

EXTENDED GENOME REPORT

Open Access



Genome sequence of *Shimia* str. SK013, a representative of the *Roseobacter* group isolated from marine sediment

Saranya Kanukollu¹, Sonja Voget², Marion Pohlner¹, Verona Vandieken¹, Jörn Petersen³, Nikos C. Kyrpides^{4,5}, Tanja Woyke⁴, Nicole Shapiro⁴, Markus Göker³, Hans-Peter Klenk⁶, Heribert Cypionka¹ and Bert Engelen^{1*}

Abstract

Shimia strain SK013 is an aerobic, Gram-negative, rod shaped alphaproteobacterium affiliated with the *Roseobacter* group within the family *Rhodobacteraceae*. The strain was isolated from surface sediment (0–1 cm) of the Skagerrak at 114 m below sea level. The 4,049,808 bp genome of *Shimia* str. SK013 comprises 3,981 protein-coding genes and 47 RNA genes. It contains one chromosome and no extrachromosomal elements. The genome analysis revealed the presence of genes for a dimethylsulfoniopropionate lyase, demethylase and the trimethylamine methyltransferase (*mttB*) as well as genes for nitrate, nitrite and dimethyl sulfoxide reduction. This indicates that *Shimia* str. SK013 is able to switch from aerobic to anaerobic metabolism and thus is capable of aerobic and anaerobic sulfur cycling at the seafloor. Among the ability to convert other sulfur compounds it has the genetic capacity to produce climatically active dimethyl sulfide. Growth on glutamate as a sole carbon source results in formation of cell-connecting filaments, a putative phenotypic adaptation of the surface-associated strain to the environmental conditions at the seafloor. Genome analysis revealed the presence of a flagellum (*fla1*) and a type IV pilus biogenesis, which is speculated to be a prerequisite for biofilm formation. This is also related to genes responsible for signalling such as N-acyl homoserine lactones, as well as quip-genes responsible for quorum quenching and antibiotic biosynthesis. Pairwise similarities of 16S rRNA genes (98.56 % sequence similarity to the next relative *S. haliotis*) and the *in silico* DNA-DNA hybridization (21.20 % sequence similarity to *S. haliotis*) indicated *Shimia* str. SK013 to be considered as a new species. The genome analysis of *Shimia* str. SK013 offered first insights into specific physiological and phenotypic adaptation mechanisms of *Roseobacter*-affiliated bacteria to the benthic environment.

Keywords: Anaerobic metabolism, Cell-connecting filaments, Quorum quenching, Flagella gene cluster, DMSP, DMSO reductase, Denitrification

Introduction

The *Roseobacter* group is known for its worldwide distribution and its broad metabolic versatility in a great variety of marine habitats [1–3]. About 25 % of all *Roseobacter* species with validly published names (42 out of 168) have a benthic origin [4]. In marine sediments, they can contribute up to 11 of all 16S rRNA genes and up to 10 % of total cell counts [5, 6], but still little is known about the specific distribution and physiology of roseobacters in this habitat.

Shimia str. SK013, analysed in the present study, was isolated from the top centimeter of Skagerrak sediments at a water depth of 114 m below sea level (mbsl) [7]. The strain is affiliated with the genus *Shimia* which was first proposed by Choi and Cho in 2006 [8] in honor of Dr. Jae H. Shim, for his contributions to marine plankton ecology in Korea. According to Pujalte et al. [4], the genus *Shimia* consists of four species, with a fifth species *Shimia sagamensis* recently included. Members of the genus *Shimia* were isolated from different marine habitats: e.g. *S. haliotis* was isolated from the intestinal tract of the abalone *Haliotis discus hannai* [9], *S. biformata* from surface sea water [10], *S. isopora* from reef building corals

* Correspondence: engelen@icbm.de

¹Institute for Chemistry and Biology of the Marine Environment (ICBM), Oldenburg, Germany

Full list of author information is available at the end of the article

[11] and *S. marina* from a fish farm biofilm [8]. The new species affiliated to the genus *Shimia* (*Shimia sagamensis*) was isolated from cold seep sediment [12]. The sequenced genome of *Shimia* str. SK013 will allow for genetic comparison between the strain and other organisms of benthic origin, additional sediment-derived roseobacters and close relatives isolated from different habitats.

Here, we present the genome of *Shimia* str. SK013 with special emphasis on the genes involved in sulfur cycling such as dimethylsulfoniopropionate (DMSP) degradation and dimethyl sulfoxide reduction, as well as other anaerobic pathways such as nitrate reduction. The second focus is on genes which may be indicative for biofilm formation (pili, flagella and quorum

sensing) as an adaptation to their surface-associated lifestyle.

Organism information

Classification and features

Sediment samples were collected in July 2011 during a cruise with the RV 'Heincke' (expedition HE361) to the eastern North Sea. The strain was isolated from surface sediment (0–1 cm) of the Skagerrak (Site 27, 57°61.28'N, 8°58.18'E) at 114 mbsl from an aerobic enrichment culture. *Shimia* str. SK013 is a Gram-negative, motile, rod shaped bacterium with a length of 1.8 to 2.0 µm and a width of approximately 0.5 µm (Table 1; Fig. 1). Colonies are small, slightly domed and white to transparent on artificial sea water medium agar plates, but cream-coloured

Table 1 Classification and general features of *Shimia* str. SK013 in accordance with the MIGS recommendations published by the Genome Standards Consortium [46]

