Microbial rhodopsins on leaf surfaces of terrestrial plants

Nof Atamna-Ismaeel[1,†], Omri M. Finkel[2,†], Fabian Glaser[3], Itai Sharon[1,4,10], Ron Schneider[2], Anton F. Post[5], John L. Spudich[6], Christian von Mering[7], Julia A. Vorholt[8], David Iluz[9], Oded Béjà[1,*], and Shimshon Belkin[2,*]

[1]*Faculty of Biology,* [3]*Bioinformatics Knowledge Unit, Lorry I. Lokey Interdisciplinary Center for Life Sciences and Engineering,* [4]*Faculty of Computer Science, Technion – Israel Institute of Technology, Haifa 32000, Israel.*

[2]*Department of Plant and Environmental Sciences, Alexander Silberman Institute of Life Sciences, Hebrew University of Jerusalem, Jerusalem 91904, Israel.*

[5]*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biology Laboratory, Woods Hole, MA 02543, USA.*

[6]*Center for Membrane Biology, Department of Biochemistry and Molecular Biology, The University of Texas Medical School, Houston, TX 77030, USA.*

[7]*Faculty of Science, Institute of Molecular Life Sciences and Swiss Institute of Bioinformatics, University of Zurich, 8057 Zurich, Switzerland.*

[8]*Institute of Microbiology, Eidgenössische Technische Hochschule Zurich, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland.*

[9]*Faculty of Life Sciences, Bar-Ilan University, Ramat Gan 52900, Israel*

[†]*N.A.-I. and O.M.F. contributed equally to this work*

[*]*To whom correspondance should be addressed. E-mails: (@tx.technion.ac.il) or S.B. (@vms.huji.ac.il).*

[10]*Present address: Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA*

Summary

The above-ground surfaces of terrestrial plants, the phyllosphere, comprise the main interface between the terrestrial biosphere and solar radiation. It is estimated to host up to 10^{26} microbial cells that may intercept part of the photon flux impinging on the leaves. Based on 454-pyrosequencing generated metagenome data, we report on the existence of diverse microbial rhodopsins in five distinct phyllospheres from tamarisk (*Tamarix nilotica*), soybean (*Glycine max*), *Arabidopsis* (*Arabidopsis thaliana*), clover (*Trifolium repens*) and rice (*Oryza sativa*). Our findings, for the first time describing microbial rhodopsins from non-aquatic habitats, point toward the potential coexistence of microbial rhodopsin-based phototrophy and plant chlorophyll-based photosynthesis, with the different pigments absorbing non-overlapping fractions of the light spectrum.

Introduction

Solar radiation is the main source of energy for both marine and terrestrial organisms, with terrestrial plants and aquatic phytoplankton performing an equivalent ecological function as chlorophyll-based photosynthetic primary producers (Field et al., 1998). Marine surface waters are now known to harbour an additional type of phototrophy; several lineages of bacteria and archaea utilize rhodopsins (Béjà et al., 2000; Béjà et al., 2001; de la Torre et al., 2003;

Balashov et al., 2005; Giovannoni et al., 2005; Sabehi et al., 2005; Frigaard et al., 2006; Gómez-Consarnau et al., 2007; Gómez-Consarnau et al., 2010; Oh et al., 2010), retinal-containing trans-membrane proteins, as light-driven proton pumps. The first microbial rhodopsin was discovered nearly four decades ago in the archaeon *Halobacterium salinarum* from hypersaline environments (Oesterhelt and Stoeckenius, 1971). Further studies revealed the existence of microbial rhodopsins in diverse habitats including freshwater, sea ice, hypersaline and brackish environments (Rusch et al., 2007; Atamna-Ismaeel et al., 2008; Sharma et al., 2008; Sharma et al., 2009; Koh et al., 2010). To date, microbial rhodopsins have been reported exclusively for aquatic habitats.

As light is an abundant resource on land, we tested the hypothesis that microbial rhodopsins also exist and play an important role in terrestrial niches. The leaf surface of terrestrial plants covers a surface area of an estimated $6.4X10^8$ km^2 and comprises the main interface between terrestrial biomass and solar photon flux. This habitat harbors an immensely diverse microbial community of up to 10^6 - 10^7 cells per cm^2 leaf surface (Lindow and Brandl, 2003). A mode of phototrophy that is compatible with the plant's photosynthesis would offer a significant ecological advantage to microbes inhabiting this environment.

