Genomic blueprint of *Hahella chejuensis*, a marine microbe producing an algicidal agent

Haeyoung Jeong, Joung Han Yim[3], Choonghwan Lee[1], Sang-Haeng Choi, Yon Kyoung Park, Sung Ho Yoon, Cheol-Goo Hur, Ho-Young Kang, Dockyu Kim[1], Hyun Hee Lee, Kyun Hyang Park, Seung-Hwan Park, Hong-Seog Park, Hong Kum Lee[3], Tae Kwang Oh[2] and Jihyun F. Kim*

Genome Research Center, [1]Division of Drug Discovery and [2]21C Frontier Microbial Genomics and Applications Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), PO Box 115, Yuseong, Daejeon 305-600, Republic of Korea and [3]Korea Polar Research Institute, KORDI, PO Box 29, Ansan, Seoul 425-600, Republic of Korea

ABSTRACT

Harmful algal blooms, caused by rapid growth and accumulation of certain microalgae in the ocean, pose considerable impacts on marine environments, aquatic industries and even public health. Here, we present the 7.2-megabase genome of the marine bacterium *Hahella chejuensis* including genes responsible for the biosynthesis of a pigment which has the lytic activity against a red-tide dinoflagellate. *H.chejuensis* is the first sequenced species in the *Oceanospiralles* clade, and sequence analysis revealed its distant relationship to the *Pseudomonas* group. The genome was well equipped with genes for basic metabolic capabilities and contained a large number of genes involved in regulation or transport as well as with characteristics as a marine heterotroph. Sequence analysis also revealed a multitude of genes of functional equivalence or of possible foreign origin. Functions encoded in the genomic islands include biosynthesis of exopolysaccharides, toxins, polyketides or non-ribosomal peptides, iron utilization, motility, type III protein secretion and pigmentation. Molecular structure of the algicidal pigment, which was determined through LC-ESI-MS/MS and NMR analyses, indicated that it is prodigiosin. In conclusion, our work provides new insights into mitigating algal blooms in addition to genetic make-up, physiology, biotic interactions and biological roles in the community of a marine bacterium.

INTRODUCTION

Accounting for >98% of the ocean's biomass, marine microbes are the major players of the biogeochemical cycles on earth. Phytoplanktons fix solar energy and provide nutrients to other marine life. On the other hand, unchecked increases in the population of certain dinoflagellates like much-blamed *Pfiesteria* spp. (1,2) in the ocean results in blooms that often threaten the marine life. These phenomena called harmful algal blooms or commonly red tides increasingly occur in the coastal waters throughout the world in recent years, affecting not only the health of human and marine organisms but regional economies and marine ecosystem (3–5). However, the only practical management strategy being employed in some places is flocculation of microalgae through clay dispersal (6,7).

Hahella chejuensis (8) is a cultivated member of the oceanic γ-*Proteobacteria* (9), which is one of the most prevalent prokaryotic groups present in marine environments (10,11). None of the members of the *Oceanospirillales* clade has yet been determined for its genome sequence (Figure 1). Originally isolated from the coastal marine sediment of the southernmost island in Korea, this red-pigmented bacterium is capable of killing *Cochlodinium polykrikoides* (J. H. Yim and H. K. Lee, unpublished results), a major red-tide dinoflagellate problematic in the western coasts of the North Pacific (6,12). Though dozens of algicidal bacteria have been isolated, genetic basis of such activity has not been thoroughly examined and thus their modes of actions remain elusive (13).

We determined for the first time in its class the complete genome sequence of *H.chejuensis* KCTC 2396[T]. It provides insights into the lifestyle of the bacterium as a heterotroph living on various available resources from mineral, to organic materials and even other aquatic organisms. In addition, we
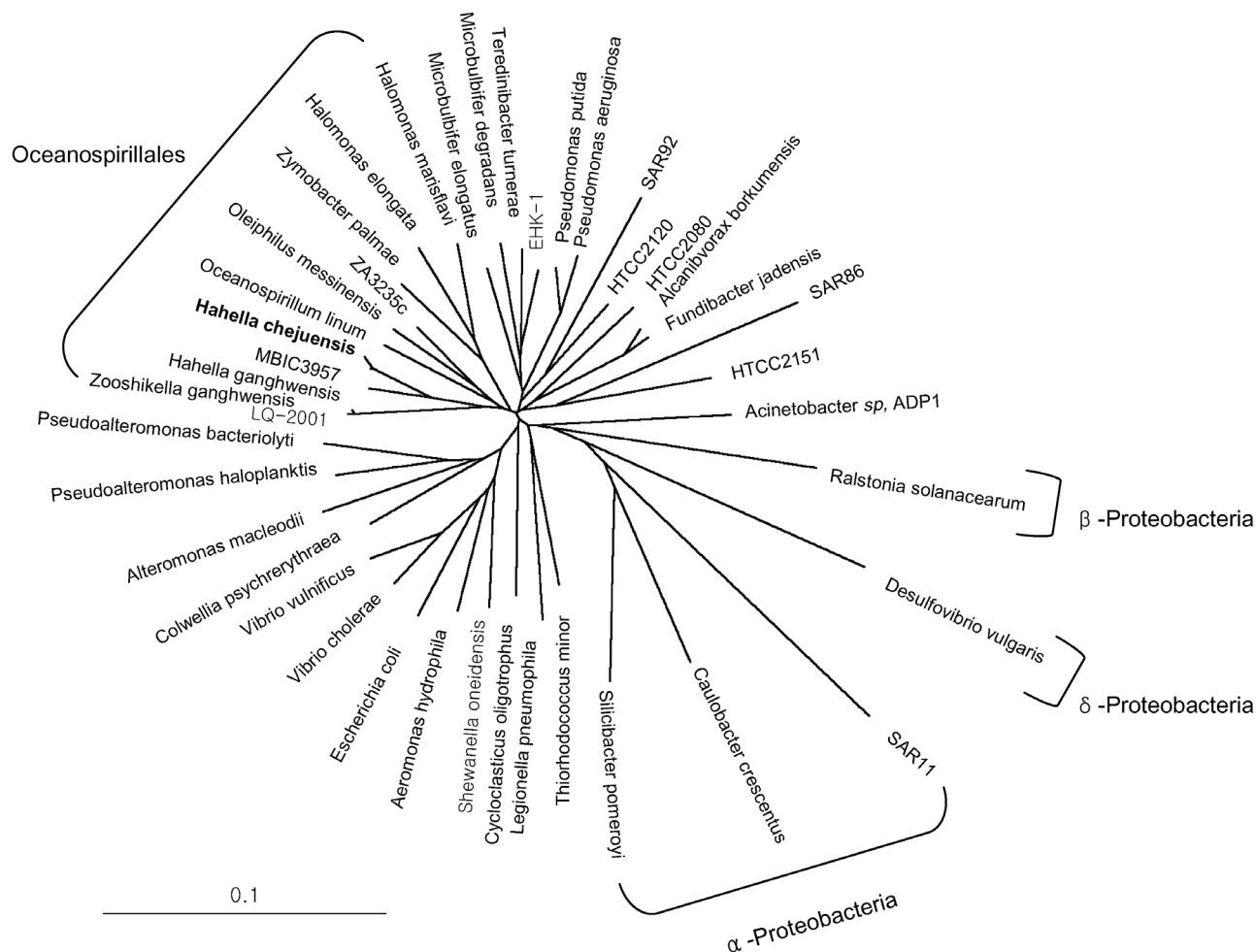
Figure 1. Phylogenetic position of *H.chejuensis* based on 16S rDNA sequences. Sequences were retrieved from GenBank and an unrooted neighbor-joining tree was calculated. A maximum parsimony analysis gave similar topology. IUB was used as DNA weight matrix and the scale bar represents number of substitutions per site.