MIGS ID	Property	Term	Evidence code ^a
	Classification	Domain <i>Bacteria</i>	TAS [47]
		Phylum <i>Proteobacteria</i>	TAS [48]
		Class <i>Alphaproteobacteria</i>	TAS [49, 50]
		Order <i>Rhodobacterales</i>	TAS [50, 51]
		Family <i>Rhodobacteraceae</i>	TAS [50, 51]
		Genus <i>Shimia</i>	TAS [8]
		Species <i>Shimia</i>	TAS [8]
			Strain SK013 (IMG2608642164)
	Gram stain	negative	IDA
	Cell shape	Rod shaped	IDA
	Motility	Motile	IDA
	Sporulation	none	NAS
	Temperature range	Mesophile; 10–35 °C	IDA
	Optimum temperature	25–30 °C	IDA
	pH range; Optimum	5–9; 7	IDA
	Carbon source	Sugars, amino acids	IDA
MIGS-6	Habitat	Marine	IDA
MIGS-6.3	Salinity	0–5 % NaCl (w/v)	IDA
MIGS-22	Oxygen requirement	Aerobic	IDA
MIGS-15	Biotic relationship	Unknown	NAS
MIGS-14	Pathogenicity	non-pathogen	NAS
MIGS-4	Geographic location	North Sea/Skagerrak area	IDA
MIGS-5	Sample collection	July 24, 2011	IDA
MIGS-4.1	Latitude	57°36.77'N	IDA
MIGS-4.2	Longitude	08°35.41'E	IDA
MIGS-4.3	Depth	114 m below sea level	IDA
MIGS-4.4	Altitude	Unknown	

^aEvidence codes - *IDA* Inferred from Direct Assay, *TAS* Traceable Author Statement (i.e., a direct report exists in the literature), *NAS* Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [52]

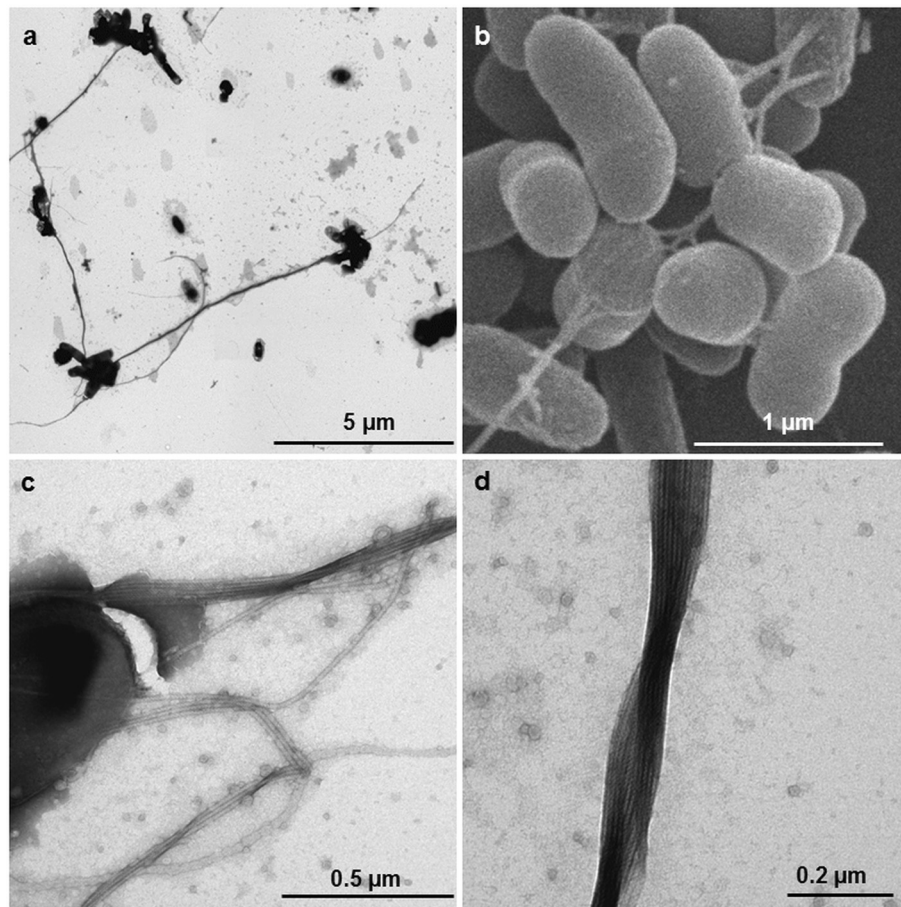


Fig. 1 Micrographs of *Shimia* str. SK013. **a** Transmission electron micrograph (TEM) showing aggregation of cells and long fibers **(b)** Scanning electron micrograph (SEM) of cells grown on glutamate with formation of cell-connecting fibers **(c)** TEM of a single cell with cell-connecting fibers **(d)** Closer view (TEM) on a bundle of fibers

or beige in marine broth medium agar plates. The strain is mesophilic (range: 10–35 °C, $T_{opt} = 30$ °C), neutrophilic (optimum pH: 6–7) and halophilic (optimum: 2–3 % w/v). *Shimia* str. SK013 grows well in liquid medium but relatively slowly on agar-solidified marine broth and artificial sea water medium. The strain is able to utilize various substrates such as glucose, lactose, glutamate, mannose, xylose, acetate and citrate. When *Shimia* str. SK013 grows in ASW medium with glutamate as sole carbon source, cell-connecting filaments that might represent bundle-forming pili or specialized flagella are induced (Fig. 1). However, these structures were not observed in cultures amended with any other tested substrate (see above). The 16S rRNA gene sequence of *Shimia* str. SK013 (1453 bp) was analysed using ARB [13] and revealed 98.56 % sequence similarity to the next relative, *Shimia haliotis*. Furthermore, in the phylogenetic tree, *Shimia* str. SK013 is branching together with the other *Shimia* species except *Shimia biformata* (Fig. 2).

