The ISME Journal (2008) 2, 656–662

www.nature.com/ismej

ORIGINAL ARTICLE

Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems

Nof Atamna-Ismaeel[1], Gazalah Sabehi[1], Itai Sharon[1,2], Karl-Paul Witzel[3], Matthias Labrenz[4], Klaus Jürgens[4], Tamar Barkay[5], Maayke Stomp[6,7], Jef Huisman[6] and Oded Beja[1]
[1]*Faculty of Biology, Technion-Israel Institute of Technology, Haifa, Israel;* [2]*Computer Science Department, Technion-Israel Institute of Technology, Haifa, Israel;* [3]*Max-Planck-Institut für Limnologie, Plön, Germany;* [4]*Department of Biological Oceanography, Leibniz Institute for Baltic Sea Research, Rostock, Germany;* [5]*Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, USA and* [6]*Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*

Proteorhodopsins (PRs) are light-driven proton pumps that have been found in a variety of marine environments. The goal of this study was to search for PR presence in different freshwater and brackish environments and to explore the diversity of non-marine PR protein. Here, we show that PRs exist in distinctly different aquatic environments, ranging from clear water lakes to peat lakes and in the Baltic Sea. Some of the PRs observed in this study formed unique clades that were not previously observed in marine environments, whereas others were similar to PRs found in non-marine samples of the Global Ocean Sampling (GOS) expedition. Furthermore, the similarity of several PRs isolated from lakes in different parts of the world suggests that these genes are dispersed globally and that they may encode unique functional capabilities enabling successful competition in a wide range of freshwater environments. Phylogenomic analysis of genes found on these GOS scaffolds suggests that some of the freshwater PRs are found in freshwater *Flavobacteria* and freshwater SAR11-like bacteria.
The ISME Journal (2008) **2,** 656–662; doi:10.1038/ismej.2008.27; published online 27 March 2008
Subject Category: microbial ecology and functional diversity of natural habitats
Keywords: microbial biodiversity; phototrophic growth; phylogenetic tree; proteobacteria; rhodopsins; spectral tuning

Introduction

Proteorhodopsins (PRs) (Béjà *et al.*, 2000, 2001) are microbial retinal-binding membrane pigments belonging to the microbial rhodopsin superfamily (Spudich *et al.*, 2000) and were suggested to have an important role in supplying light energy for microbial metabolism in different marine ecosystems (Béjà *et al.*, 2000, 2001; Sabehi *et al.*, 2005b; Gómez-Consarnau *et al.*, 2007; Martinez *et al.*, 2007; Walter *et al.*, 2007). PRs have been observed in a wide variety of ocean regions (Béjà *et al.*, 2000, 2001; de la Torre *et al.*, 2003; Sabehi *et al.*, 2003, 2004, 2005a, 2007; Venter *et al.*, 2004; Rusch *et al.*, 2007; Campbell *et al.*, 2008) and were found in diverse bacterial groups, including the ubiquitous marine gammaproteobacterial SAR86 (Béjà *et al.*, 2000; Sabehi *et al.*, 2004, 2005a) and alphaproteobacterial SAR11 groups (Giovannoni *et al.*, 2005a, b; Sabehi *et al.*, 2005a), as well as in *Roseobacter* (alphaproteobacterium HTCC2255), marine *Bacteroidetes* (Venter *et al.*, 2004; Gómez-Consarnau *et al.*, 2007), planktonic *Archaea* (Frigaard *et al.*, 2006) and other microbial taxa (de la Torre *et al.*, 2003; Sabehi *et al.*, 2003, 2005a; Venter *et al.*, 2004; McCarren and DeLong, 2007; Rusch *et al.*, 2007). While previous work did not detect light enhanced growth neither in PR-containing SAR11 (Giovannoni *et al.*, 2005a) nor in the gammaproteobacterial SAR92 (Stingl *et al.*, 2007) isolates grown in seawater, significant enhancement of both growth rate and yield was recently reported in PR-expressing marine *Flavobacteria* (Gómez-Consarnau *et al.*, 2007).