discovered that the pigment it produces has the acute lytic activity against microalgae, identified the encoding genes from the genome information that was confirmed through functional analyses, and characterized it for its molecular structure.

MATERIALS AND METHODS

Sequencing and annotation

Genome sequence was determined by standard whole-genome shotgun method. Paired end reads of ∼8.1-fold genomic coverage were produced from 2, 5 and 40 kb clones, all prepared from randomly sheared chromosomal DNA. Chromatograms were processed by the phred/phrap/consed software package (http://www.phrap.org). Gap closure and additional sequencing of low-coverage regions were done by primer walking of gap-spanning clones or PCR products. Physical structure of the final genome sequence was further confirmed by comparing the hypothetical restriction patterns of the sequence for *Pme*I, *Sma*I and *Swa*I and the patterns produced by pulsed-field gel electrophoresis after restriction digestion of the genomic DNA with the three enzymes. The putative origin

of replication was determined by GC skew analysis and by identification of genes known to cluster near the prokaryotic *oriC* site (14,15). Putative CDSs of ⩾100 bp were predicted by amalgamating the results from CRITICA (16) and GLIM-MER (17). Intergenic sequences were reanalyzed for short CDSs by running BLASTX. Functional assignment of genes was performed by searching translated CDSs against public protein databases. Manual validation of the annotation results and final refinement including sequence and feature editing were done using ARTEMIS (18). Metabolic pathways were examined using the KEGG database (19) and Pathway Tools (20).

Phylogenetic analysis

The 16s rDNA sequences or 34 concatenated protein sequences that are conserved as the genetic core of the universal ancestor (21) were retrieved from GenBank and used as common tracers of genome evolution. To identify counterparts of 34 COGs in each species, the retrieved genomes were searched with series of 34 hidden Markov models made up of each COG protein cluster. The 16s rDNA sequences or concatenated universal 34 protein sequences under *E*-value

cutoff (1.0×10^{-06}) were analyzed with neighbor-joining and maximum parsimony methods in CLUSTAL W (22) and PHYLIP (23). To correct multiple substitutions in protein residues, the Kimura 2-parameter model was used, and to evaluate the reliability of the branching patterns, 1000 random bootstrap re-samplings were executed.

Identification of genomic islands

Horizontally transferred genes (HTGs) were inferred from genomic anomalies or phylogenetic context. A gene was considered anomalous if both G+C content and codon usage are aberrant (G+C content $\geqslant 1.5\sigma$ and Mahalanobis distance as a degree of the codon usage deviation $\geqslant 80.23$) (24). In addition, genes having orthologs in other prokaryotes but not in γ-*Proteobacteria* were added to the list of HTGs. BLASTP searches were performed to find pairs of reciprocal best hits between in the *H.chejuensis* CDSs and those in each of the 208 completely sequenced prokaryotes. If the CDS in the searched genomes can be aligned with that of *H.chejuensis* ($\geqslant 80\%$ of its length and $\geqslant 30\%$ identity), the pair was considered orthologs. Then, a genome scan of a 10-gene window was run and the regions containing four or more HTGs were identified. Neighboring regions were merged into larger regions and each was manually examined to exclude fragments consisting of housekeeping genes or only small-sized hypothetical genes.