Results and Discussion

We have identified 156 microbial rhodopsin sequences in five phyllosphere metagenomes [files S1, S2, S3, S4, and S5], from different terrestrial plants, soybean (*Glycine max*) (Delmotte et al., 2009), tamarisk

(*Tamarix nilotica*), clover (*Trifolium repens*), rice (*Oryza sativa*) as well as from a wild population of the model plant *Arabidopsis thaliana*. The size of the different metagenomes obtained was 261 Mb, 448 Mb, 234 Mb, 831 Mb, and 250 Mb for soybean, tamarisk, clover, rice and *Arabidopsis* with an average read length of 235, 328, 235, 357 and 233 bp, respectively.

Phylogenetic analysis revealed that some phyllosphere microbial rhodopsins have branched away from known rhodopsin families within the bacterial and eukaryal domains (Fig. 1). Some of these sequences clustered with fungal rhodopsins, while another group clustered with xanthorhodopsins (Balashov et al., 2005; Lanyi and Balashov, 2008) and actinorhodopsins (Sharma et al., 2008; Sharma et al., 2009). However, most phyllosphere rhodopsins appear on novel branches, with no representatives from either culture-based or environmental datasets, thus rendering them with an as yet uncertain phylogenetic affiliation. In most cases, the leaf surface rhodopsins from tamarisk clustered separately from other phyllosphere rhodopsins (Fig.1) with a statistically significant phylogenetic signal [calculated using Mesquite (Maddison and Maddison, 2010)], indicating that they reside in distinct microbial taxa, probably adapted to the unique hypersaline environment of the tamarisk phyllosphere (Qvit-Raz et al., 2008).

In contrast with soil metagenomes, which do not contain any rhodopsin reads, the five phyllosphere datasets were found to contain microbial rhodopsins, however, at frequencies lower than those found in marine and freshwater metagenomes (Fig. 2). While some of the phyllosphere rhodopsins lack the retinylidene Schiff base proton donor carboxylate and are thus likely sensory

rhodopsins, others contain both proton acceptor and donor carboxylates at helix C (bacteriorhodopsin positions 85 and 96, respectively; see files S1, S2, S3, S4, and S5) and may be considered as potential proton pumps. Compared to the marine environment, where they make up only 3% of all microbial rhodopsins (Spudich, 2006), the contribution of sensory rhodopsins to phyllospheres is much higher (25-70%; Fig. 3). This suggests that microorganisms in the phyllosphere are intensively engaged in light sensing, to accommodate the effects of fluctuations in light quality, intensity and UV radiation at the leaf surface (Ballaré et al., 1990; Beattie and Lindow, 1999).

Interestingly, all phyllosphere rhodopsins detected carry a leucine residue at position 105 (Fig. 4; based on sequence reads that contain this region; not all reads cover the entire gene due to the short nature of the 454-generated sequences, ~250-300 nt on average), which renders them as putative green light absorbing pigments (Man et al., 2003), thus avoiding an overlap with the absorption spectrum of the plant's leaf and possibly even masking out the negative role of green light on plant growth (Folta and Maruhnich, 2007). This is opposed to blue light absorbing rhodopsins (Béjà et al., 2001; Sabehi et al., 2005), which contain a glutamine instead of leucine at position 105, and are abundantly found in marine habitats (Béjà et al., 2001; Rusch et al., 2007; Sabehi et al., 2007).

Another indication that this may indeed be the case in the tamarisk phyllosphere is presented by the absorption spectra in Fig. 5; it is demonstrated that the microbes washed off the leaves has an absorption maximum around 545 nm, a region of the spectrum where there is no light absorption by the

tamarisk leaves and where the absorption of microbial rhodopsins is maximal. This absorption peak however, could also be the result of the presence of pink-pigmented *Methylobacterium* spp. containing carotenoids (Kutschera, 2007) in the leaf wash.

This is the first report on microbial rhodopsins existence in terrestrial habitats; whether it portrays commensalism or mutualism should be a matter of further investigations. We show that rhodopsin sequences have been found to be abundant both in the harsh environment of the tamarisk phyllosphere (Qvit-Raz et al., 2008) as well as on the leaves of cultivated plants; furthermore, they are common to diverse leaf shapes and plant growth characteristics, but are absent from both agricultural and forest soils. This indicates that microbial rhodopsins may be selected for in the phyllosphere environment, thus conferring an important adaptive trait onto this microbial niche. We propose that rhodopsin light interception by phyllosphere bacteria needs to be taken into account in global energy balance and biomass production by the terrestrial biosphere.

