ELSEVIER

FEMS Microbiology Letters 181 (1999) 17–23

FEMS MICROBIOLOGY LETTERS

Genetic rearrangements in the *tyrB-uvrA* region of the enterobacterial chromosome: a potential cause for different class B acid phosphatase regulation in *Salmonella enterica* and *Escherichia coli*

Maria Cristina Thaller [a,b,*], Serena Schippa [b], Alessandra Bonci [c], Francesca Berlutti [b], Laura Selan [b], Gian Maria Rossolini [c]

[a] *Dipartimento di Biologia, II Università di Roma 'Tor Vergata', 00133 Rome, Italy*
[b] *Istituto di Microbiologia, Università di Roma 'La Sapienza', 00185 Rome, Italy*
[c] *Dipartimento di Biologia Molecolare, Sezione di Microbiologia, Università di Siena, 53100 Siena, Italy*

Received 13 September 1999; accepted 20 September 1999

Abstract

Unlike in *Escherichia coli*, in *Salmonella enterica* production of class B acid phosphatase (AphA) was detectable also in cells growing in the presence of glucose. Characterization of the *aphA* locus from a *S. enterica* ser. typhi strain showed that the *aphA* determinant is very similar to the *E. coli* homolog, and that its chromosomal location between the highly conserved *tyrB* and *uvrA* genes is retained. However, the *aphA* flanking regions were found to be markedly different in the two species, either between *tyrB* and *aphA* or between *aphA* and *uvrA*. The differences in the *aphA* 5′-flanking region, which in *S. enterica* is considerably shorter than in *E. coli* (183 vs. 1121 bp) and includes potential promoter sequences not present in *E. coli*, could be responsible for the different regulation of class B acid phosphatase observed in the two species.

Keywords: *Salmonella enterica*; Acid phosphatase; Regulation; Enterobacterial chromosome; Genetic rearrangement

1. Introduction

Bacterial class B acid phosphatases are periplasmic phosphohydrolases that exhibit a broad substrate profile, including 5′- and 3′-nucleotides, phenylphosphates, and other organic phosphoesters. The native enzymes are 100-kDa proteins made of four 25-kDa polypeptide subunits, and require the presence of a metal co-factor for activity [1–5]. The first enzyme of this family was purified from *Salmonella enterica* ser. typhimurium LT2, and was named AphA or nonspecific acid phosphatase II (NAP-II) to differentiate it from another NAP (the PhoN enzyme or NAP-I) produced by the same *Salmonella* strain [2,6]. Production of class B NAP was subsequently found to be widespread among enteric bac-

* Corresponding author. Tel.: +39 (6) 445 1324; Fax: +39 (6) 4991 4638; E-mail: pezzi@axrma.uniroma1.it

0378-1097/99/$20.00
PII: S0378-1097(99)00509-1

teria [1,7], and sequencing of some class B NAP genes revealed a highly conserved primary structure for these [1,4,5].

The physiological role of these enzymes remains unclear. When first described, the *Salmonella* class B NAP was considered a possible substitute for the 5′-nucleotidase and/or alkaline phosphatase that are not produced by *S. enterica* ser. typhimurium LT2 [2]. However, it was subsequently shown that production of a class B NAP is also detectable in species that are known to produce both 5′-nucleotidase and alkaline phosphatase activities [7–9]. On the other hand, an involvement of the *Escherichia coli* class B NAP in regulation of chromosomal replication has been recently proposed on the basis of its binding activity to hemi-methylated *oriC* DNA [10].

The *E. coli* class B NAP is also the only member of this family for which regulation has been analyzed in some detail. Its production appears to be influenced by the carbon source, being detectable only when cells are growing on carbon sources other than glucose [11]. Although the existence of similar or different regulatory patterns has not been specifically investigated in other species, the fact that the class B NAP of *S. enterica* ser. typhimurium was purified from cells grown in basal salt-glucose media [2–3] (i.e. under conditions that are not permissive for production of the *E. coli* homolog [11]) suggests that in *S. enterica* regulation of the class B NAP could be different from that observed in *E. coli*.

In this work we studied the influence of glucose on the regulation of class B NAP in *S. enterica*, and analyzed the structure of the *aphA* locus of this species in comparison with that of *E. coli*.