Algicidal activity assay

The medium for production of red pigment was M-RP10356 medium (5% glucose, 0.1% peptone, 0.42 g KH_2PO_4, 0.34 g K_2HPO_4, 0.5 g $MgSO_4$, 2.0 g $CaCl_2$, 0.001 g $CoCl_2 \cdot 6H_2O$, 0.001 g $MnCl_3$, 0.001 g $ZnSO_4$ and 0.001 g $NaMoO_4$, pH 7.0, where 250 ml distilled water and aged seawater were added to the final volume of one liter). Cultivation was carried out in a 5 liters jar-fermentor for 72 h at 25°C with aeration of 1.5 volume of air added to liquid volume per minute after inoculation (2.0%). From the culture broth, crude RP10356 was extracted by chloroform, which was further concentrated. Crude preparation of the pigment was purified with silica gel 60 (0.063 mm, Merck, Germany) using chloroform (100, v), and re-purified with YMC-Pack ODS-A (250 × 10 mm, YMC Co., Japan) using methanol:water:acetic acid (81:14:5, v/v/v). All algal strains were maintained in f/2 culture medium at 22.5°C, and under a light intensity of ~55 μmol/m^2/s using a 16 h light/8 h dark illumination cycle. Purified pigment was dissolved in ethanol, and aliquots of 20 μl solution were added to 980 μl of microalgal suspension in test tubes. Cell number was counted with a microscope after 1.0% Lugol's solution staining. Following formula was used to calculate the algicidal effect:

$$\text{Algicidal activity}(\%)$$
$$= \left[\frac{(\text{Number of initial cells} - \text{Number of survived cells})}{\text{Number of initial cells}} \right] \times 100.$$

Heterologous expression and mutagenesis

Five fosmid clones of *Escherichia coli* EPI300, HC81010E03, HC81008E02, HC81002H12, HC81006F09 and HC81004F05 containing either part or the whole of the pigment gene cluster were plated and incubated overnight at 37°C on Luria–Bertani agar containing 20 μg/ml chloramphenicol, CopyControl

Induction Solution (Epicentre, USA) and, sometimes, 20% crude extract of *H.chejuensis* prepared by filtration (pore size 0.22 μm). Transposon mutagenesis on a variant of HC81006F09 that constitutively expresses the red pigment was carried out using EZ::TN <KAN-2> Insertion Kit (Epicentre).

Structural determination of the pigment

Red pigment was extracted with a mixture of methanol/1N HCl (24:1, v/v) from the supernatant of *H.chejuensis* culture which was grown on Marine Broth (Difco) for 24–48 h at 30°C with vigorous shaking. Red-colored fraction was purified through high-performance liquid chromatography. Following LC using acetonitrile and water (with 0.1% formic acid) as the mobile phase at a flow rate of 0.2 ml/min, ESI-MS was carried out with a Finnigan LCQ Advantage MAX ion trap mass spectrometer equipped with a Finnigan electrospray source. To determine the molecular structure, ^{1}H NMR (CD_3OD, 300 MHz) and ^{13}C NMR (CD_3OD, 75 MHz) analyses were performed, resulting in the raw data of 6.94 (m, 1H), 6.71 (m, 1H), 6.66 (s, 1H), 6.39 (s, 1H), 6.21 (m, 1H), 6.01 (s, 1H), 3.89 (s, 3H), 2.37 (t, 2H), 2.27 (s, 3H), 1.53 (m, 2H), 1.33 (m, 4H), 0.90 (t, 3H) (^{1}H NMR) and 169.6, 160.2, 141.0, 135.8, 129.9, 129.2, 125.1, 123.0, 120.6, 115.9, 113.1, 110.9, 95.9, 58.9, 32.7, 31.8, 26.6, 23.6, 14.4, 11.5 (^{13}C NMR).

RESULTS AND DISCUSSIONS

General features

The genome of *H.chejuensis* KCTC 2396^T consists of one circular chromosome of 7 215 267 bp (Figure 2). This makes it the largest among the marine prokaryotic genomes whose genome sequences are available and also among the γ-proteobacterial genomes sequenced. Among the 6783 predicted genes, 76.3% showed significant database matches and 49.6% were assigned a putative function (Table 1). While about a quarter of the predicted genes are unique, comparison of broadly conserved *H.chejuensis* genes with those of other completely sequenced prokaryotes indicated that *H.chejuensis* is distantly affiliated to *Pseudomonas* spp. (Supplementary Figures S1 and S2).

Basic metabolic capabilities are well equipped to support a free-living, marine heterotrophic lifestyle. The bacterium has a complete repertoire of enzymes for central carbon metabolism including glycolysis, pentose phosphate pathway and TCA cycle, as well as those required for biosynthesis of nucleotides and 20 amino acids. The presence of genes for one putative carbon monoxide dehydrogenases and one hydrogenase complex without other genes for autotrophy implies that, when available organic nutrients are scarce, *H.chejuensis* might rely on the lithoheterotrophic strategy (25). Genes for inorganic sulfur oxidation, however, were not identified. We also found genes for the respiratory nitrate reductase complex.