Experimental Procedures

Phyllosphere sampling

Leaf samples were collected from a *Tamarix nilotica* tree in an oasis by the Dead Sea (31º42’41.06’’N 35º27’19.32’’E), and processed within 1 hour of sampling (Qvit-Raz et al., 2008). Briefly, 50 grams of leaves were placed inside a 250 ml sterile glass Erlenmeyer flask, immediately immersed in sterile phosphate buffered saline (1 g leaf/5 ml PBS, pH 7.4), and cavitated in a

sonication bath (Transistor/ultrasonic T7 [L&R Manufacturing Company]) for two minutes at medium intensity. The preparations were then vortexed 6 X 10 sec at 5-min intervals, and the leaf wash was separated from the leaf debris by decanting and kept for analysis. *Arabidopsis*, clover and rice phyllospheres were prepared according to the previously reported soybean phyllosphere preparation (Delmotte et al., 2009).

DNA Extraction and pyrosequencing

The leaf wash was filtered on a 0.22 µm membrane filter (Millipore), which was subjected to total community DNA extraction, using a Power Soil Microbial DNA extraction kit (MoBio). Sequencing was performed on the Genome Sequencer FLX system using 3 µg of DNA at a concentration of 17 ng/µl (as determined by a nanodrop spectrophotometer). The resulting reads were annotated using the MG-RAST rapid annotation platform (Meyer et al., 2008). Using this platform, rhodopsin-containing reads were located within each of the compared metagenomes using an e-value cutoff of 10^{-5}. For inclusion in the phylogenetic analysis, hits with higher e-values were included as well. The number of reads was normalized against the average number of selected single-copy genes found in the datasets using an e-value cutoff of 10^{-20}.

All non-phyllosphere datasets used are publicly available on the MG-RAST website. The soybean phyllosphere metagenome can be found in the genbank SRA database. The rhodopsin-containing reads from the phyllosphere metagenomes are provided in the online supporting material files S1, S2, S3, S4, and S5.

Phylogenetic tree analysis

In this work, we tried several methods for multiple sequence alignment calculation (MUSCLE, ProbCons, MAFFT and PROMALS, see references within (Kemena and Notredame, 2009)). In an effort to automatically identify the most reliable multiple sequence alignment for a given protein family, we used the AQUA protocol for automated quality improvement of multiple sequence alignments (Muller et al., 2010). We performed several alignments using MUSCLE, MAFFT, ProbCons, along with one refinement program (RASCAL) and one assessment program (NORMD). According to this protocol the MAFFT alignment refined by RASCAL produced the most reliable alignment (highest NORMD value) and was used to produce the phylogenetic tree. Following the alignment computation, we used FastTree version 2.1.1 SSE3 (Price et al., 2009) for the calculation of the phylogenetic tree using settings for high accuracy [-spr 4 (to increase the number of rounds of minimum-evolution SPR moves) and -mlacc 2 -slownni (to make the maximum-likelihood NNIs search more exhaustive)]. These parameters can produce slight increases in accuracy. To estimate the reability of each split in the tree, FastTree uses a Shimodaire-Hasegawa test on the three alternate topologies (NNIs) around that split (Guindon et al., 2009). Phyllogenetic protein trees were visualized and edited using Dendroscope software version 2.7.3 (Huson et al., 2007).

To tested if the phylogenetic signal we observe is statically significant we used the Mesquite program (Maddison and Maddison, 2010). This was done using a randomization test (to see if the observed number of changes on the tree is less than 95% of the null values). The 10,000 reshufflings of the

characters (5 different plants and other environments) allowed constructing a character chart of parsimonious changes between the 6 characters assigned.

Relative abundance of microbial rhodopsins in different metagenomes

Frequency of rhodopsin blast hits with an e-value ≤ 10^{-5} was determined for 14 metagenomes from phyllosphere (5), marine (5), freshwater (1), hypersaline (1) and soil (2) environments. Rhodopsin abundance was normalized with numbers of *rplA*, *rplC*, *rplD*, *rpoA*, *rpoB*, and *rspJ* genes (Frank and Sorensen, 2011) (blast hits with an e-value ≤ 1e-20) according to (Yutin et al., 2007; Howard et al., 2008).

Metagenomic datasets used for comparison (Fig. 2)

Freshwater: GS020, Lake Gatun, Panama (MG_RAST accession: 4441590.3)

Hypersaline: GS033, Punta Cormorant hypersaline lagoon, Galapagos (MG-RAST accession: 4441599.3)

Open Sea: GS000a, Sargasso Station 11 (MG-RAST accession: 4441570.3) and GS000b, Sargasso Station 13 (MG-RAST accession: 4441573.3)

Estuary: Monterey Bay (MG-RAST accession: 4443712.3)

Whale Fall: Whale fall Bone (MG-RAST accession: 4441619.3)

Forest Soil: Luquillo experimental forest soil, Puerto Rico (MG-RAST accession: 4446153.3) and Waseca farm soil (MG-RAST accession: 4441091.3)

Soybean: SRA accession: SRX008324 (http://www.ncbi.nlm.nih.gov/sra/SRX008324?report=full)

Absorbance spectra of tamarisk leaves and phyllosphere wash (Fig. 5)

Phyllosphere absorbance was calculated as the difference between two measurements of reflectance spectra (intact tamarisk leaves and phosphate buffered saline washed, sonicated leaves), obtained at room temperature with a Labsphere DRA-CA-30I diffuse reflectance accessory. Leaves were organized in high density on a slide and covered with another slide. Two empty slides were used as a blank. Measurements were performed on 4 different leaf samples from different dates. For chlorophylls absorbance, tamarisk leaves were grinded with acetone 90% and filtered through GFF filters. The extraction was measured using a Cary 100 spectrophotometer.