2. Materials and methods

2.1. Bacterial strains and culture conditions

S. enterica ser. typhi Sty4 [12] was used to investigate class B NAP regulation under different growth conditions, and as a source for isolation of the *aphA* locus. *E. coli* HB101 and DH5α [13] were used as hosts for recombinant plasmids. Luria-Bertani broth (LB) [13] supplemented with 50 mM inorganic phosphate (in the form of Na/K phosphate buffer) was used as a culture medium, supplemented with either 55 mM glucose (LB-U) or 135 mM glycerol (LB-Y). The final pH of the media was 7.2. In LB-U, glucose was added again when its concentration fell below 10 mM. Glucose concentration in media and culture supernatants was assayed using a commercial kit (Glucose, Sigma Diagnostics). Bacteria were always grown aerobically at 37°C.

2.2. Protein electrophoretic techniques

Zymogram detection of the AphA enzyme produced by *S. enterica* after renaturing SDS-PAGE was performed as previously described [11]. NaF (1 mM) was added to the equilibration buffer to inhibit the PhoN activity [6,7]. Phosphatase activity in zymogram bands was quantitated by scanning the developed gels with an ULTROSCAN XL Laser Densitometer (LKB) as previously described [11].

2.3. Recombinant DNA methodology, DNA sequencing and sequence analysis

Basic recombinant DNA methodology was performed essentially as described [13]. Construction of the *S. enterica* genomic library in the cosmid vector cosKT1 has been previously described [12]. The DNA probe used for screening the library was a PCR-generated amplicon containing the coding sequence of the *Morganella morganii napA* gene [1]. Screening was performed under low stringency hybridization conditions (final washes in 2×SSC, 0.1% SDS, at 60°C). Plasmid SK Bluescript (Stratagene) was used for subcloning procedures. DNA sequencing was performed on both strands by the dideoxy chain termination method [13] and custom oligonucleotides as sequencing primers. Computer analysis of sequence data was performed with an updated release (8.0.1) of the University of Wisconsin Genetic Computer Group program [14]. BLAST searches were performed at the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov), on both finished and unfinished bacterial genomes. Preliminary sequence data on the *Yersinia pestis* genome were obtained from the *Yersinia pestis* Sequencing Group at the Sanger Center website (http://www.sanger.ac.uk/DataSearch/). Sequence alignments were performed using the CLUSTAL W program [15]. Screening for

promoter sequences was performed by the Promoter Prediction by Neural Network tool [16] (website: http://www-hgc.lbl.gov/projects/promoter.html). The sequence has been submitted to the EMBL/GenBank sequence database and assigned the accession number AJ237788.

3. Results

3.1. Influence of glucose on the production of class B NAP in S. enterica

Production of the AphA enzyme by *S. enterica* ser. typhi Sty4 growing in either LB-U or LB-Y medium was assayed by means of a quantitative zymogram technique. Results of these experiments showed that production of AphA was detectable when cells were growing in the presence of glucose, although the amount of enzyme produced in the glucose-containing medium was overall lower than that produced in the glycerol-containing medium (Fig. 1). The pattern of AphA production was similar with either carbon source. AphA was already detectable in the early exponential phase of growth and its amount increased during the exponential phase. The highest values were observed when the cultures entered the stationary phase (Fig. 1).

3.2. Structure of the aphA locus of S. enterica

The region containing the *aphA* determinant was isolated from a genomic library of *S. enterica* ser. typhi Sty4 by screening the library with a probe con-