Adaptation to the marine environment

The number of genes dedicated to transcriptional regulation or environmental sensing amounts to 362, which corresponds to ~5.3% of total predicted genes (Supplementary Table S1). This is in accordance with the tendency that the number of
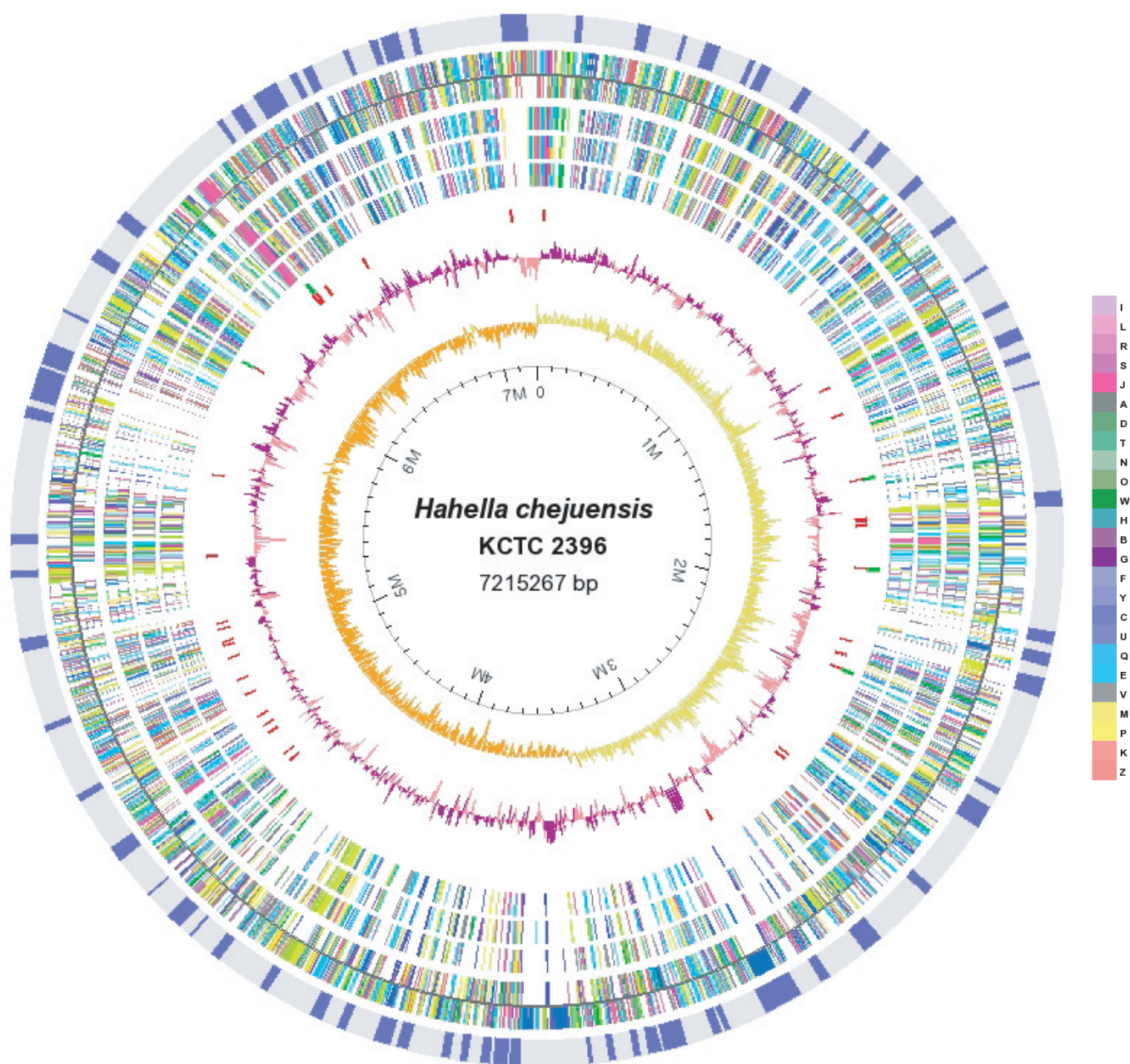
Figure 2. Circular representation of the *H.chejuensis* chromosome. Blue patches on the outermost circle indicate genomic islands. Circles 2 and 3 show CDSs transcribed clockwise and counter-clockwise, which are color-coded according to the COG functional classes as designated in the inset. Detailed function description for each one-letter classification code is available from ftp://ftp.ncbi.nih.gov/pub/COG/COG/fun.txt. Next three circles indicate the locations of homologous genes in *P.aeruginosa* PAO1, *Vibrio cholerae* El Tor N16961 and *E.coli* K-12 (reciprocal best hits, BLASTP bit score/self-score >0.3). Circles 7 and 8 denote rRNA genes and tRNA genes, respectively. Circles 9 and 10 are plots of G+C content and cumulative (G−C)/(G+C) deviation (>0, yellow; <0, orange).

regulatory genes increases as the genome size increases (26). Most common regulator types in *H.chejuesnsis* include LysR, AraC, TetR and MerR. The bacterium also possesses four major sigma-70 factors, two extracytoplamic function sigma factors and one sigma-54 factor. There are more than 20 proteins that have the sigma-54 interaction module. The number of putative two-component system (47 sensors, 103 response regulators and 23 sensor-response regulator hybrids) is overrepresented compared with other bacterial genomes. In addition, the bacterium has a complex chemosensory system with 35 genes encoding putative chemotactic sensory transducer proteins. However, the typical quorum-sensing system seems absent as no homologs of *luxI* could be found.

H.chejuensis has a wide range of transporters for sugars, peptides/amino acids, phosphate, manganese, molybdate, nickel and drugs. Sugar transport systems, however, appear highly biased to ABC transport systems as there are 11 ABC-type transporters but phosphotransferase system is incomplete. This phenomenon is rather scarce but often observed in some pathogens (27). A variety of extracellular hydrolytic enzymes

Table 1. General features of the *H.chejuensis* genome

Size (bp)	7 215 267
Number of predicted protein-coding sequences (CDS)	6783
Percentage coding	88.3
Average gene length (bp)	940
G+C content (%)	54.8
rRNA operons	5
tRNAs	67
sRNAs	22
CDS summary	
Assigned function	3362
Unknown function	1132
Conserved hypothetical[a]	683
Hypothetical[b]	1616

[a]Similar to other hypothetical proteins, or having marginal similarity without HMM matches.
[b]No substantial similarity to NCBI NR protein dataset (E-value $> 1 \times 10^{-5}$).

represented by the *H.chejuensis* genome, such as proteases, lipases, nucleases, chitinases and cellulases, could be advantageous once macromolecular nutrients become available. Along with the high portion of regulatory proteins and transporters for a variety of nutrients, these features imply the functional diversity and adaptability of *H.chejuensis* to changing marine environments.

