Figure 1. Map of the pLVPK plasmid. The circle map shows the major features of the plasmid including the replication origins (*ori*), partition regions (*par, sop*), major gene clusters, and *IS* elements (*IS*). The positions of some of the ORFs and gene clusters are depicted in box and their contents labeled. Gene clusters which are homologous unknown gene clusters reported from other species are labeled with the species name respectively. The G+C % contents along the plasmid are shown in colors from low % (red) to high % (purple).

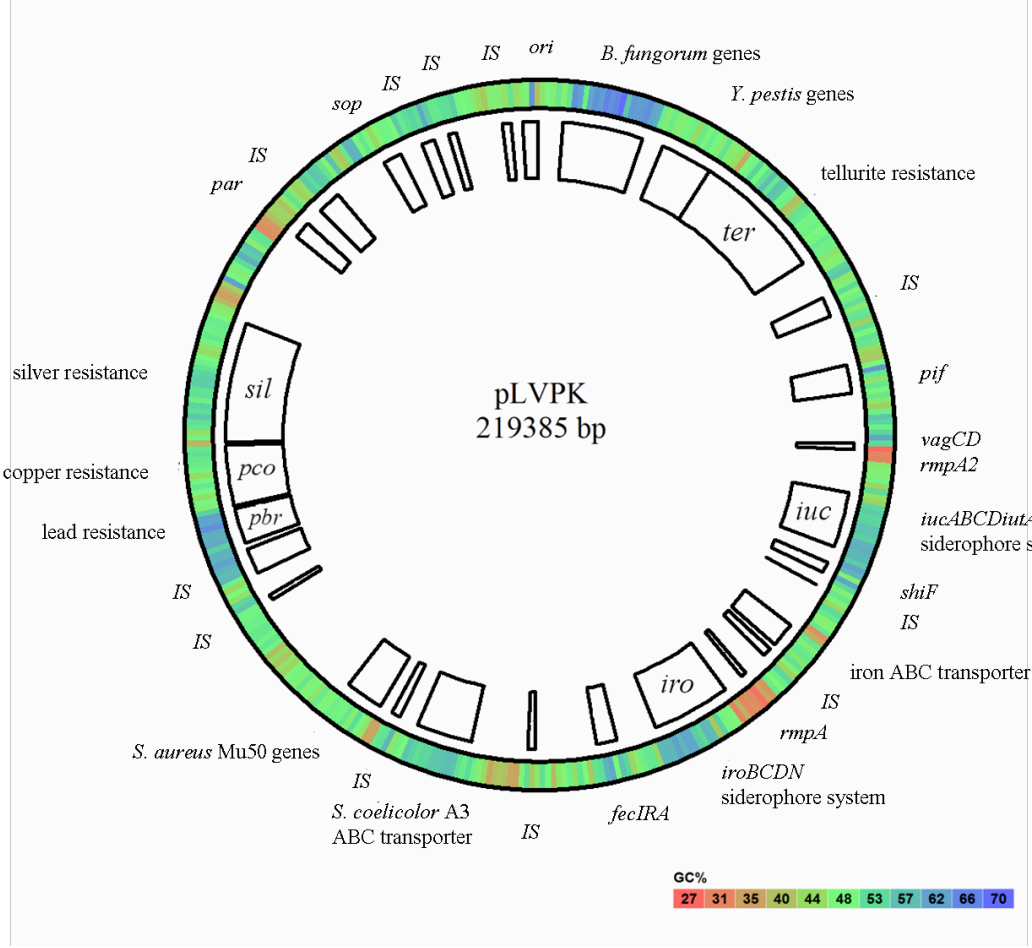


Figure 2. Base composition of pLVPK. The G+C content along the sequence of pLVPK are calculated with a window of 1000 bp. The horizontal line indicates 50 % G+C content, and the selected ORFs are shown as open boxes drawn to the exact scale. The region that contained the genes similar to that of the *B. fungorum* gene cluster is labeled. The *iut* and *iro* siderophore gene clusters, the lead-resistance gene clusters (*pbr*) and the region of *rmpA* and *rmpA2* genes respectively are also indicated.

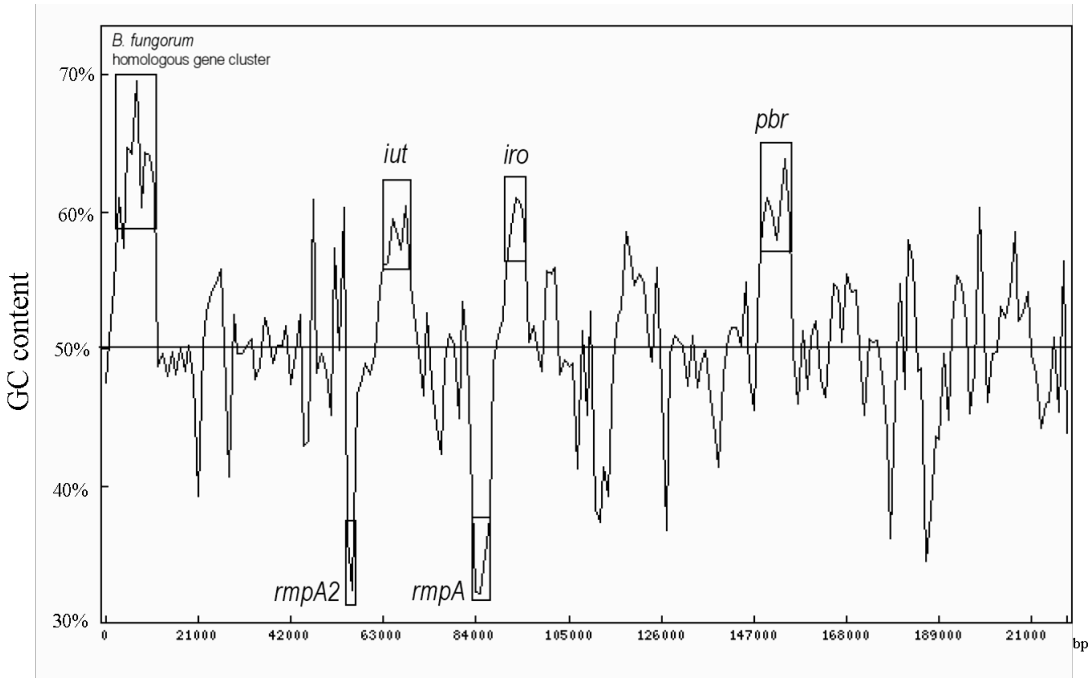


Figure 3. Comparison maps of the *rmpA* and *rmpA2* and their neighboring genes. (a) The restriction map of *rmpA2* and *rmpA*. (b) Southern analysis of pLVPK (lane 2). Southern hybridization with a probe prepared from the PCR product of the 636-bp *rmpA2* coding sequence is shown at the right. Two *Eco*RI fragments of 6.9-kb and 1.6-kb respectively representing the fragment containing *rmpA* and *rmpA2* are detected. The 3.8-kb and 14.6-kb fragments in lane 2 represent the *rmpA*- and *rmpA2*-containing *Hin*dIII fragments.

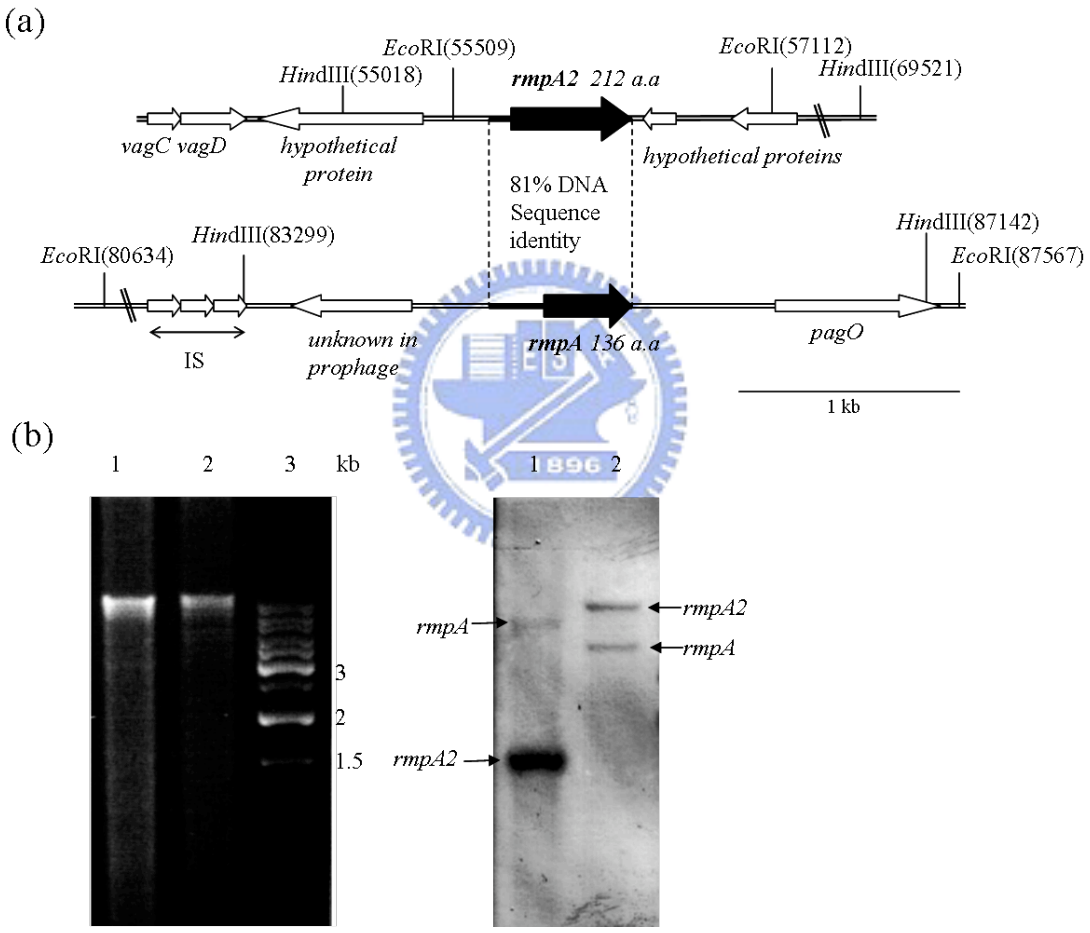


Figure 4. Comparison of the organization of *fecIRA* genes of pLVPK with similar iron transport and signaling systems. The arrows indicate the transcription orientation of the genes. The genes in the *fecIRA* gene cluster of the pLVPK are very similar to that of the *E. coli fec* operon, however lacking the *fecBCDE* genes.