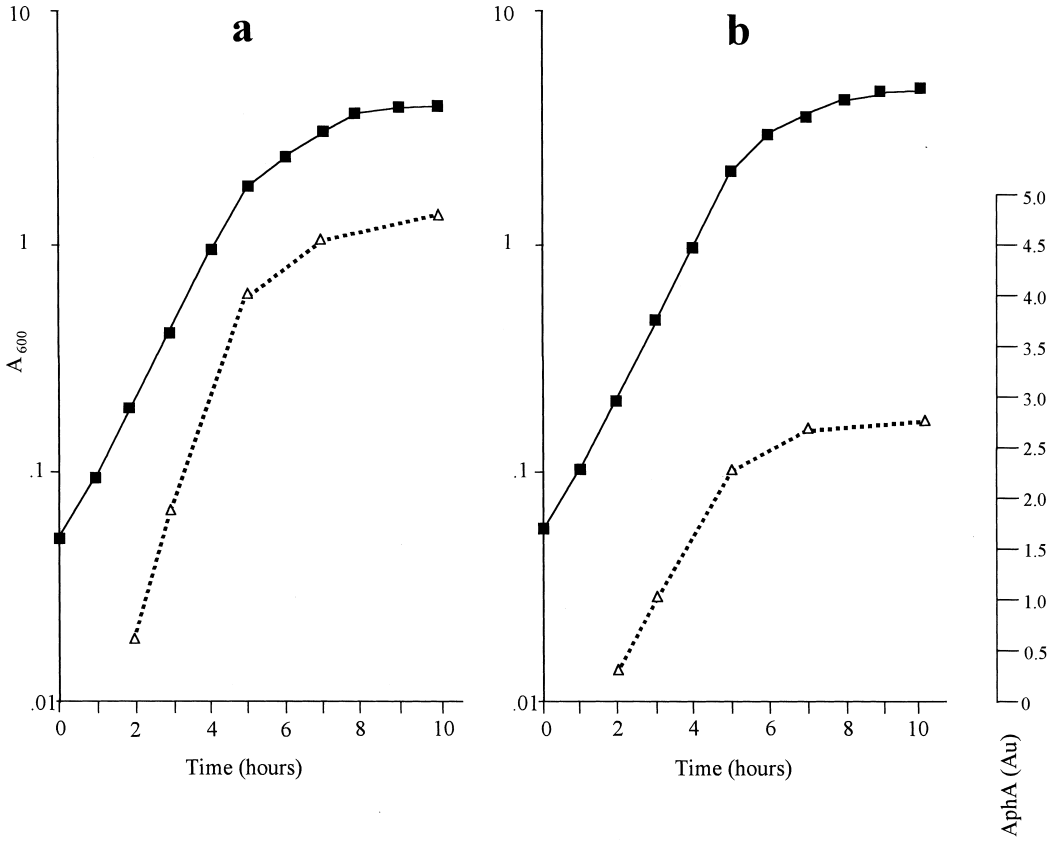

Fig. 1. Production of the AphA enzyme by *S. enterica* ser. typhi Sty4 growing in LB-U (a) or in LB-Y (b) medium. The microbial growth curve is reported (■) together with the amount of AphA activity (△) detected at different times. The AphA activity, expressed in absorbance units (Au), was detected by zymograms and measured by densitometric scanning of the developed gels as previously described [11]. The same amount of cells (0.4 OD_{600} units) was loaded for each sample.

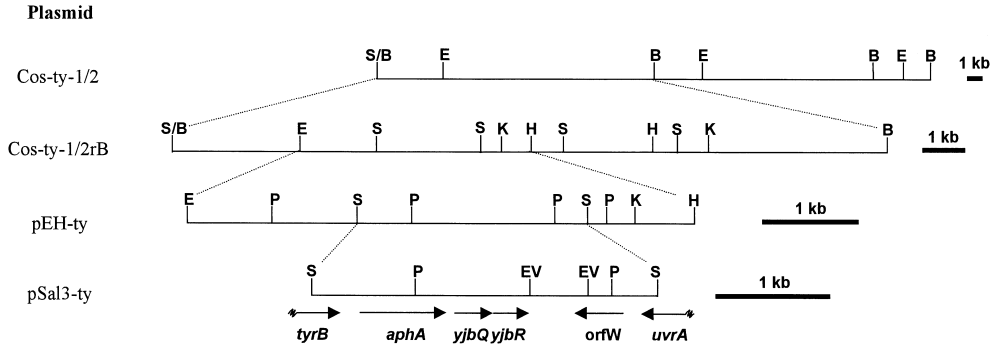

Fig. 2. Restriction map of the DNA insert of the recombinant cosmid Cos-ty-1/2, and subcloning strategy. Plasmid Cos-ty-1/2rB was obtained by removal of the 20-kb *Bam*HI fragment from Cos-ty-1/2. Plasmids pEH-ty and pSal3-ty are pBluescript SK derivatives carrying subcloned fragments. Abbreviations: B, *Bam*HI; E, *Eco*RI; EV, *Eco*RV; H, *Hin*dIII; K, *Kpn*I; P, *Pst*I; S, *Sal*I; S/B, *Sau*3AI/*Bam*HI junction. The ORFs identified in the insert of pSal3-ty are shown by arrows (interrupted arrows indicate partial ORFs). The presence of the *aphA* determinant in the subcloned fragments was monitored by hybridization experiments with the same probe used for screening of the library.

taining the class B NAP gene of *M. morganii* [1], under low-stringency hybridization conditions. A recombinant cosmid (Cos-ty-1/2) recognized by the probe was identified, which contained a 40-kb fragment of the *S. enterica* chromosome. Subcloning analysis mapped the *aphA* determinant within a 3-kb *Sal*I fragment (Fig. 2).