Like other marine bacteria, *H.chejuensis* requires 2% NaCl for optimal growth (8). Na^+ is essential for marine or halophilic bacteria as transmembrane Na^+ gradient is utilized for uptake of nutrients and flagellar rotation (28). In general, Na^+/H^+ antiporter generates the sodium motive force for these cellular processes, but Na^+-translocating respiratory NADH:ubiquinone oxidoreducatase is widely distributed among Gram-negative marine bacteria in addition to the primary H^+ pump and Na^+/H^+ antiporter. Genome analysis identified the same type of respiratory complex in *H.chejuensis* and multiple Na^+/H^+ antiporters including a multi-subunit Na^+/H^+ antiporter system.

Redundant genes and genomic islands

An interesting feature of the *H.chejuensis* genome is the multiplicity of homologous genes encoding functionally equivalent proteins (Supplementary Table S2). There are dozens of cases where the same function is redundantly encoded by two to four independent genes. As for gene sets, there are two loci each for F_0F_1-type ATP synthesis, flagellar biogenesis and type III protein secretion. When all-against-all similarity searches were performed to identify recent gene duplication within the genome, overall identities among the homologous genes were far below than those of the closest proteins from other sequenced genomes. While in many cases one member best matches to proteins in γ-*Proteobacteria*, the other members are similar to those in various other taxa. These observations support that the origin of multiplicity is likely horizontal gene transfer rather than duplication of genes in the *H.chejuensis* genome.

Like many other bacteria (29), horizontal gene transfer seems to have had essential roles in shaping the *H.chejuensis* genome. Based on genomic anomalies and phylogenetic context, the bacterium appears to have at least 69 genomic islands (GIs) constituting ∼23.0% of the chromosome (Figure 2). Genes or gene clusters contained in the islands include those involved in biosynthesis of exopolysacchrides, toxins, polyketides or non-ribosomal peptides, iron utilization, motility, type III protein secretion, or pigmentation. Of them 32 contained homologs of genes associated with mobile elements such as Rhs elements, insertion sequence elements, transposons, bacteriophages and group II introns (Supplementary Table S3). Genes encoding Rhs family proteins are the largest group in the *H.chejuensis* genome in terms of both abundance and length (∼1.7% of the chromosome). In most cases, the Rhs proteins are very closely related to those in the archaeon *Methanosarcina barkeri* or the firmicute *Clostridium thermocellum*.

Potential virulence-associated genes

H.chejuensis produces a large amount of extracellular polysaccharides (EPSs) (30). EPSs are responsible for development of biofilms, and often act as a virulence factor in pathogenic bacteria. We found five gene clusters that may be involved in the synthesis of exopolysaccharides, all of which overlap GIs partly or entirely (Supplementary Figure S3 and Table S3). Among the gene clusters is one located at the 4.8 Mb region encoding genes for key enzymes such as UDP-glucose dehydrogenase (*ugd*), Wzy-type polymerase and Wzx flippase. UGD produces UDP-D-glucuronate, which is known to be a building block for production of capsular polysaccharides in several pathogenic bacteria and colanic acid in *E.coli* (31,32).

Pore-forming hemolysin and RTX toxin play important roles in many pathogenic Gram-negative bacteria with their cytotoxic activities (33). Out of the seven RTX toxin homologs and three hemolysin homologs found in *H.chejuensis*, five are included in GIs. One of the striking findings from the *H.chejuensis* genome is the unexpected presence of two type III secretion systems (TTSSs) (34,35) located at positions 3.34–3.37 and 5.25–5.29 Mb that are similar to those present in *Yersinia* spp., *Vibrio* spp., *Pseudomonas aeruginosa* and *Aeromonas* spp. (Figure 3). While the two TTSSs in *H.chejuensis* belong to the same subfamily of TTSSs, only one of them is located in a GI. Presence of the homologs of these virulence determinants suggests that *H.chejuensis* probably is a pathogen of marine eukaryotes.

Identification of an algicidal pigment

Experiments with the extract of *H.chejuensis* indicated that both crude and purified preparations of the red pigment are responsible for the rapid cell lysis of *C.polykrikoides* (data not shown and Figure 4). Therefore, genes possibly involved in synthesis of secondary metabolites were searched for in the genome sequence of *H.chejuensis*, and a gene cluster harboring genes similar to the red genes of *Streptomyces coelicolor* A3(2) (36) was suspected to be responsible for the biosynthesis of the red pigment (Figure 5A).