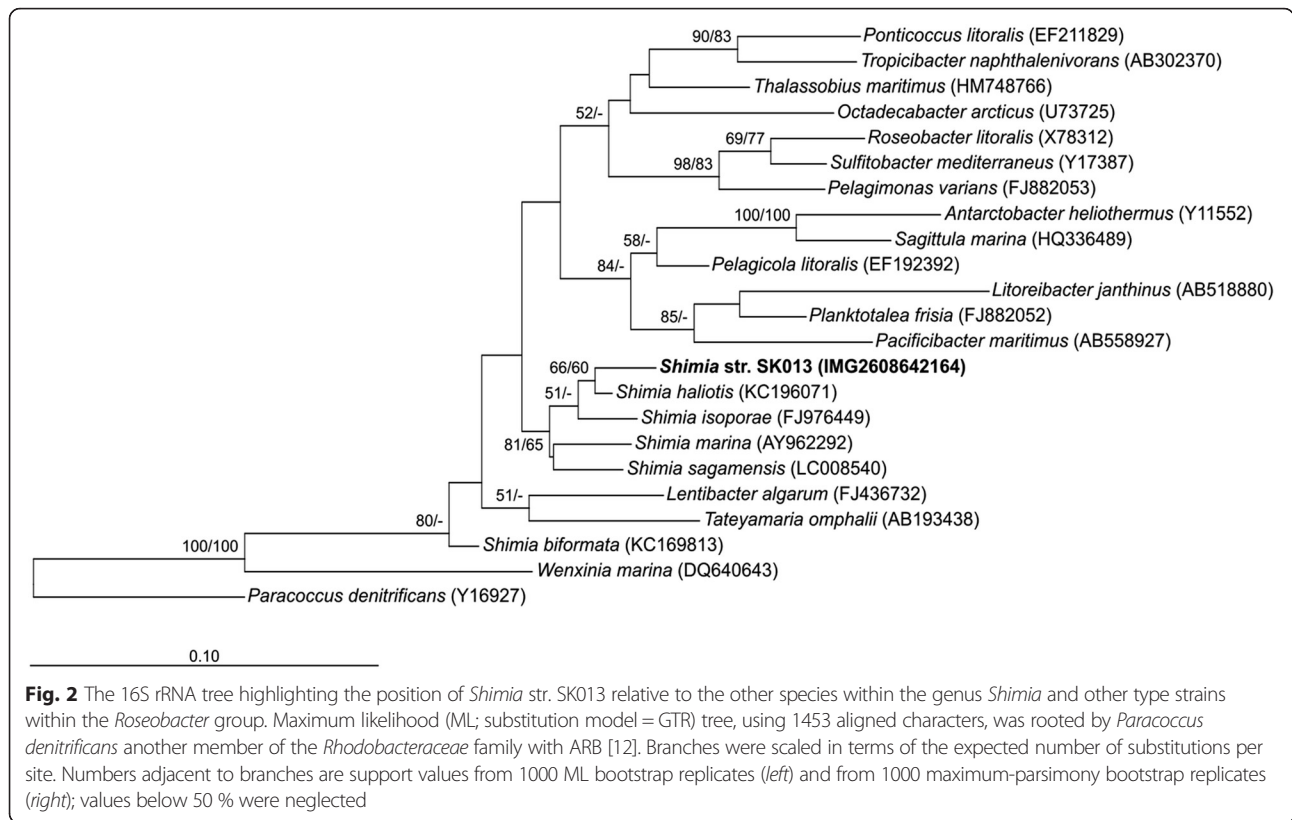
Genome sequencing information

Genome project history

Shimia str. SK013 was selected for draft genome sequencing based on its physiological and phenotypical features and its benthic origin. The information related to this project is summarized in Table 2. The draft genome is deposited in the Genomes On Line Database [14] and in the Integrated Microbial Genome database [15]. The Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LAJH00000000.1.

Growth conditions and genomic DNA preparation

Shimia str. SK013 was enriched and isolated from agar plates containing artificial sea water medium [16] with DMS (100 μM) and lactate (5 mM) as substrates, incubated at 15 °C. The genomic DNA extraction was performed using a DNA isolation kit (MO BIO, Carlsbad, CA, USA), following the manufactures instructions.



Genome sequencing and assembly

Whole-genome sequencing was performed using the Illumina technology. Preparation of paired-end sequencing library with the Illumina Nextera XT library preparation kit and sequencing of the library using the Genome Analyzer Iix were performed as described by the manufacturer (Illumina, San Diego, CA, USA). A total of 11,098,582 paired-end reads

were derived from sequencing and trimmed using Trimmomatic version 0.32 [17]. *De novo* assembly of all trimmed reads with SPAdes version 3.5.0 [18] resulted in 28 contigs and 137.9-fold coverage. A summary of project information is shown in Table 2.

Table 2 Genome sequencing project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Draft
MIGS-28	Libraries used	Nextera xt
MIGS 29	Sequencing platforms	Illumina GAii, PacBio
MIGS 31.2	Fold coverage	
MIGS 30	Assemblers	SPAdes v3.5
MIGS 32	Gene calling method	Prodigal v2.5
	Locus Tag	SHIM
	Genbank ID	LAJH00000000
	GenBank Date of Release	September 16, 2015
	GOLD ID	Gp0103193
	BIOPROJECT	PRJNA277163
MIGS 13	Source Material Identifier	SAMN03387008
	Project relevance	Environmental

Genome annotation

Protein-coding genes were identified as part of the genome annotation pipeline the Integrated Microbial Genomes Expert Review platform using Prodigal v2.50. The predicted CDS were translated and used to search the CDD, KEGG, UniProt, TIGRFam, Pfam and InterPro databases. These data sources were combined to assert a product description for each predicted protein. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [19], RNAmmer [20], Rfam [21], TMHMM [22] and SignalP [23]. Additional gene prediction analyses and functional annotation were performed within the IMG-Expert Review platform [24].