The nucleotide sequence of the insert of plasmid pSal3-ty was determined (Fig. 3). The *aphA* open reading frame (ORF), identified on the basis of sequence similarity with other class B NAP determinants, is 711 bp long and encodes a polypeptide that exhibits high-level sequence homology with the class B NAPs of *E. coli* [4] (94% identical residues) and of *M. morganii* [1] (79% identical residues). The base composition of the *Salmonella aphA* gene (G+C = 50.7%) is in agreement with the average base composition of the *Salmonella* chromosome [17]. The codon usage pattern of *aphA* was found to be similar to that of other *Salmonella* coding regions (D^2 value lower than 3 in a comparison with 46 kb of *Salmonella* coding regions for housekeeping proteins).

In the *Salmonella* chromosome, the *aphA* determinant is located between the *tyrB* and *uvrA* genes (Fig. 2). The *aphA* ORF has the same orientation as *tyrB*, starting 183 bp downstream of the *tyrB* stop codon and terminating 1817 bp downstream of the *uvrA* stop codon. In the long *aphA-uvrA* intergenic region three ORFs encoding hypothetical proteins of unknown function were identified. Two of them are apparently arranged in an operon structure, and their products exhibited significant similarity with the YjbQ and YjbR hypothetical proteins encoded by the *E. coli* and other bacterial genomes in a BLAST search. The third ORF (named orfW) has the same orientation as *uvrA*, starting 143 bp downstream of the *uvrA* stop codon and terminating 325 bp downstream of *yjbR*. It is preceded by a ribosomal binding site and followed by a 16-bp palindromic sequence that could function as a transcriptional terminator. The hypothetical orfW product, which contains an amino-terminal sequence carrying a motif typical of prokaryotic lipoprotein signal sequences, did not show any significant homology with any other (hypothetical) bacterial protein in a BLAST search.

3.3. Comparison of the aphA locus of S. enterica with that of E. coli

The *aphA* gene of *S. enterica* has overall the same chromosomal location as the *E. coli* homolog [7,18]. However, the regions flanking *aphA* and comprised between the highly conserved *tyrB* and *uvrA* enterobacterial genes are noticeably different in the two species.

The *tyrB-aphA* intergenic region, which in *Salmonella* consists of only 183 bp, in *E. coli* is 1106 bp long. This difference appears to be related to the presence, in *E. coli*, of an extra region of 921 bp

