1235 rrlA b3854 Ribosomal 12 0.1351 aV b1665 tRNA 2022 0.1233 rrlA b3854 Ribosomal 12 0.1351 aV b1665 tRNA 274 0.1239 rrlC b3854 Ribosomal 12 0.1352 eU b3352 tRNA 274 0.1239 rrlD b3275 Ribosomal 12 0.1357 rgQ b2694 tRNA <td< td=""><td>SpV</td><td>b0216</td><td>tRNA</td><td>233</td><td>0.1215</td><td>Ŧ</td><td>b3272</td><td>Ribosomal</td><td>14</td><td>0.1338</td></td<>	SpV	b0216	tRNA	233	0.1215	Ŧ	b3272	Ribosomal	14	0.1338
erU b1975 FRNA 1420 0.1221 metU b0666 FRNA 309 0.1351 aU b3174 FRNA 3814 0.1229 metT b0673 FRNA 229 0.1351 aV b1665 FRNA 2022 0.1233 rrIA b3854 Ribosomal 12 0.1351 hrU b3976 FRNA 2022 0.1235 rrIC b3758 Ribosomal 12 0.1352 eU b3277 FRNA 274 0.1239 rrIG b3758 Ribosomal 18 0.1352 eU b3277 FRNA 274 0.1239 rrIG b3758 Ribosomal 18 0.1352 rrIC b3758 Ribosomal 18 0.1352 rrIC b3758 Ribosomal 18 0.1352 rrIG b2691 FRNA 274 0.1239 rrIG b3758 Ribosomal 30 0.1355 rg2 b2692 FRNA 274 0.1239 rrIG b32589 Ribosomal 30 0.1355 rg2 b2692 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1355 rg2 b2694 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1355 rg1 b3261 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1355 rg1 b2694 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1355 rg1 b2694 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1355 rg1 b2694 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1355 rg1 b2067 FRNA 401 0.1241 rrID b3275 Ribosomal 30 0.1357 rg1 b376 rFNA 1483 0.1265 rys1 b0743 FRNA 123 0.1357 heU b4134 FRNA 1483 0.1276 rys2 b0748 FRNA 113 0.1357 heU b4134 FRNA 171 0.1285 rys0 b0749 FRNA 113 0.1357 heU b4134 FRNA 171 0.1285 rys0 b0749 FRNA 113 0.1367 rg1 b3376 rFNA 33645 rFNA 0.1367	spT	b3760	tRNA	150	0.1215	HIT	b0204	Ribosomal	4	0.1348
uU b3174 tRNA 3814 0.1229 metT b0673 tRNA 229 0.1351 alV b1665 tRNA 3814 0.1229 metT b0673 tRNA 229 0.1351 alV b1665 tRNA 2022 0.1233 rrIA b3854 Ribosomal 12 0.1351 nrU b3376 tRNA 2022 0.1235 rrIA b3854 Ribosomal 12 0.1352 eU b33277 tRNA 274 0.1239 rrIC b3758 Ribosomal 18 0.1352 eU b3275 tRNA 236 0.1239 rrIC b3757 Ribosomal 30 0.1357 rgO b2692 tRNA 255 0.1241 rrID b3275 Ribosomal 30 0.1357 rgV b2694 tRNA 765 0.1241 lysT b0745 tRNA 123 0.1357 rgV b2694 tRNA <td< td=""><td>erU</td><td>b1975</td><td>tRNA</td><td>1420</td><td>0.1221</td><td>metU</td><td>b0666</td><td>tRNA</td><td>309</td><td>0.1351</td></td<>	erU	b1975	tRNA	1420	0.1221	metU	b0666	tRNA	309	0.1351
alV b1665 tRNA 2022 0.1233 rrIA b3854 Ribosomal 12 0.1351 rrU b3976 tRNA 66 0.1235 rrIC b3758 Ribosomal 12 0.1352 eU b3277 tRNA 274 0.1235 rrIC b3758 Ribosomal 18 0.1352 eU b3277 tRNA 274 0.1239 rrIG b4009 Ribosomal 3 0.1355 rgC b2691 tRNA 236 0.1241 rrID b3275 Ribosomal 3 0.1355 rgZ b2692 tRNA 465 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 lysV b0745 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 lysV b0745 tRNA 123 0.1357 bV b2694 tRNA 3047 0.1255 lysV b0747 tRNA 94 0.1357 bV b2967 tRNA 3047 0.1285 lysV b0747 tRNA 94 0.1357 bv b2967 tRNA 1995 0.1285 lysQ b0749 tRNA 131 0.1357 bv b2967 tRNA 1995 0.1285 lysQ b0749 tRNA 131 0.1357 bv b2967 tRNA 1995 0.1285 lysQ b0749 tRNA 131 0.1357 bv b2967 tRNA 171 0.1287 proK b3545 tRNA 0.108 0.1367	UU	b3174	tRNA	3814	0.1229	metT	b0673	tRNA	229	0.1351