Acknowledgments

We thank Robert Edgar for his help and discussions over issues of short sequences alignment and Eli Geffen for his help with statistics. Our gratitude to the staff and scientists of the Josephine Bay Paul Center in Comparative Molecular Biology and Evolution at the Marine Biology Laboratory (Woods Hole, MA, USA) for their help. This work was supported in part by a grant from Bridging the Rift Foundation (O.B. & S.B.), Israel Science Foundation grant 1203/06 (O.B.), the Gruss-Lipper Family Foundation at MBL (O.M.F., S.B. & A.F.P.), a US-Israel Binational Science Foundation grant 2006324 (S.B.), and DOE National Institutes of Health Grant R37GM27750, Department of Energy Grant DE-FG02-07ER15867, and endowed chair AU-0009 from the Robert A. Welch Foundation (J.L.S.).

Figure Captions

Figure 1. A phylogenetic tree of rhodopsin-deduced amino acid sequences from the phyllospheres of tamarisk, rice, soybean, *Arabidopsis* and clover. Following alignment computation (see methods), a FastTree version 2.1.1 was used for the calculation of the approximately-maximum-likelihood phylogenetic tree using settings for high accuracy. Bootstraps above 60% are shown as black circles at the junctions. PR- proteorhodopsins, HR- halorhodopsins, BR- bacteriorhodopsins, SRI- sensory rhodopsins-I, SRII- sensory rhodopsins-II.

Figure 2. Relative abundance of microbial rhodopsins in different metagenomes. MG_RAST (Meyer et al., 2008) accession numbers of the different datasets can be found in the methods section. Abundance was normalized relative to the numbers of *rplA*, *rplC*, *rplD*, *rpoA*, *rpoB*, and *rspJ* genes (Frank and Sorensen, 2011) in each environment.

Figure 3. Sensory rhodopsins and proton pumps in different environments. Proportions of sensory rhodopsins and rhodopsin proton pumps were calculated only from reads containing the region surrounding the proton acceptor and donor carboxylates at helix C (bacteriorhodopsin positions 85 and 96, respectively); Sargasso Sea (Spudich, 2006) (n=732), tamarisk (n=13), soybean (n=31), rice (n=8), *Arabidopsis* (n=4), and clover (n=7).

Figure 4. Protein alignment of phyllosphere rhodopsins. Amino acid position 105 is marked with green or blue backgrounds according to the predicted absorption spectra of the rhodopsin pigments. Only the vicinity of amino acid 105 is shown. Examples from confirmed green absorbing proteorhodopsins *e*BAC31A08 (Béjà et al., 2000), *Dokdonia* MED134 (Gómez-Consarnau et al., 2007) and confirmed blue absorbing proteorhodopsins PAL-E6 (Béjà et al., 2001), *e*BAC49C08 (Sabehi et al., 2005) are shown for reference at the top left corner. Names of rhodopsins from the soybean phyllosphere start with SRR and from the tamarisk start with GDOVJJ. Only a subset of the phyllosphere rhodopsins is shown. See files S1, S2, S3, S4, and S5 for more variations.

Figure 5. Absorbance spectra of tamarisk leaves and phyllosphere wash. Absorbance of tamarisk chlorophylls (acetone extract) and of phyllosphere leaf buffer-wash are shown. Chlorophyll absorbance is shown for illustrative purposes only; note the different scales used.

References

Atamna-Ismaeel, N., Sabehi, G., Sharon, I., Witzel, K.-P., Labrenz, M., Jürgens, K. et al. (2008) Widespread distribution of proteorhodopsins in freshwater and brackish ecosystems. *ISME J* **2**: 656-662.

Balashov, S.P., Imasheva, E.S., Boichenko, V.A., Anton, J., Wang, J.M., and Lanyi, J.K. (2005) Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. *Science* **309**: 2061-2064.

Ballaré, C.L., Scopel, A.L., and Sánchez, R.A. (1990) Far-red radiation reflected from adjacent leaves: an early signal of competition in plant canopies. *Science* **247**: 329-332.