Genome properties

The genome analysis showed the presence of 28 scaffolds corresponding to one large chromosome with a total length of 4,049,808 bp and a G + C content of 57.22 % (Table 3). The absence of additional

Table 3 Genome statistics of *Shimia* str. SK013

Attribute	Value	% of total
Genome size (bp)	4,049,808	100.00
DNA coding (bp)	3,677,855	90.82
DNA G + C (bp)	2,317,341	57.22
DNA scaffolds	28	
Total genes	4028	100.00
Protein-coding genes	3981	98.83
RNA genes	47	1.17
Pseudo genes	0	
Genes in paralog clusters	3069	76.19
Genes with function prediction	3317	82.35
Genes assigned to COGs	2860	71.00
Genes with Pfam domains	3365	83.54
Genes with signal peptides	370	9.19
Genes with transmembrane helices	911	22.62
CRISPR repeats	0	

extrachromosomal elements was inferred based on the absence of RepABC, RepA, RepB and DnaA-like modules for plasmid replication and maintenance that are characteristic for *Rhodobacteraceae* [25]. In total, 4,028 genes were predicted, in which 3,981 were protein-coding genes and 47 RNA genes. About 82.35 % were protein-coding genes with a putative function while

those remaining were annotated as hypothetical proteins. The genome statistics are further provided in Table 3 and in Fig. 3. The distribution of genes into functional categories (clusters of orthologous groups) is shown in Table 4.

Insights from the genome sequence

The genome of *Shimia* str. SK013 contains genes for sulfur cycling that might enable anaerobic growth. Genes for quorum sensing and quorum quenching might support roseobacters to thrive in complex microbial communities found in sediments. Genome comparison (Table 5; Fig. 3) revealed that *Shimia* str. SK013 shares the respective genes with a selection of surface-associated roseobacters and the other two *Shimia* species whose genomes are available. It is well documented that roseobacters are involved in the transformation of DMSP by demethylation or by using the cleavage pathway [6, 26, 27]. Anaerobically, some roseobacters are capable of DMSO reduction resulting in the release of the climatically active DMS [5, 27]. Genes for the DMSP lyase (Shim_05930) and demethylase (Shim_7490) as well as for the DMSO reductase (Shim_34610) found in *Shimia* str. SK013 indicate their functional role in DMSP degradation and DMSO conversion. All three genes are also present within the genome of *S. haliotis* and in those of a selection of surface-associated roseobacters (Table 6): *S. haliotis* (Ga0070219_1011011, 101620, 103192), *Octadecabacter*

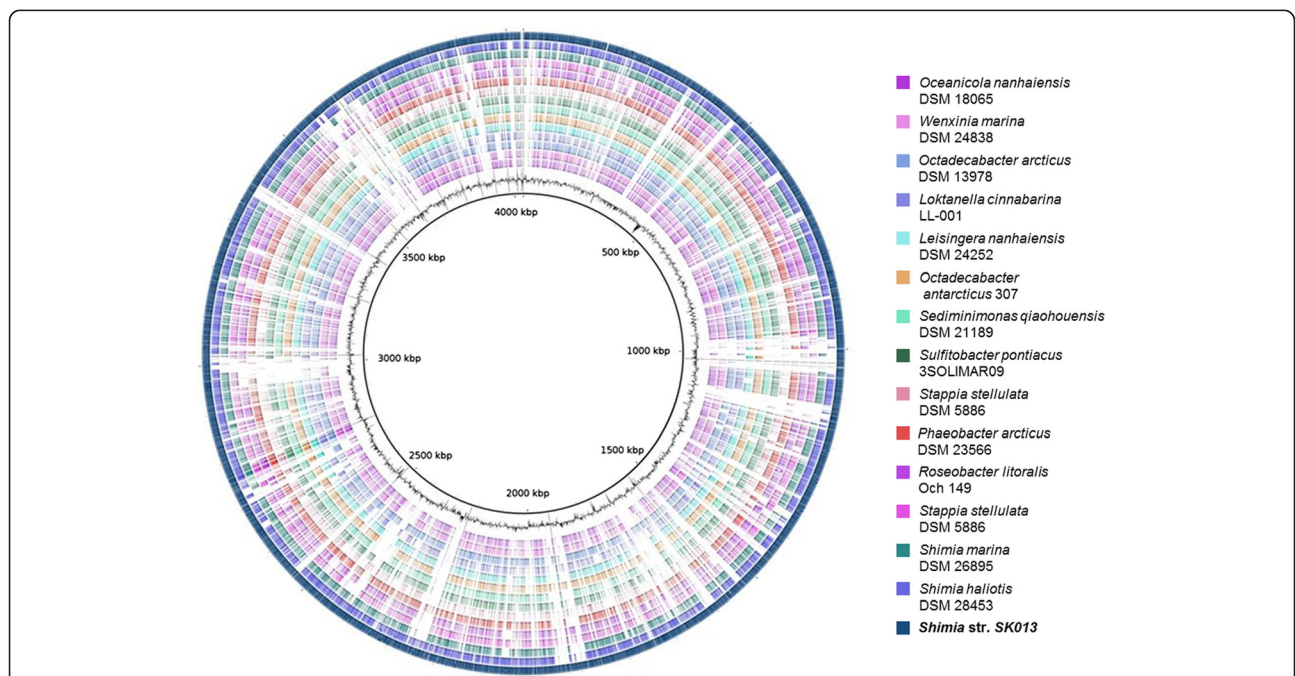


Fig. 3 Graphical representation of the genome of *Shimia* str. SK013. From outside to inside (1–15 color circles): sequence of *Shimia* str. SK013 (1st circle) is compared to the other species within the genus *Shimia* and other type strains within the *Roseobacter* group, (16th circle): G + C content of *Shimia* str. SK013. Comparisons and visualizations are performed with BRIG [53]