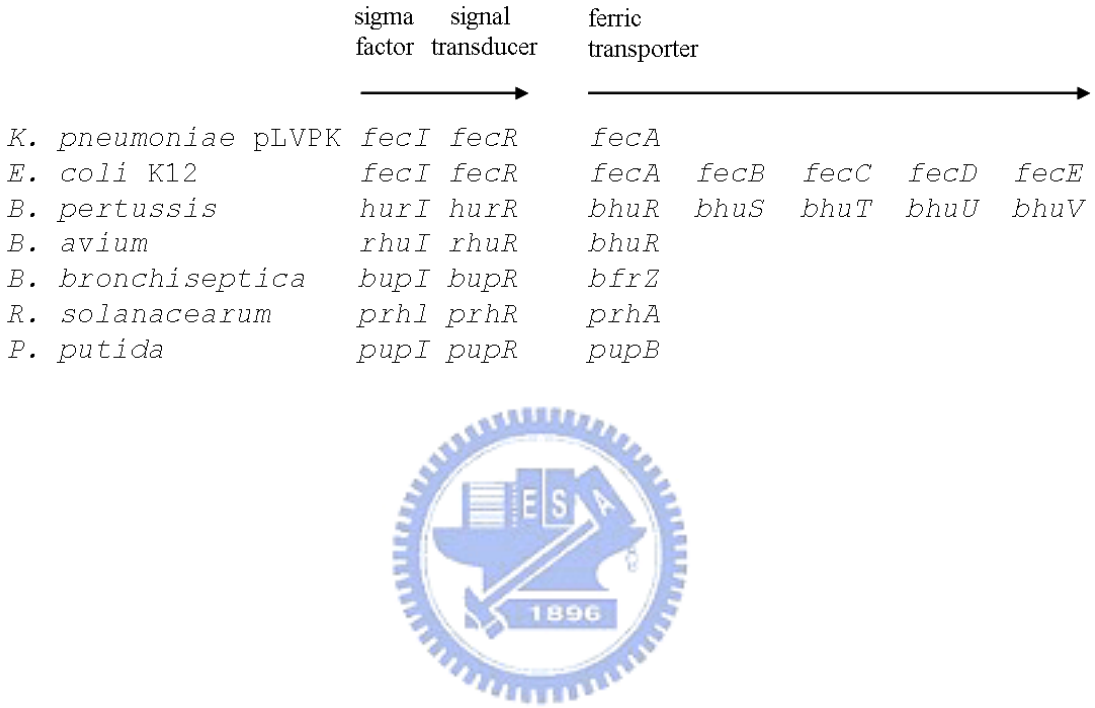


Figure 5. The heavy-metal resistant gene clusters. The silver-resistance *sil* gene cluster, in comparison with that of the *Salmonella enterica* serovar Typhimurium pMG101 (Gupta et al., 1999), the copper-resistance *pco* gene cluster, in comparison with that of the *E. coli* pRJ1004 (Brown et al., 1995), and the lead-resistance *pbr* gene cluster, in comparison with that of the *R. metallidurans* pMOL30 are shown (Borremans et al., 2001). The homologous gene clusters are depicted in solid black and the transcriptional orientation of the individual ORFs are shown by arrows.

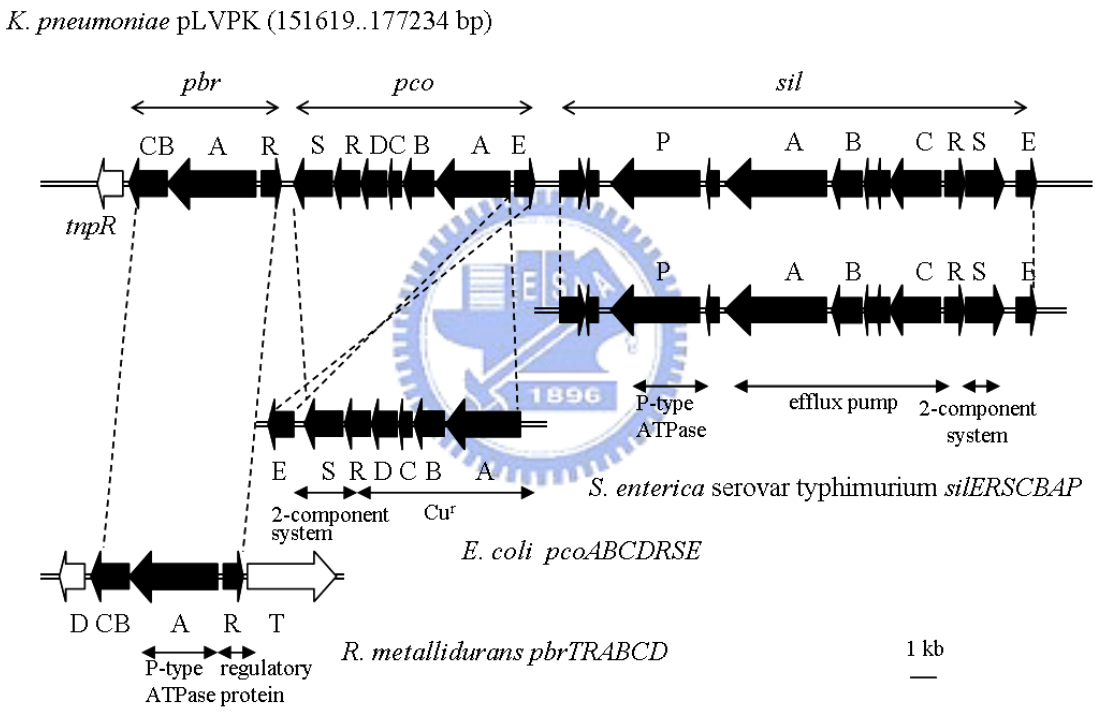
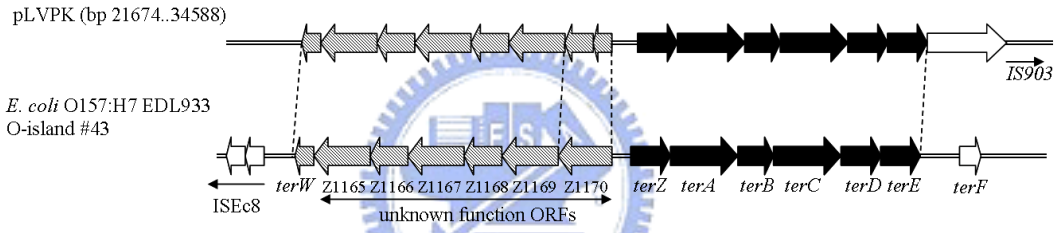


Figure 6. The tellurite resistance genes in *K. pneumoniae* CG43. (a)The *ter* gene cluster of pLVPK similar to *terZABCDE* of *E. coli* O157:H7 EDL933 O-island #43, a region of the EDL933 chromosome not homologous to *E. coli* K-12 MG1655. The homologous regions are shown in hatched and solid arrow. (b) Effects of the *tehB* deletion and curing of pLVPK on the tellurite resistance of the bacteria. Tellurite resistance assays of *K. pneumoniae* CG43S3, the pLVPK-cured bacteria (CG-101), and the *tehB* mutant CG43S3T1 using the discs immersed respectively with 20 and 40 μg K_2TeO_3 are shown. The arrows indicate the margins of the inhibition zone.

(a)



(b)

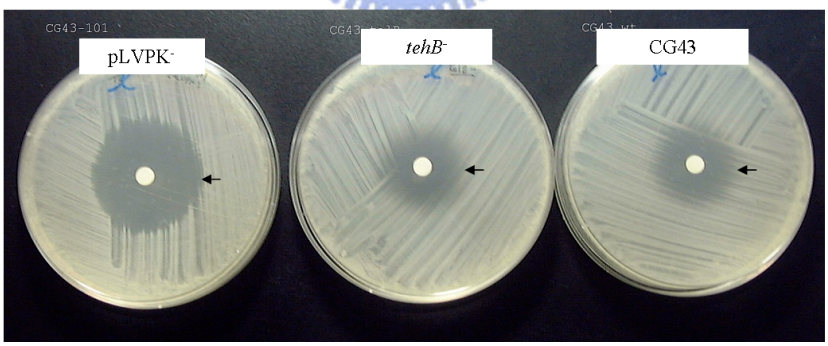


Figure 7. The *repA* and the iteron sequences at the replication origin. (a) Sequence alignment of the *repA* from pLVPK (LVPK), R27 of *S. enterica* serovar Typhi (R27), pHMT-1 of *Y. pesits* (PMT1), and pHCM2 of *S. enterica* serovar Typhi (PHCM2). Leucine residues near the N-terminus are marked with asterisks, and the two helix regions corresponding to the helix-turn-helix DNA-binding domain are shown. (b) The replication origin including the *repA* structural gene and two iteron sequences nearby of the pLVPK. The sequences of the two sets of iterons: iteron-1 (217448..218020 bp) and iteron-2 (219102.. 219203 bp) are shown with each row respectively representative for each of the adjacent repeat units.