Beattie, G.A., and Lindow, S.E. (1999) Bacterial colonization of leaves: a spectrum of strategies. *Phytopathol* **89**: 353-359.

Béjà, O., Spudich, E.N., Spudich, J.L., Leclerc, M., and DeLong, E.F. (2001) Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786-789.

Béjà, O., Aravind, L., Koonin, E.V., Suzuki, M.T., Hadd, A., Nguyen, L.P. et al. (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902-1906.

de la Torre, J.R., Christianson, L., Béjà, O., Suzuki, M.T., Karl, D., Heidelberg, J.F., and DeLong, E.F. (2003) Proteorhodopsin genes are widely distributed among divergent bacterial taxa. *Proc Natl Acad Sci U S A* **100**: 12830-12835.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R. et al. (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci U S A* **106**: 16428-16433.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237-240.

Folta, K.M., and Maruhnich, S.A. (2007) Green light: a signal to slow down or stop. *J Exp Bot* **58**: 3099-3111.

Frank, J.A., and Sorensen, S.J. (2011) Quantitative metagenomic analyses based on average genome size normalization. *Appl Environ Microbiol* **77**: 2513-2521.

Frigaard, N.-U., Martinez, A., Mincer, T.J., and DeLong, E.F. (2006) Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. *Nature* **439**: 847-850.

Giovannoni, S.J., Bibbs, L., Cho, J.-C., Stapels, M.D., Desiderio, R., Vergin, K.L. et al. (2005) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438**: 82-85.

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

Gómez-Consarnau, L., Akram, N., Lindell, K., Pedersen, A., Neutze, R., Milton, D.L. et al. (2010) Proteorhodopsin phototrophy confers enhanced survival of marine bacteria during starvation. *PLoS Biol* **8**: e1000358.

Guindon, S., Delsuc, F., Dufayard, J.F., and Gascuel, O. (2009) Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* **537**: 113–137.

Howard, E.C., Sun, S., Biers, E.J., and Moran, M.A. (2008) Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environ Microbiol* **10**: 2397-2410.

Huson, D.H., Richter, D.C., Rausch, C., Dezulian, T., Franz, M., and Rupp, R. (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8**: 460.

Kemena, C., and Notredame, C. (2009) Upcoming challenges for multiple sequence alignment methods in the high-throughput era. *Bioinformatics* **25**: 2455-2465.

Koh, E.Y., Atamna-Ismaeel, N., Martin, A., Cowie, R.O., Beja, O., Davy, S.K. et al. (2010) Proteorhodopsin-bearing bacteria in Antarctic sea ice. *Appl Environ Microbiol* **76**: 5918-5925.

Kutschera, U. (2007) Plant-associated methylobacteria as co-evolved phytosymbionts: a hypothesis. *Plant Signal Behav* **2**: 74-78.

Lanyi, J.K., and Balashov, S.P. (2008) Xanthorhodopsin: a bacteriorhodopsin-like proton pump with a carotenoid antenna. *Biochim Biophys Acta* **1777**: 684-688.

Lindow, S.E., and Brandl, M.T. (2003) Microbiology of the phyllosphere. *Appl Environ Microbiol* **69**: 1875-1883.

Maddison, W.P., and Maddison, D.R. (2010) Mesquite: a modular system for evolutionary analysis. Version 2.73 http://mesquiteproject.org.

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

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M. et al. (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**: 386.

Muller, J., Creevey, C.J., Thompson, J.D., Arendt, D., and Bork, P. (2010) AQUA: automated quality improvement for multiple sequence alignments. *Bioinformatics* **26**: 263-265.

Oesterhelt, D., and Stoeckenius, W. (1971) Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nat New Biol* **233**: 149-152.

Oh, H.M., Kwon, K.K., Kang, I., Kang, S.G., Lee, J.H., Kim, S.J., and Cho, J.C. (2010) Complete genome sequence of "*Candidatus* Puniceispirillum marinum" IMCC1322, a representative of the SAR116 clade in the *Alphaproteobacteria*. *J Bacteriol* **192**: 3240-3241.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* **26**: 1641-1650.

Qvit-Raz, N., Jurkevitch, E., and Belkin, S. (2008) Drop-size soda lakes: transient microbial habitats on a salt-secreting desert tree. *Genetics* **178**: 1615-1622.

Rusch, D.B., Halpern, A.L., Heidelberg, K.B., Sutton, G., Williamson, S.J., 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., Kirkup, B.C., Rosenberg, M., Stambler, N., Polz, M.F., and Béjà, O. (2007) Adaptation and spectral tuning in divergent marine proteorhodopsins from the eastern Mediterranean and the Sargasso Seas. *ISME J* **1**: 48-55.