Table 4 Number of genes associated with general COG functional categories

Code	Value	%age	Description
J	194	6.02	Translation, ribosomal structure and biogenesis
A	n.a.	n.a.	RNA processing and modification
K	221	6.86	Transcription
L	90	2.80	Replication, recombination and repair
B	2	0.06	Chromatin structure and dynamics
D	25	0.78	Cell cycle control, Cell division, chromosome partitioning
V	61	1.89	Defense mechanisms
T	126	3.91	Signal transduction mechanisms
M	178	5.53	Cell wall/membrane biogenesis
N	50	1.55	Cell motility
U	41	1.27	Intracellular trafficking and secretion
O	156	4.84	Posttranslational modification, protein turnover, chaperones
C	238	7.39	Energy production and conversion
G	203	6.30	Carbohydrate transport and metabolism
E	388	12.05	Amino acid transport and metabolism
F	87	2.70	Nucleotide transport and metabolism
H	174	5.40	Coenzyme transport and metabolism
I	186	5.78	Lipid transport and metabolism
P	143	4.44	Inorganic ion transport and metabolism
Q	130	4.04	Secondary metabolites biosynthesis, transport and catabolism
R	313	9.72	General function prediction only
S	192	5.96	Function unknown
-	1168	29.00	Not in COGs

The total is based on the total number of protein-coding genes in the genome

arcticus (OA238_c10540, c20430, c35930), *Roseobacter litoralis* (RLO149_c019880, c022350, c001820) and *Leisingera nanhaiensis* (Leina_00726, 01164, 02539). *Shimia marina* is missing the genes for DMSP lyase, but also contains genes for DMSP demethylase (Ga0069993_10296, 102173) and DMSO reductase (Ga0069993_106210). Interestingly, the genome of *Shimia* str. SK013 simultaneously contained genes for a sulfite reductase (Shim_12650), sulfur dehydrogenase (*SoxC*; Shim_11330), sulfur oxidizing proteins (*SoxXYZ*; Shim_11380, 11370, 11360) and sulfur oxidation (*SoxA*; Shim_11350). Other than in the *Roseobacter* group-affiliated fosmid found in German tidal-flat sediments [6], the *soxD* gene and the rDSR gene are not present.

We observed all genes necessary for the denitrification pathway such as nitrate reductase (Shim_01900), nitrite reductase (Shim_01920), nitric oxide reductase (Shim_02650) and nitrous oxide reductase (Shim_02640). *Shimia* str. SK013 contains a periplasmic nitrate reductase composed of five subunits [28] such as NapA (Shim_18270), NapB

Table 5 Genome statistics comparison with available genomes of *Shimia* species

Genome name	<i>Shimia</i> str. SK013	<i>S. haliotis</i> DSM 28453	<i>S. marina</i> DSM 26895
Genome Size	4,049,808	3,995,969	4,061,252
Gene Count	4,028	3,953	3,992
Scaffold Count	28	22	32
G + C content (%)	57.22	58.04	57.34
RNA Count	47	58	61
rRNA Count	3	5	5
COG Count	2,860	2,751	2,776
COG (%)	71.00	69.59	69.54
Pfam Count	3,350	3,300	3,365
Pfam (%)	83.17	83.48	84.29
TIGRFam Count	1,148	1,155	1,172
TIGRFam (%)	28.50	29.22	29.36
IMG Pathway Count	223	213	207
IMG Pathway (%)	5.54	5.39	5.19
Horizontally Transferred Count	223	158	135
Horizontally Transferred (%)	5.54	4.00	3.38

(Shim_18300), NapD (Shim_18260), NapE (Shim_04260) and NapG (Shim_18280). The presence of periplasmic nitrate reductase genes suggest the potential for anaerobic respiration [29] in *Shimia* str. SK013, whereas the genus *Shimia* has been described as strictly aerobic until now [8]. Interestingly, anaerobic respiration was also observed in *Leisingera nanhaiensis* [30] and *Phaeobacter inhibens* T5^T [31], which were originally described as strictly aerobic. The genes involved in nitrogen regulation (Shim_09380) and nitrogen fixation regulation (Shim_29520) were also found in the genome of *Shimia* str. SK013. Denitrification genes in *Shimia* str. SK013 showed a strong resemblance to those present in *S. haliotis*, with the exception the genes coding for nitrite reductase and nitrogen fixation regulation (nitrate reductase and subunits; Ga0070219_10142 to 10145, nitric oxide reductase; Ga0070219_106169; nitrous oxide reductase; Ga0070219_106170, nitrogen regulation; Ga0070219_101812). *S. marina* showed only the presence of genes for nitrate reduction (Ga0069993_10650), nitrite reduction (Ga0069993_10648), nitrogen regulation (Ga0069993_102260) and nitrogen fixation regulation (Ga0069993_105163). A comparative search revealed the presence of all the genes involved in the nitrogen cycle that were mentioned above for *Oceanicola nanhaiensis* (SIAM614_16412, 31426, 14520, 22007), *Roseobacter litoralis* (RLO149_c039850, c031550, c017950, c035140), *Phaeobacter arcticus* (Phaar_03838, 02837, 01419, 03079, 04163) and *Sulfitobacter pontiacus* (PM01_06655, 15855, 12625, 02530). Furthermore, the

Table 6 Highlighted genes of *Shimia* str. SK013 present in other roseobacters

Highlighted gene products and locus tags	DMSP lyaseDddP Shim_05930	DMSP Demethylase Shim_07490	DMSO reductase Shim_34610	Trimethylamine methyltransferase Shim_09600, 31260	Type IV pilus biogenesis Shim_13020	AHL acylase QuiP precursor Shim_09300	Homoserine/ homoserine lactone efflux protein Shim_16180	N-AHLs Shim_31370
<i>Shimia haliotis</i>	+	+	+	+	+	+	+	+
<i>Shimia marina</i>		+	+	+	+	+	+	
<i>Oceanicola nanhaiensis</i>	+			+			+	+
<i>Octadecabacter antarcticus</i>		+	+	+		+		
<i>Octadecabacter arcticus</i>	+	+	+	+		+		
<i>Roseobacter litoralis</i>	+	+	+	+		+	+	+
<i>Phaeobacter arcticus</i>		+	+	+	+	+	+	
<i>Stappia stellulata</i>				+			+	
<i>Leisingera nanhaiensis</i>	+	+	+		+	+		
<i>Labrenzia aggregata</i>								
<i>Loktanella cinnabarina</i>				+		+		+
<i>Sulfitobacter pontiacus</i>				+	+	+		
<i>Sediminimonas qiaohouensis</i>		+		+		+	+	+
<i>Wenxinia marina</i>				+		+	+	