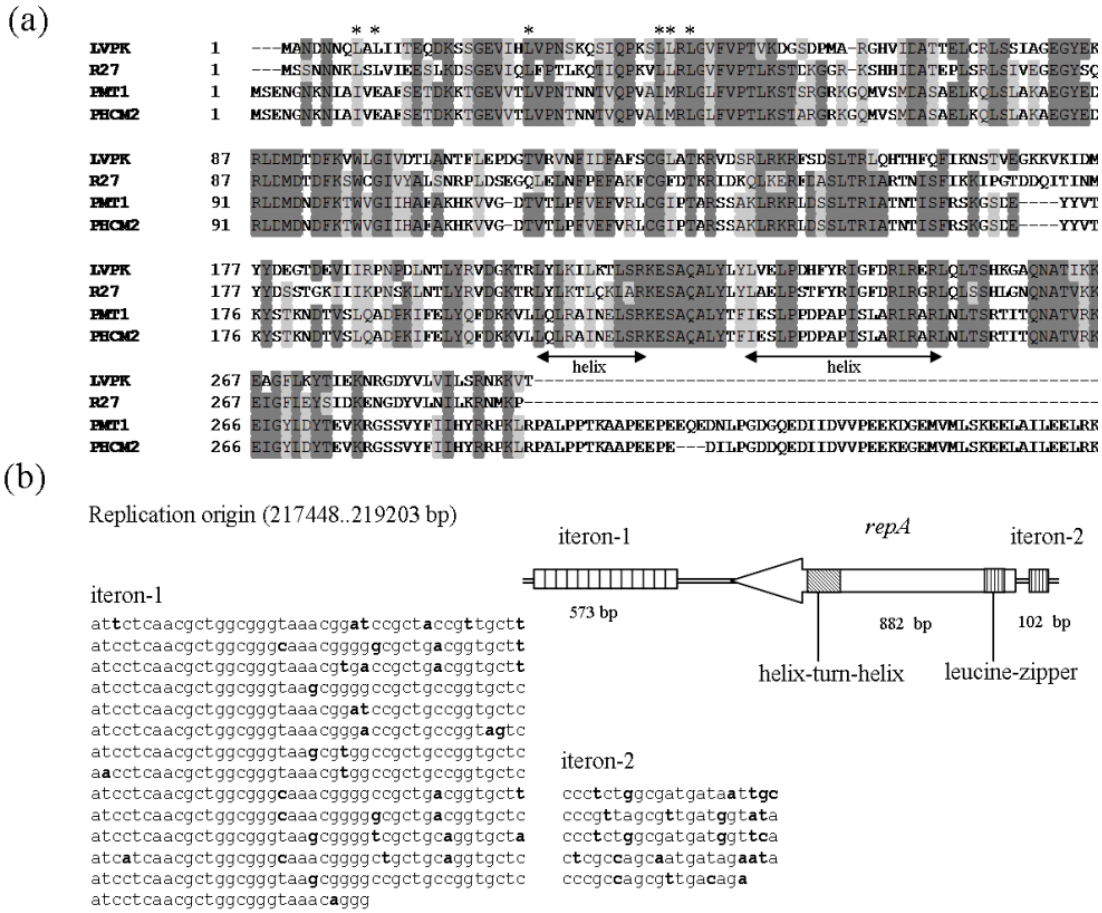


Figure 8. Phylogenetic trees of the 54 sensor kinases (A) and 58 response regulators (B) constructed by CLUSTALW using the deduced amino acid sequences. The bootstrap values generated by 1000 sets of replications of tree construction are labeled at the internal nodes of the un-rooted trees. The congruent clades respectively in the two trees are labeled from A to F. The sensors IT, ITR, and ITRO and the regulators marked with NarL, NtrC, OmpR, and UC (unclassified) are shown.