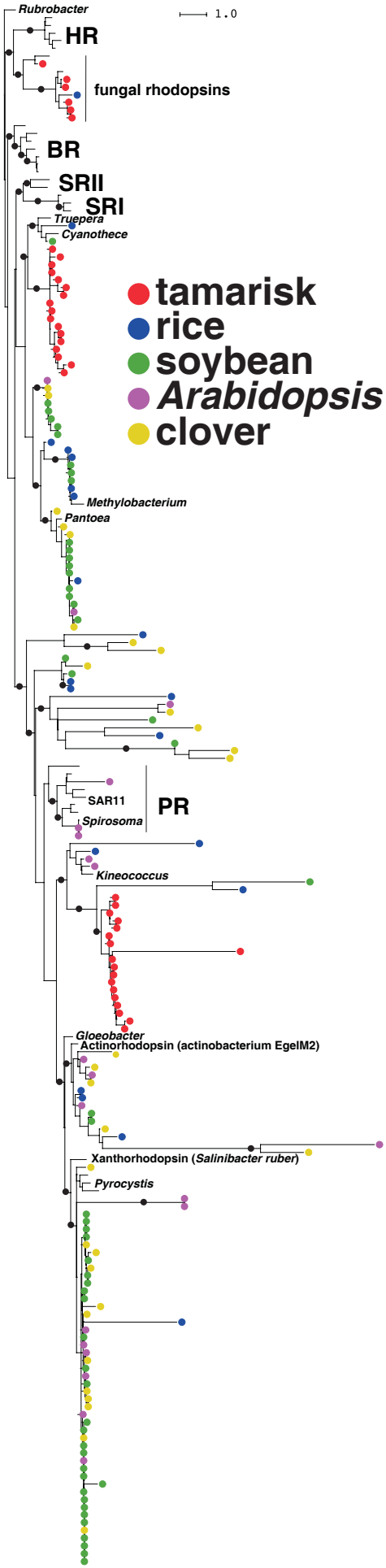
Sabehi, G., Loy, A., Jung, K.H., Partha, R., Spudich, J.L., Isaacson, T. et al. (2005) New insights into metabolic properties of marine bacteria encoding proteorhodopsins. *PLoS Biol* **3**: e173.

Sharma, A.K., Zhaxybayeva, O., Papke, R.T., and Doolittle, W.F. (2008) Actinorhodopsins: proteorhodopsin-like gene sequences found predominantly in non-marine environments. *Environ Microbiol* **10**: 1039-1056.

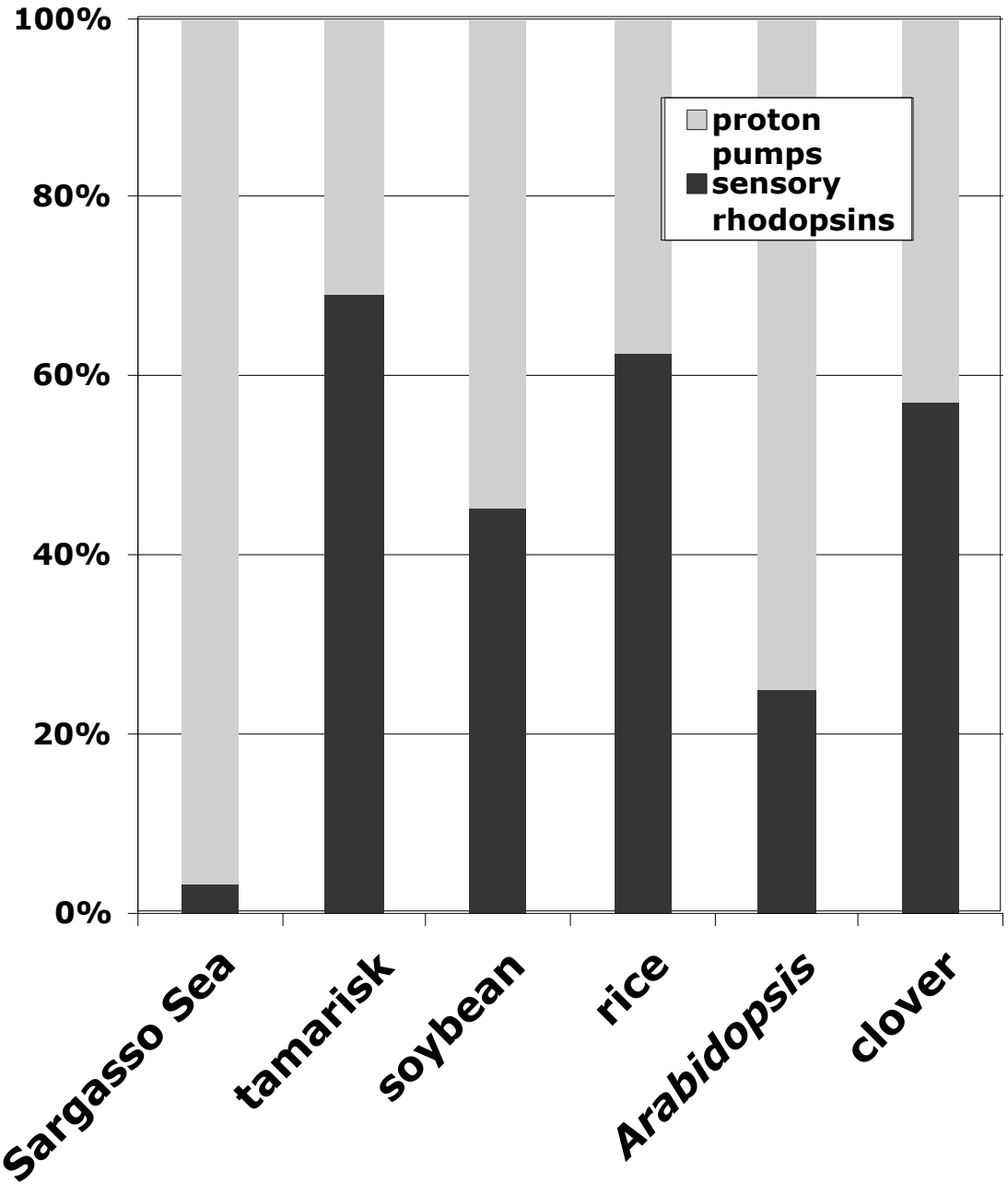
Sharma, A.K., Sommerfeld, K., Bullerjahn, G.S., Matteson, A.R., Wilhelm, S.W., Jezbera, J. et al. (2009) Actinorhodopsin genes discovered in diverse freshwater habitats and among cultivated freshwater *Actinobacteria*. *ISME J* **3**: 726-737.