When fosmid clones carrying the complete gene cluster were grown on agar plates, colonies located close to those of *H.chejuensis* or grown in a medium containing the *H.chejuensis* extract turned red, indicating that this genomic region has the gene set required for pigment biosynthesis (Figure 5A). Further supporting the observation, transposon insertions in the red homologs resulted in loss of colony color when tested with a variant clone that constitutively produces the red pigment.
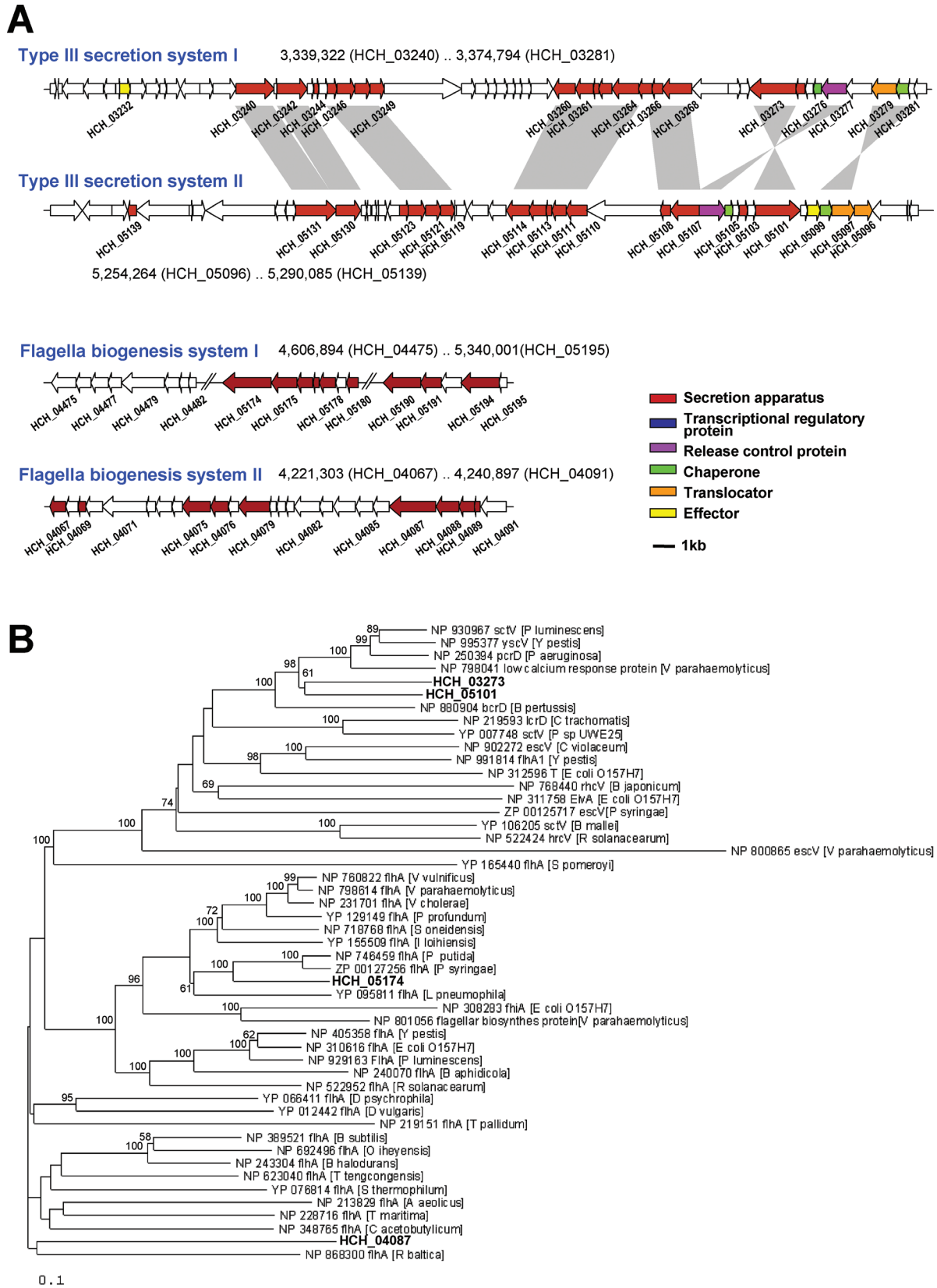
Figure 3. TTSSs and flagellar biogenesis systems in the *H.chejuensis* genome. (**A**) Genetic organizations of two TTSSs and two flagellar biogenesis systems found in *H.chejuensis*. Genes with similar functions are indicated in the same color. Genes conserved between the two TTSSs are shown by gray boxes. Homologous genes between TTSSs and the flagellar systems are shown by red color. (**B**) Phylogenetic tree based on SctV/FlhA protein sequences. Gonnet 250 was used as a protein weight matrix and proteins are neighbor-joined. Bootstrap values of >50% (for 1000 iterations) are shown and the scale bar represents number of substitutions per site.

A maximum absorbance of the purified red pigment at 535 and 470 nm in acidic and basic conditions, respectively, suggested that the pigment is a prodigiosin-like compound. Through LC-ESI-MS/MS analysis, the fragmentation pattern of a base peak [23.73 min; m/z 324.2, (M+H)$^+$] from the red pigment was shown to be identical to that of the antibiotic prodigiosin, which was further confirmed by ^{1}H NMR and ^{13}C NMR analyses (Materials and Methods and Figure 5B).

CONCLUSIONS

Sequence data suggest that *H.chejuensis* is a versatile microorganism well suited to the marine lifestyle. The large genome contains a plethora of regulatory genes, transporter-encoding

Figure 4. Algicidal activity of the purified red pigment (RP10356) of *H.chejuensis* against *C.polykrikoides* strain BWE0109 at various concentrations.

genes and those for secreted hydrolytic enzymes. The large genome size and a number of functionally redundant genes may play roles in diverse environmental conditions. Many of the niche-specific genes are situated in GIs. In addition, the presence of a list of potential virulence genes allows us to speculate that it may occupy a unique ecological position as what might be a 'predator' of microalgae.

The algicidal pigment, whose identity as prodigiosin was unveiled from gene mining and structural analyses, was shown to be highly effective against *C.polykrikoides*, a major red-tide microalga. Prodigiosin, a red pigment known for centuries, is a cytotoxic compound showing a broad range of activity (37). This tripyrrole alkaloid also induces apoptosis in human cancer cells (38). However, its activity against dinoflagellates has not been explored before. Pigments produced by marine bacteria may function as protective agents against solar radiation or protozoan grazing (39). Though the biological roles of prodigiosins in the producing organisms have not been defined, we infer that prodigiosin, as well as toxins, TTSS-delivered proteins and other virulence effectors, contributes to the pathogenic lifestyle of *H.chejuensis*.

Considering that clay dispersal has been the only practical management strategy (6,7), our work opens a possibility that biological control of harmful algal blooms with a microbe antagonistic to the causative organisms or its products might be feasible. Specifically, use of *H.chejuensis* or its bioactive substances may provide a promising alternative for controlling algal blooms caused by *C.polykrikoides*. The genome information gained through our study would function as a guide to understanding the bacteria–algae interactions and utilizing it for devising more rationale-based control measures. Finally, our work illustrates an example of accelerated discovery process through genomics and provides a research model in marine bioprospecting.