PRs were so far detected in different marine environments, mainly in the open ocean. In the recent Global Ocean Sampling (GOS) expedition (Rusch *et al.*, 2007), 44 samples were obtained, covering a wide range of distinct surface marine environments as well as a few estuarine and inland

Correspondence: O Beja, Faculty of Biology, Technion-Israel Institute of Technology, Haifa 32000, Israel.
E-mail: beja@tx.technion.ac.il
[7]Current address: WK Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060, USA
Received 23 December 2007; revised 18 February 2008; accepted 19 February 2008; published online 27 March 2008

waters. These included three estuaries (Bay of Fundy, Canada; Delaware Bay and Chesapeake Bay, USA), one freshwater lake (Lake Gatun, Panama) and a moderate hypersaline lagoon (Punta Cormorant Lagoon, Galápagos, Ecuador). Different PRs were observed in these estuaries and non-marine stations (Rusch *et al.*, 2007).

Freshwater ecosystems offer a wide range of variation in many ecologically relevant parameters, including the underwater light climate. Known PR genes encode pigments with two distinct light absorption spectra, commonly referred to as 'blue-absorbing' PR and 'green-absorbing' PR (Béjà *et al.*, 2001; Man *et al.*, 2003). Blue-absorbing PRs are particularly widespread in the blue ocean waters that have been extensively sampled in earlier PR studies (Béjà *et al.*, 2001; Sabehi *et al.*, 2007). The underwater light color of most freshwater ecosystems, however, is shifted towards the green and red part of the light spectrum (Kirk, 1994; Stomp *et al.*, 2007a, b). Thus, possibly, new PR genes encoding green-absorbing and perhaps even red-absorbing pigments might be awaiting discovery in freshwater ecosystems. Motivated by these ideas, we have designed general degenerate PR primers based on data from the non-marine GOS stations and searched for PR presence in different freshwater systems worldwide.

Materials and methods

Sample collection, DNA preparation and storage
Environmental samples were collected from Lake Kinneret in Israel, Lakes Großer Plöner See, Schöhsee, Plußsee in Germany, Lake L658 in Canada, Lake Zegerplas and Lake Westeinderplassen in the Netherlands and from the Baltic Sea (see Table 1 for details of sampling locations). Genomic DNA extractions were performed using the UltraClean Water DNA kit (MoBio Laboratories, Carlsbad, CA, USA). DNA was stored at −80 °C for further analysis.

PR PCR amplification
Proteorhodopsins were amplified from DNA extracts obtained from environmental samples using modified PR forward and reverse primers (see Table 2). The degenerate primers were designed based on known PR sequences found in NCBI and on estuary and freshwater stations in the GOS dataset (GOS stations GS06, 11, 12 and 20). A total of 48 combinations of primer sets were designed and tested against the Lake Kinneret samples. Of all the 48 combinations that were tested, only 6 combinations produced a PCR product and these 6 combinations were selected for further analysis on environmental samples. Some deep-branching actinorhodopsin-like (Sharma *et al.*, 2008) sequence variants contain the sequence Twxxyp (for example, actinorhodopsin reads 1095521408126, 1095390008418 and 1095898194153 from groups LG1 and LG2) and would, therefore, not be amplified by our reverse primer sets, which are based on a Gwxxyp protein sequence.

Polymerase chain reactions were performed using high-fidelity proof-reading polymerase mix (TaKaRa Ex Taq from Takara Shuzo). PCR amplification was carried out in a total volume of 25 μl containing 10 ng of template DNA, 200 μM dNTPs, 1.5 mM $MgCl_2$, 0.2 μM primers and 2.5 U TaKaRa Ex Taq polymerase. The amplification conditions comprised steps at 92 °C for 4 min, 35 cycles at 92 °C for 1 min, 49.8 °C for 1 min and 72 °C TaKaRa Ex Taq polymerase for 1 min. PCR products were cloned using the QIAGEN-PCR cloning kit (Qiagen, Hilden, Germany). Unique *Eco*RI and *Rsa*I restriction fragment length polymorphism groups were sequenced. All PCR products identical to samples handled or amplified previously in the lab were omitted from the analyses to avoid being influenced by possible contamination.