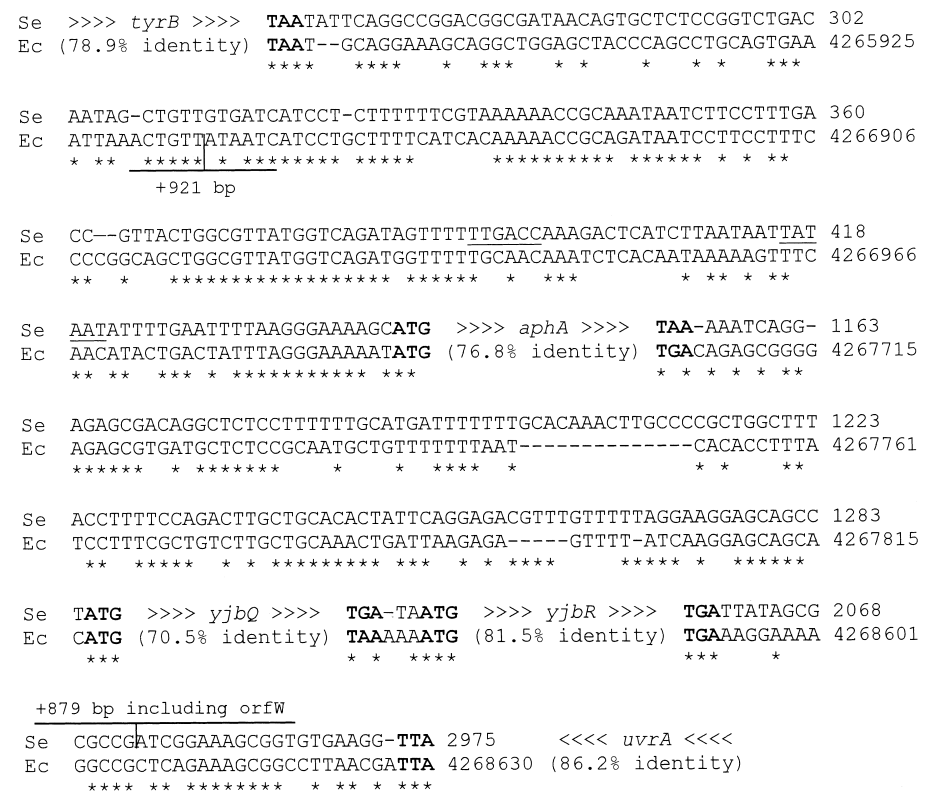

Fig. 3. Sequence alignment of the *tyrB-uvrA* regions of *S. enterica* (Se) and *E. coli* (Ec). The numbering refers to GenBank/EMBL entries AJ237788 and U00096 for the *S. enterica* and *E. coli* sequences, respectively. Only the alignment relative to the intergenic regions is shown in the figure, together with the initiation and termination codons of the various ORFs (in boldface). The degree of sequence identity between the various ORFs is also indicated. The −35 and −10 hexamers of the putative promoter identified upstream of *aphA* in *S. enterica* are underlined. Identical residues are indicated by an asterisk.

inserted approximately 130 bp upstream of the *aphA* coding sequence (Fig. 3), which has no traceable counterparts in the *Salmonella* chromosome. In the 130-bp region upstream of the *aphA* coding sequences a relatively strong sequence similarity is retained between the two species, but some discrete patches of sequence divergence are also detectable (Fig. 3). Scanning these regions for putative promoter sequences revealed the presence of a high-scoring putative promoter in *S. enterica*, which was not detectable in *E. coli* (Fig. 3).

A somewhat reversed situation is encountered in the *aphA-uvrA* intergenic region, which in *Salmonella* consists of 1817 bp and in *E. coli* of only 921 bp. This difference appears to be related to the presence, in *Salmonella*, of an extra region of 879 bp, inserted approximately 20 bp downstream of the *uvrA* termination codon (Fig. 3). This region contains an ORF (orfW) that potentially encodes a 17-kDa putative membrane lipoprotein and that apparently has no close neighbors in the *E. coli* or in other sequenced bacterial genomes.

4. Discussion

Results of this study revealed a different regulation of the class B NAP of *S. enterica*, compared to that of *E. coli*, in relation to the carbon source. In fact, unlike in *E. coli* [11], in *S. enterica* the AphA enzyme was detectable also when cells were growing in the presence of glucose, although production appeared to be somewhat enhanced in the presence of a different carbon source. A similar behavior is likely re-

lated to the structural differences existing in the region upstream of *aphA* in the two species, resulting in different promoter sequences. It would be interesting to further investigate this point by comparative transcription mapping experiments in the two species.

Comparison of the structure of the *aphA* loci of *S. enterica* and *E. coli* showed that, although the *aphA* homologs of the two species are highly conserved and maintain the same chromosomal location, genetic rearrangements have occurred in the *tyrB-uvrA* genomic region both upstream and downstream of *aphA*, since the divergence of the two species from a common ancestor. Apart from *E. coli* and *S. enterica*, the only additional enterobacterial chromosome for which the *tyrB-uvrA* region is currently known is that of *Y. pestis*. In this species the *tyrB* and *uvrA* genes and their upstream regions appear to be substantially conserved compared to those of *E. coli* and *S. enterica*, but the *tyrB-uvrA* region further differs from those of the latter species, containing only a *yjbR* homolog which is separated from *tyrB* by a 0.5-kb intergenic region that does not present any significant sequence similarity with either the *E. coli* or the *S. enterica tyrB-aphA* intergenic regions. Similar findings suggest that considerable rearrangements have occurred in the *tyrB-uvrA* region during the evolution of the enterobacterial chromosome, which could have affected both the presence and the regulation of class B NAP determinants, leading to the heterogeneous pattern of class B NAP production that is currently observed among Enterobacteriaceae [5,7].

If the involvement in chromosomal DNA replication proposed for the class B NAP of *E. coli* [10] should be retained also in other species, it would be interesting to understand the significance of the heterogeneous pattern of class B NAP gene expression observed among Enterobacteriaceae to this fundamental step of microbial physiology.