Table 7 Digital DDH similarities between *Shimia* str. SK013 and other *Shimia* species, calculated *in silico* with the GGDC server version 2.0 [45]^a

Reference species	Formula 1	Formula 2	Formula 3
<i>Shimia haliotis</i> DSM 28453	37.20 % +/- 3.44	21.20 % +/- 2.34	31.60 % +/- 3.02
<i>Shimia marina</i> DSM 26895	16.70 % +/- 3.25	19.70 % +/- 2.30	16.60 % +/- 2.75

^aThe standard deviations indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size); see [45] for details. The distance formulas are explained in [44]. Formula 2 is recommended, particularly for draft genome (like species above)

genome of *Shimia* str. SK013 revealed genes for the utilization of methylated amines, such as a trimethylamine methyltransferase (*mttB*) (Shim_09600, 31260).

The conspicuous morphological trait of cell-connecting filaments in *Shimia* str. SK013 (Fig. 1) led to the search for the presence of genes involved in the formation of pili and flagella. The bacterial flagellum is one of the signal transduction systems with complex proteins which enables the bacterial reorientation and motility [32]. So far three different types of flagella gene clusters (FGCs) were described, designated *fla1*, *fla2* and *fla3* in *Rhodobacteraceae* that originated from FGC duplications [33]. Genome analysis revealed the presence of a single compact flagella gene cluster of the *fla1*-type on the chromosome (contig_000021; Shim_33080 to Shim_33420) that contains all genes necessary for the assembly of a functional flagellum. Recently, Frank et al. [33] showed for the plasmid curing mutant of *Marinovum algicola* DG898, which is lacking the 143-kb plasmid pMaD5 with a *fla2*-type FGC, a conspicuous morphological similarity with the filamentous structures observed in the current study for *Shimia* str. SK013 (Fig. 1). The bundles of filaments were explained by the presence of an additional chromosome-encoded *fla1*-type flagellum in *Marinovum*. However, genes for type IV pilus biogenesis, which were found in *Shimia* str. SK013 (Shim_13020, Shim_37620) are also present in the genome of *M. algicola* DG898 (MALG_02262) and thus, it remains unclear if the conspicuous bundles at the cell pole are caused by pilus and/or flagellum formation.

As the described morphological traits are often related to a surface-associated lifestyle, we also searched the genome of *Shimia* str. SK013 for genes involved in the production of signalling molecules and quorum sensing as indicators for the communication within biofilms. Earlier studies showed that quorum sensing signals are mainly associated with virulence [34, 35], but recent investigations revealed that these signalling molecules play a significant role in basic metabolic processes [36, 37]. The presence of genes for the production of *N*-acylhomoserine lactones (AHLs) (Shim_31370) and homoserine lactones (Shim_16180) that are part of the quorum sensing system indicate that *Shimia* str. SK013 uses this form of bacterial communication. In contrast, the newly established genome only contains a few additional genes which interfere with quorum sensing such as quorum

quenching or antibiotic biosynthesis related genes (AHL acylase QuiP precursor; Shim_09300) [38–40]. When compared to other selected roseobacters, these three signal molecule genes were also found in *Roseobacter litoralis* (RLO149_c018030, c029420, c006500) and *Sediminimonas qiahouensis* (G568DRAFT_00799, 01106, 03483). This finding was proven by an antiSMASH analysis [41] of the *Shimia* str. SK013 genome, indicating the presence of the type I polyketide synthase (PKS), the homoserine lactone cluster and the bacteriocin gene cluster.

Pairwise similarities of 16S rRNA genes of *Shimia* str. SK013 and the next relative, *Shimia haliotis* were 98.56 %. A genome comparison of *Shimia* str. SK013 with the available draft genomes from the KMG-2 project, Genomic encyclopedia of Bacteria and Archaea (GEBA) [42, 43] of *Shimia haliotis* DSM 28453 (IMG ID: 2619619046) and *Shimia marina* DSM 26895 (IMG ID: 2619618961) was conducted using the online analysis tool “Genome-Genome-Distance Calculator” 2.0 (GGDC). The results of the *in silico* calculated DNA-DNA hybridization (DDH) of *Shimia* str. SK013 suggests that the given genome might belong to a new species based on the low percentages obtained (Table 7). According to the GGDC tool, formula 2 was recommended for the comparison between the draft genomes as it provides higher DDH correlations than Average Nucleotide Identity (ANI) implementations [44, 45]. The analysis showed that *Shimia* str. SK013 only shared a genome sequence similarity of 21 % with *Shimia haliotis* DSM 28453 and 20 % with *Shimia marina* DSM 26895 and thus represents neither a new isolate of the species *S. haliotis* nor of *S. marina*. A direct comparison with the available *Shimia* genomes revealed further differences such as the IMG pathway counts (representing the number of metabolites and macromolecular complexes) and horizontally transferred gene counts (Table 5). Until now, genome sequences of *S. bioformata*, *S. isopora* and *Shimia sagamensis* are not available for additional *in silico* calculated DNA-DNA hybridization or direct genome comparisons. However, as *S. haliotis* was identified as the closest relative by 16S rRNA gene analysis with a 66/60 % bootstrap support, the DDH data provide strong evidence that *Shimia* str. SK013 represents a new species within the genus *Shimia*.