PR phylogeny
The PR tree was constructed according to Sabehi *et al.* (2005a). PR proteins were aligned using CLUSTALx (Thompson *et al.*, 1997) and 110 amino-acids alignment was used to construct a neighbor-joining phylogenetic tree as inferred using the neighbor programs of PAUP* (Swofford, 2002). Bootstrap resampling (1000) of the distance and parsimony trees were performed in all analyses to provide confidence estimates for the inferred topologies.

Nucleotide sequence accession numbers
PR gene sequences were deposited in GenBank under accession numbers EU563268–EU563333.

Table 1 Characteristics of sampling sites

Sampling location	*Coordinates*	*Sampling dates*	*Sampling depth (m)*	*Trophic status*
Lake Kinneret, Israel	32°5′N, 35°4′E	January, May and July 2007	2, 6, 12	Meso-eutrophic
Lake L658, Canada	49°8′N, 93°8′W	August, 2002	3	Oligotrophic
Großer Plöner See, Germany	54°09′N, 10°25′E	July, 2006	1	Slightly eutrophic
Schöhsee, Germany	54°09′N, 10°27′E	July, 2006	1	Meso-eutrophic
Plußsee, Germany	54°11′N, 10°26′E	July, 2006	1	Eutrophic
Zegerplas, Netherlands	52°08′N, 4°41′E	September, 2007	0, 3	Eutrophic
Westeinderplassen, The Netherlands	52°14′N, 4°42′E	September, 2007	0, 3	Eutrophic
Baltic Sea, Scandinavia	57°19′N, 20°03′E	July, 2006	5, 15, 35, 50	Meso-eutrophic

Table 2 The proteorhodopsin forward and reverse primers used in this study, designed based on known PR sequences found in the NCBI and Global Ocean Sampling datasets

Forward	*Reverse*
RY**I**DW (5′-MGNTAYATH GAYTGG-3′)	GW**AI**YP (5′-GGNTGGGCNATH TAYCCN-3′)
RY**A**DW (5′-MGNTAYGC NGAYTGG-3′)	GW**LI**YP (5′-GGNTGGMTNATH TAYCCN-3′)
RY**F**DW (5′-MGNTAYTT YGAYTGG-3′)	GW**VI**YP (5′-GGNTGGGTNATH TAYCCN-3′)
RY**V**DW (5′-MGNTAYGT NGAYTGG-3′)	GW**II**YP (5′-GGNTGGATHATHT AYCCN-3′)
	GW**SI**YP (5′-GGNTGGWSNATH TAYCCN-3′)
	GW**GT**YP (5′-GGNTGGGGNATH TAYCCN-3′)
	GW**AV**YP (5′-GGNTGGGCNGTN TAYCCN-3′)
	GW**LV**YP (5′-GGNTGGMTNGTN TAYCCN-3′)
	GW**VV**YP (5′-GGNTGGGTNGTN TAYCCN-3′)
	GW**GV**YP (5′-GGNTGGGGNGTN TAYCCN-3′)
	GW**SV**YP (5′-GGNTGGWSNGTN TAYCCN-3′)
	GW**IV**YP (5′-GGNTGGATHGTN TAYCCN-3′)

Variable amino acids in the translated primer regions are marked in bold letters.