Figure 8. (A) continued

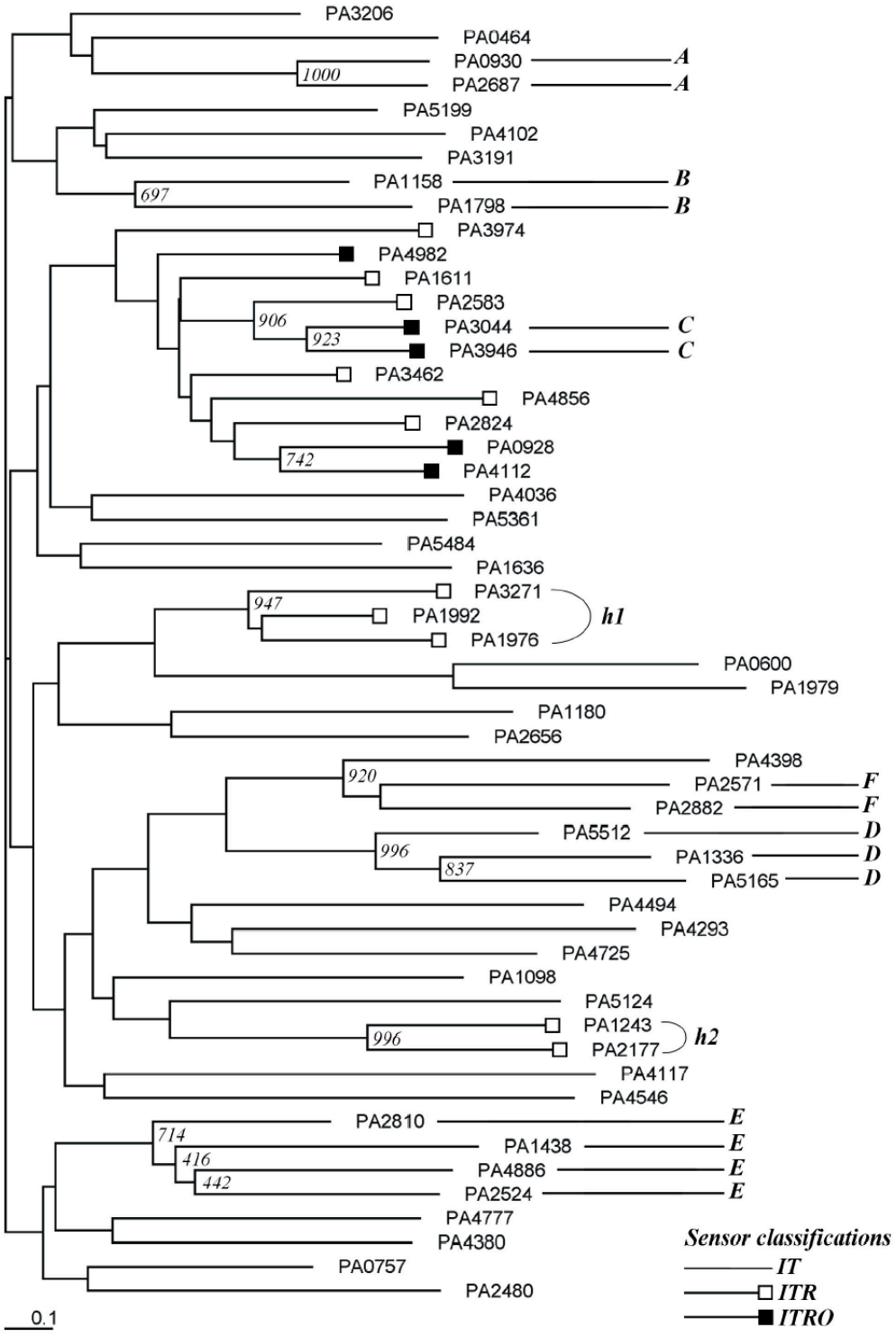


Figure 8. (B) continued

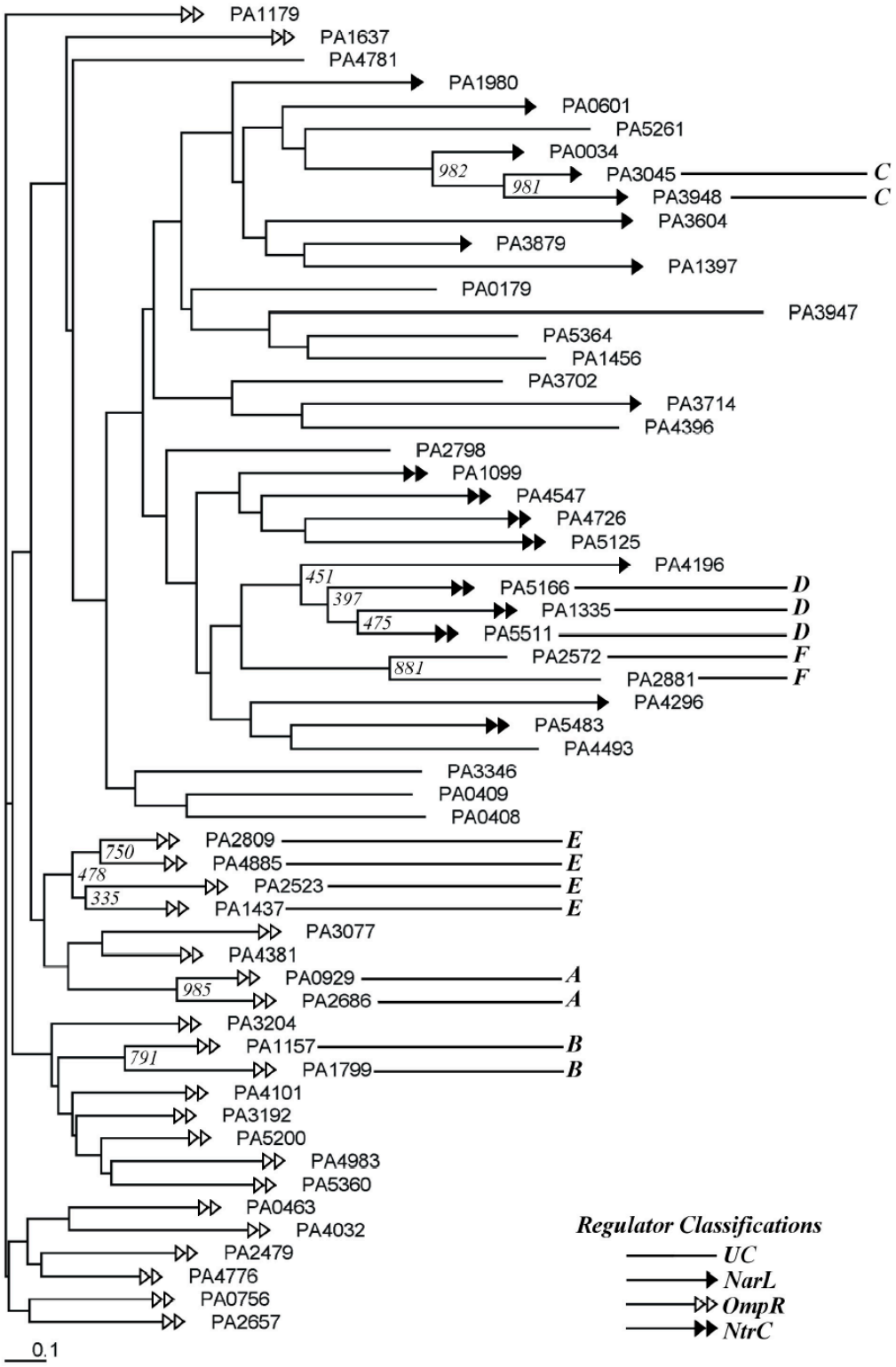


Figure 9. The ML trees of the 2CSs carrying the response regulators of OmpR- (A), NtrC- (B), and NarL-type (C) respectively. The number shown at each node of the un-rooted trees represents the bootstrap value of the operational taxonomic unit (OTU). The congruent clades in the Figure 1 are also indicated. The PA number of the OTU is shown with an underline and the PA number of the linked component is also marked.

Figure 9. continued

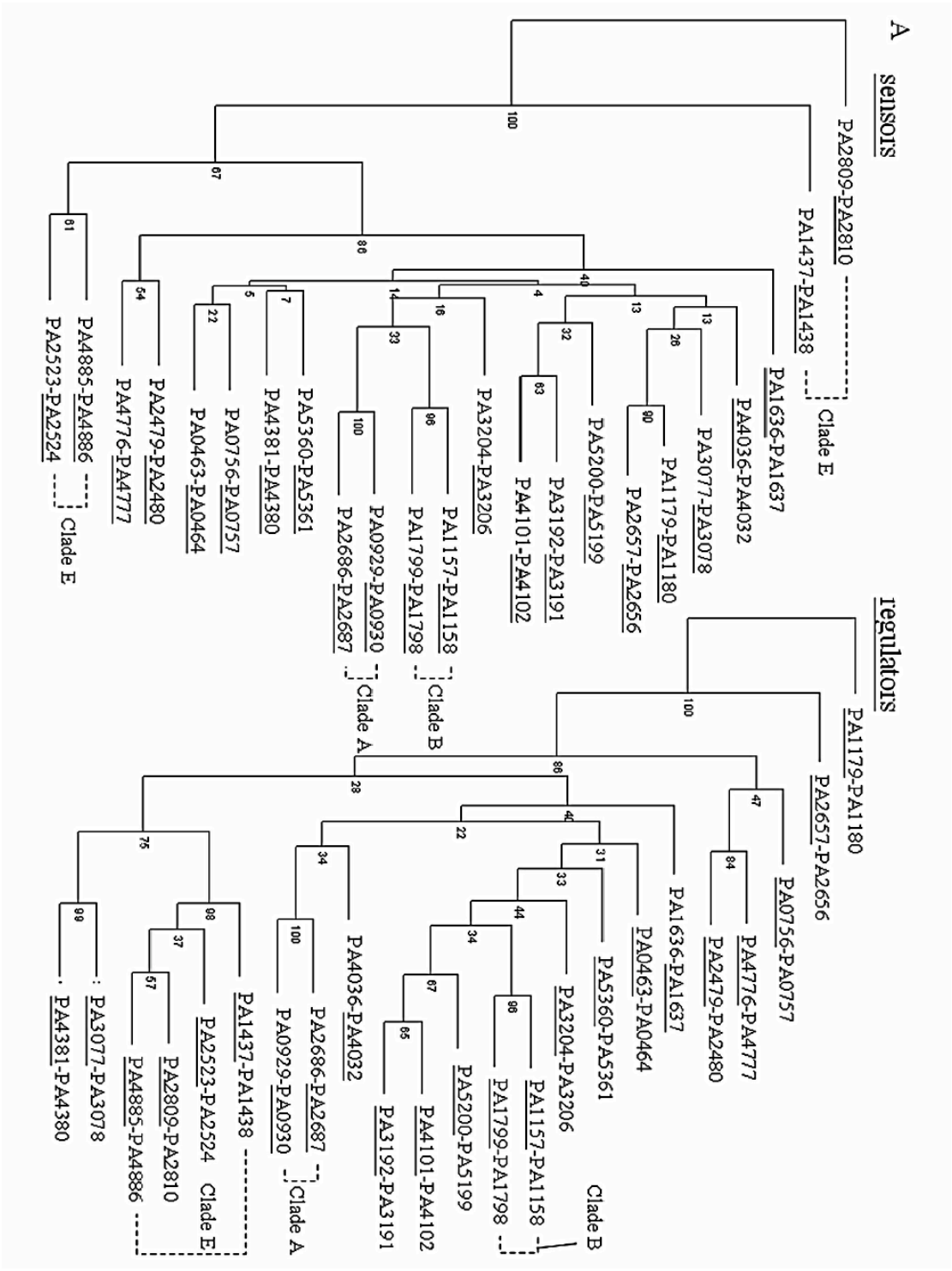
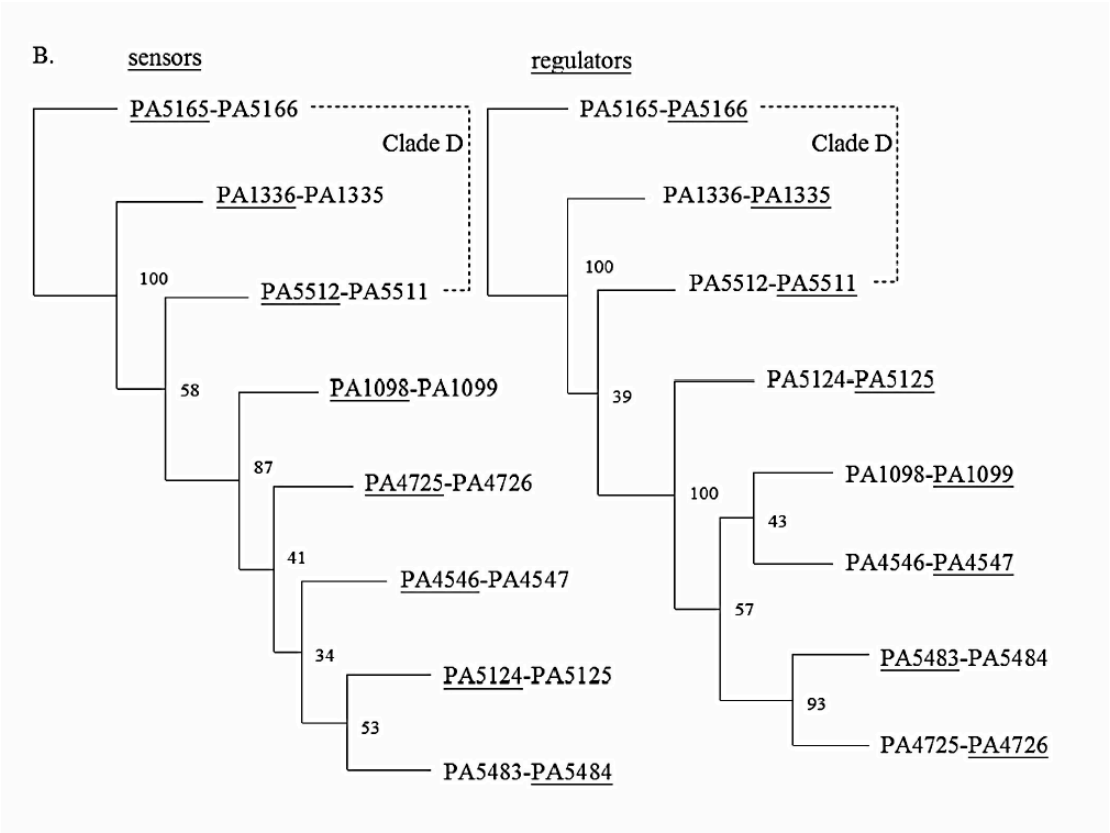