Conclusions

The genome analysis of *Shimia* str. SK013 revealed distinctive genes responsible for DMSP utilization, DMSO, nitrate and nitrite reduction which indicate that this strain is a facultative anaerobic bacterium. The presence of genes responsible for signalling can serve as a guide for identification of quorum sensing compounds, as well as antibiotics potentially responsible for quorum quenching. Based on genome comparison and DNA-DNA hybridization with the next relatives, *Shimia* str. SK013 might represent a new species and should be considered for species description.

Abbreviations

AHLs: acyl homoserine lactones; ASW: artificial sea water; BFP: bundle-forming pili; DMS: dimethyl sulfide; DMSO: dimethyl sulfoxide; DMSP: dimethylsulfoniopropionate; FGC: flagella gene clusters; GGDC: Genome-Genome-Distance Calculator; KMG: 1000 microbial genomes; mbsl: meters below sea level; PKS: polyketide synthase; SignalP: signal peptides; TMHMM: transmembrane helices hidden markov models.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SK, SV wrote the manuscript and analysed the annotated genome and comparative genomes. MP calculated the phylogenetic tree. VV contributed in phenotypic data. JP analysed the annotated genome. MG, NC, TW, NS and HPK provided sequenced, assembled and annotated genomes of *Shimia* species and other roseobacters. HC provided analytical facilities and edited the manuscript. BE planned, helped to execute the study and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors acknowledge the crew and the scientific party of the RV Heincke (expedition HE361), especially Judith Lucas for sampling and starting the enrichment cultures as well as Michael Pilzen and Jutta Graue for their assistance during sampling. Furthermore, we thank Jana Feldkamp and Frank Meyerjurgens for technical assistance and our students, Leon Dlugosch, David Nivia, Eva-Lena Nordmann and Katrin Grosse for preliminary physiological experiments and microscopy pictures. We would also extend our thanks to Marco Dogs, Sven Breider and Thorsten Brinkhoff for valuable discussions during genome analysis and Candice Raeburn for proofreading. This work was mainly conducted in the frame of the *Roseobacter* collaborative research center Transregio-SFB 51, supported by the German Research Foundation (DFG). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Author details

¹Institute for Chemistry and Biology of the Marine Environment (ICBM), Oldenburg, Germany. ²Department of Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Göttingen, Göttingen, Germany. ³Leibniz Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. ⁴Department of Energy Joint Genome Institute, Genome Biology Program, Walnut Creek, California, USA. ⁵Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia. ⁶School of Biology, Newcastle University, Newcastle upon Tyne, UK.

Received: 17 September 2015 Accepted: 18 February 2016

Published online: 12 March 2016

References

- Wagner-Döbler I, Biehl H. Environmental biology of the marine *Roseobacter* lineage. *Annu Rev Microbiol.* 2006;60:255–80.
- Brinkhoff T, Giebel HA, Simon M. Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch Microbiol.* 2008;189:531–9.

- Luo H, Moran MA. Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol Mol Biol Rev.* 2014;78:573–87.
- Pujalte MJ LT, Ruvira MA, Arahall DR, Macià MC. In: Roseonberg EDLE, Thompson F, Lory S, Stackebrandt E, editors. *The Family Rhodobacteraceae*. Berlin: Springer; 2014.
- González JM, Kiene RP, Moran MA. Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha-subclass of the class *Proteobacteria*. *Appl Environ Microbiol.* 1999;65:3810–9.
- Lenk S, Moraru C, Hahnke S, Arnds J, Richter M, Kube M, et al. *Roseobacter* clade bacteria are abundant in coastal sediments and encode a novel combination of sulfur oxidation genes. *ISME J.* 2012;6:2178–87.
- Kanukollu S, Wemheuer B, Herber J, Billerbeck S, Lucas J, Daniel R, et al. Distinct compositions of free-living, particle-associated and benthic communities of the *Roseobacter* group in the North Sea. *FEMS Microbiol Ecol.* 2016;92. doi:10.1093/femsec/fiv145.
- Choi DH, Cho BC. *Shimia marina* gen. nov., sp. nov., a novel bacterium of the *Roseobacter* clade isolated from biofilm in a coastal fish farm. *Int J Syst Evol Microbiol.* 2006;56:1869–73.
- Hyun DW, Kim MS, Shin NR, Kim JY, Kim PS, Whon TW, et al. *Shimia haliotis* sp. nov., a bacterium isolated from the gut of an abalone, *Haliotis discus hannai*. *Int J Syst Evol Microbiol.* 2013;63:4248–53.
- Hameed A, Shahina M, Lin SY, Lai WA, Hsu YH, Liu YC, et al. *Shimia biformata* sp. nov., isolated from surface seawater, and emended description of the genus *Shimia* Choi and Cho 2006. *Int J Syst Evol Microbiol.* 2013;63:4533–9.
- Chen MH, Sheu SY, Chen CA, Wang JT, Chen WM. *Shimia isoporae* sp. nov., isolated from the reef-building coral *Isopora palifera*. *Int J Syst Evol Microbiol.* 2011;61:823–7.
- Nogi Y, Mori K, Uchida H, Hatada Y. *Shimia sagamensis* sp. nov., a marine bacterium isolated from cold-seep sediment. *Int J Syst Evol Microbiol.* 2015; 65:2786–90.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadukumar ARB, et al. A software environment for sequence data. *Nucleic Acids Res.* 2004;32:1363–71.
- Lioliou K, Chen IM, Mavromatis K, Tavernarakis N, Hugenholtz P, Markowitz VM, et al. The Genomes On Line Database (GOLD) in 2009: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res.* 2010;38:D346–54.
- Markowitz VM, Mavromatis K, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics.* 2009;25:2271–8.
- Süss J, Engelen B, Cypionka H, Sass H. Quantitative analysis of bacterial communities from Mediterranean sapropels based on cultivation-dependent methods. *FEMS Microbiol Ecol.* 2004;51:109–21.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19:455–77.
- Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 1997;25:955–64.
- Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 2007;35:3100–8.
- Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, et al. Rfam 12.0: updates to the RNA families database. *Nucleic Acids Res.* 2015;43:D130–7.
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001;305:567–80.
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol.* 2004;340:783–95.
- Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Pillay M, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* 2014;42:D560–7.
- Petersen J. Phylogeny and compatibility: plasmid classification in the genomics era. *Arch Microbiol.* 2011;193:313–21.
- Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE, et al. Genome characteristics of a generalist marine bacterial lineage. *ISME J.* 2010;4:784–98.
- Moran MA, Gonzalez JM, Kiene RP. Linking a bacterial taxon to sulfur cycling in the sea: Studies of the marine *Roseobacter* group. *Geomicrobiol J.* 2003;20:375–88.