Acknowledgements

Preliminary sequence data were obtained from The Institute for Genomic Research website (http://www.tigr.org), and from the *Yersinia pestis* Sequencing Group, at the Sanger Center website (http://www.sanger.ac.uk/DataSearch/).

References

[1] Thaller, M.C., Lombardi, G., Berlutti, F., Schippa, S. and Rossolini, G.M. (1995) Cloning and characterization of the NapA acid phosphatase/phosphotransferase of *Morganella morganii*: identification of a new family of bacterial acid phosphatase-encoding genes. Microbiology 141, 147–154.

[2] Uerkvitz, W. and Beck, C.F. (1981) Periplasmic phosphatases in *Salmonella typhimurium* LT2. J. Biol. Chem. 256, 382–389.

[3] Uerkvitz, W. (1988) Periplasmic non specific acid phosphatase II from *Salmonella typhimurium* LT2. J. Biol. Chem. 263, 15823–15830.

[4] Thaller, M.C., Schippa, S., Bonci, A., Cresti, S. and Rossolini, G.M. (1997) Identification of the gene (*aphA*) encoding the class B acid phosphatase/phosphotransferase of *Escherichia coli* MG1655 and characterization of its product. FEMS Microbiol. Lett. 146, 191–198.

[5] Rossolini, G.M., Schippa, S., Riccio, M.L., Berlutti, F., Macaskie, L.E. and Thaller, M.C. (1998) Bacterial nonspecific acid phosphohydrolases: physiology, evolution, and use as tools in microbial biotechnology. Cell. Mol. Life Sci. 54, 833–850.

[6] Weppelman, R., Kier, L.D. and Ames, B.N. (1977) Properties of two phosphatases and a cyclic phosphodiesterase of *Salmonella typhimurium*. J. Bacteriol. 130, 411–419.

[7] Thaller, M.C., Berlutti, F., Schippa, S., Iori, P., Passariello, C. and Rossolini, G.M. (1995) Heterogeneous patterns of acid phosphatases containing low-molecular-mass polypeptides in members of the family *Enterobacteriaceae*. Int. J. Syst. Bacteriol. 45, 255–261.

[8] Cocks, G.T. and Wilson, A.C. (1972) Enzyme evolution in the *Enterobacteriaceae*. J. Bacteriol. 110, 793–802.

[9] Neu, H.C. (1968) The 5′-nucleotidases and cyclic phosphodiesterases (3′-nucleotidases) of the *Enterobacteriaceae*. J. Bacteriol. 95, 1732–1737.

[10] Reshetnyak, E., d'Alencon, E., Kern, R., Taghbalout, A., Guillaud, P. and Kohiyama, M. (1999) Hemi-methylated oriC DNA binding activity found in non-specific acid phosphatase. Mol. Microbiol. 31, 167–175.

[11] Rossolini, G.M., Thaller, M.C., Pezzi, R. and Satta, G. (1994) Identification of an *Escherichia coli* periplasmic acid phosphatase containing a 27-kDa polypeptide component. FEMS Microbiol. Lett. 118, 167–174.

[12] Rossolini, G.M., Muscas, P., Chiesurin, A. and Satta, G. (1993) Molecular cloning and expression in *Escherichia coli* of the *Salmonella typhi* gene cluster coding for type 1 fimbriae. In: NATO ASI Series Volume: The Biology of *Salmonella* (Cabello, F. et al., Eds.), pp. 408–412, Plenum, New York.

[13] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

[14] Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12, 387–395.
[15] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.
[16] Reese, M.G., Harris, N.L. and Eeckman, F.H. (1996) Large scale sequencing specific neural networks for promoter and splice site recognition. Proc. 1996 Pacific Symposium (Hunter, L. and Klein, T.E., Eds.) World Scientific, Singapore.
[17] Ochman, H. and Lawrence, J.G. (1996) Phylogenetics and the amelioration of bacterial genomes. In: *Escherichia coli* and *Salmonella* Cellular and Molecular Biology (Neidhardt, F.C. et al., Eds.), pp. 2627–2637, American Society for Microbiology, Washington, DC.
[18] Blattner, F.R., Burland, V., Plunkett III, G., Sofia, H.J. and Daniels, D.L. (1993) Analysis of the *Escherichia coli* genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes. Nucleic Acids Res. 21, 5408–5417.