Table 3 Pairs of primers that gave positive results in the different sites

Sampling location	*Primers*
Lake Kinneret, Israel	RYVDW+GWVIYP (6) RYIDW+GWAIYP (17) RYIDW+GWSIYP (1) RYIDW+GWAVYP (6) RYVDW+GWSIYP (3) RYVDW+GWAIYP (1)
Lake L658, Canada	RYIDW+GWAIYP (2) RYIDW+GWAVYP (7) RYVDW+GWAVYP (1)
Großer Plöner See, Germany	RYIDW+GWAIYP (1) RYIDW+GWAVYP (2) RYVDW+GWSIYP (1) RYVDW+GWVIYP (1)
Schöhsee, Germany	RYIDW+GWAIYP (2)
Plußsee, Germany	RYIDW+GWAIYP (1) RYIDW+GWSIYP (1)
Zegerplas, Netherlands	RYIDW+GWAIYP (1)
Westeinderplassen, Netherlands	RYIDW+GWAIYP (2)
Baltic Sea, Scandinavia	RYIDW+GWAIYP (3) RYIDW+GWAVYP (5)

The number of clones obtained with each pair is shown in parenthesis next to each pair.

Results and discussion

Design of new PR primers

Proteorhodopsin sequences retrieved from different estuaries and freshwater GOS stations (Stations GS06, 11, 12 and 20) were aligned and new primers were designed to conserved regions in the PR protein (amino-acid positions 94–98 and 196–201 based on PR BAC31A8 numbering (Béjà *et al.*, 2001)) to cover possible new PR diversity in brackish and freshwater ecosystems. The same PR regions were previously used to recover PRs from different marine environments and from BAC libraries (Sabehi *et al.*, 2005b, 2007; Campbell *et al.*, 2008). Four forward primers and twelve reverse primers (Table 2) were used to perform multiplex PCR reactions on Lake Kinneret samples. Six out of the possible forty-eight combinations gave positive PCR products (combinations RYIDW + GWAIYP, RYIDW + GWSIYP, RYIDW + GWAVYP, RYVDW + GWSIYP, RYVDW + GWVIYP and RYVDW + GWAVYP; see Table 3) and were later used with other samples.

A recent *in silico* search (Sharma *et al.*, 2008) for rhodopsins performed on several aquatic non-marine GOS stations has identified unique groups of deep-branching PR-like sequences linked to *Actinobacteria* (coined 'actinorhodopsins' by the authors). It is important to note that the primers used in this study would not amplify some actinorhodopsin-like sequence variants (see Materials and methods section).

Screening of different freshwater lakes and the Baltic Sea

To explore the possible existence of PRs in Lake Kinneret, Israel, we screened different season and depth samples with the 48 different primer combinations. Out of these 48 primer combinations, only 6 gave positive PCR signals. The PRs amplified from Lake Kinneret differed from known marine PRs and clustered separately or with other non-marine GOS PRs (Figure 1). Using these six primer sets we have amplified PR genes from DNA extracts from clear lakes in Germany (Lakes Großer Plöner See, Schöhsee, Plußsee) and Canada (Lake L658), peat lakes from the Netherlands (Zegerplas and Westeinderplassen), as well as from the Baltic Sea.

In total, we found a tremendous diversity of freshwater PRs that show as low as 42% identity (based on amino-acid level) between different distanced clades. Several PR groups were unique to Lake Kinneret and were not amplified in other freshwater environments (freshwater clades 2, 3 and 4 in Figure 1), whereas other groups contained PR representatives from different environments (non-marine clades 1 and 6 as well as freshwater clades 7, 8 and 9 in Figure 1). Several PR clades clustered in proximity to PRs from cultured marine *Flavobacteria* and possibly represent PRs from freshwater *Flavobacteria*. Another two PR clades (6 and 7) clustered with representatives from the marine