28. Potter LC, Cole JA. Essential roles for the products of the napABCD genes, but not napFGH, in periplasmic nitrate reduction by *Escherichia coli* K-12. *Biochem J.* 1999;344:69–76.
29. Stewart V, Lu Y, Darwin AJ. Periplasmic nitrate reductase (NapABC enzyme) supports anaerobic respiration by *Escherichia coli* K-12. *J Bacteriol.* 2002;184:1314–23.
30. Breider S, Teshima H, Petersen J, Chertkov O, Dalingault H, Chen A, et al. Genome sequence and emended description of *Leisingera nanhaiensis* strain DSM 24252(T) isolated from marine sediment. *Stand Genomic Sci.* 2014;9:687–703.
31. Dogs M, Voget S, Teshima H, Petersen J, Davenport K, Dalingault H, et al. Genome sequence of *Phaeobacter inhibens* type strain (T5(T)), a secondary metabolite producing representative of the marine *Roseobacter* clade, and emendation of the species description of *Phaeobacter inhibens*. *Stand Genomic Sci.* 2013;9:334–50.
32. Poggio S, Abreu-Goodger C, Fabela S, Osorio A, Dreyfus G, Vinuesa P, et al. A complete set of flagellar genes acquired by horizontal transfer coexists with the endogenous flagellar system in *Rhodobacter sphaeroides*. *J Bacteriol.* 2007;189:3208–16.
33. Frank O, Göker M, Pradella S, Petersen J. Ocean's twelve: Flagellar and biofilm chromids in the multipartite genome of *Marinovum algicola* DG898 exemplify functional compartmentalization. *Environ Microbiol.* 2015;17:4019–34.
34. Ramboarina S, Fernandes PJ, Daniell S, Islam S, Simpson P, Frankel G, et al. Structure of the bundle-forming pilus from enteropathogenic *Escherichia coli*. *J Biol Chem.* 2005;280:40252–60.
35. Craig L, Pique ME, Tainer JA. Type IV pilus structure and bacterial pathogenicity. *Nat Rev Micro.* 2004;2:363–78.
36. Waters CM, Bassler BL. Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol.* 2005;21:319–46.
37. An JH, Goo E, Kim H, Seo YS, Hwang I. Bacterial quorum sensing and metabolic slowing in a cooperative population. *Proc Natl Acad Sci U S A.* 2014;111:14912–7.
38. Park SY, Kang HO, Jang HS, Lee JK, Koo BT, Yum DY. Identification of extracellular N-acylhomoserine lactone acylase from a *Streptomyces* sp. and its application to quorum quenching. *Appl Environ Microbiol.* 2005;71:2632–41.
39. Lin YH, Xu JL, Hu J, Wang LH, Ong SL, Leadbetter JR, et al. Acyl-homoserine lactone acylase from *Ralstonia* strain XJ12B represents a novel and potent class of quorum-quenching enzymes. *Mol Microbiol.* 2003;47:849–60.
40. Huang JJ, Petersen A, Whiteley M, Leadbetter JR. Identification of QuiP, the product of gene PA1032, as the second acyl-homoserine lactone acylase of *Pseudomonas aeruginosa* PAO1. *Appl Environ Microbiol.* 2006;72:1190–7.
41. Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 2015;43:W237–W43.
42. Kyrpides NC, Woyke T, Eisen JA, Garrity G, Lilburn TG, Beck BJ, et al. Genomic encyclopedia of type strains, phase I: the one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci.* 2014;9:1278–84.
43. Kyrpides NC, Hugenholtz P, Eisen JA, Woyke T, Göker M, Parker CT, et al. Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol.* 2014;12:e1001920.
44. Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci.* 2010;2:117–34.
45. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *Bmc Bioinformatics.* 2013;14. doi:10.1186/1471-2105-14-60.
46. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol.* 2008;26:541–7.
47. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya. *Proc Natl Acad Sci U S A.* 1990;87:4576–9.
48. Garrity GM BJ, Lilburn T, Phylum XIV. In: DJ Brenner NK, Staley JT, Garrity GM, editors. *Proteobacteria* phyl. nov. second edition, part B (The *Gammaproteobacteria*). New York: Springer; 2005.
49. Garrity GMBJ, Lilburn T, Class I. In: Garrity GM BD, Krieg NR, Staley JT, editors. *Alphaproteobacteria* class. nov. Secondth ed. New York: Springer; 2005.
50. List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol.* 2006;56:1–6.
51. Garrity GM BJ, Lilburn T, Family I. In: Garrity GM BD, Krieg NR, Staley JT, editors. *Rhodobacteraceae* fam. nov. Secondth ed. New York: Springer; 2005.
52. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium Nat Genet. 2000;25:25–9.
53. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics.* 2011;12:402.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