hrU b3976 tRNA 66 0.1235 rrIC b3758 Ribosomal 18 0.1352 eU b3277 tRNA 274 0.1235 rrIC b3758 Ribosomal 18 0.1352 eU b3277 tRNA 274 0.1239 rrIC b3758 Ribosomal 5 0.1352 rgO b2691 tRNA 274 0.1239 rrIC b3275 Ribosomal 3 0.1354 rgO b2692 tRNA 455 0.1241 rrID b3275 Ribosomal 30 0.1355 rgV b2694 tRNA 765 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 1483 0.1241 lysV b0745 tRNA 123 0.1357 s b13761 tRNA 1483 0.12565 lysV b0747 tRNA 142 0.1357 s b33761 tRNA 1483<	alV	b1665	tRNA	2022	0.1233	rrIA	b3854	Ribosomal	12	0.1351
eU b3277 tRNA 274 0.1239 rrIE b4009 Ribosomal 5 0.1352 eT b3852 tRNA 236 0.1239 rrIG b2589 Ribosomal 3 0.1354 rgQ b2691 tRNA 455 0.1241 rrID b3275 Ribosomal 3 0.1355 rgV b2694 tRNA 401 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 lysW b0745 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 lysW b0745 tRNA 123 0.1357 bT b3761 tRNA 1483 0.1255 lysY b0747 tRNA 94 0.1357 beV b2967 tRNA 3047 0.1285 lysQ b0749 tRNA 131 0.1357 beU b4134 tRNA 1995 0.1285 lysQ b0749 tRNA 113 0.1357 beU b4134 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1361	hrU	b3976	tRNA	99	0.1235	rrIC	b3758	Ribosomal	18	0.1352
eT b3852 tRNA 236 0.1239 rrIG b2589 Ribosomal 3 0.1354 rgQ b2691 tRNA 255 0.1241 rrID b3275 Ribosomal 30 0.1355 rgQ b2692 tRNA 755 0.1241 IrrID b3275 Ribosomal 30 0.1355 rgV b2694 tRNA 765 0.1241 IysT b0745 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 IysT b0745 tRNA 123 0.1357 rgV b2694 tRNA 18NA 1241 IysV b0745 tRNA 123 0.1357 s b0455 misc_RNA 686 0.1255 IysV b0747 tRNA 142 0.1357 pV b2967 tRNA 1483 0.1276 IysQ b0749 tRNA 131 0.1357 heU b4134 tRNA 194	eU	b3277	tRNA	274	0.1239	rIE	b4009	Ribosomal	S	0.1352
rgQ b2691 tRNA 455 0.1241 rrID b3275 Ribosomal 30 0.1355 rgZ b2692 tRNA 765 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 lysW b0745 tRNA 142 0.1357 s b0455 misc_RNA 686 0.1255 lysY b0747 tRNA 94 0.1357 heV b2967 tRNA 1483 0.1276 lysZ b0748 tRNA 131 0.1357 heV b2967 tRNA 1995 0.1285 lysQ b0749 tRNA 113 0.1357 heU b4134 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367 0.1367 heU b4134 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367 0.1367 heU b4134 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367 0.1367 heU b4134 tRNA 171 0.1287 heV b3545 tRNA 3076 0.1367 0.1367 heU b4134 tRNA 171 0.1287 heV b3545 tRNA 3076 0.1367 0.1367 0.1367 heV b3545 tRNA 3076 0.1367 0.1367 0.1367 heV b3545 tRNA 3076 0.1367 0.	eT	b3852	tRNA	236	0.1239	rrlG	b2589	Ribosomal	e	0.1354