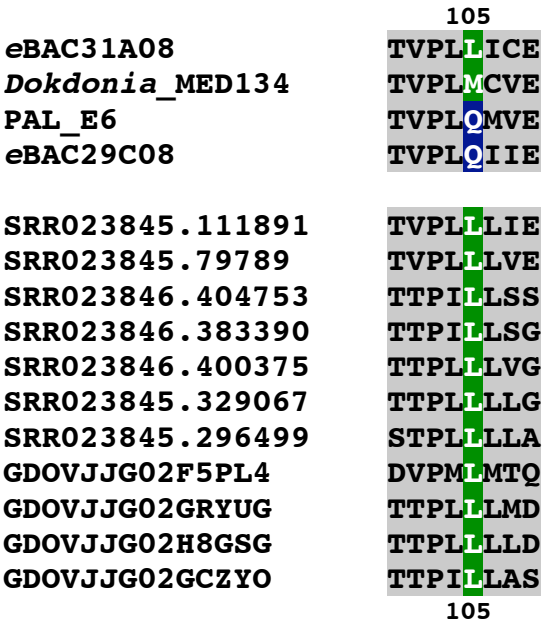
Spudich, J.L. (2006) The multitalented microbial sensory rhodopsins. *Trends Microbiol* **14**: 480-487.