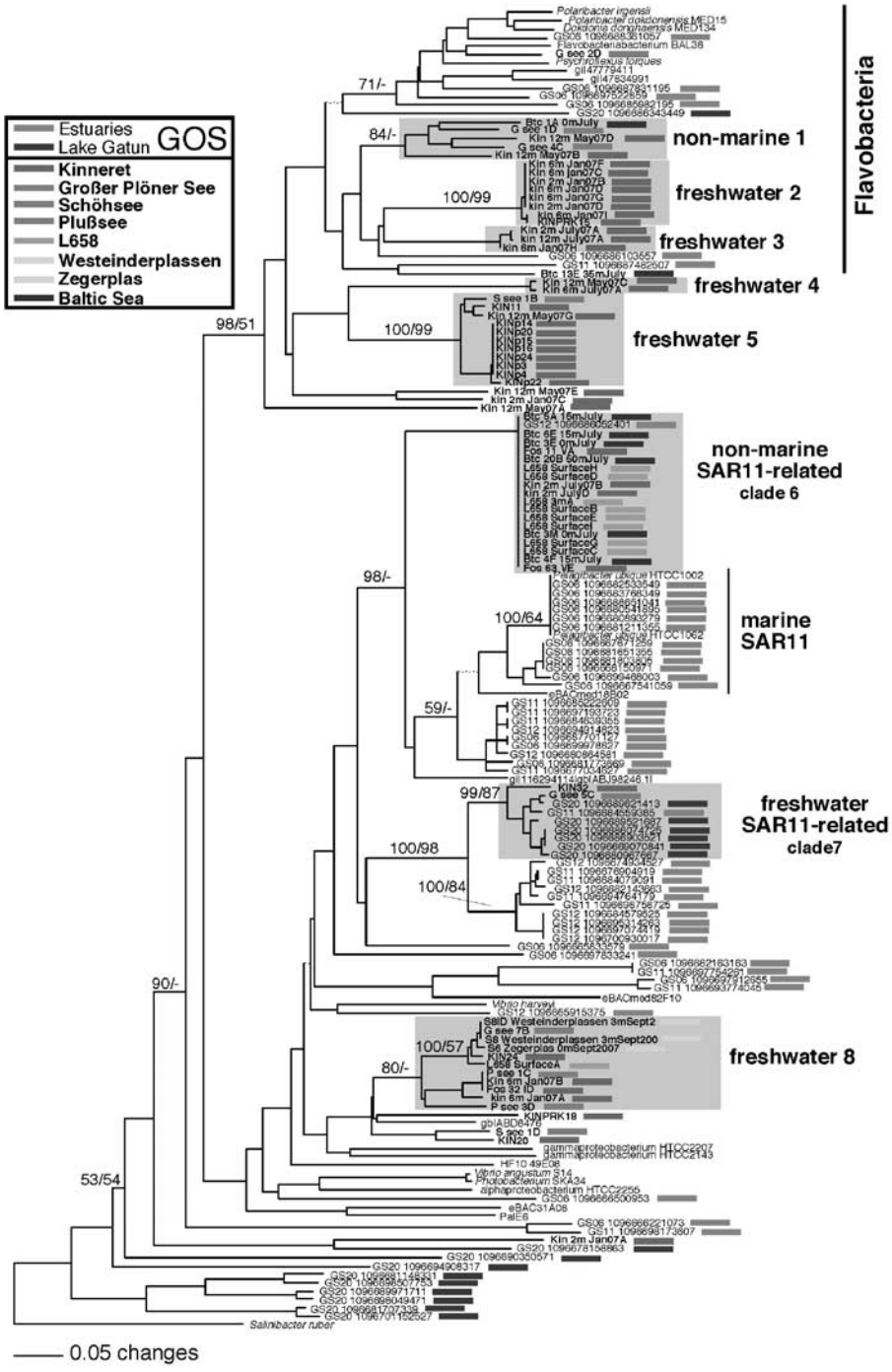

Figure 1 Neighbor-joining phylogenetic tree of aquatic non-marine proteorhodopsins (PR) proteins (102 amino acids position alignment length; without the corresponding primer regions). Estuarine global ocean-sampling (GOS) sequences are represented by green bands while freshwater GOS sequences are marked with blue. PR sequences retrieved via PCR in this study are marked by boldface letters and their station position color is indicated in the box on the top left. Bootstrap values (distance/parsimony) greater than 50% of major clades are indicated above the branches. For color figure, see online version.