rigZ b2692 tRNA 765 0.1241 lysT b0743 tRNA 123 0.1357 rgV b2694 tRNA 401 0.1241 lysW b0745 tRNA 142 0.1357 s b0455 misc_RNA 686 0.1255 lysY b0747 tRNA 94 0.1357 pT b3761 tRNA 1483 0.1276 lysZ b0748 tRNA 131 0.1357 heV b2967 tRNA 3047 0.1285 lysQ b0749 tRNA 113 0.1357 heU b4134 tRNA 1995 0.1285 dicF b1574 RNA; Cell 3004 0.1367 auT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	IngQ	b2691	tRNA	455	0.1241	rrID	b3275	Ribosomal	30	0.1355
rgV b2694 tRNA 401 0.1241 lysW b0745 tRNA 142 0.1357 s b0455 misc_RNA 686 0.1255 lysY b0747 tRNA 94 0.1357 pT b3761 tRNA 1483 0.1276 lysZ b0748 tRNA 131 0.1357 heV b2967 tRNA 3047 0.1285 lysQ b0749 tRNA 113 0.1357 heU b4134 tRNA 1995 0.1285 dicF b1574 RNA; Cell 3004 0.1361 suT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	Zpr	b2692	tRNA	765	0.1241	lvsT	b0743	tRNA	123	0.1357
s b0455 misc_RNA 686 0.1255 lýsY b0747 tRNA 94 0.1357 pT b3761 tRNA 1483 0.1276 lýsZ b0748 tRNA 131 0.1357 heV b2967 tRNA 3047 0.1285 lýsQ b0749 tRNA 113 0.1357 heU b4134 tRNA 1995 0.1285 dícF b1574 RNA; Cell 3004 0.1361 suT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	rgV	b2694	tRNA	401	0.1241	WsW	b0745	tRNA	142	0.1357
pT b3761 tRNA 1483 0.1276 lýsZ b0748 tRNA 131 0.1357 heV b2967 tRNA 3047 0.1285 lýsQ b0749 tRNA 113 0.1357 heU b4134 tRNA 1995 0.1285 dícF b1574 RNA; Cell 3004 0.1361 suT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	0	b0455	misc RNA	686	0.1255	lysy	b0747	tRNA	94	0.1357
heV b2967 tRNA 3047 0.1285 lysQ b0749 tRNA 113 0.1357 heU b4134 tRNA 1995 0.1285 dicF b1574 RNA; Cell 3004 0.1361 suT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	pTa	b3761	tRNA_	1483	0.1276	Zsv	b0748	tRNA	131	0.1357
heU b4134 tRNA 1995 0.1285 dicF b1574 RNA; Cell 3004 0.1361 euT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	Veh	b2967	tRNA	3047	0.1285	lysQ	b0749	tRNA	113	0.1357
euT b3798 tRNA 171 0.1287 proK b3545 tRNA 3076 0.1367	beu	b4134	tRNA	1995	0.1285	dicF	b1574	RNA; Cell	3004	0.1361
	euT	b3798	tRNA	171	0.1287	proK	b3545	tRNA	3076	0.1367



Conclusions

- Absolute gene expression levels are highly correlated with low PP in multiple microbial genomes.
- PP may be exploited for predicting the expression of nontranslated genes and highly expressed genes in slow growing microbes.
- Genes often encoded by DNA with low position preference values were mostly involved in 'translation, ribosomal structure and biogenesis', 'energy production and conversion', and transcription.
- For pathogens and microbes living in extreme environments, the predominant functional category was 'replication, recombination and repair'.

References

- Willenbrock & Ussery (2007) Prediction of highly expressed genes in microbes based on chromatin accessibility. *BMC Mol Biol* 8:11
- Brukner et al (1995) Sequence-dependent bending propensity of DNA as revealed by DNase I: parameters for trinucleotides. *EMBO J*, 14:1812-1818
- Willenbrock & Ussery (2004) Chromatin architecture and gene expression in *Escherichia coli*. *Gen.Biol*. 5:252
- Postow et al (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes and Dev.* 18: 1766-1779