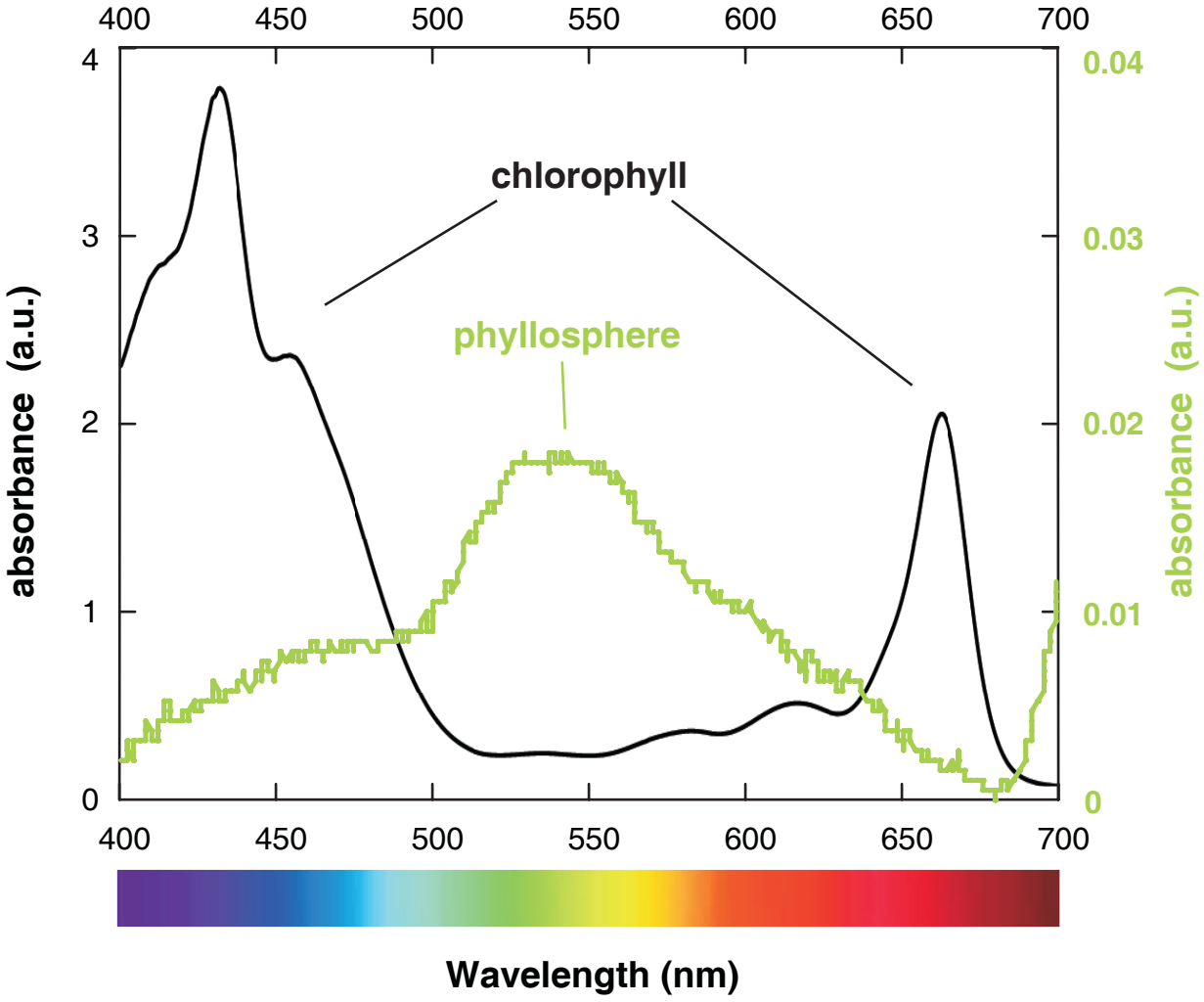
Yutin, N., Suzuki, M.T., Teeling, H., Weber, M., Venter, J.C., Rusch, D., and Béjà, O. (2007) Assessing diversity and biogeography of aerobic anoxygenic phototrophic bacteria in surface waters of the Atlantic and Pacific Oceans using the Global Ocean Sampling expedition metagenomes. *Environ Microbiol* **9**: 1464-1475.


Rubrobacter
1.0
HR
fungal rhodopsins
BR
SRII
SRI
Truepera
Cyanothece
tamarisk
rice
soybean
Arabidopsis
clover
Methylobacterium
Pantoea
SAR11
Spirosoma
PR
Kineococcus
Gloeobacter
Actinorhodopsin (actinobacterium EgelM2)
Xanthorhodopsin (Salinibacter ruber)
Pyrocystis

Relative abundance of microbial rhodopsins
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
North Atlantic
Monterey Bay
Saragasso 11
Saragasso 13
Lake Gatun
Punta Cormorant
Whale fall bone
Luquillo forest soil
Waseca farm soil
Soybean
tamarisk
Arabidopsis
Clover
Rice


100%
80%
60%
40%
20%
0%
proton pumps
sensory rhodopsins
Sargasso Sea
tamarisk
soybean
rice
Arabidopsis
clover


105
eBAC31A08 TVPLLICE
Dokdonia_MED134 TVPLMCVE
PAL_E6 TVPLQMVE
eBAC29C08 TVPLQIIE
SRR023845.111891 TVPLLLIE
SRR023845.79789 TVPLLLVE
SRR023846.404753 TTPILLSS
SRR023846.383390 TTPILLSG
SRR023846.400375 TTPLLLVG
SRR023845.329067 TTPLLLLG
SRR023845.296499 STPLLLLA
GDOVJJG02F5PL4 DVPMLMTQ
GDOVJJG02GRYUG TTPLLLMD
GDOVJJG02H8GSG TTPLLLLD
GDOVJJG02GCZYO TTPILLAS
105


400
450
500
550
600
650
700
4
3
2
1
0
0.04
0.03
0.02
0.01
0
absorbance (a.u.)
absorbance (a.u.)
chlorophyll
phyllosphere
Wavelength (nm)