Sequence data, annotations and detailed information of the genome are also accessible through the Genome Encyclopedia of Microbes (GEM), http://www.gem.re.kr/.
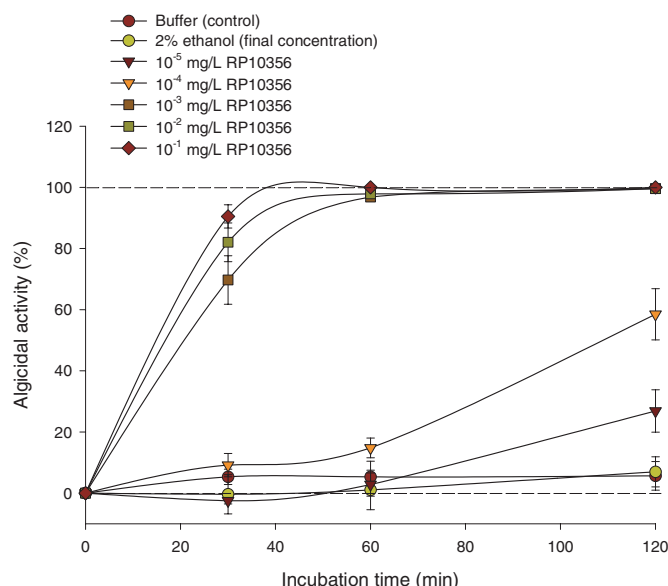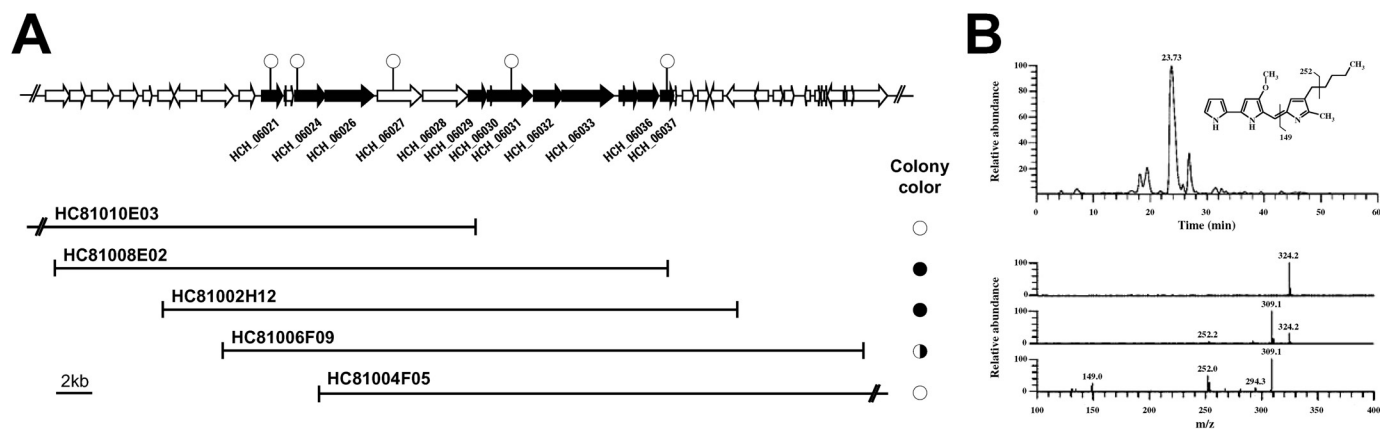

Figure 5. Biosynthesis and structure of the red pigment of *H.chejuensis*. (**A**) The genomic region involved in pigment biosynthesis. Genes homologous to those in the *S.coelicolor* A3(2) red cluster are indicated by filled arrows. Horizontal lines indicate fosmid clones containing some or all of the pigment-synthetic genes. Colony colors: red, filled circles; white, open circles; white with some constitutively red variants, half-filled circle. Open lollipops: locations of the Tn5 insertions that result in loss of the colony color in a red variant of HC81006F09. (**B**) Structural determination of the red pigment. Upper part, LC-ESI-MS in the positive-ion mode; lower part, MS/MS fragmentation pattern of the base peak (23.73 min).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