Figure 9. continued



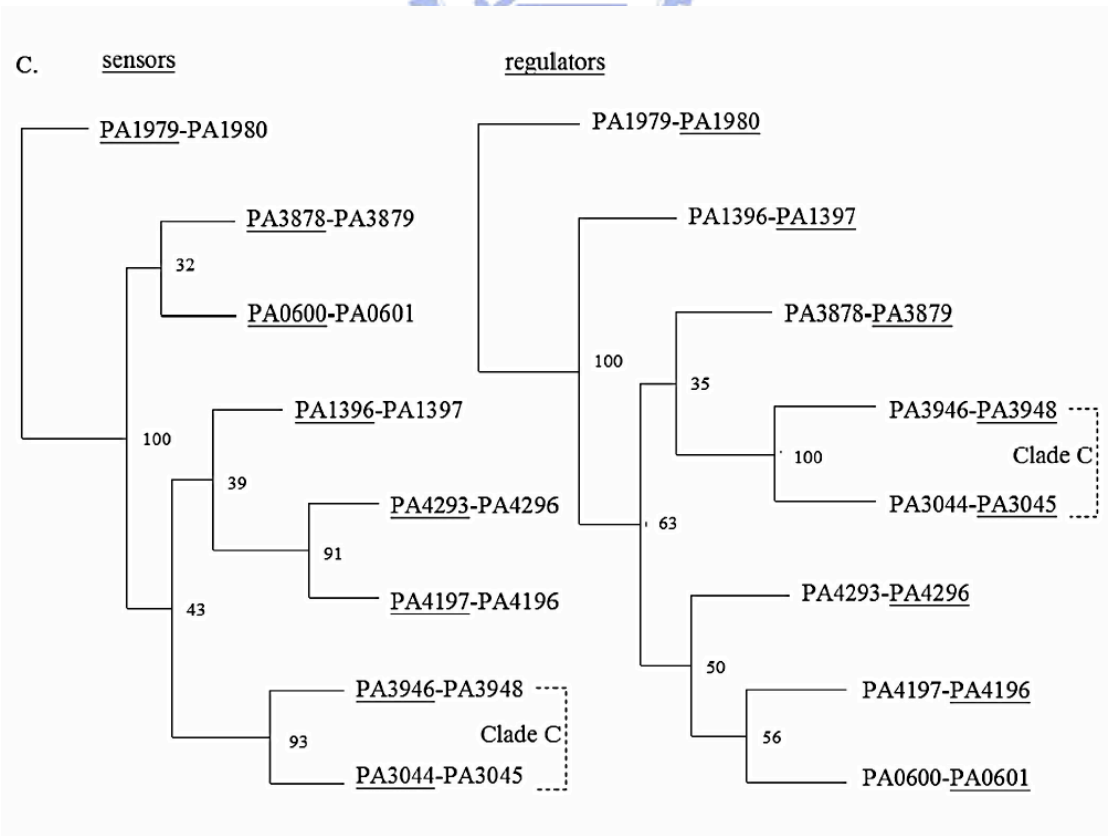


Figure 10. The distribution of the dissimilarities of ML trees built respectively with OmpR- (A), NtrC- and NarL- (B) groups measured by quartet metrics. The resolved and different quartets are scored in 300 random trees generated respectively with 23 and 8 OTUs using COMPONENT (Page, 1993). The dissimilarities of the sensor and regulator trees are indicated for each of the groups.

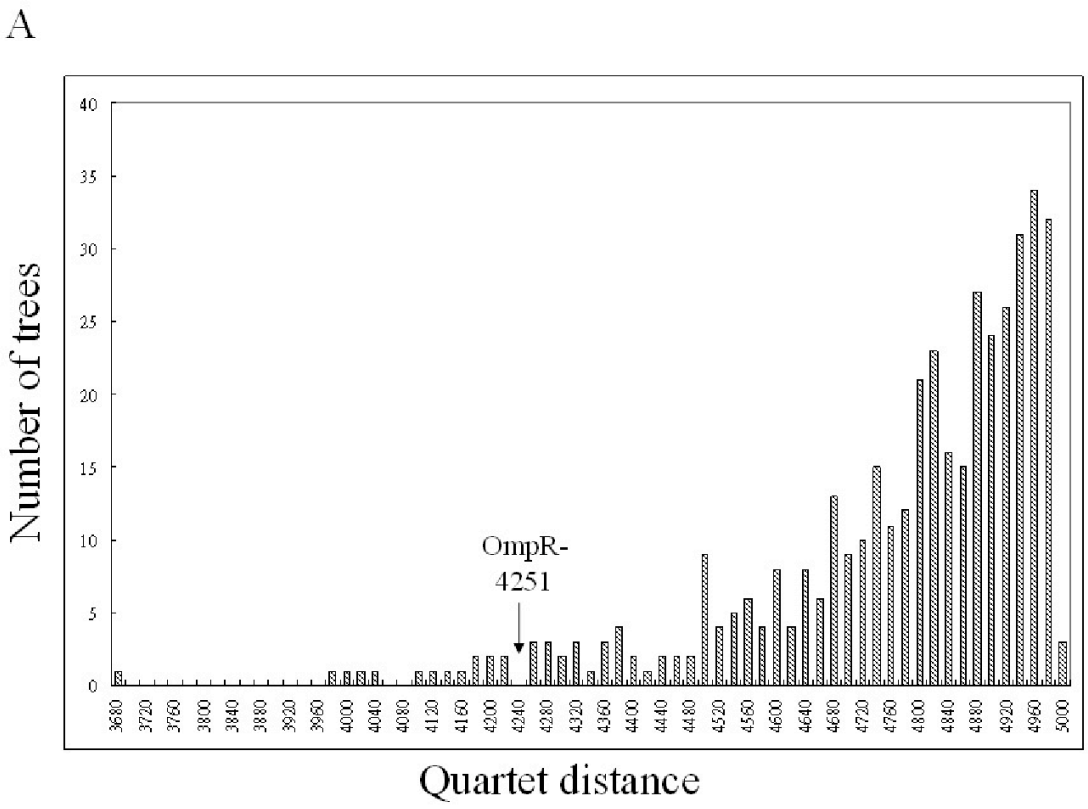


Figure 10(B). continued

B

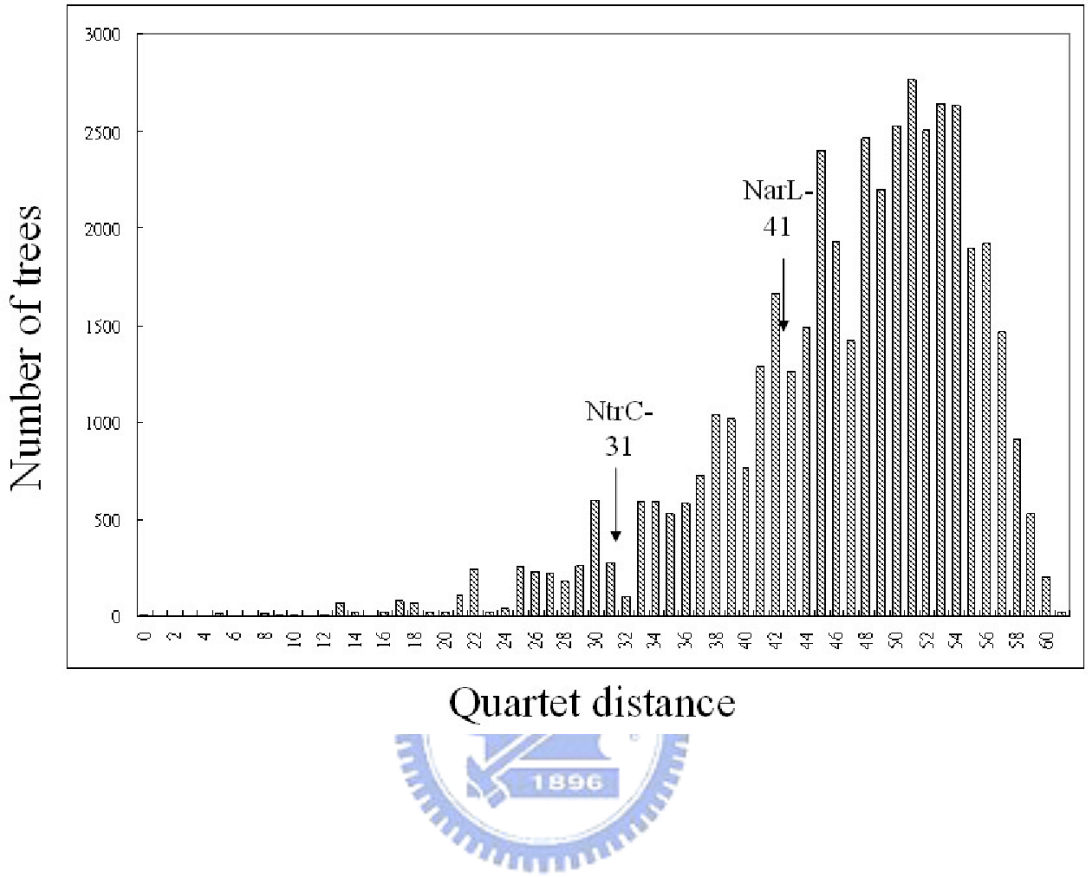


Figure 11. Classification of sensor kinases by the sequence around the histidine. (A) Neighbor-joining tree built by the multiple sequence alignment of the sequences around the phosphorylated histidine residue using CLUSTAL W. The classification of their cognate regulators, CheA-like, NarL-like, OmpR-like and NtrC-like are shown on the tree. (B) The graphical representations of the three major classes of kinases by multiple sequence alignments around the phosphorylated histidine residue are shown. The amino acid symbols are shown at each position. The sequence conservation at each position is indicated by an overall height of the stack, while the relative frequency of each amino acid is indicated by the height of symbols within the stack.