SAR11 clade. Freshwater PR clades 8 and 9 did not cluster with any known PR and their affiliations remain unknown.

All PCR products identical to samples handled or amplified previously in the lab (based on the DNA level) were omitted from the analyses to avoid influence from possible contamination. This step is extremely important when different lake samples are compared at once. In Figures 2a and b, alignments of different PRs found in different lake environments and in the Baltic Sea are shown. As can be seen in Figure 2a, non-marine clade 1 (top Figure 1 above freshwater clade 2), which contained PRs from different samples is composed of very

Figure 2 Multiple alignment of proteorhodopsins (PR) amino-acid sequences. (**a**) Non-marine clade 1; (**b**) freshwater clade 8. The phylogenetic tree shown on the left-hand side corresponds to the tree in Figure 1. Multiple alignments of PR amino-acid sequences are shown. Position 105 is marked with a red box. For color figure, see online version.

different PR proteins, which could not have been amplified from possible lab contaminations. Freshwater clade 8 includes PRs from a wide range of different and geographically distant lakes (Lake Kinneret in Israel, several clear-water lakes in Germany and Canada as well as peat lakes in the Netherlands). In contrast to non-marine clade 1, the PRs in freshwater clade 8 were closely related (Figure 2b). However, there are amino-acid differences between the PRs in this clade. This echoes similar findings of Zwart *et al.* (1998), describing 16S rRNA clades consisting of highly related proteobacteria isolated from different continents. The similarity of PRs isolated from lakes in different parts of the world suggests that these genes are dispersed globally and that they may encode unique functional capabilities enabling successful competition in a wide range of freshwater environments.

Organismal origins of PRs from freshwater lakes and the Baltic Sea

Similar to taxon prediction performed in previous studies (de la Torre *et al.*, 2003; McCarren and DeLong, 2007; Sharma *et al.*, 2008), we tried to identify the organisms containing unique non-marine PRs by carrying out a phylogenomic analysis of adjacent sequences found on similar freshwater and estuary GOS scaffolds. Using this approach, we assigned the affiliation of clades 1, 2 and 3 with *Flavobacteria* (based on GOS sequences GS06 1096686103557 and GS11 1096687482507, which correspond to GOS scaffold-1096626019272 and scaffold-1096627448042, respectively). We speculate that the same is true for clades 4 and 5 but with lower confidence.

Clades 6 and 7 form distinct clades closely related to *Pelagibacter ubique* and the SAR11-type rhodopsins. Different GOS scaffolds (GS12 singleton-1096686052401, scaffold-1096627380104) that clustered within clade 6 and 7 were highly similar to *P. ubique* based on gene content but the genomic organization was different with an operon for carotene biosynthesis adjacent to the PR gene. Similar operon organization is observed in many uncultured PR-carrying organisms (Sabehi *et al.*, 2005b, 2007) but is different from the organization in *P. ubique*, where an acetyl-CoA synthetase and a ferredoxin are adjacent to the PR gene and the retinal biosynthesis operon is found elsewhere in the genome (Giovannoni *et al.*, 2005a, b). We, therefore, assigned clades 6 and 7 as SAR11-related clades (Figure 1). Non-marine SAR11s (alphaproteobacterial cluster LD12) were first reported in an arctic lake (Bahr *et al.*, 1996) and are readily detected in other freshwater systems (Zwart *et al.*, 1998, 2003; Crump *et al.*, 1999). PRs from clades 6 and 7 could be from bacteria related to alphaproteobacterial cluster LD12. However, because of previously reported PR lateral gene transfer (Frigaard *et al.*, 2006), our predicted affiliations should be dealt with caution and are provisional.