REFERENCES

1. Burkholder,J.M. (1999) The lurking perils of *Pfiesteria*. *Sci. Am.*, **281**, 42–49.
2. Kaiser,J. (2002) Microbiology The science of *Pfiesteria*: elusive, subtle, and toxic. *Science*, **298**, 346–349.
3. Anderson,D.M., Cembella,A.D. and Hallegraeff,G.H. (eds) (1998) *Physiological Ecology of Harmful Algal Blooms*. Springer-Verlag, New York.
4. Morris,J.G.J. (1999) Harmful algal blooms: an emerging public health problem with possible links to human stress on the environment. *Annu. Rev. Energy Environ.*, **24**, 367–390.
5. Sellner,K.G., Doucette,G.J. and Kirkpatrick,G.J. (2003) Harmful algal blooms: causes, impacts and detection. *J. Ind. Microbiol. Biotechnol.*, **30**, 383–406.
6. Taylor,F.J. and Trainer,V.L. (eds) (2002) *Harmful Algal Blooms in the PICES Region of the North Pacific*. North Pacific Marine Science Organization, Sidney BC V8L 4B2.
7. Sengco,M.R. and Anderson,D.M. (2004) Controlling harmful algal blooms through clay flocculation. *J. Eukaryot. Microbiol.*, **51**, 169–172.
8. Lee,H.K., Chun,J., Moon,E.Y., Ko,S.H., Lee,D.S., Lee,H.S. and Bae,K.S. (2001) *Hahella chejuensis* gen. nov., sp. nov., an extracellular-polysaccharide-producing marine bacterium. *Int. J. Syst. Evol. Microbiol.*, **51**, 661–666.
9. González,J.M. and Whitman,W.B. (2002) *Oceanospirillum* and related genera. In Dworkin,M. *et al.* (eds), *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community*. 3rd edn. Springer-Verlag, New York. Release 3.9.
10. Giovannoni,S. and Rappé,M. (2000) Evolution, diversity, and molecular ecology of marine prokaryotes. In Kirchman,D.L. (ed.), *Microbial Ecology of the Oceans*. John Wiley & Sons, New Jersey, pp. 47–84.
11. Venter,J.C., Remington,K., Heidelberg,J.F., Halpern,A.L., Rusch,D., Eisen,J.A., Wu,D., Paulsen,I., Nelson,K.E., Nelson,W. *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, **304**, 66–74.
12. Faust,M.A. and Gulledge,R.A. (2002) *Identifying Harmful Marine Dinoflagellates*. Contributions from the United States National Herbarium.
13. Mayali,X. and Azam,F. (2004) Algicidal bacteria in the sea and their impact on algal blooms. *J. Eukaryot. Microbiol.*, **51**, 139–144.
14. Grigoriev,A. (1998) Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.*, **26**, 2286–2290.
15. Salzberg,S.L., Salzberg,A.J., Kerlavage,A.R. and Tomb,J.F. (1998) Skewed oligomers and origins of replication. *Gene*, **217**, 57–67.
16. Badger,J.H. and Olsen,G.J. (1999) CRITICA: coding region identification tool invoking comparative analysis. *Mol. Biol. Evol.*, **16**, 512–524.
17. Delcher,A.L., Harmon,D., Kasif,S., White,O. and Salzberg,S.L. (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.*, **27**, 4636–4641.
18. Rutherford,K., Parkhill,J., Crook,J., Horsnell,T., Rice,P., Rajandream,M.A. and Barrell,B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics*, **16**, 944–945.
19. Kanehisa,M., Goto,S., Kawashima,S., Okuno,Y. and Hattori,M. (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.*, **32**, D277–D280.
20. Karp,P.D., Paley,S. and Romero,P. (2002) The Pathway Tools software. *Bioinformatics*, **18**, 225–232.
21. Harris,J.K., Kelley,S.T., Spiegelman,G.B. and Pace,N.R. (2003) The genetic core of the universal ancestor. *Genome Res.*, **13**, 407–412.
22. 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.
23. Felsenstein,J. (2004) *PHYLIP (Phylogeny Inference Package) 3.6b edn*. Department of Genetics, University of Washington, Seattle.
24. Yoon,S.H., Hur,C.-G., Kang,H.-Y., Kim,Y.H., Oh,T.K. and Kim,J.F. (2005) A computational approach for identifying pathogenicity islands in prokaryotic genomes. *BMC Bioinformatics*, **6**, 184.
25. Moran,M.A., Buchan,A., Gonzalez,J.M., Heidelberg,J.F., Whitman,W.B., Kiene,R.P., Henriksen,J.R., King,G.M., Belas,R., Fuqua,C. *et al.* (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature*, **432**, 910–913.
26. Stover,C.K., Pham,X.Q., Erwin,A.L., Mizoguchi,S.D., Warrener,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M. *et al.* (2000) Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. *Nature*, **406**, 959–964.
27. Siebold,C., Flukiger,K., Beutler,R. and Erni,B. (2001) Carbohydrate transporters of the bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS). *FEBS Lett.*, **504**, 104–111.
28. Hayashi,M., Nakayama,Y. and Unemoto,T. (2001) Recent progress in the Na(+)-translocating NADH-quinone reductase from the marine *Vibrio alginolyticus*. *Biochim. Biophys. Acta*, **1505**, 37–44.
29. Doolittle,W.F. (1999) Phylogenetic classification and the universal tree. *Science*, **284**, 2124–2129.
30. Yim,J.H., Kim,S.J., Aan,S.H. and Lee,H.K. (2004) Physiochemical and rheological properties of a novel emulsifier, EPS-R, produced by the marine bacterium *Hahella chejuensis*. *Biotechnol. Bioprocess Eng.*, **9**, 405–413.
31. Dougherty,B.A. and van de Rijn,I. (1993) Molecular characterization of hasB from an operon required for hyaluronic acid synthesis in group A streptococci. Demonstration of UDP-glucose dehydrogenase activity. *J. Biol. Chem.*, **268**, 7118–7124.
32. Whitfield,C. and Roberts,I.S. (1999) Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Mol. Microbiol.*, **31**, 1307–1319.
33. Lally,E.T., Hill,R.B., Kieba,I.R. and Korostoff,J. (1999) The interaction between RTX toxins and target cells. *Trends Microbiol.*, **7**, 356–361.
34. Cornelis,G.R. (2002) The *Yersinia* Ysc-Yop 'type III' weaponry. *Nature Rev. Mol. Cell Biol.*, **3**, 742–752.
35. Kim,J.F. (2001) Revisiting the chlamydial type III protein secretion system: clues to the origin of type III protein secretion. *Trends Genet.*, **17**, 65–69.
36. Cerdeño,A.M., Bibb,M.J. and Challis,G.L. (2001) Analysis of the prodiginine biosynthesis gene cluster of *Streptomyces coelicolor* A3(2): new mechanisms for chain initiation and termination in modular multienzymes. *Chem. Biol.*, **8**, 817–829.
37. Fürstner,A. (2003) Chemistry and biology of roseophilin and the prodigiosin alkaloids: a survey of the last 2500 years. *Angew. Chem. Int. Ed. Engl.*, **42**, 3582–3603.
38. Peréz-Tomás,R., Montaner,B., Llagostera,E. and Soto-Cerrato,V. (2003) The prodigiosins, proapoptotic drugs with anticancer properties. *Biochem. Pharmacol.*, **66**, 1447–1452.
39. Matz,C., Deines,P., Boenigk,J., Arndt,H., Eberl,L., Kjelleberg,S. and Jurgens,K. (2004) Impact of violacein-producing bacteria on survival and feeding of bacterivorous nanoflagellates. *Appl. Environ. Microbiol.*, **70**, 1593–1599.