Figure 11(A). continued

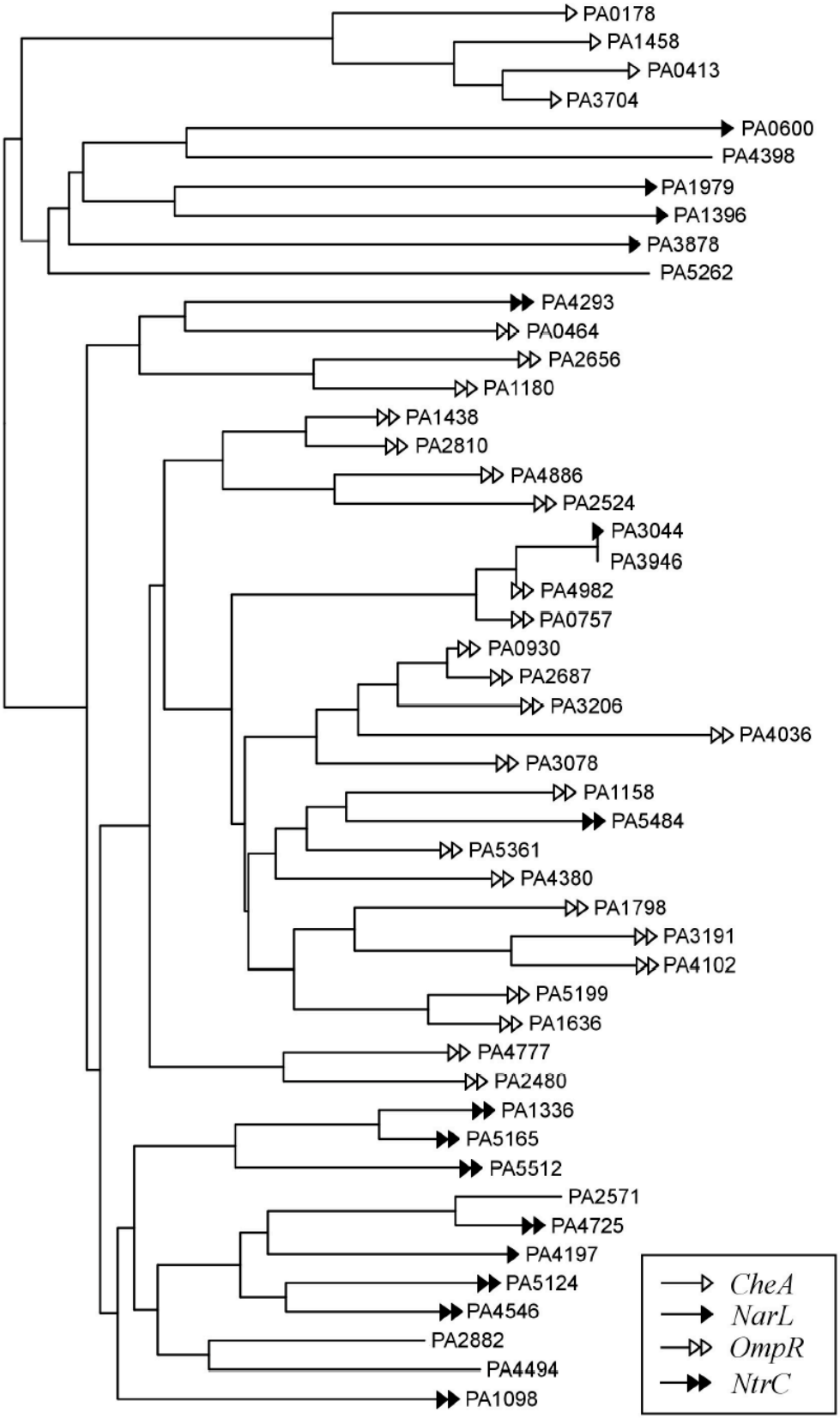


Figure 11(B). continued

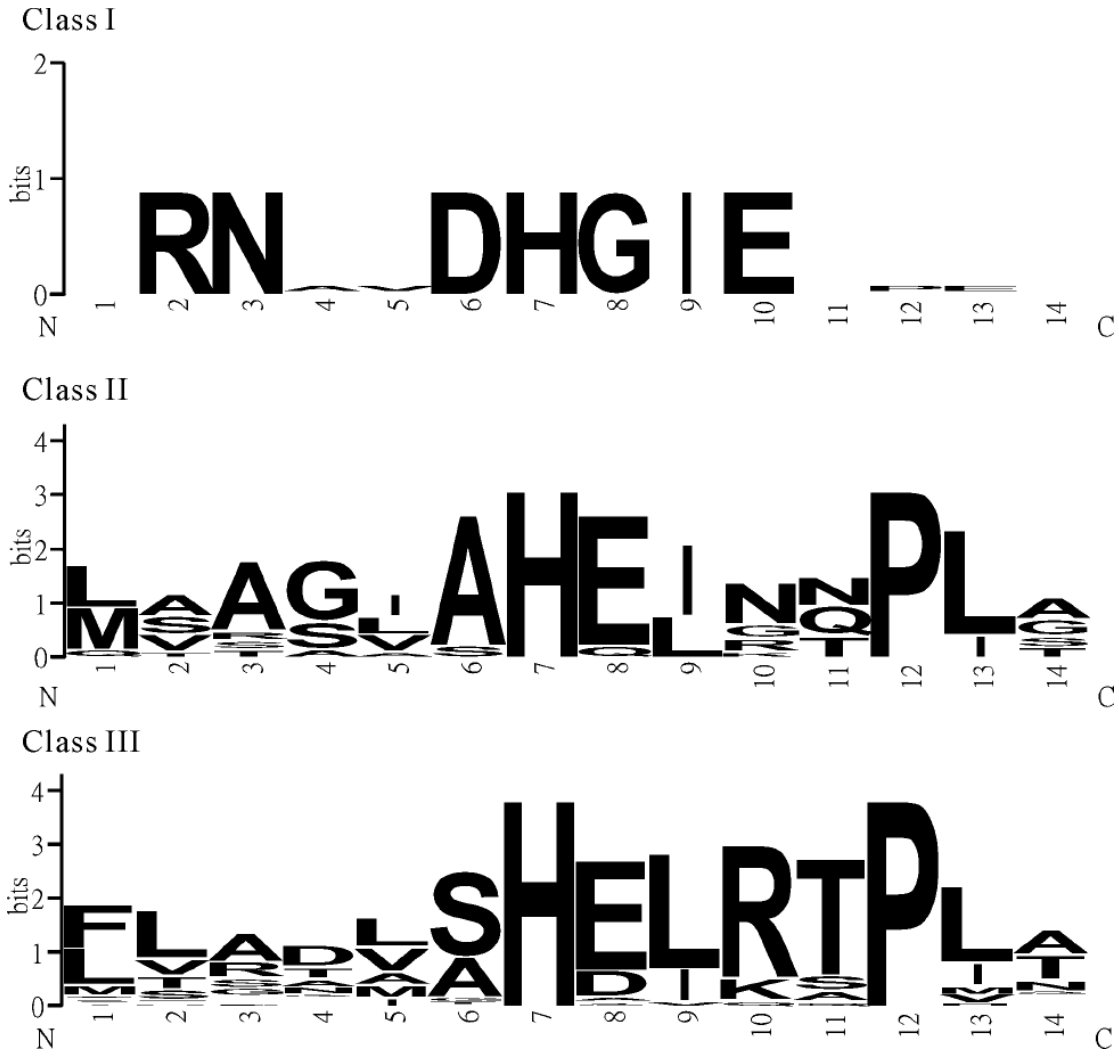
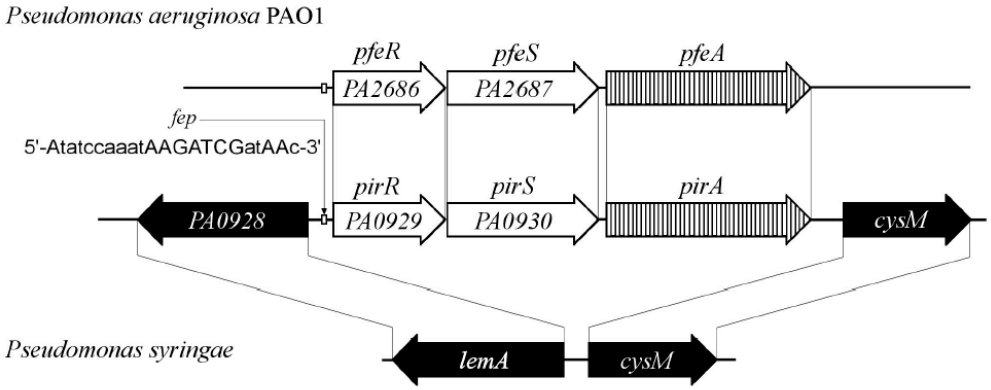


Figure 12. Schematic diagram of the gene organization of (A) the *pfeRSA* and *pirRSA* gene clusters in *P. aeruginosa* PAO1 genome. The broken likes linking genes show the regions with amino acid similarity while arrows represent the transcriptional direction of genes. The PA numbers are labeled above the arrow for the 2CS genes. A *lemA-cysM* carrying segment of *P. syringae* is also shown below. The sequence and position of a putative *fep* binding site (Wang and Church, 1992) identified in the upstream region of *pirRSA* is shown. (B) the two *bvgAS*–like gene clusters PA3044/PA3045 and PA3946/PA3947/PA3948 in *P. aeruginosa* PAO1 genome (NA: NarL-like regulators, ITRO: unorthodox sensor kinase). A region of the *P. aeruginosa* PA14 genome containing the homolog of PA3947, *pvrR*, is shown below.

Figure 12. continued

(A)



(B)

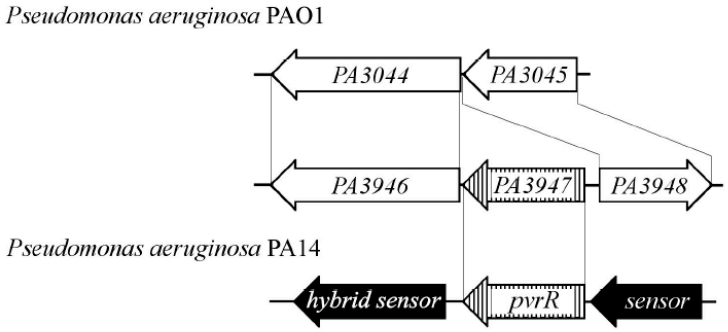
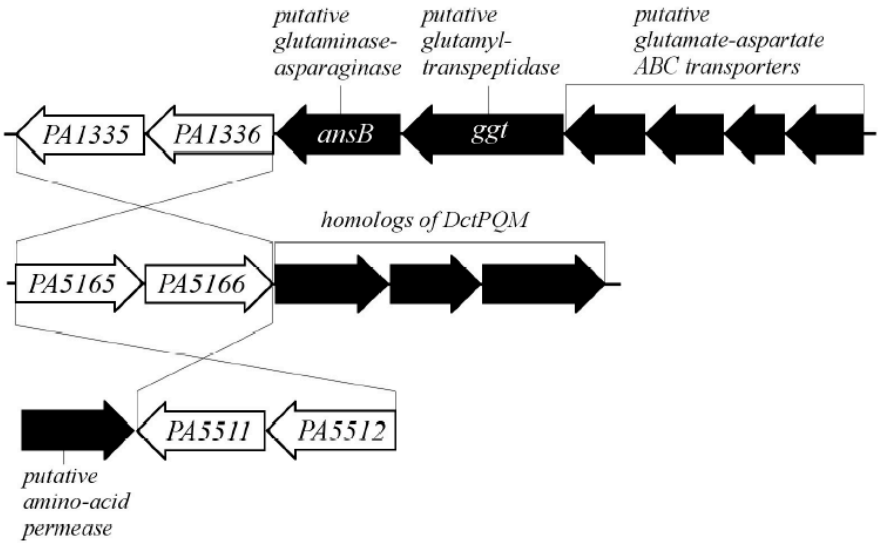