Spectral tuning in estuaries and freshwater PRs

All predicted PR proteins retrieved in this study contained leucine (L) or methionine (M) in position 105 (cf. Figure 2). This makes them suspected green-absorbing PRs (Man *et al.*, 2003). The predominance of green-absorbing PRs can be explained by the underwater light spectrum in coastal and freshwater ecosystems. In these waters, green-light conditions prevail (Stomp *et al.*, 2007b; Haverkamp *et al.*, 2008) and hence green-absorbing PRs will have a selective advantage compared to blue-absorbing PRs. Yet, the PRs described here show more variability in their protein sequence than in previous studies (Man *et al.*, 2003) and are largely distinct from the marine clades (Figure 1). Therefore, prediction of the color of light absorbed by these freshwater PRs based on a single position (such as position 105) in the overall sequence should be made with great caution.

One remaining open question is the role of the PR-like molecules. Several of the PRs reported here are phylogenetically quite distant from 'known' PRs (as low as 42% amino-acid identity). Some of these PRs might indeed function as proton pumps involved in phototrophy but others might be sensory rhodopsins in the classical sense (reacting with transducers—but may appear to be a 'proton pump' by sequence) or might have other unknown functions. Our work, along with the recent GOS observations (Rusch *et al.*, 2007; Sharma *et al.*, 2008), suggests that PR and PR-like pigments are common in different

aquatic environments worldwide, including marine, brackish and freshwater ecosystems. Thus, PR-based phototrophy might be a common theme in different aquatic environments worldwide.

Acknowledgements

We thank Mert Kumro and Mira Rosenberg for technical assistance. We also thank the students of the MSc program Limnology and Oceanography 2007 of the University of Amsterdam for their help in the Dutch lakes. This work was supported by Grant 1203/06 from the Israel Science Foundation and an EMBO YIP award to OB, by a Pioneer grant from the Earth and Life Sciences Foundation (ALW), which is subsidized by the Netherlands Organization for Scientific Research and by a US NSF grant # AER-9910268.

References

Bahr M, Hobbie JE, Sogin ML. (1996). Bacterial diversity in an arctic lake: a freshwater SAR11 cluster. *Aquat Microb Ecol* **11**: 271–277.

Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP *et al.* (2000). Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–1906.

Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. (2001). Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786–789.

Campbell BJ, Waidner LA, Cottrell MT, Kirchman DL. (2008). Abundant proteorhodopsin genes in the North Atlantic Ocean. *Environ Microbiol* **10**: 99–109.

Crump BC, Armbrust EV, Baross JA. (1999). Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl Environ Microbiol* **65**: 3192–3204.

de la Torre JR, Christianson L, Béjà O, Suzuki MT, Karl D, Heidelberg JF *et al.* (2003). Proteorhodopsin genes are widely distributed among divergent bacterial taxa. *Proc Natl Acad Sci USA* **100**: 12830–12835.

Frigaard N-U, Martinez A, Mincer TJ, DeLong EF. (2006). Proteorhodopsin lateral gene transfer between marine planktonic bacteria and archaea. *Nature* **439**: 847–850.

Giovannoni SJ, Bibbs L, Cho J-C, Stapels MD, Desiderio R, Vergin KL *et al.* (2005a). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438**: 82–85.

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D *et al.* (2005b). Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–1245.

Gómez-Consarnau L, González JM, Coll-Lladó M, Gourdon P, Pascher T, Neutze R *et al.* (2007). Light stimulates growth of proteorhodopsin-containing marine Flavobacteria. *Nature* **445**: 210–213.

Haverkamp T, Acinas SG, Doeleman M, Stomp M, Huisman J, Stal LJ. (2008). Diversity and phylogeny of Baltic Sea picocyanobacteria inferred from their ITS and phycobiliprotein operons. *Environ Microbiol* **10**: 174–188.

Kirk JTO. (1994). *Light and Photosynthesis in Aquatic Ecosystems*. Cambridge University: Cambridge, UK.

Man D, Wang W, Sabehi G, Aravind L, Post AF, Massana R *et al.* (2003). Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**: 1725–1731.