Figure 13. Gene organization of the 2CSs in clade D and E. (A) The regions containing the 2CS genes in the clade D in the *P. aeruginosa* PAO1 genome are shown. Arrows correspond to the transcription orientation of the genes. The PA numbers are labeled on the 2CS genes (open arrow). The PA5165/PA5166 is adjacent to *dctPQM*, which encodes *R. capsulatus* C4-dicarboxylates transportation system homologs (Shaw et al., 1991). The PA1336/PA1335 with the direct downstream genes of *Pseudomonas* 7A *ansB* glutaminase—asparaginase hololog, *E. coli* γ-glutamyltranspeptidase, and the four *E. coli* glutamate-aspartate ABC transporter homologs are shown. The PA5512/PA5511 is shown adjacent to the gene homolog of *S. typhimurium* putative amino acid permease. (B) Clusters of the known or putative small molecular transporter genes together with the four 2CSs in clade E are shown. The five genes in the *czc* gene cluster showed compelling similarities to the *R. eutropha czc* cation efflux system. A putative P*copH* transcription factor-binding site (Mills et al., 1994) identified at the upstream region of PA2523 is indicated. The hatched arrow represents nearby genes with known or predicted function involved in small molecular transportation.

Figure 13. continued

(A)



(B)

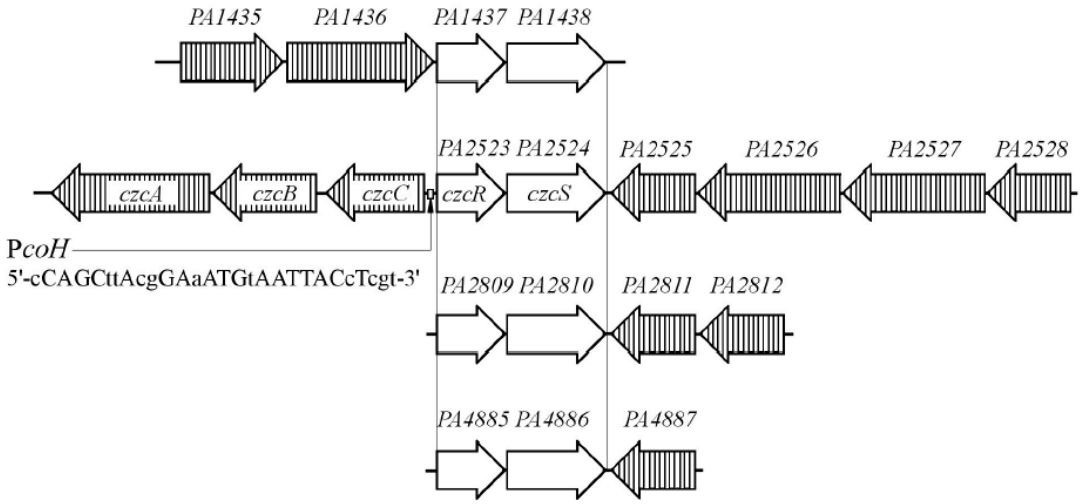


Figure 14. Phylogenetic tree of the species chosen in this study based on maximum-likelihood analysis of the 16S rRNA sequences. The tree is shown with the eukaryotic lineage as an outgroup.

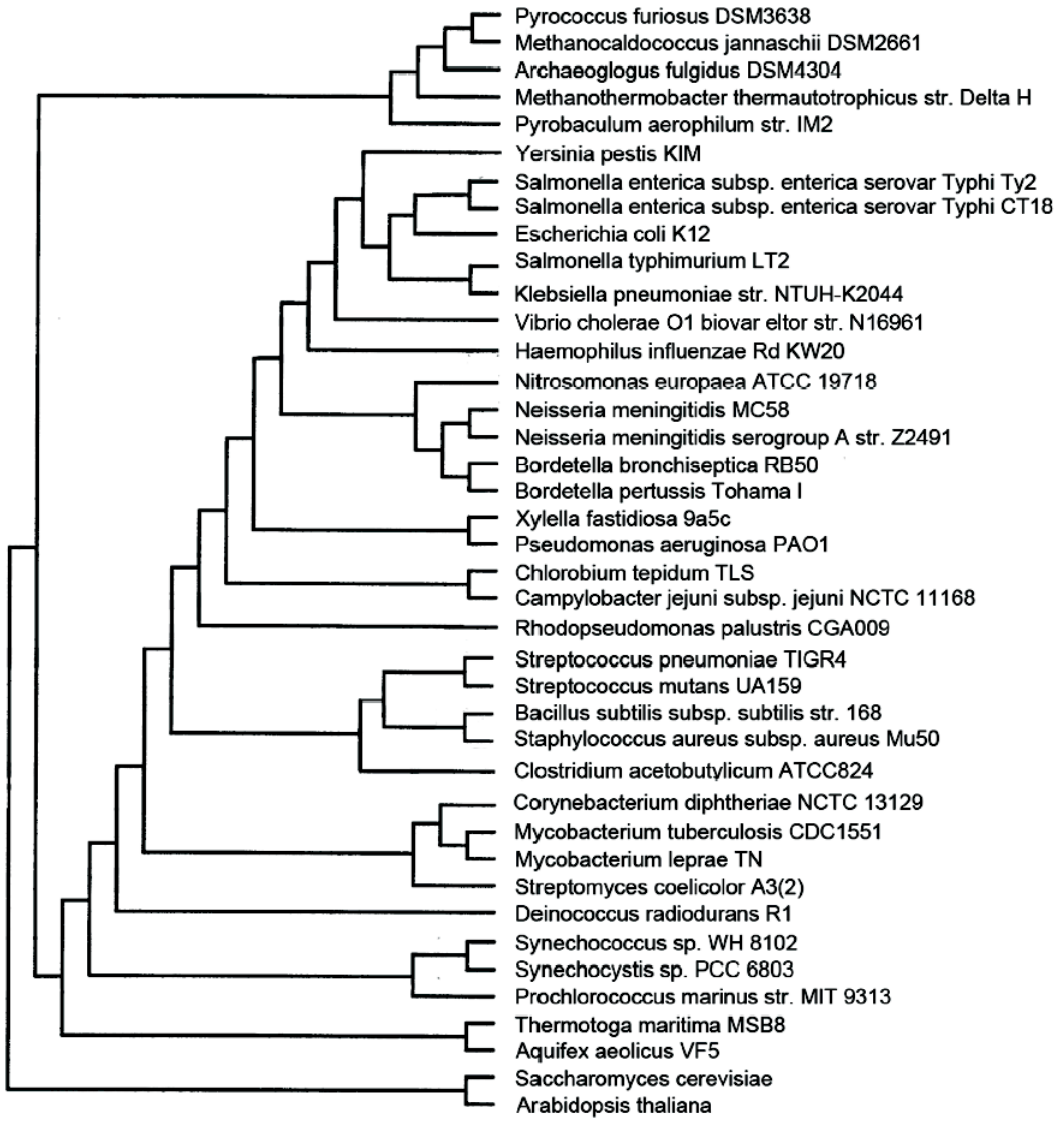


Figure 15. The relationship between the number of 2CS components, as calculated in Table 1, and the genome sizes respectively of these organisms. (A) The number of histidine kinase domain-containing ORFs versus the genome sizes. (B) The number of response regulator receiver domain-containing ORFs versus the genome sizes. (C) The number of 'OmpR-like regulator' (ORFs that contained both response regulator receiver domain and the C-terminal DNA-binding domain of OmpR regulator) versus the genome size. Note that the numbers of 2CS components were calculated based on our domain identification using HMMER, rather than the annotation results.

(A)

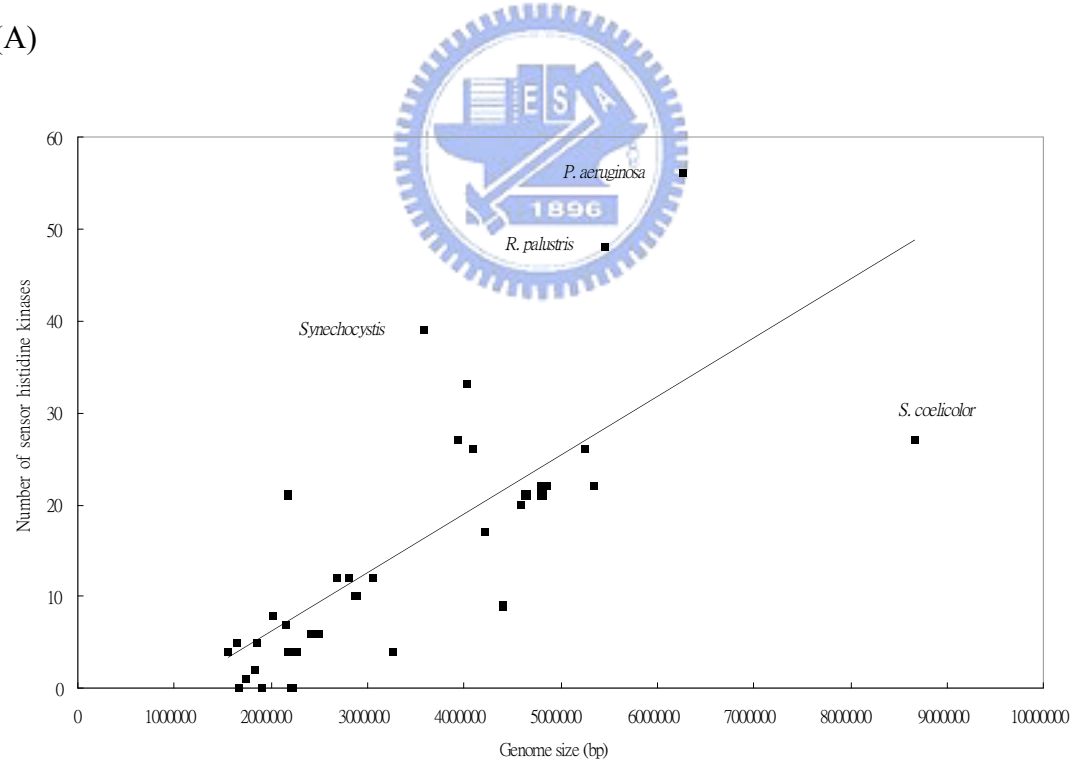
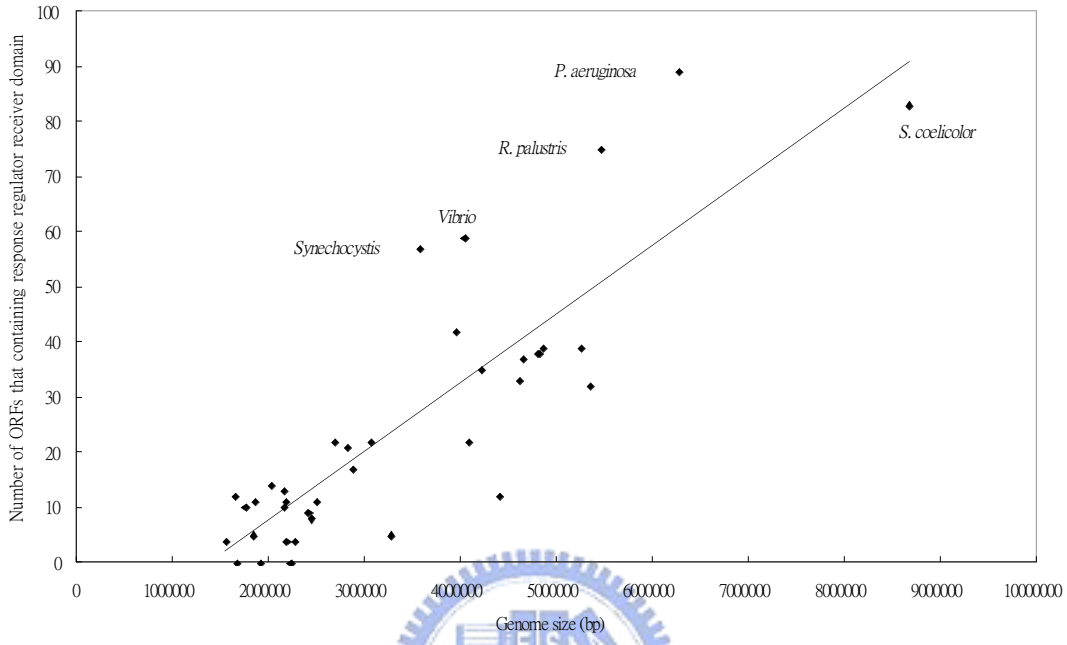


Figure 15. continued

(B)



(C)

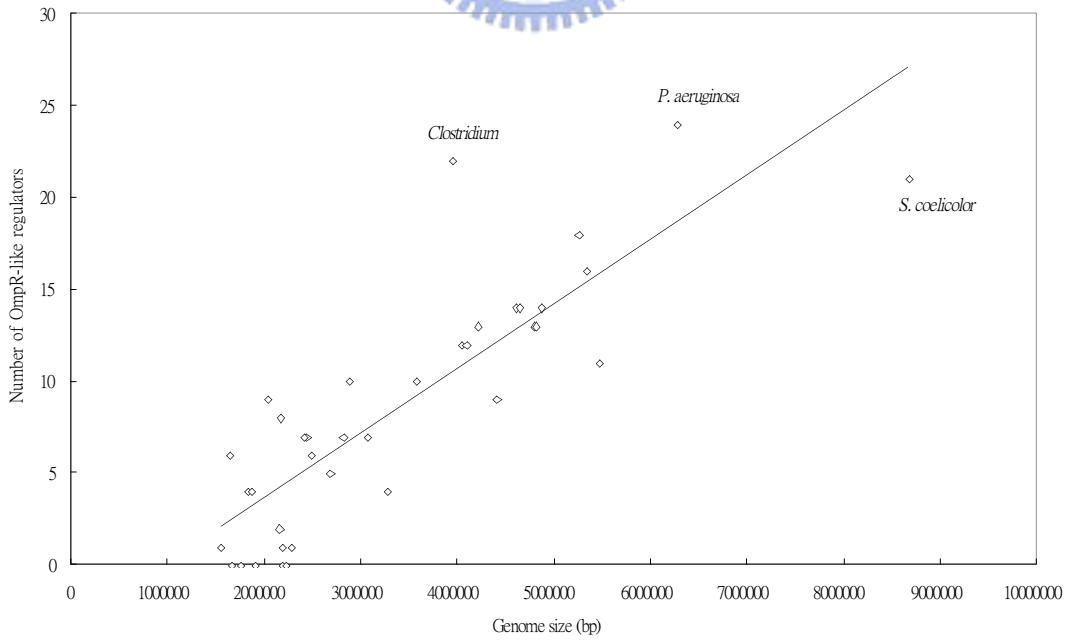


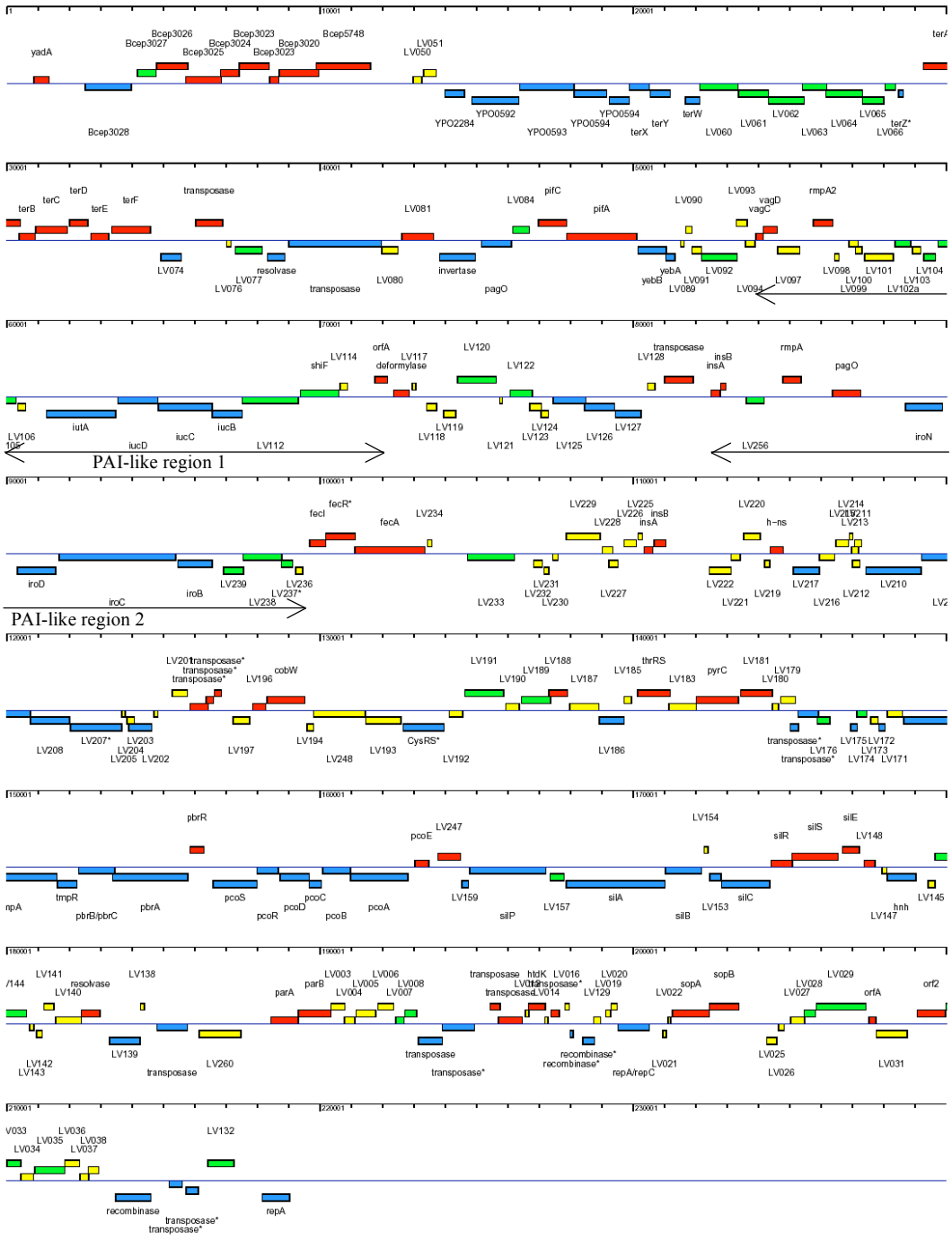
Figure 16. Sequence logos of the amino acid sequences around the phosphorylated histidine residue of the sensor kinases from (A) RS: regulator-to-sensor type and (B) SR: sensor-to-regulator type respectively according to their gene orders. The 14 amino acid residues spanning the histidine residue of the sensor kinases were aligned and their sequence logos shown by use of WebLogo (Crooks et al., 2004). The overall height of the stack indicates the sequence conservation at that position, whereas the height of symbols within the stack indicates the relative frequency of each amino acid at that position.

(A)

(B)

Appendix I

Map of the ORFs in pLVPK



Green: conserved hypothetical ORF: yellow: hypothetical ORF; Red/Blue: known ORF