Martinez A, Bradley AS, Waldbauer J, Summons RE, DeLong EF. (2007). Proteorhodopsin photosystem gene expression enables photophosphorylation in heterologous host. *Proc Natl Acad Sci USA* **104**: 5590–5595.

McCarren J, DeLong EF. (2007). Proteorhodopsin photosystem gene clusters exhibit co-evolutionary trends and shared ancestry among diverse marine microbial phyla. *Environ Microbiol* **9**: 846–858.

Rusch DB, Halpern AL, Heidelberg KB, Sutton G, Williamson SJ, Yooseph S *et al.* (2007). The sorcerer II global ocean sampling expedition: I, the northwest Atlantic through the eastern tropical Pacific. *PLoS Biol* **5**: e77.

Sabehi G, Béjà O, Suzuki MT, Preston CM, DeLong EF. (2004). Different SAR86 subgroups harbour divergent proteorhodopsins. *Environ Microbiol* **6**: 903–910.

Sabehi G, Kirkup BC, Rosenberg M, Stambler N, Polz MF, Béjà O. (2007). Adaptation and spectral tuning in divergent marine proteorhodopsins from the eastern Mediterranean and the Sargasso Seas. *ISME J* **1**: 48–55.

Sabehi G, Massana R, Bielawski JP, Rosenberg M, DeLong EF, Béjà O. (2003). Novel proteorhodopsin variants from the Mediterranean and Red Seas. *Environ Microbiol* **5**: 842–849.

Sabehi G, Loy A, Jung KH, Partha R, Spudich JL, Isaacson T *et al.* (2005a). New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *PLoS Biol* **3**: e273.

Sabehi G, Loy A, Jung KH, Partha R, Spudich JL, Isaacson T *et al.* (2005b). New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *PLoS Biol* **3**: e173.

Sharma AK, Zhaxybayeva O, Papke RT, Doolittle WF. (2008). Actinorhodopsins: proteorhodopsin-like gene sequences found predominantly in non-marine environments. *Environ Microbiol* **10**: 1039–1056.

Spudich JL, Yang CS, Jung KH, Spudich EN. (2000). Retinylidene proteins: structures and functions from archaea to humans. *Annu Rev Cell Dev Biol* **16**: 365–392.

Stingl U, Desiderio RA, Cho J-C, Vergin KL, Giovannoni SJ. (2007). The SAR92 clade: an abundant coastal clade of culturable marine bacteria possessing proteorhodopsin. *Appl Environ Microbiol* **73**: 2290–2296.

Stomp M, Huisman J, Stal LJ, Matthijs HCP. (2007a). Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. *ISME J* **1**: 271–282.

Stomp M, Huisman J, Vörös L, Pick FR, Laamanen M, Haverkamp T. *et al.* (2007b). Colourful coexistence of red and green picocyanobacteria in lakes and seas. *Ecol Lett* **10**: 290–298.

Swofford DL. (2002). PAUP*. Phylogenetic analysis using parsimony (*And Other Methods). Version 4.0b10. Sinauer Associates: Sunderland, Massachusetts.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882.

Venter JC, Remington K, Heidelberg J, Halpern AL, Rusch D, Eisen J *et al.* (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Walter JM, Greenfield D, Bustamante C, Liphardt J. (2007). Light-powering *Escherichia coli* with proteorhodopsin. *Proc Natl Acad Sci USA* **104**: 2408–2412.

Zwart G, Hiorns WD, Methé BA, Van Agterveld MP, Huismans R, Nold SC *et al.* (1998). Nearly identical 16S rRNA sequences recovered from lakes in North America and Europe indicate the existence of clades of globally distributed freshwater bacteria. *Syst Appl Microbiol* **21**: 546–556.

Zwart G, van Hannen EJ, Kamst-van Agterveld MP, Van der Gucht K, Lindström ES, Van Wichelen J *et al.* (2003). Rapid screening for freshwater bacterial groups by using reverse line blot hybridization. *Appl Environ Microbiol* **69**: 5875–5883.